
Geodätische Schriftenreihe
der Technischen Universität Braunschweig Nr. 27

Methodische Ansätze zur unscharfen Mustererkennung bei Deformationsmessergebnissen

von Dipl.-Ing. Rainer Fletling
aus Alfeld (Leine)

Herausgeber: Institut für Geodäsie und Photogrammetrie
der Technischen Universität Braunschweig

ISBN 3-926146-22-2

Von der Fakultät Architektur, Bauingenieurwesen und Umweltwissenschaften
der Technischen Universität Carolo-Wilhelmina zu Braunschweig
zur Erlangung des Grades eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Dissertation

Eingereicht am 16. Juni 2010
Disputation am 10. Dezember 2010

Berichterstatter: Univ.-Prof. Dr.-Ing. habil. Wolfgang Niemeier
Univ.-Prof. Dr.-Ing. habil. Thomas Wunderlich

Allen, die am Zustandekommen dieser Arbeit beteiligt waren, sei recht herzlich gedankt.

Summary

A fuzzy clustering method from the field of pattern recognition is adopted to tackle two problems representative for the application area of geodetic deformation measurements.

The selected problems address (a) the identification of point groups with similar movement patterns in a two-dimensional control network, and (b) the automatic classification of settlement time series.

For both examples, the manual analysis of graphical representations of the data by an expert is still predominantly used in practice at the present time. However, with an increasing number of object points the need for an automated interpretation of the data quickly arises. As a particular example, monitoring systems with automatic alarm should be able to function without human staff.

In the work at hand, automatic clustering methods are used that, based on formal criteria, classify data into homogeneous classes.

The commonly used “hard“ clustering methods assign an object to exactly one particular class. However, such a strict classification is not appropriate in many cases. Consider, for instance, a situation in which only uncertain or imprecise knowledge about the objects is available. Also, objects that are located in the immediate vicinity of a class boundary can often be assigned similarly well to both classes on either side of the boundary.

Distinctive to the work at hand is the fuzzy clustering method employed that does not assign a data point to a single unique class but to all classes, computing gradually varying degrees of membership of the point to each of the classes. If c clusters are present at the end of the clustering process, the clustering result for every object will be a vector of c degrees of membership, as opposed to a unique assignment. Viewing a problem with fuzzy sets corresponds much better to the real world and the human notion thereof. For example, we cannot only see the two colours black and white, but also arbitrarily many intermediate shades of grey.

The presented examples demonstrate that fuzzy clustering methods can be applied to geodetic deformation analysis, yielding good results. It is hence shown that for the selected problems the data analysis task can be largely automated by the fuzzy approach.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Einführung in die Thematik	1
1.2	Motivation	4
2	Vom Muster zur Mustererkennung	5
2.1	Allgemeines zur Mustererkennung	5
2.2	Muster	9
2.3	Merkmale	12
2.3.1	Merkmalskalen	14
2.3.2	Merkmalsauswahl	15
2.3.3	Merkmalsvorbehandlung	16
2.4	Klassifikation	18
2.4.1	Ähnlichkeitsmaße	19
2.4.2	Clusterverfahren	25
2.4.3	Optimale Clusteranzahl	29
3	Einführung in die Theorie unscharfer Mengen	31
3.1	Grundlegendes zur Theorie unscharfer Mengen	31
3.2	Linguistische Variablen	34
3.3	Modellierung von Zugehörigkeitsfunktionen	35
3.4	Operationen mit unscharfen Mengen	37
3.5	Mehrdimensionale Zugehörigkeitsfunktionen	40
4	Fuzzy-Clusterverfahren	42
4.1	Vorbemerkungen zur traditionellen Behandlung von Unsicherheiten in der Geodäsie	42
4.2	Vorbemerkungen zur unscharfen Klassifikation	43
4.3	Der Fuzzy-c-Means Algorithmus	45
4.4	Modifizierte Fuzzy-c-Means Algorithmen	47
4.4.1	Der Gustafson-Kessel Algorithmus	47
4.4.2	Der Gath-Geva Algorithmus	48
4.5	Possibilistische Varianten	50
4.6	Gütekriterien für unscharfe Clusterergebnisse	53
4.6.1	Globale Gütekriterien	53
4.6.2	Lokales Gütekriterium	57

5	Die Methode der Fuzzy Pattern Klassifikation	58
5.1	Beschreibung der Unschärfe	59
5.1.1	Beschreibung der Unschärfe von Objekten im Merkmalsraum	61
5.1.2	Beschreibung der Unschärfe von Klassen im Merkmalsraum	62
5.2	Aggregation der unscharfen Objekte	63
5.2.1	Transformation der Objekte einer Klasse vom Merkmalsraum in einen Klassenraum	63
5.2.2	Projektion der elementaren Zugehörigkeitsfunktionen auf die Achsen des Klas- senraumes	64
5.2.3	Berechnung der Parameter der globalen Zugehörigkeitsfunktion	66
5.3	Neueste Entwicklungen	69
6	Übertragung auf Deformationsuntersuchungen	71
6.1	Vorbemerkungen	71
6.2	Auffinden von Punktgruppen in einem Überwachungsnetz	74
6.2.1	Testnetz Delft	74
6.2.2	Isländisches Grundlagennetz ISNET	85
6.2.3	Zusammenfassung Beispiele Testnetz Delft und ISNET	104
6.3	Automatische Klassifizierung von Setzungsmessungen	107
6.3.1	Verhalten im Zeitbereich	111
6.3.2	Bemerkungen zu dem Verfahren	112
7	Zusammenfassung	114
8	Literaturverzeichnis	117
A	Tabellen	I
B	Symbolverzeichnis	XIV

Kapitel 1

Einleitung

In dieser Arbeit werden mit Hilfe eines aus der Mustererkennung abgeleiteten Ansatzes zwei typische Aufgabenstellungen aus dem Bereich der Analyse geodätischer Deformationsmessungen bearbeitet.

Die hier benutzten Ansätze bedienen sich dabei automatisch arbeitender Clusterverfahren, die unter Nutzung formaler Kriterien Daten in homogene Gruppen einordnen. Die Besonderheit sind die verwendeten Fuzzy-Clusterverfahren, bei denen die Daten nicht genau einer Gruppe zugeordnet werden, sondern jeder Gruppe mit einem graduell variierenden Zugehörigkeitswert. Diese Art der Betrachtung entspricht wesentlich besser der realen Welt und der menschlichen Denkweise als eine eindeutige Zuordnung. So gibt es in der realen Welt beispielsweise nicht nur die Farben Schwarz und Weiß, sondern dazwischen noch beliebig viele Grautöne. Die üblicherweise verwendeten scharfen Klassifikationen täuschen dagegen eine scharfe Realität vor. Sie erleichtern zwar vordergründig die Interpretation, es besteht aber in größerem Maße die Gefahr von Fehlinterpretationen.

1.1 Einführung in die Thematik

Im geodätischen Sprachgebrauch versteht man unter „Deformationsmessungen“ Messungen, „... bei denen geometrische Veränderungen an Bauwerken, technischen Anlagen oder sonstigen natürlichen oder künstlichen Objekten nachzuweisen sind.“ (PELZER, 1987). Diese Definition schließt sowohl Verformungen des Untersuchungsobjektes als auch Bewegungen gegenüber der Umgebung (Starrkörperbewegungen) ein. In neuerer Literatur spricht man auch von „Überwachungsmessungen“ (WELSCH U.A., 2000) oder „Überwachungsvermessungen“ (DIN, 2009), da Deformationen mit den üblicherweise eingesetzten Messverfahren nicht direkt gemessen werden können, sondern sich erst als abgeleitete Größen aus der Auswertung ergeben.

Typische Untersuchungsobjekte sind nach PELZER (1987):

- Bauwerke (Talsperren, Brücken, Türme, Gebäude, Tunnel, Verkehrswege, ...) oder Teile davon,
- Maschinen und Industrieanlagen,
- Teile der Erdoberfläche (Baugruben, Rutschhänge, Gletscher, Tektonische Platten, ...).

Üblicherweise führt man im Rahmen von Überwachungsmessungen am Untersuchungsobjekt eine räumliche und eine zeitliche Diskretisierung durch. Die räumliche Diskretisierung ergibt sich dadurch, dass man das Objekt durch eine gewisse Anzahl von Messpunkten abstrahiert. Die Messpunkte sollen

dabei so verteilt sein, dass ihr vermeintliches Bewegungsverhalten das Bewegungsverhalten des Objektes oder von Objektteilen repräsentiert (siehe Abbildung 1.1). Die zeitliche Diskretisierung ergibt sich dadurch, dass die messtechnische Bestimmung der Objektpunkte in bestimmten zeitlichen Abständen wiederholt durchgeführt wird, um daraus die geometrischen Veränderungen ableiten zu können. Die zu einem bestimmten Zeitpunkt ausgeführte messtechnische Bestimmung nennt man Messepoche, den zeitlichen Abstand zwischen mehreren Bestimmungen Epochenabstand. Die erste Messepoche ist dabei die sogenannte Nullepoche oder Referenzeпоche.

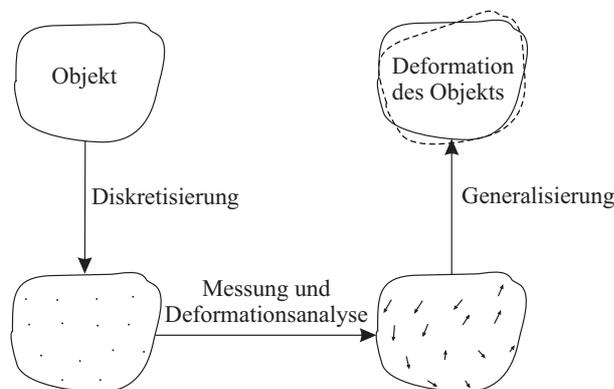


Abbildung 1.1: Das Diskretisierungs- und Generalisierungsproblem bei Überwachungsaufgaben (nach NIEMEIER, 1985a).

Die anschließende Analyse von Überwachungsmessungen, oder auch Deformationsanalyse genannt, ist die Anwendung bestimmter Auswerteprozeduren, um zielgerichtet zu Ergebnisaussagen zu kommen. Überwiegend geht man so vor, dass in jeder Messepoche aus den Originalmesswerten (den Beobachtungen) in einem der Aufgabe angepassten Koordinatensystem, ein-, zwei- oder dreidimensionale Koordinaten für die Objektpunkte berechnet werden. Anschließend unterzieht man die Koordinaten der Punkte einem Zwei- oder Mehreпоchenvergleich. Aus dem festgestellten Bewegungsverhalten der Objektpunkte wird in einem darauf folgenden Generalisierungsschritt auf das Bewegungsverhalten des Untersuchungsobjektes geschlossen.

Es ist aber auch möglich und wird in einigen Fällen auch praktiziert, die Beobachtungen direkt einem Epochenvergleich zu unterziehen. Die Verwendung von Punktkoordinaten ist allerdings bei der Darstellung von Veränderungen oft geometrisch anschaulicher und universeller anwendbar (nach NIEMEIER, 1985b). Mit dem Einsatz automatischer Messverfahren kann man heute relativ kurze Epochenabstände realisieren und erhält so in der Regel eine große Anzahl von Messeпоchen. Dadurch ist es auch möglich geworden, Verfahren der Zeitreihenanalyse (zum Beispiel CHATFIELD, 2004) im Auswerteprozess anzuwenden. Praktische Anwendungen hierzu finden sich beispielsweise in KUHLMANN (1996).

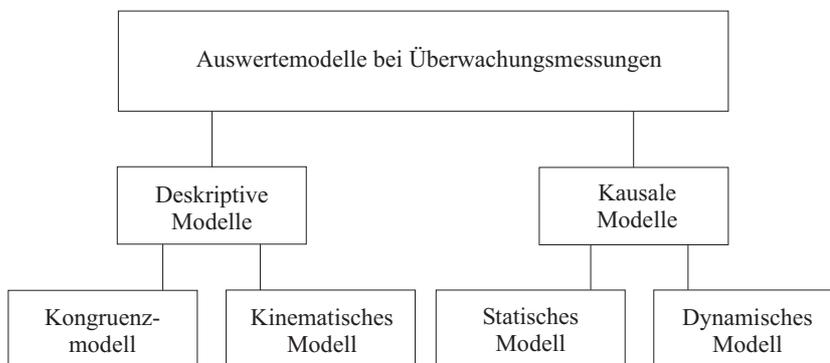


Abbildung 1.2: Auswertemodelle bei Überwachungsmessungen (aus WELSCH U.A., 2000).

Die Abbildung 1.2 zeigt verschiedene Auswertemodelle geodätischer Überwachungsmessungen, die man in neuerer Literatur in „Deskriptive Modelle“ und „Kausale Modelle“ unterscheidet (WELSCH U.A., 2000).

Bei den kausalen Modellen betrachtet man die gemessenen Objektreaktionen als Folge von einwirkenden Kräften. Sie werden auch als Ursache-Wirkungs-Modelle (engl. cause-response models) bezeichnet (HEUNECKE U.A., 1998). In den kausalen Modellen werden neben den eigentlichen geometrischen Messwerten auch die Zeit und die auf das Objekt wirkenden Einflussgrößen mit berücksichtigt. Befindet sich das Objekt zum Zeitpunkt der Messungen hinreichend in Ruhe, so dass die Zeit im Modell nicht explizit berücksichtigt werden muss, spricht man von einem „Statischen Modell“, andernfalls von einem „Dynamischen Modell“. Die dynamischen Deformationsmodelle waren in den letzten Jahren wiederholt Gegenstand der geodätischen Forschung. Zu nennen sind insbesondere: BOLJEN (1983), PFEUFER (1988), HEUNECKE (1995), HEINE (1999) und MIIMA (2002). Aufgrund der oftmals schwierigen Modellbildung und der problematischen Erfassung der Einflussgrößen haben diese Modelle allerdings in der geodätischen Praxis bisher kaum Anwendung gefunden.

In der Praxis werden fast ausschließlich die deskriptiven Modelle angewendet. Bei diesen Modellen werden die auf das Objekt einwirkenden Kräfte nicht mit berücksichtigt, und es erfolgt ein rein geometrischer Vergleich des Objektzustandes zwischen den Epochen. Die Ursache etwaiger Punktbewegungen bleibt innerhalb der Modelle außer Betracht. Man unterscheidet bei den deskriptiven Modellen ein „Kongruenzmodell“ und ein „Kinematisches Modell“. Beim Kongruenzmodell wird die Identität der Geometrie (Kongruenz) des Punktfeldes zwischen jeweils zwei Epochen untersucht. In älterer Literatur (z.B. PELZER, 1985) findet man dafür auch die Bezeichnung „Quasistatisches Modell“. Die Einflussgröße „Zeit“ findet hier nur insofern Anwendung, als dass die zu vergleichenden Messepochen in einer zeitlichen Reihenfolge geordnet werden. Ist das Ziel der Auswertung eine zeitabhängige Beschreibung des Punktfeldes, spricht man von einem kinematischen Modell. Es dient zum Beispiel dazu, ein zukünftiges Bewegungsverhalten von Objektpunkten aus den zu diskreten Zeitpunkten durchgeführten Messungen zu präzisieren.

Für die Anwendung bestimmter Auswertemodelle auf die Untersuchungsobjekte besitzen die von NIEMEIER (1988) angegebenen Kriterien nach wie vor ihre Gültigkeit. Sie lauten:

- Die Bedeutung der Deformationsuntersuchung für das Objekt.
- Die verfügbaren geodätischen und nichtgeodätischen Messinformationen.
- Die verfügbaren Vorinformationen über die vermuteten Deformationen.
- Der Aufwand, den man gegenüber einem Auftraggeber vertreten bzw. durchsetzen kann.

Bei allen Auswertemodellen ist es üblich, die Punktkoordinaten als Zufallsgrößen aufzufassen. Neben den Koordinatenvektoren aus den einzelnen Messepochen werden daher die entsprechenden Kovarianzmatrizen der Koordinaten mit in den Auswerteprozess eingeführt.

Der Ablauf bei dem in der Praxis weit verbreiteten Kongruenzmodell besteht im Wesentlichen aus den folgenden Schritten (NIEMEIER, 2008):

1. Freie Ausgleichung der Messungen jeder einzelnen Epoche.
2. Anwendung eines globalen Kongruenztests zur Beantwortung der Frage, ob es überhaupt signifikante Punktverschiebungen zwischen den einzelnen Epochen gibt. Sollten signifikante Punktverschiebungen angezeigt werden, folgt noch als nächster Schritt die
3. Lokalisierung von Einzelpunktbewegungen zur Beantwortung der Frage: Welche Punkte haben sich signifikant verschoben?

Diese gestufte Vorgehensweise gilt in der Forschung weitestgehend als abgeschlossen. Sie hat sich seit vielen Jahren in der Praxis bewährt, und es existieren ausgereifte Softwareprogramme zu ihrer Anwendung. Üblicherweise enden diese Programme unter anderem mit der Ausgabe einer graphischen Darstellung des Objektpunktfeldes mit den eingezeichneten Verschiebungsvektoren. Eine weitere Interpretation im Hinblick auf Gruppen ähnlicher Verschiebungsvektoren bleibt in der Regel dem

Bearbeiter überlassen, der durch visuelle Beurteilung der graphischen Darstellung irgendwelche Bewegungsmuster zu erkennen versucht. Für einen geschulten Bearbeiter stellt das in der Regel kein großes Problem dar. Wenn jedoch mit Hilfe automatischer Messverfahren zum Beispiel größere Berghänge (HABERLER, 2005) oder Böschungen im Tagebau (GUDER und DAHMEN, 2004) mit sehr vielen Objektpunkten überwacht werden, ergibt sich der Bedarf nach einer automatischen Interpretation der Verschiebungsvektoren. Diese Monitoringsysteme mit angeschlossener Online-Alarmierung sollen möglichst ohne einen interpretierenden Bearbeiter auskommen. Es ist nun notwendig, das Expertenwissen eines geschulten Bearbeiters in so ein System zu integrieren oder durch algorithmische Ansätze zu ersetzen. Dazu gibt es eine Reihe von Möglichkeiten, auf die teilweise im Rahmen dieser Arbeit noch näher eingegangen wird.

1.2 Motivation

Angeregt durch NAZIF und LEVINE (1984) und NIEMANN (1981) erkannten FÖRSTNER und WERNER (1986), dass viele typische Aufgabenstellungen aus der Deformationsanalyse klassische Fragestellungen der Mustererkennung in der Informatik darstellen und dass die dort benutzten Problemlösungen sich unmittelbar für die Deformationsanalyse einsetzen lassen. Insbesondere trifft das auf die automatische Segmentierung eines Punktfeldes in Punktgruppen mit ähnlichem Bewegungsverhalten zu. LEHMKÜHLER (1993) griff diese Idee auf und stellte einen speziellen Ansatz vor. Er nutzte dabei das von RISSANEN (1983) beschriebene MDL-Prinzip (engl. Minimum Description Length Principle) aus, um Diskontinuitäten in der Ähnlichkeit von Verschiebungsvektoren benachbarter Punkte aufzudecken.

Diese Arbeit soll nun einen weiteren Beitrag zur Anwendung von Mustererkennungsmethoden in der Deformationsanalyse leisten. Der Schwerpunkt dieser Arbeit liegt dabei im Gegensatz zu LEHMKÜHLER (1993) in der Einführung und Betrachtung unscharfer Clusterverfahren (engl. Fuzzy-Clustering). Die Anregung dazu kam im Wesentlichen durch die Arbeit der Spezialstudien­gruppe 4.190 „Non-probabilistic assessment in geodetic data analysis“ (www.dgfi.badw.de/ssg4.190/welcome.html) der Internationalen Assoziation für Geodäsie, die von KUTTERER in den Jahren 1999 - 2003 geleitet wurde, und insbesondere durch die Veröffentlichungen von BOCKLISCH (1987), BEZDEK und PAL (1992) sowie TILLI (1993).

Die in dieser Arbeit verwendeten Fuzzy-Clusterverfahren dienen der unscharfen Beschreibung und dem Erkennen von Situationen, Zuständen und Verhältnissen, die durch Vektoren in einem Merkmalsraum gegeben sind. Die Vorgehensweise unterscheidet sich grundlegend von den regelbasierten Fuzzy-Verfahren, an die zu allererst der Begriff „Fuzzy“ erinnert und die im ingenieur­geodätischen Kontext beispielsweise in HEINE (1999), WIESER (2002) und HABERLER (2005) Anwendung gefunden haben. Bei diesen regelbasierten Verfahren werden die Zusammenhänge zwischen den unscharf beschriebenen Merkmalsausprägungen in Form von Expertenwissen in einer Regelbasis abgelegt, während die Fuzzy-Clusterverfahren ohne formale Regeln auskommen und üblicherweise auch ohne die Vorgabe von Expertenwissen.

Kapitel 2

Vom Muster zur Mustererkennung

2.1 Allgemeines zur Mustererkennung

Zu Beginn dieser Arbeit soll zunächst der Begriff „Mustererkennung“ und was sich dahinter verbirgt, näher betrachtet werden. Leider gibt es sowohl in der deutschsprachigen als auch in der englischsprachigen Literatur keine eindeutige Definition dieses Ausdruckes. Der Autor der bedeutendsten deutschsprachigen Literaturquelle zum Thema Mustererkennung (NIEMANN, 2003) schreibt dazu: *„Die Frage, ob die Terminologie, die von „Muster“, „Mustererkennung“ und dergleichen mehr spricht, glücklich gewählt ist, sei hier zwar aufgeworfen, aber ihre Beantwortung, die eine vorherige lange und vermutlich langweilige Diskussion unterschiedlicher Definitionen erfordern würde, dem Leser anheimgestellt. Es ist aber zu erwähnen, dass Bezeichnungen wie Muster und Mustererkennung (englisch „pattern“ und „pattern recognition“) international eingeführt und in der einschlägigen Fachliteratur üblich sind. Es ist auch zu erwähnen, dass es leider immer noch keine Definition des Begriffs „Muster“ gibt, die ähnlich präzise und mathematisch verwertbar ist wie die Definition der Information durch SHANNON. Ist das ein Hinweis darauf, dass es keine gibt?“*

Eine sehr einfache und allgemeine Definition des Begriffes Mustererkennung geben BEZDEK und PAL (1992): *„Pattern recognition is the search for structure in data.“* „Mustererkennung ist die Suche nach Strukturen in Daten.“

In BEZDEK U.A. (2005) haben sich die Autoren intensiv mit verschiedenen publizierten Definitionen des Begriffs „pattern recognition“ auseinandergesetzt und sind zu dem Schluss gekommen, dass die einzelnen Definitionen stark von der fachlichen Herkunft der Autoren und den Anwendungen abhängen. Eine gemeinsame übergeordnete Definition fällt ihnen auch schwer. Sie schreiben: *„The easiest, and we think, most accurate overall description of this field is to say that it is about feature analysis, clustering, and classifier design,...“*. Aus den verschiedenen Definitionen haben sie die folgende als ihren Favoriten ausgewählt, die von DEVIJER und KITTLER (1982) stammt: *„pattern recognition is a very broad field of activities with very fuzzy borders“*.

Nach NIEMANN (2003) lassen sich die Begriffe Muster und Mustererkennung wie folgt definieren:

- **Muster:** Alle für einen bestimmten Problembereich interessierenden und durch geeignete Sensoren erfassten physikalischen Größen oder Funktionen davon.
- **Mustererkennung:** Die mathematisch-technischen Aspekte des automatischen Klassifizierens und Analysierens von Sensorsignalen.

Die Mustererkennung versucht als ein Teilgebiet der Informatik, die Fähigkeiten der menschlichen Wahrnehmung nachzubilden. Nach NIEMANN (2003) geht es hierbei allerdings nicht darum, es „*genauso*“ wie ein Mensch oder ein anderer Organismus zu machen, sondern nur „*ähnlich zuverlässig*“.

Der Prozess der Mustererkennung wird in der Literatur (zum Beispiel in DUDA U.A., 2000) vielfach vereinfacht in die Schritte

- Aufnahme (Messung) einschließlich Vorverarbeitung,
- Merkmalsextraktion und
- Klassifikation

eingeteilt. Eine detailliertere Darstellung dieser Schritte zeigt die Abbildung 2.1.

Fast allen Systemen zur Klassifikation und Analyse von Mustern liegen nach NIEMANN (2003) wenige gemeinsame Prinzipien zugrunde. Er hat hierzu mehrere Postulate aufgestellt. Die für diese Arbeit wichtigsten sind im Folgenden zitiert:

1. *„Zur Sammlung von Informationen über einen Problembereich steht eine repräsentative Stichprobe zur Verfügung.“*
2. *„Ein (einfaches) Muster besitzt Merkmale, die für seine Zugehörigkeit zu einer Klasse charakteristisch sind.“*
3. *„Die Merkmale bilden für Muster einer Klasse einen einigermaßen kompakten Bereich im Merkmalsraum. Die von Merkmalen verschiedener Klassen eingenommenen Bereiche sind einigermaßen getrennt. Dieses ist die wichtige Kompaktheitshypothese für Merkmale.“*
4. *„Ein (komplexes) Muster besitzt einfachere Bestandteile, die untereinander bestimmte Beziehungen haben. Das Muster lässt sich in diese Bestandteile zerlegen.“*
5. *„Zwei Muster sind ähnlich, wenn ihre Merkmale oder einfacheren Bestandteile sich nur wenig unterscheiden.“*

Ein allgemeines Modell der Mustererkennung zeigt die Abbildung 2.1. Das interessierende Objekt der physikalischen Welt wird durch geeignete **Sensoren** aufgenommen bzw. gemessen. Es wird dabei in einen **Messdatenvektor** überführt, der das zu analysierende Muster repräsentiert. Anschließend findet eine **Vorverarbeitung** statt, die je nach Problemstellung unterschiedlich aussieht. In der Vorverarbeitung können zum Beispiel an die Originalmesswerte Korrekturen und Reduktionen angebracht werden, es können Mittelbildungen und Transformationen stattfinden oder auch Berechnungen abgeleiteter Größen. Weiterhin können hier auch noch Schwellwertoperationen, Normierungen oder Abbildungen in einen gewünschten Wertebereich durchgeführt werden.

Im nächsten Schritt erfolgt dann eine **Merkmalsextraktion** aus den aufbereiteten Größen. Diese extrahierten wesentlichen Merkmale des Musters dienen anschließend dazu, eine Zuordnung des aufgenommenen Musters oder von Teilen des Musters zu verschiedenen Klassen zu erreichen. Die Muster innerhalb einer Klasse sollen dabei möglichst ähnliche oder gleiche Ausprägungen der Merkmale besitzen, Muster verschiedener Klassen sollen möglichst unterschiedliche Ausprägungen aufweisen. Klassen sind dabei allgemein Vereinigungen von Objekten (hier Mustern) nach inhaltlichen oder formalen Kriterien. Gleiche oder ähnliche Vorgänge innerhalb der betrachteten Problematik werden jeweils durch die Klassen repräsentiert.

Eine eventuell zwischengeschaltete **Merkmalsreduktion** dient dazu, in einem Optimierungsprozess die Anzahl der Merkmale zu reduzieren. Dies ist insbesondere in Aufgabenstellungen wichtig, bei denen eine sehr große Anzahl von Merkmalen vorliegt. Neben der Problematik eines hohen Rechenaufwandes kann eine hohe Anzahl von Merkmalen unter Umständen auch zu Interpretationsproblemen

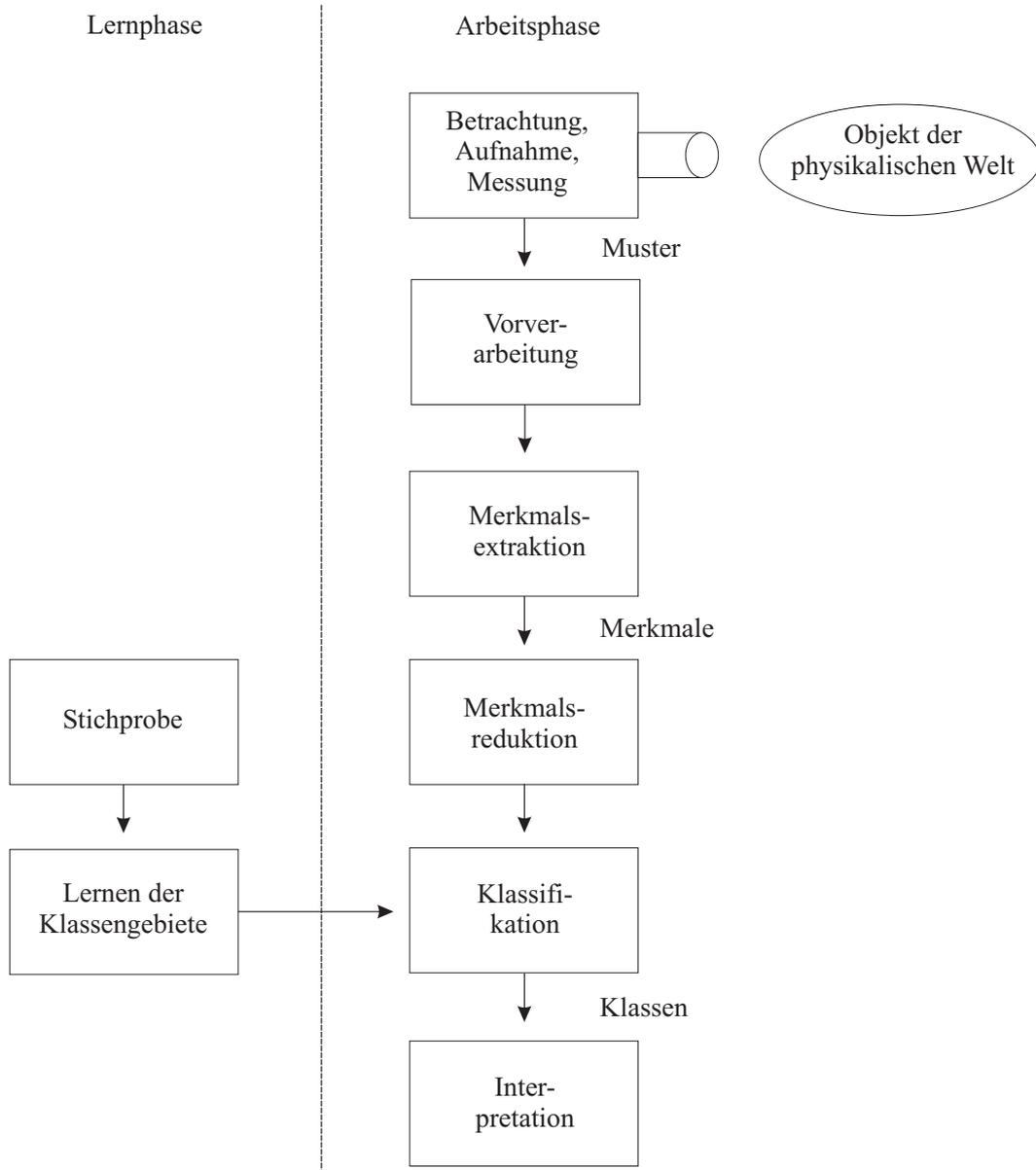


Abbildung 2.1: Allgemeines Modell der Mustererkennung (nach NIEMANN, 2003).

führen. „Ziel ist es, relevante Parameter zu gewinnen, die eine möglichst einfache anwendungsbezogene Interpretation der Messdaten gestatten.“ (KOLB, 1996)

Unter **Klassifikation** versteht man nach BOCK (1974) sowohl das Einordnen eines Objektes (hier Musters) in eine bereits bestehende Klassenstruktur als auch den strukturbildenden Prozess der Entdeckung und Lokalisierung von Klassen.

Die Parameter der Klassengebiete (Klassenzentren und Klassengrenzen) können je nach Anwendungsfall unterschiedlich bestimmt oder vorgegeben werden: (Auf Beispiele hierzu wird im weiteren Verlauf der Arbeit noch eingegangen.)

1. Liegen gänzlich unbekannte Strukturen vor, können die Parameter im Klassifizierungsprozess automatisch als Unbekannte mit geschätzt werden.
2. Liegen wiederkehrende Klassifizierungsaufgaben vergleichbarer Aufgabenstellungen vor, werden die Parameter in der Regel anhand einer repräsentativen Stichprobe von Daten vorab geschätzt. Man spricht hier auch von „Lernen“ (NIEMANN, 2003) oder „Training“ (DUDA U.A., 2000) des Klassifikators. Im späteren Klassifizierungsprozess, der auch Arbeits- oder Identifikationsphase genannt wird, werden diese Parameter dann konstant gehalten.
3. Gibt es bereits zu Beginn der Klassifikationsaufgabe Expertenwissen über die Lage und Form der zu erwartenden Klassen, können die Parameter auch direkt vorgegeben werden.

Ist unter 2. die Struktur der Stichprobe bekannt und wird der Klassifikator in einem iterativen Prozess angeleitet, bei dem die Ergebnisse mit den bekannten Strukturen der Stichprobe verglichen werden, so spricht man von „supervised learning“ („überwachtem Lernen“). Im anderen Fall spricht man von „unsupervised learning“ („unüberwachtem Lernen“). (DUDA U.A., 2000)

Nach der Klassifikation schließt sich als letzter Schritt noch eine **Interpretation** an, in der die Ergebnisse der Klassifikation durch die Verknüpfung mit Vor- und Zusatzwissen abschließend bewertet werden.

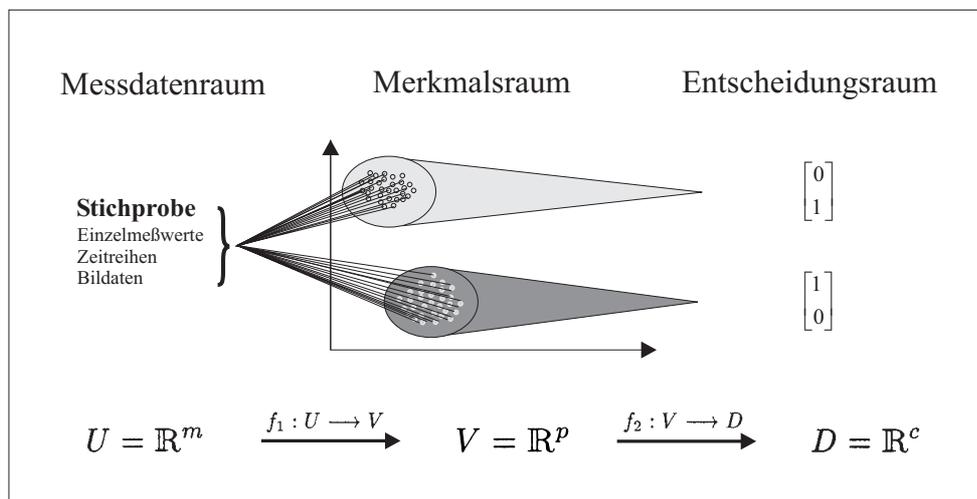


Abbildung 2.2: Darstellung der Mustererkennungsaufgabe als mathematische Abbildung (nach KOLB, 1996).

Bei dem Mustererkennungsprozess handelt es sich um eine verkettete Abbildung (siehe Abbildung 2.2). Das Muster, welches durch Messwerte im m -dimensionalen Messdatenraum U repräsentiert wird, wird in einer 1. Abbildung f_1 auf den p -dimensionalen Merkmalsraum V abgebildet. Anschließend wird der p -dimensionale Merkmalsraum durch eine 2. Abbildung (Zuordnung) f_2 weiter auf den c -dimensionalen Entscheidungsraum D abgebildet. Die Dimensionen der Räume m, p, c und damit die

Datenmenge nimmt im Laufe der Abbildungen in der Regel stark ab, und es wird ein erheblicher Datenkompressionseffekt erzielt. Durch diese Abstraktion ist bei vielen Anwendungen ein Erkennen wesentlicher Zusammenhänge oftmals überhaupt erst möglich.

2.2 Muster

Umgangssprachlich wird der Begriff „Muster“ für viele verschiedene Dinge verwendet. Unter Muster versteht man zum Beispiel:

- Einzelne Warenstücke zur Präsentation.
- Eine Vorlage, nach der etwas angefertigt wird.
- Eine sichtbare Oberflächenzeichnung oder -struktur (z.B. Teppichmuster).
- Ein bestimmtes Schema (Verhaltens- oder Denkmuster).

Wie bereits erwähnt, umfasst nach NIEMANN (2003) der Begriff „Muster“ für den Bereich der Mustererkennung in der Informatik alle für einen bestimmten Problemkreis interessierenden und durch geeignete Sensoren erfassten physikalischen Größen oder Funktionen davon.

Unter diese allgemeine Definition fallen alle physikalischen Objekte und Vorgänge (NIEMANN, 1974).

Als Sensoren sind hier nicht nur technische Sensoren denkbar, sondern auch die menschlichen Sinne. Ein Muster kann zum Beispiel auch von einem Menschen beobachtet, gefühlt oder gehört und anschließend beschrieben werden. Die menschlichen Sinne sind allerdings hinsichtlich Energie- und Frequenzumfang und hinsichtlich der erfassbaren physikalischen Größen recht beschränkt. Ein Muster kann auch von mehreren Sensoren erfasst worden sein.

Ein Muster \mathbf{M} nach dieser Definition lässt sich als vektorwertige Funktion $\mathbf{f}(\boldsymbol{\alpha})$ einer vektorwertigen Variablen beschreiben:

$$\mathbf{M} = \mathbf{f}(\boldsymbol{\alpha}) = \begin{pmatrix} f_1(\alpha_{1,1}, \dots, \alpha_{1,\lambda_1}) \\ f_2(\alpha_{2,1}, \dots, \alpha_{2,\lambda_2}) \\ \vdots \\ f_\kappa(\alpha_{\kappa,1}, \dots, \alpha_{\kappa,\lambda_\kappa}) \end{pmatrix} \quad (2.1)$$

α : unabhängige Variable
 λ : Anzahl der Variablen
 κ : Anzahl der Komponenten

Die Anzahl der Komponenten des Musters entspricht in der Regel der Anzahl der verwendeten Sensoren.

Ein einfaches Beispiel für ein Muster \mathbf{M} ist das in Abbildung 2.3 dargestellte Signal eines Weggebers als Funktion über der Zeit: $\mathbf{M} = \mathbf{f}(\mathbf{t})$ ($\kappa = 1$, $\lambda = 1$).

Für eine Messung mit dem Multisensorsystem „Tachymeter“ von einem Standpunkt (SP) zu einem Zielpunkt (ZP) besteht das Muster aus den Polarkoordinaten: Horizontalrichtung (H_z), Vertikalrichtung (V_z) und Schrägstrecke (S), die von den Variablen: Standpunktkoordinaten (X_{SP}, Y_{SP}, Z_{SP}),

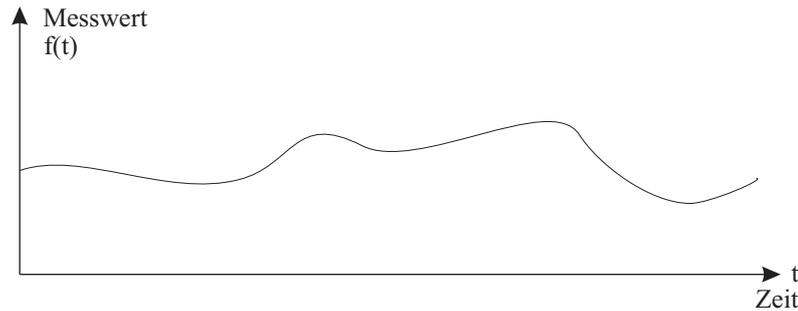


Abbildung 2.3: Darstellung einer Zeitreihe als Beispiel für ein einfaches Muster.

Zielpunktkoordinaten (X_{ZP}, Y_{ZP}, Z_{ZP}) und der Orientierungsunbekannten für den Horizontalteilkreis φ_{Hz} abhängen.¹

$$\mathbf{M}_{Tachy} = \begin{pmatrix} f_{Hz}(X_{SP}, Y_{SP}, X_{ZP}, Y_{ZP}, \varphi_{Hz}) \\ f_{Vz}(X_{SP}, Y_{SP}, Z_{SP}, X_{ZP}, Y_{ZP}, Z_{ZP}) \\ f_S(X_{SP}, Y_{SP}, Z_{SP}, X_{ZP}, Y_{ZP}, Z_{ZP}) \end{pmatrix}$$

Auch ein geodätisches Netz, das aus mehreren Stand- und Zielpunkten besteht und mit mehreren verschiedenen Sensoren gemessen wurde oder die Messwerte eines „Geosensornetzes“ (HEUNECKE, 2008) mit fest installierten Sensoren sind als Muster aufzufassen. Das Muster \mathbf{M} ist für derartige Ansätze mit dem funktionalen Modell der Ausgleichung nach vermittelnden Beobachtungen gleichzusetzen.

Synonym zum Begriff „Muster“ wird in vielen Bereichen auch der Begriff „Objekt“ verwendet. Nach BOCKLISCH (1987) sind:

Objekte: „...allgemeine diskrete Abbildungen (Modelle) von Systemen. Sie werden durch einen Satz von Informationen über das System erklärt.“.

Systeme: „...Teile der Wirklichkeit, auf die sich das Augenmerk richtet.“.

Trotz unterschiedlicher Bezeichnungsweisen und Definitionen ist aber die einführend vorgestellte prinzipielle Vorgehensweise des Mustererkennungsprozesses mit den drei Schritten

1. Messung,
2. Merkmalsextraktion,
3. Klassifikation

identisch. ZIMMERMANN (1993) schreibt in dem Zusammenhang nicht von „Mustererkennung“, sondern allgemein von „Datenanalyse“. Auch bei BOCKLISCH (1987) findet man nicht den Begriff „Mustererkennung“. Er beschreibt die Vorgehensweise im Zusammenhang mit der „Prozessanalyse“, wobei er den Begriff „Prozess“ als die Veränderung eines Systems versteht. Das entspricht im geodätischen Sprachgebrauch der Analyse von Überwachungsmessungen, bei der es darum geht, geometrische Veränderungen an Teilen der Wirklichkeit zu bestimmen.

Im Rahmen der Analyse von Überwachungsmessungen spielen abgeleitete Muster eine große Rolle. Ein Beispiel dazu sind die in Abbildung 2.4 dargestellten Verschiebungsvektoren eines Punktfeldes

¹ Zur Vereinfachung sei hier ein lokales kartesisches Koordinatensystem angenommen. Weitere Variablen (Einflussgrößen) des Musters seien durch Justierungen oder Korrekturen eliminiert oder sind in diesem Beispiel zu vernachlässigen.

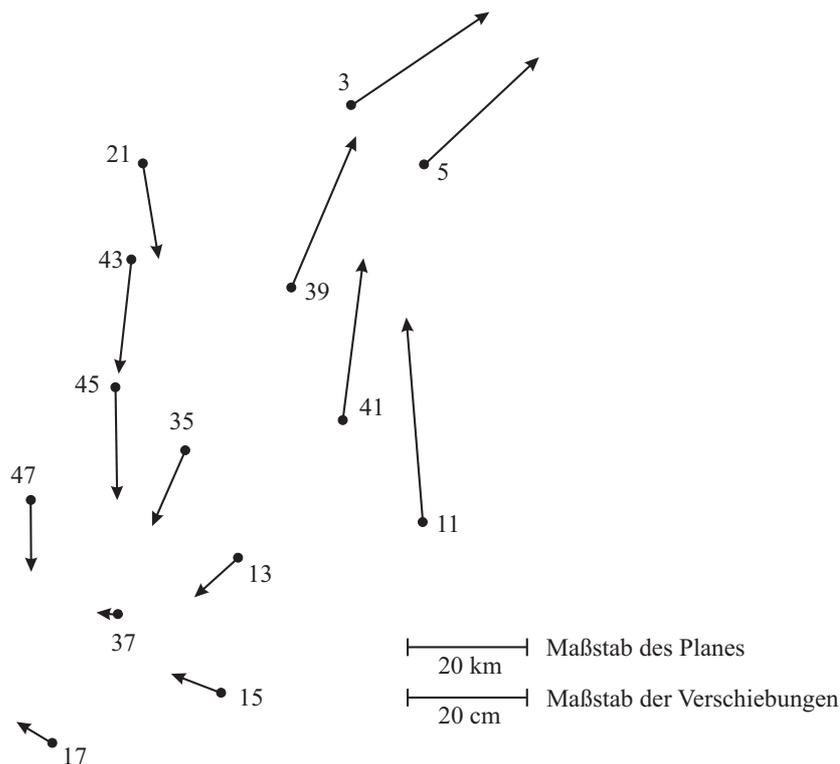


Abbildung 2.4: Verschiebungsvektoren in einem Untersuchungsgebiet. Das Testnetz Delft (WELSCH, 1983) als Beispiel für ein abgeleitetes Muster.

zwischen zwei Messepochen. Das dargestellte Punktfeld mit seinen Verschiebungsvektoren wurde aus HABERLER (2005) übernommen. Es handelt sich hierbei um Daten des Testnetzes Delft.

Das Testnetz Delft besteht aus einem zweidimensionalen Lagenetz mit simulierten Messdaten mehrerer Beobachtungsepochen, die im Rahmen des Seminars "Deformationsanalysen '83" (WELSCH, 1983) veröffentlicht wurden. Sie dienten damals als Ausgangsbasis für vergleichende Untersuchungen verschiedener Auswertestrategien und werden auch heute noch in vielen Untersuchungen als Referenz verwendet (zum Beispiel in LEHMKÜHLER (1993), HABERLER (2005) oder NIEMEIER (2008)). In der Abbildung 2.4 sind die Verschiebungsvektoren zwischen den mit identischen Näherungskoodinaten frei ausgeglichenen Epochen 1 und 3b dargestellt. Dieses Beispiel wird im weiteren Verlauf der Arbeit noch näher betrachtet werden und dient unter anderem zur anschaulichen Erläuterung verschiedener theoretischer Sachverhalte.

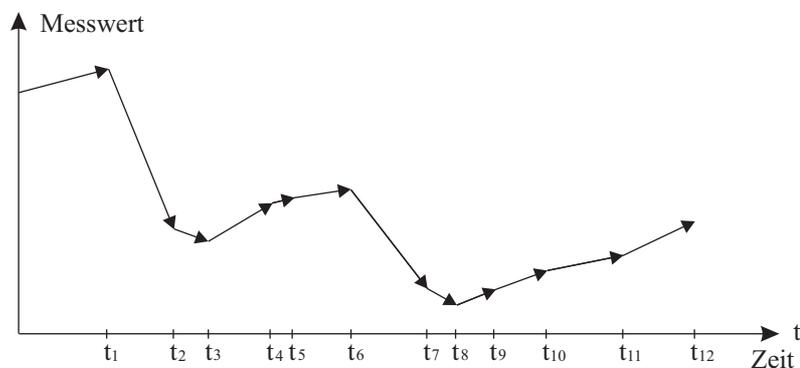


Abbildung 2.5: Zeitreihe von Verschiebungsvektoren über mehrere Epochen als Beispiel für ein abgeleitetes Muster.

Ein weiteres Beispiel für ein abgeleitetes Muster sind die in Abbildung 2.5 dargestellten Verschiebungsvektoren als Zeitreihe über mehrere Messepochen.

2.3 Merkmale

Wie einführend bereits erwähnt, besitzen Muster (oder Objekte) Merkmale, die für ihre Zugehörigkeit zu einer Klasse charakteristisch sind. Allgemein betrachtet, besitzen Muster mehr oder weniger viele Eigenschaften. Das können unterschiedliche Farbwerte sein, spezielle Strukturen, bestimmte Abmessungen und dergleichen mehr. Zur Untersuchung eines interessierenden Sachverhaltes sind einige dieser Eigenschaften aber besonders typisch. Diese Eigenschaften nennt man „Merkmale“. Die Merkmale, die die betrachteten Objekte am besten charakterisieren, fasst man zur weiteren Bearbeitung in einer sogenannten Merkmals- oder Datenmatrix \mathbf{X} zusammen. (NIEMANN, 1974)

$$\mathbf{X}_{(p,n)} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{p,1} & \cdots & x_{p,n} \end{bmatrix} \quad (2.2)$$

n : Anzahl der Objekte
 p : Anzahl der Merkmale

Ein ausgewähltes Objekt i wird durch den Spaltenvektor:

$$\mathbf{x}_i = \begin{bmatrix} x_{1,i} \\ \vdots \\ x_{p,i} \end{bmatrix} \quad (2.3)$$

beschrieben. Die Ausprägungen des Merkmals j für alle Objekte sind in der j -ten Zeile von \mathbf{X} enthalten.

Wenn p die Anzahl der für die Betrachtung eines Objektes ausgewählten Merkmale ist, dann spannen mathematisch gesehen diese Merkmale einen p -dimensionalen Merkmalsraum auf, der als p -dimensionaler euklidischer Raum verstanden werden kann. Ein Objekt lässt sich durch die Lage in diesem Merkmalsraum beschreiben (siehe Abbildungen 2.6 und 2.8).

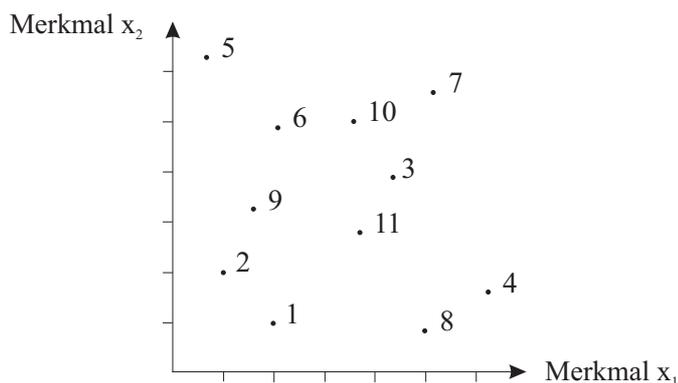


Abbildung 2.6: Objekte in einem zweidimensionalen Merkmalsraum.

Die Merkmale können unterschiedlichen Ursprungs sein. Es kann sich bei ihnen um mit technischen Sensoren objektiv gemessene Größen handeln oder auch um subjektive Einschätzungen von Menschen, die sie mit ihren Sinnen aufgenommen haben (BOCKLISCH, 1987). In der Abbildung 2.7 ist ein Beispiel für einen Merkmalsvektor angegeben.

	Objekt 1
Merkmal 1	5
Merkmal 2	1000
Merkmal 3	grün
Merkmal 4	0,7

Abbildung 2.7: Allgemeines Beispiel für einen Merkmalsvektor.

Für die Analyse von Überwachungsmessungen haben sich in den etablierten Auswerteverfahren (WELSCH U.A., 2000) als primäre Merkmale die Koordinatenunterschiede oder in polarer Form die Parameter „Länge“ und „Azimut“ der Verschiebungsvektoren identischer Punkte sehr gut bewährt². Das betrifft sowohl einen Zwei- als auch einen Mehrepochevergleich. Sollen unterschiedliche Genauigkeiten der Verschiebungsvektoren berücksichtigt werden, können (in Analogie zu den üblicherweise verwendeten statistischen Tests) die durch ihre Standardabweichung dividierten Verschiebungsvektoren oder daraus abgeleitete Größen als Merkmale verwendet werden. Bei der direkten Bestimmung von zusammenhängenden Blöcken konsistenter Punktbewegungen³, können auch die Parameter „Länge“ und „Azimut“ der relativen Verschiebungsvektoren benachbarter Punkte oder Strainparameter (WELSCH, 1982) als Merkmale Verwendung finden. Im Abschnitt 2.3.2 wird noch näher auf das Problem der Merkmalsauswahl eingegangen.

Punktnr.	3	5	11	13	15	17	...	45	47
Länge [cm]	28,0	26,5	34,5	10,0	9,0	7,0	...	19,0	12,0
Azimut [gon]	72	62	5	263	333	345	...	209	210

Tabelle 2.1: Beispiel für eine Merkmalsmatrix von Verschiebungsvektoren (Daten des Testnetzes Delft, Abbildung 2.4).

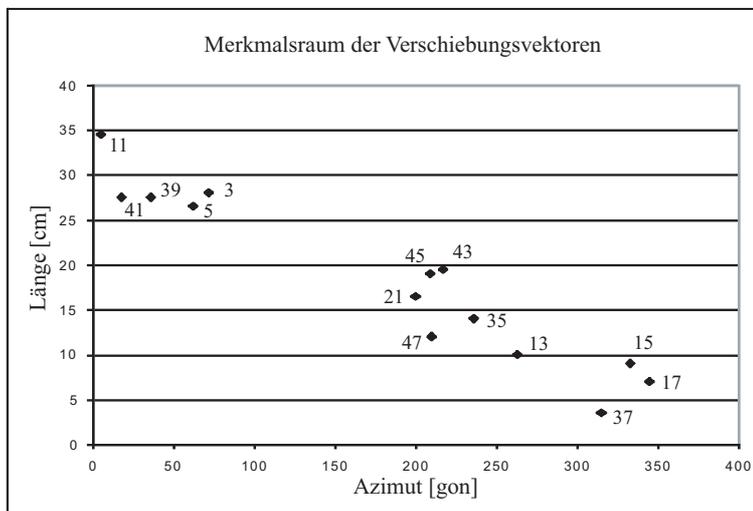


Abbildung 2.8: Die Verschiebungsvektoren des Beispiels Testnetz Delft (Abbildung 2.4) in einem zweidimensionalen Merkmalsraum.

Die Tabelle 2.1 zeigt beispielhaft einen Auszug aus der Merkmalsmatrix der Verschiebungsvektoren des Beispiels Testnetz Delft (Abbildung 2.4). In der Abbildung 2.8 ist der zugehörige Merkmalsraum dargestellt. Um zu verhindern, dass die Unstetigkeitsstelle bei 0 gon im Azimut genau durch die Punktgruppe läuft, die sich zwischen 0 gon und 100 gon befindet, wurde die Nullrichtung für die Azimute gegenüber den Originaldaten um 10 gon nach Westen verdreht⁴.

² Die Angabe bezieht sich auf ein zweidimensionales Problem. Bei einem dreidimensionalen Problem käme z.B. noch ein Neigungswinkel des Verschiebungsvektors oder ein separater Verschiebungsvektor für die Höhe hinzu.

³ Eine Aufgabenstellung, wie sie in LEHMKÜHLER (1993) und HABERLER (2005) bearbeitet wurde und die auch in dieser Arbeit im Kapitel 6 behandelt wird.

⁴ Auf dieses Problem bei der automatischen Klassifizierung der Daten wird später noch näher eingegangen.

Bei der Bearbeitung von Zeitreihen können je nach Auswerteziel als Merkmale sowohl Informationen aus dem Zeitbereich (Werte zu ausgewählten Zeitpunkten, Mittelwert, Trend, Dispersion (Varianz), Werte der Autokovarianzfunktion (Erhaltensneigung), ...) als auch Informationen aus dem Frequenzbereich (Fourierkoeffizienten, ausgewählte Frequenzen, ausgewählte Amplituden, ...) benutzt werden.

Weitere Merkmale können zum Beispiel auch Einflussgrößen auf das Objekt zu unterschiedlichen Zeitpunkten sein. Zu nennen wären da Temperaturen, Belastungszustände bei Bauwerken oder Niederschlagsmengen bzw. die Menge des Sickerwasserabflusses bei der Überwachung von Hangrutschungen. Auch können Vorabinformationen durch Experten als Merkmale eingeführt werden. Zu denken wäre hier an Informationen über Blockgrenzen an Bauwerken durch Statiker/Bauingenieure oder über vermutete Blockgrenzen an abrutschungsgefährdeten Hängen oder tektonischen Platten durch Geologen. Ein weiteres Merkmal könnte z.B. auch die Art der Vermarkung sein.

Bei vielen Problemstellungen ist es sinnvoll, eine Klasse der Punkte zu bilden, die keine signifikanten Punktverschiebungen aufweisen. Hierbei bietet es sich an, die Signifikanz des Verschiebungsvektors als ein Merkmal im Auswerteprozess mit zu betrachten.

2.3.1 Merkmalskalen

Gemessene Größen beziehen sich in der Regel auf metrische Skalen, die man in Proportionalskalen und Intervallskalen unterteilen kann. Ihre Ausprägungen geben eine Intensität wieder und sie werden daher auch als quantitative Merkmale bezeichnet. Gemeinsam ist ihnen, dass die eigentliche Maßeinheit prinzipiell beliebig gewählt sein kann. In der Regel verwendet man aber international vereinbarte Maßeinheiten.

Bei **Proportionalskalen** (auch Verhältnisskalen genannt) hat der Wert 0 eine besondere Bedeutung. Bei diesem Wert existiert die gemessene Größe nicht. Beispiele wären Längen, Winkel, elektrische Spannungen oder die Temperatur in Kelvin. Eine auf einer Proportionalskala angegebene Maßzahl ist prinzipiell immer positiv. Die Angabe eines Vorzeichens ist hier eine Zusatzinformation und drückt eine Richtung der entsprechenden Erscheinung aus (BOCKLISCH, 1987). Bei Proportionalskalen entsprechen die Maßzahlen der Stärke der Merkmalsausprägungen. Die Aussage: „Eine Strecke von 100 m ist doppelt so lang wie eine Strecke von 50 m.“ ist eine sinnvolle Aussage. Generell sinnvoll sind die folgenden Operationen: Vergleich (Identität und Größenvergleich), Addition, Subtraktion, Multiplikation und Division.

Bei **Intervallskalen** ist im Gegensatz zu den Proportionalskalen der Wert 0 eine Intensität unter den anderen. Der Nullpunkt ist hier weitgehend beliebig gewählt. Beispiele sind die Jahreszahlen oder die Temperaturskala in Grad Celsius [°C]. Die Aussage: „20°C ist doppelt so warm wie 10°C.“ ist hier nicht sinnvoll. Man kann höchstens sagen: „20°C ist wärmer als 10°C“. Sinnvoll sind bei intervallskalierten Merkmalen nur die Operationen: Vergleich (Identität und Größenvergleich), Addition und Subtraktion.

Neben den metrischen Merkmalen gibt es noch die nichtmetrischen Merkmale, die auch als qualitative Merkmale bezeichnet werden. Sie werden in Nominal- und Ordinalskalen eingeteilt.

Auf **Nominalskalen** gibt es nur die binären Aussagen „Merkmal trifft zu“ oder „Merkmal trifft nicht zu“. Möglich ist auch eine mehrklassige Einteilung in Kategorien, z.B. in Farben (rot, grün, blau), Graduierungen (schwach, mittel, stark) oder Monatsangaben (Januar, Februar, ... , Dezember). Beispiele aus dem geodätischen Bereich wären: die Vermarkungsart, die Punktart (Höhen-, Lage- oder 3D-Passpunkt, Baum, Kanaldeckel,...) oder das eingesetzte Messverfahren. Sinnvoll ist hier ausschließlich der Vergleich zweier Merkmale auf Identität. Ein Größenvergleich, eine Addition oder Multiplikation ist hier nicht möglich.

Das wesentliche Kennzeichen von **Ordinalskalen** ist eine auf- oder absteigende Ordnung der auf ihr eingeteilten Merkmale. Die Merkmale stehen in einer Rangfolge und werden zum Beispiel mit Zahlen oder Namen bezeichnet. Allgemeine Beispiele von ordinalskalierten Merkmalen sind: Schulnoten oder

Platzierungen beim Sport. Da der Skala ein Maßstab fehlt, lassen sich auf ihr aber keine Abstände berechnen. Sinnvoll sind hier nur vergleichende Operationen (Identität und Größenvergleich). Oftmals wird aber ein linearer Maßstab unterstellt und so werden zum Beispiel Durchschnittswerte von Schulnoten als arithmetisches Mittel berechnet.

proportionalskaliert:	Die Ausprägungen sind Zahlen. Eine Interpretation der Abstände ist möglich. Es existiert ein absoluter Nullpunkt.
intervallskaliert:	Die Ausprägungen sind Zahlen. Eine Interpretation der Abstände ist möglich. Es existiert kein absoluter Nullpunkt.
nominalskaliert:	Die Ausprägungen des Merkmals sind Namen. Die Ausprägung existiert oder existiert nicht. Es ist keine Ordnung möglich.
ordinalskaliert:	Die Ausprägungen des Merkmals sind Namen. Die Ausprägungen der Merkmale können geordnet werden. Ihre Abstände können in der Regel nicht interpretiert werden.

Tabelle 2.2: Merkmalstypen nach ihrem Skalenniveau sortiert (nach FAHRMEIR U.A., 2007).

Proportionalskalen besitzen den höchsten Informationsgehalt und Nominalskalen den geringsten. STEINHAUSEN und LANGER (1977) schreiben daher auch von einem hohen Skalenniveau bei Proportionalskalen und einem niedrigen Skalenniveau bei Nominalskalen. Bei bestimmten Anwendungen, auf die im Abschnitt 2.4 noch näher eingegangen wird, ist es notwendig, Merkmale auf ein anderes Skalenniveau zu transformieren. Eine Transformation von einem höheren Skalenniveau auf ein niedrigeres wird von STEINHAUSEN und LANGER (1977) als „*Niveau-Regression*“ bezeichnet. Hierbei kommt es allerdings zu einem Informationsverlust. Der umgekehrte Weg, die Transformation auf ein höheres Skalenniveau, wird als „*Niveau-Progression*“ bezeichnet und ist nur unter Zuhilfenahme von Zusatzinformationen oder Hypothesen gerechtfertigt.

Daten, die sich nicht direkt einer der Skalen zuordnen lassen, sind zum Beispiel zyklische Daten. Tageszeiten (ohne Angabe des Tagesdatums), Azimute oder Richtungen wären solche Daten. Hier sind wie bei Intervallskalen Abstände interpretierbar, es liegt aber keine eindeutige Ordnung vor. Die Frage, ob eine Horizontalteilkreisablesung von 360 gon vor oder hinter einer Ablesung von 20 gon liegt, ist nicht eindeutig zu beantworten.

In den Beispielen in dieser Arbeit werden ausschließlich metrische Merkmale verwendet, bzw. als solche angesehen. Das Problem der nicht eindeutigen Zuordnung von zyklischen Daten wird bei dem verwendeten Merkmal „Azimut eines Verschiebungsvektors“ dadurch umgangen, dass sich durch eine entsprechende Vorverarbeitung das Unstetigkeitsproblem an der Stelle 400 gon - 0 gon nicht auf das Ergebnis auswirkt.

2.3.2 Merkmalsauswahl

Die **Merkmalsextraktion**, das heißt die Auswahl der richtigen Merkmale zur Beschreibung des interessierenden Sachverhaltes, ist ein sehr wichtiger und oftmals auch komplizierter Abschnitt innerhalb des Gesamtprozesses der Mustererkennung. Eine falsche Merkmalsauswahl kann unabhängig vom eingesetzten Klassifizierungsverfahren unter Umständen zu völlig falschen Klassifizierungsergebnissen führen. Beispielsweise sind die Merkmale „Farbe“ und „Gewicht“ eines Tachymeters sicherlich keine guten Merkmale, wenn es darum geht Tachymeter verschiedenen Genauigkeitsklassen zuzuordnen. Hier wären die Merkmale „Genauigkeit der Richtungsmessungen“ und „Genauigkeit der Distanzmessung“

besser geeignet. Je komplexer die Muster aufgebaut sind und je komplexer die Problemstellung, umso schwieriger gestaltet sich üblicherweise auch die Merkmalsauswahl.

NIEMANN (2003) schreibt hierzu:

„Zur Gewinnung von Merkmalen werden hier zwei grundlegende Ansätze unterschieden:

1. *Die heuristische Methode, bei der man versucht, Merkmale aufgrund von Intuition, Phantasie und Erfahrung zu finden.*
2. *Die analytische Methode, bei der man versucht, in bestimmtem Sinne optimale Merkmale systematisch abzuleiten.“*

Bei einfacheren Aufgabenstellungen erfolgt eine sinnvolle Merkmalsauswahl überwiegend anhand der unter 1. erwähnten Vorgehensweise, durch Nutzung des Vorwissens eines Experten. Diese Vorgehensweise wird auch in den Beispielen zu dieser Arbeit genutzt. Wie bereits im Abschnitt 2.3 erwähnt, werden Merkmale verwendet, die sich bereits bei der Lösung ähnlicher Aufgabenstellungen mit anderen Auswertemethoden bewährt haben. Die unter 2. erwähnte analytische Methode funktioniert im Prinzip so, dass mit unterschiedlichen Merkmalskonfigurationen Klassifizierungen vorgenommen werden, die anschließend nach einem zu wählenden Kriterium bewertet werden. Die Merkmalskombination mit der besten Bewertung entspricht der optimalen Merkmalsauswahl. Da bei einer größeren Anzahl von Merkmalen das Bearbeiten aller möglichen Merkmalskombinationen sehr aufwändig ist, kommen besondere Selektionsverfahren zum Einsatz. Nähere Informationen hierzu finden sich zum Beispiel in OTTE (1998) oder in NIEMANN (2003).

Grundsätzlich müssen alle in die Klassifizierung einzuführenden Merkmale für alle Objekte vorliegen. Sollte das nicht der Fall sein, gibt es drei unterschiedliche Vorgehensweisen (TIMM, 2002):

- Die entsprechenden Objekte oder Merkmale werden nicht weiter berücksichtigt.
- Die fehlenden Merkmalswerte werden mittels statistischer Verfahren vorab geschätzt.
- Das Klassifizierungsverfahren muss so adaptiert werden, dass es mit fehlenden Werten umgehen kann. Ansätze hierzu finden sich zum Beispiel in TIMM (2002).

2.3.3 Merkmalsvorbehandlung

Um bestimmte Strukturen der Daten im Merkmalsraum zu erkennen, sind oftmals bestimmte Vorbehandlungen notwendig. Das können zum Beispiel Korrekturen und Reduktionen an den Originalmesswerten sein, es können Mittelbildungen und Transformationen oder auch Berechnungen abgeleiteter Größen stattfinden. Weiterhin können hier auch noch Schwellwertoperationen, Normierungen oder Abbildungen in einen gewünschten Wertebereich durchgeführt werden. Oft ergeben sich diese Vorbehandlungen aus dem Kontext der Aufgabenstellung. Als Anwender sollte man sich aber immer über die Sinnhaftigkeit und über die praktische Interpretierbarkeit der Ergebnisse Gedanken machen. (BANDEMER, 1997)

Viele automatische Klassifizierungsverfahren setzen voraus, dass die Objekte einer Klasse im Merkmalsraum eine näherungsweise hyperkugel- oder hyperellipsoidförmige Punktwolke bilden müssen. Nur dann werden sie als gemeinsame Klasse erkannt. Um dieses zu erreichen sind oftmals einige **Vorbehandlungen der Merkmale** notwendig.

Quantitative Merkmale bestehen aus dimensionsbehafteten physikalisch gemessenen Größen. Damit eine automatische Klassenbildung nicht von diesen oftmals willkürlich festgesetzten Dimensionen abhängt, werden die Merkmale in der Regel einer **Normierung** unterzogen. Das heißt, dass die entsprechenden Merkmale vorzugsweise auf ein Intervall $[0, 1]$ abgebildet werden. Die normierten Merkmale

$x_{i,j}^*$ werden für jedes Objekt x_i aus den ursprünglichen Merkmalen $x_{i,j}$ nach folgender Formel berechnet:

$$x_{i,j}^* = \frac{x_{i,j} - x_{min,j}}{x_{max,j} - x_{min,j}} \quad \forall i = 1, \dots, n; \quad \forall j = 1, \dots, p \quad (2.4)$$

- n : Anzahl der Objekte
- p : Anzahl der Merkmale
- $x_{min,j}$: Minimaler Wert des Merkmals j
- $x_{max,j}$: Maximaler Wert des Merkmals j

Es ist allerdings für jeden Einzelfall zu entscheiden, ob die Merkmale einer Normierung unterzogen werden sollen und welchen Einfluss diese Normierung auf den Informationsgehalt hat (BURMEISTER, 1997). Zu beachten ist in der Anwendung, dass einige Softwareprodukte (zum Beispiel das Programm Data Engine 4.0 (MIT GMBH, 2002)) für alle Merkmale automatisch eine Normierung durchführen.

Einen ähnlichen Effekt wie beim Normieren erhält man bei der **Wichtung** von Merkmalen durch das Multiplizieren oder Potenzieren mit Wichtungsfaktoren beziehungsweise Wichtungsexponenten. Auch hierdurch ist es möglich, die Merkmale so vorzuverarbeiten, dass im Merkmalsraum näherungsweise hyperkugel- oder hyperellipsoidförmige Punktwolken entstehen.

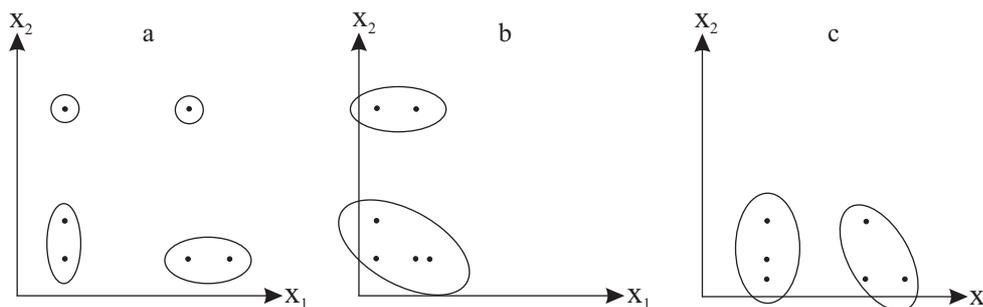


Abbildung 2.9: Beispiel zur Auswirkung unterschiedlicher Skalierungen der Merkmalsachsen auf die Klassenbildung (nach BOCKLISCH, 1987).

Ein Beispiel zur Auswirkung von Maßstabsänderungen der Achsen im Merkmalsraum zeigt die Abbildung 2.9. Dargestellt sind sechs Objekte im Merkmalsraum. In den Abbildungen b und c sind gegenüber der Abbildung a die Maßstäbe der x_1 -Achse bzw. der x_2 -Achse jeweils verändert. Man sieht, dass sich jeweils eine andere Klassenbildung ergibt.

Eine weitere Form der Merkmalsvorbehandlung ist die **Maskierung**, also die bewusste Auswahl oder Nichtberücksichtigung. Das kann zum Beispiel aus Aufwands- oder Kostengründen erfolgen, weil bestimmte Merkmale nur aufwändig oder teuer zu beschaffen sind. Wenn verschiedene Merkmale untereinander stark korreliert sind, führt das im Klassifizierungsprozess zu einer Überbetonung der entsprechenden Merkmale und sollte vermieden werden. In diesem Fall sind entsprechende Merkmale auch nicht zu berücksichtigen.

Auch das **Ableiten von Sekundärmerkmalen** aus den ursprünglichen Primärmerkmalen gehört zur Merkmalsvorbehandlung. Einzelfälle hierzu sind bereits weiter oben genannt worden. Die Berechnung von Strainparametern aus Verschiebungsvektoren gehört zum Beispiel dazu.

2.4 Klassifikation

Die Klassifikation bezeichnet einen Vorgang oder eine Methode zur Einteilung von Objekten in Klassen oder Kategorien. Eine Klasse oder Kategorie bildet dabei eine Menge von Objekten, die möglichst ähnliche oder gleiche Ausprägungen ihrer Merkmale aufweisen. Objekte verschiedener Klassen sollten möglichst unterschiedliche Ausprägungen ihrer Merkmale haben. Wie bereits im Abschnitt 2.1 erwähnt, repräsentieren die Klassen gleiche oder ähnliche Vorgänge innerhalb der betrachteten Problematik.

Eine Klassifikation kann:

- a) durch eine **Expertenvorgabe** oder
- b) durch ein **automatisches Klassenbildungsverfahren** erfolgen.

Bei der Klassifikation durch eine Expertenvorgabe erfolgt die gesamte Klassifizierung in der Regel anhand von inhaltlichen Kriterien. Es entstehen so genannte semantische Klassen, das heißt, dass die Objekte nach ihrem Sinn oder ihrer Bedeutung zusammengefasst werden (BOCKLISCH, 1987). Diese Klassen sind direkt fachkundig interpretierbar. Zu dieser Kategorie zählen auch die Verfahren, bei denen die Klassifizierung anhand einer Regel- oder Wissensbasis erfolgt, in der ein Expertenwissen abgelegt ist. Man spricht hier auch von „Expertensystemen“ (ZIMMERMANN, 1993).

Bei den automatischen Klassenbildungsverfahren⁵ erfolgt die Klassifizierung dagegen aufgrund formaler Ähnlichkeitskriterien. Auf die dabei verwendeten Ähnlichkeits- oder Distanzmaße wird im Abschnitt 2.4.1 näher eingegangen. Es entstehen sogenannte natürliche Klassen, die nicht immer fachkundig interpretierbar sind und oftmals auch erst zur Hypothesenbildung genutzt werden. Im Rahmen dieser Arbeit werden überwiegend die automatischen Klassenbildungsverfahren betrachtet.

Die automatischen Klassenbildungsverfahren werden auch als Clusterverfahren, die dabei entstehenden Klassen als Cluster bezeichnet. Diese Cluster repräsentieren idealerweise relativ dichte Punktwolken von Objekten im p -dimensionalen Merkmalsraum, die durch Regionen mit einer geringeren Dichte voneinander getrennt sind. Sind diese Grundvoraussetzungen nicht erfüllt, macht es wenig Sinn eine Clusterung durchzuführen (BACHER, 1996). Siehe dazu auch Postulat 3 im Abschnitt 2.1.

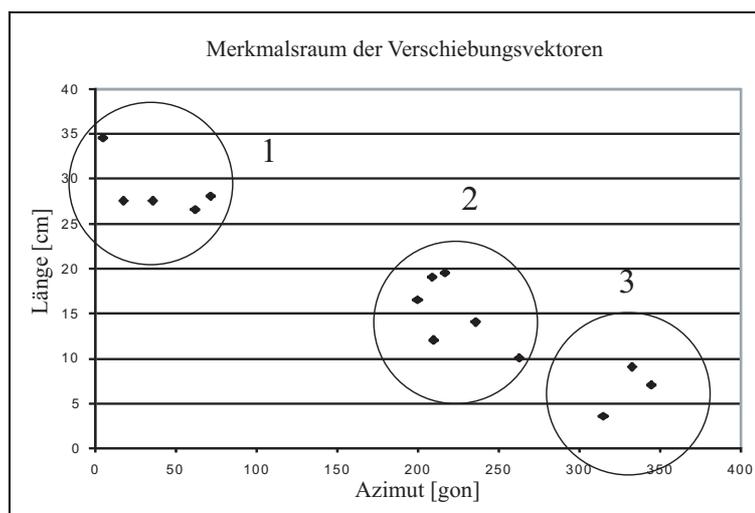


Abbildung 2.10: Die Verschiebungsvektoren des Beispiels Testnetz Delft (Abbildung 2.4) in einem zweidimensionalen Merkmalsraum und die ungefähre Lage von drei Clustern.

⁵ Der Begriff „automatische Klassenbildungsverfahren“ schließt nicht aus, dass einige Parameter dieser Verfahren im Einzelfall auch durch Expertenvorgabe festgesetzt werden können.

In der Abbildung der Verschiebungsvektoren des Testnetzes Delft im Merkmalsraum (Abbildung 2.10) können sehr gut drei Regionen von relativ dichten Punktwolken von Objekten erkannt werden, die durch Regionen mit einer geringeren Dichte voneinander getrennt sind. Diese Cluster repräsentieren Punkte mit ähnlichem Bewegungsverhalten.

2.4.1 Ähnlichkeitsmaße

Wie bereits mehrfach erwähnt, sollen Objekte innerhalb einer Klasse möglichst ähnlich und in unterschiedlichen Klassen möglichst verschieden sein. Die Ähnlichkeit ist also ausschlaggebend für die Klasseneinteilung. Sie wird bei metrischen Merkmalen oft über eine zu wählende Distanzfunktion $d_{i,k}$ im p -dimensionalen Merkmalsraum berechnet. Je kleiner die Distanzen der Merkmale, umso ähnlicher sind die Objekte.

2.4.1.1 Distanzfunktionen für metrische Merkmale

Für metrische Merkmale gibt es eine ganze Reihe von Distanzfunktionen. Die mit Abstand größte Bedeutung besitzen dabei Distanzmaße der **Minkowski-Familie** (auch L_r -Distanzen genannt (BOCK, 1974)⁶). Sie lassen sich wie folgt berechnen:

$$d_{i,k} = \left(\sum_{j=1}^p |x_{k,j} - x_{i,j}|^r \right)^{1/r} \quad (2.5)$$

- $r \geq 1$: Minkowski-Konstante
- $d_{i,k}$: Distanz
- $x_{k,j}$: Wert des j -ten Merkmals des Objektes k
- p : Dimension des Merkmalsraumes

Zwei häufig verwendete Spezialfälle sind:

- $r = 1$: Die **City-Block Distanz** (auch Manhattan oder Taxifahrer Distanz genannt). Die Distanz berechnet sich in diesem Fall aus der Summe der achsenparallelen Abstände zwischen den Objekten \mathbf{x}_i und \mathbf{x}_k .

$$d_{i,k} = \sum_{j=1}^p |x_{k,j} - x_{i,j}| \quad (2.6)$$

⁶ Es gibt in der deutschsprachigen Literatur eine ganze Reihe von Veröffentlichungen, die sich dem Thema Automatische Klassifikation respektive Clustering widmen (z.B. STEINHAUSEN und LANGER, 1977, DEICHSEL, 1985 und BACHER, 1996) beziehungsweise dieses Thema im allgemeinen Kontext der Datenanalyse behandeln (z.B. BACKHAUS U.A., 2006, FAHRMEIR U.A., 2007 und BANKHOFER und VOGEL, 2008). Es wird jedoch immer wieder auf BOCK (1974) verwiesen, die als grundlegende Literatur für diesen Themenbereich anzusehen ist.

$r = 2$: Die **Euklidische Distanz**

$$d_{i,k} = \left(\sum_{j=1}^p (x_{k,j} - x_{i,j})^2 \right)^{1/2} \quad (2.7)$$

oder in Matrixschreibweise:

$$d_{i,k} = \sqrt{(\mathbf{x}_k - \mathbf{x}_i)^T (\mathbf{x}_k - \mathbf{x}_i)}$$

Aufgrund der guten geometrischen Anschaulichkeit findet in der Praxis überwiegend die Euklidische Distanz Verwendung. Gelegentlich wird auch die quadrierte Euklidische Distanz benutzt. Bei ihr werden größere Distanzen stärker gewichtet, was zu einer geringeren Ähnlichkeit führt.

Ein weiterer Spezialfall ist die **Maximum Distanz** (auch Tschebyscheff Distanz genannt):

$r = \infty$: Die Distanz berechnet sich als größte der absoluten achsparallelen Abstände zwischen den Objekten \mathbf{x}_i und \mathbf{x}_k .

$$d_{i,k} = \max\{|x_{k,j} - x_{i,j}|, 1 \leq j \leq p\} \quad (2.8)$$

Die Maximum Distanz wird in der Praxis kaum angewendet, da sie aufgrund ihrer Maximum-Bedingung extrem empfindlich gegenüber Ausreißern in den Daten ist.

Der Einsatz eines Wichtungsfaktors q_j in Gleichung (2.5) führt auf die **gewichtete Minkowski Distanz**:

$$d_{i,k} = \left(\sum_{j=1}^p q_j |x_{k,j} - x_{i,j}|^r \right)^{1/r} \quad (2.9)$$

Durch die Verwendung individueller Gewichte q_j für jede Merkmalsachse kann sie den Einfluss bestimmter Merkmale verstärken oder abschwächen. Die Auswirkungen sind mit einer im Abschnitt 2.3.3 beschriebenen Wichtung identisch.

Die Distanzmaße der Minkowski-Familie sind translationsinvariant⁷. Sie sind allerdings nicht skaleninvariant⁸. Unterschiedliche Maßeinheiten oder vom Wertebereich nicht vergleichbare Maßskalen auf den Merkmalsachsen führen zu Verzerrungen. Das Merkmal mit dem größten Wertebereich dominiert gegenüber den anderen Merkmalen. In diesem Fall müssen die Merkmale zum Beispiel mit Hilfe einer Normierung (siehe Abschnitt 2.3.3) vergleichbar gemacht werden.

Eine weitere wichtige Distanzfunktion ist die **Mahalanobisdistanz** (MAHALANOBIS, 1936). Sie ist eine Verallgemeinerung der Euklidischen Distanz. Die Einführung der empirischen Kovarianzmatrix **Cov** der Objekte \mathbf{x}_i sorgt dafür, dass Korrelationen zwischen den einzelnen Merkmalen berücksichtigt

⁷ Translationsinvariant heißt in diesem Zusammenhang, dass die Distanzmaße unabhängig von einer merkmalsweisen Nullpunktverschiebung sind.

⁸ Skaleninvariant heißt in diesem Zusammenhang, dass die Distanzmaße unabhängig von einem merkmalsweisen Multiplikationsfaktor sind.

werden. Und dadurch, dass sich eine multiplikatorische Änderung auf einer Merkmalsachse auch auf die Kovarianzmatrix auswirkt, hat sie gegenüber den Distanzmaßen der Minkowski-Familie den Vorteil, dass sie nicht nur translations-, sondern auch skaleninvariant ist.

$$d_{i,k} = \sqrt{(\mathbf{x}_k - \mathbf{x}_i)^T \mathbf{Cov}^{-1} (\mathbf{x}_k - \mathbf{x}_i)} \quad (2.10)$$

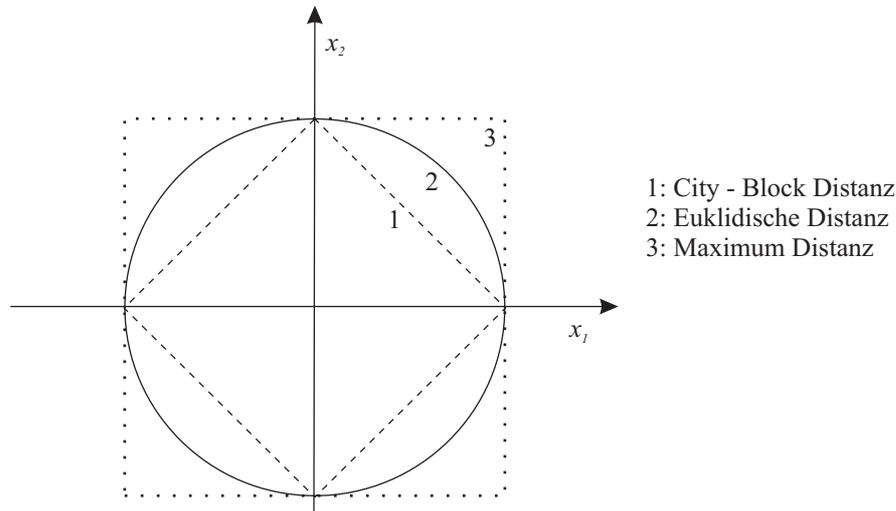


Abbildung 2.11: Linien gleicher Distanzen vom Ursprung in einem zweidimensionalen Merkmalsraum, in Abhängigkeit von der gewählten Distanz.

Die Abbildung 2.11 zeigt die Linien gleicher Distanzen vom Ursprung in einem zweidimensionalen Merkmalsraum. Bei der gewichteten Euklidischen Distanz (nach Gleichung (2.9) mit $r = 2$) wird je nach Wahl von q_j aus dem Kreis eine achsparallele Ellipse und bei der Mahalanobisdistanz ergibt sich in Abhängigkeit von \mathbf{Cov} eine beliebig gelegene Ellipse um den Ursprung.

2.4.1.2 Der Korrelationskoeffizient als Ähnlichkeitsmaß

Ein weiteres Ähnlichkeitsmaß, das zu völlig anderen Ähnlichkeiten kommt als die bisher beschriebenen Distanzmaße, ist der **Korrelationskoeffizient** $r_{i,k}$ zwischen zwei Objektvektoren (BOCK, 1974). In einigen Literaturquellen (z.B. in BACKHAUS U.A., 2006) wird er auch „Element zu Element Korrelationskoeffizient“ oder „Q-Korrelationskoeffizient“ genannt. Der Korrelationskoeffizient ist nicht ein Maß für eine Distanz zwischen den Objekten im Merkmalsraum, sondern ein Maß für die lineare Abhängigkeit der Merkmalsvektoren verschiedener Objekte untereinander.

$$r_{i,k} = \frac{\sum_{j=1}^p (x_{i,j} - \bar{x}_i)(x_{k,j} - \bar{x}_k)}{\sqrt{\sum_{j=1}^p (x_{i,j} - \bar{x}_i)^2 \cdot \sum_{j=1}^p (x_{k,j} - \bar{x}_k)^2}} \quad (2.11)$$

- $r_{i,k}$: Korrelationskoeffizient
- \bar{x} : Arithmetisches Mittel des entsprechenden Merkmals

Da für den Korrelationskoeffizienten $-1 \leq r_{i,k} \leq 1$ gilt, ist bei der Verwendung als Ähnlichkeitsmaß zu entscheiden, ob $r_{i,k} = -1$ oder $r_{i,k} = 0$ als größte Unähnlichkeit anzunehmen ist. Ein Wert $r_{i,k} = 0$ bedeutet, dass zwischen den Merkmalsvektoren der Objekte keine lineare Beziehung besteht. Ein Wert

$r_{i,k} = -1$ bedeutet eine starke negative lineare Abhängigkeit. Ein Wert $r_{i,k} = 1$ des Korrelationskoeffizienten bedeutet eine hohe lineare Abhängigkeit und damit die größte Ähnlichkeit. Entsprechend erfolgt die Umrechnung in die Ähnlichkeit $s_{i,k}$:

$$s_{i,k} = |r_{i,k}| \text{ oder } s_{i,k} = r_{i,k}^2 \quad : \quad \text{Wenn } r_{i,k} = 0 \text{ die maximale Unähnlichkeit sein soll.}$$

$$s_{i,k} = (r_{i,k} + 1)/2 \quad : \quad \text{Wenn } r_{i,k} = -1 \text{ die maximale Unähnlichkeit sein soll.}$$

In Abbildung 2.12 kann man erkennen, wie sich der Korrelationskoeffizient und zum Beispiel die euklidische Distanz als Ähnlichkeitskriterium unterschiedlich auswirken. In dieser Darstellung sind die Objekte nicht in einem Merkmalsraum dargestellt, sondern als sogenannte Merkmalsprofile. Die Merkmalsprofile der Objekte 1 und 2 verlaufen ziemlich dicht nebeneinander und besitzen daher eine ziemlich kleine Distanz. Sie unterscheiden sich aber erheblich in Bezug auf ihren Profilverlauf. Die Objekte 1 und 3 weisen dagegen einen identischen Profilverlauf auf, was zu einer hohen positiven Korrelation führt. Sie haben aber eine relativ große Distanz zueinander. In diesem Beispiel besitzen die Objekte 1 und 2 eine kleine Distanz und eine geringe Korrelation, die Objekte 1 und 3 eine große Distanz und eine hohe positive Korrelation.

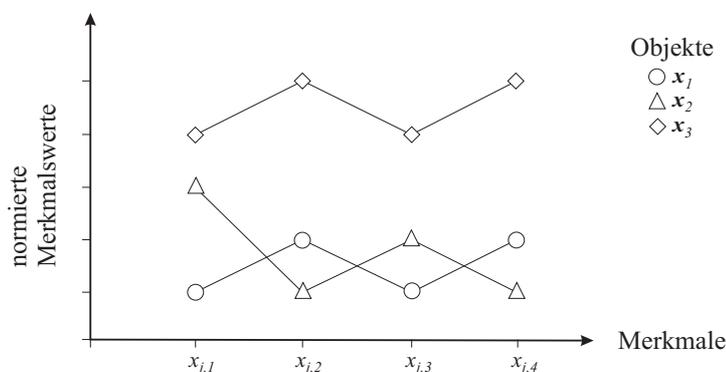


Abbildung 2.12: Merkmalsprofile von drei Objekten.

Bei Verwendung einer Distanzfunktion als Ähnlichkeitskriterium sind sich die Objekte 1 und 2 ähnlicher als die Objekte 1 und 3. Bei Verwendung des Korrelationskoeffizienten als Ähnlichkeitskriterium ist es genau umgekehrt.

Der Korrelationskoeffizient ist weder skalen- noch translationsinvariant. Daher ist eine sinnvolle Anwendung nur dann möglich, wenn vergleichbare Merkmale vorliegen oder vorher eine Normierung stattgefunden hat. Unabhängig davon ist seine Anwendung immer dann von Vorteil, wenn als Ähnlichkeit ein möglichst paralleler Verlauf der Merkmalsprofile gewünscht ist, unabhängig von der tatsächlichen Größe der Merkmalswerte.

2.4.1.3 Ähnlichkeitsfunktionen für nichtmetrische Merkmale

Aufgrund des fehlenden Maßstabes bei Ordinal- und Nominalskalen muss bei Verwendung dieser Merkmale eine andere Form der Berechnung von Ähnlichkeiten gewählt werden als bei den metrischen Skalen.

Bei **nominalskalierten Merkmalen** gibt es nur die Eigenschaften: das Merkmal existiert oder das Merkmal existiert nicht. Diese beiden Eigenschaften seien hier mit 1 und 0 bezeichnet. Wenn man jetzt die Ausprägungen eines nominalskalierten Merkmales bei zwei Objekten vergleicht, lassen sich vier Fälle unterscheiden:

- Die Eigenschaft (Ausprägung des Merkmals) ist bei beiden Objekten vorhanden.
- Nur Objekt 1 weist die Eigenschaft auf.
- Nur Objekt 2 weist die Eigenschaft auf.
- Die Eigenschaft ist bei beiden Objekten nicht vorhanden.

Auf der Grundlage der vier zu unterscheidenden Fälle von Übereinstimmungen und Nichtübereinstimmungen wurden verschiedene Ähnlichkeitsmaße entwickelt. Für ihre Berechnung bildet man vorab vier Summen:

- Σ_{11} : die Anzahl der bei beiden Objekten vorhandenen Eigenschaften,
- Σ_{00} : die Anzahl der bei beiden Objekten nicht vorhandenen Eigenschaften,
- Σ_{01} und Σ_{10} : die Anzahl der Eigenschaften, in denen die Objekte nicht übereinstimmen.

Eine Auswahl verschiedener Ähnlichkeitsmaße zeigt die Tabelle 2.3. Welches Ähnlichkeitsmaß für den konkreten Anwendungsfall auszuwählen ist, lässt sich nicht allgemeingültig sagen. Eine große Bedeutung bei dieser zu treffenden Entscheidung hat die Frage, ob das Nichtvorhandensein einer Merkmalsausprägung für die Problemstellung die gleiche Bedeutung hat wie das Vorhandensein.

Name des Ähnlichkeitsmaßes	Ähnlichkeit $s_{i,k} =$	Eigenschaften
Tanimoto	$\frac{\Sigma_{11}}{\Sigma_{11} + \Sigma_{01} + \Sigma_{10}}$	Gemeinsamer Nichtbesitz wird nicht als Ähnlichkeit betrachtet.
Simple Matching	$\frac{\Sigma_{11} + \Sigma_{00}}{\Sigma_{11} + \Sigma_{01} + \Sigma_{10} + \Sigma_{00}}$	Gemeinsamer Besitz und Nichtbesitz werden gleich gewichtet.
Russel & Rao	$\frac{\Sigma_{11}}{\Sigma_{11} + \Sigma_{01} + \Sigma_{10} + \Sigma_{00}}$	Gemeinsamer Nichtbesitz wird nicht als Ähnlichkeit betrachtet, geht aber in den Nenner ein.
Dice	$\frac{2\Sigma_{11}}{2\Sigma_{11} + (\Sigma_{01} + \Sigma_{10})}$	Gemeinsamer Nichtbesitz geht nicht in die Berechnung ein, gemeinsamer Besitz wird doppelt gewichtet.
Kulczynski	$\frac{\Sigma_{11}}{\Sigma_{01} + \Sigma_{10}}$	Gemeinsamer Nichtbesitz wird nicht als Ähnlichkeit betrachtet.

Tabelle 2.3: Ausgewählte Ähnlichkeitsmaße nominaler Merkmale (nach BACKHAUS U.A., 2006 und BACHER, 1996).

Wie bereits im Abschnitt 2.3.1 beschrieben, unterstellt man bei **Ordinalskalen** oft einen linearen Maßstab und man berechnet zum Beispiel arithmetische Mittelwerte aus Schulnoten so, als wären die Schulnoten metrisch skaliert. In ähnlicher Weise geht man auch bei der Berechnung von Distanzen zwischen ordinalskalierten Merkmalen vor. Die linear skalierten Ordinalskalen stellt man sich als Achsen eines euklidischen Merkmalsraumes vor und berechnet in der Regel Euklidische oder City-Block Distanzen zwischen den Objekten. Um die Ausprägungen eines Merkmales bei dieser Vorgehensweise unterschiedlich zu gewichten, wies zum Beispiel QU (2000) dem Merkmal „Ausstattungsstandard von Einfamilienhausgrundstücken“ mit 7 Ausprägungen nicht die Werte 1-7 zu, sondern vergab die Werte 7, 9, 10, 12, 14, 15 und 19.

In STEINHAUSEN und LANGER (1977) wird vorgeschlagen, ordinalskalierte Merkmale auf nominalskalierte Merkmale abzubilden. Jedes ordinalskalierte Merkmal mit r Ausprägungen entspricht dann $r - 1$

nominalskalierten Merkmalen. Auf diese Hilfsmerkmale können dann die in Tabelle 2.3 angegebenen Ähnlichkeitsmaße angewendet werden. Nachteilig ist bei dieser Vorgehensweise, dass das entsprechende ordinalskalierte Merkmal stark gewichtet wird. In dem Beispiel in der Tabelle 2.4 werden dadurch aus ursprünglich einem Merkmal fünf Merkmale, was zu einer fünffachen Gewichtung führt. Durch Einführung eines Gewichts von $1/(r-1)$ für jedes Hilfsmerkmal kann diese Übergewichtung allerdings vermieden werden.

Ausprägungen eines ordinalskalierten Merkmals	nominalskaliertes Merkmal				
	1	2	3	4	5
1	0	0	0	0	0
2	1	0	0	0	0
3	1	1	0	0	0
4	1	1	1	0	0
5	1	1	1	1	0
6	1	1	1	1	1

Tabelle 2.4: Abbildung eines ordinalskalierten Merkmals mit sechs Ausprägungen auf fünf nominalskalierte Hilfsmerkmale.

2.4.1.4 Ähnlichkeit bei gemischten Merkmalen

Häufig sind in einer Problemstellung nicht nur Merkmale eines Skalentypes zu verarbeiten, sondern eine Mischung aus nominal-, ordinal- und metrisch skalierten Merkmalen. Da die bisher behandelten Ähnlichkeits- oder Distanzfunktionen prinzipiell nur auf einen ausgewählten Skalentyp beschränkt sind, ist eine gemeinsame Verarbeitung aller Merkmale so nicht möglich.

Zum Umgang mit diesem Problem finden sich in der Literatur (zum Beispiel BOCK, 1974) im Wesentlichen drei unterschiedliche Vorgehensweisen:

- I. Die Merkmalsmatrix \mathbf{X} wird in einen Teil der nominalskalierten Merkmale $\mathbf{X}^{(nom)}$, einen Teil der ordinalskalierten Merkmale $\mathbf{X}^{(ord)}$ und einen Teil der metrisch skalierten Merkmale $\mathbf{X}^{(met)}$ zerlegt und die Ähnlichkeiten und Distanzen skalenweise getrennt berechnet. Die Gesamtähnlichkeit ermittelt man als gewichtetes Mittel aus den Einzelmaßen. Durch eine Gewichtung können die durch die Verschiedenartigkeiten der Maße bedingten Ungleichgewichte ausgeglichen werden. Objektive Maße zur Berechnung der Gewichte gibt es allerdings nicht, sodass die Wahl der Gewichte mehr oder weniger willkürlich erfolgt. Es bietet sich beispielsweise an, den Anteil der entsprechenden Merkmale an ihrer Gesamtzahl als Gewichte zu verwenden.
- II. Alle höherskalierten Merkmale werden auf das niedrigste vorhandene Skalenniveau transformiert (Niveau-Regression). Metrische Merkmale werden durch Diskretisierung in ordinale überführt und ordinale Merkmale (wie in Tabelle 2.4) in mehrkategoriale nominalskalierte Merkmale. Die Klassifizierung erfolgt abschließend auf der Basis des niedrigsten Skalenniveaus.
- III. Alle Merkmale werden nach der Formel (2.4) jeweils auf ein Intervall $[0, 1]$ normiert, wobei bei nominalen Merkmalen die Ausprägung „Eigenschaft existiert“ einer 1 und die Ausprägung „Eigenschaft existiert nicht“ einer 0 entspricht. Anschließend werden alle Merkmale als metrische Merkmale behandelt und die Clusterung erfolgt auf der Basis des höchsten Skalenniveaus. Eine Vorgehensweise, die zum Beispiel von QU (2000) angewendet wurde. Bei dieser Niveau-Progression ist allerdings zu prüfen, ob eine Transformation der nominalen und ordinalen in metrische Merkmale zulässig ist und nicht zu einer Sinnverfälschung führt. In der Praxis ist diese Vorgehensweise insbesondere dann problematisch, wenn relativ viele nominale Merkmale vorliegen. Das kann dazu führen, dass jede Merkmalsausprägung der nominalen Merkmale aufgrund der $[0-1]$ -Normierung zu einem eigenen Cluster führt.

2.4.2 Clusterverfahren

Die Anzahl der in der Literatur veröffentlichten Algorithmen zur Clusterung von Daten ist nahezu unüberschaubar (JAIN U.A., 2004). Sie alle vorstellen zu wollen, würde den Rahmen dieser Arbeit bei weitem sprengen. Viele dieser Verfahren ähneln sich allerdings stark und lassen sich in Gruppen zusammenfassen. Eine mögliche Einteilung anhand ihrer Vorgehensweisen zur Clusterung von Daten zeigt die Abbildung 2.13. Weitere mögliche Einteilungen findet man zum Beispiel in STEINHAUSEN und LANGER (1977) und JAIN U.A. (2004).

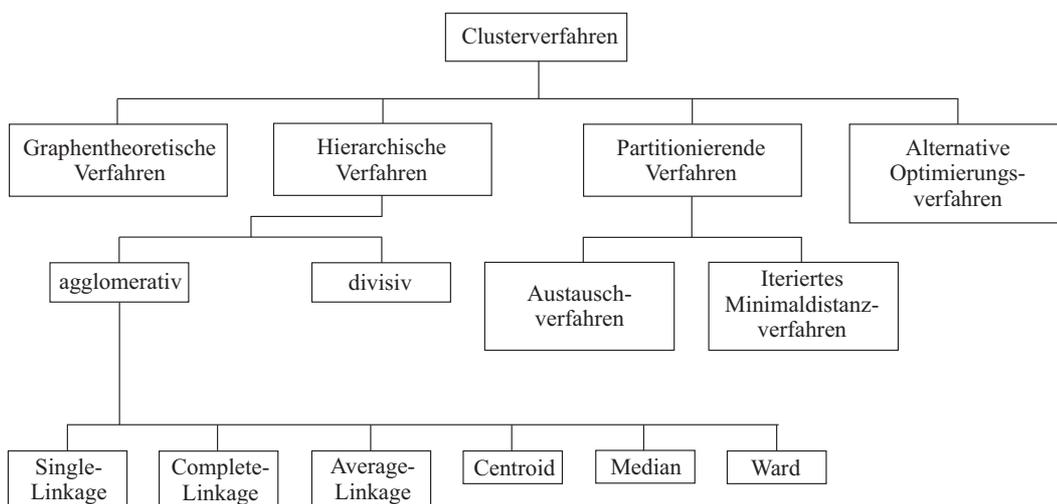


Abbildung 2.13: Überblick über ausgewählte Clusteralgorithmen (nach BACKHAUS U.A., 2006).

Die größte Bedeutung haben die hierarchischen und die partitionierenden Verfahren, auf die im Folgenden etwas ausführlicher eingegangen wird.

2.4.2.1 Hierarchische Verfahren

Der Name „Hierarchische Verfahren“ kommt dadurch zustande, dass bei diesen Verfahren Folgen von Clustern auf unterschiedlichen Distanz- oder Ähnlichkeitsebenen gebildet werden, die anschaulich in einer Hierarchie darstellbar sind.⁹ Es werden fortgesetzt durch Vergrößerung oder Verkleinerung des Distanzniveaus bereits bestehende Cluster fusioniert oder zerlegt. Der Prozess der Klassenbildung kann dabei nach zwei unterschiedlichen Konstruktionsprinzipien erfolgen, die in der Abbildung 2.14 durch die Pfeile auf der rechten Seite angedeutet sind:

- a) **Agglomeratives Verfahren:** Jedes Objekt x_i stellt zunächst ein eigenes Cluster dar. Diese Cluster werden anschließend durch schrittweise Vergrößerung des Distanzniveaus so lange miteinander fusioniert, bis eine vorgegebene Grenzdistanz zwischen zwei benachbarten Clustern überschritten wird.
- b) **Divisives Verfahren:** Hierbei bilden alle Objekte $x_1 \dots x_n$ zunächst ein gemeinsames Cluster, welches anschließend durch schrittweise Verkleinerung des Distanzniveaus so lange zerlegt wird, bis die maximale Distanz von zwei Objekten innerhalb eines Clusters kleiner ist als eine vorgegebene Grenzdistanz. Divisive Verfahren besitzen einen extrem hohen Rechenaufwand und haben deshalb in der Praxis kaum eine Bedeutung.

⁹ Insbesondere in der Photogrammetrie hat der Begriff „Hierarchische Clusterung bzw. Klassifizierung“ oft eine weitere Bedeutung. Er wird hier für gestufte Klassifizierungen benutzt, bei denen eine einmal gefundene Klassenstruktur in einem weiteren Schritt (evtl. mit anderen Merkmalen und/oder einem anderen Klassifizierungsverfahren) erneut einer Klassifizierung unterzogen wird. Diese gestufte Klassifizierung hat das Ziel, eine noch feinere Struktur zu finden.

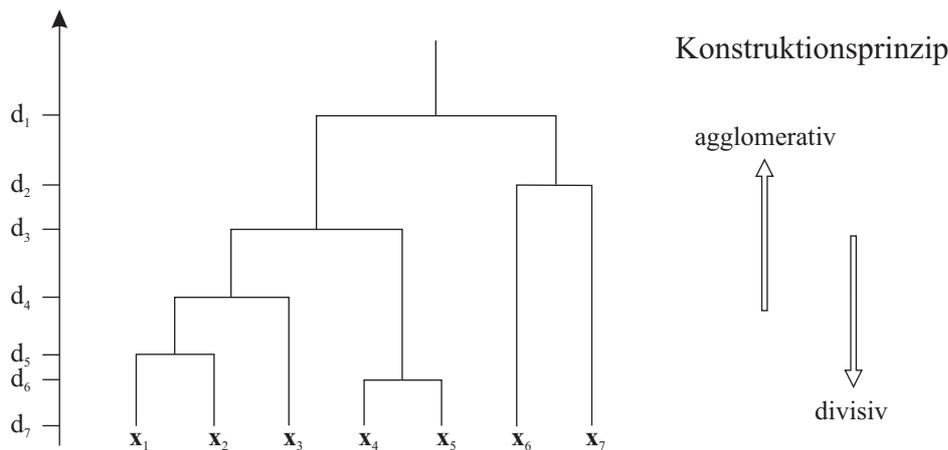


Abbildung 2.14: Dendrogramm (nach BERGS, 1981).

Der Ablauf der Zerlegung bzw. Fusionierung wird bei den hierarchischen Verfahren häufig durch ein Dendrogramm (Abbildung 2.14) verdeutlicht. Die Höhe eines Querbalkens zeigt dabei an, bei welcher Distanz Cluster zusammengelegt oder getrennt werden. Das Dendrogramm stellt somit ein wichtiges Hilfsmittel dar, um eine geeignet erscheinende Clusteranzahl festzulegen. Siehe hierzu auch Abschnitt 2.4.3.

Welche Distanzen zwischen Objekten oder Clustern zur Fusionierung herangezogen werden, kann durch unterschiedliche Fusionierungsverfahren bestimmt werden. In der Literatur (zum Beispiel BACKHAUS U.A., 2006) findet man insbesondere die folgenden:

Single Linkage: Es werden sämtliche Distanzen zwischen den Objekten des einen Clusters und den Objekten der anderen Cluster berechnet. Anschließend werden die zwei Cluster vereinigt, welche die geringste Distanz zwischen jeweils einem Objekt aus dem einen Cluster und einem Objekt aus dem anderen Cluster besitzen. Das Verfahren wird deshalb auch als „nearest neighbour“ oder Minimummethode bezeichnet. Dadurch dass bei diesem Verfahren die Fusionierung nur über jeweils zwei benachbarte Objekte gesteuert wird, kann es zu unerwünschten kettenförmigen Clusterbildungen kommen. Andererseits eignet sich dieses Verfahren gut, um Ausreißer in den Daten zu finden. Die Ausreißer bleiben in den ersten Fusionierungsschritten einzeln und werden erst auf einem relativ hohen Distanzniveau fusioniert.

Complete Linkage: Wie beim Single Linkage werden hier auch sämtliche Distanzen zwischen den Objekten des einen Clusters und den Objekten der anderen Cluster berechnet. Vereinigt werden die beiden Cluster, deren größte Distanz zwischen zwei Objekten im Vergleich zu anderen Clustern minimal ist. Das Verfahren wird deshalb auch als „furthest neighbour“ oder Maximummethode bezeichnet. Im Gegensatz zum Single Linkage werden hiermit sehr kompakte Cluster erzeugt.

Average Linkage: Wie bei den bisher beschriebenen Verfahren werden hier auch sämtliche Distanzen zwischen den Objekten des einen Clusters und den Objekten der anderen Cluster berechnet. Vereinigt werden die beiden Cluster, deren Durchschnittsdistanzen (arithmetisches Mittel) den kleinsten Abstand haben.

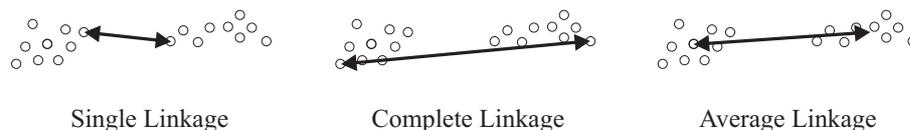


Abbildung 2.15: Graphische Illustration der Distanzen, die bei der Fusionierung von zwei Clustern berücksichtigt werden.

Centroid: In jedem Cluster wird zunächst der Clusterschwerpunkt (Centroid) als arithmetischer Mittelwert aus den ihn bildenden Objekten berechnet. Anschließend wird dann von jedem Clusterschwerpunkt zu jedem anderen Clusterschwerpunkt die Distanz berechnet. Vereinigt werden die beiden Cluster, deren Distanz zwischen den Clustercentroiden minimal ist. Der Centroid des neuen Clusters berechnet sich dann aus dem gewichteten arithmetischen Mittel der beiden Ausgangscentroide, wobei die Anzahl der Objekte in den einzelnen Clustern als Gewichte dienen.

Median: Die Median Clusterung erfolgt ähnlich der Centroid Clusterung. Der Unterschied besteht darin, dass der Centroid des neuen Clusters als ungewichtetes arithmetisches Mittel aus den Centroiden der Ausgangscluster berechnet wird.

Ward: Bei der Ward Methode werden zunächst für jedes Cluster die Clusterschwerpunkte als arithmetisches Mittel aus den sie bildenden Objekten berechnet. Anschließend werden die quadrierten euklidischen Distanzen der einzelnen Objekte aller Cluster zu allen Clusterschwerpunkten gebildet. Diese werden für jede Clusterkombination aufsummiert. Es werden die beiden Cluster vereinigt, die durch ihre Vereinigung den geringsten Zuwachs in der Gesamtsumme der Distanzen ergeben.

Nachteilig wirkt sich bei den hierarchischen Verfahren aus, dass eine einmal durchgeführte Fusion oder Zerlegung in einem späteren Schritt nicht wieder rückgängig gemacht werden kann.

2.4.2.2 Partitionierende Verfahren

Bei den partitionierenden Verfahren werden eine feste Clusteranzahl und eine genäherte Zuordnung der Objekte zu den Clustern vorgegeben. Ausgehend von dieser Startpartition wird durch schrittweises Umgruppieren der Objekte von einem Cluster zu einem anderen versucht, ein vorzugebendes Gütemaß oder eine Zielfunktion zu optimieren. In einigen Literaturquellen (zum Beispiel BOCK, 1974) werden diese Verfahren deshalb auch als „*Iterative Verfahren*“ bezeichnet.

Idealerweise könnte man eine optimale Lösung des Clusterproblems dadurch finden, dass man sämtliche mögliche Zerlegungen (Partitionen) der Menge der Objekte auf alle Cluster durchführt und jeweils ein vorzugebendes Gütemaß berechnet. Die Partition mit dem minimalen bzw. maximalen¹⁰ Wert des Gütemaßes entspricht dann der in diesem Fall optimalen Lösung. Diese Vorgehensweise scheitert in der Regel, weil die Anzahl möglicher Partitionen exponentiell mit der Anzahl der Objektpunkte ansteigt und die nötigen Rechenzeiten sehr groß werden. Schon die Partitionierung von 50 Objekten auf 3 Cluster ergibt ungefähr $1 * 10^{23}$ mögliche Aufteilungen (BOCK, 1974). Häufig ist man also aus wirtschaftlichen Gründen gezwungen, die Anzahl aller möglichen Aufteilungen erheblich einzuschränken. Das kann allerdings dazu führen, dass die optimale Lösung nicht gefunden wird.

Ein sehr gebräuchliches Gütemaß ist das Varianzkriterium g_{Var} (Formel (2.12)). Hierbei wird zunächst für jedes Objekt die (in der Regel euklidische) Distanz zum jeweiligen Clustermittelwert oder einem anderweitig definierten Clusterzentrum gebildet. Diese Distanzen werden dann innerhalb eines Clusters quadratisch aufsummiert (analog zur Berechnung der empirischen Varianz). Anschließend werden diese Summen noch einmal über alle Cluster aufsummiert. Das Varianzkriterium soll möglichst minimal werden. Wie bereits erwähnt, repräsentieren die Cluster idealerweise relativ dichte Punktwolken von Objekten im Merkmalsraum, die durch Regionen mit einer geringeren Dichte voneinander getrennt sind. Wenn das gegeben ist und die Clusterzentren in den Schwerpunkten der Punktwolken liegen und damit die Distanzen innerhalb eines Clusters relativ klein sind, führt das zu einem minimalen Wert für das Varianzkriterium.

$$g_{Var} = \sum_{j=1}^c \sum_{i=1}^{n_j} d^2(x_i, \bar{x}_j) \quad (2.12)$$

¹⁰Ob der minimale oder der maximale Wert des Gütemaßes dem Optimum entspricht, hängt von dem gewählten Gütemaß ab.

- g_{Var} : Varianzkriterium
- \mathbf{x}_i : i -tes Objekt
- $\bar{\mathbf{x}}_j$: Clusterzentrum des j -ten Clusters
- n_j : Anzahl der Objekte im j -ten Cluster
- d : Distanz
- c : Anzahl der Cluster

Nach Abbildung 2.13 werden die partitionierenden Verfahren in Austauschverfahren und iterierte Minimaldistanzverfahren unterschieden. Im Folgenden wird für das iterierte Minimaldistanzverfahren der Berechnungsablauf kurz vorgestellt, da auf ihn im Abschnitt 4.3 im Zusammenhang mit dem Fuzzy-c-Means Algorithmus noch einmal eingegangen wird.

Das **iterierte Minimaldistanzverfahren** arbeitet nach folgendem Ablauf (STEINHAUSEN und LANGER, 1977):

1. Anfangspartition vorgeben.
2. Clusterzentren berechnen.
3. Prüfe für jedes Objekt, ob das Distanzmaß zu seinem Clusterzentrum kleiner ist, als zu den anderen Clusterzentren. Wenn nicht, verschiebe das Objekt in das Cluster, zu dessen Zentrum das Distanzmaß minimal ist.
4. Fahre bei 2. fort oder beende, wenn kein Objekt in diesem Durchgang das Cluster gewechselt hat.

Durch die Neuberechnung der Clusterzentren nach jeder Verschiebung eines Objektes können hierbei die Ergebnisse von der Eingabereihenfolge der Objekte abhängig sein.

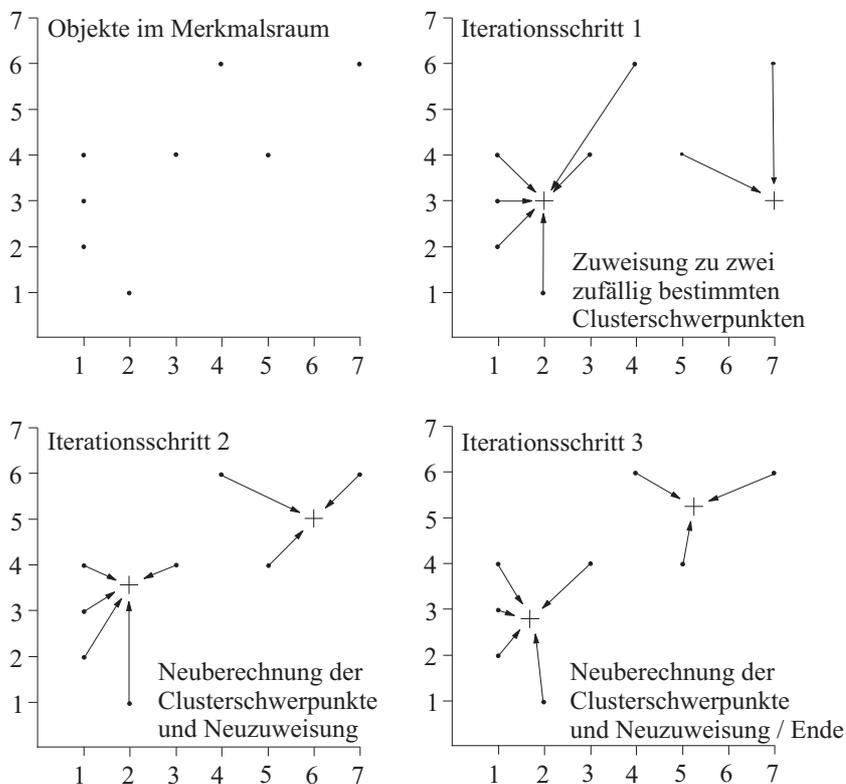


Abbildung 2.16: Iterationsschritte beim iterierten Minimaldistanzverfahren (nach DE LANGE, 2002).

Der prinzipielle Ablauf für acht Objekte und zwei Cluster bei dem iterierten Minimaldistanzverfahren ist der Abbildung 2.16 zu entnehmen.

Ein Nachteil bei den partitionierenden Verfahren ist die Tatsache, dass die Ergebnisse von der Startpartition abhängig sind. Die Ergebnisse stellen in der Regel nur lokale Optima dar. Es empfiehlt sich, die Berechnungen mit unterschiedlichen Startpartitionen durchzuführen und die Ergebnisse miteinander zu vergleichen.

Der große Vorteil der partitionierenden Verfahren gegenüber den hierarchischen Verfahren liegt in dem deutlich geringeren Rechenaufwand.

Wenn in der Literatur von „Fuzzy-Clusterverfahren“ oder „unscharfer Klassifizierung“ die Rede ist, sind üblicherweise Verfahren gemeint, die auf der Basis der iterierten Minimaldistanzverfahren arbeiten (zum Beispiel DEIMER, 1986, HÖPPNER U.A., 1997, BANDEMER, 1997, TIMM, 2002). Der bekannteste Vertreter ist der im Abschnitt 4.3 behandelte „Fuzzy-c-Means-Algorithmus“. Im Kapitel 4 wird auf die Fuzzy-Clusterverfahren noch detaillierter eingegangen.

2.4.2.3 Graphentheoretische Verfahren

Bei den graphentheoretischen Verfahren existiert für die Objektmenge eine spezielle Topologie. Es werden hier nicht Distanzen zwischen beliebigen Objekten berechnet, sondern nur zwischen Objekten, die über eine Kante in dem zugehörigen Graphen verbunden sind. Der Graph kann sowohl schon im Objektraum existieren als auch erst im Merkmalsraum. Die Graphentheorie bietet für spezielle clusteranalytische Probleme sehr effektive Lösungen. Zum Beispiel gibt es auf Basis der Graphentheorie rechentechnisch sehr effektive Verfahren um Single Linkage Cluster zu erzeugen (ECKES und ROSSBACH, 1980). Graphentheoretische Clusterverfahren bieten sich zum Beispiel für die Klassifizierung raumbezogener Daten an, bei denen die Nachbarschaftsbeziehungen eine Rolle spielen. Einen speziellen Ansatz hierzu hat ANDERS (2004) dargestellt.

2.4.2.4 Alternative Optimierungsverfahren

Unter dieser Gruppe werden alle die Verfahren zusammengefasst, die nicht zu den bisher aufgeführten Verfahren gezählt werden können. Ein Beispiel aus dieser Gruppe sind die künstlichen neuronalen Netze, auf die im Rahmen dieser Arbeit aber nicht näher eingegangen werden soll.

2.4.3 Optimale Clusteranzahl

Die Bestimmung der optimalen Anzahl an Clustern für eine ausgewählte Problemstellung gestaltet sich in der Praxis oftmals schwierig. Genauso, wie die Clusterverfahren nur lokale Optima für die Clustereinteilung finden, gibt es auch keine „wahre“ Clusteranzahl. In der Regel erfolgt die Beurteilung der Clusteranzahl aufgrund subjektiver Kriterien. Entweder aus Kenntnis über die vorhandenen Daten oder der besseren Interpretierbarkeit der Ergebnisse.

Zum Finden einer optimalen Clusteranzahl gibt es insbesondere für die weit verbreiteten hierarchischen und partitionierenden Clusterverfahren einige zumindest aus theoretischer Sicht objektive Ansätze.

Bei den partitionierenden Clusterverfahren wird üblicherweise eine Zielfunktion oder ein Gütemaß optimiert. Der Wert des Gütemaßes (zum Beispiel des Varianzkriteriums (Formel 2.12)) und/oder der Zielfunktion, der sich bei einer fest vorgegebenen Clusteranzahl c für das erreichte lokale Optimum ergibt, ist in der Regel eine mit zunehmender Clusteranzahl fallende monotone Funktion $f(c)$. Da man bei den partitionierenden Verfahren zum Start eine feste Clusteranzahl vorgeben muss, kann man das Verfahren nacheinander mit jeweils unterschiedlichen Clusteranzahlen ausführen. Den Wert

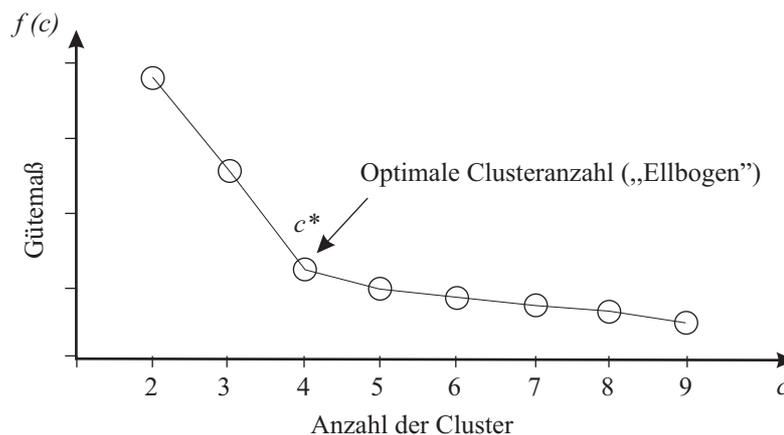


Abbildung 2.17: Bestimmung der optimalen Clusteranzahl c^* bei den partitionierenden Verfahren.

des erreichten Gütemaßes oder der Zielfunktion trägt man in einem Diagramm über der Clusteranzahl ab (Abbildung 2.17). Den optimalen Wert c^* für die Clusteranzahl findet man dort, wo der Kurvenverlauf einen erkennbaren Knick aufweist. Der Gedanke, der dahinter steht, ist, dass eine relativ große Veränderung von $f(c^* - 1)$ nach $f(c^*)$ zeigt, dass die Lösung mit $c^* - 1$ Clustern noch intern heterogene Gruppen enthält, während das bei der relativ geringen Veränderung von $f(c^*)$ nach $f(c^* + 1)$ nicht mehr der Fall ist (STEINHAUSEN und LANGER, 1977). Das Suchen nach diesem „Knick“ in der Funktion wird in der Literatur auch als „Ellbogen-Kriterium“ bezeichnet (BOCK, 1994).

Bei den hierarchischen Clusterverfahren führen ähnliche Überlegungen dazu, dass man annimmt, dass relativ große Intervalle des Distanzniveaus im Dendrogramm (Abbildung 2.14), in denen keine Fusionierungen oder Zerlegungen stattfinden, auf eine optimale Clusteranzahl hinweisen.

Die Praxis zeigt aber, dass diese theoretisch objektiven Kriterien teilweise zu schwer interpretierbaren Clusterlösungen führen. Von einer alleinigen Anwendung solcher Kriterien wird daher abgeraten. (BOCKLISCH, 1987), (HÖPPNER U.A., 1997)

Kapitel 3

Einführung in die Theorie unscharfer Mengen

Die Theorie unscharfer Mengen (engl. Fuzzy-Sets) wurde 1965 von LOTFI ZADEH (ZADEH, 1965) erstmals publiziert. Er gilt als geistiger Vater und Gründer dieser Theorie, die sowohl als Verallgemeinerung der klassischen Mengenlehre als auch der klassischen zweiwertigen Logik anzusehen ist. Die Fuzzy-Set Theorie liefert das mathematische Grundgerüst, um unscharfe Informationen mit Hilfe von Zugehörigkeitsfunktionen zu beschreiben, miteinander zu verknüpfen und zueinander in Beziehung zu setzen. Der Umgang mit unscharfen Informationen, der jedem Menschen intuitiv vertraut ist, kann mit Hilfe dieser Theorie formalisiert und menschliche Verhaltens- und Denkweisen können damit computertechnisch umgesetzt werden. (BURMEISTER, 1997)

Da sich diese Arbeit mit der unscharfen Clusterung von Daten beschäftigt, soll die Theorie der unscharfen Mengen nur soweit dargestellt werden, wie es für diesen Bereich notwendig ist. Andere Aspekte, wie insbesondere die unscharfe Logik (Fuzzy-Logic) und die Anwendung der Fuzzy Theorie in der Regelungstechnik (Fuzzy-Control) werden nicht weiter behandelt.

3.1 Grundlegendes zur Theorie unscharfer Mengen

Die Mengenlehre bezeichnet ein Teilgebiet der Mathematik und geht auf GEORG CANTOR (1845-1918) zurück. Von ihm stammt auch die Definition einer Menge (CANTOR, 1932):

„Unter einer **Menge** M verstehen wir jede Zusammenfassung von bestimmten, wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die Elemente von M genannt werden) zu einem Ganzen.“

Durch das Attribut „bestimmt“ ist diese Definition so zu verstehen, dass eindeutig entscheidbar ist, ob ein Objekt m zu einer Menge M gehört oder nicht. (BIEWER, 1997)

Eine sogenannte „Indikatorfunktion“ oder auch „charakteristische Funktion“ definiert die Zugehörigkeit eines Objektes m aus einer Grundmenge Ω zu solch einer Menge M :

$$f_M(m) = \begin{cases} 1 & \text{wenn } m \in M, \\ 0 & \text{wenn } m \notin M \end{cases} \quad \forall m \in \Omega \quad (3.1)$$

Das bedeutet, dass jedes Objekt m als Element der Menge M charakterisiert wird, indem ihm der Wert 1 zugewiesen wird. Jedes Objekt m , welches nicht Element von M ist, erhält den Funktionswert 0.

Die „klassische“ Mengenlehre ist von dieser dualen Zugehörigkeit geprägt. In ihr gibt es nur die Zustände:

- Ein Objekt m ist Element der Menge M ($m \in M$) oder
- ein Objekt m ist nicht Element der Menge M ($m \notin M$).

Die Elemente m_i einer endlichen Menge M können zum Beispiel durch eine Aufzählung angegeben werden.

$$M = \{m_1, m_2, m_3, m_4, m_5\} \quad (3.2)$$

Bei vielen praktischen Anwendungen stellte sich aber heraus, dass diese strikte Trennung nicht immer sachgerecht ist und die Frage, ob ein Element zu einer Menge gehört oder nicht, nicht immer eindeutig zu beantworten ist. „*More often than not, the classes of objects encountered in the real physical world do not have precisely defined criteria of membership.*“ (ZADEH, 1987)

Die Lösung dieses Problems führte über die auf LUKASIEWICZ (1876-1956) zurückgehende „mehrwertige Logik“ mit den Werten wahr (1), falsch (0) und möglich ($\frac{1}{2}$) (GOTTWALD, 1989) auf die **Fuzzy Theorie**, die LOTFI ZADEH 1965 (ZADEH, 1965) als Verallgemeinerung der Cantorschen Mengentheorie begründete.

Der englische Begriff „fuzzy“ bedeutet als Adjektiv in der deutschen Sprache soviel wie flaumig, flockig, kraus, struppig, zottig, unbestimmt oder verwischt. Er soll in der Fuzzy Theorie ausdrücken, dass die Grenzen der Mengen im Allgemeinen eben nicht scharf gezogen werden können, sondern unscharf sind. Fuzzy Mengen werden deshalb auch als unscharfe Mengen und als Abgrenzung dazu die „klassischen“ Mengen als scharfe Mengen bezeichnet.

In der Fuzzy Theorie sei \tilde{M} eine unscharfe Menge (engl. Fuzzy-Set). Ihre Indikatorfunktion lautet:

$$f_{\tilde{M}}(m) = \mu = \begin{cases} 1 & \text{wenn } m \in \tilde{M}, \\ \vdots & \\ 0 & \text{wenn } m \notin \tilde{M} \end{cases} \quad \forall m \in \Omega \quad (3.3)$$

Diese Funktion wird auch **Zugehörigkeitsfunktion** (engl. Membership Function) der unscharfen Menge \tilde{M} genannt. Sie wird in der Literatur in der Regel mit dem Buchstaben μ bezeichnet. Sie drückt aus, dass es neben den Werten 1 ($m \in M$) und 0 ($m \notin M$) noch beliebig viele Zwischenwerte gibt. Ein Element kann eben auch graduelle Zugehörigkeiten zu einer Menge haben. Der Übergang von Zugehörigkeit zu Nichtzugehörigkeit findet nicht abrupt statt, sondern in einem allmählichen Übergang (siehe Abbildung 3.1).

ZADEH (1965) hat vorgeschlagen, jedem Element m_i aus der Grundmenge Ω eine reelle Zahl im abgeschlossenen Intervall $[0, 1]$ als **Zugehörigkeitswert**¹ zur Menge \tilde{M} zuzuweisen².

¹ Man findet in der deutschsprachigen Literatur für den Begriff „Zugehörigkeitswert“ auch noch andere Begriffe wie z.B. „Zugehörigkeitsgrad“ (HÖPPNER U.A., 1997), oder „Sympathiewert“ (BOCKLISCH, 1987). In der englischsprachigen Literatur wird in der Regel der Begriff „degree of membership“ verwendet.

² Überwiegend werden die Zugehörigkeitswerte auf das Intervall $[0, 1]$ normiert. Grundsätzlich kann der Zugehörigkeitswert aber jede beliebige positive reelle Zahl annehmen. Ein Beispiel wird im Kapitel 5 noch behandelt.

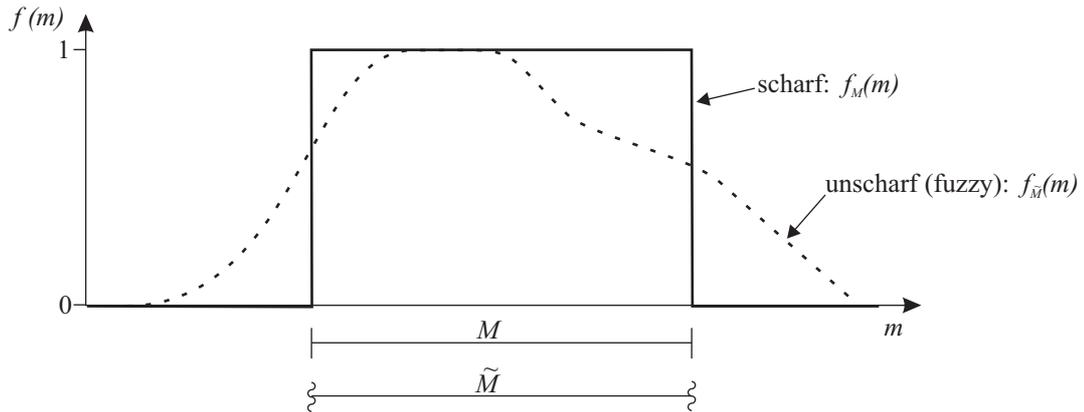


Abbildung 3.1: Zugehörigkeitsfunktionen einer scharfen und einer unscharfen Menge (nach BOCKLISCH, 1987).

Dieser Zugehörigkeitswert μ_i ist eine metrisch skalierte Eigenschaft eines Elementes m_i , mit dem es zu der entsprechenden Menge \tilde{M} gehört.

Während bei den scharfen Mengen die Menge durch Aufzählen ihrer Elemente angegeben werden kann, gehört bei einer diskreten Fuzzy Menge der entsprechende Zugehörigkeitswert des jeweiligen Elementes dazu. In dem folgenden Beispiel sind in runden Klammern jeweils Wertepaare angegeben, von denen der erste Wert das Element darstellt und der zweite den entsprechenden Zugehörigkeitswert.

Beispiel:

$$\tilde{M} = \{(m_1, 0.2), (m_2, 0.5), (m_3, 0.8), (m_4, 0.4), (m_5, 0.1)\} \tag{3.4}$$

Betrachtet man die Fuzzy Menge als kontinuierliche Fuzzy Menge, kann sie alleine durch ihre Zugehörigkeitsfunktion beschrieben werden.

Beispiel:

$$\tilde{M} = \{(m, \mu \mid m \in \mathbb{R}), \mu = \max(0, 1 - 0.1|m - 10|)\} \tag{3.5}$$

Kenngrößen einer unscharfen Menge (siehe Abbildungen 3.2 und 3.3) sind:

- Die **stützende Menge** oder der **Träger** (engl. Support) einer unscharfen Menge \tilde{M} ist gegeben durch:

$$\text{supp}(\tilde{M}) = \{m \in \tilde{M} \mid \mu(m) > 0\}.$$

- Die **Höhe** (engl. height) der unscharfen Menge \tilde{M} ist gegeben durch den Maximalwert (das Supremum) der Zugehörigkeitswerte:

$$\text{hgt}(\tilde{M}) = \sup_{m \in \tilde{M}} \mu(m)$$

Eine unscharfe Menge mit $\text{hgt} = 1$ heißt **normiert**.

- Der **Kern** (engl. core) einer unscharfen Menge \tilde{M} ist die Menge aller Elemente m , die den Zugehörigkeitswert $\mu(m)$ des Supremums haben:

$$\text{core}(\tilde{M}) = \{m \in \tilde{M} \mid \mu(m) = \text{hgt}(\tilde{M})\}$$

- Eine **Fuzzy Zahl** (unscharfe Zahl) ist eine unscharfe konvexe Menge \tilde{M} , wenn sie folgende Eigenschaften hat:

$$\mu(m) = \text{hgt}(\tilde{M}) \text{ für genau ein } m \in \mathbb{R} \text{ und } \mu(m) \text{ ist stückweise stetig.}$$

Die Zugehörigkeitsfunktion einer unscharfen Zahl besteht aus einem ansteigenden Ast, bis genau bei einem Element der Grundmenge die Zugehörigkeitsfunktion ihr Maximum annimmt. Daran anschließend fällt die Zugehörigkeitsfunktion wieder ab. Die Abbildung 3.2 zeigt das Beispiel einer unscharfen Zahl „ungefähr 10“. Das Beispiel soll ausdrücken, dass alle Zahlen in dem Intervall $[5, 15]$ ungefähr der Zahl $m = 10$ entsprechen und zwar umso mehr, umso kleiner die Distanz zu der Zahl $m = 10$ ist. Der maximale Zugehörigkeitswert ($\sup_{m \in \tilde{M}} \mu(m)$) befindet sich bei $m = 10$.

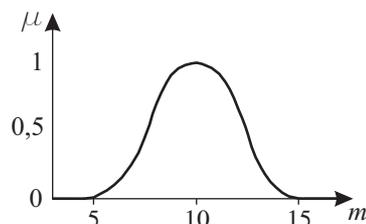


Abbildung 3.2: Beispiel einer unscharfen Zahl „ungefähr 10“.

- Der α -Schnitt einer unscharfen Menge \tilde{M} ist die (scharfe) Menge aller Elemente von \tilde{M} , deren Zugehörigkeitswert μ größer als ein Schwellwert α ist:

$$\tilde{M}_\alpha = \{m \in \tilde{M} \mid \mu(m) > \alpha\}$$

Der α -Schnitt dient insbesondere dazu, scharfe Mengen aus unscharfen Mengen abzuleiten.

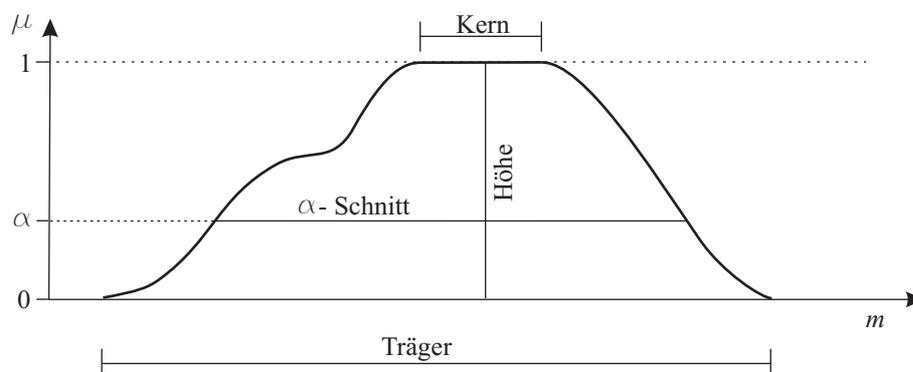


Abbildung 3.3: Kenngrößen einer unscharfen Menge.

3.2 Linguistische Variablen

Ein großer Vorteil der Fuzzy Theorie ist, dass mit dem Konzept der Zugehörigkeitswerte (oder der Zugehörigkeitsfunktionen) sehr gut Begriffe der Umgangssprache (sogenannte **linguistische Variablen**) beschrieben und einer mathematischen Behandlung zugeführt werden können. Linguistische Variablen sind Variablen, deren Werte nicht Zahlen, sondern Wörter, Phrasen oder Sätze einer natürlichen oder künstlichen Sprache sind. Die Werte dieser linguistischen Variablen bilden ihre Term-Menge (ZADEH, 1975). Beispiele sind in der Tabelle 3.1 aufgeführt.

In der geodätischen Praxis wird zum Beispiel das an sich scharfe Problem der Absteckung häufig unter Nutzung linguistischer Variablen durch den Messtrupp unscharf gelöst. Hier werden die abzusteckenden Maße oftmals nicht in einer Maßzahl ausgedrückt, sondern durch die Terme: „etwas vor“, „etwas mehr vor“, „ein ganzes Stück zurück“,

linguistische Variable	Terme
Alter	jung, mittel, alt
Betriebstemperatur	zu niedrig, gut, zu hoch
Größe	klein, mittel, groß
Zustand	sehr gut, gut, mittel, schlecht, sehr schlecht

Tabelle 3.1: Beispiele linguistischer Variablen und ihrer Terme

Ein Standardbeispiel zur Veranschaulichung des Konzepts der unscharfen Mengen und der linguistischen Variablen ist die Beschreibung unscharfer Altersangaben von Menschen (Abbildung 3.4). Ein Mensch, der zur Menge der jungen Menschen gehört, wird mit zunehmendem Alter nicht von heute auf morgen ein Mensch mittleren Alters, sondern es findet ein über mehrere Jahre dauernder (unscharfer) Übergang statt.

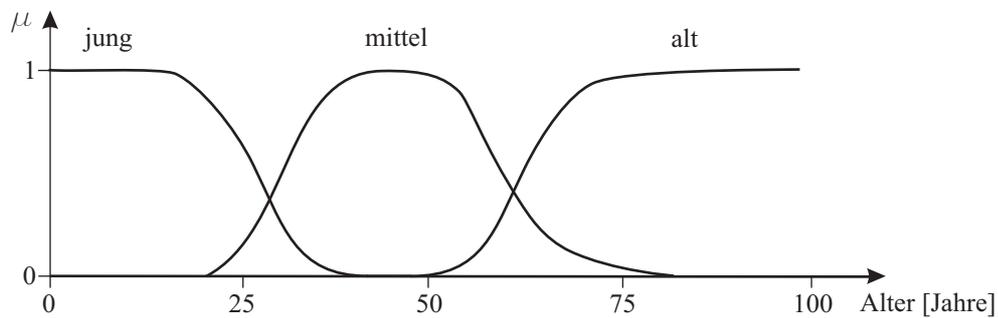


Abbildung 3.4: Beschreibung der Terme der linguistischen Variable „Alter“ durch unscharfe Zugehörigkeitsfunktionen.

Die unscharfe Menge (zum Beispiel der jungen Menschen) ist wie folgt zu interpretieren: Sie ist keine Anzahl von Personen, sondern die Funktion, die den Grad der Zugehörigkeit zu dieser Menge angibt.

Durch die Möglichkeit linguistische Variablen (beziehungsweise deren Terme) mit der Hilfe von Zugehörigkeitswerten oder -funktionen zu beschreiben, hat man eine sehr gute Möglichkeit, menschliche Erfahrungen, Aussagen, Eindrücke und Beobachtungen mathematisch zu formalisieren und für Problemstellungen heranzuziehen. (MANN, 1984)

In den Beispielen zu dieser Arbeit werden linguistische Variablen nicht als Eingangsgrößen verwendet, wie es zum Beispiel HEINE (1999), WIESER (2002) und HABERLER (2005) in ihren regelbasierten Fuzzy Anwendungen gemacht haben. In dem Beispiel „Automatische Klassifizierung von Setzungsmessungen“ dieser Arbeit (Abschnitt 6.3) werden die linguistischen Variablen (keine Setzung, beschleunigt zunehmende Setzung, lineare Setzung, konsolidierende Setzung) mit ihren Ausprägungen dagegen zur Ergebnispräsentation benutzt und erlauben dadurch eine bessere, detailliertere menschliche Interpretation als die Angabe von scharfen Klassenzuordnungen.

3.3 Modellierung von Zugehörigkeitsfunktionen

Soll das Konzept der unscharfen Mengen rechenstechnisch umgesetzt werden, sind die Zugehörigkeiten in irgendeiner Form zu modellieren. Ihre Beschreibung kann dabei durch eine Wertetabelle (nichtparametrisches Konzept) oder eine Funktion (parametrisches Konzept) erfolgen. Das nichtparametrische Konzept, in dem die Zugehörigkeitswerte diskreter Elemente durch einen numerischen Wert in einer Tabelle ausgedrückt werden, ist sehr flexibel. Unabhängig von Nachbarschaften kann jedem Element ein individueller Zugehörigkeitswert zugewiesen werden. Insbesondere bei der Bearbeitung höherdimensionaler Probleme wird dieses Konzept aber sehr aufwändig und unübersichtlich. Weitestgehend

bewährt haben sich parametrische Konzepte. Hierbei wird die Zugehörigkeit durch eine parametrisierte Funktion ausgedrückt. Die Parameter β_i werden dabei üblicherweise so gewählt, dass sie auch fachkundig interpretierbar sind. (BOCKLISCH, 1987)

Einige gebräuchliche parametrische Zugehörigkeitsfunktionen sollen in normierter Form im Folgenden kurz vorgestellt werden. Durch eine Erweiterung um zusätzliche Parameter können diese Grundformen noch modifiziert und insbesondere auch unsymmetrische Zugehörigkeitsfunktionen erzeugt werden.

Von ZADEH (1968) stammt der Vorschlag der sogenannten S - oder π -Funktionen (Abbildung 3.5). Die S -Funktion beschreibt dabei einen unscharfen Schwellwert und die π -Funktion, die sich aus zwei S -Funktionen zusammensetzt, eine unscharfe Zahl. Durch Einfügen eines Kernes mit dem Zugehörigkeitswert 1 in der Mitte lässt sich die π -Funktion leicht auf die Beschreibung eines unscharfes Intervalles erweitern.

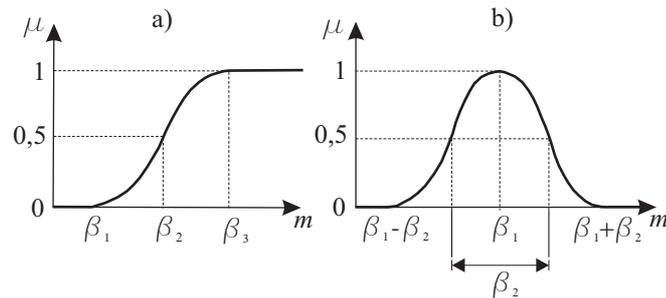


Abbildung 3.5: Zugehörigkeitsfunktionen vom Typ der a) S -Funktion und der b) π -Funktion (nach ZADEH, 1968).

Für die S -Funktion ($\beta_1 < \beta_2 < \beta_3$) gilt:

$$\mu_S(m, \beta_1, \beta_2, \beta_3) = \begin{cases} 0 & \text{für } m \leq \beta_1 \\ 2 \left(\frac{m - \beta_1}{\beta_3 - \beta_1} \right)^2 & \text{für } \beta_1 < m \leq \beta_2 \\ 1 - 2 \left(\frac{m - \beta_3}{\beta_3 - \beta_1} \right)^2 & \text{für } \beta_2 < m \leq \beta_3 \\ 1 & \text{für } m > \beta_3 \end{cases} \quad (3.6)$$

Den Parameter β_2 kann man sich als die Entsprechung eines scharfen Schwellwertes vorstellen und das Intervall β_1 bis β_3 als unscharfen Übergangsbereich.

Für die π -Funktion ($\beta_2 > 0$) gilt:

$$\mu(m, \beta_1, \beta_2) = \begin{cases} \mu_S(m, \beta_1 - \beta_2, \beta_1 - \frac{\beta_2}{2}, \beta_1) & \text{für } m \leq \beta_1 \\ 1 - \mu_S(m, \beta_1, \beta_1 + \frac{\beta_2}{2}, \beta_1 + \beta_2) & \text{für } m > \beta_1 \end{cases} \quad (3.7)$$

In einem weiten Anwendungsspektrum der Fuzzy-Theorie haben sich Zugehörigkeitsfunktionen bewährt, die auf einem stückweise linearen Verlauf basieren. Ihr großer Vorteil liegt in der einfachen rechen-technischen Handhabbarkeit, was sie insbesondere für Echt-Zeit-Anwendungen in der Regelungstechnik prädestiniert. Sie werden als Z -, Dreiecks³- oder Trapezfunktionen bezeichnet (Abbildung 3.6). Durch Einfügen eines Kernes mit dem Zugehörigkeitswert 1 in der Mitte, lässt sich auch die Dreiecksfunktion leicht auf die Beschreibung eines unscharfes Intervalles (Trapezfunktion) erweitern.

Für die Dreiecksfunktion ($\beta_2 > 0$) gilt:

$$\mu(m, \beta_1, \beta_2) = \max \left[0, 1 - \frac{1}{\beta_2} |m - \beta_1| \right] \quad (3.8)$$

³ Auch als Sägezahnfunktion bezeichnet.

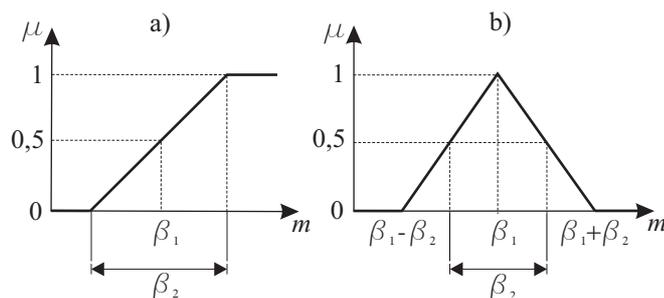


Abbildung 3.6: Stückweise lineare Zugehörigkeitsfunktionen vom Typ der a) Z-Funktion und der b) Dreiecksfunktion.

Für die Z-Funktion ($\beta_2 > 0$) gilt:

$$\mu(m, \beta_1, \beta_2) = \begin{cases} \max \left[0, 1 - \frac{1}{\beta_2} \left| m - \left(\beta_1 + \frac{\beta_2}{2} \right) \right| \right] & \text{für } m \leq \beta_1 + \frac{\beta_2}{2} \\ 1 & \text{für } m > \beta_1 + \frac{\beta_2}{2} \end{cases} \quad (3.9)$$

Ein weiterer Typ, der hier vorgestellt werden soll, ist der Potentialfunktionstyp (BOCK, 1974) (Abbildung 3.7). Er spielt in den im Abschnitt 4 behandelten unscharfen Clusterverfahren eine große Rolle. Im Gegensatz zu den bisher vorgestellten Typen ist der Potentialfunktionstyp im gesamten Definitionsbereich (hier: \mathbb{R}) echt größer als Null. Er bietet sich an, wenn ein Wirkungsbereich einer Erscheinung beschrieben werden soll, der sich im Allgemeinen nicht scharf abgrenzen lässt.

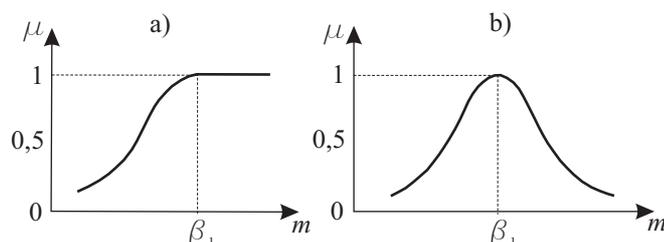


Abbildung 3.7: Zugehörigkeitsfunktionen vom Potentialfunktionstyp (nach BOCKLISCH, 1987).

Für die Potentialfunktion ($\beta_2 > 0, \beta_3 > 1$) gilt:

$$\mu(m, \beta_1, \beta_2, \beta_3) = \frac{1}{1 + \beta_2 |m - \beta_1|^{\beta_3}} \quad (3.10)$$

Der Parameter β_1 gibt, wie bei den bisher vorgestellten Typen, die Lage der Zugehörigkeitsfunktion an und die Parameter β_2 und β_3 geben die Unschärfe an. Durch den Parameter β_2 wird beim Potentialfunktionstyp die Spreizung der Funktion bestimmt und der Parameter β_3 beschreibt die Steilheit des Abfalls der Funktion zum Rande hin. Für $\beta_3 \rightarrow \infty$ strebt die Funktion gegen eine Rechteckfunktion, die der charakteristischen Funktion einer scharfen Menge entspricht.

3.4 Operationen mit unscharfen Mengen

Zur Verbindung und Modifizierung unscharfer Mengen gibt es die von den scharfen Mengen her bekannten Operationen: (ZIMMERMANN, 1993)

- Durchschnitt,
- Vereinigung und
- Komplement.

Gegeben sind zwei scharfe Mengen A und B , die Untermengen der Grundmenge Ω seien. Die **Durchschnittsmenge** $A \cap B$ (auch die Schnittmenge genannt) dieser scharfen Mengen A und B ist die Menge der Elemente, die zur Menge A **und** zur Menge B gehören.

$$A \cap B = \{m \in \Omega \mid m \in A \text{ und } m \in B\} \quad (3.11)$$

Die **Vereinigungsmenge** $A \cup B$ zweier scharfen Mengen A und B ist die Menge der Elemente, die zur Menge A **oder** zur Menge B gehören.

$$A \cup B = \{m \in \Omega \mid m \in A \text{ oder } m \in B\} \quad (3.12)$$

Wobei hier anzumerken ist, dass mit dem Begriff „oder“ das „logische oder“ gemeint ist, das auch die Elemente einschließt, die zu beiden Mengen gehören. Das „umgangssprachliche oder“ (auch „linguistisches oder“ genannt) schließt in der Regel diese Elemente aus.

Die **Komplementmenge** A^k einer Menge A über einer Grundgesamtheit Ω ist die Menge der Elemente, die zur Grundmenge Ω und **nicht** zur Menge A gehören.

$$A^k = \{m \mid m \in \Omega \text{ und } m \notin A\} \quad (3.13)$$

Bei den unscharfen Mengen finden die genannten Operationen ihre Definition und Anwendung auf Basis der Zugehörigkeitsfunktionen und nicht wie bei den scharfen Mengen auf der Basis der Elemente.

Für die Operationen Durchschnitt und Vereinigung unscharfer Mengen wurden folgende Operatoren definiert: (ZIMMERMANN, 1993)

- t-Normen,
- t-Conormen und
- mittelnde Operatoren.

Die **t-Normen** sind mathematische Modelle für den mengentheoretischen **Durchschnitt** beziehungsweise das „logische Und“. Eine t-Norm ist eine kommutative, assoziative und monoton wachsende Funktion in 2 Variablen $t : [0, 1] \times [0, 1] \rightarrow [0, 1]$, mit dem neutralen Element 1 und dem Nullelement 0. Eine t-Norm ordnet 2 Zugehörigkeitswerten $\mu_1, \mu_2 \in [0, 1]$ einen geeigneten Wert $t(\mu_1, \mu_2) \in [0, 1]$ zu.

Die **Zugehörigkeitsfunktion der Schnittmenge** $\mu_{\tilde{A} \cap \tilde{B}}(m)$ zweier unscharfer Mengen \tilde{A} und \tilde{B} mit den Zugehörigkeitsfunktionen $\mu_{\tilde{A}}(m)$ und $\mu_{\tilde{B}}(m)$ kann zum Beispiel punktweise durch den zu den t-Normen gehörenden **Minimum-Operator** definiert werden (siehe Abbildung 3.8).

$$\mu_{\tilde{A} \cap \tilde{B}}(m) = \min(\mu_{\tilde{A}}(m), \mu_{\tilde{B}}(m)) \quad \forall m \in \Omega \quad (3.14)$$

Die **t-Conormen** (auch S-Normen genannt) sind mathematische Modelle für die mengentheoretische **Vereinigung** beziehungsweise das „logische Oder“. Die t-Conormen sind ebenfalls kommutative, assoziative und monoton wachsende Funktion in 2 Variablen $s : [0, 1] \times [0, 1] \rightarrow [0, 1]$, allerdings mit dem

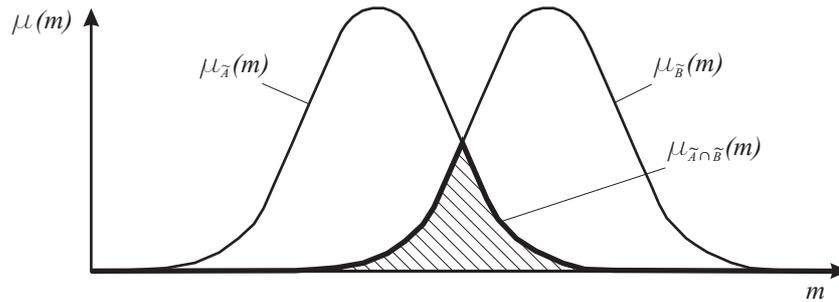


Abbildung 3.8: Der Durchschnitt zweier unscharfer Mengen mit Hilfe des Minimum-Operators.

neutralen Element 0 und dem Nullelement 1. Sie ordnen ebenfalls 2 Zugehörigkeitswerten $\mu_1, \mu_2 \in [0, 1]$ einen geeigneten Wert $s(\mu_1, \mu_2) \in [0, 1]$ zu.

Die **Zugehörigkeitsfunktion der Vereinigungsmenge** $\mu_{\tilde{A} \cup \tilde{B}}(m)$ zweier unscharfer Mengen \tilde{A} und \tilde{B} mit den Zugehörigkeitsfunktionen $\mu_{\tilde{A}}(m)$ und $\mu_{\tilde{B}}(m)$ kann zum Beispiel punktweise durch den zu den t-Conormen gehörenden **Maximum-Operator** definiert werden (siehe Abbildung 3.9).

$$\mu_{\tilde{A} \cup \tilde{B}}(m) = \max(\mu_{\tilde{A}}(m), \mu_{\tilde{B}}(m)) \quad \forall m \in \Omega \quad (3.15)$$

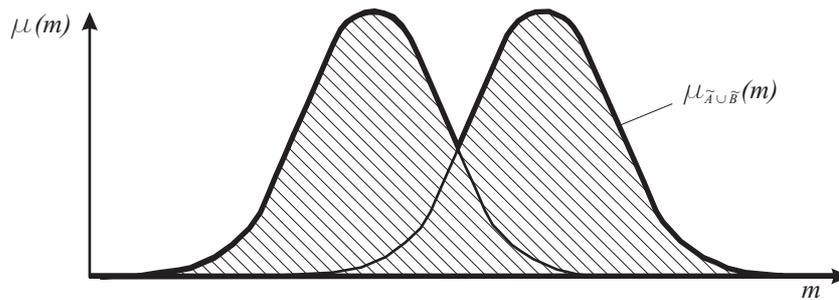


Abbildung 3.9: Die Vereinigung zweier unscharfer Mengen mit Hilfe des Maximum-Operators.

Diese bisher aufgeführten Grundoperationen für unscharfe Mengen wurden bereits von ZADEH (1965) im Zusammenhang mit der Veröffentlichung seiner „Fuzzy-Theorie“ vorgeschlagen. In der Praxis hat sich aber gezeigt, dass der Minimum- und der Maximum-Operator für den Durchschnitt und die Vereinigung in dieser Form nicht immer befriedigende Ergebnisse liefern, die der menschlichen Vorstellung entsprechen. Das „und“ bzw. das „oder“ das der Mensch in seiner normalen Kommunikation benutzt, hat für ihn inhaltlich nicht immer die gleiche Bedeutung wie das „logische und“ bzw. das „logische oder“. Deshalb wurden in den weiteren Jahren eine ganze Reihe von alternativen Möglichkeiten publiziert.

Die im Folgenden vorgestellten Beispiele für t-Normen können noch mit Hilfe eines Parameters γ individuell angepasst werden. Deshalb spricht man hier von parametrischen Normen oder von Operator Familien. Im Zusammenhang mit dieser Arbeit spielen die parametrischen t-Normen eine Rolle bei der Erweiterung von ein- auf mehrdimensionale Zugehörigkeitsfunktionen.

Für die Operator Familie T_D von DUBOIS und PRADE (1982) gilt:

$$T_D(\mu_{\tilde{A}}(m), \mu_{\tilde{B}}(m), \gamma) = \frac{\mu_{\tilde{A}}(m) \cdot \mu_{\tilde{B}}(m)}{\max(\mu_{\tilde{A}}(m), \mu_{\tilde{B}}(m), \gamma)}, \quad \gamma \in [0, 1], \quad (3.16)$$

Für die Operator Familie T_H von HAMACHER (1978) gilt:

$$T_H(\mu_{\tilde{A}}(m), \mu_{\tilde{B}}(m), \gamma) = \frac{\mu_{\tilde{A}}(m) \cdot \mu_{\tilde{B}}(m)}{\gamma + (1 - \gamma)(\mu_{\tilde{A}}(m) + \mu_{\tilde{B}}(m) - \mu_{\tilde{A}}(m)\mu_{\tilde{B}}(m))} \quad \text{für } \gamma \geq 0, \quad (3.17)$$

Für die Operator Familie T_Y von YAGER (1980) gilt:

$$T_Y(\mu_{\bar{A}}(m), \mu_{\bar{B}}(m), \gamma) = 1 - \min\{1, ((1 - \mu_{\bar{A}}(m))^\gamma + (1 - \mu_{\bar{B}}(m))^\gamma)^{1/\gamma}\} \text{ für } \gamma \geq 1, \quad (3.18)$$

Neben den t-Normen und den t-Conormen gibt es noch sogenannte **mittelnde** oder **kompensatorische Operatoren**. Weder die t-Normen noch die t-Conormen können Werte zwischen dem kleineren Zugehörigkeitswert (Minimum) und dem größeren Zugehörigkeitswert (Maximum) annehmen. Wenn man sich dagegen menschliche Entscheidungen anschaut, kann man feststellen, dass zum Beispiel niedrige Bewertungen eines Merkmals durch hohe Bewertungen eines anderen Merkmals ausgeglichen („kompensiert“) werden können. Dadurch liegt die Gesamtbewertung oftmals irgendwo in der Mitte zwischen den Einzelbewertungen, also in diesem Fall zwischen dem Minimum und dem Maximum. Für diese Fälle sind die kompensatorischen Operatoren gedacht. (BIEWER, 1997)

Die Abbildung 3.10 zeigt die Laufbereiche der parametrischen t-Normen und t-Conormen sowie der kompensatorischen Operatoren im reellen Einheitsintervall. Für die dargestellten Zugehörigkeitswerte soll gelten: $\mu_{\bar{A}}(m) < \mu_{\bar{B}}(m)$. In der Abbildung ist zu sehen, dass die t-Normen maximal den kleineren Zugehörigkeitswert und die t-Conormen minimal den größeren Zugehörigkeitswert erreichen. Die kompensatorischen Operatoren dagegen ergeben Werte zwischen dem kleineren und dem größeren Zugehörigkeitswert.

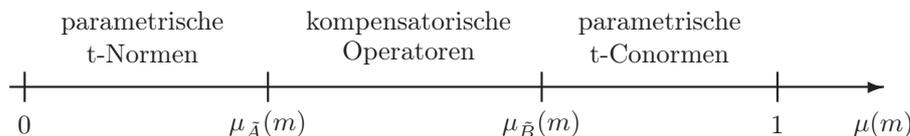


Abbildung 3.10: Laufbereiche für Durchschnitts- und Vereinigungsoperatoren (nach JAANINEH und MAIJOHANN, 1996).

Die Auswahl eines bestimmten Operators und die Einstellung der Parameter hängt von der individuellen Aufgabenstellung ab. Hinweise hierzu sind zum Beispiel in ZIMMERMANN (1993) zu finden.

Auf Beispiele zu den t-Conormen und den mittelnden Operatoren wird hier verzichtet, da sie im Rahmen dieser Arbeit keine Anwendung finden.

3.5 Mehrdimensionale Zugehörigkeitsfunktionen

Zugehörigkeitsfunktionen sind nicht auf eindimensionale Probleme begrenzt. Sie können auch ein Gebiet in einem höherdimensionalen Raum unscharf beschreiben. Die Abbildung 3.11 zeigt ein Beispiel für einen zweidimensionalen Fall.

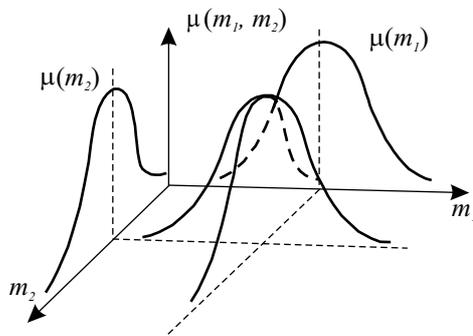


Abbildung 3.11: Zweidimensionale Zugehörigkeitsfunktion (nach BOCKLISCH, 1987).

Bevor mehrere eindimensionale Zugehörigkeitsfunktionen, die auf verschiedenen Achsen definiert sind, miteinander zu einer höherdimensionalen Zugehörigkeitsfunktion verknüpft werden können, müssen sie vorab zylindrisch erweitert werden. (KAHLERT und FRANK, 1994)

Für die Verknüpfung zweier eindimensionaler Zugehörigkeitsfunktionen $\mu(m_1)$ und $\mu(m_2)$ miteinander ist es notwendig, jedem m_1 -Wert einen Zugehörigkeitswert $\mu(m_2)$ zuzuordnen und umgekehrt. Dadurch wird jede eindimensionale Zugehörigkeitsfunktion in die jeweils zweite Dimension erweitert.

Allgemein gilt, dass eine Zugehörigkeitsfunktion $\mu(\mathbf{m}) = \mu(m_1, \dots, m_j)$, welche über einem Raum R^j definiert ist, mittels der zylindrischen Erweiterung in eine Zugehörigkeitsfunktion

$$\mu^{ze}(\mathbf{m}) = \mu(m_1, \dots, m_j, m_{j+1}, \dots, m_k) \tag{3.19}$$

über dem Raum $R^k (k > j)$ erweitert werden kann, wobei

$$\mu^{ze}(\mathbf{m}) = \mu(\mathbf{m}) = \mu(m_1, \dots, m_j) \tag{3.20}$$

für beliebige $(m_{j+1}, \dots, m_k) \in R^{k-j}$ gilt.

Die Abbildung 3.12 zeigt zwei Zugehörigkeitsfunktionen $\mu(m_1)$ und $\mu(m_2)$ mit ihren zylindrischen Erweiterungen $\mu^{ze}(m_1)$ und $\mu^{ze}(m_2)$ und die anschließende t-Norm Verknüpfung (hier mit Hilfe des Minimum-Operators) zu der zweidimensionalen Zugehörigkeitsfunktion $\mu(m_1, m_2)$.

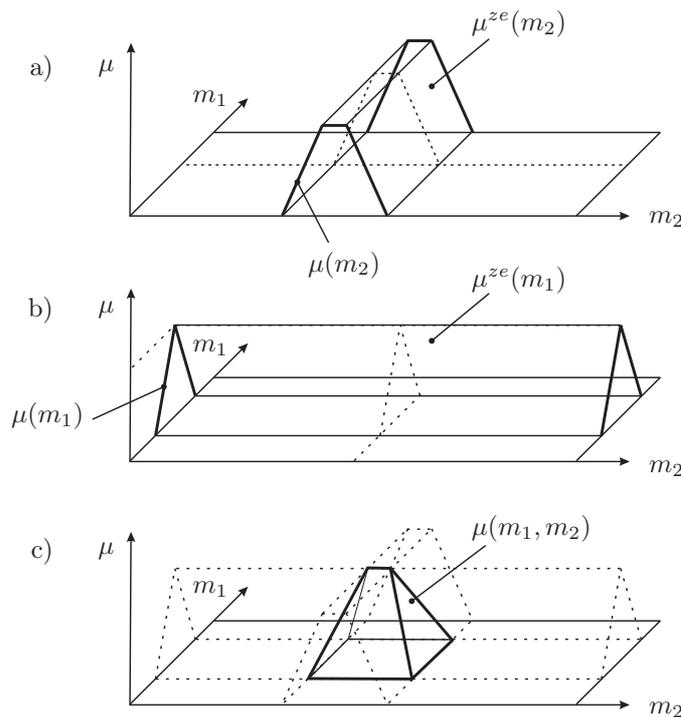


Abbildung 3.12: Zylindrische Erweiterungen der Zugehörigkeitsfunktionen $\mu(m_1)$ und $\mu(m_2)$ in a) und b) und in c) die anschließende t-Norm Verknüpfung zu der zweidimensionalen Zugehörigkeitsfunktion $\mu(m_1, m_2)$. (nach KAHLERT und FRANK, 1994).

Kapitel 4

Fuzzy-Clusterverfahren

4.1 Vorbemerkungen zur traditionellen Behandlung von Unsicherheiten in der Geodäsie

Muster und Merkmale sind nicht nur in der geodätischen Welt dadurch gekennzeichnet, dass sie eine bestimmte Unsicherheit aufweisen. Weil wir als Geodäten glauben, dass das sinnvoll ist, behandeln wir traditionell diese Unsicherheit möglichst vollständig in der Datenanalyse und Parameterschätzung mit. In anderen Fachdisziplinen ist das allerdings nicht überall so (BANDEMER, 1997).

Das klassische Modell der Betrachtung von Unsicherheiten in der Geodäsie basiert dabei auf drei Arten von sogenannten Fehlern (z.B. KAHMEN, 2006):

- **Grobe Fehler** sind Fehler im Sinne von umgangssprachlich „falschen“ Messungen. Falsche Ableesungen, Punktverwechslungen, Protokollierungsfehler, Fehlfunktionen von automatischen Messgeräten und dergleichen. Diese groben Fehler sind zu vermeiden bzw. aufzudecken und vor der Weiterverarbeitung der Daten zu eliminieren.
- **Systematische Fehler** sind die Messabweichungen, die dafür sorgen, dass der Messwert immer kleiner oder immer größer als der (unbekannte) wahre Wert ist. Dadurch verfälschen sie das Messergebnis stets in die gleiche Richtung. Sie sind begründet in einer unzureichenden Justierung, Kalibrierung oder einseitigen Handhabung der Messinstrumente sowie nicht beherrschter oder unbekannter einseitiger Einflüsse auf das Messergebnis. Die systematischen Messabweichungen sollten durch sorgfältige Justierung bzw. Kalibrierung der Messinstrumente, spezielle Messanordnungen sowie durch rechnerisches Berücksichtigen einseitiger Einflüsse eliminiert werden.
- **Zufällige Fehler** sind die Messabweichungen, die übrig bleiben, wenn die ursprünglichen Messungen von den groben Fehlern und den systematischen Abweichungen bereinigt sind. Es wird angenommen, dass die zufälligen Fehler ebenso oft ein positives wie negatives Vorzeichen annehmen und zufällig um den wahren Wert streuen. In der weiteren Datenanalyse und Parameterschätzung werden sie traditionell als (zumindest asymptotisch) normalverteilte Zufallsgrößen aufgefasst.

Insbesondere zur besseren Verständigung im interdisziplinären Raum spricht man heute bei den systematischen und zufälligen Fehlern nicht mehr von „Fehlern“, sondern von „Messabweichungen“ (DIN, 1998).

Dadurch dass man versucht, die groben und systematischen Messabweichungen zu eliminieren und man davon ausgeht, dass die übrig bleibenden Messabweichungen aus Zufallsgrößen bestehen, werden traditionell die Methoden der Stochastik zur Behandlung der Unsicherheit in der Geodäsie genutzt¹.

Dieses klassische Modell in der Geodäsie und seine Behandlung in der Theorie und Praxis wird inzwischen oftmals kritisch gesehen (z.B. CASPARY, 1988 oder WENDERLEIN, 1998). Eine sehr umfangreiche Diskussion der Fehlertheorie in der Geodäsie und möglicher Alternativen findet man in KUTTERER (2002). Der Hauptkritikpunkt ist der, dass es nicht gelingt, alle systematischen Messabweichungen zu eliminieren, da man sie nicht kennt, beherrscht bzw. mathematisch vollständig modellieren kann und daher immer Restsystematiken übrig bleiben. Diese werden anschließend fälschlicherweise als zufällige Abweichungen behandelt und führen dadurch (neben einem systematisch falschen Ergebnis) zu zu optimistischen Schätzungen für die Unsicherheit abgeleiteter Größen.

KUTTERER (2002) beschreibt (basierend insbesondere auf BANDEMER (1997) und ZIMMERMANN (1999)), dass es neben der Stochastik noch verschiedene andere mathematische Methoden zum Umgang mit der Unsicherheit von Daten gibt. Er zählt hierzu die folgenden auf:

- Approximationstheorie,
- Bayes-Theorie,
- Evidenztheorie,
- Intervallmathematik und
- Fuzzy-Theorie.

In dieser Arbeit wird die Unsicherheit in der Mustererkennung mit Hilfe der Fuzzy-Theorie, durch unscharfe (fuzzy) Klassenbeschreibungen und darauf folgende unscharfe Klassifizierungen der Muster (respektive Objekte), behandelt. Auf die anderen mathematischen Methoden wird nicht näher eingegangen. Die Fuzzy-Clusterverfahren, die zunächst vorgestellt werden, bieten keine Möglichkeit, die Unsicherheit der Merkmale mit zu berücksichtigen. Die Merkmale werden bei ihnen im klassisch geodätischen Sprachgebrauch als „fehlerfrei“ betrachtet. Die unscharfe Klassenbeschreibung ergibt sich nur aus den unterschiedlichen Distanzen zwischen den Objekten und den Clusterzentren im Merkmalsraum. Bei dem später vorzustellenden Verfahren der Fuzzy Pattern Klassifikation geht dagegen auch die Unschärfe der Merkmale mit in die unscharfen Klassenbeschreibungen ein.

4.2 Vorbemerkungen zur unscharfen Klassifikation

Die im Abschnitt 2.4.2 beschriebenen (scharfen) Clusterverfahren ordnen ein Objekt genau einem bestimmten Cluster zu. So eine strenge Zuordnung ist aber aus verschiedenen Gründen in vielen Fällen nicht sachgerecht, da z.B. nur unscharfes und/oder unsicheres Wissen über das Objekt vorliegt. Man denke nur an Objekte, die in direkter Nähe einer Clustergrenze liegen. Diese Objekte können oftmals genauso gut dem einen wie auch dem anderen Cluster zugeordnet werden. Die unscharfen Clusterverfahren ordnen dagegen ein Objekt nicht genau einem Cluster zu, sondern es werden für jedes Objekt Zugehörigkeitswerte zu jedem Cluster berechnet. Objekte, deren Merkmale eine große Ähnlichkeit mit den Merkmalen eines Clusterzentrums aufweisen, bekommen einen relativ hohen Zugehörigkeitswert zu diesem Cluster, Objekte deren Merkmale eine geringere Ähnlichkeit aufweisen, bekommen einen relativ niedrigen Zugehörigkeitswert. Das Ergebnis der Clusterung ist für jedes Objekt keine eindeutige Zuordnung, sondern bei c Clustern ein Vektor von c Zugehörigkeitswerten. Bei mehreren zu klassifizierenden Objekten erhält man eine Zugehörigkeitsmatrix wie sie beispielhaft in der Tabelle 4.1 zu sehen ist. In der Abbildung 4.1 sind beispielhaft die Zugehörigkeitsfunktionen von drei unscharfen Clustern über einem zweidimensionalen Merkmalsraum dargestellt.

¹ Man kann es auch anders herum betrachten: Da man traditionell die Stochastik zur Behandlung der Unsicherheit in der Geodäsie nutzt, müssen die groben und systematischen Fehler vorher eliminiert werden.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Objekt 1	0,1	0,7	0,1	0,1
Objekt 2	0,3	0,2	0,1	0,4
Objekt 3	0,1	0,4	0,4	0,1
Objekt 4	0,2	0,6	0,1	0,1
Objekt 5	0,1	0,1	0,2	0,6
Objekt 6	0,1	0,2	0,4	0,3

Tabelle 4.1: Beispiel einer Zugehörigkeitsmatrix von 6 Objekten zu 4 Clustern.

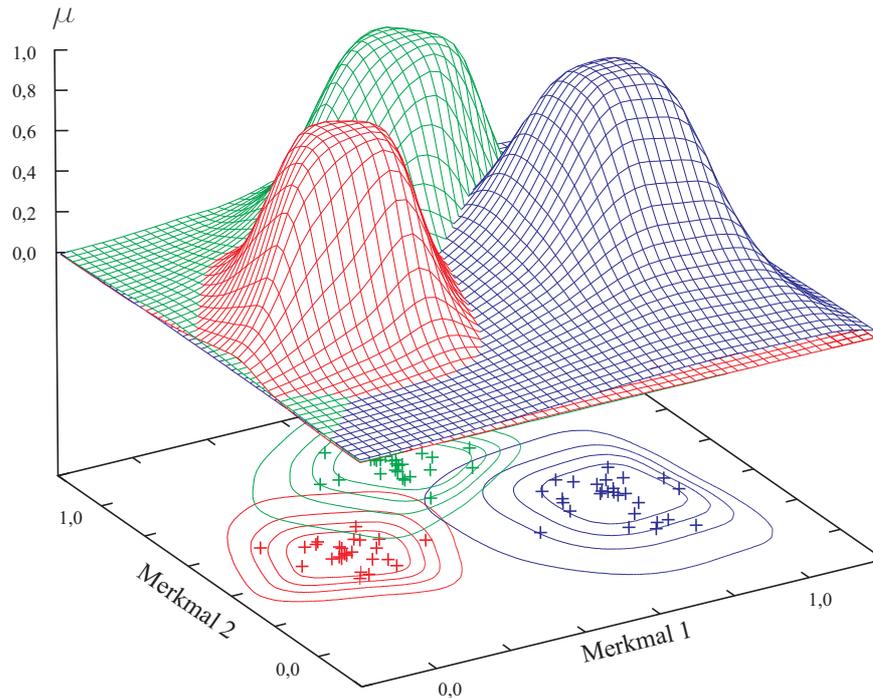


Abbildung 4.1: Die Zugehörigkeitsfunktionen von drei unscharfen Clustern über einem zweidimensionalen Merkmalsraum.

Bei vielen Clusterverfahren, z.B. dem Fuzzy-c-Means Algorithmus (siehe Abschnitt 4.3), wird vorausgesetzt, dass die Summe der Zugehörigkeitswerte eines Objektes zu allen Clustern 1 ergeben muss.

$$\sum_{i=1}^c \mu_{i,k} = 1 \quad \forall k = 1, \dots, n \tag{4.1}$$

Und für die Zugehörigkeitswerte wird üblicherweise festgelegt:

$$\mu_{i,k} \in [0, 1] \quad \forall i = 1, \dots, c; \quad \forall k = 1, \dots, n. \tag{4.2}$$

- $\mu_{i,k}$: Zugehörigkeitswert des k -ten Objekts zum i -ten Cluster
- n : Anzahl der Objekte
- c : Anzahl der Cluster

Durch die Restriktion in Gleichung (4.1) erinnern die Zugehörigkeitswerte stark an eine Wahrscheinlichkeitsverteilung. Die Fuzzy-Clusterverfahren, für die diese Restriktion gilt, bezeichnet man deshalb auch als *probabilistisch* (HÖPPNER U.A., 1997). Die Zugehörigkeitswerte $\mu_{i,k}$ sind allerdings nicht als Werte einer Wahrscheinlichkeitsverteilung zu interpretieren, sondern sie geben nur an, wie die Variablenausprägungen des k -ten Objektes mit den Charakteristika des i -ten unscharfen Clusters (in Relation zu den anderen Clustern) übereinstimmen (DEIMER, 1986).

Neben diesen probabilistischen Verfahren gibt es noch die Gruppe der *possibilistischen* Verfahren (KRISHNAPURAM und KELLER, 1993), bei denen auf die Bedingung, dass die Summe der Zugehörigkeitswerte eines Objektes über alle Cluster eins ergeben muss, verzichtet wird. Im Abschnitt 4.5 wird hierauf noch näher eingegangen.

4.3 Der Fuzzy-c-Means Algorithmus

Der Fuzzy-c-Means (FCM) Algorithmus ist sicherlich der am weitesten verbreitete Algorithmus zur unscharfen Clusterung von Objekten. In praktisch jeder Literaturquelle zu diesem Thema wird er vorgestellt. In vielen Quellen sogar als der einzige Vertreter. Auch in geodätischen Veröffentlichungen wurde er bereits erwähnt (MIIMA, 2002), beziehungsweise angewendet (QU, 2000). Er hat sich in der Praxis bei einer Vielzahl von Aufgabenstellungen als ein sehr robustes und stabiles Clusterverfahren bewährt (TIMM, 2002).

Ausgehend von einem scharfen Clusterungsverfahren (dem ISODATA-Verfahren nach BALL und HALL (1967), auch Hard-c-Means genannt (BEZDEK, 1980)) entwickelte BEZDEK (1973a) den Fuzzy-c-Means Algorithmus. Das c im Namen steht dabei stellvertretend für die Anzahl der Cluster und der Begriff *Means* steht für eine verallgemeinerte Mittelwertbildung.

Beim Fuzzy-c-Means Algorithmus wird eine Zielfunktion (engl. object function) J_{FCM} unter Vorgabe einer Clusteranzahl und unter Beachtung der beiden Nebenbedingungen $\sum_{i=1}^c \mu_{i,k} = 1$ (4.1) und $\mu_{i,k} \in [0, 1]$ (4.2) minimiert:

$$J_{FCM} = \sum_{k=1}^n \sum_{i=1}^c (\mu_{i,k})^w \cdot d_{i,k}^2(\boldsymbol{\nu}_i, \mathbf{x}_k) \quad (4.3)$$

- w : Grad der Unschärfe („Fuzzyfier“)
- $d_{i,k}$: (Euklidische) Distanz
- $\boldsymbol{\nu}_i$: Schwerpunkt des i -ten Clusters
- \mathbf{x}_k : k -tes Objekt

Der „Fuzzyfier“ $w \in [1, \infty[$ bestimmt den Grad der Unschärfe des Clusterbildungsprozesses. Je größer w ist, desto unschärfer fallen die Ergebnisse aus. Für $w \rightarrow 1$ ergibt sich eine scharfe Clustereinteilung und für $w \rightarrow \infty$ wird jedes Objekt zu gleichen Teilen jedem Cluster zugeordnet. In der Praxis haben sich Werte für w von 2 oder 3 bewährt (DEIMER, 1986). In der Regel wird der Wert $w = 2$ gewählt (TILLI, 1993).

Für die Berechnung der Distanz $d_{i,k}$ verwendet man üblicherweise die Euklidische Distanz.

Die Anzahl der Cluster muss vorgegeben werden. Zu ihrer Bestimmung können bei niedrigdimensionalen Merkmalsräumen (≤ 3 Dimensionen) Visualisierungshilfsmittel nützlich sein, bei höheren Dimensionen wendet man zum Beispiel die in Abschnitt 2.4.3 beschriebene Vorgehensweise an. In Abschnitt 4.6 wird hierauf auch noch einmal näher eingegangen.

Eine direkte analytische Lösung des Optimierungsproblems gibt es nicht. Die Minimierung der Zielfunktion erfolgt daher in einem iterativen Prozess, in dem abwechselnd die Zugehörigkeitswerte $\mu_{i,k}$ und die Clusterschwerpunkte ν_i optimiert werden.

Ablauf des Fuzzy-c-Means Algorithmus:

0. Schritt:

Vorgabe einer Clusteranzahl c , eines Unschärfeparameters w und Bestimmung von Näherungswerten für die Clusterschwerpunkte ν_i .

1. Schritt:

Berechnung von Zugehörigkeitswerten $\mu_{i,k}$ der einzelnen Objekte zu den Clusterschwerpunkten ν_i .

$$\mu_{i,k} = \frac{1}{\sum_{j=1}^c \left(\frac{d_{i,k}^2(\nu_i, \mathbf{x}_k)}{d_{j,k}^2(\nu_j, \mathbf{x}_k)} \right)^{\frac{1}{w-1}}}, \quad \forall i = 1, \dots, c; \quad \forall k = 1, \dots, n \quad (4.4)$$

2. Schritt:

Aus den aktuellen Zugehörigkeitwerten $\mu_{i,k}$ werden neue Clusterschwerpunkte ν_i^{neu} berechnet.

$$\nu_i^{neu} = \frac{\sum_{k=1}^n (\mu_{i,k})^w \cdot \mathbf{x}_k}{\sum_{k=1}^n (\mu_{i,k})^w}, \quad \forall i = 1, \dots, c \quad (4.5)$$

Die Gleichung (4.5) ist eine gewichtete Mittelbildung der n Objekte \mathbf{x}_k . Die Zugehörigkeitswerte fungieren hier als Gewichte. Dadurch wird bewirkt, dass Objekte mit hohen Zugehörigkeitswerten zu dem entsprechenden unscharfen Cluster einen größeren Einfluss auf die Lage des dazugehörenden Schwerpunktes haben als solche mit niedrigen Zugehörigkeitswerten.

3. Schritt:

Es wird geprüft, ob die Summe der komponentenweisen Distanzen zweier aufeinander folgender Clusterschwerpunktmatrizen \mathbf{N} kleiner ist als eine Konvergenzschwelle ϵ .

Falls $\|\mathbf{N}^{neu} - \mathbf{N}\| > \epsilon$, setze $\mathbf{N} = \mathbf{N}^{neu}$ und fahre fort mit Schritt 1. Andernfalls beende die Iteration.

Für den Fall, dass ein Objekt mit einem Clusterschwerpunkt identisch ist, ergibt sich eine entartete Lösung. In diesem Fall wird das Objekt dem entsprechenden Cluster scharf zugewiesen.

Der Iterationsprozess kann alternativ mit einer näherungsweise aufgestellten Zugehörigkeitsmatrix begonnen und die Schritte 1 und 2 in umgekehrter Reihenfolge durchgeführt werden. Als Abbruchkriterium ist es auch möglich, im Schritt 3 die Euklidische Distanz der Zugehörigkeitsmatrizen der letzten zwei Iterationen zu prüfen.

BEZDEK (1980) hat gezeigt, dass die Iterationsfolge des Algorithmus konvergiert. Es ist aber nicht bewiesen, dass das globale Minimum gefunden wird. Abhängig von den Näherungswerten können eventuell nur lokale Minima gefunden werden. Daher kann es sinnvoll sein, den Algorithmus mit unterschiedlichen Initialisierungen auszuführen und die Ergebnisse zu bewerten. Die Gefahr ein lokales Minimum zu erreichen, besteht aber durch die unscharfe Zuteilung der Objekte zu den Clustern nur in einem geringeren Maß als bei einer scharfen Zuteilung.

Liegen feste Clusterschwerpunkte vor, zum Beispiel aus einer Trainingsphase mit einer repräsentativen Stichprobe von Objekten oder aus Expertenwissen, können für neue Objekte nach Gleichung (4.4) in einer Identifikationsphase (manchmal auch Arbeitsphase genannt) Zugehörigkeitswerte berechnet werden. (siehe auch Abbildung 2.1)

Der größte Nachteil dieses Algorithmus ist jedoch, dass unabhängig von der tatsächlichen Verteilung der Objekte im Merkmalsraum, er bei Verwendung des euklidischen Abstandes hyperkugelförmige Cluster bildet, die alle die gleiche Größe haben. „Die Grenzen seiner Leistungsfähigkeit erreicht der Fuzzy-c-Means bei Clustern unterschiedlicher Form, Größe und Dichte.“ (HÖPPNER U.A., 1997). Ein aus geodätischer Sicht weiterer Nachteil liegt darin, dass er die Unschärfe der zu clusternden Objekte nicht mit berücksichtigen kann. Die Merkmale der Objekte werden als „fehlerfrei“ angenommen.

4.4 Modifizierte Fuzzy-c-Means Algorithmen

Um den Nachteil des Fuzzy-c-Means Algorithmus bezüglich der Bildung von hyperkugelförmigen Clustern auszugleichen, wurden im Laufe der Zeit Modifikationen entwickelt. Als Beispiele werden der Gustafson-Kessel Algorithmus und der Gath-Geva Algorithmus vorgestellt. Der iterative Berechnungsablauf, in dem abwechselnd die Zugehörigkeitswerte $\mu_{i,k}$ und die Klassenschwerpunkte ν_i optimiert werden, wird bei beiden Algorithmen beibehalten. Es werden allerdings gegenüber dem Fuzzy-c-Means Algorithmus andere Distanzfunktionen verwendet.

In HÖPPNER U.A. (1997) sind noch weitere Fuzzy-Clusterungsverfahren angegeben, die im Wesentlichen durch Modifikation der Distanzfunktion für die Erkennung ganz spezieller Strukturen (z.B. Geraden, Ebenen, Kreisränder, Rechtecke, Polygonzüge) ausgelegt sind. Da diese Verfahren oftmals sehr spezielle Vorkenntnisse über die zu suchenden Strukturen benötigen, wird auf sie in dieser Arbeit nicht weiter eingegangen.

4.4.1 Der Gustafson-Kessel Algorithmus

Durch die Verwendung einer modifizierten Mahalanobisdistanz (siehe Abschnitt 2.4.1) bildet der Gustafson-Kessel Algorithmus (GUSTAFSON und KESSEL, 1979) hyperellipsoidische Cluster gleicher Größe im Merkmalsraum. Während beim Fuzzy-c-Means Algorithmus der Schwerpunkt ν_i als Prototyp des jeweiligen Clusters i angesehen werden kann, gehört bei Verwendung des Gustafson-Kessel Algorithmus die Angabe der von GUSTAFSON und KESSEL (1979) Fuzzy-Kovarianzmatrix genannten Matrix \mathbf{Cov}_i nach Gleichung (4.7) mit dazu. Der Schwerpunkt beschreibt dabei die Lage im Merkmalsraum und die Fuzzy-Kovarianzmatrix die Form des Clusters. Analog zu dem Fuzzy-c-Means Algorithmus wird angenommen, dass alle Cluster die gleiche Größe haben. Das erreicht man dadurch, dass man die Determinante der Fuzzy-Kovarianzmatrizen für alle Cluster gleich setzt. Üblicherweise wird $\det(\mathbf{Cov}_i) = 1$ gewählt (TIMM, 2002).

Die modifizierte Mahalanobisdistanz sieht beim Gustafson-Kessel Algorithmus daher wie folgt aus:

$$d_{i,k} = \sqrt{\det(\mathbf{Cov}_i)^{\frac{1}{c}} (\mathbf{x}_k - \nu_i)^T \mathbf{Cov}_i^{-1} (\mathbf{x}_k - \nu_i)} \quad (4.6)$$

Ablauf des Gustafson-Kessel Algorithmus:

0. Schritt:

Vorgabe einer Clusteranzahl c , eines Unschärfeparameters w und Bestimmung von Näherungswerten für die Clusterschwerpunkte ν_i .

1. Schritt:

Berechnung von Zugehörigkeitswerten $\mu_{i,k}$ der einzelnen Objekte zu den Clusterschwerpunkten ν_i nach Gleichung (4.4). Hierzu wird im ersten Iterationsschritt die euklidische Distanz verwendet². In den weiteren Iterationsschritten wird die modifizierte Mahalanobisdistanz nach Gleichung (4.6) benutzt.

2. Schritt:

Aus den aktuellen Zugehörigkeitwerten $\mu_{i,k}$ werden neue Clusterschwerpunkte ν_i^{neu} nach Gleichung (4.5) berechnet.

3. Schritt:

Berechnung der aktuellen Fuzzy-Kovarianzmatrizen Cov_i^{neu} :

$$Cov_i^{neu} = \frac{\sum_{k=1}^n (\mu_{i,k})^w (\mathbf{x}_k - \nu_i)(\mathbf{x}_k - \nu_i)^T}{\sum_{k=1}^n (\mu_{i,k})^w}, \quad \forall i = 1, \dots, c \quad (4.7)$$

4. Schritt:

Es wird geprüft, ob die Summe der komponentenweisen Distanzen zweier aufeinander folgender Clusterschwerpunktmatrizen \mathbf{N} kleiner ist als eine Konvergenzschwelle ϵ .

Falls $\|\mathbf{N}^{neu} - \mathbf{N}\| > \epsilon$, setze $\mathbf{N} = \mathbf{N}^{neu}$ und fahre fort mit Schritt 1. Andernfalls beende die Iteration.

Nach GUSTAFSON und KESSEL (1979) ist auch eine Berechnung unterschiedlicher Clustergrößen möglich, in dem für jede Fuzzy-Kovarianzmatrix Cov_i eine individuelle Konstante ϱ_i eingeführt wird. In dem Fall wird $\det(Cov_i) = \varrho_i$ gesetzt. Die Wahl der Konstanten setzt jedoch Vorwissen über die Cluster voraus, da sie in dem iterativen Algorithmus nicht als Unbekannte mit bestimmt werden können. (HÖPPNER U.A., 1997)

4.4.2 Der Gath-Geva Algorithmus

Der Gath-Geva Algorithmus (GATH und GEVA, 1989) bildet hyperellipsoidische Cluster unterschiedlicher Größe im Merkmalsraum. Für jedes Cluster wird die Form und Größe individuell bestimmt.

Die Berechnung der Distanz basiert bei diesem Algorithmus auf einem wahrscheinlichkeitstheoretischen Ansatz. Die Objekte \mathbf{x}_k werden als Realisierungen p -dimensionaler normalverteilter Zufallsvariablen N_i betrachtet (p : Dimension des Merkmalsraumes). GATH und GEVA (1989) gehen davon aus, dass eine Normalverteilung N_i mit dem Erwartungswert ν_i , der Fuzzy-Kovarianzmatrix Cov_i und einer a-priori Wahrscheinlichkeit P_i zur Erzeugung eines Objektes \mathbf{x}_k verwendet wurde. Als Distanzfunktion $d_{i,k}$ wird hierbei jetzt die Wurzel aus der reziproken a-posteriori Wahrscheinlichkeit, mit der ein Objekt zu einem Cluster gehört, verwendet: (Zur Herleitung siehe auch HÖPPNER U.A., 1997.)

$$d_{i,k} = \sqrt{\frac{(2\pi)^{\frac{p}{2}} \sqrt{\det(Cov_i)}}{P_i} \exp\left(\frac{1}{2} (\mathbf{x}_k - \nu_i)^T Cov_i^{-1} (\mathbf{x}_k - \nu_i)\right)} \quad (4.8)$$

Der konstante Faktor $(2\pi)^{\frac{p}{2}}$ wird bei der Berechnung üblicherweise weggelassen.

² Alternativ kann im ersten Iterationsschritt auch mit Näherungswerten für die Fuzzy-Kovarianzmatrix und der modifizierten Mahalanobisdistanz gearbeitet werden.

Die a-priori Wahrscheinlichkeit P_i für das Auftreten eines Objektes \mathbf{x}_k in einem Cluster i berechnet sich aus den Zugehörigkeitswerten nach dem Prinzip: Summe der Zugehörigkeitswerte im Cluster i geteilt durch die Gesamtsumme aller Zugehörigkeitswerte in allen Clustern.

$$P_i = \frac{\sum_{k=1}^n (\mu_{i,k})^w}{\sum_{k=1}^n \sum_{i=1}^c (\mu_{i,k})^w}, \quad \forall i = 1, \dots, c \quad (4.9)$$

Die Fuzzy-Kovarianzmatrizen \mathbf{Cov}_i^{neu} werden in gleicher Weise wie beim Gustafson-Kessel Algorithmus nach Gleichung (4.7) berechnet.

Ablauf des Gath-Geva Algorithmus:

0. Schritt:

Vorgabe einer Clusteranzahl c , eines Unschärfeparameters w und Bestimmung von Näherungswerten für die Clusterschwerpunkte ν_i .

1. Schritt:

Berechnung von Zugehörigkeitswerten $\mu_{i,k}$ der einzelnen Objekte zu den Clusterschwerpunkten ν_i nach Gleichung (4.4). Hierzu wird im ersten Iterationsschritt die euklidische Distanz verwendet. In den weiteren Iterationsschritten wird die Distanzfunktion nach Gleichung (4.8) benutzt.

2. Schritt:

Aus den aktuellen Zugehörigkeitwerten $\mu_{i,k}$ werden neue Clusterschwerpunkte ν_i^{neu} nach Gleichung (4.5) berechnet.

3. Schritt:

Berechnung der aktuellen Fuzzy-Kovarianzmatrizen \mathbf{Cov}_i^{neu} nach Gleichung (4.7).

4. Schritt:

Berechnung der a-priori Wahrscheinlichkeit P_i für das Auftreten eines Objektes \mathbf{x}_k in einem Cluster i nach Gleichung (4.9).

5. Schritt:

Es wird geprüft, ob die Summe der komponentenweisen Distanzen zweier aufeinander folgender Clusterschwerpunktmatrizen \mathbf{N} kleiner ist als eine Konvergenzschwelle ϵ .

Falls $\|\mathbf{N}^{neu} - \mathbf{N}\| > \epsilon$, setze $\mathbf{N} = \mathbf{N}^{neu}$ und fahre fort mit Schritt 1. Andernfalls beende die Iteration.

Durch die Exponentialfunktion in der Distanzfunktion wachsen ab einer bestimmten Entfernung die Distanzen sehr stark an. Das hat zur Folge, dass beim Übergang von einem Cluster zum anderen die Zugehörigkeitswerte sehr schnell von 0 auf 1 oder umgekehrt wechseln und dadurch relativ scharfe Clustereinteilungen erfolgen. (HÖPPNER U.A., 1997)

Die Konvergenzeigenschaft der iterativen Algorithmen wird mit zunehmender Komplexität schlechter. Die Gefahr ein lokales Minimum zu finden, ist beim Gustafson-Kessel Algorithmus höher als beim Fuzzy-c-Means und beim Gath-Geva Algorithmus höher als beim Gustafson-Kessel. Um eine gute

Clustereinteilung zu erreichen, müssen insbesondere beim Gath-Geva Algorithmus die Clusterschwerpunkte in der Nähe ihrer endgültigen Lage initialisiert werden. Das kann man dadurch erreichen, indem man als Startwerte keine zufälligen Näherungswerte verwendet, sondern die Ergebnisse einer Clusterung durch den Fuzzy-c-Means bzw. Gustafson-Kessel Algorithmus benutzt (HÖPPNER U.A., 1997). Während der Bearbeitung der Beispiele in dieser Arbeit zeigte sich, dass der Fuzzy-c-Means Algorithmus mit nahezu beliebigen Initialisierungen immer gut interpretierbare Clusterlösungen fand. Die Gustafson-Kessel und Gath-Geva Algorithmen reagierten dagegen sehr empfindlich auf die Initialisierungen und konvergierten trotz vermeintlich guter Initialisierung durch den Fuzzy-c-Means sehr häufig zu Clusterlösungen, die von der menschlichen Vorstellung her als unbrauchbar angesehen wurden.

4.5 Possibilistische Varianten des Fuzzy-c-Means und Gustafson-Kessel Algorithmus

Bei den possibilistischen Fuzzy-Clusterverfahren wird auf die bei den probabilistischen Fuzzy-Clusterverfahren eingeführte Nebenbedingung $\sum_{i=1}^c \mu_{i,k} = 1$ (4.1) verzichtet.

Bei den bisher betrachteten probabilistischen Fuzzy-Clusterverfahren drückt der Zugehörigkeitswert $\mu_{i,k}$ die Zugehörigkeit eines Objektes zu einem Cluster in Relation zu den anderen Clustern aus. Der Zugehörigkeitswert zu einem Cluster hängt aufgrund der Nebenbedingung auch mit von der Lage und Anzahl der anderen Cluster ab. Er drückt nicht aus, wie typisch das Objekt für das entsprechende Cluster ist. Das kann in der Praxis bei Objekten, die eine relativ große Distanz zu allen Clusterzentren aufweisen, zu Interpretationsproblemen führen. Diese Ausreißer oder Stördaten können hohe Zugehörigkeitswerte bekommen, obwohl die Objekte nicht typisch für das Cluster sind und man intuitiv einen geringeren Zugehörigkeitswert erwarten würde. Bei einer großen Distanz eines Objektes zu allen Clusterzentren streben alle Zugehörigkeitswerte des Objektes gegen $\mu_{i,k} = 1/c$.

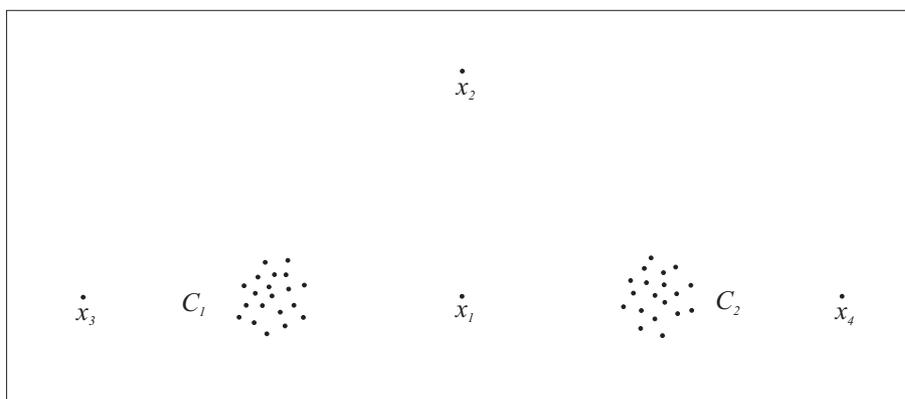


Abbildung 4.2: Zwei kompakte Cluster C_1, C_2 und vier Ausreißer x_1, x_2, x_3 und x_4 (nach KRISHNAPURAM und KELLER, 1993).

Die Abbildung 4.2 zeigt zwei kompakte Cluster C_1 und C_2 und die Objekte x_1, x_2, x_3 und x_4 . Die Objekte x_1 und x_2 erhalten bei einer probabilistischen Fuzzy-Clusterung zu beiden Clustern jeweils den Zugehörigkeitswert $\mu_{i,k} = 0,5$, obwohl das Objekt x_2 wesentlich weiter von den Clusterzentren entfernt ist. Ähnliche Interpretationsprobleme gibt es mit den Objekten x_3 und x_4 . Das Objekt x_3 erhält beispielsweise zum Cluster C_1 einen deutlich höheren Zugehörigkeitswert als das Objekt x_1 , obwohl es die gleiche Distanz zum Clusterzentrum hat. Umgekehrt gilt das Gleiche für das Objekt x_4 und dem Cluster C_2 . In dem Zusammenhang sei auch auf die noch folgenden Erläuterungen zu den Abbildungen 4.3 und 4.4 verwiesen.

Der Verzicht auf die Nebenbedingung $\sum_{i=1}^c \mu_{i,k} = 1$ (4.1) bei den possibilistischen Verfahren führt bei der Minimierung der Zielfunktion J_{FCM} des Fuzzy-C-Means Algorithmus (Gleichung (4.3))

$$J_{FCM} = \sum_{k=1}^n \sum_{i=1}^c (\mu_{i,k})^w \cdot d_{i,k}^2(\mathbf{v}_i, \mathbf{x}_k)$$

zu der trivialen Lösung:

$$\mu_{i,k} = 0 \quad \forall i = 1, \dots, c \quad \forall k = 1, \dots, n. \quad (4.10)$$

KRISHNAPURAM und KELLER (1993) haben daher vorgeschlagen, die Zielfunktion um einen Summanden zu erweitern, durch den diese triviale Lösung verhindert wird. Die Zielfunktion des possibilistischen Fuzzy-c-Means Algorithmus (PCM) sieht danach wie folgt aus:

$$J_{PCM} = \sum_{k=1}^n \sum_{i=1}^c (\mu_{i,k})^w \cdot d_{i,k}^2(\mathbf{v}_i, \mathbf{x}_k) + \sum_{i=1}^c \eta_i \sum_{k=1}^n (1 - \mu_{i,k})^w \quad (4.11)$$

η_i : Wichtungs- oder Abstandsparameter, $\eta_i \in \mathbb{R}_{>0}$

Bei der Minimierung der Zielfunktion sorgt der erste Term dafür, dass die Distanz zwischen den Objekten und den Clusterzentren minimal wird. Der zweite Term sorgt dafür, dass die Zugehörigkeitswerte möglichst groß werden, um die triviale Lösung zu verhindern. Der Parameter η_i gewichtet den zweiten Term in Gleichung (4.11) gegenüber dem ersten Term. Weiterhin gibt er die Distanz vom Clusterzentrum an, an dem der Zugehörigkeitswert eines Objektes zu diesem Cluster $\mu_{i,k} = 0,5$ beträgt. Zu seiner Berechnung haben KRISHNAPURAM und KELLER (1993) unter anderen die folgende Formel vorgeschlagen:

$$\eta_i = \frac{\sum_{k=1}^n (\mu_{i,k})^w \cdot d_{i,k}^2(\mathbf{v}_i, \mathbf{x}_k)}{\sum_{k=1}^n (\mu_{i,k})^w}, \quad \forall i = 1, \dots, c \quad (4.12)$$

Liegt Expertenwissen über die Ausdehnung der Cluster vor, kann der Parameter η_i auch a priori vorgegeben werden.

Die Minimierung der Zielfunktion erfolgt wie beim probabilistischen Fuzzy-c-Means Algorithmus durch die alternierende Optimierung der Clusterzentren und der Zugehörigkeitswerte in einem iterativen Verfahren.

Die Zugehörigkeitswerte berechnen sich dabei nach:

$$\mu_{i,k} = \frac{1}{1 + \left(\frac{d_{i,k}^2(\mathbf{v}_i, \mathbf{x}_k)}{\eta_i} \right)^{\frac{1}{w-1}}}, \quad \forall i = 1, \dots, c; \quad \forall k = 1, \dots, n \quad (4.13)$$

Man sieht, dass hier im Gegensatz zu Gleichung (4.4) des probabilistischen Fuzzy-c-Means, die Zugehörigkeitswerte nur vom Abstand zum dazugehörigen Clusterzentrum abhängen und nicht auch von den andern Clusterzentren.

Die Berechnung der Clusterzentren erfolgt in gleicher Weise wie beim probabilistischen Fuzzy-c-Means nach Gleichung (4.5).

Die Abbildungen 4.3 und 4.4 zeigen für eine eindimensionale Anwendung mit drei Clustern typische Verläufe von Zugehörigkeitsfunktionen, wie sie sich bei der Verwendung des probabilistischen und des possibilistischen Fuzzy-c-Means Algorithmus ergeben. In der Abbildung 4.3 ist deutlich zu erkennen, wie die Zugehörigkeitsfunktionen von der Lage aller Clusterzentren ν_i abhängen. Die Funktionen besitzen an den Stellen der jeweils anderen Clusterzentren ein lokales Minimum mit dem Funktionswert $\mu_i = 0$ und steigen mit zunehmender Distanz wieder an. In den Randbereichen ist ansatzweise zu sehen, dass bei einer großen Distanz eines Objektes zu allen Clusterzentren alle Zugehörigkeitswerte des Objektes gegen $\mu_{i,k} = 1/c$ streben. In der Abbildung 4.4 ist dagegen deutlich zu sehen, dass die Zugehörigkeitsfunktionen von den jeweils anderen Clusterzentren nicht beeinflusst werden.

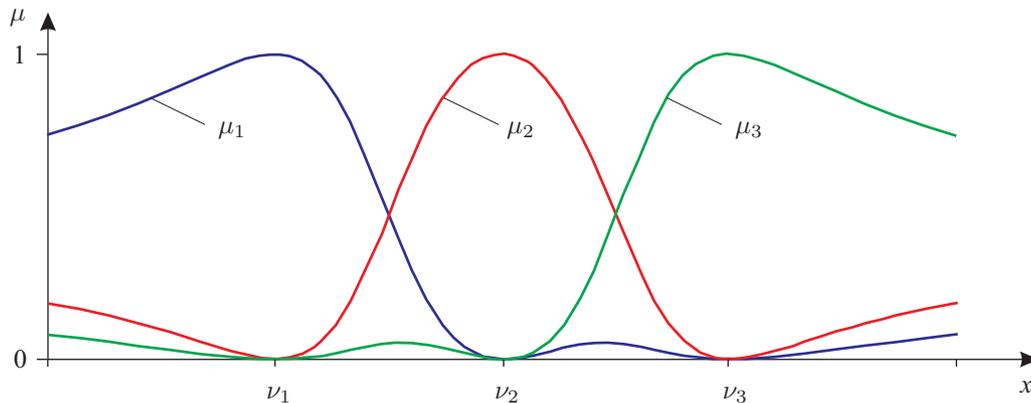


Abbildung 4.3: Drei typische Zugehörigkeitsfunktionen μ_1, μ_2 und μ_3 , wie sie sich durch Verwendung des **probabilistischen Fuzzy-c-Means** Algorithmus ergeben.

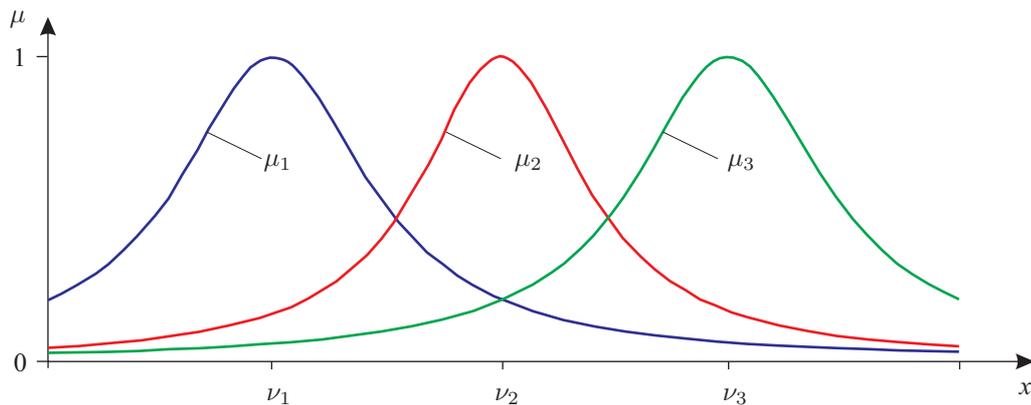


Abbildung 4.4: Drei typische Zugehörigkeitsfunktionen μ_1, μ_2 und μ_3 , wie sie sich durch Verwendung des **possibilistischen Fuzzy-c-Means** Algorithmus ergeben.

Die possibilistische Variante des Gustafson-Kessel Algorithmus erhält man, indem man die modifizierte Mahalanobisdistanz nach Gleichung (4.6) verwendet und die Fuzzy-Kovarianzmatrizen \mathbf{Cov}_i^{neu} in jedem Iterationsschritt nach Gleichung (4.7) neu berechnet. Für den Parameter η_i schlagen KRISHNAPURAM und KELLER (1993) hier $\eta_i = |\mathbf{Cov}_i|^{(1/p)}$ vor, wobei p die Dimension des Merkmalsraumes ist.

Beim Gath-Geva Algorithmus macht aufgrund des wahrscheinlichkeitstheoretischen Ansatzes für die Distanzfunktion eine possibilistische Clusterung keinen Sinn.

Die possibilistischen Clusterverfahren erhöhen die Robustheit des Clusterungsprozesses: Ihre Ergebnisse werden in wesentlich geringerem Maße von Ausreißern oder Stördaten beeinflusst. Durch den Verzicht auf die Nebenbedingung $\sum_{i=1}^c \mu_{i,k} = 1$ (4.1) kann bei Objekten, die zu mehreren Clustern gehören, die Summe der Zugehörigkeitswerte zu verschiedenen Clustern größer als 1 und bei Stördaten kleiner als 1 sein. Allerdings ist das Konvergenzverhalten der possibilistischen Clusterverfahren deutlich schlechter als das der probabilistischen. Es empfiehlt sich daher zum Beispiel zunächst eine Initialisierung mit Hilfe der probabilistischen Verfahren durchzuführen. (RUMPLER, 1999)

4.6 Gütekriterien für unscharfe Clusterergebnisse

4.6.1 Globale Gütekriterien

Unter globalen Gütekriterien werden solche Kriterien verstanden, die einen Klassifikator in seiner Gesamtheit beurteilen. Der Gegensatz dazu sind lokale Gütekriterien, die die Klassifizierung einzelner Objekte beurteilen.

Das Problem der Gütekriterien und die Bestimmung einer optimalen Clusteranzahl wurde für scharfe Clusterverfahren bereits im Abschnitt 2.4.3 kurz behandelt und insbesondere für partitionierende (iterative) Clusterverfahren wurde auf das „Ellbogen-Kriterium“ hingewiesen. Im Zusammenhang mit der Entwicklung unscharfer Clusterverfahren entstanden weitere, speziell auf die Belange der unscharfen Clusterung ausgerichtete Gütekriterien, auf die an dieser Stelle näher eingegangen werden soll. Einige Gütemaße bewerten ausschließlich die Zuordnung der Objekte zu den Clustern, andere auch die Form der Cluster oder wie gut die Cluster voneinander separiert sind.

Wie bereits im Abschnitt 2.4.3 erwähnt, führen die Gütekriterien teilweise zu schwer interpretierbaren Clusterlösungen. Sie sollten lediglich Hinweise auf eine geeignete Clusteranzahl geben. Entscheidend für die endgültige Wahl der Clusteranzahl sollte jedoch immer eine sinnvolle Interpretation durch den Anwender sein (DEIMER, 1986), (BOCKLISCH, 1987), (MIT GMBH, 2002). Als sinnvoll hat sich auch herausgestellt, mehrere Gütemaße parallel zu betrachten (HÖPPNER U.A., 1997). Ein Beispiel zeigt die Abbildung 4.5, auf die später noch näher eingegangen wird.

Partitionskoeffizient

Im Zusammenhang mit dem Fuzzy-c-Means Algorithmus schlug BEZDEK (1973b) den sogenannten Partitionskoeffizienten g_{pk} als eine Kennzahl zur Beurteilung einer unscharfen Clustereinteilung (Partitionierung) vor:

$$g_{pk} = \frac{1}{n} \sum_{k=1}^c \sum_{i=1}^n (\mu_{i,k})^2 \quad (4.14)$$

Hierbei wird davon ausgegangen, dass eine optimale Partitionierung dann gegeben ist, wenn die Objekte den Clustern möglichst eindeutig zugeordnet werden können. Das heißt, mit einem möglichst hohen maximalen Zugehörigkeitswert. Die Zugehörigkeitswerte zu den anderen Clustern sind dann dementsprechend relativ klein. Der Partitionskoeffizient ist bei einer probabilistischen Clusterung ein Wert aus dem Intervall $[\frac{1}{c}, 1]$. Umso eindeutiger (schärfer) die Einteilung ist, umso stärker strebt der Partitionskoeffizient gegen den Wert 1. Bei der unschärfsten aller möglichen Einteilungen, bei der alle Objekte gleichmäßig auf alle Cluster verteilt sind, liefert der Partitionskoeffizient den Wert $\frac{1}{c}$.

Partitionsentropie

Als eine weitere Kennzahl zur Beurteilung einer unscharfen Clustereinteilung publizierte BEZDEK (1975) die Partitionsentropie g_{pe} :

$$g_{pe} = -\frac{1}{n} \sum_{k=1}^c \sum_{i=1}^n \mu_{i,k} \cdot \ln(\mu_{i,k}) \quad (4.15)$$

Die Partitionsentropie besitzt große Ähnlichkeit mit dem Partitionskoeffizienten. Auch bei ihrer Berechnung werden die Zugehörigkeitswerte der Objekte zu allen Clustern verwendet. Die Partitionsentropie orientiert sich allerdings an der aus der Informationstheorie bekannten Shannon-Entropie (SHANNON, 1948), auf die an dieser Stelle nicht näher eingegangen werden soll, da sie für die Anwendung dieses Gütemaßes keine weitere Bedeutung hat.³

Die Partitionsentropie ist bei einer probabilistischen Clusterung ein Wert aus dem Intervall $[0, \ln(c)]$. Je eindeutiger (schärfer) die Einteilung ist, desto stärker strebt die Partitionsentropie gegen den Wert 0. Bei einer gleichmäßigen Zugehörigkeit aller Objekte zu allen Clustern liefert die Partitionsentropie den Wert $\ln(c)$.

Aus den möglichen Wertebereichen des Partitionskoeffizienten und der Partitionsentropie ist ersichtlich, dass diese von der Clusteranzahl c abhängig sind. Die Wertebereiche werden mit wachsendem c größer. Beide Kennzahlen zeigen ein monotonen Verhalten in Abhängigkeit von c . Daher sind sie nur bedingt für den Vergleich zweier Partitionen mit unterschiedlicher Clusteranzahl geeignet. In der Praxis trägt man die Kennzahlen in einem Diagramm über der Clusteranzahl ab und benutzt das in Abschnitt 2.4.3 vorgestellte „Ellbogen-Kriterium“ zur Bestimmung einer optimalen Clusteranzahl (HÖPPNER U.A., 1997).

Der Partitionskoeffizient und die Partitionsentropie sind zur Beurteilung von Partitionierungen, die mit Hilfe possibilistischer Verfahren erfolgt sind, nicht entwickelt worden und auch schlecht geeignet. Da bei den possibilistischen Clusterverfahren die Summe der Zugehörigkeitswerte eines Objektes über alle Cluster nicht 1 ergeben muss, ergeben sich für die Wertebereiche der beiden Kennzahlen, in Abhängigkeit von der jeweiligen Partitionierung, unterschiedliche Intervallgrenzen. Die Intervallgrenzen sind hierbei nicht nur von der Clusteranzahl, sondern zusätzlich auch noch von der jeweiligen Partition abhängig.

Proportionsexponent

Die Abhängigkeit der Wertebereiche des Partitionskoeffizienten und der Partitionsentropie von der Clusteranzahl, führte zu der Entwicklung des Proportionsexponenten g_{pex} (WINDHAM, 1981):

$$g_{pex} = -\log_2 \prod_{k=1}^n \sum_{i=1}^{I_k} (-1)^{i+1} \binom{c}{i} (1 - i \cdot \hat{\mu}_k)^{c-1} \quad (4.16)$$

$$\begin{aligned} \hat{\mu}_k &: \max_{i=1, \dots, c} \mu_{i,k} \quad \forall k = 1, \dots, n \\ &\quad (\text{höchster Zugehörigkeitswert eines Objektes}) \\ I_k &: \text{größte ganze Zahl} \leq \frac{1}{\hat{\mu}_k} \end{aligned}$$

Für den in der Praxis fast immer auftretenden Fall $\hat{\mu}_k > 0,5$ vereinfacht sich die Formel zu:

$$g_{pex} = -\log_2 \prod_{k=1}^n c (1 - \hat{\mu}_k)^{c-1} \quad (4.17)$$

Der Wertebereich des Proportionsexponenten ist unabhängig von der Anzahl der Cluster. Bei der unschärfsten Partitionierung, bei der alle Objekte gleichmäßig auf alle Cluster verteilt sind, liefert der Proportionsexponent den Wert 0. Mit zunehmender „Schärfe“ der Partitionierung strebt der Wert

³ „Obwohl durch die Anlehnung an Shannons Informationstheorie dieses Gütemaß einen informations-theoretischen Beigeschmack bekommt, so handelt es sich im Grunde dennoch nur um ein Maß für die Unschärfe der Clustereinteilung, sehr ähnlich zum Partitionskoeffizienten.“ (HÖPPNER U.A., 1997)

gegen ∞ . Für eine scharfe Zugehörigkeit mit einem maximalen Zugehörigkeitswert von 1 ist der Proportionsexponent nicht definiert. Bei den possibilistischen Verfahren können durchaus maximale Zugehörigkeitswerte ≥ 1 auftreten und dadurch der Summenterm in Gleichung (4.16) 0 und damit der ganze Ausdruck undefiniert werden. In diesem Fall müssten die Zugehörigkeitswerte vorab auf einen Maximalwert < 1 normiert werden.

Nachteilig ist bei diesem Gütekriterium, dass es nur den jeweils maximalen Zugehörigkeitswert eines Objektes berücksichtigt und nicht die vollständige Information aller Zugehörigkeitswerte.

Bei Anwendung des Proportionsexponenten treten im Fall von relativ hohen maximalen Zugehörigkeitswerten für den Ausdruck $(1 - i \cdot \mu_k)$ sehr kleine Werte auf, die noch mit $c - 1$ potenziert werden. Da von diesen kleinen Termen so viele multipliziert werden müssen, wie der Datensatz Elemente hat, treten für den Logarithmus sehr schnell sehr kleine Argumente auf, so dass es zu rechentechnischen Problemen kommen kann (HÖPPNER U.A., 1997). Dieses Problem zeigte sich auch bei dem in dieser Arbeit im Abschnitt 6.2.2 behandelten Beispiel.

Separationsindex

Wie bereits im Abschnitt 2.4 erwähnt, sollen Objekte innerhalb eines Clusters bezüglich ihrer Merkmalsausprägungen möglichst homogen und Objekte in verschiedenen Clustern möglichst heterogen sein. Diese Voraussetzung steht hinter dem von XIE und BENI (1991) veröffentlichten Separationsindex g_{si} :

$$g_{si} = \frac{\sum_{i=1}^c \sum_{k=1}^n (\mu_{i,k})^2 \cdot d_{i,k}^2(\mathbf{v}_i, \mathbf{x}_k)}{n \cdot \min \{d_{i,j}^2(\mathbf{v}_i, \mathbf{v}_j) \mid i, j \in \{1, \dots, c\}, i \neq j\}} \quad (4.18)$$

Der Zähler in (4.18) stellt eine Bewertung der Homogenität der Daten innerhalb der Cluster dar. Und der Nenner ist durch das Quadrat des minimalen Abstandes der Clusterzentren ein Maß der Heterogenität zwischen verschiedenen Clustern. Für eine „gute“ Partition soll der Zähler möglichst klein und der Nenner möglichst groß werden. Ein möglichst kleiner Wert des Separationsindex deutet demnach auf eine „gute“ Partition hin. Durch die Verwendung des Minimums im Nenner wird die Heterogenität allerdings tendenziell unterschätzt (TIMM, 2002).

Die Abbildung 4.5 zeigt die bisher beschriebenen Gütekriterien parallel auf Fuzzy-c-Means Clusterungen (mit $w = 2$) angewendet. Als Daten dienten die Verschiebungsvektoren des Beispiels Testnetz Delft (Abbildungen 2.4 bzw. 2.8). Bei der Suche nach der optimalen Clusteranzahl ist nach MIT GMBH (2002) diejenige zu wählen, bei der der Wert für den Partitionskoeffizienten g_{pk} über den fallenden Trend steigt bzw. bei der Partitionsentropie g_{pe} unter den steigenden Trend fällt. Das ist jeweils bei einer Clusteranzahl von drei der Fall. Der Partitionsexponent g_{pex} ist für eine gute Clusteranzahl zu maximieren. Bei einer Clusteranzahl von drei zeigt er einen Wert, der leicht über dem steigenden Trend liegt. Ein eindeutiger Hinweis auf eine optimale Clusteranzahl von drei ist das aber nicht. In dem gesamten Bereich von zwei bis sechs Clustern zeigt sich bei ihm kein deutliches (lokales) Maximum. Beim Separationsindex g_{si} ergibt sich nach HÖPPNER U.A. (1999) eine gute Clusteranzahl bei dem ersten lokalen Minimum. Das ist hier auch bei einer Clusteranzahl von drei gegeben. Eine ebenfalls denkbare Clusteranzahl von zwei Clustern lässt sich aus den hier benutzten Gütekriterien nicht ableiten. Wie bereits erwähnt, kann es bei Anwendung dieser Gütemaße zu schwer interpretierbaren Clusterlösungen kommen. Daher sollten sie nur ein Anhalt für mögliche Clusteranzahlen sein. Im Vordergrund soll immer auch eine sinnvolle Interpretation durch den Anwender stehen. In dem hier vorliegenden Fall entspricht die durch die Gütekriterien empfohlene Clusteranzahl von drei Clustern den in der Abbildung 2.8 gut zu erkennenden drei Regionen von relativ dichten Punktwolken von Objekten, die durch Regionen mit einer geringeren Dichte voneinander getrennt sind.

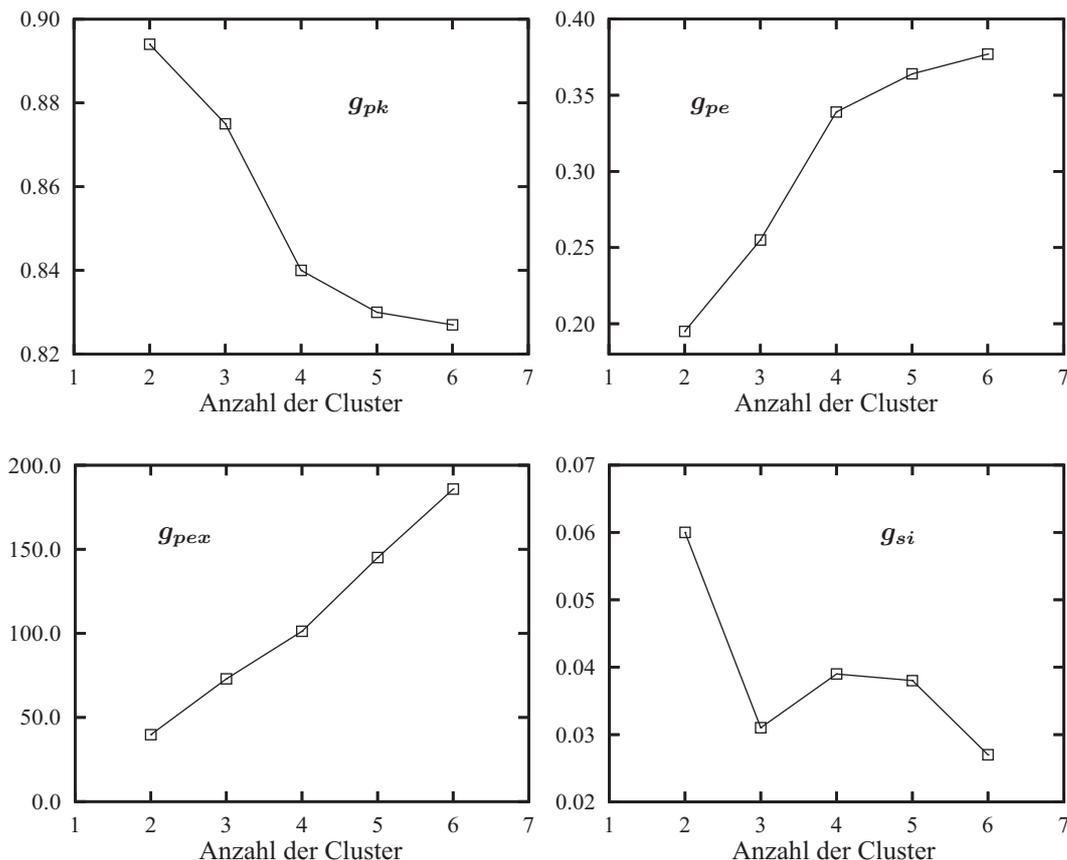


Abbildung 4.5: Gütekriterien (links oben Partitionskoeffizient g_{pk} , rechts oben Partitionsentropie g_{pe} , unten links Proportionsexponent g_{peX} und unten rechts Separationsindex g_{si}), angewendet auf Fuzzy-c-Means Clusterungen der Verschiebungsvektoren des Beispiels Testnetz Delft (Abbildungen 2.4 bzw. 2.8).

Weitere Gütekriterien

Neben den bisher vorgestellten Gütekriterien gibt es noch eine Vielzahl weiterer Gütekriterien, die in der Regel nur für bestimmte Clusterverfahren geeignet sind.

Für den Gath-Geva Algorithmus wurden zum Beispiel von GATH und GEVA (1989) auf Basis der Fuzzy-Kovarianzmatrizen der Cluster Cov_i drei Kennzahlen vorgeschlagen:

- Fuzzy-Hypervolumen,
- Partitionsdichte,
- mittlere Partitionsdichte.

Das Fuzzy-Hypervolumen basiert auf den Ausdehnungen der einzelnen Cluster. Ein Minimum dieses Gütemaßes deutet auf kleine, kompakte Cluster hin.

Für die Berechnung der beiden Partitionsdichten wird eine Hilfsgröße eingeführt, die die Anzahl der Objekte in der Nähe der Clusterzentren berücksichtigt und damit eine Häufung der Objekte in der Nähe der Clusterzentren positiv bewertet.

Eine gute Clustereinteilung im Sinne der von GATH und GEVA (1989) vorgeschlagenen Gütemaße besteht aus gut getrennten Clustern, die ein minimales Gesamtvolumen aufweisen und bei denen die Objekte sich in der Nähe der Clusterzentren häufen. (HÖPPNER U.A., 1997)

4.6.2 Lokales Gütekriterium

Während im vorhergehenden Abschnitt Gütekriterien präsentiert wurden, die einen Klassifikator in seiner Gesamtheit beurteilen, soll hier ein Gütekriterium betrachtet werden, das die Klassifizierung einzelner Objekte beurteilt.

Ein Objekt k ist gut von einem Klassifikator einem Cluster zugeordnet, wenn der höchste Zugehörigkeitswert $\hat{\mu}_k$ (Hauptzugehörigkeitswert) deutlich größer ist als die anderen Zugehörigkeitswerte (Nebenzugehörigkeitswerte).

$$\hat{\mu}_k = \max_{i=1,\dots,c} \mu_{i,k} \quad (4.19)$$

BOCKLISCH (1987) schlägt als Gütekriterium g_{lk1} den Quotienten aus dem maximalen Nebenzugehörigkeitswert und dem Hauptzugehörigkeitswert vor. Dieser Quotient sollte möglichst klein sein.

$$g_{lk1} = \max_{\substack{i=1,\dots,c \\ i \neq \hat{i}}} \mu_{i,k} / \hat{\mu}_k \quad (4.20)$$

Sofern keine scharfe Zuordnung vorliegt, kann genauso gut der Kehrwert dieses Quotienten verwendet werden. Er sollte als Kriterium für eine gute Zuordnung möglichst groß sein.

$$g_{lk2} = \hat{\mu}_k / \max_{\substack{i=1,\dots,c \\ i \neq \hat{i}}} \mu_{i,k} \quad (4.21)$$

In den Beispielen im Abschnitt 6.2 wird das Kriterium nach Gleichung (4.21) zur Beurteilung der Klassifizierung einzelner Objekte herangezogen.

Kapitel 5

Die Methode der Fuzzy Pattern Klassifikation

Wie bereits erwähnt, weisen Muster und Merkmale bestimmte Unsicherheiten auf. Bei den bisher beschriebenen iterativen Fuzzy-Clusterverfahren findet diese Unsicherheit der Objekte jedoch keinen Einzug in die Beschreibung der unscharfen Cluster. Die Unschärfe der Cluster ergibt sich rein aus der Distanz der einzelnen Objekte zum jeweiligen Clusterzentrum im Merkmalsraum.

Eine Parallelentwicklung zu den iterativen Fuzzy-Clusterverfahren ist die Methode der Fuzzy Pattern Klassifikation. Beide Verfahren unterscheiden sich grundlegend. Der erste wesentliche Unterschied besteht darin, dass hier eine strikte Trennung zwischen der Strukturbildung und der Strukturbeschreibung erfolgt. Und der zweite Unterschied besteht darin, dass diese Methode die Unschärfe der Objekte (ausgedrückt durch die Unschärfe der Merkmale) mit berücksichtigen kann und die Objekte nicht als „fehlerfrei“ behandelt werden. Einschränkend sein erwähnt, dass in dem vorliegenden Konzept die als *elementare Unschärfe* bezeichnete Unschärfe der Merkmale pro Merkmalsachse einheitlich gewählt werden muss. Eine unterschiedliche *elementare Unschärfe* für jedes Objekt ist in dem vorliegenden Konzept nicht möglich¹.

Die Methode der Fuzzy Pattern Klassifikation wurde in den 70er und 80er Jahren des letzten Jahrhunderts an der TH Karl-Marx-Stadt (heute TU Chemnitz) entwickelt und im Wesentlichen von BOCKLISCH (1981) und (1987) veröffentlicht. Das Verfahren ist sehr universell einsetzbar und hat sich in einem weiten Anwendungsspektrum in vielen Fachdisziplinen bewährt. Beispiele sind die medizinische Diagnostik, die Überwachung und Diagnose von Maschinen und Anlagen und die Analyse und Prognose von Verkehrsströmen, Energieverbräuchen und Umweltdaten. (TECHNISCHE UNIVERSITÄT CHEMNITZ, 2010).

Die **Strukturbildung** erfolgt zunächst nach einem beliebigen scharfen Clusterverfahren. In der vorliegenden Software (Modul CLUSTER der Fuzzy-Toolbox: TU Chemnitz, Professur für Systemtheorie) stehen zur Clusterung verschiedene hierarchische Möglichkeiten und das iterierte Minimaldistanzverfahren bereit.

Aufgrund der Modularität der Software und der offenen Datenschnittstellen kann die Klassenbildung aber auch mit Hilfe eigener Software nach einem beliebigen anderen Verfahren durchgeführt werden.

Die anschließende **Strukturbeschreibung** (die Aufstellung der Zugehörigkeitsfunktion) erfolgt unter Berücksichtigung der Unschärfe der Merkmale nach einem besonderen Aggregationsverfahren, auf das im Abschnitt 5.2 noch näher eingegangen wird.

¹ Während der Bearbeitungszeit dieser Arbeit wurde an der TU Chemnitz im Rahmen einer Diplomarbeit (HEMPEL, 2005) das Konzept auch auf die Bearbeitung von unterschiedlichen elementaren Unschärfen erweitert. In einer zusammengefassten Darstellung ist das Konzept in HEMPEL und BOCKLISCH (2010) veröffentlicht worden.

Bei den sich ergebenden Zugehörigkeitswerten für die zu klassifizierenden Objekte handelt es sich um possibilistische Zugehörigkeitswerte. Im Gegensatz zu den bisher beschriebenen Verfahren kann der maximale Zugehörigkeitswert auch größer als eins sein.

5.1 Beschreibung der Unschärfe

Grundlage für die Beschreibung der Unschärfe der Objekte als auch der Unschärfe der Klassen² bildet bei der Fuzzy Pattern Klassifikation eine spezielle Zugehörigkeitsfunktion vom Potentialfunktionstyp (siehe Abbildung 5.1). Es handelt sich dabei um eine unimodale Funktion, die im Allgemeinen nicht symmetrisch ist.

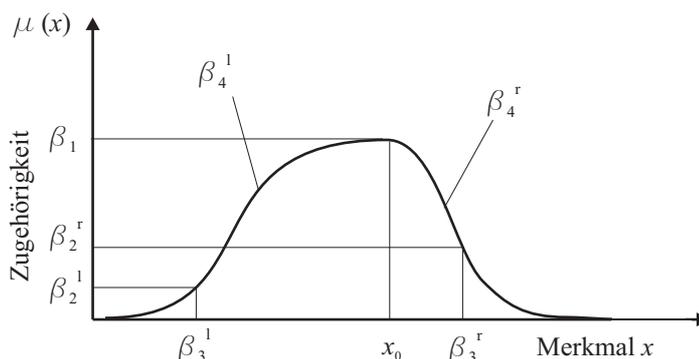


Abbildung 5.1: Eindimensionale verallgemeinerte Potentialfunktion.

Die Vorteile der hier verwendeten Zugehörigkeitsfunktion liegen nach BOCKLISCH (1987) darin, dass ihre Parameter gut interpretierbar sind, sie rechen technisch handhabbar ist und sie eine relativ gute Anpassungsfähigkeit an verschiedene Datenkonfigurationen und -strukturen bietet.

Neben dem Parameter x , der die Ausprägung des entsprechenden Merkmals darstellt, haben die Parameter β_i der Zugehörigkeitsfunktion die folgenden Bedeutungen:

- x_0 : Bei diesem Wert ist die Zugehörigkeit maximal.
Bei der **Beschreibung von Objekten** entspricht dieser Wert dem Messwert oder einem geschätzten Wert des entsprechenden Merkmals.
Bei der **Beschreibung von Klassen** entspricht der Wert dem Repräsentanten der Klasse (ν_i).
- β_1 : Höhe (Maximalwert) der Zugehörigkeitsfunktion.
- β_2 : Zugehörigkeit der Erscheinung am Rande ($x_0 \pm \beta_3$).
- β_3 : Ausdehnung der Erscheinung (Spreizung der Funktion).
- β_4 : Steilheit des Abfalls der Zugehörigkeit zum Rande hin.
Für $\beta_4 \rightarrow \infty$ ergibt sich eine Rechteckfunktion, welche der Zugehörigkeitsfunktion einer scharfen Menge entspricht.
Praktische Werte für β_4 liegen zwischen 2 und 20.

Für die Parameter β_2, β_3 und β_4 , die die Unschärfe modellieren, können unterschiedliche Werte für die jeweilige rechte und linke Seite angesetzt werden ($\beta_2^l, \beta_2^r, \beta_3^l, \beta_3^r, \beta_4^l, \beta_4^r$). Dadurch ergibt sich dann wie in der Abbildung 5.1, eine unsymmetrische Zugehörigkeitsfunktion.

² Da die Beschreibung unscharfer Klassen durch das Fuzzy Pattern Klassifikation Konzept sich nicht nur auf Klassen bzw. Cluster beschränkt, die durch Clusterverfahren gefunden wurden, soll an dieser Stelle allgemein der Begriff Klassen verwendet werden.

Die Funktionswerte der eindimensionalen Zugehörigkeitsfunktionen lassen sich nach folgender Formel berechnen (BOCKLISCH, 1981):

$$\mu(x) = \beta_1 \cdot \left[1 + \frac{1}{2} (1 + \operatorname{sgn}(x_0 - x)) \left(\frac{1}{\beta_2^l} - 1 \right) \cdot \left(\frac{|x - x_0|}{\beta_3^l} \right)^{\beta_4^l} + \frac{1}{2} (1 + \operatorname{sgn}(x - x_0)) \left(\frac{1}{\beta_2^r} - 1 \right) \cdot \left(\frac{|x - x_0|}{\beta_3^r} \right)^{\beta_4^r} \right]^{-1} \quad (5.1)$$

Diese Funktion lässt sich aus der im Abschnitt 3 vorgestellten einfachen Zugehörigkeitsfunktion vom Potentialfunktionstyp (Formel (3.10)) durch Parametersubstitution, die Einführung eines zusätzlichen Parameters für die Höhe und eine Trennung in einen linken und rechten Ast ableiten.

Der Parameter β_1 muss nicht zwangsweise 1 sein. Bei der Beschreibung von Klassen berechnet er sich aus der Anzahl der Lernobjekte, die die Klasse aufbauen. „*Er drückt die Bedeutung einer Klasse relativ zu anderen aus.*“ (BOCKLISCH, 1987). Für die in dieser Arbeit gezeigten Beispiele könnte der Parameter theoretisch auf einen konstanten Wert von $\beta_1 = 1$ oder $\beta_1 = 10$ gesetzt werden. Der Grund für die Abhängigkeit von der Anzahl der Lernobjekte ist, dass das Prinzip der Fuzzy Pattern Klassifikation nicht nur für statische Klassifikatoren³, sondern auch für dynamische Klassifikatoren⁴ konzipiert ist. Bei den dabei auftretenden Phänomenen wie Modifikation bestehender Klassen, neues Entstehen und Absterben von Klassen oder Verschmelzen und Teilen von Klassen, wird ein von der Anzahl der Lernobjekte abhängiger Parameter β_1 benötigt.

Der Übergang von einer eindimensionalen Zugehörigkeitsfunktion auf eine p -dimensionale Zugehörigkeitsfunktion geschieht dadurch, dass die p -dimensionale Funktion sich aus p eindimensionalen Funktionen über den p -Achsen des Merkmalsraumes zusammensetzt. Der Parameter β_1 wird dabei für alle p eindimensionalen Funktionen gleich gewählt.

Der Funktionswert der mehrdimensionalen Zugehörigkeitsfunktion lässt sich nach folgender Formel berechnen (BOCKLISCH, 1981):

$$\mu(\mathbf{x}) = \beta_1 \cdot \left[1 + \frac{1}{2p} \sum_{i=1}^p (1 + \operatorname{sgn}(x_{0_i} - x_i)) \left(\frac{1}{\beta_{2_i}^l} - 1 \right) \cdot \left(\frac{|x_i - x_{0_i}|}{\beta_{3_i}^l} \right)^{\beta_{4_i}^l} + \frac{1}{2p} \sum_{i=1}^p (1 + \operatorname{sgn}(x_i - x_{0_i})) \left(\frac{1}{\beta_{2_i}^r} - 1 \right) \cdot \left(\frac{|x_i - x_{0_i}|}{\beta_{3_i}^r} \right)^{\beta_{4_i}^r} \right]^{-1} \quad (5.2)$$

Zu dieser Formel gelangt man, indem p eindimensionale Zugehörigkeitsfunktionen nach Formel (5.1) mit Hilfe einer auf den p -dimensionalen Fall erweiterten parametrisierten Hamacher t-Norm (SCHEUNERT, 2002) und dem Parameter $\gamma = 0$ verknüpft werden. Für den Fall $p = 2$ befindet sich eine detaillierte Herleitung in PÄSSLER (1998).

Während sich bei den bisher vorgestellten Fuzzy-Clusterverfahren die Zugehörigkeit zu jedem Cluster nur aus den Distanzen zwischen dem Objektpunkt und dem Clusterzentrum im Merkmalsraum ergab, bestimmt sich hier die Zugehörigkeit zu der Klasse aus der Zugehörigkeit zu den Einzelmerkmalen. Hierdurch ist das Verfahren deutlich unabhängiger von den Dimensionen der einzelnen Merkmalsachsen.

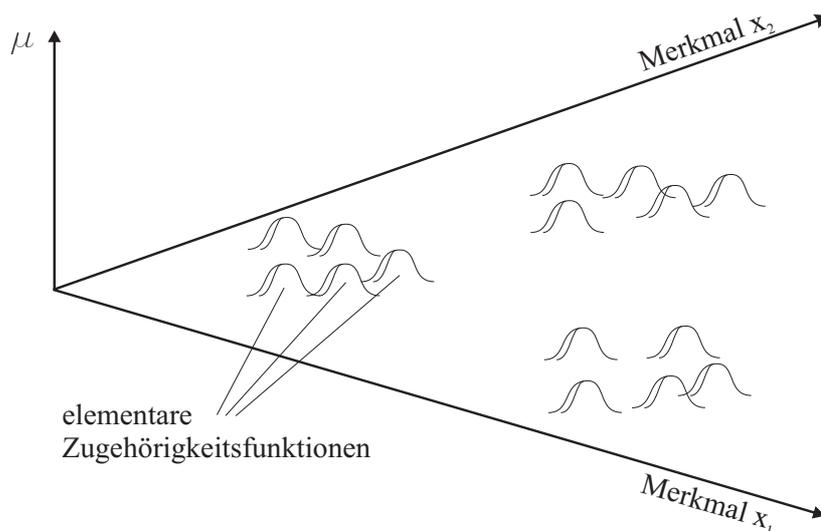


Abbildung 5.2: Unscharfe Objekte in einem zweidimensionalen Merkmalsraum (nach BURMEISTER, 1993).

5.1.1 Beschreibung der Unschärfe von Objekten im Merkmalsraum

Jedes Objekt und seine Unschärfe werden im p -dimensionalen Merkmalsraum durch eine p -dimensionale parametrische Zugehörigkeitsfunktion vom oben vorgestellten Typ beschrieben (siehe Abbildung 5.2). Diese Zugehörigkeitsfunktion wird auch als **elementare Zugehörigkeitsfunktion** bezeichnet. Da in der Regel an dieser Stelle keine Informationen vorliegen, die eine unsymmetrische Zugehörigkeitsfunktion rechtfertigen würden, werden der Einfachheit halber die Parameter für die rechte und linke Seite gleich gesetzt. Für die Parameter werden nach BOCKLISCH (1987) die folgenden Werte angenommen:

$$\begin{aligned}\beta_1 &= 1 \\ \beta_2 &= 0,5 \\ \beta_4 &= 2\end{aligned}$$

x_0 : Messwert oder geschätzter Wert des entsprechenden Merkmals. Je nach Anwendungsfall kann er aber auch subjektiv festgelegt oder anderweitig problemabhängig plausibel bestimmt werden.

β_3 : Der Parameter β_3 wird in der Literatur an dieser Stelle als *elementare Unschärfe* bezeichnet. Für geodätische Anwendungen kann hier zum Beispiel die Standardabweichung des Merkmals oder ein Vertrauensintervall angesetzt werden.

Der Funktionswert der eindimensionalen Zugehörigkeitsfunktion zur Beschreibung der Unschärfe von Objekten vereinfacht sich dadurch aus Gleichung (5.1) zu:

$$\mu(x) = \left[1 + \left(\frac{|x - x_0|}{\beta_3} \right)^2 \right]^{-1} \quad (5.3)$$

Die mehrdimensionale Zugehörigkeitsfunktion aus Gleichung (5.2) vereinfacht sich zu:

³ Statische Klassifikatoren werden einmal aufgebaut und in der späteren Identifikationsphase von Objekten nicht mehr verändert.

⁴ Bei dynamischen Klassifikatoren ändert sich die Klassenstruktur während der Identifikationsphase durch neu hinzukommende Lernobjekte oder durch zeitlich sich ändernde Lernobjekte.

$$\mu(\mathbf{x}) = \left[1 + \frac{1}{p} \sum_{i=1}^p \left(\frac{|x_i - x_{0_i}|}{\beta_{3_i}} \right)^2 \right]^{-1} \tag{5.4}$$

5.1.2 Beschreibung der Unschärfe von Klassen im Merkmalsraum

Nach einem besonderen Aggregationsverfahren, das im Abschnitt 5.2 beschrieben wird, gelangt man von der unscharfen Beschreibung der Objekte zu einer unscharfen Beschreibung für jede Klasse. Jede Klasse wird hier auch wieder durch eine p -dimensionale parametrische Zugehörigkeitsfunktion beschrieben (siehe Abbildung 5.3). Diese Zugehörigkeitsfunktion wird als **globale Zugehörigkeitsfunktion** bezeichnet.

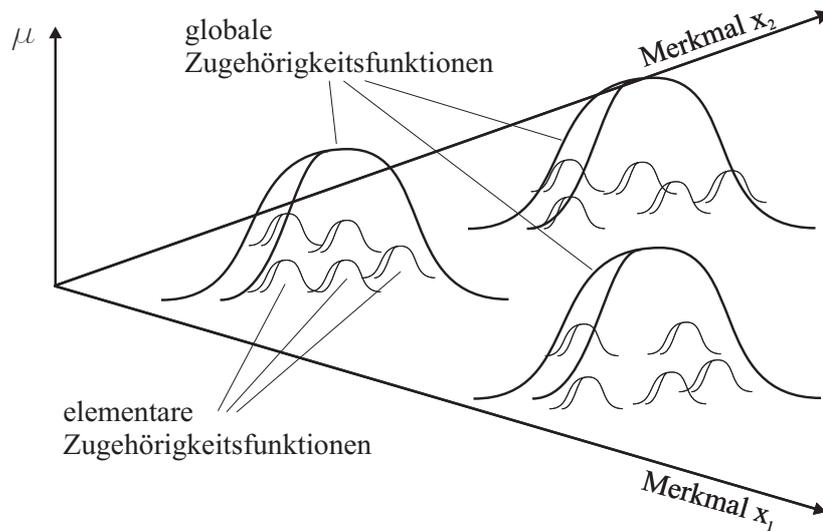


Abbildung 5.3: Unscharfe Klassen in einem zweidimensionalen Merkmalsraum (nach BURMEISTER, 1993).

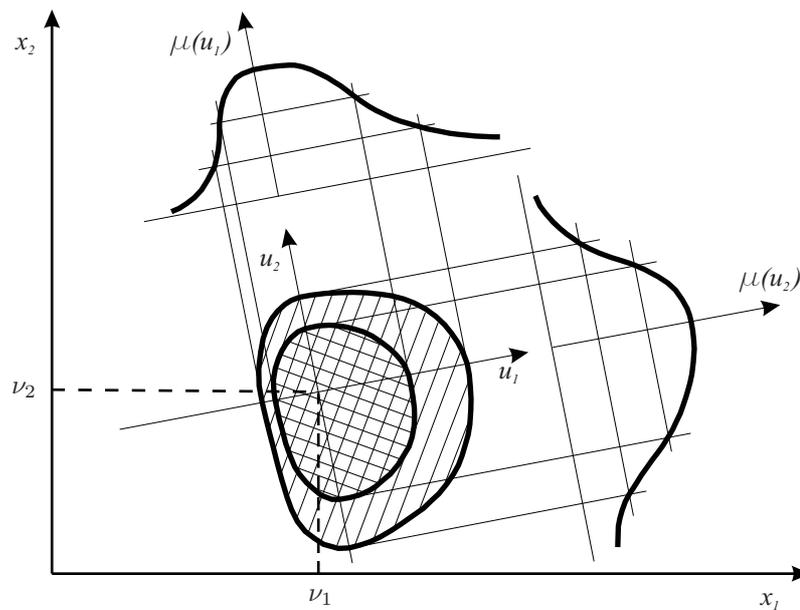


Abbildung 5.4: Grundrissdarstellung einer zweidimensionalen Zugehörigkeitsfunktion, dargestellt durch Höhenlinien und (umgeklappte) Darstellung der zwei sie stützenden eindimensionalen Zugehörigkeitsfunktionen (BOCKLISCH, 1987).

Die Beschreibung der Klassen durch die p -dimensionale Zugehörigkeitsfunktion vom oben beschriebenen Typ führt zu einer näherungsweise hyperellipsoidischen Form im Merkmalsraum, die beliebig gegenüber den Merkmalsachsen gedreht sein kann. Durch die Unabhängigkeit der Parameter in jeder Merkmalsrichtung und der links- und rechtsseitigen Unterscheidung der Funktionsäste, die zueinander nicht ähnlich sein müssen, erreicht man eine flexible unscharfe Beschreibung der Klassen. Diese Art der Beschreibung kann sich wesentlich besser an die tatsächliche Verteilung der Objekte anpassen als das bei den bisher vorgestellten Fuzzy Clusterverfahren möglich ist. Die Abbildung 5.4 zeigt hierzu ein Beispiel. Das dort abgebildete Achsensystem (u_1, u_2) kennzeichnet einen sogenannten Klassenraum, der aus den Hauptschwereachsen der Objektpunktwolke abgeleitet wird. Im Folgenden wird darauf noch näher eingegangen.

5.2 Aggregation der unscharfen Objekte zu einer unscharfen Klasse

Die Angaben in diesem Abschnitt 5.2 stützen sich auf die Veröffentlichungen von MANN (1984), BOCKLISCH (1987) und PÄSSLER (1998).

Wie bereits erwähnt, erfolgt die Strukturbildung, das heißt die Klassenbildung, zunächst nach einem beliebigen scharfen Clusterverfahren oder auf andere Art und Weise. Auf jeden Fall liegen vorerst scharfe Klasseneinteilungen mit einer eindeutigen Zuordnung der Objekte zu den Klassen vor. Die anschließende Strukturbeschreibung der unscharfen Klassen erfolgt über eine globale Zugehörigkeitsfunktion. Diese wird durch das im Folgenden beschriebene heuristische Aggregationsprinzip aus den elementaren Zugehörigkeitsfunktionen der zu einer Klasse gehörenden Objekte abgeleitet:

$$\mu_g(\mathbf{x}) = \text{agg } \mu_{e_j}(\mathbf{x}_j) \quad \forall j = 1, \dots, n_K \quad (5.5)$$

- $\mu_g(\mathbf{x})$: Zugehörigkeitsfunktion der Klasse
(globale Zugehörigkeitsfunktion)
- $\mu_{e_j}(\mathbf{x}_j)$: Zugehörigkeitsfunktion des j -ten Objektes
(elementare Zugehörigkeitsfunktion)
- n_K : Anzahl der Objekte in der Klasse

5.2.1 Transformation der Objekte einer Klasse vom Merkmalsraum in einen Klassenraum

Um die Bildung der globalen Zugehörigkeitsfunktion zu erleichtern, werden in einem vorbereitenden Schritt die zu einer Klasse gehörenden Objekte zunächst vom Merkmalsraum in den sogenannten Klassenraum transformiert. Dieser Klassenraum beschreibt die Klasse hinsichtlich ihrer Lage im Merkmalsraum. Der Ursprung des Klassenraumes u_0 liegt dabei im Schwerpunkt ν der Objektwolke. Um eine bessere Anpassung der späteren Klassenbeschreibung an die tatsächliche Form der Klasse zu erreichen, wird der Klassenraum gegenüber dem Merkmalsraum zusätzlich noch in Richtung der Hauptschwereachse der Objektwolke gedreht. Für bestimmte Anwendungen, die entweder durch die Klassenstrukturen bedingt oder rechenzeitkritisch sind, kann auf die rotatorische Anpassung unter Umständen verzichtet werden.

$$\mathbf{u} = \mathbf{R} \cdot (\mathbf{x} - \nu) \quad (5.6)$$

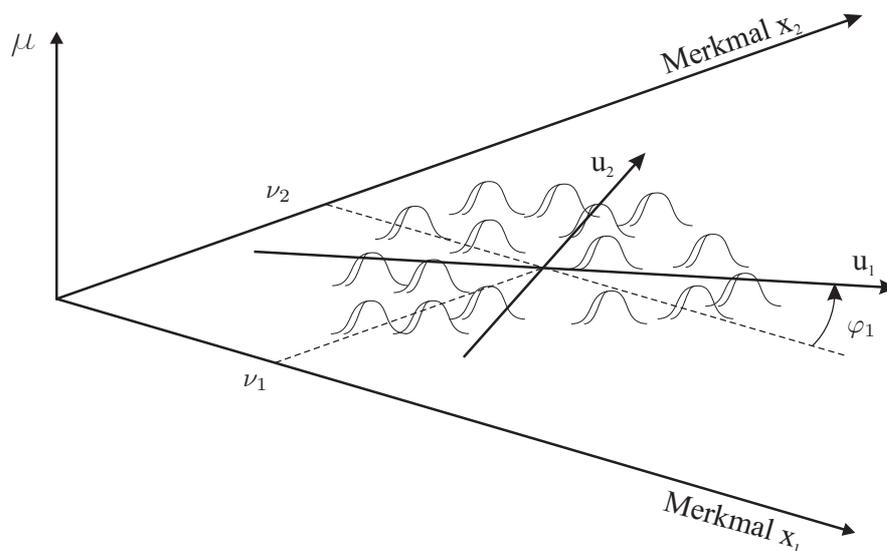


Abbildung 5.5: Achsensystem (u_1, u_2) des Klassenraumes translatorisch und rotatorisch an die Objekte angepasst (nach BURMEISTER, 1997).

- \mathbf{u} : Koordinaten im Klassenraum
- \mathbf{x} : Koordinaten im Merkmalsraum
- $\boldsymbol{\nu}$: Koordinaten des Klassenschwerpunktes
- \mathbf{R} : Rotationsmatrix

Als Transformationsparameter ergeben sich p Translationsparameter ν_i (je einer pro Merkmalsachse) und $(p - 1)$ Rotationsparameter φ_i . In der vorliegenden Software wird die Drehung mit Hilfe der Rotationsmatrix aus Gleichung (5.6) durch $p - 1$ zweidimensionale Drehschritte ersetzt. Die Hauptachse der Punktvolke wird jeweils aus $p - 1$ Projektionen der Punktvolke in die Ebenen, die vom ersten x_1 und i -ten Merkmal x_i aufgespannt werden, approximiert. In jeder Ebene wird dann der Rotationsparameter φ aus einer linearen Regression durch die in die Ebene projizierte Punktvolke geschätzt. Die jeweils anderen Merkmalsachsen werden entsprechend der Orthogonalität des Merkmalsraumes mitgedreht. Die Abbildung 5.5 zeigt in einem zweidimensionalen Merkmalsraum das gegenüber den Merkmalsachsen verschobene und gedrehte orthogonale Achsensystem (u_1, u_2) des Klassenraumes.

5.2.2 Projektion der elementaren Zugehörigkeitsfunktionen auf die Achsen des Klassenraumes

Zur Berechnung der globalen Zugehörigkeitsfunktion über den Achsen des Klassenraumes müssen die elementaren Zugehörigkeitsfunktionen auf die entsprechenden Achsen projiziert werden (Abbildung 5.6). Die Projektion der Lageparameter x_{0j} auf die Achsen des Klassenraumes ist durch die vorhergegangene Transformation bereits erfolgt. Die elementaren Unschärfen β_{3j} , die auf den Merkmalsachsen definiert sind, müssen allerdings in Abhängigkeit vom Rotationsparameter φ_i noch auf die Projektionsrichtungen umgerechnet werden. Dies geschieht mit Hilfe eines elliptischen Interpolationskonzeptes.

Die dazu notwendige allgemeine Gleichung eines Hyperellipsoides mit dem Mittelpunkt im Ursprung des Koordinatensystems lautet:

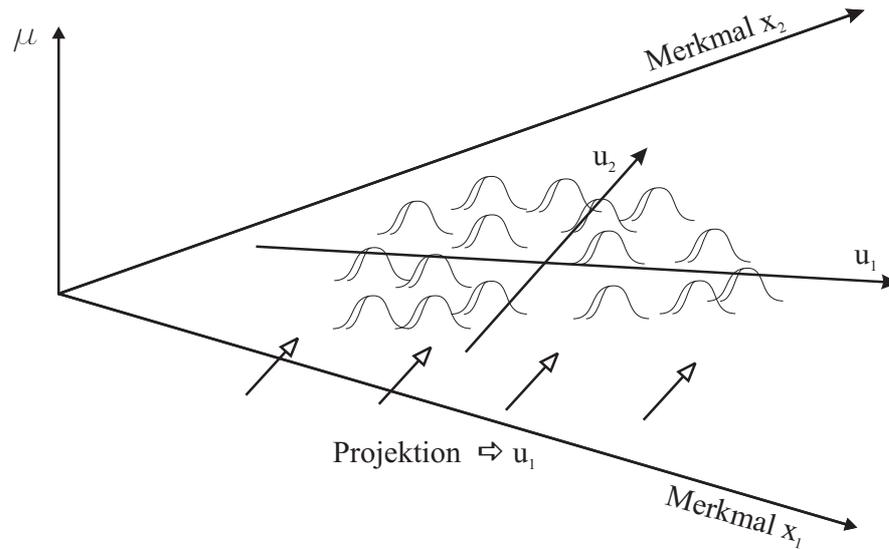


Abbildung 5.6: Veranschaulichung der Projektion der elementaren Zugehörigkeitsfunktionen auf die u_1 -Achse des Klassenraumes (nach BURMEISTER, 1997).

$$\sum_{j=1}^p \frac{x_j^2}{a_j^2} = 1 \tag{5.7}$$

a_j : Halbachsen des Hyperellipsoides

Gesucht ist die elementare Unschärfe $\beta_{3(u_k)}$ auf der entsprechenden Achse des Klassenraumes. Gedanklich wird hierzu der Ursprung des Klassenraumes in den Ursprung des Merkmalsraumes gelegt. Weiterhin wird der Mittelpunkt eines Hyperellipsoides ebenfalls in den Ursprung gelegt und zwar so, dass die Halbachsen auf den Merkmalsachsen liegen. Als Längen der Halbachsen dienen die elementaren Unschärfen β_{3_j} . Nach dem in diesem Fall benutzten Interpolationskonzeptes ist die gesuchte elementare Unschärfe $\beta_{3(u_k)}$ dann die Distanz vom Ursprung zum Schnittpunkt des Hyperellipsoides mit der entsprechenden Achse des Klassenraumes (siehe Abbildung 5.7).

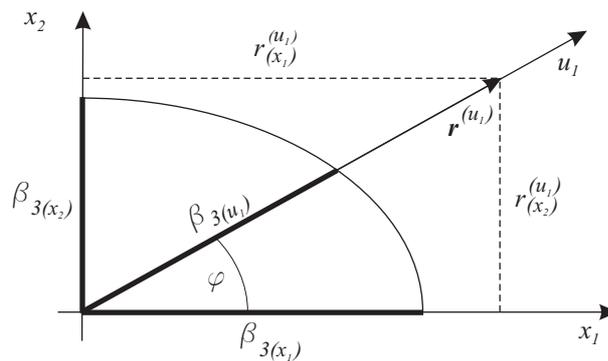


Abbildung 5.7: Skizze zur elliptischen Interpolation in einem zweidimensionalen Beispiel.

$\mathbf{r}^{u_k} = (r_1^{u_k}, \dots, r_p^{u_k})$ sei der normierte Richtungsvektor für eine Koordinatenachse u_k des Klassenraumes. Damit gilt:

$$\sum_{j=1}^p \frac{(\beta_{3(u_k)} \cdot r_j^{u_k})^2}{\beta_{3_j}^2} = 1 \tag{5.8}$$

Durch Umformung erhält man:

$$\beta_{3(u_k)} = \left(\sum_{j=1}^p \frac{(r_j^{u_k})^2}{\beta_{3_j}^2} \right)^{-\frac{1}{2}} \tag{5.9}$$

Eine rechnerische Nichtberücksichtigung der Verschiebung des Ursprungs des Klassenraumes in den Ursprung des Merkmalsraumes wirkt sich auf das Ergebnis nicht aus und ist somit unerheblich.

5.2.3 Berechnung der Parameter der globalen Zugehörigkeitsfunktion

Wie bereits erwähnt, setzt sich die p -dimensionale globale Zugehörigkeitsfunktion aus p eindimensionalen Zugehörigkeitsfunktionen mit einheitlichem funktionalen Aufbau zusammen und diese wieder aus einer jeweils rechten und linken Seite, die sich alle nur durch die numerischen Werte ihrer Parameter unterscheiden. Dadurch kann das Prinzip der Berechnung der Parameter beispielhaft an einer Seite dargelegt werden. Die Berechnung der Parameter der anderen Seite und der anderen eindimensionalen Zugehörigkeitsfunktionen erfolgt analog. Der Einfachheit halber sind in den folgenden Ausführungen die Indizes für die Achsen des Klassenraumes weggelassen. Mit dem Index g sind die Parameter der globalen Zugehörigkeitsfunktion gekennzeichnet und mit dem Index e die Parameter der elementaren Zugehörigkeitsfunktion.

Zur Berechnung der Parameter werden die auf die entsprechende Achse des Klassenraumes projizierten Objekte (siehe Abbildung 5.8) nach „linksseitigen“ und „rechtsseitigen“ Objekten vom Ursprung u_0 getrennt und anschließend nach ihrem Abstand vom Ursprung geordnet. Das Objekt mit dem größten „rechtsseitigen“ Abstand wird als u_R bezeichnet und n_R sei die Anzahl der „rechtsseitigen“ Objekte. Objekte, die direkt auf dem Ursprung liegen, werden jeweils zur Hälfte der linken und rechten Seite zugeordnet. O_r sei die Menge der „rechtsseitigen“ Objekte mit: $u_0 \leq u_{01} \leq u_{02} \leq u_{03} \dots \leq u_R$. Am Beispiel der rechten Seite sei im Folgenden die Berechnung der Parameter dargestellt.

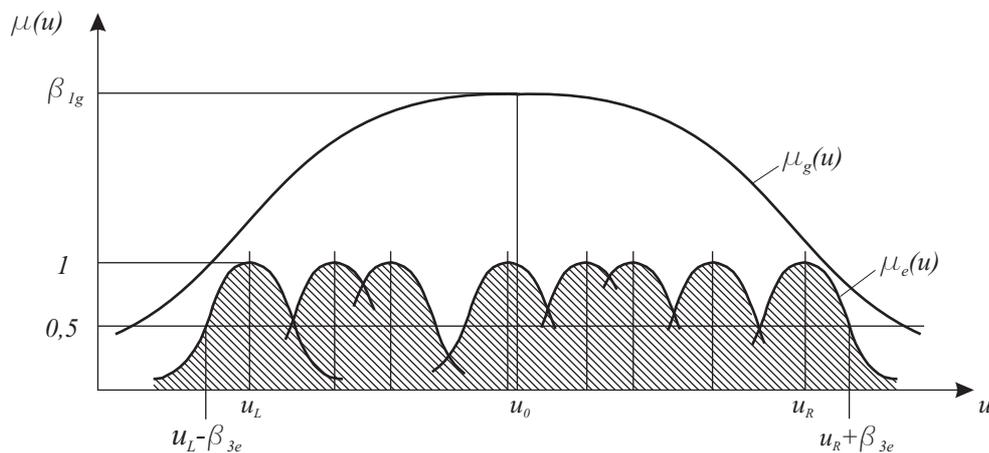


Abbildung 5.8: Die elementaren Zugehörigkeitsfunktionen μ_e der Objekte und die globale Zugehörigkeitsfunktion μ_g der Klasse über einer Achse des Klassenraumes (nach BOCKLISCH, 1987).

Ein sinnvolles Kriterium für die Beschreibung der unscharfen Klasse ist, dass die Fläche unter der globalen Zugehörigkeitsfunktion sich an der Summe der Flächen unter den elementaren Zugehörigkeitsfunktionen orientiert. Durch stückweise Integration über die ODER-verknüpften elementaren Zugehörigkeitsfunktionen kommt man zu der Summenfläche A_S .

$$\begin{aligned}
A_S = & \int_0^{\frac{u_{01}+u_{02}}{2}} \frac{du}{1 + ((u - u_{01})/\beta_{3e})^2} \\
& + \int_{\frac{u_{01}+u_{02}}{2}}^{\frac{u_{02}+u_{03}}{2}} \frac{du}{1 + ((u - u_{02})/\beta_{3e})^2} \\
& + \dots \\
& + \int_{\frac{u_{R-1}+u_R}{2}}^{u_R+\beta_{3e}} \frac{du}{1 + ((u - u_R)/\beta_{3e})^2}
\end{aligned} \tag{5.10}$$

Unter der Annahme, dass die elementaren Zugehörigkeitsfunktionen der jeweils kleinsten Objekte der rechten und linken Seite den gemeinsamen Schnittpunkt rechtsseitig von u_0 haben, kommt man durch Auflösen der Integrale und Zusammenfassen zu:

$$A_S = \beta_{3e} \cdot \left[\arctan\left(\frac{u_{01}}{\beta_{3e}}\right) + 2 \cdot \sum_{j=2}^R \arctan\left(\frac{u_{0j} - u_{0j-1}}{2\beta_{3e}}\right) + \frac{\pi}{4} \right] \tag{5.11}$$

Für den Fall, dass die elementaren Zugehörigkeitsfunktionen der jeweils kleinsten Objekte der rechten und linken Seite den gemeinsamen Schnittpunkt linksseitig von u_0 haben, kommt man zu:

$$\begin{aligned}
A_S = \beta_{3e} \cdot & \left[2 \cdot \arctan\left(\frac{u_{01} - u_{01}^{links}}{2 \cdot \beta_{3e}}\right) - \arctan\left(\frac{-u_{01}^{links}}{\beta_{3e}}\right) \right. \\
& \left. + 2 \cdot \sum_{j=2}^R \arctan\left(\frac{u_{0j} - u_{0j-1}}{2\beta_{3e}}\right) + \frac{\pi}{4} \right]
\end{aligned} \tag{5.12}$$

Berechnung des Parameters β_{1g}

Der Parameter β_{1g} , der die Höhe der globalen Zugehörigkeitsfunktion angibt, stellt in dem Aggregationsverfahren eine Besonderheit dar. Er wird für alle p eindimensionalen Zugehörigkeitsfunktionen gleichgesetzt. Wie bereits erwähnt, berechnet er sich aus der Anzahl der Objekte, die die Klasse aufgebaut haben. Zur Berechnung ist ein asymptotischer Grenzwert von β_{1g} vorzugeben.

$$\beta_{1g} = \beta_{1g \max} \cdot \left(1 - \left(\frac{\beta_{1g \max} - 1}{\beta_{1g \max}} \right)^{n_k} \right) \tag{5.13}$$

$\beta_{1g \max}$: Asymptotischer Grenzwert von β_{1g}
 n_k : Anzahl der Objekte der Klasse

Berechnung des Parameters β_{2g}

Der Parameter β_{2g} gibt den Zugehörigkeitswert an der Stelle $u = \beta_{3g}$ an. Er berechnet sich aus der Summenfläche A_S nach folgender Formel:

$$\beta_{2g} = \frac{A_S^{\beta_{4g}}}{A_S^{\beta_{4g}} + (\beta_{3g} \cdot \frac{\pi}{2})^{\beta_{4g}}} \quad (5.14)$$

Berechnung des Parameters β_{3g}

Der Parameter β_{3g} gibt die Ausdehnung der Klasse an. Er berechnet sich aus dem Randobjekt u_R zuzüglich der nach Gleichung (5.9) auf die Klassenachse projizierten elementaren Unschärfe der Objekte:

$$\beta_{3g} = u_R + \beta_{3e} \quad (5.15)$$

Berechnung des Parameters β_{4g}

Der Parameter β_{4g} , der die Steilheit des Abfalls der globalen Zugehörigkeitsfunktion zum Rande hin angibt, berechnet sich über die Verteilung der Objekte in der Klasse. Dies erfolgt hilfswise über den sogenannten Stufensprung zwischen benachbarten Objekten (Änderung des Abstandes zwischen jeweils drei benachbarten Objekten). Der Stufensprung q_j an der Stelle j berechnet sich hier nach:

$$q_j = \frac{u_{j+2} - u_{j+1} + 2\beta_{3e}/n_R}{u_{j+1} - u_j + 2\beta_{3e}/n_R} \quad \forall j = 1, \dots, n_R - 2 \quad (5.16)$$

Der Ausdruck $2\beta_{3e}/n_R$ soll unter anderem verhindern, dass der Wert von q_j unendlich groß wird, wenn zwei benachbarte Objekte u_j und u_{j+1} den selben Wert haben. Weiterhin bewirkt er noch einen Einfluss der elementaren Unschärfe (in Abhängigkeit von der Anzahl der Objekte in der Klasse) auf die Form der Zugehörigkeitsfunktion.

Der mittlere Stufensprung \bar{q} ergibt sich aus:

$$\bar{q} = \frac{1}{n_R - 2} \sum_{j=1}^{n_R-2} q_j \quad (5.17)$$

Daraus wird mit Hilfe eines empirischen Ansatzes der Wert für β_{4g} berechnet:

$$\beta_{4g} = \begin{cases} 20 & : \bar{q} \leq 1 \\ 18e^{-3(\bar{q}-1)} + 2 & : \bar{q} > 1 \end{cases} \quad (5.18)$$

Eine vom Mittelpunkt zum Rand der Klasse hin abnehmende Dichte der Objekte führt zu einem $\bar{q} > 1$. Das stellt in der Regel den Normalfall dar. Sind die Abstände der Objekte in der Klasse alle gleich, entspricht das einer Gleichverteilung mit einer scharfen Abgrenzung der Klasse. In diesem Fall ist $\bar{q} = 1$ und der Wert von β_{4g} wird auf 20 gesetzt, was nahezu zu einer Rechteckfunktion für die Zugehörigkeitsfunktion führt. Für den Fall, dass die Dichte der Objekte zum Rand hin zunimmt, was einem $\bar{q} < 1$ entspricht, wird der Wert von β_{4g} ebenfalls auf 20 gesetzt.

In PÄSSLER (1998) sind noch weitere Möglichkeiten zur Berechnung von β_{4g} angegeben, auf die hier aber nicht näher eingegangen werden soll.

5.3 Neueste Entwicklungen

Nachdem über viele Jahre hinweg die Anwendung der Methode der Fuzzy Pattern Klassifikation in einem weiten Spektrum im Vordergrund stand, wurden in den letzten Jahren einige methodische Weiterentwicklungen vorgenommen.

Wie es bereits zu Beginn des Abschnittes 5 in einer Fußnote erwähnt wurde, erfolgte für diese Methode eine Erweiterung, in der für jedes Objekt eine individuelle Unschärfe für jedes Merkmal eingeführt werden kann und nicht nur eine gemeinsame (mittlere) Unschärfe für alle Objekte (HEMPEL, 2005). Dadurch wird diese Methode für viele Anwendungen in der Geodäsie sicherlich noch interessanter. Ein Beispiel, wie sich die Berücksichtigung individueller Unschärfen, im Gegensatz zu mittleren Unschärfen, auf die Zugehörigkeitsfunktionen auswirkt, wird im Abschnitt 6.2.2 noch präsentiert.

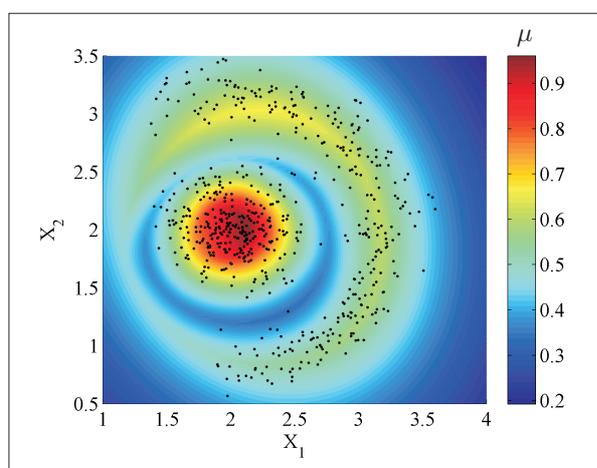


Abbildung 5.9: Nichtkonvexe Zugehörigkeitsfunktion über einem zweidimensionalen Merkmalsraum mit farbcodierten Zugehörigkeitswerten (HEMPEL und BOCKLISCH, 2009).

Aktuell wird an der TU Chemnitz daran gearbeitet, die Methode dahingehend zu erweitern, dass damit auch nichtkonvexe Zugehörigkeitsfunktionen (Abbildung 5.9) modelliert werden können (HEMPEL und BOCKLISCH, 2009). Prinzipiell funktioniert das so, dass zunächst für die gesamte Klasse eine konvexe Zugehörigkeitsfunktion nach dem oben beschriebenen Ablauf aufgestellt wird. Diese Zugehörigkeitsfunktion wird als fuzzy pattern class bezeichnet. Für die Bereiche geringerer Objektpunktdichte innerhalb der Klasse, in denen später die Zugehörigkeitsfunktion einen konkaven Verlauf aufweisen soll, werden sogenannte Antiobjekte definiert und aus ihnen eine weitere Zugehörigkeitsfunktion berechnet. Diese Zugehörigkeitsfunktion wird als fuzzy pattern anticlass bezeichnet. Die endgültige nichtkonvexe Zugehörigkeitsfunktion berechnet sich aus dem Minimum der fuzzy pattern class und dem Komplement der fuzzy pattern anticlass (Gleichung (5.19)).

$$\mu = \min(\mu_{class}, (1 - \mu_{anticlass})) \quad (5.19)$$

- μ : Zugehörigkeitswert der nichtkonvexen Zugehörigkeitsfunktion
- μ_{class} : Zugehörigkeitswert der Klassenbeschreibung für die konvexe Klasse
- $\mu_{anticlass}$: Zugehörigkeitswert der Klassenbeschreibung für die Antiklasse

In der Abbildung 5.10 ist das für eine ringförmige Klasse dargestellt. Das Hauptproblem bei dieser Vorgehensweise liegt in der entsprechenden Wahl von Antiobjekten.

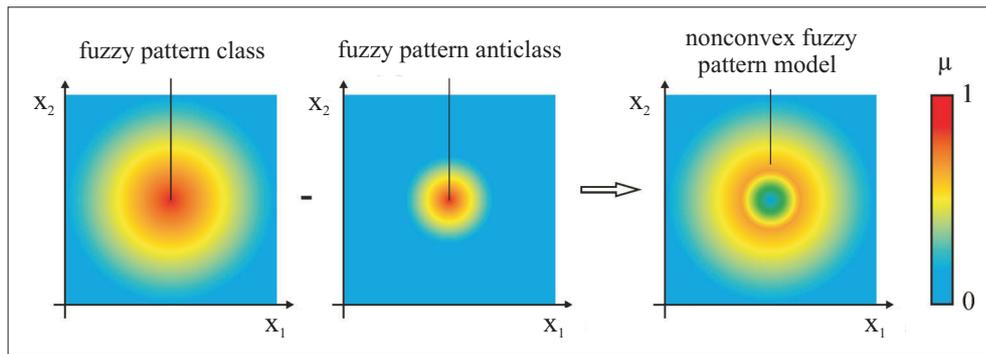


Abbildung 5.10: Erzeugung einer nichtkonvexen Zugehörigkeitsfunktion durch Subtraktion (HEMPEL und BOCKLISCH, 2009).

Kapitel 6

Übertragung auf Deformationsuntersuchungen

Zu Beginn dieses Kapitels (Abschnitt 6.1) soll kurz ein Überblick über die Anwendung von Mustererkennungsverfahren in verschiedenen Bereichen der Geodäsie gegeben werden. Anschließend folgt dann die Übertragung der bisher dargelegten Ansätze auf die Analyse von Überwachungsmessungen. Als typische Vertreter werden im ersten Beispiel Punktfelder von Verschiebungsvektoren im Hinblick auf Blöcke konsistenter Punktbewegungen untersucht (Abschnitt 6.2), und im zweiten Beispiel wird das zeitliche Setzungsverhalten von Objektpunkten über mehrere Messepochen charakteristischen Setzungsmustern zugeordnet (Abschnitt 6.3).

6.1 Vorbemerkungen zur Anwendung der Mustererkennung in der Geodäsie

In vielen geodätischen Bereichen kommen Verfahren der Mustererkennung zur Anwendung, auch wenn der Begriff „Mustererkennung“ nicht immer explizit in ihrem Zusammenhang erwähnt wird. Sie unterscheiden sich teilweise erheblich in Bezug auf die zu analysierenden Muster, die Merkmale, die Merkmalsextraktion und -vorverarbeitung sowie das eingesetzte Klassifizierungsverfahren. Bei der Klassifizierung kommen überwiegend Verfahren zum Einsatz, die in irgendeiner Art und Weise Expertenwissen als Vorabinformation nutzen. Automatische Klassifizierungen, basierend auf rein formalen Kriterien, wie sie in dieser Arbeit im Vordergrund stehen, sind eher seltener zu finden. Die dargestellten Anwendungsfelder, Beispiele und Methoden erheben keinen Anspruch auf Vollständigkeit, da die Anzahl an Publikationen zu dem Thema inzwischen unüberschaubar groß geworden ist. Die weitaus größte Anwendung finden Verfahren der Mustererkennung im Bereich der Photogrammetrie (insbesondere in der Fernerkundung) sowie in der Kartographie.

Auf eine Entwicklung von HOBROUGH (1959) geht in der **Photogrammetrie** die automatische Ableitung von Höhen aus Stereobildern durch Korrelation einander entsprechender Bildausschnitte zurück. Ein Verfahren, welches durch die Entwicklung der digitalen Photogrammetrie in den letzten Jahren stark an Bedeutung gewonnen hat (LINDNER, 2007).

Die Anfänge der automatischen Klassifikation in der Photogrammetrie finden sich in der Fernerkundung in den 1970er Jahren als Teil der automatisierten Bildinterpretation (KONECNY, 1984). Als Merkmale wurden benutzt:

- Grauwertverteilungen im Histogramm einzelner spektraler Kanäle. (Abbildung 6.1 a))

- Grauwerte mehrerer unterschiedlicher spektraler Kanäle zu gleichen Zeitpunkten. Auch multispektrale Klassifikation genannt. (Abbildung 6.1 b))
- Grauwerte gleicher spektraler Kanäle zu unterschiedlichen Zeitpunkten. Auch multitemporale Klassifikation genannt.

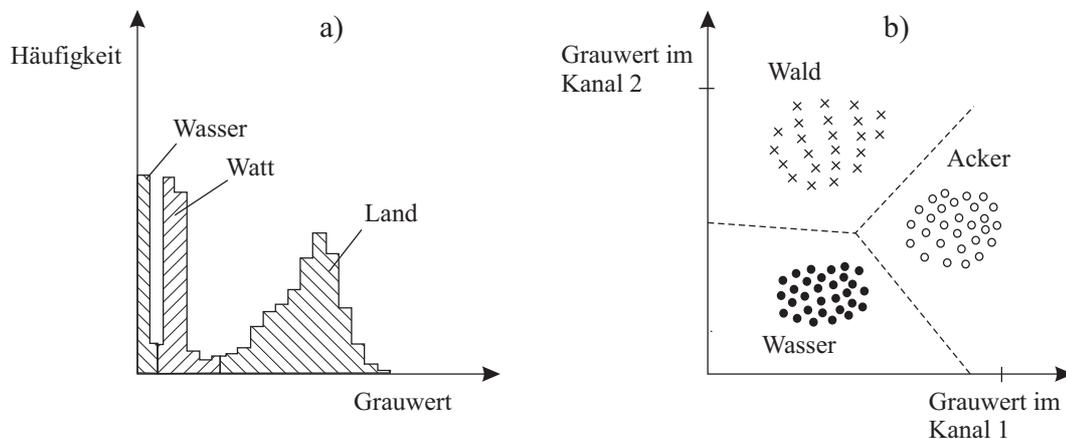


Abbildung 6.1: Klassifikation von Fernerkundungsdaten: (nach KONECNY, 1984)

a) Klassifikation anhand der Grauwertverteilung eines Kanals.

b) Multispektrale Klassifikation in einem zweidimensionalen Merkmalsraum.

Die notwendigen digitalen Daten (Bildpixel) lieferten einige Fernerkundungssensoren direkt. Analoge Photos wurden gescannt. Als Klassifizierungsverfahren wurde überwiegend das Maximum-Likelihood Verfahren auf der Basis von bekannten Trainingsgebieten verwendet (KRAUS, 1979). Später wurden auch unüberwachte Klassifizierungsverfahren (z.B. die im Abschnitt 2.4.2 beschriebenen partitionierenden Verfahren) eingesetzt (DE LANGE, 2002). Die Verfahren der multispektralen- bzw. multitemporalen Klassifikation sind heute in der Fernerkundung Standardverfahren (ALBERTZ, 2007).

Im Bereich der Merkmalsvorverarbeitung kommen verschiedene Verfahren der digitalen Bildverarbeitung zum Einsatz. Diese sind insbesondere radiometrische Verbesserungen (z.B. zur Elimination atmosphärischer Einflüsse), Entzerrungen, Georeferenzierungen, Kontrastverbesserungen, Filteroperationen zur Hervorhebung oder Unterdrückung bestimmter Eigenschaften des Bildinhaltes und evtl. die Kombination der Inhalte mehrerer Bilder oder Spektralkanäle. Zur Merkmalsreduktion dient häufig eine Hauptkomponententransformation, um die wesentlichen Informationen in wenigen Kanälen zu konzentrieren und Korrelationen zwischen Merkmalen zu eliminieren. Neben den Grauwertinformationen der Pixel werden heute auch weitere Merkmale, wie zum Beispiel Relief- oder geologische Informationen zur automatischen Klassifizierung mit einbezogen. Diese Informationen können heute direkt über Geo-Informationssysteme zur Verfügung gestellt werden (ALBERTZ, 2007).

Mit zunehmender Bedeutung der digitalen Photogrammetrie in den letzten Jahren wurden die in der Fernerkundung bewährten Methoden auch auf die Luftbildphotogrammetrie angewendet. Als problematisch erwiesen sich die pixelbasierten Verfahren bei zunehmend höherer Auflösung der Sensoren und zunehmender Anwendung im großmaßstäbigen Bereich. Es wurde dazu übergegangen, nicht nur die Informationen einzelner Pixel zu nutzen, sondern ganzer Segmente, die man aus benachbarten Pixeln mit ähnlichen Grauwerten bildet, bis hin zur direkten automatischen Ableitung von Vektordaten und daraus der automatischen Erkennung von geometrischen Objekten (z.B. Gebäuden, Bäumen, Straßen, Fahrzeugen). Die direkte Erkennung von geometrischen Objekten erfolgt nur anhand von Vorwissen, das in verschiedenen Formen bereitgestellt wird. Man nutzt dabei zum Beispiel Informationen über Geradlinig- und Rechtwinkligkeit, Symmetrien sowie einfache geometrische Primitive als Vergleichsmuster (HEIPKE, 2003, 2008). Suchräume für Objekte können durch Einführung von Zusatzwissen aus Geo-Informationssystemen stark eingegrenzt werden (z.B. LEITLOFF U.A., 2006).

Im Bereich der Nahbereichsphotogrammetrie sind die Hauptanwendungen der Mustererkennung in der automatischen Erkennung von Reseaukreuzen und verschiedener (unter Umständen codierter) Ziel-

marken für die automatische Punktmessung. Hierbei kommen Verfahren der Schwerpunktberechnung, Kantenextraktion, Korrelation und Regression zum Einsatz (LUHMANN, 2003). Für einige Spezialanwendungen im industriellen Einsatz (z.B. in der Automobilindustrie und im Anlagenbau) werden auch automatische Objekterkennungen durch sog. Konturverfahren durchgeführt, die ebenfalls auf Kantenextraktionen und anschließender Regression oder Korrelation basieren.

In der **Kartographie** wird die Mustererkennung im Wesentlichen in zwei Anwendungsgebieten eingesetzt:

1. Die Konvertierung von Raster- in Vektordaten und die anschließende Erkennung von punkt-, linien- oder flächenhaften Objekten sowie von Beschriftungszeichen aus einer gescannten Kartenvorlage. Diese Aufgabenstellung ähnelt den oben beschriebenen Aufgabenstellungen in der Photogrammetrie. Gegenüber der Photogrammetrie vereinfachen sich die Problemstellungen in der Kartographie dadurch, dass in der Regel das kartographische Zeichensystem (Symbole, Schriftzeichen, Linienformen und Strichstärken) a priori bekannt ist und die radiometrischen Variationen sehr gering sind. Wie in der Photogrammetrie, werden hier ebenfalls die verschiedensten Verfahren eingesetzt, die auf der Basis von Vorwissen arbeiten. (GRAEFF, 2002)
2. Die Modellgeneralisierung (SESTER U.A., 2008), bei der ein digitales Objektmodell mit geringerer Modellauflösung aus einem Modell höherer geometrischer und semantischer Auflösung abgeleitet wird. Bei dieser automatischen Ableitung sollen charakteristische Strukturen möglichst erhalten bleiben und zur Erkennung solcher charakteristischen Strukturen werden Verfahren der Mustererkennung eingesetzt. ANDERS (2007) beschäftigt sich in dem Kontext beispielsweise mit der Erkennung spezieller Strukturen in Straßennetzwerken oder CHRISTOPHE und RUAS (2002) mit der Erkennung von Siedlungsstrukturen. Hier wird bei der Erkennung ebenfalls Vorwissen in Form von einfachen geometrischen Strukturen als Vergleichsmuster genutzt.

Im Bereich der **Auswertung von Laserscanner - Daten** (sowohl im terrestrischen als auch im Airborn Laserscanning) stellt die Erkennung von geometrischen Objekten aus den Punktwolken einen Schwerpunkt dar. Die Problemstellung und die dabei eingesetzten Methoden sind nahezu identisch zu den oben beschriebenen in der Photogrammetrie und Kartographie. Beim Laserscanning kommen ebenfalls Verfahren der Kantenextraktion und Regression zum Einsatz (z.B. KERN, 2003). Weiterhin werden wissensbasierte Klassifikationen, zum Beispiel auf der Basis trainierter künstlicher neuronaler Netze (SCHLEINKOFER, 2007) sowie regelbasierte Verfahren (z.B. PLÜMER U.A., 2009) verwendet.

Bei der **Nutzung von GIS** (Geo-Informationssysteme) werden Klassifizierungsverfahren überwiegend als Analysewerkzeuge genutzt (REINHARD U.A., 2004, DE LANGE, 2002). Dabei kommen im Wesentlichen die im Abschnitt 2.4.2 beschriebenen hierarchischen und partitionierenden Clusterverfahren zum Einsatz.

In der **geodätischen Messtechnik** kommen Mustererkennungsverfahren beispielsweise bei der digitalen Ablesung an codierten Nivellierlatten, codierten Teilkreisen oder der elektrooptischen Entfernungsmessung zum Einsatz. Hierbei werden Korrelations-, analytische- oder geometrische Verfahren benutzt. (DEUMLICH und STAIGER, 2002)

Im Bereich der **Grundstückswertermittlung** werden für viele Bereiche Klassifizierungen zur Ableitung wertbeeinflussender Faktoren genutzt. Diese Klassifizierungen werden überwiegend durch Experten anhand inhaltlicher Kriterien durchgeführt. Clusterungen aufgrund formaler Kriterien finden nur selten Anwendung. Beispiele hierzu sind bei BRAUER (1986) und QU (2000) zu finden, die diese Methoden zur Strukturierung von Kaufpreissammlungen angewendet haben. BRAUER (1986) hat dazu die im Abschnitt 2.4.2 beschriebenen hierarchischen und partitionierenden Clusterverfahren benutzt und QU (2000) den im Abschnitt 4.3 beschriebenen Fuzzy-c-Means Algorithmus.

Die in weiten Teilen der Geodäsie benutzten statistischen Tests (z.B. NIEMEIER, 2008) können als Klassifizierungsverfahren mit binärem Entscheidungsraum ebenfalls als Verfahren der Mustererkennung bezeichnet werden.

6.2 Auffinden von Punktgruppen mit gleichförmigem Bewegungsverhalten in einem Überwachungsnetz

Wie bereits in der Einführung zu dieser Arbeit erwähnt, enden die in der geodätischen Praxis heute verwendeten Auswerteprozesse von geodätischen Überwachungsmessungen in der Regel mit der Angabe von Verschiebungsvektoren und einem Test auf Signifikanz, häufig kombiniert mit einer graphischen Darstellung, in der die Objektpunkte und ihre Verschiebungsvektoren dargestellt sind. Die weitergehende Interpretation im Hinblick auf Gruppen ähnlicher Verschiebungsvektoren bleibt in der Regel dem Bearbeiter überlassen, der durch visuelle Beurteilung der graphischen Darstellung irgendwelche Bewegungsmuster zu erkennen versucht. Für einen geschulten Bearbeiter stellt das in der Regel kein großes Problem dar. Wenn jedoch mit Hilfe automatischer Messverfahren zum Beispiel größere Berghänge oder Böschungen im Tagebau mit sehr vielen Objektpunkten überwacht werden, gibt es den Bedarf einer automatischen Interpretation der Verschiebungsvektoren. Diese Monitoringsysteme mit angeschlossener Online-Alarmierung sollen möglichst ohne einen interpretierenden Bearbeiter auskommen. Es ist nun notwendig, das Expertenwissen eines geschulten Bearbeiters in so ein System zu integrieren oder es durch formale Algorithmen zu ersetzen. Dazu gibt es eine Reihe von Möglichkeiten. Einen speziellen Ansatz hierzu stellte zum Beispiel HABERLER (2005) mit Hilfe von Fuzzy-Regeln vor, der sich aber grundlegend von der hier gezeigten Vorgehensweise unterscheidet. Während in HABERLER (2005) Expertenwissen in einer Regelbasis abgelegt ist (Zitat: „Wenn die Azimute der Vektoren ähnlich sind, liegen die Punkte wahrscheinlich auf einem gemeinsamen Block.“), wird in dieser Arbeit versucht, ohne vorliegendes Expertenwissen, gemeinsame Strukturen von Objekten zu finden, die durch ihre Merkmale beschrieben sind.

Bei dieser Aufgabenstellung geht es darum, unbekannte Strukturen in den Daten aufzudecken. Es liegen keine a priori Informationen über die Clusterzentren und Clusterformen vor, genauso wenig existiert eine repräsentative Stichprobe von Daten, mit denen ein Klassifikator „angelernt“ werden könnte. Die Parameter für die Clusterzentren und evtl. Clusterformen müssen als Unbekannte im Auswerteprozess mit bestimmt werden.

6.2.1 Testnetz Delft

Am Beispiel des bereits mehrfach erwähnten Testnetzes Delft (Abbildung 6.2) soll das Auffinden von Punktgruppen mit gleichförmigem Bewegungsverhalten mit Hilfe von Fuzzy-Clusterverfahren demonstriert werden. Dieses Beispiel wurde auszugsweise bereits in FLETTLING (2008) in englischer Sprache veröffentlicht.

Clusterung der Objektpunkte

Die Merkmale, anhand deren ein geschulter Bearbeiter die Verschiebungsvektoren beurteilt und anhand deren er auf Punktgruppen mit gleichem Bewegungsverhalten schließt, sind in einem zweidimensionalen Überwachungsnetz die Länge und das Azimut der Verschiebungsvektoren. Diese Merkmale sind auch die Eingangsgrößen für den hier vorgestellten Ansatz. Um zu verhindern, dass in diesem Beispiel die Unstetigkeitsstelle bei 0 gon im Azimut der Verschiebungsvektoren genau durch das Cluster 1 verläuft, wurde das Netz willkürlich um 10 gon in östliche Richtung gedreht.

Das Drehen des Netzes muss nicht durchgeführt werden, wenn man als Merkmale den Sinus und den Cosinus des Azimutes einführt. Das hätte für dieses Beispiel aber zur Folge, dass man einen weniger anschaulichen dreidimensionalen Merkmalsraum bekäme. Eine weitere leicht zu automatisierende Möglichkeit besteht darin, die Verschiebungsvektoren nach aufsteigenden Azimuten zu sortieren und nach dem größten Azimutabstand zwischen jeweils zwei benachbarten Vektoren zu suchen. Das Netz könnte man anschließend soweit drehen, dass die Nullrichtung in diese Lücke käme.

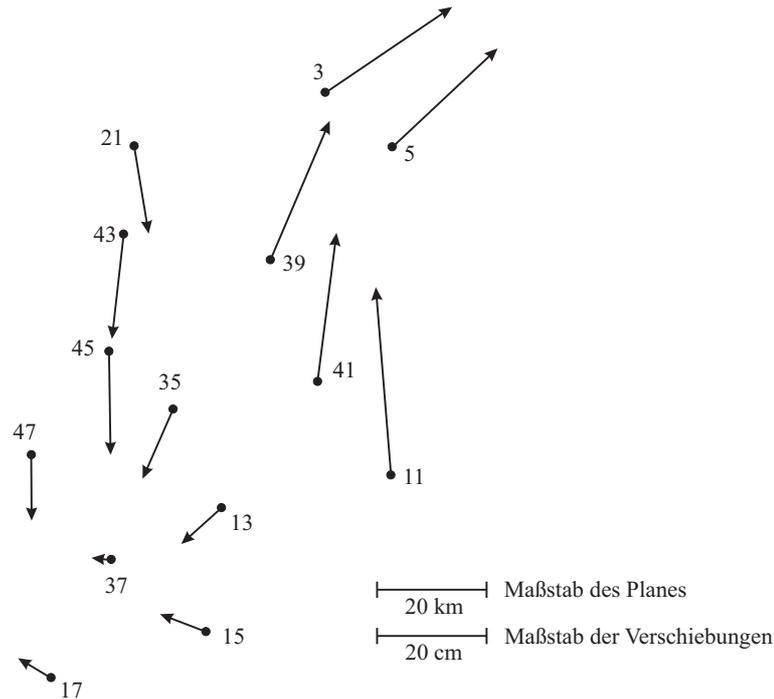


Abbildung 6.2: Testnetz Delft: Objektpunkte mit ihren Verschiebungsvektoren zwischen den Epochen 1 und 3b.

Für einen zweidimensionalen Merkmalsraum, wie er hier vorliegt, besteht noch folgende Möglichkeit (HERBST, 2010): Der Merkmalsraum wird gedanklich auf den Mantel eines geraden Kreiszyinders abgebildet. Dabei verläuft die Längsachse auf einer Mantellinie des Zylinders und die Azimutachse orthogonal dazu auf dem Mantel um den Zylinder herum und zwar so, dass der Wertebereich $[0, 400]$ gon genau dem Umfang entspricht. Die Distanzen zwischen Objektpunkten in dem Merkmalsraum könnten dann direkt, ohne Unstetigkeit auf dem Zylindermantel berechnet werden. Da die Distanz auf dem Zylindermantel zweideutig ist, müssten beide Distanzen berechnet und dann das Minimum genommen werden.

Um sinnvolle Ergebnisse im Clusterungsprozess zu bekommen, müssen die Merkmale auf den unterschiedlichen Achsen die gleiche Größenordnung haben. Sollte das nicht der Fall sein, so ist eine Normierung auf einen einheitlichen Wertebereich, insbesondere bei Anwendung des Fuzzy-c-Means Algorithmus und seiner Modifikationen, notwendig. In diesem Beispiel wurden beide Merkmalsachsen zunächst auf das Intervall $[0, 1]$ normiert¹. Bei der Bearbeitung weiterer Beispiele (insbesondere des Beispiels ISNET (Abschnitt 6.2.2)) zeigte sich, dass die Clusteralgorithmen stärker dazu tendieren, Punktgruppen in Richtung der Azimutachse zusammenzufassen als ein menschlicher Auswerter das tun würde. Von daher wurde in dieser Arbeit dazu übergegangen, die Länge der Verschiebungsvektoren auf das Intervall $[0, 0,5]$ zu normieren. Dadurch wird erreicht, dass die Algorithmen Punktgruppen in Richtung der Längsachse stärker zusammenfassen als bei einer Normierung beider Achsen auf das gleiche Intervall. Die Ergebnisse entsprechen dadurch wesentlich besser einer Bewertung eines erfahrenen Auswerters. Bei der Normierung auf die Intervalle $[0, 1]$ (Azimut) und $[0, 0,5]$ (Länge) entspricht die maximale Länge eines Verschiebungsvektors in seiner Wertigkeit einer Azimutänderung von 200 gon. In dem hier vorgestellten Beispiel des Testnetzes Delft ergeben sich durch die Änderung der Normierung keine Änderungen im Endergebnis. Lediglich die Zugehörigkeitswerte einzelner Objektpunkte ändern sich minimal.

Wie bereits im Abschnitt 4.6 dargelegt, ergibt sich für den hier behandelten Fall eine optimale Anzahl von drei Clustern (Abbildung 6.3).

¹ Für die Länge der Verschiebungsvektoren wurde das Intervall $[0, 40]$ cm und für das Azimut das Intervall $[0, 400]$ gon auf das Intervall $[0, 1]$ normiert.

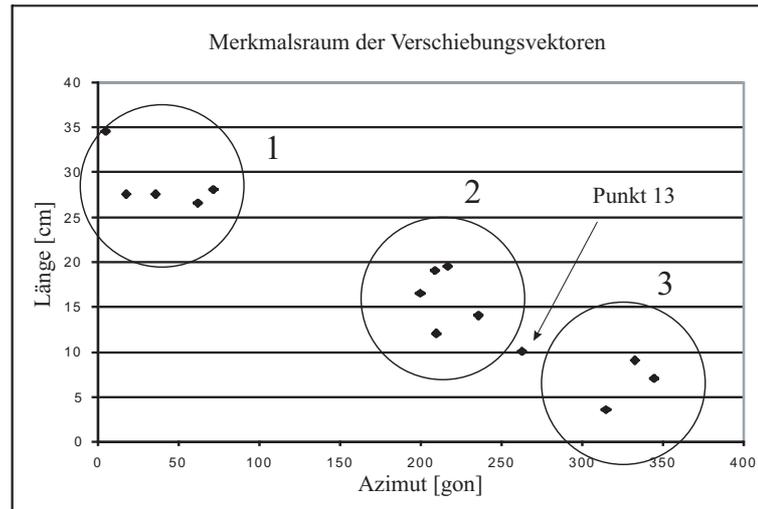


Abbildung 6.3: Testnetz Delft (gedreht): Verschiebungsvektoren zwischen den Epochen 1 und 3b im Merkmalsraum und die ungefähre Lage der drei Cluster.

In der Tabelle 6.1 sind die Ergebnisse der Clusterungen mit Hilfe des Fuzzy-c-Means Algorithmus und den modifizierten Fuzzy-Clusterverfahren nach GUSTAFSON und KESSEL (1979) und GATH und GEVA (1989) angegeben. In allen drei Fällen wurde der Unschärfeparameter $w = 2$ gewählt.

Bei der Clusterung mit Hilfe des **Fuzzy-c-Means Algorithmus** ergeben sich hier für fast alle Objektpunkte relativ hohe maximale Zugehörigkeitswerte (Hauptzugehörigkeitswerte), was für eine sehr gute Zuordnung der Objekte zu den entsprechenden Clustern spricht. Einzig der Punkt 13 weist einen Hauptzugehörigkeitswert von 0,57 auf und der nächst kleinere Zugehörigkeitswert beträgt 0,40. Für diesen Punkt wäre eine scharfe Zuordnung zu einem Cluster eventuell problematisch. In der Abbildung 6.3 ist der Punkt 13 besonders gekennzeichnet und man sieht, dass er zwischen den Clustern 2 und 3 liegt, etwas näher zum Cluster 2.

Bei der Clusterung mit Hilfe des **Gustafson-Kessel Algorithmus** ergeben sich ähnlich wie beim Fuzzy-c-Means für fast alle Objektpunkte relativ hohe Hauptzugehörigkeitswerte und damit gute Zuordnungen zu den einzelnen Clustern. Für den Punkt 13 erhöht sich der Hauptzugehörigkeitswert zum Cluster 2 auf 0,76 und der höchste Nebenzugehörigkeitswert zum Cluster 3 verringert sich auf 0,19. So resultiert hier für den Punkt 13 eine eindeutiger Zuordnung zum Cluster 2.

Wie bereits im Abschnitt 4.4.2 erwähnt, führt der **Gath-Geva-Algorithmus** zu vergleichsweise scharfen Clustereinteilungen. Der Übergangsbereich, in dem die Zugehörigkeitswerte zwischen 0 und 1 wechseln, ist sehr schmal. Das hat in diesem Beispiel zur Folge, dass sich nur Zugehörigkeitswerte von 0 oder 1 ergeben. Alle Punkte werden scharf einem Cluster zugeordnet, auch der Punkt 13 wird scharf dem Cluster 2 zugeordnet. Um mit dem Gath-Geva-Algorithmus unschärfere Ergebnisse zu bekommen, ist für den Parameter w ein größerer Wert als 2 zu wählen.

In Abhängigkeit von den Näherungswerten ergeben sich allerdings bei dem Gath-Geva Algorithmus sehr unterschiedliche Ergebnisse für die Zuordnungen. Hier reagiert der Algorithmus sehr sensibel. Kleine Änderungen bei den Näherungswerten können zu völlig anderen Clusterlösungen führen. Eine automatische Durchführung einer Clusterung mit Hilfe dieses Algorithmus ohne Berücksichtigung von Vorwissen kann nicht empfohlen werden. Das gilt etwas eingeschränkt auch für den Gustafson-Kessel Algorithmus. Der Fuzzy-c-Means Algorithmus ist dagegen als sehr robust zu bezeichnen. Bei dem hier vorliegenden Beispiel konnten die Näherungswerte nahezu beliebig gewählt werden. Es ergaben sich immer die gleichen Clusterlösungen. Es waren lediglich mehr oder weniger Iterationsschritte notwendig.

Bei dem Verfahren der **Fuzzy Pattern Klassifikation** erfolgt zunächst eine Clustereinteilung nach einem scharfen Clusterverfahren. Darauf aufbauend werden dann unter Berücksichtigung der elemen-

Punkt	Fuzzy-c-Means			Gustafson-Kessel			Gath-Geva		
	Cluster			Cluster			Cluster		
	1	2	3	1	2	3	1	2	3
3	0,94	0,04	0,02	0,95	0,04	0,01	1,00	0,00	0,00
5	0,97	0,02	0,01	0,98	0,02	0,00	1,00	0,00	0,00
11	0,95	0,04	0,01	0,96	0,03	0,01	1,00	0,00	0,00
13	0,03	0,57	0,40	0,05	0,78	0,17	0,00	1,00	0,00
15	0,00	0,02	0,98	0,00	0,02	0,98	0,00	0,00	1,00
17	0,00	0,02	0,98	0,00	0,02	0,98	0,00	0,00	1,00
21	0,01	0,97	0,02	0,03	0,96	0,01	0,00	1,00	0,00
35	0,01	0,95	0,04	0,01	0,98	0,01	0,00	1,00	0,00
37	0,01	0,03	0,96	0,01	0,05	0,94	0,00	0,00	1,00
39	1,00	0,00	0,00	0,99	0,01	0,00	1,00	0,00	0,00
41	0,99	0,01	0,00	0,97	0,02	0,01	1,00	0,00	0,00
43	0,01	0,97	0,02	0,03	0,95	0,02	0,00	1,00	0,00
45	0,01	0,97	0,02	0,02	0,97	0,01	0,00	1,00	0,00
47	0,02	0,95	0,03	0,06	0,89	0,05	0,00	1,00	0,00

Tabelle 6.1: Mit Hilfe des Fuzzy-c-Means, des Gustafson-Kessel und des Gath-Geva Algorithmus berechnete Zugehörigkeitswerte der Objektpunkte zu drei Clustern. Der jeweilige Hauptzugehörigkeitswert eines Punktes ist fett hervorgehoben.

taren Unschärfen der Merkmale die Zugehörigkeitsfunktionen für die einzelnen Cluster berechnet. Aus denen wiederum ergeben sich die possibilistischen Zugehörigkeitswerte der einzelnen Objektpunkte zu den Clustern.

Die scharfe Clustereinteilung wurde in diesem Beispiel nach dem im Abschnitt 2.4.2 beschriebenen Complete Linkage Verfahren durchgeführt. Aus dem dabei entstandenen Dendrogramm konnte ebenfalls eine gute Clusteranzahl von drei Clustern entnommen werden, wobei das Dendrogramm auch eine Clusteranzahl von zwei Clustern nahelegt. Eine ebenfalls durchgeführte Clustering nach dem Average Linkage Verfahren führt zu einem identischen Ergebnis. Eine Clustering nach dem Single Linkage Verfahren ergibt aufgrund der Neigung kettenförmige Cluster zu bilden, keine sinnvolle Lösung.

Als elementare Unschärfen wurden für die Länge des Verschiebungsvektors 50 mm und für das Azimut 15 gon angenommen². Der asymptotische Grenzwert von β_{1g} (die Höhe der Zugehörigkeitsfunktion) wurde auf 10 gesetzt.

Die geringe Anzahl von Objektpunkten und ihre Verteilung in einigen Clustern sorgen dafür, dass teilweise für diese Anwendung schwer interpretierbare Clusterformen entstehen. In der Abbildung 6.4 sind die Linien gleicher Zugehörigkeiten für die drei Cluster dargestellt. Die Berechnung des Parameters β_{4g} , der die Steilheit des Abfalls der Zugehörigkeitsfunktion zum Rande hin bewirkt und seine Auswirkung auf den Parameter β_{2g} (siehe Abschnitt 5) führt dazu, dass die Clusterformen sehr unsymmetrisch werden und dass teilweise auf der einen Seite der Hauptschwereachse des Clusters die Zugehörigkeitsfunktion ähnlich einer scharfen Clustergrenze sehr schnell abfällt, während auf der anderen Seite ein langsam abfallender kontinuierlicher Übergang vorhanden ist. Ohne Kenntnis des Berechnungsschrittes des Parameters β_{4g} erscheint diese Unsymmetrie einem Anwender nicht plausibel. Der Berechnungsschritt zur Berechnung des Parameters β_{4g} reagiert mit den hier vorliegenden

² Diese Werte entsprechen ungefähr den Standardabweichungen eines Verschiebungsvektors mittlerer Länge. Diese wurden durch Fehlerfortpflanzung aus den Standardabweichungen der Punkte ermittelt, wie sie sich aus den freien Ausgleichungen der einzelnen Epochen ergaben.

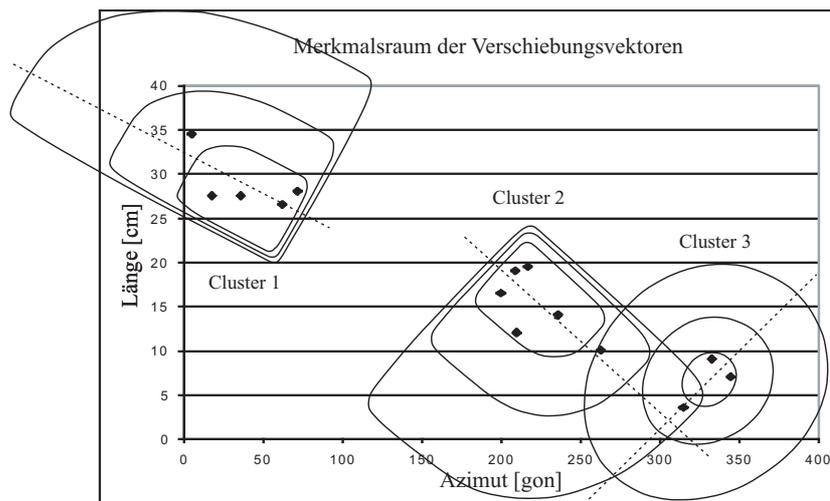


Abbildung 6.4: Mit Hilfe der Fuzzy Pattern Klassifikation berechnete Linien gleicher Zugehörigkeiten für die drei Cluster im Merkmalsraum. Die Isolinien repräsentieren von innen nach außen die Zugehörigkeitswerte $\mu = 0,9, 0,6$ und $0,3$. Die gestrichelten Linien sind die Hauptschwereachsen der Cluster.

Daten sehr sensibel auf die Lage der Punkte im Merkmalsraum. Geringe Veränderungen der Punktlagen einzelner Punkte können in diesem Beispiel, insbesondere was die Steilheit des Abfalls zum Rand hin angeht, zu einem völlig anderen Aussehen der Zugehörigkeitsfunktionen führen. Das Problem ist, dass hier in einem zweidimensionalen Merkmalsraum für jede Klassenbeschreibung, d.h. die Zugehörigkeitsfunktion, insgesamt 16 Parameter bestimmt werden müssen (2 Translationsparameter, 1 Rotationsparameter, 1 Parameter für die Höhe und 4×3 Parameter für die Form der Zugehörigkeitsfunktion (siehe Abschnitt 5.1)). Das führt in dem hier vorliegenden Fall zu unterbestimmten Aufgabenstellungen und ist nur lösbar, weil die Merkmalsausprägungen der Punkte zur Berechnung der Parameter mehrfach benutzt werden.

Das Verfahren der Fuzzy Pattern Klassifikation sollte so modifiziert werden, dass bei wenigen Objektpunkten solche starken Unsymmetrien nicht auftreten können. Zum Beispiel könnte man in der Formel (5.18) bei der Berechnung des Parameters β_{4g} eine Abhängigkeit von der Anzahl der Objektpunkte einführen.

Die Zugehörigkeitswerte bei dem Verfahren der Fuzzy Pattern Klassifikation berechnen sich nicht, wie bei den vorher beschriebenen Verfahren, direkt aus den Distanzen im Merkmalsraum, sondern aus der Zugehörigkeit der Objektmerkmale zu den Einzelmerkmalen, die das Cluster aufbauen. Dadurch ist das Verfahren theoretisch unabhängig von den Dimensionen der einzelnen Merkmalsachsen. Eine Normierung der Merkmalsausprägungen ist nicht unbedingt notwendig. In diesem Fall wurde aber mit normierten Merkmalen gearbeitet, da sie für die vorhergegangene scharfe Clusterung benötigt wurden.

Die nach dem Verfahren der Fuzzy Pattern Klassifikation berechneten Zugehörigkeitswerte (Tabelle 6.2 und Tabelle 6.4) sind nicht direkt mit den Zugehörigkeitswerten vergleichbar, die mit dem Fuzzy-C-Means Algorithmus bzw. den modifizierten Algorithmen (Tabelle 6.1) ermittelt wurden. Bei den nach dem Verfahren der Fuzzy Pattern Klassifikation berechneten Zugehörigkeitswerten handelt es sich um possibilistische Werte. Das heißt, die Summe der Zugehörigkeitswerte eines Objektes zu allen Klassen ist in der Regel $\neq 1$. Darüber hinaus ist die Höhe (d.h. das Maximum der Zugehörigkeitsfunktion) bei diesem Verfahren theoretisch nicht auf den Wert 1 begrenzt. In den Ergebnistabellen dieser Arbeit ist sie allerdings zur besseren Vergleichbarkeit auf den Wert 1 normiert worden.

Fuzzy Pattern-Klassifikation				
Punkt	C_S	Cluster		
		1	2	3
3	1	0,97	0,00	0,03
5	1	1,00	0,00	0,03
11	1	0,80	0,00	0,02
13	2	0,00	0,83	0,31
15	3	0,00	0,02	0,92
17	3	0,00	0,02	0,92
21	2	0,00	0,96	0,10
35	2	0,00	0,99	0,17
37	3	0,00	0,33	0,85
39	1	1,00	0,00	0,02
41	1	0,98	0,00	0,02
43	2	0,00	1,00	0,10
45	2	0,00	1,00	0,10
47	2	0,00	0,93	0,13

Tabelle 6.2: Mit Hilfe der Fuzzy Pattern Klassifikation berechnete Zugehörigkeitswerte der Objektpunkte zu drei Clustern. Der jeweilige Hauptzugehörigkeitswert eines Punktes ist fett hervorgehoben. In der Spalte C_S ist das Cluster angegeben, dem der Punkt mit Hilfe des Complete Linkage Verfahrens scharf zugewiesen wurde.

Berücksichtigung der Nachbarschaft

Der bisherige Auswertevorgang zeigt drei Cluster, zu denen die Objektpunkte überwiegend sehr gut zugeordnet werden können. Diese drei Cluster repräsentieren auch gut die drei Blöcke von Punkten, die ein gleichförmiges Bewegungsverhalten aufweisen. Allerdings ist in dem bisherigen Auswertevorgang die Nachbarschaft der Punkte nicht berücksichtigt worden. Theoretisch könnten zwei Punkte, die hier zu einem gemeinsamen Cluster einen Hauptzugehörigkeitswert aufweisen, räumlich völlig getrennt voneinander an unterschiedlichen Rändern des Netzes liegen und dazwischen könnte sich ein Block mit einem völlig anderen Bewegungsverhalten befinden. Deshalb ist es notwendig, in einem weiteren Schritt die Nachbarschaftsbeziehungen der Punkte mit zu berücksichtigen. Dieses erfolgt dadurch, dass zwischen den Objektpunkten mit Hilfe einer Delaunay-Triangulation eine Dreiecksvermaschung erzeugt wird und die sich dadurch ergebenden Linienverbindungen zwischen den Objektpunkten in einem weiteren Auswerteschritt betrachtet werden. Die Dreiecksvermaschung ist in der Abbildung 6.5 dargestellt.

Jede Linie wird durch zwei Endpunkte begrenzt. Aufgrund der vorhergehenden Clusterung der Objektpunkte hat jeder dieser Endpunkte gewisse Zugehörigkeitswerte zu den entsprechenden Clustern. Aus den Zugehörigkeitswerten der Endpunkte werden jetzt durch Bildung des arithmetischen Mittels Zugehörigkeitswerte für jede Linie berechnet. Haben beide Endpunkte ihre Hauptzugehörigkeitswerte zu dem gleichen Cluster, bekommt die Linie auch zu dem entsprechenden Cluster einen relativ hohen Zugehörigkeitswert. Haben beide Endpunkte ihre Hauptzugehörigkeitswerte zu verschiedenen Clustern, bekommt die Linie einen relativ niedrigen maximalen Zugehörigkeitswert. Das ist ein Zeichen dafür, dass diese Linie eine Blockgrenze schneidet. Hier liegt der Vorteil des Fuzzy-Ansatzes, weil die unscharfe Zuordnung eines Objektpunktes zu einem Cluster aufgrund der Zugehörigkeitswerte mit in den zweiten Auswerteschritt übernommen werden kann, ohne dass vorher durch eine zwangsweise scharfe Zuordnung bereits eine Vorentscheidung getroffen wird. Man denke hier an die Punkte, die dicht an einer scharfen Klassengrenze liegen.

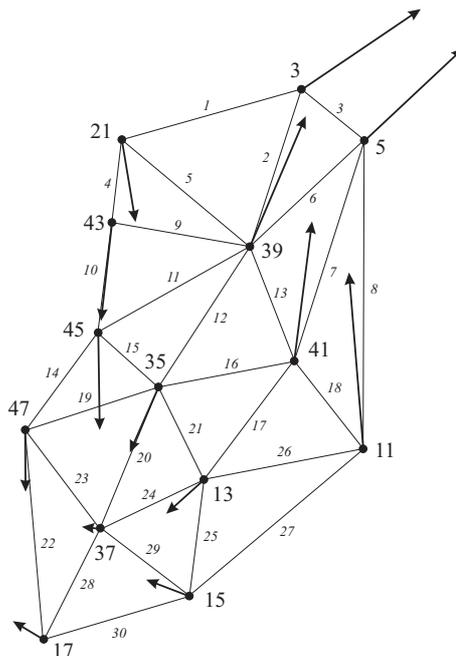


Abbildung 6.5: Testnetz Delft: Objektpunkte mit ihren Verschiebungsvektoren zwischen den Epochen 1 und 3b und einer Dreiecksvermaschung. Die kleinen Ziffern geben die Liniennummern an.

In der Tabelle 6.3 sind die Linien mit ihren beiden Endpunkten und den mittleren Zugehörigkeitswerten aus den Fuzzy-c-Means Clusterungen der Objektpunkte angegeben. In der letzten Spalte steht als Gütekriterium g der Quotient aus dem größten und dem zweitgrößten Zugehörigkeitswert (siehe Formel (4.21)). Wenn der Wert von g groß ist, deutet das auf eine gute Clusterzuordnung, wenn er klein ist, auf eine unsichere Clusterzuordnung hin. Der Wert für g sollte also möglichst groß sein. Die Ableitung einer scharfen Entscheidung (Blockgrenze ja/nein) aus den Zugehörigkeitswerten, die sogenannte „Defuzzifizierung“ erfolgt hier durch eine Schwellwertoperation. Als kleinster Wert für eine sichere Clusterzuordnung wurde der Wert $g > 2$ angenommen. Das bedeutet, dass der größte Zugehörigkeitswert mehr als doppelt so groß ist wie der zweitgrößte. In der Tabelle 6.3 sind die sicher zugeordneten Linien und ihre maximalen Zugehörigkeitswerte in fetter Schrift hervorgehoben. Bei allen Linien, die nicht als sicher zugeordnet gekennzeichnet sind, wird angenommen, dass sie von einer Blockgrenze geschnitten werden. In der Abbildung 6.6 sind in der linken Abbildung die als sicher zugeordneten Linien ebenfalls fett hervorgehoben. Die dünn gezeichneten Linien sind die Linien, die von einer Blockgrenze geschnitten werden.

Es ergeben sich hier in diesem Beispiel drei Blöcke von konsistenten Punktbewegungen. Bei einer etwas höher angesetzten Grenze für das Gütekriterium g würde die Linie 21 ebenfalls als unsicher zugeordnet gelten und der Punkt 13 als Repräsentant eines eigenständigen 4. Blockes erscheinen.

Bei Übernahme der Zugehörigkeitswerte für die Objektpunkte nach den anderen Fuzzy-Clusterverfahren (Gustavson-Kessel und Gath-Geva) aus Tabelle 6.1 sowie dem Verfahren der Fuzzy Pattern Klassifikation aus Tabelle 6.2 ergeben sich für die Blockfindung identische Ergebnisse. Die Ergebnisse dieser Verfahren sind im Anhang aus den Tabellen A.1, A.2 und A.3 zu entnehmen. Wie bereits erwähnt, können sich aber durch geringfügige Änderungen bei den Näherungswerten für die Clusterzentren beim Gustavson-Kessel Algorithmus und insbesondere beim Gath-Geva Algorithmus, völlig andere Clusterlösungen für die Punkte ergeben. Durch die nahezu scharfe Zuordnung beim Gath-Geva Algorithmus resultieren dadurch auch völlig andere Linienzugehörigkeiten und damit andere Blockgrenzen.

Linie	von Punkt 1	nach Punkt 2	Cluster			g
			1	2	3	
1	3	21	0,48	0,51	0,02	1
2	3	39	0,98	0,02	0,00	49
3	3	5	0,96	0,03	0,02	32
4	21	43	0,01	0,97	0,02	49
5	21	39	0,51	0,49	0,01	1
6	5	39	0,99	0,01	0,01	99
7	5	41	0,98	0,02	0,01	65
8	5	11	0,96	0,03	0,01	32
9	39	43	0,51	0,49	0,01	1
10	43	45	0,01	0,97	0,02	49
11	39	45	0,51	0,49	0,01	1
12	35	39	0,51	0,48	0,02	1
13	39	41	1,00	0,01	0,00	199
14	45	47	0,02	0,95	0,03	32
15	35	45	0,01	0,96	0,03	32
16	35	41	0,50	0,48	0,02	1
17	13	41	0,51	0,29	0,20	2
18	11	41	0,97	0,03	0,01	39
19	35	47	0,02	0,95	0,04	37
20	35	37	0,01	0,49	0,50	1
21	13	35	0,02	0,76	0,22	3
22	17	47	0,01	0,49	0,51	1
23	37	47	0,01	0,49	0,50	1
24	13	37	0,02	0,30	0,68	2
25	13	15	0,02	0,29	0,69	2
26	11	13	0,49	0,30	0,21	2
27	11	15	0,48	0,03	0,50	1
28	17	37	0,01	0,03	0,97	39
29	15	37	0,01	0,03	0,97	39
30	15	17	0,00	0,02	0,98	49

Tabelle 6.3: Zugehörigkeitswerte der Linien zu den drei Clustern, die aus den mit dem Fuzzy-c-Means Algorithmus berechneten Zugehörigkeitswerten der Objektpunkte abgeleitet wurden. Die Zugehörigkeitswerte sind auf zwei Nachkommastellen und das Gütemaß g auf eine ganze Zahl gerundet.

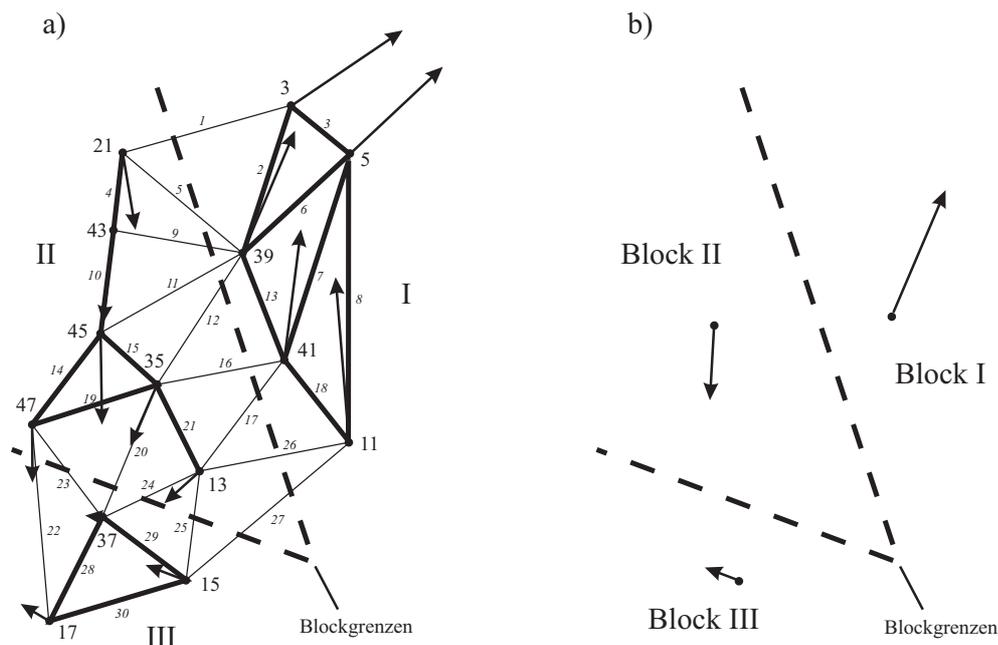


Abbildung 6.6: Testnetz Delft:

- Objektpunkte mit ihren Verschiebungsvektoren zwischen den Epochen 1 und 3b, der Dreiecksvermaschung und den eingezeichneten Blockgrenzen.
- Aus den Clusterzentren im Merkmalsraum abgeleitete Verschiebungsvektoren als Repräsentanten für die einzelnen Blöcke.

In diesem Beispiel bilden die Objektunkte, die ihren Hauptzugehörigkeitswert im gleichen Cluster haben, auch jeweils einen Block. Dadurch kann der jeweilige Clustermittelwert als Repräsentant für diesen Block angesehen werden. In der Abbildung 6.6 sind in der rechten Abbildung die drei Verschiebungsvektoren, die sich für die Clustermittelwerte im Merkmalsraum ergeben, jeweils als Blockrepräsentanten eingezeichnet.

Das hier behandelte Beispiel wurde auch in HABERLER (2005) präsentiert, sodass hier ein direkter Vergleich der Ergebnisse möglich ist. In dem Ansatz von HABERLER wurde die Blockgrenze zwischen den Blöcken I und II ebenfalls gefunden. Die Punkte 15, 17 und 37, die hier den Block III bilden, wurden zwar als nicht zum Block II gehörig detektiert, es konnte dort aber keine Aussage über einen gemeinsamen Block gemacht werden, da dazu in dem Ansatz von HABERLER mindestens vier Punkte benötigt werden. Die Punkte wurden dort als nicht zu einem Block gehörende Einzelpunkte bezeichnet. Bei dem hier vorgestellten Ansatz reicht minimal ein Punkt aus, um einen eigenständigen Block zu erkennen. Bei zwei nebeneinander liegenden Einzelpunkten kann gegenüber dem Ansatz von HABERLER eindeutig festgestellt werden, ob sie einem gemeinsamen Block zugewiesen werden können oder nicht.

In LEHMKÜHLER (1993) wurde dieses Beispiel ebenfalls benutzt. Allerdings weisen bei ihm die Verschiebungsvektoren der Punkte 15 und 17 aus nicht mehr nachvollziehbaren Gründen eine andere Richtung auf als in dem von HABERLER und dem in dieser Arbeit verwendeten Datensatz. Von daher wurde dort nur die Blockgrenze zwischen den Blöcken I und II gefunden.

Im Rahmen dieser Arbeit wurden auch Untersuchungen durchgeführt, in wie weit sich Strainparameter als Merkmale für diese Aufgabenstellung eignen. In diesen Untersuchungen konnte die Blockgrenze zwischen den Blöcken I und II ebenfalls gefunden werden. Eine Aufdeckung des Blockes III war allerdings nicht möglich.

Weitere Untersuchungen zu den Ergebnissen des Fuzzy Pattern Klassifikations Verfahrens

Wie bereits erwähnt, sorgt die geringe Anzahl von Objektpunkten und ihre Verteilung in den Clustern 1 und 2 dafür, dass, bei Anwendung des Verfahrens der Fuzzy Pattern Klassifikation, für diese Cluster schwer interpretierbare Formen entstehen (siehe Abbildung 6.4). Um die Auswirkung dieser stark unsymmetrischen Zugehörigkeitsfunktionen auf das Ergebnis der Blockfindung festzustellen, wurden in einem weiteren Untersuchungsschritt die Bereiche mit den nahezu scharfen Clustergrenzen manuell geändert. Für die Parameter β_{2g} , β_{3g} und β_{4g} , die sie stützenden eindimensionalen Zugehörigkeitsfunktionen, wurden die entsprechenden Werte der jeweils anderen Seite eingesetzt. Dadurch entstanden symmetrische Zugehörigkeitsfunktionen. In der Abbildung 6.7 sind die Linien gleicher Zugehörigkeiten für die drei Cluster dargestellt und der Tabelle 6.4 können die entsprechenden Zugehörigkeitswerte für die Objektpunkte entnommen werden.

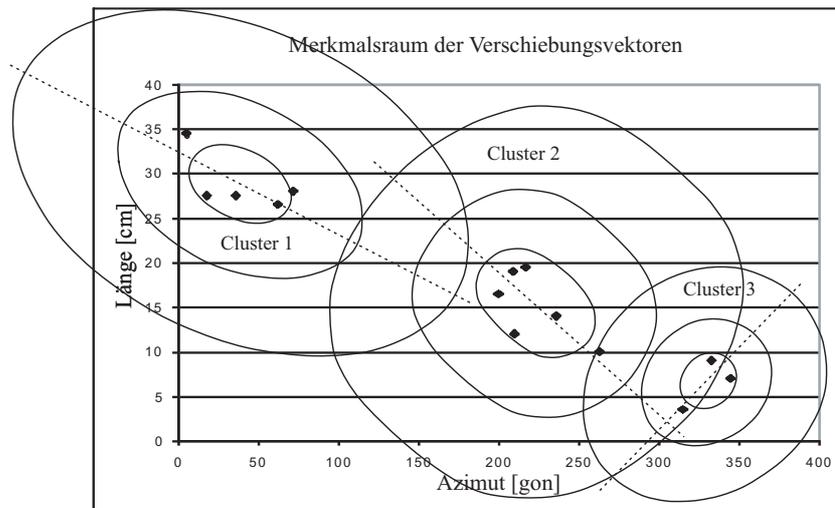


Abbildung 6.7: Mit Hilfe der Fuzzy Pattern Klassifikation und symmetrischen Zugehörigkeitsfunktionen berechnete Linien gleicher Zugehörigkeiten für die drei Cluster im Merkmalsraum. Die Isolinien repräsentieren von innen nach außen die Zugehörigkeitswerte $\mu = 0,9, 0,6$ und $0,3$. Die gestrichelten Linien sind die Hauptachsen der Cluster.

Wenn man die mit den unsymmetrischen Zugehörigkeitsfunktionen berechneten Zugehörigkeitswerte der Objektpunkte (Tabelle 6.2) mit den mit den symmetrischen Zugehörigkeitsfunktionen berechneten (Tabelle 6.4) vergleicht, ergibt sich Folgendes: Der jeweilige Hauptzugehörigkeitswert der Objektpunkte ändert sich kaum und der Hauptzugehörigkeitswert bleibt auch im gleichen Cluster. Stärkere Änderungen ergeben sich bei den Nebenzugehörigkeitswerten. Diese werden insbesondere in den Clustern 1 und 2 deutlich höher.

Die Änderung der Nebenzugehörigkeitswerte hat Auswirkungen auf die Berechnung der mittleren Zugehörigkeitswerte für die Linienverbindungen zwischen den Objektpunkten und auf den Wert des verwendeten Gütemaßes g als Quotient aus dem Haupt- und dem jeweils höchsten Nebenzugehörigkeitswert. Wenn man die beiden Tabellen A.3 und A.4 im Anhang miteinander vergleicht, zeigt sich, dass der Wert des Gütemaßes sich teilweise deutlich ändert. Bei Verwendung des Grenzwertes von $g = 2$ bleiben die als sicher zugeordneten Linien allerdings identisch und damit auch die gefundenen Blockgrenzen.

Fuzzy Pattern-Klassifikation symmetrische Zugehörigkeitsfunktionen				
Punkt	C_S	Cluster		
		1	2	3
3	1	0,88	0,15	0,03
5	1	0,93	0,14	0,03
11	1	0,80	0,06	0,02
13	2	0,14	0,83	0,31
15	3	0,09	0,34	0,92
17	3	0,08	0,26	0,92
21	2	0,24	0,96	0,10
35	2	0,17	0,99	0,17
37	3	0,09	0,33	0,85
39	1	0,99	0,10	0,02
41	1	0,92	0,09	0,02
43	2	0,21	0,95	0,10
45	2	0,23	0,96	0,10
47	2	0,20	0,93	0,13

Tabelle 6.4: Mit Hilfe der Fuzzy Pattern Klassifikation und symmetrischen Zugehörigkeitsfunktionen berechnete Zugehörigkeitswerte der Objektpunkte zu den drei Clustern. Der jeweilige Hauptzugehörigkeitswert eines Punktes ist fett hervorgehoben. In der Spalte C_S ist das Cluster angegeben, dem der Punkt mit Hilfe des Complete Linkage Verfahrens scharf zugewiesen wurde.

6.2.2 Isländisches Grundlagennetz ISNET

Bei diesem zweiten Beispiel zur Auffindung von Punktgruppen mit gleichförmigem Bewegungsverhalten handelt es sich ebenfalls um Daten eines Zweiepochevergleiches. Diese sind in VALSSON U.A. (2007) veröffentlicht worden. Die beiden Epochen bestehen jeweils aus Neumessungen des isländischen Grundlagennetzes, welche in den Jahren 1993 (ISNET93) und 2004 (ISNET2004) in groß angelegten GPS-Kampagnen durchgeführt wurden.

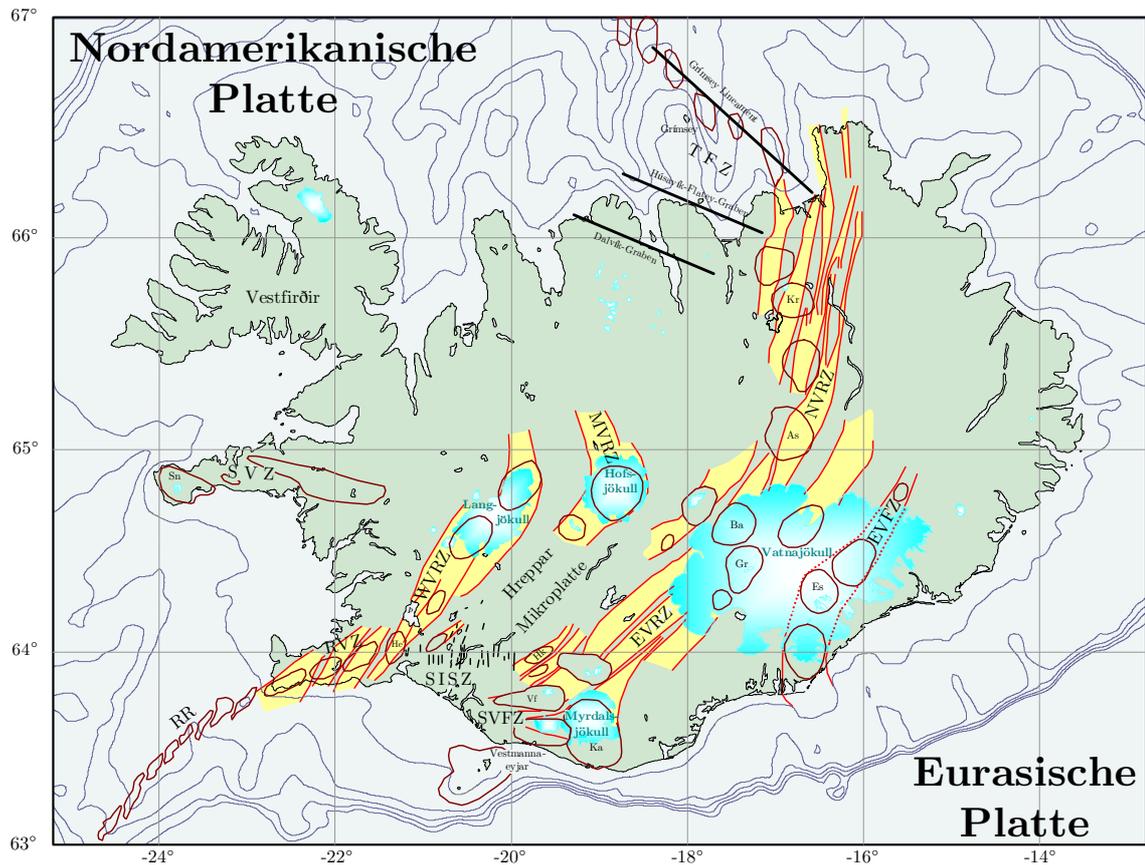


Abbildung 6.8: Die Tektonik Islands (aus HEINERT, 2008; basierend auf EINARSSON und SÆMUNDSSON, 1987).

Die Insel Island liegt auf dem mittelatlantischen Rücken, der in diesem Bereich die Plattengrenze zwischen der eurasischen und nordamerikanischen Kontinentalplatte bildet (Abbildung 6.8). Die Plattengrenze der auseinanderdriftenden Platten (der sogenannte passive oder divergente Plattenrand) geht mitten durch Island. Solch ein divergenter Plattenrand kann weltweit nur in Island an Land untersucht werden (PERLT, 2006). Am Institut für Geodäsie und Photogrammetrie (früher Institut für Vermessungskunde) der TU Braunschweig besteht seit 1964 (GERKE, 1966) eine längere Tradition an den geometrischen Untersuchungen der Plattenbewegungen in Island teilzunehmen. Sie wird durch immer wiederkehrende Beteiligungen an Messkampagnen und umfangreichen Auswertungen bis heute fortgesetzt und ist an einer Vielzahl von Veröffentlichungen abzulesen. Neuere Untersuchungen finden sich zum Beispiel in PERLT (2006) und in HEINERT (2008).

Die ursprünglichen Daten der ISNET Kampagnen wurden im ITRF (International Terrestrial Reference Frame) berechnet und über Stationen des IGS (International GPS Service) in Nordamerika und Europa gelagert. Die in diesem Beispiel benutzten Verschiebungsvektoren sind aus einer 3-Parameter Helmerttransformation der ISNET2004- auf die ISNET93-Kampagne entstanden³. Die Abbildung 6.9

³ Der ursprüngliche Epochenvergleich ist von einer globalen Bewegungskomponente von ca. 20 cm in nördliche Richtung überlagert. Die Transformation der Epochen aufeinander wäre theoretisch für die hier zu behandelnde Aufgabenstellung nicht nötig gewesen. Die aus der Transformation entstandenen Verschiebungsvektoren bilden im Merkmalsraum allerdings deutlich getrenntere Cluster.

zeigt diese Verschiebungsvektoren. Zu erkennen sind im Wesentlichen zwei große Blöcke. Einer im Osten mit einem Bewegungsverhalten in östliche Richtung und einer im Westen mit einem Bewegungsverhalten in westliche Richtung. Dazwischen gibt es einige Punkte mit relativ kurzen Bewegungsvektoren, und besonders auffällig sind im Südwesten drei Punkte (Punktnummern LM0305, OS7220, OS7225) mit einem Bewegungsverhalten in südliche Richtung.

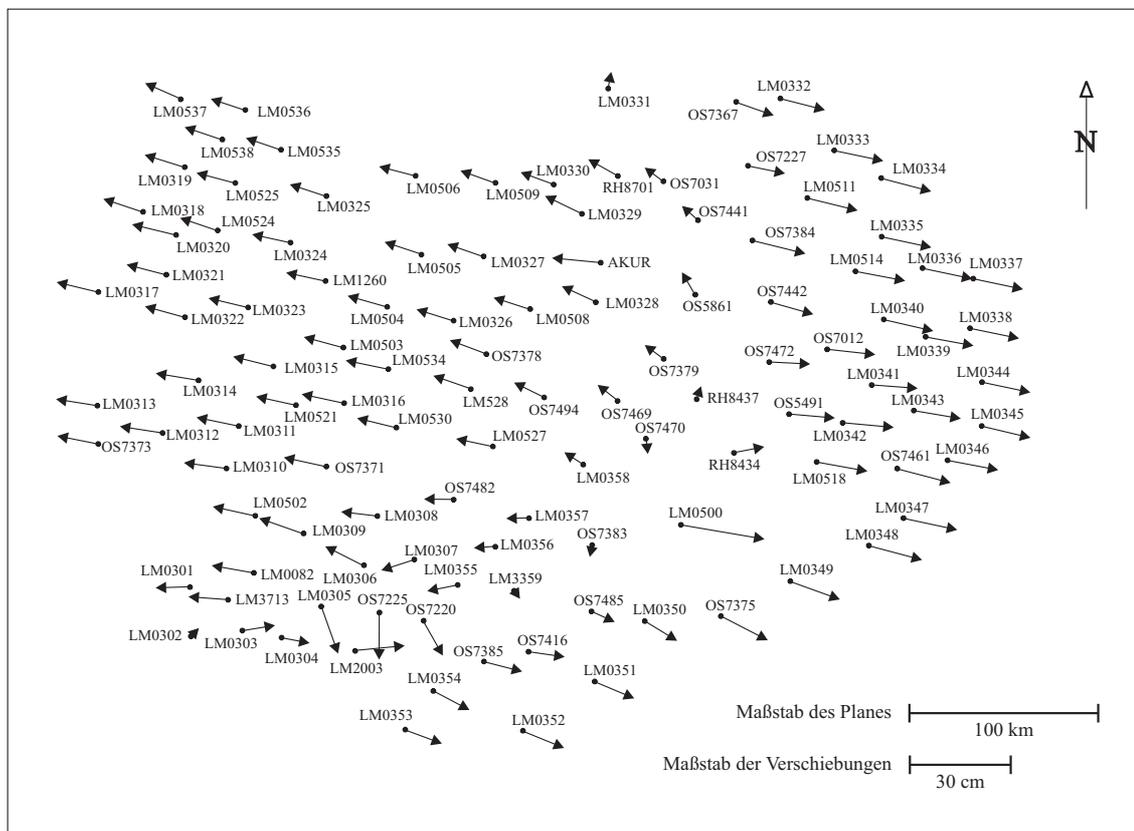


Abbildung 6.9: Die Verschiebungsvektoren des Beispiels ISNET.

Die wesentlichen Unterschiede dieses Beispiels zu dem Beispiel des Testnetzes Delft liegen darin, dass die Anzahl der Objektpunkte deutlich höher ist und darin, dass in diesem Beispiel die Verschiebungsvektoren im Merkmalsraum (Abbildung 6.10) Cluster unterschiedlicher Größe und Dichte bilden.

Das beim Beispiel Testnetz Delft aufgetretene Problem, dass die Unstetigkeitsstelle bei 0 gon im Azimut der Verschiebungsvektoren genau durch ein Cluster verlief, stellte sich in diesem Beispiel nicht. Daher wurde das Netz vor der Abbildung der Verschiebungsvektoren in den Merkmalsraum nicht gedreht.

Wie bereits im Abschnitt 4.3 erwähnt, ist der Fuzzy-c-Means Algorithmus bei der hier vorliegenden Datenstruktur (Cluster unterschiedlicher Dichte und Größe) nicht geeignet. Das Gleiche gilt für den Algorithmus nach Gustafson-Kessel. Bei dem Algorithmus nach Gath-Geva hatte sich bei dem Beispiel Delft gezeigt, dass er sehr empfindlich auf kleine Änderungen von Näherungswerten für die Clusterzentren reagiert und dadurch sehr häufig Clusterlösungen entstehen, die nicht mehr fachkundig interpretiert werden können. Das zeigte sich in diesem Beispiel ebenfalls und aus dem Grund wurde er auch nicht weiter in Betracht gezogen. Daher wurde hier nur mit dem Verfahren der Fuzzy Pattern Klassifikation gearbeitet.

Die grundsätzliche Vorgehensweise bei der Bearbeitung dieses Beispiels entspricht der des Beispiels Testnetz Delft. Auf die Unterschiede wird an den entsprechenden Stellen besonders hingewiesen.

Für die Berechnungen in diesem Beispiel wurde die Länge und das Azimut der Verschiebungsvektoren im Merkmalsraum ebenfalls normiert. Für die Länge der Verschiebungsvektoren wurde das Intervall

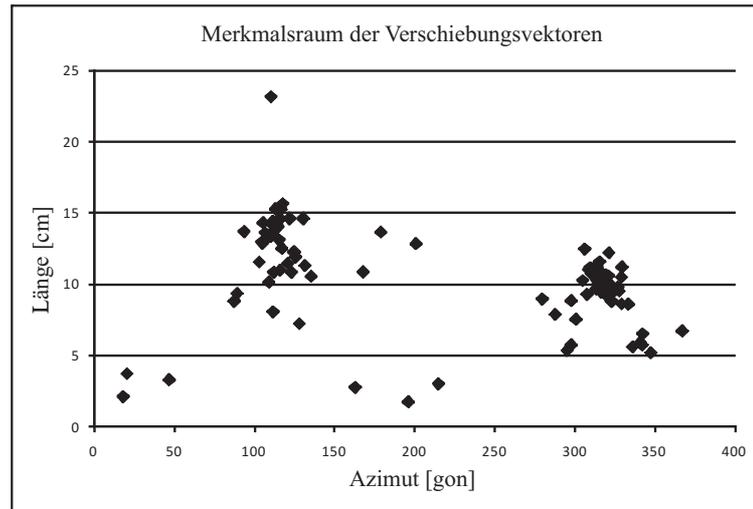


Abbildung 6.10: Die Verschiebungsvektoren des Beispiels ISNET im Merkmalsraum.

$[0, 23]^4$ cm auf das Intervall $[0, 0,5]$ und für das Azimut das Intervall $[0, 400]$ gon auf das Intervall $[0, 1]$ normiert.

Das verwendete Verfahren der Fuzzy Pattern Klassifikation benötigt zunächst scharfe Klassenzuordnungen der Objektpunkte im Merkmalsraum. Diese wurden wie bei dem Beispiel Delft nach dem Complete Linkage Verfahren bestimmt. Das Ergebnis des Average Linkage Verfahrens unterscheidet sich nur geringfügig in der Zuordnung eines Objektpunktes. Auf den Unterschied wird später noch einmal eingegangen. Alle anderen im Abschnitt 2.4.2 vorgestellten hierarchischen Verfahren ergeben keine Lösungen, die eine sinnvolle fachkundige Interpretation zulassen. Für die dort vorgestellten partitionierenden Verfahren gelten im Wesentlichen ebenfalls die für den Fuzzy-c-Means erwähnten Probleme bei Clustern unterschiedlicher Dichte und Größe.

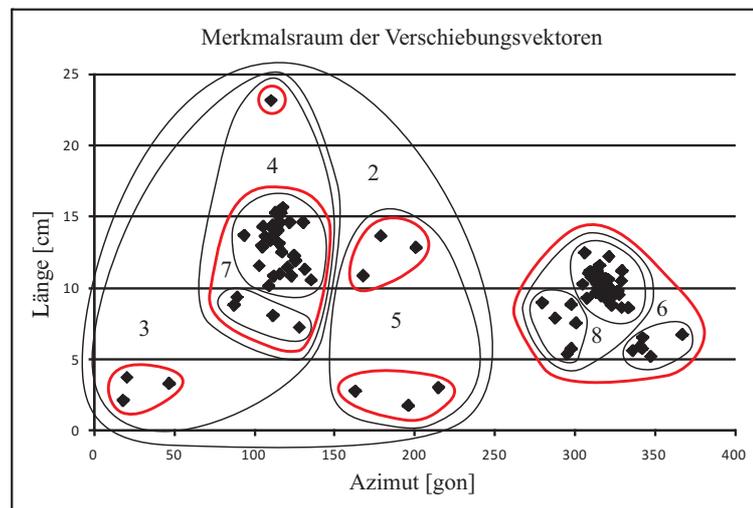


Abbildung 6.11: Die Verschiebungsvektoren des Beispiels ISNET im Merkmalsraum und Schleifen, die die Hierarchie der Fusionierung beim Complete Linkage Verfahren darstellen.

Die Abbildung 6.11 zeigt den Ablauf der Fusionierung des Complete Linkage Verfahrens ausgehend von zunächst neun Clustern bis hin zu zwei Clustern. Die Zahlen innerhalb einer Schleife geben an, auf welcher Hierarchiestufe die vorher einzelnen Cluster fusioniert werden. Die rot gezeichneten Schleifen kennzeichnen die sechs Cluster, die bei einer „optimalen“ Clusteranzahl von sechs (siehe dazu den folgenden Abschnitt) gebildet wurden.

⁴ Die 23 cm repräsentieren die Länge des längsten Verschiebungsvektors.

In der Abbildung 6.12 ist ein Ausschnitt aus dem Dendrogramm des Complete Linkage Verfahrens dargestellt. Hier sind über den Distanzen im normierten Merkmalsraum die Fusionierungsschritte von zunächst neun Clustern bis hin zu einem Cluster abgebildet.

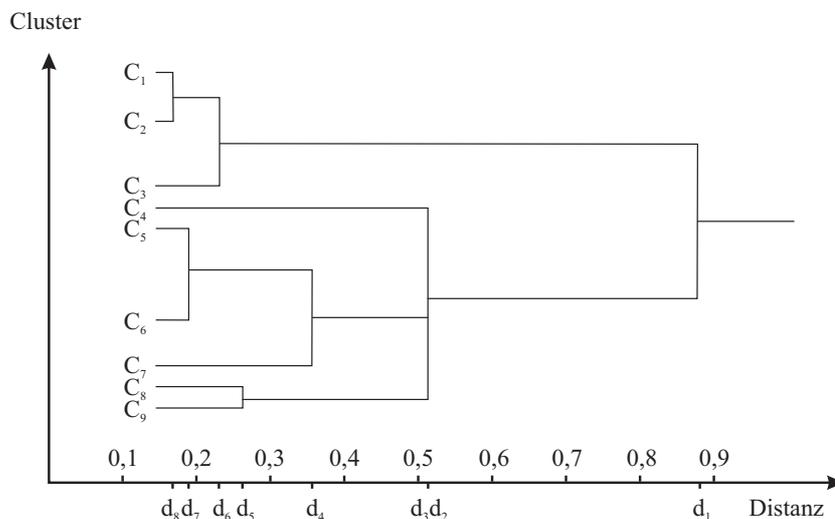


Abbildung 6.12: Ausschnitt aus dem Dendrogramm der Complete Linkage Clusterung.

Bestimmung der „optimalen“ Clusteranzahl

Zur Bestimmung der „optimalen“ Klassenanzahl wurde ein Kriterium gesucht, welches auch mit den hier vorliegenden Strukturen sinnvolle Ergebnisse liefert. Die im Abschnitt 4.6 beschriebenen Gütekriterien stehen in engem Zusammenhang mit dem Fuzzy-c-Means Algorithmus, bzw. beziehen sich auf probabilistische Clusterverfahren. Theoretisch können sie auch bei den hier vorhandenen possibilistischen Zugehörigkeitswerten verwendet werden (teilweise nach Normierung der Höhe der Zugehörigkeitsfunktion auf einen Wert ≤ 1). Bei der Anwendung stellte sich aber heraus, dass die im Abschnitt 4.6 beschriebenen Kriterien zur Ableitung einer „optimalen“ Clusteranzahl (... die Stelle, an der der Wert des Gütemaßes über den fallenden Trend steigt ...) hier nicht zu gebrauchen sind. Durch die Verwendung von possibilistischen Zugehörigkeitswerten ergeben sich für die Gütekriterien völlig andere Monotonieeigenschaften. Die Summe der Zugehörigkeitswerte eines Objektes über alle Cluster ist hierbei nicht konstant, sondern je nach Anzahl der Cluster unterschiedlich groß. Dadurch stellt sich zum Beispiel beim Partitionskoeffizienten nicht zwingend ein fallender Trend bei zunehmender Clusteranzahl ein.

Da beim Verfahren der Fuzzy Pattern Klassifikation zunächst eine scharfe Klassenzuordnung der Objekte benötigt wird, wird in diesem Beispiel bei der Bestimmung der „optimalen“ Klassenanzahl auch auf die Kriterien der scharfen Clusterung übergegangen. Im Abschnitt 2.4.3 sind hier beispielhaft das Varianzkriterium (Formel (2.12)) und das Dendrogramm des entsprechenden hierarchischen Verfahrens genannt worden.

Wie bereits im Abschnitt 2.4.3 erwähnt, deuten in einem Dendrogramm größere Distanzintervalle, in denen keine Fusionierung erfolgt, auf eine optimale Clusteranzahl hin. Das ist in dem hier vorliegenden Fall (Abbildung 6.12) zunächst einmal eine Lösung mit zwei Clustern. Das wären die in der Abbildung 6.9 gut zu erkennenden zwei großen Blöcke von Punkten, von denen sich der eine in Richtung Osten und der andere in Richtung Westen bewegt. Wenn man jedoch noch eine feinere Struktur sucht, ergeben sich aus dem Dendrogramm Lösungen mit vier bzw. fünf Clustern.

Bei der hier vorliegenden Problemstellung, dem Auffinden von Punktgruppen mit einem gleichförmigen Bewegungsverhalten, werden möglichst kompakte Cluster gesucht. Das heißt, dass die Cluster möglichst klein sein sollen. Die Abstände zwischen den Clustern sind eher von sekundärem Interesse.

Aus dem Dendrogramm des Complete Linkage Verfahrens geht bei jedem Fusionierungsschritt aber die maximale Distanz zwischen den entferntesten Objekten der fusionierten Cluster hervor und beim Average Linkage Verfahren ist es die mittlere Distanz zwischen den Objekten der fusionierten Cluster. Es sind aber keine Informationen über die Distanzen innerhalb der Cluster abzulesen.

$$g_{Var} = \sum_{j=1}^c \sum_{i=1}^{n_j} d^2(\mathbf{x}_i, \boldsymbol{\nu}_j) \quad (2.12)$$

$$g_{mqd} = \frac{1}{c} \sum_{j=1}^c \left(\frac{\sum_{i=1}^{n_j} d^2(\mathbf{x}_i, \boldsymbol{\nu}_j)}{n_j} \right) \quad (6.1)$$

- g_{Var} : Varianzkriterium
- g_{mqd} : Kriterium der mittleren quadrierten Distanzen
- \mathbf{x}_i : i -tes Objekt
- $\boldsymbol{\nu}_j$: Clusterzentrum des j -ten Clusters
- n_j : Anzahl der Objekte im j -ten Cluster
- d : Distanz
- c : Anzahl der Cluster

Ein Kriterium welches Informationen über die Distanzen innerhalb der Cluster liefert, ist das Varianzkriterium g_{Var} (Gleichung (2.12)). Es stellt die Summe der quadrierten euklidischen Distanzen (die Abweichungsquadratsumme) der Objekte von den jeweiligen Clustermittelwerten dar. Bei seiner Berechnung spielt allerdings die Anzahl der Objekte innerhalb eines Clusters keine Rolle. Wenn in einem Fusionierungsschritt zwei Cluster zusammengefasst werden, bei denen das eine Cluster deutlich mehr Objekte enthält als das andere, macht sich dieser Fusionierungsschritt in der Veränderung des Varianzkriteriums kaum bemerkbar. Ein gesuchter „Ellbogenknick“ in dem Diagramm, in dem das Varianzkriterium über der Clusteranzahl aufgetragen ist (siehe Abbildung 2.17), ist in dem Fall unter Umständen schwer zu erkennen oder zeigt sich an anderen Stellen. Von daher wurde von mir dazu übergegangen das Varianzkriterium dahingehend zu modifizieren, dass für jedes Cluster das arithmetische Mittel der quadrierten euklidischen Distanzen gebildet wird und diese noch einmal über alle Cluster gemittelt werden (Formel (6.1)). Dieses von mir genannte Kriterium der mittleren quadrierten Distanzen ist unabhängig von der Anzahl der Objekte in den einzelnen Clustern. Das Ergebnis des hier behandelten Beispiels darf nicht davon abhängig sein, wieviele Objektpunkte ein Block von Punkten mit konsistentem Bewegungsverhalten aufweist.

Von einer Einclusterlösung bis hin zu neun Clustern, die mit dem Complete Linkage Verfahren bestimmt wurden, erfolgte für jede mögliche Clusteranzahl die Berechnung der beiden vorgestellten Gütekriterien. In der Abbildung 6.13 sind die Ergebnisse dargestellt. Ein „Ellbogenknick“ zeigt sich bei beiden Kriterien zunächst an der Stelle von zwei Clustern. Das ergibt wieder die zwei großen Blöcke, die auch aus dem Dendrogramm abzulesen waren. In den jeweils rechten Darstellungen wurden die Kriteriumsachsen in einem anderen Maßstab dargestellt, dadurch können noch feinere Strukturen entdeckt werden. Hier ergeben sich „Ellbogenknick“ an unterschiedlichen Stellen. Das Varianzkriterium weist auf eine Vierclusterlösung hin, während beim Kriterium der mittleren quadrierten Distanzen der größte Knick bei sechs Clustern vorliegt. Diese Diskrepanz liegt darin begründet, dass bei der Berechnung des Varianzkriteriums nicht berücksichtigt wird, wie viele Objekte sich innerhalb eines Clusters befinden. Cluster mit vielen Objektpunkten tragen stärker zur Gesamtvarianz bei als Cluster mit wenigen Objektpunkten.

Für die weitere Bearbeitung wurde die Lösung von sechs Clustern zu Grunde gelegt. In der Abbildung 6.11 sind die Cluster durch die roten Linien bereits markiert worden. Diese Cluster sind sinnvoll interpretierbar und werden wie folgt bezeichnet:

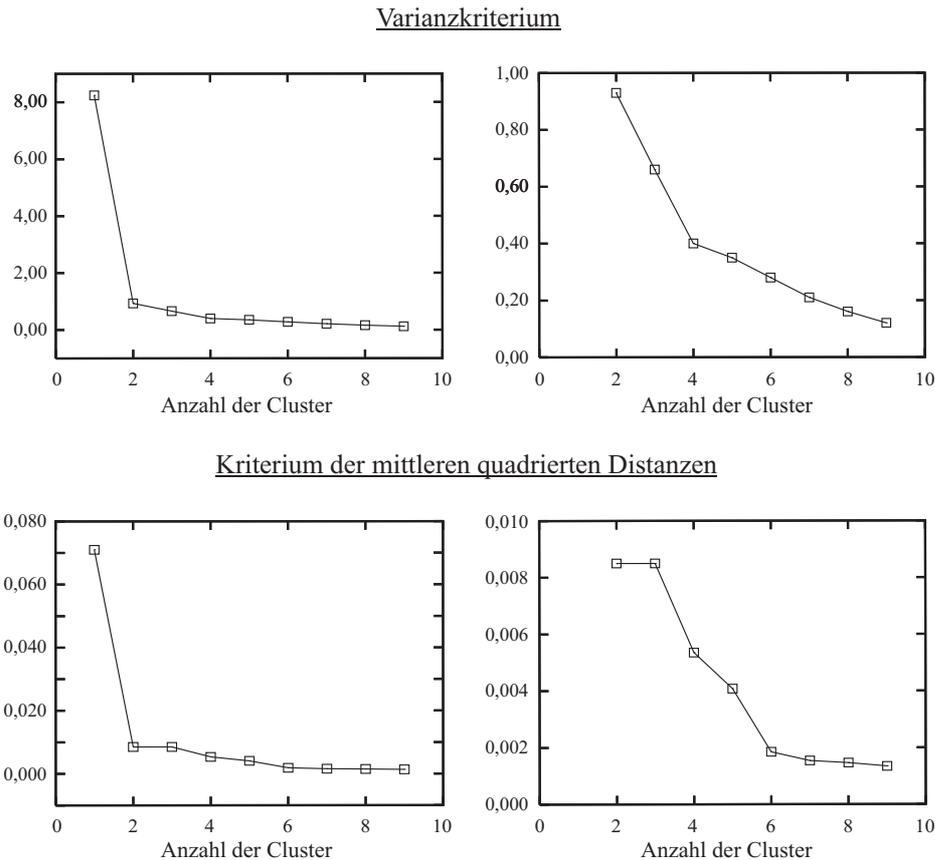


Abbildung 6.13: Oben das Varianzkriterium und unten das Kriterium der mittleren quadrierten Distanzen auf die Clusterergebnisse angewendet, die sich mit Hilfe des Complete Linkage Verfahrens ergaben. In den jeweils rechten Abbildungen sind die Werte für 2 bis 9 Cluster vergrößert dargestellt.

- Cluster 1: Kleine Verschiebung in nordnordöstliche Richtung.
- Cluster 2: Mittlere Verschiebung in östliche Richtung.
- Cluster 3: Große Verschiebung in östliche Richtung.
- Cluster 4: Mittlere Verschiebung in südliche Richtung.
- Cluster 5: Kleine Verschiebung in südliche Richtung.
- Cluster 6: Mittlere Verschiebung in westliche Richtung.

Unschärfe Beschreibung der Cluster

Aufbauend auf der vorhergehenden scharfen Clusterung und den elementaren Unschärfen der Merkmale erfolgt beim Verfahren der Fuzzy Pattern Klassifikation die Bildung der unscharfen Klassenbeschreibungen. Als elementare Unschärfen der Merkmale wurden für die Länge des Vektors 16 mm und für das Azimut 8,5 gon angesetzt⁵. Der asymptotische Grenzwert von β_{1g} (die Höhe der Zugehörigkeitsfunktion) wurde auf 10 gesetzt.

Da die Anzahl der Objekte pro Cluster sehr unterschiedlich ist, ergeben sich für die Höhen der einzelnen Zugehörigkeitsfunktionen nach Formel (5.13) auch unterschiedliche Werte. Um zu verhindern, dass in

⁵ Diese Werte entsprechen ungefähr den Standardabweichungen eines Vektors mittlerer Länge. Diese wurden durch Fehlerfortpflanzung aus den Standardabweichungen der Punkte ermittelt, wie sie sich aus den freien Ausgleichungen der einzelnen Epochen ergaben.

einem weiteren Auswerteschritt die Anzahl der Objekte pro Cluster eine Auswirkung auf die Findung der Blockgrenzen hat, wurden die Höhen der Zugehörigkeitsfunktionen auf 1,0 normiert. Die Abbildung 6.14 zeigt die Linien gleicher Zugehörigkeiten für die sechs Cluster und in der Tabelle 6.5 sind die Zugehörigkeitswerte der Objektpunkte zu den einzelnen Clustern angegeben.

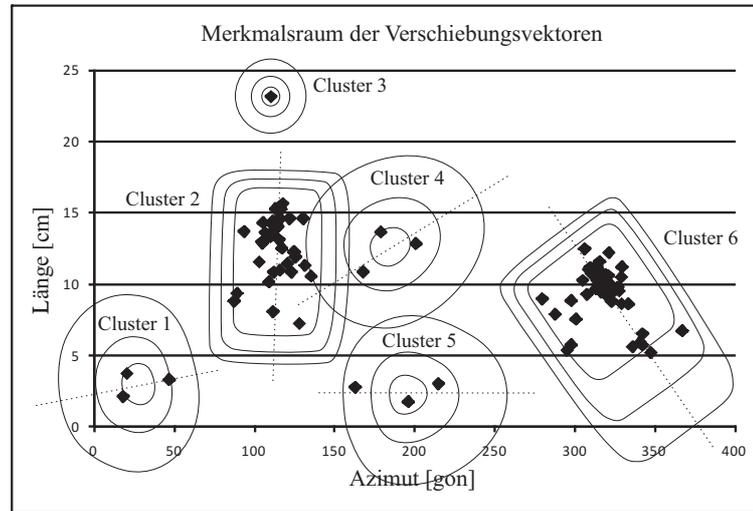


Abbildung 6.14: Linien gleicher Zugehörigkeiten für die sechs Cluster im Merkmalsraum. Die Isolinien repräsentieren von innen nach außen die Zugehörigkeitswerte $\mu = 0,9, 0,6$ und $0,3$. Die gestrichelten Linien sind die Hauptachsen der Cluster.

Zugehörigkeitswerte der Punkte							
Punkt	C_S	Cluster					
		1	2	3	4	5	6
LM0082	6	0,01	0,00	0,00	0,08	0,06	1,00
LM0301	6	0,01	0,00	0,00	0,08	0,09	1,00
LM0302	1	0,61	0,02	0,01	0,04	0,03	0,00
LM0303	2	0,09	0,95	0,01	0,10	0,05	0,00
LM0304	2	0,06	1,00	0,01	0,14	0,07	0,00
LM0305	4	0,02	0,03	0,02	0,90	0,08	0,00
LM0306	6	0,01	0,00	0,00	0,06	0,05	1,00
LM0307	6	0,01	0,00	0,00	0,10	0,11	0,98
LM0308	6	0,01	0,00	0,00	0,07	0,08	1,00
LM0309	6	0,01	0,00	0,00	0,07	0,05	1,00
LM0310	6	0,01	0,00	0,00	0,08	0,07	1,00
LM0311	6	0,01	0,00	0,00	0,07	0,06	1,00
LM0312	6	0,01	0,00	0,00	0,08	0,06	1,00
LM0313	6	0,01	0,00	0,00	0,08	0,07	1,00
LM0314	6	0,01	0,00	0,00	0,08	0,06	1,00
LM0315	6	0,01	0,00	0,00	0,07	0,06	1,00
LM0316	6	0,01	0,00	0,00	0,07	0,06	1,00
LM0317	6	0,01	0,00	0,00	0,07	0,06	1,00
LM0318	6	0,01	0,00	0,00	0,06	0,06	1,00
LM0319	6	0,01	0,00	0,00	0,06	0,06	1,00
LM0320	6	0,01	0,00	0,00	0,07	0,06	1,00

(Fortsetzung nächste Seite)

(Fortsetzung)

Punkt	C_S	Cluster					
		1	2	3	4	5	6
LM0321	6	0,01	0,00	0,00	0,07	0,06	1,00
LM0322	6	0,01	0,00	0,00	0,07	0,06	1,00
LM0323	6	0,01	0,00	0,00	0,07	0,07	1,00
LM0324	6	0,01	0,00	0,00	0,07	0,07	1,00
LM0325	6	0,01	0,00	0,00	0,06	0,07	1,00
LM0326	6	0,01	0,00	0,00	0,06	0,07	1,00
LM0327	6	0,01	0,00	0,00	0,06	0,07	1,00
LM0328	6	0,01	0,00	0,00	0,05	0,06	1,00
LM0329	6	0,01	0,00	0,00	0,06	0,06	1,00
LM0330	6	0,01	0,00	0,00	0,06	0,07	1,00
LM0331	1	0,87	0,00	0,00	0,03	0,02	0,00
LM0332	2	0,04	1,00	0,03	0,19	0,05	0,00
LM0333	2	0,04	1,00	0,03	0,17	0,04	0,00
LM0334	2	0,04	1,00	0,04	0,19	0,04	0,00
LM0335	2	0,04	1,00	0,03	0,18	0,04	0,00
LM0336	2	0,04	1,00	0,03	0,17	0,04	0,00
LM0337	2	0,04	1,00	0,03	0,17	0,04	0,00
LM0338	2	0,04	1,00	0,03	0,17	0,04	0,00
LM0339	2	0,04	1,00	0,03	0,17	0,05	0,00
LM0340	2	0,04	1,00	0,03	0,18	0,05	0,00
LM0341	2	0,05	1,00	0,03	0,15	0,04	0,00
LM0342	2	0,05	1,00	0,03	0,15	0,04	0,00
LM0343	2	0,05	1,00	0,03	0,17	0,05	0,00
LM0344	2	0,04	1,00	0,03	0,18	0,05	0,00
LM0345	2	0,04	1,00	0,03	0,18	0,05	0,00
LM0346	2	0,04	1,00	0,03	0,17	0,04	0,00
LM0347	2	0,04	1,00	0,04	0,17	0,04	0,00
LM0348	2	0,04	1,00	0,04	0,18	0,04	0,00
LM348A	2	0,04	1,00	0,03	0,21	0,05	0,00
LM0349	2	0,04	1,00	0,05	0,18	0,04	0,00
LM0350	2	0,04	0,95	0,02	0,29	0,09	0,00
LM0351	2	0,04	1,00	0,02	0,24	0,06	0,00
LM0352	2	0,04	1,00	0,02	0,24	0,06	0,00
LM0353	2	0,04	1,00	0,02	0,22	0,07	0,00
LM0354	2	0,04	0,99	0,02	0,28	0,07	0,00
LM0355	6	0,01	0,00	0,00	0,08	0,12	0,99
LM0356	6	0,01	0,00	0,00	0,06	0,13	0,90
LM0357	6	0,01	0,00	0,00	0,06	0,12	0,96
LM0358	6	0,01	0,00	0,00	0,04	0,07	0,93
LM0500	3	0,02	0,00	1,00	0,08	0,02	0,00
LM0502	6	0,01	0,00	0,00	0,07	0,06	1,00

(Fortsetzung nächste Seite)

(Fortsetzung)

Punkt	C_S	Cluster					
		1	2	3	4	5	6
LM0503	6	0,01	0,00	0,00	0,07	0,06	1,00
LM0504	6	0,01	0,00	0,00	0,07	0,06	1,00
LM0505	6	0,01	0,00	0,00	0,06	0,07	1,00
LM0506	6	0,01	0,00	0,00	0,06	0,07	1,00
LM0508	6	0,01	0,00	0,00	0,06	0,07	1,00
LM0509	6	0,01	0,00	0,00	0,06	0,07	1,00
LM0511	2	0,04	1,00	0,04	0,18	0,04	0,00
LM0514	2	0,04	1,00	0,03	0,17	0,04	0,00
LM0518	2	0,04	1,00	0,03	0,17	0,04	0,00
LM0521	6	0,01	0,00	0,00	0,07	0,07	1,00
LM0524	6	0,01	0,00	0,00	0,06	0,06	1,00
LM0525	6	0,01	0,00	0,00	0,06	0,07	1,00
LM0527	6	0,01	0,00	0,00	0,07	0,07	1,00
LM0528	6	0,01	0,00	0,00	0,06	0,06	1,00
LM0530	6	0,01	0,00	0,00	0,07	0,07	1,00
LM0534	6	0,01	0,00	0,00	0,07	0,06	1,00
LM0535	6	0,01	0,00	0,00	0,06	0,07	1,00
LM0536	6	0,01	0,00	0,00	0,06	0,07	1,00
LM0537	6	0,01	0,00	0,00	0,06	0,06	1,00
LM0538	6	0,01	0,00	0,00	0,06	0,06	1,00
LM1260	6	0,01	0,00	0,00	0,07	0,07	1,00
LM2003	2	0,06	0,98	0,03	0,12	0,04	0,00
LM3359	5	0,03	0,04	0,01	0,12	0,43	0,00
LM3713	6	0,01	0,00	0,00	0,08	0,07	1,00
OS5491	2	0,05	1,00	0,02	0,15	0,05	0,00
OS5861	6	0,00	0,00	0,00	0,03	0,05	0,81
OS7012	2	0,05	1,00	0,03	0,16	0,04	0,00
OS7031	6	0,01	0,00	0,00	0,04	0,07	0,91
OS7220	4	0,02	0,08	0,01	0,72	0,13	0,00
OS7225	4	0,01	0,00	0,01	0,82	0,10	0,00
OS7227	2	0,05	1,00	0,02	0,17	0,06	0,00
OS7367	2	0,04	1,00	0,02	0,21	0,06	0,00
OS7371	6	0,01	0,00	0,00	0,07	0,06	1,00
OS7373	6	0,01	0,00	0,00	0,07	0,07	1,00
OS7375	2	0,03	0,99	0,03	0,26	0,05	0,00
OS7378	6	0,01	0,00	0,00	0,06	0,06	1,00
OS7379	6	0,01	0,00	0,00	0,04	0,07	0,93
OS7383	5	0,02	0,00	0,00	0,11	0,75	0,00
OS7384	2	0,04	1,00	0,04	0,18	0,04	0,00
OS7385	2	0,05	1,00	0,02	0,19	0,06	0,00
OS7416	2	0,06	1,00	0,01	0,15	0,06	0,00

(Fortsetzung nächste Seite)

(Fortsetzung)

Punkt	C_S	Cluster					
		1	2	3	4	5	6
OS7441	6	0,00	0,00	0,00	0,04	0,06	0,85
OS7442	2	0,04	1,00	0,02	0,20	0,05	0,00
OS7461	2	0,04	1,00	0,04	0,18	0,04	0,00
OS7469	6	0,00	0,00	0,00	0,04	0,06	0,94
OS7470	5	0,02	0,00	0,00	0,09	0,95	0,00
OS7472	2	0,06	1,00	0,02	0,14	0,05	0,00
OS7482	6	0,01	0,00	0,00	0,07	0,10	1,00
OS7485	2	0,05	0,97	0,01	0,17	0,11	0,00
OS7494	6	0,01	0,00	0,00	0,05	0,07	1,00
RH8434	2	0,10	0,92	0,01	0,09	0,05	0,00
RH8437	1	0,93	0,00	0,00	0,03	0,02	0,00
RH8701	6	0,01	0,00	0,00	0,05	0,06	1,00
AKUR	6	0,01	0,00	0,00	0,09	0,06	0,97

Tabelle 6.5: Mit Hilfe der Fuzzy Pattern Klassifikation berechnete Zugehörigkeitswerte der Objektpunkte zu sechs Clustern. Der jeweilige Hauptzugehörigkeitswert eines Punktes ist fett hervorgehoben. In der Spalte C_S ist das Cluster angegeben, dem der Punkt mit Hilfe des Complete Linkage Verfahrens scharf zugewiesen wurde.

Berücksichtigung der Nachbarschaft

Die Berücksichtigung der Nachbarschaft erfolgt nach dem gleichen Verfahren wie beim Beispiel Testnetz Delft. Zunächst wurde eine Dreiecksvermaschung zwischen den Punkten berechnet und für die entstandenen Linienverbindungen wurden aus den Zugehörigkeitswerten der Endpunkte (Tabelle 6.5) durch Bildung des arithmetischen Mittels Zugehörigkeitswerte zu den einzelnen Clustern abgeleitet. Haben die beiden Endpunkte ihren Hauptzugehörigkeitswert im gleichen Cluster, bekommt die Linie einen relativ großen Hauptzugehörigkeitswert und die Nebenzugehörigkeitswerte sind deutlich kleiner. Haben dagegen die beiden Endpunkte ihren Hauptzugehörigkeitswert in unterschiedlichen Clustern, bekommt die Linie einen relativ niedrigen Hauptzugehörigkeitswert und der höchste Nebenzugehörigkeitswert ist relativ hoch. Als Gütekriterium g wird hier auch wieder der Quotient aus dem größten und dem zweitgrößten Zugehörigkeitswert (Formel (4.21)) benutzt. Wenn der Wert von g groß ist, deutet das auf eine gute Clusterzuordnung hin, wenn er klein ist, auf eine unsichere. Als Wert für eine sichere Clusterzuordnung wurde hier (wie im vorhergehenden Beispiel) ebenfalls ein Wert $g > 2$ angenommen.

Bei allen Linien, bei denen das Gütemaß $g \leq 2$ ist, wird angenommen, dass sie von einer Blockgrenze geschnitten werden. Aus Gründen der Übersichtlichkeit sind in der Tabelle 6.6 nur diese Linienverbindungen aufgeführt. Eine vollständige Tabelle mit allen Linienverbindungen (Tabelle A.5) befindet sich im Anhang. In der Abbildung 6.15 sind die als sicher zugeordneten Linien fett hervorgehoben. Die dünn gezeichneten Linien sind die Linien, die von einer Blockgrenze geschnitten werden. Die Blockgrenzen sind in der Abbildung manuell eingezeichnet worden. In einem automatisierten Prozess könnten sie als Mittelsenkrechte der unsicher zugeordneten Dreieckskanten gezeichnet werden, die anschließend

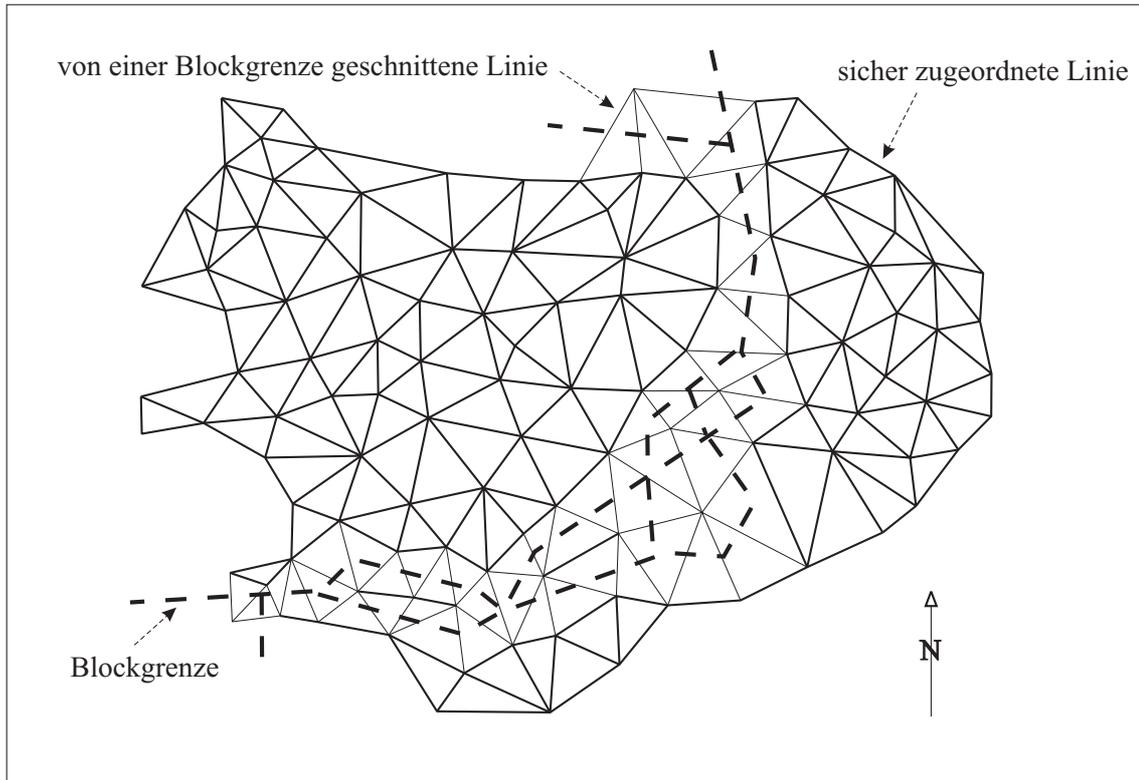


Abbildung 6.15: Die Dreiecksvermaschung des Beispiels ISNET. Die dünnen Linien sind die Linien, die von einer Blockgrenze geschnitten werden. Die dicken unterbrochenen Linien bilden die Blockgrenzen.

miteinander zum Schnitt gebracht werden. Der Ablauf ist ähnlich der in der Geoinformatik bekannten Erzeugung von Thiessen- oder Voronoi-Polygonen (z.B. DE LANGE, 2002).

Zugehörigkeitswerte der Linien, die von einer Blockgrenze geschnitten werden								
von Punkt 1	nach Punkt 2	Cluster						<i>g</i>
		1	2	3	4	5	6	
LM0082	LM0303	0,05	0,48	0,01	0,09	0,05	0,50	1
LM0082	LM0304	0,03	0,50	0,01	0,11	0,07	0,50	1
LM0082	LM0305	0,01	0,02	0,01	0,49	0,07	0,50	1
LM0301	LM0302	0,31	0,01	0,00	0,06	0,06	0,50	2
LM0302	LM0303	0,35	0,49	0,01	0,07	0,04	0,00	1
LM0302	LM3713	0,31	0,01	0,00	0,06	0,05	0,50	2
LM0303	LM3713	0,05	0,48	0,01	0,09	0,06	0,50	1
LM0304	LM0305	0,04	0,51	0,01	0,52	0,08	0,00	1
LM0305	LM0306	0,01	0,02	0,01	0,48	0,07	0,50	1
LM0305	LM0309	0,01	0,02	0,01	0,48	0,07	0,50	1
LM0305	LM2003	0,04	0,51	0,02	0,51	0,06	0,00	1
LM0306	OS7225	0,01	0,00	0,01	0,44	0,08	0,50	1
LM0307	OS7220	0,02	0,04	0,01	0,41	0,12	0,49	1

(Fortsetzung nächste Seite)

(Fortsetzung)

von Punkt 1	nach Punkt 2	Cluster						g
		1	2	3	4	5	6	
LM0307	OS7225	0,01	0,00	0,01	0,46	0,10	0,49	1
LM0330	LM0331	0,44	0,00	0,00	0,04	0,05	0,50	1
LM0331	OS7031	0,44	0,00	0,00	0,03	0,04	0,46	1
LM0331	OS7367	0,46	0,50	0,01	0,12	0,04	0,00	1
LM0331	RH8701	0,44	0,00	0,00	0,04	0,04	0,50	1
LM0349	LM0500	0,03	0,50	0,52	0,13	0,03	0,00	1
LM0350	LM0500	0,03	0,48	0,51	0,19	0,05	0,00	1
LM0350	OS7383	0,03	0,48	0,01	0,20	0,42	0,00	1
LM0354	OS7220	0,03	0,53	0,02	0,50	0,10	0,00	1
LM0355	LM3359	0,02	0,02	0,00	0,10	0,27	0,49	2
LM0355	OS7220	0,02	0,04	0,01	0,40	0,12	0,49	1
LM0355	OS7385	0,03	0,50	0,01	0,13	0,09	0,49	1
LM0356	LM3359	0,02	0,02	0,00	0,09	0,28	0,45	2
LM0357	LM3359	0,02	0,02	0,00	0,09	0,28	0,48	2
LM0357	OS7383	0,01	0,00	0,00	0,08	0,44	0,48	1
LM0358	LM0500	0,01	0,00	0,50	0,06	0,05	0,46	1
LM0358	OS7383	0,01	0,00	0,00	0,07	0,41	0,46	1
LM0358	OS7470	0,01	0,00	0,00	0,07	0,51	0,46	1
LM0500	OS7375	0,03	0,50	0,52	0,17	0,04	0,00	1
LM0500	OS7383	0,02	0,00	0,50	0,09	0,39	0,00	1
LM0500	OS7470	0,02	0,00	0,50	0,09	0,49	0,00	1
LM0500	RH8434	0,06	0,46	0,51	0,08	0,03	0,00	1
LM2003	OS7220	0,04	0,53	0,02	0,42	0,08	0,00	1
LM2003	OS7225	0,04	0,49	0,02	0,47	0,07	0,00	1
LM3359	OS7385	0,04	0,52	0,01	0,15	0,25	0,00	2
LM3359	OS7416	0,04	0,52	0,01	0,13	0,24	0,00	2
LM3359	OS7485	0,04	0,50	0,01	0,14	0,27	0,00	2
OS5491	RH8437	0,49	0,50	0,01	0,09	0,03	0,00	1
OS5861	OS7384	0,02	0,50	0,02	0,10	0,05	0,41	1
OS5861	OS7442	0,02	0,50	0,01	0,12	0,05	0,41	1
OS5861	OS7472	0,03	0,50	0,01	0,09	0,05	0,41	1
OS7031	OS7227	0,03	0,50	0,01	0,10	0,06	0,46	1
OS7031	OS7367	0,02	0,50	0,01	0,13	0,07	0,46	1
OS7220	OS7385	0,04	0,54	0,02	0,45	0,09	0,00	1
OS7227	OS7441	0,03	0,50	0,01	0,10	0,06	0,43	1
OS7379	OS7472	0,03	0,50	0,01	0,09	0,06	0,46	1
OS7379	RH8437	0,47	0,00	0,00	0,04	0,04	0,46	1
OS7383	OS7485	0,03	0,48	0,01	0,14	0,43	0,00	1
OS7384	OS7441	0,02	0,50	0,02	0,11	0,05	0,43	1

(Fortsetzung nächste Seite)

(Fortsetzung)

von Punkt 1	nach Punkt 2	Cluster						g
		1	2	3	4	5	6	
OS7469	OS7470	0,01	0,00	0,00	0,07	0,51	0,47	1
OS7469	RH8437	0,47	0,00	0,00	0,04	0,04	0,47	1
OS7470	RH8434	0,06	0,46	0,01	0,09	0,50	0,00	1
OS7470	RH8437	0,47	0,00	0,00	0,06	0,49	0,00	1
OS7472	RH8437	0,49	0,50	0,01	0,09	0,03	0,00	1
RH8434	RH8437	0,51	0,46	0,01	0,06	0,03	0,00	1

Tabelle 6.6: Linien, die durch eine Blockgrenze geschnitten werden. Angegeben sind die Zugehörigkeitswerte zu den sechs Clustern. Die Zugehörigkeitswerte sind auf zwei Nachkommastellen und das Gütemaß g auf eine ganze Zahl gerundet.

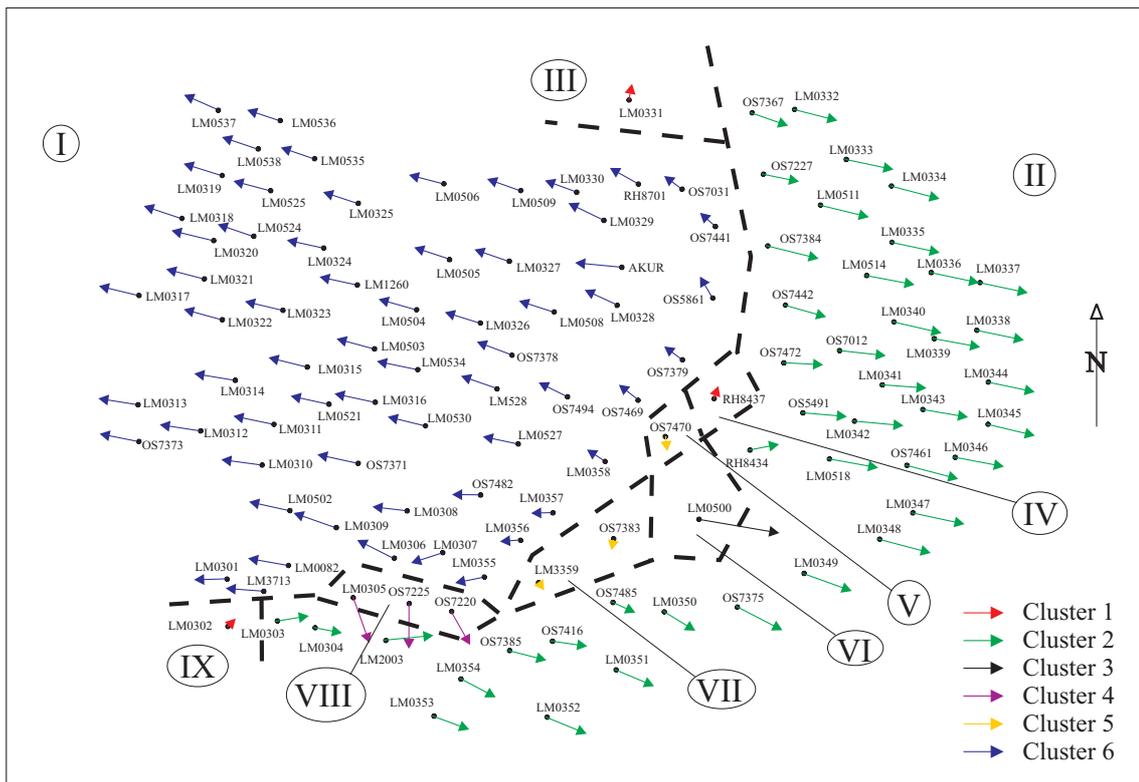


Abbildung 6.16: Die Verschiebungsvektoren des Beispiels ISNET mit den detektierten Blockgrenzen und den Blocknummern (I-IX). Verschiebungsvektoren, die zu dem gleichen Cluster gehören, sind in gleicher Farbe dargestellt.

In der Abbildung 6.16 sind die Verschiebungsvektoren, die zu einem gleichen Cluster gehören, in gleicher Farbe dargestellt. Während bei dem Beispiel Testnetz Delft die Berücksichtigung der Nachbarschaft über die Dreiecksvermaschung eigentlich nicht notwendig gewesen wäre, weil im Endergebnis jedes Cluster auch einen Block repräsentiert, sieht es hier anders aus. In diesem Beispiel liegen teilweise Punkte eines Clusters nicht in direkter Nachbarschaft, sondern zwischen ihnen befinden sich Punkte

aus anderen Clustern. Am deutlichsten wird das bei den Punkten des Clusters 1 (Verschiebungsvektoren in rot).

Durch die Berücksichtigung der Nachbarschaft sind insgesamt neun Blöcke detektiert worden. Sie sind in der Abbildung 6.16 näher bezeichnet. Es ergeben sich zwei große Blöcke, der Block I im Westen auf der nordamerikanischen Platte und der Block II im Osten auf der eurasischen Kontinentalplatte. Dazwischen, im Bereich der Plattengrenze, ergeben sich mehrere kleine Blöcke.

- **Block I** im westlichen Teil mit einer **mittleren** Verschiebung in **westliche** Richtung. Dieser Block besteht insgesamt aus 63 Punkten.
- **Block II** im östlichen Teil mit einer **mittleren** Verschiebung in **östliche** Richtung. Dieser Block besteht insgesamt aus 43 Punkten.
- **Block III** im Norden mit einer **kleinen** Verschiebung in **nordöstliche** Richtung. Dieser Block besteht nur aus dem Punkt LM0331.
- **Block IV** in der Mitte mit einer **kleinen** Verschiebung in **nordöstliche** Richtung. Dieser Block besteht nur aus dem Punkt RH8437.
- **Block V** in der Mitte mit einer **kleinen** Verschiebung in **südliche** Richtung. Dieser Block besteht nur aus dem Punkt OS7470.
- **Block VI** in der Mitte mit einer **großen** Verschiebung in **östliche** Richtung. Dieser Block besteht nur aus dem Punkt LM0500.
- **Block VII** in der Mitte mit einer **kleinen** Verschiebung in **südliche** Richtung. Dieser Block besteht aus den Punkten OS7383 und LM3359.
- **Block VIII** im Süden mit einer **mittleren** Verschiebung in **südliche** Richtung. Dieser Block besteht aus den Punkten LM0305, OS7220 und OS7225.
- **Block IX** im Süden mit einer **kleinen** Verschiebung in **nordöstliche** Richtung. Dieser Block besteht nur aus dem Punkt LM0302.

Bei dem hier vorgestellten Verfahren kann generell nicht unterschieden werden, ob eine Linienverbindung nur von einer oder von mehreren Blockgrenzen geschnitten wird. Bei einer anderen Dreiecksvermaschung (wenn die Diagonale im Viereck LM0358, OS7470, LM0500, OS7383 getauscht würde), würde die Diagonale von OS7470 zu OS7383 parallel zu den Blockgrenzen verlaufen und diese Linie nicht schneiden. Dadurch würden die beiden Blöcke V und VII einen gemeinsamen Block ergeben.

Der Block VI, der nur vom Punkt LM0500 repräsentiert wird, fällt durch einen Verschiebungsvektor auf, der ca. doppelt so lang ist, wie die benachbarten Punkte auf der eurasischen Platte. Er liegt in einem tektonisch und seismisch höchst aktiven Bereich. Hier liegt Islands aktivster Vulkan Grímsvötn, an dem es sowohl 1993 als auch 2004 zu einem Ausbruch kam (VILLEMIN U.A., 2009).

Weiter auffällig ist der Block VIII mit einer mittleren Verschiebung in südliche Richtung. Dieser Block liegt in der SISZ (Südisländische Seismische Zone), in der sich im Jahre 2000 zwei Erdbeben ereigneten (PERLT, 2006). In der SISZ befinden sich eine ganze Reihe paralleler, in Nord-Süd Richtung verlaufende Spalten, die sich gegeneinander verschieben. Die gemeinsame Bewegungskomponente der drei Punkte in südliche Richtung ist offenbar mehr oder weniger Zufall, da sich diese Punkte zufälligerweise jeweils auf der Ostseite einer Spalte befinden (HEINERT, 2010). Punkte auf der Westseite der Spalten hätten vermutlich eine Nordkomponente im Verschiebungsvektor. Beim Punkt LM0305 kommt hinzu, dass er sich im Bereich des Hegill Vulkanes befindet, in dessen Nähe in den 1990er Jahren mehrere Erdbeben stattfanden, die mit einer starken Aufwölbung des Gebietes einher gingen (FEIGL U.A., 2000).

Die gefundenen Blockgrenzen ziehen sich vom Norden, von der Tjörnes Bruchzone (TFZ) über die nördliche vulkanische Riftzone (NVRZ), die östliche vulkanische Riftzone (EVRZ), die südisländische seismische Zone (SISZ) bis zur Reykjanes Vulkanzone (RVZ) im Südwesten hin. In den anderen, in der

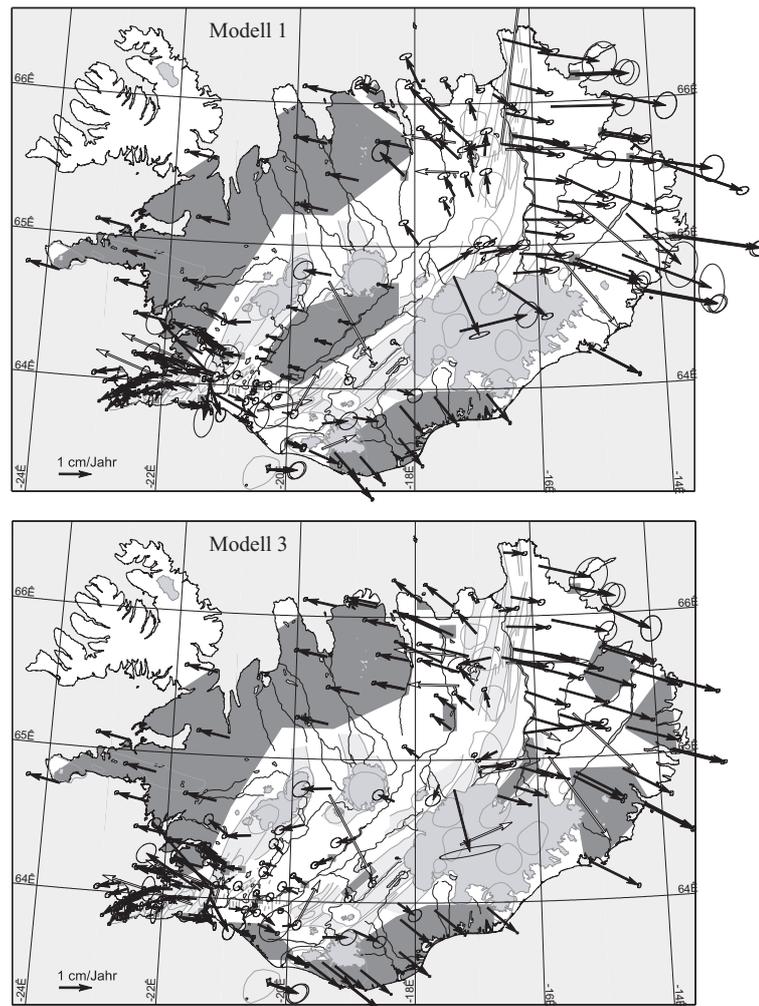


Abbildung 6.17: Die Modelle 1 (oben) und 3 (unten) aus PERLT (2006). Dunkelgrau dargestellt sind die von ihm festgelegten Blöcke synchroner Punktbewegungen.

Abbildung 6.8 gekennzeichneten Bereichen fanden entweder (in dem Zeitraum von 1993 bis 2004) keine stärkeren lokalen seismischen Aktivitäten statt oder die Punktdichte war zu gering um kleinräumige Strukturveränderungen aufzudecken.

Die gefundenen Blockgrenzen korrespondieren nur eingeschränkt mit den von PERLT (2006) beschriebenen Ergebnissen. In den Abbildungen 6.17 sind die von PERLT als Blöcke synchroner Punktbebewegungen festgelegten Bereiche dunkelgrau hinterlegt. Festsustellen ist lediglich, dass die von PERLT angenommenen Blöcke offensichtlich nicht von einer in dieser Arbeit gefundenen Blockgrenze geschnitten werden. Die Unterschiede sind im Wesentlichen dadurch begründet, dass es sich in dieser Arbeit nur um einen Zweiepochevergleich der Messkampagnen aus den Jahren 1993 und 2004 handelt und nur die Lagevektoren verglichen wurden. In die Modelle von PERLT sind dagegen insgesamt 25 Epochen aus den Jahren 1986 bis 2003 mit kompletten Lage- und Höheninformationen sowie geologische Vorinformationen eingeflossen.

Das hier vorgestellte Ergebnis stellt eine rein geometrische Interpretation dar, wie sie sich aus dem Vergleich der zwei Messepochen ergibt. Weitergehende geologische oder geophysikalische Vorinformationen sind nicht mit berücksichtigt.

Im Zusammenhang mit der zunächst durchzuführenden scharfen Clusterung wurde erwähnt, dass sich bei Verwendung des Average Linkage Verfahrens eine geringfügig andere Clusterung gegenüber der Clusterung nach dem Complete Linkage Verfahren ergeben hat. Der Unterschied liegt darin, dass

der Punkt OS7220 einem anderen Cluster zugewiesen wurde und somit sein Hauptzugehörigkeitswert ebenfalls in einem anderen Cluster liegt. Das hat für die Blockfindung die Auswirkung, dass sich im Bereich des Punktes OS7220 die Blockgrenzen ändern und der Punkt nicht mehr zu dem Block VIII, sondern zum Block II gehört. In der Abbildung 6.18 sind die Ergebnisse, wie sie sich nach den unterschiedlichen Clusterungen ergaben, zu sehen. Wenn man sich den Verschiebungsvektor des Punktes OS7220 in den Abbildungen 6.18 a) und 6.18 b) anschaut, dann fällt es nicht leicht, zu sagen, ob der Verschiebungsvektor dem Verschiebungsvektor des Punktes OS7385 oder dem des Punktes OS7225 ähnlicher ist. Zwei unterschiedliche menschliche Bearbeiter können hier durchaus auch zu unterschiedlichen Ergebnissen kommen.

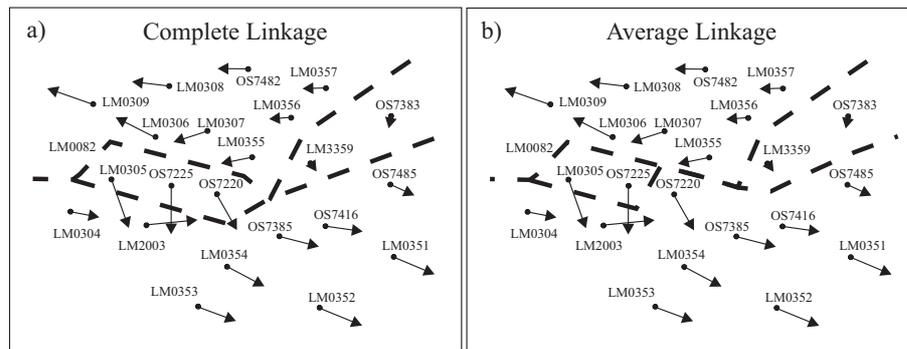


Abbildung 6.18: Die Verschiebungsvektoren und Blockgrenzen im Bereich der Südisländischen Seismischen Zone. In der Abbildung a) ist das Ergebnis der Blockgrenzen zu sehen, wie es sich nach einer Clusterung mit dem Complete Linkage Verfahren ergibt und rechts in der Abbildung b) das Ergebnis, wie es sich nach einer Clusterung mit dem Average Linkage Verfahren ergibt.

Bei der hier bearbeiteten Problemstellung, dem Auffinden von Punktgruppen mit gleichförmigem Bewegungsverhalten, ist es naheliegend, die Dreiecksvermaschung (Abbildung 6.15) zu nehmen und entlang der Dreiecksanten die Differenz der Verschiebungsvektoren der jeweiligen Endpunkte zu betrachten. Dazu wurde die Differenz im Azimut und die Differenz in der Vektorlänge berechnet. Der Gedanke war, dass eine Dreiecksante von einer Blockgrenze geschnitten wird, bei der entweder eine größere Differenz in der Vektorlänge und/oder eine größere Differenz in der Azimutrichtung auftritt. Beide Differenzen sind in einem Merkmalsraum in der Abbildung 6.19 dargestellt. Es zeigt sich in diesem Beispiel keine klare Trennung der Objektpunkte im Merkmalsraum, insbesondere in Richtung der Achse der Azimutdifferenzen. Eine Aussage über einen Schwellwert, ab welchen Differenzen eine Blockgrenze anzunehmen ist, kann nicht gemacht werden.

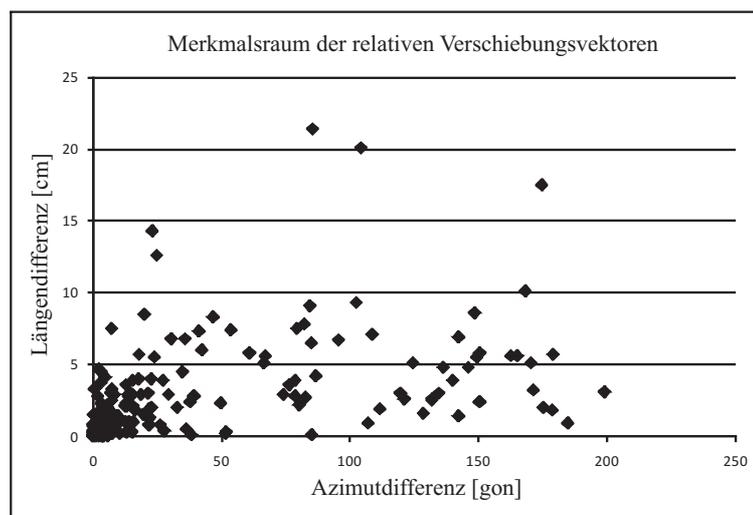


Abbildung 6.19: Die Differenzen der Verschiebungsvektoren des Beispiels ISNET entlang der Dreiecksanten.

Weiterhin wurde auf der Basis der Dreiecksvermaschung ein graphenbasiertes Clusterverfahren versucht. Die Dreiecksvermaschung wurde als Graph aufgefasst und es wurden nur die Distanzen im Merkmalsraum entlang der Kanten der Dreiecksvermaschung berechnet und nicht wie beim Complete und Average Linkage Verfahren sämtliche Distanzen zwischen allen Objekten. Die anschließende Clustering dieser Distanzen im Merkmalsraum brachte allerdings kein zufriedenstellendes Ergebnis. Bei einigen Dreieckskanten wurde zwar erkannt, dass sie von einer Blockgrenze geschnitten werden, das Ergebnis war insgesamt aber deutlich schlechter als das Ergebnis der oben beschriebenen Vorgehensweise.

Berücksichtigung unterschiedlicher elementarer Unschärfen

Durch HEMPEL (2005) wurde das Verfahren der Fuzzy Pattern Klassifikation (FPK) dahingehend erweitert, dass nicht nur für jede Merkmalsachse eine mittlere Unschärfe berücksichtigt werden kann, sondern für jeden Objektpunkt eine individuelle Unschärfe. In den bisherigen Auswertungen der Beispiele Testnetz Delft und ISNET wurden als elementare Unschärfen die Standardabweichungen eines Vektors mittlerer Länge verwendet, welche sich durch Fehlerfortpflanzung aus den Standardabweichungen der Punktkoordinaten der einzelnen Epochen ergaben. Dadurch wird die Unschärfe eines kürzeren Vektors in der Azimutrichtung zu optimistisch angenommen und die eines längeren Vektors als zu pessimistisch. Durch die Möglichkeit individuelle Unschärfen einzuführen, kann man jetzt für jeden Verschiebungsvektor die individuelle Unschärfe sowohl in der Azimutrichtung als auch in der Vektorlänge bei der Berechnung der unscharfen Klassenbeschreibung berücksichtigen. Für das Beispiel ISNET wurde es hier berechnet. Ausgehend von den in VALSSON U.A. (2007) veröffentlichten und von RENNEN (2009) mitgeteilten mittleren Standardabweichungen⁶ für die Punktkoordinaten in den Messepochen 1993 und 2004, wurden die Standardabweichungen für die Längen und die Azimute der einzelnen Verschiebungsvektoren nach dem Varianz-Fortpflanzungs-Gesetz berechnet, welche in Analogie zu den bisher durchgeführten Auswertungen als elementare Unschärfen aufgefasst wurden. Die anschließende Berechnung der unscharfen Klassenbeschreibungen wurde freundlicherweise von Herrn Arne-Jens Hempel mit der von ihm entwickelten Software durchgeführt und die Ergebnisse sowie die Grafiken in den Abbildungen 6.20 und 6.21 zu Vergleichszwecken für diese Arbeit zur Verfügung gestellt. In der Abbildung 6.20 ist die bisherige Vorgehensweise mit einheitlichen merkmalspezifischen elementaren Unschärfen zu sehen und in der Abbildung 6.21 ist für jeden Objektpunkt eine individuelle elementare Unschärfe eingeführt worden. Zwischen den beiden Abbildungen sieht man deutlich die Unterschiede. Die größere Unschärfe kurzer Vektoren in der Azimutrichtung führt bei den Clustern 1 und 5 zu deutlich ausgedehnteren Clustern in Azimutrichtung und das Cluster 3 des Punktes LM0500 mit dem längsten Verschiebungsvektor wird durch die höhere Genauigkeit in Azimutrichtung deutlich schmaler.

Auf eine erneute Berechnung von mittleren Zugehörigkeitswerten der Linien der Dreiecksvermaschung und eine weitergehende Analyse in Bezug auf Blockgrenzen wurde an dieser Stelle verzichtet. Es ist anzunehmen, dass sich in diesem Beispiel auch keine anderen Ergebnisse ergeben. Sollten aber beispielsweise die beiden Cluster 1 und 5 dichter zusammenliegen, so würden sich durch die größere Unschärfe die beiden Cluster stärker überlappen. In dem Fall wäre zu überlegen, ob eine Trennung der Cluster noch gerechtfertigt ist oder ob die beiden Cluster nicht zu einem Cluster verschmolzen werden sollten. Methoden zur Verschmelzung unscharf beschriebener Klassen nach dem Fuzzy Pattern Klassifikations Konzept finden sich in MANN (1984), SCHEUNERT (2002) und HEMPEL (2005).

Da für das Beispiel ISNET für die einzelnen Epochen nur mittlere Standardabweichungen der Punktkoordinaten vorlagen, wurden diese zur Berechnung der Standardabweichungen für die Merkmale Länge und Azimut der Verschiebungsvektoren benutzt. Sollten die kompletten Kovarianzmatrizen der Messepochen vorliegen, könnten idealerweise diese zur Berechnung herangezogen werden.

⁶ $\sigma_{x,y,1993}$ ca. 10 mm, $\sigma_{x,y,2004}$ ca. 5 mm

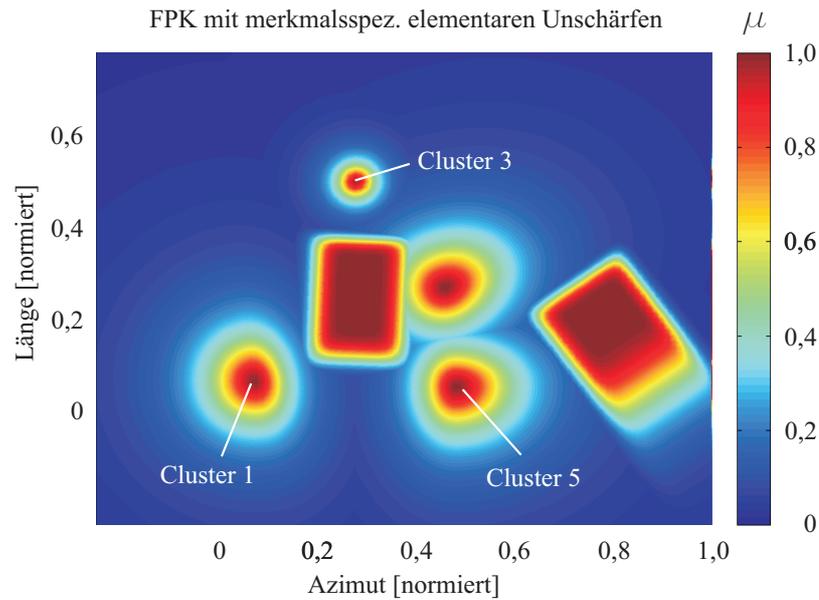


Abbildung 6.20: Die sechs unscharf beschriebenen Cluster des Beispiels ISNET im Merkmalsraum mit farbcodierten Zugehörigkeitswerten. Die unscharfen Klassenbeschreibungen sind nach dem im Abschnitt 5 beschriebenen Fuzzy Pattern Klassifikations Konzept auf der Basis jeweils einer mittleren elementaren Unschärfe pro Merkmal berechnet worden. Die Abbildung entspricht im Ergebnis der Abbildung 6.14.

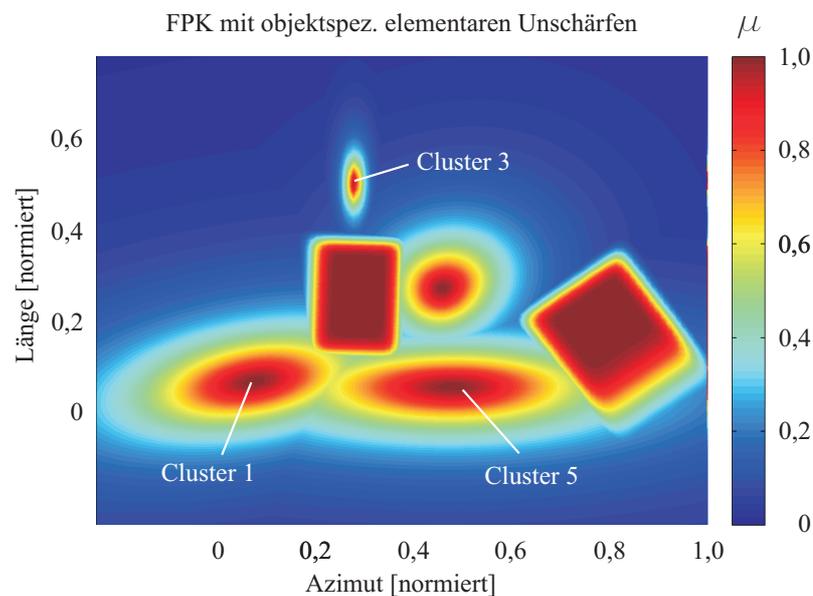


Abbildung 6.21: Die sechs unscharf beschriebenen Cluster des Beispiels ISNET im Merkmalsraum mit farbcodierten Zugehörigkeitswerten. Die unscharfen Klassenbeschreibungen sind nach dem von HEMPEL (2005) erweiterten Fuzzy Pattern Klassifikations Konzept auf der Basis individueller elementarer Unschärfen berechnet worden.

Beispielhafte Behandlung einer Klasse nichtsignifikanter Punktbewegungen

Im Abschnitt 2.3 wurde bereits erwähnt, dass es sinnvoll sein kann, eine Klasse aus den Objektpunkten zu bilden, bei denen die Länge des Verschiebungsvektors nicht signifikant ist. Das lässt sich bei dem Verfahren der Fuzzy Pattern Klassifikation relativ gut in den bisher geschilderten Auswerteprozess integrieren. Durch die Möglichkeit, die Parameter β_i der Zugehörigkeitsfunktion (Formel (5.2)) individuell zu wählen, lässt sich solch eine Klasse unabhängig vom gewählten Clusterverfahren manuell

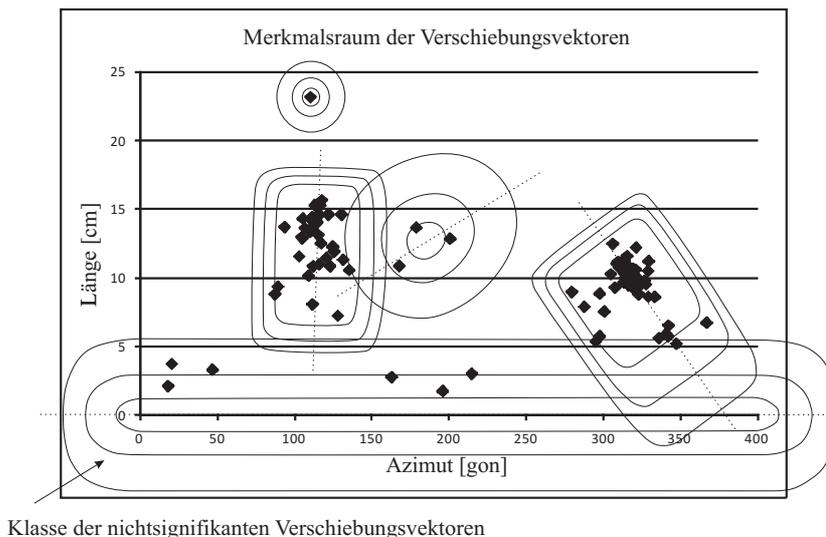


Abbildung 6.22: Linien gleicher Zugehörigkeiten für die Cluster im Merkmalsraum. Die Klasse für die Objektpunkte mit einer nichtsignifikanten Länge des Verschiebungsvektors ersetzt die in der Abbildung 6.14 abgebildeten Cluster 1 und 5.

oder auch automatisch aufbauen. Zur Demonstration ist das für das Beispiel ISNET einmal gemacht worden. Das Ergebnis ist in der Abbildung 6.22 zu sehen.

Es wurde hier angenommen, dass die Verschiebungsvektoren der sechs Objektpunkte LM0302, LM0331, LM3359, OS7383, OS7470 und RH8437, die in den Clustern 1 und 5 der bisherigen Auswertung ihren Hauptzugehörigkeitswert haben (siehe Abbildung 6.14), nicht signifikant sind.

Für die Klasse der nichtsignifikanten Verschiebungsvektoren wurden im normierten Merkmalsraum folgende Parameter⁷ gewählt:

Für die Achse des Azimutes:

$$\begin{aligned}
 x_0 &= 0,5 \text{ (Mitte der auf } [0, 1] \text{ normierten Achse)} \\
 \beta_1 &= 1,0 \text{ (Höhe der Zugehörigkeitsfunktion)} \\
 \beta_2 &= 1,0 \text{ (Zugehörigkeitswert am Rande } (x_0 \pm \beta_3)) \\
 \beta_3 &= 1,0 \text{ (Spreizung der Funktion)} \\
 \beta_4 &= 20 \text{ (steiler Abfall der Funktion)}
 \end{aligned}$$

Für die Achse der Länge:

$$\begin{aligned}
 x_0 &= 0,0 \\
 \beta_1 &= 1,0 \text{ (Höhe der Zugehörigkeitsfunktion)} \\
 \beta_2 &= 0,5 \text{ (Zugehörigkeitswert am Rande } (x_0 \pm \beta_3)) \\
 \beta_3 &= 0,05 \text{ (Spreizung der Funktion, Parameter der Unschärfe)} \\
 \beta_4 &= 2 \text{ (flacher Abfall der Funktion)}
 \end{aligned}$$

Die ehemaligen Cluster 1 und 5 (Abbildung 6.14) sind hier zugunsten der Klasse der nichtsignifikanten Verschiebungsvektoren aufgelöst worden. Das Ergebnis der Blockfindung auf der Basis der in der

⁷ Zur Bedeutung der Parameter siehe Abschnitt 5. Die Parameter für die Länge wurden in Anlehnung an die von BOCKLISCH (1987) vorgeschlagenen Parameter zur Beschreibung der Unschärfe von Objekten gewählt. Der Parameter β_3 entspricht ungefähr der normierten Standardabweichung für die Länge.

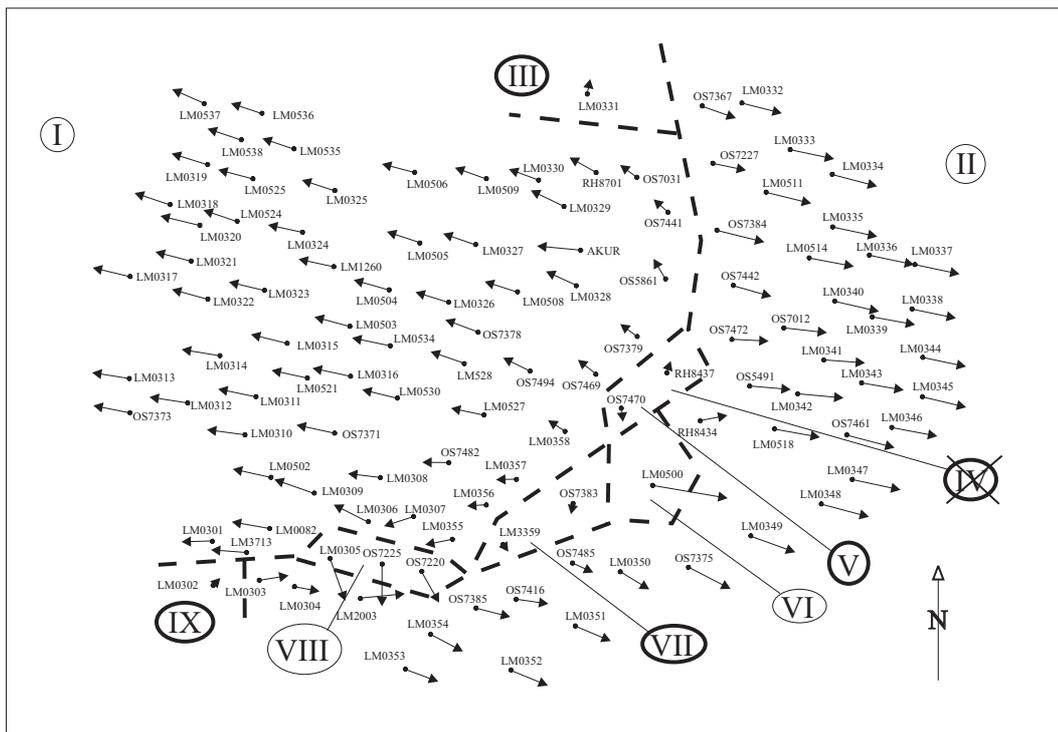


Abbildung 6.23: Die Verschiebungsvektoren des Beispiels ISNET mit den detektierten Blockgrenzen und den Blocknummern. Durch die Annahme eines Clusters nichtsignifikanter Punktverschiebungen bilden die fett umrandeten Blocknummern die Blöcke nichtsignifikanter Punktverschiebungen. Der ehemalige Block IV bildet mit dem Block V einen gemeinsamen Block.

Abbildung 6.22 dargestellten Cluster und der Klasse der nichtsignifikanter Punktverschiebungen ändert sich gegenüber der bisherigen Darstellung wie folgt:

Die bisherigen Blöcke III, IV, V, VII und IX (siehe Abbildung 6.23) bilden jeweils einen Block nichtsignifikanter Punktverschiebungen. Die nebeneinander liegenden Blöcke IV und V (siehe Abbildung 6.16) haben keine trennende Blockgrenze mehr, sondern bilden einen gemeinsamen Block. Wenn dann noch die oben erwähnte Tauschung der Diagonalen im Viereck LM0358, OS7470, LM0500, OS7383 durchgeführt würde, entstünde aus den bisherigen Blöcken IV, V und VII ein gemeinsamer Block nichtsignifikanter Punktverschiebungen.

Für reale Anwendungen ist noch zu diskutieren, wie die Parameter β_2 , β_3 und β_4 in Abhängigkeit von der Standardabweichung der Vektorlänge und eines entsprechenden Signifikanzniveaus möglichst optimal zu wählen sind. Das soll an dieser Stelle jedoch nicht weiter behandelt werden.

6.2.3 Zusammenfassung Beispiele Testnetz Delft und ISNET

In den Beispielen Testnetz Delft und ISNET wurde gezeigt, wie durch eine Abbildung der Verschiebungsvektoren eines Zweiepochevergleiches von geodätischen Überwachungsmessungen in einem Merkmalsraum und einer anschließenden unscharfen Clustering Blöcke konsistenter Punktverschiebungen gefunden werden können. Der Ablauf sei hier noch einmal kurz zusammengefasst:

1. Fuzzy Clustering der Verschiebungsvektoren von Objektpunkten im Merkmalsraum.
2. Bildung einer Dreiecksvermaschung zwischen den Objektpunkten im Objektraum durch eine Delaunay-Triangulation.

3. Berechnung mittlerer Zugehörigkeitswerte für die Dreieckskanten aus den Zugehörigkeitswerten der Endpunkte.
4. Die Ableitung einer scharfen Entscheidung (die sogenannte Defuzzifizierung) aus den mittleren Zugehörigkeitswerten auf der Basis eines empirisch ermittelten Schwellwertes.
5. Konstruktion von Blockgrenzen aus den „schlecht“ zugeordneten Dreieckskanten.

Die Ergebnisse decken sich sehr gut mit den Ergebnissen, die ein erfahrener Auswerter bei der Betrachtung von graphischen Darstellungen der Verschiebungsvektoren finden würde. Vergleiche mit anderen automatischen Auswerteverfahren sind schwierig, da lediglich für das Beispiel des Testnetzes Delft Referenzauswertungen vorliegen.

Das Verfahren ist nicht auf die vorgestellten Beispiele speziell zugeschnitten. Prinzipiell lässt sich damit jedes Vektorfeld von Verschiebungsvektoren beliebiger Größe untersuchen. Angefangen von kleinräumigen Aufgabenstellungen an Bauteilen im Rahmen von Belastungstests, über geotechnische Aufgabenstellungen, z.B. Konvergenzmessungen im Tunnelbau oder Rutschhänge, bis hin zu globalen Aufgabenstellungen im Bereich der Plattentektonik. Zur Aufdeckung von Blockgrenzen ist es nur wichtig, dass es unter den Verschiebungsvektoren Unstetigkeiten gibt, die im Merkmalsraum dazu führen, dass sich Regionen von relativ dichten Punktwolken ausbilden, die durch Regionen mit einer geringeren Dichte voneinander getrennt sind. Sollten im Vektorfeld nur kontinuierliche Übergänge vorhanden sein, so sind die Cluster im Merkmalsraum nicht getrennt und die Clusterverfahren versagen.

Eine Übertragung des Ansatzes für zwei- auf ein- und dreidimensionale Aufgabenstellungen ist machbar. Weiterhin ist eine Übertragung auf die Analyse von mehreren Epochen denkbar. Dabei gibt es zwei Möglichkeiten:

1. Man arbeitet mit einem statischen Klassifikator, bei dem die einmal bestimmten Klassen beibehalten werden und in den Folgeepochen von diesem Klassifikator für die Verschiebungsvektoren Zugehörigkeitswerte berechnet werden. Die Veränderung der Zugehörigkeitswerte kann dann zur Interpretation des Bewegungsverhaltens herangezogen werden.
2. Man benutzt einen sogenannten dynamischen Klassifikator, bei dem nach jeder Folgeepoche neue Klassenstrukturen berechnet werden. Hier können zur Interpretation des Bewegungsverhaltens sowohl die Klassenparameter (insbesondere die Klassenzentren) als auch die Zugehörigkeitswerte herangezogen werden.

Als problematisch hat sich herausgestellt, ein optimales Clusterverfahren zu finden, welches auch robust mit Clustern unterschiedlicher Größe und Dichte sinnvolle Ergebnisse liefert. Bei den hier vorliegenden Beispielen stellten sich das Complete Linkage und das Average Linkage Verfahren als die geeignetsten Verfahren heraus. Hier besteht sicherlich weiterer Forschungsbedarf. Nicht umsonst gibt es so eine große Anzahl an unterschiedlichen Verfahren, die für unterschiedliche Anwendungsfälle unterschiedliche Stärken und Schwächen haben.

Ähnliches gilt für die Findung einer „optimalen“ Clusteranzahl. Die verwendeten Kriterien führen teilweise zu schwer interpretierbaren Clusterlösungen, was aus der Literatur auch bekannt ist. Deshalb ist es bisher notwendig, die Klassenbildung im Merkmalsraum visuell zu kontrollieren, was in einem automatischen Auswerteprozess eigentlich nicht gewünscht ist. Durch die Berechnung mittlerer quadrierter euklidischer Distanzen im Merkmalsraum als Gütekriterium und deren Interpretation im Zusammenhang mit der Clusteranzahl wurden hier die besten Ergebnisse erzielt.

Ein grundsätzliches Problem bei der Clusterung in euklidischen Merkmalsräumen ist die Normierung der Merkmale. Je nach Relation der Wertebereiche für die Längen und die Azimute der Verschiebungsvektoren, wird durch die Normierung eine unterschiedliche Gewichtung vorgenommen, die sich auf die Clusterung auswirken kann. Hier ist darauf zu achten, dass die Normierung aus fachlicher Sicht sinnvoll vorgenommen wird.

Das Ergebnis der Findung von Blockgrenzen hängt im gewissen Rahmen mit von der Bildung der Dreiecksvermaschung ab. Diese Abhängigkeit von einer Dreiecksvermaschung ist in der Geodäsie ein bekanntes Problem, zum Beispiel bei der Erstellung digitaler Geländemodelle. Ein Hilfsmittel zur besseren Lokalisierung der Blockgrenzen ist hier nur über eine Erhöhung der Messpunktdichte in den Bereichen vermuteter Blockgrenzen möglich.

6.3 Automatische Klassifizierung von Setzungsmessungen

In diesem Abschnitt soll eine weitere Problemstellung aus dem Bereich der Analyse von Überwachungsmessungen betrachtet werden. Die automatische Klassifizierung von Setzungsmessungen.

Setzungsmessungen werden üblicherweise in mehreren Messepochen durchgeführt und die Ergebnisse als Zeitreihe in numerischer und/oder graphischer Form (wie zum Beispiel in der Abbildung 6.24) präsentiert. Aus der graphischen Darstellung der Zeitreihe kann ein erfahrener Auswerter relativ leicht auf das Bewegungsverhalten eines Messpunktes schließen und er ordnet das Bewegungsverhalten einem bestimmten Muster zu (z.B. lineare Setzung, beschleunigt zunehmende Setzung, konsolidierende Setzung oder keine Setzung). Bei heutigen Großprojekten werden teilweise an weit über 1000 Punkten Setzungen durch automatische Messverfahren erfasst (z.B. HEINZ und PÖNITZ, 2006), die nur mit sehr großem Aufwand manuell ausgewertet und insbesondere auch interpretiert werden können. Für die Belange des Tunnelbaus hat CHMELINA (2003) einen speziellen Ansatz vorgestellt, der auf der Basis von Regeln arbeitet, die in einer Wissensbasis abgelegt sind. In dieser Arbeit soll ein Ansatz vorgestellt werden, der die automatische Zuordnung des Bewegungsverhaltens zu einem bestimmten Muster mit Hilfe von Fuzzy-Clusterverfahren durchführt. Insbesondere die Fuzzy-Verfahren eignen sich hier, da eine eindeutige Zuordnung zu einem bestimmten Muster oftmals nicht möglich oder nicht sachgerecht ist.

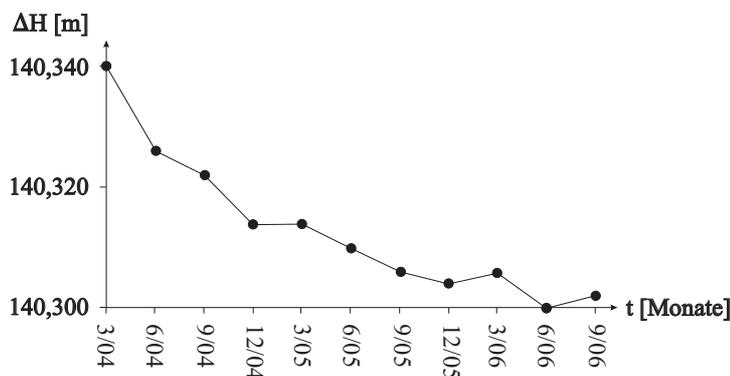


Abbildung 6.24: Graphische Darstellung der Zeitreihe von Setzungsmessungen an einem Punkt.

Die Merkmale, nach denen ein erfahrener Auswerter die Zeitreihen bestimmten Mustern zuordnet, sind z.B. die mittlere Steigung der Zeitreihe zwischen den ersten Epochen und die mittlere Steigung zwischen den letzten Epochen. Diese Vorgehensweise wird in dem hier vorgestellten automatischen Ablauf nachgebildet. Theoretisch sind beliebig viele Merkmale möglich. Da sich ein zweidimensionaler Merkmalsraum graphisch einfacher darstellen lässt als höherdimensionale Merkmalsräume, werden in diesem Beispiel nur die zwei folgenden Merkmale verwendet: Die mittlere Steigung der Zeitreihe in der ersten Zeithälfte und die mittlere Steigung der Zeitreihe in der zweiten Zeithälfte.

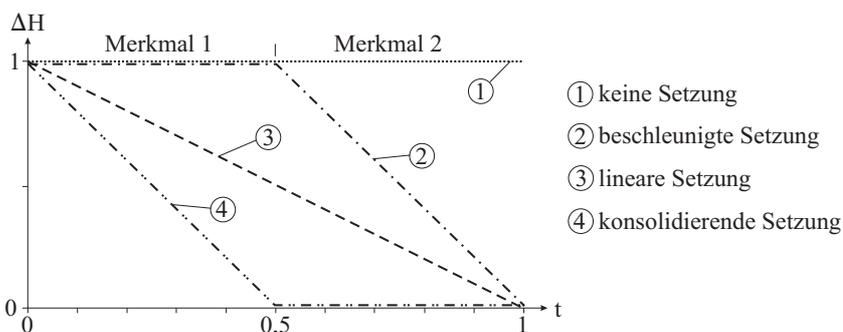


Abbildung 6.25: Einfache Prototypen von möglichen Setzungsmustern.

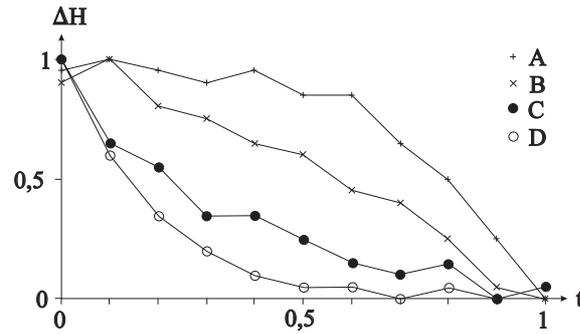


Abbildung 6.26: Normierte Zeitreihen (A, B, C, D) von Setzungsmessungen.

Der Ablauf des Klassifizierungsprozesses wird im folgenden beschrieben:

1. Durch einen Experten werden zu Beginn der Auswertung bestimmte Prototypen zu erwartender Bewegungsmuster in einem normierten Zeit- und Messwertbereich (z.B. 0-1) vorgegeben (Abbildung 6.25).

Über ihre Merkmale werden diese Prototypen (hier 1, 2, 3, 4) dann im Merkmalsraum abgebildet (Abbildung 6.27) und bilden dort die Klassenzentren, denen später die real gemessenen Zeitreihen (hier die Reihen A, B, C, D in Abbildung 6.26) zugeordnet werden.

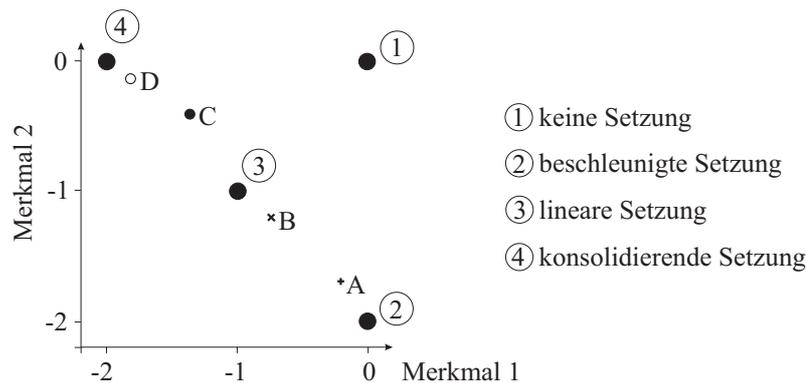


Abbildung 6.27: Zweidimensionaler Merkmalsraum mit den Clusterzentren 1 - 4 und den Messreihen A - D.

2. Um vergleichbare Zustände zu bekommen, werden nach jeder Messepoche die real gemessenen Zeitreihen (Abbildung 6.26) auf denselben Bereich wie die Prototypen normiert. Aus diesen Zeitreihen werden anschließend die gleichen Merkmale wie bei den Prototypen berechnet und ebenfalls im Merkmalsraum abgebildet.
3. Nach der Berechnung der Distanzen von der Lage der Zeitreihe im Merkmalsraum zu den Klassenzentren der Prototypen können Zugehörigkeitswerte zu den Prototypen berechnet werden. In FLETTLING (2007) und FLETTLING (2008) wurden diese z.B. nach der Formel (4.4) des Fuzzy-c-Means Algorithmus berechnet. Die Werte sind der Tabelle 6.7 zu entnehmen. Der jeweils höchste Zugehörigkeitswert (Hauptzugehörigkeitswert) einer Messreihe zu den Prototypen ist fett hervorgehoben.

Anhand der mit der Formel (4.4) berechneten Zugehörigkeitswerte (Tabelle 6.7) erkennt man die unscharfe Zuordnung der realen Messreihen zu den Prototypenmustern. Die Messreihen A, B und D werden sehr gut den entsprechenden Mustern zugeordnet. Sie weisen einen maximalen Zugehörigkeitswert auf, der deutlich größer ist als der zweitgrößte Zugehörigkeitswert dieser Messreihe. Einzig die

Messreihe	Setzungen			
	keine	beschleunigt	linear	konsolidierend
A	0,03	0,89	0,07	0,01
B	0,05	0,08	0,85	0,02
C	0,11	0,05	0,41	0,43
D	0,01	0,00	0,02	0,97

Tabelle 6.7: Probabilistische Zugehörigkeitswerte der Messreihen zu den Prototypen, nach Formel (4.5) berechnet.

Messreihe C hat sowohl zu den Mustern einer linearen Setzung als auch zu denen einer konsolidierenden Setzung ungefähr gleich große Zugehörigkeitswerte. Die hier durch den Klassifikator getroffene Zuordnung wäre sicherlich von einem erfahrenen Auswerter in gleicher Weise getroffen worden. Das Verfahren zeigt, dass mit solch einem Klassifikator manuelle Interpretationen von Messreihen durch ein automatisch arbeitendes System weitestgehend ersetzt werden können und nur wenige, nicht gut zu klassifizierende Messreihen (z.B. die Messreihe C) bei Bedarf von einem erfahrenen Auswerter interpretiert werden müssen.

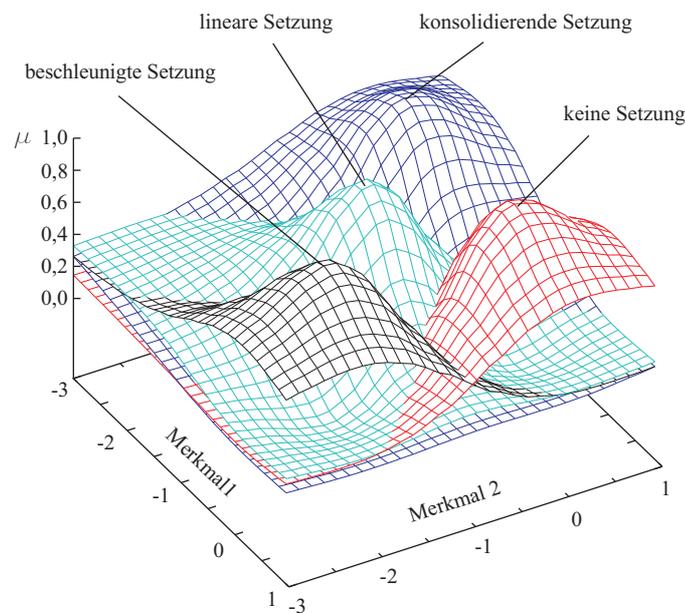


Abbildung 6.28: Die vier unscharfen probabilistischen Zugehörigkeitsfunktionen zu den vier Prototypen, nach Formel (4.4) berechnet.

Bei Verwendung der in dem Beispiel bisher benutzten Zugehörigkeitsfunktion des Fuzzy-c-Means Algorithmus nach der Formel (4.4) oder anderer probabilistischer Clusterverfahren drückt der Zugehörigkeitswert $\mu_{i,k}$ die Zugehörigkeit eines Objektes zu einem Cluster in Relation zu den anderen Clustern aus. Der Zugehörigkeitswert zu einem Cluster hängt aufgrund der Nebenbedingungen bei den probabilistischen Verfahren ($\sum_{i=1}^c \mu_{i,k} = 1$ und $\mu_{i,k} \in [0, 1]$) auch mit von der Lage und Anzahl der anderen Cluster ab. Er drückt nicht aus, wie typisch das Objekt für das entsprechende Cluster ist. Auf dieses Problem wurde bereits im Abschnitt 4.5 hingewiesen. Das kann in der Praxis bei Messreihen zu Interpretationsproblemen führen, die eine relativ große Distanz zu allen Klassenzentren aufweisen⁸. In dem Fall streben die Zugehörigkeitswerte dieser Messreihen zu allen Prototypen gegen

⁸ Das wäre zum Beispiel der Fall, wenn eine Messreihe eine Hebung repräsentiert, aber nur Prototypen für Setzungen definiert sind.

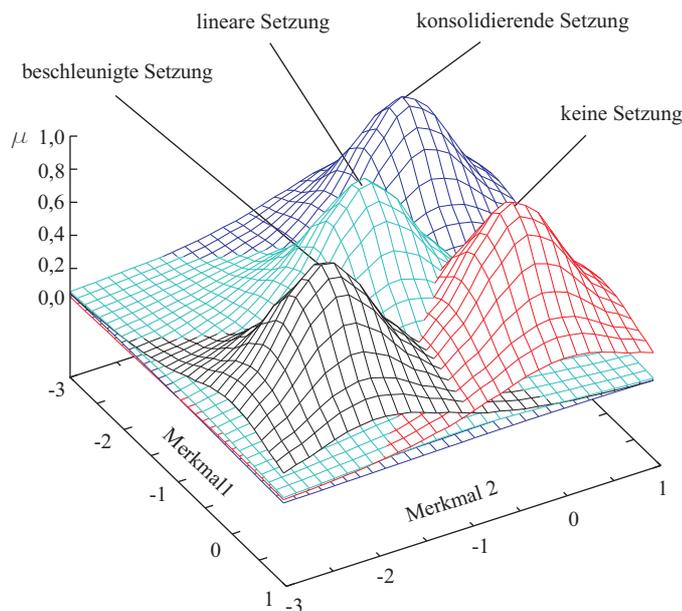


Abbildung 6.29: Die vier unscharfen possibilistischen Zugehörigkeitsfunktionen zu den vier Prototypen, nach Formel (4.13) mit $\eta_i = 0,5$ berechnet.

1/c. Andeutungsweise kann man das in der Abbildung 6.28 in der linken Ecke in der Nähe der μ -Achse erkennen.

Die Zugehörigkeitsfunktionen von den possibilistischen Clusterverfahren, zum Beispiel Formel (4.13), die auf die oben genannten Nebenbedingungen verzichten, zeigen hier ein anderes Verhalten. In Abbildung 6.29 ist dies gut zu erkennen. Hier hängt der Zugehörigkeitswert nur vom Abstand zum dazugehörigen Clusterzentrum ab und nicht auch von den anderen Clusterzentren. Bei Messreihen, die im Merkmalsraum eine relativ große Distanz zu allen Clusterzentren aufweisen, streben die Zugehörigkeitswerte alle gegen Null. Das entspricht in diesem Anwendungsfall wesentlich besser einer menschlichen Interpretation. Dadurch, dass bei diesem Verfahren in der Regel $\sum_{i=1}^c \mu_{i,k} \neq 1$ ist, kann die Summe auch zur Interpretation mit herangezogen werden. Wenn $\sum_{i=1}^c \mu_{i,k} > 1$ ist, deutet das darauf hin, dass im Merkmalsraum in der Nähe des Objektpunktes viele Prototypen liegen und eine automatische Klassifikation relativ sicher durchgeführt wird. Wenn aber $\sum_{i=1}^c \mu_{i,k} < 1$ ist, liegt der Objektpunkt relativ weit von allen Prototypen entfernt und der Klassifikator ist für diesen Fall schlecht geeignet. In diesem Fall sollte die Messreihe auf jeden Fall von einem erfahrenen Auswerter interpretiert werden.

Messreihe	Setzungen				Σ
	keine	beschleunigt	linear	konsolidierend	
A	0,14	0,84	0,29	0,07	1,34
B	0,20	0,30	0,82	0,14	1,46
C	0,19	0,10	0,49	0,48	1,26
D	0,13	0,07	0,25	0,93	1,38

Tabelle 6.8: Possibilistische Zugehörigkeitswerte der Messreihen zu den Prototypen, nach Formel (4.13), mit $\eta_i = 0,5$ berechnet.

Für das hier betrachtete Beispiel wurden ebenfalls possibilistische Zugehörigkeitswerte nach der Formel (4.13) mit $\eta_i = 0,5$ berechnet (siehe Tabelle 6.8). Die jeweils maximalen Zugehörigkeitswerte der Messreihen unterscheiden sich nicht sehr stark von den probabilistischen Zugehörigkeitswerten aus Tabelle 6.7. Lediglich bei der schlecht eindeutig zuzuordnenden Messreihe C hat sich eine geringfügige

Veränderung des maximalen Zugehörigkeitswertes von der Klasse „konsolidierende Setzung“ zu der Klasse „lineare Setzung“ ergeben. Das ändert aber nichts an der nicht eindeutigen Zuordnung, da der maximale Zugehörigkeitswert der beiden Klassen sich nur wenig unterscheidet. Größere Veränderungen hat es bei den Klassen ergeben, zu denen die Zugehörigkeitswerte bisher sehr klein waren, die sogenannten Nebenzugehörigkeitswerte. Sie sind alle größer geworden. Dadurch wird die Summe aller Zugehörigkeitswerte einer Messreihe größer als 1, was ein Zeichen dafür ist, dass sich in der Nähe der Messreihe im Merkmalsraum mehrere Prototypen befinden und von daher eine sichere Zuordnung prinzipiell möglich ist.

Neben den verwendeten Zugehörigkeitsfunktionen vom Potentialfunktionstyp wurden auch Untersuchungen mit stückweise linearen Zugehörigkeitsfunktionen vom Dreieckstyp durchgeführt. Die Ergebnisse sind allerdings aus der Betrachtung eines menschlichen Auswerters nicht so optimal. Wählt man den Träger der Zugehörigkeitsfunktion (der Bereich, in dem die Zugehörigkeitsfunktion Werte von $\mu_{i,k} > 0$ liefert) zu klein, ergeben sich überall scharfe Zugehörigkeiten zu den Prototypen, weil die Zugehörigkeitswerte zu allen anderen Prototypen zu null werden. Wählt man dagegen den Träger größer, ergeben sich schnell aus menschlicher Sicht schwer zu interpretierende hohe Zugehörigkeitswerte bei mittleren Distanzen.

6.3.1 Verhalten im Zeitbereich

In der Praxis würde der beschriebene Ablauf der Zuordnung einer konkreten Messreihe zu einem der vordefinierten Musterprototypen nach jeder Messepoche wieder neu erfolgen. Das heißt, dass nach jeder Messepoche die Messreihe wieder neu im Zeit- und Messwertbereich normiert wird und neue Merkmale und anschließend neue Zugehörigkeitswerte zu den Prototypen berechnet werden. Für die vorgestellten Beispielzeitreihen A-D wurde das einmal gemacht und die possibilistischen Zugehörigkeitswerte der Messreihen zu den Prototypen nach jeder Epoche graphisch in der Abbildung 6.30 dargestellt⁹.

In der Abbildung 6.30 kann man sehr gut erkennen, wie die Zeitreihen im Merkmalsraum wandern und sich die Zuordnung teilweise nach neu hinzugekommenen Epochen verändert. Am deutlichsten ist das bei der Messreihe A zu erkennen. Nach 5 Epochen wird sie eindeutig der Klasse „keine Setzung“ zugeordnet, obwohl beim Betrachten der Originalzeitreihe (siehe Abbildung 6.26) eine leichte lineare Setzung zu erkennen ist. Das hat seinen Grund darin, dass die mittlere Steigung der Messreihe durch die normierten Messwerte der 1., 2. und 3. Epoche die gleiche mittlere Steigung aufweisen, wie durch die normierten Messwerte der 3., 4. und 5. Epoche. Um diesen Effekt zu verhindern, müssten die Bereiche zur Berechnung der Merkmale sich noch um mindestens eine weitere Epoche überdecken. Dieser Effekt zeigt sich allerdings nur bei der Berechnung nach wenigen Epochen, wo eine Zuordnung in der Regel sowieso noch unsicher ist. Nach sieben Epochen hat die Zeitreihe einen maximalen Zugehörigkeitswert zur Klasse „lineare Setzung“. Ab der 8. Epoche wird dann ein maximaler Zugehörigkeitswert zur Klasse „beschleunigte Setzung“ angezeigt und die Zugehörigkeitswerte zu den anderen Klassen sind deutlich kleiner.

Bei der Messreihe B zeigt sich ab der 5. Epoche ein maximaler Zugehörigkeitswert zur Klasse „lineare Setzung“ und die anderen Zugehörigkeitswerte sind hier auch deutlich kleiner, was auf eine eindeutige Zuordnung zur Klasse „lineare Setzung“ hinweist.

Bei der Messreihe C sind die Zugehörigkeitswerte zu den Klassen „lineare Setzung“ und „konsolidierende Setzung“ nach allen Epochen nahezu gleich groß und die Zugehörigkeitswerte zu den beiden anderen Prototypen sind sehr klein. Hier kann in dem betrachteten Zeitraum bis zur 11. Epoche keine eindeutige Zuordnung gefunden werden.

⁹ Sinnvoll ist die Berechnung der zwei Merkmale in diesem Beispiel erst nach der 3. Epoche. Deshalb fangen die Darstellungen auch erst nach der 3. Epoche an. Im Fall einer geraden Anzahl von Epochen wurde das Merkmal 1 (Steigung in der ersten Hälfte der Zeitreihe) aus der ersten Hälfte der Epochen berechnet und das Merkmal 2 (Steigung in der zweiten Hälfte der Zeitreihe) aus der zweiten Hälfte der Epochen. Für den Fall einer ungeraden Anzahl von Epochen ging die mittlere Epoche sowohl in die Berechnung des Merkmals 1 als auch des Merkmals 2 ein.

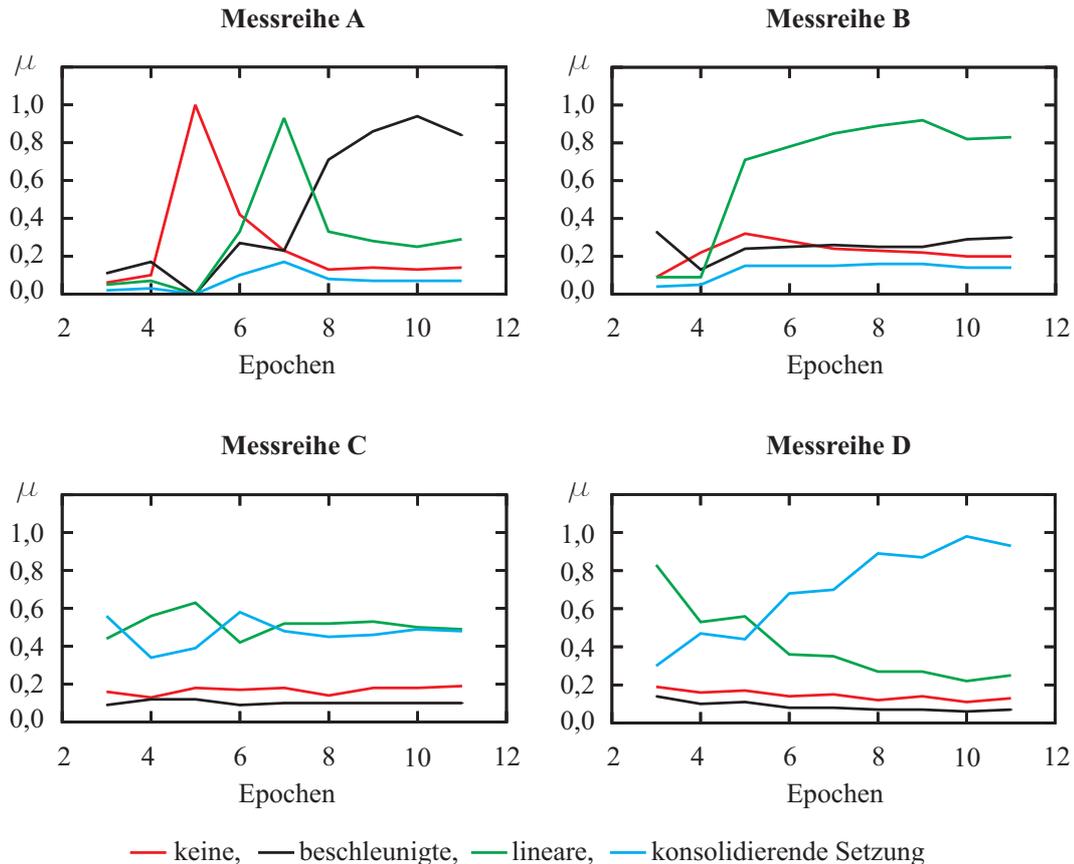


Abbildung 6.30: Veränderung der Zugehörigkeitswerte der Beispielmessreihen A, B, C und D in Abhängigkeit von den Messepochen.

Die Messreihe D hat zunächst nach 3 Epochen einen relativ hohen maximalen Zugehörigkeitswert zur Klasse „lineare Setzung“, der dann nach der 4. und 5. Epoche kleiner wird und ungefähr gleich groß ist wie der Zugehörigkeitswert zur Klasse „konsolidierende Setzung“. Dieser wird mit zunehmender Epochenanzahl immer größer und der Zugehörigkeitswert zur Klasse „lineare Setzung“ immer kleiner. Mit zunehmender Epochenanzahl wird die Messreihe D immer eindeutiger der Klasse „konsolidierende Setzung“ zugeordnet.

Da die Prototypen den langfristigen Trend der Messreihen repräsentieren, stellt sich eine sichere Zuordnung oft erst nach einer gewissen Anzahl von Epochen ein. Bei wenigen Epochen hebt sich teilweise das Signal der Messreihe vom Rauschen noch nicht deutlich genug ab.

6.3.2 Bemerkungen zu dem Verfahren

Neben der Zuordnung von Messreihen mit einem hohen Rauschen ist bei dem hier vorgestellten Ansatz die Zuordnung von Zeitreihen mit periodischen Anteilen auch problematisch. Bei dem Vorhandensein von periodischen Anteilen müssten diese eventuell in einem vorgeschalteten Berechnungsschritt eliminiert werden, um den Trend der Messreihe zu bekommen.

Durch die Normierung der Messreihen sowohl im Zeit- als auch im Messwertbereich geht der Bezug zu den absoluten Setzungsbeträgen verloren. Von daher ist neben der Angabe der Klassenzuordnung auch eine Angabe der absoluten Setzungsbeträge in der Regel notwendig. Zur Lösung des Problems ist es auch denkbar, dass man für konkrete Projekte projektbezogene Prototypen verwendet, die aufgrund

prognostizierter Setzungen definiert werden können und dass auf eine Normierung ganz verzichtet wird.

Je nach interessierendem Bereich in der Zeitreihe kann die Klassifizierung entweder über die gesamte Zeitreihe erfolgen oder aber auch nur über einen Teilbereich. Bei längeren Zeitreihen ist zum Beispiel nicht nur das Verhalten über den gesamten Zeitraum, der durchaus mehrere Jahre betragen kann, sondern insbesondere das Verhalten in den letzten Messepochen von Interesse. In dem Fall kann die Klassifizierung entweder nur über diesen Teilbereich oder über beide Bereiche durchgeführt werden.

Der hier vorgestellte Ansatz lässt sich beliebig erweitern. Denkbar sind weitere Prototypen, die z.B. auch Hebungen berücksichtigen oder für ganz spezielle Aufgaben typische Trendverläufe (z.B. GROSSAUER und LENZ, 2007) aufweisen. Der Einsatz einer größeren Anzahl von Merkmalen ist möglich. Eine Erweiterung des Ansatzes auf die Mustererkennung von Punktbewegungen in Lagenetzen oder auch in 3D-Netzen ist machbar. So wie in dem vorgestellten Beispiel Prototypen für Hebungen und Setzungen als positive und negative Bewegungen entlang der Höhen-Achse definiert werden, könnten entsprechende Prototypen auch entlang von Rechts- und Hochwertachsen definiert werden. Das könnte in einem dem Untersuchungsgebiet übergeordneten Koordinatensystem erfolgen oder um die Interpretation zu erleichtern, in einem so gedrehten Koordinatensystem, dass eine Achse in Richtung der zu erwartenden Hauptbewegungsrichtung liegt. Die Klassifizierung könnte entweder in für jede Koordinatenachse getrennten Merkmalsräumen oder in einem für alle Koordinatenachsen gemeinsamen höherdimensionalen Merkmalsraum erfolgen. Vom theoretischen Ansatz her ist das kein Problem. Auf ähnliche Beispiele aus anderen Fachdisziplinen wird in BOCKLISCH (1987) hingewiesen. Eine Aussage zu einem Punkt in einem Lagenetz könnte dann beispielsweise so aussehen: Punkt 256: beschleunigte Bewegung in X-Richtung, keine Bewegung in Y-Richtung. Wobei zur Interpretation der Aussagen die komplette Zugehörigkeitsmatrix herangezogen werden kann.

Beim Bau langestreckter Anlagen (z.B. Verkehrswege, Tunnel) werden zurzeit zur Dokumentation des Verschiebungsverhaltens von Objektpunkten häufig sogenannte LQS-Diagramme verwendet (BIERONSKI und STOLP, 2010). In denen ist das Verschiebungsverhalten achsbezogen in Längs-, Quer- und Setzungsrichtung graphisch über der Zeit dargestellt. In Analogie zu diesen Diagrammen könnte man das Setzungsverhalten für jede Richtung klassifizieren und durch die Zugehörigkeitswerte beschreiben.

Kommerzielle Softwarepakete für geodätische Monitoringsysteme bieten neben dem Ansteuern der Sensoren, der Datensammlung und Aufbereitung und der Datenspeicherung auch eine Visualisierungskomponente, mit denen die erfassten Zeitreihen dargestellt werden können. Eine weitergehende automatische Analyse der Zeitreihen, bis hin zu einer automatischen Interpretation, findet nicht statt. Es können lediglich Grenzwerte für Verschiebungsbeträge festgesetzt werden, deren Überschreitung zu automatisch generierten Alarmmeldungen führen (STEMPFHUBER, 2009). Durch den hier vorgestellten Ansatz der automatischen Klassifikation der Zeitreihen und deren Beschreibung durch die Zugehörigkeitswerte über der Zeit, sind Trends und sich anbahnende Klassenwechsel erkennbar (BOCKLISCH und BITTERLICH, 1993), bevor irgendetwelche Grenzwerte überschritten sind.

Kapitel 7

Zusammenfassung

In dieser Arbeit wurden mit Hilfe von Ansätzen, die aus der Mustererkennung abgeleitet sind, zwei typische Aufgabenstellungen aus dem Bereich der Analyse geodätischer Überwachungsmessungen bearbeitet.

Die dabei benutzten Ansätze bedienen sich automatisch arbeitender Clusterverfahren, die unter Nutzung formaler Kriterien Objekte, die durch typische Merkmale beschrieben sind, in homogene Gruppen einordnen. Die Besonderheit in den hier benutzten Ansätzen sind die verwendeten Fuzzy-Clusterverfahren, bei denen die Objekte nicht genau einer Gruppe zugeordnet werden, sondern jeder Gruppe mit einem graduell variierenden Zugehörigkeitswert. Diese Art der Betrachtung entspricht dadurch wesentlich besser der realen Welt und der menschlichen Denkweise als eine eindeutige Zuordnung. So gibt es in der realen Welt beispielsweise auch nicht nur die Farben Schwarz und Weiß, sondern dazwischen noch beliebig viele Grautöne. Die üblicherweise verwendeten scharfen Klassifikationen täuschen dagegen eine scharfe Realität vor. Sie erleichtern zwar vordergründig die Interpretation, es besteht aber die Gefahr von Fehlinterpretationen, da Daten in der Nähe der Cluster Grenzen sich oftmals nur gering unterscheiden und sowohl dem einen, als auch dem anderen Cluster zugeordnet werden können. Bei den scharfen Zuordnungen führt die binäre Entscheidung an einer Cluster Grenze zu einem Informationsverlust, der in der Regel keinen Rückschluss auf die Korrektheit der Entscheidung zulässt. Da oftmals am Ende eines Datenanalyseprozesses scharfe Entscheidungen gewünscht sind, können diese aus den unscharfen Zuordnungen zum Beispiel mit Hilfe von Schwellwert- oder Maximumoperationen abgeleitet werden.

In der vorgelegten Arbeit wurde nach einer Einführung zunächst die Mustererkennung in der Informatik mit ihren wesentlichen Bestandteilen - Extraktion typischer Merkmale des Musters und Klassifikation der Muster basierend auf den Merkmalen - betrachtet. Die Merkmale als typische Eigenschaften des Musters dienen dabei zu seiner Beschreibung. Das Muster kann als Punkt in einem, entsprechend der Anzahl der Merkmale, hochdimensionalen Raum interpretiert werden. Die Klassifikation bedeutet dabei, die vorliegenden Muster oder Teile von Mustern so zu strukturieren, dass sie eine Klasse bilden, die möglichst ähnliche oder gleiche Ausprägungen ihrer Merkmale aufweisen. Muster oder Teile von Mustern verschiedener Klassen sollten möglichst unterschiedliche Ausprägungen ihrer Merkmale haben. Bei der in dieser Arbeit betrachteten automatischen Klassifikation, der sogenannten Clusterung und den vorliegenden metrischen Merkmalen, wurden die Ähnlichkeiten über Distanzfunktionen im Merkmalsraum bestimmt.

Danach schloss sich eine Einführung in die Theorie unscharfer Mengen (engl. Fuzzy-Sets) soweit an, wie es für die in dieser Arbeit behandelten Teilaspekte notwendig ist. Darauf aufbauend erfolgte die Vorstellung verschiedener Verfahren zur automatischen Klassifizierung (Clusterung) und unscharfen Beschreibung von Clustern sowie verschiedener Kriterien zur Beurteilung einer Clusterkonfiguration.

Die zunächst theoretischen Betrachtungen wurden anschließend auf zwei typische Aufgabenstellungen aus dem Bereich der Analyse geodätischer Überwachungsmessungen übertragen. In der ersten Aufgabenstellung ging es in einem zweidimensionalen Überwachungsnetz um das Auffinden von Punktgruppen mit einem gleichförmigen Bewegungsverhalten. In der zweiten Aufgabenstellung ging es um die automatische Klassifizierung von Setzungszeitreihen. Beides sind Beispiele, die in der bisherigen Praxis nahezu ausschließlich von Menschen, anhand ihrer einschlägigen Erfahrung, durch visuelle Betrachtung graphischer Darstellungen bearbeitet wurden. Für beide Aufgabenstellungen besteht der Bedarf nach automatischen Lösungen, die den menschlichen Bearbeiter möglichst vollständig oder zumindest für ausgewählte Fälle ersetzen sollen.

Die erste Aufgabenstellung, das Auffinden von Punktgruppen mit einem gleichförmigen Bewegungsverhalten, wurde anhand von zwei numerischen Beispielen behandelt. Als Merkmale für die Clusterung wurden in normierter Form die Länge und das Azimut der Verschiebungsvektoren benutzt. Das erste numerische Beispiel (Testnetz Delft) war relativ einfach aufgebaut. Es gab nur wenige Objektpunkte, bei denen die Verschiebungsvektoren gut getrennte Cluster nahezu gleicher Größe im Merkmalsraum bildeten. Dieses sehr anschauliche Beispiel wurde auch bereits im theoretischen Teil der Arbeit zur Verdeutlichung theoretischer Sachverhalte genutzt.

Das zweite Beispiel (isländisches Grundlagennetz ISNET) war deutlich komplexer. Einerseits bestand es aus deutlich mehr Punkten und andererseits bildeten die Verschiebungsvektoren im Merkmalsraum Cluster unterschiedlicher Größe und Dichte.

In beiden Beispielen wurde ein zweistufiger Ansatz gewählt, um die Blockgrenzen aufzudecken. In der ersten Stufe erfolgte eine unscharfe Clusterung der Verschiebungsvektoren im Merkmalsraum und im zweiten Schritt die Berücksichtigung der Nachbarschaft im Objektraum durch eine Dreiecksvermaschung. Ausgehend von den Zugehörigkeitswerten der Verschiebungsvektoren wurden mittlere Zugehörigkeitswerte für die Dreiecksseiten berechnet, aus denen dann anhand einer Schwellwertoperation die Blockgrenzen abgeleitet wurden. Der Vorteil des Fuzzy-Ansatzes lag hier darin, dass die Unschärfe der Clusterung aus dem ersten Schritt über die Zugehörigkeitswerte mit in den zweiten Schritt übernommen werden kann, ohne dass schon eine Vorentscheidung stattfindet. Darüber hinaus können die mittleren Zugehörigkeitswerte für die Dreiecksseiten zur weiteren Analyse der Blockgrenzen genutzt werden.

Die gefundenen Blockgrenzen korrespondieren sehr gut mit dem, was ein menschlicher Auswerter aus einer graphischen Darstellung der Verschiebungsvektoren erkennt. Und sie decken sich ebenfalls gut mit den teilweise vorliegenden Referenzbewertungen nach anderen Verfahren.

Die zweite Aufgabenstellung, das Identifizieren von Setzungszeitreihen, ist ebenfalls eine Anwendung, die bisher in der Praxis fast ausschließlich durch das Betrachten einer graphischen Darstellung von einem erfahrenen menschlichen Auswerter durchgeführt wurde. Aus Anschauungsgründen wurde in dieser Arbeit ein relativ einfaches Modell gewählt, welches nur mit zwei Merkmalen pro Zeitreihe auskommt. Aufbauend auf Expertenwissen wurde ein Klassifikator für typische Zeitreihenverläufe realisiert. Für zu identifizierende Zeitreihen können mit seiner Hilfe Zugehörigkeitswerte zu den einzelnen Klassen berechnet werden, die anschließend eine praktikable Hilfe für weitergehende Entscheidungen liefern. Der vorgestellte relativ einfache Klassifikator liefert erstaunlich gute Ergebnisse, insbesondere bei der Verwendung possibilistischer Zugehörigkeitsfunktionen, die denen eines menschlichen Auswerters entsprechen.

Die hier vorgestellten Lösungen für die Beispiele sind nicht nur auf diese beschränkt, sondern können auch auf andere Beispiele mit abweichenden Ausdehnungen und anderen Dimensionen übertragen werden. Gedanken dazu sind in den entsprechenden Abschnitten der Arbeit vorgestellt worden.

Weitergehender Forschungsbedarf wird insbesondere für die erste Aufgabenstellung, dem Auffinden von Punktgruppen mit einem gleichförmigen Bewegungsverhalten, gesehen. Um das Verfahren für beliebige Aufgabenstellungen vollständig zu automatisieren, wird ein universelles Clusterverfahren benötigt, welches auch bei Vorliegen von Clustern unterschiedlicher Größe und Dichte robuste sinnvolle Ergebnisse liefert. Aufbauend auf die in dieser Arbeit mit dem Complete Linkage und dem Average

Linkage Verfahren gemachten Erfahrungen, sollten weitere Beispiele und Clusterverfahren untersucht werden. Weiterhin wird ein robustes Verfahren zur Bestimmung von „optimalen“ Klassenanzahlen benötigt, ohne dass diese, vor Einsatz des Klassifikators, fachkundig interpretiert werden müssen.

Die relativ junge Fuzzy-Theorie hat in den letzten Jahren in weiten Teilen der Praxis Einzug gehalten. Insbesondere in der Regelungstechnik hat sie sich etabliert, da mit ihr auch komplexe Regelungsaufgaben relativ einfach gelöst werden können. Kritisiert wird an ihr insbesondere, dass ihre Theorie, im Gegensatz zur Stochastik, nicht als einheitliches Gedankengebäude darstellbar ist (MAMMITZSCH, 1999) und bei ihrer Anwendung viele Parameter frei gewählt werden können, die gravierende Auswirkungen auf die Ergebnisse haben, z.B. die Formen der Zugehörigkeitsfunktionen, Parameter der Operationen von unscharfen Mengen, Schwellwerte zur Defuzzifizierung, Unbestreitbar sind allerdings die Erfolge in ihrer Anwendung.

Kapitel 8

Literaturverzeichnis

- Albertz, J. (2007):** Einführung in die Fernerkundung. Wissenschaftliche Buchgesellschaft Darmstadt, 2007.
- Anders, F. (2007):** Mustererkennung in Straßennetzwerken - Verfahren zur Interpretation von Vektordaten. Wissenschaftliche Arbeit der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover, Nr. 266, Hannover, 2007.
- Anders, K.-H. (2004):** Parameterfreies hierarchisches Graph-Clustering-Verfahren zur Interpretation raumbezogener Daten. Dissertation, Universität Stuttgart 2004.
Verfügbar unter der URL: <http://elib.uni-stuttgart.de/opus/volltexte/2004/2024/>
Letzter Zugriff am 19.1.2011
- Bacher, J. (1996):** Clusteranalyse. Oldenbourg Verlag, München, 1996.
- Backhaus, K., Erichson, B., Plinke, W., Weiber, R. (2006):** Multivariate Analysemethoden. 11. Auflage, Springer Verlag, Berlin, 2006.
- Ball, G. H., Hall, D. J. (1967):** A Clustering Technique for Summarizing Multivariate Data. In Behav. Sci. 12, 1967, pp. 153 - 155.
- Bandemer, H. (1997):** Ratschläge zum mathematischen Umgang mit Ungewissheit - Reasonable Computing. B.G. Teubner Verlag, Leipzig, 1997.
- Bankhofer, U., Vogel, J. (2008):** Datenanalyse und Statistik. Gabler Verlag, Wiesbaden, 2008.
- Bergs, S. (1981):** Optimalität bei Clusteranalysen. Dissertation im Fachbereich Wirtschafts- und Sozialwissenschaften der Westfälischen Wilhelms-Universität Münster, Münster, 1981.
- Bezdek, J. C. (1973a):** Fuzzy Mathematics in Pattern Classification. Cornell University, Ph.D. Thesis, Ann Arbor, 1973.
- Bezdek, J. C. (1973b):** Cluster validity with fuzzy sets. Cybernetics, Vol. 3, pp. 58 - 73, 1973.
- Bezdek, J. C. (1975):** Mathematical models for systematics and taxonomy. In: Estabrook, G., ed., Proceedings 8th Int. Conference on Numerical Taxonomy, Freeman Co., San Francisco, 1975.
- Bezdek, J. C. (1980):** A Convergence Theorem for the Fuzzy ISODATA Clustering Algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence 2, 1980.
- Bezdek, J. C., Pal, S. K. (1992):** Fuzzy Models for Pattern Recognition. IEEE Press, New York, 1992.
- Bezdek, J. C., Keller, J. M., Krishnapuram, R., Pal, S. K. (2005):** Fuzzy Models and Algorithms for Pattern Recognition and Image Processing. Springer Science and Business Media, New York, 2005.

- Bieronski, R., Stolp, T. (2010):** Bewegungs- und Deformationsmonitoring im Informationssystem GKSPRO[®] beim Bau des Jagdbergtunnels. In: Schwarz, W. (Red.) Interdisziplinäre Messaufgaben im Bauwesen - Weimar 2010. Schriftenreihe des Deutschen Vereins für Vermessungswesen, Bd. 62, Wißner Verlag, Augsburg, 2010.
- Biewer, B. (1997):** Fuzzy-Methoden. Springer Verlag, Berlin, 1997.
- Bock, H. H. (1974):** Automatische Klassifikation. Vandenhoeck und Ruprecht Verlag, Göttingen, 1974.
- Bock, H. H. (1994):** Classification and Clustering: Problems for the Future. In: E. Diday, Y. Lechevallier, M. Schrader, P. Bertrand, B. Burtshy (eds.), New Approaches in Classification and Data Analysis. Springer Verlag, Berlin, 1994.
- Bocklisch, S. F. (1981):** Experimentelle Prozessanalyse mit unscharfer Klassifikation. Dissertation B, TH Karl-Marx-Stadt, Fakultät Elektroingenieurwesen, 1981.
- Bocklisch, S. F. (1987):** Prozessanalyse mit unscharfen Verfahren. Verlag Technik, Berlin, 1987.
- Bocklisch, S. F., Bitterlich, N. (1993):** Fuzzy Pattern Classification. Spektrum der Wissenschaft, März 1993, S. 99 - 101.
- Boljen, J. (1983):** Ein dynamisches Modell zur Analyse und Interpretation von Deformationen. Wissenschaftliche Arbeit der Fachrichtung Vermessungswesen der Universität Hannover, Nr. 122, Hannover, 1983.
- Brauer, H. (1986):** Clusteranalytische Methoden zur Strukturierung einer städtischen Bodenpreissammlung. Veröffentlichung des Geodätischen Instituts der Rheinisch-Westfälischen Technischen Hochschule Aachen Nr. 39, Aachen, 1986.
- Burmeister, J. (1993):** Klassifikation via Fuzzy. Zeitschrift Design & Elektronik 1/2, Markt & Technik Verlag, Haar bei München, 1993.
- Burmeister, J. (1997):** Tutorial Grundlagen der Fuzzy-Pattern-Klassifikation. TAT Transfer von Automatisierungstechnologien GmbH, Chemnitz, Berlin, 1997.
- Cantor, G. (1932):** Gesammelte Abhandlungen mathematischen und philosophischen Inhalts. Springer Verlag, Berlin, 1932 (Neudruck, 1980).
- Caspary, W. (1988):** Fehlerverteilungen, Methode der kleinsten Quadrate und robuste Alternativen. Zeitschrift für Vermessungswesen, 3/1988, S. 123 - 133.
- Chatfield, C. (2004):** The Analysis of Time Series. Chapman & Hall, Boca Raton, 2004.
- Chmelina, K. (2003):** Wissensbasierte Analyse von Verschiebungsdaten im Tunnelbau. Veröffentlichung des Instituts für Geodäsie und Geophysik der TU Wien, Heft Nr. 62, Wien, 2003.
- Christophe, S., Ruas, A. (2002):** Detecting Building Alignments for Generalisation Purposes. In: Symposium on Geospatial Theory - Processing and Applications, Ottawa, 2002.
Verfügbar unter der URL: <http://www.isprs.org/proceedings/XXXIV/part4/paper.html>
Letzter Zugriff am 19.1.2011
- Deichsel, G. (1985):** Clusteranalyse und Diskriminanzanalyse. Gustav Fischer Verlag, Stuttgart, 1985.
- Deimer, R. (1986):** Unscharfe Clusteranalysemethoden. Schulz-Kirchner Verlag, Idstein, 1986.
- Deumlich, F., Staiger, R. (2002):** Instrumentenkunde der Vermessungstechnik. Wichmann Verlag, Heidelberg, 2002.
- Devijer, P., Kittler, J. (1982):** Pattern Recognition: A Statistical Approach. Prentice-Hall, Englewood Cliffs, NJ, 1982.
- DIN - Deutsches Institut für Normung e.V. (1998):** Vermessungswesen. DIN-Taschenbuch 111, Beuth Verlag, Berlin, 1998.

- DIN - Deutsches Institut für Normung e.V. (2009):** DIN 18710 Teil 4 (Entwurf): Überwachung, Beuth Verlag, Berlin, 2009.
- Dubois, D., Prade, H. (1982):** A class of fuzzy measures based on triangular norms. *Journal of general Systems*, 8, 43-6, 1982.
- Duda, R. O., Hart, P. E., Stork, D. G. (2000):** *Pattern Classification*. 2. Edition, Wiley Interscience, New York, 2000.
- Eckes, T., Roßbach, H. (1980):** *Clusteranalysen*. Kohlhammer Verlag, Stuttgart, 1980.
- Einarsson, P., Sæmundsson, K. (1987):** Earthquake epicenters 1982 – 1985 and volcanic systems in Iceland (map). In *Í Hlutarsins Eðli: Festschrift for Þorbjörn Sigurgeirsson*, edited by T. Sigfússon, Menningarsjóður, Reykjavik, 1987.
- Fahrmeir, L., Künstler, R., Pigeot, I., Tutz, G. (2007):** *Statistik*. 6. Auflage, Springer Verlag, Berlin, 2007.
- Feigl, K. L., Gasperi, J., Sigmundsson, F., Rigo, A. (2000):** Crustal deformation near Hengill volcano, Iceland 1993-1998: Coupling between magmatic activity and faulting inferred from elastic modeling of satellite radar interferograms. *Journal of Geophysical Research*, Vol. 105, No. B11, S. 25,655-25,670, November 10, 2000.
- Fletling, R. (2007):** Fuzzy-Clustering zur Analyse von Überwachungsmessungen. In: Brunner, F.K. (Hrsg.): *Ingenieurvermessung 07, Beiträge zum 15. Internationalen Ingenieurvermessungskurs*, Wichmann Verlag, Heidelberg, 2007.
- Fletling, R. (2008):** Application of Fuzzy Clustering in Deformation Analysis. In: Reiterer, A., Egly, U. (Hrsg.): *First Workshop on Applications of Artificial Intelligence in Engineering Geodesy (AIEG 2008)*, Vienna University of Technologie, Wien, 2008.
- Förstner, W., Werner, H. (1986):** Programmgesteuerte Deformationsanalyse. *Zeitschrift für Vermessungswesen*, 6/1986, S. 236 - 247.
- Gath, I., Geva, A. B. (1989):** Unsupervised Optimal Fuzzy Clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11, 1989.
- Gerke, K. (1966):** Deutsche Geo-Forschungsarbeiten auf Island 1964/67. *Polarforschung*, 36 (1/2), S. 99 - 102, 1966.
- Gottwald, S. (1989):** *Mehrwertige Logik*. Akademie Verlag, Berlin, 1989.
- Graeff, B. (2002):** Abfragesprache für geometrische und semantische Information aus rasterbasierten topographischen Karten. Institut für Geodäsie und Photogrammetrie der Eidgenössisch Technischen Hochschule Zürich, Mitteilungen Nr. 77, Zürich, 2002.
- Großauer, K., Lenz, G. (2007):** Is it Possible to Automate the Interpretation of Displacement Monitoring Data? *Felsbau* 25 (2007), Nr. 5, S. 99 - 106.
- Guder, W., Dahmen, W. (2004):** Kontinuierliche Überwachung der Standsicherheit von Tagebaurandböschungen am Beispiel des Tagebaus Hambach. In: Schwarz, W. (Red.) *Interdisziplinäre Messaufgaben im Bauwesen - Weimar 2004*. Schriftenreihe des Deutschen Vereins für Vermessungswesen, Bd. 46, Wißner Verlag, Augsburg, 2004.
- Gustafson, E. E., Kessel, W. C. (1979):** Fuzzy Clustering with a Fuzzy Covariance Matrix. IEEE CDC, San Diego, Kalifornien, 1979.
- Haberler, M. (2005):** Einsatz von Fuzzy Methoden zur Detektion konsistenter Punktbewegungen. Veröffentlichung des Instituts für Geodäsie und Geophysik der TU Wien, Heft Nr. 73, Wien, 2005.
- Hamacher, H. (1978):** Über logische Aggregation nicht-binär expliziter Entscheidungskriterien. G. Fischer Verlag, Frankfurt (Main), 1978.

- Heine, K. (1999):** Beschreibung von Deformationsprozessen durch Volterra- und Fuzzy-Modelle sowie Neuronale Netze. Deutsche Geodätische Kommission, Reihe C, Heft Nr. 516, München, 1999.
- Heinert, M., Ritter, B., Niemeier, W. (2004):** Angepasste Methoden der Deformationsanalyse für die Geodätischen Messungen in Südwestisland. Zeitschrift für Vermessungswesen, 6/2004, S. 399 - 406.
- Heinert, M. (2008):** Systemanalyse der seismisch bedingten Kinematik Islands. Geodätische Schriftenreihe der Technischen Universität Braunschweig, Nr. 22, Braunschweig, 2008.
- Heinert, M. (2010):** Persönliche Mitteilungen.
- Heinz, D., Pönitz, S. (2006):** Bewegungs- und Deformationsmonitoring am Projekt City-Tunnel Leipzig. In : Schwarz, W. (Red.) Interdisziplinäre Messaufgaben im Bauwesen - Weimar 2006. Schriftenreihe des Deutschen Vereins für Vermessungswesen, Bd. 50, Wifner Verlag, Augsburg, 2006.
- Heipke, C. (2003):** Photogrammetrie & Fernerkundung - vom Elektronenmikroskop bis zur Planetenbeobachtung. Photogrammetrie Fernerkundung Geoinformation, Organ der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation, 3/2003, S. 165 - 180.
- Heipke, C. (2008):** Editorial: 10 Jahre IPI unter neuer Leitung. Photogrammetrie Fernerkundung Geoinformation, Organ der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation, 5/2008, S. 313 - 318.
- Hempel, A.-J. (2005):** Aggregation und Identifikation von Fuzzy-Objekten mit unterschiedlicher elementarer Unschärfe. Diplomarbeit (unveröffentlicht), TU Chemnitz, Fakultät für Elektrotechnik, 2005.
- Hempel, A.-J., Bocklisch, S. F. (2009):** Parametric Fuzzy Modelling Framework for Complex Data-Inherent Structures. In: Carvalho, J. P., Dubois, D., Kaymak, U., Sousa J. M. C., Proceedings of the Joint 2009 International Fuzzy Systems Association World Congress and 2009 European Society of Fuzzy Logic and Technology Conference, Lisbon, Portugal, July 20-24, 2009. Verfügbar unter der URL: <http://archiv.tu-chemnitz.de/pub/2009/0148>
Letzter Zugriff am 19.1.2011
- Hempel, A.-J., Bocklisch, S. F. (2010):** Fuzzy Pattern Modelling of Data Inherent Structures Based on Aggregation of Data with heterogeneous Fuzziness. In: Gregorio Romero Rey and Luisa Martinez Muneta (Ed.), Modelling Simulation and Optimization. ISBN: 978-953-307-048-3, INTECH. Verfügbar unter der URL:
<http://sciyo.com/articles/show/title/fuzzy-pattern-modelling-of-data-inherent-structures-based-on-aggregation-of-data-with-heterogeneous->
Letzter Zugriff am 19.1.2011
- Hempel, A.-J. (2010):** Persönliche Mitteilungen.
- Herbst, G. (2010):** Persönliche Mitteilungen.
- Heunecke, O. (1995):** Zur Identifikation und Verifikation von Deformationsprozessen mittels adaptiver KALMAN-Filterung (Hannoversches Filter). Wissenschaftliche Arbeit der Fachrichtung Vermessungswesen der Universität Hannover, Nr. 208, Hannover, 1995.
- Heunecke, O., Pelzer, H., Welsch, W. (1998):** On the Classification of Deformation Models and Identification Methods in Engineering Surveying. XXI. International FIG Congress, Commission 6, Brighton, 1998.
- Heunecke, O. (2008):** Geosensornetze im Umfeld der Ingenieurvermessung. Forum, Zeitschrift des Bundes der öffentlich bestellten Vermessungsingenieure e.V., 2/2008, S. 357 - 364.
- Hobrough, G. L. (1959):** Automatic Stereo Plotting. Phot. Eng. (763 -769), 1959.
- Höppner, F., Klawonn, F., Kruse, R. (1997):** Fuzzy-Clusteranalyse. Vieweg Verlag, Braunschweig, 1997.

- Höppner, F., Klawonn, F., Kruse, R., Runkler, T. (1999):** Fuzzy Cluster Analysis. Wiley& Sons, Chichester, 1999.
- Jaanineh, G., Maijohann, M. (1996):** Fuzzy-Logik und Fuzzy-Control. Vogel Verlag, Würzburg, 1996.
- Jain, A. K., Topchy, A., Law, M. H. C., Buhmann, J. M. (2004):** Landscape of Clustering Algorithms. Proceedings of the 17th International Conference on Pattern Recognition, Cambridge UK, 2004, pp. 49 - 56.
- Kahlert, J., Frank, H. (1994):** Fuzzy-Logik und Fuzzy-Control. 2. Auflage, Vieweg Verlag, Braunschweig, 1994.
- Kahmen, H. (2006):** Vermessungskunde. 20. Auflage, De Gruyter Verlag, Berlin, New York, 2006.
- Kern, F. (2003):** Automatisierte Modellierung von Bauwerksgeometrien aus 3D-Laserscanner-Daten. Geodätische Schriftenreihe der Technischen Universität Braunschweig, Nr. 19, Braunschweig, 2003.
- Kolb, H.-J. (1996):** Signalverarbeitung und neuronale Netze in der Qualitätssicherung und Maschinendiagnose. In: Bocklisch, S. F., Haas, U. L., Bitterlich, N., Protzel, P. (Hrsg.): Fuzzy Technologien und Neuronale Netze in der Praxis. Tagungsband des 10. Chemnitzer Kolloquiums 29./30. November 1995, Shaker Verlag, Aachen, 1996 .
- Konecny, G. (1984):** Photogrammetrie. 4. Auflage, De Gruyter Verlag, Berlin, New York, 1984.
- Kraus, K. (1979):** Zur Theorie der Klassifizierung multispektraler Bilder, Bildmessung und Luftbildwesen 4/1979, Organ der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung, S. 119 - 128.
- Krishnapuram, R., Keller, J. M. (1993):** A possibilistic approach to clustering. IEEE Transactions on Fuzzy Systems, 1(2), S. 98 - 110, 1993.
- Kuhlmann, H. (1996):** Ein Beitrag zur Überwachung von Brückenbauwerken mit kontinuierlich registrierten Messungen. Wissenschaftliche Arbeit der Fachrichtung Vermessungswesen der Universität Hannover, Nr. 218, Hannover, 1996.
- Kutterer, H. (2002):** Zum Umgang mit Ungewissheit in der Geodäsie. Deutsche Geodätische Kommission, Reihe C, Heft Nr. 553, München, 2002.
- Lange, N. de (2002):** Geoinformatik in Theorie und Praxis. Springer Verlag, Berlin, 2002.
- Lehmkuhler, H. (1993):** Die geodätische Deformationsanalyse als Mustererkennungsaufgabe. Veröffentlichung des Geodätischen Instituts der RWTH Aachen, Nr. 49, 1993.
- Leitloff, J., Hinz, S., Stilla, U. (2006):** Detection of Vehicle Queues in QuickBird Imagery of City Areas. Photogrammetrie Fernerkundung Geoinformation, Organ der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation, 4/2006, S. 315-325.
- Lindner, W. (2007):** Digitale Photogrammetrie am PC - Aktuelle Entwicklungen und Anwendungen in den Geowissenschaften. Habilitationsschrift an der Mathematisch- Naturwissenschaftlichen Fakultät der Heinrich Heine Universität Düsseldorf, 2007.
- Luhmann, T. (2003):** Nahbereichsphotogrammetrie. 2. Auflage, Wichmann Verlag, Heidelberg, 2003.
- Mahalanobis, P. C. (1936)** On the generalised distance in statistics, Proceedings of the National Institute of Science of India, 12, 1936, S. 49 - 55.
- Mammitzsch, V. (1999):** Fuzzy Theorie als Alternative zur Stochastik - Was heißt hier: Eine Alternative? In: Seising, R. (Hrsg.): Fuzzy Theorie und Stochastik, Vieweg Verlag, Braunschweig, 1999.
- Mann, S. (1984):** Ein Lernverfahren zur Modellierung zeitvarianter Systeme mittels unscharfer Klassifikation. Dissertation A, TH Karl-Marx-Stadt, Fakultät Elektroingenieurwesen, 1984.

- Miima, J.-B. (2002):** Artificial Neural Networks and Fuzzy Logic Techniques for the Reconstruction of Structural Deformations. Geodätische Schriftenreihe der TU Braunschweig, Band 18, Braunschweig, 2002.
- MIT GmbH (2002):** Data Engine 4.0 Benutzerhandbuch. Band: Tutorials und Grundlagen, MIT-Management Intelligenter Technologien GmbH, Aachen, 2002.
- Nazif, A., Levine, M. D. (1984):** Low Level Image Segmentation: An Expert System. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, No. 5, 9/1984.
- Niemann, H. (1974):** Methoden der Mustererkennung. Akademische Verlagsgesellschaft, Frankfurt (Main), 1974.
- Niemann, H. (1981):** Pattern Analysis. Springer Verlag, Berlin, 1981.
- Niemann, H. (2003):** Klassifikation von Mustern. (2. überarbeitete Auflage im Internet). Verfügbar unter der URL:
<http://www5.informatik.uni-erlangen.de/fileadmin/Persons/NiemannHeinrich/klassifikation-von-mustern/m00-www.pdf>
Letzter Zugriff am 19.1.2011
- Niemeier, W. (1985a):** Anlage von Überwachungsnetzen. In: Pelzer (Hrsg.): Geodätische Netze in Landes- und Ingenieurvermessung II, Wittwer Verlag, Stuttgart, 1985.
- Niemeier, W. (1985b):** Deformationsanalyse. In: Pelzer (Hrsg.): Geodätische Netze in Landes- und Ingenieurvermessung II, Wittwer Verlag, Stuttgart, 1985.
- Niemeier, W. (1988):** Deformationsanalyse - Aktueller Stand in Theorie und Praxis. In: Schnädelbach, K., Ebner, H. (Hrsg.): Ingenieurvermessung 88, Beiträge zum X. Internationalen Kurs für Ingenieurvermessung, Band 2, Dümmler Verlag, Bonn, 1988.
- Niemeier, W. (2008):** Ausgleichsrechnung. 2. Auflage, De Gruyter Verlag, Berlin, New York, 2008.
- Otte, C. (1998):** Fuzzy-Prototyp-Klassifikatoren und deren Anwendung zur automatischen Merkmalsselektion. Dissertation, Universität Oldenburg, Fachbereich für Informatik, 1998.
- Päßler, M. (1998):** Mehrdimensionale Zeitreihenmodellierung und -prognose mittels Fuzzy Pattern Modellen. Diplomarbeit, TU Chemnitz, Fakultät für Mathematik, 1998.
Verfügbar unter der URL: <http://www.familiepaessler.de/>
Letzter Zugriff am 19.1.2011
- Pelzer, H. (1985):** Statische, kinematische und dynamische Punktfelder. In: Pelzer (Hrsg.): Geodätische Netze in Landes- und Ingenieurvermessung II, Wittwer Verlag, Stuttgart, 1985.
- Pelzer, H. (Hrsg.) (1987):** Ingenieurvermessung: Deformationsmessungen; Massenberechnung; Ergebnisse des Arbeitskreises 6 des Deutschen Vereins für Vermessungswesen (DVW) e.V., Wittwer Verlag, Stuttgart, 1987.
- Perlt, J. (2006):** Ein geokinematisches Modell für Island. Geodätische Schriftenreihe der Technischen Universität Braunschweig, Nr. 20, Braunschweig, 2006.
- Pfeufer, A. (1988):** Beitrag zur Identifikation und Modellierung von Deformationsprozessen. Dissertation der Fakultät für Bau-, Wasser- und Forstwesen der Technischen Universität Dresden, 1988.
- Plümer, L., Schmittwilken, J., Dörschlag, D., Krückhans, M. (2009):** Grammatikbasierte Fassadenrekonstruktion. In: Godhoff, F., Staiger, R. (Red.) Terrestrisches Laserscanning (TLS 2009). Yes, we Scan! Beiträge zum 91. DVW-Seminar am 19. und 20. November 2009 in Fulda. Schriftenreihe des Deutschen Vereins für Vermessungswesen, Bd. 60, Wißner Verlag, Augsburg, 2009.
- Qu, W. (2000):** Zur Anwendung der Fuzzy-Clusteranalyse in der Grundstückswertermittlung. Wissenschaftliche Arbeit der Fachrichtung Vermessungswesen der Universität Hannover, Nr. 240, Hannover, 2000.

- Reinhardt u.a. (2004):** Raumbezogene Informationssysteme. In: Möser, M., Müller, G., Schlemmer, H., Werner, H. (Hrsg.): Handbuch Ingenieurgeodäsie, Wichmann Verlag, Heidelberg, 2004.
- Rennen, M. (2009):** Persönliche Mitteilungen.
- Rissanen, J. (1983):** A universal prior for integers and estimation by minimum description length. The Annals of Statistics, Vol. 11, No. 2, pp 416 - 431, 1983.
- Rumpler, T. A. (1999):** Probabilistische und Fuzzy Methoden für die Clusteranalyse. In: Seising, R. (Hrsg.): Fuzzy Theorie und Stochastik, Vieweg Verlag, Braunschweig, 1999.
- Scheunert, U. (2002):** Fuzzy-Mengen-Verknüpfung und Fuzzy-Arithmetik zur Sensor-Daten-Fusion. Fortschritt-Berichte VDI, Reihe 8, Nr. 941, VDI Verlag, Düsseldorf, 2002.
- Schleinkofer, M. (2007):** Wissensbasierte Unterstützung zur Erstellung von Produktmodellen im Baubestand. Dissertation an der Technischen Universität München, Lehrstuhl für Bauinformatik, 2007.
- Sester, M., Haurert, J.-H., Anders, K.-H. (2008):** Modell- und kartographische Generalisierung von topographischen und thematischen Informationen. Kartographische Nachrichten, 6/2008, S. 307 - 314.
- Shannon, C. E. (1948):** A Mathematical Theory of Communication. Bell System Technical Journal, Vol. 27, pp 379 - 423, 1948.
- Steinhausen, D., Langer, K. (1977):** Clusteranalyse. Verlag de Gruyter, Berlin, 1977.
- Stempfhuber, W. (2009):** Geodätische Monitoringsysteme - Stand der Technik und Abgrenzung der gegenwärtigen Systeme. In: Foppe, K., Knaak, L., Pawel, B. (Red.) Zeitabhängige Messgrößen - Verborgene Schätze in unseren Daten. Schriftenreihe des Deutschen Vereins für Vermessungswesen, Bd. 59, Wifner Verlag, Augsburg, 2009.
- Technische Universität Chemnitz (2010):** Fakultät für Elektrotechnik und Informationstechnik - Professur Systemtheorie. URL: <http://www.tu-chemnitz.de/etit/systh/>
Letzter Zugriff am 19.1.2011
- Tilli, T. (1993):** Mustererkennung mit Fuzzy-Logik. Franzis Verlag, München, 1993.
- Timm, H. (2002):** Fuzzy-Clusteranalyse: Methoden zur Exploration von Daten mit fehlenden Werten sowie klassifizierten Daten. Magdeburg, 2002.
Verfügbar unter der URL: <http://diglib.uni-magdeburg.de/Dissertationen/2002/heitimm.pdf>
Letzter Zugriff am 19.1.2011
- Valsson, G. Þ., Sigurdsson, Þ., Völksen, Chr., Rennen, M. (2007):**
ISNET 2004 Niðurstöður úr endurmælingum grunnstöðvanets Íslands. Landmælingar Íslands, Akranes, 2007.
- Villemin, Th., Geirsson, H., Sturkell, E., Jounanne, F. (2009):** Active deformation at Grímsvötn subglacial volcano: a composite evolution to be deciphered. Geophysical Research Abstracts, Vol. 11, EGU2009-11177, 2009.
- Welsch, W. (1982):** Einige Erweiterungen der Deformationsermittlung in geodätischen Netzen durch Methoden der Strainanalyse. In: Joo, I., Detreköi, A. (Editors): III. International Symposium on Deformation Measurements by Geodetic Methods. Budapest, 25.-27.08.1982. Proceedings, pp. 83-97. Akademiai Kiado, Budapest, 1983.
- Welsch, W. (Hrsg.) (1983):** Deformationsanalysen '83. Schriftenreihe Studiengang Vermessungswesen Hochschule der Bundeswehr München, Heft Nr. 9, München, 1983.
- Welsch, W., Heunecke, O., Kuhlmann, H. (2000):** Auswertung geodätischer Überwachungsmessungen. In: Möser, M., Müller, G., Schlemmer, H., Werner, H. (Hrsg.): Handbuch Ingenieurgeodäsie, Wichmann Verlag, Heidelberg, 2000.
- Wenderlein, W. (1998):** Die Nichtverlässlichkeit von Messungen - und wie wir damit umgehen. Allgemeine Vermessungsnachrichten, 1/1998, S. 1 - 4.

- Wieser, A. (2002):** Robust and Fuzzy Techniques for Parameter Estimation and Quality Assessment in GPS. Shaker Verlag, Aachen, 2002.
- Windham, M. P. (1981):** Cluster Validity for Fuzzy Clustering Algorithms. Fuzzy Sets and Systems, 5, 1981, pp. 177 - 185.
- Xie, X. L., Beni, G. (1991):** A Validity Measure for Fuzzy Clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 13(8), pp. 841 - 847, 1991.
- Yager, R. R. (1980):** On a General Class of Fuzzy Connectives. Fuzzy Sets and Systems, 4, 1980, pp. 221 - 229.
- Zadeh, L. (1965):** Fuzzy Sets. In: Information and Control, Vol. 8, S. 338 - 353, 1965 (Reprint in: BEZDEK, J. C., PAL, S. K., 1992).
- Zadeh, L. (1968):** Fuzzy algorithms. In: Information and Control, Vol. 12, S. 94 - 102, 1968.
- Zadeh, L. (1975):** The concept of a linguistic variable and its application to approximate reasoning. I, Information Science, Vol. 8, S. 199 - 250, 1975 .
- Zadeh, L. (1987):** In: Yager, R. R., Ovchinnikov, S., Tong, R. M., Nguyen, H. T. (Hrsg.) (1987): Fuzzy Sets and Applications. Selected Papers by L. A. Zadeh. Wiley, New York, 1987.
- Zimmermann, H.-J. (1991):** Fuzzy Set Theorie and its Applications (2nd ed.). Kluwer Academic Publishers, Boston, 1991.
- Zimmermann, H.-J. (Hrsg.) (1993):** Fuzzy Technologien. VDI Verlag, Düsseldorf, 1993.
- Zimmermann, H.-J. (1999):** Zur Modellierung von Unsicherheit realer Probleme. In: Seising, R. (Hrsg.): Fuzzy Theorie und Stochastik, Vieweg Verlag, Braunschweig, 1999.

Zur Berechnung der Beispiele wurde folgende Software benutzt:

- **Caplan V 2.7**, Cremer Programmentwicklung GmbH, Pfaffenhofen
- **Cluster Analysis Manager V 1.33**, TransferTech GmbH, Braunschweig
- **Data Engine V 4.0**, MIT-Management Intelligenter Technologien GmbH, Aachen
- **Fuzzy Toolbox**, TU Chemnitz, Professur für Systemtheorie, Chemnitz
- **Panda V 4.0**, GeoTec GmbH, Laatzen
- **Predictive Analytics Software (PASW) Statistics V 18**, SPSS Inc., Chicago

sowie diverse selbst erstellte Programme.

Anhang A

Tabellen

Beispiel Delft: Zugehörigkeitswerte Linien						
Linie	von	nach	Cluster			g
	Punkt 1	Punkt 2	1	2	3	
1	3	21	0,49	0,50	0,01	1
2	3	39	0,97	0,02	0,01	42
3	3	5	0,97	0,03	0,00	36
4	21	43	0,03	0,95	0,02	31
5	21	39	0,51	0,48	0,01	1
6	5	39	0,99	0,01	0,00	110
7	5	41	0,98	0,02	0,00	58
8	5	11	0,97	0,02	0,01	41
9	39	43	0,51	0,48	0,01	1
10	43	45	0,03	0,96	0,01	35
11	39	45	0,51	0,49	0,00	1
12	35	39	0,50	0,49	0,01	1
13	39	41	0,98	0,01	0,01	76
14	45	47	0,04	0,93	0,03	23
15	35	45	0,02	0,97	0,01	63
16	35	41	0,49	0,50	0,01	1
17	13	41	0,51	0,40	0,09	1
18	11	41	0,97	0,03	0,00	35
19	35	47	0,04	0,93	0,03	27
20	35	37	0,01	0,51	0,48	1
21	13	35	0,03	0,88	0,09	10
22	17	47	0,03	0,46	0,51	1
23	37	47	0,04	0,47	0,49	1
24	13	37	0,03	0,42	0,55	1
25	13	15	0,03	0,40	0,57	1
26	11	13	0,50	0,41	0,09	1
27	11	15	0,48	0,03	0,49	1
28	17	37	0,01	0,03	0,96	27
29	15	37	0,01	0,03	0,96	27
30	15	17	0,00	0,02	0,98	59

Tabelle A.1: Beispiel Delft. Zugehörigkeitswerte der Linien zu den drei Clustern, die aus den mit dem Gustafson-Kessel Algorithmus berechneten Zugehörigkeitswerten der Objektpunkte abgeleitet wurden. Die Zugehörigkeitswerte sind auf zwei Nachkommastellen und das Gütemaß g auf eine ganze Zahl gerundet.

Beispiel Delft: Zugehörigkeitswerte Linien						
Linie	von Punkt 1	nach Punkt 2	Cluster			g
			1	2	3	
1	3	21	0,50	0,50	0,00	1
2	3	39	1,00	0,00	0,00	*
3	3	5	1,00	0,00	0,00	*
4	21	43	0,00	1,00	0,00	*
5	21	39	0,50	0,50	0,00	1
6	5	39	1,00	0,00	0,00	*
7	5	41	1,00	0,00	0,00	*
8	5	11	1,00	0,00	0,00	*
9	39	43	0,50	0,50	0,00	1
10	43	45	0,00	1,00	0,00	*
11	39	45	0,50	0,50	0,00	1
12	35	39	0,50	0,50	0,00	1
13	39	41	1,00	0,00	0,00	*
14	45	47	0,00	1,00	0,00	*
15	35	45	0,00	1,00	0,00	*
16	35	41	0,50	0,50	0,00	1
17	13	41	0,50	0,50	0,00	1
18	11	41	1,00	0,00	0,00	*
19	35	47	0,00	1,00	0,00	*
20	35	37	0,00	0,50	0,50	1
21	13	35	0,00	1,00	0,00	*
22	17	47	0,00	0,50	0,50	1
23	37	47	0,00	0,50	0,50	1
24	13	37	0,00	0,50	0,50	1
25	13	15	0,00	0,50	0,50	1
26	11	13	0,50	0,50	0,00	1
27	11	15	0,50	0,00	0,50	1
28	17	37	0,00	0,00	1,00	*
29	15	37	0,00	0,00	1,00	*
30	15	17	0,00	0,00	1,00	*

Tabelle A.2: Beispiel Delft. Zugehörigkeitswerte der Linien zu den drei Clustern, die aus den mit dem Gath-Geva Algorithmus berechneten Zugehörigkeitswerten der Objektpunkte abgeleitet wurden. Die Zugehörigkeitswerte sind auf zwei Nachkommastellen gerundet. Das Zeichen * für das Gütemaß g kennzeichnet eine scharfe Zuordnung.

Beispiel Delft: Zugehörigkeitswerte Linien						
Linie	von Punkt 1	nach Punkt 2	Cluster			g
			1	2	3	
1	3	21	0,48	0,49	0,06	1
2	3	39	0,98	0,00	0,02	43
3	3	5	0,98	0,00	0,03	39
4	21	43	0,00	0,99	0,10	10
5	21	39	0,50	0,49	0,06	1
6	5	39	1,00	0,00	0,02	44
7	5	41	0,99	0,00	0,02	45
8	5	11	0,90	0,00	0,02	44
9	39	43	0,50	0,50	0,06	1
10	43	45	0,00	1,00	0,10	10
11	39	45	0,50	0,50	0,06	1
12	35	39	0,50	0,50	0,09	1
13	39	41	0,99	0,00	0,02	50
14	45	47	0,00	0,96	0,11	9
15	35	45	0,00	1,00	0,13	8
16	35	41	0,49	0,50	0,09	1
17	13	41	0,49	0,42	0,16	1
18	11	41	0,89	0,00	0,02	51
19	35	47	0,00	0,96	0,15	7
20	35	37	0,00	0,66	0,51	1
21	13	35	0,00	0,91	0,24	4
22	17	47	0,00	0,47	0,52	1
23	37	47	0,00	0,63	0,49	1
24	13	37	0,00	0,58	0,58	1
25	13	15	0,00	0,42	0,61	1
26	11	13	0,40	0,42	0,16	1
27	11	15	0,40	0,01	0,47	1
28	17	37	0,00	0,18	0,89	5
29	15	37	0,00	0,18	0,89	5
30	15	17	0,00	0,02	0,92	48

Tabelle A.3: Beispiel Delft. Zugehörigkeitswerte der Linien zu den drei Clustern, die aus den mit dem Verfahren der Fuzzy Pattern Klassifikation berechneten Zugehörigkeitswerten der Objektpunkte abgeleitet wurden. Die Zugehörigkeitswerte sind auf zwei Nachkommastellen und das Gütemaß g auf eine ganze Zahl gerundet.

Beispiel Delft: Zugehörigkeitswerte Linien						
Linie	von Punkt 1	nach Punkt 2	Cluster			g
			1	2	3	
1	3	21	0,56	0,56	0,06	1
2	3	39	0,93	0,13	0,02	7
3	3	5	0,91	0,15	0,03	6
4	21	43	0,23	0,96	0,10	4
5	21	39	0,61	0,53	0,06	1
6	5	39	0,96	0,12	0,02	8
7	5	41	0,93	0,11	0,02	8
8	5	11	0,87	0,10	0,02	9
9	39	43	0,59	0,53	0,06	1
10	43	45	0,22	0,96	0,10	4
11	39	45	0,61	0,53	0,06	1
12	35	39	0,58	0,55	0,09	1
13	39	41	0,95	0,09	0,02	10
14	45	47	0,22	0,95	0,11	4
15	35	45	0,20	0,98	0,13	5
16	35	41	0,55	0,54	0,09	1
17	13	41	0,53	0,46	0,16	1
18	11	41	0,86	0,07	0,02	12
19	35	47	0,19	0,96	0,15	5
20	35	37	0,13	0,66	0,51	1
21	13	35	0,16	0,91	0,24	4
22	17	47	0,14	0,60	0,52	1
23	37	47	0,15	0,63	0,49	1
24	13	37	0,11	0,58	0,58	1
25	13	15	0,11	0,58	0,61	1
26	11	13	0,47	0,44	0,16	1
27	11	15	0,45	0,20	0,47	1
28	17	37	0,09	0,30	0,89	3
29	15	37	0,09	0,33	0,89	3
30	15	17	0,09	0,30	0,92	3

Tabelle A.4: Beispiel Delft. Zugehörigkeitswerte der Linien zu den drei Clustern, die aus den mit dem Verfahren der Fuzzy Pattern Klassifikation berechneten Zugehörigkeitswerten der Objektpunkte abgeleitet wurden. Hier wurden im Gegensatz zu Tabelle A.3 symmetrische Zugehörigkeitsfunktionen verwendet. Die Zugehörigkeitswerte sind auf zwei Nachkommastellen und das Gütemaß g auf eine ganze Zahl gerundet.

Beispiel ISNET: Zugehörigkeitswerte Linien								
von	nach	Cluster						
Punkt 1	Punkt 2	1	2	3	4	5	6	<i>g</i>
LM0082	LM0301	0,01	0,00	0,00	0,08	0,08	1,00	13
LM0082	LM0303	0,05	0,48	0,01	0,09	0,05	0,50	1
LM0082	LM0304	0,03	0,50	0,01	0,11	0,07	0,50	1
LM0082	LM0305	0,01	0,02	0,01	0,49	0,07	0,50	1
LM0082	LM0309	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0082	LM0502	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0082	LM3713	0,01	0,00	0,00	0,08	0,07	1,00	13
LM0301	LM0302	0,31	0,01	0,00	0,06	0,06	0,50	2
LM0301	LM3713	0,01	0,00	0,00	0,08	0,08	1,00	12
LM0302	LM0303	0,35	0,49	0,01	0,07	0,04	0,00	1
LM0302	LM3713	0,31	0,01	0,00	0,06	0,05	0,50	2
LM0303	LM0304	0,07	0,98	0,01	0,12	0,06	0,00	8
LM0303	LM3713	0,05	0,48	0,01	0,09	0,06	0,50	1
LM0304	LM0305	0,04	0,51	0,01	0,52	0,08	0,00	1
LM0304	LM2003	0,06	0,99	0,02	0,13	0,05	0,00	8
LM0305	LM0306	0,01	0,02	0,01	0,48	0,07	0,50	1
LM0305	LM0309	0,01	0,02	0,01	0,48	0,07	0,50	1
LM0305	LM2003	0,04	0,51	0,02	0,51	0,06	0,00	1
LM0305	OS7225	0,02	0,02	0,01	0,86	0,09	0,00	9
LM0306	LM0307	0,01	0,00	0,00	0,08	0,08	0,99	12
LM0306	LM0308	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0306	LM0309	0,01	0,00	0,00	0,06	0,05	1,00	16
LM0306	OS7225	0,01	0,00	0,01	0,44	0,08	0,50	1
LM0307	LM0308	0,01	0,00	0,00	0,09	0,10	0,99	10
LM0307	LM0355	0,01	0,00	0,00	0,09	0,11	0,98	9
LM0307	OS7220	0,02	0,04	0,01	0,41	0,12	0,49	1
LM0307	OS7225	0,01	0,00	0,01	0,46	0,10	0,49	1
LM0307	OS7482	0,01	0,00	0,00	0,08	0,11	0,99	9
LM0308	LM0309	0,01	0,00	0,00	0,07	0,07	1,00	14
LM0308	LM0530	0,01	0,00	0,00	0,07	0,07	1,00	14
LM0308	OS7371	0,01	0,00	0,00	0,07	0,07	1,00	14
LM0308	OS7482	0,01	0,00	0,00	0,07	0,09	1,00	11
LM0309	LM0502	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0309	OS7371	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0310	LM0311	0,01	0,00	0,00	0,08	0,07	1,00	13
LM0310	LM0312	0,01	0,00	0,00	0,08	0,07	1,00	13
LM0310	LM0502	0,01	0,00	0,00	0,08	0,06	1,00	13
LM0310	OS7371	0,01	0,00	0,00	0,08	0,06	1,00	13
LM0311	LM0312	0,01	0,00	0,00	0,08	0,06	1,00	13
LM0311	LM0314	0,01	0,00	0,00	0,07	0,06	1,00	13

(Fortsetzung nächste Seite)

(Fortsetzung)

von Punkt 1	nach Punkt 2	Cluster						g
		1	2	3	4	5	6	
LM0311	LM0315	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0311	LM0521	0,01	0,00	0,00	0,07	0,07	1,00	14
LM0311	OS7371	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0312	LM0313	0,01	0,00	0,00	0,08	0,07	1,00	13
LM0312	LM0314	0,01	0,00	0,00	0,08	0,06	1,00	13
LM0312	OS7373	0,01	0,00	0,00	0,07	0,07	1,00	13
LM0313	LM0314	0,01	0,00	0,00	0,08	0,07	1,00	13
LM0313	OS7373	0,01	0,00	0,00	0,07	0,07	1,00	14
LM0314	LM0315	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0314	LM0322	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0314	LM0323	0,01	0,00	0,00	0,07	0,07	1,00	14
LM0315	LM0323	0,01	0,00	0,00	0,07	0,07	1,00	15
LM0315	LM0503	0,01	0,00	0,00	0,07	0,06	1,00	15
LM0315	LM0521	0,01	0,00	0,00	0,07	0,07	1,00	14
LM0315	LM1260	0,01	0,00	0,00	0,07	0,07	1,00	14
LM0316	LM0503	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0316	LM0521	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0316	LM0530	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0316	LM0534	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0316	OS7371	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0317	LM0318	0,01	0,00	0,00	0,07	0,06	1,00	15
LM0317	LM0321	0,01	0,00	0,00	0,07	0,06	1,00	15
LM0317	LM0322	0,01	0,00	0,00	0,07	0,06	1,00	15
LM0318	LM0319	0,01	0,00	0,00	0,06	0,06	1,00	16
LM0318	LM0320	0,01	0,00	0,00	0,07	0,06	1,00	15
LM0318	LM0321	0,01	0,00	0,00	0,07	0,06	1,00	15
LM0319	LM0320	0,01	0,00	0,00	0,07	0,06	1,00	15
LM0319	LM0524	0,01	0,00	0,00	0,06	0,06	1,00	16
LM0319	LM0525	0,01	0,00	0,00	0,06	0,06	1,00	15
LM0319	LM0537	0,01	0,00	0,00	0,06	0,06	1,00	16
LM0319	LM0538	0,01	0,00	0,00	0,06	0,06	1,00	16
LM0320	LM0321	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0320	LM0524	0,01	0,00	0,00	0,07	0,06	1,00	15
LM0321	LM0322	0,01	0,00	0,00	0,07	0,06	1,00	15
LM0321	LM0323	0,01	0,00	0,00	0,07	0,07	1,00	15
LM0321	LM0524	0,01	0,00	0,00	0,06	0,06	1,00	15
LM0322	LM0323	0,01	0,00	0,00	0,07	0,07	1,00	15
LM0323	LM0324	0,01	0,00	0,00	0,07	0,07	1,00	15
LM0323	LM0524	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0323	LM1260	0,01	0,00	0,00	0,07	0,07	1,00	15
LM0324	LM0325	0,01	0,00	0,00	0,06	0,07	1,00	15

(Fortsetzung nächste Seite)

(Fortsetzung)

von Punkt 1	nach Punkt 2	Cluster						g
		1	2	3	4	5	6	
LM0324	LM0524	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0324	LM0525	0,01	0,00	0,00	0,07	0,07	1,00	15
LM0324	LM1260	0,01	0,00	0,00	0,07	0,07	1,00	14
LM0325	LM0505	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0325	LM0506	0,01	0,00	0,00	0,06	0,07	1,00	14
LM0325	LM0525	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0325	LM0535	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0325	LM1260	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0326	LM0327	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0326	LM0504	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0326	LM0505	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0326	LM0508	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0326	LM0528	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0326	LM0534	0,01	0,00	0,00	0,07	0,06	1,00	15
LM0326	OS7378	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0327	LM0329	0,01	0,00	0,00	0,06	0,06	1,00	16
LM0327	LM0330	0,01	0,00	0,00	0,06	0,07	1,00	14
LM0327	LM0505	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0327	LM0508	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0327	LM0509	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0327	AKUR	0,01	0,00	0,00	0,07	0,06	0,98	13
LM0328	LM0508	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0328	OS5861	0,01	0,00	0,00	0,04	0,06	0,91	16
LM0328	OS7379	0,01	0,00	0,00	0,05	0,07	0,96	15
LM0328	OS7469	0,01	0,00	0,00	0,05	0,06	0,97	15
LM0328	OS7494	0,01	0,00	0,00	0,05	0,07	1,00	15
LM0328	AKUR	0,01	0,00	0,00	0,07	0,06	0,98	14
LM0329	LM0330	0,01	0,00	0,00	0,06	0,06	1,00	16
LM0329	RH8701	0,01	0,00	0,00	0,05	0,06	1,00	16
LM0329	AKUR	0,01	0,00	0,00	0,07	0,06	0,98	14
LM0330	LM0331	0,44	0,00	0,00	0,04	0,05	0,50	1
LM0330	LM0509	0,01	0,00	0,00	0,06	0,07	1,00	14
LM0330	RH8701	0,01	0,00	0,00	0,05	0,07	1,00	15
LM0331	OS7031	0,44	0,00	0,00	0,03	0,04	0,46	1
LM0331	OS7367	0,46	0,50	0,01	0,12	0,04	0,00	1
LM0331	RH8701	0,44	0,00	0,00	0,04	0,04	0,50	1
LM0332	LM0333	0,04	1,00	0,03	0,18	0,05	0,00	5
LM0332	OS7227	0,05	1,00	0,02	0,18	0,05	0,00	6
LM0332	OS7367	0,04	1,00	0,02	0,20	0,06	0,00	5
LM0333	LM0334	0,04	1,00	0,03	0,18	0,04	0,00	6
LM0333	LM0511	0,04	1,00	0,03	0,18	0,04	0,00	6

(Fortsetzung nächste Seite)

(Fortsetzung)

von Punkt 1	nach Punkt 2	Cluster						g
		1	2	3	4	5	6	
LM0333	OS7227	0,05	1,00	0,02	0,17	0,05	0,00	6
LM0334	LM0335	0,04	1,00	0,03	0,18	0,04	0,00	6
LM0334	LM0336	0,04	1,00	0,03	0,18	0,04	0,00	6
LM0334	LM0337	0,04	1,00	0,03	0,18	0,04	0,00	6
LM0334	LM0511	0,04	1,00	0,04	0,18	0,04	0,00	5
LM0335	LM0336	0,04	1,00	0,03	0,17	0,04	0,00	6
LM0335	LM0511	0,04	1,00	0,03	0,18	0,04	0,00	6
LM0335	LM0514	0,04	1,00	0,03	0,17	0,04	0,00	6
LM0336	LM0337	0,04	1,00	0,03	0,17	0,04	0,00	6
LM0336	LM0338	0,04	1,00	0,03	0,17	0,04	0,00	6
LM0336	LM0339	0,04	1,00	0,03	0,17	0,04	0,00	6
LM0336	LM0340	0,04	1,00	0,03	0,18	0,04	0,00	6
LM0336	LM0514	0,04	1,00	0,03	0,17	0,04	0,00	6
LM0337	LM0338	0,04	1,00	0,03	0,17	0,04	0,00	6
LM0338	LM0339	0,04	1,00	0,03	0,17	0,04	0,00	6
LM0338	LM0344	0,04	1,00	0,03	0,18	0,04	0,00	6
LM0339	LM0340	0,04	1,00	0,03	0,18	0,05	0,00	6
LM0339	LM0341	0,05	1,00	0,03	0,16	0,04	0,00	6
LM0339	LM0343	0,04	1,00	0,03	0,17	0,05	0,00	6
LM0339	LM0344	0,04	1,00	0,03	0,18	0,05	0,00	6
LM0340	LM0341	0,05	1,00	0,03	0,16	0,04	0,00	6
LM0340	LM0514	0,04	1,00	0,03	0,18	0,04	0,00	6
LM0340	OS7012	0,04	1,00	0,03	0,17	0,04	0,00	6
LM0341	LM0342	0,05	1,00	0,03	0,15	0,04	0,00	7
LM0341	LM0343	0,05	1,00	0,03	0,16	0,05	0,00	6
LM0341	OS7012	0,05	1,00	0,03	0,15	0,04	0,00	7
LM0342	LM0343	0,05	1,00	0,03	0,16	0,04	0,00	6
LM0342	LM0518	0,04	1,00	0,03	0,16	0,04	0,00	6
LM0342	OS5491	0,05	1,00	0,03	0,15	0,04	0,00	7
LM0342	OS7012	0,05	1,00	0,03	0,15	0,04	0,00	7
LM0342	OS7461	0,04	1,00	0,04	0,16	0,04	0,00	6
LM0343	LM0344	0,04	1,00	0,03	0,17	0,05	0,00	6
LM0343	LM0345	0,04	1,00	0,03	0,18	0,05	0,00	6
LM0343	LM0346	0,04	1,00	0,03	0,17	0,04	0,00	6
LM0343	OS7461	0,04	1,00	0,03	0,17	0,04	0,00	6
LM0344	LM0345	0,04	1,00	0,03	0,18	0,05	0,00	6
LM0345	LM0346	0,04	1,00	0,03	0,18	0,04	0,00	6
LM0346	LM0347	0,04	1,00	0,04	0,17	0,04	0,00	6
LM0346	OS7461	0,04	1,00	0,04	0,17	0,04	0,00	6
LM0347	LM0348	0,04	1,00	0,04	0,17	0,04	0,00	6
LM0347	LM0518	0,04	1,00	0,04	0,17	0,04	0,00	6

(Fortsetzung nächste Seite)

(Fortsetzung)

von Punkt 1	nach Punkt 2	Cluster						g
		1	2	3	4	5	6	
LM0347	OS7461	0,04	1,00	0,04	0,17	0,04	0,00	6
LM0348	LM0349	0,04	1,00	0,04	0,18	0,04	0,00	6
LM0348	LM0518	0,04	1,00	0,04	0,17	0,04	0,00	6
LM0349	LM0500	0,03	0,50	0,52	0,13	0,03	0,00	1
LM0349	LM0518	0,04	1,00	0,04	0,17	0,04	0,00	6
LM0349	OS7375	0,03	0,99	0,04	0,22	0,05	0,00	4
LM0349	RH8434	0,07	0,96	0,03	0,14	0,04	0,00	7
LM0350	LM0351	0,04	0,98	0,02	0,27	0,08	0,00	4
LM0350	LM0500	0,03	0,48	0,51	0,19	0,05	0,00	1
LM0350	OS7375	0,03	0,97	0,02	0,28	0,07	0,00	3
LM0350	OS7383	0,03	0,48	0,01	0,20	0,42	0,00	1
LM0350	OS7485	0,04	0,96	0,01	0,23	0,10	0,00	4
LM0351	LM0352	0,04	1,00	0,02	0,24	0,06	0,00	4
LM0351	OS7416	0,05	1,00	0,02	0,20	0,06	0,00	5
LM0351	OS7485	0,04	0,98	0,01	0,21	0,09	0,00	5
LM0352	LM0353	0,04	1,00	0,02	0,23	0,06	0,00	4
LM0352	LM0354	0,04	0,99	0,02	0,26	0,07	0,00	4
LM0352	OS7385	0,04	1,00	0,02	0,21	0,06	0,00	5
LM0352	OS7416	0,05	1,00	0,02	0,20	0,06	0,00	5
LM0353	LM0354	0,04	0,99	0,02	0,25	0,07	0,00	4
LM0353	LM2003	0,05	0,99	0,02	0,17	0,05	0,00	6
LM0354	LM2003	0,05	0,98	0,02	0,20	0,06	0,00	5
LM0354	OS7220	0,03	0,53	0,02	0,50	0,10	0,00	1
LM0354	OS7385	0,04	0,99	0,02	0,23	0,07	0,00	4
LM0355	LM0356	0,01	0,00	0,00	0,07	0,12	0,94	8
LM0355	LM3359	0,02	0,02	0,00	0,10	0,27	0,49	2
LM0355	OS7220	0,02	0,04	0,01	0,40	0,12	0,49	1
LM0355	OS7385	0,03	0,50	0,01	0,13	0,09	0,49	1
LM0355	OS7482	0,01	0,00	0,00	0,07	0,11	0,99	9
LM0356	LM0357	0,01	0,00	0,00	0,06	0,13	0,93	7
LM0356	LM3359	0,02	0,02	0,00	0,09	0,28	0,45	2
LM0356	OS7482	0,01	0,00	0,00	0,06	0,11	0,95	8
LM0357	LM0358	0,01	0,00	0,00	0,05	0,10	0,95	10
LM0357	LM0527	0,01	0,00	0,00	0,06	0,10	0,98	10
LM0357	LM3359	0,02	0,02	0,00	0,09	0,28	0,48	2
LM0357	OS7383	0,01	0,00	0,00	0,08	0,44	0,48	1
LM0357	OS7482	0,01	0,00	0,00	0,06	0,11	0,98	9
LM0358	LM0500	0,01	0,00	0,50	0,06	0,05	0,46	1
LM0358	LM0527	0,01	0,00	0,00	0,05	0,07	0,96	13
LM0358	OS7383	0,01	0,00	0,00	0,07	0,41	0,46	1
LM0358	OS7469	0,01	0,00	0,00	0,04	0,07	0,94	14

(Fortsetzung nächste Seite)

(Fortsetzung)

von Punkt 1	nach Punkt 2	Cluster						g
		1	2	3	4	5	6	
LM0358	OS7470	0,01	0,00	0,00	0,07	0,51	0,46	1
LM0358	OS7494	0,01	0,00	0,00	0,05	0,07	0,96	14
LM0500	OS7375	0,03	0,50	0,52	0,17	0,04	0,00	1
LM0500	OS7383	0,02	0,00	0,50	0,09	0,39	0,00	1
LM0500	OS7470	0,02	0,00	0,50	0,09	0,49	0,00	1
LM0500	RH8434	0,06	0,46	0,51	0,08	0,03	0,00	1
LM0502	OS7371	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0503	LM0504	0,01	0,00	0,00	0,07	0,06	1,00	15
LM0503	LM0521	0,01	0,00	0,00	0,07	0,07	1,00	15
LM0503	LM0534	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0503	LM1260	0,01	0,00	0,00	0,07	0,07	1,00	15
LM0504	LM0505	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0504	LM0534	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0504	LM1260	0,01	0,00	0,00	0,07	0,07	1,00	15
LM0505	LM0506	0,01	0,00	0,00	0,06	0,07	1,00	14
LM0505	LM0509	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0505	LM1260	0,01	0,00	0,00	0,07	0,07	1,00	15
LM0506	LM0509	0,01	0,00	0,00	0,06	0,07	1,00	14
LM0506	LM0535	0,01	0,00	0,00	0,06	0,07	1,00	14
LM0508	OS7378	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0508	OS7494	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0508	AKUR	0,01	0,00	0,00	0,07	0,06	0,98	13
LM0511	LM0514	0,04	1,00	0,03	0,18	0,04	0,00	6
LM0511	OS7227	0,05	1,00	0,03	0,18	0,05	0,00	6
LM0511	OS7384	0,04	1,00	0,04	0,18	0,04	0,00	6
LM0514	OS7012	0,04	1,00	0,03	0,16	0,04	0,00	6
LM0514	OS7384	0,04	1,00	0,04	0,17	0,04	0,00	6
LM0514	OS7442	0,04	1,00	0,03	0,18	0,05	0,00	5
LM0518	OS5491	0,05	1,00	0,03	0,16	0,04	0,00	6
LM0518	OS7461	0,04	1,00	0,04	0,17	0,04	0,00	6
LM0518	RH8434	0,07	0,96	0,02	0,13	0,04	0,00	7
LM0521	OS7371	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0524	LM0525	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0525	LM0535	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0525	LM0538	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0527	LM0528	0,01	0,00	0,00	0,06	0,07	1,00	14
LM0527	LM0530	0,01	0,00	0,00	0,07	0,07	1,00	14
LM0527	OS7482	0,01	0,00	0,00	0,07	0,09	1,00	11
LM0527	OS7494	0,01	0,00	0,00	0,06	0,07	1,00	14
LM0528	LM0530	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0528	LM0534	0,01	0,00	0,00	0,07	0,06	1,00	15

(Fortsetzung nächste Seite)

(Fortsetzung)

von Punkt 1	nach Punkt 2	Cluster						g
		1	2	3	4	5	6	
LM0528	OS7378	0,01	0,00	0,00	0,06	0,06	1,00	15
LM0528	OS7494	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0530	LM0534	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0530	OS7371	0,01	0,00	0,00	0,07	0,06	1,00	14
LM0530	OS7482	0,01	0,00	0,00	0,07	0,08	1,00	12
LM0535	LM0536	0,01	0,00	0,00	0,06	0,07	1,00	14
LM0535	LM0538	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0536	LM0537	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0536	LM0538	0,01	0,00	0,00	0,06	0,07	1,00	15
LM0537	LM0538	0,01	0,00	0,00	0,06	0,06	1,00	16
LM2003	OS7220	0,04	0,53	0,02	0,42	0,08	0,00	1
LM2003	OS7225	0,04	0,49	0,02	0,47	0,07	0,00	1
LM3359	OS7383	0,02	0,02	0,01	0,11	0,59	0,00	5
LM3359	OS7385	0,04	0,52	0,01	0,15	0,25	0,00	2
LM3359	OS7416	0,04	0,52	0,01	0,13	0,24	0,00	2
LM3359	OS7485	0,04	0,50	0,01	0,14	0,27	0,00	2
OS5491	OS7012	0,05	1,00	0,03	0,15	0,04	0,00	7
OS5491	OS7472	0,05	1,00	0,02	0,15	0,05	0,00	7
OS5491	RH8434	0,07	0,96	0,02	0,12	0,05	0,00	8
OS5491	RH8437	0,49	0,50	0,01	0,09	0,03	0,00	1
OS5861	OS7379	0,01	0,00	0,00	0,04	0,06	0,87	15
OS5861	OS7384	0,02	0,50	0,02	0,10	0,05	0,41	1
OS5861	OS7441	0,00	0,00	0,00	0,03	0,06	0,83	14
OS5861	OS7442	0,02	0,50	0,01	0,12	0,05	0,41	1
OS5861	OS7472	0,03	0,50	0,01	0,09	0,05	0,41	1
OS5861	AKUR	0,01	0,00	0,00	0,06	0,05	0,89	15
OS7012	OS7442	0,04	1,00	0,03	0,18	0,05	0,00	6
OS7012	OS7472	0,05	1,00	0,02	0,15	0,05	0,00	7
OS7031	OS7227	0,03	0,50	0,01	0,10	0,06	0,46	1
OS7031	OS7367	0,02	0,50	0,01	0,13	0,07	0,46	1
OS7031	OS7441	0,01	0,00	0,00	0,04	0,07	0,88	13
OS7031	RH8701	0,01	0,00	0,00	0,04	0,07	0,95	14
OS7031	AKUR	0,01	0,00	0,00	0,06	0,06	0,94	15
OS7220	OS7225	0,02	0,04	0,01	0,77	0,11	0,00	7
OS7220	OS7385	0,04	0,54	0,02	0,45	0,09	0,00	1
OS7227	OS7367	0,05	1,00	0,02	0,19	0,06	0,00	5
OS7227	OS7384	0,04	1,00	0,03	0,17	0,05	0,00	6
OS7227	OS7441	0,03	0,50	0,01	0,10	0,06	0,43	1
OS7378	OS7494	0,01	0,00	0,00	0,06	0,07	1,00	15
OS7379	OS7469	0,01	0,00	0,00	0,04	0,07	0,94	14
OS7379	OS7472	0,03	0,50	0,01	0,09	0,06	0,46	1

(Fortsetzung nächste Seite)

(Fortsetzung)

von Punkt 1	nach Punkt 2	Cluster						<i>g</i>
		1	2	3	4	5	6	
OS7379	RH8437	0,47	0,00	0,00	0,04	0,04	0,46	1
OS7383	OS7485	0,03	0,48	0,01	0,14	0,43	0,00	1
OS7384	OS7441	0,02	0,50	0,02	0,11	0,05	0,43	1
OS7384	OS7442	0,04	1,00	0,03	0,19	0,05	0,00	5
OS7385	OS7416	0,05	1,00	0,02	0,17	0,06	0,00	6
OS7416	OS7485	0,05	0,98	0,01	0,16	0,08	0,00	6
OS7441	AKUR	0,01	0,00	0,00	0,06	0,06	0,91	15
OS7442	OS7472	0,05	1,00	0,02	0,17	0,05	0,00	6
OS7469	OS7470	0,01	0,00	0,00	0,07	0,51	0,47	1
OS7469	OS7494	0,01	0,00	0,00	0,05	0,07	0,97	15
OS7469	RH8437	0,47	0,00	0,00	0,04	0,04	0,47	1
OS7470	RH8434	0,06	0,46	0,01	0,09	0,50	0,00	1
OS7470	RH8437	0,47	0,00	0,00	0,06	0,49	0,00	1
OS7472	RH8437	0,49	0,50	0,01	0,09	0,03	0,00	1
RH8434	RH8437	0,51	0,46	0,01	0,06	0,03	0,00	1
RH8701	AKUR	0,01	0,00	0,00	0,07	0,06	0,98	15

Tabelle A.5: Zugehörigkeitswerte der Linien zu den sechs Clustern, die aus den berechneten Zugehörigkeitswerten der Objektpunkte abgeleitet wurden. Die Zugehörigkeitswerte sind auf zwei Nachkommastellen und das Gütemaß *g* auf eine ganze Zahl gerundet.

Anhang B

Symbolverzeichnis

a	:	Parameter
A	:	Fläche
c	:	Anzahl der Cluster
c^*	:	Optimale Anzahl der Cluster
C	:	Cluster
Cov	:	Kovarianzmatrix
D	:	Entscheidungsraum
$d_{i,k}$:	Distanz
$f()$:	Funktion
g	:	Gütemaß
M	:	Muster
M	:	Menge
\tilde{M}	:	unscharfe Menge
m	:	Dimension des Messdatenraumes, Objekte
n	:	Anzahl der Objekte
N	:	Matrix der Clusterschwerpunkte ν_i
p	:	Anzahl der Merkmale
P	:	Wahrscheinlichkeit
pe	:	Partitionsentropie
pe_x	:	Proportionsexponent
pk	:	Partitionskoeffizient
q	:	Gewichtsfaktor, Stufensprung
R	:	Rotationsmatrix
r	:	normierter Richtungsvektor
r	:	Minkowski-Konstante, Anzahl der Ausprägungen eines ordinalskalierten Merkmals, Komponenten eines normierten Richtungsvektors
$r_{i,k}$:	Korrelationskoeffizient
s	:	t-Conorm (auch s-Norm)
$s_{i,k}$:	Ähnlichkeit
si	:	Separationsindex

t	:	Zeit, t-Norm
U	:	Messdatenraum
u	:	Kordinaten im Klassenraum
V	:	Merkmalsraum
w	:	Grad der Unschärfe („Fuzzyfier“)
\mathbf{X}	:	Merkmals- oder Datenmatrix
X	:	X-Koordinate
\mathbf{x}_i	:	Merkmalsvektor des Objekts i
x	:	Merkmal
x_0	:	Merkmalswert, bei dem der Zugehörigkeitswert μ maximal ist
Y	:	Y-Koordinate
Z	:	Z-Koordinate
α	:	Variable eines Musters
β	:	Parameter einer Zugehörigkeitsfunktion
η	:	Wichtungsparameter
κ	:	Anzahl der Komponenten eines Musters
λ	:	Anzahl der Variablen eines Musters
μ	:	Zugehörigkeitsfunktion, -wert
ν_i	:	Schwerpunkt eines Clusters (Klassenschwerpunkt)
φ	:	Orientierungsunbekannte, Rotationsparameter
Ω	:	Grundmenge