

INAUGURAL - DISSERTATION  
zur  
Erlangung der Doktorwürde  
der  
Naturwissenschaftlich-Mathematischen  
Gesamtfakultät  
der Ruprecht-Karls-Universität  
Heidelberg

vorgelegt von  
Dipl.-Phys. Kristina Giske  
aus Almaty  
Tag der mündlichen Prüfung: 19.01.2011



GPU basierte Parallelisierung von  
deformierbaren Registrierungsverfahren als  
Grundlage für dynamische Anpassungen von  
Therapieplänen in der adaptiven  
Strahlentherapie

Gutachter: Prof. Dr. Uwe Oelfke  
Prof. Dr. Rolf Bendl



## **GPU basierte Parallelisierung von deformierbaren Registrierungsverfahren als Grundlage für dynamische Anpassungen von Therapieplänen in der adaptiven Strahlentherapie**

**Zusammenfassung** - Hochkonforme Strahlentherapietechniken ermöglichen eine gezielte Tumorbestrahlung bei maximaler Schonung des benachbarten gesunden Gewebes und der Risikoorgane. Anatomische Veränderungen während des Bestrahlungsverlaufs können jedoch zu einer Abweichung von der verschriebenen Dosisverteilung führen. Durch Adaptionen der Patientenposition oder des Bestrahlungsplans können diese Abweichungen kompensiert werden. Dafür werden Transformationen benötigt, die diese Veränderungen beschreiben. Die sich verändernde Anatomie wird unter der Therapie in Kontrollaufnahmen abgebildet. Um die Transformationen aus den Aufnahmen zu ermitteln, können Bildregistrierungsverfahren eingesetzt werden. Zur Berücksichtigung von lokalen Veränderungen der Anatomie werden deformierbare Registrierungsverfahren benötigt.

Um während der Therapie auf beobachtete Veränderungen eingehen zu können, steht nur ein sehr enges Zeitfenster zur Verfügung. Existierende Registrierungsverfahren sind zu langsam, um diese Aufgabe erfüllen zu können. Durch Parallelisierung der algorithmischen Komponenten lassen sich die Registrierungsverfahren signifikant beschleunigen, ohne die Genauigkeit zu reduzieren.

In dieser Arbeit wurde ein hoch parallelisierbares Verfahren zur Ermittlung der Deformationen aus Bildern entwickelt und auf einer parallelen Architektur der GPU Hardware umgesetzt. Die resultierende Rechendauer von unter einer bis wenigen Minuten bei guter Registrierungsqualität kann die Umsetzung von adaptiven Korrekturstrategien noch während der Behandlung ermöglichen.

## **GPU based parallelisation of deformable image registration methods as foundation for dynamic adaptations of treatment plans in adaptive radiation therapy**

**Abstract** - Highly conformal radiation therapy techniques enable a targeted tumour irradiation with minimal impact on nearby healthy tissue and organs at risk. Changes in anatomy of the patient during the treatment course, which might lead to deviation from the previously prescribed dose distribution, can be compensated by adapting the position of the patient or the plan itself. For such adaptations the transformations describing those anatomical changes are needed. The current anatomy is acquired during the treatment course with control images. The transformations between the images can be extracted using image registration methods.

Only little time is available to cope with observed changes during the therapy. Existing registration methods are too slow to fulfil this task. Parallelisation of the algorithmic components leads to a significant speed-up of registration methods without compromising the accuracy.

In this project we developed and implemented a highly parallel deformable image registration method using the parallel architecture of a GPU hardware. The resulting duration of computations lies between less than one to a few minutes and the good accuracy of the results can enable the implementation of adaptive strategies within the same treatment fraction.



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Hintergrund . . . . .	1
1.2	Motivation . . . . .	2
1.3	Ziel der Arbeit . . . . .	3
<b>2</b>	<b>Material und Methoden</b>	<b>5</b>
2.1	Strahlentherapie . . . . .	6
2.2	Grundlagen der Registrierung . . . . .	17
2.3	Hardwareparallelisierung . . . . .	35
<b>3</b>	<b>Resultate</b>	<b>41</b>
3.1	Konzeptionierung des Registrierungsverfahrens . . . . .	42
3.2	Evaluierung der Registrierungsergebnisse . . . . .	52
3.3	Anwendung auf Patientendatensätze . . . . .	62
<b>4</b>	<b>Diskussion</b>	<b>67</b>
4.1	Vor- und Nachteile des Registrierungsverfahrens . . . . .	67
4.2	Anwendbarkeit in der Strahlentherapie . . . . .	70
4.3	Ausblick . . . . .	72
<b>5</b>	<b>Schlussfolgerung</b>	<b>73</b>



# Abkürzungen und Akronyme

ALARA	.....	As low as reasonably achievable
ALU	.....	Arithmetical and Logical Units
API	.....	Application Programming Interface
ART	.....	Adaptive Radiation Therapy
CC	.....	Cross Correlation
CPU	.....	Central Precessing Unit
CT	.....	Computed Tomography
CTV	.....	Clinical Tumor Volume
CUDA	.....	Compute Unified Device Architecture
DGRT	.....	Dose Guided Radiation Therapy
DIR	.....	Deformable Image Registration
DVH	.....	Dose Volume Histogram
FEM	.....	Finite Element Modell
FFD	.....	Free Form Deformation
FPGA	.....	Field Programmable Gate Array
GPGPU	.....	General Purpose calculation using a Graphical Processing Unit
GPU	.....	Graphics Processing Unit
HU	.....	Hounsfield Units
ID	.....	IDentifikation
IGRT	.....	Image Guided Radiation Therapy
IMRT	.....	Intensity Modulated Radiation Therapy
JISP	.....	Joint Intensity Scatter Plot
LC	.....	Local Correlation
Linac	.....	Linear Accelerator
MI	.....	Mutual Information
MLC	.....	Multi Leaf Collimator
MR	.....	Magnet Resonance
NCC	.....	Normalized Cross Correlation
NMI	.....	Normalized Mutual Information
NTCP	.....	Normal Tissue Complication Probability
PC	.....	Personal Computer
PDE	.....	Partial Differential Equation
PET	.....	Positron Emission Tomography
RBF	.....	Radial Basis Functions
SIMD	.....	Single Instructure, Multiple Data
SISD	.....	Single Instruction, Single Data
SSD	.....	Sum of Squared Differences
TCP	.....	Tumor Control Probability

TM .....	Template Matcher
TPS .....	Thin Plate Splines
VIRTUOS ...	VIRTUal RadiOtherapy Simulator
VOI .....	Volume of Interest

# Kapitel 1

## Einleitung

### Inhaltsangabe

1.1 Hintergrund	1
1.2 Motivation	2
1.3 Ziel der Arbeit	3

Die Bildregistrierung ist ein Verfahren mit dem man eine räumliche Transformation zwischen zwei Abbildungen desselben Objekts bestimmen kann, so dass beide Abbildungen geometrisch korrekt übereinander projiziert werden können. Vom Einsatz der Registrierungsverfahren in der Strahlentherapie wird eine weitere Steigerung der Präzision der Behandlung erwartet. Die Anforderungen an Registrierungsverfahren bezüglich der Laufzeit und Qualität sind hoch. In dieser Arbeit wird eine Lösung vorgestellt, die es ermöglicht, die aufwändigen Berechnungen mit hoher Qualität innerhalb möglichst kurzer Zeit durchzuführen.

### 1.1 Hintergrund

Die Strahlentherapie ist, neben dem chirurgischen Eingriff und der Chemotherapie, eine wichtige Form der Krebstherapie. Bereits 60% aller Krebspatienten werden im Verlauf der Therapie mit ionisierenden Strahlen behandelt.

Das Ziel der Strahlentherapie ist es, durch gezielte Applikation der therapeutischen Dosis im Tumolvolumen die Tumorkontrollwahrscheinlichkeit (TCP<sup>1</sup>) zu erhöhen. Gleichzeitig müssen die Nebenwirkungen durch die meist unvermeidliche Belastung des gesunden Gewebes und der Risikoorgane (NTCP<sup>2</sup>) reduziert werden. Mithilfe hochkonformer Therapietechniken ist es heute möglich, Dosisverteilungen so zu formen, dass hohe Dosen im Tumorbereich appliziert werden und dennoch angrenzende Risikoorgane und gesundes Gewebe durch steile räumliche Dosisgradienten geschont werden können.

Die Bestrahlungsplanung für hochkonforme Therapietechniken beruht auf der Simulation der erwünschten Dosisverteilung in einem virtuellen Patientenmodell. Die Dosisberechnung benötigt in der Regel die Verteilung der Elektronendichten in diesem Modell, die durch das Planungs-CT repräsentiert wird. Die Dosisverschreibung für die im CT ab-

<sup>1</sup>engl. tumor control probability

<sup>2</sup>engl. normal tissue complication probability

gegrenzten Tumorumfängen und Risikoorganen dienen in der inversen Planung als Nebenbedingungen für die Optimierung des Plans. Die Planqualität wird aufgrund der simulierten Dosisvolumenhistogramme (DVHs) beurteilt, die widerspiegeln, in wie viel Prozent eines Volumens mehr als eine bestimmte Dosis erhalten.

Die Positionierungsungenauigkeiten und die Veränderungen der Anatomie im Verlauf der Therapie führen jedoch dazu, dass die geplante Dosisverteilung bei der Anwendung des ursprünglichen Plans nicht mehr realisiert werden kann. Strategien, diese Veränderung durch rigide Fixierungssysteme zu vermeiden zeigen nur eine begrenzte Wirkung. Deshalb versucht man Abweichungen von der geplanten Position mit Kontrollaufnahmen zu detektieren und anschließend zu korrigieren.

Um wirkungsvolle Korrekturen vornehmen zu können, müssen die anatomischen Positions- und Formänderungen, die man in den Kontrollaufnahmen sieht, erst quantifiziert werden. Hierzu werden häufig Bildregistrierungsverfahren eingesetzt. Diese ermöglichen alle zusätzlichen Kontrollaufnahmen mit dem Planungs-CT räumlich zu korrelieren, d.h. eine eindeutige Transformationsvorschrift der Bildstrukturen zwischen den beiden Zeitpunkten zu finden, an denen die Bilder entstanden sind.

Die resultierenden Abbildungen werden meist mittels Vektorfeldern beschrieben. Die Vektorfelder dienen der Positions- und Verformungsanalyse von anatomischen Strukturen, so dass mit der resultierenden Information Korrekturstrategien entwickelt werden können. Strategien, die es ermöglichen, Anpassungen aufgrund von beobachteten Veränderungen zu realisieren, werden unter adaptiver Strahlentherapie (ART) zusammengefasst.

## 1.2 Motivation

Es existiert eine Vielzahl von deformierbaren Bildregistrierungsverfahren (DIRs<sup>3</sup>). Trotz der vielfältigen entwickelten Registrierungsverfahren werden derzeit klinisch kaum DIRs eingesetzt. Und das, obwohl der Bedarf und die Erwartungen an die Algorithmen, die Therapien durch Adaptationen aufgrund von Bildinformationen zu verbessern, steigen. Der Grund dafür ist einerseits die schwierige Validierung und andererseits die hohe Rechendauer der Algorithmen. Die Validierbarkeit von DIRs ist ein prinzipielles noch ungelöstes Problem, da die wirklich stattgefundenen Veränderungen der Anatomie nicht bekannt sind und die aufgenommenen Bilder lediglich einen momentanen Zustand abbilden. Diese Arbeit konzentriert sich auf das Problem der hohen Rechendauer der DIRs. Durch eine Beschleunigung wollen wir die Möglichkeit haben, genauer zu rechnen und damit die Qualität zu erhöhen.

Trotz der schnellen Weiterentwicklung der Rechner wurden bisherige Anstrengungen zur Beschleunigung von Registrierungsverfahren häufig durch zunehmend steigende Anforderungen an Bildauflösung und Qualität kompensiert. Das Potential für eine Beschleunigung von Prozessoren durch weitere Miniaturisierung der Rechenstrukturen scheint derzeit weitgehend ausgeschöpft. Eine vielversprechende Alternative ist es, die Hardwareparallelisierung zu nutzen. Für bildverarbeitende Algorithmen sind Graphikprozessoren (GPUs<sup>4</sup>) wegen ihrer historischen Entwicklung und breiter Verfügbarkeit am ehesten geeignet.

Dieses Potential kann aber nur dann effizient genutzt werden, wenn Algorithmen so

---

<sup>3</sup>engl. deformable image registration

<sup>4</sup>engl. graphics processing unit

designiert werden, dass sie die Vorteile dieser hochparallelen Prozessoren nutzen können.

### **1.3 Ziel der Arbeit**

In dieser Arbeit stellen wir ein Konzept vor, wie deformierbare Registrierungsverfahren für den Einsatz in der adaptiven Strahlentherapie beschleunigt werden können. Hierfür haben wir die notwendigen algorithmischen Komponenten konzipiert und ein speziell auf die GPU zugeschnittenes deformierbares Registrierungsverfahren entwickelt.

Besonderes Augenmerk - im Vergleich zu vorhandenen Registrierungsverfahren - wird darauf gerichtet, durch den potentiellen Zeitgewinn die maximal mögliche lokale Bewegungsinformation aus den Bildern extrahieren zu können, um die Unsicherheit des Registrierungsergebnisses so weit wie möglich zu reduzieren.

Nur wenn die Laufzeitperformance gesteigert werden kann, ohne dabei Einbußen in der Genauigkeit zu erfordern, können adaptive Korrekturkonzepte im klinischen Einsatz der Strahlentherapie realisiert werden.



## Kapitel 2

# Material und Methoden

### Inhaltsangabe

---

<b>2.1 Strahlentherapie</b> . . . . .	<b>6</b>
2.1.1 Entwicklung der Strahlentherapie . . . . .	6
2.1.2 Einsatz der Bildregistrierung . . . . .	12
<b>2.2 Grundlagen der Registrierung</b> . . . . .	<b>17</b>
2.2.1 Registrierungskonzept . . . . .	17
2.2.2 Multi-Resolution Schema . . . . .	20
2.2.3 Algorithmische Komponenten . . . . .	20
2.2.4 Evaluierung und das Validierungsproblem . . . . .	30
<b>2.3 Hardwareparallelisierung</b> . . . . .	<b>35</b>
2.3.1 Parallele Architekturen . . . . .	35
2.3.2 Grundlagen der GPGPU . . . . .	37

---

Die hochkonformen Strahlentherapietechniken ermöglichen eine hochpräzise Tumorabdeckung und die Aussparung der Risikoorgane mit der applizierten Dosisverteilung. Um diese Techniken zur erfolgreichen Therapie einsetzen zu können, müssen die Positionen des Tumors und der Risikoorgane zu jedem Zeitpunkt der Bestrahlung bekannt sein. Bildbasiert lässt sich diese Information mithilfe von Registrierungsmethoden ermitteln. Hinter den Registrierungsmethoden stecken jedoch viele verschiedene Strategien und Annahmen. Oft reichen rigide Ansätze nicht aus, so dass deformierbare Algorithmen eingesetzt werden müssen. Eine universelle und verlässliche Methode zur Verifikation von Ergebnissen deformierbarer Registrierungsverfahren gibt es jedoch nicht. Um die Unsicherheit der Resultate abzuschätzen, haben sich einige Evaluierungsmethoden etabliert. Algorithmen, die kleinere Registrierungsunsicherheiten zeigen, sind meist rechenintensiv und benötigen sehr viel Rechenzeit, so dass sie nicht vernünftig im heutigen Strahlentherapieablauf integriert werden können. Ein optimierterer Strahlentherapieablauf bedarf eines Registrierungsverfahrens, das einen vertretbaren Kompromiß zwischen Laufzeitperformance und Ergebnisgenauigkeit liefert. Rechenintensive Algorithmen können durch Hardwareparallelisierung beschleunigt werden. Unter den unterschiedlichen parallelen Hardwarearchitekturen, die genutzt werden können, eignen sich Graphikkartenprozessoren besonders gut für die Abbildung von bildverarbeitenden Rechenaufgaben, wie der Bildregistrierung.

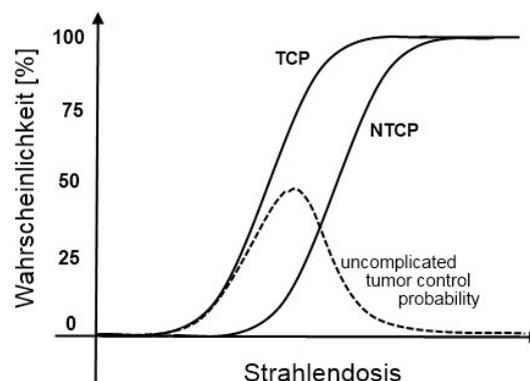
## 2.1 Strahlentherapie

Die Wirkung von ionisierenden Strahlen auf gesunde und tumoröse Gewebszellen wurde von Holthusen [1] beschrieben. Seine *in-vitro*<sup>1</sup>-Experimente zeigten, dass die Anzahl der abgetöteten Zellen mit der Dosis wie in Abbildung 2.1 ansteigt. Das Absterben der Tumorzellen der nicht strahlenresistenten Tumore im Vergleich zum Absterben der gesunden Zellen setzt in der Regel bei geringeren verabreichten Strahlendosen ein. Dieses, so genannte therapeutische Fenster, ermöglichte schon früh nach der Entdeckung der Röntgenstrahlen den therapeutischen Einsatz von ionisierenden Strahlen in der Krebsbehandlung. Bei der perkutanen<sup>2</sup> Bestrahlung von Patienten konnten so mithilfe der optimal gewählten Dosis sichtbare therapeutische Erfolge erzielt werden. Jedoch waren mit dem Rückgang des Tumors durch die unvermeidbare Strahlenbelastung des gesunden umgebenden Gewebes auch sichtbare Nebenwirkungen, wie Hautrötungen bis -verbrennungen, verbunden. Heute weiß man, dass die Belastung des gesunden Gewebes nicht nur zu akuten aber temporären Nebenwirkungen, sondern auch zu Spätfolgen – unter anderem erhöhter Wahrscheinlichkeit für Zweitumore – führt und daher, wenn möglich, zu vermeiden ist. Im Besonderen muss auch die Belastung von Risikoorganen minimiert werden, da sonst mit der zu hoch applizierten Dosis ein Ausfall von physiologischen Funktionen der Organe verbunden sein kann.

### 2.1.1 Entwicklung der Strahlentherapie

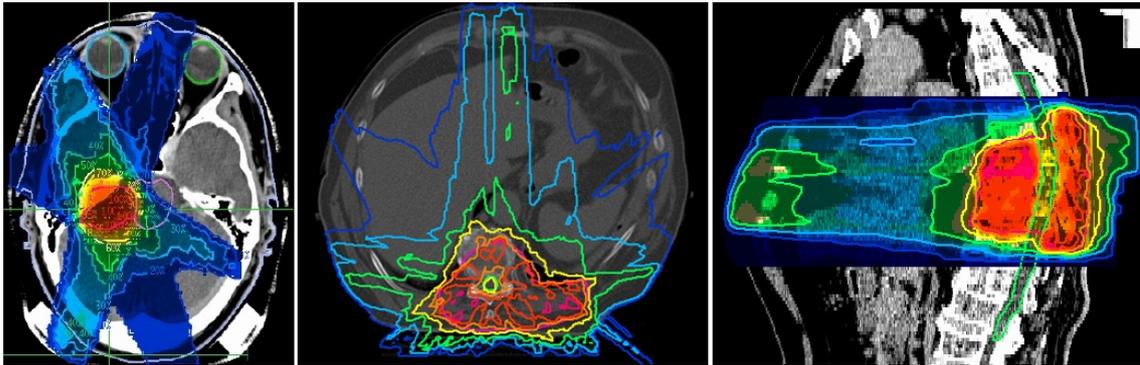
Der Kompromiss zwischen Tumorkontrolle – Applikation möglichst hoher Dosis im Tumorgewebe – und Schonung der unbetroffenen gesunden Anatomie steht seit den Anfängen im Fokus der Strahlentherapie. Sowohl therapeutische als auch technische Weiterentwicklungen in der Bestrahlungsplanung tragen dazu bei, die Nachteile des Kompromisses zu verringern. Therapeutische Konzepte, wie die Kombination mit Chemotherapie oder Variation der Fraktionierungsschemata, helfen das therapeutische Fenster zu vergrößern, so dass weniger Dosis benötigt wird, um eine hohe TCP zu erzielen. Die Grenzwerte der Dosisverträglichkeit von Risikoorganen sind so leichter einzuhalten, so dass ein geringerer Teil der gesunden Anatomie vom Hochdosisbereich erfasst und belastet wird. Technische Entwicklungen ermöglichen eine räumlich fokussiertere Applikation der Strahlendosis im Gewebe, so dass weniger von gesunder Anatomie durch den Hochdosisbereich der Dosisverteilung abgedeckt und belastet wird.

**Abbildung 2.1:** Wahrscheinlichkeit für Tumorkontrolle (TCP) und Komplikationen durch Nebenwirkungen (NTCP) als Funktion der Bestrahlungsdosis (nach Holthusen [1]). Die gestrichelte Linie zeigt den Verlauf der Wahrscheinlichkeit für Tumorkontrolle ohne starke Nebenwirkungen. Die Breite dieser Wahrscheinlichkeitsverteilung ist das therapeutische Fenster.



<sup>1</sup>lat. im Glas, im Gegensatz zu *in-vivo*: lat. im Lebendigem

<sup>2</sup>lat. durch die Haut, sinngemäß: von außerhalb des Körpers hinein



**Abbildung 2.2:** Beispiele für simulierte Dosisverteilung als Ergebnis der Bestrahlungsplanung. Von links nach rechts: Resultierende Verteilung aus einer Vorwärtsplanung mit drei Einstrahlrichtungen (entnommen aus [5], Colorwash-Darstellung), einer inversen Planung für die IMRT in transversaler (Darstellung mit Isodosenlinien) und sagittaler Ansicht (Colorwash-Darstellung).

### 2.1.1.1 Techniken

In der konventionellen Strahlentherapie definiert man die Form der Bestrahlungsfelder so, dass die gesunde Anatomie möglichst ausgeblendet wird. Durch opponierende Strahlengeometrie mehrerer Einstrahlrichtungen erreicht man sowohl homogenere Dosisabdeckung im Tumervolumen, als auch geringere Belastung des verbliebenen gesunden Gewebes im Strahlengang parallel zur Einstrahlrichtung. Der Preis für dieses Vorgehen ist, dass man nun mehr Gewebsanteil im Strahlengang hat, aber dafür die Strahlenintensität einzelner Felder reduzieren kann und somit zwar mehr gesunde Anatomie aber mit einer geringeren Dosis belastet. Auch die Form der Strahlungsfelder adaptiert man an die geometrischen Konturen der Tumorprojektionsflächen aus der Strahlrichtung. Hierzu wurden früher individuelle Blöcke für jeden Patienten und jedes Feld gegossen, die das konforme Feld definierten.

Was damals noch aufwändig war, wurde durch die Entwicklung der Multi-Leaf Kollimatoren (MLCs) [2] vereinfacht. Die MLCs sind aus mehreren jeweils gegenständigen und einzeln ansteuerbaren Metallblättern konstruiert, die durch individuelle Positionierung der Lamellen eine beliebige Öffnung formen und damit die aufwändigen gegossenen Bleiblöcke ersetzen. Trotz der Anpassbarkeit der Feldformen lässt sich die Form der Dosisverteilung in der konventionellen Strahlentherapie meist nicht an konkav geformte Zielvolumina anpassen. Dies kann aber mit der intensitätsmodulierten Strahlentherapie (IMRT) [3] erreicht werden, die auch mit MLCs realisiert werden kann. Mit dem Aufkommen der IMRT wurden selbst doughnut-geformte Dosisverteilungen applizierbar. Gewünschte Dosisverteilungen lassen sich durch die starken Dosisgradienten mithilfe der IMRT realisieren und ermöglichen damit wiederum weitere Entwicklungen der Therapie z.B. integrierte Boost-Bestrahlungen oder Dosepainting [4].

### 2.1.1.2 Bestrahlungsplanung

In der konventionellen Strahlentherapie nutzt man für die Bestrahlungsplanung die Vorwärtsplanung. Hier werden die Einstrahlrichtungen und Felder vorgegeben und die Verteilung der notwendigen Strahlenintensitäten an die gewünschte Dosis im Tumor anhand

eines Patientenmodells simuliert. Dabei gibt man basierend auf Rechenschätzungen die Strahlenintensitäten vor und simuliert die Dosisverteilung in der Anatomie des Patienten, die durch das Planungs-CT repräsentiert wird. In einem Trial-and-error-Verfahren kann man so einen optimierten Plan erstellen.

Bei der Intensitätsmodulation, die wesentlich mehr Freiheitsgrade bezüglich Strahlenintensität und Felder hat, nutzt man die inverse Planung. Hier gibt man neben den Einstrahlrichtungen nur die gewünschten Dosisgrenzwerte in einzelnen definierten Strukturen (VOIs<sup>3</sup>) vor. Durch Dosisoptimierungsalgorithmen lassen sich dann unter den eingegebenen Nebenbedingungen die optimalen Fluenzverteilungen ermitteln [6]. Beispielhafte simulierte Dosisverteilungen für beide Methoden sind in Abbildung 2.2 zu sehen.

Beide Methoden basieren auf dem Planungs-CT als 3-dimensionales Patientenmodell. Doch das Planungs-CT enthält nicht die Positionierungsunsicherheit der anatomischen Strukturen, die während des Therapieverlaufs durch z.B. Bewegungen auftreten kann. Da sowohl Repositionierung des Patienten als auch auftretende anatomische Veränderungen zu geometrischen Abweichungen des genutzten Patientenmodells führen, ist mit Abweichungen der tatsächlich applizierten von der geplanten Dosisverteilung zu rechnen.

### 2.1.1.3 Umgang mit anatomischen Veränderungen

Im Wesentlichen basiert die heutige hochkonforme Strahlentherapie auf einem stationären Plan, der wiederum auf dem statischen Planungs-CT des Patienten beruht. Die Repositionierung des Patienten, das erwünschte Anschlagen der Therapie und damit verbundene Tumervolumenabnahme oder möglicher Gewichtsverlust des Patienten unter der Therapie sind mit anatomischen Veränderung gegenüber der Planungssituation verbunden. Diese führen zu einer Verschlechterung der applizierten Dosisverteilung. Bei hochkonformen Plänen, die keine Sicherheitssäume für mögliche Bewegungen berücksichtigen, können bereits kleine Organverschiebungen zu relevanten Abweichungen von der verschriebenen Dosis führen.

Man unterscheidet einerseits zwischen inter- und intrafraktionellen anatomischen Veränderungen und damit Positionierungsungenauigkeiten und andererseits zwischen rigiden Veränderungen und Verformungen der Anatomie.

Interfraktionelle Veränderungen treten in der fraktionierten Strahlentherapie, bei der die gesamte verschriebene Dosis auf mehrere Fraktionen verteilt wird, auf. Hierbei können sowohl rigide Veränderungen als auch Verformungen der Anatomie zwischen den Fraktionen stattfinden. Rigide Positionierungsfehler können z.B. durch die notwendige Repositionierung des Patienten oder auch schon durch Fehlpositionierung in der Lagerungs- und Fixierhilfe herrühren. Verformungen, die zwischen Fraktionen auftreten, können ihre Ursache z.B. in der Gewichtsabnahme des Patienten haben, die zu einer Fehllagerung in der Fixierungshilfe führen kann.

Intrafraktionelle anatomische Veränderungen treten in den meisten Fällen auf: Sowohl innerhalb jeder Fraktion während der fraktionierten Behandlung als auch in Einzeitbehandlungen, bei denen die gesamte verschriebene Dosis nur in einer Fraktion appliziert wird. Intrafraktionelle Fehler treten aufgrund von unvermeidbaren Patientenbewegungen auf, wie z.B. der Atmung. Meistens treten damit Verformungen der Anatomie auf. Nur wenn die inneren Organstrukturen keine wesentlichen Verformungen erlauben, können auch teilrigide Veränderungen der Anatomie beobachtet werden, wie z.B. die Verschie-

---

<sup>3</sup>engl. Volume Of Interest



**Abbildung 2.3:** Beispiel für Patientenfixierungssysteme: Stereotaktischer Rahmen kombiniert mit einer festen Kopfmaske und einem Torso aus Scotchcast [7](links) und Einsatz einer Vakuummattze [8](rechts).

bung der Nieren bei der Atmung.

### Patientenfixierungen

Patientenlagerungs- und -fixierungshilfen dienen in erster Linie dazu, den Patienten optimal zu lagern und über den Zeitraum der Behandlung in der geplanten Position zu halten. Nur wenn die tatsächliche Anatomie des Patienten mit dem Patientenmodell, das auf dem Planungs-CT basiert, übereinstimmt, kann man davon ausgehen, dass die geplante Dosisverteilung verabreicht wird. Daher werden die Planungs-CTs im Unterschied zu diagnostischen CTs immer in Bestrahlungsposition, d.h. auch mit Fixierungs- und Lagerungshilfen aufgenommen.

Um möglichst gute Positionierungsreproduzierbarkeit der Anatomie zu ermöglichen werden unterschiedliche Lagerungs- und Fixierungshilfen eingesetzt. Für die Fixierung des Schädels haben sich in der Vergangenheit stereotaktische Kopfringe allein oder in Kombination mit harten Masken bewährt. Da stereotaktische Positionierungsstrategien des Schädels eine hohe Positionierungsgenauigkeit erzielten, wurden ähnliche Fixierungen auch für andere Körperbereiche eingesetzt (siehe Abbildung 2.3). Harte Masken aus Scotchcast<sup>4</sup> oder Vakuummattzen wurden mit stereotaktischen Rahmen verbunden und zur Lagerung und Fixierung eingesetzt. Heute werden häufig Vakuummattzen ohne stereotaktische Rahmen zur Lagerung genutzt. Zur Positionierungslokalisierung werden dann radioopake Marker oder Tattoos verwendet. Bei Kopf-Hals-Patienten dienen häufig Aquaplast Masken mit radioopaken Markern als Fixierung. Knie- und Fußkissen oder Mamma-schalen finden je nach Tumorlokalisierung als Lagerungshilfen Verwendung. Um Atembewegungen einzuschränken, können z.B. Bauchpressen eingesetzt werden.

Während die Cranio-Stereotaxie wegen der naturgemäßen Rigidität des Schädelknochens eine hohe Positionierungspräzision erreicht und durch die Fixierung erhält, gelingt das nicht mehr bei Fixierungen anderer Körperbereiche. Die Beweglichkeit der Anatomie einerseits und die langfristigen anatomischen Veränderungen durch z.B. Gewichtsabnahme andererseits werden gar nicht oder nur teilweise durch die mehr oder minder rigorosen Fixierungen eingeschränkt [7].

Trotz guter Fixierungshilfen, gelingt es nur teilweise Veränderungen der Anatomie gegenüber der geplanten Situation einzuschränken, so dass immer noch Abweichungen

<sup>4</sup>Gipsähnliches Material aus Kunststoff

möglich sind, die anders kompensiert werden müssen. Ein Verzicht auf jegliche Fixierungen aus Patientenkomfortgründen ist dennoch nicht ratsam. Ohne Fixierungshilfen würden noch größere Abweichungen auftreten, für die es noch keine alternativen Kompensierungsstrategien gibt.

### **Vergrößerte Sicherheitssäume**

Zum derzeitigen Standard gehört es, alle möglichen Verschiebungen oder Verformungen der Anatomie gegenüber der geplanten Situation, die man nicht korrigieren kann, auf Kosten des gesunden Gewebes durch Sicherheitssäume zu berücksichtigen [9]. Hierbei wird nur die mögliche Form- und Positionsänderung des Tumorareals durch Vergrößerung der Säume berücksichtigt, um die TCP durch die sichere Tumorabdeckung durch die Dosis zu jedem Zeitpunkt zu gewährleisten. Sicherheitssäume für Risikoorgane [10] haben sich in der klinischen Routine nicht etabliert. Befindet sich ein Risikoorgan in der unmittelbaren Nachbarschaft zum Tumorareal, kann die Vergrößerung durch Sicherheitssäume um das Tumolvolumen dazu führen, dass das Risikoorgan von ihnen mitüberdeckt wird. Hier müssen die Therapeuten eine Entscheidung treffen, welches Risiko in Kauf genommen wird: erhöhte Wahrscheinlichkeit für Nebenwirkungen oder erniedrigte Tumorkontrolle.

Um Risikoorgane besser zu schonen und den Anteil des mitbestrahlten gesunden Gewebes um das Tumolvolumen zu reduzieren, sind alle Positionierungsfehler so weit wie möglich zu reduzieren, statt diese durch weiter vergrößerte Sicherheitssäume in die Planung miteinzubeziehen. Denn ein scheinbar kleiner Saumradius bedeutet die Mitbestrahlung eines wesentlichen Volumenanteils des gesunden Gewebes, da das Volumen in der dritten Potenz mit dem Radius wächst.

### **Bildgestützte Therapie**

Um Sicherheitssäume möglichst klein zu halten, werden bildgeführte Therapiekonzepte (IGRT<sup>5</sup>) entwickelt. Mittels integrierter Bildgebungsmodalitäten am therapeutischen Linearbeschleuniger (Linac<sup>6</sup>) lässt sich die aktuelle Anatomie des Patienten in Kontrollaufnahmen abbilden. Kann man die benötigte Information über die Veränderung aus diesen Kontroll-CTs extrahieren, lassen sich Korrekturstrategien entwickeln, wie möglichst viel von der stattgefundenen Veränderung wieder kompensiert werden kann [11–14].

Die IGRT ist eine adaptive Strahlentherapie und setzt ein bildbasiertes Monitoring des Patienten während der Behandlung voraus. Doch sind zusätzliche CT Bilder auch mit zusätzlicher Strahlenbelastung für den Patienten verbunden [15]. Ist der Informationsgewinn trotz der Belastung gerechtfertigt, weil er zum besseren Therapieergebnis wesentlich beiträgt, handelt man im Sinne des ALARA Prinzips: „As low as reasonably achievable“. Da hochkonforme Bestrahlungstechniken nur in Kombination mit präziser Positionierung der anatomischen Strukturen ihre Vorteile entfalten [16], ist die zusätzliche Bildgebung gerechtfertigt, da derzeit nur so die aktuelle Position aller anatomischer Strukturen gemessen werden kann.

**Zielpunktkorrektur:** Im Verlauf der fraktionierten Bestrahlung wird der Patient für jede Fraktion neu positioniert und gegenüber der Linac-Geometrie ausgerichtet, um die Pla-

---

<sup>5</sup>engl. Image Guided Radiation Therapy

<sup>6</sup>engl. LINear ACcelerator

nungssituation zu reproduzieren. Mögliche interfraktionelle Positionierungsungenauigkeiten der Patientenanatomie lassen sich durch eine rigide Verschiebung des Zielpunktes reduzieren. Offset Positionierungsfehler können mittels rigider Registrierungsverfahren aus den akquirierten Kontrollaufnahmen ermittelt werden. Der resultierende Verschiebungsvektor des Isozentrums entspricht dann dem Positionierungsfehler, der entweder durch die Planzielpunktkorrektur am Bestrahlungsgerät selbst oder durch eine Tischverschiebung korrigiert werden kann. Die Zielpunktkorrektur wird häufig durch die Tischverschiebung realisiert, da die Verschiebung des Planzielpunktes am Gerät durch technischen Limitierungen, z.B. durch die nachlassende Homogenität des offenen Bestrahlungsfeldes zum Rand hin, auf kleine Werte begrenzt ist. Durch die Zielpunktkorrektur allein können weder Rotationsabweichungen noch Deformationen der Anatomie kompensiert werden.

**Rotationsfehlerausgleich:** Das rigide Registrierungsverfahren kann sowohl die Verschiebung des Isozentrums als auch die Rotation des rigiden Körpers um das Isozentrum gegenüber der Planungssituation ermitteln. Doch nicht alle Rotationen um die drei Raumachsen können technisch durch eine Tischdrehung realisiert werden. Um einen kompletten Rotationsfehlerausgleich zu erreichen, muß auch der Bestrahlungsplan modifiziert werden [17]. Alternativ können kleine Rotationskorrekturen mithilfe von auf einem Roboterarm gelagerten Patientenlagerungstisch vorgenommen werden. Solche Patientendrehungen könnten allerdings nur mit ausreichender Fixierung realisiert werden, um ihrerseits keine relevanten anatomischen Veränderungen zu verursachen.

**Umplanung:** Genügt eine rigide Zielpunktkorrektur nicht, um die geplante Situation wiederherzustellen, muss in den Bestrahlungsverlauf durch eine Umplanung eingegriffen werden. Eine Neuplanung ist aufwändig. Zur Definition, zur Optimierung und zur Beurteilung des Plans ist ein Modell der aktuellen Patientenanatomie notwendig. Da die Konturen des Tumors und der Risikoorgane, die auf dem ursprünglichen Planungs-CT definiert worden sind, nicht mehr zur Kontroll-Aufnahme passen, ist eine neue Segmentierung aller therapierelevanter Strukturen auf dem Kontroll-CT oder auf einem neuen Planungs-CT notwendig.

### Gating und Tracking

Intrafraktionelle Bewegungen limitieren ebenfalls die theoretisch erreichbare Präzision im Kompromiss zur optimalen Tumorabdeckung einerseits und Schonung des angrenzenden Gewebes andererseits. Die Atembewegung, im Falle von Lungen-, Leber- oder Bauchspeicheldrüsentumoren, kann eine Lageabweichung der entsprechenden Organe um mehrere Zentimeter verursachen. Um solche Tumorpositionsänderungen zu berücksichtigen, werden zur Zeit vergrößerte Sicherheitssäume auf Basis von 4D-CTs angewandt. Die wesentlich vergrößerten Zielvolumina enthalten dann jedoch auch einen großen Volumenanteil des eigentlich gesunden Gewebes. Um auf die Sicherheitssäume verzichten und z.B. möglichst viel Lungengewebe schonen zu können, werden Gating- bzw. Trackingmethoden entwickelt. Hierbei schaltet man den Behandlungsstrahl entweder nur in einem vordefinierten Bewegungsverlaufsfenster ein, wenn sich der Tumor im Strahlenfeld befindet und aus, wenn er sich aus dem Strahlenfeld hinaus bewegt, oder man versucht der Bewegung des Tumors mit dem Strahl zu folgen [18]. Die Detektion der Tumorbewegung kann hier durch im Tumorgewebe implantierte Marker erfolgen. Das Calypso System bietet z.B. solche Marker an [19]. Die Möglichkeiten eines Tumor-Trackings werden zur Zeit intensiv

erforscht [20–22], allerdings nicht an Patienten eingesetzt.

Gating- und Trackingstrategien ermöglichen den Verzicht auf vorher genutzte bewegungsinduzierte Sicherheitssäume und damit die Reduktion des bestrahlten Gewebes auf nahezu das klinische Tumolvolumen (CTV<sup>7</sup>) selbst. Nachteilig sind allerdings, neben den verlängerten Bestrahlungszeiten beim Gating, vor allem die Blindheit gegenüber Bewegungen von Risikoorganen, die nicht mit Markern versehen sind. Hier können bildgestützte Lokalisierungsverfahren helfen.

### Dosisgeführte Therapie

Während sich die IGRT Konzepte in der Regel darauf beschränken eine beobachtbare Lageabweichung durch direkte Korrektur der Patientenposition zu kompensieren, versucht man in der dosisgeführten Radiotherapie (DGRT<sup>8</sup>) Methoden zu entwickeln, welche die tatsächlich applizierte Dosis bei der Beurteilung der Abweichung und bei möglichen Planadaptionen berücksichtigen.

Um die Adaptionen aufgrund der Dosisveränderung zu ermöglichen, benötigt man erst die Rekonstruktion der tatsächlich applizierten Dosis. Eine direkte dosimetrische Messung am Patienten ist – abgesehen von einzelnen zugänglichen Messpunkten z.B. Hautoberfläche – nicht vertretbar. Daher benötigt man eine modellbasierte Rekonstruktion. Hierzu kann man die Kontrollaufnahmen nutzen, um die derzeitige Patientenanatomie abzubilden und nicht-rigide Registrierungsalgorithmen eingesetzt werden, um die anatomischen Abweichungen zu den vorherigen Zuständen zu ermitteln. Damit kann man theoretisch die in vorherigen Sitzungen bereits applizierte Dosisverteilungen auf die aktuelle anatomische Situation verformen und mit der hinzukommenden Fraktionsverteilung aufsummieren [23–26]. Erst mit der kontinuierlichen Dosisüberwachung während des Strahlentherapieverlaufs kann man Abweichungen von der geplanten Dosisverteilung detektieren und diese durch eine Planmodifikation ausgleichen.

#### 2.1.2 Einsatz der Bildregistrierung

Die Bildregistrierung, als Methode zwei Bilder räumlich zu korrelieren, wird in der Strahlentherapie häufig eingesetzt. Rigide Registrierungsverfahren helfen den Therapeuten bei der Definition von Tumor und Risikoorganen, indem sie die räumliche Kombination von Bildern verschiedener Modalitäten ermöglichen. Interfraktionelle globale und lokale Positionierungsungenauigkeiten werden mithilfe von rigiden Registrierungsverfahren quantifiziert. Diese Information kann sowohl in Form von Sicherheitssäumen in die Bestrahlungsplanung einfließen, als auch zur Reduktion der Positionierungsfehler durch Positionierungskorrekturen herangezogen werden. Deformierbare Registrierungsverfahren sollen in Zukunft noch genauere Informationen über die stattgefundenen Verformungen der Patientenanatomie liefern. Mit diesen Informationen lassen sich Patientenmodelle an den aktuellen Stand anpassen und die Akkumulation der verabreichten Dosisverteilung realisieren. Erst durch die Simulation der applizierten Dosisverteilung im Verlauf der Therapie kann auf die Ursache der Wirksamkeit der Behandlung zurückgeschlossen werden. Auch die Entwicklung neuer Korrekturstrategien und möglicher Planadaptionen, um den Abweichungen vom Geplanten entgegen zu wirken, benötigt dieses Wissen als Ausgangspunkt.

---

<sup>7</sup>engl. Clinical Tumor Volume

<sup>8</sup>engl. Dose Guided Radiation Therapy

### 2.1.2.1 Bildfusion zur Erleichterung der Volumendefinition

Für die Therapieplanung werden meist native<sup>9</sup> CT Bilder in Bestrahlungsposition benötigt. Die im CT dargestellten Grauwerte, gemessen in Hounsfield-Einheiten (HU), spiegeln die Elektronendichte des Gewebes wieder. Auf Basis dieser Information ist es möglich die zu erwartende Dosisverteilung im Gewebe zu berechnen. Oft haben aber im CT Tumore sowie einige Risikoorgane, z.B. der Hirnstamm, nicht den gewünschten Kontrast, um deren Grenzen exakt zu identifizieren. Hier können Bilder anderer Modalitäten, die andere Kontraste darstellen, zur Segmentierung herangezogen werden. Ein Beispiel für unterschiedliche Kontraste ist in der Abbildung 2.4 zu sehen. Auch wenn die Interpretation der zusätzlichen Information meist nicht trivial ist, hilft das Vorgehen die Unsicherheit, die aufgrund der schweren Abgrenzbarkeit entstehen, zu reduzieren. Jedoch ist hierfür unerlässlich, dass die Bilder der unterschiedlichen Modalitäten zu einander räumlich eindeutig korreliert sind. Diese Bildregistrierung kann entweder durch Kombination der Hardware verschiedener Modalitäten in einem Gerät wie beim PET/CT<sup>10</sup> erfolgen, oder muss, auf software-technischer Basis geschehen [27]. In diesem Zusammenhang haben sich die rigiden Registrierungsverfahren im klinischen Einsatz etabliert. Doch werden auch Fragestellungen immer wichtiger, bei denen die anatomische Verformung eine nicht zu vernachlässigende Rolle spielt. Hier benötigt man deformierbare Registrierungsalgorithmen, um die Transformation zwischen den Abbildungen zu ermitteln.

Gegenüber der hardware-basierten Registrierung der Bildinformation durch Vereinigung verschiedener Geräte ist die software-basierte Registrierung flexibler, wenn auch schwerer zu validieren. Hier können theoretisch die unterschiedlichsten Modalitäten korreliert werden, ohne technische Konstruktionsprobleme der fusionierten Geräte überwinden zu müssen. Weiterhin könnten im Idealfall auch die bereits aufgenommenen diagnostischen Bildinformationen für die Planungsunterstützung herangezogen werden.

### 2.1.2.2 Evaluierung von Sicherheitssäumen

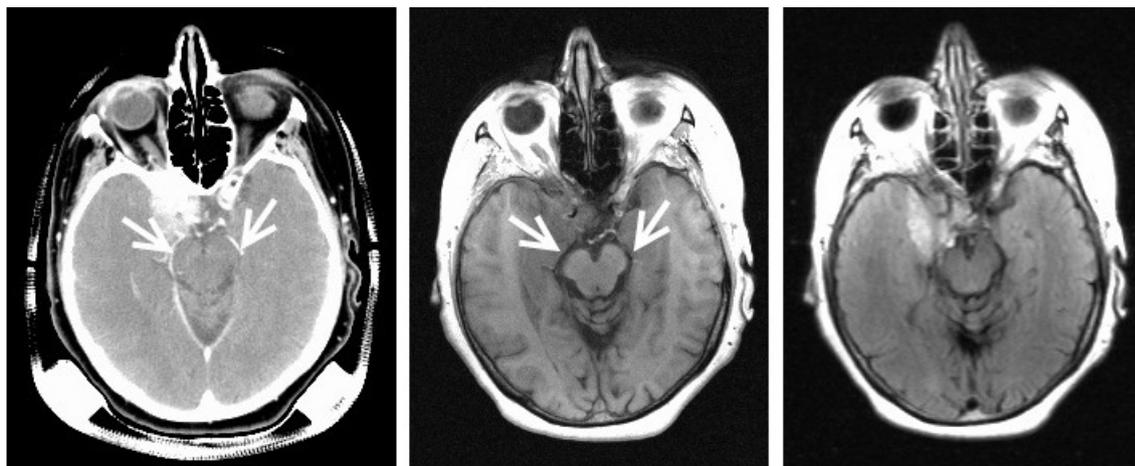
Wenn die Behandlung eines Tumors keine engmaschige bildbasierte Überwachung benötigt oder keine möglich ist, müssen bereits bei der Planung alle möglichen Positionierungs- und Bewegungsunsicherheiten durch Sicherheitssäume Beachtung finden.

Zur Ermittlung adäquater Säume hat sich das van-Herk-Rezept [28] durchgesetzt. Dieses gibt den Saumradius an, der eine Abdeckung von mindestens  $x\%$  des Volumens bei mindestens  $y\%$  der Patienten mit der verschriebenen Dosis im DVH garantiert. Die freien Parameter  $x$  und  $y$  werden meist als 95% bzw. 90% gewählt. Das Rezept wurde allerdings nur unter der Annahme eines Massenpunktes hergeleitet und deckt damit nicht die Unsicherheiten, die durch Rotationsfehler oder Verformungen auftreten können. Ein Rezept für deformierbare Zielvolumina gibt es noch nicht.

Das van-Herk-Rezept zur Berechnung des Saumradius benötigt allerdings die Ermittlung von statistischen Positionierungsunsicherheiten eines Patientenkollektivs. Hierzu werden Registrierungsverfahren eingesetzt, um die anatomischen Abweichungen aus den Kontrollaufnahmen im betrachteten Kollektiv zu ermitteln. Bisher wurden nur lokale rigide

<sup>9</sup> ohne Kontrastmittel aufgenommene Tomographische Aufnahmen

<sup>10</sup> räumlich kombiniertes Positron Emissions Tomograph mit einem Computer Tomographen.



**Abbildung 2.4:** Registrierte Schichtbildaufnahmen eines Kopfes: CT-Aufnahme mit Kontrastmitteln (links),  $T_1$  gewichtete MR-Aufnahme (mitte) und  $T_2$ -Flair MR Aufnahme (rechts). Die multimodale Bildfusion ermöglicht es unterschiedliche Informationen zu kombinieren. Im CT lässt sich der Hirnstamm (markiert mit weißen Pfeilen) nahezu nicht erkennen, hier kann man zur Abgrenzung die  $T_1$  gewichtete MR-Aufnahme wesentlich besser nutzen. Die  $T_2$ -Flair Aufnahme gibt Anhaltspunkte welche Bildbereiche zum Tumorareal gehören. Die Interpretation ist jedoch nicht trivial, denn derzeit gibt es keine Bildgebung, die spezifisch den Tumor abbildet.

Verfahren angewandt, um Schätzungen für Positionierungsfehler verformbarer Anatomie zu erhalten [8, 29, 30]. Denn gerade in den Fällen, bei denen anatomische Deformationen nicht vernachlässigt werden können, ist der Einsatz von optimierten Sicherheitssäumen unerlässlich.

Aber auch hier führen Sicherheitssäume zu ähnlichen Problemen, wie bereits unter 2.1.1.3 beschrieben. Je größer die lokalen Deformationen oder Verschiebungen sind, desto größere Säume werden benötigt. Damit steigt jedoch sowohl der Volumenanteil des mitbestrahlten gesunden Gewebes sowie die Wahrscheinlichkeit, dass Risikoorgane, die ebenfalls Deformationen und Verschiebungen zeigen, in den Hochdosisbereich rücken.

### 2.1.2.3 Positionierungskorrekturstrategien durch Image-Guidance

Bei Image-Guidance Strategien werden Kontrollaufnahmen während des Therapieverlaufs akquiriert. Mit einem rigiden Registrierungsverfahren werden die Kontroll-CTs mit dem Planungs-CT räumlich korreliert. Der resultierende Transformationsparametersatz unter dem die Bildähnlichkeit beider verglichenen CTs am größten ist, wird zu einem Korrekturvektor umgerechnet. Meist wird der Patientenlagerungstisch mitsamt dem Patienten vor der Bestrahlung um diesen Korrekturvektor verschoben, um die Planungssituation durch eine Zielpunktkorrektur so gut wie möglich wiederherzustellen.

Da mit jeder zusätzlichen CT-Aufnahme eine zusätzliche Dosisbelastung der gesamten abgebildeten Anatomie einhergeht, ist nicht immer bei jeder Behandlung eine tägliche Positionierungsüberwachung sinnvoll. Um für verschiedene Tumorlokalisationen bestimmen zu können, wie viel Kontrolle sinnvoll ist, müssen vorhandene Kontrollaufnahmen auf die Variation der Anatomie unter der Therapie untersucht werden. Hier bieten rigide Registrierungsverfahren ein geeignetes Werkzeug. Die gute Reproduzierbarkeit der

Schädelpositionierung mit dem stereotaktischen Rahmen mit einer durch Registrierung ermittelten Fehler von unter  $1.80 \pm 0.60$  mm [31], rechtfertigt z.B. keine tägliche Bildkontrolle, wenn Risikoorgane weiter vom Hochdosisbereich der Verteilung entfernt sind. Die Positionierungsunsicherheit von Patienten mit paraspinalen Tumoren allerdings, bei denen das Rückenmark als Risikoorgan fast vollständig vom Tumor umschlossen ist und sehr steile Dosisgradienten in der Verteilung geplant werden müssen, kann trotz möglichst rigider Patientenfixierung zu hoch sein [7]. Hier ist eine tägliche Bildüberwachung mit Zielpunkt-korrektur ratsam.

Auch wenn zusätzlich zu Offsetfehlern Deformationen der Patientenanatomie hinzukommen, wie z.B. beim Patientenkollektiv mit Kopf-Hals-Tumoren, können rigide Registrierungsverfahren eingesetzt werden, um, fokussiert auf interessante Strukturen, einen minimalen lokalen Positionierungsfehler der jeweiligen Struktur zu ermitteln. Auch wenn die IGRT keine Kompensationsmöglichkeit für diese Restfehler bietet, können die ermittelten Fehler wenigstens durch Sicherheitssäume in der Planung berücksichtigt werden [8, 29, 30].

Die Zielpunktkorrektur auf Basis von Kontrollaufnahmen und rigiden Registrierungsverfahren wird immer häufiger eingesetzt. Meistens wird das gesamte aufgenommene Bild der rigiden Registrierung unterworfen. Beim Vorhandensein von nicht vernachlässigbaren Deformationen werden aber so ermittelte Zielpunktkorrekturvektoren die lokale Positionierungsungenaugigkeit der VOIs unterschätzen [32]. Alternativen für die Berechnung eines optimaleren globalen Korrekturvektors aus lokalen rigiden Registrierungen mit unterschiedlichen Gewichtungen werden weiter erarbeitet [33].

Dennoch bleibt die Zielpunktkorrektur eine globale Korrekturstrategie, die lediglich den rigiden Offsetfehler gegenüber der Planungssituation zu kompensieren vermag. Deformationen bleiben als Restfehler erhalten, die zu dosimetrischen Abweichungen von der verschriebenen Dosis führen können. Auch die aufgrund von rigiden Registrierungs-algorithmen hergeleiteten Sicherheitssäume sind nur minimale Schätzungen der eigentlich benötigten Säume, um die Abdeckung der Deformationen mit in die Planung integrieren zu können. Der nächste Schritt, um diese Limitierungen zu überwinden, kann der Einsatz von deformierbaren Registrierungsverfahren zur Herleitung von Sicherheitssäumen bei deformierbarer Anatomie sein.

Die Häufigkeit, mit der Kontrollaufnahmen generiert werden, wird zur Zeit wesentlich durch den Wunsch limitiert, die zusätzliche Dosisbelastung des Patienten durch die Röntgenstrahlung gering zu halten. Projekte, die Bildmodalitäten mit dem Linac fusionieren wollen, die nicht mit zusätzlicher Dosisbelastung verbunden sind, wie der MR-Tomograph [34], versprechen die Möglichkeiten für ein Monitoring in Zukunft zu erweitern und eine kontinuierliche Überwachung der Patientenanatomie während des Therapieverlaufs zu ermöglichen. Um die geometrischen Veränderungen aus den Bildern zu extrahieren, wird man auch hier auf deformierbare Registrierungs-algorithmen zurückgreifen wollen.

#### 2.1.2.4 Registrierung bei Umplanung

In der fraktionierten Strahlentherapie werden manchmal im Therapieverlauf Umplanungen nötig. Da IGRT Strategien keine lokale Verformungen kompensieren können, aber größere Deformationen z.B. wegen starkem Gewichtsverlust des Patienten auftreten können, wird ein neuer Plan auf dem Patientenmodell des letzten Kontroll-CTs erstellt. Für

die Planerstellung an sich wird ein adaptiertes Patientenmodell benötigt. Die Lage und Form der aktuellen VOIs kann entweder auf der Kontrollaufnahmen aufwendig neu segmentiert werden oder man setzt deformierbare Registrierungsalgorithmen ein, um das Deformationsfeld zwischen dem ursprünglichen Planungsmodell und der aktuellen Anatomie zu ermitteln. Dann lassen sich die bereits eingezeichneten VOIs schnell nach der Transformationsvorschrift in das neue Patientenmodell transferieren [35, 36]. Der Einsatz der Bildregistrierung verkürzt die Umplanungszeit enorm, da Neukonturierungen meist annähernd eine Stunden benötigen.

#### **2.1.2.5 Registrierungsalgorithmen beim Gating und Tracking**

Bei Gating- oder Trackingstrategien ist es wesentlich die Position des Tumorumens zu jedem Zeitpunkt zu kennen. Um die schnellen und auch plötzlichen Bewegung kontinuierlich detektieren zu können, können z.B. trackbare Marker in den Tumor implantiert werden [19]. Jedoch sind solche nicht bildbasierte Gating- und Trackingsysteme blind gegenüber Bewegungen von Risikoorganen. Auch wenn 2- oder 3-dimensionale Bildgebungssysteme nicht mit der gleichen Rate Bilder aufnehmen können, so lassen sich damit Informationen über die Risikoorganbewegungen ermitteln. Registrierungsalgorithmen können hier zur Bewegungsextraktion eingesetzt werden, um den Strahl zu unterbrechen, wann auch immer Risikoorgane in den sich bewegenden Strahlengang rücken. Solche Anwendungen für Registrierungsalgorithmen erfordern eine besonders hohe Geschwindigkeit, um mit den Änderungen der Anatomie und der Frame-Rate der Bildgebungsmethode Schritt halten zu können.

#### **2.1.2.6 Simulation der tatsächlich applizierten Dosis**

##### **Umplanungen und Re-planungen**

Bei der Umplanung in der fraktionierten Therapie wird nicht beachtet, wieviel Dosis bereits in welchen VOIs durch die bisherigen Fraktionen appliziert worden ist, da es bislang keine etablierten Verfahren gibt, wie Einzeldosen über die Zeit zu akkumulieren sind, wenn sich Lage, Größe und Form der anatomischen Strukturen ändert. Daher verschreibt man im neuen Plan die noch theoretisch fehlende Dosis, um die ursprüngliche Verschreibung zu erreichen. Mithilfe von täglicher Fraktionsbildgebung und deformierbarer Registrierung kann die tatsächlich bereits applizierte Dosis genauer ermittelt werden und somit in der Optimierung des neuen Plans berücksichtigt werden.

Die Berechnung der tatsächlich applizierten Dosis auf veränderlicher Anatomie kann auch bei Re-Bestrahlungen eingesetzt werden. Ein denkbarer Einsatz von deformierbaren Registrierungsalgorithmen ist die räumliche Korrelation von Bildern der vorangegangenen Bestrahlungsplanung mit der aktuellen Anatomie des Patienten zum Zeitpunkt der Re-Bestrahlung. Für den Therapeuten ist es hierbei wichtig zu wissen, welche Gewebsareale in der Nähe des Rezidivs oder des neuen Tumorumens durch die ehemalige Behandlung vorbelastet sind, um weitere bestrahlungsinduzierte Toxizitäten im der aktuellen Planung zu vermeiden. Heutige Registrierungsverfahren können natürlich nur die geometrische Abweichungen erfassen. Eine biologische Modellierung des Gewebes über einen häufig jahrelangen Zeitraum kann noch nicht sinnvoll erstellt werden. Dennoch kann die geometrische Anpassung den ersten Verbesserungsschritt darstellen, um die Wahrscheinlichkeit für den Patienten zu minimieren, dass Komplikationen auftreten.

### Online Planadaption

Theoretisch kann die Akkumulation der bisher applizierten Dosisverteilungen auf veränderlicher Anatomie auch sofort für eine Adaption des Plans eingesetzt werden. Dazu benötigt man, neben zuverlässigen Deformationsfeldern, schnelle Registrierungs-, Planoptimierungs- und Dosisberechnungsverfahren.

Bisher haben nur retrospektive Untersuchungen die möglichen Vorteile der DGRT aufgezeigt [37–40]. Alle setzen deformierbare Registrierungsalgorithmen ein, um die Dosisakkumulation während der Therapie zu realisieren. Die Fehler der Registrierungsverfahren stellen den größten Beitrag zur Unsicherheit dieser Berechnungen dar [41]. Da die großflächige Validierung der Registrierungsergebnisse vor allem in homogenen Bildbereichen nicht möglich ist, läßt sich der Fehler nicht genau quantifizieren. Unsicherheitsschätzungen durch z.B. Vergleiche mit expertendefinierten Landmarken können den wahren Fehler unterschätzen, da auch visuelle Identifikation von Landmarken nur in kontrastreichen Bildbereichen sinnvoll ist. Methoden, um Konfidenzintervalle für Transformationsergebnisse zu ermitteln gibt es wenige. Sie sind meist rechenintensiv und bisher limitiert auf rigide Transformationsmodelle [42].

Trotz der Unsicherheit der resultierenden Transformationen von reinen Registrierungsverfahren, bieten diese die einzige Möglichkeit um die benötigten Deformationen aus den Bildern zu extrahieren, weshalb Registrierungsverfahren weiterhin verbessert werden müssen. Die Rechengeschwindigkeit der Registrierungsverfahren muß gesteigert werden, um in adaptiven online Strategien schnell die notwendigen anatomischen Veränderungen zu ermitteln.

## 2.2 Grundlagen der Registrierung

Unter Registrierung versteht man eine Methode zwei verschiedene Bilder so räumlich zueinander zu korrelieren, dass der gleiche Informationsgehalt in beiden Bildern eindeutig einander zugewiesen werden kann. Da man die wahre Lösung des Problems in der Regel nicht kennt, formuliert man die Registrierungsaufgabe als einen Optimierungsprozess. Um das Risiko zu minimieren, während des Optimierungsprozesses in lokale Extrema zu konvergieren, werden häufig Multi-Resolution Ansätze verfolgt.

Ein Registrierungsverfahren besteht aus verschiedenen algorithmischen Komponenten, die nach Bedarf unterschiedlich zusammengestellt werden und damit zu unterschiedlichen Lösungen führen können. Die Validierung der Registrierungsergebnisse ist schwierig und nur wenige Evaluierungsmethoden haben sich etabliert, um die Fehler abzuschätzen.

### 2.2.1 Registrierungskonzept

Während sich rigide Registrierungsverfahren im klinischen Alltag bereits etabliert haben, finden deformierbare Verfahren in der Anwendung noch keine Akzeptanz. Denn unabhängig von genutzten Strategien stellt die deformierbare Registrierung ein schlecht definiertes Problem dar. Eine eindeutige Lösung ist unbekannt und verschiedene Registrierungsalgorithmen werden bei gleichen Inputdaten häufig zu verschiedenen Ergebnissen führen [43, 44]. Eine Methode zur Validierung der Ergebnisse ist trotz der langen Erforschung des Themas nicht etabliert. Häufig ist das verwendete Transformationsmodell das ausschlaggebende Kriterium für die Glaubwürdigkeit des Ergebnisses, doch je bio-

physikalisch realistischer die Modelle werden sollen, desto rechentechnisch aufwendiger wird der Lösungsprozess. Der Anspruch an die Qualität von Registrierungsergebnissen treibt die Entwicklungen von bio-physikalischen Modellierung der Patientenanatomie im Bereich der Finiten Elementmodellierung voran. Neben den auch hierfür schwer zu ermittelnden Parametern bringen alle FEM-basierte Lösungsansätze einen für den klinischen Einsatz immensen Nachteil mit: ihre lange Rechendauer im Bereich mehrerer Stunden. Der Anspruch vieler Registrierungsverfahren ist nicht ein vergleichbares Ergebnis mithilfe der rein grauwertbasierten Registrierungsmethoden zu erzielen. Der Anspruch von Verfahren, die sich nur auf die reine Auswertung der Grauwertinformationen verlassen, beschränkt sich oft darauf, visuell glaubhafte Resultate in kurzer Zeit zu ermöglichen. Aufgrund der knappen Zeit, die man häufig nur zur Verfügung hat, nimmt man Fehler die dabei entstehen in Kauf, damit man überhaupt eine Chance hat, die wesentlichen Auswirkungen der Deformationen zu berücksichtigen.

### 2.2.1.1 Formulierung

Der Registrierungsprozess kann als ein Optimierungsprozess formuliert werden. Optimiert wird hierbei eine Zielfunktion  $M$ , die die Ähnlichkeit beider Bilder  $R$  und  $D$  zu einander unter einer Transformation  $T$  beschreibt

$$M(T_{opt}) = \max M[R(i), T(D(i))]. \quad (2.1)$$

Hierbei sind  $R$  das statische Referenzbild,  $D$  das durch die Transformation verformte deformierte Template und  $i$  der Index über alle Voxel. Äquivalent zur Nutzung von Ähnlichkeitsmaßen ist die Verwendung von Distanzmaßen, die so definiert sind, dass im Optimierungsprozeß ihr Minimum angestrebt wird. In der Registrierung finden unterschiedliche Ähnlichkeitsmaße Verwendung. Dementsprechend unterscheidet man je nach eingesetztem Ähnlichkeitsmaß, ob ein Registrierungsverfahren nur Bilder gleicher Modalität (monomodal) korrelieren kann. Oder ob es die Voraussetzung, dass gleiche Gewebe durch gleiche Grauwerte abgebildet werden, nicht unterstellt und damit zu multimodalen Registrierungsverfahren zählt.

Genauso wie bei Ähnlichkeitsmaßen gibt es auch bei verwendeten Transformationsmodellen unterschiedliche Strategien. Wird ein Transformationsmodell eines festen Körpers abgebildet, nennt man das Registrierungsverfahren rigide. Ein nicht-rigides oder auch deformierbares Verfahren liegt vor, wenn das Transformationsmodell die Nebenbedingungen des festen Körpers nicht erfüllt und sich damit auch Verformungen des Raums darstellen lassen. Es gibt viele verschiedene Ansätze ein nicht-rigides Transformationsmodell zu parametrisieren.

Als dritte einflussreiche Komponente des Registrierungsverfahrens kann man die Optimierungsstrategie ansehen. Denn eine der größten Herausforderungen der nicht-rigiden Registrierung ist die hohe Anzahl der Freiheitsgrade aller Parametrisierungen, was die Optimierung wesentlich erschwert und das Risiko in lokalen Extrema zu konvergieren erhöht. Um dieses Risiko zu minimieren, können unterschiedliche Optimierungsverfahren gewählt werden. Einen Überblick über verschiedene Registrierungsverfahren in der medizinischen Bildverarbeitung geben [27, 44–46].

### 2.2.1.2 Globale vs. lokale Optimierung

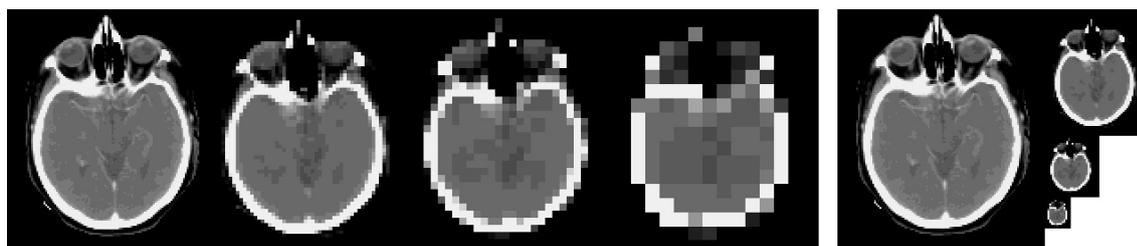
Der Optimierungsprozess kann unterschiedlich angewandt werden. Im globalen Ansatz bestimmt man die Transformation des gesamten verformbaren Bildes und das gesamte Bild trägt zur Bestimmung der Zielfunktion bei. Damit tragen alle Bildvoxel zum Optimierungsprozess bei. Auch das Transformationsmodell ist so der Bewertung durch das Ähnlichkeitsmaß unterworfen. Der optimale Wert der Zielfunktion kann so unerreicht bleiben, wenn die tatsächliche Transformation nicht durch das Transformationsmodell darstellbar ist. Da nur ein globaler Wert der Zielfunktion optimiert wird, kann man nicht die Beiträge einzelner lokaler Transformationen erkennen, so kann z.B. eine Verschlechterung der Ähnlichkeit an einer Stelle im Bild die Verbesserung an anderer Stelle betragsmäßig kompensieren.

Ein alternativer Ansatz ist eine lokale Suchstrategie der Optimierung. Bei z.B. einem Block-Matching-Verfahren definiert man mehrere lokale Bildausschnitte in Blockform und optimiert nur deren Ähnlichkeitsmaß bezüglich der lokalen Transformation [35, 47, 48]. Beim Block-Matching Verfahren werden häufig lokal-rigide Transformationen eingesetzt, so dass einzelne *per se* rigide Bildausschnitte sich unabhängig von einander im Suchraum bewegen können. Ein Vorteil eines lokalen Verfahrens ist, dass man die Optimierung einzelner Blöcke unabhängig von einander beenden kann und so insgesamt schneller einzelne robuste über das Bild verteilte Korrespondenzen erhält. Da die Korrespondenzen lokal beschränkt sind, benötigt man bei diesem Ansatz aber eine mathematische Interpolation der Bewegung zwischen den Punktkorrespondenzen, um diese zu einem dichten Feld von Bewegungsvektoren zu vervollständigen. Hierzu können verschiedene Transformationsmodelle eingesetzt werden. In Kapitel 2.2.3.2 werden einige vorgestellt.

### 2.2.1.3 Brute Force vs. Optimierung

Das Problem mit allen Optimierungsstrategien ist, dass sie bei nicht konvexen Zielfunktionen in lokale Extrema konvergieren können. Für deformierbare Registrierungsverfahren, die sich durch eine Vielzahl von Freiheitsgraden auszeichnen, und deren Einsatz in der medizinischen Bildverarbeitung, wo wenige eindeutige lokale Strukturen vorhanden sind, ist dieses Problem besonders wichtig. Um dieses Risiko zu minimieren, benötigt man ein bio-physikalisch realistisches Modell, um den Suchraum der Lösungen auf die realistischen einzuschränken. Doch die Parameter bio-physikalischer Transformationsmodelle sind wenig bekannt und schwierig am lebenden Probanden auszumessen. Weiterhin benötigt die Lösung des Optimierungsprozesses unter Einsatz solcher Modelle viel Zeit.

Da dieser Weg für zeitkritische Anwendungen keine Lösung darstellt, braucht man andere Strategien, die den Optimierungsprozess unweigerlich verlängern. Bei lokalen Registrierungsansätzen wie dem Block-Matching kann man auf eine Optimierungsstrategie ganz verzichten und den ganzen definierten Suchraum der Transformationen berechnen, will man sicher sein, das globale Extremum zu finden. Denn da einzelne Bildausschnitte unabhängig von einander optimiert werden, reduziert sich die Gesamtzahl der Freiheitsgrade. Eine Implementierung des Block-Matching-Verfahrens ohne den Einsatz einer Optimierungsstrategie wurde von Malsch *et. al.* präsentiert [35] und erreicht neben guter Genauigkeit auch eine verhältnismäßig gute Geschwindigkeitsperformance.



**Abbildung 2.5:** Realisierung des Multi-Resolution Schemas: Die Einträge einer Gaußpyramide für eine beispielhafte CT-Schichtaufnahme eines Schädels. Das linke Paneel zeigt die sukzessive Unterabtastung nach dem Tiefpaßfilter. Zur Vergleichbarkeit der Bilddetails wurde in die jeweiligen Stufen hineingezoomt. Das rechte Paneel zeigt die tatsächlichen Größenverhältnisse der Daten im Speicher.

### 2.2.2 Multi-Resolution Schema

Da gerade bei der deformierbaren Registrierung ein Optimierungsproblem mit extrem vielen Freiheitsgraden vorliegt, haben sich Multi-Resolution Schemata zur Vereinfachung etabliert [49–51]. Man versucht sich die Problemstellung hierbei auf unterschiedlichen Auflösungsstufen anzuschauen. Große Verformungen können auf gröberer Bildauflösungen schneller gefunden werden und das Risiko auf feinerer Auflösung in lokalen Extrema während der Optimierung zu konvergieren wird reduziert.

Es gibt verschiedene Ansätze das Multi-Resolution Schema umzusetzen. Die beliebteste Methode ist es, das Schema mit Hilfe einer Gaußpyramide, die die Bildinhalte in verschiedene Frequenzanteile zerlegt, zu realisieren. Die Erstellung der einzelnen reduzierten Bilder wird iterativ mittels eines Tiefpaßfilters und einer reduzierten Neuabtastung durchgeführt. Während das Originalbild das gesamte ursprüngliche Frequenzspektrum enthält, wird bei steigender Stufe in der Gaußpyramide der enthaltene Wellenzahlbereich immer mehr eingeschränkt. Die Bildzeilen und -spalten im 2D Fall können hierbei ohne weiteren Informationsverlust nach der Filterung reduziert werden. Da bei diesem Prozess die hohen Frequenzen weg gefiltert werden, verschwinden in den so erzeugten Bildern sowohl scharfe Kanten als auch kleine Strukturen (siehe Abbildung 2.5). Diese sind für große Verschiebungen im Falle der Registrierungsaufgabenstellung zunächst nicht relevant. Findet man die grobe Transformation zwischen den Bildern, kann man sukzessive zur immer feineren Granularität zurückkehren und immer mehr Details zur Transformation hinzufügen.

### 2.2.3 Algorithmische Komponenten

Registrierungsalgorithmen können sehr unterschiedlich aufgebaut sein. Dennoch lassen sich grundlegende algorithmische Komponenten immer wiederfinden. Zu den wesentlichen Komponenten der Registrierung gehören Ähnlichkeitsmaße, Transformationsmodelle und eine Suchstrategie, d.h. ein Optimierungsverfahren.

#### 2.2.3.1 Ähnlichkeitsmaß

Ähnlichkeitsmaße dienen dazu die Abhängigkeit eines abgebildeten Objekts bezüglich eines Modells zu quantifizieren. Im Falle von Bildregistrierung stellt das statische Referenzbild das Modell dar, während das zu transformierte Bild dem Objekt entspricht. Ein Ähn-

lichkeitsmaß ist also eine Funktion die beiden Zuständen der zu vergleichenden Bilder einen Ähnlichkeitswert zuweist. Damit ist das Ähnlichkeitsmaß auch die Zielfunktion im Optimierungsproblem der Registrierung.

Je nach dem, welche Arten von Bildern miteinander registriert werden sollen, benötigt man unterschiedliche Ähnlichkeitsmaße. Einige sind nur für monomodale Registrierungsaufgaben geeignet, andere auch für multimodale Aufgaben.

Weiterhin kann man zwischen intensitäts- und featurebasierten Ähnlichkeitsmaßen unterscheiden. Die intensitätsbasierten Maße benötigen keinen weiteren Vorverarbeitungsschritt, da nur direkt die Bildintensitätswerte allein in die Berechnung mit einfließen. Anders ist es bei featurebasierten Maßen, die zunächst die Features aus den Bildern erkennen müssen, seien es Grenzkonturen oder eindeutig identifizierbaren Landmarken [52]. Bei den letztgenannten Methoden lassen sich allerdings besser die Nebenbedingungen formulieren. Die intensitätsbasierten Maße werden dafür häufig wegen ihrer Geschwindigkeit favorisiert.

Das Ähnlichkeitsmaß bildet eine wesentliche Komponente eines Registrierungsalgorithmus und entscheidet über die Qualität des Registrierungsergebnisses. Doch sind die meisten Ähnlichkeitsmaße keine im Suchraum konvexe Funktionen, da einerseits sich im Patientenmodell - z.B. einer CT Aufnahme - wenige lokal eindeutige Strukturen befinden und andererseits sich bei der Berechnung zusätzliche Artefakte ergeben können [53, 54]. Während die letzteren durch zusätzliche Rechnungen verringert werden können, kann man das erstgenannte Problem nur mit einem guten Patientenbewegungsmodell reduzieren, das den Suchraum auf nur realistische Transformationen beschränkt.

Im Folgenden sollen einige häufig eingesetzte einfache intensitätsbasierte Ähnlichkeitsmaße vorgeteilt werden, um die Struktur und den Einfluss auf den Registrierungsprozess zu vermitteln.

**Sum of Squared Differences (SSD)** ist ein intensitätsbasiertes Maß, das nur von Intensitätswerten  $I(i)$ , bzw. Grauwertdifferenzen der Bilder direkt abhängt und kein Wissen über die Eigenschaften des dargestellten Objekts benötigt

$$SSD = \sum_{i=1}^n (I_R(i) - I_D(i))^2. \quad (2.2)$$

Die Indizes  $R$  und  $D$  stehen für das Referenzbild und das deformierte Template. Da SSD direkt die Intensitätsdifferenzen vergleicht, kann man nur Bilder gleicher Modalität zum Ähnlichkeitsvergleich heranziehen. SSD ist damit ein monomodales Ähnlichkeitsmaß. Rechentechisch ist der Einsatz von SSD sehr beliebt, da sich quadratische Formen gut für schnelle Optimierungsmethoden wie das Newton-Verfahren (siehe Abschnitt 2.2.3.3) zur Lösung anbieten. Ebenfalls eignet sich SSD für Parallelisierungsansätze, da große Datenmengen nahezu unabhängig von einander verarbeitet werden können.

**Cross Correlation (CC)** kann ebenfalls im monomodalen Kontext für Registrierungen eingesetzt werden. Auch CC hängt nur von Grauwertintensitäten der Bilder ab:

$$CC = \sum_{i=1}^n (I_R(i) - \bar{I}_R)(I_D(i) - \bar{I}_D), \quad (2.3)$$

hierbei bezeichnen  $\bar{I}$  den jeweiligen Mittelwert der Intensitätsverteilung der Bilder. Meistens wird die normierte Form der CC eingesetzt:

$$NCC = \frac{\sum_{i=1}^n (I_R(i) - \bar{I}_R)(I_D(i) - \bar{I}_D)}{\sqrt{\sum_{i=1}^n (I_R(i) - \bar{I}_R)^2} \sqrt{\sum_{i=1}^n (I_D(i) - \bar{I}_D)^2}}. \quad (2.4)$$

Ein wesentlicher Vorteil der CC ist, dass sich ihre Berechnung über ihre Darstellung im Fourierraum beschleunigen lässt, was bereits in Registrierungsverfahren von z.B. Malsch *et al.* [35] Verwendung fand. Modifikationen dieses Maßes erlauben flexibleren Einsatz: z.B. lässt sich die Kreuzkorrelation von lokalisierten Subregionen im Bild in Form der lokalen Korrelation (LC) zu einem multimodalen Maß erweitern [55].

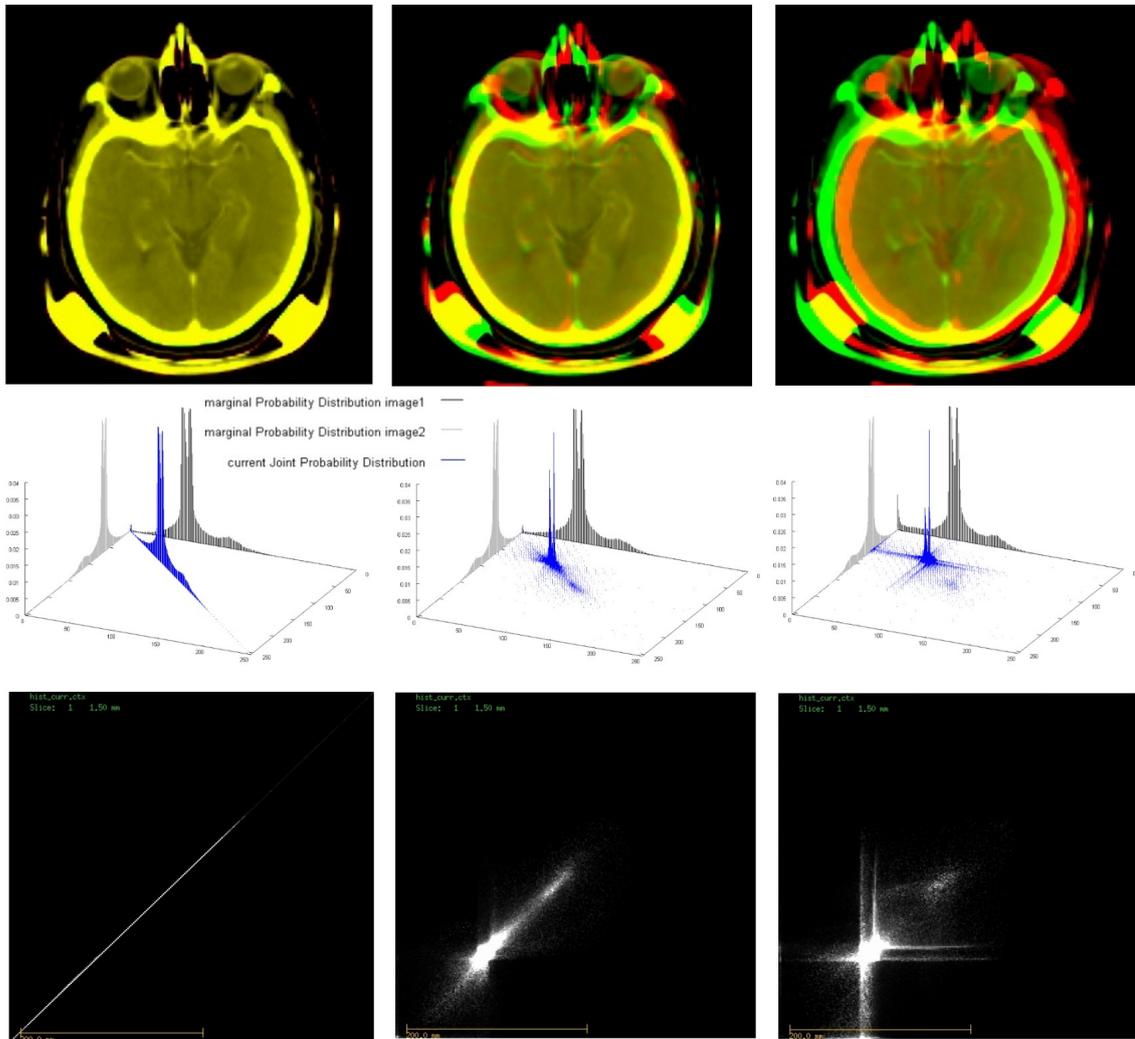
**Mutual Information (MI)** ist ein weit verbreitetes entropiebasiertes Ähnlichkeitsmaß für multimodale Registrierung [45, 56, 57]. Der grundlegende Gedanke ist die statistisch signifikante Entsprechung der Grauwerte in beiden Bildern zu untersuchen. Die gegenseitige Wahrscheinlichkeitsverteilung, die durch das gemeinsame Histogramm beschrieben wird, zeigt im Fall einer besseren räumlichen Übereinstimmung der Bildinhalte eine geringere Unordnung und damit, statt einer breiten Verteilung von gegenseitigen Grauwertentsprechungen, ausgeprägte Häufungspunkte im Histogramm. Dies ist in Abbildung 2.6 dargestellt. Bei der 2-dimensionalen Darstellung des Histogramms in Abbildung 2.6 kann man die Häufigkeiten, die als Intensitätswerte dargestellt sind, nicht direkt vergleichen, da hier um die Breite der Verteilung darzustellen eine Grauwertskalierung durch eine Grauwertfensterung vorgenommen wurde. Hier ist nur die Verbreiterung der Verteilung bei der Zunahme der Abweichung sichtbar. Die MI berechnet sich durch

$$MI = - \sum_{r,d=1}^h (p_{r,d} \log \frac{p_{r,d}}{p_r p_d}), \quad (2.5)$$

wobei  $p_{r,d}$  die gemeinsame Wahrscheinlichkeitsverteilung in Form des 2D gemeinsamen Histogramms ist mit  $h$  Grauwertbins auf der  $x$  und  $y$  Achse, und  $p_r$  sowie  $p_d$  die Einzelhistogramme der beiden Bilder bezeichnen. Eine Erweiterung mit äquivalentem Verhalten stellt die normierte Version der MI, die so genannte Normalized Mutual Information (NMI), dar. Die MI bzw. NMI gilt heute als das am meisten eingesetzte Ähnlichkeitsmaß, weshalb es darauf basierende Modifikationen und Erweiterungen gibt [58]. Die Nachteile der MI liegen einerseits im größeren Berechnungsaufwand verglichen zu z.B. SSD und andererseits auch in seiner statistischen Herleitung. Während MI ein robustes globales Maß ist, gelingt es sie nur bis zur einer Grenze von lokalen Bereichen auf lokale Schemata anzuwenden, was gerade für deformierbare Registrierungsansätze wichtig wäre. Einige Versuche wurden unternommen, um räumliche Information mit zu integrieren [59]. Diese erfordern jedoch einen erhöhten Rechenaufwand.

**Joint Intensity Scatter Plot (JISP) basierte Maße** sind Ähnlichkeitsmaße, die direkt auf der gemeinsamen Wahrscheinlichkeitsverteilung der Grauwerte beider Inputbilder basieren. Anders als bei MI, wird die lokale Zusammensetzung des gemeinsamen Histogramms, das die gemeinsame Wahrscheinlichkeitsverteilung annähert, direkt optimiert. Auch bei der MI wird indirekt das Phänomen ausgenutzt, dass registrierte Bilder mehr Ordnung, d.h. auch schärfere Häufungspunktverteilungen, im JISP<sup>11</sup> aufweisen. Doch im Endeffekt

<sup>11</sup>engl. Joint Intensity Scatter Plot



**Abbildung 2.6:** Das gemeinsame Histogramm als Basis für die Mutual Information: Obere Zeile: eine Serie von überlagerten CT Bildern. Eine CT Schädelaufnahme sich selbst überlagert (links), selbe CT Schicht überlagert mit einer um  $10^\circ$  rotierten Kopie (mitte) und selbe Schicht überlagert mit einer um  $10^\circ$  rotierten und um 10 mm verschobenen Kopie (rechts). Zweite Zeile: Die gemeinsamen Histogramme der jeweiligen Bildpaare in 3D Ansicht. Dritte Zeile: gemeinsame Histogramme in 2D Ansicht als Helligkeitsverteilung. Auf den jeweiligen Achsen sind die Grauwertstufen der jeweiligen Inputbilder aufgetragen.

ist MI ein skalares Maß, das die Dispersion im Histogramm global wertet. Während JISP-basierte Maße auch lokale Veränderungen des Histogramms detaillierter auflösen können. Ansätze für solche Maße wurden von Leventon *et. al.* [60] und Orchard [61] entwickelt. Leventon's Maß: Joint Intensity Distribution, basiert auf einer Maximum Likelihood Registrierung unter der Bedingung, dass zuvor trainierte JISP mit bereits optimal registrierten Bildern, durch die Transformationsparameter angestrebt wurden. Orchard dagegen, führte punkt- und linienförmige Regressoren in das JISP ein, um die Häufungspunktverteilungen darin zu modellieren. Die Bewertung der Voxelpositionen basiert bei ihm auf der Summe der Abweichungsquadrate von den nächsten definierten Regressoren. Beide Ähnlichkeitsmaße sind nicht weit verbreitet und wurden nur für den Einsatz mit rigiden Transformationsmodellen getestet.

### 2.2.3.2 Transformationsmodell

Abgesehen vom Ähnlichkeitsmaß beeinflusst das Transformationsmodell das Ergebnis der Registrierung. Das Transformationsmodell dient der Interpolation der Bewegung zwischen den kontrastreichen Bildbereichen, wie den sichtbaren Gewebsgrenzen, deren Übereinstimmung durch das Ähnlichkeitsmaß quantifizierbar ist. Der Wert der Bewertungsfunktion für Ähnlichkeit ändert sich nicht, wenn sich das zugrunde liegende Modell in homogenen Bereichen ändert. Damit liegt die größte Fehlerquelle des Registrierungsprozesses in der Fähigkeit die wirklich stattgefundene Bewegung durch das Transformationsmodell abzubilden.

Mithilfe des Transformationsmodells wird ebenfalls der Suchbereich im Optimierungsprozess auf die modellierten Basisfunktionen, die dem Modell zugrunde liegen, eingeschränkt. Das führt zur Reduktion von Freiheitsgraden im hochdimensionalen Problem, was die Konvergenz des Optimierungsprozesses beschleunigt und die Wahrscheinlichkeit, zum globalen Extremum zu konvergieren erhöht.

Jedoch ein korrektes bio-physikalisches Transformationsmodell aufzustellen erweist sich bei der Fülle von unterschiedlichen Geweben im Körper und den möglichen Freiheitsgraden der Bewegung als nicht einfach. Auch das rechen-technische Lösen eines solchen Systems, wie im Rahmen von FEM Simulation angestrebt wird, benötigt viel zu lange Rechenzeiten und große -kapazitäten. In der Registrierung haben sich daher verschiedene mathematische Hilfsmodelle entwickelt, die nicht mehr den Anspruch der korrekten Modellierung erfüllen, sondern im Wesentlichen eine glaubhafte Verformung nachahmen. Als glaubhaft wird die Transformation bezeichnet, wenn sie keine auffälligen Singularitäten und Unstetigkeiten enthält, sowie eher rigide Organe nicht unrealistisch beliebig verzerrt. Eine diffeomorphe Abbildung soll die Strukturen räumlich intakt belassen und die reale Verformung annähern. Dieses Herangehen ist im Wesentlichen für das Validierungsproblem von Registrierungsergebnissen verantwortlich, da meist die unterschiedlichen Parametrisierungen der mathematischen Modelle nicht eindeutig bio-physikalischen Eigenschaften zugeordnet und damit keinem Vergleich zu empirischen Messungen unterzogen werden können.

In der klinischen Routine hat sich bisher nur das Transformationsmodell des rigiden Körpers durchgesetzt. Ausgehend von Bestrahlungen des rigoros fixierten Schädels, wo die rigide Registrierung die für die Strahlentherapie benötigte Genauigkeit erreicht, setzt man sie auch als erste Näherung der Transformation bei Bestrahlungen anderer extra-cranieller Lokalisationen ein. Obgleich man in der Regel keine Verformungen der Patienten-anatomie

durch Positionierungskorrektur ausgleichen kann, kann man dennoch mithilfe rigider Registrierung den systematischen Positionierungsfehler z.B. in der fraktionierten Therapie durch eine entsprechende ebenfalls rigide Tischverschiebung reduzieren.

Das rigide Transformationsmodell ist trotz des fehlenden Goldstandards in der Registrierung wegen der starken Konditionierung leichter nachvollziehbar und dadurch als eine Näherung auch glaubhaft. Schwieriger wird es beim Übergang zu verformbaren Modellen. Erste Ansätze erlaubten ein sogenanntes Free Form Deformation (FFD) Modell ohne jeglicher plausibler Nebenbedingungen. Während ein solches Registrierungsverfahren die Konturen eines Apfel auf die einer Birne perfekt mappen kann, möchte man im Falle eines Patienten unrealistische Transformationen aus dem Suchraum ausschließen. Dazu führt man in der Regel weitere Regularisierungsterme ins Transformationsmodell ein, die jedoch meist auch nicht bio-physikalische Eigenschaften nachbilden, sondern globale Glattheitsnebenbedingungen erzwingen.

Da keines solcher mathematischen Hilfsmodelle einen eindeutigen Vorteil mit sich bringt, haben sich viele Modifikationen je nach spezialisierter Fragestellung entwickelt. Im Folgenden sollen einige der verbreiteten Transformationsmodelle, die in Registrierungsverfahren zum Einsatz kommen, beschrieben werden.

### Von rigide...

Das rigide Transformationsmodell besteht aus der Annahme, dass die Distanzen zwischen jeweils allen Massenpunkten eines festen Körpers bei beliebiger Transformation erhalten bleiben. Mit dieser Bedingung lassen sich die  $3N$  Freiheitsgrade bei  $N$  Massenpunkten auf 6 Freiheitsgrade reduzieren. Der Suchraum im Optimierungsprozess wird somit immens verkleinert und die Transformation selbst ist leicht nachvollziehbar und bei möglichst rigide fixierten Patienten in erster Näherung realistisch.

Die 6 Freiheitsgrade bilden die 3 Translationen des Schwer- bzw. Bezugspunkts entlang der drei Raumachsen und die 3 Rotationen des rigiden Körpers um die entsprechenden Raumachsen durch den Bezugspunkt. Mithilfe der homogenen Koordinaten lässt sich die Gesamttransformation durch eine  $4 \times 4$ -Matrix-Vektor-Multiplikation formulieren:

$$T_{rigide}(\vec{v}) = \begin{pmatrix} v'_x \\ v'_y \\ v'_z \\ 1 \end{pmatrix} = \left( \begin{array}{ccc|c} r_{00} & r_{01} & r_{02} & t_x \\ r_{10} & r_{11} & r_{12} & t_y \\ r_{20} & r_{21} & r_{22} & t_z \\ \hline 0 & 0 & 0 & 1 \end{array} \right) \begin{pmatrix} v_x \\ v_y \\ v_z \\ 1 \end{pmatrix}. \quad (2.6)$$

Die Translation  $\vec{t}$  steht hierbei in der vierten Spalte. Die vierte Reihe wird um Null-elemente und eine 1 in der Diagonale erweitert. Im übrigen  $3 \times 3$  Block befindet sich die Rotationsmatrix  $R = (r_{ij}) = R_x(\alpha)R_y(\beta)R_z(\gamma)$ , die wiederum das Produkt der einzelnen Rotationsmatrizen um die 3 Raumachsen  $\vec{e}_x, \vec{e}_y$  und  $\vec{e}_z$  ist.

Die rigide Registrierung lässt sich im Rahmen der Strahlentherapie gut mithilfe von Phantomexperimenten validieren [62]. Im Schädelbereich erreicht sie hohe Präzession [31]. Weiterhin wird sie eingesetzt, um systematische offset Fehler bei der Positionierung zu minimieren, auch wenn Deformationen im Bildern vorhanden sind. Um die Präzision der adaptiven Therapie zu steigern benötigt man allerdings deformierbare Transformationsmodelle.

### ... zu deformierbar

Im Folgenden sollen einige der am häufigsten eingesetzten Transformationsmodelle für deformierbare Registrierung vorgestellt werden. In der Regel lassen sich alle Modelle weiter spezialisieren und zu beliebigen Variationen erweitern. Hier soll lediglich das Augenmerk auf den Grundgedanken gelegt werden, um den Eindruck zu vermitteln, wie Verformungen der Eingangsbilder zustande kommen.

Die jüngeren Entwicklungen in der Registrierung zielen häufig darauf, mehr realistische Nebenbedingungen in die bereits vorhandenen Modelle einzubauen. Dennoch werden meistens nicht bio-physikalisch motivierte Bedingungen formuliert, sondern Regularisierungsterme, die meist globale Glattheit erzwingen.

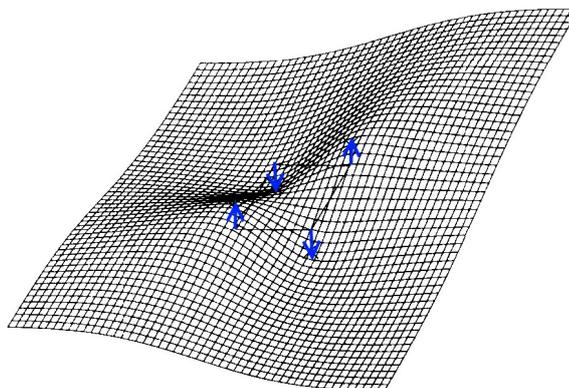
**Free Form Deformation** ist der einfachste Ansatz um eine Raumverformung abzubilden [63, 64]. Ein meist reguläres Kontrollpunktgitter wird verformt, indem man alle Kontrollpunkte uneingeschränkt im Raum bewegen kann. Die restlichen Raumpunkte werden dann mithilfe mathematischer Funktionen zwischen den Kontrollpunkten interpoliert. Hierbei werden meist polynomische Interpolationen eingesetzt. Die FFD zeichnet sich im wesentlichen durch das Fehlen von Regularisierungstermen aus. Meist tragen alle resultierende Grauwerte des verformten Bildes nach der Transformation zum Ähnlichkeitsmaß bei. Die Stärke des Modells liegt in der Fähigkeit eine sehr lokale Deformation abzubilden und der kurzen Rechendauer, die für die lokale Interpolation auf einem regulären Gitter benötigt wird. Nachteilig ist das Fehlen von Nebenbedingungen, dadurch werden beliebige Deformationen und damit auch nicht glaubhafte verzerrte Lösungen zugelassen.

**Radial Basis Functions und Thin Plate Splines** (RBF und TPS) bilden eine beliebte Klasse von Verformungsmodellen [65–67]. Hier nutzt man RBFs, um die Transformation zwischen den Kontrollpunkten zu interpolieren. Die Kontrollpunkte können hier irregulär im Raum verteilt sein, da die Interpolation an jedem Punkt nur vom Abstandsradius zu den Kontrollpunkten abhängt. Die Interpolation zwischen den Stützstellen ergibt sich dann mit

$$f(x) = a_1 + a_2x + \sum w_i U(|P_i - x|), \quad (2.7)$$

mit  $a_j$  und  $w_i$  als Interpolationskoeffizienten, die durch die Bedingungen der bekannten Stützstellen  $P_i$  ermittelt werden. Hierbei ist ein quadratisches lineares Gleichungssystem zu lösen, das mit der Zahl der bekannten Stützstellen wächst. Die Koeffizienten  $a_j$  bestimmen den rigiden Transformationsanteil. Das Potential  $U(r = |P_i - x|)$  kann unterschiedliche Formen annehmen. Ein immer lösbares Gleichungssystem entsteht beim Einsatz von

**Abbildung 2.7:** Schematische Darstellung der Auswirkung von der TPS Interpolation. Die vier symmetrisch angeordneten Stützstellen sind als blaue Pfeile markiert. Abbildung entnommen aus Bookstein *et. al.* [65]. Die blaue Pfeilmarkierungen sind zur besseren Sichtbarkeit nachträglich eingefügt worden.



Hardy's Multiquadratics [68]

$$U(r) = \sqrt{r^2 + c}. \quad (2.8)$$

Ein beliebiger alternativer Spezialfall ist die TPS, die das Potential von dünnen Metallplättchen modelliert. Hierbei wird das Biegepotential der Form

$$-U(r) = -r^2 \log r \quad (2.9)$$

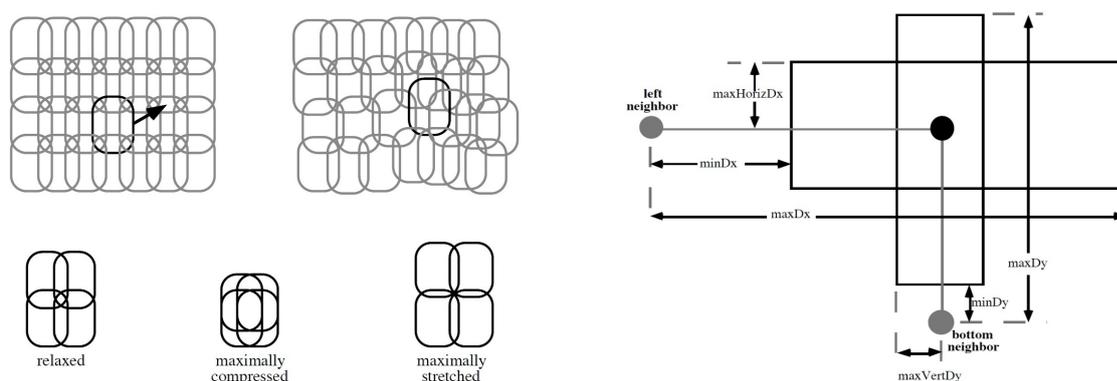
angesetzt. Das Verformungsverhalten ist in Abbildung 2.7 anhand von vier symmetrischen Stützstellen dargestellt. Bei dieser Klasse handelt es sich meist um globale Interpolationen. Damit ist der Schwachpunkt die relativ lange Berechnungsdauer verglichen zu lokalen Methoden. Dagegen ergeben sich dadurch aber glattere Deformationsfelder, die ebenfalls – anders als bei lokalen Methoden – einen rigiden Transformationsanteil abbilden können.

**Fluid-basiertes Verfahren** ist ein nicht parametrischer Ansatz in der Registrierung mit der Motivation besonders große auftretende Verformungen im Patientenmodell darzustellen [69,70]. Die Verformung wird analog zu einer Navier-Poisson Newtonschen Flüssigkeit modelliert, die erlaubt, die durch die Verformung aufgebaute Spannung mit der Zeit wieder abzubauen, anders als bei linearen elastischen Modellen nach Hookschem Gesetz. Eine für kleine Reynoldszahlen vereinfachte Navier-Stokes Gleichung wird häufig angesetzt:

$$\mu \nabla^2 \vec{v} + (\lambda + \mu) \vec{\nabla}(\vec{\nabla} \cdot \vec{v}) + \vec{F}(\vec{u}) = \vec{0}. \quad (2.10)$$

Der letzte Term  $\vec{F}(\vec{u})$  beschreibt die Körperkräfte, die den Bildinhalt des verformten Bildes zum stationären leiten. Der erste Term  $\mu \nabla^2 \vec{v}$  beschreibt den zähen Fluss mit konstanten Volumen, während der zweite Term mit der Divergenz nur in den Regionen nicht verschwindet, die wachsen oder schrumpfen. Die Koeffizienten  $\lambda$  und  $\mu$  kontrollieren die Scherspannung und den lokalen Volumenänderungsbetrag. Die Lösung der partiellen Differentialgleichungen bei diesem Transformationsmodell benötigt in der Regel längere Rechenzeiten als die zuvor vorgestellten parametrischen Modelle, liefert aber glatte Vektorfelder auch für große Verformungen. Trotz des physikalischen Modells für die Transformation beschreibt die Navier-Stokes Gleichung das Verhalten von zähen Flüssigkeiten und nicht das des biologischen Gewebes.

**Finite Elemente Modell** zählt nicht mehr zu den Registrierungsmethoden, sondern zu der Simulation der Verformung, da diese nicht als ein Optimierungsprozess abgebildet wird [71]. Die Grenzen sind jedoch fließend, da es viele Ansätze gibt FE Modelle in die Registrierung und umgekehrt mitaufzunehmen. Die finite Elemente Methode ist eine Strategie um partielle Differentialgleichungen (PDE), wie z.B. im Falle der fluiden Registrierung zu lösen, indem eine Lösungsanäherung in finiten Elementen angestrebt wird, die den kontinuierlichen Raum mithilfe eines Meshes in diskrete Domänen aufteilen. Die PDEs können so in äquivalente normale Differentialgleichungen umgewandelt und gelöst werden. Die Vorteile der Methode sind die Flexibilität bezüglich des durch das Mesh abgebildeten Modells und die numerische Stabilität des Verfahrens. Während man hier auch biophysikalische Modelle mithilfe der PDEs abbilden kann, hängt die Qualität der Lösung dennoch stark von der Güte des verwendeten Meshes ab. Weiterhin schränkt die Rechenzeit deren Anwendbarkeit in Routinesituationen ein, da diese Laufzeiten von mehreren Stunden aufweisen können.



**Abbildung 2.8:** Links: Schematische Darstellung des ChainMail Mechanismus, um Bewegungen durch eine 2D Fläche zu propagieren [72]. Links: Eine anfängliche Bewegung ist Ursache für die Deformation der Fläche. Unten sind die maximal erlaubte Kompression bzw. Expansion der Volumenelemente dargestellt. Rechts: Eine detaillierte Zeichnung des erlaubten Bewegungsraums für ein Element [72]. Betritt ein Nachbarelement diesen Raum, muß das betrachtete Element ausweichen, bis die Mindestanforderungen erfüllt sind, die durch die Kompressions- bzw. Expansionsparameter festgelegt sind.

**ChainMail** ist ein Verfahren aus der computergestützten Simulation zur Schulung von chirurgischen Eingriffen. Es wurde von Gibson im Mitsubishi Electric Research Lab entwickelt [72]. Der 3D ChainMail Algorithmus dient der schnellen Verformung eines volumetrischen Objektes aufgrund der Einwirkung einer Kraft. Da bei der Simulation von chirurgischen Eingriffen schnelle und einigermaßen realistische Deformationen aufgrund von interaktiven Aktionen modelliert werden müssen, wurde ein Algorithmus entwickelt, der zwar nicht bio-physikalischer Modellierung entspricht, jedoch über die Anwendung einfacher mathematischer Modelle in Fragen des Realismus hinausgeht. Der Algorithmus basiert auf der Simulation der Bewegungspropagation einzelner Volumenelemente, die aufgrund der Bewegung ihrer Nachbar in einem für sie definierten Bewegungsbereich ihre Position ändern. Die Abbildung 2.8 zeigt schematisch das Vorgehen. Hierbei werden zwei Parameter eingeführt, die sowohl die maximal erlaubte Kompression bzw. Expansion als auch die Scherspannung zwischen den Volumenelementen festlegen. Bewegt sich ein Nachbarelement entsprechend den Parameter zu nah oder zu weit vom betrachteten Element, muss dieses mit einer Bewegung reagieren. Sind die maximalen Parameter nicht überschritten, braucht keine Adaption der Elementposition zu erfolgen.

### 2.2.3.3 Optimierungsverfahren

Gewöhnlich wird die Registrierungsaufgabe als ein Optimierungsproblem wie in Gleichung 2.1 formuliert. Die Optimierung ist ein iterativer Prozess, bei dem der Parametersatz für die unterstellte Transformation  $T$  so variiert wird, dass die Werte der Zielfunktion – hier das Ähnlichkeitsmaß  $M$  der beiden Bilder – sich möglichst effizient dem globalen Funktionsextremum annähern. Für das rigide Transformationsmodell z.B. können die Parameter aus den drei Translationsvektoren und drei Rotationswinkeln bestehen, wie es in Gleichung 2.6 formuliert ist. Die Anzahl der freien Parameter bestimmt die Dimension des Optimierungsproblems. Die Dimension wiederum bestimmt den Rechenaufwand und Gutartigkeit der Optimierung. Bei der deformierbaren Registrierung wächst die Di-

mension mit der Anzahl der Stützstellen enorm, was sie zusammen mit dem fehlenden Goldstandard zu einem schlecht definierten Problem macht.

Es gibt unterschiedliche Klassen von Optimierungsstrategien. In der Registrierung hat man es in der Regel mit einem skalaren Optimierungsproblem zu tun. Ein weiteres Unterteilungskriterium ist, ob und welche Ableitung der Zielfunktion für die Methode benötigt wird. Ableitungsfreie Strategien sind meistens robuster gegenüber Unstetigkeiten der Zielfunktion und können auch bei nicht differenzierbaren Zielfunktionen eingesetzt werden. Einerseits ist die Berechnung innerhalb einer Iteration schneller, da keine Gradienten berechnet werden müssen, andererseits benötigen sie aber mehr Iterationsschritte, da das Entlanghangeln in Richtung des Gradienten meist den effizientesten Optimierungsansatz darstellt.

Im Falle der Registrierung von klinischen Bilddaten hat man es häufig mit nicht-konvexen Zielfunktionen zu tun. Das führt bei den meisten Optimierungsmethoden dazu, dass sie in lokalen statt globalen Extrema konvergieren können. Interpolationsartefakte können das Problem zusätzlich schon bei wenigen Freiheitsgraden verschlimmern [53]. Daher sollten Methoden eingesetzt werden, um dem entgegenzuwirken, was allerdings mit einem größerem Rechenaufwand verbunden ist und die bereits langen Rechenzeiten weiter verlängert.

Wie Optimierungsstrategien in Registrierungsverfahren aussehen können, soll anhand einiger Beispiele erläutert werden. Die Wahl der eingesetzten Optimierungsmethode hängt stark von der genutzten Zielfunktion ab.

**Das Simplex Verfahren** – oder auch Downhill-Simplex-Verfahren – ist ein beliebtes ableitungsfreies Optimierungsverfahren nach Nelder und Mead [73]. Der Algorithmus basiert auf einem Simplex, der das einfachste Volumen im  $N$ -dimensionalen Raum mit  $N + 1$  Punkten aufspannt. Im 2-dimensionalen Raum, einer Fläche, wäre das ein Dreieck, dessen Punkte jeweils einem Parametersatz entsprechen. Die Kostenfunktionswerte werden an allen Punkten bestimmt und verglichen. Die schlechtesten Punkte werden durch eine Reihe von Reflexionen, Expansionen oder Kontraktionen am Schwerpunkt des Dreiecks in jeder Iteration durch neue Parametersätze ersetzt, bis das Dreieck sich um das Extremum zusammenzieht. Das Verfahren benötigt zwar relativ viele Iterationen, ist aber einfach zu implementieren und robust. Weiterhin kann es leicht modifiziert werden, um das Risiko zu minimieren in lokalen Extrema zu konvergieren.

**Das Gradientenverfahren** benötigt die Berechnung der ersten Ableitung der Zielfunktion  $f(x)$  nach ihren Variablen in jeder Iteration  $i$

$$d^{(i)} = -\nabla f(x^{(i)}). \quad (2.11)$$

Ausgehend vom Startpunkt  $x^{(0)}$  wird der nächste Parametersatz mithilfe einer Schrittweite  $\alpha^{(i)}$  bestimmt

$$x^{(i+1)} = x^{(i)} + \alpha^{(i)} d^{(i)}. \quad (2.12)$$

Der Schrittweitenparameter  $\alpha^{(i)}$  muß ebenfalls in jeder Iteration neu bestimmt werden. Anschaulich bedeutet die Suchstrategie, dass man an jedem erreichten Punkt die Richtung des steilsten Abstiegs ermittelt, die Schrittweite meist an diesen anpasst und sich so iterativ dem Minimum der Zielfunktion entlang des Gradienten nähert. Der Schwachpunkt des Verfahrens kann sein, dass in der Nähe des Minimums viele Iterationen benötigt werden, da auch die Schrittweite mit dem Gradienten abnimmt. Da der Gradient bestimmt werden

muß, muß die Zielfunktion stetig differenzierbar sein. Um mit stark verrauschten Daten umgehen zu können, werden auch Gradientenannäherungsverfahren eingesetzt.

**Das Newton-Verfahren** ist ein beliebtes, schnell konvergierendes Optimierungsverfahren zur Bestimmung von Nullstellen von nicht-linearen Funktionen. Zur Berechnung wird die Zielfunktion lokal linearisiert, indem im 1-dimensionalen Raum die Funktion durch eine Tangente ersetzt wird. Hierzu ist eine Gradientenbestimmung notwendig. Die Nullstelle der Tangente wird als Näherung der Funktionsnullstelle bestimmt. Für diesen Parameter wird nun in der nächsten Iteration wiederum die Steigung der Funktion bestimmt und so weiter. Im Mehrdimensionalen wird der Gradient durch die Jacobi-Matrix ersetzt. Da man in der Registrierung aber keine Nullstellen, sondern Extrempunkte der Zielfunktion sucht, modifiziert sich das Verfahren insoweit, dass man die Nullstellen der Ableitung der Funktion bestimmen muß, da diese mit den Extrempunkten der Funktion verbunden sind. Damit benötigt man auch die zweite Ableitung der Zielfunktion, die im Mehrdimensionalen durch die Hesse-Matrix dargestellt wird. Die Berechnung der Hesse-Matrix ist rechenintensiv, so dass Verfahren entwickelt wurden, die unter bestimmten Bedingungen die Hesse-Matrix ohne explizite Berechnung annähern lassen. Besonders für quadratische Formen der Zielfunktion ist das so angenäherte Newton-Verfahren sehr schnell. Der Nachteil ist die geringe Robustheit des Verfahrens, wenn die Zielfunktion viele lokale Extrema aufweist.

## 2.2.4 Evaluierung und das Validierungsproblem

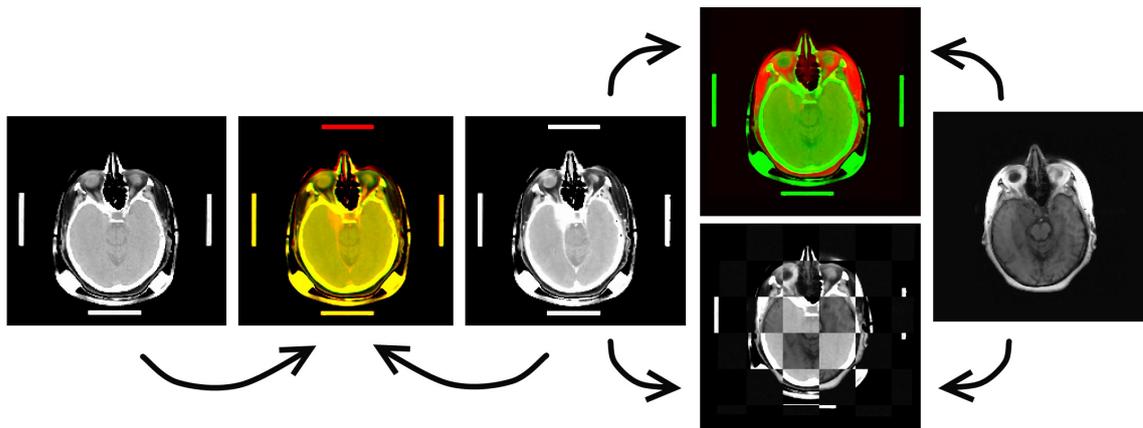
Die Validierung der Registrierungsalgorithmen steht heutzutage im Mittelpunkt aller betroffenen Anwendungen. Während rigide Registrierungsergebnisse visuell im Rahmen des klinisch Relevantem relativ einfach interpretierbar sind, entziehen sich die elastische Verformungen dem visuellen Empfinden und auch dem Erfahrungsempfinden des Experten. Häufig kann man nur indirekt Unterschiede im Ergebnisbild nach Verformung durch das Vektorfeld werten. Solche visuellen Überprüfungen sind heute der Standard bei vielen wissenschaftlichen Applikationen. Aber unterschiedliche Vektorfelder können zum selben Ergebnis führen, da man lediglich Konturen mit starken Kontrasten vergleichen kann.

### 2.2.4.1 Fehlende Validierung

Das führt dazu, dass Vergleiche zwischen unterschiedlichen deformierbaren Algorithmen schwer fallen. In der Literatur findet man kaum multi-institutionelle Vergleiche verschiedener Algorithmen [74–76] oder ganz selten Vergleiche zwischen unterschiedlichen Modifikationen desselben Verfahrens oder nur wenigen alternativen Algorithmen [77, 78]. Auch die Evaluierung der Verfahren ist nicht eindeutig standardisiert, so dass die Autoren meist individuelle von der Anwendung abhängige Kriterien anlegen. Auch wenn sich einige Hilfsmethoden – wie der Vergleich zu Expertenlandmarken – etabliert haben, so hängen sie doch vom visuellen Empfinden des Experten ab und geben meist nur Schätzungen der Unsicherheit an wenigen Positionen im Bild.

### 2.2.4.2 Evaluierungsmethoden um Unsicherheit abzuschätzen

Meistens werden mehrere Evaluierungsmethoden für die Ergebnisse der Registrierungsverfahren herangezogen. Da keine endgültig überzeugende Evaluierungsmethode existiert,



**Abbildung 2.9:** Bildfusion nach der erfolgten Registrierung: Das erste und mittlere sowie mittlere und letzte Bild in der Reihe werden zum zweiten und vorletzten Bild fusioniert (die jeweiligen Zusammenhänge sind durch die Pfeile visualisiert). Monomodale Bildfusion (zweites Bild) in der Rot-Grün Kombination ist leicht interpretierbar: rote und grüne Areale zeigen unterschiedliche Intensitäten im Bild an, während gleiche Intensitäten und damit Bildstrukturen in der Summe gelb erscheinen. Bei multimodaler Fusion (vorletztes Bild oben) ist die Deutung schwerer, da unterschiedliche Kontraste trotz guter Überlagerung rot oder grün erscheinen. In der Checkerboard Darstellung (vorletztes Bild unten) lassen sich die Grenzen der Strukturen zwischen den einzelnen Kästchen leichter auf Konsistenz prüfen.

behilft man sich damit das Ergebnis möglichst vielen Konsistenzprüfungen zu unterziehen. Dazu gehören visuelle Überprüfungen der Objektübereinstimmung, Vergleiche gegen gesetzte eindeutig wiedererkennbare anatomische Landmarken oder Vergleiche zu künstlichen Verformungen.

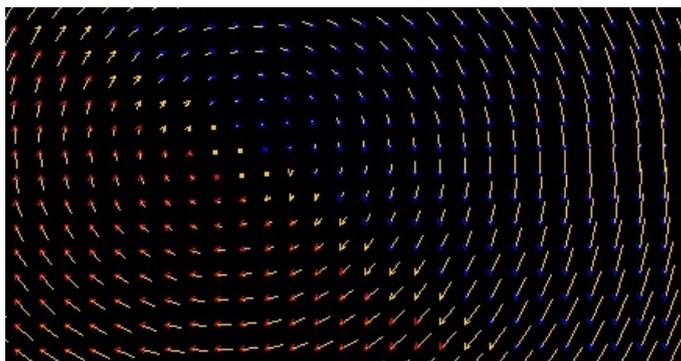
Besonders schwierig ist die Evaluierung von multimodalen Registrierungsverfahren. Hier ist selbst die visuelle Abschätzung oder Landmarkenidentifikation durch die unterschiedliche Grauwertverteilungen und unterschiedlichen Bildinhalt immens erschwert.

### Visuelle Hilfsmittel

Bei der Evaluierung dienen visuelle Hilfsmittel in erster Linie nicht der Fehlerschätzung des Registrierungsergebnisses, sondern der Vertrauensbildung zum Ergebnis. Da eine Validierung des Resultats wegen des fehlenden Goldstandards nicht möglich ist, will man visuell überprüfen, ob das Ergebnis der Verformung realistisch ist. Alleine können die visuellen Überprüfungen nicht zur Unsicherheitsschätzung genutzt werden, da eine Quantifizierung des Fehlers so nicht machbar ist. Einfacher ist es, hier eindeutige Abweichungen und unglaubhafte Deformationen überhaupt erst zu identifizieren. Auch hier gilt, dass man optisch nur Bereiche mit hohem Kontrast ob der Richtigkeit der Transformation einschätzen kann. Über Bereiche mit homogenem Bildinhalt kann man keine Aussage treffen, auch wenn die visuelle Interpolation glaubhaft ist.

**Differenzbilder** werden als die Differenz beider Intensitäten aus dem jeweiligen zu vergleichenden Bild Voxel für Voxel berechnet. Das neu entstehende Bild zeigt im Falle monomodaler Bildpaare wo Intensitätsunterschiede nicht verschwinden. Auch hier ist bei der Interpretation Vorsicht geboten, denn verschwindende Differenzen in homogenen Bildbereichen deuten nicht auf eine verifizierte Transformation hin. Wieder gilt: nur heterogene

**Abbildung 2.10:** Beispiel für die Visualisierung eines Vektorfeldes in der transversalen Schnittbilddarstellung. Die dargestellte Transformation ist eine rigide Rotation. Der Vektoranteil senkrecht zum Bild ist farblich kodiert: rot sind Anteile, die ins Bild hinein, und blau solche, die hinaus zeigen.

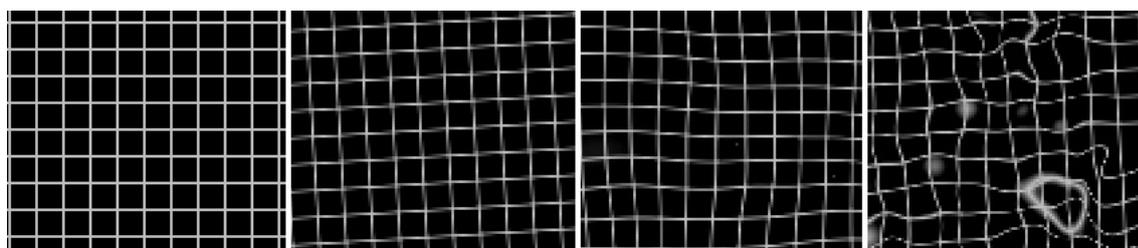


kontrastreiche Bereiche sollten betrachtet werden. Außerdem können die Beträge der Intensitätsdifferenzen als Maß für die Registrierungsfehler misgedeutet werden.

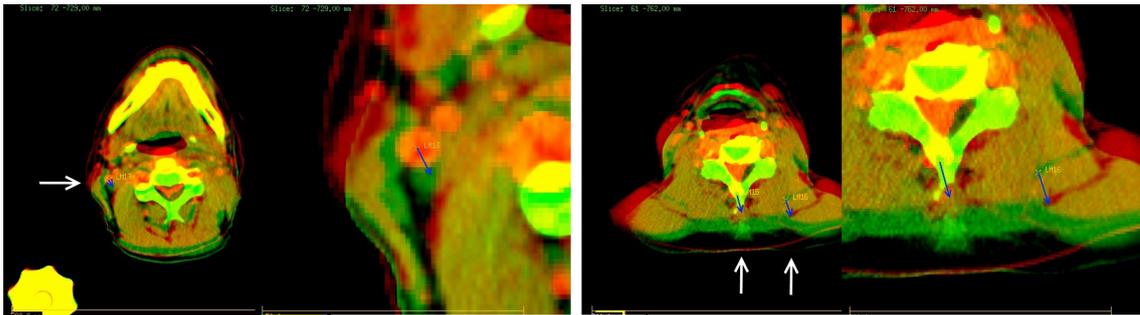
**Bildfusionierung** ist die Kombination zweier Bilder in eine Ansicht. Hierbei können bei monomodaler Bildfusion die beiden Bilder in unterschiedlichen Farbkanälen addiert dargestellt werden. Häufig werden z.B. rot-grün Fusionen wie in Abbildung 2.9 zur Ansicht eingesetzt. So erscheinen nicht übereinstimmende Bildbereiche rot und grün, während gleiche Intensitätsbeiträge aus den beiden Bildern gelb erscheinen. Beim Vergleich von Bildern verschiedener Modalitäten lohnt sich der Farbfusionsmodus nicht, da unterschiedliche Bildinformation an gleichen Positionen die Interpretation erschweren. Hier wird häufig das Schachbrettmuster genutzt, um schachbrettartig die Bildinformation beider Bilder abwechselnd in die Ansicht einzublenden.

**Vektorfeldvisualisierung** wird meistens durch Vektoren auf einem virtuellen Gitter im Bildraum realisiert. In den für Strahlentherapie üblichen Schnittbilddarstellungen kann man die einzelnen Projektionsbeiträge der Vektorkomponenten farblich kodieren, wie in Abbildung 2.10 zu sehen ist. In solchen Darstellungen lassen sich die Transformationen selbst visualisieren im Gegensatz zu Differenzbildern, wo die Bewegungsinformation fehlt. So kann die Transformation auf z.B. Glattheit oder Volumenänderungen überprüft werden.

**Verformung eines Gitters** kombiniert die beiden Möglichkeiten, einerseits die resultierende interpolierte Bildinformation und andererseits auch die Transformation selbst interpretierbar zu visualisieren. Hier wird anstelle des zu transformierenden Bildes eine regelmäßige Struktur – meist ein Gitter – aufgrund des Vektorfeldes deformiert. So lässt sich die Verformung besser interpretieren, da man die ursprüngliche Struktur sich gut merken



**Abbildung 2.11:** Visualisierung von Verformungsabbildungen anhand einer regelmäßigen Struktur. Von rechts nach links: Ausgangsstruktur, Transformationsergebnis nach einer rigiden Transformation, einer globalen und einer lokalen Deformation.



**Abbildung 2.12:** Beispiele für definierte Landmarken in der Schichtbilddarstellung im Bereich des Kopf-Halses im Fusionsmodus. Die Landmarken werden von den blauen Pfeilen repräsentiert. Die weißen Pfeile zeigen zur leichteren Auffindung auf die Landmarkpositionen. Das erste und dritte Bild zeigen unterschiedliche Schichten, während das zweite und vierte Bild die jeweilige Schicht in Vergrößerung zeigen.

kann und diese sich regelmäßig über den gesamten Bildraum verteilt. Anhand des verformten Gitters lässt sich die Stärke und Gutartigkeit (z.B. Glattheit) der Transformation darstellen. Die Abbildung 2.11 zeigt einige beispielhafte Transformationen anhand eines verformten Gitters.

### Vergleich mit künstlichen Transformationen

Die einfachste und sicherste Möglichkeit einen Goldstandard einzuführen ist, bekannte künstlich erzeugte Transformationen einzusetzen. Damit kann man am Ende der Registrierung die resultierende Transformation direkt Punkt für Punkt im gesamten Bildbereich vergleichen. Damit ist aber noch nicht bewiesen, dass das Verfahren bei beliebigen Transformationen die gleichen Leistungen erreicht. Erwünscht wird ein Verfahren, das einen individuell zutreffenden Fehler abschätzen kann, da Bilder unterschiedlicher Qualität und unterschiedlichen Inhalts nicht zu gleichwertigen Registrierungsergebnissen führen müssen. Die Methode hat den Nachteil, dass man erst nachweisen müßte, dass die künstlichen Verformungen den Bereich der tatsächlichen Verformungen im Patienten umfassen. Außerdem dürfen nicht die selben Transformationsmodelle zur Initiierung der Verformung genutzt werden, wie für die Suche im Registrierungsprozess.

### Vergleich zu Experten-Landmarken

Diese Variante gilt als eine Standardmethode [79], denn ohne die Nutzung von Registrierungsalgorithmen entscheidet nach wie vor der Experte anhand des Bildes was passiert sein kann. Dieser nutzt hierzu im Bild markante wieder zu erkennende Punkte, die als Landmarken bezeichnet werden. In Abbildung 2.12 sind solche Landmarken gezeigt. Die Nachteile hierbei sind der hohe Aufwand für das Identifizieren von Landmarken und ebenfalls die mögliche geringe Anzahl. Denn eindeutig erkennbare Punkte im 3D Tomographien sind nicht leicht zu finden, da wenige punktförmige Objekte im Körper vorhanden sind. Weiterhin ist die Reproduzierbarkeit stark abhängig von der Erfahrung des Experten und vorhandenen Protokollen.

### **Vergleich zu rigiden Teilregistrierungen: Grenzfallbetrachtung**

Wenn man knöcherne Strukturen als rigide betrachtet und sie dann auch als Teilstücke rigide registriert, setzt man aus einigen rigiden Bildbereichen eine quasi-elastische Verformung im Großen zusammen. Ein deformierbarer Registrierungsalgorithmus liefert nichts anderes als viele dichtliegende Bildkorrelationen als Ergebnis. Das Resultat des Registrierungsalgorithmus sollte also im Grenzfall von fallender Vektordichte gegen die resultierenden Translationsparameter der rigiden Teilregistrierungen konvergieren. Die Ergebnisse der rigiden Teilregistrierungen sind durch die eingeschränkte Zahl der Bewegungsfreiheitsgrade leichter auf ihre Richtigkeit zu prüfen.

### **Phantom Evaluierung**

Zur Verifizierung der rigiden Registrierungsverfahren haben sich Phantomexperimente glaubhaft einsetzen lassen. Verformbare Phantome sind allerdings rar und bilden meist nicht annähernd das Deformationsspektrum eines realen Patienten ab [80]. Die Messpunkte sind auch in einem Phantom meist nicht engmaschig verteilt, womit diese Methode eher mit der landmarkenbasierten Evaluierung vergleichbar ist. Internationale und -institutionelle Studien versuchen mithilfe von verformbaren Phantomen die Leistung der einzelnen Algorithmen vergleichbar zu quantifizieren [74–76, 81]. Beteiligung an solchen Studien ist erstrebenswert, kann jedoch nicht im Rahmen dieser Arbeit durchgeführt werden, da solche Studien über längere Zeiträume geplant werden.

### **Vergleich mit FEM Simulationen**

Ein vielversprechender Ansatz ist es, Registrierungsalgorithmen gegen die FEM Simulationsergebnisse von vorgegebenen Verformungen zu testen [82, 83]. Die FEM Simulation basiert auf bio-physikalischen Modellen des betrachteten Gewebes und ist damit die der wahren Transformation am nächsten kommende Beschreibung der Verformung von biologischer Materie, denn bio-physikalische Parameter lassen sich theoretisch aus messbaren Proben ableiten und mit Probandendaten auf Konsistenz testen. Doch auch FEM Simulationen benötigen einen Input an verformenden Kräften, die als Startbedingungen für die Simulation dienen, will man reale Patientendatensätze zum Vergleich heranziehen. Meistens bekommt man diese Information wiederum nur aus Registrierungsalgorithmen. Wie stabil die FE Modelle gegenüber den variablen Startbedingungen sind, bleibt noch zu zeigen.

### **Zusammenfassung**

Die Validierung von deformierbaren Registrierungsalgorithmen ist immer noch ein nicht gelöstes Problem. Trotz dazugehöriger Evaluierung der Resultate entscheidet letztendlich der visuelle Eindruck beim Anwender, ob das Ergebnis akzeptabel ist. Das kann mitunter zu einer Fehleinschätzung führen, vor allem wenn der Anwender sich nicht die Annahmen des Registrierungsverfahrens bewußt macht.

## 2.3 Hardwareparallelisierung

Die Einprozessor-Rechner der von-Neumann Architektur, die zu „Single Instruction, Single Data“ (SISD) Architekturen zählt und damit alle Rechenaufgaben seriell verarbeitet, wird heute immer mehr von parallelen Hardware-Architekturen ersetzt.

Die Gründe hierfür sind im steigenden Bedarf an sowohl Sicherheit als auch Rechengeschwindigkeit zu finden. Die Sicherheit vom Einsatz der parallelen Architekturen liegt in der Möglichkeit der Kreuzvalidierung von redundanten parallelen Rechensträngen. Bei ergebniskritischen Anwendungen wie der Steuerung von Chemigroßanlagen oder Energiekraftwerken oder auch beim Einsatz in der Medizintechnik können so fehlerhafte Berechnungen einzelner Rechner durch Vergleich mit übrigen Resultaten detektiert werden. Wollte man serielle Architekturen dafür einsetzen, wäre das mit Verlust an Rechengeschwindigkeit verbunden. Der Bedarf an Real-Time-Applikationen in diesem und anderen Bereichen ist jedoch auch hoch: denn was nützt einem die Detektion einer fehlerhaften Prozessentscheidung, wenn der Prozess, in den man eingreifen wollte, schon längst abgelaufen ist. Bei performancekritischen Anwendungen, wo es auf die Genauigkeit und Geschwindigkeit der benötigten Rechnungen ankommt, bieten unterschiedliche parallelen Architekturen eine vielversprechende Alternative zu seriellen Architekturen.

In der Bildverarbeitung, insbesondere in der deformierbaren Bildregistrierung, sind die Algorithmen häufig sehr rechenintensiv. Die hohe Auflösung der detaillierten medizinischen Aufnahmen und die zeitlich engmaschige Überwachung der dort enthaltenen Strukturinformation erfordert den Einsatz besonders schneller Registrierungsverfahren [84]. Dennoch darf die Anforderung an die Genauigkeit nicht gesenkt werden, im Gegenteil: die immer besser werdenden Abbildungstechniken liefern immer mehr therapeutisch relevante Details, die in immer kürzeren Zeitabständen verarbeitet und interpretiert werden müssen, um noch während der Therapie auf Abweichungen vom Plan eingehen zu können.

### 2.3.1 Parallele Architekturen

Es gibt verschiedene Technologien, die die Parallelisierung von Algorithmen ermöglichen. Unterschiedliche Hardware-Architekturen eignen sich für unterschiedliche Aufgabenstellungen unterschiedlich gut. Um die Eignung der Hardware für eine bestimmte Aufgabe zu ermitteln, müssen die Vorteile und Nachteile der Architektur abgewogen werden.

#### Hyperthreading

Hyperthreading ist eine Technologie in Intel Prozessoren, die es erlaubt auch einem Einkern-Prozessor mehrere Rechenstränge (Threads) zuzuweisen. Diese werden insoweit synchron verarbeitet, als dass die vorhandenen Hardware-Ressourcen effizienter aufgeteilt werden können. Während der laufende Thread meistens nicht die gesamten Ressourcen nutzt, da zwischen den Rechenschritten auch Daten aus dem Speicher gelesen werden müssen und es dabei zusätzlich zu Cache-Misses<sup>12</sup> kommen kann, werden Rechenteile eines zweiten Threads auf den so lange freien arithmetischen und logischen Einheiten verarbeitet und umgekehrt.

---

<sup>12</sup>falls die benötigten Daten nicht im vorgeschalteten Bereich liegen, verliert man zusätzliche Zeit zum Abfragen und Auslesen des langsameren Caches nächster Hierarchie.

Die Aufteilung und Überwachung der Ressourcen wird hierbei von System erledigt, so dass der Programmierer nur verschiedene Threads implementieren muß, um die Technologie zu nutzen. Diese Technologie wird in dieser oder ähnlicher Form mittlerweile auch in anderen Mehrkern- oder -prozessor-Architekturen eingesetzt.

### PC Cluster

PC Cluster oder Farmen sind Multiprozessorsysteme, die mehrere CPUs enthalten. Die einzelnen CPUs sind nicht auf einem Chip integriert, so können sie als eigenständige Einheiten betrachtet werden, die keinen gemeinsamen geteilten Speicher haben. Austausch von Daten zwischen den einzelnen CPUs müssen über externe Verbindungen und Schnittstellen realisiert werden.

Mit diesem Parallelisierungskonzept lassen sich besonders gut unterschiedliche Aufgaben parallel abbilden, denn die einzelnen Recheneinheiten verfügen über eigene Ressourcen. Austausch von Information zwischen den Einheiten ist aber aufwändig und kostet durch die Auftrennung mehr Zeit. Ausserdem werden Ressourcen nicht so effizient genutzt, da kein gemeinsamer Zugriff auf einzelne Speicherressourcen möglich ist. Auch die Synchronisation einzelner Einheiten bringt einen zusätzlichen Aufwand, da der Stand jeder Einheit zunächst abgefragt werden muß. Einen weiteren Nachteil bildet der Platzverbrauch, denn hier hat man meist mehrere Rechner oder Racks unterzubringen und zu kühlen, was eigens für diesen Zweck gestellte Räume erfordert.

### Multi-core CPUs

Mehrkern CPUs unterscheiden sich von Multiprozessorsystemen dadurch, dass mehrere CPUs auf dem selben Chip integriert sind. Dementsprechend ist es möglich ihnen gemeinsame Speicherressourcen zur Verfügung zu stellen. So teilen sie sich den Cache und die Speicherverwaltungseinheit. Die Kommunikationswege sind kürzer und damit kann der Austausch von Daten gegenüber Multiprozessorsystemen beschleunigt werden.

Durch die integriertere Technologie hat der Programmierer mehr Flexibilität, denn die Aufteilung einer Aufgabe in mehrere parallele Teile kann nun realisiert werden. Die Programmierschnittstelle wird dadurch jedoch komplexer. Auf einen PC Chip können heute bis zu acht CPU Kerne realisiert werden.

### FPGAs

Field Programmable Gate Arrays (FPGAs) sind integrierte Schaltkreise, in denen eine Rechenaufgabe als logische Schaltung in der Hardware selbst programmiert werden kann. Die Hardware ist konfigurierbar, so dass unterschiedliche Rechenschritte ohne Hardwareänderung modelliert werden können. Jedoch verfügen die FPGAs nicht über ein Betriebssystem, so dass sie für hochspezialisierte Anwendungen immer neu konzipiert werden.

Durch die Flexibilität bei der Konzipierung können *per se* parallele Architekturen aufgebaut werden, die nicht um Ressourcen mit einander konkurrieren müssen. So bleibt der Overhead durch die Synchronisation verschiedener Teilaufgaben oder Daten erspart.

FPGAs werden häufig bei rechenintensiven hochparallelisierbaren Aufgaben eingesetzt. In der Bildverarbeitung und -registrierung ist die Rechengeschwindigkeit der FPGAs

durch eine andere Architektur kaum zu übertreffen [85]. Dennoch sind das so spezielle Bauteile, dass sie nur eine geringe Verbreitung haben. Abgesehen von verbauten FPGAs in z.B. Bildgebungsgeräten, stellt eine extra für eine Aufgabe spezialisierte Hardware für Nutzer eher eine Verschwendung von Platz und Anschaffungskosten dar.

## GPUs

Die graphics processing unit (GPU) wurde ursprünglich im Gegensatz zur CPU entwickelt, um die 3D Visualisierung und das Rendern von Bildern zu optimieren. Während also die CPU zu einem Alleskönnen avancierte, konnte sich die GPU komplett auf eine einzige Aufgabe fokussiert entwickeln. Diese Aufgabe bestand hauptsächlich darin eine große Menge Datenpunkte möglichst schnell nach vorgegebenen Instruktionen zu berechnen. Die GPU Architektur entspricht dem „Single Instruction, Multiple Data“ (SIMD) Parallelisierungskonzept. Hoher Datenthroughput und die massive Datenparallelität waren und sind das Hauptanliegen des GPU Designs, was zur Entwicklung heutiger Multikern Prozessoren führte. Die GPU Architektur widmet mehr Transistoren dem Rechenvorgang und verzichtet auf die anspruchsvolle Instruktionsverarbeitung sowie die Cache-Elemente [86].

Eigene Transistoreinheiten konnten eingesetzt werden, um einzelne Aufgaben in der graphischen Pipeline zu übernehmen. Damit konnten spezialisierte Aufgaben der Visualisierung effizient und schnell berechnet werden. Doch Weiterentwicklungen der neuen GPU Architektur ermöglichen die Erweiterung der abbildbaren Aufgaben. Die General Purpose calculation using GPUs (GPGPU<sup>13</sup>) entwickelte sich schnell, da auch nicht graphische Aufgaben durch die Parallelisierung an Rechengeschwindigkeit gewinnen konnten.

### 2.3.2 Grundlagen der GPGPU

Auch seit der Entwicklung von GPGPU hat sich viel getan. Technische Neuerungen wie die Öffnung der Programmierschnittstelle, ermöglichen einen breiteren und kosteneffizienteren Einsatz. NVIDIAs Eintritt in den High Performance Markt treibt die Entwicklung weiterer Rechenleistung an. Vor allem rechenintensive Aufgaben profitieren von der Architektur und der Beschleunigung, die sich erreichen lässt. Einige Beispiele für solche Anwendungen werden auf der Homepage des Chipherstellers vorgestellt [88].

Für zeitkritische mathematischen Berechnungen, die in selben Instruktionen auf einer großen Menge Daten stattfinden sollten, wurden die GPUs schon früh genutzt. Doch die Programmierschnittstellen (APIs<sup>14</sup>) erlaubten keinen einfachen Zugang um diese Rechenkapazität auszunutzen. Die vorhandenen APIs waren in der Regel dem gleichen Ziel wie die GPU selbst geweiht, um die Visualisierung und das Rendern zu ermöglichen. Das Mappen von beliebigen Algorithmen, die nichts mit der graphischen Pipeline zu tun hatten und sich nicht intuitiv in diese einbetten ließen, sowie das Umgehen von zweckgebundenen Limitierungen, stellte eine Kunst für sich dar und verhinderte einen breiteren Einsatz der GPU als Rechenprozessor. Entwicklungen im Forschungsbereich ermöglichten vereinfachte Programmierkonzepte und -schnittstellen [89, 90], erreichten jedoch keine nennenswerte Verbreitung außerhalb der Forschung.

Die GPGPU erlebte einen neuen Aufschwung mit NVIDIAs Neuausrichtung der Hard-

---

<sup>13</sup>Information zur und um die GPGPU können auf der Homepage der Dachorganisation für die GPGPU-interessierte Gemeinschaft gefunden werden [87].

<sup>14</sup>engl. Application Programming Interface

warearchitektur CUDA<sup>15</sup>. Mit einer geeigneten, leicht zu erlernenden Programmierschnittstelle und abgeänderten Architektur gelang es die Popularität der GPGPU fast schlagartig zu steigern. Viele Algorithmen wurden in kurzer Zeit auf die GPU mittels CUDA erfolgreich portiert und zeigten einen immensen Speed-up Faktor gegenüber den alten CPU Implementierungen [91, 92]. Auch Algorithmen aus der Bildverarbeitung sowie Bildregistrierung wurden parallelisiert und zeigen viel versprechende Resultate [93–100].

Doch auch wenn große Teile der meisten Algorithmen sich gut für die SIMD Architektur parallelisieren lassen, gibt es immer auch andere Teile im Algorithmus, die das Beschleunigungspotential limitieren. Da die meisten Algorithmen ursprünglich für die CPU konzipiert waren, sind sie stellenweise für die Portierung auf die GPU ungeeignet. Wünschenswert sind daher speziell für die Ausführung auf der GPU zugeschnittene Algorithmen, die schlecht parallelisierbare Teilaufgaben vermeiden, oder zu mindestens so auslagern, dass die Hauptlast des Datenthrouputs ungehindert bearbeitet werden kann.

### 2.3.2.1 CUDA Programmiermodell

Die Hauptaspekte des Programmiermodells der CUDA Architektur sind die Aufteilung der Rechenaufgabe auf die unabhängigen Threadsgruppierungen, die Ausnutzung des gemeinsamen Speichers für kooperierende Threads innerhalb der Gruppe und der Synchronisationsmechanismus, um divergierende Rechenpfade wieder zusammen zu führen. Das ermöglicht eine Skalierbarkeit aller CUDA Programme auf beliebige CUDA-fähigen Hardware, unabhängig von der Anzahl der physikalisch vorhandenen Prozessoren.

Die parallel ausgeführten Threads sind in Gruppen zu sogenannten Warps<sup>16</sup> zusammengefasst. Diese sind die Einheiten, die unterschiedlichen Prozessoren in einem Multiprozessor zur Ausführung zugeteilt werden. Warps wiederum werden zu Blöcken (Block) zusammengefügt, die verschiedenen Multiprozessoren zugewiesen werden können. Die Blöcke wiederum sind in einem Grid organisiert. Das heißt, daß man je nach Algorithmus, die Rechenaufgaben auch auf diese Struktur abbilden muß. Hierbei ist zu beachten, daß kooperierende Stränge unbedingt im selben Block anzuordnen sind.

Die Aufgabenverteilung muß nach dem SIMD Konzept stattfinden. Nur Stränge, die den selben Instruktionsset zugewiesen bekommen, können tatsächlich parallel ausgeführt werden.

### 2.3.2.2 Algorithmus Mapping

Beim Mappen der Rechenschritte auf die GPU ist die Einhaltung des SIMD Prinzips im Parallelisierungsschritt wesentlich. Alle datenintensiven Teile, die klassisch in einer Schleifenstruktur berechnet werden, lassen sich in der Regel auf die GPU auslagern. Wesentlich ist, dass die Datenelemente unabhängig von einander und mit gleichen Instruktionen berechnet werden.

---

<sup>15</sup>engl. Compute Unified Device Architecture

<sup>16</sup>engl. [textil.] Webkette; Der Begriff ist aus der Webkunst entliehen - der ersten Technologie zur parallelen Strang- (Thread) Verarbeitung [86].

### Warps, Blocks, Grids - Kernel Konfiguration

Die eigentliche Parallelisierung der Rechenaufgabe besteht darin die notwendigen Rechenschritte auf das CUDA Modell zu transferieren. Prinzipiell gibt es immer verschiedene Möglichkeiten eine Aufgabe auf das Modell zu mappen. Erfüllt ein Algorithmus nicht perfekt die Modellanforderungen, läßt sich nur empirisch ermitteln, welches Mapping zu einer performanteren Lösung führt.

Ein einfaches Beispiel soll die wählbare Konfiguration veranschaulichen. Die Aufgabe soll ein Mittelwertfilter in 2D sein. Die Parallelisierung: jeder Ausgabewert wird von einem Strang berechnet. Das Schreiben des Endergebnisses auf den globalen Speicher erfüllt die Bedingungen für die parallele Ausführung, da jedes nachfolgende Element von einem nachfolgenden Strang geschrieben wird, wenn build-in Typen verwendet werden. Das Lesen von der notwendigen Information am Anfang sollte allerdings nicht direkt vom globalen Speicher stattfinden, da mindestens  $3 \times 3$  Lesezugriffe per Thread notwendig wären. Stattdessen werden mehrere Threads zu Blöcken gruppiert, die Synergieeffekte beim Lesen vom globalen Speicher ausnutzen können. Zwei benachbarte Threads z.B. benötigen die gleichen 6 von 9 benötigten Datenwerten. Zusammen müssten sie also keine  $3 \times 3 \times 2 = 18$  Lesezugriffe machen, sondern nur 12. Gruppiert man also immer mehr Threads zusammen kann man immer mehr Zugriffe auf den globalen Speicher vermeiden. Da sie aber nur über den schnellen shared memory (wird im folgenden Unterkapitel erklärt) nach dem Lesen die Datenwerte austauschen können, müssen diese Threads im selben Block gruppiert werden, da ein anderer Block auf einen anderen Multiprozessor zugewiesen werden könnte. Ein Block ist aber limitiert auf maximal 512 Threads, was bedeutet, dass man z.B.  $22 \times 23$  Threads pro Block wählen könnte und die Griddimension, also die Anzahl der Blöcke durch die Bilddimensionen bestimmt wäre  $Griddim_x = Bilddim_x/22$  und  $Griddim_y = Bilddim_y/23$ . Die Identifikationsnummern (IDs) der Threads in einem Block können 1-, 2- oder 3-dimensional gewählt werden, während das Grid nur 1- oder 2-dimensional initialisiert werden kann. Hier ist ein ummappen von Block IDs nötig, um 3-dimensionale Bildaten gezielt auszulesen.

### Speicher und Speicherzugriffsmuster

Die GPUs verfügen über mehrere unterschiedliche Speichertypen, die für unterschiedliche Aufgaben optimiert sind. In der Regel ist der schnellste Speicher auch der kleinste, so dass geschickte Lesestrategien notwendig sind, um die benötigten Daten im Speicher zu arrangieren. Speicherzugriffsmuster sind hierbei genau so wichtig, um eine optimale Parallelisierung der Berechnungen zu erreichen.

**Global memory** - der globale Speicher - ist der langsamste Speichertyp, aber auch der einzige, auf den vom CPU Thread zugegriffen werden kann. Damit befinden sich alle zur Berechnung nötigen Daten zu Beginn der GPU Funktion, genannt Kernel, auf dem globalen Speicher. Er ist nicht gecacht und erlaubt allen Threads sowohl Lese- als auch Schreibzugriff. Da der Zugriff auf diesen Speicher besonders zeitintensiv ist, sollte man darauf achten, alle Datenzugriffe darauf auf das Minimum zu beschränken. Es ist prinzipiell keine direkte Synchronisation zwischen Threads verschiedener Blöcke über diesen Speicher möglich. Ein Ausnahme bieten die implementierten „atomic operations“, die allerdings erst auf Hardware der Version mit compute capability von 1.2 zur Verfügung stehen [86]. Verfügt man nicht über die entsprechende Hardwareversion, muß man die Synchroni-

sation über den CPU Thread - sogenannten Host Thread - implementieren. Damit die Parallelität der Lesezugriffe tatsächlich erhalten bleibt, müssen die Zugriffsmuster im Wesentlichen zwei Bedingungen genügen: das Alignment in 128-bit oder 64-bit Worten der verwendeten Datentypen und die Zugriffshierarchie der Threads auf die Speicheradressen. Bei build-in Typen ist die erste Bedingung automatisch erfüllt, während die Zugriffshierarchie so realisiert sein muss, dass nachfolgende Threads auch auf nacheinander folgende Speicheradressen zugreifen sollten. Einen so geordneten Zugriff nennt man koaleszent.

**Texture memory** - der Texturspeicher - ist ein speziell arrangierter Speicherbereich auf dem globalen Speicher, der einen hardware-optimierten Zugriff für linear angeordnete Daten, die 2D Verteilungen beschreiben, regelt. Dieser Speichertyp bietet allerdings nur Lesezugriff und kann im Wesentlichen für bilineare Dateninterpolationen effizient genutzt werden.

**Constant memory** - der konstante Speicher - ist ebenfalls Teil des globalen Speichers, der nur Lesezugriff bietet. Dieser Typ eignet sich hervorragend um einzelne Konstanten zu speichern, auf die alle Threads gleichzeitig Zugriff benötigen.

**Shared memory** - der geteilte Speicher - ist neben den eigentlichen Registereinheiten der schnellste physikalische Speicher. Er ist im Gegensatz zum globalen Speicher auf dem Graphikchip selbst positioniert. Er ist auf 16 KB pro Multiprozessor beschränkt, bietet aber das einzige Medium, um Daten zwischen den einzelnen Threads des selben Blocks auszutauschen, da alle Threads sowohl Schreib- als auch Lesezugriff haben und es zusätzlich einen Synchronisationsmechanismus gibt. Auch hier sind die Speicherzugriffsmuster für optimale Performance zu beachten [86].

**Local memory** - der lokale Speicher - ist ein spezieller Teil des globalen Speichers, der für jeden Thread fest lokalisiert ist und damit etwas schneller als der Zugriff auf den normalen Bereich des globalen Speichers gemanagt werden kann - jedoch immer noch viel langsamer als der Zugriff auf den shared memory. Dieser Speichertyp wird automatisch eingeteilt, falls Datenarrays in Registereinheiten allokiert werden, die die zur Verfügung stehende Registermenge überschreiten. Hier lohnt es sich oft, diese Datenarrays auf den shared memory auszulagern, falls möglich.

### Datentransfer

Für Performanceüberlegungen sind natürlich nicht nur die puren Rechenzeiten ausschlaggebend, sondern auch die benötigten Datentransfers auf den GPU Speicher. Man sollte stets versuchen alle benötigten Daten und Zwischenergebnisse auf dem globalen Speicher so lange es geht zu halten. Für großen Datenthrouput gibt es ebenfalls asynchrone Kernelstarts, was dem CPU Thread ermöglicht weiter zu laufen, während die GPU Threads noch rechnen, und somit Transferzeiten hinter Rechenzeiten zu verstecken. Mehrere kleine Transfers sind im Allgemeinen ungünstiger als zusammengefaßte größere Datenarrays zu transferieren.

## Kapitel 3

# Resultate

### Inhaltsangabe

---

<b>3.1 Konzeptionierung des Registrierungsverfahrens . . . . .</b>	<b>42</b>
3.1.1 Anforderungen . . . . .	42
3.1.2 Registrierungsverfahren . . . . .	42
3.1.3 Algorithmische Komponenten . . . . .	44
3.1.4 Zusammenfassung des Registrierungsalgorithmus . . . . .	51
<b>3.2 Evaluierung der Registrierungsergebnisse . . . . .</b>	<b>52</b>
3.2.1 Vergleich mit künstlichen Deformationen . . . . .	52
3.2.2 Vergleich zu Landmarken und rigiden Teilregistrierungen . . . . .	55
3.2.3 Laufzeitperformance . . . . .	58
3.2.4 Limitierungen . . . . .	59
<b>3.3 Anwendung auf Patientendatensätze . . . . .</b>	<b>62</b>
3.3.1 Vergleich zur Verformungsanalyse alternativer Algorithmen . . . . .	62
3.3.2 Adaption von Patientenmodellen . . . . .	62

---

Die Anforderungen der heutigen ART an deformierbare Registrierungsalgorithmen sind sowohl bezüglich der Genauigkeit als auch der Rechengeschwindigkeit sehr hoch. Da mit der wachsenden Genauigkeit meist auch eine erhöhte Rechenintensität verbunden ist, muß Hardwareparallelisierung genutzt werden, um die Rechenzeit zu erhalten oder gar zu reduzieren. Um das volle Potential der Hardwareparallelisierung effizient nutzen zu können, benötigt man effizient parallelisierbare Algorithmuskomponenten und auf die Hardware zugeschnittene Registrierungsverfahren. Diese Arbeit trägt durch die Konzipierung von geeigneten algorithmischen Komponenten für die Ausführung auf einer meist bereits in PCs vorhandenen Hardware bei, die Registrierung zu beschleunigen. Damit soll der Einsatz von deformierbaren Registrierungsverfahren in der straff organisierten klinischen Routine erleichtert werden, um die Qualität der Behandlung durch die Strahlentherapie weiter zu steigern.

## 3.1 Konzeptionierung des Registrierungsverfahrens

### 3.1.1 Anforderungen

Um die Verformungs- und Bewegungsinformation während der Strahlentherapie für Adaptionen nutzen zu können, müssen diese möglichst schnell aus den vorhandenen Kontrollaufnahmen extrahiert werden. Das notwendige Zeitfenster für diese Aufgabe wird je nach Behandlungsstrategie entweder durch den möglichst kurzen, straff organisierten Fraktionsverlauf oder gar interne Bewegungsabläufe des Patienten wie beim Gating oder Tracking (siehe Kapitel 2.1.1.3) vorgegeben.

Um diesen Anforderungen gerecht zu werden, soll die GPU Technologie genutzt werden, um die erforderliche Rechengeschwindigkeit zu erreichen. GPUs verfügen über Multiprozessoren, die für bildverarbeitenden Aufgaben optimiert sind und erreichen gegenüber klein-skaligen Multikern-Prozessoren höhere Rechenraten. Gegenüber FPGAs sind sie leichter zugänglich, da nur weitverbreitete normale Graphikkarten, die meist bereits Teil der PCs sind, zur Ausführen benötigt werden.

Die Aufgabenstellung hat nicht einfach nur eine Beschleunigung durch die Parallelisierung zum Ziel, sondern auch die Extraktion möglichst vollständiger Information aus den vorhandenen Daten, um auch Genauigkeitsansprüchen gerecht zu werden. Somit liegt das Hauptaugenmerk bei der Konzeptionierung des Registrierungsverfahrens in der möglichst optimalen Abbildung der einzelnen Komponenten auf die genutzte GPU Hardware, ohne rechenintensive informationsverarbeitende Schritte zu vermeiden.

### 3.1.2 Registrierungsverfahren

#### 3.1.2.1 Aufbau

Als Ausgangsschritt für das Registrierungsverfahren wurde das bereits in die in-house entwickelte und eingesetzte Bestrahlungsplanungssoftware VIRTUOS<sup>1</sup> [101] implementierte Registrierungsverfahren mit dem Namen Template Matcher (TM) genutzt. Das Verfahren entspricht einer im Kapitel 2.2.1.2 beschriebenen Block-Matching Strategie, bei dem zunächst lokale Korrespondenzen mittels CC gesucht und nachfolgend als Stützpunkte für das TPS Modell zur Interpolation des kompletten Vektorfeldes genutzt werden [35].

Ein Vorteil des Verfahrens ist die Einschränkung der Freiheitsgrade im Optimierungsproblem der deformierbaren Registrierung. Da die rigiden Bildausschnitte unabhängig von einander bewegt und durch ein lokales Ähnlichkeitsmaß bewertet werden, ist der Suchraum der möglichen Lösungen stark eingeschränkt. Die Bewertung der einzelnen Ähnlichkeiten der Templates ist robust, denn die einzelnen Templatebewegungen können das Bewertungskriterium nicht gegenseitig kompensieren. Ebenfalls von Vorteil ist die Brute-Force-Suchstrategie in einem realistisch eingeschränkten Suchbereich, denn so entgeht man der sonst möglichen Konvergenz in ein lokales Extremum. Durch die schnelle Fouriertransformation läßt sich die Berechnung des CC Faltungintegral im Ortsraum in ein Integral über schneller zu berechnenden Multiplikationen im Frequenzraum umwandeln und so die Brute-Force-Suchstrategie in vernünftiger Zeit umsetzen. Ein weiterer Vorteil liegt in der globalen Interpolation durch die TPS, die auch einen rigiden Transformationsanteil enthält.

---

<sup>1</sup>engl. VIRTUal RadiOtherapy Simulator

Nachteile des Verfahrens sind in der geringen Anzahl der Korrespondenzen zu finden, die nur begrenzte Information über die stattgefundene anatomische Veränderung liefern. Die meisten lokalen Positionen des Vektorfeldes basieren auf einer mathematischen globalen Interpolation. Damit sind Deformationen auf kleinen Skalen nicht detektierbar. Weiterhin beeinflussen auch weit entfernte Korrespondenzen die lokale Transformation wegen der globalen Natur des Interpolationsverfahrens. Was für die Erfassung des rigiden Anteils gut ist, ist für lokale Deformationen nicht realistisch.

In der Konzeption des neuen Registrierungsverfahrens sollen Vorteile des letzteren möglichst beibehalten und die Nachteile kompensiert werden. Dafür wird das gleiche Aufbauprinzip genutzt:

- Zunächst wird eine Korrespondenzsuche durchgeführt, die möglichst viel Bewegungsinformation extrahieren soll.
- Nachfolgend, sollen die etablierten Korrespondenzen auf einen globalen rigiden Transformationsanteil untersucht und dieser gegebenenfalls abgetrennt werden.
- Der verbleibende Anteil mit der Verformung soll möglichst schnell und lokal für die Vektorfeldvervollständigung interpoliert werden.

### 3.1.2.2 Parallelisierungsaspekte

#### Zeitkritische Teile identifizieren

Nicht alle Teile eines Registrierungsalgorithmus lassen sich selbstverständlich auf eine parallele Architektur abbilden. Einige Schritte benötigen nach wie vor serielle Abarbeitung. Daher stellt sich die Frage, welche Algorithmusteile sich einerseits prinzipiell parallelisieren lassen und andererseits bei welchen sich eine Parallelisierung lohnt. Teile zu parallelisieren, die nur geringen Zeitaufwand des gesamten Registrierungsprozesses einnehmen, macht wenig Sinn. Statt dessen muß mit den rechenintensivsten Schritten, die auch den wesentlichen Zeitverbrauch bestimmen, angefangen werden. Die zeitaufwändigsten Schritte im deformierbaren Registrierungsverfahren sind gewöhnlich die Berechnung des Ähnlichkeitsmaßes und die Transformation des deformierten Templates, da das Maß unter vielen Transformationen im Optimierungsprozess neu berechnet werden muß. In unserem Ansatz erfolgt die Berechnung ohne einen Optimierer, sondern durch eine lokale Bestimmung des Ähnlichkeitsmaßes und der erschöpfenden Suche in einem begrenzten Bereich. Deshalb werden diese Schritte ebenfalls den rechenintensivsten Teil darstellen. Auch die Interpolation der vielen Korrespondenzen, die möglichst zahlreich aus den Bildern extrahiert werden sollen, kann je nach Schema einen weiteren zeitkritischen Schritt darstellen. Daher werden für diese Aufgaben Algorithmen konzipiert, die ein lokales Schema widerspiegeln und möglichst gut parallelisierbar sein sollen.

#### Design für effiziente Parallelisierung

Um eine effiziente Abbildung auf die SIMD Architektur der GPU zu ermöglichen, werden möglichst lokale von einander unabhängige Datenverarbeitungsschritte benötigt. Daher sollen lokale Ähnlichkeitsmaße und Interpolationsschemata eingesetzt werden. Um einen globalen Transformationsanteil zu bestimmen, müssen Bildinformationen aus dem

gesamten Bildbereich ausgetauscht werden, daher wird dieser Schritt der Korrespondenzsuche nachfolgend ausgelagert. Die Vektorfeldvervollständigung des Deformationsanteils der Transformation wird wieder mithilfe eines lokalen Schemas parallelisiert. Da die Vektorfeldvervollständigung in unserem Ansatz nicht mehr durch ein Ähnlichkeitsmaß bewertet wird, wird hier ein Methode benötigt, die möglichst realistische Verformung beschreibt und eine Möglichkeit bietet, lokale Materialinformationen in das Modell miteinfließen zu lassen.

Die meisten genutzten Ähnlichkeitsmaße, wie MI oder CC, bieten nur für große Bildbereiche eine ausreichende Stabilität [102, 103], da sonst der enthaltene Bildinhalt und die Strukturinformation in den Templates reduziert wird. Die Wahrscheinlichkeit, ähnliche Voxelnachbarschaftsverteilungen zu finden, steigt mit der abnehmenden individuellen Information an mehreren, falschen Positionen im Suchbereich an. Da für die effiziente Parallelisierung der Datenaustausch zwischen Voxelnachbarschaften möglichst minimieren werden sollte, benötigt man ein sehr lokales Ähnlichkeitsmaß.

Auch bei der Korrespondenzsuche in der definierten Suchregion für eine kleine Voxelnachbarschaft muß die Speicherlimitierung beachtet werden. Der schnellste Speicher auf der GPU, der bei diesem zeitkritischen Schritt eingesetzt werden sollte, ist auf 16 KB pro Multiprozessor eingeschränkt [86]. Das bedeutet, dass nur eine eingeschränkte Suchregion in einer Iteration abgesucht werden kann. Um aber auch größere Abweichungen in der Anatomie finden zu können wird ein multi-resolution Ansatz benötigt.

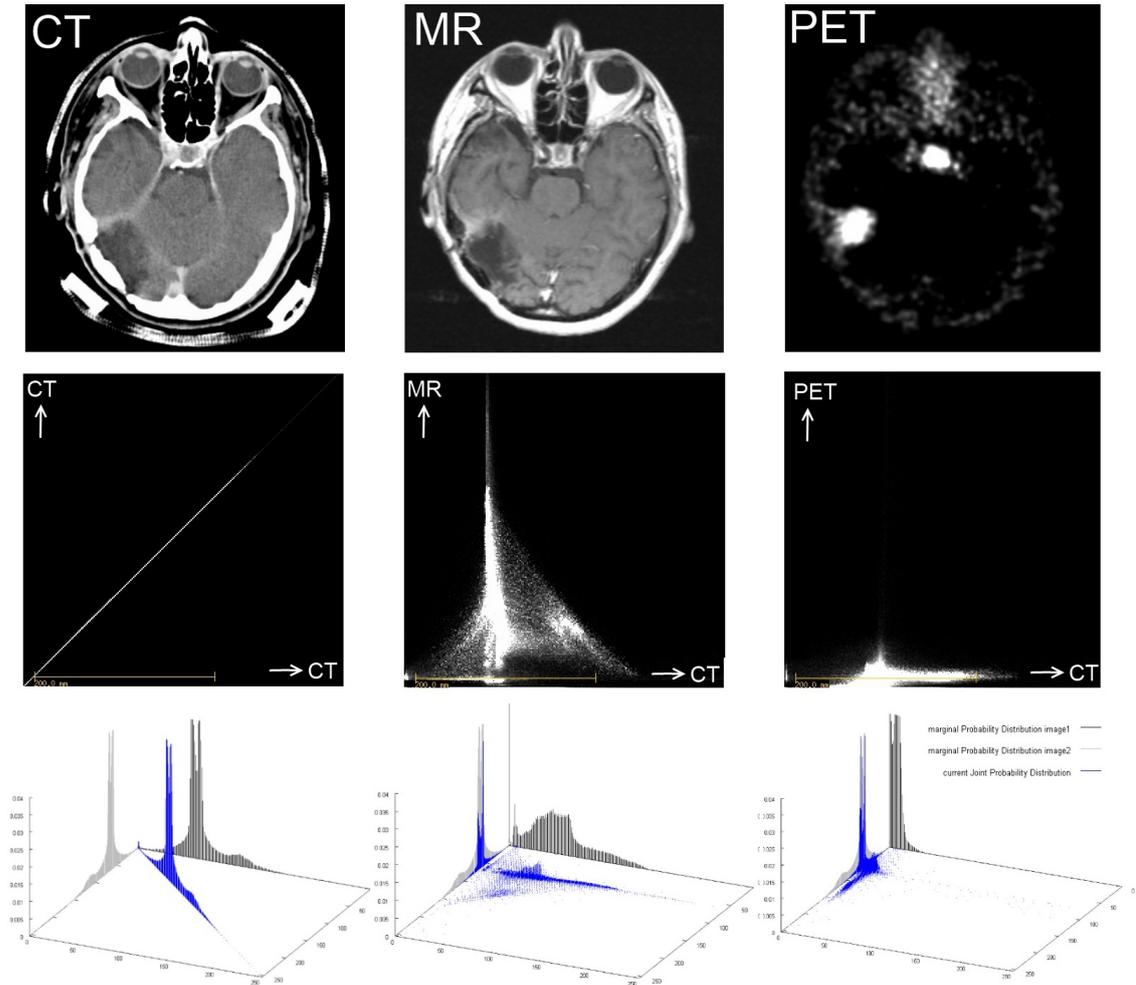
### 3.1.2.3 Multi-resolution Schema

Um die Einschränkung des Suchbereichs durch die lokale Suchstrategie und Hardwarelimitierungen zu überwinden, wird ein multi-resolution Schema mithilfe einer Gaußpyramide implementiert, wie in Kapitel 2.2.2 erläutert. Die Gaußpyramide wird nicht insgesamt dauerhaft gespeichert, sondern vor jeder Iteration wird die benötigte Gaußebene  $j$  aus dem originalen Template  $D$  erzeugt und anschließend mit der Transformation aus der letzten Iteration  $T^{(j-1)}$  in das Input-Template  $D'_{(j)}$  verformt. Die Iteration im Registrierungsprozess wird von der groben Auflösung der Bilder zur immer feiner aufgelösten Bildern durchlaufen, um das Vektorfeld nach und nach mit feineren Strukturen zu vervollständigen.

## 3.1.3 Algorithmische Komponenten

### 3.1.3.1 Ähnlichkeitsmaß

Das benötigte Ähnlichkeitsmaß muß für möglichst lokale Bildbereiche robust sein, um eine effiziente Abbildung auf die GPU Hardware zu erreichen. Ein weiteres Wunschkriterium ist, dass das Ähnlichkeitsmaß auch multimodale Bildpaare in ihrer räumlichen Übereinstimmung bewerten kann. Da bereits vorhandene Ähnlichkeitsmaße nur bedingt einige der Kriterien erfüllen, wird ein alternatives Ähnlichkeitsmaß auf der Basis von gemeinsamen Wahrscheinlichkeitsverteilung – ähnlich der MI – entwickelt. Bisher entwickelte JISP-basierte Maße (siehe Kapitel 2.2.3.1) benötigen ein Vorwissen über die tatsächliche Wahrscheinlichkeitsverteilung oder definieren Regressorstrukturen auf dem gemeinsamen Histogramm, was immer zusätzliche Vorverarbeitungsschritte benötigt. Außerdem wenden sie die entwickelten Ähnlichkeitsmaße nur in rigiden Registrierungsverfahren an.



**Abbildung 3.1:** Gemeinsame Histogramme von multimodalen Bildaufnahmen eines Schädels. Von links nach rechts: (CT,CT),(CT,MR<sub>T1</sub>) und (CT,PET) Paar. Oberste Zeile: Schnittbilder des genutzten Datensatzes. Mittlere Zeile: 2D Verteilung des gemeinsamen Histogramms. Untere Zeile: 3D Histogrammverteilung.

### Herleitung

Das gemeinsame Histogramm  $h_{r,d}$  beschreibt die Wahrscheinlichkeitsverteilung  $p_{r,d}$ , dass bestimmte Grauwertintensitäten  $I_R$  im Referenzbild  $R$  den Grauwertintensitäten  $I_{D'}$  im deformierten Template  $D'$  unter einer Transformation  $T$  entsprechen

$$p_T : (I_R(i), I_{D'}(i)) \mapsto p_{r,d}(T). \quad (3.1)$$

Sind z.B. in beiden Bildern alle Knochenstrukturen einander gut überlagert, erreicht die Wahrscheinlichkeit für den entsprechenden Grauwert im Referenzbild ihren höchsten möglichen Wert. Im monomodalen Fall würde dieser dem gleichen Grauwert entsprechen und somit würde sich das gemeinsame Histogramm als eine Diagonale darstellen. Im multimodalen Fall kann es passieren, dass ein Referenzgrauwert mehreren Grauwerten im Template entspricht, z.B. der Intensitätswert des Knochen im CT entspricht in einer MR Aufnahme nicht nur dem kaum sichtbaren Knochen, sondern auch dem zum Rand des Knochen hin im MR anders ansteigenden Signal des benachbarten Fettgewebes, da Bil-

der unterschiedlicher Modalitäten unterschiedliche Informationen enthalten. Abbildung 3.1 zeigt die gemeinsamen Histogramme von unterschiedlichen Modalitäten gegenüber einem CT. Die Bilder enthalten eine Aufnahme von einem rigide registrierten Schädel, so dass ihre räumlichen Überlagerungen als optimal angesehen werden können.

Die Streuung im Histogramm nimmt mit einer zunehmenden räumlichen Abweichung zu. Betrachtet man hierbei einzelne Voxel, so entfernen sich ihre Positionen im Histogramm im monomodalen Fall immer weiter von der Diagonale (siehe Abbildung 3.2). Schafft man es, einzelne Voxel wieder in Richtung der Diagonale zu schieben, verbessert sich umgekehrt also die globale Ordnung im Histogramm. Im ersten Schritt des entwickelten Ähnlichkeitsmaßes werden für alle Voxel  $i$  Wahrscheinlichkeitskarten  $\omega_{i+\Delta t}$  unter der Translation  $\Delta t$  einzelner Voxel in einem definierten Suchbereich ermittelt

$$\omega_T : (I_{R_{256}}(i), I_{D'_{256}}(i)) \mapsto \omega(i + \Delta t) \in p_{r,d}^{256}. \quad (3.2)$$

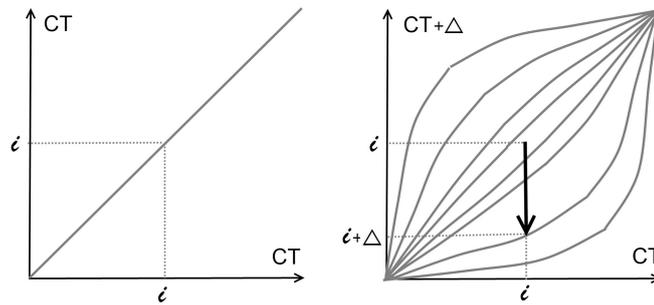
Hierbei bestimmt der Abtastabstand, mit dem die Translation des Voxels durchgeführt wird, die Größe der Karte. Die Karteneinträge sind dann so zu interpretieren, dass in der Mitte der 3-dimensionalen Karte die Wahrscheinlichkeit steht, dass der Grauwert des Voxels  $i$  im Referenzbild dem Grauwert des Voxels  $i$  im Template entspricht, also die Nulltransformation beschreibt. Der nächste Eintrag enthält die Wahrscheinlichkeit, dass der Grauwert des Voxels  $i$  in  $R$  dem Voxel  $i + 1$  in  $D'$  entspricht, falls der Abtastabstand eine Voxellänge beträgt, und so weiter. Der maximale Eintrag in dieser Karte zeigt die Translation des Voxels an, die die Häufungspunkte im gemeinsamen Histogramm lokal weiter einengen würde. Die gemeinsame Wahrscheinlichkeitsverteilung  $p_{r,d}$  errechnet man in erster Näherung durch das gemeinsame Histogramm  $h_{r,d}$ . Um eine gute Näherung zu bekommen, muß man das Histogramm mit statistisch relevanter Binbreite erstellen. Die Anzahl der Meßpunkte – in unserem Fall Bildvoxeln – bestimmt im Wesentlichen die Grauwertaufösung zur Bestimmung des gemeinsamen Histogramms. Wählt man größere Binbreiten, erhöht sich der Füllungsgrad einzelner Histogrammeinträge und die Bewertung der Ähnlichkeit wird stabiler. Dennoch verliert man dabei aber immer mehr Grauwertunterschiede und damit die Unterscheidung zwischen Gewebsarten. Um noch genügend Gewebsunterschiede zu erhalten und dennoch genügend Statistik zu haben, wird im vorgestellten Algorithmus immer mit 256-Bin Histogrammen gerechnet.

Die Transformation im ersten Schritt, der durch die Gleichung 3.2 beschrieben wird, ist beliebig, da einzelne Voxel unabhängig von ihren Nachbarn bewegt werden, um ihr lokales Optimum zu erreichen. Da außerdem meistens gleiche Grauwerte geklustert in einem Bildbereich auftreten, wird man sehr selten eindeutige Maxima in den Wahrscheinlichkeitskarten finden. Erst die unterschiedlichen Grauwerte einer Voxelumgebung enthalten überhaupt Strukturinformation. Um einen lokalen Zusammenhang der Voxelnachbarschaft miteinzubeziehen, müssen ihre Wahrscheinlichkeiten in Abhängigkeit der Transformation kombiniert werden. Abbildung 3.3 veranschaulicht das Vorgehen. Hierbei werden jeweils die Wahrscheinlichkeitskarten  $\omega(i + \Delta t)$  von nächsten  $n$  Nachbarn unter der Annahme rigider Translation für diese kleine Voxelnachbarschaft  $n \in \pm NN$  addiert

$$\omega_i^{NN} = \omega(i + \Delta t) + \sum_{\forall n \in \pm NN} \omega((i + n) + \Delta t). \quad (3.3)$$

Der maximale Eintrag der resultierenden Karte  $\omega_i^{NN}$  zeigt nun die optimale Voxeltranslation  $\vec{\kappa}_i$  an, die nun mit Nachbarschaftsvoxeltranslationen gemittelt ist

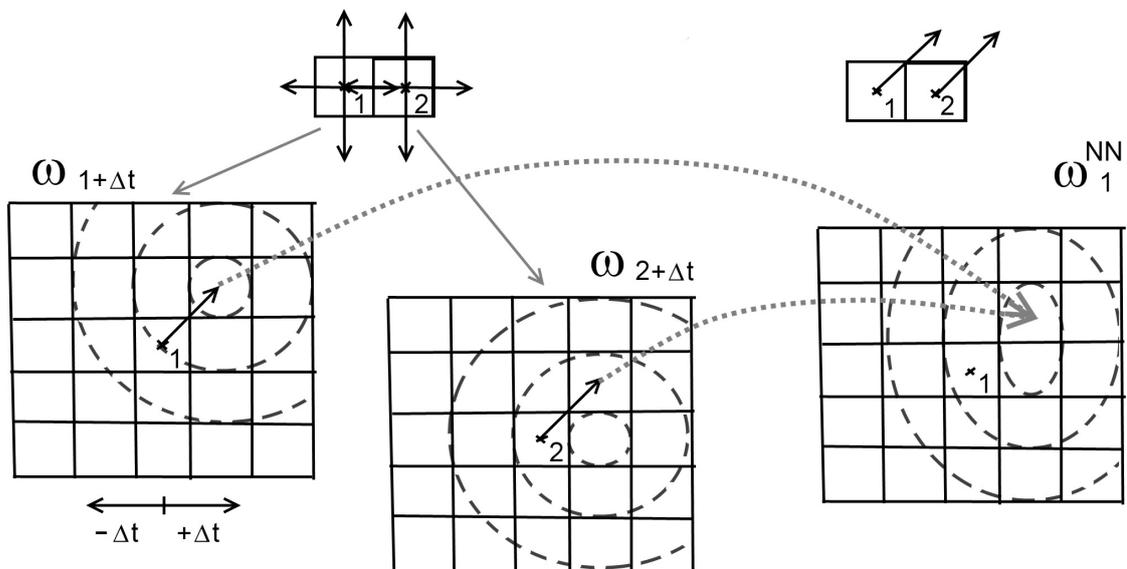
$$\max \omega_i^{NN} = \omega_i(\vec{\kappa}). \quad (3.4)$$



**Abbildung 3.2:** Schematische Darstellung des sich auffächernden gemeinsamen Histogramms eines monomodalen perfekt überlappenden Bildpaares unter einer Transformation. Links: vor der Transformation erscheint eine Diagonale. Rechts: nach der Transformation stimmen die Intensitäten vieler Voxelpaare nicht mehr überein und entfernen sich von der Diagonale entlang der Achse des verformten Bildes. Die Intensität  $i$  eines ausgewählte Voxels im statischen Referenzbild entspricht nun einer anderen Intensität  $i + \Delta$  im verformten Template.

Die Anzahl der Voxel, die ein eindeutiges Maximum in der Wahrscheinlichkeitskarte  $\omega_i^{NN}$  haben, wird in der Regel kleiner sein, als die Voxelanzahl im zu registrierenden Bildbereich. Damit ergeben sich auch weniger etablierte Korrespondenzen  $\vec{\kappa}_i$  mit  $i = 1, 2, \dots, N$  und  $N < \text{Voxelanzahl}$ .

Bei der Berechnung der Wahrscheinlichkeitskarten  $\omega_{i+\Delta t}$  (Gleichung 3.2) kann man bei nicht starken räumlichen Abweichungen jeweils für die aktuelle Transformation das aktuelle gemeinsame Histogramm benutzen. Das Ähnlichkeitsmaß ist so konstruiert, dass möglichst viele Voxelpaare ihre Histogrammeinträge in Richtung der Häufungspunkte – d.h. der Häufung von maximal besetzten Zustände im Histogramm – bewegen, da hier



**Abbildung 3.3:** Vereinfachte Darstellung der Berechnung der hier entwickelten Korrespondenzsuche. Die Wahrscheinlichkeitskarten  $\omega_{i+\Delta t}$  benachbarter Voxel 1 und 2 werden unter der Annahme addiert, dass kleine Voxelnachbarschaften als rigide Körper nur Translationen ausführen. Eine beispielhafte Translation ist durch Pfeile dargestellt.

starke Gradienten in der Wahrscheinlichkeitsverteilung vorhanden sind. So verbessert sich die Schätzung der wahren Wahrscheinlichkeitsverteilungsfunktion durch das aktuelle Histogramm in jeder Iteration. Im monomodalen Fall ist diese Schätzung gar nicht notwendig, da hier die wahre Verteilung bekannt ist und als eine Diagonale im gemeinsamen Histogramm dargestellt werden kann.

Es gibt einen weiteren Punkt, der bei der Berechnung von  $\omega_{i+\Delta t}$  beachtet werden sollte. Das Ähnlichkeitsmaß belohnt die Transformation, wenn diese die Häufungspunkte in der Wahrscheinlichkeitsverteilungen im gemeinsamen Histogramm weiter einengt, also die meistbesetzten Zustände weiter besetzt. Durch Interpolation entstehen aber in kontrastreichen Regionen neue Grauwerte, die die Ordnung im gemeinsamen Histogramm künstlich verschlechtern. Diese Artefakte sind auch bei der Berechnung der MI beschrieben worden [53]. Dieser Effekt führt in unserem Verfahren dazu, dass neuentstehende Grauwerte im Template vor allem an kontrastreichen Kanten zu falschen Korrespondenzen bezüglich des unveränderten Referenzbildes führen können. Die Anzahl dieser Fehlerkorrespondenzen kann minimiert werden, wenn man die Wahrscheinlichkeitsverteilung im gemeinsamen Histogramm entlang der Achse mit Grauwerten des Templates künstlich verbreitert. Die Verbreiterung des Histogramms steigert jedoch auch die Verlustwahrscheinlichkeit von richtigen Korrespondenzen in Umgebungen mit geringem Bildkontrast. Derzeit wird nur ein Histogramm mit einer festen Verbreiterung genutzt. Eine Lösung könnte sein, eine abhängig vom Umgebungskontrast variierende Verbreiterung einzusetzen, was allerdings mehr Speicherressourcen verbrauchen würde.

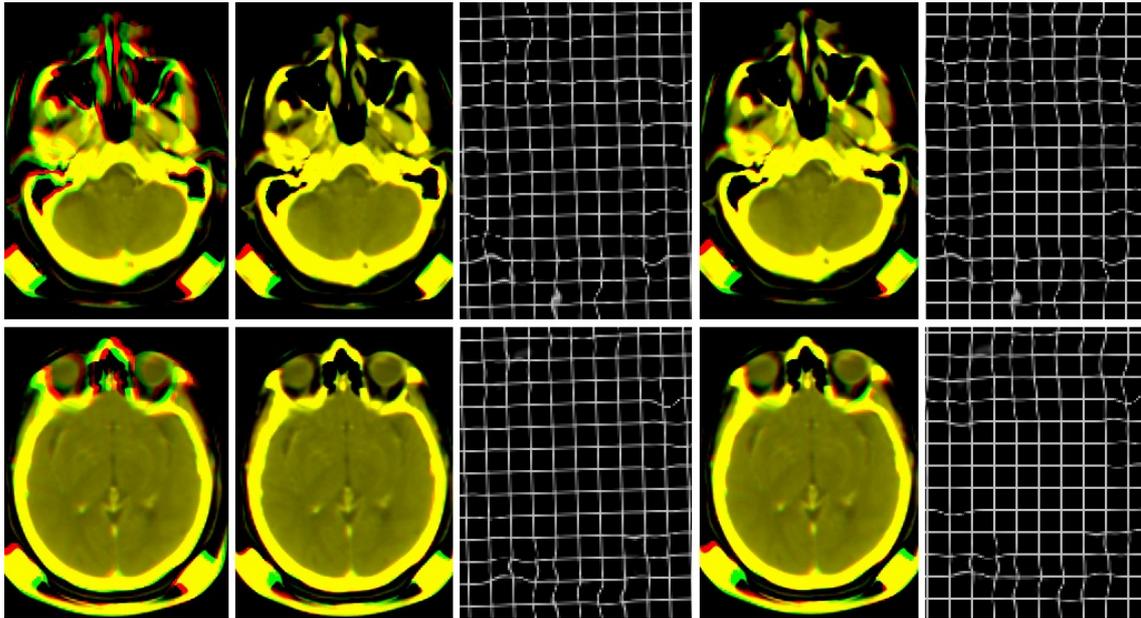
## Implementierung

Die Berechnungen im ersten Schritt in Gleichung 3.2 sind nahezu ohne Einschränkungen auf die CUDA Architektur übertragbar, da die Wahrscheinlichkeitskarten  $\omega_{i+\Delta t}$  für jeden Voxel  $i$  von je einem Thread unabhängig berechnet werden können. Eine kleine Einschränkung in der optimalen Parallelisierbarkeit stellt das nicht-koaleszente<sup>2</sup> Auslesen der gemeinsamen Wahrscheinlichkeiten aus dem gemeinsamen Histogramm dar. Hier werden in heterogenen Bildbereichen, die auf demselben Multiprozessor verarbeitet werden, divergierende Branches entstehen, die aber nach diesem Schritt innerhalb der jeweiligen Blocks (siehe Kapitel 2.3.2.2) synchronisiert werden können. Die Suchregion und damit auch die Größe aller Karten  $\omega_i$ , die im local memory zwischengespeichert werden, ist durch die Begrenzung der Hardwareressourcen auf  $11 \times 11 \times 5$  Voxellängen in  $x, y$  und  $z$  Richtung pro Iteration  $j$  im Multi-Resolution Kontext eingeschränkt.

Der zweite Schritt, der den Austausch der Wahrscheinlichkeitskarten der benachbarten Voxel benötigt (Gleichung 3.3), wird über den schnellen shared Memory der CUDA Hardware ermöglicht. Die Größe der Voxelnachbarschaft kann auf  $3 \times 3 \times 3$  Voxel reduziert werden. Die Stabilität der Korrespondenzen steigt jedoch mit der Vergrößerung dieser Voxelnachbarschaft. Dennoch schränkt die Wahl einer zu großen Nachbarschaft die darstellbare Deformation ein, denn für die Nachbarschaft wird eine rigide Translation modelliert. Für alle nachfolgenden Evaluierungen und Zeitmessungen wird eine Nachbarschaft von  $5 \times 5 \times 3$  Voxel festgelegt.

Da der Austausch der Karten  $\omega_{i+\Delta t}$  zwischen den Nachbarvoxeln der zeitaufwändigste Schritt ist, ist es sinnvoll, die Voxel in homogenen Bildbereichen, die keine eindeutigen

<sup>2</sup>Zugriffsmuster auf den Speicher nicht mit nachfolgenden Threads auf nachfolgende Speicherbereiche (siehe 2.3.2.2).



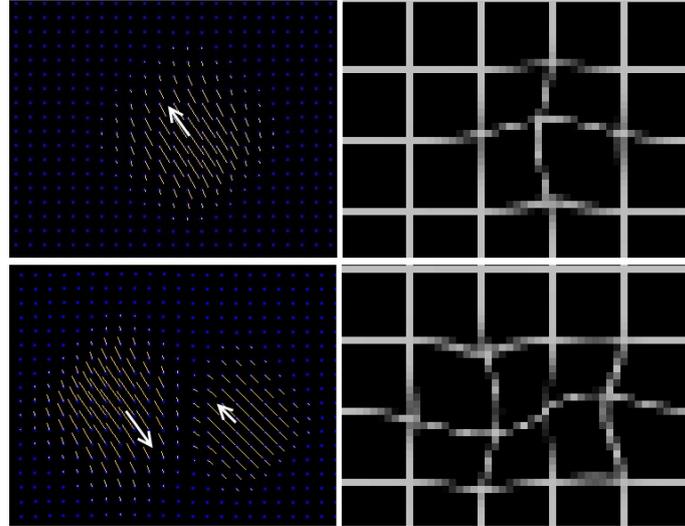
**Abbildung 3.4:** Einfluß der rigiden Transformationskomponente auf die resultierende Abbildung einer CT-Schädelaufnahme. Ersten Spalte: unterschiedliche Schichten eines CTs überlagert mit dessen um  $3^\circ$  gedrehten Kopie. Zweite Spalte: Rot-grün Fusion der registrierten Aufnahme mit dem originalen CT mit Einsatz einer globalen rigiden Transformationskomponente, die auf  $2.5^\circ$  bestimmt worden ist. Dritte Spalte: Die dazugehörige Transformation: das gesamte Gitter ist um  $2.5^\circ$  gegen den Uhrzeigersinn verdreht. Vierte Spalte: Überlagerung der registrierten Aufnahme dieses mal ohne eine globale rigide Transformationskomponente. Letzte Spalte: Die dazugehörige Transformation: nur wo Bildkontraste vorhanden sind, konnte eine Transformation bestimmt werden. Die homogenen Bereiche wurden nicht transformiert. Der globale Anteil der Drehung wird im homogenen Bereich durch die lokale Deformation ohne die Extraktion der globalen rigidem Komponente falsch beschrieben, obwohl in der Überlagerung des Endergebnisses kein Unterschied zu sehen ist.

Lösungen für Korrespondenzen liefern können, vor dem Austauschschritt auszuschließen. Hierzu wird die Varianz über die im ersten Schritt berechneten Wahrscheinlichkeitskarten  $\omega_{i+\Delta t}$  bestimmt. Durch einen Schwellwert lassen sich Voxel in homogenen Bildbereichen der Suchregion detektieren und ausmustern.

### 3.1.3.2 Extraktion des globalen Transformationsanteils

Da die Suchregion im ersten Schritt des Registrierungsverfahrens lokal eingeschränkt ist, kann man mithilfe eines lokalen Interpolationsverfahrens, das zur Parallelisierung optimal geeignet ist, keine Information über den rigidem Anteil der Gesamttransformation extrahieren. Jedoch benötigt man gerade in homogenen Regionen eine Bewegungspropagation, die aus benachbarten kontrastreichen Bildbereichen interpoliert ist, um wenigstens rigide Verschiebung der nicht sichtbaren Strukturen zu ermöglichen. Ohne einen rigidem Transformationsanteil können nur Kanten eines abgebildeten Objekts angepasst werden, wobei das Objekt selbst stark verformt wird, auch wenn es in Wirklichkeit nur rigide verschoben wurde. Abbildung 3.4 zeigt eine solche rigide Verschiebung eines Objekts und die aus der Registrierung resultierende Transformation einmal mit und einmal ohne einer rigidem Transformationskomponente.

**Abbildung 3.5:** Obere Zeile: Auswirkung einer etablierten Korrespondenz auf die Propagation der Bewegung in ihrer Umgebung nach dem eingesetzten 3D ChainMail Modell. Unterer Zeile: Auswirkung zwei benachbarter entgegen gerichteter Korrespondenzen. Die weißen Pfeile deuten die Position der Korrespondenzen an.



Um einen rigiden Transformationsanteil zu extrahieren wird ein rigides Transformationsmodell  $T_{rigid}$  an die etablierten Korrespondenzen angefügt. Die Summe der quadratischen Betragsabweichungen der Vektoren  $\vec{\kappa}_i$  und  $\vec{t}_i^{rigid}$  an den Stützstellen  $i$  wird als Zielfunktion unter Einsatz des Simplex Optimierungsverfahrens (siehe 2.2.3.3) eingesetzt. Die resultierenden Transformationsparameter werden nun von den etablierten Korrespondenzen entsprechend abgezogen, um von der Gesamttransformation, die durch die etablierten Korrespondenzen  $\vec{\kappa}_i$  beschrieben wurde, den rigiden Anteil abzukoppeln. Der verbleibende reine Deformationsanteil  $T_{deform}$  kann nun lokal interpoliert werden.

### 3.1.3.3 Vervollständigung des Vektorfeldes

Die Vervollständigung des Vektorfeldes über den gesamten Bildraum erfolgt lokal mittels des ChainMail Algorithmus [72], der in Kapitel 2.2.3.2 beschrieben wurde. Da eine effiziente Parallelisierung unabhängige Datenverarbeitung erfordert, wird die Bewegungsinformation zwischen den benachbarten Stützstellen, die beliebig im Raum verteilt sein können, nicht ausgetauscht. Als erste Näherung wird die Verschiebung jeder einzelnen Stützstelle  $\vec{\kappa}_i$  in der ihr zugehörigen Voronoizelle als Bewegung propagiert. Dadurch kann sich an Zellgrenzen eine nicht stetig differenzierbare Transformation ergeben. Im zweiten Schritt wird mittels eines Gaußfilters die Transformation geglättet, so dass starke Unstetigkeiten im Vektorfeld abgemildert werden. Der schnelle dem Propagationsschritt nachgelagerte Gaußfilter ersetzt den langsameren Informationsaustauschschritt zwischen benachbarten Voronoizellen während der Bewegungsausbreitung. Nachteilig ist hierbei, dass der Gaußfilter als Tiefpaßfilter wirkt und damit kleinskalige Deformationen wegfiltert.

Für die Bewegungspropagation, die von der jeweilige Stützstelle  $\vec{\kappa}_i$  in der Voronoizelle  $i$  ausgeht, werden empirisch bestimmte globale Kompressions- bzw. Expansions- und Scherspannungsparameter gewählt. Abbildung 3.5 zeigt exemplarisch die resultierende Transformation für eine bzw. zwei Stützstellen. Die Bewegungspropagation ist derzeit nicht von der Bildinformation abhängig. Verbesserungsmöglichkeiten bieten sich, wenn schnelle Verfahren entwickelt werden, die aus der lokalen Bildinformation beliebiger Modalitäten die Materialeigenschaften für Scher- und Kompressionsparameter realistisch schätzen können.

Der Vorteil des ChainMail Mechanismus für die Registrierungsaufgabe liegt im Wesent-

lichen darin, dass eine makroskopische physikalische Eigenschaft der Materie simuliert wird. Die Simulation basiert im Vergleich zur FE Modellierung auf einem sehr vereinfachten Kriterium, bietet aber im Vergleich zu rein mathematischen Interpolationsmethoden eine Möglichkeit lokale physikalisch-motivierte Bewegungsabläufe mitzuintegrieren. Außerdem ist die Berechnungsdauer mit der von mathematischen Methoden vergleichbar, während FEM Methoden Größenordnungen mehr an Zeit benötigen.

### 3.1.4 Zusammenfassung des Registrierungsalgorithmus

Das Referenzbild  $R(i)$  und das zu deformierende Template  $D(i)$  werden dem Algorithmus übergeben. Das Transformationsfeld wird initialisiert

$$T^0(i) = \vec{0}.$$

Die Iteration  $j$  beginnt mit der Transformation des Templates

$$D' = T^{(j-1)}(D(i)).$$

Die Datensätze werden auf die Auflösungsstufe reduziert und auf die Helligkeitsstufen des gemeinsamen Histogramms linear normiert

$$R(i), D'(i) \mapsto R_{256}^{(j)}(i), D'_{256}^{(j)}(i).$$

Die Korrespondenzsuche der Iteration  $j$  wird durchgeführt. Im ersten Schritt werden für alle Voxel die Wahrscheinlichkeitskarten erstellt

$$\omega(\Delta t \in \Omega_{search}) : I_{R_{256}^{(j)}}(i), I_{D'_{256}^{(j)}}(i) \mapsto \omega(i + \Delta t).$$

Im zweiten Schritt werden die Karten einer Nachbarschaft unter der Annahme der rigiden Translation addiert

$$\omega_i^{NN} = \omega(i + \Delta t) + \sum_{\forall n \in \pm NN} \omega((i + n) + \Delta t).$$

Der maximale Eintrag der resultierenden Karten entspricht der gesuchten Voxelkorrespondenz

$$\max \omega_i^{NN} = \omega_i(\vec{\kappa}).$$

Da nicht alle Karten ein eindeutiges Maximum zeigen, werden weniger Korrespondenzen  $\vec{\kappa}_i$  mit  $i = 1, 2, \dots, N$  als Voxel im Datensatz gefunden.

Vor der lokalen Propagation der Deformation wird ein rigider Anteil  $T_{rigid}^{(j)}$  aus den etablierten Korrespondenzen  $\vec{\kappa}_i$  abgetrennt. Hierzu wird ein optimaler Fit mithilfe eines rigidem Transformationsmodells durchgeführt. Es wird der Simplex Optimierer eingesetzt. Als Zielfunktion dient die Quadratsumme aller Abweichungen der Korrespondenzen von den Transformationsvektoren

$$Optimum(T_{rigid}^{(j)}) = \min \left( \sum_{i=1}^N (t_i - \vec{\kappa}_i)^2 \right).$$

Die Transformationsvektoren errechnen sich aus den 6 Parametern des rigidem Transformationsmodells  $trans_x, trans_y, trans_z, rot_x, rot_y, rot_z$ .

Die verbleibenden Korrespondenzen  $\vec{\kappa}'_i = \vec{\kappa}_i - \vec{t}_i^{rigid}$  werden mithilfe des ChainMail Algorithmus in ihren Voronoizellen propagiert, um die Deformation  $T_{deform}^{(j)}$  übers komplette Bild zu erhalten. Mithilfe eines Gaußfilters werden Unstetigkeiten an den Grenzen der Voronoizellen geglättet.

Die resultierende Transformation in der Iteration  $j$  setzt sich aus den beiden ermittelten Komponenten zusammen

$$T^j(i) = T_{rigid}^j(i) + T_{deform}^j(i).$$

Beginn der nächsten Iteration.

## 3.2 Evaluierung der Registrierungsergebnisse

### 3.2.1 Vergleich mit künstlichen Deformationen

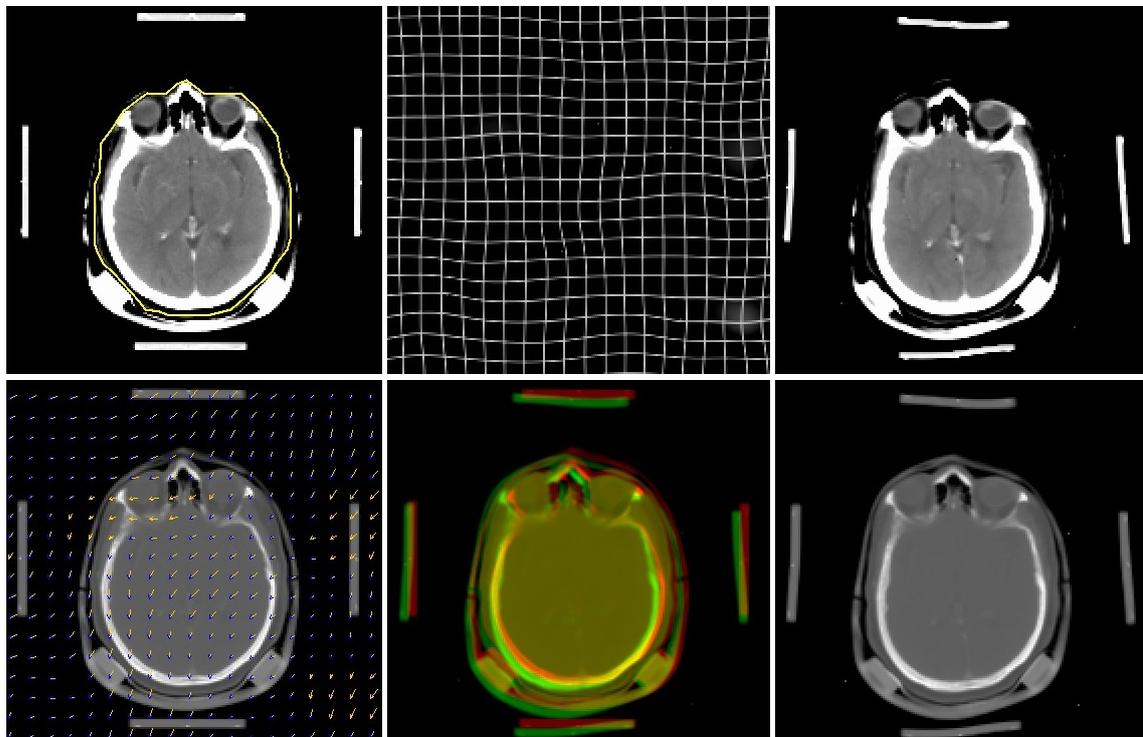
Die einzige Möglichkeit die Registrierungsgenauigkeit eines deformierbaren Registrierungsalgorithmus über den gesamten Bildraum zu überprüfen, ist es das Transformationsergebnis mit einem künstlichen bekannten Transformationsfeld zu vergleichen. Um die Datensätze mit der künstlichen Transformation herzustellen, darf natürlich nicht dasselbe Transformationsmodell genutzt werden. Wir vergleichen die Ergebnisse des entwickelten deformierbaren Registrierungsverfahrens gegen die durch TPS Interpolation erzeugten Bilddatensätze.

Die Stützstellen für die TPS Interpolation nach Gleichung 2.7 werden zufällig initiiert: sowohl deren Position im Bild, als auch die Richtungskomponenten für die Positionsabweichung. Die so erzeugten verformten Bilddatensätze werden wieder mit dem Ausgangsbild registriert und die resultierende und die bekannte Transformationen können Voxel für Voxel im gesamten interessierenden Bereich verglichen werden. Eine Schichtaufnahme eines exepharischen Bilddatensatzes vor und nach der Verformung durch die künstliche TPS Transformation ist in der Abbildung 3.6 gezeigt. Die globale TPS Interpolation erzeugt ein glattes Vektorfeld, das auch eine rigide Feldkomponente enthalten kann. Damit deckt eine solche Transformation auch einen großen Teil der realistisch möglichen Veränderungen der Patientenanatomie ab, lediglich lokal stark beschränkte Verformungen können so nicht abgebildet werden. Dennoch stellt der Vergleich mit den künstlichen TPS Transformationen eine gute Registrierungsfehlerschätzung dar, da gerade globale Veränderungen durch das eingesetzte lokale Transformationsmodell des entwickelten Registrierungsverfahrens eine potentiell große Fehlerquelle bilden.

Die Ergebnisse der Vergleiche der ermittelten mit den künstlichen Transformationen für zwei Datensätze und mehrere Felder sind in der Tabelle 3.1 zusammengetragen. Die grauunterlegten Zeilen beschreiben das ursprüngliche Ausmaß der Verformung durch die TPS. Die Mittelwerte der einzelnen Komponenten deuten eine rigide Offsettranslation in die jeweilige Richtung an, während die Standardabweichung ein Maß für die Verformung ist. Darunter sind die Abweichungen nach der Registrierung für das jeweilige Feld  $f_i$  gelistet. Die Mittelung ist über alle verbleibenden Voxeltransformationen innerhalb der in Abbildung 3.6 angedeuteten VOI (gelbe Linie) durchgeführt. Die VOI wurde als Bereichseingrenzung für die Registrierung benutzt, damit keine unnötige Bildinformation,

wie z.B. das stereotaktische Rahmensystem, die Registrierungsergebnisse beeinflussen. Sowohl die Mittelwerte als auch die Standardabweichung liegen unter der Voxelgröße von  $1.92 \times 1.92 \times 3 \text{ mm}^3$ , so dass davon ausgegangen werden kann, dass die lokale Transformation die künstliche TPS Transformation genügend gut kompensieren kann, wenn gleiche Bildinformation verglichen wird. Das statistische Rauschen in den Bildern spielt bei nicht stark verrauschten Aufnahmen keine große Rolle, da die Bilder im Prozess der Registrierung und die Berechnung des Ähnlichkeitsmaßes jeweils auf 256 Grauwerte linear reduziert werden. Damit verschwinden kleine Rauschbeiträge nahezu vollständig. Geringe Änderungen der Bildinformation durch Interpolation werden durch die Konzeption des Ähnlichkeitsmaßes ebenfalls kompensiert. Probleme für das Registrierungsverfahren können auftreten, wenn Bildstrukturen komplett verschwinden.

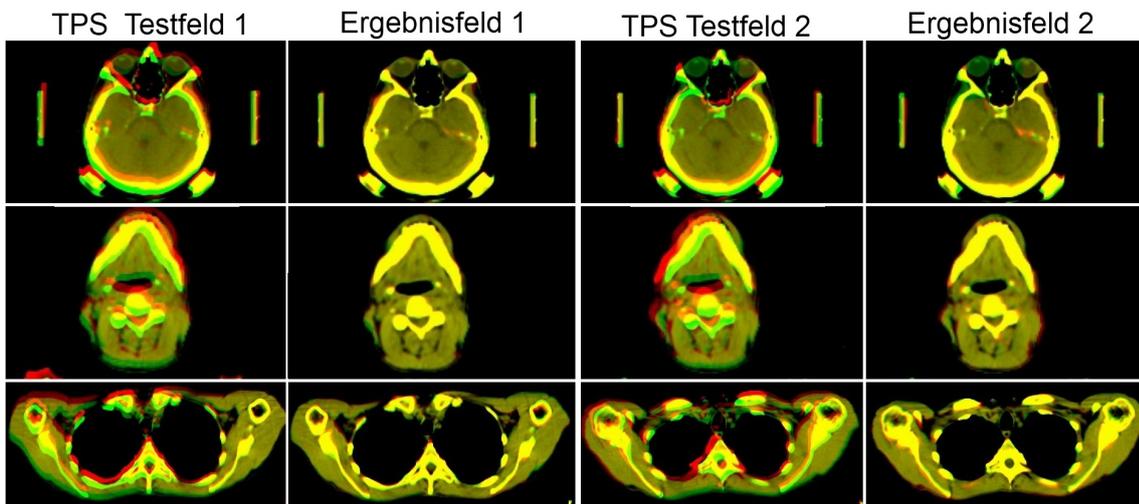
Die Abbildung 3.7 zeigt beispielhaft ausgewählte rot-grün überlagerten Schnittbilder der Ausgangsbilder und der registrierten Ergebnisse. Die Ausgangsbilder wurden mit den nach dem TPS-Modell künstlich initiierten Feldern erstellt. Die Ergebnisse durch die resultierende Transformation aus dem Registrierungsverfahren verformen die künstlich verformten Ausgangsbilder wieder zurück. Die meisten Abweichungen können durch das Registrierungsverfahren ermittelt und kompensiert werden, nur selten sieht man noch stellenweise rote Bereiche in den Ergebnisfusionierung.



**Abbildung 3.6:** Beispiel für ein künstlich erzeugtes Testfeld mithilfe TPS Interpolation und zufällig induzierten Stützstellen. Der Bildinhalt dient hier lediglich zur Visualisierung der Verformung. Ähnliche Transformationen wurden benutzt, um das hier entwickelte deformierbare Registrierungsverfahren zu testen.

	# voxel	$\Delta x \pm \sigma_x$	$\Delta y \pm \sigma_y$	$\Delta z \pm \sigma_z$	$\Delta \vec{r}  \pm \sigma_{ \vec{r} }$
<i>p</i> <sub>1</sub>					
$\Delta_{100}$	623217	$-1.83 \pm 1.10$	$2.47 \pm 1.22$	$-0.53 \pm 0.35$	$3.38 \pm 1.04$
$f_{100}$		$-0.50 \pm 0.95$	$0.11 \pm 0.93$	$0.02 \pm 0.35$	$1.26 \pm 0.76$
$\Delta_{300}$		$2.02 \pm 1.09$	$2.18 \pm 0.94$	$-0.68 \pm 0.42$	$3.24 \pm 1.01$
$f_{300}$		$0.32 \pm 0.98$	$0.13 \pm 0.82$	$-0.22 \pm 0.40$	$1.21 \pm 0.70$
$\Delta_{800}$		$-2.01 \pm 0.97$	$1.84 \pm 0.96$	$-0.35 \pm 0.36$	$2.95 \pm 0.91$
$f_{800}$		$-0.29 \pm 0.89$	$0.27 \pm 0.89$	$-0.04 \pm 0.36$	$1.15 \pm 0.73$
$\Delta_{rigid}$		$2.52 \pm 0.77$	$-0.78 \pm 0.92$	$-0.02 \pm 0.01$	$2.80 \pm 0.73$
$f_{rigid}$		$0.27 \pm 0.52$	$-0.15 \pm 0.50$	$0.06 \pm 0.11$	$0.59 \pm 0.54$
<i>p</i> <sub>2</sub>					
$\Delta_{300}$	582263	$1.89 \pm 0.99$	$2.59 \pm 0.89$	$0.72 \pm 0.37$	$3.43 \pm 0.97$
$f_{300}$		$0.06 \pm 0.86$	$0.20 \pm 0.77$	$0.02 \pm 0.45$	$1.11 \pm 0.59$
$\Delta_{800}$		$-1.98 \pm 1.18$	$2.14 \pm 1.08$	$-0.53 \pm 0.37$	$3.22 \pm 1.02$
$f_{800}$		$-0.20 \pm 1.03$	$0.36 \pm 0.92$	$-0.09 \pm 0.40$	$1.28 \pm 0.78$
$\Delta_{rigid}$		$-0.60 \pm 0.78$	$3.55 \pm 1.30$	$-0.00 \pm 0.02$	$3.70 \pm 1.27$
$f_{rigid}$		$0.08 \pm 0.69$	$0.65 \pm 0.90$	$-0.04 \pm 0.21$	$1.05 \pm 0.81$

**Tabelle 3.1:** Residuelle Abweichungen der aus der Bildregistrierung ermittelten Transformationen im Vergleich zu den bekannten künstlichen Verformungen. Die Bezeichnungen  $p_j$  stehen für Patient  $j$ ,  $f_i$  für die resultierende Abweichung nach der Registrierung des Testfeldes  $f$  mit  $i$  zufällig initiierten Stützstellen und  $\Delta_i$  für die ursprüngliche unregistrierte Fehlerverteilung des entsprechenden Feldes. Jeweils ein zufällig initiiertes rigides Transformationsfeld  $f_{rigid}$  als Testfeld ist mit aufgeführt. Alle Angaben sind nicht in mm, sondern in Voxellängen angegeben ( $1.92 \times 1.92 \times 3 \text{ mm}^3$ ).



**Abbildung 3.7:** Fusionierte Ausgangs- und Resultatbilder der Registrierung anhand drei beispielhaften transversalen Schichten des gleichen Datensatzes (jeweils in der Spalte). Jeweils die ungeraden Spalten zeigen die überlagerten CTs mit zwei beispielhaften künstlich erzeugten Verformungen mittels TPS (TPS Testfeld 1 und 2) und die geraden Spalten die Ergebnisse nach der Registrierung mit dem hier entwickelten Registrierungsverfahren (Ergebnisfeld 1 und 2).

### 3.2.2 Vergleich zu Landmarken und rigiden Teilregistrierungen

Die aus der deformierbaren Registrierung resultierenden Transformationen wurden an einzelnen Raumkoordinaten mit vordefinierten Landmarken verglichen. Die Landmarken wurden auf den originalen CT Aufnahmen mit der Auflösung von  $0.96 \times 0.96 \times 3.00 \text{ mm}^3$  von einem geschulten Experten vor der Registrierung im Behandlungsplanungssystem VIRTUOS definiert. Es ist zu beachten, dass auch Expertenlandmarken nur in heterogenen Bildregionen definiert werden können, wo der Helligkeitskontrast die Identifikation von eindeutig wieder erkennbaren Strukturen erlaubt. Da die CTs eine Schichtdicke von 3 mm aufweisen, machten sich schon während der Landmarkendefinition Fehlerquellen bemerkbar: Der Experte war gezwungen die Landmarken auf den definierten Schichten zu setzen, obwohl ersichtlich war, dass die im Planungs-CT markierte Struktur im Fraktions-CT so weit verrutscht war, dass sie zwischen den rekonstruierten Schichten zu lokalisieren sein mußte. Somit ist der Vergleich der z-Koordinate mit größerem intrinsischen Fehler behaftet, der nicht quantifiziert werden kann. Die  $x$ - und  $y$ -Koordinaten waren prinzipiell mit Fließkommazahlgenauigkeit definierbar. Um den Fehler der Landmarkendefinition abzuschätzen, wurden in einem Planungs-CT mehrmals die gleichen Landmarken definiert. Der über alle Landmarken gemittelte Fehler beträgt  $0.42 \pm 0.31 \text{ mm}$ . Zur leichteren Identifikation derselben Landmarken wurde eine visuelle Darstellung der bereits einmal definierten Landmarken genutzt.

Man weiß, dass die Definition von selben Landmarken sich von einem Experten zum anderen ebenfalls unterscheiden kann. Da die Landmarkendefinition eine sehr aufwändige und zeitintensive Aufgabe darstellt und dennoch keine optimale Evaluationsmethodik für Registrierungsalgorithmen bietet, wurde auf den Einsatz eines zweiten Experten verzichtet.

Die Landmarken wurden auf Patientendaten mit Kopf-Hals-Tumoren definiert. Dieses Patientenkollektiv zeigt ausgeprägte Verformungen während des Therapieverlaufs, da diese Tumorlokalisierung schwierig zu fixieren ist und die Patienten häufig auch wegen Ge-

		# LM	$\Delta x \pm \sigma_x [mm]$	$\Delta y \pm \sigma_y [mm]$	$\Delta z \pm \sigma_z [mm]$	$\Delta  \vec{r}  \pm \sigma_{ \vec{r} } [mm]$
$p_1$	$\Delta_{LM}$	32	$-0.67 \pm 1.71$	$-4.07 \pm 3.05$	$-1.41 \pm 4.38$	$6.21 \pm 3.35$
	$\Delta_{GPU-DIR}$		$-0.51 \pm 2.51$	$-0.50 \pm 1.94$	$-0.53 \pm 3.58$	$4.05 \pm 2.60$
	$\Delta_{TM}$		$-0.43 \pm 1.55$	$-0.39 \pm 2.26$	$0.12 \pm 3.82$	$3.97 \pm 2.50$
$p_2$	$\Delta_{LM}$	32	$-2.01 \pm 2.30$	$-7.59 \pm 4.19$	$-1.50 \pm 2.85$	$8.95 \pm 3.77$
	$\Delta_{GPU-DIR}$		$-0.01 \pm 2.23$	$-0.71 \pm 2.29$	$0.17 \pm 2.48$	$3.49 \pm 2.08$
	$\Delta_{TM}$		$-0.33 \pm 1.96$	$-0.65 \pm 1.99$	$-0.03 \pm 2.04$	$3.15 \pm 1.51$
$p_3$	$\Delta_{LM}$	32	$1.28 \pm 1.93$	$2.95 \pm 3.21$	$-0.19 \pm 3.49$	$5.45 \pm 2.50$
	$\Delta_{GPU-DIR}$		$0.35 \pm 1.68$	$-0.42 \pm 3.10$	$-0.78 \pm 3.48$	$3.70 \pm 3.37$
	$\Delta_{TM}$		$-0.11 \pm 1.60$	$-0.39 \pm 3.12$	$-0.59 \pm 3.83$	$3.70 \pm 3.66$

**Tabelle 3.2:** Residuelle Abweichungen der aus der Bildregistrierung ermittelten Transformationen im Vergleich zu einzelnen vom Experten gesetzten Landmarken. Grau unterlegt sind die gemittelten Abweichungen von unregistrierten Bildern definiert durch die Landmarken. Die Datensätze wurden mit dem hier entwickelten Registrierungsverfahren (GPU-DIR) und einem alternativen Verfahren [35] (TM) registriert. Beide Registrierungsalgorithmen schaffen es die Offset-Verschiebungen der Patientendatensätze in  $y$ -Richtung zu ermitteln, was durch die kleinen Mittelwerte der Restabweichungen beschrieben wird. Die Standardabweichung von  $1\text{-}\sigma$  beschreibt die Verteilung der Restfehler, die durch die Deformations- und rigide Rotationsanteile herrühren.

		# LM	$\Delta x \pm \sigma_x$ [mm]	$\Delta y \pm \sigma_y$ [mm]	$\Delta z \pm \sigma_z$ [mm]	$\Delta  \vec{r}  \pm \sigma_{ \vec{r} }$ [mm]
$p_1$	$\Delta_{sR}$	7	-0.31 $\pm$ 0.56	-3.79 $\pm$ 4.24	-0.38 $\pm$ 1.45	4.28 $\pm$ 4.00
	$\Delta_{GPU-DIR}$		0.87 $\pm$ 0.65	-2.08 $\pm$ 2.23	0.08 $\pm$ 0.69	2.71 $\pm$ 1.81
	$\Delta_{TM}$		0.32 $\pm$ 0.43	-1.71 $\pm$ 2.02	0.53 $\pm$ 1.04	2.43 $\pm$ 1.52
$p_2$	$\Delta_{sR}$	7	-1.68 $\pm$ 1.53	-5.51 $\pm$ 4.96	-1.00 $\pm$ 1.54	6.22 $\pm$ 4.91
	$\Delta_{GPU-DIR}$		0.22 $\pm$ 2.11	-0.55 $\pm$ 1.27	-0.53 $\pm$ 1.92	2.59 $\pm$ 1.64
	$\Delta_{TM}$		-0.36 $\pm$ 0.70	-0.09 $\pm$ 0.69	0.15 $\pm$ 0.80	1.21 $\pm$ 0.29

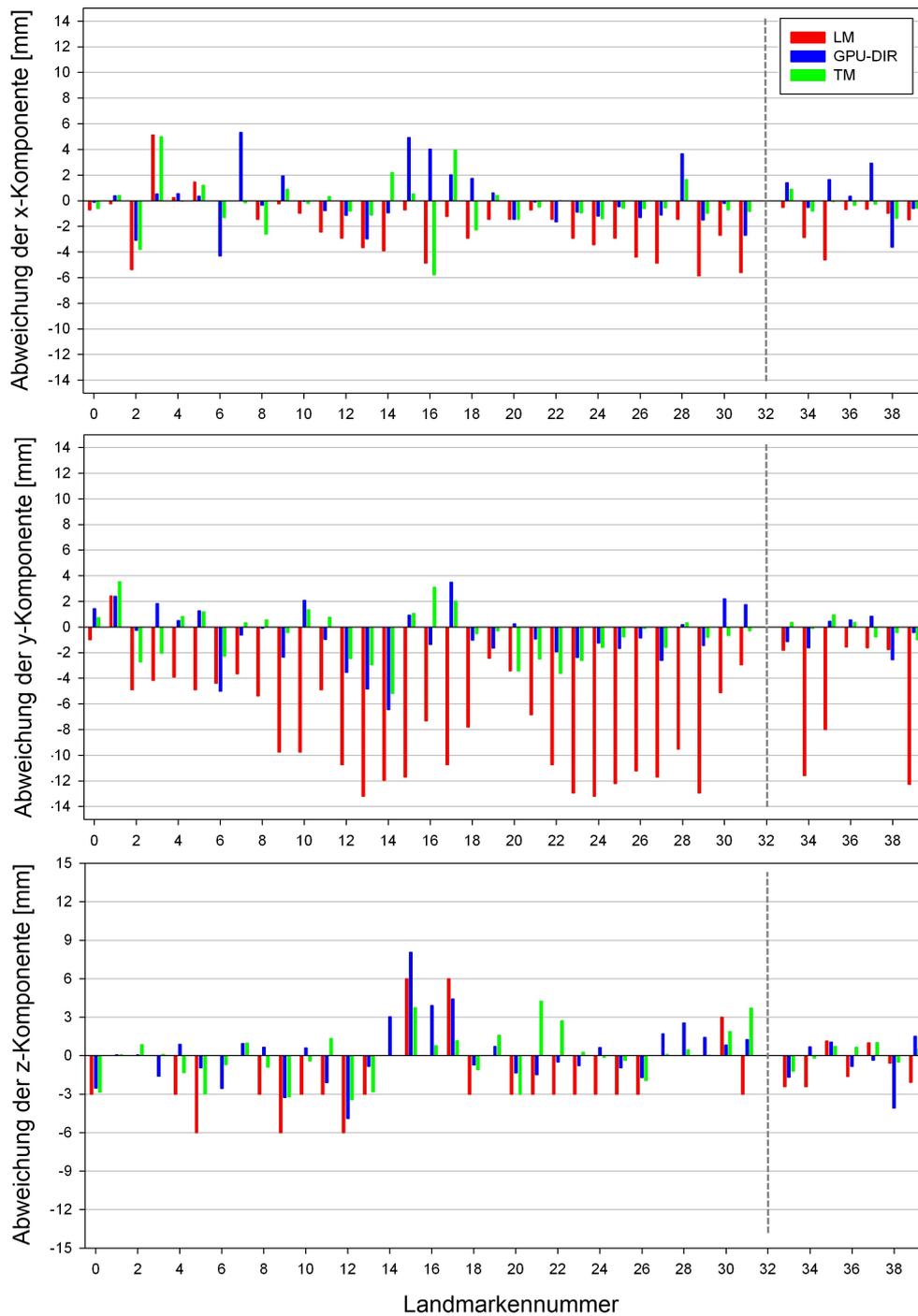
**Tabelle 3.3:** Residuelle Abweichungen der aus der Bildregistrierung ermittelten Transformationen im Vergleich zu rigiden Teilregistrierungen mittels Mutual Information. Grau unterlegt sind die ermittelten Abweichungen durch die rigide Registrierung der Subregionen. Die restlichen Bezeichnungen sind die gleichen wie in Tabelle 3.2

wichtsverlust nicht mehr optimal in der Fixierungshilfe gelagert werden können. Es wurden immer 32 Landmarken definiert, wobei die ersten 22 bewußt in Bereichen mit potentiell starken Deformationen unterworfenen Anatomie im Weichteilbereich gewählt wurden. Die übrigen 10 Landmarken wurden an knöchernen Strukturen gesetzt.

Die Positionsunterschiede an den Expertenlandmarken zwischen dem Planungs-CT eines exemplarisch gewählten Patienten und einer Kontrollaufnahme vor der 24. Fraktion sind in der Abbildung 3.8 in  $x, y$  und  $z$  Richtung aufgespalten als rote Balken dargestellt. Zu beachten ist, dass die Korrespondenzsuche auf einer reduzierten Bildauflösungsstufe von doppelter Voxelgröße in  $x$  und  $y$  Richtung stattfand. Die über alle Landmarken gemittelten Abweichungen und deren Standardabweichung für drei Bildpaare sind in der Tabelle 3.2 in den mit grau unterlegten Zeilen zusammengetragen. Wobei der Mittelwert der einzelnen Komponenten eher die rigiden Offsetverschiebung der Patienten-anatomie beschreibt und die  $1\sigma$ -Abweichung als ein Maß für die Deformation interpretiert werden kann.

Die verbliebenen Abweichungen der registrierten Fraktionen bezüglich der Expertenlandmarken nach der Anwendung des entwickelten deformierbaren Registrierungsverfahrens sind als blaue Balken in der Abbildung 3.8 dargestellt. Sieht man die Landmarken als Goldstandard an, müßten bei perfekter Registrierung alle verbleibenden Positionsabweichungen verschwinden, da sich die gleichen Strukturen überlagern müßten. In derselben Abbildung sind als grüne Balken die Abweichungen nach einer alternativen Transformation eingezeichnet, darauf soll später im Vergleich mit alternativen Algorithmen (Kapitel 3.3.1) eingegangen werden. Die grauen horizontalen Linien markieren die Voxelgröße in die jeweilige Richtung. Die verbliebenen gemittelten Abweichungen der übrigen Bildpaare sind ebenfalls in der Tabelle 3.2 zusammengefaßt. Während der Mittelwert über alle Landmarken Subvoxelgröße erreicht, zeigt die Abweichung eine größere Schwankung als eine Voxellänge, was dafür spricht, dass nur etwa 70% der Landmarkenpositionen im registrierten Bild genügend genau mit dem Referenzbild übereinstimmen. Der restliche Anteil kann aufgrund von Registrierungsfehlern oder den Landmarkdefinitionsfehlern entstehen.

Ein ähnliches Bild zeigt der Vergleich von der ermittelten Transformation gegenüber den lokalen Transformationen einer rigiden Subregistrierung. Kleine Subregionen derselben Patientendaten wurden mithilfe eines rigiden Transformationsmodells registriert. So kann eine globale Deformation durch mehrere lokale rigide Transformationen angenähert werden. Wählt man die Subregionen klein genug, sollte die Transformation des repräsentativen Mittelpunktes der Subregion dem Vektor an derselben Position im ermittelten Vek-



**Abbildung 3.8:** Detaillierter Vergleich der resultierenden Verformung aus deformierbaren Registrierungsverfahren mit vordefinierten Landmarken und rigiden Teilregistrierungen. Die Landmarken links der gestrichelten vertikalen Linie sind durch einen Experten per Hand gesetzt. Die übrigen sieben Landmarken sind die Mittelpunkte der kleinen Subregionen, die als rigide Templates registriert worden sind. Die Graphen zeigen die verbleibende Abweichung in die x (oben), y (mitte) und z (unten) Richtung ohne eine Registrierung (rot), nach der Registrierung mit dem hier entwickelten Registrierungsverfahren (blau) und mit dem TM Verfahren [35] (grün). Alle Restfehler beziehen sich auf die Abweichung von den vordefinierten Landmarken.

		$p_1$ : Zeit [s]		$p_2$ : Zeit [s]	
TM	Templateauswahl	12.2		8.8	
	Korrespondenzsuche	12.8		9.3	
	Koeffizientenberechnung	96.9		88.2	
	Feldinterpolation	99.0		90.0	
	Korrespondenzanzahl	<1500		<1300	
Summe		220.9		186.3	
		1. Schritt	2. Schritt	1. Schritt	2. Schritt
GPU-DIR	Korrespondenzsuche	6.1	31.6	5.7	29.6
	globale Komponente	12.9	12.9	6.3	7.3
	Feldvervollständigung	6.4	14.2	5.0	11.8
	Korrespondenzanzahl	11242	52537	10691	46876
	Summe	25.4	58.7	17.0	48.7
Gesamtsumme		84.1		65.7	

**Tabelle 3.4:** Dauer der Registrierungsprozesse für zwei exemplarische Bildpaare  $p_1$  und  $p_2$ . Das hier entwickelte Verfahren wurde in zwei Auflösungsstufen durchlaufen (1. und 2. Schritt).

torfeld der deformierbaren Registrierung entsprechen. Die Abweichungen von den so ermittelten Positionskorrespondenzen sind für einen exemplarischen Patienten zusammen mit den Landmarkenvergleichen in Abbildung 3.8 nach der gestrichelten Linie dargestellt. Die gemittelten Werte sind in Tabelle 3.3 zusammengetragen. Hier ist zu beachten, dass die rigide Subregistrierungen den Positionsfehler des Mittelpunktes unterschätzen können, da eine große Nachbarschaftsumgebung zum möglichst guten Überlappen der abgebildeten Strukturen gebracht wird. Die Vergleiche zeigen, dass große Abweichungen mit vergleichbarer Genauigkeit wie im Landmarkenvergleich gefunden werden. Das TM-Registrierungsverfahren erreicht in diesem Vergleich kleinere Abweichungen verglichen zu dem hier entwickelten GPU-DIR. Dies kann daran liegen, dass die im TM eingesetzte globale TPS Interpolation besser rigide Transformationen darstellen kann, als im GPU-DIR eingesetzte lokale Verfahren. Da auch die Positionsfehler nur mithilfe einer rigiden Transformation ermittelt worden sind, erwartet man vom TM erzeugten Resultat kleinere Abweichungen.

### 3.2.3 Laufzeitperformance

Die Laufzeitperformance des vorgestellten deformierbaren Registrierungsverfahrens wurde auf einem System erfasst, das mit einer NVIDIA GTX 8800 (760 MB D-RAM) und einer 3,3 GHz CPU (3 MB RAM) ausgerüstet ist. Die Laufzeiten werden mit denen des bereits in VIRTUOS implementierten schnellen TM Algorithmus auf dem selben PC gegenübergestellt. Die Ergebnisse sind in Tabelle 3.4 zusammengetragen. Die Anforderungen an die Rechengeschwindigkeit für die fraktionierte Therapie ist mit der Gesamtlaufzeit von unter einer Minute für die Registrierung von 3D Tomographien auf der eingesetzten Standardhardware (GeForce 8800 GTX) gut erfüllt.

Da insbesondere die Korrespondenzsuche und die Interpolation als rechenintensivste Schritte eines deformierbaren Registrierungsverfahrens gelten, sind in Tabelle 3.4 auch die Laufzeiten nach Komponenten aufgeschlüsselt. Verglichen mit dem TM Verfahren, benötigt die parallelisierte lokale Vektorfeldvervollständigung viel weniger Zeit als die

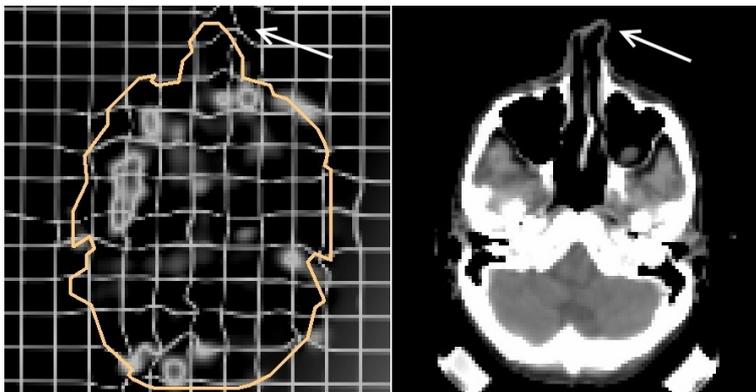
globale TPS Interpolation, obwohl viel mehr etablierte Korrespondenzen im vorgestellten Verfahren gefunden wurden. Diese Beschleunigung ist im Wesentlichen auf die lokale Natur der Propagation der Bewegung im Volumen durch den 3D ChainMail Algorithmus und dessen gute Abbildbarkeit auf die GPU Hardware zurückzuführen. Die Korrespondenzsuche nimmt beim vorgestellten Algorithmus verglichen zum TM wesentlich mehr Zeit in Anspruch, was hauptsächlich daran liegt, dass wesentlich mehr Templates als Korrespondenzkandidaten in der Suche eingesetzt werden. Nahezu jeder Voxel kann mit seiner Umgebung als Kandidat interpretiert werden. Nur die im Vorverarbeitungsschritt ausgemusterten Voxelpositionen, die keine Variation im Ähnlichkeitsmaß gezeigt haben, sowie implementierungstechnisch ausgemusterten Voxel, die keine vollständige Nachbarschaft in ihrem Block aufweisen, wurden von der Suche als Kandidaten ausgeschlossen. Auch die Anzahl der endgültigen etablierten Korrespondenzen ist der Tabelle 3.4 zu entnehmen.

Wollte man die benötigten Zeiten für eine ähnliche Menge an Korrespondenzen, die im hier entwickelten Verfahren ermittelt wird, im TM Verfahren abschätzen, könnte man grob einen linearen Zusammenhang annehmen. Denn die Korrespondenzsuche im TM hängt nur von der Anzahl der Templates ab – bei bleibender Template- und Suchregiongröße. Auch die blockweise aufgeteilte TPS Interpolation hängt nur von der Blockanzahl ab, wenn die Stützstellenzahl pro Bildblock gleich bleibt. Die benötigte Zeit für die Verarbeitung der gleichen Menge an Korrespondenzen mit dem TM Verfahren wäre dann 9388 und 8253 Sekunden für  $p_1$  und  $p_2$  – entsprechend 2 Stunden und 36 Minuten bzw. 2 Stunden und 17 Minuten.

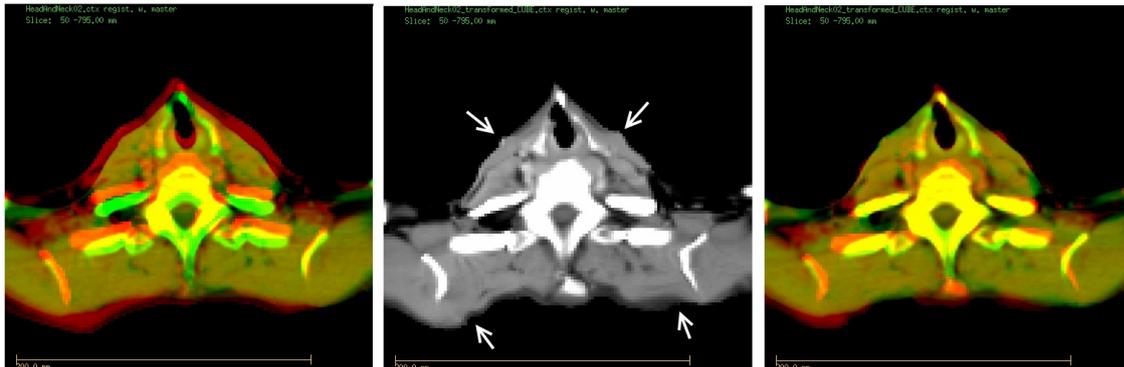
### 3.2.4 Limitierungen

#### 3.2.4.1 Monomodale Registrierung

Die Limitierungen des entwickelten Registrierungsalgorithmus liegen wie bei anderen auch im noch fehlenden physikalisch zusammenhängenden Transformationsmodell. Die Vektorfeldvervollständigung findet allein auf Basis der etablierten Korrespondenzen statt. Die Bildinformation an den zu interpolierenden Stellen fließt nicht in die Berechnungen mit ein. Fehlkorrespondenzen werden nicht als solche erkannt und durch das Bild – auch wenn nur lokal – propagiert, wie in Abbildung 3.9 zu sehen ist. Abhängig von der Dichte der etablierten Korrespondenzen, können durch die lokale Bewegungspropagation Lücken in einer tatsächlich zusammenhängenden Bewegung entstehen. Das führt zu unrealistisch wirkenden kantigen Verformungen wie in Abbildung 3.10 durch Pfeile angedeutet.



**Abbildung 3.9:** Beispiel für die Auswirkung einer Fehlkorrespondenz in der Korrespondenzsuche. Die betroffene Position ist durch den weißen Pfeil markiert.

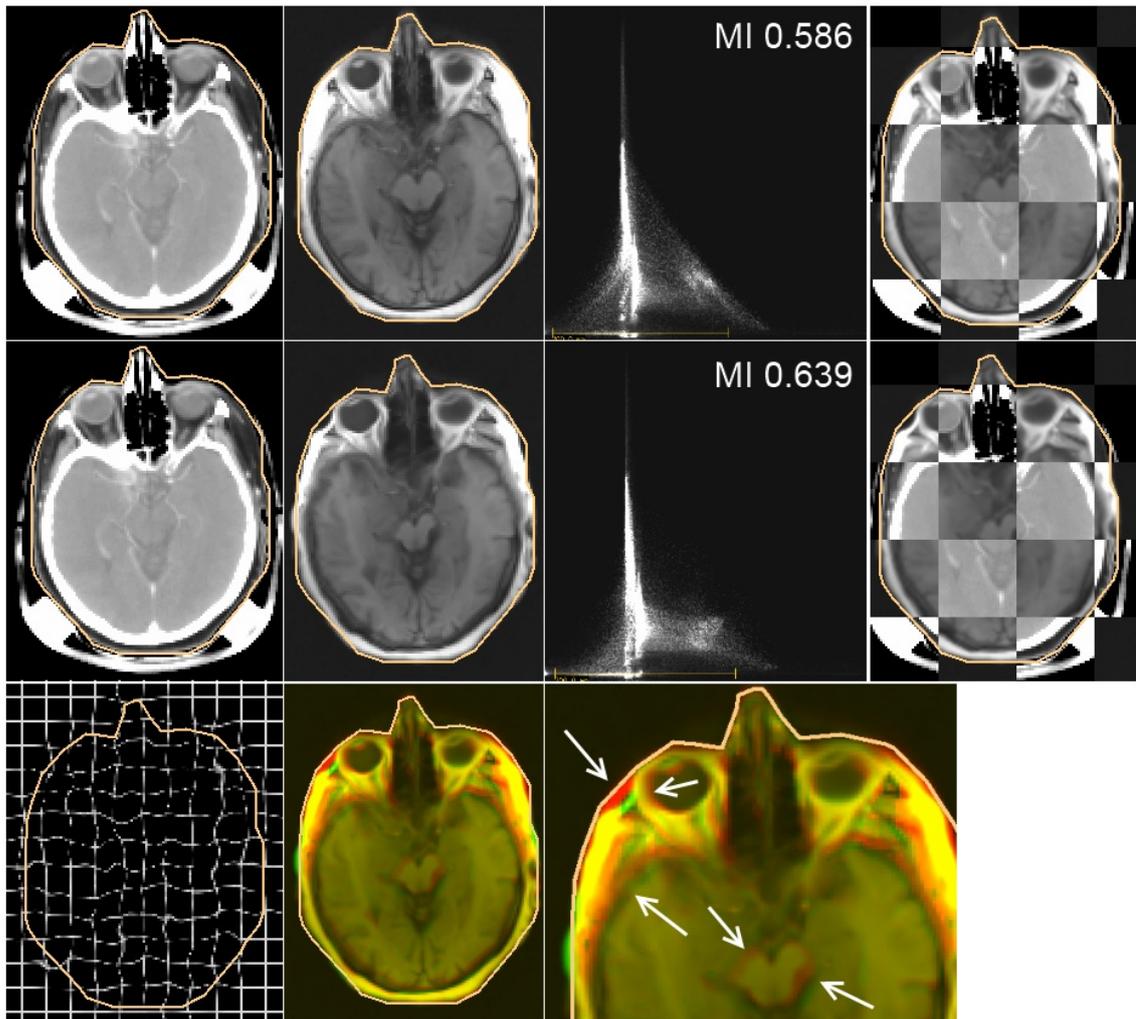


**Abbildung 3.10:** Beispiel für die Auswirkung der lokal eingeschränkten Transformation: wenn nicht genügend Korrespondenzen gefunden werden, entstehen Lücken in der Transformation, die nur durch globale oder modellbasierte Ansätze abgefangen werden können.

Die Einführung eines zusammenhängenden Transformationsmodells würde allerdings zusätzlichen Synchronisationsoverhead bedeuten, da auch weiter von einander im Bild entfernten Daten zwischen den Rechensträngen ausgetauscht werden müssten. Dennoch braucht man solche Modelle, da im Patienten sowohl lokal limitierte, als auch global zusammenhängende Bewegungen vorkommen können. Ein Zwischenschritt kann sein, die lokalen Bewegungspropagationsparameter von der lokalen Bildinformation abhängig zu implementieren. So kann man einen lokal rigiden Bildausschnitt rigide belassen, wenn die Kompression bzw. Expansion der Volumenelemente in diesem Bildbereich verboten wird, indem man die lokalen Kompressions- und Expansionsparameter auf null setzt. Dennoch wird dann zusätzlicher Rechenaufwand benötigt um diese Information aus Bildern zu extrahieren.

### 3.2.4.2 Multimodale Registrierung

Registrierung verschiedener Modalitäten sollte prinzipiell mit dem entwickelten Ähnlichkeitsmaß möglich sein, wenn man durch eine rigide Vorregistrierung eine gute Schätzung der wahren gemeinsamen Wahrscheinlichkeitsverteilung für Grauwertentsprechung erreichen kann. Um das zu testen wird ein Datensatz mit abgebildetem Schädelbereich genutzt, wo keine Deformationen zu erwarten sind, und man davon ausgehen kann, dass eine rigide Registrierung in diesem Fall die wahre gemeinsame Verteilung ergeben wird. Ein CT eines Schädels und eine  $T_1$  gewichtete Schädelaufnahme desselben Patienten werden nun mit dem entwickelten deformierbaren Registrierungsverfahren registriert. Die Erwartung ist, dass nur weniger bis gar keine Verschiebung gefunden wird, da alle Grauwertentsprechungen auch im gemeinsamen Histogramm gefüllt sind. Doch ist das Maß so geartet, dass die Streuung im Histogramm, falls sich ein Voxelpaar in der Nähe eines Häufungspunktes befindet, durch Transformation des Referenzvoxels weiter abnehmen soll. Bei einem freien Bewegungsmodell führt das zur Erhöhung des lokalen Ähnlichkeitsmaßes, aber zu einer falschen Transformation, da keine globale Nebenbedingung im gemeinsamen Histogramm beachtet wird. Übrigens gilt dasselbe für den Einsatz der MI als Ähnlichkeitsmaß, da diese auch nur die Streuung des gemeinsamen Histogramms wertet. Wenn auch unrealistische Transformationen möglich sind, die die Häufungspunkte weiter schärfen, wird auch die MI diese falschen Transformationen als die besseren im Registrierungsprozess be-



**Abbildung 3.11:** Multimodale Registrierung. Erste Zeile von links nach rechts: Eine CT-Schädelaufnahme, eine rigide vorregistrierte  $T_1$  gewichtete MR Aufnahme, deren gemeinsames Histogramm mit eingetragenem MI Wert, Fusion beider Aufnahmen in Checkboard Darstellung. Zweite Zeile: selbes CT, transformierte MR Aufnahme nach der deformierbaren Registrierung auf maximaler Auflösungsstufe, gemeinsamen Histogramm mit MI Wert, Fusion. Dritte Zeile: aus DIR resultierende Transformation, Rot-Grün-Überlagerung des originalen und transformierten MR Datensatzes und ein vergrößerter Ausschnitt. Die weißen Pfeile deuten einige Verformungen an. Trotz der unrealistischen Transformation (z.B. sind die Augen etwas vergrößert) erreicht man einen geordneteren Zustand im gemeinsamen Histogramm und einen höheren MI Wert.

werten. Die Abbildung 3.11 zeigt das Ergebnis der multimodalen Registrierung und die entsprechenden gemeinsamen Histogramme.

Beim Einsatz multimodaler Ähnlichkeitsmaße, wie der MI, in der deformierbaren Registrierung, muß auf die Wahl des Transformationsmodells besonders geachtet werden. Nur ein stark konditioniertes Modell, wie das des rigiden Körpers oder ein vollständiges biophysikalisches Patientenmodell, kann hier zu einer akzeptablen Lösung führen.

### 3.3 Anwendung auf Patientendatensätze

Die Anwendung der deformierbaren Registrierungsverfahren in der ART soll die Anpassung des Patientenmodells im Verlauf der Therapie ermöglichen. Das Patientenmodell besteht einerseits aus den volumetrischen anatomischen Informationen, welche durch die Bildgebung gemessen werden und andererseits aus dem Wissen, wie man diese Grauwertverteilung interpretieren kann. Dazu muß man therapeutisch relevanten Strukturen in den Bildern finden, abgrenzen und mit der Zeit nachverfolgen können. Durch dieses Wissen ist es möglich, die Gesamtdosis, die während einer fraktionierten Therapie in einem Volumen appliziert wird, korrekt zu berechnen, auch wenn sich die Form der Struktur über die Zeit verändert.

#### 3.3.1 Vergleich zur Verformungsanalyse alternativer Algorithmen

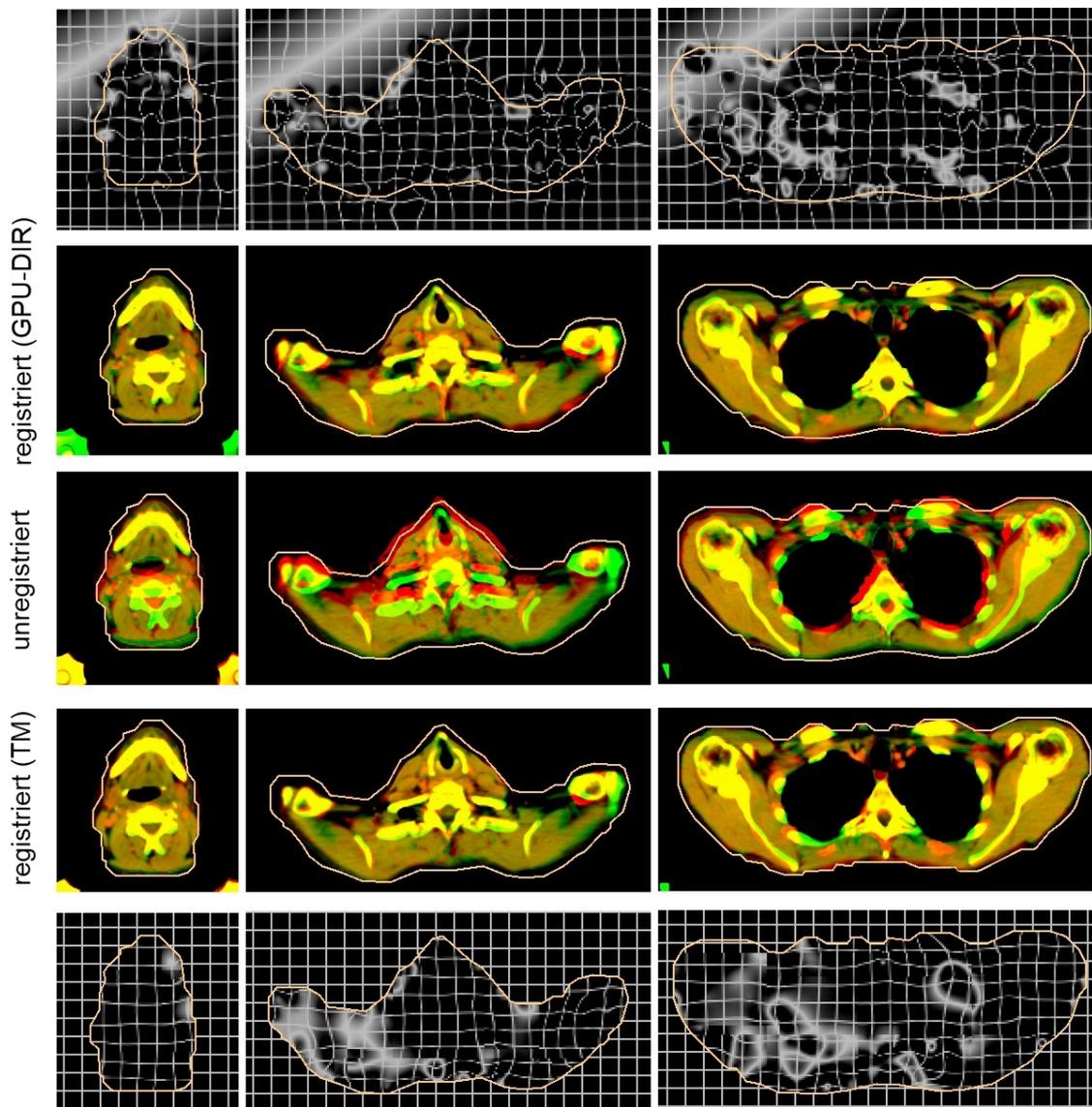
Im Vergleich zu dem TM Algorithmus [35] zeigt der vorgestellte Algorithmus ähnliche Resultate. Die überlagerten Darstellungen der registrierten Datensätze in Abbildung 3.12 lassen kaum qualitative Unterschiede erkennen. Der Landmarkenvergleich auf den gleichen Datensätzen für einen exemplarischen Patienten und Fraktion zeigt (Abbildung 3.8), dass abhängig von der individuellen Landmarke manchmal schlechtere (z.B. Landmarke #7,  $x$  und  $y$ -Komponente) manchmal bessere (z.B. Landmarke #4) Resultate vom vorgestellten Verfahren zu erwarten sind. Im Mittel über alle Landmarken (Tabelle 3.2) schneidet der vorgestellte Algorithmus vergleichbar mit dem TM ab, wobei beide Algorithmen im Schnitt zwar die Subvoxelgenauigkeit erreichen, aber bei beiden die Streuung der Abweichungen von  $1-\sigma$  bereits über der Voxelgröße von  $1.92 \times 1.92 \times 3.00 \text{ mm}^3$  liegt.

Zu bedenken sind aber auch die Ungenauigkeiten der Landmarkendefinition selbst, die bereits bei der Identifikation zu Schwierigkeiten geführt hat, insoweit dass anatomische Strukturen zwischen den dargestellten Schichten verschwanden und dennoch nur auf den Schichten definiert werden konnten. Die Reproduzierbarkeit der Landmarken im selben CT betrug  $0.42 \pm 0.31 \text{ mm}$  innerhalb der Schicht, wobei bei der wiederholten Landmarkendefinition die bereits definierten Landmarken als Unterstützung bei der Identifizierung ausgedruckt vorlagen. In der  $z$  Komponente konnte wegen der großen Schichtdicke und visueller Unterstützung keine Abweichung festgestellt werden.

In diversen Publikationen über deformierbare Algorithmen wird eine Registrierungsungenauigkeit im Subvoxelbereich angegeben [48,52,64]. In interinstitutionellen Studien im Phantomexperiment wurde aber gezeigt, dass diese Fehlerangaben stark unterschätzt sind [76]. Vor diesem Hintergrund sind Landmarkenvergleiche mit Vorsicht zu genießen, da die Landmarkenqualität –, ob diese z.B. nur in kontrastreichen, rigiden Strukturen gesetzt sind oder nicht – sich stark unterscheiden kann.

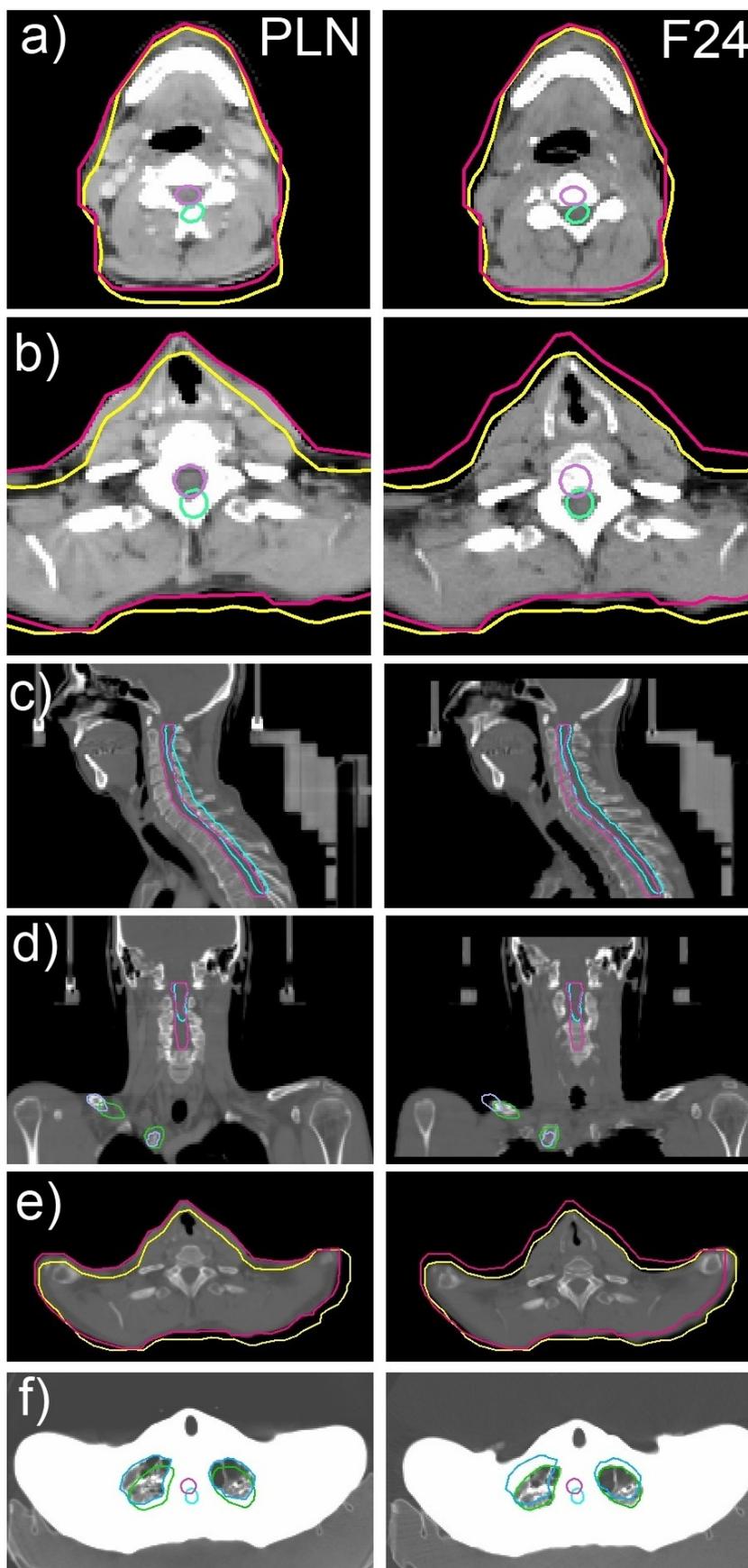
#### 3.3.2 Adaption von Patientenmodellen

Das Patientenmodell, das die segmentierten Strukturen enthält, wird derzeit in der Strahlentherapie Tage oder gar Wochen vor der ersten Bestrahlungsfraction anhand eines Planungs-CTs definiert. Gelingt es, den Patienten durch gute Lagerungs- und Fixierungshilfen wie geplant in allen Fraktionen zu positionieren, reichen diese statischen Konturen und der statische Plan aus. Treten jedoch während der Behandlung anatomische Veränderungen auf, braucht man eine Anpassung der konturierten VOIs, um sowohl dosimetrische Änderungen des angewandten Plans zu rekonstruieren, um diese z.B. als neue Nebenbe-



**Abbildung 3.12:** Visuelle Resultate der DIR-Verfahren an Patientendaten: Die mittlere Zeile zeigt die Überlagerung der unregistrierten Ausgangsbilder (Planungs- und Kontroll-CT in der 24sten Fraktion) in drei verschiedenen transversalen Ebenen. Die Zeile darüber (darunter) zeigt die Überlagerung der registrierten Datensätze mittels des hier entwickelten Verfahrens (alternativen TM-Verfahren). Jeweils Zeile darüber (darunter) sind die entsprechenden Transformationen mittels Gitterverformung visualisiert. Beide Verfahren kompensieren die Bildunterschiede, zeigen aber auch noch stellenweise verbleibende Abweichungen. Die weißen ellipsoiden Strukturen in den Gitterdarstellungen deuten die Verformung senkrecht zur Transversalebene an. Da das Gitter 3-dimensional aufgebaut ist, äußert sich eine lokale z-Verschiebung durch die Projektion einer weißen Gitterebene in die darunter oder darüber liegende Schichten je nach Verschiebungslänge. Diese Strukturen können ähnlich den Höhenlinien interpretiert werden.

**Abbildung 3.13:** Beispiele für die Adaption der Konturen an die veränderte Anatomie. Neue Konturen wurden anhand des Vektorfeldes nach der deformierbaren Registrierung erzeugt. Alle Bildzeilen zeigen jeweils eine Schicht vom Planungs-CT (links) und der 24sten Kontrollaufnahme (rechts) überlagert mit Planungskonturen des Patientenmodells und den adaptierten Konturen: a) Transversale Schicht (CT-Weichteilfenster) im Kieferbereich mit der originalen (pink) und adaptierten (gelb) Hautkontur sowie der originalen (violett) und der adaptierten (türkis) Rückenmarkskontur. b) Transversale Schicht im Larynxbereich mit selben Konturen wie a). c) Sagittale Schicht (CT-Knochenfenster) mit den beiden Rückenmarkskonturen wie a) und b). d) Frontale Schicht mit den originalen und adaptierten Konturen des Rückenmarks und zusätzlich die originale (blau) und adaptierte (grün) Konturen des rechten Schlüsselbeinknochens. e) Transversale Schicht im Schulterbereich mit den Konturen der Haut (wie a)). f) Transversale Schicht durch die obere Lunge (CT-Lungenfenster) und die originalen (blau) sowie adaptierten (grün) Lungenkonturen.



dingungen für die Planoptimierung einzusetzen, falls eine Umplanung nötig ist. Auch ein Dosismonitoring für die aktuelle Patientenanatomie vor der Fraktion mit dem applizierten Plan bedarf angepasster Konturen, so kann man überprüfen, ob und wie das DHV einzelner VOIs von der geplanten Situation abweicht.

Um diese geometrischen Anpassungen zu realisieren, kann die resultierende Transformation von deformierbaren Registrierungsverfahren genutzt werden. Da gewöhnlich das Planungs-CT als das Referenzbild  $R$  im Registrierungsprozess dient, verformt die resultierende Abbildung das aktuelle Fraktionsbild als Template  $D$  zum neuen deformierten Template  $D'$

$$T : D(i) \mapsto D'(i) \sim R(i). \quad (3.5)$$

Als Ergebnis wird  $D'$  nun dem Planungs-CT  $R$  ähnlich sein. Um die VOI Konturen aus dem Planungs-CT auf die Anatomie aus dem aktuellen Fraktions-CT zu übertragen benötigt man die zu  $T$  inverse Transformation  $T^{-1}$ . Abhängig von der VOI Repräsentation, muß  $T^{-1}$  an unterschiedlichen Positionen interpoliert werden, da  $T$  durch ein diskretes Vektorfeld repräsentiert wird.

Abhängig von der internen Repräsentation segmentierter Strukturen ergeben sich mehrere Möglichkeiten, wie diese angepasst werden können. Wenn die Strukturen (wie in der Bestrahlungsplanungssoftware VIRTUOS) als triangulierte Oberflächen definiert sind, so kann man die Punkte, welche die Oberflächendreiecke repräsentieren, transformieren. Da sich die Topologie in der Regel nicht verändert, können nach der Anpassung des Oberflächenmodells neue Konturen als Schnitt des Modells mit den CT Schichten berechnet werden. Alternativ können die Voxel, die zu einer bestimmten Struktur gehören in einem leeren Datensatz markiert werden. Nach der Transformation dieser künstlichen Bildserie können die neuen Strukturgrenzen durch eine Schwellwertbildung leicht extrahiert werden.

Für die Darstellung der adaptierten VOIs eines Kopf-Hals-Tumor Patienten in Abbildung 3.13 wurde die volumetrische Transformation ohne einer Triangulation der VOIs verwendet. Das inverse Transformationsfeld  $T^{-1}$  wurde durch die schnelle Nächste-Nachbar-Interpolation aus dem Ergebnis  $T$  des hier entwickelten deformierbaren Registrierungsverfahrens erzeugt. VIRTUOS wurde zur Visualisierung der Ergebnisse eingesetzt. Die Planungs-Konturen der VOIs sind in der Regel per Hand definiert worden. Die adaptierten Konturen sind ohne Nachverarbeitung in der Abbildung 3.13 farblich unterschiedlich den CTs überlagert dargestellt.

Die neuen Konturen zeigen die stattgefundenen lokale Positionsänderungen und globale Verformungen. Die Genauigkeit der Transformation ist für solche Anwendungen ausreichend gut, da Registrierungsfehler meistens in kontrastarmen Bildregionen zu lokalisieren sind. Die Organkonturen jedoch befinden sich meist an stark kontrastierten Gewebegrenzen, wo auch intensitätsbasierte Registrierungsverfahren genauere Ergebnisse liefern, da hier die Bildinformation leichter extrahiert werden kann. Der Einsatz des Registrierungsverfahrens erspart die Neukonturierung per Hand und benötigt abhängig von der Anzahl der zu transformierenden Konturen weniger als drei Minuten inklusive der Konturpropagation.



## Kapitel 4

# Diskussion

### Inhaltsangabe

<b>4.1 Vor- und Nachteile des Registrierungsverfahrens</b>	<b>67</b>
4.1.1 Registrierungsgenauigkeit	68
4.1.2 Rechendauer	69
4.1.3 Limitationen	69
<b>4.2 Anwendbarkeit in der Strahlentherapie</b>	<b>70</b>
4.2.1 Positionierungsunterstützung	70
4.2.2 Kontourpropagation	71
4.2.3 Voxel-to-Voxel-Mapping	71
<b>4.3 Ausblick</b>	<b>72</b>

Das hier entwickelte deformierbare Registrierungsverfahren dient der Ermittlung einer Transformation, welche die Änderung der Patientenanatomie unter der Strahlentherapie zwischen zwei aufgenommenen Zeitpunkten beschreibt. Diese Transformation kann die Grundlage für Adaptionstrategien im Therapieverlauf sein. Da Adaptionen meist sofort nach der erkannten Verschlechterung der Organposition durchgeführt werden müssen, müssen die notwendigen Rechenoperationen zur Ermittlung der Transformation möglichst schnell durchgeführt werden können. Die Anforderung an die Rechendauer liegt abhängig von der eingesetzten Therapieform im Bereich von mehreren Sekunden bis wenigen Minuten. Die Geschwindigkeit darf aber nicht auf Kosten der Genauigkeit erreicht werden. Das hier entwickelte deformierbare Registrierungsverfahren ist optimiert für die hochparallele Hardwarearchitektur der GPU, um die Rechengeschwindigkeit so weit zu steigern, dass der Zeitgewinn zur Extraktion zusätzlicher Information, die sonst aus Zeitgründen vernachlässigt wird, genutzt werden kann.

### 4.1 Vor- und Nachteile des Registrierungsverfahrens

Das hier vorgestellte deformierbare Registrierungsverfahren hat den gleichen Nachteil für den Einsatz in der ART, den alle intensitätsbasierten deformierbaren Registrierungsverfahren teilen. Es schließt kein bio-physikalisches Transformationsmodell ein, das den Suchraum allein auf realistische Transformationen einschränkt. Da die Bewertung der errechneten Transformation nach wie vor schwierig ist, bleibt die Unsicherheit beim Einsatz der Registrierungsergebnisse für nachfolgende Rechenschritte bestehen. Eine orts aufgelöste

Quantifizierung der Fehler für die Fehlerfortpflanzung ist nicht möglich. Der Vorteil allerdings, den das Fehlen eines bio-physikalischen Modells bietet, ist die kürzere Rechendauer der intensitätsbasierten Algorithmen, die trotz der höheren Unsicherheit eine stattgefundene Transformation zwischen zwei Bildern bestimmen können. So lassen sich innerhalb von Minuten Transformationen aus Bildern der aktuellen Patientenanatomie ermitteln, ohne den Therapieverlauf wesentlich zu verlängern.

Durch den Einsatz einer hochparallelen Hardwarearchitektur, wie der GPU mit CUDA, limitiert man sich auf die Nutzung einer propriäteren Hardware. Nachteilig ist hierbei die potentielle Gefahr, dass der Hersteller die Architektur wesentlich ändert oder die Hardware nicht mehr produziert. Denn eine effiziente Implementierung der algorithmischen Komponenten auf dieser Hardware ist darauf spezialisiert und kann nicht auf beliebiger Hardware zum Einsatz gebracht werden. Den immensen Vorteil jedoch bringt die Parallelisierbarkeit der Rechenschritte auf dieser Hardware. Damit gelingt eine signifikante Beschleunigung der rechenintensiven Registrierungsalgorithmen. Durch die noch kürzere Rechendauer wird der Einsatz von zeitkritischen Adaptionen im Verlauf der Therapie in jeder einzelnen Fraktion ermöglicht.

Auch wenn sich prinzipiell alle intensitätsbasierten Registrierungsverfahren stellenweise mittels GPU parallelisieren und damit beschleunigen lassen, kann die potentiell erreichbare Beschleunigung durch die Begrenzung der Abbildbarkeit der einzelnen Rechenschritte auf die Hardware unerreicht bleiben. Das hier entwickelte Registrierungsverfahren ist speziell an die Architektur der Hardware angepaßt und ermöglicht möglichst viel Information in möglichst geringer Zeit zu verarbeiten. Der Zeitgewinn durch die effiziente Parallelisierung der Rechenaufgabe, kann eingesetzt werden, um zusätzliche Informationen aus den Bildern zu extrahieren. Das wiederum kann zur Verbesserung der Genauigkeit der resultierenden Transformation beitragen.

#### 4.1.1 Registrierungsgenauigkeit

Alle Evaluierungsmethoden im Kapitel 2.2.4.1 zeigen, dass der entwickelte Algorithmus in der Strahlentherapie eingesetzt werden kann. Sowohl Landmarkenvergleiche als auch Vergleiche mit künstlichen bekannten Transformationen zeigen eine Subvoxelgenauigkeit bei 60 bis 70 % der Ortstransformationsvektoren. Verbleibende Abweichungen nach der Registrierung sind im Mittel nicht größer als 2 Voxellängen, und zeigen damit meist in das dem Zielvoxel direkt benachbarte Voxel. Wobei der Vergleich mit den künstlichen bekannten Transformationen eine etwas besser Registrierungsgenauigkeit verglichen zu dem mit den Landmarken zeigt.

Natürlich muß beachtet werden, dass Vergleiche mit Expertenlandmarken nur eingeschränkt die Genauigkeit der Registrierungsalgorithmen widerspiegeln. Denn auch die Definition von Landmarken weist eine Ungenauigkeit auf. Auch interpretieren unterschiedliche Experten gleiche Grauwertverteilungen der Bilder nicht exakt gleich, so dass Landmarkenvergleiche keine sichere Quantifizierung der Registrierungsfehler erlauben. Landmarken haben gegenüber künstlichen Deformationen der Bilder einen weiteren Nachteil, denn meist ist die Zahl der eindeutig identifizierbaren Strukturen in einer Patientenaufnahme sehr gering. Auch befinden sich Landmarken naturgemäß ausschließlich an kontrastreichen Stellen im Bild. Künstliche Deformationen können dagegen auch in homogenen Bereichen die Unsicherheit der Registrierung angeben. Ob die eingesetzten künstlichen Deformationen den tatsächlichen Bewegungsbereich der realen Patientenbewegung abdeckt,

kann nicht nachgewiesen werden, denn dafür wäre ein genaues Patientenmodell nötig, das – wenn es vorhanden wäre – auch schon in der Registrierung selbst eingesetzt werden könnte.

Auch im Vergleich zu einem alternativen Registrierungsverfahren, dem Template Matcher, der bereits gute Ergebnisse in der Registrierung deformierbarer Anatomie erzielte, zeigt der hier entwickelte Algorithmus vergleichbare Genauigkeit, wenn man den Vergleich zu Landmarken als Beschreibung der wahren Transformation heranzieht.

Alle Evaluierungstests und visuelle Eindrücke zeigen, dass der entwickelte Registrierungsalgorithmus überall dort in der Strahlentherapie eingesetzt werden kann, wo eine Registrierungsgenauigkeit von  $\pm 1$ -2 Voxellängen ausreicht. Hierbei bleibt zu untersuchen, wie man ortsaufgelöst über das gesamte Vektorfeld die Unsicherheit bestimmen kann.

#### 4.1.2 Rechendauer

Die Rechendauer des vorgestellten deformierbaren Registrierungsverfahrens kann durch die Parallelisierung der rechenintensivsten Schritte mithilfe der GPU wesentlich reduziert werden. Registrierungen von 3-dimensionalen Datensätzen, mit in der Strahlentherapie typischen Auflösungen, können, abhängig von der Größe der Datensätze in ein bis zwei Minuten durchgeführt werden. Dabei können eine hohe Zahl von etablierten Korrespondenzen berücksichtigt werden. Dadurch werden verlässliche Ergebnisse erzielt. Es können auch stark lokale Deformationen erfaßt werden, was bei globalen Interpolationsmethoden nicht möglich ist.

Die Rechenzeiten werden beim Einsatz heute verbreiteter und zukünftiger GPU Hardware weiter reduziert werden, da die Implementierung mittels CUDA mit der Anzahl der physikalisch vorhandenen Multiprozessoren linear skaliert. Alternativ liessen sich höher aufgelöste Bilder registrieren, falls eine höhere Vektorfeldauflösung benötigt wird.

#### 4.1.3 Limitationen

##### **monomodal**

Limitationen des entwickelten Registrierungsalgorithmus bei der Korrelation monomodaler Bilder sind mit dem fehlenden globalen Transformationsmodell verbunden. Da alle etablierten Korrespondenzen unabhängig vom Bildinhalt oder benachbarten Korrespondenzen propagiert werden, können Fehlkorrespondenzen die Transformation verfälschen. Unrealistische Strukturverformungen sind dann die Folge, die allerdings lokal stärker eingeschränkt sind, als beim Einsatz globaler Interpolationsmethoden. Auch Lücken in der Positionsverteilung der gefundenen Korrespondenzen können zu unrealistischen Verformungen führen, denn die benachbarte Landmarken tauschen keine Information miteinander aus.

Das stark lokale Propagationsmodell benötigt möglichst viele Korrespondenzen, um eine global zusammenhängende Bewegung möglichst gut zu beschreiben. Sind nur wenige eindeutige Korrespondenzen vorhanden, kann das Verfahren keine zufriedenstellende Resultate liefern. Auch die Extraktion der globalen rigiden Komponente, wie es in diesem Verfahren realisiert ist, benötigt möglichst viele über das gesamte Bild verteilte Korrespondenzen, soll eine realistische Transformation erfaßt werden.

### **multimodal**

Multimodale Bilder können mit diesem Verfahren, trotz des multimodal-fähigen Ähnlichkeitsmaßes, nicht registriert werden. Denn die lokale Optimierung des Ähnlichkeitsmaßes mithilfe von uneingeschränkter lokaler Bewegungen einzelner Voxel führt zwar zur eindeutigeren aber nicht zur wahren Zuweisung der Grauwertentsprechungen beider Modalitäten. Die derzeit genutzten multimodale Ähnlichkeitswerte favorisieren die eindeutige Zuteilung der Grauwerte zu einander. Ist in Bild A eine Kante mit benachbarten Grauwerten 1 und 2, und im Bild B drei verschiedene Grauwerte 1,2 und 3, sowie zwei parallele Kanten vorhanden, würden sowohl das entwickelte Ähnlichkeitsmaß als auch z.B. MI versuchen eine eindeutige Zuordnung der Werte 3 aus Bild B entweder dem Wert 1 oder 2 aus Bild A zuzuweisen, statt jeweils einen Teil von diesen Grauwerten sowohl der Gruppe 1 als auch 2 zuzuweisen. In Wirklichkeit können aber unterschiedliche Modalitäten eben unterschiedliche Information darstellen, so dass auch entgegen der Annahme im Ähnlichkeitsmaß eine geteilte Zuordnung richtig sein könnte.

## **4.2 Anwendbarkeit in der Strahlentherapie**

Die adaptive Strahlentherapie stellt hohe Anforderungen an die Genauigkeit und Geschwindigkeit der deformierbaren Registrierungsalgorithmen. Bei der Erhöhung der Geschwindigkeit kann keine Reduzierung der Genauigkeit toleriert werden. Unterschiedliche Anwendungen jedoch erfordern unterschiedliche Genauigkeit der resultierenden Transformation. Die Frage, die es stets zu beantworten gilt, ist: Ist der Gewinn durch den Einsatz der Registrierung und darauf basierender Adaption trotz Unsicherheit größer als das Risiko das durch die veränderte Anatomie für den Patienten entsteht? In vielen Fällen können deformierbare Registrierungsalgorithmen dem Therapeut helfen, die Situation schneller einzuschätzen und darauf zu reagieren.

### **4.2.1 Positionierungsunterstützung**

In der fraktionierten Therapie ist es wichtig, den Patienten vor jeder Bestrahlung richtig zu positionieren. Bildbasiert kann man die Positionierung kontrollieren. Deformierbare Registrierungsalgorithmen können eingesetzt werden, um sowohl den rigiden Transformationsanteil als auch die Deformation zu ermitteln. Daraufhin kann der rigide Transformationsanteil durch eine Zielpunktkorrektur kompensiert werden. Der verbleibende Deformationsanteil kann den Therapeut bei der Entscheidung unterstützen, ob der Patient neu gelagert werden muß oder gar eine Neuplanung erforderlich ist.

Eine Neuplanung in der Therapie wäre notwendig, wenn die geometrischen Änderungen der Patienten-anatomie starke Änderungen der geplanten Dosisverteilung nachschieben würden. Aber nicht alle Änderungen führen zu relevanten Abweichungen. Daher benötigt man räumlich verteilte präzise Informationen über die stattgefundenen geometrische Veränderungen, um im ersten Schritt einschätzen zu können, ob relevante Strukturen sich verformt oder verschoben haben. Da die Einschätzung hier auf rein geometrischen Kriterien beruht, reicht die Genauigkeit des entwickelten Registrierungsverfahrens aus, um solche Verformungsanalysen durchzuführen. Auch die Rechendauer von etwa einer Minute verlängert nicht wesentlich den Gesamtverlauf.

### 4.2.2 Kontourpropagation

Oft kann man allein aus geometrischen Verschiebungen und Verformungen der Anatomie nicht einfach auf Abweichungen von der geplanten Dosisverteilung schließen. Man benötigt eine Neuberechnung der Dosisverteilung auf der aktuellen Patientenanatomie. Die DVHs der relevanten Organe können dann mit den geplanten DVHs verglichen werden. Um die neuen DVHs auszurechnen benötigt man jedoch die neuen veränderten Konturen der relevanten Organe und Strukturen. Eine Neukonturierung per Hand kann mehrere Minuten einnehmen und lässt sich in der kurzen Zeit, die für eine Fraktion zur Verfügung steht, nicht realisieren. Ein schneller Registrierungsalgorithmus kann innerhalb von einer Minute die Deformation auf die Konturen, die auf dem Planungs-CT definiert wurden, übertragen, sodass nur noch die Dosisverteilung und die DVHs berechnet werden müssten. Die gesamte Berechnungen würden dann wenige Minuten einnehmen und sich noch in den Therapieverlauf integrieren lassen. Die Anforderungen an die Genauigkeit des Registrierungsverfahrens sind auch hier mit dem entwickelten Verfahren erfüllt, da die Konturierungen von Hand vergleichbare Schwankungen in der Konturdefinition haben.

Im Falle einer Umplanung benötigt man ebenfalls die neuen adaptierten Konturen als Input für die Neuoptimierung. Auch hier kann ein schnelles deformierbares Registrierungsverfahren die Arbeit dem Therapeuten erleichtern.

### 4.2.3 Voxel-to-Voxel-Mapping

Um auch mögliche Deformationen der Patientenanatomie während der Behandlung kompensieren zu können, benötigt man Planadaptionen. Solche Strategien können nur auf dem Wissen basieren, wie die bisher tatsächlich applizierte Dosisverteilung aussieht. Man benötigt die Transformationen, die zwischen den Kontrollaufnahmen stattgefunden haben, um die einzelnen Dosisverteilungen aus jeder Fraktion zu einer Gesamtdosisverteilung aufsummieren zu können. Anders als bei der Konturpropagation, benötigt man hier eine hohe Genauigkeit der gesamten Vektorverteilung, auch in homogenen Regionen, wo Registrierungsfehler potentiell am größten sind. Die Rekonstruktion der applizierten Dosisverteilung beruht auf einem sogenannten Voxel-to-Voxel-Mapping. Die Unsicherheit der Registrierungsergebnisse für die Transformation einzelner Voxel beeinflusst die resultierende Dosisverteilung stark, da sich mit jeder applizierten Fraktion die Fehler akkumulieren. Besonders die fehlende Quantifizierung der Registrierungsfehler in homogenen Bildbereichen können zu Unterschätzung der Fehler in der Dosisaufsummierung führen. In Bereichen mit steilen Dosisgradienten können so große Unsicherheiten entstehen.

Ob ein solches Voxel-to-Voxel-Mapping mit dem entwickelten Registrierungsverfahren sinnvoll durchgeführt werden kann, lässt sich nicht sagen, da die orts aufgelöste Fehlerquantifizierung nicht ermittelt werden kann. Um diese Barriere zu überwinden, müssen bessere Validierungsmethoden für Registrierungsverfahren entwickelt werden. Auch die Einbeziehung von bio-physikalischen Patientenmodellen in die Registrierung kann die Fehlergrenzen weiter reduzieren.

Mit dem hier entwickelten Registrierungsverfahren ist allerdings die Einbeziehung eines solchen Modells möglich. Auch die große Menge an extrahierten Bildinformationen ist eine Voraussetzung für genauere Modellierung der Bewegung, die unser Verfahren erfüllt. Insoweit kann das entwickelte Verfahren als geeignete Basis für eine Weiterentwicklung dienen. Vor allem die erreichte Rechengeschwindigkeitssteigerung durch Parallelisierung

für intensitätsbasierte Registrierung zeigt auch das Potential für die Beschleunigung einer physikalischen Patientenbewegungsmodellierung.

### 4.3 Ausblick

Die Bildregistrierung ist trotz der langen Entwicklungsgeschichte noch kein abgeschlossenes Gebiet. Nach wie vor gibt es großen Bedarf nach Weiterentwicklungen. Die Berücksichtigung der anatomischen Informationen im Bild und Patientenbewegungsmodelle werden immer mehr in den Vordergrund treten.

Mit dem hier entwickelten Ansatz lokale Korrespondenzen aus Bildern zu extrahieren und diese als Bewegung lokal zu propagieren ist es möglich, lokale Grauwertinformation in die Propagationsmodellierung mit zu integrieren. So können z.B. Grauwerte, die Knochen repräsentieren, durch die Parameter des ChainMail Algorithmus mit hoher Steifigkeit propagiert werden. Das vereinfachte physikalische Modell des ChainMail Algorithmus kann durch lokal optimierte Parameter an den anatomischen Bildinhalt ohne nennenswerten Geschwindigkeitsverlust angepaßt werden. Die benachbarten etablierten Korrespondenzen können über globalere Bedingungen miteinander gekoppelt werden, um auch die Propagation globaler Bewegungsanteile zu gewährleisten. Auch die Anzahl der etablierten Korrespondenzen kann durch lokale Optimierungen der gemeinsamen Wahrscheinlichkeitsverteilung im Ähnlichkeitsmaß weiter gesteigert werden.

Die Einbeziehung von bio-physikalischen Modellen kann bei dem entwickelten Registrierungsverfahren einfacher realisiert werden als bei intensitätsbasierten Verfahren, die als ein globales Optimierungsproblem formuliert sind. Die zahlreichen Korrespondenzen können als Inputkräfte für die Simulation genutzt werden und während der Simulation schnell mit der aktualisierten Anatomie gegengeprüft werden, falls die Parameter der Modellierung zu Abweichungen führen.

Adaptive Therapieansätze können von unserem Verfahren profitieren. Die Analyse der ermittelten Transformationen aus Patientendaten selbst und die resultierenden applizierten Dosisverteilungen können zur Ableitung von Sicherheitssäumen bei deformierbarer Anatomie bereits in der Planung dienen, wenn keine Kompensationsmethoden während der Therapie eingesetzt werden können. Hierzu müssen Registrierungsfehler in Phantomexperimenten weiter quantifiziert werden. Schnelle Registrierungsalgorithmen können auch bei Gating und Trackingstrategien in Kombination mit Tumormarkern eingesetzt werden. So können auch Bewegungen von Risikoorganen detektiert werden und der Strahl kann abgeschaltet werden, wenn ein Risikoorgan in den Strahlengang rückt. Besonders die Transformationsmodellierung einzelner Regionen, für die Gating oder Trackingstrategie geeignet sind, kann im Registrierungsverfahren aufgebaut werden.

## Kapitel 5

# Schlussfolgerung

Das entwickelte deformierbare Registrierungsverfahren erfüllt die wesentlichen Anforderungen der adaptiven Strahlentherapie. Die Laufzeit der rechenintensiven algorithmischen Komponenten kann durch die Hardware-basierte Parallelisierung auf der GPU signifikant verringert werden, ohne dabei die erreichbare Genauigkeit rein intensitätsbasierter Registrierungsverfahren einzuschränken. Anders als bisher entwickelte intensitätsbasierte Registrierungsverfahren ermöglicht das vorgestellte Verfahren die zukünftige Einbindung von bio-physikalischen Modellen, um die verbleibenden Limitationen aufzuheben.

Viele Anwendungen in der adaptiven Strahlentherapie können von der Geschwindigkeit des Verfahrens profitieren. In der bildgestützten Therapie ermöglicht es die schnelle Lagerungsanalyse der Patienten noch vor der Bestrahlung, sowie die schnelle Adaption der Konturen zur dosimetrischen Lagerungskontrolle oder für eine Neuplanung. Durch die Anwendung von schnellen deformierbaren Registrierungsverfahren zur Ermittlung der anatomischen Veränderungen noch während der Therapie ermöglicht man adaptive Korrekturstrategien und damit die Vollausschöpfung der Bildinformationen im Sinne des ALARA-Prinzips. Durch die erhöhte Präzision der Positionsbestimmung relevanter anatomischer Strukturen können auch hochkonforme Strahlentherapietechniken effizienter eingesetzt werden.

Die Entwicklung optimal parallelisierter deformierbarer Registrierungsverfahren auf der GPU eröffnet die Möglichkeit mehr Informationen gleichzeitig zu verarbeiten und so dem Therapeut schneller die wesentlichen Informationen zur Verfügung zu stellen, die seine Entscheidungen im Strahlentherapieverlauf zugunsten des Patienten beeinflussen.

# Abbildungsverzeichnis

2.1	Holthusenplot: TCP vs. NTCP als Funktion der Dosis . . . . .	6
2.2	Beispiele für Dosisverteilungen in der Bestrahlungsplanung . . . . .	7
2.3	Beispiel für Patientenfixierungssysteme . . . . .	9
2.4	Multimodale Bildfusion als Hilfe für Volumendefinition . . . . .	14
2.5	Gaußpyramide im Multi-Resolution Schema . . . . .	20
2.6	Das gemeinsame Histogramm als Basis für die Mutual Information . . . . .	23
2.7	Anschauliches Beispiel für TPS Interpolation . . . . .	26
2.8	Funktionsweise des ChainMail Algorithmus . . . . .	28
2.9	Bildfusion als Registrierungsevaluierungshilfe . . . . .	31
2.10	Vektorfeldvisualisierung in der Registrierungsevaluierung . . . . .	32
2.11	Verformtes Gitter zur Registrierungsevaluierung . . . . .	32
2.12	Beispiele für Landmarkenidentifikation . . . . .	33
3.1	Gemeinsame Histogramme multimodaler Aufnahmen . . . . .	45
3.2	Schema eines gemeinsamen Histogramms unter Transformation . . . . .	47
3.3	Schematische Zeichnung der Korrespondenzsuche . . . . .	47
3.4	Einfluß der rigiden Transformationskomponente . . . . .	49
3.5	Bewegungspropagation nach dem 3D ChainMail Modell . . . . .	50
3.6	Visualisierung einer künstlich erzeugten TPS Transformation . . . . .	53
3.7	Registrierungsergebniss der künstlich erzeugten TPS Bilder . . . . .	54
3.8	Quantitative Registrierungsergebnisse im Landmarkenvergleich . . . . .	57
3.9	Limitierung durch Fehlkorrespondenzen . . . . .	59
3.10	Limitierung durch Lücken der lokalen Deformation . . . . .	60
3.11	Limitierung bei multimodaler Registrierung . . . . .	61
3.12	Deformationsanalyse von Patientendaten . . . . .	63
3.13	Organrekonturierung auf deformierter Anatomie mittels DIR . . . . .	64

# Tabellenverzeichnis

3.1	Registrierungsunsicherheit im Vergleich zu künstlichen Feldern . . . . .	54
3.2	Registrierungsunsicherheit im Vergleich zu definierten Landmarken . . . . .	55
3.3	Registrierungsunsicherheit im Vergleich zu rigiden Teilregistrierungen . . .	56
3.4	Registrierungslaufzeiten . . . . .	58

# Literaturverzeichnis

- [1] H. Holthusen, “Erfahrungen über die Verträglichkeitsgrenze für Röntgenstrahlen und deren Nutzen zur Verhütung von Schäden,” *Strahlentherapie und Onkologie*, vol. 57, pp. 30–36, 1936.
- [2] R. Boesecke, J. Doll, B. Bauer, W. Schlegel, O. Pastyr, and W. J. Lorenz, “Treatment planning for conformation therapy using a multi-leaf collimator,” *Strahlentherapie Onkologie*, vol. 164, no. 3, pp. 151–154, 1988.
- [3] T. Bortfeld, “IMRT: a review and preview,” *Physics in Medicine and Biology*, vol. 51, pp. R363–379, 2006.
- [4] A. Sovik, E. Malinen, and D. R. Olsen, “Strategies for biologic image-guided dose escalation: A review,” *International Journal of Radiation Oncology, Biology and Physics*, vol. 73, no. 3, pp. 650–658, 2009.
- [5] W. Schlegel and A. Mahr, eds., *3D Conformal Radiation Therapy: A multimedia introduction to methods and techniques*. Springer Verlag Berlin Heidelberg.
- [6] U. Oelfke and T. Bortfeld, “Inverse planning for photon and proton beams,” *Medical Dosimetry*, vol. 26, no. 2, pp. 113–124, 2001.
- [7] E. M. Stoiber, G. Lechsel, K. Giske, M. W. Muenther, A. Hoess, R. Bendl, J. Debus, P. E. Huber, and C. Thieke, “Quantitative assessment of image-guided radiotherapy for paraspinal tumors,” *International Journal of Radiation Oncology, Biology and Physics*, vol. 75, no. 3, pp. 933–940, 2009.
- [8] K. Giske, E. M. Stoiber, M. Schwarz, A. Stoll, M. W. Muenther, C. Timke, F. Roeder, J. Debus, P. E. Huber, C. Thieke, and R. Bendl, “Local setup errors in image-guided radiotherapy for head and neck cancer patients immobilized with a custom-made device,” *International Journal of Radiation Oncology, Biology and Physics*, vol. in press, 2010.
- [9] J. C. Stroom and B. J. Heijmen, “Geometrical uncertainties, radiotherapy planning margins, and the ICRU-62 report,” *Radiotherapy and Oncology*, vol. 64, no. 1, pp. 75–83, 2002.
- [10] A. McKenzie, M. van Herk, and B. Mijnheer, “Margins for geometric uncertainty around organs at risk in radiotherapy,” *Radiation and Oncology*, vol. 62, pp. 299–307, 2002.

- [11] T. R. Mackie, J. Kapatoes, K. Ruchala, W. Lu, C. Wu, G. Olivera, L. Forrest, W. Tome, J. Welsh, R. Jeraj, P. Harari, P. Reckwerdt, B. Paliwal, M. R. abd Harry Keller, J. Fowler, and M. Mehta, "Image guidance for precise conformal radiotherapy," *International Journal of Radiation Oncology, Biology and Physics*, vol. 56, no. 1, pp. 89–105, 2003.
- [12] D. Verellen, M. D. Ridder, N. Linthout, K. Tournel, G. Soete, and G. Storme, "Innovations in image-guided radiotherapy," *Nature Reviews Cancer*, vol. 7, pp. 949–960, 2007.
- [13] M. van Herk, "Different styles of image-guided radiotherapy," *Seminars in Radiation Oncology*, vol. 17, no. 4, pp. 258–267, 2007.
- [14] M. van Herk, "Will IGRT live up to its promise?," *Acta Oncologica*, vol. 47, pp. 1186–1187, 2008.
- [15] G. X. Ding and C. W. Coffey, "Radiation dose from kilovoltage cone beam computed tomography in an image-guided radiotherapy procedure," *International Journal of Radiation Oncology, Biology and Physics*, vol. 73, no. 2, pp. 610–617, 2009.
- [16] M. Kapanen, J. Collan, K. Saarilahti, J. Heikkonen, K. Kairemo, and M. Tenhunen, "Accuracy requirements for head and neck intensity-modulated radiation therapy based on observed dose response of the major salivary glands," *Radiotherapy and Oncology*, vol. 93, pp. 109–114, 2009.
- [17] N. J. Yue, J. P. S. Knisely, H. Song, and R. Nath, "A method to implement full six-degree target shift corrections for rigid body in image-guided radiotherapy," *Medical Physics*, vol. 33, no. 1, pp. 21–31, 2006.
- [18] S. Webb, "Motion effects in (intensitiy modulated) radiation therapy: a review," *Physics in Medicine and Biology*, vol. 51, pp. R403–R425, 2006.
- [19] T. R. Willoughby, P. A. Kupelian, J. Pouliot, K. Shinohara, M. Aubin, M. Roach, L. L. Skrumeda, J. M. Balter, D. W. Litzenberg, S. W. Hadley, T. Wei, and H. M. Sandler, "Target localization and real-time tracking using the calypso 4d localization system in patients with localized prostate cancer," *International Journal of Radiation Oncology, Biology and Physics*, vol. 65, no. 2, pp. 528–534, 2006.
- [20] E. Nioutsikou, Y. Seppenwoolde, J. R. Symonds-Tayler, B. Heijmen, P. Evans, and S. Webb, "Dosimetric investigation of lung tumor motion compensation with a robotic respiratory tracking system: an experimental study," *Medical Physics*, vol. 35, no. 4, pp. 1232–40, 2008.
- [21] M. B. Tacke, S. Nill, A. Krauss, and U. Oelfke, "Real-time tumor tracking: automatic compensation of target motion using the siemens 160 mlc," *Medical Physics*, vol. 37, no. 2, pp. 753–61, 2010.
- [22] P. R. Poulsen, B. Cho, A. Sawant, D. Ruan, and P. J. Keall, "Detailed analysis of latencies in image-based dynamic mlc tracking," *Medical Physics*, vol. 37, no. 9, pp. 4998–5005, 2010.

- [23] R. Mohan, X. Zhang, H. Wang, Y. Kang, X. Wang, H. Liu, K. Ang, D. Kuban, and L. Dong, "Use of deformed intensity distributions for on-line modification of image-guided IMRT to account for interfractional anatomic changes," *International Journal of Radiation Oncology, Biology and Physics*, vol. 61, no. 4, pp. 1258–1266, 2005.
- [24] G. G. Zhang, T.-C. Huang, K. M. Forster, K.-P. Lin, C. Stevens, E. Harris, and T. Guerrero, "Dose mapping: Validation in 4D dosimetry with measurements and application in radiotherapy follow-up evaluation," *Computer Methods and Programs in Biomedicine*, vol. 90, pp. 25–37, 2008.
- [25] G. Starkschall, K. Britton, M. F. McAleer, M. D. Jeter, M. R. Kaus, K. Bzdusek, R. Mohan, and J. D. Cox, "Potential dosimetric benefits of four-dimensional radiation treatment planning," *International Journal of Radiation Oncology, Biology and Physics*, vol. 73, no. 5, pp. 1560–1565, 2009.
- [26] G. Janssens, J. O. de Xivry, S. Fekkes, A. Dekker, P. Lambin, and W. van Elmpt, "Evaluation of nonrigid registration models for interfraction dose accumulation in radiotherapy," *Medical Physics*, vol. 36, no. 9, pp. 4268–4276, 2009.
- [27] P. J. Slomka and R. P. Baum, "Multimodality image registration with software: state-of-the-art," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 36, no. Supplement 1, pp. 44–55, 2009.
- [28] M. van Herk, P. Remeijer, C. Rasch, and J. V. Lebesque, "The probability of correct target dosage: Dose-population histograms for deriving treatment margins in radiotherapy," *International Journal of Radiation Oncology, Biology and Physics*, vol. 47, no. 4, pp. 1121–1135, 2000.
- [29] B. Polat, J. Wilbert, K. Baier, M. Flentje, and M. Guckenberger, "Nonrigid patient setup errors in the head-and-neck region," *Strahlentherapie und Onkologie*, vol. 9, pp. 506–511, 2007.
- [30] S. van Kranen, S. van Beek, C. Rasch, M. van Herk, and J.-J. Sonke, "Setup uncertainties of anatomical sub-regions in head-and-neck cancer patients after offline CBCT guidance," *International Journal of Radiation Oncology, Biology and Physics*, vol. 73, no. 5, pp. 1566–1573, 2009.
- [31] C. Karger, O. Jaekel, J. Debus, S. Kuhn, and G. H. Hartmann, "Three-dimensional accuracy and interfractional reproducibility of patient fixation and positioning using a stereotactic head mask system," *International Journal of Radiation Oncology, Biology and Physics*, vol. 49, pp. 1493–1504, 2001.
- [32] P. M. Evans, C. Coolens, and E. Nioutsikou, "Effects of averaging over motion and the resulting systematic errors in radiation therapy," *Physics in Medicine and Biology*, vol. 51, pp. N1–N7, 2006.
- [33] S. van Beek, S. van Kranen, A. Mencarelli, P. Remeijer, C. Rasch, M. van Herk, and J.-J. Sonke, "First clinical experience with a multiple region of interest registration and correction method in radiotherapy of head-and-neck cancer patients," *Radiotherapy and Oncology*, vol. 94, no. 2, pp. 213–7, 2010.

- [34] B. W. Raaymakers, J. J. Lagendijk, J. Overweg, J. G. Kok, A. J. Raaijmakers, E. M. Kerkhof, R. W. van der Put, I. Meijnsing, S. P. C. nad F. Benedosso, M. van Vulpen, C. H. de Graaff, J. Allen, and K. J. Brown, "Integrating a 1.5 T MRI scanner with a 6 MV accelerator: proof of concept," *Physics in Medicine and Biology*, vol. 54, no. 12, pp. N229–N237, 2009.
- [35] U. Malsch, C. Thieke, P. E. Huber, and R. Bendl, "An enhanced block matching algorithm for fast elastic registration in adaptive radiotherapy," *Physics in Medicine and Biology*, vol. 51, no. 19, pp. 4789–4806, 2006.
- [36] S. Thrnqvist, J. B. Petersen, M. Hyer, L. N. Bentzen, and L. P. Muren, "Propagation of target and organ at risk contours in radiotherapy of prostate cancer using deformable image registration," *Acta Oncologica*, vol. 49, no. 7, pp. 1023–32, 2010.
- [37] E. E. Ahunbay, C. Peng, A. Godley, C. Schultz, and X. A. Li, "An on-line replanning method for head and neck adaptive radiotherapy," *Med Phys*, vol. 36, no. 10, pp. 4776–90, 2009.
- [38] P. M. af Rosenschld, M. C. Aznar, D. E. Nygaard, G. F. Persson, S. S. Korreman, S. A. Engelholm, and H. Nystrm, "A treatment planning study of the potential of geometrical tracking for intensity modulated proton therapy of lung cancer," *Acta Oncol. 2010 Oct;49(7):1141-8.*, vol. 49, no. 7, pp. 1141–8, 2010.
- [39] M. Velec, J. L. M. and C. L. Eccles, T. Craig, M. B. Sharpe, L. A. Dawson, and K. K. Brock, "Effect of breathing motion on radiotherapy dose accumulation in the abdomen using deformable registration," *International Journal of Radiation Oncology, Biology and Physics*, vol. in press, 2010.
- [40] M. Guckenberger, J. Wilbert, A. Richter, K. Baier, and M. Flentje, "Potential of adaptive radiotherapy to escalate the radiation dose in combined radiochemotherapy for locally advanced non-small cell lung cancer," *International Journal of Radiation Oncology, Biology and Physics*, vol. in press, 2010.
- [41] H. Zhong, Elisabeth Weiss, and J. V. Siebers, "Assessment of dose reconstruction errors in image-guided radiation therapy," *Physics in Medicine and Biology*, vol. 53, pp. 719–736, 2008.
- [42] R. Bansal, L. H. Staib, A. F. Laine, D. Xu, J. Liu, L. F. Posecion, and B. S. Peterson, "Calculation of the confidence intervals for transformation parameters in the registration of medical images," *Medical Image Analysis*, vol. 13, pp. 215–233, 2009.
- [43] W. Crum, L. Griffin, D. Hill, and D.J.Hawkes, "Zen and the art of medical image registration: correspondence, homology, and quality," *NeuroImage*, vol. 20, pp. 1425–1437, 2003.
- [44] B. Fischer and J. Modersitzki, "Ill-posed medicine - an introduction to image registration," *Inverse Problems*, vol. 24, pp. 16–31, JUN 2008.
- [45] J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, no. 1, pp. 1–37, 1998.
- [46] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Physics in Medicine and Biology*, vol. 46, pp. R1–R45, 2001.

- [47] V. Walimbe and R. Shekhar, "Automatic elastic image registration by interpolation of 3D rotations and translations from discrete rigid-body transformations," *Medical Image Analysis*, vol. 10, pp. 899–914, 2006.
- [48] M. Shn, M. Birkner, Y. Chi, J. Wang, D. Yan, B. Berger, and M. Alber, "Model-independent, multimodality deformable image registration by local matching of anatomical features and minimization of elastic energy," *Medical Physics*, vol. 35, no. 3, pp. 866–878, 2008.
- [49] R. Bajcsy and S. Kovacic, "Multiresolution elastic matching," *Computer Vision, Graphics, and Image Processing*, vol. 46, pp. 1–21, 1989.
- [50] P. Thevenaz and M. Unser, "Optimization of mutual information for multiresolution image registration," *IEEE Transactions on Image Processing*, vol. 9, no. 12, pp. 2083–2099, 2000.
- [51] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual information matching in multiresolution contexts," *Image and Vision Computing*, vol. 19, pp. 45–52, 2001.
- [52] W. Peckar, C. Schnrr, K. Rohr, and H. S. Stiehl, "Non-rigid image registration using a parameter-free elastic model," in *British Machine Vision Conference*, 1998.
- [53] J. P. Pluim, J. A. Maintz, and M. A. Viergever, "Interpolation artefacts in mutual information-based image registration," *Computer Vision and Image Understanding*, vol. 77, pp. 211–232, 2000.
- [54] O. Salvado and D. L. Wilson, "Removal of local and biased global maxima in intensity-based registration," *Medical Image Analysis*, vol. 11, no. 2, pp. 183–196, 2007.
- [55] J. Weese, P. Roesch, T. Netsch, T. Blaffert, and M. Quist, "Gray-value based registration of CT and MR images by maximization of local correlation," *MICCAI99, Lecture notes in computer science*, vol. 1679, pp. 656–663, 1999.
- [56] F. Maes, A. Collignon, D. Vandeemeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions in Medical Imaging*, vol. 16, pp. 187–198, 1997.
- [57] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.
- [58] E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens, "A viscous fluid model for multimodal non-rigid image registration using mutual information," *Medical Image Analysis*, vol. 7, pp. 565–575, 2003.
- [59] M. R. Sabuncu and P. J. Ramadge, "Spatial information in entropy-based image registration," in *Biomedical Image Registration (WBIR 2003)*, pp. 132–141, 2003.
- [60] M. E. Leventon and W. E. L. Grimson, "Multi-modal volume registration using joint intensity distributions," *Lecture Notes on Computer Science*, vol. MICCAI98, p. 10571066, 1998.

- [61] J. Orchard, "Multimodal image registration using floating regressors in the joint intensity scatter plot.," *Medical Image Analysis*, vol. 12, pp. 385–96, August 2008.
- [62] S. Clippe, D. Sarrut, C. Malet, S. Miguet, C. Ginestet, and C. Carrie, "Patient setup error measurement using 3D intensity-based image registration techniques," *International Journal of Radiation Oncology, Biology and Physics*, vol. 56, no. 1, pp. 259–265, 2003.
- [63] T. Sederberg and S. Parry, "Free form deformation of solid geometry models," *SIGGRAPH*, vol. 20, pp. 151–60, 1986.
- [64] W. Lu, M.-L. Chen, G. H. Olivera, K. J. Ruchala, and T. R. Mackie, "Fast free-form-deformable registration via calculus of variations," *Physics in Medicine and Biology*, vol. 49, pp. 3067–3087, 2004.
- [65] F. L. Bookstein, "Principal warps: Thin-plate splines and decomposition of deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 6, pp. 567–585, 1989.
- [66] M. Fornefett, K. Rohr, and H. S. Stiehl, "Radial basis functions with compact support for elastic registration of medical images," *Image and Vision Computing*, vol. 19, pp. 87–96, 2001.
- [67] E. M. Vasques-Osorio, M. S. Hoogeman, L. Bondar, P. C. Levendag, and B. J. M. Heijmen, "A novel flexible framework with automatic feature correspondence optimization for nonrigid registration in radiotherapy," *Medical Physics*, vol. 36, no. 7, pp. 2848–2859, 2009.
- [68] R. L. Hardy, "Multiquadric equations of topography and other irregular surfaces," *Journal of Geophysical Research*, vol. 76, no. 8, pp. 1905–1915, 1971.
- [69] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Transactions on Image Processing*, vol. 5, no. 10, pp. 1435–1447, 1996.
- [70] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "3D brain mapping using a deformable neuroanatomy," *Physics in Medicine and Biology*, vol. 39, pp. 609–618, 1994.
- [71] K.-S. Choi, H. Sun, P.-A. Heng, and J. Zou, "Deformable simulation using force propagation model with finite element optimization," *Computers & Graphics*, vol. 28, pp. 559–568, 2004.
- [72] S. F. F. Gibson, "3D chainmail: a fast algorithm for deforming volumetric objects," tech. rep., Mitsubishi Electric Research Laboratories, 1996.
- [73] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Computer Journal*, vol. 7, pp. 308–313, 1965.
- [74] G. E. Christensen, X. Geng, J. G. Kuhl, J. Bruss, T. Grabowski, I. A. Pirwani, M. W. Vannier, J. S. Allen, and H. Damasio, "Introduction to the non-rigid image registration evaluation project (NIREP)," in *3rd International Workshop on Biomedical Image Registration*, vol. LCNS 4057, pp. 128–135, Springer-Verlag, 2006.

- [75] R. Kashani, M. Hub, J. M. Balter, M. L. Kessler, L. Dong, L. Zhang, L. Xing, Y. Xie, D. Hawkes, J. A. Schnabel, J. McClelland, S. Joshi, Q. Chen, and W. Lu, "Objective assessment of deformable image registration in radiotherapy: A multi-institution study," *Medical Physics*, vol. 35, no. 12, pp. 5944–5953, 2008.
- [76] K. K. Brock and the Deformable Registration Accuracy Consortium, "Results of a multi-institution deformable registration accuracy study (MIDRAS)," *International Journal of Radiation Oncology, Biology, Physics*, vol. 76, no. 2, pp. 583–596, 2010.
- [77] P. Castadot, J. A. Lee, A. Parraga, X. Geets, B. Macq, and V. Gregoire, "Comparison of 12 deformable registration strategies in adaptive radiation therapy for the treatment of head and neck tumors," *Radiotherapy and Oncology*, vol. 89, pp. 1–12, October 2008.
- [78] A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, J. H. Song, M. Jenkinson, C. Lepage, D. Riuekert, P. Thompson, T. Vercauteren, R. P. Woods, J. J. Mann, and R. V. Parsey, "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration," *Neuroimage*, vol. 46, no. 3, pp. 786–802, 2009. Epub 2009 Jan 13.
- [79] R. Castillo, E. Castillo, R. Guerra, V. E. Johnson, T. McPhail, A. K. Garg, and T. Guerrero, "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," *Physics in Medicine and Biology*, vol. 54, pp. 1849–1870, 2009.
- [80] M. Serban, E. Heath, G. Stroian, D. L. Collins, and J. Seuntjens, "A deformable phantom for 4D radiotherapy verification: Design and image registration evaluation," *Medical Physics*, vol. 35, no. 3, pp. 1094–1102, 2008.
- [81] J. Vandermeulebroucke, D. Sarrut, and P. Clarysse, "The POPI-model, a point-validated pixel-based breathing thorax model.," in *XVth International Conference on the Use of Computers in Radiation Therapy (ICCR), Toronto, Canada, 2007*.
- [82] H. Zhong, T. Peters, and J. V. Siebers, "FEM-based evaluation of deformable image registration for radiation therapy," *Physics in Medicine and Biology*, vol. 52, no. 16, pp. 4721–4738, 2007.
- [83] H. Zhong, J. Kim, and I. J. Chetty, "Analysis of deformable image registration accuracy using computational modeling," *Medical Physics*, vol. 37, pp. 970–979, 2010.
- [84] S. S. Samant, J. Xia, P. Muyan-zcelik, and J. Owens, "High performance computing for deformable image registration: Towards a new paradigm in adaptive radiotherapy," *Med Phys*, vol. 35, no. 88, pp. 3546–3553, 2008.
- [85] S. Asano, T. Maruyama, and Y. Yamaguchi, "Performance comparison of FPGA, GPU and CPU in image processing," in *FPL 2009*, pp. 126–131, 2009.
- [86] *NVIDIA CUDA Programming Guide*.
- [87] <http://gpgpu.org/>.
- [88] [http://www.nvidia.de/object/cuda\\_home\\_new\\_de.html](http://www.nvidia.de/object/cuda_home_new_de.html).

- [89] M. D. McCool, Z. Qin, and T. S. Popa, “Shader metaprogramming,” in *SIGGRAPH*, 2002.
- [90] I. Buck, T. Foley, D. Horn, J. Sugerman, K. Fatahalian, M. Houston, and P. Hanrahan, “Brook for GPUs: Stream computing on graphics hardware,” in *SIGGRAPH*, 2004.
- [91] G. I. Egri, Z. Fodor, C. Hoelbling, S. D. Katz, D. Nogradi, and K. Szabo, “Lattice QCD as a video game,” *Computer Physics Communications*, vol. 177, no. 8, p. 631639, 2007. arXiv:heap-lat/0611022v2.
- [92] Z. Taylor, O. Comas, M. Cheng, J. Passenger, D. Hawkes, D. Atkinson, and S. Ourselin, “On modelling of anisotropic viscoelasticity for soft tissue simulation: Numerical solution and GPU execution,” *Medical Image Analysis*, vol. 13, pp. 234–244, 2009.
- [93] T. Rohlfing and C. R. Maurer, “Nonrigid image registration in shared-memory multiprocessor environment with application to brains, breast, and bees,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 7, no. 1, pp. 16–25, 2003.
- [94] R. Strzodka, M. Droske, and M. Rumpf, “Image registration by a regularized gradient flow. a streaming implementation in DX9 graphics hardware,” *Computing*, vol. 73, pp. 373–389, 2004.
- [95] G. C. Sharp, H. Kandasamy, H. Singh, and M. Folkert, “GPU-based streaming architectures for fast cone-beam CT image reconstruction and demons deformable registration,” *Phys Med Biol*, vol. 52, pp. 5771–5783, 2007.
- [96] K. O. Noe, B. D. DeSenneville, U. V. Elstrom, T. Sorensen, and T. S. Sorensen, “Acceleration and validation of optical flow based deformable registration for image-guided radiotherapy,” *Acta Oncologica*, vol. 47, pp. 1286–1293, 2008.
- [97] K. O. Noe, K. Tanderup, J. C. Lindegaard, C. Grau, and T. S. Sorensen, “GPU accelerated viscous-fluid deformable registration for radiotherapy,” *Studies in Health Technology and Informatics*, vol. 132, pp. 327–32, 2008.
- [98] T. ur Rehman, E. Haber, G. Pryor, J. Melonakos, and A. Tannenbaum, “3D nonrigid registration via optimal mass transport on the GPU,” *Medical Image Analysis*, vol. 13, pp. 931–940, 2009.
- [99] X. Gu, H. Pan, Y. Liang, R. Castillo, D. Yang, D. Choi, E. Castillo, A. Majumdar, T. Guerrero, and S. B. Jiang, “Implementation and evaluation of various demons deformable image registration algorithms on a GPU,” *Phys*, vol. 55, pp. 207–219, 2010.
- [100] R. Shams, P. Sadeghi, R. Kennedy, and R. Hartley, “Parallel computation of mutual information on the GPU with application to real-time registration of 3D medical images,” *Computer Methods and Programs in Biomedicine*, vol. 99, pp. 133–146, 2010.
- [101] R. Bendl, A. Hoess, and W. Schlegel, *Computer Vision, Virtual Reality and Robotics in Medicine*, vol. 905 of *Lecture Notes in Computer Science*, ch. Virtual simulation in radiotherapy planning, pp. 287–292. Springer Berlin / Heidelberg, 1995.

- [102] V. Pekar, E. Gladilin, and K. Rohr, "An adaptive irregular grid approach for 3D deformable image registration," *Physics in Medicine and Biology*, vol. 51, pp. 361–377, 2006.
- [103] A. Andronache, M. von Siebenthal, G. Szekely, and P. Cattin, "Non-rigid registration of multi-modal images using both mutual information and cross-correlation," *Medical Image Analysis*, vol. 12, no. 1, pp. 3–15, 2007.

# Danksagung

Mein besonderer Dank gilt Prof. Rolf Bendl für die gute Betreuung: Danke für ein spannendes Thema, für die unzähligen wertvollen Diskussionen, für die Freiheit meine Vorstellungen durchsetzen zu können, für die aufgebrauchte Geduld, wo sie mir gefehlt hat, für die Tipps und Tricks mit VIRTUOS, für das kurzfristige Gegenlesen von Beiträgen, für die Unterstützung bei der Auseinandersetzung mit Koautoren und Reviewern, sowie viele der schwer in Begriffe fassbaren Dinge, die eben eine gute Betreuung ausmachen und die ich nie hätte missen wollen.

An dieser Stelle danke ich auch Prof. Uwe Oelfke für die Übernahme der Begutachtung und Fakultätsvertretung sowie kritische Fragen und aufmunternde Kommentare beim Freitagsseminar.

Prof. Schlegel danke ich für die freundliche Aufnahme in seine Abteilung. Aber im Besonderen für seine Unterstützung und die Ratschläge bezüglich meiner Weiterentwicklung und der kurzfristigen Gespräche trotz seines vollen Terminkalenders.

Ich danke meinen Zimmerkollegen: Gerhard Lechsel, der mich in meiner Anfangszeit eingearbeitet hat und aus dessen fortgeführten Projekten ich ebenfalls profitieren konnte; Michael Schwarz für die gegenseitigen Diskussionen am Whiteboard, sowie solche die über das Fachliche hinausgingen. Besonders sei auch Eva Stoiber gedankt: ich habe die fruchtbare Kooperation nicht nur sehr genossen, sondern von den Einblicken von der strahlentherapeutisch-ärztlichen Seite immens profitiert. Herrn Prof. Bendl an dieser Stelle vor allem für das Überlassen des großen Tisches samt des neueren PCs ☺.

Silvia Handlos, Andrea Fränze, Armin Stoll, Gernot Wurst, Jan Rüppell, Sven Skobovsky, Jens Schwanke, Janina Bräunling und die schon genannten e0403er habe ich für ganz unterschiedliche Dinge zu danken, wie das Korrekturlesen und die Mittagskaffeesgespräche, die Ausflüge und Mensagänge, die Kuchen, die Tapas, die Dinosaurier, rege Beteiligung am Cake&Discuss, sowie Spiel- und Grillabende.

Sven Skobovsky, Florian Kronberger, Franziska Herget und Florian Dittmann danke ich für ihre guten Studien-, Bachelor- und Diplomarbeiten - für ihre Kreativität und vor allem für ihre Selbständigkeit ☺.

Auch gruppenübergreifend habe ich am DKFZ nur eine freundliche Atmosphäre erfahren. Danke an Peter Ziegenhein, Andrea Schwahofer, Asja Pfaffenberger, Mathies Breithaupt und Wolfram Stiller für die hilfreichen und die mehr oder weniger sinnvollen Ge-

sprache.

Mein Dank gebührt auch den Strahlentherapeuten der KKE um Prof. Huber für die Zusammenarbeit und das zur-Verfügung-stellen von Patientendatensätzen.

Viele Sorgen und Datentransfer- sowie Computerprobleme haben mir andere Leute am DKFZ abgenommen. An dieser Stelle sei unserem IT-Team, Herrn Müller und Herrn Döttling, herzlich gedankt, wo ich mit meinen Problemen stets offene Türen vorfand.

Ich danke Stefanie Lindemann, Silvia Handlos und Simone Barthold-Beß, die bei vielen organisatorischen Fragen mit Rat und Tat und guter Laune zur Seite standen.

Ich danke dem Medizin-Physiker-Team der Kopfklinik um Frau Sroka-Perez, vor allem Gerald und Oliver für die geduldigen Erklärungen und Ratschläge bei der QA des alten LB2. Kai danke ich für die Mühe mit den Tomodatensätzen.

Ich danke meiner Familie für die ewig-währende Unterstützung. Im Besonderen gilt mein Dank Christian - nicht nur für die Geduld und liebevolle Unterstützung, sondern auch für den Feinschliff meiner  $\TeX$ -Styles.