

# Medizinische Bildanalyse <sup>1</sup>

Annika Schlechter und Stephan Lachnit

Abstract: Thanks to Medical Image Analysis doctors can have a detailed look inside the human body. In these experiments we take a look at how we can transform the data from CT and MRT scans into one 3-dimensional picture. As it turns out, the quality and computing time of the transformations are highly dependent on a variety of parameters, which we had to optimize to get a satisfying result.

---

<sup>1</sup>Versuch F95, ausgeführt am 3. & 4.2.19, Betreuer Rene-Philipp Austermuehl, lange besondere Auswertung

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>2</b>
<b>2</b>	<b>Grundlagen</b>	<b>2</b>
2.1	Bildmodalitäten . . . . .	2
2.1.1	MRT-Bildgebung . . . . .	2
2.1.2	CT-Bildgebung . . . . .	3
2.2	Bildregistrierung . . . . .	3
2.2.1	Ähnlichkeitsmaße . . . . .	3
2.2.2	Transformationen . . . . .	4
<b>3</b>	<b>Versuchsdurchführung</b>	<b>5</b>
3.1	Ermittlung einer Kamerabewegung . . . . .	5
3.1.1	Aufbau . . . . .	5
3.1.2	Bildaufnahme mit der Kamera . . . . .	5
3.1.3	Paarweise Registrierung der Aufnahmen . . . . .	5
3.2	Ähnlichkeit bei monomodalen und multimodalen Bilddaten . . . . .	5
3.2.1	Aufnahme der Ähnlichkeit im monomodalen Fall . . . . .	5
3.2.2	Aufnahme der Ähnlichkeit im multimodalen Fall . . . . .	6
3.3	Rigide 3D-Registrierung von medizinischen Bilddaten . . . . .	6
3.3.1	Implementierung des Algorithmus . . . . .	6
3.3.2	Anwendung auf die Testdaten . . . . .	6
3.4	Deformierbare Registrierung . . . . .	6
<b>4</b>	<b>Auswertung</b>	<b>7</b>
4.1	Ermittlung der Kamerabewegung aus einer Bildfolge . . . . .	7
4.2	Ähnlichkeit bei monomodalen Bilddaten . . . . .	7
4.3	Ähnlichkeit bei multimodalen Bilddaten . . . . .	8
4.4	Rigide 3D-Bildregistrierung von medizinischen Bilddaten . . . . .	9
4.5	Deformierbare Registrierung . . . . .	10
<b>5</b>	<b>Diskussion</b>	<b>11</b>
<b>6</b>	<b>Literatur</b>	<b>12</b>
<b>7</b>	<b>Appendix</b>	<b>12</b>

# 1 Einleitung

In der Medizin wird die medizinische Bildanalyse benötigt, um aus verschiedenen Aufnahmetechniken ein möglichst genaues Bild des Patienten zu gewinnen. Diese Informationen sind wichtig, um das Krankheitsbild der Patienten möglichst gut zu diagnostizieren.

In diesem Versuch soll von einfachen Translationen einer Kamera bis zu einer Auswertung von echten Patientendaten die Grundlagen verschiedener Techniken der medizinischen Bildanalyse erarbeitet werden.

## 2 Grundlagen

### 2.1 Bildmodalitäten

In der medizinischen Bildgebung gibt es verschiedene Bildgebungstechniken, welche je nach Technik unterschiedliche Informationen besser darstellen. Besonders wichtig sind die Magnetresonanztomographie (MRT) und Computertomographie (CT) zur Diagnostizierung, welche wir im folgenden genauer betrachten wollen.

#### 2.1.1 MRT-Bildgebung

Für die Magnetresonanztomographie sind der Spin, der Drehimpuls und das magnetisches Moment des Protons grundlegend. Wenn ein Magnetfeld angelegt wird, können durch diese Eigenschaften die Lage der Rotationsachse des Protons, der induzierte Magnetfeldvektor sowie dessen Änderungen beobachtet werden.

Die Ausrichtung der Spins wird mit der Larmorfrequenz  $\omega_0$  beschrieben. Diese ist die Frequenz, welche die Präzessionsbewegung eines Protons in einem äußeren Magnetfeld  $B_0$  beschreibt:  $\omega_0 = \gamma_0 \cdot B_0$ , wobei  $\gamma_0$  das gyromagnetische Verhältnis beschreibt.

Mit einer elektromagnetischen Welle mit Larmorfrequenz kann dem System Energie hinzugefügt werden. Da in einem äußeren Magnetfeld in  $z$ -Richtung fast alle Spins parallel in  $z$ -Richtung ausgerichtet sind, kippen diese durch solch eine Anregung aus der  $z$ -Achse. Das MR-Signal, auf welchem die MRT-Bildgebung beruht, entsteht durch einen  $90^\circ$ -Puls. Dabei kippen die Spins in die  $x$ - $y$ -Ebene und induzieren eine Wechselspannung mit der Larmorfrequenz in der Empfangsspule.

Da der Ausgangszustand (parallele Ausrichtung in  $z$ -Richtung) stabiler ist, fallen die Spins in diesen mit der Zeit zurück. Dies geschieht durch die Spin-Gitter- und die Spin-Spin-Wechselwirkung oder auch T1- beziehungsweise T2-Relaxation genannt.

Die T1-Relaxation wird auch longitudinale Relaxation genannt, da durch die Ausrichtung der Spins entlang des Magnetfeldes die longitudinale Magnetisierung zunimmt. Die zugehörige Zeitkonstante T1 ist abhängig von der Magnetfeldstärke und vom Material. Die Spin-Spin Wechselwirkung mit der zugehörigen Zeitkonstante T2 trägt ebenso zur Abnahme der transversalen Magnetisierung bei. Hierbei entsteht eine Dephasierung der Spins durch einen Energieaustausch der Spins untereinander. Auch Inhomogenitäten des Magnetfeldes verstärken den Zerfall der Phasenkohärenz. Diese entstehen einerseits durch das Gerät andererseits durch den Patienten im Magnetfeld. Mit Hilfe von spezielle Sequenzen kann dieser Effekt unterdrückt werden.

Die Anzahl anregbarer Spins pro Volumeneinheit, die T1-Zeit (Zeit bis sich die Spins nach einer Anregung wieder entlang des Magnetfeldes ausgerichtet sind) und die T2-Zeit (Zeit bis das Signal nach einer Anregung abklingt aufgrund der aus der Phase laufenden Spins) bestimmen

die Intensität des Gewebes auf einem MRT-Bild. Somit entsteht je nach Parameter eine andere Eigenschaft auf dem MRT-Bild. Für die Aufnahme eines dreidimensionalen Bildes wird den MR-Signalen noch Koordinaten zugeordnet: Man benötigt eine zusätzliche Magnetspule für einen Gradienten entlang der  $z$ -Richtung um eine Auflösung der  $z$ -Koordinate zu gewährleisten. Da die Larmorfrequenz abhängig von dem angelegten Magnetfeld ist, sind entlang dieses Gradienten auch die Larmorfrequenzen unterschiedlich. Je nach Frequenz wird nun nur eine Schicht des Körpers angeregt und somit können je nach Stärke des Gradienten feinere oder gröbere Schichten aufgelöst werden. Auch in der  $x$ - und  $y$ -Richtung wird ein zusätzlicher Gradient mit einer Magnetspule erzeugt. In  $y$ -Richtung kreisen durch die unterschiedlichen Larmorfrequenzen die Spins unterschiedlich schnell und erzeugen einen Phasengradienten. Wenn nun kurz danach dieser Gradient ausgeschaltet wird, bewegen sich die Spins gleich schnell, allerdings in unterschiedlicher Phase und können somit ihrer Position in  $y$ -Richtung zugeordnet werden. Durch einen Gradienten in  $x$ -Richtung besteht das MR-Signal nicht mehr aus nur aus einer Frequenz sondern einem Frequenzspektrum, da die Spins unterschiedlich schnell präzedieren.

### 2.1.2 CT-Bildgebung

Der Computertomograph (CT) besteht aus einer Röntgenröhre mit Blenden, welche die Röntgenstrahlung kontrolliert zum Patienten leiten und einem Detektor. Dieser detektiert die Energie der Strahlung, welche durch das Gewebe abgeschwächt wurde. Nun werden viele Bilder aus unterschiedlichen Perspektiven aufgenommen durch Drehungen der Röntgenröhre und des Detektors. Die unterschiedlichen Absorptionscharakteristiken der Gewebeschichten resultiert in Kontrasten auf den Bildern.

## 2.2 Bildregistrierung

Die Informationen dieser zwei verschiedenen Bildgebungssysteme werden miteinander kombiniert um eine Diagnose zu stellen oder medizinische Eingriffe zu planen. Um diese Daten geometrisch zu verknüpfen benötigt es die sogenannte Bildregistrierung. Die Bildregistrierung beruht darauf, dass ein Datensatz als Referenz verwendet wird und für den zweiten Datensatz eine Transformation gesucht wird, die diesen mit dem ersten optimal überlagert. In dieser Transformation wird für jeden Bildpunkt berechnet wohin dieser im transformierten Bild liegt. Daneben benötigt man ein Interpolationsverfahren um die Intensitätswerte an den Pixel-Positionen im transformierten Bild, da ein Bild aufgrund der Pixel diskretisiert ist.

Da es sich hier um ein inverses Problem handelt, weil die Transformation nicht bekannt ist, sondern nur das Referenzbild und das transformierte Bild (bewegtes Bild), kann dies nur mit Hilfe eines Optimierungsverfahren gelöst werden. Der Algorithmus der Bildregistrierung basiert auf einer Metrik, welche die Ähnlichkeit zweier Bilddatensätze beschreibt und einer Transformation, die die Ähnlichkeit zwischen bewegtem Bild und Referenzbild erhöht. Hinzu kommen das Interpolationsverfahren und ein Optimierungsalgorithmus.

### 2.2.1 Ähnlichkeitsmaße

Zwar gibt es neben dem intensitätsbasierten auch das landmarkenbasierte Maß, jedoch wollen wir uns hier auf Ersteres konzentrieren. Hierbei werden die Intensitätswerte der Bilder genutzt um die Ähnlichkeit zu bestimmen. Es wird nun zwischen einer Metrik, welche die Bildintensitäten vergleicht, und einer, welche nur statistische Informationen der Bilder nutzt, unterschieden. Es muss jedoch beachtet werden, dass die erstere Metrik nur im monomodalen Fall genutzt werden

kann da nur dort garantiert ist, dass gleiche Informationen durch gleiche Intensitätskontraste dargestellt werden.

Mit Hilfe der mittleren quadratischen Abweichung (Mean Squared Difference, MSD) kann man die Bildintensitäten direkt vergleichen:

$$M_{MSD} = \frac{1}{N} \sum_{i=1}^N (R(x_i) - M^{\mathcal{T}}(x_i))^2 \quad (1)$$

$N$  ist die Anzahl betrachteter Punkte,  $R$  das Referenzbild,  $x_i$  die Voxelkoordinaten und  $M^{\mathcal{T}}$  das bewegte Bild, welches mit der Transformation  $\mathcal{T}$  in das Koordinatensystem des Referenzbildes transformiert wurde.

Ein anderes Ähnlichkeitsmaß ist Mutual Information (MI), welches genutzt wird falls CT- und MRT-Aufnahmen miteinander verglichen werden sollen, da hier die Kontraste der Bildintensitäten nicht direkt verglichen werden können. Bei diesem Ähnlichkeitsmaß werden nur statistische Informationen der Bildintensitäten verwendet:

$$M_{MI} = H(I_A) + H(I_B) - H(I_A, I_B) \quad (2)$$

wobei die marginale Entropie  $H(I)$  der Bilddaten  $I$  wie folgt definiert ist:

$$H(I) = -K \sum_a p_I(a) \cdot \log p_I(a) \quad (3)$$

mit der Wahrscheinlichkeit  $p(a)$  der Intensität  $a$  und einer positiven Konstante  $K$ , die abhängig von der gewählten Basis des Logarithmus ist. Die gemeinsame Entropie  $H(I_A, I_B)$  ist analog definiert, allerdings ist die Wahrscheinlichkeit  $p_{I_A, I_B}(a, b)$  nun entsprechend zweidimensional. Die Wahrscheinlichkeit wird mit Hilfe von einem (zweidimensionalen) Histogramms berechnet, auf dem die Häufigkeiten der Intensitäten  $a$  in gruppierte Bins eingetragen werden.

### 2.2.2 Transformationen

Man unterscheidet zwischen der Transformation der Bildintensitäten und der geometrischen Transformation, welche wir genauer betrachten wollen. Hier wird zusätzlich differenziert zwischen einer lokalen oder globalen Transformation der Bilddaten. Die globale Transformation basiert darauf, dass sie für alle Bildpunkte gleich ist, die lokale Transformation ist für jeden Punkt verschieden. Ein einfaches Modell der globalen Transformation ist eine Translation mit einer Rotation. Hierbei gibt es 3 Freiheitsgrade für die Translation und 3 für die Rotation. Da sich hier der Abstand zwischen zwei Bildpunkten nicht ändert, bezeichnet man eine solche Transformation als rigide. Erweitert man diese Transformation um eine mögliche Änderung der Skalierungen und Scherungen, so erhöht sich die Zahl der zu optimierenden Parameter auf 12 und man spricht von einer affinen Transformation. Für die transformierten Koordinaten  $(x', y', z')$  eines Punktes  $(x, y, z)$  gilt nun folgende Darstellung:

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \\ p_{13} & p_{14} & p_{15} & p_{16} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (4)$$

Eine deformierbare Transformation wird hingegen beschrieben mittels eines Transformationsfeldes. Hier ist die Transformation eines Voxels abhängig von dessen Position im Bildvolumen. Um

nicht für jedes Voxel eigene Transformationsparameter wählen zu müssen, kann ein grobes Gitter definiert werden. An den Gitterpunkten wird die Transformation berechnet und anschließend werden für die dazwischenliegenden Punkte die Parameter interpoliert, da die Transformation kontinuierlich sein soll. Um eine solche deformierbare Transformation zu beschleunigen, wird zunächst ein grobes Gitter gewählt und später diese Gitterpunkte genutzt um auf einem feineren Gitter zu optimieren.

## **3 Versuchsdurchführung**

### **3.1 Ermittlung einer Kamerabewegung**

Im ersten Versuchsteil soll mit Hilfe einer Bildfolge die Translationsbewegung einer Kamera ermittelt werden.

#### **3.1.1 Aufbau**

Der Versuchsaufbau besteht aus einer USB-Kamera, welche auf einer Schiene für die Kameraführung befestigt ist. Die Kamera ist an einem Computer angeschlossen, wie auch im Versuchsprotokoll zu erkennen. An diesem sollen nun mit Hilfe von Matlab die Bilder ausgewertet werden. Die Kamerafunktionalität wird anschließend getestet indem das Bild als Video angezeigt wird und die Kamera auf der Schiene bewegt wird.

#### **3.1.2 Bildaufnahme mit der Kamera**

Die Kamera wird auf eine Wand mit möglichst markanten Gegenständen ausgerichtet. Von dieser Wand werden mehrere Bilder mit verschiedenen Positionen der Kamera auf der Schiene aufgenommen. Diese Abstände werden im Versuchsprotokoll notiert.

#### **3.1.3 Paarweise Registrierung der Aufnahmen**

Diese Bilder werden nun in Matlab mit Hilfe eines Algorithmus für eine reine Translation paarweise zueinander registriert und die resultierenden Bilder gespeichert. Zusätzlich werden die durch Matlab ermittelten Translationparameter notiert.

### **3.2 Ähnlichkeit bei monomodalen und multimodalen Bilddaten**

Nun sollen Ähnlichkeitsmaße genauer betrachtet werden. Mit der MSD- und MI-Metrik wird untersucht wie sich die Ähnlichkeit von der verwendeten Transformation ändert. Dafür werden 3D-Bilddatensätze aus CT und MRT Aufnahmen des Kopfes genutzt. Mit Hilfe des Trainingsdatensatzes wurden bereits die optimalen Transformationsparameter bestimmt. Diese Parameter werden nun sowohl im monomodalen als auch im multimodalen Fall betrachtet werden:

#### **3.2.1 Aufnahme der Ähnlichkeit im monomodalen Fall**

Zunächst ist im monomodalen Fall das Referenzvolumen und das zu registrierende Volumen dasselbe, in unserem Fall ist das zu untersuchende Volumen das CT-Volumen. Im Programm wird dafür auf die optimalen Werten ein Shift addiert und anschließend die Metrikwerte berechnet.

### **3.2.2 Aufnahme der Ähnlichkeit im multimodalen Fall**

Nun soll als zu registrierendes Volumen das MR-T2-Volumen gewählt werden. Neben der Variation der Parameter wird zusätzlich die Anzahl der Histogramm-Bins im Code verändert.

## **3.3 Rigide 3D-Registrierung von medizinischen Bilddaten**

Mit einem Algorithmus sollen in diesem Versuchsteil CT- und MR-T2 Daten registriert werden um anschließend eine optimale Transformation zu ermitteln.

### **3.3.1 Implementierung des Algorithmus**

In diesem Algorithmus wird zur Registrierung Mutual Information zusammen mit einem Quasi-Newton-Optimierer verwendet. Der Algorithmus kann mit verschiedenen Parametern optimiert werden. Wir betrachten dabei die Anzahl der Registrierungsstufen, wobei bei jeder Stufe ein Auflösungsfaktor und Glättungsfaktor gesetzt werden. Anhand eines Trainingbildes haben wir diese Parameter von Hand optimiert.

### **3.3.2 Anwendung auf die Testdaten**

Der Registrierungsalgorithmus mit den optimalen Parameter der Trainingsdaten soll nun auf die Testdaten angewandt werden, um ein Bild davon zu bekommen wie gut die bestimmten Parametern bei verschiedenen Daten funktionieren.

## **3.4 Deformierbare Registrierung**

Nachdem die ersten drei Versuchsteile sich nur mit rigiden Transformationen beschäftigt haben, soll nun die deformierbare Registrierung anhand zweier Beispiele betrachtet werden.

Zuerst soll ein weißes Rechteck auf schwarzem Hintergrund zu einem weißen Kreis auf schwarzem Hintergrund deformiert werden. Besonders soll bei diesem Teil der Einfluss der Anzahl der Gitterpunkte des Transformationsgitters auf das Registrierungsergebnis und die Laufzeit der Registrierung untersucht werden. Danach soll das optimale Transformationsfeld auf ein Schachbrettmuster angewandt werden. Ein Schachbrettmuster ist besonders hilfreich, da man gut erkennen kann, wohin einzelne Bereiche des Bildes transformiert werden.

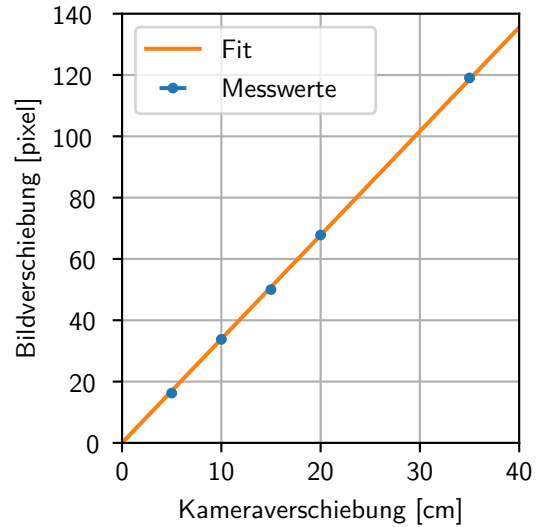
## 4 Auswertung

### 4.1 Ermittlung der Kamerabewegung aus einer Bildfolge

Das Translationsverhältnis zwischen Bewegung der Kamera und Verschiebung in Pixeln wurde über die Transformationsmatrix in Matlab ausgelesen. Zusätzlich haben wir das Ergebnis testweise an einer Bildüberlagerung überprüft. Die Messergebnisse des Verhältnisses sind in Abbildung 1 dargestellt. Wir haben  $(3.39 \pm 0.01)$  pixel/cm gemessen, wobei der Fehler rein aus der Regression stammt, für welche mögliche Fehler der Verschiebungen und des Transformationsverfahrens nicht in Betracht gezogen wurden.

Betrachtet man Abbildung 8 (Appendix), so ist leicht zu sehen, dass die Transformation bei 65 cm nicht mehr funktioniert hat. Dies könnte unter anderem dadurch entstanden sein, dass sich bei einer Translation von einer 3-dimensionalen Umgebung auch die Perspektive ändert, je entfernter die Umgebung im Vergleich zur Translation, desto geringer ist der Effekt. Den Effekt kann man gut sehen wenn man zum Beispiel den Stuhl im Bild bei 35 cm Translation mit dem Bild bei 5 cm vergleicht. Im Ersteren erkennt man an der Lehne deutliche Unterschiede in der Überlagerung (lila beziehungsweise grün), was im Letzteren kaum sichtbar ist. Der grüne Balken am rechten Rand kommt daher, da dieser Bereich nicht im Aufnahmebereich des bewegten Bildes liegt.

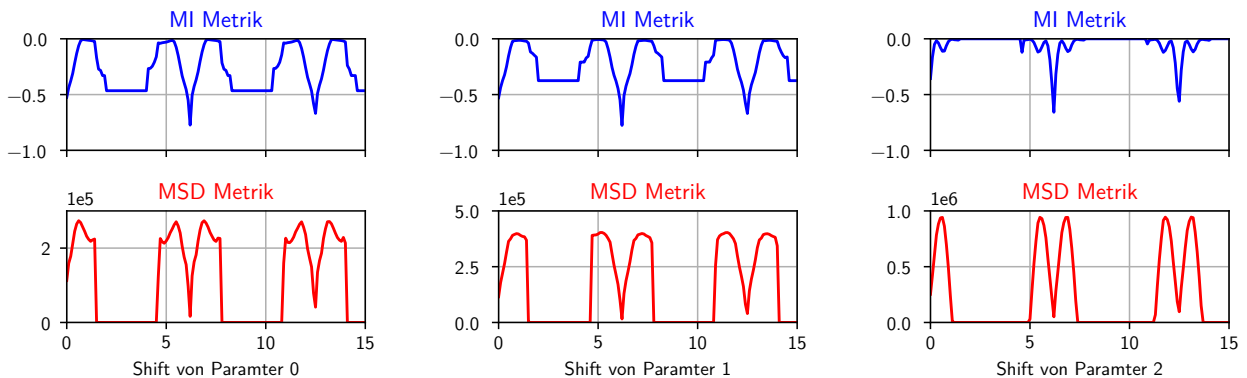
Abbildung 1: Die Translation in Pixeln eines entfernten Objektes im Bild ist linear zur parallelen Verschiebung einer Kamera



### 4.2 Ähnlichkeit bei monomodalen Bilddaten

In Abbildung 2 sind die Variationen der drei verschiedenen Rotationsparameter dargestellt. An der Periodizität von  $2\pi$  der Parameter 0, 1 und 2 erkennt man, dass es sich hierbei um die Rotationsparameter handelt. Ebenfalls gut zu erkennen ist, dass das Minimum der Metrik bei 0 bzw.  $2\pi$  und  $4\pi$  ist, wobei das Plateau zwischen den Minima vermutlich dadurch entstanden ist, dass der Algorithmus keine Ähnlichkeit mehr feststellen konnte.

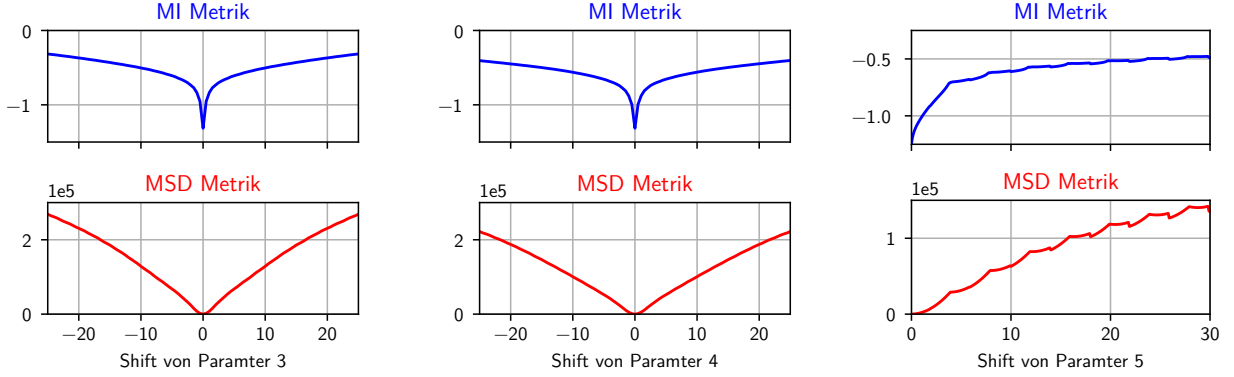
Abbildung 2: Metrik in Abhängigkeit der Rotationsparameter





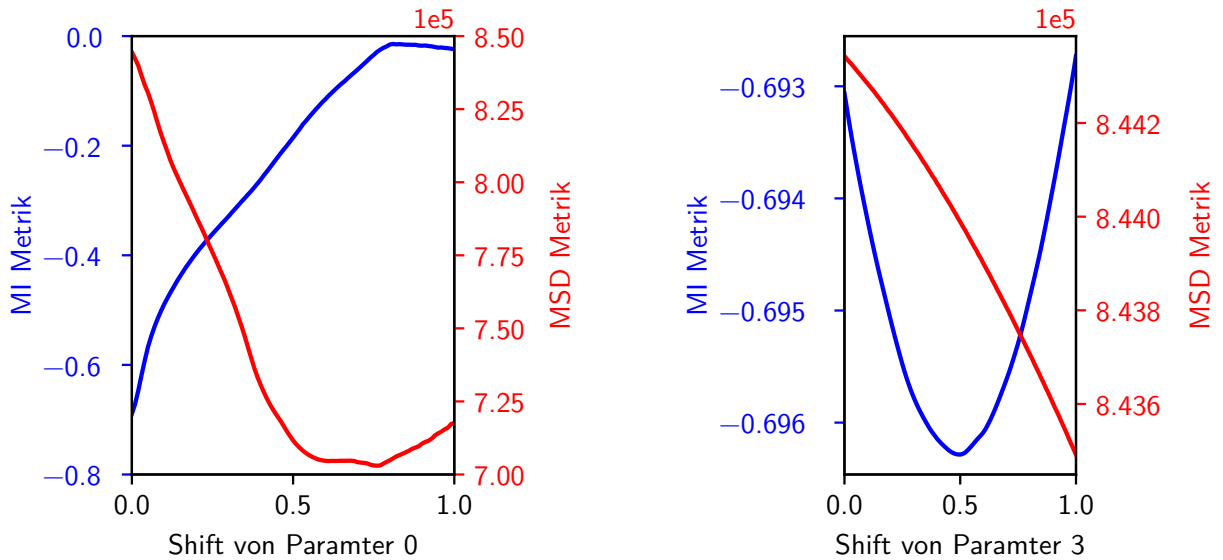
In Abbildung 3 sind die Variationen der drei Translationsparameter dargestellt. Die Parameter 3 und 4 sind die Translationsparameter für die x- und y-Achse, während Parameter 5 der Translationsparameter für die z-Achse ist. Dies erkennt man daran, dass die z-Achse kleine Plateaus aufweist, die auf die geringere Auflösung im Vergleich zur x- bzw. y-Achse zurückzuführen sind. Hier bleibt eine kleine Verschiebung der Metrik im gleichen Voxel.

Abbildung 3: Metrik in Abhängigkeit der Translationsparameter



### 4.3 Ähnlichkeit bei multimodalen Bilddaten

Abbildung 4: Metrikverhalten bei Überlagerung von CT und MRT Daten



In Abbildung 4 ist das Verhalten der beiden Metriken für einen Rotationsparameter und einen Translationsparameter dargestellt. Da CT und MRT Scans unterschiedliche Kontrastverhältnisse haben, erwarten wir keine sinnvollen Ergebnisse von der MSD Metrik (rot). Dies ist bei Parameter 0 gut erkennbar, da der Metrikwert bei stärkerer Abweichung vom optimalen Parameter abnimmt. Der Wert der MI Metrik (blau) steigt dagegen bei größerem Shift, was bedeutet, dass die MI Metrik trotz unterschiedlicher Messmethoden korrekt funktioniert. Bei Parameter 3 handelt es sich um einen Translationsparameter, was bedeutet, dass wir lediglich um einen Voxel insgesamt verschieben, was sehr wenig ist. Die MI

Metrik hat ihr Minimum bei 0.5 statt 0.0, allerdings liegt der Unterschied des Metrikwertes bei etwa 0.003, was verglichen mit den Werten beim Translationsparameter sehr gering ist. Die MSD Metrik kann bei so einer kleinen Skala nicht sinnvoll analysiert werden.

Verändert man die Anzahl der Bins der MI Metrik, das heißt man verfeinert das Histogramm zur Bestimmung der Wahrscheinlichkeit für Gleichung 3. Wie man an Abbildung 5 erkennt, verschiebt sich vor allem das Minimum nahe des optimalen Wertes bei 0.0, außerdem erkennt man, dass die Kurven mit hoher Anzahl an Bins etwas mehr Konturen in der Nähe des Wertes 0.75 aufweisen. Es ist anzumerken, dass bei einer Anzahl von 5 Bins die Metrik keine Ergebnisse mehr geliefert hat.

Verwendet man mehr Bins, so kann der Algorithmus genauer zwischen den Intensitätswerten unterscheiden, so dass die Metrik besser minimiert werden kann. Je höher die Anzahl der Bins, desto näher kann man an das Optimum gelangen, daher verbessert sich das Ergebnis bei hoher Binzahl nur noch geringfügig.

#### 4.4 Rigide 3D-Bildregistrierung von medizinischen Bilddaten

In diesem Teil haben wir MRT Scans auf CT Scans transformiert, und dabei die Qualität der Transformation und die benötigte Rechenzeit beobachtet.

Wir fanden gute Ergebnisse bei drei Registrierungsstufen mit Shrinkfaktoren von 4, 2 & 1, so wie einem Glättungsfaktor von 2 für alle Stufen bei einer Sampling-Rate von 50 %.

In Abbildung 9 sind Überlagerungen mit der genannten guten Strategie, so wie einer Berechnung ohne zusätzliche Registrierungsstufen bei einer vollen Samplingrate zu sehen, die deutlich länger zum Berechnen benötigt hat. Bei genauen Betrachten erkennt man Qualitätsunterschiede, etwa im Fenster links oben im unteren Bereich des Gehirns, oder im Fenster oben rechts an den Augen. Es stellt sich also heraus, dass die Qualität nicht durch die Samplingrate bestimmt wird, obwohl diese einen großen Einfluss auf die Rechenzeit hat, denn an Abbildung 6 erkennt man, dass die Zeit zur Berechnung in etwa linear zur Samplingrate (also zur Anzahl der Bildpunkte) ist.

Abbildung 5: Variation der Binanzahl beim MI Algorithmus führt bei hoher Anzahl nur noch zu geringen Veränderungen der Metrik

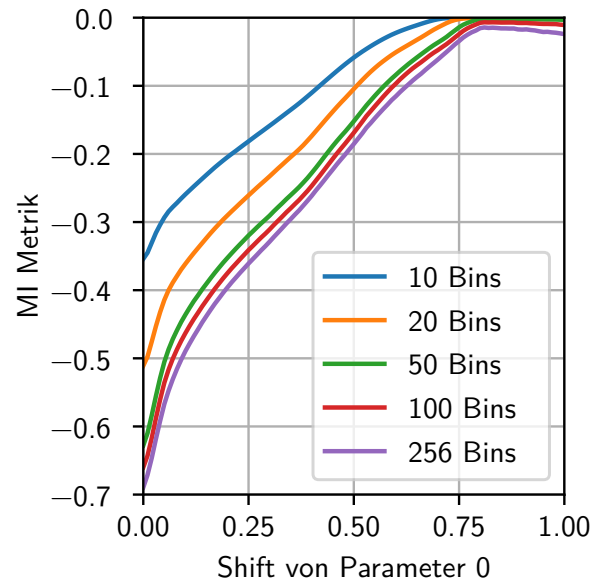
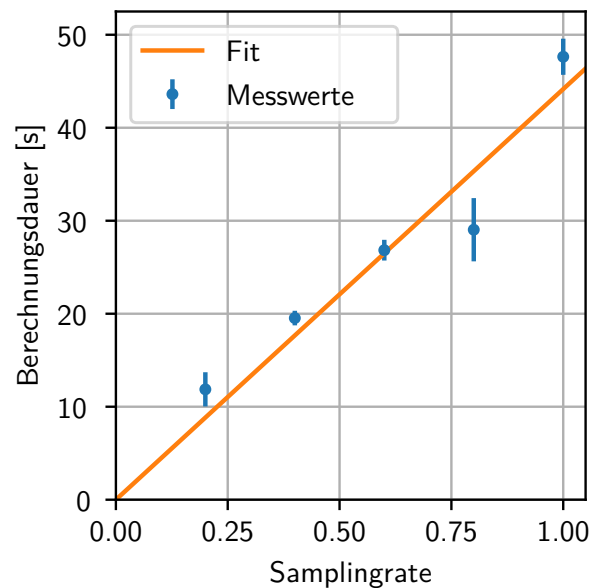


Abbildung 6: Die Berechnungsdauer steigt quasi linear mit der Samplingrate des Algorithmus



## 4.5 Deformierbare Registrierung

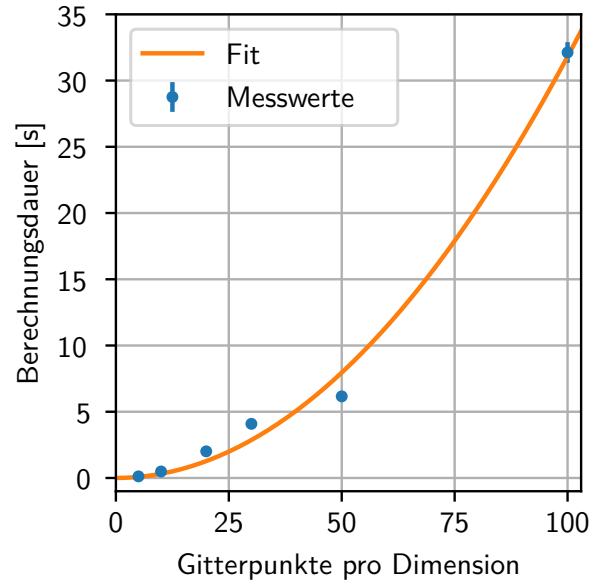
In Abbildung 10 (Appendix) sind das bewegte Bild und das fixierte Zielbild abgebildet. Das optimale Registrierungsverfahren um das Quadrat auf den Kreis zu transformieren lag bei uns bei 30 gppd (grid points per dimension, Gitterpunkte pro Dimension), wobei die gleichen Registrierungsstufen wie zuvor verwendet wurden. Es ist erkennbar, dass bei niedrigen Auflösungen der Kreis nicht rund ist, und am Rand graue Artefakte zu sehen sind. In diesen Bereichen kann der Algorithmus keine Werte finden. Im Fall von 5 gppd kann man sich etwa vorstellen, dass in den Ecken nur ein einzelner Gitterpunkt im schwarzen Bereich ist und dieser in Richtung Mitte gezogen wird, so dass im transformierten Bild am Rand keine Informationen verfügbar sind. Ist die Anzahl der Gitterpunkte zu hoch, so ergibt sich ein anderes Problem, die Gitterpunkte werden zu nah aneinander transformiert. Gut erkennbar ist das bei der Transformation mit 100 gppd, also wo jeder Pixel auch ein Gitterpunkt ist.

An den Ecken des Quadrats müssen mehrere Pixel auf die Ränder des Kreises transformiert werden, was zu Informationsverlust an diesen Stellen führt, weshalb die grauen Bereiche um den transformierten Kreis entstehen.

Um die Transformation besser sichtbar zu machen, haben wir die Transformation mit den optimalen Parametern auf ein Schachbrett angewendet. Man erkennt gut, dass die äußeren und inneren Quadrate von der Transformation nicht beeinflusst werden. Da unser Schachbrett 10 Quadrate pro Dimension hat, kommen auf jedes Quadrat 9 Gitterpunkte. An den Quadraten in den Ecken, die noch transformiert wurden, erkennt man gut, dass zwischen den Gitterpunkten interpoliert wird.

Wir haben auch die benötigte Rechenzeit mit der Anzahl der Gitterpunkte pro Dimension verglichen, die Messergebnisse sind in Abbildung 7 dargestellt. Wie nach dem vorherigen Teil zu erwarten, ist die benötigte Rechenzeit des Algorithmus linear zur Anzahl der Gitterpunkte. Bei einem 2-dimensionalen Bild ist der Verlauf also quadratisch zur Anzahl der Gitterpunkte pro Dimension.

Abbildung 7: Berechnungsdauer der deformierbaren Registrierung in Abhängigkeit der Anzahl der Gitterpunkte pro Dimension



## 5 Diskussion

Im ersten Versuchsteil war sehr anschaulich zu sehen, wie die Bildregistrierung einer einfachen Translation funktionieren. Wir haben herausgefunden, dass die Registrierung bei kleinen Translationen am besten funktioniert, da bei größeren Abständen aufgrund des räumlichen Effekts die Ähnlichkeit schlechter festgestellt werden kann.

Das Verschiebungsverhältnis zwischen fixiertem und bewegtem Bild ist konstant und lag bei uns bei  $(3.39 \pm 0.01)$  pixel/cm.

Im zweiten Teil haben wir die verschiedenen Metriken untersucht. Im monomodalen Fall konnten wir zwischen Rotations- und Translationsparametern anhand der Periodizität unterscheiden. Bei den Rotationsparametern ist zu sehen, dass die Metriken um Vielfache von  $2\pi$  ihr Minimum haben und dazwischen der Algorithmus keine Ähnlichkeit mehr feststellen kann.

Im multimodalen Fall konnten wir am Rotationsparameter gut erkennen, dass die MSD Metrik wie zu erwarten keine sinnvollen Ergebnisse mehr liefert, während die MI Metrik noch einen sinnvollen Verlauf hat.

Beim Translationsparameter haben wir leider nur einen sehr kleinen Shift berechnet, dadurch konnten wir dies dort nicht wirklich erkennen. Allerdings kann man am Verlauf der MI Metrik feststellen, dass der vorgegebene optimale Wert um etwa einen halben Voxel von dem Metrikminimum abweicht. Das liegt daran, dass die optimalen Parameter von Hand bestimmt wurden. Bei der Variation der Anzahl der Bins der MI Metrik, ist zu erkennen, dass eine größere Anzahl positive Auswirkungen auf die Genauigkeit des Algorithmus hat, wobei der Effekt mit steigender Anzahl abnimmt.

Im Versuchsteil über die Rigid 3D-Registrierungen war das Finden der optimalen Registrierungsstrategie die Hauptaufgabe. Aufgrund der begrenzten Zeit, die wir hatten um uns in die Programmbibliothek einzulesen, bestand dieser Teil hauptsächlich aus Ausprobieren. Hatte man einmal eine gute Strategie, hat diese auch bei unterschiedlichen Patientendaten gute Ergebnisse geliefert.

Im letzten Teil wurde die deformierbare Transformation genauer betrachtet. Zunächst haben wir eine gute Auflösung des Transformationsfeldes gesucht, und dies anhand einer Transformation eines Rechtecks auf einen Kreis untersucht. Dabei haben wir festgestellt, dass zu viele oder zu wenige Gitterpunkte bei der Transformationen zu unerwünschten Ergebnissen führen. Nachdem wir einen guten Wert gefunden haben, haben wir diese Transformation auf ein Schachbrettmuster angewandt, um sie besser zu verdeutlichen.

Als Fazit können wir festhalten, dass der Versuch gut durchzuführen war und gut aufgezeigt, wie Physik in anderen Bereichen, wie in diesem Fall der Medizin, Anwendungen findet. Besonders realitätsnah war hierbei die Verwendung von medizinischen Patientenbildern. Der Teil mit der deformierbaren Transformation hat besonders anschaulich gezeigt, dass das Themengebiet aber auch darüber hinausgehen kann.

## 6 Literatur

[1] F95-Versuchsanleitung Version 1.5

## 7 Appendix

Abbildung 8: Überlagerung bei verschiedenen Translationen der Kamera



(a) 5 cm



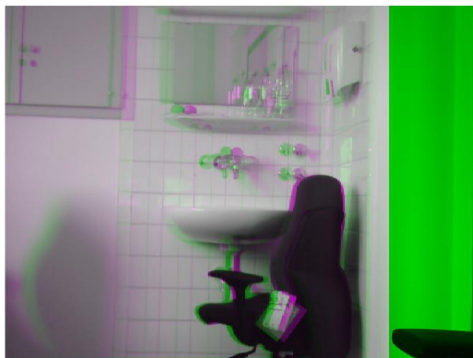
(b) 10 cm



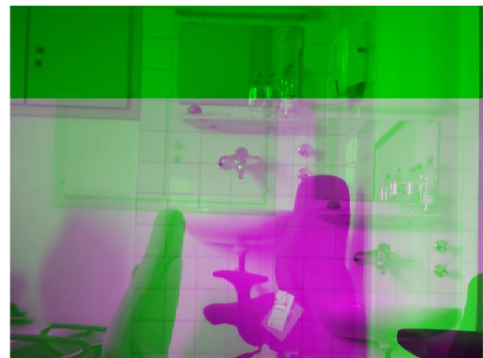
(c) 15 cm



(d) 20 cm

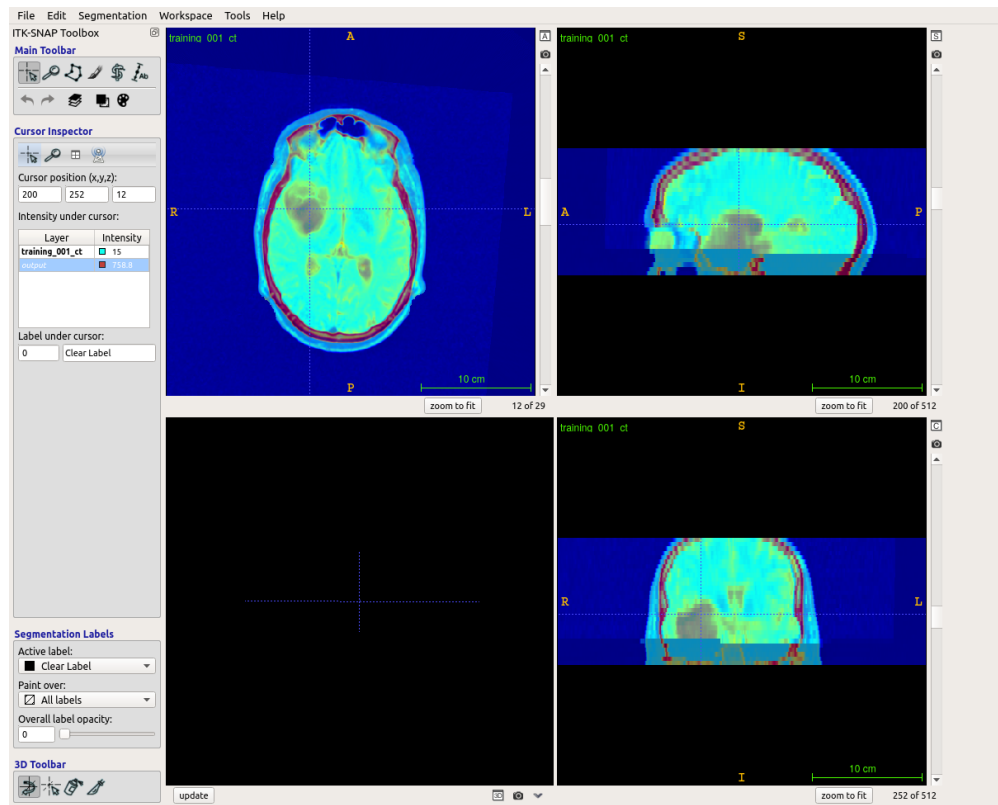


(e) 35 cm

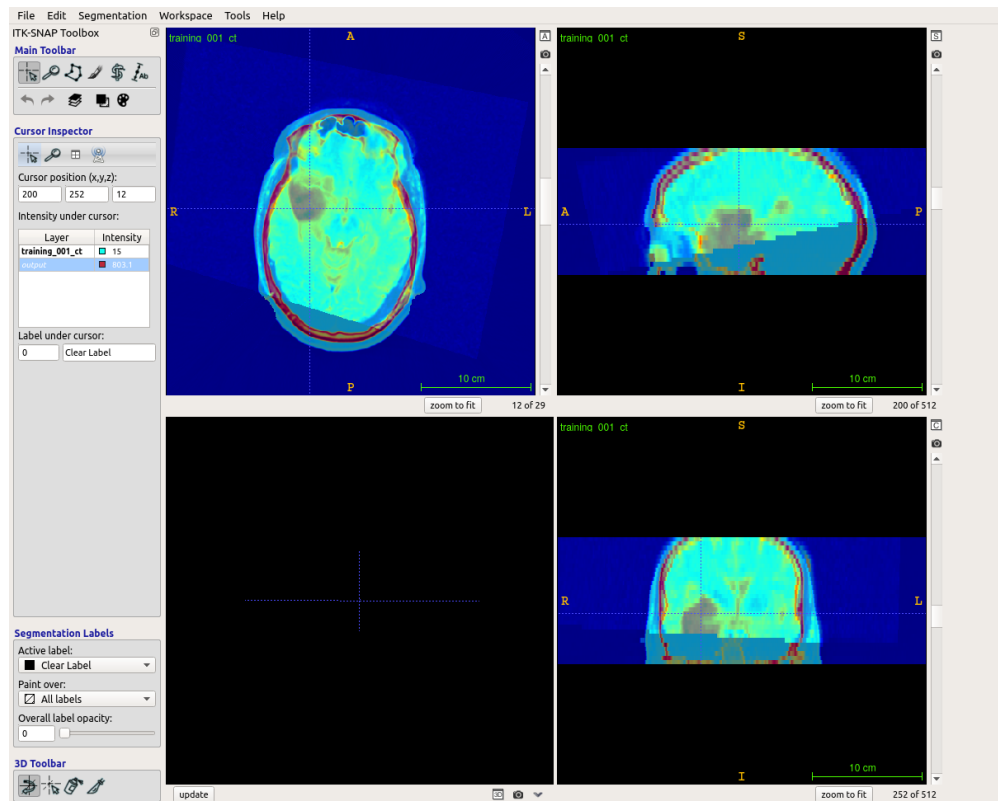


(f) 65 cm

Abbildung 9: Qualitätsunterschiede bei verschiedenen Registrierungsstrategien

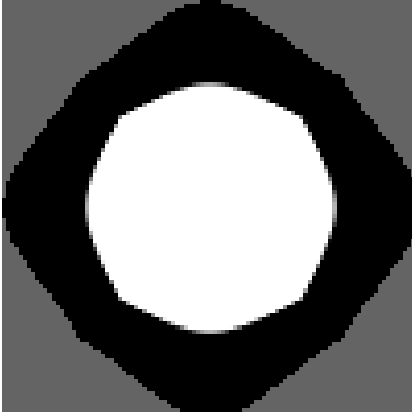


(a) Überlagerung mit optimierten Parametern

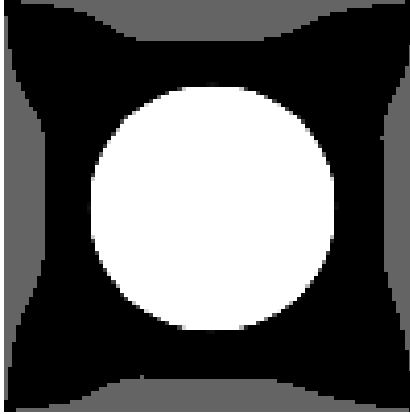


(b) Überlagerung mit weniger optimalen Parametern

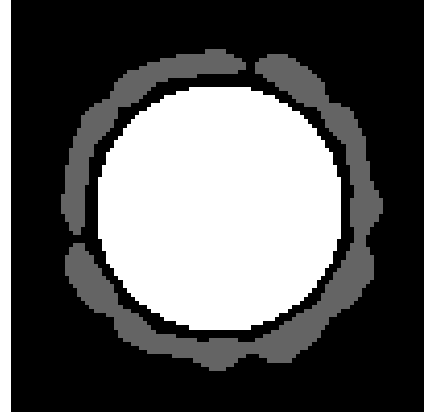
Abbildung 10: Bilder der deformierbaren Registrierung



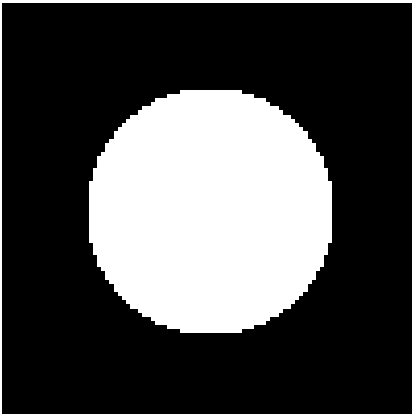
(a) 5 gppd



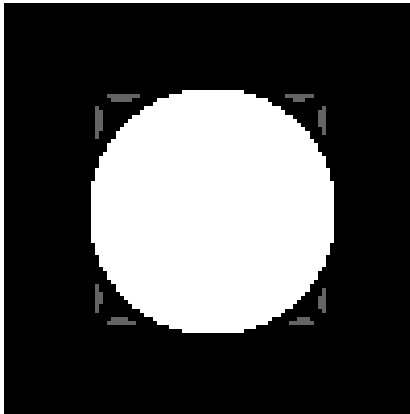
(b) 10 gppd



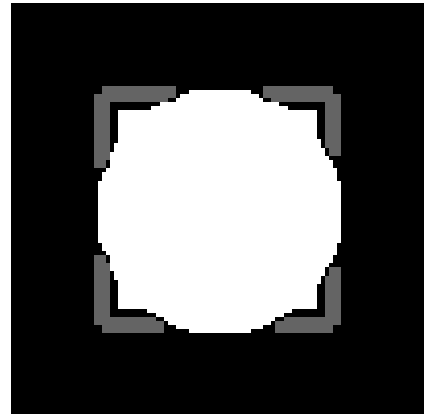
(c) 20 gppd



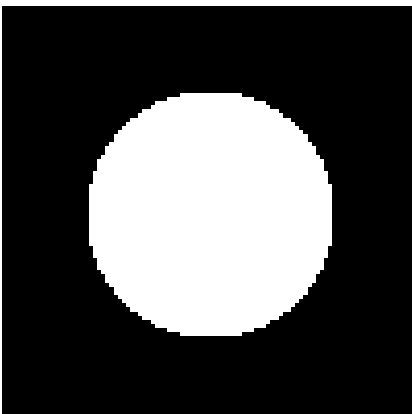
(d) 30 gppd



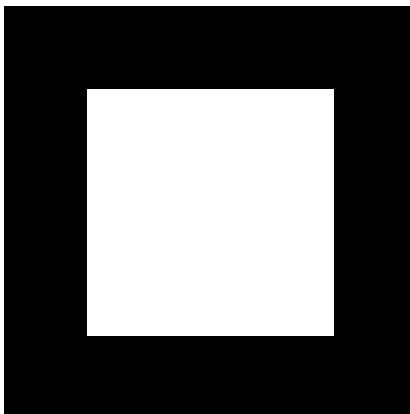
(e) 50 gppd



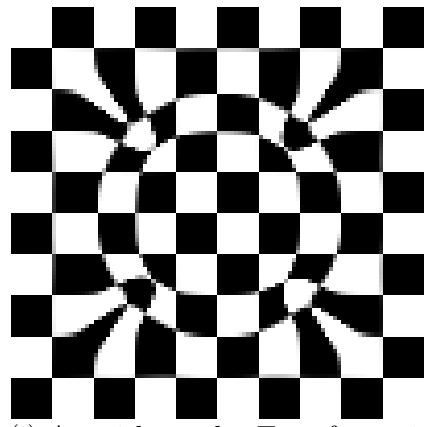
(f) 100 gppd



(g) Zielbild



(h) bewegtes Bild



(i) Auswirkung der Transformation auf ein Schachbrett. 30 gppd



## Versuchsteil 1: Ermittlung der Kamerabewegung aus einer Bildfolge

3.02.2020

### 1.1 Versuchsaufbau

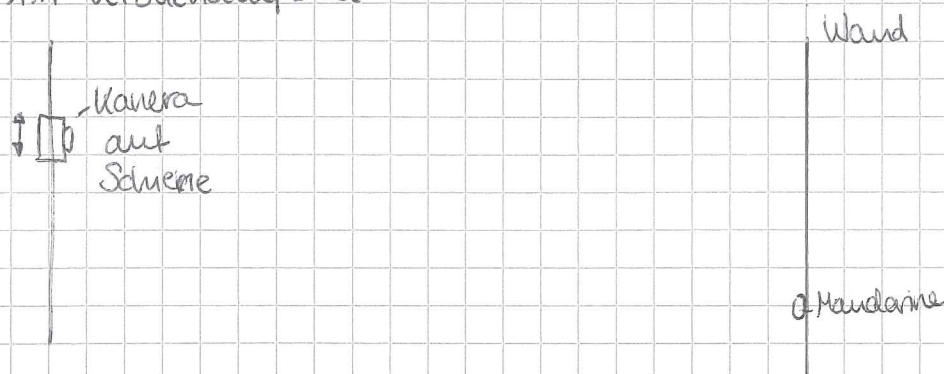


Abb. 1 Versuchsaufbau

Position der Kamera ist relativ weit weg vom messenden Objekt gewählt, damit Rotationen eine geringe Rolle spielen.

### 1.2 Bildaufnahme mit der Kamera

Neben der Ausgangsposition der Kamera links auf der Schiene (0cm) wurden noch weitere Bilder mit den Abständen 5cm, 15cm, 35cm und 65cm von der Ausgangsposition.

### 1.3 Paarweise Registrierung der Aufnahmen

Matlab berechnet mit seiner Transformationsmatrix die Bildverschiebung. Diese Matrix setzt sich folgendermaßen zusammen:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x+a \\ y+b \\ 1 \end{pmatrix}$$

Da wir hier nur 2d arbeiten ist die 3. Zeile auf 1 gesetzt.

Für die verschiedenen Abstände zwischen Referenzbild und bewegtem Bild erhalten wir folgende Abstände:

± 0,5cm ←

Abstand [cm]	5	10	15	20	30	35	65
a [Pixel]	-16,24	-33,77	-50,01	-67,81	0,02	-119,06	0,02
b [Pixel]	-1,6	0,4	-1,1	-2,1	0,02	-3,6	0,02

Tab. 1 Translationswerte

Diese Translationswerte können auch von Hand bestimmt werden, indem das translatierte Bild mit dem ursprünglichen Bild übereinandergelegt werden und die Verschiebung in Pixel abgezeichnet werden. Diese Methode lieferte experimentell ähnliche Ergebnisse.



## Versuchsteil 2: Ähnlichkeit bei monomodalen und multimodalen Bilddaten

03.02.2020

→ siehe C++ code und Python code

Für den multimodalen Teil wählen wir Parameter 0 und 3

Die verschiedenen Bins sind: 5, 10, 10, 50, 100

## Versuchsteil 3: Rigide 3D-Registrierung von medizinischen Bilddaten

03.02.2020

Unsere optimale Einstellung, welche nach Betrachten des CT Bildes unter dem MRT-Bild gefunden wurden lautet:

3 levels (number of levels)

shrink factors per level: 4, 2, 1

smoothing sigmas per level: 2 für jedes

Bildpunkte: 50% Samplingstrategie: Regular

Mit den oben angegebenen Optimalwerten werden nun die Anzahl der Bildpunkte und die Samplingstrategie verändert und das Zeitverhalten notiert.

REGULAR Sampling	Bildpunkte	Zeit [s]		
	100%	47,2	45,5	50,2
	80%	33,8	27,2	26,1
	60%	25,6	28,3	26,6
	40%	19,7	18,5	20,4
	20%	9,7	11,7	14,2
RANDOM Sampling	40%	28,1	29,9	33,1
	80%	55,4	50,7	58,7

Tab. 2 Zeitverhalten rigide 3D-Registrierung

unser Optimum benötigte eine Zeit von 23,5s

Der „Worst-Case“ mit den Einstellungen

1 level, shrink factor = 1, smoothing sigma = 0  
und unter Verwendung aller Bildpunkte

ergab eine Zeit von 114,8s



## Registrierungsgenauigkeit:

Je mehr Level man wählt <sup>umso</sup> desto besser ist die Überlagerung von CT und MRT Bild.

Bei einem Vergleich zwischen hoher und niedriger sampling Rate (10 und 100 %) lässt sich kein großer Unterschied bei sonst „optimalen“ Bedingungen erkennen.

Bei der Wahl eines zu hohen smoothing sigmas (20) ist die Zeit sehr schnell aber dafür leidet die Bildgenauigkeit fast nicht mehr vorhanden. Die Überlagerung ist deutlich schlechter.

## Patientenanalyse:

Patient-101: mit unseren optimalen Einstellungen erhalten wir bei der Überlagerung von CT und MRT Bild eine sehr gute Registrierungsgenauigkeit

Patient-102: auch hier erhalten wir mit unseren Einstellungen sehr gute Ergebnisse

Train reg error: • Patient-101: 171,7

• Patient-102: 155,8

## Aufgabenteil 4: Deformierbare Registrierung

grid points per dim	Time [s]		
5	0,12	0,12	0,12
10	0,49	0,49	0,49
20	2,01	2,01	2,01
30	4,04	4,13	4,09
50	6,16	6,18	6,14
100	31,61	31,53	33,22

Tab. 3 Zeitverhalten  
Deformierung  
(Optimizer stop  
condition failure)

Bei 30 ~~war die~~ grid points per dim war die optimale Deformation. Das Registrierergebnis wurde mit den gespeicherten Bildern dokumentiert.  
Es wurden die gleichen „optimalen“ Registrierstufen wie davor verwendet