

Information Clustering in der Umweltinformatik

Vom Fachbereich Informatik der Technischen Universität Kaiserslautern
zur Verleihung des akademischen Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)
genehmigte Dissertation
von

Maja Ruby

Dekan	Prof. Dr. Reinhard Gotzhein
Promotionskommission	
Vorsitz	Prof. Dr. Karsten Berns
1. Gutachter	Prof. Dr. Hans Hagen
2. Gutachter	Prof. Dr. Jörg Meyer
Wissenschaftliche Aussprache	27. Oktober 2006

Danksagung

Mein besonderer Dank gilt meinem Doktorvater, Herrn Prof. Dr. Hans Hagen, der die vorliegende Arbeit nicht nur grundsätzlich ermöglicht und die interessante Thematik vorgeschlagen, sondern die Entwicklung der Arbeit mit hilfreichen Gesprächen und fruchtbaren Anregungen stets konstruktiv begleitet hat.

Herrn Prof. Dr. Jörg Meyer möchte ich für die Bereitschaft danken, diese Arbeit als Zweitgutachter anzunehmen sowie für die intensiven Diskussionen in Zusammenhang mit der Arbeit.

Meiner Kollegin und Freundin Frau Dipl.-Ing. Inga Scheler danke ich für die wissenschaftliche Begleitung bei der Arbeit, die Motivation, die sie mir mit Anregungen und positiver Kritik vermittelt hat und nicht zuletzt für das mühsame und sorgfältige lesen des Manuskriptes.

Frau Mady Gruys möchte ich herzlich für die beratende Anteilnahme und die Zuversicht danken, die sie mir stets entgegen gebracht hat.

Ganz besonders danke ich meiner Familie für die unermüdliche Unterstützung während dieser Jahre. Besonders meine Mutter hat mich in allen Umbrüchen und Veränderungen gestärkt und mir vieles ermöglicht, was ohne ihre Hilfe nicht möglich gewesen wäre. Insbesondere danke ich meinem Ehemann Herbert Ruby für die Geduld und Zuversicht während der gesamten Zeit.

Zusammenfassung

Die Computerisierung der Gesellschaft bedingt ein ständiges Zunehmen der Geschwindigkeit, mit der neue Daten erzeugt werden. Parallel zu dieser Entwicklung steigt der Bedarf an geeigneten Analyseverfahren, die in diesen großen und oftmals heterogenen Datenmengen Muster finden, Zusammenhänge entdecken und damit Wissen erzeugen. Das in dieser Arbeit entwickelte Verfahren findet die passende Struktur in einer ungeordneten, abstrakten Datenmenge, ordnet die zugrunde liegenden Informationen und bündelt diese somit für eine gezielte Anwendung. Dieser Prozess des Information Clustering ist zweistufig, es erfolgt zuerst ein generelles Clustering, an das sich eine interpretierende Visualisierung anschließt. Für das Clustering wird das Verfahren der Voronoidiagramme entscheidend erweitert. Durch den Einsatz einer generellen Distanzfunktion wird die Modellierung der durch die großen Datenmengen entstehenden multidimensionalen Parameter sowie weiterer Gewichte ermöglicht. Gleichzeitig werden hierdurch gravierende Probleme beim Einsatz von Metriken vermieden. Eine anschließende Visualisierung mit Techniken aus dem Bereich der Informationsvisualisierung unterstützt die Interpretation der neu gewonnenen Informationen.

Für die praktische Anwendung der in dieser Arbeit entwickelten Verfahren wird die Stadtplanung betrachtet. In der Stadtplanung wird das Modell des Planungsablaufes eingesetzt, mit dem verschiedene Planungsalternativen erzeugt werden. Dieses Modell ist jedoch zu starr, um den dynamischen Anforderungen in der Realität gerecht zu werden. Das Information Clustering erweitert den klassischen Planungsablauf, die Flexibilität des Modells wird dadurch erhöht und die Komplexität reduziert. Das Ergebnis der Berechnung ist genau eine Planungsalternative, die sämtliche Eingabeparameter kanalisiert.

Inhaltsverzeichnis

1	Einführung	1
2	Grundlagen	5
2.1	Clustering	6
2.1.1	Clusteranalyseverfahren	7
2.1.2	Metriken	8
2.2	Visualisierung	11
2.2.1	Informationsvisualisierung	13
2.2.2	Techniken	15
2.2.3	Systeme	21
2.3	Umweltinformatik	23
2.3.1	Grundlagen	24
2.3.2	Stadtplanung	25
2.3.3	Informationssysteme	30
3	Information Clustering im Planungsablauf	36
3.1	Planerische Entscheidungsfindung	37
3.1.1	Systematischer Ansatz	37
3.1.2	Modellbildung	40
3.2	Profilerstellung	42
3.2.1	Akteure und Sichtweisen	43
3.2.2	Parameterauswahl	46
3.3	Information Clustering	49
3.3.1	Generelle Zielfunktion	49
3.3.2	Visualisierung	52
4	Clusteranalyse	54
4.1	Verfahren	54
4.1.1	Voronoidiagramme	55
4.1.2	Topologie	57
4.1.3	Geometrische Eigenschaften	58

4.2	Distanzfunktion	60
4.2.1	Konvexe Distanzfunktion	61
4.2.2	Gewichtete Distanzfunktion	65
4.2.3	Genereller Ansatz	69
4.3	Generischer Wachstumsalgorithmus	73
4.3.1	Herkömmliche Verfahren	74
4.3.2	Gewichtetes Wachstum	75
4.3.3	Richtungsabhängiges Wachstum	77
5	Visualisierungsstrategien	79
5.1	Anforderungen	79
5.2	Picklock-Icon	81
5.2.1	Stick-Figure- und Velcro Icon	81
5.2.2	Konzept	83
5.2.3	Umsetzung	85
5.3	Spike-Tree	88
5.3.1	Botanische Visualisierung	89
5.3.2	Konzept	94
5.3.3	Umsetzung	98
6	Informationssystem für die Stadtplanung	103
6.1	Konzept	104
6.2	Systemaufbau	105
6.2.1	Datenverarbeitung	107
6.2.2	Client	108
6.2.3	Server	110
6.3	Anwendung in der Stadtplanung	111
6.3.1	IKone	112
6.3.2	GIS-Bericht	116
6.3.3	Virtueller Rundgang	117

Kapitel 1

Einführung

In der modernen, von Computern geprägten Gesellschaft werden immer mehr Vorgänge, seien es geschäftliche Transaktionen, wissenschaftliche Experimente oder andere, Daten generierende Prozesse, rein elektronisch erfasst und in großen Datenbanken gespeichert. Dabei wächst nicht nur die Größe der Datenbanken, sondern auch die Geschwindigkeit, mit der neue Daten produziert werden. Bei der steigenden Informationsflut werden Techniken und Tools immer bedeutsamer, die in großen Mengen von Daten Muster finden, Zusammenhänge entdecken und damit Wissen erzeugen. Durch die stetige Erhöhung der Komplexität der Aufgabenstellungen und der Ausweitung des Analyse- und Informationsbedarfes für viele Anwendungen werden Systeme benötigt, die mit diesen neuen Anforderungen umgehen können. Diese Systeme müssen in der Lage sein, die großen und oftmals heterogenen Datenmengen bearbeiten zu können. Zusätzlich müssen sie die geeigneten Techniken und Tools für die Auswertung der Daten zur Verfügung stellen.

Im Rahmen dieser Arbeit werden solche Techniken und Tools entwickelt. Basis dieses Verfahrens bildet ein flexibles Clustering, das die zugrunde liegenden Informationen ordnet und somit für eine gezielte Auswertung bündelt. Durch den Einsatz einer generellen Distanzfunktion wird hierbei ein freier Ähnlichkeitsbegriff entwickelt, der auch die Integration multidimensionaler Parameter und weiterer Gewichte zulässt. Ein weiterer Vorteil der generellen Distanzfunktion ist die Vermeidung der Probleme, die bei dem klassischen Ansatz mit Metriken auftreten. An das Clustering schließt sich eine Visualisierung an, mit der die neu gewonnenen Informationen für eine anschließende Evaluation dargestellt werden. Für die Visualisierung werden Techniken aus dem Bereich der Informationsvisualisierung entwickelt, mit denen die hochdimensionalen Parameter nachgebildet werden können.

Besonders wirkungsvoll lassen sich die Verfahren in bestehende Entscheidungsprozesse integrieren. Durch die Kombination des Clustering mit der Visualisierung werden Entscheidungsabläufe beschleunigt und die Qualität der Ergebnisse erhöht. Die

Potentiale agierender, auch gegebenenfalls gegengerichteter Gruppen werden mit diesem Verfahren kanalisiert. Durch die Integration der Profile einer speziellen Aufgabenstellung werden die Verfahren für eine konkrete Anwendung konfiguriert.

Für die praktische Anwendung der in dieser Arbeit entwickelten Verfahren wird die Umweltinformatik betrachtet. Die Umweltinformatik ist eine relativ neue Teildisziplin der angewandten Informatik und beschäftigt sich interdisziplinär mit der Analyse und Bewertung von Umweltsachverhalten mit Mitteln der Informatik. Sie verfolgt damit das Ziel, einen Beitrag zur Untersuchung, Behebung, Vermeidung oder Minimierung von Umweltbelastungen und Umweltschäden zu leisten. Neben der vorrangig ökologisch orientierten Umweltinformatik hat sich in den letzten 10 Jahren eine stärkere Verlagerung der Schwerpunkte der Umweltinformatik in Richtung der informationstechnischen Unterstützung des betrieblichen Umweltschutzes ergeben. Diese Ausweitung der Sichtweise der Umweltinformatik ist vor allem auch vor dem Hintergrund der Diskussion zur Nachhaltigkeit zu sehen, die neben der rein ökologischen Sicht auch ökonomische und soziale Fragen mit einbezieht.

Die Diskussion um die Nachhaltigkeit hat u.a. auch Auswirkungen auf die Flächenneuinanspruchnahme. Für Deutschland hat dies zur Folge, dass bis zum Jahr 2020 der Neufächenverbrauch von derzeit ca. 105 ha/Tag auf 30 ha/Tag gesenkt werden soll. Zum Erreichen dieses Zieles wird es vermehrt notwendig, bereits vorgenutzte, brachgefallene Flächen einer Neu- oder Umnutzung zuzuführen sowie Baulücken zu schließen. Damit kommt der Konversion allgemein und mit den Teilbereichen der militärischen und der zivilen Konversion von Liegenschaften im speziellen eine besondere Bedeutung zu.

Diese Arbeit konzentriert sich zunächst auf die Flächennutzung im innerstädtischen Bereich, umgesetzt durch die Stadtplanung. Die Stadtplanung ist eine Teildisziplin der Raumplanung und beschäftigt sich mit der Analyse der Stadt und darauf folgend der Erarbeitung von Planungskonzepten unter Abwägung aller relevanten Interessen mit dem Ziel der Konfliktminimierung. Stadtplanung dient der Steuerung der Bodennutzung für gesamte Stadt- oder Gemeindegebiete oder für Teilbereiche davon. Durch die Ausweitung des Informations- und Analysebedarfs erhöht sich die Komplexität in der Stadtplanung stetig, Planungs- und Entscheidungsabläufe beschleunigen sich. Die räumliche Entscheidungsfindung ist somit ein hochgradig komplexer und dynamischer Prozess. Bisherige Ansätze dieser Komplexität zu begegnen sind zum einen zu starr, um den dynamischen Anforderungen in der Realität zu genügen und zum anderen sind sie in ihrer Umsetzung zu langsam.

Die Nachbildung der Realität mit Modellen ist ein notwendiges Hilfsmittel zum Erkennen von Zusammenhängen und zum Treffen von Entscheidungen. Die Transparenz und die Nachvollziehbarkeit komplexer Entscheidungsprozesse wird durch eine modellhafte Darstellung erheblich erhöht. Auch in Entscheidungssituationen

im Zusammenhang mit hochgradig interdependenten räumlichen Systemen, ist die unterstützende Einschaltung von Modellen unerlässlich. Das maßgebliche Instrument zur Begegnung von Komplexität in der Stadtplanung ist der Planungsablauf. Der Planungsablauf legt die Gesetzmäßigkeiten bei Planungsvorhaben fest und wird zur Entscheidungsfindung zwischen verschiedenen Planungsalternativen eingesetzt. Die so gewonnene modellhafte Darstellung erhöht zwar die Transparenz der Planungsabläufe, erfaßt jedoch nicht deren dynamische Interdependenz. Der Starrheit und der zunehmenden Geschwindigkeit bei der Umsetzung von Planungen wird hierdurch nicht entgegengewirkt. Die Einbindung von Verfahren zur Unterstützung der Entscheidungsfindung bei dem Planungsablauf ist an dieser Stelle sinnvoll und erfolgversprechend. Die in dieser Arbeit entwickelten Verfahren werden hierfür an die Rahmenbedingungen der Planungsabläufe in der Stadtplanung angepaßt. Das Ziel ist, die Flexibilität der Planinhalte zu erhöhen und eine Effizienz- und Effektivitätssteigerung bei Planungen zu erzielen.

Hierzu werden die in der Stadtplanung vorliegenden Informationen - die Ist-Situation - gezielt ermittelt und einem Prozess - dem Information Clustering - zugeführt. Im Laufe der Arbeit werden die charakteristischen Punkte des Planungsablaufes ermittelt. An diesen Punkten setzt das Information Clustering unterstützend an. Durch diese Erweiterung ist es möglich genau eine Planungsalternative zu berechnen und somit den Entscheidungsprozess maßgeblich zu unterstützen.

Für die Aufnahme der Ist-Situation werden zunächst die Profile ermittelt, die bei jedem Planungsablauf entstehen. Diese sind durch die Standortfaktoren einer Fläche und die agierenden Akteure geprägt. Die Menge der einflußnehmenden Parameter ist sehr groß und heterogen, so dass die Datenbasis eine sehr große und unstrukturierte Datenmenge darstellt. Das Ziel des Information Clustering ist eine Planungsalternative zu erstellen - in dem Fall der Konversion entspricht das der Unterteilung einer Fläche -, die das Zusammenspiel der Parameter widerspiegelt und die Anforderungen der einzelnen Akteure kanalisiert. Mit der neu berechneten Flächenunterteilung wird eine aussagekräftige Struktur in die Daten gebracht - es wird somit neues Wissen aus den zugrunde liegenden Daten erzeugt.

Um diesen Anforderungen gerecht zu werden besteht das Information Clustering aus einem zweistufigen Prozess. Zunächst erfolgt das Clustering, an das sich eine Visualisierung anschließt. Für das Clustering wird das Verfahren der Voronoidiagramme entscheidend erweitert. Die Eigenschaft der Voronoidiagramme lokale Probleme zu globalisieren, ist für den geplanten Einsatzbereich besonders wertvoll. In dem Bereich der Planung werden globale Verfahren benötigt, da kleine lokale Änderungen im Planungsablauf eine große globale Wirkung haben. In dem Bereich der Stadtplanung bestehen zudem spezielle Anforderungen an die Topologie der Flächenunterteilungen. Besonders der Zusammenhang ist ein wichtiges Kriterium. Auch an

die geometrischen Eigenschaften eines Clusteranalyseverfahrens - wie beispielsweise die Konvexität einer Unterteilung - werden Anforderungen gestellt. Für die Implementation des generellen Clusterings wurde ein generischer Wachstumsalgorithmus entworfen.

An das Clustering schließt sich die Visualisierung an, für die zwei verschiedene Techniken aus dem Bereich der Informationvisualisierung entwickelt wurden. Das Picklock-Icon gehört zu der Iconbasierten Informationvisualisierung, welche sich vor allem für die Darstellung von diskreten, multivariaten Daten eignet. Die Iconbasierten Visualisierungstechniken übertragen nicht sichtbare und unentdeckte statistische Strukturen der Daten in visuelle Strukturen wie Inseln, Streifen oder Gradienten. Der Vorteil des Picklock-Icons, die Daten räumlich zusammenhängend darstellen zu können, birgt aber auch Nachteile. Durch den räumlichen Zusammenhang überdeckt die Visualisierung möglicherweise eine darunter liegende Karte. Das zweite Verfahren aus dem Bereich der Botanischen Informationsvisualisierung, der Spike-Tree, arbeitet sowohl dreidimensional als auch punktorientiert und vermeidet dadurch dieses Problem. Der Vorteil der kontextsensitiven Darstellung der Daten in Bezug zu einer Karte macht diese Technik äußerst vorteilhaft für die behandelte Anwendung.

Es wurde ein System entwickelt, das Lösungen für verschiedene, komplexe Anwendungen in der Stadtplanung integriert. Mit dem System ist die Erstellung eines GIS-Berichtes oder eines virtuellen Rundganges durch ein Stadtgebiet möglich. Das Kernstück dieses Systemes bildet das Tool IKone, welches sämtliche in dieser Arbeit entwickelte Verfahren anhand eines Konversionsprojektes in der Stadtplanung beispielhaft umsetzt.

In dem Kapitel 2 erfolgt zunächst eine allgemeine Einleitung in die Themenfelder dieser Arbeit. Hierzu werden die grundlegenden Begriffe und Ansätze der Bereiche Umweltinformatik, Clustering und Visualisierung vorgestellt. Die Arbeiten in den Kapitel 3 - 6 setzen auf diesen Grundlagen auf. Das Kapitel 3 konzentriert sich auf das strukturelle Vorgehen bei dem Planungsablauf und arbeitet die Vorteile und Anforderungen eines Analyseverfahrens heraus. Hierbei werden die Anknüpfungsmöglichkeiten an den Planungsablauf ermittelt und dessen Struktur erarbeitet. In Kapitel 4 wird das Clustering behandelt und die hierfür benötigten Verfahren entwickelt und in Kapitel 5 die Visualisierungstechniken. Den Abschluss dieser Arbeit bildet der Entwurf des Informationssystems für die Stadtplanung in Kapitel 6.

Kapitel 2

Grundlagen

Es werden immer mehr Verfahren und Techniken benötigt, mit denen aus großen, oftmals unstrukturierten Datenmengen Informationen und Zusammenhänge abgeleitet werden können. Das Ziel dieser Techniken ist eine (semi) automatische Extraktion von Wissen aus den Datenmengen, um so neues Wissen aus den vorhandenen Daten zu erschließen.

Hierbei ist es wichtig, grundlegende Eigenschaften des zu erzeugenden Wissens zu berücksichtigen. Es sollte

- gültig
- neuartig
- potentiell nützlich und
- verständlich

sein.

Ein wichtiger Prozess umfasst hierbei den Bereich der Datenhandhabung, der in einem eigenen Forschungsgebiet mit speziellen Algorithmen und Techniken behandelt wird. Die hier verwendeten Techniken unterstützen den effizienten Zugriff auf die Daten sowie deren sinnvolle Speicherung.

Die Extraktion des Wissens ist eine multidisziplinäre Aktivität, deren Techniken losgelöst von einem speziellen Anwendungsgebiet sind. Die Anwendungen müssen an einen solchen Prozess adaptiert und geeignete Algorithmen definiert werden. In dieser Arbeit wird der Prozess der Wissensextraktion aus großen, unstrukturierten Datenmengen im Bereich der Umweltinformatik betrachtet. Hierbei ist besonders der Bereich des Clustering der zugrunde liegenden Daten und der Evaluation in Form der Visualisierung von Interesse. Das Ziel des Clustering ist es, die zugrunde liegenden Daten anhand von Merkmalen in Gruppen mit einer hohen Ähnlichkeit

zu bündeln und somit die Auswertung der Daten zu erleichtern. Mit Hilfe der Visualisierung wird der kognitive Zugang der Nutzer zu dem neu gewonnenen Wissen erleichtert.

In den folgenden drei Kapiteln werden die grundlegenden Begriffe und Ansätze der Themenfelder Umweltinformatik, Clustering und Visualisierung vorgestellt. Die Arbeiten in den Kapitel 3, 4, 5 und 6 setzen auf diesen Grundlagen auf. In Kapitel 2.1.1 werden gängige Verfahren behandelt, mit denen ein Clustering von Daten durchgeführt werden kann und in Kapitel 2.1.2 werden Metriken vorgestellt. Metriken sind ein häufig verwendetes Verfahren, um die Ähnlichkeit zwischen Objekten zu modellieren.

Kapitel 2.2 behandelt das Themenfeld der Visualisierung. Es wird der Begriff der Informationsvisualisierung eingeführt (Kapitel 2.2.1), eine geeignete Methode für die Visualisierung räumlich nicht angeordneter Daten. Im Anschluss daran werden in Kapitel 2.2.2 Techniken dieses Teilgebietes der Visualisierung vorgestellt und schließlich in Kapitel 2.2.3 Systeme analysiert, die in diesem Bereich zum Einsatz kommen.

Kapitel 2.3 behandelt das in dieser Arbeit betrachtete Anwendungsfeld der Umweltinformatik mit den Grundlagen in Kapitel 2.3.1. Kapitel 2.3.2 beschreibt ein Teilgebiet der Umweltinformatik, die Stadtplanung und Kapitel 2.3.3 Systeme, die in diesem Teilgebiet zum Einsatz kommen.

2.1 Clustering

Unter dem Begriff Clustering (von englisch cluster = Schwarm, Haufen, Häufung) werden Verfahren verstanden, mit denen sich Strukturen in großen Datenbeständen finden lassen. Das Ziel des Clustering ist es, eine Datenmenge anhand von Merkmalen in Gruppen mit einer hohen Ähnlichkeit zu bündeln und somit die Auswertung der Daten zu erleichtern.

Der Begriff Clustering kann auf die folgenden zwei Arten beschrieben werden: Zerlegung einer Menge von Objekten (bzw. Daten) so in Teilmengen (Cluster), dass

- die Ähnlichkeit der Objekte innerhalb eines Clusters maximiert
- die Ähnlichkeit der Objekte verschiedener Cluster minimiert wird.

Üblicherweise wird die Ähnlichkeit zweier Objekte über ihre *Distanz* zueinander approximiert. Je größer die Distanz, desto unähnlicher sind sich die Objekte. In Abbildung 2.1 sind 2-dimensionale Clusterstrukturen mit verschiedenen Charakteristika dargestellt. Der Parameterraum ist jedoch häufig hochdimensional. Durch die Abstraktion in Parameter ist eine einheitliche Behandlung von Daten verschiedenster Anwendungsdomänen möglich.



Abbildung 2.1: Clusterstrukturen mit verschiedenen Charakteristika [ES00]

Es existieren verschiedene Verfahren zum clustern von Daten, die im folgenden Kapitel dargestellt werden.

2.1.1 Clusteranalyseverfahren

Clusteranalyseverfahren gewinnen in der Forschungspraxis zunehmend an Bedeutung. Sie werden in zahlreichen Wissenschaftsdisziplinen zur Lösung von Klassifikationsaufgaben eingesetzt. Primäres Ziel der Verfahren ist das Zusammenfassen einer Menge von Klassifikationsobjekten in homogene Gruppen (= Klassen, Cluster, Typen), wobei die Homogenität im Gegensatz zu der Heterogenität die Ähnlichkeit beschreibt.

Mit dem Begriff der homogenen Gruppe sind folgende Vorstellungen verbunden [Bac02]:

1. Die Klassifikationsobjekte, die einer homogenen Gruppe angehören, sollen untereinander ähnlich sein. Es soll Homogenität innerhalb der Cluster vorliegen.
2. Die Klassifikationsobjekte, die unterschiedlichen homogenen Gruppen angehören, sollen verschieden sein. Es soll Heterogenität zwischen den Clustern vorliegen.

Diese beiden Grundvorstellungen werden in der Literatur unterschiedlich bezeichnet. Cormack [Cor71] spricht von externer Isolierung (= Heterogenität zwischen den Clustern) und von interner Kohäsion (= Homogenität innerhalb der Cluster), die eine empirische Klassifikation erfüllen soll. Everitt [Eve01] spricht von „natürlichen“ Clustern, wenn beide Forderungen erfüllt sind. Diese werden als dichte Punktwolken in einem n -dimensionalen Raum beschrieben, die durch Regionen mit geringer Dichte voneinander getrennt sind. Neben diesen beiden Grundvorstellungen werden häufig aus forschungspraktischen, aber auch inhaltlichen Gründen weitere Anforderungen an die gesuchte Klassifikation gestellt.

Bacher [Bac02] hingegen unterscheidet unvollständige, deterministische und probabilistische Clusteranalyseverfahren. Unvollständige Clusteranalyseverfahren werden in der Literatur auch als geometrische Methoden [Gor81] bzw. Repräsentations- oder

Projektionsverfahren [JD88], [Opi80] bezeichnet.

Die Bildung von Clustern und die Zuordnung der Klassifikationsobjekte zu den Clustern erfolgt vom Anwender bei der Interpretation der räumlichen Darstellung. Deterministische Verfahren ordnen mit einer Wahrscheinlichkeit von 1 die Klassifikationsobjekte einem oder mehreren Clustern zu. Beispiele sind das K-Means-Verfahren oder hierarchische Verfahren. Bei den probabilistischen Verfahren werden die Klassifikationsobjekte mit einer zwischen 0 und 1 liegenden Wahrscheinlichkeit zugeordnet. Die unvollständigen Clusteranalyseverfahren sind auf einen maximal dreidimensionalen Raum beschränkt und somit ungeeignet, wenn ein höherdimensionaler Parameterraum vorliegt.

Han und Kamber [HK01] teilen Clusteranalyseverfahren in partitionierende, hierarchische, dichte-basierte, gitter-basierte und modell-basierte Verfahren ein. Diese Einteilung beruht auf der Vorgehensweise, die die Verfahren verwenden, um Cluster zu finden.

Ester und Sander [ES00] unterteilen dagegen Clusteranalyseverfahren nur in partitionierende und hierarchische Verfahren. Diese Einteilung beruht nicht auf der Vorgehensweise der jeweiligen Algorithmen, sondern auf der Art der Cluster, die die Algorithmen in den Eingabedaten finden. Es lassen sich alle in Han und Kamber vorgestellten Verfahren auch danach unterscheiden, ob sie ein Objekt nur jeweils einem Cluster zuordnen, oder ob sie eine Hierarchie von Clustern erzeugen. Es gibt beispielsweise dichte-basierende Verfahren, die ein partitionierendes Cluster erzeugen, und solche, die ein hierarchisches erzeugen.

Ein sehr effektives, häufig eingesetztes partitionierendes Clusteranalyseverfahren sind die *Voronoidiagramme* (siehe hierzu auch Kapitel 4.1.1). Partitionierende Clusteranalyseverfahren zerlegen eine Datenmenge unter folgenden Prämissen in k Cluster:

- jeder Cluster enthält mindestens ein Objekt
- jedes Objekt gehört zu genau einem Cluster

Für die Anwendung eines Clusteringverfahrens ist jedoch zunächst eine geeignete Modellierung der Ähnlichkeit zwischen Datenobjekten erforderlich, die im folgenden Kapitel eingeführt wird.

2.1.2 Metriken

Diese Modellierung der Ähnlichkeit erfolgt durch eine *Distanzfunktion*, die für Paare von Objekten definiert ist. Zur Definition der Distanz zwischen zwei Objekten werden direkte oder abgeleitete Eigenschaften der Objekte verwendet.

Unabhängig von der konkreten Form der Funktion müssen folgende mathematische

Bedingungen der Distanzfunktion d erfüllt sein, wobei d eine *Metrik* (griechisch: Zählung, Messung) beschreibt:

Definition 1 Sei X eine beliebige Menge. Eine Abbildung $d : X \times X \rightarrow \mathbb{R}^2$ heißt **Metrik**, wenn für beliebige Elemente $x, y, z \in X$ die folgenden axiomatischen Bedingungen erfüllt sind:

A1 : $d(x, x) = 0$;

A2 : aus $d(x, y) = 0$ folgt $x = y$ (Definitheit);

A3 : $d(x, y) = d(y, x)$ (Symmetrie);

A4 : $d(x, y) \leq d(x, z) + d(z, y)$ (Dreiecksungleichung);

Das Paar (X, d) bezeichnet einen **metrischen Raum**.

Typische Beispiele für Distanzfunktionen mit numerischen Attributswerten sind die *Minkowski-Metriken*. Je nach Wahl von p ergeben sich verschiedene, in der Praxis häufig verwendete Metriken.

Definition 2 Seien x und y Punkte im \mathbb{R}^2 , dann heißt

$$d_{L_p} = \sqrt[p]{|x_1 - y_1|^p + |x_2 - y_2|^p} \quad (2.1)$$

die **Minkowski-Metrik** von (x, y) .

Im Fall $d_{L_2} = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2}$ ergibt sich die *Euklidische Metrik*. Dies ist die in der Praxis am häufigsten verwendete Metrik.

In einem Metrischen Raum (M, d) sind die Raumeigenschaften der Menge M durch die Metrik d festgelegt. Durch die Herleitung einer *Umgebung* lassen sich die topologischen Eigenschaften aus der Metrik herausziehen, siehe hierzu auch Abbildung 2.2.

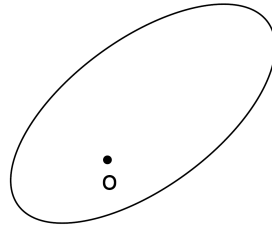
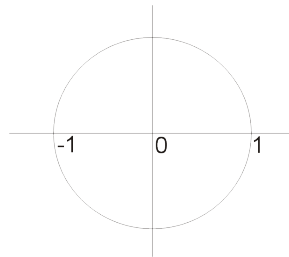
Definition 3 Für alle $x_0 \in M$, $\delta > 0$ bezeichnet

$$U_\delta(x_0) := \{x \in M \mid d(x, x_0) < \delta\} \quad (2.2)$$

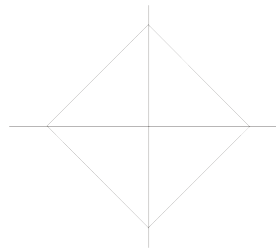
die δ -*Umgebung* von x_0 .

Zur graphischen Veranschaulichung der Metriken wird ein Koordinatensystem dargestellt, in welchem um den Nullpunkt der jeweilige Einheitsabstand als sogenannte *Einheitssphäre* aufgetragen wird. Die Einheitssphäre ist die Grenzlinie einer Einheitsumgebung um den Nullpunkt. Die Einheitsumgebung der Euklidischen Metrik hat die in Abbildung 2.3 dargestellte Struktur.

Die Euklidische Metrik eignet sich aber nicht für jedes Einsatzgebiet. Beispielsweise liegt in Manhattan nur ein orthogonales Straßengitter vor. Dadurch werden Abstände durch die Längen der Straßen definiert und nicht, wie beim Euklidischen Abstand, durch die Luftlinie. Hierfür eignet sich die in Abbildung 2.4

Abbildung 2.2: Konturlinie der Umgebung $U_\delta(0)$ Abbildung 2.3: Einheitsumgebung $U_1(0)$ der Euklidischen Metrik

dargestellte *Manhattan-Metrik*, in der der Abstand zwischen zwei Punkten gegeben ist durch $d_{L_1} = |x_1 - y_1| + |x_2 - y_2|$. Im Fall, dass p gegen Unendlich geht,

Abbildung 2.4: Einheitsumgebung $U_1(0)$ der L_1 -Metrik

wird von der *Maximums-Norm* (siehe Abbildung 2.5) gesprochen, wobei $d_{L_\infty} = \max(|x_1 - y_1|, |x_2 - y_2|)$ entspricht.

Die vorgestellten Minkowski-Metriken sind immer streng symmetrisch zum Nullpunkt. Abbildung 2.6 zeigt die Entwicklung der Minkowski-Metriken für einen ansteigenden p -Wert. Es ist für ein wachsendes p deutlich zu erkennen, wie sich die Konturlinien um den Mittelpunkt von der Manhattan-Metrik über die Euklidische Metrik nach außen biegen, bis sie schließlich die Rechtecksform der Maximum-Norm für $p \rightarrow \infty$ asymptotisch nähern. Des weiteren ist zu erkennen, dass die Einheitsumgebungen der Minkowski-Metriken stets konvex sind.

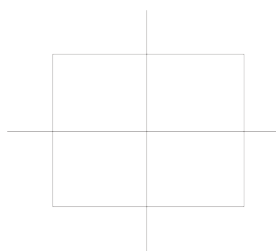
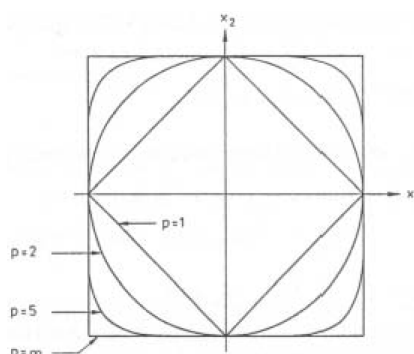
Abbildung 2.5: Einheitsumgebung $U_1(0)$ der Maximumnorm

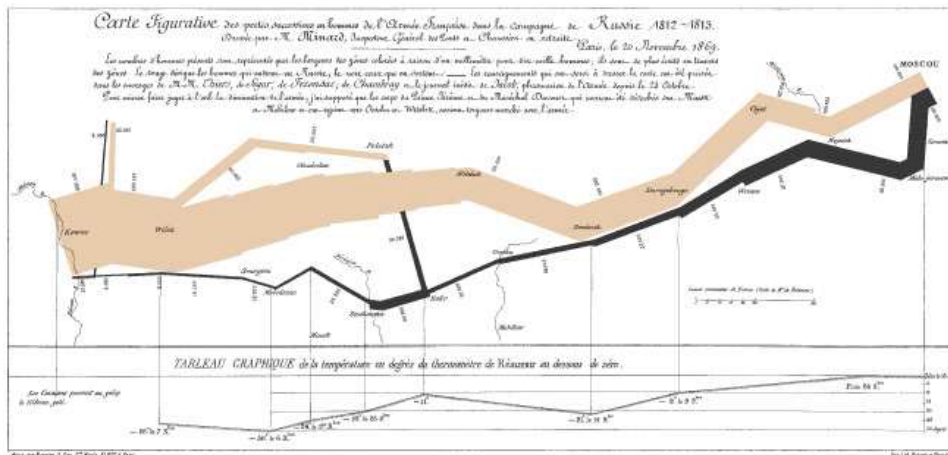
Abbildung 2.6: Minkowski-Konturen

2.2 Visualisierung

Seit Beginn der Menschheit spielt die Visualisierung eine zentrale und bedeutende Rolle bei der Übertragung und dem Verständnis von Information. Die Gründe für dieses Phänomen beruhen dabei auf der Physiologie des menschlichen Auges und den mit ihm verbundenen visuellen Cortex. Von allen Sinnesorganen des Menschen besitzt der Sehapparat die größte Bandbreite bei der Aufnahme von Informationen. Diese Tatsache wird durch die allgemein bekannte Redewendung „ein Bild sagt mehr als tausend Worte“ hervorgehoben.

Welche Bedeutung visuelle Metaphern für den Informationsaustausch des Menschen besitzen, ist nicht erst seit der Entwicklung der Computergraphik und der computergestützten Visualisierung bekannt.

Charles Joseph Minard (1781-1870), ein französischer Ingenieur, zeigte bereits vor einhundert Jahren in einer graphischen Darstellung (siehe Abbildung 2.7) des Rußlandfeldzuges von Napoleon von 1812-1813, wie multivariate Daten graphisch dargestellt werden können. Fünf verschiedene Parameter sind in einer einzigen Graphik visuell miteinander kombiniert: die Größe der Armee, der Aufenthaltsort der Truppen zu einem bestimmten Zeitpunkt, die Bewegungsrichtung der Armee und die Temperatur auf dem Rückzug der Armee nach der Belagerung von Moskau. Mit kei-



Napoleon's March to Moscow The War of 1812

Charles Joseph Minard

This chart of Charles Joseph Minard (left, above), the French engineer, shows the terrible fate of Napoleon's army in Russia. Described by E. J. Mearns in working into the pen of the historian in his limited obituary, this combination of data map and time-series diagram in 1869, portrays the devastating losses suffered by Napoleon's Russian campaign of 1812. Beginning at the left on the Polish-Russian border near the Niemen River, the thick black band shows the size of the army (120,000 men) as it moved toward Moscow in June 1812. The width of the band indicates the size of the army at each place on the map. In September the army reached Moscow, which was by then isolated and encircled with Russian men. The path of Napoleon's retreat from Moscow is depicted by the thick, orange band, which is marked by a temperature

scale and dots at the bottom of the chart. It was a bitterly cold winter, and many froze to the south east of Russia. As the graphic shows, the crossing of the Berezina River was a disaster, and the army finally straggled back into Poland with only 10,000 men remaining. Also shown are the movements of auxiliary troops, as they sought to protect the rear and the flank of the retreating army. Minard's graphic tells a rich, colorful story with no extraneous data, far more enlightening than just a single number heading along over time. The variables are plotted: the size of the army, its location on a two-dimensional surface, direction of the army's movement, and temperature on various dates during the retreat from Moscow. It may well be the best example of graphic over time.

Edward R. Tufte, The Visual Display of Quantitative Information, Cheshire Press, Box 602, Cheshire, Connecticut 06610.

Abbildung 2.7: Joseph Minard: Karte des Rußlandfeldzuges Napoleons von 1812-1813

ner anderen als der visuellen Darstellung gelingt es, so viele verschiedene Aspekte und Beziehungen auf einen Blick sichtbar zu machen.

Nach Däßler [DäB99] lassen sich unabhängig vom Anwendungszweck und der Präsentationsform vier elementare Aufgaben der Visualisierung formulieren:

1. Symbole, Diagramme oder Animationen helfen komplexe Prozessabläufe und Objektbeziehungen in der Realwelt zu veranschaulichen und gegebenenfalls zu vereinfachen.
2. Visualisierung vereinfacht den Zugang zu Massendaten, z.B. durch Klassifikation und Datenstrukturierung.
3. Visualisierung hilft bei der Analyse und Interpretation von Daten, bei der Sichtbarmachung verborgener Trends sowie bei der Mustererkennung.
4. Visualisierung entspricht der Neigung der menschlichen Spezies und unserer Kultur, visuelle Informationsprozesse und Repräsentationsformen zu bevorzugen. Aus der Gehirnforschung ist darüber hinaus bekannt, dass sich Visualisierung positiv auf die Gedächtnisleistung und auf die menschliche Informationsaufnahme auswirkt.

Däßlers Definition des Ausgabenfeldes ist relativ allgemein gehalten und macht keinerlei Aussage darüber, welche Daten, Objekte oder Prozesse visualisiert werden.

Daher stellt sich die Frage, in welchen Bereichen heutzutage rechnergestützte Visualisierung zur Anwendung kommt. Scheuermann [Sch04a] gibt hierzu eine Unterteilung der Visualisierung in drei Teilgebiete an:

1. **Scientific Visualization**

Visualisierung von Daten und Modellen in Natur-, Technik- und Lebenswissenschaften. Die Daten sind hierbei numerisch und anwendungsspezifisch im ein-, zwei-, drei- oder vierdimensionalen Raum angeordnet. Beispiele: Skalar-, Vektor-, Tensorfeldvisualisierung.

2. **Software Visualization**

Visualisierung der Struktur und des Ablaufes von Algorithmen und Softwaresystemen. Beispiele: Klassen-, Sequenzdiagramme in UML.

3. **Information Visualization**

Visualisierung von räumlich nicht angeordneten Daten (oftmals: Tabellen, Graphen oder Datenbanken). Beispiele: Fahrpläne, meteorologische Wetterreihen.

Die in dieser Arbeit behandelten Daten fallen in die dritte Kategorie. Im folgenden wird deshalb auf die Informationsvisualisierung näher eingegangen.

2.2.1 Informationsvisualisierung

Zu einem relativ neuen, interdisziplinären Forschungsgebiet zählt die Informationsvisualisierung. Sie beschäftigt sich mit der Darstellung abstrakter Informationen. Da der Umfang verfügbarer Daten stetig zunimmt, wird es immer schwieriger, diese sehr großen und heterogenen Datenmengen zu untersuchen und daraus sinnvolle Informationen abzuleiten. Die Informationsvisualisierung verwendet hierbei Methoden und Erkenntnisse der Informatik, Statistik, Data Mining und der Kognitionswissenschaft. Von der Informationsvisualisierung wird weiterhin angestrebt, die Mensch-Computer-Interaktion zu verbessern. In [Wik06] wird der Begriff Informationsvisualisierung folgendermaßen definiert:

Definition 4 *Information visualization is the use of interactive, sensory representations, typically visual, of abstract data to reinforce cognition*

Informationsvisualisierung beinhaltet demnach die computergestützte Aufbereitung und die visuelle Repräsentation abstrakter Informationen mit dem Ziel, den kognitiven Zugang zu elektronisch gespeicherten Daten zu erleichtern.

Die Aufgabe der Informationsvisualisierung wird durch die folgenden vier Zitate zu Gegenstand und Zielen der Informationsvisualisierung noch etwas genauer beschrieben:

- Information Visualization is the use of computer-supported interactive visual representation of abstract data to amplify cognition [Däß99].
- Information Visualization combines aspects of scientific visualization, human-computer interfaces, data mining, imaging and graphics [GE97].
- Information Visualization is the process of binding information that is not inherently spatial into a medium which the user can view and perceive [Däß99].
- The greatest value of a picture is when it forces us to notice what we never expected to see [Tuk04].

Die Informationsvisualisierung stellt Methoden, Konzepte und Werkzeuge bereit, um große Datenmengen visuell erforschen und auswerten zu können. Dabei wird die Eigenschaft des Menschen ausgenutzt, Muster und Strukturen visuell schnell zu erfassen. Dieser kognitive Zugang zu den Daten wird durch spezielle Visualisierungstechniken weiter unterstützt. Hauptziel der Informationsvisualisierung ist somit das Erkennen von Mustern, Strukturen, Klassen und Relationen im mehrdimensionalen abstrakten Datenraum anhand von Bildern.

Die Methoden der Informationsvisualisierung lassen sich in drei Kategorien unterteilen:

1. **Präsentation**

Es existieren fest definierte Fakten, welche durch geeignete Visualisierungstechniken dargestellt werden.

2. **Bestätigende Analyse**

Es existieren vom Benutzer definierte Hypothesen zu den jeweiligen Datensätzen. Aufgabe der Visualisierung ist es, die Daten in einer geeigneten Form darzustellen, die eine Bestätigung oder Entkräftung der aufgestellten Hypothese erlaubt.

3. **Erkundende Analyse**

Ausgangspunkt sind Datensätze, über die weder Fakten noch Hypothesen bekannt sind. Durch eine visuelle, interaktive und im Allgemeinen ungerichtete Suche nach Trends, Strukturen und Mustern ist der Benutzer in der Lage, neue Informationen aus den Daten zu gewinnen und letztendlich eine Hypothese zu definieren.

Nach Shneiderman [Shn96] können Datensätze ein-, zwei- oder auch multidimensional sein oder auch aus komplexeren Datentypen wie Texten, Hypertexten, Hierarchien, Graphen oder Algorithmen bestehen. Alternativ werden diese Datensätze auch als uni-, bi-, oder multivariate Datensätze bezeichnet.

Typische Vertreter von eindimensionalen Datensätzen sind zeitabhängige Daten. Dabei können jedem Punkt auf der Zeitskala ein oder mehrere Messwerte zugeordnet werden.

Zwei- oder dreidimensionale Datensätze bestehen aus zwei bzw. drei definierten Variablen. Ein typisches Beispiel für solche Datensätze stellen geo-graphische Karten dar. Zwei- und dreidimensionale Datensätze werden in der Regel durch einfache x-y- bzw. x-y-z-Plots visualisiert. Auch wenn diese Darstellungsform einfach erscheint, können mit zunehmender Datengröße bzw. -dichte schnell übersichtliche und teilweise unbrauchbare Visualisierungen resultieren.

Häufig bestehen Datensätze aus mehr als drei Dimensionen und können daher nicht mit Hilfe von zwei- oder dreidimensionalen Plots dargestellt werden. Multidimensionale Datensätze enthalten in der Regel mehrere Hundert bis Tausend Dateneinträge. Diese Daten können nur mit Hilfe weiter entwickelter Visualisierungstechniken dargestellt werden, da das effektive Mapping der zahlreichen Dimensionen auf einen zweidimensionalen Bildschirm ein schwieriger Vorgang ist.

Nicht alle Datentypen können durch Angabe der Dimensionalität beschrieben werden. Dies trifft beispielsweise auf digitale Texte und Hypertexte zu, deren Analyse vor allem im Bereich des World Wide Web hohe Bedeutung beizumessen ist. Da diese Datentypen nicht sinnvoll in Form von Zahlen dargestellt werden können, eignen sich viele Visualisierungstechniken nicht zur Darstellung dieser Daten. Eine weitere Klasse von Datensätzen stellen Hierarchien und Graphen dar, die auf Beziehungen zwischen einzelnen Datenpunkten basieren. Algorithmen zählen ebenfalls zu der Klasse der speziellen Datenformen und erfordern besondere Visualisierungstechniken.

2.2.2 Techniken

Im Laufe der letzten dreißig Jahre wurden zahlreiche Techniken zur Visualisierung von Informationen und Daten entwickelt, wobei die Anzahl der Visualisierungstechniken auch in Zukunft weiter ansteigen wird.

Nach Keim [Kei97] werden multivariate Visualisierungstechniken in Abhängigkeit ihrer zugrunde liegenden Darstellungsprinzipien in fünf Kategorien unterteilt:

1. Geometrie-basierte Techniken
2. Icon- und Glyph-basierte Techniken
3. Pixel- und Voxel-basierte Systeme
4. Hierarchische Techniken
5. Techniken, die auf Graphen basieren

Darüber hinaus existieren zahlreiche hybride Ansätze, die sich durch Kombination verschiedener Visualisierungstechniken aus den genannten Bereichen ergeben.

Geometrie-basierte Techniken

Die Grundidee geometrischer Visualisierungstechniken basiert auf der Nutzung geometrischer Transformationen und Projektionen, um so eine bestmögliche Darstellung multidimensionaler Daten zu realisieren. Dabei können sowohl zwei- als auch dreidimensionale Darstellungstechniken zum Einsatz kommen.

Zur Klasse dieser Techniken zählen zum Beispiel Liniengraphen [HG02], Scatterplots bzw. Scatterplot-Matrizen [HG02],[Cle93], Histogramm-Darstellungen [HG02], Landscapes [Wri95], Prosection Views [FB94], HyperSclice [JvW93], Hyperbox [AC91], Parallele Koordinaten [ID90], RadViz und GridViz [Hof99]. Bei den *Scatterplots*

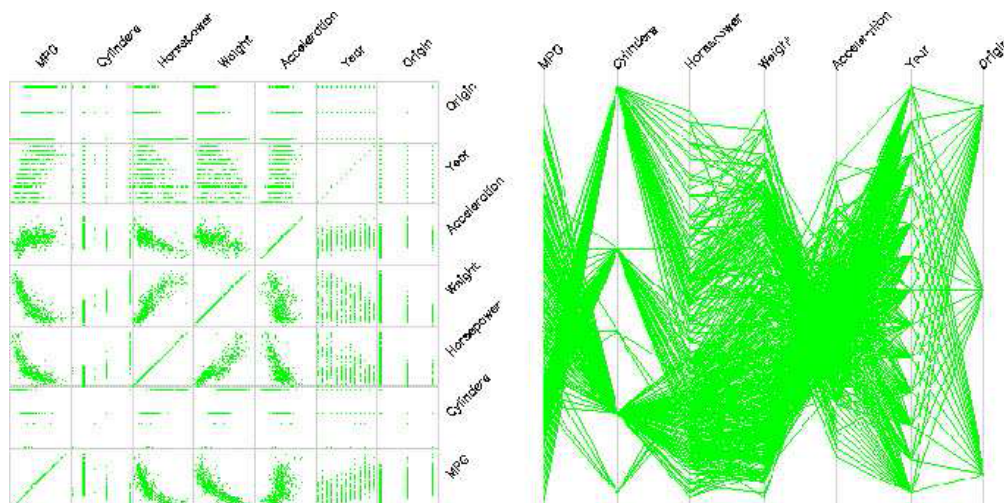


Abbildung 2.8: Scatterplot-Matrizen und Parallele Koordinaten

kommen sowohl zwei- als auch dreidimensionale Darstellungen zum Einsatz. Im Regelfall können mittels Scatterplots zwei bzw. drei Datendimensionen dargestellt werden, wobei jede Datendimension auf eine der zwei bzw. drei orthogonalen Achsen abgebildet werden. Um Datensätze mit höherer Dimensionalität zu visualisieren, werden unter anderem sogenannte *Scatterplot-Matrizen* verwendet. Dabei kann beispielsweise ein vierdimensionaler Datensatz durch eine 4x4-Matrix von Scatterplots dargestellt werden.

Die *Parallelen Koordinaten* bilden den k -dimensionalen Raum auf die zwei Bildschirmdimensionen ab, indem jeder Datensatz als polygonale Linie in einem System paralleler Achsen dargestellt wird. Jede der k äquidistanten parallelen Achsen entspricht einer Dimension und ist linear skaliert vom Minimum bis zum Maximum der

auftretenden Datenwerte der Dimension. Jeder Datensatz wird durch eine polygonale Linie repräsentiert, welche die vertikalen Achsen (Dimensionen) an der Stelle schneidet, die dem jeweiligen Datenwert der dazugehörigen Dimension entspricht.

Mit beiden Methoden lassen sich sehr viele Dimensionen in einer einzigen Darstellung kombinieren und erzeugen zudem einen guten Überblick über die Daten. Des Weiteren kann der Benutzer leicht absolute Werte wahrnehmen. Probleme können jedoch bei beiden Methoden dadurch entstehen, dass identische Datenrekorde nur einen einzigen Eintrag bzw. eine einzige polygonale Linie erzeugen.

Icon- und Glyph-basierte Techniken

Eine andere Klasse von explorativen Visualisierungstechniken stellen die so genannten Icon- bzw. Glyph-basierten Ansätze dar. Diese Techniken werden vor allem zur Darstellung von diskreten, multivariaten Daten eingesetzt.

Ein Glyph ist dabei ein graphisches Objekt, welches ein einzelnes multivariates Datenobjekt repräsentiert. Bei der Generierung der Icons bzw. Glyphen werden die diversen Datendimensionen eines Datensatzes in systematischer Weise den verschiedenen graphischen Attributen wie Form, Farbe, Größe, Orientierung, Textur, etc. des graphischen Objektes zugeordnet. Dieses Mapping der Dimensionen auf die sogenannten retinalen Eigenschaften wird auch als visuelles Mapping bezeichnet [Ber83]. Abbildung 2.9 enthält eine Übersicht der wichtigsten retinalen Eigenschaften. Neben der Anzahl der möglichen Dimensionen, die mit Hilfe der einzelnen graphischen Eigenschaften beschrieben werden können, enthält die Abbildung auch Beispiele, die den Gebrauch von Glyphen bei kontinuierlichen, numerischen sowie diskreten bzw. kategorischen Daten zeigt. Viele der graphischen Attribute hängen dabei direkt voneinander ab. So setzen beispielsweise Texturen mindestens eine Farbe zur Darstellung voraus und Blinken kann mit der Darstellung von Bewegungen wechselwirken. Im Allgemeinen ist daher eine sinnvolle Darstellung auf acht Dimensionen beschränkt.

Insbesondere bei Vorlage diskreter Daten muss eine gute Unterscheidung der einzelnen Glyphen gewährleistet sein, um eine sinnvolle Visualisierung zu erhalten. Dabei müssen unter anderem verschiedene psychologische (z.B. unterschiedliche Wahrnehmung von Farben und Orientierungen) und psychologische Aspekte (z.B. Rot-Grün-Blindheit) berücksichtigt werden. Eine gute Einführung in die Thematik liefert das Buch von Ware [War99].

Zu der Klasse der Icon- und Glyph-basierten Techniken gehören beispielsweise Chernoffs Faces [Che71], Stick-Figures [RP88], Shape Coding [Bed90], Color Icons [Lev91] sowie Star Glyphs [CCKT83].

Mit den *Chernoff Faces* lassen sich durch Abbildung der einzelnen Datendimensionen auf markante Merkmale eines Gesichtes wie z.B. Größe der Augen, Neigung der







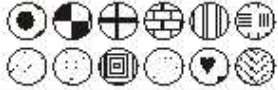


Graphisches Attribut	Dimensionalität	Kontinuierliche Daten Quantitatives Mapping	Diskrete Daten Nominales Mapping
Farbe	max. 3 Dimensionen (3 bei color opponent)		
Form	max. 3 Dimensionen (x, y, z)		
Orientierung	3 Dimensionen (x, y, z)		
Textur	3 Dimensionen (Kontrast, Größe, Orientierung)	Textur-Morphing (ungünstig)	
Bewegung	mind. 2-3 Dimensionen		
Blinken	1 Dimension	Stufenlose Blinkgeschwindigkeit	Blinken, Nicht-Blinken, klar definierte Stufen

Abbildung 2.9: Retinale Eigenschaften in der Glyph-basierten Visualisierung

Augenbrauen, Größe und Form des Mundes usw. sehr viele Dimensionen darstellen. Da Menschen dazu neigen die resultierenden Gesichter als Ganzes wahrzunehmen, eignet sich diese Technik sehr gut zur Darstellung relativer Werte und als Überblick. Diese Art der Repräsentation basiert auf der Metapher von Emotionen, da die resultierenden Gesichter unterschiedliche Stimmungen vermitteln. Die Darstellung spezifischer Datensätze erzeugt fröhliche, traurige oder verärgerte Gesichter.

Eine Interpretation dieser Gesichter kann aber auch einige Probleme mit sich bringen. Soll beispielsweise der Zustand eines Patienten visualisiert werden, kann der Gesichtsausdruck des Chernoff Faces den Benutzer dazu verleiten jedes fröhliche Gesicht mit einem guten Zustand und jedes traurige Gesicht mit einem schlechten Zustand des Patienten zu verbinden. Der Gesichtsausdruck hängt dabei vor allem von der Kombination der gewählten Fragen durch den Benutzer und der Abbildung der Attribute des Gesichtes ab. Die Daten eines einzelnen Patienten können so sehr unterschiedliche Gesichtsausdrücke erzeugen.

Bei den *Star Glyphs* werden die einzelnen Dimensionen durch gleichwinklige, im Uhrzeigersinn angeordnete Speichen repräsentiert. Die äußeren Enden der Speichen entsprechen dabei den maximalen Werten der Datendimensionen. Schließlich werden die äußeren Enden noch durch eine polygonale Linie verbunden.

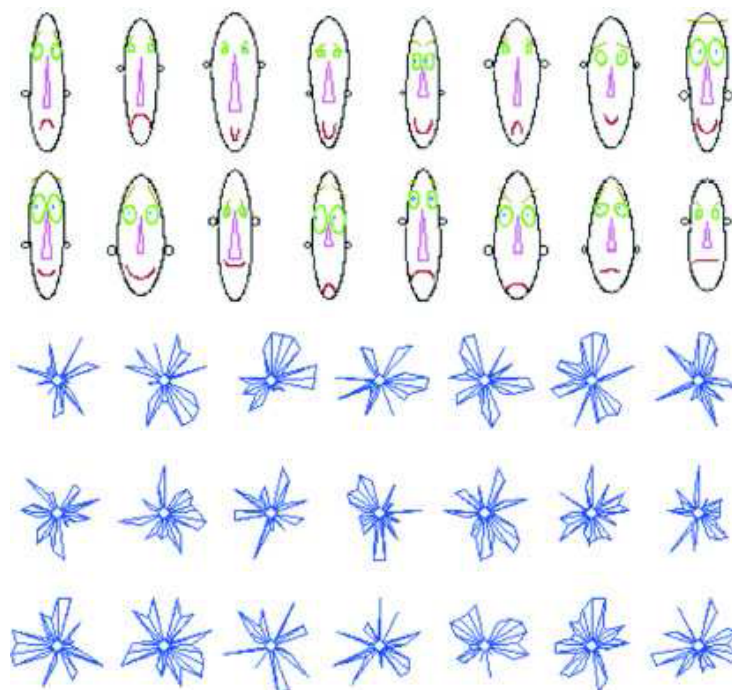


Abbildung 2.10: Chernoff Faces und Star Glyphs

Pixel- und Voxel-basierte Techniken

In Pixel-basierten Ansätzen wird jeder einzelne Dimensionswert einem farbigen Pixel zugeordnet. Die Pixel werden darüber hinaus in Abhängigkeit von den jeweiligen Dimensionen gruppiert und in separate Regionen dargestellt [Kei97]. Die Werte der einzelnen Dimensionen werden durch die Farbe des Pixels repräsentiert. Da lediglich ein Pixel pro Datenobjekt benötigt wird, kann mit Hilfe dieser Technik die derzeit höchste Anzahl an Datenpunkten gleichzeitig dargestellt werden.

Die bekanntesten Vertreter dieser Visualisierungsform sind die so genannte Recursive Pattern-Technik [KKA95] und die Circle Segment-Technik [AK96]. Die *Recursive Pattern-Technik* basiert auf einer rekursiven Verallgemeinerung einer zeilen- und spaltenorientierten Anordnung der Pixel. Auf einer Rekursionsstufe werden die Pixel von links nach rechts, dann eine Zeile tiefer von rechts nach links usw. angeordnet. Die Anordnung der Pixel erfolgt nach dem gleichen Prinzip auf allen Rekursionsstufen mit dem Unterschied, dass die Basiselemente, die auf der nächsthöheren Rekursionsstufe angeordnet werden, aus Pixelarrays bestehen, die sich aus den tieferen Rekursionsstufen ergeben. Abbildung 2.11 links zeigt die tagesgenauen Kurse der 100 Aktien des FAZ-Index über einen Zeitraum von 20 Jahren (Januar 1974 - April 1995). Die Abbildung der Kurswerte auf der Farbskala wurde so gewählt, dass helle Farben hohen Kursen und dunkle Farben niedrigen Kursen entsprechen.

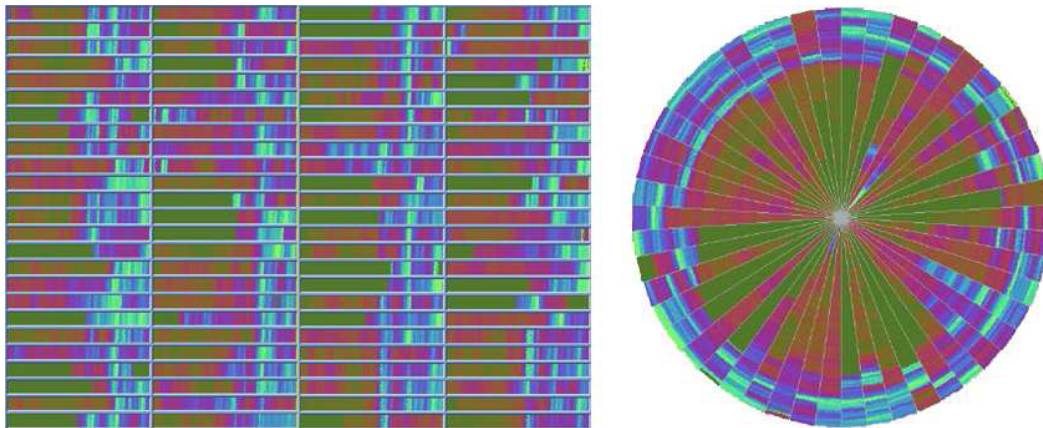


Abbildung 2.11: Recursive Pattern und Circle Segment

Die *Circle Segment-Technik* ordnet die Pixel, die zu einem Attribut gehören, in den Segmenten eines Kreises an. Die Anordnung der Pixel innerhalb eines Segmentes beginnt im Zentrum des Kreises, endet am Kreisrand und erfolgt nach einem links-rechts Schema orthogonal zur Segment-Halbierenden. Die Visualisierung in Abbildung 2.11 rechts zeigt die tagesgenauen Kurse von 50 Aktien des FAZ-Index über einen Zeitraum von 20 Jahren. Deutlich zu erkennen sind ähnliche Aktienkursverläufe, die Hochpreisphasen (helle kreisförmige Ringe) sowie Aktien, die sich gegen den Trend verhalten.

Hierarchische und Graph-basierte Techniken

Hierarchische Techniken stellen Daten in Form von hierarchisch aufgeteilten Unter-einheiten dar. Im Fall von multidimensionalen Datensätzen dienen dabei selektierte Dimensionen zur Aufteilung des Datensatzes und zum Aufbau der Hierarchie.

Zur Klasse dieser Techniken zählen zum Beispiel das Dimensional Stacking [KKA95] und der Cone Tree [GR91]. Die Idee der *Dimensional Stacking-Technik* (Abbildung 2.12 links) ist die Einbettung eines Koordinatensystems in ein anderes Koordinatensystem. Die visuelle Darstellung wird durch die Aufteilung des äußersten Koordinatensystemes in rechteckige Zellen erzeugt. Innerhalb dieser rechteckigen Zellen spannen zwei weitere Attribute ein inneres Koordinatensystem auf. Das innere Koordinatensystem wird dann wiederum in rechteckige Zellen zerlegt usw.

Cone Trees (Abbildung 2.12 rechts) stellen hierarchische Informationen auf einer Baumstruktur dar, die in drei Dimensionen ausgelegt ist. Die räumliche Tiefe wird ausgenutzt, indem Kegelkreise teils im Vordergrund, teils im Hintergrund abgebildet werden. Die Elemente einer Hierarchieebene werden bei den Cone Trees entlang des Umfanges eines Kreises angeordnet. Dadurch werden Knoten und Kanten der Kegel

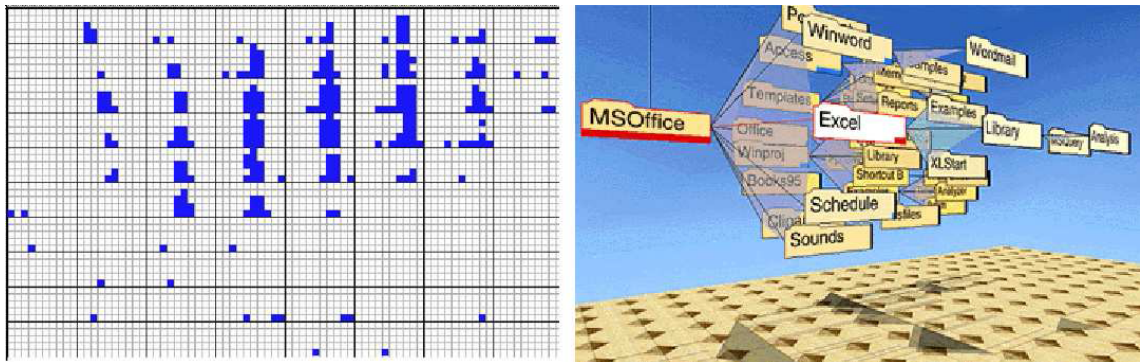


Abbildung 2.12: Dimensional Stacking und Cone Tree

teilweise überlagert. Die dritte Dimension des Darstellungsraumes wird bestmöglich genutzt, indem die Kegel selbst transparent sind. Der Anwender hat außerdem noch Interaktionsmöglichkeiten, um lokale Details sowie die gesamte Informationsstruktur zu betrachten.

2.2.3 Systeme

Noch vor der Einführung des Begriffs der Informationsvisualisierung wurden Mitte der 80er Jahre an der Universität von Singapur Überlegungen angestellt, wie neue Virtual Reality Techniken benutzt werden könnten, um Informationsräume aufzubauen, in denen der Benutzer abstrakt navigieren kann. Aus diesen Überlegungen entwickelte Kim Fairchild im Jahre 1988 SemNet [FPF88], eines der ersten Informationsvisualisierungssysteme überhaupt. Bis zum heutigen Zeitpunkt wurde eine Vielzahl von Visualisierungstechniken für die unterschiedlichsten Anwendungsfelder entwickelt. Dementsprechend groß ist auch die Menge an Systemen, die diese Techniken umsetzten. An dieser Stelle werden einige interessante Anwendungen vorgestellt.

Webvisualisierung

Seit 1994 wurde eine große Anzahl von Anwendungen zur Visualisierung des World-Wide-Web entwickelt. Dabei stand die Darstellung von Hyperlinkschematas bestimmter Informationsdomänen im Netz mittels geeigneter Navigationsoberflächen im Vordergrund. An der University of Minnesota wurde dazu 1994 das Tool WebViz [Mun95] vorgestellt. Dabei werden die gewünschten Informationsstrukturen durch die Projektion sogenannter Hyperbolischer Bäume (eine Variation des beschriebenen Cone Trees) auf die Oberfläche einer Kugel projiziert. Diese Projektion erlaubt die saubere Darstellung einer Informationsmenge, die im euklidischen Raum hingegen sehr unübersichtlich angeordnet wäre. Die Informationsstruktur liegt im VRML-Format vor und wird durch den frei verfügbaren 3D-Webbrowser WebOOGL [Web94]

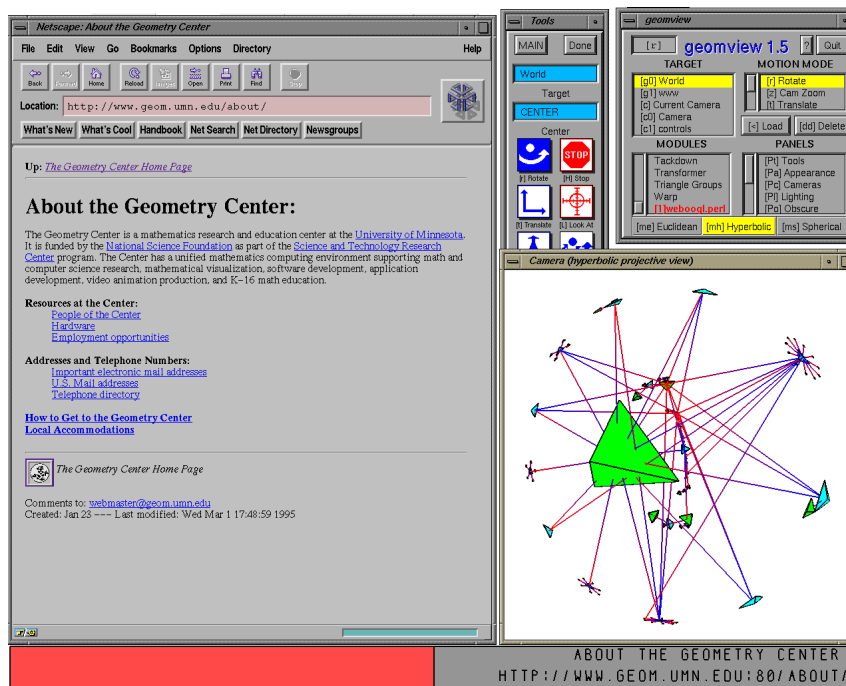


Abbildung 2.13: Webvisualisierung durch hyperbolische Darstellung

dargestellt. Abbildung 2.13 zeigt die gleichzeitige Darstellung eines Dokumentes in einem normalen Webbrowser und die hyperbolische dreidimensionale Darstellung. Interaktionen in der 3D-Darstellung beeinflussen dabei die 2D-Darstellung und umgekehrt.

Fly Through-Navigation

Besondere Popularität erlangten in den letzten zwei Jahren interaktive Benutzerschnittstellen, bei denen der Anwender die Darstellung des Informationsraumes stufenlos, z.B. durch Zoomen verändern kann. Die visuelle Exploration wird als *Fly Through-Navigation* bezeichnet. Das erste Interface, das solch eine Möglichkeit vorsah, war das 1995 von Apple entwickelte *Hotsauce*. Die Entwicklung wurde 1997 von Apple eingestellt, das Interface wurde allerdings in modifizierter Form kommerziell im Browser *Perspecta View* der Firma Perspecta angeboten. Grundidee ist, die zu visualisierenden Daten sinnvoll zu strukturieren und in das Datenformat MCF (*Meta Content Format*) zu überführen. Diese Daten werden dann visualisiert.

Visuelle Repräsentation von Textdokumenten

Zu den interessantesten aber auch komplexesten Gebieten in der Informationsvisualisierung gehört die visuelle Repräsentation von Textdokumenten. Vor allem in Arbeitsgruppen in den USA wird an solchen grafischen Textretrievaloberflächen ge-

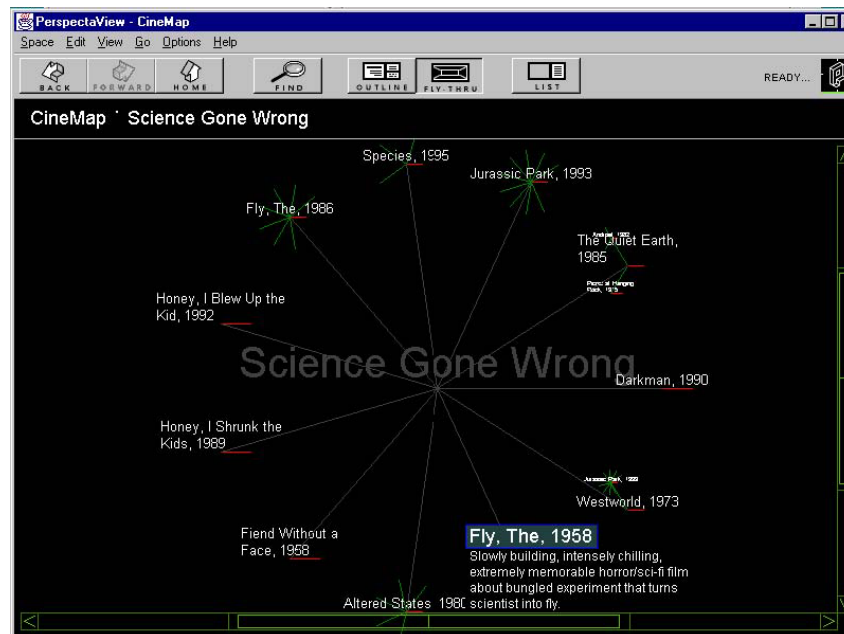


Abbildung 2.14: Flug durch den Informationsraum von Kaufvideokassetten mit Perspecta View

arbeitet. Die häufigste und auch bekannteste visuelle Metapher ist dabei die sogenannte *Dokumentenkarte*. Dabei werden Dokument-Dokument Relationen als Cluster entweder in einer Punktwolkendarstellung oder als thematische Karte dargestellt. Die Software IN-SPIRE (Spatial Pradigm for Information Retrieval) des Pacific Northwest National Laboratory [Lab06] verwendet beide Darstellungstechniken. Links ist eine Punktwolkendarstellung (*Galaxy*) einer Dokumentensammlung zu sehen. Nahe beieinander liegende Dokumentenpunkte entsprechen inhaltlich ähnlichen Dokumenten. Auf der rechten Seite ist eine thematische Kartendarstellung (*ThemeScope*) einer Dokumentensammlung zu sehen. Die hellen Bereiche in dem Terrain symbolisieren als Berge eine große Dichte von Dokumenten in der Umgebung charakteristischer Begriffe. In beiden Darstellungen kann der Anwender zoomen, die Darstellung drehen und verschieben sowie gewünschte Details näher betrachten.

2.3 Umweltinformatik

Das in dieser Arbeit betrachtete Anwendungsgebiet ist die Umweltinformatik. In diesem Kapitel wird deshalb zunächst der Begriff der Umweltinformatik erläutert und ihre Schwerpunkte vorgestellt (Kapitel 2.3.1). Für die praktische Anwendung konzentriert sich diese Arbeit auf die Flächennutzung im innerstädtischen Bereich.

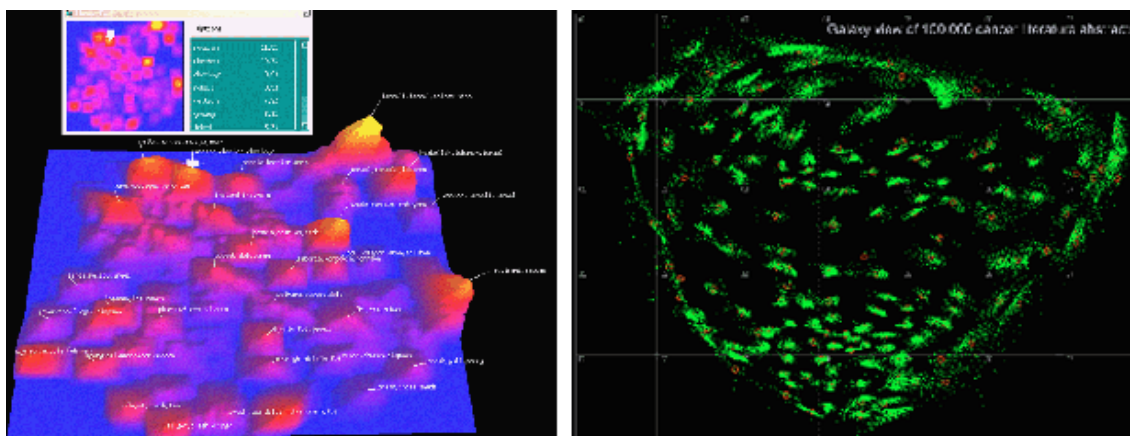


Abbildung 2.15: Punktwolkendarstellung und thematische Karte einer Dokumentensammlung durch IN-SPIRE

Hierfür wird in Kapitel 2.3.2 der Begriff der Stadtplanung eingeführt und deren Aufgaben und Ziele dargestellt. Eine aktuelle Tendenz innerhalb der Stadtplanung beschäftigt sich mit der Konversion von Flächen, die ebenfalls in diesem Kapitel eingeführt wird. Am Beispiel der Konversion werden in Kapitel 6.3.1 die in dieser Arbeit entwickelten Methoden und Prozesse umgesetzt.

In Kapitel 2.3.3 werden gegenwärtig eingesetzte raumbezogene Informationssysteme vorgestellt.

2.3.1 Grundlagen

Die Umweltinformatik ist eine relativ neue Teildisziplin der angewandten Informatik und beschäftigt sich interdisziplinär mit der Analyse und Bewertung von Umweltsachverhalten mit Mitteln der Informatik. Sie verfolgt damit das Ziel, einen Beitrag zur Untersuchung, Behebung, Vermeidung oder Minimierung von Umweltbelastungen und Umweltschäden zu leisten.

Schwerpunkte sind die Verwendung von Geographischen Informationssystemen (GIS), Datenbanksystemen und Simulationsprogrammen.

Typische Anwendungen der Umweltinformatik sind:

- Informationsauswertung für ökonomische und ökologische Optimierung
- Ausbreitungsmodelle von Schadstoffen
- Simulation von solaren und energiewandelnden Systemen

Die betriebliche Umweltinformatik greift besonders die ökonomischen Aspekte der Umweltinformatik auf und steht somit inhaltlich zwischen der Umwelt- und Wirtschaftsinformatik.

Neben der vorrangig ökologisch orientierten Umweltinformatik (beispielsweise Umweltdatenbanken für die verschiedenen Umweltmedien, Umweltinformationssysteme, ökologische Anwendungen raumbezogener Informationssysteme (GIS), Visualisierungsmethoden für komplexe Umweltdaten) hat sich in den letzten 10 Jahren eine stärkere Verlagerung der Schwerpunkte der Umweltinformatik in Richtung der informationstechnischen Unterstützung des betrieblichen Umweltschutzes ergeben. Diese Ausweitung der Sichtweise der Umweltinformatik ist vor allem auch vor dem Hintergrund der Diskussion zur Nachhaltigkeit zu sehen, die neben der rein ökologischen Sicht auch ökonomische und soziale Fragen mit einbezieht.

Auf der UN-Konferenz für Umwelt und Entwicklung in Rio de Janeiro 1992 wurde das Leitbild der nachhaltigen Entwicklung international von den beteiligten 178 Staaten anerkannt. Als zentrales Ergebnis wurde das Aktionsprogramm Agenda 21 mit Handlungsempfehlungen für Umweltvorsorge und Raumentwicklung ratifiziert. Auf dieser Basis wurden für Deutschland ein nationales Aktionsprogramm zur Umsetzung der Agenda 21 und zahlreiche lokale Agenda 21 Prozesse mit konkreten Zielen und zugehörigen Indikatoren entwickelt.

Dem folgte ein Umdenken auch im Zusammenhang des Flächenmanagements. Einer der in der Agenda 21 festgelegten Indikatoren aus dem Bereich der nachhaltigen Raumentwicklung zum Schutz der natürlichen Lebensgrundlagen ist die Flächenneuanspruchnahme.

Für Deutschland hat dies zur Folge, dass bis zum Jahr 2020 der Neufächenverbrauch von derzeit ca. 105 ha/Tag auf 30 ha/Tag gesenkt werden soll. Zum Erreichen dieses Zieles wird es vermehrt notwendig, bereits vorgenutzte, brachgefallene Flächen einer Neu- oder Umnutzung zuzuführen sowie Baulücken zu schließen. Damit kommt der Konversion allgemein und mit den Teilbereichen der militärischen und der zivilen Konversion von Liegenschaften im Speziellen eine besondere Bedeutung zu.

Diese Arbeit konzentriert sich für die praktische Anwendung zunächst auf die Flächenutzung im innerstädtischen Bereich. Der hiermit verbundene Planungsablauf ist der Ansatzpunkt der in Kapitel 4 und 5 beschriebenen Verfahren für das Clustering und die Visualisierung. Zunächst werden die generellen Aufgaben und Ziele der Stadtplanung erläutert und anschließend auf das Themenfeld der Konversion eingegangen. Im Anschluss daran werden die Informationssysteme vorgestellt, die in diesem Teilgebiet der Umweltinformatik zum Einsatz kommen.

2.3.2 Stadtplanung

Städtebau und Stadtplanung regeln eine vorausschauende Ordnung der baulichen und sonstigen Nutzung von Grund und Boden in Städten und Siedlungsbereichen. Stadtplanung beschäftigt sich mit der Lenkung der räumlichen Entwicklung und der Nutzung der Flächen einer Stadt, wohingegen Aufgabe des Städtebaus ist, die

Planung umzusetzen und die bauliche Gestaltung städtischen Lebensraumes vorzunehmen.

Die Stadtplanung ist eine Teildisziplin der Raumplanung und beschäftigt sich mit der Analyse der Stadt und darauf folgend der Erarbeitung von Planungskonzepten unter Abwägung aller relevanten Interessen mit dem Ziel der Konfliktminimierung. Stadtplanung dient der Steuerung der Bodennutzung für gesamte Stadt- oder Gemeindegebiete oder für Teilbereiche davon, sog. Quartiere.

Aufgabe der Stadtplanung ist die Erzielung einer nachhaltigen städtebaulichen Entwicklung der Städte. Dabei sind die sozialen, wirtschaftlichen und umweltschützenden Anforderungen miteinander in Einklang zu bringen. Eine dem Wohl der Allgemeinheit dienende sozialgerechte Bodennutzung ist zu gewährleisten. Stadtplanung soll dazu beitragen, eine menschenwürdige Umwelt zu sichern und die natürlichen Lebensgrundlagen zu schützen und zu entwickeln, auch in Verantwortung für den allgemeinen Klimaschutz. Darüber hinaus sollen die städtebauliche Gestalt und das Orts- und Landschaftsbild baukulturell erhalten und entwickelt werden.

Gesetzliche Grundlage für stadtplanerisches Handeln bildet das Baugesetzbuch (BauGB). Im BauGB werden förmliche Verfahren zur Aufstellung verschiedener Pläne geregelt. Den höchsten Stellenwert nimmt die Bauleitplanung ein, die zwei Planwerke von unterschiedlicher Detailschärfe und Verbindlichkeit unterscheidet:

Als vorbereitender Bauleitplan dient der Flächennutzungsplan (FNP), der das gesamte Stadt- oder Gemeindegebiet umfasst und als Grundlage für die Ausarbeitung von detaillierten Plänen für Teile des Gemeindegebietes dient. Im FNP werden Aussagen über die zukünftig beabsichtigte Verteilung von Bodennutzungen getroffen, also die Verteilung und Zuordnung von Wohn-, Gewerbe-, Frei- und Sonderflächen sowie die Lage wichtiger Verkehrsstrassen.

Für Teilbereiche eines Gemeindegebietes werden als verbindliche Bauleitpläne Bebauungspläne aufgestellt, die neben den Aussagen zur Verteilung der Bodennutzungen auch gestalterische Festsetzungen und bestimmte Grundstücksrechte enthalten. Bebauungspläne sind parzellenscharf und werden nach Abschluss umfangreicher Verfahren zur Beteiligung der Öffentlichkeit als Satzungen beschlossen und sind unmittelbar rechtswirksam.

Abbildung [Bra99] beschreibt das System der Raumplanung und die Einordnung der Stadtplanung in dieses System. Nach Braam [Bra99] lassen sich drei grundlegende Prinzipien der Raumplanung ausmachen, die sich an der Forderung des Grundgesetzes nach Wahrung der Einheitlichkeit der Lebensverhältnisse über das Gebiet eines Landes hinaus (Artikel 72 Abs. 2 Satz 3 GG) orientieren:

Prinzip der Nachhaltigkeit

„Erarbeitung von längerfristigen Konzeptionen und Leitbildern, planerischen Vorgaben und ordnerischen Maßnahmen, um Zersiedelung und Überlastung zu vermeiden

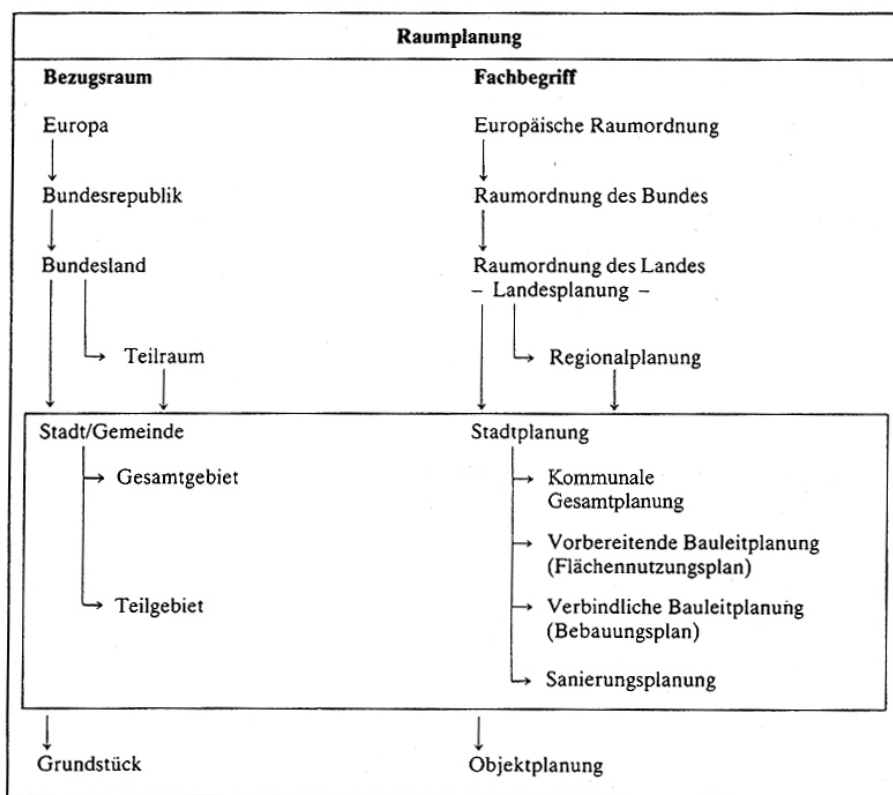


Abbildung 2.16: System der Raumplanung

und Ressourcen zu sichern.“

Prinzip der Regionalisierung

„Entwicklung der Regionen in ökonomischer, ökologischer und sozialer Hinsicht. Damit soll größeren Abwanderungen aus den wirtschaftsschwachen peripheren Räumen entgegengewirkt werden.“

Prinzip der Gleichwertigkeit

„Ausgleich zwischen den wirtschaftsstarken und den wirtschaftsschwächeren Regionen, um eine möglichst gleichmäßige räumliche Entwicklung zu gewährleisten, wobei dies nicht mit Gleichartigkeit verwechselt werden darf; vielmehr gilt es, die verschiedenen Regionen Deutschlands unter Beachtung ihrer jeweiligen Potentiale und Eigenarten zu entwickeln.“

Der Begriff der Stadtplanung setzt sich aus den beiden Elementen der Stadt und der Planung zusammen, die die Inhalte der Stadtplanung prägen. Die Stadt ist ein komplexes Gebilde, das „den Bedürfnissen und Wünschen materieller und immaterieller Art einer vielschichtigen Bürgerschaft Erfüllung bietet; das den sozialen und

wirtschaftlichen, kulturellen und politischen Verhältnissen der jeweiligen Generation Entsprechung gibt; das aufgrund der lokalen Bedingungen von Standort und Klima seine Eigenart besitzt; das mit den technischen Mitteln der Zeit geschaffen wird...“[Hil75]. Als Planung wird die „systematische Vorbereitung eines vernunftgemäßen Handelns von Einzelnen oder Gemeinschaften, um ein Ziel unter den gegebenen Verhältnissen auf die beste Weise zu erreichen“ [fSuV98] bezeichnet. Beide Begriffe zusammen stellen die Rahmenbedingungen und Konfliktbereiche, mit denen sich die Stadtplanung auseinandersetzt. Die Aufgabe der Stadtplanung ist eine fundierte Vorbereitung einer möglichen Entscheidung. Hierbei müssen alle relevanten Kriterien, die für die Beurteilung eines Sachverhaltes benötigt werden, zusammengebracht werden und die möglichen Alternativen einer Maßnahme aufgezeigt werden.

Das Ziel der Stadtplanung ist es, „sachlich richtige, inhaltlich ausgewogene und organisatorisch aufeinander abgestimmte Maßnahmen vorzubereiten und damit die Stadt den jeweiligen Erfordernissen, Vorstellungen und Möglichkeiten entsprechend anzupassen oder weiterzuentwickeln“[Bra99]. Das Instrument der Stadtplanung ist ein differenzierter Planungsablauf.

Aktuelle Tendenzen in der Stadtplanung beschäftigen sich mit der Zunahme des Verkehrsaufkommens, der räumlich-funktionalen Entmischung von Nutzungen, der räumlichen Ausdehnung der Siedlungsflächen im Umland der Großstädte, der Umstrukturierung der Landwirtschaft sowie mit der Konversion von Flächen [Bra99]. Im folgenden werden die Konversionsflächen näher betrachtet, da in den Kapitel 3 und 6.3.1 das Themenfeld der militärischen Konversion als Anwendungsbeispiel für das entwickelte Information Clustering Verfahren verwendet wird.

Konversion

Der Begriff Konversion kommt aus dem Lateinischen und bedeutet Umwandlung oder Abkehr. Unter Konversion wird die Umnutzung alter Gebäude oder Flächen vorwiegend militärischer, industrieller oder infrastruktureller Natur verstanden. Oft wird die Konversion im Zusammenhang mit der Umnutzung alter Militärfächen in zivil genutzte Flächen verwendet. Ähnlich geschieht dies bei ehemaligen Industrie- und Eisenbahn-Anlagen. Unterschieden wird deshalb die militärische von der zivilen Konversion.

Durch zahlreiche Umstrukturierungen der in Europa stationierten Streitkräfte zur Anpassung an die veränderten Sicherheitsbedürfnisse europa- und weltweit hat die Wiedernutzung freigewordener militärischer Flächen seit Jahren in Europa eine hohe Bedeutung für die Stadt-, Raum- und Umweltentwicklung. Das wird sich in der Zukunft verstärkt fortsetzen.

Die militärische Konversion umfasst insbesondere

- die Umnutzung von ehemals militärisch genutzten Flächen für zivile Zwecke,



Abbildung 2.17: Konversion in Kaiserslautern: PRE Park [Gmb02]

- die Umstellung von Rüstungsbetrieben auf zivile Produktion,
- die Beseitigung oder Abschwächung der negativen wirtschaftlichen Folgen durch kleinere oder größere Projekte in den Konversionsräumen auf der Grundlage einer Gesamt- oder von regionalen Strategien,
- auf die Anfangsphase befristete, den Arbeitsmarkt stabilisierende Maßnahmen im Wesentlichen für die Bauwirtschaft in den Konversionsräumen.

Ein grundsätzlicher Überblick über das Themenfeld Konversion militärischer Liegenschaften wird von Steinebach et al. [SJ97] und zu Herangehensweisen an die Umnutzung einer solchen Fläche in der durch das Bundesland Rheinland-Pfalz [MdIufS02] veröffentlichten Arbeitshilfe gegeben.

Die zivile Konversion umfasst insbesondere

- die Umwandlung aller stadt- bzw. regionalentwicklungspolitisch bedeutsamen brachgefallenen nichtmilitärischen Flächen (z. B. von Bahn, Post, Gewerbe und Industrie); Auslöser hierfür ist der normale strukturelle Wandel,
- die raum- bzw. gebietsbezogene Bewältigung des generellen Strukturwandels unter wirtschaftlichen, städtebaulichen und sozialen Gesichtspunkten (z.B. in Sanierungs- und Entwicklungsgebieten, neuestens auch unter dem Gesichtspunkt der Sozialen Stadt).

Die Bedingungen für die Bewältigung der militärischen und zivilen Konversion sind vergleichbar und erfordern ähnliche Strategien und den Einsatz entsprechender Instrumente. Durch raumplanerische Konzepte sollen die Flächen in die städtebauliche Entwicklung wieder einbezogen werden. Zurzeit erfolgt eine Neuplanung der Konversionsflächen noch fast vollständig ohne Computerunterstützung und geeignete Programme. Grundsätzlich basieren die Entscheidungen auf empirischen Daten,

sonstigen Bestandsaufnahmen, Analysen, Zielvorgaben und der Bewertung von Planungsalternativen.

Im Folgenden wird ein Einblick in die gegenwärtig eingesetzten raumbezogenen Informationssysteme gegeben.

2.3.3 Informationssysteme

Informationssysteme halten immer mehr Einzug in die verschiedensten Bereiche des alltäglichen Lebens. Ihr Anwendungsgebiet ist vielfältig: von der Versicherungsagentur, die mögliche Schäden und Kosten kalkulieren will, bis hin zur Polizei, die ihre Verbrecherjagd besser koordinieren möchte. Obwohl in der Literatur verschiedene Definitionen des Begriffes existieren, ist eine klare Abgrenzung von Informationssystemen sehr schwierig. Auch ein Visualisierungssystem, das „nebenbei“ eine Landkarte darstellen kann, darf zum Beispiel als Geoinformationssystem (GIS) bezeichnet werden.

De Lange [dL02] definiert den Begriff Informationssystem wie folgt:

Definition 5 *Ein System, das auf einen Datenbestand zurückgreift und Auswertungen dieser Daten zulässt, so dass Informationen abgeleitet und wiedergegeben werden können, kann allgemein als Informationssystem bezeichnet werden.*

Reine Auskunftssysteme, die nur bereits vorhandene Daten verarbeiten, werden nicht als Informationssysteme bezeichnet, da bei solchen Systemen die Datenaufnahme (z.B. Neuaufnahme oder Aktualisierung) fehlt. Informationssysteme allgemein lassen sich nach Art der gespeicherten Informationen klassifizieren. So gibt es beispielsweise (alpha-)numerische, textliche, bildhafte und multimediale Informationen. Ausgehend von den verschiedenen Informationsarten werden Informationssysteme mit speziell auf den jeweiligen Informationstyp abgestimmten Verarbeitungsmethoden benötigt. Beispiele für solche Informationssysteme finden sich in Banken (Kundenstamm, Kontoführung,...), Reisebüros (Hotelbelegungen, Buchungen,...) und Bibliotheken (Benutzerdaten, Buchreservierungen,...).

Der Begriff Geoinformationssystem erweitert die Definition von Informationssystemen um raumbezogene Daten. Geoinformationssysteme modellieren Geobjekte der realen Welt und bilden diese Objekte in ein digitales Informationssystem ab. Dies ist ein wesentlicher Unterschied zu den übrigen Informationssystemen. Eine kompakte Definition des Begriffs Geoinformationssystem liefert Bill [Bil99]:

Definition 6 *Ein GIS ist ein rechnergestütztes System, das aus Hardware, Software, Daten und den Anwendungen besteht. Mit ihm können raumbezogene Daten digital erfasst und redigiert, gespeichert und reorganisiert, modelliert und analysiert sowie alphanumerisch und geographisch repräsentiert werden.*

Lindner [Lin99] definiert den Begriff des Geoinformationssystems etwas anders:

Definition 7 *Unter einem Geo-Informationssystem versteht man die Verarbeitung und Verwaltung raumbezogener Daten (GeoDaten) mithilfe der elektronischen Datenverarbeitung. Als raumbezogen können dabei solche Informationen bezeichnet werden, die sich eindeutig mithilfe von Koordinaten in einem Untersuchungsgebiet verorten lassen.*

Es liegen bereits einige Ansätze vor, existierende GIS-Systeme mit zusätzlichen Informationen zu kombinieren. Bei Kanzler [KdL96] wird eine Übersicht über den Einsatz von GIS-Systemen in der Regionalplanung gegeben.

Alle Definitionen des Begriffs Geoinformationssystem sagen nichts über die Art der raumbezogenen Daten aus. Dies ist durchaus beabsichtigt, da erst durch die Festlegung der Aufgaben und Einsatzgebiete eines GIS die raumbezogenen Daten genauer bestimmt werden. Dies führt zu verschiedenen Ausprägungen von GIS. So haben sich im Laufe der Zeit die verschiedenen Fachdisziplinen ihre eigenen Geoinforma-

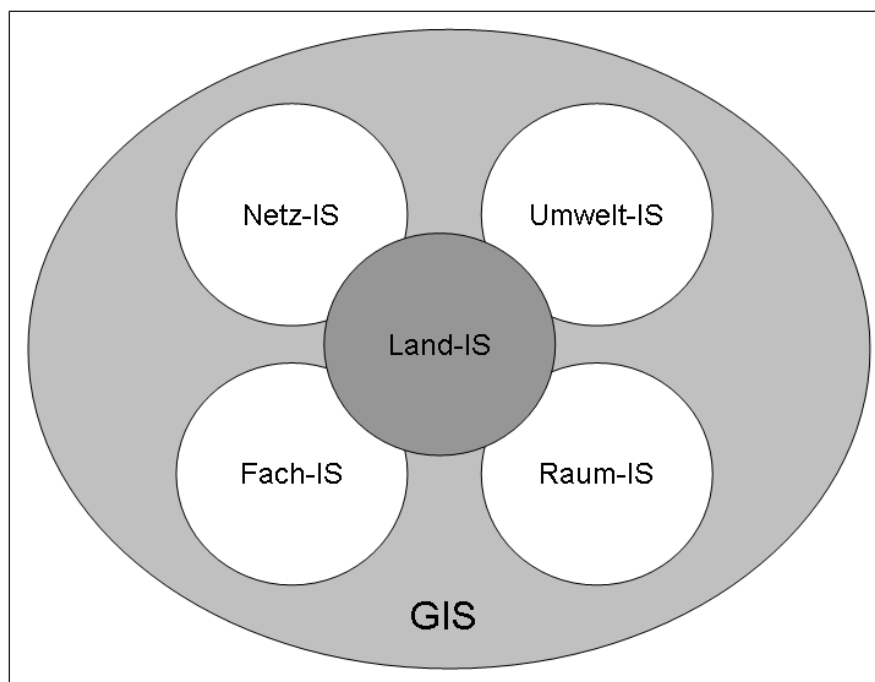


Abbildung 2.18: Ausprägungen von GIS

tionssysteme aufgebaut. In Abbildung 2.18 sind die fünf großen Gruppen von GIS-Ausprägungen dargestellt.

Landinformationssysteme

Das Vermessungswesen kann als Stein des Anstoßes für die Entwicklung von Landinformationssystemen angesehen werden. Bei diesen Systemen steht die geometrische Erfassung und Laufendhaltung des Grund und Bodens und die damit verbundenen Sachdaten im Vordergrund. Ein Beispiel hierfür stellt das Liegenschaftskataster dar, in dem alle Flurstücke nach ihrer Lage, Nutzung, Größe usw. verzeichnet und dargestellt werden. Eine formale Definition von Landinformationssystemen wurde 1982 auf dem „Federation Internationale des Geometres“ Kongress in Montreux gefasst [Bil99]:

Definition 8 *Ein Landinformationssystem (LIS) ist ein Instrument zur Entscheidungsfindung in Recht, Verwaltung und Wirtschaft sowie ein Hilfsmittel für Planung und Entwicklung. Es besteht einerseits aus einer Datensammlung, welche auf Grund und Boden bezogene Daten einer bestimmten Region enthält, andererseits aus Verfahren und Methoden für die systematische Erfassung, Aktualisierung, Verarbeitung und Umsetzung dieser Daten. Die Grundlage eines LIS bildet ein einheitliches, räumliches Bezugssystem für die gespeicherten Daten, welches auch eine Verknüpfung der im System gespeicherten Daten mit anderen bodenbezogenen Daten erleichtert.*

LIS befassen sich mit geometrischen Fragestellungen in einem großmaßstäblichen Bereich (1:500 - 1:10000). Wichtige Aufgabengebiete von Landinformationssystemen sind das Vermessungswesen und die Liegenschaftskataster. Somit lassen sich folgende Aufgaben eines LIS spezifizieren:

- Vermessungswesen
 - Liegenschaftskataster (Liegenschaftskarte, Liegenschaftsbuch)
 - Landesvermessung (topographische Landesaufnahme)
 - Kommunale Vermessung
 - Ingenieursvermessung (lokale Vermessungsgrundlagen)
- Grundbuch
 - Bestandsverzeichnis
 - Eigentumsnachweis
 - Lasten und Beschränkungen
 - Finanzielle Belastungen

Rauminformationssysteme

Geographen, Raumplaner und Statistiker können als Urheber für die Entwicklung

von Rauminformationssystemen (RIS) genannt werden. Aufgrund der Vielfalt dieser Disziplinen ergibt sich ein breit gefächertes Anwendungsbereich. RIS können sowohl bei der Erfassung der Bevölkerungsentwicklung, der Wirtschaft und den Siedlungen als auch bei der amtlichen Statistik oder zur Aufstellung von Entwicklungsprogrammen eingesetzt werden. Mit folgender Definition werden LIS und GIS gegeneinander abgegrenzt [Bil99]:

Definition 9 *Ein Rauminformationssystem ist ein Instrument zur Entscheidungsfindung sowie ein Hilfsmittel für Planung und Entwicklung. Es besteht aus einer Datensammlung zur Bevölkerungs-, Wirtschafts- und Siedlungsentwicklung, zum Infrastrukturausbau, zur Flächennutzung und den Ressourcen, die in regionale Entwicklungsprogramme und raumbedeutsame Vorhaben einfließen. Ebenso sind die Verfahren und Methoden zur Erfassung, Aktualisierung und Umsetzung dieser Daten wesentlicher Bestandteil des Informationssystems. Die Grundlage bildet der einheitliche Raumbezug, der die verschiedenartigen Daten miteinander verknüpft.*

Die Erfassung, Verwaltung, Analyse und Präsentation als Komponenten sowie die Permanentdatenhaltung und die hohe Interaktivität sind einige Merkmale für Rauminformationssysteme. Bei diesen Systemen steht weiterhin die Analyse im Vordergrund. Es ist anzumerken, dass es sich bei Rauminformationssystemen um hybride GIS handelt, da diese Systeme sowohl Vektor- als auch Rasterdaten verarbeiten können. Die Stärken von RIS liegen in der thematisch-kartographischen Visualisierung.

Als konkrete Hauptanwendungsgebiete eines RIS sind die Raumordnung, die Landes-/Regionalplanung, die Kommunalplanung und die amtliche Statistik zu nennen. Dies führt zu folgender Untergliederung:

- Raumordnung und Landesplanung (Raumordnungsprogramm,...)
- Kommunalplanung (Flächennutzungsplan, Bebauungsplan)
- Fachplanungen wie z.B. Landschaftsplanung/Verkehrsplanung
- Amtliche Statistik (Infrastrukturkataster,...)

Umweltinformationssysteme

Als weitere große Gruppe der Spezialisierungen von GIS sind die Umweltinformationssysteme (UIS) zu nennen. Diese Systeme lassen sich durch folgende formale Definition fassen [Bil99]:

Definition 10 *Ein Umweltinformationssystem ist ein erweitertes GIS, das der Erfassung, Speicherung, Verarbeitung und Präsentation von raum-, zeit- und inhaltsbezogenen Daten zur Beschreibung des Zustandes der Umwelt hinsichtlich Belastungen und Gefährdungen dient und Grundlagen für Maßnahmen des Umweltschutzes bildet.*

Es gibt auch hier eine Vielzahl von Anwendungsgebieten. So erfolgt der Einsatz solcher Systeme bei der Erfassung von Radioaktivität, bei Biotopkartierungen und bei der Erhaltung der Artenvielfalt. Weitere Aufgaben für UIS lassen sich im Forstwesen, in der Abfallentsorgung oder im Gewässerschutz finden. Mithilfe von UIS erfolgen beispielsweise auch Umweltverträglichkeitsprüfungen, deren Sinn und Zweck darin besteht zu prüfen, welche Einflüsse ein geplantes Bauvorhaben auf die Umwelt hat. Die verschiedenen Aufgabengebiete lassen sich folgendermaßen zusammenfassen:

- Erfassung und Aktualisierung von Daten zur Qualität von Luft, Boden und Wasser
- Erkennung von Pflanzenschäden und Gesundheitsrisiken
- Überprüfung der Einwirkung von Radioaktivität und chemischen Stoffen auf die Umwelt
- Erhaltung der Artenvielfalt durch Gründung und Pflege von Biotopen und Schutzgebieten

Bei Umweltinformationssystemen handelt es sich, wie auch bei RIS, um hybride GIS. Diese Systeme erlauben außerdem die Modellierung, Simulation und Animation von Umweltprozessen und verknüpfen unterschiedliche Thematiken über den gemeinsamen Raumbezug. Dabei wird oft auf sich zeitlich schnell ändernden Daten gearbeitet. Der Aufbau von Umweltinformationssystemen erfolgt sowohl auf internationaler Ebene als auch auf Bundes-, Landes-, und kommunaler Ebene sowie in privaten Unternehmen. Speziell im Bereich der privaten Unternehmen erfolgte in den letzten Jahren ein rasanter Zuwachs.

Netzinformationssysteme

Netzinformationssysteme (NIS) dokumentieren und bearbeiten Betriebsmitteldaten. Hierbei erstreckt sich der Bereich der Betriebsmitteldaten von Kundendaten über die Leitungen bis hin zu Anlagen zur Ver- und Entsorgung. Folgende Definition 11 fasst den Begriff des Netzinformationssystems formal [Bil99]:

Definition 11 *Ein Netzinformationssystem ist ein Instrument zur Erfassung, Verwaltung, Analyse und Präsentation von Betriebsmitteldaten. Diese beziehen sich auf die Netzwerktopologie, die in einem einheitlichen Bezugsrahmen gegeben sein muss.*

NIS beschränken sich normalerweise auf eine lokale bis regionale Gebietsausdehnung, wobei Vektordaten dominieren und Rasterdaten nur als Hintergrundgrafik verwendet werden. Die Erfassung, Verwaltung, Analyse und Präsentation werden dabei als Komponenten aufgefasst. In der Praxis werden von vielen GIS-Herstellern so genannte „Fachschaalen“ für Gas, Strom, Wasser, Kanal, usw. dem entsprechenden Kunden angeboten. Dies hat den Vorteil, dass das System optimal auf die jeweiligen

Ver- und Entsorgungssparten abgestimmt ist. Um nun beispielsweise einen neuen Kanal zu planen, muss das NIS auf die geometrische Darstellung des Liegenschaftskatasters zurückgreifen. Dabei ist es wichtig, dass diese Daten aktuell sind. Somit ist ein NIS eng mit dem entsprechenden LIS, das für die Laufendhaltung der Daten sorgt, verbunden.

Fachinformationssysteme

Aus Sichtweise der jeweiligen Fachdisziplin könnten die bis jetzt vorgestellten Ausprägungen von Geoinformationssystemen alle als Fachinformationssysteme (FIS) bezeichnet werden. In der Literatur bleibt aber der Begriff des FIS der Gruppe von Geoinformationssystemen vorbehalten, die bis jetzt noch nicht erwähnt wurden. In dieser Kategorie finden sich unter anderem Spezialanwendungen, die sich zu keiner der zuvor genannten Gruppen zuordnen lassen. Die Umstellung der analogen Luftfahrkarten für die Navigation von Flugzeugen ist ein Beispiel für solch eine spezielle Anwendung. Als weiteres Beispiel ist die Einrichtung von digitalen Straßenkarten als Grundlage für die autonome Fahrzeugnavigation zu nennen. Es gibt viele weitere Anwendungen im Bereich der Telekommunikation und im Verkehrs-/Logistiksektor. Aufgrund der Tatsache, dass es sich bei Fachinformationssystemen um sehr spezielle Systeme handelt, müssen die GIS-Funktionalitäten solcher Systeme natürlich entsprechend der geplanten Anwendung erweitert und ergänzt werden. Dadurch wird es aber gleichzeitig schwer, den Begriff FIS formal zu fassen, da es nicht möglich ist, für die vielen verschiedenen Systeme einheitliche Charakteristika anzugeben.

Im Bereich der Raum- und Umweltplanung werden zurzeit verstärkt den Planungsprozess unterstützende Systeme entwickelt (planning support systems, PSS), deren Kernstück wiederum das Geoinformationssystem ist. Planungssysteme stellen die Informations-Infrastruktur für die Planung mit dem Ziel bereit, die Interaktion zwischen den Akteuren zu unterstützen. Sie kombinieren analytische Tools und computergestützte Simulationsmodelle mit visuellen Darstellungen.

Ein Planungssystem muss strukturierte und zugreifbare Informationen zur Verfügung stellen, zusammen mit einer weit gestreuten Zahl von Analyse-, Prognose- und Entscheidungsfindungstools. Das Ziel ist es einen kontinuierlichen und interaktiven Analyse-, Design- und Evaluationsprozess unterstützen zu können. Hierfür sind zusätzlich zu dem GIS-System die traditionellen planerischen Tools eingebunden.

Kapitel 3

Information Clustering im Planungsablauf

Die räumliche Entscheidungsfindung ist ein hochgradig komplexer, interdependenter Prozess, zu deren Unterstützung Modelle eingesetzt werden. Maßgebliches Modell der Stadtplanung ist der Planungsablauf, der die Gesetzmäßigkeiten bei Planungsvorhaben festlegt. Dieses Kapitel konzentriert sich auf die strukturellen Komponenten des Planungsablaufes. Es werden neue Anforderungen an den Planungsablauf beschrieben, die eine Ausweitung des bestehenden Modells durch computergestützte Tools nötig machen. Die Anknüpfungsmöglichkeiten einer solchen Modellerweiterung werden analysiert und der Aufbau und die Inhalte der benötigten Tools erarbeitet.

Im folgenden Kapitel 3.1 wird der traditionelle Planungsablauf analysiert und die Schwierigkeiten bei dessen Handhabung untersucht. Kapitel 3.1.1 beschreibt hierfür zunächst die Anforderungen an ein Planungsvorhaben und den praktischen Einsatz des Planungsablaufes. Kapitel 3.1.2 setzt auf diesen Untersuchungen auf und arbeitet die Ansatzpunkte für die in dieser Arbeit entwickelten Verfahren heraus, das Information Clustering.

Im Anschluss an die Analyse des Planungsablaufes werden die Einflußfaktoren betrachtet, die den Planungsablauf prägen. In 3.2 werden diese Profile zunächst allgemein beschrieben und anschließend am Beispiel von Konversionsflächen konkretisiert. Hierfür werden die Akteure in Kapitel 3.2.1 und die Parameter in Kapitel 3.2.2 ermittelt.

Zum Abschluss dieses Kapitels wird eine generelle Zielfunktion entwickelt, die die gefundenen Profile integriert und die konkreten Vorgaben an die Modellerweiterung, das Information Clustering, mathematisch beschreibt. Schließlich wird der Bedarf für eine Visualisierung in Kapitel 3.3.2 dargelegt, die Anforderungen an diese ermittelt und Lösungsansätze für die Visualisierung erarbeitet. Die theoretische Erarbeitung der Verfahren für das Information Clustering erfolgt in den Kapiteln 4 und 5.

3.1 Planerische Entscheidungsfindung

Der Planungsablauf ist das maßgebliche Instrument der Stadtplanung. Er beinhaltet verschiedene Komponenten, zwischen denen es diverse Rückkopplungen gibt. Der Planungsablauf ist somit ein dynamischer Prozess, der u.a. durch die große Anzahl der mitwirkenden Personen sowie dem großen Informationsvolumen zu einem hochgradig komplexen Prozess wird.

In Kapitel 3.1.1 werden die Gesetzmäßigkeiten des Planungsablaufes beschrieben. Hierfür wird der Planungsablauf detailliert dargestellt und seine einzelnen Schritte erläutert. Kapitel 3.1.2 analysiert die Gründe für die auftretende Komplexität und die neu entstehenden Anforderungen an den Planungsablauf. Es wird der Prozess des Information Clustering entwickelt, mit dem die bestehenden Mängel reduziert werden und die Komplexität abgebaut wird.

3.1.1 Systematischer Ansatz

Die Stadtplanung beinhaltet das Zusammenspiel von vielen Einzelkomponenten, die sich in dem räumlichen Bereich der Stadt abspielen. Für die Planung neuer Vorhaben wird deswegen ein Instrument benötigt, mit dem diese einzelnen Komponenten koordiniert werden können. Dieses Instrument ist der Planungsablauf. Jede Planung unterliegt gewissen Gesetzmäßigkeiten, die in einem generellen Planungsablauf festgeschrieben sind und den Erfolg einer Planung sicher stellen sollen. Der systematische Planungsprozess wird durch gewisse regelhafte, sich grundsätzlich wiederholende Phasen bestimmt:

1. Aufgabenstellung
2. Bestandsbeschreibung
3. Situationsbewertung
4. Lösungsfindung
5. Umsetzung

Die Aufgabenstellung umfasst die Entstehung der Planungsaufgabe, die Aufgabengrenzung und die Problembeschreibung. Bei der Bestandsbeschreibung wird der aktuelle Status aufgenommen, in der Situationsbewertung werden Kriterien, Entwicklungsmöglichkeiten und mögliche Konflikte herausgearbeitet. Die Lösungsfindung stellt methodische Ansätze und andere Lösungsstrategien vor, die in der Umsetzung zur Anwendung kommen.

In der klassischen Form nach Braam [Bra99] (siehe Abbildung 3.1) gliedert sich der

Planungsablauf im Wesentlichen in die Bereiche Entwicklung von Zielvorstellungen, Entwicklung von Planungsalternativen, Entscheidung für eine Alternative und abschließend die Umsetzung der gewählten Alternative. Zwischen diesen einzelnen

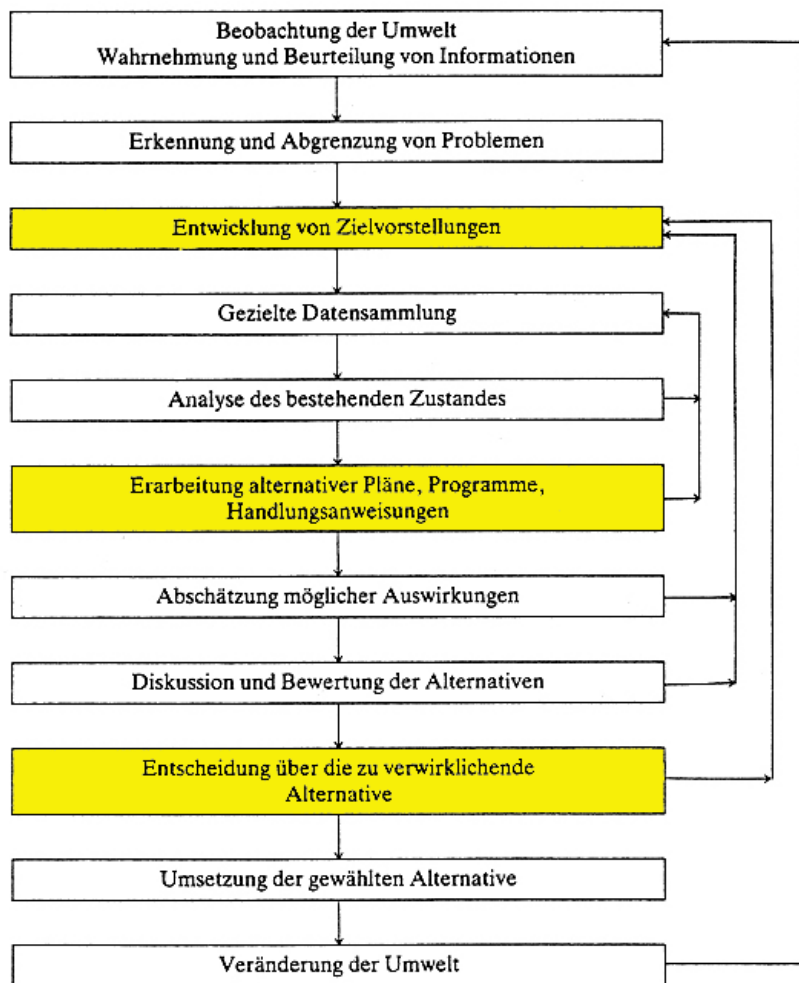


Abbildung 3.1: Ablaufschema einer Planung

Komponenten gibt es diverse Rückkopplungen, so dass sich ein einer dynamischer Prozess entwickelt. Ein besonderer Diskussionsbedarf zwischen den Mitwirkenden Personen besteht bei der Entwicklung der Zielvorstellungen bis hin zur der Entwicklung der Planungsalternativen. Die Entwicklung von Planungsalternativen richtet sich auf die strukturellen Nutzungskonzepte der Fläche.

In der praktischen Durchführung wird der systematische Planungsablauf dem jeweiligen Anwendungsfall angepasst und hierfür in den einzelnen Punkten bedarfsgerecht

modifiziert. Im folgenden wird ein Beispiel aus dem Bereich der Verkehrsentwicklung vorgestellt.

Auch die Verkehrsentwicklung als ein Teil der Stadtentwicklung unterliegt den Gesetzmäßigkeiten des Planungsablaufes. Der Verkehr ist vereinfacht die Ortsveränderung von Personen und Gütern, z.B. von der Wohnung zur Arbeit.

Die Arbeitsschritte in der Verkehrsplanung charakterisieren den Ablauf und die Arbeitsphasen bei der Erstellung von Verkehrsentwicklungsplänen für Städte. Der Verkehrsplanungsprozess hat dabei den in Abbildung 3.2 dargestellten prinzipiellen Ablauf: Wie sämtliche Planungsprozesse ist auch der Verkehrsplanungsprozess ein

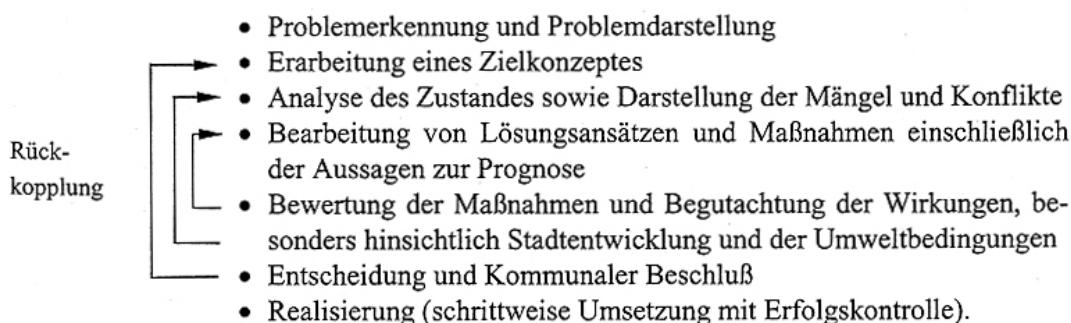


Abbildung 3.2: Verkehrsplanungsprozess [Kor97]

interaktiver Prozess, der in die verschiedenen Ebene zurückspringt, um mit einem neuen oder veränderten Ansatz die gestellten Ziele zu erreichen.

Ein konkretes Beispiel eines Planungablaufes in der Stadtplanung kommt aus dem Bereich der Verkehrsberuhigung.

Die enorme Zunahme des motorisierten Verkehrs hat in den 50er und 60er Jahren zu einer wesentlichen Minderung der Aufenthalts- und Wohnqualität von Erschließungs- und Hauptnetzstraßen geführt. Die negativen Folgen der vermeintlichen Freiheit des Autoverkehrs führten in den hochmotorisierten Ländern zu einem Wertewandel in der Stadtverkehrsplanung [GS94], in dem die bisher vernachlässigten Verkehrsarten Fußgänger, Radfahrer und öffentlicher Verkehr gleichberechtigt in die verkehrsplanerischen Zielsetzungen eingeführt wurden [Kor97].

Beispiele für Ziele der Verkehrsberuhigung sind:

- verkehrlich: Erhöhung der Verkehrssicherheit und Dämpfung des Geschwindigkeitsniveaus
- umweltlich: Minderung der Lärm- und Schadstoffemissionen und Erhöhung des Anteils an Großgrün- und Freiraumflächen

- städtebaulich: Verbesserung der Aufenthaltsqualität im Straßenraum

Maßnahmen zur Umsetzung dieser Zielvorstellungen können nach verschiedenen Ge-

Arbeitsschritte	Maßnahmen/Aspekte
1. Erarbeitung einer Aufgabenstellung	Problem- und Zielstellung, Gebietsabgrenzung, Definition von Teilbereichen
2. Fixierung von übergeordneten Rahmenbedingungen	Straßennetzkonfiguration, Erschließung ÖPNV, gesamtstädtische Planungen (Flächennutzungsplan, Teilverkehrskonzepte)
3. Analyse und Bestandsaufnahme	Verkehrsbeobachtungen, Erhebungen, Befragungen, städtebauliche Funktions- und Gestaltanalyse (einschließlich Freiraum)
4. Bewertung der Situation	Wertung und Besichtigung der Situation, Ableitung von Lösungsansätzen
5. Planung und Entwurf	Straßennetzplanung, Führung des öffentlichen Verkehrs, Führung des Rad- und Fußgängerverkehrs, Detailentwurf von Knotenpunkten, Querschnitten, Straßenräumen
6. Diskussion	Diskussion und Wertung der Lösungsvorschläge, Beteiligung der Bürger, Fachplaner, Kommunalpolitiker
7. Realisierung	möglichst in nutzbaren Teilgebieten, ggf. begleitende Beobachtungen
8. Vorher-Nachher-Untersuchung	abschließende Bewertung, Schlussfolgerungen
9. Übertragbarkeit	Änderung, Rücknahme oder Erweiterung der Maßnahmen, Übertragung der Ergebnisse auf andere Bereiche der Stadt

Abbildung 3.3: Planungsschritte für die Verkehrsberuhigung [Kor97]

sichtspunkten geordnet und in der Planung abgearbeitet werden. Abbildung 3.3 zeigt die Planungsschritte für die Verkehrsberuhigung. Auch hier orientieren sich die einzelnen Arbeitsschritte an das Ablaufschema der Planung nach Braam (siehe Abbildung 3.1).

3.1.2 Modellbildung

Die räumliche Entscheidungsfindung ist ein hochgradig komplexer Prozess, der nach Leung [Leu97] durch folgende Punkte charakterisiert wird:

- hoher Komplexitätsgrad
- sehr unterschiedlich strukturierte Problemstellungen
- großes Informationsvolumen
- das Einfließen von Wissen und Werturteilen
- die Existenz von Unschärfen (Unvollständigkeit, Dynamik, ...)

Daraus resultieren hohe Anforderungen an das menschliche Kognitionsvermögen, aber auch erhebliche Hindernisse für die Problem-Modellierung.

In Entscheidungssituationen, im Zusammenhang mit hochgradig interdependenten räumlichen Systemen, ist die unterstützende Einschaltung von Modellen unerlässlich. Die Nachbildung der Realität mit Modellen ist ein notwendiges Hilfsmittel zum Erkennen von Zusammenhängen und zum Treffen von Entscheidungen. Die Transparenz und die Nachvollziehbarkeit komplexer Entscheidungsprozesse wird durch eine modellhafte Darstellung erheblich erhöht. Aufgrund des Dokumentationsbedarfs in der Planung sollten aber auch kognitiv erfassbare Entscheidungsabläufe modellhaft dargestellt werden.

Jede Planung verläuft nach dem Modell des Planungsablaufes, wie er durch Braam [Bra99] beschrieben wird und in Kapitel 3.1.1 eingeführt wurde. Mit Hilfe dieses Ablaufes werden die globalen Auswirkungen kleiner lokaler Änderungen in der Planung modelliert. Die modellhafte Darstellung erhöht zwar insgesamt die Transparenz der Abläufe, erfasst jedoch nicht deren dynamische Interdependenzen. Zu diesen Problemen kommen neue Anforderungen an den Planungsablauf hinzu. Die Aufgabenstellungen in der Stadtplanung unterliegen einer steigenden Komplexität, parallel dazu weitet sich der Informations- und Analysebedarf ständig aus. Dies hat eine Beschleunigung der Planungs- und Entscheidungsabläufe zur Folge, denen der klassische Planungsablauf nicht mehr gerecht wird. Er ist zu starr, um den dynamischen Anforderungen der Realität zu genügen [SM06], die Umsetzung dauert zu lange. Zudem wird der in der Stadtentwicklung involvierte Personenkreis größer und zunehmend heterogener, was zu einer weiteren Verstärkung der Komplexität führt.

Mit dem in dieser Arbeit entwickelten Information Clustering wird die Flexibilität des Planungsablaufes erhöht. Das Information Clustering modelliert die dynamischen Abläufe, so dass die Komplexität reduziert wird. Es werden die Potentiale agierender, auch gegebenenfalls gegengerichteter Gruppen kanalisiert und die großen raumbezogenen Datenmengen strukturiert. Eine visuelle Aufbereitung der Ergebnisse unterstützt zusätzlich den kognitiven Zugang der Beteiligten zu den Ergebnissen. [SHaMR06]

Eine besondere Dynamik weist der Planungsablauf in den folgenden drei Punkten auf, zwischen denen im Wesentlichen die Rückkopplungen erfolgen:

1. Entwicklung der Zielvorstellungen
2. Erarbeitung alternativer Pläne
3. Entscheidung über eine Alternative

Abbildung 3.4 zeigt die sich daraus ergebenden Ansatzpunkte für das Information Clustering. Das Information Clustering setzt an den dynamischen Punkten des

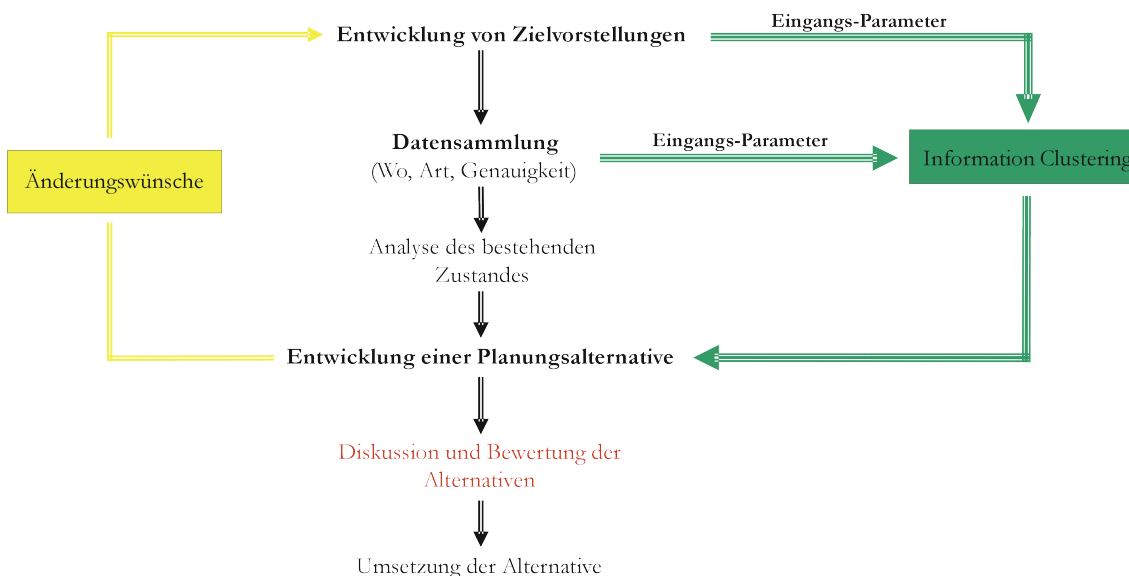


Abbildung 3.4: Ansatzpunkt des Information Clustering im Modell des Planungsablaufs

Planungsablaufes an, verarbeitet die heterogenen Eingabeparameter und liefert als Ausgabe genau eine Planungsalternative. Das Information Clustering unterscheidet sich von bestehenden Informationssystemen (siehe Kapitel 2.3.3) dadurch, dass es nicht nur Fakten sammelt, systematisiert und speichert, sondern die Fakten miteinander vernetzt, daraus Schlussfolgerungen zieht und somit zur Unterstützung bei der Entscheidungsfindung nachhaltig beiträgt.

Die Eingabeparameter für das Information Clustering ergeben sich durch die Entwicklung der Zielvorstellungen. Hierbei werden die konkreten Bedürfnisse der beteiligten Akteure ermittelt und in die Berechnung integriert. Zusätzlich werden die Standortfaktoren ermittelt, die für das betrachtete Vorhaben wichtig sind. Bei der nachfolgenden Datensammlung werden die konkreten Werte der Standortfaktoren erhoben. Der Punkt *Diskussion und Bewertung der Alternativen* entfällt in diesem erweiterten Modell des Planungsablaufes, da das Information Clustering eine Alternative bietet, die alle Beteiligten zufrieden stellt.

In den folgenden beiden Kapiteln werden die Eingabeparameter analysiert und dadurch die Anforderungen an das Information Clustering konkretisiert.

3.2 Profilerstellung

Der Beginn des Planungsablaufes ist geprägt durch die Profilerstellung. In dieser Phase werden die Zielvorstellungen der Planung ermittelt und die benötigten Da-

ten erhoben. Diese Daten umfassen im Wesentlichen die handelnden Akteure und die Standortfaktoren. Sämtliche Daten beziehen sich auf zugrunde liegende Nutzungsbereiche. Abbildung 3.5 verdeutlicht diesen Zusammenhang am Beispiel einer Konversionsfläche. Die bei der Profilerstellung erhobenen Daten fließen in Form eines

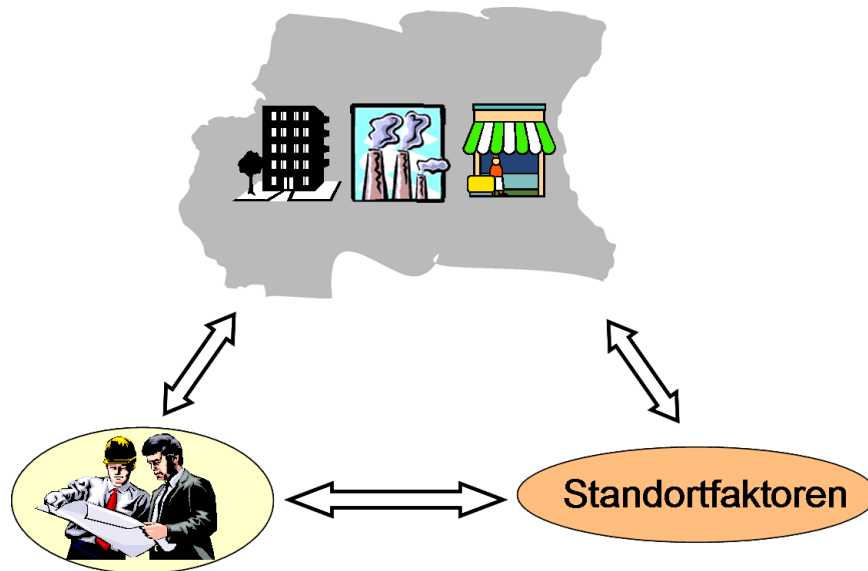


Abbildung 3.5: Profilerstellung einer Konversionsfläche

zweistufigen Prozesses in das Information Clustering ein.

3.2.1 Akteure und Sichtweisen

Der Planungsablauf ist stets eine Abstimmung zwischen Planungsfachleuten unterschiedlicher fachlicher Herkunft, den politischen Gremien mit ihren demokratisch gewählten Vertretern und den betroffenen Bürgern.

Der Verlauf des Planungsprozesses wird stark durch diese Akteure bestimmt. Die einflussnehmenden Akteure sind beispielsweise:

- die Träger der Planung, die den Planungsauftrag vergeben, die Bauherrenfunktion übernehmen und den Planungsrahmen und die Hauptziele festlegen, das Monitoring und die Datenpflege durchführen
- die Planer des Projektes, die das zu bearbeitende Projekt im Auftrage des Trägers/Bauherren fachlich erarbeiten
- die zu beteiligenden Träger Öffentlicher Belange und Öffentlichen Planungsträger und Verbände, deren Beiträge eingefordert werden

- die betroffenen Bürger, deren individuelle oder Gruppenmeinungen berücksichtigt werden sollen
- die Aufsichtsbehörden, die die Kontrolle des Aufstellungsverfahrens, der Planungsinhalte und die Genehmigung durchzuführen haben

Jeder der Akteure hat eine individuelle Sicht auf den Planungsablauf mit eigenen Interessen an der Planung. Für die Umsetzung des Information Clusterings müssen zunächst alle Akteure für das aktuelle Planungsvorhaben ermittelt werden.

Nach Isenhöfer [Ise99] lassen sich z.B. für den Bereich der Projektentwicklung die in Abbildung 3.6 dargestellten Akteure in Deutschland identifizieren.

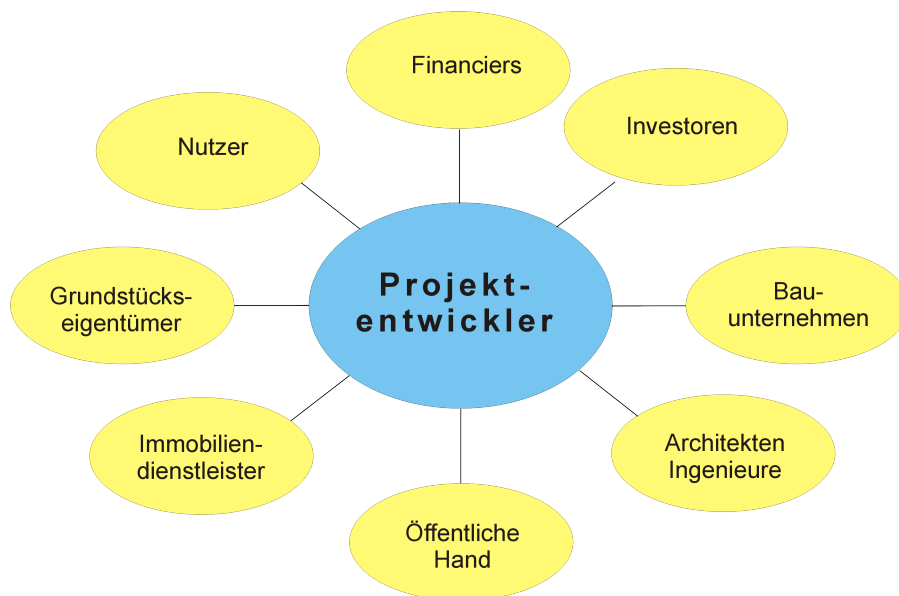


Abbildung 3.6: Akteure der Projektentwicklung

Der Nutzer kann ein Unternehmen sein, welches beispielsweise ein Gebäude zu Zwecken der Verwaltung oder Forschung nachfragt oder eine private Person, die Wohnimmobilien zur Bewohnung nachfragt. Financiers sind Institutionen, die Fremdkapital zur Verfügung stellen, wohingegen Investoren, ähnlich den Nutzern, als Anbieter oder Nachfrager von Projektentwicklungen auftreten.

Die Zielsetzungen der handelnden Akteure orientiert sich dabei an einer angemessenen Rentabilität, Sicherheit und Liquidität. Beispielsweise entscheidet sich die Geschäftsführung aufgrund der Vorgaben des Bebauungsplans, den Standortfaktoren sowie den Gründen der Wirtschaftlichkeit für die Entwicklung eines Bürogebäudes.

In Abhängigkeit von dem Anwendungsfeld der Planung unterscheiden sich die Anzahl und die Art der Akteure sowie die Sicht der Akteure auf das Planungsvorhaben. Das Information Clustering kanalisiert diese Sichtweisen der Akteure auf den Planungsablauf. Im einzelnen nehmen die Akteure eine Gewichtung der sie betreffenden Eingangsparameter vor und stellen somit heraus, welche Faktoren für sie von übergeordneter oder eher nachrangiger Bedeutung sind.

Diese Gewichtung geht in Form von Eingabeparametern in das Information Clustering ein. Hierbei muss jeder Akteur festlegen, welches Gewicht er welchem Faktor zukommen lässt.

Als praktisches Anwendungsfeld wird hier die Konversion von Flächen betrachtet. Dazu wurden die Hauptakteure des Konversionsprozesses [HRS05] extrahiert und deren Interessen untersucht. Abbildung 3.7 zeigt die ermittelten drei Hauptakteure

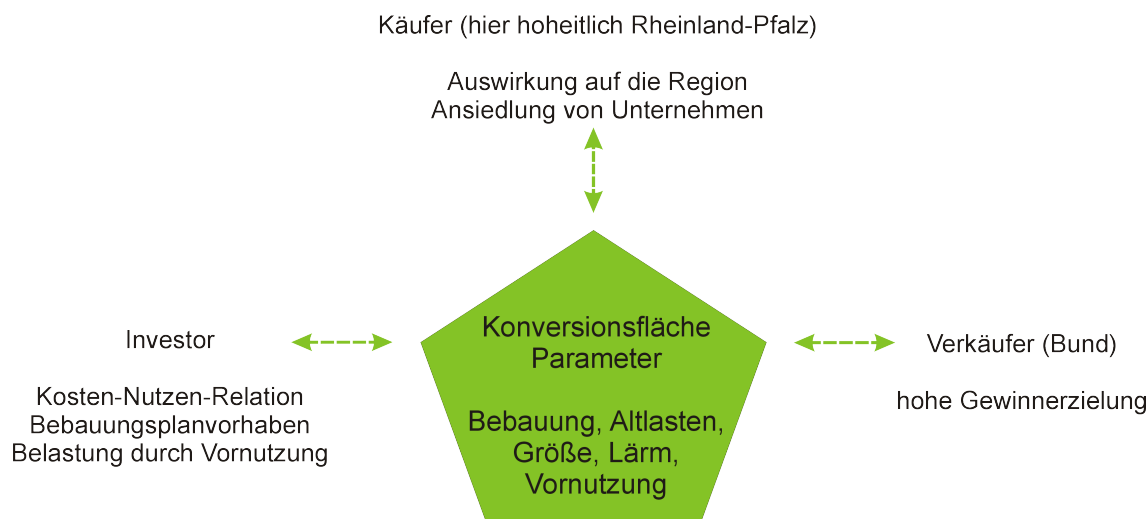


Abbildung 3.7: Hauptakteure und deren Interessen

Investor, Käufer und Verkäufer einer Konversionsfläche.

Dieser Personenkreis greift entscheidend in den Planungsablauf ein. Dem Verkäufer wird das Interesse zugeordnet, die Fläche mit möglichst hohem Gewinn zu veräußern. Der (hier hoheitliche) Käufer möchte mit dem Erwerb der Fläche eine größtmögliche positive Auswirkung auf die Region erzielen und dazu nachhaltig Unternehmen ansiedeln. Der Investor hingegen entscheidet nach Kosten-Nutzen-Relationen, überprüft alle Vorgaben des Bebauungsplanes und überprüft die Fläche auf Altlasten. In dieser Aufstellung zeigt sich, dass die drei Akteure verschiedene Interessen bei der Umnutzung einer Konversionsfläche besitzen, die prinzipiell gleichzeitig verfolgt werden. Der darin enthaltene Interessensausgleich wird im günstigsten Fall bei einer gleichrangigen Berücksichtigung aller Interessen zu einem positiven Ergebnis der Konversion für alle führen. In dem Information Clustering werden diese Interessen

mittels der beschriebenen Gewichte modelliert.

3.2.2 Parameterauswahl

Die Parameterauswahl stellt das Kernstück der Profilerstellung dar, sie bestimmt die weitere Ausrichtung des Planungsvorhabens. Es ist wichtig sämtliche Parameter zu berücksichtigen, die in einem Bezug zu der Planung stehen. Zunächst müssen diese Parameter identifiziert und anschließend erhoben werden.

Die Identifikation der Parameter ist ein aufwendiger Prozess, da es eine sehr hohe Zahl von Parametern gibt, die einem Planungsvorhaben zugeordnet werden können. In der Stadtplanung lassen sich beispielsweise folgende funktionale Aspekte ausmachen, die das Zusammenleben der Menschen prägen und somit als potentielle Standortfaktoren betrachtet werden können [Kor97]:

- Wohnen: möglichst ungestört in Verbindung mit der Landschaft
- Arbeiten: Produktionsbetriebe außerhalb der Stadt, verkehrsgünstig, mit Erweiterungsmöglichkeiten für Produktion und Lagerung
- Freizeit/Erholung: wohnungsnah im Wohnungsumfeld, wohnungsfern in Siedlungsbereichen bzw. als Fernerholung
- Bildung: gewöhnlich zentral, z.T. aber auch dezentralisiert, dem Wohngebiet folgend, von den Wohngebieten gut erreichbar
- Versorgung: dezentral als Stadtteilversorgung, bzw. zentral im Stadtkern

Weitere Aspekte einer Bestandsaufnahme können folgende Parameter darstellen [Kor97]:

- Natürliche Gegebenheiten: Topologie, Geologie, Wasser, Klima
- Bevölkerung und Besiedlung: Bevölkerungsentwicklung, flächenrelevante Aussagen
- Bebauung
- Einrichtung des Gemeindebedarfs der sozialen Versorgung, der Ver- und Entsorgung, Handel- und Dienstleistungen
- Freiflächen: Landschaftsrahmenplanung, Grünflächenstruktur, Freiflächenutzung, ...
- Wirtschaft: Einpendlerzahlen, Gewerbeansiedlungspolitik, ...

- Verkehr: Belastungsprognosen, Schienenverkehr, Straßenverkehr, ...

Es ist leicht zu sehen, dass die Anzahl der Parameter extrem hoch und zusätzlich teilweise nicht eindeutig messbar sind. Generell kann zwischen einer einfachen und einer komplexen Bestandsaufnahme unterschieden werden. Bei der einfachen Bestandsaufnahme werden die gemessenen Parameter direkt dargestellt, wohingegen bei der komplexen Bestandsaufnahme mehrere Meßgrößen zu einem abgeleiteten (Summen-) Parameter zusammen gefasst werden [SHSR06b].

Des weiteren können qualitative von quantitativen Parametern unterschieden werden. Bei den quantitativen Parametern lassen sich beispielsweise folgende Abhängigkeiten charakterisieren:

- Zeitabhängigkeit
- Streckenabhängigkeit
- Flächenabhängigkeit
- Raumabhängigkeit

Diese sind in der Regel direkt bewertbar, während für die Bewertung der qualitativen Parameter verschiedene Techniken zur Verfügung stehen.

Als Modellierungstechniken für vorwiegend qualitative Fakten existieren u.a. Expertensysteme oder Methoden des Soft Computings, wie die Fuzzy Logic [BM96], [Her00]. Ziel des Soft Computing ist u.a. die rechnergestützte Erfassung und Simulation menschlich-intelligenten Verhaltens und die Bildung intelligenter Systeme von realen Gegebenheiten.

Der Oberbegriff Soft Computing fasst die neuen, intelligenten Paradigmen, wie die Fuzzy Logic, Neuronale Netze und genetische Algorithmen zusammen. Neuronale Netze bieten die Möglichkeit aus Erfahrungen der Vergangenheit bei Simulationsabläufen selbständig hinzuzulernen [ABA00]. Expertensysteme und ihre Kombination mit Fuzzy-Methoden eignen sich in der Form von Fuzzy-Expertensystemen zur Erfassung und zum Nachvollziehen (Simulation) menschlichen Verhaltens zur rechnergestützten Strukturierung und Formalisierung von Konflikts- und Entscheidungssituationen [GRS00], [Pup99]. Das Ziel dieser Modellierungstechniken ist es, die einzelnen Parameter untereinander vergleichbar zu machen, so dass es möglich ist diese zu bewerten.

Auch die Standortfaktoren einer Konversionsfläche unterteilen sich in qualitative und quantitative Standortfaktoren. Bei den quantitativen Standortfaktoren handelt es sich vorrangig um natürlich-technische und von Menschen geschaffene Faktoren. Beispiele sind die Verfügbarkeit von Naherholungs- und Freizeiteinrichtungen und die Erreichbarkeit von Arbeitsplätzen. Sie werden getrennt nach Faktoren aus den

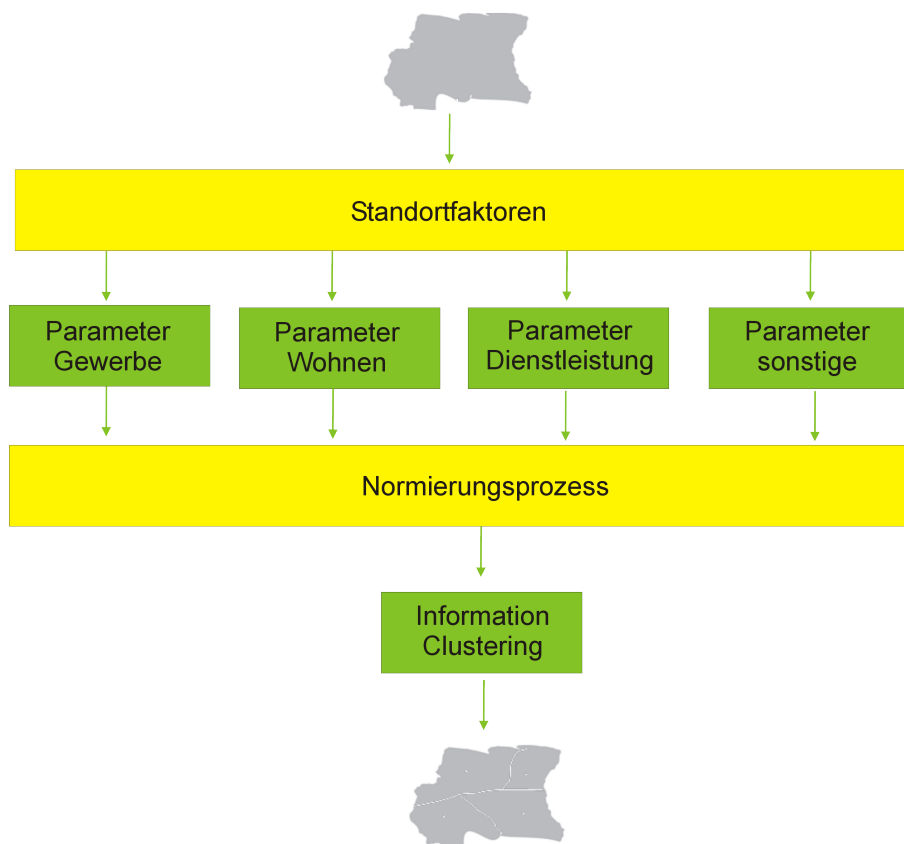


Abbildung 3.8: Parameterauswahl zum Entwurf des Nutzungskonzeptes

Bereichen Gewerbe, Dienstleistung, Wohnen und sonstige Parameter erfasst. Es hat sich im Laufe der Arbeit herausgestellt, dass in jedem der vier Bereiche die Themenfelder Raumstruktur, Nutzung, Stadtplanung, Ökologie, Ökonomie und Recht die quantitativ relevanten Parameter liefern. Beispiele für qualitative Standortfaktoren sind die Lebensqualität und das Stadt- und Regionsimage. Diese Parameter werden ebenfalls getrennt nach den Bereichen Gewerbe, Dienstleistung, Wohnen und Sonstigem erfasst und zählen zu den qualitativen Faktoren.

Sämtliche Eingabeparameter für das Information Clustering durchlaufen einen Normierungsprozess, der die Vergleichbarkeit der Parameter untereinander sichert [SHSR06a]. Die Parameter sind so zu bestimmen, dass mit ihnen die gesamte Konversionsfläche definiert wird und sie untereinander keine Abhängigkeiten aufweisen.

Abbildung 3.9 zeigt die Integration der Gewichte und der Standortfaktoren in das Information Clustering am Beispiel von Konversionsflächen. Die Planungsalternative stellt hierbei eine Flächenunterteilung dar.

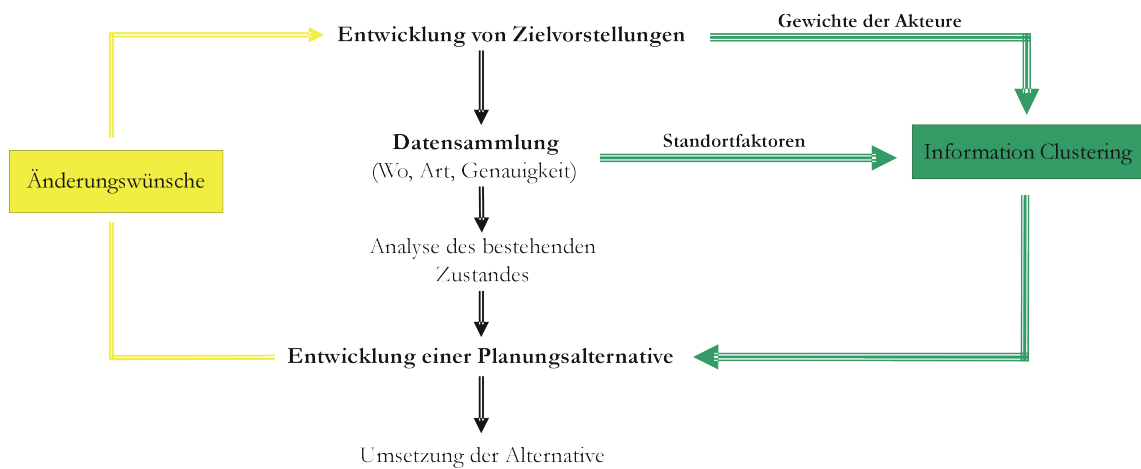


Abbildung 3.9: Integration der Gewichte und Akteure

3.3 Information Clustering

Das Ziel des Information Clustering ist es anhand der Eingabeparameter (die Standortfaktoren und die Akteure) eine Planungsalternative zu berechnen. Diese Planungsalternative repräsentiert eine Flächenunterteilung, die sämtlichen Anforderungen gerecht wird und die Interessen der Akteure kanalisiert. In dem letzten Kapitel wurde beschrieben, wo genau der Prozess des Information Clustering am Planungsablauf ansetzt. In diesem Kapitel wird nun der mathematische Aufbau des Information Clustering beschrieben.

Für das Information Clustering wird zunächst ein geeignetes Berechnungsverfahren benötigt, das anhand der Profile eine Planungsalternative berechnet. Dieses Verfahren wird in Kapitel 3.3.1 entwickelt. In Kapitel 3.3.2 wird die Notwendigkeit einer anschließenden Visualisierung erläutert sowie die Anforderungen an die Visualisierung ermittelt.

3.3.1 Generelle Zielfunktion

Die Menge der Einfluss nehmenden Parameter im Planungsablauf ist sehr groß, sodass bei der Wahl einer geeigneten Berechnungsmethode von einer großen und unstrukturierten Datenmenge auszugehen ist. Das Ziel der Berechnung ist eine Unterteilung der Fläche zu erreichen, die das Zusammenspiel der Parameter untereinander widerspiegelt und den Anforderungen der einzelnen Akteure gerecht wird. Mit Hilfe der so gewonnenen Flächenunterteilung wird eine aussagekräftige Struktur in die Daten gebracht und somit die Erstellung einer Planungsalternative ermöglicht.

Für die Berechnung eignen sich prinzipiell die in Kapitel 2.1.1 beschriebenen Clu-

steranalyseverfahren, wobei die Cluster so zu bilden sind, dass

- jeder Cluster mindestens ein Objekt enthält und
- jedes Objekt genau zu einem Cluster gehört.

Aufgrund dieser Prämissen eignen sich die partitionierenden Clusteranalyseverfahren für das in dieser Arbeit behandelte Problemfeld. Ein bewährtes Verfahren aus diesem Bereich sind die Voronoidiagramme, die in den unterschiedlichsten Anwendungsgebieten bereits mit Erfolg eingesetzt werden. Das Voronoidiagramm liefert eine Unterteilung einer Fläche in verschiedene Regionen anhand einer endlichen Menge verschiedener, isolierter Punkte. Das Konzept der Voronoidiagramme ist sehr intuitiv und flexibel und dadurch äußerst ansprechend auch für den Einsatz in der Stadtplanung.

In Kapitel 3.1.2 wurden die globalen Auswirkungen kleiner lokaler Änderungen im Planungsablauf beschrieben. Mit den Voronoidiagrammen lassen sich diese Änderungen modellieren. Das Voronoidiagramm bildet die Variation einzelner Parameter sichtbar nach. Das grundlegende Verfahren der Voronoidiagramme wird in Kapitel 4.1.1 beschrieben.

Eine besonders wichtige und zu berücksichtigende Eigenschaft bei der Berechnung ist der Zusammenhang der Flächenunterteilung. Kein Akteur hat ein Interesse an einem Gebiet, das keiner speziellen Nutzung zugeordnet werden kann. Die geometrischen und topologischen Eigenschaften der Voronoidiagramme werden deswegen in den Kapiteln 4.1.2 und 4.1.3 untersucht und die konkreten Bedingungen bei der Berechnung analysiert.

Einer der Vorteile von Voronoidiagrammen ist die Verarbeitung auch einer sehr großen Menge von Eingabeparametern. Sie eignen sich somit auch für die Verarbeitung der in Kapitel 3.2 beschriebenen multidimensionalen und heterogenen Parameter, die bei dem Planungsablauf entstehen. Es müssen sämtliche Akteure in Form einer Gewichtung integriert werden sowie die Standortfaktoren, die direkt in die Berechnung einfließen. Die Akteure gewichten für sie interessantere Parameter stärker und weniger interessante Parameter schwächer. Sämtliche Parameter werden im Umfeld eines Referenzpunktes ermittelt. Hierfür wird beispielsweise bei den Konversionsflächen eine Liste von Bezugspunkten gewählt und alle Parameter in Bezug zu diesen Referenzpunkten aufgenommen.

Als Eingabekriterien für das Clustering ergeben sich hierdurch zum einen eine Liste der Referenzpunkte

$$RP = \{RP_1, RP_2, \dots, RP_k\} \quad (3.1)$$

mit k Anzahl der Referenzpunkte und eine Liste der Parameter

$$P = \{p_1, p_2, \dots, p_m\} \quad (3.2)$$

mit m Anzahl der Parameter. Da jeder Akteur jeden Parameter gewichtet, werden die Gewichte in Form einer Matrix dargestellt:

$$G_{n,m} := \begin{pmatrix} g_{1,1} & \cdots & g_{1,m} \\ \vdots & \ddots & \vdots \\ g_{n,1} & \cdots & g_{n,m} \end{pmatrix} \quad (3.3)$$

Für die Anwendung der Voronoidiagramme ist jedoch zunächst eine geeignete Modellierung der Ähnlichkeit zwischen den Datenobjekten erforderlich. Wie bereits in Kapitel 2.1.2 beschrieben, erfolgt diese Modellierung durch eine Distanzfunktion. Für die Berechnung der Distanz werden direkte oder abgeleitete Eigenschaften der Objekte verwendet, im vorliegenden Fall also die Referenzpunkte mit den zugehörigen Parametern. In der Regel wird für die Distanzberechnung eine Metrik verwendet, die jedoch für die Modellierung der Ähnlichkeit zwischen den multidimensionalen Parametern unzureichend ist (siehe hierzu auch Kapitel 4.2). Es wird ein wesentlich flexiblerer Abstandsbegriff benötigt. In Kapitel 4.2.3 wird ein genereller Ansatz einer solchen Distanzfunktion entwickelt, mit dem es möglich ist die Ähnlichkeit zwischen den Parametern zu modellieren.

Die Parameter gehen direkt in diese Distanzberechnung ein. Pro Referenzpunkt ergibt sich eine Liste mit Parametern. Die Matrix $G_{n,m}$ wird mit der Distanzfunktion verknüpft, um so die unterschiedlichen Ansichten der Akteure zu modellieren.

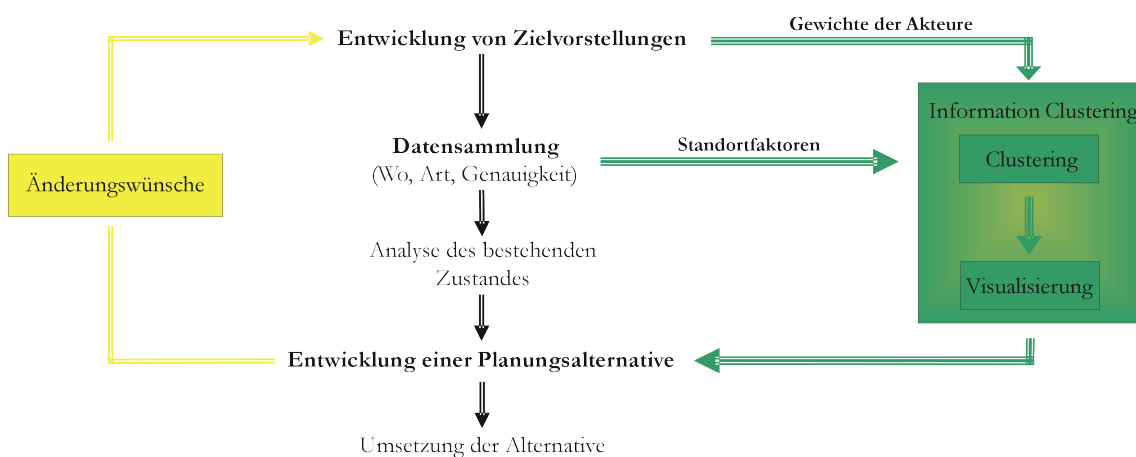


Abbildung 3.10: Information Clustering im Planungsablauf

Für die Verknüpfung existieren unterschiedliche Alternativen, die einen wesentlichen Einfluss auf die entstehende Unterteilung, das Voronoidiagramm, ausüben. Kapitel 4.2.2 entwickelt verschiedene Ansätze für eine in diesem Anwendungsfeld geeignete Verknüpfungsart.

Die Akteure werden ebenfalls in Form einer Liste erfasst:

$$A = \{a_1, a_2, \dots, a_n\} \quad (3.4)$$

mit n Anzahl der Akteure. Die Zielfunktion setzt sich also folgendermaßen zusammen:

$$ZFK = f(RP_k, P_m, A_n, G_{n,m}) \quad (3.5)$$

Abbildung 3.10 beschreibt den endgültigen Ausbau des Information Clusterings und die Einordnung im Planungsablauf.

Die Zielfunktion ZFK liefert eine Planungsalternative in Form einer Flächenunterteilung. Es entstehen mindestens k Regionen im n -dimensionalen Raum, wobei jeder Region ein Referenzpunkt RP_k zugeordnet sein sollte. Die Akteure diskutieren diese Alternative und nehmen evtl. Änderungen an den Eingabeparametern vor, was aufgrund der hohen Parameteranzahl jedoch schwierig ist. Für eine solche Evaluation des Ergebnisses eignet sich besonders der Einsatz einer Visualisierung. Das folgende Kapitel 3.3.2 beschreibt, wie durch eine Visualisierung das Problem beseitigt wird.

3.3.2 Visualisierung

Wie bereits in Kapitel 2.2 erläutert hilft eine Visualisierung bei der Analyse und Interpretation von Daten, da sie sich u.a. positiv auf die Gedächtnisleistung und die menschliche Informationsverarbeitung auswirkt. Die Daten, die im Rahmen des Planungsablaufes entstehen, lassen sich dem Gebiet der Informationsvisualisierung zuordnen, da sie im wesentlichen in abstrakter Form vorliegen und räumlich nicht angeordnet sind.

Im Bereich der Informationsvisualisierung existieren unterschiedlichste Techniken, die meistens für einen konkreten Anwendungsfall entwickelt wurden. Die meisten dieser Techniken lassen sich nur schwer auf den vorliegenden Fall übertragen, da sie entweder nicht fähig, sind die enorme Menge der Parameter zu verarbeiten bzw. diese anschaulich darzustellen oder aber schwer zu interpretieren sind.

Die gesuchte Visualisierung sollte fähig sein die einzelnen Parameter darzustellen und zusätzlich die Gewichte der Akteure zu repräsentieren. Von Vorteil wäre es, mit der Visualisierung eine natürliche Nähe zu der Fläche herstellen zu können. Hierfür eignen sich die Techniken aus dem Bereich der Iconbasierten Visualisierungstechniken. Diese sind besonders vorteilhaft bei der Darstellung von diskreten, multivariaten Daten, welche in visuelle Texturen umgewandelt werden. Ein Beispiel hierfür ist in Abbildung 3.11 zu sehen. In Kapitel 5 wird auf der Basis dieser Technik eine geeignete Visualisierung entwickelt.

Der Vorteil der Iconbasierten Techniken die Daten räumlich zusammenhängend darstellen zu können, birgt aber auch Nachteile. Durch den räumlichen Zusammenhang

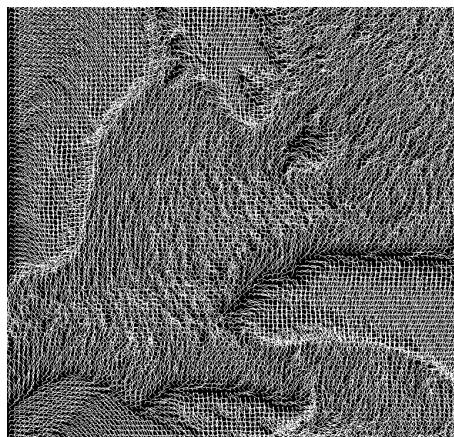


Abbildung 3.11: Visualisierung mit dem Velcro-Icon [RP88]

überdeckt die Visualisierung eine darunter liegende Karte. Zudem verliert die Technik bei einer extrem hohen Parameteranzahl den visuellen Effekt. Es ist somit sinnvoll eine weitere Visualisierungstechnik in Betracht zu ziehen. Vor allem die Techniken der Botanischen Informationvisualisierung sind aufgrund der kontextsensitiven Darstellung der Daten in Bezug zu einer Karte erfolgversprechend. Diese Techniken orientieren sich an botanischen Vorgaben und übertragen somit geübte sensorische Fähigkeiten der Anwender auf die Visualisierung. In Kapitel 5.3 wird diese Technik näher erläutert und eine passende Variante für den vorliegenden Fall entwickelt.

Kapitel 4

Clusteranalyse

Im letzten Kapitel wurde die Struktur des Information Clustering aufgebaut. Es hat sich gezeigt, dass dies ein zweistufiger Prozess ist, der sowohl ein Clusteringverfahren als auch eine Visualisierung beinhaltet. Für das benötigte Clustering wird in diesem Kapitel das Verfahren der Voronoidiagramme entscheidend erweitert. Basis dieser Erweiterungen bilden die in Kapitel 2 dargestellten Grundlagen sowie die in Kapitel 3.3.1 ermittelten Anforderungen. Die anschließende Visualisierung erfolgt in Kapitel 5.

In Kapitel 4.1 wird das Konzept der Voronoidiagramme eingeführt. Hierzu werden deren Topologie und deren geometrische Eigenschaften analysiert.

Durch die Entwicklung einer generellen Distanzfunktion werden gravierende Mängel beim Einsatz von Metriken vermieden. Dieses Verfahren wird in Kapitel 4.2 entwickelt. Eine besondere Beachtung in diesem Zusammenhang muss der Topologie der mit den Distanzfunktionen berechneten Voronoidiagrammen geschenkt werden. Herkömmliche Algorithmen zur Berechnung von Voronoidiagrammen dienen in erster Linie einer effizienten Berechnung in der Ebene. In Kapitel 4.3 werden neue Verfahren vorgestellt, die flexibel genug sind die in 4.2 entwickelten Distanzfunktionen effizient zu integrieren.

4.1 Verfahren

Das Konzept der Voronoidiagramme ist sehr intuitiv und dadurch äußerst ansprechend für die unterschiedlichsten Anwendungsfelder, in denen es mit Erfolg eingesetzt wird. Die Voronoidiagramme zählen zu der Klasse der partitionierenden Clusteranalyseverfahren, da jedes Objekt genau einem Cluster zugeordnet wird (siehe hierzu auch Kapitel 2.1.1).

Der Vorteil der Nutzung dieser generellen Struktur, für den in Kapitel 3.3 beschriebenen Prozess des Information Clustering ist zum einen deren natürliche Nähe zu

Abstandsbeschreibungen und zum anderen die Möglichkeit eine große und heterogene Menge Parameter verarbeiten zu können.

Ein weiterer enormer Vorteil für die in dieser Arbeit betrachtete Anwendung ist die Globalisierung der Voronoidiagramme. In dem Bereich der Planung werden globale Analyseverfahren benötigt, da kleine lokale Änderungen im Planungsablauf eine große globale Auswirkung haben (siehe Kapitel 3.1). Mit Voronoidiagrammen lassen sich diese Anforderungen umsetzen, da durch die Möglichkeit der Integration der multidimensionalen Parameter (siehe Kapitel 3.2 und 3.3) leicht lokale Änderungen im Planungsablauf berücksichtigt werden können und deren globale Auswirkungen direkt sichtbar sind.

Im folgenden Kapitel 4.1.1 wird ein kurzer formaler Überblick über Voronoidiagramme gegeben. Anschließend wird in Kapitel 4.1.2 die Topologie der Voronoidiagramme untersucht. Für deren Einsatz in der Stadtplanung sind besonders die Erhaltung einiger topologischer Attribute bei der Erzeugung der Voronoidiagramme wichtig. In Kapitel 4.1.3 werden geometrische Eigenschaften der Voronoidiagramme analysiert, die in einem direkten Zusammenhang zu deren Einsatz in der Stadtplanung stehen.

4.1.1 Voronoidiagramme

Das Voronoidiagramm unterteilt einen Raum in unterschiedliche Regionen anhand einer endlichen Menge verschiedener, isolierter Punkte im Raum. Aufgrund der variablen Einsatzmöglichkeit dieser Prämisse ist es nicht verwunderlich, dass das Konzept der Voronoidiagramme mehrfach 'erfunden' und mit diversen Namen titulierte wurde. Prinzipiell kann zwischen einfachen und generellen Voronoidiagrammen unterschieden werden. Hierbei bezieht sich ein einfaches Voronoidiagramm auf den zweidimensionalen Raum bei dem sämtliche Punkte gleichwertig sind, also identisch bis auf ihre Lage. Für die Abstandsdefinition werden bei einem einfachen Voronoidiagramm Metriken (siehe Kapitel 2.1.2) verwendet.

Generelle Voronoidiagramme stellen Verallgemeinerungen der einfachen Voronoidiagramme dar und finden ihre Begründung in den verschiedenen Anwendungsfeldern. Von besonderem Interesse im Rahmen dieser Arbeit sind die gewichteten Voronoidiagramme und die Voronoidiagramme mit einem erweiterten Abstandsbegriff.

Ein einfaches Voronoidiagramm unter der euklidischen Metrik besteht aus einer Anzahl von Polygonen. Da Polygone mittels Halbebenen definiert werden können, lässt sich auch das Voronoidiagramm durch sie definieren. Der hierfür benötigte Bisektor b beschreibt die Gerade, die senkrecht auf der Strecke zwischen zwei Voronoi-Punkten $\overline{p_i p_j}$ steht. Die Funktion d beschreibt den Abstand zwischen zwei Punkten.

$$b(p_i, p_j) = \{x | d(x, p_i) = d(x, p_j)\}, j \neq i \quad (4.1)$$

Der Bisektor teilt eine Ebene in zwei Halbebenen:

$$H(p_i, p_j) = \{x | d(x, p_i) < d(x, p_j)\}, j \neq i \quad (4.2)$$

Die Voronoiregion (siehe Abbildung 4.1) wird nun basierend auf den Halbebenen

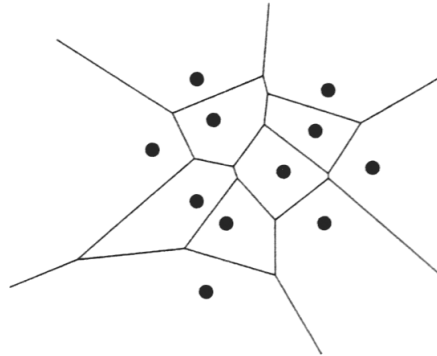


Abbildung 4.1: Einfaches Voronoidiagramm

(Abbildung 4.2) folgendermaßen definiert:

Definition 12 Sei $P = \{p_1, \dots, p_n\} \subset \mathbb{R}^2$ mit $2 \leq n < \infty$ und $p_i \neq p_j$. Die Region

$$V(p_i) = \bigcap_{p_j \in P \setminus \{p_i\}} H(p_i, p_j) \quad (4.3)$$

heißt **Voronoiregion**.

Das Voronoidiagramm wiederum ist definiert als das Komplement zu der Vereinigung aller Voronoiregionen in der Ebene. Der Spezialfall mit $n = 1$ wird hier nicht berücksichtigt, da das entstehende Voronoidiagramm trivial ist. Es besteht in diesem Fall aus nur einem Voronoipolygon, nämlich der gesamten Ebene. Des weiteren werden ausschließlich Voronoidiagramme mit endlich vielen Voronoipunkten betrachtet, da in dem betrachteten Anwendungsgebiet stets endlich viele Bezugsquellen vorliegen. Voronoidiagramme mit $n = \infty$, wie z.B. das Poisson Voronoidiagramm, können beispielsweise in dem Buch von de Berg, van Kreveld, Overmars und Schwarzkopf [dBvKOS00] nachgelesen werden.

Ein besonderer Pluspunkt der Voronoidiagramme und dadurch eine Grundlage für deren kontinuierlichen Erfolg, ist die Möglichkeit das Konzept der Voronoidiagramme zu generalisieren. Nicht zuletzt deswegen eignen sie sich auch besonders für den Prozess des Information Clustering in der Stadtplanung. Die Generalisierung kann in unterschiedlicher Weise erfolgen. Interessant für diese Arbeit sind besonders Voronoidiagramme, bei denen die einzelnen Punkte unterschiedlich stark gewichtet

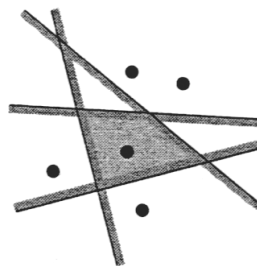


Abbildung 4.2: Voronoizelle aus Schnitt von Halbräumen

werden. Dies ermöglicht die Integration von Attributen an den einzelnen Voronoi-punkten und somit die Einbindung der in Kapitel 3.2 bestimmten Parameter und Gewichte der Akteure. Eine detaillierte Beschreibung der Einbindung der Parameter und der sich hierdurch ergebenden geometrischen Eigenschaften der gewichteten Voronoidiagramme findet sich in Kapitel 4.2.2.

Es existieren diverse weitere Generalisierungsmöglichkeiten für Voronoidiagramme, die ihren Sinn in den verschiedenen Anwendungen haben. Order-k-Voronoidiagramme eignen sich u.a. für Ortslokalisierungsprobleme. Ist zum Beispiel der k-te nächste Notarzt bereits im Einsatz oder es werden mehr Notärzte benötigt, können durch diese Art der Voronoidiagramme weitere Ärzte k-ter Ordnung lokalisiert werden. Andere Einsatzgebiete für Order-k-Voronoidiagramme sind multivariate Dichteschätzungen oder die Kartographie. Dynamische Voronoidiagramme eignen sich besonders für sich ändernde Situationen, beispielsweise bei einer mobilen Radiostation oder um im Bereich des motion planning einen kollisionsfreien Weg eines Roboters um sich bewegend Objekte zu finden.

Ein detaillierter Überblick über die Entstehung der Voronoidiagramme und die verschiedenen Ansätze findet sich in dem Buch von Okabe et al. [OBSC99].

4.1.2 Topologie

Topologie ist die Lehre von allgemeinen räumlichen Beziehungen und Eigenschaften des Raumes und wird spezifisch in verschiedenen Wissenschaften gebraucht. Die Topologie als Teilgebiet der Mathematik beschäftigt sich mit denjenigen Eigenschaften eines Gegenstandes, die unter kontinuierlichen Verformungen erhalten bleiben. Die Summe dieser Eigenschaften wird auch die „Topologie“ des Gegenstandes genannt. Eine Menge, auf der eine Topologie definiert ist, heißt topologischer Raum.

Die Topologie untersucht Eigenschaften geometrischer Gebilde, die unter bijektiven, umkehrbar stetigen Abbildungen erhalten bleiben. Eine solche Eigenschaft ist etwa „zusammenhängend“ zu sein, während Längen und Volumina unter allgemeinen stetigen Abbildungen nicht zwingend erhalten bleiben.

Der axiomatische Aufbau der modernen Topologie beruht auf dem grundlegenden Konzept der Nachbarschaft, formalisiert als offene Umgebung. Neben offen und abgeschlossen gibt es als weitere fundamentale topologische Attribute stetig, kompakt, separabel, zusammenhängend, dicht und abzählbar. Neben der Algebra kann die Topologie als zweiter Stützpfiler für alle anderen Felder der Mathematik angesehen werden. Sie ist besonders wichtig für die Geometrie.

Definition 13 : Eine abgeschlossene Menge G heißt **konvex**, wenn für je zwei Elemente $p, q \in G$ jeder Punkt der Strecke \overline{pq} in G enthalten ist.

Es lässt sich einfach zeigen, dass der Schnitt konvexer Mengen wieder konvex ist. Hingegen ist die Vereinigung konvexer Mengen nicht notwendigerweise konvex, selbst wenn der Schnitt nicht leer ist. Damit ist entgegen dem üblichen Verständnis eine konvexe Hülle immer eine gefüllte Menge. Um von der Umhüllung im umgangssprachlichen Sinne zu reden, wird der Terminus „Konturlinie“ gewählt.

Definition 14 : Ein Gebiet G heißt **sternförmig**, wenn es mindestens einen Punkt p in G gibt, von dem aus alle anderen Punkte $q \in G$ direkt sichtbar sind, das heißt jede Strecke \overline{pq} liegt vollständig in G .

Die Gesamtheit aller Punkte p wird als **Kern** von G , kurz $\ker(G)$, bezeichnet.

Aus diesen beiden Definitionen ist ersichtlich, dass folgender Zusammenhang gilt:

$$G \text{ ist konvex} \iff \ker(G) = G \quad (4.4)$$

„Topologieerhaltend“ bedeutet in diesem Kontext, dass die entstehenden Voronoidiagramme dieser Arbeit zusammenhängend bezüglich der Topologie sind. Es sollen also keine Gebiete entstehen, welche von ihrem Referenzpunkt topologisch getrennt sind. Solche Gebiete werden auch als Löcher bezeichnet und bedarfen in geometrischen Anwendungen einer besonderen Beachtung. Bei den betrachteten planerischen Vorhaben beispielsweise bei Konversionsprojekten bedeuten Löcher in der Unterteilung Flächen, die keiner späteren Nutzung zuzuordnen sind. Diese Art der Flächen sind zu vermeiden (siehe Kapitel 3)!

Als natürliche Abstandsmaße dienen die in Kapitel 2.1.2 definierten Metriken. Jedoch schränkt eine Metrik, wie in Kapitel 4.2 näher erläutert wird, den Abstandsbe-griff zu stark ein. Eine Metrik definiert auf einem Raum eine Topologie, umgekehrt ist einer Topologie nicht unbedingt eine Metrik zuzuordnen. Daher ist die Topologie der allgemeinere Begriff, der wesentlich mehr Modellierungsfreiräume zulässt.

4.1.3 Geometrische Eigenschaften

Nach der formalen Einführung der Voronoidiagramme und deren Topologie werden nun ihre geometrischen Eigenschaften behandelt. Wie in Kapitel 3 dargestellt, sind

besonders die Konvexität und der Zusammenhang der Voronoidiagramme interessant. Es gibt jedoch noch weitere Eigenschaften der Voronoidiagramme, die für den Anwendungsbereich der Stadtplanung zu beachten sind und im Folgenden untersucht werden.

Einfache Voronoidiagramme haben die Eigenschaft, dass die einzelnen Voronoiregionen eine konvexe Struktur aufweisen und in jeder Region ein Schwerpunkt vorhanden ist. Es existieren keine leeren Voronoiregionen. Es gilt somit folgender Satz, dessen Beweis bei Okabe, Boots, Sugihara und Chiu [OBSC99] nachgelesen werden kann:

Satz 1 Sei $P = \{p_1, \dots, p_n\} \subset \mathbb{R}^2$ eine Menge $2 \leq n < \infty$ verschiedener Punkte in der euklidischen Ebene. Die Menge $V(p_i) = \{x \mid \|x - p_i\| \leq \|x - p_j\| \text{ für } i \neq j, i, j \in I_n\}$ ist ein **nicht-leeres konvexes Polygon**.

Bei der Betrachtung von Voronoidiagrammen unter anderen Metriken trifft dieser Satz nicht unbedingt zu. Wird für die Distanzberechnung beispielsweise die Manhattan-Metrik (siehe Kapitel 2.1.2) anstatt der euklidischen Metrik verwendet, dann ist die Konvexität des Voronoidiagrammes nicht mehr garantiert. Die Konvexität hängt also sehr stark von dem eingesetzten Abstandsbegriff ab. In Kapitel 4.2 wird dieser Zusammenhang untersucht und Kriterien für dessen Einsatz entwickelt.

Auch die zweite Eigenschaft in Satz 1 lässt sich ohne Einschränkungen nicht verallgemeinern. Durch die Einbindung von Gewichten bei generellen Voronoidiagrammen können einzelne, leere Polygone entstehen. Für den praktischen Einsatz sind

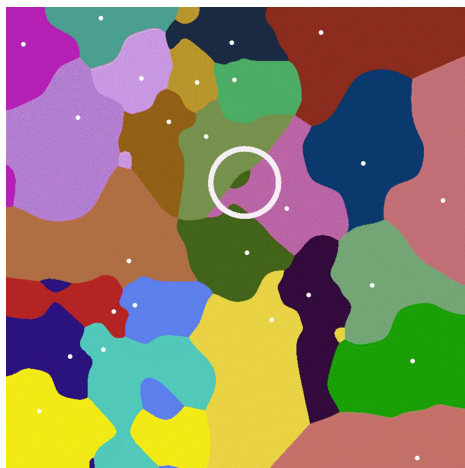


Abbildung 4.3: Gewichtetes Voronoidiagramm mit leerer Voronoiregion

diese Gebiete allerdings nicht wünschenswert, da ihnen nicht direkt ein Schwerpunkt mit den zugehörigen Attributen zuzuordnen ist. Auf Grund dessen werden in Kapitel 4.2.2 Eigenschaften der Voronoidiagramme aufgezeigt, die das Auftre-

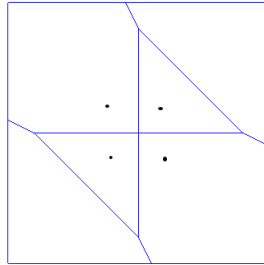


Abbildung 4.4: Degeneriertes Voronoidiagramm

ten leerer Polygone indizieren. Abbildung 4.3 zeigt ein Voronoidiagramm mit einer nicht-zusammenhängenden Region. Für den praktischen Einsatz bei der Konversion würde dies einem Gebiet entsprechen, dem keine Nutzung zugeordnet werden kann!

Eine weitere problematische Eigenschaft stellen degenerierte Voronoidiagramme dar, die auftreten können, wenn die Voronoi punkte regulär verteilt sind. In manchen Anwendungen benötigen degenerierte Voronoidiagramme spezielle Längenbehandlungen. Um diese Schwierigkeit zu vermeiden, werden die degenerierten Fälle durch die Festlegung der Anzahl der Voronoi kanten ausgeschlossen. Hierzu sollte jede Voronoi ecke genau drei Voronoi kanten haben. In Abbildung 4.4 wird ein degeneriertes Voronoidiagramm dargestellt, bei dem ein Voronoi punkt mit vier Kanten verbunden ist. In Abhängigkeit von dem verwendeten Abstandsbegriff kann die Prämisse der maximalen Kantenzahl nicht eingehalten und somit können degenerierte Voronoidiagramme nicht ausgeschlossen werden.

Die Konvexität der Einheitsumgebung steht in keinem direkten Zusammenhang mit der Konvexität der Voronoi regionen. Die in Kapitel 2.1.2 vorgestellten Minkowski-Konturen besitzen alle eine konvexe Einheitsumgebung, jedoch ist das Voronoidiagramm beispielsweise der Manhattan-Metrik und der Supremum-Metrik nicht unbedingt konvex. In Kapitel 4.2.1 wird beschrieben, dass die Konvexität des Voronoidiagramms wesentlich durch die Bisektoren bestimmt werden.

4.2 Distanzfunktion

Metriken führen bei dem Clustern mit Voronoidiagrammen zu dem Problem, dass die Bisektoren Flächen aufspannen können. In diesem Fall lassen sich die Voronoi regionen nur dann berechnen, wenn den Bisektoren ein eindeutiger Wert zugeordnet

wird. Um dieses Problem zu vermeiden hat sich der Ansatz mit Distanzfunktionen bewährt.

Da die Distanzfunktionen jedoch weniger Axiome erfüllen als die Metriken (z.B. muss die Symmetrie nicht eingehalten werden), ist es nötig die sich dadurch ergebenden Auswirkungen zu untersuchen. Eine besondere Bedeutung kommt der Topologie der Voronoidiagramme zu, die mit den Distanzfunktionen berechnet werden. Die Topologie der Voronoidiagramme unter Metriken wurden in Kapitel 4.1 untersucht und die Anforderungen an die entstehende Unterteilung im praktischen Einsatz bei Planungsvorhaben in der Stadtplanung herausgearbeitet. Im folgenden werden diese Überlegungen auf Distanzfunktionen ausgeweitet und ihr Einfluss auf die Topologie der Voronoidiagramme untersucht.

Alternativ zu Definition 12 wird eine Voronoiregion $V(p_i)$ mittels einer **Distanzfunktion** d folgendermaßen definiert:

$$V(p_i) = \{p | d(p, p_i) < d(p, p_j), j \neq i\} \quad (4.5)$$

In Kapitel 4.2.1 werden zunächst konvexe Distanzfunktionen eingeführt und deren geometrische Eigenschaften bei dem Clustern mit Voronoidiagrammen untersucht. In dieser Arbeit wird eine Erweiterung der Distanzfunktionen für die Integration der in Kapitel 3.3.1 beschriebenen Gewichte-Matrix der Akteure dringend benötigt. Um diese Attribute berücksichtigen zu können, werden in Kapitel 4.2.2 Gewichte ω_i in die Distanzfunktion eingebunden und deren Einfluss beschrieben. In Kapitel 4.2.3 wird ein genereller Ansatz dieser Arbeiten entwickelt.

4.2.1 Konvexe Distanzfunktion

Konvexe Distanzfunktionen bieten eine gute Alternative zu dem Einsatz von Metriken. Da sie weniger Axiome erfüllen als die Metriken sind sie flexibler einsetzbar und sie vermeiden zudem das Auftreten flächiger Bisektoren und die damit verbundenen Probleme.

In Abbildung 4.5 sind mögliche Bisektoren der Manhattan Metrik dargestellt. Das linke Bild zeigt den typischen Fall, der immer dann auftritt, wenn zwei Punkte p und q ein echtes Rechteck aufspannen, das also kein Liniensegment und kein Quadrat ist. Der Bisektor besteht aus zwei Halbgeraden und einem Liniensegment mit Winkel $\pi/4$ zu den Koordinatenachsen. Haben p und q hingegen identische x- oder y-Koordinaten, so stimmt ihr Bisektor mit dem euklidischen Bisektor überein. Spannen p und q aber ein Quadrat auf, so enthält ihr Bisektor zwei volle Viertelebenen. Das Phänomen flächiger Bisektoren tritt nur dann auf, wenn die Einheitssphäre C bezüglich der jeweiligen Metrik teilweise abgeplattet ist, d.h. Liniensegmente in ihrem Rand enthält, und wenn außerdem die Punkte p und q auf einer Geraden liegen,

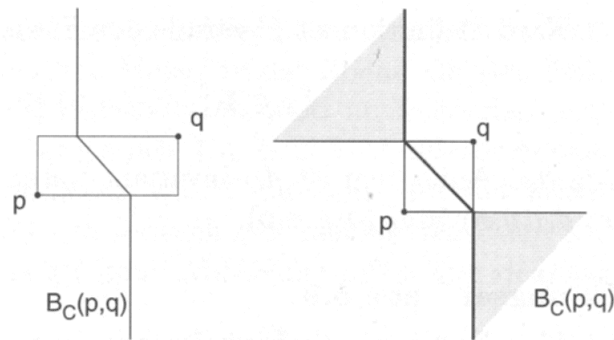


Abbildung 4.5: Bisektor der Manhattan-Metrik

die parallel zu solch einem Randsegment von C ist. Ist dagegen die Norm, welche die Metrik induziert, streng konvex, dann hat die Einheitssphäre keine linearen Segmente und die Bisektoren können keine Flächenstücke enthalten.

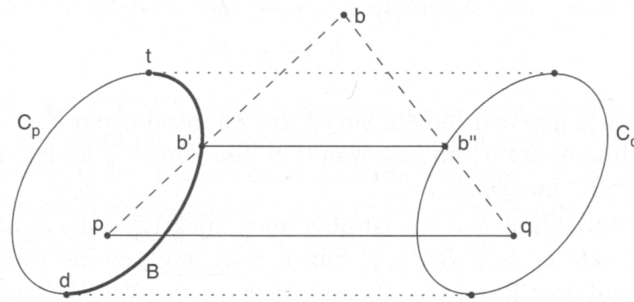


Abbildung 4.6: Ist ein Einheitskreis C streng konvex, so ist der Bisektor von zwei Punkten immer ein Weg.

Satz 2 Sei C die Einheitssphäre einer streng konvexen Metrik in der Ebene. Dann ist jeder Bisektor $B_C(p, q)$ bzgl. der konvexen Distanzfunktion homöomorph zu einer Geraden.

Beweis: Es werden zwei in p und q zentrierte Kopien C_p, C_q von C betrachtet, wie in Abbildung 4.6 dargestellt. Damit sie disjunkt sind, sollen sie gegebenenfalls hinreichend verkleinert werden. Sei b ein Punkt außerhalb von C_p und C_q ; mit b' und b'' werden die Schnittpunkte von pb und pq mit den Rändern von C_p und C_q bezeichnet.

Es gilt:

$$b \in B_C(p, q) \Leftrightarrow \quad (4.6)$$

$$d_C(p, b) = d_C(q, b) \Leftrightarrow \frac{|bp|}{|b'p|} = \frac{|bq|}{|b''q|} \Leftrightarrow \quad (4.7)$$

$$b'b'' \parallel pq. \quad (4.8)$$

Von unten nach oben gelesen folgt aus obiger Äquivalenz: Seien b', b'' zwei Punkte auf den einander zugewandten Seiten von C_p und C_q mit $b'b'' \parallel pq$. Dann schneiden sich die Strahlen von p durch b' und von q durch b'' in einem Punkt b , der zu $B_C(p, q)$ gehört. Seien $t, d \in \partial C_p$ die Berührungspunkte der oberen und unteren gemeinsamen Tangenten von C_p und C_q , und sei B der C_q zugewandte Teil von $\partial C_p \setminus \{t, d\}$. Die oben eingeführte Abbildung vom Bisektor $B_C(p, q)$ zum Randstück B ist ein Homöomorphismus.

$$f : B_C(p, q) \rightarrow B, b \rightarrow b' \quad (4.9)$$

(i) *f* ist surjektiv: Zu jedem b' läßt sich ein b'' im zugewandten Teil von ∂C_p finden so, dass $b'b''$ parallel zu pq ist.

(ii) *f* ist injektiv: Gäbe es zwei Punkte $b_1, b_2 \in B_C(p, q)$ mit $b'_1 = b'_2$, so müssten die Punkte b''_1 und b''_2 auf dem Rand von C_q verschieden sein. Weil $b'_1 b''_1$ und $b'_1 b''_2$ parallel zu pq sind, müssten b''_1, b''_2 auf einer Geraden liegen. Und weil b''_1 und b''_2 auf dem C_p zugewandtem Teil des Randes von C_q liegen, zwischen den Berührungspunkten der oberen Tangente, muss der Rand von C_q zwischen b''_1 und b''_2 verlaufen, was im Widerspruch zur strengen Konvexität steht; siehe Abbildung 4.7.

Da diese Zuordnung $b \rightarrow q$ stetig ist, ist die Homöomorphie zu B gezeigt. B seinerseits ist aber homöomorph zu einer Geraden. Damit ist der Satz bewiesen. Die

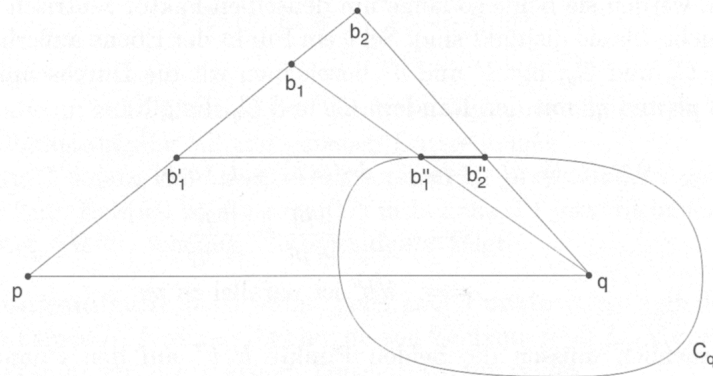


Abbildung 4.7: Wenn $b''_1, b''_2 \in \partial C_p$ auf der C_p zugewandten Seite und mit b'_1 auf einer zu pq parallelen Geraden liegen, enthält ∂C_p das Liniensegment b''_1, b''_2 .

geometrische Anschauung illustriert Abbildung 4.8, der ausführliche Beweis kann bei H. Mattes [Mat02] nachgelesen werden.

Um den Abstand von p nach q bezüglich C zu definieren, wird die Einheitssphäre C um den Vektor p verschoben. Der Strahl von p durch q schneidet den Rand von C in genau einem Punkt q' . Eine **konvexe Distanzfunktion** von p und einem Punkt q

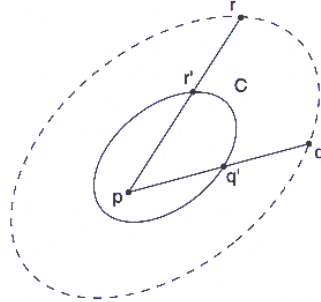


Abbildung 4.8: konvexe Distanzfunktion

wird wie folgt definiert:

$$d_C(p, q) = \frac{|pq|}{|pq'|} \quad (4.10)$$

Zusätzlich sei $d_C(p, p) = 0$. d_C erfüllt die positive Definitheit und die Dreiecksungleichung. Die Symmetrie ist nur erfüllt, wenn C in Bezug zu dem Nullpunkt symmetrisch ist. In diesem Fall wäre d_C auch eine Metrik. Somit sind die Minkowski-Metriken Spezialfälle der konvexen Distanzfunktion. Das Voronoidiagramm ändert seine topologische Eigenschaften nicht in Abhängigkeit von der Symmetrie der Distanzfunktion, d.h. die Symmetrie hat keinen Einfluss auf den Zusammenhang und die Konvexität des Voronoidiagrammes.

Da die Bisektoren bei allgemeinen streng konvexen Distanzfunktionen nicht notwendigerweise Geraden sind, sind die Voronoiregionen auch nicht notwendigerweise konvex, wie das bei der euklidischen Metrik der Fall war. Sie sind aber sternförmig.

Satz 3 Für eine gegebene Distanzfunktion d sind die möglichen Voronoiregionen star-shaped, wenn die Dreiecksungleichung für d gilt.

Beweis: Angenommen, es gibt eine Voronoiregion mit dem zugehörigen Voronoi-punkt p die nicht sternförmig ist. Dann existiert ein Punkt y in der Region p , und ein Punkt x auf dem Segment py , der näher zu einem anderen Voronoi-punkt q als zu p ist. Es gilt also $d(q, x) < d(p, x)$. Dies bedeutet, dass $d(q, x) + d(x, y) < d(p, x) + d(x, y) = d(p, y)$ ist. Aus der Dreiecksungleichung folgt jedoch $d(q, y) \leq d(q, x) + d(x, y)$. Es folgt direkt, dass $d(q, y) < d(p, y)$ ist. Das ist ein Widerspruch zu der Annahme, dass y in der Voronoiregion von p liegt [CD85].

Das bedeutet, obwohl sowohl die Manhattan- als auch die Supremum-Metrik (-Distanzfunktion) nicht konvex sind, sind ihre Regionen doch zumindest star-shaped.

4.2.2 Gewichtete Distanzfunktion

Im allgemeinen sind die Referenzpunkte eines Voronoidiagramms gleichmäßig gewichtet und nicht mit variierenden Gewichten kombiniert. Um jedoch einen flexibleren Ansatz zu erhalten, können Attribute in Form von zusätzlichen Gewichten in die Distanzfunktion integriert werden. Gewichte erlauben die Simulation von wissenschaftlichen Phänomenen oder unterstützen die Visualisierung bisher unsichtbarer Datencluster in vielen Anwendungsgebieten, wie auch der Umweltinformatik. In Kapitel 3 wurde der Aufbau der generellen Zielfunktion beschrieben. Hierbei ist die Art und Weise, wie die Parameter und die Akteure integriert werden, besonders wichtig. In Abhängigkeit von der mathematischen Verknüpfung z.B. der Gewichte mit der Distanzfunktion ändern sich die topologischen Eigenschaften des Voronoidiagrammes. Die am häufigsten verwendeten Gewichte sind „additiv“, „multiplikativ“ und „zusammengesetzt“. Im folgenden werden diese vorgestellt und ihre charakteristischen Eigenschaften beschrieben.

Additive Gewichtung

In der Praxis werden die Referenzpunkte eines Voronoidiagrammes oft mit wichtigen Einrichtungen (Dienstleistern) assoziiert. Erscheinen diese Einrichtungen zeitlich versetzt, beispielsweise neu hinzukommende Versorgungsstationen, so folgt die Gebietsaufteilung dem Johnson-Mehl Model [OBSC99]. Dies entspricht einem Voronoidiagramm basierend auf einer Distanzfunktion, welche zusätzlich zum Abstandsmaß noch das Gewicht des Referenzpunktes additiv verknüpft.

Definition 15 Die *additiv gewichtete Distanzfunktion* wird definiert als:

$$d_{aw}(p, p_i) = \|p - p_i\| - \omega_i \quad (4.11)$$

Eine Möglichkeit zur Darstellung der Gewichte sind Gewichtskreise um die Referenzpunkte, wobei die Radien genau den Gewichten entsprechen (siehe Abbildung 4.9).

Das additiv gewichtete Voronoidiagramm hat drei entscheidende Eigenschaften:

1. Die additiv gewichtete Voronoi Zelle mit Gewichtskreis σ neben n anderen Gewichtskreisen $\{\sigma_i, 0 \leq i < n\}$ ist leer, genau dann, wenn ein i existiert mit $0 \leq i < n$, so dass der Gewichtskreis um σ vollständig im Inneren einer anderen Voronoi Zelle V_i liegt.

Mit anderen Worten bedeutet dies, sobald die Gewichtung umliegender Referenzpunkte einen Schwellenwert übersteigt, trägt die verhältnismäßig schwache

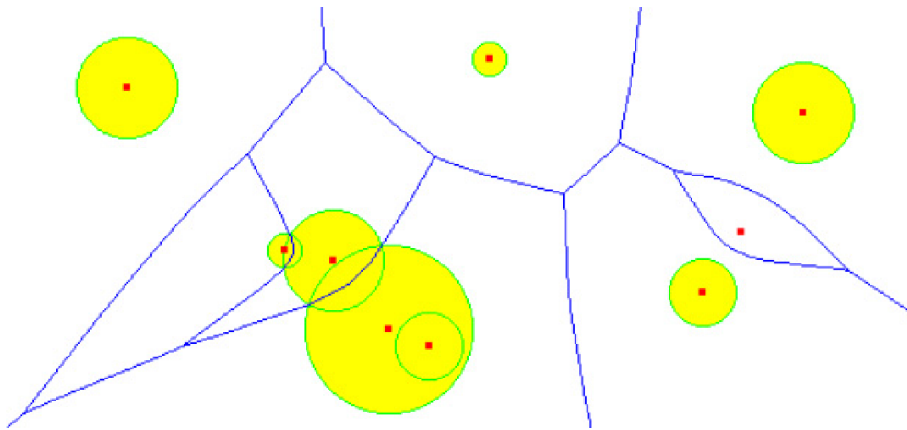


Abbildung 4.9: Additiv gewichtetes euklidisches Voronoidiagramm mit gewichteten Kreisen

- Gewichtung im Inneren nichts mehr zum Voronoidiagramm bei und die Zelle bleibt leer.
2. Die additiv gewichtete Voronoizelle ist nicht zwingend konvex, aber sternförmig. Daraus resultieren immer topologisch zusammenhängende, additiv gewichtete Voronoigebiete, falls die gewählte Konturlinie konvex ist.
 3. Der Bisektor zweier Referenzpunkte ist
 - (a) leer, falls Eigenschaft 1 erfüllt ist,
 - (b) eine Halblinie, falls der kleinere Gewichtskreis im Inneren des größeren liegt und die Kreislinien sich berühren,
 - (c) eine Gerade bei gleichen Punktgewichten,
 - (d) in allen anderen Fällen ein Teilstück einer Hyperbel, gekrümmt um den kleineren Gewichtskreis.

Durch die Eigenschaften dieser Art der Gewichtung wird deutlich, dass die additive Gewichtung als solches immer topologieerhaltend ist. Lediglich die Wahl einer nicht-konvexen Distanzfunktion kann diesen Zusammenhang verhindern [Sch04b].

Multiplikative Gewichtung

Naturwissenschaftliche Wachstumsprozesse, beispielsweise das Kolonienwachstum von Bakterienkulturen oder Kristallisierungsprozesse in Mineralien, zeichnen sich oft durch unterschiedlich schnelles Wachstum in den einzelnen Keimpunkten aus. Eine Simulation dieser Phänomene bietet das Apollonius-Modell [OBSC99]. Dabei wachsen die Gebiete, ausgehend von den Referenzpunkten, radial, aber mit unterschiedlicher Geschwindigkeit. Dies entspricht einem Voronoidiagramm, basierend auf

einer Distanzfunktion, welche zusätzlich zum Abstandsmaß noch das Gewicht des Referenzpunktes multiplikativ verknüpft.

Definition 16 Die *multiplikativ gewichtete Distanzfunktion* wird definiert als:

$$d_{mw}(x, x_i) = \frac{1}{\omega_i} \|x - x_i\|, \omega_i > 0. \quad (4.12)$$

Das multiplikativ gewichtete Voronoidiagramm (siehe Abbildung 4.10) hat fünf wich-

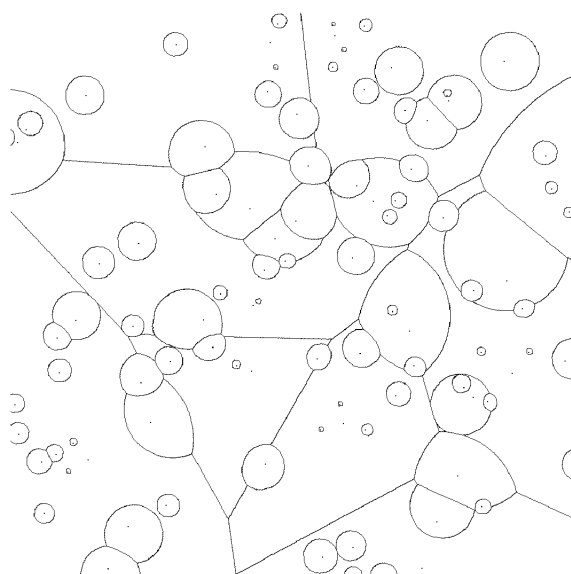


Abbildung 4.10: Multiplikativ gewichtetes euklidisches Voronoidiagramm

tige Eigenschaften:

1. Eine multiplikativ gewichtete Voronoiregion ist nicht leer und nicht notwendigerweise konvex oder zusammenhängend. Dadurch können Löcher im Voronoidiagramm entstehen, das heißt Teilregionen, die von ihrer Referenzregion topologisch getrennt sind.
2. Eine Voronoiregion ist genau dann konvex, wenn alle Gewichte der umliegenden Regionen größer oder gleich dem aktuellen Gewicht sind.
3. Eine Voronoiregion ist unbegrenzt, wenn ihr Gewicht gleich dem Maximalgewicht aller Referenzpunkte ist.
4. Zwei Voronoiregionen können sich unverbundene Kanten teilen.
5. Der Bisektor beschreibt

- (a) einen Kreisbogen genau dann, wenn unterschiedliche Gewichte an den entsprechenden Referenzpunkten vorliegen und
- (b) eine Halbgerade, wenn die beiden Gewichte gleich sind.

Durch die Eigenschaften dieser Art der Gewichtung wird deutlich, dass multiplikative Gewichtung nicht immer topologieerhaltend ist. Dennoch stellt diese Gewichtung eine fundamentale Form dar, welche verschiedenste Prozesse sehr gut simuliert [Sch04b], [HRSS05], [HSS⁺05].

Zusammengesetzte Gewichtung

Eine Verschmelzung der additiven und multiplikativen Gewichtung bringt eine Verallgemeinerung dieser Verfahren.

Die aus additiver und multiplikativer Distanzfunktion zusammengesetzte, gewichtete Distanzfunktion wird nun folgendermaßen definiert:

Definition 17 Die **zusammengesetzt gewichtete Distanzfunktion** wird definiert als:

$$d_{cw}(x, x_i) = \frac{1}{\omega_{i1}} \|x - x_i\| - \omega_{i2}, \omega_i > 0. \quad (4.13)$$

Werden alle Gewichte w_{i1} auf 1 gesetzt, so resultiert gerade das additive Voronoidiagramm und beim Setzen der Gewichte w_{i2} auf 0 gerade das multiplikative. Somit gelten im zusammengesetzt gewichteten Voronoidiagramm alle Eigenschaften aus additiver und multiplikativer Gewichtung.

Zusätzlich hat das zusammengesetzt gewichtete Voronoidiagramm die Eigenschaft, dass die Bisektoren Teile einer jeweiligen polynomialen Kurve vierten Grades sind. Durch den Einfluss des multiplikativen Gewichts kann auch bei dieser Gewichtung

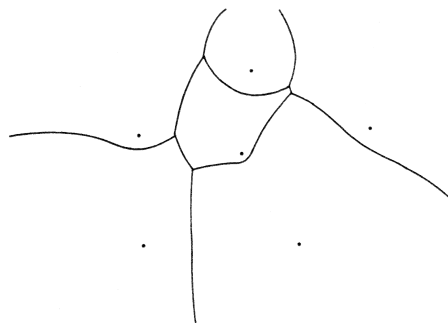


Abbildung 4.11: Zusammengesetzt gewichtetes Voronoidiagramm

der topologische Zusammenhang verloren gehen.

In Abhängigkeit von dem konkreten Einsatz ist also wichtig zwischen der Einbindung

der Gewichte zu unterscheiden. Sollte das berechnete Voronoidiagramm ein zusammenhängendes Gebiet ergeben, darf die Gewichtung nicht multiplikativ verknüpft werden. Generell kann bei einer Integration verschiedener Akteure über Gewichte die Konvexität der entstehenden Unterteilung nicht garantiert werden.

4.2.3 Genereller Ansatz

Im folgenden wird auf die Möglichkeiten einer Erweiterung der bisher beschriebenen Ansätze durch eine freie Gewichtung eingegangen. Auf diese Weise kann die Gewichtung von Distanzfunktionen sehr allgemein vorgenommen werden, wodurch ein Verfahren entsteht, das losgelöst von einer speziellen Anwendung ist.

Die **frei gewichteten Distanzfunktionen** seien definiert durch

$$d_{fw}(p, p_i) = f(\|p - p_i\|, w_i), \quad w_i = \{w_{i1}, \dots, w_{im}\} \quad (4.14)$$

Dabei ist f eine frei wählbare Funktion abhängig von einer Metrik bezüglich p und p_i und einem Gewichtsvektor w_i .

Die Praxis zeigt, dass dieser Gewichtsvektor oft hochdimensional werden kann, denn bei der Konversion sind viele, manchmal abhängige, manchmal unabhängige Parameter zu betrachten (siehe Kapitel 3.2).

Die geschickte Wahl der Funktion f ist bei der freien Gewichtung die größte Herausforderung. Ist eine geeignete Funktion bestimmt, so kann sie in ihre Parameterform überführt werden. Diese liefert schließlich die parametrisierte Kurve, die als Konturlinie für die Generatorgebiete dient [HSRS06].

Als Beispiel einer nicht-konvexen Distanzfunktion kann die Konturlinie eines Kar-



Abbildung 4.12: Konturlinie eines Kardioiden

dioiden (siehe Abbildung 4.12) gewählt werden. Dabei ist folgende Parametrisierung gegeben:

$$f(t) = (x(t), y(t)) = (2 \cos(t) + \cos(2t), 2 \sin(t) + \sin(2t)) \quad t \in [0, 2\pi] \quad (4.15)$$

Dies ist ein sehr universeller Ansatz, der es ermöglicht, eine beliebige geschlossene Kurve als Distanzfunktion zu interpretieren.

Für die praktische Umsetzung dieser Ansätze eignet sich eine Jordan Kurve ohne Verschneidungen. Die Modellierung der Kurve mit B-Splines oder Non-Uniform Rational B-Splines (NURBS) liefert eine äußerst effiziente Möglichkeit die Kurve zu erzeugen. Sind das Kontrollpolygon und die Gewichte gegeben, so lässt sich die Kurve mit dem bewährten De-Boor Algorithmus [dB72], [Far88] berechnen. Der Vorteil dieser Methode ist die Möglichkeit die Distanzfunktion interaktiv zu verändern, so dass die Resultate direkt im Voronoidiagramm sichtbar sind [HSS⁺05]. Abbildung 4.13 zeigt die Möglichkeit zur interaktiven Modifikation einer beispielhaften Distanzfunktion, die mit dieser Methode erzeugt wurde.

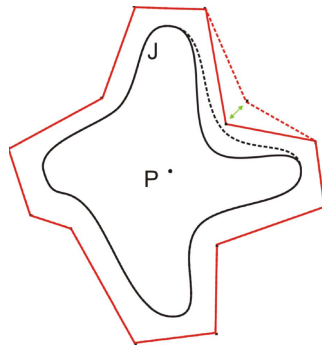


Abbildung 4.13: Interaktive Modifikation einer Distanzfunktion

Im Gegensatz zu der *freien Variation der Distanzfunktion* ist mit diesem Ansatz auch eine *beliebige Gewichtung der Distanzfunktion* möglich. Die Grundidee ist die Verwendung einer NURBS-Kurve höherer Dimension. Für ein zweidimensionales Voronoidiagramm in der xy -Ebene wird eine Kurve in der xz -Ebene konstruiert. Wird die Kurve nun um die z -Achse gedreht liefert dieses Verfahren einen Rotationskörper, der den Voronoi-punkt enthält. Die Projektion aller Rotationskörper in die xy -Ebene liefert die Basis für das beliebig gewichtete Voronoidiagramm. Die einzelnen Gewichte werden durch Funktionen beschrieben die mit den NURBS-Kurven modelliert werden (siehe hierzu auch Abbildung 4.14).

Hoff [HCK⁺99] beschreibt die Erzeugung des euklidischen Voronoidiagrammes durch die Projektion von Kegeln. Additive und multiplikative Gewichte können nun jeweils durch die Position (Höhe) und die Winkel der Kegel kontrolliert werden, die jeweilige Distanzfunktion entsprechend durch das Kontrollpolygon [HSS⁺05].

In Abbildung 4.15 oben ist eine modellierte Distanzfunktion sowie der Fall zu sehen, wenn alle Voronoi-punkte dieselbe Distanzfunktion verwenden. Die obere Zeile zeigt den ungewichteten Fall, in der Mitte sind jeweils ein additiv und ein multiplikativ

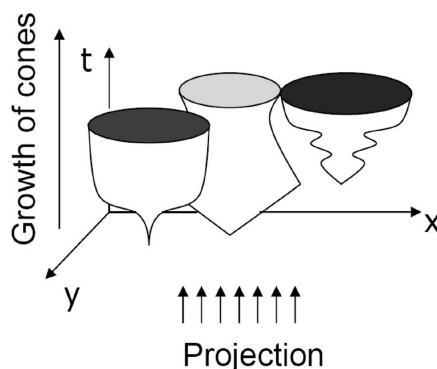


Abbildung 4.14: Generation eines beliebig gewichteten Voronoidiagrammes

gewichtetes Voronoidiagramm abgebildet. Unten links ist ein zusammengesetzt gewichtetes Voronoidiagramm zu sehen, bei dem keine unzusammenhängenden Gebiete entstehen. Unten links ist ein multiplikativ gewichtetes Voronoidiagramm abgebildet, bei dem ein topologisch getrenntes Gebiet hervorgehoben ist.

Eine *generelle Distanzfunktion* wird folgendermaßen aufgebaut: Schneidet der ausgehende Strahl einen nicht-konvexen Teil von C , liefert die Distanzfunktion d_C das gleiche Resultat für unterschiedliche Punkte q_i .

Beispielsweise ist in Abbildung 4.16

$$d_c(p, q_1) = d_c(p, q_2) = d_c(p, q_3). \quad (4.16)$$

Das bedeutet, dass die wichtige Eigenschaft der Dreiecksungleichung der Metrik (siehe Definition 1) nicht eingehalten wird. Für den Einsatzbereich der Konversion ist es jedoch nur von Bedeutung, dass d_c die positive Definitheit erfüllt. Dies wird direkt durch die Definition von d_C garantiert [HSRS06].

Aussagen darüber, wann das Voronoidiagramm den Zusammenhang verliert sind besonders wichtig in dem Anwendungsfeld der Konversion. Eine interessante Aussage betrifft die Konvexität der Konturlinie. Wenn die Konturlinie nicht konvex ist, wie das beispielsweise bei der Verwendung einer generellen Distanzfunktion (siehe Abbildung 4.16) sein kann, ist der Zusammenhang der Voronoiregionen nicht garantiert!

Satz 4 *Wenn die Einheitssphäre C nicht konvex ist, dann sind die Voronoiregionen nicht notwendigerweise zusammenhängend.*

Beweis: siehe [CD85].

Eine Aussage zu der Konvexität der verwendeten Konturlinie lässt sich über die Dreiecksungleichung machen:

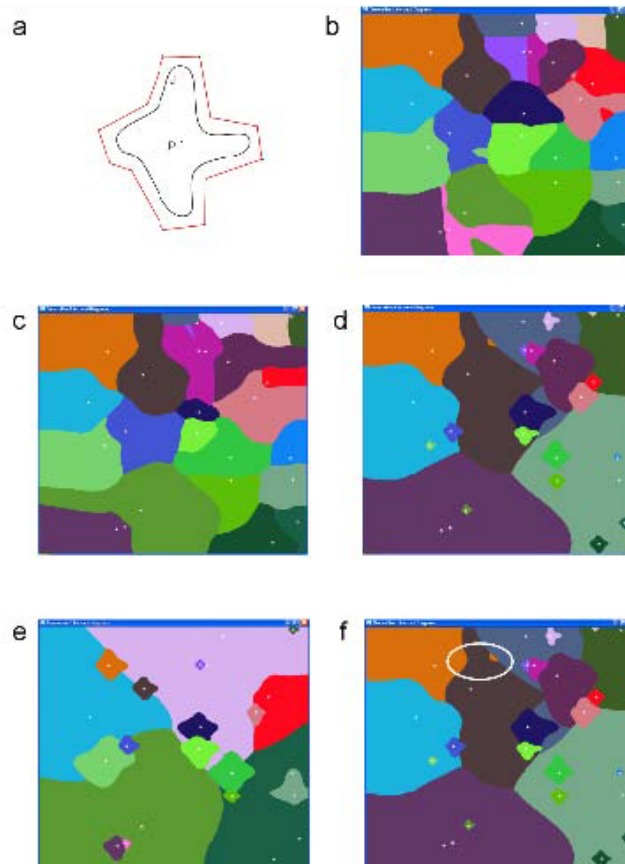


Abbildung 4.15: Schnelle Modellierung eines Voronoidiagrammes mit einer kontrollierbaren Distanzfunktion(a), Resultate für ungewichtete Fälle (b), additiv (c), multiplikativ (d+f) und zusammengesetzt (e) gewichtete Fälle. Eine topologisch unzusammenhängendes Gebiet ist in (f) hervorgehoben

Satz 5 Für eine gegebene Einheitskugel C und für eine Distanzfunktion d , die durch C determiniert wird, wird die Dreiecksungleichung genau dann erfüllt, wenn C konvex ist.

Beweis: siehe [Cas59].

Die konvexe Distanzfunktion ist symmetrisch, sofern C symmetrisch zum Ursprung ist. Die mit einer asymmetrischen Distanzfunktion berechneten Voronoiregionen weisen allerdings die gleichen Eigenschaften wie bei einer Metrik auf. Lediglich die Schwerpunkte der Voronoiregionen sind entsprechend der asymmetrischen Form der Distanzfunktion verschoben.

Zusätzlich erfüllt d_C folgende Eigenschaften:

- C ist der Einheitskreis von d_C , d.h. $\{x | d_C(x, 0) = 1\}$

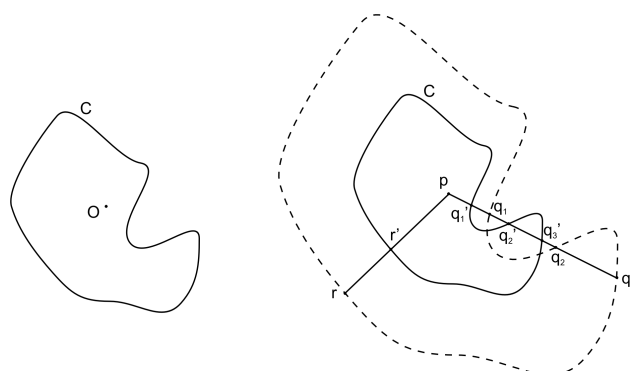


Abbildung 4.16: Konstruktion einer generellen Distanzfunktion

- d_C ist translationsinvariant

4.3 Generischer Wachstumsalgorithmus

Der große Vorteil bei der Verwendung eines Wachstumsalgorithmus für die Umsetzung des Clusterings ist der hohe Grad an Allgemeinheit, den dieser Algorithmus mit sich bringt. Durch die Möglichkeit, jede geschlossene, parametrisierte Kurve als Distanzfunktion zu verwenden, wird er sehr flexibel und damit besonders interessant für den Einsatz. Er ist somit ein zukunftsorientierter Algorithmus, in den sich auch bisher noch nicht entwickelte Abstandsbegriffe einbinden lassen. Die zusätzlich Möglichkeit, Gewichtungen einzubinden, verstärkt die Flexibilität des Algorithmus entscheidend.

Ein wichtiges Kriterium für den Einsatz von Distanzfunktionen bei einem Wachstumsalgorithmus ist die Topologieerhaltung. Das heißt, dass bei der Verwendung bestimmter Distanzfunktionen im vorgestellten Wachstumsalgorithmus keine nicht-zusammenhängenden Gebiete (Löcher) resultieren. Es ist dabei generell zu unterscheiden, ob Löcher aufgrund des gewählten Abstandbegriffes oder aber einer Gewichtung entstehen. Werden beim Wachstumsalgorithmus Gewichtungen mit multiplikativen Anteilen verwendet, so kann grundsätzlich nicht garantiert werden, dass die Topologie erhalten bleibt.

In Kapitel 4.3.1 wird ein kurzer Überblick über herkömmliche Algorithmen gegeben. Kapitel 4.3.2 beschreibt anschließend die Umsetzung des Wachstumsalgorithmus für ein gewichtetes Wachstum und in Kapitel 4.3.3 wird eine Erweiterung durch ein richtungsabhängiges Wachstum vorgestellt.

4.3.1 Herkömmliche Verfahren

Herkömmliche Algorithmen zur Berechnung von Voronoidiagrammen, wie sie z.B. bei Klein [LWW90] zu finden sind, dienen in erster Linie einer effizienten Berechnung von Voronoidiagrammen in der Ebene. Bei der inkrementellen Konstruktion wird das Voronoidiagramm durch die sukzessive Hinzunahme der Referenzpunkte inkrementell berechnet. Dies ist eine sehr nahe liegende Idee, auf der die ältesten Konstruktionsverfahren beruhen. Hierbei wird das Voronoidiagramm schrittweise aufgebaut, beginnend mit dem ersten Punkt aus der Punktmenge. Beim Einfügen eines neuen Punktes aus der Punktmenge wird dann die Struktur des Voronoidiagramms wieder hergestellt. Es kann dabei ausgenutzt werden, dass sich die Struktur grundsätzlich nur lokal ändert, also nicht das gesamte Voronoidiagramm. Dies wird wiederholt, bis alle Punkte im Voronoidiagramm eingefügt sind. Dieses Verfahren eignet sich insbesondere für Situationen, bei denen vermehrt Punkte entfernt und wieder eingefügt werden müssen, da speziell das Einfügen von Punkten ein charakteristischer Bestandteil des Algorithmus ist. Die Laufzeit des Algorithmus liegt in $O(n^2)$.

Sweepverfahren hingegen „fegen“ den gegebenen Raum mit einem Sweep-Hyperraum aus und berechnen sukzessive das Ergebnis. Für die Berechnung des Voronoidiagramms in der Ebene bedeutet dies, dass sich eine Sweepgerade von links nach rechts durch die Ebene bewegt. Bei jedem Berühren eines Punktes aus der Punktmenge wird dieser in das bis dahin errechnete Voronoidiagramm eingefügt. Dieses wird so von links nach rechts vervollständigt. Eine detaillierte Beschreibung des Verfahrens ist bei [Wad04b] zu finden. Ein nachträgliches Einfügen von beliebigen Punkten in das Voronoidiagramm führt zu einer vollständigen Neuberechnung - ein wesentlicher Unterschied zur inkrementellen Konstruktion. Die Laufzeit des Sweepverfahrens liegt in $O(n \log n)$.

Ein häufig praktiziertes Verfahren zur Lösung komplexer Probleme ist das „Divide and Conquer“ Verfahren. Zuerst wird das Problem solange in kleinere Teilprobleme zerlegt bis die entstandenen Teilprobleme leicht lösbar sind. Zum Einsatz in der Ebene eignen sich hierbei Quadrees, also ein Teilen der Ebene in vier ungefähr gleich große Bereiche. Der Algorithmus unterteilt die Bereiche wiederum solange rekursiv, bis nur noch zwei Referenzpunkte im Gebiet vorliegen. Die Lösungen der Teilprobleme sind sehr schnell und einfach zu berechnen und werden dann schrittweise zur Gesamtlösung zusammengesetzt. Die Laufzeit liegt wie schon beim Sweepverfahren in $O(n \log n)$.

Ein sehr interessantes Verfahren ist die geometrische Transformation. Hier wird ein anderes aber ähnliches geometrisches Problem betrachtet, welches einfach oder zumindest einfacher zu berechnen ist. Die Lösung wird dann in die des eigentlichen Problems transformiert. Das Voronoidiagramm kann auf diese Art durch die Berechnung der unteren konvexen Hülle einer Punktmenge im Raum berechnet werden

[ES86]. Die Laufzeit richtet sich nach der Laufzeit des Verfahrens zur Berechnung der konvexen Hülle und kann in $O(n \log n)$ liegen. Die Transformation lässt sich dann in Linearzeit $O(n)$ durchführen. Das heißt, dass durch diese geschickte Transformation der Aufwand erheblich reduziert werden kann, so dass dieses Verfahren insgesamt in $O(n \log n)$ liegt.

Neben den herkömmlichen Verfahren werden nun neue Verfahren zur Verallgemeinerung von Voronoidiagrammen vorgestellt.

Die bisher betrachteten Verfahren setzen meistens die Euklidische Metrik voraus und lassen sich in Einzelfällen auf die Minkowski-Metriken verallgemeinern. Der Grundgedanke für noch allgemeinere Verfahren ist die Verwendung der in Kapitel 4.2.1 vorgestellten allgemeinen konvexen Distanzfunktionen. Zur Realisierung dieser Verallgemeinerung wird in diesem Kapitel ein Wachstumsalgorithmus vorgestellt. Die notwendigen Distanzfunktionen können beispielsweise durch geschlossene parametrisierte Kurven, welche Einheitsgebiete bilden, dargestellt werden. Der jeweilige Referenzpunkt liegt dabei stets im Inneren eines solchen Gebietes. Der vorgestellte Wachstumsalgorithmus kann dann auf diesen Gebieten operieren.

4.3.2 Gewichtetes Wachstum

Die Konturlinien der gewählten Distanzfunktion starten beim Wachstum mit einer Distanz von Null und schneiden sich nach einer endlichen Anzahl von Iterationen. Die so gebildeten Schnittpunkte haben alle den gleichen Abstand von den Referenzpunkten und formen in ihrer Gesamtheit das Voronoidiagramm. Ein Beispiel hierzu ist in Abbildung 4.17 zu sehen. Bei den bekannten Verfahren zur Konstruktion von

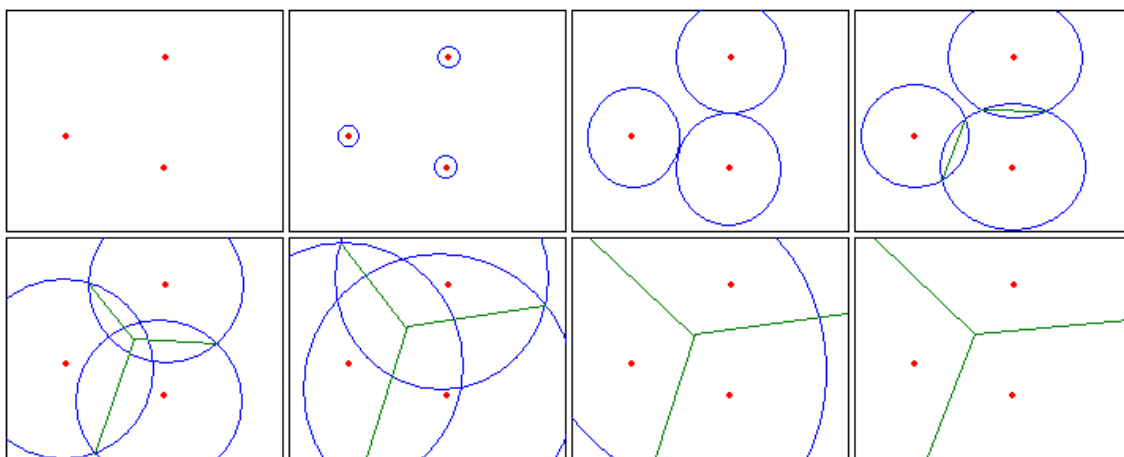


Abbildung 4.17: Schrittweises Wachstum.

Voronoidiagrammen liegt die Priorität auf der Geschwindigkeit zur Erstellung des

Diagramms. Dabei werden immer einfache mathematische Metriken, welchen keinerlei Gewichtung unterliegt, benutzt. Diese Verfahren scheitern aber meistens aufgrund ihrer Konstruktion bei Verwendung von komplexeren Distanzfunktionen. Die Flexibilität dieses Wachstumsalgorithmus bildet die grundlegende Eigenschaft für deren Einsatz in der Konversion. Alle folgenden Konzepte werden durch Varianten dieses Wachstumskonzeptes gebildet.

Additiv gewichtetes Wachstum

Wachsen die Generatorgebiete der Referenzpunkte mit gleicher Wachstumsrate aber unterschiedlichen Startzeiten, dann resultieren daraus die additiv gewichteten Voronoidiagramme, wobei die Startzeiten gerade den Gewichten w_i entsprechen (siehe

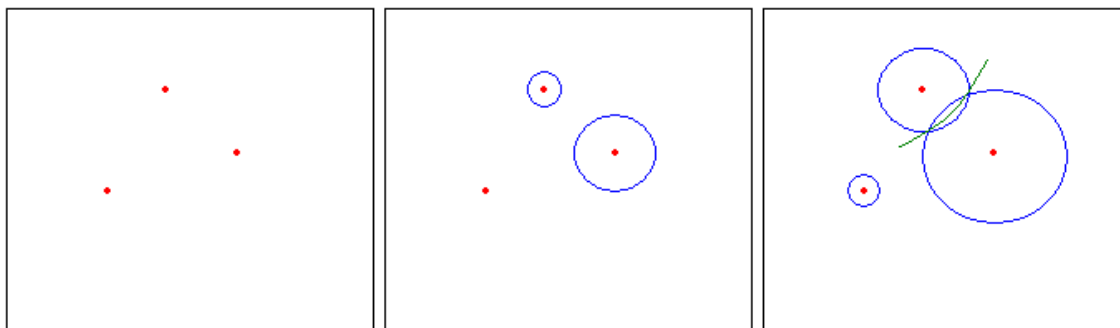


Abbildung 4.18: Additives Wachstum zu den Zeitpunkten t_0, t_5 und t_{10} .

Abbildung 4.18).

Okabe et al. [OBSC99] erwähnen, dass die Johnson-Mehl Tessellation ein Spezialfall der additiv gewichteten Voronoidiagramme ist. Dabei entsprechen die w_i gerade den Zeiten t_i , bei denen die Referenzpunkte erstmals auftauchen.

Multiplikativ gewichtetes Wachstum

Multiplikative Voronoidiagramme (siehe Abbildung 4.19) werden erzeugt, indem die Generatoren zwar zur selben Zeit, aber mit unterschiedlichen Wachstumsgeschwindigkeiten, entsprechend den w_i , gestartet werden.

Jeder Referenzpunkt besitzt ein Gewicht; wird nun der Bisektorpunkt auf der Strecke von p_i und p_j bestimmt, so teilt dieser Punkt gerade die Strecke im Verhältnis $w_j : w_i$. Die Bisektoren sind somit zusammengesetzt aus Teilkreisbögen.

Satz 6 *Das Wachstum der Generatorgebiete C_i mit der jeweiligen Geschwindigkeit w_i entspricht gerade der Konstruktion von Voronoidiagrammen basierend auf multiplikativ gewichteten Distanzfunktionen.*

Beweis: Ohne Beschränkung der Allgemeinheit wachse der Generator C_i mit Geschwindigkeit x und C_j mit Geschwindigkeit y . Dann bringt eine Gewichtung der

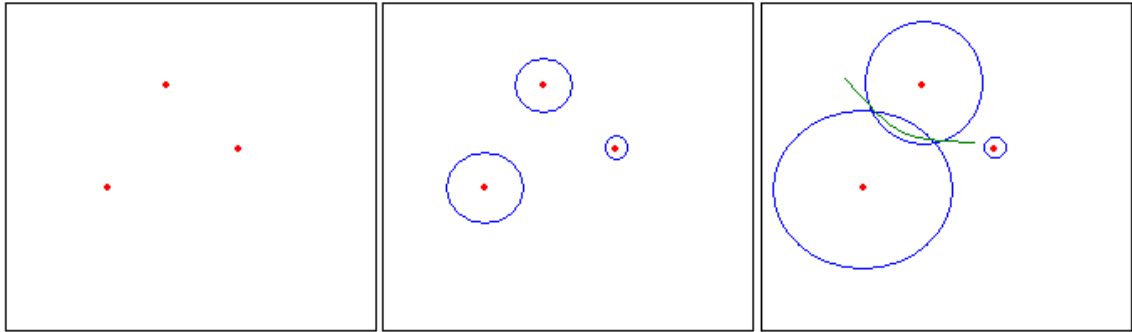


Abbildung 4.19: Multiplikatives Wachstum zu den Zeitpunkten t_0, t_5 und t_{10} .

Distanzfunktion von C_i mit $1/x$ und von C_j mit $1/y$ gerade ein Teilungsverhältnis der Strecke $p_i p_j$ von $y : x$. Sei nun $w_j = y$ und $w_i = x$. Daraus folgt die obige Behauptung [HRSS05], [HSS⁺05].

4.3.3 Richtungsabhängiges Wachstum

Eine mögliche Erweiterung der betrachteten Voronoialgorithmen besteht in der Ausnutzung der beiden orthogonalen Koordinatenachsen im R^2 .

Wird zum Beispiel eine Minkowski-Metrik durch ein Gewicht w_i am Punkt q_i multiplikativ skaliert, so ergibt sich:

$$d(q_i, x) = \frac{1}{w_i} \sqrt[p]{|q_{i_1} - x_1|^p + |q_{i_2} - x_2|^p} \quad (4.17)$$

Diese Gleichung lässt sich umformen zu:

$$d(q_i, x) = \sqrt[p]{\left(\frac{1}{w_i}\right)^p |q_{i_1} - x_1|^p + \left(\frac{1}{w_i}\right)^p |q_{i_2} - x_2|^p} \quad (4.18)$$

Stehen nun in jedem q_i zwei Gewichte w_{i_1} und w_{i_2} zur Verfügung, so lässt sich durch obige Formel leicht eine richtungsabhängige Distanzfunktion d_{ds} definieren:

$$d_{ds}(q_i, x) = \sqrt[p]{\left(\frac{1}{w_{i_1}}\right)^p |q_{i_1} - x_1|^p + \left(\frac{1}{w_{i_2}}\right)^p |q_{i_2} - x_2|^p} \quad (4.19)$$

Dadurch können nun zwei Datenaspekte gleichzeitig durch Voronoidiagramme dargestellt werden. Diese Datenaspekte werden durch die beiden Raumrichtungen x und y visualisiert, indem die Ausdehnung in x - und y -Richtung mit den jeweiligen Teilgewichten w_{i_1} und w_{i_2} multiplikativ verknüpft sind. Das heißt, die Gewichtung ist über die Koordinatenachsen kodiert.

Dies kann im Wachstumsalgorithmus durch separates Wachstum in beide Achsenrichtungen mit den jeweiligen Geschwindigkeiten w_{i_1} und w_{i_2} realisiert werden.

Die Idee des richtungsabhängigen Wachstums, welche exemplarisch für Minkowski-Metriken hergeleitet wurde, ist durch die Art der Implementierung nicht mehr von der jeweiligen Metrik abhängig und somit auf alle denkbaren Distanzfunktionen anwendbar. Aus den Betrachtungen geht insgesamt hervor, dass Topologieerhaltung bei konvexen Distanzfunktionen mit einer Gewichtung ohne multiplikative Anteile garantiert werden kann. Bei nicht-konvexen Distanzfunktionen hingegen kann es zu nicht zusammenhängenden Gebieten kommen.

Kapitel 5

Visualisierungsstrategien

In dem letzten Kapitel wurde das Verfahren der Voronoidiagramme generalisiert, so dass es für das Information Clustering eingesetzt werden kann. In diesem Kapitel wird nur der Bereich der Visualisierung behandelt. Es werden zwei grundsätzliche Strategien entwickelt, die sich für die Visualisierung der heterogenen Daten im Bereich der Stadtplanung eignen. Entscheidende Kriterien für die Auswahl der Techniken waren deren Eignung multivariate Daten zu handhaben und diese in einen Bezug zu einer Karte darstellen zu können. Besonders wichtig ist es mit Hilfe dieser Techniken die Ergebnisse interpretieren zu können.

In Kapitel 5.1 wird zunächst auf bestehende Techniken eingegangen, um deren Vorteile zu sondieren, deren Eignung für den gegebenen Einsatzbereich zu testen und konkrete Anforderungen an die für diese Arbeit benötigten Visualisierungsstrategien herzuleiten.

In Kapitel 5.2 wird das *Picklock-Icon* aus dem Bereich der Iconbasierten Visualisierungstechniken vorgestellt und anschließend in Kapitel 5.3 der *Spike-Tree* aus dem Bereich der Botanischen Informationsvisualisierung.

5.1 Anforderungen

Die Techniken der Iconbasierten Informationsvisualisierung werden in die Klasse der explorativen Visualisierungstechniken (siehe Kapitel 2.2.1) eingeordnet, welche vor allem für die Darstellung von diskreten, multivariaten Daten eingesetzt werden und diese in visuelle Texturen verwandeln. Diese Iconbasierten Visualisierungen übertragen nicht sichtbare und unentdeckte statistische Strukturen der Daten in visuelle Strukturen wie Inseln, Streifen oder Gradienten.

Der Mensch ist in der Lage Unterschiede in Texturen sehr effizient zu unterscheiden. Des Weiteren nutzt er Abweichungen in Texturen als Quelle wichtiger Informationen für die Erkennung und Wahrnehmung von Objekten. Daher nimmt der Anwender

die statistischen Eigenschaften der Daten als Eigenschaften der dargestellten Textur sehr gut wahr. Die Iconbasierten Techniken eignen sich somit sehr gut für den beabsichtigten Einsatz in der Stadtplanung.

Durch die Übertragung der Parameter auf die visuellen Strukturen der Icons liefern diese Techniken die benötigte Flexibilität, die zur Exploration der multivariaten, räumlich zusammenhängenden Daten benötigt wird.

Die bisher aus diesem Bereich eingesetzten Techniken haben den Nachteil, dass sie für spezielle Anwendungen entwickelt wurden und damit für den hier vorliegenden Einsatzbereich nicht genügend Flexibilität aufweisen. Beispielsweise sind sie auf maximal sieben Dimensionen beschränkt und nur teilweise dreidimensional ausgerichtet. Ihr Vorteil liegt in der prinzipiellen Erweiterbarkeit der Techniken und ihre grundsätzliche Eignung für den vorliegenden Anwendungsbereich. Das in Kapitel 5.2 entwickelte *Picklock-Icon* nutzt diese Vorteile und kombiniert sie mit einer geeigneten Ausrichtung auf den Bereich der Stadtplanung. Durch die räumlich zusammenhängende Visualisierung des *Picklock-Icons* ist es mit dieser Methode möglich das zugrundeliegende generelle Voronoidiagramm implizit mit zu visualisieren. In Abhängigkeit von dem Mapping der Daten auf die visuellen Strukturen des *Picklock-Icons* ähnelt die Visualisierung einer Graslandschaft (siehe Abbildung 5.4) und spiegelt somit die Tessellierung des berechneten Voronoidiagrammes sehr gut wieder.

Der Vorteil der Iconbasierten Visualisierungstechniken, die Daten räumlich zusammenhängend darstellen zu können, birgt aber auch Nachteile. Durch den räumlichen Zusammenhang überdeckt die Visualisierung eine darunter liegende Karte. In Abhängigkeit davon, welche Strukturen der Daten auf welche visuellen Strukturen des Icons übertragen werden, kann das ein sehr unerwünschter Nebeneffekt sein.

Es ist somit sinnvoll weitere Visualisierungstechniken zu diskutieren, die prinzipiell punktorientiert arbeiten und damit diese Probleme verhindern. Hierarchische Techniken stellen Daten in Form von hierarchisch aufgeteilten Untereinheiten dar. Im Fall von multidimensionalen Datensätzen dienen dabei selektierte Dimensionen zur Aufteilung des Datensatzes und zum Aufbau der Hierarchie.

Es bieten sich Techniken an, die sowohl dreidimensional orientiert sind als auch punktorientiert arbeiten und somit die multivariaten Daten optimal repräsentieren können. Besonders die Techniken der Botanischen Informationsvisualisierung sind in diesem Zusammenhang erfolgversprechend. Der Vorteil der kontextsensitiven Darstellung der Daten in Bezug zu einer Karte machen sie sehr interessant für den Anwendungsbereich der Planung. Bei diesem Verfahren werden bereits geübte sensorische Fähigkeiten der Akteure optimal für die Visualisierung genutzt.

Eine Variante der botanischen Visualisierung ist die Darstellung als dreidimensionaler Baum. Ein enormer Vorteil dieser Variante ist der intuitive Zugang der Anwender zu der Visualisierung. Durch die Ähnlichkeit mit den biologischen Vorbildern ist es

dem Benutzer intuitiv möglich die einzelnen Elemente und ihre Anordnungen visuell extrahieren zu können.

Mögliche Probleme hinsichtlich der Interpretierbarkeit der Darstellung lassen sich durch das Hinzunehmen einzelner Kriterien vermeiden. Beispiele hierfür sind sehr lange, dünne Äste und ein Überlappen der Blätter. Aus symmetrischen Gründen wurden die Früchte an den oberen Ästen des in Kapitel 5.3.1 eingeführten *Spike-Trees* kugelförmig gewählt.

5.2 Picklock-Icon

In diesem Kapitel werden zunächst zwei Iconbasierte Visualisierungstechniken vorgestellt, die die Grundlage für das entwickelte *Picklock-Icon* bilden. Dieses neue Icon unterstützt die Visualisierung der multivariaten Informationen aus dem Bereich der Stadtplanung optimal.

Ausschlaggebend für die Wahl der Icon- und Glyph-basierten Techniken sind deren Vorteile in Bezug auf die Darstellung der Informationen im Kontext zu einer Karte. Dazu wird zunächst auf das Stick-Figure-Icon von Ronald M. Picket und Georges Grinstein [RP88] und anschließend das Velcro-Icon von Rob Erbacher und Georges Grinstein [EG95] eingegangen.

5.2.1 Stick-Figure- und Velcro Icon

Das von Picket und Grinstein entwickelte Stick-Figure-Icon besteht aus fünf miteinander verbundenen Liniensegmenten. Dabei ist eines der Liniensegmente der Rumpf des Icons und die restlichen vier sind die Glieder. Die gesamte Icon-Familie (siehe Abbildung 5.1) besteht aus zwölf Mitgliedern, von denen jedes durch seine Art und Weise, wie es mit dem Rumpf oder den anderen Gliedern verbunden ist, definiert wird. Dabei eignen sich die einzelnen Mitglieder der Familie für unterschiedlichste Datensätze - insbesondere aber für räumlich zusammenhängende Daten.

Die zu untersuchenden Daten werden auf das Icon abgebildet, indem sie den einzelnen Gliedern zugeordnet und dessen Winkel zum Rumpf kontrolliert werden. Dadurch können mit jedem Mitglied der Familie Daten bis zu vier Dimensionen dargestellt werden - und durch die Orientierung des Rumpfes sogar noch eine fünfte Dimension.

Abbildung 5.2 zeigt ein integriertes Bild von fünf Datenkanälen des Wettersatelliten NOAA-7 AVHRR, das das westliche Ende des Lake Ontario und die östliche Spitze des Lake Erie in Nord-Amerika darstellt. Die untere Mitte des Bildes liegt ungefähr über den Niagarafällen.

Bei der Berechnung des Bildes wurde das Mitglied 12 der Stick-Figure Familie (siehe Abbildung 5.1) verwendet. Die Abbildung der Daten auf den Rumpf und die Ge-

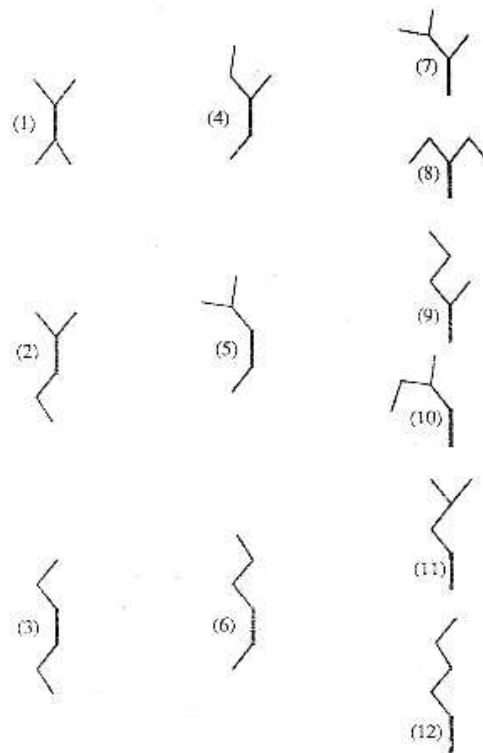


Abbildung 5.1: Stick-Figure-Icon-Familie

gestaltung der einzelnen Glieder wurden so gewählt, dass einer der visuellen Kanäle die Orientierung des Rumpfes kontrolliert und die restlichen visuellen Kanäle die beiden Glieder, die am nächsten zum Rumpf liegen. Die beiden IR-Kanäle kontrollieren letztlich die beiden äußersten Glieder.

Insgesamt entsteht trotz fehlender Grauwert-Kodierung ein auffallend gut strukturiertes Bild. Dieses Bild demonstriert sehr schön, wie stark sich die einzelnen Regionen in der Anordnung der Icons auf Basis von Textur-Unterschieden differenzieren lassen.

In gegenwärtigen Anwendungen wird typischerweise nach Regionen gesucht, die aus den Daten nicht direkt ersichtlich sind.

Das von Erbacher und Grinstein entwickelte Velcro-Icon (Abbildung 5.3) besteht aus einer geraden Basis und einer gekrümmten oder hakenförmigen Spitze. Während die Icons selbst dreidimensional sind, werden sie in einem zweidimensionalen Gitter platziert. Dabei werden Orientierung, Form und Größe des resultierenden Icons von den datenabhängigen Parametern Basislänge, Krümmung, Azimuth, Höhe, Rotation der Spitze, Basisintensität und Intensität der Spitze kontrolliert. Die Länge der Spitze ist dabei für alle Icons gleich.

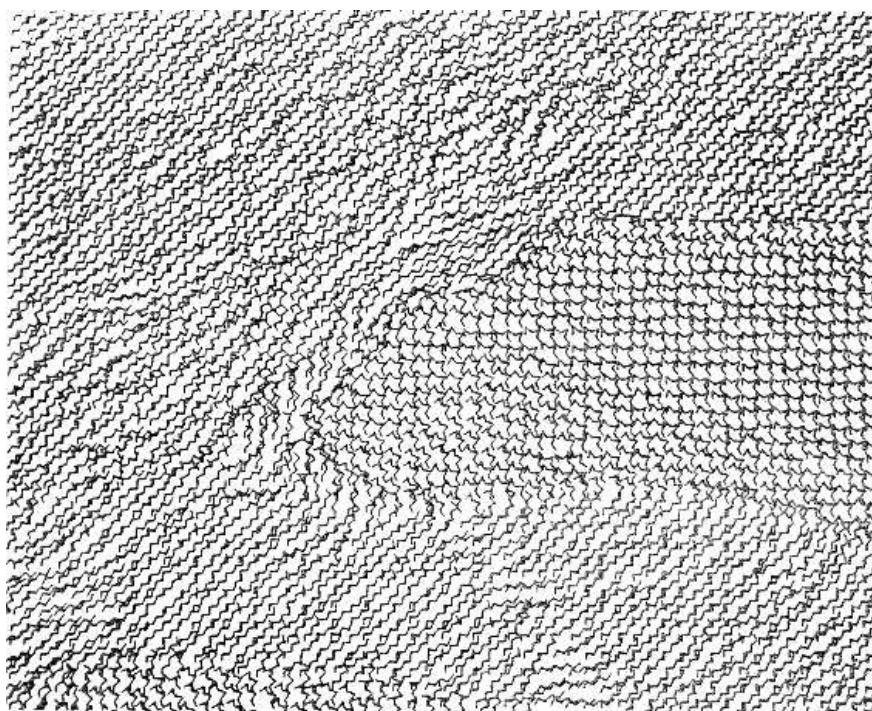


Abbildung 5.2: Visualisierung multispektraler Bilddaten des NOAA-7 AVHRR Wettersatelliten unter Verwendung des Stick-Figure-Icons

Diese und weitere, durch den Benutzer kontrollierbare Parameter, liefern jene Flexibilität, die zur Exploration multivariater Daten benötigt wird. Abbildung 5.4 zeigt eine Sequenz von vier Satellitenbildern der Great Lakes Region. Der Datensatz besteht aus zwei Bildern im sichtbaren Spektrum und zwei Bildern im Infrarotspektrum. Die Darstellung wurde dabei um zwei Achsen rotiert, wodurch die Regionen der Great Lakes durch dreidimensionale Effekte klar zu erkennen sind. Seen führen dabei zu konvexen und Gebirgsregionen zu niedrigen Merkmalen. Durch eine Rotation um 180 Grad können die Seen auch konkav werden.

5.2.2 Konzept

Mit dem in Kapitel 5.2.1 beschriebenen Stick-Figure-Icon lassen sich lediglich bis zu fünf Informationen darstellen, das Velcro-Icon eignet sich für maximal sieben Informationen. Die Icons selbst sind zweidimensional und werden zur Generierung einer ebenfalls zweidimensionalen Textur verwendet. Die datenabhängigen Parameter kontrollieren dabei die Winkel der vier Glieder zum Rumpf und die Orientierung des Rumpfes.

Keines der beiden Icons eignet sich zur Visualisierung der gegebenen hochdimensio-

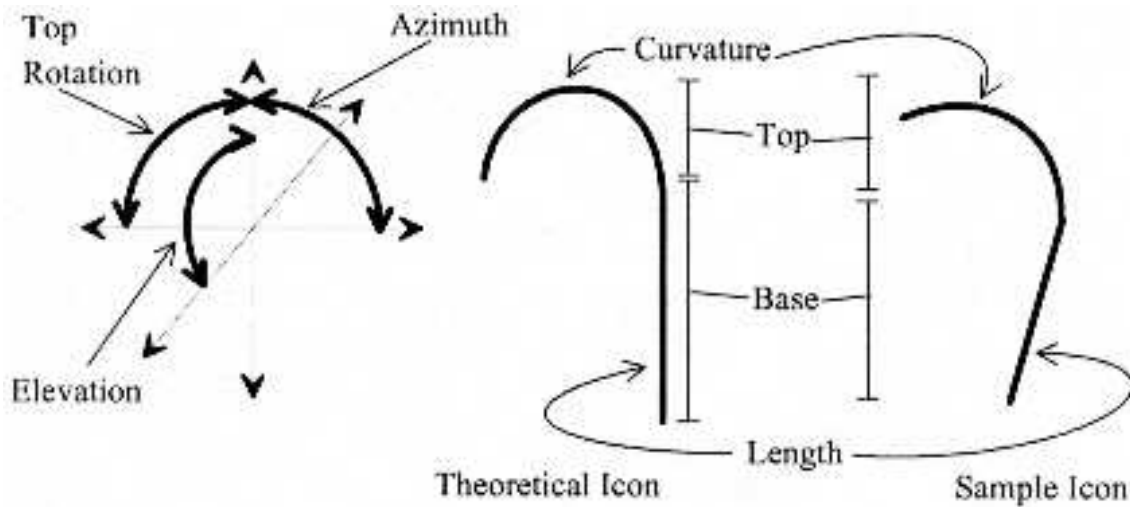


Abbildung 5.3: Eigenschaften des Velcro-Icons

nenalen Informationen im Bereich der Stadtplanung, speziell bei der Konversion von Flächen. Es wird ein wesentlich flexibleres und aussagekräftigeres Icon benötigt, um die wachsende Anzahl an Informationen sinnvoll zu bewältigen. Das neu entwickelte *Picklock-Icon* (siehe Abbildung 5.5) hat diese Eigenschaften. Es ist flexibel genug, um eine wachsende Anzahl von Informationen darzustellen.

Das *Picklock-Icon* ist verwandt mit dem Mitglied 12 der Stick-Figure-Familie (siehe Abbildung 5.1). Für die Erhöhung der Flexibilität wurde es stark modifiziert und den gegebenen Bedürfnissen angepasst. Dabei wurde zunächst das vierte Glied entfernt und anschließend wurde es entsprechend dem Velcro-Icon in die dritte Dimension erhoben, um so weitere Freiheitsgrade zu erlangen.

Das *Picklock-Icon* besitzt neben dem Rumpf drei Glieder. Die Länge des Rumpfes und der Glieder lassen sich durch datenabhängige Parameter kontrollieren. Zusätzlich können der Rumpf und jedes Glied durch jeweils zwei Parameter um die y-Achse und die z-Achse rotiert werden. So lassen sich bis zu zwölf Informationen visualisieren. Mit dem Hinzufügen weiterer Glieder am oberen Ende des Icons können mit jedem neuen Glied weitere drei Informationen visualisiert werden.

Die maximale bzw. maximal sinnvolle Anzahl der Glieder hängt dabei stark von den zu visualisierenden Daten ab. Ebenso hängen auch die Gesamtgröße der Icons, die Länge des Rumpfes und der Glieder sowie die Rotationswinkel um die y- und z-Achse von den Daten ab.

Um die gegebenen Informationen mit dem *Picklock-Icon* zu visualisieren, werden zunächst die einzelnen Informationen auf die Parameter abgebildet, die die Länge und Winkel des Rumpfes und der Glieder kontrollieren. Anschließend werden die

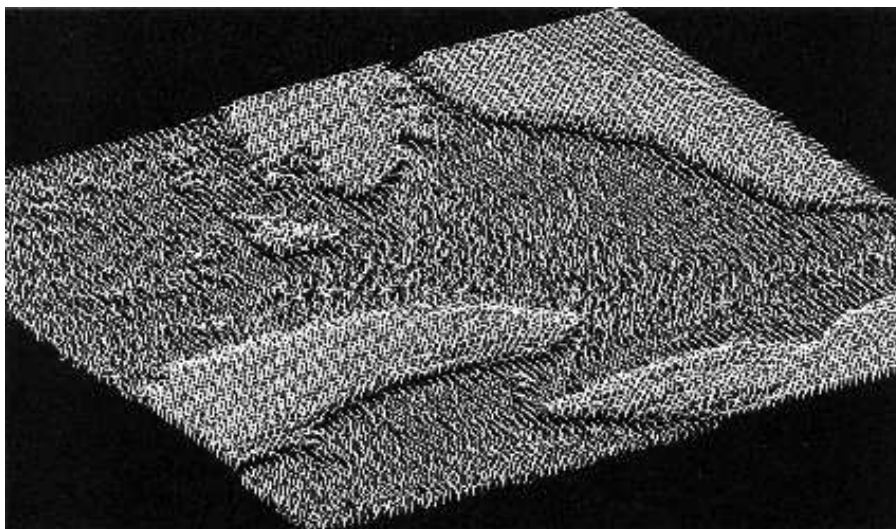


Abbildung 5.4: Visualisierung eines Satellitenbildes unter Verwendung des Velcro-Icons [EG95]

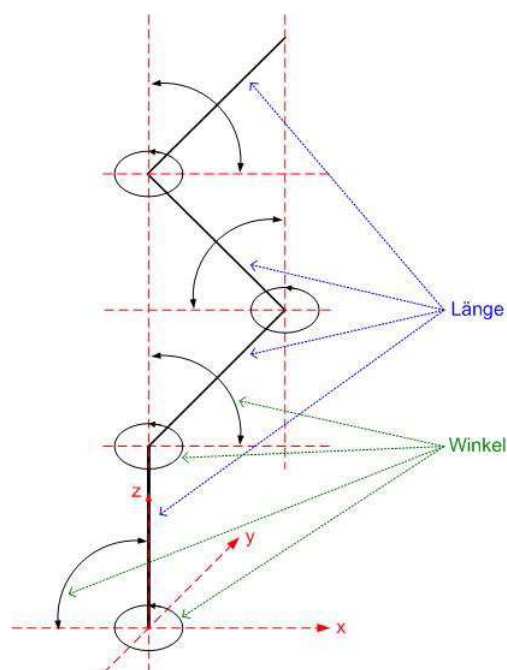
Picklock-Icons anhand eines 2D-Rasters in der Ebene positioniert. Die Auflösung des 2D-Rasters entspricht dabei gerade der Auflösung der Textur, die das gewichtete Voronoidiagramm enthält.

5.2.3 Umsetzung

In Abbildung 5.6 ist der Ausschnitt einer topographischen Karte (TK 25) der Stadt Koblenz zu sehen. Sie ist einer historischen Sammlung von TK-25 Ausgaben des „Landesamt für Vermessung und Geobasisinformation Rheinland-Pfalz, Koblenz“ [LVe04] entnommen und mit einem Bildbearbeitungsprogramm überarbeitet worden. Der abgebildete Kartenausschnitt ist eindeutig durch die Universal Transverse Meractor- bzw. Gauß-Krüger-Koordinaten der Eckpunkte definiert.

Die zu den Konversionsflächen gehörenden Daten enthalten Informationen über die Lage, die Größe und den Nutzungsgrad der Flächen. Zusätzlich gibt es drei unterschiedliche Sichtweisen auf die Daten bzw. drei Interessengruppen, die die Daten unterschiedlich stark gewichten: das Land, der Verkäufer und der Investor.

Zusätzlich zu diesen Daten kommen verschiedene Parameter von Seiten der Akteure hinzu. Dieser ist beispielsweise auf der Suche nach einer Fläche mit einer bestimmten Größe und einem bestimmten Nutzungsgrad. Für erste Testzwecke wurden die Parameter auf diese Informationen begrenzt. Zusammen ergeben sich zunächst neun verschiedene Informationen, die mit dem *Picklock-Icon* visualisiert werden sollen. Die Gewichte der einzelnen Voronoiregionen ergeben sich dabei anhand der unter-

Abbildung 5.5: Das *Picklock-Icon*

schiedlichen Sichtweisen (siehe Abbildung 5.7) auf die Daten.

Bei der Berechnung der Visualisierung wird mit der Darstellung der Karte begonnen, gefolgt von dem gewichteten Voronoidiagramm. Sie endet mit der Berechnung des *Picklock-Icons*. Wie in Abbildung 5.8 zu sehen ist, wird das Voronoidiagramm über die Karte und die *Picklock-Icons* wiederum über das Voronoidiagramm gelegt. Die *Picklock-Icons* und das Voronoidiagramm sind dabei in der Transparenz variabel, um dem Anwender einen genügend großen Spielraum bei der Betrachtung der Visualisierung zu geben.

Nachdem die Berechnung des Voronoidiagramms und der *Picklock-Icons* gestartet wurde, wird die Karte in einer der Auflösung des Bildes nahe kommenden Textur geladen und im Visualisierungsfenster dargestellt (siehe Abbildung 5.9). Das Voronoidiagramm wird ebenfalls in einer Textur gespeichert, dessen Auflösung in der Höhe und der Breite identisch ist. Die Auflösung entspricht dem Maximum von der Breite und der Höhe der Textur, die die Karte enthält.

Wie aus Abbildung 5.10 ersichtlich ist, setzen sich die Gewichte ω_i der einzelnen Voronoiorte aus den gegebenen Informationen und den drei Sicht- bzw. Interessengruppen zusammen. Zu jedem Voronoiort existieren drei Informationen: Lage, Größe und Nutzungsgrad. Zusätzlich sind für jede Interessengruppe drei Parameter vorhanden, die Lage, Fläche und Nutzungsgrad beeinflussen und in der Summe jeweils eins

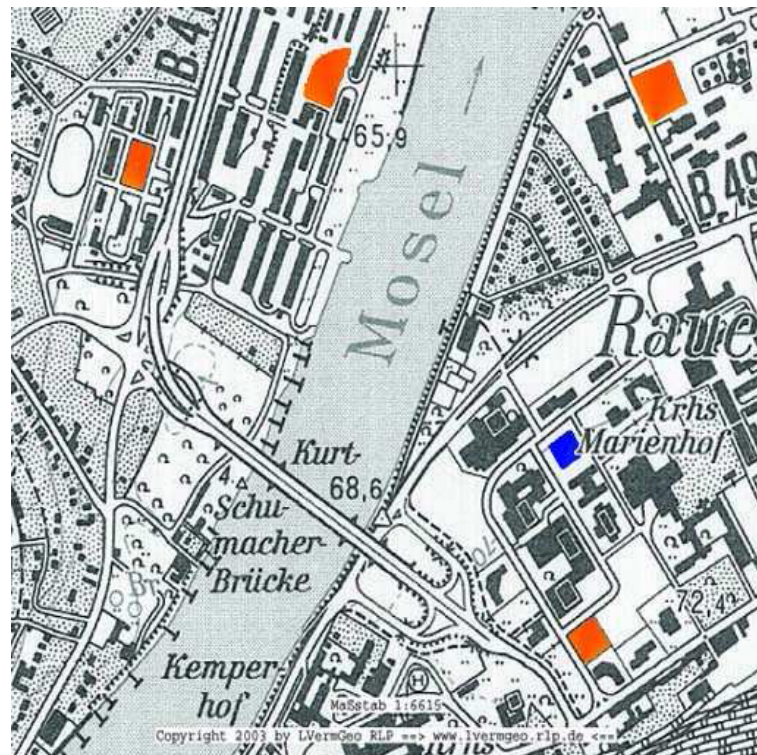


Abbildung 5.6: Ausschnitt einer topographischen Karte der Stadt Koblenz

ergeben. Insgesamt setzen sich die Gewichte also aus neun Parametern zusammen. Die in Abbildung 5.10 markierten Informationen haben aus Sichtweise der jeweiligen Interessengruppe einen großen, mittleren und geringen Einfluss auf die Bestimmung des Gesamtgewichtes.

Für die Berechnung der *Picklock-Icons* werden die gegebenen Informationen auf die Kontrollparameter der einzelnen Glieder des Icons abgebildet. Dazu werden analog zu Abbildung 5.10 die Parameter der Interessengruppen auf die Kontrollparameter des *Picklock-Icons* abgebildet.

Dabei kontrollieren Parameter mit mittlerem oder geringem Einfluss die Länge bzw. den Rotationswinkel um die z-Achse des Rumpfes oder des Gliedes. Parameter mit großem Einfluss kontrollieren die Rotation um die y-Achse, da dies eine größere visuelle Änderung bewirkt.

Des weiteren besitzen Rumpf und Glieder initiale Längen, die sich aus der Gesamtgröße des Icons ergeben. Die Größe des Icons hängt direkt von der Größe der geladenen Karte ab. Die Winkel des Rumpfes und der Glieder zueinander sind zu Beginn ebenfalls vorgegeben.

Die Abbildung der einzelnen Sichtweisen bzw. Interessengruppen auf den Rumpf und

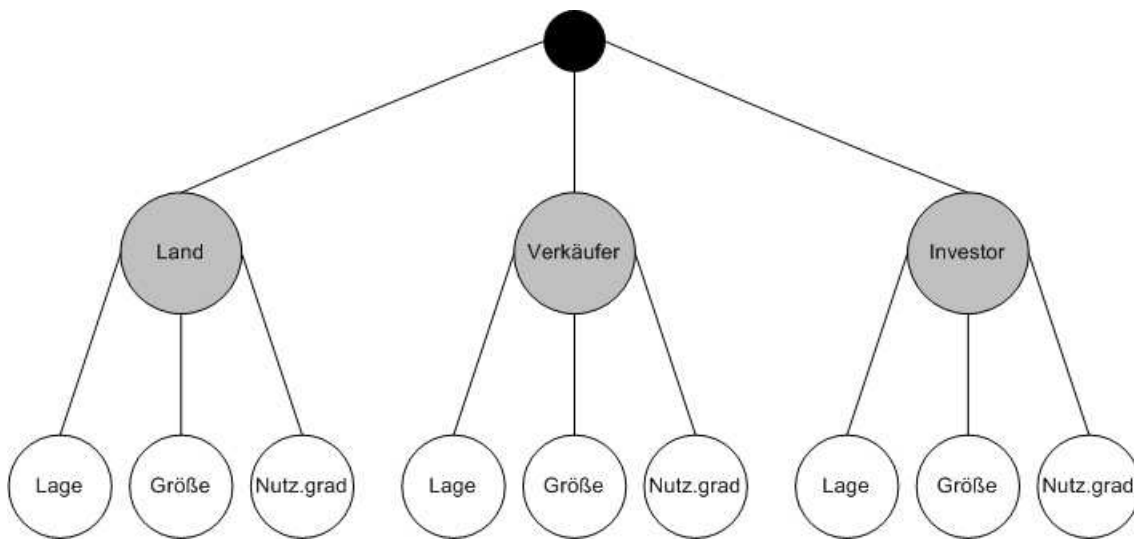


Abbildung 5.7: Sichtweisen auf die Daten

die Glieder ist variabel und kann zum Zeitpunkt der Berechnung ausgewählt werden. Dies ermöglicht dem Benutzer, die Daten aus unterschiedlichen Blickwinkeln zu betrachten, um so möglichst alle interessanten Eigenschaften der Daten visuell zu erfassen.

Die *Picklock-Icons* werden mittels des Voronoidiagramms so in einem 2D-Raster platziert, dass sie sich nicht überschneiden.

Eine zufällige Anordnung der Icons führte automatisch zu Überschneidungen und erzeugt Strukturen, die in den Daten nicht vorhanden sind. Aus diesen Gründen werden die Icons in regelmäßigen Abständen angeordnet (siehe Abbildung 5.11).

5.3 Spike-Tree

Wie bereits gesehen sind Visualisierungstechniken, die letztendlich in einer zweidimensionalen Darstellung der Daten münden, in diesem Zusammenhang problematisch: Die Darstellung überlagert die ebenfalls zweidimensionale Karte, wodurch es schwierig ist, die Daten in einen räumlichen Bezug zur Karte zu setzen. Besonders vorteilhaft sind deshalb Techniken, die in drei Dimensionen arbeiten.

Von den in Kapitel 2.2.2 angeführten Visualisierungstechniken nutzen lediglich die in Kapitel 5.2.1 beschriebenen Stick-Figures und die Cone-Trees sowie die Technik der Botanischen Informationsvisualisierung die dritte Dimension.

Die Stärke der Cone-Tree-Technik [GR91] liegt in der optimalen Ausnutzung der dritten Dimension des Darstellungsraumes und in der für den Anwender gegebenen Möglichkeit, die Kegel zu drehen und heranzuzoomen. Hierdurch können neben der

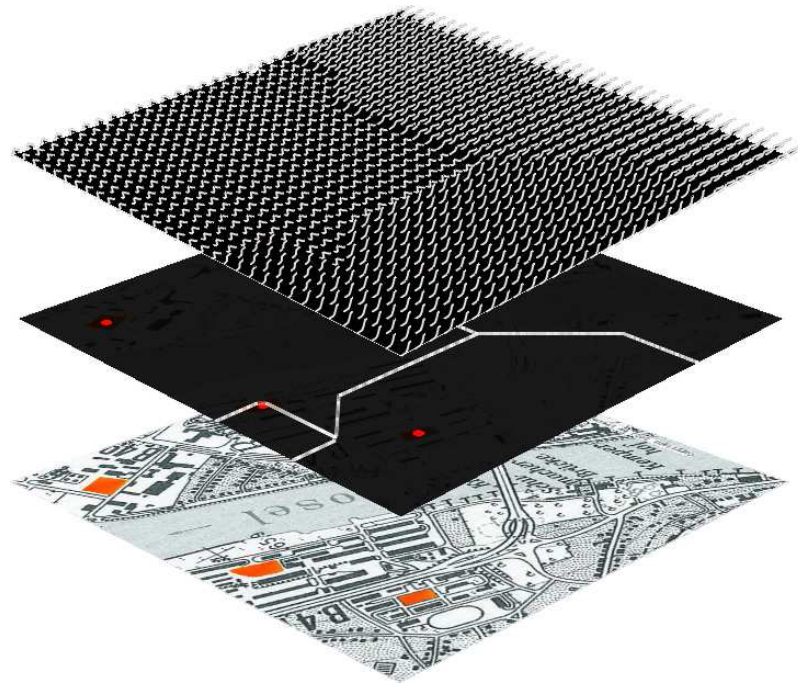


Abbildung 5.8: Vorgehensweise bei der Berechnung und Darstellung von Karte, Voronoidiagramm und *Picklock-Icons*

gesamten Informationsdarstellung auch lokale Details betrachtet werden. Allerdings erweist sich diese nur dann als notwendig, wenn die darzustellende Informationsstruktur sehr umfangreich und verschachtelt ist, wie zum Beispiel bei der Darstellung der Daten und Programme einer Festplatte. Diese extrem hohe Komplexität ist bei den vorliegenden Daten nicht gegeben.

Die Landscape-Technik nutzt zwar ebenfalls die dritte Dimension, sie eignet sich jedoch nicht zur Darstellung graph-basierter, hierarchischer Informationen und scheidet dadurch für den gegebenen Anwendungszweck ebenfalls aus.

5.3.1 Botanische Visualisierung

Der Vorteil der botanischen Informationsvisualisierung liegt in der kontextsensitiven Darstellung der Daten in Bezug zu einer Karte. Streng genommen bietet sich diese Technik für die Darstellung hierarchischer Daten an, sie hat jedoch auch in anderen Bereichen ihre Vorteile. Die gewünschte Visualisierungstechnik sollte sich räumlich mit dem Ausschnitt einer Landkarte überlagern, wie das auch bei dem *Picklock-Icon* der Fall ist.



Abbildung 5.9: In die Textur geladene Karte

Kleiberg, van de Wetering und van Wijk stellen in [KvdWvW01] eine Methode vor, wie auf effektive Art und Weise eine große Anzahl von Elementen gruppiert und dargestellt werden kann. Dabei verwenden sie einen rekursiven Ansatz.

Für die Darstellung hierarchischer Daten werden bevorzugt sogenannte *trees* verwendet, in dem Kontext der Darstellung von Dateisystemen wird auch von einem *Verzeichnisbaum* gesprochen. Diese Methoden haben das Problem, dass sie bezüglich der Anzahl der gleichzeitig darstellbaren Elemente limitiert sind, da der zur Verfügung stehende Platz nicht effektiv genutzt wird. Eine Lösung für das Problem in 2D bieten die in Kapitel 2.2.1 vorgestellten Treemaps [Shn92], [Joh93]. Diese nutzen den zur Darstellung verfügbaren Platz effizient aus.

Eine weitere Möglichkeit ist die Erweiterung des Darstellungsraumes um eine weitere Dimension, d.h. die räumliche, dreidimensionale Darstellung der Bäume. Das Ziel hierbei ist, die Elemente und ihre Anordnung trotz ihrer großen Anzahl möglichst einfach durch den Betrachter visuell extrahieren zu können. Baumdarstellungen, bei denen eine große Übereinstimmung ihres Aussehens mit dem ihrer echten, biologischen Vorbildern erreicht wird, erfüllen diese Anforderungen.

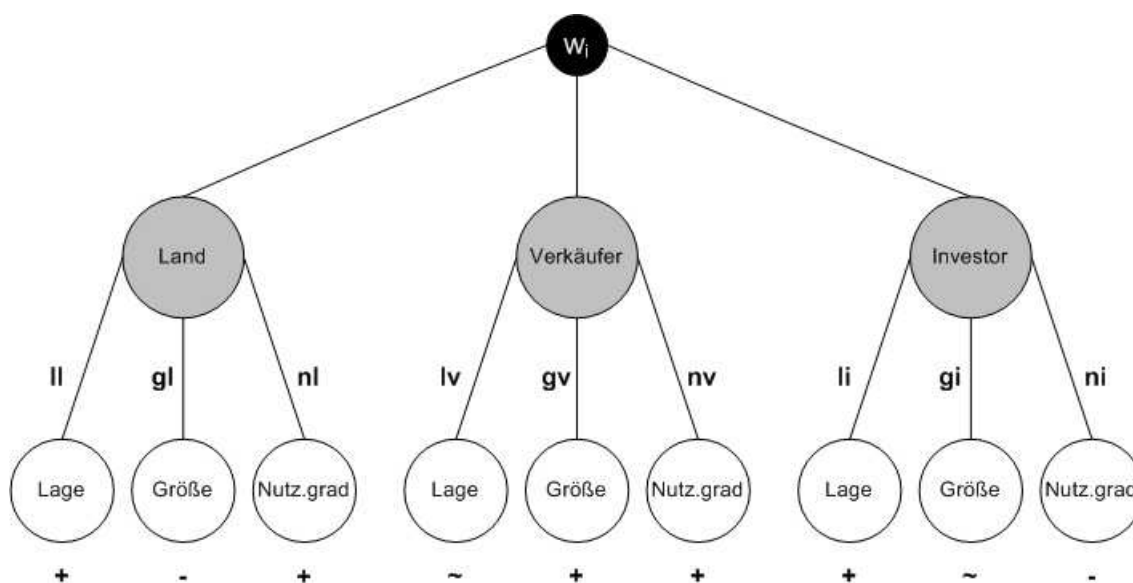


Abbildung 5.10: Bestimmung der Gewichte

Ader-Modell botanischer Bäume

Bei der Visualisierung ist die Struktur der Bäume durch die darzustellenden Daten determiniert, wodurch sich die Modellierung relativ einfach gestaltet. Dazu wird das *strand model* zur Darstellung natürlich aussehender Bäume von Holton [Hol94] in vereinfachter Form verwendet. Die Adern werden benutzt, um die innere Gefäßstruktur eines Baumes nachzuahmen. Dabei wird der kleinste Ast durch eine einzige Ader und jeder andere Ast durch die Anzahl von Adern dargestellt, die der Summe der Adern der Äste entsprechen, in die er weiter unterteilt wird. Die Adernbündel, welche die verschiedenen Äste repräsentieren, werden in Form eines Zylinders dargestellt.

Ader-Modell von Verzeichnisbäumen

Um mit dem *strand model* ein dreidimensionales Modell eines Verzeichnisbaumes zu generieren, muss zuerst festgelegt werden, wie die Unterverzeichnisse und Dateien an den Verzweigungspunkten abgespalten werden. Damit das Ergebnis einem echten Baum oder einer Pflanze ähnelt, werden in absteigender Reihenfolge zuerst die größten Unterverzeichnisse und Dateien abgespalten. Abbildung 5.12 verdeutlicht dies. In der Abbildung ist oben ein Node-Link-Diagramm einer Verzeichnisstruktur und darunter das korrespondierende Ader-Modell zu sehen. Die Zahlen entsprechen der Größe der Dateien und Verzeichnisse.

Obwohl ein Verzeichnisbaum im Allgemeinen kein Binärbaum ist, ist das Ergebnis der Visualisierung nach diesem Modell dennoch ein binärer Baum, denn bei ab-

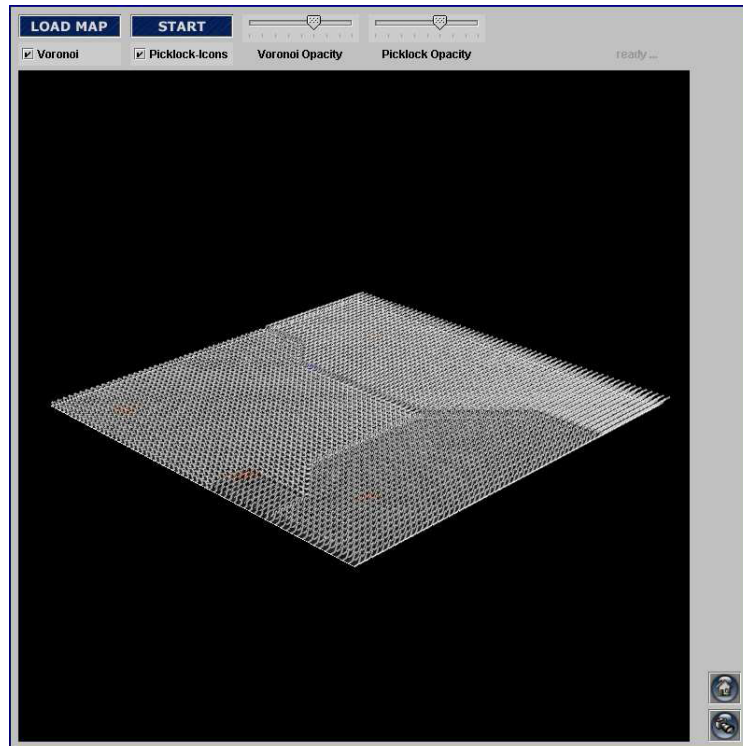


Abbildung 5.11: Visualisierung multivariater Daten unter Verwendung der Manhattan-Metrik und des *Picklock-Icons*

und aufsteigender Rekursion wird jeweils ein Ast abgespalten. Die zur Modellierung notwendige Anzahl von Adern wird durch die rekursive Funktion

$$strands(dir) = \sum_{d \in dl} strands(d) + \sum_{f \in fl} f.size \quad (5.1)$$

berechnet. Die Funktion traversiert die Verzeichnisse bis hinunter zur Dateiebene und liefert die Größe der Dateien f in den jeweiligen Verzeichnissen d zurück. Nach Beendigung der Rekursion liefert $strands(dir)$ eine zur Größe des Verzeichnisses dir proportionale Anzahl von Adern zur weiterführenden Modellierung. Der fortlaufende Ast eines großen Verzeichnisbaumes besitzt dann vergleichsweise viele Adern, dargestellt durch ein dickes und langes erstes Segment.

Die grafische Darstellung des Modells in Abbildung 5.13 weist Probleme hinsichtlich der Interpretierbarkeit durch den Benutzer auf. So ist z.B. die Verzweigungsstruktur unübersichtlich. Verzeichnisse mit einer großen Zahl von Unterverzeichnissen und Dateien führen zu sehr langen und dünnen Ästen. Außerdem ist es möglich, dass die Blätter sich zu stark überlappen und somit die Visualisierung behindern.

Das erste Problem wird durch eine Glättung der Übergänge zwischen zwei Zylind-

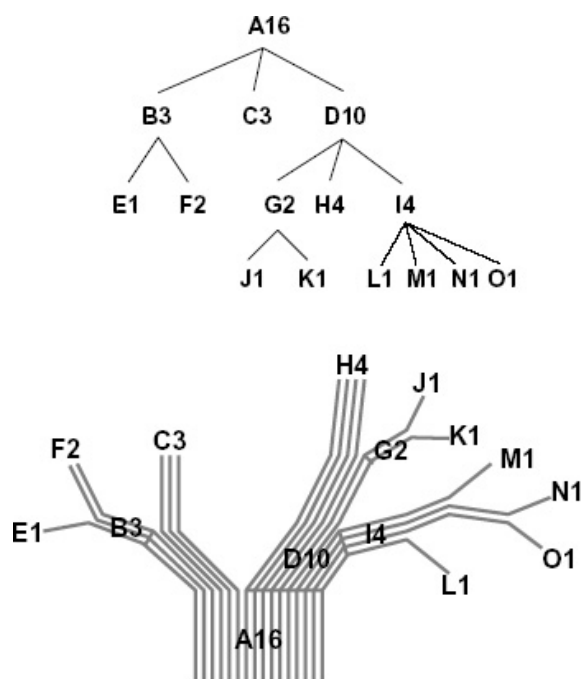


Abbildung 5.12: Node-Link-Diagramm und entsprechendes Ader-Modell

den mittels Extrusion behoben. Um zu verhindern, dass die Äste zu lang und zu dünn werden, wird zusätzlich ein Entscheidungskriterium (Kontraktion) eingeführt, welches die Darstellung von neuen Stämmen dann verhindert, wenn nicht mindestens eine vorgegebene Anzahl von Adern vom übergeordneten Stamm abgespalten werden.

Wichtig ist in diesem Zusammenhang, die Überlappung der Blätter zu verhindern. Hierfür wird ein Symbol verwendet, welches eine Frucht darstellt und aus einer Kugel und einem darauf aufsetzenden Kegel besteht. Dabei wird die Kugel am Ende eines Astes in so viele Schichten aufgeteilt, wie Elemente dargestellt werden sollen. Die so unterteilte Kugel wird in Anlehnung an [LD99] *Phi-Ball* genannt.

Die Fläche und die Höhe der einzelnen Schichten ist dabei proportional zur Größe der entsprechenden Dateien. Anschließend wird auf jeder Schicht ein Kegel platziert und um den Winkel β zur übergeordneten Scheibe gedreht. Um ein optisch ansprechendes Ergebnis zu erzielen, werden die Kegel so berechnet, dass sie tangential an der Kugeloberfläche aufsetzen.

Im Querschnitt betrachtet verläuft die Achse eines Kegels entlang der z -Achse. Die Hypotenuse des Schnittdreiecks besitzt die Länge c . Aus diesen Werten können der Punkt T an der Spitze des Kegels, der Basispunkt B und der Radius der Bodenfläche berechnet werden (siehe Abbildung 5.14). Mit diesen drei Parametern wird der Kegel

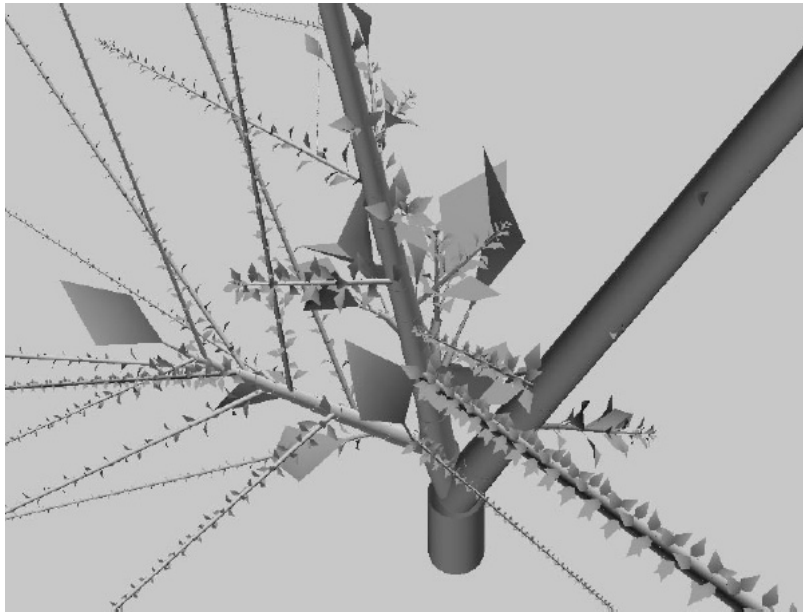


Abbildung 5.13: Verzeichnisbaum

positioniert und anschließend um den Winkel $\angle POQ$ um die x -Achse rotiert. In den folgenden Schleifendurchläufen wird ein Kegel mit dem Winkel $\beta = 360/\varphi$ um die y -Achse rotiert. Der Vorteil bei der Verwendung von Kegeln anstelle von Farbklecken auf den Kugeloberflächen ist die Ausnutzung der dritten Dimension: Kegel, die große Dateien symbolisieren, treten entsprechend größer hervor, wohingegen kleine Dateien eher als flache Kegel erscheinen. Durch die zusätzliche Verwendung farbiger Kegel können verschiedene Dateitypen unterschieden werden.

5.3.2 Konzept

In Kapitel 5.3.1 wurde dargestellt, dass sich die Technik der Botanischen Informationsvisualisierung besonders zur Darstellung hierarchischer Daten eignet. Wie beim *Picklock-Icon* wird die sichtspezifische, hierarchische Anordnung der Daten in Form eines Graphen (siehe Abbildung 5.7) erfasst. Deswegen wird bei dem *Spike-Tree* die Struktur des Graphen mittels einer Baumdarstellung in drei Dimensionen visualisiert.

In diesem Kapitel wird der Aufbau des eigentlichen *Spike-Trees* beschrieben. Die darzustellenden Informationen wurden zunächst ausschließlich in den Baumfrüchten kodiert. Für die vorliegende Eignungsprüfung der Methoden reicht diese Codierung der Informationen aus, eine Erweiterung auf eine Codierung in Stamm und Ästen stellt jedoch kein Problem dar.

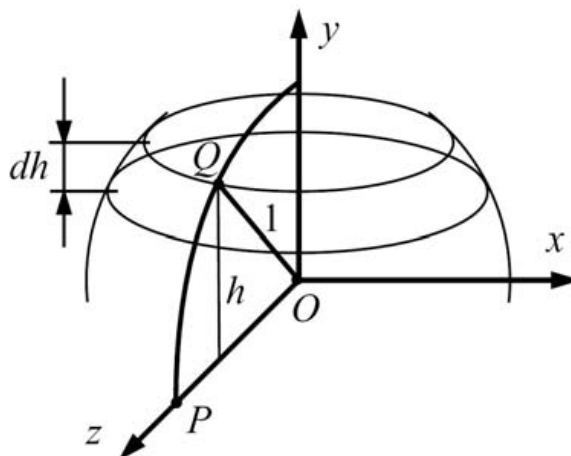


Abbildung 5.14: Konstruktion des Phi-Balls

Der untere Teil des Baumes besteht aus einem auf dem Schwerpunkt der jeweiligen Voronoiregion stehenden Stamm. Am oberen Ende des Stammes zweigt dann für jede Interessengruppe jeweils ein Ast ab. Von diesen drei Ästen zweigen dann wieder drei Äste ab, die eine Beschreibung der einzelnen Parameter darstellen und jeweils in der Farbe der jeweiligen Interessengruppe markiert sind.

Die Früchte an den Enden der neun oberen Äste sind aus Symmetriegründen kugelförmig gewählt. Die hierarchische Struktur des Baumes ist somit vollkommen analog zu der des Graphen in Abbildung 5.7. Die Informationen aus Abbildung 5.18 werden dabei durch die Größe der Früchte visualisiert. Je nach Einfluss der Parameter (+, o, -) wird der initial vorgegebene Radius der Kugeln dementsprechend größer oder kleiner skaliert.

Der eigentliche *Spike-Tree* entsteht jedoch erst durch die Hinzunahme eines Kegels zu jeder kugelförmigen Frucht. Genau wie die Radien der Kugeln, wird auch die Höhe dieser *Spikes* durch die Parameter kontrolliert. Aus ästhetischen Gründen und damit es für den Benutzer möglichst einfach ist, die Informationen durch die Visualisierung zu interpretieren, werden die *Spikes* so platziert, dass deren gekrümmte Oberfläche tangential mit denen der Kugeln abschließen. Dies verdeutlicht in Abbildung 5.17 der rechte Winkel auf der linken Seite der Frucht.

Zur Darstellung des *Spikes* muss die Höhe s und der Radius x des *Spike*-Kegels berechnet werden. Der Radius r der Frucht und der Wert h , der die Höhe des *Spikes* beeinflusst, sind durch die Kontrollparameter gegeben. Zusätzlich ist es für den Benutzer während des Visualisierungsprozesses möglich, den Wert von h und damit die Höhe des *Spike*-Kegels interaktiv zu variieren.

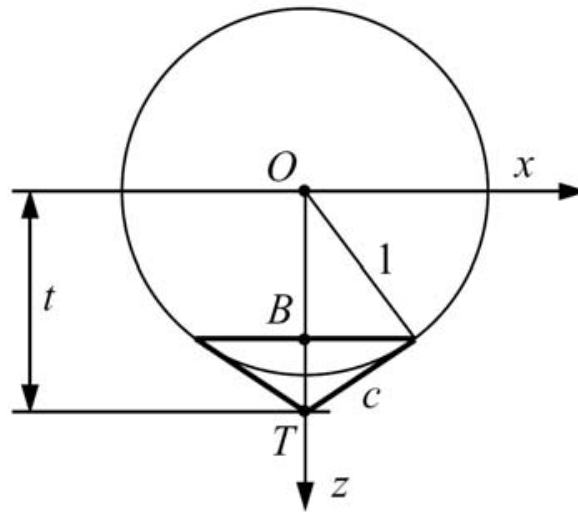


Abbildung 5.15: Berechnung der Kegelparameter

Um den Radius x des Kegels zu berechnen, wird zunächst der Winkel α bestimmt:

$$\alpha = \arcsin \frac{h}{r} \quad (5.2)$$

Anschließend wird über den Winkel α der Radius x wie folgt berechnet:

$$x = r \cdot \cos(\alpha) \quad (5.3)$$

Die Forderung, dass der Kegel tangential auf der kugelförmigen Frucht aufsetzen soll, liefert schließlich die Höhe s des Kegels:

$$s = \frac{r}{\sin(\alpha)} - h \quad (5.4)$$

Vor der Platzierung auf der Kartentextur werden die *Spike-Trees* soweit verkleinert, dass es zwischen ihnen zu keiner Überschneidung kommt. Der entsprechende Skalierungsfaktor aller *Spike-Trees* wird in vier Schritten bestimmt:

1. Berechnung des minimalen Abstandes der Voronoiorte
2. Bestimmung des maximalen Skalierungsparameters der Baumfrüchte
3. Ermittlung der maximalen Baumausdehnung in einer Richtung mittels dieses Skalierungsparameters

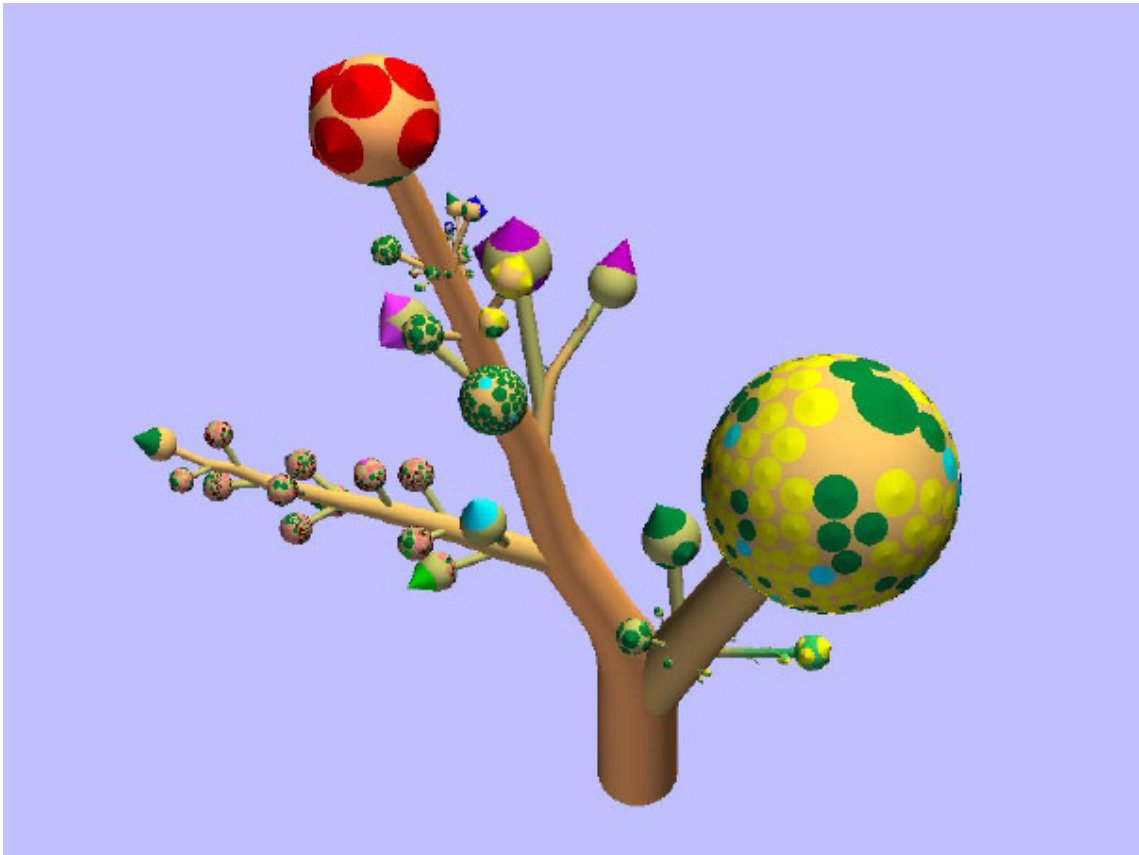
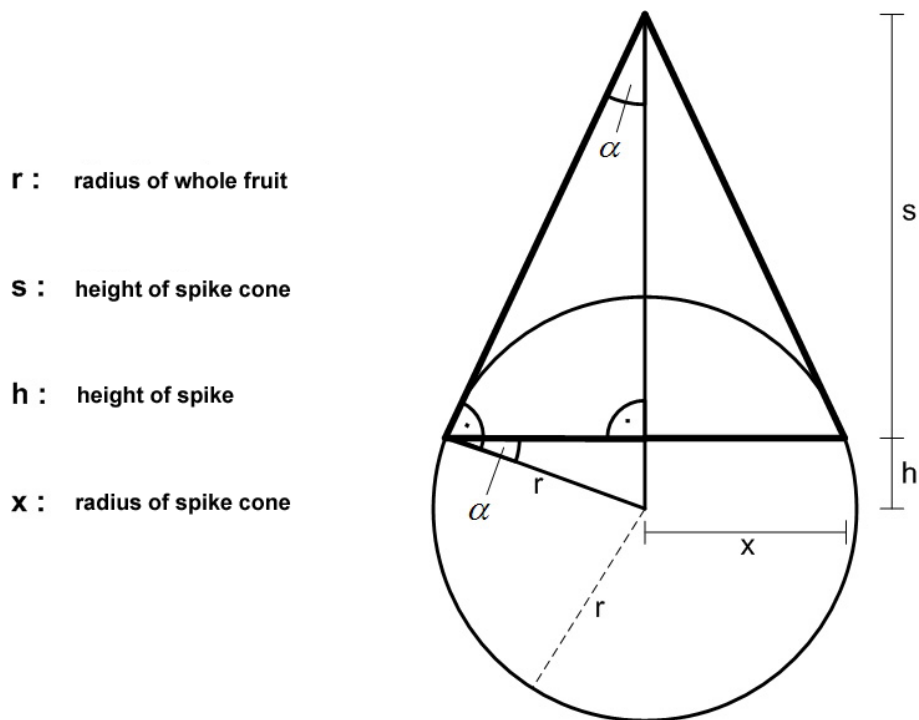


Abbildung 5.16: Modell mit Kontraktion, Extrusion und Phi-Ball

4. Berechnung des *Spike-Tree*-Skalierungsfaktors, so dass die Bäume unter Annahme maximaler Ausdehnung und minimalen Abstandes keine Überschneidungen aufweisen

Dazu muss gelten: die maximale Baumausdehnung in einer Richtung darf höchstens der Hälfte des minimalen Abstandes der Voronoiorte entsprechen. Durch dieses Worst-Case-Szenario fallen die Bäume unter Umständen relativ klein aus, jedoch gestaltet sich die Skalierungsberechnung unter diesen vereinfachten Annahmen unkompliziert und schnell. Ein weiterer Vorteil der Skalierungsberechnung ist die Anpassung der Größe der *Spike-Trees* an die Ausmaße der Kartentextur. Da die Berechnung in Canvas-Koordinaten von den Abständen der Voronoiorte in der Kartentextur abhängt, sind zum Beispiel die Bäume bei einer niedriger aufgelösten Karte dementsprechend kleiner. Wurden die Bäume skaliert, werden sie oberhalb der Texturen von Karte und Voronoidiagramm an den zugehörigen Orten platziert (Abbildung 5.19).

Abbildung 5.17: Parameter und Berechnung eines *Spike*-Kegels

5.3.3 Umsetzung

Als Grundlage für ein konkretes Beispiel dient wie bei dem *Picklock-Icon* die Topographische Karte der Stadt Koblenz (siehe Abbildung 5.6). Durch den hohen Detaillierungsgrad der Karte eignet sie sich auch besonders für die Umsetzung des *Spike-Trees*. Neben jeder Art von Straßen sind auch kleinere Wege und die für den vorliegenden Zweck notwendigen Flächen -bebaut oder unbebaut- verzeichnet.

Die grafische Oberfläche der zugrundeliegenden CD-ROM erlaubt es, von jedem anwählbaren Ort eines Kartenausschnittes die Gauß-Krüger-Koordinaten oder die Ortskoordinaten nach der universalen Mercator-Projektion zu ermitteln. Zusätzlich besteht die Möglichkeit Entfernungen und Flächeninhalte zu messen. Dadurch wird den Konversionsflächen eine eindeutig definierte Ortsposition und Flächengröße zugeordnet. Dazu kommen Informationen über den Nutzungsgrad der Fläche und die Parameter, welche die Daten der Konversionsfläche aus den drei verschiedenen Sichtweisen heraus gewichten. Die Sichtweisen entsprechen dabei denen von Land, Verkäufer und Investor. Wie beim *Picklock-Icon* sind die darzustellenden Informationen somit zunächst auf neun begrenzt (siehe Abbildung 5.7).

Diese Abbildung beschreibt prinzipiell eine hierarchische Anordnung von Informationen bzw. Daten. Der Wurzelknoten entspricht dabei einer allgemeinen Gesamtsicht

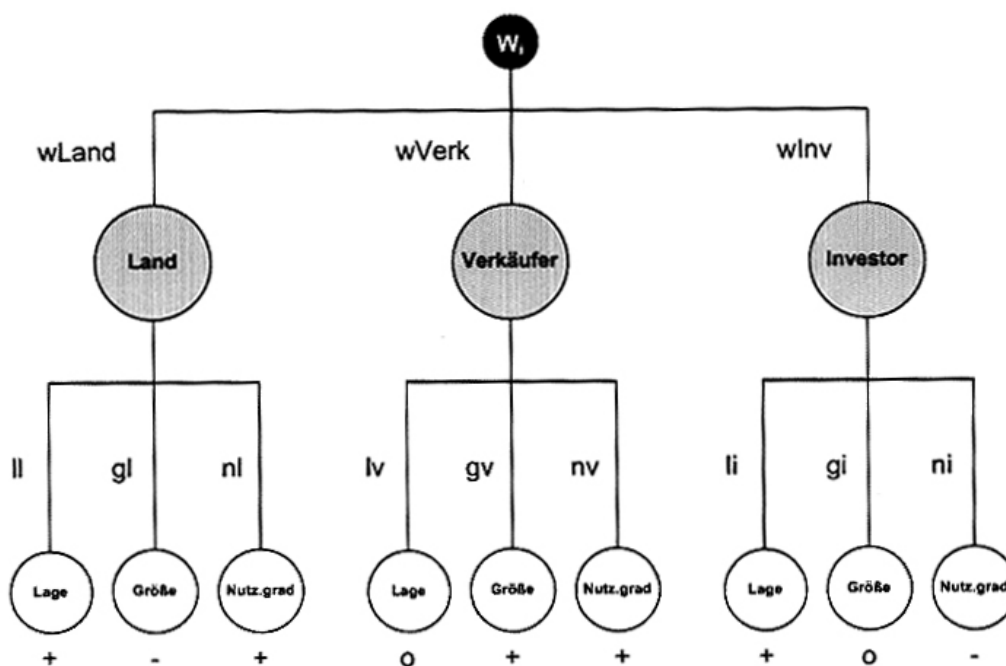


Abbildung 5.18: Bestimmung der Gewichte eines Voronoiortes

auf die Daten einer Konversionsfläche. Auf der nächsten Ebene wird die Sichtweise durch die drei an der Konversion beteiligten Interessengruppen Land, Verkäufer und Investors genauer spezifiziert. Dazu kommen die sichtweisenabhängigen Parameter Lage, Größe und Nutzungsgrad.

Für die vorliegende Anzahl von neun Informationen ist die Komplexität der Darstellung begrenzt, was sich vorteilhaft auf die visuelle Interpretation der Daten durch den Benutzer auswirkt. Die gewählte Methode lässt eine Erweiterung auf eine sehr große Anzahl an Dimensionen ohne Probleme zu. In Abbildung 5.20 ist die Baumdarstellung des Graphen zu sehen. Den Informationen, welche in Abbildung 5.7 auf unterster Ebene durch die Blätter repräsentiert werden, entsprechen in der Baumdarstellung kugelförmige Früchte an den Enden der oberen Äste. Um diese den einzelnen Sichtweisen optisch zuzuordnen zu können, wurden nicht nur die Früchte, sondern auch die dazugehörigen Teilbäume farblich markiert: die Farbe Rot entspricht der Sichtweise des Landes, Grün dem Verkäufer und Blau ist dem Investor zugeordnet.

Für den vorliegenden prinzipiellen Test für die Eignung der Methode war es zunächst ausreichend das Layout des Baumes statisch zu gestalten, d.h. es wurden mit Ausnahme der Früchte keinerlei Informationen in der geometrischen Konstruktion des Baumes codiert. Somit bestehen der Stamm an der Baumbasis und die Äste auf

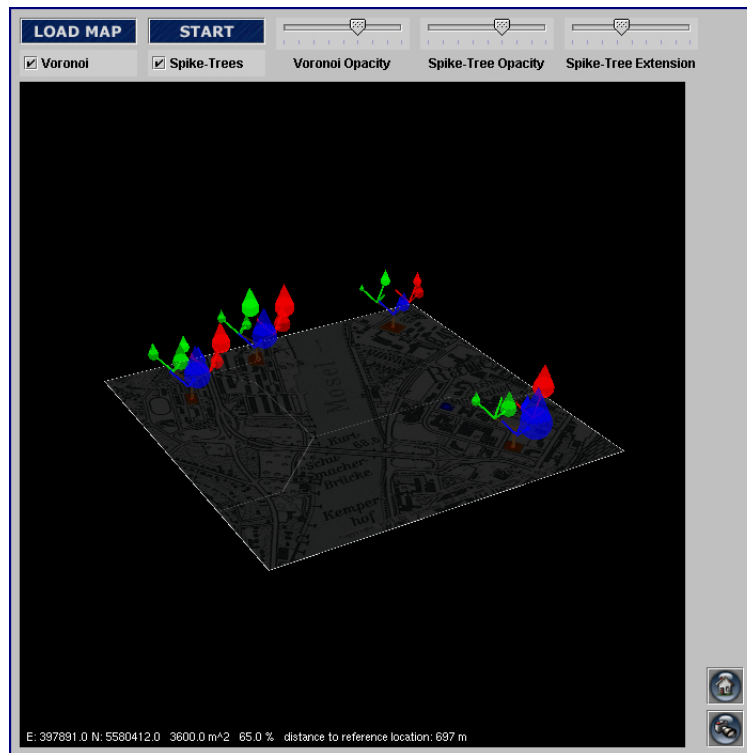


Abbildung 5.19: Informationsvisualisierung multivariater Daten mittels *Spike-Tree*

der ersten und zweiten Ebene aus Geometrien mit festen Ausmaßen. Ebenfalls wurden die Winkel, welche die relative Lage des Stammes, der Äste und der einzelnen Teilbäume zueinander beschreiben, fest gewählt. Abbildung 5.20 zeigt die Basisdarstellung des endgültigen Baumes, den *Spike-Tree*. Bei diesem wurde die Darstellung der Früchte dahingehend erweitert, dass sich die visuelle Interpretation durch den Benutzer noch intuitiver und einfacher gestaltet.

Da den Konversionsflächen und jedem Voronoiort ein *Spike-Tree* zugeordnet wurde, lässt sich die Eignung der Unterteilung der Konversionsfläche in Abhängigkeit der Sichtweise durch einen optischen Vergleich der zugehörigen *Spike-Trees* durch den Benutzer feststellen. Hierfür ist lediglich ein einfacher Größenvergleich der Baumfrüchte notwendig. Da als Teilform der Früchte des *Spike-Trees* die Kugel gewählt wurde, ist dies, bedingt durch die Symmetrie in allen drei Raumrichtungen, sehr einfach und effizient möglich.

Wie auch beim *Picklock-Icon* wurden die Implementierungsschritte und die Berechnung in folgender Reihenfolge vorgenommen.

1. Laden und Anzeigen der Karte als Vorschau

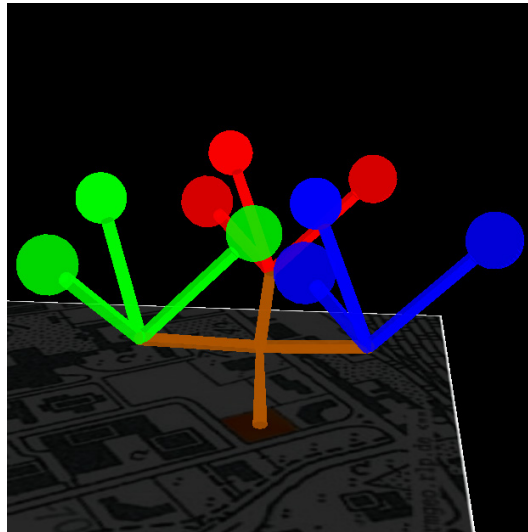


Abbildung 5.20: Basisdarstellung des Spike-Tree

2. Berechnung und Visualisierung des Voronoidiagrammes
3. Berechnung und Darstellung des *Spike-Trees*

Die daraus resultierende, endgültige Visualisierung umfasst die gleichzeitige Darstellung der Karte, des Voronoidiagrammes und des *Spike-Trees* (siehe Abbildung 5.21). Dabei wird die das Voronoidiagramm enthaltende Textur über die Kartentextur gelegt und schließlich die *Spike-Trees* auf diesen beiden Texturen räumlich platziert. Um dem Benutzer eine größtmögliche Freiheit bei dem Betrachten und Interpretieren der Visualisierung zu gewährleisten, sind die Textur des Voronoidiagramms und die Grafik der *Spike-Trees* in ihrer Transparenz bzw. Opazität variabel.

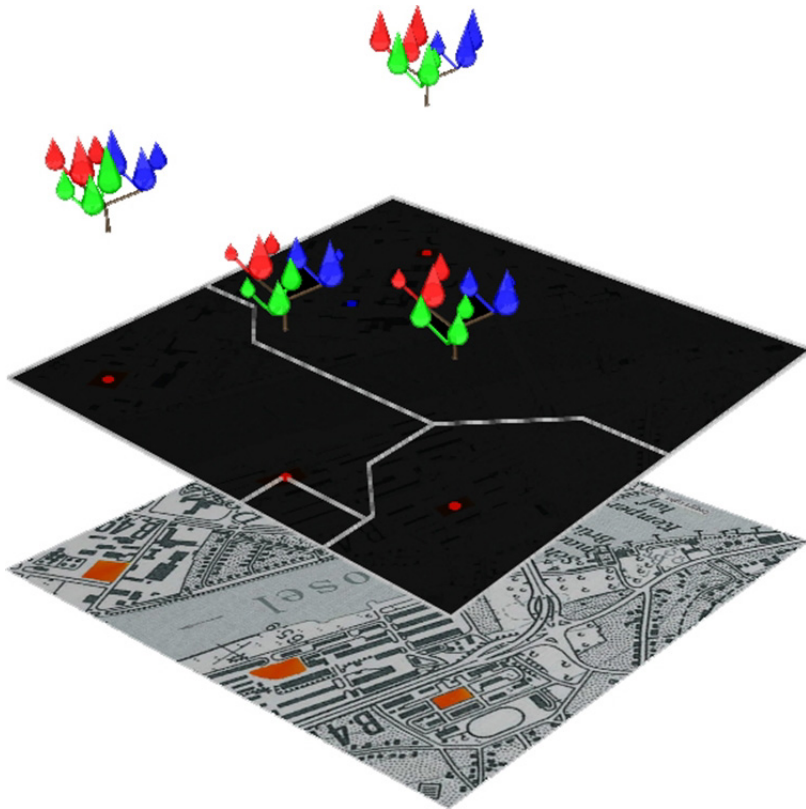


Abbildung 5.21: Vorgehensweise bei der Visualisierung von Karte, Voronoidiagramm und *Spike-Tree*

Kapitel 6

Informationssystem für die Stadtplanung

Informationssysteme halten immer mehr Einzug in die Stadtplanung. Oftmals werden hierbei einfach existierende GIS-Systeme mit zusätzlichen Informationen kombiniert. Die Lösung komplexer Fragestellungen, wie sie in der Stadtplanung vermehrt auftreten und in dieser Arbeit behandelt werden, bieten diese Systeme allerdings nicht an.

In diesem Kapitel wird ein Informationssystem für die Stadtplanung vorgestellt, das Lösungen für verschiedene, komplexe Anwendungen integriert. Mit dem System ist beispielsweise die Erstellung eines GIS-Berichtes und eines virtuellen Rundganges möglich. Im Gegensatz zu existierenden Systemen ist in dem vorgestellten System DaMaViS auch die Bearbeitung komplexer Fragestellungen in dem Bereich des Planungsablaufes möglich.

Hierzu wird das in Kapitel 3.1 erweiterte Modell des traditionellen Planungsablaufes am Beispiel von Konversionsflächen umgesetzt. Für das Clustering werden die in Kapitel 4 gefundenen Verfahren eingesetzt. Für die Visualisierung werden das Picklock-Icon und der Spike-Tree aus Kapitel 5 verwendet. Das so entstandene Tool IKone ermöglicht die praktische Umsetzung der in dieser Arbeit entwickelten Verfahren und bietet eine Möglichkeit, auch die komplexen Fragestellungen in der Stadtplanung zu lösen.

In Kapitel 6.1 werden zunächst grundsätzliche Überlegungen zu Systemen in der Stadtplanung entwickelt sowie das Konzept des im Rahmen dieser Arbeit implementierten Systems DaMaViS vorgestellt.

Kapitel 6.2 beschreibt den Systemaufbau von DaMaViS mit dem zugehörigen Client-Server-Modell. In Kapitel 6.2.1 wird das Tool DaVEa vorgestellt, das für die Verwaltung und Aktualisierung der Daten eingesetzt wird. In Kapitel 6.2.2 wird die Struktur des Clients mit der Clientapplikation CliVViS beschrieben. Der in Kapitel

6.2.3 beschriebene Server ist die zentrale Komponente des Systems, der die Anfragen der Clients koordiniert und den Zugriff auf die Datenbanken steuert.

In Kapitel 6.3 werden verschiedene Anwendungen vorgestellt, die im Bereich der Stadtplanung eingesetzt werden. Zunächst wird in Kapitel 6.3.1 das Tool IKone vorgestellt, in Kapitel 6.3.2 wird ein GIS-Bericht erstellt und in Kapitel 6.3.3 wird schließlich ein virtueller Rundgang erzeugt.

6.1 Konzept

In der heutigen Zeit produziert fast jede Anwendungsdomäne sehr große, heterogene Datenmengen. Für den Bereich der Stadtplanung eröffnet sich daraus ein immer breiteres Themenfeld, in dem vermehrt rechnergestützte Entscheidungshilfen Anwendung finden. Als Beispiel hierfür kann ein virtueller Rundgang durch ein geplantes Stadtgebiet genannt werden. Mit Hilfe dieses Verfahrens können Entscheidungen in der Stadtentwicklungspolitik zu einem Ziel geführt werden, da die visuelle Aufbereitung der Plandaten die Vorstellungskraft der Entscheidungsträger stärkt. Viele Auswertungen in der Stadtplanung werden noch von Hand erledigt, da für Rechneranwendungen keine geeigneten Module auf dem Markt angeboten werden. Die auf dem Markt befindlichen Module sind vielfach Insellösungen, die in getrennten Arbeitsschritten, nacheinander ausgeführt, zu einer Problemlösung führen.

Mit Hilfe von Datenmanagementsystemen wird in den einzelnen Fachdisziplinen der Komplexität der Daten begegnet. Aufgrund der stark wachsenden Datenmengen bei der computergestützten Verarbeitung von Daten erfahren diese Systeme eine immer stärkere Verbreitung im beruflichen sowie im privaten Umfeld. Es existieren diverse Systeme zur Bewältigung dieser Datenflut. Diese Systeme sind aber meistens auf einen bestimmten Anwendungsbereich spezialisiert. Im Bereich der Geoinformationssysteme findet sich keine umfassende Lösung für komplexe Fragestellungen, obwohl auch hier immer größere Datenmengen anfallen.

Der Begriff des Geoinformationssystems wird in der Literatur in vielfältiger Weise definiert. Allgemein wird hiermit die Verarbeitung und Verwaltung raumbezogener Daten (Geo-Daten) mithilfe der elektronischen Datenverarbeitung bezeichnet. Als raumbezogen werden solche Informationen bezeichnet, die sich eindeutig mithilfe von Koordinaten in einem Untersuchungsgebiet verorten lassen. In dem Anwendungsfeld Raum- und Umweltplanung existieren diverse spezialisierte Informationssysteme wie beispielsweise Landinformationssysteme, Rauminformationssysteme und Umweltinformationssysteme (siehe auch Kapitel 2.3.3). Diese Systeme bieten spezifische Auswertungen einzelner Anwendungsbereiche an.

Das hier vorgestellte Datenmanagement- und Visualisierungssystem *DaMaViS* (**D**aten-**M**anagement und **V**isualisierung**S**ystem) [HSM⁺05] vereinigt verschiedene Insellösun-

gen in einem System für die Anwendung in der Stadtplanung. *DaMaVis* benutzt generische Datenstrukturen, wodurch es in der Lage ist, Daten aus den unterschiedlichsten Anwendungsdomänen zu verarbeiten.

Das Datenmanagementsystem *DaMaVis* ist ein komponentenbasiertes Client-Server System [HSM⁺04]. Neben seiner generischen Datenstruktur zeichnet sich das System durch die Verwendbarkeit für andere Anwendungen und die Kompatibilität zu anderen Systemen bei schneller bedarfsgerechter Anpassungsfähigkeit aus. Für die Kommunikation zwischen Client und Server wird CORBA (Common Object Request Broker Architecture) eingesetzt. *DaMaVis* basiert auf Open Source Bausteinen, was flexible Einsatzmöglichkeiten zulässt und einen weiteren bedarfsgerechten Ausbau in der Zukunft ermöglicht.

DaMaVis verknüpft eine große Zahl verteilt vorliegender, heterogener, sehr großer Datenbanken miteinander und ermöglicht eine gezielte Auswertung der gesammelten Daten im Bereich der nachhaltigen Stadtteilentwicklung (z.B. Bahnhofsumgebung Kaiserslautern) und der Konversion im weiteren Sinne. Die Basis bilden verschiedenste Datenbanken mit Statistiken, Texten und Bildern sowie Graphiken. Insgesamt erfolgt eine Visualisierung der Ergebnisse mittels geeigneter Techniken aus dem Bereich der Informationsvisualisierung. Neben den vorab genannten spezifischen Anwendungsmöglichkeiten bietet das System grundlegende Möglichkeiten, Texte zu bearbeiten und einzelne Textpassagen in einen Report zu integrieren.

Beispielsweise kann es eingesetzt werden, um die Daten eines Geographischen Informationssystems (GIS) zu verwalten und zu visualisieren. Ein weiteres Anwendungsfeld ist die Erstellung eines virtuellen Walkthrough durch eine Stadt oder kleinere Teilräume.

Ein besonderer Schwerpunkt des Systems liegt auf dem Tool *IKone* [HSRS06], das die in dieser Arbeit entwickelten Verfahren am praktischen Beispiel der Umnutzung von Konversionsflächen umsetzt.

6.2 Systemaufbau

Ziel des Projektes *DaMaVis* (*DatenManagement Visualisierungssystem*) war die Erstellung eines verwaltungsspezifischen Datenmanagement- und Visualisierungssystems für die gezielte Anwendung in der Stadtplanung. Ein Schwerpunkt im Bereich Datenmanagement liegt auf der Koordination der heterogenen Datenmengen, ein weiterer Schwerpunkt liegt im Bereich Visualisierung. Das Ziel der Visualisierung ist es eine Auswertung und Präsentation der Daten vorzunehmen.

Die Komplexität und Heterogenität der vorhandenen Datenbanken, die sowohl verteilt über mehrere Standorte, als auch sehr umfangreich sind, erfordert einen erhöhten Aufwand für die Datenkoordination und -haltung.

Das Entwicklungskonzept für das Projekt *DaMaVis* [HMS⁺05] basiert daher auf einem Client-Server-System, wobei die Hauptaufgaben auf dem Server ausgeführt werden, so dass nur die Ergebnisse auf den Arbeitsplatzrechner transferiert werden. Abbildung 6.1 beschreibt die Systemstruktur sowie den prinzipiellen Aufbau des Ser-

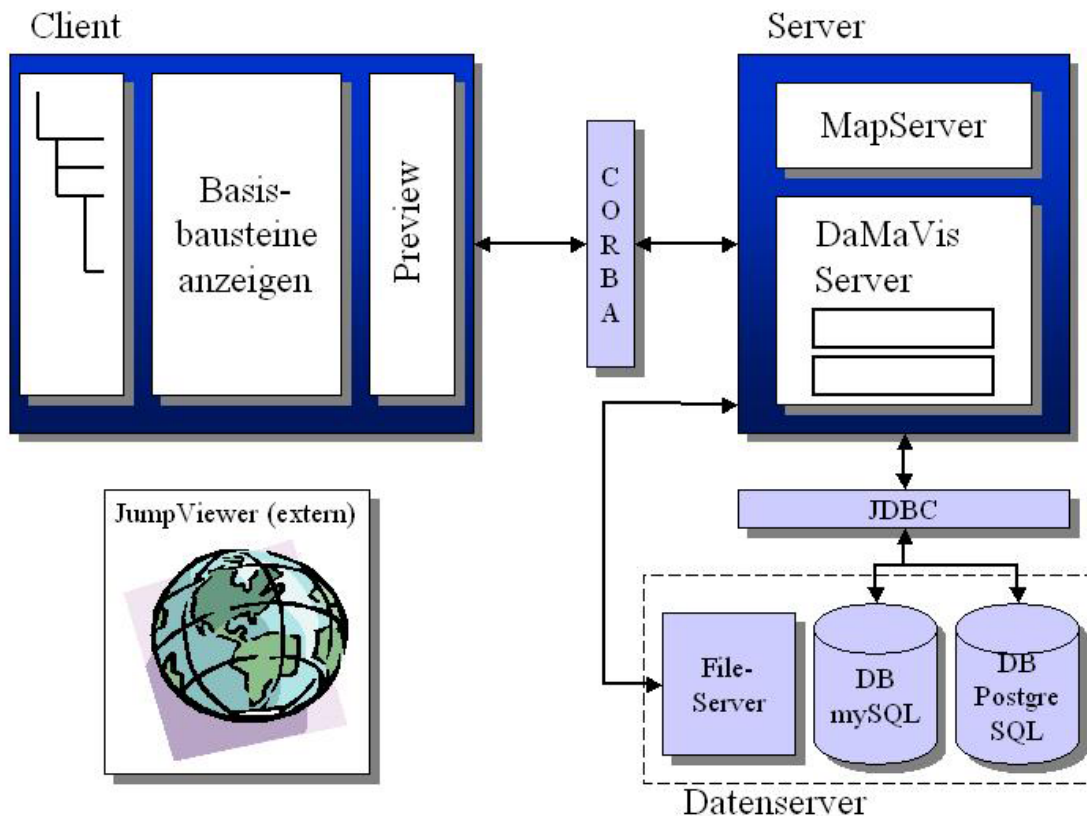


Abbildung 6.1: Client-Server-Modell

vers und des Clients. Das System ist flexibel gestaltet, was die einfache und effiziente Einbindung weiterer externer Programme ermöglicht (z.B. das externe Programm JUMP-Viewer).

Zur Gewährleistung der Plattformunabhängigkeit wird für die Kommunikation zwischen den einzelnen Komponenten die weit verbreitete Middleware CORBA der OMG (Object Management Group) eingesetzt. Aus Sicherheitsgründen erfolgt der Zugriff vom Client auf den Server mittels Passwortschutz und gesicherter CORBA-Verbindung. Die Funktionalitäten können vom Client aus so genutzt werden, als würden diese auf dem Client ausgeführt.

Aufgrund der Nutzung der JDBC-Technologie (*Java Data Base Communicator*)

unterstützt *DaMaVis* eine Vielzahl verschiedener Datenbanken und erreicht somit eine weitgehende Datenbankunabhängigkeit. Einzige Reglementierung bei der Einbindung der Datenbanken ist die Existenz eines JDBC-Treibers für die gewählte Datenbank.

Zum Laden und Speichern von Projekten wird die Metasprache XML (Extensible Markup Language) verwendet. Mit XML und einem geeigneten Style Sheet ist es möglich, Projektdaten in einem Format zu exportieren, welches unter anderem von Microsoft Office 2003-Produkten gelesen werden kann.

6.2.1 Datenverarbeitung

An den verschiedenen Standorten existieren Datenbanken mit sehr heterogenen Daten. Dieses sind sowohl numerische Daten, als auch Vektor- und Rasterdaten. Numerische Daten sind beispielsweise Gemeinde-, Förder-, Finanzdaten und statistische Daten. Vektor- und Rasterdaten liegen in Form kartografischer Daten vor. Deswei-

Abbildung 6.2: Datenbankadministrationstool *DaVeA*

teren existieren Textdateien, Bilder und Videos. Die Auswertung der Daten ist eine Mischung aus Standard- und fallspezifischen Prozessen. Für die Standardprozesse wurde ein Automatismus entwickelt, für die fallspezifischen Prozesse muss auf existierende Ressourcen zurückgegriffen werden. Das Datenbankadministrationstool *DaVeA* (Daten-Verwaltung-Entwicklungsagentur)

wird für die Verwaltung und Aktualisierung der existierenden Datenbanken sowie zum Einspielen neuer Daten eingesetzt (siehe Abbildung 6.2). Mit Hilfe dieses Tools ist es möglich die Pflege der Datenbanken anwenderfreundlich ohne direkte SQL-Anweisungen durchzuführen. *DaVEa* ist ein externes Programm, welches direkt auf die Datenbanken zugreift.

6.2.2 Client

Die Clientapplikation *ClivViS* (**C**lient: **V**iewer **V**isualisierungstool **S**trukturbaum) umfasst das eigentliche, auf dem Arbeitsrechner laufende Programm. Der Client dient zum Erzeugen neuer Projekte und zum Zugriff der auf dem Server abgelegten Projekte.

Der Client unterstützt das Erstellen von Dokumenten durch die Verwendung von Vorlagen. Soll ein neues Projekt gestartet werden, ist es möglich aus mehreren Vorlagen eine auszuwählen. Danach wird der Anwender durch die einzelnen Objekte (Tabellen, Bilder, Texte, Datenbankabfragen) geführt, um diese mit dem Inhalt aus den hinterlegten Daten zu füllen.

Der Client setzt sich aus drei Bereichen zusammen. Dem *Strukturbaum*, mit dessen

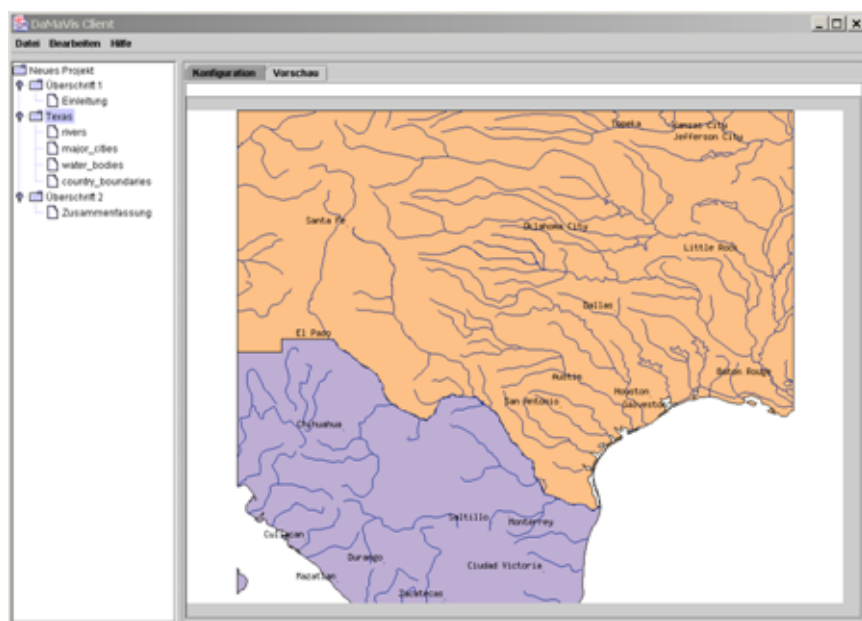


Abbildung 6.3: DaMaViS Client

Hilfe ein Überblick über das aktuell bearbeitete Projekt vermittelt wird, dem *Basisbaustein*, in dem eine direkte Bearbeitung einer einzelnen Komponente des Projektes erfolgen kann (eine detaillierte Bearbeitung ist mit den Office2003 Programmen

möglich) und dem *Preview*, durch das der Anwender einen visuellen Überblick über das Projekt erhält.

Der Client unterstützt das Erstellen von Dokumenten durch die Verwendung von Projektvorlagen. Soll ein neues Projekt gestartet werden, ist es möglich, eine Vorlage auszuwählen oder ein neues Projekt ohne Vorlage anzufertigen. Danach wird der Anwender durch die einzelnen Objekte (Tabellen, Bilder, Texte, Datenbankabfragen) geführt, um diese basierend auf den hinterlegten Daten mit Inhalt zu füllen. Der Client besteht aus einer Bedienoberfläche, mit der die Funktionen auf dem Server koordiniert werden, einem Fenster, in dem die Basisbausteine angezeigt werden und einem Vorschaufenster (*Preview*) der einzelnen Objekte.

Eine große Stärke des Systems ist die Benutzerfreundlichkeit. Diese wurde mit Hilfe einer intuitiven Menüführung und der Definition verschiedener Sichtweisen verwirklicht. Diese Sichtweisen entsprechen einem pyramidenförmigen Aufbau und reduzieren die Informationsvielfalt auf die benutzerspezifischen Bedürfnisse. Der ausgebildete Spezialist erhält die Möglichkeit, auf dem Server gespeicherte Daten und Abfragemodi zu verändern und hat somit einen umfassenden Zugriff auf das Gesamtsystem. Personen, die keine vertieften Kenntnisse des Systems besitzen, haben lediglich die Möglichkeit, Standardberichte aus bereits gespeicherten Vorlagen zu generieren.

Auf der linken Seite der Abbildung 6.2.2 wird der strukturelle Aufbau des aktuellen Projektes in Form eines Baumes visualisiert. Jedes Objekt des Projektes wird als ein Knoten bzw. ein Blatt dargestellt. Ein Knoten untergliedert sich in weitere Unterobjekte, ein Blatt stellt die tiefste Ebene der jeweiligen Hierarchie dar.

Innerhalb des Strukturbaumes ist es möglich, Objekte zu löschen oder zu generieren. Beispielsweise kann ein GIS-Report erzeugt werden, in dem unterschiedliche Textbausteine mit Tabellen und einer mit dem System erzeugten Karte kombiniert werden.

Die rechte Seite des Clients gliedert sich in zwei Karteikarten (*tabbed pane*). Die erste Karteikarte besteht aus einem Konfigurationspanel, für das aktuell selektierte Objekt des Strukturbaumes. Jedem Objekt des Strukturbaumes ist ein solches Panel zugeordnet. Die zweite Karteikarte beinhaltet einen *Preview*-Bereich, in dem eine Vorschau der aktuell ausgewählten Komponente des Strukturbaumes dargestellt wird. Auch hier ist jeder Komponente des Strukturbaumes ein eigener *Preview* zugeordnet. Wird beispielsweise eine GIS-Komponente in dem Strukturbaum selektiert, wird in dem *Preview*-Bereich die entsprechend generierte Karte (siehe Abbildung 6.3) eingeblendet.

6.2.3 Server

Der Server ist die zentrale Komponente des *DaMaViS*-Systems. Er koordiniert die Anfragen der Clients und steuert den Zugriff auf die Datenbanken. Um die Datensicherheit zu gewährleisten, ist die Funktionalität lediglich auf dem Server vorhanden. Der Server untergliedert sich in verschiedene Komponenten, damit das System flexibel gegenüber neuen Anforderungen bleibt.

Die Implementierung des Servers erfolgt mittels JAVA, was zur Flexibilität des Systems beiträgt. Für die Kommunikation zwischen dem Server und den Datenbanken werden JDBC-Treiber eingesetzt. Hierdurch wird die Flexibilität und Wiederverwendbarkeit des Systems erhöht sowie die Integration unterschiedlicher Datenbanksysteme ermöglicht.

Aus der Vielzahl von Datenbanksystemen wurde für den *DaMaViS*-Server die Datenbanksoftware PostgreSQL ausgewählt. Ausschlaggebend für die Wahl dieser Datenbank-Software ist die Erweiterbarkeit von PostgreSQL durch das Paket PostGIS welches es ermöglicht, GIS-Objekte zu verwalten.

PostGIS stellt eine Erweiterung für die relationale Datenbank PostgreSQL dar. Zusätzlich zu alphanumerischen Daten und Attributen können so auch geometrische Daten in der Datenbank verwaltet werden.

Die Organisation der geometrischen Daten in der Datenbank orientiert sich an den Richtlinien des OGC (OpenGIS Consortium). Die Standards des OGC sind weltweit verbreitet und sichern somit die Interoperabilität des Systems. Des Weiteren beherrscht PostgreSQL den Umgang mit Transaktionen bei der Verwaltung der Daten. Diese Eigenschaft ist besonders für den Mehrbenutzerbetrieb der Datenbank wichtig.

Ein weiterer Punkt für die Wahl von PostgreSQL ist die freie Verfügbarkeit der Software im Netz (Open-Source). Die Verwendung von Open-Source-Produkten sichert die Unabhängigkeit von kommerziellen Anbietern.

Das externe Programm MapServer ist in den Server integriert und bietet die Möglichkeit, basierend auf GIS-Daten Karten zu generieren, d.h. aus den Metadaten verschiedener Layer einer GIS Komponente werden Bilder erzeugt.

Die Konfiguration der zu erzeugenden Bilder erfolgt über so genannte Mapfiles. Dieser Mechanismus stellt die Basiskonfiguration des MapServers dar. Das Mapfile enthält die Informationen über das Aussehen des Bildes sowie die Orte an denen die benötigten Daten zu finden sind. Es beinhaltet keine GIS Daten. Das Endprodukt ist ein normales Bild, z.B. im jpeg-Format und kann an den Client zur weiteren Bearbeitung im aktuellen Projekt übertragen werden.

Im Folgenden wird die Datenstruktur des Servers kurz beschrieben. Eine detaillierte Beschreibung kann [Wad04a] entnommen werden.

Die Schnittstelle *DaMaVisBaseObject* wurde entwickelt, um die Konformität der

verschiedenen Objekte im Strukturbaum des *CliVViS* zu gewährleisten. Diese Schnittstelle gibt neben Operationen zur Wartung des Strukturbaumes auch XML-Operationen zur Speicherung der Daten vor. Des Weiteren wird eine Methode definiert, die zu jedem Objekt das zugehörige Konfigurationspanel liefert. Alle Objekte, die in den Strukturbaum aufgenommen werden sollen, müssen diese Schnittstelle implementieren. Ein Ausschnitt des Vererbungsdiagramms ist in Abbildung 6.4 dargestellt. Durch die enorme Variationsbreite der Objekte, die in die *DaMaViS*-Objekthierarchie

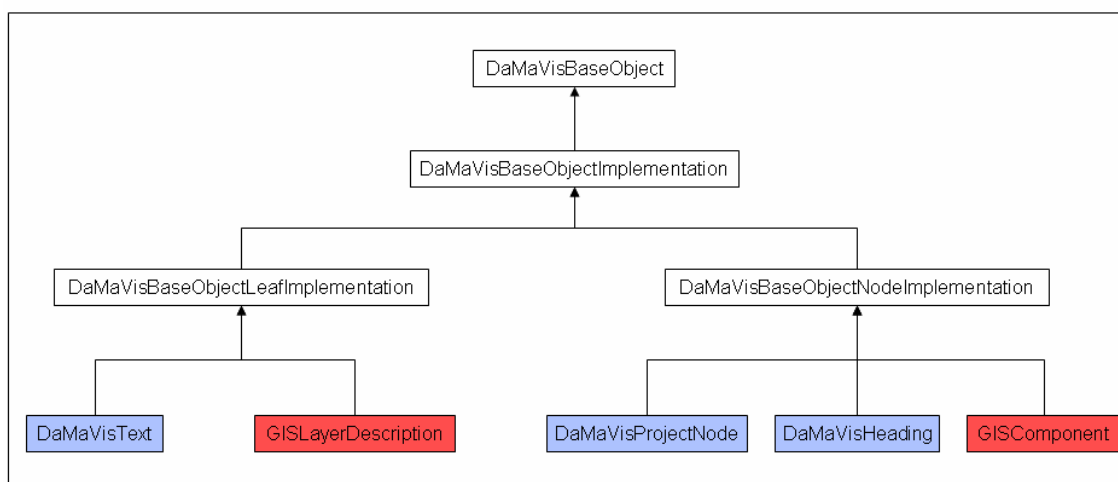


Abbildung 6.4: Objekthierarchie

integriert werden können, wird die Flexibilität des Systems weiter gesteigert.

Die in Abbildung 6.4 rot gekennzeichnete GISComponent wurde zur Speicherung der Referenzen der verschiedenen Layer entwickelt. Da sie von der DaMaVisBaseObjectNodeImplementation abgeleitet wurde, implementiert sie die Schnittstelle DaMaVisBaseObject. Dies bildet die Grundlage für das Einfügen von Instanzen der GIS-Komponente in den Strukturbaum des Clients. Die einzelnen Layer werden als Söhne an den Knoten GIS-Komponente angehängt. Abbildung 6.5 zeigt eine solche GIS-Komponente. Sie enthält in diesem Fall 5 Layerreferenzen, die im linken Bereich des Fensters im Strukturbaum angezeigt werden.

6.3 Anwendung in der Stadtplanung

In der Stadtplanung öffnet sich ein immer breiteres Themenfeld in dem rechnerunterstützte Entscheidungshilfen zukünftig benötigt werden. In der heutigen Zeit werden noch viele Auswertungen in diesem Anwendungsbereich von Hand erledigt,

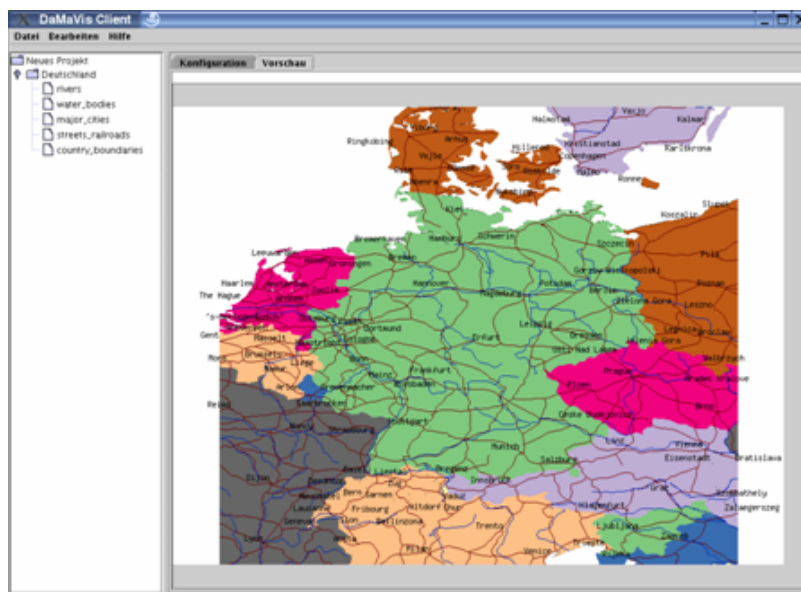


Abbildung 6.5: GIS Komponente mit 5 Layerreferenzen

da für Rechneranwendungen keine geeigneten Module auf dem Markt angeboten werden. Es existieren vielfach Insellösungen, die in getrennten Arbeitsschritten, nacheinander ausgeführt, zu einer Problemlösung führen.

Die Flexibilität des Datenmanagementsystems *DaMaViS* erlaubt es bei Bedarf neue Funktionalitäten in das System einzubinden und somit eine einzige Basis für eine Vielzahl von Entscheidungen zu erhalten. Im Folgenden werden anhand von den drei Beispielen *IKone*, GIS-Bericht und virtueller Rundgang die bereits implementierten Anwendungsfelder näher beschrieben.

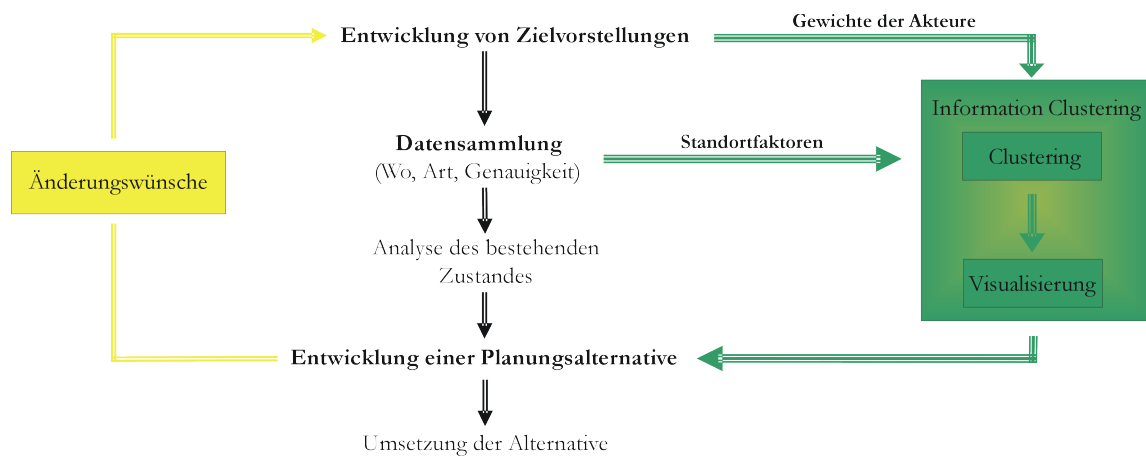
6.3.1 *IKone*

Mit der hier vorgestellten computergestützten Entscheidungshilfe *IKone* (Interpretation von Konversionsflächen) [HSRS06], [SHSR06b] wird die Möglichkeit geschaffen, die im Umnutzungsprozess der Konversionsflächen Involvierten schneller und zielorientierter zu einer Entscheidungsfindung zu führen.

IKone setzt die in dieser Arbeit vorgestellten theoretischen Überlegungen für das Anwendungsgebiet der Stadtplanung um.

Das Verfahren untergliedert sich in die folgenden drei Prozessabschnitte:

1. Profilerstellung der Fläche
2. Clustering

Abbildung 6.6: Prozessablauf bei *IKone*

3. Visualisierung

Während des Verfahrens werden die im Laufe der Bestandsaufnahme eines Planungsprozesses anfallenden Informationen (siehe auch Kapitel 3.2) gesammelt und mittels des in Kapitel 4 beschriebenen Clusteringverfahrens sichtweisenbezogen für die einzelnen Akteure ausgewertet. Das Ergebnis dieser beiden Prozesse ist eine Gebietsunterteilung der Konversionsfläche anhand der vorher ermittelten Standortfaktoren.

Die beteiligten Akteure erhalten somit eine direkte Rückkopplung, welche Gebietseinteilungen für sie von besonderem Vorteil sind. An den Clusteringprozess schließt sich die Visualisierung mit Verfahren aus dem Bereich der Informationsvisualisierung an, die schließlich die Interpretation der zu Grunde liegenden Daten ermöglicht (siehe hierzu auch Kapitel 5).

Profilerstellung

Ausgangspunkt für den gesamten in Abbildung 6.6 beschriebenen Prozessablauf ist eine Analyse des Konversionsgebietes. Hierbei werden die Standortfaktoren bestimmt, gefolgt von der Erhebung der Daten.

Diese Datenerhebung impliziert eine Definition der Konversionsfläche durch unabhängige Parameter. Abbildung 6.7 beschreibt den Ablauf der Parameteridentifikation bei Konversionsflächen. Aus den ermittelten Standortfaktoren ergeben sich eine Vielzahl objektiver und subjektiver Einflussfaktoren, die die weitere Nutzung der Fläche indizieren.

Die Standortfaktoren werden gegliedert nach Faktoren, die für alle Nutzungsarten von Interesse sind und solche, die lediglich in den drei Nutzungsarten Wohnen, Dienstleistung und Gewerbe benötigt werden.

Objektive Einflussfaktoren sind alle harten Standortfaktoren. Sie umfassen hauptsächlich

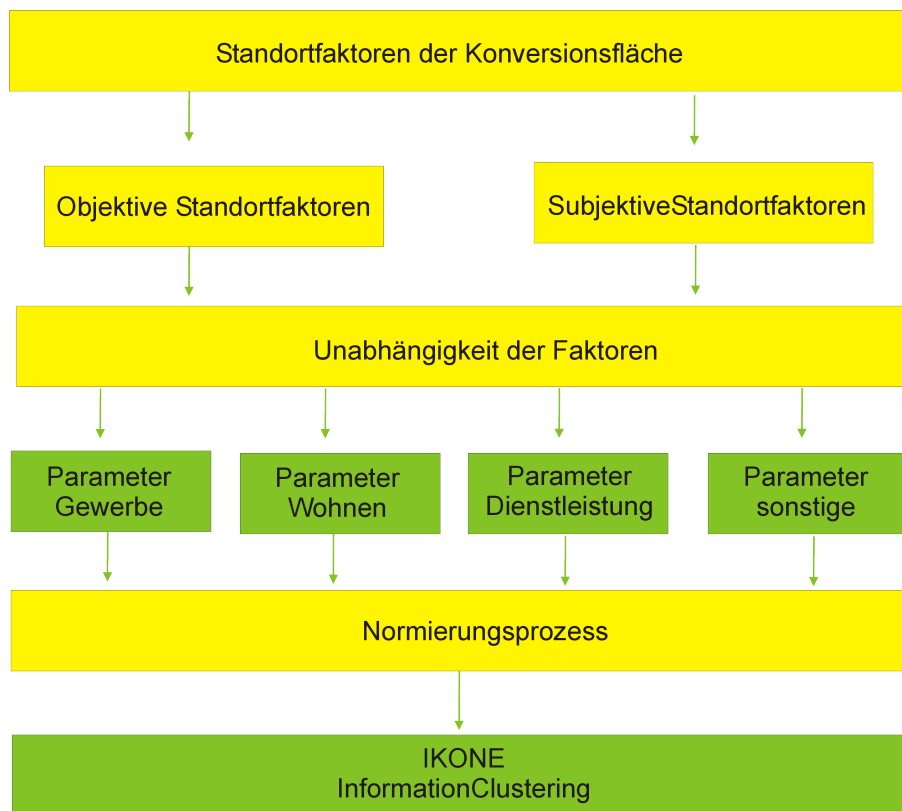


Abbildung 6.7: Parameteridentifikation bei Konversionsflächen

natürlich-technische und von Menschen geschaffene Faktoren. Hierzu zählen beispielsweise die Verfügbarkeit von Naherholungs- und Freizeiteinrichtungen, Erreichbarkeit von Arbeitsplätzen, Aufbau des Bodens, das Arbeitskräftepotential, das Bildungsangebot und die Nähe zu einem Flughafen. Bei diesen quantifizierbaren Daten der Flächen handelt es sich sowohl um statistische Daten wie auch um Messdaten, die gezielt an bestimmten Punkten vor Ort oder durch statistische Umfragen erhoben werden.

Darüber hinaus werden die subjektiven Einflussfaktoren untersucht. Hierzu zählen alle qualitativen Faktoren. Als Beispiel für subjektive Einflussfaktoren können die Lebensqualität, das Stadt- und Regionsimage, das Image des Betriebsstandortes etc. gesehen werden. Diese Parameter sind nicht direkt quantifizierbar, sondern werden im Zuge der Bestandsaufnahme ermittelt.

Die Gesamtheit der Einflussfaktoren ergibt eine Fülle voneinander abhängiger und unabhängiger Parameter. Für die Klassifizierung der Flächen ergibt sich die unbedingte Forderung nach der Unabhängigkeit der Parameter. Da die Parameter in verschiedenen Einheiten und Dimensionen vorliegen, erfolgt ein Normierungsprozess.

Hierdurch wird die Vergleichbarkeit der Parameter als Eingangsgrößen der Zielfunktion garantiert. Die unabhängigen Parameter werden der Zielfunktion (siehe Kapitel 3) zugeführt.

Clusteringprozess

Als Methode für das Clusteringverfahren werden verallgemeinerte Voronoidiagramme verwendet. Durch die Wahl dieses Verfahrens ist es möglich, die zahlreichen Parameter der Zielfunktion (siehe Kapitel 3.3.1) zuzuführen. Eine weitere nützliche Option der verallgemeinerten Voronoidiagramme bietet die Integration der Gewichte (Kapitel 4.2.2) in das Clusteringverfahren. Mit Hilfe der Gewichte ist es möglich die verschiedenen Sichten der Akteure auf die Konversionsfläche zu verwalten. Hierbei müssen die topologischen Eigenschaften der Zielfunktion berücksichtigt werden, da beispielsweise der Zusammenhang der Unterteilung eine besondere Anforderung bei der Konversion darstellt. Löcher in der Struktur sind spezielle Artefakte, die entweder unterdrückt werden müssen, oder aber toleriert werden können. Auf Grund dessen werden topologieerhaltende Distanzfunktionen verwendet, mit der die Gewichte additiv und multiplikativ verknüpfen werden. Im Anschluss an das Clustering erfolgt eine Visualisierung der Objekte mittels Techniken aus dem Gebiet der Informationsvisualisierung.

Visualisierungsprozess

Nachdem der Clusterprozess abgeschlossen ist, liegen die neu geordneten Daten in Form von mehrdimensionalen Objekten vor. Eine Interpretation der Ergebnisse an dieser Stelle ist jedoch mühsam und verwirrend, da die Objekte zu viele Informationen enthalten. Auf Grund dessen schließt sich eine Visualisierung an, für die Techniken aus dem Bereich der Informationsvisualisierung verwendet werden. Diese eignen sich besonders für eine Interpretation der abstrakten Objekte.

Die Grundidee hierbei ist, die zugrunde liegende Karte der Konversionsfläche mit der Visualisierung zu überlagern und somit einen direkten Bezug zwischen diesen Informationen herzustellen. Die einzelnen Parameter müssen hierbei durch die Visualisierung besonders repräsentiert werden.

Um diesen Anforderungen gerecht zu werden, werden zwei Visualisierungsmethoden verwendet, eine Icon-basierte und eine hierarchische Methode.

Die Icon-basierte Methode hat den Vorteil, dass die Resultate graphisch auf den Flächen dargestellt werden können. Somit lässt sie sich sehr leicht mit den Karten kombinieren, indem einfach verschiedene Layer übereinander gelegt werden. Mit den beiden Standardtechniken *stick-figure icon* und *velco-icon* [8] ist es jedoch nicht möglich eine größere Zahl von Parametern darzustellen. Deswegen wird das in Kapitel 5.2 beschriebene *Picklock-Icon* verwendet. Zusätzlich wird eine hierarchische Visualisierungstechnik eingesetzt. Der Vorteil dieser Methode ist, dass die Visualisierung nicht die Karte überdeckt. Der in Kapitel 5.3 vorgestellte *Spike-Tree* gehört

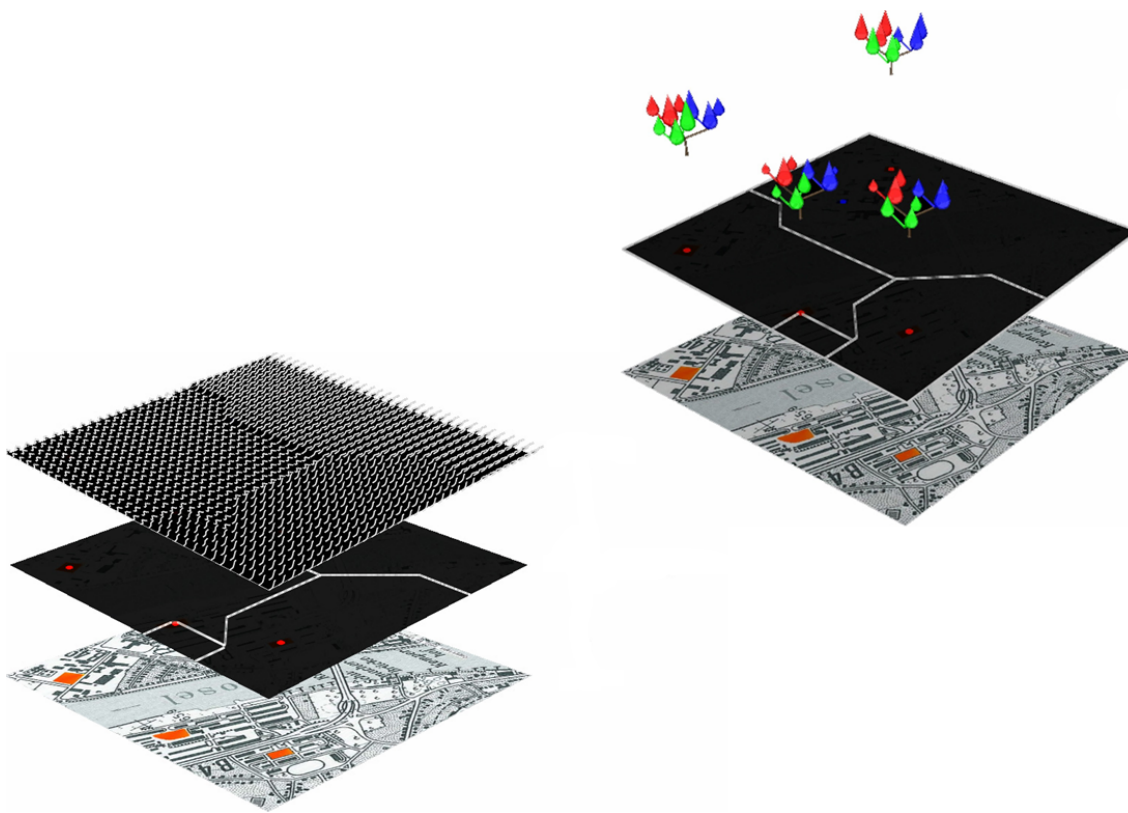


Abbildung 6.8: Picklock-Icon und Spike-Tree Visualisierung

zur Botanischen Informationsvisualisierung und ist punkt-orientiert. Der Spike-Tree ist eine Weiterentwicklung der Ansätze von Kleiberg, van de Wetering und van Wijk [9].

Zwei Beispiele dieser Visualisierungstechniken sind in Abbildung 6.8 zu sehen. Basis ist die zugrunde liegende Karte, darüber wird das Clusterergebnis gelegt, auf das die Visualisierung aufgesetzt wird.

6.3.2 GIS-Bericht

Basierend auf den in den Datenbanken gespeicherten Metadaten können verschiedene Berichte erstellt werden. In einem GIS Bericht werden, wie in einem herkömmlichen Geoinformationssystem, raumbezogene Daten mit Sachinformationen gekoppelt und diese Daten in Form einer Karte in 2D ausgegeben.

Zum Erstellen eines GIS Bericht muss der *DaMaViS* Client gestartet werden. Anschließend wird eine neue GIS Komponente angelegt. Innerhalb dieser Komponente werden alle benötigten Daten referenziert und innerhalb der Karte dargestellt. In

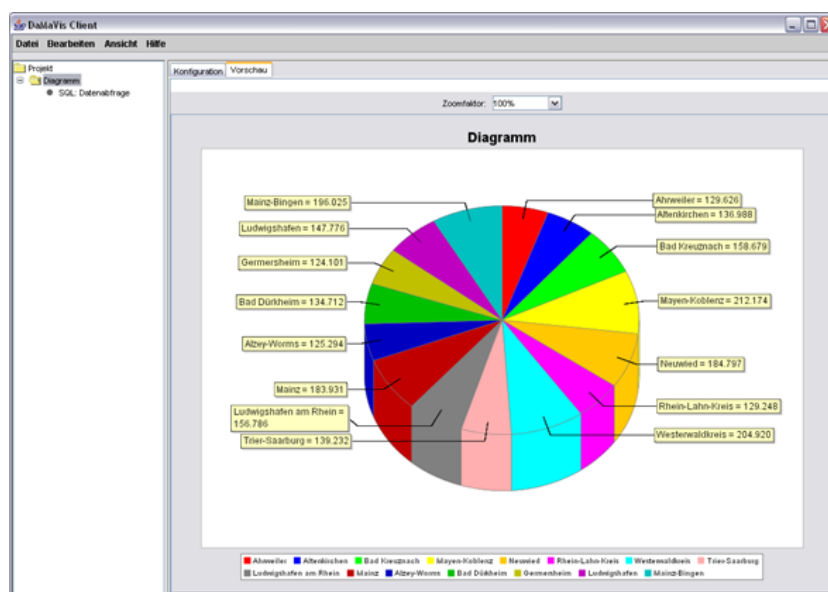


Abbildung 6.9: Auswertung in DaMaViS

dieser Karte können flexibel und bedarfsgerecht die Informationen geladen und bearbeitet werden, die im Einzelfall benötigt werden. Der GIS Bericht enthält dann Informationen, wie sie auch ein herkömmliches GIS-System erstellen kann.

Nach Erstellen des GIS Berichtes werden die zusammengestellten Daten an den Server übergeben. Auf dem Server können verschiedene Ausgabeformate für diesen GIS Bericht erstellt werden. Mittels des MapServers wird beispielsweise ein Rasterbild erstellt, welches sich in einem weiteren Schritt problemlos in einen zu erstellenden Bericht oder ein neu zu erstellendes Dokument einfügen lässt.

Abbildung 6.10 zeigt ein Beispiel eines solchen GIS Berichtes für den Bereich Kaiserslautern und Umgebung. Im Preview-Fenster wird als Basis eine Karte des Großraumes angezeigt, in die Zusatzinformationen wie Städtenamen oder Flüsse eingezeichnet werden. Ein weiterer Layer gibt eine farbkodierte Übersicht über die verschiedenen Länder. Mit diesem Bild wird aufgezeigt, wie die geometrischen Daten wie Position und Lage mit den Metadaten wie Text und Farbe verknüpft werden.

6.3.3 Virtueller Rundgang

Ein weiteres Anwendungsfeld im Bereich der Stadtplanung ist die Erstellung virtueller Rundgänge. Basierend auf einem digitalen Geländemodell kann ein virtueller Rundgang durch ein für den Anwender interessantes Stadtgebiet erstellt werden. Dadurch ist es möglich, einen Eindruck des Stadtbildes zu vermitteln. Darüber hin-

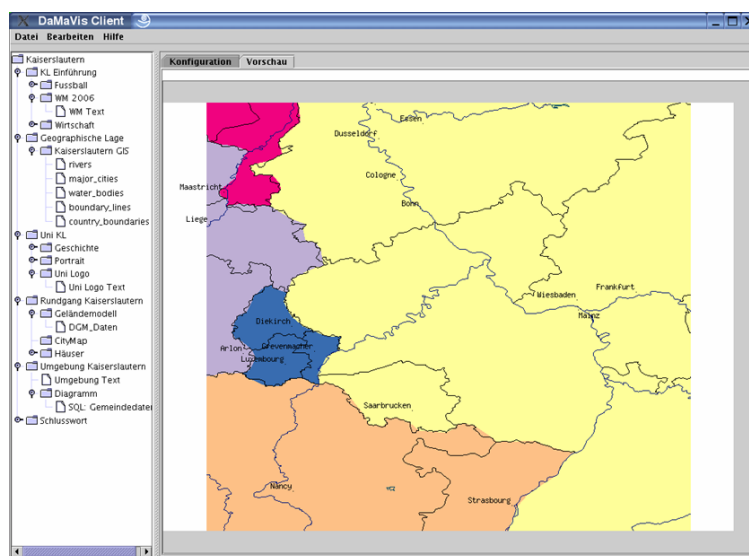


Abbildung 6.10: GIS Report über Kaiserslautern und Umgebung

aus können Veränderungen in der Bausubstanz visuell dargestellt werden und so zur Entscheidungsfindung, ob ein Bauvorhaben verwirklicht wird, beitragen.

Im Folgenden wird die Erstellung eines virtuellen Stadtrundgangs unter Verwendung des Datenmanagementsystems beschrieben. Der *DaMaVis* Client wird gestartet und eine neue Komponente *walkthrough* hinzugefügt. Automatisch werden die erforderlichen Knoten für den virtuellen Rundgang im Strukturbaum angelegt. [HMS⁺06]

Hierbei enthält der erste Knoten das digitale Geländemodell, wobei das relevante digitale Geländemodell durch den Nutzer ausgewählt werden muss. Alle für die Modellierung des Geländes benötigten Höhendaten im digitalen Geländemodell sind in Gauss-Krüger-Koordinaten angegeben. Innerhalb dieses Knotens erfolgt die Modellierung des Geländes und der Gebäudehöhen.

Der zweite Knoten enthält einen Stadtplan mit den genauen Grundrissinformationen aller Gebäude. Die Datenquelle für den Stadtplan können XML oder ESRI-Daten sein. Innerhalb des Stadtplans erfolgt die Selektion der Gebäude, die Teil des virtuellen Rundgangs werden sollen. Abbildung 6.11 zeigt einen Stadtplan von Kaiserslautern im Preview-Bereich.

Der dritte Knoten enthält die weitergehenden Informationen der selektierten Gebäude. In diesem Knoten sind die Unterknoten angelegt, die die Informationen der einzelnen Gebäude repräsentieren. Hierbei handelt es sich z.B. um die Lage, Adresse, die Höhe des Gebäudes und ähnliches.

Zum Erreichen einer photorealistischen Darstellung der Gebäude werden zusätzlich

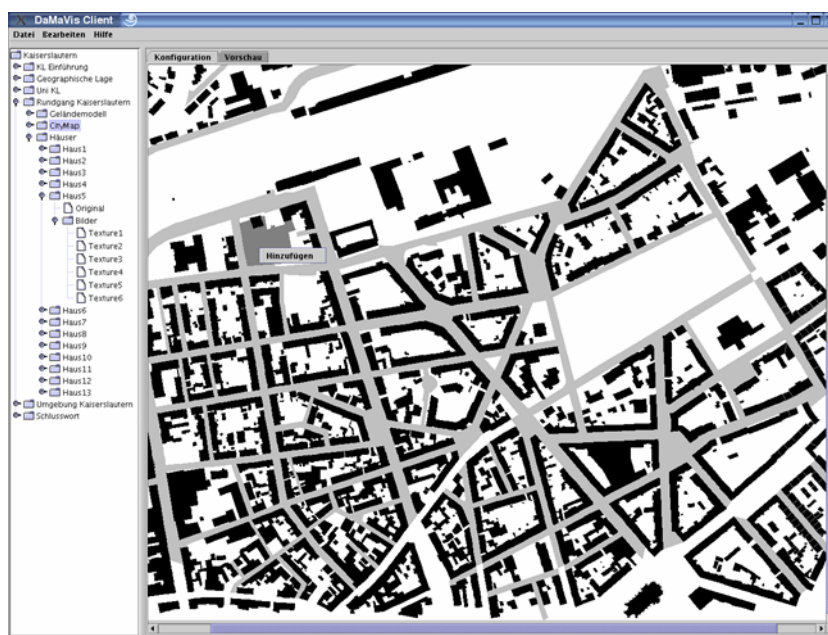


Abbildung 6.11: Stadtplan von Kaiserslautern

Texturen für die Gebäude benötigt. Diese werden aus hochauflösenden Bildern extrahiert. Zum Einpassen der Texturen an der richtigen Stelle auf dem Gebäude werden die Geometrieinformationen im Grundriss des Bauwerks um eine Nummerierung der Kanten ergänzt, was in Abbildung 6.12 exemplarisch für ein Gebäude gezeigt wird. Jeder Kantenummer wird eine Textur zugeordnet und diese damit verknüpft.

Für das Stadtbild interessante Gebäude wie zum Beispiel der Bahnhof oder die Post, können als *point of interest* gekennzeichnet werden. Durch diese Kennzeichnung ergibt sich beim später generierten Rundgang ein Pfeil der auf dieses Gebäude hinweist, eine Orientierung ermöglicht und alle Zusatzinformationen zu diesem *point of interest* liefert. Im Anschluss an die Selektion der relevanten Gebäude wird das 3D-Modell generiert. Alle Daten für das 3D-Modell werden als Open Scene Graph files abgelegt.

Zur Visualisierung des virtuellen Rundgangs enthält *DaMaVis* einen externen Viewer, der auf dem OSG (OpenSceneGraph) graphics toolkit basiert. Hierfür wurde der Original OSG-Viewer zusätzlich mit der Fähigkeit ausgestattet, alle *points of interest* anzuzeigen. Hierbei weist ein Pfeil auf einen möglichen interessanten Punkt in einer bestimmten Richtung hin, was eine einfache Navigation innerhalb des Rundgangs hin zu möglichen Zielen zulässt.

Abbildungsverzeichnis

2.1	Clusterstrukturen mit verschiedenen Charakteristika [ES00]	7
2.2	Konturlinie der Umgebung $U_\delta(0)$	10
2.3	Einheitsumgebung $U_1(0)$ der Euklidischen Metrik	10
2.4	Einheitsumgebung $U_1(0)$ der L_1 -Metrik	10
2.5	Einheitsumgebung $U_1(0)$ der Maximumnorm	11
2.6	Minkowski-Konturen	11
2.7	Joseph Minard: Karte des Rußlandfeldzuges Napoleons von 1812-1813	12
2.8	Scatterplot-Matrizen und Parallele Koordinaten	16
2.9	Retinale Eigenschaften in der Glyph-basierten Visualisierung	18
2.10	Chernoff Faces und Star Glyphs	19
2.11	Recursive Pattern und Circle Segment	20
2.12	Dimensional Stacking und Cone Tree	21
2.13	Webvisualisierung durch hyperbolische Darstellung	22
2.14	Flug durch den Informationsraum von Kaufvideokassetten mit Perspecta View	23
2.15	Punktwolkendarstellung und thematische Karte einer Dokumentensammlung durch IN-SPIRE	24
2.16	System der Raumplanung	27
2.17	Konversion in Kaiserslautern: PRE Park [Gmb02]	29
2.18	Ausprägungen von GIS	31
3.1	Ablaufschema einer Planung	38
3.2	Verkehrsplanungsprozess [Kor97]	39
3.3	Planungsschritte für die Verkehrsberuhigung [Kor97]	40
3.4	Ansatzpunkt des Information Clustering im Modell des Planungslaufs	42
3.5	Profilerstellung einer Konversionsfläche	43
3.6	Akteure der Projektentwicklung	44
3.7	Hauptakteure und deren Interessen	45
3.8	Parameterauswahl zum Entwurf des Nutzungskonzeptes	48
3.9	Integration der Gewichte und Akteure	49

3.10	Information Clustering im Planungsablauf	51
3.11	Visualisierung mit dem Velcro-Icon [RP88]	53
4.1	Einfaches Voronoidiagramm	56
4.2	Voronozelle aus Schnitt von Halbräumen	57
4.3	Gewichtetes Voronoidiagramm mit leerer Voronoiregion	59
4.4	Degeneriertes Voronoidiagramm	60
4.5	Bisektor der Manhattan-Metrik	62
4.6	Ist ein Einheitskreis C streng konvex, so ist der Bisektor von zwei Punkten immer ein Weg.	62
4.7	Wenn $b''_1, b''_2 \in \partial C_p$ auf der C_p zugewandten Seite und mit b'_1 auf einer zu pq parallelen Geraden liegen, enthält ∂C_p das Liniensegment b''_1, b''_2	63
4.8	konvexe Distanzfunktion	64
4.9	Additiv gewichtetes euklidisches Voronoidiagramm mit gewichteten Kreisen	66
4.10	Multiplikativ gewichtetes euklidisches Voronoidiagramm	67
4.11	Zusammengesetzt gewichtetes Voronoidiagramm	68
4.12	Konturlinie eines Kardioiden	69
4.13	Interaktive Modifikation einer Distanzfunktion	70
4.14	Generation eines beliebig gewichteten Voronoidiagrammes	71
4.15	Schnelle Modellierung eines Voronoidiagrammes mit einer kontrollierbaren Distanzfunktion(a), Resultate für ungewichtete Fälle (b), additiv (c), multiplikativ (d+f) and zusammengesetzt (e) gewichtete Fälle. Eine topologisch unzusammenhängendes Gebiet ist in (f) hervorgehoben	72
4.16	Konstruktion einer generellen Distanzfunktion	73
4.17	Schrittweises Wachstum.	75
4.18	Additives Wachstum zu den Zeitpunkten t_0, t_5 und t_{10}	76
4.19	Multiplikatives Wachstum zu den Zeitpunkten t_0, t_5 und t_{10}	77
5.1	Stick-Figure-Icon-Familie	82
5.2	Visualisierung multispektraler Bilddaten des NOAA-7 AVHRR Wettersatelliten unter Verwendung des Stick-Figure-Icons	83
5.3	Eigenschaften des Velcro-Icons	84
5.4	Visualisierung eines Satellitenbildes unter Verwendung des Velcro-Icons [EG95]	85
5.5	Das <i>Picklock-Icon</i>	86
5.6	Ausschnitt einer topographischen Karte der Stadt Koblenz	87
5.7	Sichtweisen auf die Daten	88
5.8	Vorgehensweise bei der Berechnung und Darstellung von Karte, Voronoidiagramm und <i>Picklock-Icons</i>	89

5.9	In die Textur geladene Karte	90
5.10	Bestimmung der Gewichte	91
5.11	Visualisierung multivariater Daten unter Verwendung der Manhattan-Metrik und des <i>Picklock-Icons</i>	92
5.12	Node-Link-Diagramm und entsprechendes Ader-Modell	93
5.13	Verzeichnisbaum	94
5.14	Konstruktion des Phi-Balls	95
5.15	Berechnung der Kegelparameter	96
5.16	Modell mit Kontraktion, Extrusion und Phi-Ball	97
5.17	Parameter und Berechnung eines <i>Spike</i> -Kegels	98
5.18	Bestimmung der Gewichte eines Voronoiortes	99
5.19	Informationsvisualisierung mutlivariater Daten mittels <i>Spike-Tree</i>	100
5.20	Basisdarstellung des Spike-Tree	101
5.21	Vorgehensweise bei der Visualisierung von Karte, Voronoidiagramm und <i>Spike-Tree</i>	102
6.1	Client-Server-Modell	106
6.2	Datenbankadministrationstool <i>DaVEa</i>	107
6.3	DaMaViS Client	108
6.4	Objekthierarchie	111
6.5	GIS Komponente mit 5 Layerreferenzen	112
6.6	Prozessablauf bei <i>IKone</i>	113
6.7	Parameteridentifikation bei Konversionsflächen	114
6.8	Picklock-Icon und Spike-Tree Visualisierung	116
6.9	Auswertung in DaMaViS	117
6.10	GIS Report über Kaiserslautern und Umgebung	118
6.11	Stadtplan von Kaiserslautern	119
6.12	Gebäudegrundriss mit nummerierten Kanten	120

Literaturverzeichnis

- [ABA00] R. Aliev, K.-W. Bonfig, and F. Aliev. *Soft Computing: Eine grundlegende Einführung*. Berlin, 2000.
- [AC91] B. Alpern and L. Carter. *Hyperbox*. IEEE Visualization 1991, San Diego, CA, USA, 1991.
- [AK96] M. Ankerst and D.A. Keim. *Circle Segments: A Technique for Visually Exploring Large Multidimensional Data Sets*. Visualization 1996, Hot Topic Session, 1996.
- [Bac02] J. Bacher. *Clusteranalyse*. Oldenburg, 2002.
- [Bed90] J. Beddow. *Shape Coding of Multidimensional Data on a Microcomputer Display*, volume 1990. IEEE Visualization, San Francisco, CA, USA, 1990.
- [Ber83] J. Bertin. *Semiology of Graphics*. The University of Wisconsin Press, Madison, WI, USA, 1983.
- [Bil99] R. Bill. *Grundlagen der Geoinformationssysteme*. Wichmann Verlag, 1999.
- [BM96] B. Breckling and F. Müller. *Modellbildung und Simulation im Projektzentrum Ökosystemforschung. Ökosysteme und ökologische Prozesse im Bereich der Bornhöveder Seenkette*. EcoSys Beiträge zur Ökosystemforschung, Kiel, 1996.
- [Bra99] W. Braam. *Stadtplanung, Aufgabenbereiche, Planungsmethodik, Rechtsgrundlagen*. Werner Verlag, 1999.
- [Cas59] J.W.S. Cassels. *An Introduction to the Geometry of Numbers*. Springer Verlag, Berlin, 1959.

- [CCKT83] J.M. Chambers, W.S. Cleveland, B. Kleiner, and P.A. Tukey. *Graphical Methods for Data Analysis*. Wadsworth Press, Belmont, CA, USA, 1983.
- [CD85] L.P. Chew and R.L. Drysdale. *Voronoi diagrams based on convex distance functions*. 1st Annu. ACM Sympos. Comput. Geom., 1985.
- [Che71] H. Chernoff. *The Use of Faces to Represent Points in n-Dimensional Space Graphically*, volume 71. Department of Statistics, Stanford University, CA, USA, 1971.
- [Cle93] W.S. Cleveland. *Visualizing Data*. Hobart Press, Summit, NJ, USA, 1993.
- [Cor71] R.M. Cormack. *A Review of Classification*, volume 134. Journal of the Royal Statistical Society, 1971.
- [Däß99] R. Däßler. *Informationsvisualisierung: Stand, Kritik und Perspektiven*. Fachhochschule Potsdam, 1999.
- [dB72] C. de Boor. *A Practical Guide to Splines*. Springer Verlag, Berlin, 1972.
- [dBvKOS00] M. de Berg, M. van Kreveld, M. Overmars, and O. Schwarzkopf. *Computational Geometry: Algorithms and Applications*. Springer Verlag, 2000.
- [dL02] N. de Lange. *Geoinformatik in Theorie und Praxis*. Springer Verlag, Berlin, 2002.
- [EG95] R.F. Erbacher and G. Grinstein. *Issues in the Development of 3D Icons*. Visualization in Scientific Computing, 1995.
- [ES86] H. Edelsbrunner and R. Seidel. *Voronoi diagrams and arrangements*. Discrete Comput. Geom., 1986.
- [ES00] M. Ester and J. Sander. *Knowledge discovery in databases: Techniken und Anwendungen*. Springer, Berlin, 2000.
- [Eve01] B.S. Everitt. *Cluster Analysis*. 4nd Edition, London, 2001.
- [Far88] G. Farin. *Curves and Surfaces for Computer Aides Geometric Design: A Practical Guide*. Academic Press, 1988.
- [FB94] G.W. Furnas and A. Buja. *Prosection Views: Dimensional Inference Through Sections and Projections*, volume 3(4). J. Comp. Stat., 1994.

- [FPF88] K.M. Fairchild, S.E. Poltrock, and G.W. Furnas. *SemNet: Three-dimensional graphic representations of large knowledge bases*. Cognitive Science and its applications to human computer interaction, 1988.
- [fSuV98] Forschungsgesellschaft für Straßen-und Verkehrswesen. *Begriffsbestimmungen -Teil: Straßenplanung und Straßenverkehrstechnik*. FGSV-Verlag, 1998.
- [GE97] N. Gershon and S. Eick. *Information Visualization*. IEEE Computer Graphics and Applications, 1997.
- [Gmb02] PRE Regionalentwicklung GmbH. *PRE Park in Kaiserslautern*. <http://www.pre-park.de>, 2002.
- [Gor81] A.D. Gordon. *Classification*. London-New York, 1981.
- [GR91] S.K. Card G.G. Robertson, J.D. Mackinlay. *Cone Trees: Animated 3D Visualizations of Hierarchical Information*. Human Factors in Computing Systems CHI 1991 Conference, New Orleans, LA, USA, 1991.
- [GRS00] G. Görz, C.R. Rollinger, and J. Schneeberger. *Handbuch der Künstlichen Intelligenz*. Oldenbourg Verlag, 2000.
- [GS94] H.D. Kühne G. Steierwald. *Stadtverkehrsplanung. Grundlagen, Methoden, Ziele*. Springer Verlag, 1994.
- [HCK⁺99] K.E. Hoff, T. Culver, J. Keyser, M. Lin, and D. Manocha. *Fast computation of generalized voronoi diagrams using graphics hardware*. Siggraph99, 1999.
- [Her00] C. Herzog. *Ein Methodenpaket zur Modellierung und integrativen Bewertung von Raumnutzungsänderungen und dem Einsatz von Fuzzy-Expertensystemen*, volume 3. Zeitschrift für Angewandte Umweltforschung, Berlin, 2000.
- [HG02] P.E. Hoffman and G.G. Grinstein. *A Survey of Visualizations for High-Dimensional Data Mining*. Morgan Kaufman Publishers, 2002.
- [Hil75] R. Hillebrecht. *Städtebau als Herausforderung-Ausgewählte Schriften und Vorträge*. Neue Schriften des Deutschen Städtetages, Heft 30, Köln, 1975.

- [HK01] J. Han and M. Kamber. *Data Mining, Concepts and Techniques*. Morgan Kaufmann, 2001.
- [HMS⁺05] H. Hagen, M. Münchhofen, I. Scheler, M. Ruby, F. Michel, and M. Wadlé. *DaMaViS - Data Management and Visualization System*. WSCG, Plzen, Czech Republic, 2005.
- [HMS⁺06] H. Hagen, M. Münchhofen, I. Scheler, M. Ruby, and F. Michel. *Interaktive Rekonstruktion von Gebäudemodellen für virtuelle Rundgänge*. AGIT 2005, Salzburg, Österreich, 2006.
- [Hof99] P. E. Hoffman. *Table Visualizations: A Formal Modal and its Applications*. Computer Science Department, University of Massachusetts Lowell, 1999.
- [Hol94] M. Holton. *Strands, gravity and botanical tree imagery*, volume 13(1). Computer Graphics Forum, 1994.
- [HRS05] H. Hagen, M. Ruby, and I. Scheler. *Information Clustering in the context of Urban Planning*. International Symposium on Generalization of Information, Berlin, 2005.
- [HRSS05] H. Hagen, M. Ruby, I. Scheler, and M. Schneider. *Topology-preserving distance functions for tessellations*. IASTED International Conference on Visualization, Imaging and Image Processing, 2005.
- [HSM⁺04] H. Hagen, G. Steinebach, M. Münchhofen, M. Ruby, I. Scheler, and M. Wadlé. *DaMI - Data Management for Multimedial Information Systems*. IEEE Visualization, 2004.
- [HSM⁺05] H. Hagen, G. Steinebach, M. Muenchhofen, M. Ruby, I. Scheler, M. Wadlé, and F. Michel. *Datenmanagementsystem fuer die Stadtplanung*. CORP – Competence Center of Urban and Regional Planning, Wien, 2005.
- [HSRS06] H. Hagen, G. Steinebach, M. Ruby, and I. Scheler. *IKone - Computergestützte Auswertung von Konversionsflächen mithilfe von Voronoi Diagrammen*. CORP – Competence Center of Urban and Regional Planning, Wien, 2006.
- [HSS⁺05] H. Hagen, M. Schneider, M. Schlemmer, M. Ruby, and I. Scheler. *Fast Voronoi Modeling*. Dagstuhl Seminar on Geometric Modelling, 2005.

- [ID90] A. Inselberg and B. Dimsdale. *Parallel Coordinates: A Tool for Visualizing Multi-Dimensional Geometry*. IEEE Visualization, 1990.
- [Ise99] B. Isenhöfer. *Strategisches Management von Projektentwicklungsunternehmen*, volume 4. Schriften zur Immobilienökonomie, 1999.
- [JD88] A.K. Jain and R.C. Dubes. *Algorithms for Clustering Data*. Prentice Hall, New Jersey, 1988.
- [Joh93] B. Johnson. *Visualizing Hierarchical and Categorical Data*. Department of Computer Science, University of Maryland, 1993.
- [JvW93] R. van Liere J.J. van Wijk. *HyperSclice*. IEEE Visualizations, 1993.
- [KdL96] K. Kanzler and N. de Lange. *Entscheidungsunterstützung in der Regionalplanung; Bewertung von Ansiedlungsvorhaben des großflächigen Einzelhandels*. 17. AGIT-Symposium, 1996.
- [Kei97] D.A. Keim. *Visual Techniques for Exploring Databases*. Institut für Informatik, Martin-Luther-Universität Halle-Wittenberg, 1997.
- [KKA95] D.A. Keim, H.P. Kriegel, and M. Ankerst. *Recursive Pattern: A Technique for Visualizing Very Large Amounts of Data*. Visualization 1995, 1995.
- [Kor97] M. Korda. *Städtebau*. Teubner Verlag, 1997.
- [KvdWvW01] E. Kleiberg, H. van de Wetering, and J.J. van Wijk. *Botanical Visualization of Huge Hierarchies*. IEEE Symposium on Information Visualization, 2001.
- [Lab06] Pacific Northwest National Laboratory. *In-Spire Visual Document Analysis*. <http://in-spire.pnl.gov/index.html>, 2006.
- [LD99] B. Lintermann and O. Deussen. *Interactive modeling of plants*, volume 19(1). IEEE Computer Graphics and Applications, 1999.
- [Leu97] Y. Leung. *Intelligent Spatial Decision Support Systems*. Springer Verlag, Berlin, 1997.
- [Lev91] H. Levkowitz. *Color Icons: Merging Color and texture Perception for Integrated Visualization of Multiple Parameters*. IEEE Visualization 1991, 1991.
- [Lin99] W. Lindner. *Geo-Informationssysteme*. Springer Verlag, Berlin, 1999.

- [LVe04] LVermGeo. *CD-ROM Serie Landschaft im Wandel*. Landesamt für Vermessung und Geobasisinformation Rheinland-Pfalz, Koblenz, 2004.
- [LWW90] J. Leblanc, M.O. Ward, and N. Wittels. *Exploring N-Dimensional Databases*. IEEE Visualization 1990, 1990.
- [Mat02] H. Mattes. *Voronoi-Daigramme und deren Berechnung*. Institut für Informatik III, Universität Bonn, 2002.
- [MdIufS02] Rheinland-Pfalz Ministerium des Innern und für Sport. *Arbeitshilfe zu den rechtlichen, planerischen und finanziellen Aspekten der Konversion militärischer Liegenschaften*. www.konversion.com, 2002.
- [Mun95] T. Munzer. *Visualizing the Structure of the World Wide Web in 3D Hyperbolic Space*. The Geometry Center, University of Minnesota, 1995.
- [OBSC99] A. Okabe, B. Boots, K. Sugihara, and S. Nok Chiu. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. John Wiley and sons, LTD, 1999.
- [Opi80] O. Opitz. *Numerische Taxonomie*. Stuttgart-New York, 1980.
- [Pup99] F. Puppe. *XPS-99: Knowledge-Based Systems Survey and Future Directions*. 5th biannual German Conference on Knowledge-Based Systems, 1999.
- [RP88] G.G. Grinstein R.M. Pickett. *Iconographic Displays for Visualizing Multidimensional Data*. Procedure of IEEE Conference on Systems, Man and Cybernetics, IEEE Press, Piscataway (NJ), 1988.
- [Sch04a] G. Scheuermann. *Visualisierungssysteme in Naturwissenschaft und Technik*. AG Graphische Datenverarbeitung, Technische Universität Kaiserslautern, 2004.
- [Sch04b] M. Schneider. Topologieerhaltende Voronoi-Algorithmien. Master's thesis, University of Kaiserslautern, 2004.
- [SHaMR06] G. Steinebach, H. Hagen, and I. Scheler ands M. Ruby. *Visualisierung des Planungsablaufes am Beispiel von Konversionsflächen*. AGIT 2006, Salzburg, Österreich, 2006.
- [Shn92] B. Shneiderman. *Tree Visualization with Treemaps: A 2D Space-Filling Approach*, volume 11/2. ACM Transaction on Graphics, 1992.

- [Shn96] B. Shneiderman. *The Eyes Have It: A Task by Data-type Taxonomy for Information Visualization*. Proceedings of Visual Languages, IEEE Computer Science Press, Los Alamitos, CA, USA, 1996.
- [SHSR06a] G. Steinebach, H. Hagen, I. Scheler, and M. Ruby. *Generic approach to develop planning alternatives*. Brownfields, Boston, to be published, 2006.
- [SHSR06b] G. Steinebach, H. Hagen, I. Scheler, and M. Ruby. *ICONE - computer supported evaluation of military conversion areas using Voronoi Diagrams*. WPSC 06, World Planning Schools Congress, Mexico City, 2006.
- [SJ97] G. Steinebach and A. Jacob. *Konversion - Stadtplanung auf Militärflächen*. Bundesministerium für Raumordnung, Bauwesen und Städtebau, 1997.
- [SM06] G. Steinebach and P. Müller. *Dynamisierung und Flexibilisierung von Planverfahren der Stadtplanung durch Informations- und Kommunikationssysteme*. Schriften zur Stadtplanung, Band 4, 2006.
- [Tuk04] J. Tukey. *Quotations by John Tukey*. www-history.mcs.st-andrews.ac.uk/Mathematicians/Tukey.html, 2004.
- [Wad04a] M. Wadlé. *Konzeptionelle OpenGIS-Integration in Visualisierungssysteme*. University of Kaiserslautern, 2004.
- [Wad04b] M. Wadlé. *Voronoidiagramme*. Master's thesis, University of Kaiserslautern, 2004.
- [War99] C. Ware. *Information Visualization - Perception for Design*. Morgan Kaufmann Publishers, San Francisco, USA, 1999.
- [Web94] WebOOGL. www.geom.uiuc.edu/software/weboogl/. Integrating 3D graphics and the Web, 1994.
- [Wik06] Free Encyclopedia Wikipedia. *Information Visualization*. <http://en.wikipedia.org/wiki/Information-visualization>, 2006.
- [Wri95] W. Wright. *Information Animation in the Capital Markets*. Int. Symp. On Information Visualization 1995, Atlanta, GA, USA, 1995.

Lebenslauf

Persönliche Angaben

Name: Maja Ruby
Anschrift: Fliederweg 42
67657 Kaiserslautern
Kontakt: E-mail: ruby@rhrk.uni-kl.de
Geburtsdatum/-ort: 14. August 1968/Hamburg
Familienstand: verheiratet, zwei Kinder
Staatsangehörigkeit: deutsch

Berufstätigkeit

Seit 04/2001 Researcherin am Deutschen Forschungszentrum für künstliche Intelligenz
09/1997-06/2005 Wissenschaftliche Mitarbeiterin an der Technischen Universität Kaiserslautern
Seit 03/2005 Kollegiatin des Internationalen Graduiertenkollegs „Visualisierung großer unstrukturierter Datensätze“

Hochschulstudium

10/1988-08/1997 Diplom-Informatik an der Technischen Universität Kaiserslautern

Schulausbildung

1976-1988 Grundschule und Staatl. Gymnasium Lohbrügge, Hamburg