

Kapitel 3

Aktuelle Konzepte der Bildverarbeitung

Die Zielsetzung der in dieser Arbeit entwickelten Verfahren ist das Bestimmen eines präzisen anatomischen Modells des Patientenauges. Die Grundlage für die Rekonstruktion dieses Bestrahlungsmodells ist die Auswertung der in Abschnitt 2.4.1 beschriebenen Bilddaten, die im Rahmen der Diagnostik vom Patienten angefertigt werden. Die Konzeption der Bildverarbeitungsalgorithmen hängt maßgeblich von den Eigenschaften des Bildmaterials ab.

Bevor im nachfolgenden Kapitel die neuen Bildverarbeitungskonzepte zur Lösung der gegebenen Aufgabenstellung erläutert werden, sollen in diesem Kapitel die verschiedenen Aspekte der Problemstellung analysiert und die aktuellen Konzepte der medizinischen Bildverarbeitung erläutert werden.

Die in dieser Arbeit beschriebene modellbasierte Bildverarbeitung verfolgt eine von aktuellen Bildverarbeitungsansätzen abweichende Herangehensweise. Um einerseits das Verwerfen aktueller Bildverarbeitungsansätze, andererseits die Konzeption der modellbasierten Bildverarbeitung nachvollziehen zu können, wurde in den nachfolgenden Abschnitten zunächst die grundsätzliche Problematik bei der Bildverarbeitung beleuchtet. Hierzu werden in den Abschnitten 3.1 und 3.2 Überlegungen zu der Herangehensweise an die Bildverarbeitung und das Bildmaterial vorgenommen. Anschließend werden in den Abschnitten 3.3 bis 3.5 abstrakte Konzepte zur Auswertung von Bildmaterial, die von den visuellen Fähigkeiten des Menschen abgeleitet oder im Rahmen der Forschungsarbeiten im Bereich der künstlichen Intelligenz entwickelt wurden, erläutert. Abschnitt 3.6 sammelt grundlegende Lösungsansätze, die bei dem Entwurf eines Bildverarbeitungsansatzes, der die vorangegangenen Überlegungen berücksichtigt, einbezogen werden müssen. Dabei werden Kriterien herausgearbeitet, die anschließend auf die aktuellen Ansätze der medizinischen Bildverarbeitung, beschrieben in den Abschnitten 3.7 bis 3.11, angewandt werden, um die Vor- und Nachteile der bisherigen Herangehensweisen herauszuarbeiten.

3.1 Das Problem „Bildverarbeitung“

Unter dem Begriff der Bildverarbeitung sind sämtliche Verfahren zu verstehen, die angewandt werden, um Bildmaterial mit einer vorgegebenen Zielsetzung nutzbar zu machen. Im Hinblick auf die medizinische Nutzung von Bildmaterial dient die Bildverarbeitung der Unterstützung des Arztes bei Diagnose, Therapie und Nachsorge von pathologischen Gewebeveränderungen bei einem Patienten. Welche Möglichkeiten die Auswertung der Daten bietet und welche Verfahren hierzu angewandt werden müssen, hängt in erster Linie von dem Bildmaterial ab. Um eine inhaltliche Interpretation aus einem gegebenen Bildmaterial zu gewinnen, müssen die Informationen des Bildes in eine abstrakte inhaltliche Form gebracht werden. Das Problem bei diesem Schritt besteht darin, daß die Informationen implizit in der Konstellation der Intensitätswerte der einzelnen Bildpunkte kodiert sind. Welche Probleme und Hürden bei dem Dekodieren dieser Intensitätsverteilungen eines Bildes auftreten können, sollen an einem lebensnahen Beispiel, dem Erkennen eines Stuhles, und anschließend an dem zu bearbeitenden Bildmaterial erläutert werden.

Abbildung 3.1 zeigt verschiedene Fotografien von Stühlen. Ein normaler Beobachter wird ohne jede Zeitverzögerung die dargestellten Objekte als Stühle erkennen und vom Hintergrund abgrenzen können. Für einen Computer ist dies eine außerordentlich schwere Aufgabe, die nur durch einen sehr großen Entwicklungsaufwand bewältigt werden könnte.



Abbildung 3.1: Veranschaulichung von Bildeigenschaften anhand von vier Stuhl-Bildern

Während alle Bilder den gleichen Gegenstand, einen Stuhl, darstellen, könnte aufgrund der veränderten Bild- und Darstellungseigenschaften die Problemstellung für das visuelle Verarbeitungssystem des Menschen kaum unterschiedlicher sein.

Betrachten wir das ganz rechte Bild von Abbildung 3.1 mit dem Holzstuhl. In den ersten Hundertstel Sekunden wird der Beobachter das Bild an sich wahrnehmen. Dabei werden farblich zusammengehörige Bildpunkte zusammengefaßt. Da Holz eine Maserung ausweisen kann, müssen die verschieden farbigen Bildpunkte des Holzes erkannt und zu einem Objekt zusammengefaßt werden. Schatten und perspektivische Veränderungen der Farbe müssen ebenso mit einbezogen werden wie unter Umständen die Farbe der Lichtquelle. Anschließend muß der allgemeine Kontext des Bildes ermittelt werden: handelt es sich um eine Landschaftsaufnahme, tritt ein bestimmtes Objekt in den Vordergrund, gibt es einen Hinter- und Vordergrund und vieles mehr.

Nachdem festgestellt wurde, daß es sich um ein Foto mit einem Objekt in der Mitte handelt, muß das Objekt als Stuhl erkannt werden, indem Form, Farbe/Maserung, Geometrie, Erscheinungsbild und viele andere Faktoren mit zuvor erworbenen Wissen über Stühle verglichen werden. Erst wenn die notwendigen Kriterien als erfüllt festgestellt wurden, kann der Mensch annehmen, daß es sich bei dem betrachteten Objekt möglicherweise um einen Stuhl handelt. Dieser in so einfachen Schritten beschriebene Erkennungsprozeß ist in Wirklichkeit ein überaus komplexer kognitiver und bis heute nur sehr wenig verstandener Wahrnehmungsprozeß, den der Mensch millionenfach jeden Tag in unendlicher Vielfalt unentwegt durchführt.

Wie komplex dieser Erkennungsprozeß, Perzeption genannt, in Wirklichkeit ist, sollen die weiteren Überlegungen veranschaulichen: Bei dem Zusammenfassen von zusammengehörigen Bildpunkten stoßen wir auf das Problem, daß keine Regeln bezüglich Farbe, Kontrast oder sonstiger Parameter aufgestellt werden können, um die Stühle vom Hintergrund abzugrenzen. Ein weiterer Effekt, der für den Menschen zwar hilfreich ist, aber die computergestützte Erkennung erschwert, ist die Veränderung von Objekteigenschaften durch Schatten, perspektivische Verkürzungen, Maserungen, Farben auf der Objektfläche und Gegenstände im Hintergrund. Bei der Identifizierung des Objektes als Stuhl muß berücksichtigt werden, daß das Objekt aus den verschiedensten Perspektiven betrachtet werden kann und damit zu rechnen ist, daß Teile des Stuhls durch andere Objekte verdeckt werden. Hinzu kommen grundsätzliche Unterschiede in der Form, in der Konstruktion von Sitzfläche, Rückenlehne und gegebenenfalls Armlehnen.

Die Problemstellungen bei der Verarbeitung von medizinischem Bildmaterial stimmen mit denen von realen Bildern weitgehend überein. Erschwerend kommt für den ungeübten Betrachter hinzu, daß zur

Interpretation der CT- und MRT-Bilder medizinisches Vorwissen notwendig ist, um eine korrekte Interpretation vornehmen zu können. Abbildung 3.2 stellt beispielhafte Aufnahmen des Auges in unterschiedlichen Bildmodi mit variierenden Bildeigenschaften dar. Schon das Erkennen der linken oberen CT-Aufnahme setzt viel Vorwissen über die Anatomie des Schädels und über die bildgebenden Eigenschaften der Computertomographie voraus. Der Beobachter muß wissen, daß der schwarze obere Bereich Luft und damit der Hintergrund ist. Das Erkennen der Knochenstrukturen erfordert zunächst das ungefähre Einordnen des Bildes in die menschliche Anatomie. Das heißt, der Benutzer muß wissen, wo ungefähr die Schicht innerhalb des Schädels positioniert ist. Erst nachdem diese grundsätzliche Interpretation vorgenommen wurde, kann der Betrachter die grauen Weichgewebestrukturen genauer untersuchen und beispielsweise die Augen oberhalb des Schädels erkennen.

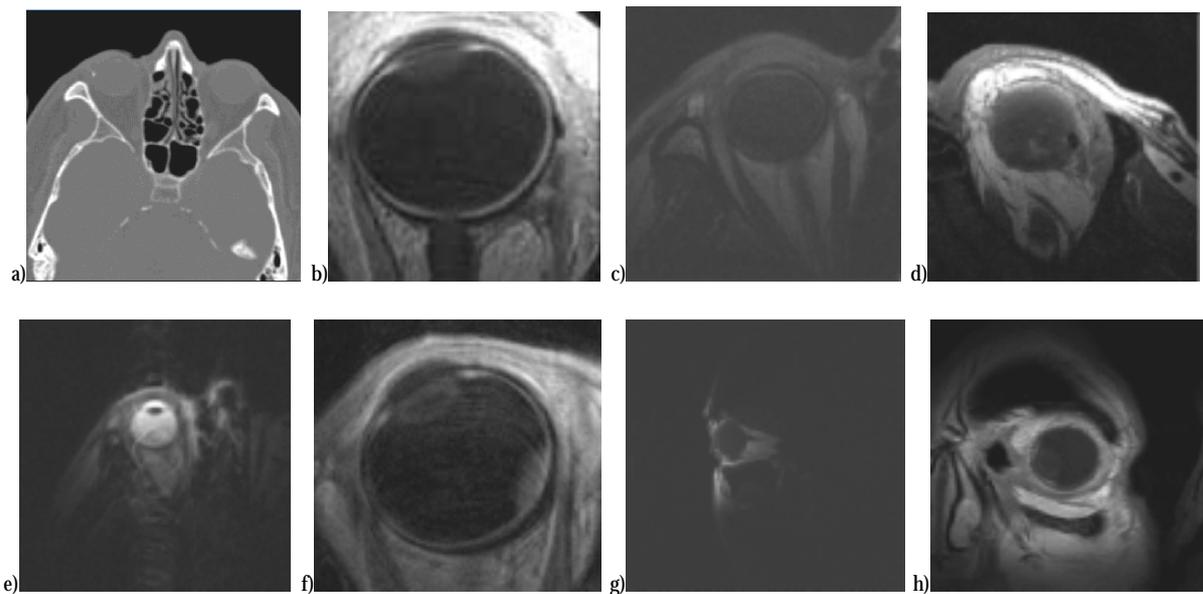


Abbildung 3.2: Die ophthalmologischen Bildmodi und ihre Darstellungseigenschaften

Die Abbildungen zeigen typisches Bildmaterial, auf dem die Therapie von Augentumoren aufsetzt. In allen Fällen wird das Auge eines Patienten dargestellt. Oben links: CT-Aufnahme mit dem Schädel (a); Zweites Bild von links: axiale T1-Aufnahme des Auges (b); Zweites Bild von rechts: axiale T1-Aufnahme mit Auges mit anderen Grauwerten und mit Bewegungsartefakten (c); Rechts: axiale T1-Aufnahme des Auges in einer Randschicht des Auges (d). Unten links: T2-Aufnahme des Auges (e); Zweite von links: axiale T1-Aufnahme des Auges mit deutlich anderen Grauwerteigenschaften als bei der Grafik darüber (f); Zweite von rechts: sagittale T1-Aufnahme des Auges (g); Rechts: dorsale T1-Aufnahme des Auges (h).

Die Fähigkeit zur Interpretation der CT-Aufnahme hilft jedoch nicht bei der Auswertung der Kernspindaten rechts daneben. Die Darstellung des Weichgewebes, der Bildausschnitt, der Kontrast und viele weitere Eigenschaften der Darstellung weichen so stark von der CT-Aufnahme ab, daß anderes Vorwissen benötigt wird, um den dargestellten Bildinhalt zu erkennen. Erschwerend kommt hinzu, daß nicht nur die Form und die Orientierung der Organe innerhalb des Bildes aufgrund der wechselnden Anatomie der Patienten stark variiert. Darüber hinaus können aufgrund von Artefakten auch die Bildeigenschaften sehr unterschiedlich sein. Die beiden Bilder rechts oben zeigen ebenfalls das Auge. Jedoch werden in einem Bild alle Strukturen dunkler dargestellt und in dem ganz rechten Bild ist die Bildschicht bezogen auf das Auge anders lokalisiert.

In der unteren Reihe von Abbildung 3.2 werden weitere Aspekte verdeutlicht, auf die ein Bildverarbeitungsalgorithmus eingehen muß. Ganz links ist anstatt einer T1- eine T2-Aufnahme des Auges dargestellt. Wie leicht zu erkennen ist, wird in diesem Fall der Bulbus hell und die Linse dunkel dargestellt. Bei den T1-Aufnahmen ist der Bulbus hingegen dunkel und die Linse dunkelgrau. Während sich der Mensch

ohne weiteres auf solche Veränderungen einstellen kann, machen sie bei der computergestützten Bildverarbeitung ein grundsätzliches Anpassen der Algorithmen erforderlich, da der Computer nicht in der Lage ist, solche Änderungen logisch von sich aus einzuordnen.

Eine weitere sehr wichtige Fähigkeit des Menschen zur Wahrnehmung visueller Informationen ist das Ergänzen von fehlenden Informationen und das Herausfiltern von fehlerhaften Bildfragmenten. Sowohl in der T2-Aufnahme ganz links unten als auch in der T1-Aufnahme rechts daneben sind Bewegungsartefakte in Form von hellen Reproduktionen der Augenkontur dem Bild überlagert. Der Mensch ist in der Lage diese herauszufiltern. Um diese Fähigkeit in einem Algorithmus zu realisieren, müßten sehr komplexe kognitive Prozesse nachgebildet werden, die in der Lage wären, den Gesamtinhalt des Bildes zu erkennen und anschließend aufgrund dieser Erkenntnis die überlagerten Fehler als solche auszumachen.

Die letzten beiden Bilder rechts unten veranschaulichen, wie stark die Darstellung des Auges variiert, wenn die Orientierung der Schnittebene verändert wird. In dem Bild rechts unten ist ein sagittaler, in dem Bild links daneben ein dorsaler Schnitt durch das Auge zu sehen. Die übrigen Aufnahmen in Abbildung 3.2 sind axiale Aufnahmen. Selbst für den Menschen erschwert diese veränderte Schichtorientierung die Interpretation der Daten. Anhand des Bildmaterials wird ersichtlich, welche komplexen Problemstellungen bei dem Entwurf eines computergestützten Algorithmus berücksichtigt werden müssen.

Diese Erläuterungen veranschaulichen die Komplexität der Interpretation von medizinischen Bilddaten. Der Mensch ist in der Lage, komplexe visuelle Problemstellungen durch unbewußte und bewußte intellektuelle Verarbeitungsschritte zu analysieren und miteinander zu assoziieren. Sofern erforderlich, kann er ohne Zeitverzögerung spezifisches und unspezifisches Vorwissen zur Unterstützung der Interpretation heranziehen. Dadurch ist die Interpretation von Bildern mitunter vom persönlichen Erfahrungshorizont des Betrachters abhängig. Die Strategien, mit der der Mensch die einzelnen Problemstellungen bewältigt, sind jedoch nur unvollständig verstanden.

Die Umsetzung eines Systems zur Auswertung des oben dargestellten medizinischen Bildmaterials ist deswegen so schwierig, weil einerseits der Computer nicht über die intellektuellen Fähigkeiten eines Menschen verfügt. Andererseits hat es die Wissenschaft noch nicht geschafft, ein System zu entwickeln, das diese Fähigkeiten auch nur annäherungsweise nachbilden kann.

Um sich dennoch der Lösung des Problems Bildverarbeitung anzunähern, soll in einem ersten Schritt die zu bearbeitende Aufgabenstellung, in diesem Fall das Bildmaterial, genauer untersucht werden. Erst wenn die Eigenschaften, Problemstellungen und Variabilitäten des Bildmaterials verstanden sind, kann ein Bildverarbeitungsansatz zur Analyse des Bildmaterials entwickelt werden.

3.2 Bildeigenschaften

Um den möglichen Erfolg eines Bildverarbeitungsansatzes einschätzen zu können, muß der Ansatz im Zusammenhang mit dem Bildmaterial, das bearbeitet werden soll, gesehen werden. Im Unterschied zum Menschen sind computergestützte Bildverarbeitungsalgorithmen nicht in der Lage, sich dynamisch auf variierende Bildeigenschaften einzustellen. Bevor also die Frage beantwortet werden kann, wie ein Bildverarbeitungsansatz zu entwerfen ist, muß untersucht werden, welches Bildmaterial ausgewertet werden soll. Zwei wichtige Aspekte, die in diesem Zusammenhang berücksichtigt werden müssen, sind:

- der Bildinhalt und
- die Darstellungseigenschaften.

Die Zielsetzung der Bildverarbeitung besteht darin, spezifische Aspekte des Bildinhaltes zu bestimmen. Um diesen Bildinhalt extrahieren zu können, müssen die angewandten Bildverarbeitungsalgorithmen

die Darstellungseigenschaften kennen, um aus den Intensitätsverläufen des Bildes eine abstrakte Interpretation ableiten zu können.

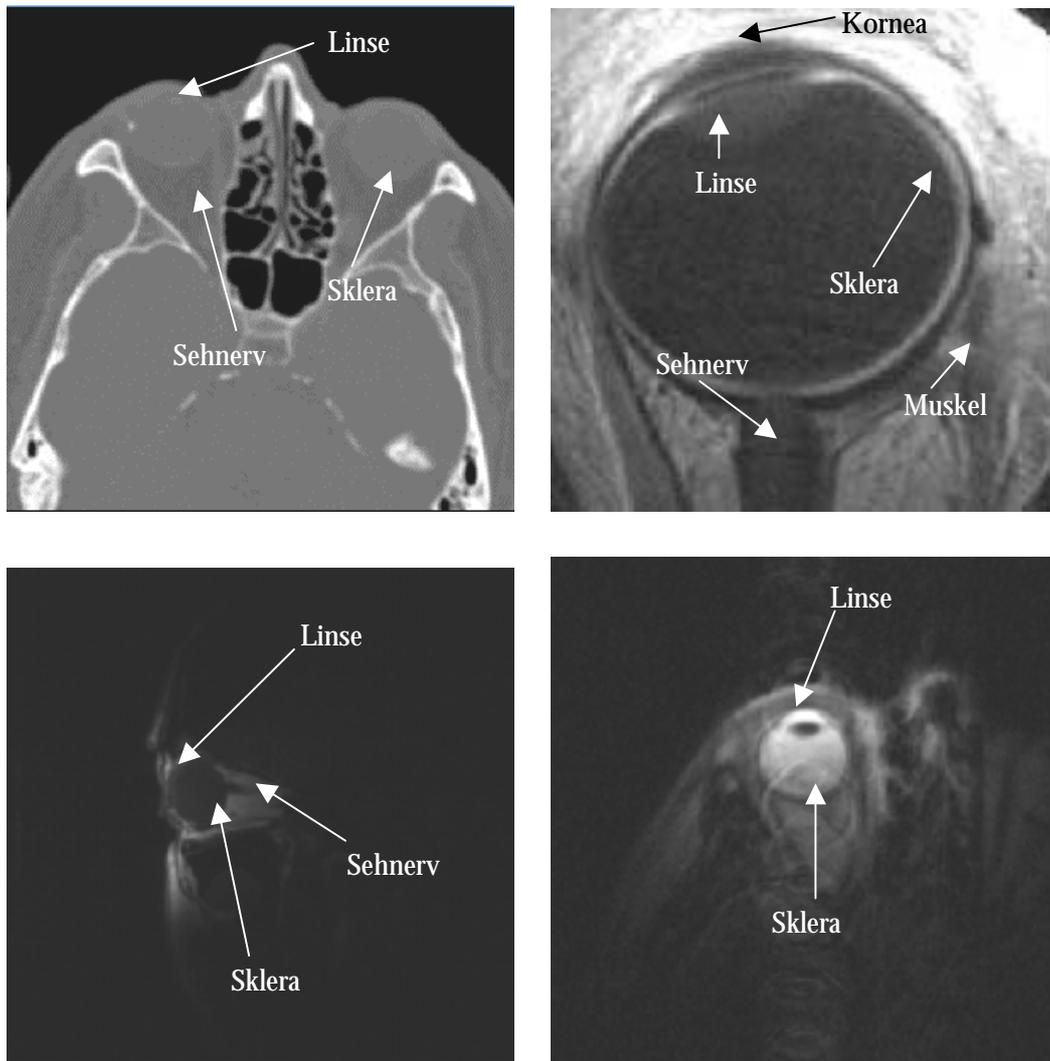


Abbildung 3.3: Veranschaulichung des Bildinhalts

Die Bilder zeigen zwar immer den gleichen Gegenstand, das Auge. Die Darstellung der enthaltenen Organe Bulbus, Sklera, Linse, Sehnerv und Muskel ist jedoch von Datensatz zu Datensatz grundsätzlich verschieden. Ihr Erscheinungsbild wird durch die Aufnahmeparameter bestimmt.

3.2.1 Bildgegenstand und Bildinhalt

Der Bildgegenstand ist das Organ, das in dem Bildmaterial gezielt untersucht werden soll. Abbildung 3.3 zeigt verschiedene Bilddaten, die jeweils das Auge als Bildgegenstand beinhalten. Es sind die typischen Aufnahmearten, die für die Therapie von Augentumoren aufgenommen werden.

Der Bildinhalt beschreibt, was das Bild darstellt. Während die CT-Aufnahmen zusätzlich zum Auge noch den Schädel abbilden, sind auf den MRT-Aufnahmen nur das Auge und das umgebende Fett- und Muskelgewebe zu sehen.

Unterschiede im Bildinhalt äußern sich darin, daß unterschiedliche Organe, verschiedene Ausschnitte oder verschiedene Schnitorientierungen eines Organs dargestellt werden. Auch die Wahl unterschiedlicher Aufnahmesequenzen bei Kernspinaufnahmen führt zu unterschiedlichen Bildinhalten, da sich bei-

spielsweise in T1-Aufnahmen ganz andere Organe durch andere Grauwerte voneinander abgrenzen lassen, als es bei den T2-Aufnahmen der Fall ist.

3.2.2 Darstellungseigenschaften

Die Darstellungseigenschaften äußern sich vor allem in den Grauwert- und Textureigenschaften, durch die sich Objekte von ihrer Umgebung unterscheiden. Die Variationen der Bildeigenschaften führen zu deutlichen Änderungen in den Grauwerteigenschaften der einzelnen Bildelemente. Obgleich die Bilder in Abbildung 3.4 in ihrem Bildinhalt übereinstimmen, weisen sie sehr unterschiedliche Darstellungseigenschaften auf. Die Unterschiede bestehen vor allem in dem Grauwertkontrast, dem Rauschniveau und Texturen bei den einzelnen anatomischen Strukturen wie Fettgewebe oder Glaskörper. Einen starken Einfluß auf die Darstellungseigenschaften haben auch Artefakte wie beispielsweise die Bewegungsartefakte im rechten Bild. Alle Variationen des dargestellten Auges machen klar, daß ein Verfahren, das den Bulbus segmentieren soll, nicht allein auf den Grauwerten aufsetzen kann.

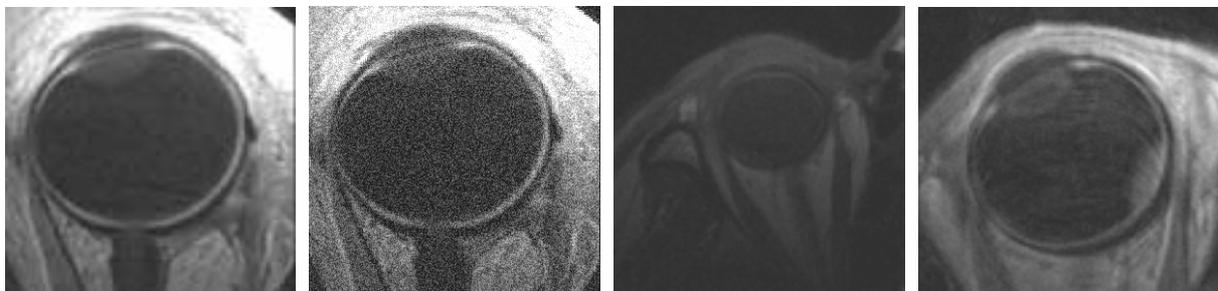


Abbildung 3.4: Variationen der Darstellungseigenschaften

Trotz der Verwendung des gleichen Bildmodus (axiale T2-MRT-Aufnahmen in einer zentralen Schicht des Auges) sind die Darstellungseigenschaften grundsätzlich unterschiedlich.

3.3 Anatomie und Physiologie des Auges

Die vorangegangenen Abschnitte haben gezeigt, wie komplex die Problemstellungen sind, die bei einer Interpretation von visuellen Informationen berücksichtigt werden müssen. Ein möglicher Ansatz, die Problemstellungen zu verstehen und mögliche Lösungskonzepte für die Bildverarbeitung zu entwerfen, besteht in dem bionischen Ansatz, von der Natur zu lernen. Aus diesem Grund soll zunächst betrachtet werden, wie das Problem des "Sehens" beim Menschen gelöst wurde.

Die wichtigste Rolle spielt beim Menschen hierbei das Sehorgan, zu dem die beiden Augen, die Sehbahnen und das Sehzentrum (Visueller Cortex) gehören. Der grundsätzliche Aufbau des Sehorgans ist in Abbildung 3.5 dargestellt.

Die Augen übernehmen in erster Linie den rezeptiven Teil bei der Wahrnehmung visueller Informationen. Die Lichtphotonen werden durch den Linsenapparat gebündelt und auf die Retina fokussiert. Dort wird das einfallende Licht durch lichtempfindliche Neuronen, den Stäbchen und Zäpfchen, registriert. Über ein komplexes, mehrschichtiges System von verschiedenen Nervenzellen werden die Signalimpulse der Stäbchen und Zäpfchen an die Axone der Ganglienzellen weitergeleitet.

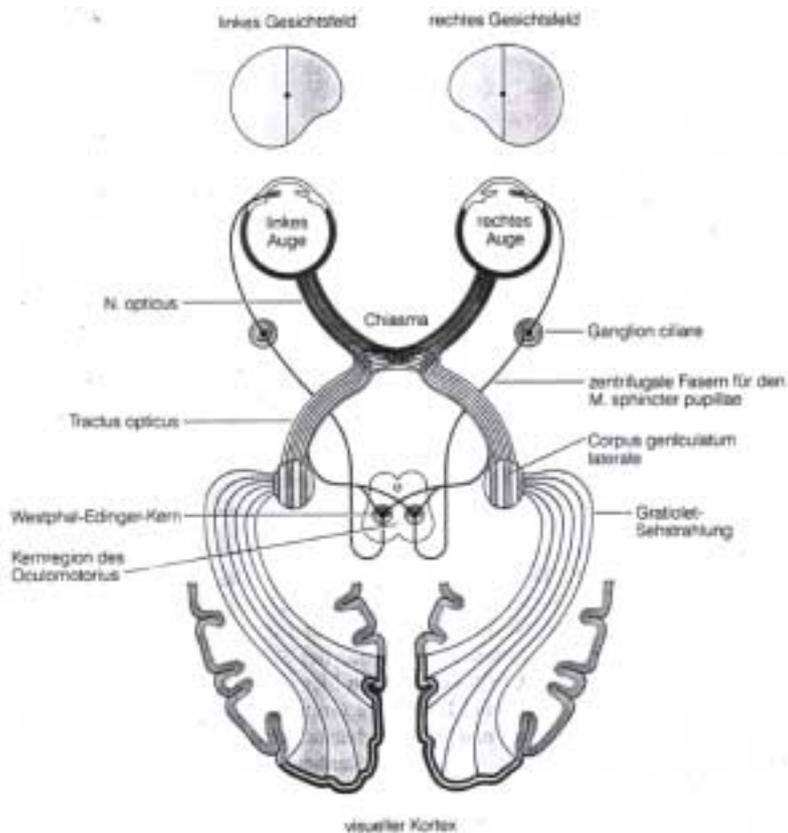


Abbildung 3.5: Strukturen des menschlichen Sehorgans mit seinen Sehbahnen

Die Abbildung wurde entnommen aus [Ley93].

Diese Axone der retinalen Ganglienzellen vereinigen sich zu den Sehnerven, über die die visuellen Informationen über den Chiasma opticum an den Visuellen Cortex weitergeleitet werden. Im Chiasma opticum vereinigen sich die Sehnerven beider Augen. Anschließend werden die Reize über den Thalamus (Corpus geniculatum laterale) und die Gratiolet-Sehstrahlung zur Hinterhaupttrinde weitergeleitet. Von dem primären Sehzentrum gehen die Signale in übergeordneten Bahnen zu anderen Rindengebieten für die Weiterverarbeitung der optischen Eindrücke. Dort findet eine Einordnung der verarbeiteten Signale in den Gesamtkomplex der Erfahrungen statt (siehe [Ley93]).

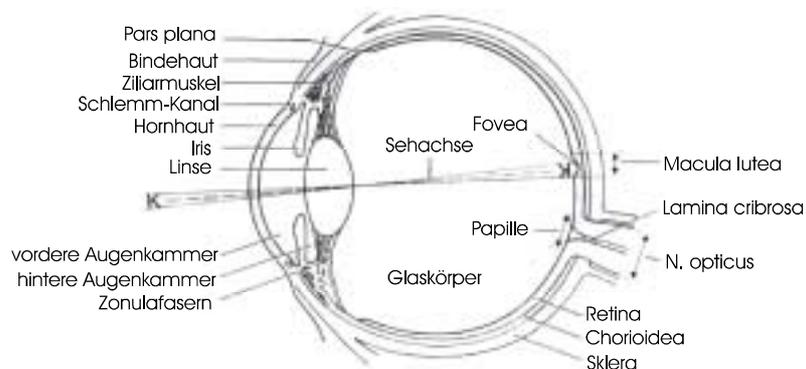


Abbildung 3.6: Waagerechter schematischer Schnitt durch den rechten Augapfel

Die aus [Ley93] entnommene Grafik zeigt die wichtigsten anatomischen Bestandteile des Auges.

Abbildung 3.6 zeigt die grundsätzliche Anatomie des menschlichen Auges. Der Augapfel ist eine sphärische Hülle aus festem Bindegewebe und besteht aus der durchsichtigen Hornhaut (Kornea) und der weißen Lederhaut (Sklera). Die Linse dient der Entfernungseinstellung (Akkommodation). Sie ist ein kristallklarer Körper, dessen Brechungsindex größer ist als der des Kammerwassers und des Glaskörpers. Durch Kontraktion und Relaxation des Ziliarmuskels kann der Zug der Zonularfasern auf die Linse übertragen und somit deren Wölbung modifiziert werden, um die Brechkraft so anzupassen, daß der Brennpunkt auf der Netzhaut (Retina) liegt. Die Linse wird durch keine Blutgefäße oder Nerven durchzogen. Sie wird durch das Kammerwasser ernährt. Die Aderhaut hält die Temperatur des Auges konstant und ernährt die äußeren Schichten der Netzhaut. Die Netzhaut ist das bildaufnehmende Organ und setzt sich aus den Stäbchen und Zapfen zusammen. Der Glaskörper füllt den Raum hinter der Linse und besteht aus einem feinen Fasergerüst und einem darin eingebetteten Hydrogel. Der Brechungsindex des Glaskörpers entspricht ungefähr dem des Kammerwassers. Die Netzhaut stellt eine Fortsetzung des Gehirns dar. Sie umfaßt ungefähr 127 Millionen Rezeptoren, die Lichtreize aufnehmen und ordnen. Die Axone der Ganglienzellen laufen in der Sehnervscheibe (Papilla nervi optici) zum Sehnerv zusammen. Die Regenbogenhaut (Iris), der Strahlenkörper (Corpus ciliare) und die Aderhaut (Chorioidea) bilden den Tractus uvealis, kurz Uvea.

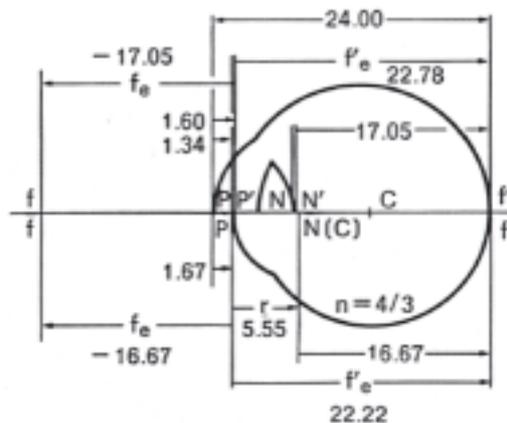


Abbildung 3.7: Das Gullstrandsche Normalauge

Die Abbildung zeigt die statistisch ermittelten Größenmaße des menschlichen Auges. Die Werte der oberen Hälfte zeigen die Größenmaße des erweiterten Gullstrand Modell, die Werte der unteren Hälfte die des reduzierten Gullstrandschen Normalauges (Grafik stammt aus [Mic85]).

Das Auge ist ein vergleichsweise kleines Organ. Der mittlere Durchmesser liegt bei ca. 24 Millimetern, obgleich diese Ausmaße interindividuell schwanken. Die relativen Proportionen der einzelnen Teilstrukturen des Auges wurden gegen 1920 statistisch untersucht und in dem Gullstrandschen Normalauge zusammengefaßt (siehe Abbildung 3.7). Dieses Modell wurde sehr lange als Standard in der Ophthalmologie akzeptiert. Durch die Möglichkeiten moderner bildgebender Verfahren konnte jedoch festgestellt werden, daß es gerade bei starker Kurz- oder Weitsichtigkeit zu beträchtlichen Abweichungen von der Kugel-form dieses Modells kommt.

Der erste neuronale Verarbeitungsschritt der visuellen Informationen findet in den Fotorezeptoren statt. Trifft ein Lichtquant auf einen Photorezeptor, wird durch die Umsetzung des Pigmentfarbstoffs Rhodopsin ein Aktionsimpuls ausgelöst. Die Fotorezeptoren werden in die funktional unterschiedlichen Zapfen und Stäbchen unterteilt. Die Zapfen dienen der Farberkennung, die Stäbchen der Kontrasterkennung. Die Netzhaut beinhaltet ungefähr 120 Millionen Stäbchen, die sich alle außerhalb der Fovea befinden, und 6 Millionen Zapfen, die in erster Linie innerhalb der Fovea liegen. Die Zapfen unterteilen sich in

drei Gruppen von Photorezeptoren mit jeweils unterschiedlichen Sensitivitätsmaxima, die für das Erkennen der Farben blau, grün und rot verantwortlich sind [Sha93].

Die Fovea ist der für die Aufnahme visueller Informationen wichtigste Teil der Netzhaut. Hier findet sich die höchste Konzentration von Zapfen und somit die höchste Sehschärfe des Menschen. Die Motorik des Auges ist so ausgerichtet, daß die gerade beobachteten Zielobjekte immer mit den Photorezeptoren der Fovea aufgenommen werden. Die Peripherie der Netzhaut weist eine geringere Dichte an Photorezeptoren auf und dient in erster Linie der unbewußten Erkennung der Umgebung.

Die erste Auswertung der Signale von den Photorezeptoren wird nicht erst im Visuellen Cortex, sondern bereits in der Retina vorgenommen. In der Retina befindet sich neben den Photorezeptoren ein dichtes Netz von Bipolar-, Horizontal-, Amacrine- und Ganglienzellen, die die angrenzenden Photorezeptoren in einem komplexen Geflecht miteinander verbinden. Durch sie werden physiologische Phänomene wie laterale Inhibition, rezeptive Felder (On-, OFF-Zentrum) und die Detektion von Kanten verursacht [Sha93]. Eine weitere Verarbeitungsinstanz besteht in dem Zusammenfassen der Signale der 120 Millionen Photorezeptoren auf 1 Million Nervenfasern im Sehnerv.

Eine weitergehende Beschreibung der Anatomie und Physiologie des Auges und des Sehorgans liegt außerhalb des Rahmens dieser Arbeit und kann in der ophthalmologischen Standardliteratur nachgelesen werden [Ley93, Sha93]. Die skizzierten Beschreibungen verdeutlichen jedoch bereits, daß bereits in der Netzhaut wichtige Verarbeitungsschritte für die Aufnahme und Erkennung von visuellen Informationen stattfinden. Eine Therapie, die die Sehfähigkeit des Patienten erhalten will, muß daher beispielsweise eine Bestrahlung der Fovea so weit wie möglich vermeiden, da sie funktional für das menschliche Sehvermögen unverzichtbar ist. Aber auch andere essentielle Strukturen innerhalb des Auges sollten so weit wie möglich geschützt werden, damit die visuelle Verarbeitungspipeline nicht unterbrochen wird. Essentielle Strukturen stellen in diesem Zusammenhang vor allem die Linse, die Fovea und die Papille dar.

3.4 Visuelle Perzeption

Die Natur hat beim Menschen und vielen Tieren ein ausgefeiltes System entwickelt, um visuelle Informationen aufzunehmen und inhaltlich auszuwerten. Nur ein sehr kleiner Teil des Bilderkennungsprozesses beim Menschen kann durch die Anatomie und Physiologie des Sehapparats erklärt werden. Um weitergehende Erkenntnisse zu gewinnen, versucht sich der Forschungsbereich "Visuelle Perzeption" auf kognitiver Ebene der Funktionsweise des menschlichen Sehapparats zu nähern.

Im Unterschied zu Insekten oder Kaltblütern findet die Verarbeitung von visuellen Informationen beim Menschen nicht allein im Auge statt, sondern ist auf unterschiedliche anatomische Strukturen verteilt [Sha93]. Für die Verarbeitung der visuellen Information sind vor allem die folgenden Aspekte der Bilder wichtig: die Kontrastempfindlichkeit (Unterscheiden unterschiedlicher Grautöne), die Sehschärfe (örtliche Auflösungsvermögen) und die Gewichtung von Inhalten. Der letzte Punkt äußert sich beim Auge durch die unterschiedliche Auflösung innerhalb der Retina (zwischen Fovea und Peripherie), in der Motorik und dem Fokussieren des Auges [Sha93].

Weitere Erkenntnisse lassen sich aus der Sequenz der Wahrnehmungsschritte ableiten. In den ersten Zehntelsekunden des Betrachtens eines Bildes spielt die Peripherie der Netzhaut die entscheidende Rolle. In dieser ersten, weitgehend unbewußten Erkennungsphase findet die sogenannte "pre-attentive evaluation" statt. In dieser Phase erkennt der Mensch den grundsätzlichen Kontext des Bildes, der die Grundlage darstellt für eine weitergehende, intellektuelle Verarbeitung der visuellen Informationen. Dieser Schritt, der aufgrund des evolutionären Überlebenskampfes des Menschen innerhalb der freien Natur sehr schnell ablaufen mußte, zeichnet sich durch eine parallelisierte Verarbeitung der Impulse aus [Gal93]. Nach diesem ersten Erkennen des Kontextes schließt sich die teilweise bewußte "attentive evaluation" an. In dieser

Phase werden in erster Linie die Informationen von der Fovea ausgewertet. Durch das gezielte Betrachten einzelner Bildelemente wird die Kontext-Hypothese des ersten parallelen Bildverarbeitungsschritts verfeinert und in ihren Details bestätigt oder verworfen. Diese bewußte und selektive Aufnahme von Bildinformationen zeichnet sich durch eine weitgehend serielle Bildauswertung aus. Diese Tatsache verdeutlicht, daß der Mensch Verarbeitungsschritte auf hoher intellektueller Ebene seriell bearbeitet.

Sowohl das Auge als auch der visuelle Kortex besitzen die Fähigkeit zum hervorheben und unterdrücken verschiedener Teile des aufgenommenen Bildes. Diesen Vorgang bezeichnet man als "selektive Filterung". So werden Kanten nicht nur in der Netzhaut (laterale Inhibition), sondern auch im visuellen Kortex besonders stark wahrgenommen. Für den Beleg dieser Tatsache erhielten Hubel und Wiesel 1981 den Nobelpreis [Lin96]. Dabei werden vertikale Linien deutlich stärker detektiert als horizontale oder Linien eines beliebigen anderen Winkels.

Darüber hinaus laufen beim Menschen Interpretationsautomatismen ab, die selbst durch bewußte Überlegungen kaum beeinflusst werden können. Ein Beispiel für einen solchen Automatismus ist die Mondillusion: Der Mond scheint kurz über dem Horizont sehr viel größer zu sein, als er es ist, wenn er hoch oben am Sternenhimmel steht. Die Erklärung, daß dies durch die Linsenfunktion der Atmosphäre verursacht wird, stimmt nicht, wie leicht durch Fotografien nachgewiesen werden kann. Schaut man sich den Mond durch ein aus Papier gerolltes Fernrohr an, verschwindet dieser visuelle Effekt. Grund für diesen Effekt ist das automatische In-Beziehung-Setzen des Mondes mit dem Horizont. Objekte die weiter entfernt sind, sind in ihrer Projektion deutlich kleiner als Objekte, die sehr nah sind. Das Auge kompensiert diesen visuellen Eindruck automatisch. Der Sinn solcher Automatismen ist die Beschleunigung der visuellen Interpretationsschritte und offensichtlich hat die Natur solcherlei Fehlinterpretationen in Kauf genommen, da sie in der natürlichen Umgebung nur selten vorkamen und nicht von lebenswichtiger Bedeutung waren.

Ein weiterer Ansatz zur Beschleunigung von visuellen Verarbeitungsschritten wird beim Menschen erreicht, indem spezifische Reize wie Farben, Muster, Kantenformen oder Kontraste durch spezifische Gehirnbereiche ausgewertet werden. Bei Affen konnten in dem Sulcus temporalis inferior, einem Schläfenlappen des Gehirns, Bereiche festgestellt werden, die allein für das Erkennen von Gesichtern zuständig sind. Ähnlich verhält es sich bei der Ausführung von spezifischen Bewegungen, die in leichten Modifikationen immer wieder auftreten. So existieren hierfür in der motorischen Großhirnrinde Bereiche, in denen das Gehirn Grundmuster für Befehle wie beispielsweise von Armbewegungen speichert. Für die individuellen Handlungen sind diese Befehlsmuster, die im allgemeinen nur aus wenigen hundert Neuronen bestehen, nur noch an den aktuellen Kontext anzupassen. Ein weiterer automatischer visueller Mechanismus übernimmt das schnelle Gruppieren von Mustern [Hen93].

Solche automatisierten Verarbeitungsmechanismen können auf unterschiedlichen intellektuellen Ebenen stattfinden. Bei geübten Korrekturlesern kann beispielsweise beobachtet werden, daß sie Texte deutlich schneller nach Fehlern durchsuchen können, als es theoretisch bei einem bewußten und seriellen Verarbeiten der Informationen möglich wäre [Hen93, Kel93]. Dieses Beispiel deutet unterschiedliche Instanzen bei der Auswertung von Bildmaterial an, die oft unbewußt und sehr schnell durchgeführt werden. Solche beschleunigten Erkennungsmechanismen sind bekannt für das Erkennen von Kontrasten, Kanten, Ecken, Balken und Bewegungsrichtungen.

Nur sehr wenig verstanden sind die Verarbeitungsschritte, die auf einer höheren Abstraktionsebene stattfinden und eine intelligente Auswertung der Informationen erforderlich machen. Hierzu gehören vor allem das Erfassen komplexer Farbverläufe, inhaltlicher Strukturen und die Fähigkeit, die zweidimensionalen Informationen ohne Verzögerungen in eine dreidimensionale Interpretation und Rekonstruktion der Umwelt zu überführen. Um die einzelnen Verarbeitungsschritte zur Interpretation visueller Informationen

besser zu verstehen, wurden verschiedene Modelle entwickelt, die versuchen, das Bilderkennungsprinzip in einzelnen Verarbeitungseinheiten zu formalisieren.

3.5 Ansätze zur Beschreibung des Bilderkennungsprinzips

Es wurde in vielen wissenschaftlichen Bereichen versucht, Erklärungen für die Funktionsweise des visuellen Wahrnehmungssystems des Menschen zu finden. Es entstanden in den Bereichen Philosophie, künstliche Intelligenz und Medizin verschiedene Theorien, die zunächst versuchten, den kognitiven Prozeß des Erkennens zu beschreiben. Auf diese Weise sollte das Gesamtproblem in Einzelprobleme auf verschiedenen Abstraktionsebenen untergliedert werden. Der Sinn einer Formalisierung des Bilderkennungsprinzips besteht einerseits im besseren Verständnis der Vorgänge, andererseits im Schaffen einer theoretischen Grundlage für die Realisierung eines korrespondierenden, computergestützten Bilderkennungsverfahrens.

Ein bedeutender Ansatz zur Beschreibung des Bilderkennungsprinzips wurde von Marr, einem der einflußreichsten Psychologen, der sich in diesem Forschungsbereich etablierte, 1982 entwickelt. Sein Ansatz ging von drei Verarbeitungsschritten aus (vergleiche Abbildung 3.8), in denen jeweils unterschiedliche Eigenschaften des Bildes analysiert werden [Mar82, Hen93]:

1. Aufnehmen eines primären Eindrucks vom Bild, der allgemeinen Struktur und Intensitätseigenschaften,
2. Bestimmen wichtiger Details wie Unregelmäßigkeiten in Oberflächen und Tiefeninformationen und
3. Mentale Rekonstruktion des 3D-Modells der Gesamtszene und der einzelnen Objekte.

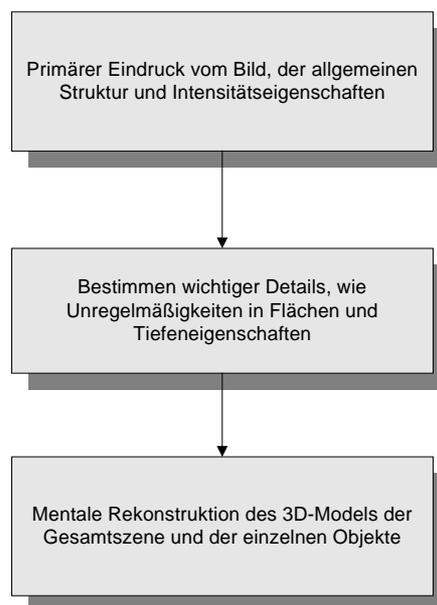


Abbildung 3.8: Darstellung des Bilderkennungsprinzips nach Marr [Mar82]

Das Bilderkennungsprinzip nach Marr besteht aus den drei Erkennungsschritten: primärer Eindruck, Detailinterpretation und dreidimensionale Rekonstruktion.

Wenngleich der von Marr entwickelte Ansatz bereits die Unterteilung zwischen dem Erkennen des allgemeinen Kontextes, des speziellen Kontextes und anschließend dem Rekonstruieren der dreidimensionalen Situation vornimmt, ignoriert dieser Ansatz einige Aspekte des menschlichen Sehverhaltens. Besonders die Untersuchungen von Gregory [Gre74] bezüglich der Wahrnehmung von optischen Täuschungen

haben gezeigt, daß die Bilderkennung ein rekursiver Prozeß ist, bei dem die Informationen weitgehend seriell verarbeitet werden und zu sich immer wieder ändernden Hypothesen über den Bildinhalt führen. Weitere Untersuchungen haben gezeigt, daß die Interpretationen des Bildes auch durch Vorwissen oder andere individuelle Faktoren beeinflußt werden. Diese Untersuchungen belegen, daß trotz der zuvor erwähnten Automatismen auch Interpretationsmechanismen existieren, die unabhängig von physiologischen Vorgaben sind.

Anschaulich werden diese Erkenntnisse beim Betrachten von Sinnestäuschungen wie beispielsweise den Bildern von Escher in Abbildung 3.9. Der Betrachter stellt jeweils Hypothesen auf, über die dreidimensionale Situation der beiden Konstrukte. Bei einer genaueren Untersuchung der Treppen oder des Wasserlaufes stellt er jedoch fest, daß die erste Hypothese falsch war und stellt eine modifizierte Hypothese auf. Verfolgt er die Leiter oder die Wasserläufe weiter, stellt er fest, daß auch die zweite Hypothese falsch war, und geht wieder zu der ersten Hypothese über. In dieser Weise wechseln sich beide Hypothesen immer wieder ab, ohne daß es eine endgültige Erkenntnis über die tatsächliche dreidimensionale Struktur der dargestellten Bauten gibt. Sequentielles und iteratives Analyseverhalten des Menschen führen dazu, daß das Konzept dieser Sinnestäuschungen funktioniert und der Mensch zwischen den zwei Hypothesen hin und her wechselt. Es gibt eine Vielzahl von weiteren Sinnestäuschungen, die einen Einblick über die Verarbeitung von visuellen Informationen durch den Menschen gewähren.

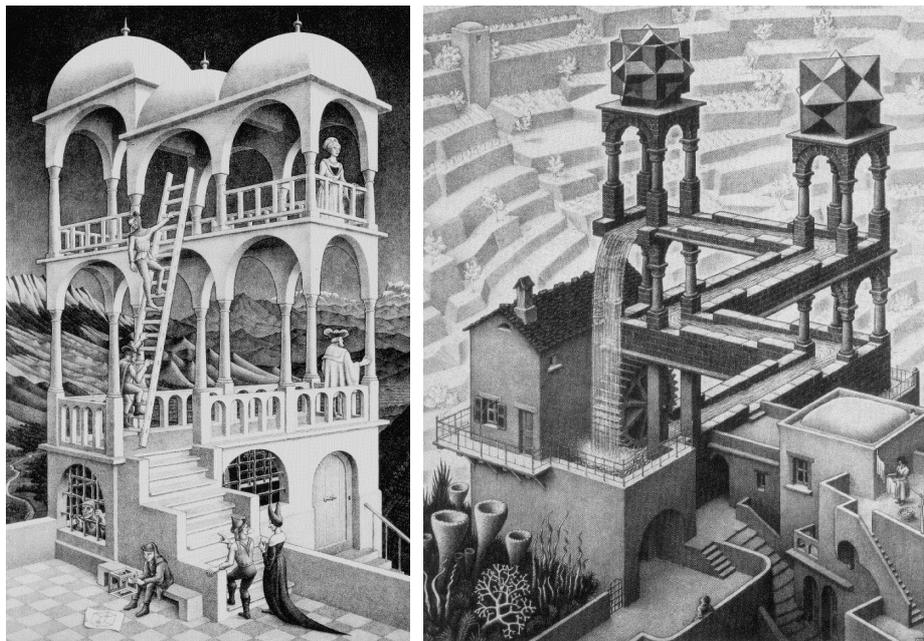


Abbildung 3.9: Von MC Escher angefertigte Bilder mit optischen Täuschungen

Die optischen Täuschungen in den Bildern von Escher veranschaulichen das iterative und serielle Verarbeiten der Bildinformationen.

Zwei weitere Ansätze zur Beschreibung des menschlichen Bilderkennungsverfahrens wurden im Bereich der Künstlichen Intelligenz entwickelt und werden als „bottom-up“ und „top-down“ Ansatz bezeichnet. Dabei versuchen die beiden Ansätze den kognitiven Prozeß des abstrakten Erkennens einer Bildsituation zu beschreiben. Bei dem „bottom-up“ Ansatz werden die einzelnen Elemente des Bildes separat erkannt und dann zu einer Gesamtinterpretation zusammengefügt. Die Verarbeitungsschritte werden weitgehend parallel durchgeführt und durch selektive Filterung unterstützt. Dem entgegen steht jedoch die Tatsache, daß selbst für sehr komplexe Szenen innerhalb von wenigen Zehntelsekunden eine Hypothese für den Bildinhalt bereit steht und erst anschließend die einzelnen Bildelemente genauer unter-

sucht werden [Win84]. Bei dem „top-down“ Ansatz gewinnt der Betrachter zunächst einen Eindruck von der Gesamtsituation, um anschließend die Interpretationen für die Einzelobjekte des Bildes schrittweise zu verfeinern, indem durch detailliertes Betrachten das Wissen schrittweise vermehrt wird.

Bei dem Ansatz nach Alastair G. Gale werden die einzelnen Aspekte des Bilderkennungsprinzips in den Zusammenhang mit dem Vorwissen des Beobachters gebracht und der iterative Charakter der Global- und Lokal-Hypothesen berücksichtigt. Darüber hinaus wird eine klare Unterscheidung zwischen „pre-attentive“ und „attentive stage“ vorgenommen (siehe Abbildung 3.10):

- Bei dem „pre-attentive stage“ oder „distributed stage“ wird eine parallele Verarbeitung der gesamten Informationen der Netzhaut vorgenommen („first glimpse“ innerhalb von 20-200 ms). Dieser erste Schritt führt zu mehreren Stimuli, die in die zweite Phase überleiten. Es werden vor allem Helligkeiten, Größen, Neigungen, Kanten und geschlossene Formen gefunden.
- Die zweite Phase wird „focal attention“ genannt. Sie wird durch das Ausführen von seriellen Suchbewegungen und das Verarbeiten der daraus resultierenden Informationen charakterisiert. Die einzelnen Details der Szene werden nicht gezielt und chronologisch, sondern in ungeordneter Form durch die voluntären Bewegungen des Auges über das Bild aufgenommen. Dabei kommt es häufig dazu, daß einzelne Details mehrfach betrachtet werden. Der Bildausschnitt (Field of View) wird bei diesem Schritt oft variiert. Da das Auge nur geringe Möglichkeiten zum Zoomen besitzt, die allein auf kognitiven Prozessen aufbauen, tritt beispielsweise der Betrachter eines Bildes im Museum oft an das Bild heran oder davon weg, um Details des Bildes aufzunehmen. Ein solches Sehverhalten kann ebenfalls bei Ärzten bei der Auswertung von Röntgenbildern festgestellt werden [Hen93, Gal93].

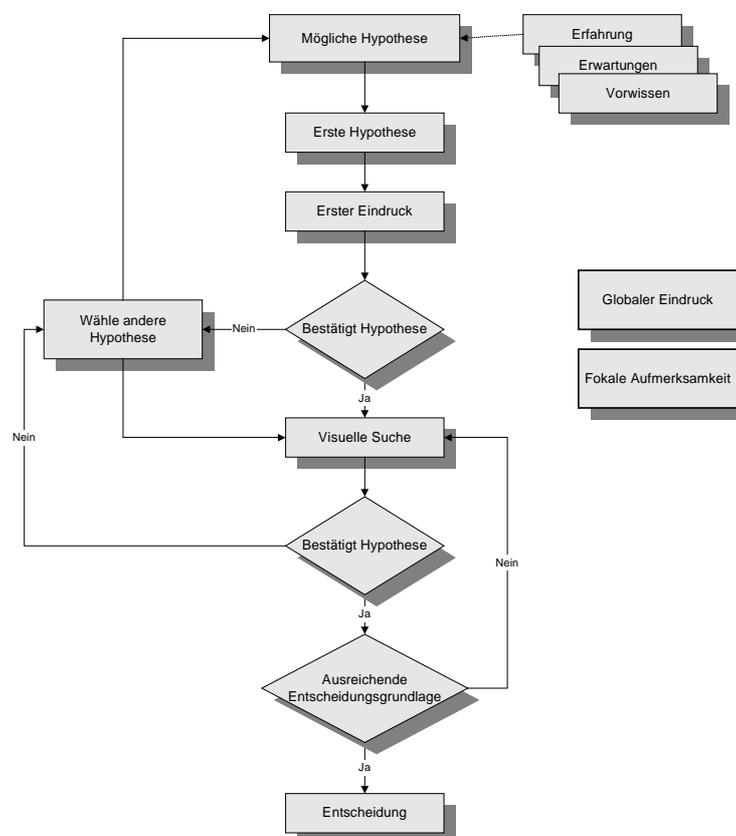


Abbildung 3.10: Ansatz zur Beschreibung der Bildverarbeitungspipeline nach Gale [Gal93]

Bei dem Bilderkennungsprinzip nach Gale werden iteratives und sequentielles Verarbeiten visuellen Informationen berücksichtigt.

Der Bilderkennungsansatz nach Gale weist jedoch immer noch das Problem auf, daß eine hierarchische Hypothesenstruktur, wie sie durch die Künstliche Intelligenz angenommen wird, nicht berücksichtigt wird. Darüber hinaus wird auch das Verhältnis von bewußtem zu unbewußtem Erkennen nicht geklärt, welches bei dem Beispiel des Korrekturlesers von oben durchaus eine Rolle spielt. Die Frage, in welchem Verhältnis das zweidimensionale und dreidimensionale Erkennen steht, bleibt offen.

3.6 Wissensbasierte Bildverarbeitung

Die im vorangegangenen Abschnitt erwähnten Ansätze zur Beschreibung des Bilderkennungsprinzips erlauben die Rekapitulation einiger grundsätzlicher Aspekte des Sehens. Sie bieten jedoch noch keinen Ansatzpunkt, um den komplexen Erkennungsprozeß in einem Algorithmus auf den Computer zu übertragen. In diesem Abschnitt sollen Kriterien entwickelt werden, die ein Bildverarbeitungsansatz erfüllen müßte, um die gesetzten Ziele und Aufgabenstellungen einer automatisierten Bildverarbeitung zu erfüllen.

Die wissenschaftlichen Arbeiten in den Bereichen der künstlichen Intelligenz und der medizinischen Bildverarbeitung versuchen seit geraumer Zeit, Konzepte für die automatisierte Bildverarbeitung zu entwerfen. Die bisher erzielten Ergebnisse können noch nicht als Durchbruch zur Erreichung dieser Zielsetzungen angesehen werden. Vielmehr ist weitgehend die Einsicht eingekehrt, daß die intellektuellen Fähigkeiten des Menschen zur Verarbeitung von visuellen Informationen so komplex sind, daß sie in absehbarer Zeit nicht durch intelligente Computerprogramme erreicht werden können. Nichtsdestotrotz wurden eine Vielzahl von Ansätzen entwickelt, die sich der Lösung der Problemstellungen annähern.

Um die Komplexität der Problemstellung zu reduzieren und sich so den Aufgabenstellungen der Bildverarbeitung anzunähern, wurden in den letzten Jahrzehnten verschiedenste Bildverarbeitungsalgorithmen entwickelt, die für spezifisches Bildmaterial konzipiert wurden und so von konkreten Eigenschaften des Bildmaterials ausgehen und gezielt nach erwarteten Bildinhalten suchen konnten. Auf diese Weise wurde die Komplexität der zu entwickelnden Algorithmen so weit reduziert, daß zumindest für das angenommene Bildmaterial die gesetzten Aufgabenstellungen innerhalb eines begrenzten Zeithorizonts realisiert werden konnten. Bei dem Verständnis aktueller Bildverarbeitungsalgorithmen spielen die Eigenschaften des zugrunde gelegten Bildmaterials daher eine entscheidende Rolle. Aus diesem Grund sollen in den folgenden Abschnitten das Bildmaterial und die jeweils relevanten Eigenschaften genauer untersucht werden.

Medizinisches Bildmaterial erweist sich unter den Aspekten der Auswertung und Interpretation als außerordentlich komplex. Dies veranschaulichten einerseits die Betrachtungen der Abschnitte 3.1 und 3.2. Andererseits stellt die Auswertung von medizinischem Bildmaterial selbst für den Menschen ein so komplexes Problem dar, daß selbst nach Jahren der radiologischen Ausbildung die Interpretation von medizinischem Bildmaterial interindividuellen Schwankungen unterliegt.

Die Interpretation von medizinischem Bildmaterial erfordert Vorwissen über Bildgebungsprozeß und medizinischen Kontext. Das Wissen über den Bildgebungsprozeß ist wichtig, damit der Betrachter versteht, wie das Bildmaterial in den medizinischen Kontext einzuordnen ist. Gerade bei MRT-Aufnahmen ist die Kenntnis der Aufnahmeparameter wie beispielsweise Aufnahmegerät, Aufnahmesequenz, Schichtpositionen und Darstellungseigenschaften von Artefakten wichtig, um eine Korrelation des Betrachteten mit anatomischem Vorwissen herzustellen. Die medizinischen Vorkenntnisse dürfen andererseits nicht auf anatomische Kenntnisse beschränkt bleiben. Eine genaue Kenntnis der Variabilität der Organe, Veränderungen der Darstellungseigenschaften des Gewebes im Krankheitsfall und der Darstellungsmöglichkeiten der jeweiligen Bildmodi sind unverzichtbar, um das in der klinischen Routine anfallende Bildmaterial auswerten zu können. Neben dem faktischen Wissen über Bildgebungsprozeß und medizinischen Kontext muß sich die medizinische Fachkraft außerdem in Jahren der Ausbildung algorithmisches (prozedurales) Wissen aneignen, um die beiden Komplexe faktischen Wissens miteinander zu korrelieren.

Im Bereich der Künstlichen Intelligenz wurden verschiedenste Expertensysteme und Wissensbasierte Systeme zur Kodierung von faktischem und algorithmischem Wissen entwickelt. Expertensysteme sind Programme, die Lösungen zu komplexen Problemstellungen in vorgegebenen Domänen suchen, indem das System intelligente Entscheidungs- und Überlegungsmechanismen von Experten zu simulieren sucht. Diese Simulationen setzen auf Wissen über die Domäne und Verfahren zur Anwendung dieses Wissens auf. Hierzu muß das Wissen zunächst formalisiert und repräsentiert werden. Anschließend müssen Verfahren zur Manipulation des Wissens für das jeweils angewandte Problemlösungskonzept umgesetzt werden [Pup93]. Die Expertensysteme unterscheiden sich von den Wissensbasierten Systemen lediglich in der Tatsache, daß bei Expertensystemen das Wissen durch Experten beigetragen wird [Pup93].

Nachdem in der Anfangszeit der Expertensysteme die Problemlösemechanismen im Vordergrund standen, verschob sich in den 70er Jahren der Fokus des wissenschaftlichen Interesses auf die Repräsentation des Wissens: „*The fundamental problem of understanding intelligence is not the identification of a few powerful techniques, but rather the question of how to represent large amounts of knowledge in a fashion that permits their effective use and interaction....*“ [Dud83]. Während der Mensch das erworbene Wissen durch die Verschaltung der Neuronen in Gehirn und zentralem Nervensystem speichert, muß für computergestützte, wissensbasierte Systeme eine symbolische Sprache für die Wissensrepräsentation geschaffen werden. Man spricht in diesem Zusammenhang von der Modellierung von Wissen [Sch91].

Entsprechend der Erkenntnis, daß einer der wichtigsten Faktoren eines wissensbasierten Systems die Wissensrepräsentation ist, war das erste weltweit akzeptierte Expertensystem das von der Stanford Universität entwickelte medizinische Expertensystem MYCIN, das die Diagnose und Therapie von bakteriologischen Erkrankungen des Blutes und Meningitis unterstützt. Im Unterschied zum Großteil der anderen Entwicklungen dieser Zeit bestand eines der wichtigsten Kriterien beim Entwurf dieses Programms in der klaren Trennung von Wissensrepräsentation und Problemlösungsmethoden. Das nachfolgende Expertensystem INTERNIST basierte auf dem gleichen Grundsatzkonzept wie MYCIN. Es erlaubt die Diagnose von ca. 75% der internistischen Erkrankungen [Pup93, S. 8]. Die Erfolge dieser Expertensysteme bestätigten den Ansatz, daß das zentrale Problem bei dem Entwurf eines wissensbasierten Systems die Entwicklung einer umfassenden Wissensrepräsentation und deren Übertragung auf den Problemlösungsmechanismus darstellt.

In diesem Zusammenhang stellt sich die Frage, was Wissen eigentlich ist. Das Wort „Wissen“ ist in der Umgangssprache ein vielseitig genutztes Wort, bei dem nur schwer festzustellen ist, was es wirklich bedeutet. Es gibt die unterschiedlichsten Ansätze, das Wort Wissen sprachlich, intellektuell oder wissenschaftlich zu definieren. Umgangssprachlich verbindet man Wissen mit Erkenntnis, Erfahrung und Bildung. Aber diese Worte helfen einem nicht weiter, da sie in ihrer Beschreibung der Wissensakquisition genauso schwammig sind wie das Wort „Wissen“ selbst. Etwas konkreter umschreiben die Worte „Fakten, Tatsachen, Methoden, Handfertigkeiten und Verhaltensschemen“ den Begriff des Wissens. Die Psychologie unterscheidet zusätzlich bewußtes und unterbewußtes Wissen. Das Wissen, das ein Individuum besitzt und das Gemeinschaftswissen, das sich aus vielen Wissensfragmenten der Individuen zusammensetzt und unter Umständen für sich allein keine Bedeutung besitzt. Ein Beispiel hierfür ist ein Ameisenstamm, in dem unterschiedliche Ameisengruppen existieren, die jeweils unterschiedliche Aufgaben erfüllen. Jede Gruppe allein für sich kann nicht überleben. Nur durch das Zusammenwirken stellen die Ameisen eine komplexe und flexible Lebensform dar, die effektiv den Überlebenskampf meistert.

In der philosophischen Tradition werden drei Bedingungen gestellt, die erfüllt sein müssen, damit gesagt werden kann „A weiß p“: A glaubt, daß p gilt; p ist wahr; A kann seinen Glauben, daß p gilt, rechtfertigen oder beweisen. Dementsprechend wird angenommen, daß *echtes Wissen* nur in einem Glauben besteht, den man beweisen oder begründen kann. Der Skeptizismus hingegen bezweifelt, daß es ein sicheres

Wissen geben kann, da er davon ausgeht, daß man Glauben nicht beweisen kann. Edmund Gettier stellt daher die These auf, daß Glauben nicht bewiesen, sondern belegt oder gerechtfertigt werden muß.

Auch der Wechsel in andere Sprachen hilft nicht weiter. Das Wort „know“ im Englischen bringt beispielsweise noch zusätzliche Bedeutungen wie „Kennen“ mit in die Bedeutung vom Wissen. „I know the smell of pineapples“ würde im Deutschen wohl kaum übersetzt werden mit den Worten „Ich weiß den Geruch von Ananas“. Eine weitere Bedeutung von Wissen wird in dem Satz „Ich weiß, wie man Fahrrad fährt“ deutlich. Wissen ist also nicht nur etwas statisches (faktisches Wissen), sondern kann auch einen dynamischen Vorgang einschließen. Eine besondere Bedeutung gewinnt dieser Aspekt, da wir nicht nur Wissen über das Bildmaterial benötigen, sondern auch darüber, wie wir das Bildmaterial mit dem Computer interpretieren können. Ähnlich verhält es sich mit dem Erkennen von visuellen Informationen durch den Menschen. Der Mensch weiß, wie man sieht und Gegenstände erkennt, er weiß jedoch nicht, wie er es macht. Er kann nicht ohne Weiteres erklären, warum er einen Hocker als Stuhl erkennt und nicht als Blumentisch [Nae97, Get94].

In den aktuellen Segmentierungsverfahren ist das Wissen, mit dem das Bildmaterial interpretiert werden soll, meist implizit, selten explizit integriert. Es wird beispielsweise vorausgesetzt, daß ein Organ einen einheitlichen Grauwert oder eine durchgehende Kante zu dem umgebenden Gewebe aufweist. Wissen ist einerseits in den Bildpunkten enthalten, da sie mit ihren Grauwerten physikalische Eigenschaften beschreiben. Bei den CT-Aufnahmen ist es die Absorption von Röntgenstrahlen, bei der Kernspintomographie die Dichte von Elementarmagneten. Andererseits wird die Anatomie des Patienten nur durch die Konstellation der einzelnen Bildpunkte beschrieben. Doch ist medizinisches Bildmaterial auch Wissen oder sind es statt dessen Informationen. In welchem Verhältnis stehen Wissen und Informationen? Eine mögliche Verbindung stellt die Annahme dar, daß Wissen eine Einheit Information und gleichzeitig die Verfahren zur Interpretation der Information umfaßt. Die künstliche Intelligenz definiert sich den Begriff „Wissen“ wiederum anders. Sie definiert, daß Wissen die Fähigkeit ist, zu denken, Wissen zu erwerben und anzuwenden, die Fähigkeit wahrzunehmen und zu handeln [Win85].

Diese Betrachtungen zeigen, daß die Verwendung des Begriffs "Wissen" in wissenschaftlicher Hinsicht problematisch ist. Für die Behandlung der Bildverarbeitung spielen philosophische und psychologische Aspekte nur eine geringe Rolle. Daher soll im weiteren Verlauf dieser Arbeit der Begriff "Wissen" als Zusammenfassung für die Faktenkenntnisse über das Bildmaterial und die Methoden zur Interpretation der Bilder definiert werden.

Die vorangegangenen Erläuterungen haben gezeigt, wie wichtig das Wissen für die Bildverarbeitung ist, und daß die Definition dessen, was Wissen ist und wie es genutzt werden kann, zunächst geklärt werden muß, bevor ein Bildverarbeitungsalgorithmus entworfen oder beurteilt werden kann. Bevor daher in den Abschnitten 3.7 bis 3.11 der aktuelle Stand der wissenschaftlichen Arbeiten im Bereich der medizinischen Bildverarbeitung beschrieben wird, soll der Leser zunächst für die Anforderungen an ein Bildverarbeitungssystem sensibilisiert werden. Hierzu müssen zunächst die Fragen geklärt werden, was für Wissen in Bildern enthalten ist, wie man das Wissen kodieren und repräsentieren kann, woher das Wissen bezogen werden kann und wie das Wissen durch einen Problemlösungsprozeß genutzt werden kann.

3.6.1 Das Wissen in Bildern

J. MacCarthy, einer der Mitbegründer der Künstlichen Intelligenz [Sch91], postulierte: „Ein intelligentes System muß eine Repräsentation der Welt besitzen, d.h. auf eine Menge von Fakten zurückgreifen, das es manipulieren und erweitern kann.“ Für die medizinische Bildverarbeitung besteht diese Welt aus dem medizinischen Bildmaterial. Es ist daher von entscheidender Bedeutung, daß vor dem Entwurf eines wissensbasierten Systems definiert wird, welche Wissenswelt repräsentiert und wie die Repräsentation strukturiert wird. Aufgrund von zwei Faktoren existiert a priori Wissen über das medizinische Bildmaterial:

1. Der Mensch weist eine weitgehend einheitliche Anatomie auf.
2. Es werden weitgehend standardisierte Bildmodi verwendet. Das Wissen über die Bildmodi spiegelt sich in der Kenntnis wider, wie die verschiedenen Organe in den jeweiligen Bildmodi dargestellt werden.

Die Kenntnis dieser beiden Punkte stellt die Grundlage dar, um medizinisches Bildmaterial interpretieren zu können. In der klinischen Routine kommen jedoch Faktoren hinzu, die den Interpretationsprozeß erschweren oder zumindest bei der Auswertung der Bilddaten berücksichtigt werden müssen. Die Anatomie variiert von Patient zu Patient innerhalb einer physiologischen Schwankungsbreite. Während Augen eine vergleichsweise einheitliche Form aufweisen, unterliegt die Oberflächenform von Organen wie dem Dünndarm beträchtlichen Schwankungen. Die Kenntnis der Anatomie umfaßt dementsprechend sowohl die durchschnittliche Form der Organe einschließlich der möglichen Abweichungen davon. Zusätzlich variieren die bildgebenden Eigenschaften der medizinischen Geräte innerhalb eines gerätespezifischen Toleranzbereiches. Artefakte, Rauschen und Probleme bei der Bildberechnung können darüber hinaus zu einem Verfälschen der Bildinformationen führen. Ein Betrachter muß diese verschiedenen Einflüsse wahrnehmen und bei der Interpretation des Bildmaterials einbeziehen.

Um die Organe in dem Bildmaterial zu erkennen und in den medizinischen Kontext einzuordnen, muß der Betrachter oft viele Faktoren gleichzeitig auswerten. Hierzu gehören die Grauwerteigenschaften, der mittlere Grauwert, die Varianz der Grauwerte, die Kanteneigenschaften, der Kontrast zum umgebenden Gewebe, das Kantenprofil, die Kantenstärke, Formeigenschaften, die mittlere Fläche/Volumen, der Umfang/Oberfläche, die Krümmung der Kontur, die Orientierung und viele Faktoren mehr.

Durch Artefakte gehen Informationen im Bildmaterial verloren. Der Betrachter muß diese fehlenden Informationen durch kognitive Prozesse ausgleichen, um zu einer vollständigen Interpretation des Bildinhaltes zu gelangen. Sind die Grauwerte des Bildes überlagert, ist der Mensch in der Lage, fehlende Kanten- und Grauwertinformationen durch sinnvolle Ergänzung zu kompensieren. Darüber hinaus kann ein medizinisch ausgebildeter Betrachter sein Vorwissen für eine inhaltliche Überprüfung der Interpretation einbeziehen, sobald er das Material in den anatomischen Kontext eingeordnet hat. Ist ein Betrachter sehr geübt bei der Interpretation des Bildmediums erhöht sich die Fähigkeit, Fehler im Bildmaterial mental zu kompensieren. Grundsätzlich benötigt der Betrachter Wissen über

1. die anatomischen Organeigenschaften und
2. die Darstellungseigenschaften der Organe.

Die anatomischen Organeigenschaften umfassen Vorkenntnisse zu der Anatomie und Physiologie der Organe. Es beschreibt die dreidimensionale Form, Position, Relation zu anderen Organen und die Variabilität dieser Aspekte. Die Darstellungseigenschaften der Organe beschreiben, wie das jeweilige Organ in dem Bildmaterial erwartungsgemäß auszusehen hat und welche Abweichungen durch Artefakte auftreten können.

3.6.2 Wie kann Wissen kodiert und angewandt werden?

Nachdem im vorangegangenen Abschnitt das notwendige Wissen in die Kategorien anatomische Eigenschaften und Darstellungseigenschaften unterteilt wurden, ist die Frage zu beantworten, wie sich das Wissen dieser Kategorien kodieren läßt. Die im Bereich Expertensysteme entwickelten Ansätze basieren auf drei unterschiedlichen Wissenskonzepten [Pup93]:

- Regeln,
- Probabilistik und

- Fallbeispiele.

Regelbasierte Verfahren wenden Regeln oder ganze Regelsysteme zur Lösung von Problemen an. Regelsysteme stützen sich auf unterschiedlichen Grundlagen, die unterteilt werden können in Logik, Regeln, Objekte/Frames und Constraints [Pup96]. Eine beispielhafte Regel bezüglich der Darstellungseigenschaften des gegebenen medizinischen Bildmaterials könnten wie folgt lauten: in T1 MRT-Aufnahmen ist der Bulbus dunkel, das Fettgewebe hell und die Linse weist einen mittleren Grauwert auf.

Ein Nachteil solcher regelbasierter Verfahren ist ihr problemspezifischer Charakter. Für jedes Bildmaterial muß ein eigenes Netzwerk von Regeln entworfen und verifiziert werden. Die Schwierigkeit dieser Regelsysteme besteht weniger in dem Aufstellen der Regeln, sondern vor allem in dem Sicherstellen der Konsistenz der Regeln zueinander und in dem Warten des Systems über einen längeren Zeitraum hinweg. Je größer das Regelsystem wird, desto komplexer sind die Verflechtungen der Regeln untereinander. Viele regelbasierte Systeme sind gerade an diesem Punkt gescheitert [Pup93].

Erschwerend kommt hinzu, daß Bildmaterial keine wirklich regelmäßigen Bildeigenschaften aufweist, so daß die Anwendung eines Regelwerkes von vornherein in Frage gestellt werden muß. Bildeigenschaften variieren innerhalb von Organen, sei es durch physiologische Inhomogenitäten, durch Artefakte oder aufgrund des bildgebenden Verfahrens. Ein weiteres Problem stellt die Beschreibung der anatomischen Organeigenschaften dar, die nicht direkt durch Regeln darstellbar sind. Da es sich hier um dreidimensionale Oberflächenformen und die Konstellation von mehreren Oberflächen zueinander handelt, wäre eine Parametrisierung der Oberflächen, ihrer Variabilität und der räumlichen Anordnung mit anderen Organen vonnöten.

Diese Probleme des regelbasierten Ansatzes lassen sich auf die Ansätze mit statistischem Wissen übertragen. Im Unterschied zum regelbasierten Ansatz ließen sich die Bildeigenschaften automatisch von segmentierten Datensätzen ableiten, indem beispielsweise die Grauwerteigenschaften der einzelnen Organen statistisch untersucht oder durch neuronale Netze angelernt werden. Das Problem der Variabilität innerhalb des Organs bleibt jedoch bestehen. Für die Repräsentation vom anatomischen Organwissen wäre wie für ein regelbasiertes System eine Parametrisierung der Oberflächen erforderlich.

Bei beiden Wissensformen kommt ein weiteres Problem hinzu, daß ich mit dem Begriff der „Gleichzeitigkeit des Wissens“ bezeichnen möchte. Das Wissen über die Darstellungseigenschaften kann nur dann sinnvoll angewandt werden, wenn das anatomische Wissen mit den Darstellungseigenschaften in Beziehung steht. Darstellungs- und anatomische Eigenschaften müssen direkt miteinander korreliert sein, damit sie vollständig genutzt werden können. Es reicht nicht festzustellen, daß der Bulbus dunkel ist und bei der Kornea der helle Bereich der Linse ist, wenn nicht bekannt ist, wo sich die Kornea befindet. Position, Orientierung und die dreidimensionale räumliche Anordnung mit anderen Organen spielen eine wesentliche Rolle bei der Feststellung, welche Grauwerte welchem Organ entsprechen. Ein Segmentierungsverfahren sollte in der Lage sein, beide Wissensformen gleichzeitig und in synergetischer Wirkung anzuwenden. Der grundsätzlich unterschiedliche Charakter von Regeln und statistischen Methoden, die einerseits anatomische Organeigenschaften, andererseits Darstellungseigenschaften beschreiben, erschwert eine Integration beider Wissensformen.

Ein weiteres Verfahren zur Repräsentation und Anwendung von Wissen bietet das Fallbasierte Schließen (Case-based Reasoning). Dieser Ansatz bietet sich an, wenn eine große Sammlung von realen oder konstruierten Fallbeispielen einschließlich der jeweils korrekten Lösungen zur Verfügung steht. Ein neuer Fall wird bei dem „Case-based Reasoning“ gelöst, indem aus einer Fallsammlung (Falldatenbank) jener Fall ausgewählt wird, der mit dem aktuellen Fall am besten übereinstimmt. Anschließend wird die Lösung dieses Vergleichsfalls auf den aktuellen Fall übertragen. Um einen Fall in der Falldatenbank zu finden, der mit dem aktuellen Fall am besten übereinstimmt, muß zunächst festgelegt werden, worin sich

diese Übereinstimmung äußert. Es muß ein sogenanntes Ähnlichkeitskriterium aufgestellt werden, das die Ähnlichkeit des aktuellen Falls mit den Fällen der Datenbank quantifiziert. Mit jedem zusätzlichem Fall in der Datenbank steigt die Wahrscheinlichkeit, eine optimale Lösung für den aktuellen Fall zu finden. Ein Vorteil der Fallbasierten Klassifikation gegenüber der statistischen Klassifikation besteht darin, daß Daten und Ähnlichkeitskriterien weniger stringent definiert sein können. Die Schritte zur Bestimmung eines Vergleichsfalls sind dabei wie folgt:

1. Vorauswahl der zu vergleichenden Fälle,
2. Bestimmen der Ähnlichkeit zwischen den Vergleichsfällen und dem aktuellen Fall,
3. Interpretation der Fallunterschiede und Auswahl des ähnlichsten Falls oder ablehnen sämtlicher Fälle.

Eine Vorauswahl der Vergleichsfälle im ersten Schritt ist gerade bei großen Datenbanken aus Effizienzgründen notwendig und wird in der Regel anhand von wenigen Schlüsseigenschaften vorgenommen. Das Ergebnis dieser Vorauswahl ist dann ein Satz von Fällen, die zumindest in diesen Schlüsseigenschaften mit dem aktuellen Fall übereinstimmen. Ein grundlegender Faktor für die Qualität der Problemlösung stellt die Definition des Ähnlichkeitskriteriums dar. Ähnlichkeiten können durch ja-nein-Entscheidungen bei dem Feststellen von qualitativen Übereinstimmungen, durch quantitative Ähnlichkeiten (z.B. Abweichungen von einem erwarteten Grauwert) oder durch Kombinationen von diesen definiert werden. Ebenso kann der Auswahlprozeß in eine Sequenz von einzelnen Auswahlritten unterteilt werden, bei denen Ähnlichkeitskriterien mit unterschiedlicher Wichtung oder Ausrichtung angewandt werden.

Aufgrund der Komplexität der Ähnlichkeitsbestimmung bietet sich die Anwendung eines einfachen Expertensystems an [Pup96]. Welche Informationen die einzelnen Fälle beinhalten, hängt sehr stark von dem Anwendungsbereich und dem Charakter der Fälle ab. Die Wissensakquisition besteht in erster Linie aus dem Sammeln von Fällen und deren Lösungen. In der Regel wird am Anfang ein Satz von typischen Fällen in die Datenbank eingefügt.

Der noch vergleichsweise junge Problemlösemechanismus Case-based Reasoning hat erst in den letzten Jahren einen zunehmenden Zuspruch gefunden. Als der schwierigste Punkt bei der Realisierung erwies sich der Aufbau umfangreicher Falldatenbanken. Ein Vorteil des Case-based Classification ist die vergleichsweise einfache Wissensakquisition, da lediglich zuvor bearbeitete Fälle der Datenbank hinzugefügt werden müssen. Einige Beispiele für die Anwendung des Case-based reasoning finden sich in [Kol85, Kot88, Por90, Pup96].

Der Ansatz des Fallbasierten Schließens weist nicht die Probleme bei der Kodierung des Wissens auf wie regelbasierte und statistische Verfahren. Ein anwendbarer Fall für die Segmentierung medizinischer Bilddaten bestünde aus einem segmentierten dreidimensionalen Beispieldatensatz. Die Darstellungseigenschaften werden durch das Bildmaterial repräsentiert. Diese Darstellungsform hat den Vorteil, daß noch keine pauschalisierenden Annahmen über die Einheitlichkeit der Darstellungseigenschaften innerhalb der Organe gemacht werden müssen. Die anatomischen Organeigenschaften werden durch die Segmentierung beschrieben.

Da die Segmentierung von Datensätzen bei therapeutischen Verfahren, die über die einfache Diagnose durch betrachten der Bilder hinausgehen, sowieso durchgeführt werden muß, ist für die Beschreibung der anatomischen Organeigenschaften kein zusätzlicher Arbeitsaufwand erforderlich. Werden die Segmentierungen zu einem dreidimensionalen Objekt zusammengefaßt, enthält der Datensatz ebenfalls das Oberflächenwissen. Auch für diesen Schritt sind in der Regel keine zusätzlichen Arbeiten gegenüber der herkömmlichen Therapieplanung notwendig. Da die segmentierte Oberfläche direkt mit dem Datensatz korreliert ist, ist ebenfalls das Problem der Integration von beiden Wissensformen gelöst.

Die fallbasierte Darstellung des Wissens durch zuvor bearbeitete Fälle stellt somit eine Wissensrepräsentation dar, die nur einen geringen bis keinen zusätzlichen Arbeitsaufwand erfordert, beide Wissensformen, Bild- und Oberflächeneigenschaften in sich vereint und gleichzeitig miteinander integriert.

3.6.3 Wissensrepräsentation für die medizinische Bildverarbeitung

Um Wissen in einem Wissensbasierten Bildverarbeitungsverfahren anwenden zu können, muß zunächst die Frage geklärt werden, wie beschreibt oder repräsentiert man das Wissen über die Anatomie und die Bildeigenschaften (Wissenszustand [Sch91]). Dabei ist die Frage nach dem „Wie“ noch zu unpräzise, da es in der Regel mehrere Formen der Repräsentation des gleichen Wissens gibt. Eine grundlegende Anforderung an die Wissensrepräsentation ist, daß das Wissen so einfach und unverfälscht wie möglich repräsentiert sein sollte [Pup96]. Von Bedeutung ist ebenfalls die Effizienz, mit der auf relevante Daten zugegriffen werden kann.

Daraus leitet sich die Frage ab: „Wie können die Wissenszustände adäquat beschrieben werden“. Das adäquat bezieht sich auf die jeweils gegebene Aufgabenstellung, zu der das Wissen genutzt werden soll. Erst wenn die Frage nach der Wissensrepräsentation beantwortet ist, können die Inferenzverfahren (beispielsweise Segmentierungsverfahren) spezifiziert werden, die das Wissen anwenden.

Für die Auswertung von medizinischem Bildmaterial wird das alleinige Auswerten von Faktenwissen, das anhand von Beispielen zeigt, wie die korrekte Segmentierung eines Organs aussieht, nicht ausreichen. Für die Durchführung einer Segmentierung ist vor allem das algorithmische Wissen von großer Bedeutung, das angibt, wie die gesuchten Organe im Bildmaterial eingegrenzt werden können. Das Wissen wird demnach nach Habel [Hab85] unterschieden in deklaratives und prozedurales Wissen. Das deklarative Wissen unterteilt sich wiederum in faktuelles Wissen, Regelwissen und Objektwissen. Das prozedurale Wissen unterscheidet Wissen über Vorgänge, Vorgehen, Strategien beim Problemlösen, Inferenzprozesse und Metawissen über prozedurales und deklaratives Wissen. Um zunächst die Anforderungen an das deklarative Wissen zu spezifizieren, bieten sich die Kriterien für die Wissensrepräsentation nach Fikes [Fik85] an:

- **Object and relationship:** Die Anatomie ist einer der wichtigsten Anhaltspunkte zur Interpretation des medizinischen Bildmaterials. Hierzu gehören die Kenntnis der dreidimensionalen Form und räumlichen Anordnung der Organe.
- **Typical situations:** Für die Bildverarbeitung stellen beispielhafte Bilddaten eine gute Beschreibungsform der typischen Problemsituation dar. Für die Ausbildung von medizinischem Personal in der Radiologie sind beispielhafte Datensätze das wichtigste Lehrmedium. Ähnlich könnte für ein Expertensystem zuvor bearbeitetes Bildmaterial eine prägnante Beschreibungsform von Problemlösungen darstellen.
- **Processes:** Liegt das Bildmaterial vor und ist Wissen zur Identifikation von Organen verfügbar, müssen Vorgehensweisen und Inferenzmechanismen bekannt sein, um die Interpretation durchzuführen. Unter dem Begriff Vorgehensweisen lassen sich in erster Linie jene Algorithmen fassen, die eine Verarbeitung der Bilddaten vornehmen. Dabei werden die subsymbolischen Informationen in den Bildpunkten in symbolische oder quantitative Informationen überführt. Anschließend können eine Reihe von Entscheidungsmechanismen angewandt werden, um den Erkennungsprozeß anzustoßen. Hierzu gehören vor allem Verfahren wie Heuristiken (Regeln, die sich aus zuvor untersuchten Fällen ableiten lassen), Hypothesengeneration (im Bereich der künstlichen Intelligenz wurden eine Vielzahl von Verfahren entwickelt, die Hypothesen aus einer Sequenz von Entscheidungsschritten ableiten), disjunctive facts (Entscheidungsmechanismen, die auf sich aus-

schließenden Fakten basieren) und Constraints (Entscheidungsmechanismen, die darauf basieren, daß Beschränkungen bezüglich der Variabilität der Organe und des Bildmaterials existieren).

- Behaviour descriptions: Die Variationsmöglichkeiten der Organe müssen bekannt sein. Hierunter sind in erster Linie die Größen-, Form- und Positionsvariationen zu verstehen. In einzelnen Fällen kann es auch die Anzahl sein. Ein Beispiel hierfür ist die Tatsache, daß einige Menschen zwei, anstatt drei Nieren besitzen. Bei Kindern muß zusätzlich die Entwicklung der Organe über die Zeit hinweg berücksichtigt werden, da sich Form, Größe und Proportionen kontinuierlich verändern können.
- Decision rules: Selbst bei Kenntnis der Anatomie ist die Unterscheidung von Organen nicht immer eindeutig allein aufgrund von Form und Position möglich. Besonders wenn Krankheiten das Erscheinungsbild der betroffenen Organe beeinflussen, müssen zusätzliche Entscheidungsmechanismen zur Identifikation angewandt werden. Ärzte erarbeiten sich diese Entscheidungsmechanismen in ihrer Ausbildung und in jahrelanger praktischer Erfahrung. Soll ein Bildverarbeitungssystem in der Lage sein, ebenfalls solche Entscheidungen zu treffen, müssen diese zusätzlichen Entscheidungsmechanismen algorithmisch übertragen werden.
- Uncertain facts: Um das in der Praxis anfallende Bildmaterial bearbeiten zu können, muß das Bildverarbeitungssystem nicht nur in der Lage sein, erwartetes Bildmaterial zu interpretieren. Es muß darüber hinaus mit unerwarteten Veränderungen wie beispielsweise Artefakten und Rauschen umgehen können, ohne dramatische Einbrüche in der Ergebnisqualität hinnehmen zu müssen.
- Vocabulary definitions: Es ist wichtig, die erkannten Bildelemente nach einer einheitlichen Nomenklatur zu bezeichnen. Dieser Punkt ist besonders für die Strahlentherapie von Bedeutung, da klar zwischen den verschiedenen Zielvolumina unterschieden werden muß. So kann das Zielvolumen in dem einen Fall das krankhafte Gewebe umschließen, während im anderen Fall zusätzlich ein Toleranzbereich aus therapeutischen Gründen einbezogen wurde.

Um zu entscheiden, ob eine gewählte Wissensrepräsentation den Anforderungen genügt, müssen Adäquanzkriterien erfüllt werden. Nach Schefe [Sch91] müssen die folgenden beiden Kriterien durch das System abgedeckt werden:

1. Epistemologische Adäquanz: Werden alle das Wissen betreffenden Gegebenheiten abgebildet? Wurden alle notwendigen Wissensinhalte modelliert?
2. Ergonomische Adäquanz: Ist das Wissen leicht handhabbar und wartbar? Sind die Wissensstrukturen verständlich?

Das für die Bildverarbeitung notwendige prozedurale Wissen charakterisiert sich durch den Segmentierungsalgorithmus und kann an dieser Stelle noch nicht genauer beschrieben werden. In den Abschnitten 3.8 bis 3.11 werden die aktuellen Ansätze zur Bildverarbeitung genauer bezüglich des vorausgesetzten deklarativen und prozeduralen Wissens betrachtet.

3.6.4 Die Wissensquelle

Nach dem Festlegen der grundsätzlichen Kriterien für die Wissensrepräsentation sind die zur Verfügung stehenden Wissensquellen zu untersuchen. In der Regel stehen für die medizinische Bildauswertung zwei Wissensquellen zur Verfügung: Experten und Lehrbücher. Bei der Entwicklung eines wissensbasierten Systems besteht die erste Problemstellung in der Transformation des Wissens von diesen Quellen in eine Form, die ein entsprechend entwickeltes Bildverarbeitungsprogramm nutzen kann.

In fast allen Fällen wird der medizinische Experte nicht die Informatikkenntnisse haben, um sein Wissen in einem komplexen Expertensystem umzusetzen. Auf der anderen Seite verfügt der Informatiker in der Regel nicht über das medizinische Wissen, um das Expertensystem korrekt zu konfigurieren. Auch der heutzutage oft angewandte Ansatz, einen Wissensvermittler einzusetzen, den sogenannten Knowledge-Engineer, ist aufgrund des nicht immer eindeutigen Informationsaustausches fehleranfällig. Es ist daher von besonderer Wichtigkeit, eine Form der Wissensrepräsentation zu finden, die eine Verfälschung vermeidet.

Wieder bietet sich der Ansatz des fallbasierten Schließens an, bei dem die bisher bearbeiteten Fälle die Wissensquelle darstellen. Indem die in der Vergangenheit segmentierten Fälle in einer Falldatenbank aufgenommen und gesammelt werden, stellt der behandelnde Arzt die direkte Wissensquelle dar, ohne daß ein zusätzlicher Arbeitsaufwand durch den Arzt aufgebracht werden müßte.

Eine mögliche Schwachstelle dieses Ansatzes stellt die Tatsache dar, daß sich die Qualität der Segmentierung in den Fallbeispielen direkt auf die bearbeiteten Fälle überträgt. Ist die Segmentierung eines Falls in der Datenbank schlecht und wird dieser Fall für einen aktuellen Datensatz ausgewählt, so wird auch die Segmentierung des aktuellen Falls schlecht sein. Fügen mehrere Ärzte Fälle der Falldatenbank hinzu und vertreten diese Ärzte unterschiedliche Kriterien für die Segmentierung von Datensätzen, so werden auf den aktuellen Fall jeweils die Kriterien übertragen, die bei der Segmentierung des Fallbeispiels angewandt wurden. Eine nicht zu unterschätzende Schwierigkeit besteht demnach darin, eine konsistente Falldatenbank aufzubauen.

3.6.5 Der Problemlösungsprozeß

Ist die Wissensrepräsentation definiert, kann ein Problemlösungsprozeß entworfen werden, der unter Ausnutzung des gegebenen Wissens die aufgeworfenen Fragestellungen zu lösen sucht. Der Problemlösungsprozeß besteht aus drei Schritten:

1. Der Definition der Problemlage,
2. dem Suchen und Konzipieren eines Problemlösungsverfahrens und
3. dem Erreichen und Bewerten der Problemlösung.

Diesen Prozeßschritten entsprechen die drei Problemlösungszustände: Anfangszustand, Transformationsphase und Zielzustand [Sch91]. Wird das notwendige Wissen zur Verfügung gestellt, ist in der Transformationsphase zunächst eine Problemklassifikationsmethode nötig. In ihr wird zunächst der Zugriff auf das Wissen durchgeführt, um festzustellen, was für ein Problem bearbeitet werden muß. Im Falle der Auswertung von medizinischem Bildmaterial muß festgestellt werden, um was für ein Bildmaterial es sich handelt.

Konnte die Problemstellung konkretisiert werden, sind mögliche Lösungen zu der Problemstellung zu identifizieren. Bei der von Clancey [Cla85] beschriebenen heuristischen Problemklassifikation (heuristic classification) wird die Lösung aus einer Auswahl von Alternativen ausgewählt. Dabei können die Lösungen nicht aus den Daten direkt, sondern über interpretative Zwischenschritte, der sogenannten Daten- und Lösungsabstraktion, abgeleitet werden. Bei diesem Abstraktionsschritt werden datenkontrolliert Regeln zur Verwertung des Wissens angewandt.

Übertragen auf die Bildverarbeitung würde dieser Schritt bedeuten, daß zunächst das Bildmaterial klassifiziert werden muß. Nachdem bestimmt wurde, um welchen Bildmodus es sich handelt und welche Organe allem Anschein nach dargestellt werden, wird ein Segmentierungsverfahren ausgesucht, das für die Auswertung von diesen Bilddaten geeignet ist.

Problemlösungsmethoden können in weiche und starke Problemlösungsverfahren unterschieden werden. Weiche Problemlösungsverfahren können für eine Vielfalt von Aufgabenstellungen und Wissensformen angewandt werden. Typische Beispiele für diese Kategorie sind grundlegende Wissensformen mit ihren Inferenzmechanismen wie Regeln, Frames oder Constraints [Pup96]. Starke Problemlösungsmethoden sind weniger flexibel, jedoch besser angepaßt an den jeweiligen Anwendungsbereich und unterstützen die Wissensakquisition.

Die Ergebnisse aktueller Veröffentlichungen zeigen, daß je präziser die Parameter und Interdependenzen der Wissensrepräsentation definiert sind, desto zielsicherer und zuverlässiger arbeitet das System. Ein Großteil der aktuellen Bildverarbeitungsalgorithmen sind weiche Problemlösungsverfahren, da sie auf unterschiedliches Bildmaterial angewandt werden können. Im Fall der Segmentierung von medizinischem Bildmaterial könnte der Problemlösungsprozeß durch die folgenden Zustände umschrieben werden:

1. Es liegt ein aktueller Datensatz vor, der segmentiert werden soll.
2. Es wird ein Segmentierungsverfahren ausgewählt, angepaßt und angewandt, um die gesuchten Organe zu segmentieren.
3. Die Segmentierung der Organe wird durch einen Arzt betrachtet, gegebenenfalls korrigiert und abschließend als korrekt bestätigt.

Aktuelle Bildverarbeitungsansätze beschränken sich auf die Durchführung der zweiten Hälfte des zweiten Schrittes, der Anwendung eines Problemlösungsverfahrens. Da keine Auswahl und Anpassung eines geeigneten Verfahrens vorgenommen wird, müssen fest definierte Bildeigenschaften vorausgesetzt werden, für die der jeweilige Bildverarbeitungsansatz sinnvolle Ergebnisse produzieren kann. Demgegenüber steht der Vorsatz, einen Bildverarbeitungsansatz zu entwickeln, der auf möglichst viel unterschiedliches Bildmaterial angewandt werden kann. Aus diesem Grund versuchen aktuelle Bildverarbeitungsansätze die Grundlagen des Problemlösungsmechanismus möglichst allgemein zu halten, indem beispielsweise vorausgesetzt wird, daß Organe einheitliche Grauwerte aufweisen oder vollständig vom umgebenden Gewebe durch eine Kante abgegrenzt werden. Dieser Versuch zur Verallgemeinerung der Verfahren führt dazu, daß keine konkreten Vorgaben für den Zielzustand der Segmentierung gemacht werden oder gemacht werden können. Unterschiedlichste Formen und Grauwerteigenschaften können eine korrekte Segmentierung ausmachen. Aufgrund der variablen Bildeigenschaften, Patientenanatomien und dem Interpretationsspielraum, der zwischen den Segmentierungen unterschiedlicher Ärzte auftreten können, existiert bei aktuellen Bildverarbeitungsalgorithmen keine präzise Zieldefinition, die im dritten Schritt oder noch während der Segmentierung auf Konsistenz getestet werden könnte.

Analysiert man den Problemlösungsprozeß der medizinischen Bildverarbeitung nach den Kriterien von Scheffe [Sch91], kann festgestellt werden, daß es sich bei der Segmentierung von CT- und MRT-Daten um einen zerlegbaren, kommutativen und reversiblen Problemlösungsprozeß handelt. Die Segmentierung eines dreidimensionalen Datensatzes ist zerlegbar, da der Segmentierungsvorgang in die Segmentierung der einzelnen Organe und dieser Schritt wiederum in die Auswertung der zweidimensionalen Schichten aufgeteilt werden kann. Die Segmentierung ist kommutativ, da es unerheblich ist, in welcher Reihenfolge die Organe segmentiert werden, sofern das Segmentierungsverfahren nicht versucht, Segmentierungsergebnisse untereinander zu nutzen. Der Segmentierungsprozeß ist außerdem reversibel, da sämtliche Segmentierungen wieder korrigiert und aufgehoben werden können, ohne die zuvor bestimmten Ergebnisse ungültig zu machen.

Problemlöser für reversible Prozesse sind in der Regel so strukturiert, daß sämtliche Lösungsmöglichkeiten durchsucht und durch Überprüfen der Zielbedingung die finale Lösung bestimmt wird [Sch91, Seite 48]. Für die Bildverarbeitung ist dieser Ansatz nicht praktikabel, da der Suchraum für die Segmentierung eines Organs zu groß ist. Es gibt zu viele mögliche Segmentierungen, die getestet werden müßten.

Statt dessen muß eine zielgerichtete Suche durchgeführt werden. Aktuelle Bildverarbeitungsansätze verfolgen in der Regel den „Hill climbing“-Ansatz (Bergsteigermodell). Bei diesem Ansatz wird die Lösung gesucht, indem jeweils in der Richtung gesucht wird, in der der Gradient der Lösungsfunktion am steilsten ist. Diese Lösungsfunktion beschreibt die Übereinstimmung der aktuellen Lösung mit der Zieldefinition. Die Zieldefinition besteht in diesem Fall wiederum in den Annahmen über die Einheitlichkeit der Grauwerte oder Kanten bei einem Organ.

3.7 Konzepte der Bildverarbeitung

Die Aufgabenstellungen im Bereich der medizinischen Bildverarbeitung sind so vielseitig, daß nur schwer ein umfassender Überblick über die Konzepte der Bildverarbeitung gewährt werden kann, ohne wichtige Anwendungsbereiche unberücksichtigt zu lassen. Abbildung 3.11 beschreibt die klassischen Verarbeitungsschritte der Bildverarbeitung von der Bildakquisition bis zum interpretierten Datensatz nach Jähne. Unter einem interpretierten Datensatz sei hierbei ein Datensatz zu verstehen, in dem den Bildpunkten des Bildes eindeutige Identitäten zugeordnet wurden. Das bedeutet, daß zu jedem Pixel bekannt ist, zu welchem Organ es gehört. Diesen Vorgang des Identifizierens der Bildpunkte wird Segmentierung genannt und kann entweder auf das gesamte Bild oder auf Teile des Bildes angewandt werden. Die Segmentierung eines Datensatzes wird erreicht, indem die im Bild dargestellten Objekte anhand von Konturen vom umgebenden Gewebe abgegrenzt werden oder alle Pixel, die eine gemeinsame Identität besitzen aber nicht miteinander verbunden sind, klassifiziert werden.

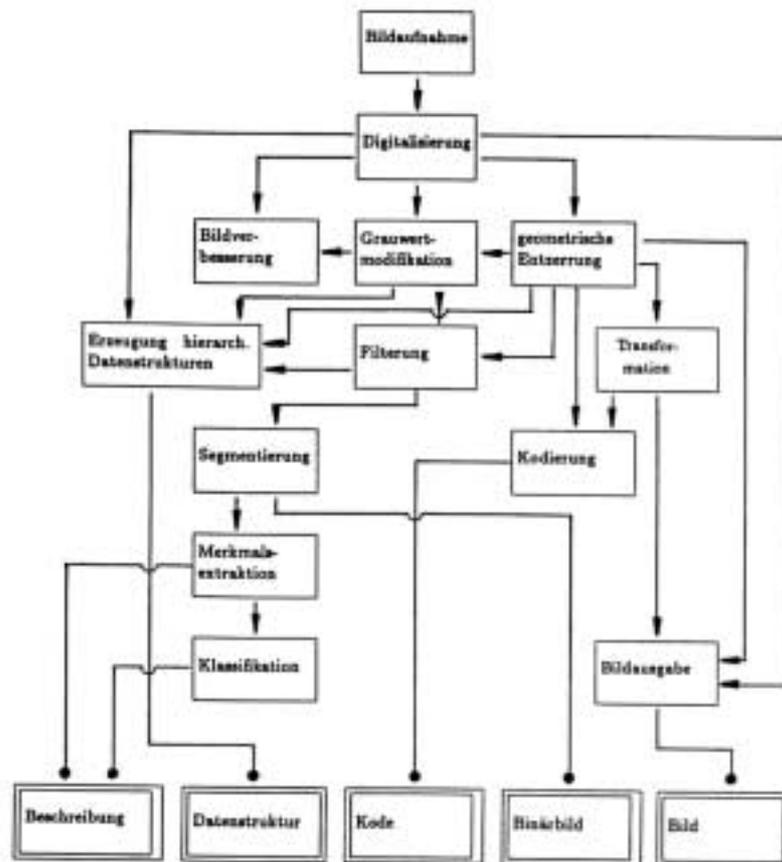


Abbildung 3.11: Klassische Verarbeitungsschritte der Bildverarbeitung nach Jähne [Jae92]

Das Bilderkennungsprinzip nach Jähne versucht so viele Bildverarbeitungs-konzepte wie möglich in der strukturierten Grafik festzuhalten. Der aktuellen Vielseitigkeit in den Anwendungsbereichen und Bildverarbeitungstechniken kann sie jedoch nicht mehr in vollem Umfang Rechnung tragen.

Um von dem aufgenommenen Bild zu einem segmentierten Bild zu gelangen, können die unterschiedlichsten Ansätze verfolgt werden. Die Digitalisierung, bei der ein analoges Bild in ein digitales überführt wird, fällt bei medizinischen bildgebenden Verfahren wie der Computer- und der Kernspintomographie weg, da die Daten bereits digital erzeugt werden. Durch Verfahren wie Bildverbesserung, Grauwertmodifikation, geometrische Entzerrung, Filtern und Transformationen können Fehler im Bildmaterial ausgeglichen und relevante Bildinhalte hervorgehoben werden. Übliche Aufgabenstellungen solcher Verfahren sind die Elimination von Artefakten, die Verbesserung der Darstellungseigenschaften wie Kontrast und Signal-Rausch-Verhältnis oder das Hervorheben von Kanten oder anderen Bildelementen. Segmentierung, Merkmalsextraktion und Klassifikation sind Verfahren zur Interpretation des Bildmaterials, bei der einzelnen Pixeln oder Pixelgruppen die Organidentitäten zugeordnet werden.

Zusätzlich zu den in der Grafik dargestellten Verfahren zur Bildverarbeitung haben sich eine Reihe von wissenschaftlichen Forschungsbereichen entwickelt, die auf den segmentierten Datensätzen aufsetzen und weitergehende Aufgabenstellungen behandeln. Hierzu gehören die Visualisierung und Kombination (Registrierung) von unterschiedlichen Bilddaten, Virtuelle Realität und Operationssimulationen wie auch die computergestützte Strahlentherapie. Die dreidimensionale Visualisierung der Daten stellt eine Erweiterung der herkömmlichen Auswertung medizinischer Bilddaten dar und versucht durch eine realitätsnahe oder intuitive Darstellung der Bildinhalte das Verständnis komplexer medizinischer Sachverhalte zu vereinfachen. Die Kombination von unterschiedlichen Bilddaten hat seinen Ursprung in der Tatsache, daß die verschiedenen bildgebenden Verfahren in der Medizin spezifische Darstellungseigenschaften aufweisen. Damit verbunden sind jeweils voneinander abweichende Vor- und Nachteile eines jeden Bildmodus, die in Abhängigkeit von der jeweiligen medizinischen Aufgabenstellung mehr oder weniger zum Tragen kommen. Durch die Kombination von unterschiedlichen Datensätzen wird versucht, die Informationen synergetisch zusammenzuführen, während sich Nachteile gegenseitig kompensieren können.

Ein Forschungsbereich, der die Medizin in Zukunft grundlegend verändern wird, ist die Anwendung von medizinischen Bilddaten für Therapiesimulationen und computergestützte Therapiedurchführungen. Während die bisherigen Punkte vor allem die Diagnostik betrafen, ermöglichen die Therapiesimulation und die computergestützte Therapiedurchführung eine grundlegende Verbesserung des Behandlungsablaufes. Therapien können noch vor der Anwendung am Patienten optimiert werden. Sofern darüber hinaus die Gerätschaften für einen mit der Patientenanatomie kalibrierten Eingriff existieren, kann die Präzision der Eingriffe deutlich verbessert, das Risiko für den Patienten hingegen deutlich reduziert werden. Da sich diese Arbeit mit anderen Themen der Bildverarbeitung beschäftigt, würde ein genaueres Eingehen auf die Möglichkeiten der Virtuellen Realität über den Rahmen dieser Arbeit hinausgehen. Aus diesem Grund sei der Leser auf die entsprechende Fachliteratur verwiesen [Kal93, Mor97, Weg96].

Segmentierung, Visualisierung, Registrierung, Virtuelle Realität und computergestützte Strahlentherapie sind eng miteinander verwoben, da die Algorithmen des einen Verfahrens oft auf den Ergebnissen eines anderen Verfahrens aufsetzen. Um Organe zu visualisieren oder Datensätze miteinander zu kombinieren, müssen zunächst die Organe segmentiert oder eine andere Form der Interpretation der Grauwerte vorgenommen worden sein. Verfahren der Virtuellen Realität hingegen basieren sowohl auf segmentierten Datensätzen als auch auf schnellen Verfahren zur Visualisierung der Organe. Der Fokus dieser Arbeit liegt bei der Interpretation und Registrierung von Bilddaten. Die nachfolgenden Beschreibungen der Problemstellungen und Lösungskonzepte aktueller Ansätze für die Bildverarbeitung beschränken sich somit auf diese vorgegebenen Themenbereiche.

3.8 Segmentierung

Die Interpretation eines medizinischen Datensatzes besteht in erster Linie in der Segmentierung der enthaltenen Organe. Segmentierung ist das Zusammenfassen von inhaltlich zusammengehörigen Bildpunkten

zu Bildpunktgruppen. Diese Bildpunktgruppen können einerseits zusammenhängend ein abgeschlossenes Organ beschreiben. Andererseits können die einzelnen Pixel einer Bildpunktgruppe getrennt voneinander über das Bildmaterial verteilt sein und jene Punkte innerhalb des Bildes beschreiben, die bestimmte Gewebeeigenschaften aufweisen.

In der aktuellen Literatur werden die unterschiedlichsten Verfahren zur Segmentierung von Bildmaterial beschrieben [Pal93]. Um zu entscheiden, welche Segmentierungskonzepte für eine gegebene Aufgabenstellung am besten geeignet sind, müssen zwei grundlegende Punkte untersucht werden: Einerseits müssen die zu erwartenden Eigenschaften des Bildmaterials, andererseits die Möglichkeiten des Segmentierungsverfahrens mit diesen Bildeigenschaften umzugehen, analysiert werden. Die grundlegenden Eigenschaften des zu bearbeitenden Bildmaterials wurden im Abschnitt 2.3 beschrieben.

Um die Möglichkeiten der in der Literatur beschriebenen Segmentierungskonzepte beurteilen zu können, sind die Verfahren zunächst nach ihren maßgeblichen Vorgehensweisen zu kategorisieren. Diese Aufgabe gestaltet sich jedoch recht schwierig, da sich keine stringenten Kriterien für die Kategorisierung von Segmentierungsverfahren etabliert haben. Viele Entwicklungen leiten ihre Konzepte aus mehreren Ansätzen ab und kombinieren sie, um Nachteile des einen Ansatzes durch Vorteile des anderen zu kompensieren. Durch diese Vermischung der Ansätze muß eine Kategorisierung nach der im jeweiligen Verfahren angewandten Technik fehlschlagen. So werden in der Regel mehr oder weniger willkürliche Hauptkategorien eingeführt, nach denen die Verfahren dann sortiert werden (vergleiche [Pal93, Mei91]).

Ein Aspekt, der im weiteren Verlauf dieser Arbeit als wichtigster Anhaltspunkt für die Möglichkeiten eines Verfahrens ausgearbeitet werden soll, ist das bei der Konzeption der Verfahren vorausgesetzte Vorwissen über das Bildmaterial und die dargestellten Organe. Dementsprechend wurden die nachfolgenden Kategorien von Segmentierungsverfahren aufgestellt, die sich an der Art und dem Wirkungsradius der Identitätszuordnungen orientieren:

1. Interaktive Segmentierungsverfahren,
2. Punktorientierte Klassifikationsverfahren,
3. Regionenorientierte Verfahren,
4. Kantenorientierte Verfahren,
5. Transformationsverfahren,
6. Modellbasierte Verfahren und
7. Wissensbasierte Verfahrenskombinationen.

In den nachfolgenden Abschnitten werden diese Kategorien und die ihnen zugeordneten Ansätze der aktuellen Literatur erläutert.

3.8.1 Interaktive Segmentierungsverfahren

Die interaktive Segmentierung umfaßt alle Verfahren, bei der ein Benutzer die Segmentierung interaktiv mit der Maus oder einem anderen Eingabemedium maßgeblich vornimmt. Hierunter sind in erster Linie die manuelle Segmentierung und manuell unterstützte Verfahren wie beispielsweise das Live-Wire-Verfahren zu verstehen.

Sowohl bei der manuellen Segmentierung als auch bei den Live-Wire-Verfahren werden die gesuchten Organgrenzen eingezeichnet [Udu82]. Während bei der manuellen Segmentierung alle Punkte der eingezeichneten Kontur durch den Benutzer gesetzt werden, fügt das Live-Wire-Verfahren automatisch weitere Punkte zwischen den Punkten, die der Benutzer setzt, ein [Fal97, Fal98, Mor95]. Die Position der zusätzli-

chen Zwischenpunkte ermittelt der Algorithmus automatisch auf der Basis von im Bild detektierten Kanten oder spezifischen lokalen Grauwertverläufen. Durch diesen Ansatz versucht das Live-Wire-Verfahren den Arbeitsaufwand für den Benutzer zu reduzieren, ohne dabei die Kontrolle über den Segmentierungsvorgang abzugeben.

Die Ergebnisse des Live-Wire-Verfahrens erwiesen sich jedoch als nicht sehr hilfreich. Der Live-Wire-Ansatz setzt voraus, daß die Kanteneigenschaften eines Objektes einheitlich sind. Statt dessen ist bei medizinischem Bildmaterial immer wieder festzustellen, daß Kantenstärke und Kantenprofil selbst innerhalb einer Schicht sehr stark variieren. Dadurch werden in vielen Fällen die automatisch hinzugefügten Punkte falsch positioniert. Der Zeitaufwand für die daraus resultierenden Korrekturen der Fehlpositionierungen war in der Regel so groß, daß kein entscheidender Gewinn gegenüber der reinen manuellen Segmentierung gewonnen werden konnte.

Manuelle Segmentierungsverfahren setzen voraus, daß der Benutzer die Kompetenz zum Durchführen dieser Aufgabe besitzt. Diese Vorgabe wird in der Regel durch das medizinische Personal erfüllt sein, schränkt jedoch den Benutzerkreis auf eben dieses Personal ein. Darüber hinaus wird angenommen, daß die Ergebnisqualität für alle Datensätze und Schichten einheitlich ist. Das bedeutet, daß die Konzentration, mit der die Arbeiten durchgeführt werden, immer gleich bleiben muß. In Anbetracht der Tatsache, daß moderne bildgebende Verfahren wie Kernspintomographie und Computertomographie immer größere Mengen an Bildmaterial produzieren, die nicht selten um die hundert Bildschnitte pro Patient umfassen, ist diese Vorgabe kaum zu erfüllen. Hinzu kommt die Tatsache, daß ein menschlicher Benutzer nicht wie ein Computer bei dem gleichen Bildmaterial auch die gleichen Ergebnisse produzieren wird. Die nicht gleichbleibende Konzentration und die Schwankungen der menschlichen Handlungsweisen führen zu einer intraindividuellen Variabilität der segmentierten Organgrenzen. Die Ergebnisse der interaktiven Segmentierung sind daher in der Regel nicht reproduzierbar (vergleiche [Gal93]).

Die intraindividuellen Schwankungen bei der Segmentierung wirken sich auch auf die Qualität von rekonstruierten dreidimensionalen Oberflächen aus. Werden aus den zweidimensionalen manuellen Segmentierungen die dreidimensionalen Organoberflächen rekonstruiert, so muß man in der Regel feststellen, daß die Oberflächen kantig und unregelmäßig sind. Die Ursachen hierfür sind vor allem in der Nicht-Reproduzierbarkeit der Segmentierung und in erschwerenden Faktoren wie Artefakte im Bildmaterial zu suchen, die zu inkonsistenten Segmentierungen zwischen den einzelnen Schichten führen. Der Effekt der intraindividuellen und interindividuellen Variabilität der Segmentierungen wird durch Artefakte wie beispielsweise dem Partialvolumen-Effekt deutlich verstärkt. Die Folge ist, daß selbst bei größter Konzentration keine vollständig korrekten und übereinstimmenden Eingaben der Organgrenzen erzielt werden können. Ein weiterer Nachteil, in dem sich alle interaktiven Segmentierungsverfahren gleich sind, ist der große Arbeitsaufwand, der für ihre Anwendung erforderlich ist.

3.8.2 Punktorientierte Klassifikationsverfahren

Punktorientierte Klassifikationsverfahren umfassen sämtliche Verfahren, die eine Zuordnung der Identität zu jedem Pixel allein auf der Information zu dem betreffenden Bildpunkt oder einem kleinem Bildausschnitt vornehmen und dabei globale Kriterien anwenden. Dieser Kategorie gehören Thresholding-Verfahren [Bel89] und komplexe Klassifikationsverfahren an, die eine globale Identitätszuordnung durch Optimierung von statistischen oder anderen Parametern vornehmen [Hsi89, Hur96, Sny92, Sol97, Tsa92]. Ansätze mit neuronalen Netzen sind auch dieser Kategorie zuzuordnen, sofern die Eingänge auf wenige Pixel beschränkt sind [Leu91, Tou92, Zhu97].

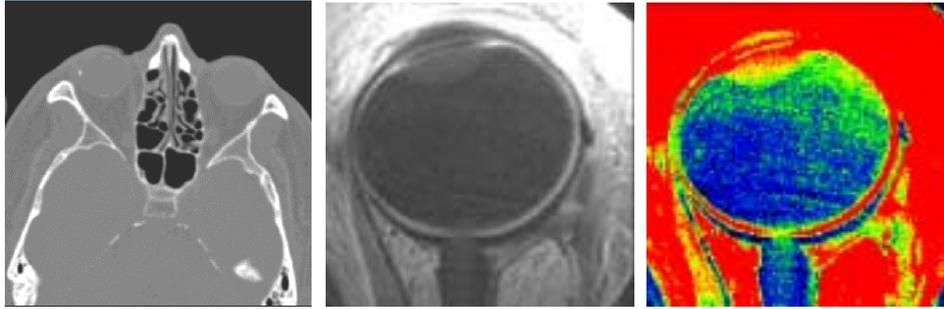


Abbildung 3.12: Problemstellungen bei der Punktorientierten Segmentierung

Links: Die punktorientierte Klassifikation der weißen Knochenstrukturen kann vergleichsweise problemlos durchgeführt werden, da der Grauwertunterschied zu allen anderen Strukturen sehr groß ist. Mitte: Kernspinaufnahme des Auges. Während der Mensch keine Probleme bei dem Erkennen des dunklen Bulbus hat, zeigt die Falschfarbendarstellung des gleichen Bildes (rechts), wie unterschiedlich die Grauwerte innerhalb des Bulbus sind und daß sich die Grauwerte des Bulbus mit denen von mehreren anderen Strukturen wie dem Sehnerv und den Muskelansätzen überschneiden.

Punktorientierte Verfahren setzen voraus, daß sich die Bildpunkte eindeutig durch ihren Grauwert von anderen Organen und vom umgebenden Gewebe abgrenzen lassen. Diese Annahme trifft jedoch nur für die Segmentierung von Knochen in CT-Aufnahmen zu (siehe Abbildung 3.12, links). Knochen absorbieren deutlich mehr Röntgenstrahlen als andere Gewebearten. Dadurch unterscheiden sich Knochenstrukturen durch deutlich hellere Grauwerte vom Weichgewebe und können dadurch problemlos durch ein Thresholding mit dem sogenannten Knochenfenster segmentiert werden.

Für nahezu alle anderen Organe und Bildmodi gilt diese Voraussetzung nicht (vergleiche Abbildung 3.12 Mitte und rechts, und Abbildung 3.13). Aus diesem Grund bleibt der Anwendungsbereich des Thresholdings bei der Auswertung von medizinischem Bildmaterial auf die Segmentierung von Knochen in CT-Aufnahmen beschränkt.

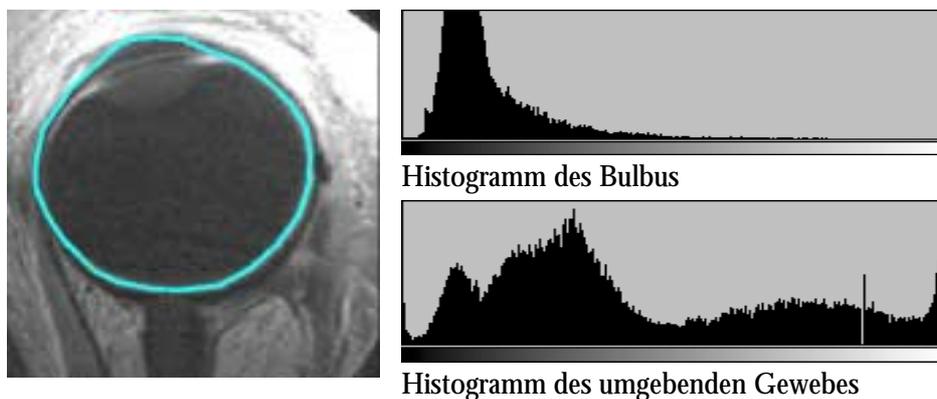


Abbildung 3.13: Punktorientierte Segmentierung des Bulbus

Vergleicht man die Grauwerte, die innerhalb des Bulbus liegen mit jenen die außerhalb liegen, stellt man fest, daß sich die Grauwertbereiche überschneiden. Eine eindeutige Segmentierung anhand der Grauwerte ist somit nicht möglich.

Eine neue Bedeutung hat die punktorientierte Klassifikation seit der Möglichkeit gewonnen, multimodale Bildauswertungen vorzunehmen. Der gängigste Anwendungsbereich ist die Auswertung von CT und MRT-Daten oder von mehreren MRT-Daten unterschiedlicher Aufnahmesequenzen [Ber96,Ber97]. Dabei wird die punktorientierte Klassifikation multimodal gleichzeitig auf das unterschiedliche Bildmaterial angewandt. Entscheidungsgrundlage ist in diesem Fall nicht mehr nur der Grauwert des einen Bildes,

sondern die gleichzeitige Auswertung der Grauwerte des gleichen Volumensegments in den unterschiedlichen Bildmodi.



Abbildung 3.14: Typische Ergebnisse des Thresholding

Links ist eine MRT-Aufnahme der Orbita zu sehen. Die Aufgabenstellung besteht in der Segmentierung des Bulbus. Grundlegendes Problem bei der Segmentierung des Bulbus mittels Thresholding ist die Festlegung des Schwellwertes. Mitte: Der Schwellwert ist zu hoch angesetzt. Dadurch werden viele Bildelemente segmentiert (weiß), die nicht zum Bulbus gehören. Rechts: Der Schwellwert ist zu niedrig, Teile des Bulbus sind nicht mehr in der Segmentierung enthalten. Trotzdem sind immer noch Bereiche eingeschlossen, die nicht zum Bulbus gehören. Eine korrekte Segmentierung mittels Thresholding ist aufgrund der Eigenschaften des Bildmaterials nicht möglich.

3.8.3 Regionenorientierte Verfahren

Regionenorientierte Verfahren treffen die Entscheidung über die Identität eines Pixels im Kontext der Identitäten der umgebenden Pixel und/oder des bisher segmentierten Gesamtobjektes. In dieser Kategorie werden Region Growing [Cha95, Fun90, Jus97, San91], Clustering-Verfahren [Dek92, Xu97], texturbasierte Verfahren [Bou91, Dem91, Der87, Dun95, Fun90, Kas86; Kim96, Man91, Uns89] und adaptive Verfahren [Wel96] zusammengefaßt. Die in [Ahu96, Del96, Had90, Pav90] beschriebenen Ansätze versuchen die Vorteile von punkt- und regionenorientierten Verfahren zu kombinieren. Diese erweiterten Ansätze werden ebenfalls zu den regionenorientierten Verfahren gerechnet, da der Betrachtungsbereich nicht nur auf das aktuelle Pixel beschränkt bleibt.

Die Schwachstellen der regionenorientierten Verfahren stimmen mit denen der punktorientierten Klassifikationsverfahren weitgehend überein (siehe Abbildung 3.15). Die grundsätzliche Annahme besteht wiederum darin, daß sich die Grauwerte des Objektes von denen des umgebenden Gewebes unterscheiden. Außerdem wird davon ausgegangen, daß die Grauwerte innerhalb des Objektes weitgehend einheitlich sind.



Abbildung 3.15: Typische Ergebnisse des Region Growing

In der mittleren Abbildung sind die Grauwertvorgaben zu weit gefaßt, so daß der Algorithmus aus dem Bulbus wächst. Rechts: Die Schwellwerte sind zu klein, so daß im Randbereich die Segmentierung unpräzise ist. Die Vorderkammer bleibt bei beiden Schwellwerten unberücksichtigt. Sofern die Fehler korrigiert werden sollen, müßte auf die manuelle Segmentierung zurückgegriffen werden. Dieses Korrektur würde jedoch aller Voraussicht nach länger dauern, als wenn der Bulbus von vornherein manuell segmentiert worden wäre.

In spezifischen Anwendungsbereichen lassen sich mit den Regionenorientierten Verfahren zwar akzeptable Ergebnisse erzielen. In den meisten dieser Fälle müssen diese jedoch mit einem hohen Aufwand zum Einstellen der richtigen Schwellwerte erkaufte werden. Als Einschränkung bleibt in jedem Fall jedoch die Tatsache erhalten, daß das segmentierte Organ in sich homogen sein muß. Bei der Segmentierung des Bulbus bedeutet dies, daß die Segmentierung nicht über die Linse hinaus kommt (siehe Abbildung 3.15, rechts). Medizinisch gesehen gehört jedoch die Vorderkammer in jedem Fall zu dem Bulbus dazu.

3.8.4 Kantenorientierte Verfahren

Kantenorientierte Verfahren segmentieren Objekte, indem nach den Umrißlinien der Objekte gesucht wird. Reine Kantendetektionsverfahren [Bom90,Che89,Tan91,Tor90] werden in der aktuellen Literatur immer seltener behandelt. Der Hauptgrund hierfür besteht in der Erkenntnis, daß nur die wenigsten Organe in medizinischem Bildmaterial durchgehende Kanten zum umgebenden Gewebe aufweisen. In vielen Fällen muß mit weichen oder diffusen Grauwertübergängen gerechnet werden. In einigen Bilddaten sind aufgrund geringer Unterschiede in den bildgebenden Eigenschaften zumindest partiell keine Kanten zu erwarten (vergleiche Abbildung 3.16, links). Aus diesem Grund ist man zur modellbasierten Kantenerkennung übergegangen [Gos95, Han97, Mai96, She96]. Dabei zeichnen sich die Kanten nicht mehr nur durch Intensitäts- oder Texturunterschiede aus, sondern durch charakteristische Kantenprofile oder -formen.

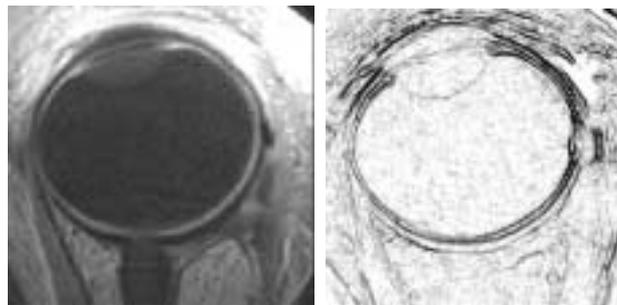


Abbildung 3.16: Typisches Eingangsbild für ein Kantensuchverfahren

Große Probleme bereiten die vielen Kanten, die nicht mit Organkanten übereinstimmen.

Ein weiteres Problem der Kantendetektion ist die Auswahl der richtigen Kanten aus den Hochpaßbildern (vergleiche Abbildung 3.16, rechts). Um den Kantensuchprozeß zu stabilisieren wurden für unterschiedliche Anwendungsbereiche spezialisierte Verfahren entwickelt. Bei dem von [Kut96] beschriebenen Ansatz werden beispielsweise zur Extraktion der Arterien in Angiographiebildern die Kanten durch Verwendung von Direction Fields verstärkt und detektiert. Dabei werden lediglich jene Kanten zusammengefügt, die sich gegenseitig gut ergänzen und eine für Arterien sinnvolle Form und Breite aufweisen. Ähnlich proprietäre Ansätze gibt es auch für verschiedenste andere Anwendungsbereiche.

Kantenorientierte Verfahren setzen voraus, daß sich Organe mit durchgängigen Kanten von dem umgebenden Gewebe abgrenzen. Außerdem setzen sie voraus, daß das Organ eine Kante zum umgebenden Gewebe besitzt. Probleme treten dann auf, wenn das Organ fließende oder weiche Organgrenzen hat oder zumindest partiell die Kanten weich oder diffus sind. Darüber hinaus ist dieser Ansatz sehr empfindlich gegenüber Rauschen und Artefakten, da diese zu einer deutlichen Verschlechterung der Kantenerkennung führen. Grundsätzlich muß bei medizinischem Bildmaterial davon ausgegangen werden, daß die Forderung nach durchgängigen, gleich strukturierten Kanten nicht erfüllt wird.

3.8.5 Transformationsverfahren

Transformationsverfahren ist keine gängige Gruppierung von Segmentierungsverfahren. Hierunter sollen jene Verfahren gruppiert werden, die eine Transformation der Daten in ein anderes Raum- oder Koordinatensystem vornehmen, um gesuchte Bildinhalte oder -bereiche leichter von umgebenden Bildbereichen abgrenzen zu können. Zu dieser Gruppe sind Verfahren wie Fourier-, Z- oder andere Transformationen [Sol92] zu zählen, ebenso wie der Scale-Space-Ansatz [Ayd96, Flo95, Hon84] und die Spatial/Spatial-Frequency Repräsentation [Ree90, Ree91]. Genaugenommen stellt diese Gruppe von Verfahren keine Segmentierung, sondern eine Vorverarbeitung dar, an die sich der eigentliche Interpretationsschritt, in der Regel eine punktorientierte Klassifikation, anschließt.

3.8.6 Modellbasierte Verfahren

Modellbasierte Verfahren lassen bei der Segmentierung detailliertes Wissen über das gesuchte Objekt mit einfließen. Dieses Wissen wird in Form von Modellen festgehalten, die an das aktuelle Bildmaterial angepaßt werden. Regionen-Matching-Verfahren [Lep97, Scl95, Ste92] gehen davon aus, daß Objekte eine spezifische globale Form und bestimmte Grauwerteigenschaften besitzen. Die Interpretation wird dadurch vorgenommen, daß das Modell im aktuellen Bildmaterial zunächst gesucht wird. Anschließend werden Position und Skalierung des Objektes an das aktuelle Bildmaterial angepaßt.

Andere Verfahren passen dreidimensionale Oberflächenmodelle an die Daten an [Bat97, Abo97, Ami97, Chr96, Han97, Hin97, Jai96, Jol96, Kau97, Sta96]. Durch die Verwendung von dreidimensionalen Modellen können in einzelnen Schichten fehlende Informationen kompensiert werden, indem die Parametrisierung des Modells anhand der zur Verfügung stehenden Informationen vorgenommen wird. Auf diese Weise können Organoberflächen vollständig rekonstruiert werden, selbst wenn Partialvolumeneffekte oder Metallartefakte Teile der Oberfläche überdecken oder nicht bestimmbar machen. Als Referenzen für die Anpassung der Modelle werden in der Regel Landmarken genutzt. Verfahren, die zweidimensionale Bilder interpretieren sollen, korrelieren statt dessen Segmentierungskonturen mit dem Bildmaterial, indem angepaßte Suchmechanismen nach bestimmten Bildelementen suchen [Gei95, Hei97, Jos97, Vos97, Wol90, Xin96].

In einigen Artikeln wird auch der in Abschnitt 3.8.9 detailliert beschriebene Snake-Ansatz zu der Gruppe der modellbasierten Segmentierungsverfahren gerechnet [Min96]. Hierzu gehören die in Abschnitt 3.8.9 detailliert beschriebenen zweidimensionalen Snake-Ansätze, aber auch der dreidimensionale, der sich lediglich durch die Definition der Oberfläche von den zweidimensionalen Verfahren unterscheidet. Bei dem dreidimensionalen Ansatz wird die Oberflächendarstellung durch Fourier-Repräsentationen oder durch Superquadratics erreicht [Min96]. Sowohl beim zweidimensionalen als auch beim dreidimensionalen Ansatz sind interne und externe Energie auf die Annäherung von möglichst starken und runden Kanten ausgerichtet. Ob in diesem Zusammenhang tatsächlich von einer modellbasierten Herangehensweise gesprochen werden sollte, ist nach Ansicht des Autors fraglich.

Die modellbasierte Segmentierung ist in den letzten Jahren intensiv untersucht worden, da sie durch die Integration von Vorwissen eine Stabilisierung der Segmentierungsergebnisse verspricht. Ein wichtiger Aspekt stellte in diesem Zusammenhang der Umgang mit fehlenden Bildinformationen dar. Immer wieder können im medizinischen Bildmaterial einige Bildteile nicht ausgewertet werden, weil durch Bildartefakte die Informationen lokal verzerrt oder gänzlich überlagert werden. Sofern im übrigen Bild genug Informationen zur Anpassung der Modelle enthalten sind, können durch die Verwendung von Modellen diese fehlenden Informationen kompensiert werden. Der Nachteil der bisherigen Ansätze besteht meist in einer aufwendigen Konfigurierung und Anpassung der Verfahren an sich verändernde Bildinhalte oder -eigenschaften. Darüber hinaus werden in der Regel als einzige Segmentierungsreferenzen eingegebene Landmarken genutzt. Eine Korrelation mit dem Bildmaterial wird durch die oben erwähnten Algorithmen

nicht durchgeführt. Darüber hinaus wird lediglich ein Modell der Segmentierung zugrunde gelegt. Dadurch kann es schnell zu Diskrepanzen zwischen dem Modell und dem tatsächlichen Inhalt des Bildmaterials kommen. Solche Diskrepanzen wirken sich direkt auf die Ergebnisqualität aus.

3.8.7 Wissensbasierte Verfahrenskombinationen

Wissensbasierte Verfahrenskombinationen sind Segmentierungsansätze, die aufgrund von detailliertem Wissen über das Bildmaterial spezifische Kombinationen von Segmentierungsverfahren aus den bereits beschriebenen Kategorien und unspezifische Manipulationsverfahren wie Morphologische Filter zu einer Verarbeitungspipeline zusammenstellen. Die Konzeptionen der Verfahren basieren auf dem Wissen über die Bildeigenschaften [Bru93, Fis97, Li93, Lim89], über spezifische Grauwerteigenschaften [Cli90, Joh93, Man92, Par96, Pul93, San97, Son96, Wel96] oder anatomische Formeigenschaften [Cha96, Gra97, Kut96, Son96].

Ebenfalls zu dieser Kategorie zählen die multimodale Auswertung von Kernspinbildern [Ger92] oder Ansätze der Gestaltanalyse und Perceptual Organization [Moh92, Sab93]. Die Bezeichnung „wissensbasierte Verfahrenskombination“ leitet sich ursprünglich aus der in der Literatur verwendeten Bezeichnung „Wissensbasierte Verfahren“ ab. Meiner Meinung nach ist die gängige Bezeichnung unpräzise, da alle Segmentierungsverfahren auf Wissen oder Annahmen über das Bildmaterial aufsetzen. Durch die Verwendung des Begriffs „Verfahrenskombination“ wird die von den bisher beschriebenen Kategorien abweichende Vorgehensweise verdeutlicht. Sie besteht in der Ausnutzung von Vorwissen über das Bildmaterial, indem verschiedene Verfahren in einer angepaßten Sequenz kombiniert angewandt werden.

Da diese Verfahren von sehr spezifischem Bildmaterial mit klar definierten Bildeigenschaften und Bildinhalten ausgehen, ist eine Segmentierung von Bildmaterial, das diesen Annahmen nicht entspricht, in der Regel nicht möglich. Da bei medizinischem Bildmaterial immer mit einer Toleranz sowohl in den Bildeigenschaften als auch in der individuellen Anatomie des Patienten zu rechnen ist, stoßen diese Verfahren selbst für das gegebene Bildmaterial schnell an die Grenzen der Anwendbarkeit.

3.8.8 Neuronale Netze

Neuronale Netze, in der anglistischen Literatur oft als künstliche Neuronale Netze (artificial neural networks – ANN) bezeichnet, versuchen die logischen und kombinatorischen Fähigkeiten von biologischen Neuronen nachzuahmen. Das menschliche Gehirn und Zentrale Nervensystem ist ein gutes Beispiel für ein funktionales Netz von vergleichsweise einfachen Funktionseinheiten, den Neuronen, die durch ihr Zusammenspiel in der Lage sind, innerhalb von kürzester Zeit komplizierte Problemstellungen zu lösen. Durch mathematische Nachbildung von funktionalen Einheiten wie Synapsen, Dendriten, Nervenkörper und Axon werden die Fähigkeiten von Neuronen nachgebildet. Die Einzelneuronen werden in Netzen zusammengefaßt und anschließend durch eine Trainingsmenge „angelernt“. Die Zielsetzungen dieses Anlernens ist einerseits die möglichst gute Abbildung der Trainingseingangswerte auf die Trainingsausgangswerte und andererseits eine sinnvolle Interpolation der möglichen Ausgabewerte auf die erwarteten Eingabewerte. In dieser Funktionsweise ähnelt der Ansatz neuronaler Netze dem statistischen Ansatz der punktorientierten Klassifikationsverfahren. Da die Ergebnisse neuronaler Netze in der Regel nicht-lineare Approximationen der Lösungsfunktion produzieren, stellen sie eine Erweiterung gegenüber den herkömmlichen statistischen Verfahren dar.

Neuronale Netze werden in der Medizin in erster Linie als Klassifikatoren eingesetzt. Die Ansätze zur Realisierung von Klassifikatoren mittels neuronaler Netze sind so vielseitig, daß ein kurzer Überblick in wenigen Worten nicht möglich ist. Diese Ansätze unterscheiden sich in der Regel durch den Aufbau des Netzes und durch das Lernverfahren [Cha97, Fuk97, Law97, Lee97, Lee97b, Leu91, Lin97, Rid97, Set97, Vou97].

In der aktuellen Literatur zur medizinischen Bildverarbeitung basieren nur wenige Verfahren auf der Verwendung von Neuronalen Netzen. Ein Grund hierfür mag sein, daß insgesamt von dem Ansatz der globalen Segmentierung von Bildern auf der Basis der Grauwerte Abstand genommen wurde. Der Anwendungsbereich von neuronalen Klassifikatoren in der medizinischen Bildverarbeitung blieb daher vor allem auf zwei Themenbereiche begrenzt: der Detektion von Tumorkanten [Zhu97a, Zhu97b] und der multimodalen Segmentierung von CT und MRT-Daten [Gla98, Kar99, Tou92].

3.8.9 Snake

Der Snake-Ansatz wurde von Kass [Kas87] und Terzopoulos [Ter88] entwickelt. Bei diesem Ansatz wird eine geschlossene Kontur in einem Bild plaziert, die sich anschließend unter der Einwirkung von zuvor definierten Kräften in Form und Position verändert. Das Verfahren läßt sich mit einem Gummi vergleichen, das im Bild plaziert wird und sich dann um das gesuchte Objekt zusammenzieht. Das Ziel ist es, daß sich die Kontur so lange zusammenzieht, bis sie an den Kanten des gesuchten Objektes angelangt ist und die Objektgrenzen auf dem gesamten Umfang beschreibt.

Der Ursprung dieses Segmentierungsverfahrens liegt bei den kantenbasierten Verfahren. Ein grundsätzliches Problem, das die normalen kantenbasierten Verfahren nicht lösen konnten, ist die Tatsache, daß die kantenbasierte Segmentierung eines Organs deutlich erschwert wird, wenn die Kanten des Organs zur Umgebung nicht durchgehend ausgeprägt sind. Bei medizinischem Bildmaterial ist ein zumindest teilweises Verschwinden der Organkanten eher die Regel als die Ausnahme. Durch die Verwendung einer geschlossenen Kontur, die sich als ganzes an die Kanten im Bild heranbewegt wird, können beim Snake-Ansatz fehlende Kantensegmente problemlos überbrückt werden.

Die Bewegung der Kontur wird beim Snake-Verfahren durch zwei Kräfte koordiniert (siehe Abbildung 3.17): die interne Kraft und die externe Kraft, die sich jeweils aus der internen und externen Energie ableiten. Die interne Kraft leitet sich von Form und Größe der Kontur ab und versucht eine möglichst runde und kleine Kontur anzustreben. Die externe Kraft bezieht das Bildmaterial ein, indem sie die Kontur in Richtung zuvor detektierter Kanten im Bildmaterial zieht. Die Kontur bewegt sich iterativ so lange, bis sie das sogenannte globale Minimum erreicht hat, das als die gewünschte Segmentierung angesehen wird.

Das Snake-Verfahren ist ein weit verbreiteter Ansatz, der zur Segmentierung [Cha96, Fal97, Hon97b, Ley93, Lob95, Sch97, Tag97, Xu97a, Xu97b, Yez97, Zhu96], Kantendetektion [Kas87], Formmodellierung [Ter88] und Bewegungsverfolgung [Ter92] genutzt wird.

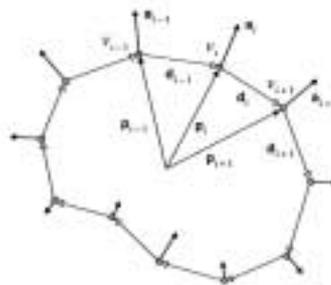


Abbildung 3.17: Die Snake-Kontur V mit den Einzelpunkten V_i

Die Kräfte a_i wirken auf die jeweiligen Einzelpunkte V_i .

Der Snake-Ansatz wird im weiteren Verlauf dieser Arbeit in einer modifizierten Form wieder aufgegriffen werden. Aus diesem Grund wird dieser Segmentierungsansatz in den folgenden Abschnitten detailliert beschrieben.

3.8.9.1 Interne Energie

Die interne Energie versucht eine möglichst weiche Snake-Kontur herzustellen. Sie kann in die Komponenten der ersten und zweiten Ordnung unterteilt werden. Der Anteil erster Ordnung verhindert wie ein Gummi das Ausweiten der Kontur. Sie bestimmt die „Elastizität“ der Snake-Kontur. Der Anteil zweiter Ordnung verhindert die Beugung der Kontur und beschreibt somit die „Rigidität“.

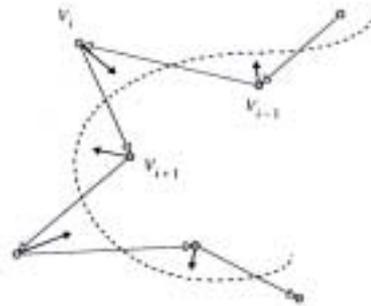


Abbildung 3.18: Veranschaulichung der internen Kraft

Die Zielsetzung der internen Energie besteht vor allem darin, eine möglichst runde und kleine Kontur herzustellen.

Die Formel für die Elastizität lautet: Sei \mathbf{v} eine geschlossene Kontur und $\mathbf{v}(s) = [x(s), y(s)]$, $s \in [0,1]$. Dann berechnet sich die interne Energie zu:

$$E_{\text{int},1} = \int_{s=0}^1 |\mathbf{v}'(s)|^2 ds \quad (1)$$

Dabei ist $\mathbf{v}'(s)$ die erste Ableitung von \mathbf{v} nach den Raumkoordinaten x und y . Beim Übergang in die diskreten Raumkoordinaten von digitalisiertem Bildmaterial verändert sich die Formel leicht. Sei \mathbf{v} eine diskrete geschlossene Kontur mit n Punkten und $\mathbf{v}_i = [x_i, y_i]$, $0 \leq i \leq n$. Dann berechnet sich die Elastizität zu:

$$E_{\text{int},1} = \sum_{i=0}^{n-1} |\mathbf{v}_{i+1} - \mathbf{v}_i|^2 \quad (2)$$

Die Rigidität gilt als die interne Energie zweiter Ordnung und ist folgendermaßen definiert:

$$E_{\text{int},2} = \int_{s=0}^1 |\mathbf{v}''(s)|^2 ds \quad (3)$$

Dabei ist $\mathbf{v}''(s)$ die zweite Ableitung von \mathbf{v} nach den Raumkoordinaten x und y . Für diskrete Raumkoordinaten gilt die folgende Formel:

$$E_{\text{int},2} = \sum_{i=0}^{n-1} |v_{i+1} - 2v_i + v_{i-1}|^2 \quad (4)$$

dabei ist $|\cdot|$ die Standard Euklidische Norm. Die gesamte interne Energie wird dementsprechend berechnet wir folgt:

$$E_{\text{int}} = \int_{s=0}^1 w_1 |v'(s)|^2 + w_2 |v''(s)|^2 ds \quad (5)$$

für kontinuierliche Raumkoordinaten, und

$$E_{\text{int}} = \sum_{i=0}^{n-1} w_1 |v_{i+1} - v_i|^2 + w_2 |v_{i+1} - 2v_i + v_{i-1}|^2 \quad (6)$$

für diskrete Raumkoordinaten.

3.8.9.2 Problem der internen Energiedefinition

Bei der Definition der internen Energie wird davon ausgegangen, daß das zu segmentierende Objekt nahezu rund ist. Die überwiegende Mehrzahl der Objekte ist jedoch nicht rund, sondern weist deutlich davon abweichende Formen auf, die gerade im medizinischem Bildmaterial in ihrer Individualität wichtige Informationen in sich bergen (vergleiche Abbildung 3.19).

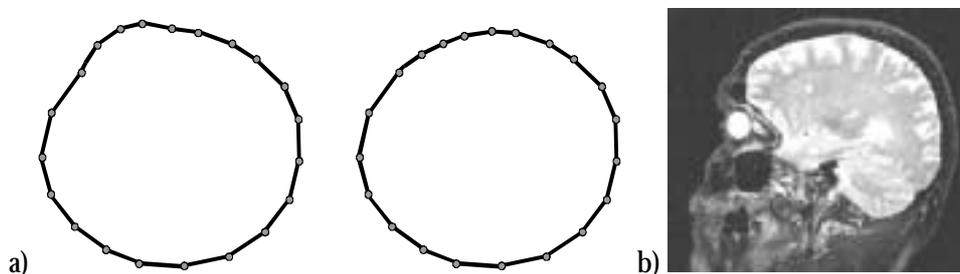


Abbildung 3.19: Schwachstellen bei der Definition der internen Energie

Die Abbildungen zeigen zwei Beispiele für die Tatsache, daß die gesuchten Konturen in der Regel nicht rund sind. a) Wenngleich das Auge weitgehend rund ist, stellt die Ausbuchtung am oberen Ende, die der Kornea entspricht, eine wichtige anatomische Struktur dar. Die korrekte Segmentierung eines Auges könnte beispielsweise wie die linke Kontur aussehen. Die interne Energie des Snake-Verfahrens würde eine vollständig runde Kontur anstreben, wie sie rechts daneben dargestellt ist. Die Kornea würde demnach fälschlicherweise „ausgebügelt“ werden. b) Noch offensichtlicher wird der Fehler bei der Annahme von runden Konturen bei der Segmentierung des Gehirns. Soll die graue Gehirnmasse von der weißen getrennt werden, so sind die unterschiedlichen Sulci von großer Bedeutung. Der Snake-Ansatz würde versuchen, die Sulci abzurunden.

Durch die Annahme von runden Konturen weist die interne Energie eine falsche Zieldefinition auf. Das bedeutet, daß selbst unter optimalen Voraussetzungen die interne Energie ein in jeder Hinsicht falsches Ergebnis produzieren würde.

3.8.9.3 Externe Energie

Die externe Energie paßt die Kontur an das Bildmaterial an, indem es die Konturpunkte unter Verwendung von Potentialkurven zu bestimmten Bildmerkmalen zieht. In der Regel wurden in den aktuellen Verfahren als Merkmal die im Bildmaterial vorhandenen Kanten verwendet. Die Verschiebungsvektoren

der Potentialfelder sind auf die lokal stärkste Kante ausgerichtet und verschoben dementsprechend die Konturpunkte zu den Kanten hin.

Abbildung 3.20 zeigt beispielhaft das Potentialfeld für eine U-förmige Figur. Die Potentialpfeile zeigen zu den Kanten. Ein Punkt, der sich direkt zwischen U-Bögen befindet, besitzt einen Verschiebungsvektor mit der Länge null, da sich die entgegengesetzten Vektoren annullieren. Bereits bei dieser einfachen Figur wird die Schwäche der externen Energiedefinition offensichtlich.

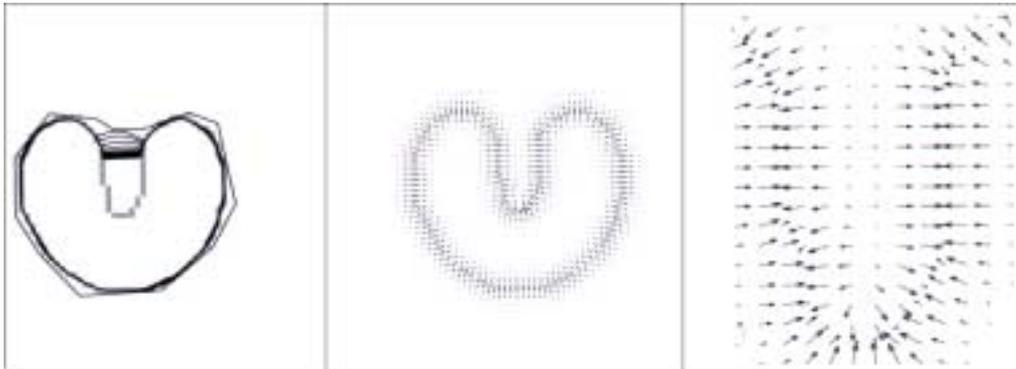


Abbildung 3.20: Potentialkurven bei dem Snake-Ansatz

Die Abbildungen zeigen beispielhaft die Potentialkurve für einen U-förmigen Körper. Die unterschiedlichen Längen der Pfeile veranschaulichen die Stärke der Kraft, mit der die Kontur in die jeweilige Richtung gezogen wird. Die sehr kurzen Pfeile in der Mitte des U's veranschaulichen, warum die Snake-Kontur das U nicht richtig segmentiert.

Die mathematische Definition der externen Energie lautet [Kas87, Ter88]:

$$E_{ext}^{(1)} = -|\nabla I(x, y)|^2 \quad (7)$$

Da der reine Kantendetektor sehr empfindlich auf hochfrequente Signale reagiert, wird zur Stabilisierung der externen Energie vielfach zuvor ein Tiefpaß angewandt. Die Formel für die so veränderte externe Energie lautet:

$$E_{ext}^{(2)} = -|\nabla G_\sigma * I(x, y)|^2 \quad (8)$$

wobei G_σ der Gauß-Weichzeichnungsfilter mit der Standardabweichung σ ist. Im diskreten Bereich werden zu Berechnung Hochpaßfilter angewandt. Im einfachsten Fall kann der LaPlace-Filter für die Berechnung der externen Energie genutzt werden:

$$E_{ext}(x, y) = h * I(x, y) \quad (9)$$

$$h = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix} \quad (10)$$

Die Ergebnisse des LaPlace-Operators sind oft von unzureichender Qualität, da dieser Filter sehr empfindlich gegenüber Rauschen reagiert. Andere Kantenoperatoren, wie beispielsweise der Canny-

Operator, machen einen größeren Rechenaufwand erforderlich, weisen dafür jedoch bessere Eigenschaften bezüglich der Kantenerkennung auf.

3.8.9.4 Probleme der externen Energiedefinition

Die Definition der externen Energie basiert auf der Annahme, daß Objekt und umgebendes Gewebe durch eine deutliche Kante voneinander abgegrenzt werden. Oft ist diese Annahme nur zum Teil erfüllt oder es treten zusätzliche Schwierigkeiten auf. Nicht immer werden Organe durch abrupte Kanten oder Grauwertsprünge vom umgebenden Gewebe abgegrenzt. Häufig weist medizinisches Bildmaterial weiche oder kontinuierliche Grauwertübergänge auf, weil Gewebeübergänge diffus sind oder weil durch Partialvolumeneffekte die Kanten verwischt werden. Darüber hinaus wird beispielsweise eine korrekte Erkennung der richtigen Kante bei der Segmentierung des Bulbus dadurch erschwert, daß direkt neben der eigentlichen Kante zusätzlich weitere Kanten liegen, die gegebenenfalls sogar stärker ausgeprägt sind als die eigentliche Organkante (siehe Abbildung 3.21). Die Folge sind Fehlpositionierungen der Segmentierungskontur und das Hin-und-her-Wechseln zwischen den anatomisch unterschiedlichen Kanten.

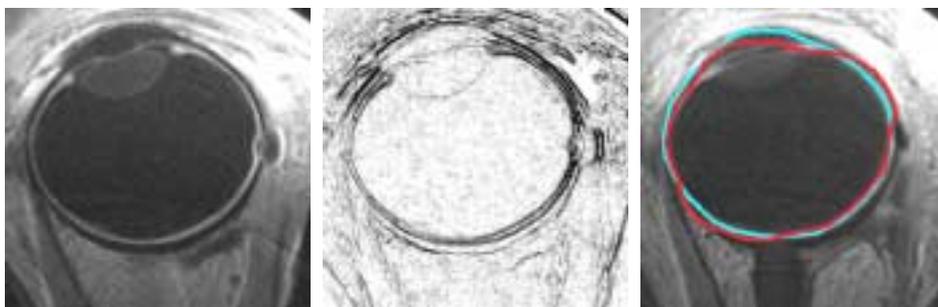


Abbildung 3.21: Problemstellung bei der Erkennung der Organkontur durch die externe Energie

Links: T1-Aufnahme des Auges; Mitte: Hochpaßbild der Aufnahme. Es können drei parallel verlaufende Kanten erkannt werden. Die eine beschreibt die Innenkante der Sklera, die zweite die Außenkante der Sklera, die dritte Kante wird durch den Chemical Shift zwischen Sklera und umgebenden Fettgewebe erzeugt. Rechts: Die gewünschte Segmentierung ist blau dargestellt. Ein typisches Ergebnis für den Snake-Ansatz ist rot dargestellt. Die Kontur wechselt zwischen den einzelnen Kanten hin und her, da kein Kriterium vorgibt, welche der Kanten für die Segmentierung genutzt werden soll. Die Kontur verläuft jeweils dort, wo im Bildmaterial jeweils die Kante am stärksten ausgeprägt ist.

3.8.9.5 Gesamtenergie

Bei der Anwendung des Snake-Verfahrens wirken interne und externe Energie gleichzeitig. Die Gesamtenergie berechnet sich durch die folgende Formel:

$$E_{Snake} = w_{ext} E_{ext} + w_{int,1} E_{int,1} + w_{int,2} E_{int,2} \quad (11)$$

$$E_{Snake} = \int_{s=0}^1 w_{ext} E_{ext}(v(s)) + w_{int,1} |v'(s)|^2 + w_{int,2} |v''(s)|^2 ds \quad (12)$$

Die Energien können mit den Faktoren w_{ext} , $w_{int,1}$ und $w_{int,2}$ gewichtet werden, um auf Bild- und Objekteigenschaften eingehen zu können. Die Snake-Energie hat ein Minimum erreicht, wenn die Euler-Lagrangsche Gleichung erfüllt ist (siehe [Bro89], Beispiel Seite 388):

$$w_{int,1} v''(s) - w_{int,2} v''''(s) - w_{ext} \nabla E_{ext} = 0 \quad (13)$$

Das bedeutet, daß bei dieser Kontur die internen und externen Kräfte im Gleichgewicht stehen:

$$F_{\text{int}} + F_{\text{ext}} = 0 \quad (14)$$

wobei

$$F_{\text{int}} = w_{\text{int},1} v''(s) - w_{\text{int},2} v''''(s) \quad (15)$$

und

$$F_{\text{ext}} = -w_{\text{ext}} \nabla E_{\text{ext}} \quad (16)$$

gilt. Um eine Lösung zu finden, ist v eine Funktion über die Zeit t und der Kontur, abhängig von s : $v(t,s)$. Die partielle Ableitung nach t ergibt dann die folgende Gleichung:

$$x_t(s,t) = w_{\text{int},1} v''(s,t) - w_{\text{int},2} v''''(s,t) - w_{\text{ext}} \nabla E_{\text{ext}}. \quad (17)$$

Wenn die Lösung $x(s,t)$ sich stabilisiert, verschwindet der Term $x_t(s,t)$. Eine numerische Lösung zu dem Problem wird in der Praxis durch die Diskretisierung der Aufgabenstellung und iterative Anpassung der Kontur erreicht. Um die Suche nach dem globalen Minimum zu stabilisieren werden entweder die Einzelparameter der Gesamtenergie und/oder x_t durch eine Skalar-Multiplikation modifiziert.

3.8.9.6 Modifikationen des Snake-Ansatzes

Aufgrund der oben beschriebenen Schwächen konnten mit den in der aktuellen Literatur beschriebenen Snake-Ansätzen noch keine zufriedenstellenden Segmentierungsergebnisse erzielt werden. In einer Reihe von Veröffentlichungen wurde der Snake-Ansatz durch verschiedene Ergänzungen modifiziert, um diese Schwächen für spezifisches Bildmaterial zu kompensieren.

Um die Segmentierung von Zellen zu verbessern stellte [Gun97] eine Snake-Verfahren vor, bei dem die Zelle gleichzeitig durch zwei Konturen angenähert wird. Die eine Kontur nähert sich von innen, die zweite von außen an die gesuchte Kontur an. Um die Bewegung der Konturpunkte besser kontrollieren zu können, stellt [Tag97] einen Ansatz vor, bei dem mittels orthogonaler Kurven die Freiheitsgrade für die Snake-Kontur koordiniert wird. Einen grundsätzlichen Durchbruch konnte man jedoch auch mit diesen Ansätzen nicht erreichen.

3.8.10 Zusammenfassung

Die Betrachtungen der vorangegangenen Abschnitte über die verschiedenen Segmentierungsverfahren verdeutlichen, daß bei aktuellen Segmentierungsansätzen offensichtlich die folgenden Nachteile in Kauf genommen werden müssen:

1. Die Segmentierung von dreidimensionalen Bilddaten ist arbeitsintensiv, weil die Parametrisierung des Algorithmus, die Durchführung des Algorithmus oder die notwendige Nachbearbeitung der primären Segmentierungsergebnisse aufwendig ist.
2. Die Verfahren lassen sich nicht ohne weiteres an unterschiedliche Organe oder Bildmodi anpassen, weil entweder die Anpassung aufwendig oder aufgrund der Verfahrenskonzeption gar nicht möglich ist.

Aktuelle Segmentierungsansätze weisen damit zwei grundlegende Probleme auf, die ihre Anwendung in der medizinischen Praxis behindern: Einerseits können aufgrund der hohen Kosten für die Bildverarbeitung die verbesserten diagnostischen Verfahren nicht für Routinefälle eingesetzt werden. Andererseits müssen viele proprietäre Bildverarbeitungsansätze entwickelt werden, die bei einer Veränderung der Bild-
daten wieder überarbeitet oder gar nicht mehr genutzt werden können. Die Ursachen dafür, daß die aktuellen Segmentierungsverfahren es bisher nicht geschafft haben, den Arbeitsaufwand zu reduzieren und gleichzeitig auf spezifische Anwendungsbereiche beschränkt bleiben, sehe ich vor allem in den folgenden Punkten:

- Es wird nur ein geringer Anteil des zur Verfügung stehenden Vorwissens über das Bildmaterial und die anatomischen Gegebenheiten bei der Konzeption und Durchführung der Segmentierungsalgorithmen ausgenutzt.
- Viele der Verfahren wenden globale Annahmen über die Darstellungseigenschaften des Bildmaterials zur Segmentierung der Organe an. Aufgrund der bereits beschriebenen Darstellungseigenschaften sind diese Annahmen falsch oder zumindest unpräzise.
- Die Segmentierungsverfahren wenden die Annahmen und das Vorwissen über das Bildmaterial implizit durch die Konzeption des Verfahrens an. Bei der Anwendung des Region Growing wird beispielsweise vorausgesetzt, daß die zu segmentierenden Objekte homogene Grauwerteigenschaften aufweisen. Auf Bildmaterial oder Organe mit davon abweichenden Eigenschaften kann das Region Growing nicht angewandt werden.
- Es existiert keine Schnittstelle zwischen dem Vorwissen über das Bildmaterial und dem Segmentierungsverfahren. Dadurch kann nicht durch Anpassung des Vorwissens das Segmentierungsverfahren an Änderungen im Bildmaterial angepaßt werden.

Korrespondierend zu den Erkenntnissen angrenzender Forschungsbereiche wie der Künstlichen Intelligenz deuten die aufgezählten Probleme nach meiner Meinung an, daß der entscheidende Faktor bei der Bildverarbeitung das Vorwissen über das Bildmaterial ist. Von Relevanz ist dabei der Umfang und die Anwendung des zur Verfügung stehenden Vorwissens. Ein Bildverarbeitungsansatz, der die Probleme vorangegangener Ansätze vermeiden soll, muß vor allem die folgenden Fragestellungen intensiv untersuchen:

- Welches Wissen steht zur Verfügung?
- Wie läßt sich dieses Wissen kodieren, damit der Segmentierungsalgorithmus dieses Wissen verstehen kann?
- Wie wird das Wissen ausgewählt, das auf den aktuellen Fall am besten angewandt werden kann?
- Wie wird das Wissen angewandt?
- Wieviel Wissen ist notwendig, um das Bildmaterial, das in der klinischen Routine anfällt, bearbeiten zu können?

Die Antworten auf diese Fragen soll der in den nachfolgenden Kapiteln beschriebene modellbasierte Bildverarbeitungsansatz liefern.

3.9 Oberflächenrekonstruktion

Das Ergebnis einer Segmentierung besteht üblicherweise aus den zweidimensionalen Konturen der segmentierten Organe. Um die segmentierten Organe dreidimensional weiter verarbeiten zu können, beispielsweise durch eine dreidimensionale Visualisierung, Interpolation oder Registrierung von mehreren

Datensätzen, müssen die zweidimensionalen Konturen zu dreidimensionalen Oberflächen zusammengefaßt werden. Dieser Vorgang wird als Oberflächenrekonstruktion bezeichnet.

Für die gegebene medizinische Aufgabenstellung des zugrunde liegenden Projektes müssen ausgehend von den zweidimensionalen Segmentierungen der CT- und MRT-Daten die Oberflächen der Organe wie Bulbus und Linse rekonstruiert werden.

Die grundsätzliche Aufgabenstellung der Verfahren zur Oberflächenrekonstruktion besteht in der Korrelation der Konturverläufe von angrenzenden Schichten. Die Schwierigkeit besteht darin, aus der Form und Anordnung der Konturverläufe möglichst sinnvolle Verbindungen der einzelnen Konturen zu berechnen. Gelöst wird diese Problematik durch eine Gruppe von Verfahren, die mit dem Begriff Triangulationsverfahren bezeichnet werden, weil sie zwischen den angrenzenden Konturen Dreiecke zur Definition der Oberfläche konstruieren.

Es werden Dreiecke und nicht andere geometrische Formen wie Rechtecke oder Sechsecke konstruiert, da Dreiecke für beliebige Punktekongstellationen definiert werden können. Bei der Oberflächenrekonstruktion werden die Punkte der einen Kontur durch Dreiecke mit den Punkten der anliegenden Kontur verbunden. Wichtig ist dabei die Vorgehensweise, nach der die Konstruktion dieser Dreiecke und damit die Korrelation der Konturen untereinander vonstatten geht.

Von den Triangulationsverfahren sind Ansätze wie Marching Cubes abzugrenzen, die aus Voxeldaten Oberflächen erstellen. Dabei werden nach vorgegebenen Konstruktionsprinzipien binäre Voxelobjekte in den Randbereichen in Polygonoberflächen überführt. Bei der Korrelation von Konturen müssen hingegen zweidimensionale Konturen miteinander korreliert werden, um eine möglichst realitätsnahe und sinnvolle Rekonstruktion der Oberflächen zu erreichen.

Eine der Schwierigkeiten bei der Konzeption von Triangulationsverfahren besteht darin, daß es keine objektiv korrekte Lösung gibt, zwei Konturen inhaltlich miteinander zu korrelieren [Hal90]. Grundlage für die Aussage ist die Annahme, daß keine umfassenden Informationen für die Relation zwischen zwei Konturen und dem von ihnen aufgespannten Volumen existieren. Zwischen den Schichten werden in jedem Fall Daten hinzugefügt, die nicht der Realität entsprechen, sondern auf der Basis von Plausibilitätsüberlegungen und aus der Notwendigkeit heraus erfunden sind. Jedes Verfahren versucht dementsprechend mittels Heuristiken oder regulierenden Regeln eine Korrelation zwischen den Konturen zu ermitteln, bei dem das Ergebnis von vornherein in eine gewünschte Ergebnisqualität gelenkt wird. Die tatsächliche Korrektheit der Ergebnisse hängt somit davon ab, wie stark erwartete und tatsächliche Objekteigenschaften übereinstimmen.

Die Heuristiken für das Finden von Korrespondenzen können auf Kontur-, Volumen- oder anderen Eigenschaften aufsetzen. In vielen Fällen werden durch die Konzeption des Verfahrens implizit Vorgaben zu den globalen Korrelationsstrategien vorgenommen, die die Ergebnisqualität von vornherein manifestieren. Ein Beispiel hierfür ist die Frage, ob Objektkonturen zwischen zwei Schichten gegeneinander rotiert sein können oder ob nur vertikale Konturdeformationen möglich sind. In der Literatur werden drei unterschiedliche Ansätze zur Oberflächenrekonstruktion beschrieben:

- Triangulation mittels Optimierung,
- Dreidimensionale Delaunay-Triangulation und
- Topologiebasierte Triangulation.

Bei dem ersten Verfahren wird zwischen jeweils zwei Konturen eine Triangulation durchgeführt, die auf der Basis von Flächen- oder Volumenparametern optimiert wird [Fuc77,Gan82]. Bei dem zweiten wird eine dreidimensionale, durch Voronoi-Diagramme koordinierte Delaunay-Triangulation durchgeführt

[Boi92,Lis94]. Bei dem dritten Ansatz werden topologische Informationen über das Objekt benutzt, um die Triangulation zwischen den beiden Konturen auszurichten [Anj87]. In den nachfolgenden Abschnitten werden die drei Verfahren zur Oberflächentriangulation detailliert beschrieben.

3.9.1 Triangulation mittels Optimierung

Bei der Triangulation von planaren Konturen mittels Optimierung [Fuc77] wird im ersten Schritt eine tabellarische Graphenrepräsentation der Konturen erzeugt. Auf den beiden Achsen der Tabelle werden die Punkte beider Konturen aufgetragen (siehe Abbildung 3.22). In dem Graphen werden die Korrespondenzen, die eine Triangulation zwischen den beiden Konturen aufspannen, als Vektoren zwischen den Knotenpunkten der Tabelle dargestellt. Durch Verbinden der aneinander liegenden Korrespondenzen wird ein Weg innerhalb des Graphen erzeugt, wobei jeder Vektor ein Dreieck darstellt. Die zwei Konturen sind vollständig miteinander trianguliert, wenn der Vektorpfad sämtliche Spalten und Zeilen mindestens einmal gekreuzt hat. Der Weg durch den gesamten Graphen definiert die Dreiecke der Oberfläche zwischen den beiden Konturen. Beim Sprung vom letzten zum ersten Punkt der Konturen kann der Graph auch in mehrere Untergraphen zerlegt werden.

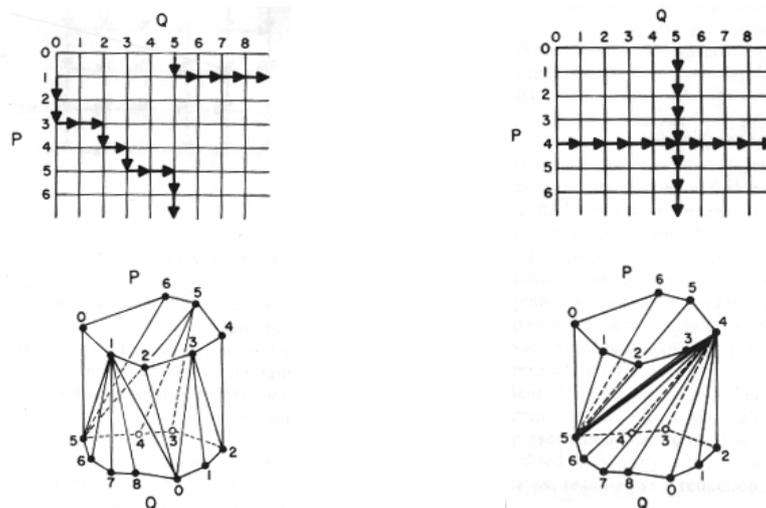


Abbildung 3.22: Veranschaulichung der Triangulation mittels Optimierung

Die Konturkorrespondenzen werden anhand eines tabellarischen Graphen nach [Fuc77] bestimmt. Graphiken entnommen aus [Fuc77].

Die Qualität eines Graphen und damit der Triangulation wird durch eine Kostenfunktion innerhalb der tabellarischen Repräsentation bestimmt. Jedem Dreieck wird ein Kostenwert mittels der Kostenfunktion zugeordnet. Diese Kostenfunktion berechnet sich aus der Fläche des Dreiecks. Durch die Summation der Einzelkosten eines Weges durch den Graphen werden die Gesamtkosten der Triangulation ermittelt. Um die optimale Triangulation zu finden wird zunächst für alle Knotenpunkte des Graphen der Kostenwert berechnet. Anschließend wird durch ein iteratives Suchverfahren ein möglichst kostengünstiger Pfad berechnet.

Der von Ganapathy verfolgte Ansatz stimmt mit dem von Fuchs überein. Einziger Unterschied ist eine Kostenfunktion, die nicht über die Oberfläche, sondern über die von den Dreiecken und einem Punkt innerhalb des Objektes eingeschlossenen Volumens definiert wird.

Grundlegender Nachteil beider Ansätze ist die Tatsache, daß keine Verzweigungen berücksichtigt werden können. Es kann immer nur eine Kontur mit einer anderen korreliert werden. Darüber hinaus läßt

sich durch die abstrakte Zuweisungsmethodik die Ergebnisqualität nur vergleichsweise schlecht kontrollieren. Würden beispielsweise zwei Objekte mit langen, geraden und parallelen Kanten miteinander korreliert werden, könnten die Dreiecke in diesen Bereichen beliebig angelegt werden, da weder das Volumen noch die Oberfläche bei gescherten Dreiecken variieren. Darüber hinaus können keine Aussagen gemacht werden, inwieweit geometrische Korrespondenzen zwischen den Konturen berücksichtigt werden, da die Verbindungen zwischen den Konturen implizit konstruiert werden.

3.9.2 3D Delaunay-Triangulation

Bei der Delaunay-Triangulation werden zunächst die Konturen in Dreiecke unterteilt. Anschließend werden basierend auf diesen Dreiecken Tetraeder konstruiert, indem die Verbindungen zwischen den Konturen der unterschiedlichen Schichten als Spitzen für die Tetraeder genutzt werden [Boi92, Lis94]. Durch diese Tetraeder wird aus den Konturen eine voxelähnliche Objektdarstellung berechnet, von der dann die Oberfläche extrahiert wird. Die erste Unterteilung der Kontur in Dreiecke wird durch Voronoi-Diagramme und Konsistenzprüfungen bestimmt (vergleiche Abbildung 3.23). Dabei muß zunächst eine Triangulation erzeugt werden, die alle tatsächlichen Kanten des Objektes beinhaltet. Bei nicht konvexen Konturverläufen müssen an einzelnen Stellen Punkte hinzugefügt werden, um dies zu erreichen (siehe Abbildung 3.24).

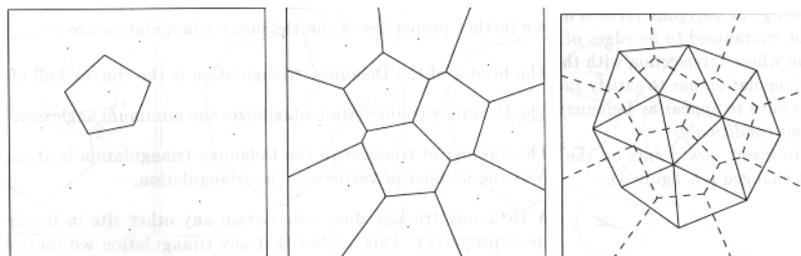


Abbildung 3.23: Berechnung eines Voronoi-Diagramms

Die Abbildungen zeigen die Berechnung der Dreiecke (rechts) für ein gegebenens Polygon (links) durch das Voronoi-Diagramm (Mitte). Das Ergebnis der Verarbeitungsschritte ist ein Netzwerk von Dreiecken, das für die Berechnung der Tetraeder und der Oberfläche herangezogen werden kann. Grafik aus [Boi92] entnommen.

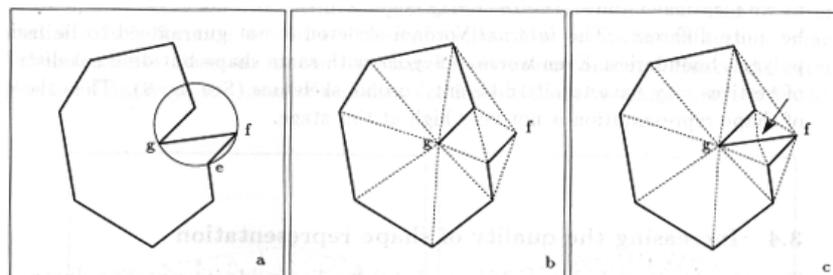


Abbildung 3.24: Behandlung von nicht-konvexen Polygonen

Bei nicht-konvexen Objekten müssen zusätzliche Punkte eingefügt werden, um eine korrekte Delaunay-Triangulation berechnen zu können. Grafik aus [Boi92] entnommen.

Bei der Konstruktion des Voronoi-Diagramms wird zwischen externen und internen Skeletten unterschieden. Diese Differenzierung ist notwendig, um das Aufteilen des zu rekonstruierenden Objektes in Unterarme zu koordinieren. Externe Voronoi-Skelette trennen die Unterarme untereinander. Interne Voronoi-Skelette ermitteln eine interne Skelettstruktur, die auf die anliegende Schicht übertragen wird (siehe

Abbildung 3.25). Die internen Voronoi-Skelette müssen immer innerhalb des Objektes liegen. Bei nicht konvexen Objekten müssen wiederum Punkte eingefügt werden, falls diese Bedingung nicht eingehalten wird (siehe Abbildung 3.26).

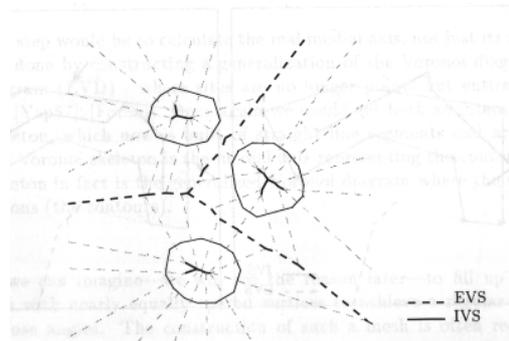


Abbildung 3.25: Behandlung von mehreren Objektarmen

Durch die externen Voronoi-Skelette werden einzelne Arme eines Objektes voneinander getrennt. Durch interne Voronoi-Skelette wird für jeden Arm eine interne Grundstruktur angelegt, die für die Konstruktion der Tetraeder genutzt wird. Grafik aus [Boi92] entnommen.

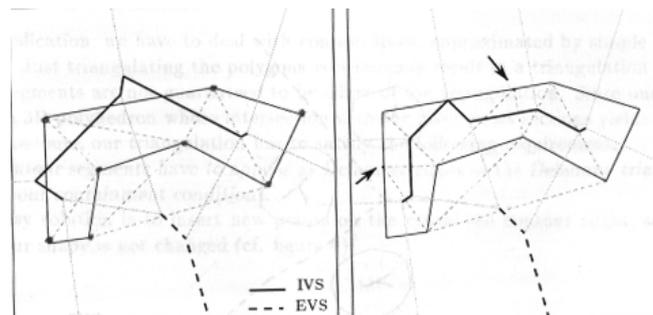


Abbildung 3.26: Hinzufügen von Punkten beim internen Voronoi-Skelett

Falls die Voronoi-Kontur teilweise außerhalb der Objektkontur liegt, müssen zusätzlich einige Punkte beim internen Voronoi-Skelett hinzugefügt werden. Grafik aus [Boi92] entnommen.

Anschließend wird die Delaunay-Triangulation zwischen den Schichten durchgeführt. Es werden die nächsten Punkte für die Konstruktion der Delaunay-Triangulation herangezogen. Anschließend müssen Tetraeder außerhalb der Objekte und nicht solide Verbindungen eliminiert werden. Die externen Voronoi-Skelette werden in die anliegende Schicht projiziert und zerschneiden die Objekte in Unterobjekte, die dann den einzelnen Unterarmen zugewiesen werden. Die Trennungslinie liegt durch die Definition des Voronoi-Diagramms immer auf halber Strecke zwischen den beiden Unterobjekten (vergleiche Abbildung 3.27).

Nach Aussage des Autors arbeitet dieser Algorithmus für einfache Objekte sehr gut. Probleme treten auf, wenn komplizierte Verästelungen und Löcher in den zu rekonstruierenden Objekten enthalten sind. Komplexe Trennlinien innerhalb von Objekten können nicht berücksichtigt werden, da keine Innenkanten erzeugt werden. Darüber hinaus wird auch keine Aussage über das Verhalten der Innenflächen gemacht (siehe Abbildung 3.28). Wie bewegen sich die Bildpunkte bei einer Verästelung? Kommen die Bildwerte einfach mit der nächsten Schicht oder Verengen sich die Arme und tragen zu der Veränderung der Fläche bei?

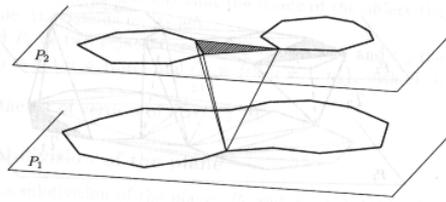


Abbildung 3.27: Elimination von Tetraedern ohne Objektbezug und Trennen in Unterobjekte

Nach der einfachen Delaunay-Triangulation können immer noch Tetraeder existieren, die ohne tatsächlichen Objektbezug sind. Diese sind im letzten Schritt zu entfernen. Grafik aus [Boi92] entnommen.

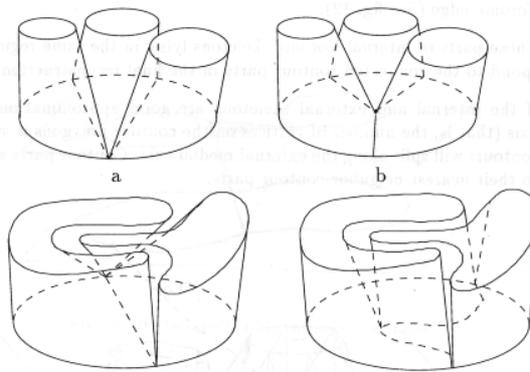


Abbildung 3.28: Problemfälle der Delaunay-Triangulation

Einige Problemstellungen lassen sich durch die Delaunay-Triangulation nicht bearbeiten. So besteht keine Möglichkeit, den Verlauf von Innenkanten vorzugeben. Grafik aus [Boi92] entnommen.

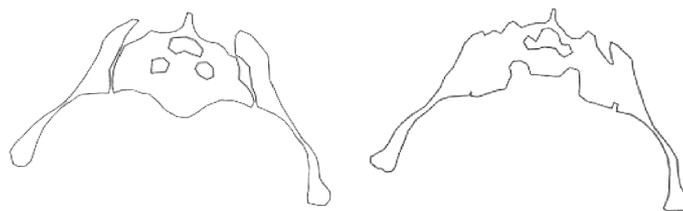


Abbildung 3.29: Problemstellung bei der Fortführung von Konturverläufen

Ein weiteres Problem stellt die eingeschränkte Möglichkeit zur Einflußnahme auf den Rekonstruktionsprozeß dar. Der Algorithmus ist nicht in der Lage, die unteren Einbuchtungen in der rechten Kontur in Verbindung mit der Trennlinie zwischen den einzelnen Segmenten im linken Bild zu setzen. Das manuelle Vorgeben solcher Korrespondenzen wird durch den Algorithmus nicht unterstützt. Grafik aus [Boi92] entnommen.

Doch nicht nur bei den Innenkanten stellt die mangelnde Kontrolle über den Korrelationsprozeß ein Problem dar. Auch bei kleinen, lokalen Translationen der Außenkontur werden bei der Konstruktion der Triangulation keine inhaltlichen Korrespondenzen berücksichtigt. In den Schichten leicht zueinander verschobene Konturen und Profile werden generell vertikal „ausgebügelt“ (siehe Abbildung 3.29).

3.9.3 Topologiebasierte Triangulation

Der in [Anj87] beschriebene Ansatz zur topologiebasierten Triangulation versucht eine möglichst realistisch erscheinende Oberflächenform zu erzeugen, indem die Form und Position der Konturen zueinander berücksichtigt werden (siehe Abbildung 3.30). Das Gewicht des Ansatzes liegt nach Aussage des Autors weniger auf der tatsächlichen Korrektheit der Korrelation, sondern auf der Berechnung einer möglichst runden und geschmeidigen Oberfläche. Hierzu werden die Konturen in den Schichten einander zugeordnet und das Korrelationsverhältnis, dem sogenannten count, festgestellt. Dabei werden die zugeordneten Regionen gezählt. Hat eine Region den count Null, wird sie als geschlossenes Objekt erklärt. Überlagern sich zwei Regionen, wird dies als einfache Situation bezeichnet und direkt per Triangulation verbunden (siehe Abbildung 3.31).

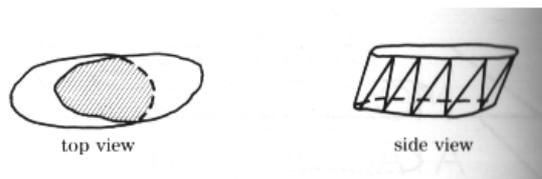


Abbildung 3.30: Darstellung von zwei topologisch ähnlichen verschobenen Kurven

Die Überschneidungen zwischen den beiden Konturen fließen in den Rekonstruktionsprozeß ein [Anj87].

Treten zwischen mehr als zwei Objekten Überschneidungen auf, werden die Schnittflächen bestimmt. Entsprechend den Flächenverhältnissen wird dann die Gesamtregion unterteilt. Nach welchem Verfahren die Untergliederung vorgenommen wird, ist in dem Artikel nicht spezifiziert. Im letzten Schritt werden die Arme den unterteilten Regionen zugeordnet und trianguliert (siehe Abbildung 3.31). Auf diese Weise werden auch komplexere Regionenkonstellationen bearbeitet. Innen- und Außenkonturen werden unterschieden und getrennt abgearbeitet.

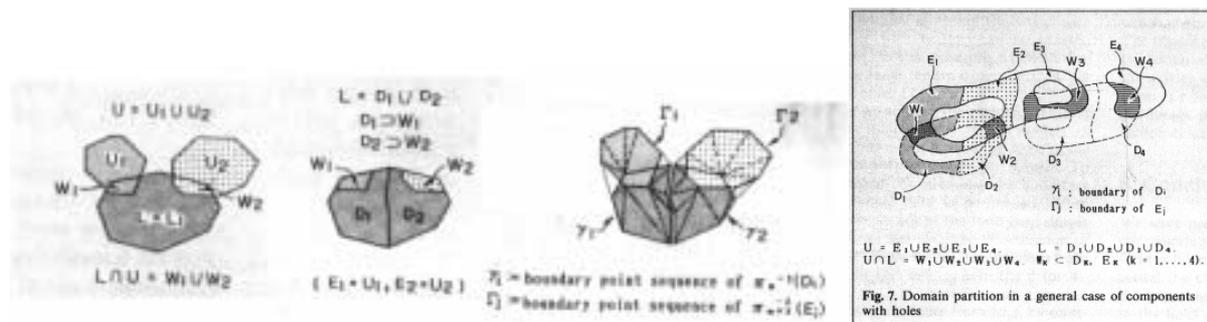


Abbildung 3.31: Untergliedern in Unterobjekte aufgrund der Konturüberlagerung [Anj87]

Bei Mehrfachüberschneidungen werden die Überschneidungsflächen als Grundlage für die Aufteilung der Gesamtkontur herangezogen.

Ein Schwachpunkt des Ansatzes besteht in der Untergliederung des Objektes durch die Bestimmung der Überschneidungen. Treten beträchtliche Translationen zwischen den Konturen auf, wird eine Rekonstruktion der Verbindung erschwert oder sogar vollständig unterlassen. Bei der Rekonstruktion von mehreren Armen stellt die Überlagerung der Regionen das einzige Korrelationskriterium dar, anhand derer die flächenbasierte Untergliederung der Regionen durchgeführt wird. Gerade bei komplex verästelten Organen oder dem Zusammensetzen eines Organs aus mehreren Lappen erscheint dieses Unterteilungskriterium mittels relativer Überschneidungsflächen als unzureichend.

3.9.4 Zusammenfassende Beurteilung

Alle beschriebenen Verfahren favorisieren bei der Korrelation der Umrißlinien vertikale Verbindungen. Für eine korrekte Konfiguration der Korrespondenzen muß jedoch vor allem der jeweilige Objektkontext berücksichtigt werden. Keines der Verfahren berücksichtigt beispielsweise den Fall, daß Konturen einfach zueinander verdreht sein könnten. Aufgrund der Konzeption der Verfahren sind mit Ausnahme des topologischen Ansatzes keine Modifikationen möglich, die ein Anpassen auf individuelle Objekteigenschaften erlaubt. Dieser Nachteil ist um so schwerer zu gewichten, da die Integration von a priori Wissen über das Objekt die Qualität der Oberflächenrekonstruktion deutlich verbessern kann. Gerade bei zweideutigen Situationen kann Vorwissen über die zu rekonstruierenden Objekte darüber entscheiden, ob die Rekonstruktion korrekt ist oder nicht (vergleiche Abbildung 3.32).

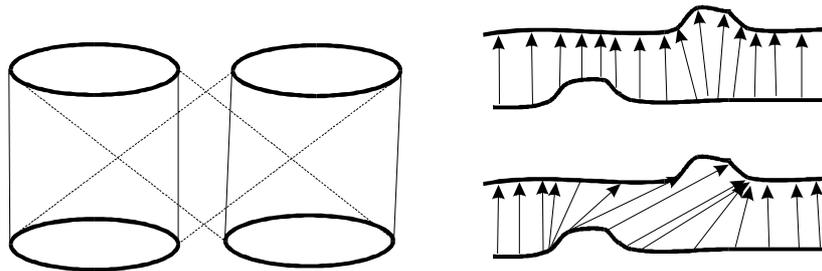


Abbildung 3.32: Zwei Problemstellungen bei der Korrelation von Konturen

Links: Es existieren mehrere Möglichkeiten zur Korrelation der Konturen, gerade und schräg. Rechts: Gehören die Ausbuchtungen zusammen oder nicht?

Im Falle des Auges ist bei der Rekonstruktion des Bulbus eine möglichst gleichmäßige sphärische Form der Oberfläche anzustreben. Wenn hingegen die aufgenähten Marker rekonstruiert werden sollen, ist aufgrund der gegenwärtigen Artefakte eine direkte Rekonstruktion nicht möglich. Die tatsächliche Form der Marker wird bei den MRT-Bildern durch die verursachten Artefakte deutlich zu groß dargestellt. Bei CT-Aufnahmen bilden hingegen die Metallartefakte strahlenförmige Ausläufer aus. Da andererseits die Form der Marker einheitlich und von vornherein bekannt ist, kann der Rekonstruktionsschritt statt dessen mit deutlich größerer Präzision realisiert werden, indem die bekannte dreidimensionale Form in die Daten eingepaßt wird.

Bei genauerer Betrachtung kann sogar festgestellt werden, daß in den Daten lediglich das Massenzentrum bestimmt werden muß, um eine korrekte Rekonstruktion der Marker zu erreichen. Die präzisen Ausmaße der Marker sind bekannt. Darüber hinaus sind die Marker mit geringen Abweichungen stets platt auf der Oberfläche des Bulbus aufgenäht, wodurch die Orientierung bezüglich der Bulbusoberfläche bekannt ist. Wurde die Oberfläche des Bulbus in einem vorangegangenen Schritt rekonstruiert, kann allein aufgrund der Kenntnis des Massenzentrums des Markers eine vollständige Rekonstruktion erreicht werden.



Abbildung 3.33: Rekonstruktion des Bulbus und der aufgenähten Marker

Während der Bulbus weitgehend sphärisch und regelmäßig ist, kann der Marker aufgrund seiner definierten Ausmaße durch Berücksichtigung des Vorwissens trotz der durch ihn verursachten Artefakte stabil rekonstruiert werden.

Wiederum ganz anders müßte an die Rekonstruktion des Gehirns von CT- oder MRT-Daten heran gegangen werden. Bei einer Anwendung der oben beschriebenen Verfahren, könnten durch die vertikal und rein geometrische ausgerichteten Korrelationsverfahren schnell anatomisch divergierende Sulci miteinander verbunden werden. Um auch hier eine korrekte Rekonstruktion zu erreichen, müßte anatomisches Wissen gegebenenfalls auch durch interaktive Eingaben durch einen Benutzer hinzugezogen werden. Interaktive Modifikationen der Rekonstruktion unterstützten jedoch keiner der oben beschriebenen Ansätze.

Die in diesem Abschnitt beschriebenen Überlegungen sollten veranschaulichen, daß die Berücksichtigung von objektspezifischen Oberflächen- und Variationseigenschaften auch bei der Rekonstruktion von Organoberflächen zu einer deutlichen Stabilisierung der Rekonstruktion führen kann. Der entscheidende Faktor zur Berücksichtigung solcher objektspezifischer Eigenschaften besteht in der Nutzung von Vorwissen über die zu rekonstruierenden Objekte. Darüber hinaus sind Objektschichten mit normalen Darstellungseigenschaften abzugrenzen von jenen mit unerwünschten Modifikationen wie beispielsweise Artefakten. Bei der Rekonstruktion des Bulbus sollten außerdem an der Ober- und Unterkante die durch die Partialvolumeneffekte verwischten Konturen durch Splines kompensiert werden.

3.10 Interpolation

Medizinische MRT- oder CT-Bilddatensätze weisen in der Regel innerhalb der Schichten (x- und y-Achse) deutliche höhere Auflösungen auf als zwischen den Schichten (z-Achse) (siehe Abbildung 3.34 links). Sobald die Betrachtungsebenen nicht mehr in den ursprünglichen Schichtebenen liegen, wirkt sich die geringe Auflösung in der z-Achse negativ auf die Darstellung der Daten aus. Aufgrund der asymmetrischen Voxelgröße werden die Organkanten eckig und verschwommen dargestellt (siehe Abbildung 3.34 rechts).

Ein schwerwiegendes Problem stellt die schlechte Auflösung des Datensatzes nicht nur für die Visualisierung der Daten, sondern auch für die Rekonstruktion der Organe und die Kombination von unterschiedlichen Datensätzen (Registrierung) dar. Um die Folgen einer schlechten Auflösung zu reduzieren, werden medizinische Bilddatensätze interpoliert, um die Auflösung künstlich zu erhöhen. Interpolationsverfahren versuchen den Verlauf der gegebenen Grauwerte möglichst sinnvoll fortzusetzen, indem nach vorgegebenen Gesetzmäßigkeiten Grauwerte für den Bereich zwischen den existierenden Voxeln berechnet werden [Gre98].

Je nach Aufgabenstellung wird die Interpolation angewandt, um die Auflösung lediglich in der z-Achse zu erhöhen oder um den Datensatz bezogen auf ein beliebig definiertes Koordinatensystem neu zu

berechnen. Die Ansätze zur Interpolation von medizinischen Datensätzen können dabei in szenebasierte und objektbasierte Verfahren unterteilt werden [Gre99].

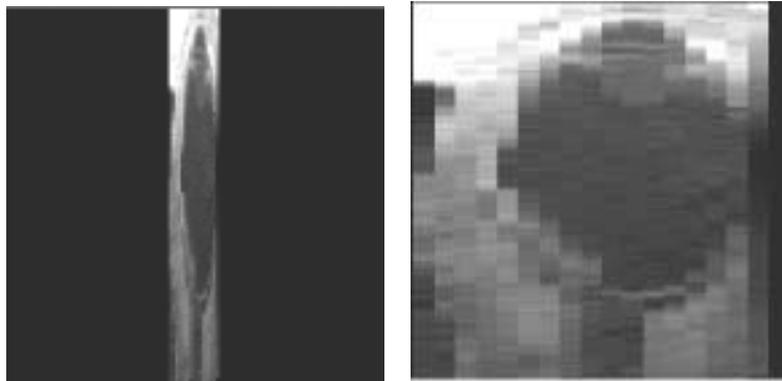


Abbildung 3.34: Darstellung eines MRT-Datensatzes mit einer Auflösung von 256x256x17

Links: Darstellung des Datensatzes auf die y-z-Ebene. Während die Auflösung in der y-Achse 256 beträgt, können in der z-Achse (horizontal) nur 17 Pixel eingetragen werden. Wird die tatsächliche Voxelgröße auf den Datensatz angewandt, erhält man die Darstellung rechts. Jedes Voxel ist 1x1x8 Einheiten groß. Durch die anisotrope Auflösung werden die Organkanten kantig und verschwommen dargestellt.

Bei den szenebasierten Verfahren wird die Interpolationsfunktion alleine aufgrund der Grauwertinformationen des Datensatzes berechnet. Der wohl einfachste Vertreter dieser Kategorie von Interpolationsverfahren stellt die lineare Interpolation dar. Bei der linearen Interpolation werden die Grauwerte der Zwischenebenen berechnet, indem die Grauwerte des jeweils darüber und darunter liegenden Voxels mittels Inverse-distance-weighted-Interpolation interpoliert werden. Ein Nachteil dieses Ansatzes stellt jedoch die Tatsache dar, daß durch das direkte Interpolieren von übereinander liegenden Voxeln in vielen Fällen anatomisch nicht zusammengehörige Voxel miteinander korreliert werden. Gerade in Randbereichen von Organen gehören übereinander liegende Voxel oft nicht zum gleichen Organ. Während beispielsweise das eine Voxel dunkel ist und zum Bulbus gehört, ist das darüber liegende Voxel hell und gehört zum Fettgewebe. Bei der Interpolation dieser Voxel wird in den Zwischenschichten das gesamte Grauwertspektrum zwischen diesen Grauwerten abgebildet. Dadurch kommt es zu einem Verwischen der Randbereiche und einer verfälschten Darstellung der Organgrenzen (siehe Abbildung 3.35 links).

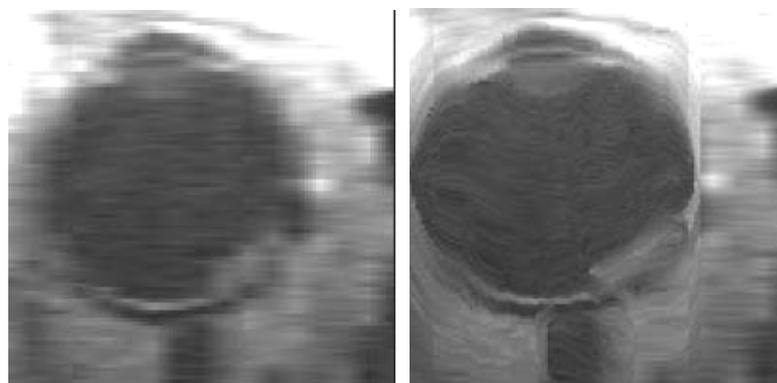


Abbildung 3.35: Ergebnisse der linearen und oberflächenbasierten Interpolation

Links: Ergebnis der linearen Interpolation des Datensatzes aus Abbildung 3.34. Durch das Interpolieren von anatomisch nicht zusammengehörenden Voxeln ist der Randbereich sämtlicher Organe stark verwischt. Rechts: Der Bulbus wurde als Referenz für eine oberflächenbasierte Interpolation genutzt. Dadurch wurden die Kanten des Bulbus besser erhalten als bei der linearen Interpolation.

Bei den objektbasierten Interpolationsverfahren werden Objekte als Referenz für die Definition der Interpolationsfunktion genutzt. Diese Referenzen beschreiben anatomisch markante Strukturen und vermeiden die Korrelation von Voxeln unterschiedlicher anatomischer Identität bei der Berechnung der Interpolationspfade. Als Referenzobjekte werden in der Regel segmentierte Organe genutzt. Für die Interpolation eines MRT-Datensatzes der Orbita bietet es sich beispielsweise an, den Bulbus zu segmentieren, um dadurch zu verhindern, daß die dunklen Voxel des Glaskörpers mit den hellen Voxeln des Fettgewebes verwischt werden (siehe Abbildung 3.35, rechts).

Für die Vermeidung dieser Vermischung von unterschiedlichen anatomischen Strukturen bietet der objektbasierte Interpolationsansatz unterschiedliche Verfahren an. Grundsätzlich wird unterschieden zwischen shape-based Interpolation und objektbasierte Interpolation mittels Warping. Bei den shape-based Interpolationsverfahren werden zunächst auf den segmentierten Organen aufsetzend die Bilddaten binarisiert. Anschließend werden über Distanztransformationen Verbindungsvektoren zwischen den Voxeln der einzelnen Schichten konstruiert. Anhand dieser Verbindungsvektoren werden dann mittels Inverse-distance-weighted-Interpolation der ursprünglichen Schichtdaten die Grauwerte für die gesuchten Zwischenschichten berechnet [Gre96, Her92]. Ein Nachteil der Verwendung von Distanzfunktionen für die Berechnung von Interpolationspfaden besteht darin, daß durch die Distanzfunktionen nur gerade Interpolationspfade berechnet werden können (siehe Abbildung 3.36). Da die Organe statt dessen jedoch weitgehend runde Formen aufweisen, sollten die Interpolationspfade nicht nur auf der Basis von zwei aneinander grenzenden Schichten, sondern auch anhand der danach folgenden Schichten berechnet werden.

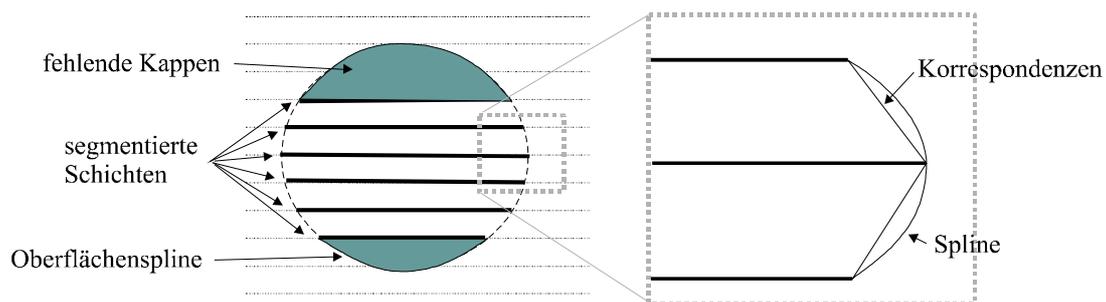


Abbildung 3.36: Weiche Oberflächeninterpolation durch Anwendung von Splines

Bei der objektbasierten Interpolation werden zunächst Korrespondenzen zwischen den Schichten bestimmt. Diese Korrespondenzen werden anschließend genutzt, um die Organoberfläche zu rekonstruieren. Die bei shape-based Interpolationsverfahren angewandten Distanztransformationen erlauben lediglich die Repräsentation von geraden Oberflächenkorrespondenzen (siehe rechte Abbildung). Sinnvoller ist jedoch eine runde, die globale Organform nachbildende Rekonstruktion der Oberfläche. Die Korrespondenzen werden durch Splines abgerundet, um die runden Organoberflächen besser nachzuempfinden. Abschließend werden die Grauwerte entlang der Korrespondenzpfade interpoliert.

Dieses Problem wird durch den zweiten Ansatz der objektbasierten Interpolation behoben, indem die Interpolationspfade nicht durch Distanztransformationen, sondern direkt aus der Korrespondenz von Landmarken oder Oberflächenpunkten berechnet werden. Die Korrespondenzen in Form von Vektoren werden auf das gesamte Datenvolumen interpoliert und beschreiben die Interpolationspfade, entlang welcher die Grauwerte der ursprünglichen Schichtdaten mittels Inverse-distance-weighted-Interpolation interpoliert werden. Ein sehr wichtiger Parameter bei der Berechnung der Interpolationspfade stellt die Auswahl der Interpolationsfunktion zur Interpolation der Korrespondenzvektoren auf den gesamten Datensatz dar. Gängige Verfahren verwenden Gauß-Funktionen [Ap96], B-Splines [Ham97, Lee97], Radial Basis Interpolation [Rup94, Rup95a, Rup95b] und andere Techniken [Cha98, Hon96, Kar95, Tho96, Wan97, Zha96].

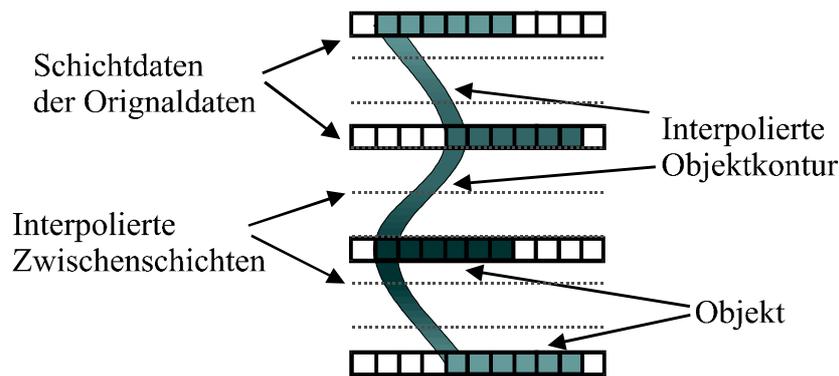


Abbildung 3.37: Berechnung der Grauwerte entlang der Interpolationspfade

Während die Interpolationspfade durch Interpolationsfunktionen wie Radial-Basis oder B-Splines berechnet werden, können die Grauwerte entlang der Interpolationspfade mit der einfachen Inverse-distance-weighted-Interpolation berechnet werden.

Die in der Literatur beschriebenen Ergebnisse zeigen, daß objektbasierte Interpolationspfade deutlich bessere Ergebnisse produzieren als szenenbasierte Verfahren [Gre99]. Darüber hinaus erwiesen sich objektbasierte Verfahren mittels Warping als flexibler und besser steuerbar als die Verfahren mittels Distanzfunktionen. Bei der Berechnung der Interpolationspfade können beliebige Referenzobjekte (Punkte, Linien oder Oberflächen) in beliebiger Anzahl ausgenutzt werden. Dabei kann als Regel festgestellt werden, daß je mehr Referenzen genutzt werden, desto besser ist in der Regel das Interpolationsergebnis.

3.11 Registrierung

Registrierung ist in der medizinischen Bildverarbeitung der Überbegriff für Verfahren, die mehrere Bilddatensätze miteinander kombinieren. Einerseits kann es sich um Datensätze handeln, die den gleichen Bildmodus aufweisen (z.B. CT-Aufnahmen) und die von dem gleichen Patient zu unterschiedlichen Zeitpunkten aufgenommen wurden. Eine solche Aufgabenstellung besteht vor allem dann, wenn die Entwicklung von Gewebestrukturen über die Zeit hinweg beobachtet werden soll. Durch die Korrelation der Organoberflächen von Datensatz zu Datensatz kann die Entwicklung des betreffenden Organs detailliert beobachtet werden. Andererseits, und dies ist der in der medizinischen Praxis deutlich häufigere Anwendungsbereich der Registrierungsverfahren, werden von einem Patienten Datensätze mit unterschiedlichen Bildmodi aufgenommen (beispielsweise CT-Daten und MRT-Daten). Diese Kombination von unterschiedlichen Bilddaten hat seinen Ursprung in der Tatsache, daß die verschiedenen bildgebenden Verfahren spezifische, voneinander abweichende Darstellungseigenschaften aufweisen. Damit verbunden sind divergierende Vor- und Nachteile und verschiedenen Informationsinhalte im Bildmaterial.

Diese unterschiedlichen Informationen können je nach medizinischer Aufgabenstellung mehr oder weniger wichtig für eine korrekte Diagnose und Therapie sein. Durch die Kombination von unterschiedlichen Datensätzen wird versucht, die disjunkten Informationen synergetisch zusammenzuführen, während sich Nachteile der verschiedenen Bildmodi gegenseitig kompensieren sollen. Ein aktuell intensiv angewandter Ansatz ist die Kombination von T1- und T2-Aufnahmen, um die Verlässlichkeit von Weichgeweidentifikationen zu verbessern. Beispielhafte Anwendung ist die Unterscheidung von alten und neuen Schlaganfällen oder die Abgrenzung und Unterscheidung von Tumoren [Ber95, Ber96, Ber97]. Die Kombination von CT- und MRT-Daten ist vor allem für die Strahlentherapie von großer Bedeutung, da die CT-Daten die Absorptionseigenschaften des Gewebes beschreiben, während die MRT-Daten eine bessere Abgrenzung der Weichgewebestrukturen ermöglichen [Nöh97, Nöh98a, Nöh98b].

Um die Informationen in unterschiedlichen Datensätzen gleichzeitig auswerten zu können, müssen sie zunächst anatomisch korrekt miteinander korreliert werden. Das bedeutet, daß jedes Voxel des einen

Datensatzes dem korrespondierenden Voxel des anderen Datensatzes zugeordnet werden muß. Da sich Datensätze fast immer in Aufnahmeparametern wie Orientierung, Auflösung oder Verzerrung voneinander unterscheiden, sind zunächst eine Reihe von Verfahren auf die Daten anzuwenden, um diese Korrelation von Voxeln zu ermöglichen. Bei dem Großteil der medizinischen Verfahren werden die folgenden Verarbeitungsschritte durchgeführt:

1. Inhaltliche Interpretation der Daten (Segmentierung),
2. Transformation des einen oder beider Datensätze in eine einheitliche Datenstruktur (Interpolation) und
3. Korrelation der Daten unter Anwendung der zuvor ermittelten inhaltlichen Interpretation (Registrierung).

Eine inhaltliche Interpretation von Bilddaten wird bei zwei- oder dreidimensionalen Schnittbildern durch die Segmentierung erreicht. Bevor dann die inhaltlichen Interpretationen der unterschiedlichen Bilddaten miteinander korreliert werden können, sind die Daten in eine einheitliche Datenstruktur zu transformieren. Die Datenstruktur zeichnet sich durch grundlegende Eigenschaften wie Dimension (ein-, zwei- oder dreidimensional), Datendichte (Auflösung) und Datenrepräsentation (Grauwerte, Farbwerte oder Amplitudeninterpretation) aus. Durch die Anwendung von Interpolationsverfahren können die Datenstrukturen abgeglichen werden [App96, Cha98, Gre98, Gre99, Kar95, Lee97, Rup94, Rup95a, Rup95b, Tho96, Zha96, Wan97]. Im letzten Schritt werden die inhaltlichen Interpretationen durch die Registrierung miteinander korreliert. In der wissenschaftlichen Literatur werden die unterschiedlichsten Verfahren beschrieben. Die Vielseitigkeit der entwickelten Ansätze ist dabei ebenso umfassend, wie es die zu registrierenden Bilddaten sind. Mainz kategorisiert Registrierungsverfahren nach den folgenden Kriterien [Mai98]:

- Dimensionalität (2D, 3D, 2+1D, 3+1D),
- Registrierungsbasis (extrinsisch, wie beispielsweise mittels künstlicher Landmarken, oder intrinsisch auf der Basis der Patientendaten oder mittels kalibrierter Scanner-Koordinaten),
- Transformationsmodus (rigide, affine, projektive oder Kurventransformation),
- Transformationsdomäne (global, lokal),
- Grad der Benutzerinteraktion (automatisch, semi-automatisch),
- Optimierungsfunktion (Mutual Information, Kreuzkorrelation, Cepstrum...),
- Bildmodalitäten (CT, MRT, PET...),
- Bildgegenstand (Intrasubjekt oder intersubjekt) und
- Anatomischer Bildgegenstand.

Da sich diese Arbeit in erster Linie mit der Verarbeitung von dreidimensionalen Schichtdaten beschäftigt, und ein zentraler Punkt für die Bildverarbeitung das vorausgesetzte Wissen über das Bildmaterial ist, sollen die aktuellen Registrierungsverfahren durch die folgenden Kategorien untergliedert werden (vergleiche [Bro92]):

- Korrelations- und statistische Registrierungsverfahren,
- Referenzbasierte Registrierung und
- Modellbasierte Registrierung.

3.11.1 Korrelations- und statistische Registrierungsverfahren

Korrelations- und statistische Registrierungsverfahren benötigen, nicht wie oben dargestellt, zunächst eine inhaltliche Interpretation der Bilddaten. Die inhaltliche Interpretation besteht statt dessen implizit in der Annahme, daß entweder die Grauwerteigenschaften, die Kantenstrukturen oder Texturen bei den zur registrierenden Datensätzen in einem globalen und gleichbleibenden Verhältnis stehen. Die Qualität der Übereinstimmung der Bildcharakteristiken wird durch unterschiedliche Korrelationsfunktionen quantifiziert und dienen als Maßstab für die Qualität der Registrierung [Chr95, Kov96, Mae97, Mit96, Red96, Sat95].

Ein Schwachpunkt dieser Verfahren ist das Voraussetzen von globalen Beziehungen zwischen den Grauwerten der beiden Bilder. Vor allem bei den MRT-Daten können durch die starke Varianz der Bildeigenschaften keine definierten Entsprechungen von Grauwerten und Bildinhalten festgestellt werden, so daß die Ergebnisse dieser Ansätze einen systematischen Fehler beinhalten. Dieser Fehler würde gerade bei dem zugrunde liegenden Projekt auftreten, da Artefakte wie beispielsweise der Chemical Shift beim Bulbus nahezu immer die MRT-Daten lokal verfälschen. Darüber hinaus setzen diese Verfahren voraus, daß die Bildqualität einheitlich ist, daß die Schichten die gleiche Orientierung und Auflösung aufweisen, und daß eine direkte Beziehung zwischen den Grauwerten existiert. Diese Voraussetzungen stellen jedoch eine Einschränkung des Anwendungsbereichs dar und werden in vielen Fällen vom Bildmaterial nicht vollständig erfüllt.

Beispielsweise könnten nur dreidimensionale Schichtdaten mit dreidimensionalen Schichtdaten, die eine annähernd gleiche Auflösung und einen weitgehend gleichen Bildinhalt aufweisen, miteinander korreliert werden. Sind diese Anforderungen an Auflösung und Orientierung nicht erfüllt, muß in jedem Fall vor Durchführung der Registrierung eine Interpolation durchgeführt werden. Für die Oberflächenbasierte Interpolation müßte jedoch zunächst eine Segmentierung jener Strukturen vorgenommen werden, die für die Interpolation herangezogen werden kann. Die Durchführung einer Segmentierung würde wiederum die Anwendung eines statistischen Registrierungsverfahrens an sich in Frage stellen, da in diesem Fall eine referenzbasierte Registrierung aller Voraussicht nach bessere Ergebnisse produzieren würde.

Bei den statistischen Registrierungsverfahren können darüber hinaus nur rigide Transformationen vorgenommen werden. Das würde die Anwendung dieser Verfahren auf exakt gleich große und orientierte Datensätze beschränken. Um dieses Problem zu vermeiden wurde der grundsätzliche Verfahrensansatz erweitert, indem Bilder in Fensterbereiche unterteilt werden, die anschließend separat registriert werden. Die regionalen rigiden Transformationen können anschließend zu einer globalen flexiblen Registrierung zusammengefaßt werden. Dabei werden die Einzelergebnisse auf die gesamte Bildmatrix interpoliert [Mai98, Sat95, Chr96].

Für die Realisation des in dieser Arbeit entwickelten Bildverarbeitungsansatzes wird eine Korrelationsfunktion genutzt werden, die ebenfalls oft im Zusammenhang mit statistischen Registrierungsverfahren erwähnt wird. Diese Funktion, Mutual Information genannt, sei in dem folgenden Abschnitt detailliert beschrieben.

3.11.2 Mutual Information

Mutual Information ist die englische Bezeichnung eines informationstheoretischen Parameters, der derzeit vielfach für die Registrierung von medizinischen Bildern Anwendung findet [Mai98, Mey96, Plu99, Rod98, Stu98]. Da der Großteil der wissenschaftlichen Literatur in Englisch verfaßt ist, konnte sich besonders für die Statistik keine allgemein akzeptierte deutsche Nomenklatur durchsetzen. Um die Eindeutigkeit der hier verwendeten Nomenklatur sicherzustellen, wird im weiteren Verlauf dieser Arbeit die englische Nomenklatur übernommen.

$$H(X | Y) = \sum_{0 \leq k \leq l} \left(\sum_{0 \leq i \leq j} p(x_i | y = y_k) \cdot \log \left(\frac{1}{p(x_i | y_i = b)} \right) \right) \quad (22)$$

$$H(X | Y) = \sum_{0 \leq k \leq l} p(y_k) \left(\sum_{0 \leq i \leq j} p(x_i | y = y_k) \cdot \log \left(\frac{1}{p(x_i | y_i = y_k)} \right) \right) \quad (23)$$

$$H(X | Y) = \sum_{\substack{0 \leq i \leq j \\ 0 \leq k \leq l}} p(x_i, y_k) \cdot \log \left(\frac{1}{p(x_i | y_i = y_k)} \right) \quad (24)$$

Wenn es keine Korrelation zwischen x_i und y_i gibt, dann ist $H(X | y=y_k) = H(X)$. Sind auf der anderen Seite x_i und y_i direkt miteinander korreliert, das heißt, wenn x_i bekannt ist, dann ist auch y_i bekannt, dann ist $H(X | y_i=b) = 0$. Somit beschreibt die Differenz zwischen der Entropie beider Ereignisse und der Conditional Entropy, wie stark die Stichprobe Y die Unsicherheit von X reduziert.

$$MI(X, Y) = H(X) - H(X | Y) \quad (25)$$

Diese Differenz wird Mutual Information (oder Transinformation) genannt. Wird diese Formel ausformuliert und zusammengefaßt erhält man durch Einsetzen der Definition der bedingten Wahrscheinlichkeit [Bro91, Seite 659]

$$p(x | y) = \frac{p(x, y)}{p(y)} \quad (26)$$

die folgende Formel für die Mutual Information:

$$MI(X, Y) = \sum_{\substack{0 \leq i \leq j \\ 0 \leq k \leq l}} p(x_i, y_k) \cdot \log \left(\frac{p(x_i, y_k)}{p(x_i) p(y_i)} \right) \quad (27)$$

Um zwei Bilder A und B mittels Mutual Information zu registrieren, wird die Mutual Information für die unterschiedlichen Transformationen von B auf Bild A berechnet. Die optimale Registrierung wird dann angenommen, wenn die Grauwerte des Bildes A am stärksten mit den Grauwerten in Bild B korrelieren. Es muß also jene Transformation gefunden werden, für die die Mutual Information ein globales Maximum hat. Die Idee hinter der Korrelationsfunktion Mutual Information besteht darin, daß auch, wenn die Grauwerte von Bild A und Bild B für die gleichen Organe unterschiedlich sind, so sind die dargestellten Organe doch aller Wahrscheinlichkeit nach gleich groß und gleich geformt. Die Anzahl der Pixel die den Bulbus ausmachen sind in den MRT-T1-Aufnahmen bei gleicher Skalierung genauso groß, wie bei der MRT-T2- oder CT-Aufnahme.

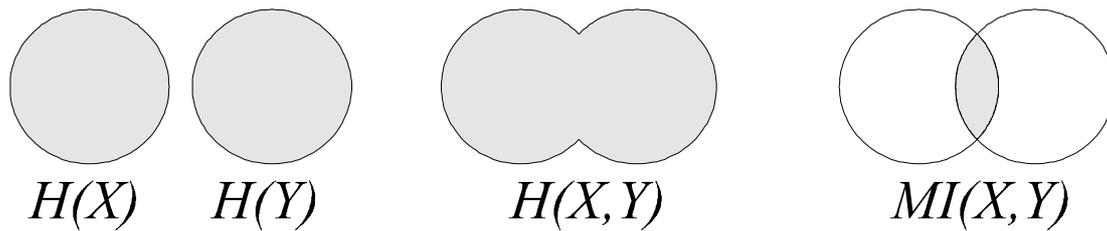


Abbildung 3.39: Veranschaulichung der Mutual Information

Die Abbildungen erläutern die marginalen Entropien $H(X)$ und $H(Y)$, die Joint Entropy $H(X,Y)$ und das Verhältnis zwischen Mutual Information $MI(X,Y)$ und den marginalen Entropien $H(X)$ und $H(Y)$.

Nachfolgend soll die Berechnung der Mutual Information für ein Bild A , das n mal m Pixel groß ist und das transformierte Bild B , das k mal l Pixel groß ist, erläutert werden. Das Histogramm eines Bildes A lautet [Leh97] :

$$h : A \rightarrow N \text{ mit } h(g) = |\{(x, y) \in A \text{ mit } A(x, y) = g\}| \quad (28)$$

Von diesem Histogramm lassen sich die jeweiligen marginalen Entropien ableiten. Dementsprechend lautet die Mutual Information für das aktuelle Bild A und das transformierte Bild B_T :

$$B_T(x', y') = B(x, y) * T \quad (29)$$

$$p(a) = \frac{|\{(x, y) \in A \text{ mit } A(x, y) = a\}|}{n \cdot m} \quad (30)$$

$$p(b_T) = \frac{|\{(x, y) \in B_T \text{ mit } B_T(x, y) = b_T\}|}{l \cdot k} \quad (31)$$

$$p(a, b_T) = \frac{|\{(x, y) \in A \wedge (x, y) \in B_T \text{ mit } A(x, y) = a \wedge B_T(x, y) = b_T\}|}{|A \cap B_T|} \quad (32)$$

$$MI(A, B_T) = \sum_{\substack{\text{für alle } a \\ \text{für alle } b_T}} p(a, b_T) \cdot \log\left(\frac{p(a, b_T)}{p(a)p(b_T)}\right) \quad (33)$$

Sofern nach der Transformation von B eine Interpolation angewandt wird, gilt in der Regel $p(b) \neq p(b_T)$. Eine Interpolation der transformierten Bilddaten wird vor allem dann notwendig sein, wenn eine Transformation angewandt wird, die zu deutlichen Unterschieden bezüglich der diskretisierten Koordinatensystemen von A und B_T führt. Ein Beispiel für eine solche Transformation ist die Rotation von Bildern.

Indem das Maximum von $MI(A, B_T)$ für alle möglichen Transformationen T bestimmt wird, kann jene Transformation T identifiziert werden, die die beiden Datensätze am besten in Übereinstimmung bringt.

3.11.3 Referenzbasierte Registrierungsverfahren

Die referenzbasierte Registrierung verwendet einander zugeordnete Referenzobjekte, die die Transformation des einen Datensatzes in den anderen koordinieren. Die Transformation an sich wird berechnet, indem die Verschiebungsvektoren zwischen den Referenzvektoren auf das gesamte Datenvolumen interpoliert werden. Abschließend werden diese Verschiebungsvektoren auf jenen Datensatz angewandt, der in den anderen Datensatz projiziert werden soll.

Als Referenzobjekte werden in der Regel in beiden Datensätzen die gleichen anatomischen Strukturen segmentiert und anschließend einander zugeordnet. Der Vorteil dieses Ansatzes besteht in der flexiblen Vorgabe der Transformationsfunktion durch Punkte, Linien, Flächen, oder Oberflächen [Dea97, Fan97, Liu97, Mai96, Mau96, Mau97, Mey95].

Um den gesamten Registrierungsprozeß zu automatisieren, versuchen einige Ansätze die Bestimmung der notwendigen Referenzlandmarken automatisch durchzuführen. Bei dem von [Ami97] entwickelten Ansatz werden im aktuellen Datensatz Grauwertbereiche gesucht, die mit Vergleichslandmarken in einem Referenzdatensatz übereinstimmen. Die dabei bestimmten Landmarken werden aufgrund ihrer räumlichen Anordnung identifiziert und für die Registrierungsfunktion angewandt. Ein weiteres Verfahren ermittelt die Landmarken durch die Vorgabe und Detektion von Kanteneigenschaften mittels Orientation-Maps [Sim96]. Die Zuverlässigkeit solcher automatisch vorgegebenen Landmarken ist jedoch aufgrund der Variabilität und Komplexität des medizinischen Bildmaterials instabil.

3.11.4 Modellbasierte Registrierungsverfahren

Modellbasierte Registrierungsverfahren verwenden ein-, zwei- oder dreidimensionale Modelle, die an die aktuellen Daten angepaßt werden. Die grundsätzlichen Unterschiede der Verfahren bestehen in der Definitionsform des Modells über Funktionen oder Oberflächenmodellen und in der Anpassungsroutine [Ami96, Bur95, Dav96, Fis97, Mei98, Ris96, San97, Sha96, Shi97, Tho96].

Neben der oben beschriebenen softwarebasierten Registrierung existiert auch die Registrierung anhand von stereotaktischen Rahmen, die die anschließende softwarebasierte Registrierung deutlich vereinfacht. Für die Registrierung von Bilddaten des Auges kommt die Registrierung mittels stereotaktischer Rahmen nicht in Betracht, da das Field of View der Bilddaten in der Regel zu klein ist, um einen Rahmen sichtbar zu installieren.