

Otto-von-Guericke-Universität Magdeburg
Fakultät für Informatik
Institut für Simulation und Grafik



Diplomarbeit

Implementierung und Evaluierung von Visualisierungs- und Interaktionstechniken in Patientendatenbanken

Maik Mewes

20. Dezember 2007

Otto-von-Guericke-Universität Magdeburg
Fakultät für Informatik
Institut für Simulation und Grafik

Diplomarbeit

Implementierung und Evaluierung von Visualisierungs- und Interaktionstechniken in Patientendatenbanken

Verfasser: Maik Mewes
Matrikelnummer: 16 25 12

1. Gutachter: Prof. Dr.-Ing. BERNHARD PREIM
2. Gutachter: Jun.-Prof. Dr.-Ing. RAIMUND DACHSELT

Betreuer: Dipl.-Ing. JEANETTE CORDES
Dipl.-Ing. RAGNAR BADE

Bearbeitungszeitraum: Juni bis Dezember 2007

Selbstständigkeitserklärung

Hiermit erkläre ich, Maik Mewes (Matrikel-Nr. 162512), dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Quellen angefertigt habe.

Maik Mewes
Magdeburg, den 10. Dezember 2007

Danksagung

An dieser Stelle möchte ich den Menschen danken, ohne die diese Arbeit nicht möglich gewesen wäre.

Besonderer Dank gilt meiner Frau Christiane, die mich in dieser Zeit ertragen musste. Danke, dass es dich gibt.

Ich möchte meinen beiden Betreuern, Dipl.-Ing. Jeanette Cordes und Dipl.-Ing. Ragnar Bade, dafür danken, dass sie mir die gesamte Zeit über mit ihrer sehr guten Hilfe und ihren nützlichen Ratschlägen zur Seite standen.

Meinen Freunden danke ich dafür, dass sie bei der Evaluierung so zahlreich mitgemacht und dass sie nicht allzu böse darüber sind, die letzten Monate kaum etwas von mir gehört zu haben.

Auch meinen Mitdiplomanden, den Mitarbeitern des Lehrstuhls für Visualisierung und meinen Kollegen bei icubic möchte ich dafür danken, dass sie die Evaluierung durch ihre rege Teilnahme möglich gemacht haben. Meinen Chefs bin ich besonders dankbar, da sie mir bei icubic die Freiräume ließen, die nötig sind, um eine solche Arbeit zu schreiben.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation und Zielsetzung	2
1.2	Aufbau der Arbeit	2
2	Grundlagen	5
2.1	Visualisierungstechniken	5
2.1.1	Tabellen	6
2.1.2	TableLens	8
2.1.3	InfoZoom	8
2.1.4	Parallele Koordinaten	10
2.1.5	Weitere Techniken	12
2.2	Evaluierungsmethoden	14
2.2.1	Formale Evaluierung	15
2.2.2	Heuristische Evaluierung	15
2.2.3	Empirische Evaluierung	16
2.3	Verwandte Arbeiten	17
2.3.1	Fallstudien	18
2.3.2	Vergleichende Studien	21
2.4	Kapitelzusammenfassung	24
3	Entwurf	25
3.1	Anforderungen an eine effiziente Fallauswahl	25
3.1.1	Die Falldaten	26
3.1.2	Die Fallanzahl	27
3.1.3	Die Fallauswahl	27
3.1.4	Stellenwert der Fallauswahl	29
3.2	Geeignete Techniken für die Fallauswahl	29
3.2.1	Tabellen	30
3.2.2	TableLens	30
3.2.3	InfoZoom	31
3.2.4	Parallele Koordinaten	31
3.2.5	Weitere Techniken	32
3.2.6	Thesen zur Benutzung der Visualisierungen	32
3.3	Rahmenbedingungen für eine Evaluierung	34

3.3.1	Falldaten	34
3.3.2	Testpersonen	35
3.3.3	Testaufgaben für die Techniken	35
3.3.4	Zu erhebende Daten	37
3.4	Konzept des VisualizationEvaluators	39
3.5	Kapitelzusammenfassung	41
4	Umsetzung	43
4.1	Werkzeuge	43
4.2	Techniken	44
4.2.1	Datenaquirierung	44
4.2.2	Einbettung der Techniken	45
4.2.3	Aufbau der Techniken allgemein	45
4.2.4	Tabellen	47
4.2.5	TableLens	48
4.2.6	InfoZoom	49
4.2.7	Parallele Koordinaten	51
4.3	Der VisualizationEvaluator	54
4.4	Kapitelzusammenfassung	56
5	Durchführung, Auswertung & Ergebnisse der Evaluierung	59
5.1	Durchführung	59
5.1.1	Ablauf	60
5.2	Eingegangene Daten	63
5.2.1	Verwertbarkeit der erhobenen Daten	63
5.3	Auswertung der gemessenen Daten	65
5.3.1	Untersuchung der gemessenen Zeiten	65
5.3.2	Korrektheit	68
5.4	Auswertung der subjektiven Bewertung der Techniken	71
5.4.1	Eignung der Techniken für die Aufgaben	71
5.4.2	Bewertung der Techniken nach verschiedenen Aspekten	72
5.4.3	Kommentare und Verbesserungsvorschläge	73
5.4.4	Empfehlung für eine Technik	75
5.5	Ergebnisse und Kapitelzusammenfassung	78
5.5.1	Eignung der Techniken für bestimmte Aufgaben	79
5.5.2	Empfehlung einer Visualisierung	80
6	Zusammenfassung & Ausblick	83
6.1	Zusammenfassung	83
6.2	Weitergehende Untersuchungen	85
6.2.1	Charakterisierung der Fallauswahl	85
6.2.2	Weiterentwicklung der Techniken und mögliche Folgen für das Evaluierungsergebnis	85
6.2.3	Mischung von Aufgaben verschiedener Kategorien	86

6.2.4 Weitere Evaluierungsszenarien	87
Literaturverzeichnis	89
Tabellenverzeichnis	95
Abbildungsverzeichnis	97

Inhaltsverzeichnis

1 Einleitung

In den letzten Jahren haben Computersysteme für die Operations- und Behandlungsplanung in der Medizin immer mehr an Bedeutung gewonnen. Diese Computer-Based-Training-Systeme (CBT) sparen Kosten und minimieren Risiken [48], da angehende Ärzte durch sie Erfahrung sammeln können, ohne am realen Patienten aktiv zu werden. Der weit verbreitete Einsatz von fallbasierten Systemen wie CAMPUS [19] und CASUS [2] ist ein Beispiel für die effiziente Integration von Computersoftware in die Ausbildung von Medizinerinnen und zeigt wie diese effektiv unterstützt werden kann. Anhand von realen oder künstlich erstellten Krankheitsfällen wird die Vorgehensweise bei der Behandlung oder der Operationsdurchführung im Vorhinein trainiert und gefestigt.

Allen diesen fallbasierten Systemen ist eines gemein: Sie behandeln die Auswahl des zu trainierenden Falles derzeit noch sehr unzureichend. In den meisten Fällen dient hier eine simple Listenauswahl der eher zufälligen Selektion eines Trainingsfalles.

Falldaten besitzen jedoch eine Vielzahl an Parametern, die für die Wahl eines Trainingsfalles relevant sein können. Neben einer Fallnummer und einem Patientennamen stehen in einem aufbereiteten Trainingsfall auch Informationen zur Diagnose oder der bisherigen Behandlung zur Verfügung. Neuartige Systeme wie der LIVERSURGERYTRAINER [11] verwenden Metadaten, um z. B. den Bearbeitungsstand oder den Schwierigkeitsgrad eines Trainingsfalles bereitzustellen. Auch solche Daten scheinen geeignet im Auswahlmechanismus eine Rolle zu spielen. Diese und zusätzliche Informationen zur speziellen Ausprägung der Krankheit bilden den Informationsraum eines Falles. Es muss möglich sein, anhand dieser speziellen Charakteristik einen Fall auszuwählen und diese Charakteristik im Kontext aller verfügbaren Fälle einzuschätzen und zu vergleichen [33]. Nur so kann garantiert werden, dass auch der richtige Fall trainiert wird. Die aktuelle Auswahl über eine einfache Listendarstellung bietet dazu nicht die nötigen Möglichkeiten. Das wissenschaftliche Feld der Informationsvisualisierung befasst sich mit Methoden zur geeigneten Darstellung eben solcher multivariaten Datensätze. Spezielle Techniken werden entwickelt, um möglichst viele Daten gleichzeitig für das menschliche Auge bearbeitbar darzustellen und somit einen Überblick über die Daten zu vermitteln. Neue Werkzeuge und Interaktionsmöglichkeiten werden geschaffen, um große Datenmengen zu explorieren und so neue Erkenntnisse und verborgene Zusammenhänge zwischen den verschiedenen Datenattributen zu erkennen.

1.1 Motivation und Zielsetzung

Die Auswahl eines zu trainierenden Falles geschieht gezielt anhand einer speziellen Charakteristik, einer speziellen Motivation oder vor einem speziellen Hintergrund. So will sich ein Student eventuell auf eine Prüfung vorbereiten oder der angehende Fachchirurg vor einer anstehenden Operation noch einmal möglichst analoge Fälle trainieren, um optimal vorbereitet zu sein. Aus diesen Gründen wird mit der vorliegenden Arbeit eine Analyse des Vorgangs der Fallauswahl in einem CBT-System vorgenommen und die Eignung verschiedener Techniken aus der Informationsvisualisierung für eben diese untersucht. Geeignete Techniken werden umgesetzt und mit Hilfe einer empirischen Evaluierung anhand von Beispielfalldaten, wie sie im LIVERSURGERYTRAINER vorkommen, verglichen.

Durch die Evaluierung soll herausgefunden werden, ob sich Visualisierungstechniken aus der Datenanalyse und -exploration zur Präsentation medizinischer Falldaten eignen, ob Vorteile einer Technik eine eventuell erschwerte Benutzung rechtfertigen und inwieweit die Techniken von potentiellen Nutzern akzeptiert werden. Das Hauptziel der Arbeit ist es, für die Fallauswahl geeignete Visualisierungstechniken zu identifizieren und zu implementieren. Mit Hilfe des Evaluierungsergebnisses wird eine Empfehlung bezüglich der einzusetzenden Techniken und ihrer Eignung für bestimmte Aufgabenbereiche ausgesprochen.

1.2 Aufbau der Arbeit

Grundlagen Im Grundlagenkapitel werden einige Techniken aus dem Feld der Informationsvisualisierung vorgestellt, die hinsichtlich der Fallauswahl geeignet scheinen und ihre Interaktionsmöglichkeiten werden erläutert. Des Weiteren werden theoretische Grundlagen verschiedener Evaluierungsmethoden dargestellt und schließlich, durch einen Überblick der verwandten Arbeiten im Gebiet der Evaluierung von Informationsvisualisierungen, der Stand der Wissenschaft herausgearbeitet. Schlussfolgernd können bestehende Forschungslücken aufgedeckt werden, welche die vorliegende Arbeit abzudecken sucht. So muss untersucht werden, welchen Nutzen komplexe Visualisierungsansätze in einem von der Datenanalyse entfernten Kontext, wie der Fallauswahl, gegenüber einfachen Techniken wie der Tabelle bieten. Auch die Akzeptanz dieser erweiterten Visualisierungen muss in diesem Umfeld neu erschlossen werden.

Konzept Die in den Grundlagen gewonnenen Erkenntnisse werden im dritten Kapitel in ein wissenschaftliches Konzept gewandelt. Eine Analyse des Vorgangs der Fallauswahl wird genutzt, um geeignete Techniken für eine Implementierung auszuwählen. Anhand von Beispielszenarien werden Aufgaben, Daten und andere Rahmenbedingungen für eine Evaluierung zusammengestellt und es wird diskutiert, wie eine solche durchgeführt werden kann.

Umsetzung Im Umsetzungskapitel wird näher auf die Implementierung der ausgewählten Techniken und ihrer Interaktionen eingegangen. Es werden wichtige Werkzeuge und Schnittstellen erläutert und der Aufbau des VisualizationEvaluators, dem Programm zur vollautomatischen Durchführung der Evaluierung, beschrieben.

Durchführung, Auswertung & Ergebnisse Im fünften Kapitel wird die Durchführung der Evaluierung zusammengefasst und die Ergebnisse der statistischen Auswertung der erhobenen Daten beschrieben. Zudem wird eine Diskussion und Interpretation der subjektiven Bewertungen der Testpersonen vorgenommen. Aus diesen beiden Bereichen werden dann die Ergebnisse der Evaluierung extrahiert und erläutert.

Zusammenfassung und Ausblick Im letzten Abschnitt der Arbeit wird ein zusammenfassender Rückblick gegeben, es wird herausgestellt ob die Ziele der Arbeit erreicht wurden und welche Thematiken in Zukunft noch zu untersuchen sind.

1 Einleitung

2 Grundlagen

In diesem Kapitel wird eine kurze Übersicht über verschiedene Visualisierungstechniken und deren Interaktionsformen gegeben. Im weiteren Verlauf werden Methoden zur Evaluierung von Informationsvisualisierungen besprochen und in einem Überblick der Stand der Forschung gezeigt. Es wird angegeben, welche Evaluierungen im Bezug auf die gewählten Techniken stattgefunden haben und welche Erkenntnisse sie hervorgebracht haben. Schlussfolgernd werden Forschungslücken, im Bezug auf den Einsatz dieser Techniken in medizinischen Datenbanken und den Transfer ihres Einsatzes in Aufgabengebiete abseits der Datenanalyse, beispielsweise in einer Fallauswahl eines medizinischen CBT-Systems, offengelegt. Aus diesen wird im folgenden Kapitel der konkrete Kontext dieser Arbeit abgeleitet.

2.1 Visualisierungstechniken

Techniken der Informationsvisualisierung werden dort eingesetzt, wo extrem viele Daten anfallen. So werden z. B. groß angelegte Datenbanken, wie Automobilkataloge, mit Tausenden von Einträgen über solche Techniken visualisiert (Abbildung 2.1). Diese lassen dabei Interaktionen zu, welche das Explorieren und Auswerten der präsentierten Datensätze ermöglichen. Wie in Abschnitt 1.1 erwähnt werden Fallauswahlen und ihre Bedeutung in medizinischen Trainingssystemen zur Zeit noch unterschätzt und die bisher eingesetzte Variante der Listendarstellung der Fälle ohne Interaktionsmöglichkeiten ist für diese Aufgabe nur unzureichend (Abbildung 2.2). Auch wenn die einzelnen Falldatenbanken eines Trainingssystems noch überschaubar groß sind, so wachsen diese Datenbanken jedoch stetig an, oder werden sogar in Verbundprojekten gesammelt [14]. Spätestens dann können auch diese Datenbanken unübersichtlich werden.

Mit der vorliegenden Arbeit soll auch überprüft werden, welche Visualisierungstechnik für den Einsatz in einer Fallauswahl besonders geeignet ist. Aus diesem Grund werden in diesem Abschnitt einige Techniken erläutert, welche für die Fallauswahl geeignet scheinen.

2 Grundlagen



Abbildung 2.1: Suche im Automobilkatalog im ADAC-Kundenportal mit INFOZOOM (modifiziert nach [1]).

2.1.1 Tabellen

Tabellen finden in vielen Kalkulationsprogrammen oder auch in relationalen Datenbanken Anwendung. Die einfache und intuitive Anordnung von Daten in Zellen einer Gitterstruktur wurde aber auch vor dem Computerzeitalter schon z. B. in Ahnentafeln genutzt.

Bei einer Tabelle werden Daten in Zellen einer Gitterstruktur hinterlegt. Sie kann eine Vorspalte und eine Kopfzeile besitzen und meist besteht ein semantischer Zusammenhang zwischen einer Datenzelle und der Spalte oder Zeile in der sie eingeordnet ist. Abbildung 2.3 zeigt eine übliche Tabelle in einem Computerprogramm.

Am Computer wird die Tabelle erst durch Interaktionsmöglichkeiten als Visualisierungstechnik interessant. Möglichkeiten hierfür sind das Sortieren, oft auch nach mehreren Spalten bzw. Zeilen oder auch das Umordnen von Spalten. Eine weitere Möglichkeit der Interaktion und klare Erweiterung bei Tabellen stellt unter anderem auch die hierarchische Ordnung von Zeilen dar, welche dann wie in einer Baumstruktur ein- und ausklappbar sind.

CASUS®

→ Ausloggen ? Hilfe

Fallauswahl

Fälle:	
1	48-jähriger Psychologe mit starkem Brustschmerz Mögliche Differentialdiagnosen bei Brustschmerz und entsprechende Therapieansätze. Autor(en): A. Meier, E. Catherwood, H. Lyon, B. Aulinger
2	9-jähriger Junge mit Fieber Ein Fall zur Infektiologie. Autor(en): Sylvia Handl, Tim Müller

(a) CASUS Fallauswahl

← CAMPUS Lehr- und Lernsystem ▶ CAMPUS ▶ Classic-Player ▶
[Online-Version](#)

Säugling mit Fieber

Fachgebiet(e): Kinderheilkunde

Autor(en): Dr. med. S. Huwendiek, Dr. med. B. Höcker, Dr. med. C. Seidel

Kurzinfo: Frau M. stellt Ihnen in der Ambulanz der Kinderklinik ihre fünf Monate alte Tochter Katrin vor, die seit zwei Tagen Fieber habe und heute zunehmend schlapp wirke.

[Fall starten](#)

Säugling mit Erbrechen

Fachgebiet(e): Kinderheilkunde

Autor(en): Dr. med. S. Huwendiek, Dr. med. S. Köpf

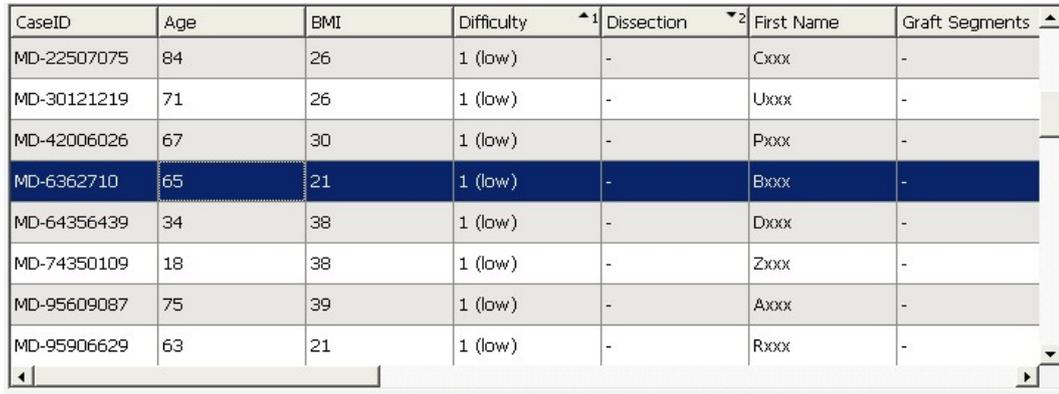
Kurzinfo: Es ist 14.00 Uhr. Sie sind Ärztin/Arzt in der Ambulanz der Kinderklinik und werden zu dem sieben Wochen alten männlichen Säugling "Lars" gerufen, der seit fünf Tagen immer wieder erbreche.

[Fall starten](#)

(b) CAMPUS Fallauswahl

Abbildung 2.2: Fallauswahlen verschiedener medizinischer Trainingssysteme. Es wird auf bewährte, aber unübersichtliche Listendarstellung gesetzt. (a) zeigt die Fallauswahl des CASUS Lernsystem (modifiziert nach [3]), (b) die Entsprechung im System CAMPUS der Uni Heidelberg (modifiziert nach [21]).

2 Grundlagen



CaseID	Age	BMI	Difficulty	Dissection	First Name	Graft Segments
MD-22507075	84	26	1 (low)	-	Cxxx	-
MD-30121219	71	26	1 (low)	-	Uxxx	-
MD-42006026	67	30	1 (low)	-	Pxxx	-
MD-6362710	65	21	1 (low)	-	Bxxx	-
MD-64356439	34	38	1 (low)	-	Dxxx	-
MD-74350109	18	38	1 (low)	-	Zxxx	-
MD-95609087	75	39	1 (low)	-	Axxx	-
MD-95906629	63	21	1 (low)	-	Rxxx	-

Abbildung 2.3: Darstellung einer üblichen Tabellenansicht, wie sie in vielen Computerprogrammen Anwendung findet.

2.1.2 TableLens

Die TABLENS, erstmals vorgestellt von Rao und Card [40], ist eine Visualisierungstechnik, welche basierend auf einer normalen Tabellendarstellung die Fokus & Kontext-Metapher¹ verwendet. Einzelne Datenzellen werden verkleinert dargestellt, und die textuelle Darstellung des Inhaltes, abhängig vom Datentyp, durch eine grafische Repräsentation ersetzt (Abbildung 2.4). Damit wird es möglich, im Verhältnis deutlich mehr Datenzellen gleichzeitig auf dem Bildschirm darzustellen. So werden z. B. numerische Daten auf einen Balken abgebildet, der die Zelle entsprechend des Wertes ausfüllt. So entsteht über eine gesamte Spalte ein Balkendiagramm, das durch eine Sortierung in eine Form gebracht werden kann, aus der grob die statistische Verteilung der Daten im Wertebereich abgelesen werden kann. Interessante Bereiche der Tabelle können in eine Fokuslinse gebracht werden, um sie näher zu untersuchen. Hierbei wird die entsprechende textuelle Repräsentation wieder mit angeboten. Verschiedene Interaktionsformen ermöglichen das Verschieben, Vergrößern oder Verkleinern der Linse. Abbildung 2.5 zeigt eine Darstellung von Baseballstatistiken mit der TABLENS und verdeutlicht schematisch das Fokus & Kontext-Prinzip.

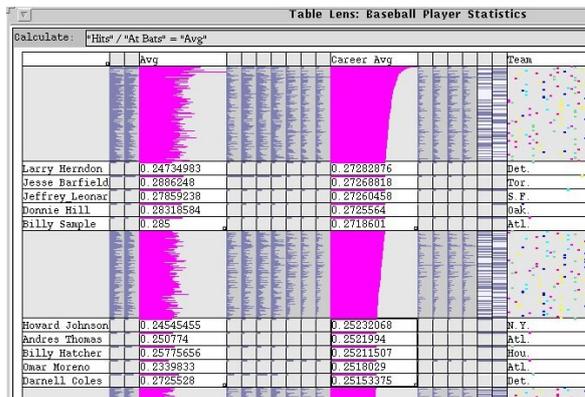
2.1.3 InfoZoom

Der INFOZOOM beschreibt eine weitere Technik, die auf einem tabellarischen Ansatz beruht. Beim INFOZOOM wird allerdings zu Anfang der gesamte Inhalt der Datenbank in einer komprimierten Tabelle dargestellt. Dies wird durch den Austausch von Zeilen und Spalten erreicht, die Variablen werden auf Tabellenzeilen und die

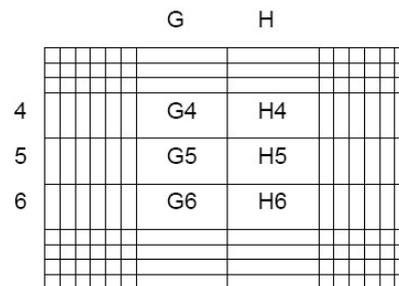
¹Bei der Fokus & Kontext-Metapher geht es darum, sowohl Detail- als auch Umgebungsinformationen von betrachteten Elementen gleichzeitig und übersichtlich darzustellen.



Abbildung 2.4: Darstellung einer TABLE LENS-Visualisierung ohne Linsenbereich (erstellt mit einer Online-Demonstration der TABLE LENS [24]).



(a) Darstellung mit Linsenbereichen



(b) Fokusprinzip schematisch

Abbildung 2.5: Darstellung (a) zeigt eine TABLE LENS-Visualisierung mit mehreren Fokusbereichen und in (b) die schematische Verdeutlichung des Fokusprinzips (beide modifiziert nach [40]). Die Zellen G4 - H6 befinden sich im fokussierten Bereich.

Datensätze auf Tabellenspalten abgebildet. Spalten können je nach Größe der Datenbank auch eine Größe im Subpixelbereich besitzen. Die starre Gitterstruktur einer Tabelle wird aufgebrochen und benachbarte Zellen gleichen Inhalts in einer Zeile zusammengefasst, um auch bei großen Datenbanken eine textuelle Beschreibung zu ermöglichen. Ähnlich der TABLE LENS werden numerische Daten zusätzlich grafisch in einer Art Balkendiagramm erfasst, die es ermöglicht über Sortieroperationen die Werteverteilung einer Variable zu beurteilen. Zusätzlich zu den gewöhnlichen Inter-

2 Grundlagen

aktionen einer Tabelle, dem Sortieren, dem Selektieren und der Variablenanordnung durch Spaltenverschiebung, können uninteressante Daten über die Zoomoperation ausgeblendet werden. So wird es möglich, sich Schritt für Schritt dem für die aktuelle Situation wichtigen Dateninhalt anzunähern. Wenn der betrachtete Bereich genügend weit eingegrenzt ist, bietet INFOZOOM die Möglichkeit, die übrigen Daten auch unkomprimiert, in der für Tabellen typischen Gitterstruktur, anzuzeigen. Zur weiteren Datenanalyse bietet INFOZOOM die Möglichkeit weitere Variablen einzufügen, die Berechnungen in den vorhandenen Daten darstellen und es können statistische Größen wie Mittel- und Medianwert hervorgehoben werden. Eine weitere Überblicksansicht dient der näheren Beurteilung von Datenmengen und statistischen Verteilung im Vergleich der Variablen. Hier werden die Datensätze komplett vom tabellarischen Ansatz getrennt. Abbildung 2.6 beinhaltet INFOZOOM die drei möglichen Ansichten.

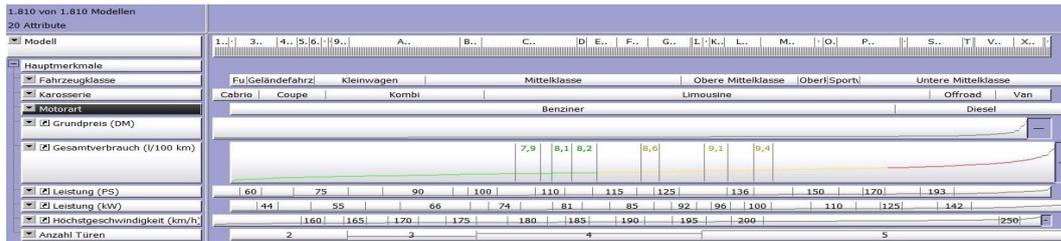
2.1.4 Parallele Koordinaten

PARALLELE KOORDINATEN sind ein Ansatz um hochdimensionale Daten auf den zweidimensionalen Raum zu projizieren um sie für die weitere Verarbeitung besser interpretieren zu können. Sie können in der Robotik, der Flugverkehrsüberwachung und insbesondere auch auf dem Gebiet der Statistik zur Anwendung kommen [22].

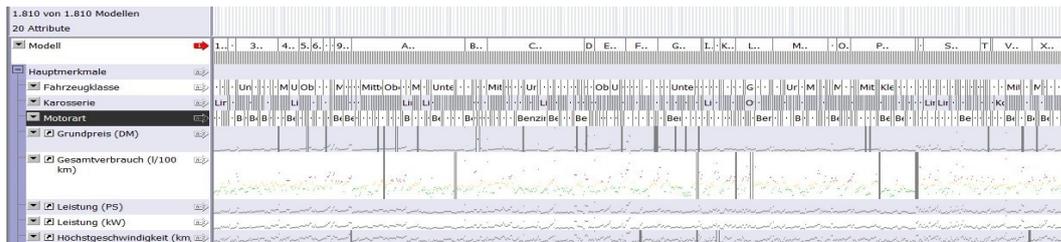
Bei dieser Technik werden die Variablen bzw. Dimensionen auf parallel angeordneten Achsen abgebildet. Die Achsen befinden sich in einem einheitlichen Abstand zueinander und sie enthalten sortiert alle im gesamten Datensatz vorkommenden Werte der ihnen entsprechenden Variable. Ein Punkt im hochdimensionalen Raum wird dann abgebildet auf einen Linienzug, welcher die Koordinatenachsen jeweils an dem Punkt schneidet, an dem der Wert des abgebildeten Datensatzes auf der Achse abgetragen ist.

Das Brushing, also das Eingrenzen der einbezogenen Werte hebt Untermengen der Datenbank hervor und hilft so die interessierenden Bereiche in den Daten zu finden [32]. Weiterführende Interaktionen wie das Dimensionzooming, welches einem Neuskalieren einer Dimensionsachse entspricht oder ein hierarchisches Clustering [16] verhelfen zu einer verbesserten Übersicht in der weiteren Exploration und begegnen so dem größten Problem dieser grafischen Darstellung von Datensätzen: Der Überzeichnung. Abbildung 2.7 zeigt verschiedene Darstellungen von PARALLELEN KOORDINATEN.

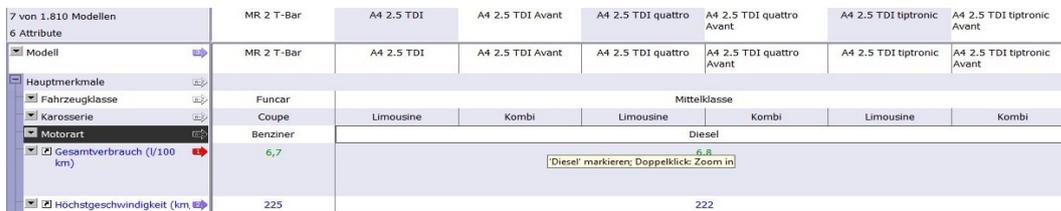
Farbhistogramme, Dichte- sowie Sättigungsplots [51, 50] sind weitere Möglichkeiten, die Überzeichnung zu begrenzen. Farbhistogramme legen Farbgradienten über eine Koordinatenachse, die Liniensegmente, welche die Achse schneiden werden entsprechend eingefärbt (Abbildung 2.8 (a)). Dichtefunktionen, die auf die Daten angewendet werden, oder das einfache Zählen der Überzeichnungen einzelner Achsenabschnit-



(a) Überblicksansicht



(b) Komprimiert



(c) Komprimierte Sicht eingegrenzt



(d) Eingegrenzt und unkomprimiert

Abbildung 2.6: Die verschiedenen Sichten der INFOZOOM-Technik. (a) Überblicksansicht einer Datenbank mit Automobilkennzahlen wie Größe, Höchstgeschwindigkeit und Verbrauch (1810 Einträge). (b) Die Datenbank in einer komprimierten Tabelle. (c) Der betrachtete Bereich auf Fahrzeuge begrenzt, die einen Durchschnittsverbrauch von unter sieben Litern pro 100 gefahrenen Kilometern aufweisen und eine Höchstgeschwindigkeit von über 220 km/h erlauben. Diese Eigenschaften werden von sieben Fahrzeugen erfüllt. (d) Abschnitt der Datenbank aus (c) unkomprimiert (Alle Darstellungen erstellt mit InfoZoomViewer [49]).

2 Grundlagen

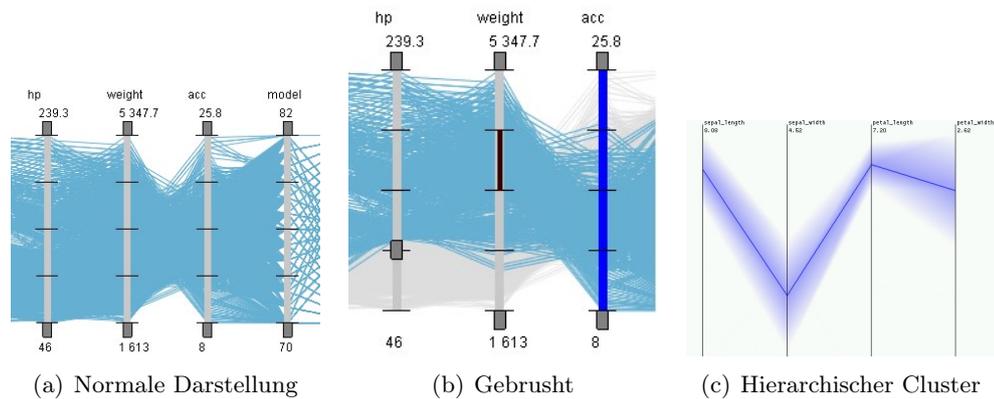


Abbildung 2.7: Darstellung von PARALLELEN KOORDINATEN mit und ohne verschiedene Erweiterungen. (a) zeigt eine normale Darstellung, (b) eine Darstellung mit einem ausgegrauten, sprich gebrushten Bereich (beide aus [13]). (c) zeigt eine Darstellung eines hierarchischen Clusters. Die Größe des Clusters wird über die Breite des Bandes kodiert (aus [16]).

te können auf Farbsättigungen angewendet werden, so dass stark geclusterte Daten stärker hervortreten (Abbildung 2.8 (b)).

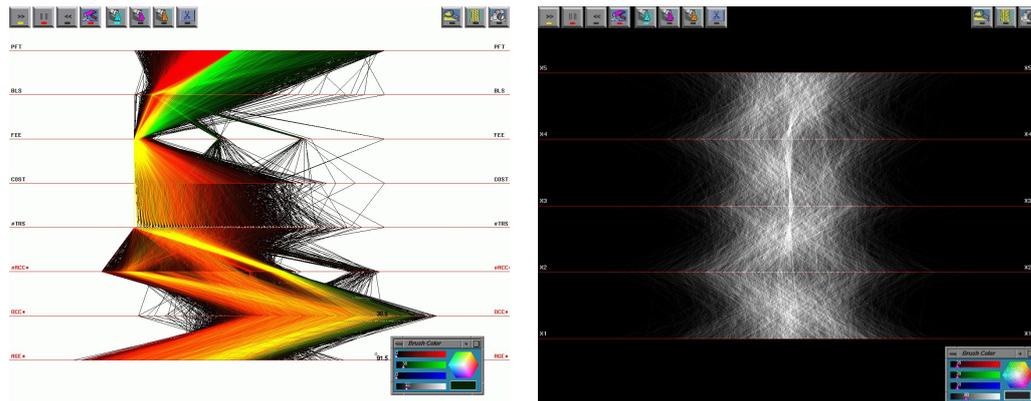
2.1.5 Weitere Techniken

Im folgenden Abschnitt sollen in Kürze weitere Visualisierungsansätze erklärt werden, die in der folgenden Argumentation erwähnt werden, aber von untergeordneter Bedeutung sind, da sie sich im Kontext dieser Arbeit nicht einsetzen lassen.

PROJEKTION PURSUIT beschreibt einen Visualisierungsansatz, welcher die Projektion von hochdimensionalen Daten auf die zweidimensionale Ebene auf ein Optimierungsproblem einer zu definierenden Indexfunktion abbildet [15]. Es entsteht eine dem Scatterplot² ähnliche Darstellung, deren Interpretation zunächst unklar ist. Aufgrund der mathematischen Grundlagen dieses Ansatzes ist es jedoch möglich, auch eine Interpretationshilfe darzustellen (Abbildung 2.9).

RADVIS [10] stellt einen weiteren Ansatz dar, multivariate Daten in der Ebene zu visualisieren. Er basiert auf einem physikalischen Federkraftmodell. Die Dimensionen der Daten werden auf Kreissegmente abgebildet. Die Datenpunkte werden nach einer

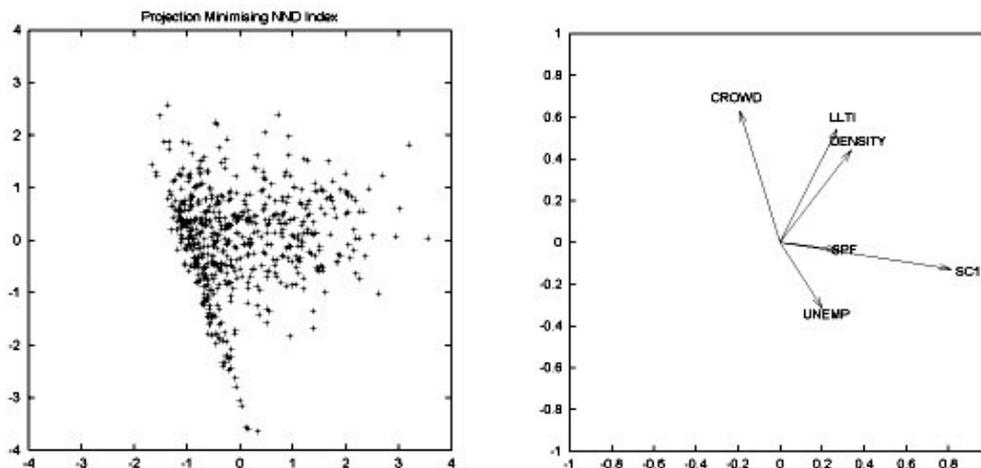
²Ein Scatterplot ist eine grafische Repräsentation von typischerweise zwei Variablen eines Datenbestandes. Jeder Datensatz des Datenbestandes wird als Punkt in einem kartesischen Koordinatensystem, welches durch die entsprechenden Daten-Dimensionen aufgespannt wird, dargestellt. Daraus resultiert eine Visualisierung in einem Streudiagramm, dem Scatterplot.



(a) Farbhistogrammdarstellung

(b) Sättigungsdarstellung

Abbildung 2.8: Farbhistogramm(a) und Sättigungsplot(b) von PARALLELEN KOORDINATEN (beide aus [52]).



(a) Optimale Projektion

(b) Interpretation der Projektion

Abbildung 2.9: Darstellung einer optimalen Projektion (a) und ihrer Interpretationshilfe (b) in PROJEKTION PURSUIT (beide aus [10]).

Normalisierung entsprechend ihres Wertes in den Dimensionen in den Kreissegmenten abgetragen (Abbildung 2.10).

Bei IVEE [4] bzw. SPOTFIRE, seinem Nachfolger, handelt es sich um ein vollintegriertes System zur Datenanalyse und -exploration. Es vereint mehrere Diagrammtypen und die Grundvisualisierung des Scatterplots mit der tabellarischen Darstellung der geladenen Daten. Die Fokussierung auf interessierende Untermengen in

2 Grundlagen

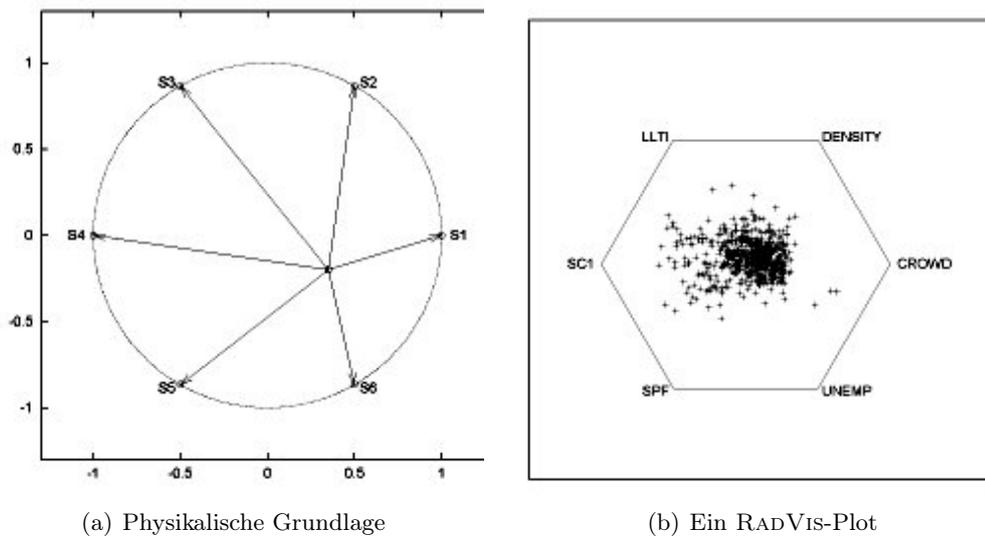


Abbildung 2.10: RADVIS-Visualisierungsansatz: (a) veranschaulicht das zugrunde liegende physikalische Modell, (b) zeigt einen typischen RADVIS-Plot (beide aus [10]).

den Daten geschieht durch bekannte Oberflächenelemente wie Slider und Checkboxes.

2.2 Evaluierungsmethoden

Zur Evaluierung der Usability von Software gibt es verschiedene Methoden. So werden die Testpersonen bei der Evaluation anhand von Software Guidelines, mit von Experten veröffentlichten definierten Richtlinien, wie z. B. bei Brown [9] bekannt gemacht, nach denen die Benutzung eines neuen Softwareproduktes bewertet werden soll [25]. Bei der Methode des Cognitive Walkthroughs von Lewis und Wharton [30] werden typische Handlungsabläufe des zu bewertenden Produktes von Experten ausgeführt, welche versuchen, sich in die Denkprozesse der Endnutzer hineinzusetzen. Sowohl die Methodik des Cognitive Walkthroughs als auch die Erstellung von Software Guidelines beruhen auf Erkenntnissen der Kognitionsforschung. Im Folgenden wird auf weitere Formen der Evaluierung eingegangen. Die im Rahmen der vorliegenden Arbeit durchgeführte Evaluierung ist eine empirische Studie, weshalb diese Methode näher erläutert wird.

2.2.1 Formale Evaluierung

Bei der formalen Evaluierung nach dem GOMS-Modell wird die Benutzung eines Programms anhand von Zielen, Operatoren, Methoden, und Selektionsregeln charakterisiert.

Ziele (Goals) charakterisieren dabei Systemzustände, welche der Benutzer durch Interaktionen erreichen möchte, z. B. wenn eine Spalte einer Tabelle sortiert werden soll.

Operatoren beschreiben die elementaren Interaktionstechniken, die der Nutzer in einem System verwenden kann. So ein System kann beispielsweise Mausbewegungen und -klicks oder Tastatureingaben verarbeiten, beispielsweise ein Mausklick auf einen Spaltenkopf.

Methoden stellen die Verknüpfung mehrerer Operatoren dar, welche nötig sind, um ein Ziel zu erreichen. So muss der Benutzer für eine Sortierung die Hand zur Maus führen, die Maus über den Spaltenkopf bewegen und einen Klick ausführen.

Selektionsregeln kommen zum Einsatz wenn mehrere Methoden zur Erreichung eines Zieles gewählt werden können. Beispielsweise könnte die Sortierung einer Spalte auch durch eine Tastaturkombination veranlasst werden.

Von diesem Modell ausgehend kann die zu erwartende Effizienz eines Systems gemessen werden. Hierzu werden für alle benötigten Operatoren Zeiten festgelegt, welche sich auf empirische Erhebungen stützen. So kann für alle Methoden, Ziele und Selektionsregeln eine erwartete Zeit errechnet werden. Durch geübte Benutzer wird dann eine Analyse der Effizienz des zu untersuchenden Systems anhand von Zeitmessungen durchgeführt. Dadurch lassen sich quantitative Aussagen darüber treffen, um wie viel besser eine Variante im Vergleich zu einer anderen Variante ist. Im Vergleich zu anderen Evaluationsverfahren ist die Aussagekraft allerdings eher begrenzt, zudem die Vorbereitung und Durchführung dieser Art von Tests sehr aufwendig ist [39].

2.2.2 Heuristische Evaluierung

Bei der heuristischen Evaluierung wird die Usability eines Systems von Experten methodisch analysiert, um Schwachstellen in der Benutzungsschnittstelle aufzuzeigen. Experten auf dem Gebiet der Usability sind gegenüber Endnutzern eher in der Lage unvollkommene Systeme einzuschätzen [39]. Nach Nielsen und Molich [36] wird dabei der größte Teil der Probleme mit der Benutzungsschnittstelle (bis zu 95%), schon bei einer geringen Anzahl von drei bis fünf Evaluatoren gefunden. Die Untersuchungen werden dabei anhand von festgelegten Gestaltungsprinzipien, den Heuristiken, durchgeführt. Im Gegensatz zu vollständigen Software Guidelines handelt es sich bei

2 Grundlagen

diesen Prinzipien um eine überschaubare Anzahl von Anhaltspunkten. Beispielsweise haben Nielsen und Molich 1990 [34] neun wichtige Kriterien herausgearbeitet, welche bis heute immer wieder erweitert oder angepasst wurden [35]. Zu diesen Kriterien zählen u. a. Konsistenz über das gesamte System, z. B. bei Beschriftungen, und ständiges Feedback über den Systemstatus für den Benutzer.

2.2.3 Empirische Evaluierung

Bei einer empirischen Evaluierung wird ein Systemtest durch eine repräsentative Gruppe von Endnutzern durchgeführt. Oft wird solch eine Studie erst in den späten Phasen der Entwicklung bzw. kurz vor der Veröffentlichung des Softwareproduktes durchgeführt.

Die wichtigsten Kriterien für die Vorbereitung und Durchführung einer empirischen Evaluierung sind [39]:

Die Testpersonen Die Probanden für die Evaluierung sollten repräsentativ der späteren Nutzergruppe entsprechen. Geht es darum Probleme der Usability aufzudecken, reicht eine geringere Anzahl von bis zu zehn Personen, ähnlich der heuristischen Evaluierung, aus um einen Großteil der Probleme zu finden. Sollen allerdings mehrere Systeme verglichen und eine Empfehlung für ein bestimmtes System gegeben werden, sollten möglichst viele Testpersonen gewonnen werden, um die statistische Stichprobe zu stabilisieren.

Das Testsetting Die Umgebung, in der ein Test durchgeführt wird bringt verschiedene Vor- und Nachteile mit sich. Am Arbeitsplatz des Anwenders entsteht ein reales Setting, hier können aber Störfaktoren von den Evaluatoren nicht gut kontrolliert werden. Die Alternative besteht in speziell für solche Zwecke eingerichteten Usability Laboren. Hier können die Probanden genau beobachtet und der Test zur späteren Analyse aufgezeichnet werden. So ist es im Gegensatz zu anonymen Studien möglich, Probleme und Fehler zu bemerken, welche der Nutzer nicht als solche identifiziert hat. Der Nachteil besteht im geringeren Realismus des Testes, weshalb die Ergebnisse schwer zu übertragen sind. Über Techniken wie Videokonferenzen ist es allerdings möglich, ein Usability Lab über ein Netzwerk oder das WWW auch am realistischeren Arbeitsplatz aufzubauen [20].

Testaufgaben Den Testpersonen werden Testaufgaben gestellt, die diese mit dem zu bewertenden System bearbeiten.

Messwerte & Fragebögen Bei der Bearbeitung der Aufgaben werden objektive, zur statistischen Auswertung vorgesehene Kriterien, automatisch durch die Protokollierung der Nutzeraktionen aufgezeichnet. Die am häufigsten gemessenen Merkmale sind die Zeit zur Erledigung der Aufgabe, die Korrektheit der Bearbeitung, die Fehlerrate aber auch die Erlernbarkeit von Aufgaben. Subjektive

Kriterien wie die Zufriedenheit der Nutzer werden durch Fragebögen oder Interviews aufgenommen.

Statistische und interpretative Auswertung Die automatisch gemessenen Werte und die Ergebnisse aus den Umfragen können beispielsweise über eine Varianzanalyse, einem χ^2 -Test oder auch t-Tests statistisch ausgewertet werden. Welcher Test angewendet werden sollte, hängt zum einen von der Werteverteilung der gemessenen Daten und zum Anderen von der Fragestellung ab, die mit dem Test beantwortet werden soll. Die Aufzeichnungen der subjektiven Meinungen zu verschiedenen Aspekten des Systems werden zusätzlich zur statistischen Auswertung, unter Beachtung der persönlichen Daten der Testpersonen, interpretiert und zur Relativierung oder Bestätigung der statistischen Ergebnisse genutzt.

Eindeutige Ergebnisse sichern jedoch nicht den Nutzen einer durchgeführten Studie. Wichtige Kriterien zur Einschätzung der Verwendbarkeit der Ergebnisse einer empirischen Studie sind nach [39]:

Glaubwürdigkeit Eine Studie gilt als glaubwürdig, wenn die gestellten Aufgaben einen Bezug zur Fragestellung der Studie haben und auf einem realistischen Nutzungsszenario aufbauen. Die Testpersonen müssen repräsentativ zur späteren Nutzergruppe gewählt werden und in wichtigen Kriterien, wie Vorkenntnissen, mit ihr übereinstimmen. Bei einem Vergleich mehrerer Systeme können eventuell Lerneffekte auftreten, wenn alle Systeme von allen Testpersonen (Within Subjects-Design [29]) bewertet wurden. Dem kann mit einem Aufteilen der Testpersonen in Gruppen (Between Subjects-Design [29]), von denen jeweils ein System getestet wird, oder aber mit einer zufälligen Bearbeitungsreihenfolge entgegengewirkt werden.

Übertragbarkeit ist bei einer empirischen Erhebung gegeben, wenn die Ergebnisse der Studie gut auf andere Testumgebungen angewendet werden können. Für die Übertragbarkeit ist auch der realistische Hintergrund einer Testaufgabe ein wichtiger Aspekt.

Zuverlässigkeit bedeutet für eine empirische Studie die Stabilität der gewonnenen Erkenntnisse. Wenn also angenommen werden kann, dass die Ergebnisse sich bei anderen Versuchsparametern, wie der Testlänge oder der Anzahl der Testpersonen, nicht wesentlich verändern würden.

2.3 Verwandte Arbeiten

In diesem Abschnitt wird auf verwandte Arbeiten zum Thema Evaluierungen im Feld der Informationsvisualisierung eingegangen. Neben verbreiteten Evaluierungsmethoden, wie der empirischen Evaluierung, sind in der Literatur zu den in Abschnitt 2.1

2 Grundlagen

besprochenen Techniken auch immer wieder Fallstudien zu finden, welche zumeist beispielhafte Anwendungen einer Technik oder eines Systems anhand eines konkreten Problems darstellen. Im folgenden Abschnitt wird ein Überblick über diese Arbeiten gegeben. Da die im Rahmen der vorliegenden Arbeit durchgeführte Evaluierung als vergleichende Studie geplant ist, wird in einem eigenen Abschnitt insbesondere auch auf bereits durchgeführte Studien zu den beschriebenen Techniken eingegangen. Dabei wird dargelegt welche Technik in welchem Zusammenhang mit welchen Daten getestet wurde. Die Ergebnisse der Studien werden zusammengefasst und das Umfeld der Testpersonen beschrieben.

2.3.1 Fallstudien

Zu beinahe jeder Visualisierungstechnik finden sich in den zu ihnen gehörenden wissenschaftlichen Veröffentlichungen auch Fallstudien, an denen die einzelnen Fähigkeiten der Technik demonstriert werden oder gezeigt wird, dass diese Technik für eine bestimmte Art von Aufgabe besonders geeignet sind.

2.3.1.1 TableLens

Bei der Einführung der TABLELENS-Technik führen Rao et al. [40] eine Analyse in einer Statistik über Baseballspieler und deren Leistung, z. B. ihre Trefferquote und Homerunanzahl, durch. Es wird ausgeführt die Technik sei in einem halben Dutzend weiterer Datendomänen getestet worden und mit einigen wenigen Sortieroperationen sei es in allen Fällen schnell möglich gewesen, interessante Zusammenhänge in den Daten zu finden. Es wird darauf aufmerksam gemacht, dass ein wesentlicher Vorteil der Technik im Gegensatz zu anderen Visualisierungspaketen, wie z. B. der statistischen Analysesoftware SPLUS [42], im geringen Lernaufwand liegt, da die Entwickler die grafische Repräsentation in ein allgemeines Format, die Tabellenform, integriert haben. Dies wird dadurch verdeutlicht, dass die Versuchspersonen, in diesem Fall Baseballfans, bei den Demonstrationen die eigentliche Präsentation schnell vergaßen und eigenständig Zusammenhänge in den Daten aufdeckten. Daraus lässt sich auf eine erhöhte Akzeptanz dieser Technik auch bei Anwendern außerhalb des Feldes der Datenanalyse schließen. Auf die Notwendigkeit formeller Analysen und weiterer Evaluierungen wurde trotzdem hingewiesen.

2.3.1.2 InfoZoom

Verschiedene Fallstudien für die INFOZOOM-Technik zeigen deren Nutzen in der Datenanalyse und -exploration in verschiedenen Datenkategorien, so z. B. im Finanzsektor [45], der Sportstatistik [46] und auch in medizinischen Daten [44]. In der zuletzt aufgeführten Studie wurden Korrelationen zwischen dem Auftreten von Thrombose

und bestimmten Blutwerten nachgewiesen. Im Hinblick auf eine mögliche Anwendung dieser Technik im Bereich der Medizin durch die vorliegende Arbeit, ist es wichtig hervorzuheben, dass die Analyse der Daten von Medizinern selbstständig durchgeführt wurde. In einer weiteren Studie wird anhand von biologischen Klassifizierungsdaten gezeigt, dass der INFOZOOM auch für die Darstellung stark hierarchisch gegliederten Daten geeignet ist, da es eine eindeutige Abbildung der Baumblätter auf die Spalten und der Pfade zu den Blättern auf die Zeilen der Tabelle gibt [47].

In zwei weiteren Arbeiten von González und Kobsa [18], [17] wurde in einer Langzeitstudie das INFOZOOM-System untersucht. Die Studie wurde mit fünf spezialisierten Datenanalysten aus den Bereichen Finanzen, Forschung und Verwaltung durchgeführt. Es wurde mit allen Probanden ein ausführliches Interview zu ihrem Hintergrund geführt bevor sie anhand von Beispieldaten 90 Minuten in das System eingeführt wurden. Danach sollten die Testpersonen INFOZOOM sechs Wochen lang benutzen und in ihren bisherigen Arbeitsplatz integrieren. Jede Woche wurden sie kurz interviewt und konnten zu diesem Zeitpunkt weitere Hilfestellung erhalten. Es hat sich gezeigt, dass mit INFOZOOM schnell und flexibel neue Datensätze exploriert werden können. Für wiederkehrende Aufgaben und Anfragen in ähnlichen Datensätzen, z. B. für die gleiche Erhebung von Daten in zwei aufeinanderfolgenden Jahren stellte sich das Programm aber als unzureichend heraus. Auch die Integration mit anderen Datenanalysetools war nicht gut möglich, es fehlten Import- und Exportschnittstellen, um mit mehreren Analyseprogrammen synchron auf demselben Datensatz zu arbeiten.

Es bleibt herauszufinden, ob der INFOZOOM auch zur Suche bestimmter Datensätze in kleineren Datenbanken geeignet ist oder ob hier die Verwirrung durch die anfangs ungewöhnliche Darstellung überwiegt. Weiterhin ist zu untersuchen ob seine Akzeptanz bei potentiellen Nutzern groß genug ist, beispielsweise in einer Fallauswahl angewendet zu werden.

2.3.1.3 Parallele Koordinaten

Es zeigt sich in der Arbeit von Inselberg [23], dass PARALLELE KOORDINATEN geeignet sind um auf visuelle Art und Weise sehr große Datenmengen zu explorieren und dabei wichtige Zusammenhänge ausfindig zu machen. In der Arbeit wird die Technik theoretisch vorgestellt und es werden zwei beispielhafte Datenanalysen durchgeführt. Zum einen wurden aus den Produktionsdaten eines Mikrochips Zusammenhänge zur Optimierung seiner Herstellung herausgearbeitet. Zum Anderen wird ein Modell einer Volkswirtschaft mit Daten zu verschiedenen Wirtschaftssektoren nachgestellt und gezeigt, dass PARALLELE KOORDINATEN dazu verwendet werden können, um Trade off- und Sensitivitätsanalysen auf Daten interaktiv durchzuführen und somit eine

2 Grundlagen

Entscheidungshilfe darstellen. Da in dieser Arbeit zwei Fallstudien durchgeführt werden, fehlen Vergleiche zu anderen Visualisierungstechniken. Dadurch wurden auch keine Aussagen zur Performanz der Technik gemacht.

E. Wegmann hat sich in verschiedenen Arbeiten mit den PARALLELEN KOORDINATEN und möglichen Erweiterungen beschäftigt. Die Interpretation von Daten in der Darstellung PARALLELER KOORDINATEN anhand der zugrunde liegenden geometrischen Zusammenhänge wird in [51] exemplarisch anhand von Automobil Daten, u.a. Länge, Gewicht und Verbrauch, durchgeführt. Es wird hier auch darauf hingewiesen, dass eine Darstellung mit Hilfe von PARALLELEN KOORDINATEN in einer Gesamtansicht von nur 86 Datensätzen schon sehr überzeichnet wirkt, weshalb verschiedene Erweiterungen, wie Dichteplots, Sättigungsplots und Farbhistogramme (siehe Abschnitt 2.1.4), vorgeschlagen werden. Mit der beispielhaften Anwendung der Sättigungsplots und Achsenpermutationen anhand eines künstlich erstellten Datensatzes in dem das Wort EUREKA in 99 von 3848 Daten kodiert wurde, zeigen Wegmann und Luo [50] wie man in PARALLELEN KOORDINATEN Datencluster aufspüren kann.

In der Arbeit von Siirtola [43] werden zusätzliche Darstellungen und Interaktionen vorgestellt. Die Durchschnittslinien von zu einem Cluster gehörenden Datensätzen und der Dimensionszoom, eine Neuskalierung der Achsen um einzelne Abschnitte hervorzuheben, werden in einer Beispielanwendung auf Automobil Daten wie Verbrauch, Motorengröße und Herkunft eingeführt (Abbildung 2.11 (b)).

Auch in den Fallstudien zu verschiedenen Implementierungen werden Automobil Daten genutzt [5, 13]. Beide Implementierungen werden auch durch unterstützende Visualisierungen wie den Scatterplot begleitet. Ericson et al. [13] verwenden in ihrer Studie neben einem synthetisch vergrößerten Autodatensatz auch Luftverschmutzungsdaten und eine Verkaufsbestandsdatenbank. Es werden Beispielzusammenhänge in den Daten herausgearbeitet, die verdeutlichen, dass eine zusätzliche Darstellung der Veränderung statistischer Daten, wie in Abbildung 2.11 (a) dargestellt, durch Interaktionen das Verständnis der präsentierten Daten erleichtern und intensivieren kann. So kann z. B. verfolgt werden um wie viel sich der Medianwert durch ein Brushing verschiebt.

Zusammenfassend ist zu sagen, dass die PARALLELEN KOORDINATEN bisher noch nicht vergleichend präsentiert wurden. Fallstudien zeigen, dass die Visualisierung sich für die interaktive Datenanalyse und -exploration in verschiedensten Datendomänen eignet, ohne Aussagen über die Performanz oder auftretende Probleme zu treffen. Im Hinblick auf die vorliegende Arbeit ist zu bemerken, dass die PARALLELEN KOORDINATEN bisher nicht außerhalb der Datenanalyse eingesetzt wurden und somit keine Fakten zur Akzeptanz der Technik durch Nicht-Datenanalysten vorliegen.

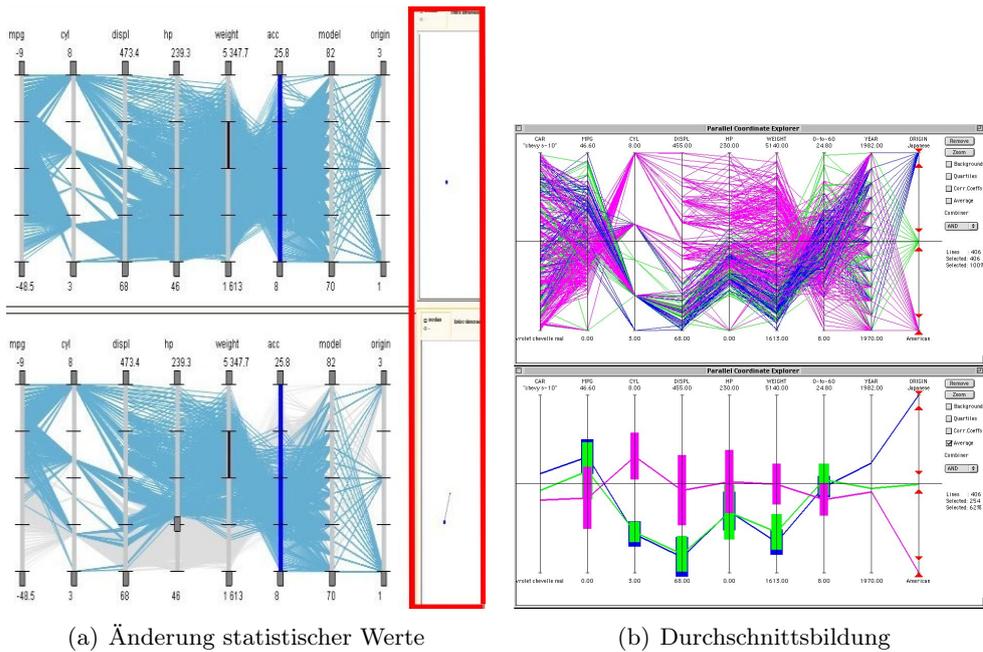


Abbildung 2.11: (a) zeigt die Darstellung der Änderung von statistischen Werten in PARALLELEN KOORDINATEN durch ein Brushing (modifiziert nach [13]). (b) zeigt die Durchschnittsbildung von geclusterten Daten (modifiziert nach [43]).

2.3.2 Vergleichende Studien

In der Literatur finden sich neben Fallstudien auch vergleichende Studien zu den verschiedensten Visualisierungstechniken. Über diese wird im Folgenden ein Überblick gegeben. Da auch die mit dieser Arbeit durchzuführende Evaluierung eine vergleichende Studie sein wird, wird auch jeweils eine Wertung im Hinblick auf die zu entwerfende Studie gegeben werden.

In ihrer Arbeit untersuchen Brunson et al. [10] verschiedene Ansätze zur Informationsvisualisierung an zwei verschiedenen Datensätzen. Zum einen werden sechs sozioökonomische Faktoren einer Population verwendet, als zweites ein synthetisch erstellter Datensatz, welcher benutzt wird um die Qualität der Ansätze im Finden von Ausreißern zu untersuchen. Der Datensatz beinhaltet sechs Variablen und einen Ausreißer dessen Besonderheit erst im Zusammenspiel aller sechs Dimensionen auffällt.

Drei Ansätze wurden auf beiden Datensätzen getestet. Dabei handelte es sich um PROJEKTION PURSUIT, RADVIS und PARALLELE KOORDINATEN. Es wurde gezeigt, dass die PARALLELEN KOORDINATEN von den getesteten die intuitivste Technik ist und die einzige, mit welcher der Ausreißer hoher Ordnung ausfindig gemacht werden

2 Grundlagen

konnte. Dies war allerdings auch erst möglich nachdem eine Untermenge der Daten bestimmt wurde, da der Plot der gesamten Datenmenge in PARALLELEN KOORDINATEN schnell unübersichtlich werden kann. Weiterhin ist zu bemerken, dass die Arbeit vergleichsweise alt ist und somit keinen Vergleich der PARALLELEN KOORDINATEN mit moderneren Visualisierungsansätzen zulässt.

Pirolli und Rao haben das Konzept der TABLELENS-Technik in einer formellen Analyse mit der GOMS Methode dem SPLUS-System gegenübergestellt [37]. Dafür wurden zwei repräsentative Hauptaufgaben der explorativen Datenanalyse, das initiale Verschaffen eines Überblickes des gesamten Datensatzes und im zweiten Schritt das Finden interessanter Korrelationen in den Daten, herausgegriffen. Ihre Einzelschritte wurden aufgeführt und die Zeitkosten mit Hilfe von empirischen Werten entsprechend des GOMS-Modells abgeschätzt und danach empirisch gemessen. Es wurde herausgearbeitet, dass TABLELENS dem Vergleich zu SPLUS mit den genannten Aufgaben standhält, einige Verfeinerungen vorgenommen werden könnten und die Technik eine sehr viel kürzere Einarbeitungszeit hat. Es bleibt zu bemerken, dass die Autoren die Analyse im Selbsttest durchgeführt haben und zumindest einer der Erfinder der TABLELENS-Technik beteiligt war, was auf eine Verfälschung der Zeitmessungen zugunsten der Technik schließen lassen könnte. Außerdem wurden für die Analyse ausschließlich numerische Variablen verwendet, da man sich bei den durchzuführenden Aufgaben auf das Finden von statistischen Kennzahlen wie Median- und Extremwerte beschränkt hat. Es fehlt demnach eine Untersuchung über die Verwendung nominaler und ordinaler Daten.

Alfred Kobsa vergleicht in seiner Studie [26] drei kommerziell vertriebene Visualisierungssysteme, INFOZOOM, Eureka (TABLELENS) und SPOTFIRE, ein Visualisierungssystem welches eine dynamische Anfragesprache mit mehreren Visualisierungsmethoden verbindet, darunter auch die PARALLELEN KOORDINATEN. Sein Ziel war es, herausfinden, ob es signifikante Unterschiede im Hinblick auf die Korrektheit und die Bearbeitungszeit von Aufgaben in den drei Systemen gibt. Das als Benutzerstudie gestaltete Experiment wurde mit drei verschiedenen Datensätzen durchgeführt, darunter Daten einer Partnervermittlung, Aufzeichnungen über Schwermetallvorkommen in Schweden und technische Daten von Oldtimern. Es wurden 81 Studenten als Versuchspersonen ausgewählt und auf die drei Systeme verteilt. Sie wurden mit Übungsaufgaben eine Stunde lang in das jeweilige Programm eingearbeitet und sollten danach 36 Aufgaben lösen. Diese waren komplexer Natur, involvierten immer mehrere Variablen und teilweise sollten Korrelationen herausgefunden werden. So wurde z.B in der Domäne der Partnervermittlung nach dem Verhältnis zwischen Katholiken und Menschen evangelischen Glaubens gefragt, denen Sex in einer festen Beziehung wichtig ist. Zur Erfüllung einer solchen Aufgabe wird eine Form der grafischen Repräsentation beinahe zwingend notwendig, da hier die Anzahl der zutreffenden Datensätze geschätzt werden muss.

Es werden in der Arbeit von Kobsa mehrere Faktoren herausgearbeitet, die für die Umsetzung eines Visualisierungssystems im Hinblick auf die Usability wichtig sein

können.

1. *Die Anzahl der zur Verfügung stehenden Visualisierungen und ihre Einordnung im Gesamtsystem.* So ist beispielsweise bei SPOTFIRE der interaktive Scatterplot die Standardvisualisierung, so dass die Testpersonen auch bei Aufgaben bei denen eine andere Visualisierung geeigneter war, erst eine Lösung mit dieser Technik versuchten, was zu Performanzverlusten bei der Aufgabebearbeitung führte.
2. *Die Interaktionsmöglichkeiten auf den Techniken, ihre Anzahl und ihre Intuitivität.* So sind die Brushingwerkzeuge bei SPOTFIRE zumeist Standardslider, die dank ihrer Verbreitung intuitiv zu bedienen sind, während Eureka und INFOZOOM völlig neue Interaktionen einführen. Hierbei muss auf den meist erst später deutlich werdenden Nutzen der Technik geachtet werden, da ein Effizienzgewinn den Lernaufwand für neue Interaktionsformen erhöht.
3. *Die direkte Umsetzung des visuellen Paradigmas.* Eureka und INFOZOOM entsprechen beide einer tabellarischen Sicht mit graphischen Repräsentationen. Bei Eureka jedoch geschieht diese in der Breite einer Spalte, so dass sich das Verteilungsmuster der Daten gut erkennen lässt. Bei INFOZOOM muss hierfür die Höhe einer Zeile genügen.
4. *Die Usability des die Techniken umgebenden Systems.* Gemeint ist die begleitende Benutzungsoberfläche oder auch Möglichkeiten des Datenimports und -exports.

Es wird darauf hingewiesen, dass eine Mischung komplexer und einfacher Aufgaben in einer Evaluierung angebracht sind, um die Evaluierungsergebnisse sinnvoll auswerten und interpretieren zu können, da bei komplexen Aufgaben allein die Fehlerursachen zu vielfältig sein können. Obwohl alle drei Systeme die Möglichkeiten bieten die gestellten Aufgaben adäquat zu lösen, stellte sich heraus, dass der INFOZOOM mit wenigen Interaktionen einen sehr guten Überblick über die Daten ermöglicht. Bei der TABLENS stellten sich immer dann Probleme ein, wenn zu viele Variablen involviert waren, so dass deren Beschriftung im Tabellenkopf verdeckt war.

Im Hinblick auf die durchzuführende Evaluierung ist zu sagen, dass in der Studie keine medizinischen Daten verwendet wurden und obwohl alle der untersuchten Systeme eine Tabellensicht in natürlicher oder abgewandelter Form darstellen oder beinhalten, die reine Tabellensicht nicht als eine eigenständige Visualisierungstechnik untersucht wurde.

Eine weitere Studie von Mark, Carpenter und Kobsa [31] vergleicht INFOZOOM mit SPOTFIRE, um herauszufinden, ob Gruppen besser als Individuen mit den Visualisierungsprogrammen arbeiten. Als Versuchspersonen dienten 100 Studenten und es wurden wieder Daten eines Partnervermittlers genutzt. Es wurde gezeigt, dass Gruppen eine höhere Anzahl korrekter Antworten geben aber langsamer arbeiten. Weiterhin

stellte sich heraus, dass INFOZOOM-Benutzer besser abschneiden. Ein Grund hierfür könnte die überladene Oberfläche von SPOTFIRE sein. Sie bietet mehrere Visualisierungen und mehrere UI-Elemente zum Eingrenzen der beobachteten Datensets an, während INFOZOOM drei Sichten für verschiedene Aufgaben bietet und über die Zoominteraktion immer der gesamte Platz ausgenutzt wird.

2.4 Kapitelzusammenfassung

Zusammenfassend ist zu bemerken, dass alle vorgestellten Techniken bereits im Bereich der Datenexploration und -analyse eingesetzt wurden. Fallstudien zeigen, wie gut die einzelnen Techniken verwendet werden können, um Eigenheiten und Zusammenhänge auch in extrem großen Datenmengen zu finden. Einige wenige vergleichende Evaluierungen haben schon stattgefunden, allerdings nicht im medizinischen Kontext. Auch wurde die Eignung der Techniken abseits der reinen Datenanalyse, beispielsweise auf kleinen Datenbanken oder zur Präsentation einer Auswahlmenge, noch nicht untersucht. Mit der vorliegenden Arbeit wird untersucht, inwieweit die erklärten Visualisierungen auf einen solchen Kontext adaptiert werden können oder ob sie eventuell einen zu komplexen Ansatz verfolgen. Sind die Techniken einsetzbar, bleibt zu untersuchen, ob komplexe Techniken, wie PARALLELE KOORDINATEN, im Gegensatz zu einfachen Techniken, wie der gewöhnlichen Tabelle, in einer Fallauswahlsituation einen deutlichen Vorteil bringen. Zwar haben einige Untersuchungen gezeigt, dass die verschiedenen Informationsvisualisierungen im Allgemeinen als „gut“ empfunden werden. Dennoch finden sie im alltäglichen Gebrauch von Computerprogrammen bisher kaum Anwendung. Aus diesem Grund wird zusätzlich untersucht, inwieweit potentielle Nutzer, im Fall der vorliegenden Arbeit sind dies Mediziner, die Techniken akzeptieren würden.

3 Entwurf

Wie im vorangehenden Kapitel herausgestellt, wurden die besprochenen Visualisierungen noch nicht abseits der Datenanalyse eingesetzt. Aus diesem Grund gibt es noch keine Untersuchungen dazu, wie sie sich auf kleineren Datenbanken anwenden lassen, ob sie auch dort von erheblichen Nutzen sein können, und ob sie von den anderen Endnutzern akzeptiert werden. Es muss untersucht werden inwieweit komplexere Visualisierungsansätze bei der Fallauswahl gegenüber einfachen Techniken, wie den Tabellen, von Vorteil sind.

Aus diesen Gründen soll nun in diesem Kapitel der Arbeit herausgearbeitet werden, welche Techniken aus dem Bereich der Informationsvisualisierung geeignet scheinen, um für die Fallauswahl in einem medizinischen fallbasierten Lernsystem eingesetzt zu werden. Des Weiteren wird gezeigt, wie diese Techniken hinsichtlich dieses Einsatzgebietes effektiv evaluiert werden können. Zu diesem Zweck werden typische Eigenschaften einer schnellen und intuitiven Fallauswahl in einem medizinischen Trainingssystem charakterisiert und daraus die Anforderungen an eine Visualisierungstechnik herausgearbeitet.

Anhand dieser Anforderungen wird diskutiert, welche Techniken hinsichtlich ihres Visualisierungsparadigmas und ihrer Interaktionsmöglichkeiten für eine Fallauswahl geeignet scheinen und welche Rahmenbedingungen sich für eine vergleichende Studie aus den Eigenschaften der Fallauswahl ergeben.

Im weiteren Verlauf des Kapitels werden aus den gewählten Techniken und den Eigenschaften der durchzuführenden Evaluierung Thesen über die Benutzung der aufgeführten Techniken im Bereich der medizinischen Lernsysteme aufgestellt, welche später anhand der Ergebnisse aus der Evaluierung diskutiert werden.

Am Ende des Kapitels wird die Konzeption der Durchführung der Evaluierung stehen. Diese wurde mit Hilfe eines alleinstehenden Programmes durchgeführt, dessen inhaltlicher Aufbau im letzten Abschnitt erläutert werden soll.

3.1 Anforderungen an eine effiziente Fallauswahl

Um in einem fallbasierten Trainingssystem eine geeignete Visualisierungstechnik für die Fallauswahl einsetzen zu können, muss zuerst herausgearbeitet werden, wodurch

3 Entwurf

eine geeignete bzw. effiziente Fallauswahl charakterisiert ist, um somit die Anforderungen an die zu verwendene Technik darzustellen und festzulegen, unter welchen Rahmenbedingungen eine Evaluierung stattfinden sollte.

3.1.1 Die Falldaten

Die Daten in der Fallauswahl eines Trainingssystems lassen sich unabhängig vom Einsatzgebiet des Systems grundsätzlich in drei Bereiche unterteilen. Dazu zählen Daten zur Identifikation eines Falles bzw. eines Patienten, wie z. B. ein Foto oder der Name. Die zweite Kategorie beinhaltet die Metadaten des Trainingsprogrammes, z. B. die wahrscheinliche Bearbeitungszeit eines Falles, seinen Schwierigkeitsgrad, Bearbeitungsstand oder auch seine Ausstattung mit Expertenkommentaren [33]. In die dritte Kategorie zählen domänenspezifische Daten des Trainingsfalles. Sie lässt sich in der Medizin in verschiedene Anwendungsgebiete teilen, z. B. ob es um die Behandlung im Hals- oder Leberbereich geht. Die dritte Kategorie stellt somit das allgemeine Anwendungsgebiet dar, während ihre Unterteilung den speziellen Anwendungsbereich einer Trainingssoftware abgrenzt. Abbildung 3.1 zeigt eine beispielhafte Kategorisierung von Daten eines Falles in einem Lernsystem.

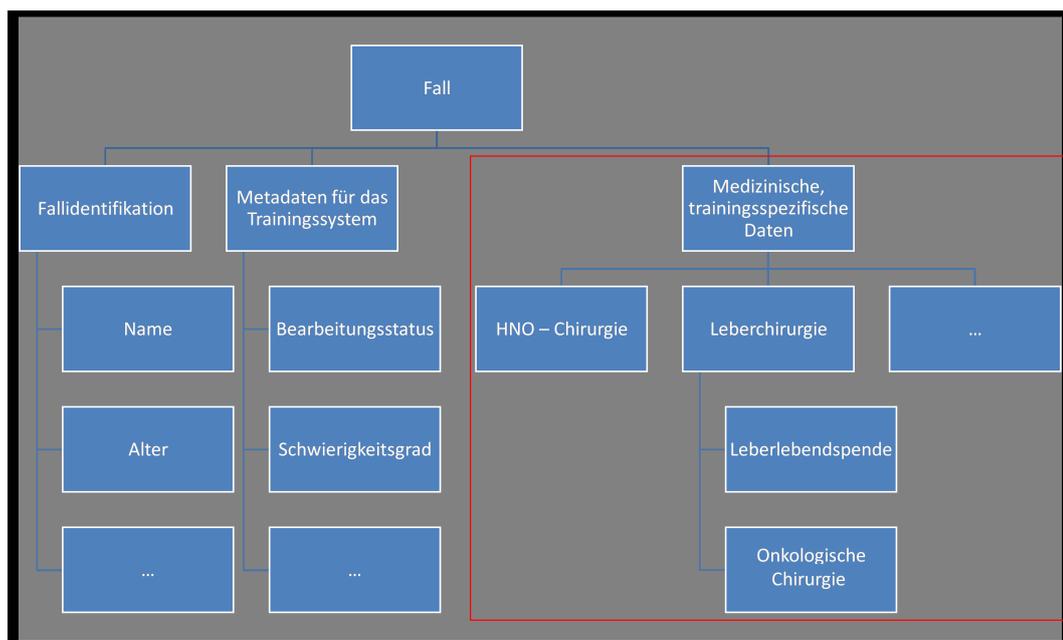


Abbildung 3.1: Beispiel einer Datenkategorisierung für ein fallbasiertes Lernsystem in der Medizin. Der Teil, welcher den krankheitsspezifischen Einsatz des Lernsystems abgrenzt, wird durch den Rahmen hervorgehoben.

Eine einzusetzende Technik unterstützt schlussfolgernd idealerweise eine Gruppier-

rung oder hierarchische Anordnung von Variablen, um diese Einteilung in Datenkategorien möglichst genau abbilden zu können.

3.1.2 Die Fallanzahl

Die Fallanzahl in einem Trainingssystem umfasst initial meist eine kleine, repräsentative Auswahl. Die Datenbanken wachsen jedoch stetig mit der Einsatzzeit und Popularität eines Trainingssystems. Die Erstellung von Fällen ist durch die notwendige Anreicherung mit Expertenwissen und Aufbereitung der multimedialen Inhalte sehr aufwendig [33]. So wird die Falldatenbank trotz Wachstum immer auch eine repräsentative Auswahl bleiben, weshalb in einem System zu einem Fachgebiet mehr als 100 Fälle eher unüblich sind. Dies bedeutet, dass einer eventuell hohen Anzahl von Variablen eine im Gegensatz zur Datenanalyse in anderen Datendomänen eher geringe Anzahl von Fällen gegenübersteht.

Eine geeignete Technik sollte folglich mit hochdimensionalen Daten umgehen können sowie bei einer mittleren Menge an Fällen eine gute Übersicht über den gesamten Datenbestand zulassen.

3.1.3 Die Fallauswahl

Die Fallauswahl als solches wird nicht zufällig durchgeführt. Der Trainierende wählt den zu trainierenden Fall nach ganz bestimmten Vorstellungen aus. Er wird sich bei seiner Suche in der Datenbank an ausgewählte Variablen halten, um so schnell die für seine Situation interessanten Fälle zu finden. Die Suche erfolgt z. B. nach dem zuletzt bearbeitetem Fall, einem Fall der dem zur Verfügung stehenden Zeitfenster entspricht oder einen Fall welcher einem realen Patienten mit dessen Daten am ehesten entspricht.

Einige wahrscheinliche Nutzungsszenarien sind:

Erstnutzung Ein Nutzer welcher zum ersten Mal ein Trainingsprogramm benutzt, wird sich erst einmal einen Überblick über die vorhandenen Fälle schaffen wollen. Eine eingesetzte Technik wird somit eher explorativ genutzt, um zu sehen welches Fallspektrum das System anbietet. Bei der Auswahl eines ersten Trainingsfalles ist anzunehmen, dass der Auswählende sich an diejenigen Metadaten des Systems hält, welche im allgemeinen die gesamte Datenbank grob unterteilen, beispielsweise nach dem Schwierigkeitsgrad oder dem Themengebiet des Falles, bei der Leberchirurgie beispielsweise ob es sich um einen onkologischen Fall oder eine Leberlebendspende handelt.

Facharztausbildung/ Prüfungsvorbereitung Ein fortgeschrittener Nutzer eines Trainingssystems könnte beispielsweise ein Arzt in der Facharztausbildung sein. Dieser hat mehrmals mit dem System gearbeitet, mehrere Arten von Fällen

3 Entwurf

bearbeitet und sich im Hinblick auf eine eher geringe Datenbankgröße mit weniger als 100 Fällen einen Überblick über das angebotene Fallspektrum verschafft. Ein solcher Nutzer würde zumeist nach medizinischen Parametern das Fallspektrum eingrenzen um dann anhand des Bearbeitungsstands diejenigen Fälle suchen die noch nicht trainiert wurden, um daraus den spezifischen Fall auszuwählen, den er trainieren möchte.

Im besonderen Falle einer Prüfungsvorbereitung werden am ehesten die komplizierten Fälle mit medizinischen Besonderheiten herausgesucht werden, um diese der Reihe nach mit dem Ziel einer stetigen Verbesserung (wiederholend) zu trainieren.

Arzt vor einer Operation Ein Arzt, welcher eine Operation durchzuführen hat wird immer diejenigen Fälle trainieren wollen, die den Merkmalen des zu operierenden Patienten am ehesten entsprechen. Er wird sich folglich den Fall herausuchen, der einem aktuellen realen Fall in entscheidenden Attributen gleicht, beispielsweise also die gleiche Diagnose und die gleiche Therapieform aufweist.

Fallautoren Ein fallbasiertes Trainingssystem besitzt zumeist einen Autorenmodus in welchem neue Fälle von Benutzern selbst in das System eingepflegt werden können. Da die Falldatenbank eine repräsentative Auswahl darstellen soll, muss davon ausgegangen werden, dass:

- Sich nicht zu viele zu ähnliche Fälle im System befinden
- Gleichzeitig das gesamte Spektrum der Ausprägungen des im Trainingssystem behandelten Gebietes abgedeckt wird.

Für einen Fallautor ist es folglich erforderlich, sich sich einen Überblick darüber zu verschaffen, welche Art von Fällen im System noch nicht bzw. zu häufig vorhanden sind.

Einsatz in Prüfung/Seminar Eine weitere Einsatzmöglichkeit eines Trainingsprogrammes ergibt sich für Prüfungen oder Seminare. Hierfür wird der Seminarleiter nicht nur einen, sondern eine Reihe von Fällen suchen, die zur jeweiligen Thematik des Seminartermines passen.

Ein Prüfer wird auch mehrere Fälle suchen, die er zur Prüfung bearbeiten lässt. Diese werden aller Wahrscheinlichkeit nach alle Schwierigkeitsgrade abdecken, so dass er z. B. für eine besonders schwierige Aufgabe einen Fall benötigt der aufgrund von besonderen Eigenschaften eine spezielle Herangehensweise benötigt. Dieser Nutzer einer Fallauswahl muss demnach neben der Gruppierung von Variablen auch in der Lage sein, außergewöhnliche Fälle schnell in einer Datenbank zu finden.

Die geeignete Visualisierung bietet folglich idealerweise:

- Die Möglichkeit zur Fokussierung auf die in einer gegebenen Situation interessierenden Untermenge durch das Ausblenden nicht interessierender Datenbereiche

- Einen Gesamtüberblick über die Daten und die Möglichkeit die Datenbank explorativ nach Fällen zu erforschen, welche eventuell Besonderheiten aufweisen

3.1.4 Stellenwert der Fallauswahl

Der Stellenwert der Fallauswahl in einem medizinischen Trainingssystem ist grundsätzlich nicht unerheblich. Wenn der gewählte Fall nicht den gewünschten Parametern entspricht, dies aber erst zu einem späten Zeitpunkt im Ablauf einer Trainingssession klar wird, kann dem Trainierenden viel Zeit verloren gehen und die Motivation das System weiterzubnutzen könnte sinken. Es sei darauf hingewiesen, dass die Fallauswahl trotz der genannten Gründe ein sehr kleiner Teil des Programmes ist und losgelöst vom eigentlichen Training statt findet.

Eine geeignete Visualisierung sollte genaue und zügige Entscheidung herbeiführen und gleichzeitig intuitiv bedienbar sein. Das bedeutet, der Lernaufwand sollte sehr gering sein und der Aufbau der Technik sollte den Benutzer nicht verwirren, so dass die Vorteile der Technik ab der ersten Benutzung zur Geltung kommen ohne das eigentliche Training aufzuhalten.

3.2 Geeignete Techniken für die Fallauswahl

Anhand der herausgearbeiteten Eigenschaften einer Fallauswahl können die Anforderungen an eine für ein medizinisches Lernsystem geeignete Visualisierung folgendermaßen zusammengefasst werden:

- Zulassen einer Gruppierung bzw. Kategorisierung von Daten
- Eignung für den Einsatz in hochdimensionalen Daten bei gleichzeitig eher geringer Fallzahl
- Vermitteln eines Gesamtüberblicks über das Fallspektrum
- Realisierung einer Fokussierung auf interessante Bereiche
- Ausgeprägte Besonderheiten in den Daten sollten erkennbar sein
- Bereitstellung intuitiver Interaktionsmöglichkeiten

Im Folgenden wird diskutiert, welche Techniken im Bezug auf diese Anforderungen als besonders geeignet scheinen, in einer Fallauswahl eingesetzt zu werden.

Alle Visualisierungen haben gemeinsam, dass sie selbsterklärender werden je mehr Zusatzinformationen geboten werden. Um die eigentlichen Visualisierungen zu entlasten, kann, neben klassischen Fokuginformationen wie Tooltips, bei einer Fallauswahl

eine umgebende GUI zur Verfügung gestellt werden. In dieser können beispielsweise Auswahlsets näher untersucht werden und bebilderte Beschreibungen der Fälle eingeblendet werden.

3.2.1 Tabellen

Tabellen haben als Visualisierungstechnik den großen Nachteil, dass sie aufgrund ihrer ineffizienten Platzausnutzung nur einen unzureichenden Gesamtüberblick über die betrachteten Daten geben. Ein weiterer Nachteil ist die bei Tabellen übliche rein textuelle Repräsentation der Daten, welche vom Menschen nur bewusst und mit hoher Aufmerksamkeit verarbeitet werden kann. Vergleichsweise leichter fällt dies bei grafischen Mustern [27]. Dadurch wird ein schnelles Verständnis der Gesamtheit der Daten behindert. Werden grafische Hilfsmittel wie z. B. Fotos als visuelle Stütze zu Namen integriert, geht noch einmal weiterer Platz am Bildschirm verloren, so dass ein einzelner Datensatz nicht mehr im Kontext aller Datensätze zu erforschen ist. Extremwerte in den Daten werden durch die einheitliche Darstellung nicht ohne weitere Interaktionen auffallen. Diesen Nachteilen stehen im Hinblick auf die angeführten Anforderungen aber diverse deutliche Vorteile gegenüber. So ist die Darstellungsform der Tabelle sehr verbreitet und wird in vielen Computerprogrammen im Alltag verwendet, so dass Interaktionen, wie eine Selektion oder eine Sortierung, für nahezu jeden Nutzer intuitiv ausführbar sind. Für die Gruppierung und Fokussierung von Daten stehen einfache Interaktionen wie die Umordnung von Spalten und die Sortierung der Datensätze zur Verfügung. Bei der zu erwartenden Datenbankgröße bei einer medizinischen Fallauswahl kann immer noch ein Großteil der Daten auf einer Bildschirmgröße sichtbar bleiben, so dass nicht zuviel Zeit beim Scrollen verloren geht.

Der tabellarische Visualisierungsansatz scheint nach Abwägung der Vor- und Nachteile gut geeignet und wird deshalb in die Evaluierung mit einbezogen. Er dient zusätzlich als Vergleich zu erweiterten Techniken, um deren Nutzen gegenüber etablierten Techniken messen zu können.

3.2.2 TableLens

Die TABLENS verwendet ebenfalls einen tabellarischen Darstellungsansatz. Hierbei werden Vorteile der Tabellen, wie die Gruppierung von Spalten und das Sortieren um die Fähigkeit komplette Datenbanken auf einem Bildschirm darzustellen erweitert. Erreicht wird dies durch den Verzicht auf durchgängig textuelle Präsentation der Daten und deren Ersatz durch eine integrierte grafische Darstellung, bei numerischen Daten z. B. durch eine Art Balkendiagramm. So ist diese Technik besonders geeignet die Falldatenbank als Ganzes zu untersuchen und eventuelle Besonderheiten aufzudecken. Durch die zusätzliche textuelle Darstellung in einer Fokuslinse ist die erweiterte Untersuchung interessierender Bereiche möglich.

Im Gegensatz zu den Vorteilen des Gesamtüberblickes steht die verlorene textuelle Darstellung, was eine Fokussierung und Untersuchung einzelner Datensätze allein mit dem Auge ohne weitere Interaktionen unmöglich macht. Durch die Einführung neuer Interaktionen durch die Linse geht auch die Intuitivität verloren. Dies kann dazu führen, dass ein Nutzer für die Fallauswahl ohne Hilfe mehr Zeit benötigt, um diese vorteilhaft zu bedienen. Inwieweit eine Präsentation zusätzlicher Informationen in einer begleitenden GUI hilft, diesen Nachteil wett zu machen, wird die Evaluierung zeigen.

3.2.3 InfoZoom

INFOZOOM bedient sich in seiner komprimierten Ansicht ebenfalls einer tabellarischen Grunddarstellung. Verwirren kann einen ungeübten Nutzer hier, dass die Variablen, normalerweise auf die Spalten abgebildet, bei dieser Technik auf die Zeilen abgebildet werden und dass Zellen mit gleichem Inhalt zusammengefasst werden. Dadurch wird ein direktes Zuordnen eines bestimmten Wertes zu einem Fall, also einer Spalte, visuell erschwert. Jedoch wird durch dieses Zusammenfassen dort, wo genügend Zellen zusammengefasst werden auch die textuelle Darstellung der Daten möglich obwohl auch bei dieser Technik die komplette Datenbank auf der Bildfläche zusammengefasst dargestellt wird. Durch eine integrierte grafische Repräsentation wie bei der TABLENS gehen auch hier durch den Verzicht auf textuelle Darstellung nicht alle Informationen verloren. Durch Sortieren und durch die leicht zu bedienende Zoominteraktion kann ein Fokus auf die interessierenden Bereiche gelegt werden. Ein großer Vorteil dabei ist, dass für den eingegrenzten Bereich wiederum die gesamte zur Verfügung stehende Fläche genutzt wird und auf diese Weise mit jedem Zoom mehr Platz für eine textuelle Darstellung zur Verfügung steht.

Da der Nutzer eines Trainingssystems nicht schon bei der Fallauswahl mit mehreren Sichten überfordert werden darf, soll die Implementierung des INFOZOOM nur eine von ursprünglich drei Ansichten umsetzen. Die dekomprimierte Tabelle ist zu unübersichtlich und der Übersichtsmodus ist überwiegend für die Abschätzung der Verteilung der Werte in den Attributen gedacht und wird für die meisten Nutzungsszenarien unbenutzt bleiben, so dass für die Implementierung nur die komprimierte Tabellensicht vorgesehen ist.

3.2.4 Parallele Koordinaten

PARALLELE KOORDINATEN nutzen im Gegensatz zu anderen Techniken einen Ansatz der Projektion. Auch hier wird die gesamte Falldatenbank dargestellt und die Technik erlaubt einen schnellen Überblick über die vorhandenen Daten. Die Technik scheint auch am ehesten dafür geeignet Ausreißer oder nicht abgedeckte Wertebereiche zu entdecken, da diese direkt in einzelnen Falllinien bzw. leeren Achsenbereichen sichtbar werden.

3 Entwurf

Ein großer Nachteil ist, wie für die meisten Visualisierungsmethoden, der geringe Bekanntheitsgrad im Alltag und damit eine wahrscheinlich geringe Akzeptanz der Nutzer. Ein weiteres Problem liegt in der Intuitivität der Technik. So wird ein ungeübter Nutzer im Hinblick auf den komplett anderen Visualisierungsansatz eventuell irritiert sein und auch die Selektion einzelner Fälle in einer überfüllten Ansicht ist ohne einen gewissen Lernaufwand kaum durchführbar. Durch das Eingrenzen der Wertebereiche ist zwar eine Ausfilterung nicht interessierender Bereiche möglich aber die geringe Achsenbeschriftung wird dies dem unerfahrenen Benutzer erschweren. Des Weiteren kann durch die Darstellung der gesamten Daten die Ansicht unübersichtlich werden. Dies gilt sowohl bei einer hohen Anzahl von Variablen als auch bei einer hohen Anzahl von Fällen.

Durch die Zusatzinformationen in der umgebenden Benutzungsschnittstelle und auch in den Tooltips ist es trotz der genannten Nachteile möglich, in den Daten gut zu navigieren und auch die zu erwartende Menge zwischen 30 und höchstens 100 Fällen wird die Technik wahrscheinlich nicht visuell überlasten. Inwieweit PARALLELE KOORDINATEN mit den anderen Techniken in einer Situation wie der Fallauswahl besser geeignet sind, soll in der Evaluierung herausgefunden werden.

3.2.5 Weitere Techniken

Andere genannte Techniken werden nicht weiter verfolgt, da sie wie z. B. RADVIS oder PROJECTION PURSUIT einen zu komplizierten Ansatz verfolgen und durch die bei jeder neuen Darstellung neu zu erlernende Interpretation der Daten als zu wenig intuitiv eingeschätzt werden, um in einer Zielgruppe von Nicht-Datenanalysten präsentiert zu werden. Vollintegrierte Visualisierungssysteme wie SPOTFIRE oder das Statistikprogramm SPLUS sind zu komplex und umfangreich und würden den Rahmen einer Fallauswahl sprengen. Hier müssten einzelne Visualisierungen herausgelöst werden, um sie mit den besprochenen Techniken vergleichen zu können. Der Nutzen dieser Systeme indes liegt im Verbund von mehreren interaktiven Visualisierungen, bei der Interaktionen in einer Technik Auswirkungen auf eine andere haben (Linking & Brushing). Eine Mischung aus mehreren Visualisierungen wäre aber anhand des Stellenwertes einer Fallauswahl in einem Gesamtsystem nicht gerechtfertigt (zu hoher Lernaufwand) und ist somit inakzeptabel.

3.2.6 Thesen zur Benutzung der Visualisierungen

Anhand der ausgewählten Techniken und ihrer Vor- und Nachteile werden nun Thesen zu deren Benutzung und ihres Nutzens bezüglich einer Fallauswahl aufgestellt, welche im weiteren Verlauf der Arbeit mit den Ergebnissen der durchzuführenden Evaluierung aufgegriffen und diskutiert werden sollen.

Folgende Thesen sind zu überprüfen:

Tabellen

1. Tabellen sind ungeeignet einen Gesamtüberblick über das angebotene Fallspektrum zu bekommen und erlauben schwerlich das Erkennen von Ausreißern oder fehlenden Daten.
2. Tabellen benötigen den geringsten Lernaufwand.
3. Für einfache Aufgaben wie das Auswählen von Fällen nach gegebenen Kriterien sind Tabellen mit den Interaktionen Sortieren und Spalten umordnen ausreichend.
4. Für das Auswählen von Fällen nach komplex zusammenhängenden Kriterien bzw. für das Erkennen globaler Zusammenhänge oder Werteverteilungen sind Tabellen nicht ausreichend.

TableLens

5. TABLELENS erlaubt einen schnellen Überblick über die Daten und ermöglicht, schnell Aussagen über Eigenheiten der Daten zu treffen.
6. TABLELENS benötigt einen geringen Lernaufwand, der durch die fehlende textuelle Repräsentation und zusätzliche Interaktionsformen bedingt ist.
7. TABLELENS ist für einfache und komplexe Auswahlsszenarien geeignet.
8. Das Fehlen von Text wirkt sich bei einer geringen Menge von Fällen störend aus.

InfoZoom

9. INFOZOOM ermöglicht mit einem gewissen Aufwand einen Überblick über die Datenbank und das Finden von seltenen Parameterkombinationen.
10. Durch die Darstellung im Zellverbund lassen sich in Kombination mit den möglichen Interaktionen Ausreißer oder umgekehrt die Menge sehr ähnlicher Fälle gut mit der INFOZOOM-Visualisierung bestimmen.
11. INFOZOOM benötigt durch die befremdliche Anordnung in einer Tabelle und das Kombinieren von Zellen einen höheren Lernaufwand, lässt sich dann aber ohne Probleme bedienen.
12. Durch die Zoomfunktion ist INFOZOOM für die Suche nach einzelnen Fällen sehr gut geeignet. Das übergreifende Suchen in verschiedenen Clustern wird durch das Ausblenden von Fällen beim Zoom erschwert, da Zoomoperationen für den Clusterwechsel rückgängig gemacht werden müssen.

Parallele Koordinaten

13. PARALLELE KOORDINATEN benötigen durch den weniger populären Visualisierungsansatz den höchsten Lernaufwand
14. Auswahlaufgaben sind bei PARALLELEN KOORDINATEN mehr als bei anderen Techniken von der Anzahl der Variablen und der Fälle abhängig. Für die Suche einzelner Fälle in beliebiger Komplexität sind die parallelen Koordinaten sehr gut geeignet. Das übergreifende Suchen in verschiedenen Clustern erfordert das Rückgängigmachen von einzelnen Brushingoperationen für den Clusterwechsel und ist somit nur erschwert möglich.
15. PARALLELE KOORDINATEN eignen sich am besten für den Datenüberblick und das Finden von außergewöhnlichen Daten.

3.3 Rahmenbedingungen für eine Evaluierung

Durch die Charakterisierung einer Fallauswahl in medizinischen Lernsystemen, der Auswahl geeigneter Techniken und das Aufstellen von Thesen zur Benutzung derselben, lassen sich nun die Rahmenbedingungen für eine Benutzerstudie aufstellen. Dazu werden die zu verwendenden Testdaten, die zu befragenden Testpersonen und zu lösende Testaufgaben zusammengetragen.

3.3.1 Falldaten

Für Evaluierungen ist es wichtig, realistische Anwendungsszenarien und real zu erwartende Daten zu präsentieren, um für Nutzer aus der Anwendungsdomäne überzeugend zu sein [38]. Um diesem Anspruch zu genügen, müssen die Daten für die Evaluierung eine realistische medizinische Falldatenbank eines Trainingssystems möglichst genau abbilden. So werden Variablen aus allen in 3.1.1 beschriebenen Kategorien zu finden sein. Idealerweise werden für die spezifischen Krankheitsattribute bei verschiedenen Anwendungsaufgaben Datensätze aus unterschiedlichen Anwendungsgebieten benutzt, um die Anwendbarkeit der Techniken in verschiedenen Datendomänen zu testen bzw. sicherzustellen.

Es werden 30 bis 80 Fälle in einem Fallspektrum für eine Aufgabenstellung vorhanden sein, welche sich repräsentativ über die Wertebereiche der einzelnen Fallattribute verteilen. Ausgehend von einigen realen Datensätzen können künstlich entworfene Fälle hinzukommen um die Datenbank aufzufüllen. Diese können so gestaltet werden, dass Ausreißer in die Daten eingeführt werden oder einzelne Wertebereiche nicht abgedeckt werden.

3.3.2 Testpersonen

Die Zielgruppe eines medizinischen Trainingsprogrammes sind Ärzte und Medizinstudenten. Aus diesem Grund sind auch Menschen aus dieser Zielgruppe als Testpersonen für eine Evaluierung einzusetzen. Fallbasierte Trainingsprogramme gibt es allerdings in vielen Bereichen, z. B. auch im Maschinenbau. Eine zu testende Fallauswahl kann folglich auch mit Personen außerhalb der Medizin evaluiert werden. Da im Rahmen dieser Arbeit speziell die Fallauswahl in medizinischen Falldatenbanken behandelt wird, ist es wünschenswert, hier auch Teilnehmer aus der medizinischen Zielgruppe für eine Evaluierung zu gewinnen.

3.3.3 Testaufgaben für die Techniken

Anhand der Anwendungsszenarien lassen sich verschiedene Kategorien von Testaufgaben für eine Evaluierung ableiten. Diese sollen in der Evaluierung von den Testpersonen mit den unterschiedlichen Visualisierungsmethoden abgearbeitet werden. In Abhängigkeit von den zu bearbeitenden Aufgaben werden verschiedene Eigenschaften der Visualisierung gefordert und der Einsatz verschiedener möglicher Interaktionen nötig gemacht. Entsprechend dem Interaktionsaufwand können die Aufgaben mit steigendem Schwierigkeitsgrad in die folgenden fünf Kategorien eingeteilt werden:

1. *Einfaches Finden und Auswählen*

In dieser Kategorie ist ein einfaches Finden eines einzelnen Falles anhand von ein bis drei Variablen gefordert. Dies erfordert gewöhnlich eine oder mehrere Sotieroperationen und eventuell ein Ordnen bzw. Zoomen.

2. *Einfaches Finden und Auswählen mehrerer Fälle*

Wie (1), jedoch sollen mehrere Fälle ausgewählt werden.

3. *Komplexes Finden und Auswählen*

In eine solche Auswahl sind viele Attribute involviert, sie erfordert ein Zusammenspiel aller möglichen Interaktionen und teilweise müssen Fälle in Beziehung zueinander gestellt werden. Hier zeigt sich, wie gut die im Konzept der Technik integrierten Interaktionen bei gleichzeitiger Erhaltung der Kontextinformationen funktionieren. Hier wird folglich das Gesamtpotential der Technik getestet.

4. *Komplexes Finden und Auswählen mehrerer Fälle*

Wie (3), jedoch sollen mehrere Fälle ausgewählt werden.

5. *Überblick/Eigenheiten der Daten*

Diese Aufgabenkategorie zielt auf das Erkennen von Zusammenhängen zwischen Variablen ab und sie hat zum Inhalt die Verteilung der Fälle auf die Wertebereiche der Variablen abzuschätzen und auf diese Weise eventuell nicht vorhandene oder außergewöhnliche Fälle zu finden.

3 Entwurf

Die beiden ersten Aufgabenarten in (1) und (2) entsprechen einem einfachen Auswahlscenario und sollen in der Evaluierung die häufigste Verwendung finden. Beispielaufgaben aus der Leberchirurgie wären:

- Suchen Sie einen mittelschweren Fall, welcher durch eine Resektion behandelt wurde.
- Wählen Sie fünf Fälle bei denen eine Leberlebenspende vorgenommen wurde und die als einfach eingestuft sind.

Die beiden Kategorien in (3) und (4) entsprechen in etwa dem in Abschnitt 3.1.3 beschriebenen Szenario einer Prüfungs- oder Seminarvorbereitung. Beispielaufgaben aus der Leberchirurgie könnten hier lauten:

- Wählen Sie einen schweren, noch nicht bearbeiteten Fall, bei dem ein nicht rezidierbares CCC mittels Kryoablation behandelt wurde.
- Suchen Sie diejenigen schon einmal bearbeiteten Fälle einer Leberresektion eines HCC, bei denen Sie schlechter als befriedigend abgeschnitten haben.

Die in (5) erklärte Aufgabenkategorie entspricht etwa einem Muster, in welchem ein Fallautor die Fallauswahl eines Lernsystems benutzt. Eine Aufgabenstellung könnte lauten:

- Geben sie für jeden Schwierigkeitsgrad den im Bezug auf Tumortyp, Behandlung und Tumorage außergewöhnlichsten Fall an.

Um alle Anwendungsszenarien zu testen, sollten alle dieser Aufgabentypen in einer Evaluierung vorkommen. Bei den Kategorien ist von (1) bis (5) ein steigender Schwierigkeitsgrad zu verzeichnen, während das entsprechende Anwendungsszenario weniger häufig zur Anwendung kommt. Bei der Evaluierung werden deshalb mehr Aufgaben aus den leichteren Aufgabenkategorien und entsprechend seltener schwierige Aufgaben zu erfüllen sein.

Um diese Verteilung zu gewährleisten sollten also idealerweise zehn oder mehr Aufgaben in einer Evaluierung bearbeitet werden. Zu beachten ist dabei die maximale Bearbeitungszeit. Die Testpersonen können ermüden oder die Motivation verlieren, darum sollten nach jeweils drei Aufgaben kurze Pausen angeboten werden. Auch wird es mit steigender maximaler Bearbeitungszeit schwieriger genügend Teilnehmer für die Evaluierung zu gewinnen, da z. B. bei der Zielgruppe der Ärzte mit einem kleinen zur Verfügung stehenden Zeitfenster gerechnet werden muss. Eine Evaluierung muss so angelegt werden, dass im Idealfall aus nur zum Teil durchgeführten Experimenten nutzbare Daten gewonnen werden können, auch wenn Probanden nur einen Teil der Aufgaben bearbeiten oder nur teilweise Aufgaben abarbeiten aber dennoch vorhandene Fragebögen ausfüllen. Die fehlenden Ergebnisse können über eine Gewichtung der Ergebnisse ausgeglichen werden.

Von Probanden der direkten Zielgruppe (Mediziner) gewonnene Daten sollten bei der Auswertung stärker ins Gewicht fallen. Vor allem bei der Interpretation der subjektiven Aussagen und Bewertungen durch die Probanden sollte diese Gruppe gesondert untersucht werden.

3.3.4 Zu erhebende Daten

Die Frage, welche Größen gemessen werden sollen und welche Daten z. B. in einem Fragebogen erhoben werden wird durch das Ziel der Evaluierung bestimmt. Im Fall dieser Arbeit ist dies, eine Entscheidungsempfehlung geben zu können, welche Visualisierungstechnik sich für die Fallauswahl in einem medizinischen Lernsystem eignet. Diese eignet sich dann besonders gut wenn mit ihr schnell richtige Entscheidungen getroffen werden. Das Hauptaugenmerk liegt also auf der Performanz, der Bearbeitungszeit und der Korrektheit der Auswahl. Weiterhin sollte die Technik intuitiv zu bedienen sein, woraus sich schließen lässt, dass auch die Zeit bis eine Technik effizient genutzt werden kann, der Lernaufwand, eine Rolle bei der Evaluierung spielt. Die genannten Eigenschaften können automatisch gemessen werden. Neben diesen automatischen Messungen muss in einer Evaluierung immer auch genug Raum für persönliche Kommentare gelassen werden, um so ein Feedback über die Nutzerzufriedenheit zu erhalten.

3.3.4.1 Automatische Messungen

Die Performanz betreffende Größen können durch verschiedene automatische Messungen aufgezeichnet werden. So kann die Bearbeitungszeit einer Aufgabe für jede Technik gemessen werden. Um Unterschiede in der Bearbeitungszeit zu erklären, können Mauswege in Pixel und die Anzahl der benötigten Mausklicks aufgezeichnet werden, um so eventuelle Zusammenhänge zu erkennen. So wird eine Technik, welche oft eine umgebende GUI benutzt, wie es beim INFOZOOM der Fall ist um Zoomoperationen rückgängig zu machen und zu wiederholen, mehr Mauswege in Anspruch nehmen als eine Technik, welche alle Interaktionen integriert anbieten kann. Falls also der INFOZOOM längere Bearbeitungszeit in Anspruch nimmt, wäre dies eine Erklärung.

Der Lernfortschritt kann durch das Variieren und Wiederholen bestimmter, wie in Abschnitt 3.3.3 beschriebenen Aufgabentypen gemessen werden. So werden bei stetig gesteigertem Schwierigkeitsgrad am Anfang, in der Mitte und am Ende des Experiments die Bearbeitungszeiten für eine leichte Aufgabe gemessen. Dadurch kann im Nachhinein untersucht werden, bei welchen Techniken ein Unterschied zu verzeichnen ist.

Auch die Korrektheit kann automatisch registriert werden. Aufgezeichnet werden hierbei die ausgewählten Fälle bei normalen Fallauswahlszenarien oder die eingegebenen Antworten bei Fragen zu Inhalten der Datenbank.

Bei Fallauswahlen können alle Auswahlen mitgeschnitten werden, so dass auch Selektion und Deselektion falscher Fälle im Evaluierungsprotokoll verzeichnet sind. Dies kann eine weitere Ursache für Performanzverluste sein und weiteren Aufschluss über die Verwendbarkeit der jeweiligen Visualisierung geben.

Bei Fragestellungen zum Fallspektrum werden die getippten Antworten ausgewertet und auf Fehler untersucht. Dabei sollten Entscheidungsfragen bevorzugt werden, da Antworten in diesem Fall leichter registriert und zudem noch nach falsch positiv, richtig positiv, falsch negativ und richtig negativ aufgeschlüsselt werden können.

3.3.4.2 Befragungen

Fragebögen sollen einerseits Daten über die Testpersonen und andererseits über deren subjektive Meinung zu den evaluierten Visualisierungen gewinnen. So können Schlüsse über das Nutzungsverhalten von Personengruppen mit Techniken aus der Informationsvisualisierung gezogen werden. Diese fließen dann in eine spätere Weiterentwicklung ein. Auf jedem Fragebogen sollte zusätzlich Platz für Anmerkungen und Kommentare der Testpersonen geboten werden, falls die Antwortmöglichkeiten nicht den nötigen Spielraum für die Meinung der Teilnehmer lassen.

Persönliche Daten werden erhoben um die Teilnehmer einer Evaluierung in Gruppen zu unterteilen, welche bei der Auswertung verschieden gewichtet werden. Für die vorliegende Evaluierung ist insbesondere die Berufsgruppe wichtig, da die Eignung in medizinischen Falldatenbanken getestet wird. Folglich sind hier die Meinungen der Teilnehmer aus der direkten Zielgruppe der Mediziner besonders wichtig. Eine weitere Unterteilung wird nach Computerkenntnissen, hier speziell nach Vorkenntnissen mit den vorgestellten Techniken, unternommen. Dies ist eine weitere Methode um auffällige Unterschiede in den Bearbeitungszeiten zu erklären, so dass deren Auswertung differenzierter betrachtet werden kann, da zu erwarten ist, dass geübte Nutzer, welchen die Techniken bekannt sind, schneller zu Ergebnissen kommen werden und dabei weniger Fehlselektionen vornehmen. Auch wird für diese Nutzer kein merklicher Lernerfolg messbar sein, da für sie auch kein Lernaufwand mehr erforderlich ist.

Die subjektive Bewertung der evaluierten Visualisierungen spielt eine wichtige Rolle für die Einsatzempfehlung einer Technik, da Menschen immer die Technik bevorzugen welche ihnen subjektiv am meisten zusagt, obwohl sie eventuell messbar nicht die geeignetste Technik ist. Dabei werden Meinungen zu verschiedenen Aspekten der

jeweiligen Technik aufgezeichnet. Diese Angaben können von den Erkenntnissen aus den automatisch gemessenen Werten stark abweichen und werden benutzt, um sie zu bestätigen bzw. zu relativieren.

- **Intuitivität** Spiegelt das persönliche Empfinden des Lernerfolges wieder.
- **Interaktionen** Damit wird ausgedrückt wie ausreichend Interaktionsmöglichkeiten vorhanden waren und wie gut damit gearbeitet werden konnte.
- **Überblick/Ausreißer** Im welchem Maß konnte die Technik subjektiv einen Überblick über das Fallspektrum geben und wie gut waren Ausreißer erkennbar.
- **Effizienz der Fallauswahl** Gibt an, wie die Performanz subjektiv eingeschätzt wurde. Zum Beispiel kann durch eine sehr oft wiederholte Interaktion für eine Technik der Eindruck vermittelt werden, redundant bedient zu werden, obwohl sie objektiv gemessen die beste Performanz erzielt.
- **Gesamteindruck** Hier wird angegeben, wie eine einzelne Technik allgemein bewertet wird, also wie gut der Visualisierungsansatz der Technik als Ganzes empfunden wird. Es wird erwartet, dass der Gesamteindruck dem Mittel der vorhergehenden Eigenschaften entspricht.
- **Empfehlung** Gibt an, welche Technik vom Nutzer selbst am geeignetsten empfunden wird. Hier wird dem Probanden die Möglichkeit gegeben, eine subjektive Antwort auf die Frage zu geben, welche mit der vorliegenden Arbeit objektiv beantwortet werden soll.

Zusatzfragen Zu den Daten innerhalb eines Fallspektrums können neben den Bearbeitungsaufgaben auch innerhalb der Fragebögen noch einmal Fragen gestellt werden. So kann beurteilt werden wie exakt eine Technik auch Eigenschaften von Fällen vermittelt, welche bei der Bearbeitung einer bestimmten Aufgabe nicht von Belang waren. Weiterhin ist auch hier eine Möglichkeit gegeben nach Besonderheiten oder gar fehlenden Fällen, sprich nicht abgedeckten Wertebereichen zu fragen. Durch diese Fragestellungen außerhalb der Techniken wird folglich gemessen wie gut sie unbewusst einen Überblick vermitteln, während Bearbeitungsaufgaben ein bewusstes Auseinandersetzen mit den Daten fordern. Ein Einsatz beider Varianten wird angestrebt, auch weil durch Fragebögen die Gesamtdauer eines Experimentes reduziert werden kann.

3.4 Konzept des VisualizationEvaluators

Nachdem durch die Charakterisierung einer Fallauswahl geeignete Techniken identifiziert und die Rahmenbedingungen für eine Evaluierung bestimmt werden konnten,

wird in diesem Abschnitt der VisualizationEvaluator konzeptioniert. Der VisualizationEvaluator wird in diesem Zusammenhang implementiert, um ein Programm zur vollautomatischen Durchführung der empirischen Studie zur Verfügung stellen zu können. Der Aufbau und die Konzepte des Programms werden an folgendem Ablauf erläutert.

Begrüßung, Aufnahme persönlicher Daten Die Aufnahme persönlicher Daten, der Berufsgruppe, der Computerkenntnisse und der Vorkenntnisse mit den Techniken ist ein wichtiger Bestandteil für die zielgruppengerechten Interpretation (siehe Abschnitt 3.3.4.2) der Kommentare und einer eventuellen Gewichtung der Ergebnisse der elektronischen Messungen (siehe Abschnitt 3.3.4.1).

Bearbeitung der Aufgaben Die Bearbeitung der Testaufgaben ist in vier Blöcke geteilt, einer für jede Technik. Da Lerneffekte über die Aufgaben von Technik zu Technik auftreten können, wird sicher gestellt, dass für jede Testperson die Reihenfolge der Blöcke variiert. Der Ablauf jedes Blocks ist folgendermaßen:

- 1. Vorstellung der Technik** Auch die Intuitivität bzw. der Lerneffekt wird gemessen. Um die Unterschiede deutlich messen zu können wird die Bearbeitung ungelernt erfolgen. Hierbei muss dafür Sorge getragen werden, dass die Testpersonen die Techniken, trotz ungelernter Benutzung, diese vor der Nutzung kennen gelernt haben. Als bester Kompromiss zwischen statischer Präsentation und dem Kennenlernen durch Benutzung werden hier bearbeitete Bildschirmvideos angesehen, welche die Interaktionsmöglichkeiten kurz vorführen und sicher stellen, dass Testpersonen diese schon einmal in Aktion gesehen haben.
- 2. Bearbeitung der Aufgaben 1 bis 5** Es werden fünf Aufgaben zur Bearbeitung vorgegeben. Die geringe Anzahl der Aufgaben wurde gewählt, um die maximale Bearbeitungszeit der Evaluierung einzugrenzen. Die Aufgaben folgen der Aufteilung und den Kategorien, wie sie in Abschnitt 3.3.3 ausgearbeitet wurden. So folgen auf eine sehr einfache Aufgabe zwei mittelschwere Aufgaben und eine sehr schwere Aufgabe, welche testet, wie gut die Technik einen Überblick über die Daten zulässt. Die Bearbeitung der Testaufgaben wird durch eine weitere, sehr einfache Aufgabe abgeschlossen. Nach der Bearbeitung jeder einzelnen Aufgabe wird ein Fragebogen angeboten, welcher die subjektive Einschätzung der Eignung durch den Nutzer aufzeichnet. Auf eine zufällige Reihenfolge der Aufgaben zur weiteren Vermeidung der erwähnten Lerneffekte wird hier verzichtet, um einen ansteigenden Schwierigkeitsgrad sicherzustellen. Dieser ist nötig, um den Lernerfolg innerhalb des Bearbeitungsblockes einer Technik im Nachhinein zwischen den Techniken vergleichen zu können und um Frustration zu vermeiden, welche durch einen sehr hohen Schwierigkeitsgrad am Beginn der Bearbeitung entstehen kann. Diese könnte so auch die erwartet

unvoreingenommene Meinung zu einzelnen Techniken beeinflussen, was eine vergleichende Interpretation der subjektiven Einschätzung verfälschen bzw. unmöglich machen könnte.

Abschließende Fragebögen Hier werden zusammenfassend die subjektiven Einschätzungen der Nutzer abgefragt. Hierzu sollen die Testpersonen die Interaktionen der Techniken, die Vermittlung eines Überblicks über die Daten und die Effizienz der Techniken bewerten. Die Testpersonen sollen hier auch eine Präferenz für eine der Techniken angeben.

Versenden der Logdatei per Email Das aufgezeichnete Protokoll der Nutzung aller Techniken und aller Antworten aus den Fragebögen werden in Form einer Logdatei per Email versandt.

Die Evaluierung wird mit Hilfe des Programmes anonym und vollautomatisch ablaufen, es kann über das WWW von interessierten Testpersonen heruntergeladen werden. Dies hat den Vorteil, dass potentiell sehr viele Versuchspersonen gewonnen werden können, was potentiell zu (statistisch) stabilen Ergebnissen führt. Werden Personen aus der direkten Zielgruppe, müssen diese nicht zu einem extra vereinbarten Termin an einem extra eingerichteten Ort erscheinen, sondern können als potentielle Nutzer der Techniken diese auch am potentiellen Einsatzort testen. Das Experiment wird als WITHINSUBJECTS-Experiment aufgebaut, d. h. jede Testperson bearbeitet jede Aufgabe mit allen Techniken, so wird die maximale Anzahl an Testfällen garantiert. Die direkte Integration von Fragebögen in das Programm erhöht die Wahrscheinlichkeit, dass Versuchspersonen diese auch ausfüllen. Die Erinnerung an die Bearbeitung der Aufgaben ist direkt danach stärker als wenn beispielsweise erst eine bestimmte Webseite besucht werden muss, um dort Fragebögen auszufüllen.

Das beschriebene Vorgehen birgt jedoch Probleme. So verliert der Durchführende der Evaluation einen Teil seiner Kontrolle über den Testvorgang, da unvorhergesehene Störungen, z. B. durch andere Personen, auftreten können. Auch ist es ihm nicht möglich auf eventuelle Fragen einzugehen. Deshalb ist zu erwarten, dass mehr Datensätze als bei anderen Testsettings nicht verwendet werden können. Zur sicheren Unterscheidung zwischen verwertbaren und nicht nutzbaren Daten ist eine sehr detaillierte Protokollierung der Nutzung der Techniken vorgesehen. Um sicherzugehen, dass die Messdaten nicht durch Wiederholungen verfälscht werden, ist der Ablauf streng linear, es kann nicht in der Bearbeitung zurückgegangen werden.

3.5 Kapitelzusammenfassung

Es wurden in diesem Kapitel geeignete Visualisierungstechniken für die Fallauswahl diskutiert. Mit TABLENS, INFOZOOM und PARALLELE KOORDINATEN wurden

3 Entwurf

diejenigen Techniken ausgewählt, welche für eine Evaluierung umgesetzt werden. Tabellen scheinen ebenfalls geeignet und werden umgesetzt um einen Vergleich zwischen etablierten und neuartigen Visualisierungstechniken zu ermöglichen. Der Vorgang der Fallauswahl wurde anhand wichtiger Kriterien, wie der Datenbankstruktur, den Nutzern und realistischen Nutzungsszenarien, charakterisiert und daraus Rahmenbedingungen für die durchzuführende Evaluierung und Thesen zur Benutzung der Techniken abgeleitet. Es konnten geeignete Testaufgaben und Testpersonen sowie die zu erhebenden Daten bestimmt werden. Weiterhin wurde die Konzeption des VisualizationEvaluators, eines Programmes zur vollständigen und anonymen Durchführung einer empirischen Studie, vorgestellt.

4 Umsetzung

In diesem Kapitel wird näher auf die Implementierung der ausgewählten Visualisierungstechniken eingegangen. Hierbei werden insbesondere die umgesetzten Interaktionsmöglichkeiten erläutert. Des Weiteren werden die Unterschiede zu den jeweiligen Originalimplementationen dargelegt und begründet, warum bei den Techniken Eigenschaften hinzugefügt oder auch ausgelassen wurden. Auch die Umsetzung des Programms zur Durchführung der empirischen Evaluierung wird näher erläutert.

4.1 Werkzeuge

Die Implementierung der Techniken sowie des Evaluationsprogramms erfolgte in der Entwicklungsumgebung MEVISLAB [41]. Diese unterstützt einen modularen Aufbau und beinhaltet eine Vielzahl vorgefertigter Funktionsmodule zur Erstellung komplexer Funktionsnetzwerke insbesondere zur medizinischen Bildverarbeitung mit Hilfe der integrierten Funktionsbibliothek für Bildverarbeitung, der ML-Bibliothek. Netzwerke können in Macromodule zusammengefasst werden um Funktionseinheiten zu bilden.

Weiterhin bietet MEVISLAB eine fast vollständige Umhüllung der 3D-Bibliothek OpenInventor zur Erstellung von dreidimensionalen Szenen mit MEVISLAB-Modulen. In OpenInventor werden Formprimitive objektorientiert zu komplexen Szenegraphen zusammengefasst. Eine nähere Beschreibung der Konzepte von OpenInventor findet sich in [53].

Module kommunizieren untereinander über eine Feld-API. Auch Datenein- und Ausgänge der Module sind über Felder abgebildet. In ihnen werden alle Eigenschaften von Modulen gespeichert, welche anderen Modulen oder einer Benutzungsschnittstelle kommuniziert werden sollen. Die Entwicklungsumgebung selbst ist mit Hilfe des Qt-Frameworks der Firma Trolltech umgesetzt. Für die Anbindung eigener Qt-Komponenten werden ebenfalls Schnittstellen im MEVISLAB angeboten. Eigene Module können in der Programmiersprache C++ entwickelt und in die Umgebung integriert werden. Eine weiterführende Übersicht über das Qt-Framework findet sich in [7].

Zur Gestaltung der Benutzungsschnittstelle eigener Module bietet MEVISLAB in der internen Modulbeschreibungssprache MDL zahlreiche Komponenten an, darunter

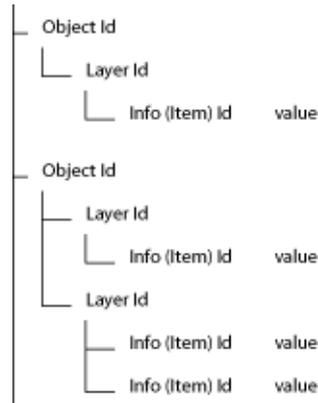


Abbildung 4.1: Datenablageformat des ObjMgr in MEVISLAB [12].

selbst komplexe Elemente, etwa einen Videoplayer. Die erstellte Oberfläche kann über verschiedene Skriptsprachen gesteuert werden. Um eigene Anwendungen kommerziell vertreiben zu können, ist es möglich alleinstehende Anwendungen mit MEVISLAB zu erstellen.

4.2 Techniken

Im Folgenden wird näher auf die Umsetzung der Techniken und ihrer Interaktionsmöglichkeiten eingegangen.

4.2.1 Datenaquirierung

Für die Evaluierung wurden bestehende Fälle aus der Datenbank des an Otto-von-Guericke-Universität entwickelten LIVERSURGERYTRAINER verwendet, welche in Form von XML-Dateien vorlagen. Die vorhandene Datenbank wurde synthetisch um weitere Fälle vergrößert.

Für die Bereitstellung der darzustellenden Daten wurde eine Schnittstelle (Modul CaseScanner) entwickelt, welche mit Hilfe eines Python-Skriptes die Variablen und Fälle der Falldatenbank aus allen in einem gegebenen Verzeichnis gefundenen XML-Dateien extrahiert und diese allen verbundenen Visualisierungsmodulen in einem einheitlichen Format zur Verfügung stellt. Verwendet wurde hier das MEVISLAB-interne Datenverwaltungsmodell des ObjectManagers [12]. Dieser legt arbiträre Informationen in einer dreischichtigen Form nach Objekten, Layern und Informationszellen ab (Abbildung 4.1). Mit Hilfe der CaseScanner-Schnittstelle wurde die Datengewinnung gekapselt und ein einheitlicher Anschlusspunkt für alle Visualisierungsmodule geschaffen (Abbildung 4.2).

Alle der nachfolgenden Umsetzungen der verschiedenen Visualisierungstechniken nutzen die ObjectManager-API, indem sie die ObjectManagerClient-Schnittstelle implementieren. So entsteht die Möglichkeit, diese Visualisierungsmodule direkt mit einer Datenbank in Form eines ObjectManagers zu koppeln, wie sie in der CaseScanner-Schnittstelle enthalten ist.

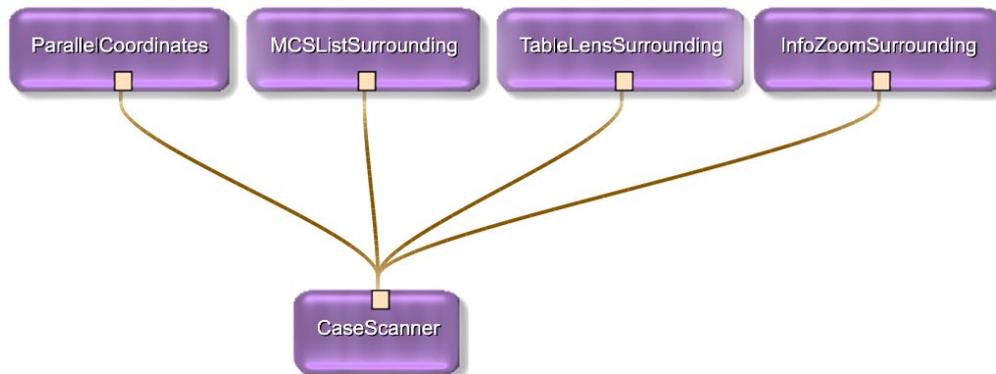


Abbildung 4.2: CaseScanner-Schnittstelle zur Datenbereitstellung mit mehreren verbundenen Visualisierungsmodulen

4.2.2 Einbettung der Techniken

Eine Oberfläche wurde in MEVISLAB-MDL implementiert, welche Visualisierungsmodule einbetten kann (Abbildung 4.3). So können Zusatzinformationen zum Fall unter dem Mauszeiger angezeigt werden. In einer zusätzlichen Tabelle wird die Auswahlmenge der Visualisierung gezeigt und es besteht die Möglichkeit, die Benutzungsanleitung der eingebetteten Technik anzuzeigen.

4.2.3 Aufbau der Techniken allgemein

Alle implementierten Visualisierungen folgen in großen Teilen einem abgewandelten ModelViewController UserInterface-Paradigma [28], welches eine Trennung von Datenmodell, Ansicht und Eingabeschnittstelle vorsieht, die eventgesteuert kommunizieren. Ein einheitliches Datenmodell wird von der CaseScanner-Schnittstelle bereitgestellt. Der Modellteil des Paradigmas wird bei den Tabellen, dem INFOZOOM und der TABLELENS durch in Qt geschriebene Tabellendatenmodelle realisiert. Bei den PARALLELEN KOORDINATEN wurden in Standard-C++ eigene Klassen zur Abbildung der erstellten ObjMgr-Objekte, der Dimensionen und Fälle, definiert. Die Datenmodelle werden von den Visualisierungsteilen der implementierten Module dargestellt. View-

4 Umsetzung

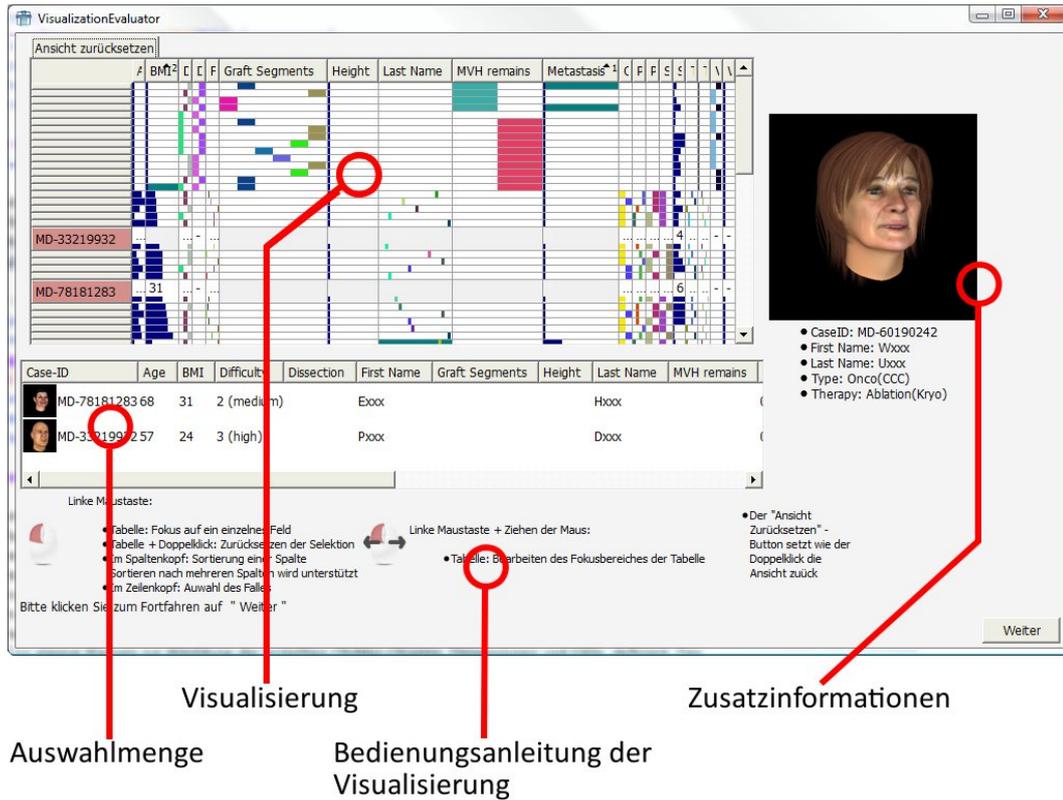


Abbildung 4.3: Darstellung einer Visualisierung einbettende Oberfläche am Beispiel der TABLENS-Visualisierung

und Controllerteil werden nicht komplett voneinander getrennt, da diese Trennung weder im Qt-Framework noch in der OpenInventor-Bibliothek vorgesehen ist. Im Qt-Framework werden View- und Controllerteil des Paradigmas zusammengefasst, wobei die Ansicht beim Zeichnen die geometrische Aufteilung des Bildschirms übernimmt, während das konkrete Zeichnen von Delegatenklassen übernommen wird (Abbildung 4.4). Für alle Teile des Paradigmas werden in Qt Standardimplementationen angeboten. Diese wurden für die Techniken, wie in den folgenden Abschnitten beschrieben, erweitert. Die OpenInventor-Bibliothek konzentriert sich vollkommen auf die Darstellung der 3D-Szene und der Interaktion mit dieser. Die Interaktionen sind wie in Qt eventgesteuert. Zur Datenverwaltung stellt OpenInventor keine API zur Verfügung (was allerdings auch nicht Intention der Bibliothek ist).

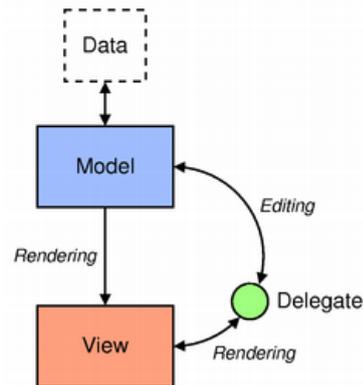


Abbildung 4.4: Model-View-Architektur des Qt-Frameworks nach [6]

4.2.4 Tabellen

Um die Tabellenvisualisierung im Gegensatz zu den anderen gewählten Visualisierungen konkurrenzfähig zu halten und ein gezieltes Suchen nach mehreren Kriterien zu ermöglichen, wurde eine Sortierung nach mehreren Tabellenspalten umgesetzt. Da die intern, in der MDL von MEVISLAB angebotenen Oberflächenelemente nicht erweitert werden können, wurde für die Tabellen eine eigene Qt-Komponente geschaffen. Hierfür wurde eine im Qt-Framework bereits vorhandene Sortierfunktionalität in einem eigenen, vom Qt-Standardtabellenmodell abgeleiteten Tabellenmodell auf die Mehrfachsortierung erweitert. Qt-intern wird für die Sortierung eine StableSort-Implementierung verwendet. Dafür werden die zu sortierenden Zeilen separiert und zusammen mit einer Implementation eines Vergleichsobjektes an den Algorithmus übergeben. Dieses Vergleichsobjekt wurde so erweitert, dass bei zwei Tabellenindices gleichen Wertes in einer Sortierspalte die Sortierung an die nächste Sortierspalte weitergegeben wird, solange kein eindeutiges Ergebnis vorliegt oder nach allen Sortierspalten verglichen wurde.

Ein eigener Tabellenkopf, welcher den Standardtabellenkopf von Qt erweitert, wurde implementiert um die zusätzlichen Informationen visuell darzustellen (Abbildung 4.5). Eine Interaktion, welche das Umordnen von Spalten ermöglicht, wird in der eigentlichen QHeaderView bereits angeboten und wird in der Technik verwendet. Weiterhin bietet Qt zur Selektion in Tabellen, Baumansichten oder generell allen Sichten, die im ItemView Framework angesiedelt sind, mehrere Selektionsstrategien an. Für die Tabellen wurde die multiple Toggle-Strategie für ganze Zeilen bzw. Fälle gewählt. Ein Klick auf einen Fall in der Tabelle schaltet den Selektionsstatus um, ein weiterer macht die Selektion rückgängig. So können mehrere Fälle selektiert werden, ohne dass die bisherige Auswahl davon beeinträchtigt wird.

CaseID	Age ¹	BMI ²	Difficulty ³	
MD-80425421	85	19	2 (medium)	-
MD-43438950	85	21	2 (medium)	-
MD-74598556	85	43	2 (medium)	-
MD-40052695	84	36	1 (low)	-

Abbildung 4.5: Tabellenkopf einer nach mehreren Spalten sortierten Tabelle

4.2.5 TableLens

Die Implementation baut konsequent auf der der Tabellen auf. Es wurden alle Interaktionen übernommen. Zusätzlich wurde eine eigene Delegatenklasse implementiert, welche abhängig vom Status das Zeichnen der Zellen übernimmt. Hierdurch ergibt sich die eigentliche Änderung gegenüber einer gewöhnlichen Tabelle. Es sind hier vier verschiedene Zustände (siehe Abbildung 4.6) zu nennen:

Fokussiert Eine Zelle im wird grafisch, textuell und mit blauen Hintergrund gezeichnet.

Spaltenfokussiert Eine Zelle, die eine Spalte mit einer fokussierten Zelle teilt, wird grafisch minimiert gezeichnet, wobei die Spalte breiter ist.

Zeilenfokussiert Eine Zelle die eine Zeil mit einer fokussierten Zelle teilt wird grafisch und textuell gezeichnet. Dabei wird nicht garantiert, dass die textuelle Repräsentation lesbar ist.

Minimiert Eine Zelle außerhalb aller Fokusbereiche wird minimiert dargestellt. Ihre Breite berechnet sich nach der Breite aller anderen Spalten. Zehn Pixel sind jedoch das Minimum.

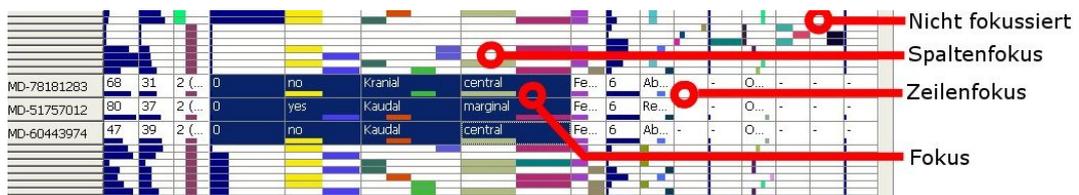


Abbildung 4.6: Mögliche Zustände der Zellen in der TABLELENS-Implementation

Abhängig vom Datentyp wird bereits bei der Initialisierung das Aussehen jeder einzelnen Zelle festgelegt.

	TABLELENS nach Rao [40]	Implementation
Erstellen neuer Fokusbereiche	Klicken und Ziehen der Maus über die gewünschten Zellen	
Verschieben eines Fokusbereiches	Selektion des gewünschten Fokusbereiches und Ziehen mit der Maus an die gewünschte Stelle	Deselektion der fokussierten Zellen und Selektion der gewünschten Zellen im neuen Fokusbereich
Größe eines Fokusbereiches ändern	Handle am Rand des Fokusbereiches nach innen/außen bewegen	Deselektion bzw. Selektion von Zellen am Rand des Fokusbereiches

Tabelle 4.1: Vergleich der Operationen der TABLELENS im Original und der implementierten Version

In Spalten mit numerischen Daten wird das Minimum und Maximum innerhalb der Spalte gesucht, die grafische Darstellung der einzelnen Dateneinträge wird mit einem Balken auf die Breite der Zelle gezeichnet. So entsteht über eine gesamte Spalte der Eindruck eines Balkendiagramms, anhand dessen z. B. die Verteilung der Werte abgeschätzt werden kann.

Bei ordinalen oder nominalen Daten werden so genannte COLOR SWATCHES eingesetzt. Jedem Wert, der in der Spalte vorkommt, wird dabei eine Farbe zugewiesen und abhängig von der Anzahl und der aktuellen Breite der Spalte berechnet die Delegatenklasse den Ort und die Breite eines Swatches.

Die Interaktionen zur Handhabung des Fokus, das Erstellen eines Fokusbereiches, das Verschieben eines Fokusbereiches und das Verändern seiner Größe, wurde auf tabelleninterne Markierung von Zellen abgebildet. Es können beliebig viele Zellbereiche markiert werden, so dass mehrere Fokuslinsen verwendet werden können. Da die Steuerung der Visualisierung die tabelleninterne Markierung ausnutzt, wird ein Fall über einen Klick auf den entsprechenden Zeilenkopf in die Lösungsmenge übernommen. Tabelle 4.1 stellt die Interaktionen der Implementation dem Original nach Rao [40] gegenüber. Der wesentliche Nachteil besteht beim Verschieben von Fokusbereichen. Die vorgenommene Implementierung benötigt zwei separate Schritte: das Löschen des alten und das Öffnen des neuen Fokusbereiches. Dies kann bei der Nutzung umständlich wirken, so dass durch diese Implementation ein Nachteil für diese Technik in der Evaluierung entstehen kann.

4.2.6 InfoZoom

Da auch der INFOZOOM einen tabellarischen Visualisierungsansatz verfolgt, konnte auch hier die Implementation (Abbildung 4.7) auf dem ItemView-API von Qt aufbauen, welches auch den Verbund von Zellen unterstützt, der für diese Visualisierung

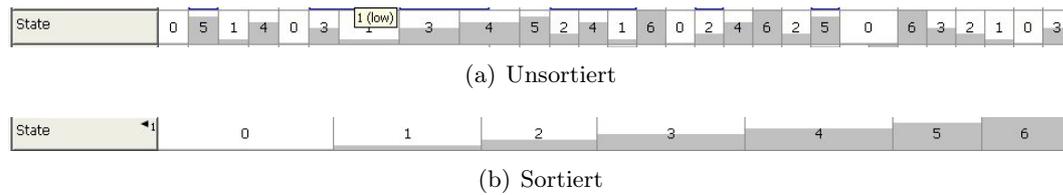


Abbildung 4.8: Quantitätsabschätzung in einer INFOZOOM-Zeile: Durch die Sortierung in (b) wird die Verteilung der Werte abschätzbar (hier ungefähr gleichverteilt).

4.2.7 Parallele Koordinaten

Die PARALLELEN KOORDINATEN (Abbildung 4.9) wurden in OpenInventor bzw. Standard C++ implementiert. Eigene Datenstrukturen zur Abbildung der Variablen

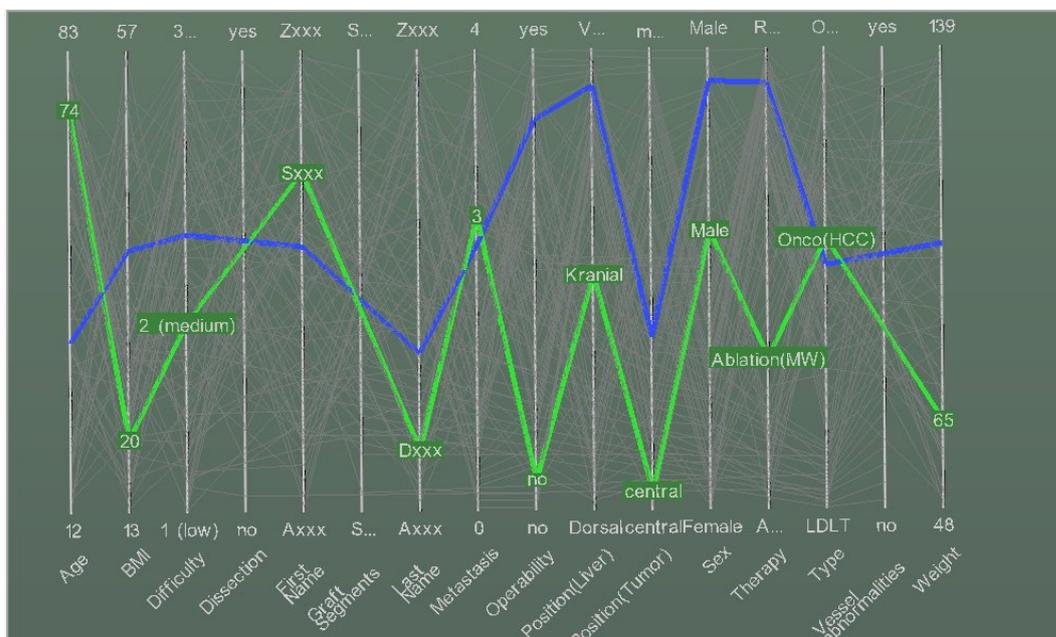


Abbildung 4.9: Implementierung der PARALLELEN KOORDINATEN: Ausgewählte Fälle werden blau dargestellt, der Fall unter dem Mauszeiger wird grün hervorgehoben, während alle anderen sichtbaren Falllinien verblasen.

und Fälle unterstützen die 3D-Szene, welche aus Inventorknoten aufgebaut ist. Statt Spalten- und Zeilenindices unterstützen diese Strukturen die Szene mit Informationen zur Positionierung im Raum. Die Implementation wurde für nominale und solche Daten angepasst, die nur wenige diskrete Werte enthalten. Daraus resultieren größere

	Maus über	Mausklick auf
Falllinie	Falllinie wird grün hervorgehoben, es werden an den Schnittpunkten mit den Dimensionsachsen die Werte des Falles in der Dimension als Tooltip angezeigt. Alle anderen sichtbaren Falllinien werden blasser und dünner dargestellt.	Auswahl des Falles, er wird blau markiert und bleibt auch bei weiteren Mausinteraktionen sichtbar.
Dimensionsachse	Anzeige des Wertes, welcher durch diesen Achsenabschnitt abgebildet ist, als Tooltip. Der Abschnitt wird farblich rot hervorgehoben. Alle Fälle, die in dieser Variable diesen Wert nicht besitzen werden blasser dargestellt. So entsteht eine Vorschau auf die Sicht, wie sie aussehen würde, wenn dieses Achsenelement selektiert würde.	Selektion des Abschnittes, hierbei werden alle Falllinien ausgeblendet die diesen Achsenabschnitt nicht schneiden.

Tabelle 4.2: Darstellungsvarianten der PARALLELEN KOORDINATEN bei Mausinteraktionen.

Achsenabschnitte pro Wert. Um diesen zusätzlichen Platz auszunutzen werden die Falllinien, welche einen Achsenabschnitt schneiden, gleichmäßig auf den vorhandenen Platz aufgeteilt, um mehr Sichtbarkeit zu erlangen. Zu diesem Zweck werden in den Strukturen auch die Anzahl der Teilstücke der Achsen und die Anzahl der Treffer pro Teilstück gespeichert, um die Berechnung der Koordinaten der Linienpunkte eines Falles zu erleichtern. Auch die für Tooltips benötigten Texte sind hier hinterlegt. Die OpenInventor-Szene leitet sich von einem SoSeparator ab. Der Aufbau der Szene wird durch Abbildung 4.10 verdeutlicht.

Das größte Problem der Visualisierung der PARALLELEN KOORDINATEN ist die Übersicht. Selbst bei kleinen Datenbanken mit ca. 100 Einträgen kommt es zu zahlreichen Überzeichnungen. Um eine höhere Übersichtlichkeit zu ermöglichen wurden Techniken zur interaktiven Exploration der Daten umgesetzt. Tabelle 4.2 fasst die implementierten Varianten zusammen.

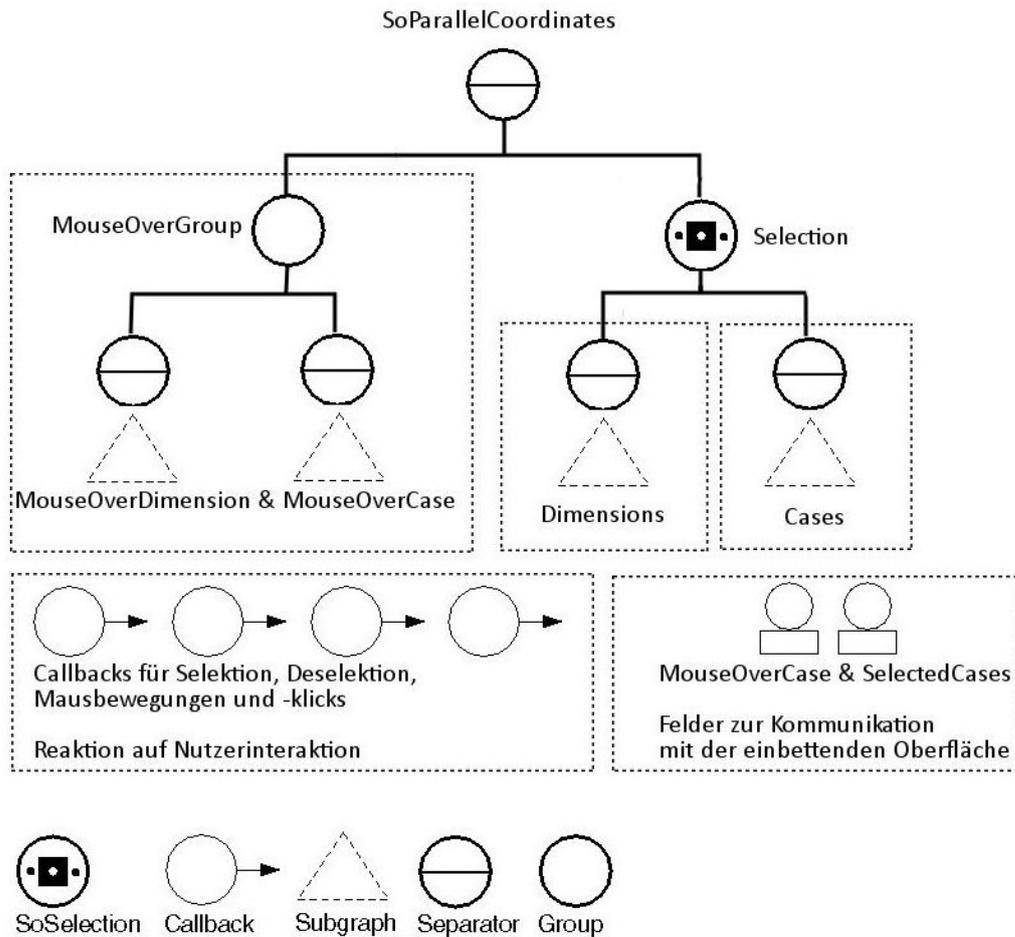


Abbildung 4.10: Aufbau der OpenInventor-Szene des SoParallelCoordinates-Moduls: Der linke Subgraph blendet bei Mausbewegungen über Fällen bzw. Dimensionsachsen Tooltips ein, der rechte Subgraph beinhaltet die eigentliche Szene aus Dimensionsachsen und Falllinien. Beides kann selektiert werden, durch den SoSelection-Knoten Selection wird die in OpenInventor angebotene Toggle-Selektion benutzt. Callbacks werden benutzt um bei Mausbewegungen das Aussehen der Szene zu ändern oder die Selektion von Achsen oder Fällen darzustellen. Über StringFields kommuniziert die Technik mit der einbettenden Oberfläche.

4.3 Der VisualizationEvaluator - Das Programm zur Evaluierung

In Abschnitt 3.4 wurde der VisualizationEvaluator als Programm zur anonymen und vollautomatischen Durchführung der Evaluierung konzipiert. Der Grad der Automatisierung macht es nötig, die Techniken aus der Entwicklungsumgebung zu lösen und einheitlich in ein Programm einzubetten. So wurde mit Hilfe von MEVISLAB eine Installationsroutine erzeugt, welche alle zur Evaluierung nötigen Daten enthält und wie ein gewöhnliches Programm von der Testperson installiert werden kann.

Der Ablauf und Aufbau des Programmes ist in Abbildung 4.11 noch einmal skizziert. Es besteht aus einem Hauptfenster mit einem Button, der den Nutzer zum nächsten Schritt weiterleitet. Der Übergang von einem Schritt zum nächsten ist dabei immer identisch: das Hauptfenster wird ausgeblendet, dafür ein anderes eingeblendet, welches den Nutzer auffordert, kurz zu warten. In der Zwischenzeit werden eingegebene Daten, z. B. Antworten von Fragebögen, in eine Logdatei geschrieben. In diese werden auch Steuerinformationen für das Programm eingetragen. Falls es unterbrochen wird, kann so die gleiche Bearbeitungsreihenfolge der Aufgabenblöcke wiederhergestellt werden, zum richtigen Schritt gesprungen werden. Wenn in einem Schritt eine Aufgabenbearbeitung stattfindet, werden bei Einblendung des Fensters und bei der Beendigung der Bearbeitung Zeitstempel im Logfile festgehalten, welche es ermöglichen, die benötigte Zeit für die Aufgabe zu berechnen. Die ausgewählten Fälle werden ebenfalls in die Logdatei geschrieben. So wird es möglich zu überprüfen, ob die korrekten Fälle gewählt wurden.

Die Oberfläche wurde komplett in der MDL von MEVISLAB umgesetzt und wird durch ein Python-Skript gesteuert. Umgesetzt wurde hier das Hauptfenster, welches die Automatik zum Weiterschalten zum nächsten Schritt beinhaltet. Zusätzlich sind Fragebögen zur Aufnahme persönlicher Daten, zur Nutzbarkeit der Technik nach jeder Aufgabenbearbeitung und für die Aufnahme der subjektiven Bewertung der Techniken integriert. Weiterhin wurde die einbettende Oberfläche für alle Techniken integriert.

Über das Python-Skript wird zu Beginn eines Durchlaufs die Reihenfolge der Aufgabenblöcke festgelegt, welche bei Unterbrechung des Durchlaufs in der Logdatei gespeichert werden.

Die Aufgabenbearbeitungsblöcke für die vier Techniken sind jeweils gleich aufgebaut. Sie beginnen mit einem Video, in dem die Interaktionsmöglichkeiten der Technik vorgestellt werden. Die Videos wurden zusätzlich nachbearbeitet, um die Arbeitsabläufe mit der Technik zu verdeutlichen (Abbildung 4.12). Es folgt ein schematischer Aufbau der Bearbeitungsoberfläche, so dass sich der Nutzer sofort zurechtfinden kann. Danach folgen für fünf Aufgaben jeweils die Bearbeitungsoberfläche und nach der Bearbeitung ein Fragebogen, welcher unmittelbar aufzeichnet, wie der Nutzer die

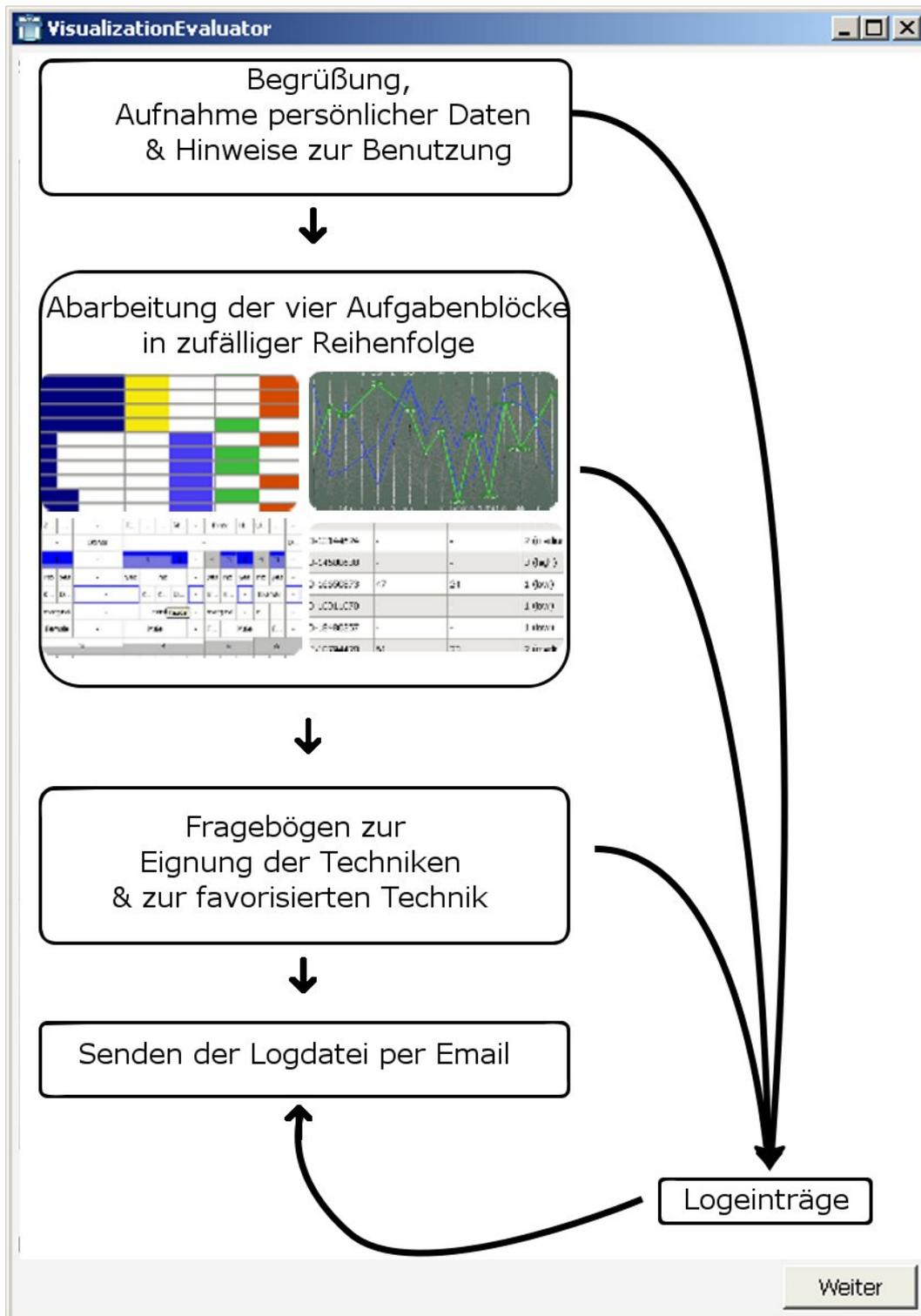


Abbildung 4.11: Aufbau/Ablauf des VisualizationEvaluators

4 Umsetzung

The screenshot shows the 'VisualizationEvaluator' application window. At the top, it displays the title 'Vorstellung der Techniken - InfoZoom'. Below this, there is a task description in German: 'Aufgabe 1 Suchen Sie einen Techniken, Originalen FallType = Onco) welcher mit einer Reaktion behaftet(Chilaps = Reaction) wurde. Auch der InfoZoom ist eine Art erweiterte Tabelle. Allerdings werden die Attribute in den Zeilen und die eigentlichen Fälle in den Spalten abgebildet, sprich die Zeilen mit den Spalten vertauscht. Es ist zunächst immer die gesamte Datenbank auf dem Bildschirm zu sehen, über Sortier- und Zoomoperationen kann diese Fallmenge geordnet und verkleinert werden. Bei Mausbewegungen über der Technik werden Zusatzinformationen zum jeweiligen Fall angeboten und alle Zeilen welche den selben Wert besitzen blau unrandet. Sind Fälle ausgewählt, werden diese in Spaltenkopf rot angezeigt und in die Auswahlmenge übernommen.'

The main area of the window displays a data table with columns: Case-ID, Age, BMI, Difficulty, Description, First Name, Graft Segments, Height, Last Name, Mini-remans, Metastasi, Operability, Sex, Therapy, Type, and Risk. The table is annotated with red text and arrows. The text 'Umordnen von Zeilen' (Reordering rows) is written in red above the table. Below it, 'Klicken und Ziehen der linken Maustaste' (Click and drag left mouse button) is written in red, with a red arrow pointing to the left mouse button icon in the legend below. A small portrait of a man is visible on the right side of the table.

Below the table, there is a legend titled 'Folgende Interaktionen sind möglich:' (The following interactions are possible:). It lists several actions:

- Linke Maustaste:**
 - In der Tabelle: Selektion eines Wertes
 - In der Tabelle + Doppelklick: Zoom - Es werden alle Fälle ausblendet welche nicht dem geklicktem Wert oder den Werten der Selektion in dieser Zeile entsprechen.
 - Im Spaltenkopf: Auswahl des Falles
 - Im Zellenkopf: Sortierung der Zeile
- Rechte Maustaste:**
 - Im Zellenkopf: Umkehren der Sortierung
 - In der Sicht + Doppelklick: "Zoom(Zurück)"
- Linke Maustaste + Ziehen der Maus:**
 - Im Spaltenkopf: Umordnen der Zeilen

At the bottom of the window, there is a 'Weiter' button and a small note: 'Über die Buttons "<-Zoom(Zurück)", "Zoom(Vor)->" bzw. "Reset" ist ein Rückgängigmachen und Wiederholen von Zoomoperationen möglich und es kann die Ansicht auf Anfang zurückgesetzt werden. Bitte klicken Sie zum Fortfahren auf "Weiter"'

Abbildung 4.12: Die Visualisierungen wurden mit Hilfe von bearbeiteten Bildschirmvideos vorgestellt. Die Abbildung zeigt eine Szene aus der Beschreibung zum INFOZOOM.

Eignung der verwendeten Technik für die bearbeitete Aufgabe empfunden hat. Jeweils nach Aufgabe 4, der schwersten Aufgabe, wurde eine Zusatzfrage gestellt, welche sich auf den Überblick über die präsentierten Daten richtete. Abbildung 4.13 zeigt die Umsetzung der Bearbeitungsoberfläche.

Zum Ende des Programmes wird die gesamte Logdatei automatisch per Email übermittelt.

4.4 Kapitelzusammenfassung

Nach dem Finden geeigneter Techniken, der Anforderungsanalyse und der Konzeption der Evaluierung wurde in diesem Kapitel näher auf die Implementierung der Techniken und des Evaluierungskonzeptes eingegangen. Damit sind alle Vorausset-

Aufgabe 3
Wählen Sie die schwierigen (Difficulty = 3), noch nicht bearbeiteten (State = 0) Fälle, bei denen ältere Leute ab 60 Jahren nicht durch Resektion therapiert wurden

Table columns: MD-, MD-, MD-, CE-, MD-, CE-, MD-, CE-, HH-, HH-, MD-, MD-, MD-, MD-, MD-, LD-, HL-, MD-, MD-, CE-, MD-, LD-, MD-, MD-, CE-, MD-, MD-, MD-

Summary Table:

Case-ID	Age	BMI	Difficulty	Dissection	First Name	Graft Segments	Height	Last Name	MVH remains	Metastasis	Operability	Position(Liver)	Positi
MD-967404	136		3 (high)		Woox			Moox	1	yes	Dorsal	margr	
MD-375555	68		1 (low)		Woox			Woox	3	no	Ventral	centra	

Labels and Annotations:

- Aufgabenstellung**: Points to the task title 'Aufgabe 3'.
- Auswahlmenge**: Points to the selected row for Case-ID MD-375555 in the summary table.
- Benutzungsanleitung**: Points to the mouse control instructions at the bottom left.
- Einbettung der Technik**: Points to a cell in the main data table.
- Zusatzinfo**: Points to the patient photo and details on the right.
- Weiter zum nächsten Schritt**: Points to the 'Weiter' button at the bottom right.

Abbildung 4.13: Umsetzung der Bearbeitungsoberfläche des Evaluators am Beispiel einer Bearbeitung mit dem INFOZOOM

zungen für die Durchführung der Evaluation gegeben, welche im nächsten Kapitel im Mittelpunkt stehen wird.

4 Umsetzung

5 Durchführung, Auswertung & Ergebnisse der Evaluierung

Im Zuge dieser Arbeit wurden mehrere Visualisierungstechniken für den Einsatz zur Fallauswahl in einem medizinischen Trainingssystem analysiert und implementiert. Um den Vorteil einer solchen Technik in Gegensatz zu simplen Visualisierungen zu erforschen und damit den Einsatz zu rechtfertigen, wurde eine Evaluierung anhand einer Charakterisierung einer Fallauswahl konzeptioniert. Die Durchführung und insbesondere die Auswertung dieser Evaluierung ist Gegenstand des folgenden Kapitels.

5.1 Durchführung

Die Durchführung der Studie wurde in zwei Schritten durchgeführt. Im ersten Schritt wurde ein überwachter Evaluierungsdurchgang durchgeführt, um letzte Fehler im Evaluator (siehe Abschnitt 4.3) zu finden und um zu kontrollieren inwieweit mit dem Programm eine vollautomatische Evaluierung vorgenommen werden kann. Hierzu wurden Studenten der Medizin und einige Mitarbeiter des Institutes für Biometrie und medizinische Informatik des Universitätsklinikums Magdeburg eingeladen, einmal überwacht mit dem Evaluator zu arbeiten. Im Gegensatz zur anonymen Bearbeitung hatten die Teilnehmer hier die Möglichkeit Fragen zu den Techniken, dem Programm und den Testaufgaben zu stellen. Bei diesem Durchgang haben sechs Personen teilgenommen. Da nur vereinzelt Fragen oder Zweifel aufkamen, die zudem auf noch vorhandene Fehler in der Implementierung zurückzuführen waren, konnte angenommen werden, dass mit dem Programm eine empirische Studie auch anonym und vollautomatisch möglich ist.

Aufgrund der Erkenntnisse des ersten Durchganges wurden letzte Fehler im Evaluator beseitigt und dieser über Mailinglisten und Downloadmöglichkeiten einer größeren Gruppe von Testpersonen zugänglich gemacht. Auch hier wurde versucht, möglichst Testpersonen aus dem medizinischen Bereich zu erreichen. Unter anderem wurde die Downloadmöglichkeit des Evaluators in Verbindung mit einigen Hinweisen zur Installation und dem Hintergrund der Studie Kooperationspartnern des Lehrstuhls für

Visualisierung und einigen Ärzten bekanntgegeben. Diese wurden wiederum gebeten, den Weblink zum Evaluator auch an Kollegen weiterzuleiten.

5.1.1 Ablauf

Im Mittelpunkt eines Evaluierungsdurchlaufs stand die Bearbeitung von fünf vorbereiteten Aufgaben, welche in diesem Abschnitt beschrieben werden. Die jeweils zur Lösung nötigen Schritte werden erläutert. Die Aufgaben wurden in fester Reihenfolge bearbeitet, um eventuelle Lernerfolge messen zu können. Die Bearbeitung fand nacheinander mit allen Techniken statt, wobei die Testpersonen den Lauf unterbrechen durften. Da sich die Aufgaben von Technik zu Technik wiederholen, können Lerneffekte auftreten, welche den Vergleich der Techniken verfälschen könnten. Die Aufgaben sind nach der ersten Technik bekannt und die Testperson kennt potentiell die Lösung, nach der zu suchen ist. Um diese Effekte auszugleichen wurde die Reihenfolge der Techniken für jeden Durchlauf zufällig generiert.

5.1.1.1 Testaufgaben

Die Testaufgaben waren im Einzelnen:

1. *Suchen Sie einen leichten, onkologischen Fall (Type = Onco) welcher mit einer Resektion behandelt (Therapy = Resection) wurde.*

Die erste Aufgabe ist eine leichte Aufgabe aus Kategorie eins (Abschnitt 3.3.3). Sie beinhaltet zwei Variablen, anhand denen ein einzelner aus einer Menge aus 46 Fällen ausgewählt werden sollte.

2. *Sie wollen in Kürze eine 83-jährige (Age = 83) Frau mit einem zentral gelegenen (Position Tumor = central), nicht rezidivierbaren Lebertumor (HCC) (Type = Onco (HCC)) mittels RF-Ablation behandeln (Therapy = Ablation(RF)). Suchen Sie zwei möglichst ähnliche onkolog. Fall zum Trainieren heraus.*

Die zweite Aufgabe beinhaltet eine gezielte Auswahl nach vier Parametern. Gesucht waren jene zwei Fälle des Fallspektrums, die in den meisten Parametern den gesuchten Werten entsprechen oder nahe kommen. Sie entspricht der Aufgabenkategorie vier und besitzt einen mittleren bis hohen Schwierigkeitsgrad. Das dargebotene Spektrum umfasst hier 86 Fälle.

3. *Wählen Sie die schwierigen (Difficulty = 3), noch nicht bearbeiteten (State = 0) Fälle, bei denen ältere Leute ab 60 Jahren nicht durch Resektion therapiert wurden.*

Aufgabe 3 zählt durch eine geforderte Negativeingrenzung des Parameters Therapy (Therapy ungleich Resection), eine Bereichsauswahl und eine mögliche Mehrfachauswahl ebenfalls zur vierten Kategorie, besitzt aber durch weniger

zu beachtende Parameter und ein kleineres Fallspektrum (31 Fälle) einen abgeschwächten Schwierigkeitsgrad.

4. *Suchen Sie für die Therapieformen Resektion und RF-Ablation den bezüglich Erkrankung (Type) und Lage des Tumors/der Metastasen seltensten Fall heraus.*

Aufgabe vier besitzt den höchsten Schwierigkeitsgrad und zählt somit zur fünften Kategorie. Hier mussten für zwei Untermengen des Parameters Therapy jeweils die Parameter Tumorposition und Typ im Zusammenhang betrachtet werden und die seltenste Kombination selektiert werden. Dabei spielte die Fähigkeit einer Technik, einen Überblick über die angebotenen Daten zu liefern, eine entscheidende Rolle. Gleichzeitig geht diese Aufgabe zu einem verhältnismäßig geringen Anteil in die Gesamtauswertung ein, da sie einem eher seltenem Nutzungsszenario (siehe Abschnitt 3.1.3) entspricht. Die Datenbank in dieser Aufgabe umfasste mit 91 Einträgen die meisten Fälle.

5. *Wählen Sie die nicht leichten Fälle vom Typ Leberlebendspende (LDLT), welche Gefäßanomalien (Vessel abnormalities = yes) aufweisen.*

Die letzte Aufgabe besitzt wieder einen sehr leichten Schwierigkeitsgrad, der dem der ersten Aufgabe entspricht. Durch einen Vergleich dieser mit der ersten Aufgabe sollte der Lernfortschritt für eine Technik aufgezeigt werden. Das Fallspektrum beinhaltete hier 41 Fälle.

5.1.1.2 Lösungsschritte

Die gestellten Testaufgaben finden ihre Entsprechung in den in Abschnitt 3.3.3 beschriebenen Kategorien und teilen sich grob in Aufgaben, welche eine gezielte Selektion erfordern und solche Aufgaben, bei denen ein Überblick über die Daten bzw. Untermengen der Daten nötig ist. Im Folgenden werden die allgemeinen Lösungswege dieser beiden Aufgabentypen erklärt. Tabelle 5.1 zeigt die für einzelne Lösungsschritte nötigen Interaktionen in den implementierten Techniken.

Gezieltes Auswählen Vier der fünf gestellten Aufgaben fallen in diese Kategorie. Die Schritte, welche zur Lösung einer solchen Aufgabe nötig waren, sind:

1. Ausfindig machen der entscheidenden Variablen
2. Optional: Anordnung der betreffenden Variablen nebeneinander für eine bessere Übersicht
3. Sortieren nach Variablen oder/und Filtern nach den interessierenden Werten

5 Durchführung, Auswertung & Ergebnisse der Evaluierung

	Tabellen	TABLELENS	INFOZOOM	PARALLELE KOORDINATEN
Finden entscheidener Variablen	Suche der entsprechenden Spalte			Suche der entsprechenden Koordinatenachse
Anordnen der Variablen	Verschieben von Spalten		Verschieben von Zeilen	-
Sortieren nach Variablen	Sortierinteraktion auf Spalten		Sortierinteraktion auf Zeilen	Achsen sind vortriert
Filtern	-	Fokussieren der entsprechenden Werte	Selektion der entsprechenden Werte und Zoom	Selektion der entsprechenden Achsenabschnitte
Abschätzung von Fallmengen	-	Orientierung an COLOR SWATCHES	Orientierung an Zellbreite	Maus über einem Achsenabschnitt hebt die Liniendurchgänge hervor

Tabelle 5.1: Gegenüberstellung der Interaktionen in den implementierten Techniken, die zur Durchführung einzelner Lösungsschritte einer Aufgabe verwendet werden können.

4. Eventuelle Wiederholung des letzten Schrittes, wenn das Fallspektrum offensichtlich zu weit eingeschränkt wurde. So gab es beispielsweise in der zweiten Aufgabe nur eine 83-jährige Frau, bei der jedoch andere Parameter nicht übereinstimmten, so dass hier eine Lockerung der Eingrenzung erwartet wurde, um die, für die in der Aufgabenstellung erwarteten zwei Lösungsfälle finden zu können.
5. Finden des/der zutreffenden Fälle in der geordneten Ansicht und ihre Auswahl

Überblick und Abschätzen von Quantitäten Zu diesem Aufgabentyp wurde eine Aufgabe gestellt. Es gibt hier Hauptvariablen zur Eingrenzung von Submengen und Nebenvariablen, welche als Kriterien zur Abschätzung der Fallmengen in den Submengen dienen. Um diese Aufgabe zu lösen, waren folgende Schritte nötig:

1. Finden der Haupt- und Nebenvariablen
2. Anordnung der betreffenden Variablen nebeneinander für eine bessere Übersicht

3. Sortieren/Filtern nach Hauptvariablen, um die interessierenden Submengen zu finden
4. Sortieren/Filtern in den Submengen nach Nebenvariablen
5. Vergleich der Fallmenge nach den Nebenvariablen und die Auswahl der treffenden Fälle
6. Bei Techniken die einen Zoom durch Ausblenden von Fällen unterstützen, muss die Eingrenzung wieder gelockert werden und der letzte Schritt für die nächste interessante Submenge wiederholt werden.

5.2 Eingegangene Daten

An der Evaluierung haben sich insgesamt 42 Testpersonen beteiligt. Aus der direkten Zielgruppe der Mediziner konnten fünf Testpersonen gewonnen werden. Zusätzlich haben 20 Personen teilgenommen, die in die Berufsgruppe Informatiker zusammengefasst werden können und 17 Personen, welche weder der Gruppe der Mediziner noch der Informatiker angehören. Die Computerkenntnisse sowie die Vorkenntnisse mit den Visualisierungstechniken sind in Abbildung 5.1 nach diesen drei Gruppen aufgeschlüsselt. Demnach sind die Testpersonen aus allen Gruppen in der Benutzung von Tabellen geübt, während die Techniken TABLELENS, INFOZOOM und PARALLELE KOORDINATEN weitestgehend unbekannt sind. Lediglich ein Teil der Informatiker gab an, PARALLELE KOORDINATEN zu kennen, wobei auch hier in keinem Fall mittlere Kenntnisse überschritten wurden.

Alle Testpersonen haben eine Präferenz für eine der Techniken angegeben, wobei aus technischen Gründen vier Votierungen nicht in die Interpretation einfließen, da hier die Evaluierung nicht mit allen Techniken durchgeführt werden konnte. Zwölf Teilnehmer haben zusätzlich eine Alternativpräferenz für eine Technik angegeben, für die sie sich entschieden hätten, wenn bestimmte Interaktionen oder Konzepte der Visualisierung auf anderem Wege umgesetzt gewesen wären. Eine nähere Auswertung der Alternativen und der Kommentare findet sich in Abschnitt 5.4.

5.2.1 Verwertbarkeit der erhobenen Daten

Aufgrund der anonymen Vorgehensweise bei der Evaluierung wurde für jeden eingegangenen Datensatz eine intensive Analyse über die Verwendbarkeit der Daten durchgeführt. Hier wurden die gemessenen Bearbeitungszeiten auf Ausreißer untersucht und im Zusammenhang mit dem entsprechenden Kommentar zu der jeweiligen Aufgabe entschieden, ob die Bearbeitung dieser in die Bewertung der Zeit und der Korrektheit einfließt.

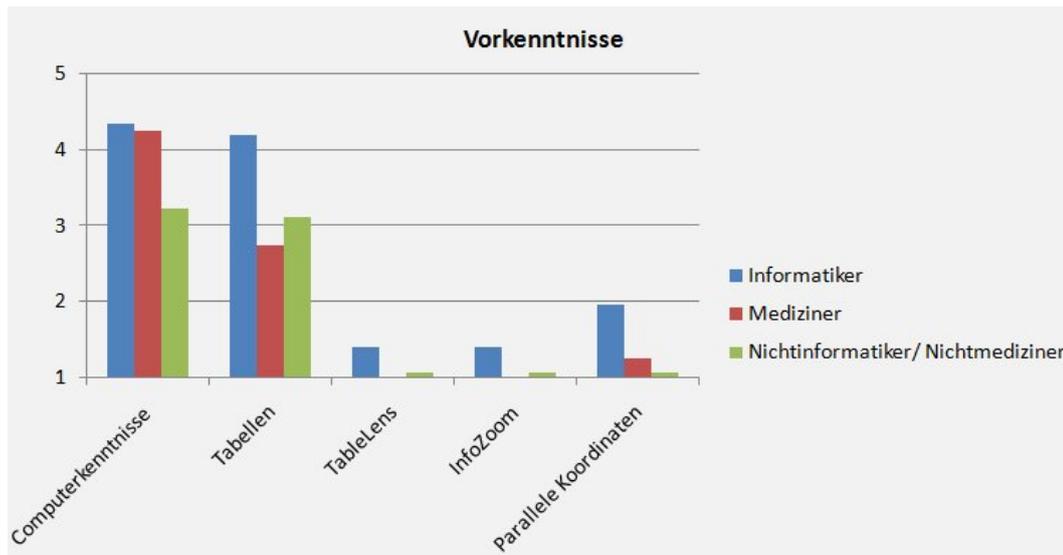


Abbildung 5.1: Vorkenntnisse der Testpersonen am Computer sowie mit den einzelnen Visualisierungstechniken auf einer Skala von 1 bis 5 (Durchschnitt)

Sind bei der Bearbeitung technische Probleme mit einer Technik aufgetreten, so wurden die Messergebnisse für die Bearbeitungszeiten nicht mit ausgewertet. Auch die Votierung der Präferenz für eine Technik wurde in diesem Fall nicht gewertet, da der Nutzer keinen objektiven Vergleich ziehen konnte. Die aufgezeichnete subjektive Einschätzung der einzelnen Aspekte der betroffenen Technik wurden dann ebenfalls nicht gewertet, da auch hier von einem verzerrten Eindruck ausgegangen werden kann.

War erkenntlich, dass eine Aufgabe nicht verstanden wurde, wurden sowohl Zeit- als auch Korrektheitsmessungen nicht berücksichtigt, da hier davon ausgegangen werden muss, dass die Aufgabe nicht mit dem erwarteten Ziel bearbeitet wurde. Indikatoren für diesen Fall waren:

- Falsche Ergebnisse der Aufgabe mit allen Techniken
- Kommentar, der deutlich ein Unverständnis ausdrückte
- Bei teilweise richtiger Auswahl wurde bei extrem kurzer Bearbeitungszeit davon ausgegangen, dass hier zufällig ein zur Lösung gehörender Fall ausgewählt war, aber keine Bearbeitung stattfand. Wenn im Vergleich mit allen Bearbeitungen dieser Aufgabe mit einer bestimmten Technik extrem lange Bearbeitungszeit auftrat, wurde auch dabei von Unsicherheit über die Aufgabenstellung ausgegangen.

Auch die Bewertung der Techniken wurde bei Unverständnis der Aufgabe nicht mit in die statistische Auswertung aufgenommen. Da keine der Aufgabenstellungen bei allen Probanden unklar war, wurden diese Daten für die interpretative Auswertung getrennt betrachtet. In einigen Fällen ließen auch diese gesondert betrachteten Wertungen wertvolle Schlüsse zu, z. B. wenn in Kommentaren vermerkt war, wie die Aufgabe verstanden wurde, sie mit allen Techniken bearbeitet wurde und eine Technik als besonders geeignet bewertet wurde.

Die Korrektheit der gewählten Fälle wurde nicht gewertet sofern ein Unverständnis der Technikinteraktionen vorlag. Dies war der Fall wenn positive Wertungen für die Eignung nach der Bearbeitung abgegeben wurden aber keinerlei Fälle ausgewählt wurden. Vorgekommnisse dieser Art gab es in einigen Fällen bei der TABLELENS und in einem Fall beim INFOZOOM, da hier die Selektion in der Tabellensicht zur speziellen Steuerung der Technik dient und nicht der eigentlichen Auswahl.

Besonders häufig konnten die Ergebnisse der vierten Aufgabe nicht gewertet werden. Hier wurde in etwa der Hälfte aller Evaluierungsdurchläufe die Aufgabe nicht richtig interpretiert. Die Auswertungen für diese Aufgabe wurde deshalb, sofern sie gewertet wurden, immer im Zusammenspiel von Bearbeitungszeit, Korrektheit der Antworten (bei den gewerteten Durchgängen waren häufig teilweise richtige Antworten gegeben worden) und den abgegebenen Wertungen und Kommentaren betrachtet.

5.3 Auswertung der gemessenen Daten

Bei der Evaluierung wurden für jede einzelne Aufgabenbearbeitung die benötigte Zeit und die gegebenen Antworten aufgezeichnet. Diese Daten wurden nach der in Abschnitt 5.2.1 beschriebenen Verwendbarkeitsanalyse zur weiteren Auswertung in ein Kalkulationsprogramm übertragen. Im folgenden Abschnitt werden die gewonnenen Daten näher erläutert.

5.3.1 Untersuchung der gemessenen Zeiten

Für alle drei Berufsgruppen wurden die jeweils verwertbaren Bearbeitungszeiten pro Aufgabe zusammengefasst und der Mittelwert sowie die Standardabweichung berechnet. Extrem hohe Antwortzeiten wurden im Zuge der Verwendbarkeitsanalyse ausgeschlossen. Da sich die Mittelwerte und auch die Standardabweichungen der einzelnen Personengruppen kaum unterschieden, sind Mittelwerte und Standardabweichungen nicht nach den Gruppen getrennt aufgeführt.

5.3.1.1 Bearbeitungszeiten

Die mittleren Bearbeitungszeiten sind in Abbildung 5.2 dargestellt. Es lassen sich hier mehrere interessante Beobachtungen machen.

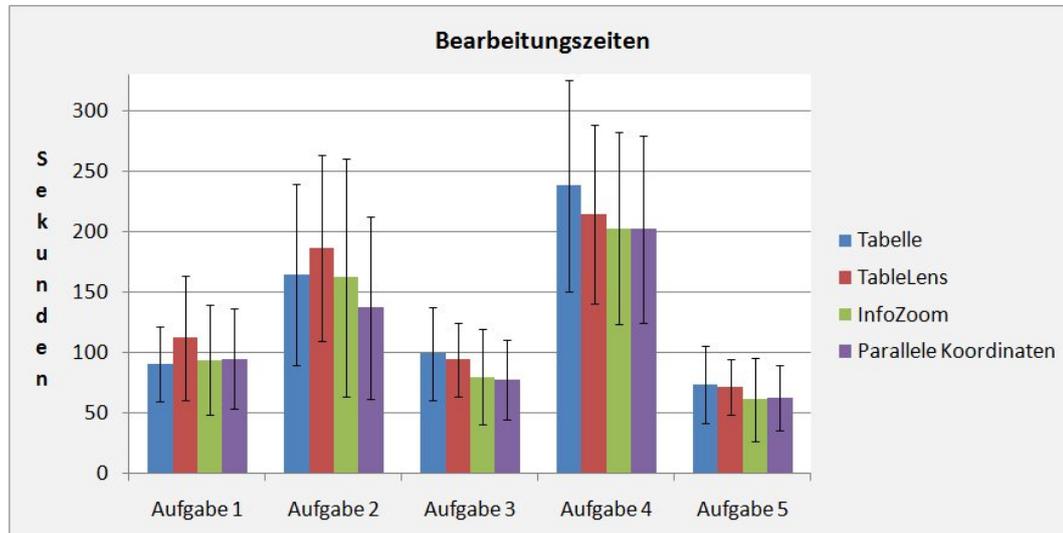


Abbildung 5.2: Durchschnittliche Bearbeitungszeit pro Aufgabe und Visualisierung

Zum Ersten fällt auf, dass bei Aufgabe 1 nur die TABLELENS mit 110 Sekunden eine wesentlich höhere mittlere Bearbeitungszeit aufweist, während sich die anderen Techniken bei ca. 90 Sekunden ansiedeln. Dies kann entweder ein Indiz für einen höheren Eingewöhnungsaufwand oder eine insgesamt schlechtere Leistung mit dieser Technik sein. Da sich die Bearbeitungszeit der TABLELENS ab der Bearbeitung der dritten Aufgabe auf bzw. leicht unter dem Niveau der Tabellen einordnet, wird hier auf eine erhöhte Eingewöhnungszeit geschlossen.

Eine zweite Auffälligkeit besteht in der sehr viel niedrigeren Bearbeitungszeit der zweiten Aufgabe mit den parallelen Koordinaten, im Schnitt 20 Sekunden weniger als die zweitschnellste Visualisierungstechnik, der INFOZOOM. Dies lässt vermuten, dass sich PARALLELE KOORDINATEN besser als die übrigen Techniken für Auswahlaufgaben mit sehr vielen Kriterien eignen.

Ab Aufgabe 3 liegen die Bearbeitungszeiten des INFOZOOM und der PARALLELEN KOORDINATEN, beides Techniken, die für eine erhöhte Übersichtlichkeit Datensätze ausblenden, unter denen der Tabellen und der TABLELENS. Daraus lässt sich schließen, dass Techniken, welche ein Filtern von Datensätzen erlauben, in einer Fallauswahl performanter als Techniken sind, die keinerlei Einschränkungen zulassen (Tabellen) oder bei denen eingeschränkte Daten trotzdem sichtbar bleiben (TABLELENS).

Die sehr viel höhere Bearbeitungszeit der Tabellen bei Aufgabe 4 weist darauf hin, dass Aufgaben mit einem solch hohen Schwierigkeitsgrad ohne visuelle oder interaktive Hilfsmittel, wie sie die anderen Techniken bieten, ungleich schwerer zu bewältigen sind. Es muss hierbei allerdings berücksichtigt werden, dass die Bearbeitung dieser Aufgabe sehr komplex war und hier die wenigsten Ergebnisse verwendet werden konnten. In ca. 50% der Testdurchläufe konnten die Antworten und die Zeiten hier nicht gewertet werden. So können die Bearbeitungszeiten bei dieser Aufgabe nur in Zusammenhang mit der Korrektheit der Antworten und der Gesamtheit der gewerteten Durchgänge betrachtet werden. So weisen auch hier die Techniken einen Zeitvorteil auf, welche ein Filtern erlauben. Jedoch wurde mit diesen Techniken auch gerade bei den Medizinerinnen und Nichtinformatikerinnen die Bearbeitung am häufigsten abgebrochen, während sie mit den Tabellen und der TABLELENS häufiger zum Ende geführt wurde. Dabei weist die TABLELENS die wenigsten falschen Antworten auf (siehe Abschnitt 5.3.2 und Abbildung 5.5).

5.3.1.2 Erlernen der Techniken

Nicht nur die Zeitwerte, auch die Entwicklung der zur Lösung benötigten Zeit von Aufgabe zu Aufgabe birgt interessante Beobachtungen. Das Diagramm in Abbildung 5.3 zeigt den Verlauf der Bearbeitungszeiten von Aufgabe 1 über Aufgabe 3 zu Aufgabe 5. Dies waren die Problemstellungen mit den geringsten Schwierigkeitsgraden. Hieran kann der Lernerfolg im Umgang mit den Techniken gemessen werden.

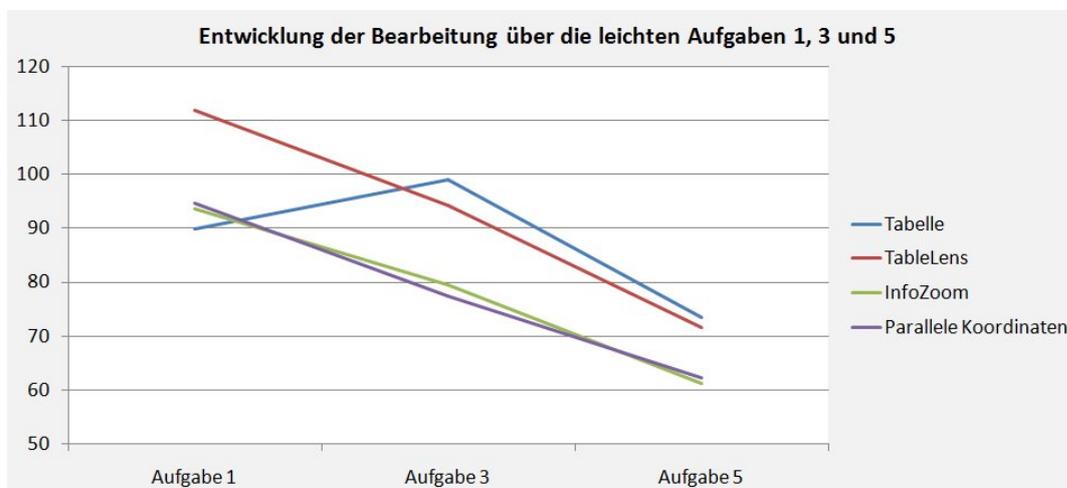


Abbildung 5.3: Verlauf der Bearbeitungszeiten über die drei einfachsten Aufgaben

Interessant ist, dass der Umgang mit INFOZOOM und PARALLELEN KOORDINATEN im Vergleich zur Tabelle kaum mehr Zeit benötigt. Es kann also davon ausgegangen

werden, das hier nur eine extrem kurze Gewöhnungsphase benötigt wird, welche ungefähr die Zeitspanne ausmacht, die durch den Einsatz dieser Techniken im Gegensatz zur Tabelle gespart werden kann.

Im weiteren Verlauf der Kurven bestätigt sich, dass die „unbekannten“ Visualisierungstechniken in etwa gleich schnell zu erlernen sind. Dies erklärt sich durch die im Vergleich zur Tabelle stärker fallenden Kurven für TABLELENS, INFOZOOM und den PARALLELEN KOORDINATEN, die näherungsweise den gleichen Anstieg besitzen. So wird der erwähnte Vorteil der Techniken mit Filtermöglichkeit ab Aufgabe 3 sichtbar, die anfangs erhöhte Bearbeitungszeit der TABLELENS gleicht der der Tabelle.

Die Visualisierung der Tabellen war allen Teilnehmern bekannt, so dass hier erwartungsgemäß kein messbarer Lernerfolg eintritt. Das auch hier dennoch ein Unterschied zu verzeichnen ist, kann mit der Gewöhnung an die Mehrfachsortierung und an das Themengebiet der Aufgabenstellungen erklärt werden.

5.3.2 Korrektheit der gewählten Fälle

Zusätzlich zu den Zeiten, welche die Testpersonen zum Lösen der Problemstellungen benötigten, wurden die letztlich gewählten Fälle pro Aufgabe und Technik für jeden Evaluierungsdurchlauf aufgezeichnet. In Abbildung 5.4 ist die prozentuale Verteilung der richtigen, teilweise richtigen und falschen Antworten für alle Techniken über alle Aufgaben aufgezeigt. Eine Antwort wurde als teilweise richtig angesehen, wenn die Antwort gewertet werden konnte (Abschnitt 5.2.1) und ein Teil der Fälle aus der Lösungsmenge ausgewählt war. Bearbeitungen, welche aufgrund extremer Nichteignung einer Technik abgebrochen wurden, wurden von der Wertung ausgeschlossen. Eine Interpretation dieser Vorkommnisse findet sich im nächsten Abschnitt.

Es lassen sich in dem Diagramm in Abbildung 5.4 verschiedene Beobachtungen machen. Zunächst fällt auf, dass mit 77% mit dem INFOZOOM im Durchschnitt etwa 6% mehr richtige Antworten gegeben wurden als mit allen anderen Techniken. Gleichzeitig weist diese Visualisierung mit 7% die geringste Fehlerquote auf.

Während die Fehlerquoten bei den Tabellen und den PARALLELEN KOORDINATEN bei etwa 10% liegen, treten bei der TABLELENS mit 15% die meisten falschen Antworten auf. Gleichzeitig sind hier mit 15% weniger teilweise richtigen Antworten aufgetreten, als bei den restlichen Techniken. Diese beiden Beobachtungen lassen im Zusammenhang vermuten, dass ein Nutzer, welcher den falschen Weg bei der Lösung einer Problemstellung eingeschlagen hat, durch die Visualisierung der TABLELENS endgültig fehlgeleitet wird. Eine genauere Analyse hat gezeigt, dass der Großteil des aufgezeigten Unterschieds bei Aufgabe 2 auftrat (siehe Abbildung 5.5). Dies lässt darauf schließen, dass die TABLELENS bei Aufgaben mit sehr vielen Auswahlkriterien in einer großen Falldatenbank weniger geeignet ist. Weiterhin ist zu beobachten,

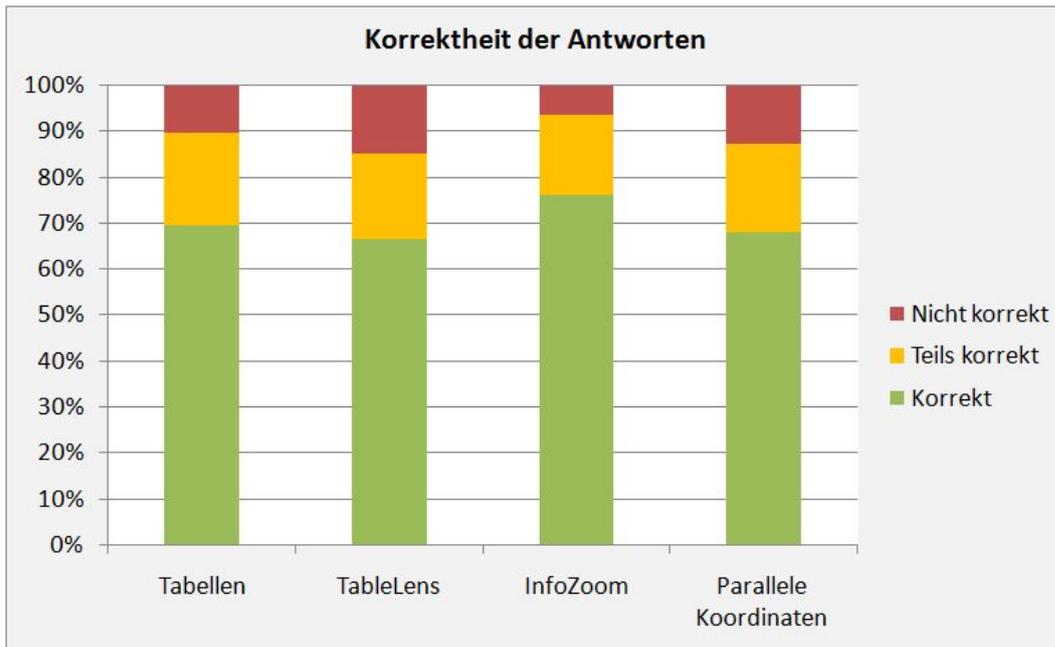


Abbildung 5.4: Verteilung der Korrektheit der Antworten für alle Techniken

dass die Korrektheit der Antworten bei den PARALLELEN KOORDINATEN geringer ist als bei den Tabellen. Dies könnte am ungewöhnlichen Visualisierungsansatz liegen, der exakte Grund ließ sich aber anhand der erhobenen Daten nicht bestimmen.

Für die vierte und schwerste Aufgabe sind geringe Vorteile für die rein tabellari-schen Ansätze der Tabelle und der TABLELENS zu verzeichnen, wobei hier mit der TABLELENS durch die geringste Fehlerquote und die häufigsten richtigen Antworten noch einmal geringfügig bessere Ergebnisse erzielt wurden (Abbildung 5.5). Eine Erklärung dafür könnte die regelmäßige Gitterstruktur sein. Dabei kann die TABLELENS im Gegensatz zur Tabelle durch ihre Unterscheidung von Fokus und Kontext mehr Zellen auf gleicher Fläche darstellen, die sich aufgrund der visuellen Hilfsmittel der COLOR SWATCHES zudem besser unterscheiden lassen. Durch die hohe Zahl der nicht verwendbaren Messungen für diese Aufgabe können die Ergebnisse für die vierte Aufgabe aber auch rein zufällig entstanden sein.

5.3.2.1 Zusatzaufgaben

Zusätzlich zu den eigentlichen Problemstellungen wurde nach jeder Bearbeitung der vierten Aufgabe im Bewertungsbildschirm eine Zusatzfrage gestellt, welche über den

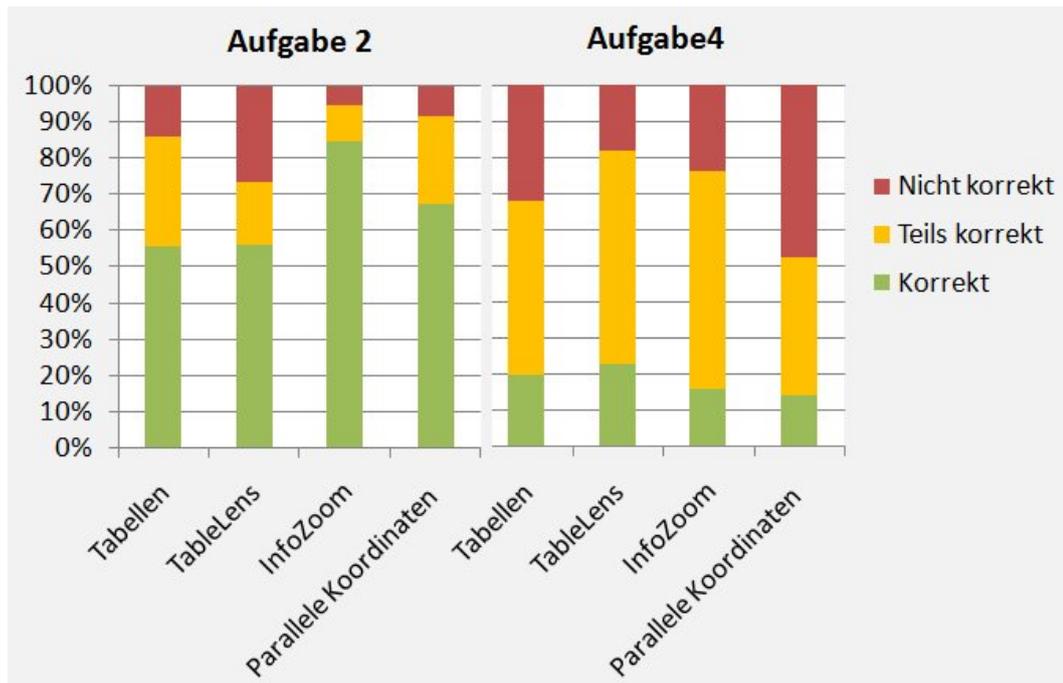


Abbildung 5.5: Verteilung der Korrektheit der Antworten für die Aufgaben 2 und 4.

vermittelten Überblick über die präsentierten Daten weiteren Aufschluss geben sollte. Aus technischen Gründen konnten hier nur 22 der 42 Evaluierungsdurchgänge betrachtet werden. Im Ergebnis wurden zu diesen Fragen nur sehr wenige Antworten gegeben. Dies kann mehrere Ursachen haben. Zum Einen wurden diese Fragen nach Aufgabe 4 präsentiert, bei welcher die Bearbeitung oft abgebrochen wurde. Zum anderen wurden sie nur sehr selten überhaupt beantwortet¹. Wenn sie beantwortet wurden, lautete die Antwort zumeist „Weiß nicht“. Hier kann aber aufgrund der allgemein eingeschränkten Teilnahme an dieser zusätzlichen Überprüfung nicht unterschieden werden ob dies daran lag, dass trotz diverser Hinweise im Evaluator keine weitere Untersuchung der Datensätze unternommen wurde oder ob wirklich die verwendete Technik keinen genügenden Eindruck der Daten vermitteln konnte. Bei den abgegebenen Antworten ließ sich trotzdem ein Trend feststellen, nachdem die TABLELENS den besten Eindruck vermitteln konnte. Hier wurden 17 Antworten gegeben, von denen 7 „Weiß nicht“ lauteten. Acht Testpersonen gaben bei dieser Technik die richtige Antwort. Dies würde die restlichen Beobachtungen bestätigen, bei denen die TABLELENS als besonders geeignet erscheint, einen Überblick zu vermitteln. Bei den übrigen Techniken wurden weniger bzw. in der Mehrzahl falsche Antworten gegeben. Dieser Trend kann aber auch rein zufällig entstanden sein. Die

¹Es konnte jeweils mit „Ja“, „Nein“ und „Weiß nicht“ geantwortet werden, nicht gegebene Antworten wurden mit der Implementierung extra für die Auswertung kenntlich gemacht.

selteneren Antworten auf diese Fragen und auch die wenigen verwertbaren Messungen von Aufgabe 4 lassen vermuten, dass eine Evaluierung, die den Überblick über die Daten bei verschiedenen Techniken prüfen möchte, gesondert durchgeführt werden sollte.

5.4 Auswertung der subjektiven Bewertung der Techniken

Neben den objektiv messbaren Kriterien wurde den Testpersonen nach jeder Bearbeitung einer Problemstellung die Möglichkeit geboten, die Eignung der genutzten Technik für die bearbeitete Aufgabe anzugeben. Dies konnte über eine Bewertungsskala geschehen, zusätzlich gab es jeweils ein Kommentarfeld für ergänzende Anmerkungen. Am Ende eines Evaluierungsdurchgangs wurden noch einmal Skalen für die Gesamtbewertung der Techniken angeboten. Hier sollte einzeln nach den Aspekten Interaktionsmöglichkeiten, Vermitteln eines Überblickes über die Daten und Effizienz bewertet werden. Die Auswertung dieser Daten steht im Mittelpunkt dieses Abschnittes.

5.4.1 Eignung der Techniken für die Aufgaben

Die subjektive Einschätzung der Eignung der eingesetzten Technik pro Aufgabe stellt ein wichtiges Maß für die wahrgenommene erbrachte Leistung mit einer Technik dar. So kann es vorkommen, dass eine Testperson den Umgang mit einer Technik als hervorragend einstuft, während die objektiv gemessene Bearbeitungszeit mit einer anderen Technik deutlich niedriger lag. Andererseits ist es möglich, dass sich ein Nutzer nach der Bearbeitung mit einer Technik über die Ergebnisse nicht sicher ist, tatsächlich aber mit nur mit dieser Technik eine korrekte Antwort gegeben wurde.

Durch die Auswertung dieser direkten Einschätzungen lassen sich wichtige Rückschlüsse auf die Akzeptanz der einzelnen Visualisierungen ziehen. Zu diesem Zweck sind in Abbildung 5.6 die durchschnittlich abgegebenen Werte für jede Aufgabe mit jeder Technik dargestellt.

Auffällig ist der enorme Vorsprung der Visualisierung des INFOZOOM bei Aufgabe 1. Eine Analyse der Kommentare zu dieser Aufgabe in Bezug auf dieses Phänomen ergab, dass viele Nutzer überrascht waren, wie leicht die Bearbeitung mit dieser Technik auch bei der Erstnutzung schon fällt. Einige hatten den Eindruck durch die Zoomfunktion von der Technik zum Ziel geleitet zu werden.

Ein ähnlicher Effekt konnte für die PARALLELEN KOORDINATEN festgestellt werden. Dieser wird allerdings erst ab Aufgabe 2 offensichtlich. Ein großer Teil der Testpersonen konnte erst hier das Darstellungsprinzip dieser Visualisierung endgültig

5 Durchführung, Auswertung & Ergebnisse der Evaluierung

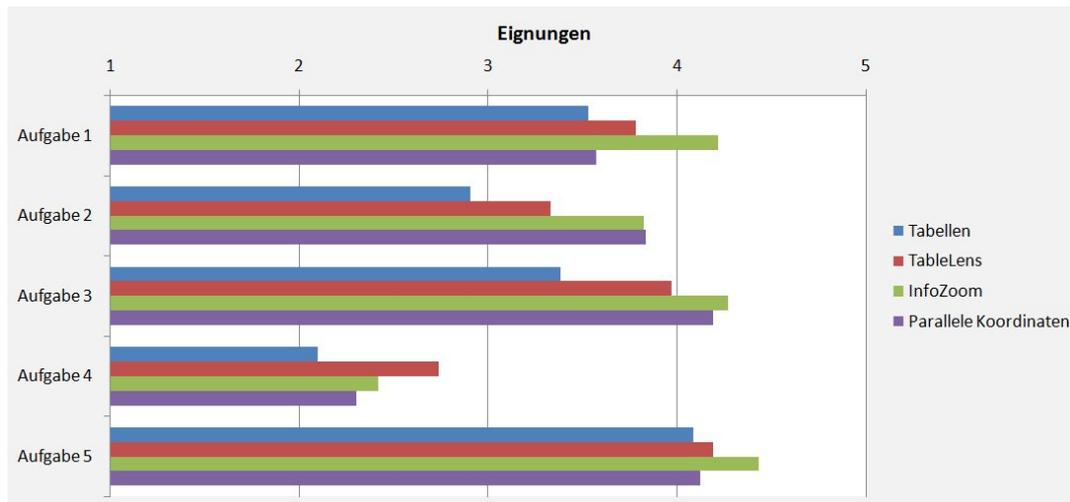


Abbildung 5.6: Durchschnittliche Wertung für die Eignung einer Technik für eine Aufgabe

verstehen, war dann aber genau wie beim INFOZOOM erstaunt über die effiziente Fallauswahl.

Die TABLELENS wurde für jede Aufgabe besser als die Tabellen bewertet, schien aber mit Ausnahme von Aufgabe 4 nicht so geeignet wie diejenigen Techniken, welche uninteressante Fälle komplett ausblenden. Dies bestätigt die in Abschnitt 5.3.1.1 getätigte Annahme, dass solche Techniken für eine gezielte Fallauswahl anhand festgelegter Kriterien besser geeignet sind. In Aufgabe 4 wurde die TABLELENS als am besten geeignet empfunden. Dies untermauert den in Abschnitt 5.3.2 gezogenen Schluss, dass diese Technik sich besonders für Aufgaben eignet, welche einen Überblick über die Daten und ein Abschätzen von Quantitäten erfordern.

Insgesamt wurden die Visualisierungen der TABLELENS, des INFOZOOMS und der PARALLELEN KOORDINATEN durchschnittlich bei allen Aufgaben als geeigneter als die Tabellen eingeschätzt, was ein wichtiges Indiz für eine breite Akzeptanz dieser Techniken ist. Auch bei dieser Auswertung war zwischen den Berufsgruppen kein bemerkenswerter Unterschied feststellen.

5.4.2 Bewertung der Techniken nach verschiedenen Aspekten

Das Diagramm in Abbildung 5.7 zeigt die durchschnittlichen Bewertungen der Interaktionsmöglichkeiten, des vermittelten Überblickes über die Daten und der empfundenen Effizienz der Techniken durch die Testpersonen. In den Bewertungen der

5.4 Auswertung der subjektiven Bewertung der Techniken

verschiedenen Berufsgruppen gab es bei den Tabellen und den PARALLELEN KOORDINATEN einen bemerkenswerten Unterschied. So wurde der Überblick der Tabellen bei den Informatikern im Mittel als „weniger gut“ (2 Punkte), bei Medizinern und sonstigen Nichtinformatikern als „mittel“ (3 Punkte) eingestuft. Bei den PARALLELEN KOORDINATEN ist dieses Verhältnis umgekehrt. Eine Erklärung hierfür lässt sich in den höheren Vorkenntnissen mit den parallelen Koordinaten bei den Informatikern finden.

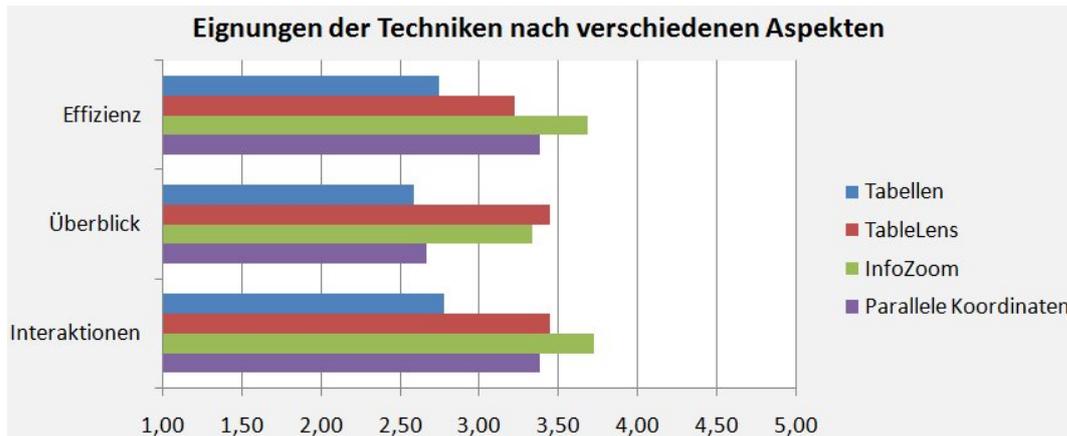


Abbildung 5.7: Durchschnittliche Bewertung der Visualisierungen nach den Aspekten Interaktion, Überblick und Effizienz

Das Gesamtergebnis bestätigt hier ein weiteres Mal die in vorherigen Auswertungen gewonnenen Kenntnisse. So erhält die TABLELENS für den Überblick die beste Wertung, während der INFOZOOM besonders bei der Effizienz einen deutlichen Vorsprung aufweisen kann. Der Vorsprung der PARALLELEN KOORDINATEN fällt hier weniger deutlich aus, da insbesondere Testpersonen aus der Berufsgruppe der Mediziner, aber auch einige andere Nichtinformatiker mit dieser Technik unüberwindbare Schwierigkeiten hatten, was zu einer Verschlechterung der Gesamtwertung führte. Die Interaktionen wurden bei den erweiterten Techniken zwischen „mittel“ und „gut“ bewertet, bei der Tabelle wurde häufig eine Eingabemaske zum Filtern von Einträgen vermisst und das nötige Scrollen sehr häufig als sehr aufwendig betrachtet. Am besten wurde hier insgesamt der INFOZOOM bewertet.

5.4.3 Kommentare und Verbesserungsvorschläge

Die Testpersonen hatten nach jeder Bearbeitung und zum Ende die Möglichkeit, zusätzlich zur skalenmäßigen Erfassung ihrer Einschätzung auch freie Kommentare abzugeben. Sie wurden hier gebeten, auf die Aufgabenstellung und die Techniken einzugehen. So sollten zum Einen ein eventuelles Unverständnis der Aufgaben und

zum Anderen Verbesserungsvorschläge und Handicaps der Techniken aufgenommen werden. Im Folgenden werden die häufigsten Aussagen zu den einzelnen Techniken zusammengefasst.

5.4.3.1 Tabellen

Bei den Tabellen wurde bei allen Berufsgruppen eine mangelnde Übersicht bemerkt. Da keinerlei visuelle Hilfsmittel geboten werden, müssen Ergebnisse mit steigender Kriterienanzahl sehr häufig gegengeprüft werden. Sehr häufig vermisst wurde auch eine Filtermaske, durch die sich Fälle ausblenden lassen. Dies als Alternative zu den Vorteilen der übrigen Techniken hätte in zwei Fällen zu einer endgültigen Entscheidung für die Tabellen geführt. Das oft nötige horizontale und vertikale Scrollen wurde als extrem störend empfunden, hier wurde öfter die Möglichkeit gewünscht, Spalten ausblenden zu können. Positiv bemerkt wurde der logische, bekannte Aufbau. Man könne ohne Weiteres in die Bearbeitung einsteigen. Die angebotenen Interaktionen wurden zwar als „mittel“ bewertet, wurden aber für die dargebotenen Aufgaben als nicht ausreichend empfunden. Der Vorteil der übrigen Techniken, der durch mögliche Unterscheidung zwischen uninteressanten und interessanten Bereichen der Datenbank entsteht, ist in allen Berufsgruppen deutlich geworden.

5.4.3.2 TableLens

Bei der TABLELENS gab es geteilte Meinungen. Einerseits wurde sie durch die große Menge der gezeigten Daten als sehr übersichtlich empfunden, was auch die höchste Wertung für den Überblick erklärt. Andererseits wurde es als störend empfunden, immer erst Daten markieren zu müssen, um Details sehen zu können. So konnten auch Spalten nicht manuell vergrößert werden. Weiterhin störten bei konkreten Auswahlen das Verbleiben der minimierten Zellen, hier hätten einige Nutzer die nicht benötigten Fälle vorzugsweise vollständig ausgeblendet. Die ColorSwatches wurden zwar als sehr nützlich bewertet, aber durch eine fehlende Legende als nicht sehr intuitiv eingestuft. Hier wurde eine Legende im Spaltenkopftooltip vorgeschlagen. Auch zu den Daten passende Farben hätten einigen Nutzern geholfen. Als Beispiel wurde hier der Farbverlauf von grün nach rot für den Schwierigkeitsgrad erwähnt. Bei dieser Technik wurden die meisten Kommentare zur Bedienung abgegeben, was etwas im Gegensatz zur durchschnittlichen Bewertung für die Interaktionen steht (siehe Abschnitt 5.4.2), in der die Interaktionen als überdurchschnittlich eingestuft wurden. Dies könnte auch eine Erklärung dafür sein, dass diese Technik eher als Alternative zu einer der anderen Techniken empfohlen wurde.

5.4.3.3 InfoZoom

Über den INFOZOOM gab es im Vergleich eine große Überzahl an positiven Kommentaren. Die Zoomfunktion wurde sehr häufig als sehr gut zu bedienen und sehr nützlich eingestuft. Die Probanden waren sich häufig bezüglich der Korrektheit ihrer gemachten Auswahlen sehr sicher. Das Arbeiten mit der Technik wurde als effizient und angenehm empfunden. Einige Nutzer waren am Anfang aufgrund des Tauschs von Zeilen und Spalten verwirrt. Auch wurde eine Hervorhebung von Fällen bzw. Spalten gewünscht, die das Zuordnen von Daten zu ihren Fällen trotz zusammengefasster Zellen erleichtert.

5.4.3.4 Parallele Koordinaten

Die PARALLELEN KOORDINATEN wurden von den meisten Testpersonen ebenso positiv bewertet wie der INFOZOOM. So waren die Nutzer auch hier nach anfänglicher Verwirrung über die Effizienz der Technik überrascht, welche sich bei dieser Technik ebenfalls an der Fähigkeit, Fälle auszublenden, festmachen lässt. Die Nutzer waren sicherer über ihre Entscheidungen und das Arbeiten wurde als sehr schnell empfunden, wobei objektiv mit dieser Technik die wenigsten richtigen Antworten gemessen wurden. Zusätzlich gab es bei dieser Technik mehrere Testpersonen, welche mit dem Visualisierungsansatz gar nicht arbeiten konnten und Bearbeitungen abgebrochen haben. Direkte Gründe hierfür ließen sich nicht bestimmen, viele der Probanden empfanden die Technik allgemein als zu befremdlich und zu verwirrend. Die Interaktionen wurden als verbesserungswürdig eingestuft, insbesondere die Auswahl von Achsenabschnitten und Falllinien. Auch war es bei einem großen Fallspektrum durch Überzeichnungen schwierig, die MouseOver-Informationen für eine bestimmte Linie abzurufen, da die Trefferzonen zu klein waren. Zudem wurde die Beschriftung der Achsen als weniger gut bewertet, wobei auch dafür keine Verbesserungsvorschläge gemacht wurden.

5.4.4 Empfehlung für eine Technik

Jede Testperson hatte am Ende des Evaluierungslaufes die Möglichkeit, eine Antwort auf die zentrale Frage dieser Arbeit zu geben und eine Empfehlung einer Technik für den Einsatz in einem medizinischen Trainingssystem abzugeben. In Abbildung 5.8 ist das Ergebnis aufgezeigt.

Vier Votierungen konnten nicht berücksichtigt werden, da hier aus technischen Gründen nicht mit allen Techniken gearbeitet werden konnte. Hierbei wurde zweimal der INFOZOOM gewählt, einmal ohne Bearbeitung mit der TABLENS und einmal ohne Bearbeitung mit den PARALLELEN KOORDINATEN. Einmal wurden die PARALLELEN KOORDINATEN gewählt, ohne mit den Tabellen gearbeitet zu haben. Im vierten

5 Durchführung, Auswertung & Ergebnisse der Evaluierung

Fall wurde für die TABLELENS votiert, ohne dass die Möglichkeit bestand, mit den PARALLELEN KOORDINATEN zu arbeiten.

Anhand der Grafik zeigt sich eine Präferenz für die Visualisierung des INFOZOOM. Bei der TABLELENS und den PARALLELEN KOORDINATEN würden einige Änderungen an der Technik in mehreren Fällen zur Auswahl dieser führen, hier wurde der Umgang mit der Technik als zu umständlich empfunden.

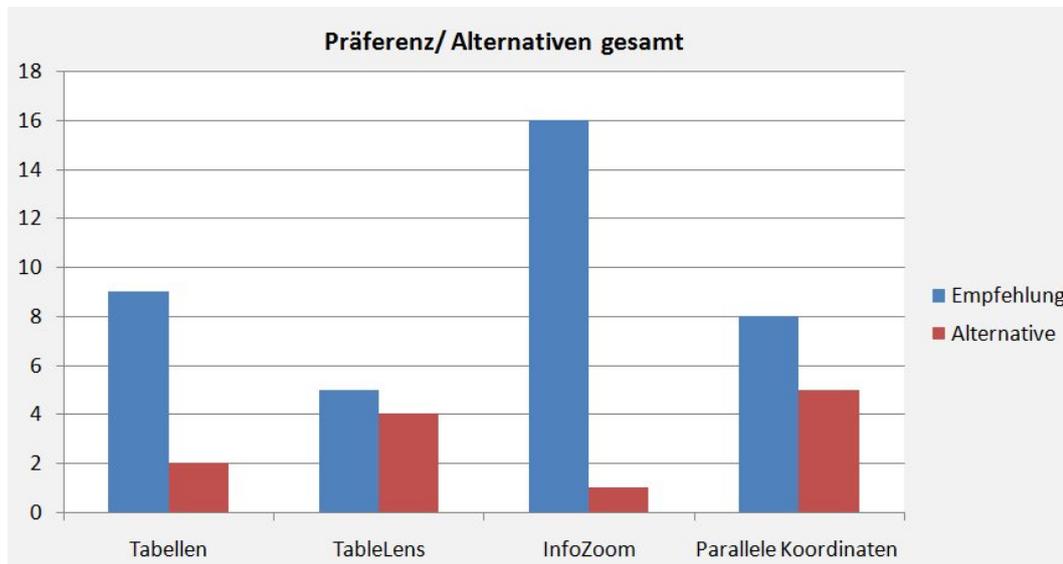


Abbildung 5.8: Votierungen der Testpersonen für die Techniken

Insgesamt setzt sich in dieser Auswertung in ähnlicher Weise der Trend fort, welcher bereits bei Auswertung der Bearbeitungszeiten (siehe Abschnitt 5.3.1.1) sowie den Eignungen (siehe Abschnitt 5.4.1) zu bemerken war. Auch bei der endgültigen Empfehlung setzte sich mit dem INFOZOOM eine derjenigen Techniken durch, welche Fälle ausblenden können. Eben dieses Filtern scheint der entscheidende Vorteil dieser Techniken zu sein.

Zudem lässt sich bei dieser Auswertung ein großer Unterschied zwischen den Berufsgruppen ausmachen. Abbildung 5.9 zeigt die prozentuale Aufteilung der Votierungen in den verschiedenen Gruppen. Während die Gruppe der Informatiker eindeutig, auch anhand der abschließenden Kommentare, den INFOZOOM bevorzugt, ist dies bei den Medizinerinnen nicht der Fall. Drei von fünf Testpersonen favorisierten die klassische Tabelle, auch mit besonderem Verweis auf technisch nicht so versierte Kollegen, die bei den „neuen“ Techniken schnell frustriert wären. Weiterhin war in dieser Berufsgruppe die Akzeptanz der erweiterten Techniken geringer. Die integrierte „Such- bzw. Filtermaske“, welche häufig von den Probanden in Kommentaren gewünscht wurde, wurde hier eher als verwirrend und den Überblick verringernd angesehen. Dieser Trend lässt sich auch bei den Nichtinformatikern fortsetzen. Hier gab es zwar

5.4 Auswertung der subjektiven Bewertung der Techniken

ebenfalls eine eindeutige Präferenz für den INFOZOOM, jedoch wurde auch häufig die Tabelle gewählt. Zusätzlich wären aus dieser Berufsgruppe zwei weitere Votierungen für die Tabelle abgegeben worden, wenn diese eine Suchmaske bieten würde.

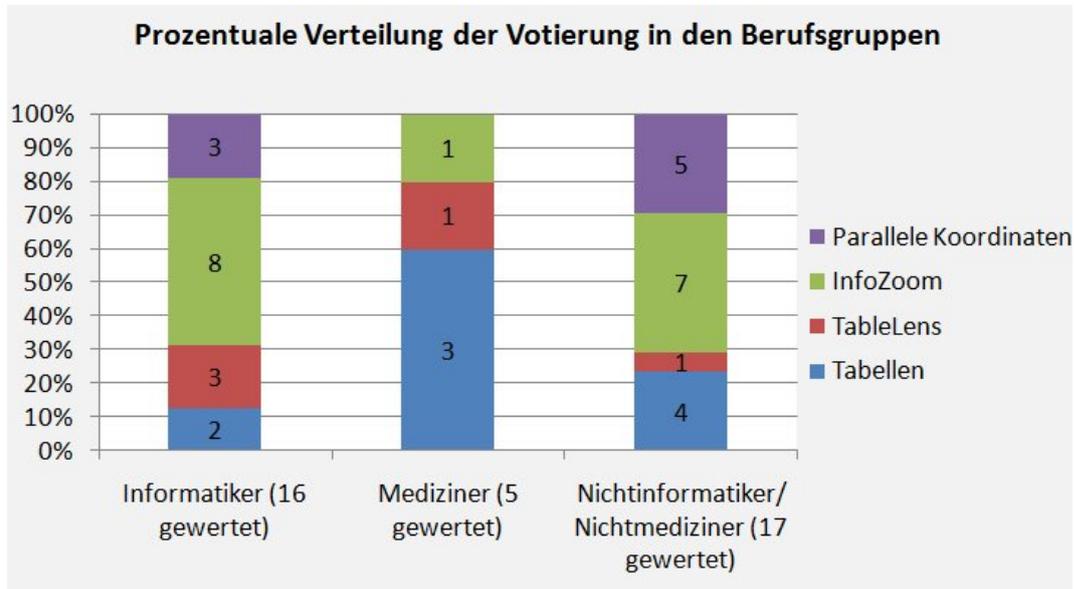


Abbildung 5.9: Prozentuale Verteilung der Votierungen in den Berufsgruppen

Auffällig ist, dass mehr Personen die Tabellen als die TABLELENS bevorzugten. Auch mit eingerechneten Alternativvotierungen würde dieses Ergebnis bestehen. Die TABLELENS wurde oft als zu wenig intuitiv eingeschätzt. In Abschnitt 5.3.1.1 wurde auf eine erhöhte Eingewöhnungszeit für diese Technik geschlossen. Es ist zu vermuten, dass diese Technik bei einer Evaluierung, in der die Testpersonen sich selbstständig mit den Techniken vertraut machen können, ein besseres Ergebnis erzielt hätte. Allerdings rechtfertigt der Stellenwert der Fallauswahl keine Technik, bei der eine längere Einarbeitung nötig ist (siehe Abschnitt 3.1.4).

5.4.4.1 Statistische Überprüfung

Die angenommene gleiche Chance für jede Technik gewählt zu werden, steht im Gegensatz zur Beobachtung, welche einen eindeutigen Vorteil für den INFOZOOM aufzeigt. Statistisch wird hier von praktischer Relevanz gesprochen [8]. Durch einen χ^2 -Test kann die angenommene Gleichverteilung überprüft werden, welche als Nullhypothese für den Test dient. Der Chi-Quadrat-Test folgt der Formel:

$$\chi^2 = \sum_{i=1}^m \frac{(f_{io} - f_{ie})^2}{f_{ie}}$$

Hierbei stehen f_{ie} und f_{io} für die erwartete und beobachteten Häufigkeiten der Votierungen für eine Technik. Es werden bei diesem Test die Abweichungen der beobachteten von der angenommenen Häufigkeit berechnet, quadriert und durch die angenommene Häufigkeit einer Votierung geteilt. Dieser Wert wird über die Alternativen, hier die verschiedenen Visualisierungstechniken, aufsummiert und ergibt den berechneten χ^2 -Wert [8]. Bei einem Signifikanzniveau² von 0,05 und drei Freiheitsgraden³ beträgt der kritische Wert für den χ^2 -Test 7,815. Wäre der nach der oben beschriebenen Methode berechnete Wert größer, müsste die Nullhypothese verworfen werden. Es könnte dann mit 95%iger Wahrscheinlichkeit davon ausgegangen werden, dass die Votierungen der Techniken nicht einer Gleichverteilung unterliegen, somit würde auch eine statistisch relevante Präferenz bezüglich des Vorteils des INFOZOOMs bestehen. Abbildung 5.10 zeigt das Ergebnis des Tests. Es ergibt sich ein Wert von 6,84. Da dieser nicht größer als der kritische Wert ist, kann die Nullhypothese nicht verworfen werden, es bleibt bei der praktischen Relevanz der Mehrheit der Votierungen für den INFOZOOM.

Beobachtete Häufigkeit	Erwartete Häufigkeit	$(f_o - f_e) * (f_o - f_e) / f_e$	Test-Parameter	
9	9,50	0,026315789	Freiheitsgrade	3
5	9,50	2,131578947	Signifikanzniveau	0,05
16	9,50	4,447368421	Kritischer Wert der χ^2 -Verteilung	7,81472776
8	9,50	0,236842105		
Σ		6,842105263	<	7,81472776

Abbildung 5.10: χ^2 -Test über die Votierungen. Die Nullhypothese, die Votierungen unterlägen einer Gleichverteilung, konnte nicht abgelehnt werden.

5.5 Ergebnisse und Kapitelzusammenfassung

Eine empirische Untersuchung der ausgewählten und implementierten Techniken wurde durchgeführt, um die Eignung von Tabellen, TABLENS, INFOZOOM und PARALLELEN KOORDINATEN für die Fallauswahl in einem medizinischen Trainingssystem

²Ein Signifikanzniveau von 0,05 entspricht einer 95%igen Sicherheit der getroffenen Aussage.

³Für die Berechnung des kritischen Wertes werden m-1 Freiheitsgrade verwendet, in diesem Fall die Auswahlmöglichkeiten für eine Technik.

auch empirisch zu überprüfen. Die Auswertung der erhobenen Daten hat folgende wichtige Erkenntnisse hervorgebracht:

- Im Vergleich zu simplen Visualisierungen, wie der Tabelle, sind erweiterte Techniken sehr gut geeignet in einer Fallauswahl genutzt zu werden.
- Die verschiedenen Erweiterungen der Techniken bringen auch in vergleichsweise kleinen Datenbanken für verschiedene Aufgabentypen Vorteile.
- Die Bearbeitungszeiten haben gezeigt, dass die fortschrittlicheren Techniken schnell erlernbar sind und innerhalb einer Spanne von fünf Bearbeitungen das Performanzniveau der Tabellen erreichen bzw. deutlich überbieten können und zum Teil mehr korrekte Antworten erzielen.
- Techniken wie der INFOZOOM, welche Datensätze filtern, vermitteln Sicherheit bei der Lösungsfindung. Sie sind besonders für Auswahlen anhand konkreter Kriterien geeignet. Auch bei einer hohen Anzahl von bei der Auswahl relevanten Parameter leiten diese Techniken sicher durch die Bearbeitung.
- Die neuen Techniken zeigen allgemein eine breite Akzeptanz. Aufgrund der befremdlichen Darstellungsweise durch die Lösung vom tabellarischen Visualisierungsansatz müssen bei den PARALLELEN KOORDINATEN hierbei allerdings Abstriche gemacht werden, insbesondere bei Nichtinformatikern.

5.5.1 Eignung der Techniken für bestimmte Aufgaben

Bezug nehmend auf die in Abschnitt 3.2.6 aufgestellten Thesen zur Benutzung der Techniken in einer Fallauswahl, können anhand der Untersuchungsergebnisse folgende Aussagen bestätigt bzw. verworfen werden:

Tabellen Die Annahmen über die Nutzung der Tabellen können empirisch bestätigt werden. Die Testpersonen hatten keinerlei Schwierigkeiten mit der Bedienung der Technik und waren bekannt mit der Darstellungsweise (These 2). Die Bearbeitungszeiten und Korrektheitsverteilung zeigen die gute Nutzbarkeit der Tabellen für einfache Fallauswahlen (These 3) anhand weniger Parameter aber auch die Schwierigkeit bei vielen Parametern oder einem großen Fallspektrum den Überblick zu behalten (Thesen 1 und 4, siehe Abbildung 5.5).

TableLens Die Annahmen zu dieser Technik lassen sich nur in wenigen Punkten bestätigen. Die visuellen Hilfsmittel verhelfen laut der Testpersonen zu einem schnellen und guten Überblick, dies bestätigt These 4, sind aber wenig intuitiv. Auch die deutlich erhöhten anfänglichen Bearbeitungszeiten zeigen einen hohen Einarbeitungsaufwand, wodurch These 5 verworfen wird. Die Hilfsmittel, die durch die TABLELENS zur Verfügung gestellt werden, ermöglichen ein erleichtertes Lösen von komplexen Auswahlenszenarien (bestätigt zum Teil These 6).

5 Durchführung, Auswertung & Ergebnisse der Evaluierung

Bei einfachen Aufgaben wirkt sich der Verzicht auf textuelle Repräsentation unabhängig von der Größe des Fallspektrums sehr störend aus (verwirrt These 7 und teilweise These 6).

InfoZoom INFOZOOM erzielt in der Gesamtwertung den zweitbesten Wert für den Überblick und weist den zweitgeringsten Fehlerwert für Aufgabe 4 auf. These 8 über den Überblick kann somit bestätigt werden, auch wenn keine wirkliche Aussage darüber getroffen werden kann, wie aufwendig es ist, sich den Überblick über die Daten zu verschaffen. Sicher ist, dass mehrere Interaktionen notwendig sind, um Mengen mit dieser Technik abschätzen und vergleichen zu können. Thesen 9 und 10 können vollständig bestätigt werden. So waren Nutzer wegen des ungewohnten Spalten-Zeilen-Tauschs und der Zusammenfassung von Zellen häufig anfänglich verwirrt, dann aber erstaunt, wie gut sich mit der Technik arbeiten lässt. Die zentrale Interaktion, der Zoom, wurde als sehr effizient eingeschätzt, allerdings gehen Informationen dabei verloren, so dass hier bei komplexen Szenarien oft nur ein Teil der Lösung ausgewählt wurde (siehe Abbildung 5.5).

Parallele Koordinaten PARALLELE KOORDINATEN stellen die effizienteste Möglichkeit dar, unabhängig von der Kriterienanzahl selektive Fallauswahlen vorzunehmen, wodurch These 13 in einem Teil bestätigt wird. Bei hoher Fallanzahl geht durch Überzeichnung sehr viel vom Überblick über die Technik und ihre Möglichkeiten verloren. These 14 muss hier verworfen werden, die PARALLELEN KOORDINATEN wurden für den Zweck des Überblickes über die Daten durchschnittlich ähnlich ungünstig wie die Tabellen bewertet. Entgegen der Annahme in These 12 benötigt die Visualisierung der PARALLELEN KOORDINATEN nicht die höchste Einarbeitung. Allerdings ist die ungewöhnliche Darstellungsweise der Grund für eine vergleichsweise geringe Akzeptanz der Technik.

5.5.2 Empfehlung einer Visualisierung

Die durchgeführte Evaluierung konnte Daten erheben, die eine Auswertung und Bewertung verschiedenster Aspekte einer Visualisierungstechnik zulassen. Die verschiedenen Auswertungen nach objektiven Kriterien und subjektiven Einschätzungen lassen eine relativ eindeutige Empfehlung für die Technik des INFOZOOM zu. So zeigen die Bearbeitungszeiten, dass mit dieser Technik eine Fallauswahl sehr effektiv durchführbar ist. Dabei ist die Einarbeitungszeit auch ohne Übung so gering, dass bereits ab dem zweiten Einsatz das Leistungsniveau der bekannten Tabellen eingeholt und nach fünf Bearbeitungen deutlich besser ist. Auch wurden mit dieser Visualisierung subjektiv sehr sichere und objektiv mehr richtige Antworten gegeben. Es ist bemerkenswert, wie angenehm die Arbeit mit der Technik auch bei schwierigen Auswahlen mit vielen Kriterien, wie in Aufgabe 2, bewertet wurde und wie groß bei der Korrektheit der Vorsprung zu den anderen Techniken ist (siehe Abbildung 5.5).

Weit wichtiger als objektive Bewertungskriterien werden aber die subjektiven Bewertungen und Kommentare der Testpersonen angesehen. Mit 42 Personen, von denen mehr als die Hälfte einen nichtinformatischen Hintergrund besitzen, wird von einer ausreichend großen Stichprobe ausgegangen, um Rückschlüsse auf die Vorlieben der Zielgruppe ziehen zu können. Betrachtet man die Gruppen der Nichtinformatiker und Mediziner zusammen, sind zwei wichtige Trends zu verzeichnen. Zum Einen werden die traditionellen Tabellen als ausreichend für eine Fallauswahl angesehen, auch wenn die Interaktionen nicht ausreichend sind und keine Übersicht über die Daten gegeben werden kann. Eine Eingabemaske für Filter, durch die das Fallspektrum eingegrenzt werden kann, wird hier häufig als die optimale Lösung angesehen. Zum anderen wird der INFOZOOM bevorzugt, in dem mit der Zoominteraktion ein solcher Filter bereits integriert ist. Nimmt man die Gruppe der Informatiker hinzu, wird das Ergebnis eindeutig. So wurde im Durchschnitt der INFOZOOM in allen Kategorien besser bewertet als die Tabelle. Allerdings wurde für diese Evaluierung keine Filtermaske für die Tabellen implementiert.

Durch ein verändertes Testsetting kann das Ergebniss der Evaluation in die eine oder andere Richtung variieren. Aber gerade durch die gewählten Rahmenbedingungen des anonymen Arbeitens ohne Vorkenntnis und ohne Hilfe ist deutlich geworden, welche Techniken real für ein Fallauswahlszenario in der Medizin anwendbar sind. So zeigt sich am Beispiel der PARALLELEN KOORDINATEN, dass die objektiv performanteste bzw. schnellste Technik nicht die geeignetste sein muss. Mit dieser Technik wurden, gerade von Nichtinformatikern, die meisten Bearbeitungen abgebrochen, auch in den Kommentaren wurde hier häufig von sehr großer Verwirrung gesprochen. Zwar wurde auch für diese Technik votiert, allerdings nur in Ausnahmefällen von Personen, die ihre Computerkenntnisse als „mittel“ und darunter eingestuft haben. So kann für die PARALLELEN KOORDINATEN im Hinblick auf die Zielgruppe der Mediziner keine Eignung ausgesprochen werden.

5 Durchführung, Auswertung & Ergebnisse der Evaluierung

6 Zusammenfassung & Ausblick

Im abschließenden Kapitel soll nun die gesamte Arbeit noch einmal anhand der zentralen Fragestellungen zusammengefasst werden. Zusätzlich sollen Anregungen für weiterführende Untersuchungen gegeben werden.

6.1 Zusammenfassung

Die vorliegende Arbeit hatte zum Ziel, mehrere zentrale Fragestellungen anhand einer empirischen Evaluierung zu beantworten.

Es sollte herausgefunden werden, ob erweiterte Interaktions- und Visualisierungstechniken aus dem Gebiet der Informationsvisualisierung für die Fallauswahl in einem medizinischen Trainingssystem einsetzbar sind.

Zu diesem Zweck wurde eine solche Fallauswahl nach ihren verschiedenen Aspekten charakterisiert. Es wurden die zugrunde liegenden Datenbanken, der Stellenwert einer Fallauswahl innerhalb eines Trainingssystems und der Vorgang der Fallauswahl analysiert. Dabei wurden einige zu erwartende Nutzungsszenarien anhand der Zielgruppe der Medizinstudenten und der sich weiterbildenden Ärzte erläutert.

Aus dieser vorgenommenen Charakterisierung konnten zum Einen Anforderungen an auszuwählende Techniken gestellt werden, nach denen die Eignung verschiedener Visualisierungen diskutiert wurde. Als Ergebnis wurden die Tabelle, die TABLELENS, der INFOZOOM und die PARALLELEN KOORDINATEN ausgewählt und im Rahmen einer Evaluation verglichen. Des Weiteren konnten durch die Charakterisierung der Fallauswahl Rahmenbedingungen für eine Evaluierung bestimmt werden, welche möglichst realistische Nutzungsszenarios vorgibt.

Nach einem objektiven Vergleich der Anforderungen durch die Fallauswahl und der gebotenen Möglichkeiten konnte diese erste Frage bejaht werden. Die Antwort sollte durch eine Evaluierung bestätigt werden. Eine weitere Fragestellung war, ob der Einsatz einer solchen Technik gegenüber der Nutzung von einfachen Listendarstellungen oder Tabellen, wie sie gegenwärtig in Fallauswahlen verbreiteter Trainingssysteme eingesetzt werden, einen Vorteil bringt. Ein weiteres Ziel der Evaluierung war es, die

Akzeptanz der einzelnen Techniken bei der Zielgruppe zu messen. Anhand der Ergebnisse sollte eine Empfehlung für den Einsatz einer bestimmten Technik in einem Trainingssystem gegeben werden können.

Die Implementierung der ausgewählten Techniken wurde über die prototypische Umsetzung hinaus soweit vorangetrieben, dass sie in einer empirischen Studie für Endnutzer einsetzbar sind. Um die teils sehr ungewöhnlichen und unterschiedlichen Techniken in einem homogenen Kontext präsentieren zu können, wurde eine die Techniken einbettende GUI umgesetzt, die gleichzeitig die Grundlage der Bearbeitungsschritte in der Evaluierung bildete.

Für die eigentliche Evaluierung wurde der VisualizationEvaluator entwickelt, ein Programm welches ohne großen Aufwand eine anonyme und eigenständige Durchführung eines Evaluierungsdurchlaufs erlaubt. Hierfür wurden aus den erarbeiteten Rahmenbedingungen für verschiedene Kategorien und Schwierigkeitsgrade, Testaufgaben und -daten generiert, welche durch die Testpersonen bearbeitet werden sollten. Diese Aufgaben wurden in den linearen Ablauf des Evaluators integriert. Dabei wurden die benötigte Zeit und die Korrektheit der gegebenen Antworten für alle Bearbeitungen in der Evaluierung als objektive Messvariablen für den Vergleich der Techniken elektronisch protokolliert. Um auch die subjektive Einschätzung zur Arbeit mit den Techniken aufzunehmen wurden Fragebögen zur Aufnahme persönlicher Daten, der direkten Einschätzung nach einer Testaufgabenbearbeitung und für eine abschließende Umfrage direkt in den Verlauf des Evaluators eingebunden.

Die Evaluierung wurde mit der Verbreitung des VisualizationEvaluators über das Internet und die Weitergabe an Mediziner als Angehörige der direkten Zielgruppe durchgeführt.

Die Auswertung der Evaluierungsergebnisse konnte alle oben genannten Fragen mehr oder weniger deutlich beantworten. So wurde anhand der objektiv gemessenen Werte für die Bearbeitungszeiten (INFOZOOM, PARALLELE KOORDINATEN) und der Korrektheit der Auswahlen (INFOZOOM) gezeigt, dass der Einsatz solcher Techniken sehr wohl einen Vorteil bringt. Eine weitere wichtige Erkenntnis war, dass alle erweiterten Techniken innerhalb einer sehr kurzen Zeitspanne erlernt und im Vergleich zur einfachen Tabelle effizienter genutzt werden können. Die Auswertung der Kommentare und der subjektiven Einschätzungen konnte einen deutlichen Vorteil für die Technik des INFOZOOM zeigen, da dieser überaus positiv bewertet wurde, die breiteste Akzeptanz erfuhr und auch die häufigsten Votierungen in der Frage nach der Empfehlung einer Technik bekam.

So konnte mit der Empfehlung des INFOZOOM auch die Frage nach der am ehesten einzusetzenden Visualisierungstechnik beantwortet werden. Statistisch konnte die Antwort allerdings nicht bestätigt werden.

6.2 Weitergehende Untersuchungen

Neben der erfolgreichen Beantwortung der Hauptfragestellungen konnten durch die Bearbeitung der Thematik und die Durchführung der Evaluierung Bereiche für weitergehende Untersuchungen identifiziert werden.

6.2.1 Charakterisierung der Fallauswahl

Die Charakterisierung der Fallauswahl und die Wahl der Testaufgaben für die Evaluierung basieren in dieser Arbeit auf angenommenen Nutzungsszenarien. Diese sollten in Rücksprache mit Medizinern und Medizinstudenten noch einmal überprüft werden, da sich durch die Evaluierung der Techniken gezeigt hat, dass den Testpersonen beispielsweise weniger der Überblick über die Daten als vielmehr eine übersichtliche Darstellung wichtig war, was aus zahlreichen Votierungen für die Tabellen geschlossen werden kann.

6.2.2 Weiterentwicklung der Techniken und mögliche Folgen für das Evaluierungsergebnis

6.2.2.1 Tabellen

Es hat sich klar gezeigt, dass der Einsatz der erweiterten Techniken sinnvoll ist. Es ging aus den Nutzungsprotokollierungen sowie den Befragungen deutlich hervor, dass der zentrale Vorteil der erweiterten Techniken in der teilweisen oder vollständigen Filterung nicht interessanter Datenbereiche liegt. Der dadurch begründete Verlust von Kontext- und Vergleichsinformationen war für die meisten Testpersonen nicht von Interesse. So wurde eine solche Funktion sehr häufig bei der Tabelle vermisst und eine Filtermaske zur Ergänzung vorgeschlagen, wie sie in Onlinesuchsystemen beispielsweise bei der Durchsuchung von Immobiliendatenbanken angewendet werden. So könnten nicht interessierende Spalten und uninteressante Fälle ausgeblendet werden. Auch das Umordnen von Spalten könnte durch automatisches Scrollen bei Verlassen des Sichtbarkeitsbereiches optimiert werden. Tatsächlich könnte sich das Ergebnis der subjektiven Meinungsumfrage aufgrund der erhöhten Akzeptanz stark in Richtung Tabelle verschieben, hierbei sollte dann jedoch erneut überprüft werden, um wie viel die Tabelle durch den Einsatz zusätzlicher Komponenten effizienter wird.

6.2.2.2 TableLens

Die TABLENS hat sich als am wenigsten intuitiv gezeigt, auch das nicht vollständige Ausblenden von Fällen wurde eher als störend empfunden. Insbesondere ist hier

die Farbkodierung der COLOR SWATCHES passend zu wählen, so dass ein intuitiver Bezug zu den Daten entsteht. Weiterhin könnte die Zeilenhöhe minimierter Zellen noch so weit verringert werden, dass ein Fall nicht mehr als einzelne Zeile auszumachen ist, sondern durch seine farbliche Präsentation nur noch zur Darstellung des Kontextes beiträgt. Es müsste dann aber noch einmal überprüft werden, ob der Einsatz in einer relativ kleinen Datenbank, wie dem Fallspektrum eines Trainingssystems, noch sinnvoll ist.

6.2.2.3 InfoZoom

Beim INFOZOOM wurde die Schwierigkeit bemängelt, Datenwerte eindeutig einem Fall zuzuordnen. Dies könnte durch eine weitere ToOLTIP-Technik erreicht werden, bei der die gesamte entsprechende Spalte eines Falles unabhängig von Zellverbänden eingefärbt wird.

6.2.2.4 Parallele Koordinaten

Die PARALLELEN KOORDINATEN wurden insbesondere von Personen mit mittleren oder geringen Computerkenntnissen nicht akzeptiert. Es kann dies auf die Implementierung mit einer 3D-Grafikbibliothek zurückgeführt werden, die insbesondere Defizite für die Mausinteraktionen sowie ToOLTIP-Informationen und allgemein Beschriftungen aufweist. Eine Reimplementierung in einem homogenen Kontext mit den übrigen Techniken, beispielsweise im Qt-Framework, könnte die Interaktionen erheblich erweitern und wahrscheinlich die Akzeptanz deutlich erhöhen. So wären dann durch die Nutzung des Model-View-API auch manuelle Sortierungen und das Umordnen von Spalten bzw. Achsen möglich, was auch die Lösung komplexer Problemstellungen mit dieser Technik effizienter gestalten würde. Eine möglicherweise erhöhte Akzeptanz muss in diesem Fall mit einer weiteren Evaluierung bestätigt werden.

6.2.3 Mischung von Aufgaben verschiedener Kategorien

Es hat sich in der Evaluierung gezeigt, dass die Mischung verschiedener Aufgabenkategorien nicht von Vorteil ist. So wurde die sehr komplexe Aufgabe 4, die eine andere Vorgehensweise verlangte, häufig nicht verstanden. Auch die Zusatzaufgaben, welche ein Verständnis und einen Überblick der Daten überprüfen, wurden oft trotz Hinweisen nicht genügend beachtet. Es ist deshalb anzunehmen, dass für solch komplexe Aufgabenstellung eine vollständig anonyme Evaluierung nicht geeignet ist. Diese Aufgabenstellungen benötigen die Möglichkeit, Rückfragen stellen zu können und sollten in einer überwachten Evaluierung überprüft werden. Durch so eine zusätzliche Studie könnte wahrscheinlich der Vorteil, den die TABLENS bei solchen Aufgaben anhand

der in dieser Arbeit erhaltenen Ergebnisse zu haben scheint, deutlicher nachgewiesen werden. Da diese Aufgabenkategorie nur im geringen Maße und hauptsächlich im Nutzungsszenario eines Fallautors auftritt, sollten unterschiedliche Visualisierungen für Fallauswahl und Fallerstellung erwogen werden.

6.2.4 Weitere Evaluierungsszenarien

Insgesamt hat sich eine breite Akzeptanz der verschiedenen Visualisierungen gezeigt. Um zusätzlich andere bzw. zukünftige Techniken für die Verwendung in einer Fallauswahl mit in den vorhandenen Testablauf einzugliedern, sollte jeweils eine vorgelagerte Studie durchgeführt werden, welche die Akzeptanz durch die zukünftige Nutzergruppe sichert. Hierzu bietet es sich an, generierte Daten aus dem medizinischen Kontext zu verwenden und von einer kleineren Gruppe, welche ausschließlich aus Medizinerinnen besteht, testen zu lassen. Eine solche Studie würde zu einer für die Zielgruppe akzeptablen Implementierung führen und ließe damit einen genaueren Vergleich von Visualisierungsmethoden im bestehenden Testframework zu.

Die Evaluierung hat, neben einem deutlichen Ergebnis für den INFOZOOM, auch einen Trend zur Tabelle aufzeigen können. Eine Überlegung wäre, die um eine Filtermaske erweiterten Tabellen und den INFOZOOM in einem Trainingssystem über längere Zeit parallel anzubieten und somit einer Langzeitstudie zu unterziehen. Mit diesem Vorgehen wäre es möglich zu messen, inwieweit die in der durchgeführten Studie gezeigte Effizienz des INFOZOOM auch nach längeren Benutzungspausen noch präsent ist, oder ob die Technik dann immer wieder „neu erlernt“ werden muss.

6 Zusammenfassung & Ausblick

Literaturverzeichnis

- [1] ADAC. Automobilsuche mit InfoZoom, Abgerufen am 15.11.2007. Automobilsuche im Kundenportal des ADAC, 2007. https://www.adac.de/Auto_Motorrad/Autokatalog/default.asp.
- [2] M. Adler, I. Hege, M. Holzer, and M. Fischer. CASUS - new implementations and challenges. In *Slice of Life Conference*, Lausanne, 2006.
- [3] Klinikum Innenstadt der Ludwig-Maximilians-Universität München AG Medizinische Lernprogramme. Fallauswahl des CASUS-Systems, Abgerufen am 10.11.2007. CASUS Website, 2007. http://www.casus.net/vhb_demo/data/cases4.html.
- [4] C. Ahlberg and E. Wistrand. IVEE: an Information Visualization and Exploration Environment. In *INFOVIS '95: Proceedings of the 1995 IEEE Symposium on Information Visualization*, page 66, Washington, DC, USA, 1995. IEEE Computer Society.
- [5] A. Alexakis, M. Deftereos, and Y. Samiotakis. Visualisation Techniques in Statistics. The Parallel Coordinates Visualiser(PARCOVI).
- [6] Trolltech ASA. An Introduction to Model/View Programming, Abgerufen am 10.11.2007. Online Dokumentation zu Schlüsseltechniken des Qt-Frameworks, 2007. <http://doc.trolltech.com/4.3/model-view-introduction.html>.
- [7] Trolltech ASA. Qt Reference Documentation (Open Source Edition). Online Dokumentation des Qt-Frameworks, 2007. <http://doc.trolltech.com/4.3/index.html>.
- [8] F. Brosius. *SPSS 14*. Mitp-Verlag, Bonn, 2006.
- [9] C. M. Brown. *Human-computer interface design guidelines*. Ablex Publishing Corp., Norwood, NJ, USA, 1988.
- [10] C. Brunsdon, A. S. Fotheringham, and M. E. Charlton. An Investigation of Methods for Visualising Highly Multivariate Datasets. Technical report, AGCOG, February 1995.

- [11] J. Cordes, K. Muehler, and B. Preim. Die Konzeption des LiverSurgery-Trainers. In *GI-Workshop „Softwareassistenten - Computerunterstützung für die medizinische Diagnose und Therapieplanung“ im Rahmen der GI-Jahrestagung 2006*, pages 514–521. Springer, 2006.
- [12] MeVisLab Dokumentation. Hilfeseite zur Benutzung des ObjMgr-Konzepts, Abgerufen am 08.11.2007. <http://www.mevislab.de/fileadmin/MLabModules/std/help/html/private/ObjMgrFrame.html>.
- [13] D. Ericson, J. Johansson, and M. Cooper. Visual Data Analysis Using Tracked Statistical Measures within Parallel Coordinate Representations, 2005.
- [14] Caseport-Portal für die fallbasierte Lehre in der Medizin. Fallauswahl des Caseport-Portals, Abgerufen am 10.11.2007. Caseport Website, 2007. <http://link.caseport.de/caseport/indexjsp.htm>.
- [15] J. Friedman. Exploratory Projection Pursuit. *Journal of the American Statistical Association*, 82(397):249–266, 1987.
- [16] Fua, Ward, and Rundensteiner. Hierarchical parallel coordinates for exploration of large datasets. In *Visualization '99. Proceedings*, October 1999.
- [17] V. González and A. Kobsa. A Workplace Study of the Adoption of Information Visualization Systems. In *Proceedings of I-KNOW'03: 3rd International Conference on Knowledge Management*, pages 92–102, 2003.
- [18] V. González and A. Kobsa. Benefits of Information Visualization Systems for Administrative Data Analysts. In *IV '03: Proceedings of the Seventh International Conference on Information Visualization*, page 331, Washington, DC, USA, 2003. IEEE Computer Society.
- [19] M. Haag and F. J. Leven. *Computerunterstützte Lehr- und Lernsysteme in der Medizin in Kramme R (Hrsg.): Medizintechnik. 3. Aufl.*, pages 803–813. Springer Medizin Verlag, 2007.
- [20] H. Hartson, J. Castillo, C. Jose, J. Kelso, W. Neale, C Wayne, and J. Kamler. Remote evaluation: the network as an extension of the usability laboratory. In *The 1996 Conference on Human Factors in Computing Systems, CHI 96*, pages 228–235, Vancouver, BC, Canada, 1996.
- [21] Universität Heidelberg. Fallauswahl des CAMPUS-Systems, Abgerufen am 10.11.2007. CAMPUS Website, 2007. <http://www.medicase.de>.
- [22] A. Inselberg. Parallel coordinates: a tool for visualizing multi-dimensional geometry. In *Visualization '90., Proceedings of the First IEEE Conference on Visualization*, pages 361–378, September 1990.
- [23] A. Inselberg. Multidimensional detective. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 100–107, September 1997.

- [24] inxight Software. Online-Demonstration der TableLens-Visualisierung, Abgerufen am 08.12.2007. inxight Software Website, 2007. http://www.inxight.com/demos/tl_college/tl_college.html.
- [25] R. Jeffries, J. R. Miller, C. Wharton, and K. Uyeda. User interface evaluation in the real world: a comparison of four techniques. In *CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 119–124, New York, NY, USA, 1991. ACM.
- [26] A. Kobsa. An Empirical Comparison of Three Commercial Information Visualization Systems. In *INFOVIS '01: Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01)*, page 123, Washington, DC, USA, 2001. IEEE Computer Society.
- [27] A. Kobsa and J. Kim. Market Penetration of Visual Interfaces: An Analysis of Impediments. Industry Advisory Board IAB-2002-2, 2002.
- [28] G. Krasner and S. Pope. A Description of the Model-View-Controller User Interface Paradigm in the Smalltalk-80 system. *Journal of Object Oriented Programming*, 1(3):26–49, 1988.
- [29] D. Lane. HyperStat Online Statistics Textbook, Abgerufen am 10.11.2007. <http://davidmlane.com/hyperstat/index.html>.
- [30] C. Lewis, P. G. Polson, C. Wharton, and J. Rieman. Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. In *CHI '90: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 235–242, New York, NY, USA, 1990. ACM.
- [31] G. Mark, K. Carpenter, and A. Kobsa. Are there benefits in seeing double?: a study of collaborative information visualization. In *CHI '03: CHI '03 extended abstracts on Human factors in computing systems*, pages 840–841, New York, NY, USA, 2003. ACM Press.
- [32] A. R. Martin and M. O. Ward. High Dimensional Brushing for Interactive Exploration of Multivariate Data. In *VIS '95: Proceedings of the 6th conference on Visualization '95*, page 271, Washington, DC, USA, 1995. IEEE Computer Society.
- [33] S. Mirschel. Erstellung eines Prototypen für ein fallbasiertes Lernsystem in der Leberchirurgie. Master's thesis, Dept. of Computer Science, Mai 2004. Betreut durch Ragnar Bade.
- [34] R. Molich and J. Nielsen. Improving a human-computer dialogue. *Commun. ACM*, 33(3):338–348, 1990.
- [35] J. Nielsen. Ten Usability Heuristics, Abgerufen am 09.11.2007. http://www.useit.com/papers/heuristic/heuristic_list.html.

- [36] J. Nielsen and R. Molich. Heuristic evaluation of user interfaces. In *CHI '90: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 249–256, New York, NY, USA, 1990. ACM.
- [37] P. Pirolli and R. Rao. Table lens as a tool for making sense of data. In *Proceedings of the workshop on Advanced visual interfaces*, pages 67–80. ACM Press, 1996.
- [38] C. Plaisant. The challenge of information visualization evaluation. In *AVI '04: Proceedings of the working conference on Advanced visual interfaces*, pages 109–116, New York, NY, USA, 2004. ACM Press.
- [39] B. Preim. *Entwicklung Interaktiver Systeme*. Springer Verlag, 1999.
- [40] R. Rao and S. K. Card. The table lens: merging graphical and symbolic representations in an interactive focus + context visualization for tabular information. In *CHI '94: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 318–322, New York, NY, USA, 1994. ACM Press.
- [41] MeVis Research. Webauftritt zur Entwicklungsumgebung MeVisLab, Abgerufen am 06.12.2007. <http://www.mevislab.de>.
- [42] S-Plus. Website des Vertreibers, Abgerufen am 17.09.2007. <http://www.insightful.com/products/splus/default.asp>.
- [43] H. Siirtola. Direct Manipulation of Parallel Coordinates. In *Proceedings. IEEE International Conference on Information Visualization*, pages 373–378, July 2000.
- [44] M. Spenke. Visualization and interactive analysis of blood parameters with InfoZoom. *Artificial Intelligence in Medicine*, 22(2):159–172, 2001.
- [45] M. Spenke and C. Beilken. Visual, Interactive Data Mining with InfoZoom - the Financial Data Set, September 1999. Contribution to the 'Discovery Challenge' at the 3rd European Conference on Principles and Practice of Knowledge Discovery in Databases.
- [46] M. Spenke and C. Beilken. Infozoom - analysing formula one racing results with an interactive data mining and visualisation tool, Juli 2000.
- [47] M. Spenke and C. Beilken. Visualization of trees as highly compressed tables with InfoZoom, 2003. Poster compendium of IEEE Symposium on Information Visualization.
- [48] P. Starkloff and P. Reimann. VIROR TP3 Evaluationsbericht: Das Computerlernprogramm CAMPUS in der Mediziner Ausbildung der Pädiatrie an der Universität Heidelberg, 2001.
- [49] InfoZoomViewer von humanIT. Downloadsite des Vertreibers, Abgerufen am 08.11.2007. <http://www.infozoom.com/deu/download/index.htm>.

- [50] E. Wegman and Q. Luo. High dimensional clustering using parallel coordinates and the grand tour. Technical Report 124, Center for Computational Statistics George Mason University George Mason University, Fairfax, Virginia 22030, U.S.A., 1996.
- [51] E. Wegmann. Hyperdimensional Data Analysis Using Parallel Coordinates. *Journal of the American Statistical Association*, 85(411):464–475, September 1990.
- [52] E. Wegmann. Data Mining and Visualization: Some Strategies. *Bull Int Statist Inst*, (52):223–226, 1999.
- [53] J. Wernecke. *The Inventor Mentor: Programming Object-Oriented 3D Graphics with Open Inventor, Release 2*. Addison-Wesley Publishing Group, 1994.

Literaturverzeichnis

Tabellenverzeichnis

4.1	Vergleich der Operationen der TABLELENS im Original und der implementierten Version	49
4.2	Darstellungsvarianten der PARALLELEN KOORDINATEN bei Mausinteraktionen.	52
5.1	Gegenüberstellung der Interaktionen in den implementierten Techniken, die zur Durchführung einzelner Lösungsschritte einer Aufgabe verwendet werden können.	62

Tabellenverzeichnis

Abbildungsverzeichnis

2.1	Suche im Automobilkatalog im ADAC-Kundenportal mit INFOZOOM	6
2.2	Fallauswahlen verschiedener medizinischer Trainingssysteme	7
2.3	Übliche Tabellendarstellung	8
2.4	Darstellung einer TableLens-Visualisierung ohne Linsenbereich	9
2.5	TABLELENS-Darstellung mit mehreren Linsenbereichen	9
2.6	Verschiedene Sichten von INFOZOOM	11
2.7	Darstellung von PARALLELEN KOORDINATEN mit und ohne verschiedene Erweiterungen	12
2.8	Farbhistogramm und Sättigungsplot von PARALLELEN KOORDINATEN	13
2.9	PROJEKTION PURSUIT	13
2.10	RADVIS-Visualisierungsansatz	14
2.11	Änderung von statistischen Werten durch Brushing und Durchschnittsbildung in PARALLELEN KOORDINATEN	21
3.1	Beispiel einer Datenkategorisierung in einem medizinischen fallbasierten Lernsystem	26
4.1	Datenablageformat des ObjMgr in MEVISLAB	44
4.2	CaseScanner-Schnittstelle zur Datenbereitstellung	45
4.3	Darstellung einer einbettenden Visualisierungsoberfläche	46
4.4	Model-View-Architektur des Qt-Frameworks	47
4.5	Tabellenkopf einer nach mehreren Spalten sortierten Tabelle	48
4.6	Mögliche Zustände der Zellen in der TABLELENS-Implementation	48
4.7	Implementierung des INFOZOOM	50
4.8	Quantitätsabschätzung in einer INFOZOOM-Zeile	51
4.9	Implementierung der PARALLELEN KOORDINATEN	51
4.10	Aufbau der OpenInventor-Szene des SoParallelCoordinates-Moduls	53
4.11	Aufbau/Ablauf des VisualizationEvaluators	55
4.12	Videos zur Vorstellung der Visualisierungen	56
4.13	Bearbeitung einer Aufgabe	57
5.1	Vorkenntnisse der Testpersonen im Durchschnitt	64
5.2	Durchschnittliche Bearbeitungszeit pro Aufgabe und Technik	66
5.3	Verlauf der Bearbeitungszeiten über die drei einfachsten Aufgaben	67
5.4	Verteilung der Korrektheit der Antworten für alle Techniken	69

Abbildungsverzeichnis

5.5	Verteilung der Korrektheit der Antworten für die Aufgaben 2 und 4 .	70
5.6	Durchschnittliche Wertung für die Eignung einer Technik für eine Aufgabe	72
5.7	Durchschnittliche Bewertung der Visualisierungen nach den Aspekten Interaktion, Überblick und Effizienz	73
5.8	Votierungen der Testpersonen für die Techniken	76
5.9	Prozentuale Verteilung der Votierungen in den Berufsgruppen	77
5.10	χ^2 -Test über die Votierungen	78