

Aus dem Institut für Neuroradiologie
der Universität zu Lübeck
Direktor: Prof. Dr. med. Dirk Petersen

In Zusammenarbeit mit
Prof. Arvid Lundervold, M.D. Ph.D.
Abteilung Neuroinformatik und Bildanalyse
Institut für Physiologie
Universität Bergen, Norwegen

Gehirnvolumenberechnung aus multispektralen MRT-Bildern

Inauguraldissertation
zur
Erlangung der Doktorwürde
der Universität zu Lübeck
- Aus der Medizinischen Fakultät -

vorgelegt von
Frederick Klauschen
aus Itzehoe

Lübeck 2004

1. Berichterstatter	Prof. Dr. med. Dirk Petersen
2. Berichterstatter	Priv.-Doz. Dr. med. Oliver Schmitt
Tag der mündlichen Prüfung	13.01.2005
Zum Druck genehmigt. Lübeck, den	13.01.2005

Inhaltsverzeichnis

1	Einleitung	6
1.1	Die Gehirngröße - entwicklungsgeschichtlicher Überblick	6
1.2	Medizinische Bedeutung von Gehirnvolumenveränderungen und deren Messung	8
1.3	Methoden zur Bestimmung des Gehirnvolumens	10
1.3.1	Anatomische Verfahren	10
1.3.2	Konventionell-radiologische Verfahren	11
1.3.3	Computertomographie	12
1.3.4	Magnetresonanztomographie	13
1.3.5	Auswertung der Bilddaten	15
1.4	Fragestellung	17
2	Material und Methoden	19
2.1	Learning Vector Quantization <i>LVQ</i>	19
2.1.1	Einleitung	19
2.1.2	Datenformate und Konvertierungen	22
2.1.3	Trainingsprozedur	23

<i>INHALTSVERZEICHNIS</i>	3
2.1.4 Klassifizierungsvorgang	25
2.1.5 Fehlerkorrektur und Volumenberechnung	25
2.2 Statistical Parametric Mapping <i>SPM</i>	27
2.2.1 Einleitung	27
2.2.2 Datenformat und Konvertierung	28
2.3 Durchführung	29
2.4 Freesurfer	30
2.4.1 Einleitung	30
2.4.2 Datenformat und Konvertierung	30
2.4.3 Durchführung	31
2.5 AFNI	34
2.6 Datenmaterial	34
2.7 Computersystem und Software	36
3 Ergebnisse	42
3.1 BrainWeb	42
3.2 LVQ	43
3.3 SPM	47
3.4 Freesurfer	52
3.5 Vergleich der Ergebnisse	56
3.5.1 Graue Substanz	56
3.5.2 Weiße Substanz	57
3.5.3 Graue + weiße Substanz	57
3.5.4 Zusammenfassung	58
3.5.5 Liquor	59
4 Diskussion	61
4.1 LVQ	63

<i>INHALTSVERZEICHNIS</i>	4
4.1.1 Analyse der Ergebnisse der Volumenberechnung	64
4.2 SPM	66
4.2.1 Analyse der Ergebnisse der Volumenberechnung	67
4.3 Freesurfer	68
4.3.1 Analyse der Ergebnisse der Volumenberechnung	68
4.4 Datenmaterial	69
4.5 Schlussbetrachtungen	71
4.5.1 Bedeutung der Ergebnisse für die Gehirnvolumetrie	72
4.5.2 Ausblick	73
Zusammenfassung	75
Abkürzungen	77
Anhang	78
Programm-Listing Nr. 1: lvqtraining	78
Programm-Listing Nr. 2: lvqprepare	79
Programm-Listing Nr. 3: lvqperform	80
Programm-Listing Nr. 4: lvqreadclassified	80
Programm-Listing Nr. 5: readraw	81
Programm-Listing Nr. 6: afni2spm	82
Programm-Listing Nr. 7: corscan	82
Programm-Listing Nr. 8: volcount	83
Programm-Listing Nr. 9: calcvol	83
Programm-Listing Nr. 10: embedrotflip	84
Programm-Listing Nr. 11: perfit	85
Programm-Listing Nr. 12: lvqcorr2	85
Programm-Listing Nr. 13: deletepix	87

<i>INHALTSVERZEICHNIS</i>	5
Programm-Listing Nr. 14: cvol	88
Programm-Listing Nr. 15: spmvol	89
Literaturverzeichnis	97
Internet-Adressen	98
Danksagung	99
Lebenslauf	100

Kapitel 1

Einleitung

1.1 Die Gehirngröße - entwicklungsgeschichtlicher Überblick

Was können morphologische Merkmale eines Lebewesens über dessen Fähigkeiten bzw. funktionelle Eigenschaften aussagen? Diese Frage erscheint einfach, denkt man allein an die Vielfalt der im Tierreich bestehenden Körperformen und Merkmale, die direkt mit den Notwendigkeiten der jeweiligen ökologischen Nische in Zusammenhang gebracht werden können und deren genaues Studium dem Naturforscher Charles Darwin als Grundlage und Begründung für seine Evolutionstheorie gedient haben.

Aber auch wenn man nur die Entwicklung des einzelnen Individuum betrachtet, so treten im Körper strukturelle Veränderungen in jedem Lebensalter auf und geben Auskünfte über physiologische wie auch pathologische Zustände. Besonders eindrucksvoll ist die morphologische Entwicklung des Embryos, die uns sogar Rückschlüsse auf unsere Abstammung erlaubt. Betrachtet man die Entwicklung des menschlichen Gehirns in Abhängigkeit vom Lebensalter, stellt man fest, dass

das Volumen der weißen Substanz zunimmt, während das Volumen der grauen Substanz in der Kindheit erst zunimmt, aber bereits vor Erreichen des Erwachsenenalters wieder beginnt abzunehmen (Durston et al., 2001).

Aber inwieweit kann man einen Zusammenhang zwischen Morphologie und Volumen des Gehirns und dessen Funktion bzw. den intellektuellen Fähigkeiten eines Lebewesens feststellen?

Vergleicht man die Gehirnvolumina unterschiedlicher Arten, so erkennt man, dass nicht allgemein behauptet werden kann, dass ein größeres Gehirn auch mehr Intelligenz impliziert, auch wenn sicherlich ein positiver Zusammenhang zwischen Körper- und Gehirngröße berücksichtigt werden muss. So hat ein Elefant ein ungefähr viermal größeres Gehirn als ein Mensch, ein Rind ein Gehirn, dessen Volumen dem eines Gorillas entspricht, das fast 20 mal so groß wie das einer Katze und über 250 mal so groß wie das einer Ratte ist (Roth, 2000).

Hinweise für eine positive Korrelation zwischen Gehirngröße und zerebralen Fähigkeiten findet man allerdings in entwicklungsgeschichtlichen Untersuchungen. So erfolgt in der Entwicklung der Säugetiere eine deutliche Zunahme der Größe des Gehirnschädels und auch innerhalb der Evolution der Hominiden ist eine starke kontinuierliche Zunahme des Gehirns über aufeinanderfolgende Entwicklungsschritte zu verzeichnen. Z. B. ist das Gehirnschädelvolumen des *Homo sapiens* rund dreimal so groß wie das des frühen Menschen der Gattung *Australopithecus*, der vor ungefähr einer Jahrmillion lebte, und circa 50 Prozent größer als das des *Homo erectus*. Kein anderes Organ machte in der Entwicklung des Menschen eine so deutliche relative Größenveränderung durch (Herrmann und Ullrich, 1991).

1.2 Medizinische Bedeutung von Gehirnvolumenveränderungen und deren Messung

Für die Medizin sind insbesondere krankheitsbedingte Veränderungen des Gehirnvolumens von Interesse. Es ist allgemein bekannt, dass organische Krankheiten oft typische mikro- und makroskopische morphologische Korrelate aufweisen. Es gibt aber auch Untersuchungen über psychische Erkrankungen nichtorganischer Ursache, bei denen morphologische Auffälligkeiten des Gehirns beobachtet werden (s. u.). Die quantitative Beschreibung von morphologischen Eigenschaften des gesunden und des kranken Gehirns kann dabei nicht nur in der Grundlagenforschung wichtige Hinweise auf Auswirkungen oder Ursachen von Krankheiten liefern. Das Erkennen und die Abschätzung des Ausmaßes solcher Veränderungen stellt außerdem eine der wichtigsten Säulen der Diagnostik und Verlaufsbeurteilung vieler Erkrankungen dar und ist Voraussetzung für die Wahl effizienter Therapieoptionen.

Als Beispiel für Untersuchungen nichtorganischer psychiatrischer Erkrankungen können Veröffentlichungen genannt werden, die auf eine Abnahme des Volumens der grauen Substanz bei Patienten mit Schizophrenie hinweisen (Job et al., 2002; Lee et al., 2002). Auch Drogen wie Marihuana (Block et al., 2000) oder Alkohol (Götze et al., 1978; Graff-Radford et al., 1982; Moselhy et al., 2001) können zu Gehirnatrophie führen. Bei der Untersuchung des Gehirnvolumens bei einer Komorbidität von Schizophrenie und Alkoholismus konnte im Vergleich zum Vorliegen einer der beiden Erkrankungen eine stärkere Volumenabnahme der präfrontalen grauen Substanz festgestellt werden (Mathalon et al., 2003). Ein weiteres, besonders anschauliches Beispiel für den Zusammenhang zwischen Volumen und Funktion ist der Nachweis, dass die abnehmende Gehirnleistungsfähigkeit bei De-

menzerkrankungen mit einer Abnahme des Gehirnvolumens einhergeht (Bogdanoff et al., 1975; Bigler, 2001; Bigler und Tate, 2001; Bigler et al., 2000, 2002). Beim Morbus Alzheimer konnte eine Verringerung des Volumens der grauen Substanz im Bereich des Temporallappens festgestellt werden (Busatto et al., 2003). Bei Kindern kommt der Untersuchung des Gehirns eine besondere Bedeutung zu, da sich das Gehirn bei ihnen in einer steten Entwicklung befindet, die anatomische und funktionelle Fortschritte verbindet (Giedd et al., 1996; Sowell et al., 2002). Eine Synthese morphologischer und psychologischer Betrachtungen kann so Rückschlüsse auf die normale Kindesentwicklung und Abweichungen davon erlauben. Es konnte z. B. gezeigt werden, dass Kinder und Jugendliche mit hyperkinetischem Syndrom (ADHD) ein global kleineres Gehirnvolumen haben (Castellanos et al., 2002). Bei Kindern mit Sprachentwicklungsstörungen konnte eine Verringerung der weißen Substanz und Erweiterung der Ventrikel nachgewiesen werden (Trauner et al., 2000).

Die Messung des gesamten Gehirnvolumens ist aber nicht nur für direkte Aussagen über morphologische und psychopathologische Zusammenhänge von Bedeutung, sondern spielt auch bei der Untersuchung der relativen Größe von Teilen des Gehirns wie z. B. des Frontal- oder Temporallappens oder des Balkens eine Rolle, da hier die betrachteten Größen ins Verhältnis zum Gesamtvolumen gesetzt werden müssen. Als Beispiel kann die Forschung über die selektive Lese- und Rechtschreibschwäche (Legasthenie) genannt werden, bei der bei betroffenen Kindern Unterschiede der Flächen des Corpus callosum im sagittalen Schnittbild festgestellt wurden (von Plessen et al., 2002).

Eine ebenfalls wichtige Rolle spielt die Volumetrie bei Erkrankungen des Gehirns, die auch zu lokalisierten Veränderungen innerhalb eines Gewebetyps führen. Beispiele sind Neoplasien und entzündliche Prozesse wie die Multiple Sklerose, bei der Entzündungsreaktionen der weißen Substanz und Zeichen einer globalen Ge-

hirnatrophie auftreten (Bermel et al., 2003). Hier kann die Bestimmung der Größe der von der Entzündungsreaktion betroffenen Areale wichtige Aussagen über einen Zusammenhang mit der klinischen Symptomatik und über den Verlauf der Krankheit machen.

Praktische Anwendungen dieser und ähnlicher Ergebnisse ergeben sich jedoch nicht nur aus der Verbesserung der diagnostischen Möglichkeiten, sondern auch für die Beurteilung und den Vergleich verschiedener Therapiestrategien. Quantitative Zusammenhänge zwischen Gehirnvolumen bzw. der Größe von pathologischen Befunden und dem Ausprägungsgrad von Krankheiten liefern objektive Kriterien zur Bewertung des Einflusses verschiedener Therapien auf den Verlauf. Es ist daher anzunehmen, dass solche quantitativen Methoden im Zuge der evidenzbasierten Medizin in Zukunft an Bedeutung gewinnen werden. Ein Beispiel hierfür ist die Multiple Sklerose, bei der man die Parameter „gesamtes Gehirnvolumen“ und „Läsionsvolumen“ zur Beurteilung der Wirksamkeit neuer Medikamente wie des Interferon-beta benutzen kann (Gasperini et al., 2002; Paolillo et al., 2002).

Alle hier genannten Untersuchungen basieren auf der Bildgebung des Gehirns und auf Methoden, mit denen man aus den erstellten Bildern möglichst genau und zuverlässig das Volumen der interessierenden Gewebetypen bestimmen kann.

1.3 Methoden zur Bestimmung des Gehirnvolumens

1.3.1 Anatomische Verfahren

Die Bestimmung des Gehirnvolumens beschäftigt die Wissenschaft bereits seit Ende des 19. Jahrhunderts. Anfangs wurden dabei post mortem Präparate in For-

malin fixiert und deren Volumen durch die Messung der Flüssigkeitsverdrängung bestimmt (Marshall, 1892).

Eine Weiterentwicklung stellte die computergestützte Auswertung von Fotografien von Scheiben des post mortem Gehirns dar, weil so die Volumina von grauer und weißer Substanz getrennt quantifiziert werden konnten (Miller et al., 1980).

Des weiteren ist die Punktzählmethode in der Stereologie erwähnenswert, mit der sich effizient die Flächen von Substrukturen in Schnittbildern und damit auch aus Serien von Schnittbildern und bekannter Schichtdicke deren Volumina berechnen lassen (Howard et al., 2003).

1.3.2 Konventionell-radiologische Verfahren

Die erste radiologische *in vivo* Methode zur Darstellung und Quantifizierung von Strukturen des Zentralen Nervensystems (ZNS) war die Pneumenzephalographie (PEG), die es eingeschränkt möglich machte, die Liquorräume *in vivo* zu untersuchen. Bei dieser Methode wurden die Liquorräume durch Injektion von Luft kontrastiert und durch konventionelle Röntgenaufnahmen in verschiedenen Projektionen dargestellt. Diese Technik wurde schon in den 50er Jahren u. a. für Studien zur Analyse des Einflusses der Alkoholabhängigkeit auf das Liquorraumvolumen benutzt (Pluinage, 1954; Tumarkin et al., 1955), wobei Erweiterungen der Ventrikel sowie kortikale Atrophien festgestellt werden konnten. Methodische Mängel wie Luftfüllungsartefakte und die Abhängigkeit der Messungen von der injizierten Luftmenge sowie die aus der Zweidimensionalität resultierenden Projektionseffekte lassen allerdings bei der Pneumenzephalographie keine exakten quantitativen Volumenbestimmungen zu. Außerdem erlaubt dieses Verfahren keine Aussagen über das Verhältnis von weißer und grauer Substanz.

Grundsätzlich stehen als konventionell neuroradiologische, kontrastmittelabhängige Verfahren noch die Angiographie und die Myelographie zur Verfügung, die

allerdings wie die Pneumenzephalographie nur indirekte und quantitativ ungenaue Rückschlüsse über das Gehirngewebe zulassen und daher zur Gehirnvolumenberechnung unbrauchbar sind. Außerdem sind sie aufgrund ihrer Invasivität nicht risikolos und daher zur Anwendung bei größeren Patientenkollektiven ungeeignet.

1.3.3 Computertomographie

Das Erkennen und insbesondere die Quantifizierbarkeit von normalen Strukturen und Veränderungen des Gehirns ist erst durch die Ende der 60er Jahre von dem englischen Physiker Godfrey N. Hounsfield entwickelte und seit Anfang der 70er Jahre in der klinischen Anwendung eingesetzte Computertomographie (CT) zufriedenstellend möglich geworden (Hounsfield, 1973; Hounsfield, 1980). Es handelt sich um eine nichtinvasive Methode, die auch ohne Kontrastmittel eine Differenzierung der verschiedenen intrakraniellen Gewebetypen nach deren Dichte und damit eine präzisere Berechnung des Gehirnvolumens erlaubt. Erste Studien, die die Computertomographie benutzten, um quantitative Aussagen über Zusammenhänge zwischen dem Volumen von Gehirnstrukturen und Krankheiten bzw. deren Ausprägungsgrad zu machen, wurden Mitte der 70er Jahre durchgeführt. Fox et al. (1975) konnten auf diese Weise bei Patienten mit fokaler Epilepsie Zeichen von Gehirnatrophie feststellen. In den Untersuchungen von Bogdanoff et al. (1975) wurde sogar eine Korrelation zwischen dem Ausmaß von Hirnrindenatrophie und dem Grad seniler Demenz nachgewiesen.

Auch wenn diese Untersuchungen einen großen Fortschritt gegenüber der Pneumenzephalographie darstellten, waren zu der Zeit genaue quantitative Aussagen über das Gehirnvolumen nur begrenzt möglich, da die Auswertung der Bilddaten aufgrund unzureichender Rechenleistung und fehlender Algorithmen schwierig war. In neueren Studien konnten mit weiterentwickelter Technik und Software

die Aussage von Bogdanoff et al. (1975) bestätigt und verfeinert werden (Hatazawa et al., 1982; Wanifuchi et al., 2002).

1.3.4 Magnetresonanztomographie

Erst mit der Entwicklung der Magnetresonanztomographie (MRT) für die medizinische Bildgebung (Damadian et al., 1977; Mansfield und Maudsley, 1977; Lauterbur, 1980) stand eine Technik zur Verfügung, mit der ohne Strahlenbelastung Schnittbilder des menschlichen Körpers angefertigt werden konnten.

Der Vorteil gegenüber der Computertomographie besteht im viel höheren Weichteilkontrast der MRT, der eine genaue Differenzierung des Gehirngewebes in graue und weiße Substanz möglich macht und die Unterscheidbarkeit des Gehirns von nicht-Gehirngewebe verbessert.

Die Grundlagen der Magnetresonanztomographie und der physikalische Ursprung der Bildinformation lassen sich exakt nur quantenmechanisch verstehen. Einige Atomkerne (z. B. Wasserstoff) besitzen einen Eigendrehimpuls, genannt *Spin*, durch den sie ein magnetisches Moment erhalten. In einem äußeren statischen Magnetfeld können sich die Kernspins nicht mehr beliebig ausrichten, sondern nur in ganz bestimmten Winkeln relativ zur Magnetfeldrichtung, so dass die Spins um die Magnetfeldachse mit einer bestimmten Frequenz präzedieren. Beim Wasserstoff gibt es zwei verschiedene Orientierungen, wobei ein Spin in der gleichen Orientierung wie das Magnetfeld einen energetisch günstigeren Zustand für die Kerne darstellt als ein Spin gegen die Magnetfeldrichtung. Auch wenn im thermischen Gleichgewicht (bei Körpertemperatur) die Besetzungszahlen der unterschiedlichen Zustände nur gering voneinander abweichen, gibt es eine Nettomagnetisierung des Gewebes in Richtung des Magnetfeldes (Längsmagnetisierung). Mit Einstrahlung eines hochfrequenten magnetischen Wechselfeldes in Form von Radiowellen einer Energie, die dem Energieunterschied der verschiedenen Zu-

stände entspricht, können nun Übergänge vom energetisch niedrigen zum höheren Zustand angeregt werden (Resonanz), die nach Beendigung der Hochfrequenzeinstrahlung nach einer bestimmten Zeit wieder unter Strahlungsabgabe in den Grundzustand übergehen (Relaxation). Wählt man jedoch nur einen kurzen Hochfrequenzpuls, ändert sich die Magnetisierung nicht vollständig, also um 180 Grad, sondern um einen kleineren Winkel. Die Spins präzedieren nun um die Richtung der neuen Magnetisierung. Insbesondere lässt sich die Pulsdauer so wählen, dass die Magnetisierung nach dem Puls im rechten Winkel zum ursprünglichen, konstanten Magnetfeld steht, die Längsmagnetisierung also verschwunden, aber eine transversale Komponente aufgetreten ist (Quermagnetisierung); man spricht dann von einem 90-Grad-Puls.

Nach Beendigung des Hochfrequenz-Pulses kehren die Spins unter Abgabe von Strahlung in den Ausgangszustand zurück. Die Signalintensität wird bestimmt durch den Anteil der Wasserstoffatome im Gewebe (dies ist näherungsweise der Wassergehalt); man nennt dies *Protonendichte*-Wichtung.

Die Zeit, die die Kernspins brauchen, um wieder in die ursprüngliche Orientierung zurückzukehren, ist die Spin-Gitter-Relaxationszeit (*T1-Wert*), da sie von Wechselwirkungen zwischen Spins und den sie umgebenden Strukturen abhängt.

Die Spin-Spin-Relaxationszeit (*T2-Wert*) ist ein Maß für die Abnahme der Quermagnetisierung aufgrund von sich gegenseitig beeinflussenden magnetischen Momenten, die zu einem Auseinanderlaufen (Verlust der Phasenkohärenz) der präzedierenden Spins führt.

Da die Energiedifferenz zwischen den möglichen Zuständen und damit die Resonanzfrequenz abhängig von der magnetischen Feldstärke ist, kann man durch Umschalten von Magnetfeldgradienten anhand der Frequenz der bei der Relaxation erzeugten Strahlung den Ursprungsort feststellen und erhält so die für die Tomographie notwendige Ortsinformation (Haken und Wolf, 1993; Hornak, 2002).

Während also die Computertomographie nur die lokale Elektronendichte des Gewebes abbildet, ermöglicht die MRT durch die Verwendung unterschiedlicher Pulssequenzen die Messung verschiedener Gewebeeigenschaften. Neben den hier vorgestellten und in dieser Arbeit verwendeten Gewichtungen T1, T2 und Protonendichte gibt es noch viele andere Sequenzen (Uhl et al, 1998).

Es besteht also bei der MRT die Möglichkeit das zu untersuchende Gewebe mit verschiedenen Gewichtungen abzubilden, was weitere Vorteile für die Qualität der Gewebeklassifizierung mit sich bringt, da sich bei der Kombination von mehreren Bildsequenzen die Differenzierbarkeit zwischen verschiedenen Gewebetypen erhöht bzw. überhaupt erst möglich wird.

Wegen dieser Eigenschaften wird die MRT heute als Methode der Wahl für Aufgaben wie die Gehirnvolumenberechnung angesehen.

1.3.5 Auswertung der Bilddaten

Bisher werden bei der klinischen Auswertung von radiologischem Bildmaterial Befunde hauptsächlich qualitativ oder nur grob quantitativ beschrieben. Dies hat verschiedene Gründe: Zum einen liegt es daran, dass erst seit relativ kurzer Zeit ausreichend leistungsstarke Computersysteme zu günstigen Preisen verfügbar sind, um die rechen- und speicheraufwendigen Bildanalysen durchzuführen, die eine Gewebeklassifizierung und damit quantitative Auswertung der Bilddaten ermöglichen. Ein anderer Grund besteht darin, dass dies im klinischen Alltag für die Diagnose und Verlaufsbeurteilung vieler Erkrankungen als ausreichend angesehen wird. Es ist allerdings wahrscheinlich, dass durch die Weiterentwicklung der Diagnostik in Richtung einer differenzierteren Betrachtung von Pathologien und Bildung neuer Krankheitsentitäten in Zukunft quantitative Aussagen an praktischer Bedeutung gewinnen werden. Auch im Zusammenhang mit der Entwick-

lung von neuen Möglichkeiten der MRT, neben morphologischen auch funktionelle Eigenschaften darzustellen, kommt quantitativen rechnergestützten Methoden eine besondere Bedeutung zu, da hier zur Bildauswertung verschiedenste Informationen zu einem Bild kombiniert werden müssen.

Noch ohne Computer kommen die einfachsten Ansätze zur quantitativen Auswertung von Bilddaten aus. Sie bestehen in der manuellen Messung von Durchmessern des Gehirns, der Ventrikel oder pathologischer Veränderungen wie Tumoren. In den oben genannten Studien (Pluvinage, 1954; Tumarkin et al., 1955) werden die Abstände zwischen bestimmten, einfach identifizierbaren Punkten im Gehirn benutzt. Diese Methode kann jedoch nur in einigen Anwendungen als gute Näherung für Volumenänderungen benutzt werden. Es gelingt so z. B. nicht, die Volumina von weißer und grauer Substanz getrennt zu bestimmen. Ein anderes Vorgehen besteht darin, manuell die Grenzen des betrachteten Volumens in CT- oder MRT-Schichtbildern mit speziellen Computerprogrammen nachzuzeichnen. Dieses Verfahren ist jedoch nicht nur sehr zeitaufwendig, sondern auch durch subjektive Faktoren der durchführenden Person beeinflusst, so dass es sich höchstens für kleinere, gut abgrenzbare Bereiche eignet (von Plessen et al., 2002).

Die Gehirnvolumetrie ist daher auf Methoden angewiesen, die automatisch, schnell und präzise arbeiten, damit die Auswertung der großen Datenmengen möglich ist. Grundsätzlich kommen dafür verschiedenste Algorithmen in Frage, denen unterschiedliche mathematische Ansätze zugrunde liegen.

Einer der ersten Ansätze geht auf den Klassifikationsalgorithmus nach *Haslett* zurück, der Bildelemente anhand der Wahrscheinlichkeit der Zugehörigkeit zu einer Klasse zuordnet (Haslett et al., 1985). Neuere Methoden kommen u. a. aus den Bereichen (*Fuzzy*) *C-Means*, *neuronale Netzwerke*, *Maximum Likelihood* und *k-Nearest Neighbor* (Bezdek et al., 1993). In vielen Ansätzen werden zur Verbesserung des Klassifikationsergebnisses unterschiedliche Methoden kombiniert (Chiu

et al., 2001). Man unterscheidet dabei auch sog. „supervised“ und „unsupervised methods“, d. h. Methoden, die Arbeitsschritte beinhalten, bei denen der Untersucher über die Festlegung von Trainingsklassen in den Klassifizierungsvorgang eingreifen muss, von weitestgehend automatischen Verfahren. Welche Methode sich am besten für die MRT-Datenklassifizierung eignet, ist bisher noch unzureichend untersucht worden. Außerdem ist zu vermuten, dass die Auswahl der Methode abhängig von Art und Qualität der Bilddaten und von der Fragestellung der Volumenberechnung ist. So kann z. B. für die Betrachtung von Gehirnvolumina bei Patienten, die an Demenzerkrankungen leiden und bei denen es wichtig ist, das Volumen der grauen im Vergleich zur weißen Substanz zu betrachten, eine andere Methode geeignet sein, als bei Fragestellungen, die nur das gesamte Volumen des Gehirnparenchyms im Vergleich zum Liquorraumvolumen betrachten. Des Weiteren ist anzunehmen, dass Art und Anzahl der verwendeten MRT-Bildsequenzen Einfluss auf das Klassifikationsergebnis haben.

1.4 Fragestellung

Für die Anwendbarkeit der Gehirnvolumetrie in Studien und in zukünftigen klinischen Anwendungen ist es essentiell, dass praktikable Methoden vorhanden sind, die das Volumen mit hoher Genauigkeit und weitestgehend unbeeinflusst von Störfaktoren der Bildakquisition bestimmen können. Als Beitrag zu dieser Diskussion werden in dieser Arbeit drei Methoden der Gehirnsegmentierung und folgenden Gehirnvolumenberechnung evaluiert und verglichen.

Es wird der Klassifizierungsalgorithmus *Learning Vector Quantization* aus dem Bereich der neuronalen Netzwerke für die multispektrale MRT-Bildanalyse implementiert und verwendet. Die beiden anderen Algorithmen sind Bestandteil der Programmpakete *SPM99 (Statistical Parametric Mapping)* und *Freesurfer* und

wurden ausgewählt, weil sie in der Forschung bereits weit verbreitet und leicht zugänglich sind und über die Segmentierung hinausgehende Funktionen im Bereich der funktionellen MRT besitzen.

Es sollen die Fragen geklärt werden, wie präzise und robust diese Algorithmen das Gehirnvolumen und die Teilvolumina von grauer und weißer Substanz und von Liquor bestimmen können und wie diese Leistungsmerkmale von den Eigenschaften des verwendeten Bildmaterials abhängen. Dabei werden die Methoden anhand von Bilddaten bekannter Volumina mit verschiedenen Bildqualitäten und unterschiedlichen Kombinationen der Bildsequenzen *T1*, *T2* und *Protonendichte* getestet.

Kapitel 2

Material und Methoden

In den ersten drei Abschnitten werden die hier verwendeten Klassifizierungsalgorithmen mit den entsprechenden Programmpaketen und den in dieser Arbeit entwickelten Erweiterungen vorgestellt.

Begriffsklärung: Die Verwendung der Begriffe „*Spektrum*“ und „*multi-spektral*“ in dieser Arbeit hat keinen Bezug zur Magnetresonanztomographie, sondern steht im Zusammenhang mit der Abbildung des Gehirns durch unterschiedliche Bildsequenzen (Spektrum = Gewichtung = Bildsequenz).

2.1 Learning Vector Quantization *LVQ*

2.1.1 Einleitung

Bei den Learning-Vector-Quantization-Verfahren (*LVQ*) handelt es sich um eine Gruppe von Algorithmen, die vom *LVQ* Programming Team Prof. T. Kohonen des Laboratory of Computer and Information Science, Helsinki University of Technology, zur Bearbeitung verschiedenster Aufgaben der Mustererkennung

und Klassifizierung, wie z. B. Bildverarbeitung und Spracherkennung, entwickelt wurden. *LVQ* basiert auf *neuronalen Netzen*, Algorithmen, die in ihrer Funktionsweise Ähnlichkeiten zu biologischen Gehirnen aufweisen, da sie es ermöglichen, mit Hilfe von Computern unbekannte Objekte aufgrund von Ähnlichkeiten mit bekannten Objekten zu erkennen bzw. zu klassifizieren (Kohonen, 1995; Kohonen et al., 1996).

Ein neuronales Netz besteht aus Knoten und Kanten, die jeweils zwei Knoten verbinden. Man unterscheidet Knoten der Eingabeschicht, einer oder mehrerer Zwischenschichten und der Ausgabeschicht. Die Kanten unterscheiden sich voneinander dadurch, dass sie verschieden stark gewichtet sein können. Die Anzahl an Knoten und ihre Verteilung auf die möglichen Schichten hängt von Art und Komplexität des Klassifizierungsproblems ab. Die Anwendung von neuronalen Netzen auf Mustererkennungsaufgaben besteht im Wesentlichen aus zwei Prozessen.

Lernen anhand bekannter Muster

Am Anfang des *Lern-* oder *Trainingsvorgangs* legt der Anwender fest, wie viele verschiedene Klassen er benutzen möchte, d. h. hier, in wie viele verschiedene Gewebetypen das Gehirn unterteilt werden soll.

Dann muss für jede dieser Klassen ein Lerndatensatz zusammengestellt und jeder Klasse ein eindeutiger Name (Label) gegeben werden. Dabei ist es wichtig, dass die Muster so gewählt werden, dass sie die jeweilige Klasse möglichst optimal repräsentieren. Wenn die Mengen der Merkmale verschiedener Klassen ausreichend disjunkt sind (d. h. gar nicht oder nur wenig überlappen), wenn man geeignete Beispielmuster ausgesucht hat und die Anzahl der gewählten Klassen der Anzahl der im Merkmalsraum wirklich vorhandenen disjunkten Klassen ähnlich ist, wird eine Klasse A eine andere Konfiguration von Wichtungen erzeugen als alle anderen Klassen. Dies ist Voraussetzung für eine erfolgreiche spätere Klassifizierung

unbekannter Objekte.

Die verschiedenen Lerndaten mit entsprechendem Label werden nun dem neuronalen Netzwerk als Beispieleingaben präsentiert. Der Algorithmus funktioniert so, dass (bei anfänglich gleich starken Verbindungen) die Kanten zwischen verschiedenen Netzknoten den Lerndaten entsprechend unterschiedlich gewichtet werden. Dadurch wird es möglich, unbekannte Muster, die Ähnlichkeiten mit gelernten Mustern aufweisen, einer Klasse zuzuordnen.

Klassifizieren von unbekanntem Mustern

Im *Klassifizierungsvorgang* wird das „angelernte“ neuronale Netz schließlich auf unbekannte Muster angewendet, die jetzt anhand der während des Trainings gewonnenen Informationen einer Klasse zugeordnet werden können. Die Güte der Klassifizierung hängt neben den im vorangehenden Abschnitt genannten Voraussetzungen von der Größe der Ähnlichkeit des zu klassifizierenden Objektes mit den Trainingsobjekten ab (Details zur Durchführung sind ab Abschnitt 2.1.3 beschrieben).

Die *LVQ*-Algorithmen¹ sind für verschiedenste Klassifizierungsaufgaben geeignet und können über Parameteranpassungen für spezielle Anwendungen optimiert werden. In dieser Arbeit wird die Methode auf multispektrale MRT-Daten zur Gewebeklassifizierung angewendet. In den folgenden Abschnitten wird die genaue Vorgehensweise dieser Arbeit dargestellt.

¹Eine vollständige Dokumentation des *LVQ*-Programm-Pakets findet sich unter der Internet-Adresse <http://www.cis.hut.fi/research/>, wo auch das Softwarepaket heruntergeladen werden kann. Außerdem ist eine komplette Beschreibung als Buch veröffentlicht (Kohonen, 1985).

2.1.2 Datenformate und Konvertierungen

An dieser Stelle werden die unterschiedlichen Datenformate und Konvertierungen beschrieben, die in den verschiedenen Bearbeitungsschritten notwendig sind.

Als erstes wird der zu untersuchende Gehirndatensatz mit dem *Matlab*-Skript `readraw` (Programm-Listing Nr. 5) geladen. Da die ursprünglichen Daten aus dem *MRI Simulator* (vgl. Kapitel 2.6) im Format $(181 \times 217 \times 181)$ vorliegen und hier aber ein Format von $(256 \times 256 \times 256)$ Volumenelementen (Voxel) von Vorteil ist², werden die Daten mit dem *Matlab*-Programm `embedrotflip` (Programm-Listing Nr. 10) konvertiert. Es findet gleichzeitig eine Rotation und Spiegelung der Bilder statt, damit unter *Matlab* an sagittalen Bildern Trainingsdaten ausgewählt werden können. Die gesamte zu bearbeitende Datenmenge bei drei Spektren ist dann $3 \times 256^3 \approx 50\text{MByte}$ bei 8Bit pro Voxel. Während der Datenbearbeitung innerhalb der *Matlab*-Umgebung wird der Datensatz scheibenweise in der Form (Anzahl der Bildpunkte in Y-Richtung) \times (Anzahl der Bildpunkte in X-Richtung) \times (Anzahl der Spektren), also $(256 \times 256 \times 3)$ gespeichert, wobei die z-Werte (Scheiben) 1, 2, 3 den Spektren T1, T2, PD entsprechen. Gespeichert werden diese Daten in dem *Matlab*-spezifischen Format (Suffix `mat`). Die für die Trainings- und Testdaten verwendeten Bilder haben das Format $(256 \times 256 \times 4)$, denn in der vierten Scheibe befinden sich die vom Anwender festgelegten Klassenbezeichnungen. Für den *LVQ*-Klassifizierungsvorgang müssen diese Daten ins ASCII-Format konvertiert werden, und zwar so, dass alle Bildpunkte in der Datei zeilenweise aufgelistet sind. Jede Spalte enthält die Bildwerte eines Spektrums und dementsprechend jede Zeile die drei verschiedenen Spektralwerte eines Bildpunktes, so dass sich folgende dreispaltige Matrix ergibt:

²Bei *Freesurfer* und *SPM* vereinfacht ein 256^3 -Format die Bearbeitung, da so die Daten standardisiert vorliegen und leichter weitere Untersuchungen, wie z. B. Geschwindigkeitsvergleiche durchgeführt werden können.

$$\begin{pmatrix} T1_1 & T2_1 & PD_1 \\ T1_2 & T2_2 & PD_2 \\ \vdots & \vdots & \vdots \\ T1_{256^2} & T2_{256^2} & PD_{256^2} \end{pmatrix}$$

Hier hat so ein Bild also die Form ($256^2 = 65536 \times 3$) und wird im ASCII-Format (Suffix `dat`) gespeichert, wobei in der ersten Zeile noch die Anzahl der Spektren als Dimensionszahl eingefügt werden muss.

2.1.3 Trainingsprozedur

Die Trainingsprozedur wird im Wesentlichen in zwei Teilschritten durchgeführt. Im ersten wird die Anzahl der unterschiedlichen Klassen als $k_{max} = 7$, in die das Gehirn³ eingeteilt werden soll, mit den in Tabelle 2.1 aufgeführten Zuweisungen gewählt.

Die Auswahl der Trainingsdaten wird interaktiv in *Matlab* mit dem Skript `lvqttraining` (Programm-Listing Nr. 1) durchgeführt. Es wurde dafür die sagittale Scheibe $z = 122$ gewählt und jeweils die linke obere und rechte untere Begrenzung der repräsentativen Klassen markiert. Für jede Klasse wurden ein bis fünf verschiedene Bereiche gewählt, um möglichst repräsentative Trainingsdaten zu haben. Anhand von einer dreidimensionalen Darstellung der Trainingsdaten nach den Variablen $T1$, $T2$ und PD (siehe Abbildung 2.1) kann man überprüfen, wie gut die Unterteilung des Ereignisraums in verschiedene Klassen durchgeführt wurde. Wie man sehen kann, ist die hier gewählte Unterteilung akzeptabel, denn

³Es gibt natürlich kein Hirngewebe „Luft“, allerdings wird Luft im MRT-Datensatz abgebildet und da sie von Liquor zu unterscheiden sein soll, muss es für eine Klasse für Luft geben.

Tabelle 2.1: Klassenzuweisungen der Bilddaten

Klasse k	Beschreibung
1	Gesichtsweichteile
2	graue Substanz
3	Luft
4	Knochen
5	weiße Substanz
6	Liquor
7	Muskulatur

die verschiedenen Punktwolken überlappen kaum. Eingabedatei ist ein unbearbeitetes MRT-Bild im *Matlab*-Format, die Ausgabedateien `testlvqascii.dat` und `trainlvqascii.dat` werden im *LVQ*-ASCII-Format gespeichert.

Danach wird mit dem im *LVQ*-Paket enthaltenen Programm `lvq_run` der für den Klassifizierungsvorgang benötigte Code-Vektor erstellt, indem den Bildschirmanweisungen für die Erstellung eines neuen „Classifiers“ gefolgt wird und zunächst die vorgeschlagenen Einstellungen verwendet werden. Der Code-Vektor enthält die Werte der Kantenwichtungen. Am Ende des ersten Durchlaufs wird anhand von Testdaten (aus `testlvqascii.dat`) die Klassifizierungsgüte (engl. *recognition accuracy*) bestimmt, die anzeigt, ob bereits ausreichend trainiert wurde oder ob noch eine Feinabstimmung (engl. *fine-tuning*) angeschlossen werden sollte.

Als Trainings- bzw. Testdatei werden `trainlvqascii` und `testlvqascii` verwendet. Am Ende des Trainings erhält man zwei Dateien `<Name>.cod` und `<Name>.lvq`, die die Code-Vektoren enthalten, also die für den Klassifizierungsvorgang notwendigen Parameter des neuronalen Netzwerkes bereitstellen.

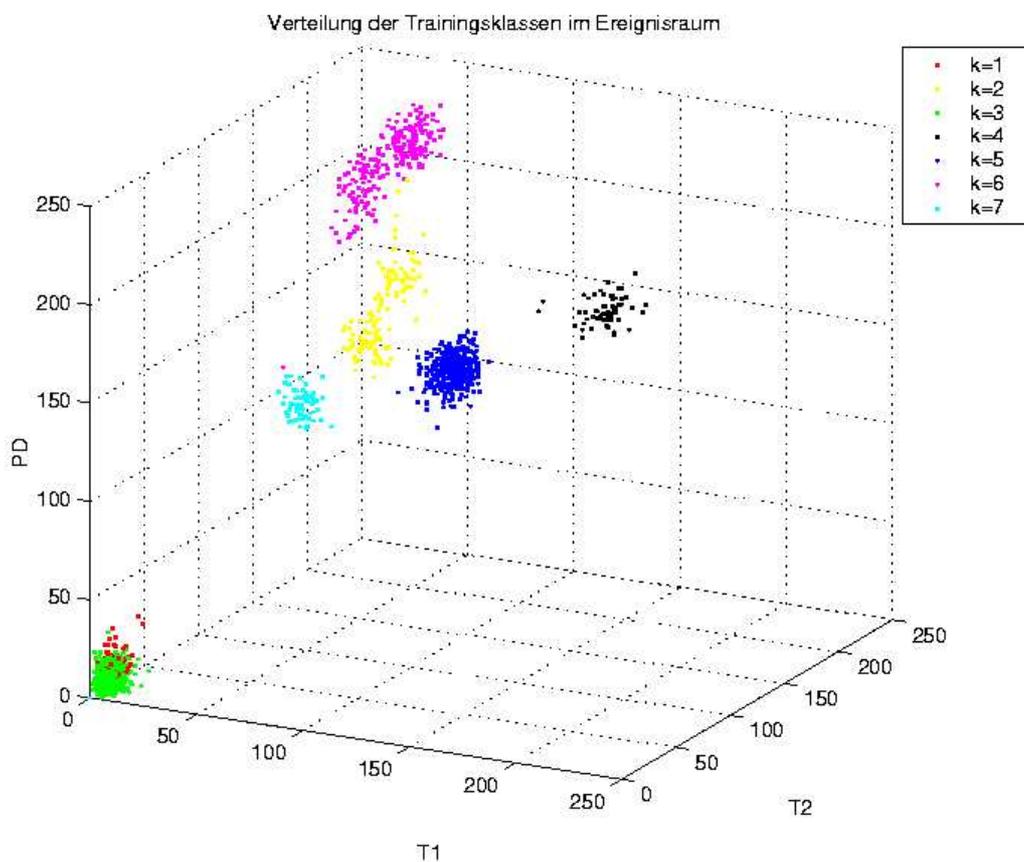


Abbildung 2.1: Darstellung der Trainingsdaten im Ereignisraum. Jede Punktwolke zeigt die Verteilung der Daten einer Klasse in den drei Spektren $T1$, $T2$ und PD . Die Klassen k entsprechen den Zuweisungen in Tabelle 2.1.

Es wurde mit einem Feinabstimmungsdurchgang eine Güte von 97,8% erreicht. Jetzt kann der Klassifizierungsvorgang begonnen werden.

2.1.4 Klassifizierungsvorgang

Mit dem Skript `lvqprepare` (Programm-Listing Nr. 2) werden die unbearbeiteten MRT-Bilddateien der Form $(256 \times 256 \times 3)$ im *Matlab*-Format in *LVQ*-lesbare ASCII-Dateien umgewandelt.

Das Linux-C-Shell-Skript `lvqperform` (Programm-Listing Nr. 3) initiiert den eigentlichen Klassifizierungsvorgang der Bilddateien, wobei scheibenweise vorgegangen wird.

Nach der Klassifizierung werden die Ergebnisdateien im *LVQ*-ASCII-Format zur weiteren Bearbeitung wieder in das *Matlab*-Format überführt, wobei eine das gesamte Gehirn enthaltende Datei der Form $(256 \times 256 \times 256)$ erstellt wird. Dies erledigt das *Matlab*-Skript `lvqreadclassified` (Programm-Listing Nr. 4). Das Ergebnis der Klassifikation und die verwendeten Spektren des MRT-Datensatzes sind in Abbildung 2.2 dargestellt.

2.1.5 Fehlerkorrektur und Volumenberechnung

Im letzten Abschnitt wurde der Gehirndatensatz klassifiziert, d. h. jedes Voxel wurde einer der nach den Trainingsbedingungen möglichen Klasse zugeordnet. Obwohl zur Klassifizierung MRT-Daten in drei verschiedenen Gewichtungen benutzt wurden, gibt es Gewebeteile, die fälscherlicherweise dem Gehirn zugeordnet werden, da ihre Signaturen⁴ denen des Gehirns gleichen.

Bevor also das Gehirnvolumen berechnet werden kann, müssen diese Fehler behoben oder zumindest bestmöglich reduziert werden. Dazu dienen zwei Gruppen im Folgenden vorgestellter Algorithmen:

Korrektur durch Nachbarschaftsvergleich

Hier werden Fehler aufgrund von empirischen Annahmen über Wahrscheinlichkeiten der Existenz von Nachbarschaftsbeziehungen und der Konnektivität von Gehirngewebe korrigiert.

Dabei kommen zwei Überlegungen zur Anwendung:

⁴Mit „Signatur“ eines Gewebetyps sind hier die drei (mittleren) Intensitätswerte gemeint, die ein bestimmtes Gewebe typischerweise in den Gewichtungen T1, T2 und Protonendichte aufweist.

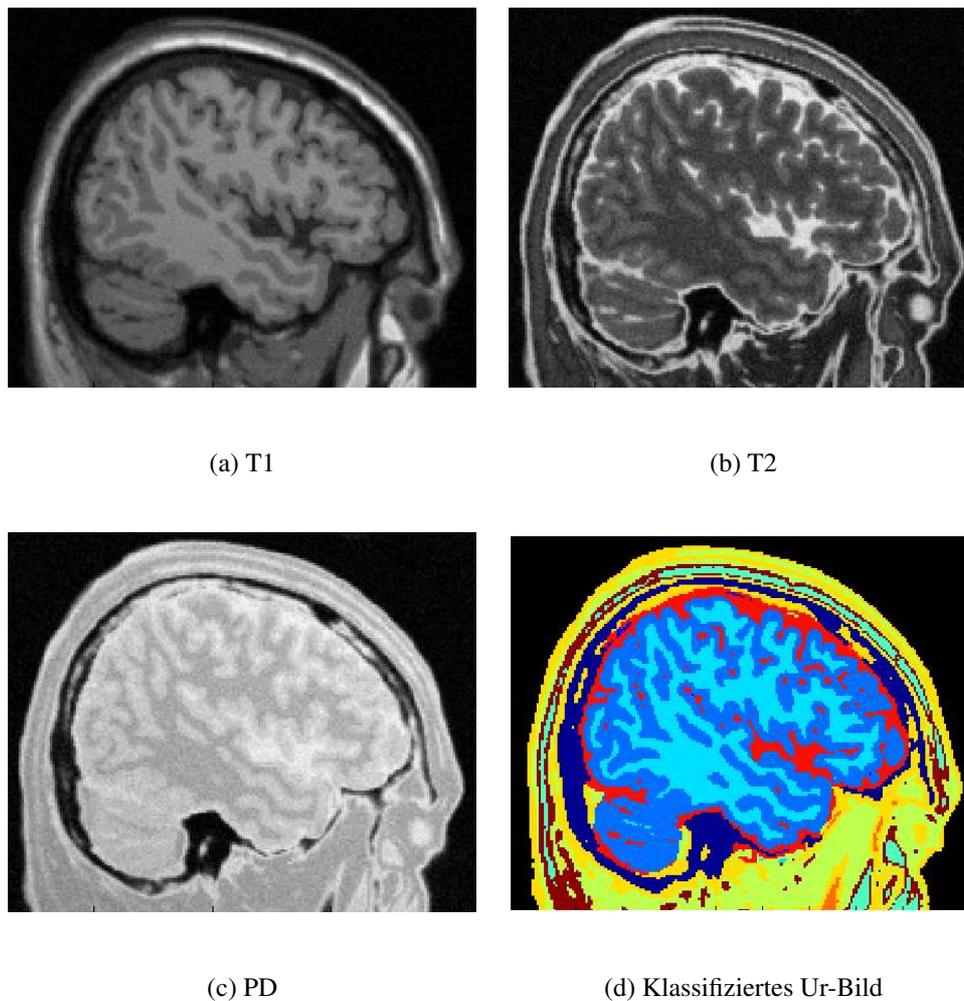


Abbildung 2.2: Beispielbilder der zur Klassifizierung verwendeten Spektren (a-c) und das Ergebnis der *LVQ*-Klassifizierung (d). Dabei lautet die Farbkodierung wie folgt: blau: GM; hellblau: WM; rot: Liquor u.ä.; gelb, grün: Binde-, Knochen-, Fett- und Muskelgewebe; schwarz: Luft

Im ersten Teil wird ausgenutzt, dass weiße Substanz nicht in direkter Nachbarschaft zu Knochengewebe vorkommt. Dabei wird der Datensatz voxelweise bearbeitet und für jedes Voxel, das als Knochen klassifiziert wurde, die Nachbarschaft nach weißer Substanz untersucht. Nachbarschaft bedeutet hier unmittelbare Nachbarschaft, so dass für jedes Voxel sechs nächste Nachbarn betrachtet werden. Findet sich ein Nachbar-Voxel aus weißer Substanz, wird angenommen, dass es sich um eine Fehlklassifikation handelt und das Voxel wird in „Knochen“ umbenannt. Zur Durchführung wird das *Matlab*-Skript `lvqcorr2` (Programm-Listings Nr. 12) benutzt.

Nach diesem Schritt werden aus dem klassifizierten Datensatz nur die für die weitere Bearbeitung notwendigen Gewebeklassen „weiße Substanz“ und „graue Substanz“ mit dem Skript `transform2bw3d` extrahiert. Außerdem macht man sich noch die Eigenschaft von grauer und weißer Substanz zu Nutzen, dass beide untereinander oder zu sich selbst eine hohe Konnektivität aufweisen, d. h. es ist sehr unwahrscheinlich, ein richtig als weiße oder graue Substanz klassifiziertes Voxel zu finden, dessen Menge aller unmittelbaren Nachbarn in x - und y -Richtung zwei oder weniger Voxel weißer oder grauer Substanz enthält. Werden solche Voxel vom Skript `deletepix` (Programm-Listing Nr. 13) erkannt, so werden sie als fehlerklassifiziert angesehen und gelöscht.

Das in den letzten beiden Absätzen beschriebene Vorgehen wird mit dem Skript `perfit` (Programm-Listing Nr. 11) durchgeführt, in dem die einzelnen Schritte zusammengefasst sind.

Connected-Component-Algorithmus

In den vorangegangenen Schritten wurde versucht, im Klassifizierungsvorgang entstandene Fehler zu korrigieren. Dies funktioniert nicht vollständig, jedoch erreicht man dadurch mit guter Zuverlässigkeit, dass das *falsche* keine Verbindung

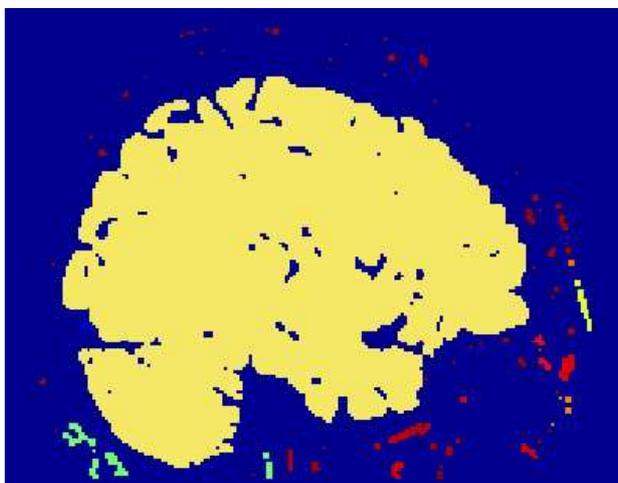


Abbildung 2.3: Dargestellt ist das Resultat nach Anwendung der Korrekturalgorithmen und des Connected-Component-Algorithmus. Die gelbe Komponente ist das Gehirn, die kleineren andersfarbigen Areale sind damit im dreidimensionalen nicht verbunden und daher falsch als Gehirn klassifiziertes Gewebe.

zum *richtigen* Gehirngewebe mehr besitzt. Sucht man nun im dreidimensionalen Datensatz nach allen verbundenen Komponenten (engl. connected components) des Datensatzes, so findet man einige kleinere und genau eine größere Komponente. Genau diese Komponente ist das Gehirn (siehe Abbildung 2.3).

Verwendet wird die in der *Matlab Image Processing Toolbox* bereitgestellte Funktion *bwlabeln*. Anschließend kann unmittelbar das Volumen von weißer und grauer Substanz berechnet werden. (siehe Skript `cv01` (Programm-Listing Nr. 14)).

2.2 Statistical Parametric Mapping *SPM*

2.2.1 Einleitung

Die Herangehensweise des *SPM* unterscheidet sich wesentlich von der des *LVQ*, da die einzelnen Voxel nun nicht mehr unabhängig von ihrer weiteren Umgebung

betrachtet werden, sondern Informationen über die relative Lage der Voxel im Datensatz in den Klassifizierungsvorgang einbezogen werden.

Dazu bedient sich das Programm eines aus 152 Gehirnen gemittelten Gehirndatensatzes, der in einem anatomischen Standardraum (hier: Talairach) vorliegt und als Vorlage (engl. *template*) dient. Der zu untersuchende Datensatz muss vom Benutzer an die Orientierung und die Lage des Nullpunktes der Vorlage angepasst werden. Vom Programm werden dann bei ausreichender Genauigkeit der manuellen Nullpunktwahl und Ähnlichkeit des Testgehirns mit der Vorlage die Daten durch die sog. „räumliche Normalisierung“ (engl. *spatial normalisation*) in den Raum des Standardgehirns abgebildet (Friston, 1995).

Dieses Vorgehen ermöglicht dann eine Bildklassifizierung, in die Informationen über die Position der Bildelemente in die statistische Analyse einbezogen werden (Ashburner und Friston, 1997)⁵.

2.2.2 Datenformat und Konvertierung

Die im `raw`-Format vorliegenden Daten müssen erst in ein für *SPM* lesbares Format konvertiert werden. In diesem Format gibt es eine Datei, die die eigentlichen Daten enthält (Suffix `img`) und eine weitere Datei (Suffix `hdr`), in der Informationen über den Aufbau des Datensatzes gespeichert sind. Die Konvertierung wird wie folgt durchgeführt.

- Die Daten im `raw`-Format werden mit dem Skript `readraw.m` (Programm-Listing Nr. 5) in *Matlab* eingelesen.
- Es werden dann mit dem Skript `afni2spm.m` (Listing Nr. 6) die beiden

⁵Eine genaue Beschreibung der *SPM* zugrundeliegenden Methoden und Dokumentation des Programmpakets finden sich unter <http://www.fil.ion.ucl.ac.uk/spm/dox.html>. Zur Verwendung von *SPM* benötigt man *Matlab*. Die hier benutzte Version ist *SPM99* und kann heruntergeladen werden unter <ftp://ftp.fil.ion.ucl.ac.uk/spm>.

SPM-Dateien erzeugt. Dabei wird das Skript `spm_hwrite.m` benutzt, das zum Programmpaket *SPM* gehört.

2.3 Durchführung

Nachdem die Daten in das geeignete Format transformiert worden sind, wird *SPM* von *Matlab* aus gestartet. Man wählt „fMRI time-series“ und öffnet damit die graphische Benutzeroberfläche. Bevor mit dem Klassifizierungsvorgang begonnen werden kann, muss der zu untersuchende Datensatz so modifiziert werden, dass Koordinatenursprung und räumliche Orientierung mit den *Template*-Gehirn-daten übereinstimmen. Um die entsprechenden Werte herauszufinden, wird der *Template*-Datensatz geöffnet, indem man über „Display“ aus dem Verzeichnis `template` die Datei `T1.img` auswählt. Jetzt kann man die graphische Darstellung und angegebenen Parameterwerte mit dem zu untersuchenden Datensatz vergleichen. Die Anpassung der Parameterwerte erfolgt durch die Modifizierung der „Header“-Information des Datensatzes (Suffix `hdr`) mit „HDR edit“ von der Benutzeroberfläche aus. Der Koordinaten-Ursprung wird auf die Werte $[x, y, z] = [92, 126, 72]$ gesetzt.

Danach wird der eigentliche Klassifizierungsvorgang mit „Segment“ gestartet. Nach Auswahl des Dateinamens gibt man an, dass eine räumliche Normalisierung („spatial normalization“) und eine starke Korrektur von Intensitätsinhomogenitäten durchgeführt werden soll. Sind diese Parameter gesetzt, beginnt *SPM* mit dem Klassifizierungsvorgang dessen Resultat (siehe Abbildung 2.4) in drei neuen Datensätzen für die Klassen „weiße Substanz“, „graue Substanz“ und „Liquor“ gespeichert wird. Die klassifizierte Daten werden in *Matlab* ausgewertet und das Volumen bestimmt. Dazu benutzt man `spmvol` (Programm-Listing Nr. 15).

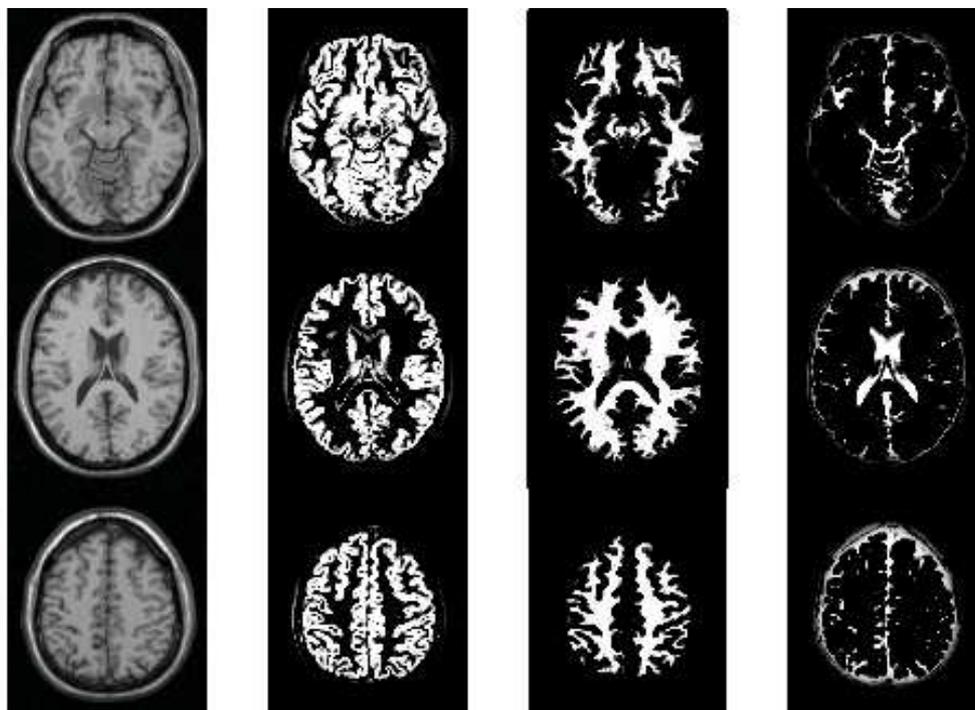


Abbildung 2.4: Dargestellt ist der T1-Ausgangs-Datensatz (links) und die Ergebnisse der Klassifizierung mit *SPM* (von links nach rechts): graue Substanz, weiße Substanz, Liquor.

2.4 Freesurfer

2.4.1 Einleitung

Man kann sich die Kortexoberfläche des Gehirns als zweidimensionale, gefaltete Struktur vorstellen. Mit dem Programmpaket *Freesurfer* wurde eine Methode entwickelt, die aus dreidimensionalen MRT-Datensätzen die Oberfläche des Gehirns rekonstruiert und die Darstellung der Oberflächen von weißer Substanz und Kortex ermöglicht (Dale et al., 1999). Dadurch ist auch die Berechnung der Volumina

von weißer und grauer Substanz möglich (Fischl et al., 1999)⁶.

Freesurfer kann im Gegensatz zu *SPM* unabhängig von *Matlab* benutzt werden. Allerdings muss auch hier die Vorbereitung der Datensätze und die Auswertung der Daten zur Volumenberechnung außerhalb der *Freesurfer*-Umgebung vorgenommen werden, was hier ebenfalls in *Matlab* gemacht wurde. Zu bemerken ist noch, dass sich mit *Freesurfer* nur Gehirnbereiche oberhalb des Pons und ohne das Cerebellum darstellen lassen, was in den folgenden Unterabschnitten beachtet werden muss, wenn von den Volumina von grauer und weißer Substanz die Rede ist.

2.4.2 Datenformat und Konvertierung

Es gibt mehrere mögliche Datenformate, die von *Freesurfer* lesbar sind. Hier wird das *AFNI*-Datenformat verwendet.

Da die Testdatensätze in einem einfachen Dateiformat (*raw*) vorliegen, das keine Hinweise über die Bildaufteilung enthält und daher nicht von *Freesurfer* gelesen werden kann, müssen sie erst in das *AFNI*-Format konvertiert werden.

Dazu werden die Daten erst in das *SPM*-Format konvertiert, wie in Kapitel 2.2.2 beschrieben ist. Danach wird wie folgt fortgefahren, indem das Programm `to3d` aus dem *AFNI*-Paket verwendet wird:

```
to3d <Dateiname>.hdr
```

Zur Konvertierung werden die Parameter „Datentyp“ und „Anatomietyp“ auf `3DIM_HEAD_ANAT` und `MRI Anatomy` gesetzt und ein geeigneter Dateiname gewählt. Man erhält so die beiden Dateien des *AFNI*-Formats.

⁶Das Programmpaket und eine ausführliche Dokumentation findet man unter <http://surfer.nmr.mgh.harvard.edu>.

2.4.3 Durchführung

im Folgenden sind die einzelnen Schritte zur Durchführung der Berechnung des Gehirnvolumens mit *Freesurfer* dargestellt.

Zu allererst werden die Datensätze wie in Unterkapitel "Datenformate und Konvertierungen" beschrieben für die Weiterverarbeitung vorbereitet. Dann wird die Hauptkonsole von *Freesurfer* mit `csurf` gestartet, ein neuer Datensatz angelegt (*File* → *NewSubject*) und die beiden Dateien im *AFNI*-Format in das entsprechende, vom Programm angelegte Unterverzeichnis kopiert (`<PFAD>/subjects/<NAME>/mri/orig/001/`).

Jetzt werden die strukturellen Informationen über die Bilder analysiert und die Daten in das *Freesurfer* eigene Format konvertiert (*SubjectTools* → *SetupStructuralScans*, ↘ *StructuralScanDir*, *convert/average*).

Bevor die weiteren Berechnungen gestartet werden, muss ein Fehler in der hier benutzten Version behoben werden, indem in der Datei `COR-.info` im Verzeichnis `<PFAD>/mri/orig` der Parameter `fov=0.256` gesetzt wird. Dann wird der Hauptvorgang gestartet, der aus folgenden vier einzelnen Prozessen besteht (*SubjectTools* → *Process/Create/Fix/Final*):

1. Volumenprozessierung (Process Volume)
2. Berechnung der Oberfläche (Create Surface)
3. Fehlerbehebung der Oberflächentopologie (Fix Surface Topology)
4. Berechnung der endgültigen Oberfläche (Make Final Surface)

Dabei ist Teilschritt 3 nicht obligatorisch, er dient nur zur Fehlerkorrektur von Fehlern in der Oberflächentopologie, die hier jedoch höchstens in nicht sichtbarem Ausmaß auftraten.

Die Aufnahmequalität der Daten ist allerdings schon vorher von Bedeutung und

kann bereits bei der Berechnung der Oberfläche (Schritt 2) Probleme verursachen, indem die Trennung von Cerebellum und Pons nicht einwandfrei funktioniert. Während dieser Vorgang bei Datensatz 1 automatisch ablief, fand bei Datensatz 2 eine unvollständige Abtrennung statt und die Position der Trennlinie musste manuell durch Suchen eines geeigneten Punktes vorgenommen werden. Dazu geht man gemäß *Freesurfer*-Anleitung vor. Die Berechnungen müssen in diesem Fall ab Teilschritt 2 wiederholt werden.

Als Resultat dieses Vorgangs ergeben sich zwei verschieden strukturierte Datensätze für weiße und graue Substanz.

Volumenberechnung: weiße Substanz

Die weiße Substanz wird während des Vorgangs *Create Surface* berechnet, die Daten zur Ableitung des Volumens der grauen Substanz sind Resultat des letzten Vorganges.

Die Daten der weißen Substanz werden im Verzeichnis `<PFAD>/mri/filled` gespeichert. Die Kodierung ist dabei im Gegensatz zum Datenformat der grauen Substanz konventionell, d. h. die Daten liegen als dreidimensionaler Block vor, in dem jedes Voxel, das zur weißen Substanz gehört, den Wert 255 erhält, alle anderen haben den Wert 0. Dieses Teilvolumen ergibt sich also aus der Summe aller Voxel, deren Werte größer als 0 sind.

Zur Berechnung des Volumens der weißen Substanz werden die in 256 (256×256)-Matrizen vorliegenden Daten in *Matlab* geladen und in einen dreidimensionalen Array der Größe ($256 \times 256 \times 256$) gespeichert. Dies überimmt das *Matlab*-Skript `corscan` (Programm-Listing Nr. 7). Zur Ermittlung des Volumens der weißen Substanz für linke und rechte Hemisphäre dient das Skript `volcount` (Programm-Listing Nr. 8).

Volumenberechnung: graue Substanz

Zur Berechnung des Volumens der grauen Substanz werden die vier Dateien `lh.thickness`, `rh.thickness`, `lh.area` und `rh.area` benötigt. Sie finden sich im Verzeichnis `<PFAD>/surf`. An diesem Dateiformat erkennt man die Besonderheit von *Freesurfer*. Es findet sich hier nur noch die Information über die Koordinaten der Oberfläche der weißen Substanz, die aus n kleinen Dreiecken verschiedener Fläche a zusammengesetzt ist (`lh.area`) und der Dicke d des Kortex für jedes dieser elementaren Dreiecke. Das Volumen der grauen Substanz erhält man nun einfach als Summe über die Produkte der Teilflächen mit den entsprechenden Dicken:

$$Volumen_{GM} = \sum_{i=1}^n a_i d_i \quad (2.1)$$

Nach Abschluss der bisher besprochenen Prozesse werden die Daten in *Matlab* eingelesen, wo dann die Volumenberechnung nach Gleichung (2.1) durchgeführt wird. Dazu wird ein Hilfsskript benutzt, das Daten im Curvature-Format von *Freesurfer* direkt in *Matlab* lädt, das Bestandteil der EEG-Toolbox⁷ für *Matlab* ist.

Die hier dargestellten Vorgänge übernimmt das *Matlab*-Skript `calcvol` (Programm-Listing Nr. 9), das direkt die Volumenwerte der grauen Substanz für rechte und linke Hemisphäre ausgibt.

2.5 AFNI

Bei *AFNI* handelt es sich um ein weiteres Programmpaket, das zur Analyse von anatomischen und funktionellen MRT-Daten benutzt werden kann⁸. Es findet hier allerdings nur Erwähnung, da in dieser Arbeit das *AFNI*-Datenformat benutzt

⁷erhältlich unter <http://sourceforge.net/projects/eeg/>

⁸Details und Download unter <http://afni.nimh.nih.gov/afni/>

wird, das von verschiedenen anderen Programmen gelesen und konvertiert werden kann. Dabei sind die eigentlichen Daten in einer Datei (Suffix `BRIK`) gespeichert und in einer weiteren Datei (Suffix `HEAD`) werden Informationen über das Format abgelegt. Benutzt wird hier die Version 2.40e (Cox und Hyde, 1997).

2.6 Datenmaterial

Als Datenmaterial werden simulierte MRT-Bilddatensätze verwendet, die mit Hilfe des *MRI Simulators* vom *McConnell Brain Imaging Centre* erstellt worden sind. Dabei handelt es sich um Datensätze, die auf einem anatomischen Modell eines normalen aus realen MRT-Daten gemittelten Gehirns basieren, in dem die unterschiedlichen Gewebetypen eindeutig klassifiziert sind. Im Simulationsvorgang werden die Werte der einzelnen Voxel so verändert, dass der resultierende Datensatz dem eines im Magnetresonanztomographen gemessenen äquivalent ist. Neben den Fehlerwerten „Rauschen“ und „Intensitätsinhomogenität“ werden auch Partialvolumeneffekte realistisch simuliert (Collins et al., 1998; Kwan et al., 1999). Die in Prozent angegebenen Werte sind folgendermaßen zu interpretieren: Der Wert für „Rauschen“ gibt das prozentuelle Verhältnis der Standardabweichung des weißen gaussverteilten Rauschen zum Signal des Referenzgewebes an und „Intensitätsinhomogenität“ bezeichnet den Wertebereich des Intensitätsinhomogenitätsfeldes (z. B. bedeutet 20%, dass Werte zwischen 0,90 und 1,10 vorkommen).

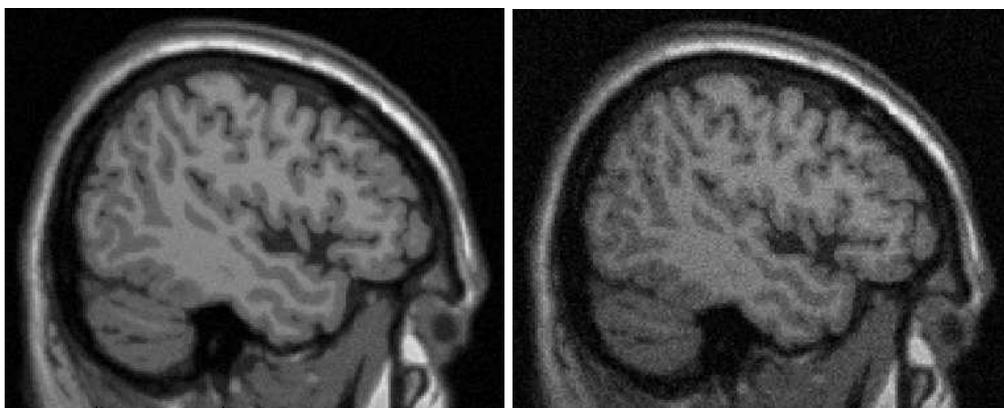
Es wurden T1-, T2- und PD(=Protonendichte)-gewichtete Datensätze der Schichtdicke 1 mm mit zwei verschiedenen Werten für die Parameter „Rauschen (Noise)“ und „Intensitätsinhomogenität (intensity non-uniformity, RF)“ gewählt, und zwar:

	Datensatz 1	Datensatz 2
Rauschen	3%	7%
Intensitätsinhomogenität	20%	40%
Bezeichnung	<i>n3rf20</i>	<i>n7rf40</i>

Beide Datensätze werden jeweils mit oder ohne dem Klassifikationsvorgang vorausgehende Anwendung des in *SPM* enthaltenen Inhomogenitätskorrekturalgorithmus verwendet, so dass insgesamt vier qualitativ verschiedene Datensätze zur Verfügung stehen, wobei Datensatz 1 relativ niedrige und Datensatz 2 relativ hohe Werte für Rauschen und Intensitätsinhomogenitäten aufweist (siehe Abbildung 2.5). Da die mit *BrainWeb* simulierbaren Parameter die real vorkommenden Werte abdecken (Collins et al., 1998; Kwan et al., 1999), kann davon ausgegangen werden, dass mit dieser Auswahl die am häufigsten vorkommenden realen Messvorgänge angemessen repräsentiert sind und sich die Daten für eine Bestimmung der Robustheit der verschiedenen Methoden gegenüber unterschiedlicher Bildqualität eignen.

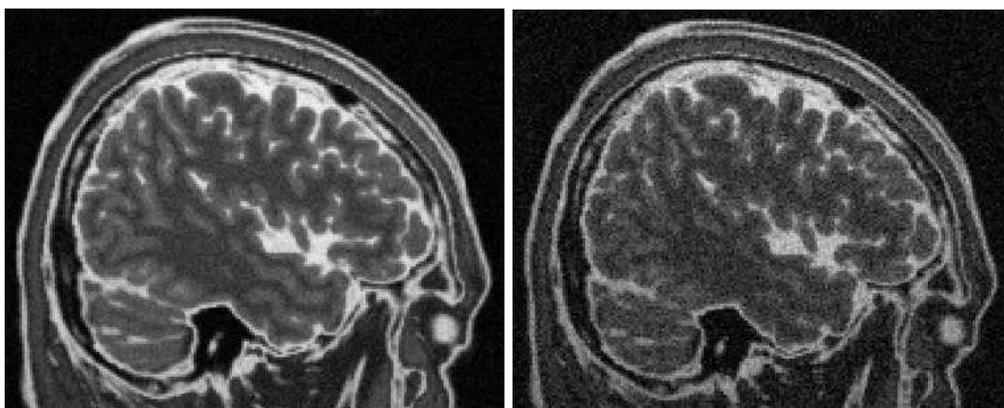
Die Datensätze haben eine Größe von $(181 \times 217 \times 181)$ Bildpunkten in 8 Bit Graustufen. Allerdings liegen sie in einem Dateiformat vor, das keine Information über die Aufteilung der Bildpunkte auf die verschiedenen Koordinaten enthält. Daher müssen die Daten hier zur Benutzung erst in ein Format konvertiert werden, das sie für die Programmpakete lesbar macht (siehe Kapitel 2.2.2)⁹.

⁹Nähere Informationen und die Möglichkeit zum Download findet man im Internet unter <http://www.bic.mni.mcgill.ca/brainweb/>. Zu beachten ist, dass die Daten mit GNUzip komprimiert sind, aber das Suffix `.gz` fehlt!



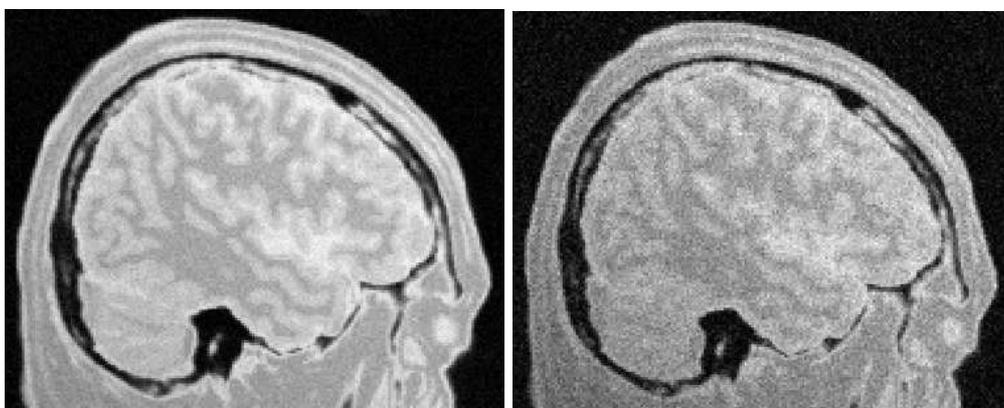
(a) T1 n3rf20

(b) T1 n7rf40



(c) T2 n3rf20

(d) T2 n7rf40



(e) PD n3rf20

(f) PD n7rf40

Abbildung 2.5: Vergleich der Datensätze 1 und 2

2.7 Computersystem und Software

Die Berechnungen wurden mit einem Intel Pentium III 1GHz/512MB RAM mit dem Betriebssystem SuSE Linux 7.2 (Kernel Version 2.4.4) durchgeführt. Die in dieser Arbeit erstellten Programme und Hilfsskripte sind in *Matlab 5.3/6.0* und der Unix-Shell-Sprache *csh* programmiert worden.

Kapitel 3

Ergebnisse

3.1 BrainWeb

Tabelle 3.1 enthält die Werte für die (Teil-)Volumina des den Simulationen zugrundeliegenden Datensatzes, die als Referenzwerte benutzt werden.

Tabelle 3.1: *BrainWeb*-Volumina in 10^5 mm^3 für das gesamte Gehirn (*BrainWeb*) und für das Gehirn ohne Kleinhirn (*BrainWebnc*).

Datensatz	Volumen GM	Volumen WM	Volumen CSF	Volumen GM+WM
BrainWeb	9,0291	6,7478	3,7195	15,7769
BrainWebnc	7,3435	6,3724	—	13,7159

3.2 LVQ

Die Ergebnisse der Berechnungen der Volumina für graue, weiße und weiße+graue Substanz und Liquor mit *LVQ* inklusive Nachbearbeitung sind in Tabelle 3.2 aufgelistet und in Abbildung 3.1 dargestellt. Abbildung 3.2 zeigt das Klassifikationsergebnis im Vergleich zu den Referenzdaten und das Ergebnis der Berechnung der größten 3D-verbundenen Komponente.

Für die graue Substanz ergibt sich eine Streuung der Werte von -11,4 bis +4,4 Prozent, wobei der Datensatz *n3rf20* mit Vorkorrektur der Intensitätsinhomogenitäten mit 2,6 Prozent die geringste Abweichung vom Referenzwert zeigt.

Die Abweichungen der weißen Substanz bewegen sich im Bereich von -41,4 bis 4 Prozent. Die geringste hat ebenfalls der Datensatz *n3rf20+c* mit -1,3%, die größte der Datensatz mit starkem Rauschen und großen Intensitätsinhomogenitäten (*n7rf40*).

Bei den Summen aus grauer und weißer Substanz erhält man eine Streuung von -24,1 bis 1,3 Prozent.

In diesen drei Fällen ist die Rangfolge der Beträge der Abweichungen gleich und korreliert mit der Bildqualität, für die Spearman'schen Korrelationskoeffizienten der Abweichungsbeträge errechnet sich für alle Daten $r_{SP} = 1,0$. Der Datensatz der relativ guten Qualität mit Vorkorrektur (*n3rf20+c*) weicht am geringsten ab, gefolgt von *n3rf20*. Wiederum deutlich schlechter (ca. Faktor 2) schneidet der Datensatz *n7rf40+c* ab. Das Schlußlicht bildet der Datensatz relativ schlechter Qualität ohne Vorkorrektur (*n7rf40*).

Die Abweichungen der Volumina des Liquor sind deutlich größer und korrelieren nicht positiv mit der Bildqualität. Die geringste Abweichung ergibt sich sogar für den schlechtesten Datensatz, die größte Abweichung für den besten, man erhält $r_{SP} = -1,0$.

Tabelle 3.2: *LVQ*-Ergebnis in 10^5 mm^3 bei Benutzung der Bildsequenzen T1, T2 und PD; die Werte in Klammern geben die prozentuale Abweichung von den Referenzwerten an. Bei den Datensätzen „+c“ wurde vor der Klassifizierung eine Reduktion der Intensitätsinhomogenitäten mit dem Korrekturalgorithmus von *SPM* durchgeführt (GM: graue Substanz; WM: weiße Substanz)

Datensatz	GM	WM	WM+GM	Liquor
n3rf20	9,4267 (4,4)	6,5591 (-2,8)	15,9858 (1,3)	2,2880 (-38,5)
n3rf20+c	9,2660 (2,6)	6,6592 (-1,3)	15,9252 (0,9)	1,9849 (-46,6)
n7rf40	7,9965 (-11,4)	3,9781 (-41,1)	11,9746 (-24,1)	3,0413 (-18,2)
n7rf40+c	8,2633 (-8,5)	7,0355 (+4,3)	15,2988 (-3,0)	4,5795 (23,1)

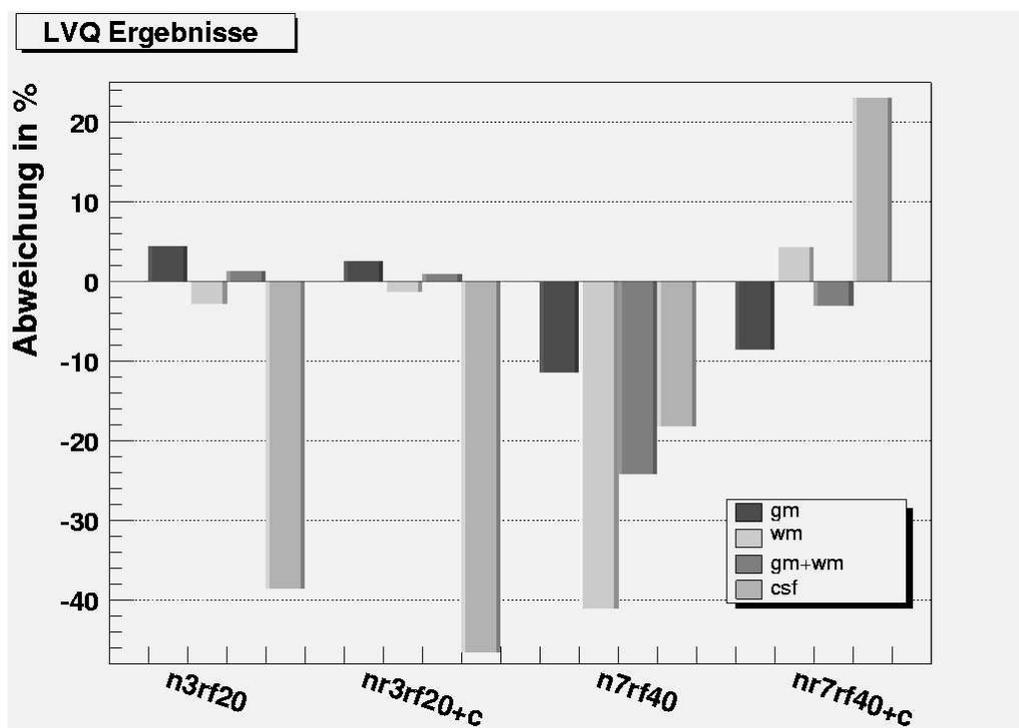
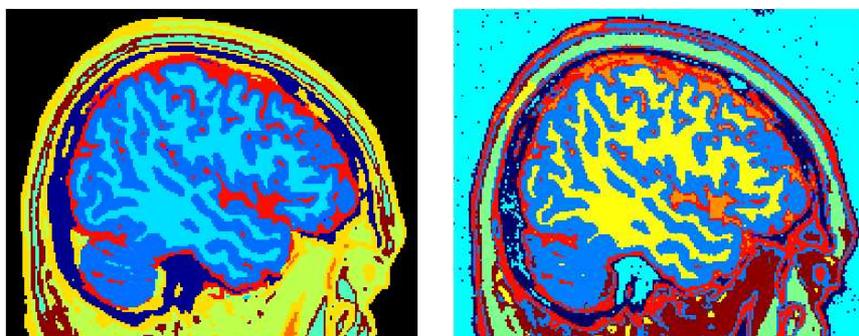
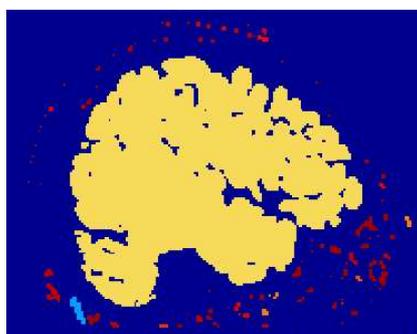


Abbildung 3.1: LVQ-Ergebnisse: Abweichungen der Volumina von grauer und weißer Substanz von den Referenzwerten in Prozent für die Datensätze *n3rf20*, *n3rf20+c*, *n7rf40*, *n7rf40+c* (+c: Vorkorrektur von Intensitätsinhomogenitäten)



(a) BrainWeb

(b) LVQ-Ergebnis



(c) größte verbundene Komponente

Abbildung 3.2: Beispielbilder des LVQ-Klassifizierungsergebnisses [b] im Vergleich zu den Ursprungsdaten aus *BrainWeb* [a] (vgl. Abbildung 2.2). Farbkodierungen in [b] sind wie folgt: *blau*: graue Substanz; *gelb*: weiße Substanz; *orange*: Liquor; *rot*: Auge; *grün*: Knochen; *braun*: Muskelgewebe; *hellblau*: Luft; *dunkelblau*: Bindegewebe. In [a] sind die relevanten Gewebetypen so farbkodiert: *blau*: graue Substanz; *hellblau*: weiße Substanz; *rot*: Liquor. In [c] ist das Ergebnis der Berechnung der in 3D verbundenen Komponenten dargestellt.

3.3 SPM

Die Ergebnisse der Berechnungen mit *SPM* für die drei verschiedenen Kombinationen T1, T1+T2 und T1+T2+PD mit und ohne Inhomogenitäts-Korrektur sind in Tabelle 3.4 aufgelistet und in den Abbildungen 3.4 mit und 3.5 ohne Korrektur dargestellt. In Abbildung 3.3 sind graphische Beispiele der Ergebnisse für einen sagittalen Schnitt dargestellt, wobei für jede Klasse (GM, WM und CSF) ein Datensatz erstellt wird, dessen Voxel-Werte die Wahrscheinlichkeiten für das Vorliegen der jeweiligen Klasse enthalten.

Für die graue Substanz findet sich eine Streuung von -7,4 bis +8,0%, Werte für die weiße Substanz liegen zwischen -11,8 bis -7,4%, die CSF-Werte von -19,1 bis +20,5%. Untersucht man die Korrelation zwischen Bildqualität (abnehmende Qualität: $n3rf20+c^1$, $n3rf20$, $n7rf40+c$, $n7rf40$) und Größe der Beträge der Abweichungen anhand des Spearman'schen Rangkorrelationskoeffizienten, so kommt man zu den in Tabelle 3.3 dargestellten Ergebnissen.

Tabelle 3.3: Korrelationskoeffizient r_{SP}

Bildsequenz	GM	WM	WM+GM	CSF
T1	-1,0	-1,0	0,8	1,0
T1+T2	-0,11	0	-0,74	-0,8
T1+T2+PD	-0,8	1,0	-0,6	-0,8

Man erkennt keinen einheitlichen Trend eines Zusammenhanges zwischen Bildqualität und Betrag der Abweichung von den Referenzwerten, im T1-Spektrum nehmen sogar die Abweichungen mit schlechter werdender Bildqualität ab. Lediglich beim Liquor für das T1-Spektrum und bei der weißen Substanz (T1+T2+PD)

¹„+c“: mit Inhomogenitätskorrektur

findet man $r_{SP} = 1$. Besonders deutlich zeigt sich die Uneinheitlichkeit der Korrelationen an der weißen Substanz, bei der der Korrelationskoeffizient r_{SP} die Werte -1, 0 und +1 für T1, T1+T2 und T1+T2+PD annimmt.

Tabelle 3.4: SPM-Berechnungen der Volumina in 10^5 mm^3 bei verschiedenen Kombinationen der Bildsequenzen. Die Berechnungen wurden jeweils mit und ohne Korrektur der Intensitätsinhomogenitäten durchgeführt. Die Werte in Klammern sind die relativen Abweichungen vom Referenzwert in Prozent (GM: graue Substanz; WM: weiße Substanz; „+c“: mit Inhomogenitätsvorkorrektur).

Datensatz	Sequenz	GM	WM	GM+WM	Liquor
n3rf20+c	T1	9,7508 (+8,0)	5,9515 (-11,8)	15,702 (-0,5)	3,4459 (-7,4)
	T1+T2	8,8230 (-2,3)	6,0696 (-10,1)	14,893 (-5,6)	4,2628 (+14,6)
	T1+T2+PD	8,6220 (-4,5)	6,2480 (-7,4)	14,870 (-5,7)	4,2410 (+14,0)
n7rf40+c	T1	9,5060 (+5,3)	6,1163 (-9,4)	15,622 (-1,0)	3,0267 (-18,6)
	T1+T2	9,1773 (+1,6)	6,1121 (-9,4)	15,289 (-3,1)	3,5217 (-5,3)
	T1+T2+PD	9,3448 (+3,5)	6,1732 (-8,5)	15,518 (-1,6)	3,6203 (-2,7)
n3rf20	T1	9,7167 (+7,6)	5,9655 (-11,6)	15,682 (-0,6)	3,2326 (-13,1)
	T1+T2	8,8396 (-2,1)	5,9845 (-11,3)	14,824 (-6,0)	4,2787 (+15,0)
	T1+T2+PD	8,3608 (-7,4)	6,1961 (-8,2)	14,557 (-7,7)	4,4826 (+20,5)
n7rf40	T1	9,4221 (+4,4)	6,2135 (-7,9)	15,636 (-0,9)	3,0079 (-19,1)
	T1+T2	9,2329 (+2,3)	6,0565 (-10,2)	15,289 (-3,1)	3,5849 (-3,6)
	T1+T2+PD	9,1397 (+1,2)	6,0797 (-9,9)	15,219 (-3,5)	3,7337 (+0,4)

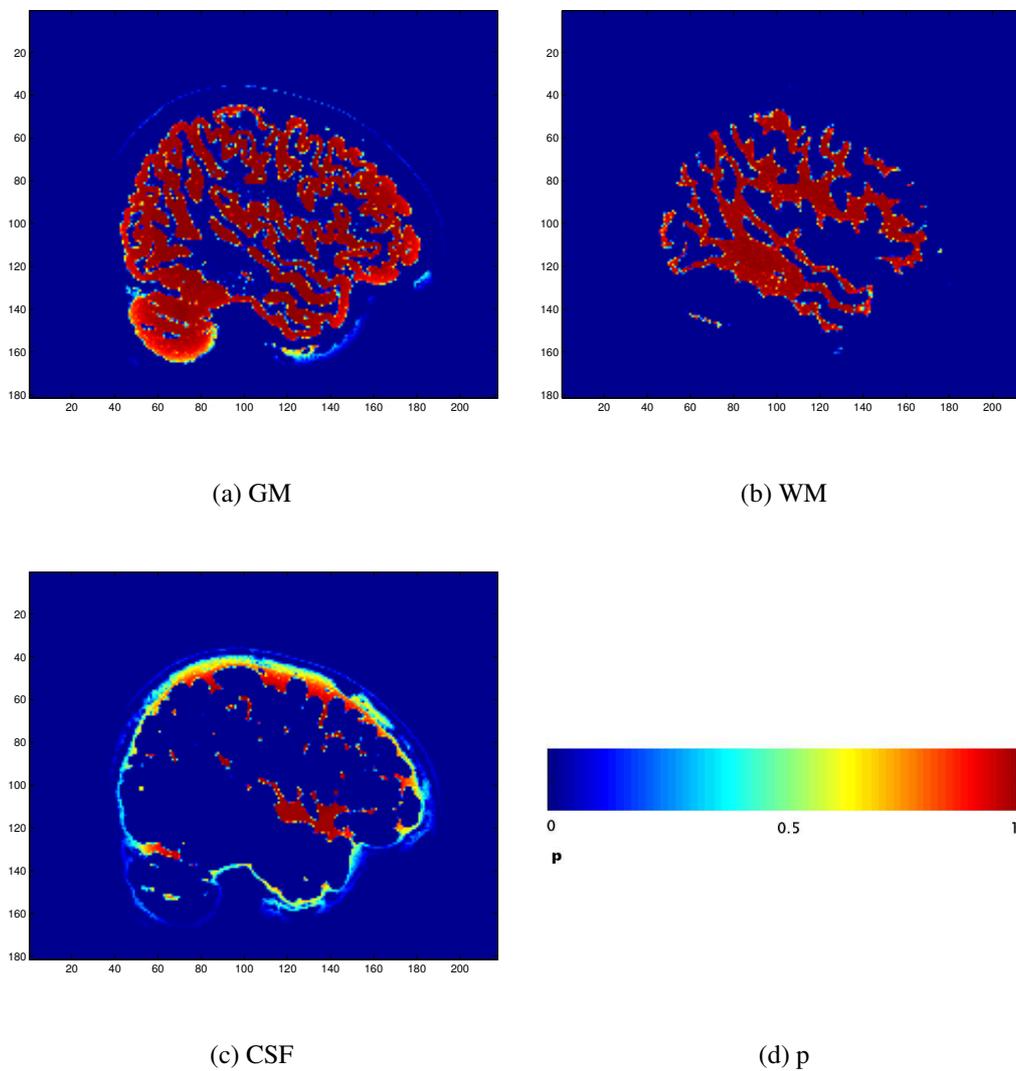


Abbildung 3.3: Beispielbilder (Sagittalschnitt $z = 44$) für die Ergebnisse der *SPM*-Klassifikation für graue und weiße Substanz und Liquor. Die Bildelemente sind mit den Wahrscheinlichkeiten für das Vorliegen des jeweiligen Gewebetyps gewichtet, von „dunkelblau“ ($p = 0$) bis „dunkelrot“ ($p = 1$) (siehe Abbildung (d)).

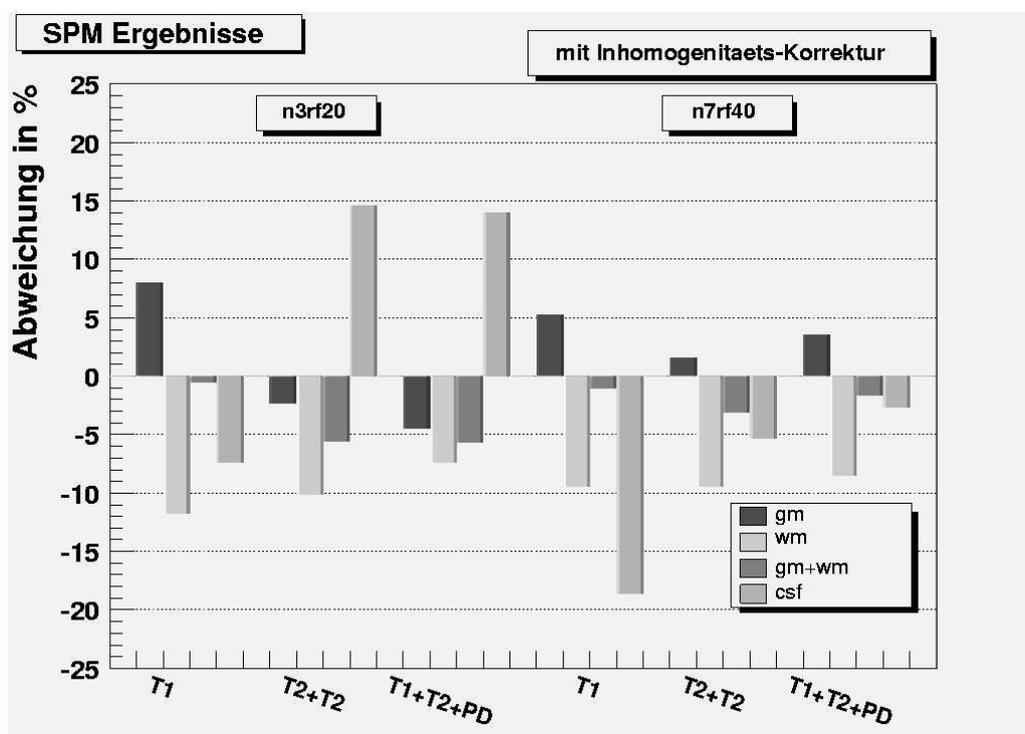


Abbildung 3.4: SPM-Ergebnisse: Abweichungen der Volumina von grauer und weißer Substanz und Liquor von den Referenzwerten in Prozent für Datensätze *n3rf20* und *n7rf40*, jeweils für drei verschiedene Kombinationen der Spektren T1, T2 und PD *mit* Korrektur der Intensitätsinhomogenitäten.

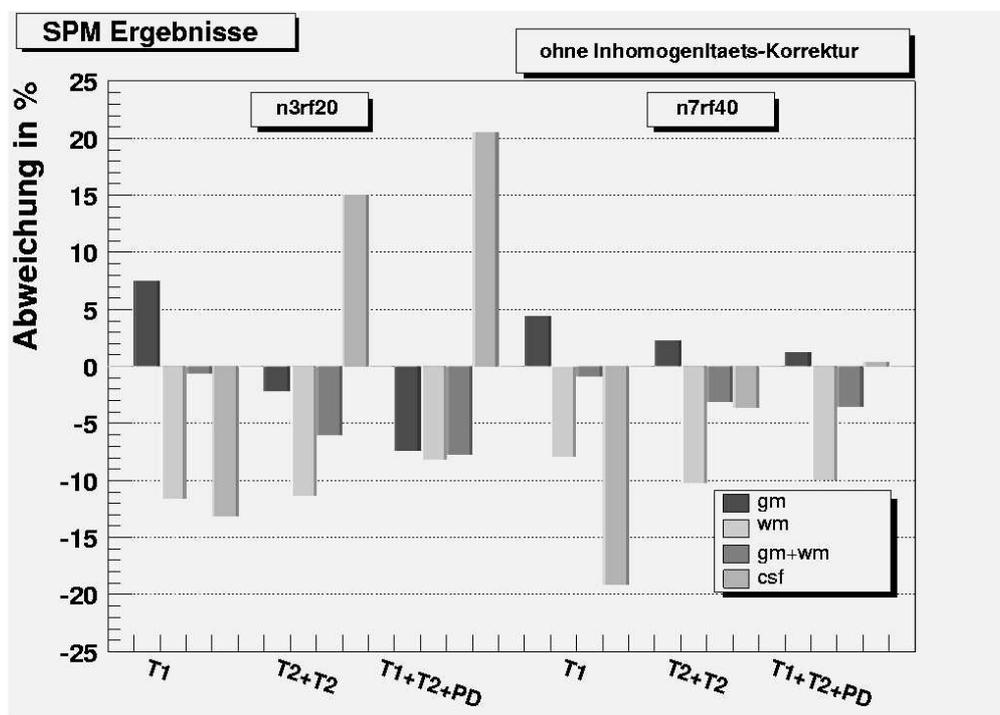


Abbildung 3.5: SPM-Ergebnisse: Abweichungen der Volumina von grauer und weißer Substanz und Liquor von den Referenzwerten in Prozent für Datensätze *n3rf20* und *n7rf40*, jeweils für drei verschiedene Kombinationen der Spektren T1, T2 und PD *ohne* Korrektur der Intensitätsinhomogenitäten.

3.4 Freesurfer

Die Ergebnisse sind graphisch in Abbildung 3.7 als dreidimensionale Oberflächen dargestellt, während in Abbildung 3.8 in einem sagittalen Schnitt die zweidimensionalen Projektionen der Oberflächen von weißer und grauer Substanz zu sehen sind.

Die Volumina für die graue Substanz sind für alle Bildqualitäten (abnehmend von links nach rechts: *n3rf20*, *n3rf20nofix*, *n7rf40*, *n7rf40nofix*) unterbestimmt und

schwanken zwischen -6,6 und -3,0 Prozent. Für die weiße Substanz erhält man nur zwei Werte, da die Behebung von Oberflächenfehlern nur auf die graue Substanz angewendet wird. Im Gegensatz zur grauen wird die weiße Substanz zu hoch berechnet, so dass sich für die Summe aus grauer und weißer Substanz relativ wenig abweichende Werte ergeben (-0,7 bis 2,0%). Eine vollständige Zusammenstellung der Werte findet sich in Tabelle 3.5 und in Abbildung 3.6. Aus den Berechnungen des Spearman'schen Rangkorrelationskoeffizienten r_{SP} erkennt man, dass kein Zusammenhang zwischen Bildqualität und Genauigkeit der Volumenberechnung festgestellt werden kann, es ergeben sich folgende Werte:

	GM	WM	WM+GM
r_{SP}	0	-0,89	0,4

Tabelle 3.5: *Freesurfer*-Volumina in 10^5 mm^3 . *Mit* oder *ohne* (nofix) Korrektur von Fehlern in der Oberflächentopologie. Die Werte in Klammern bezeichnen die prozentualen Abweichungen vom Referenzwert.

Datensatz	Volumen GM	Volumen WM	GM+WM
n3rf20	6,9341 (-5,6)	6,8609 (+7,7)	13,7950 (+0,6)
n3rf20nofix	7,1259 (-3,0)	6,8609 (+7,7)	13,9868 (+2,0)
n7rf40	6,8616 (-6,6)	6,7558 (+6,0)	13,6174 (-0,7)
n7rf40nofix	7,0642 (-3,8)	6,7558 (+6,0)	13,8200 (+0,8)

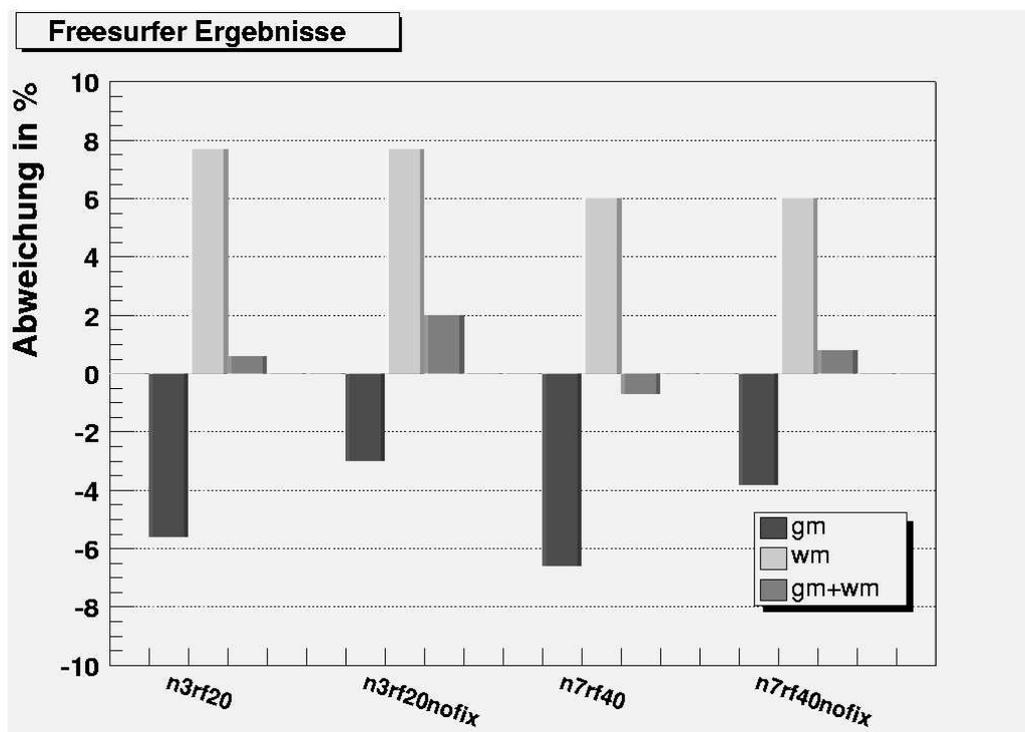
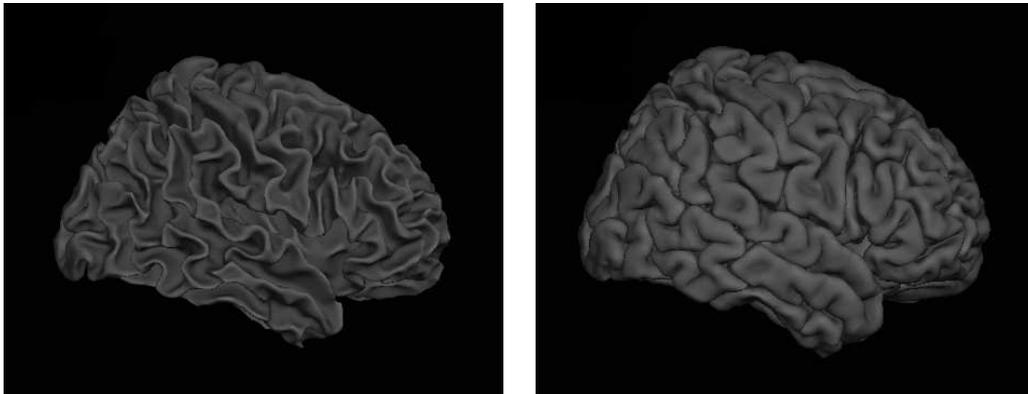


Abbildung 3.6: *Freesurfer*-Ergebnisse: Abweichung der Volumina von grauer und weißer Substanz von den Referenzwerten in Prozent für die Datensätze *n3rf20* und *n7rf40* mit und ohne (nofix) Korrektur von Fehlern in der Oberflächentopologie.



(a) weiße Substanz

(b) graue Substanz

Abbildung 3.7: 3D-Darstellung der Oberflächen von weißer und grauer Substanz mit *Freesurfer*.

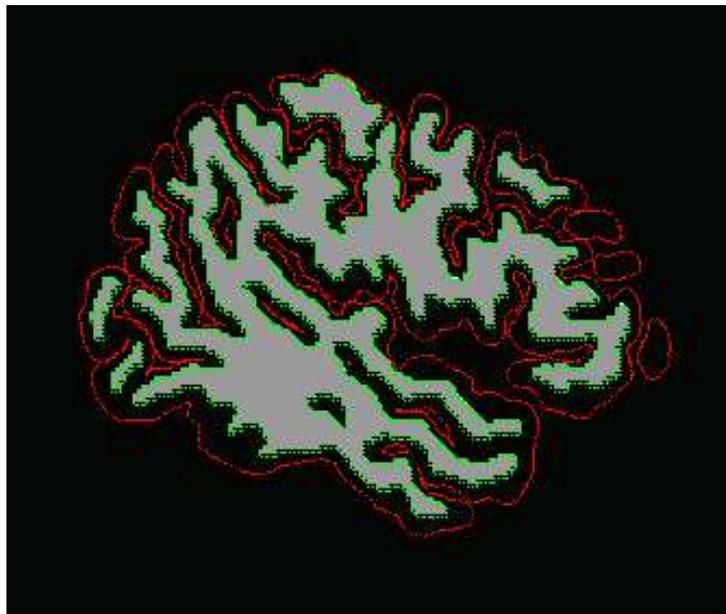


Abbildung 3.8: Sagittalschnitt mit Darstellung der Oberflächen von weißer (grün) und grauer (rot) Substanz. Grau dargestellt ist das als weiße Substanz klassifizierte Gewebe.

3.5 Vergleich der Ergebnisse

Abbildung 3.9 zeigt einen Vergleich der Ergebnisse der drei verschiedenen Methoden.

3.5.1 Graue Substanz

Für die graue Substanz ergeben sich beim Datensatz *n3rf20* positive Abweichungen für *LVQ*, während *Freesurfer* das Volumen zu niedrig berechnet. Mit *SPM* erhält man je nach verwendeten Bildsequenzen positive oder negative Abweichungen. Beim Datensatz *n7rf40* verstärken sich die Abweichungen für *LVQ* (Faktor ≈ 2 , negativ Abweichung) und *Freesurfer*, während die Berechnungen mit *SPM* geringere Abweichungen (positiv für alle Spektren) ergeben.

3.5.2 Weiße Substanz

Die Berechnung des Volumens der weißen Substanz von Datensatz *n3rf20* mit *LVQ* ergibt die kleinsten Abweichungen ($< 3\%$). Mit *SPM* erhält man hier – weitgehend unabhängig von der verwendeten Bildsequenz-Kombination – ein zu geringes Volumen (Abweichung $\approx -10\%$), während *Freesurfer* zu hohe Werte berechnet ($< 8\%$). Datensatz *n7rf40* ergibt ähnliche Werte, nur das *LVQ*-Resultat ohne Vorkorrektur der Intensitätsinhomogenitäten ergibt ein extrem zu niedrig bestimmtes Volumen. Mit Korrektur ergibt sich ein etwas zu hoch bestimmter Wert. Vergleicht man bei *SPM* die Volumenwerte von Datensatz *n7rf40* mit und ohne Intensitätsinhomogenitätskorrektur, finden sich bei alleiniger Verwendung der T1-gewichteten Bildsequenz Abweichungen mit entgegengesetztem Vorzeichen, so dass eine Diskrepanz von ca. 18% entsteht.

3.5.3 Graue + weiße Substanz

Bei den Summen aus weißer und grauer Substanz (*n3rf20*) ergeben sich die besten Werte (Abweichung $< 1\%$) aus *LVQ* mit Korrektur, aus *SPM* mit und ohne Korrektur bei alleiniger Verwendung von T1-gewichteten Daten und aus *Freesurfer* mit Oberflächenkorrektur. Bei Datensatz *n7rf40* erhält man nur mit *Freesurfer* Abweichungen, die kleiner sind als 1%.

3.5.4 Zusammenfassung

Obwohl es schwierig ist, eine allgemeine Aussage über die Güte einer Methode zu treffen, da die Berechnungen für weiße, graue bzw. weiße+graue Substanz sehr unterschiedliche Abweichungen aufweisen, soll im Folgenden mit Hilfe eines einfachen statistischen Parameters (mittlere Abweichung $\bar{\delta}$) untersucht werden, welche Methode *im Mittel* für welchen Datensatz die genauesten Ergebnisse liefert. Dabei ist $\bar{\delta}$ der Mittelwert der Beträge der prozentualen Abweichungen der drei Werte (GM, WM und GM+WM) von den Referenzwerten und wird für jeden Datensatz und für jede Methode berechnet.

Anhand der Ergebnisse in Tabelle 3.6 erkennt man, dass *LVQ* für Daten relativ guter Qualität die geringsten $\bar{\delta}$ aufweist, während für die relativ schlechteren Datensätze *SPM* und *Freesurfer* bessere Ergebnisse liefern. Außerdem ergeben sich für alle untersuchten Datensätze kleinere mittlere Abweichungen für *Freesurfer* als für *SPM* und die Abweichungen sind bei *Freesurfer* am stabilsten gegenüber Bildqualitätsschwankungen.

Tabelle 3.6: Vergleich der mittleren Abweichungen $\bar{\delta}$ von den Referenzwerten in Prozent. Datensatznamen mit Suffix „corr” bezeichnen dabei jeweils die mit dem für die jeweilige Methode verwendeten Vorkorrekturalgorithmus bearbeiteten Datensätze. Für *SPM* liest sich die Tabelle folgendermaßen: $\bar{\delta}(T1+T2+PD)/\bar{\delta}(T1)$.

$\bar{\delta}$	<i>LVQ</i>	<i>SPM</i>	<i>Freesurfer</i>
n3rf20.corr	1,6	5,9/6,8	4,6
n3rf20	2,8	7,8/6,6	4,2
n7rf40.corr	5,3	4,5/5,2	4,4
n7rf40	25,5	4,9/4,4	3,5

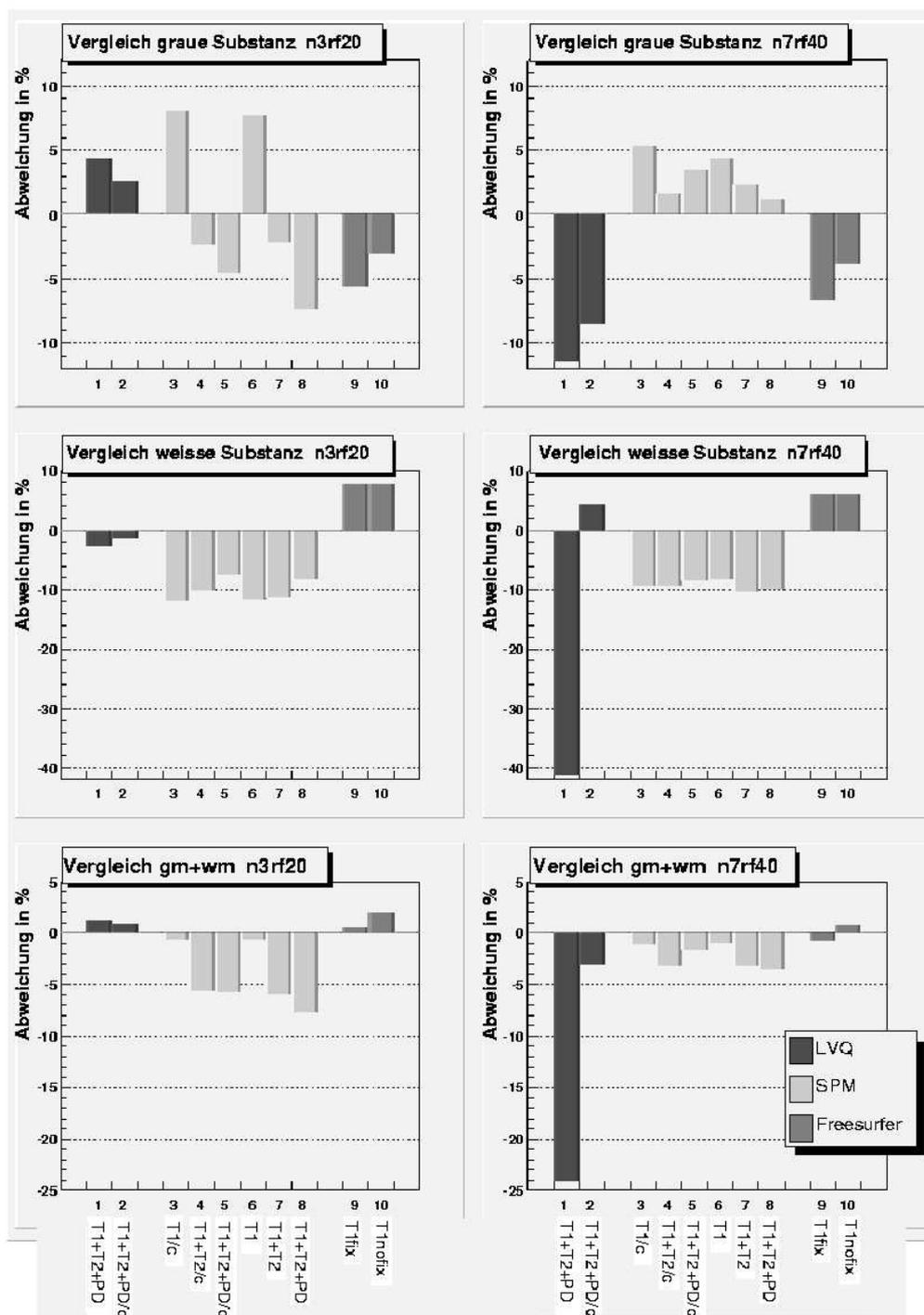


Abbildung 3.9: Vergleich der Ergebnisse für weiße, graue und Summe aus weißer und grauer Substanz; „/c“: mit Inhomogenitäts-Vorkorrektur; „fix“ bzw. „nofix“: mit bzw. ohne Oberflächenfehlerkorrektur.

3.5.5 Liquor

Die größten Abweichungen von den Referenzwerten insgesamt ergeben sich bei den Volumina des Liquors. Wiederum findet man hier die größten Abweichungen bei *LVQ* mit Datensatz *n3rf20*. Durchschnittlich kleinere Abweichungen erhält man bei Datensatz *n7rf40*, die allerdings für *LVQ* immer noch in der Größenordnung 20% liegen. Lediglich mit *SPM* kann bei *n7rf40* unter Verwendung der Kombination T1+T2+PD mit Korrektur der Fehler deutlich (bis unter 1%) reduziert werden (siehe Abbildung 3.10). Für *Freesurfer* liegen keine Daten vor.

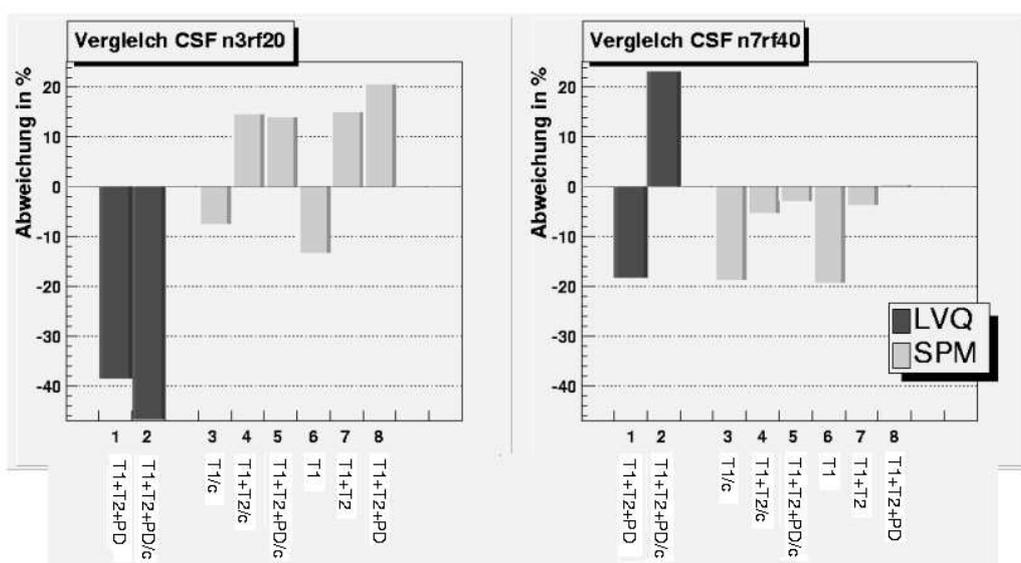


Abbildung 3.10: Vergleich der Ergebnisse für Liquor. „/c“: mit Inhomogenitäts-Vorkorrektur.

Kapitel 4

Diskussion

Die Gehirnvolumetrie findet in der medizinischen Forschung und Praxis zahlreiche Anwendungen. Sie ermöglicht die quantitative Analyse sowohl der die Kindesentwicklung als auch das Altern begleitenden gehirnmorphologischen Veränderungen und hilft bei der Suche nach Zusammenhängen zwischen Krankheiten und morphologischen Auffälligkeiten (Pfefferbaum et al., 1994; Giedd et al., 1996; Kidron et al., 1997; Coffey et al., 1998). Daraus kann man nicht nur Hinweise auf Krankheitsursachen gewinnen, sondern solche Erkenntnisse können auch in der klinischen Praxis diagnostisch zur Früherkennung und Verlaufskontrolle von Krankheiten genutzt oder in Therapiestudien als objektive quantitative Parameter zur Erfolgsbeurteilung herangezogen werden. Letzteres kann besonders im Zuge der evidenzbasierten Medizin Bedeutung erlangen (Evidence-Based Radiology Working Group, 2001).

Dabei reicht das Spektrum der möglichen Anwendungen von der Detektion einer Zu- oder Abnahme des gesamten Gehirnvolumens oder der Teilvolumina von grauer und weißer Substanz oder des Liquorraumes bis zur Vermessung von pathologischen Gewebeveränderungen, bedingt durch z. B. entzündliche Prozesse oder Tumoren.

Zur gewebedifferenzierenden Gehirnvolumetrie *in vivo* ist die Magnetresonanztomographie aufgrund ihres hohen Weichteilkontrastes und der fehlenden Strahlenbelastung als Methode der Wahl anzusehen. Um die in den Bilddaten gespeicherte Information über die Gehirn(teil-)volumina effizient analysieren zu können, braucht man computergestützte Methoden, die eine Bearbeitung der Bilddaten größtenteils automatisch und mit hoher Geschwindigkeit ermöglichen und grundsätzlich auch geringe Volumenveränderungen zuverlässig erkennen. Dafür bedarf es Algorithmen, die die Bilddaten klassifizieren, d. h. die jedem Bildpunkt genau einen Gewebetyp zuordnen. Wie geeignet ein Verfahren zur Gehirnvolumenberechnung ist, hängt u. a. davon ab, wie gut die im Folgenden beschriebenen Fehlerquellen korrigiert werden. Es muss berücksichtigt werden, dass es unterschiedliche Gewebetypen gibt, die in einer Gewichtung identische Signale haben, d. h. isointens sind. Auch kommt es, wie bei jeder Messung, zu Schwankungen der Messwerte, die um einen Mittelwert normalverteilt sind und außerdem von gerätespezifischen Fehlerquellen überlagert werden. Hierbei spielen insbesondere Magnetfeldinhomogenitäten eine Rolle, die von der Bauweise des Geräts und Art des Magneten abhängen. Auch beeinflussen die Eigenschaften der verwendeten elektronischen Bauelemente die Abbildungsqualität.

Die hier verglichenen Methoden bedienen sich unterschiedlicher Herangehensweisen, deren Erfolg – wie in dieser Arbeit gezeigt werden konnte – u. a. auch davon abhängt, ob das gesamte Gehirnvolumen oder nur die Teilvolumina weiße oder graue Substanz oder Liquor bestimmt werden sollen und welche Bildsequenzen verfügbar sind.

In den folgenden Abschnitten werden die Vor- und Nachteile der hier benutzten Methoden, wie sie sich aus ihren unterschiedlichen Herangehensweisen an die Gehirnvolumetrie ergeben, gegenübergestellt und die Ergebnisse der Volumenberechnungen diskutiert.

4.1 LVQ

Learning Vector Quantization basiert auf neuronalen Netzwerken, weshalb es vor der Durchführung des Klassifizierungsvorganges erforderlich ist, dass der Anwender entscheidet, in wie viele verschiedene Klassen die zu untersuchenden Daten eingeteilt werden sollen. Dann muss das neuronale Netz anhand repräsentativer Daten trainiert werden. Da es bei der Auswahl der Trainingsdaten der subjektiven Beurteilung des Anwenders unterliegt, was er als repräsentativ für die jeweilige Klasse ansieht, kann es zu unterschiedlichen Trainingserfolgen kommen. Allerdings konnte in dieser Arbeit festgestellt werden, dass Trainingsdaten für die Klassen graue Substanz, weiße Substanz und Liquor, die aus verschiedenen Bereichen der jeweiligen Gewebetypen ausgewählt wurden, gleiche Klassifikationsergebnisse lieferten, subjektive Unterschiede bei der Trainingsprozedur sich also nicht auf das Endergebnis der Klassifikation von Gehirndatensätzen der hier verwendeten Qualität auswirken, vorausgesetzt, dass die Trainingsdaten nicht zu große Intensitätsinhomogenitäten aufweisen.

Der Klassifikationsvorgang bei *LVQ* läuft voxelweise ab, ohne dass dabei Informationen über die Position des Volumenelements im Bild in die Entscheidung, zu welcher Klasse es gezählt wird, einbezogen werden. Dies hat den Nachteil, dass es auch bei Verwendung von Daten aller drei Gewichtungen T1, T2 und PD zu unvermeidbaren Fehlklassifikationen im Bereich des Gesichtsschädels kommt, da gewisse Gewebe nicht allein aufgrund ihrer Intensitätswerte, sondern nur unter Einbeziehen von Lageinformationen korrekt klassifiziert werden können. Es gelingt allerdings, diese Fehlklassifikationen nachträglich fast vollständig zu korrigieren, indem man Eigenschaften von grauer und weißer Substanz, wie in Kapitel 2 beschrieben, ausnutzt, die sich aus deren möglichen räumlichen Verteilung ergeben. Größere fehlerklassifizierte Areale liegen im Allgemeinen im Bereich des Gesichtsschädels. Eine vollständige Korrektur ist möglich, wenn diese Gebiete

keine oder eine nur sehr geringe Verbindung zum eigentlichen Gehirngewebe aufweisen, wie es bei den Datensätzen $n3rf20$, $n3rf20+corr$ und $n7rf40+corr$ der Fall ist. Bei schlechter Bildqualität kann es allerdings Verbindungen geben, die so groß sind, dass das Gehirnvolumen fälschlicherweise zu hoch bestimmt wird. Man erkennt diese Problematik an dem Ergebnis für den hier verwendeten Datensatz schlechter Qualität ohne Vorkorrektur ($n7rf40$). Bei den anderen verwendeten Daten waren die Korrekturmechanismen effektiv.

Der Vorteil des *LVQ* liegt in seiner hohen Flexibilität und Anpassungsfähigkeit. Durch die Trainierbarkeit können leicht neue Klassen hinzugefügt werden, was notwendig sein kann, wenn pathologische Strukturen untersucht werden sollen, deren MR-Signaturen von denen der normalen Gehirngewebetypen deutlich abweichen. Die voxelweise Bearbeitung bietet dabei den Vorteil, dass keine Anforderungen an die Struktur der Bilddaten gestellt werden. D. h. man kann z. B. ohne Modifikation des Algorithmus Kindergehirne untersuchen und Raumforderungen, die die anatomische Struktur verändern, „stören“ den Algorithmus nicht. Bei den anderen Methoden, die relative Lageinformationen der Gehirngewebetypen zur Klassifizierung nutzen, muss die Anatomie des zu untersuchenden Gehirns an eine Standard-Anatomie angepasst sein und zu große Abweichungen davon führen zum Versagen des Klassifikationsvorgangs.

4.1.1 Analyse der Ergebnisse der Volumenberechnung

Betrachtet man die Ergebnisse der Volumenberechnung mit *LVQ*, fällt auf, dass die Abweichungen der Werte von weißer und grauer Substanz von den Referenzwerten mit schlechter werdender Bildqualität zunehmen. Beim Datensatz schlechtester Qualität im Vergleich zu den anderen finden sich besonders deutliche Diskrepanzen zu den Referenzwerten. Dies ist, wie oben beschrieben, darauf zurückzuführen, dass die Bildqualität, relativ betrachtet, so schlecht geworden ist und die

Fehlklassifikationen so groß geworden sind, dass die Korrekturalgorithmen versagen.

Man erkennt weiterhin bei den drei relativ besten Datensätzen, dass die Abweichungen von weißer und grauer Substanz jeweils in entgegengesetzte Richtungen zeigen, so dass die Abweichung ihrer Summe geringer ausfällt als die der Teilvolumina. Dies kann man so erklären, dass ein Großteil der Fehler in der Grenzschicht zwischen grauer und weißer Substanz gemacht wird, wo eine richtige Klassifizierung, bedingt durch Partialvolumeneffekte, schwierig ist. Plausibel ist auch, dass die Vorkorrektur von Intensitätsinhomogenitäten einen umso größeren positiven Effekt auf die Bestimmung des Volumens von grauer und weißer Substanz hat, je schlechter die Bildqualität ist.

Vergleicht man die Abweichungen von den Referenzwerten jeweils einer Gewebe-klasse bei verschiedenen Bildqualitäten, so kann man feststellen, dass sie wechselnde Vorzeichen haben. Dies ist besonders relevant, da man deshalb keinen funktionalen Zusammenhang zwischen Abweichung und Bildqualität herstellen kann. Könnte nämlich nicht nur eine Korrelation der Beträge der Abweichungen festgestellt werden, sondern hätten die Abweichungen auch gleiche Vorzeichen, so wäre es u. U. möglich, einen Korrekturfaktor als Funktion der Bildqualität zu ermitteln, der eine verbesserte Volumenabschätzung erlaubt.

Besonders schlecht mit Abweichungen bis über 40% vom Referenzwert sind die Ergebnisse für das Liquorraumvolumen. Dies kann u. a. damit erklärt werden, dass der Korrekturalgorithmus zur Suche der verbundenen Komponenten beim Liquorraum nicht wirksam angewendet werden kann, da die enge Nachbarschaft und die daraus resultierenden Verbindungen von richtig und falsch als Liquor klassifizierten Bildpunkten eine Separierung unmöglich machen. Es fällt auf, dass die Größe der Abweichung nicht mit der Bildqualität korreliert.

Neben der eingeschränkten Verwendbarkeit bei schlechter Bildqualität besteht ein

Nachteil darin, dass *LVQ* nur angewendet werden kann, wenn die Daten in den drei beschriebenen Bildsequenzen vorliegen. Die Flexibilität und Anpassungsfähigkeit und die hohe Genauigkeit bei guten Bildqualitäten, die *LVQ* auszeichnen, fordern also eine relativ große Datenmenge. Dies ist allerdings nicht schwerwiegend, da es sich bei T1, T2 und PD um Standardbildsequenzen handelt, die auch in der klinischen Routine aufgenommen werden.

4.2 SPM

Bei dieser Methode werden neben den Intensitätswerten der einzelnen Bildpunkte auch Lagebeziehungen der Daten in den Klassifizierungsvorgang einbezogen. Dazu wird ein Referenz-Gehirn (engl. template-brain) benutzt, auf das der zu untersuchende Datensatz abgebildet wird. Vorteil dieses Ansatzes ist, dass eine Volumenberechnung bei Verwendung von nur einer Bildsequenz (T1) möglich ist und folglich weniger Daten gebraucht werden, ohne dass sich dies negativ auf die Qualität der Klassifizierung auswirkt. Nachteilig ist, dass bei der Untersuchung von Gehirnen, die eine andere Größe (z. B. Kindergehirne) als der Referenzdatensatz oder besondere anatomische Eigenschaften (z. B. bedingt durch Raumforderungen) aufweisen, Fehler bei der Klassifizierung entstehen. Man muss in solchen Fällen besondere, an die jeweilige Fragestellung angepasste Referenzdatensätze benutzen. Kann ein passender Datensatz bereitgestellt werden, funktioniert *SPM*, ohne dass eine Trainingsprozedur durchlaufen werden muss. Allerdings muss der Anwender die Datensätze vor der Volumenberechnung so modifizieren, dass eine dem Referenz-Gehirn entsprechende Orientierung vorliegt und der Ursprungspunkt im Bereich der vorderen Kommissur liegt.

4.2.1 Analyse der Ergebnisse der Volumenberechnung

Obwohl bei *SPM* eine Bildsequenz ausreichend ist, wurde hier zusätzlich anhand der Kombinationen T1+T2 und T1+T2+PD die Vermutung überprüft, ob sich mit mehreren Gewichtungen die Genauigkeit der Volumenberechnung steigern lässt. Betrachtet man die Summe aus grauer und weißer Substanz, so stellt man jedoch wider Erwarten fest, dass das beste Ergebnis für alle Bildqualitäten bei Verwendung von nur einer Bildsequenz erzielt wird und die Abweichungen bei mehreren Spektren deutlich zunehmen. Da die Simulation der MR-Daten aber keine Bewegungsartefakte beinhaltet, können dafür hier nicht Registrierungsunterschiede zwischen Bildern verschiedener Gewichtung verantwortlich gemacht werden, wie sie bei realen Daten durch minimale Bewegungen des Patienten zwischen zeitlich nacheinander aufgenommenen Bildsequenzen entstehen können.

Auffällig bei dieser Methode ist, dass im Vergleich zu *LVQ* der Betrag der Abweichungen von den Referenzwerten nicht den erwarteten Zusammenhang von steigenden Abweichungen bei sinkender Bildqualität zeigt. Es ergeben sich geringe Abweichungen von den Referenzwerten von ca. 1% nur für die Summen aus grauer und weißer Substanz bei alleiniger Verwendung der T1-Daten. Dieses Ergebnis ist allerdings unabhängig von der Qualität der Datensätze und dem Einsatz einer Vorkorrektur. Alle anderen Werte, wie graue und weiße Substanz oder Liquor alleine, zeigen ein von verwendeten Bildsequenzen und Qualitäten unabhängiges Verhalten, das nur unregelmäßig Ähnlichkeiten bei gleicher Spektrenkombination zeigt und relativ große, stark variierende Abweichungen der Beträge und variierende Vorzeichenwechsel aufweist. Da die Abweichungen außer für die Summe aus grauer und weißer Substanz keinem einheitlichen Muster folgen, relativ groß und offenbar sehr sensitiv gegenüber Veränderungen der Bildqualität und Kombination der Bildsequenzen sind, ist es nicht möglich, einen wie unter 4.1.1 beschriebenen Korrekturfaktor zu ermitteln und es muss festgestellt werden,

dass *SPM* nur eingeschränkt zur Bestimmung der Teilvolumina grauer und weißer Substanz und des Liquorraumvolumens geeignet ist.

4.3 Freesurfer

Das Programm *Freesurfer* basiert auf einer Methode, die wie *SPM* keinen Trainingsprozess erfordert, bei der der Benutzer aber kein Referenzgehirn auszuwählen braucht. Insofern ist *Freesurfer* besonders anwenderfreundlich. Es kann allerdings auch hier notwendig sein, dass der Anwender die Daten manuell modifizieren muss, um eine saubere Abtrennung des Cerebellum zu ermöglichen. Wenn das automatische Bestimmen der Trennlinie im Hirnstamm versagt, muss eine passende Stelle manuell gefunden und markiert werden, was allerdings i. d. R. erst nach einigen Versuchen gelingt.

4.3.1 Analyse der Ergebnisse der Volumenberechnung

Da bei *Freesurfer* in den Daten vorhandene Intensitätsinhomogenitäten automatisch korrigiert werden, unterscheidet sich die Auswahl der Bildqualitäten von der der anderen Methoden. Aufgrund des besonderen Ansatzes über die Berechnung der Oberflächen von grauer und weißer Substanz, besteht bei *Freesurfer* die Möglichkeit zur Korrektur von Fehlern in der Oberflächentopologie, so dass dies hier als weiteres Unterscheidungskriterium herangezogen wurde.

Wie bei *SPM* fällt die Tatsache auf, dass Bildqualität und Größe der Abweichungen unkorreliert sind und dass die Werte der Teilvolumina von grauer und weißer Substanz jeweils relativ hohe Abweichungen von den Referenzwerten zeigen, wobei allerdings alle Abweichungen der weißen positiv und die der grauen Substanz negativ sind, so dass sie sich zu einer relativ geringen Abweichung ihrer Summe von -0,7 bis 2% ausgleichen.

Auch wenn die Höhe der Abweichungen von grauer und weißer Substanz von den Referenzwerten eher gegen einen Einsatz dieser Methode zur Bestimmung der Teilvolumina sprechen, besteht aufgrund der relativ geringen Schwankungsbreite der Abweichungen von unter 2% und der jeweils gleichen Vorzeichen bei verschiedenen Bildqualitäten eher als bei *SPM* und *LVQ* die Möglichkeit, Korrekturen zur Abschätzung der wirklichen Volumina zu finden.

4.4 Datenmaterial

Bei den verwendeten Daten handelt es sich um simulierte Daten, also nicht um Daten aus einer realen Magnetresonanztomographie-Aufnahme. Der Grund dafür liegt darin, dass es sehr schwierig bzw. praktisch nicht möglich ist, aus (echten) MRT-Bildern die Volumina der einzelnen Gewebe objektiv zu bestimmen, da man entweder auf eine andere, potentiell fehlerbehaftete Bildanalyse-Methode zurückgreifen oder die Bestimmung der Grenzen zwischen den Gewebetypen manuell durchführen muss, was auch bei mehrfacher Ausführung durch unterschiedliche Anwender nur eine „subjektive“ Näherung darstellt. Die simulierten Daten können jedoch als sehr realistisch und für die in dieser Arbeit durchgeführten Untersuchungen als echten klinischen Daten gleichwertig angesehen werden, denn ihnen liegt ein anatomischer Atlas zugrunde, der aus einer großen Anzahl realer Daten gemittelt wurde. Daraus erzeugt ein Algorithmus Bilddaten, der die Abbildung eines Gehirns durch ein MRT simuliert. Deshalb sind die Volumina der betrachteten Gewebe exakt bekannt und gleichzeitig werden realistische Daten verwendet. Darüber hinaus entsprechen die Werte der simulierten Parameter „Rauschen“ und „Intensitätsinhomogenität“ den in realen Daten vorhandenen Werten (Collins et al., 1998; Kwan et al., 1999).

Allerdings repräsentieren die hier beschriebenen Daten nur einen durchschnittlichen Gehirndatensatz, in dem einerseits nicht alle real möglichen Parameter simuliert werden und der andererseits nicht als Beispiel für Gehirne mit besonderen anatomischen Merkmalen herangezogen werden kann. So ist anzunehmen, dass Gehirne mit besonderen, stark ausgeprägten pathologischen Veränderungen oder Gehirne von Kindern und alten Menschen einer gesonderten Betrachtung bedürfen. Außerdem variieren auch bei übereinstimmender Anatomie in Wirklichkeit mehr Parameter als die hier simulierten. Nicht erfasst werden z. B. Bewegungsartefakte oder technische Eigenschaften der zur Aufnahme verwendeten Geräte, die sich teilweise deutlich auf die Bildqualität auswirken. Interessant wäre es auch zu untersuchen, inwieweit unterschiedliche Voxel-Größen die Ergebnisse beeinflussen, denn die in dieser Arbeit verwendeten Daten haben eine Auflösung von 1mm^3 , während z. B. in der klinischen Radiologie auch mit geringeren Auflösungen gearbeitet wird oder mit Hochfeld-Magnetresonanztomographen noch höhere Auflösungen erreicht werden. Hier besteht noch weiterer Forschungsbedarf, insbesondere um zu versuchen, den Einfluss verschiedener Faktoren auf die Volumetrieergebnisse zu quantifizieren.

Dieses „Defizit“ der hier verwendeten Daten stellt jedoch die Ergebnisse und die daraus im Folgenden gezogenen Schlüsse nicht in Frage, sondern legt sogar eher nahe, dass die beschriebenen Effekte bei realen Aufnahmen noch deutlicher ausfallen, weil die simulierten Daten „besser“ sind.

In diesem Zusammenhang stellt sich auch die grundsätzliche Frage, wie man die Qualität von Bilddaten definiert bzw. misst und inwieweit beurteilt werden kann, ob MRT-Bilder einen gewissen qualitativen Anspruch erfüllen, um für Analyseverfahren wie die Volumetrie geeignet zu sein. Denn nur wenn es standardisierte Qualitätsrichtlinien gibt, sind vergleichende Untersuchungen sinnvoll; dies gilt erst recht, wenn Daten verschiedener Herkunft verwendet werden sollen (Tofts,

1998). Ein Ansatz kann darin bestehen, bei realen Bilddaten die Parameter „Rauschen“ und „Intensitätsinhomogenitäten“ zu schätzen und so nur Daten relativ guter bzw. ähnlicher Qualität für volumetrische Analysen zu verwenden (Sijbers et al., 1996, 1998; Styner et al., 2000).

4.5 Schlussbetrachtungen

Stellt man die Frage, welche der Methoden die „beste“ ist, d. h. welche Methode die genaueste Volumenbestimmung erlaubt und größte Robustheit gegenüber Bildqualitätsunterschieden besitzt, wird aus dem oben Gesagten deutlich, dass keine einfache Aussage getroffen werden kann, sondern jede Methode Stärken und Schwächen in bestimmten Bereichen aufweist.

Mit Hilfe der unter 3.5.4 eingeführten „mittleren Abweichung“ $\bar{\delta}$ kann eine etwas allgemeinere Aussage über eine „globale“ Güte der einzelnen Methoden gemacht werden, allerdings muss dabei beachtet werden, dass dieser Parameter unter den Ergebnissen für weiße und graue Substanz und der Summe aus beiden mittelt und eine so als relativ schlechter eingestufte Methode besser sein kann, wenn es z. B. nur um ein Teilvolumen geht. Es zeigt sich, dass *LVQ* nach $\bar{\delta}$ bei guten Bildqualitäten eine bessere Leistungsfähigkeit hat als die etablierten Methoden *SPM* und *Freesurfer* und bei der Bestimmung des Volumens der weißen Substanz sogar deutlich bessere Resultate liefert (ca. um Faktor 4 genauer), während *LVQ* bei den Daten der schlechtesten Bildqualität unterlegen ist (Abweichungen bis 40%). Auch zeigt $\bar{\delta}$ eine leichte Überlegenheit von *Freesurfer* gegenüber *SPM* bei allen Datensätzen.

4.5.1 Bedeutung der Ergebnisse für die Gehirnvolumetrie

Insgesamt kann man feststellen, dass bei Abweichungen der Volumetrie-Ergebnisse von den Referenzwerten bis über 10% bei *SPM* und deutlich über 5% bei *Freesurfer* auch bei guten Bildqualitäten die Genauigkeit als relativ niedrig angesehen werden muss und daher Vergleiche von Volumina, die mit unterschiedlichen Methoden bestimmt wurden, nicht sinnvoll sind. Vergleicht man z. B. Ergebnisse für die weiße Substanz von *Freesurfer* und *SPM*, so kann es zu Diskrepanzen von knapp 20% kommen.

Die Ergebnisse zeigen aber auch, dass Untersuchungen zum Vergleich von Gehirnvolumina, die mit *derselben* Methode berechnet worden sind, vorsichtig betrachtet werden müssen, da offenbar relativ geringe Qualitätsunterschiede der Bilddaten einen großen Einfluss auf die Abweichungen der Volumetrieergebnisse haben können. Besonders eindrucksvoll äußert sich dies bei den Ergebnissen der Liquorraumbestimmung mit *SPM*, bei denen Volumenbestimmungen zwischen zwei Bildqualitäten um bis zu 20% abweichen, oder bei den Werten für die graue Substanz (bis zu 8% Diskrepanz). Als vergleichsweise robust kann *Freesurfer* angesehen werden, da hier die geringsten Unterschiede der Ergebnisse untereinander auftreten.

Die Bedeutung dieser Resultate liegt darin, dass eine mögliche klinische Anwendung der Gehirnvolumetrie z. B. als Unterstützung in der Diagnostik und Verlaufskontrolle von Krankheiten aufgrund der sensitiven Abhängigkeit der Volumetrieergebnisse von Bildqualitätsparametern problematisch ist. Auch können sich Auswirkungen auf Studien ergeben, die solche Methoden zur Volumenberechnung nutzen, um Zusammenhänge von Krankheiten und Volumenänderungen nachzuweisen (Job et al., 2002), weil es bei MRT-Untersuchungen leicht zu Qualitätsunterschieden kommen kann, die bei der Datenauswertung die eben besprochenen

Effekte hervorrufen können. Da oft nur kleine Patientenkollektive untersucht werden und die zu erwartenden Änderungen der Gehirnvolumina i. d. R. gering sind, muss in Betracht gezogen werden, dass diese Effekte für die Volumenberechnung nicht vernachlässigbar sein könnten und sie im Rahmen einer Fehlerrechnung in die Signifikanzanalyse einbezogen werden müssten.

Zwar wurde in dieser Arbeit der Einfluss anderer Faktoren wie Bewegungsartefakte oder gerätespezifische Unterschiede nicht untersucht, es ist aber zu vermuten, dass dadurch die Schwierigkeiten der Vergleichbarkeit eher noch erhöht werden. In jedem Fall sollte bei der Gehirnvolumetrie besonderer Wert auf eine übereinstimmende Bildqualität der zu vergleichenden Daten gelegt werden.

4.5.2 Ausblick

Da anzunehmen ist, dass in Zukunft die Bedeutung der Gehirnvolumetrie und der exakten Bestimmung der Größe von pathologischen Veränderungen steigen wird, sollten weitere Untersuchungen über Genauigkeit und Robustheit der verwendeten Methoden auch unter Einbeziehung anderer Parameter durchgeführt und Qualitätsstandards für Bilddaten entwickelt werden, um u. U. quantitative Aussagen über die Abhängigkeit der Fehler der Volumetrie von Bildqualität und verwendeter Methode machen und zur Korrektur nutzen zu können.

Eine Möglichkeit, die Evaluierung von volumetrischen Methoden auf eine breitere Grundlage zu stellen, liegt in der Verwendung von realen Daten. Insbesondere könnte hier das *Visible Human Project*¹ interessant sein, bei dem MR-Daten mit digitalisierten anatomischen Gewebsschnitten des gesamten Gehirns eines Individuums verglichen werden können (Banvard, 2002). Auch sollten weitere Unter-

¹Über das *Visible Human Project* der National Library of Medicine (National Institutes of Health, USA) finden sich nähere Informationen im Internet unter http://www.nlm.nih.gov/research/visible/visible_human.html.

suchungen an größeren Kollektiven vorgenommen werden, in denen die Bilddaten eines Individuums zu unterschiedlichen Zeitpunkten mit einem indentischen oder anderen Gerät aufgenommen werden, da variierende Aufnahmebedingungen Einfluss auf die Bildqualität haben können. Dabei hat man zwar i. d. R. keine objektiven Informationen über das Volumen wie bei den hier verwendeten Daten, aber relative Abweichungen der Volumetrieergebnisse geben Aufschluss über die Robustheit der Verfahren.

Zusammenfassung

Veränderungen des gesamten Gehirnvolumens und der Teilvolumina von grauer und weißer Substanz und des Liquorraumes treten nicht nur während der Entwicklung vom kindlichen zum erwachsenen Gehirn und beim Altern auf, sondern können auch mit pathologischen Prozessen verschiedener Ursachen (z. B. Alkoholismus, Schizophrenie, M. Alzheimer, Hirnödeme) einhergehen.

Die *in vivo* Gehirnvolumetrie kann daher dazu benutzt werden, Zusammenhänge zwischen Krankheiten und morphologischen Veränderungen zu untersuchen, Hilfestellung in der Diagnostik zu geben oder objektive Verlaufskontrollparameter in Therapiestudien bereitzustellen.

In dieser Arbeit werden drei verschiedene Methoden zur Gehirnvolumenberechnung miteinander verglichen und anhand von Referenzdatensätzen unterschiedlicher Bildqualität (Variation der Parameter „Rauschen“ und „Intensitätsinhomogenität“) auf ihre Genauigkeit und Robustheit untersucht. Die Methode *Learning Vector Quantization (LVQ)* wird hier für die Analyse multispektraler Magnetresonanztomographie-Daten implementiert, während *Statistical Parametric Mapping (SPM)* und *Freesurfer* in der bildgebenden Hirnforschung weit verbreitete Methoden sind.

Es konnte gezeigt werden, dass *LVQ* qualitativ vergleichbar mit den etablierten Methoden arbeitet, wenn die Bilddaten von guter Qualität sind. Insbesondere aber *Freesurfer* zeigt eine bessere Robustheit gegenüber Daten schlechterer Qualität. Vergleicht man die Volumetrieergebnisse der drei Methoden mit den Referenzdaten, fallen von 2 bis über 10% schwankende Abweichungen der Volumina für graue und weiße Substanz von den Referenzwerten in Abhängigkeit von Methode, Bildqualität und Kombination der Bildsequenzen auf. Noch deutlich größere Diskrepanzen bis über 20% ergeben sich im direkten Vergleich von Volumina, die

mit unterschiedlichen Methoden anhand identischer Daten berechnet wurden, da die Abweichungen teilweise entgegengesetzte Vorzeichen haben. Die Ergebnisse für Liquor weisen die größten Abweichungen von ca. -40% bis $+20\%$ von den Referenzwerten auf.

Darüber hinaus wurde festgestellt, dass die *gleiche* Methode bei Unterschieden in der Bildqualität variierende Volumina berechnet, ohne dass dabei die Änderungen von Bildqualität und Volumenberechnung korreliert sind.

Diese Ergebnisse machen deutlich, dass bei der Gehirnvolumetrie sowohl Art der verwendeten Methode als auch Bildqualität einen nicht zu vernachlässigenden Einfluss auf die Ergebnisse haben können und dass dies bei der Auswertung und Interpretation von Volumetrieergebnissen – insbesondere angesichts der oft geringen nachzuweisenden Effekte – berücksichtigt werden sollte.

Abkürzungen

CSF	Liquor cerebrospinalis (engl. cerebrospinal fluid)
CT	Computertomographie
GM	Graue Substanz (engl. gray matter)
lh	Linke Hemisphäre
LVQ	Learning Vector Quantization
MRT	Magnetresonanztomographie
PD	Protonendichte
rh	Rechte Hemisphäre
SPM	Statistical Parametric Mapping
T1	Kernspin-Relaxationszeit T1
T2	Kernspin-Relaxationszeit T2
WM	Weißer Substanz (engl. white matter)

Anhang - Programm-Listings

Programm-Listing Nr. 1: lvqtraining

```
% Matlab Skript lvqtraining.m
% ERSTER Schritt des Trainingsprozesses:
% Auswahl der f"ur das Training benutzten Bildbereiche
% durch den Anwender

disp('Programm zur Erstellung der Trainings- und Testdatei')
    z=0; t=0;
Trainingsende=0;
while not(Trainingsende)
    loadstr=input('Geben Sie Pfad und Dateinamen des f"ur den
        ... Testvorgang benutzten Bildes ein: ','s');
    load(loadstr);
    img=im;
    Bildende=0;
    while not(Bildende)
        kstr=input('F"ur welche Klasse soll trainiert werden? (0
            ... beendet das Training f"ur dieses Bild) ','s');
        k=str2num(kstr);
        if or(isempty(k),(k==0))
            Bildende=1;
            break;
        end
    end
    figure(1);
    imagesc(img(:,:,1));
    colormap(gray);
    disp('Waehlen Sie mit dem Fadenkreuz die linke obere Ecke
        ... des Trainingsbereiches aus!');
    [x1,y1]=ginput(1);
    disp('W"ahlen Sie mit dem Fadenkreuz die rechte untere Ecke
```

```

        ... des Trainingsbereiches aus!');
    [x2,y2]=ginput(1);
% der Klassenlabel wird an den entsprechenden Stellen im
% Bild eingetragen und die Daten werden der Trainings-
% und Testmatrix im LVQ-Format hinzugef"ugt
    for x=round(x1):round(x2)
        for y=round(y1):round(y2)
            img(y,x,4)=k;
            z=z+1;
            if not(mod(z,10))
t=t+1;
                testascii(t,:)=img(y,x,:);
            else
                trainascii(z,:)=img(y,x,:);
            end
        end
    end
end %while Bildende
disp('Training f"ur dieses Bild abgeschlossen.')
ende=input('Moechten Sie das Training forsetzen?(j/n)', 's');
if ende=='n'
    Trainingsende=1;
end
end %while Trainingsende
dim=3;
save testlvqascii.dat dim testascii -ascii
save trainlvqascii.dat dim trainascii -ascii

```

Programm-Listing Nr. 2: lvqprepare

```

% Matlab-Skript lvqprepare.m: Datenvorbereitung
% Konvertierung Matlab-Format --> LVQ-ASCII-Format
dim=3; % drei Bildimensionen [x,y,z]=[181,217,181]
for z=1:181
    iy=0;
    z
    imasc=zeros(217*181,3);
    for y=1:217
        for x=1:181
            iy=iy+1;
            imasc(iy,1)=img1(x,y,z); % T1
            imasc(iy,2)=img2(x,y,z); % T2

```

```

        imasc(iy,3)=img3(x,y,z); % PD
    end
end
savestr=strcat('/li2/dat/n7rf40/imasc',num2str(z),'.dat');
save(savestr,'dim','imasc','-ascii');
end

```

Programm-Listing Nr. 3: lvqperform

```

#!/bin/csh
# Linux-C-Shell-Skript lvqperform.csh
# f"uhrt die LVQ-Klassifikation f"ur alle Scheiben des Gehirns
# durch und speichert die Ergebnisse
# classify ist das Klassifizierungsprogramm des lvq_pak
# -din folgt die Input-Datei
# -cin folgt den zu benutzenden Code-Vektor
# -dout folgt die Output-Datei

set x=10
foreach i (*)
    <PFAD>/lvq_pak-3.1/classify -din imasc$x.dat -cin
... <PFAD>/lvq_pak-
    3.1/CodeVektorDatei.lvq -dout
        ... ./lvqErgebnis/classified$x.cla
    @ x++
end

```

Programm-Listing Nr. 4: lvqreadclassified

```

% lvqreadclassified.m
imgfullplus=uint8(zeros(256,256,122));
for z=10:131
    loadcla2=strcat('<PFAD>classified',num2str(z),'.cla');
    fid=fopen(loadcla,'r');
    status=fseek(fid,1,'bof');
% die Bilddatei wird ge"offnet, die erste Zeile wird nicht
% gelesen; nach Rotation um 90° und up-down-flip wird ...
    bild=flipud(rot90(fscanf(fid,'%g %g %g %g',[4 inf])));
    iy=0;
% ... aus den einzelnen Scheiben eine Datei der Form

```

```
% 256x256x122 erstellt, die das gesamte klassifizierte
% Gehirn enth"alt
for y=1:256
    for x=1:256
        iy=iy+1;
% nur die Werte der 4. Spalte der aus dem Klassifizierungsteil
% hervorgegangenen Dateien werden weiterbenutzt, denn sie
% sind die den Bildpunkten zugeordneten Klassen
        imgfullplus(y,x,z-9)=uint8(bild(iy,4));
    end
end
    status=fclose(fid);
    clear fid;
end % for z
clear bild
```

Programm-Listing Nr. 5: readraw

```
% readraw.m
% Skript liest .raw Daten und erstellt eine 181x217x181-Bildmatrix
fid = fopen('/li2/refbrain/brainwebnormal.raw');
f = uint8(fread(fid));
img=(zeros(181,217,181));
x=1;
y=1;
z=1;
for i=1:7109137
    if ((x==182) & (y~=217))
        x=1;
        y=y+1;
    end
    if ((x==182) & (y==217))
        x=1;
        y=1;
        z=z+1;
    end
    img(x,y,z)=(f(i));
    x=x+1;
end
```

Programm-Listing Nr. 6: afni2spm

```
% afni2spm.m
% Skript speichert 181x217x181-Bildmatrix im SPM-Format
DIM(1)=181; % setzt die Matrixgr"oÙe
DIM(2)=217;
DIM(3)=181;
VOX(1)=1; % setzt die Voxel-Gr"oÙe
VOX(2)=1; % auf 1mm^3
VOX(3)=1;
SCALE=1; % Skalierungsfaktor
TYPE=4;
OFFSET=0;
ORIGIN = [1; 1; 1]; % setzt die Ursprungskordinaten
prefix = '<PFAD><Dateiname>';
DESCRIP='spm compatible';
spm_hwrite(prefix,DIM,VOX,SCALE,TYPE,OFFSET,ORIGIN,DESCRIP);
fid=fopen('<PFAD><Dateiname>.img','w');
fwrite(fid,img,'int16');
fclose(fid);
```

Programm-Listing Nr. 7: corscan

```
% Matlab-Skript corscan.m
% ldt Freesurfer COR-xxx Daten nach Matlab
corimg=uint8(zeros(256,256,256));
for z=1:256
    y=1;
    x=1;
    if (z<10)
        fl=['<PFAD>/mri/orig/COR-00',int2str(z)];
        fid = fopen(fl,'r');
    end
    if ((z>9) & (z<100))
        fl=['<PFAD>/mri/orig/COR-0',int2str(z)];
        fid = fopen(fl,'r');
    end
    if (z>99)
        fl=['<PFAD>/mri/orig/COR-',int2str(z)];
        fid = fopen(fl,'r');
    end
    f=fread(fid);
```

```
    for i=1:65536
        corimg(y,x,z)=f(i);
        x=x+1;
        if (x==257)
            x=1;
            y=y+1;
        end
    end % for i
    fclose(fid);
end % for z
```

Programm-Listing Nr. 8: volcount

```
% Matlab-Skript volcount.m
% Berechnet das Volumen der weißen Substanz für linke
% und rechte Hemisphäre

wmrh=0; % Volumen rechte Hemisphäre
wmlh=0; % " linke "
for z=1:256
    for y=1:256
        for x=1:256
            if (corimg(y,z,x)==80) %re.H.
                wmrh=wmrh+1;
            end
            if (corimg(y,z,x)==255) %li.H.
                wmlh=wmlh+1;
            end
        end
    end
end
wmrh
wmlh
wmges=wmrh+wmlh
```

Programm-Listing Nr. 9: calcvol

```
%Matlab-Skript calcvol.m
%zur Berechnung des Volumens der Grauen Substanz
%benutzt fs_read_curv.m aus der EEG_Toolbox
```

```
[rhthick, face]=fs_read_curv('<PFAD>/surf/rh.thickness');
[lhthick, face]=fs_read_curv('<PFAD>/surf/lh.thickness');
[rharea, face]=fs_read_curv('<PFAD>/surf/rh.area');
[lharea, face]=fs_read_curv('<PFAD>/surf/lh.area');

volgmrh=sum(rhthick .* rharea)
volgmlh=sum(lhthick .* lharea)
volgmgges=sum(rhthick .* rharea)+sum(lhthick .* lharea)
```

Programm-Listing Nr. 10: embedrotflip

```
%Matlab-Skript embedrotflip.m
%bettet das primäre 181x217x181-Bild in einen
%256x256x256-Array, rotiert und spiegelt, um die
%Daten auf den Trainingsvorgang vorzubereiten.

img2PD=uint8(zeros(256,256,256));
img2T1=uint8(zeros(256,256,256));
img2T2=uint8(zeros(256,256,256));
img2=uint8(zeros(256,256,256));

for i=1:3
    if i==1
        img=imgT1;
    end
    if i==2
        img=imgT2;
    end
    if i==3
        img=imgPD;
    end

    for x=1:181
        for y=1:217
            for z=1:181
                img2(z,y,x)=img(x,y,z);
            end
        end
    end

    if i==1
        img2T1=img2;
```

```
end
if i==2
    img2T2=img2;
end
if i==3
    img2PD=img2;
end
end

for z=1:256
    img=img2PD(:,:,z);
    img2PD(:,:,z)=flipud(img);
end
for z=1:256
    img=img2T1(:,:,z);
    img2T1(:,:,z)=flipud(img);
end
for z=1:256
    img=img2T2(:,:,z);
    img2T2(:,:,z)=flipud(img);
end
```

Programm-Listing Nr. 11: perfit

```
% Matlab-Skript perfit.m
% fasst die Korrekturalgorithmen zusammen
lvqcorr2
transform2bw3d
a3=zeros(256,256);
a3dfull(:,:,1)=a3;
a3dfull(:,:,2)=a3;
a3dfull(:,:,3)=a3;
a3dfull(:,:,4)=a3;
a3dfull(:,:,5)=a3;
deletepix
connectedcomponent
```

Programm-Listing Nr. 12: lvqcorr2

```
% Matlab-Skript lvqcorr2.m
% führt die Korrektur von fehlklassifizierter weisser Substanz
```

```
% durch, basierend auf der Annahme, dass weisse Substanz nicht
% in Nachbarschaft von Knochen liegt.
%
% li->re
for z=2:255
  for y=2:255
    for x=2:255
      if imgfullplus(x,y,z)==4
        if (imgfullplus(x,y-1,z)==5)
          imgfullplus(x,y-1,z)=4;
        end
        if (imgfullplus(x,y+1,z)==5)
          imgfullplus(x,y+1,z)=4;
        end
        if (imgfullplus(x+1,y,z)==5)
          imgfullplus(x+1,y,z)=4;
        end
      end
    end
  end
end

% re->li
for z=255:-1:2
  for y=255:-1:2
    for x=255:-1:2
      if imgfullplus(x,y,z)==4
        if (imgfullplus(x,y-1,z)==5)
          imgfullplus(x,y-1,z)=4;
        end
        if (imgfullplus(x,y+1,z)==5)
          imgfullplus(x,y+1,z)=4;
        end
        if (imgfullplus(x-1,y,z)==5)
          imgfullplus(x-1,y,z)=4;
        end
      end
    end
  end
end

% top->bottom
for z=2:255
```

```

for x=2:255
  for y=2:255
    if imgfullplus(x,y,z)==4
      if (imgfullplus(x-1,y,z)==5)
        imgfullplus(x-1,y,z)=4;
      end
      if (imgfullplus(x+1,y,z)==5)
        imgfullplus(x+1,y,z)=4;
      end
      if (imgfullplus(x,y+1,z)==5)
        imgfullplus(x,y+1,z)=4;
      end
    end
  end
end
end

% bottom->top
for z=255:-1:2
  for x=255:-1:2
    for y=255:-1:2
      if imgfullplus(x,y,z)==4
        if (imgfullplus(x-1,y,z)==5)
          imgfullplus(x-1,y,z)=4;
        end
        if (imgfullplus(x+1,y,z)==5)
          imgfullplus(x+1,y,z)=4;
        end
        if (imgfullplus(x,y-1,z)==5)
          imgfullplus(x,y-1,z)=4;
        end
      end
    end
  end
end
end
end

```

Programm-Listing Nr. 13: deletepix

```

% Skript deletepix.m
% Korrigiert Voxel weisser/grauer Substanz, die nicht
% die Bedingung der Konnektivität erfüllen
a3dfullalt=a3dfull;
for z=2:255

```

```

for y=2:255
  for x=2:255
    if (a3dfull(x,y,z))>0
      if ((a3dfull(x-1,y,z)==0) & (a3dfull(x+1,y,z)==0))
        a3dfull(x,y,z)=0;
      end
      if ((a3dfull(x,y+1,z)==0) & (a3dfull(x,y-1,z)==0))
        a3dfull(x,y,z)=0;
      end
    end
  end
end
end
end
end

```

Programm-Listing Nr. 14: cvol

```

% Matlab-Skript cvol.m
% Zur Berechnung des Volumen nach
% Anwendung des ConnectedComponent-Algorithmus

% 1)Bestimmung der größten verbundenen Komponente in 3D
connectivity=6 % 6 nächste Nachbarn
[L,NUM]= bwlabeln(a3dfull,connectivity)
component_vol=zeros(NUM,1);
min=0;
% Berechnet die Volumina der einzelnen Komponenten
for i=1:NUM
  for z=1:256
    for y=1:256
      for x=1:256
        if (L(x,y,z)==i)
          component_vol(i,1)=component_vol(i,1)+1;
        end
      end
    end
  end
end
end
end
end
[d,biggest_comp]=max(component_vol);
% biggest_comp enthält die Index-Nummer der größten Komponente
wm=0; gm=0; %Zählvariablen für weisse/grau Substanz
for z=1:256
  for y=1:256
    for x=1:256

```

```
        if (L(y,x,z)==biggest_comp)
            if (imgfullplus(y,x,z)==2)
                gm=gm+1;
            end
            if (imgfullplus(y,x,z)==5)
                wm=wm+1;
            end
        end
    end
end
end
end
```

Programm-Listing Nr. 15: spmvol

```
% Matlab-Skript spmvol.m
% zur Berechnung der Volumina der mit SPM segmentierten Daten
readraw; % einlesen der segmentierten Daten
img1=img; % graue Substanz
readraw;
img2=img; % weisse Substanz
readraw;
img3=img; % Liquor
gm=0; wm=0; csf=0;
% Berechnung der einzelnen Volumina:
for z=1:181
    for y=1:217
        for x=1:181
            if (img1(x,y,z)>0)
                gm=gm+double(img1(x,y,z))/255;
            end
            if (img2(x,y,z)>0)
                wm=wm+double(img2(x,y,z))/255;
            end
            if (img3(x,y,z)>0)
                csf=csf+double(img3(x,y,z))/255;
            end
        end
    end
end
end
```

Literaturverzeichnis

- [1] J. Ashburner und K. J. Friston. Multimodal image coregistration and partitioning—a unified framework. *Neuroimage*, 6(3):209–217, Oct 1997.
- [2] R. A. Banvard. The Visible Human Project Image Data Set From Inception to Completion and Beyond. *Proceedings CODATA 2002: Frontiers of Scientific and Technical Data, Track I-D-2: Medical and Health Data, Montréal, Canada*, 2002.
- [3] R. A. Bermel, J. Sharma, C. W. Tjoa, S. R. Puli, R. Bakshi. A semiautomated measure of whole-brain atrophy in multiple sclerosis. *J Neurol Sci*, 208(1-2):57–65, Apr 2003.
- [4] J. C. Bezdek, L. O. Hall, L. P. Clarke. Review of MR image segmentation techniques using pattern recognition. *Med Phys*, 20(4):1033–1048, Jul 1993.
- [5] E. D. Bigler. Premorbid brain volume and dementia. *Arch Neurol*, 58(5):831–833, May 2001. Letter.
- [6] E. D. Bigler, B. Kerr, J. Victoroff, D. F. Tate, J. C. S. Breitner. White matter lesions, quantitative magnetic resonance imaging, and dementia. *Alzheimer Dis Assoc Disord*, 16(3):161–170, Jul 2002.

- [7] E. D. Bigler, C. M. Lowry, C. V. Anderson, S. C. Johnson, J. Terry, M. Steed. Dementia, quantitative neuroimaging, and apolipoprotein E genotype. *AJNR Am J Neuroradiol*, 21(10):1857–1868, Nov 2000.
- [8] E. D. Bigler und D. F. Tate. Brain volume, intracranial volume, and dementia. *Invest Radiol*, 36(9):539–546, Sep 2001.
- [9] R. I. Block, D. S. O’Leary, J. C. Ehrhardt, J. C. Augustinack, M. M. Ghoneim, S. Arndt, J. A. Hall. Effects of frequent marijuana use on brain tissue volume and composition. *Neuroreport*, 11(3):491–496, Feb 2000.
- [10] B. M. Bogdanoff, C. R. Stafford, L. Green, C. F. Gonzalez. Computerized transaxial tomography in the evaluation of patients with focal epilepsy. *Neurology*, 25(11):1013–1017, Nov 1975.
- [11] G. F. Busatto, G. E. Garrido, O. P. Almeida, C. C. Castro, C. H. Camargo, C. G. Cid, C. A. Buchpiguel, S. Furuie, C. M. Bottino. A voxel-based morphometry study of temporal lobe gray matter reductions in Alzheimer’s disease. *Neurobiol Aging*, 24(2):221–231, Mar 2003.
- [12] M. J. Chiu, C. C. Lin, K. H. Chuang, J. H. Chen, K. M. Huang. Tissue segmentation-assisted analysis of fMRI for human motor response: an approach combining artificial neural network and fuzzy C means. *J Digit Imaging*, 14(1):38–47, Mar 2001.
- [13] C. E. Coffey, J. F. Lucke, J. A. Saxton, G. Ratcliff, L. J. Uritas, B. Billig, R. N. Bryan. Sex differences in brain aging: a quantitative magnetic resonance imaging study. *Arch Neurol*, 55(2):169–179, Feb 1998.
- [14] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, A. C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Trans Med Imaging*, 17(3):463–468, Jun 1998.

- [15] R. W. Cox und J. S. Hyde. Software tools for analysis and visualization of fMRI data. *NMR Biomed*, 10(4-5):171–178, Jun 1997.
- [16] A. M. Dale, B. Fischl, M. I. Sereno. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*, 9(2):179–194, Feb 1999.
- [17] R. Damadian, M. Goldsmith, L. Minkoff. NMR in cancer: FONAR image of the live human body. *Physiol Chem Phys*, 9(1):97–100, 108, 1977.
- [18] S. Durston, H. E. Hulshoff, B. J. Casey, J. N. Giedd, J. K. Buitelaar, H. van Engeland. Anatomical MRI of the developing human brain: what have we learned. *J Am Acad Child Adolesc Psychiatry*, 40(9):1012–1020, Sep 2001.
- [19] B. Fischl, M. I. Sereno, A. M. Dale. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage*, 9(2):195–207, Feb 1999.
- [20] J. H. Fox, J. L. Topel, M. S. Huckman. Use of computerized tomography in senile dementia. *J Neurol Neurosurg Psychiatry*, 38(10):948–953, Oct 1975.
- [21] K. J. Friston. Spatial Normalisation: A New Approach. *Proceedings of BrainMap '95 UTHSC San-Antonio, Texas*, 1995.
- [22] C. Gasperini, A. Paolillo, E. Giugni, S. Galgani, F. Bagnato, C. Mainero, E. Onesti, S. Bastianello, C. Pozzilli. MRI brain volume changes in relapsing-remitting multiple sclerosis patients treated with interferon beta-1a. *Mult Scler*, 8(2):119–123, Apr 2002.
- [23] J. N. Giedd, J. W. Snell, N. Lange, J. C. Rajapakse, B. J. Casey, P. L. Kozuch, A. C. Vaituzis, Y. C. Vauss, S. D. Hamburger, D. Kaysen, J. L. Rapoport. Quantitative magnetic resonance imaging of human brain development: ages 4-18. *Cereb Cortex*, 6(4):551–560, Jul 1996. Clinical Trial.

- [24] N. R. Graff-Radford, R. K. Heaton, M. P. Earnest, J. C. Rudikoff. Brain Atrophy and Neuropsychological Impairment in Young Alcoholics. *Stud Alcohol*, 43:859–868, 1982.
- [25] Evidence-Based Radiology Working Group. Evidence-based radiology: a new approach to the practice of radiology. *Radiology*, 220(3):566–575, Sep 2001.
- [26] P. Götze, D. Kühne, J. Hansen, H. P. Knipp. Hirnatrophische Veränderungen bei chronischem Alkoholismus. *Archiv für Psychiatrie u. Nervenkrankheiten*, 226:137–156, 1978.
- [27] H. Haken und Wolf H. C. Atom- und Quantenphysik. *Springer-Verlag*, Kapitel 20:351–375, 1993.
- [28] J. Haslett. Maximum likelihood discriminant analysis on the plane using a Markovian model of spatial context. *Pattern Recognition*, 18:287–296, 1985.
- [29] J. Hatazawa, M. Ito, H. Yamaura, T. Matsuzawa. Sex difference in brain atrophy during aging; a quantitative study with computed tomography. *J Am Geriatr Soc*, 30(4):235–239, Apr 1982.
- [30] H. Herrmann und J. Ullrich. Menschwerdung. *Akademie-Verlag*, pages 712–720, Dez 1991.
- [31] J. P. Hornak. The basics of MRI. <http://www.cis.rit.edu/htbooks/mri>, Kapitel 3, 2002.
- [32] G. N. Hounsfield. Computerized transverse axial scanning (tomography). 1. Description of system. *Br J Radiol*, 46(552):1016–1022, Dec 1973.

- [33] G. N. Hounsfield. Computed medical imaging. *Science*, 210(4465):22–28, Oct 1980. Historical Article.
- [34] M. A. Howard, N. Roberts, M. Garcia-Finana, P. E. Cowell. Volume estimation of prefrontal cortical subfields using MRI and stereology. *Brain Res Brain Res Protoc*, 10(3):125–138, Feb 2003.
- [35] D. E. Job, H. C. Whalley, S. McConnell, M. Glabus, E. C. Johnstone, S. M. Lawrie. Structural gray matter differences between first-episode schizophrenics and normal controls using voxel-based morphometry. *Neuroimage*, 17(2):880–889, Oct 2002. Clinical Trial.
- [36] D. Kidron, S. E. Black, P. Stanchev, B. Buck, J. P. Szalai, J. Parker, C. Szekely, M. J. Bronskill. Quantitative MR volumetry in Alzheimer’s disease. Topographic markers and the effects of sex and education. *Neurology*, 49(6):1504–1512, Dec 1997.
- [37] T. Kohonen. Self-Organizing Maps. *Springer Series in Information Sciences*, 30:362pp, 1985.
- [38] T. Kohonen. Learning vector quantization for pattern recognition. *Report TKK-F-A, Helsinki University of Technology, Espoo, Finland*, 1995.
- [39] T. Kohonen, J. Hynninen, J. Kangas, J. Laaksonen, K. Torkkola. LVQ_PAK: the learning vector quantization programming package. *Report A30, Helsinki University of Technology, Laboratory of Computer and Information Science, Espoo, Finland*, 1996.
- [40] R. K. Kwan, A. C. Evans, G. B. Pike. MRI simulation-based evaluation of image-processing and classification methods. *IEEE Trans Med Imaging*, 18(11):1085–1097, Nov 1999.

- [41] P. C. Lauterbur. Progress in n.m.r. zeugmatography imaging. *Philos Trans R Soc Lond B Biol Sci*, 289(1037):483–487, Jun 1980.
- [42] C. U. Lee, M. E. Shenton, D. F. Salisbury, K. Kasai, T. Onitsuka, C. C. Dickey, D. Yurgelun-Todd, R. Kikinis, F. A. Jolesz, R. W. McCarley. Fusiform gyrus volume reduction in first-episode schizophrenia: a magnetic resonance imaging study. *Arch Gen Psychiatry*, 59(9):775–781, Sep 2002.
- [43] P. Mansfield und A. A. Maudsley. Medical imaging by NMR. *Br J Radiol*, 50(591):188–194, Mar 1977.
- [44] J. Marshall. On the relations between the weight of the brain and its parts and stature of the body in man. *J. Anat. Physiol.*, 26:445–500, 1892.
- [45] D. H. Mathalon, A. Pfefferbaum, K. O. Lim, M. J. Rosenbloom, E. V. Sullivan. Compounded brain volume deficits in schizophrenia-alcoholism comorbidity. *Arch Gen Psychiatry*, 60(3):245–252, Mar 2003.
- [46] A. K. Miller, R. L. Alston, J. A. Corsellis. Variation with age in the volumes of grey and white matter in the cerebral hemispheres of man: measurements with an image analyser. *Neuropathol Appl Neurobiol.*, 6(2):119–32, 1980.
- [47] H. F. Moselhy, G. Georgiou, A. Kahn. Frontal lobe changes in alcoholism: a review of the literature. *Alcohol Alcohol*, 36(5):357–368, Sep 2001.
- [48] A. Paolillo, C. Pozzilli, E. Giugni, V. Tomassini, C. Gasperini, M. Fiorelli, C. Mainero, M. Horsfield, S. Galgani, S. Bastianello, C. Buttinelli. A 6-year clinical and MRI follow-up study of patients with relapsing-remitting multiple sclerosis treated with Interferon-beta. *Eur J Neurol*, 9(6):645–655, Nov 2002.

- [49] A. Pfefferbaum, D. H. Mathalon, E. V. Sullivan, J. M. Rawles, R. B. Zipursky, K. O. Lim. A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Arch Neurol*, 51(9):874–887, Sep 1994.
- [50] K. von Plessen, A. Lundervold, N. Duta, E. Heiervang, F. Klauschen, A. I. Smievoll, L. Ersland, K. Hugdahl. Less developed corpus callosum in dyslexic subjects—a structural MRI study. *Neuropsychologia*, 40(7):1035–1044, 2002.
- [51] M. R. Pluvinage. Les atrophies cerebrales des alcooliques. *Bull. Mem. Soc. Med. Paris*, 70:524–526, 1954.
- [52] G. Roth. Das Gehirn und seine Wirklichkeit. *Suhrkamp Verlag*, pages 80–96, Aug 2000.
- [53] J. Sijbers, A. J. den Dekker, J. Van Audekerke, M. Verhoye, D. Van Dyck. Estimation of the noise in magnitude MR images. *Magn Reson Imaging*, 16(1):87–90, 1998.
- [54] J. Sijbers, P. Scheunders, N. Bonnet, D. Van Dyck, E. Raman. Quantification and improvement of the signal-to-noise ratio in a magnetic resonance image acquisition procedure. *Magn Reson Imaging*, 14(10):1157–1163, 1996.
- [55] E. R. Sowell, D. A. Trauner, A. Gamst, T. L. Jernigan. Development of cortical and subcortical brain structures in childhood and adolescence: a structural MRI study. *Dev Med Child Neurol*, 44(1):4–16, Jan 2002.
- [56] M. Styner, C. Brechbuhler, G. Szekely, G. Gerig. Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE Trans Med Imaging*, 19(3):153–165, Mar 2000.

- [57] P. S. Tofts. Standardisation and optimisation of magnetic resonance techniques for multicentre studies. *J Neurol Neurosurg Psychiatry*, 64 Suppl 1:S37–43, May 1998.
- [58] D. Trauner, B. Wulfeck, P. Tallal, J. Hesselink. Neurological and MRI profiles of children with developmental language impairment. *Dev Med Child Neurol*, 42(7):470–475, Jul 2000.
- [59] B. Tumarkin, J. D. Wilson, G. Snyder. Cerebral Atrophy due to Alcoholism in Young Adults. *United States Armed Forces medical Journal*, 6:67–74, 1955.
- [60] M. Uhl, M. P. Hauer, K. H. Allmann, H. Gufler, J. Laubenberger, J. Hennig. Recent developments and applications of MRI sequence technique. I: turbo spin echo, HASTE, turbo inversion recovery, turbo gradient echo, turbo gradient spin sequences. *Aktuelle Radiol*, 8(1):4–10, Jan 1998.
- [61] H. Wanifuchi, T. Shimizu, T. Maruyama. Age-related changes in the proportion of intracranial cerebrospinal fluid space measured using volumetric computerized tomography scanning. *J Neurosurg*, 97(3):607–610, Sep 2002.

Internet-Adressen

AFNI	http://afni.nimh.nih.gov/afni/
Freesurfer	http://surfer.nmr.mgh.harvard.edu
LVQ	http://www.cis.hut.fi/research/
Matlab-EEG-Toolbox	http://sourceforge.net/projects/eeg/
MRI Simulator	http://www.bic.mni.mcgill.ca/brainweb/
SPM99	http://www.fil.ion.ucl.ac.uk/spm/dox.html
The basics of MRI	http://www.cis.rit.edu/htbooks/mri
The Visible Human Project	http://www.nlm.nih.gov/research/visible/visible_human.html

Danksagung

Ich danke Herrn Prof. Dr. Dirk Petersen für die Betreuung und konstruktive Kritik bei der Fertigstellung der Arbeit. Des Weiteren möchte ich mich bei Herrn Prof. Dr. Arvid Lundervold für die freundliche Aufnahme in seiner Arbeitsgruppe, die Vergabe des Dissertationsthemas und die Anregungen zur Bearbeitung bedanken.

CURRICULUM VITAE

NAME Frederick Klauschen
ADRESSE Domhof 23
D-23909 Ratzeburg
E-mail: fklusach@physnet.uni-hamburg.de
GEBURTSTAG/-ORT 2.7.1974 in Itzehoe, Schleswig-Holstein
FAMILIENSTAND ledig
NATIONALITÄT deutsch

SCHULAUSBILDUNG

1981-85 Grundschule Nord, Halstenbek, Kreis Pinneberg
1985-94 Wolfgang-Borchert-Gymnasium, Halstenbek, Abitur

WEHRDIENST

1994-95 Marinesanitätsdienst, Versorgungsschiff Donau, Warnemünde

UNIVERSITÄTSAUSBILDUNG

ab 10/1995 Medizinstudium Universität Lübeck
seit 10/1997 Kurse in Mathematik und Physik, Universität Lübeck/Hamburg
01/1999 Erstes Staatsexamen in Medizin
07-10/1999 Praktikum in medizinischer Informatik, Beginn der Bearbeitung des Dissertationsthemas, Prof. A. Lundervold MD PhD, Institut für Physiologie, Universität Bergen, Norwegen
10/2000 Zweites Staatsexamen in Medizin
10/00-02/01 Praktisches Jahr, Innere Medizin, Medizinische Klinik I, Universität Lübeck

02-05/2001	Praktisches Jahr, Chirurgie, Centre Hospitalier Sud Réunion, service de chirurgie infantile et orthopédique, St. Pierre, Ile de la Réunion, Frankreich
05/2001	Teilnahme an der <i>Bergen</i> -Fahrt der Universität Lübeck, Vorstellung der Zwischenergebnisse der Doktorarbeit an der Universität Bergen, Norwegen
05-07/2001	Praktisches Jahr, Radiologie, Universität Lübeck
11/2001	Drittes Staatsexamen in Medizin
07-10/2001	Praktikum in Computational Biology, Dr. W.E. Paul, Laboratory
07-09/2002	of Immunology, National Institutes of Health, Bethesda, USA
seit 04/2002	Physikstudium Universität Hamburg
02/2003	Diplomvorprüfung in Physik, Universität Hamburg
bis 03/2004	Diplomprüfungen in Physik, Universität Hamburg
seit 04/2004	Visiting Fellow, National Institutes of Health, USA, und Diplomand, II. Institut für Theoretische Physik, Universität Hamburg