



UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR MEDIZINISCHE INFORMATIK

Aus dem Institut für Medizinische Informatik
der Universität zu Lübeck
Direktor: Prof. Dr. rer. nat. habil. Heinz Handels

**Effiziente Schätzung der Atembewegung mittels
statistischer Bewegungs- und regressionsbasierter
Korrespondenzmodelle**

Inauguraldissertation
zur
Erlangung der Doktorwürde
der Universität zu Lübeck

Aus der Sektion Informatik/Technik

vorgelegt von
Matthias Wilms
aus Wedel

Lübeck, 2018

1. Berichtstatter: Prof. Dr. rer. nat. habil. Heinz Handels
2. Berichtstatter: Prof. Dr.-Ing. habil. Alfred Mertins

Tag der mündlichen Prüfung: 30.10.2018

Zum Druck genehmigt. Lübeck, den 05.04.2019

Zusammenfassung

Atmungsbedingte Positionsänderungen und Deformationen anatomischer und pathologischer Strukturen stellen ein zentrales Problem während der Behandlung von Tumoren in Thorax und Abdomen mittels externer Strahlentherapie oder nichtinvasiver Thermoablation dar. Um dennoch eine präzise Behandlung zu ermöglichen, werden vielfach technische Lösungen zur Kompensation der Atembewegung eingesetzt. Für eine effektive technische Bewegungskompensation muss die intrainerventionelle Atembewegung möglichst präzise in Echtzeit geschätzt werden. Dies kann entweder auf Grundlage von Signalen externer Bewegungsindikatoren (z. B. Bauchgurt) oder durch Nutzung intrainerventionell aufgenommener Bilddaten der zu behandelnden Zielregion erfolgen. Entsprechende Bewegungsschätzungsverfahren setzen zumeist gelernte (statistische) Modelle der Atembewegung ein: Bei der Nutzung von Bewegungsindikatorsignalen, auch surrogatbasierte Bewegungsschätzung genannt, liefert ein Korrespondenzmodell die zu einer gegebenen Indikatormessung korrespondierende interne Atembewegung. In Bildregistrierungsverfahren für die Bewegungsschätzung mit intrainerventionellen Bilddaten dienen statistische Bewegungsmodelle als Vorwissen bezüglich plausibler Atembewegungen, um Genauigkeit, Robustheit und Effizienz zu erhöhen.

Diese Arbeit fokussiert sich in diesem Kontext auf die Entwicklung, Erweiterung und Evaluation von Verfahren zur Modellierung und präzisen, zeiteffizienten modellbasierten Schätzung der variablen intrainerventionellen Atembewegung ganzer Körperregionen. Hierfür werden neue Verfahren im Bereich der surrogatbasierten Bewegungsschätzung, der modellbasierten Bildregistrierung und der statistischen Bewegungsmodellierung entwickelt, welche u. a. eine höhere Genauigkeit bei Atmungsvariabilität aufweisen oder vielseitiger einsetzbar sind als bisher verfügbare Methoden. Somit leistet diese Arbeit einen Beitrag zur zukünftigen Verbesserung der intrainerventionellen Bewegungsschätzung in der klinischen Praxis.

Der erste Teil dieser Arbeit konzentriert sich auf die surrogatbasierte Bewegungsschätzung mithilfe von regressionsbasierten Korrespondenzmodellen. In diesem Zusammenhang wird zunächst ein bestehendes Framework zur patientenspezifischen Korrespondenzmodellierung durch zusätzliche Regressionsmethoden erweitert, um die Nutzung von aktuellen hochdimensionalen Indikatorsignalen (z. B. Abtastungen der Körperoberfläche) potenziell zu verbessern. Dieses Framework wird anschließend mithilfe einer neuen Umgebung zur bildbasierten Simulation von Bewegungsindikatoren und praxisnahen Anwendungsszenarien evaluiert. Die bildbasierte Simulation ermöglicht so trotz der schlechten Verfügbarkeit realer Evaluationsdaten Aussagen hinsichtlich

der Eignung verschiedener Bewegungsindikatoren und Modellierungsansätze. Ein zentrales Ergebnis dieser Studie ist die Feststellung, dass sich die Schätzungsgenauigkeit stark verringert, wenn es zu einer deutlichen Veränderung der Bewegung zwischen den Trainingsdaten des Modells und der eigentlichen Intervention (interfraktionelle Bewegungsvariabilität) kommt. Motiviert durch diese Feststellung wird deshalb ein neues Verfahren zur Generierung patientenspezifischer Korrespondenzmodelle entwickelt, welches den Mangel an Informationen hinsichtlich der patientenindividuellen Bewegungsvariabilität durch die Nutzung von Bewegungsinformationen anderer Patienten umgeht und so eine Verbesserung der Schätzungsgenauigkeit bei interfraktioneller Bewegungsvariabilität erreicht. Weiterhin beschäftigt sich diese Arbeit auch mit dem innovativen Einsatz von Korrespondenzmodellen außerhalb der intrainerventionellen Bewegungsschätzung und präsentiert in diesem Zusammenhang ein neues modellbasiertes Verfahren zur realistischen Atembewegungssimulation, das u. a. in Virtual-Reality-Simulatoren zur Erhöhung der Immersion bei abdominalen Nadelpunktionen eingesetzt werden kann.

Der zweite Teil dieser Arbeit befasst sich mit der modellbasierten Bildregistrierung zur Bewegungsschätzung unter Verwendung intrainerventionell aufgenommener Bilddaten. Hierbei stellen insbesondere die Verarbeitung der Daten in Echtzeit und der Umgang mit Dimensionsunterschieden zwischen Referenz- und Zielbild eine Herausforderung dar. Diese Arbeit präsentiert in diesem Zusammenhang zwei neue modellbasierte Registrierungsverfahren, bei denen sich die eigentliche Bildregistrierung zunächst auf eine geringe Anzahl von Merkmalspunkten beschränkt, was sie schnell und vielseitig einsetzbar macht. Im Rahmen der Bildregistrierung wird in beiden Verfahren ein statistisches Bewegungsmodell zur Regularisierung eingesetzt, welches auch als patientenspezifischer Interpolator dient, um die Atembewegung fernab der Merkmalspunkte schätzen zu können. Für beide entwickelten Verfahren wird gezeigt, dass sie echtzeitfähig sind und vielfach eine mit deutlich langsameren klassischen Registrierungsverfahren vergleichbare Genauigkeit erreichen.

Der abschließende dritte Teil dieser Arbeit widmet sich der statistischen Bewegungsmodellierung mittels Hauptkomponentenanalyse bei wenigen Trainingsdaten. Wenn repräsentative Trainingsdaten fehlen, führt dies in der Praxis häufig zu unflexiblen Modellen mit schlechter Generalisierungsfähigkeit. Für diese Szenarien wird im Rahmen dieser Arbeit ein neues Verfahren zur Generierung flexibler statistischer Bewegungsmodelle vorgestellt. Mit diesen flexiblen Modellen ist bei der modellbasierten Bewegungsschätzung eine im Vergleich zu Standardmodellen signifikant höhere Genauigkeit erreichbar. Kern des Verfahrens ist eine Lokalitätsannahme, nach der die gegenseitige Abhängigkeit von Bewegungen räumlich entfernter Körperregionen gering sein sollte. Durch distanzbasierte Manipulationen der Kovarianzmatrix lässt sich diese Annahme in den Standardmodellierungsansatz integrieren und sorgt so u. a. für eine verbesserte Repräsentation lokaler Bewegungsvariationen.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Medizinischer Kontext	1
1.2	Ziele, wissenschaftlicher Beitrag und Aufbau dieser Arbeit	6
1.3	Publikationen	9
2	Framework für die patientenspezifische Korrespondenzmodellierung	11
2.1	Hintergrund und Vorgehen	12
2.1.1	Literaturübersicht	13
2.1.2	Einordnung des weiteren Vorgehens	17
2.1.3	Notation	18
2.2	Diffeomorphe Bildregistrierung zur Schätzung der Atembewegung in 4D-Bilddaten	18
2.2.1	Gütemaße für die Evaluation diffeomorpher Bildregistrierungsverfahren	21
2.3	Regressionsbasierte diffeomorphe Korrespondenzmodellierung	23
2.4	Erweiterung des Frameworks durch projektionsbasierte Regressionsmethoden	25
2.4.1	Hauptkomponentenregression (PCR)	26
2.4.2	Partial Least Squares (PLS)	28
2.4.3	Kanonische Korrelationsanalyse (CCA)	28
2.5	Zusammenfassung	29
3	Simulationsbasierte Evaluation patientenspezifischer Korrespondenzmodelle	31
3.1	Studiendesign	33
3.1.1	Verwendete Bilddaten	33
3.1.2	Simulierte Surrogate	35
3.1.3	Experimente und Evaluationsmaße	38
3.2	Ergebnisse	43
3.3	Zusammenfassung und Einordnung der Ergebnisse	50
4	Patientenspezifische Simulation der Atembewegung mit Korrespondenzmodellen	55
4.1	B-Spline-basierte Modellierung von Surrogatsignalen	58
4.1.1	Zerlegung des Surrogatsignals in Atemzyklen	59

4.1.2	B-Spline-basierte Approximation einzelner Atemzyklen	60
4.1.3	Von der statistischen Modellierung zur Signal- und Bewegungs- simulation	62
4.1.4	Evaluation der Simulationsmethode	64
4.1.5	Zusammenfassung	67
4.2	Anwendung 1: 4D-Dosissimulation in der Strahlentherapieplanung . . .	67
4.2.1	4D-Dosissimulation durch modellbasierte Dosisakkumulation . .	69
4.2.2	Statistische Untersuchung atmungsbedingter Effekte auf die Do- sisverteilung	70
4.2.3	Experimente mit IMRT-Bestrahlungsplänen	72
4.2.4	Ergebnisse	73
4.2.5	Zusammenfassung	74
4.3	Anwendung 2: Virtual-Reality-Training von Punktionseingriffen	77
4.3.1	Vom statischen zum atmenden virtuellen Patienten	78
4.3.2	Initiale Experimente und deren Ergebnisse	81
4.3.3	Zusammenfassung und weitere Überlegungen	83
4.4	Zusammenfassung	84
5	Subpopulationsbasierte Erweiterung des Korrespondenzmodellierungs- frameworks	85
5.1	Hintergrund und vergleichbare Arbeiten	86
5.2	Subpopulationsbasierte Korrespondenzmodellierung	88
5.2.1	Von der patientenspezifischen zur populationsbasierten Korre- spondenzmodellierung im Atlasraum	91
5.2.2	Sparse Manifold Clustering zur Identifikation von Subpopulationen	94
5.3	Experimente	97
5.3.1	Verwendete Bilddaten	98
5.3.2	Design der Experimente	99
5.4	Ergebnisse	101
5.5	Zusammenfassung und Einordnung der Ergebnisse	104
5.6	Alternative Methode basierend auf affinen Unterräumen	107
6	Patch- und modellbasierte Registrierungsverfahren für die zeiteffi- ziente Bewegungsschätzung mit intrainerventionellen Bilddaten	111
6.1	Literaturübersicht	113
6.2	Problembeschreibung und generelles Vorgehen	115
6.3	Registrierung durch modellgeführtes Block-Matching	120
6.3.1	Methodenbeschreibung	121
6.3.2	Zusammenfassung	123
6.4	Registrierung mit einem Patch-AAM	125
6.4.1	Methodenbeschreibung	126

6.4.2 Zusammenfassung	128
6.5 Experimente	129
6.5.1 Verwendete Bilddaten	129
6.5.2 Design der Experimente	130
6.6 Ergebnisse	135
6.7 Zusammenfassung und Diskussion	138
7 Generierung flexibler statistischer Bewegungsmodelle mit wenigen Trainingsdaten	141
7.1 Hintergrund, Vorgehen und verwandte Arbeiten	142
7.2 Statistische Bewegungsmodellierung mit Lokalitätsannahme	145
7.2.1 Integration der Lokalitätsannahme in den Modellierungsprozess	145
7.2.2 Fusionierung verschiedener lokaler Modelle	148
7.2.3 Interpretation des Ansatzes mit Kernelfunktionen	150
7.2.4 Anmerkungen zur praktischen Umsetzung	152
7.3 Experimente	153
7.4 Ergebnisse	155
7.5 Zusammenfassung und Diskussion	158
8 Zusammenfassung und Ausblick	161
8.1 Wissenschaftlicher Beitrag	161
8.2 Ausblick	165
A Ergänzende Ergebnisse zu Kapitel 3	169
A.1 Ergebnisse für DIRLAB- und CREATIS-Daten	169
B Eigene Publikationen mit Bezug zur Dissertation	171
Literaturverzeichnis	175

Kapitel 1

Einleitung

Kapitelinhalt

1.1 Medizinischer Kontext	1
1.2 Ziele, wissenschaftlicher Beitrag und Aufbau dieser Arbeit	6
1.3 Publikationen	9

1.1 Medizinischer Kontext

Die menschliche Atmung ist ein komplexer und dynamischer Prozess, bei dem es durch das Zusammenspiel verschiedener Muskeln und einer damit verbundenen Veränderung des Thoraxvolumens zu einem Gasaustausch in der Lunge kommt. Diese Volumenveränderungen während der Ein- und Ausatmung sorgen für eine atemungsbedingte Bewegung bzw. Deformation der meisten thorakalen und abdominalen Strukturen, was im Folgenden als Atembewegung bezeichnet wird (siehe Abb. 1.1 für ein Beispiel). Die Atembewegung ist in Abhängigkeit von der Körperregion unterschiedlich stark ausgeprägt, wobei die maximale Bewegung typischerweise im Bereich des Zwerchfells zu finden ist und die Bewegung in kраниокаудaler Richtung die anderen Bewegungsrichtungen dominiert [Brandner u. a., 2006; Bussels u. a., 2003; Liu u. a., 2007]. Für zwerchfellnahe Tumoren werden in der Literatur beispielsweise Bewegungsamplituden von bis zu 5 cm genannt [Chen u. a., 2001].

In der klinischen Praxis stellt die Atembewegung u. a. eines der Hauptprobleme bei nichtinvasiven onkologischen Interventionen im Thorax oder Abdomen mittels externer Strahlentherapie [Keall u. a., 2006] oder durch Thermoablation mit hochintensivem fokussiertem Ultraschall (HIFU; [Muller u. a., 2013]) dar. Die Atembewegung führt während dieser Interventionen zu einer kontinuierlichen Positionsänderung/Deformation der pathologischen Zielstruktur (Tumor), was eine präzise Behandlung erschwert. Bei der Strahlentherapie werden die Tumorzellen mit einem (zumeist) durch einen externen Teilchenbeschleuniger erzeugten Therapiestrahл mittels ionisierender Strahlung geschädigt [Wannenmacher u. a., 2013], wohingegen bei HIFU die durch die Bündelung von Ultraschallwellen im Fokuspunkt lokal entstehende Hitze hierfür verantwortlich ist

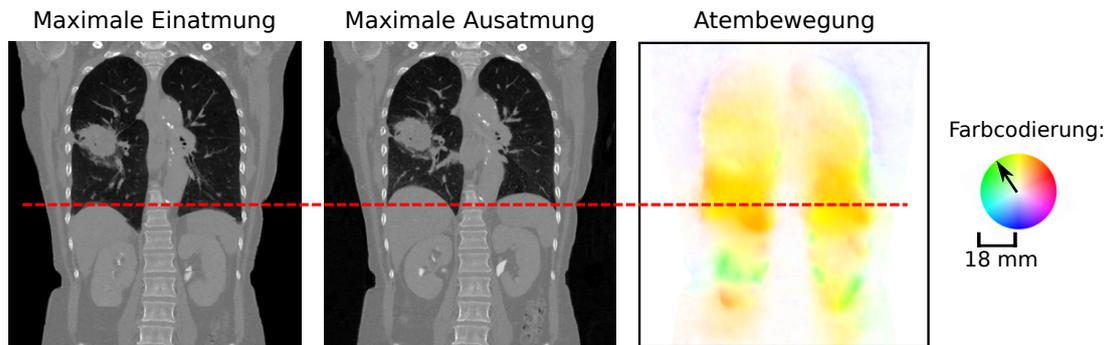


Abb. 1.1: Visualisierung der Atembewegung zwischen den Phasen maximaler Ein- und Ausatmung des 4D-CT-Bilddatensatzes eines Lungentumorpatienten für eine koronare Schicht, welche Thorax und Abdomen zeigt. Die Atembewegung wurde mittels nichtlinearer Bildregistrierung der beiden Bilder geschätzt. Die Bewegungsdarstellung erfolgt farbcodiert im Koordinatensystem des Bildes maximaler Einatmung. Die Codierung gibt die Richtung der Bewegung jedes Punktes (Farbwert) und deren Amplitude (Sättigung; höhere Sättigung entspricht höherer Amplitude) an.

[Haar, 2016]. Um trotz der Atembewegung eine ausreichende Bestrahlung bzw. Erhitzung der Zielstruktur zu gewährleisten, ohne dass es durch eine einfache Vergrößerung der zu bestrahlenden/erhitzenden Region zu einer übermäßigen Belastung von Nachbarstrukturen kommt, muss die Bewegung intrainerventionell kompensiert werden. Da die Patienten vielfach ihre Atmung nicht über längere Zeit anhalten können, haben sich in der Praxis technische Kompensationsmechanismen etabliert. Diese können grob in aktive und passive Verfahren unterteilt werden. Bei aktiven Verfahren wird der Therapiestrahlfokus mit der Zielstruktur mitbewegt (Tracking; [Booth u. a., 2016; Muller u. a., 2013; Schweikard u. a., 2000]), während die Strahlen- bzw. Schallquelle bei passiven Verfahren dynamisch ein- und ausgeschaltet wird, wenn sich die Zielstruktur im Bereich des ortsfesten Therapiestrahls/Fokuspunktes befindet bzw. diesen wieder verlässt (Gating; [Kubo u. a., 1996; Muller u. a., 2013]). Aktives Tracking ist technisch aufwendiger umzusetzen, wohingegen das technisch einfachere Gating durch die häufigen Pausen u. a. die Behandlungsdauer verlängert.

Unabhängig von der konkreten technischen Umsetzung der Bewegungskompensation erfordern diese Techniken eine kontinuierliche, intrainerventionelle Kenntnis der Position der Zielstruktur. Bei einer zusätzlichen Überwachung und Analyse der Dosis- oder Temperaturverteilung im Körper sind zudem Informationen bezüglich der atmungsbedingten Bewegung umgebender Strukturen notwendig [Tanner u. a., 2012]. Diese Erfordernisse motivieren die Entwicklung und den Einsatz von Verfahren zur intrainerventionellen Schätzung der komplexen Atembewegung ganzer Körperregionen in Echtzeit. Die Bewegung wird hierbei in der Regel mit Bezug zu einem zu Planungszwecken akquirierten 3D-Bild (Referenzbild) der entsprechenden Körperregion geschätzt.

Idealerweise ist das Resultat der Bewegungsschätzung dann eine nichtlineare Transformation, welche die atmungsbedingte Bewegung aller relevanten Strukturen zwischen Referenzbild und aktuellem Behandlungszeitpunkt beschreibt (siehe Abb. 1.2). Diese Transformation dient dann als Grundlage für die Bewegungskompensation, wobei die Überbrückung technisch bedingter Latenzzeiten, die sich u. a. bei der Neuausrichtung des Therapiestrahls ergeben, den Einsatz zusätzlicher Bewegungsprädiktionsverfahren [Ernst u. a., 2013] erfordert.

Vielfach erfolgt die intrainerventionelle Bewegungsschätzung im Kontext der Strahlentherapie atmungsbewegter Tumoren im klinischen Alltag mithilfe indirekter Verfahren, bei denen mittels Informationen externer Bewegungsindikatoren (z. B. Bauchgurte) auf die interne Atembewegung geschlossen wird (Surrogatbasierte Bewegungsschätzung) [Keall u. a., 2006; Korreman, 2015; McClelland u. a., 2013]. Dies ergibt sich aus der Tatsache, dass die (alleinige) direkte bildgestützte Beobachtung der Zielregion während der Intervention mit projektiven Standardtechniken wie Fluoroskopie/kV-Bildgebung oftmals nur unzureichend bzw. durch die zusätzliche Strahlenbelastung zeitlich sehr eingeschränkt umsetzbar ist und/oder der risikobehafteten Implantation von Markern bedarf. Aktuelle technische Entwicklungen u. a. hin zur Integration von Magnetresonanztomographiesystemen in Linearbeschleuniger (MRT-Linearbeschleuniger; [Keall u. a., 2014; Lagendijk u. a., 2014; Menten u. a., 2017; Mutic u. a., 2014]) oder der Einsatz von 4D-Ultraschallsystemen [O’Shea u. a., 2016a; Western u. a., 2015] eröffnen allerdings neue Möglichkeiten zur rein bildbasierten Bewegungskompensation anhand in Echtzeit intrainerventionell akquirierter, überlagerungsfreier Bilddaten. Die gesuchte Transformation kann hierbei durch Bildregistrierung von Referenzbild und intrainerventionellen Bilddaten bestimmt werden, wobei letztere die Zielregion oftmals nur unvollständig abbilden (z. B. wenige 2D-Bilder), was die Registrierungsverfahren berücksichtigen müssen. Auch für HIFU-Anwendungen ist bildbasierte Bewegungskompensation möglich [Denis de Senneville u. a., 2012]. Des Weiteren versprechen direkte Verfahren durch die Beobachtung der Zielregion verglichen mit indirekten Verfahren eine höhere Bewegungsschätzungsgenauigkeit [Paganelli u. a., 2015; Park u. a., 2018].

Die genauen Echtzeitanforderungen, die indirekte und direkte Bewegungsschätzungsverfahren erfüllen sollten, hängen u. a. vom Therapiesystem und der Frequenz ab, mit der neue Daten geliefert werden können (z. B. Bauchgurt: 40 Hz [Otani u. a., 2010] und 2D-MRT: 10 Hz [Denis de Senneville u. a., 2015]). Ziel sollte es sein, die Daten zumindest so schnell verarbeiten zu können, wie das System neue Daten akquiriert. Die AAPM Task Group 76 [Keall u. a., 2006] fordert zudem, dass der komplette Prozess der Bewegungskompensation 0.5 s nicht überschreiten sollte, um die Notwendigkeit zur Bewegungsprädiktion zeitlich zu begrenzen.

Verfahren zur indirekten oder direkten Bewegungsschätzung (siehe Abb. 1.2) nutzen in der Regel gelernte (statistische) Modelle der Atembewegung [McClelland u. a., 2013; Tanner u. a., 2012]. Diese Modelle repräsentieren die komplexe Atembewegung und ihre Variabilität auf kompakte Weise mit wenigen Freiheitsgraden und werden ent-

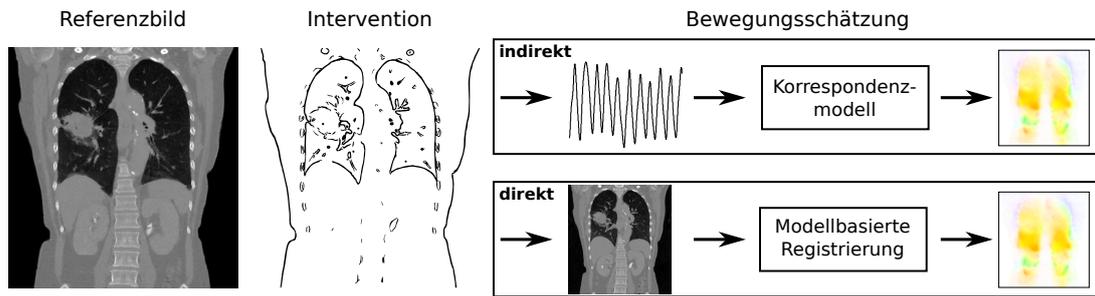


Abb. 1.2: Schematische Illustration der intrainerventionellen Schätzung der Atembewegung zwischen einem bekannten Referenzbild und der hier abstrakt dargestellten Position der internen Strukturen zu einem spezifischen Zeitpunkt einer Behandlungssitzung mittels indirekter (oben; basierend auf Indikatorsignal) oder direkter Verfahren (unten; basierend auf intrainerventioneller Bildgebung).

weder auf Basis patientenspezifischer Bild- bzw. Bewegungsdaten (patientenspezifische Modelle) oder unter Verwendung von Informationen verschiedener Patienten (Populationsmodelle) trainiert. Bei der surrogatbasierten Bewegungsschätzung werden diese Modelle auch Korrespondenzmodelle genannt. Diese beschreiben den Zusammenhang zwischen Bewegungsindikatorsignal und interner Atembewegung und ermöglichen so die Bewegungsschätzung anhand einer Signalmessung. Im Rahmen der bildbasierten Bewegungskompensation werden häufig sogenannte statistische Bewegungsmodelle bei der Bildregistrierung eingesetzt. Diese Modelle integrieren anhand von Trainingsdaten erlerntes Vorwissen bezüglich plausibler Bewegungen in den dann modellbasierten Registrierungsprozess, um eine genaue, robuste und schnelle Bewegungsschätzung zu ermöglichen. Das Modelltraining kann beispielsweise anhand vor der eigentlichen Behandlung in einer Planungssitzung aufgenommener Daten oder zu Beginn einer/jeder Behandlungssitzung (Fraktion) akquirierter Daten erfolgen.

Die Generalisierungsfähigkeit der eingesetzten Modelle und damit ihre Praxistauglichkeit hängt u. a. wesentlich von der Angemessenheit der zugrunde liegenden mathematischen Modellierungsannahmen und der Repräsentativität der genutzten Trainingsdaten ab. In der Praxis stellt insbesondere die ausgeprägte Variabilität der Atembewegung eine Herausforderung dar, die bei der Modellierung beachtet werden muss, um eine genaue modellbasierte Bewegungsschätzung auch mit kompakten Modellen zu ermöglichen. Die menschliche Atmung ist zwar prinzipiell durch die sich stets wiederholende Abfolge von Atemzyklen bestehend aus Ein- und Ausatmungsphase gekennzeichnet. Allerdings ist dieser Prozess einer gewissen Variabilität unterworfen [Benchetrit, 2000], die nicht nur zu Unterschieden zwischen verschiedenen Patienten (Interpatientenvariabilität) führt, sondern auch in einer patientenindividuellen Variabilität der Bewegung resultiert. Abseits von reinen Frequenz- oder Amplitudenvariationen, die beispielsweise durch Veränderungen der körperlichen Belastung entstehen und relativ

einfach kompensiert werden können, ist u. a. der variable Einfluss, den die einzelnen an der Atmung beteiligten Muskelgruppen (u. a. Zwerchfell und Zwischenrippenmuskeln) haben können [McClelland u. a., 2013], eine Ursache für zum Teil fundamentale Variationen in der Bewegungscharakteristik (z. B. Wechsel von Bauch- auf Brustatmung), welche die Bewegungsschätzung erschweren. Weitere Bewegungsveränderungen sind beispielsweise durch natürliche Drifts von Strukturen oder behandlungsbedingte Veränderungen der Zielregion möglich. Drifts resultieren u. U. in einer relativen Veränderung der Bewegung verschiedener Körperbereiche zueinander [Siebenthal u. a., 2007; Sonke u. a., 2008] und Volumenveränderungen des Tumors durch die Behandlung können dessen Mobilität erhöhen [Britton u. a., 2007].

Nach McClelland u. a. [2013] lässt sich die patientenindividuelle Variabilität der Atembewegung in insgesamt vier Kategorien einteilen: (1) Intrazyklische Variabilität, (2) interzyklische Variabilität, (3) intrafraktionelle Variabilität und (4) interfraktionelle Variabilität. Diese Kategorien unterscheiden sich hauptsächlich durch die verschiedenen Beobachtungszeiträume und die damit verbundenen Unterschiede in den beobachtbaren Effekten. Mit der intrazyklischen Bewegungsvariabilität wird hauptsächlich der Effekt der sogenannten Hysterese (Unterschiede des Druck-/Volumenverhältnisses zwischen Ein- und Ausatmung; [Keall u. a., 2006]) bezeichnet, wodurch ein Punkt bei der Einatmung einen anderen Weg als während der Ausatmung nimmt. Die maximale Distanz zwischen Ein- und Ausatmungsposition ist variabel und kann mehrere Millimeter betragen (bis zu 5 mm in [Seppenwoolde u. a., 2002]). Mit der interzyklischen Bewegungsvariabilität werden Variationen der Punkttrajektorien zwischen einzelnen (aufeinanderfolgenden) Atemzyklen zusammengefasst, welche von einfachen Amplitudenveränderungen bis hin zu fundamentalen Veränderungen der Bewegungscharakteristik reichen können. Die intra- und interfraktionelle Variabilität charakterisieren zusätzlich auch Veränderungen, die sich erst über einen längeren Zeitraum innerhalb der Behandlungssitzung (u. a. Drifts) oder zwischen verschiedenen Behandlungssitzungen bzw. zwischen Planungs- und Behandlungssitzung (u. a. Veränderung der Tumormobilität) ergeben.

Die konkret bei der Modellierung bzw. Bewegungsschätzung zu berücksichtigende Variabilität hängt u. a. auch vom zeitlichen Umfang der Behandlung selbst ab (insbesondere Anzahl und Dauer der einzelnen Behandlungssitzungen) und wann die Trainingsdaten aufgenommen werden. So fällt bei der Aufnahme von neuen Trainingsdaten direkt zu Beginn der Behandlungssitzung die Berücksichtigung interfraktioneller Variabilität weg. Mit steigender Behandlungszeit ist beispielsweise mit Drifts zu rechnen, sodass hiervon HIFU-Behandlungen, die zumeist in einer einzigen Sitzung mit relativ langer Dauer (teilweise mehrere Stunden) durchgeführt werden [Denis de Senneville u. a., 2012], potenziell stärker betroffen sind als konventionelle Strahlentherapiebehandlungen, wo die Gesamtdosis in mehreren kurzen Einzelsitzungen appliziert wird [Wannenmacher u. a., 2013].

Zusammenfassend lässt sich festhalten, dass die intrainterventionelle Bewegungsschätzung integraler Bestandteil moderner Therapieprozesse ist und ihre Genauigkeit direkten Einfluss auf das Therapieergebnis hat. Verbesserungen in der Schätzungs-genauigkeit erlauben eine effektivere Bestrahlung/Ablation der Zielstruktur bei gleichzeitig geringeren Belastungen umliegender Risikostrukturen, wodurch Komplikationen für den Patienten verringert werden können. (Modellbasierte) Verfahren zur intrainterventionellen Schätzung der Atembewegung sollten optimalerweise drei Eigenschaften aufweisen: (1) Eine hohe Genauigkeit, welche auch durch Bewegungsvariabilität nicht wesentlich beeinträchtigt wird, (2) Echtzeitfähigkeit und (3) die Fähigkeit zur Schätzung der Atembewegung ganzer Körperregionen.

1.2 Ziele, wissenschaftlicher Beitrag und Aufbau dieser Arbeit

Die vorliegende Dissertation beschäftigt sich mit der Entwicklung, Erweiterung und Evaluation von Verfahren zur Atembewegungsmodellierung und modellbasierten Schätzung der intrainterventionellen Atembewegung im Kontext der Strahlentherapie bzw. Ablation atmungsbewegter Tumoren. Generelles Ziel ist es, die präzise modellbasierte Schätzung der Atembewegung für ganze Körperregionen in Echtzeit zu ermöglichen. Die Arbeit zeichnet hierbei gewissermaßen den schon in Kapitel 1.1 diskutierten Übergang von indirekten, surrogatbasierten Verfahren hin zu direkten, bildbasierten Verfahren nach und lässt sich in drei Teile unterteilen:

Surrogatbasierte Bewegungsschätzung: Von der patientenspezifischen zur populationsbasierten Korrespondenzmodellierung Ausgangspunkt für die Arbeiten zur surrogatbasierten Bewegungsschätzung in dieser Dissertation bildet ein bestehendes regressionsbasiertes Framework zur (diffeomorphen) Korrespondenzmodellierung (Kapitel 2). Generell fehlt es für die Bewegungsschätzung mit patientenspezifischen Korrespondenzmodellen an umfassenden, systematischen Vergleichen aktueller Modellierungsansätze und Bewegungsindikatoren in einem kontrollierten Umfeld. Dies gilt insbesondere für die Bewegungsschätzung ganzer Körperregionen, den Einsatz hochdimensionaler Indikatorsignale kamerabasierter Systeme zur Erfassung der Körperoberfläche und den Einfluss verschiedener Arten von Bewegungsvariabilität. Ein Ziel dieser Arbeit ist es, einen Beitrag in diese Richtung zu leisten. Hierfür erfolgt in Kapitel 2 zunächst die Erweiterung des Korrespondenzmodellierungsframeworks um projektionsbasierte Regressionsmethoden, die eine verbesserte Nutzung von hochdimensionalen Indikatorsignalen ermöglichen sollen. Da das Fehlen von Vergleichsstudien u. a. in der schlechten Verfügbarkeit realer Trainings- und Evaluationsdaten begründet ist, wird in Kapitel 3 eine neue Evaluationsumgebung vorgestellt und eingesetzt. Diese erlaubt durch die Verwendung räumlich-zeitlicher Bilddaten eine bildbasierte Simulati-

on verschiedener Bewegungsindikatoren und Anwendungsszenarien und ermöglicht so einen effizienten und systematischen Vergleich verschiedener Regressionsmethoden und Bewegungsindikatoren anhand etablierter Evaluationsprinzipien der Bildregistrierung.

Werden Korrespondenzmodelle abstrakt als Generatoren betrachtet, die basierend auf (beliebigen) Indikatorsignalmessungen Atembewegungen erzeugen, ergibt sich die Möglichkeit sie zur Bewegungssimulation einzusetzen. Dies erlaubt die Erschließung bisher kaum bzw. gar nicht untersuchter Anwendungsszenarien von Korrespondenzmodellen abseits der intrainerventionellen Bewegungsschätzung. Ein Problem hierbei ist die Abhängigkeit der zeitlichen Länge und der Bewegungsvariabilität der Simulation von einem realen Indikatorsignal. Zur Reduzierung dieser Abhängigkeit wird in Kapitel 4 ein neues Verfahren zur Simulation von Indikatorsignalen vorgestellt, welches in Kombination mit einem Korrespondenzmodell realistische, variable Bewegungssimulationen beliebiger zeitlicher Länge ermöglicht. Weiterhin wird gezeigt, wie diese simulierte Atembewegung gewinnbringend zur Dosissimulation in der Strahlentherapieplanung und zur Atmungssimulation in einem Virtual-Reality-Simulator eingesetzt werden kann. Virtual-Reality-Simulatoren stellen damit eine neue Anwendung für Korrespondenzmodelle dar.

Die Genauigkeit der Bewegungsschätzung mit patientenspezifischen Korrespondenzmodellen wird u. U. stark negativ durch interfraktionelle Bewegungsvariationen zwischen der Planungssitzung, deren Daten zur Modellgenerierung genutzt werden, und den Behandlungssitzungen beeinflusst. Ein Ziel dieser Arbeit ist es, diesem Problem methodisch zu begegnen und mögliche interfraktionelle Bewegungsvariabilität bei der Modellierung zu berücksichtigen. In Kapitel 5 wird hierfür ein neues Verfahren entwickelt, das ausgewählte Daten einer großen Patientenpopulation in das Korrespondenzmodelltraining integriert. Die Annahme hierbei ist, dass die Interpatientenvariabilität ähnlich atmender Populationspatienten Ähnlichkeiten zur interfraktionellen Bewegungsvariabilität dieser Patienten aufweist und die Integration dieser Populationsdaten in den Modellierungsprozess deshalb hilfreich ist, um die Genauigkeit der Bewegungsschätzung bei interfraktionellen Variationen zu verbessern.

Modellbasierte Bildregistrierung für die Bewegungsschätzung mit intrainerventionellen Bilddaten Bei der direkten Bewegungsschätzung im Kontext MRT-basierter Therapiesysteme ergibt sich häufig ein Problem durch u. a. Unterschiede in der Datendimension zwischen Referenzbild und intrainerventionell akquirierten Bilddaten (z. B. 3D vs. 2D), welches verfügbare Registrierungsverfahren zumeist nur unzureichend oder stark situationsabhängig durch Begrenzung der Bewegungsschätzung auf den sichtbaren Bereich lösen. Ein Ziel dieser Arbeit ist es, dieses Problem allgemeiner zu lösen. Hierfür werden in Kapitel 6 zwei modellbasierte Registrierungsverfahren entwickelt, welche durch eine patchbasierte Definition des Registrierungsproblems so vielseitig sind, dass sie unabhängig von z. B. Dimensionsunterschieden oder

Modalitätsunterschieden zwischen Referenzbild und intrainerventionellen Bilddaten eingesetzt werden können. Kernelement dieser Verfahren sind patientenspezifische statistische Bewegungsmodelle, die bei der Registrierung zur Regularisierung und Bewegungsinterpolation zum Einsatz kommen, um eine genaue, zeiteffiziente und robuste Bewegungsschätzung ganzer Körperregionen zu ermöglichen.

Statistische Bewegungsmodellierung mit wenigen Trainingsdaten Ein häufiges Problem bei der Erstellung statistischer Bewegungsmodelle mittels Hauptkomponentenanalyse ist die mangelnde Verfügbarkeit von repräsentativen Trainingsdaten. Dies resultiert in zu restriktiven Modellen mit mangelnder Generalisierungsfähigkeit und schränkt die Genauigkeit der sie einsetzenden Bewegungsschätzungsverfahren ein. Ein Ziel dieser Arbeit ist es, dieses Problem durch methodische Veränderungen des Modellierungsprozesses zu verringern. Zu diesem Zweck wird in Kapitel 7 ein neues Verfahren zur Generierung von flexiblen statistischen Bewegungsmodellen mit wenigen Trainingsdaten entwickelt und im Kontext der populationsbasierten Bewegungsschätzung evaluiert. Dieses Verfahren basiert u. a. auf einer Lokalitätsannahme, nach der die Bewegung räumlich weit entfernter Bereiche (relativ) unabhängig sein soll, was durch eine Modifikation des Modellierungsprozesses abgebildet wird.

Die fünf Hauptbeiträge dieser Arbeit sind:

- Methodische Erweiterung eines diffeomorphen Korrespondenzmodellierungsframeworks um zusätzliche Regressionsmethoden sowie eine umfangreiche Evaluation des Frameworks mithilfe einer neuen Umgebung zur bildbasierten Simulation verschiedener Bewegungsindikatoren und Anwendungsszenarien (Kapitel 2 und Kapitel 3).
- Entwicklung eines neuen Verfahrens zur Simulation von realistischer, variabler Atembewegung durch Korrespondenzmodelle, was u. a. umfangreiche Dosisimulationen und zeitlich unbegrenzte Atmungssimulation in Virtual-Reality-Simulatoren ermöglicht (Kapitel 4).
- Entwicklung eines neuen Verfahrens zur Integration von Populationsinformationen in den Trainingsprozess patientenspezifischer Korrespondenzmodelle, um die Genauigkeit der Bewegungsschätzung bei interfraktioneller Bewegungsvariabilität zu verbessern (Kapitel 5).
- Entwicklung zweier neuer echtzeitfähiger patch- und modellbasierter Registrierungsverfahren für die intrainerventionelle Bildregistrierung, die sich durch ihre hohe Genauigkeit und Geschwindigkeit auszeichnen und im Vergleich zu anderen Verfahren wesentlich vielseitiger einsetzbar sind (Kapitel 6).

- Entwicklung eines neuen Verfahrens zur Erzeugung von flexiblen statistischen Bewegungsmodellen mit wenigen Trainingsdaten, welche zu einer signifikanten Verbesserung der Bewegungsschätzung gegenüber Standardmodellen führen (Kapitel 7).

In Kapitel 8 werden die Entwicklungen und Ergebnisse dieser Arbeit abschließend zusammengefasst und mögliche Anknüpfungspunkte für zukünftige Arbeiten aufgezeigt. In den einzelnen Kapiteln erfolgt zudem jeweils eine Diskussion der Forschungsliteratur, eine umfangreiche und genaue Beschreibung der Ziele der einzelnen Entwicklungen und eine Zusammenfassung und Diskussion der Ergebnisse mit kontextspezifischem Ausblick.

1.3 Publikationen

Ein großer Teil der in dieser Dissertation präsentierten Inhalte ist in Fachzeitschriften oder Konferenzbänden unter Erstautorenschaft des Autors dieser Arbeit veröffentlicht worden. Nachfolgend werden die zentralen Publikationen zu den in Kapitel 1.2 definierten Teilen der Dissertation aufgeführt. Eine vollständige Liste aller 31 Publikationen, die im direkten Zusammenhang mit dieser Dissertation veröffentlicht wurden (inkl. Koautorenschaften), ist in Anhang B zu finden. In den Einleitungen der einzelnen Kapitel werden zudem die jeweiligen Schlüsselpublikationen genannt. Das Symbol * weist auf eine geteilte Erstautorenschaft hin.

Das erweiterte diffeomorphe Korrespondenzmodellierungsframework zur surrogatbasierten Bewegungsschätzung (Kapitel 2), das simulationsbasierte Evaluationsframework (inkl. Evaluation; Kapitel 3), das Bewegungssimulationsverfahren und das Verfahren zur Integration von Populationsinformationen (Kapitel 5) sind u. a. in

- [Wilms u. a., 2014a] Wilms, M. *, Werner, R. *, Ehrhardt, J., Schmidt-Richberg, A., Schlemmer, H.-P. und Handels, H. “Multivariate regression approaches for surrogate-based diffeomorphic estimation of respiratory motion in radiation therapy”. *Phys Med Biol* 59 (5), 2014, S. 1147–1164.
- [Wilms u. a., 2014b] Wilms, M., Ehrhardt, J., Werner, R., Marx, M. und Handels, H. “Statistical Analysis of Surrogate Signals to Incorporate Respiratory Motion Variability into Radiotherapy Treatment Planning”. In: *SPIE Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and Modeling*. Bd. 9036. Proc SPIE. 2014, 90360J.
- [Wilms u. a., 2017c] Wilms, M. *, Werner, R. *, Yamamoto, T., Handels, H. und Ehrhardt, J. “Subpopulation-based correspondence modelling for improved respiratory motion estimation in the presence of inter-fraction motion variations”. *Phys Med Biol* 62 (14), 2017, S. 5823–5839.

veröffentlicht worden. Teile der modellbasierten Registrierungsverfahren aus Kapitel 6 und deren Grundlagen wurden in

- [Ha u. a., 2019] Ha, I. Y. *, Wilms, M. *, Handels, H. und Heinrich, M. P. “Model-based Sparse-to-dense Image Registration for Realtime Respiratory Motion Estimation in Image-guided Interventions”. *IEEE Trans Biomed Eng* 66 (2), 2019, S. 302–310.
- [Wilms u. a., 2016] Wilms, M., Ha, I. Y., Handels, H. und Heinrich, M. P. “Model-based Regularisation for Respiratory Motion Estimation with Sparse Features in Image-guided Interventions”. In: *Med Image Comput Comput Assist Interv – MICCAI 2016*. Bd. 9902. LNCS. Springer, 2016, S. 89–97.
- [Wilms u. a., 2017b] Wilms, M., Handels, H. und Ehrhardt, J. “Representative Patch-based Active Appearance Models Generated from Small Training Populations”. In: *Med Image Comput Comput Assist Interv – MICCAI 2017*. Bd. 10433. LNCS. Springer, 2017, S. 152–160.

veröffentlicht. In [Wilms u. a., 2017b] werden Teilaspekte bzw. Grundlagen eines der modellbasierten Registrierungsverfahren anhand des verwandten, aber weniger aufwendigen Problems der Segmentierung mit statistischen Form-/Erscheinungsmodellen eingeführt, welche dann in dieser Arbeit im Kontext der Bewegungsschätzung genutzt werden. [Wilms u. a., 2017b] beschreibt ebenso auch Teile der Grundlagen des Verfahrens zur Erzeugung flexibler Bewegungsmodelle mit wenigen Trainingsdaten aus Kapitel 7, welches initial ebenfalls im Kontext der statistischen Formmodellierung in

- [Wilms u. a., 2017a] Wilms, M., Handels, H. und Ehrhardt, J. “Multi-resolution multi-object statistical shape models based on the locality assumption”. *Med Image Anal* 38 (5), 2017, S. 17–29.

publiziert wurde.

Kapitel 2

Framework für die patientenspezifische Korrespondenzmodellierung

Kapitelinhalt

2.1	Hintergrund und Vorgehen	12
2.1.1	Literaturübersicht	13
2.1.2	Einordnung des weiteren Vorgehens	17
2.1.3	Notation	18
2.2	Diffeomorphe Bildregistrierung zur Schätzung der Atembewegung in 4D-Bilddaten	18
2.2.1	Gütemaße für die Evaluation diffeomorpher Bildregistrierungsverfahren	21
2.3	Regressionsbasierte diffeomorphe Korrespondenzmodellierung	23
2.4	Erweiterung des Frameworks durch projektionsbasierte Regressionsmethoden	25
2.4.1	Hauptkomponentenregression (PCR)	26
2.4.2	Partial Least Squares (PLS)	28
2.4.3	Kanonische Korrelationsanalyse (CCA)	28
2.5	Zusammenfassung	29

Im ersten Teil dieses Kapitels wird, nach einer generellen Einführung in das Thema Korrespondenzmodellierung (Kapitel 2.1), das in [Werner, 2012] entwickelte Verfahren bzw. Framework zur surrogatbasierten Schätzung der Atembewegung mittels patientenspezifischer Korrespondenzmodelle vorgestellt (Kapitel 2.2–2.3). In [Werner, 2012] wird ein linearer Zusammenhang zwischen externer und interner Atembewegung angenommen und das patientenspezifische Korrespondenzmodell mittels Regression bestimmt. Die methodische Besonderheit des Verfahrens besteht darin, dass die durch das Modell zu schätzende interne Atembewegung als diffeomorphe Transformation beschrieben wird. Diffeomorphe Transformationen haben den Vorteil, dass sie u. a. topo-

gieerhaltend sind, was für die physiologisch plausible Beschreibung der Atembewegung eine wichtige Eigenschaft darstellt.

Wurden in [Werner, 2012] nur relativ niedrigdimensionale Surrogatsignale genutzt (Dimension < 30), so beschäftigt sich Kapitel 2.4 mit den im Rahmen dieser Arbeit vorgenommenen methodischen Erweiterungen des Frameworks, die speziell die Probleme (z. B. Informationsredundanz/Multikollinearität) bei der regressionsbasierten Korrespondenzmodellierung mit sehr hochdimensionalen (multivariaten) Surrogatsignalen durch Nutzung von Projektionsmethoden adressieren. Diese Erweiterungen sollen im weiteren Verlauf die verbesserte Nutzung von hochdimensionalen Tiefenbildern der Körperoberfläche des Patienten mit mehreren Hundert oder Tausend Bildpunkten als Surrogatsignal ermöglichen (siehe Kapitel 3).

Insgesamt bildet dieses Kapitel die Grundlage für die nachfolgenden beiden Kapitel, die sich auf ausführliche Evaluationen, neuartige Anwendungen und methodische Erweiterungen des generellen Frameworks konzentrieren. Der Fokus der Darstellung in diesen Kapiteln liegt zumeist auf einer Anwendung der Verfahren im Kontext der Strahlentherapie atembewegter Tumoren. Die Verfahren sind aber ebenso auch in entsprechenden HIFU-Szenarien nutzbar. Die wesentlichen Teile dieses Kapitels sind unter maßgeblicher Beteiligung des Autors dieser Dissertation in [Wilms u. a., 2013; Wilms u. a., 2014a,c] veröffentlicht worden, wobei sich die Darstellung hauptsächlich an [Wilms u. a., 2014a] orientiert.

2.1 Hintergrund und Vorgehen

Wie in Kapitel 1 ausgeführt, basieren klinisch verfügbare Verfahren zur intrainerventionellen Schätzung des Atemzustandes bzw. der kompletten Atembewegung vielfach auf der Verwendung von sogenannten Bewegungsindikatoren (z. B. Bauchgurte), mit denen zumeist einfach extern beobachtbare Effekte der Atembewegung gemessen werden (indirekte Verfahren). Die zeitliche Veränderung des Indikatorsignals beschreibt dann die externe Atembewegung, die als Surrogat¹ der eigentlich zu bestimmenden internen Atembewegung dient. Um auf Basis dieses Surrogats die interne Bewegung zu schätzen, werden patientenspezifische Korrespondenzmodelle genutzt, die mathematisch den (angenommenen) Zusammenhang (z. B. linear) zwischen externer und interner Bewegung beschreiben. Der intrainerventionellen Nutzung eines Korrespondenzmodells geht eine entsprechende Trainingsphase voraus, in der die Parameter des Modells auf Basis von Trainingsdaten initial bestimmt werden (siehe Abb. 2.1 für eine Illustration). Als Trainingsdaten dienen im Kontext der Strahlentherapie atembewegter Tumoren für die Bestrahlungsplanung aufgenommene 4D-CT-Bilddaten² mit

¹Nachfolgend werden Surrogat und Bewegungsindikator im Kontext der Korrespondenzmodellierung als Synonyme verwendet.

²CT: Computertomographie

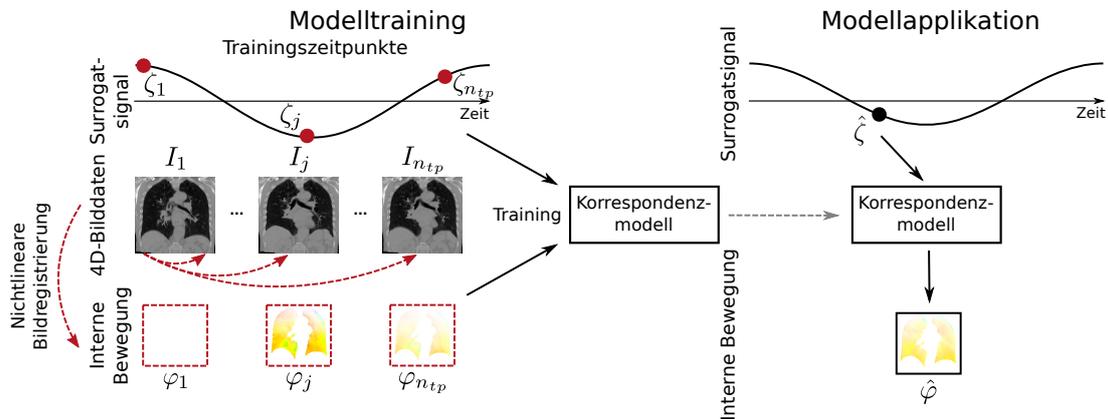


Abb. 2.1: Schematische Illustration von Training und Applikation eines Korrespondenzmodells zur surrogatbasierten Bewegungsschätzung. Links: Training des Korrespondenzmodells basierend auf Surrogat-Signalmessungen und der korrespondierenden internen Atembewegung (hier nur Lunge). Die interne Atembewegung wird durch dichte Verschiebungsfelder repräsentiert, die mittels nichtlinearer Bildregistrierung aus 4D-Bilddaten bestimmt wurden. Rechts: Intra-interventionelle Applikation des gelernten Korrespondenzmodells. Auf Basis der intrainerventionellen Surrogat-Signalmessung wird durch das Korrespondenzmodell die zugehörige interne Atembewegung geschätzt. Die verwendeten Formelzeichen werden im Text erläutert.

zugehörigen Surrogat-Signalmessungen. Mittels Bildregistrierung wird auf Basis der 4D-Bilddaten die interne Atembewegung geschätzt und mit den korrespondierenden Signalmessungen zum Training des Korrespondenzmodells genutzt.

2.1.1 Literaturübersicht

In den letzten 10–20 Jahren sind eine Vielzahl von Publikationen zur surrogatbasierten Bewegungsschätzung mithilfe von Korrespondenzmodellen entstanden. Eine umfassende Übersicht bezüglich methodischer Unterschiede verschiedener Ansätze ist in [McClelland u. a., 2013] zu finden. Eine grobe Unterscheidung der verschiedenen Verfahren bzw. Veröffentlichungen ist anhand der drei Kernkomponenten des jeweiligen Ansatzes zur Korrespondenzmodellierung möglich: (1) Das verwendete Surrogat, um die externe Atembewegung zu erfassen. (2) Die Parametrisierung der zu schätzenden internen Bewegung. (3) Die eigentliche mathematische Beschreibung des angenommenen Zusammenhangs zwischen externer und interner Bewegung.

2.1.1.1 Surrogate

Klinisch verfügbare Standardsurrogate sind u. a. Spirometriesysteme [Hoisak u. a., 2004; Zhang u. a., 2003], die mittels Masken die Menge der ein- bzw. ausgeatmeten Luft

messen, Bauchgurte [Li u. a., 2006; Otani u. a., 2010], welche die atmungsbedingte Expansion/Kontraktion des Bauches über Drucksensoren ermitteln und kamerabasierte Systeme zur Verfolgung der atmungsbedingten Bewegung auf der Patientenoberfläche angebrachter optischer Marker [Beddar u. a., 2007; Li u. a., 2006; Otani u. a., 2010; Ozhasoglu u. a., 2008] (siehe [Keall u. a., 2006] für eine erweiterte Übersicht und Diskussion). Nachteil von Spirometrysystemen und Bauchgurten ist, dass diese Surrogate nur ein eindimensionales Atemsignal liefern. Aufgrund der schon in Kapitel 1 diskutierten Variabilität der Atmung und der damit verbundenen Komplexität der internen Atembewegung erscheint deren präzise Schätzung anhand einer einzelnen Kenngröße schwierig. Es hat sich deshalb als günstig erwiesen, neben dem eigentlichen eindimensionalen Signal auch Informationen über den zeitlichen Verlauf der externen Bewegung (z. B. zeitliche Ableitung) bei der Korrespondenzmodellierung zu beachten, um so der Variabilität besser gerecht zu werden und beispielsweise zwischen Ein- und Ausatmung unterscheiden zu können [McClelland u. a., 2013].

Auch wenn optische Markersysteme vielfach in der Lage sind mehrdimensionale (multivariate) Surrogatsignale zu liefern und somit mehr Informationen für die Schätzung der internen Bewegung bereitzustellen als Spirometrie und Bauchgurte, stellt sich bei diesen Systemen immer die Schwierigkeit der korrekten Positionierung der Marker. Um Nachteile der Standardsurrogate (Eindimensionalität, Positionierung, teils fehlende Patiententoleranz) zu umgehen, hat in den letzten Jahren die Verwendung von markerlosen Kamerasystemen (z. B. Time-of-Flight- oder Stereokameras), die in der Lage sind, die gesamte relevante Patientenoberfläche (Thorax und Abdomen) in 2.5D bzw. 3D zu erfassen, verstärkt Aufmerksamkeit erhalten (siehe z. B. [Fassi u. a., 2014; Fayad u. a., 2012; McClelland u. a., 2011; Ortmüller u. a., 2015; Schaller u. a., 2008; Wasza u. a., 2016; Wentz u. a., 2012]). Ein Vorteil dieser Verfahren ist, dass mithilfe der pixelweise erfassten bzw. rekonstruierten Oberflächenbewegung auf einfache Art und Weise flexibel sehr hochdimensionale Surrogatsignale abgeleitet werden können, die die zu erwartende Variabilität der Atmung potenziell besser abbilden als eine begrenzte Menge an Markern [Fayad u. a., 2011].

Abschließend ist allerdings festzuhalten, dass es trotz der Fülle an Literatur zur Verwendung verschiedener Surrogate, keine wirklich aussagekräftigen Vergleichsstudien gibt, die mehrere klinisch relevante Surrogat-/Surrogattypen in einer kontrollierten Umgebung bei unterschiedlicher Bewegungsvariabilität vergleichen, um definitive Aussagen bezüglich deren Eignung machen zu können.

2.1.1.2 Interne Bewegung

Bei der Parametrisierung der internen Bewegung reichen die verwendeten Techniken von der rigiden und unabhängigen Beschreibung der Bewegung einzelner ausgezeichneter Punkte bis hin zur Verwendung von nichtlinearen Transformationen, die die komplexe Bewegung großer zusammenhängender Bereiche des Körpers beschreiben.

Die Verfolgung weniger ausgezeichneter Punkte ist derzeitiger klinischer Standard in kommerziell verfügbaren Systemen zur adaptiven Bestrahlung mittels Tumorverfolgung wie Cyberknife [Fu u. a., 2007; Schweikard u. a., 2000] und Vero [Depuydt u. a., 2013]. Bei diesen Systemen wird entweder die Bewegung des Tumors bzw. des Tumorschwerpunktes selbst oder die Bewegung von in dessen Nähe eingebrachter Marker verfolgt.

Bei der isolierten Verfolgung einzelner Punkt wird allerdings die mögliche Deformation des Tumors selbst und die potenziell komplexe Bewegung umgebender Risikoorgane ignoriert. Die Kenntnis der Tumordeformation und die Bewegung der Risikoorgane ist allerdings beispielsweise für eine möglichst genaue intrainventionelle Dosisakkumulation bei einer Strahlentherapie oder im Rahmen der räumlich-zeitlichen Temperaturbestimmung bei HIFU elementar [Tanner u. a., 2012]. Aus diesem Grund hat sich im Forschungskontext in den letzten 10 Jahren die Nutzung nichtlinearer Transformationen zur Repräsentation der internen Bewegung etabliert [McClelland u. a., 2013]. Für die Parametrisierung der nichtlinearen Transformationen werden dieselben Techniken eingesetzt, die auch bei der Parametrisierung nichtlinearer Transformationen in nichtlinearen Registrierungsverfahren eingesetzt werden (z. B. dichtes Verschiebungsfeld [Liu u. a., 2010], B-Splines [McClelland u. a., 2011], dichtes Geschwindigkeitsfeld [Geneser u. a., 2011]). Die für die Korrespondenzmodellierung gewählte Technik richtet sich oftmals nach dem Registrierungsverfahren, welches zur Bestimmung der internen Bewegung in den Trainingsdaten genutzt wird.

2.1.1.3 Korrespondenzmodellierung

Die mathematische Beschreibung des Zusammenhanges zwischen externer und interner Bewegung beeinflusst im Zusammenspiel mit der Dimension des genutzten Surrogatsignals (Freiheitsgrade des Modells) die Fähigkeit des Korrespondenzmodells die verschiedenen Arten von Atmungsvariabilität abzubilden (siehe Kapitel 1 und [McClelland u. a., 2013]). Ausführliche Übersichten über verschiedene mathematische Korrespondenzbeschreibungstechniken sind in [McClelland u. a., 2013] und [Dürichen, 2015] zu finden, an denen sich die folgende Zusammenfassung orientiert.

Nach Dürichen [2015] und McClelland u. a. [2013] ist die Annahme, dass die interne Bewegung als Linearkombination der Surrogatsignaldimensionen beschrieben werden kann (lineares Modell), der am häufigsten eingesetzte Modellierungsansatz (siehe z. B. [Cerviño u. a., 2010; Klinder u. a., 2010; Martin u. a., 2013; McClelland u. a., 2011; Zhang u. a., 2007]). Wird hierfür nur ein 1D-Surrogatsignal eingesetzt (siehe u. a. [Cerviño u. a., 2010]), limitiert dies die durch das Modell beschreibbare Atmungsvariabilität auf die Abdeckung unterschiedlicher Tiefen von Ein-/Ausatmung (einfache Skalierung). Eine Abbildung der Hysterese und anderer Variabilitäten ist somit ausgeschlossen, weshalb verschiedene Autoren dazu übergegangen sind, höherdimensionale Surrogatsignale für die Modellierung einzusetzen. Dies kann einerseits durch sogenannte Erweiterung

des Surrogatsignals um aus den eigentlich gemessenen Informationen abgeleitete Größen wie Vorgänger und zeitliche Ableitungen [Ruan u. a., 2008] oder durch die direkte Nutzung multivariater Signale erfolgen (siehe die Diskussion verschiedener Surrogatsignale weiter oben). Alternativ können auch zwei lineare Modelle genutzt werden, um die Hysterese zu modellieren: Jeweils ein Modell für Ein- bzw. Ausatmung (duales Modell; siehe z. B. Cyberknive [Sayeh u. a., 2007]).

Neben der Nutzung von linearen Modellen haben viele Autoren auch die Nutzung verschiedener parametrischer nichtlinearer oder nichtparametrischer Modelle vorgeschlagen und untersucht, um die Schätzungsgenauigkeit im Vergleich zu klassischen linearen Modellen zu verbessern. Die eingesetzten Techniken reichen hier von einfachen quadratischen Modellen [Ruan u. a., 2008; Sayeh u. a., 2007] über höhergradige Polynome [Ernst u. a., 2012; McClelland u. a., 2011; McClelland, 2008] und B-Splines [McClelland u. a., 2011] bis hin zu Support Vector Regression (SVR) [Ernst u. a., 2009, 2012], neuronalen Netzen [Seregni u. a., 2013] und Gaussian Process Models (GPM) [Dürichen u. a., 2015]. Nachteil dieser Modelle ist, dass mit wachsender Komplexität des Modells/des Modellierungsverfahrens und damit der potenziell höheren Genauigkeit oftmals auch die Komplexität des Trainingsprozesses stark erhöht wird und die Anzahl an benötigten Trainingsdaten ansteigt. Letzteres führt somit u. U. zu dem Problem auf der Basis von klinisch verfügbaren Daten (z. B. 4D-CT-Datensatz mit 10–14 Zeitpunkten) repräsentative Modelle erzeugen zu können. Beispielsweise nutzen Ernst u. a. [2012] für ihren SVR-Ansatz 20 Datenpunkte für das Training und Seregni u. a. [2013] für ihr neuronales Netz sogar 100 Datenpunkte.

Wie schon in [McClelland u. a., 2013] erwähnt, ist es schwierig auf Basis der verfügbaren Veröffentlichungen den besten Modellierungsansatz zu identifizieren, da die u. a. in [Dürichen u. a., 2015; Ernst u. a., 2012; McClelland u. a., 2011; Poels u. a., 2015; Ruan u. a., 2008; Seregni u. a., 2013] verfügbaren Evaluationen nur bestimmte Untermengen an Ansätzen abbilden, unterschiedliche/unterschiedlich große Trainings- und Testdatensammlungen einsetzen, unterschiedliche Testszenarien definieren (z. B. dichte Verschiebungsfelder vs. Tracking einzelner Punkte; Unterschiede abzubildender Atmungsvariabilität) und unterschiedliche Maße zur Quantifizierung der Schätzungsgenauigkeit einsetzen. Es lassen sich allerdings einige Tendenzen identifizieren (siehe auch [McClelland u. a., 2013]):

1. Die Verwendung eines einfachen linearen Modells mit 1D-Surrogatsignal ist schon von der theoretischen Betrachtung her zu einschränkend im Bezug auf die abbildbare Bewegungsvariabilität (siehe Diskussion in [McClelland u. a., 2013]), auch wenn es Studien gibt (u. a. [Poels u. a., 2015]), die keine klinisch relevanten Genauigkeitsunterschiede zwischen Modellen mit und ohne Hystereseberücksichtigung festgestellt haben.

2. Die Nutzung von nichtlinearen Modellen mit Polynomgrad ≥ 4 ist nicht sinnvoll, da es hier u. a. leicht zu einer Überanpassung an die Trainingsdaten kommt [McClelland, 2008].
3. Die Studien in [Dürichen u. a., 2015; Ernst u. a., 2012] zeigen, dass komplexe nicht-parametrische Korrespondenzmodelle basierend auf SVR oder GPM das Potenzial haben, deutlich höhere Genauigkeiten als klassische (nicht-)lineare Modelle zu erreichen, deren Nutzung aber auch mit einem erhöhten Trainingsaufwand einhergeht.

2.1.2 Einordnung des weiteren Vorgehens

Aufbauend auf der Kategorisierung der Kernkomponenten von Korrespondenzmodellierungsansätzen in Kapitel 2.1.1, soll nun das Framework aus [Werner, 2012], welches dieser Arbeit zugrunde liegt, entsprechend eingeordnet werden.

In [Werner, 2012] wird u. a. aufbauend auf den grundlegenden Arbeiten von Zhang u. a. [2007] ein lineares Korrespondenzmodell definiert, welches flexibel Surrogatsignale beliebiger Dimension verarbeiten kann und so die erweiterte Berücksichtigung verschiedener Typen von Atmungsvariabilität ermöglicht. Die interne Bewegung ganzer Körperregionen wird hierbei als nichtlineare Transformation repräsentiert, welche wiederum durch ein dichtes, stationäres Geschwindigkeitsfeld parametrisiert wird, um die Erzeugung von topologieerhaltenden diffeomorphen Transformationen zu garantieren. Die Fokussierung auf die Generierung von diffeomorphen Transformationen ist hierbei eine Besonderheit des Ansatzes, die sonst bisher nur in [Geneser u. a., 2011; Hinkle u. a., 2009] aufgegriffen wurde. In [Geneser u. a., 2011; Hinkle u. a., 2009] wird allerdings mit einem stückweise-linearen Modellierungsansatz und zeitabhängigen Geschwindigkeitsfeldern gearbeitet. Insbesondere Letzteres führt im Vergleich zur Nutzung von stationären Geschwindigkeitsfeldern zu einem stark erhöhtem Rechenaufwand. Die Beschränkung auf die Erstellung linearer Korrespondenzmodelle wird für diese Arbeit übernommen, da diese beim Einsatz multivariater Surrogatsignale einen guten Kompromiss zwischen erzielbarer Genauigkeit und Trainingsaufwand, sowohl zeitlich als auch bezüglich der Anzahl an benötigten Trainingsdaten, bieten.

In [Werner, 2012] wurde das Korrespondenzmodellierungsframework schon testweise mit (simulierten) multivariaten Surrogatsignalen (Dimension < 30) genutzt, wobei für das regressionsbasierte Modelltraining nur ein unspezifisches Regularisierungsverfahren eingesetzt wurde, um die Probleme (Unterbestimmtheit des zu lösenden Gleichungssystems, Informationsredundanz/Multikollinearität), die bei der regressionsbasierten Korrespondenzmodellierung mit sehr hochdimensionalen (multivariaten) Surrogatsignalen in der Praxis zwangsläufig auftreten, zu adressieren. In dieser Arbeit wird das Framework deshalb um Projektionsmethoden zur Regularisierung/zum Umgang mit hochdimensionalen Surrogatsignalen erweitert, welche datengetrieben arbeiten und somit potenziell eine spezifischere Regularisierung ermöglichen. Hierdurch

soll die effektive Nutzung von hochdimensionalen Tiefenbildern der Körperoberfläche des Patienten als Surrogatsignal ermöglicht und deren Eignung analysiert werden.

2.1.3 Notation

Die folgenden Abschnitte verwenden eine auf [Werner, 2012; Wilms u. a., 2014a; Wilms u. a., 2017c] basierende Notation, um die diffeomorphe Bildregistrierung und Korrespondenzmodellierung mathematisch zu beschreiben. Sei $(I_j)_{j=1}^{n_{tp}}$ ein aus n_{tp} 3D-Bildern $I_j : \Omega \rightarrow \mathbb{R}$ (mit $\Omega \subset \mathbb{R}^3$) bestehender räumlich-zeitlich aufgelöster 4D-Bilddatensatz, der für das Training eines Korrespondenzmodells genutzt wird. Jedes 3D-Bild I_j bildet hierbei die Anatomie des (frei-)atmenden Patienten zu einem Zeitpunkt j ab. Die zeitliche Auflösung der Daten ermöglicht demnach die Abbildung des patientenspezifischen Atemprozesses, wobei die genaue physiologische Bedeutung der einzelnen Zeitpunkte u. a. von der genutzten Bildgebungstechnik bzw. der konkreten Rekonstruktionstechnik abhängt. Für die weiteren Betrachtungen wird zudem ein Zeitpunkt (hier: $j = 1$) als Referenzzeitpunkt gewählt. Dieser Referenzzeitpunkt dient als Ausgangspunkt für die Definition der nichtlinearen diffeomorphen Transformationen $(\varphi_j : \Omega \rightarrow \Omega)_{j=1}^{n_{tp}}$, welche die (interne) atmungsbedingte Bewegung der Patientenanatomie zwischen Referenzbild I_1 und allen n_{tp} Zeitpunkten (inkl. $j = 1$) mathematisch beschreiben. Die Sequenz $(\zeta_j \in \mathbb{R}^{n_{sur}})_{j=1}^{n_{tp}}$ beschreibt zudem die n_{sur} -dimensionalen Surrogatsignale, deren Aufnahmezeitpunkte denen der korrespondierenden Bilder entsprechen. Siehe Abb. 2.1 für eine Illustration der Notation.

2.2 Diffeomorphe Bildregistrierung zur Schätzung der Atembewegung in 4D-Bilddaten

Dieser Abschnitt beschäftigt sich mit der Atembewegungsschätzung in den für das Modelltraining bereitgestellten 4D-Bilddaten mittels intensitätsbasierter Bildregistrierung. Der Prozess der Bildregistrierung bezeichnet hierbei die Suche bzw. Schätzung der in Kapitel 2.1 definierten diffeomorphen Transformationen φ_j , welche die jeweiligen Zielbilder I_j auf physiologisch plausible Weise auf das Referenzbild I_1 abbilden, sodass die Ähnlichkeit zwischen Referenz- und Zielbild maximiert wird ($I_1 \approx I_j \circ \varphi_j$).

Die Plausibilität der Transformation φ_j wird u. a. durch das gewählte Transformationsmodell bzw. den zulässigen Transformationsraum beeinflusst. Das hier präsentierte Framework setzt auf diffeomorphe Transformationen, die durch dichte Geschwindigkeitsfelder parametrisiert werden. Eine diffeomorphe Transformation $\varphi : \Omega \rightarrow \Omega$ ist eine bijektive, differenzierbare Abbildung, deren Inverse φ^{-1} ebenfalls differenzierbar ist. Weiterhin gilt, dass die Identitätsabbildung Id eine diffeomorphe Transformation ist. Gleiches gilt ebenso für die Komposition zweier diffeomorpher Transformationen und die Inverse einer diffeomorphen Transformation [Younes u. a., 2009]. Diese Eigenschaf-

ten implizieren u. a. den erwähnten Topologieerhalt (z. B. kein Zerreißen, Verschmelzen oder Verschwinden von Strukturen). Die Menge der diffeomorphen Transformationen über Ω sei im Folgenden als $\text{Diff}(\Omega)$ bezeichnet. Zusammen mit der Komposition \circ bildet sie eine Gruppe $(\text{Diff}(\Omega), \circ)$, deren neutrales Element die Identitätsabbildung Id ist.

Nach Beg u. a. [2005] kann eine diffeomorphe Transformation $\varphi \in \text{Diff}(\Omega)$ als Lösung der Transportgleichung

$$\frac{\partial}{\partial t} \phi_t(\mathbf{x}) = v(\phi_t(\mathbf{x}), t) \quad \text{mit} \quad \phi_0(\mathbf{x}) = \mathbf{x} \quad \text{und} \quad t \in [0, 1] , \quad (2.1)$$

zum Zeitpunkt $t = 1$ durch

$$\varphi(\mathbf{x}) = \phi_1(\mathbf{x}) = \phi_0(\mathbf{x}) + \int_0^1 v(\phi_t(\mathbf{x}), t) dt \quad (2.2)$$

beschrieben werden. Hierbei bezeichnet $\phi_t : \Omega \times [0, 1] \rightarrow \Omega$ ausgehend von der Identitätsabbildung $\phi_0 = Id$ einen Pfad, der durch das zeitabhängige und glatte Geschwindigkeitsfeld $v : \Omega \times [0, 1] \rightarrow \mathbb{R}^3$ parametrisiert wird [Hernandez, 2008]. Für nähere Informationen zur Glattheit von v sei der Leser auf [Beg u. a., 2005] verwiesen. Die Konstruktion diffeomorpher Transformationen auf der Basis zeitabhängiger Geschwindigkeitsfelder (siehe Glg. 2.1) ist der Standardweg bei diffeomorphen Registrierungsverfahren, die auf dem Large Deformation Diffeomorphic Metric Mapping (LDDMM)-Framework nach Beg u. a. [2005] aufbauen. Nachteil dieser Verfahren ist ihr relativ hoher Speicher- und Zeitbedarf, der durch die Zeitabhängigkeit der Geschwindigkeitsfelder begründet ist, auch wenn neuere Arbeiten in diesem Bereich (u. a. [Ashburner u. a., 2011; Polzin u. a., 2016]) versuchen dieses Problem zu lösen.

Ein alternativer Weg ist die Parametrisierung von Diffeomorphismen über stationäre (zeitunabhängige) Geschwindigkeitsfelder $v : \Omega \rightarrow \mathbb{R}^3$, wie u. a. in [Arsigny, 2006] vorgeschlagen:

$$\frac{\partial}{\partial t} \phi_t(\mathbf{x}) = v(\phi_t(\mathbf{x})) \quad \text{mit} \quad \phi_0(\mathbf{x}) = \mathbf{x} \quad \text{und} \quad t \in [0, 1] . \quad (2.3)$$

In [Arsigny, 2006] wird zur Lösung von Glg. 2.3 für $t = 1$, $(\text{Diff}(\Omega), \circ)$ als Lie-Gruppe aufgefasst. Dies bedeutet u. a., dass $\text{Diff}(\Omega)$ eine glatte Mannigfaltigkeit ist und es eine zugehörige Lie-Algebra \mathfrak{g} gibt. Die Lie-Algebra \mathfrak{g} ist ein Vektorraum, der in diesem Fall isomorph zum Tangentialraum $\mathcal{T}_{Id}\text{Diff}(\Omega)$ der Lie-Gruppe $(\text{Diff}(\Omega), \circ)$ am neutralen Element Id ist [Ehrhardt u. a., 2011; Hernandez, 2008; Schmidt-Richberg, 2013]. Zudem gilt $v \in \mathcal{T}_{Id}\text{Diff}(\Omega)$. Lie-Algebra und Lie-Gruppe sind durch die Gruppenexponentialabbildung

$$\exp : \mathcal{T}_{Id}\text{Diff}(\Omega) \rightarrow \text{Diff}(\Omega) \quad (2.4)$$

und den Logarithmus

$$\log : \text{Diff}(\Omega) \rightarrow \mathcal{T}_{Id}\text{Diff}(\Omega) \quad (2.5)$$

verbunden [Arsigny, 2006; Hernandez, 2008; Schmidt-Richberg, 2013]. Darüber hinausgehend sind die Pfade, die sich durch Glg. 2.3 ergeben, Ein-Parameter-Untergruppen der Lie-Gruppe, welche über die Exponentialabbildung in Glg. 2.4 durch $\phi_t = \exp(tv)$ beschrieben werden können [Arsigny, 2006; Hernandez, 2008]. Demnach ergibt sich die Lösung von Glg. 2.3 für $t = 1$ durch $\varphi = \phi_1 = \exp(v)$. Auf Basis dieses Zusammenhangs und der Scaling-and-Squaring-Methode zur Berechnung der Exponentialabbildung von Vektorfeldern [Arsigny, 2006] lassen sich sehr laufzeit- und speichereffiziente diffeomorphe Registrierungsalgorithmen realisieren (u. a. [Schmidt-Richberg u. a., 2014; Vercauteren u. a., 2009]). Weiterhin lässt sich durch $\varphi^{-1} = \exp(-v)$ auf einfache Weise die inverse Transformation φ^{-1} zu φ erzeugen.

Es ist anzumerken, dass die hier vorstehende Beschreibung der Parametrisierung von Diffeomorphismen durch stationäre Geschwindigkeitsfelder einige mathematische Details (z. B. der Unterschied zwischen endlich- und unendlich-dimensionalen Lie-Gruppen) auslöst und u. a. bestehende theoretische Probleme nicht diskutiert, da sie für die folgende praktische Nutzung des Ansatzes keine wesentliche Rolle spielen. Eine Ausnahme stellt hierbei der sich direkt aus der Beschränkung auf Ein-Parameter-Untergruppen ergebende theoretische Nachteil dar, dass deshalb gegenüber zeitabhängigen LDDMM-Varianten weniger Diffeomorphismen parametrisiert werden können. Bisherige Untersuchungen u. a. von Hernandez [2008], Vercauteren u. a. [2009] und Lorenzi u. a. [2013] haben allerdings gezeigt, dass dieser theoretische Unterschied keine Auswirkung auf die durch stationäre Geschwindigkeitsfelder parametrisierte Registrierungsverfahren erreichbare Genauigkeit hat. Lorenzi u. a. [2013] argumentieren zudem, dass es nicht wichtig ist, alle Diffeomorphismen generieren zu können, sondern im Kontext der medizinischen Bildregistrierung nur anatomisch bzw. physiologisch plausible Transformationen überhaupt als gewünschte Lösung infrage kommen und Effekte der praktischen Implementierung (z. B. räumliche oder zeitliche Diskretisierung der Felder, Glattheit der Felder, usw.) große Unterschiede in vielen Anwendungen verhindern. Für weitergehende Informationen zu mathematischen Details beider Ansätze und entsprechender Probleme bzw. Nachteile sei der interessierte Leser u. a. auf [Arsigny, 2006; Beg u. a., 2005; Ehrhardt u. a., 2011; Hernandez, 2008; Lorenzi u. a., 2013] verwiesen.

Unter Ausnutzung des Zusammenhangs $\varphi_j = \exp(v_j)$ lässt sich nun die gesuchte optimale diffeomorphe Transformation φ_j zwischen den Bildern I_1 und I_j durch Minimierung des Energiefunktional

$$\mathcal{J}[v_j] = \mathcal{D}[I_1, I_j \circ \varphi_j] + \alpha \mathcal{S}[v_j] \quad (2.6)$$

bestimmen [Werner u. a., 2014]. Hierbei parametrisiert das stationäre Geschwindigkeitsfeld v_j die gesuchte diffeomorphe Transformation φ_j , \mathcal{D} quantifiziert als Ähnlichkeitsmaß die (Un)Ähnlichkeit zwischen I_1 und $I_j \circ \varphi_j$ und \mathcal{S} ist ein Regularisierungsterm, der für die geforderte Glattheit des Geschwindigkeitsfeldes sorgt und entscheidend zur Plausibilität der geschätzten Transformation beiträgt.

Zur praktischen Lösung des Registrierungsproblems nach Glg. 2.6 wird in dieser Arbeit, wenn nicht anders angegeben, auf die am Institut für Medizinische Informatik der Universität zu Lübeck entwickelte und frei verfügbare *VariationalRegistration-Toolbox* zurückgegriffen [Schmidt-Richberg u. a., 2014; Werner u. a., 2014]. Diese Toolbox ist sehr flexibel aufgebaut und erlaubt die Kombination diverser Ähnlichkeitsmaße und Regularisierer in einem variationellen Framework, in dem Glg. 2.6 durch Lösung der korrespondierenden Euler-Lagrange-Gleichungen iterativ mittels Gradientenabstieg minimiert wird. Für eine umfangreiche Untersuchung der durch Kombination unterschiedlicher Ähnlichkeitsmaße und Regularisierungsterme mit dieser Toolbox erreichbaren Registrierungsgenauigkeit bei der patientenspezifischen Bewegungsschätzung in 4D-CT-Bilddaten sei auf [Werner u. a., 2014] verwiesen.

2.2.1 Gütemaße für die Evaluation diffeomorpher Bildregistrierungsverfahren

Da die Genauigkeit und physiologische Plausibilität einer nach Glg. 2.6 bestimmten Transformation φ_j direkten Einfluss auf die nachfolgenden Verfahren hat, kommt der Evaluation des Registrierungsverfahrens bzw. der von ihm geschätzten Transformationen in der Praxis eine enorme Bedeutung zu. Im Laufe der Zeit wurden im Bereich der Evaluation von nichtlinearen Registrierungsverfahren zur Schätzung der Atembewegung verschiedenste Gütemaße vorgeschlagen (siehe Übersicht in [Kabus u. a., 2013]), wobei sich die folgende Darstellung auf drei ausgewählte Maße beschränkt, die in dieser Arbeit praktisch zum Einsatz kommen und in der Literatur weitverbreitet sind.

Target Registration Error Als Standard zur Genauigkeitsevaluation in größeren Vergleichsstudien (siehe z. B. [Murphy u. a., 2011; Werner u. a., 2014]) hat sich die Nutzung von (manuell gesetzten) korrespondierenden Landmarken als Grundwahrheit durchgesetzt. Eine Landmarke ist hierbei typischerweise eine anatomisch oder pathologisch eindeutig bestimmbare Position (Gefäßbaum-bifurkationen, Tumorzentren, usw.). Zur landmarkenbasierten Evaluation der Transformation φ_j werden dann in Referenzbild I_1 und Zielbild I_j insgesamt n_{lmk} Tupel korrespondierender Landmarken $(\mathbf{l}_{I_1}^i, \mathbf{l}_{I_j}^i)$ mit $i \in \{1, \dots, n_{lmk}\}$ und $\mathbf{l}_{I_1}^i, \mathbf{l}_{I_j}^i \in \Omega$ bestimmt. Der Target Registration Error (TRE) ist dann definiert als der euklidische Abstand zwischen transformierter Referenzbildlandmarke und der korrespondierenden Landmarke im Zielbild [Kabus u. a., 2013]:

$$\text{TRE}(\mathbf{l}_{I_1}^i, \mathbf{l}_{I_j}^i, \varphi_j) = \|\varphi_j(\mathbf{l}_{I_1}^i) - \mathbf{l}_{I_j}^i\|_2 \quad \text{und} \quad (2.7)$$

$$\overline{\text{TRE}}\left((\mathbf{l}_{I_1}^i, \mathbf{l}_{I_j}^i)_{i=1}^{n_{lmk}}, \varphi_j\right) = \frac{1}{n_{lmk}} \sum_{i=1}^{n_{lmk}} \text{TRE}(\mathbf{l}_{I_1}^i, \mathbf{l}_{I_j}^i, \varphi_j) . \quad (2.8)$$

Da das Ergebnis einer einzelnen Landmarke wenig Aussagekraft hat, wird bei Evaluationen oft der mittlere TRE ($\overline{\text{TRE}}$) über alle n_{lmk} Landmarken angegeben. In dieser Arbeit werden die Begriffe mittlerer TRE und TRE zumeist synonym benutzt.

Symmetrische Oberflächendistanz Ist die Identifikation von geeigneten Landmarkenpunkten beispielsweise aufgrund geringer Bildauflösung nicht möglich, bietet es sich an mit den Oberflächen in den Bildern (manuell) segmentierter, korrespondierender anatomischer oder pathologischer Strukturen (Organe, Knochen, Tumoren, usw.) als Grundwahrheit zu arbeiten, um die Güte der Transformation φ_j zu evaluieren. Seien hierfür $S_{I_1} = \{\mathbf{x}_{I_1}^i \in \Omega\}_{i=1}^{n_{I_1}}$ und $S_{I_j} = \{\mathbf{x}_{I_j}^i \in \Omega\}_{i=1}^{n_{I_j}}$ die Mengen der Rand-/Oberflächen voxel, die die segmentierte Struktur in Referenzbild I_1 und Zielbild I_j begrenzen. Werden die Voxel aus dem Referenzbild S_{I_1} dann mittels $\varphi_j(S_{I_1}) = \{\varphi_j(\mathbf{x}_{I_1}^i)\}_{i=1}^{n_{I_1}}$ in das Zielbild abgebildet, kann auf Basis dieser transformierten Voxel die mittlere symmetrische Oberflächendistanz zwischen $\varphi_j(S_{I_1})$ und S_{I_j} als

$$\begin{aligned} \bar{D}(\varphi_j(S_{I_1}), S_{I_j}) &= \frac{1}{2} \left(\frac{1}{n_{I_1}} \sum_{i=1}^{n_{I_1}} \min_{\mathbf{x} \in S_{I_j}} \|\varphi_j(\mathbf{x}_{I_1}^i) - \mathbf{x}\|_2 \right) \\ &\quad + \frac{1}{2} \left(\frac{1}{n_{I_j}} \sum_{i=1}^{n_{I_j}} \min_{\mathbf{x} \in \varphi_j(S_{I_1})} \|\mathbf{x}_{I_j}^i - \mathbf{x}\|_2 \right) \end{aligned} \quad (2.9)$$

berechnet werden [Kabus u. a., 2013]. Hierfür wird zunächst für jeden Punkt der einen Oberfläche der nach euklidischer Distanz nächste Punkt auf der anderen Oberfläche bestimmt, um die einseitigen mittleren Oberflächendistanzen zu ermitteln. Die einseitigen Distanzen werden anschließend durch Mittelwertbildung symmetrisiert.

Jacobi-Determinante Die zuvor definierten Gütemaße ermöglichen zwar die Quantifizierung der Genauigkeit der Bewegungsschätzung, geben allerdings keinen Aufschluss über die physiologische Plausibilität der ermittelten Transformation. Als physiologisch unplausibel kann eine Transformation φ_j , die die Atembewegung beschreibt, u. a. dann angesehen werden, wenn sie nicht bijektiv ist bzw. lokale Faltungen (Orientierungswechsel) enthält [Kabus u. a., 2013; Murphy u. a., 2011]. Beides sind Eigenschaften, die bei der Einschränkung auf Diffeomorphismen ausgeschlossen sein sollten und deren Überprüfung mithilfe der Determinante der Jacobi-Matrix von φ_j möglich ist. Die Jacobi-Determinante lässt sich pro Voxel $\mathbf{x} := (x, y, z)^T \in \Omega$ durch

$$\text{Jac}(\varphi_j, \mathbf{x}) = \det(\nabla \varphi_j(\mathbf{x})) = \begin{vmatrix} \frac{\partial \varphi_{j,x}(\mathbf{x})}{\partial x} & \frac{\partial \varphi_{j,x}(\mathbf{x})}{\partial y} & \frac{\partial \varphi_{j,x}(\mathbf{x})}{\partial z} \\ \frac{\partial \varphi_{j,y}(\mathbf{x})}{\partial x} & \frac{\partial \varphi_{j,y}(\mathbf{x})}{\partial y} & \frac{\partial \varphi_{j,y}(\mathbf{x})}{\partial z} \\ \frac{\partial \varphi_{j,z}(\mathbf{x})}{\partial x} & \frac{\partial \varphi_{j,z}(\mathbf{x})}{\partial y} & \frac{\partial \varphi_{j,z}(\mathbf{x})}{\partial z} \end{vmatrix} \quad (2.10)$$

berechnen [Rey u. a., 2002; Werner, 2012]. Hierbei gibt $\varphi_{j,x}$ die Transformation in x -Richtung an. Der Wert von $\text{Jac}(\varphi_j, \mathbf{x})$ gibt nun Auskunft über die lokalen physikalischen Auswirkungen der Transformation auf die betrachtete Struktur. Bei $\text{Jac}(\varphi_j, \mathbf{x}) = 1$ kommt es zu einem Volumenerhalt, wohingegen $0 < \text{Jac}(\varphi_j, \mathbf{x}) < 1$ eine Kontraktion und $\text{Jac}(\varphi_j, \mathbf{x}) > 1$ eine Expansion anzeigt. Für die hier angestrebte Plausibilitätsprüfung entscheidender sind die beiden anderen Situationen: Bei $\text{Jac}(\varphi_j, \mathbf{x}) = 0$ werden

durch die Transformation mehrere Punkte auf einen einzigen Punkt abgebildet und bei $\text{Jac}(\varphi_j, \mathbf{x}) < 0$ kommt es zu einer lokalen Änderung der Orientierung (Faltung). Wie eingangs beschrieben, sind diese beiden Eigenschaften physiologisch unplausibel und zeigen außerdem an, dass es sich nicht um eine diffeomorphe Transformation handelt, da die Transformation nicht invertierbar ist. In [Murphy u. a., 2011] wird das Auftreten von $\text{Jac}(\varphi_j, \mathbf{x}) \leq 0$ deshalb als Singularität der Transformation bezeichnet. Als eigentliches Gütemaß wird dann die Anzahl von Voxeln mit Singularität verwendet, wobei 0 (keine Singularitäten vorhanden) der angestrebte Wert ist.

2.3 Regressionsbasierte diffeomorphe Korrespondenzmodellierung

Basierend auf den geschätzten diffeomorphen Transformationen $(\varphi_j)_{j=1}^{n_{tp}}$ und den zeitlich korrespondierenden Surrogatsignalmessungen $(\zeta_j)_{j=1}^{n_{tp}}$, soll nun mittels multivariater linearer Regression ein diffeomorphes Korrespondenzmodell geschätzt werden, welches den angenommenen linearen Zusammenhang zwischen interner Atembewegung und Surrogatsignal mathematisch beschreibt. Das u. a. in [Klinder u. a., 2010] beschriebene Standardvorgehen zur Generierung von Korrespondenzmodellen basiert hierbei auf der Parametrisierung der nichtlinearen Transformationen φ_j mithilfe von Verschiebungsfeldern $u_j : \Omega \rightarrow \mathbb{R}^3$ ($\varphi_j = Id + u_j$). Die Verschiebungsfelder werden dabei als Elemente eines Vektorraums aufgefasst, um normale euklidische/lineare Statistik für die Schätzung der Parameter des Regressionsmodells anwenden zu können.

Da das Ziel hier allerdings die Erstellung eines diffeomorphen Korrespondenzmodells ist, welches nur diffeomorphe Transformationen (Elemente von $\text{Diff}(\Omega)$) generiert, muss beachtet werden, dass $\text{Diff}(\Omega)$ nicht abgeschlossen bezüglich Addition ist und demnach kein Vektorraum ist [Avants u. a., 2010; Durrleman, 2010; Joshi u. a., 2004]. Die Addition zweier Verschiebungsfelder, die diffeomorphe Transformationen parametrisieren, führt demnach nicht zwingend wieder zu einer diffeomorphen Transformation [Vercauteren u. a., 2009]. Entsprechend ist die Nutzung euklidischer/linearer Statistik auf Basis der Verschiebungsfelder keine valide Option zur Erzeugung eines diffeomorphen Korrespondenzmodells. Vielmehr muss zu diesem Zweck die Struktur von $\text{Diff}(\Omega)$ beachtet werden (glatte Mannigfaltigkeit; siehe Kapitel 2.2), um garantieren zu können, dass das Modell nur diffeomorphe Transformationen generiert.

Statistische Berechnungen auf allgemeinen Riemannschen Mannigfaltigkeiten können effizient durch die Tangentialraumrepräsentationen der beteiligten Elemente durchgeführt werden (siehe u. a. [Freifeld, 2014; Pennec u. a., 2015]). Exponentialabbildung und Logarithmus sorgen für den Transport zwischen Tangentialraum und Mannigfaltigkeit. Dasselbe Grundprinzip lässt sich auch auf mittels stationärer Geschwindigkeitsfelder parametrisierte diffeomorphe Transformationen anwenden. Basierend auf der in Kapitel 2.2 eingeführten Lie-Gruppe $(\text{Diff}(\Omega), \circ)$ lassen sich zwei diffeomor-

Die Transformationen $\varphi_i = \exp(v_i)$ und $\varphi_j = \exp(v_j)$ mittels Exponentialabbildung und Logarithmus (siehe Glg. 2.4 und Glg. 2.5) zwischen $\text{Diff}(\Omega)$ und Tangentialraum $\mathcal{T}_{Id}\text{Diff}(\Omega)$ transferieren.

Statistische Betrachtungen setzen weiterhin die Definition einer Distanz zwischen zwei Diffeomorphismen voraus, die hier auf Basis der euklidischen Norm von $\mathcal{T}_{Id}\text{Diff}(\Omega)$ erfolgt [Arsigny, 2006]:

$$d(\varphi_i, \varphi_j) := \|\log(\varphi_i) - \log(\varphi_j)\| = \|v_i - v_j\|. \quad (2.11)$$

Auf Basis dieser Distanzdefinition lassen sich entsprechend Operationen/Methoden der klassischen euklidischen Statistik auf Diffeomorphismen anwenden, indem hierfür deren Logarithmen (Geschwindigkeitsfelder) genutzt werden, was auch Log-Euklidische Statistik bzw. Log-Euklidisches-Framework genannt wird [Arsigny, 2006]. Beispielsweise kann die mittlere Log-Euklidische-Transformation von φ_i und φ_j mittels $\bar{\varphi} = \exp(\frac{1}{2}\log(\varphi_i) + \frac{1}{2}\log(\varphi_j))$ berechnet werden, wobei $\bar{\varphi}$ selbst wieder Element von $\text{Diff}(\Omega)$ ist.

Zusammenfassend lässt sich festhalten, dass durch Nutzung des Log-Euklidischen-Frameworks statistische Berechnungen im Tangentialraum am neutralen Element Id vorgenommen werden und das Resultat zurück auf die Mannigfaltigkeit abgebildet wird. Hierauf aufbauend lässt sich dann ein diffeomorphes Korrespondenzmodell definieren, wobei hervorzuheben ist, dass die Logarithmierung der Transformationen nicht explizit durchgeführt werden muss, da diese Teil des Ergebnisses des in Kapitel 2.2 vorgestellten Registrierungsverfahrens sind.

Im weiteren Verlauf dieses Abschnitts werden die Geschwindigkeitsfelder v_j der diffeomorphen Transformationen φ_j und die Surrogatsignale ζ_j als Spaltenvektoren repräsentiert: $\mathbf{v}_j = \text{vec}(v_j) \in \mathbb{R}^{3n_{vx}}$ und $\mathbf{z}_j = \zeta_j \in \mathbb{R}^{n_{sur}}$. Hierbei bezeichnet $\text{vec}(\cdot)$ einen Vektorisierungsoperator, der das Vektorfeld v_j in einen Spaltenvektor umwandelt und n_{vx} die Anzahl der Voxel des zugehörigen Bildes I_j .

Ziel ist es nun, basierend auf einer Surrogatsignalmessung $\hat{\mathbf{z}}$ den korrespondierenden Geschwindigkeitsfeldvektor $\hat{\mathbf{v}}$ (und damit auch $\hat{\varphi} = \exp(\hat{\mathbf{v}})$) durch ein lineares Modell zu bestimmen, welches anhand der Trainingsdaten $(\mathbf{v}_j)_{j=1}^{n_{tp}}$ und $(\mathbf{z}_j)_{j=1}^{n_{tp}}$ trainiert wurde. Es handelt sich hierbei um ein klassisches lineares Regressionsproblem. Die Basisform des linearen³, diffeomorphen Korrespondenzmodells wird hierfür definiert als

$$\hat{\mathbf{v}} = \bar{\mathbf{v}} + \mathbf{B}(\hat{\mathbf{z}} - \bar{\mathbf{z}}) \quad \text{mit } \bar{\mathbf{v}} = \frac{1}{n_{tp}} \sum_{j=1}^{n_{tp}} \mathbf{v}_j \quad \text{und } \bar{\mathbf{z}} = \frac{1}{n_{tp}} \sum_{j=1}^{n_{tp}} \mathbf{z}_j. \quad (2.12)$$

Die Matrix $\mathbf{B} \in \mathbb{R}^{3n_{vx} \times n_{sur}}$ repräsentiert hier die während des Trainings zu schätzenden Regressionskoeffizienten. Für deren Schätzung werden die Trainingsdaten zuerst mittelwertbereinigt ($\mathbf{v}_j^c = \mathbf{v}_j - \bar{\mathbf{v}}$ und $\mathbf{z}_j^c = \mathbf{z}_j - \bar{\mathbf{z}}$) und in Matrizen $\mathbf{V} = [\mathbf{v}_1^c \cdots \mathbf{v}_{n_{tp}}^c] \in$

³Die Bezeichnung linear bezieht sich hier streng genommen nur auf den Tangentialraum $\mathcal{T}_{Id}\text{Diff}(\Omega)$.

$\mathbb{R}^{3n_{vx} \times n_{tp}}$ (Regressanden) und $\mathbf{Z} = [\mathbf{z}_1^c \cdots \mathbf{z}_{n_{tp}}^c] \in \mathbb{R}^{3n_{vx} \times n_{tp}}$ (Regressoren) zusammengeführt. Die Schätzung der Koeffizienten erfolgt dann durch Minimierung der Summe der quadrierten Residuen/Fehlerquadrate der Trainingsdaten mit zusätzlicher Frobeniusnorm-Regularisierung der Koeffizientenmatrix (Thikonov-Regularisierung, Ridge-Regression; RR), wodurch sich für die Lösung eine geschlossene Form ergibt [Hastie u. a., 2009]:

$$\mathbf{B}^{\text{RR}} = \arg \min_{\mathbf{B}} \|\mathbf{V} - \mathbf{B}\mathbf{Z}\|_F^2 + \lambda \|\mathbf{B}\|_F^2 = \mathbf{V}\mathbf{Z}^T (\mathbf{Z}\mathbf{Z}^T + \lambda \mathbf{I})^{-1} = \mathbf{V} (\mathbf{Z}^T \mathbf{Z} + \lambda \mathbf{I})^{-1} \mathbf{Z}^T . \quad (2.13)$$

Hierbei ist $\lambda \in \mathbb{R}_{\geq 0}$ der zu wählende Regularisierungsparameter (größere Werte entsprechen einer stärkeren Regularisierung) und \mathbf{I} eine Einheitsmatrix. Die Frobeniusnorm-Regularisierung hat im Vergleich zur unregularisierten Variante mit $\lambda = 0$ neben der generell weniger starken Anfälligkeit für Überanpassungen u. a. auch den Vorteil, dass die (numerisch stabile) Invertierbarkeit der Streuungsmatrix $\mathbf{\Sigma}_{\mathbf{Z}\mathbf{Z}} = \mathbf{Z}\mathbf{Z}^T$ bzw. $\mathbf{Z}^T \mathbf{Z}$ gegeben ist, auch wenn diese keinen vollen Rang aufweist, oder zumindest nahezu singular ist. Dieses Problem kann beispielsweise im Kontext dieser Arbeit auch im überbestimmten Fall ($n_{tp} > n_{sur}$) mit $\mathbf{\Sigma}_{\mathbf{Z}\mathbf{Z}}$ relativ schnell bei der Verwendung von multivariaten Surrogatsignalen durch die starke Korrelation einzelner Signaldimensionen (die sog. Multikollinearität) auftreten. Diese Korrelationen sind insbesondere bei hochdimensionalen Signalen wie Tiefenbildern der Körperoberfläche überaus wahrscheinlich, da räumlich benachbarte Körperregionen (= einzelne Signaldimensionen) relativ ähnliche Bewegungen aufweisen. Letztendlich würde die starke Korrelation von Regressoren zu einer verminderten Präzision der Schätzung der Koeffizientenmatrix führen und damit die Generalisierungsfähigkeit des Modells beeinträchtigen [Hastie u. a., 2009].

2.4 Erweiterung des Frameworks durch projektionsbasierte Regressionsmethoden

Das in Kapitel 2.3 vorgestellte Framework entspricht größtenteils der Formulierung in [Werner, 2012] und erlaubt durch die Verwendung der regularisierten Regression auch die Nutzung von multivariaten Surrogatsignalen mit korrelierten Signaldimensionen. Nachteile des vorgestellten Verfahrens sind allerdings u. a. die nichttriviale Wahl des Regularisierungsparameters λ und die fehlende Berücksichtigung der Struktur der durch die Regressorenmatrix \mathbf{Z} repräsentierten Surrogatdaten durch simple Addition der gewichteten Einheitsmatrix. Aus diesem Grund wurde von verschiedenen Autoren vorgeschlagen, projektionsbasierte Regressionsmethoden einzusetzen, um $\mathbf{\Sigma}_{\mathbf{Z}\mathbf{Z}}$ zu berechnen [Gao u. a., 2008; Klinder u. a., 2010; Liu u. a., 2010; McClelland u. a., 2013]. Generell deutet das Auftreten von Multikollinearitäten in einem multivariaten Surrogatsignal darauf hin, dass es eine Redundanz in den Signaldimensionen gibt und der

darunterliegende Prozess eine geringere Anzahl von Parametern hat. Projektionsbasierte Regressionsmethoden entfernen diese Redundanz durch Projektion der hochdimensionalen Surrogatdaten auf einen niedrigdimensionalen linearen Unterraum, sodass $\Sigma_{\mathbf{Z}\mathbf{Z}}$ invertiert werden kann (siehe Abb. 2.2 für eine Illustration des Unterschiedes zwischen projektionsbasierter und klassischer Regression).

Sei $\mathbf{U}_k \in \mathbb{R}^{n_{sur} \times k}$ eine Matrix, deren orthonormale Spaltenvektoren eine Orthonormalbasis eines k -dimensionalen Unterraums des $\mathbb{R}^{n_{sur}}$ darstellen. Dann lässt sich die Streuungsmatrix $\Sigma_{\mathbf{Z}\mathbf{Z}}$ durch Projektion der Surrogatdaten \mathbf{Z} auf diesen Unterraum ($\mathbf{Z} \approx \mathbf{U}_k \mathbf{U}_k^T \mathbf{Z}$) approximieren als

$$\Sigma_{\mathbf{Z}\mathbf{Z}} \approx \mathbf{C}_k = \mathbf{U}_k \mathbf{U}_k^T \mathbf{Z} \mathbf{Z}^T \mathbf{U}_k \mathbf{U}_k^T . \quad (2.14)$$

Sei $\mathbf{U}_k^+ = (\mathbf{U}_k^T \mathbf{U}_k)^{-1} \mathbf{U}_k^T = \mathbf{U}_k^T$ die Pseudoinverse von \mathbf{U}_k , dann lässt sich basierend auf Glg. 2.14 eine Pseudoinverse

$$\mathbf{C}_k^+ = \mathbf{U}_k (\mathbf{U}_k^T \mathbf{Z} \mathbf{Z}^T \mathbf{U}_k)^{-1} \mathbf{U}_k^T \quad (2.15)$$

bilden. Es geht nun bei der projektionsbasierten Regression zunächst darum, einen k -dimensionalen Unterraum bzw. eine Matrix \mathbf{U}_k zu bestimmen, sodass die Redundanz in den Daten entfernt wird und hierdurch die Streuungsmatrix $\mathbf{U}_k^T \mathbf{Z} \mathbf{Z}^T \mathbf{U}_k$, ausgedrückt in Koordinaten der Orthonormalbasis des Unterraums, vollen Rang hat bzw. problemlos invertiert werden kann [Blanc u. a., 2012b]. Weiterhin kann über die Wahl der Dimension k des Unterraums der Grad der Regularisierung bestimmt werden (niedriger Wert für k entspricht starker Regularisierung). Im Folgenden werden drei verschiedene Methoden zur Bestimmung von \mathbf{U}_k in das Framework zur diffeomorphen Korrespondenzmodellierung integriert. Deren Beschreibungen folgen den Definitionen und Darstellungen der Verfahren in [Blanc u. a., 2012b].

2.4.1 Hauptkomponentenregression (PCR)

Eine einfache und gängige Möglichkeit zur Bestimmung eines Unterraums, sodass die (approximierte) Streuungsmatrix $\Sigma_{\mathbf{Z}\mathbf{Z}}$ invertierbar ist, stellt die Anwendung der Hauptkomponentenanalyse (engl. Principal Component Analysis bzw. PCA) auf die Surrogatdaten dar, was in einer Hauptkomponentenregression (engl. Principal Component Regression bzw. PCR) resultiert. Die mittels PCA ermittelte Basis spannt einen linearen Unterraum maximaler Variation der Daten auf, in dem die Datendimensionen dekorelliert sind. Neben der Gewährleistung der Invertierbarkeit der Streuungsmatrix, ist zudem durch zusätzliche Beschränkung der Dimension des Unterraums eine weitere Regularisierung des Regressionsproblems möglich.

Praktisch umgesetzt wird die PCA zunächst durch Singulärwertzerlegung (SVD) der Regressorenmatrix $\mathbf{Z} = \mathbf{U}\mathbf{S}\mathbf{W}^T$ in die unitären Matrizen $\mathbf{U} \in \mathbb{R}^{n_{sur} \times n_{sur}}$ und $\mathbf{W} \in \mathbb{R}^{n_{tp} \times n_{tp}}$, deren Spaltenvektoren (Links- bzw. Rechts-)Singulärvektoren genannt

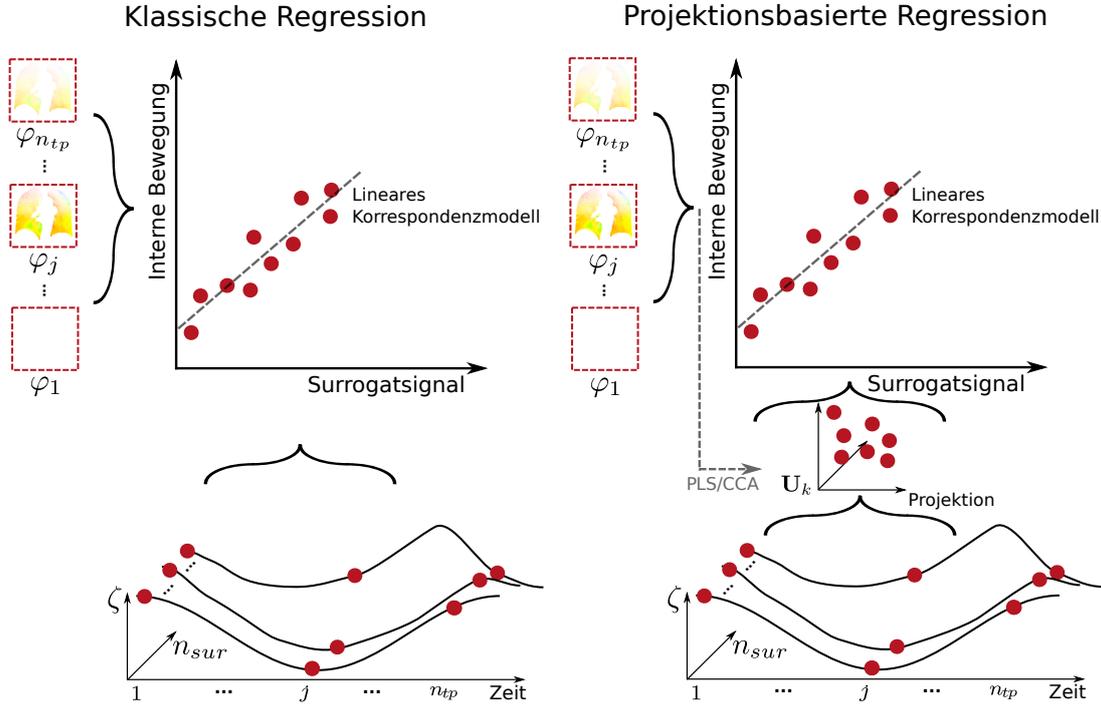


Abb. 2.2: Schematische Illustration des Unterschiedes zwischen klassischer Regression (links) und projektionsbasierter Regression (rechts) für das Training eines linearen Korrespondenzmodells. Bei der projektionsbasierten Regression wird das multivariate Surrogatsignal zuerst auf einen linearen Unterraum projiziert, welcher bei PLS und CCA zusätzlich auch mithilfe der internen Bewegung bestimmt wird.

werden, sowie die Diagonalmatrix von Singulärwerten $\mathbf{S} \in \mathbb{R}^{n_{sur} \times n_{tp}}$. Anschließend werden nur die ersten k Links-Singulärvektoren $\mathbf{U}_k \in \mathbb{R}^{n_{sur} \times k}$ mit den größten positiven Singulärwerten für die Approximation der Streuungsmatrix $\Sigma_{\mathbf{Z}\mathbf{Z}} \approx \mathbf{U}_k \mathbf{S}_k \mathbf{S}_k^T \mathbf{U}_k^T$ genutzt (siehe Glg. 2.14 mit $\mathbf{Z} \approx \mathbf{U}_k \mathbf{S}_k \mathbf{W}^T$), um die Invertierbarkeit zu gewährleisten. Die orthonormalen Spaltenvektoren in \mathbf{U}_k bilden hierbei Basisvektoren eines k -dimensionalen Unterraums maximaler Datenvariation, wobei k den Grad der Regularisierung steuert. Integriert man die Approximation der Streuungsmatrix in das unregularisierte ($\lambda = 0$) Regressionsproblem in Glg. 2.13, kann die Koeffizientenmatrix berechnet werden durch

$$\mathbf{B}^{\text{PCR}} = \Sigma_{\mathbf{V}\mathbf{Z}} \mathbf{U}_k (\mathbf{S}_k \mathbf{S}_k^T)^{-1} \mathbf{U}_k^T, \text{ mit } \Sigma_{\mathbf{V}\mathbf{Z}} = \mathbf{V}\mathbf{Z}^T. \quad (2.16)$$

Werden nur die Links-Singulärvektoren mit Singulärwerten gleich Null entfernt, so ist Glg. 2.16 Lösung des unregularisierten ($\lambda = 0$) Regressionsproblems in Glg. 2.13 [Hastie u. a., 2009]. Auch für $\lambda > 0$ lässt sich die Lösung für Glg. 2.13 mittels SVD darstellen ($\Sigma_{\mathbf{V}\mathbf{Z}} \mathbf{U} (\mathbf{S}\mathbf{S}^T + \lambda \mathbf{I})^{-1} \mathbf{U}^T$; [Hastie u. a., 2009]). In diesem Fall werden die

einzelnen (quadrierten) Singulärwerte um λ erhöht, wobei der Effekt mit der Größe der Singulärwerte abnimmt.

2.4.2 Partial Least Squares (PLS)

Die isolierte Anwendung der PCA auf die Surrogatdaten in Kapitel 2.4.1 sorgt zwar für eine Regularisierung des Regressionsproblems und für die numerisch stabile Invertierbarkeit der Streuungsmatrix, hat aber gleichzeitig das Problem, dass durch die Projektion der Surrogatdaten auf den gewählten Unterraum Informationen verloren gehen können, die für die präzise Schätzung der internen Bewegung notwendig sind. Bei der Verwendung von Partial Least Squares (PLS) finden deshalb neben den Regressoren auch die Regressanden Beachtung, um einen Unterraum für die Projektion der Surrogatdaten zu bestimmen, sodass ausreichend Informationen für die Schätzung der internen Bewegung erhalten bleiben.

Mathematisch wird dies durch die Bestimmung von zwei linearen Unterräumen umgesetzt, die durch die orthonormalen Matrizen $\mathbf{U}_k := [\mathbf{u}_1 \dots \mathbf{u}_k]$ und $\mathbf{P}_k := [\mathbf{p}_1 \dots \mathbf{p}_k]$ beschrieben werden. Die Basisvektoren $\mathbf{u}_i \in \mathbb{R}^{n_{sur}}$ und $\mathbf{p}_i \in \mathbb{R}^{3n_{vx}}$ bilden hierbei Paare und maximieren die Kreuzkovarianz $\rho_i = \mathbf{u}_i^T \boldsymbol{\Sigma}_{\mathbf{ZV}} \mathbf{p}_i$. Hierdurch wird die Verbindung zwischen Regressoren und Regressanden erreicht, die verhindern soll, dass durch die Projektion der Surrogatdaten wertvolle Informationen entfernt werden. Wie in [Borga u. a., 1997] gezeigt, lassen sich diese Basisvektorenpaare durch Lösung des Eigenwertproblems

$$\begin{cases} \boldsymbol{\Sigma}_{\mathbf{ZV}} \boldsymbol{\Sigma}_{\mathbf{VZ}} \mathbf{u}_i &= \rho^2 \mathbf{u}_i \\ \boldsymbol{\Sigma}_{\mathbf{VZ}} \boldsymbol{\Sigma}_{\mathbf{ZV}} \mathbf{p}_i &= \rho^2 \mathbf{p}_i \end{cases} \quad (2.17)$$

finden. Für die Lösung dieses Eigenwertproblems wird in dieser Arbeit der Non-linear Iterative Partial Least Squares Algorithm (NIPALS) eingesetzt (siehe [Rosipal u. a., 2006] für nähere Informationen zu diesem Verfahren).

Die k berechneten Basisvektoren \mathbf{U}_k beschreiben dann Richtungen maximaler Kovariation (in Bezug auf die Regressanden) und lassen sich wieder nach Glg. 2.14 zur Approximation der Streuungsmatrix nutzen. Eingesetzt in das unregularisierte ($\lambda = 0$) Regressionsproblem in Glg. 2.13 ergibt sich dann wiederum die Koeffizientenmatrix

$$\mathbf{B}^{\text{PLS}} = \boldsymbol{\Sigma}_{\mathbf{VZ}} \mathbf{U}_k (\mathbf{U}_k^T \boldsymbol{\Sigma}_{\mathbf{ZZ}} \mathbf{U}_k)^{-1} \mathbf{U}_k^T . \quad (2.18)$$

Wie bei der PCA-Variante in Kapitel 2.4.1, ist auch hier die Anzahl von berechneten Basisvektorpaaren k der vom Nutzer zu bestimmende Parameter, dessen Wert den Grad der Regularisierung bestimmt.

2.4.3 Kanonische Korrelationsanalyse (CCA)

Die Idee bei der Nutzung der kanonischen Korrelationsanalyse (engl. Canonical Correlation Analysis bzw. CCA) ist prinzipiell äquivalent zur Nutzung von PLS-basierter

Datenprojektion: Verwendung von Regressoren und Regressanden bei der Bestimmung von Unterräumen, um weiterhin eine Schätzung der abhängigen Variablen auf Basis der reduzierten Regressorinformationen zu ermöglichen. Werden bei PLS Richtungen maximaler Datenkovariation bestimmt, so geht es bei der CCA um die Bestimmung von Basisvektorenpaaren $(\mathbf{u}_i, \mathbf{p}_i)$, welche die Korrelation der Daten $\rho_i = \frac{\mathbf{u}_i^T \boldsymbol{\Sigma}_{\mathbf{ZV}} \mathbf{p}_i}{\sqrt{\mathbf{u}_i^T \boldsymbol{\Sigma}_{\mathbf{ZZ}} \mathbf{u}_i \mathbf{p}_i^T \boldsymbol{\Sigma}_{\mathbf{VV}} \mathbf{p}_i}}$ maximieren [Borga u. a., 1997]. Auch diese Basisvektorenpaare ergeben sich als Lösung eines Eigenwertproblems:

$$\begin{cases} \boldsymbol{\Sigma}_{\mathbf{ZZ}}^{-1} \boldsymbol{\Sigma}_{\mathbf{ZV}} \boldsymbol{\Sigma}_{\mathbf{VV}}^{-1} \boldsymbol{\Sigma}_{\mathbf{VZ}} \mathbf{u}_i &= \rho^2 \mathbf{u}_i \\ \boldsymbol{\Sigma}_{\mathbf{VV}}^{-1} \boldsymbol{\Sigma}_{\mathbf{VZ}} \boldsymbol{\Sigma}_{\mathbf{ZZ}}^{-1} \boldsymbol{\Sigma}_{\mathbf{ZV}} \mathbf{p}_i &= \rho^2 \mathbf{p}_i \end{cases} \quad (2.19)$$

Hier ergibt sich allerdings die Problematik, dass sowohl $\boldsymbol{\Sigma}_{\mathbf{ZZ}}$ als auch $\boldsymbol{\Sigma}_{\mathbf{VV}} = \mathbf{V}\mathbf{V}^T$ invertiert werden müssen. Da hierbei die gleichen Probleme auftreten, die den Einsatz projektionsbasierter Regressionsmethoden erst motivieren, werden in dieser Arbeit beide Matrizen, wie u. a. in [Blanc u. a., 2012b; Gao u. a., 2008] vorgeschlagen und analog zum Vorgehen in Kapitel 2.4.1, mittels PCA/SVD-basierter Projektion approximiert, um die Invertierbarkeit zu gewährleisten. Nach der Berechnung von k Basisvektorenpaaren entspricht das Verfahren zur Berechnung der Koeffizientenmatrix

$$\mathbf{B}^{\text{CCA}} = \boldsymbol{\Sigma}_{\mathbf{VZ}} \mathbf{U}_k (\mathbf{U}_k^T \boldsymbol{\Sigma}_{\mathbf{ZZ}} \mathbf{U}_k)^{-1} \mathbf{U}_k^T \quad (2.20)$$

dem Vorgehen in Kapitel 2.4.2 mit den orthonormalen Matrizen $\mathbf{U}_k := [\mathbf{u}_1 \dots \mathbf{u}_k]$ und $\mathbf{P}_k := [\mathbf{p}_1 \dots \mathbf{p}_k]$. Auch hier bestimmt der Parameter k den wählbaren Grad der Regularisierung.

2.5 Zusammenfassung

Die surrogatbasierte Schätzung der Atembewegung ist ein auch in der klinischen Praxis weitverbreiteter Ansatz, um intrainterventionelle Bewegungskompensationstechniken zu steuern. Kernelement der surrogatbasierten Bewegungsschätzung ist zumeist ein patientenspezifisches, anhand von Trainingsdaten gelerntes Korrespondenzmodell, welches basierend auf externen Surrogatsignalmessungen die interne atmungsbedingte Bewegung schätzt. Sind im klinischen Alltag hauptsächlich Systeme verfügbar, die nur einzelne bzw. wenige Punkte (bzw. Marker) im Körperinneren verfolgen können, so wurde in diesem Kapitel ein Framework zur Korrespondenzmodellierung vorgestellt, dessen methodischer Kern initial in [Werner, 2012] skizziert wurde und die Schätzung dichter Verschiebungsfelder für ganze Körperregionen ermöglicht. Die Möglichkeit zur Schätzung dichter Verschiebungsfelder anhand eines Surrogatsignals ermöglicht dabei erst über die reine Marker-/Tumorverfolgung hinausgehende Applikationen wie die intrainterventionelle Dosisakkumulation oder diverse Analyse- und Simulationsanwendungen, für die weitergehende Informationen über die atmungsbedingte Bewegung/Verformung größerer Bereiche des Körperinneren notwendig sind.

Das vorgestellte Framework baut methodisch auf der Annahme eines linearen Zusammenhangs zwischen externer und interner Atembewegung auf. Diese Annahme ist in der Literatur weitverbreitet und scheint auch aufgrund der für die Schätzung dichter Verschiebungsfelder oftmals nur in geringer Menge zur Verfügung stehender Trainingsdaten ein angemessener Ausgangspunkt für etwaige nichtlineare Erweiterungen zu sein. Methodische Besonderheit des vorgestellten Frameworks ist die Beschränkung auf die Schätzung diffeomorpher Transformationen, die wegen ihrer inhärenten mathematischen Eigenschaften besonders gut für die physiologisch plausible Beschreibung der Atembewegung geeignet sind. Die explizite Beschränkung auf diffeomorphe Transformationen erfolgt hierbei durch Formulierung des dem Korrespondenzmodelltraining zugrunde liegenden Regressionsproblems im sogenannten Log-Euklidischen-Framework.

Methodisch erweitert wurde das Framework in dieser Arbeit durch die Integration drei verschiedener projektionsbasierter Regressionstechniken (PCR, PLS und CCA). Diese Techniken haben zum Ziel das beim Einsatz multivariater Surrogatsignale auftretende Problem der Multikollinearität/Redundanzen zu lösen bzw. generell für eine angemessene Regularisierung des Regressionsproblems zu sorgen, um eine Überanpassung zu vermeiden. Im Vergleich zu der in [Werner, 2012] genutzten, unspezifischen Frobeniusnorm-Regularisierung versuchen die projektionsbasierten Techniken die durch Redundanzen potenziell vorhandene niedrigdimensionale Struktur der Daten (sowohl des multivariaten Surrogatsignals als auch der internen Bewegung) für die Regularisierung zu bestimmen. Dies ermöglicht dann eine verstärkte Fokussierung auf die Informationen, die für eine präzise Bestimmung des Korrespondenzmodells am erfolgversprechendsten sind.

Insgesamt stellt dieses Kapitel die Basis für die nachfolgenden drei Kapitel dar, die sich mit der ausführlichen Evaluation des Frameworks, Anwendungen der Methoden über die intrainerventionelle Bewegungsschätzung hinaus und methodischen Erweiterungen des Frameworks beschäftigen.

Kapitel 3

Simulationsbasierte Evaluation patientenspezifischer Korrespondenzmodelle

Kapitelinhalt

3.1 Studiendesign	33
3.1.1 Verwendete Bilddaten	33
3.1.2 Simulierte Surrogate	35
3.1.3 Experimente und Evaluationsmaße	38
3.2 Ergebnisse	43
3.3 Zusammenfassung und Einordnung der Ergebnisse	50

Dieses Kapitel beschäftigt sich mit der umfangreichen und systematischen Evaluation des in Kapitel 2 vorgestellten Frameworks zur patientenspezifischen Korrespondenzmodellierung, die über die initiale Studie in [Werner, 2012] anhand von zehn 4D-CT-Bilddaten deutlich hinausgeht und auch die im Rahmen dieser Arbeit eingebrachten Erweiterungen umfasst. Die hier präsentierte Studie basiert hauptsächlich auf [Wilms u. a., 2014a], enthält aber auch Elemente aus [Wilms u. a., 2013; Wilms u. a., 2014c] und verfolgt insgesamt fünf verschiedene Ziele:

1. Generelle Untersuchung der mithilfe des Frameworks erreichbaren Genauigkeit bei der surrogatbasierten Bewegungsschätzung der atmungsbedingten Lungenbewegung im Vergleich zur Genauigkeit des für die Erstellung der Trainingsdaten genutzten Bildregistrierungsverfahrens.
2. Systematischer Vergleich von verschiedenen Surrogaten/Surrogatsignalen und Regressionsmethoden mit speziellem Fokus auf die Nutzung der Patientenoberfläche als Bewegungsindikator in Verbindung mit dem Einsatz projektionsbasierter Regression.

3. Untersuchung des Einflusses verschiedener relevanter Arten von patientenspezifischer Bewegungsvariabilität (intra-/interzyklische und intra-/interfraktionelle Variabilität) auf die Genauigkeit der Bewegungsschätzung.
4. Vergleich von diffeomorpher und nicht-diffeomorpher Korrespondenzmodellierung in Bezug auf die erreichbare Genauigkeit der Bewegungsschätzung und physiologische Plausibilität der geschätzten Bewegung.
5. Wenn möglich, die Nutzung frei verfügbarer Daten, um die Reproduzierbarkeit und einen studienübergreifenden Ergebnisvergleich zu ermöglichen.

Die genannten Dinge sind nach dem Wissen des Autors dieser Dissertation in dieser Gesamtheit und Kombination noch in keiner anderen Studie/Arbeit systematisch untersucht worden. Hauptproblem bei einer derartigen Studie ist die schlechte Verfügbarkeit geeigneter Trainings- und Evaluationsdaten, da für das Training der patientenspezifischen Korrespondenzmodelle 4D-Bilddaten zur Extraktion der internen Atembewegung und zugehörige Surrogatsignale notwendig sind. 4D-CT- oder 4D-MRT-Bilddaten werden zwar zumeist unter Zuhilfenahme eines zeitgleich aufgenommenen Surrogatsignals rekonstruiert, die zusätzliche zeitsynchrone Aufnahme von Signalen weiterer Bewegungsindikatoren ist aber zumindest im klinischen Umfeld unüblich und nur schwer möglich. Aus diesem Grund werden die verschiedenen Bewegungsindikatoren/Surrogatsignale in dieser Arbeit auf Basis der jeweiligen 4D-Bilddaten simuliert. Hierdurch wird ein systematischer und fairer Vergleich verschiedener Surrogate auf einfache Art und Weise möglich.

Insgesamt baut die Studie auf 34 4D-Bilddatensätzen (31 4D-CT- und 3 4D-MRT-Bilddatensätze) von 31 verschiedenen Patienten aus öffentlichen und institutseigenen Datenbanken auf. Durch die räumlich und zeitlich hohe Auflösung der 4D-CT-Bilddaten, die allerdings jeweils nur einen Atemzyklus abbilden, wird eine genaue Untersuchung der Genauigkeit der Bewegungsschätzung unter Einfluss intrazyklischer Bewegungsvariabilität möglich. Die 4D-MRT-Bilddaten umfassen eine größere Anzahl von Atemzyklen und erlauben somit auch die Untersuchung der Genauigkeit der Bewegungsschätzung unter Einfluss von interzyklischer und intrafraktioneller Bewegungsvariabilität. Da für 3 Patienten sowohl 4D-CT- als auch 4D-MRT-Bilddaten zur Verfügung stehen, die in unterschiedlichen Sitzungen aufgenommen wurden, ist zudem die Untersuchung des Einflusses von interfraktioneller Bewegungsvariabilität möglich.

Um die Genauigkeit und weitere Eigenschaften der durch die Korrespondenzmodelle geschätzten nichtlinearen Transformationen zu ermitteln, wird auf etablierte Evaluationstechniken aus dem Bereich der Evaluation nichtlinearer Registrierungsverfahren zurückgegriffen (siehe Kapitel 2.2.1). Hauptelement ist hierbei die Nutzung von manuell gesetzten, korrespondierenden Landmarken in den für die Evaluation genutzten Bilddaten zur Ermittlung des TRE. Die korrespondierenden Landmarken bilden hierbei die Grundwahrheit für die zu schätzende Bewegung. Deren Nutzung erklärt zudem

die Fokussierung auf die Lungenregion in dieser Studie, da es im Zuge des großen Forschungsaufwandes im Bereich der registrierungsbasierten Schätzung der Lungenbewegung in thorakalen (4D-)CT-Bilddaten in den letzten zehn Jahren (siehe z. B. EMPIRE10-Studie [Murphy u. a., 2011]) einige frei verfügbare thorakale 4D-CT-Bilddatensätze mit von Experten gesetzten korrespondierenden Landmarken gibt [Castillo u. a., 2009, 2010; Vandemeulebroucke u. a., 2011], die für diese Studie eingesetzt werden.

Insgesamt entsteht durch die bildbasierte Simulation verschiedener Bewegungsindikatoren und Anwendungsszenarien und die quantitative Bestimmung der Schätzungsgenauigkeit anhand etablierter Evaluationsprinzipien der Bildregistrierung eine komplette Evaluationsumgebung für Korrespondenzmodelle. Diese kann in Zukunft (bei Nutzung von frei zugänglichen Bilddaten) auch den studien-/publikationsübergreifenden Vergleich von Ergebnissen ermöglichen und als Grundlage für öffentliche Challenges dienen. Bis jetzt ist ein sinnvoller Ergebnisvergleich zwischen verschiedenen Studien bzw. Publikationen kaum möglich, da jede Publikation/Forschungsgruppe ihre eigenen Bilddaten und Daten für die Grundwahrheit nutzt (siehe auch Kapitel 2.1.1.3).

3.1 Studiendesign

In den folgenden Unterkapiteln wird das genaue Studiendesign inklusive verwendeter Bilddaten (Kapitel 3.1.1), simulierter Surrogate (Kapitel 3.1.2) und der verschiedenen Experimente (Kapitel 3.1.3) näher erläutert.

3.1.1 Verwendete Bilddaten

Die 34 verwendeten 4D-Bilddatensätze teilen sich auf vier unterschiedliche Datenquellen auf:

- **WUSTL:** Diese zwölf thorakalen 4D-CT-Bilddatensätze von Lungentumorpatienten stammen aus einer Kooperation des Instituts für Medizinische Informatik der Universität zu Lübeck mit der Washington University in St. Louis, USA und sind nicht frei verfügbar. Die Bilddaten sind mittels Step-&Shoot-Technik [Low u. a., 2003] aufgenommen und mit dem in [Ehrhardt u. a., 2007] beschriebenen strukturerhaltenden Interpolationsverfahren rekonstruiert worden. Für vier der zwölf Patienten wurden zehn Atemphasen rekonstruiert, wohingegen die 4D-Daten der anderen acht Patienten aus jeweils 14 Atemphasen bestehen. Die einzelnen 3D-Bilder bestehen aus $512 \times 512 \times 190 - 270$ Voxeln (Reihenfolge in dieser Arbeit immer: rechts-links (RL) \times anterior-posterior (AP) \times kraniokaudal (KK)) mit einer räumlichen Auflösung von $1 \times 1 \times 1.5 \text{ mm}^3$. Zudem stehen für diese Daten die für die retrospektive Bildrekonstruktion genutzten Spirometrie- und Bauchgurtdaten zur Verfügung. Zu Evaluationszwecken sind für diese Daten weiterhin ca. 70 manuell gesetzte, korre-

spondierende Landmarken für vier Atemphasen (EI, EE, MI, ME)⁴ verfügbar. Diese Landmarken sind gleichmäßig im Lungeninneren verteilt und befinden sich zumeist an Gefäßbifurkationen.

- **DIRLAB:** Diese zehn thorakalen 4D-CT-Bilddatensätze von Patienten mit verschiedenen Tumorerkrankungen sind öffentlich verfügbar und werden vom DIR-Lab, The University of Texas M. D. Anderson Cancer Center, USA für die Evaluation von Registrierungsverfahren zur Verfügung gestellt [Castillo u. a., 2009, 2010]. Sie bestehen aus jeweils zehn Bildern/Atemphasen, welche aus $256 - 512 \times 256 - 512 \times 94 - 136$ Voxeln mit einer räumlichen Auflösung von $0.97 - 1.16 \times 0.97 - 1.16 \times 2.5 \text{ mm}^3$ zusammengesetzt sind. Diese Bilddaten wurden ebenfalls mithilfe einer Step-&Shoot-Technik aufgenommen und sind mit dem in [Pan u. a., 2004] beschriebenen Verfahren unter Verwendung von Surrogatdaten des Real-time Position Management (RPM)-System von Varian Medical Systems rekonstruiert worden. Für diese Daten sind jeweils 300 manuell gesetzte, korrespondierende Landmarken für die Atemphasen EI und EE öffentlich verfügbar.
- **CREATIS:** Diese sechs thorakalen 4D-CT-Bilddaten von Lungentumorpatienten sind ebenfalls öffentlich verfügbar und zur Evaluation von Registrierungsverfahren gedacht. Sie werden vom Léon Bérard Cancer Center & CREATIS lab Lyon in Frankreich [Vandemeulebroucke u. a., 2011] bereitgestellt. Die Bilddaten wurden im Spiralmodus aufgenommen und mithilfe eines Bauchgurtsignals rekonstruiert. Für jeden Patienten stehen jeweils zehn Bilder/Atemphasen mit einer Größe von $512 \times 512 \times 138 - 187$ Voxeln zur Verfügung. Die räumliche Auflösung der Voxel beträgt $0.879 - 1.172 \times 0.879 - 1.172 \times 2 \text{ mm}^3$. Für diese Daten sind jeweils 100 manuell gesetzte, korrespondierende Landmarken für die Atemphasen EI und EE öffentlich verfügbar.
- **DKFZ:** Diese drei 4D-CT- und 4D-MRT-Bilddatensätze stammen aus einer Kooperation des Instituts für Medizinische Informatik der Universität zu Lübeck mit dem Deutschen Krebsforschungszentrum (DKFZ) in Heidelberg und sind nicht öffentlich verfügbar. Für jeden der drei Lungentumorpatienten wurden in unterschiedlichen Sitzungen ein thorakaler 4D-CT-Bilddatensatz und ein 4D-MRT-Bilddatensatz, welcher Thorax und Abdomen abbildet, aufgenommen. Die genaue Aufnahme- und Rekonstruktionstechnik der CT-Bilddaten ist dem Autor dieser Dissertation unbekannt. Die MRT-Bilddaten wurden mittels TWIST-Technik unter freier Atmung mit einem 1.5-Tesla-Tomographen aufgenommen (genaue Beschreibung der Sequenz in [Dinkel u. a., 2009]). Insgesamt besteht jeder 4D-CT-Bilddatensatz aus sieben Bildern/Atemphasen mit einer räumlichen Auflösung von $512 \times 512 \times 94 - 235$ Voxeln

⁴EI: maximale Einatmung; EE: maximale Ausatmung; MI: mittlere Einatmung; ME: mittlere Ausatmung

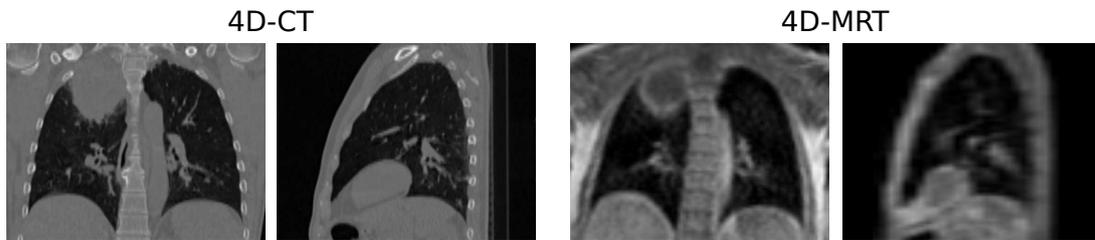


Abb. 3.1: Anatomisch korrespondierende koronare (links) und sagittale (rechts) Schicht eines CT-/MRT-Bildes eines Patienten der DKFZ-Daten bei ähnlicher Einatmungstiefe. Die Bilder illustrieren u. a. die deutlich geringere räumliche Auflösung der MRT-Daten (siehe Text für Details).

und einer Voxelgröße von $0.977 \times 0.977 \times 1.5 - 3 \text{ mm}^3$. Die 4D-MRT-Bilddaten bestehen jeweils aus insgesamt 157 3D-Bildern mit einer räumlichen Auflösung von $128 \times 36 \times 128$ Voxeln und einer Voxelgröße von $3.91 \times 10 \times 3.91 \text{ mm}^3$, die bei freier Atmung im zeitlichen Abstand von 0.5 s aufgenommen wurden. Im Gegensatz zu den 4D-CT-Bilddaten bilden die 4D-MRT-Bilddaten mehrere Atemzyklen (20 – 30, patientenspezifisch) ab. Eine beispielhafte Illustration der räumlichen Auflösungsunterschiede zwischen den beiden Bilddatentypen ist in Abb. 3.1 zu finden.

3.1.2 Simulierte Surrogate

Für diese Evaluationsstudie werden insgesamt vier verschiedene Surrogatsignalkategorien bildbasiert simuliert, die nachfolgend einzeln vorgestellt werden. Die Auswahl der Kategorien erfolgt hierbei einerseits mit Blick auf die derzeitige klinische Praxis und die dort zum Gating und zur Bildrekonstruktion eingesetzten Standardsurrogate, andererseits soll aber auch der Trend hin zu hochdimensionalen Surrogatsignalen basierend auf tiefenbildgebenden Kamerasystemen abgebildet werden (siehe auch die Übersicht in Kapitel 2.1.1.1).

3.1.2.1 Spirometrie

Ein Spirometer misst das durch den Patienten ein- bzw. ausgeatmete Luftvolumen (Spirometervolumen) und ist ein klassisches System zur Generierung von Atemsignalen [Keall u. a., 2006]. Über den zeitlichen Verlauf dieser Messung kann dann die Veränderung des Luftvolumens in der Lunge beobachtet werden.

Bildbasiert kann das Luftvolumen in der Lunge in 4D-CT-Bilddaten durch voxelweise Auswertung der Hounsfield-Werte der Lungenvoxel unter Berücksichtigung des Partialvolumeneffekts ermittelt werden (siehe [Lu u. a., 2005b] für die Berechnung). In [Lu u. a., 2005b] wurde darüber hinausgehend experimentell gezeigt, dass, nach Korrektur des Spirometersignals bezüglich Drift und Phasenverschiebung, ein linearer

Zusammenhang zwischen dem bildbasiert bestimmten Luftvolumen und dem Spirometervolumen besteht. Aufgrund dieses Zusammenhangs erscheint die Nutzung des bildbasiert bestimmten Luftvolumens als Spirometersignal gerechtfertigt. In den nachfolgenden Experimenten wird sowohl das so simulierte 1D-Spirometersignal als auch ein hieraus abgeleitetes 2D-Signal (1D-Spirometersignal + zeitliche Ableitung) eingesetzt (d.h. $n_{sur} = 1$ oder 2). Hierbei wird die zeitliche Ableitung mittels finiter Differenzen approximiert.

Da für MRT-Bilddaten kein zu den CT-Bilddaten vergleichbarer Zusammenhang zwischen Bildinhalt und Luftgehalt in der Lunge existiert, werden die Spirometersignale für die drei 4D-MRT-Bilddatensätze vereinfacht durch Berechnung des kompletten Lungenvolumens simuliert. Für beide Modalitäten werden die jeweiligen Lungenvoxel automatisiert durch die in [Wilms u. a., 2011, 2012] beschriebenen, selbst entwickelten Lungensegmentierungsverfahren bestimmt.

3.1.2.2 Bauchgurt

Bauchgurte sind ein klassisches Instrument zur Generierung von Atemsignalen im klinischen Alltag (siehe [Keall u. a., 2006]). Ein Bauchgurt misst typischerweise über einen Drucksensor die durch Änderung des Luftvolumens in der Lunge induzierte Volumenveränderung der durch den Gurt umspannten Körperregion.

Bildbasiert wird das Bauchgurtsignal durch automatisierte schwellwertbasierte Segmentierung der Körperregion in einer kleinen Anzahl von axialen Bildschichten des Abdomens simuliert. Der Signalwert entspricht dann der Anzahl der segmentierten Körpervoxel. Die Anzahl der axialen Schichten wird hierbei so gewählt, dass sie analog zu einem realen Bauchgurt circa 1 cm der abdominalen Körperoberfläche in kranio-kaudaler Richtung umfassen. Die Position der Schichten wird analog zur klinischen Praxis patientenspezifisch so gewählt, dass über die Zeit eine klare atmungsbedingte Volumenveränderung sichtbar ist.

Da anders als beim Spirometersignal in der Literatur keine Experimente zur Verifizierung der durchgeführten bildbasierten Simulation verfügbar sind, erfolgt die experimentelle Verifikation anhand von drei zufällig ausgewählten Patienten der WUSTL-Daten. Da für diese sowohl die für verschiedene Tischpositionen aufgenommenen Originalbilddaten als auch das zusätzlich für die Rekonstruktion aufgezeichnete Bauchgurtsignal zur Verfügung steht, lässt sich anhand dieser Daten die Korrelation des bildbasiert bestimmten abdominalen Körpervolumens und des echten Bauchgurtsignals analysieren. Bei dieser Analyse ergab sich über durchschnittlich fünf Atemzyklen pro Patient eine Korrelation von > 0.98 (Patient 1: 0.996; Patient 2: 0.997, Patient 3: 0.984), was den Realismus der bildbasierten Simulation belegt. Analog zum Spirometersignal werden in den Experimenten das simulierte 1D-Bauchgurtsignal und ein zusammengesetztes 2D-Signal (1D-Bauchgurtsignal + zeitliche Ableitung) genutzt (d.h. $n_{sur} = 1$ oder 2).

3.1.2.3 Zwerchfellbewegung

Das Zwerchfell ist beim Menschen der wichtigste Atemmuskel, weshalb die hauptsächlich in kraniokaudaler Richtung verlaufende Zwerchfellbewegung ein sehr guter (direkter) Indikator der gesamten Atembewegung ist. Die Messung der Zwerchfellbewegung ist, verglichen mit Spirometrie und Bauchgurt, wesentlich aufwendiger, lässt sich aber in der klinischen Praxis bildbasiert in Echtzeit z. B. mittels Fluoroskopie (in 2D) oder schneller MRT-Sequenzen (in 1D oder 2D) durchführen [Cerviño u. a., 2010, 2011; Keall u. a., 2006].

Für diese Studie wird auf Basis der jeweiligen Bilddaten die Translation der linken und rechten Zwerchfellkuppel in kraniokaudaler Richtung über die Zeit verfolgt, ohne eine bestimmte Bildaufnahmetechnik exakt zu simulieren. Um ein 1D-Atemsignal zu erhalten, werden die Bewegungen beider Zwerchfellkuppeln gemittelt. Die subvoxelgenaue Bestimmung der Position der Kuppeln erfolgt nach manueller Auswahl geeigneter kraniokaudaler Bildschichten automatisiert mithilfe von Kantenerkennungsverfahren. Da das Zwerchfell durch die Bildauflösung der genutzten Bilddaten zumeist nicht klar abgrenzbar ist, wird durch die Kantenerkennung die klar sichtbare Grenze zwischen Lunge (Luft) und Abdomen (Gewebe) bestimmt. Dieses Vorgehen ist in beiden Modalitäten problemlos möglich. Zusätzlich zu dem so gewonnen 1D-Atemsignal wird auch hier wieder ein zusammengesetztes 2D-Signal (1D-Signal + zeitliche Ableitung) erstellt (d.h. $n_{sur} = 1$ oder 2).

3.1.2.4 Tiefenbilder der Körperoberfläche

Die aus Ein- und Ausatmung resultierende Veränderung des thorakalen und abdominalen Körpervolumens lässt sich u. a. auch durch Tiefenbilder der Körperoberfläche messen (siehe Übersicht in Kapitel 2.1.1.1). Gegenüber der Verwendung des klassischen Bauchgurts ergibt sich hierbei der Vorteil, dass auf Basis der pixelweise erfassten Oberflächenbewegung sowohl 1D- als auch sehr hochdimensionale Surrogatsignale abgeleitet werden können.

Für diese Studie erfolgt die bildbasierte Simulation von Tiefenbildern des Thorax und Abdomens durch ein relativ einfaches, approximatives Verfahren, welches optische und anderweitige Effekte realer Kamerasysteme weitestgehend ignoriert. Hierfür wird ein virtueller Kamerasensor parallel zu den koronaren Schichten der Bilddaten in anteriorer Richtung oberhalb des Patientenkörpers platziert. Der virtuelle Kamerasensor hat dabei die gleiche Auflösung wie die koronaren Schichten. Ein Tiefenbild des Körpers wird dann erzeugt, indem von jedem Pixel des Kamerasensors orthogonal zur Sensorebene in anterior-posteriorer Richtung ein Strahl ausgesendet wird, der solange verfolgt wird, bis er auf die Körperoberfläche trifft. Die hierbei zurückgelegte Entfernung ist der simulierte Tiefenwert. Der Treffpunkt zwischen Strahl und Körperoberfläche wird anhand eines Grauwertschwellwerts im Bereich des Übergangs zwischen Luft und Gewebe (für CT: -100 HU; für MRT: patientenspezifisch) durch lineare Interpolation

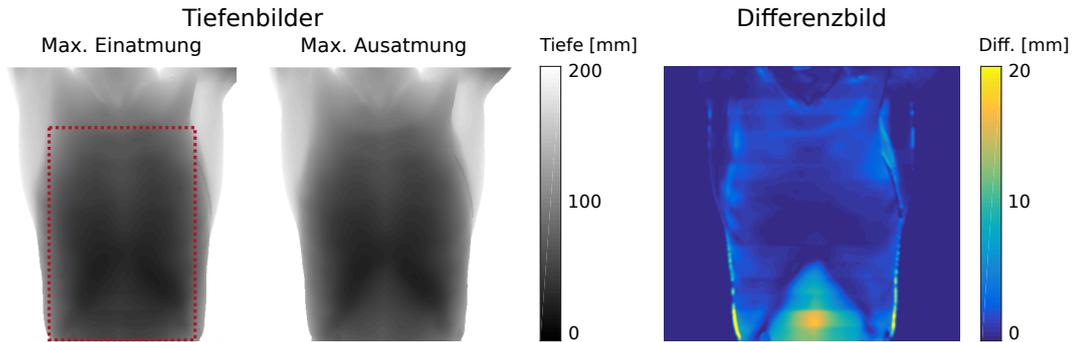


Abb. 3.2: Zwei aus 4D-CT-Bilddaten simulierte Tiefenbilder des Thorax und Abdomens zu den Zeitpunkten maximaler Ein- und Ausatmung für einen Patienten der WUSTL-Daten. Zusätzlich illustriert das aus den beiden Tiefenbildern errechnete Differenzbild (Betrag) die mittels Tiefenbildgebung detektierbare Atembewegung, welche für diesen Patienten hauptsächlich durch Bauchatmung gekennzeichnet ist. Die rote Box definiert die in den Experimenten eingesetzte rechteckige Region-of-Interest des Thorax/Abdomens.

subvoxelgenau bestimmt. Abbildung 3.2 zeigt beispielhaft simulierte Tiefenbilder für einen Patienten der WUSTL-Daten. Auf Basis des so simulierten Tiefenbildes lassen sich beliebige Abtastmuster bzw. Surrogatsignale definieren. Da die MRT-Bilddaten in anterior-posteriorer Richtung einen Schichtabstand von 10 mm aufweisen, werden diese auf die Auflösung der zugehörigen CT-Bilddaten hochskaliert.

Für die Evaluation werden aus den Tiefenbildern insgesamt sechs verschiedene Abtastmuster bzw. Surrogatsignale generiert: (1) Ein Punkt ($n_{sur} = 1$) zentriert im Bereich des Schwertfortsatzes des Brustbeins. In diesem Bereich wird im klinischen Alltag oft der zum RPM-System von Varian gehörende Markerblock platziert. (2) Ein Punkt ($n_{sur} = 1$) zentriert im Bereich des simulierten Bauchgurts. (3-6) Vier Abtastungen einer rechteckigen Region-of-Interest des Thorax/Abdomens mit $n_{sur} = 10, 100, 1000, 10000$ gleichmäßig verteilten Abtastpunkten. Für die beiden 1D-Signale wird zusätzlich noch ein zusammengesetztes Signal mit zeitlicher Ableitung ($n_{sur} = 2$) getestet. Bei den hochdimensionalen Signalen mit $n_{sur} \geq 10$ wird auf die Verwendung der zeitlichen Ableitung verzichtet, da erwartet wird, dass die Informationen über die jeweilige Atemphase implizit im Signal codiert ist.

3.1.3 Experimente und Evaluationsmaße

Die Experimente wurden mit Blick auf die eingangs genannten Ziele dieser Studie entworfen und werden nach einigen allgemeinen Anmerkungen nachfolgend genauer vorgestellt. Eine schematische Übersicht bezüglich des Vorgehens bei den Experimenten 1–3 ist ergänzend in Abb. 3.3 zu finden.

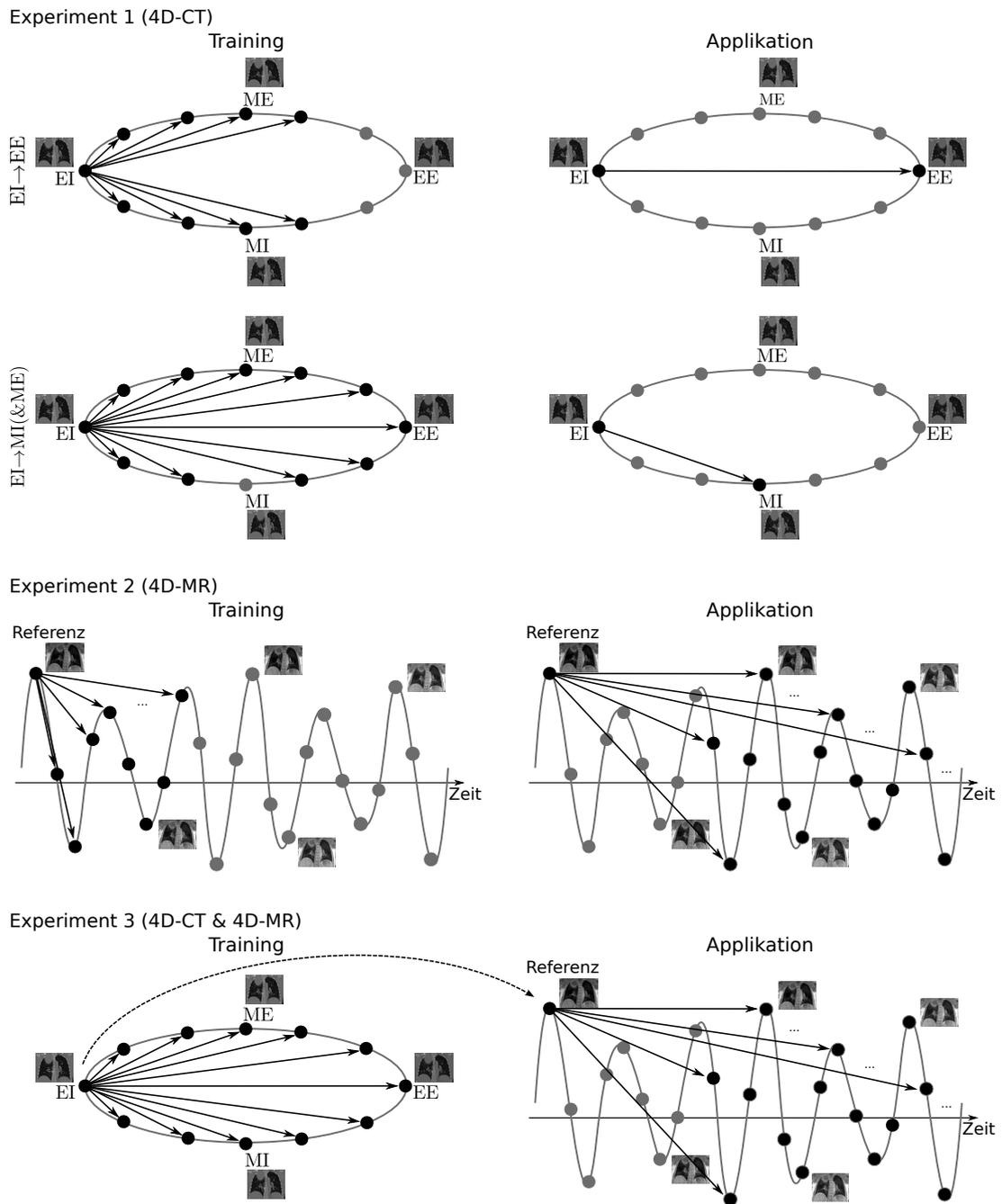


Abb. 3.3: Schematische Darstellung der für das Training und die Applikation der Korrespondenzmodelle verwendeten (Bild-)Daten in den Experimenten 1–3 (siehe Text für Details). Verwendete Daten sind durch schwarze Punkte/Pfeile gekennzeichnet, wohingegen im jeweiligen Schritt nicht genutzte Daten grau dargestellt sind.

Wenn nicht anders angegeben, werden in den Experimenten alle in Kapitel 3.1.2 beschriebenen Surrogate eingesetzt. Von den in Kapitel 2.3 und Kapitel 2.4 beschriebenen Regressionsmethoden kommt für die 1D- und 2D-Surrogatsignale sinnvollerweise nur das Standardverfahren mit Frobeniusnorm-Regularisierung zum Einsatz (Glg. 2.13). Für die durch Tiefenbildgebung erzeugten höherdimensionalen Signale werden zudem auch die in Kapitel 2.4 vorgestellten projektionsbasierten Regressionsmethoden (Glg. 2.16–2.20) eingesetzt.

Die optimalen Regularisierungsparameter der einzelnen Regressionsmethoden werden für jedes Experiment patientenspezifisch durch Leave-One-Out-Kreuzvalidierung auf den jeweiligen Trainingsdaten ermittelt. Da für diese Trainingsdaten zumeist keine Landmarken zur Verfügung stehen, wird für die Kreuzvalidierung der mittlere Fehler der modellbasierten Bewegungsschätzung mit Bezug zur jeweiligen Trainingstransformation über alle Lungenvoxel bestimmt (euklidische Norm der Vektordifferenz). Es wird der Regularisierungsparameter für die Experimente genutzt, der im Mittel über alle Teilmengen der Kreuzvalidierung den geringsten Fehler erzeugt.

Die für die Korrespondenzmodellierung benötigten diffeomorphen Transformationen und stationären Geschwindigkeitsfelder werden für alle Bilddatensätze mithilfe der *VariationalRegistration*-Toolbox geschätzt [Schmidt-Richberg u. a., 2014; Werner u. a., 2014] (siehe auch Kapitel 2.2). Da das dort implementierte Registrierungsverfahren die Gleitbewegungen am Lungenrand nicht abbilden kann, werden die schon zur Surrogatsimulation in Kapitel 3.1.2 generierten Lungensegmentierungen auch eingesetzt, um die Registrierung durch Maskierung auf das Lungeninnere zu beschränken. Gleiches gilt auch für die anschließende Korrespondenzmodellierung. Die Parameter des Registrierungsverfahrens wurden basierend auf eigenen Experimenten und den Erfahrungen aus [Werner u. a., 2014] so gewählt, dass ein guter Kompromiss zwischen Rechenzeit und niedrigem Landmarkenfehler erreicht wird (\mathcal{D} : normalisierte SSD-Kräfte; \mathcal{S} : diffuser Regularisierer; $\alpha = 0.7/\alpha = 1.0$; siehe Glg. 2.6).

3.1.3.1 Experiment 1: Intra- und interzyklische Bewegungsvariabilität (4D-CT-Bilddaten)

Um den Einfluss von intra- und interzyklischer Bewegungsvariabilität auf die Genauigkeit der modellbasierten Bewegungsschätzung zu ermitteln, werden die insgesamt 28 4D-CT-Datensätze der WUSTL-, DIRLAB- und CREATIS-Daten eingesetzt. Ausgehend von der Phase maximaler Einatmung (EI) als Referenzphase werden analog zu dem Vorgehen in [Werner, 2012] für jeden Datensatz insgesamt drei verschiedene Korrespondenzmodelle pro Kombination aus Surrogatsignal und Regressionsmethode erstellt (siehe Abb. 3.2). Je einmal wird die Phase mittlerer Ein- (MI) bzw. Ausatmung (ME) beim Training weggelassen und dann versucht die Bewegung zwischen EI und MI (EI→MI) bzw. ME (EI→ME) durch das jeweilige Korrespondenzmodell zu schätzen. Mithilfe dieser beiden Modelle soll insbesondere der Einfluss intrazyklischer

Bewegungsvariabilität untersucht werden, da dieser in den Atemmittellagen am deutlichsten sichtbar sein sollte. Zusätzlich wird ein Modell gebaut, bei dem die Phase der maximalen Ausatmung (EE) sowie die beiden Nachbarphasen (EE-1 und EE+1) beim Training nicht verwendet werden, um anschließend die Bewegung zwischen EI und EE (EI→EE) modellbasiert zu schätzen. Hierbei wird approximativ interzyklische Bewegungsvariabilität simuliert, da die Trainingsdaten nun nur noch eine wesentlich geringere Atmungstiefe beinhalten.

Zur Evaluation der modellbasiert geschätzten Transformationen wird der mittlere Target-Registration-Error (TRE) (siehe Glg. 2.7) auf Basis manuell gesetzter, korrespondierender Landmarken im Lungeninneren berechnet. Da für die DIRLAB- und CREATIS-Daten keine Landmarken für die beiden mittleren Atemphasen verfügbar sind, werden die beiden Modelle zur Evaluation der intrazyklischen Bewegungsvariabilität nur für die zwölf WUSTL-Datensätze erstellt.

3.1.3.2 Experiment 2: Intrafraktionelle Bewegungsvariabilität (4D-MRT-Bilddaten)

Da die drei 4D-MRT-Bilddatensätze jeweils 20 – 30 komplette Atemzyklen umfassen, eignen sich diese besonders zur Untersuchung des Einflusses intrafraktioneller Variabilität auf die Schätzungsgenauigkeit. Hierfür werden die Daten von jeweils zwei zusammenhängenden Atemzyklen als Trainingsdaten zur Modellbildung genutzt, wobei der verbleibende Teil der Daten zur Evaluation genutzt wird (siehe Abb. 3.2). Die beiden für das Training ausgewählten Atemzyklen wurden so gewählt, dass sie in Länge und Tiefe der Atmung ungefähr dem patientenspezifischen Durchschnitt entsprechen, sodass extreme, singuläre Ereignisse (Husten etc.) nicht das Training beeinflussen. Hierdurch wird ein typischer klinischer Workflow simuliert, bei dem zu Beginn jeder Behandlungssitzung ein neues Korrespondenzmodell erzeugt wird.

Da die geringe Auflösung der drei 4D-MRT-Bilddatensätze keine landmarkenbasierte Evaluation zulässt, werden für diese Daten die schon vorhandenen Lungensegmentierungen zur Evaluation eingesetzt. Mithilfe der jeweils geschätzten Transformation wird die Lungensegmentierung des zur Surrogatmessung korrespondierenden MRT-Bildes auf das Referenzbild transformiert. Dies ermöglicht den quantitativen Vergleich zwischen der Lungensegmentierung des Referenzbildes und der transformierten Segmentierung mittels mittlerer symmetrischer Oberflächendistanz (siehe Glg. 2.9). Die Werte der Einzelbilder eines 4D-Datensatzes werden jeweils gemittelt.

3.1.3.3 Experiment 3: Interfraktionelle Bewegungsvariabilität (4D-CT- & 4D-MRT-Bilddaten)

Die kombinierten 4D-MRT-/4D-CT-Bilddaten der drei DKFZ-Patienten werden für die Untersuchung interfraktioneller Variabilitätseffekte eingesetzt. Angenommen wird

hierfür ein möglicher klinischer Workflow, bei dem basierend auf den in einer Planungssitzung aufgenommenen Daten (hier: 4D-CT-Bilddaten) ein Korrespondenzmodell geschätzt wird, welches im Anschluss in einer Behandlungssitzung zum Einsatz kommt (siehe Abb. 3.2). Die Behandlungssitzung wird hierbei durch die MRT-Bilddaten simuliert, wobei nur die Atemzyklen zur Evaluation herangezogen werden, die auch in Experiment 2 zum Einsatz kommen, um eine Vergleichbarkeit der Ergebnisse zu gewährleisten.

Da die CT- und MRT-Bilddaten in unterschiedlichen Sitzungen und Scannern aufgenommen wurden, müssen für die Evaluation zumindest Lage- und Deformationsunterschiede der Lungen ausgeglichen werden. Hierfür wird in den 4D-MRT-Bilddaten zunächst ein 3D-Bild als Referenzbild ausgewählt, dessen Lungenvolumen ungefähr dem der Referenzphase der 4D-CT-Bilddaten entspricht. Zwischen diesen beiden Bildern wird auf Basis der Lungensegmentierungen eine affine Transformation $\varphi_{aff} : \Omega_{MRT} \rightarrow \Omega_{CT}$ geschätzt, um Lageunterschiede und globale Deformationen auszugleichen. Aufgrund der relativ geringen räumlichen Auflösung der MRT-Bilddaten, hat sich eine weitere nichtlineare Registrierung der Daten als nicht notwendig bzw. sinnvoll erwiesen. Die affine Transformation kann im Anschluss genutzt werden, um die durch das Korrespondenzmodell geschätzten Transformationen $\hat{\varphi}_{CT} : \Omega_{CT} \rightarrow \Omega_{CT}$ vom Raum der CT-Bilddaten Ω_{CT} in den Raum der MRT-Bilddaten Ω_{MRT} zu transportieren:

$$\hat{\varphi}_{MRT} = \varphi_{aff}^{-1} \circ \hat{\varphi}_{CT} \circ \varphi_{aff} . \quad (3.1)$$

Die simulierten Surrogatsignale der MRT-Bilddaten werden so verschoben, dass sie für die Referenzphase den jeweiligen Werten der CT-Daten entsprechen. Die manuelle Platzierung von Bauchgurt, Abtastpunkten, usw. erfolgt in CT- und MRT-Bilddaten an anatomisch korrespondierenden Positionen.

Die finale Evaluation der Genauigkeit der Bewegungsschätzung erfolgt analog zum vorangegangenen Experiment anhand der transformierten Lungensegmentierungen der MRT-Bilddaten durch Berechnung von symmetrischen Oberflächendistanzen.

3.1.3.4 Experiment 4: Vergleich diffeomorpher/nicht-diffeomorpher Korrespondenzmodelle

In Kapitel 2.3 wird der Ansatz zur diffeomorphen Korrespondenzmodellierung damit motiviert, dass bei dem in der Literatur beschriebenen nicht-diffeomorphen Standardvorgehen (direkte Regression zwischen Surrogatdaten und Verschiebungsfeldern) die resultierenden Korrespondenzmodelle nicht unbedingt diffeomorphe Transformationen schätzen. Um die Unterschiede zwischen diffeomorpher und nicht-diffeomorpher Korrespondenzmodellierung praktisch zu belegen und genauer zu untersuchen, werden auf Basis einer Untermenge der Bilddaten (WUSTL- und DKFZ-Daten) und einer Auswahl von Surrogattypen für alle drei zuvor beschriebenen Experimentkategorien

jeweils zusätzlich auch nicht-diffeomorphe Korrespondenzmodelle gebaut. Die Evaluation erfolgt hierbei wieder anhand der in den vorherigen Experimenten beschriebenen Gütemaße. Zusätzlich findet auch ein Vergleich anhand der durch Berechnung der aus dem Wert der Jacobi-Determinanten ermittelten Singularitäten in den geschätzten Verschiebungsfeldern statt (siehe Glg. 2.10), wobei diese Auswertung wieder nur innerhalb der Lungenregion erfolgt.

3.2 Ergebnisse

Tabelle 3.1 fasst die Ergebnisse der Experimente 1–3 kompakt für alle Datenquellen zusammen. Nach Datenquelle aufgeteilte Ergebnisse für die frei verfügbaren DIRLAB- und CREATIS-Daten sind zusätzlich in Anhang A.1 zu finden, um den zukünftigen Vergleich mit anderen Verfahren auf diesen Daten zu ermöglichen. Alle Ergebnisse in Tab. 3.1 sind Mittelwerte der (patientenspezifischen) Einzelergebnisse. Um die statistische Signifikanz der Unterschiede dieser Gesamtergebnisse zu beurteilen, wurden gepaarte t-Tests (zweiseitig; Signifikanzniveau: 5%) auf Basis der Einzelergebnisse durchgeführt. Für die Experimente auf den MRT-Daten wurden hierfür zuerst die Ergebnisse der jeweils ca. 130 Testbilder pro Patient gemittelt und anschließend die Signifikanztests durchgeführt⁵. Zur besseren Übersicht erfolgt die Präsentation der umfangreichen Ergebnisse dieser Studie anhand aus den eingangs formulierten Zielen abgeleiteter Fragen.

Frage 1: Ist die durch surrogatbasierte Bewegungsschätzung erreichbare Genauigkeit vergleichbar mit der Registrierungsgenauigkeit? Für die CT-Daten in Experiment 1 sind alle vom genutzten Registrierungsverfahren *VariationalRegistration* erreichten mittleren TRE-Werte signifikant besser (alle $p < 0.05$), als die durch surrogatbasierte Bewegungsschätzung erreichten Werte, unabhängig von Surrogat und eingesetztem Regressionsverfahren. Im Mittel ist der TRE der modellbasierten Schätzung um ca. 18% (EI→MI/ME) bzw. ca. 51% (EI→EE) höher als der von *VariationalRegistration*.

Für die MRT-Daten in den Experimenten 2–3 ergeben sich mit Ausnahme des 1D-Tiefenbildsignals an der Bauchgurtposition in Experiment 2 keine signifikanten Unterschiede zwischen den vom Registrierungsverfahren erreichten Oberflächenabstände und

⁵Es ist zu beachten, dass dieses Vorgehen dazu führt, dass bei den Experimenten auf Basis der MRT-Daten immer nur drei Ergebnisse miteinander verglichen werden, um statistische Signifikanz zu ermitteln. Die Aussagekraft der Signifikanztests ist hierdurch eingeschränkt (geringe Teststärke). Als Alternative bietet sich u. a. die patientenspezifische Untersuchung der Ergebnisse an, die so in [Wilms u. a., 2014a] durchgeführt wurde. Hier wird auf diese Darstellung mit Ausnahme von Frage 5 verzichtet, da sich die Grundaussagen der Studie hierdurch nicht ändern und eine bessere Übersicht für den Leser erreicht wird. Interessierte Leser seien für die patientenspezifischen Ergebnisse auf [Wilms u. a., 2014a] verwiesen.

Tabelle 3.1: Ergebnisse der Experimente 1–3 zur surrogatbasierten Bewegungsschätzung angegeben als Mittelwert \pm Standardabweichung des mittleren TRE oder der mittleren symmetrischen Oberflächendistanz. Die Mittelung erfolgte über alle für das jeweilige Experiment verfügbaren Einzelergebnisse (siehe Text für Details). Die Ergebnisse sind als Blöcke nach genutztem Regressionsverfahren/Dimension des Surrogatsignals geordnet. Tabelle modifiziert aus [Wilms u. a., 2014a].

Bewegungsschätzung	Mittlerer TRE [mm]		Mittlere Oberflächendistanz [mm]	
	Experiment 1		Experiment 2	Experiment 3
	4D-CT		4D-MRT	4D-CT & 4D-MRT
	EI \rightarrow MI&ME	EI \rightarrow EE		
Mittlere Bewegung	3.59 \pm 1.46	7.59 \pm 2.72		2.81 \pm 1.54
<i>VariationalRegistration</i>	1.52 \pm 0.15	1.40 \pm 0.49		1.12 \pm 0.42
<i>Surrogatbasierte Bewegungsschätzung mit \mathbf{B}^{RR} und simuliertem 1D-Surrogatsignal</i>				
Spirometrie	1.87 \pm 0.31	2.05 \pm 0.85	1.31 \pm 0.51	1.94 \pm 1.16
Bauchgurt	1.94 \pm 0.32	2.03 \pm 0.82	1.40 \pm 0.54	2.00 \pm 1.19
Zwerchfellbewegung	1.90 \pm 0.32	2.01 \pm 0.88	1.27 \pm 0.50	2.13 \pm 1.25
Tiefenbild Bauchgurt	1.92 \pm 0.27	2.08 \pm 0.91	1.46 \pm 0.62	2.31 \pm 1.31
Tiefenbild Brustbein	2.18 \pm 0.66	3.76 \pm 2.02	1.58 \pm 0.54	2.17 \pm 0.75
<i>Surrogatbasierte Bewegungsschätzung mit \mathbf{B}^{RR} und simuliertem 2D-Surrogatsignal</i>				
Spirometrie	1.73 \pm 0.19	1.96 \pm 0.81	1.28 \pm 0.50	1.88 \pm 1.08
Bauchgurt	1.72 \pm 0.19	1.98 \pm 0.73	1.37 \pm 0.52	1.94 \pm 1.08
Zwerchfellbewegung	1.74 \pm 0.19	2.04 \pm 0.86	1.24 \pm 0.48	2.23 \pm 1.46
Tiefenbild Bauchgurt	1.72 \pm 0.17	2.04 \pm 0.84	1.38 \pm 0.58	2.31 \pm 1.27
Tiefenbild Brustbein	2.25 \pm 0.71	3.12 \pm 1.91	1.52 \pm 0.59	2.56 \pm 1.20
<i>Surrogatbasierte Bewegungsschätzung mit \mathbf{B}^{RR} und simulierten Tiefenbildern</i>				
10 Punkte	1.74 \pm 0.21	2.12 \pm 0.86	1.44 \pm 0.59	1.99 \pm 0.89
100 Punkte	1.71 \pm 0.22	1.86 \pm 0.63	1.41 \pm 0.57	2.01 \pm 0.89
1000 Punkte	1.71 \pm 0.20	1.88 \pm 0.59	1.40 \pm 0.56	1.94 \pm 0.80
10000 Punkte	1.80 \pm 0.35	2.02 \pm 0.68	1.40 \pm 0.56	1.95 \pm 0.81
<i>Surrogatbasierte Bewegungsschätzung mit \mathbf{B}^{PCR} und simulierten Tiefenbildern</i>				
10 Punkte	1.74 \pm 0.20	2.36 \pm 1.49	1.48 \pm 0.62	1.99 \pm 0.88
100 Punkte	1.71 \pm 0.23	1.87 \pm 0.61	1.41 \pm 0.58	1.97 \pm 0.82
1000 Punkte	1.68 \pm 0.19	1.87 \pm 0.63	1.40 \pm 0.57	1.96 \pm 0.81
10000 Punkte	1.69 \pm 0.20	1.82 \pm 0.62	1.40 \pm 0.56	1.96 \pm 0.82
<i>Surrogatbasierte Bewegungsschätzung mit \mathbf{B}^{PLS} und simulierten Tiefenbildern</i>				
10 Punkte	1.73 \pm 0.20	2.25 \pm 1.16	1.48 \pm 0.61	2.00 \pm 0.89
100 Punkte	1.70 \pm 0.24	1.86 \pm 0.61	1.41 \pm 0.57	1.96 \pm 0.82
1000 Punkte	1.68 \pm 0.19	1.84 \pm 0.60	1.40 \pm 0.56	1.96 \pm 0.81
10000 Punkte	1.69 \pm 0.20	1.79 \pm 0.59	1.39 \pm 0.56	1.96 \pm 0.82
<i>Surrogatbasierte Bewegungsschätzung mit \mathbf{B}^{CCA} und simulierten Tiefenbildern</i>				
10 Punkte	1.75 \pm 0.20	2.22 \pm 0.86	1.45 \pm 0.66	2.21 \pm 1.15
100 Punkte	1.77 \pm 0.23	2.04 \pm 0.96	1.34 \pm 0.52	2.10 \pm 1.01
1000 Punkte	1.72 \pm 0.19	2.02 \pm 0.91	1.32 \pm 0.51	1.98 \pm 0.84
10000 Punkte	1.72 \pm 0.19	1.88 \pm 0.67	1.48 \pm 0.42	1.97 \pm 0.83

Tabelle 3.2: Ergebnisse von Experiment 4 zum Vergleich diffeomorpher und nicht-diffeomorpher Korrespondenzmodelle für verschiedene Surrogate in Kombination mit klassischer Regression. Die Tabelle ist in zwei Blöcke (oben und unten) aufgeteilt, wobei der obere Block die Ergebnisse auf den genutzten Teildaten von Experiment 1 zusammenfasst und der untere Block die Ergebnisse auf den Daten von Experiment 2 und 3 zeigt. Es sind jeweils die mittleren TRE bzw. Oberflächendistanzen und die mittlere Anzahl der Voxel mit Singularitäten angegeben. Singularitäten wurden nur für Voxel innerhalb der Lunge evaluiert (4D-CT: ≈ 3000000 Lungenvoxel; 4D-MRT: ≈ 300000 Lungenvoxel). Tabelle modifiziert/erweitert aus [Wilms u. a., 2014a].

Experiment 1				
4D-CT				
Bewegungsschätzung	EI→MI&ME		EI→EE	
	TRE [mm]	Singularitäten	TRE [mm]	Singularitäten
Diffeomorphes Korrespondenzmodell:				
\mathbf{B}^{RR} & Spirometrie (1D)	1.87 ± 0.31	0	1.84 ± 0.45	0
\mathbf{B}^{RR} & Spirometrie (2D)	1.73 ± 0.19	0	1.81 ± 0.44	0
\mathbf{B}^{RR} & Bauchgurt (1D)	1.94 ± 0.32	0	1.79 ± 0.34	0
\mathbf{B}^{RR} & Bauchgurt (2D)	1.72 ± 0.19	0	1.79 ± 0.34	0
\mathbf{B}^{RR} & Tiefenb. 1000 P.	1.71 ± 0.20	0	1.67 ± 0.29	0
Nicht-diffeomorphes Korrespondenzmodell:				
\mathbf{B}^{RR} & Spirometrie (1D)	1.86 ± 0.30	0	1.83 ± 0.44	0.33 ± 1.15
\mathbf{B}^{RR} & Spirometrie (2D)	1.72 ± 0.18	0	1.80 ± 0.43	0.33 ± 1.15
\mathbf{B}^{RR} & Bauchgurt (1D)	1.93 ± 0.32	0	1.80 ± 0.34	109.58 ± 357.53
\mathbf{B}^{RR} & Bauchgurt (2D)	1.71 ± 0.18	0	1.79 ± 0.34	127.17 ± 416.26
\mathbf{B}^{RR} & Tiefenb. 1000 P.	1.70 ± 0.20	0	1.69 ± 0.28	29.75 ± 71.90
Experiment 2				
4D-MRT				
Bewegungsschätzung	Mittlere Oberfl.-	Singularitäten	Mittlere Oberfl.-	Singularitäten
	distanz [mm]		distanz [mm]	
Diffeomorphes Korrespondenzmodell:				
\mathbf{B}^{RR} & Spirometrie (1D)	1.31 ± 0.51	0	1.94 ± 1.16	0.13 ± 0.22
\mathbf{B}^{RR} & Spirometrie (2D)	1.28 ± 0.50	0	1.88 ± 1.08	0.10 ± 0.18
\mathbf{B}^{RR} & Bauchgurt (1D)	1.40 ± 0.54	0	2.00 ± 1.19	0.01 ± 0.01
\mathbf{B}^{RR} & Bauchgurt (2D)	1.37 ± 0.52	0	1.94 ± 1.08	0.02 ± 0.03
\mathbf{B}^{RR} & Tiefenb. 10000 P.	1.40 ± 0.56	0	1.95 ± 0.81	2.32 ± 3.81
Nicht-diffeomorphes Korrespondenzmodell:				
\mathbf{B}^{RR} & Spirometrie (1D)	1.29 ± 0.49	0.61 ± 1.06	1.90 ± 1.13	257.93 ± 442.30
\mathbf{B}^{RR} & Spirometrie (2D)	1.26 ± 0.49	0.90 ± 1.56	1.86 ± 1.07	251.93 ± 428.93
\mathbf{B}^{RR} & Bauchgurt (1D)	1.38 ± 0.53	0.18 ± 0.18	1.97 ± 1.13	2.65 ± 3.76
\mathbf{B}^{RR} & Bauchgurt (2D)	1.35 ± 0.51	0.42 ± 0.52	1.91 ± 1.07	8.10 ± 11.69
\mathbf{B}^{RR} & Tiefenb. 10000 P.	1.38 ± 0.55	2.32 ± 3.81	1.91 ± 0.79	173.25 ± 169.22

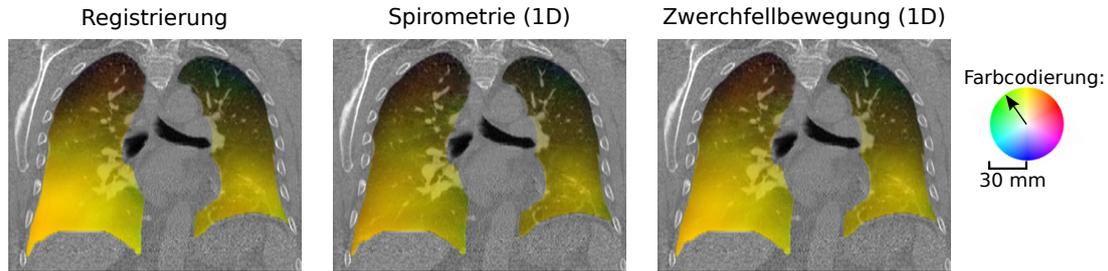


Abb. 3.4: Farbcodierte Darstellung der mittels direkter Bildregistrierung und Korrespondenzmodell (1D-Spirometersignal, 1D-Zwerchfellbewegung) geschätzten Atembewegung zwischen maximaler Ein- und Ausatmung für Case02 der CREATIS-Daten. Die mittlere zu kompensierende Bewegung für diesen Fall liegt bei 14.04 mm und der jeweils erreichte TRE bei 1.36 mm (Registrierung), 4.17 mm (Spirometrie) und 2.02 mm (Zwerchfellbewegung). Nimmt man die durch Registrierung ermittelte Bewegung als approximative Grundwahrheit an, zeigt sich insbesondere im unteren Bereich des rechten Lungenflügels (im Bild links) eine deutlich bessere Schätzung der Bewegungsamplitude bei Nutzung der Zwerchfellbewegung im Vergleich zum Spirometersignal.

den Ergebnissen der Korrespondenzmodelle, obwohl die Abstände der modellbasierten Schätzung im Mittel ca. 25% (Experiment 2) bzw. ca. 83% (Experiment 3) höher als die von *VariationalRegistration* sind.

Frage 2: Gibt es Genauigkeitsunterschiede zwischen den einzelnen Surrogattypen? Vergleicht man die vier simulierten Surrogattypen bei gleicher Dimension des generierten Atemsignals (1D bzw. 2D), lassen sich zwischen den einzelnen Signalen in den Experimenten 1–3 keine signifikanten Genauigkeitsunterschiede bezüglich der Mittelwerte von TRE bzw. Oberflächendistanz feststellen ($p \geq 0.05$). Einzige Ausnahme bildet der per Tiefenbildgebung verfolgte Punkt im Bereich des Brustbeins. Dessen Ergebnisse sind für Experiment 1 nahezu vollständig signifikant schlechter ($p < 0.05$) als die der anderen Surrogate (Ausnahme hier: 1D-Bauchgurtsignal vs. 1D-Tiefenbild Brustbein für EI→MI/ME mit $p = 0.09$).

Auch wenn die über alle Patienten gemittelten Ergebnisse überwiegend keine signifikanten Unterschiede zwischen den einzelnen Surrogattypen aufzeigen, so gibt es dennoch patientenspezifisch teilweise deutliche Genauigkeitsunterschiede. Als Beispiel sei hier Case02 der CREATIS-Daten genannt, für den Spirometrie- und Bauchgurtsignale im Vergleich zur Zwerchfellbewegung einen wesentlich höheren TRE für die geschätzte Bewegung EI→EE aufweisen (1D-/2D-Spirometrie: 4.17 mm/4.15 mm; 1D-/2D-Bauchgurt: 4.14 mm/4.32 mm; 1D-/2D-Zwerchfell: 2.02 mm/2.49 mm; siehe Abb. 3.4).

Frage 3: Welchen Einfluss hat die Signaldimension auf die Genauigkeit?

Beim Vergleich der 1D- und 2D-Atemsignale (1D-Signal + zeitliche Ableitung) des gleichen Surrogats zeigen sich nur für die Ergebnisse von Experiment 1 bei Schätzung der Bewegung zu den Atemmittellagen (EI→MI/ME) nahezu vollständig signifikant höhere Genauigkeiten für die 2D-Signale ($p < 0.001$; Ausnahme: 1D/2D-Tiefenbild Brustbein mit $p = 0.56$). Alle anderen Unterschiede zwischen 1D- und 2D-Signalen sind nicht statistisch signifikant.

Erhöht man die Dimension des Surrogatsignals weiter (Tiefenbildgebung mit $n_{sur} \geq 10$) unter Beibehaltung der Regressionsmethode \mathbf{B}^{RR} , so ergeben sich zwar insbesondere in Experiment 1 bei Schätzung der Bewegung zu den Atemmittellagen (EI→MI/ME) nahezu für alle Kombinationen signifikant bessere Ergebnisse für die höherdimensionalen Signale gegenüber den 1D-Signalen ($p < 0.01$; Ausnahmen: 10000 Punkte vs. 1D-Spirometrie und 1D-Brustgurt mit $p = 0.30$ und $p = 0.12$). Gegenüber den 2D-Signalen sind allerdings keine signifikanten Unterschiede festzustellen. Für den anderen Teil von Experiment 1 zeigt sich wieder kein klarer Unterschied zwischen 1D- und mehrdimensionalen Signalen. Dies gilt auch für die Experimente 2 und 3.

Vergleicht man die mittels simulierter Tiefenbildgebung generierten Signale mit unterschiedlicher Dimension mit $n_{sur} \geq 10$ und \mathbf{B}^{RR} untereinander, so ähneln sich die Ergebnisse zwischen den einzelnen Signalen sehr stark. Signifikante Unterschiede zwischen den Signalen sind nur in Einzelfällen für Experiment 1 vorhanden (EI→EE: 10 Punkte vs. 1000 Punkte mit $p < 0.05$; EI→MI/ME: 100 bzw. 1000 Punkte vs. 10000 Punkte mit $p < 0.05$). Bei den Ergebnissen von Experiment 1 ist zudem zu beobachten, dass die TRE-Werte bei zunehmender Anzahl von Abtastpunkten nicht monoton abfallen. In den anderen beiden Experimenten treten überhaupt keine signifikanten Unterschiede zwischen den Signalen auf.

Frage 4: Erhöhen projektionsbasierte Regressionsmethoden die Genauigkeit?

Eine signifikant höhere Genauigkeit gegenüber \mathbf{B}^{RR} erreichen die drei projektionsbasierten Regressionsmethoden bei gleichem Surrogatsignal nur in vier von 32 möglichen Vergleichen in Experiment 1 (EI→EE: 10000 Punkte mit \mathbf{B}^{RR} vs. \mathbf{B}^{PCR} und \mathbf{B}^{PLS} ; EI→MI/ME: 1000 Punkte mit \mathbf{B}^{RR} vs. \mathbf{B}^{CCA} sowie 10000 Punkte mit \mathbf{B}^{RR} vs. \mathbf{B}^{PCR} und \mathbf{B}^{PLS} ; alle $p < 0.05$). In Experiment 2 und 3 ergeben sich keine signifikant besseren Ergebnisse für die projektionsbasierten Methoden.

Statistisch signifikante Unterschiede zwischen den Ergebnissen der drei projektionsbasierten Verfahren bei gleichem Surrogatsignal ergeben sich nur in vier von 24 Vergleichen für Experiment 1 (EI→MI/ME: 100 und 1000 Punkte mit \mathbf{B}^{CCA} vs. \mathbf{B}^{PCR} und \mathbf{B}^{PLS} ; alle $p < 0.05$). In allen diesen Fällen liefert \mathbf{B}^{CCA} schlechtere Ergebnisse. In Experiment 2 und 3 ergeben sich keine signifikanten Unterschiede zwischen den drei Verfahren.

Frage 5: Inwieweit kompensieren die Korrespondenzmodelle die verschiedenen Arten von Bewegungsvariabilität? In Experiment 1 (EI→MI/ME) hat sich gezeigt, dass insbesondere die Kompensation der intrazyklischen Variabilität entscheidend von der Dimension des Surrogatsignals abhängt und hier mehrdimensionale Signale zumeist signifikante Vorteile gegenüber den eindimensionalen Signalen haben (siehe auch Frage 3). Betrachtet man beispielhaft die Ergebnisse des Spirometersignals, zeigt sich, dass bei intrazyklischer Variabilität in Experiment 1 (EI→MI/ME) durch die Nutzung des 2D-Signals eine Verbesserung des mittleren TRE um ca. 0.14 mm im Vergleich zum 1D-Signal erreicht werden kann. Für das 2D-Signal ist der TRE hierbei um ca. 0.21 mm schlechter als der des Registrierungsverfahrens. Bei simulierter interzyklischer Variabilität (EI→EE) ergibt sich für das 2D-Spirometersignal ein um ca. 0.56 mm höherer TRE als für das Registrierungsverfahren. Hierbei ist allerdings zu beachten, dass die zu kompensierende Bewegung bei EI→EE im Vergleich zu EI→MI/ME auch mehr als doppelt so hoch ist (3.59 mm vs. 7.59 mm). Für die anderen Surrogatsignale ergeben sich ähnliche Ergebnisse. Insgesamt lässt sich festhalten, dass die Modelle für $n_{sur} \geq 2$ bei intra- und interzyklischer Variabilität in der Lage sind einen Großteil der Atembewegung zu kompensieren.

In Experiment 2 kompensieren die surrogatbasierten Verfahren mit $n_{sur} \geq 2$ bei intrafraktioneller Bewegungsvariabilität ca. 46–56% der mittleren Bewegung (Kompensation: TRE geteilt durch mittlere Bewegung). Dies entspricht ungefähr der auch in Experiment 1 für die rein intrazyklische Variabilität erreichten prozentualen Bewegungskompensation. Im Vergleich zu den Ergebnissen bei simulierter intrazyklischer Variabilität in Experiment 1 ist die prozentuale Bewegungskompensation hier geringer. Insgesamt zeigt sich aber, dass die Oberflächenabstände bei intrafraktioneller Bewegungsvariabilität, insbesondere für die besten Surrogate (z. B. 2D-Spirometrie), nur knapp unterhalb der durch direkte Registrierung der Bilder erreichten Abstände liegen.

Wie in [Wilms u. a., 2014a] gezeigt, ergeben sich bei einem patientenspezifischen Vergleich der Ergebnisse der Experimente 2 und 3 für gleiche Surrogatsignale und Regressionsmethoden zumeist signifikant bessere Ergebnisse für die Korrespondenzmodelle in Experiment 2. Dies deutet darauf hin, dass die Korrespondenzmodelle in Experiment 3 Probleme haben die im Vergleich zu Experiment 2 zusätzlich auftretende interfraktionale Variabilität zu kompensieren. Im Mittel steigen die Oberflächenabstände zwischen Experiment 2 und 3 um ca. 47%.

Frage 6: Inwieweit unterscheiden sich die Ergebnisse diffeomorpher und nicht-diffeomorpher Korrespondenzmodelle? Die Ergebnisse von Experiment 4 zum Vergleich diffeomorpher und nicht-diffeomorpher Korrespondenzmodelle sind in Tab. 3.2 zusammengefasst. Es zeigt sich, dass die Unterschiede bezüglich der erreichbaren mittleren Genauigkeit der Bewegungsschätzung von diffeomorphen und nicht-

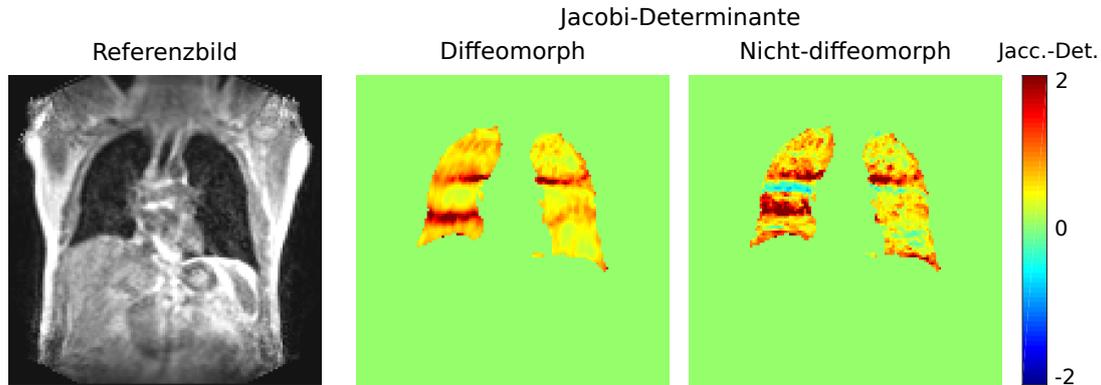


Abb. 3.5: Farbcodierte Darstellung der pixelweise berechneten Jacobi-Determinante für eine koronare Schicht von geschätzten Verschiebungsfeldern (diffeomorph vs. nicht-diffeomorph; 1D-Spirometriesignal) in Experiment 3 (4D-CT & 4D-MRT) für einen Patienten der DKFZ-Daten. Das dargestellte MRT-Bild zeigt dieselbe koronare Schicht und dient zur anatomischen Orientierung. Beide Korrespondenzmodelle (diffeomorph bzw. nicht-diffeomorph) erreichen eine ähnliche Genauigkeit (Oberflächendistanz: 3.51 mm vs. 3.60 mm), unterscheiden sich aber deutlich in der Anzahl von Singularitäten (0 vs. 1833 in 3D).

diffeomorphen Korrespondenzmodellen sehr gering sind (≈ 0.01 – 0.02 mm bei TRE und ≈ 0.02 – 0.04 mm bei den Oberflächendistanzen). Die meisten Differenzen zwischen den beiden Modellarten in Tab. 3.2 sind nicht statistisch signifikant. Ausnahmen bilden hier allein die Ergebnisse der 4D-CT-Daten für EI→MI/ME bei Nutzung des Spirometersignals (1D und 2D) oder des 2D-Bauchgurtsignals (jeweils $p < 0.05$).

Deutliche Unterschiede zeigen sich aber, wie erwartet, bei der Anzahl der Singularitäten in den geschätzten Feldern (siehe Abb. 3.5). Für die diffeomorphen Korrespondenzmodelle sind hier insbesondere bei den 4D-CT-Bilddaten und 4D-MRT-Bilddaten aus den Experimenten 1 und 2 keine Singularitäten zu finden. Für die nicht-diffeomorphen Korrespondenzmodelle treten nur bei der Schätzung der Bewegung zur Atemmittellage in den 4D-CT-Bilddaten keine Singularitäten auf. Dies lässt darauf schließen, dass nicht-diffeomorphe Korrespondenzmodelle nur dann Felder ohne Singularitäten produzieren, wenn das Signal/die Bewegung im Bereich der in den Trainingsdaten repräsentierten Signalwerte/Bewegung liegt und das Modell dann interpoliert. Verlässt das Signal/die Bewegung den Bereich der Trainingsdaten, erfolgt eine Extrapolation durch das Modell, bei der der Grad der Extrapolation vermeintlich die Anzahl der Singularitäten bestimmt. Im Endeffekt tritt dieses Phänomen auch bei den hier präsentierten diffeomorphen Korrespondenzmodellen auf (siehe Ergebnisse auf den Daten von Experiment 3 in Tab. 3.2), auch wenn hierfür der Extrapolationsfaktor entsprechend groß sein muss (> 4 für die Werte von Experiment 3 in Tab. 3.2), wodurch das gesuchte Geschwindigkeitsfeld entsprechend weit vom Ursprung des Tangentialraums

entfernt ist, was nach Anwendung der Scaling-and-Squaring-Methode dann zu einer geringen Anzahl von Singularitäten in den Verschiebungsfeldern führen kann. Extrapolationsfaktoren > 4 sind in der Praxis allerdings eher selten und treten in diesem Experiment vor allem durch den relativ rudimentären Ausgleich der Lage- und Deformationsunterschiede zwischen den Modalitäten und vereinzelte Hustensequenzen in den MRT-Bilddaten auf. Die Ergebnisse zeigen zudem, dass bei diffeomorphen Korrespondenzmodellen auch für große Extrapolationsfaktoren bedeutend weniger Singularitäten auftreten als bei dem Einsatz nicht-diffeomorpher Modelle. Weiterhin ist anzumerken, dass die Anzahl der Singularitäten bei diffeomorphen Korrespondenzmodellen teilweise auch durch eine Anpassung der Parameter der Scaling-and-Squaring-Methode (Anzahl der Schritte; hier: sechs Schritte) noch verbessert werden kann [Werner u. a., 2012b].

Ein größerer Unterschied zwischen beiden Verfahren ergibt sich allerdings bei einem Vergleich der jeweils benötigten Rechenzeiten. Für Laufzeitmessungen wurde ein Rechner mit Intel i7-5775C 3.30 GHz-Prozessor und 32 GB RAM Arbeitsspeicher genutzt und die selbst entwickelte, nicht-laufzeitoptimierte C++-Software (reine CPU-Nutzung) eingesetzt, welche für diese Evaluation genutzt wurde. Die eigentliche Schätzung des Geschwindigkeitsfeldes bzw. des Verschiebungsfeldes benötigt je nach Verfahren bzw. Surrogatdimension 15 ms bis 80 ms, wobei bei dem diffeomorphen Framework dann noch der zeitintensive Einsatz (≈ 5 Sekunden) der Scaling-and-Squaring-Methode folgt, um die endgültige Transformation zu erhalten. Diese Werte können allerdings durch eine Verlagerung der Berechnungen auf eine GPU (Nvidia GeForce GTX 1080 Ti; Testimplementierung in MATLAB-Software) deutlich reduziert werden. Die eigentliche Schätzung des Geschwindigkeits- bzw. Verschiebungsfeldes dauert auf diese Weise zwischen < 1 ms und 5 ms. Wird das Scaling-and-Squaring ebenfalls auf die GPU verlagert, ergibt sich für diesen Schritt eine zusätzliche Rechenzeit von ca. 75 ms. Bei diesen Rechenzeiten ist allerdings zu berücksichtigen, dass sie sich auf die in dieser Evaluation durchgeführte Schätzung der Lungenbewegung beziehen. Soll die Bewegung größerer Regionen geschätzt werden, werden sich die Rechenzeiten erhöhen.

3.3 Zusammenfassung und Einordnung der Ergebnisse

In diesem Kapitel wurde das in Kapitel 2 beschriebene Framework zur patientenspezifischen diffeomorphen Korrespondenzmodellierung umfangreich und systematisch evaluiert. Der Fokus lag hierbei nicht nur auf der generellen Untersuchung der mittels surrogatbasierter Bewegungsschätzung erreichbaren Genauigkeit unter Einfluss verschiedener Arten von Bewegungsvariabilität und dem Vergleich diffeomorpher und nicht-diffeomorpher Korrespondenzmodellierung, sondern auch auf dem systematischen Vergleich verschiedener Surrogattypen und Regressionsmethoden.

Um die genannten Anforderungen an die Evaluation umzusetzen, wurden auf der Basis verschiedener Arten von räumlich-zeitlichen 4D-CT- und 4D-MRT-Bilddaten

ein- und mehrdimensionale Surrogatsignale und verschiedene praxisnahe Szenarien zur surrogatbasierten Bewegungsschätzung mit unterschiedlichen Atmungsvariabilitäten simuliert. Neben der großen Flexibilität in Bezug auf die untersuchten Surrogate und Atmungsvariabilitäten ist ein weiterer großer Vorteil der durchgeführten bildbasierten Evaluation die Möglichkeit des Rückgriffs auf etablierte, landmarkenbasierte Techniken zur Genauigkeitsbestimmung aus dem Bereich der Bildregistrierung. Zudem wurden verschiedene öffentlich verfügbare 4D-CT-Bilddaten in diese Studie integriert, was zumindest in Teilen die Reproduzierbarkeit der Ergebnisse und insbesondere einen zukünftigen studienübergreifenden Ergebnisvergleich ermöglicht. Auch wenn einzelne Aspekte dieser Studie schon in der Literatur von anderen Autoren untersucht bzw. umgesetzt wurden, ist die hier durchgeführte Evaluation in ihrer Tiefe und ihrem Umfang neu und nach Kenntnis des Autors dieser Arbeit so nirgendwo in der Literatur zu finden.

Die Ergebnisse der Studie lassen sich in sieben Kernaussagen zusammenfassen:

1. Auch wenn durch die surrogatbasierte Bewegungsschätzung ein Großteil der Atembewegung kompensiert werden konnte, ist die erreichbare Schätzungsgenauigkeit im Vergleich zur direkten Registrierung der korrespondierenden Bilder teilweise deutlich geringer. Es konnten je nach Surrogat und vorhandener Atmungsvariabilität in Extremfällen Unterschiede größer Faktor 2 bei TRE/Oberflächendistanz festgestellt werden, wenngleich der mittlere Genauigkeitsunterschied ungefähr dem 1.4-fachen des per Registrierung erreichten TRE bzw. der erreichten Oberflächendistanz entspricht.
2. Bis auf vereinzelte Ausnahmen und patientenspezifische Unterschiede führen im Mittel alle untersuchten 1D-/2D-Surrogatsignale (Spirometer, Bauchgurt, Zwerchfellbewegung, punktweise Bewegung der Körperoberfläche) bei gleicher Signaldimension zu sehr ähnlichen Ergebnissen bei der surrogatbasierten Bewegungsschätzung.
3. Die Genauigkeit der surrogatbasierten Bewegungsschätzung lässt sich durch die Nutzung mehrdimensionaler Surrogatsignale (≥ 2) signifikant erhöhen. Hierbei ist es unerheblich, ob die Dimensionserhöhung durch das Hinzufügen zeitlicher Information (zeitliche Ableitung) oder durch direkte Erhebung mehrdimensionaler Signale (Tiefenbildgebung der Körperoberfläche) erfolgt.
4. Die für die Verarbeitung der sehr hochdimensionalen Tiefenbildgebungssignale (10–10000 Punkte/Dimensionen) in das Framework integrierten projektionsbasierten Regressionsmethoden haben im Vergleich zur klassischen Regression mit Frobeniusnorm-Regularisierung/Ridge-Regression nur in Einzelfällen zu einem Genauigkeitsvorteil geführt und ihr Einsatz scheint somit im Allgemeinen nicht zwingend notwendig.

5. Es hat sich gezeigt, dass beim Vorliegen von intra- und interzyklischer Bewegungsvariabilität bzw. zusätzlicher intrafraktioneller Bewegungsvariabilität bei Verwendung mehrdimensionaler Surrogatsignale ein Großteil der Atembewegung durch surrogatbasierte Bewegungsschätzung kompensiert werden kann. Die Genauigkeit sinkt allerdings erheblich, wenn interfraktionelle Variabilität hinzukommt (Fehler der Bewegungsschätzung erhöht sich im Mittel um 47%).
6. Die Unterschiede bezüglich der erreichbaren Genauigkeit von diffeomorphen und nicht-diffeomorphen Korrespondenzmodellen sind im Mittel sehr gering (kleiner 0.05 mm) und das diffeomorphe Framework verhindert in den meisten praktisch relevanten Fällen erfolgreich die beim Standardvorgehen auftretenden, physiologisch unplausiblen Singularitäten in den geschätzten Transformationen.
7. Rechenzeiten von deutlich unter 10 ms bzw. unter 100 ms für die nicht-diffeomorphen/diffeomorphen Korrespondenzmodelle belegen die Praxistauglichkeit der präsentierten Verfahren.

Insgesamt zeigt sich, dass die hier untersuchten linearen Korrespondenzmodelle auch bei Vorliegen verschiedener Arten von Bewegungsvariabilität generell in der Lage sind, einen großen Teil der intrainterventionellen Atembewegung zu kompensieren und so in der Praxis zu einer Verbesserung der jeweiligen Therapie beitragen können. Betrachtet man allerdings die Genauigkeit, die durch die direkte Registrierung der für die Evaluation genutzten Bilddaten erreicht werden kann, ist klar, dass hier noch Potenzial für Verbesserungen besteht. Diese Tatsache motiviert somit u. a. in weiterführenden Arbeiten die Integration nichtlinearer Ansätze in das vorliegende Framework, unter der Annahme, dass der lineare Zusammenhang der limitierende Faktor ist.

Bezüglich der Ergebnisse in den einzelnen Experimenten mit verschiedenen Bewegungsvariabilitäten ist festzuhalten, dass insbesondere die interfraktionelle Bewegungsvariabilität im zugehörigen Experiment mit 4D-CT- und 4D-MRT-Bilddaten nur unzureichend kompensiert werden konnte. Dies weist darauf hin, dass der potenzielle klinische Workflow, bei dem das Korrespondenzmodell auf Basis von Daten einer Planungssitzung erstellt wird und dann in nachfolgenden Behandlungssitzungen zum Einsatz kommt, mit Problemen behaftet ist. Es bleibt hier allerdings unklar, ob diese Probleme ausschließlich einer starken Veränderung der Atmung zwischen den Aufnahmen geschuldet sind oder auch durch die für die Evaluation nur sehr rudimentär durchgeführte Angleichung der Koordinatensysteme von CT- und MRT-Bilddaten begründet sind. Auch dieses Ergebnis motiviert somit weitere Experimente, bei denen u. a. auch die Anpassung der Korrespondenzmodelle mittels z. B. fraktionsspezifisch bildbasiert ermittelter Parameter (siehe z. B. [Fassi u. a., 2014]) untersucht werden sollte, um die Auswirkungen interfraktioneller Bewegungsvariabilität besser kompensieren zu können.

Auffälliges Ergebnis der Evaluation ist die generell weitgehende Äquivalenz der Ergebnisse für die verschiedenen Surrogattypen bei gleicher Signaldimension. Diese lässt sich allerdings u. a. anhand der starken Korrelation der verschiedenen Surrogatsignale (oftmals lineare Korrelation ≥ 0.9) erklären, die andeutet, dass die einzelnen Signale zumeist nur einen geringen Anteil exklusiver Information beinhalten, der zudem nicht zwingend mit der internen Bewegung zusammenhängen muss. Einen Teil der hohen Korrelation der Signale ist durch die bildbasierte Simulation erklärbar, da hierdurch u. a. Phasenverschiebungen verloren gehen können und technische Details wie unterschiedliche Signal-Rausch-Verhältnisse nicht abgebildet werden. Andererseits wurden hohe Signalkorrelationen in der Literatur auch für verschiedenste reale Surrogate beobachtet (siehe z. B. [Dürichen, 2015; Lu u. a., 2005a; Ortmüller u. a., 2015]), was generell schon deshalb plausibel erscheint, da sie alle den gleichen Prozess abbilden. Die verbleibenden Unterschiede zwischen den Signalen motivieren allerdings eine genauere Untersuchung der (patientenspezifischen) Kombination verschiedener Surrogate (z. B. verschiedene Typen oder verschiedene Positionen), um die Bewegungsschätzung durch die Nutzung komplementärer Anteile in den Signalen zu verbessern. Dies wurde schon in [Koch u. a., 2004] vorgeschlagen und u. a. in [Dürichen, 2015; Savill u. a., 2011] erfolgreich experimentell untersucht. Die hier vorgestellte Simulationsumgebung bietet hierfür sehr gute Voraussetzungen und erlaubt die automatisierte Auswahl der patientenspezifisch optimalen Kombination während des Planungsprozesses. Dieses Vorgehen wurde unter maßgeblicher Mitwirkung des Autors dieser Dissertation in [Ortmüller u. a., 2014] schon testweise umgesetzt.

Dass durch die Nutzung mehrdimensionaler Surrogatsignale die Genauigkeit der Bewegungsschätzung gegenüber den eindimensionalen Signalen verbessert werden konnte, entspricht der generellen theoretischen Erwartung und dem Wissen aus anderen Studien [McClelland u. a., 2013]. Interessant ist allerdings, dass beim Einsatz von Tiefenbildgebung die Anzahl der abgetasteten Punkte keinen großen Einfluss auf die Genauigkeit hat, was ebenso für unterschiedliche (optimierte) Abtastmuster (Regionen, Linien, etc.) gilt, wie sich in zusätzlichen Experimenten in [Wilms u. a., 2014c] gezeigt hat. Es ist allerdings anzumerken, dass tiefenbildgebende Verfahren durch ihre berührungslose, nichtinvasive Aufnahmetechnik weiterhin große Vorteile im praktischen Einsatz gegenüber Standardsurrogaten haben. Zudem ist interessant, dass sich durch die Nutzung projektionsbasierter Regressionsmethoden kein nachweisbarer Vorteil gegenüber der klassischen Frobeniusnorm-Regularisierung/Ridge-Regression gezeigt hat, was durch die breite Nutzung dieser Verfahren in der Literatur doch überrascht und dessen genaue Gründe in weiteren Experimenten untersucht werden sollten.

Die geringen Genauigkeitsunterschiede zwischen diffeomorphen und nicht-diffeomorphen Korrespondenzmodellen zeigen, dass die zumindest in der Theorie mit potenziellen Limitierungen behaftete Restriktion auf zeitunabhängige Geschwindigkeitsfelder in der Praxis kein grundsätzliches Problem darstellt und vernachlässigt werden kann. Hinsichtlich der Rechenzeiten sollte insbesondere bei der GPU-Implementierung des

Scaling-and-Squaring noch Potenzial für Verbesserungen vorhanden zu sein. Diese sind auch notwendig, wenn sehr große Regionen in Echtzeit verarbeitet werden sollen.

Unabhängig von den schon aufgeführten Anmerkungen lassen sich insgesamt drei generelle Kritikpunkte an der präsentierten Studie formulieren, die in Folgearbeiten aufgegriffen werden sollten:

1. Die Nutzung von bildbasiert simulierten Surrogatsignalen hat zwar den großen Vorteil, dass hierdurch leicht ein umfangreicher Vergleich verschiedener Surrogattypen möglich wird, die Simulation birgt aber auch das Risiko, dass in der Praxis entscheidende (technische) Eigenschaften der einzelnen Surrogate nicht ausreichend genau erfasst werden. Dies ist auch für die in dieser Studie simulierten Surrogate nicht auszuschließen, da beispielsweise keinerlei Rauschen implementiert wurde, auch wenn bis auf das stark vereinfacht implementierte Tiefenbildgebungsverfahren, alle anderen Surrogatsimulationen mittels realer Daten verifiziert wurden. Es erscheint deshalb sinnvoll, die Studie mit realen Signalen zu wiederholen und/oder die Tiefenbildgebungssimulation zu verifizieren.
2. Die Nutzung der verschiedenen 4D-CT- und 4D-MRT-Bilddatensätze ermöglichte in dieser Studie u. a. den Einsatz bekannter landmarkenbasierter Evaluationstechniken und die Simulation unterschiedlicher Atmungsvariabilitäten. Bei der Bewertung der Aussagekraft der Ergebnisse, die auf den drei 4D-MRT-Bilddatensätzen beruhen, muss allerdings die geringe Anzahl an Daten und deren geringe räumliche Auflösung berücksichtigt werden. Es wäre deshalb wünschenswert für die Evaluation von intra- und interfraktioneller Bewegungsvariabilität weitere und höher aufgelöste Daten hinzuzuziehen. Für die Experimente auf Basis der 4D-CT-Bilddaten standen zwar 28 Datensätze zur Verfügung, es muss allerdings kritisch angemerkt werden, dass die 4D-CT-Bilddaten, bedingt durch die Rekonstruktionstechniken, eigentlich eine Mischung aus verschiedenen Atemzyklen darstellen, was theoretisch auch einen Einfluss auf die Studienergebnisse haben könnte.
3. Alle Untersuchungen in dieser Studie beziehen sich ausschließlich auf die Lungenregion, auch wenn andere Tumorregionen oder Risikoorgane (z. B. die Leber) ebenso von der Atmung betroffen sind. Diese Einschränkung ist hauptsächlich der Datenverfügbarkeit geschuldet, da das präsentierte Framework prinzipiell mit Verschiebungsfeldern bzw. Transformationen beliebiger Größe umgehen kann. Im Fall von diffeomorphen Transformationen ist das einzige methodische Hindernis bei einer Erweiterung auf andere Strukturen die korrekte Behandlung von Unstetigkeiten in der Bewegung, die beispielsweise durch Gleitbewegungen in Bereichen zwischen zwei Strukturen auftreten können. Zur Lösung dieses Problems ist aber die Konstruktion von stückweise diffeomorphen Korrespondenzmodellen möglich, da die stückweise Behandlung der einzelnen Regionen ein Weg ist, der auch bei der diffeomorphen Registrierung bei Gleitbewegungen vorgeschlagen wurde [Risser u. a., 2013].

Kapitel 4

Patientenspezifische Simulation der Atembewegung mit Korrespondenzmodellen

Kapitelinhalt

4.1 B-Spline-basierte Modellierung von Surrogatsignalen	58
4.1.1 Zerlegung des Surrogatsignals in Atemzyklen	59
4.1.2 B-Spline-basierte Approximation einzelner Atemzyklen	60
4.1.3 Von der statistischen Modellierung zur Signal- und Bewegungssimulation	62
4.1.4 Evaluation der Simulationsmethode	64
4.1.5 Zusammenfassung	67
4.2 Anwendung 1: 4D-Dosissimulation in der Strahlentherapieplanung	67
4.2.1 4D-Dosissimulation durch modellbasierte Dosisakkumulation	69
4.2.2 Statistische Untersuchung atmungsbedingter Effekte auf die Dosisverteilung	70
4.2.3 Experimente mit IMRT-Bestrahlungsplänen	72
4.2.4 Ergebnisse	73
4.2.5 Zusammenfassung	74
4.3 Anwendung 2: Virtual-Reality-Training von Punktionseingriffen	77
4.3.1 Vom statischen zum atmenden virtuellen Patienten	78
4.3.2 Initiale Experimente und deren Ergebnisse	81
4.3.3 Zusammenfassung und weitere Überlegungen	83
4.4 Zusammenfassung	84

Auch wenn patientenspezifische Korrespondenzmodelle bisher hauptsächlich für die intrainterventionelle Schätzung der Atembewegung eingesetzt werden, sind auch andere

Einsatzbereiche denkbar. Als Beispiel zeigt dieses Kapitel auf, wie Korrespondenzmodelle gewinnbringend zur surrogatbasierten, patientenspezifischen Simulation variabler und realistischer Atembewegung in medizinischen Planungs- und Simulationsumgebungen zum Einsatz kommen können.

Wie in Kapitel 1 einführend erläutert, führt die Atmung zu einer Deformation bzw. Bewegung vieler thorakaler und abdominaler Strukturen und beeinflusst damit nahezu alle diese Regionen betreffende Bildgebungs- und Therapieprozesse. Dies bedeutet, dass die in der Praxis zur Simulation, Planung oder Verifikation dieser Prozesse eingesetzten computerbasierten Verfahren die Atembewegung berücksichtigen sollten, um ein realitätsnahes Ergebnis zu erhalten. Die folgenden praktischen Beispiele dienen hierbei als Motivation für das in diesem Kapitel präsentierte Simulationsverfahren.

Im Bereich der Simulation nuklearer Bildgebungsprozesse werden u. a. virtuelle 4D-Phantome einer Referenzanatomie genutzt (siehe [Ehrhardt u. a., 2013; Zaidi u. a., 2009] für Übersichten und Einordnungen), die neben der Anatomie auch die Atembewegung modellieren. Bekanntes Beispiel ist das 4D-XCAT-Phantom [Segars u. a., 2010], welches aufbauend auf den Daten des Visible Human Project [Ackerman, 1998] die Oberflächen anatomischer Strukturen modelliert, die dann zur Atmungssimulation deformiert werden. Das hierfür benötigte Atemmodell wurde durch Analyse der Atembewegung eines Probanden in einem 4D-Bilddatensatz manuell definiert. Nachteil der Nutzung einer bestimmten Referenzanatomie ist die fehlende anatomische Variabilität, wodurch patientenspezifische Simulationen ohne weitere Anpassungen nicht möglich sind. Gleiches gilt ebenso für die Atembewegung, die in der Regel nur sehr eingeschränkt variiert oder angepasst werden kann. Weiterhin ist durch die oberflächenbasierte Anatomiebeschreibung eine aufwendige und stark vereinfachende Bewegungsinterpolation notwendig, wenn dichte Verschiebungsfelder benötigt werden.

Eine Möglichkeit zur Verifikation und Analyse von Bestrahlungsplänen für atmungsbewegte thorakale oder abdominale Tumoren stellt die 4D-Dosisberechnung durch Dosisakkumulation dar (Beispiele: [Admiraal u. a., 2008; Guerrero u. a., 2005; Sothmann u. a., 2017b; Velec u. a., 2011; Werner u. a., 2012a; Wu u. a., 2008]). Die Basis für diese Verfahren bilden in der Regel ein 4D-CT-Bilddatensatz des Patienten und eine auf einer Referenzatemphase geplante 3D-Dosisverteilung. Durch Bildregistrierung wird dann die Atembewegung in den 4D-Bilddaten ausgehend von der Referenzphase bestimmt und die ermittelten Transformationen werden genutzt, um zu ermitteln, welche Strahlendosis ein Punkt in den einzelnen Atemphasen erhält. Durch zeitliche Akkumulation dieser Informationen ergibt sich dann eine 4D-Dosisverteilung, die die in den 4D-Bilddaten repräsentierte Atembewegung berücksichtigt [Werner, 2012]. Die Nutzung eines 4D-CT-Bilddatensatzes erlaubt zwar einerseits die Integration patientenspezifischer Bewegungsinformationen in die Dosisberechnung, andererseits wird aber die mögliche intrafraktionelle Bewegungsvariabilität vollständig ignoriert. Ähnliche Simulationen sind beispielsweise auch für HIFU möglich, um die Temperaturverteilung im Gewebe patientenspezifisch zu simulieren. Auch wenn beispielsweise in [Rijkhorst u. a.,

2011] für die HIFU-Simulation eine patientenspezifische Simulation der Atembewegung mithilfe eines Korrespondenzmodells erfolgt, so ist auch hier die simulierte Variabilität der Atmung auf die Variabilität der direkt aufgenommenen Daten beschränkt.

Die Berücksichtigung der Atembewegung spielt auch eine wichtige Rolle bei der Erhöhung der Immersion in medizinischen Virtual-Reality-Simulatoren. Bei der Simulation von Eingriffen im Thorax oder Abdomen wie der Nadelpunktion der Leber [Fortmeier u. a., 2013; Villard u. a., 2014] erscheint es deswegen sinnvoll, atmungsbelegte Deformationen der Körperstrukturen und die Bewegung der eingesetzten Instrumente in die Visualisierung sowie das für den Nutzer über spezielle Eingabegeräte fühlbare haptische Feedback zu integrieren. Verfügbare Simulatoren mit integrierter Atmungssimulation [Ni u. a., 2011; Villard u. a., 2014] haben hierbei gemeinsam, dass die Atmung ähnlich zu den 4D-Phantomen durch manuell definierte, generische Atemmodelle repräsentiert wird, die teilweise stark vereinfachende Annahmen beinhalten (z. B. Periodizität in [Ni u. a., 2011]) und deren patientenspezifische Adaption [Villard u. a., 2012] aufwendig ist.

Diese Beispiele zeigen, dass es für Verfahren, die eine realitätsnahe, patientenspezifische Simulation variabler Atembewegung ermöglichen, vielfältige Einsatzmöglichkeiten gibt, verfügbare Methoden aber den Anforderungen der Anwendungen selten wirklich gerecht werden. Deshalb wird in diesem Kapitel ein Verfahren zur patientenspezifischen Simulation variabler Atembewegung mit Korrespondenzmodellen vorgestellt. Den Ausgangspunkt bildet hierbei das in [Werner, 2012] vorgestellte Verfahren zur korrespondenzmodellbasierten 4D-Dosisimulation. Grundgedanke des dort skizzierten Ansatzes ist, dass nach dem Training eines patientenspezifischen Korrespondenzmodells ein (aufgezeichnetes) Surrogatsignal genutzt werden kann, um eine zeitliche Folge von Transformationen zu generieren, die die Atembewegung beschreiben.

Bei diesem Simulationsansatz ist neben den für das Modelltraining benötigten patientenspezifischen Daten (4D-Bilddaten mit Surrogatsignalmessungen) nur noch ein vergleichsweise einfach zu akquirierendes Surrogatsignal der gewünschten zeitlichen Simulationslänge nötig. Ein großer Vorteil des in [Werner, 2012] beschriebenen Verfahrens ist, dass durch ein variables Surrogatsignal (u. a. zeitliche Änderung der Ein-/Ausatemungslänge und Änderung der Signalamplitude) einfach eine Variabilität der Atmung erreicht werden kann. Weiterhin werden bei der Simulation dichte Verschiebungsfelder generiert, sodass sowohl Volumendaten (z. B. Bilder oder Dosis- und Temperaturverteilungen) als auch Oberflächen deformiert/animiert werden können. Die Berechnung der einzelnen Transformationen ist zudem in Sekundenbruchteilen möglich, wodurch der Einsatz in Echtzeitumgebungen wie Virtual-Reality-Simulatoren möglich ist. Speziell die fehlende Echtzeitfähigkeit ist ein Grund dafür, dass der Einsatz von Bewegungsmodellen basierend auf der Finite-Elemente-Methode (FEM) [Fuerst u. a., 2015; Giroux u. a., 2017; Werner u. a., 2009] in diesen Simulatoren schwierig ist.

Ein Nachteil des beschriebenen Vorgehens ist die direkte Abhängigkeit der Simulation vom aufgenommenen Surrogatsignal. Dies bezieht sich auf die zeitliche Länge und

auf die Variabilität der zeitlichen Abfolge verschiedener Atemzyklen. Aus diesem Grund wird das Bewegungssimulationsverfahren in dieser Arbeit um ein neues Verfahren zur Modellierung und anschließender Simulation von Surrogatsignalen auf Basis von B-Splines erweitert (Kapitel 4.1). Aufbauend auf einem aufgenommenen Surrogatsignal kann das Verfahren variable Surrogatsignale beliebiger zeitlicher Länge mit typischen patientenspezifischen Variationen von Zykluslänge oder Ein-/Ausatemungstiefe simulieren. Diese simulierten Surrogatsignale können dann wiederum zur korrespondenzmodellbasierten Simulation variabler interner Atembewegung eingesetzt werden. Dieses kombinierte Simulationsverfahren wird in Kapitel 4.2 dann in Anlehnung an [Werner, 2012] zur 4D-Dosisimulation in der Strahlentherapieplanung eingesetzt, um die Auswirkungen von variabler Atembewegung auf die geplante Dosisverteilung statistisch zu untersuchen. Kapitel 4.3 befasst sich zudem mit der Integration dieses Simulationsverfahrens in einen Virtual-Reality-Simulator, womit ein neuer Einsatzbereich für Korrespondenzmodelle erschlossen wird. Die Darstellungen in Kapitel 4.1 und Kapitel 4.2 basieren hauptsächlich auf [Wilms u. a., 2014b], wohingegen Kapitel 4.3 [Fortmeier u. a., 2015; Wilms u. a., 2015a] als Grundlage hat.

4.1 Simulation variabler Atembewegung durch B-Splinebasierte Modellierung von Surrogatsignalen

Dieses Teilkapitel beschreibt ein neues Verfahren zur Modellierung und Simulation von 1D-Surrogatsignalen (siehe Kapitel 2.1.1.1) basierend auf B-Splines. Ziel dieses Verfahrens ist es, variable und gleichzeitig realistische patientenspezifische Surrogatsignale beliebiger zeitlicher Länge für die korrespondenzmodellbasierte Simulation der Atembewegung zu generieren. Das Verfahren besteht aus vier Arbeitsschritten (siehe Abb. 4.1) und setzt ein aufgenommenes 1D-Surrogatsignal voraus: (1) In einem ersten Schritt wird das Originalsignal in einzelne Atemzyklen zerlegt, die anschließend (2) durch B-Splines approximiert werden, womit jeder Atemzyklus durch seinen Vektor von B-Spline-Koeffizienten parametrisiert ist. Anschließend werden (3) auf Basis der Koeffizientenvektoren aller Atemzyklen die Parameter einer multivariaten Normalverteilung geschätzt. (4) Durch die Ziehung von Koeffizientenvektoren aus der Normalverteilung können dann realistische, variable Surrogatsignale synthetisiert werden.

Die explizite Simulation von Surrogatsignalen hat bisher nur wenig Aufmerksamkeit erfahren. Die einzige vergleichbare Arbeit, die dem Autor dieser Dissertation bekannt ist, ist [Quirk u. a., 2012]. Dort wird ein Verfahren zu Generierung von Surrogatsignalen auf Basis einer Population von Beispielsignalen verschiedener Patienten beschrieben, wobei die Eigenschaften der generierten Signale aus den Eigenschaften der Population abgeleitet werden, um realistische Signale zu erhalten. Im Unterschied zu dem in diesem Kapitel präsentierten Verfahren, werden die einzelnen Atemzyklen in [Quirk u. a., 2012] allerdings mittels Sigmoidfunktionen repräsentiert, wobei Ein- und Ausat-

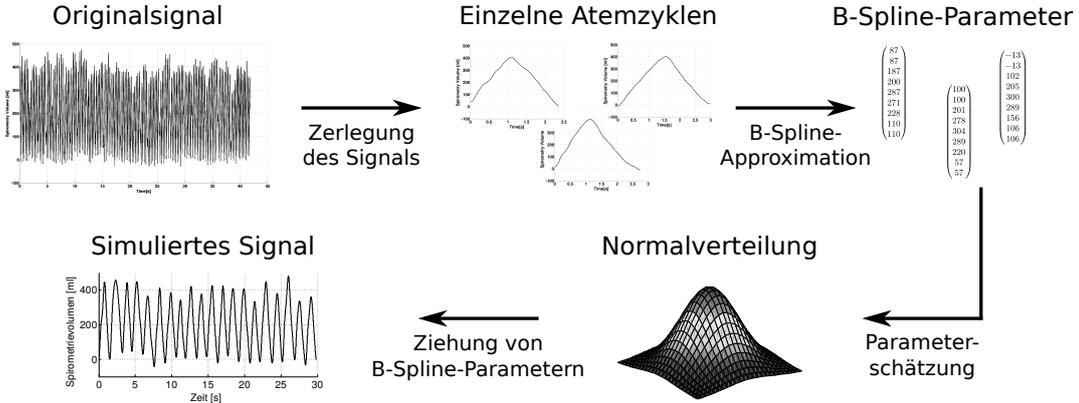


Abb. 4.1: Schematische Darstellung der einzelnen Schritte des B-Spline-basierten Verfahrens zur Modellierung und Simulation von Surrogatsignalen (hier: Spirometer). Die einzelnen Schritte sind im Text näher erläutert.

mung getrennt modelliert werden. Der hier propagierte Einsatz von B-Splines hat den Vorteil, dass jeder Atemzyklus durch eine einzelne Kurve repräsentiert werden kann, was beispielsweise Unstetigkeiten zwischen den Ein- und Ausatmungsteilen des Signals ohne weitere Maßnahmen verhindert.

Im weiteren Verlauf werden die einzelnen Schritte des Verfahrens näher beschrieben. Hierbei wird vorausgesetzt, dass ein zeitkontinuierliches 1D-Surrogatsignal $s(t) : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ mit n_{cyc} vollständigen Atemzyklen zur Verfügung steht. Sei weiterhin $S = (s_i)_{i=0}^{n_{sp}-1}$ die diskrete Version von $s(t)$ bestehend aus n_{sp} Signalwerten $s_i = s(i\Delta t)$, die durch zeitlich äquidistante Abtastung des kontinuierlichen Signals mit einer Abtastrate von $1/\Delta t$ erzeugt wurde.

4.1.1 Zerlegung des Surrogatsignals in Atemzyklen

In einem ersten Schritt erfolgt die Zerlegung des diskreten Originalsignals S in einzelne Atemzyklen (siehe Abb. 4.2). In Anlehnung an [Marx u. a., 2014] wird ein Atemzyklus hierbei als der Signalabschnitt $C_j = (s_i)_{i=l_j}^{r_j}$ zwischen zwei zeitlich aufeinander folgenden Zuständen maximaler Ausatmung (EE) definiert (hier: l_j und r_j mit $l_j < r_j$). Diese Zustände stellen sich als lokale Minima im Signal dar und können durch ableitungsbasierte Steigungsanalyse identifiziert werden. Zwei zeitlich aufeinander folgende lokale Minima bilden dann jeweils einen Atemzyklus. Da durch technische Effekte oder singuläre Ereignisse (z. B. Husten) während der Signalaufnahme nicht alle so extrahierten Atemzyklen wirklich vollständig oder für weitere Analysen brauchbar sind, wird im Anschluss eine Nachverarbeitung durchgeführt, bei der Atemzyklen entfernt werden, deren zeitliche Länge oder Amplitude außerhalb eines Toleranzbereichs liegt. Dieser Toleranzbereich wird patientenspezifisch anhand aller extrahierter Atemzyklen auf den Bereich der mittleren Länge/Amplitude \pm zweifache Standardabweichung fest-

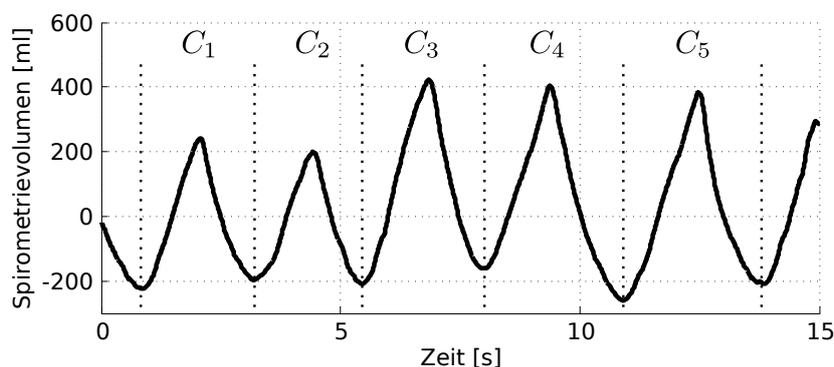


Abb. 4.2: Zerlegung eines 1D-Spirometersignals in fünf vollständige Atemzyklen. Die gestrichelten Linien markieren die detektierten lokalen Minima des Signals (= Zustände maximaler Ausatmung), die die Grenze zwischen zwei Atemzyklen bilden.

gelegt. Diese patientenspezifischen Werte haben sich als robuster erwiesen als die festen Grenzwerte in [Marx u. a., 2014].

4.1.2 B-Spline-basierte Approximation einzelner Atemzyklen

Die einzelnen Atemzyklen C_j werden jeweils durch eine eindimensionale, offene, nicht-periodische und nichtuniforme B-Spline-Kurve vom Grad $d = 3$ mit insgesamt $n_{cp} = 9$ Kontrollpunkten/Koeffizienten approximiert (siehe [Piegl u. a., 2005] für eine Einführung zu B-Spline-Kurven). Diese konsistente, niedrigdimensionale und kontinuierliche Parametrisierung erlaubt im Anschluss eine einfache und effiziente Weiterverarbeitung der Atemzyklen. Ein approximierter Atemzyklus \tilde{C}_j ist dann gegeben durch

$$\tilde{C}_j(t) = \sum_{h=1}^{n_{cp}} N_{h,d}(t) p_{j,h}, \text{ mit } t \in [l_j, r_j]. \quad (4.1)$$

Hierbei bezeichnet $N_{h,d}$ die B-Spline-Basisfunktion des Kontrollpunktes/Koeffizienten $p_{j,h} \in \mathbb{R}$ und das Intervall $[l_j, r_j]$ gibt den Definitionsbereich der Approximation mit zugehörigem Knotenvektor $[l_j, l_j, l_j, l_j, l_j + \frac{1}{6}(r_j - l_j), l_j + \frac{2}{6}(r_j - l_j), l_j + \frac{3}{6}(r_j - l_j), l_j + \frac{4}{6}(r_j - l_j), l_j + \frac{5}{6}(r_j - l_j), r_j, r_j, r_j, r_j]$ bestehend aus $d + n_{cp} + 1 = 13$ Elementen an. Die bei der Approximation eines Atemzyklus zu bestimmenden Parameter sind die neun Kontrollpunkte $\{p_{j,1}, \dots, p_{j,n_{cp}}\}$ (siehe auch Abb. 4.3).

Auch wenn die B-Spline-Kurve den gemessenen Atemzyklus generell nur approximiert, erfolgt durch die Wahl des Knotenvektors (offene B-Spline-Kurve, $d+1$ -fache Wiederholung an Anfang und Ende) eine Interpolation der Endpunkte, sodass $\tilde{C}_j(l_j) = s_{l_j} = p_{j,1}$ und $\tilde{C}_j(r_j) = s_{r_j} = p_{j,n_{cp}}$ gilt. Hintergrund dieser Wahl ist die Tatsache, dass die beiden Endpunkte zu den charakteristischsten Punkten eines Atemzyklus gehören. Da die Endpunkte zudem lokale Minima des Signals darstellen, soll auch diese Eigenschaft in der Approximation berücksichtigt werden. Hierzu werden der

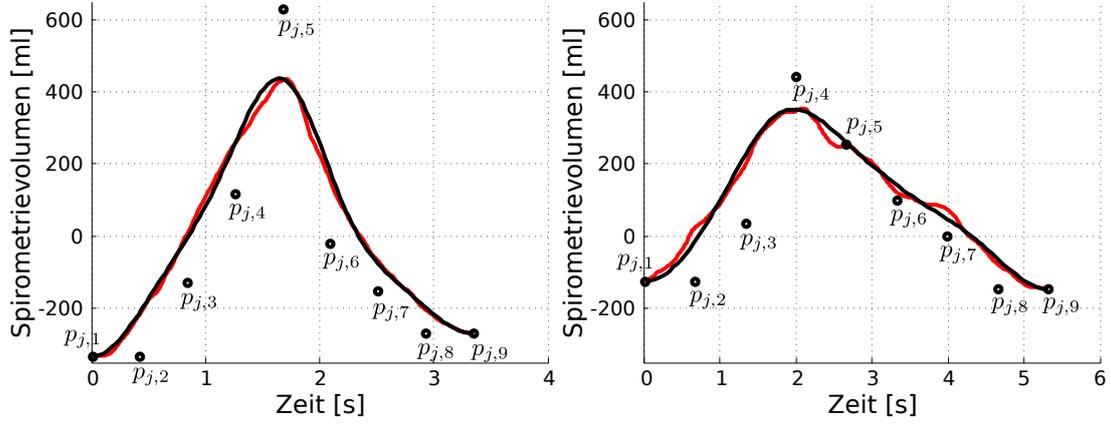


Abb. 4.3: Approximation zweier Atemzyklen eines Patienten mittels offener, nichtperiodischer und nichtuniformer B-Spline-Kurven vom Grad $d = 3$ mit jeweils $n_{cp} = 9$ Kontrollpunkten. Die schwarze Kurve zeigt die Approximation, wohingegen die rote Kurve das aufgenommene Spirometersignal darstellt. Die schwarzen Kreise symbolisieren die Kontrollpunkte. Hierbei ist zu beachten, dass jeweils die ersten und letzten beiden Kontrollpunkte denselben Wert haben. Insbesondere die B-Spline-Kurve des rechten Atemzyklus zeigt die Glättungseigenschaften der Approximation, da hier die mutmaßlich technisch bedingten Schwankungen des Originalsignals nicht vorhanden sind. Weiterhin sind beide Atemzyklen Beispiele für die Atmungsvariabilität eines Patienten, die hier gut durch die unterschiedlichen Verhältnisse von Ein- und Ausatemungsdauer und die unterschiedliche Gesamtdauer der beiden Zyklen zu erkennen ist.

zweite und der vorletzte Kontrollpunkt $p_{j,2}$ und $p_{j,n_{cp}-1}$ auf den Wert des vorherigen bzw. nachfolgenden Kontrollpunktes gesetzt: $p_{j,2} = p_{j,1}$ und $p_{j,n_{cp}-1} = p_{j,n_{cp}}$. Dies führt dazu, dass die erste Ableitung von \tilde{C}_j in den Endpunkten 0 ist [Piegl u. a., 2000], womit diese Punkte auch lokale Extrema der Approximation sind. Die verbleibenden fünf Kontrollpunkte werden so gewählt, dass die Summe der quadratischen Abstände zu C_j minimiert wird [Piegl u. a., 2000; Piegl u. a., 2005]:

$$\{p_{j,3}, \dots, p_{j,7}\} = \arg \min_{p_{j,3}, \dots, p_{j,7} \in \mathbb{R}} \sum_{i=l_j+1}^{r_j-1} \left\| s_i - \left(\sum_{h=1}^{n_{cp}} N_{h,d}(i) p_{j,h} \right) \right\|^2. \quad (4.2)$$

Da $p_{j,2} = p_{j,1}$ und $p_{j,n_{cp}-1} = p_{j,n_{cp}}$ gilt, kann der Signalverlauf jedes Atemzyklus C_j durch sieben Parameter beschrieben werden. Durch zusätzliche Berücksichtigung der zeitlichen Länge $t_j = r_j - l_j$ ergibt sich ein Vektor $\mathbf{c}_j = [p_{j,1} = p_{j,2}, \dots, p_{j,n_{cp}-1} = p_{j,n_{cp}}, t_j]^T$ mit acht Einträgen, der einen Atemzyklus vollständig beschreibt.

4.1.3 Von der statistischen Modellierung zur Signal- und Bewegungssimulation

Nachdem die gemessenen Atemzyklen jeweils durch einen Parametervektor \mathbf{c}_j repräsentiert werden, können diese Vektoren genutzt werden, um die Zyklen statistisch zu analysieren, zu modellieren und anschließend zu simulieren. Beispielsweise kann ein Atemzyklus mittlerer Form und zeitlicher Länge berechnet werden durch

$$\boldsymbol{\mu}_{\mathbf{c}} = \frac{1}{n_{cyc}} \sum_{j=1}^{n_{cyc}} \mathbf{c}_j . \quad (4.3)$$

Abbildung 4.4 zeigt beispielhaft drei auf diese Weise ermittelte mittlere Atemzyklen.

Um für die Simulation neuer, realistischer Atemzyklen aus den vorliegenden Parametervektoren ein generatives Modell zu erstellen, wird angenommen, dass die Parametervektoren einer multivariaten Normalverteilung $\mathcal{N}(\boldsymbol{\mu}_{\mathbf{c}}, \boldsymbol{\Sigma}_{\mathbf{c}})$ mit Mittelwertvektor $\boldsymbol{\mu}_{\mathbf{c}} \in \mathbb{R}^8$ und Kovarianzmatrix $\boldsymbol{\Sigma}_{\mathbf{c}} \in \mathbb{R}^{8 \times 8}$ folgen. Durch das Ziehen neuer Parametervektoren $\tilde{\mathbf{c}}_j$ aus $\mathcal{N}(\boldsymbol{\mu}_{\mathbf{c}}, \boldsymbol{\Sigma}_{\mathbf{c}})$ können neue Atemzyklen generiert werden. Auch wenn die Normalverteilungsannahme u. a. durch die Vermischung von Form- und Zeitparametern mutmaßlich nur sehr approximativ zutrifft, eröffnet sie nachfolgend eine sehr einfache Möglichkeit zur Simulation einer Folge von Atemzyklen mit kontinuierlichem Übergang.

Um Kontinuität zwischen zwei aufeinanderfolgenden simulierten Atemzyklen zu garantieren, wird, beginnend mit dem zweiten Atemzyklus einer zu simulierenden Folge $G = (\tilde{\mathbf{c}}_1, \dots, \tilde{\mathbf{c}}_j, \tilde{\mathbf{c}}_{j+1}, \dots, \tilde{\mathbf{c}}_{n_{cyc}})$ von n_{cyc} Atemzyklen, eine bedingte multivariate Normalverteilung $\mathcal{N}(\boldsymbol{\mu}_{\mathbf{b}|\mathbf{a}}, \boldsymbol{\Sigma}_{\mathbf{b}|\mathbf{a}})$ für das Ziehen der Zyklusparameter benutzt, da die Werte der ersten beiden Kontrollpunkte eines gezogenen Parametervektors/Atemzyklus $\tilde{\mathbf{c}}_{j+1}$ mit den letzten beiden Kontrollpunkten des vorherigen Zyklus $\tilde{\mathbf{c}}_j$ übereinstimmen müssen, um eine C^1 -Stetigkeit⁶ am Übergang zwischen den Zyklen zu erreichen. Hierfür wird der 8-dimensionale Zufallsvektor der Zyklusparameter $\mathbf{c} = [p_1 = p_2, \dots, p_{n_{cp}-1} = p_{n_{cp}}, t]^T$ zuerst in die Teilvektoren $\mathbf{a} = [p_1 = p_2]^T$ und $\mathbf{b} = [p_3, \dots, p_{n_{cp}-1} = p_{n_{cp}}, t]^T$ zerlegt. Die bedingte Normalverteilung beschreibt dann die Verteilung der übrigen Zyklusparameter bei Kenntnis des Wertes der ersten beiden Kontrollpunkte. Unter Berücksichtigung der Zerlegung des Zufallsvektors, lassen sich die Parameter der multivariaten Normalverteilung $\mathcal{N}(\boldsymbol{\mu}_{\mathbf{c}}, \boldsymbol{\Sigma}_{\mathbf{c}})$ als

$$\boldsymbol{\mu}_{\mathbf{c}} = \begin{pmatrix} \boldsymbol{\mu}_{\mathbf{a}} \\ \boldsymbol{\mu}_{\mathbf{b}} \end{pmatrix} \text{ und } \boldsymbol{\Sigma}_{\mathbf{c}} = \begin{pmatrix} \boldsymbol{\Sigma}_{\mathbf{aa}} & \boldsymbol{\Sigma}_{\mathbf{ab}} \\ \boldsymbol{\Sigma}_{\mathbf{ba}} & \boldsymbol{\Sigma}_{\mathbf{bb}} \end{pmatrix} \quad (4.4)$$

schreiben, aus der dann die Parameter der bedingten Normalverteilung $\mathcal{N}(\boldsymbol{\mu}_{\mathbf{b}|\mathbf{a}}, \boldsymbol{\Sigma}_{\mathbf{b}|\mathbf{a}})$ abgeleitet werden können (siehe [Bishop, 2006]). Sollen nun für einen konkreten Atemzyklus $\tilde{\mathbf{c}}_{j+1}$ die restlichen sieben Parameter $\{p_{j+1,3}, \dots, p_{j+1,n_{cp}-1}, p_{j+1,n_{cp}}, t_{j+1}\}$ unter

⁶ C^1 -Stetigkeit ist streng genommen nur bei gleicher zeitlicher Länge der aufeinanderfolgenden Atemzyklen gegeben, da sich sonst die Knotenvektoren unterscheiden.

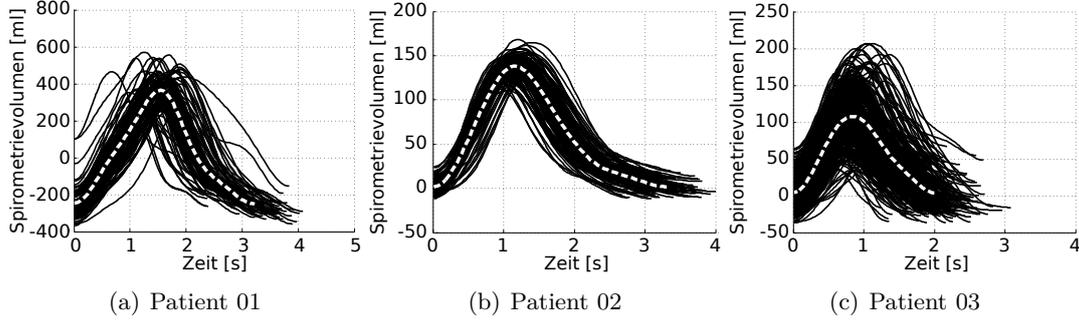


Abb. 4.4: Darstellung der mittleren Atemzyklen (weiße gestrichelte Kurve; mittlere Form und Länge) von drei verschiedenen Patienten. Zusätzlich sind jeweils die zur Berechnung der mittleren Zyklen genutzten approximierten Originalzyklen des Spirometersignals zu sehen. Die drei Patienten zeigen zudem die große Intra- und Interpatientenvariabilität der Atmung. Während bei Patient 01 (a) Ein- und Ausatemungsdauer im Mittel relativ gleich lang sind, zeigt sich bei Patient 02 (b) eine deutlich kürzere mittlere Einatemungsphase im Vergleich zur Ausatemungsphase. Bei Patient 03 (c) lässt sich eine sehr große Variabilität bezüglich zeitlicher Länge der einzelnen Atemzyklen und der Einatemungstiefe erkennen.

Kenntnis von $\mathbf{a} = [p_{j+1,1} = p_{j+1,2} = p_{j,n_{cp}}]^T$ bestimmt werden, ergeben sich die Parameter der für das Ziehen genutzten bedingten Normalverteilung als [Bishop, 2006]:

$$\boldsymbol{\mu}_{\mathbf{b}|\mathbf{a}} = \boldsymbol{\mu}_{\mathbf{b}} + \boldsymbol{\Sigma}_{\mathbf{ba}}\boldsymbol{\Sigma}_{\mathbf{aa}}^{-1}(\mathbf{a} - \boldsymbol{\mu}_{\mathbf{a}}) \text{ und } \boldsymbol{\Sigma}_{\mathbf{b}|\mathbf{a}} = \boldsymbol{\Sigma}_{\mathbf{bb}} - \boldsymbol{\Sigma}_{\mathbf{ba}}\boldsymbol{\Sigma}_{\mathbf{aa}}^{-1}\boldsymbol{\Sigma}_{\mathbf{ab}}. \quad (4.5)$$

Unter der Annahme, dass mit der skizzierten Methodik eine Folge von n_{cyc} Atemzyklen $G = (\tilde{\mathbf{c}}_1, \dots, \tilde{\mathbf{c}}_j, \tilde{\mathbf{c}}_{j+1}, \dots, \tilde{\mathbf{c}}_{n_{cyc}})$ generiert worden ist, können die einzelnen Parametervektoren in Verbindung mit Glg. 4.1 genutzt werden, um ein zeitkontinuierliches, simuliertes 1D-Surrogatsignal $g(t) : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ zu erzeugen. Ist zusätzlich mithilfe des Originalsurrogatsignals und zugehöriger 4D-Bilddaten ein patientenspezifisches diffeomorphes Korrespondenzmodell⁷

$$M(\zeta_t) : \mathbb{R}^{n_{sur}} \rightarrow \text{Diff}(\Omega) \quad (4.6)$$

erzeugt worden (siehe Kapitel 2.3), so lässt sich mit $\hat{\varphi}_t = M(\zeta_t)$ und $\zeta_t = [g(t), \frac{d}{dt}g(t)]^T$ die diffeomorphe nichtlineare Transformation $\hat{\varphi}_t : \Omega \rightarrow \Omega$ generieren, die die interne Atembewegung zum Zeitpunkt t des simulierten Surrogatsignals $g(t)$ beschreibt. Hierbei wurde $n_{sur} = 2$ angenommen, wobei die zweite Dimension der Surrogatsignalmessung ζ_t die zeitliche Ableitung des 1D-Signals ist, um die Hysterese modellieren zu können. Durch die B-Spline-basierte Beschreibung des simulierten Signals kann die zeitliche Ableitung analytisch ermittelt werden (siehe [Piegl u. a., 2005]).

⁷Die hier genutzte Schreibweise $M(\zeta_t)$ kapselt das eigentliche lineare Modell und die Umwandschritte, die nötig sind, um vom Surrogatsignalwert zur zugehörigen nichtlinearen Transformation zu gelangen.

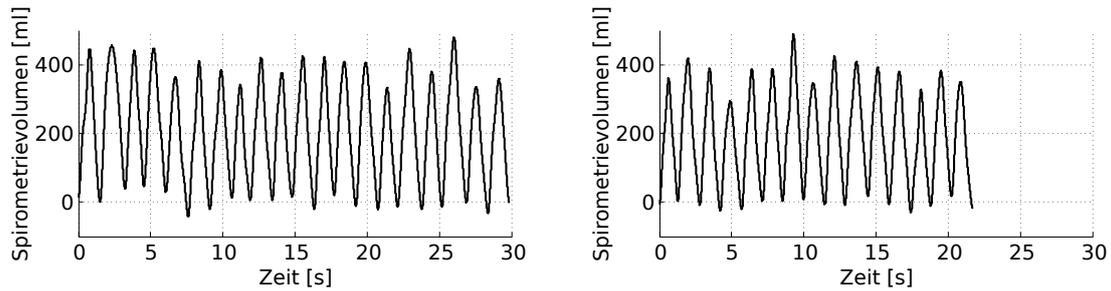


Abb. 4.5: Zwei Beispiele für simulierte 1D-Spirometersignale unterschiedlicher Länge für einen Patienten.

Die Kombination der normalverteilungsbasierten Simulation einer Folge von Atemzyklen (siehe Glg. 4.5 und Glg. 4.4) mit der B-Spline-basierten zeitkontinuierlichen Beschreibung des simulierten Surrogatsignals (siehe Glg. 4.1) und eines patientenspezifischen diffeomorphen Korrespondenzmodells (siehe Glg. 4.6) ermöglicht mithilfe der vorgestellten Methodik die Simulation realistischer und variabler Atembewegungssequenzen beliebiger zeitlicher Länge. Aufgrund der zeitkontinuierlichen Beschreibung des simulierten Signals kann zudem ohne Mehraufwand ein auf die konkrete Simulationsanwendung abgestimmtes zeitliches Diskretisierungsschema genutzt werden.

4.1.4 Evaluation der Simulationsmethode

Bei der Evaluation der präsentierten Simulationsmethode müssen prinzipiell drei verschiedene Methodenteile evaluiert werden: (1) Die B-Spline-basierte Zyklusapproximation. Wie gut ist Approximation? (2) Das generative Modell zur Erzeugung von Surrogatsignalen. Wie realistisch sind die simulierten Signale? (3) Das patientenspezifische Korrespondenzmodell im Zusammenspiel mit simulierten Signalen. Wie realistisch ist die simulierte interne Atembewegung? Nachfolgend werden die zur Evaluation der einzelnen Teile durchgeführten Experimente beschrieben, Ergebnisse präsentiert und etwaige Schwierigkeiten diskutiert. Hierfür wurden für die Rekonstruktion von 4D-CT-Bilddaten aufgenommene 1D-Spirometersignale von insgesamt 19 Lungentumorpatienten eingesetzt. Diese Daten stammen aus einer Kooperation des Instituts für Medizinische Informatik der Universität zu Lübeck mit der Washington University in St. Louis, USA, wobei Daten von zwölf dieser Patienten auch in Kapitel 3 zum Einsatz kommen (WUSTL-Daten; siehe Kapitel 3.1.1). Die verfügbaren Signale sind jeweils 2–14 Minuten lang.

B-Spline-basierte Signalapproximation Die Güte der Signalapproximation ist im Wesentlichen von der Wahl der Parameter für die B-Spline-Kurven abhängig. Nach Festlegung des generellen Modells (nichtperiodisch, nichtuniform, offen, Grad $d = 3$) verbleibt hier als Parameter die Anzahl von Kontrollpunkten n_{cp} , durch die ein Atem-

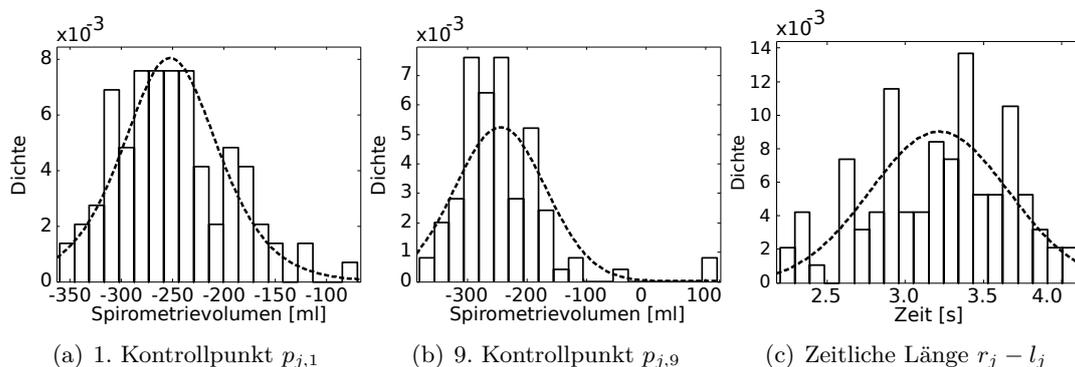


Abb. 4.6: Normalisierte Histogramme für die Werte des 1. Kontrollpunktes, des 9. Kontrollpunktes und der zeitlichen Länge aller approximierten Atemzyklen des Spirometersignals von Patient 01 mit jeweils zusätzlich geschätzter univariater Normalverteilungskurve (gestrichelte Kurve). Die Parameter dieses Patienten sind nicht multivariat normalverteilt. Es ist bei allen drei Parametern schon eine Asymmetrie der univariaten Verteilung sichtbar.

zyklus repräsentiert werden soll. Da durch die angestrebte (approximative) C^1 -Kontinuität an Zyklusübergängen jeweils die ersten und letzten beiden Kontrollpunkte den gleichen Wert haben müssen, erscheint die Wahl von $n_{cp} \leq 4$ wenig sinnvoll. Weiterhin ist klar, dass die Approximation umso besser wird, je näher der Wert von n_{cp} der Anzahl von Abtastpunkten des Atemzyklus kommt. Hier gilt es einen entsprechenden Kompromiss aus niedriger Parameteranzahl und hoher Approximationsgenauigkeit zu finden. In [Quirk u. a., 2012] wurde zu diesem Zweck das Akaike Information Criterion [Bishop, 2006] genutzt. Tests mit diesem Maß wurden auch für diese Arbeit durchgeführt, wobei sich hierbei insbesondere die sinnvolle quantitative Beschreibung der Approximationsgenauigkeit als problematisch erwiesen hat. Beispielsweise sollen große Abweichungen vom globalen Maximum des Zyklus unbedingt vermieden werden, um die Einatmungstiefe und die Dauer von Ein- und Ausatmung korrekt zu repräsentieren, wohingegen mutmaßlich durch technische Effekte verursachte lokale Extrema geglättet werden können (siehe Abb. 4.3 als Beispiel). Aus diesem Grund wurde der in dieser Arbeit genutzte Wert $n_{cp} = 9$ durch visuelle Beurteilung der Approximationsgüte verschiedener Atemzyklen aller 19 Patienten ermittelt. Für die Zukunft ist eine quantitativ untermauerte Wahl des Parameters allerdings wünschenswert.

Simulation von Surrogatsignalen Beim Versuch der Beurteilung des Realismus der simulierten Surrogatsignale ergibt sich analog zur Bestimmung der Approximationsgüte das Problem der sinnvollen Quantifizierung von Realismus. Greift man hierzu auf die Standardmethoden zur Evaluation von statistischen Formmodellen (u. a. Spezifität; siehe [Davies, 2002]) zurück, kann beispielsweise die Spezifität des generativen

Modells durch Erzeugung von neuen Atemzyklen und anschließender Bestimmung des Abstandes zum ähnlichsten Originalzyklus ermittelt werden. Dieses Vorgehen ist allerdings mit zwei Problemen behaftet: (1) Die sinnvolle Quantifizierung des Abstandes/der Ähnlichkeit von zwei Atemzyklen ist schwierig (siehe Approximationsgüte). (2) Die isolierte Betrachtung einzelner Atemzyklen ignoriert den Fakt, dass durch die Methode zumeist längere Signalsequenzen erzeugt werden sollen. Diese Probleme haben dazu geführt, dass im Rahmen dieser Arbeit im ersten Schritt eine visuelle Beurteilung simulierter Surrogatsignale aller 19 Patienten erfolgt ist (siehe Beispiele für simulierte Signale in Abb. 4.5). Hierbei erfolgte hauptsächlich eine Beurteilung der Form der Atemzyklen, deren zeitlicher Länge und der Übergänge zwischen den Zyklen. Bei dieser Beurteilung wurden alle simulierten Signale als realistisch erachtet.

Neben der Untersuchung der simulierten Signale bietet es sich zudem an die dem generativen Modell zugrunde liegende Annahme einer multivariaten Normalverteilung der acht Zyklusparameter zu untersuchen. Hierfür wurden in der Literatur viele verschiedene Tests vorgeschlagen (siehe z. B. Übersicht und Diskussion im Kontext statistischer Formmodellierung in [Gollmer, 2015]), wobei es keinen allgemeingültigen und allgemein anerkannten Test auf multivariate Normalverteilung gibt und deshalb in der Regel eine Kombination aus mehreren Tests eingesetzt wird, um verschiedene Eigenschaften der Verteilung der Testdaten prüfen. In dieser Arbeit werden der Schiefe- und Kurtosis-Test von Mardia [1970] und der Henze-Zirkler-Test [Henze u. a., 1990] eingesetzt. Letzterer basiert auf der charakteristischen Funktion einer multivariaten Normalverteilung. Auf eine nähere mathematische Definition dieser Verfahren wird zugunsten der Übersichtlichkeit der Darstellung in diesem Kapitel verzichtet und der interessierte Leser sei auf die genannten Originalarbeiten und [Gollmer, 2015] verwiesen.

Der Test auf potenziell vorhandene Normalverteilung der Parametervektoren wird für jeden der 19 Patienten einzeln mit allen drei Tests durchgeführt. Für alle Tests wird ein Signifikanzniveau von $\alpha = 0.05$ gewählt und die Parametervektoren eines Patienten werden nur dann als normalverteilt aufgefasst, wenn alle drei Tests positiv auf eine Normalverteilung ausgewertet werden. Dies trifft allerdings nur auf einen einzigen der 19 untersuchten Patienten zu. Die Parametervektoren der 18 anderen Patienten sind dementsprechend nicht multivariat normalverteilt (siehe auch Abb. 4.6). Dieses Ergebnis zeigt, dass die Normalverteilungsannahme in der Regel theoretisch nicht gerechtfertigt ist. Dies ist im Bereich der statistischen Modellierung anatomischer Formen ebenfalls oftmals der Fall [Gollmer, 2015]. Nichtsdestotrotz haben sich dort Modelle, die nicht auf einer Normalverteilungsannahme aufbauen, bisher u. a. wegen mangelnder Effizienz nicht durchsetzen können. Diese Argumentation lässt sich auch auf die hier vorgestellte Methode übertragen. Die Normalverteilungsannahme ist zwar theoretisch nur in Ausnahmefällen wirklich korrekt, ermöglicht aber beispielsweise eine einfache Modellierung. Zukünftig erscheinen aber Untersuchungen bezüglich anderer Modellierungsmöglichkeiten wünschenswert.

Simulierte interne Bewegung Auf eine explizite Evaluation des Realismus der simulierten internen Atembewegung wird im Rahmen dieser Arbeit verzichtet. Dieses Vorgehen erfolgt ausgehend von der Annahme, dass die simulierten Surrogatsignale bei visueller Bewertung realistisch sind, auch wenn die Normalverteilungsannahme nicht zutrifft, und der Tatsache, dass die patientenspezifische Korrespondenzmodellierung ausgiebig und erfolgreich in Kapitel 3 evaluiert wurde.

4.1.5 Zusammenfassung

In diesem Teil des vorliegenden Kapitels wurde ein Verfahren zur modellbasierten Simulation realistischer und variabler Atembewegung vorgestellt. Zentrales Element dieses Ansatzes ist die hier neu vorgestellte Methode zur B-Spline-basierten Erzeugung von variablen Surrogatsignalen beliebiger zeitlicher Länge mit typischen patientenspezifischen Variationen. Hierdurch wird in Verbindung mit einem Korrespondenzmodell die Simulation variabler interner Atembewegung beliebiger zeitlicher Länge möglich.

Die Surrogatsignalsimulation basiert auf einer konsistenten B-Spline-basierten Approximation der einzelnen Atemzyklen eines real vorliegenden 1D-Surrogatsignals. Die Atemzyklen des realen Signals werden jeweils durch einen spezifischen Vektor von B-Spline-Koeffizienten parametrisiert, welche genutzt werden, um die Parameter einer multivariaten Normalverteilung zu schätzen. Durch die Ziehung neuer Koeffizienten können auf diese Weise realistische, variable Surrogatsignale simuliert werden.

Als Einschränkung des Verfahrens ist eine gewisse Abhängigkeit vom realen Surrogatsignal zu nennen, da Effekte, die in diesem Signal nicht vorhanden sind oder vor der B-Spline-Approximation entfernt werden, (sehr wahrscheinlich) auch nicht in den simulierten Signalen auftreten werden. Weiterhin zeigt die Evaluation, dass die Annahme einer multivariaten Normalverteilung der Zyklusparameter streng genommen nicht zutreffend ist, auch wenn die auf diese Weise simulierten Signale realistisch aussehen. Diese Feststellung sollte deshalb in der Zukunft weitere Untersuchungen hinsichtlich alternativer Verteilungen motivieren, um die Modellierung weiter zu verbessern.

4.2 Anwendung 1: 4D-Dosissimulation in der Strahlentherapieplanung

In der Einleitung zu diesem Kapitel wurde die 4D-Dosissimulation als mögliches Einsatzgebiet für simulierte Atembewegung genannt, um die intrafraktionelle Variabilität der Atmung bei Bestrahlungsplanung und Planverifikation stärker berücksichtigen zu können. Dies ist u. a. vor dem Hintergrund des Trends hin zu modernen stereotaktischen und/oder intensitätsmodulierten Bestrahlungsverfahren von Bedeutung.

Stereotaktische Bestrahlung (engl. Stereotactic Body Radiation Therapy; SBRT) bezeichnet hierbei keine konkrete Technik, sondern ist ein Sammelbegriff für Thera-

pieansätze, bei denen die Bestrahlung im Vergleich zu konventionellen Verfahren in wenigen Fraktionen mit sehr hoher lokaler Einzeldosis möglich ist [Sahgal u. a., 2012]. Diese Therapieform stellt hohe Anforderungen an die Präzision des kompletten Therapieablaufs von der Planung über die Verifikation bis hin zur eigentlichen Bestrahlung. Technisch kommen bei der Bestrahlung deshalb häufig intensitätsmodulierte Verfahren (engl. Intensity Modulated Radiation Therapy; IMRT) zum Einsatz, bei denen die Intensität der Bestrahlung präzise an die Tumorform angepasst werden kann. Dies erfolgt beispielsweise durch Modulation des Therapiestrahls mithilfe von Multileaf-Kollimatoren, bei denen die Lamellen so verschoben werden, dass sich ein zugehöriges inhomogenes Bestrahlungsfeld ergibt. Durch die Nutzung einer Vielzahl verschiedener inhomogener Bestrahlungsfelder, die aus unterschiedlichen Winkeln mit entsprechenden Lamellenkonfigurationen für eine jeweils exakt berechnete Dauer bzw. mit einer vorbestimmten Dosisrate erzeugt werden, wird so über den zeitlichen Verlauf einer Fraktion durch Überlagerung der Einzeldosen die gewünschte homogene Dosisverteilung im Zielgebiet erreicht [Thilmann u. a., 2013].

Da die vielen Einzelfelder bei IMRT-basierter Bestrahlung jeweils nur für sehr kurze Zeit erzeugt werden, kann es durch die Atembewegung zu einer ungeplanten Bestrahlung eines falschen Gebietes kommen, sodass im Endeffekt eine zu geringe Dosis im Tumor appliziert wird, aber gleichzeitig angrenzende Risikostrukturen übermäßig stark bestrahlt werden. Diese (negative) Wechselwirkung von intrafraktioneller Atembewegung und den inhomogenen Einzelfeldern wird Interplay-Effekt genannt [Bortfeld u. a., 2002] und stellt ein zentrales Risiko der IMRT-basierten Bestrahlung thorakaler und abdominaler Tumoren dar.

Um den Interplay-Effekt in Simulationsstudien patientenspezifisch zu untersuchen, wurde in [Werner u. a., 2012c; Werner, 2012] ein modellbasiertes 4D-Dosissimulationsverfahren vorgeschlagen. Bei diesem Verfahren wird ausgehend von einem 4D-CT-Bilddatensatz ein patientenspezifisches diffeomorphes Korrespondenzmodell erzeugt, welches anschließend genutzt wird, um mithilfe eines aufgenommenen Surrogatsignals die Atembewegung während einer Bestrahlungsfraktion zu simulieren. Die modellbasiert erzeugten diffeomorphen Transformationen werden dann zur 4D-Dosisakkumulation eingesetzt, wodurch sich die Interplay-Effekte der durch das Surrogatsignal repräsentierten Atembewegung untersuchen lassen. Dieses Verfahren stellt damit eine Verbesserung gegenüber 4D-Dosissimulationsverfahren dar, welche ausschließlich auf der in einzelnen 4D-CT-Bilddatensätzen repräsentierten Bewegung arbeiten, da es auch die intrafraktionelle Variabilität der Atembewegung berücksichtigt.

Eine Limitation des Verfahrens aus [Werner u. a., 2012c; Werner, 2012] stellt die explizite Beschränkung der Untersuchung des Interplay-Effekts auf die durch das aufgenommene Surrogatsignal repräsentierte Bewegungssequenz dar. Sollen weitere Fraktionen oder Szenarien simuliert/untersucht werden, müssen entsprechende Signalaufnahmen vorliegen. An dieser Stelle kann durch Integration des in Kapitel 4.1 vorgestellten Bewegungssimulationsverfahrens eine Verbesserung der 4D-Dosissimulation

erreicht werden, da dann ausgehend von einer aufgenommenen Surrogatsignalsequenz beliebig viele unterschiedliche Bestrahlungsfractionen simuliert werden können. Diese Integration wird in Kapitel 4.2.1 beschrieben, worauf aufbauend in Kapitel gezeigt wird, wie dieses erweiterte Verfahren zur statistischen Untersuchung atmungsbedingter Effekte auf die Dosisverteilung eingesetzt werden kann. In Kapitel 4.2.3 und Kapitel 4.2.4 erfolgt dann eine initiale experimentelle Erprobung des erweiterten 4D-Dosissimulationsverfahren anhand von IMRT-Bestrahlungsplänen von Lungentumorpatienten.

4.2.1 4D-Dosissimulation durch modellbasierte Dosisakkumulation

Die Basis für die hier beschriebene Methodik zur modellbasierten Dosisakkumulation mit simulierten Surrogatsignalen bildet das in [Werner u. a., 2012a,c; Werner, 2012] ausführlich beschriebene Dosisakkumulationsschema, welches hier nur zusammengefasst dargestellt wird. Grundvoraussetzung für die angestrebte modellbasierte Dosisakkumulation ist das Vorliegen eines patientenspezifischen (diffeomorphen) Korrespondenzmodells M (siehe Glg. 4.6), eines aufgenommenen 1D-Surrogatsignals $s(t)$ des Patienten und eines 3D-Bestrahlungsplans mit n_{fields} Bestrahlungsfeldern. Hierbei wird davon ausgegangen, dass der Bestrahlungsplan auf Basis des CT-Bildes berechnet wurde, welches auch als Referenzphase für das Training des Korrespondenzmodells dient. Der Bestrahlungsplan ist dann konkret gegeben durch die Dosisverteilungen $(D_k : \Omega \rightarrow \mathbb{R}_{\geq 0})_{k=1}^{n_{fields}}$ der einzelnen Bestrahlungsfelder berechnet auf dem Referenzbild. Weiterhin seien die Anfangs- und Endzeiten ($t_{k,0}$ und $t_{k,end}$) der Bestrahlung mit Einzelfeld $k \in \{1, \dots, n_{fields}\}$ während der Fraktion mit Länge $\Delta t_{fract} = t_{n_{fields},end} - t_{1,0}$ bekannt.

Wird ausgehend vom Originalsignal $s(t)$ mithilfe des in Kapitel 4.1 beschriebenen Verfahrens ein simuliertes 1D-Surrogatsignal $g(t)$ mit Parametervektor \mathbf{g} erzeugt, dessen zeitliche Länge der Fraktionsdauer Δt_{fract} entspricht und das mit Abtastrate $1/\Delta t_g$ diskretisiert wird, lässt sich die modellbasierte Dosisakkumulation nach [Werner u. a., 2012c; Werner, 2012] vereinfacht als

$$D^{fract}(\mathbf{g}, \mathbf{x}) = D_g^{fract}(\mathbf{x}) = \sum_{k=1}^{n_{fields}} \frac{\Delta t_g}{t_{k,end} - t_{k,0}} \sum_l D_k(\hat{\varphi}_l(\mathbf{x})) \quad (4.7)$$

mit

$$\hat{\varphi}_l = M(\zeta_{t_l}) \text{ und } \zeta_{t_l} = [g(t_l), \frac{d}{dt}g(t_l)]^T \quad (4.8)$$

schreiben. $D^{fract}(\mathbf{g}, \mathbf{x})$ ist dann die Dosis, die der Punkt $\mathbf{x} \in \Omega$ des Referenzbildes über die mit Parametervektor \mathbf{g} simulierte Fraktion erhalten hat (4D-Dosissimulation). Hierbei bezeichnet l die einzelnen Samplingpunkte des simulierten Surrogatsignals mit zugehöriger Zeit t_l , zu der das jeweilige Einzelfeld aktiv ist.

4.2.2 Statistische Untersuchung atmungsbedingter Effekte auf die Dosisverteilung

Mithilfe des 4D-Dosissimulationsverfahrens in Kapitel 4.2.1 lassen sich nun einzelne Fraktionen simulieren und die entstehenden Dosisverteilungen untersuchen. Für eine Verifikation eines Bestrahlungsplans ist aber weniger die durch eine konkrete Atembewegungssequenz in einer Fraktion erzeugte Dosisverteilung von Interesse, als Kenntnisse über die zu erwartenden Variationen der Dosisverteilung bei unterschiedlichen, aber realistischen Bewegungssequenzen. Aus diesem Grund haben Geneser u. a. [2011] die modellbasierte statistische Analyse der Auswirkungen verschiedener Bewegungssequenzen auf die Dosisverteilung vorgeschlagen. Konkret werden hierzu der Erwartungswert und die Standardabweichung der in einem Punkt akkumulierten Dosis errechnet. Anhand dieser Werte wird es dann möglich zu entscheiden, ob der Bestrahlungsplan bei der zu erwartenden patientenspezifischen Atmungsvariabilität geeignet ist oder nicht. Konkret könnte bei einer solchen Entscheidung z. B. darauf geachtet werden, dass der Erwartungswert der Dosis in jedem Punkt des Zielvolumens bei geringer Standardabweichung der verschriebenen Dosis entspricht und zugleich die umliegenden Risikoorgane keine unzulässig hohe Dosis erhalten.

Für das hier präsentierte Modellierungs- und Dosisakkumulationsverfahren lassen sich Erwartungswert und Standardabweichung aufbauend auf Glg. 4.7 berechnen. Unter der Annahme, dass \mathbf{G} ein Zufallsvektor von Surrogatsignalparametern mit zugehöriger Dichte $f_{\mathbf{G}}(\mathbf{g})$ ist, lässt sich der Erwartungswert der Dosis mittels Monte-Carlo-Approximation des zugrunde liegenden Integrals vereinfacht als

$$E[D^{\text{fract}}(\mathbf{G}, \mathbf{x})] \approx \mu_{D^{\text{fract}}}(\mathbf{x}) = \frac{1}{n_{\text{sim}}} \sum_{i=1}^{n_{\text{sim}}} D^{\text{fract}}(\mathbf{g}_i, \mathbf{x}) \quad (4.9)$$

schreiben und die Approximation der Standardabweichung ergibt sich dann durch

$$\sqrt{\text{Var}\left(D^{\text{fract}}(\mathbf{G}, \mathbf{x})\right)} \approx \sigma_{D^{\text{fract}}}(\mathbf{x}) = \sqrt{\frac{1}{n_{\text{sim}} - 1} \sum_{i=1}^{n_{\text{sim}}} \left(D^{\text{fract}}(\mathbf{g}_i, \mathbf{x}) - \mu_{D^{\text{fract}}}(\mathbf{x})\right)^2}. \quad (4.10)$$

$\mu_{D^{\text{fract}}}(\mathbf{x})$ und $\sigma_{D^{\text{fract}}}(\mathbf{x})$ geben dann den Mittelwert und die Standardabweichung der in Punkt $\mathbf{x} \in \Omega$ des Referenz-CT pro Fraktion unter Berücksichtigung der Atembewegung akkumulierten Dosis an. Hierbei wird jede Fraktion durch den jeweiligen Parametervektor \mathbf{g}_i des simulierten Surrogatsignals parametrisiert. Diese Parametervektoren werden aus der Verteilung mit Dichte $f_{\mathbf{G}}$ gezogen, wobei dies praktisch durch Nutzung der einzelnen, zykluspezifischen Normalverteilungen geschieht (siehe Kapitel 4.1.3 für deren Definition).

Tabelle 4.1: Ergebnisse der Experimente zur Evaluation der 4D-Dosissimulation mit vier IMRT-Bestrahlungsplänen. Die Auswertung erfolgt CTV-zentriert durch Analyse der Dosisabdeckung im CTV gemäß: (1) Geplanter Dosisverteilung, (2) Dosisakkumulation basierend auf einem patientenspezifischen 4D-CT-Bilddatensatz und (3) modellbasierter Dosisakkumulation über 100 simulierte Fraktionen (Angabe: Mittlere Dosis und Bereich der Werte der einzelnen Fraktionen). Die Angaben beziehen sich hierbei auf die pro Fraktion zu applizierende Dosis von 2 Gy. $D_{99}(\text{CTV})$: Minimale Dosis, die für 99% der Voxel des CTV erreicht wurde; $D_{\min/\max}(\text{CTV})$: minimale bzw. maximale Dosis, die für die CTV-Voxel erreicht wurde. Tabelle/Ergebnisse aus [Wilms u. a., 2014b].

	$D_{99}(\text{CTV})$ [%]	$D_{\min}(\text{CTV})$ [%]	$D_{\max}(\text{CTV})$ [%]
<i>Geplante Dosisverteilung</i>			
Patient 01	96.3	95.5	110.2
Patient 02	97.0	95.4	109.9
Patient 03	97.5	96.1	107.3
Patient 04	99.9	98.1	109.7
<i>Dosisakkumulation basierend auf 4D-CT-Bilddatensatz</i>			
Patient 01	97.1	96.4	109.0
Patient 02	97.2	94.1	108.2
Patient 03	97.3	95.7	106.5
Patient 04	95.4	94.3	108.6
<i>Mittlere akkumulierte Dosis bei $n_{sim} = 100$ simulierten Fraktionen</i>			
Patient 01	97.1 (96.3 – 97.5)	95.8 (95.6 – 95.8)	109.3 (109.2 – 110.9)
Patient 02	96.2 (94.1 – 97.3)	92.9 (91.4 – 94.3)	107.2 (106.7 – 112.1)
Patient 03	96.9 (95.6 – 97.8)	95.4 (94.1 – 96.7)	106.4 (105.0 – 107.8)
Patient 04	89.6 (85.4 – 94.3)	88.1 (80.2 – 92.5)	106.0 (105.0 – 108.7)

4.2.3 Experimente mit IMRT-Bestrahlungsplänen

Das vorgestellte 4D-Dosissimulationsverfahren bzw. die damit mögliche statistische Analyse der Dosisverteilung wird im Folgenden anhand von vier IMRT-Bestrahlungsplänen (Quelle: [Werner, 2012]) verschiedener Lungentumorpatienten evaluiert. Diese vier Patienten sind eine Untermenge der WUSTL-Daten (siehe Kapitel 3.1.1 bzw. Kapitel 4.1.4), für die neben thorakalem 4D-CT-Bilddatensatz und Spirometersignal auch jeweils ein 3D-IMRT-Bestrahlungsplan (Step-&Shoot-Technik) zu Verfügung steht. Step-&Shoot-Technik bedeutet hierbei, dass der Therapiestrahl bei Bewegung der Lamellen und/oder Gantry des Linearbeschleunigers nicht eingeschaltet ist [Thilmann u. a., 2013].

Eine genaue Beschreibung der eingesetzten Bestrahlungspläne findet sich in [Werner, 2012]⁸. Die folgende Beschreibung beschränkt sich deshalb auf die für das Verständnis der Experimente wichtigsten Aspekte. Die Bestrahlungspläne sind generell jeweils auf eine Therapie mit 25 Fraktionen und einer Dosis pro Fraktion von 2 Gy ausgelegt (= 50 Gy Gesamtdosis) und wurden auf Basis der Phase mittlerer Einatmung (MI) des 4D-CT-Bilddatensatzes erstellt. Der Sicherheitssaum um den Lungentumor wurde hierbei so dimensioniert, dass das PTV (Planning Target Volume), welches zur eigentlichen Planung genutzt wird, in alle Richtungen isotrop 10 mm größer als das CTV (Clinical Target Volume) ist, welches den sichtbaren Tumor und die mutmaßlich mit Tumorgewebe infiltrierte Umgebung (≈ 5 mm isotroper Saum) enthält⁹. Die Gesamtbestrahlungszeit (= Zeit, in der der Therapiestrahl aktiv ist) pro Fraktion beträgt patientenspezifisch jeweils zwischen 45 s und 75 s, wobei die Pausen zwischen den Einzelfeldern (patientenabhängig 11–24 Felder bei drei Einstrahlwinkeln) bei der Step-&Shoot-Bestrahlung analog zu [Werner, 2012] mit jeweils 1 s veranschlagt werden.

Die Berechnung der mittleren Dosis $\mu_{D^{\text{fract}}}$ pro Fraktion und der zugehörigen Standardabweichung $\sigma_{D^{\text{fract}}}$ erfolgt auf Basis von $n_{\text{sim}} = 100$ simulierten Fraktionen, für die jeweils ein Surrogatsignal passender Länge erzeugt und mit einem Abtastintervall von $\Delta = 0.1$ s diskretisiert wird. Die Beschränkung auf nur 100 simulierte Fraktionen erfolgt hierbei aufgrund des erheblichen Rechenaufwands: Auf dem genutzten Testsystem mit Intel W3520 2.67 GHz-Prozessor und 24 GB RAM dauert die Simulation einer Fraktion inkl. Dosisakkumulation durchschnittlich ca. 30 Minuten. Aus demselben Grund wird auch abweichend von Glg. 4.8 bzw. Glg. 4.6 für die Bewegungssimulation kein diffeomorphes Korrespondenzmodell nach Kapitel 2.3 erzeugt, da die in Kapitel 3 erwähnte GPU-Implementierung der Scaling-und-Squaring-Methode für diese Evaluation noch nicht zur Verfügung stand. Deshalb wird hier eine direkte Regression zwischen Surrogatdaten und Verschiebungsfeldern durchgeführt (siehe auch Kapitel 3.1.3.4), um bei

⁸Patient 01 in dieser Arbeit entspricht WashU 01 in [Werner, 2012]; Patient 02 → WashU 02; Patient 03 → WashU 04; Patient 04 → WashU 36

⁹Genaue klinische Definitionen der verschiedenen bei der Bestrahlung und Bestrahlungsplanung genutzten Volumina sind in [IRCU, 2010] zu finden.

ähnlicher Genauigkeit das zeitaufwendige Scaling-und-Squaring zu vermeiden (siehe Kapitel 3.2). Zudem wird die eigentliche Bewegungsschätzung auf die Lungenregion eingeschränkt, um Gleitbewegungen bei der Registrierung zur Korrespondenzmodellgenerierung nicht berücksichtigen zu müssen.

4.2.4 Ergebnisse

Die Ergebnisse der Experimente sind in Tab. 4.1, Abb. 4.7 und Abb. 4.8 zusammengefasst. Die quantitative Auswertung der (simulierten) Dosisverteilungen in Tab. 4.1 und Abb. 4.7 erfolgt hierbei grundsätzlich CTV-zentriert durch Analyse der Dosisabdeckung im CTV in Abhängigkeit von der verschriebenen Dosis. Im CTV sollte nach ICRU-Empfehlungen [ICRU, 1999] eine Dosis erreicht werden, die zwischen 95% und 107% der tatsächlich pro Fraktion verschriebenen Einzeldosis (hier: 2 Gy) liegt, um Unter- oder Überdosierungen zu verhindern. Die folgende Analyse der Ergebnisse wird sich hierbei allerdings auf die für das Therapieziel schwerwiegendere Unterdosierung fokussieren. Zusätzlich zu den Ergebnissen der modellbasierten Dosisakkumulation über 100 simulierte Fraktionen finden sich in Tab. 4.1, Abb. 4.7 und Abb. 4.8 zu Vergleichszwecken auch die Auswertungen der Dosisabdeckung der geplanten, statischen 3D-Dosisverteilungen und einer Dosisakkumulation basierend auf den einzelnen, patientenspezifischen 4D-CT-Bilddatensätzen.

Bei patientenspezifischer Analyse der Ergebnisse der simulierten Fraktionen zeigt sich, dass für Patient 01 in allen Fraktionen eine ausreichende Abdeckung des kompletten CTV erreicht wird ($> 95\%$ der verschriebenen Dosis; siehe D_{\min} in Tab. 4.1 und das Dosis-Volumen-Histogramm in Abb. 4.7). Für diesen Patienten scheint die zu erwartende Atembewegung keine Neuplanung des anhand von statischen Daten erzeugten Bestrahlungsplans notwendig zu machen. Dies zeigt sich auch in dem sehr geringen Unterschied von mittlerer Dosisverteilung und geplanter Dosisverteilung in Abb. 4.8 und der zudem sehr geringen Standardabweichung.

Für Patient 02 zeigt sich eine Gefahr der Unterdosierung im CTV bei der zu erwartenden Atembewegung. Dies lässt sich u. a. an der mittleren Minimaldosis von 92.9% in Tab. 4.1 und des partiellen Verlaufs der 95%-Isodosislinie der mittleren Dosisverteilung im CTV in Abb. 4.8 erkennen. Für diesen Patienten scheint demnach die Erzeugung eines neuen Bestrahlungsplans angebracht zu sein. Es ist allerdings zu beachten, dass diese potenzielle Unterdosierung auch bei reiner Nutzung der Dosisakkumulation mittels 4D-CT-Bilddaten aufgedeckt worden wäre. Die simulierten Fraktionen zeigen aber zusätzlich auf, dass dies kein singuläres Ergebnis des spezifischen Bewegungsmusters ist, sondern mutmaßlich alle Fraktionen betroffen sind.

Die Ergebnisse für Patient 03 sind grob mit denen von Patient 01 vergleichbar. Auch für diesen Patienten sind mittlere Dosisverteilung und geplante Dosis in Abb. 4.8 bei zugleich geringer Standardabweichung kaum zu unterscheiden. Es zeigt sich allerdings an den Ergebnissen in Tab. 4.1, dass für diesen Patienten dennoch eine potenzielle Ge-

fahr der Unterdosierung besteht, da in einzelnen Fraktionen teilweise eine Dosis $<95\%$ im CTV ermittelt wurde. Die geringen Abweichungen von der Mindestdosis ($<1\%$) in Einzelfällen, rechtfertigen generell noch keine Neuplanung, die simulierten Fraktionen weisen aber im Gegensatz zur Nutzung einer rein 4D-CT-basierten Dosisakkumulation auf die mögliche atemungsbedingte Gefahr der Unterdosierung hin.

Für Patient 04 zeigen sich verglichen zu der durch den Bestrahlungsplan angestrebten Dosisabdeckung im CTV die größten Abweichungen in Tab. 4.1 und Abb. 4.7. Für diesen Patienten liegt die minimale Dosis im CTV in den 100 simulierten Fraktionen zwischen $80.2\% - 92.5\%$, was klar für eine Neuplanung spricht. Dies lässt sich ebenso aus der unvollständigen Abdeckung des CTV durch die 95% -Isodosislinie der mittleren Dosisverteilung in Abb. 4.8 in Verbindung mit der sehr hohen Standardabweichung folgern. Auch für diesen Patienten ist wieder eine Ermittlung des Unterdosierungspotenzials durch Beachtung der Atembewegung in den 4D-CT-Bilddaten möglich. Allerdings zeigen die Werte in Tab. 4.1 und das zugehörige Dosis-Volumen-Histogramm in Abb. 4.7, dass die Auswirkungen der Atmung dann zumindest stark unterschätzt würden, was zu einem weiteren ungünstigen Bestrahlungsplan führen könnte.

4.2.5 Zusammenfassung

Diese Anwendung hat beispielhaft aufgezeigt, wie der präsentierte Ansatz zur modellbasierten Atembewegungssimulation im Kontext der Strahlentherapieplanung eingesetzt werden kann, um potenziell ungeeignete Bestrahlungspläne zu identifizieren. Hierzu wurde das modellbasierte Atembewegungssimulationsverfahren mit einem Dosisakkumulationsschema gekoppelt, wodurch die Simulation der atemungsbedingten Effekte auf die Dosisverteilung einer Fraktion möglich wird. Werden auf diese Weise mehrere Fraktionen simuliert, lassen sich so mittlere Dosisverteilungen und Standardabweichungen schätzen. Diese statistische Quantifizierung der atemungsbedingten Auswirkungen auf die Dosisverteilungen ermöglicht eine umfassendere Einschätzung der Qualität eines Bestrahlungsplans hinsichtlich potenzieller Über- oder Unterdosierungen, als dies durch die Berücksichtigung der Atembewegung aus einem patientenspezifischen 4D-CT-Bilddatensatz möglich wäre. Dies wurde in der durchgeführten Evaluation exemplarisch für Patienten mit IMRT-Bestrahlungsplänen gezeigt.

Die grundlegende Idee der modellbasierten 4D-Dosisimulation ist nicht auf das hier genutzte Step-&Shoot-IMRT als Bestrahlungstechnik beschränkt. Durch den Austausch des Dosisakkumulationsschemas können auch beliebige andere Techniken in die Simulation integriert werden. Dies wurde unter Mitwirkung des Autors dieser Dissertation in [Sothmann u. a., 2017a] gezeigt, wo modellbasierte Dosisakkumulation für VMAT-Bestrahlung (engl. Volumetric Arc Therapy; IMRT mit kontinuierlicher Bestrahlung) realisiert wurde. Es ist zudem eine Weiterentwicklung hinsichtlich der Integration interfraktioneller Bewegungsvariabilität in das Simulationsframework wünschenswert, damit auch deren Effekte in die statistische Analyse einfließen können.

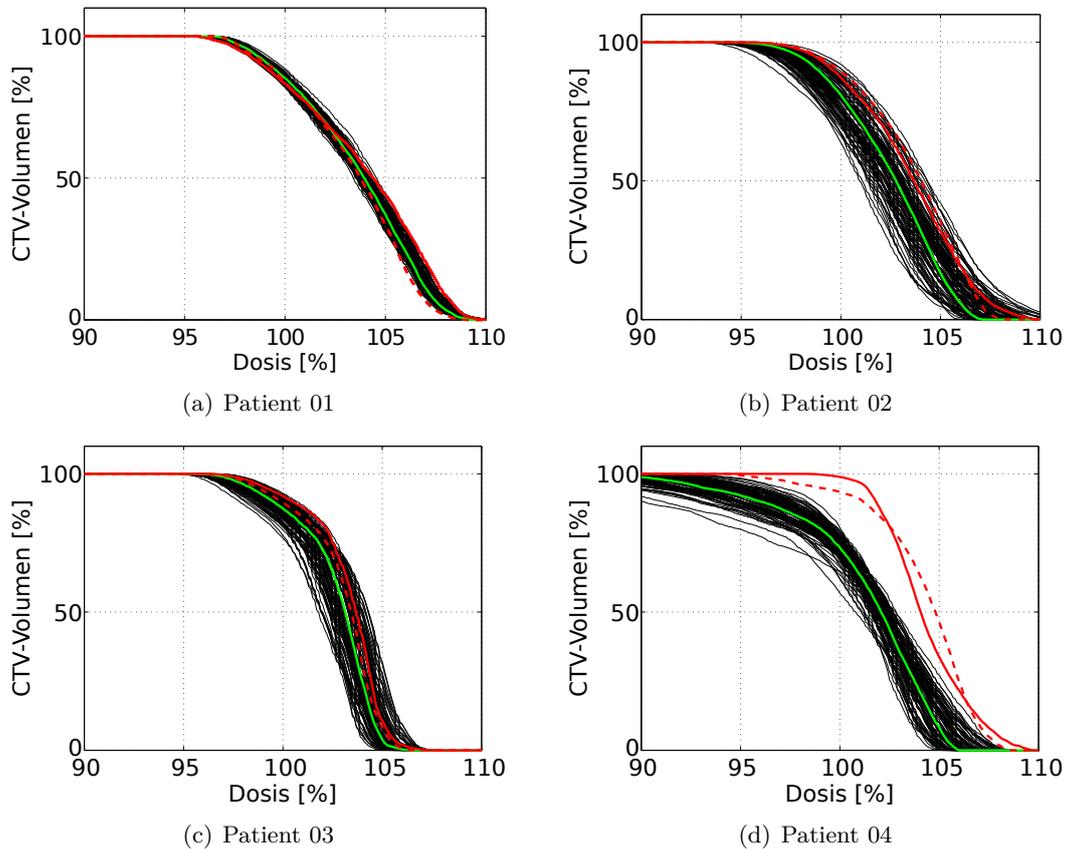


Abb. 4.7: Dosis-Volumen-Histogramme, welche die Dosisabdeckung des CTV bei unterschiedlicher Berechnung der Dosisverteilung für alle vier Patienten zeigen. Das Diagramm gibt hierbei an, wie viel Prozent des Gesamtvolumens des CTV die jeweilige Dosis (gemessen in Prozent der verschriebenen Dosis) erhalten haben. Nach ICRU-Empfehlungen [ICRU, 1999] tritt eine Unterdosierung bei weniger als 95% der verschriebenen Dosis ein. Rote durchgezogene Linie: Geplante Dosisverteilung bei statischer 3D-Planung. Rote gestrichelte Linie: Dosisakkumulation basierend auf 4D-CT-Bilddatensatz. Grüne Linie: Mittlere akkumulierte Dosis nach $n_{sim} = 100$ simulierten Fraktionen. Schwarze Linien: Einzelergebnisse für die $n_{sim} = 100$ simulierten Fraktionen. Erweiterte Abbildung nach [Wilms u. a., 2014b].

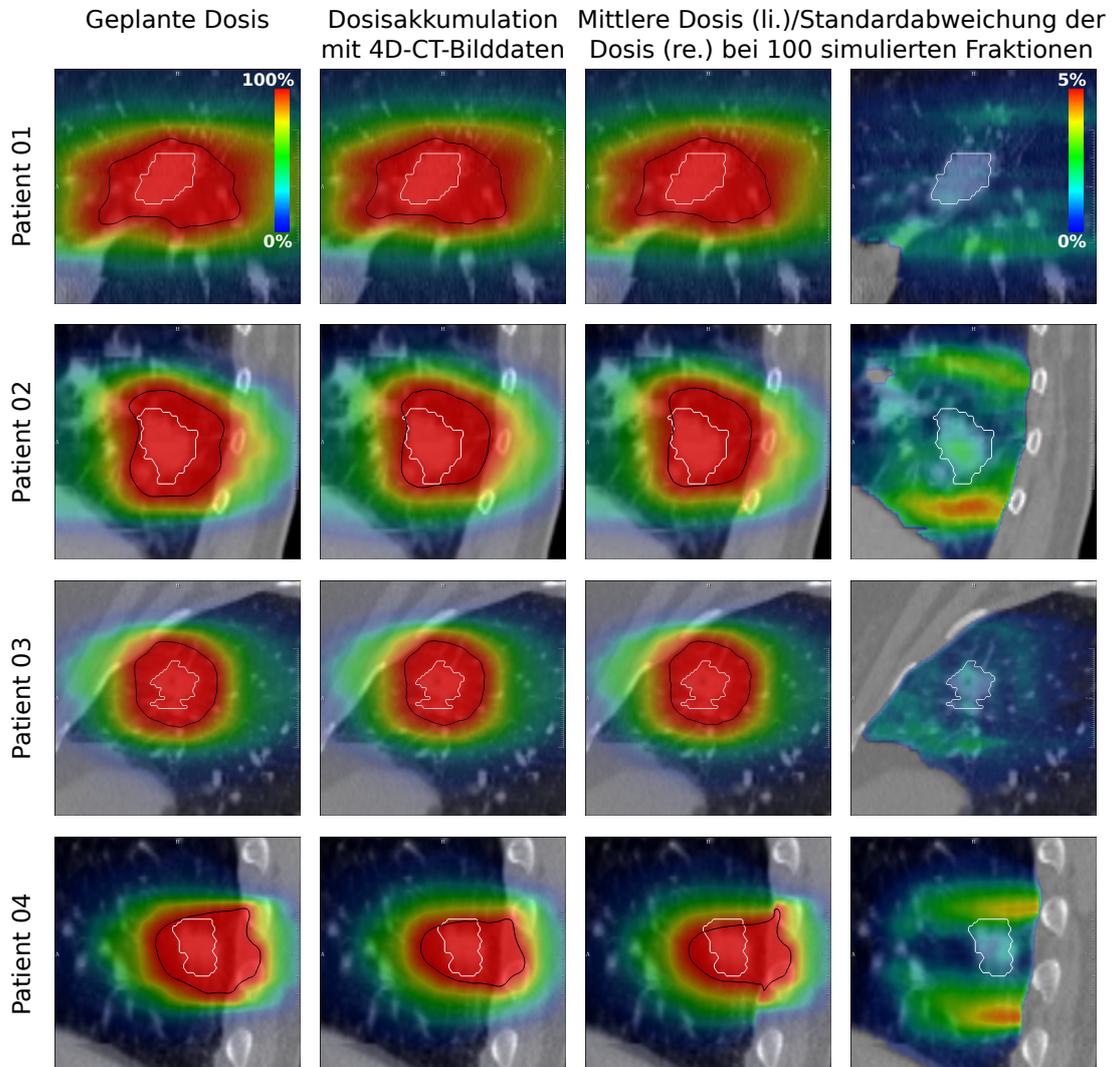


Abb. 4.8: Illustration der Dosisverteilung in einer ausgewählten sagittalen Schicht des Referenz-CT für alle vier Patienten (Zeilen) bei unterschiedlicher Berücksichtigung der Atembewegung zur Abschätzung der Dosisverteilung (Spalten). Alle Angaben sind in Prozent der verschriebenen Fraktionsdosis angegeben. Die weiße Linie kennzeichnet das CTV. Die schwarze Linie ist die 95%-Isodosislinie, welche Bereiche im CTV sichtbar macht, die eine zu geringe Dosis erhalten haben. Abbildung modifiziert nach [Wilms u. a., 2014b].

4.3 Anwendung 2: Virtual-Reality-Training von Punktionseingriffen

Moderne Virtual-Reality-Systeme bieten Medizinern die Möglichkeit ihre Fertigkeiten hinsichtlich des simulierten Eingriffes zu verbessern oder einen Eingriff patientenspezifisch zu üben, ohne hierbei direkt an Patienten arbeiten zu müssen (siehe Übersicht in [Ruthenbeck u. a., 2015] und Einführung in [Handels, 2009; Preim u. a., 2014]). Hierdurch werden potenziell gefährliche Situationen für den Patienten verringert und im Optimalfall wird durch das maßgeschneiderte, steuerbare und beliebig oft wiederholbare Training eine verbesserte Behandlung möglich. Für eine starke Immersion bieten Virtual-Reality-Simulatoren im Regelfall sowohl eine stereoskopische 3D-Darstellung des virtuellen Körpermodells und des simulierten Eingriffs über spezielle Brillen und/oder Bildschirme als auch die Möglichkeit zur haptischen Interaktion mit der Szene (visuelles und haptisches Rendering). Die hierfür genutzten Eingabegeräte erlauben einerseits die Steuerung der für den simulierten Eingriff benötigten Werkzeuge, können aber andererseits auch die durch die Interaktion mit dem virtuellen Patienten an den Werkzeugen auftretenden Kräfte durch Motoren an den Benutzer zurückgegeben, sodass der Eingriff fühlbar wird.

Ein Schwerpunkt im Bereich der medizinischen Virtual-Reality-Systeme lag in den vergangenen zehn Jahren auf der Entwicklung von Simulatoren für nadelbasierte Eingriffe. Beispiele hierfür sind Prostatapunktionen [Goksel u. a., 2013], die Nervus-fermoralis-Blockade [Ullrich u. a., 2012], Lumbalpunktionen [Färber u. a., 2009] und Leberpunktionen bzw. -biopsien [Ni u. a., 2011; Villard u. a., 2009, 2014]. Die Simulation der Nadelpunktion der Leber und ihrer Gefäße ist auch zentraler Anwendungsfall des am Institut für Medizinische Informatik der Universität zu Lübeck entwickelten Punktionssimulators [Fortmeier u. a., 2016; Fortmeier, 2016], der in dieser Arbeit als Ausgangspunkt dient.

Auch wenn Punktionssimulatoren in der Regel die durch die Nadel oder andere simulierte Gegenstände erzeugten lokalen Weichteildeformationen realistisch visuell und haptisch simulieren [Preim u. a., 2014], ist der Rest des virtuellen Patienten zumeist statisch. Diese Annahme greift insbesondere bei der Simulation von Eingriffen zu kurz, die im Thorax oder Abdomen stattfinden, da diese Regionen stark von der Atembewegung beeinflusst werden. Dies bedeutet, dass beispielsweise für eine realistische Leberpunktion die atmungsbedingte Bewegung und Deformation der thorakalen und abdominalen Strukturen für den Nutzer sichtbar sein muss und ebenso die durch die Atmung auf die Nadel einwirkenden Kräfte am Eingabegerät spürbar sein müssen.

Die Integration der Atmung in Virtual-Reality-Simulatoren hat bisher nur mit starken Vereinfachungen stattgefunden. So werden in [Ni u. a., 2011; Villard u. a., 2009, 2014] zwar Effekte der Atmung in die Visualisierung integriert, es erfolgt allerdings keine entsprechende Rückgabe der Kräfte an den Nutzer. Zudem werden in diesen

Arbeiten relativ einfache, weitgehend manuell definierte, generische Atemmodelle genutzt, die mit einem Sinussignal gesteuert werden und zumindest in [Ni u. a., 2011] nur eine periodische Atembewegung erzeugen. Darüber hinaus ist die sichtbare Atembewegung in [Ni u. a., 2011] auf einen kleinen Bereich um das Einsatzgebiet beschränkt, was auch für den in [Wu u. a., 2007] beschriebenen Ansatz gilt.

Ziel ist es nun, realistische Atembewegung in den institutseigenen Nadelpunktionsimulator zu integrieren. Hierbei soll die Atmung im Gegensatz zu verfügbaren Systemen sowohl in die Visualisierung als auch die Krafrückgabe integriert werden. Zudem soll das Gesamtsystem die Möglichkeit bieten patientenspezifische, variable Atmung in Echtzeit zu simulieren. Hauptbeitrag dieser Arbeit zur Erreichung dieses Ziels ist die Idee, die Atembewegung basierend auf Kapitel 4.1.3 oder unter Nutzung eines Originalsurrogatsignal mithilfe eines patientenspezifischen Korrespondenzmodells zu simulieren. In Verbindung mit dem Promotionsprojekt von Dirk Fortmeier [2016], welches sich u. a. mit der Entwicklung der für die Visualisierung und Krafrückgabe notwendigen Methoden beschäftigt, ist ein Gesamtsystem entstanden [Fortmeier u. a., 2015; Wilms u. a., 2015a], das die genannten Anforderungen erfüllt und im Folgenden aus Sicht der Bewegungsmodellierung zusammenfassend erläutert wird.

4.3.1 Vom statischen zum atmenden virtuellen Patienten

Der Hardwareaufbau des institutseigenen Nadelpunktionsssimulators [Fortmeier u. a., 2016; Fortmeier, 2016] entspricht weitgehend dem Standard in diesem Bereich und besteht neben dem Basisrechner aus einem Geomagic Phantom Premium 1.5 6DOF als haptischem Eingabegerät und einer Shutterbrille mit entsprechendem Monitor zur stereoskopischen 3D-Darstellung des simulierten Eingriffs. Für die Simulation des (virtuellen) Patienten kommt ein 3D-CT-Bilddatensatz $I_{CT} : \Omega \rightarrow \mathbb{R}$ zum Einsatz, wobei die eigentliche Visualisierung des Körpers durch direktes Volumenrendering erfolgt, was die vorherige Erzeugung von Oberflächenmodellen nahezu überflüssig macht. Zudem dienen die Bilddaten als Grundlage für die Simulation eingriffsspezifischer, zusätzlicher Modalitäten wie Ultraschall. Die Informationen in den Bilddaten dienen weiterhin ebenfalls als Grundlage zur Berechnung der an den Nutzer zurückgegebenen Kräfte, die bei Interaktion von Nadel und Körper auftreten. Zusammenfassend lässt sich festhalten, dass der virtuelle Patientenkörper durch die Informationen in I_{CT} gegeben ist und mit der Ausnahme von lokal stark begrenzten Weichteildeformationen, die u. a. bei Interaktion von Nadel und Gewebe visualisiert werden, statisch ist.

Da alle Informationen, die für die Visualisierung und die Krafrückgabe benötigt werden, direkt aus dem 3D-CT-Bilddatensatz I_{CT} abgeleitet werden, kann eine Integration der Atembewegung in den Simulator einfach durch Deformation des CT-Bildes in Echtzeit erreicht werden (siehe Abb. 4.9). Steht für den zu simulierenden Patienten ein 4D-CT-Bilddatensatz und ein zugehöriges 1D-Surrogatsignal $s(t)$ zur Verfügung, kann ein patientenspezifisches (diffeomorphes) Korrespondenzmodell M (siehe Glg.

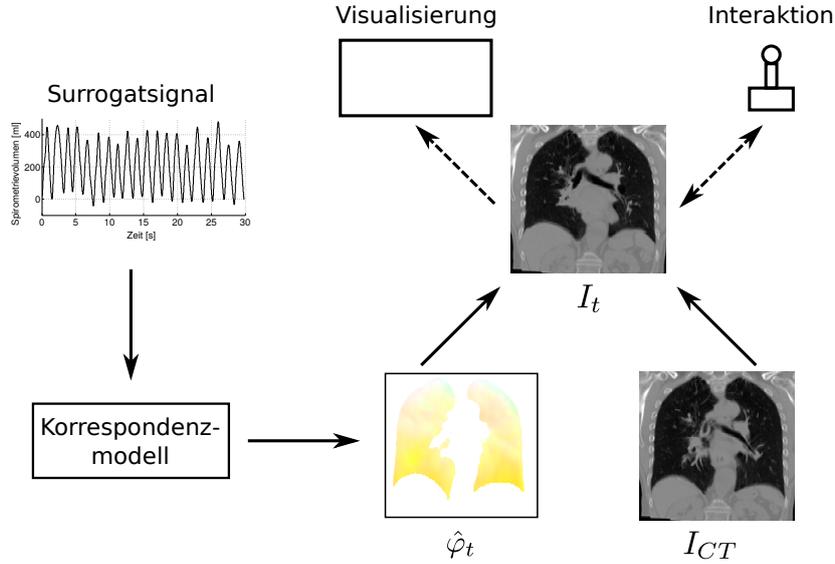


Abb. 4.9: Schematische Illustration der modellbasierten Atembewegungssimulation in einem Virtual-Reality-Simulator, bei der das statische Referenzbild I_{CT} mithilfe einer durch das Korrespondenzmodell erzeugten Transformation $\hat{\varphi}_t$ in jedem Simulationsschritt t deformiert wird.

4.6) erstellt werden, welches anschließend zur Bewegungssimulation genutzt werden kann (siehe Kapitel 4.1.3). Hierbei wird davon ausgegangen, dass I_{CT} das zur Erzeugung des Korrespondenzmodells genutzte Referenzbild ist. Das durch die Atembewegung deformierte Referenzbild $I_t : \Omega \rightarrow \Omega$ zu einem Zeitpunkt t der Simulation ergibt sich dann als

$$I_t = I_{CT} \circ \hat{\varphi}_t^{-1} \text{ mit } \hat{\varphi}_t^{-1} = M(\zeta_t)^{-1} \text{ und } \zeta_t = [s(t), \frac{d}{dt}s(t)]^T. \quad (4.11)$$

Hierbei bezeichnet $\hat{\varphi}_t^{-1}$ die Inverse der durch das Korrespondenzmodell geschätzten Transformation $\hat{\varphi}_t = M(\zeta_t)$. Bei einem diffeomorphen Korrespondenzmodell M lässt sich diese inverse Transformation durch Negation des $\hat{\varphi}_t$ zugrunde liegenden Geschwindigkeitsfeldes \hat{v} erzeugen (siehe Kapitel 2.2). Das in Glg. 4.11 eingesetzte Originalsurrogatsignal $s(t)$ kann zudem auch durch ein simuliertes Signal ersetzt werden (siehe Kapitel 4.1.3). Das deformierte Referenzbild I_t kann dann anschließend dem Volumenderer übergeben werden, um das atmende Patientenmodell zu visualisieren. Ebenso kann die Transformation $\hat{\varphi}_t$ genutzt werden, um möglicherweise eingesetzte Oberflächenmodelle punktweise zu deformieren oder die Position und Form der simulierten Werkzeuge zu verändern. Auch Informationen zur Krafrückgabe können so gewonnen werden. Prinzipiell ist das hier skizzierte Vorgehen direkt in die Praxis umsetzbar, wobei die größte Herausforderung in der Berechnung der Transformationen $\hat{\varphi}_t$ und $\hat{\varphi}_t^{-1}$ in Kombination mit entsprechender Visualisierung und Krafrückgabe in Echtzeit liegt.

Dieses und weitere Probleme, die für den Praxiseinsatz gelöst werden müssen, werden ebenso wie die für eine erste praktische Umsetzung vorgenommenen Vereinfachungen und Einschränkungen im Folgenden diskutiert.

4.3.1.1 Einschränkungen der ersten praktischen Umsetzung

Generell sollte das Ziel sein, für die modellbasierte Simulation der Atembewegung in dieser Anwendung ein diffeomorphes Korrespondenzmodell zu nutzen, insbesondere weil die inverse Transformation $\hat{\varphi}_t^{-1}$ benötigt wird. Da das als Patientenmodell dienende Referenz-CT allerdings einen möglichst großen Teil des Körpers abdecken sollte, um die Simulation realistisch zu gestalten, müssen bei der Schätzung der Transformationen Unstetigkeiten in der Atembewegung durch Gleitbewegungen korrekt behandelt werden. Dies setzt u. a. ein entsprechendes Registrierungsverfahren für das Modelltraining voraus (siehe auch Diskussion in Kapitel 3.3). Da das in dieser Arbeit für das Modelltraining eingesetzte diffeomorphe Registrierungsverfahren (siehe Kapitel 2.2) Gleitbewegungen nicht korrekt berücksichtigen kann und es nach Wissen des Autors dieser Dissertation kein entsprechend frei verfügbares Verfahren gibt, wird für die erste praktische Umsetzung im Simulator auf die Erzeugung eines diffeomorphen Korrespondenzmodells verzichtet. Für die Registrierung der 4D-Bilddaten im Rahmen des Modelltrainings wird das in [Schmidt-Richberg u. a., 2012; Schmidt-Richberg, 2013] beschriebene Registrierungsverfahren eingesetzt, welches ebenfalls am Institut für Medizinische Informatik der Universität zu Lübeck entwickelt wurde und Gleitbewegungen korrekt behandelt, aber nicht zwingend diffeomorphe Transformationen schätzt und nicht auf der Grundlage statischer Geschwindigkeitsfelder arbeitet. Das entsprechende Korrespondenzmodell wird deshalb rein auf den geschätzten Transformationen aufgebaut. Durch Nutzung eines nicht-diffeomorphen Korrespondenzmodells kann bei Nutzung eines 2D-Surrogatsignals jede Richtungskomponente eines Bewegungsvektors von $\hat{\varphi}_t$ mit 4(+2) arithmetischen Standardoperationen¹⁰ effizient auf der GPU des Simulators berechnet werden.

Das direkte Volumenrendering setzt weiterhin die Transformation (Warping) des kompletten Referenzbildes I_{CT} voraus. Dieses Vorgehen erscheint bezogen auf die hierfür benötigte Rechenzeit selbst bei einer GPU-basierten Implementierung nicht sinnvoll, da im Allgemeinen aufgrund von Verdeckungen nicht das komplette Bildvolumen für das Rendering genutzt wird. Aus diesem Grund wird im Simulator aufbauend auf [Kurzion u. a., 1995] ein Renderingverfahren eingesetzt, bei dem nicht das zu visualisierende Volumen deformiert wird, sondern die für das Raycasting ausgesendeten Strahlen (siehe [Fortmeier u. a., 2015; Fortmeier, 2016]). Da auch zur Berechnung der deformierten Strahlen zumindest für die Punkte auf dem Strahl die inverse Transformation $\hat{\varphi}_t^{-1}$

¹⁰Mittelwertbereinigung des Surrogatsignals (2 Subtraktionen); Multiplikation der jeweiligen Modellkoeffizienten mit den Signaldimensionen (2 Multiplikationen); Summation der skalierten Modellkoeffizienten und der mittleren Bewegung (2 Additionen)

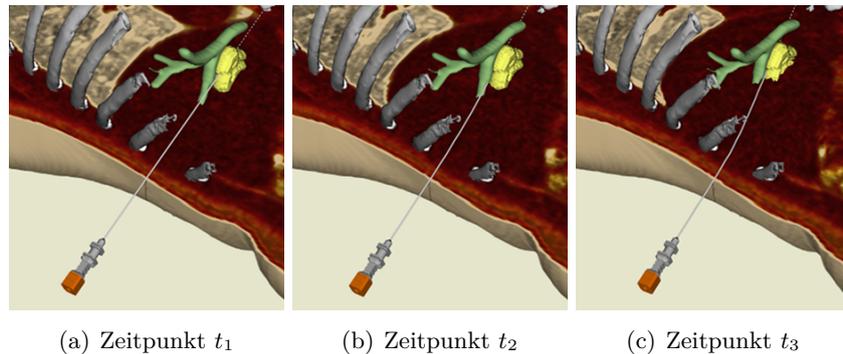


Abb. 4.10: Ausschnitt des virtuellen Körpers bei simulierter Leberpunktion. Die einzelnen Bilder geben den zeitlichen Verlauf einer Ausatemungssequenz wieder. Hierbei ist neben der generellen Bewegung der visualisierten Organe/Strukturen u. a. auch die durch die Atembewegung verursachte Biegung der Nadel sichtbar. Die Darstellung der Rippen (grau), Gallengänge (grün) und des simulierten Tumors (gelb) basieren zur besseren Unterscheidung auf einem oberflächenbasierten Rendering. Abbildung aus [Wilms u. a., 2015a].

von $\hat{\varphi}_t$ benötigt wird, muss diese entsprechend effizient lokal erzeugt bzw. approximiert werden können. Hierfür wird das in [Chen u. a., 2008] beschriebene Fixpunktverfahren eingesetzt, wobei für eine Steigerung der Laufzeiteffizienz das Ergebnis des jeweils von der Sichtebene aus vorher auf dem Strahl gelegenen Punktes als Startwert für den nachfolgenden Punkt eingesetzt wird. Eine genaue Beschreibung der technischen Fakten der Implementierung ist in [Fortmeier, 2016] zu finden.

4.3.2 Initiale Experimente und deren Ergebnisse

Für die initiale experimentelle Verifikation des beschriebenen Verfahrens wird eine Leberpunktion simuliert. Als Modell dient hierfür ein Patient der WUSTL-Daten (siehe Kapitel 3.1.1), für den ein 4D-CT-Bilddatensatz mit 14 Atemphasen und ein 1D-Spirometersignal zur Verfügung stehen. Der 4D-CT-Bilddatensatz bildet sowohl den kompletten Thorax als auch das komplette Abdomen ab, sodass bis auf den Kopf und die Extremitäten alle Bereiche des Körpers dargestellt werden können (siehe Abb. 4.11). Die einzelnen 3D-Bilder bestehen aus $512 \times 512 \times 460$ Voxeln mit einer räumlichen Auflösung von $1 \times 1 \times 1.5 \text{ mm}^3$, wobei sie auf eine Auflösung von $256 \times 256 \times 256$ Voxeln herunterskaliert werden, damit sowohl das Referenzbild als auch das Korrespondenzmodell in den Speicher der Grafikkarte des verwendeten Testsystems (Intel i7 970 3.20 GHz-Prozessor mit 24 GB RAM und einer Nvidia GTX 680-Grafikkarte mit 3 GB RAM) passen. Als Surrogatsignal zur Steuerung der Bewegungssimulation wird eine Sequenz (60 s lang) mit maximaler Variabilität des aufgenommenen Originalsignals

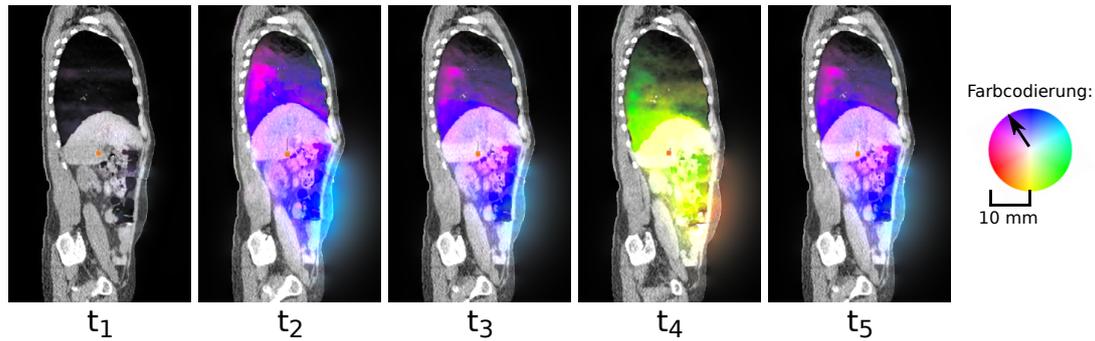


Abb. 4.11: Sagittale Schicht des deformierten Referenzbildes I_{CT} zu fünf verschiedenen Zeitpunkten der Simulation mit farbcodierter Darstellung der simulierten Atembewegung. Abbildung modifiziert aus [Wilms u. a., 2015a].

verwendet. Hier ist aber alternativ auch die Nutzung eines simulierten Surrogatsignals möglich.

Eine visuelle Darstellung verschiedener Zeitpunkte der Punktionsimulation ist in Abb. 4.10 und Abb. 4.11 zu finden. Eine Bewegtbilddarstellung der Simulation ist in dem zur Publikation [Fortmeier u. a., 2015] gehörenden Video¹¹ verfügbar. Bei visueller Begutachtung zeigt sich insgesamt eine durchaus realistische Bewegungssimulation, die nicht nur visuell wahrnehmbar ist, sondern durch das haptische Eingabegerät auch durch den Nutzer fühlbar wird (siehe Video). Für die Visualisierung werden auf dem Testsystem Renderingzeiten von < 30 ms pro Einzelbild (> 33 Einzelbilder pro Sekunde) bei einer Ausgabegröße von 1224×1014 Pixeln erreicht, was insgesamt die Echtzeitfähigkeit des Ansatzes belegt.

Auch wenn die Bewegungssimulation insgesamt realistisch ist, sind in Regionen, in denen es zu einer Gleitbewegung zwischen verschiedenen Strukturen kommt, teilweise Artefakte in der Darstellung wahrzunehmen. Beispielsweise ist in Abb. 4.12 eine lokale Deformation der direkt an die Lungen angrenzenden Rippenteile sichtbar, die so nicht realistisch ist. Da das zur Korrespondenzmodellgenerierung genutzte Registrierungsverfahren diese Unstetigkeiten korrekt in den Transformationen repräsentiert, ist es wahrscheinlich, dass diese Artefakte durch eine Kombination von Problemen der nachfolgenden Verarbeitungsschritte verursacht werden: Durch zeitliche Inkonsistenzen bei der Registrierung kann es zu einer Verschmierung des Gleitrandes im Korrespondenzmodell kommen. Die Artefakte können ihre Ursache ebenso im genutzten Invertierungsverfahren haben. Zusätzlich sind auch Interpolationseffekte als Mitverursacher denkbar. Eine weitere Diskussion dieser Probleme ist auch in [Fortmeier, 2016] zu finden.

¹¹<http://ieeexplore.ieee.org/document/7124506/media> (Stand: 18.01.2018)

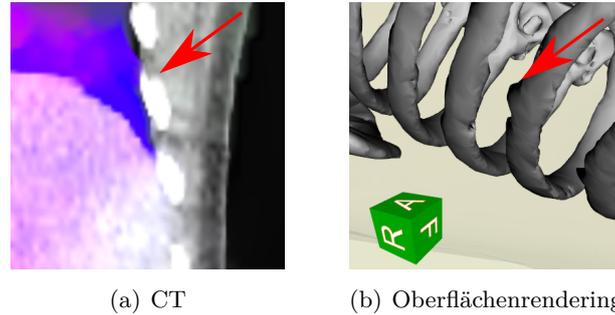


Abb. 4.12: Beispielhafte Darstellung der visuellen Artefakte, die durch die Bewegungsvisualisierung an Gleiträndern auftreten können (hier: Rippen und Lunge). Der rote Pfeil markiert jeweils einen Bereich, in dem der Knochen durch die Lungenbewegung lokal ebenfalls bewegt wird. Abbildung modifiziert nach [Fortmeier u. a., 2015].

4.3.3 Zusammenfassung und weitere Überlegungen

In diesem Teilkapitel wurde beispielhaft gezeigt, wie mithilfe eines patientenspezifischen Korrespondenzmodells simulierte Atembewegung gewinnbringend in die Virtual-Reality-Simulation von Nadelpunktionseingriffen integriert werden kann. Die präsentierte Methodik erlaubt nicht nur die Integration der Atembewegung in die Visualisierung des virtuellen Patienten, sondern auch die Rückgabe der durch die Bewegung beispielsweise an den simulierten Instrumenten entstehenden Kräfte an den Nutzer, was ein entscheidender Vorteil gegenüber den in der Literatur beschriebenen Lösungen ist. Ein weiterer Vorteil des hier beschriebenen Verfahrens ist die Möglichkeit patientenspezifische Atembewegung zu simulieren, was bei den zumeist generischen Modellen in anderen Simulatoren nicht möglich ist. Als Grundlage wird allerdings ein 4D-Bilddatensatz benötigt, dessen Aufnahme im Fall einer durchzuführenden Leberpunktion aus Kosten- und/oder Strahlenschutzgründen sicherlich nicht realistisch ist. Aus diesem Grund wird aktuell in weiteren Arbeiten (u. a. [Mastmeyer u. a., 2018]) unter Mitwirkung des Autors dieser Arbeit untersucht, inwieweit beispielsweise die Korrespondenzmodelle ähnlich atmender Patienten zur patientenspezifischen Simulation der Atembewegung eingesetzt werden können, auch wenn keine 4D-Bilddaten verfügbar sind. Neben den 4D-Bilddaten ist das Vorhandensein eines Atemsignals weitere Voraussetzung für die Nutzung des präsentierten Systems. Das hier genutzte Spirometersignal kann hierbei prinzipiell durch ein beliebiges anderes Signal ersetzt werden. Insbesondere ist auch die Nutzung von Kamerasystemen (siehe 3.1.2) denkbar, die beispielsweise auch zur Steuerung des Virtual-Reality-Simulators durch den Patienten in Echtzeit eingesetzt werden könnten. Die hier präsentierte Evaluation des Verfahrens ist als initiales Proof-of-Concept zu verstehen und sollte in zukünftigen Arbeiten um weitere (virtuelle) Patienten und eine ausführliche Nutzerstudie mit Medizinern ergänzt werden.

4.4 Zusammenfassung

Im vorliegenden Kapitel wurde aufbauend auf dem in Kapitel 2 präsentierten Framework zur patientenspezifischen Korrespondenzmodellierung eine Methode zur modellbasierten Simulation realistischer und variabler Atembewegung vorgestellt. Hierfür wurde eine neue Methode zur B-Spline-basierten Erzeugung von variablen Surrogatsignalen beliebiger zeitlicher Länge mit typischen patientenspezifischen Variationen entwickelt. Durch Kombination eines Korrespondenzmodells mit einem (simulierten) Surrogatsignal können dann Atmungssequenzen beliebiger Länge patientenspezifisch simuliert werden. Dieses Bewegungssimulationsverfahren wurde anschließend zur Untersuchung und Bewertung der Auswirkung der Atembewegung in der Strahlentherapieplanung eingesetzt. Weiterhin wurde gezeigt, wie die Bewegungssimulation in einen medizinischen Virtual-Reality-Simulator integriert werden kann, um die Immersion bei simulierten Eingriffen im Thorax oder im Abdomen entscheidend zu verbessern.

Abseits der beiden präsentierten Anwendungen ist aber auch ein Einsatz des Bewegungssimulationsverfahrens in vielfältigen anderen Planungs-, Simulations- oder Verifikationsszenarien möglich. Beispielsweise könnte das in [Hecht u. a., 2013] beschriebene Verfahren zur Planung von Nadelpunktionspfaden bei Eingriffen mit zu erwartender Atembewegung um eine statistische Auswertung der Atembewegungseffekte analog zu Kapitel 4.2.1 ergänzt werden. Das entwickelte Verfahren zur B-Spline-basierten Surrogatsignalsimulation kann weiterhin auch losgelöst von der Nutzung mit einem Korrespondenzmodell flexibel eingesetzt werden. Mögliche Anwendungen sind hier u. a. die Ansteuerung der im Bereich der Qualitätssicherung in der Strahlentherapie eingesetzten Hardwarephantome [Grohmann u. a., 2015], für die derzeit entweder aufgenommene Atemsignale oder einfache Sinussignale eingesetzt werden. Bezüglich Erweiterungen oder Verbesserungen des Surrogatsignalsimulationsverfahrens ist neben der Untersuchung von Alternativen zur für die Modellierung eingesetzten Normalverteilungsannahme der B-Spline-Parameter auch die Integration von interfraktioneller Bewegungsvariabilität und die Integration verschiedener Patienten ähnlich zu [Quirk u. a., 2012] sinnvoll.

Kapitel 5

Subpopulationsbasierte Erweiterung des Korrespondenzmodellierungsframeworks

Kapitelinhalt

5.1	Hintergrund und vergleichbare Arbeiten	86
5.2	Subpopulationsbasierte Korrespondenzmodellierung	88
5.2.1	Von der patientenspezifischen zur populationsbasierten Korrespondenzmodellierung im Atlasraum	91
5.2.2	Sparse Manifold Clustering zur Identifikation von Subpopulationen	94
5.3	Experimente	97
5.3.1	Verwendete Bilddaten	98
5.3.2	Design der Experimente	99
5.4	Ergebnisse	101
5.5	Zusammenfassung und Einordnung der Ergebnisse	104
5.6	Alternative Methode basierend auf affinen Unterräumen	107

Eines der zentralen Ergebnisse der in Kapitel 3 durchgeführten Evaluation des patientenspezifischen Korrespondenzmodellierungsframeworks ist die Feststellung, dass die Genauigkeit der modellbasierten Bewegungsschätzung bei interfraktioneller Bewegungsvariabilität deutlich sinkt. Dieser Effekt lässt sich mutmaßlich durch die mangelnde Repräsentativität der Trainingsdaten des Modells erklären, wodurch dieses entsprechend schlecht generalisiert. Da ausgeprägte interfraktionelle Variationen der Atmung in verschiedenen Studien immer wieder beobachtet werden (siehe u. a. [Dhont u. a., 2017; McClelland u. a., 2011; Ruben u. a., 2016]; u. a. 46% der Patienten mit interfraktionellen Bewegungsamplitudenunterschieden > 5 mm in [Dhont u. a., 2017]), scheint die auch in dieser Arbeit propagierte Nutzung eines auf Planungsdaten gelernten Korrespondenzmodells für die intrainterventionelle Bewegungsschätzung zumindest fragwürdig [McClelland u. a., 2011]. Ein möglicher Ausweg ist die fraktionsspezifische Adaption

des Modells anhand weiterer (Bild-)Daten des Patienten [Fassi u. a., 2014; Yuan u. a., 2016]. In diesem Kapitel wird in diesem Zusammenhang allerdings ein innovativer, neuer populationsbasierter Modellierungsansatz (siehe Abb. 5.1) vorgestellt, der zwei Grundannahmen unterliegt: (1) Eine große Patientenpopulation kann in Subpopulationen von Patienten mit ähnlichen Bewegungsmustern unterteilt werden. (2) Es gibt Subpopulationen von Patienten mit ähnlicher Atembewegung, deren Interpatientenvariabilität ungefähr der interfraktionellen Bewegungsvariabilität eines der enthaltenen Patienten entspricht. Die Idee hierbei ist es, die Repräsentativität der Trainingsdatensmenge des patientenspezifischen Korrespondenzmodells durch die Nutzung von Daten ähnlich atmender Populationspatienten zu erhöhen, um im Mittel eine Verbesserung der Schätzungsgenauigkeit zu erreichen (siehe auch Abb. 5.3).

In diesem Kapitel wird gezeigt, wie Subpopulationen von Patienten mit ähnlicher Bewegungscharakteristik durch Nutzung eines Manifold Clustering-Verfahrens identifiziert werden können und auf Basis dieser Subpopulationen erstellte Korrespondenzmodelle bei Vorliegen von interfraktionellen Bewegungsvariationen eine signifikant höhere Genauigkeit als patientenspezifische oder allgemeine populationsbasierte Korrespondenzmodelle erzielen können. Hierfür liefert Kapitel 5.1 eine generelle Einordnung des Verfahrens in den Forschungskontext, bevor in Kapitel 5.2 die methodische Umsetzung der (sub-)populationsbasierten Korrespondenzmodellierung erfolgt. Das neue Verfahren wird anschließend in Kapitel 5.3 und Kapitel 5.4 umfangreich evaluiert. Die Darstellungen in den genannten Unterkapiteln basieren auf [Wilms u. a., 2015b, 2017c].

5.1 Hintergrund und vergleichbare Arbeiten

Während die Nutzung patientenspezifischer Korrespondenzmodelle oder patientenspezifischer Atembewegungsmodelle im Allgemeinen sowohl in der Literatur als auch in der klinischen Praxis weitverbreitet ist (siehe u. a. Kapitel 2.1), werden populationsbasierte Modelle in der Literatur seltener erwähnt und haben im klinischen Alltag nahezu keine Relevanz. Bei populationsbasierten Modellen erfolgt die Modellierung im Gegensatz zu einem patientenspezifischen Modell unter Verwendung der Daten einer Menge verschiedener Patienten, der sogenannten Population [Ehrhardt u. a., 2013]. In der Literatur werden neben mittleren Bewegungsmodellen [Ehrhardt u. a., 2011] auch Modelle vorgestellt, welche versuchen die Bewegungsvariabilität innerhalb der Population (statistisch) zu erfassen und für die Bewegungsschätzung oder -analyse nutzbar zu machen [Ehrhardt u. a., 2010; He u. a., 2010; Klinder u. a., 2009, 2010, 2012; Preiswerk u. a., 2014] (erweiterte Übersichten in [Ehrhardt u. a., 2013; McClelland u. a., 2013]).

Ein großer Vorteil populationsbasierter Modelle ist, dass es mit ihnen möglich wird, auch repräsentative Modelle für Patienten zu erzeugen bzw. eine entsprechende modellbasierte Bewegungsschätzung durchzuführen, für die keine oder nur wenige der zur Erstellung eines patientenspezifischen Modells erforderlichen Daten vorliegen. Bei-

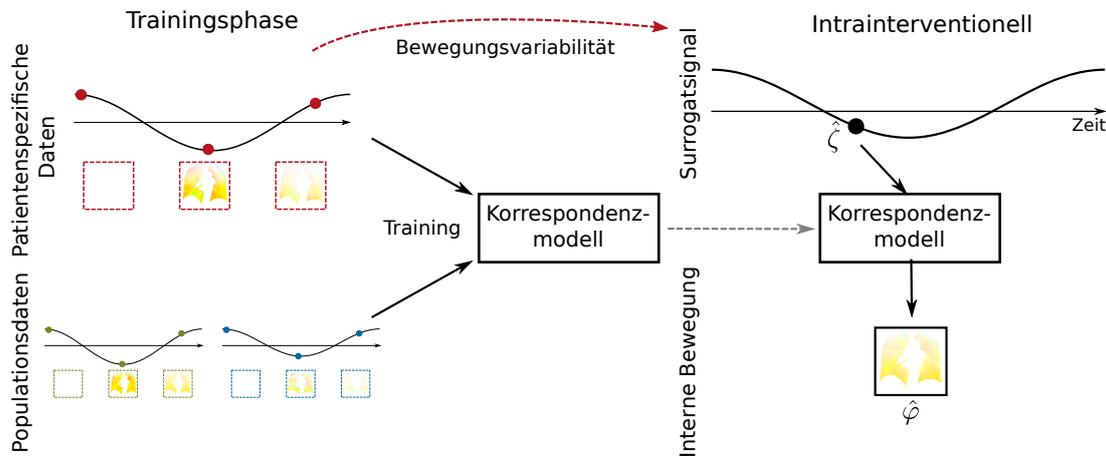


Abb. 5.1: Illustration des generellen Workflows, des in diesem Kapitel präsentierten (sub-) populationsbasierten Korrespondenzmodellierungsverfahrens, bei dem sowohl patientenspezifische Planungsdaten als auch Populationsinformationen eingesetzt werden, um die intrainerventionelle Bewegungsschätzung zu verbessern, da Bewegungsvariationen zu erwarten sind. Siehe auch Abb. 2.1 für die Unterschiede zu rein patientenspezifischen Korrespondenzmodellen.

spielsweise wird das mittlere Bewegungsmodell aus [Ehrhardt u. a., 2011] in [Werner, 2012] eingesetzt, um eine 4D-Dosisakkumulation für Lungentumorpatienten durchzuführen (siehe auch Kapitel 4), für die nur ein statischer 3D-CT-Bilddatensatz verfügbar ist. Hierzu wird das Modell an den jeweiligen (neuen) Patienten angepasst, indem geometrische Unterschiede der Anatomien und Unterschiede bezüglich der Einatmungstiefe zwischen Modellreferenz und Patientenbild ausgeglichen werden. Diese patientenspezifische Adaption des Modells ist zentraler Schritt bei der Verwendung nahezu aller Populationsmodelle und die Genauigkeit dieses Prozesses hat zusammen mit der Populationsgröße bzw. der Repräsentativität der Stichprobe großen Einfluss auf die Genauigkeit der modellbasierten Bewegungsschätzung oder -analyse.

Eine mögliche Erklärung für die geringe Relevanz von Populationsmodellen in der klinischen Praxis ist die in einigen initialen Evaluationen [Ehrhardt u. a., 2011; Klinder u. a., 2009] festgestellte (teilweise deutlich) geringere Genauigkeit der populationsbasierten Bewegungsschätzung im Vergleich zur Nutzung patientenspezifischer Modelle. Wird, wie üblich, die gesamte Population zur Erzeugung und patientenspezifischen Adaption des Populationsmodells genutzt, können diese Genauigkeitsdefizite wiederum ihre Ursache in der Heterogenität der Bewegungseigenschaften der verwendeten Gesamtpopulation haben, sodass die im Einzelfall benötigten Eigenschaften in der Gesamtheit untergehen. Diese Erklärung ist eng verbunden mit der diesem Kapitel zugrunde liegenden Annahme, dass es in einer großen Population von Patienten, Subpopulationen von Patienten mit ähnlichen Bewegungsmustern gibt. Dieser Annahme

folgend, sollte ein Modell basierend auf der passenden Subpopulation bessere Ergebnisse liefern als ein Modell der Gesamtpopulation. Die Ergebnisse in [Samei u. a., 2012; Tanner u. a., 2015] für die atmungsbewegte Leber und in [Peressutti u. a., 2013a; Peressutti u. a., 2014] für die weniger komplexe atmungsbedingte Verschiebung/Deformation des Herzens zeigen, dass durch die Identifikation und Nutzung von Subpopulationen von Patienten oder Modellen tatsächlich eine Verbesserung der Bewegungsschätzung gegenüber normalen Populationsmodellen erreicht werden kann. Die Auswahl oder Gewichtung der im Einzelfall für einen neuen Patienten zu verwendenden Populationspatienten erfolgt in [Samei u. a., 2012; Tanner u. a., 2015] anhand der Surrogatsignale der Patienten, wohingegen in [Peressutti u. a., 2013a; Peressutti u. a., 2014] ein Bewegungs-Clustering in Kombination mit bildbasiert ermittelten anatomischen Merkmalen zum Einsatz kommt.

Von der Grundidee dem Bewegungs-Clustering in [Peressutti u. a., 2013a; Peressutti u. a., 2014] ähnelnd, erfolgt auch in dem in diesem Kapitel präsentierten Ansatz zur subpopulationsbasierten Korrespondenzmodellierung die Identifikation der Subpopulation mittels Clustering. Die neue Methode unterscheidet sich aber sonst deutlich von den diskutierten Arbeiten von Peressutti u. a. [2013a], Peressutti u. a. [2014], Samei u. a. [2012] und Tanner u. a. [2015]: (1) Das hier präsentierte Verfahren setzt Manifold Clustering unter Verwendung von Sparse Coding ein, um die Struktur der Daten bei der Ähnlichkeitsbestimmung von Bewegungsmustern einfließen zu lassen. (2) Das Clustering erfolgt basierend auf komplexen Bewegungssequenzen repräsentiert durch dichte Verschiebungsfelder. (3) Dies ist der erste Ansatz, der subpopulationsbasierte Korrespondenzmodellierung zur Berücksichtigung von interfraktioneller Bewegungsvariabilität nutzt.

5.2 Subpopulationsbasierte Korrespondenzmodellierung

In diesem Teilkapitel wird für die Beschreibung der Korrespondenzmodellierung die in Kapitel 2 und Kapitel 4.1.3 eingeführte Notation eingesetzt und erweitert. Für die angestrebte populationsbasierte Korrespondenzmodellierung sei eine Population bestehend aus n_{pat} Patienten $p \in \{1, \dots, n_{pat}\}$ gegeben. Für jeden Patienten p seien gemäß Kapitel 2 weiterhin ein 4D-Bilddatensatz $(I_{p,j})_{j=1}^{n_{tp}}$ mit n_{tp} einzelnen 3D-Bildern $I_{p,j} : \Omega_p \rightarrow \mathbb{R}$ (mit $\Omega_p \subset \mathbb{R}^3$) und zugehörige korrespondierende Surrogatsignalmessungen $(\zeta_{p,j} \in \mathbb{R}^{n_{sur}})_{j=1}^{n_{tp}}$ als Grundlage für das Modelltraining verfügbar. Mittels diffeomorpher Bildregistrierung (siehe Kapitel 2.2) seien zudem ausgehend vom jeweiligen Referenzbild $I_{p,1}$ ($j = 1$) die nichtlinearen, diffeomorphen Transformationen $(\varphi_{p,j} : \Omega_p \rightarrow \Omega_p)_{j=1}^{n_{tp}}$ erzeugt worden, die die Atembewegung im 4D-Bilddatensatz beschreiben.

Zusätzlich zu den n_{pat} Populationspatienten sei ein neuer Patient $p = 0$ gegeben, für den im Folgenden (subpopulationsbasierte) Korrespondenzmodelle erstellt werden

sollen. Auch für diesen Patienten sind ein 4D-Bilddatensatz $(I_{0,j})_{j=1}^{n_{tp}}$ mit n_{tp} einzelnen 3D-Bildern $I_{0,j} : \Omega_0 \rightarrow \mathbb{R}$ (mit $\Omega_0 \subset \mathbb{R}^3$), korrespondierende Surrogatsignalmessungen $(\zeta_{0,j} \in \mathbb{R}^{n_{sur}})_{j=1}^{n_{tp}}$ und nichtlineare, diffeomorphe Transformationen $(\varphi_{0,j} : \Omega_0 \rightarrow \Omega_0)_{j=1}^{n_{tp}}$ verfügbar. Wobei auch hier die Bildregistrierung ausgehend vom Referenzbild $I_{0,1}$ durchgeführt wurde.

Für die Definition der populationsbasierten Korrespondenzmodelle müssen die Daten (Bilddaten, Surrogatsignal und Transformationen) aller $n_{pat} + 1$ Patienten $p \in \{0, \dots, n_{pat}\}$ (inkl. $p = 0$) räumlich und zeitlich zueinander ausgerichtet sein. Hierfür wird zunächst vereinfachend angenommen, dass ein Zeitpunkt $j \in \{1, \dots, n_{tp}\}$ jeweils für alle Patienten $p \in \{0, \dots, n_{pat}\}$ den gleichen Atemzustand bezeichnet und somit eine zeitliche Korrespondenz aller Daten für alle Patienten besteht. Dies kann bei der Nutzung von 4D-CT-Bilddaten beispielsweise leicht durch die Nutzung des gleichen Rekonstruktionsschemas erreicht werden. Alternativ ist aber auch die explizite Herstellung der zeitlichen Korrespondenz durch Inter- bzw. Extrapolationsmethoden verschiedenster Komplexität und Genauigkeit möglich (siehe z. B. [Avants u. a., 2005; Singh u. a., 2015; Sundaram u. a., 2004]). Ausgehend von dieser Annahme ist weiterhin klar, dass das jeweils für die patientenspezifische Registrierung der 4D-Bilddaten genutzte Referenzbild $I_{p,1}$ für alle Patienten den gleichen Atemzustand abbildet. Da gleicher Atemzustand (z. B. maximale Einatmung) und damit zeitliche Korrespondenz aber nicht gleichzusetzen ist mit anatomischer Korrespondenz, erfolgt im nächsten Schritt eine Angleichung der anatomischen Unterschiede der verschiedenen Patienten.

Zur Herstellung der anatomischen Korrespondenz wird auf die Referenzbilder der Population $(I_{p,1} : \Omega_p \rightarrow \mathbb{R})_{p=1}^{n_{pat}}$ des Atemzustandes $j = 1$ zurückgegriffen, auf deren Grundlage ein (mittleres) Atlasbild $\bar{I}_1 : \Omega_A \rightarrow \mathbb{R}$ erzeugt wird, welches die mittlere Anatomie (und mittlere Intensität) der Population für $j = 1$ abbildet. Hierfür wird konkret in dieser Arbeit das in [Guimond u. a., 2000] zur registrierungsbasierten, iterativen Atlasgenerierung vorgeschlagene Verfahren in Kombination mit der *Variational-Registration-Toolbox* [Schmidt-Richberg u. a., 2014; Werner u. a., 2014] zur diffeomorphen Bildregistrierung eingesetzt (siehe auch Kapitel 2.2). Die Atlasgenerierung mit dieser Methodenkombination wird ausführlich in [Ehrhardt u. a., 2011; Werner, 2012] beschrieben, weshalb hier auf weitere Erläuterungen verzichtet wird.

Durch die Erzeugung des Atlasbildes liegt nun ein Atlaskoordinatensystem Ω_A vor. In dieses Atlaskoordinatensystem werden die für die Korrespondenzmodellierung genutzten patientenspezifischen Transformationen und Surrogatsignale, welche in den jeweiligen Patientenkoordinatensystemem Ω_p definiert sind, abgebildet, um alle weiteren Berechnungen im einheitlichen Atlaskoordinatensystem durchführen zu können (siehe auch Abb. 5.2). Hierfür wird zunächst wieder mittels diffeomorpher Bildregistrierung für jeden Patienten $p \in \{0, \dots, n_{pat}\}$ eine Transformation $\psi_p : \Omega_A \rightarrow \Omega_p$ zwischen \bar{I}_1 und $I_{p,1}$ geschätzt, die die Punkte im Atlaskoordinatensystem auf die anatomisch korrespondierende Position im Patientenkoordinatensystem abbildet. Die patientenspezifischen Transformationen $(\varphi_{p,j} : \Omega_p \rightarrow \Omega_p)_{j=1}^{n_{tp}}$ im Patientenkoordinatensystem Ω_p

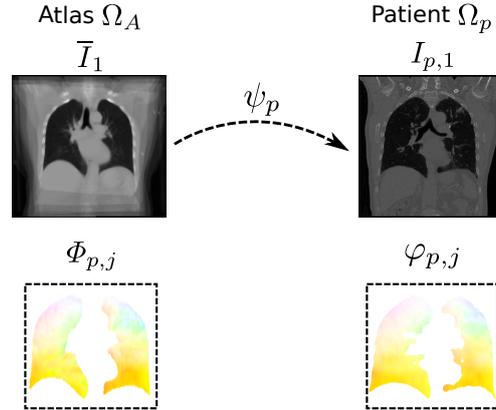


Abb. 5.2: Illustration des Zusammenhangs zwischen Atlas- und Patientenkoordinatensystem und der verwendeten Notation. Das mittlere Atlasbild \bar{I}_1 entstammt der Evaluation zu diesem Kapitel und wurde nur für die Lunge erstellt. Auch die Transformationen sind auf diesen Bereich beschränkt. Siehe Text für weitere Details.

werden dann analog zu dem Vorgehen in Kapitel 3.1.3.3 mittels Konjugation (engl. Conjugate Action; siehe [Lorenzi u. a., 2014] für eine Diskussion dieses Ansatzes) in das Atlaskoordinatensystem Ω_A transportiert:

$$\Phi_{p,j} = \psi_p^{-1} \circ \varphi_{p,j} \circ \psi_p . \quad (5.1)$$

Resultat dieses Vorgehens ist das Vorliegen von Transformationen $(\Phi_{p,j} : \Omega_A \rightarrow \Omega_A)_{j=1}^{n_{tp}}$ für jeden Patienten $p \in \{0, \dots, n_{pat}\}$, welche die jeweilige patientenspezifische Atembewegung im Atlaskoordinatensystem Ω_A beschreiben.

Durch die Abbildung der Atembewegung auf die Atlasanatomie ist auch eine entsprechende Anpassung der Surrogatsignalmessungen $(\zeta_{p,j} \in \mathbb{R}^{n_{sur}})_{j=1}^{n_{tp}}$ notwendig. Die Umsetzung dieser Anpassung ist nicht allgemeingültig beschreibbar, da sie stark von dem verwendeten Surrogat und dem Zusammenhang zwischen Referenzbild und Signalwert abhängig ist. Ist deren Zusammenhang bekannt, kann ein 1D-Surrogatsignal ($n_{sur} = 1$) möglicherweise durch einfache Skalierung an die Atlasanatomie angepasst werden. Für komplexere Surrogate wie hochdimensionale Oberflächenmessungen bietet sich u.U. die Abbildung der gemessenen Oberfläche in den Atlasraum mithilfe der Transformation ψ_p an, bevor die eigentlichen Signalwerte ermittelt werden. Ein ähnliches Vorgehen wird in den Experimenten zu diesem Kapitel verfolgt, da analog zu Kapitel 3 ausschließlich bildbasiert simulierte Surrogatsignale zum Einsatz kommen (konkret: 2D-Spirometerignal; siehe Kapitel 3.1.2.1 für die Simulation). Um diese korrekt im Atlasraum simulieren zu können, werden die 4D-Bilddaten jedes Patienten mittels ψ_p in den Atlasraum abgebildet, bevor die eigentliche Surrogatsimulation erfolgt. Im Folgenden bezeichnet $(\nu_{p,j} \in \mathbb{R}^{n_{sur}})_{j=1}^{n_{tp}}$ die für den jeweiligen Patienten $p \in \{0, \dots, n_{pat}\}$ in den Atlasraum abgebildeten Surrogatsignalmessungen.

5.2.1 Von der patientenspezifischen zur populationsbasierten Korrespondenzmodellierung im Atlasraum

Aus den beschriebenen Vorverarbeitungsschritten stehen nun die patientenspezifische Atembewegung charakterisierende diffeomorphe Transformationen $(\Phi_{p,j})_{p=1,j=1}^{n_{pat},n_{tp}}$ und zugehörige Surrogatsignalmessungen $(\nu_{p,j})_{p=1,j=1}^{n_{pat},n_{tp}}$ einer Population von n_{pat} Patienten im Atlasraum zur Verfügung. Transformationen $(\Phi_{0,j})_{j=1}^{n_{tp}}$ und Surrogatsignalmessungen $(\nu_{0,j})_{j=1}^{n_{tp}}$ sind ebenso für einen neuen Patienten $p = 0$ verfügbar, für den im Folgenden Korrespondenzmodelle erzeugt werden sollen.

Grundlage für die Korrespondenzmodellierung in diesem Kapitel bildet grundsätzlich das in Kapitel 2.3 präsentierte (diffeomorphe) Modellierungsframework, wobei auf den diffeomorphen Aspekt der Modellierung verzichtet wird. Dies bedeutet, dass wie in Kapitel 3.1.3.4 (und teilweise in den Experimenten zu Kapitel 4) eine direkte Regression zwischen Surrogatdaten und Verschiebungsfeldern erfolgt. Die mithilfe der so trainierten Korrespondenzmodelle erzeugten Transformationen sind folglich nicht mehr zwingend Elemente von $\text{Diff}(\Omega)$ (siehe Diskussion in Kapitel 2.3), sondern es wird allgemeiner angenommen, dass sie Elemente von $C^2(\Omega)$ sind¹². Die Abkehr von der diffeomorphen Korrespondenzmodellierung basierend auf stationären Geschwindigkeitsfeldern in diesem Kapitel hat, analog zu den Experimenten in Kapitel 4, ausschließlich Effizienzgründe, da für die aufwendigen Experimente zu diesem Kapitel die schnelle GPU-Implementierung der Scaling-and-Squaring-Methode noch nicht verfügbar war (siehe auch Kapitel 3.2). Die hier präsentierte Methodik lässt sich aber ohne Veränderungen auch auf diffeomorphe Korrespondenzmodelle anwenden.

Analog zum in Kapitel 2.3 beschriebenen Standardvorgehen bei der nicht-diffeomorphen Korrespondenzmodellierung seien die einzelnen nichtlinearen Transformationen $\Phi_{p,j}$ in diesem Kapitel jeweils durch Verschiebungsfelder $u_{p,j} : \Omega_A \rightarrow \mathbb{R}^3$ parametrisiert, sodass $\Phi_{p,j} = Id + u_{p,j}$ gilt. Weiterhin werden diese Verschiebungsfelder $u_{p,j}$ und die zugehörigen Surrogatsignalmessungen $\nu_{p,j}$ für die Regression als Spaltenvektoren $\mathbf{u}_{p,j} = \text{vec}(u_{p,j}) \in \mathbb{R}^{3n_{vx}}$ und $\mathbf{n}_{p,j} = \nu_{p,j} \in \mathbb{R}^{n_{sur}}$ repräsentiert. Anschließend sollen nun, wie in Kapitel 2.3, die Parameter eines linearen Modells

$$\hat{\mathbf{u}}_{\Omega_A} = \bar{\mathbf{u}} + \mathbf{B}_{\Omega_A}(\hat{\mathbf{n}} - \bar{\mathbf{n}}) \quad (5.2)$$

geschätzt werden, mit dem basierend auf einer neuen Surrogatsignalmessung $\hat{\mathbf{n}}$ für Patient $p = 0$ der Vektor $\hat{\mathbf{u}}_{\Omega_A} \in \mathbb{R}^{3n_{vx}}$ bzw. das korrespondierende Verschiebungsfeld $\hat{u}_{\Omega_A} : \Omega_A \rightarrow \mathbb{R}^3$ und damit auch die Transformation $\hat{\Phi}_{\Omega_A} = Id + \hat{u}_{\Omega_A}$ im Atlasraum Ω_A erzeugt werden können. Diese Transformation kann anschließend mittels $\hat{\Phi} = \psi_0 \circ \hat{\Phi}_{\Omega_A} \circ \psi_0^{-1}$ in das patientenspezifische Koordinatensystem Ω_0 transportiert werden. Die Matrix $\mathbf{B}_{\Omega_A} \in \mathbb{R}^{3n_{vx} \times n_{sur}}$ ist wieder die während des Trainings zu schätzende

¹² $C^2(\Omega)$ bezeichnet die Menge der (mindestens) zweimal stetig differenzierbaren Funktionen über Ω .

Die hierzu notwendige Glattheit der Transformation wird typischerweise durch den eingesetzten Regularisierer des Registrierungsverfahrens gewährleistet (siehe z. B. [Modersitzki, 2004]).

Koeffizientenmatrix, wohingegen $\bar{\mathbf{u}} \in \mathbb{R}^{3n_{vx}}$ und $\bar{\mathbf{n}} \in \mathbb{R}^{n_{sur}}$ die Mittelwertvektoren der für das Training verwendeten Verschiebungsfelder bzw. Surrogatvektoren sind.

Patientenspezifisches Modell Der Methodik um Glg. 2.13 zur Regression mit Frobeniusnorm-Regularisierung/Ridge-Regression folgend¹³, lässt sich die Koeffizientenmatrix \mathbf{B}_{Ω_A} für ein patientenspezifisches Korrespondenzmodell für Patient $p = 0$ im Atlasraum durch

$$\mathbf{B}_{\Omega_A} = \arg \min_{\mathbf{B}} \|\mathbf{U} - \mathbf{B}\mathbf{N}\|_F^2 + \lambda \|\mathbf{B}\|_F^2 = \mathbf{U}\mathbf{N}^T(\mathbf{N}\mathbf{N}^T + \lambda\mathbf{I})^{-1} = \mathbf{U}(\mathbf{N}^T\mathbf{N} + \lambda\mathbf{I})^{-1}\mathbf{N}^T \quad (5.3)$$

ermitteln. Hierbei bestehen bei einem patientenspezifischen Modell die Matrizen $\mathbf{U} = [\mathbf{u}_{0,1}^c \dots \mathbf{u}_{0,n_{tp}}^c] \in \mathbb{R}^{3n_{vx} \times n_{tp}}$ und $\mathbf{N} = [\mathbf{n}_{0,1}^c \dots \mathbf{n}_{0,n_{tp}}^c] \in \mathbb{R}^{n_{sur} \times n_{tp}}$ aus den mittelwertzentrierten Verschiebungsfeldern $\mathbf{u}_{0,j}^c = \mathbf{u}_{0,j} - \bar{\mathbf{u}}$ bzw. Surrogatsignalmessungen $\mathbf{n}_{0,j}^c = \mathbf{n}_{0,j} - \bar{\mathbf{n}}$, mit $\bar{\mathbf{u}} = \frac{1}{n_{tp}} \sum_{j=1}^{n_{tp}} \mathbf{u}_{0,j}$ und $\bar{\mathbf{n}} = \frac{1}{n_{tp}} \sum_{j=1}^{n_{tp}} \mathbf{n}_{0,j}$.

Populationsmodell Die patientenspezifische Formulierung in Glg. 5.3 lässt sich auch für die Erzeugung eines populationsbasierten Korrespondenzmodells für Patient $p = 0$ nutzen. Die einzige notwendige Änderung betrifft die Zusammensetzung der für das Modelltraining eingesetzten Matrizen \mathbf{U} und \mathbf{N} (bzw. $\bar{\mathbf{u}}$ und $\bar{\mathbf{n}}$), die nun die Bewegungsinformationen $(\mathbf{u}_{p,j}^c)_{p=0,j=1}^{n_{pat},n_{tp}}$ und Surrogatdaten $(\mathbf{n}_{p,j}^c)_{p=0,j=1}^{n_{pat},n_{tp}}$ aller $n_{pat} + 1$ Patienten enthalten:

$$\begin{aligned} \mathbf{U} &= [\mathbf{u}_{0,1}^c \dots \mathbf{u}_{0,n_{tp}}^c \dots \mathbf{u}_{p,1}^c \dots \mathbf{u}_{p,n_{tp}}^c \dots \mathbf{u}_{n_{pat},1}^c \dots \mathbf{u}_{n_{pat},n_{tp}}^c] \in \mathbb{R}^{3n_{vx} \times n_{tp}(n_{pat}+1)} \text{ und} \\ \mathbf{N} &= [\mathbf{n}_{0,1}^c \dots \mathbf{n}_{0,n_{tp}}^c \dots \mathbf{n}_{p,1}^c \dots \mathbf{n}_{p,n_{tp}}^c \dots \mathbf{n}_{n_{pat},1}^c \dots \mathbf{n}_{n_{pat},n_{tp}}^c] \in \mathbb{R}^{n_{sur} \times n_{tp}(n_{pat}+1)}. \end{aligned}$$

In der Einleitung zu diesem Kapitel wird die Nutzung von Populationsinformationen mit der hierdurch potenziell verbesserten Genauigkeit der Bewegungsschätzung bei interfraktionellen Bewegungsvariationen begründet. Basierend auf dieser Argumentation erscheint deshalb eine erste Diskussion der Auswirkungen dieser Variationen auf das bisher in diesem Abschnitt im Atlasraum definierte patientenspezifische Modell und das (allgemeine) Populationsmodell hilfreich, um die weiteren Schritte zu motivieren. Hierfür seien die Daten für Patient $p = 0$ während einer Planungssitzung aufgenommen worden und das zu erstellende Korrespondenzmodell für diesen Patienten soll während einer zeitlich getrennten Behandlungssitzung eingesetzt werden. Die Genauigkeit der Bewegungsschätzung durch das patientenspezifische Modell wird bei diesem Vorgehen u. a. stark von den Unterschieden in den Bewegungseigenschaften zwischen Planungs- und Behandlungssitzung (interfraktionelle Bewegungsvariabilität) abhängen. Bei großen Unterschieden wird die Genauigkeit entsprechend stark sinken. Bei

¹³Hier ist alternativ auch die Nutzung der in Kapitel 2.4 vorgestellten projektionsbasierten Regressionsmethoden oder anderer Regressionsverfahren möglich.

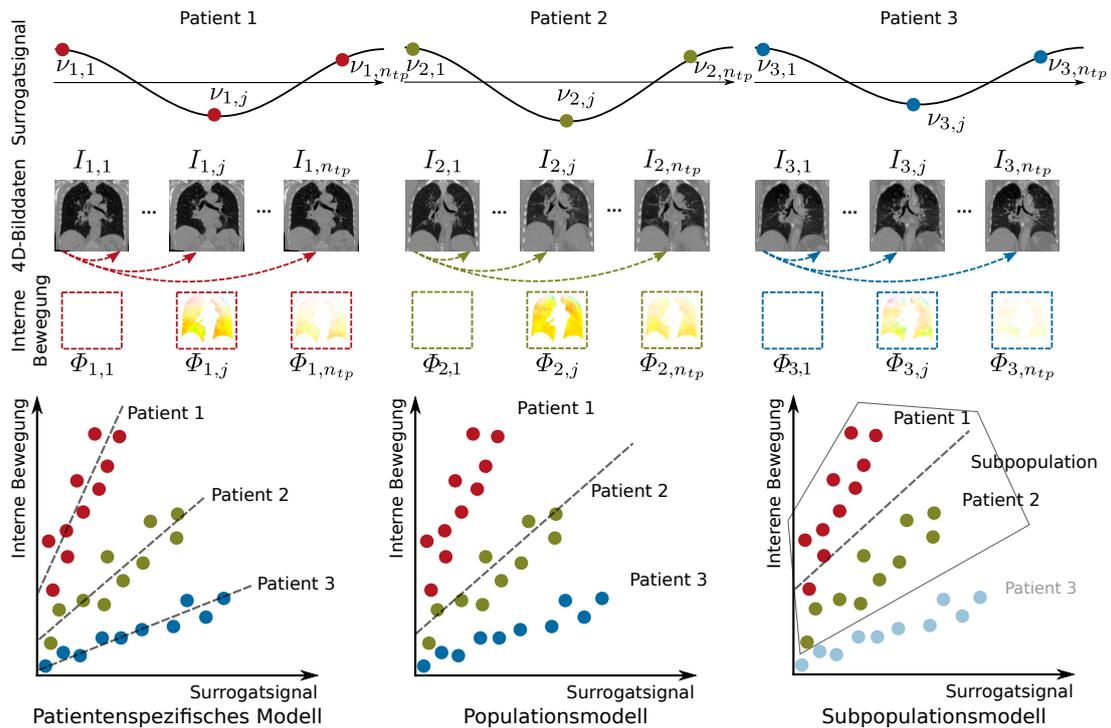


Abb. 5.3: Illustration der drei in diesem Kapitel diskutierten Korrespondenzmodellierungsverfahren (unterste Reihe) anhand von drei fiktiven Patienten. Die gestrichelten Linien zeigen jeweils einen Abschnitt der durch das lineare Korrespondenzmodell abgebildeten Regressionsgeraden in den Trainingsdaten der drei Patienten. Bei der patientenspezifischen Modellierung entstehen entsprechend drei verschiedene Modelle, wohingegen Populations- und Subpopulationsmodellierung jeweils ein Modell erzeugen, wobei sich die Modelle hinsichtlich der Lage der Regressionsgeraden unterscheiden. Abbildung adaptiert aus [Wilms u. a., 2017c].

einem Populationsmodell, bei dem die patientenspezifischen Daten mit in das Modelltraining eingeflossen sind, wird durch die Minimierung der Summe der Fehlerquadrate in Glg. 5.3 ein Mittelungseffekt erreicht, sodass die Bewegungseigenschaften der patientenspezifischen Daten je nach Populationsgröße in der Masse der Populationsdaten verschwinden bzw. zumindest verringert werden. Resultat der populationsbasierten Modellbildung ist eine Art mittleres Korrespondenzmodell (siehe auch Abb. 5.3), welches die Bewegungsschätzung gegenüber dem patientenspezifischen Modell durch die Mittelung beispielsweise dann verbessern kann, wenn die patientenspezifischen Bewegungseigenschaften in den Planungsdaten und während der Behandlungssitzung stark gegensätzlich sind, was auch bei starken Artefakten in den zur initialen Bewegungsschätzung genutzten 4D-Bilddaten gegeben sein kann.

Ausgehend von der Annahme, dass ein Populationsmodell genau dann am effektivsten ist, wenn es ungefähr die mittleren Bewegungsmuster des zu modellierenden Patienten abdeckt, da dies einen guten Kompromiss zwischen Passgenauigkeit für einzelne Behandlungssitzungen und Abdeckung der gesamten interfraktionellen Bewegungsvariabilität darstellt, stellt sich die Frage, ob die Gesamtpopulation hierfür die geeignete Datenbasis ist? In Kapitel 5.1 wird die geringere Genauigkeit von Populationsmodellen gegenüber patientenspezifischen Modellen u. a. mit der Heterogenität der Bewegungen der einzelnen Patienten (= Interpatientenvariabilität) begründet. Aufgrund dieser Heterogenität ist auch anzunehmen, dass die Interpatientenvariabilität der Gesamtpopulation nicht unbedingt der interfraktionellen Bewegungsvariabilität von Patient $p = 0$ entspricht. Es scheint deshalb wesentlich erfolgversprechender eine weniger heterogene Subpopulation $S \subset \{1, \dots, n_{pat}\}$ zu ermitteln, deren Patienten eine ähnliche Atembewegung wie Patient $p = 0$ aufweisen. Die Hoffnung ist dann, dass die Interpatientenvariabilität in dieser Subpopulation S eher der interfraktionellen Bewegungsvariabilität von Patient $p = 0$ entspricht als die der Gesamtpopulation. Somit sollten die auf dieser Datengrundlage erstellten Subpopulationsmodelle eine Verbesserung der Bewegungsschätzung gegenüber patientenspezifischen und (allgemeinen) populationsbasierten Modellen bei Interfraktionsvariabilität ermöglichen (siehe auch Abb. 5.3). Voraussetzung für die Umsetzung dieser Idee ist die Zerlegung der Gesamtpopulation in Subpopulationen ähnlich atmender Patienten, was in diesem Kapitel durch Clustering erfolgt.

5.2.2 Sparse Manifold Clustering zur Identifikation von Subpopulationen

Ziel ist es nun, per Clustering die Patientenpopulation aller $n_{pat} + 1$ Patienten (inkl. $p = 0$) in disjunkte Subpopulationen $S \subset \{0, \dots, n_{pat}\}$ ähnlich atmender Patienten zu zerlegen. Das Subpopulationsmodell für Patient $p = 0$ wird anschließend basierend auf der Subpopulation gebaut, der dieser Patient angehört. Die Wahl eines geeigneten Clustering-Verfahrens und der grundsätzliche Ablauf der Zerlegung der Population hängen u. a. wesentlich von zwei Faktoren ab:

1. Der Wahl einer geeigneten Repräsentationsform der patientenspezifischen Atembewegung, die insbesondere die charakteristischen Merkmale des Bewegungsmusters für einen Vergleich nutzbar machen sollte.
2. Der Wahl eines geeigneten Distanz- bzw. Ähnlichkeitsmaßes zur Quantifizierung der Ähnlichkeit der Bewegungsmuster zweier Patienten.

Die Wahl von Repräsentationsform und Distanzmaß kann nicht unabhängig voneinander erfolgen, da es wechselseitige Abhängigkeiten gibt. Die Literatur zur Repräsentation und Quantifizierung der Ähnlichkeit zeitlicher Entwicklungen/Bewegungen im medizinischen Bereich bewegt sich ähnlich zur Korrespondenzmodellierung hauptsächlich

zwischen zwei Extremen: (1) Die Verschiebungsfelder werden als Elemente eines Vektorraums angesehen, in dem mittels euklidischem Abstand zwischen Paaren von vektorisierten Feldern eine Ähnlichkeit ermittelt wird (siehe z. B. [Peressutti u. a., 2014]). (2) Die Repräsentation der Bewegung oder des zeitlichen Verlaufs erfolgt explizit über den Einsatz von Diffeomorphismen und entsprechend auf diese Struktur abgestimmte Maße werden zur Ähnlichkeitsbestimmung genutzt (Computational Functional Anatomy; siehe [Miller u. a., 2009] für einen Einstieg).

Die Repräsentation der patientenspezifischen Atembewegung in diesem Kapitel erfolgt ähnlich zu [Peressutti u. a., 2014] durch Vektorisierung der Verschiebungsfelder, wobei alle Felder $(\mathbf{u}_{p,j})_{j=1}^{n_{tp}}$ eines Patienten p zu einem einzelnen Vektor $\mathbf{m}_p = [\mathbf{u}_{p,1}^T, \dots, \mathbf{u}_{p,n_{tp}}^T]^T \in \mathbb{R}^d$ zusammengefasst werden. Die einzelnen Bewegungsvektoren \mathbf{m}_p sind dann Elemente des \mathbb{R}^d , mit $d = 3n_{vx}n_{tp}$, was die Ähnlichkeitsquantifikation rein über euklidische Abstände als Grundlage für ein Clustering ermöglicht, wie dies in [Peressutti u. a., 2014] erfolgt. Nachteilig bei der ausschließlichen Nutzung des euklidischen Abstandes als Distanzmaß ist dessen fehlende Skalierungsinvarianz. Da aber beispielsweise das resultierende Korrespondenzmodell Skalierungsunterschiede (unterschiedliche Ein- bzw. Ausatmungstiefe) selbstständig ausgleichen kann, wenn sie sich im Surrogatsignal widerspiegeln, sollte auch das verwendete Distanzmaß diese Eigenschaft entsprechend abbilden können. Im Gegensatz zu [Peressutti u. a., 2014], wo diesem Problem mit einer expliziten Normalisierung der Bewegungen begegnet wird, erfolgt in diesem Kapitel die Ähnlichkeitsbestimmung und das nachfolgende Clustering mit einem integrierten Ansatz (Sparse Manifold Clustering and Embedding – SMCE [Elhamifar u. a., 2011]), der eine Kombination aus euklidischer Distanz und Sparse Coding einsetzt und hierdurch u. a. das Skalierungsproblem vermeidet.

Die Grundannahme bei der Nutzung von SMCE zum Clustering der Atembewegungen ist, dass die Bewegungsvektoren \mathbf{m}_p zwar Elemente eines sehr hochdimensionalen Vektorraums \mathbb{R}^d sind, der abgebildete Atmungsprozess aber wesentlich weniger als d Freiheitsgrade besitzt und es somit niedrigdimensionale Struktur(en)¹⁴ in den Daten gibt. Weiterhin seien diese niedrigdimensionalen Strukturen in diesem Anwendungskontext Indikatoren bzw. Repräsentanten verschiedener Bewegungsmuster und damit Subpopulationen. Folglich sollten Patienten mit ähnlicher Atembewegung Elemente derselben niedrigdimensionalen Struktur sein. Das Clustering der Bewegungsdaten hat demnach rein technisch das Ziel diese niedrigdimensionalen Strukturen in den Daten zu ermitteln und die einzelnen Bewegungsvektoren der jeweiligen Struktur zuzuordnen. SMCE geht dabei von der allgemeinsten Annahme aus: Die Datenpunkte bzw. Bewegungsvektoren entstammen einer Menge von nichtlinearen Mannigfaltigkeiten von möglicherweise verschiedener, niedriger Dimension, wodurch es sich bei dem Clustering um ein sogenanntes Manifold Clustering handelt. Weiterhin beinhaltet diese Annahme

¹⁴Siehe auch [Li u. a., 2011] für eine Diskussion niedrigdimensionaler Strukturen im Kontext der Atembewegung.

die häufig zur Atmungsmodellierung genutzten linearen Unterräume (siehe z. B. [Li u. a., 2011]) als Spezialfall.

Grundlegende Eigenschaft einer n -dimensionalen Mannigfaltigkeit ist, dass sie lokal homöomorph zum \mathbb{R}^n ist [Lee, 2011]. Diesen Umstand nutzt SMCE zur Definition eines Ähnlichkeitsgraphen für das Clustering aus: Es wird für jeden Bewegungsvektor \mathbf{m}_p mittels euklidischer Distanz eine Menge von $n + 1$ Nachbarvektoren $\{\mathbf{m}_i \in \mathbb{R}^d\}_{i \neq p}$ gesucht, die approximativ einen n -dimensionalen affinen Unterraum bilden, der \mathbf{m}_p enthält (siehe auch Abb. 5.4). Bei entsprechender Datendichte entspricht dieser dem Tangentialraum am Punkt \mathbf{m}_p [Elhamifar u. a., 2011]. Um kein explizites n für die Nachbarschaftgröße festlegen zu müssen, wird für jeden Bewegungsvektor \mathbf{m}_p ein sparses Optimierungsproblem definiert [Elhamifar u. a., 2011]:

$$\min_{\mathbf{c}_p} \gamma \|\mathbf{Q}_p \mathbf{c}_p\|_1 + \frac{1}{2} \|\mathbf{M}_p \mathbf{c}_p\|_2^2 \quad \text{s.t.} \quad \mathbf{1}^T \mathbf{c}_p = 1. \quad (5.4)$$

Nach der Minimierung von Glg. 5.4 geben die Indizes der von Null verschiedenen Elemente des Vektors $\mathbf{c}_p \in \mathbb{R}^{n_{pat}}$ die durch das Verfahren ausgewählten (wenigen) Nachbarn von \mathbf{m}_p an, die optimalerweise auf derselben Mannigfaltigkeit liegen. Hierbei ist γ für die Gewichtung der beiden wesentlichen Bestandteile des Optimierungsproblems verantwortlich ($\|\mathbf{Q}_p \mathbf{c}_p\|_1$: Auswahl weniger, naher Nachbarn; $\|\mathbf{M}_p \mathbf{c}_p\|_2^2$: Bestmögliche Konstruktion des affinen Unterraums). Zur Wahl eher naher Nachbarn werden die Elemente der Diagonalmatrix $\mathbf{Q}_p \in \mathbb{R}^{n_{pat} \times n_{pat}}$ mit einer invers gewichteten euklidischen Distanz $\frac{\|\mathbf{m}_i - \mathbf{m}_p\|_2}{\sum_{t \neq p} \|\mathbf{m}_t - \mathbf{m}_p\|_2} \in (0, 1]$ zwischen \mathbf{m}_p und den anderen n_{pat} Bewegungsvektoren besetzt und die Matrix $\mathbf{M}_p \in \mathbb{R}^{d \times n_{pat}}$ besteht aus den normierten Richtungsvektoren zwischen \mathbf{m}_p und den restlichen Bewegungsvektoren [Elhamifar u. a., 2011]:

$$\mathbf{M}_p \triangleq \begin{bmatrix} \frac{\mathbf{m}_0 - \mathbf{m}_p}{\|\mathbf{m}_0 - \mathbf{m}_p\|_2} & \cdots & \frac{\mathbf{m}_{n_{pat}} - \mathbf{m}_p}{\|\mathbf{m}_{n_{pat}} - \mathbf{m}_p\|_2} \end{bmatrix}. \quad (5.5)$$

Da durch die Konstruktion des Optimierungsproblems in Glg. 5.4 die Affinkombination $\mathbf{M}_p \mathbf{c}_p \approx 0$ sein sollte, lässt sich im Umkehrschluss \mathbf{m}_p durch die Linearkombination seiner Nachbarn approximieren: $\mathbf{m}_p \approx [\mathbf{m}_0 \dots \mathbf{m}_{n_{pat}+1}] \mathbf{w}_p$. Der Gewichtungsvektor $\mathbf{w}_p \in \mathbb{R}^{n_{pat}+1}$ wird hierfür aus \mathbf{c}_p abgeleitet (siehe [Elhamifar u. a., 2011] für die exakte Herleitung) und seine Elemente geben die inverse Distanz zwischen \mathbf{m}_p und diesen Nachbarn auf der Mannigfaltigkeit an. Die Distanz zu allen anderen Bewegungsvektoren und \mathbf{m}_p selbst ist dementsprechend 0. Nachfolgend werden diese Gewichtungsvektoren zu einer symmetrischen Matrix

$$\mathbf{W} = \max([\mathbf{w}_0] \dots [\mathbf{w}_{n_{pat}}], [\mathbf{w}_0] \dots [\mathbf{w}_{n_{pat}}]^T) \in \mathbb{R}^{(n_{pat}+1) \times (n_{pat}+1)} \quad (5.6)$$

konkateniert, wobei $[\mathbf{w}_0]$ den Vektor \mathbf{w}_0 nach elementweiser Betragsbildung bezeichnet. Diese Matrix \mathbf{W} dient dann als Ähnlichkeitsmatrix eines Ähnlichkeitsgraphen zwischen

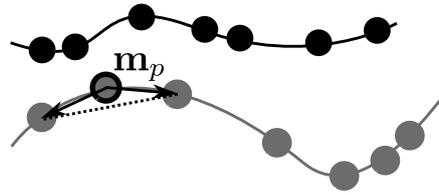


Abb. 5.4: Illustration des Ziels und des Vorgehens beim durchgeführten Manifold Clustering (SMCE). Die Abbildung zeigt Ausschnitte zweier glatter, eindimensionaler Mannigfaltigkeiten im \mathbb{R}^2 (graue und schwarze Kurve) und durch Kreise markierte diskrete Elemente dieser Strukturen (= Bewegungsvektoren), die jeweils im zu erstellenden Ähnlichkeitsgraphen ein einzelnes Cluster bilden sollen. Der schwarz umrandete Punkt der grauen Mannigfaltigkeit zeigt beispielhaft, wie bei SMCE jeder Bewegungspunkt \mathbf{m}_p mithilfe seiner Nachbarn auf der gleichen Struktur bestmöglich approximiert wird (orthogonale Projektion auf die gestrichelte Linie). Siehe Text für weitere Details.

den $n_{pat} + 1$ Bewegungsvektoren, der mittels Spectral Clustering (hier: Laplace-Matrix nach [Ng u. a., 2002] und k -means Clustering) in k Cluster zerlegt wird.

Die Anzahl der Zusammenhangskomponenten des Graphen sollte durch die Konstruktion von \mathbf{W} (zumindest theoretisch) der Anzahl der verschiedenen Mannigfaltigkeiten entsprechen, wobei jede Mannigfaltigkeit eine Subpopulation für die Korrespondenzmodellierung bildet. Parameter k ist neben γ der zweite und wichtigere der zu wählenden Parameter des Verfahrens. In der Praxis sollte sich die Wahl von k an der erwarteten Anzahl von Mannigfaltigkeiten orientieren, die u.U. auch durch Analyse der Eigenwerte der für das Clustering erstellten Laplace-Matrix automatisiert bestimmt werden kann. Generell gilt hierbei, dass je größer k wird, die durchschnittliche Subpopulationsgröße sinkt und bei $k = 1$ der Spezialfall des allgemeinen Populationsmodells vorliegt.

5.3 Experimente

Die Evaluation des in diesem Kapitel entwickelten subpopulationsbasierten Korrespondenzmodellierungsverfahrens versucht zwei essenzielle Fragen zu beantworten:

1. Kann bei Vorliegen von interfraktioneller Bewegungsvariabilität mithilfe von Subpopulationsmodellen wirklich eine höhere Genauigkeit der Bewegungsschätzung erreicht werden, als dies durch Nutzung patientenspezifischer oder allgemeiner Populationsmodelle möglich ist?
2. Lässt sich der für das Clustering entscheidende Parameter k (= Anzahl der Cluster; beeinflusst u. a. die Subpopulationsgröße) automatisiert, patienten- und fraktionsspezifisch so ermitteln, dass bezogen auf die Genauigkeit ein optimales Subpopulationsmodell entsteht?

Zur Klärung dieser Fragen kommen hauptsächlich Daten zweier verschiedener Patientengruppen (Gruppe 1 + Gruppe 2) mit insgesamt 24 Lungentumorpatienten zum Einsatz, für die jeweils im Rahmen von klinischen Studien zwei 4D-CT-Bilddatensätze in zeitlichem Abstand aufgenommen wurden (siehe Kapitel 5.3.1). Diese 4D-Bilddaten von zwei Aufnahmezeitpunkten ermöglichen in der Evaluation eine realistische Untersuchung der Effekte von interfraktioneller Bewegungsvariabilität, da jeweils ein Zeitpunkt für die Korrespondenzmodellbildung und die Modellapplikation verwendet werden kann (siehe 5.3.1).

5.3.1 Verwendete Bilddaten

Insgesamt werden für die Evaluation 73 thorakale 4D-CT-Bilddatensätze von (hauptsächlich) Lungentumorpatienten verwendet, welche sich in drei verschiedene Gruppen aufteilen:

- **Gruppe 1:** Diese Gruppe umfasst 28 thorakale 4D-CT-Bilddatensätze von 14 Lungentumorpatienten, die nicht frei verfügbar sind und im Rahmen einer Ventilationsstudie am Department of Radiation Oncology der Stanford University School of Medicine, USA aufgenommen wurden. Die Daten wurden dem Institut für Medizinische Informatik der Universität zu Lübeck im Rahmen einer Kooperation zur Verfügung gestellt. Für jeden der 14 Patienten stehen insgesamt zwei 4D-CT-Datensätze zur Verfügung (Aufnahme 1 und Aufnahme 2), wobei jeweils einer vor der Bestrahlungsbehandlung des Patienten aufgenommen wurde. Die zweite Aufnahme wurde entweder ebenfalls vor der Bestrahlungsbehandlung (fünf Patienten), während der mehrwöchigen Behandlung (acht Patienten) oder nach der letzten Bestrahlungsbehandlung (ein Patient) aufgenommen. Für alle Patienten gilt, dass die Datenaufnahme für Aufnahme 1 und Aufnahme 2 an unterschiedlichen Tagen erfolgte. Jeder 4D-Datensatz besteht aus zehn Bildern/Atemphasen, die auf eine einheitliche räumlichen Auflösung von $1 \times 1 \times 2.5 \text{ mm}^3$ skaliert wurden. Die Daten wurden mittels Step-&-Shoot-Technik aufgenommen und anhand eines Signals des Real-time Position Management (RPM)-System von Varian Medical Systems rekonstruiert (siehe [Yamamoto u. a., 2012] für weitere Informationen).
- **Gruppe 2:** Diese Gruppe umfasst 20 thorakale 4D-CT-Bilddatensätze von zehn Lungentumorpatienten, die ebenfalls nicht frei verfügbar sind und am Department of Radiation Oncology der Stanford University School of Medicine, USA im Rahmen einer Studie zur Untersuchung der Atmung bei audiovisuellem Feedback aufgenommen wurden. Die technischen Daten der 4D-CT-Bilddatensätze und die Datenverfügbarkeit pro Patient entsprechen denen von Gruppe 1. Einziger Unterschied ist, dass die Aufnahme 1 und Aufnahme 2 bei dieser Gruppe für jeden Patienten während der gleichen Sitzung aufgenommen wurden. Bei Aufnahme 1 erfolgte dies jedoch unter Einsatz des audiovisuellen Feedbacksystems, wohingegen der Patient bei Aufnahme

2 frei atmen durfte. Dies führt dazu, dass sich die Bewegungen in den Aufnahmen teilweise stark unterscheiden und somit für die Evaluation zumindest eine Ähnlichkeit der Variabilität zu realen interfraktionellen Bewegungsvariationen unterstellt werden kann.

- **Weitere Populationspatienten:** Um für die Evaluation eine möglichst große Patientenpopulation verfügbar zu haben, werden neben den 48 Datensätzen/24 Patienten aus Gruppe 1 und Gruppe 2 zusätzlich noch 15 Datensätze/Patienten der WUSTL-Daten und zehn Datensätze/Patienten der DIRLAB-Daten genutzt. Diese Datenquellen sind in Kapitel 3.1.1 ausführlich beschrieben. Auch diese Daten wurden auf eine einheitliche räumlichen Auflösung von $1 \times 1 \times 2.5 \text{ mm}^3$ skaliert. Bei den WUSTL-Daten wurden bei den Patienten mit 14 verfügbaren Atemphasen zehn ausgewählt, die ungefähr den Atemzuständen der Daten von Gruppe 1/2 entsprechen, um eine zeitliche Korrespondenz der Daten zu gewährleisten. Für die DIRLAB-Daten ist diese zeitliche Korrespondenz durch ein ähnliches Rekonstruktionsverfahren gegeben.

5.3.2 Design der Experimente

Da das vorrangige Ziel dieser Evaluation die Analyse der Auswirkungen interfraktioneller Bewegungsvariabilität auf die Korrespondenzmodellierung ist, erfolgt die quantitative Auswertung der getesteten Korrespondenzmodelle alleinig anhand der Daten aus Gruppe 1 und Gruppe 2. Die WUSTL- und DIRLAB-Daten dienen nur als Populationspatienten. Weiterhin beschränken sich alle Registrierungen/Bewegungsschätzungen analog zu Kapitel 3 auf die Lungenregion. Dies gilt ebenso für das Atlasbild, welches zur Abbildung aller Patienten in den Atlasraum Ω_A genutzt wird, und für das die WUSTL-Daten die Grundlage bilden. Als Referenzphase für die Registrierung wird wie in Kapitel 3 die Phase maximaler Einatmung (EI) genutzt. Alle Korrespondenzmodelle werden im Atlasraum gebaut und als Surrogatsignal kommt ein bildbasiert simuliertes 2D-Spirometersignal ($n_{sur} = 2$; 1D-Signal + zeitliche Ableitung; siehe Kapitel 3.1.2.1) zum Einsatz. Alle Korrespondenzmodelle basieren auf der Methodik in Glg. 5.3, wobei $\lambda = 0$ (= keine Regularisierung) gewählt wird, was sich in Vortests als ausreichend gezeigt hat.

Für jeden der 48 4D-CT-Bilddatensätze von Gruppe 1 und Gruppe 2 werden insgesamt vier Korrespondenzmodellierungsansätze getestet:

1. **Patientenspezifisches Intrafraktionsmodell:** Bei diesem Ansatz werden analog zu dem Vorgehen in Experiment 1 in Kapitel 3.1.3.1 drei einzelne patientenspezifische Korrespondenzmodelle (siehe Kapitel 5.2.1) auf Basis des einen 4D-CT-Datensatzes gebaut, um die Bewegungen zwischen EI→MI, EI→ME und EI→EE in diesem Datensatz durch das jeweilige Korrespondenzmodell schätzen zu lassen. Dieses Modell/diese Modelle basieren demnach auf Bewegungsdaten, die in derselben

Sitzung aufgenommen wurden, wie die Bewegungen, die geschätzt werden sollen. Da damit keine interfraktionelle Bewegungsvariabilität auftritt, werden diese Modelle im weiteren Verlauf als (patientenspezifische) Intrafraktionsmodelle bezeichnet. Die durch diese Modelle erreichte Genauigkeit bei der modellbasierten Bewegungsschätzung bildet (im Normalfall) eine untere Schranke für die nachfolgenden Modelle.

2. **Patientenspezifisches Interfraktionsmodell:** Bei diesem Ansatz wird jeweils auf der Basis der einen Aufnahme (Aufnahme 1 bzw. Aufnahme 2) des Patienten ein patientenspezifisches Korrespondenzmodell (siehe Kapitel 5.2.1) erzeugt, welches dann genutzt wird, um die Bewegungen zwischen $EI \rightarrow MI$, $EI \rightarrow ME$ und $EI \rightarrow EE$ in der anderen Aufnahme (Aufnahme 2 bzw. Aufnahme 1) modellbasiert zu schätzen. Dieses Vorgehen simuliert das in der Einleitung zu diesem Kapitel erwähnte Standardszenario, wo Planungsdaten für das Lernen des für die intrainterventionelle Bewegungsschätzung eingesetzten Korrespondenzmodells genutzt werden. Es ist zu erwarten, dass diese Modelle durch die interfraktionelle Bewegungsvariabilität schlechtere Ergebnisse liefern als die Intrafraktionsmodelle.
3. **Populationsmodell:** Bei diesem Ansatz wird für jede Aufnahme eines Patienten ein (allgemeines) Populationsmodell (siehe Kapitel 5.2.1) erzeugt, um die die Bewegungen zwischen $EI \rightarrow MI$, $EI \rightarrow ME$ und $EI \rightarrow EE$ in diesem Datensatz modellbasiert zu schätzen. Die Population enthält in diesem Fall die andere Aufnahme des Patienten, alle anderen Aufnahmen der jeweils anderen Gruppe (Gruppe 1 oder Gruppe 2) und die zusätzlichen Populationsdaten (WUSTL und DIRLAB). Die Daten der eigenen Gruppe werden nicht für die Population verwendet, um einen möglichen Biaseffekt durch gruppenspezifische Eigenschaften, die u. a. durch die gleiche Datenaufnahme vorliegen könnten, bei der Auswertung auszuschließen. Insgesamt ergibt sich somit eine Populationsgröße von 45 für Gruppe 1 bzw. 61 für Gruppe 2.
4. **Subpopulationsmodell:** Bei diesem Ansatz wird für jede Aufnahme eines Patienten ein Subpopulationsmodell nach Kapitel 5.2.2 erstellt. Die Population wird hierfür wie bei dem Populationsmodell aufgebaut. Der Vektor \mathbf{m}_0 , dessen Cluster für das Training des Subpopulationsmodells genutzt wird, wird aus den Daten der anderen Aufnahme des Patienten erstellt. Generell beinhalten die Bewegungsvektoren \mathbf{m}_p für diese Evaluation abweichend von der Beschreibung in Kapitel 5.2.2 nur die Bewegungen zwischen $EI \rightarrow MI$, $EI \rightarrow ME$ und $EI \rightarrow EE$. Der Gewichtungssparameter in Glg. 5.4 wird nach Vortests als $\gamma = 10$ gewählt, wobei er sich ähnlich zu den Erfahrungen in [Elhamifar u. a., 2011] als nicht sehr sensitiv erwiesen hat. Da sich in Vortests die automatisierte Bestimmung von k basierend auf dem Eigenspektrum der Laplace-Matrix als unpraktikabel erwiesen hat, werden für diese Evaluation zunächst Subpopulationsmodelle für alle möglichen Werte von k erzeugt und ausgewertet, um anschließend sowohl den jeweils optimalen Wert und Möglichkeiten der automatisierten Wahl zu analysieren.

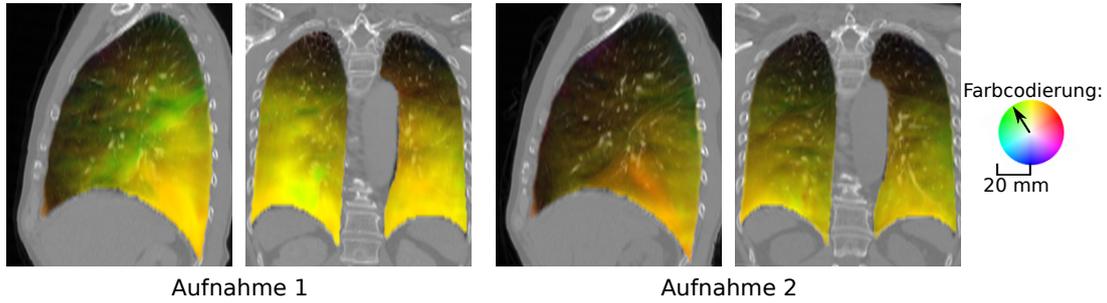


Abb. 5.5: Farbcodierte Darstellung der durch Registrierung ermittelten Atembewegung $EI \rightarrow EE$ in einer sagittalen und koronaren Schicht eines Patienten aus Gruppe 1 für Aufnahme 1 (links) und Aufnahme 2 (rechts). Die Unterschiede in Farbgebung und Intensität illustrieren die interfraktionellen Bewegungsdifferenzen. Abbildung modifiziert aus [Wilms u. a., 2017c].

Die Auswertung der Korrespondenzmodelle findet aus Gründen der Laufzeiteffizienz ausschließlich im Atlasraum Ω_A statt. Dies bedeutet, dass auf den Rücktransport der modellbasiert geschätzten Transformationen $\hat{\Phi}_{\Omega_A}$ in das jeweilige Patientenkoordinatensystem Ω_p verzichtet wird. Da anders als in Kapitel 3 für die meisten der hier verwendeten 4D-CT-Bilddaten keine korrespondierenden Landmarken zur Verfügung stehen, kann die Quantifizierung der modellbasierten Schätzungsgenauigkeit in diesem Kapitel nicht anhand des TRE erfolgen. Als Alternative hierzu werden die durch das Registrierungsverfahren geschätzten und in den Atlasraum transportierten Transformationen $\Phi_{p,j}$ bzw. die zugehörigen Verschiebungsfelder $u_{p,j}$ als (Pseudo-)Grundwahrheit für die reale Atembewegung genutzt. Anschließend wird dann voxelweise der Betrag der Abweichung zwischen modellbasiert geschätztem Verschiebungsfeld $\hat{u}_{\Omega_A} : \Omega_A \rightarrow \mathbb{R}^3$ und des durch das Registrierungsverfahren ermittelten Verschiebungsfeldes $u : \Omega_A \rightarrow \mathbb{R}^3$ durch $\|\hat{u}_{\Omega_A}(\mathbf{x}) - u(\mathbf{x})\|_2$ berechnet. Dieser Wert wird in der Auswertung gemittelt über alle Lungenvoxel angegeben und als mittlerer Fehler der Bewegungsschätzung bezeichnet. Dieses Vorgehen ist durch den durch das Registrierungsverfahren erzeugten Bias als nicht optimal zu bewerten, wird in der Literatur aber häufiger eingesetzt, wenn keine geeigneteren Objekte als Grundwahrheit verfügbar sind (siehe z. B. [Denis de Senneville u. a., 2015; Glitzner u. a., 2015]).

5.4 Ergebnisse

Die quantitativen Ergebnisse der durchgeführten Experimente sind kompakt in Tab. 5.1 zusammengefasst. Diese Tabelle beinhaltet auch die jeweils in den Datensätzen auftretende mittlere Atembewegung (= durch bildbasierte Registrierung geschätzte Bewegung) und Angaben bezüglich der interfraktionellen Bewegungsdifferenz zwischen Aufnahme 1 und Aufnahme 2 für die Patienten von Gruppe 1 und Gruppe 2. Die-

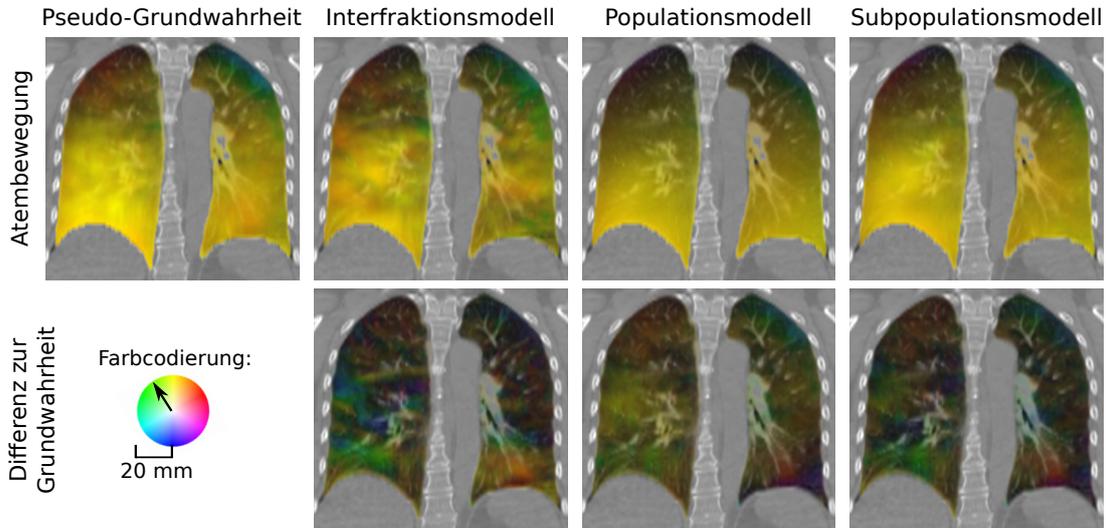


Abb. 5.6: Farbcodierte Darstellung der Unterschiede in der Bewegungsschätzung der einzelnen Modelle für einen Beispielpatienten aus Gruppe 1. Obere Reihe: Atembewegung; Untere Reihe: Differenzvektor zur Pseudo-Grundwahrheit. Abbildung modifiziert aus [Wilms u. a., 2017c].

ser Wert wird im weiteren Verlauf dieses Kapitels als Indikator für das Vorliegen von interfraktioneller Bewegungsvervariabilität verwendet und berechnet sich über die registrierungsbasiert geschätzten Verschiebungsfelder $u_j^1 : \Omega_A \rightarrow \mathbb{R}^3$ und $u_j^2 : \Omega_A \rightarrow \mathbb{R}^3$ von Aufnahme 1 bzw. Aufnahme 2 für jede Atemphase/jeden Zeitpunkt j analog zum mittleren Fehler der Bewegungsschätzung in Kapitel 5.3.2 als $\|u_j^1(\mathbf{x}) - u_j^2(\mathbf{x})\|_2$. Auch diese Größe wird für jede Aufnahme über alle Lungenvoxel gemittelt angegeben. Abbildung 5.5 illustriert die interfraktionellen Bewegungsunterschiede für einen Beispielpatienten aus Gruppe 1. Um die statistische Signifikanz der Unterschiede zwischen den einzelnen mittleren Ergebnissen in Tab. 5.1 zu beurteilen, werden gepaarte t-Tests (zweiseitig; Signifikanzniveau: 5%) auf Basis der aufnahmespezifischen Einzelergebnisse durchgeführt. Als Ergänzung zu den numerischen Resultaten in Tab. 5.1 illustriert Abb. 5.6 die Unterschiede in der Bewegungsschätzung der einzelnen Modelle für einen Beispielpatienten aus Gruppe 1.

Die Ergebnisse in Tab. 5.1 zeigen, dass die Intrafraktionsmodelle, wie angenommen, für beide Gruppen und über alle Zeitpunkte hinweg die mit Abstand geringsten Fehler produzieren. Alle Unterschiede zu den anderen getesteten Modellen sind statistisch signifikant (alle $p < 0.01$). Die Unterschiede zwischen den Interfraktionsmodellen und den (allgemeinen) Populationsmodellen sind inkonsistent. Während der Unterschied dieser Modelle bezüglich der Mittelwerte über alle Zeitpunkte für Gruppe 1 nicht signifikant ist ($p = 0.30$), zeigt sich für Gruppe 2 hier eine signifikant bessere Bewegungsschätzung durch die Interfraktionsmodelle ($p < 0.05$). Diese Feststellung

kollidiert prinzipiell mit der Annahme aus Kapitel 5.1, dass patientenspezifische Modelle in der Regel besser als Populationsmodelle sind. Bei einer genaueren Analyse der Ergebnisse dieser beiden Modelltypen zeigt sich, dass eine starke (positive) Korrelation ρ zwischen interfraktioneller Bewegungsdifferenz und dem Fehler der Bewegungsschätzung für das Interfraktionsmodell existiert (Gruppe 1: $\rho = 0.90$ Gruppe 2: $\rho = 0.67$; beides ermittelt durch partielle Korrelation¹⁵ nach Pearson). Dieser Zusammenhang besteht bei den Populationsmodellen nicht bzw. in einem sehr viel geringeren Ausmaß (Gruppe 1: $\rho = 0.54$ Gruppe 2: $\rho = 0.16$; beides ermittelt durch partielle Korrelation nach Pearson). Allerdings besteht eine zum Teil starke Korrelation zwischen der durch das Populationsmodell erreichbaren Verbesserung der Schätzungsgenauigkeit gegenüber dem Interfraktionsmodell und der interfraktionellen Bewegungsdifferenz (Gruppe 1: $\rho = 0.85$ Gruppe 2: $\rho = 0.58$; beides ermittelt durch partielle Korrelation nach Pearson). Diese Ergebnisse zeigen zusammen interpretiert, dass die Genauigkeit der Interfraktionsmodelle erwartungsgemäß bei höherer interfraktioneller Bewegungsvariabilität abnimmt und in diesen Fällen Populationsmodelle in der Lage sind, eine verbesserte Bewegungsschätzung zu liefern.

Für die Subpopulationsmodelle mit optimaler Wahl von k für jeden Datensatz durch retrospektive Analyse aller Ergebnisse zeigt sich, dass diese Modelle eine signifikant höhere Genauigkeit erzielen als die Interfraktions- und Populationsmodelle (Mittelwerte und einzelne Zeitpunkte; alle $p < 0.01$). Auch für die optimalen Subpopulationsmodelle existiert erwartungsgemäß eine teilweise starke Korrelation zwischen der durch die Subpopulationsmodelle erreichten Verbesserung gegenüber den Interfraktionsmodellen und der interfraktionellen Bewegungsdifferenz (Gruppe 1: $\rho = 0.84$ Gruppe 2: $\rho = 0.53$; beides ermittelt durch partielle Korrelation nach Pearson). Die im Mittel genutzte Anzahl von Subpopulationspatienten n_{sub} unterscheidet sich, trotz sehr ähnlicher mittlerer interfraktioneller Bewegungsdifferenz, sehr stark zwischen beiden Gruppen (Gruppe 1: 7.36 ± 12.33 Gruppe 2: 2.30 ± 1.49).

Nachdem durch die vorherigen Ergebnisse gezeigt wurde, dass zumindest durch mit optimalem k erzeugte Subpopulationsmodelle bei Vorliegen von interfraktioneller Bewegungsvariabilität tatsächlich eine höhere Genauigkeit der Bewegungsschätzung erreicht werden kann als durch Nutzung patientenspezifischer oder allgemeiner Populationsmodelle, stellt sich nun die Frage, ob der Parameter k auch automatisiert, datengetrieben ermittelt werden kann? Hierfür ist es sinnvoll zunächst noch einmal den Zusammenhang zwischen Interfraktions-, Populations- und Subpopulationsmodellen für verschiedene k zu verdeutlichen: Wird $k = 1$ gewählt, so entspricht das Subpopulationsmodell dem Populationsmodell. Wird k so gewählt, dass sich $n_{sub} = 1$ ergibt,

¹⁵Die partielle Korrelation wurde ermittelt über die aufnahmespezifischen Mittelwerte über alle Zeitpunkte. Partielle Korrelation wurde eingesetzt, um den Einfluss der Bewegungsamplitude, repräsentiert durch die mittlere Bewegung, und die datensatzspezifischen Einflüsse durch beispielsweise Artefakte auf die Bewegungsschätzung, repräsentiert durch die Ergebnisse der Intrafraktionsmodelle, herausrechnen zu können.

entspricht das Subpopulationsmodell dem Interfraktionsmodell. Gleichzeitig besteht ein Zusammenhang zwischen interfraktioneller Bewegungsdifferenz und Verbesserung durch das Populationsmodell gegenüber dem Interfraktionsmodell. Hieraus lässt sich ableiten, dass k u. a. so zu wählen ist, dass die für die interfraktionelle Bewegungsdifferenz angemessene Anzahl von Subpopulationspatienten n_{sub} vorhanden ist. Hierfür wird in dieser Arbeit angenommen, dass es einen direkten Zusammenhang zwischen interfraktioneller Bewegungsdifferenz gemittelt über alle Zeitpunkte d_{frak} und n_{sub} gibt, welcher sich durch ein Polynom vom Grad e als $n_{sub} = a(d_{frak} - b)^e$ approximieren lässt (siehe auch Abb. 5.7). In diesem Kapitel werden hierfür testweise die Polynomgrade von $e = 1$ bis $e = 5$ ausprobiert, wobei sich für $e = 2$ bis $e = 4$ die besten Ergebnisse ergeben haben, die auch in Tab. 5.1 zu finden sind. Nachdem auf Basis des patientenspezifischen Wertes für d_{frak} ein n_{sub} ermittelt wurde, wird für k der kleinste Wert gewählt, bei dem die Clustergröße kleiner oder gleich n_{sub} ist. Als Trainingsdaten für die Schätzung der Parameter des Polynoms (a und b) dienen die Werte der optimalen Subpopulationsmodelle für beide Gruppen, wobei die eigentliche Parameterschätzung durch Minimierung der Fehlerquadrate erfolgt.

Die Ergebnisse für die Subpopulationsmodelle mit automatisierter Wahl von k in Tab. 5.1 und Abb. 5.7 deuten zwar an, dass es potenziell einen Zusammenhang zwischen interfraktioneller Bewegungsdifferenz und n_{sub} gibt, die Polynome aber nur sehr rudimentär in der Lage sind, diesen abzubilden. Für alle drei in Tab. 5.1 abgebildeten Polynomgrade gilt, dass es zwischen Subpopulationsmodellen mit automatisierter k -Wahl und Interfraktionsmodellen für beide Gruppen keine signifikanten Unterschiede gibt (z. B. $e = 3$; Gruppe 1: $p = 0.18$ Gruppe 2: $p = 0.44$). Die optimalen Subpopulationsmodelle sind allerdings signifikant besser als die mittels automatisierter k -Wahl erzeugten Modelle (alle $p < 0.01$), wohingegen diese signifikant besser als die Populationsmodelle sind (alle $p < 0.01$). Diese Ergebnisse zeigen, dass die rudimentäre automatisierte k -Wahl mittels Polynomen überdacht werden sollte, da so keine Verbesserung gegenüber Interfraktionsmodellen erreicht werden kann.

5.5 Zusammenfassung und Einordnung der Ergebnisse

Bei der klassischen patientenspezifischen Korrespondenzmodellierung für die intrainterventionelle Schätzung der Atembewegung wird das Korrespondenzmodell oftmals anhand von Planungsdaten erzeugt und dann mit zeitlichem Abstand während der eigentlichen Behandlungssitzungen eingesetzt. Dabei hängt die intrainterventionelle Genauigkeit der Bewegungsschätzung ganz entscheidend von der interfraktionellen Bewegungsvariabilität ab. Gibt es große Unterschiede in der Atembewegung zwischen Planungsdaten und Behandlung, wird sich dies negativ auf die Genauigkeit des patientenspezifischen Korrespondenzmodells auswirken. In diesem Kapitel wurde versucht dieses Problem durch die Einführung sogenannter Subpopulationsmodelle zu lösen.

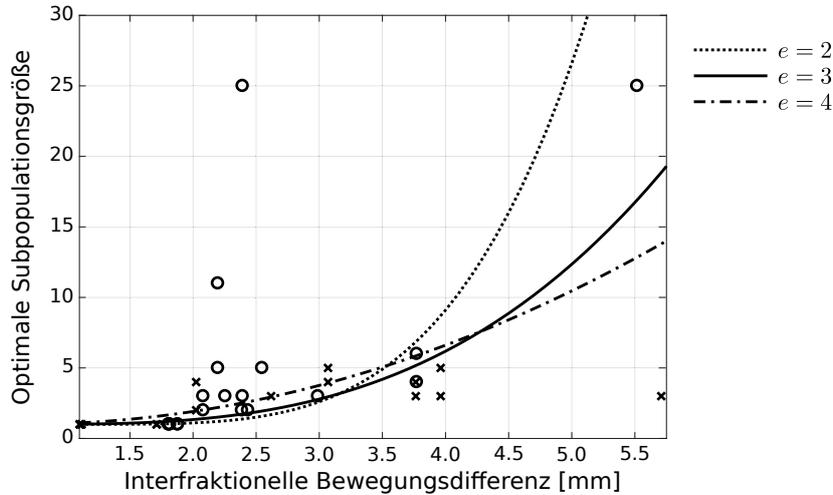


Abb. 5.7: Ausschnittweiser Plot von interfraktioneller Bewegungsdifferenz d_{frak} gegen optimale Subpopulationsgröße n_{sub} für die Patienten aus Gruppe 1 (Kreis) und Gruppe 2 (Kreuz). Abgebildet sind zusätzlich drei der basierend auf diesen Daten für die automatisierte k -Wahl ermittelten Funktionen ($n_{sub} = a(d_{frak} - b)^e$) mit $e = 2 - 4$. Abbildung angelehnt an [Wilms u. a., 2017c].

Grundlage dieser Subpopulationsmodelle ist die Hypothese, dass es in einer großen Patientenpopulation Subpopulationen von Patienten mit ähnlicher Atembewegung gibt, deren Interpatientenvariabilität ungefähr der interfraktionellen Bewegungsvariabilität eines der enthaltenen Patienten entspricht. Weiterhin wird dann angenommen, dass ein Korrespondenzmodell, welches auf Grundlage dieser Subpopulation erzeugt wurde, interfraktionelle Bewegungsvariationen besser ausgleichen kann als ein reines patientenspezifisches Modell.

Um diese Hypothesen zu testen, wurde ein neues Verfahren zur Erstellung von Subpopulationsmodellen in das in Kapitel 2 vorgestellte Korrespondenzmodellierungsframework integriert. Subpopulationen mit ähnlich atmenden Patienten werden bei diesem Verfahren durch ein Ähnlichkeits-Clustering der Bewegungen einer Patientenpopulation identifiziert. Für das Clustering kommt ein Manifold Clustering-Verfahren zum Einsatz, welches die Ähnlichkeit von Atemmustern u. a. mithilfe von Sparse Coding und unter Berücksichtigung niedrigdimensionaler Strukturen in den Daten ermittelt.

In einer initialen Evaluation des entwickelten Verfahrens konnte anhand von zwei Gruppen von Patienten mit wiederholt aufgenommenen 4D-CT-Bilddatensätzen die initiale Annahme bestätigt werden: Korrespondenzmodelle, die auf der Basis einer Subpopulation von ähnlich atmenden Patienten erstellt wurden, liefern bei dem Vorliegen interfraktioneller Bewegungsvariabilität eine signifikant genauere Bewegungsschätzung als rein patientenspezifische Korrespondenzmodelle. Dies konnte im Rahmen der Evaluation für Subpopulationsmodelle gezeigt werden, bei denen die Wahl der zu bestim-

menden Cluster k beim k -means-basierten Clustering retrospektiv datensatzspezifisch optimiert wurde.

Die automatisierte Wahl des Parameters k erwies sich hingegen als problematisch. Auch wenn es durch die erzielten Ergebnisse durchaus Hinweise auf einen Zusammenhang zwischen optimalem k bzw. der Subpopulationsgröße und der interfraktionellen Bewegungsvariabilität gibt, konnten die eingesetzten Polynomfunktionen diesen nicht ausreichend genau abbilden, sodass mit den so erzeugten Subpopulationsmodellen keine Verbesserung gegenüber der Genauigkeit der patientenspezifischen Modelle erreicht werden konnte. Eine automatisierte, datengetriebene k -Wahl ist allerdings entscheidend für die Praxistauglichkeit des vorgestellten Verfahrens, weshalb diese Problematik in weiterführenden Arbeiten weiterbearbeitet werden sollte, um festzustellen, ob der angenommene Zusammenhang wirklich existiert und wie dieser geeignet modelliert werden kann. Hierbei ist insbesondere auch zu beachten, dass die in diesem Kapitel für die Parametrisierung des Zusammenhangs berechnete Größe zur Quantifizierung der interfraktionellen Bewegungsvariabilität die eigentlich zu schätzende Bewegung als gegeben voraussetzt. Hier ist dementsprechend ebenfalls eine praxistauglichere Lösung erforderlich. Eine Möglichkeit wäre die in [Tanner u. a., 2015] propagierte Aufnahme zweier Breath-Hold-Bilder bei maximaler Ein- bzw. Ausatmung zu Beginn der jeweiligen Behandlungssitzung, um basierend auf der Atembewegung zwischen diesen Bildern eine Quantifizierung der Interfraktionsvariabilität zu erreichen. Ebenso ist zu prüfen, ob sich diese Variabilität möglicherweise auch unter Zuhilfenahme hochdimensionaler Surrogatsignale kostengünstig und schnell ermitteln lässt (z. B. Patientenoberfläche; siehe auch Kapitel 2.1.1.1).

Einen Schwachpunkt der durchgeführten Evaluation ergibt sich durch die Nutzung von registrierungsbasiert ermittelten Transformationen als Grundwahrheit für die modellbasierte Bewegungsschätzung. Hier ist es wünschenswert in der Zukunft, ähnlich zu der Studie in Kapitel 3, eine landmarkenbasierte Auswertung durchzuführen, um den durch das eingesetzte Registrierungsverfahren eingeführten Bias zu entfernen. Diese landmarkenbasierte Auswertung sollte zudem im jeweiligen Patientenkoordinatensystem und nicht im Atlasraum erfolgen, um so auch eine für die Praxis relevante Einschätzung der Ungenauigkeiten zu erhalten, die durch den Transport der geschätzten Transformationen zwischen diesen Räumen entstehen.

Ein grundsätzlicher Nachteil der präsentierten Methodik zur Subpopulationsidentifikation ist die ausschließliche Fokussierung auf die interne Atembewegung bei der Ähnlichkeitsbestimmung. Diese Einschränkung ist, insbesondere vor dem Hintergrund der intendierten Erstellung optimierter Korrespondenzmodelle, als suboptimal zu bewerten, da sich gleiche interne Bewegungen verschiedener Patienten nicht unbedingt in einem ähnlichen Zusammenhang zum Surrogatsignal niederschlagen müssen und so u. U. das eigentliche Ziel des Clustering nicht (optimal) erreicht wird. Aus diesem Grund wäre in zukünftigen Arbeiten die Integration von Surrogatsignalinformationen

in das Clustering bzw. die Ähnlichkeitbestimmung sinnvoll (siehe auch Kapitel 5.6 für weitere Überlegungen zu diesem Thema).

Abseits der Auswertung des reinen Nutzens der Cluster bei der Korrespondenzmodellierung wären beispielsweise auch Analysen hinsichtlich deren einzelner Bedeutung und Zusammensetzung interessant. Möglicherweise ergeben sich auf diesem Wege auch noch weitere Ansätze für die automatisierte k -Wahl, wenn klarer ist, wie die einzelnen Patienten in einem Cluster zusammenhängen. Für diese und ähnliche Untersuchungen erscheint es zudem angebracht auch die Anzahl an Populationspatienten deutlich zu erhöhen, um die Aussagekraft der Ergebnisse zu verbessern. Eine gute Ausgangsbasis für die Erweiterung der Datenbasis wären die frei verfügbaren 4D-CT- und 4D-CBCT-Bilddaten¹⁶ von 20 Patienten aus [Hugo u. a., 2017]. Für diese Patienten liegen teilweise deutlich mehr als zwei Aufnahmen pro Patient vor, was zudem noch weitere Analysen der interfraktionellen Variabilität und ihrer Auswirkungen auf die Bewegungsschätzung erlaubt.

5.6 Alternative Methode basierend auf affinen Unterräumen

Die in Kapitel 5.2.2 vorgestellte Methode zur subpopulationsbasierten Korrespondenzmodellierung setzte einige Designentscheidungen und Annahmen voraus, die die konkrete Realisierung der einzelnen Methodenschritte maßgeblich und teilweise auch wechselseitig beeinflusst haben. In diesem Abschnitt werden diese Designentscheidungen unter Berücksichtigung der Erfahrungen aus der Evaluation noch einmal kritisch hinterfragt und es wird auf Basis der Erkenntnisse dieser Diskussion eine neue Alternativmethode vorgeschlagen, die allerdings im Rahmen dieser Dissertation nicht mehr praktisch umgesetzt wurde.

Folgende Designentscheidungen bzw. Annahmen sind wichtig für die präsentierte Subpopulationsmethode:

1. Die Zerlegung der Patientenpopulation erfolgt durch ein allgemeines Verfahren zum (Manifold) Clustering von Daten im \mathbb{R}^d , welches keinen konkreten Bezug zur zu modellierenden Atembewegung hat.
2. Um das gewählte (allgemeine) Clustering-Verfahren zur bewegungsbasierten Zerlegung der Patientenpopulation einsetzen zu können, wird die Bewegung jedes Patienten bzw. 4D-Datensatzes als einzelner Vektor im \mathbb{R}^d durch Konkatination der vektorisierten Verschiebungsfelder repräsentiert.
3. Die automatisierte Auswahl der patienten- bzw. datensatzspezifisch optimalen Anzahl von Clustern k erfolgt über die Annahme eines Zusammenhangs zwischen inter-

¹⁶CBCT: Cone Beam CT

fraktioneller Bewegungsdifferenz und der Anzahl der Patienten im zur Modellierung genutzten Cluster.

Ausgehend von diesen drei Punkten stellen sich insbesondere zwei Fragen: (1) Ist die durch das eingesetzte Clustering-Verfahren strukturell vorgegebene Repräsentation der patientenspezifischen Atembewegung sinnvoll und adäquat bezogen auf das eigentlich zu lösende Problem (= Erstellung von optimalen linearen Korrespondenzmodellen, die den Zusammenhang zwischen Surrogat und interner Bewegung beschreiben)? (2) Ist basierend auf der Annahme, dass es einen Zusammenhang zwischen der interfraktionellen Bewegungsdifferenz und der optimalen Subpopulationsgröße gibt, ein Clustering der Daten, bei dem nur ein indirekter Zusammenhang zwischen Parameter k und Clustergröße existiert, die richtige Strategie? Ausgehend von diesen beiden Fragen wird im Folgenden ein Alternativansatz zur Ermittlung einer passenden Subpopulation für einen Patienten $p = 0$ präsentiert. Basis dieser Alternative ist die Überlegung, dass ein patientenspezifisches, lineares Korrespondenzmodell eine kompakte Repräsentation der Atembewegung des jeweiligen Patienten ist und die Auswahl ähnlich atmender Patienten deshalb auf Basis dieser Modelle erfolgen kann bzw. sollte. Dies gilt insbesondere dann, wenn das Ziel wiederum die Erzeugung eines (sub-)populationsbasierten Korrespondenzmodells ist.

Sei nachfolgend die Atembewegung eines Patienten p durch das Parametertupel $(\bar{\mathbf{u}}_p, \mathbf{B}_p)$ eines linearen Korrespondenzmodells im Atlasraum (siehe Kapitel 5.2.1 und insbesondere Glg. 5.2¹⁷) repräsentiert. Hierbei ist $\bar{\mathbf{u}}_p \in \mathbb{R}^{3n_{vx}}$ die mittlere Bewegung von Patient p und $\mathbf{B}_p \in \mathbb{R}^{3n_{vx} \times n_{sur}}$ bezeichnet die zugehörige Matrix der Regressionskoeffizienten. Weiterhin gilt, dass die durch das mittels $(\bar{\mathbf{u}}_p, \mathbf{B}_p)$ parametrisierte Modell in Glg. 5.2 generierten Vektoren, Elemente des n_{sur} -dimensionalen¹⁸ affinen Unterraums $\mathcal{A}_p = \{\bar{\mathbf{u}}_p + \mathbf{w} \mid \mathbf{w} \in \text{span}(\mathbf{B}_p)\} \subseteq \mathbb{R}^{3n_{vx}}$ sind, der durch den linearen Unterraum $\text{span}(\mathbf{B}_p)$ ¹⁹ und den Verschiebungsvektor $\bar{\mathbf{u}}_p$ beschrieben wird (siehe auch [Lim u. a., 2018]). Dies bedeutet, dass die Atembewegung von p kompakt durch \mathcal{A}_p beschrieben werden kann. Diese Unterraumrepräsentation soll nun zur Quantifizierung der Ähnlichkeit der Atmung verschiedener Patienten zum Einsatz kommen. Gegeben Patient $p = 1$ mit affinem Unterraum \mathcal{A}_1 und Patient $p = 2$ mit affinem Unterraum \mathcal{A}_2 , dann soll die Ähnlichkeit der Bewegungsmuster über die Distanz zwischen \mathcal{A}_1 und \mathcal{A}_2 ermittelt werden.

Zur Berechnung von Distanzen zwischen linearen Unterräumen werden diese Unterräume insbesondere in Applikationen der Computer Vision seit einiger Zeit als Elemente der entsprechenden Graßmann-Mannigfaltigkeit $Gr(m, n)$ (= Mannigfaltigkeit, die alle m -dimensionalen linearen Unterräume des \mathbb{R}^n parametrisiert) betrachtet, wodurch

¹⁷Zur Vereinfachung der Notation wird in diesem Abschnitt auf die explizite Erwähnung von Ω_A verzichtet und stattdessen der Index p für die Patientenidentifizierung eingeführt.

¹⁸Um die Dimension n_{sur} zu erreichen, muss der Spaltenrang von \mathbf{B}_p n_{sur} entsprechen, was voraussetzt, dass die Surrogatsignaldimensionen in den Trainingsdaten linear unabhängig sind.

¹⁹ $\text{span}(\mathbf{B}_p)$ bezeichnet hier die lineare Hülle des Spaltenraums von \mathbf{B}_p .

die Distanz zwischen zwei Unterräumen als Distanz auf $Gr(m, n)$ definiert werden kann (siehe z. B. [Chang, 2008; Turaga u. a., 2008] für Übersichten und allgemeinere, mathematisch genauere Definitionen). Analog zu $Gr(m, n)$ als Mannigfaltigkeit der linearen Unterräume lässt sich mit $GrA(m, n)$ auch eine affine Graßmann-Mannigfaltigkeit definieren, die die m -dimensionalen affinen Unterräume des \mathbb{R}^n parametrisiert [Lim u. a., 2018]. Über diese Struktur können dann entsprechend auch Distanzen zwischen zwei affinen Unterräumen $\mathcal{A}_1, \mathcal{A}_2 \in GrA(n_{sur}, 3n_{vx})$ definiert werden. Nach [Lim u. a., 2018] lässt sich beispielsweise die geodätische Distanz entlang der Mannigfaltigkeit definieren als

$$d_{GrA}(\mathcal{A}_1, \mathcal{A}_2) = \left(\sum_{i=1}^{n_{sur}+1} \cos^{-1}(\sigma_i)^2 \right)^{\frac{1}{2}}, \quad (5.7)$$

wobei die σ_i die Singulärwerte von $\mathbf{Y}_{\mathcal{A}_1}^T \mathbf{Y}_{\mathcal{A}_2}$ mit

$$\mathbf{Y}_{\mathcal{A}_1} = \begin{bmatrix} \mathbf{O}_{\mathcal{A}_1} & \bar{\mathbf{u}}_1 / \sqrt{1 + \|\bar{\mathbf{u}}_1\|^2} \\ 0 & 1 / \sqrt{1 + \|\bar{\mathbf{u}}_1\|^2} \end{bmatrix} \quad \text{und} \quad \mathbf{Y}_{\mathcal{A}_2} = \begin{bmatrix} \mathbf{O}_{\mathcal{A}_2} & \bar{\mathbf{u}}_2 / \sqrt{1 + \|\bar{\mathbf{u}}_2\|^2} \\ 0 & 1 / \sqrt{1 + \|\bar{\mathbf{u}}_2\|^2} \end{bmatrix} \quad (5.8)$$

sind. Zur Konstruktion von $\mathbf{Y}_{\mathcal{A}_1}, \mathbf{Y}_{\mathcal{A}_2} \in \mathbb{R}^{3n_{vx}+1 \times n_{sur}+1}$, werden die beiden Matrizen $\mathbf{O}_{\mathcal{A}_1}, \mathbf{O}_{\mathcal{A}_2} \in \mathbb{R}^{3n_{vx} \times n_{sur}}$ benötigt, deren Spaltenvektoren eine Orthonormalbasis für $\text{span}(\mathbf{B}_1)$ bzw. $\text{span}(\mathbf{B}_2)$ bilden. Über Glg. 5.7 lässt sich nun die Ähnlichkeit der Atembewegung von zwei Patienten quantifizieren. Durch die Orthogonalisierung und Normierung der Koeffizientenmatrix gehen allerdings Informationen über das Verhältnis zum Surrogatsignal verloren, was u.U. auch Auswirkungen auf die Ähnlichkeitsbestimmung haben kann.

Analog zu Kapitel 5.2.2 ließe sich durch paarweise Berechnung der Distanz zwischen den Atemmustern aller $n_{pat} + 1$ Patienten wieder ein Ähnlichkeitsgraph für ein k -means-basiertes Spectral Clustering der Bewegungen erstellen. Hierbei wäre allerdings weiterhin der fehlende direkte Zusammenhang zwischen k und entstehender Clustergröße hinderlich. Als Alternative bietet sich deshalb die direkte Berechnung der k nächsten Nachbarn des Modells von Patient $p = 0$ als Subpopulation an, sodass $k = n_{sub}$.

Tabelle 5.1: Ergebnisse zu den Experimenten zur surrogatbasierten Bewegungsschätzung mit Subpopulationsmodellen angegeben als Mittelwert \pm Standardabweichung der mittleren Abweichung/des mittleren Fehlers bezogen auf die registrierungsbasiert ermittelten Transformationen (Pseudo-Grundwahrheit). Die Ergebnisse werden einzeln für die drei getesteten Atemphasen/Zeitpunkte und als Mittelwert dieser drei Ergebnisse für jede der beiden Gruppen von 4D-CT-Bilddaten und alle getesteten Modellen angegeben. Die vier Zeilen für die Subpopulationsmodelle beinhalten neben den Ergebnissen für ein Subpopulationsmodell mit datensatzspezifisch optimal gewähltem k auch die Ergebnisse für drei verschiedene Polynome mit Polynomgrad e zur automatisierten k -Wahl (siehe Text für Details zur k -Wahl). Tabelle modifiziert und erweitert aus [Wilms u. a., 2017c].

Bewegungsschätzung	Mittlerer Fehler [mm]			Mittelwert
	EI \rightarrow EE	EI \rightarrow MI	EI \rightarrow ME	
Gruppe 1				
Mittlere Bewegung	7.70 \pm 3.41	3.94 \pm 2.15	6.44 \pm 2.79	6.02 \pm 2.70
Interfraktionelle Bewegungsdifferenz	2.88 \pm 1.36	2.35 \pm 0.90	2.51 \pm 0.99	2.58 \pm 0.97
Intrafraktionsmodell	0.91 \pm 0.37	1.39 \pm 0.60	1.02 \pm 0.54	1.11 \pm 0.39
Interfraktionsmodell	2.93 \pm 1.63	2.05 \pm 0.83	2.59 \pm 1.29	2.52 \pm 1.16
Populationsmodell	3.21 \pm 0.95	1.96 \pm 0.68	2.97 \pm 1.06	2.71 \pm 0.77
Subpopulationsmodell (optimales k)	2.44 \pm 0.82	1.68 \pm 0.56	2.23 \pm 0.70	2.18 \pm 0.61
Subpopulationsmodell (k mit $e = 2$)	2.74 \pm 1.03	1.90 \pm 0.68	2.57 \pm 1.03	2.40 \pm 0.83
Subpopulationsmodell (k mit $e = 3$)	2.71 \pm 0.99	1.94 \pm 0.67	2.44 \pm 0.87	2.36 \pm 0.74
Subpopulationsmodell (k mit $e = 4$)	2.68 \pm 0.69	1.91 \pm 0.64	2.41 \pm 0.76	2.33 \pm 0.64
Gruppe 2				
Mittlere Bewegung	7.14 \pm 3.34	3.83 \pm 2.08	5.76 \pm 2.66	5.58 \pm 2.62
Interfraktionelle Bewegungsdifferenz	3.14 \pm 1.82	2.60 \pm 1.33	2.65 \pm 1.16	2.79 \pm 1.39
Intrafraktionsmodell	0.83 \pm 0.26	1.18 \pm 0.51	0.86 \pm 0.32	0.95 \pm 0.29
Interfraktionsmodell	2.39 \pm 1.50	2.10 \pm 1.40	2.03 \pm 0.77	2.17 \pm 1.19
Populationsmodell	3.27 \pm 0.92	2.09 \pm 0.75	2.97 \pm 0.77	2.77 \pm 0.71
Subpopulationsmodell (optimales k)	2.01 \pm 0.65	1.65 \pm 0.69	1.84 \pm 0.53	1.89 \pm 0.60
Subpopulationsmodell (k mit $e = 2$)	2.38 \pm 1.00	1.95 \pm 0.90	2.17 \pm 0.89	2.17 \pm 0.89
Subpopulationsmodell (k mit $e = 3$)	2.24 \pm 0.89	1.88 \pm 0.86	2.08 \pm 0.81	2.07 \pm 0.81
Subpopulationsmodell (k mit $e = 4$)	2.27 \pm 0.87	1.92 \pm 0.84	2.11 \pm 0.83	2.10 \pm 0.79

Kapitel 6

Patch- und modellbasierte Registrierungsverfahren für die zeiteffiziente Bewegungsschätzung mit intrainterventionellen Bilddaten

Kapitelinhalt

6.1	Literaturübersicht	113
6.2	Problembeschreibung und generelles Vorgehen	115
6.3	Registrierung durch modellgeführtes Block-Matching . . .	120
6.3.1	Methodenbeschreibung	121
6.3.2	Zusammenfassung	123
6.4	Registrierung mit einem Patch-AAM	125
6.4.1	Methodenbeschreibung	126
6.4.2	Zusammenfassung	128
6.5	Experimente	129
6.5.1	Verwendete Bilddaten	129
6.5.2	Design der Experimente	130
6.6	Ergebnisse	135
6.7	Zusammenfassung und Diskussion	138

Dieses Kapitel beschäftigt sich mit der zeiteffizienten Schätzung der Atembewegung auf Grundlage intrainterventionell aufgenommener Bilddaten. Durch die aktuellen technischen Entwicklungen hin zur Integration echtzeitfähiger, nicht-projektiver intrainterventioneller Bildaufnahmetechniken (zumeist US und MRT; [Denis de Senneville u. a., 2012; Lagendijk u. a., 2014; Menten u. a., 2017; O’Shea u. a., 2016a; Western u. a., 2015]) in moderne Systeme der Strahlentherapie (z. B. kombinierte MRT-Linearbeschleuniger) und der Ultraschallablation (HIFU) ist die effiziente Nutzung dieser Bilddaten ein hochaktuelles Thema. Im Gegensatz zu der in den vergangenen Kapiteln diskutierten indirekten Bewegungsschätzung anhand von externen Surrogaten

bieten die mittels MRT oder US aufgenommenen Bilddaten die Möglichkeit zumindest Teile der Therapieregion direkt zu beobachten, was zu einer erhöhten Präzision der Bewegungsschätzung und damit letztlich auch der eigentlichen Behandlung führen sollte [Paganelli u. a., 2015; Park u. a., 2018]. Zudem ermöglicht die Integration dieser nicht-ionisierenden Bildgebungstechniken in die Therapiesysteme die Aufnahme tagesaktueller Planungsdaten, womit beispielsweise die Effekte interfraktioneller Bewegungsveränderungen auf die Behandlung reduziert oder gänzlich ausgeschlossen werden können. Diesen Vorteilen stehen mindestens zwei Herausforderungen gegenüber:

1. Die intrainterventionelle Bewegungsschätzung muss trotz der erhöhten Datenmenge in Echtzeit erfolgen, um die extrahierten Informationen zur Bewegungskompensation in den Therapieprozess integrieren zu können. Die genauen Echtzeitanforderungen hängen von der verwendeten Bildgebungs- bzw. Therapietechnik ab (siehe auch Kapitel 1.1). Es erscheint aber sinnvoll zu fordern, dass die Verarbeitungsfrequenz mindestens der Bildaufnahmefrequenz des Therapiesystems entspricht.
2. Unterschiede zwischen den intrainterventionellen Bildern und dem Referenzbild hinsichtlich der abgebildeten Körperregionen bzw. ein Dimensionsunterschied zwischen intrainterventionellen Daten und Referenzbild (z. B. Referenz: 3D-MRT; intrainterventionell: 2D-MRT). Trotz dieser Einschränkungen ist weiterhin die Schätzung dichter Verschiebungsfelder größerer Körperregionen in 3D erstrebenswert.

Diese Herausforderungen können in ihrer Kombination in der Regel nicht mit den klassischen und mittlerweile sehr präzisen Bildregistrierungsverfahren zur voxelweisen Schätzung der Atembewegung in Planungsdaten bewältigt werden. Der aktuelle Stand der Entwicklungen in diesem Bereich spiegelt sich u. a. für 4D-CT-Bilddaten in den Ergebnissen der EMPIRE10-Challenge²⁰ [Murphy u. a., 2011] und den Ergebnissen auf den Daten des DIR-Lab²¹ [Castillo u. a., 2009, 2010, 2013] wider. So geben beispielsweise Rühaak u. a. [2017] Rechenzeiten von ca. 5 Minuten auf einem Standard-PC für ihr Verfahren zur Schätzung dichter Verschiebungsfelder der Lungenbewegung an, welches im März 2018 u. a. das Ranking der EMPIRE10-Challenge anführt. Es ist zu erwarten, dass sich diese Rechenzeiten für potenziell geringer aufgelöste intrainterventionelle Bilddaten noch reduzieren lassen, wobei für einen Echtzeiteinsatz Geschwindigkeitssteigerungen um 1–2 Größenordnungen notwendig wären. Für viele klassische Verfahren ist zudem unklar, wie sie für die Registrierung von Daten mit unterschiedlicher Dimension oder teilweiser Überlappung adaptiert werden können, sodass trotzdem ein dichtes, dreidimensionales Verschiebungsfeld der kompletten Therapieregion bestimmt werden kann. Dies motiviert letztlich die Entwicklung spezieller intrainterventioneller Registrierungsverfahren.

²⁰<http://empire10.isi.uu.nl>

²¹<https://www.dir-lab.com>

In diesem Kapitel werden zwei neue Registrierungsverfahren vorgeschlagen, welche die beiden Herausforderungen direkt adressieren (siehe Kapitel 6.3 und Kapitel 6.4). Beide Verfahren haben gemeinsam, dass sie sowohl patch- als auch modellbasiert arbeiten. Patchbasiert bedeutet in diesem Kapitel, dass im direkt vor der Intervention aufgenommenen Referenzbild eine geringe Menge von markanten Punkten detektiert wird (Merkmalspunkte), die mit hoher Wahrscheinlichkeit auch in den intrainerventionell akquirierten Bildern sichtbar sind. Hierbei wird jeder Merkmalspunkt durch die Bildinformationen seiner unmittelbaren Umgebung (Patch) repräsentiert und die eigentliche Registrierung bezieht sich dann nur noch auf diese wenigen Merkmalspunkte/Patches. Dies verringert gegenüber einer voxelweisen Registrierung den Rechenaufwand erheblich und erlaubt einen problemlosen Umgang mit unvollständigen/partiellen Daten durch Anpassung der Patches. In beiden Ansätzen kommt zudem ein PCA-basiertes, patientenspezifisches statistisches Bewegungsmodell zum Einsatz, welches auf dichten Verschiebungsfeldern aufbaut, die aus kurz vor der Intervention aufgenommenen Bilddaten der kompletten Therapieregion extrahiert wurden. Das statistische Bewegungsmodell dient während des Registrierungsprozesses als patientenspezifischer Regularisierer des nur an den Merkmalspunkten definierten, dünnbesetzten Verschiebungsfeldes und wird zudem als patientenspezifischer Interpolator eingesetzt, um aus dem dünnbesetzten Verschiebungsfeld ein dichtes Feld zu generieren.

Eine grobe Unterscheidung beider Verfahren ist u. a. anhand des konkreten Aufbaus und der Lösung des Registrierungsproblems möglich. Während das Verfahren in Kapitel 6.3 auf diskrete Optimierung setzt und traditionelles Block-Matching mit modellbasierter Regularisierung kombiniert, wird in Kapitel 6.4 ein patchbasiertes aktives Erscheinungsmodell (engl. Patch-based Active Appearance Model; Patch-based AAM) für die intrainerventionelle Bewegungsschätzung adaptiert. Hierbei wird die Bewegungsschätzung ableitungsbasiert iterativ durchgeführt, wobei Vorwissen aus einem Bewegungsmodell und einem Modell der Bildinformationen kombiniert wird.

Teilkapitel 6.1 gibt im Anschluss eine Literaturübersicht hinsichtlich vorgeschlagener Verfahren für die intrainerventionelle Schätzung der Atembewegung basierend auf MRT- und US-Bilddaten. In Kapitel 6.2 erfolgt dann eine allgemeine Beschreibung des Registrierungsproblems und des statistischen Bewegungsmodells. Die Kapitel 6.3 und 6.4 beschreiben anschließend die beiden Methoden, welche in Kapitel 6.5 und Kapitel 6.6 umfangreich auf der Basis von MRT-/US-Bilddaten evaluiert werden. Die Darstellungen in Kapitel 6.1–6.3 basieren hierbei größtenteils auf [Ha u. a., 2019; Wilms u. a., 2016], wohingegen Kapitel 6.4 in Teilen [Wilms u. a., 2017b] als Basis nutzt.

6.1 Literaturübersicht

Diese Literaturübersicht fokussiert sich auf zeiteffiziente Registrierungsverfahren für die intrainerventionelle Schätzung der (dreidimensionalen) Atembewegung anhand

von MRT- oder US-Daten. Es erfolgt aus Gründen der Übersichtlichkeit zudem keine explizite Diskussion genereller patch- und/oder modellbasierter Registrierungsansätze. Eine entsprechende Übersicht ist u. a. in [Sotiras u. a., 2013] zu finden. Ebenso werden andere Modalitäten für die Aufnahme der intrainerventionellen Daten nicht behandelt.

Vielfach haben Autoren bisher klassische Registrierungsverfahren für die intrainerventionelle Bewegungsschätzung adaptiert. So wurde von Somphone u. a. [2014] im Rahmen der CLUST-Challenge²² ein Sparse Demons-Algorithmus präsentiert, der die Optimierung auf ausgewählte Bildpunkte beschränkt, um eine Echtzeitfähigkeit (≈ 100 ms für 3D-US-Registrierung) zu erreichen. Neben angepassten Demons-Verfahren gibt es auch entsprechende Optical Flow-Versionen [Denis de Senneville u. a., 2015; Seregini u. a., 2018; Zachiu u. a., 2015], die echtzeitfähig sind (z. B. < 50 ms in [Seregini u. a., 2018; Zachiu u. a., 2015] für 2D-MRT-Registrierung), was ebenso für die variationellen Verfahren in [König u. a., 2014; Zachiu u. a., 2017] gilt. Bei den zitierten Verfahren wird hauptsächlich von einer gleichen Dimension von Referenz- und Zielbild ausgegangen, sodass keine Lösung für den Umgang mit unvollständigen Daten existiert.

Viele Verfahren in der Literatur setzen auf Block-Matching oder verwandte Techniken (siehe u. a. [Banerjee u. a., 2015; Bjerre u. a., 2013; Brix u. a., 2014; Cerviño u. a., 2011; De Luca u. a., 2012, 2013; Ipsen u. a., 2016; Paganelli u. a., 2015; Shepard u. a., 2017]), um eine zeiteffiziente Bewegungsschätzung realisieren zu können. Vorteil dieser Techniken ist, dass u. a. durch den Einsatz von GPU-Implementierungen relativ genaue Bewegungsschätzungen in Echtzeit möglich sind, ohne die Probleme mit lokalen Minima variationeller Ansätze. Es ergeben sich in der Regel aber auch zwei Nachteile: (1) Die Bewegungsschätzung bezieht sich zumeist nur auf wenige Punkte oder kleine Regionen. Dichte Verschiebungsfelder der gesamten Therapieregion können nicht erzeugt werden. (2) Viele der aufgeführten Verfahren (u. a. [Brix u. a., 2014]) betrachten die Verfolgung mehrerer Punkte/Regionen als unabhängige Probleme, ohne dass Vorwissen über die wechselseitigen Abhängigkeiten der Bewegungen genutzt wird. Somit sind für diese Verfahren u. a. räumlich widersprüchliche und nichtglatte Bewegungen zu erwarten, die in klassischen Registrierungsverfahren durch eine räumliche Regularisierung verhindert werden.

Um ausgehend von dünnbesetzten Verschiebungsfeldern dichte Verschiebungsfelder bestimmen zu können, bieten sich verschiedene Interpolationstechniken²³ an. Unter den üblichen Verfahren befinden sich u. a. allgemeine Interpolationstechniken wie Thin-Plate-Splines [Lee u. a., 2011; Miao u. a., 2016], B-Splines [Nakamoto u. a., 2006] oder stückweise definierte affine Transformationen [Royer u. a., 2017]. Nachteil dieser all-

²²Challenge on Liver Ultrasound Tracking (CLUST); <http://clust.ethz.ch/>; [De Luca u. a., 2015]

²³Das Wort Interpolationstechniken wird hier als Synonym für Interpolations- und Extrapolationstechniken genutzt und bei der Nutzung von Bewegungsmodellen auch als Synonym für die modellbasierte Rekonstruktion eines Verschiebungsfeldes.

gemeinen, unspezifischen Techniken ist die Tatsache, dass patientenspezifische Eigenschaften der Atembewegung nur dann in der Interpolation abgebildet werden können, wenn sie schon im dünnbesetzten Verschiebungsfeld auffällig sind. Demgegenüber stehen (patientenspezifische) modellbasierte Interpolationstechniken, die ein statistisches Bewegungsmodell für die Interpolation nutzen [Boye u. a., 2013; Harris u. a., 2016; King u. a., 2012; Klinder u. a., 2012; Preiswerk u. a., 2014; Stemkens u. a., 2016], welches auf der Grundlage von Trainingsdaten erzeugt wird. Mit diesen Modellen wird bei der Interpolation eine Abbildung von patientenspezifischen Eigenheiten der Atembewegung möglich, auch wenn diese nicht direkt im dünnbesetzten Verschiebungsfeld zu erkennen sind. Diese Bewegungsmodelle können zudem nicht nur als patientenspezifische Interpolatoren eingesetzt werden, sondern sind aufgrund der modellierten Zusammenhänge zwischen den einzelnen Punkten auch als Vorwissen für eine Regularisierung bzw. Beschränkung der Bewegungsschätzung selbst nützlich (siehe u. a. [Denis de Senneville u. a., 2015; King u. a., 2012; Stemkens u. a., 2016]), was in Block-Matching-Verfahren bisher aber kaum genutzt wird. Neben statistischen Bewegungsmodellen sind zudem auch andere Arten von Modellen der Atembewegung möglich. So wird beispielsweise in [Royer u. a., 2017] ein biomechanisches Modell (Masse-Feder-System) für die Regularisierung eingesetzt.

6.2 Problembeschreibung und generelles Vorgehen

Ziel dieses Teilkapitels ist die methodenübergreifende Darstellung des zu lösenden Registrierungsproblems. Um die Übersichtlichkeit der Notation zu gewährleisten, wird hierfür angenommen, dass das Referenzbild und die intrainerventionell akquirierten Bilddaten jeweils 3D-Bilder sind, auch wenn beide Verfahren ebenso mit Daten anderer und unterschiedlicher Dimension arbeiten können. Die zur Illustration genutzten Bildbeispiele beziehen sich allerdings auf 2D-Bilder.

Registrierungsproblem Sei $I_R : \Omega \rightarrow \mathbb{R}$ (mit $\Omega \subset \mathbb{R}^3$) das 3D-Referenzbild, welches unmittelbar vor der eigentlichen Intervention aufgenommen wurde und welches als Ausgangspunkt für die Schätzung der atmungsbedingten Bewegung der abgebildeten Strukturen dient. Weiterhin sei mit $I_t : \Omega \rightarrow \mathbb{R}$ ein während der Intervention zum Zeitpunkt t aufgenommenes 3D-Zielbild. Ziel ist es dann, eine Transformation $\varphi_t = Id + u_t$ ($\varphi_t : \Omega \rightarrow \Omega$) mit Verschiebungsfeld $u_t : \Omega \rightarrow \mathbb{R}^3$ zu schätzen, welche die Atembewegung zwischen I_R und I_t für jeden Punkt des Referenzbildes beschreibt. Weiterhin sei $\Omega_{fp} = \{\mathbf{x}_1, \dots, \mathbf{x}_{n_{fp}}\} \subset \Omega$ ($n_{fp} = |\Omega_{fp}|$) eine Menge von Merkmalspunkten, die markante Punkte in I_R beschreiben und mit hoher Wahrscheinlichkeit auch in I_t sichtbar sind. Diese Punkte werden in dieser Arbeit automatisiert mittels Förstner- bzw. Harris-Operator [Förstner u. a., 1987; Harris u. a., 1988] in I_R detektiert. Das

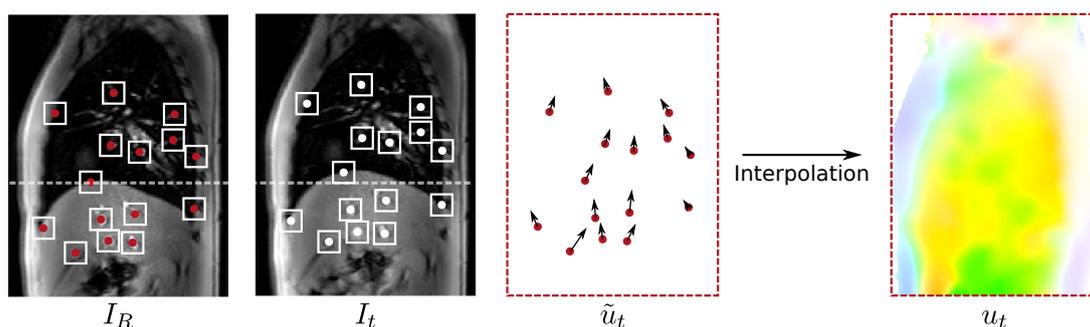


Abb. 6.1: Illustration des Registrierungsproblems zwischen I_R und I_t . Die eigentliche Registrierung bezieht sich nur auf eine geringe Anzahl von Patches (weiße Blöcke), welche um in I_R definierte Merkmalspunkte (rote Punkte) zentriert sind. Für diese Patches wird die entsprechende Position in I_t gesucht (weiße Punkte), wodurch das dünnbesetzte Verschiebungsfeld \tilde{u}_t entsteht. Aus diesem wird anschließend mittels modellbasierter Interpolation das dichte Verschiebungsfeld u_t generiert.

zu u_t korrespondierende dünnbesetzte (= nur an diesen Merkmalspunkten definierte) Verschiebungsfeld wird als $\tilde{u}_t : \Omega_{fp} \rightarrow \mathbb{R}^3$ bezeichnet.

Die Registrierung beschränkt sich aus Effizienzgründen in einem ersten Schritt auf die Suche nach einem dünnbesetzten Verschiebungsfeld \tilde{u}_t (siehe Abb. 6.1). Hierzu wird die Güte der Abbildung über ein Ähnlichkeitsmaß \mathcal{D} quantifiziert, welches entsprechend der Definition von \tilde{u}_t nur lokal für Patches mit Zentrum $\mathbf{x} \in \Omega_{fp}$ ausgewertet wird. Um \tilde{u}_t auf glatte bzw. physiologisch plausible Lösungen einzuschränken, wird ähnlich zu Glg. 2.6 ein Regularisierer eingesetzt, der in diesem Kapitel durch ein patientenspezifisches statistisches Bewegungsmodell repräsentiert wird. Das Bewegungsmodell wird zudem auch genutzt, um aus dem dünnbesetzten Verschiebungsfeld \tilde{u}_t ein dichtes Verschiebungsfeld u_t zu interpolieren.

Das dünnbesetzte Verschiebungsfeld \tilde{u}_t lässt sich dann ähnlich zu Glg. 2.6 durch Minimierung einer Kostenfunktion

$$E(\tilde{u}_t) = \sum_{\mathbf{x} \in \Omega_{fp}} \mathcal{D}(I_R, I_t, \tilde{u}_t, \mathbf{x}) + \alpha \mathcal{S}(\tilde{u}_t) \quad (6.1)$$

schätzen. \mathcal{S} bezeichnet hier den additiven, modellbasierten Regularisierungsterm, der un plausible Felder lediglich bestraft und mit $\alpha \in \mathbb{R}_{\geq 0}$ gewichtet wird. Alternativ lässt sich die Plausibilität des Feldes auch durch die explizite Restriktion möglicher Lösungen des Optimierungsproblems auf durch das Bewegungsmodell parametrisierbare Felder erzwingen.

Das Ähnlichkeitsmaß wird für dieses Kapitel definiert als

$$\mathcal{D}(I_R, I_t, \tilde{u}_t, \mathbf{x}) = \|\theta(I_R, \mathbf{x}) - \theta(I_t, \mathbf{x} + \tilde{u}_t(\mathbf{x}))\|_2^2, \quad (6.2)$$

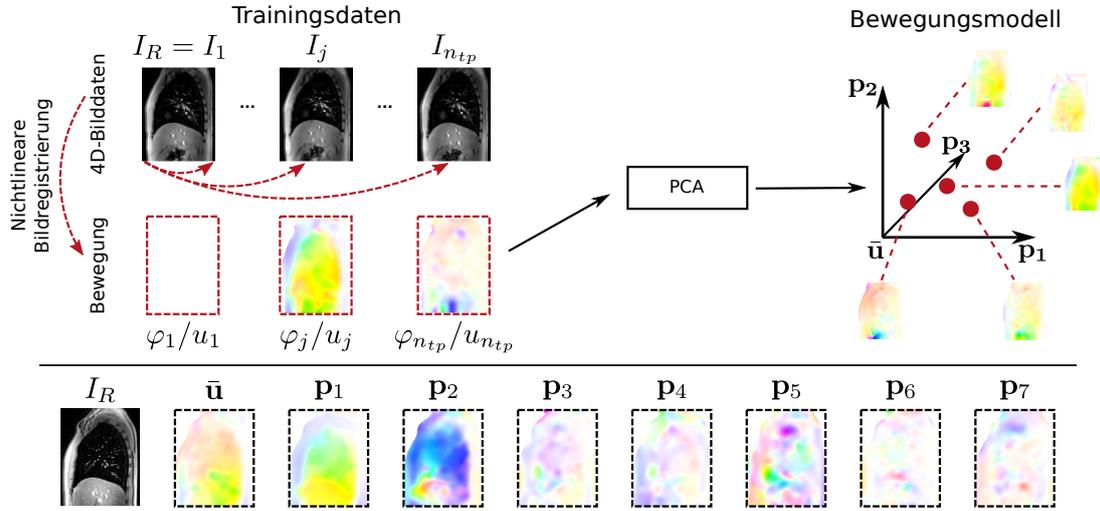


Abb. 6.2: Schematische Darstellung der Generierung des patientenspezifischen statistischen Bewegungsmodells. In den für das Modelltraining aufgenommenen räumlich-zeitlichen Bilddaten wird registrierungsbasiert die Atembewegung geschätzt. Aus den ermittelten Transformationen bzw. Verschiebungsfeldern wird mittels PCA ein affiner Unterraum geschätzt, der plausible Verschiebungsfelder des Patienten beschreibt. Im unteren Teil der Abbildung sind zudem die mittlere Bewegung und die ersten sieben Eigenvektoren/Basisvektoren ($2\sqrt{\lambda_i}\mathbf{p}_i$) des statistischen Bewegungsmodells eines Beispielpatienten dargestellt. Die Sättigung der Farbcodierung in Abhängigkeit von der Bewegungsamplitude ist für jeden Eigenvektor/Basisvektor zur besseren Darstellung unterschiedlich gewählt. Siehe Abb. 6.3 für die Richtungsbedeutung der Farben.

wobei $\theta(\cdot, \cdot)$ eine abstrakte Funktion zur Merkmals- und Patchextraktion ist, die für ein gegebenes Bild I_R und einen gegebenen Punkt \mathbf{x} einen Patchvektor $\mathbf{f}_{\mathbf{x}, I_R} \in \mathbb{R}^{n_{psz}^3 n_{pch}}$ erzeugt. Der Patchvektor $\mathbf{f}_{\mathbf{x}, I_R}$ enthält die Bildinformationen des kubischen Patches mit n_{psz} Abtastpunkten pro Raumrichtung (n_{psz}^3 insgesamt), dessen Zentrum \mathbf{x} ist. n_{pch} gibt hierbei die Anzahl der extrahierten Merkmale pro Abtastpunkt an. Bei einer intensitätsbasierten Registrierung ist $n_{pch} = 1$ und die Elemente von $\mathbf{f}_{\mathbf{x}, I_R}$ und $\mathbf{f}_{\mathbf{x}, I_t}$ sind die Grauwerte der beteiligten Bilder. Entsprechend ist $\|\mathbf{f}_{\mathbf{x}, I_R} - \mathbf{f}_{\mathbf{x}, I_t}\|_2^2$ die Summe der quadrierten Grauwertdifferenzen (SSD) der Patches. Durch entsprechende Wahl von $\theta(\cdot, \cdot)$ und n_{pch} ist auch die Nutzung von verschiedensten Bilddescriptoren (z. B. MIND [Heinrich u. a., 2012], SSC [Heinrich u. a., 2013b] oder SIFT [Lowe, 2004]) möglich. $\mathcal{D}(I_R, I_t, \tilde{u}_t, \mathbf{x})$ ist zumeist nichtlinear von \tilde{u}_t abhängig und nicht konvex.

Statistisches Bewegungsmodell Das statistische Bewegungsmodell wird in einer Trainingsphase erzeugt. Hierfür sei zusätzlich zu I_R ein ebenfalls unmittelbar vor der Intervention aufgenommener 4D-Bilddatensatz $(I_j)_{j=1}^{n_{tp}}$ bestehend aus n_{tp} Einzelbil-

den $I_j : \Omega \rightarrow \mathbb{R}$ gegeben, der die Strukturen des Referenzbildes²⁴ zu verschiedenen Atemzuständen abbildet. Durch nichtlineare Registrierung von I_R und allen Einzelbildern lassen sich dann wieder Transformationen ($\varphi_j = Id + u_j$) _{$j=1$} ^{n_{tp}} schätzen, welche die Atembewegung in dem 4D-Datensatz beschreiben. Auf Basis der Verschiebungsfelder (u_j) _{$j=1$} ^{n_{tp}} wird dann analog zu [Klinder u. a., 2012; Preiswerk u. a., 2014] mittels Hauptkomponentenanalyse (PCA, siehe auch Kapitel 2.4.1) ein statistisches Bewegungsmodell erstellt, welches den Raum plausibler patientenspezifischer Transformationen/Verschiebungsfelder kompakt beschreibt (siehe Abb. 6.2). Hierfür werden die Verschiebungsfelder in Spaltenvektoren $\mathbf{u}_j = \text{vec}(u_j) \in \mathbb{R}^{3n_{vx}}$ überführt, wobei n_{vx} wieder die Anzahl der Voxel des Referenzbildes angibt.

Eine PCA der vektorisierten Verschiebungsfelder (\mathbf{u}_j) _{$j=1$} ^{n_{tp}} ist durch Eigenwertzerlegung der Kovarianzmatrix

$$\mathbf{C}_{\mathbf{u}} = \frac{1}{(n_{tp} - 1)} \sum_{j=1}^{n_{tp}} (\mathbf{u}_j - \bar{\mathbf{u}})(\mathbf{u}_j - \bar{\mathbf{u}})^T = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^T \quad (6.3)$$

möglich. Hierbei bezeichnet $\bar{\mathbf{u}} = \frac{1}{n_{tp}} \sum_{j=1}^{n_{tp}} \mathbf{u}_j$ die mittlere Bewegung in den Trainingsdaten. Die Eigenvektoren von $\mathbf{C}_{\mathbf{u}}$ sind durch die Spaltenvektoren der orthonormalen Matrix $\mathbf{P} = [\mathbf{p}_1 \cdots \mathbf{p}_{3n_{vx}}] \in \mathbb{R}^{3n_{vx} \times 3n_{vx}}$ gegeben und die Diagonalelemente der Diagonalmatrix $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_{3n_{vx}}) \in \mathbb{R}_{\geq 0}^{3n_{vx} \times 3n_{vx}}$ sind die korrespondierenden Eigenwerte. Die Eigenwerte geben die Varianz entlang des jeweiligen Eigenvektors in den Trainingsdaten an und werden typischerweise genutzt, um eine Reduktion der Eigenvektoren für die kompakte Modellierung zu erreichen. Sei hierfür $\mathbf{P}_k \in \mathbb{R}^{3n_{vx} \times k}$ die orthonormale Matrix bestehend aus den k größten Eigenvektoren von $\mathbf{C}_{\mathbf{u}}$, die einen gewählten Teil (hier immer 98%) der Gesamtvarianz in den Trainingsdaten abdecken. Für die Übersichtlichkeit der Notation wird der Index k in der Folge weggelassen.

Die Kombination aus mittlerer Bewegung $\bar{\mathbf{u}}$ und \mathbf{P} parametrisiert dann einen k -dimensionalen affinen Unterraum $\{\bar{\mathbf{u}} + \mathbf{w} \mid \mathbf{w} \in \text{span}(\mathbf{P})\}$ des $\mathbb{R}^{3n_{vx}}$ (siehe auch Kapitel 5.6), der basierend auf den Trainingsdaten plausible Verschiebungsfelder beschreibt. Elemente des Unterraums lassen sich durch

$$\hat{\mathbf{u}} = \bar{\mathbf{u}} + \mathbf{P}\mathbf{S}\mathbf{b} \quad (6.4)$$

erzeugen ($\hat{\mathbf{u}} = \text{vec}^{-1}(\hat{\mathbf{u}})$). Hierbei bezeichnet $\mathbf{b} \in \mathbb{R}^k$ den zu wählenden Parametervektor (= Koordinatenvektor des Feldes bezogen auf die Orthonormalbasis) und $\mathbf{S} = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_k}) \in \mathbb{R}_{\geq 0}^{k \times k}$ ist eine zur Gewichtung genutzte Diagonalmatrix, die die Standardabweichungen aus den Trainingsdaten enthält.

Die Anwendung der schon für die Korrespondenzmodellierung eingesetzten Thikonov-Regularisierung (siehe Glg. 2.13) bietet eine einfache Möglichkeit zur regularisier-

²⁴Typischerweise ist I_R ein Element des 4D-Bilddatensatzes.

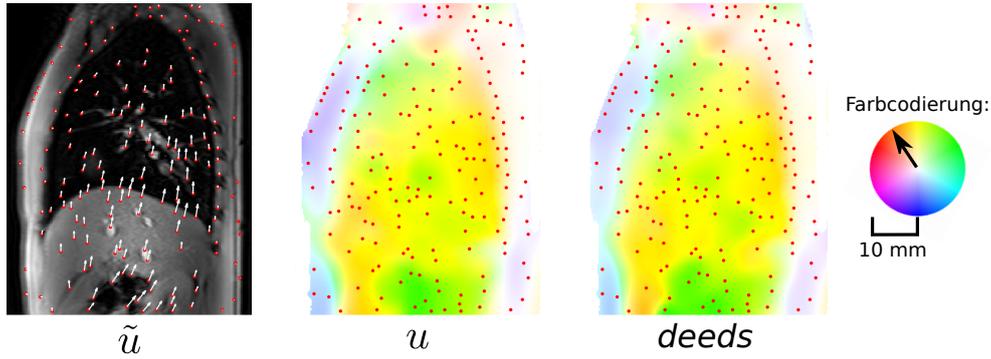


Abb. 6.3: Modellbasierte Interpolation eines dünnbesetzten Verschiebungsfeldes \tilde{u} (links; nur definiert an den 176 roten Punkten) mithilfe von Glg. 6.5 und Glg. 6.4. Das interpolierte Feld u (mittig; farbcodierte Darstellung) zeigt überwiegend dieselben Bewegungseigenschaften, wie das durch pixelweise Registrierung zweier 2D-MRT-Bilder mit *deeds* [Heinrich u. a., 2013a] erzeugte Referenzfeld (rechts; farbcodierte Darstellung), aus dem \tilde{u} extrahiert wurde. Das Referenzfeld war kein Teil der Trainingsdaten des Modells.

ten Projektion eines Verschiebungsfeldes u bzw. \mathbf{u} auf den affinen Unterraum basierend auf der Minimierung der Summe der Fehlerquadrate [Blanz u. a., 2002]:

$$\mathbf{b}_{\mathbf{u}} = \arg \min_{\mathbf{b}} \|\mathbf{P}\mathbf{S}\mathbf{b} - (\mathbf{u} - \bar{\mathbf{u}})\|_2^2 + \eta \|\mathbf{b}\|_2^2 . \quad (6.5)$$

Das Minimierungsproblem kann analog zu Glg. 2.13 geschlossen gelöst werden und mittels Einsetzen des optimalen Parametervektors $\mathbf{b}_{\mathbf{u}}$ in Glg. 6.4 ergibt sich das gesuchte projizierte Verschiebungsfeld. Der Regularisierungsparameter $\eta \in \mathbb{R}_{\geq 0}$ gewichtet den Rekonstruktionsfehler und die quadrierte Norm des Parametervektors gegeneinander. Durch die Multiplikation mit \mathbf{S} entspricht $\|\mathbf{b}\|_2^2$ der (quadrierten) Mahalanobis-Distanz des approximierten Verschiebungsfeldes zu $\bar{\mathbf{u}}$ [Blanz u. a., 2002].

Das Minimierungsproblem in Glg. 6.5 lässt sich ebenso zur regularisierten Approximation bzw. Interpolation von dünnbesetzten Feldern einsetzen (siehe Abb. 6.3). Sei hierfür $\tilde{\mathbf{u}} \in \mathbb{R}^{3n_{fp}}$ der Verschiebungsfeldvektor des dünnbesetzten Feldes \tilde{u} . Weiterhin seien $\tilde{\mathbf{P}} \in \mathbb{R}^{3n_{fp}}$ und $\tilde{\mathbf{u}} \in \mathbb{R}^{3n_{fp}}$ die zu \mathbf{P} korrespondierende Matrix und der zu $\bar{\mathbf{u}}$ korrespondierende Vektor, welche nur noch die zu den Merkmalspunkten gehörenden Elemente enthalten. Durch Austausch der entsprechenden Teile in Glg. 6.5 durch die reduzierten Versionen ergibt sich dann die Möglichkeit den optimalen Parametervektor $\mathbf{b}_{\mathbf{u}}$ zu dem Vektor $\tilde{\mathbf{u}}$ eines dünnbesetzten Verschiebungsfeldes $\tilde{u} : \Omega_{fp} \rightarrow \mathbb{R}^3$ zu ermitteln und in Kombination mit Glg. 6.4 ein dichtes Verschiebungsfeld \mathbf{u} zu generieren [Blanz u. a., 2002]. Diese Projektion ist in der Literatur ein Standardverfahren zur Regularisierung und Interpolation von dünnbesetzten Verschiebungsfeldern (siehe u. a. [Klinder u. a., 2012; Preiswerk u. a., 2014]).

Durch diese Technik kann zu einem beliebigen (nicht glatten, physiologisch unplausiblen) dünnbesetzten oder dichten Verschiebungsfeld ein Feld des durch das Modell repräsentierten affinen Unterraums erzeugt werden, welches das gegebene Feld optimal approximiert. Unter der Annahme, dass das Modell nur Felder generieren kann, die sehr ähnlich zu den für das Training eingesetzten Verschiebungsfeldern sind, sollten diese auch die charakteristischen Eigenschaften des für das Training genutzten Registrierungsverfahrens u. a. hinsichtlich Glattheit und physiologischer Plausibilität (z. B. Berücksichtigung von Gleitbewegungen) aufweisen.

6.3 Registrierung durch modellgeführtes Block-Matching

Ziel dieses Teilkapitels ist die Entwicklung eines Verfahrens, welches das patchbasierte Registrierungsproblem in Kapitel 6.2 (approximativ) durch ein modellgeführtes bzw. -regularisiertes Block-Matching löst und bei Implementierung auf heutiger Standardgrafikhardware echtzeitfähig ist. Anders als die in Kapitel 6.1 diskutierten Block-Matching-Verfahren, soll das Bewegungsmodell nicht nur in einem finalen Schritt zur Interpolation eines dichten Verschiebungsfeldes dienen, sondern wie in [Denis de Senneville u. a., 2015; Harris u. a., 2016; King u. a., 2012; Stemkens u. a., 2016] schon während des Registrierungsprozesses als Vorwissen hinsichtlich glatter und plausibler Felder eingesetzt werden.

In [Denis de Senneville u. a., 2015; Harris u. a., 2016; King u. a., 2012; Stemkens u. a., 2016] wird allerdings kein diskretes Block-Matching eingesetzt, sondern das jeweilige nichtlineare Optimierungsproblem wird klassisch mittels iterativer, ableitungsbasierter Optimierungsverfahren (kontinuierliche Optimierung) gelöst. Dies führt zu den für diese Verfahrensklasse üblichen Nachteilen (u. a. langsame Konvergenz und die Konvergenz gegen lokale Optima). Im Bereich klassischer Registrierungsverfahren sind deshalb einige diskrete Verfahren vorgeschlagen worden (z. B. [Glocker u. a., 2008, 2011; Heinrich u. a., 2013a; Heinrich u. a., 2016]; auch patchbasiert: [Dalca u. a., 2016; Heinrich u. a., 2015]), die den Raum der zulässigen Verschiebungen pro Voxel diskretisieren und die optimale Auswahl des jeweiligen Verschiebungsvektors dann beispielsweise als Labelingproblem in einem Graphen beschreiben, welches global optimal gelöst werden kann. Hierdurch werden lokal optimale Lösungen verhindert, die Optimierung kombinierter Kostenfunktionen (siehe Glg. 6.2) bleibt allerdings auch mit diesen graphenbasierten Verfahren rechenaufwendig. Aus diesem Grund wird in [Heinrich u. a., 2014] basierend auf [Steinbrücker u. a., 2009] vorgeschlagen für diskrete Registrierungsverfahren die Optimierung von Ähnlichkeitsmaß und Regularisierung in zwei Teilprobleme aufzuteilen, die sich einzeln effizient und global optimal lösen lassen. Durch die alternierende Optimierung dieser Teilprobleme, welche durch einen Kopplungsterm verbunden werden, ist so eine Approximation des globalen Optimums des Originalproblems möglich. Dieses generelle Grundprinzip wird schon lange erfolgreich in nichtdiskreten Verfah-

ren der medizinischen Bildregistrierung verwendet [Cachier u. a., 2003; Miao u. a., 2016; Vercauteren u. a., 2009]. Das hier präsentierte Verfahren kann grob als Erweiterung von [Heinrich u. a., 2014] (voxelweise diskrete Registrierung mit diffusivem Regularisierer) um die Möglichkeit zur Nutzung von unvollständigen Daten und eines modellbasierten Regularisierers aufgefasst werden.

6.3.1 Methodenbeschreibung

Sei

$$E(\tilde{u}_t, \tilde{w}_t) = \sum_{\mathbf{x} \in \Omega_{fp}} \left(\mathcal{D}(I_R, I_t, \tilde{u}_t, \mathbf{x}) + \kappa \|\tilde{u}_t(\mathbf{x}) - \tilde{w}_t(\mathbf{x})\|_2^2 \right) + \alpha \mathcal{S}(\tilde{w}_t) \quad (6.6)$$

die nach [Heinrich u. a., 2014; Steinbrücker u. a., 2009] um ein zusätzliches Verschiebungsfeld $\tilde{w}_t : \Omega_{fp} \rightarrow \mathbb{R}^3$ und einen Kopplungsterm $\kappa \|\tilde{u}_t(\mathbf{x}) - \tilde{w}_t(\mathbf{x})\|_2^2$ erweiterte Kostenfunktion aus Glg. 6.2. Zusätzlich sei der modellbasierte Regularisierungsterm $\mathcal{S}(\tilde{w}_t)$ aufbauend auf Glg. 6.5 (hier zur Vereinfachung mit $\eta = 0$) definiert als

$$\begin{aligned} \mathcal{S}(\tilde{w}_t) &= \left\| \tilde{\mathbf{P}}\mathbf{S}\mathbf{b} - (\tilde{\mathbf{w}}_t - \tilde{\mathbf{u}}) \right\|_2^2 \text{ mit } \tilde{\mathbf{w}}_t = \text{vec}(\tilde{w}_t) \\ \text{und } \mathbf{b} &= \arg \min_{\mathbf{b}} \left\| \tilde{\mathbf{P}}\mathbf{S}\mathbf{b} - (\tilde{\mathbf{w}}_t - \tilde{\mathbf{u}}) \right\|_2^2 . \end{aligned} \quad (6.7)$$

Der Regularisierungsterm quantifiziert in diesem Fall die Entfernung von \tilde{w}_t zum nächsten durch das Bewegungsmodell parametrisierbaren Verschiebungsfeld.

Die Kostenfunktion in Glg. 6.6 soll dann analog zu [Heinrich u. a., 2014; Steinbrücker u. a., 2009] in wenigen Iterationen durch alternierende Optimierung von \tilde{u}_t und \tilde{w}_t minimiert werden, wobei $\kappa \in \mathbb{R}_{\geq 0}$ ausgehend von $\kappa = 0$ in jeder Iteration erhöht wird, sodass die beiden Felder gegeneinander konvergieren ($\tilde{u}_t = \tilde{w}_t$, wenn $\kappa \rightarrow \infty$). Für \tilde{u}_t ergibt sich bei gegebenem \tilde{w}_t das Minimierungsproblem

$$\min_{\tilde{u}_t} \sum_{\mathbf{x} \in \Omega_{fp}} \left(\mathcal{D}(I_R, I_t, \tilde{u}_t, \mathbf{x}) + \kappa \|\tilde{u}_t(\mathbf{x}) - \tilde{w}_t(\mathbf{x})\|_2^2 \right) . \quad (6.8)$$

Dieses Problem lässt sich durch die Unabhängigkeit der einzelnen Punkte $\mathbf{x} \in \Omega_{fp}$ und der Diskretisierung des Raumes der möglichen Verschiebungen global optimal mittels Block-Matching lösen. Sei hierfür $D_{\mathbf{x}} = \{0, \pm\delta_d, \pm 2\delta_d, \dots, \pm n_d \delta_d\}^3 \subset \mathbb{R}^3$ die diskrete Menge der zulässigen Verschiebungsvektoren für einen Punkt $\mathbf{x} \in \Omega_{fp}$ mit Abtastdistanz δ_d und n_d als Anzahl der Abtastpunkte pro Richtung. Beim Block-Matching wird dann für jeden Punkt $\mathbf{x} \in \Omega_{fp}$ (unter geringfügiger Vernachlässigung der Notation) durch erschöpfende Suche die optimale Verschiebung $\mathbf{d} \in D_{\mathbf{x}}$ bestimmt:

$$\tilde{u}_t(\mathbf{x}) \leftarrow \arg \min_{\mathbf{d} \in D_{\mathbf{x}}} \mathcal{D}(I_R, I_t, \mathbf{d}, \mathbf{x}) + \kappa \|\mathbf{d} - \tilde{w}_t(\mathbf{x})\|_2^2 . \quad (6.9)$$

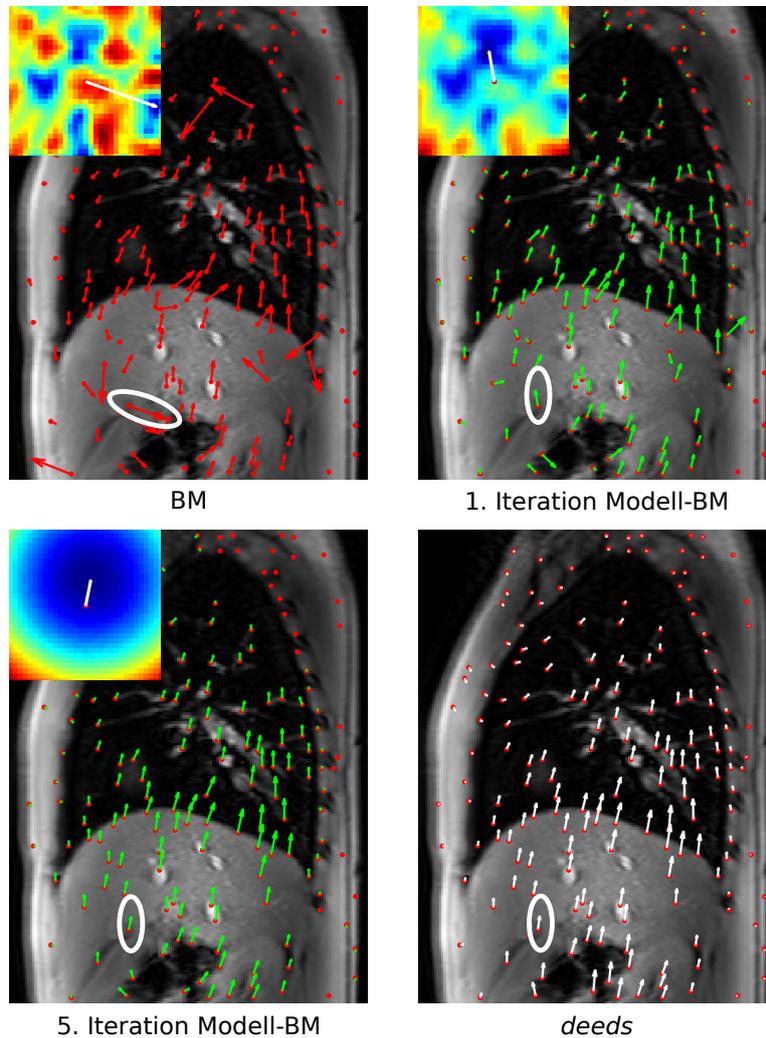


Abb. 6.4: Illustration der dünnbesetzten Verschiebungsfelder bei einer Registrierung zweier 2D-MRT-Bilder mit reinem Block-Matching (BM; Glg. 6.9 mit $\kappa = 0$) und nach der 1./5. Iteration des modellgeführten Block-Matching-Verfahrens (Modell-BM). Als Vergleich dienen die durch pixelweise Registrierung mit *deeds* [Heinrich u. a., 2013a] generierten Verschiebungsvektoren (letztes Bild). In der linken oberen Ecke ist für den markierten Merkmalspunkt in der Leber zudem das Kostenbild bei Auswertung von Glg. 6.9 für alle Verschiebungsvektoren $D_{\mathbf{x}}$ (warm/kalt: hohe/niedrige Kosten; Pfeil: globales Minimum) visualisiert. BM erzeugt sehr viele Verschiebungsvektoren, die nicht konsistent mit der generellen Atembewegung sind. Modell-BM ist in der Lage deren Ausrichtung iterativ zu korrigieren, was auch im Kostenbild sichtbar ist.

Für \tilde{w}_t ergibt sich dann bei festem \tilde{u}_t in Vektorschreibweise mit $\tilde{\mathbf{u}}_t = \text{vec}(\tilde{u}_t)$ und $\tilde{\mathbf{w}}_t = \text{vec}(\tilde{w}_t)$ entsprechend das Minimierungsproblem

$$\min_{\tilde{\mathbf{w}}_t, \mathbf{b}} \kappa \|\tilde{\mathbf{u}}_t - \tilde{\mathbf{w}}_t\|_2^2 + \alpha \|\tilde{\mathbf{P}}\mathbf{S}\mathbf{b} - (\tilde{\mathbf{w}}_t - \tilde{\mathbf{u}})\|_2^2, \quad (6.10)$$

welches sich global optimal mittels

$$\begin{aligned} \tilde{\mathbf{w}}_t &\leftarrow \frac{\kappa}{\kappa + \alpha} \tilde{\mathbf{u}}_t + \frac{\alpha}{\kappa + \alpha} (\tilde{\mathbf{u}} + \tilde{\mathbf{P}}\mathbf{S}\mathbf{b}) \\ \text{mit } \mathbf{b} &= \arg \min_{\mathbf{b}} \|\tilde{\mathbf{P}}\mathbf{S}\mathbf{b} - (\tilde{\mathbf{u}}_t - \tilde{\mathbf{u}})\|_2^2 \end{aligned} \quad (6.11)$$

lösen lässt. Die alternierende Minimierung von Glg. 6.6 führt bei entsprechender Wahl von α dazu, dass das resultierende Feld \tilde{u}_t nicht notwendigerweise perfekt durch das Bewegungsmodell abgebildet werden kann. Diese Eigenschaft ist einerseits wünschenswert im Sinne einer erhöhten Flexibilität und das angestrebte Resultat bei Einführung einer derartigen flexiblen Einschränkung (soft constraint), sorgt aber andererseits dafür, dass bei einer modellbasierten Interpolation des eigentlich gesuchten dichten Feldes u_t basierend auf \tilde{u}_t mittels Glg. 6.5 diese Abweichungen weiterhin nicht durch das Modell abgebildet werden können und sie deshalb nicht im finalen dichten Feld repräsentiert sein werden. Da das eigentliche Ziel die Schätzung eines dichten Feldes auf Basis partieller Informationen ist, wird Glg. 6.11 hier im Sinne einer starken Vereinfachung deshalb durch

$$\begin{aligned} \tilde{\mathbf{w}}_t &\leftarrow \tilde{\mathbf{u}} + \tilde{\mathbf{P}}\mathbf{S}\mathbf{b} \\ \text{mit } \mathbf{b} &= \arg \min_{\mathbf{b}} \|\tilde{\mathbf{P}}\mathbf{S}\mathbf{b} - (\tilde{\mathbf{u}}_t - \tilde{\mathbf{u}})\|_2^2 \end{aligned} \quad (6.12)$$

approximiert. Diese Vereinfachung erzielt in der Praxis gute Ergebnisse (siehe Abb. 6.4) und hat zur Folge, dass das finale dünnbesetzte Verschiebungsfeld $\tilde{u}_t = \tilde{w}_t$ auf die Menge der durch das Modell generierbaren (Teil-)Felder beschränkt ist. Diese Approximation ersetzt den additiven Regularisierungsterm durch eine Restriktion der erlaubten Lösungen auf durch das Modell generierbare Felder. Dies führt dazu, dass der Parametervektor \mathbf{b} von $\tilde{u}_t = \tilde{w}_t$ ebenso das gesuchte dichte Verschiebungsfeld u_t generiert, ohne dass es abseits von Diskretisierungseffekten Inkonsistenzen zwischen den durch das Block-Matching geschätzten Verschiebungen und dem Rest des Feldes gibt. In der Praxis erfolgt die Bestimmung von \mathbf{b} in Glg. 6.12 in dieser Arbeit mit zusätzlicher Thikonov-Regularisierung ($\eta > 0$ in Glg. 6.5).

6.3.2 Zusammenfassung

Das präsentierte Verfahren besteht aus zwei Schritten, die iterativ und alternierend ausgeführt werden und zusammen ein modellgeführtes Block-Matching-Verfahren ergeben (siehe Abb. 6.5):

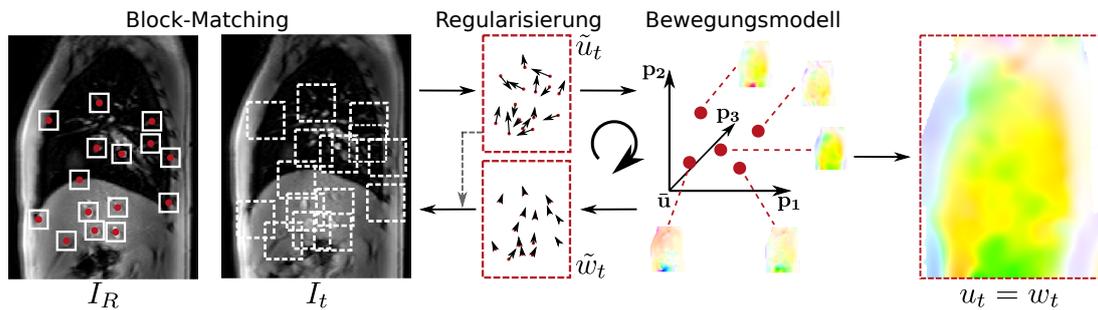


Abb. 6.5: Schematische Darstellung des modellgeführten Block-Matching-Verfahrens. Abwechselnd werden durch das Block-Matching und das Bewegungsmodell dünnbesetzte Verschiebungsfelder \tilde{u}_t und \tilde{w}_t erzeugt. Hierbei beeinflusst \tilde{w}_t iterativ die Bestimmung der Verschiebungsvektoren in \tilde{u}_t . Abschließend wird das Bewegungsmodell genutzt, um ein dichtes Verschiebungsfeld zu erzeugen.

1. Erzeugung eines optimalen dünnbesetzten Verschiebungsfeldes durch Auswertung des Ähnlichkeitsmaßes (+ Kopplungsterm) mittels Block-Matching (Glg. 6.9).
2. Glättung bzw. Einschränkung des mittels Block-Matching erzeugten Verschiebungsfeldes auf physiologisch plausible Lösungen durch Approximation mithilfe des Bewegungsmodells (Glg. 6.12).

Schritt 1 entspricht dem korrespondierenden Schritt im Ausgangsverfahren in [Heinrich u. a., 2014], welches allerdings auf dichten Verschiebungsfeldern arbeitet. Schritt 2 besteht in [Heinrich u. a., 2014] aus einer Filterung des aus dem 1. Schritt resultierenden Feldes mit einem Gauß-Filter, um eine diffusive Regularisierung zu realisieren. Dies wird in dem hier skizzierten Verfahren durch die modellbasierte Approximation des Feldes ersetzt. Während in [Heinrich u. a., 2014] der Gewichtungsfaktor α der Regularisierung die Glättung des Feldes reguliert, so gibt es in der hier eingesetzten Approximation keine direkte Entsprechung mehr. Die Glattheit eines durch das Modell generierten Feldes wird implizit u. a. durch die Glattheit der Trainingsdaten und k mitbestimmt.

Beide Schritte des Verfahrens lassen sich sehr effizient implementieren. Die Auswertung des Ähnlichkeitsmaßes für alle Punkte und alle möglichen Verschiebungsvektoren muss analog zu [Heinrich u. a., 2014] nur einmal durchgeführt werden. Das resultierende Kostenvolumen kann für jeden Punkt gespeichert und in jeder Iteration in Kombination mit dem Kopplungsterm neu ausgewertet werden.

Im Vergleich zu Klinder u. a. [2012] und Preiswerk u. a. [2014], die das Modell nur zur finalen Regularisierung/Interpolation eines anderweitig bestimmten dünnbesetzten Verschiebungsfeldes einsetzen, integriert das hier vorgestellte Verfahren die Schätzung des dünnbesetzten Feldes und die modellbasierte Regularisierung/Interpolation in einem einzelnen Ansatz. Durch diese Kopplung kann das modellbasierte Vorwissen schon

bei der Schätzung des dünnbesetzten Feldes eingesetzt werden, was eine erhöhte Robustheit verglichen zu entkoppelten Verfahren erwarten lässt.

6.4 Registrierung mit einem Patch-AAM

Prinzipiell geht es bei dem in Kapitel 6.2 definierten Registrierungsproblem um die Lokalisierung der im Referenzbild definierten Merkmalspunkte im intrainerventionellen Zielbild mit anschließender Interpolation eines dichten Verschiebungsfeldes aus den Punktkorrespondenzen. Das Problem der Punktlokalisierung in neuen Bildern tritt in der medizinischen Bildverarbeitung oftmals auch im Rahmen kontur- bzw. oberflächenbasierter Segmentierungsverfahren (z. B. mit statistischen Formmodellen [Heimann u. a., 2009]) auf und ist Hauptelement der verwandten Verfahren zur Landmarkenerkennung in Gesichtern in der Computer Vision (siehe z. B. Überblick in [Zafeiriou u. a., 2015]). Bei dem modellgeführten Block-Matching-Verfahren in Kapitel 6.3 wird jeder Merkmalspunkt einzig durch die Bildinformationen seiner unmittelbaren Umgebung im Referenzbild repräsentiert, wohingegen die potenziellen Merkmalspunktverschiebungen bzw. neuen -positionen durch ein globales, statistisches Modell beschrieben werden, welches die Lösungen des Registrierungs-/Lokalisierungsproblems analog zu statistischen Formmodellen beschränkt und aus Trainingsdaten erzeugt wird. Von den Bildinformationen der Trainingsbilder wird abseits des Referenzbildes kein Gebrauch gemacht.

In diesem Teilkapitel wird ein für die Landmarkenerkennung in Gesichtern entwickeltes patchbasiertes Active Appearance Model (AAM) [Tzimiropoulos u. a., 2014] für die Schätzung dichter Verschiebungsfelder unter Verwendung partieller Bildinformationen adaptiert, welches neben den Bewegungsinformationen der Trainingsdaten auch die zugehörigen Bildinformationen nutzt, um den Registrierungsprozess robuster zu gestalten. Ein AAM nach Cootes u. a. [2001] besteht neben einem globalen statistischen Formmodell, welches die zulässigen Merkmalspunktkonfigurationen modelliert, auch aus einem globalen statistischen Erscheinungsmodell. Das Erscheinungsmodell modelliert die zu erwartenden Variationen der Bildinformationen in der Merkmalspunktumgebung oder innerhalb des kompletten durch die Punkte repräsentierten Objektes (z. B. Gesicht oder Organ). Beide Modelle werden PCA-basiert aus denselben Trainingsdaten erzeugt und während der Merkmalspunktdetektion in einem neuen Bild zusammen als Vorwissen eingesetzt. Diese Kombination verspricht prinzipiell eine höhere Robustheit als die alleinige Nutzung von Formmodellen, führt aber auch zu einem komplexeren Optimierungsproblem, für das es lange nur heuristische [Cootes u. a., 2001] oder ineffiziente [Matthews u. a., 2004] Lösungsverfahren gab. Dies hat sich zuletzt insbesondere durch methodische Entwicklungen im Bereich der Computer Vision (patch-basierte Repräsentation der Bildinformationen [Tzimiropoulos u. a., 2014], effiziente Optimierungsverfahren [Tzimiropoulos u. a., 2017]) entscheidend geändert.

Der Einsatz eines statistischen Erscheinungsmodells zusätzlich zu einem Bewegungsmodell folgt der Annahme, dass hiermit eine robustere Bewegungsschätzung möglich wird als bei der alleinigen Nutzung der Bildinformationen des Referenzbildes analog zu Kapitel 6.3. Dies gilt insbesondere vor dem Hintergrund zeitlich variierender intrainventioneller Bildqualität oder physiologisch bedingter Änderungen der Bildinformationen durch lokale Deformationen. Das patchbasierte AAM aus [Tzimiropoulos u. a., 2014] wird hier aufgrund der sehr guten Ergebnisse bei der Landmarkendetektion in Gesichtern (siehe z. B. [Tzimiropoulos u. a., 2017]) und des erfolgreichen Einsatzes zur Segmentierung von medizinischen Bilddaten durch den Autor dieser Arbeit in [Wilms u. a., 2017b] als Basis gewählt. Im Vergleich zu [Tzimiropoulos u. a., 2014; Wilms u. a., 2017b] sind drei Adaptionen notwendig: (1) Generelle Erweiterung auf 3D-Daten (nur 2D in [Tzimiropoulos u. a., 2014; Wilms u. a., 2017b]). (2) Austausch des Formmodells durch ein Bewegungsmodell. (3) Umgang mit dünnbesetzten Verschiebungsfeldern während der Optimierung.

6.4.1 Methodenbeschreibung

Das hier verwendete Bewegungsmodell entspricht dem in Kapitel 6.2 definierten und in Kapitel 6.3 eingesetzten Modell. Zusätzlich wird auf Basis der in Kapitel 6.2 eingeführten Trainingsdaten bestehend aus 4D-Bilddatensatz $(I_j)_{j=1}^{n_{tp}}$ und Verschiebungsfeldern $(u_j)_{j=1}^{n_{tp}}$ ein statistisches Erscheinungsmodell erzeugt. Hierfür wird angenommen, dass das Referenzbild I_R ein Bild des 4D-Bilddatensatzes ist. Aufbauend auf den Verschiebungsfeldern $(u_j)_{j=1}^{n_{tp}}$ lässt sich für jeden Punkt $\mathbf{x} \in \Omega_{fp}$ in I_R durch $\mathbf{x} + u_j(\mathbf{x})$ die korrespondierende Position in Bild I_j ermitteln. Mit diesen Korrespondenzen können dann mithilfe der in Kapitel 6.2 definierten Funktion zur Merkmals- und Patchextraktion $\theta(\cdot, \cdot)$ Patchvektoren $\mathbf{f}_{\mathbf{x}, I_j} \in \mathbb{R}^{n_{ps}^3 n_{pch}}$ für alle in I_R definierten n_{fp} Merkmalspunkte erzeugt werden, die die jeweilige Punktumgebung in den Trainingsbildern beschreiben. Sei zudem $\mathbf{a}_j \in \mathbb{R}^{n_{av}}$ ($n_{av} = n_{ps}^3 n_{pch} n_{fp}$) der Erscheinungsvektor von Bild I_j , der alle n_{fp} Patchvektoren zusammenfasst.

Basierend auf den Erscheinungsvektoren aller Trainingsbilder $(\mathbf{a}_j)_{j=1}^{n_{tp}}$ kann nun analog zum statistischen Bewegungsmodell in Kapitel 6.2 ein statistisches Erscheinungsmodell erzeugt werden. Sei hierfür $\bar{\mathbf{a}} = \frac{1}{n_{tp}} \sum_{j=1}^{n_{tp}} \mathbf{a}_j$ die mittlere Erscheinung und $\mathbf{Q} \in \mathbb{R}^{n_{av} \times l}$ die Matrix der orthonormalen Eigenvektoren der Kovarianzmatrix $\mathbf{C}_{\mathbf{a}} = \frac{1}{(n_{tp}-1)} \sum_{j=1}^{n_{tp}} (\mathbf{a}_j - \bar{\mathbf{a}})(\mathbf{a}_j - \bar{\mathbf{a}})^T$ mit den l größten Eigenwerten (hier immer 90% der Variabilität). Die Kombination aus $\bar{\mathbf{a}}$ und \mathbf{Q} parametrisiert einen affinen Unterraum des $\mathbb{R}^{n_{av}}$, der plausible Bildinformationskombinationen an den Merkmalspunkten beschreibt. Elemente des Unterraums können erzeugt werden durch:

$$\hat{\mathbf{a}} = \bar{\mathbf{a}} + \mathbf{Q} \mathbf{c} . \quad (6.13)$$

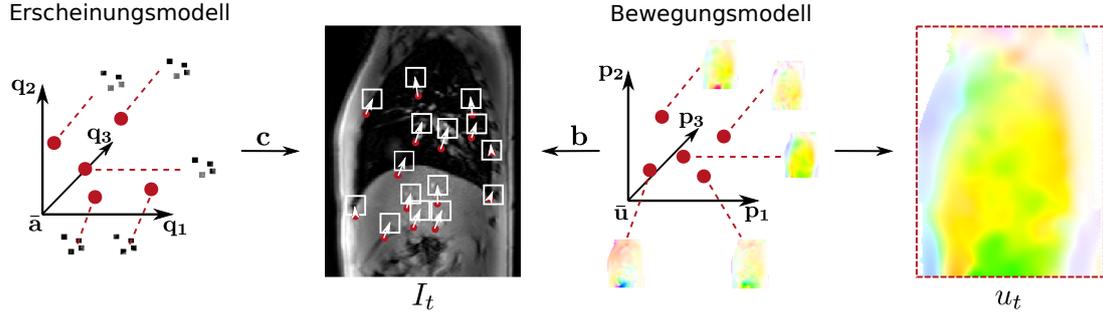


Abb. 6.6: Illustration der Registrierung mit einem Patch-AAM. Es wird für die Merkmalspunkte ein dünnbesetztes Verschiebungsfeld gesucht (Parameter \mathbf{b}), welches sich durch das Bewegungsmodell ausdrücken lässt. Dieses soll gleichzeitig zu einer Kombination von Patches an den durch das Verschiebungsfeld spezifizierten Bildpunkten führen, deren Bildinformationen sich bestmöglich durch das Erscheinungsmodell approximieren lassen (Parameter \mathbf{c}).

In Analogie zu [Tzimiropoulos u. a., 2014; Wilms u. a., 2017b] und basierend auf Glg. 6.1 wird die Kostenfunktion für das zu lösende Registrierungsproblem als

$$E(\mathbf{b}, \mathbf{c}) = \sum_{\mathbf{x} \in \Omega_{fp}} \|\bar{\mathbf{a}}_{\mathbf{x}} + \mathbf{Q}_{\mathbf{x}}\mathbf{c} - \theta(I_t, \mathbf{x} + \bar{\mathbf{u}}_{\mathbf{x}} + \mathbf{P}_{\mathbf{x}}\mathbf{b})\|_2^2 \quad (6.14)$$

definiert, wobei $\bar{\mathbf{a}}_{\mathbf{x}} \in \mathbb{R}^{av}$, $\mathbf{Q}_{\mathbf{x}} \in \mathbb{R}^{n_{ps}^3 n_{pch} \times l}$, $\bar{\mathbf{u}}_{\mathbf{x}} \in \mathbb{R}^3$ und $\mathbf{P}_{\mathbf{x}} \in \mathbb{R}^{3 \times k}$ nur die zum jeweiligen Merkmalspunkt $\mathbf{x} \in \Omega_{fp}$ korrespondierenden Elemente der gleichnamigen vollen Vektoren/Matrizen enthalten. Im Vergleich zu Glg. 6.1 wird hier der jeweilige Patch des Referenzbildes I_R durch das Erscheinungsmodell (siehe Glg. 6.13) ausgetauscht, wobei das eigentliche Ähnlichkeitsmaß \mathcal{D} aus Glg. 6.2 sonst erhalten bleibt. Die Verschiebung der Merkmalspunkte/das dünnbesetzte Verschiebungsfeld wird zudem direkt durch das Bewegungsmodell parametrisiert (siehe Glg. 6.4²⁵ bzw. $\tilde{u}_t = \text{vec}^{-1}(\tilde{\mathbf{u}} + \tilde{\mathbf{P}}\mathbf{b})$), wodurch der additive Regularisierungsterm \mathcal{S} entfällt. Wie bei der AAM-basierten Segmentierung wird ein Tupel von Modellparametern (\mathbf{b}, \mathbf{c}) gesucht, welches die Kostenfunktion minimiert (siehe Abb. 6.6). Der resultierende Parametervektor \mathbf{b} kann dann genutzt werden, um mithilfe des Bewegungsmodells in Glg. 6.4 das eigentlich gesuchte dichte Verschiebungsfeld u_t zu generieren.

Zur Minimierung von Glg. 6.14 wird hier das in [Tzimiropoulos u. a., 2017] vorgeschlagene Fast-SIC-Verfahren genutzt, welches eine durch Umsortierung der zentralen Schritte (Linearisierung und Projektion) effizienter gestaltete Variante des klassischen Simultaneous Inverse Compositional (SIC)-Algorithmus [Baker u. a., 2004; Matthews u. a., 2004] ist. Fast-SIC löst das in Glg. 6.14 definierte nichtlineare Optimierungspro-

²⁵Zur Vereinfachung hier ohne \mathbf{S} .

blem im Sinne eines Gauß-Newton-Verfahrens durch iterative, alternierende Anpassung der Parametervektoren mit $\mathbf{b} \leftarrow \mathbf{b} + \Delta\mathbf{b}$ und $\mathbf{c} \leftarrow \mathbf{c} + \Delta\mathbf{c}$. Hierfür sei

$$(\Delta\mathbf{b}, \Delta\mathbf{c}) = \arg \min_{\Delta\mathbf{b}, \Delta\mathbf{c}} \sum_{\mathbf{x} \in \Omega_{fp}} \|\bar{\mathbf{a}}_{\mathbf{x}} + \mathbf{Q}_{\mathbf{x}}\mathbf{c} - \mathbf{Q}_{\mathbf{x}}\Delta\mathbf{c} - \mathbf{J}_{\mathbf{x}}\Delta\mathbf{b} - \theta(I_t, \mathbf{x} + \bar{\mathbf{u}}_{\mathbf{x}} + \mathbf{P}_{\mathbf{x}}\mathbf{b})\|_2^2 \quad (6.15)$$

das Optimierungsproblem bezüglich $(\Delta\mathbf{b}, \Delta\mathbf{c})$, welches u. a. nach Linearisierung des Erscheinungsmodells für $\mathbf{b} = \mathbf{0}$ aus Glg. 6.14 folgt (siehe [Tzimiropoulos u. a., 2014, 2017]²⁶). Hierbei bezeichnet $\mathbf{J}_{\mathbf{x}} = \nabla(\bar{\mathbf{a}}_{\mathbf{x}} + \mathbf{Q}_{\mathbf{x}}\mathbf{c}) \frac{\partial}{\partial \mathbf{b}}(\mathbf{x} + \bar{\mathbf{u}}_{\mathbf{x}} + \mathbf{P}_{\mathbf{x}}\mathbf{b})|_{\mathbf{b}=\mathbf{0}} \in \mathbb{R}^{n_{ps}^3 n_{pch} \times k}$ die zum Merkmalspunkt \mathbf{x} korrespondierenden Teile der kompletten Jacobi-Matrix $\mathbf{J} \in \mathbb{R}^{n_{av} \times k}$, welche Erscheinungsänderungen zu Parameteränderungen des Bewegungsmodells in Beziehung setzt (siehe auch [Tzimiropoulos u. a., 2014]). Sei weiterhin $\mathbf{i}_{\mathbf{b}} \in \mathbb{R}^{n_{av}}$ der Erscheinungsvektor, der alle n_{fp} Patchvektoren $\theta(I_t, \mathbf{x} + \bar{\mathbf{u}}_{\mathbf{x}} + \mathbf{P}_{\mathbf{x}}\mathbf{b})$ von I_t zu einem gegebenen Parametervektor \mathbf{b} zusammenfasst. Dann ist nach [Tzimiropoulos u. a., 2014, 2017] die geschlossene Lösung für das Tupel $(\Delta\mathbf{b}, \Delta\mathbf{c})$ gegeben durch

$$\Delta\mathbf{b} = (\mathbf{K}^T \mathbf{K})^{-1} \mathbf{K}(\mathbf{i}_{\mathbf{b}} - \bar{\mathbf{a}}) \quad \text{mit } \mathbf{K} = \mathbf{J} - \mathbf{Q}\mathbf{Q}^T \mathbf{J} \quad (6.16)$$

$$\Delta\mathbf{c} = \mathbf{Q}^T(\mathbf{i}_{\mathbf{b}} - \bar{\mathbf{a}} - \mathbf{Q}\mathbf{c} - \mathbf{J}\Delta\mathbf{b}) . \quad (6.17)$$

6.4.2 Zusammenfassung

Das hier präsentierte Verfahren nutzt die sonst zur robusten modellbasierten Segmentierung mit AAMs eingesetzte Kombination aus Form- bzw. Bewegungsmodell und Erscheinungsmodell für die patchbasierte Registrierung. Hierbei wird die Regularisierung des geschätzten Verschiebungsfeldes durch die explizite Einschränkung auf durch das Bewegungsmodell parametrisierbare Felder realisiert. Ähnlich zum modellgeführten Block-Matching in Kapitel 6.3 wird das resultierende Optimierungsproblem ebenfalls iterativ und alternierend gelöst. Hier wird allerdings ein ableitungsbasiertes Verfahren eingesetzt, was die schon in Kapitel 6.3 diskutierten potenziellen Nachteile (u. a. Konvergenz gegen lokale Minima und Abhängigkeit von der Initialisierung) hat. Bezüglich der effizienten Umsetzbarkeit des Verfahrens lässt sich festhalten, dass die aufwendigsten Schritte die Berechnung der Jacobi-Matrix \mathbf{J} und die Berechnung der Patchvektoren für $\mathbf{i}_{\mathbf{b}}$ sind. Für \mathbf{J} können allerdings die räumlichen Ableitungen von \mathbf{Q} vorberechnet werden und die Berechnung der Patchvektoren für die einzelnen Merkmalspunkte kann wiederum problemlos parallelisiert werden.

²⁶Die exakte Herleitung des Fast-SIC-Verfahrens ist in [Tzimiropoulos u. a., 2014, 2017] zu finden. Auf eine Replikation wird aus Gründen der Übersichtlichkeit verzichtet. Die einzigen Unterschiede zwischen [Tzimiropoulos u. a., 2014, 2017] und dem hier diskutierten Problem sind die Erweiterung auf 3D und die zusätzliche Punktkoordinate \mathbf{x} durch den Wechsel von einem Form- auf ein Bewegungsmodell. Letztere ist als Konstante allerdings für die Optimierung irrelevant.

6.5 Experimente

Die Evaluation der beiden in diesem Kapitel vorgeschlagenen patch- und modellbasierten Verfahren zur intrainerventionellen Schätzung der Atembewegung verfolgt insbesondere die folgenden drei Ziele:

1. Untersuchung beider Verfahren hinsichtlich erreichbarer Genauigkeit und Verarbeitungsgeschwindigkeit mit starkem Fokus auf MRT-geführte Interventionen. Zudem Vergleich mit einem aktuellen Verfahren aus der Literatur.
2. Untersuchung der Unterschiede beider Verfahren in Bezug auf die erreichte Genauigkeit insbesondere vor dem Hintergrund der unterschiedlichen Optimierungsmethoden und Nutzung der Trainingsdaten.
3. Soweit möglich die Nutzung frei verfügbarer Daten, um die Reproduzierbarkeit und einen zukünftigen studienübergreifenden Ergebnisvergleich zu ermöglichen.

Diese Evaluation teilt einen großen Teil des Aufbaus und Inhalts mit der Evaluation in [Ha u. a., 2019], erweitert diese aber um zwei wichtige Aspekte: (1) Evaluation des Patch-AAMs (in [Ha u. a., 2019] nur modellgeführtes Block-Matching). (2) Vergleich der Nutzung der reinen Grauwertinformationen und des SSC-Deskriptors zur Beschreibung der Bildinformationen eines Patches (in [Ha u. a., 2019] nur SSC-Deskriptor). Zudem gibt es teilweise Unterschiede in der Parameterwahl, der Anzahl der genutzten Landmarken und des Aufbaus der Experimente, was die (geringen) Unterschiede der quantitativen Ergebnisse beider Studien erklärt. Die folgenden beiden Unterkapitel beschreiben die verwendeten Bilddaten und die durchgeführten Experimente näher.

6.5.1 Verwendete Bilddaten

In Ermangelung realer intrainerventionell aufgenommener MRT- oder US-Bilddaten erfolgt die Evaluation in diesem Kapitel anhand von für andere Zwecke akquirierten räumlich-zeitlichen Bilddaten. Diese Bilder werden genutzt, um verschiedene intrainerventionelle Bewegungsschätzungsszenarien zu simulieren. Insgesamt basiert die Evaluation auf acht 2D+t-MRT-Bilddatensätzen, zwei 4D-MRT-Bilddatensätzen und neun 4D-US-Bilddatensätzen aus öffentlichen Datenquellen (siehe auch Abb. 6.7):

- **2D+t-MRT:** Diese thorakalen/abdominalen Bilddaten von insgesamt vier Probanden wurden im Rahmen der Publikation [Baumgartner u. a., 2017] aufgenommen. Insgesamt wurden für jeden Probanden jeweils 30 T1-gewichtete 2D+t-MRT-Bilddatensätze bestehend aus 40 2D-Bildern (sagittale Ausrichtung; freie Atmung) mit einem 3T-Scanner aufgenommen. Die 30 2D-Serien decken in äquidistanten Abständen den kompletten Körper entlang der Querachse (rechts-links) ab. Zusätzlich ist für jede der 30 Querachsenpositionen ein zusätzliches (statisches) 2D-Bild verfügbar,

welches bei maximaler Ausatmung (Breath-Hold) aufgenommen wurde. Jedes 2D-Bild besteht aus 215×288 Pixeln (Pixelgröße: $1.4 \times 1.4 \text{ mm}^2$). Für die Experimente in dieser Arbeit werden für jeden Probanden von den Querachsenpositionen jeweils die gewählt, die in der Mitte des rechten bzw. linken Lungenflügels liegen. Somit stehen für die Evaluation insgesamt acht Datensätze von vier Probanden bestehend aus einem statischen 2D-Bild bei maximaler Ausatmung und einer 2D-Serie mit 40 Aufnahmen bei freier Atmung zu Verfügung. Für die quantitative Evaluation wurden im Rahmen von [Ha u. a., 2019] in jedem der acht statischen 2D-Bilder 10–20 Landmarken manuell in Lunge/Leber ausgewählt und deren korrespondierende Positionen in vier zufällig ausgewählten Bildern der zweiten Hälfte der zeitlichen 2D-Serie bestimmt. Diese Landmarken sind öffentlich verfügbar (siehe [Ha u. a., 2019]).

- **4D-MRT:** Diese thorakalen/abdominalen 4D-MRT-Bilddaten von zwei Personen wurden im Rahmen der Publikation [Boye u. a., 2013] veröffentlicht. Jeder 4D-Datensatz besteht aus 200 3D-Bildern aufgenommen bei freier Atmung mit einer zeitlichen Auflösung von ≈ 300 ms. Jedes 3D-Bild besteht aus 50/52 sagittalen Schichten mit 224×224 Pixeln (Pixelgröße: isotrop $1.2/1.3 \text{ mm}^2$). Der Abstand zwischen den sagittalen Schichten beträgt 5 mm. Für diese Daten wurden im Rahmen von [Ha u. a., 2019] in jeweils ersten 3D-Bild des Datensatzes manuell 27 Landmarken in Lunge und Leber ausgewählt und deren korrespondierende Positionen in weiteren neun/zehn zufällig ausgewählten 3D-Bildern jedes Datensatzes bestimmt.
- **4D-US:** Die neun Datensätze sind eine Teilmenge (SMT-01–SMT-09) der Daten der Challenge on Liver Ultrasound Tracking (CLUST) 2014 und 2015 [De Luca u. a., 2015]. Die Daten wurden durch SINTEF Medical Technology, Trondheim, Norwegen aufgenommen und zeigen Teile der Leber von neun Patienten/Probanden bei freier Atmung. Jeder 4D-Datensatz besteht aus 92–96 3D-Bildern (zeitliche Auflösung: 125 ms) mit einer Größe von $227 \times 229 \times 229$ Voxeln (Voxelgröße: $0.7 \times 0.7 \times 0.7 \text{ mm}^3$). Für diese Daten wurde im Rahmen von [Ha u. a., 2019] im ersten 3D-Bild jedes Datensatzes ein markanter Punkt als Landmarke manuell ausgewählt, dessen Position ebenso in 3–12 weiteren zufällig ausgewählten 3D-Bildern bestimmt wurde.

6.5.2 Design der Experimente

Die einzelnen Experimente bzw. die simulierten Szenarien orientieren sich an den drei verschiedenen Typen von Bilddaten aus Kapitel 6.5.1 und versuchen möglichst realitätsnah eine klinische Behandlungssitzung bei MRT/US-geführter Bestrahlung bzw. Ablation bestehend aus einer Trainingsphase zur Modellbildung und der eigentlichen Anwendungsphase zu imitieren. Die quantitative Auswertung der Genauigkeit der zu untersuchenden Verfahren erfolgt für alle Experimente mittels TRE (siehe Kapitel 2.2.1) anhand der Landmarken. Alle Experimente haben zudem gemein, dass die für

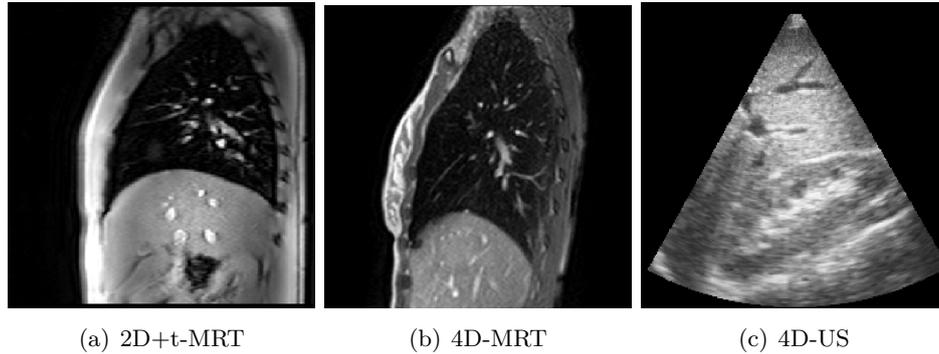


Abb. 6.7: Beispiele für Referenzbilder der für die Evaluation genutzten räumlich-zeitlichen Bilddaten.

das Training der Bewegungsmodelle benötigten Verschiebungsfelder mithilfe von *deeds* [Heinrich u. a., 2013a] mit SSC-Deskriptor [Heinrich u. a., 2013b] erzeugt werden. Die Nutzung dieses diskreten Registrierungsverfahrens ist u. a. in der nachgewiesenen hohen Genauigkeit bei der Atembewegungsschätzung (siehe [Heinrich u. a., 2013a]), der Fähigkeit zum Umgang mit Gleitbewegungen, der einfachen Nutzbarkeit von Bilddeskriptoren zur Ähnlichkeitsbestimmung und der freien Verfügbarkeit begründet.

Insgesamt werden in den drei Experimenten fünf verschiedene Verfahren zur modellbasierten intrainerventionellen Schätzung dichter Verschiebungsfelder verglichen:

- **BM:** Schätzung eines dünnbesetzten Verschiebungsfeldes durch einfaches, nicht-iteratives und patchbasiertes Block-Matching basierend auf den detektierten Merkmalspunkten und Glg. 6.9 mit $\kappa = 0$. Anschließend wird aus dem dünnbesetzten Feld mithilfe von Glg. 6.5 ($\eta = 0$) ein dichtes Feld erzeugt. Dieses Verfahren dient als obere Schranke hinsichtlich des zu erwartenden TRE, da das dünnbesetzte Verschiebungsfeld in den meisten Fällen widersprüchliche Informationen enthält, die durch die implizite modellbasierte Regularisierung bei der Interpolation nicht adäquat ausgeglichen werden.
- **TR-BM:** Gleiches Vorgehen wie bei BM allerdings mit $\eta > 0$ (Thikonov-Regularisierung) bei der Interpolation des dichten Feldes. Hiermit sollten plausiblere Felder bzw. genauere Ergebnisse erreichbar sein, da starke Abweichungen der widersprüchlichen Informationen im dünnbesetzten Feld von der mittleren Bewegung bestraft werden. Dieses Verfahren dient als genereller Vertreter der in Kapitel 6.1 diskutierten entkoppelten Verfahren, bei denen das Bewegungsmodell nur zur finalen Regularisierung/Interpolation eines anderweitig bestimmten dünnbesetzten Verschiebungsfeldes zum Einsatz kommt (u. a. [Klinder u. a., 2012; Preiswerk u. a., 2014]).

- **Modell-BM:** Modellgeführtes Block-Matching wie in Kapitel 6.3 beschrieben. Im Vergleich zur TR-BM wird das Bewegungsmodell hier schon während des Block-Matchings eingesetzt.
- **Patch-AAM:** Registrierung mit einem Patch-AAM (Kapitel 6.4).
- **RealTITracker:** Aktuelles und frei verfügbares Optical Flow-Verfahren aus [Denis de Senneville u. a., 2015] (Closed Source Software; Implementation in C++ und MATLAB) zur intrainerventionellen Atembewegungsschätzung auf der Basis von MRT-Bilddaten. Dieses Verfahren integriert u. a. ein PCA-basiertes Bewegungsmodell zur Regularisierung in den klassischen Optical Flow-Ansatz (siehe [Denis de Senneville u. a., 2015; Zachiu u. a., 2015]). *RealTITracker* ermöglicht die Registrierung vollständiger 3D- bzw. 2D-Bildpaare, bietet aber abseits von der Möglichkeit multiple 2D-Bilder/Schichten (parallel) zu registrieren keine Möglichkeit zum Umgang mit partiellen/unvollständigen Bildinformationen. Für dieses Verfahren werden für alle Experimente die in der Software voreingestellten Parameter verwendet.

Alle patchbasierten Verfahren nutzen in den Experimenten dieselben automatisiert detektierten Merkmalspunkte. Die Anzahl der Merkmalspunkte n_{fp} hat großen Einfluss auf die Genauigkeit der Bewegungsschätzung und die benötigte Rechenzeit (siehe Abb. 6.8). Für diese Evaluation wird n_{fp} so gewählt, dass ein guter Kompromiss zwischen Genauigkeit und Rechenzeit entsteht. Alle Verfahren werden jeweils einmal mit der Verwendung von Grauwertpatches ($n_{pch} = 1$) und bei Einsatz des SSC-Deskriptors ($n_{pch} = 6$ für 2D; $n_{pch} = 12$ für 3D) evaluiert. Der Suchbereich für die Block-Matching-Verfahren (Abtastpunkte n_d) wird jeweils datensatzspezifisch anhand der maximalen Bewegung in den Trainingsdaten festgelegt, wobei die Abtastdistanz δ_d immer der Pixel- bzw. Voxelgröße entspricht. Patch-AAM wird jeweils mit der mittleren Bewegung initialisiert. Informationen über die anderen Parameter sind in den Experimentbeschreibungen angegeben. Alle patchbasierten Verfahren wurden in MATLAB für die CPU-Nutzung implementiert. Als Testsystem kommt ein Rechner mit Intel i7-5775C 3.30 GHz-Prozessor und 32 GB RAM Arbeitsspeicher zum Einsatz. Teilweise werden zur Illustration des Potenzials der vorgeschlagenen Verfahren Rechenzeiten der GPU-Implementierungen aus [Ha u. a., 2019] genannt, deren Entwicklung aber nicht Teil dieses Promotionsprojektes war.

6.5.2.1 Experiment 1: 2D+t-MRT

Die erste Hälfte der 40 2D-Bilder jeder 2D-Serie wird zusammen mit dem statischen 2D-Bild (= Referenzbild) für das Modelltraining des jeweiligen Verfahrens eingesetzt. Die restlichen 20 2D-Bilder imitieren eigentlich intrainerventionell aufgenommene Bilddaten, anhand derer die Atembewegung mit Bezug zum Referenzbild geschätzt werden soll (2D-2D-Registrierung; siehe Abb. 6.9). Die eigentliche bildbasierte Bewegungsschätzung findet allerdings nur für die vier mit Landmarken versehenen 2D-

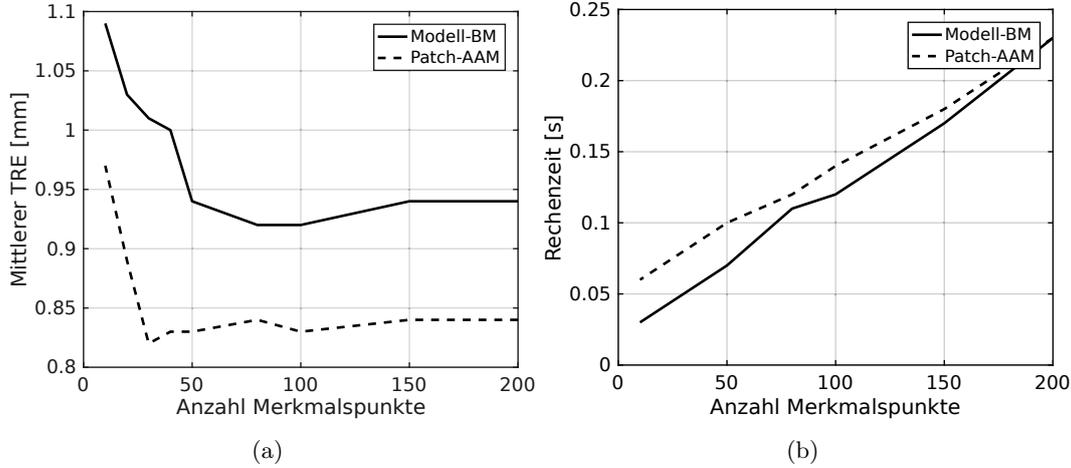


Abb. 6.8: Darstellung des Einflusses der gewählten Anzahl von Merkmalspunkten n_{fp} auf die Genauigkeit der Bewegungsschätzung (mittlerer TRE; (a)) und die Rechenzeit (b) für Modell-BM und Patch-AAM in Experiment 1 (2D+t-MRT-Daten; SSC-Deskriptor; alle Patienten gemittelt). Auffällig: Erst bei $n_{fp} < 50$ kommt es zu einer deutlichen Verschlechterung des TRE für beide Verfahren.

Bilder statt (Testbilder) und wird mittels schwellwertbasierter Segmentierung auf das Körperinnere beschränkt. Für alle patchbasierten Verfahren wurden die Parameter TRE-basiert für einen der acht Datensätze optimiert (bester mittlerer TRE über alle vier Testbilder) und dann ebenso für die anderen sieben Datensätze angewendet. Anzahl der Merkmalspunkte $n_{fp} = 50$ (siehe auch Abb. 6.8); Patchgröße $n_{psz} = 5$ Pixel (Grauwert)/ $n_{psz} = 3$ Pixel (SSC); $\eta = 10$; $\kappa = 10^{\{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\}}$ (= 5 Iterationen); ≤ 20 Iterationen bei Patch-AAM.

6.5.2.2 Experiment 2: 4D-MRT

Mit den beiden 4D-MRT-Datensätzen wird ein Szenario simuliert, bei dem wie bei der MRT-geführten Strahlentherapie durch den MRT-Linearbeschleuniger kurz vor der Intervention 4D-Bilddaten für das Modelltraining aufgenommen werden können, während der Bestrahlung aber intrainerventionell nur wenige 2D-Schichten (z. B. max. drei parallele sagittale Schichten bei ViewRay [Mutic u. a., 2014]) zur Verfügung stehen auf deren Basis dennoch ein dichtes Verschiebungsfeld für die 3D-Daten geschätzt werden soll (3D-2D-Registrierung bzw. 3D-3D-Registrierung mit unvollständigen Daten; siehe Abb. 6.9). Die ersten 66 3D-Bilder jedes 4D-Datensatzes werden für das Modelltraining verwendet (erstes Bild ist Referenzbild). Die restlichen Bilder dienen als Testbilder für die Schätzung dichter Verschiebungsfelder in 3D, wobei die intrainerventionelle Bewegungsschätzung nur für die neun bzw. zehn Bilder mit Landmarken durchgeführt wird.

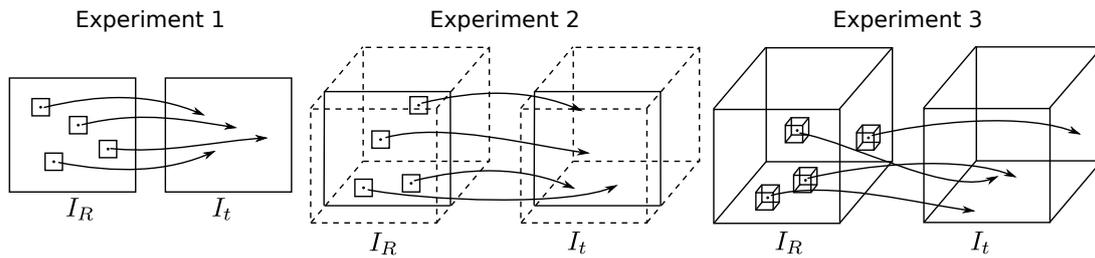


Abb. 6.9: Schematische Darstellung der verschiedenen Registrierungsprobleme in den drei Experimenten. Experiment 1: 2D-2D-Registrierung durchgeführt anhand der ausgewählten 2D-Patches. Experiment 2: 3D-2D-Registrierung bzw. 3D-3D-Registrierung mit unvollständigen Daten, bei der ein dichtes Verschiebungsfeld für das komplette Volumen geschätzt wird. Das Block-Matching etc. erfolgt aber nur mit 2D-Patches auf den ausgewählten Schichten (hier: nur eine Schicht). Experiment 3: 3D-3D-Registrierung anhand der ausgewählten 3D-Patches.

Die Bewegungsschätzung basiert hier analog zum ViewRay-System auf drei parallelen sagittalen Schichten, die über die beiden Lungenflügel verteilt werden. Für das Block-Matching/Patch-AAM werden 2D-Patches eingesetzt und die Suche der neuen Position eines Merkmalspunktes bezieht sich auf die korrespondierende 2D-Schicht der intrainerventionellen Bilder. Der Vorgang selbst lässt sich somit auch als simultane 2D-2D-Registrierung der korrespondierenden Schichten interpretieren. Da der Abstand der sagittalen Schichten in den 3D-Bildern 5 mm beträgt, wird der TRE des dichten 3D-Verschiebungsfeldes nur für die Bewegung innerhalb der Schichten ausgewertet, da die mittels *deeds* erzeugten Trainingsfelder die Bewegung zwischen den Schichten aufgrund der Schichtdicke nur sehr ungenau abbilden. Für alle patchbasierten Verfahren wurden die Parameter TRE-basiert für einen der beiden Datensätze optimiert (bester mittlerer TRE über alle Testbilder) und dann für den anderen Datensatz übernommen. Anzahl der Merkmalspunkte $n_{fp} = 110$; Patchgröße $n_{psz} = 5$ Pixel (Grauwert)/ $n_{psz} = 3$ Pixel (SSC); $\eta = 10$; $\kappa = 10^{\{-1.5, -1.4, -1.3, -1.2, -1.1, -1.0\}}$ (= 6 Iterationen); ≤ 12 Iterationen bei Patch-AAM. *RealTITracker* wird hier nicht eingesetzt, da das Verfahren nicht in der Lage ist auf Basis einzelner Schichten ein dichtes 3D-Verschiebungsfeld zu erzeugen.

6.5.2.3 Experiment 3: 4D-US

Eine Hälfte der 3D-Bilder jedes 4D-Datensatzes wird für das Modelltraining des jeweiligen Verfahrens eingesetzt. Die restlichen 3D-Bilder imitieren wieder intrainerventionell aufgenommene Bilddaten anhand derer die Atembewegung mit Bezug zum Referenzbild (erstes 3D-Bild des 4D-Datensatzes) geschätzt werden soll (3D-3D-Registrierung; siehe Abb. 6.9). Die eigentliche bildbasierte Bewegungsschätzung erfolgt allerdings nur für die mit Landmarken versehenen Bilder. Für Modell-BM wird allerdings aufgrund der hohen zeitlichen Auflösung der Daten in Kombination mit den großen

Bewegungen zusätzlich die Integration des additiven zeitlichen Regularisierungsterms $\kappa_2 \|\mathbf{d} - \tilde{u}_{t-1}(\mathbf{x})\|_2^2$ in Glg. 6.9 getestet, der zusätzlich die Abweichung vom zeitlichen Vorgängerfeld bestraft. Hierfür müssen alle nicht zum Training verwendeten Bilder ausgewertet werden. Für alle patchbasierten Verfahren wurden die Parameter TRE-basiert für einen der neun Datensätze optimiert (bester mittlerer TRE über alle Testbilder) und dann ebenso für die anderen acht Datensätze angewendet. Anzahl der Merkmalspunkte $n_{fp} = 70$; Patchgröße Block-Matching $n_{psz} = 7$ Pixel (Grauwert)/ $n_{psz} = 7$ Voxel (SSC); Patchgröße Patch-AAM $n_{psz} = 7$ Voxel (Grauwert)/ $n_{psz} = 5$ Voxel (SSC); $\eta = 10$; $\kappa = 10^{\{-1.70, -1.22, -0.74, -0.26, 0.22, 0.7\}}$ (= 6 Iterationen); ≤ 15 Iterationen bei Patch-AAM; $\kappa_2 = \{0.50, 0.25, 0, 0, 0, 0\}$. *RealTITracker* wird hier nicht eingesetzt, da die Entwickler das Verfahren nur für die MRT-basierte Bewegungsschätzung entwickelt haben.

6.6 Ergebnisse

Tabelle 6.1 fasst die Ergebnisse der drei Experimente kompakt zusammen. Zusätzlich enthält Tab. 6.1 für Experiment 1 die Ergebnisse einer surrogatbasierten Bewegungsschätzung mit einem (nichtdiffeomorphen) Korrespondenzmodell nach Kapitel 2 als Vergleichsgröße ($n_{sur} = 2$; 1D-Spirometer+Ableitung; Surrogatsignal simuliert anhand von Lungensegmentierungen). Alle aufgeführten Ergebnisse in Tab. 6.1 sind Mittelwerte der (patientenspezifischen) Einzelergebnisse. Die statistische Signifikanz der Unterschiede wird durch gepaarte t-Tests (zweiseitig; Signifikanzniveau: 5%) auf Basis der Einzelergebnisse ermittelt. Da für Experiment 2 nur zwei Patienten zur Verfügung stehen, werden die gepaarten t-Tests hier jeweils pro Patient auf die patientenspezifischen Ergebnisse der Testbilder angewendet. Eine statistische Signifikanz der Unterschiede wird dann nur angenommen, wenn beide Tests dies anzeigen.

Die Ergebnisse der surrogatbasierten Bewegungsschätzung bestätigen zumindest im Vergleich mit den Ergebnissen von Modell-BM und Patch-AAM die Grundhypothese dieses Kapitels (Einsatz von Bilddaten erhöht Genauigkeit der intrainerventionellen Bewegungsschätzung) experimentell. Die Ergebnisse der Experimente bestätigen zudem für die Block-Matching-Verfahren die Hypothese, dass Modell-BM durch die Integration des Bewegungsmodells in den Block-Matching-Prozess Genauigkeitsvorteile gegenüber den entkoppelten Verfahren (BM und TR-BM) hat. Die Unterschiede zwischen (TR-)BM und Modell-BM sind generell statistisch signifikant ($p < 0.05$) bei ähnlichen Rechenzeiten. Einzige Ausnahme bilden hierbei die Ergebnisse zwischen Modell-BM und TR-BM von Experiment 3 bei der Verwendung von Grauwertpatches ($p = 0.19$ bzw. $p = 0.13$ bei zeitlicher Regularisierung). Die zusätzliche Integration der zeitlichen Regularisierung bei Modell-BM für Experiment 3 zeigt zwar eine deutliche Reduktion des mittleren TRE über alle Patienten gegenüber der zeitlich unregulierten Variante (z. B. SSC 2.02 mm vs. 1.73 mm), die Verbesserungen sind aber über

Tabelle 6.1: Ergebnisse zu den drei Experimenten für die untersuchten Verfahren angegeben als Mittelwert \pm Standardabweichung des mittleren TRE. Die Ergebnisse sind jeweils für die Nutzung von Grauwertpatches und dem SSC-Deskriptor zur Ähnlichkeitsbestimmung aufgeführt. Die mittlere Rechenzeit pro Einzelbild wird in Klammern hinter dem TRE-Wert angegeben. Für Modell-BM ist bei Experiment 3 zusätzlich der TRE bei Nutzung des zeitlichen Regularisierungsterms angegeben (zusätzliche Zeile). Siehe Text für weitere Erläuterungen.

	Experiment 1 2D+t-MRT	Experiment 2 4D-MRT	Experiment 3 4D-US
Bewegungsschätzung			
Mittlere Bewegung	6.04 \pm 2.52	3.23 \pm 1.89	5.97 \pm 2.77
<i>deeds</i>	0.77 \pm 0.19 (9.150 s)	1.05 \pm 0.08 (62.541 s)	1.71 \pm 1.73 (66.230 s)
<i>RealTITracker</i>	0.99 \pm 0.45 (0.071 s)	-/-	-/-
<i>Korrespondenzmodell</i>	1.51 \pm 0.36 (<0.001 s)	-/-	-/-
<i>Grauwertpatches</i>			
BM	2.35 \pm 0.77 (0.027 s)	8.61 \pm 1.30 (0.111 s)	3.06 \pm 2.27 (0.244 s)
TR-BM	1.88 \pm 0.64 (0.027 s)	3.12 \pm 1.77 (0.112 s)	2.60 \pm 1.97 (0.245 s)
Modell-BM	1.02 \pm 0.23 (0.033 s)	1.21 \pm 0.23 (0.124 s)	2.02 \pm 1.94/ 1.73 \pm 1.56 (0.257 s)
Patch-AAM	0.83 \pm 0.23 (0.031 s)	1.13 \pm 0.08 (0.084 s)	2.01 \pm 1.93 (0.419 s)
<i>SSC-Deskriptor</i>			
BM	2.18 \pm 0.59 (0.070 s)	3.36 \pm 0.14 (0.160 s)	2.80 \pm 1.95 (3.598 s)
TR-BM	1.77 \pm 0.55 (0.074 s)	1.66 \pm 0.25 (0.161 s)	2.73 \pm 2.02 (3.597 s)
Modell-BM	0.94 \pm 0.21 (0.076 s)	1.13 \pm 0.12 (0.175 s)	2.08 \pm 1.60/ 1.71 \pm 1.26 (3.612 s)
Patch-AAM	0.85 \pm 0.30 (0.098 s)	1.11 \pm 0.11 (0.183 s)	2.20 \pm 2.17 (2.352 s)

alle Patienten gesehen nicht statistisch signifikant (SSC: $p = 0.11$; Grauwert: $p = 0.25$). Bei einer detaillierten Analyse der Ergebnisse zeigt sich, dass durch die zeitliche Regularisierung im Einzelfall allerdings drastische Ausreißer bezüglich der Genauigkeit der Bewegungsschätzung für einzelne Bilder verhindert werden können.

Bei Vergleich von Modell-BM und Patch-AAM zeigen sich über alle Experimente hinweg relativ ähnliche Ergebnisse. Für Experiment 1 ergeben sich allerdings für beide Patchvarianten signifikant bessere Ergebnisse für Patch-AMM als für Modell-BM (z. B. Grauwert: 0.83 mm vs. 1.02 mm; $p < 0.05$). Die Unterschiede zwischen diesen beiden Verfahren in den anderen beiden Experimenten sind nicht statistisch signifikant. Aus diesem Grund sind generelle Aussagen hinsichtlich der praktischen Auswirkungen der spezifischen Vor- oder Nachteile der einzelnen Verfahren schwierig, da auch die Rechenzeiten der beiden Verfahren in ähnlichen Größenordnungen liegen. Für Experiment 1 werden im Mittel 9.25 ± 2.34 Eigenvektoren für das Bewegungsmodell und

15.25 ± 0.53 (Grauwert) bzw. 13.50 ± 0.53 (SSC) Eigenvektoren für das Erscheinungsmodell eingesetzt.

Die Verwendung des SSC-Deskriptors oder der Einsatz von einfachen Grauwertpatches hat für Modell-BM und Patch-AAM nur einen sehr geringen Einfluss auf die Genauigkeit. Bis auf den Unterschied in Experiment 1 für Modell-BM (1.02 mm vs. 0.94 mm; $p < 0.05$) sind alle anderen Unterschiede zwischen diesen beiden Patchvarianten bei gleicher Bewegungsschätzungsmethode (Modell-BM bzw. Patch-AAM) nicht signifikant. Es ist allerdings anzumerken, dass sich für die SSC-Varianten durch die zusätzliche Berechnung des Deskriptors und die zusätzlichen Bildinformationen teilweise drastische Erhöhungen der Rechenzeit (bis zu Faktor 15) im Vergleich zur Grauwertvariante ergeben.

Bei direktem Vergleich von Modell-BM und Patch-AAM mit *deeds* zeigt sich, dass zumindest Patch-AAM für Experiment 1 und 3 eine vergleichbare Genauigkeit wie das klassische Registrierungsverfahren erreicht (alle Unterschiede nicht signifikant; z. B. Exp. 1 bei Grauwert 0.83 mm vs. 0.77 mm; $p = 0.11$), bei gleichzeitig deutlicher Reduktion des Rechenaufwands (zumeist Faktor 100 und mehr). Für Experiment 3 sind zudem auch die Unterschiede zwischen Modell-BM und *deeds* nicht signifikant und in Experiment 2 ist *deeds* zwar signifikant besser als beide Vergleichsverfahren, die Unterschiede sind aber relativ gering.

Die Ähnlichkeit der Ergebnisse von Modell-BM und Patch-AAM zu denen von *RealTITracker* für Experiment 1 (alle Unterschiede nicht statistisch signifikant) zeigt, dass die beiden in diesem Kapitel präsentierten Verfahren in der Lage sind, eine Genauigkeit vergleichbar mit aktuellen Verfahren aus der Literatur zu erreichen. Im Vergleich zum *RealTITracker* sind die hier präsentierten Verfahren allerdings in der Lage die Gleitbewegungen besser abzubilden (siehe Abb. 6.10). Weiterhin ist auffällig, dass die Rechenzeit bei Verwendung der Grauwertvarianten der hier implementierten Verfahren deutlich geringer ist als die der öffentlich verfügbaren Implementierung von *RealTITracker*. Generell lässt sich bezüglich der Rechenzeit der implementierten Verfahren sagen, dass die mittleren Rechenzeiten in Experiment 1 alle unterhalb 100 ms für eine 2D-2D-Registrierung liegen. Bei der Nutzung von Gauwertpatches liegen die mittleren Rechenzeiten für dieses Szenario sogar unterhalb von 35 ms und damit in einem ähnlichen Bereich wie andere aktuelle Verfahren für die 2D-2D-Registrierung (z. B. [Seregini u. a., 2018]: ≈ 50 ms; [Zachiu u. a., 2015]: ≈ 25 ms; [Denis de Senneville u. a., 2015]: ≈ 30 ms; alles CPU-Implementierungen). Auch die Rechenzeiten für die 3D-2D-Registrierung in Experiment 2 übersteigen im Mittel die 100 ms, insbesondere für die Grauwertvarianten, (wenn überhaupt) nur marginal. Als Vergleichswert aus der Literatur seien hier die 3 s in [Stemkens u. a., 2016] genannt. Die Rechenzeiten für die 3D-3D-Registrierung in Experiment 3 liegen für die Grauwertvarianten unterhalb von 500 ms und sind grob mit denen des aktuellen 3D-3D-US-Registrierungsverfahrens in [Royer u. a., 2017] (350 ms) vergleichbar.

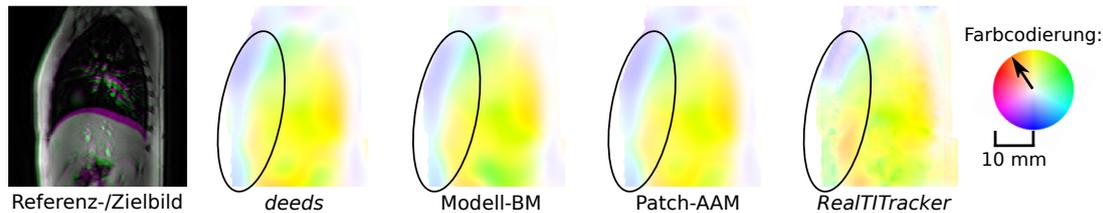


Abb. 6.10: Farbcodierte Darstellung der durch die entsprechenden Verfahren erzeugten Verschiebungsfelder, die zur Kompensation der Atembewegung zwischen zwei 2D-MRT-Bildern (ganz links; dargestellt durch Überlagerung von Referenz- und Zielbild) geschätzt wurden. Die markierte Region in den Verschiebungsfeldern zeigt die Fähigkeit von Modell-BM und Patch-AAM (und *deeds*) zur deutlich besseren Behandlung von Gleitbewegungen zwischen Lunge/Leber und den umliegenden Strukturen im Vergleich zum Verschiebungsfeld von *RealTITracker*, welches eine physiologisch unplausible starke Glättung dieser Übergänge zeigt.

Bei Beurteilung der Rechenzeiten hinsichtlich der Echtzeitfähigkeit der präsentierten Verfahren ist zu beachten, dass durch eine optimierte GPU-Implementierung (siehe Experimente in [Ha u. a., 2019]) eine Reduktion der Rechenzeiten um den Faktor 10 und mehr durchaus realistisch ist. Neben der Optimierung der Implementation ist die Rechenzeit u. a. auch von der verwendeten Anzahl von Merkmalspunkten abhängig (siehe Abb. 6.8), sodass auch hier u.U. ein patientenspezifisches Optimierungspotenzial hinsichtlich der Rechenzeit besteht.

6.7 Zusammenfassung und Diskussion

Moderne Systeme der Strahlentherapie (MR-Linearbeschleuniger) und der Ultraschallablation (HIFU) ermöglichen seit einiger Zeit die intrainterventionelle Akquisition von Bilddaten, die u. a. zur Schätzung und letztlich Kompensation der Atembewegung in diesen Therapieprozessen eingesetzt werden können. Durch die technischen Limitierungen der Systeme und die gleichzeitigen Erfordernisse der Therapieprozesse ergeben sich zwei spezifische Herausforderungen für die eingesetzten Registrierungsverfahren, die klassische Verfahren nicht erfüllen: (1) Echtzeitfähigkeit und (2) flexibler Umgang mit unvollständigen Daten (z. B. Registrierung von 3D-Bild und multiplen 2D-Bildern).

In diesem Kapitel wurden zwei neue patch- und modellbasierte Registrierungsverfahren vorgeschlagen, die diesen Herausforderungen gerecht werden und eine ähnliche Genauigkeit und Geschwindigkeit wie verfügbare Verfahren aus der Literatur erreichen, aber deutlich vielseitiger einsetzbar sind. Durch die patchbasierte und auf wenige Merkmalspunkte beschränkte Registrierung ergibt sich ein erheblicher Geschwindigkeitsvorteil gegenüber klassischen Registrierungsverfahren und diese Vorgehensweise erlaubt zudem den einfachen Umgang mit unvollständigen Daten. Um auf Basis von wenigen Merkmalspunkten trotzdem dichte Verschiebungsfelder schätzen zu können,

kommt in beiden Verfahren ein statistisches Bewegungsmodell zum Einsatz. Dieses Bewegungsmodell wird mittels PCA aus patientenspezifischen Trainingsdaten erzeugt und parametrisiert den Raum plausibler Verschiebungsfelder des Patienten. Es dient damit einerseits als patientenspezifischer Interpolator zur Generierung dichter Felder auf Basis von Merkmalspunktkorrespondenzen und wird andererseits auch als Regularisierer in den Registrierungsprozess eingebunden, um die Robustheit zu erhöhen. Das Bewegungsmodell ermöglicht hierbei die unkomplizierte und korrekte Handhabung von Gleitbewegungen. Beide Verfahren sind zudem so konstruiert, dass neben Grauwertpatches auch beliebige Bilddeskriptoren zur Beschreibung der Bildinformationen eingesetzt werden können, was somit u. a. auch problemlos die multimodale Registrierung mit diesen Verfahren ermöglicht (z. B. US zu MRT; siehe [Preiswerk u. a., 2014]).

Die beiden entwickelten Verfahren lösen das definierte Registrierungsproblem durch unterschiedliche alternierende, iterative Optimierungsstrategien. Das erste Verfahren (Modell-BM) setzt auf ein diskretes Block-Matching, welches durch das Bewegungsmodell (alternierend und iterativ) regularisiert bzw. geführt wird. Diese direkte Integration des Bewegungsmodells in den Registrierungsprozess hat sich im Rahmen der Evaluation als sehr hilfreich und förderlich für die Genauigkeit erwiesen und unterscheidet das präsentierte Verfahren deutlich von anderen Block-Matching-Methoden. Das zweite Verfahren (Patch-AAM) stellt eine Adaption eines für die Landmarkendetektion in Gesichtern entwickelten patchbasierten Active Appearance Models dar, welches ein Gauß-Newton-Verfahren für die Optimierung nutzt. Neben dem Bewegungsmodell kommt in diesem Verfahren auch ein Modell der Bildinformationen (Erscheinungsmodell) zum Einsatz, um den Registrierungsprozess robuster gegenüber Veränderungen der Bilddaten zu gestalten. Optimale Parameter für Bewegungs- und Erscheinungsmodell werden hier ebenfalls alternierend und iterativ ermittelt.

Beide Verfahren haben in der durchgeführten Evaluation in den verschiedenen realitätsnahen Szenarien (2D-2D-Registrierung, 3D-2D-Registrierung und 3D-3D-Registrierung mit vollen Daten) zumeist ähnliche Genauigkeiten bei ebenfalls ähnlichen Rechenzeiten erreicht, wobei sich insbesondere für die 2D-2D-Registrierung signifikante Vorteile für Patch-AAM ergeben haben. Die Genauigkeit beider Verfahren hat sich für die 2D-2D-Registrierung als vergleichbar mit einem aktuellen Verfahren für die in-trainterventionelle Bewegungsschätzung aus der Literatur gezeigt. Für alle Szenarien haben sich zudem nur geringe (teilweise unsignifikante) Genauigkeitsunterschiede im Vergleich zu einem klassischen pixelweisen Registrierungsverfahren ergeben. Generell ist hervorzuheben, dass die beiden präsentierten Verfahren diese Genauigkeit bei einer deutlich reduzierten Rechenzeit (Faktor 100 und mehr) erreichen. Im Fall der 2D-2D- und der 3D-2D-Registrierung liegt der TRE im Mittel zudem unterhalb (2D-2D) bzw. im Bereich (3D-2D) der Pixel-/Voxelgröße. Weiterhin hat sich gezeigt, dass zumindest bei der Verwendung von Grauwertpatches beide Verfahren schon bei Nutzung der für diese Arbeit erstellten CPU-Implementierungen auf Standardhardware für viele Szenarien

rien echtzeitfähig sind und die Rechenzeiten durch eine GPU-Implementierung zudem perspektivisch noch drastisch gesenkt werden können.

Auf Basis der Ergebnisse der durchgeführten Evaluation ergibt sich keine eindeutige Empfehlung für oder gegen die Nutzung eines der beiden Verfahren. Aus theoretischer Sicht könnte man allerdings argumentieren, dass das diskrete Block-Matching bei großen Bewegungen Genauigkeitsvorteile haben sollte, da es nicht gegen lokale Minima konvergiert. Hierfür muss allerdings der komplette diskretisierte Suchraum evaluiert werden, was bei kleineren Bewegungen zu Geschwindigkeitsvorteilen für Patch-AAM führen sollte. Zudem sollte sich für Patch-AAM durch den Einsatz des Erscheinungsmodells eine größere Robustheit gegenüber Veränderungen der Bildinformationen zwischen Referenz- und Zielbild zeigen als bei einem reinen Block-Matching. Potenziell lässt sich aber auch Modell-BM um ein Erscheinungsmodell erweitern.

Bei den für die Evaluation genutzten Bilddaten handelt es sich mit Ausnahme der Ultraschalldaten nicht um intrainerventionell aufgenommene Bilder. Kern von Verbesserungen der durchgeführten Evaluation sollte deshalb die Integration von wirklich intrainerventionell akquirierten Daten sein, um beispielsweise die spezifischen Charakteristiken dieser Daten (z. B. nur 0.35T-Scanner bei ViewRay) bei der Auswertung berücksichtigen zu können. Zudem sollten die Verfahren auf zeitlich deutlich längeren Bildsequenzen evaluiert werden. Hierbei sind u. U. deutliche Auswirkungen von intrainfractioneller Bewegungsvariabilität auf die Genauigkeit der Bewegungsschätzung zu erwarten, da diese Bewegungen möglicherweise nicht gut durch das Bewegungsmodell repräsentiert werden können. Dies ist ein entscheidender Nachteil der expliziten Beschränkung der Lösungen des Registrierungsproblems auf durch das Bewegungsmodell generierbare Transformationen. Diesem potenziellen Problem könnte u. U. durch eine intraininterventionelle Anpassung des Modells an diese Variabilität (siehe z. B. [Baumgartner u. a., 2017] als Vorbild) begegnet werden.

Da die statistische Modellierung und das Registrierungsproblem anhand von Verschiebungsfeldern definiert wurden, sollte in Zukunft eine diffeomorphe Formulierung beider Teile beispielsweise im Log-Euklidischen-Framework angestrebt werden, um ähnlich zu den Korrespondenzmodellen in Kapitel 2 eine diffeomorphe Bewegungsschätzung zu gewährleisten. Generell erscheint zudem eine weitere Integration von zeitlichen Regularisierungsmechanismen in beide Ansätze sinnvoll, um nicht nur eine räumliche, sondern auch zeitliche Glattheit der Bewegungsschätzung zu ermöglichen. In diesem Kontext sollte zudem über die Integration bzw. Adaption von Deep Learning-basierten Techniken aus dem verwandten Bereich der Gesichtserkennung/-verfolgung in Videos nachgedacht werden, die die echtzeitfähige, modellbasierte räumlich und zeitlich glatte Verfolgung von Gesichtsländern ermöglichen (siehe z. B. [Hou u. a., 2018]).

Kapitel 7

Generierung flexibler statistischer Bewegungsmodelle mit wenigen Trainingsdaten

Kapitelinhalt

7.1	Hintergrund, Vorgehen und verwandte Arbeiten	142
7.2	Statistische Bewegungsmodellierung mit Lokalitätsannahme	145
7.2.1	Integration der Lokalitätsannahme in den Modellierungsprozess	145
7.2.2	Fusionierung verschiedener lokaler Modelle	148
7.2.3	Interpretation des Ansatzes mit Kernelfunktionen	150
7.2.4	Anmerkungen zur praktischen Umsetzung	152
7.3	Experimente	153
7.4	Ergebnisse	155
7.5	Zusammenfassung und Diskussion	158

Kapitel 6 hat gezeigt, dass PCA-basierte statistische Bewegungsmodelle gewinnbringend im Rahmen der intrainerventionellen Bildregistrierung als Vorwissen und Interpolator eingesetzt werden können. Hierbei stand im letzten Kapitel der Einsatz patientenspezifischer Bewegungsmodelle im Fokus, die anhand direkt vor der Intervention aufgenommener Bilddaten generiert werden. Durch dieses Vorgehen können zwar die in Kapitel 5 diskutierten Effekte interfraktioneller Bewegungsvariabilität vermieden werden, gleichwohl hängt auch mit diesen Bewegungsmodellen die erzielbare Genauigkeit der Bewegungsschätzung u. a. von der intrafraktionellen Bewegungsvariabilität ab und zudem bedeutet die Aufnahme neuer Trainingsdaten vor jeder Intervention einen zusätzlichen Aufwand. Deshalb kommt auch für die in Kapitel 6 diskutierte Anwendung der Einsatz populationsbasierter statistischer Bewegungsmodelle infrage.

Wie in Kapitel 5.1 ausführlicher diskutiert, sind populationsbasierte (und PCA-basierte) statistische Bewegungsmodelle schon vielfach in der Literatur vorgeschlagen und untersucht worden (u. a. [Ehrhardt u. a., 2010; He u. a., 2010; Klinder u. a., 2009, 2012; Preiswerk u. a., 2014]). Diese erreichen in der Praxis allerdings oftmals eine

weitaus geringere Genauigkeit als patientenspezifische Modelle, was für eine zu geringe Generalisierungsfähigkeit der Populationsmodelle spricht. In diesem Zusammenhang wurden in Kapitel 5.1 schon die unter Umständen große Interpatientenvariabilität der Atembewegung und die gleichzeitig mangelnde Repräsentativität der Trainingsstichprobe als Gründe genannt. Dieser Mangel an Repräsentativität ergibt sich zumeist schon aus der geringen Anzahl verwendeter Trainingsdaten (z. B. 12 Datensätze in [Ehrhardt u. a., 2010] oder 20 Datensätze in [Preiswerk u. a., 2014]) und der großen Anzahl zu modellierender Verschiebungsvektoren (z. B. mehrere Millionen), wodurch ein HDLSS-Problem (engl. High Dimension Low Sample Size) entsteht. Auch wenn durch den zunehmenden klinischen Einsatz von räumlich-zeitlichen Bildgebungstechniken die Beschaffung von Ausgangsdaten für das Modelltraining einfacher geworden ist, erscheint es dennoch u. a. hinsichtlich einer Aufwandsreduzierung und aufgrund der Interpatientenvariabilität der Atmung sinnvoll, sich mit der Generierung repräsentativer statistischer Bewegungsmodelle auf Basis weniger Trainingsdaten zu beschäftigen. In diesem Kapitel wird deshalb ein neues Modellierungsverfahren präsentiert und im Kontext der populationsbasierten Bewegungsmodellierung und -schätzung evaluiert, wobei die Methodenbeschreibung größtenteils auf [Wilms u. a., 2017a,b] basiert.

7.1 Hintergrund, Vorgehen und verwandte Arbeiten

Sei als Ausgangspunkt für das Training eines PCA-basierten statistischen Bewegungsmodells analog zu Kapitel 6.2 eine Menge von n_{ts} Verschiebungsfeldern ($u_j \in \Omega$) $_{j=1}^{n_{ts}}$ verfügbar. Unabhängig davon, ob diese Felder von einem Patienten (patientenspezifisches Modell) oder von mehreren Patienten (populationsbasiertes Modell mit $\Omega \equiv \Omega_A$ ²⁷) stammen, basiert die Modellgenerierung auf der Kovarianzmatrix

$$\mathbf{C} = \frac{1}{(n_{ts} - 1)} \sum_{j=1}^{n_{ts}} (\mathbf{u}_j - \bar{\mathbf{u}})(\mathbf{u}_j - \bar{\mathbf{u}})^T = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^T \in \mathbb{R}^{3n_{vx} \times 3n_{vx}} \text{ mit } \mathbf{u}_j = \text{vec}(u_j) \in \mathbb{R}^{3n_{vx}}. \quad (7.1)$$

Hierbei gibt $3n_{vx}$ die Größe der Verschiebungsfelder in Abhängigkeit von der Anzahl der Bildvoxel n_{vx} des 3D-Referenzbildes an, $\bar{\mathbf{u}} = \frac{1}{n_{ts}} \sum_{j=1}^{n_{ts}} \mathbf{u}_j$ bezeichnet die mittlere Bewegung in den Trainingsdaten, die Spaltenvektoren von $\mathbf{P} = [\mathbf{p}_1 \cdots \mathbf{p}_{3n_{vx}}] \in \mathbb{R}^{3n_{vx} \times 3n_{vx}}$ sind die orthonormalen Eigenvektoren von \mathbf{C} und die Diagonalelemente der Diagonalmatrix $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_{3n_{vx}}) \in \mathbb{R}_{\geq 0}^{3n_{vx} \times 3n_{vx}}$ entsprechen den korrespondierenden Eigenwerten (siehe auch Kapitel 6.2). Das Bewegungsmodell wird dann durch das Tupel $(\bar{\mathbf{u}}, \mathbf{P}_k)$ bestehend aus mittlerer Bewegung $\bar{\mathbf{u}}$ und den k orthonormalen Eigenvektoren $\mathbf{P}_k \in \mathbb{R}^{3n_{vx} \times k}$ mit den größten positiven Eigenwerten repräsentiert, welches einen k -dimensionalen affinen Unterraum des $\mathbb{R}^{3n_{vx}}$ parametrisiert. Da die Anzahl der

²⁷ Ω_A entspricht einem gemeinsamen Referenz-/Atlasraum, in den die Felder/Transformationen aller Populationspatienten transportiert werden. Siehe Kapitel 5.2.1 für eine detaillierte Beschreibung.

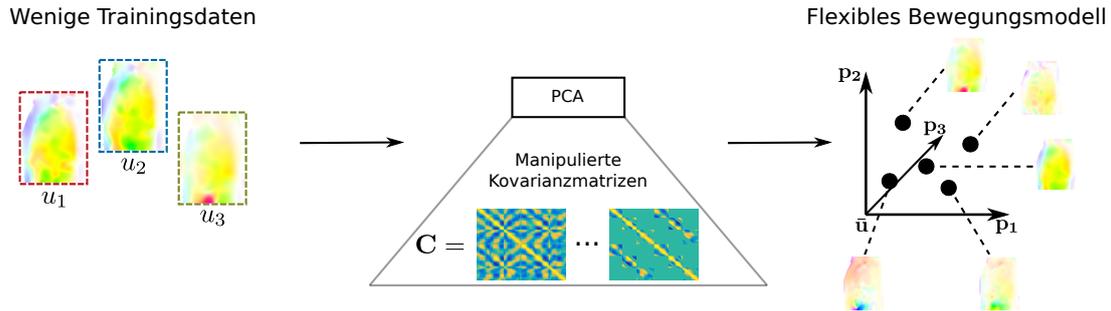


Abb. 7.1: Schematische Illustration des in diesem Kapitel vorgeschlagenen Verfahrens zur Erhöhung der Flexibilität PCA-basierter statistischer Bewegungsmodelle bei der Nutzung weniger Trainingsdaten durch Kovarianzmatrixmanipulationen.

positiven Eigenwerte einer Kovarianzmatrix deren Rang angibt, ist der maximal wählbare Wert für k abhängig vom Rang $\text{rang}(\mathbf{C})$ von \mathbf{C} . In der Regel gilt $n_{ts} \ll 3n_{vx}$ und somit $k \leq \text{rang}(\mathbf{C}) \leq (n_{ts} - 1)$, wodurch die Generalisierungsfähigkeit bzw. Flexibilität des resultierenden Bewegungsmodells insbesondere bei kleiner Trainingsstichprobe stark eingeschränkt ist (je kleiner k , desto weniger Flexibilität).

Soll die Flexibilität eines Bewegungsmodells in HDLSS-Szenarien erhöht werden, muss hierfür grundsätzlich die Dimension des affinen Unterraums erhöht werden. Entsprechende Verfahren in der Literatur beziehen sich mehrheitlich auf statistische Formmodelle und können grob in drei verschiedene Kategorien eingeteilt werden: (1) Künstliche Vergrößerung der Menge der Trainingsdaten durch u. a. zufällige Deformation der Ausgangsdaten, wodurch der Rang der Kovarianzmatrix erhöht wird [Koikkalainen u. a., 2008]. (2) Direkte Manipulation der Kovarianzmatrix durch u. a. Einbringung FEM-basierter Deformationsinformationen [Cootes u. a., 1995] oder Erhöhung/Reduktion der Kovarianzen aufgrund von Nachbarschaftsbeziehungen [Cootes u. a., 1996; Wang u. a., 2000]. Auch durch diese Verfahren wird der Rang der Kovarianzmatrix verändert. (3) Wavelet-basierte Verfahren, die eine Hierarchie von Modellen verschiedener Detaillevel erzeugen, die zusammen flexibler sind als ein normales Modell [Cerrolaza u. a., 2012; Davatzikos u. a., 2003].

Nachteil der Verfahren der ersten beiden Kategorien ist die oftmals mangelnde Spezifität der in den Modellierungsprozess eingebrachten Zusatzinformationen, die teilweise nur sehr geringen Bezug zum zu modellierenden Problem aufweisen (zufällige Deformationen). Bei den Wavelet-basierten Verfahren ist die aufwendige Wavelet-Transformation als negativ anzusehen, die bei der Modellgenerierung und auch bei der Erzeugung von Elementen durch das Modell genutzt wird. Zudem parametrisiert das Modell keinen einzelnen affinen Unterraum, wodurch die direkte Kompatibilität mit Standardmodelle nutzenden Anwendungen nicht gegeben ist.

Verfahren zur Erhöhung der Flexibilität in HDLSS-Szenarien, die sich explizit auf die Modellierung von dichten Verschiebungsfeldern beziehen, sind in der Literatur selten

zu finden. Eine Ausnahme ist u. a. [Xue u. a., 2006], wo aufbauend auf dem Wavelet-basierten Verfahren von [Davatzikos u. a., 2003] flexible Modelle für die modellbasierte Registrierung generiert werden. Weiterhin wird in [Ehrhardt u. a., 2016] unter Mitwirkung des Autors dieser Arbeit ein Verfahren präsentiert, welches auf einer Lokalisierungsannahme (vgl. Lokalisierungsprinzip in der Physik) basiert. Hierbei wird angenommen, dass generelle bzw. globale Bewegungsvariationen (z. B. die kraniokaudale Richtung der Atembewegung) schon bei Nutzung weniger Trainingsdaten relativ zuverlässig geschätzt werden können und es somit darum geht auch lokale Variationen, die nur einen lokal begrenzten Wirkungsbereich haben und durch die Trainingsdaten nicht abgebildet werden, in das Modell zu integrieren. Um dies zu erreichen, werden die realen Trainingsdaten zerteilt und durch Kombination räumlich weit voneinander entfernter Teile verschiedener realer Datensätze zu einer großen Menge neuer, virtueller Trainingsdaten zusammengesetzt. Um Inkonsistenzen zwischen den Teilen zu vermeiden, werden Lücken gelassen, wodurch ein Low-rank Matrix Completion Problem entsteht. Dieses kann durch entsprechende Completion-Verfahren gelöst werden, welche ein flexibleres Modell erzeugen. Das Verfahren lässt sich der ersten Kategorie der obigen Kategorisierung zuordnen, wobei sich die Erzeugung neuer Trainingsdaten durch die Kombination vorhandener Informationen deutlich spezifischer an den realen Daten orientiert als die anderen Verfahren. Es ist allerdings unklar, wie viele virtuelle Trainingsdaten erzeugt werden müssen und wie die Kombination der Teile konkret umgesetzt werden muss, um eine gewünschte Flexibilität zu erreichen.

Das in diesem Kapitel präsentierte Verfahren basiert ebenfalls auf der Lokalisierungsannahme, setzt diese aber durch Manipulationen der Kovarianzmatrix um. Die Grundidee ist, die Kovarianzen weit entfernter Elemente zu entfernen (Nullsetzen), um so den Rang der Kovarianzmatrix zu erhöhen und mehr Eigenvektoren generieren zu können. Da in der Praxis die konkrete Definition eines einzelnen Distanzschwellenwertes zur Definition von Lokalität schwierig ist, werden Modelle mit unterschiedlichen Distanzschwellenwerten/unterschiedlicher Lokalität generiert. Diese werden anschließend zu einem einzelnen flexiblen Modell fusioniert, welches direkt in Kapitel 6 eingesetzt werden kann (siehe Abb. 7.1). Dieses Verfahren wurde erstmals in [Wilms u. a., 2017a] für die Formmodellierung und in [Wilms u. a., 2017b] für die Erscheinungsmodellierung publiziert und wird in diesem Kapitel für die Bewegungsmodellierung mit Verschiebungsfeldern beschrieben (Kapitel 7.2.1–7.2.2) und evaluiert (Kapitel 7.3–7.4).

Grundsätzlich neu ist die Integration von Lokalisierungsannahmen in den Modellierungsprozess durch Kovarianzmatrixmanipulationen nicht. Diese Möglichkeit wurde schon in [Cootes u. a., 1996; Wang u. a., 2000] für die Formmodellierung beschrieben, wobei das hier präsentierte Verfahren diese Idee wesentlich konsequenter und umfangreicher (u. a. Kombination von Modellen verschiedener Lokalität) umsetzt. Die größte Ähnlichkeit besitzt die hier präsentierte Umsetzung zu dem zeitgleich entwickelten Verfahren in [Lüthi u. a., 2018], welches in [Jud u. a., 2017] zur modellbasierten Schätzung der Atembewegung genutzt wird. Lüthi u. a. [2018] erhöhen die Flexibilität von Gaussian

Process Morphable Models für die Form-/Deformationsmodellierung ebenfalls durch die distanzbasierte Manipulation von Kovarianzen und präsentieren auch einen Ansatz zur Kombination verschiedener Lokalitätslevel in einem Modell. Eine ausführliche Darstellung der Gemeinsamkeiten dieser Verfahren erfolgt in Kapitel 7.2.3.

7.2 Statistische Bewegungsmodellierung mit Lokalitätsannahme

Die nachfolgenden Teilkapitel beschreiben das neue Verfahren zur PCA-basierten Modellierung der Atembewegung mit Lokalitätsannahme bei Nutzung von Verschiebungsfeldern. Für die grafische Veranschaulichung der einzelnen Schritte des Verfahrens wird allerdings, wie in [Wilms u. a., 2017a], auf punktbasiert repräsentierte Handkonturen mit unterschiedlichen Fingerbewegungen/-stellungen zurückgegriffen. Auch wenn dies ein Beispiel aus dem Bereich der klassischen Formmodellierung ist, so eignet es sich u. a. aufgrund der Bewegungskomponente sehr gut zur Illustration der Effekte des Verfahrens im Kontext der Bewegungsmodellierung (siehe z. B. Abb. 7.2).

7.2.1 Integration der Lokalitätsannahme in den Modellierungsprozess

Die Kovarianzmatrix

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{11} & \cdots & \mathbf{C}_{1n_{vx}} \\ \vdots & \ddots & \vdots \\ \mathbf{C}_{1n_{vx}} & \cdots & \mathbf{C}_{n_{vx}n_{vx}} \end{pmatrix} \in \mathbb{R}^{3n_{vx} \times 3n_{vx}} \quad (7.2)$$

in Glg. 7.1 besteht aus n_{vx}^2 Untermatrizen $\mathbf{C}_{ij} \in \mathbb{R}^{3 \times 3}$, deren Elemente die Kovarianzen der drei Verschiebungsvektorkomponenten des jeweiligen Punkttupels ($\mathbf{x}_i \in \Omega$, $\mathbf{x}_j \in \Omega$) beschreiben und damit den linearen Zusammenhang der Bewegungen zwischen den Punkten \mathbf{x}_i und \mathbf{x}_j codieren. Zur Umsetzung der Lokalitätsannahme gilt es, diesen Zusammenhang zu entfernen. Hierfür wird zunächst eine symmetrische Distanzfunktion $d : \Omega \times \Omega \rightarrow \mathbb{R}_{\geq 0}$ mit $d(\mathbf{x}, \mathbf{x}) = 0$ zur Quantifizierung der Entfernung zweier Punkte im Referenzkoordinatensystem definiert (z. B. euklidische Distanz). Zudem wird ein Distanzschwellenwert $\tau \in \mathbb{R}_{\geq 0}$ festgelegt, sodass $d(\mathbf{x}_i, \mathbf{x}_j) > \tau$ anzeigt, dass die Bewegung der Punkte \mathbf{x}_i und \mathbf{x}_j unabhängig sein soll.

Auf Basis von \mathbf{C} , $d(\cdot, \cdot)$ und τ wird nun eine symmetrische Matrix $\mathbf{C}_\tau \in \mathbb{R}^{3n_{vx} \times 3n_{vx}}$ erzeugt, deren Elemente für nahe Punkte denen von \mathbf{C} entsprechen, wohingegen die Untermatrizen $(\mathbf{C}_\tau)_{ij} \in \mathbb{R}^{3 \times 3}$ entfernter Punkte durch Nullmatrizen ersetzt werden:

$$(\mathbf{C}_\tau)_{ij} = \begin{cases} \mathbf{C}_{ij} & , \text{ wenn } d(\mathbf{x}_i, \mathbf{x}_j) \leq \tau \\ \mathbf{0}_{3 \times 3} & , \text{ wenn } d(\mathbf{x}_i, \mathbf{x}_j) > \tau \end{cases} \quad (7.3)$$

Die Matrix \mathbf{C}_τ entspricht letztlich einer distanzgefilterten Version von \mathbf{C} . Im Vergleich zu \mathbf{C} ist \mathbf{C}_τ durch die Manipulationen aber nicht zwingend positiv semidefinit und damit keine valide Kovarianzmatrix mehr, was entsprechende Auswirkungen auf die Eigenwertzerlegung hat (negative Eigenwerte). Um diese Problematik zu umgehen, wird zu \mathbf{C}_τ zunächst eine möglichst ähnliche valide Kovarianzmatrix $\hat{\mathbf{C}}_\tau$ gesucht.

Die nach Frobeniusnorm nächste positiv semidefinite Matrix zu \mathbf{C}_τ lässt sich durch Entfernung der Eigenvektoren mit negativen Eigenwerten bei einer Eigenwertzerlegung von \mathbf{C}_τ erzeugen [Boyd u. a., 2004]. Durch dieses einfache Vorgehen werden aber u. a. potenziell auch die Hauptdiagonalelemente der Matrix verändert, wodurch sich die Varianzen der einzelnen Verschiebungsfeldkomponenten ändern und auch die Gesamtvarianz ändert ($\text{tr}(\mathbf{C}_\tau) \neq \text{tr}(\hat{\mathbf{C}}_\tau)$). Um dies zu verhindern, wird \mathbf{C}_τ zunächst in eine Korrelationsmatrix $\mathbf{R}_\tau = \mathbf{S}^{-1}\mathbf{C}_\tau\mathbf{S}^{-1}$ umgewandelt, wobei $\mathbf{S} = \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_{3n_{vx}}}) \in \mathbb{R}^{3n_{vx} \times 3n_{vx}}$ eine Diagonalmatrix ist, deren Hauptdiagonalelemente den Standardabweichungen der Verschiebungsfeldkomponenten in den Trainingsdaten entsprechen. Anschließend wird anstatt der nächsten Kovarianzmatrix die nächste (postiv semidefinite) Korrelationsmatrix $\hat{\mathbf{R}}_\tau$ zu \mathbf{R}_τ gesucht:

$$\min_{\hat{\mathbf{R}}_\tau} \|\mathbf{R}_\tau - \hat{\mathbf{R}}_\tau\|_F \quad \text{s.t.} \quad \hat{\mathbf{R}}_\tau = \hat{\mathbf{R}}_\tau^T, \det(\hat{\mathbf{R}}_\tau) \geq 0 \text{ und } \text{diag}(\hat{\mathbf{R}}_\tau) = (1, \dots, 1)^T. \quad (7.4)$$

Die dritte Nebenbedingung sorgt explizit dafür, dass die Hauptdiagonalelemente nicht verändert werden (alle Elemente sind 1). Dieses Optimierungsproblem wird hier iterativ durch das in [Higham, 2002] beschriebene Verfahren gelöst, welches $\hat{\mathbf{R}}_\tau$ ausgehend von $\hat{\mathbf{R}}_\tau = \mathbf{R}_\tau$ alternierend auf die konvexe Menge symmetrischer positiv semidefiniter Matrizen und die konvexe Menge von Matrizen mit ausschließlich Einsen auf der Hauptdiagonale abbildet. Mithilfe von \mathbf{S} ergibt sich schließlich eine positiv semidefinite Kovarianzmatrix $\hat{\mathbf{C}}_\tau = \mathbf{S}\hat{\mathbf{R}}_\tau\mathbf{S}^T$, deren Hauptdiagonalelemente denen von \mathbf{C}_τ entsprechen. Somit bleibt durch die Manipulation die Gesamtvarianz erhalten.

Analog zu Kapitel 7.1 parametrisiert die Matrix $\mathbf{P}_\tau \in \mathbb{R}^{3n_{vx} \times k_\tau}$ der k_τ orthonormalen Eigenvektoren von $\hat{\mathbf{C}}_\tau$ mit den k_τ größten positiven Eigenwerten (Diagonalelemente der Diagonalmatrix $\mathbf{\Lambda}_\tau \in \mathbb{R}^{k_\tau \times k_\tau}$) zusammen mit $\bar{\mathbf{u}}$ wieder einen k_τ -dimensionalen affinen Unterraum des $\mathbb{R}^{3n_{vx}}$. Abhängig von τ gilt hier durch die Manipulationen $\text{rang}(\hat{\mathbf{C}}_\tau) \geq \text{rang}(\mathbf{C})$, wodurch im Vergleich zu Kapitel 7.1 entsprechend größere Werte für k_τ gewählt werden können und das Modell entsprechend mehr Flexibilität aufweist (siehe Abb. 7.2). Das durch das Tupel $(\bar{\mathbf{u}}, \mathbf{P}_\tau)$ repräsentierte Bewegungsmodell ist ein lokales Bewegungsmodell für den Schwellenwert τ . Wird τ so groß gewählt, dass es zu keiner Manipulationen der Kovarianzmatrix kommt, entspricht das lokale Modell dem globalen, unflexiblen Modell in Kapitel 7.1, bei dem die Bewegungsvariationen entlang der einzelnen Basisvektoren/Eigenvektoren hauptsächlich globaler Natur sind. Durch Reduktion von τ steigt der Rang der Kovarianzmatrix und es können mehr Basisvektoren/Eigenvektoren ermittelt werden, die zudem lokale Bewegungsvariationen

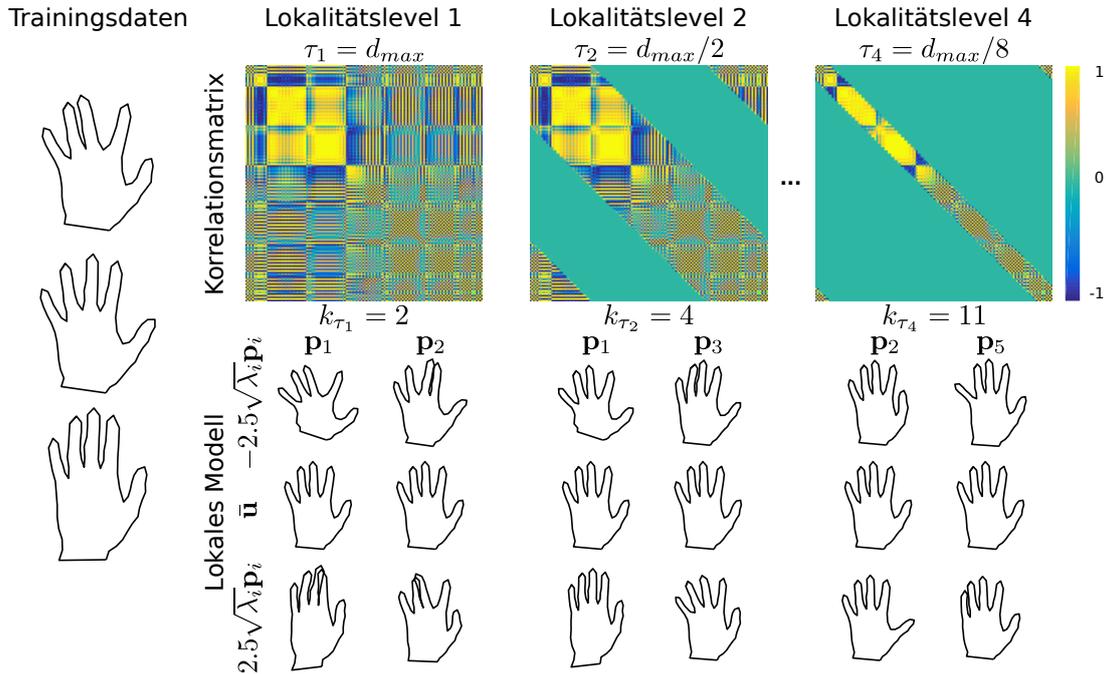


Abb. 7.2: Illustration der Auswirkung verschiedener Distanzschwellenwerte (τ_1 , τ_2 und τ_4) auf die resultierende Korrelationsmatrix und die hieraus entstehenden lokalen Modelle. Alle Modelle wurden basierend auf den drei am linken Rand abgebildeten Handformen erzeugt. d_{max} bezeichnet die maximale Distanz zwischen zwei Punkten entlang der Kontur, weshalb das lokale Modell für $\tau_1 = d_{max}$ dem Standardmodell entspricht (keine Manipulation). Für jedes lokale Modell sind die Formdeformationen/Fingerbewegungen entlang zweier ausgewählter Basisvektoren/Eigenvektoren abgebildet. Es ist erkennbar, dass die lokalen Modelle teilweise redundante Informationen abbilden (z. B. \mathbf{p}_1 für τ_1 und \mathbf{p}_1 für τ_2) und die modellierten Effekte mit abnehmendem Schwellenwert lokaler werden (für τ_4 wird die Bewegung einzelner Finger modelliert). Abbildung modifiziert und ergänzt aus [Wilms u. a., 2017a].

beschreiben. Das entsprechende Extremum der Lokalität ist erreicht, wenn alle Punkte unabhängig sind, was in einem extrem flexiblen, aber nutzlosen Modell resultiert.

Die Wahl des richtigen Schwellenwertes τ , um die gewünschte Flexibilität/Lokalität zu erreichen, stellt nun die zentrale Herausforderung bei Einsatz des Verfahrens dar. Da ein einzelner korrekter Schwellenwert in der Praxis zumeist unbekannt bzw. schwierig zu bestimmen ist, wird hier vorgeschlagen, n_l lokale Modelle mit abnehmendem Schwellenwert $\tau_1 > \tau_2 > \dots > \tau_l > \dots > \tau_{n_l}$ zu erzeugen, welche verschiedene Lokalitätslevel (von global nach lokal) beschreiben, aber ebenso auch redundante Informationen enthalten (siehe Abb. 7.2). Um diese Redundanz zu entfernen und die n_l Modelle beispielsweise in Kapitel 6 einsetzen zu können, werden sie im nächsten Teilkapitel zu einem einzelnen Modell fusioniert, welches die verschiedenen Lokalitätslevel integriert.

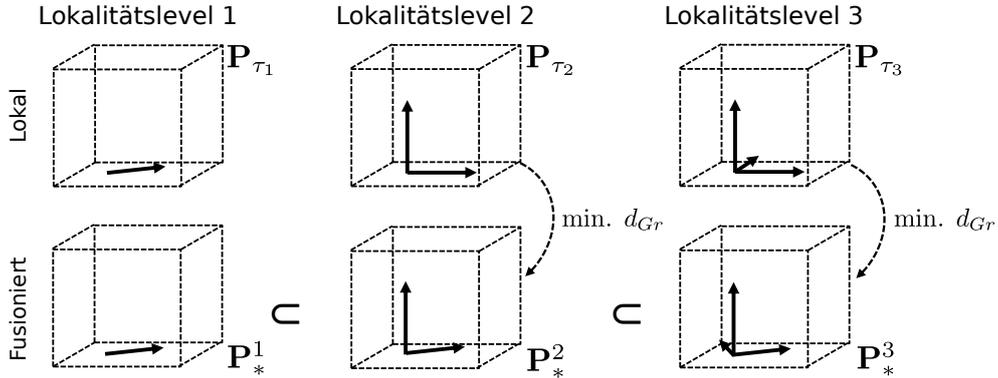


Abb. 7.3: Vereinfachte und schematische Illustration des Fusionierungsprozesses für drei lokale Modelle repräsentiert durch die Basen der jeweiligen Unterräume. Die erste Zeile zeigt die Basen der lokalen Unterräume verschiedener Dimension, wohingegen die zweite Zeile die Basen der optimalen Unterräume nach der Kombination darstellt.

7.2.2 Fusionierung verschiedener lokaler Modelle

Seien $(\mathbf{P}_{\tau_1}, \dots, \mathbf{P}_{\tau_l}, \dots, \mathbf{P}_{\tau_{n_l}})$ die Matrizen orthonormaler Eigenvektoren zur Schwellenwertsequenz $\tau_1 > \tau_2 > \dots > \tau_l > \dots > \tau_{n_l}$ (siehe Kapitel 7.2.1), welche zusammen mit $\bar{\mathbf{u}}$ insgesamt n_l lokale Modelle definieren. Formal parametrisiert jedes lokale Modell einen k_{τ_l} -dimensionalen affinen Unterraum $\{\bar{\mathbf{u}} + \mathbf{w} \mid \mathbf{w} \in \text{span}(\mathbf{P}_{\tau_l})\} \subseteq \mathbb{R}^{3n_{vx}}$ bestehend aus Verschiebungsvektor $\bar{\mathbf{u}}$ und dem linearen Unterraum $\mathcal{L}_{\tau_l} = \text{span}(\mathbf{P}_{\tau_l}) \subseteq \mathbb{R}^{3n_{vx}}$ (siehe Kapitel 5.6 für die Notation). Da der Verschiebungsvektor $\bar{\mathbf{u}}$ in allen Modellen enthalten ist, unterscheiden sich diese nur durch die linearen Unterräume. Durch die abnehmende Definition der Schwellenwerte mit ansteigendem Lokalisierungslevel und unter Nutzung der gleichen Kriterien zur Wahl von k_{τ_l} (z. B. 98% der Gesamtvarianz) steigt die Dimension der Unterräume mit zunehmender Lokalität (zumeist) an ($k_{\tau_1} \leq \dots \leq k_{\tau_l} \leq \dots \leq k_{\tau_{n_l}}$).

Ziel ist es nun, die n_l lokalen Modelle zu einem einzelnen Modell, repräsentiert durch das Tupel $(\bar{\mathbf{u}}, \mathbf{P}_*)$, zusammenzufassen. Hierbei sind die orthonormalen Spaltenvektoren von $\mathbf{P}_* \in \mathbb{R}^{3n_{vx} \times k_{\tau_{n_l}}}$ die Basis des $k_{\tau_{n_l}}$ -dimensionalen linearen Unterraums $\mathcal{L}_* = \text{span}(\mathbf{P}_*)$. Der gesuchte Unterraum \mathcal{L}_* soll so konstruiert sein, dass dabei die globalen Eigenschaften des Modells mit der geringsten Lokalität als charakteristische Eigenschaften erhalten bleiben und die Variationen der anderen Modelle jeweils bestmöglich approximiert werden können. Da eine direkte Bestimmung von \mathcal{L}_* bzw. $\text{span}(\mathbf{P}_*)$ schwierig ist, wird das Problem für die n_l Lokalisierungslevel nacheinander optimal gelöst. Ausgehend von $\mathbf{P}_*^1 = \mathbf{P}_{\tau_1}$ werden für das jeweilige Level l optimale Basisvektoren $\mathbf{P}_*^l \in \mathbb{R}^{3n_{vx} \times k_{\tau_l}}$ bestimmt, sodass der k_{τ_l} -dimensionale Unterraum $\mathcal{L}_*^l = \text{span}(\mathbf{P}_*^l)$ den optimalen $k_{\tau_{l-1}}$ -dimensionalen Unterraum $\mathcal{L}_*^{l-1} = \text{span}(\mathbf{P}_*^{l-1})$ für das vorherige Level enthält (Erhalt globaler Informationen) und möglichst geringen Abstand zum Unterraum $\mathcal{L}_{\tau_l} = \text{span}(\mathbf{P}_{\tau_l})$ des lokalen Modells dieses Levels hat (Ap-

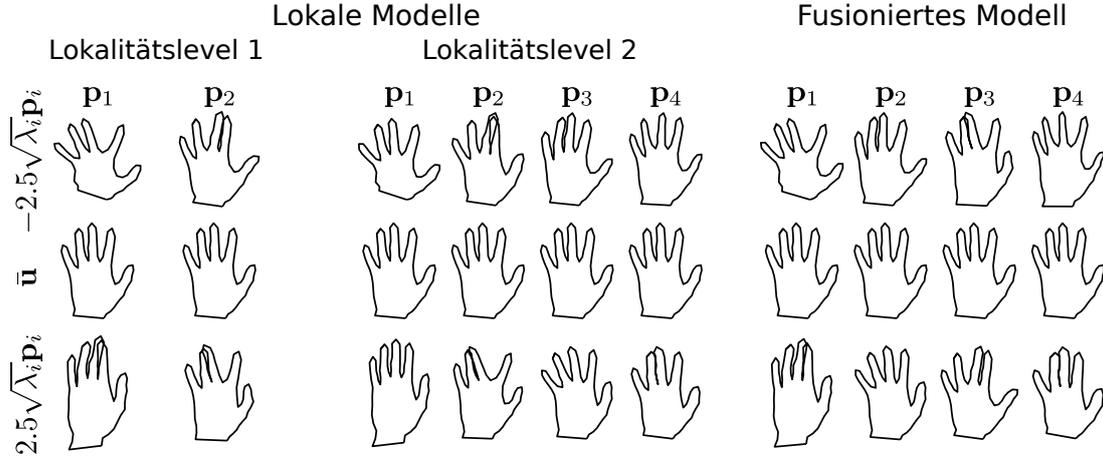


Abb. 7.4: Beispielhafte Illustration des Ergebnisses des Fusionierungsprozesses für die ersten beiden Lokalitätslevel aus Abb. 7.2. Das fusionierte Modell enthält die Eigenschaften beider Teilmodelle, entfernt aber Redundanzen. Abbildung modifiziert aus [Wilms u. a., 2017a].

proximation lokaler Variationen). Hierdurch entsteht eine Folge von n_l Unterräumen, sodass $\mathcal{L}_*^1 \subseteq \dots \subseteq \mathcal{L}_*^l \subseteq \dots \subseteq \mathcal{L}_*^{n_l}$ gilt (siehe Abb. 7.3 und Abb. 7.4). Ist das letzte Level erreicht, gilt $\mathcal{L}_* = \mathcal{L}_*^{n_l}$ bzw. $\mathbf{P}_* = \mathbf{P}_*^{n_l}$. Dieser Fusionierungsprozess lässt sich für jedes Level $l > 1$ durch das Optimierungsproblem

$$\mathcal{L}_*^l = \arg \min_{\mathcal{L} \in Gr(k_{\tau_l}, 3n_{vx})} d_{Gr}(\mathcal{L}, \mathcal{L}_{\tau_l}) \quad \text{s.t.} \quad \mathcal{L}_*^{l-1} \subseteq \mathcal{L} \quad (7.5)$$

beschreiben. Hierbei quantifiziert $d_{Gr}(\mathcal{L}, \mathcal{L}_{\tau_l})$ die geodätische Distanz zwischen \mathcal{L} und \mathcal{L}_{τ_l} entlang der Graßmann-Mannigfaltigkeit $Gr(k_{\tau_l}, 3n_{vx})$ der k_{τ_l} -dimensionalen Unterräume des $\mathbb{R}^{3n_{vx}}$ (siehe auch Kapitel 5.6).

Gegeben $\mathcal{L}_*^{l-1} \in Gr(k_{\tau_{l-1}}, 3n_{vx})$ und $\mathcal{L}_{\tau_l} \in Gr(k_{\tau_l}, 3n_{vx})$, lässt sich das Optimierungsproblem in Glg. 7.5 basierend auf dem Verfahren zur Berechnung von Distanzen zwischen Unterräumen unterschiedlicher Dimension (hier angenommen: $m = k_{\tau_{l-1}} < n = k_{\tau_l}$) aus [Ye u. a., 2016] lösen. Sei $(\mathbf{P}_*^{l-1})^T \mathbf{P}_{\tau_l} = \mathbf{U} \mathbf{D} \mathbf{V}^T$ die Singulärwertzerlegung des Produkts der beiden Matrizen von Basisvektoren der beteiligten Unterräume. Dann definieren $\mathbf{U} \in \mathbb{R}^{m \times m}$ und $\mathbf{V} \in \mathbb{R}^{n \times n}$ jeweils Rotationsmatrizen, sodass die rotierten Basen $\mathbf{P}_*^{l-1} \mathbf{U} = [\mathbf{a}_1 \dots \mathbf{a}_m]$ und $\mathbf{P}_{\tau_l} \mathbf{V} = [\mathbf{b}_1 \dots \mathbf{b}_n]$ genutzt werden können, um $\mathbf{P}_*^l = [\mathbf{a}_1 \dots \mathbf{a}_m, \mathbf{b}_{m+1} \dots \mathbf{b}_n] \in \mathbb{R}^{3n_{vx} \times n}$ zu bilden. Die Spaltenvektoren von \mathbf{P}_*^l sind orthonormal und $\mathcal{L}_*^l = \text{span}(\mathbf{P}_*^l)$ ist der gesuchte Unterraum aus der Menge der m -dimensionalen Unterräume, die \mathcal{L}_*^{l-1} enthalten, der die geringste geodätische Distanz $d_{Gr}(\mathcal{L}_*^l, \mathcal{L}_{\tau_l})$ aufweist (siehe [Ye u. a., 2016] für detaillierte Herleitungen und Beweise).

Werden bei der Anwendung des Bewegungsmodells auch die Eigenwerte der Kovarianzmatrix (= Varianzen) für die Interpretation der Wichtigkeit der einzelnen Basisvektoren benötigt (siehe z. B. Glg. 6.5), so müssen auch diese bei der Fusion der lokalen

Modelle mitberücksichtigt werden. Sei hierfür $\mathbf{\Lambda}_*^{l-1} \in \mathbb{R}^{m \times m}$ die Diagonalmatrix von Varianzen des fusionierten Modells des vorherigen Levels und $\mathbf{\Lambda}_{\tau_l} \in \mathbb{R}^{n \times n}$ die Diagonalmatrix von Varianzen/Eigenwerten zu den Eigenvektoren \mathbf{P}_{τ_l} der Kovarianzmatrix des aktuellen Levels. Gesucht ist dann eine entsprechende Diagonalmatrix von Varianzen $\mathbf{\Lambda}_*^l \in \mathbb{R}^{n \times n}$ für die optimale Basis \mathbf{P}_*^l . Hierfür werden die Diagonalmatrizen analog zu den Basen (siehe oben) zunächst mit \mathbf{U} bzw. \mathbf{V} transformiert: $\hat{\mathbf{\Lambda}}_*^{l-1} = \mathbf{U}^T \mathbf{\Lambda}_*^{l-1} \mathbf{U}$ bzw. $\hat{\mathbf{\Lambda}}_{\tau_l} = \mathbf{V}^T \mathbf{\Lambda}_{\tau_l} \mathbf{V}$. Die Matrix

$$\mathbf{\Lambda}_*^l = \begin{pmatrix} \hat{\mathbf{\Lambda}}_*^{l-1} & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{\Lambda}}_{\tau_l}[m+1:n, m+1:n] \end{pmatrix} \quad (7.6)$$

wird dann aus $\hat{\mathbf{\Lambda}}_*^{l-1}$ und $\hat{\mathbf{\Lambda}}_{\tau_l}$ zusammengesetzt, wobei $\hat{\mathbf{\Lambda}}_{\tau_l}[m+1:n, m+1:n]$ die Untermatrix von $\hat{\mathbf{\Lambda}}_{\tau_l}$ bei Entfernung der ersten m Spalten und Zeilen ist. Ist $\mathbf{\Lambda}_*^l$ keine Diagonalmatrix, liegen Korrelationen zwischen den Basisvektoren bezüglich der (manipulierten) Daten vor. Unkorrelierte Basisvektoren und eine Diagonalmatrix von Varianzen können bei Bedarf durch Eigenwertzerlegung von $\mathbf{P}_*^l \mathbf{\Lambda}_*^l (\mathbf{P}_*^l)^T$ erzeugt werden.

7.2.3 Interpretation des Ansatzes mit Kernelfunktionen

In diesem Teilkapitel wird das in den beiden vorangegangenen Teilkapiteln vorgestellte Verfahren zur Erzeugung flexibler Bewegungsmodelle mit Kernelfunktionen interpretiert. Dies ermöglicht einerseits ein besseres Verständnis für die Gründe der möglichen Indefinitheit der in Kapitel 7.2.1 manipulierten Kovarianzmatrizen, andererseits wird so auch ein direkter Vergleich zur Umsetzung der Lokalitätsannahme bei den Gaussian Process Morphable Models in [Jud u. a., 2017; Lüthi u. a., 2018] möglich, welche mit Kernelfunktionen definiert werden. Die folgende Darstellung lässt einige mathematische Details aus, weshalb interessierte Leser für eine ausführliche Betrachtung von Kernelfunktionen und -methoden auf [Shawe-Taylor u. a., 2004] verwiesen seien.

Eine symmetrische Funktion $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ wird in diesem Kapitel valide Kernelfunktion bzw. valider Kernel genannt, wenn sie die Bedingung

$$\sum_{i=1}^n \sum_{j=1}^n c_i c_j k(\mathbf{x}_i, \mathbf{x}_j) \geq 0 \quad (7.7)$$

für alle Sequenzen $(\mathbf{x}_i \in \mathcal{X})_{i=1}^n$, alle Sequenzen $(c_i \in \mathbb{R})_{i=1}^n$ und alle $n \in \mathbb{N}_{>0}$ erfüllt (positive Semidefinitheit; siehe [Hofmann u. a., 2008]). In diesem Fall ist die zugehörige symmetrische Kernelmatrix

$$\mathbf{K} = \begin{pmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \cdots & k(\mathbf{x}_1, \mathbf{x}_n) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_n, \mathbf{x}_1) & \cdots & k(\mathbf{x}_n, \mathbf{x}_n) \end{pmatrix} \in \mathbb{R}^{n \times n} \quad (7.8)$$

positiv semidefinit [Hofmann u. a., 2008].

Die distanzbasierte Kovarianzmatrixmanipulation in Kapitel 7.2.1 lässt sich nun in diesem Kernelkontext interpretieren, um u. a. aufzuzeigen, warum die manipulierten Kovarianzmatrizen nicht zwingend positiv semidefinit sind. Sei hierfür

$$k_S(\mathbf{x}_i, \mathbf{x}_j) = \begin{cases} 1 & , \text{ wenn } d(\mathbf{x}_i, \mathbf{x}_j) \leq \tau \\ 0 & , \text{ wenn } d(\mathbf{x}_i, \mathbf{x}_j) > \tau \end{cases} \quad \text{mit } \mathbf{x}_i, \mathbf{x}_j \in \Omega \quad (7.9)$$

eine Kernelfunktion mit zugehöriger Kernelmatrix \mathbf{K}_S , die aus Glg. 7.3 zur Distanzfilterung abgeleitet wird. Zusammen mit einer Einsmatrix $\mathbf{1} \in \mathbb{R}^{3 \times 3}$ lässt sich die distanzgefilterte Kovarianzmatrix \mathbf{C}_τ dann ausdrücken durch

$$\mathbf{C}_\tau = \mathbf{C} \odot (\mathbf{K}_S \otimes \mathbf{1}) . \quad (7.10)$$

Hierbei bezeichnet \otimes das Kronecker- und \odot das Hadamard-Produkt zweier Matrizen. Da \mathbf{C} und $\mathbf{1}$ aufgrund ihrer Konstruktion positiv semidefinit sind und das Kronecker- und das Hadamard-Produkt zweier positiv semidefiniter Matrizen wiederum positiv semidefinit ist [Hom u. a., 1991], kann \mathbf{C}_τ nur dann negative Eigenwerte haben, wenn \mathbf{K}_S nicht positiv semidefinit ist bzw. k_S kein valider Kernel ist. Die Funktion k_S ist für praxisrelevante Werte von τ ($0 < \tau < d_{max}$; d_{max} : maximale Distanz in den Daten) und unabhängig von der verwendeten Distanzfunktion im Allgemeinen keine valide Kernelfunktion, da die Definitheitsbedingung in Glg. 7.7 nicht zwingend erfüllt wird²⁸.

Dieses Resultat bedeutet, dass bei Ersetzung von k_S durch einen validen positiv semidefiniten Kernel die manipulierte Kovarianzmatrix \mathbf{C}_τ in Glg. 7.10 positiv semidefinit ist und auf die aufwendige Bestimmung der nächsten Kovarianz- bzw. Korrelationsmatrix in Kapitel 7.2.1 verzichtet werden kann. Ein in kernelbasierten Verfahren vielfach genutzter Kernel ist der auch in Lüthi u. a. [2018] eingesetzte Gaußkernel

$$k_G(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma d(\mathbf{x}_i, \mathbf{x}_j)^2) \quad \text{mit } \mathbf{x}_i, \mathbf{x}_j \in \Omega \text{ und } \gamma \in \mathbb{R}_{>0} , \quad (7.11)$$

bei dem z. B. $\gamma = 1/(2\tau^2)$ gewählt werden kann, um eine Lokalitätsannahme in Glg. 7.10 umzusetzen. Der Gaußkernel hat hierbei den Vorteil, dass er zumindest bei der Nutzung von Distanzmaßen $d(\cdot, \cdot)$, die isometrisch zu einer L_2 -Norm sind, positiv semidefinit ist [Haasdonk u. a., 2004] und somit auch die manipulierte Kovarianzmatrix diese Eigenschaft aufweist. Es ist zu beachten, dass k_G in Abhängigkeit von der Punktdistanz eine andere Modifikation der Kovarianzmatrix vornimmt als k_S (siehe auch Abb. 7.5). Das hier vorgestellte Verfahren als Kombination aus Kovarianzmatrixmanipulation und Modellfusionierung lässt sich problemlos auch mit k_G nutzen.

Bei direktem Vergleich des Vorgehens zur Integration einer Lokalitätsannahme in den Modellierungsprozess in [Jud u. a., 2017; Lüthi u. a., 2018] und diesem Kapitel lässt

²⁸Gegenbeispiel für Glg. 7.7 mit k_S bei $n = 3$: Wenn $d(\mathbf{x}_1, \mathbf{x}_2) > \tau \rightarrow k_S(\mathbf{x}_1, \mathbf{x}_2) = 0$, $d(\mathbf{x}_1, \mathbf{x}_3) \leq \tau \rightarrow k_S(\mathbf{x}_1, \mathbf{x}_3) = 1$ und $d(\mathbf{x}_2, \mathbf{x}_3) \leq \tau \rightarrow k_S(\mathbf{x}_2, \mathbf{x}_3) = 1$ mit $c_1 = 1$, $c_2 = 1$ und $c_3 = -1$, dann evaluiert Glg. 7.7 zu $2c_1c_3 + 2c_2c_3 + c_1^2 + c_2^2 + c_3^2 = -1$.

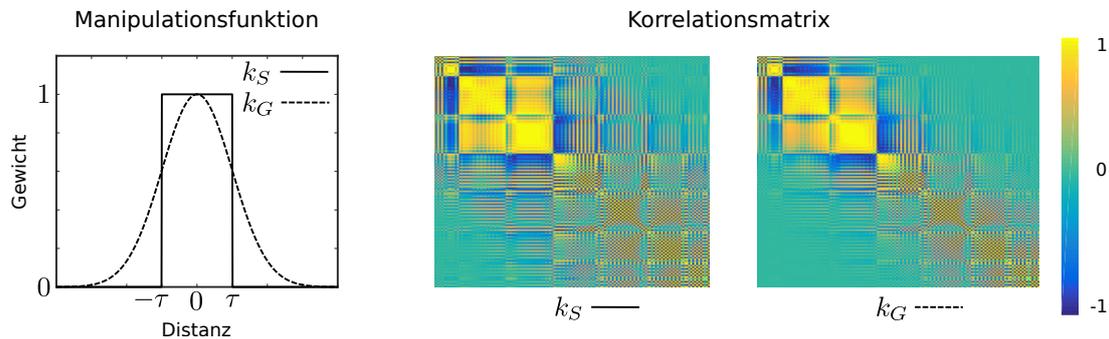


Abb. 7.5: Visualisierung der Unterschiede des Effekts, den die Nutzung der nicht-validen Kernelfunktion k_S und der validen Kernelfunktion k_G mit $\gamma = 1/(2\tau^2)$ auf die resultierenden Korrelationsmatrizen haben. Für k_S ist die nächste valide Korrelationsmatrix abgebildet. Die Visualisierung der Korrelationsmatrizen erfolgt aufgrund der im Vergleich zu den Kovarianzmatrizen besseren Sichtbarkeit der Unterschiede.

sich festhalten, dass die Umsetzung in [Lüthi u. a., 2018] durch einen Gaußkernel und in [Jud u. a., 2017] durch einen ähnlichen Wendland-Kernel den Vorteil hat, dass hierdurch die positive Semidefinitheit der Kovarianzmatrix erhalten bleibt, ohne dass ein Nachverarbeitungsschritt notwendig ist. Der in diesem Kapitel vorgeschlagene Nachverarbeitungsschritt (Bestimmung der nächsten Kovarianz- bzw. Korrelationsmatrix) ermöglicht aber beispielsweise die Nutzung beliebiger Funktionen als Distanzfunktion d , unabhängig von deren mathematischen Eigenschaften. Hierdurch lässt sich einfach beliebiges anatomisches oder physiologisches Vorwissen punktweise in die Distanzfunktion integrieren (z. B. komplexe Zusammenhänge zwischen Strukturen). Inwieweit das Manipulationsergebnis von k_G mit der nächsten Kovarianzmatrix nach Einsatz von k_S zusammenhängt, ist unklar. In [Jud u. a., 2017; Lüthi u. a., 2018] werden zudem Gaußkernel mit verschiedenen Werten für γ gewichtet aufsummiert, um unterschiedliche Lokalisierungslevel in einem Modell abbilden zu können. Es ist unklar, inwieweit dieses Vorgehen mit der Fusionierung verschiedener lokaler Modelle in diesem Kapitel zusammenhängt.

7.2.4 Anmerkungen zur praktischen Umsetzung

Hinsichtlich der praktischen Umsetzung des in den Kapiteln 7.2.1 und 7.2.2 skizzierten Ansatzes für die Bewegungsmodellierung stellt die Notwendigkeit der Eigenwertzerlegung von Matrizen mit $(3n_{vx})^2$ Elementen eine besondere Herausforderung dar. Diese wird sowohl bei der Berechnung der Eigenvektoren von $\hat{\mathbf{C}}_\tau$ als auch (mehrfach) bei der Bestimmung der nächsten Korrelationsmatrix benötigt. Bei der Bewegungsmodellierung müssen regelmäßig mehrere Millionen Bildpunkte/Bewegungsvektoren (n_{vx}) berücksichtigt werden, was eine explizite Repräsentation der entsprechenden Kovarianz- bzw. Korrelationsmatrix im Arbeitsspeicher unmöglich macht und auch die Bere-

chenbarkeit verhindert (Zeitkomplexität einer naiven Eigenwertzerlegung bei Annahme eines vollen Rangs: $\mathcal{O}(n_{vx}^3)$). Ist keine Bestimmung der nächsten Korrelationsmatrix nötig, weil $\hat{\mathbf{C}}_\tau = \mathbf{C}_\tau$ positiv semidefinit ist (siehe auch Kapitel 7.2.3 für entsprechende Voraussetzungen), wird hier die Nutzung des Verfahrens zur approximativen Eigenwertzerlegung einer positiv semidefiniten Matrix mit niedrigem Rang aus [Wang u. a., 2016] vorgeschlagen. Dieses Verfahren kombiniert die Vorteile bisher im Bereich des maschinellen Lernens weitverbreiteter Approximationsverfahren²⁹ und ermöglicht die Eigenwertzerlegung (theoretisch) in linearer Zeit bezogen auf n_{vx} und benötigt hierbei nur einen Bruchteil der $(3n_{vx})^2$ Matrixelemente ($< 0.01\%$ in den Experimenten in Kapitel 7.3).

7.3 Experimente

Die Evaluation des vorgeschlagenen Verfahrens zur Erhöhung der Flexibilität von statistischen Bewegungsmodellen bei Nutzung weniger Trainingsdaten verfolgt als Hauptziel die Untersuchung der Modelleigenschaften (u. a. Generalisierungsfähigkeit) und den Vergleich dieser mit denen eines Standardmodells für unterschiedlich große Trainingspopulationen. Auf einen Vergleich mit anderen Ansätzen zur Flexibilisierung (siehe Kapitel 7.1) wird hier verzichtet und auf die ausführliche Evaluation im Kontext der Formmodellierung in [Wilms u. a., 2017a] verwiesen, in der das vorgeschlagene Verfahren in fast allen Situationen bessere Ergebnisse erzielt als die anderen Verfahren.

Verwendete Daten Für die Evaluation werden Populationsmodelle mit verschiedenen großen Trainingsmengen erzeugt. Als Datengrundlage dienen hierfür die in Kapitel 5.3.1 beschriebenen 73 4D-CT-Bilddatensätze (Gruppe 1+2 und weitere Populationspatienten) mit den zugehörigen durch Bildregistrierung geschätzten Transformationen/Verschiebungsfeldern, welche alle in einem Atlasraum ausgerichtet sind. Von diesen Daten wird nur die Bewegung zwischen den Phasen maximaler Ein- und Ausatmung modelliert (EI→EE). Zudem wird die Bewegungsmodellierung analog zu Kapitel 5.3 auf die Lungenregion beschränkt. Die Verschiebungsfelder von zehn zufällig ausgewählten Datensätzen dienen für die Evaluation als Testdaten (Pseudo-Grundwahrheit). Alle restlichen 63 Datensätze bilden die Gesamtpopulation für das Modelltraining.

Durchgeführte Experimente Im Hauptteil der Evaluation werden Standardmodelle (siehe Kapitel 7.1 und Kapitel 6.2) und Modelle mit umgesetzter Lokaltätsannahme (Lokaltätsmodelle; Kapitel 7.2) für verschieden große Mengen von Trainingsdaten mit $n_{ts} \in \{5, 10, 20, 30, 40, 60\}$ erzeugt. Hierfür werden die Trainingsmengen

²⁹Nyström-Approximation: Hohe Effizienz möglich, nur Teile der Matrix notwendig, relativ geringe Genauigkeit bei hoher Effizienz; Randomisierte Verfahren (z. B. randomisierte SVD): Hohe Genauigkeit, setzt Kenntnis der kompletten Matrix voraus; siehe Diskussion in [Wang u. a., 2016].

jeweils zufällig aus der Gesamtpopulation ausgewählt. Um den durch die Zufallsauswahl eingeführten Bias zu reduzieren, wird diese zufällige Auswahl und die zugehörige Modellerzeugung und -auswertung abhängig von n_{ts} mehrfach wiederholt ($30\times$). Da die Lungenregion der Daten im Atlasraum ca. $n_{vx} = 3 \times 10^7$ Voxel umfasst, ist die Bestimmung der nächsten Korrelationsmatrix in Kapitel 7.2.1 praktisch nicht möglich (siehe auch Kapitel 7.2.4). Aus diesem Grund wird für den Hauptteil der Evaluation in der Lungenregion im Atlasbild (siehe Kapitel 5.2) eine Menge von $n_{vx} = 500$ Merkmalspunkten ausgewählt, die mittels Förstner-Operator (siehe Kapitel 6.2) automatisiert detektiert werden und zumeist an Bifurkationen der Gefäßbäume platziert sind. Diese Reduktion erlaubt eine effiziente Modellgenerierung und -auswertung im Rahmen der Evaluation.

Für die Manipulation der Kovarianzmatrix zur Integration der Lokalitätsannahme wird einerseits die in Kapitel 7.2.1 eingeführte Distanzfilterung (Lokalität-S; k_S in Kapitel 7.2.3 mit Suche nach nächster Korrelationsmatrix) eingesetzt, zum Vergleich aber auch der in Kapitel 7.2.3 diskutierte (positiv semidefinite) Gaußkernel (Lokalität-G; k_G mit $\gamma = 1/(2\tau^2)$) für eine Distanzgewichtung der Kovarianzen genutzt. Die Lokalitätsmodelle werden jeweils mit $n_l = 3$ Lokalitätsleveln erzeugt. Als Schwellenwertsequenz wird $\tau_l = (d_{max}, \frac{4}{9}d_{max}, \frac{8}{27}d_{max})$ verwendet, wobei d_{max} die maximale Distanz zwischen zwei Merkmalspunkten im Atlasbild bezeichnet. Als Distanzmaß $d(\cdot, \cdot)$ wird die euklidische Distanz im Atlasraum genutzt. Der Wert für k_{τ_l} (Dimension des Unterraums) wird so gewählt, dass jeweils 95% der Gesamtvarianz abgedeckt werden. Für alle Lokalitätsmodelle wird der Fusionierungsansatz aus Kapitel 7.2.2 verwendet.

Die Evaluation der Modelle erfolgt anhand von drei verschiedenen Kriterien, die aus dem Bereich der Formmodellierung übernommen werden (siehe [Davies, 2002]): (1) Generalisierungsfähigkeit, (2) Spezifität und (3) Kompaktheit. Die Generalisierungsfähigkeit beschreibt die Fähigkeit des Modells zur Repräsentation unbekannter Bewegungen. Für diesen Test werden die Testdaten auf den durch das Modell aufgespannten Unterraum projiziert (Glg. 6.4 und Glg. 6.5 mit $\eta = 0$) und der Approximationsfehler berechnet. Der Approximationsfehler ist analog zu Kapitel 5.3 gegeben als die mittlere vektorweise Abweichung des approximierten Verschiebungsfeldes vom mittels Bildregistrierung geschätzten Feld (Pseudo-Grundwahrheit). Die Spezifität quantifiziert die Plausibilität der durch das Modell generierbaren Bewegungen. Hierfür werden zufällig 1000 Verschiebungsfelder durch das Modell generiert (Glg. 6.4; Komponenten von \mathbf{b} gezogen aus Normalverteilung mit $\sigma = 1$) und die Distanz zum nächsten Verschiebungsfeld in der Gesamtpopulation bestimmt. Die Distanz wird hier auf die gleiche Weise wie der Approximationsfehler berechnet. Die Kompaktheit ist ein Maß für die Anzahl der Modellparameter, was hier der Dimension des Unterraums entspricht.

Um neben den generellen Modelleigenschaften auch die Praxistauglichkeit des Ansatzes zu zeigen, wird zusätzlich für $n_{ts} = 10$ eine modellbasierte Registrierung der Bilder der Phasen EI und EE für die Testdaten durch das modellgeführte Block-Matching-Verfahren (Modell-BM) aus Kapitel 6.3 durchgeführt. Die Registrierung erfolgt

anhand einer zufällig ausgewählten Untermenge von $n_{fp} = 100$ Merkmalspunkten aus der Menge der $n_{vx} = 500$ Merkmalspunkte, die für die Modellierung verwendet werden. Es kommen dieselben Parameter zur Modellgenerierung und dieselbe Auswertungsstrategie (Abweichung des geschätzten Verschiebungsfeldes zum Ausgangsfeld) zum Einsatz wie bei der Messung der Generalisierungsfähigkeit. Durch diesen Test wird zwar die generelle Praxistauglichkeit der Lokalitätsmodelle im Rahmen der modellbasierten Registrierung gezeigt, da aber nur $n_{vx} = 500$ Merkmalspunkte modelliert werden, muss in der Praxis das erwünschte dichte Verschiebungsfeld der kompletten Körperregion mit einem zusätzlichen (unspezifischen) Interpolationsverfahren erzeugt werden. Aus diesem Grund werden in einem weiteren Test Lokalitätsmodelle mit Gaußkernel für die modellbasierte Registrierung mit Modell-BM erzeugt und angewendet, die alle $n_{vx} = 3 \times 10^7$ Lungenvoxel umfassen und relativ effizient mittels approximativer Eigenwertzerlegung berechnet werden können (siehe Diskussionen in Kapitel 7.2.3 und Kapitel 7.2.4). Für die Registrierung werden wieder $n_{fp} = 100$ Merkmalspunkte verwendet. Die Registrierungsexperimente werden jeweils $15 \times$ wiederholt, um den Einfluss der Zufallsauswahl der Trainingsdaten zu verringern.

7.4 Ergebnisse

Tabelle 7.1, Abb. 7.6 und Abb. 7.7 fassen die Ergebnisse der Experimente zusammen. Alle aufgeführten Ergebnisse in Tab. 7.1 sind Mittelwerte über alle zehn Testdatensätze und Wiederholungen. Die statistische Signifikanz der Unterschiede zwischen den Modellen in Tab. 7.1 wird für die gleiche Anzahl von Testdatensätzen durch gepaarte t-Tests (zweiseitig; Signifikanzniveau: 5%) anhand der Einzelergebnisse der zehn Testdatensätze ermittelt, welche über alle (bis zu 30) Wiederholungen gemittelt werden.

Die Ergebnisse der Experimente zur Generalisierungsfähigkeit in Tab. 7.1 und Abb. 7.7 zeigen, dass durch die Integration der Lokalitätsannahme (Schwellenwert oder Gaußkernel) verglichen mit dem Standardmodellierungsverfahren für alle untersuchten Trainingsdatensätze Modelle mit deutlich verbesserter Generalisierungsfähigkeit erzeugt werden. Beispielsweise wird für Lokalität-S schon mit zehn Trainingsdatensätzen im Mittel ein Approximationsfehler erreicht, der für das Standardmodell erst bei 30–40 Trainingsdatensätzen erreichbar ist. Alle Unterschiede bezüglich der Generalisierungsfähigkeit zwischen dem Standardmodell und Lokalität-S/Lokalität-G in Tab. 7.1 sind statistisch signifikant ($p < 0.001$). Diese Verbesserung der Generalisierungsfähigkeit durch die Lokalitätsannahme geht allerdings mit einer Verringerung der Spezifität der jeweiligen Modelle einher (Erhöhung des mittleren Fehlers in Abb. 7.7). Hierbei ist allerdings zu beachten, dass bei einer genaueren Analyse des Zusammenhangs von Generalisierungsfähigkeit und Spezifität (siehe Abb. 7.7; Generalisierung vs. Spezifität) sichtbar wird, dass die Lokalitätsmodelle auch bei zum Standardmodell vergleichbarer Spezifität eine bessere Generalisierungsfähigkeit aufweisen. Eine reduzierte Spezifität

Tabelle 7.1: Ergebnisse zu den Experimenten zur Generalisierungsfähigkeit und modellbasierten Bildregistrierung mit Standardmodell und Lokali tätsmodellen in Abhängigkeit von der Anzahl der Trainingsdatensätze n_{ts} . Die Fehler sind angegeben als Mittelwert \pm Standardabweichung der mittleren Abweichung/des mittleren Fehlers bezogen auf die registrierungsbasiert ermittelte Pseudo-Grundwahrheit für die zehn Testdatensätze in den (bis zu) 30 Wiederholungen der Experimente (siehe Text für Details). Für die modellbasierte Registrierung mit $n_{vx} = 500$ Merkmalspunkten ist zusätzlich noch das Ergebnis bei anschließender TPS-Interpolation zur Erzeugung eines Verschiebungsfeldes für die komplette Lungenregion angegeben. Die mittlere Bewegung in den Testdaten beträgt für $n_{vx} = 500$ 10.34 ± 2.29 mm und für die komplette Lungenregion 9.24 ± 2.19 mm. Die zufällig gewählten Trainingsdaten sind in jeder Zeile für alle Modelle gleich.

#Trainingsdaten	Mittlerer Fehler [mm]		
	Standardmodell	Lokalität-S	Lokalität-G
Generalisierungsfähigkeit			
5	4.08 ± 1.37	3.22 ± 0.93	3.05 ± 0.86
10	3.38 ± 1.02	2.63 ± 0.67	2.38 ± 0.55
20	2.93 ± 0.85	2.38 ± 0.60	2.02 ± 0.44
30	2.68 ± 0.75	2.28 ± 0.56	1.88 ± 0.40
40	2.53 ± 0.68	2.25 ± 0.55	1.81 ± 0.37
60	2.33 ± 0.59	2.15 ± 0.50	1.71 ± 0.34
Modellbasierte Registrierung			
10 ($n_{vx} = 500$)	$3.57 \pm 1.21/3.46 \pm 1.09$	$2.89 \pm 0.81/2.90 \pm 0.74$	$2.68 \pm 0.71/2.71 \pm 0.64$
10 ($n_{vx} = 3 \times 10^7$)	3.47 ± 1.09	-/-	2.73 ± 0.68

kann potenziell bei der Modellnutzung in der Praxis zu Problemen führen, da beispielsweise Überanpassungseffekte auftreten können. Die Experimente zur modellbasierten Registrierung mit den Lokali tätsmodellen (siehe Tab. 7.1) zeigen allerdings, dass auch für diese Praxisbeispiele die Lokali tätsmodelle in der Lage sind, deutlich verbesserte Ergebnisse als das Standardmodell zu erzielen (alle Unterschiede statistisch signifikant mit $p < 0.001$). Dies zeigt die Praxistauglichkeit des entwickelten Verfahrens.

Bei direktem Vergleich der Ergebnisse von Lokalität-S (Schwellwert für Lokalitätsannahme) und Lokalität-G (Gaußkernel für Lokalitätsannahme) zeigt sich, dass die Nutzung des Gaußkernels zur Manipulation der Kovarianzmatrizen zu besseren Ergebnissen in den Experimenten zur Generalisierungsfähigkeit und bei der modellbasierten Registrierung führt. Alle Unterschiede zwischen Lokalität-S und Lokalität-G in Tab. 7.1 sind statistisch signifikant mit $p < 0.001$. Dieses Resultat ist für die Generalisierungsfähigkeit generell erwartungsgemäß, da durch den Gaußkernel tendenziell mehr Einträge der Kovarianzmatrix verändert werden als bei der Nutzung eines harten Schwellwertes (siehe Abb. 7.7), was sich u. a. auch an der deutlich geringeren

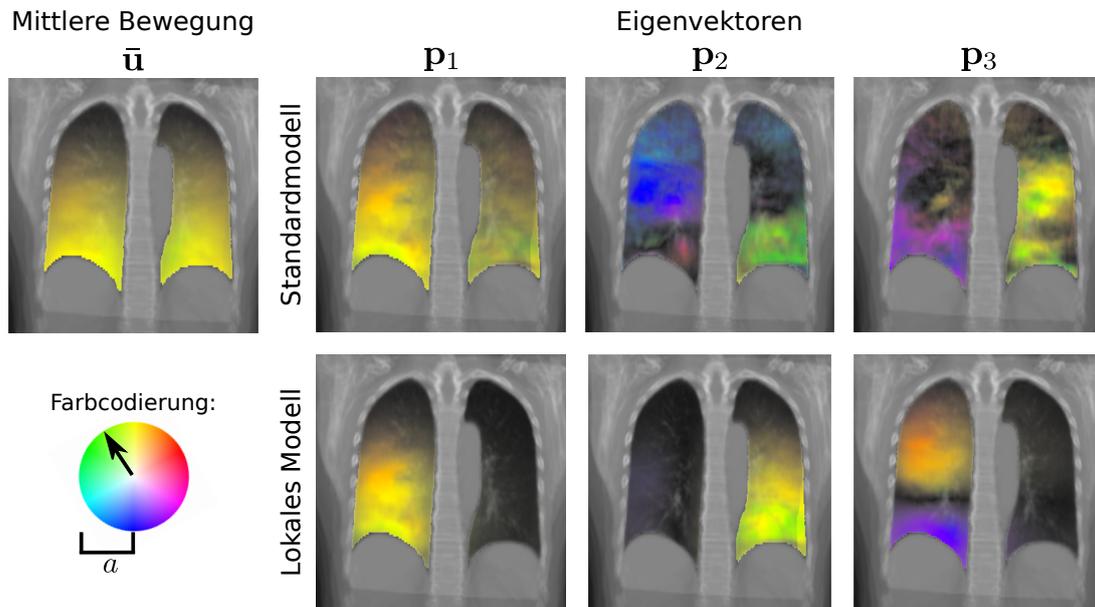


Abb. 7.6: Beispielhafte Visualisierung der modellierten Bewegung in einer ausgewählten koronaren Schicht des Atlasbildes bei Nutzung von $n_{ts} = 10$ Trainingsdatensätzen und aller $n_{vx} = 3 \times 10^7$ Lungenvoxel. Oben: Standardmodell. Unten: Lokales Modell (Lokalität-G für $\tau = \frac{4}{9}d_{max}$; 2. Lokalitätslevel in den Experimenten). Neben der mittleren Bewegung, die für beide Modelle gleich ist, werden spaltenweise auch die ersten drei Eigenvektoren/Basisvektoren der beiden Modelle visualisiert. Die dargestellte Bewegung entspricht einer Skalierung des jeweiligen Eigenvektors \mathbf{p}_i mit $2\sqrt{\lambda_i}$. Die Sättigung bzw. Deckkraft der Farbcodierung ist bildspezifisch gewählt. Volle Sättigung/volle Deckkraft ist erreicht für Amplituden $\geq a$ ($\bar{\mathbf{u}}$: $a = 16.89$ mm; Standardmodell: $a = 14.11/6.87/6.06$ mm für $\mathbf{p}_1/\mathbf{p}_2/\mathbf{p}_3$; Lokalität-G: $a = 12.50/7.99/8.00$ mm für $\mathbf{p}_1/\mathbf{p}_2/\mathbf{p}_3$). Das Standardmodell beschreibt hauptsächlich globale Bewegungsmuster, wohingegen im lokalen Modell eine Trennung der Lungenflügel sichtbar ist.

Kompaktheit der Modelle erkennen lässt. Diese deutlich geringere Kompaktheit geht allerdings nicht mit einer Verringerung der Spezifität gegenüber Lokalität-S einher und Lokalität-G erreicht bei ähnlichem Spezifitätslevel eine bessere Generalisierungsfähigkeit (siehe Abb. 7.7; Generalisierung vs. Spezifität). Weiterhin erlaubt die Nutzung des positiv definiten Gaußkerns die direkte Erzeugung von Modellen für die komplette Lungenregion. Interessanterweise führt die Nutzung von Lokalität-S für $n_{vx} = 500$ mit TPS-basierter Interpolation eines Verschiebungsfeldes für die komplette Lungenregion bei der modellbasierten Registrierung allerdings nur zu leicht schlechteren Ergebnissen.

7.5 Zusammenfassung und Diskussion

PCA-basierte statistische Bewegungsmodelle werden u. a. als Vorwissen über plausible Bewegungsmuster in Verfahren zur intrainerventionellen Bildregistrierung eingesetzt. Die Flexibilität bzw. Generalisierungsfähigkeit dieser Modelle und damit auch die Genauigkeit des sie nutzenden Verfahrens wird in der Praxis oftmals durch die mangelnde Verfügbarkeit geeigneter Trainingsdaten negativ beeinflusst. In diesem Kapitel wurde deshalb ein Verfahren zur Generierung flexibler statistischer Bewegungsmodelle auf Basis weniger Trainingsdaten vorgestellt, welches sich im Vergleich zu anderen Verfahren durch seinen hohen Bezug zu den vorhandenen Trainingsdaten auszeichnet.

Bei der PCA-basierten Generierung von Bewegungsmodellen mit wenigen Trainingsdaten werden hauptsächlich globale Zusammenhänge in den Daten modelliert, da lokale Variationen in einer kleinen Datenmenge kaum isoliert auftreten. Um diese lokalen Variationen im Modell besser berücksichtigen zu können, und somit die Flexibilität des Modells zu erhöhen, müssen diese lokalen Variationen während des Modellierungsprozesses stärker isoliert werden. Dies erfolgt in dem hier präsentierten Verfahren durch die Annahme eines lokal begrenzten Wirkungsbereichs von Bewegungseffekten. Diese sogenannte Lokalitätsannahme wird durch die Manipulation der Kovarianzmatrix in den Standardmodellierungsprozess integriert. Hierbei werden die Kovarianzen der Bewegungen im Körper weit entfernter Punkte komplett entfernt (Distanzfilterung) oder alternativ auch nur verringert (Distanzgewichtung). Auf diese Weise können in Abhängigkeit von den gewählten Gewichtungparametern bzw. dem Distanzschwellenwert lokale Modelle mit unterschiedlichen Lokalitätsleveln erzeugt werden. Um eine Nutzung dieser Modelle in Verfahren zu ermöglichen, die sonst Standardmodelle einsetzen, wird ein Fusionierungsprozess vorgeschlagen, der lokale Modelle mittels Optimierung auf der Graßmann-Mannigfaltigkeit linearer Unterräume zu einem einzelnen Modell zusammenfasst. Dieses Gesamtmodell integriert dann globale und lokale Variationen.

Die durchgeführte Evaluation im Kontext der populationsbasierten Bewegungsmodellierung konnte bestätigen, dass die durch das vorgestellte Verfahren erzeugten Modelle eine deutlich höhere Flexibilität aufweisen, als bei gleicher Anzahl von Trainingsdaten erzeugte Standardmodelle. Es konnte zudem in einem Praxisbeispiel gezeigt werden, dass diese Flexibilität zu einer höheren Genauigkeit der Bewegungsschätzung bei der modellbasierten Registrierung führt. Die Evaluation hat weiterhin gezeigt, dass die Distanzgewichtung mit einem Gaußkernel im Vergleich zur Distanzfilterung zu teilweise deutlich besseren Ergebnissen führt, ohne hierbei die Plausibilität der durch das Modell generierbaren Bewegungen zu verringern. Ein Grund hierfür könnte die Nutzung der registrierungsbasiert ermittelten Verschiebungsfelder als Pseudo-Grundwahrheit sein, da die für die Registrierung genutzte diffusive Regularisierung Ähnlichkeiten zu einer Distanzgewichtung mittels Gaußkernel aufweist, woraus sich ein entsprechender Bias ergeben könnte. Generell ist hinsichtlich der Evaluation zu beachten, dass diese nur ein eingeschränktes Szenario (Bewegung zwischen maximaler Ein- und Ausatmung

in einem Populationsszenario) untersucht hat. Deshalb sollten in zukünftigen Studien weitere Daten verwendet werden, um auch zusätzliche Bewegungsvariationen abbilden und deren Auswirkungen auf die Flexibilitätsverbesserungen untersuchen zu können.

Eine detaillierte Analyse des Verfahrens im Kontext von Kernelfunktionen hat ausführlich die Gründe für die bei einer binären Distanzfilterung zumeist auftretende Indefinitheit/negative Definitheit der manipulierten Kovarianzmatrix aufgezeigt. Wurde diese Problematik zunächst durch eine Korrektur der Kovarianzmatrix gelöst, so hat die Analyse gezeigt, dass dies u. a. auch durch die Nutzung einer positiv semidefiniten Kernelfunktion (z. B. Distanzgewichtung mittels Gaußkernel) möglich ist. Hierdurch konnte zudem ein direkter Bezug zur sehr ähnlichen und weitestgehend zeitgleich entstandenen Methodik von Lüthi u. a. [2018] hergestellt werden, welche die Betrachtung der Distanzgewichtung in diesem Kapitel letztlich motiviert hat.

Die Betrachtung des Verfahrens mittels Kernelfunktionen sollte zudem als Ansatzpunkt für weitere Analysen und daraus ableitbare methodische Verbesserungen sein. Die Nutzung der Distanzgewichtung mittels Gaußfunktion hat in der durchgeführten Evaluation bei ähnlicher Spezifität zu besseren Ergebnissen als die binäre Distanzgewichtung geführt. Hier wäre es interessant zu untersuchen, ob dies u. a. auch in der gewählten Konfiguration der Modelle (Lokalitätslevel) begründet ist. Je nach Ergebnis einer solchen Untersuchung könnte es dann weiterhin sinnvoll sein die Distanzfilterung direkt durch eine positiv semidefinite Kernelfunktion zu ersetzen, die die binäre Filterung möglichst gut nachahmt, aber deutlich weniger Einträge manipuliert als eine gaußbasierte Filterung. Hierdurch würde die explizite Speicherung der Kovarianzmatrix wegfallen können. Genereller und anwendungsunabhängiger Vorteil der hier vorgestellten Kovarianzmatrixmanipulation mit anschließender Korrektur, um die Semidefinitheit wieder herzustellen, ist allerdings die extrem große Flexibilität dieses Vorgehens, da hierdurch beliebiges anatomisches oder physiologisches Vorwissen in den Manipulationsprozess integriert werden kann. Dieses Potenzial wurde in diesem Kapitel zwar nicht genutzt, hat sich aber beispielsweise bei der Formmodellierung in [Wilms u. a., 2017a] als sehr hilfreich erwiesen. Insgesamt wäre es deshalb interessant Möglichkeiten zur Bestimmung der nächsten positiv semidefiniten Kernelfunktion zu einer beliebigen Filterung-/Gewichtungsfunktion zu untersuchen. Weiterhin erscheint eine genauere Analyse der Unterschiede/Ähnlichkeiten des in diesem Kapitel vorgeschlagenen expliziten Fusionierungsansatzes und der eher impliziten und kernelbasierten Lösung zur Integration mehrerer Lokalitätslevel in [Lüthi u. a., 2018] interessant.

Weiterhin wäre auch die Untersuchung anderer Anwendungsszenarien für die flexiblen Modelle sinnvoll. Neben der Form- und Erscheinungsmodellierung [Wilms u. a., 2017a,b] konnte das Verfahren unter Mitwirkung des Autors dieser Dissertation in [Uzunova u. a., 2017] auch schon erfolgreich zur Erzeugung von Trainingsdaten für die Deep Learning-basierte Bildregistrierung eingesetzt werden. Im Kontext des Deep Learnings sollte es deshalb noch weiteres Einsatzpotenzial geben.

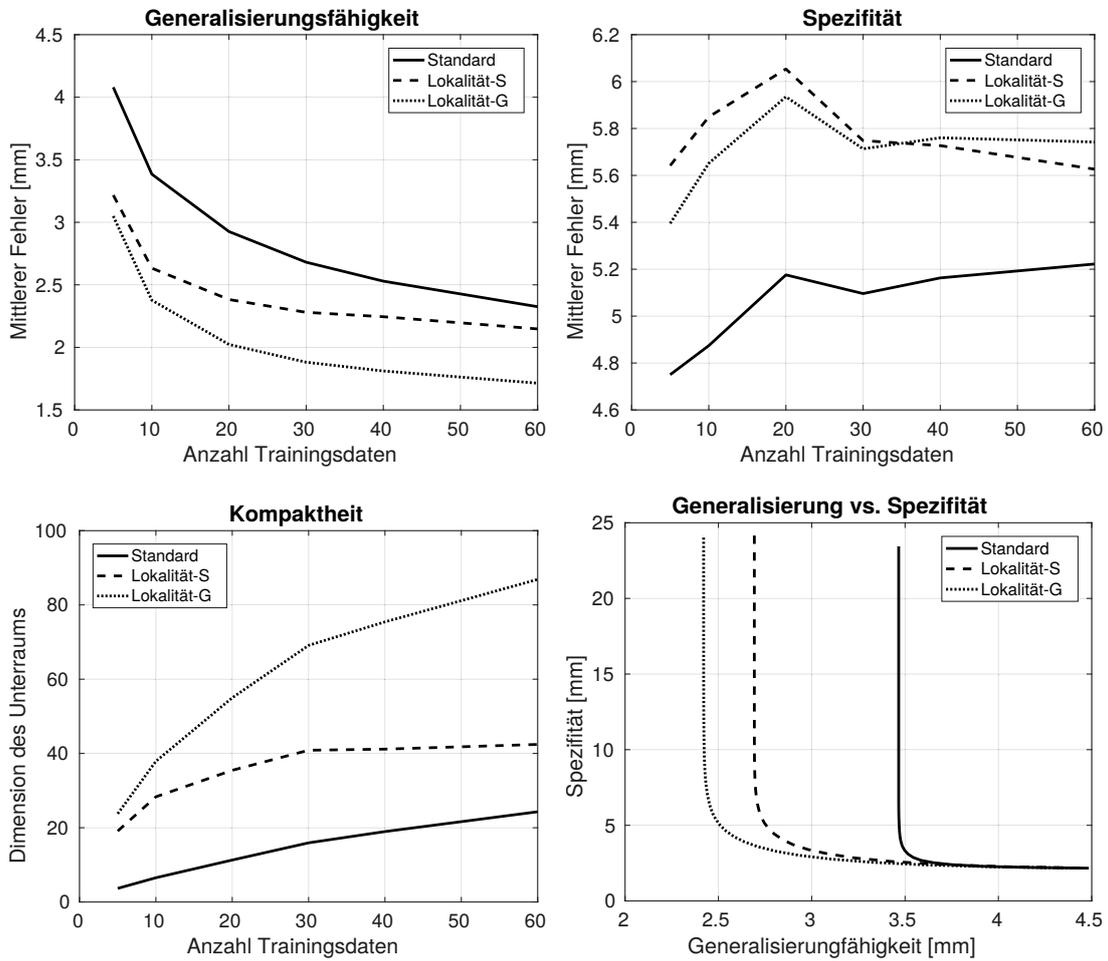


Abb. 7.7: Grafische Darstellung der Ergebnisse zu den Experimenten zur Generalisierungsfähigkeit, Spezifität und Kompaktheit in Abhängigkeit von der Anzahl der Trainingsdatensätze n_{ts} für die drei verglichenen Modelle. Alle Werte sind Mittelwerte über alle Testdatensätze und Wiederholungen. Kleinere Werte weisen auf bessere Modelle hin. Zusätzlich zu den drei Einzelkriterien zeigt die grafische Darstellung der Spezifität in Abhängigkeit von der Generalisierungsfähigkeit für $n_{ts} = 10$ (rechts unten) auf, inwieweit sich die Modelle bei gleicher Spezifität/Generalisierungsfähigkeit unterscheiden. Verschiedene Spezifitätslevel wurden durch Veränderung des Parameters σ bei Ziehung der Parametervektoren \mathbf{b} erzeugt (siehe Text für Details).

Kapitel 8

Zusammenfassung und Ausblick

Kapitelinhalt

8.1	Wissenschaftlicher Beitrag	161
8.2	Ausblick	165

Die Atembewegung stellt ein zentrales Problem bei der Behandlung thorakaler und abdominaler Tumoren mittels externer Strahlentherapie oder nichtinvasiver Thermoablation dar. Für eine effektive technische Bewegungskompensation muss die Atembewegung intrainerventionell in Echtzeit geschätzt werden. Hierfür werden u. a. Signale externer Bewegungsindikatoren (Surrogatbasierte Bewegungsschätzung) oder intrainerventionell akquirierte Bilddaten als Ausgangsbasis verwendet, welche dann zumeist von modellbasierten Verfahren weiterverarbeitet werden. Bei der surrogatbasierten Bewegungsschätzung erzeugt ein gelerntes Korrespondenzmodell aus einer Signalmessung die korrespondierende interne Atembewegung. Bildregistrierungsverfahren für die Bewegungsschätzung mit intrainerventionellen Bilddaten nutzen häufig statistische Bewegungsmodelle als Vorwissen über plausible Atembewegungen.

Die vorliegende Dissertation beschäftigt sich mit der Entwicklung, Erweiterung und Evaluation von Verfahren zur Atembewegungsmodellierung und modellbasierten Schätzung der intrainerventionellen Atembewegung in den o.g. Anwendungsbereichen. Generelles Ziel war es, modellbasierte Verfahren zu entwickeln bzw. durch Veränderungen so zu verbessern, dass diese (1) eine hohe Genauigkeit bei der Bewegungsschätzung erreichen, (2) echtzeitfähig sind und (3) die Atembewegung ganzer Körperregionen schätzen können, um komplexe intrainerventionelle Anwendungen bzw. Analysen (z. B. Dosisakkumulation) zu ermöglichen. Im Folgenden werden die Entwicklungen und Untersuchungen dieser Arbeit zusammengefasst und einige Ideen für Folgearbeiten diskutiert.

8.1 Wissenschaftlicher Beitrag

Die methodischen Entwicklungen und Untersuchungen in dieser Dissertation lassen sich in drei Teile unterteilen: (1) Surrogatbasierte Bewegungsschätzung, (2) modellbasierte

Bildregistrierung für die Bewegungsschätzung mit intrainerventionellen Bilddaten und (3) statistische Bewegungsmodellierung mit wenigen Trainingsdaten. Im Folgenden werden die Ergebnisse in diesen drei Teilbereichen zusammengefasst.

Surrogatbasierte Bewegungsschätzung: Von der patientenspezifischen zur populationsbasierten Korrespondenzmodellierung In letzter Zeit haben sich im Bereich der surrogatbasierten Bewegungsschätzung einige vielversprechende neue Entwicklungen ergeben (u. a. Nutzung von diffeomorphen Transformationen und Einsatz kamerabasierter Systeme zur hochdimensionalen Abtastung der Körperoberfläche). Bisher fehlen allerdings umfassende und systematische Vergleiche bezüglich der erreichbaren Bewegungsschätzungsgenauigkeit bei Einsatz verschiedener aktueller Modellierungsansätze und Bewegungsindikatoren in realistischen Anwendungsszenarien, um besser einschätzen zu können, welche Kombinationen in der Praxis die höchste Genauigkeit versprechen.

Ein Ziel dieser Arbeit war es, einen ersten Schritt in diese Richtung zu leisten. Den Ausgangspunkt für die Arbeiten zur surrogatbasierten Bewegungsschätzung in dieser Dissertation bildete ein bestehendes regressionsbasiertes Framework zur (diffeomorphen) Korrespondenzmodellierung, welches zunächst um projektionsbasierte Regressionstechniken (PCR, PLS, CCA) erweitert wurde, die eine verbesserte Verarbeitung hochdimensionaler Indikatordaten gegenüber Ridge-Regression versprechen. Weiterhin wurde eine Evaluationsumgebung entwickelt, welche durch die Verwendung räumlich-zeitlicher Bilddaten eine bildbasierte Simulation verschiedener Bewegungsindikatoren (u. a. Bauchgurte, Spirometrie und Oberflächenabtastungen) und Anwendungsszenarien ermöglicht, ohne dass z. B. reale Indikatordaten vorhanden sein müssen. Mithilfe dieser Umgebung wurden verschiedenste Kombinationen aus Modellierungstechnik (nicht-diffeomorph/diffeomorph und Regressionstechnik) und Indikatordaten für praxisnahe Szenarien mit unterschiedlichen Bewegungsvariabilitäten quantitativ evaluiert. Durch die Nutzung von räumlich-zeitlichen Bilddaten ergab sich die Möglichkeit etablierte landmarkenbasierte Evaluationsmaße der Bildregistrierung einzusetzen. Die Ergebnisse der durchgeführten Studie lassen folgende Aussagen zu:

- Die untersuchten patientenspezifischen Korrespondenzmodelle konnten auch bei Bewegungsvariabilität einen Großteil der intrainerventionellen Atembewegung in Echtzeit kompensieren (Rechenzeiten bei GPU-Implementierung < 10 ms (nicht-diffeomorph) bzw. < 100 ms (diffeomorph)). Die Genauigkeit verringert sich allerdings bei interfraktioneller Variabilität erheblich.
- Die Wahl des Bewegungsindikators hat (nahezu) keinen Einfluss auf die mittlere Genauigkeit, solange ein mehrdimensionales Signal (≥ 2) genutzt wird. Ob zu einem 1D-Signal zeitliche Informationen (Ableitung) hinzugefügt werden oder mehrdimensionale Signale (Abtastung Körperoberfläche) direkt erhoben werden, ist unerheblich.

- Projektionsbasierte Regressionstechniken liefern im Mittel keine genaueren Modelle als Ridge-Regression.
- Die Genauigkeitsunterschiede nicht-diffeomorpher und diffeomorpher Korrespondenzmodelle sind vernachlässigbar. Diffeomorphe Modelle verhindern aber zumeist erfolgreich physiologisch unplausible Singularitäten in den Transformationen.

Die Evaluationsumgebung lässt sich in Zukunft flexibel erweitern und bietet insbesondere bei Verwendung von öffentlich verfügbaren Bilddaten eine ausgezeichnete Grundlage für weitere Studien und öffentliche Challenges.

Wie auch die Studienergebnisse zeigen, haben interfraktionelle Bewegungsvariationen zwischen der Planungssitzung, deren Daten zum Modelltraining genutzt wurden, und den eigentlichen Behandlungssitzungen oftmals einen starken negativen Einfluss auf die Genauigkeit der surrogatbasierten Bewegungsschätzung mit patientenspezifischen Korrespondenzmodellen. In dieser Arbeit wurde ein Verfahren entwickelt, welches versucht dieses Problem zu verringern, indem populationsbasierte Informationen über mögliche interfraktionelle Bewegungsvariationen in den Trainingsprozess patientenspezifischer Korrespondenzmodelle integriert werden. Dieses Verfahren unterliegt zwei Annahmen: (1) In einer großen Patientenpopulation gibt es Subpopulationen von Patienten mit ähnlicher Atembewegung, deren Interpatientenvariabilität ungefähr der interfraktionellen Bewegungsvariabilität der betrachteten Patienten entspricht. (2) Ein Korrespondenzmodell, welches auf Grundlage dieser Subpopulation generiert wurde, kann interfraktionelle Bewegungsvariationen besser ausgleichen als ein patientenspezifisches Modell. Methodisch werden diese Subpopulationen mittels Ähnlichkeits-Clustering der Bewegungen einer Patientenpopulation identifiziert. Es konnte gezeigt werden, dass dieses neue Verfahren Korrespondenzmodelle generiert, welche bei dem Vorliegen interfraktioneller Bewegungsvariabilität eine signifikant genauere Bewegungsschätzung als rein patientenspezifische Modelle ermöglichen, wobei die Bestimmung der optimalen Parameter für das Ähnlichkeits-Clustering in der Praxis noch ein Problem darstellt. Insgesamt ist dies der erste Ansatz, der subpopulationsbasierte Korrespondenzmodellierung zur Berücksichtigung von interfraktioneller Bewegungsvariabilität nutzt.

Bisher wurden Korrespondenzmodelle hauptsächlich zur intrainerventionellen Bewegungsschätzung eingesetzt. Betrachtet man sie aber prinzipiell als Generatoren, die basierend auf einem Indikatorsignal eine zeitliche Folge von Transformationen erzeugen, die die Atembewegung beschreiben, lassen sich diese Modelle auch für die Atmungssimulation nutzen. Werden für die Atmungssimulation nur aufgenommene Indikatorsignale eingesetzt, ist die zeitliche Länge der Simulation an die Länge des Indikatorsignals gebunden und auch die Abfolge von Bewegungsvariationen wird durch das Signal begrenzt. Um diese Abhängigkeit zu reduzieren, wurde in dieser Arbeit ein Verfahren zur statistischen Modellierung und anschließenden Simulation von patientenspezifischen Indikatorsignalen auf der Basis von B-Splines entwickelt. Aufbauend auf einem realen Indikatorsignal kann das Verfahren variable Signale beliebi-

ger zeitlicher Länge simulieren, die u. a. typische Variationen bezüglich Zykluslänge und Ein-/Ausatmungstiefe enthalten. Hierdurch ergibt sich in Kombination mit einem Korrespondenzmodell die Möglichkeit zur Simulation von realistischen, variablen Bewegungssequenzen beliebiger zeitlicher Länge. Dieses Verfahren wurde exemplarisch zur 4D-Dosisimulation in der Strahlentherapieplanung eingesetzt, wodurch die Auswirkungen variabler intrafraktioneller Atembewegung auf die geplante Dosisverteilung statistisch untersucht werden konnten. Weiterhin wurde gezeigt, wie die Bewegungssimulation in einen Virtual-Reality-Simulator integriert werden kann, um die Immersion bei simulierten Punktionseingriffen in Thorax oder Abdomen deutlich zu verbessern.

Modellbasierte Bildregistrierung für die Bewegungsschätzung mit intrainerventionellen Bilddaten Die Integration schneller, überlagerungsfreier Bildgebungstechniken wie Magnetresonanztomographie und Ultraschall in Strahlentherapie- oder Ultraschallablationsgeräte eröffnet neue Möglichkeiten zur direkten, bildbasierten Schätzung der intrainerventionellen Atembewegung. Für entsprechende Bildregistrierungsverfahren ergeben sich allerdings zwei spezielle Herausforderungen: (1) Echtzeitfähigkeit trotz u. U. großer zu verarbeitender Datenmenge und (2) flexibler Umgang mit unvollständigen Daten (z. B. Registrierung zwischen 3D-Bild und 2D-Bildern). Verfügbare Verfahren lösen diese Probleme meist nur eingeschränkt oder stark situationsabhängig. In dieser Arbeit wurden deshalb zwei neue patch- und modellbasierte Registrierungsverfahren vorgeschlagen, die diesen Herausforderungen gerecht werden und sehr vielseitig einsetzbar sind. Hierbei wird die Registrierung initial auf die Umgebung (Patch) weniger ausgewählter Merkmalspunkte beschränkt, die mit großer Wahrscheinlichkeit in Referenz- und intrainerventionellem Bild sichtbar sind. Durch diese Datenreduktion ergibt sich eine hohe Verarbeitungsgeschwindigkeit und ein einfacher Umgang mit unvollständigen Daten. Die Atembewegung abseits der Merkmalspunkte wird mithilfe eines patientenspezifischen statistischen Bewegungsmodells interpoliert, welches zudem bei der Registrierung als Regularisierer bzw. Vorwissen genutzt wird.

Die beiden entwickelten Verfahren unterscheiden sich hauptsächlich durch ihre Lösung des patchbasierten Registrierungsproblems. Das erste Verfahren setzt diskretes Block-Matching ein, welches durch das Bewegungsmodell iterativ regularisiert wird. Das zweite Verfahren basiert auf einem eigentlich für Segmentierungszwecke entwickelten Active Appearance Model, welches neben der Bewegung auch Bildinformationen statistisch modelliert und für dessen Anpassung kontinuierliche Optimierung eingesetzt wird. In der durchgeführten Evaluation konnte gezeigt werden, dass beide Verfahren eine ähnliche, hohe Genauigkeit erreichen, die nur gering (teilweise insignifikant) von der Genauigkeit eines klassischen pixelweisen Registrierungsverfahrens abweicht. Beide Verfahren sind schon bei der Nutzung von CPU-Implementierungen für viele Szenarien echtzeitfähig (deutlich unter 50 ms für 2D-2D-Registrierung; \approx 100 ms für die Registrierung eines 3D-Referenzbildes und multiplen 2D-Bildern; $<$ 500 ms für 3D-3D-

Registrierung; alle Zeiten bei Nutzung von Grauwertpatches) und die Rechenzeiten können durch GPU-Implementierungen perspektivisch noch drastisch gesenkt werden.

Statistische Bewegungsmodellierung mit wenigen Trainingsdaten Ein häufiges Problem bei der Nutzung von (PCA-basierten) statistischen Bewegungsmodellen ist die mangelnde Verfügbarkeit von repräsentativen Trainingsdaten, was die Flexibilität bzw. Generalisierungsfähigkeit der Modelle und damit auch die Genauigkeit der entsprechenden modellbasierten Bewegungsschätzungsverfahren einschränkt. Deshalb wurde im Rahmen dieser Arbeit ein Verfahren zur Generierung flexibler statistischer Bewegungsmodelle mit wenigen Trainingsdaten entwickelt. Werden Bewegungsmodelle mit wenigen Trainingsdaten generiert, so modellieren diese hauptsächlich globale Zusammenhänge. Um auch lokale Variationen im Modell besser berücksichtigen zu können, wurde eine sogenannte Lokalitätsannahme eingeführt, nach der die Bewegungen räumlich weit entfernter Bereiche (relativ) unabhängig voneinander sein sollen. Diese Lokalitätsannahme lässt sich durch distanzbasierte Kovarianzmatrixmanipulationen ausdrücken und so in den Standardmodellierungsprozess integrieren. Durch die Nutzung verschiedener Distanzwerte zur Definition der räumlichen Unabhängigkeit können lokale Modelle mit unterschiedlichen Lokalitätsleveln generiert werden. Um eine Nutzung dieser Modelle in Verfahren zu ermöglichen, die sonst Standardmodelle einsetzen, wurde zudem ein Fusionierungsprozess entwickelt, der mehrere lokale Modelle zu einem einzelnen flexiblen Modell zusammenfassen kann, welches dann globale und lokale Variationen repräsentiert. In einer Evaluation zur populationsbasierten Bewegungsschätzung konnte u. a. gezeigt werden, dass durch Nutzung dieser flexiblen Modelle, verglichen mit Standardmodellen, eine signifikante Genauigkeitsverbesserung der modellbasierten Registrierung erreichbar ist.

Insgesamt konnten in dieser Arbeit mehrere neue Verfahren im Kontext der surrogatbasierten Bewegungsschätzung und der modellbasierten Registrierung intrainerventioneller Bilddaten entwickelt werden, die spezifische Verbesserungen gegenüber bisher verfügbaren Verfahren u. a. hinsichtlich Genauigkeit oder Vielseitigkeit aufweisen und so einen Beitrag zur Verbesserung der intrainerventionellen Bewegungsschätzung in der klinischen Praxis leisten.

8.2 Ausblick

Im Folgenden werden einige allgemeine Ideen für mögliche weiterführende Untersuchungen und Entwicklungen diskutiert. Detaillierte Diskussionen möglicher Weiterentwicklungen der einzelnen in dieser Arbeit präsentierten Verfahren bzw. Untersuchungen sind in den jeweiligen Einzelkapiteln zu finden.

Umfassende, vergleichende Untersuchung von indirekter und direkter Bewegungsschätzung Die in dieser Arbeit entwickelten Methoden zur indirekten, surrogatbasierten und direkten, registrierungsbasierten Bewegungsschätzung sind hier hauptsächlich isoliert voneinander untersucht worden. Um eine bessere Einschätzung der durch die direkten Verfahren erreichbaren Genauigkeitsverbesserungen im Vergleich zu indirekten Verfahren zu ermöglichen, könnte in Anlehnung bzw. als Fortführung der Studie in [Paganelli u. a., 2015] die in Kapitel 3 für die surrogatbasierte Bewegungsschätzung präsentierte Evaluationsumgebung genutzt werden, um indirekte und direkte Methoden systematisch miteinander zu vergleichen.

Zeitliche Regularisierung und Prädiktion der Bewegung Vorwissen bezüglich der in weiten Teilen räumlichen Glattheit der Atembewegung wird in klassische Registrierungsverfahren zur Bewegungsschätzung in 4D-Bilddaten u. a. durch den Einsatz eines entsprechenden Regularisierers integriert. Die in dieser Arbeit präsentierten Verfahren schätzen ebenfalls räumlich glatte Transformationen, da die genutzten Modelle auf Basis glatter Transformationen trainiert werden. Für die atmungsbedingte Bewegung anatomischer oder pathologischer Strukturen ist auch von einer gewissen zeitlichen Glattheit auszugehen. Dieses Vorwissen wird in dieser Arbeit allerdings bis auf eine Ausnahme (Experimente in Kapitel 6) im Rahmen der Bewegungsschätzung nicht genutzt, da die Bewegungsschätzung für jeden Interventionszeitpunkt isoliert betrachtet wird. Dies trifft auch auf viele klassische Registrierungsverfahren bei der Verarbeitung von 4D-Bilddaten oder andere Verfahren für die intrainventionelle Bewegungsschätzung zu (Ausnahme u. a. [O’Shea u. a., 2016b]).

Für die surrogatbasierte Bewegungsschätzung kann man insbesondere bei der Nutzung sehr niedrigdimensionaler Indikatorsignale argumentieren, dass bei Vorliegen eines zeitlich glatten (oder geglätteten) Signals auch die geschätzte interne Atembewegung eine entsprechende Glattheit aufweist. Für die hier vorgestellten bildbasierten Bewegungsschätzungsverfahren scheint allerdings eine explizite Berücksichtigung der zeitlichen Glattheit im Verfahren selbst perspektivisch sinnvoll, um beispielsweise die Robustheit gegenüber Bildartefakten zu erreichen. In diesem Kontext erscheint es auch sinnvoll die Bewegungsschätzung und die für den Ausgleich technisch bedingter Latenzzeiten zusätzlich notwendige Bewegungsprädiktion zu kombinieren (siehe auch [Dürichen, 2015]), um so beispielsweise bei der patch-basierten Registrierung die Informationen über zukünftige Positionen von Punkten auch für eine die Rechenzeit reduzierende Initialisierung und Beschränkung des Suchbereichs nutzen zu können.

Von linearer zu nichtlinearer statistischer Bewegungsmodellierung Die in dieser Arbeit genutzten statistischen Bewegungsmodelle basieren auf der Annahme, dass Transformationen, die plausible Atembewegungen beschreiben, in einem niedrigdimensionalen linearen bzw. affinen Unterraum eines sehr hochdimensionalen Vek-

torraums liegen. Diese Annahme führt zur relativ kompakten Modellen und u. a. der einfachen Berechenbarkeit von Projektionen auf diese Unterräume. Es stellt sich aber die Frage, ob diese Darstellung der (angenommenen) niedrigdimensionalen Struktur der Daten angemessen ist? In der statistischen Formmodellierung wurden deshalb u. a. schon mittels Kernel PCA generierte nichtlineare Modelle untersucht und es wurde gezeigt, dass diese Modelle eine höhere Spezifität bzw. eine höhere Robustheit bei der Modellanpassung aufweisen können [Kirschner u. a., 2011]. Kernel PCA wurde in [He u. a., 2010] auch schon erfolgreich für die Atmungsmodellierung eingesetzt und Baumgartner u. a. [2017] nutzen mittels Manifold Learning geschätzte niedrigdimensionale, nichtlineare Mannigfaltigkeiten für die Atmungsmodellierung. Entsprechende Untersuchungen scheinen deshalb auch im Kontext der hier präsentierten Verfahren sinnvoll.

Integration von Informationen über die Unsicherheit der Schätzung Die in dieser Arbeit präsentierten Verfahren liefern zu jeder gegebenen Indikatormessung oder einem intrainerventionellen Bild genau eine Transformation. Was hierbei fehlt, sind Angaben hinsichtlich der Unsicherheiten dieser Schätzung, die u. a. während der Intervention genutzt werden können, um diese zu unterbrechen, wenn die Schätzung zu unsicher ist [Samei u. a., 2013]. Große Unsicherheiten in der Schätzung können beispielsweise durch starke Abweichungen der Indikatormessung von den Trainingsdaten oder Ambiguitäten in Bereichen mit niedrigem Kontrast bei der modellbasierten Bildregistrierung begründet sein. Dieses Thema ist bisher trotz der hohen praktischen Relevanz kaum im Kontext der intrainerventionellen Bewegungsschätzung bearbeitet worden. In den Gaussian Process Models von Dürichen u. a. [2015] für die surrogatbasierte Bewegungsschätzung und -prädiktion werden Normalverteilungen geschätzt, deren Varianz als Maß für die Unsicherheit der Schätzung genutzt werden kann. Peressutti u. a. [2013b] präsentieren ein Bayessches Verfahren zur kombinierten Bewegungsschätzung mittels Indikatormessung und Bildern, welches ebenfalls ein Maß für die Unsicherheit der Schätzung mitliefert. Samei u. a. [2013] stellen ein Verfahren zur Ermittlung von Konfidenzregionen der Bewegungsschätzung mittels PCA-basierter Bewegungsmodellen vor, wenn nur wenige Teile des Verschiebungsfeldes bekannt sind (vgl. dünnbesetzte Verschiebungsfelder in Kapitel 6), welches in Teilen auf einem ähnlichen Ansatz für die PCA-basierte Formmodellierung basiert [Blanc u. a., 2012a]. Insbesondere die Ansätze aus [Blanc u. a., 2012a; Samei u. a., 2013] bilden eine gute Grundlage für eine erste Ermittlung von Unsicherheiten bei der modellbasierten Registrierung mit den Verfahren aus Kapitel 6.

Nutzung von Deep Learning-Techniken Deep Learning-Techniken haben in den letzten Jahren sehr erfolgreich Einzug in die medizinische Bildverarbeitung gehalten und wurden zuletzt auch verstärkt für die schnelle und genaue Bewegungsschätzung/Registrierung genutzt [Litjens u. a., 2017]. Im Kontext dieser Arbeit sind diese

Techniken insbesondere für die Bewegungsschätzung mit intrainerventionellen Bilddaten interessant. So ließe sich das Registrierungsproblem zwischen Referenz- und intrainerventionellem Bild ähnlich zu z. B. [Uzunova u. a., 2017] als Regressionsproblem auffassen, welches mit einem Convolutional Neural Network (CNN) gelöst werden kann. Vorteil eines solchen Vorgehens ist u. a. die durch das Netz optimierte Extraktion für das Endergebnis relevanter Bildmerkmale, sodass die manuelle Definition von passenden Bilddeskriptoren wegfällt. Da diese CNNs sehr viele Parameter haben, muss eine entsprechend große Menge von Trainingsdaten vorhanden sein, um eine gute Generalisierung zu erreichen. Neben der Nutzung von künstlich erzeugten Daten (z. B. Verfahren aus Kapitel 7 in [Uzunova u. a., 2017]), kann auch wieder eine patchbasierte Definition verwendet werden, um die Komplexität des Regressionsproblems zu verringern [Yang u. a., 2017]. Durch die üblicherweise genutzten Encoder/Decoder-Netzarchitekturen findet auch in diesen Verfahren (implizit) die Abbildung der Daten auf eine niedrigdimensionale Struktur statt. Explizit können auch PCA-basierte Modelle als Vorwissen in diese Verfahren integriert werden [Milletari u. a., 2017; Yu u. a., 2016]. Beispielsweise Generative Adversarial Models (GANs) erlauben es allerdings auch explizit mittels Deep Learning nichtlineare Bewegungsmodelle zu erzeugen, die dann als Vorwissen weiterverwendet werden können [Hu u. a., 2017]. Weiterhin existieren auch Deep Learning-Techniken zur Integration zeitlicher Informationen, sodass auch eine zeitliche Regularisierung der Bewegungsschätzung auf diese Weise umgesetzt werden kann (siehe z. B. [Hou u. a., 2018]). Es ist zu erwarten, dass diese Techniken in der nächsten Zeit auch für die intrainerventionelle Bewegungsschätzung und -modellierung eingesetzt werden und so neue und komplett lernbasierte Verfahren entstehen.

Einsatz der patch- und modellbasierten Registrierungsverfahren außerhalb der Bewegungsschätzung PCA-basierte statistische Modelle plausibler Transformationen (statistische Deformationsmodelle) können auch außerhalb der Bewegungsschätzung in anderen Anwendungen (Populationsstudien, Atlas-zu-Patient-Registrierung, usw.) als Vorwissen bei der Registrierung eingesetzt werden, um eine robuste und schnelle Schätzung der gesuchten Transformation zu ermöglichen (Übersicht in [Sotiras u. a., 2013]). Ein konkretes Beispiel hierfür ist die in den Neurowissenschaften notwendige Abbildung von Patientenbildern des Gehirns auf entsprechende Standardanatomien (z. B. MNI brain), um Populationsstudien in einem einheitlichen Referenzsystem durchführen zu können [Klein u. a., 2009]. Robustheit ist hier insbesondere gegenüber den pathologischen Veränderungen in den Patientendaten nötig und da in Populationsstudien entsprechend viele Transformationen geschätzt werden müssen, ist ebenso auch die zeitliche Effizienz der Registrierung wichtig. Es wäre deshalb interessant die hier präsentierten patch- und modellbasierten Registrierungsverfahren in einem solchen Szenario einzusetzen, insbesondere auch in Kombination mit dem präsentierten Verfahren zur Erzeugung flexibler statistischer Modelle mit wenigen Trainingsdaten.

Anhang A

Ergänzende Ergebnisse zu Kapitel 3

A.1 Ergebnisse für DIRLAB- und CREATIS-Daten

Um in der Zukunft einen Vergleich des in Kapitel 2 präsentierten diffeomorphen Korrespondenzmodellierungsverfahrens mit anderen Verfahren zur (surrogatbasierten) in-trainterventionellen Bewegungsschätzung zu ermöglichen, werden in Tab. A.1 die Einzelergebnisse aus Tab. 3.1 für die frei verfügbaren DIRLAB- und CREATIS-Datensätze aufgelistet. Hierbei ist zu beachten, dass diese Daten nur für einen Teil von Experiment 1 eingesetzt wurden (siehe Kapitel 3.1.3).

Tabelle A.1: Teilergebnisse von Experiment 1 zur surrogatbasierten Bewegungsschätzung angegeben als Mittelwert \pm Standardabweichung des mittleren TRE oder der mittleren symmetrischen Oberflächendistanz getrennt nach DIRLAB- und CREATIS-Daten. Die Mittelung erfolgte über alle für das jeweilige Experiment verfügbaren Einzelergebnisse. Die Ergebnisse sind als Blöcke nach genutztem Regressionsverfahren/Dimension des Surrogatsignals geordnet.

Bewegungsschätzung	Mittlerer TRE [mm]	
	Experiment 1 (4D-CT)	
	DIRLAB	CREATIS
	EI \rightarrow EE	EI \rightarrow EE
Mittlere Bewegung	8.46 \pm 3.33	8.12 \pm 3.09
<i>VariationalRegistration</i>	1.67 \pm 0.71	1.05 \pm 0.18
<i>Surrogatbasierte Bewegungsschätzung mit \mathbf{B}^{RR} und 1D-Surrogatsignal</i>		
Spirometrie	2.27 \pm 1.06	2.08 \pm 1.11
Bauchgurt	2.27 \pm 1.03	2.13 \pm 1.11
Zwerchfellbewegung	2.38 \pm 1.38	1.71 \pm 0.42
Tiefenbild Bauchgurt	2.33 \pm 1.00	2.19 \pm 1.42
Tiefenbild Brustbein	4.06 \pm 2.30	3.52 \pm 2.26
<i>Surrogatbasierte Bewegungsschätzung mit \mathbf{B}^{RR} und 2D-Surrogatsignal</i>		
Spirometrie	2.19 \pm 1.06	2.08 \pm 1.13
Bauchgurt	2.09 \pm 0.79	2.21 \pm 1.15
Zwerchfellbewegung	2.41 \pm 1.30	1.89 \pm 0.51
Tiefenbild Bauchgurt	2.06 \pm 0.74	2.49 \pm 1.44
Tiefenbild Brustbein	3.29 \pm 1.99	2.81 \pm 2.26
<i>Surrogatbasierte Bewegungsschätzung mit \mathbf{B}^{RR} und Tiefenbildern</i>		
10 Punkte	2.26 \pm 0.86	2.55 \pm 1.39
100 Punkte	2.21 \pm 0.87	1.56 \pm 0.43
1000 Punkte	2.11 \pm 0.76	1.90 \pm 0.66
10000 Punkte	2.15 \pm 0.73	2.11 \pm 1.00
<i>Surrogatbasierte Bewegungsschätzung mit \mathbf{B}^{PCR} und Tiefenbildern</i>		
10 Punkte	2.19 \pm 0.82	3.64 \pm 2.81
100 Punkte	2.17 \pm 0.85	1.64 \pm 0.39
1000 Punkte	2.16 \pm 0.85	1.74 \pm 0.62
10000 Punkte	2.15 \pm 0.85	1.50 \pm 0.46
<i>Surrogatbasierte Bewegungsschätzung mit \mathbf{B}^{PLS} und Tiefenbildern</i>		
10 Punkte	2.23 \pm 0.84	3.11 \pm 2.39
100 Punkte	2.17 \pm 0.85	1.60 \pm 0.36
1000 Punkte	2.08 \pm 0.81	1.77 \pm 0.61
10000 Punkte	2.08 \pm 0.82	1.47 \pm 0.39
<i>Surrogatbasierte Bewegungsschätzung mit \mathbf{B}^{CCA} und Tiefenbildern</i>		
10 Punkte	2.19 \pm 0.77	2.52 \pm 1.56
100 Punkte	2.37 \pm 1.14	2.07 \pm 1.40
1000 Punkte	2.27 \pm 1.01	2.19 \pm 1.41
10000 Punkte	2.09 \pm 0.86	1.78 \pm 0.81

Anhang B

Eigene Publikationen mit Bezug zur Dissertation

Anmerkung: Geteilte Erstautorenschaften des Autors dieser Dissertation sind mit dem Symbol * nach dem Vornamen gekennzeichnet.

Beiträge in Fachzeitschriften

1. Fortmeier, D., **Wilms, M.**, Mastmeyer, A. und Handels, H. “Direct Visuo-Haptic 4D Volume Rendering Using Respiratory Motion Models”. *IEEE Trans Haptics* 8 (4), 2015, S. 371–383.
2. Ha, I. Y.* , **Wilms, M.***, Handels, H. und Heinrich, M. P. “Model-based Sparse-to-dense Image Registration for Realtime Respiratory Motion Estimation in Image-guided Interventions”. *IEEE Trans Biomed Eng* 66 (2), 2019, S. 302–310.
3. Mastmeyer, A., **Wilms, M.** und Handels, H. “Interpatient Respiratory Motion Model Transfer for Virtual Reality Simulations of Liver Punctures”. *Journal of World Society of Computer Graphics - WSCG* 25 (1), 2017, S. 1–10.
4. Ortmüller, J., Gauer, T., **Wilms, M.**, Handels, H. und Werner, R. “Respiratory Surface Motion Measurement by Microsoft Kinect: Implementation and Evaluation of a Clinical Setup”. *Current Directions in Biomedical Engineering* 1, 2015, S. 270–273.
5. Sothmann, T., Gauer, T., **Wilms, M.** und Werner, R. “Correspondence model-based 4D VMAT dose simulation for analysis of local metastasis recurrence after extracranial SBRT”. *Phys Med Biol* 62 (23), 2017, S. 9001–9017.
6. **Wilms, M.***, Werner, R.* , Ehrhardt, J., Schmidt-Richberg, A., Schlemmer, H.-P. und Handels, H. “Multivariate regression approaches for surrogate-based diffeomorphic estimation of respiratory motion in radiation therapy”. *Phys Med Biol* 59 (5), 2014, S. 1147–1164.

7. **Wilms, M.**, Werner, R., Blendowski, M., Ortmüller, J., Handels, H. u. a. “Simulation of Range Imaging-based Estimation of Respiratory Lung Motion”. *Methods Inform Med* 53 (4), 2014, S. 257–263.
8. **Wilms, M.**, Handels, H. und Ehrhardt, J. “Multi-resolution multi-object statistical shape models based on the locality assumption”. *Med Image Anal* 38 (5), 2017, S. 17–29.
9. **Wilms, M.***, Werner, R.* , Yamamoto, T., Handels, H. und Ehrhardt, J. “Subpopulation-based correspondence modelling for improved respiratory motion estimation in the presence of inter-fraction motion variations”. *Phys Med Biol* 62 (14), 2017, S. 5823–5839.

Beiträge in Konferenzbänden und Konferenzabstracts

1. Blendowski, M., **Wilms, M.**, Werner, R. und Handels, H. “Simulation of Range-Imaging-Based Prediction of Respiratory Organ and Tumor Motion Using 4D CT Data: Influence of Signal Dimensionality and Sampling Patterns”. In: *58. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS)*. 2013, S. 216–217.
2. — “Simulation und Evaluation tiefenbildgebender Verfahren zur Prädiktion atemungsbedingter Organ- und Tumorbewegungen”. In: *Bildverarbeitung für die Medizin, BVM 2013*. Springer, 2013, S. 350–355.
3. Ehrhardt, J., **Wilms, M.** und Handels, H. “Patch-based low-rank matrix completion for learning of shape and motion models from few training samples”. In: *European Conference on Computer Vision – ECCV 2016*. Bd. 9908. LNCS. 2016, S. 712–727.
4. Mastmeyer, A., **Wilms, M.**, Fortmeier, D., Schröder, J. und Handels, H. “Real-Time Ultrasound Simulation for Training of US-Guided Needle Insertion in Breathing Virtual Patients”. In: *Stud Health Technol Inform – MMVR22*. Bd. 220. 2016, S. 219–26.
5. Mastmeyer, A., **Wilms, M.** und Handels, H. “Interpatientenübertragung von Atemmodellen für das Virtual-Reality-Training von Punktionseingriffen”. In: *Bildverarbeitung für die Medizin 2017*. Informatik aktuell. Springer, 2017, S. 340–345.
6. Mastmeyer, A., **Wilms, M.** und Handels, H. “Population-based respiratory 4D motion atlas construction and its application for VR simulations of liver punctures”. In: *SPIE Medical Imaging 2018: Image Processing*. Bd. 10574. Proc SPIE. 2018, S. 1057417.

-
7. Mastmeyer, A., **Wilms, M.** und Handels, H. “Populationsbasierte 4D Bewegungsatlanten für VR Simulationen”. In: *Bildverarbeitung für die Medizin 2018*. Informatik aktuell. Springer, 2018, S. 200.
 8. Ortmüller, J., **Wilms, M.**, Werner, R. und Handels, H. “Kombination von Atemsignalen zur Optimierung der Prädiktion komplexer atmungsbedingter Organ- und Tumorbewegungen”. In: *Bildverarbeitung für die Medizin 2014*. Informatik aktuell. Springer, 2014, S. 72–77.
 9. Ortmüller, J., Werner, R., **Wilms, M.**, Handels, H. und Gauer, T. “Einsatz der Microsoft Kinect zur Analyse zeitsynchron erfasster thorakaler und abdominaler Atemsignale”. In: *21. Jahrestagung der Deutschen Gesellschaft für Radioonkologie*. Bd. 191 (Suppl. 1). Strahlentherapie und Onkologie. 2015, S65.
 10. Uzunova, H.* , **Wilms, M.*** , Handels, H. und Ehrhardt, J. “Training CNNs for Image Registration from Few Samples with Model-based Data Augmentation”. In: *Med Image Comput Comput Assist Interv – MICCAI 2017*. Bd. 10433. LNCS. Springer, 2017, S. 223–231.
 11. Werner, R., Ehrhardt, J., Schmidt-Richberg, A., **Wilms, M.**, Blendowski, M. und Handels, H. “A Diffeomorphic MLR Framework for Surrogate-based Motion Estimation in Radiation Therapy: Concept and First Evaluation”. In: *Proc. Informatik 2012*. 2012, S. 1774–1785.
 12. Werner, R., **Wilms, M.**, Ehrhardt, J., Schmidt-Richberg, A., Blendowski, M. und Handels, H. “A Diffeomorphic MLR Framework for Surrogate-based Motion Estimation and Situation-adapted Dose Accumulation”. In: *Image-Guidance and Multimodal Dose Planning in Radiation Therapy: a MICCAI Workshop*. 2012, S. 42–49.
 13. Werner, R., **Wilms, M.**, Schmidt-Richberg, A., Ehrhardt, J., Gauer, T., Petersen, C. und Handels, H. “Incorporation of Surrogate-Based Information about Breathing Motion Variations into 4D Dose Calculation: Concept and First Result”. In: *43. Jahrestagung der Deutschen Gesellschaft für Medizinische Physik (DGMP)*. 2012, S. 390–393.
 14. Werner, R., Gauer, T., Petersen, C., **Wilms, M.** und Handels, H. “Respiratory Motion and its Variability in Lung and Liver SBRT: DVH Analysis Exploiting 4DCT-Based Dose Accumulation”. In: *2nd ESTRO Forum 2013 of the European Society for Radiotherapy and Oncology*. Bd. 106 (Suppl. 2). Radiotherapy and Oncology. 2013, S166–S167.

15. **Wilms, M.**, Werner, R., Ehrhardt, J., Schmidt-Richberg, A., Blendowski, M. und Handels, H. “Surrogate-based Diffeomorphic Motion Estimation for Radiation Therapy: Comparison of Multivariate Regression Approaches”. In: *SPIE Medical Imaging 2013: Image Processing*. Bd. 8669. Proc SPIE. 2013, S. 866915.
16. **Wilms, M.**, Ehrhardt, J., Werner, R., Marx, M. und Handels, H. “Statistical Analysis of Surrogate Signals to Incorporate Respiratory Motion Variability into Radiotherapy Treatment Planning”. In: *SPIE Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and Modeling*. Bd. 9036. Proc SPIE. 2014, 90360J.
17. **Wilms, M.**, Fortmeier, D., Mastmeyer, A. und Handels, H. “Modellbasierte Simulation der Atembewegung für das Virtual-Reality-Training von Punktionseingriffen”. In: *Bildverarbeitung für die Medizin 2015*. Informatik aktuell. Springer, 2015, S. 317–322.
18. **Wilms, M.**, Werner, R., Yamamoto, T., Handels, H. und Ehrhardt, J. “Population-based correspondence models for respiratory motion estimation in the presence of inter-fraction motion variations”. In: *MICCAI Workshop on Imaging and Computer Assistance in Radiation Therapy*. 2015, S. 81–88.
19. **Wilms, M.**, Ha, I. Y., Handels, H. und Heinrich, M. P. “Model-based Regularisation for Respiratory Motion Estimation with Sparse Features in Image-guided Interventions”. In: *Med Image Comput Comput Assist Interv – MICCAI 2016*. Bd. 9902. LNCS. Springer, 2016, S. 89–97.
20. **Wilms, M.**, Handels, H. und Ehrhardt, J. “Learning of Representative Multi-Resolution Multi-Object Statistical Shape Models from Small Training Populations”. In: *Bildverarbeitung für die Medizin 2017*. Informatik aktuell. Springer, 2017, S. 359–360.
21. — “Patch-Based Learning of Shape, Appearance, and Motion Models from Few Training Samples by Low-Rank Matrix Completion”. In: *Bildverarbeitung für die Medizin 2017*. Informatik aktuell. Springer, 2017, S. 215–216.
22. — “Representative Patch-based Active Appearance Models Generated from Small Training Populations”. In: *Med Image Comput Comput Assist Interv – MICCAI 2017*. Bd. 10433. LNCS. Springer, 2017, S. 152–160.

Literaturverzeichnis

Anmerkung: Geteilte Erstautorenschaften des Autors dieser Dissertation sind mit dem Symbol * nach dem Vornamen gekennzeichnet.

- [Ackerman, 1998] Ackerman, M. J. “The Visible Human Project”. *Proc IEEE* 86 (3), 1998, S. 504–511.
- [Admiraal u. a., 2008] Admiraal, M. A., Schuring, D. und Hurkmans, C. W. “Dose calculations accounting for breathing motion in stereotactic lung radiotherapy based on 4D-CT and the internal target volume”. *Radiother Oncol* 86 (1), 2008, S. 55–60.
- [Arsigny, 2006] Arsigny, V. “Processing Data in Lie Groups: An Algebraic Approach. Application to Non-Linear Registration and Diffusion Tensor MRI”. Thèse de sciences. École polytechnique, 2006.
- [Ashburner u. a., 2011] Ashburner, J. und Friston, K. J. “Diffeomorphic registration using geodesic shooting and Gauss–Newton optimisation”. *NeuroImage* 55 (3), 2011, S. 954–967.
- [Avants u. a., 2005] Avants, B. B., Epstein, C. L. und Gee, J. C. “Geodesic image interpolation: Parameterizing and interpolating spatiotemporal images”. In: *Variational, Geometric, and Level Set Methods in Computer Vision*. Springer, 2005, S. 247–258.
- [Avants u. a., 2010] Avants, B. B., Yushkevich, P., Pluta, J., Minkoff, D., Korczykowski, M., Detre, J. und Gee, J. C. “The optimal template effect in hippocampus studies of diseased populations”. *NeuroImage* 49 (3), 2010, S. 2457–2466.
- [Baker u. a., 2004] Baker, S. und Matthews, I. “Lucas-kanade 20 years on: A unifying framework”. *Int J Comput Vis* 56 (3), 2004, S. 221–255.
- [Banerjee u. a., 2015] Banerjee, J., Klink, C., Peters, E. D., Niessen, W. J., Moelker, A. und Walsum, T. van. “Fast and robust 3D ultrasound registration – Block and game theoretic matching”. *Med Image Anal* 20 (1), 2015, S. 173–183.
- [Baumgartner u. a., 2017] Baumgartner, C. F., Kolbitsch, C., McClelland, J. R., Rueckert, D. und King, A. P. “Autoadaptive motion modelling for MR-based respiratory motion estimation”. *Med Image Anal* 35 (1), 2017, S. 83–100.
- [Beddar u. a., 2007] Beddar, A. S., Kainz, K., Briere, T. M., Tsunashima, Y., Pan, T., Prado, K., Mohan, R., Gillin, M. und Krishnan, S. “Correlation between

- internal fiducial tumor motion and external marker motion for liver tumors imaged with 4D-CT". *Int J Radiat Oncol Biol Phys* 67 (2), 2007, S. 630–638.
- [Beg u. a., 2005] Beg, M. F., Miller, M. I., Trounev, A. und Younes, L. "Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms". *Int J Comput Vis* 61 (2), 2005, S. 139–157.
- [Benchetrit, 2000] Benchetrit, G. "Breathing pattern in humans: diversity and individuality". *Respir Physiol* 122 (2), 2000, S. 123–129.
- [Bishop, 2006] Bishop, C. M. *Pattern Recognition and Machine Learning*. 1. Auflage. Springer, 2006.
- [Bjerre u. a., 2013] Bjerre, T., Crijns, S., Rosenschöld, P. af, Aznar, M., Specht, L., Larsen, R. und Keall, P. "Three-dimensional MRI-linac intra-fraction guidance using multiple orthogonal cine-MRI planes". *Phys Med Biol* 58 (14), 2013, S. 4943–4950.
- [Blanz u. a., 2002] Blanz, V. und Vetter, T. "Reconstructing the Complete 3D Shape of Faces from Partial Information (Rekonstruktion der dreidimensionalen Form von Gesichtern aus partieller Information)". *it-Information Technology* 44 (6), 2002, S. 295–302.
- [Blanc u. a., 2012a] Blanc, R. und Szekely, G. "Confidence Regions for Statistical Model Based Shape Prediction From Sparse Observations". *IEEE Trans Med Imag* 31 (6), 2012, S. 1300–1310.
- [Blanc u. a., 2012b] Blanc, R., Seiler, C., Székely, G., Nolte, L.-P. und Reyes, M. "Statistical model based shape prediction from a combination of direct observations and various surrogates: Application to orthopaedic research". *Med Image Anal* 16 (6), 2012, S. 1156–1166.
- [Booth u. a., 2016] Booth, J. T., Caillet, V., Hardcastle, N., O'Brien, R., Szymura, K., Crasta, C., Harris, B., Haddad, C., Eade, T. und Keall, P. J. "The first patient treatment of electromagnetic-guided real time adaptive radiotherapy using MLC tracking for lung SABR". *Radiother Oncol* 121 (1), 2016, S. 19–25.
- [Bortfeld u. a., 2002] Bortfeld, T., Jokivarsi, K., Goitein, M., Kung, J. und Jiang, S. B. "Effects of intra-fraction motion on IMRT dose delivery: statistical analysis and simulation". *Phys Med Biol* 47 (13), 2002, S. 2203–2220.
- [Borga u. a., 1997] Borga, M., Landelius, T. und Knutsson, H. *A Unified Approach to PCA, PLS, MLR and CCA*. Technical Report LiTH-ISY-R-1992. Linköping University, The Institute of Technology, 1997.
- [Boyd u. a., 2004] Boyd, S. und Vandenberghe, L. *Convex Optimization*. 1. Auflage. Cambridge University Press, 2004.
- [Boye u. a., 2013] Boye, D., Samei, G., Schmidt, J., Székely, G. und Tanner, C. "Population based modeling of respiratory lung motion and prediction from partial

- information". In: *SPIE Medical Imaging 2013: Image Processing*. Bd. 8669. Proc SPIE. 2013, 86690U-1-7.
- [Brandner u. a., 2006] Brandner, E. D., Wu, A., Chen, H., Heron, D., Kalnicki, S., Komanduri, K., Gerszten, K., Burton, S., Ahmed, I. und Shou, Z. "Abdominal organ motion measured using 4D CT". *Int J Radiat Oncol Biol Phys* 65 (2), 2006, S. 554-560.
- [Britton u. a., 2007] Britton, K. R., Starkschall, G., Tucker, S. L., Pan, T., Nelson, C., Chang, J. Y., Cox, J. D., Mohan, R. und Komaki, R. "Assessment of Gross Tumor Volume Regression and Motion Changes During Radiotherapy for Non-Small-Cell Lung Cancer as Measured by Four-Dimensional Computed Tomography". *Int J Radiat Oncol Biol Phys* 68 (4), 2007, S. 1036-1046.
- [Brix u. a., 2014] Brix, L., Ringgaard, S., Sorensen, T. S. und Poulsen, P. R. "Three-dimensional liver motion tracking using real-time two-dimensional MRI". *Med Phys* 41 (4), 2014, S. 042302-1-10.
- [Bussels u. a., 2003] Bussels, B., Goethals, L., Feron, M., Bielen, D., Dymarkowski, S., Suetens, P. und Haustermans, K. "Respiration-induced movement of the upper abdominal organs: a pitfall for the three-dimensional conformal radiation treatment of pancreatic cancer". *Radiother Oncol* 68 (1), 2003, S. 69-74.
- [Cachier u. a., 2003] Cachier, P., Bardinet, E., Dormont, D., Penneç, X. und Ayache, N. "Iconic feature based nonrigid registration: the PASHA algorithm". *Comput Vis Image Underst* 89 (2-3), 2003, S. 272-298.
- [Castillo u. a., 2009] Castillo, R., Castillo, E., Guerra, R., Johnson, V. E., McPhail, T., Garg, A. K. und Guerrero, T. "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets". *Phys Med Biol* 54 (7), 2009, S. 1849-1870.
- [Castillo u. a., 2010] Castillo, R., Castillo, E., Martinez, J., Shenoy, M. und Guerrero, T. "Four-dimensional deformable image registration using trajectory modeling". *Phys Med Biol* 55 (1), 2010, S. 305-327.
- [Castillo u. a., 2013] Castillo, R., Castillo, E., Fuentes, D., Ahmad, M., Wood, A. M., Ludwig, M. S. und Guerrero, T. "A reference dataset for deformable image registration spatial accuracy evaluation using the COPDgene study archive". *Phys Med Biol* 58 (9), 2013, S. 2861-2877.
- [Cerviño u. a., 2010] Cerviño, L. I., Jiang, Y., Sandhu, A. und Jiang, S. B. "Tumor motion prediction with the diaphragm as a surrogate: a feasibility study". *Phys Med Biol* 55 (9), 2010, N221.
- [Cerviño u. a., 2011] Cerviño, L. I., Du, J. und Jiang, S. B. "MRI-guided tumor tracking in lung cancer radiotherapy". *Phys Med Biol* 56 (13), 2011, S. 3773-3785.

- [Cerrolaza u. a., 2012] Cerrolaza, J. J., Villanueva, A. und Cabeza, R. “Hierarchical statistical shape models of multiobject anatomical structures: application to brain MRI”. *IEEE Trans Med Imag* 31 (3), 2012, S. 713–724.
- [Chang, 2008] Chang, J.-M. “Classification on the grassmannians: theory and applications”. PhD Thesis. Colorado State University, 2008.
- [Chen u. a., 2001] Chen, Q., Weinhaus, M. S., Deibel, F. C., Ciezki, J. P. und Macklis, R. M. “Fluoroscopic study of tumor motion due to breathing: Facilitating precise radiation therapy for lung cancer patients”. *Med Phys* 28 (9), 2001, S. 1850–1856.
- [Chen u. a., 2008] Chen, M., Lu, W., Chen, Q., Ruchala, K. J. und Olivera, G. H. “A simple fixed-point approach to invert a deformation field”. *Med Phys* 35 (1), 2008, S. 81–88.
- [Cootes u. a., 2001] Cootes, T. F., Edwards, G. J. und Taylor, C. J. “Active appearance models”. *IEEE Trans Pattern Anal Mach Intell* 23 (6), 2001, S. 681–685.
- [Cootes u. a., 1995] Cootes, T. F. und Taylor, C. J. “Combining point distribution models with shape models based on finite element analysis”. *Image Vis Comput* 13 (5), 1995, S. 403–409.
- [Cootes u. a., 1996] ———. “Data Driven Refinement of Active Shape Model Search.” In: *British Machine Vision Conference – BMVC 1996*. 1996, S. 1–10.
- [Dalca u. a., 2016] Dalca, A. V., Bobu, A., Rost, N. S. und Golland, P. “Patch-Based Discrete Registration of Clinical Brain Images”. In: *Patch Based Tech Med Imaging – Patch-MI 2016*. Bd. 9993. LNCS. Springer, 2016, S. 60–67.
- [Davatzikos u. a., 2003] Davatzikos, C., Tao, X. und Shen, D. “Hierarchical active shape models, using the wavelet transform”. *IEEE Trans Med Imag* 22 (3), 2003, S. 414–423.
- [Davies, 2002] Davies, R. H. “Learning shape: optimal models for analysing natural variability”. Dissertation. University of Manchester, 2002.
- [De Luca u. a., 2012] De Luca, V., Tanner, C. und Székely, G. “Speeding-up image registration for repetitive motion scenarios”. In: *IEEE International Symposium on Biomedical Imaging – ISBI 2012*. 2012, S. 1355–1358.
- [De Luca u. a., 2013] De Luca, V., Tschannen, M., Székely, G. und Tanner, C. “A Learning-based Approach for Fast and Robust Vessel Tracking in Long Ultrasound Sequences”. In: *Med Image Comput Comput Assist Interv – MICCAI 2013*. Bd. 8149. LNCS. Springer, 2013, S. 518–525.
- [De Luca u. a., 2015] De Luca, V., Benz, T., Kondo, S. und et. al. “The 2014 liver ultrasound tracking benchmark”. *Phys Med Biol* 60 (14), 2015, S. 5571–5599.
- [Denis de Senneville u. a., 2012] Denis de Senneville, B., Ries, M., Bartels, L. W. und Moonen, C. T. W. “Interventional Magnetic Resonance Imaging”. In: Springer,

2012. Kap. MRI-Guided High-Intensity Focused Ultrasound Sonication of Liver and Kidney, S. 349–366.
- [Denis de Senneville u. a., 2015] Denis de Senneville, B., El Hamidi, A. und Moonen, C. “A direct PCA-based approach for real-time description of physiological organ deformations”. *IEEE Trans Med Imag* 34 (4), 2015, S. 974–982.
- [Depuydt u. a., 2013] Depuydt, T., Poels, K., Verellen, D., Engels, B., Collen, C., Haverbeke, C., Gevaert, T., Buls, N., Gompel, G. V., Reynders, T., Duchateau, M., Tournel, K., Boussaer, M., Steenbeke, F., Vandenbroucke, F. und Ridder, M. D. “Initial assessment of tumor tracking with a gimbaled linac system in clinical circumstances: A patient simulation study”. *Radiother Oncol* 106 (2), 2013, S. 236–240.
- [Dhont u. a., 2017] Dhont, J., Vandemeulebroucke, J., Burghelca, M., Poels, K., Depuydt, T., Van Den Begin, R., Jaudet, C., Collen, C., Engels, B., Reynders, T. u. a. “The long- and short-term variability of breathing induced tumor motion in lung and liver over the course of a radiotherapy treatment”. *Radiother Oncol* 126 (2), 2017, S. 339–346.
- [Dinkel u. a., 2009] Dinkel, J., Tetzlaff, R., Hintze, C., Thieke, C., Rochet, N., Delorme, S. und Biederer, J. “4D Lung Imaging with Time-resolved 3D MRI: Comparison of Different Sequence Parameters for the Assessment of Tumor Size and Image Quality in a Ventilated Ex-vivo Setup”. *Int J Radiat Oncol Biol Phys* 75 (3, Supplement), 2009. Proceedings of the American Society for Radiation Oncology 51st Annual Meeting, S145–S146.
- [Dürichen u. a., 2015] Dürichen, R., Pimentel, M. A. F., Clifton, L., Schweikard, A. und Clifton, D. A. “Multitask Gaussian Processes for Multivariate Physiological Time-Series Analysis”. *IEEE Trans Biomed Eng* 62 (1), 2015, S. 314–322.
- [Durrleman, 2010] Durrleman, S. “Statistical models of currents for measuring the variability of anatomical curves, surfaces and their evolution”. Dissertation. Université Nice Sophia Antipolis, 2010.
- [Dürichen, 2015] Dürichen, R. “From univariate to multivariate respiratory motion compensation. A Bayesian way to increase treatment accuracy in robotic radiotherapy”. Dissertation. Universität zu Lübeck, 2015.
- [Ehrhardt u. a., 2007] Ehrhardt, J., Werner, R., S., D. u. a. “An optical flow based method for improved reconstruction of 4D CT data sets acquired during free breathing.” *Med Phys* 34 (2), 2007, S. 711–721.
- [Ehrhardt u. a., 2010] Ehrhardt, J., Werner, R., Schmidt-Richberg, A. und Handels, H. “A Statistical Shape and Motion Model for the Prediction of Respiratory Lung Motion”. In: *SPIE Medical Imaging 2010: Image Processing*. Bd. 7623. Proc SPIE. 2010, S. 531–539.

- [Ehrhardt u. a., 2011] Ehrhardt, J., Werner, R., Schmidt-Richberg, A. und Handels, H. “Statistical Modeling of 4D Respiratory Lung Motion Using Diffeomorphic Image Registration.” eng. *IEEE Trans Med Imag* 30 (2), 2011, S. 251–65.
- [Ehrhardt u. a., 2013] Ehrhardt, J., Klinder, T. und Lorenz, C. “Computational motion phantoms and statistical models of respiratory motion”. In: *4D Modeling and Estimation of Respiratory Motion for Radiation Therapy*. Springer, 2013, S. 215–247.
- [Ehrhardt u. a., 2016] Ehrhardt, J., Wilms, M. und Handels, H. “Patch-based low-rank matrix completion for learning of shape and motion models from few training samples”. In: *European Conference on Computer Vision – ECCV 2016*. Bd. 9908. LNCS. 2016, S. 712–727.
- [Elhamifar u. a., 2011] Elhamifar, E. und Vidal, R. “Sparse manifold clustering and embedding”. *Conference on Neural Information Processing Systems – NIPS 2011*, 2011, S. 55–63.
- [Ernst u. a., 2009] Ernst, F., Martens, V., Schlichting, S., Beširević, A., Kleemann, M., Koch, C., Petersen, D. und Schweikard, A. “Correlating Chest Surface Motion to Motion of the Liver Using ϵ -SVR – A Porcine Study”. In: *Med Image Comput Comput Assist Interv – MICCAI 2009*. Bd. 5762. LNCS. Springer, 2009, S. 356–364.
- [Ernst u. a., 2012] Ernst, F., Bruder, R., Schlaefer, A. und Schweikard, A. “Correlation between external and internal respiratory motion: a validation study”. *Int J Comput Assist Radiol Surg* 7 (3), 2012, S. 483–492.
- [Ernst u. a., 2013] Ernst, F., Dürichen, R., Schlaefer, A. und Schweikard, A. “Evaluating and comparing algorithms for respiratory motion prediction”. *Phys Med Biol* 58 (11), 2013, S. 3911–3929.
- [Färber u. a., 2009] Färber, M., Hummel, F., Gerloff, C. und Handels, H. “Virtual Reality Simulator for the Training of Lumbar Punctures”. *Methods Inform Med* 48 (5), 2009, S. 493–501.
- [Fassi u. a., 2014] Fassi, A., Schaerer, J., Fernandes, M., Riboldi, M., Sarrut, D. und Baroni, G. “Tumor Tracking Method Based on a Deformable 4D CT Breathing Motion Model Driven by an External Surface Surrogate”. *Int J Radiat Oncol Biol Phys* 88 (1), 2014, S. 182–188.
- [Fayad u. a., 2011] Fayad, H., Pan, T., François Clement, J. und Visvikis, D. “Technical Note: Correlation of respiratory motion between external patient surface and internal anatomical landmarks”. *Med Phys* 38 (6Part1), 2011, S. 3157–3164.
- [Fayad u. a., 2012] Fayad, H., Pan, T., Pradier, O. und Visvikis, D. “Patient specific respiratory motion modeling using a 3D patient’s external surface”. *Med Phys* 39 (6), 2012, S. 3386–3395.

- [Fortmeier u. a., 2013] Fortmeier, D., Mastmeyer, A. und Handels, H. “Image-based Soft Tissue Deformation Algorithms for Real-time Simulation of Liver Puncture”. *Curr Med Imaging Rev* 9 (2), 2013, S. 154–165.
- [Fortmeier u. a., 2015] Fortmeier, D., Wilms, M., Mastmeyer, A. und Handels, H. “Direct Visuo-Haptic 4D Volume Rendering Using Respiratory Motion Models”. *IEEE Trans Haptics* 8 (4), 2015, S. 371–383.
- [Fortmeier u. a., 2016] Fortmeier, D., Mastmeyer, A., Schröder, J. und Handels, H. “A Virtual Reality System for PTCd Simulation using Direct Visuo-haptic Rendering of Partially Segmented Image Data”. *IEEE J Biomed Health Inform* 20 (1), 2016, S. 355–366.
- [Förstner u. a., 1987] Förstner, W. und Gülch, E. “A fast operator for detection and precise location of distinct points, corners and centres of circular features”. In: *ISPRS intercommission conference on fast processing of photogrammetric data*. 1987, S. 281–305.
- [Fortmeier, 2016] Fortmeier, D. “Direct volume rendering methods for needle insertion simulation”. Dissertation. Universität zu Lübeck, 2016.
- [Freifeld, 2014] Freifeld, O. “Statistics on Manifolds with Applications to Modeling Shape Deformations”. PhD Thesis. Brown University, 2014.
- [Fu u. a., 2007] Fu, D., Kahn, R., Wang, B., Wang, H., Mu, Z., Park, J., Kuduvalli, G. und Maurer, C. R. “Xsight Lung Tracking System: A Fiducial-Less Method for Respiratory Motion Tracking”. In: *Treating Tumors that Move with Respiration*. Springer, 2007, S. 265–282.
- [Fuerst u. a., 2015] Fuerst, B., Mansi, T., Carnis, F., Sälzle, M., Zhang, J., Declerck, J., Boettger, T., Bayouth, J., Navab, N. und Kamen, A. “Patient-specific biomechanical model for the prediction of lung motion from 4-D CT images”. *IEEE Trans Med Imag* 34 (2), 2015, S. 599–607.
- [Gao u. a., 2008] Gao, G., McClelland, J., Tarte, S., Blackall, J. und Hawkes, D. “Modelling the respiratory motion of the internal organs by using Canonical Correlation Analysis and dynamic MRI”. In: *The First International Workshop on Pulmonary Image Analysis*. 2008, S. 145–154.
- [Geneser u. a., 2011] Geneser, S., Hinkle, J., Kirby, R., Wang, B., Salter, B. und Joshi, S. “Quantifying variability in radiation dose due to respiratory-induced tumor motion”. *Med Image Anal* 15 (4), 2011, S. 640–649.
- [Giroux u. a., 2017] Giroux, M., Ladjal, H., Beuve, M. und Shariat, B. “Biomechanical Patient-Specific Model of the Respiratory System Based on 4D CT Scans and Controlled by Personalized Physiological Compliance”. In: *Med Image Comput Comput Assist Interv – MICCAI 2017*. Bd. 10434. LNCS. Springer, 2017, S. 216–223.

- [Glitzner u. a., 2015] Glitzner, M., Senneville, B. D. de, Lagendijk, J. J. W., Raaymakers, B. W. und Crijs, S. P. M. “On-line 3D motion estimation using low resolution MRI”. *Phys Med Biol* 60 (16), 2015, N301–10.
- [Glocker u. a., 2008] Glocker, B., Komodakis, N., Tziritas, G., Navab, N. und Paragios, N. “Dense image registration through MRFs and efficient linear programming”. *Med Image Anal* 12 (6), 2008. Special issue on information processing in medical imaging 2007, S. 731–741.
- [Glocker u. a., 2011] Glocker, B., Sotiras, A., Komodakis, N. und Paragios, N. “Deformable Medical Image Registration: Setting the State of the Art with Discrete Methods”. *Annu Rev Biomed Eng* 13 (1), 2011, S. 219–244.
- [Goksel u. a., 2013] Goksel, O., Sapchuk, K., Morris, W. und Salcudean, S. “Prostate Brachytherapy Training With Simulated Ultrasound and Fluoroscopy Images”. *IEEE Trans Biomed Eng* 60 (4), 2013, S. 1002–1012.
- [Gollmer, 2015] Gollmer, S. T. “3D-Bildsegmentierung mittels statistischer Formmodelle : Korrespondenzfindung, Modellierung, Segmentierung und ihre wechselseitigen Abhängigkeiten”. Dissertation. Universität zu Lübeck, 2015.
- [Grohmann u. a., 2015] Grohmann, C., Frenzel, T., Werner, R. und Cremers, F. “Design, performance characteristics and application examples of a new 4D motion platform”. *Z Med Phys* 25 (2), 2015, S. 156–167.
- [Guerrero u. a., 2005] Guerrero, T., Zhang, G., Segars, W., Huang, T.-C., Bilton, S., Ibbott, G., Dong, L., Forster, K. und Lin, K. P. “Elastic image mapping for 4-D dose estimation in thoracic radiotherapy”. *Radiat Prot Dosimetry* 115 (1-4), 2005, S. 497–502.
- [Guimond u. a., 2000] Guimond, A., Meunier, J. und Thirion, J.-P. “Average Brain Models: A Convergence Study”. *Comput Vis Image Underst* 77 (2), 2000, S. 192–210.
- [Ha u. a., 2019] Ha, I. Y.* , Wilms, M.* , Handels, H. und Heinrich, M. P. “Model-based Sparse-to-dense Image Registration for Realtime Respiratory Motion Estimation in Image-guided Interventions”. *IEEE Trans Biomed Eng* 66 (2), 2019, S. 302–310.
- [Haasdonk u. a., 2004] Haasdonk, B. und Bahlmann, C. “Learning with Distance Substitution Kernels”. In: *DAGM 2004: Pattern Recognition*. Bd. 3175. LNCS. Springer, 2004, S. 220–227.
- [Haar, 2016] Haar, G. ter. “HIFU Tissue Ablation: Concept and Devices”. In: *Therapeutic Ultrasound*. Springer, 2016, S. 3–20.
- [Handels, 2009] Handels, H. *Medizinische Bildverarbeitung: Bildanalyse, Mustererkennung und Visualisierung für die computergestützte ärztliche Diagnostik und Therapie*. 2. Auflage. Vieweg+Teubner, 2009.

- [Harris u. a., 2016] Harris, W., Ren, L., Cai, J., Zhang, Y., Chang, Z. und Yin, F.-F. “A Technique for Generating Volumetric Cine-Magnetic Resonance Imaging”. *Int J Radiat Oncol Biol Phys* 95 (2), 2016, S. 844–853.
- [Harris u. a., 1988] Harris, C. und Stephens, M. “A combined corner and edge detector.” In: *Alvey Vision Conference*. 1988, S. 147–151.
- [Hastie u. a., 2009] Hastie, T., Tibshirani, R. und Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2. Auflage. Springer, 2009.
- [He u. a., 2010] He, T., Xue, Z., Xie, W. und Wong, S. “Online 4-D CT Estimation for Patient-Specific Respiratory Motion Based on Real-Time Breathing Signals”. In: *Med Image Comput Comput Assist Interv – MICCAI 2010*. Bd. 6363. LNCS. Springer, 2010, S. 392–399.
- [Hecht u. a., 2013] Hecht, T., Mastmeyer, A., Fortmeier, D. und Handels, H. “4D-Planung von Nadelpfaden für Punktionseingriffe mit der Ray-Casting-Methode”. In: *Bildverarbeitung für die Medizin, BVM 2013*. Informatik aktuell. Springer, 2013, S. 9–14.
- [Heimann u. a., 2009] Heimann, T. und Meinzer, H.-P. “Statistical shape models for 3D medical image segmentation: A review”. *Med Image Anal* 13 (4), 2009, S. 543–563.
- [Heinrich u. a., 2012] Heinrich, M. P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F. V., Brady, S. M. und Schnabel, J. A. “MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration”. *Med Image Anal* 16 (7), 2012. Special Issue on the 2011 Conference on Medical Image Computing and Computer Assisted Intervention, S. 1423–1435.
- [Heinrich u. a., 2013a] Heinrich, M. P., Jenkinson, M., Brady, M. und Schnabel, J. A. “MRF-Based Deformable Registration and Ventilation Estimation of Lung CT”. *IEEE Trans Med Imag* 32 (7), 2013, S. 1239–1248.
- [Heinrich u. a., 2013b] Heinrich, M. P., Jenkinson, M., Papież, B. W., Brady, M. und Schnabel, J. A. “Towards realtime multimodal fusion for image-guided interventions using self-similarities”. In: *Med Image Comput Comput Assist Interv – MICCAI 2013*. Bd. 8149. LNCS. 2013, S. 187–194.
- [Heinrich u. a., 2014] Heinrich, M. P., Papież, B. W., Schnabel, J. A. und Handels, H. “Non-parametric Discrete Registration with Convex Optimisation”. In: *Workshop on Biomedical Image Registration – WBIR 2014*. Bd. 8545. LNCS. Springer, 2014, S. 51–61.
- [Heinrich u. a., 2015] Heinrich, M. P., Handels, H. und Simpson, I. J. A. “Estimating Large Lung Motion in COPD Patients by Symmetric Regularised Correspondence Fields”. In: *Med Image Comput Comput Assist Interv – MICCAI 2015*. Bd. 9350. LNCS. Springer International Publishing, 2015. Kap. Estimating Lar-

- ge Lung Motion in COPD Patients by Symmetric Regularised Correspondence Fields, S. 338–345.
- [Heinrich u. a., 2016] Heinrich, M. P., Simpson, I. J., Papież, B. W., Brady, S. M. und Schnabel, J. A. “Deformable image registration by combining uncertainty estimates from supervoxel belief propagation”. *Med Image Anal* 27 (1), 2016. Discrete Graphical Models in Biomedical Image Analysis, S. 57–71.
- [Henze u. a., 1990] Henze, N. und Zirkler, B. “A class of invariant consistent tests for multivariate normality”. *Commun Stat Theory Methods* 19 (10), 1990, S. 3595–3617.
- [Hernandez, 2008] Hernandez, M. “Variational techniques with applications to segmentation and registration of medical images”. PhD Thesis. University of Zaragoza, 2008.
- [Higham, 2002] Higham, N. J. “Computing the nearest correlation matrix – a problem from finance”. *IMA Journal of Numerical Analysis* 22 (3), 2002, S. 329–343.
- [Hinkle u. a., 2009] Hinkle, J., Fletcher, P. T., Wang, B., Salter, B. und Joshi, S. “4D MAP Image Reconstruction Incorporating Organ Motion”. In: *Inf Process Med Imaging – IPMI 2009*. Bd. 5636. LNCS. Springer, 2009, S. 676–687.
- [Hofmann u. a., 2008] Hofmann, T., Schölkopf, B. und Smola, A. J. “Kernel methods in machine learning”. *The Annals of Statistics* 36 (3), 2008, S. 1171–1220.
- [Hoisak u. a., 2004] Hoisak, J. D., Sixel, K. E., Tirona, R., Cheung, P. C. und Pignol, J.-P. “Correlation of lung tumor motion with external surrogate indicators of respiration”. *Int J Radiat Oncol Biol Phys* 60 (4), 2004, S. 1298–1306.
- [Hom u. a., 1991] Hom, R. A. und Johnson, C. R. *Topics in matrix analysis*. 1. Cambridge University Press, 1991.
- [Hou u. a., 2018] Hou, Q., Wang, J., Bai, R., Zhou, S. und Gong, Y. “Face alignment recurrent network”. *Pattern Recognit* 74 (2), 2018, S. 448–458.
- [Hu u. a., 2017] Hu, Y., Gibson, E., Vercauteren, T., Ahmed, H. U., Emberton, M., Moore, C. M., Noble, J. A. und Barratt, D. C. “Intraoperative Organ Motion Models with an Ensemble of Conditional Generative Adversarial Networks”. In: *Med Image Comput Comput Assist Interv – MICCAI 2017*. Bd. 10434. LNCS. Springer, 2017, S. 368–376.
- [Hugo u. a., 2017] Hugo, G. D., Weiss, E., Sleeman, W. C., Balik, S., Keall, P. J., Lu, J. und Williamson, J. F. “A longitudinal four-dimensional computed tomography and cone beam computed tomography dataset for image-guided radiation therapy research in lung cancer”. *Med Phys* 44 (2), 2017, S. 762–771.
- [ICRU, 1999] ICRU. *ICRU Report 62: Prescribing, Recording, and Reporting Photon Beam Therapy (Supplement to ICRU Report 50)*. International Commission on Radiation Units and Measurements (ICRU), 1999.

- [Ipsen u. a., 2016] Ipsen, S., Bruder, R., O'Brien, R., Keall, P. J., Schweikard, A. und Poulsen, P. R. "Online 4D ultrasound guidance for real-time motion compensation by MLC tracking". *Med Phys* 43 (10), 2016, S. 5695–5704.
- [IRCU, 2010] IRCU. *ICRU Report 83: Prescribing, Recording, and Reporting Photon-Beam Intensity-Modulated Radiation Therapy*. International Commission on Radiation Units and Measurements (ICRU), 2010.
- [Joshi u. a., 2004] Joshi, S., Davis, B., Jomier, M. und Gerig, G. "Unbiased diffeomorphic atlas construction for computational anatomy". *NeuroImage* 23, 2004, S151–S160.
- [Jud u. a., 2017] Jud, C., Giger, A., Sandkühler, R. und Cattin, P. C. "A Localized Statistical Motion Model as a Reproducing Kernel for Non-rigid Image Registration". In: *Med Image Comput Comput Assist Interv – MICCAI 2017*. Bd. 10434. LNCS. Springer, 2017, S. 261–269.
- [Kabus u. a., 2013] Kabus, S., Klinder, T., Murphy, K., Werner, R. und Sarrut, D. "Validation and Comparison of Approaches to Respiratory Motion Estimation". In: *4D Modeling and Estimation of Respiratory Motion for Radiation Therapy*. Springer, 2013, S. 159–183.
- [Keall u. a., 2006] Keall, P. J., Mageras, G. S., Balter, J. M., Emery, R. S., Forster, K. M., Jiang, S. B., Kapatoes, J. M., Low, D. A., Murphy, M. J., Murray, B. R., Ramsey, C. R., Van Herk, M. B., Vedam, S. S., Wong, J. W. und Yorke, E. "The management of respiratory motion in radiation oncology report of AAPM Task Group 76". *Med Phys* 33 (10), 2006, S. 3874–3900.
- [Keall u. a., 2014] Keall, P. J., Barton, M. und Crozier, S. "The Australian Magnetic Resonance Imaging–Linac Program". *Semin Radiat Oncol* 24 (3), 2014. Magnetic Resonance Imaging in Radiation Oncology, S. 203–206.
- [King u. a., 2012] King, A., Buerger, C., Tsoumpas, C., Marsden, P. und Schaeffter, T. "Thoracic respiratory motion estimation from MRI using a statistical model and a 2-D image navigator". *Med Image Anal* 16 (1), 2012, S. 252–264.
- [Kirschner u. a., 2011] Kirschner, M., Becker, M. und Wesarg, S. "3D Active Shape Model Segmentation with Nonlinear Shape Priors". In: *Med Image Comput Comput Assist Interv – MICCAI 2011*. Bd. 6892. LNCS. Springer, 2011, S. 492–499.
- [Klein u. a., 2009] Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M.-C., Christensen, G. E., Collins, D. L., Gee, J., Hellier, P., Song, J. H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R. P., Mann, J. J. und Parsey, R. V. "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration". *NeuroImage* 46 (3), 2009, S. 786–802.

- [Klinder u. a., 2009] Klinder, T., Lorenz, C. und Ostermann, J. “Free-Breathing intra- and intersubject respiratory motion capturing, modeling, and prediction”. In: *SPIE Medical Imaging 2009: Image Processing*. Bd. 7259. Proc. of SPIE. 2009, 72590T-1–11.
- [Klinder u. a., 2010] — . “Prediction framework for statistical respiratory motion modeling.” In: *Med Image Comput Comput Assist Interv – MICCAI 2010*. Bd. 6363. LNCS. Springer, 2010, S. 327–334.
- [Klinder u. a., 2012] Klinder, T. und Lorenz, C. “Respiratory motion compensation for image-guided bronchoscopy using a general motion model”. In: *IEEE International Symposium on Biomedical Imaging – ISBI 2012*. 2012, S. 960–963.
- [Koch u. a., 2004] Koch, N., Liu, H., Starkschall, G., Jacobson, M., Forster, K., Liao, Z., Komaki, R. und Stevens, C. W. “Evaluation of internal lung motion for respiratory-gated radiotherapy using MRI: Part I—correlating internal lung motion with skin fiducial motion”. *Int J Radiat Oncol Biol Phys* 60 (5), 2004, S. 1459–1472.
- [Koikkalainen u. a., 2008] Koikkalainen, J., Tölli, T., Lauerma, K., Antila, K., Mattila, E., Lilja, M. und Lötjönen, J. “Methods of Artificial Enlargement of the Training Set for Statistical Shape Models”. *IEEE Trans Med Imag* 27 (11), 2008, S. 1643–1654.
- [König u. a., 2014] König, L., Kipshagen, T. und Rühaak, J. “A non-linear image registration scheme for real-time liver ultrasound tracking using normalized gradient fields”. In: *MICCAI Challenge on Liver Ultrasound Tracking – CLUST 2014*. 2014, S. 29–36.
- [Korreman, 2015] Korreman, S. “Image-guided radiotherapy and motion management in lung cancer”. *Br J Radiol* 88 (1051), 2015, S. 20150100.
- [Kubo u. a., 1996] Kubo, H. D. und Hill, B. C. “Respiration gated radiotherapy treatment: a technical study”. *Phys Med Biol* 41 (1), 1996, S. 83–91.
- [Kurzion u. a., 1995] Kurzion, Y. und Yagel, R. “Space Deformation using Ray Deflectors”. In: *Rendering Techniques 1995*. Springer, 1995, S. 21–30.
- [Lagendijk u. a., 2014] Lagendijk, J. J. W., Raaymakers, B. W., Berg, C. A. T. V. den, Moerland, M. A., Philippens, M. E. und Vulpen, M. van. “MR guidance in radiotherapy”. *Phys Med Biol* 59 (21), 2014, R349–R369.
- [Lee u. a., 2011] Lee, D. und Krupa, A. “Intensity-based visual servoing for non-rigid motion compensation of soft tissue structures due to physiological motion using 4D ultrasound”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems – IROS 2011*. 2011, S. 2831–2836.
- [Lee, 2011] Lee, J. *Introduction to topological manifolds*. 2. Auflage. Springer, 2011.

- [Li u. a., 2006] Li, X. A., Stepaniak, C. und Gore, E. “Technical and dosimetric aspects of respiratory gating using a pressure-sensor motion monitoring system”. *Med Phys* 33 (1), 2006, S. 145–154.
- [Li u. a., 2011] Li, R., Lewis, J. H., Jia, X., Zhao, T., Liu, W., Wuenschel, S., Lamb, J., Yang, D., Low, D. A. und Jiang, S. B. “On a PCA-based lung motion model”. *Phys Med Biol* 56 (18), 2011, S. 6009–6030.
- [Lim u. a., 2018] Lim, L.-H., Wong, K. S.-W. und Ye, K. “Numerical Algorithms on the Affine Grassmannian”. *arXiv Preprint arXiv:1607.01833v2*, 2018.
- [Litjens u. a., 2017] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciampi, F., Ghafoorian, M., Laak, J. A. van der, Ginneken, B. van und Sánchez, C. I. “A survey on deep learning in medical image analysis”. *Med Image Anal* 42, 2017, S. 60–88.
- [Liu u. a., 2007] Liu, H. H., Balter, P., Tutt, T., Choi, B., Zhang, J., Wang, C., Chi, M., Luo, D., Pan, T., Hunjan, S., Starkschall, G., Rosen, I., Prado, K., Liao, Z., Chang, J., Komaki, R., Cox, J. D., Mohan, R. und Dong, L. “Assessing Respiration-Induced Tumor Motion and Internal Target Volume Using Four-Dimensional Computed Tomography for Radiotherapy of Lung Cancer”. *Int J Radiat Oncol Biol Phys* 68 (2), 2007, S. 531–540.
- [Liu u. a., 2010] Liu, X., Oguz, I., Pizer, S. M. und Mageras, G. S. “Shape-correlated deformation statistics for respiratory motion prediction in 4D lung”. In: *SPIE Medical Imaging 2010: Visualization, Image-Guided Procedures, and Modeling*. Bd. 7625. Proc. of SPIE. 2010, S. 76252D.
- [Lorenzi u. a., 2013] Lorenzi, M. und Pennec, X. “Geodesics, Parallel Transport & One-Parameter Subgroups for Diffeomorphic Image Registration”. *Int J Comput Vis* 105 (2), 2013, S. 111–127.
- [Lorenzi u. a., 2014] — . “Efficient parallel transport of deformations in time series of images: from Schild’s to pole ladder”. *J Math Imaging Vis* 50 (1-2), 2014, S. 5–17.
- [Low u. a., 2003] Low, D. A., Nystrom, M., Kalinin, E., Parikh, P., Dempsey, J. F., Bradley, J. D., Mutic, S., Wahab, S. H., Islam, T., Christensen, G., Politte, D. G. und Whiting, B. R. “A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing”. *Med Phys* 30 (6), 2003, S. 1254–1263.
- [Lowe, 2004] Lowe, D. G. “Distinctive image features from scale-invariant keypoints”. *Int J Comput Vis* 60 (2), 2004, S. 91–110.
- [Lu u. a., 2005a] Lu, W., Low, D. A., Parikh, P. J., Nystrom, M. M., El Naqa, I. M., Wahab, S. H., Handoko, M., Fooshee, D. und Bradley, J. D. “Comparison of spirometry and abdominal height as four-dimensional computed tomography metrics in lung”. *Med Phys* 32 (7), 2005, S. 2351–2357.

- [Lu u. a., 2005b] Lu, W., Parikh, P. J., Naqa, I. M. E., Nystrom, M. M., Hubenschmidt, J. P., Wahab, S. H., Mutic, S., Singh, A. K., Christensen, G. E., Bradley, J. D. und Low, D. A. “Quantitation of the reconstruction quality of a four-dimensional computed tomography process for lung cancer patients”. *Med Phys* 32 (4), 2005, S. 890–901.
- [Lüthi u. a., 2018] Lüthi, M., Gerig, T., Jud, C. und Vetter, T. “Gaussian Process Morphable Models”. *IEEE Trans Pattern Anal Mach Intell* 40 (8), 2018, S. 1860–1873.
- [Martin u. a., 2013] Martin, J., McClelland, J., Yip, C., Thomas, C., Hartill, C., Ahmad, S., O’Brien, R., Meir, I., Landau, D. und Hawkes, D. “Building motion models of lung tumours from cone-beam CT for radiotherapy applications”. *Phys Med Biol* 58 (6), 2013, S. 1809–1822.
- [Marx u. a., 2014] Marx, M., Ehrhardt, J., Werner, R., Schlemmer, H.-P. und Handels, H. “Simulation of spatiotemporal CT data sets using a 4D MRI-based lung motion model”. *Int J Comput Assist Radiol Surg* 9 (3), 2014, S. 401–409.
- [Mardia, 1970] Mardia, K. V. “Measures of multivariate skewness and kurtosis with applications”. *Biometrika* 57 (3), 1970, S. 519–530.
- [Mastmeyer u. a., 2018] Mastmeyer, A., Wilms, M. und Handels, H. “Population-based respiratory 4D motion atlas construction and its application for VR simulations of liver punctures”. In: *SPIE Medical Imaging 2018: Image Processing*. Bd. 10574. Proc SPIE. 2018, S. 1057417.
- [Matthews u. a., 2004] Matthews, I. und Baker, S. “Active appearance models revisited”. *Int J Comput Vis* 60 (2), 2004, S. 135–164.
- [McClelland u. a., 2011] McClelland, J. R., Hughes, S., Modat, M., Qureshi, A., Ahmad, S., Landau, D. B., Ourselin, S. und Hawkes, D. J. “Inter-fraction variations in respiratory motion models”. *Phys Med Biol* 56 (1), 2011, S. 251–272.
- [McClelland u. a., 2013] McClelland, J. R., Hawkes, D. J., Schaeffter, T. und King, A. P. “Respiratory motion models: A review”. *Med Image Anal* 17 (1), 2013, S. 19–42.
- [McClelland, 2008] McClelland, J. R. “Registration Based Respiratory Motion Models for use in Lung Radiotherapy”. Dissertation. University College London, 2008.
- [Menten u. a., 2017] Menten, M. J., Wetscherek, A. und Fast, M. F. “MRI-guided lung SBRT: Present and future developments”. *Phys Med* 44, 2017, S. 139–149.
- [Miao u. a., 2016] Miao, S., Wang, Z., Pan, L., Butler, J., Moran, G. und Liao, R. “Scatter to volume registration for model-free respiratory motion estimation from dynamic MRIs”. *Comput Med Imaging Graph* 52, 2016, S. 72–81.
- [Miller u. a., 2009] Miller, M. I. und Qiu, A. “The emerging discipline of computational functional anatomy”. *Neuroimage* 45 (1), 2009, S16–S39.

- [Milletari u. a., 2017] Milletari, F., Rothberg, A., Jia, J. und Sofka, M. “Integrating Statistical Prior Knowledge into Convolutional Neural Networks”. In: *Med Image Comput Comput Assist Interv – MICCAI 2017*. Bd. 10433. LNCS. Springer, 2017, S. 161–168.
- [Modersitzki, 2004] Modersitzki, J. *Numerical methods for image registration*. 1. Auflage. Oxford University Press, 2004.
- [Muller u. a., 2013] Muller, A., Petrusca, L., Auboiroux, V., Valette, P. J., Salomir, R. und Cotton, F. “Management of Respiratory Motion in Extracorporeal High-Intensity Focused Ultrasound Treatment in Upper Abdominal Organs: Current Status and Perspectives”. *Cardiovasc Intervent Radiol* 36 (6), 2013, S. 1464–1476.
- [Murphy u. a., 2011] Murphy, K. u. a. “Evaluation of Registration Methods on Thoracic CT: The EMPIRE10 Challenge”. *IEEE Trans Med Imag* 30, 2011, S. 1901–1920.
- [Mutic u. a., 2014] Mutic, S. und Dempsey, J. F. “The ViewRay System: Magnetic Resonance-Guided and Controlled Radiotherapy”. *Semin Radiat Oncol* 24 (3), 2014. Magnetic Resonance Imaging in Radiation Oncology, S. 196–199.
- [Nakamoto u. a., 2006] Nakamoto, M., Hirayama, H., Sato, Y., Konishi, K., Kakeji, Y., Hashizume, M. und Tamura, S. “Recovery of Liver Motion and Deformation Due to Respiration Using Laparoscopic Freehand 3D Ultrasound System”. In: *Med Image Comput Comput Assist Interv – MICCAI 2006*. Bd. 4191. LNCS. Springer, 2006, S. 372–379.
- [Ng u. a., 2002] Ng, A. Y., Jordan, M. I., Weiss, Y. u. a. “On spectral clustering: Analysis and an algorithm”. *Conference on Neural Information Processing Systems – NIPS 2002*, 2002, S. 849–856.
- [Ni u. a., 2011] Ni, D., Chan, W. Y., Qin, J., Chui, Y.-P., Qu, I., Ho, S. S. und Heng, P.-A. “A virtual reality simulator for ultrasound-guided biopsy training”. *IEEE Comput Graph Appl* 31 (2), 2011, S. 36–48.
- [Ortmüller u. a., 2014] Ortmüller, J., Wilms, M., Werner, R. und Handels, H. “Kombination von Atemsignalen zur Optimierung der Prädiktion komplexer atmungsbedingter Organ- und Tumorbewegungen”. In: *Bildverarbeitung für die Medizin 2014*. Informatik aktuell. Springer, 2014, S. 72–77.
- [Ortmüller u. a., 2015] Ortmüller, J., Gauer, T., Wilms, M., Handels, H. und Werner, R. “Respiratory Surface Motion Measurement by Microsoft Kinect: Implementation and Evaluation of a Clinical Setup”. *Current Directions in Biomedical Engineering* 1, 2015, S. 270–273.
- [O’Shea u. a., 2016a] O’Shea, T., Bamber, J., Fontanarosa, D., Meer, S. van der, Verhaegen, F. und Harris, E. “Review of ultrasound image guidance in exter-

- nal beam radiotherapy part II: intra-fraction motion management and novel applications”. *Phys Med Biol* 61 (8), 2016, R90–R137.
- [O’Shea u. a., 2016b] O’Shea, T., C., B. J. und J., H. E. “Temporal regularization of ultrasound-based liver motion estimation for image-guided radiation therapy”. *Medical Physics* 43 (1), 2016, S. 455–464.
- [Otani u. a., 2010] Otani, Y., Fukuda, I., Tsukamoto, N., Kumazaki, Y., Sekine, H., Imabayashi, E., Kawaguchi, O., Nose, T., Teshima, T. und Dokiya, T. “A comparison of the respiratory signals acquired by different respiratory monitoring systems used in respiratory gated radiotherapy”. *Med Phys* 37 (12), 2010, S. 6178–6186.
- [Ozhasoglu u. a., 2008] Ozhasoglu, C., Saw, C. B., Chen, H., Burton, S., Komanduri, K., Yue, N. J., Huq, S. M. und Heron, D. E. “Synchrony – Cyberknife Respiratory Compensation Technology”. *Med Dosim* 33 (2), 2008. Image-Guided Radiation Therapy: S. 117–123.
- [Paganelli u. a., 2015] Paganelli, C., Seregini, M., Fattori, G., Summers, P., Bellomi, M., Baroni, G. und Riboldi, M. “Magnetic Resonance Imaging–Guided versus Surrogate-Based Motion Tracking in Liver Radiation Therapy: A Prospective Comparative Study”. *Int J Radiat Oncol Biol Phys* 91 (4), 2015, S. 840–848.
- [Pan u. a., 2004] Pan, T., Lee, T.-Y., Rietzel, E. und Chen, G. T. Y. “4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT”. *Med Phys* 31 (2), 2004, S. 333–340.
- [Park u. a., 2018] Park, S., Farah, R., Shea, S. M., Tryggestad, E., Hales, R. und Lee, J. “Simultaneous tumor and surrogate motion tracking with dynamic MRI for radiation therapy planning”. *Phys Med Biol* 63 (2), 2018, S. 025015.
- [Pennec u. a., 2015] Pennec, X. und Fillard, P. “Statistical Computing on Non-Linear Spaces for Computational Anatomy”. In: *Handbook of Biomedical Imaging: Methodologies and Clinical Research*. Springer, 2015, S. 147–168.
- [Peressutti u. a., 2013a] Peressutti, D., Penney, G. P., Kolbitsch, C. und King, A. P. “Personalising cross-population respiratory motion models using anatomical features”. In: *Conference on Medical Image Understanding and Analysis – MIUA 2013*. 2013, S. 45–50.
- [Peressutti u. a., 2013b] Peressutti, D., Penney, G. P., Housden, R. J., Kolbitsch, C., Gomez, A., Rijkhorst, E.-J., Barratt, D. C., Rhode, K. S. und King, A. P. “A novel Bayesian respiratory motion model to estimate and resolve uncertainty in image-guided cardiac interventions”. *Med Image Anal* 17 (4), 2013, S. 488–502.
- [Peressutti u. a., 2014] Peressutti, D., Penney, G. P., Kolbitsch, C. und King, A. P. “Personalising population-based respiratory motion models of the heart using neighbourhood approximation based on learnt anatomical features”. *Med Image Anal* 18 (7), 2014, S. 1015–1025.

- [Piegl u. a., 2000] Piegl, L. und Tiller, W. “Least-Squares B-Spline Curve Approximation with Arbitrary End Derivatives”. *Engineering with Computers* 16 (2), 2000, S. 109–116.
- [Piegl u. a., 2005] Piegl, L. und Tiller, W. *The NURBS Book*. 1. Auflage. Springer, 2005.
- [Poels u. a., 2015] Poels, K., Dhont, J., Verellen, D., Blanck, O., Ernst, F., Vandemeulebroucke, J., Depuydt, T., Storme, G. und Ridder, M. D. “A comparison of two clinical correlation models used for real-time tumor tracking of semi-periodic motion: A focus on geometrical accuracy in lung and liver cancer patients”. *Radiother Oncol* 115 (3), 2015, S. 419–424.
- [Polzin u. a., 2016] Polzin, T., Niethammer, M., Heinrich, M. P., Handels, H. und Mordersitzki, J. “Memory Efficient LDDMM for Lung CT”. In: *Med Image Comput Comput Assist Interv – MICCAI 2016*. Bd. 9902. LNCS. Springer, 2016, S. 28–36.
- [Preim u. a., 2014] Preim, B. und Botha, C. “Visual Computing for Medicine”. In: *Visual Computing for Medicine*. 2. Auflage. Morgan Kaufmann, 2014. Kap. Computer-Assisted Medical Education, e101–e169.
- [Preiswerk u. a., 2014] Preiswerk, F., Luca, V. D., Arnold, P., Celicanin, Z., Petrusca, L., Tanner, C., Bieri, O., Salomir, R. und Cattin, P. C. “Model-guided respiratory organ motion prediction of the liver from 2D ultrasound”. *Med Image Anal* 18 (5), 2014, S. 740–751.
- [Quirk u. a., 2012] Quirk, S., Becker, N. und Smith, W. “External respiratory motion: Shape analysis and custom realistic respiratory trace generation”. *Med Phys* 39 (8), 2012, S. 4999–5003.
- [Rey u. a., 2002] Rey, D., Subsol, G., Delingette, H. und Ayache, N. “Automatic detection and segmentation of evolving processes in 3D medical images: Application to multiple sclerosis”. *Med Image Anal* 6 (2), 2002, S. 163–179.
- [Rijkhorst u. a., 2011] Rijkhorst, E.-J., Rivens, I., Haar, G. ter, Hawkes, D. und Barratt, D. “Effects of Respiratory Liver Motion on Heating for Gated and Model-Based Motion-Compensated High-Intensity Focused Ultrasound Ablation”. In: *Med Image Comput Comput Assist Interv – MICCAI 2011*. Bd. 6891. LNCS. Springer, 2011, S. 605–612.
- [Risser u. a., 2013] Risser, L., Vialard, F.-X., Baluwala, H. Y. und Schnabel, J. A. “Piecewise-diffeomorphic image registration: Application to the motion estimation between 3D CT lung images with sliding conditions”. *Med Image Anal* 17 (2), 2013, S. 182–193.
- [Rosipal u. a., 2006] Rosipal, R. und Krämer, N. “Overview and Recent Advances in Partial Least Squares”. In: *Subspace, Latent Structure and Feature Selection*. Bd. 3940. LNCS. Springer, 2006, S. 34–51.

- [Royer u. a., 2017] Royer, L., Krupa, A., Dardenne, G., Bras, A. L., Marchand, E. und Marchal, M. “Real-time target tracking of soft tissues in 3D ultrasound images based on robust visual information and mechanical simulation”. *Med Image Anal* 35 (1), 2017, S. 582–598.
- [Ruan u. a., 2008] Ruan, D., Fessler, J. A., Balter, J. M., Berbeco, R. I., Nishioka, S. und Shirato, H. “Inference of hysteretic respiratory tumor motion from external surrogates: a state augmentation approach”. *Phys Med Biol* 53 (11), 2008, S. 2923–2936.
- [Ruben u. a., 2016] Ruben, J. D., Seeley, A., Panettieri, V. und Ackerly, T. “Variation in Lung Tumour Breathing Motion between Planning Four-dimensional Computed Tomography and Stereotactic Ablative Radiotherapy Delivery and its Dosimetric Implications: Any Role for Four-dimensional Set-up Verification?” *Clin Oncol (R Coll Radiol)* 28 (1), 2016, S. 21–7.
- [Rühaak u. a., 2017] Rühaak, J., Polzin, T., Heldmann, S., Simpson, I., Handels, H., Modersitzki, J. und Heinrich, M. P. “Estimation of Large Motion in Lung CT by Integrating Regularized Keypoint Correspondences into Dense Deformable Registration”. *IEEE Trans Med Imag* 36 (8), 2017, S. 1746–1757.
- [Ruthenbeck u. a., 2015] Ruthenbeck, G. S. und Reynolds, K. J. “Virtual reality for medical training: the state-of-the-art”. *Journal of Simulation* 9 (1), 2015, S. 16–26.
- [Sahgal u. a., 2012] Sahgal, A., Roberge, D., Schellenberg, D., Purdie, T., Swaminath, A., Pantarotto, J., Filion, E., Gabos, Z., Butler, J., Letourneau, D., Masucci, G., Mulroy, L., Bezjak, A., Dawson, L. und Parliament, M. “The Canadian Association of Radiation Oncology Scope of Practice Guidelines for Lung, Liver and Spine Stereotactic Body Radiotherapy”. *Clin Oncol (R Coll Radiol)* 24 (9), 2012, S. 629–639.
- [Samei u. a., 2012] Samei, G., Tanner, C. und Székely, G. “Predicting liver motion using exemplar models”. In: *Abdominal Imaging. Computational and Clinical Applications – 4th International Workshop at MICCAI 2012*. Bd. 7601. LNCS. Springer, 2012, S. 147–157.
- [Samei u. a., 2013] Samei, G., Chlebus, G., Székely, G. und Tanner, C. “Adaptive Confidence Regions of Motion Predictions from Population Exemplar Models”. In: *Abdominal Imaging. Computation and Clinical Applications – MICCAI 2013*. Bd. 8198. LNCS. Springer, 2013, S. 231–240.
- [Savill u. a., 2011] Savill, F., Schaeffter, T. und King, A. P. “Assessment of input signal positioning for cardiac respiratory motion models during different breathing patterns”. In: *IEEE International Symposium on Biomedical Imaging – ISBI 2011*. 2011, S. 1698–1701.

- [Sayeh u. a., 2007] Sayeh, S., Wang, J., Main, W. T., Kilby, W. und Maurer, C. R. “Respiratory Motion Tracking for Robotic Radiosurgery”. In: *Treating Tumors that Move with Respiration*. Springer, 2007, S. 15–29.
- [Schweikard u. a., 2000] Schweikard, A., Glosser, G., Bodduluri, M., Murphy, M. J. und Adler, J. R. “Robotic motion compensation for respiratory movement during radiosurgery”. *Comput Aided Surg* 5 (4), 2000, S. 263–277.
- [Schaller u. a., 2008] Schaller, C., Penne, J. und Hornegger, J. “Time-of-flight sensor for respiratory motion gating.” *Med Phys* 35 (7), 2008, S. 3090–3093.
- [Schmidt-Richberg u. a., 2012] Schmidt-Richberg, A., Werner, R., Handels, H. und Ehrhardt, J. “Estimation of Slipping Organ Motion by Registration with Direction-Dependent Regularization”. *Med Image Anal* 16 (1), 2012, S. 150–159.
- [Schmidt-Richberg u. a., 2014] Schmidt-Richberg, A., Werner, R., Handels, H. und Ehrhardt, J. “A Flexible Variational Registration Framework”. *Insight Journal*, 2014.
- [Schmidt-Richberg, 2013] Schmidt-Richberg, A. “Integration of morphological and physiological knowledge into registration methods for pulmonary image analysis”. Dissertation. Universität zu Lübeck, 2013.
- [Segars u. a., 2010] Segars, W. P., Sturgeon, G., Mendonca, S., Grimes, J. und Tsui, B. M. W. “4D XCAT phantom for multimodality imaging research”. *Med Phys* 37 (9), 2010, S. 4902–4915.
- [Seppenwoolde u. a., 2002] Seppenwoolde, Y., Shirato, H., Kitamura, K., Shimizu, S., Herk, M. van, Lebesque, J. V. und Miyasaka, K. “Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy.” *Int J Radiat Oncol Biol Phys* 53 (4), 2002, S. 822–834.
- [Seregini u. a., 2013] Seregini, M., Pella, A., Riboldi, M., Orecchia, R., Cerveri, P. und Baroni, G. “Real-time tumor tracking with an artificial neural networks-based method: A feasibility study”. *Phys Med* 29 (1), 2013, S. 48–59.
- [Seregini u. a., 2018] Seregini, M., Paganelli, C., Summers, P., Bellomi, M., Baroni, G. und Riboldi, M. “A hybrid image registration and matching framework for real-time motion tracking in MRI-guided radiotherapy”. *IEEE Trans Biomed Eng* 65 (1), 2018, S. 131–139.
- [Shawe-Taylor u. a., 2004] Shawe-Taylor, J. und Cristianini, N. *Kernel methods for pattern analysis*. 1. Auflage. Cambridge University Press, 2004.
- [Shepard u. a., 2017] Shepard, A. J., Wang, B., Foo, T. K. und Bednarz, B. P. “A Block Matching Based Approach with Multiple Simultaneous Templates for the Real-Time 2D Ultrasound Tracking of Liver Vessels”. *Med Phys* 44 (11), 2017, S. 5889–5900.

- [Siebenthal u. a., 2007] Siebenthal, M. von, Székely, G., Lomax, A. J. und Cattin, P. C. “Systematic errors in respiratory gating due to intrafraction deformations of the liver”. *Med Phys* 34 (9), 2007, S. 3620–3629.
- [Singh u. a., 2015] Singh, N., Vialard, F.-X. und Niethammer, M. “Splines for diffeomorphisms”. *Med Image Anal* 25 (1), 2015, S. 56–71.
- [Somphone u. a., 2014] Somphone, O., Allaire, S., Mory, B. und Dufour, C. “Live feature tracking in ultrasound liver sequences with sparse demons”. In: *MICCAI Challenge on Liver Ultrasound Tracking – CLUST 2014*. 2014, S. 53–60.
- [Sonke u. a., 2008] Sonke, J.-J., Lebesque, J. und Herk, M. van. “Variability of Four-Dimensional Computed Tomography Patient Models”. *Int J Radiat Oncol Biol Phys* 70 (2), 2008, S. 590–598.
- [Sotiras u. a., 2013] Sotiras, A., Davatzikos, C. und Paragios, N. “Deformable Medical Image Registration: A Survey”. *IEEE Trans Med Imag* 32 (7), 2013, S. 1153–1190.
- [Sothmann u. a., 2017a] Sothmann, T., Gauer, T., Wilms, M. und Werner, R. “Correspondence model-based 4D VMAT dose simulation for analysis of local metastasis recurrence after extracranial SBRT”. *Phys Med Biol* 62 (23), 2017, S. 9001–9017.
- [Sothmann u. a., 2017b] Sothmann, T., Gauer, T. und Werner, R. “4D dose simulation in volumetric arc therapy: Accuracy and affecting parameters”. *PLoS one* 12 (2), 2017, e0172810.
- [Steinbrücker u. a., 2009] Steinbrücker, F., Pock, T. und Cremers, D. “Large displacement optical flow computation without warping”. In: *International Conference on Computer Vision – ICCV 2009*. 2009, S. 1609–1614.
- [Stemkens u. a., 2016] Stemkens, B., Tijssen, R., Senneville, B. de, Lagendijk, J. und Berg, C. van den. “Image-driven, model-based 3D abdominal motion estimation for MR-guided radiotherapy”. *Phys Med Biol* 61 (14), 2016, S. 5335.
- [Sundaram u. a., 2004] Sundaram, T., Avants, B. und Gee, J. “A Dynamic Model of Average Lung Deformation Using Capacity-Based Reparameterization and Shape Averaging of Lung MR Images”. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004*. Bd. 3217. LNCS. Springer, 2004, S. 1000–1007.
- [Tanner u. a., 2012] Tanner, C., Boye, D., Samei, G. und Székely, G. “Review on 4D models for organ motion compensation”. *Crit Rev Biomed Eng* 40 (2), 2012, S. 135–154.
- [Tanner u. a., 2015] Tanner, C., Samei, G. und Székely, G. “Robust exemplar models of respiratory liver motion and individualization using an additional breath-

- hold image”. In: *IEEE International Symposium on Biomedical Imaging – ISBI 2015*. 2015, S. 1576–1579.
- [Thilmann u. a., 2013] Thilmann, C., Oelfke, U. und Sterzing, F. “Intensitätsmodulierte Strahlentherapie”. In: *Strahlentherapie*. 2. Auflage. Springer, 2013, S. 271–286.
- [Turaga u. a., 2008] Turaga, P., Veeraraghavan, A. und Chellappa, R. “Statistical analysis on Stiefel and Grassmann manifolds with applications in computer vision”. In: *IEEE Conference on Computer Vision and Pattern Recognition – CVPR 2008*. 2008, S. 1–8.
- [Tzimiropoulos u. a., 2014] Tzimiropoulos, G. und Pantic, M. “Gauss-newton deformable part models for face alignment in-the-wild”. In: *IEEE Conference on Computer Vision and Pattern Recognition – CVPR 2014*. 2014, S. 1851–1858.
- [Tzimiropoulos u. a., 2017] —. “Fast Algorithms for Fitting Active Appearance Models to Unconstrained Images”. *Int J Comput Vis* 122 (1), 2017, S. 17–33.
- [Ullrich u. a., 2012] Ullrich, S. und Kuhlen, T. “Haptic Palpation for Medical Simulation in Virtual Environments”. *IEEE Trans Vis Comput Graph* 18 (4), 2012, S. 617–625.
- [Uzunova u. a., 2017] Uzunova, H. *, Wilms, M. *, Handels, H. und Ehrhardt, J. “Training CNNs for Image Registration from Few Samples with Model-based Data Augmentation”. In: *Med Image Comput Comput Assist Interv – MICCAI 2017*. Bd. 10433. LNCS. Springer, 2017, S. 223–231.
- [Vandemeulebroucke u. a., 2011] Vandemeulebroucke, J., Rit, S., Kybic, J., Clarysse, P. und Sarrut, D. “Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs”. *Med Phys* 38 (1), 2011, S. 166–178.
- [Velec u. a., 2011] Velec, M., Moseley, J. L., Eccles, C. L., Craig, T., Sharpe, M. B., Dawson, L. A. und Brock, K. K. “Effect of Breathing Motion on Radiotherapy Dose Accumulation in the Abdomen Using Deformable Registration”. *Int J Radiat Oncol Biol Phys* 80 (1), 2011, S. 265–272.
- [Vercauteren u. a., 2009] Vercauteren, T., Pennec, X., Perchant, A. und Ayache, N. “Diffeomorphic demons: Efficient non-parametric image registration”. *NeuroImage* 45 (1, Supplement 1), 2009, S61–S72.
- [Villard u. a., 2009] Villard, P. F., Vidal, F. P., Hunt, C., Bello, F., John, N. W., Johnson, S. und Gould, D. A. “A prototype percutaneous transhepatic cholangiography training simulator with real-time breathing motion”. *Int J Comput Assist Radiol Surg* 4 (6), 2009, S. 571–578.
- [Villard u. a., 2012] Villard, P. F., Vidal, F., Bello, F. und John, N. “A method to compute respiration parameters for patient-based simulators”. In: *MMVR-Medicine Meets Virtual Reality 19*. 2012, S. 529–533.

- [Villard u. a., 2014] Villard, P. F., Vidal, F. P., Cenydd, L., Holbrey, R., Pisharody, S., Johnson, S., Bulpitt, A., John, N. W., Bello, F. und Gould, D. “Interventional radiology virtual simulator for liver biopsy”. English. *Int J Comput Assist Radiol Surg* 9 (2), 2014, S. 255–267.
- [Wang u. a., 2000] Wang, Y. und Staib, L. H. “Boundary finding with prior shape and smoothness models”. *IEEE Trans Pattern Anal Mach Intell* 22 (7), 2000, S. 738–743.
- [Wannenmacher u. a., 2013] Wannenmacher, M., Debus, J., Wenz, F. und Bahnsen, J. “Allgemeine Grundlagen”. In: *Strahlentherapie*. 2. Auflage. Springer, 2013, S. 3–10.
- [Wang u. a., 2016] Wang, S., Zhang, Z. und Zhang, T. “Towards more efficient SPSP matrix approximation and CUR matrix decomposition”. *J Mach Learn Res* 17 (210), 2016, S. 1–49.
- [Wasza u. a., 2016] Wasza, J., Fischer, P., Leutheuser, H., Oefner, T., Bert, C., Maier, A. und Hornegger, J. “Real-Time Respiratory Motion Analysis Using 4-D Shape Priors”. *IEEE Trans Biomed Eng* 63 (3), 2016, S. 485–495.
- [Wentz u. a., 2012] Wentz, T., Fayad, H., Bert, J., Pradier, O., Clément, J.-F., Vourch, S., BouSSION, N. und Visvikis, D. “Accuracy of dynamic patient surface monitoring using a time-of-flight camera and B-spline modeling for respiratory motion characterization”. *Phys Med Biol* 57 (13), 2012, S. 4175–93.
- [Werner u. a., 2009] Werner, R., Ehrhardt, J., Schmidt, R. und Handels, H. “Patient-specific finite element modeling of respiratory lung motion using 4D CT image data”. *Med Phys* 36 (5), 2009, S. 1500–1511.
- [Werner u. a., 2012a] Werner, R., Ehrhardt, J., Schmidt-Richberg, A., Albers, D., Frenzel, T., Petersen, C., Cremers, F. und Handels, H. “Towards Accurate Dose Accumulation for Step-&Shoot IMRT: Impact of Weighting Schemes and Temporal Image Resolution on the Estimation of Dosimetric Motion Effects”. *Z Med Phys* 22 (2), 2012, S. 109–122.
- [Werner u. a., 2012b] Werner, R., Ehrhardt, J., Schmidt-Richberg, A., Wilms, M., Blendowski, M. und Handels, H. “A Diffeomorphic MLR Framework for Surrogate-based Motion Estimation in Radiation Therapy: Concept and First Evaluation”. In: *Proc. Informatik 2012*. 2012, S. 1774–1785.
- [Werner u. a., 2012c] Werner, R., Wilms, M., Ehrhardt, J., Schmidt-Richberg, A., Blendowski, M. und Handels, H. “A Diffeomorphic MLR Framework for Surrogate-based Motion Estimation and Situation-adapted Dose Accumulation”. In: *Image-Guidance and Multimodal Dose Planning in Radiation Therapy: a MICCAI Workshop*. 2012, S. 42–49.
- [Werner u. a., 2014] Werner, R., Schmidt-Richberg, A., Handels, H. und Ehrhardt, J. “Estimation of Lung Motion Fields in 4D CT Data by Variational Non-Linear

- Intensity-Based Registration: A Comparison and Evaluation Study”. *Phys Med Biol* 59 (15), 2014, S. 4247–4260.
- [Werner, 2012] Werner, R. “Bewegungsfeldschätzung und Dosisakkumulation anhand von 4D-Bilddaten für die Strahlentherapie atmbewegter Tumoren”. Dissertation. Universität zu Lübeck, 2012.
- [Western u. a., 2015] Western, C., Hristov, D. und Schlosser, J. “Ultrasound imaging in radiation therapy: From interfractional to intrafractional guidance”. *Cureus* 7 (6), 2015, e280.
- [Wilms u. a., 2011] Wilms, M., Ehrhardt, J. und Handels, H. “Automatische Segmentierung der Lungenflügel in CT-Daten”. In: *Bildverarbeitung für die Medizin 2011*. Informatik aktuell. Springer, 2011, S. 119–123.
- [Wilms u. a., 2012] — . “A 4D Statistical Shape Model for Automated Segmentation of Lungs with Large Tumors”. In: *Med Image Comput Comput Assist Interv – MICCAI 2012*. Bd. 7511. LNCS. Springer, 2012, S. 347–354.
- [Wilms u. a., 2013] Wilms, M., Werner, R., Ehrhardt, J., Schmidt-Richberg, A., Blendowski, M. und Handels, H. “Surrogate-based Diffeomorphic Motion Estimation for Radiation Therapy: Comparison of Multivariate Regression Approaches”. In: *SPIE Medical Imaging 2013: Image Processing*. Bd. 8669. Proc SPIE. 2013, S. 866915.
- [Wilms u. a., 2014a] Wilms, M.* , Werner, R.* , Ehrhardt, J., Schmidt-Richberg, A., Schlemmer, H.-P. und Handels, H. “Multivariate regression approaches for surrogate-based diffeomorphic estimation of respiratory motion in radiation therapy”. *Phys Med Biol* 59 (5), 2014, S. 1147–1164.
- [Wilms u. a., 2014b] Wilms, M., Ehrhardt, J., Werner, R., Marx, M. und Handels, H. “Statistical Analysis of Surrogate Signals to Incorporate Respiratory Motion Variability into Radiotherapy Treatment Planning”. In: *SPIE Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and Modeling*. Bd. 9036. Proc SPIE. 2014, 90360J.
- [Wilms u. a., 2014c] Wilms, M., Werner, R., Blendowski, M., Ortmüller, J., Handels, H. u. a. “Simulation of Range Imaging-based Estimation of Respiratory Lung Motion”. *Methods Inform Med* 53 (4), 2014, S. 257–263.
- [Wilms u. a., 2015a] Wilms, M., Fortmeier, D., Mastmeyer, A. und Handels, H. “Modellbasierte Simulation der Atembewegung für das Virtual-Reality-Training von Punktionseingriffen”. In: *Bildverarbeitung für die Medizin 2015*. Informatik aktuell. Springer, 2015, S. 317–322.
- [Wilms u. a., 2015b] Wilms, M., Werner, R., Yamamoto, T., Handels, H. und Ehrhardt, J. “Population-based correspondence models for respiratory motion estimation in the presence of inter-fraction motion variations”. In: *MICCAI Work-*

- shop on Imaging and Computer Assistance in Radiation Therapy*. 2015, S. 81–88.
- [Wilms u. a., 2016] Wilms, M., Ha, I. Y., Handels, H. und Heinrich, M. P. “Model-based Regularisation for Respiratory Motion Estimation with Sparse Features in Image-guided Interventions”. In: *Med Image Comput Comput Assist Interv – MICCAI 2016*. Bd. 9902. LNCS. Springer, 2016, S. 89–97.
- [Wilms u. a., 2017a] Wilms, M., Handels, H. und Ehrhardt, J. “Multi-resolution multi-object statistical shape models based on the locality assumption”. *Med Image Anal* 38 (5), 2017, S. 17–29.
- [Wilms u. a., 2017b] — . “Representative Patch-based Active Appearance Models Generated from Small Training Populations”. In: *Med Image Comput Comput Assist Interv – MICCAI 2017*. Bd. 10433. LNCS. Springer, 2017, S. 152–160.
- [Wilms u. a., 2017c] Wilms, M.* , Werner, R.* , Yamamoto, T., Handels, H. und Ehrhardt, J. “Subpopulation-based correspondence modelling for improved respiratory motion estimation in the presence of inter-fraction motion variations”. *Phys Med Biol* 62 (14), 2017, S. 5823–5839.
- [Wu u. a., 2007] Wu, X., Thongphiew, D., Wang, Z., Chankong, V. und Yin, F. “Real-Time Modeling of Vascular Flow for Angiography Simulation”. In: *Med Image Comput Comput Assist Interv – MICCAI 2007*. Bd. 4791. LNCS. Springer, 2007, S. 557–565.
- [Wu u. a., 2008] Wu, Q., Allard, J. und Cotin, S. “The impact of respiratory motion and treatment technique on stereotactic body radiation therapy for liver cancer”. *Med Phys* 35 (4), 2008, S. 1440–1451.
- [Xue u. a., 2006] Xue, Z., Shen, D. und Davatzikos, C. “Statistical representation of high-dimensional deformation fields with application to statistically constrained 3D warping”. *Med Image Anal* 10 (5), 2006, S. 740–751.
- [Yamamoto u. a., 2012] Yamamoto, T., Kabus, S., Berg, J. von, Lorenz, C., Chung, M. P., Hong, J. C., Loo Jr, B. W. und Keall, P. J. “Reproducibility of four-dimensional computed tomography-based lung ventilation imaging”. *Acad Radiol* 19 (12), 2012, S. 1554–65.
- [Yang u. a., 2017] Yang, X., Kwitt, R., Styner, M. und Niethammer, M. “Quicksilver: Fast predictive image registration – A deep learning approach”. *NeuroImage* 158 (9), 2017, S. 378–396.
- [Ye u. a., 2016] Ye, K. und Lim, L.-H. “Schubert Varieties and Distances between Subspaces of Different Dimensions”. *SIAM J Matrix Anal Appl* 37 (3), 2016, S. 1176–1197.

- [Younes u. a., 2009] Younes, L., Arrate, F. und Miller, M. I. “Evolutions equations in computational anatomy”. *NeuroImage* 45 (1, Supplement 1), 2009. Mathematics in Brain Imaging, S40–S50.
- [Yu u. a., 2016] Yu, X., Zhou, F. und Chandraker, M. “Deep Deformation Network for Object Landmark Localization”. In: *European Conference on Computer Vision – ECCV 2016*. LNCS. Springer, 2016, S. 52–70.
- [Yuan u. a., 2016] Yuan, A., Wei, J., Gaebler, C. P., Huang, H., Olek, D. und Li, G. “A Novel Respiratory Motion Perturbation Model Adaptable to Patient Breathing Irregularities”. *Int J Radiat Oncol Biol Phys* 96 (5), 2016, S. 1087–1096.
- [Zachiu u. a., 2015] Zachiu, C., Papadakis, N., Ries, M., Moonen, C. und Senneville, B. de. “An improved optical flow tracking technique for real-time MR-guided beam therapies in moving organs”. *Phys Med Biol* 60 (23), 2015, S. 9003.
- [Zachiu u. a., 2017] Zachiu, C., Ries, M., Ramaekers, P., Guey, J.-L., Moonen, C. T. W. und Senneville, B. D. de. “Real-time non-rigid target tracking for ultrasound-guided clinical interventions”. *Phys Med Biol* 62 (20), 2017, S. 8154–8177.
- [Zafeiriou u. a., 2015] Zafeiriou, S., Zhang, C. und Zhang, Z. “A survey on face detection in the wild: Past, present and future”. *Comput Vis Image Underst* 138 (9), 2015, S. 1–24.
- [Zaidi u. a., 2009] Zaidi, H. und Tsui, B. M. W. “Review of Computational Anthropomorphic Anatomical and Physiological Models”. *Proc IEEE* 97 (12), 2009, S. 1938–1953.
- [Zhang u. a., 2003] Zhang, T., Keller, H., O’Brien, M. J., Mackie, T. R. und Paliwal, B. “Application of the spirometer in respiratory gated radiotherapy”. *Med Phys* 30 (12), 2003, S. 3165–3171.
- [Zhang u. a., 2007] Zhang, Q., Pevsner, A., Hertanto, A., Hu, Y.-C., Rosenzweig, K. E., Ling, C. C. und Mageras, G. S. “A patient-specific respiratory model of anatomical motion for radiation treatment planning”. *Med Phys* 34 (12), 2007, S. 4772–4781.

