



Institut für Prozessrechentechnik, Automation und
Robotik

Prof. Dr.-Ing. H. Wörn
Prof. Dr.-Ing. R. Dillmann
Universität Karlsruhe
Fakultät für Informatik

Bewertung und Optimierung von Schnitten in Weichgewebe

Ein risikobasierter Ansatz

Studienarbeit
von

Tim Weyrich

15.10.98 – 15.01.99

Referent : Prof. Dr.-Ing. R. Dillmann
Betreuer : Dipl.-Inform. T. Salb

Ich versichere hiermit, die vorliegende Studienarbeit selbständig und ohne fremde Hilfe angefertigt zu haben. Die verwendeten Hilfsmittel und Quellen sind im Literaturverzeichnis vollständig aufgeführt.

Karlsruhe, den 15. Januar 1999

Tim Weyrich

Inhaltsverzeichnis

Inhaltsverzeichnis	i
Abbildungsverzeichnis	iii
Tabellenverzeichnis	v
1 Einführung	1
2 Stand der Forschung	3
2.1 Risikobewertung	3
2.2 Schnittplanung	4
3 Mathematische und medizinische Grundlagen	5
3.1 Mathematische Grundlagen	5
3.1.1 Voxelbilder	5
3.1.2 Abstandstransformation	5
3.1.3 Optimierungsverfahren	7
3.2 Medizinische Grundlagen	11
3.2.1 Bildgebende Verfahren	11
3.2.2 Segmentierung und Klassifizierung	13
3.2.3 Der Visible Man Datensatz	13
4 Bewertung von Schnitten	15
4.1 Gegenstand der Bewertung	15
4.2 Ziel der Bewertung	15
4.3 Spezifikation der Problemstellung	16
4.4 Lösungsansatz	16
4.4.1 Das Risikopotential	16
4.4.2 Das Gewebefeld	18
4.4.3 Integration	19
4.4.4 Die Bewertungsfunktion	20
5 Schnittoptimierung	21
5.1 Problemstellung	21
5.2 Ansatz	21
5.3 Mögliche Verfahren	22
6 Implementierung	23
6.1 Automatische Schnittoptimierung	23
6.1.1 Risikopotentiale	23
6.1.2 Bewertungsfunktion	24

6.1.3	Optimierungsalgorithmen	24
6.1.4	Visualisierung	24
6.2	Grundlegende Methoden	24
7	Ergebnisse	27
7.1	Der Datensatz	27
7.2	Gewebespezifische Risikofunktionen	27
7.3	Überlagerung der Risikopotentiale	28
7.4	Führungspolygone	30
7.5	Bewertungsfunktion	31
7.6	Optimierungsläufe	31
8	Diskussion	35
8.1	Schnittbewertung	35
8.2	Schnittoptimierung	35
8.2.1	Startkonfigurationen	36
8.2.2	Erweiterungen der Optimierungsaufgabe	37
9	Zusammenfassung und Ausblick	39
9.1	Zusammenfassung	39
9.2	Ausblick	39
9.2.1	Risikomodellierung	40
9.2.2	Bewertungsfunktion	40
9.2.3	Schnittoptimierung	40
9.2.4	weitere Anwendungen	41
A	Die Job-Datei	43
	Literaturverzeichnis	45

Abbildungsverzeichnis

3.1	Elementare Schritte des Simplexabstiegsverfahrens am Beispiel des Tetraeders	9
4.1	Beispiele eines Fasermodells	19
7.1	Transversaler Schnitt durch den linken Oberschenkel des Visible Man Datensatzes	28
7.2	Schnittbilder des Oberschenkels und des korrespondierenden Risikopotentials	29
7.3	Verschiedene Überlagerungsfunktionen zur Bestimmung des Gesamtrisikopotentials	30
7.4	Unerwünschtes Optimierungsergebnis	31
7.5	Verschiedene Stadien des Optimierungsprozeß'	33
7.6	Ein dem Risikopotential überlagerter Ergebnisschnitt	34

Tabellenverzeichnis

3.1	Aufnahmearten und Parameter des Visible Man Datensatzes	13
7.1	Gewebespezifische Parameter der Risikofunktion 7.1	28

Kapitel 1

Einführung

Chirurgische Eingriffe sind immer mit der Verletzung von Weichgewebe verbunden. Wenngleich große Fortschritte im Bereich der minimalinvasiven Chirurgie gemacht worden sind, so sind selbst hier Zerstörungen gesunden Gewebes nicht zu vermeiden. Es ist aber das Ziel, daß jeder notwendige Eingriff für den Patienten minimal verletzend und damit mit möglichst wenig Risiko verbunden ist.

Die vorliegende Arbeit versucht nun, erste Grundlagen zu schaffen, die es ermöglichen, eine Risikoabschätzung des chirurgischen Eingriffs rechnergestützt durchführen zu können. Es wird ein risikobasiertes Modell des zu operierenden Gewebes aufgebaut, über das man dann versucht, zur Abschätzung konkreter Eingriffe zu gelangen. Dabei werden im Rahmen dieser Arbeit nur Schnitte, wie sie von einem Skalpell verursacht werden, untersucht. Andersartige Eingriffe wie Bohren/Fräsen, das Schneiden mit einer Schere oder das Einfügen eines Implantats könnten durchaus mit Hilfe des gleichen Modells ausgewertet werden, sind aber nicht Gegenstand dieser Arbeit.

Vielfältige Anwendungen einer rechnergestützten Risikoabschätzung sind denkbar:

- Präoperativ ließe sich, aufbauend auf den tomographischen Aufnahmen des Patienten, eine optimale, patientenspezifische Operationsplanung realisieren.
- Während einer Operation könnte ein Rechner, der die Vorgänge über eine entsprechende Sensorik mitverfolgt, eine permanente Risikoabschätzung durchführen, um den Arzt zu informieren oder sogar zu warnen, wenn er im Begriff ist, ernsthafte unerwünschte Verletzungen herbeizuführen.
- In Ausbildung und Lehre ließe sich ein System einsetzen, das Schnitte, die ein Auszubildender an einem computerbasierten Gewebemodell durchführt, bewerten und kommentieren kann.
- Unser Konzept der Risikopotentiale kann dem Mediziner außerdem eine Sicht der menschlichen Physiologie geben, wie sie nur selten erfahren werden kann. Ausgestattet mit einem haptischen Eingabegerät kann der Mensch über die Krafrückkopplung einen taktilen Eindruck der Risikoverteilung in einem gegebenen Gewebedatensatz gewinnen und sich so auf intuitive Weise ein Bild der Verhältnisse machen¹.

Über die Modellbildung hinaus wurde im Rahmen der Studienarbeit zu Demonstrations- und Evaluationszwecken eine Anwendung erstellt, die vollautomatisch versucht, die optimale Lage eines Schnitts zu finden, der ein bestimmtes Zielgebiet in einer Gewebeprobe erreichen soll.

Im folgenden sei kurz der Inhalt dieser Arbeit beschrieben: In Kapitel 2 werde ich kurz schildern, inwieweit bereits Arbeiten vorliegen, die mit der gegebenen Thematik verwandt

¹Parallel zu dieser Studienarbeit wurde bereits ein erster Prototyp der Anwendung realisiert.

sind. Kapitel 3 ist den mathematischen und medizintechnischen Grundlagen gewidmet, auf denen diese Studienarbeit aufbaut. Kapitel 4 stellt das vorgeschlagene Konzept der Schnittbewertung vor, und Kapitel 5 entwirft einen automatischen Schnittoptimierer, der auf diesem Konzept aufsetzt. Kapitel 6 gibt einen Überblick über das System, das von mir innerhalb der Studienarbeit implementiert worden ist und in dem die Ansätze aus den beiden vorangegangenen Kapiteln teilweise umgesetzt worden sind. Die Ergebnisse und ersten Erfahrungen, die mit diesem System gesammelt werden konnten, finden sich in Kapitel 7 und werden in Kapitel 8 diskutiert. Schließlich findet sich in Kapitel 9 eine Zusammenfassung der Arbeit. Darüberhinaus wird ein Ausblick gegeben, in welcher Richtung noch weitere Arbeiten nötig sind.

Kapitel 2

Stand der Forschung

2.1 Risikobewertung

Wenn in dieser Arbeit Schnitte bewertet werden, ist ein elementarer Begriff das *Operationsrisiko*. Dies erweckt den Eindruck, daß es sich dabei um eine wohldefinierte Größe handelt, die man ohne weiteres zahlenmäßig fassen kann. In Wirklichkeit liegt aber einer skalaren Erfassung dieses Risikos immer eine gewisse Willkür zugrunde, die aber durch erfolgreiche Evaluation im medizinischen Alltag gerechtfertigt werden kann.

Tatsächlich spielt die Risikoabschätzung in der Chirurgie eine wichtige Rolle und wird im präoperativen Bereich routinemäßig eingesetzt. In [6] heißt es etwa „*Die Abschätzung des Operationsrisikos ist nach der Indikationsstellung zur OP die wichtigste präoperative Maßnahme.*“

In der Praxis werden dabei je nach erwarteter Operationsdauer einfache *Scoresysteme* eingesetzt, um das Risiko einer Operation abzuschätzen. In ein solches System können Patientenmerkmale wie

- Lungen-, Herz-, Leber- und Nierenfunktion,
- Allgemeinzustand des Patienten und
- die bereits erfolgte Vorbehandlung

eingehen. Diese Risikoscores erlauben zumeist durch simple Addition von spezifischen, diese Merkmale beschreibenden „Risikopunkten“ eine grobe Abschätzung des individuellen Risikos ohne die Notwendigkeit einer eigenen aufwendigen Datenanalyse. Nach Abgleich der Punktesumme mit einer systemspezifischen Punkteskala kommt man zu einer Einschätzung, ob die Operation zum Beispiel ohne weiteres, nur unter erhöhtem Monitoringaufwand oder überhaupt nicht durchgeführt werden kann [6][7].

Im postoperativen Bereich hingegen sind Scoresysteme zur Risikoüberwachung noch nicht allgemein verbreitet. Aktuelle Studien befassen sich mit der Aufgabe, für bestimmte Operationen durch anschließende Beobachtung des Patienten Standards für eine Beurteilung des postoperativen Verlaufs zu schaffen [8].

Die vorliegende Arbeit hingegen befaßt sich mit dem interoperativen Verlauf, und der Fragestellung, wie sich verschiedene operative Schnittführungen auf das Operationsrisiko auswirken. Leider haben wir noch kein existierendes interoperatives Risikoschema gefunden, auf das wir in diesem Falle sinnvoll aufbauen könnten. Daher haben wir die Risikobewertung (siehe Abschnitt 4.4.4) direkt an unserem Risikomodell orientiert. Es bedarf noch des Abgleichs mit der Medizin, um von hier aus zu einer Skala mit entsprechenden Indikationen zu kommen.

2.2 Schnittplanung

Zu Planung und Optimierung von Eingriffen in den menschlichen Körper sind vor allem Arbeiten in der Neurologie und Laparoskopie anzutreffen, wo die Eingriffe gezielter, endoskopischer Natur sind. Meist handelt es sich dabei aber um rechnergestützte Systeme, die dem Menschen nur die visuelle, manchmal auch taktile Rückkopplung während eines endoskopischen Eingriffs simulieren; Planung und Optimierung werden wieder dem Menschen überlassen. Eine automatische Risikoabschätzung wird dabei nur eingeschränkt angeboten.

So wird zum Beispiel das in [9] beschriebene Operationsplanungssystem *NeurOPS* für die neurochirurgische Operationszugangsplanung eingesetzt. Dabei kann ein virtuelles Endoskop innerhalb und außerhalb des Schädels positioniert werden, der aus CT- und MRT-Daten (siehe 3.2.1) rekonstruiert wurde. Damit können verschiedene potentielle Operationszugänge inspiziert, auf ihre Risiken hin untersucht, miteinander verglichen und patientenindividuell der Zugangsweg optimiert werden.

Ein System, das eine vollautomatische Zugangsfindung durch den Computer bietet, ist uns aber nicht bekannt.

Leider konnten wir unter den Systemen, die Schnitt- und Zugangsplanung erlauben, nur Arbeiten finden, die sich mit der Einführung eines nadelförmigen Fremdkörpers, wie eben eines Endoskops, befaßt haben. Schnitttrajektorien, bei denen eine Skalpellklinge eine Fläche im Raum durchstreift, wurden dort nicht untersucht. Arbeiten, die Schnittflächen modellieren, befassen sich weniger mit dem Risiko, daß durch den Eingriff entsteht, sondern versuchen, die mechanische Reaktion des Gewebes zu simulieren. Ein solches System findet sich zum Beispiel in [10].

Kapitel 3

Mathematische und medizinische Grundlagen

3.1 Mathematische Grundlagen

3.1.1 Voxelbilder

Die hier vorgestellten Verfahren arbeiten auf dreidimensionalen Datenfeldern. Ein solches Feld enthält in gewohnter Anordnung eine Menge von Grauwertelementen, *Voxel* (von engl. *volume element*, in Analogie zu *Pixel* für *picture element*) genannt.

Die Grauwerte dieser Elemente bewegen sich innerhalb eines wohldefinierten Intervalls. Bei medizinischen Datensätzen sind ganzzahlige Werte zwischen 0 und 255 üblich. Nehmen die Grauwerte nur die Grenzen des Intervalls an, spricht man von *Binärbildern*.

Jedem Voxel ist eindeutig eine Position im Raum zugeordnet. Diese ergibt sich aus der Konvention, daß die Voxelmittelpunkte auf einem regelmäßigen kartesischen Gitter liegen, dessen „Schrittweite“ in allen drei Raumrichtungen bekannt ist. Stammt der Datensatz aus einem bildgebenden Verfahren (siehe Abschnitt 3.2.1), bestimmt dessen Abtastgeometrie diese Weiten. Ferner bezeichnet der Grauwert eines Voxels dann üblicherweise die Intensität an der gegebenen Stelle der Aufnahme.

Inwieweit Voxel eine Ausdehnung besitzen, wird unterschiedlich spezifiziert. In dieser Arbeit habe man sich ein Voxel als infinitesimal kleinen „Meßpunkt“ innerhalb eines Kontinuums vorzustellen¹. Der Anschaulichkeit halber sprechen wir dennoch von *Voxelmittelpunkten*, wenn wir deren Position meinen.

Die Menge aller Voxel in einem Datensatz bildet den *Voxelraum* V .

Vektoren \mathbf{v} kennzeichnen oft eine Position in einem Voxelraum V . Ist der Grauwert der Stelle \mathbf{v} gemeint, schreiben wir $V(\mathbf{v})$. Fällt \mathbf{v} nicht auf einen Voxelmittelpunkt, muß das Ergebnis über eine geeignete Interpolation definiert werden. Üblich sind die Nächster-Nachbar-Methode, trilineare oder trikubische Interpolation ([11]).

3.1.2 Abstandstransformation

Aus der mathematischen Morphologie ist die *Abstandstransformation* bekannt. Ohne auf deren theoretischen Hintergrund einzugehen, sei sie für unsere Anwendung wie folgt definiert:

Die Abstandstransformation bildet eine Punktmenge $\mathcal{M} \subset \mathbb{R}^3$ auf ein Potential $D: \mathbb{R}^3 \rightarrow \mathbb{R}$ ab. Dabei gilt

$$D(\mathbf{x}) = \text{dist}(\mathbf{x}, \mathcal{M}) \tag{3.1}$$

¹Es hilft die Vorstellung einer Funktion $\mathbb{R}^3 \rightarrow \mathbb{R}$, die nur an den Voxelmittelpunkten ausgewertet werden kann.

mit

$$\text{dist}(\mathbf{x}, \mathcal{M}) := \min_{\mathbf{p} \in \mathcal{M}} \text{dist}(\mathbf{x}, \mathbf{p}), \quad (3.2)$$

wobei $\text{dist}(\mathbf{x}, \mathbf{p})$ der Abstand zwischen den Punkten \mathbf{x} und \mathbf{p} ist. Oft nennt man auch D selbst die Abstandstransformation von \mathcal{M} . In unserem Fall werden \mathbf{x} und \mathbf{p} ausschließlich in Voxelmittelpunkten liegen, da \mathcal{M} in Form eines binären Voxelbilds gegeben sein wird und wir auch D wieder in Form eines Voxelbilds tabellieren werden.

Für die Abstandstransformation eines Voxelbilds existiert eine Reihe von Algorithmen, je nachdem welche Norm dem Abstand zugrundeliegt. Häufig finden Normen wie die Summennorm $\|\cdot\|_1$, die euklidische Norm $\|\cdot\|_2$ und die Maximumsnorm $\|\cdot\|_\infty$ Verwendung. Beliebte ist auch eine Näherung der euklidischen Abstandstransformation, die *chamfer transformation* ([12]). Um auch weniger gebräuchliche Abstandsnormen zulassen zu können, verwende ich einen Algorithmus, den ich zusammen mit R. Glantz entwickelt habe (1998).

Da für die Studienarbeit nur die Abstandstransformation bezüglich $\|\cdot\|_2$ verwendet worden ist, stelle ich hier das auf euklidische Abstände spezialisierte Verfahren von Saito und Toriwaki [12] vor, welches das gleiche Laufzeitverhalten wie unser Verfahren hat, aber leichter zu implementieren ist:

Das Verfahren von Saito und Toriwaki

Sei V ein binäres Voxelbild und $v_{b,r,c} := V((c, r, b)^\top)$. Dann gilt

$$v_{b,r,c} \in \{0, 1\}, \quad 0 \leq b \leq n_b, \quad 0 \leq r \leq n_r, \quad 0 \leq c \leq n_c, \quad (3.3)$$

wobei n_b , n_r und n_c jeweils die Anzahl Schichten, Zeilen und Spalten des Voxelbilds sind. Voxel mit $v_{b,r,c} = 0$ werden *Hintergrundvoxel* genannt. Ist $v_{b,r,c} = 1$, so gehört das Voxel zum *Vordergrund*. Ziel der Abstandstransformation ist nun, daß in jedem Hintergrundvoxel anschließend der Abstand zum nächstgelegenen Vordergrundvoxel steht.

Der Algorithmus von Saito und Toriwaki verfährt in drei Prozessschritten. In jedem Schritt werden Abstände entlang einer der drei Hauptachsen berechnet.

Im ersten Schritt wird das Ausgangsbild v wie folgt Voxel für Voxel in ein Zwischenbild f umgesetzt:

$$f_{b,r,c} = \min_{\substack{v_{b,r,i}=0 \\ 0 \leq i \leq n_c}} \{(c-i)^2\}. \quad (3.4)$$

Anschließend steht in jedem $f_{b,r,c}$ der quadrierte Abstand zum nächstgelegenen Vordergrundvoxel in v , das in der gleichen Spalte (b, r) liegt.

Im zweiten Transformationsschritt wird das Zwischenbild f zeilenweise in ein weiteres Voxelbild g umgeformt:

$$g_{b,r,c} = \min_{0 \leq j \leq n_r} \{f_{b,j,c} + (r-j)^2\}. \quad (3.5)$$

Nach diesem Schritt steht in jeder Schicht von g die quadrierte zweidimensionale Abstandstransformation der korrespondierenden Schicht in v . Dies läßt sich leicht nachweisen, indem man Gleichung 3.4 in 3.5 einsetzt:

$$\begin{aligned} g_{b,r,c} &= f_{b,j,c} + \min_{0 \leq j \leq n_r} \{(r-j)^2\} \\ &= \min_{\substack{v_{b,r,i}=0 \\ 0 \leq i \leq n_c}} \{(c-i)^2\} + \min_{0 \leq j \leq n_r} \{(r-j)^2\} \\ &= \min_{\substack{v_{b,j,i}=0 \\ 0 \leq j \leq n_r, \quad 0 \leq i \leq n_c}} \{(r-j)^2 + (c-i)^2\}. \end{aligned} \quad (3.6)$$

Schritt drei erfolgt analog zum zweiten Schritt, und mit

$$h_{b,r,c} = \min_{0 \leq k \leq n_b} \{g_{k,r,c} + (b - k)^2\} \quad (3.7)$$

erhält man die quadrierte dreidimensionale Abstandstransformation. Einsetzen von Gleichung 3.6 in 3.7 ergibt analog zur vorhergehenden Substitution

$$h_{b,r,c} = \min_{\substack{v_{k,j,i}=0 \\ 0 \leq i \leq n_c, 0 \leq j \leq n_r, \\ 0 \leq k \leq n_b}} \{(c - i)^2 + (r - j)^2 + (b - k)^2\}, \quad (3.8)$$

was zu zeigen war.

Letzter Arbeitsschritt ist, aus diesen Abstandsquadraten durch Wurzelziehen euklidische Abstände zu berechnen.

Das Verfahren kann auch zur Transformation anisotroper Voxelbilder eingesetzt werden. Sind α , β und γ die Gitterabstände zwischen den Spalten, Zeilen und Schichten des Voxelbilds, ergeben sich die drei Schritte des Algorithmus' zu

$$f_{b,r,c} = \min_{\substack{v_{b,r,i}=0 \\ 0 \leq i \leq n_c}} \{\alpha^2(c - i)^2\}, \quad (3.9)$$

$$g_{b,r,c} = \min_{0 \leq j \leq n_r} \{f_{b,j,c} + \beta^2(r - j)^2\} \quad \text{und} \quad (3.10)$$

$$h_{b,r,c} = \min_{0 \leq k \leq n_b} \{g_{k,r,c} + \gamma^2(b - k)^2\}. \quad (3.11)$$

Ein Vorteil dieses Algorithmus' ist, daß er weitestgehend *in place* arbeiten kann: Benötigt man das ursprüngliche Binärbild nach der Abstandstransformation nicht mehr, können f , g und h nacheinander im gleichen Speicherbereich wie v abgelegt werden. Dabei ist zu bemerken, daß nur ganzzahlige Werte als Abstandskvadrat auftreten können.

3.1.3 Optimierungsverfahren

Viele Probleme lassen sich als Minimierungsaufgabe einer Funktion $F: \mathbb{R}^n \rightarrow \mathbb{R}$ formulieren. Oftmals läßt sich das Minimum von F aber nicht analytisch bestimmen. Insbesondere ist F oft nicht einmal in geschlossener Form gegeben.

Eine häufig auftretende Problemstellung ist es, die „Kosten“, die mit einer bestimmten Konfiguration eines Systems verbunden sind, zu minimieren. Dabei versucht man, eine Konfiguration eindeutig durch einen Vektor aus \mathbb{R}^n auszudrücken. \mathbb{R}^n wird dann auch der *Konfigurationsraum* genannt. F angewandt auf solch einen Konfigurationsvektor soll eine Maß für die Kosten des Systems liefern. Man nennt F daher *Kosten-, Bewertungs-, Qualitäts- oder Energiefunktion*².

In der Numerik sind nun einige Verfahren bekannt, die auch bei beliebig gegebenem F das Minimierungsproblem lösen oder zumindest fast-optimale Lösungen finden. Oft ist es nämlich unmöglich in akzeptabler Zeit das *globale* Minimum einer Funktion zu finden, so daß man sich schon mit einem *lokalen* Minimum zufriedengibt, von dem man weiß, daß es dem globalen Minimum schon recht nahe kommt.

Abstiegsverfahren

Stellt man sich F als „Gebirge“ über dem \mathbb{R}^n vor, liegt es nahe, ein Minimum der Funktion zu finden, indem man von einem Punkt des Gebirges ausgehend immer „bergab“ läuft, bis

²In Anlehnung an natürliche Optimierungsprozesse, bei denen stets die Gesamtenergie eines physikalischen Systems minimiert wird.

keine Position mit kleinerem F mehr eingenommen werden kann. Algorithmen, denen diese Idee zugrunde liegt, bezeichnet man als *Abstiegsverfahren*.

Der einfachste Vertreter eines Abstiegsverfahrens ist das *Gradientenabstiegsverfahren*, das ausgehend von einem Startpunkt \mathbf{x}_0 im Konfigurationsraum stets kleine Schritte in entgegengesetzter Richtung des Gradienten macht, also in Richtung kleinerer Werte von F . Iterativ werden also \mathbf{x}_i mit

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \lambda \nabla F(\mathbf{x}_i), \quad i \in \mathbb{N}, \mathbf{x}_i \in \mathbb{R}^n, \lambda \in \mathbb{R}^+ \quad (3.12)$$

gebildet. Bricht man das Verfahren ab, erhält man mit dem letzten \mathbf{x}_i eine Näherung für die Stelle, an der F minimal wird. Die Terminierung des Verfahrens kann unterschiedlich bedingt sein. So kann man den Algorithmus zum Beispiel abbrechen, sobald sich der bisher beste Wert für eine bestimmte Anzahl Iterationen nicht mehr verändert hat³. Verfeinerungen des Gradientenabstiegsverfahrens verkleinern λ mit der Zeit oder setzen λ in Abhängigkeit von $|\nabla F(\mathbf{x}_i)|$, um etwa in sanft abfallenden Gebieten größere Schrittweiten einzusetzen.

Ist F nicht analytisch gegeben, ist der Gradient in Gleichung 3.12 im Normalfall nicht direkt zu bestimmen. In den meisten Fällen bestimmt man daher $\nabla F(\mathbf{x})$, indem man den Gradienten über den Differenzenquotienten hinreichend annähert.

In manchen Situationen konvergiert das Gradientenabstiegsverfahrens sehr langsam. Häufigste Ursache dafür ist, daß das Verfahren oszilliert. Dieser Fall tritt auf, wenn sich das Verfahren anschaulich gesprochen in einem Tal mit langsam abfallender Talsohle am Hang befindet. Hier tendiert es dazu, bei jedem Schritt vom Hang zum Gegenhang zu pendeln und nur langsam eine Verschiebung der \mathbf{x}_i in Richtung des abfallenden Talgrunds zu machen. Dem versucht man entgegenzuwirken, indem man bei jedem Schritt die vorherige Gradientenrichtung miteinbezieht, um dem Gradientenabstieg eine gewisse „Trägheit“ zu geben:

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \lambda \nabla F(\mathbf{x}_i) - \mu \nabla F(\mathbf{x}_{i-1}), \quad i \in \mathbb{N} \setminus 0, \mathbf{x}_i \in \mathbb{R}^n, \lambda, \mu \in \mathbb{R}^+, \quad (3.13)$$

wobei zweckmäßigerweise $\mu < \lambda$ ist.

Trotz solcher Vorkehrungen gibt es Fälle, in denen das Gradientenabstiegsverfahren zu unflexibel ist und durch die mangelnde Anpassung an die Topographie der zu minimierenden Funktion nur langsam konvergiert.

Ein anderes Abstiegsverfahren wurde von Nelder und Mead in [13] vorgestellt: Das *Simplexabstiegsverfahren* (engl. *downhill simplex method*) verfügt über ein besseres Anpassungsvermögen an die Funktionstopographie und kommt darüberhinaus ohne die Bestimmung eines Gradienten aus. Der Algorithmus ist nicht mit dem Simplexverfahren der linearen Optimierung zu verwechseln!

Statt in einem einzelnen Punkt im \mathbb{R}^n beginnt das Verfahren mit $n+1$ Punkten $\mathbf{p}_1, \dots, \mathbf{p}_{n+1}$ die ein n -dimensionales Simplex aufspannen⁴.

Das Verfahren verlangt, daß die \mathbf{p}_i sich in allgemeiner Lage befinden, d.h., daß das Simplex ein Volumen größer Null hat.

Unter dieser Voraussetzung verläuft eine Iteration des Simplexabstiegsverfahrens wie folgt: Die Funktion F wird an allen Ecken \mathbf{p}_i des Simplex' ausgewertet. So werden die Punkte \mathbf{p}_{\min} und \mathbf{p}_{\max} mit minimalem und maximalem F bestimmt. Da es in Richtung von \mathbf{p}_{\max} offensichtlich „bergauf“ geht, versucht das Verfahren zuerst, das Simplex in die entgegengesetzte Richtung zu verlagern. Das erfolgt durch Schritt a) in Abbildung 3.1: \mathbf{p}_{\max} wird an der entgegengesetzten Simplexfläche gespiegelt. Wenn der neue Punkt in F ein besseres Ergebnis

³Tatsächlich kann das Gradientenabstiegsverfahren kurzzeitig auch einmal bergauf laufen, wenn ein Abwärtsschritt über ein Minimum hinaus geht.

⁴Ein *Simplex* der Dimension n ist das einfachste nicht-leere Polyeder im n -dimensionalen Raum. Seine Seitenflächen werden durch $n+1$ $n-1$ -dimensionale Simplizes gebildet. Die Simplizes der Dimensionen 0, 1, 2 und 3 sind *Punkt*, *Kante*, *Dreieck* und *Tetraeder*.

liefert als \mathbf{p}_{\min} , heißt das, daß es in der gewählten Richtung bergab geht. Daher verwirft das Verfahren den eben generierten Punkt zugunsten einer Position, die noch weiter in der eingeschlagenen Richtung liegt (Schritt b)).

Hat der Punkt nach Schritt a) aber zu einem schlechteren Ergebnis als \mathbf{p}_{\min} geführt, hängt die weitere Vorgehensweise davon ab, wie schlecht er ausgefallen ist. Besitzt F hier noch einen günstigeren Wert als im zweitschlechtesten Punkt des Simplex', wird der Punkt aus Schritt a) beibehalten. Andernfalls muß man schließen, daß das Simplex zwischen zwei „Hängen“ liegt. Darum versucht der Algorithmus dann, das Simplex sich in das dazwischenliegende Tal einschniegen zu lassen. Das geschieht über Schritt c), bei dem \mathbf{p}_{\max} auf die gegenüberliegende Simplexfläche zubewegt wird. Nur wenn aber auch hier der neu gewonnene Punkt ein schlechteres Ergebnis als der zweitschlechteste Simplexeckpunkt liefert, wird auch dieser Punkt verworfen und auf Schritt d) ausgewichen, bei dem sich das Simplex in Richtung \mathbf{p}_{\min} zusammenzieht.

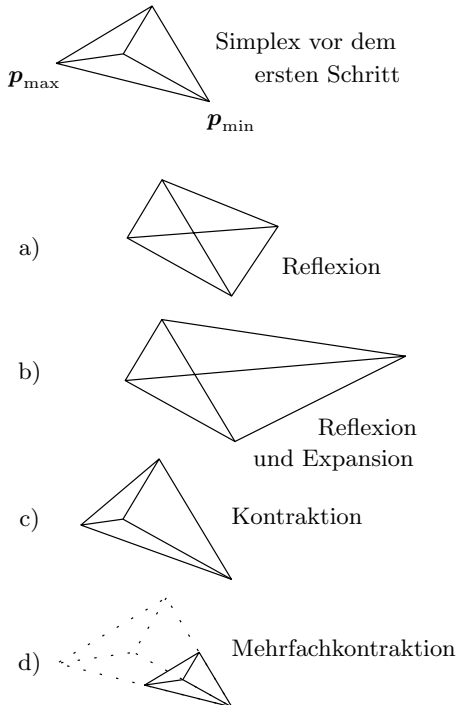


Abbildung 3.1: Elementare Schritte des Simplexabstiegsverfahrens am Beispiel des Tetraeders

Diese Heuristik bewirkt, daß die Simplexeckpunkte mit der Zeit immer günstigere Werte annehmen. Die Simplexgeometrie paß sich dabei automatisch der Topographie von F an. Dadurch führen die Schritte aus Abbildung 3.1 zu Bewegungen, die in den Richtungen die größten Schrittweiten aufweisen, in denen F am wenigsten variiert. Umgekehrt kann sich das Simplex durch die Kontraktionsschritte durch „Nadelöhre“ der Bewertungsfunktion zwingen. Man beachte auch, daß keiner der Schritte das Simplex degenerieren läßt: Das Simplex behält stets ein Volumen größer Null.

Reine Abstiegsverfahren bestehen durch ihre verhältnismäßig schnelle und gesicherte Konvergenz. Allerdings finden diese Verfahren nur das nächstgelegene lokale Minimum zum Startpunkt; wie gut diese Lösung im Vergleich zum globalen Minimum ist, kann nicht festgestellt werden. Der Einsatz dieser Verfahren eignet sich also nur für Funktionen, die möglichst keine lokalen Minima neben dem globalen besitzen.

Manchmal kann man sich noch behelfen, indem man das Verfahren an verschiedenen Startpunkten losschickt und hofft, daß es jeweils in einem anderen Minimum terminieren wird. Doch stößt auch diese Herangehensweise an ihre Grenzen, sobald die Funktion zu viele lokale Minima besitzt.

Stochastische Algorithmen

Es gibt aber auch eine Klasse von Algorithmen, die in gewissen Grenzen aus lokalen Minima auszubrechen vermögen. Das sind die stochastischen Optimierungsalgorithmen. Stellvertretend sollen hier die beiden in der Studienarbeit zum Einsatz gekommenen Verfahren vorgestellt werden: Das *Schwellwertverfahren* und *simuliertes Tempern* (engl. *simulated annealing*).

In beiden Fällen wird wie beim Gradientenabstiegsverfahren ausgehend vom Startpunkt \mathbf{p}_0 eine Folge von Punkten im Konfigurationsraum generiert, die gegen das Minimum der Funktion F konvergieren soll. Diesmal ist die Iterationsvorschrift jedoch nicht deterministisch: In beiden Fällen wird zu einem Punkt \mathbf{p}_i ein Punkt $\hat{\mathbf{p}}_{i+1}$ bestimmt, indem \mathbf{p}_i zufällig verändert

wird. Das könnte zum Beispiel über

$$\hat{\mathbf{p}}_{i+1} = \mathbf{p}_i + \lambda \frac{\mathbf{r}}{|\mathbf{r}|}, \quad \lambda \in \mathbb{R}^+ \quad (3.14)$$

erfolgen, wobei \mathbf{r} ein zufällig gewählter Vektor aus $\mathbb{R}^n \setminus \mathbf{o}$ ist. Ist nun $F(\hat{\mathbf{p}}_{i+1})$ kleiner als $F(\mathbf{p}_i)$, wird $\hat{\mathbf{p}}_{i+1}$ als neuer Folgenpunkt \mathbf{p}_{i+1} übernommen, denn das heißt, daß die zufällige Bewegung im Konfigurationsraum zu einer Verbesserung geführt hat.

Hat eine Verschlechterung stattgefunden, so kann abhängig vom verwendeten Verfahren $\hat{\mathbf{p}}_{i+1}$ doch noch übernommen werden.

Beim Schwellwertverfahren wird der neue Punkt trotzdem weiterverfolgt, wenn

$$F(\hat{\mathbf{p}}_{i+1}) - F(\mathbf{p}_i) < \delta, \quad \delta > 0, \quad (3.15)$$

die Verschlechterung also unterhalb einer festen Schranke δ liegt.

Beim simulierten Tempern hingegen hängt die Annahme des $\hat{\mathbf{p}}_{i+1}$ von einer *Temperatur* $T \in [0, 1]$ ab: Der neue Punkt wird mit der Wahrscheinlichkeit T akzeptiert. Im Laufe des Verfahrens senkt man T langsam gegen Null ab, wodurch mit der Zeit fast nur noch Verbesserungen der Konfiguration zugelassen werden.

Beide Verfahren verfolgen die Strategie, durch gezielte Inkaufnahme von Verschlechterungen während der Minimumsuche aus lokalen Minima und Tälern auszubrechen, in der Hoffnung auf diesem Wege in tiefergelegene Minima zu gelangen. Daß dies durch Einführung einer Zufallskomponente erfolgt, ist eine Anleihe bei der Natur, bei der die Energieminimierung eines Systems meist von zufälligen Störungen unterstützt wird. Am deutlichsten wird das am Kristallisationsprozeß, der auch dem simulierten Tempern Pate gestanden hat: Möchte man einen möglichst regelmäßigen Kristall, also eine energetisch möglichst günstige Form, aus einer Schmelze erhalten, muß man den Abkühlungsprozeß sehr langsam stattfinden lassen. Durch die Brownsche Molekularbewegung werden dabei immer wieder Moleküle aus dem sich bildenden Kristallgitter herausgerissen und haben damit die Möglichkeit, vielleicht noch günstigere Positionen einzunehmen. Kühlt man die Schmelze hingegen zu schnell ab, bildet sich ein weitestgehend amorpher Feststoff mit einer recht hohen freien Enthalpie.

Stochastische Algorithmen eröffnen die Möglichkeit, fast-optimale Lösungen von Problemen mit vielen lokalen Optima zu finden, an denen reine Abstiegsverfahren hängenbleiben würden. Dafür verläuft die Konvergenz recht langsam und ist zudem nicht immer gesichert. Sind Schwellwert oder Verlauf der Temperaturkonstante schlecht bemessen, kann es passieren, daß der Algorithmus sich entweder unkontrolliert im Konfigurationsraum bewegt oder doch in lokalen Minima hängenbleibt. Daher muß beim Einsatz stochastischer Algorithmen verstärk Augenmerk auf die Parametrisierung gelegt werden.

Welche der beiden vorgestellten stochastischen Methoden nun besser ist, ist stark anwendungsabhängig. Interessant ist, daß sie einfacher zu implementieren sind als die Abstiegsverfahren.

Evolutionäre Algorithmen

In gewisser Weise zählen evolutionäre Algorithmen auch zu den stochastischen Algorithmen. Da man sie allerdings selten in diesem Kontext vorfindet, sei ihnen hier ein eigener Abschnitt gewidmet.

Evolutionäre Algorithmen sind in Anlehnung an Prozesse der natürlichen Evolution entstanden. Sie betrachten jeweils eine größere Menge Konfigurationen (*Individuen*) gleichzeitig. Diese Menge wird *Population* genannt. Ein Iterationsschritt besteht aus *Selektion* (auch *Auslese*) und *Mutation*. Die Selektion sortiert Konfigurationen mit einer schlechten Bewertung aus; nur die μ Besten überleben. Aus diesen μ Individuen werden durch Mutation λ neue Nachkommen erzeugt.

Diese Nachkommen halten Einzug in eine neue Population. Je nach *Evolutionsstrategie* (ES) gilt das auch für die μ „Eltern“. Man unterscheidet zwischen der $(\mu + \lambda)$ -ES, bei der die Eltern in der Population verbleiben, und der (μ, λ) -ES, bei der die Eltern „vergessen“ werden und die Nachkommen alleine die neue Population bilden.

So führt zum Beispiel die $(1, 2)$ -Evolutionsstrategie zu einem Algorithmus, der aus zwei Anfangskonfiguration die beste auswählt, um daraus durch Mutation wieder zwei neue Konfigurationen zu schaffen, daraus wieder die beste auswählt, und so fort.

Auch hier sind wieder vielfältige Abbruchkriterien denkbar, z.B. wie stark sich die beste Konfiguration in den letzten n Iterationen verändert hat.

Evolutionäre Algorithmen unterscheiden sich von den bereits genannten stochastischen Algorithmen vor allem dadurch, daß sie stets mehrere Konfigurationen gleichzeitig betrachten. Dadurch wird der Suchraum bei der Minimumsfindung besser ausgeleuchtet, dafür ist die Laufzeit meist länger.

Evolutionäre Algorithmen sind nicht mit *genetischen* Algorithmen zu verwechseln. Für diese muß jede Konfiguration über eine Datenstruktur, die sogenannte DNS, beschrieben werden, die es ermöglicht, zwei Konfigurationen zu „kreuzen“. Daraufhin kann die Analogie zu biologischen Evolutionsmechanismen noch weiter getrieben werden und geschlechtliche Fortpflanzung erfolgreicher Konfigurationen implementiert werden. Eine umfangreiche Diskussion evolutionärer und genetischer Algorithmen findet sich in [14].

Das Simplexabstiegsverfahren in Kombination mit simuliertem Tempern

Die genannten Verfahren lassen sich durchaus kombinieren. So findet man in [15] den Vorschlag, das Simplexabstiegsverfahren mit simuliertem Tempern zu verbinden. Da diese Kombination in der vorliegenden Arbeit Verwendung fand, sei sie hier kurz angerissen.

Der Kern des Mischverfahrens ist das Simplexabstiegsverfahren. Zusätzlich läßt man aber wie beim simulierten Tempern eine Temperaturkonstante T laufen. Werden nun beim Simplexabstiegsverfahren die Eckpunkte des Simplex evaluiert, so wird anschließend noch das Produkt aus T und einer logarithmisch verteilten Zufallszahl *addiert*. Wenn aber nach einem der Schritte aus Abbildung 3.1 der neu konstruierte Punkt ausgewertet wird, wird solch eine Zahl zusätzlich *subtrahiert*. Dadurch werden die Bewertungsvergleiche verfälscht, die der Wahl des nächsten Schritts zugrunde liegen.

Das hat zur Folge, daß das Verfahren Verbesserungen weiterhin akzeptiert, nun aber auch mit einer gewissen Wahrscheinlichkeit in Abhängigkeit von T Verschlechterungen zuläßt. Dies entspricht soweit den Eigenschaften der Standardmethode des simulierten Temperns. Interessant wird diese Kombination aber dadurch, daß sie weiterhin die hohe Anpassungsfähigkeit und gezielte Konvergenz des Simplexabstiegsverfahren aufweist.

3.2 Medizinische Grundlagen

3.2.1 Bildgebende Verfahren

Diese Arbeit setzt voraus, daß das Zielgebiet der Operation in Form eines tomographischen Datensatzes, eines Voxelbilds (siehe 3.1.1), vorliegt. Solche Aufnahmen können unterschiedlichen bildgebenden Verfahren entstammen oder auch aus der Kombination mehrerer Bildquellen entstanden sein.

Heutzutage eingesetzte Verfahren zur dreidimensionalen Bildgewinnung sind unter anderem die *Computertomographie* (CT), die *Kernspintomographie* (MRT) und die *Positron-Emissionscomputertomographie* (PET). Außerdem hat man in der räumlichen Sonographie Fortschritte gemacht, bei der durch Ultraschallaufnahmen gewonnene Schnittbilder zu einer räumlichen Aufnahme kombiniert werden.

CT und MRT, die derzeit wichtigsten bildgebenden Verfahren, seien im folgenden kurz beschrieben; für Informationen zu den anderen Methoden sei auf [16] und [17] verwiesen.

Computertomographie

Bei der Computertomographie wird das Objekt aus verschiedenen Winkeln mit einem Fächer von Röntgenstrahlen durchleuchtet. Das Intensitätsprofil, das sich hinter der Probe ergibt, wird durch eine Reihe von Detektoren aufgenommen und digitalisiert.

Aus den Intensitätsprofilen, die sich für jeden Bestrahlungswinkel ergeben, können die Dichteverhältnisse innerhalb des Objekts durch gefilterte Rückprojektion der lokalen Schwächungskoeffizienten rekonstruiert werden. Dies erfolgt entlang der zweidimensionalen Schicht, in der sich der Strahlenfächer bewegt hat. Verschiebt man nun die Anordnung senkrecht zu dieser Ebene, können weitere Schichtbilder des Objekts aufgenommen werden, die „gestapelt“ eine dreidimensionale Aufnahme der Probe ergeben.

Die CT eignet sich besonders gut, um Knochen von Weichgewebe zu unterscheiden, hat aber den Nachteil, in der Auflösung durch die maximale Fokussierung des Röntgenstrahls begrenzt zu sein: Mit herkömmlichen Apparaturen sind Schichtdicken bis zu einem halben Millimeter die Regel. Der Einsatz der CT am lebenden Menschen wird wenn möglich auf kleine Bereiche des Körpers beschränkt, um die Strahlenbelastung möglichst gering zu halten.

Eine geringere Strahlenbelastung und höhere Auflösungen lassen sich unter Verwendung von Synchrotronstrahlung erzielen [18]. Der große Platzbedarf eines Synchrotrons und die hohen Kosten beschränken aber den Einsatz bisher auf wenige Forschungsprojekte.

Kernspintomographie

Die Kernspintomographie (auch Magnetresonanztomographie, MRT) beruht auf der magnetischen Kernresonanz (NMR). Im Gegensatz zur Computertomographie wird hierbei keine ionisierende Strahlung verwendet. Die Kernspins im Körper werden durch einen kurzen elektromagnetischen Hochfrequenzimpuls angeregt.

Unter Einfluß eines von außen angelegten Magnetfelds wird das bei der Relaxation der Kernspins wieder emittierte Hochfrequenzsignal gemessen. Durch Überlagerung des homogenen magnetischen Hauptfeldes mit einem Gradientenfeld werden Magnetresonanzmessungen ermöglicht, bei denen aus den von der Feldstärke abhängigen Resonanzsignalen zusätzlich auf deren Entstehungsort geschlossen werden kann. Die Frequenz dieser Signale gibt die Ortsinformation; das Abklingverhalten liefert den Bildkontrast [1][19].

Die Ortsauflösung des Verfahrens unterliegt keiner geometrischen Beschränkung, wie sie bei der CT durch Detektorgröße und die Strahlbreite gegeben ist. Nach dem Abtasttheorem ist die minimale Größe der Pixel also nur durch die Dauer der maximalen Abtastzeit bestimmt [16].

Der Hochfrequenzimpuls kann auf die Resonanzfrequenz bestimmter Elemente abgestimmt werden. Meist verwendet man die Resonanzfrequenz des Wasserstoffkerns, da das Element Wasserstoff in fast jeder organischen Verbindung vorkommt und damit eine ausreichende Meßgrundlage geschaffen wird. Außerdem ist die kernmagnetische Nachweisempfindlichkeit für Wasserstoff am höchsten.

Die Kernspintomographie zeichnet sich dadurch aus, daß sie keine Strahlenbelastung für den Patienten bedeutet und Gewebearten gut unterscheiden kann, die im CT kaum einen Kontrastunterschied aufweisen würden. Die im Prinzip mögliche sehr hohe Auflösung wird in der Praxis selten erreicht, da sie nur über eine sehr lange Meßdauer erreicht werden kann.

Aufnahmeart	Schicht- anzahl	Schicht- abstand [mm]	Kanäle	Pixelanzahl	Pixellänge [mm]
CT (fresh)	512	1	1	512 × 512	0,49–0,94
CT (frozen)	1870	1	1	512 × 512	0,49–0,94
MR	–	4	3	256 × 256	1,0–4,0
Photo	1876	1	3	2048 × 1216	0,33

Tabelle 3.1. Aufnahmearten und Parameter des Visible Man Datensatzes [23]

3.2.2 Segmentierung und Klassifizierung

Oft ist es von Interesse, für eine tomographischen Aufnahme zu jedem Voxel zu wissen, zu welcher Gewebeart es gehört. Dazu segmentiert und klassifiziert man das Voxelbild.

Als *Segmentierung* wird die Unterteilung eines Bildes in homogene Regionen bezeichnet. Die *Klassifizierung* ist die Zuordnung dieser Regionen zu Klassen. Bei einer Gebeweklassifizierung sind dies die Gewebeklassen.

Diese Schritte müssen heute noch weitgehend interaktiv erfolgen. Dabei sind Thresholding- und Region-Growing-Techniken weit verbreitet [20]. Zu leistungsfähigeren Segmentierungsansätzen sei auf [21] und [22] verwiesen.

Thresholding

Thresholding ist ein Verfahren, bei dem alle Voxel einer Klasse zugeordnet werden, deren zugeordneten Merkmale (z.B. Farb- oder Intensitätswerte) in einem vorgegebenen Bereich liegen.

Region-Growing

Beim Region-Growing (auch Keimzellenwachstumsverfahren) werden Bereiche der Aufnahme ausgehend von einer Keimzelle (engl. *seed-point*) mit Voxeln einer Klassifizierung „gefüllt“, solange sie gegebene Kriterien erfüllen. Diese können z.B. über Thresholding definiert sein.

3.2.3 Der Visible Man Datensatz

Die Versuche der vorliegenden Arbeit wurden auf dem *Visible Man Datensatz* durchgeführt, der im Rahmen des Visible Human Projektes der National Library of Medicine (NLM), Bethesda, Maryland (USA) erstellt wurde.

Der Visible Man Datensatz besteht aus Schichtaufnahmen eines menschlichen Leichnams, der mittels Computertomographie, Kernspintomographie und Dünnschicht-Farbphotographie erzeugt wurden (siehe Tabelle 3.1). Für letztere wurde der Leichnam tiefgefroren und in dünne transversale Scheiben geschnitten.

Der Leichnam wurde aufgrund strenger Kriterien ausgewählt: Kandidaten mit Infektions- und Krebskrankheiten wurden ebenso als ungeeignet eingestuft, wie Personen, an denen anatomieverändernde operative Eingriffe vorgenommen worden waren. Die Wahl fiel auf einen 38-jährigen Mannes, welcher zum Tode verurteilt worden war und durch eine Giftinjektion hingerichtet wurde. Der Mann war 180 cm groß, sein Gewicht betrug 92 kg [23].

In dieser Studienarbeit haben wir auf die segmentierte Form dieses Datensatzes zurückgegriffen, in der jedem Voxel Luft oder eine von 28 Gewebeklassen zugeordnet wurde. In der verwendeten Oberschenkelprobe waren dann nur noch die Klassifizierungen *Luft*, *Knochenmark*, *Knochen*, *Fett*, *Muskel* und *Blut* anzutreffen. Die Voxelkantenlänge des segmentierten Datensatzes ist in jeder Raumrichtung ein Millimeter.

Kapitel 4

Bewertung von Schnitten

4.1 Gegenstand der Bewertung

Soll das Operationsrisiko abgeschätzt werden, das durch Schnitte in Weichgewebe entsteht, können viele Faktoren berücksichtigt werden. So fließen in das allgemeine Risiko sowohl das Ausmaß der Verletzung ein, als auch die Gefahr, durch versehentliche Abweichung von der geplanten Schnittführung unerwünschte Verletzungen herbeizuführen — ein Fall, den man bei der Arbeit in Weichgewebe immer berücksichtigen muß. Außerdem sind weitere Kriterien denkbar, etwa die durch den Schnitt geschaffene Zugänglichkeit des Zielgewebes für weiterführende Operationen und andere verfahrenstechnische Faktoren.

Um solche Aspekte zu bewerten, bedarf es teilweise chirurgischen Fachwissens und Ansätzen, wie man dieses Wissen im Rechner repräsentiert. Da hier wenig Vorarbeiten vorliegen, haben wir uns bemüht, Bewertungsgrundlagen zu finden, die eine Risikoabschätzung alleine aufgrund tomographischer Aufnahmen des Patienten erlauben sollen.

Formal gehen dabei folgende zwei Komponenten in die Risikoabschätzung ein:

- Eine Risikokomponente, die invariant gegenüber der Ausrichtung des Schnitts relativ zur Geweberichtung ist; wir nennen sie daher *isotrope Risikokomponente*. Sie beschreibt das Risiko, das der Eingriff durch sein bloßes Vorhandensein an einer bestimmten Stelle verursacht.
- Die andere Komponente gibt Auskunft über das Risiko, das dadurch entsteht, daß der Eingriff an relevanten Stellen gegen die Geweberichtung verläuft. Wir nennen sie die *anisotrope Risikokomponente*.

Diese Begriffe werden im Laufe dieses Kapitels noch präzisiert werden.

Im folgenden sollen vorerst nur einzelne Schnitte in einem gegebenen Gewebestück untersucht werden.

4.2 Ziel der Bewertung

Ziel der Schnittbewertung ist eine Risikoabschätzung, die ein für den Mediziner intuitiv zu verstehende Maß darstellt und nicht zu komplex ist. Es sollte jederzeit möglich sein, dem Chirurgen zu erklären, *warum* ein spezieller Schnitt als ungewöhnlich risikoreich eingeschätzt wird. Nur so ist es möglich, auf lange Sicht ein System zu schaffen, das von Ärzten auch angenommen wird.

Andererseits sollte die Bewertung so beschaffen sein, daß ein Rechner in Planungen und Simulationen daraus selbständig weitere Informationen über günstige Schnittführungen ziehen kann.

Beim Entwurf des Ansatzes wurde immer versucht, diese Punkte zu berücksichtigen.

4.3 Spezifikation der Problemstellung

Um einen Schnitt in Weichgewebe zu modellieren, benötigt man sowohl ein Modell des betroffenen Gewebes, als auch eine Spezifikation des Schnitts selbst.

Die vorliegende Arbeit setzt voraus, daß das Gewebe durch eine bereits vollständig nach Gewebeklassen segmentierte Voxelaufnahme beschrieben ist, so daß sich ohne weiteres Binärbilder generieren lassen, die die Verteilung eines einzelnen Gewebetyps wiedergeben (siehe auch 3.2.2).

Für die Erstellung der unten vorgestellten Gewebefelder müssen zudem noch anatomische Atlanten zu Rate gezogen werden. Dieser Schritt konnte allerdings noch nicht ausgearbeitet werden, da dies den Rahmen der Studienarbeit gesprengt hätte.

Der Schnitt selbst wird als Fläche im Raum spezifiziert. Der Begriff Fläche ist dabei durchaus im mathematischen Sinne zu verstehen. Daß der Eingriff damit eine infinitesimal dünne Schicht erfaßt, widerspricht nicht der Tatsache, daß das Schneidewerkzeug eine räumliche Dicke hat, schließlich erfolgt die entstehende Verletzung — bei hinreichend scharfem Werkzeug — nur entlang einer flächenförmigen Trajektorie der Klinge. Zwar wird das betroffene Weichgewebe durch den Eingriff verschoben, doch sehen wir darin keinen Beitrag zur Verletzung des Gewebes. Deformationen, die so stark sind, daß sie Verletzungen in entfernteren Geweberegionen verursachen, sind nicht Gegenstand der Arbeit.

4.4 Lösungsansatz

Ausgehend von den Gewebedaten und der Schnittspezifikation möchte man nun zu einer Bewertung des Schnitts kommen. Dazu muß zuerst ein Modell aufgebaut werden, an dem Operationsrisiken abgelesen werden können.

Entsprechend den zwei bereits genannten Risikokomponenten, die in unsere Bewertung einfließen — die isotrope Risikoverteilung, die die allgemeine Relevanz bestimmter Gewebetypen und die Gefahr, durch versehentliche Abweichung von der geplanten Schnittführung unerwünschte Verletzungen herbeizuführen, beschreibt, und das Ausmaß der notwendigen Verletzung, das sich aus der vorliegenden Geweberichtung ergibt — haben wir zwei Modelle eingeführt: Das Risikopotential und das Gewebefeld. Beide führen durch geschickte Integration entlang der Schnittfläche zu den Komponenten einer Bewertungsfunktion. Im folgenden wird näher auf diese Modelle eingegangen und der Weg zur eigentlichen Bewertungsfunktion beschrieben werden.

4.4.1 Das Risikopotential

Um die isotrope Risikokomponente des Eingriffs bestimmen zu können, führen wir den Begriff des Risikopotentials ein. Die Definition geht davon aus, daß das zu untersuchende Gebiet bereits nach Gewebetypen klassifiziert worden ist:

Das Risikopotential eines bestimmten Gewebetyps ordnet jedem Punkt im Raum das Operationsrisiko zu, das an diesem Gewebe entsteht, wenn ein Eingriff an dieser Stelle stattfindet.

Wie schon in 2.1 aufgezeigt, ist der Begriff des Operationsrisikos in unserem Falle schwer zu fassen. Deshalb haben wir den Versuch gemacht, das Risikopotential intuitiv aufzubauen, und hoffen darauf, die Plausibilität unserer Annahmen in Zusammenarbeit mit Medizinern erhärten zu können.

In diesem Sinne leiten wir aus obenstehender Definition folgende Forderungen an das Potential ab:

- Im Inneren des Gewebetyps muß das Potential das Risiko beschreiben, das entsteht, wenn dieses Gewebe verletzt wird.
- Das Risikopotential soll mit steigendem Abstand zum Bezugsgewebe abnehmen.
- Die Abnahme soll, da wir das ausrichtungsunabhängige Risiko beschreiben wollen, in allen Raumrichtungen gleichmäßig erfolgen.

Sucht man nach Methoden, ein solches Potential aufzubauen, bieten sich beispielsweise folgende Möglichkeiten: Man könnte das Potential ähnlich einem elektrischen Potential berechnen, aufbauend auf der Poissonschen Potentialgleichung [19]. Es wäre aber auch denkbar, das Potential als Funktion des Abstands zur Gewebeoberfläche zu definieren, was auf die Berechnung einer euklidischen Abstandstransformation (siehe Abschnitt 3.1.2) hinauslaufen würde.

Während der Ansatz über die Poissonsche Gleichung dadurch besticht, daß das entstehende Potential physikalische Anschaulichkeit besitzt und beliebig oft differenzierbar ist, sehe ich bei der Verwendung der Abstandstransformation folgende wesentliche Vorteile:

- Topographische Merkmale der Gewebeoberfläche werden weiter in den Raum fortgepflanzt, was sich insbesondere für eine Optimierungsanwendung als Informationsgewinn darstellen kann. Felder nach der Poissonschen Gleichung „verschmieren“ solche Details mit wachsender Entfernung zur Gewebeoberfläche.
- Die Berechnung der Abstandstransformation ist nach heutigem Stand wesentlich schneller: Während man die Lösung der Poissonschen Potentialgleichung über ein dreidimensionales Faltungsintegral berechnen würde [15], was auch unter Verwendung der schnellen Fourier-Transformation noch einen Zeitaufwand von $O(n^3 \log^3 n)$ bedeutet, läßt sich eine euklidische Abstandstransformation in $O(n^2)$ berechnen, wenn n jeweils die mittlere Kantenlänge des Voxelbilds in Voxeln ist.

Aus diesem Grunde schlage ich für die Herleitung des Risikopotentials folgenden Weg vor:

Das Potential eines bestimmten Gewebetyps ergibt sich durch Anwendung einer gewebespezifischen Risikofunktion ρ auf die Abstandstransformation des Gewebes \mathcal{G} (vgl. Gl. 3.1):

$$r_{\mathcal{G}}(\mathbf{x}) = \rho(\text{dist}(\mathbf{x}, \mathcal{G})). \quad (4.1)$$

ρ gibt damit an, wie sich das Risiko mit dem Abstand zur Gewebeoberfläche verändert. Im allgemeinen wird man daher für ρ eine monoton fallende Funktion wählen. Da im übrigen $\text{dist}(\mathbf{x}, \mathcal{G}) = 0$ für alle $\mathbf{x} \in \mathcal{G}$ ist, werden alle Punkte innerhalb des Gewebes mit dem konstanten Risiko $\rho(0)$ beaufschlagt.

Daher muß ρ zwei Bedingungen erfüllen:

- $\rho(0)$ muß ein Maß für das Risiko ausdrücken, wenn ein Punkt dieses Gewebes verletzt wird.
- $\rho(d)$ sollte ungefähr wiedergeben, wie groß das Operationsrisiko ist, wenn sich ein Operationswerkzeug der Gewebeoberfläche bis auf eine Entfernung von $d > 0$ genähert hat.

Für unsere Experimente wählten wir für ρ abfallende e-Funktionen, die je nach Gewebetyp verschieden parametrisiert wurden (siehe Anhang 7.2). Andere Risikofunktionen sind aber denkbar.

Das Gesamtisikopotential r ergibt sich nun als Überlagerung der einzelnen Risikopotentiale. Unter der Annahme, daß die Überlagerung punktweise erfolgt, läßt sich dies als

$$r(\mathbf{x}) = \Omega(r_{\mathcal{G}_1}(\mathbf{x}), \dots, r_{\mathcal{G}_n}(\mathbf{x})) \quad (4.2)$$

schreiben. Als Überlagerungsfunktion Ω sind einige Möglichkeiten denkbar. Mein Vorschlag ist, zu berücksichtigen, daß die Art der Überlagerung mit dem Verfahren, das zur Bestimmung des Einzelpotentials r_{G_i} verwendet wurde, harmonisieren soll. Legt man wie in unserem Falle die Abstandstransformation zugrunde, muß man beachten, daß sich bei deren Bildung, die Einflüsse zweier Punkte des selben Gewebetyps durch Minimumsbildung überlagern. Für die Bestimmung von r liegt es nun nahe, die gleiche Überlagerung zu verwenden. Da ρ in Gl. 4.1 aber monoton fallend ist, entspricht ein Minimum der Abstandstransformation einem Maximum von r_{G_i} . Es empfiehlt sich daher, $\Omega = \max$ zu setzen¹.

Diese Wahl hat den Vorteil, daß sie widersprüchliche Situationen umgeht. Schließlich wäre es unerwünscht, wenn das Gesamtrisikopotential zwischen zwei Gewebebereichen, die über unterschiedliche Klassifizierung, aber gleiche Risikofunktion verfügen, anders ausfiele, als wenn die Bereiche zum gleichen Gewebetyp gehörten.

Dennoch hat sie auch Nachteile: So dominiert an jeder Stelle jeweils nur der Einfluß der hier am stärksten gefährdeten Region. Alle weniger gefährdeten Gewebetypen gehen an diesem Punkt nicht mehr in die Schnittbewertung ein. Welche Wahl nun in welchen Anwendungen am zweckdienlichsten ist, muß noch herausgefunden werden.

Auch diese Aspekte zeigen, daß es diverse unterschiedliche Ansätze zur Festlegung des Risikopotentials gibt. Es ist hier die Herausforderung zu sehen, in Zusammenarbeit mit der Medizin brauchbare Heuristiken zu dessen Bildung zu finden. Aber auch Kriterien wie die Geschwindigkeit seiner Berechnung werden eine Rolle spielen. In keinem Fall darf aber die Schnittbewertung allein aufgrund des Risikopotentials erfolgen.

4.4.2 Das Gewebefeld

Das Risikopotential erlaubt es, eine grobe Abschätzung des Operationsrisikos zu machen. Einfache Prioritäten und Gefährdungen der einzelnen Gewebetypen können hier modelliert werden.

Benötigt man allerdings eine höher aufgelöste Abschätzung des Ausmaßes einer eingriffsbedingten Verletzung, muß zusätzlich ein anatomisch exakteres Modell zu Rate gezogen werden.

Es reicht in keinem Fall mehr aus, lediglich ein Voxelbild zugrunde zu legen, das die Gewebeklassifizierung wiedergibt. Insbesondere benötigt man noch Informationen über die konkrete Geweberichtung in jedem Voxel (Abbildung 4.1). So macht es durchaus einen Unterschied, ob man Muskelgewebe in Faserrichtung schneidet, oder ob die Schnittfläche senkrecht zu den einzelnen Muskelfasern verläuft — auch bei gleicher Größe und Form des Eingriffs ist die Verletzung im einen Fall als minimal, im anderen als recht schwerwiegend einzuschätzen.

Wir führen daher die *anisotrope Risikokomponente* ein, die genau solche Effekte beschreiben soll. Zugrundeliegendes Modell ist das *Gewebefeld*. Es handelt sich dabei um ein Vektorfeld $v: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, das die Verhältnisse im Gewebe beschreiben soll. Dabei modelliert an jeder Stelle

- die *Vektorrichtung* die vorliegende Faserausrichtung des Gewebes und
- der *Vektorbetrag* die Relevanz, die diese Geweberichtung für die Risikobestimmung hat. Liegt an einer bestimmten Stelle keine bestimmbar Geweberichtung vor, ist der Betrag Null.

Der Betrag spielt eine wichtige Rolle, da es Gewebetypen gibt, die zwar eine ausgeprägte Geweberichtung haben, welche aber für das Ausmaß einer Verletzung irrelevant ist (das gilt in vielen Fällen zum Beispiel für kollagenes Bindegewebe [3]).

¹Hätte man eine Poissonsche Feldverteilung zugrundegelegt, würde die gleiche Argumentation zu einer Überlagerung durch Addition führen.

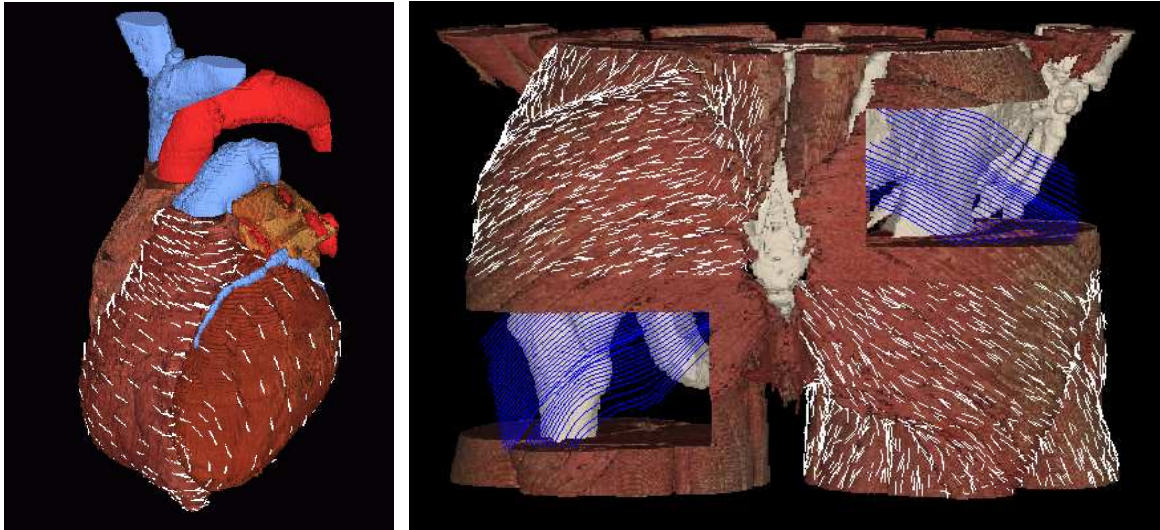


Abbildung 4.1. Beispiele eines Fasermodells: Muskelfaserrekonstruktion am menschlichen Herzen (*Bild links*) und im Beckenbereich (*rechts*). Die Modelle wurden am Institut für Biomedizinische Technik, Universität Karlsruhe erstellt [24].

Das Gewebefeld kann nur unter Zuhilfenahme anatomischer Atlanten gebildet werden. Herkömmliche Aufnahmeverfahren wie CT, MRT oder Ultraschall sind nicht in der Lage, die benötigte Faserinformation zu liefern. Das Hauptproblem ist es dabei, einen gegebenen anatomischen Atlas mit der Physiologie eines konkreten Patienten zu überlagern. Leider konnte dieser Problematik noch nicht nachgegangen werden, da auch diese über den Rahmen der Studienarbeit hinausgegangen wäre.

Wir sind aber der Meinung, daß es sich bei dem Gewebefeld um einen Ansatz handelt, der ein unverzichtbarer Teil einer realistischen Schnittbewertung ist, auch wenn unsere Experimente bis jetzt nur auf dem skalaren Risikopotential aufbauen konnten.

4.4.3 Integration

Um aus den beiden Modellen unter einer konkreten Schnittvorgabe zu einer Einschätzung des Schnitttrisikos zu gelangen, muß man Risikomodellierung und Schnittspezifikation verknüpfen.

Integration über das Risikopotential

Zur Schnittauswertung über dem Risikopotential bietet es sich an, entlang der Schnittfläche über das Risikopotential zu integrieren:

$$R(\mathcal{C}) = \int_{\mathcal{C}} r \, d\sigma \quad (4.3)$$

Dabei entsteht eine Risikofunktion R in Abhängigkeit vom konkreten Schnitt. Durch das Integral summieren sich die Einzelrisiken in allen Punkten, die vom Skalpell durchzogen werden.

R ist damit das oben erwähnte *isotrope Operationsrisiko*, das mit dem Schnitt \mathcal{C} verbunden ist.

Integration über die Geweberichtung

Auch die Bestimmung des *anisotropen Operationsrisikos* V erfolgt über eine Integration. Wenn \mathbf{n} jeweils der Normalenvektor auf der Schnittfläche ist, ergibt sich

$$V(\mathcal{C}) = \int_{\mathcal{C}} (\mathbf{v} \cdot \mathbf{n})^2 \, d\sigma. \quad (4.4)$$

Anschaulich gesprochen summieren sich damit die Sinusquadrate der Winkel, die der Schnitt an jeder Stelle mit der Faserrichtung einschließt, gewichtet mit der Faserrelevanz in diesem Punkt.

Das führt dazu, daß ein hohes V anzeigt, daß der Schnitt an vielen Stellen gegen die Geweberichtung verläuft.

Nach dieser Gleichung würde aber ein Schnitt, der exakt mit der Faser des Gewebes verläuft, keinen Schaden verursachen. Bei manchen Gewebetypen ist diese Annahme sogar annähernd richtig, so zum Beispiel bei einigen Teilen des Muskelapparats. Es gibt aber auch Teile, wie zum Beispiel Blutgefäße oder die Darmwand, deren Verletzung zu einem erhöhten Operationsrisiko führt, unabhängig von der Frage, wie die Geweberichtung vorgelegen hat. In solchen Fällen muß man diesem Umstand wieder mit Anpassung des Risikopotentials dieses Gewebetyps Rechnung zu tragen.

4.4.4 Die Bewertungsfunktion

Mit R und V (Gl. 4.3 und 4.4) liegen nun zwei Risikogrößen vor, die geeignet zu einer Gesamtbewertung des Schnitts zusammengeführt werden müssen.

Wie oben gezeigt, kann eine Schnittbewertung in sehr unterschiedlichen Situationen eingesetzt werden. Das führt aber auch zu unterschiedlichen Anforderungen: Je nach Anwendungsfall mag der Mediziner unterschiedliche Kriterien an die Bewertung eines Schnitts stellen. Insbesondere wird es im gegebenen Fall von der konkreten Pathologie des Patienten abhängen, wie stark ein Risikofaktor bei der Einschätzung des Eingriffs miteinbezogen wird.

Man kann daher nur eine allgemeine Form der Bewertungsfunktion vorschlagen, und da schon R und V als Summen einzelner Risikogrößen entstanden sind, ist es sinnvoll, die Gesamtbewertung E wieder als gewichtete Summe dieser Größen aufzubauen:

$$E(\mathcal{C}) = \mu_1 R(\mathcal{C}) + \mu_2 V(\mathcal{C}), \quad \mu_i \in \mathbb{R}^+. \quad (4.5)$$

über die μ_i lassen sich die R und V je nach Anforderung kombinieren. Kommen später noch weitere Risikokomponenten dazu, können sie auf die gleiche Weise zu E hinzugefügt werden. Genaueres zur Wahl der μ_i wird in 5.2 gegeben.

Kapitel 5

Schnittoptimierung

5.1 Problemstellung

In Kapitel 4 haben wir eine Methode zur Bewertung von Schnitten in Weichgewebe vorgestellt. Nun stellt sich die Frage, inwieweit die Modelle im realen Einsatz zu vernünftigen Risikoabschätzungen führen. Wie in 2.1 dargelegt, mangelt es aber an konkreten Vergleichsmöglichkeiten, um die Qualität unserer Risikoangaben einzuschätzen. Daher war die, den Plausibilitätstest indirekt über folgende Anwendung durchzuführen:

Dem Rechner wird zusätzlich zu den Gewebedaten eine Zielregion vorgegeben, die über einen operativen Eingriff von außen erreicht werden soll. Die Maschine soll nun eigenständig einen optimalen Zugang zu der anvisierten Region finden. Optimalitätskriterium sei dabei, daß das Operationsrisiko minimiert wird.

Um das mit dieser Aufgabe verbundene Optimierungsproblem vorerst so einfach wie möglich zu halten, soll davon ausgegangen werden, daß der Eingriff über einen einzigen Schnitt erfolgen soll, der zudem noch eine fest vorgegebene Form hat.

Sollte es möglich sein, allein mit diesen Voraussetzungen, aufbauend auf der Schnittbewertung aus Kapitel 4 algorithmisch zu Schnittanweisungen zu kommen, die mit unserer intuitiven Vorstellung von einem risikominimalen Zugang zur Zielregion übereinstimmen, so wäre dies ein starkes Indiz, daß der vorgestellte Bewertungsansatz praxisrelevante Aussagen zuläßt.

5.2 Ansatz

Ein Algorithmus, der das gegebene Problem löst, hat nun zwei Aufgaben zu erfüllen:

- Das Operationsrisiko, das mit dem Schnitt verbunden ist, zu minimieren und
- zu gewährleisten, daß die vorgegebene Zielregion erreicht wird.

Vergleichbare Probleme finden sich bei Navigationsaufgaben in der Robotik. Insbesondere wäre es naheliegend, die von dort bekannte Methode der künstlichen Potentialfelder [25] aufzugreifen, bei der Anziehungskräfte zum Ziel und Abstoßungskräfte von Hindernissen dafür sorgen, daß ein Roboter kollisionsfrei zum Ziel geleitet wird. Ausgehend von der Zielregion wären optimale Pfade, das heißt Pfade mit minimalem Risiko, durch das Risikopotential nach außen zu finden. Zum einen aber würde damit das Gewebefeld noch nicht berücksichtigt, was erst möglich ist, wenn die genaue Lage des Schnitts bekannt ist, und zum anderen wäre es schwierig, die Schnittfläche im Nachhinein durch optimale Pfade zu legen, die im allgemeinen auch noch gekrümmt wären. Danach noch sicherzustellen, daß die Bewertung des Schnitts gewisse Optimalitätskriterien erfüllt, wäre denkbar schwierig.

Mein Ansatz geht daher in eine andere Richtung: Grundidee ist es, die gesamte Aufgabe zu einem Optimierungsproblem einer einzigen Funktion zu machen. Dies wird erreicht, indem man das Erreichen der Zielregion ebenfalls zu einer skalaren, zu optimierenden Größe macht. Diese Größe in eine erweiterte Bewertungsfunktion mitaufzunehmen ermöglicht es, einen gebräuchlichen Optimierungsalgorithmus einzusetzen, um alle Kriterien parallel zu optimieren.

Ausgehend von Gleichung 4.5 kommt man dann zur erweiterten Bewertungsfunktion \hat{E} :

$$\hat{E}(C) = \mu_1 R(C) + \mu_2 V(C) + \mu_3 \text{dist}^2(C, \text{Zielregion}) \quad (5.1)$$

$\text{dist}(C, \text{Zielregion})$ sei dabei eine beliebige Gütefunktion, die ausdrückt, wie weit der Schnitt C von der Zielvorgabe abweicht. Betragsmäßig kleinere Werte von dist sollen dabei eine bessere Vorgabenerfüllung ausdrücken.

Das verbleibende Problem ist es nun, diese Bewertungsfunktion zu minimieren.

Dabei empfiehlt es sich, die μ_i mit der Zeit zu variieren. Insbesondere mag es von Vorteil sein, während der Optimierung die Zielvorgabe über ein kleines μ_3 erst schwächer zu berücksichtigen, um dem Algorithmus mehr Freiheitsgrade auf der Suche nach einer optimalen Schnittposition zu lassen. Anschließend legt man ihn allmählich stärker auf die Zielvorgabe fest, indem man μ_3 stark anhebt.

5.3 Mögliche Verfahren

Nun gilt es, ein Verfahren zu finden, das geeignet ist, $\hat{E}(C)$ zu minimieren. Dazu muß man sich erst einmal über die Eigenschaften von \hat{E} klar werden.

Sei C der Konfigurationsraum, der alle möglichen Schnittführungen C enthält, dann ist \hat{E} eine Abbildung $C \rightarrow \mathbb{R}$. Unter den gegebenen Einschränkungen besitzt jeder Schnitt aus C sechs Freiheitsgrade: Drei Translationen und drei Rotationen im Raum. Ein einzelner Schnitt läßt sich also über sechs Werte aus ganz \mathbb{R} eindeutig festlegen, womit C isomorph zu \mathbb{R}^6 und $\dim C = 6$ wäre.

Dem Optimierungsverfahren stellt sich das Problem also als Minimumsuche über einer Funktion $\mathbb{R}^6 \rightarrow \mathbb{R}$ dar.

Diese Funktion enthält im allgemeinen eine Vielzahl lokaler Minima, weshalb reine Abstiegsverfahren (siehe 3.1.3) ungeeignet sind. Man sollte daher ein stochastisches Optimierungsverfahren einsetzen.

Aber auch in diesem Fall will die Wahl gut überlegt sein. Da die Weichgewebedatensätze komplizierte und sehr vielfältige Topographien aufweisen werden, je nachdem welche Körperregion sie beschreiben, wird auch die sechsdimensionale Topographie von $\hat{E}(C)$ recht komplex sein. Im Simplexabstiegsverfahren in Kombination mit simuliertem Tempern [15], das in Abschnitt 3.1.3 vorgestellt wurde, findet man ein adaptives Verfahren, das meiner Meinung nach für diesen Fall sehr gut geeignet ist.

Kapitel 6

Implementierung

Zu den Zielen der Studienarbeit gehörte auch die Aufgabe, die erarbeiteten Ansätze teilweise zu implementieren. In diesem Umfeld ist ein System entstanden, das hier im Detail vorzustellen den Rahmen sprengen würde. Ich gehe daher nur auf die wichtigsten Punkte der Implementierung ein und verweise für genauere Informationen auf die Programmdokumentation.

6.1 Automatische Schnittoptimierung

Um ein Testumfeld des Bewertungsansatzes zu schaffen, wurde die in Kapitel 5 vorgeschlagene Anwendung eines automatischen Schnittoptimierers erstellt. Hier werden anfangs die Risikopotentiale des Gewebedatensatzes berechnet und zum Gesamtrisikopotential zusammengefaßt, das dann Datengrundlage der Bewertungsfunktion ist.

Anschließend wird ein Optimierungsalgorithmus gestartet, der versucht einen Schnitt zu finden, der zu einem minimalen Wert der Bewertungsfunktion führt. Dieser Suchvorgang kann wahlweise in Echtzeit visualisiert werden.

Terminiert der Optimierungsalgorithmus oder wird das Programm vorzeitig unterbrochen, wird das bisher beste Ergebnis in eine Datei geschrieben. Zudem wird ein Voxelbild erzeugt, in dem das Risikopotential mit dem Ergebnisschnitt überlagert ist.

6.1.1 Risikopotentiale

Wie angedeutet wurden aus Zeitmangel noch keine praktischen Arbeiten zum Gewebefeld durchgeführt. Daher basiert die Implementierung der Schnittbewertung bis jetzt nur auf dem Risikopotential.

Die Bestimmung des Risikopotentials erfolgt innerhalb der Optimierungsanwendung und läuft wie folgt ab:

Die Gewebedaten, die den Potentialen zugrunde gelegt werden, stammen aus dem bereits segmentierten Visible Human Datensatz (siehe 3.2.3). Damit kann sofort die Verteilung der verschiedenen Gewebetypen entnommen werden. Für alle Gewebeklassen werden abfallende e-Funktionen als Risikofunktionen verwendet. Deren Parametrisierung ist in Kapitel 7 beschrieben. Die Parameter werden in Form einer Job-Datei (siehe Anhang A) an den Optimierer übergeben, der die ebenfalls in dieser Datei spezifizierten Gewebetypen einzeln abstandstransformiert und entsprechend der spezifischen Risikofunktion daraus die Risikopotentiale der Gewebeklassen bildet.

Die Überlagerung der einzelnen Potentiale zum Gesamtrisikopotential kann wahlweise per Maximumsbildung, Addition oder Multiplikation erfolgen.

Anschließend werden noch die Führungspolygone (siehe 7.4) wie in der Job-Datei spezifiziert in das Risikopotential eingetragen.

6.1.2 Bewertungsfunktion

Da die Bewertungsfunktion innerhalb der automatischen Schnittoptimierung *sehr* häufig ausgewertet werden wird — einige zehntausendmal für eine einfache Optimierungsaufgabe —, wurde auf eine möglichst effiziente Realisierung Wert gelegt. Für die ersten Experimente wurde die Spezifikation der Schnittfläche auf ein planares, konvexes Polygon im Raum beschränkt. Dies stellt keine nennenswerte Einschränkung dar, da es mit einem naturgemäß planen Skalpell ohnehin schwierig ist, nicht-planare Schnitte durchzuführen, und konkave Schnitttrajektorien durch mehrere konkave Polygone beschrieben werden können.

Durch diese Einschränkung war es mir möglich, einen sehr schnellen Algorithmus zu entwickeln, um das Oberflächenintegral in Gleichung 4.3 zu berechnen. Da die Schnittfläche im allgemeinen nicht durch Voxelmittelpunkte verläuft, wird zudem zwischen den Voxeln linear interpoliert, um die Zwischenwerte des Risikopotentials zu berechnen.

Sollte aber doch einmal Bedarf an der Spezifikation gekrümmter Schnittflächen bestehen, schlage ich vor, diese durch ein Dreiecksnetz anzunähern, dessen Facetten wieder über die bestehende Integrationsroutine abgearbeitet werden können.

6.1.3 Optimierungsalgorithmen

Der Optimierer ist so modular gehalten, daß dem Benutzer eine ganze Reihe verschiedener Optimierungsalgorithmen angeboten werden kann. In der gegebenen Zeit habe ich allerdings nur das *Schwellwertverfahren* und *simuliertes Tempern* in Kombination mit dem *Simplexabstiegsverfahren* implementiert (siehe 3.1.3).

6.1.4 Visualisierung

Um den Optimierungsprozeß geeignet parametrisieren zu können, war eine Visualisierung des Vorgangs unabdinglich. Dazu wurde ein Programm entwickelt, das über einen Kommunikationskanal mit dem Optimierer verbunden wird und in regelmäßigen Abständen den aktuellen Zustand des Algorithmus' abfragt.

Die zur Zeit untersuchten Schnittkonfigurationen und die bisher beste Schnittführung werden als transparente Schnittflächen in ein animiertes Oberflächenmodell des Gewebedatensatzes eingeblendet.

Da dieses Oberflächenmodell aus einer sehr hohen Anzahl Dreiecksflächen gebildet wird — das in unseren Versuchen verwendete Modell besaß fast 90.000 Facetten —, mußten einige Optimierungen am Darstellungsverfahren vorgenommen werden, um eine Echtzeitvisualisierung zu ermöglichen. Deshalb konnte ich nicht auf bestehende, allgemeinere Visualisierungsprogramme zurückgreifen, sondern mußte einen Oberflächenvisualisierer neu implementieren, um alle Optimierungsmöglichkeiten ausschöpfen zu können. Dabei kam die Grafiksprache OpenGL zum Einsatz, wodurch es möglich geworden ist, den Visualisierer auf den unterschiedlichsten Systemen zu verwenden.

6.2 Grundlegende Methoden

Im Vorfeld der Arbeiten habe ich eine Bibliothek zur Handhabung von Voxelbildern entwickelt, die auch für die Arbeit grundlegende Schritte wie die Abstandstransformation und einfache Bildbearbeitungsfunktionen umfaßt. Bei der Erstellung wurde Wert auf größtmögliche Allgemeinheit gelegt, damit die Arbeit auch für weitere Anwendungen genutzt werden

kann. So können die bearbeiteten Voxelbilder durchwegs beliebig dimensional sein. Neben einem proprietären Dateiformat wird auch das PIC-Format [26] unterstützt, das am DKFZ Heidelberg entwickelt wurde. Im Gegensatz zur derzeitigen Version der offiziellen PIC-Library unterstütze ich schon die Speicherung reellwertiger Voxeldaten, was insbesondere benutzt wird, um vorberechnete Risikopotentiale abzulegen.

Ferner sind noch kleinere Hilfsprogramme entstanden, die vor allem Dateikonvertierungen vornehmen.

Kapitel 7

Ergebnisse

Im vorangegangenen Kapitel wurde kurz das System dargestellt, das der Auswertung der vorgestellten Ansätze zugrunde lag. Im Folgenden werde ich Ergebnisse und Erfahrungen mit dem System vorstellen.

7.1 Der Datensatz

Für die Versuche wollten wir zuerst Schnitte in einer möglichst einfach strukturierten Körperregion untersuchen, um die Topologie der Bewertungsfunktion überschaubar zu halten. Deshalb fiel die Wahl auf den Oberschenkel. Hier sind große, lineare Strukturen vorherrschend, und die Blutgefäße bilden ein natürliches Hindernis bei der Schnittfindung. Abbildung 7.1 zeigt zwei Schnittbilder durch die Oberschenkelprobe, die dem Visible Human Datensatz entnommen ist.

Der Ausschnitt ist ein Voxelbild aus $210 \times 212 \times 186$ Voxeln und enthält damit 8.280.720 Werte.

7.2 Gewebespezifische Risikofunktionen

Als Risikofunktion wählten wir für unsere Experimente jeweils eine abfallende e-Funktion der Form

$$\rho(d) = ae^{\left(\frac{1}{b} \log \frac{1}{2}\right)d}, \quad d, b [\text{mm}]. \quad (7.1)$$

a ist demnach der isometrische Risikofaktor im Inneren des Gewebetyps, und b ist die Entfernung zur Gewebeoberfläche, ab der das Risikopotential auf die Hälfte abgefallen ist, denn es gilt: $\rho(0) = a$ und $\rho(b) = \frac{a}{2}$. a ist dimensionslos, d und b werden in Millimetern gemessen.

Tabelle 7.1 zeigt a und b für unterschiedliche Gewebeklassen, so wie sie sich in Experimenten zur automatischen Schnittoptimierung als plausibel ergeben haben. Es sind nur Gewebetypen aufgelistet, die im verwendeten Oberschenkel Datensatz vorkommen.

In die Wahl der Parameter sind folgende Überlegungen eingeflossen: *Fett* wurde als weitestgehend irrelevant für das Operationsrisiko eingeschätzt, weshalb eine verhältnismäßig kleine, schnell abfallende Risikofunktion verwendet wird. *Knochenmark* und *Muskelgewebe* wurden gleichermaßen mit einem durchschnittlichen Operationsrisiko versehen: Das Risikopotential ist innerhalb der beiden Gewebeklassen konstant Eins¹ und fällt nach 5 mm auf die Hälfte

ab. Als besonders gefährdet werden die *Blutgefäße* eingeordnet. Entsprechend fällt ihr Risikopotential recht hoch aus und hat eine „Halbwertsdistanz“ von 12 Millimeter. Ein Sonderfall

¹Entsprechend hat man sich beim Entwurf der Risikofunktionen an 1,0 als „durchschnittlichem“ Risikofaktor orientiert.

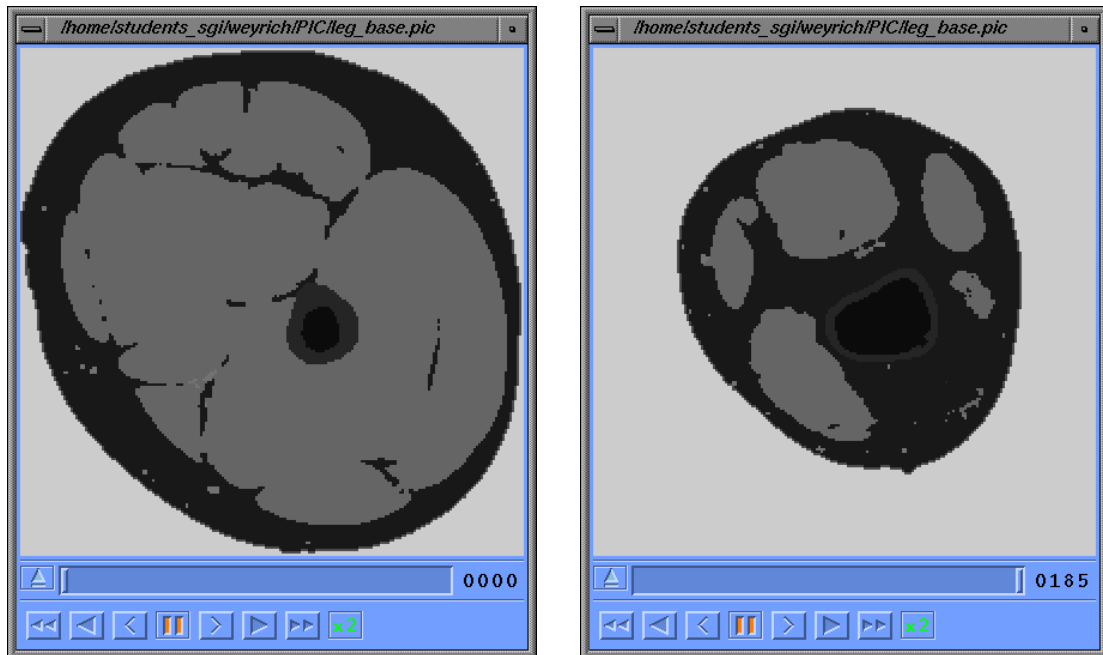


Abbildung 7.1. Transversaler Schnitt durch den linken Oberschenkel des Visible Man Datensatzes: Kurz unterhalb des Hodens (*linkes Bild*) und oberhalb der Kniescheibe (*rechts*). Die verschiedenen Grauwerte bezeichnen die Gewebeklassen. In den Bildern sind Knochen, Knochenmark, Muskeln, Fettgewebe und Blutbahnen zu erkennen.

ist der *Knochen*: Die Reichweite des Risikopotentials ist während der Operation gut lokalisiert werden kann, also auch präzise Schnitte in Knochennähe möglich sind. Dafür wurde der Risikofaktor im Inneren der Knochensubstanz extrem hoch angesetzt, da die automatische Optimierung in keinem Fall Schnitte finden soll, die durch den Knochen verlaufen würden.

Diese Parameter wurden über die Job-Datei (Anhang A) an den Optimierer gereicht. Dieser berechnet das Risikopotential für jede Voxelposition im Ursprungsbild vor, damit die Auswertung der Oberflächenintegrale schnell erfolgen kann. Abbildung 7.2 zeigt Schnittbilder des Oberschenkel Datensatzes mit jeweils einem Schnitt durch das korrespondierende Risikopotential, wie es im Optimierer Verwendung findet.

äußerst gering, da der Knochen

Gewebeklasse	a	b [mm]
Knochenmark	1,0	5,0
Fett	0,1	1,0
Knochen	100,0	0,75
Muskel	1,0	5,0
Blutgefäße	5,0	12,0

Tabelle 7.1: Gewebespezifische Parameter der Risikofunktion 7.1

7.3 Überlagerung der Risikopotentiale

Wie in 4.4.1 schon diskutiert, stehen mehrere Methoden zur Auswahl, um die einzelnen Risikopotentiale zu einem Gesamtpotential zu überlagern. Abbildung 7.3 zeigt das resultierende Risikopotential für drei verschiedene Überlagerungsfunktionen. Wie zu sehen ist, unterscheidet sich das Ergebnis der Addition von dem der Maximumsbildung bis auf Skalierung nur marginal: Die Addition führt zu etwas weicheren Übergängen zwischen den Einflußbereichen der gefährdeten Gewebeklassen, hat aber einen entscheidenden Nachteil: Entfernt man den Einfluß eines Gewebetyps — möchte man zum Beispiel den Einfluß des Fettgewebes vernachlässigen und setzt die Fett-Risikofunktion auf Null —, verschiebt sich der Wertebereich

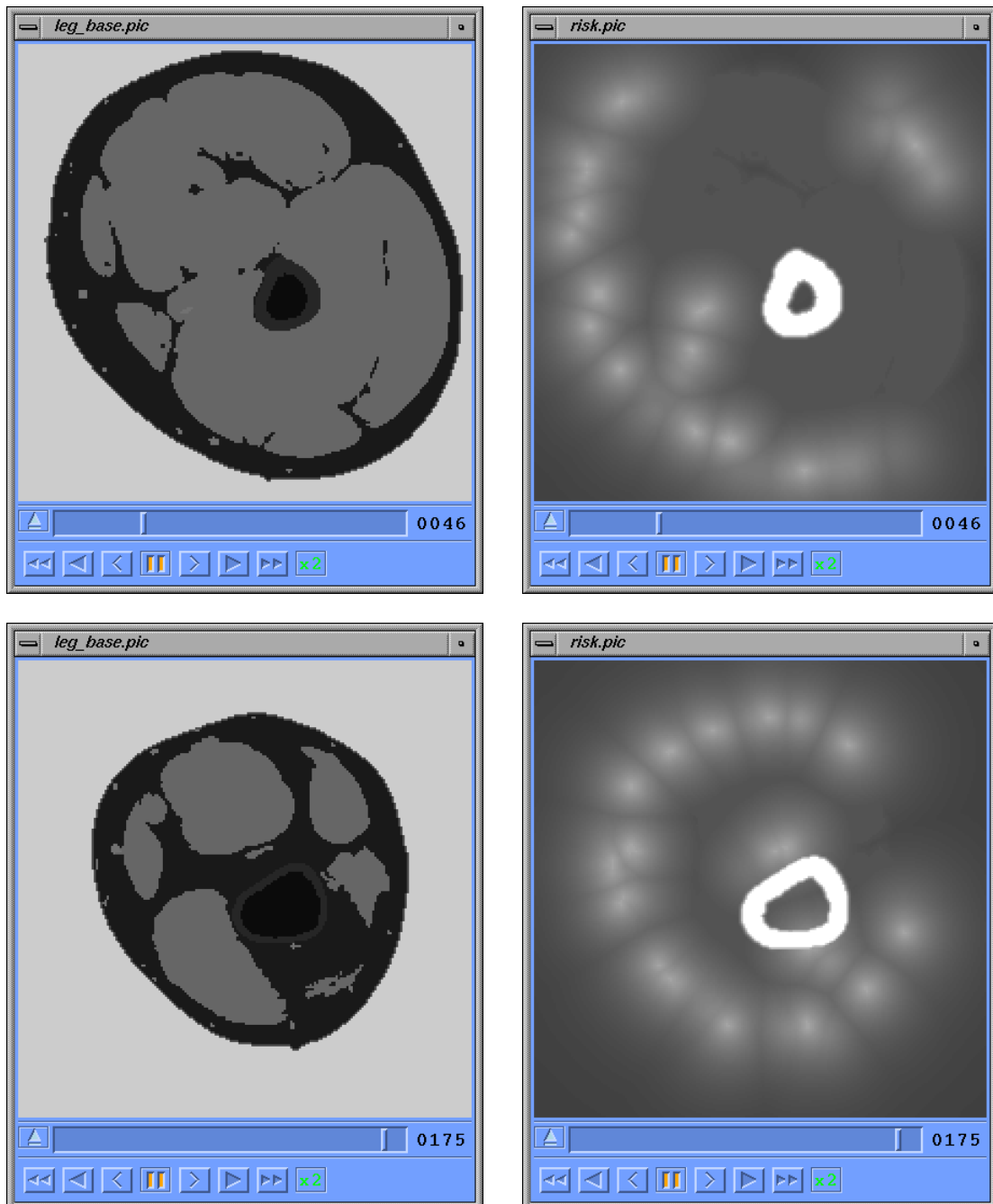


Abbildung 7.2. Schnittbilder des Oberschenkels (*linke Spalte*) und des korrespondierenden Risikopotentials (*rechts*): Helle Punkte im Risikopotential entsprechen einem hohen Risiko. Die hohen Risikopotentials des Knochens und der Blutgefäße sind deutlich zu erkennen.

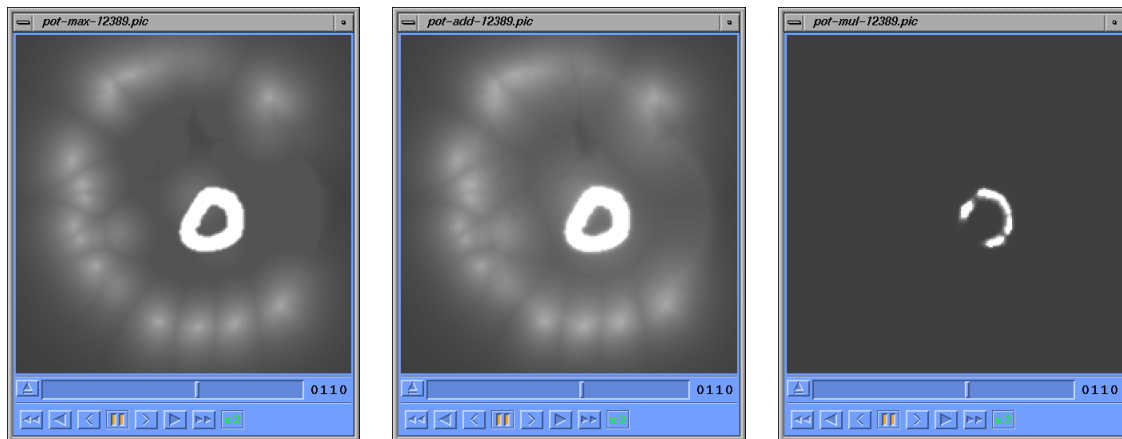


Abbildung 7.3. Verschiedene Überlagerungsfunktionen zur Bestimmung des Gesamtrisikopotentials: Maximumsbildung (*linkes Bild*), Addition (*Mitte*) und Multiplikation (*rechts*). Da das resultierende Potential jeweils sehr unterschiedlich skaliert ist, wurden die Potentiale für den Vergleich linear in den gleichen Wertebereich abgebildet.

des Gesamtpotentials, was eine einheitliche Risikobewertung verhindert².

Die Multiplikation liefert ein eher unbrauchbares Ergebnis, da die verwendeten Risikofunktionen für den Fall der Maximumsbildung ausgelegt sind. In jedem Fall tritt aber auch hier das Skalierungsproblem auf.

Aufgrund dieser Argumente und der Diskussion aus 4.4.1 haben wir uns für die weiteren Schritte für die Maximumsbildung als Überlagerungsfunktion entschieden.

7.4 Führungspolygone

Bei der Verwendung von Ausschnittsdaten aus dem menschlichen Körper treffen zwangsläufig innere Bereiche des Körpers auf die Grenzflächen der Aufnahme. Nun ist es vorgekommen, daß der Optimierer dazu tendiert hat, von diesen „offenen Stellen“ aus Schnitte in das Innere des Datensatzes zu setzen (siehe Abschnitt 7.6). Da ein Chirurg aber nie im Inneren des Gewebes anfangen kann zu schneiden, ohne zuvor schon durch den geschaffenen Zugang Verletzungen erzeugt zu haben, müssen solche Lösungen ausgeschlossen werden.

Dies erreichen wir, indem an den Schnittflächen, an denen die Grenzflächen der Aufnahme Gewebe durchlaufen, das Risikopotential extrem angehoben wird. Jeder Schnitt, der nun durch eine dieser „verbotenen“ Flächen stößt, wird mit einem ungewöhnlich hohen Risiko beaufschlagt, was den Optimierer eine andere Richtung einschlagen läßt.

Da sich dieser Mechanismus auch nutzen läßt, um weitere Randbedingungen an den Zugang zur Zielregion zu stellen, wird die Risikoanhebung manuell vorgenommen: In der Job-Datei (siehe Anhang A) können sogenannte Führungspolygone definiert werden. Entlang dieser Polygone wird das Risikopotential auf einen frei wählbaren Wert gesetzt. Diese Polygone können flexibel als Barrieren für die Schnittführung eingesetzt werden.

Die Verrasterung der Polygone erfolgt *antialiased*, um den Optimierungsvorgang störende Treppenstrukturen zu minimieren.

²In der Darstellung 7.3 tritt das Skalierungsproblem nicht zutage, da die Bilder normiert wurden.

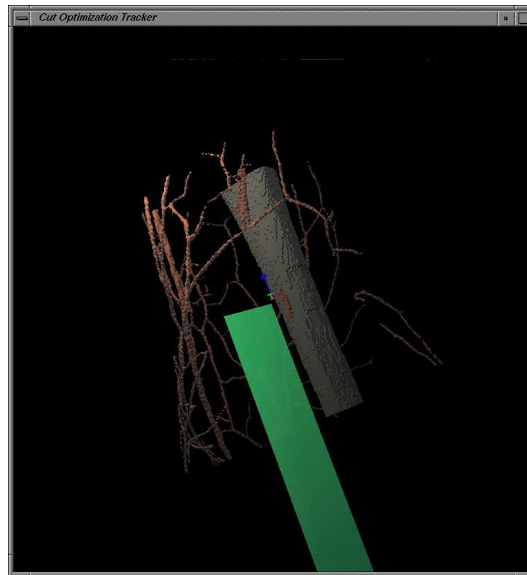


Abbildung 7.4. Unerwünschtes Optimierungsergebnis: Ohne die Einführung der Führungspolygone führt der Optimierer den Schnitt parallel zum Knochen. In der Oberflächendarstellung sind der Röhrenknochen des Oberschenkels, die Blutgefäße und, halbtransparent gezeichnet, der rechteckige Lösungsschnitt zu sehen.

7.5 Bewertungsfunktion

Die Optimierung erfolgte über die Bewertungsfunktion \hat{E} aus Gleichung 5.1. Da noch kein Gewebefeld zur Verfügung stand, ist μ_2 konstant Null. Die weiteren Parameter waren $\mu_1 = 1,0$ und $\mu_3 = 10,0$. Da die Zielregion bis jetzt nur durch einen einzigen Punkt gegeben war, wurde für $\text{dist}(\mathcal{C}, \text{Zielpunkt})$ der Abstand des nächstgelegenen Eckpunkts des Schnittpolygons zum Zielpunkt eingesetzt.

Diese Vorgaben waren mehr oder weniger willkürlich, wurden aber nicht mehr verändert, um vergleichbare Ergebnisse zu bekommen.

7.6 Optimierungsläufe

Als Optimierungsaufgabe wurde die Positionierung einer rechteckigen Trajektorie einer Klinge gefordert, die mit einer Ecke genau die Mitte des Datensatzes erreichen sollte. Die Trajektorie ist ca. 21,2 cm lang und 4,2 cm breit. Diese Aufgabe wurde im Laufe der Versuche nicht mehr verändert, um vergleichbare Ergebnisse zu bekommen.

Die ersten Optimierungsläufe erfolgten unter Verwendung des Schwellwertverfahrens (siehe 3.1.3). Dieses Verfahren betrachtet zu einem Zeitpunkt jeweils nur eine aktuelle Schnittkonfiguration. Diese wird von der Visualisierung in Form einer halbtransparenten Schnitttrajektorie in ein Oberflächenmodell des Oberschenkeldatensatzes eingeblendet und dem gerade untersuchten Punkt des Algorithmus' nachgeführt. Darüber hinaus wird laufend der bisher beste Schnitt in einer anderen Farbe überlagert. Auf diese Weise ist es möglich, den Fortgang des Algorithmus zu verfolgen.

Bei den ersten Versuchen bewegte sich die Schnittfläche noch sehr unkontrolliert durch den Raum. Ob Hindernisse mit hohem Risikofaktor vorhanden waren, hatte keinen Einfluß auf die Bahn, die der Algorithmus verfolgte. Das ließ auf einen zu hohen Schwellwert schließen — das Verfahren ließ fast beliebige Verschlechterungen in der Schnittbewertung zu. Außerdem

waren die Veränderungen, die der Algorithmus von Schritt zu Schritt an der Schnittposition vornahm, recht groß, so daß es fast unmöglich war, eine Feinoptimierung der Schnittposition zu erlangen.

Eine Anpassung von Schwellwert und Schrittweite führte zu etwas besseren Ergebnissen. Dabei durfte man den Schwellwert aber nicht zu weit nach unten korrigieren, da das Verfahren sonst frühzeitig in lokalen Minima terminiert hat. Es war abzusehen, daß dieser mühsame Anpassungsprozeß für andere Gewebeproben und Schnittformen wieder neu begonnen werden müßte. Daher wurden die Experimente mit dem Schwellwertverfahren nicht weitergeführt.

Stattdessen bin ich zum Simplexabstiegsverfahren übergegangen, das sich mit simuliertem Tempern gekoppelt besser an unterschiedliche Umgebungen anpassen kann (siehe 3.1.3, Seite 11).

Schon während der Versuche mit dem Schwellwertverfahren trat aber das Problem zutage, daß das Verfahren Schnitte lieferte, die den Oberschenkeldatensatz über Bereiche verließen, die noch im Inneren des Körpers lagen (Abbildung 7.4). Das waren im Sinne der zu optimierenden Funktion durchaus gute Ergebnisse. Dennoch waren sie unerwünscht, da sie keine medizinische Relevanz hatten. Daher wurden, wie in Abschnitt 7.4 beschrieben, im folgenden Führungspolygone eingeführt.

Nachdem nun zwei Führungspolygone an die Ränder der Aufnahme gesetzt wurden, konnte mit dem neuen Algorithmus fortgefahren werden.

Da das Simplexabstiegsverfahren im n -dimensionalen Konfigurationsraum nicht nur einen Punkt, sondern $n + 1$ Punkte zur Zeit betrachtet, müssen in der Visualisierung auch $n + 1$ Schnittflächen dargestellt werden. In unserem Fall sind deshalb sieben halbtransparente Flächen zu sehen, die mit den Ecken des sechsdimensionalen Simplex' korrespondieren. Je stärker sich diese Flächen voneinander in Rotation und Translation unterscheiden, desto größer ist das Simplex des Optimierungsverfahrens, und desto größer sind auch die Änderungen, die die Schnitte in jedem Optimierungsschritt erfahren. So läßt sich auch das Konvergenzverhalten des Verfahrens gut verfolgen (Abbildung 7.5).

Die Ergebnisse fielen um einiges besser aus, als mit dem Schwellwertverfahren. Am augenfälligsten war die schnelle Konvergenz der Optimierung. Während das Schwellwertverfahren erst nach größenordnungsmäßig zehn bis zwanzig Minuten³ eine verhältnismäßig stabile Lösung gefunden hatte, fand das neue Verfahren nach ein bis zwei Minuten Schnitte, die nicht mehr sichtbar variierten. Bei genauer Betrachtung schmiegt sich diese Schnitte auch sehr schön in Regionen des Risikopotentials mit minimalem Risiko ein (Abbildung 7.6).

Diese Ergebnisse lassen sich in der Qualität gut reproduzieren, wengleich sich die Lösungstrajektorien jedesmal leicht unterscheiden, da es sich um einen stochastischen Algorithmus handelt. Dabei fällt auf, daß die unterschiedlichen Lösungen immer in einer kleinen Gruppe von Lösungsklassen liegen, die sich durch die Einstiegsrichtung unterscheiden.

³Alle Zeitangaben beziehen sich auf Berechnungen auf einem der mit 195 MHz getakteten MIPS R10000 Prozessoren einer SGI Onyx2.

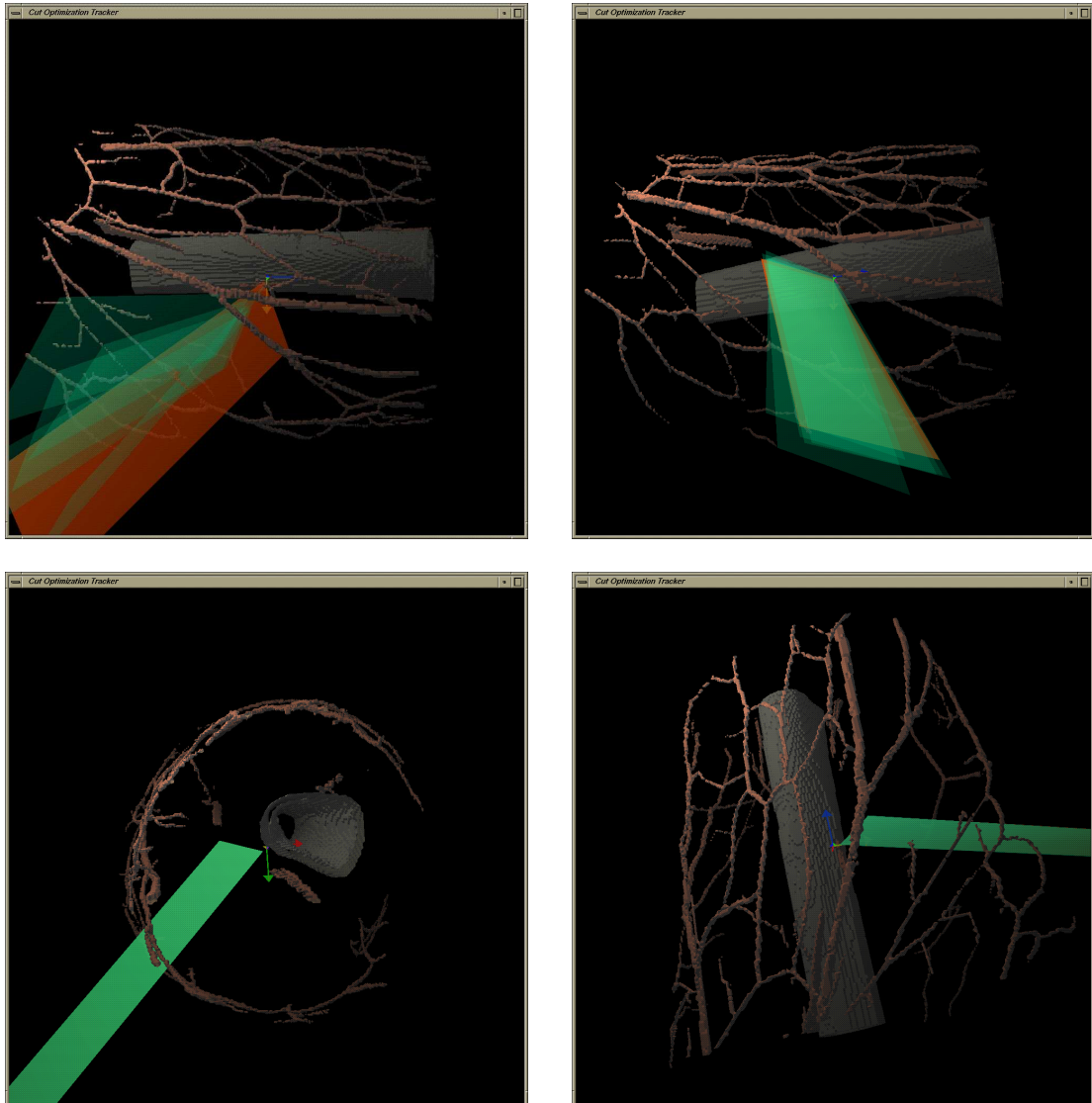


Abbildung 7.5. Verschiedene Stadien des Optimierungsprozeß: Die ersten drei Bilder zeigen den selben Optimierungslauf zu aufeinanderfolgenden Zeitpunkten (*oben links, oben rechts, unten links*). Das Bild unten rechts zeigt eine andere Lösung, die einmal mehr verdeutlicht, daß die Geweberichtung bei der Optimierung noch nicht berücksichtigt wird. In allen Bildern bestand die Zielvorgabe darin, den Ursprung des farblich hervorgehobenen Achsenkreuz' zu erreichen.

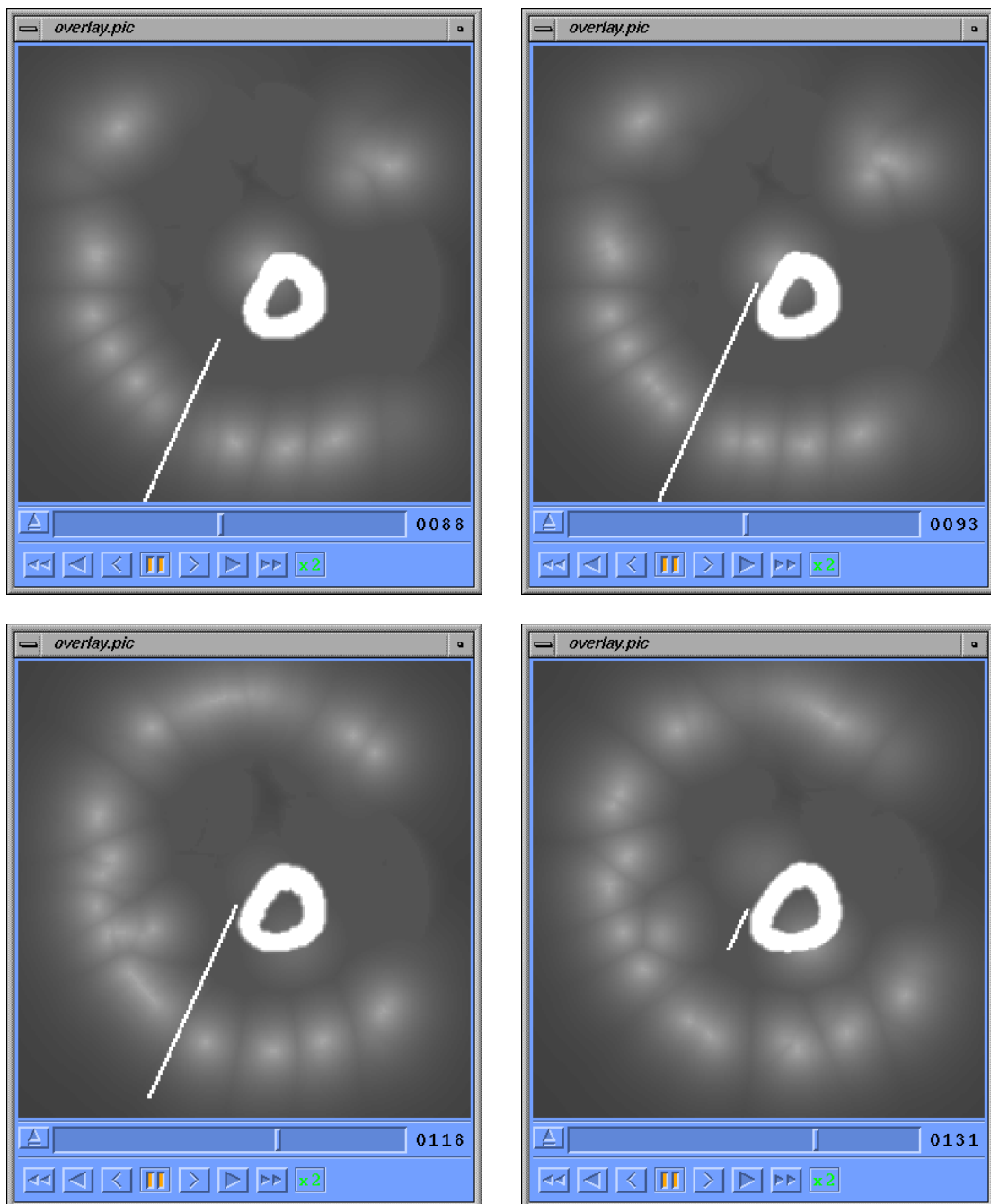


Abbildung 7.6. Ein dem Risikopotential überlagerter Ergebnisschnitt: Die Schnitttrajektorie wurde in das Risikopotential gezeichnet. In den horizontalen Schnitten durch das Risikopotential tritt sie als helle Linie zutage. Bemerkenswert ist, wie sich die Schnittfläche einerseits an den Knochen anschmiegt und andererseits recht mittig zwischen den Potentialen der Blutgefäße hindurch verläuft.

Kapitel 8

Diskussion

8.1 Schnittbewertung

Die guten Ergebnisse der Schnittoptimierung sind ein Hinweis dafür, daß das Konzept der Bewertungsfunktion ein gangbarer Ansatz ist. Inwieweit die Parametrisierung von Risikopotential und Bewertungsfunktion aber mit der medizinischen Realität korrespondiert, ist noch unklar. Allein aufgrund der Optimierungsergebnisse schon Rückschlüsse auf die chirurgische Aussagekraft ziehen zu wollen, ist aber noch verfrüht, solange das Gewebefeld nicht in die Bewertung einbezogen wird.

Da das Risikopotential aber im Gegensatz zum Gewebefeld in jedem Punkt auch Information zu seinem Umfeld enthält, kann man unter alleiniger Verwendung des Risikopotentials schon Aussagen treffen, wie gut die Grobplanung durch den Schnittoptimierer erfolgt. Das heißt, ob der Optimierer selbständig einen Zugang findet, der in möglichst großem Abstand zu Gefahrenzonen verläuft. Lokale, durch den Winkel zwischen Geweberichtung und Schnittfläche verursachte Risikofaktoren seien dabei außen vor gelassen. Ein Beispiel eines Einschnitts, der zwar viele Muskelfasern durchtrennen würde, aber schon in einer sehr günstigen „Schneise“ liegt, zeigt Abbildung 7.5, unten rechts, auf Seite 33. Innerhalb dieser Schneise wäre der Schnitt unter Berücksichtigung der Geweberichtung noch ohne Probleme in eine optimale Position zu bringen.

Dies war zwar für alle gefundenen Lösungen der Fall, doch darf man das nicht überbewerten. Die Tatsache, daß das Haupthindernis bei der Wegfindung, die Blutbahnen, in etwa in gleicher Richtung wie die Muskelfasern des Oberschenkels verläuft, führt nämlich häufig zu Lösungen, die bereits günstig in bezug auf die Faserrichtung liegen. Fundierte Aussagen über die Qualität der Grobplanung lassen sich daher erst nach Einsatz in anderen Körperregionen treffen.

Schließlich hat aber auch der verwendete Optimierungsalgorithmus großen Einfluß auf die Qualität der Grobplanung. Insbesondere spielen die in 8.2.1 genannten Effekte eine große Rolle.

8.2 Schnittoptimierung

Die Implementierung der automatischen Schnittoptimierung will nur als erster Prototyp verstanden sein. Es gibt noch einige Möglichkeiten, sowohl die Qualität der gefundenen Lösungen als auch die Geschwindigkeit, in der der Algorithmus konvergiert, zu verbessern. Ein Schlüssel zu solchen Verbesserungen ist zum Beispiel die Art und Weise, wie man die μ_i aus Gleichung 5.1 zeitlich variiert, und die Frage, welche Basis man für den Konfigurationsraum wählt.

Darüberhinaus sind auch essentielle Erweiterungen des Optimierungsvorgangs denkbar:

8.2.1 Startkonfigurationen

Unabhängig vom verwendeten Optimierungsalgorithmus fängt der Minimierungsprozeß an einem zufällig bestimmten Punkt im Konfigurationsraum an. Beim Simplexabstiegsverfahren sind es in unserem Fall sogar sieben zufällige Startkonfigurationen, die im Laufe des Verfahrens in Richtung einer besseren Schnitfführung variiert werden. Da aber kein Optimierungsverfahren in der Lage ist, den kompletten Konfigurationsraum abzusuchen, erfaßt die Optimierung bei jedem Programmstart daher einen anderen Teil des Suchraums, der von diesen Startpunkten abhängt, weshalb sich die Lösungen der Optimierungsaufgabe jedesmal unterscheiden können.

Bei den Experimenten auf dem Oberschenkel Datensatz fiel allerdings auf, daß es nur wenige Klassen von Lösungen gibt, innerhalb derer sich die gefundenen Schnitfführungen nur marginal unterscheiden. Wohl unterscheidet sich aber das resultierende Schnitttrisiko der Klassen untereinander. Die Klassen liegen also in lokalen Minima der Bewertungsfunktion, die in tiefen, voneinander getrennten „Tälern“ der Funktion liegen. Stochastische Optimierungsalgorithmen verfügen zwar über Mechanismen, aus lokalen Minima herauszukommen, ganz läßt sich dieser Effekt aber schon allein wegen der Größe des Suchraums nicht vermeiden.

Dieser Zustand ist allerdings unbefriedigend, da man nie sicher sein kann, ob es nicht noch eine wesentlich bessere Lösung gibt, als die, die von der Schnittoptimierung beim ersten Lauf geliefert worden ist.

Möchte man diese Unsicherheit verringern, liegt es nahe, den Schnittoptimierer einfach mehrmals aufzurufen, um sich dann für die beste der gefundenen Lösungen zu entscheiden. Dieser Weg kann aber eine Vielzahl von Optimierungsläufen erfordern und ist daher nicht besonders effizient.

Eine andere Herangehensweise baut auf der Vermutung auf, daß Schnitte, die in der gleichen Lösungsklasse liegen, von Startpunkten aus gefunden worden sind, die aus einem gemeinsamen Gebiet des Konfigurationsraums stammen. Das bedeutet, daß man über die Startkonfiguration die Lösungsklasse relativ direkt auswählen kann. Würde man also den Konfigurationsraum nur hinreichend dicht mit Startpunkten ausfüllen, von denen aus man jeweils einen Optimierungsvorgang startet, sollte man alle relevanten Lösungen erhalten. Das unterscheidet sich aber nicht wesentlich vom oben vorgeschlagenen Mehrfachaufruf des Optimierers. Mein Vorschlag ist nun, die von den Startpunkten ausgehenden Optimierungsläufe *parallel* zu verfolgen. So ist es möglich, in regelmäßigen Abständen den Fortgang der unterschiedlichen Prozesse zu vergleichen. Verfolgen zwei Optimierungsprozesse Punkte im Konfigurationsraum, deren Abstand dort unter eine gewisse Schwelle fällt, kann einer der beiden Prozesse abgebrochen werden, da man davon ausgehen kann, daß beide auf die gleiche Lösungsklasse zustreben. Auf diese Weise werden anfangs eine Unzahl Optimierungsvorgänge angestoßen, um den Suchraum möglichst vollständig auszuleuchten. Mit fortschreitender Zeit wird deren Zahl geringer, so daß gegen Ende nur noch wenige Optimierungsläufe bis zum Ende durchgerechnet werden müssen.

Ähnliche Effekte würde man auch mit einem evolutionären Algorithmus erreichen, bei dem man die Populationsgröße mit der Zeit absenkt. Die Selektion würde über die Güte des Schnitts erfolgen, und statt der Mutation ließe man jeden Schnitt sich mit einem „herkömmlichen“ Optimierungsalgorithmus weiterentwickeln. Hier kann es aber im Gegensatz zum von mir vorgeschlagenen Verfahren passieren, daß Wege frühzeitig abgebrochen werden, die eventuell noch zu einer brauchbaren Lösung geführt hätten.

Wie dieser Diskussion zu entnehmen ist, gibt es noch viele Möglichkeiten, den bestehenden Optimierungsalgorithmus zu verfeinern. Die Kombination aus Simplexabstiegsverfahren und simuliertem Tempern als Kern beizubehalten, empfiehlt sich aber in jedem Fall, da das Verfahren bereits auf Anhieb vielversprechende Ergebnisse geliefert hat.

8.2.2 Erweiterungen der Optimierungsaufgabe

In Kapitel 5 wurde die Optimierungsaufgabe bewußt so einfach wie möglich gehalten, um sich bei den ersten Versuchen auf die Grundstruktur des Problems konzentrieren zu können. Daher findet man aber auch schnell weitere Anforderungen, die man an die Schnittfindung stellen könnte.

Insbesondere ist die Festlegung auf eine feste Schnittform- und -größe eine ungünstige Einschränkung. Man könnte daher fordern, daß das Verfahren von sich aus eine optimale Gestalt des Schnitts finden soll.

Dazu müssen aber auch neue Randbedingungen für die Schnittgestaltung gestellt werden, da eine Optimierung, die die Schnittform beliebig verändern darf, um \hat{E} zu minimieren, innerhalb kürzester Zeit einen infinitesimal dünnen Schnitt liefern würde. Dies minimiert zwar erfolgreich die Oberflächenintegrale, die der Bewertungsfunktion zugrunde liegen, stellt aber einen Chirurgen, der mit einem solchen Schnitt operieren soll, vor unlösbare Probleme. Die Zugänglichkeit der Zielregion und der nötige Arbeitspielraum im Gewebe wären nicht mehr gegeben.

Möchte man solche Randbedingungen mit in die Optimierung aufnehmen, ist es am elegantesten, dies durch Erweiterung der zu optimierenden Funktion zu tun. So könnte man in \hat{E} noch einen Formfaktor F einführen, der beschreibt, wie wohlgeformt der Schnitt ist. Gleichung 5.1 würde dann wie folgt angepaßt:

$$\hat{E}(\mathcal{C}) = \mu_1 R(\mathcal{C}) + \mu_2 V(\mathcal{C}) + \mu_3 \text{dist}^2(\mathcal{C}, \text{Zielregion}) + \mu_4 F(\mathcal{C}) \quad (8.1)$$

Es bleibt aber das Problem, daß in großem Umfang chirurgisches Fachwissen in die Bestimmung von F einfließen muß. — Wissen, welche Schnittführung dem Chirurgen genügend Platz bietet, um im Zielgebiet zu operieren, oder wo die Grenzen des Machbaren liegen, was bestimmte Schnittformen unter Verwendung eines konkreten Werkzeugs angeht.

Bevor man also den Rechner automatisch Formfaktoren bestimmen lassen kann, müssen erst noch grundlegende Arbeiten in der Repräsentation chirurgischen Fachwissens abgeschlossen werden.

Das gleiche Problem ergibt sich, wenn man den Rechner *Schnittfolgen*, d.h. eine Aneinanderreihung einzelner Operationsschritte, finden lassen will. Auch hier muß zusätzliches Wissen akquiriert werden, um zu brauchbaren Ergebnissen zu kommen.

Die bisher vorgenommene Einschränkung, daß Schnitte nur planare Flächen sein dürfen, mag auch aufgehoben werden. Es wird für einen Arzt nur schwierig werden, mit einer planen Skalpellklinge gekrümmte Trajektorien zu ziehen, die ihm ein Rechner vorschreibt. Andererseits gibt es auch gekrümmte Instrumente wie Löffel, Küvetten, etc., für die die minimalinvasive Schnittführung eher der Oberfläche eines Rotationskörpers entsprechen dürfte.

Kapitel 9

Zusammenfassung und Ausblick

9.1 Zusammenfassung

Mit den in Kapitel 4 behandelten Bewertungsgrundlagen wurde ein Ansatz vorgestellt, der eine erste Risikoabschätzung für Schnitte in Weichgewebe ermöglicht.

In diesem Zusammenhang sind die isotrope und die anisotrope Risikokomponente eingeführt worden.

Die isotrope Risikokomponente wird aus den segmentierten Gewebedaten gewonnen und leistet eine grobe Abschätzung des Operationsrisikos. Sowohl die Geweberelevanz, als auch das Risiko, durch Abweichung von der geplanten Schnittführung unerwünschte Verletzungen herbeizuführen, werden beschrieben. Sie dient als Komponente der Bewertung dazu, die Grobplanung eines Schnitts zu bewerten.

Demgegenüber erlaubt die anisotrope Risikokomponente eine genauere Abschätzung des Risikos durch Berücksichtigung der Geweberichtung. Dafür fließen hier aber weniger globale Zusammenhänge ein, da das zugrundeliegende Gewebefeld nur lokale Information hält, während innerhalb des Risikopotentials jede Gewebeklasse eine gewisse „Fernwirkung“ entfaltet. Die anisotrope Risikokomponente findet daher Verwendung bei der Bewertung der Feinplanung eines Schnitts. Für die Bestimmung der anisotropen Risikokomponente müssen anatomische Atlanten zu Rate gezogen werden, die die Faserrichtung des Gewebes modellieren.

Um diesen Ansatz evaluieren zu können, wurde eine automatische Schnittoptimierung implementiert. Dafür ist das Problem, eine optimale Schnittführung zu finden, die ein Zielgebiet unter möglichst geringem Risiko erreicht, auf die Minimierung einer skalarwertigen Funktion abgebildet worden (siehe Kapitel 5). Die Implementierung basiert bis jetzt nur auf der isotropen Risikokomponente, das Gewebefeld wurde noch nicht behandelt.

An die Implementierung sind erste Versuche mit dem Schnittoptimierer angeschlossen worden. Dabei wurde eine Schnitttrajektorie konstanter, planarer Form gewählt, die im Datensatz eines menschlichen Oberschenkels risikominimal ausgerichtet werden sollte. Vorgabe war, einen Zielpunkt im Inneren des Schenkels zu erreichen.

Die ersten Ergebnisse waren vielversprechend und führten zu subjektiv guten Schnittführungen. Eine medizinische Auswertung der Ergebnisse muß noch folgen.

9.2 Ausblick

Angesichts der Komplexität des Themas konnte nur ein Teil der Problematik in Angriff genommen werden. In vielen Bereichen besteht noch Handlungsbedarf. Im folgenden seien gezielt einige Bereiche genannt.

9.2.1 Risikomodellierung

Es ist zu untersuchen, ob die bestehende Form der Risikomodellierung alle Anforderungen erfüllt.

Insbesondere können für die Bildung des Risikopotentials Alternativen ausprobiert werden. Zum einen ist der Einsatz von e-Funktionen als Risikofunktionen nicht verbindlich und deren Parametrisierung noch zu evaluieren; in Anlehnung an das elektrische Potential könnte man auch Hyperbelfunktionen verwenden. Zum anderen ist die Frage nach einer optimalen Überlagerung der Einzelpotentiale zum Gesamtrisikopotential noch nicht geklärt. Es wäre zum Beispiel denkbar, zu einer additiven Überlagerung überzugehen, um an jeder Position im Gewebe *alle* Risikoeinflüsse zu berücksichtigen. Dazu müßte noch das dadurch entstehende Skalierungsproblem gelöst werden (siehe 7.3). Schließlich ist noch zu klären, ob es sinnvoll ist, die dem Risikopotential zugrundeliegende Abstandstransformation durch ein Feld gemäß der Poissonschen Gleichung zu ersetzen. Dabei ist zu untersuchen, ob damit ein „Qualitätsgewinn“ des Risikomodells verbunden ist und ob dies den größeren Zeitbedarf der Herleitung rechtfertigt.

Da die Berücksichtigung der Geweberichtung für eine realistische Risikoabschätzung unverzichtbar ist, müssen möglichst bald Arbeiten zum Gewebefeld durchgeführt werden. Es müssen geeignete anatomische Atlanten gefunden werden, die Rückschlüsse auf die Geweberichtung zulassen. Dabei wird die Hauptschwierigkeit darin liegen, die dort enthaltene Information mit nach Gewebeklassen segmentierten Aufnahmen des Patienten zu korrelieren (sog. *matching*).

9.2.2 Bewertungsfunktion

Die Erfahrungen mit der automatischen Schnittoptimierung haben gezeigt, daß der Ansatz auf den ersten Blick zu brauchbaren Ergebnissen führt. Insbesondere scheint die Parametrisierung der Bewertungsfunktion schon recht günstig zu sein.

Es ist aber unbedingt notwendig, nun eine Evaluation der Bewertungsfunktion anzuschließen. Mediziner müssen befragt werden, inwieweit sich medizinisch relevante Aussagen aus den Ergebnissen der Bewertungsfunktion ableiten lassen. Es ist auch weiterhin eine indirekte Evaluation über die automatische Schnittoptimierung sinnvoll: In diesem Falle sollte man dem Rechner Schnittaufgaben geben, zu denen chirurgische Standardmethoden existieren, um diese dann mit den Lösungen zu vergleichen, die der Rechner geliefert hat.

9.2.3 Schnittoptimierung

Soll die Anwendung einer automatischen Schnittoptimierung weiter verfolgt werden, sind Untersuchungen notwendig, welchen Einfluß die Startbedingung der Optimierung auf das Ergebnis hat. Das Verfahren ist dahingehend anzupassen, daß dieser Einfluß minimiert wird. Dazu könnte zum Beispiel, wie in Kapitel 8 diskutiert, der Optimierungsalgorithmus mit evolutionären Methoden oder einer modifizierten Breitensuche kombiniert werden, um eine vollständigere Erfassung des Suchraums zu gewährleisten.

Damit der Optimierer möglichst allgemein bleibt und man ihn nicht zu stark an den verwendeten Beispieldatensatz anpaßt, muß das Verfahren möglichst bald auf unterschiedliche Körperregionen angewendet werden.

Um die chirurgische Relevanz der Optimierungsergebnisse zu verbessern, muß sowohl eine veränderlicher Form der Schnitttrajektorie zugelassen werden, als auch eine Hintereinanderausführung von Schnitten als zulässige Lösung betrachtet werden. Beides erfordert das Einbringen chirurgischen Fachwissens in den Optimierungsprozeß.

Schließlich machen einzelne Schnitte noch keine Operation aus, so daß auch noch wei-

tere Arbeitsschritte eines Chirurgen modelliert und in den Suchraum aufgenommen werden müssen.

9.2.4 weitere Anwendungen

Desweiteren wäre es interessant, noch weitere Anwendungen der Schnittbewertung zu implementieren. Exemplarisch wurden bereits in Kapitel 1 mögliche Anwendungsfelder genannt. Erste Arbeiten zur haptischen Darstellung eines Risikopotentials haben wir bereits durchgeführt.

Anhang A

Die Job-Datei

Die Bildung des Risikopotentials und die Parameterwahl des Schnittoptimierers werden über eine ASCII-Datei gesteuert. Im Folgenden sei exemplarisch eine Datei wiedergegeben, die in unseren Experimenten verwendet wurde. Zur ausführlichen Syntaxbeschreibung sei auf die Programmdokumentation verwiesen. Der interessierte Leser mag sich an der Kommentierung orientieren.

```
;;;
;;; Title: op.job
;;; Created: Sat Nov 28 00:14:37 1998
;;; Author: Tim Weyrich <unj8@rz.uni-karlsruhe.de>
;;; $Id: op.job,v 1.3 1998/12/23 17:31:49 weyrich Exp $
;;;
;;; copyright (c) 1998 by quasi.mod0 software, Tim Weyrich
;;;
;;; $Log: op.job,v $
;;; Revision 1.3 1998/12/23 17:31:49 weyrich
;;; added further comments
;;;
;;; Revision 1.2 1998/12/04 23:05:52 weyrich
;;; found reasonable values
;;;
;;; Revision 1.1 1998/11/28 01:16:09 weyrich
;;; Initial revision
;;;
;;;

;;; *** Job-Datei fuer COP, den Schnittoptimierer fuer Weichgewebe ***

;; Alle Dateinamen werden an 'path' angehaengt, wenn sie nicht
;; wie angegeben gefunden werden koennen.

;; Dateinamen sind ohne Suffix anzugeben, damit die Endungen .pol,
;; .pol.gz, .edt.pic, .edt.pic.gz, .pic und .pic.gz in der angegebenen
;; Reihenfolge ausprobiert werden koennen.

;; Angaben hinter der tissue-Direktive:
;;
;; 'estmaxdist' ist der Wert, den dil bei der EDT ausspuckt! (Ok, ich
;; geb' zu, das ist nicht gerade der sauberste Weg, aber es muss grade
;; mal schnell gehen -- dieser Zustand wird geaendert!)
;;
;; Auf 'abcd' folgen die Paramter a, b, c und d der Risikofunktion --
```

```

;; je nach Implementierung muss diese nicht alle Parameter nutzen.

path          PIC/
source        leg_base
view          leg_9

overlaymode   max          ; es sind max, add und mul moeglich

;; Risikofunktionen zur Maximumsueberlagerung
;tissue 1     leg_1 estmaxdist 258 abcd 1.0 5.0 0 0   ; Knochenmark
;tissue 2     leg_2 estmaxdist 260 abcd 0.1 1.0 0 0   ; Fett
;tissue 3     leg_3 estmaxdist 294 abcd 100 0.75 0 0  ; Knochen
;tissue 8     leg_8 estmaxdist 286 abcd 1.0 5.0 0 0   ; Muskel
;tissue 9     leg_9 estmaxdist 402 abcd 5.0 12.0 0 0  ; Blut

;; Parameter der Bewertungsfunktion
mu[0] = 1.0
mu[1] = 0.0
mu[2] = 10.0

;; Optimierungsparameter
thresh = 1500.0          ; wird nur vom Schwellwertverfahren genutzt

;; Fuer meine Oberschenkelprobe: Die 'Schnittflaechen' des
;; Oberschenkelstumpfs muessen noch als undurchdringbar gekennzeichnet
;; werden. Hierzu sollte ein Wert verwendet werden, der wesentlich
;; groesser als 'VERY_HIGH_VALUE' ist (definiert in op.c und
;; pic2handypic.c), damit diese in der Lage sind, diese hohen Werte bei
;; der min/max-Analyse auszublenden.

;; Bei der Wahl der Koordinaten ist zudem zu beachten, dass
;; Schnittflaechen, die mit den Raendern der Aufnahme zusammenfallen
;; (sollte der Normalfall sein ;-), nicht /exakt/ auf diese Raender
;; gesetzt werden duerfen, da aus numerischen Gruenden das
;; Clipping-Volume fuer alle Polygone/Polygonintegrale in Voxelbildern
;; etwas kleiner als das eigentliche Voxelvolumen ist. (Und zwar ist es
;; an jeder Seite um 'SHRINK_CLIPPING_VOLUME_AMOUNT' (definiert in
;; op.c) geschrumpft.)

pot_draw_polygon  {-10000.0, -10000.0, 0.1,          \
                  10000.0, -10000.0, 0.1,          \
                  10000.0,  10000.0, 0.1,          \
                  -10000.0,  10000.0, 0.1} 1e32    ; wesentlich groesser als 1e30

pot_draw_polygon  {-10000.0, -10000.0, 184.9, \
                  10000.0, -10000.0, 184.9, \
                  10000.0,  10000.0, 184.9, \
                  -10000.0,  10000.0, 184.9} 1e32 ; wesentlich groesser als 1e30

```

Literaturverzeichnis

- [1] W. Pschyrembel. *Klinisches Wörterbuch*. de Gruyter, 257. Aufl., Berlin · New York 1994.
- [2] R. F. Schmidt und G. Thews. *Physiologie des Menschen*. Springer Verlag, 27. Aufl., Berlin · Heidelberg · New York 1997.
- [3] H. Lippert. *Lehrbuch Anatomie*. Urban & Schwarzenberg, 4. Aufl., München · Wien · Baltimore 1996.
- [4] I. N. Bronstein, A. Semendjajew, G. Musiol und H. Mühlig. *Taschenbuch der Mathematik*. Verlag Harri Deutsch, 3. Aufl., Frankfurt am Main · Thun 1997.
- [5] R. Dillmann und M. Huck. *Informationsverarbeitung in der Robotik*. Springer Verlag, Berlin · Heidelberg · New York 1991.
- [6] TU München. *Chirurgisches Klinikmanual*. <http://nt1.chir.med.tu-muenchen.de/manual/>, Stand Januar 1999.
- [7] B. R. Oswald, U. Tochtermann, P. Schweiger, G. Thomas, C. F. Vahl und S. Hagl. *Lassen sich konkrete Fragen des Patienten zu Risiko und Überleben mit statistischen Methoden beantworten?* Abstraktband der Tagung „Rechner- und sensorgestützte Chirurgie: neue Ansatzpunkte für eine wissenschaftliche Qualitätssicherung“, Heidelberg 1998.
- [8] G. Trummer, R. Hillmann, C. Wuestenberg, M. Becker, A. Geiger und F. Beyersdorf. *Etablierung eines postoperativen Risikoscores in der herzchirurgischen Intensivmedizin*. Abstraktband der Tagung „Rechner- und sensorgestützte Chirurgie: neue Ansatzpunkte für eine wissenschaftliche Qualitätssicherung“, Heidelberg 1998.
- [9] U. Jendrysiak, S. Gregg und J. Weinert. *Strukturspezifische Segmentierung mit NeurOPS für die computergestützte Operationsplanung*. Tagungsband zum Workshop „Digitale Bildverarbeitung in der Medizin“, Klinikum Freiburg, S. 63, Freiburg 1997.
- [10] A. Mazura. *Virtuelles Schneiden in Volumendaten*. Shaker Verlag, Aachen 1997.
- [11] Ch. Schneider. *Resampling medizinischer Bilddaten*. Studienarbeit, Karlsruhe 1998.
- [12] G. Lohmann. *Volumetric Image Analysis*. John Wiley & Sons / B.G. Teubner, Chichester / Stuttgart 1998.
- [13] J. A. Nelder und R. Mead. *Computer Journal*, Vol. 7, S. 391–398, 1965.
- [14] E. Schöneburg, F. Heinzmann und S. Feddersen. *Genetische Algorithmen und Evolutionsstrategien*. Addison-Wesley, Bonn · Paris · Reading, Mass. 1994.
- [15] W. H. Press, S. A. Teukolsky, W. T. Vetterling und B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 2nd edition, Cambridge · New York · Melbourne 1992.

- [16] T. Lehmann, W. Oberschelp, E. Pelikan und R. Repges. *Bildverarbeitung für die Medizin*. Springer Verlag, Berlin · Heidelberg · New York 1997.
- [17] G. Sakas. *Dreidimensionale Bildrekonstruktion aus Ultraschall-Daten*. Spektrum der Wissenschaft, Dossier: Perspektiven in der Medizintechnik, S. 28–24, 1999.
- [18] U. Kleuker. Dissertation, Konstanz 1998.
- [19] H. Vogel. *Gerthsen Physik*. Springer Verlag, 19. Aufl., Berlin · Heidelberg · New York 1997.
- [20] U. Jendrysiak. *Segmentierung von Schnittbildern*. Spektrum der Wissenschaft, Dossier: Perspektiven in der Medizintechnik, S. 24–29, 1999.
- [21] M. Fathi, A. Schönwald und I. Hoffmann. *Ein fuzzyunterstütztes System zur Segmentierung dreidimensionaler NMR-Tomogramme*. Spektrum der Wissenschaft, Dossier: Perspektiven in der Medizintechnik, S. 29–31, 1999.
- [22] S. Wegner, H. Oswald, P. Wust und E. Fleck. *Segmentierung mit der Wasserscheiden-transformation*. Spektrum der Wissenschaft, Dossier: Perspektiven in der Medizintechnik, S. 31–33, 1999.
- [23] F. B. Sachse, Ch. Werner, M. Müller und K. Meyer-Waarden. *MEET Man: Erstellung und technische Parameter*. Institut für Biomedizinische Technik, Universität Karlsruhe, Karlsruhe 1997.
- [24] Institut für Biomedizinische Technik, Universität Karlsruhe. *MEET Man Projekt: Models for Simulation of Electromagnetic, Elastomechanic and Thermic Behaviour of Man*. http://www-ibt.etec.uni-karlsruhe.de/MEETMan_de.htm, Stand Januar 1999.
- [25] E. Rimon und D. E. Koditschek. *Exact Robot Navigation in a Dynamic World Using an Unsteady Diffusion Equation Strategy*. IEEE Trans. on Robotics and Automation, Vol. 8, S. 501–518, 1992.
- [26] U. Engelmann und A. Schröter. *PIC Image File Standard*. Technischer Bericht, Deutsches Krebsforschungszentrum, Abt. MBI, Heidelberg 1993.