



RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT BONN
INSTITUT FÜR GEODÄSIE UND GEOINFORMATION
FACHBEREICH PHOTOGRAMMETRIE



DIPLOMSTUDIENGANG VERMESSUNGSWESEN

DIPLOMARBEIT

von

Kerstin Herms

EXPLORATION DES SKALENRAUMES
BEZÜGLICH DER GEBÄUDEEXTRAKTION
IN TERRESTRISCHEN FARBBILDERN

BONN 2007

BETREUER:

PROF. DR.-ING. WOLFGANG FÖRSTNER
DIPL. INFORM. MARTIN DRAUSCHKE

INSTITUT FÜR GEODÄSIE UND GEOINFORMATION

PROFESSUR FÜR PHOTOGRAMMETRIE, NUSSALLEE 15, 53115 BONN

Diplomaufgabe

für Frau cand. geod. Kerstin Herms

Exploration des Skalenraumes bezüglich der Gebäudeextraktion in terrestrischen Farbbildern

Die Gebäudedetektion in digitalen Bildern stellt wegen der Komplexität der Objekte ein schwieriges Problem der Mustererkennung dar. In neueren Ansätzen zur Gebäudeextraktion wird das Bild in verschiedenen Auflösungsstufen, im sog. Skalenraum analysiert. Auf diese Weise können für die Bildinterpretation hinderliche Details ausgeblendet werden. Dabei spielen stabile Regionen, d. s. Regionen die sich bei Veränderung der Auflösung wenig ändern, eine besondere Rolle. Von stabilen Regionen im Skalenraum kann man auf kontraststarke Übergänge zwischen Objekten im Bild schließen [Drauschke et al. 2006: Stabilität von Regionen im Skalenraum].

Diese Diplomarbeit soll untersuchen, ob über stabilen Bildregionen eine Klassifikation von Gebäuden und anderen Objekten durchgeführt werden kann. Dazu sollen Merkmale der stabilen Regionen ausgewählt und bestimmt werden und diese Merkmale auf ihre Skalenabhängigkeit hin überprüft werden. Mit Hilfe eines geeignet gewählten Klassifikators sollen Gebäude und andere Objekte identifiziert werden. An Hand von terrestrischen Bildern soll bewertet werden, ob die u. U. skalenabhängigen Merkmale für die Gebäudeextraktion geeignet sind.

Prof. Wolfgang Förstner

Betreuer: W. Förstner, M. Drauschke

Ausgegeben am: 8. Januar 2007

Abgabetermin: 8. April 2007

Abgegeben am:

ERKLÄRUNG GEMÄSS § 22 (7) DPO:

Hiermit versichere ich, dass ich die vorliegende Diplomarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Bonn, 26. März 2007

(Kerstin Herms)

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	1
1.2	Aufgabenstellung und Zielsetzung	2
1.3	Gliederung	2
2	Grundlagen	4
2.1	Verfahren des Vorverarbeitungsschrittes	5
2.1.1	Segmentierung von Bildern	5
2.1.2	Aufbau eines Skalenraumes	8
2.1.3	Stabile Regionen im Skalenraum	9
2.2	Klassifikationsmethoden	12
2.2.1	Bayes-Klassifikation	12
2.2.2	Lineare Diskriminanzanalyse nach Fisher	13
2.2.3	Supportvektormaschinen (SVM)	15
2.2.4	Kriterien zur Auswahl einer Klassifikation und Fehler	19
3	Analyse des Skalenraumes	21
3.1	Werkzeuge	21
3.2	Ausgangsdaten	22
3.3	Vorverarbeitung der Bilddaten	22
3.3.1	Entzerrung	23
3.3.2	Skalenwahl und Auflösung	23
3.3.3	Beschreibung der Klassenzugehörigkeit der Trainingsdaten	25
3.4	Untersuchte Merkmale	25
3.4.1	Merkmale der stabilen Regionen	26
3.4.1.1	Umfang, Flächeninhalt und Formfaktor	26
3.4.1.2	Homogenität	28
3.4.2	Merkmale direkter Nachbarn	29
3.4.2.1	Flächeninhalt der direkten Nachbarn	29
3.4.2.2	Farbtöne der direkten Nachbarn	30
3.4.3	Merkmale indirekter Nachbarn	32
4	Auswertung der Ergebnisse	35
4.1	Skalenabhängigkeit der Merkmale	35
4.2	Ergebnisse der Analyse einzelner Merkmale	38

4.3	Der Einfluss von Rauschen	41
4.4	Klassifikation	44
4.4.1	Auswahl des Klassifikators	44
4.4.2	Klassifikationsergebnisse	49
5	Zusammenfassung und Ausblick	53
A	Bedienung des in Matlab implementierten Programms	55
A.1	Vorverarbeitungsschritt	55
A.2	Merkmalsextraktion	56
A.3	Ausgabe der Ergebnisse	57

Kapitel 1

Einleitung

1.1 Motivation

Untersuchungen zur Gebäudeextraktion stehen schon seit einiger Zeit im Fokus der Forschung zur Bildanalyse. Zu den Gebieten, die an einer vollautomatischen Gebäudeextraktion interessiert sind, zählt die Geoinformation. Im Bereich dreidimensionaler Stadtmodelle geht der Trend zu einer realitätsnahen Darstellung der Objekte. Hierzu werden Fassadendarstellungen benötigt. Eine effiziente Methode, um in kurzer Zeit viele Fassaden zu erfassen, stellt die Aufnahme der Gebäude durch terrestrische Farbbilder dar. Deren Extraktion ist jedoch noch ein ungelöstes Problem.

Ein weiteres Einsatzgebiet stellt die Verwaltung von Bildmaterial (beispielsweise städtebaulicher Planung) in Datenbanken dar. Ein Bild vermittelt meist einen guten Überblick des Sachverhaltes. Es enthält in der Regel aber mehr Information als eine wörtliche Beschreibung des Sachverhaltes. Nach einem bekannten Sprichwort heißt es ja auch: “Ein Bild sagt mehr als tausend Worte.“ Daher ist es sinnvoll, die Inhalte direkt dem Bild entnehmen zu können.

Für den Menschen erweist sich die Aufgabe des Erkennens von Objekten in Bildern als sehr einfach. Wir können einem Bild wie Abbildung 1.1 direkt entnehmen, ob es ein Gebäude enthält und wo sich dieses im Bild befindet. Selbst mit Objekten, die von der üblichen Vorstellung eines Hauses mit vier Wänden und einem Dach abweichen, als auch mit einem gewissen Maß an Verdeckung oder Unschärfe können wir umgehen. Dies ist möglich, da der Mensch Merkmale von Gebäuden kennt und diese schnell extrahieren und kombinieren kann. Der Klassifikationsvorgang beruht hier auf der menschlichen Erfahrungen. Unsere Intention ist es, diesen Vorgang auf die Maschine zu übertragen.



Abbildung 1.1: Darstellung eines Gebäudes mit Verdeckung.

Den Prozess der Merkmalsextraktion maschinell zu automatisieren, gestaltet sich jedoch als wesentlich komplexer. Hierfür wird ein mathematisches Modell benötigt, bestehend aus mehreren Merkmalen, durch das Gebäude beschrieben werden können. Es sollte möglichst allgemeingültig sein und daher mit der Formen- und Farbenvielfalt der Gebäude umgehen können. Gebäudeteile unterschiedlicher Größe erfassen zu wollen, führt uns auf die Betrachtung verschiedener Auflösungsstufen im Skalenraum. Auch gegenüber Rauschen im Bild sollte das Modell unempfindlich sein.

1.2 Aufgabenstellung und Zielsetzung

Ein noch ungelöstes Problem in der Bildverarbeitung ist die automatische Extraktion von Gebäuden aus terrestrischen Farbbildern. Ziel der Arbeit ist die Erforschung neuer Möglichkeiten für die Gebäudeextraktion in terrestrischen Farbbildern durch Beobachtung von Bildmerkmalen in unterschiedlichen Stufen des Skalenraumes.

Allgemeine Ansätze zur Skalenraumanalyse finden sich bei [Lindeberg 1990] und [Koenderink 1984]. Eine Betrachtung des Skalenraumes mit besonderem Augenmerk auf Gebäudeextraktion aus Luftbildern behandeln [Drauschke et al. 2006(1)].

Wir stellen speziell Merkmale für diejenigen Regionen auf, die mit Hilfe eines Wasserscheidenalgorithmus gewonnen werden und die sich im Skalenraum stabil verhalten. Über den Skalenraum hinweg wird auch das Verhalten der Merkmale verfolgt.

Einen guten Überblick über das Verfahren des Wasserscheidenalgorithmus bieten [Vincent, Soille 1991] und [Olsen, Nielsen 1997].

Ein Indikator dafür, ob sich ein Merkmal für die Klassifikation von Gebäuden und anderen Objekten eignet, stellt dessen Vermögen dar, die Daten unterschiedlicher Klassen zu trennen. Wir wollen zusätzlich das Verhalten der einzelnen Merkmale bei verrauschten Bildern mit in die Untersuchung einfließen lassen.

Als geeignet erscheinende Merkmale werden kombiniert und dienen der Aufstellung von Trainingsdaten. Diese fließen in eine Klassifikation ein. Es sei jedoch bereits an diese Stelle darauf hingewiesen, dass manche Merkmale alleine genommen schlechte Ergebnisse liefern, in der Kombination mit anderen Merkmalen allerdings einen wichtigen Beitrag leisten. Die Güte der Merkmalskombination wird auch hier darüber bestimmt, wie gut sie sich für die Separation der Daten unterschiedlicher Klassen eignet.

Die Auswahl einer geeigneten Klassifikationsmethode stützt sich unter anderem auf die Arbeiten von [Duda et al. 2001] und [Franc, Hlaváč 2004].

1.3 Gliederung

Kapitel 2 stellt Konzepte vor, die in diese Arbeit einfließen. Das Kapitel gliedert sich in zwei Teile. Zum einen werden Grundlagen der Skalenraumanalyse vorgestellt. Hier geht es um den Aufbau eines Skalenraumes sowie der Segmentierung der skalierten Bilder in Wasserscheidenregionen, die für die Gebäudeextraktion näher betrachtet werden. Des Weiteren werden ausgewählte Klassifikationsmethoden vorgestellt, die zur Auswertung der Merkmale herangezogen werden. Dazu gehört

die Bayes-Klassifikation, die Diskriminanzanalyse nach Fisher sowie die Supportvektormaschinen.

Einen Einblick in die Datengrundlage sowie die für die Vorverarbeitung und Klassifikation der Daten verwendeten Werkzeuge bietet Kapitel 3. Anschließend werden Klassifikationsmerkmale vorgestellt, die wir auf ihre Eignung für die Gebäudeextraktion aus terrestrischen Farbbildern überprüfen.

Kapitel 4 umfasst die Resultate der Skalenraumanalyse bezüglich der in Kapitel 3 beschriebenen Merkmale. Betrachtet wird ihr Verhalten im Skalenraum, ihre Fehleranfälligkeit im Rahmen einer Klassifikation mit den in Kapitel 2 beschriebenen Methoden sowie der Einfluss von Rauschen im Bild auf das Ergebnis.

Einen Abschluss dieser Diplomarbeit bildet Kapitel 5 mit einer Zusammenfassung und einem Ausblick.

Im Anhang befindet sich eine CD, die sowohl das Bildmaterial dieser Arbeit als auch die in Matlab implementierten Methoden umfasst.

Kapitel 2

Grundlagen

In diesem Kapitel wird ein Konzept vorgestellt, das eine Basis für unsere nachfolgenden Untersuchungen zur Gebäudeextraktion bildet.

Der für den Menschen offensichtlich einfache Vorgang des Erkennens von Gebäuden in Bildern, gestaltet sich für eine Maschine als äußerst komplex. Für diese Aufgabe stellt die Mustererkennung ein mächtiges Werkzeug dar, um die für eine Objektklasse relevanten Teile eines Bildes vom restlichen Bild zu trennen.

Der grundlegende Prozess der Mustererkennung besteht nach Duda [Duda et al. 2001] aus drei Schritten, wie sie in Abbildung 2.1 am Beispiel einer Gebäudeaufnahme dargestellt sind.

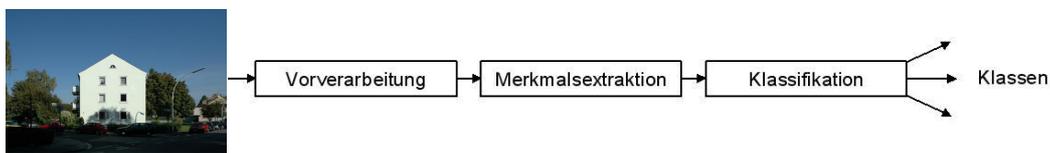


Abbildung 2.1: Prozess der Mustererkennung

Der Schritt der Vorverarbeitung der Bilddaten dient dazu, die Bildanalyse zu vereinfachen. Wir betrachten in dem Zusammenhang speziell die Bildsegmentierung sowie den Aufbau eines Skalenraumes (Kapitel 2.1). Dabei werden für diese Bereiche grundlegende Begriffe eingeführt.

Anschließend können die Merkmalsausprägungen der vorverarbeiteten Bilder für zuvor festgelegte Merkmale bestimmt werden. Dies geschieht im Schritt der Merkmalsextraktion. Die speziell auf das Problem der Gebäudeextraktion angepasste Merkmalsauswahl wird später in Kapitel 3 vorgestellt.

Eine Auswertung der Ergebnisse der Merkmalsextraktion erfolgt im Klassifikationsschritt. Hier wird die Entscheidung über die Klassenzugehörigkeit getroffen. Wir betrachten verschiedene grundlegende Techniken aus dem Bereich der Klassifikation (Kapitel 2.2). Hierbei liegt der Fokus auf der Bayes-Klassifikation, der linearen Diskriminanzanalyse nach Fisher sowie den Supportvektormaschinen.

2.1 Verfahren des Vorverarbeitungsschrittes

Als Ausgangsdaten für den Schritt der Vorverarbeitung betrachten wir Farbbilder. Diese werden im Folgenden mit $\mathbf{b}(x_1, x_2)$ bezeichnet. Es handelt sich dabei um 3-Vektoren, deren Komponenten je einen der drei RGB-Farbkanäle $b_c(x_1, x_2)$, $c = 1, \dots, 3$ enthalten.

Arbeiten aus dem Bereich der Gebäudeextraktion beschäftigen sich im großen Umfang mit der Analyse von Luftbildern. Dieser Ansatz soll hier nicht verfolgt werden. Stattdessen basieren die Untersuchungen auf terrestrischen Aufnahmen.

Der grundlegende Unterschied zwischen Luftbildern und terrestrischen Bildern liegt in der Komplexität, die in der unterschiedlichen Auflösung begründet ist. Ist die Gebäudeextraktion aus Luftbildern aus diesem Grund bereits ein komplexes Problem, so ist die Extraktion von Gebäuden aus terrestrischen Bildern nochmals komplexer. Diese Eigenschaft spielt bei der später behandelten Merkmalsauswahl sowie bei der Festlegung der Parameter der angewandten Methoden eine Rolle. Die Methodik der Verfahren des Vorverarbeitungsschrittes ist für beide Bildarten jedoch vergleichbar. Aus diesem Grund müssen wir bei der Bezeichnung der Farbbilder mit $\mathbf{b}(x_1, x_2)$ nicht weiter zwischen Luftbildern und terrestrischen Bildern differenzieren.

Wir betrachten in diesem Kapitel zwei ausgewählte Verfahren, die wir im Zuge der Vorverarbeitung für die Gebäudeextraktion anwenden. Dabei sei an dieser Stelle auf die Austauschbarkeit der Methoden hingewiesen.

Wir verwenden zur Bildsegmentierung einen Wasserscheidenalgorithmus auf dem Bild der quadratischen Gradienten. Dieser wird in Kapitel 2.1.1 näher erläutert. Die anschließende Merkmalsextraktion wird zudem nicht auf das Originalbild beschränkt, sondern auf unterschiedliche Auflösungsstufen des Bildes ausgeweitet. Eine zunehmende Skalierung und die damit einhergehende Verringerung der Auflösung des Originalbildes erweist sich als vorteilhaft für die Merkmalsextraktion, da feinere Bildstrukturen ausgeblendet werden können. Die Auflösungsstufen ergeben zusammengefasst einen Skalenraum, mit der geringsten Auflösung als höchste Skala. Der Aufbau eines Skalenraumes wird in Kapitel 2.1.2 beschrieben. In Kapitel 2.1.3 verknüpfen wir den Wasserscheidenalgorithmus mit dem Skalenraum.

2.1.1 Segmentierung von Bildern

Die Merkmalsextraktion aus Bildern erfordert eine Segmentierung dieser, um eine Beschreibung der Form der Bildstruktur zu erhalten [Olsen, Nielsen 1997]. Als Segmentierungsalgorithmus wird an dieser Stelle der Wasserscheidenalgorithmus vorgestellt, der eine Grundlage dieser Arbeit darstellt. Nachstehende Definitionen beziehen sich auf eine Arbeit von Vincent und Soille [Vincent, Soille 1991] sowie auf eine Arbeit von Olsen und Nielsen [Olsen, Nielsen 1997].

Der Wasserscheidenalgorithmus basiert auf der Vorstellung, dass Bilder als Gebirge von Grauwerten, Farbintensitäten oder anderen Bildfunktionen interpretiert werden können. In dem Zusammenhang lässt sich der Begriff der Wasserscheide als Grenze eines Einzugsgebietes von der Topographie ins Bild übertragen.

Betrachten wir im Folgenden die Wasserscheiden des quadratischen Gradientenbildes h . Dieses lässt sich mit Hilfe der quadratischen Gradienten der einzelnen Farbkanäle $\|\nabla b_c(x_1, x_2)\|^2$ und einem Gewicht $\sigma_{n_c}^2$, der Rauschvarianz des jeweiligen Farbkanals, bestim-

men [Brügelmann, Förstner 1992].

$$h(x_1, x_2) = \sqrt{\sum_{c=1}^3 \frac{\|\nabla b_c(x_1, x_2)\|^2}{\sigma_{n_c}^2}} \quad (2.1)$$

Ein Einzugsgebiet bezieht sich dann jeweils auf ein regionales Minimum dieses quadratischen Gradientenbildes. Wenden wir uns zur besseren Anschauung nochmals der Interpretation als Gebirge zu. Abbildung 2.2 zeigt einen Querschnitt durch ein solches Gebirge.

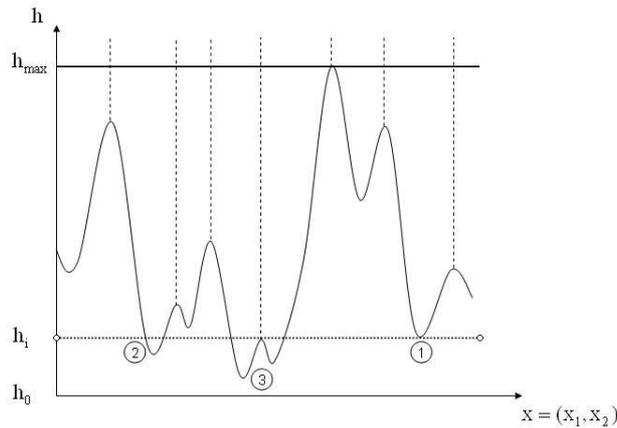


Abbildung 2.2: Schematische Darstellung der Wasserscheiden. Das Bild zeigt ein Gebirge (der Werte des quadratischen Gradienten). Die waagerechten Linien zeigen den Füllstand des Wassers an (gepunktet = Füllhöhen h_i und h_j , durchgezogen = maximale Füllhöhe h_{max}). Es treten drei Fälle von Punkten auf. (1): Es liegt ein lokales Minimum vor, das ein Einzugsgebiet beschreibt. (2): Der Punkt wird dem benachbarten Minimum in der Höhe h_j zugeordnet. (3): Hier stoßen Wassermassen aus zwei unterschiedlichen Einzugsgebieten zusammen. In gestrichelten senkrechten Linien sind die Wasserscheiden markiert, die an diesen Stellen entstehen.

Nun beginnen wir den Querschnitt von unten her mit Wasser zu füllen. Der Ausgangswasserstand liegt in der Höhe $h_0 = 0$, der maximale Wasserstand h_{max} ist durch die Höhe des Gebirges bestimmt. Für jeden Punkt x_i des Querschnitts, der in einer Höhe $h_i = h(x_i)$, $0 \leq h_i \leq h_{max}$ liegt, werden drei Fälle der Zugehörigkeit zu einem Einzugsgebiet unterschieden.

1. $\nexists x_j$ aus der Nachbarschaft von x_i mit $h_j \leq h_i$
In diesem Fall ist x_i ein lokales Minimum, das ein Einzugsgebiet definiert.
2. $\exists x_j$ aus der Nachbarschaft von x_i mit $h_j \leq h_i$
Dann ist x_i dem gleichen Minimum zugeordnet wie x_j . Alle Punkte die zum gleichen Minimum gehören, beschreiben ein Einzugsgebiet.
- 3.a $h_j \leq h_i \forall x_j$ aus der Nachbarschaft von x_i
In diesem Fall liegt bei x_i eine Wasserscheide vor. Das bedeutet, dass an dieser Stelle des Querschnitts Wassermassen aus zwei Bereichen zusammenstoßen.

Die ersten beiden Fälle sind vom Querschnitt auf das gesamte Bild übertragbar. Im Falle einer Wasserscheide haben wir es hier jedoch nicht mit einem lokalen Maximum sondern mit einem Grat zu tun. Daher führen wir den alternativen dritten Fall ein.

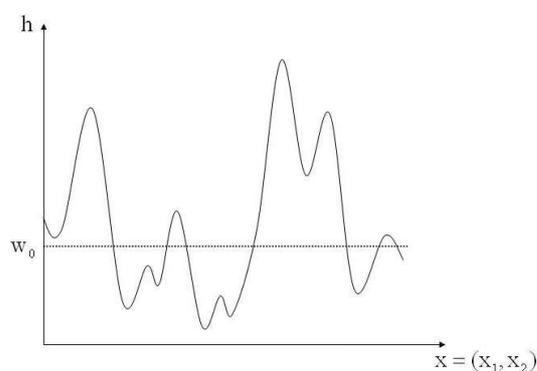
3.b $\exists (x_1, x_2)_j$ und $(x_1, x_2)_k$ aus der Nachbarschaft von $(x_1, x_2)_i$ mit $h_j \leq h_i$ und $h_k \leq h_i$ und zu zwei verschiedenen Einzugsgebieten gehörig

In diesem Fall macht $(x_1, x_2)_i$ einen Teil der Wasserscheide aus.

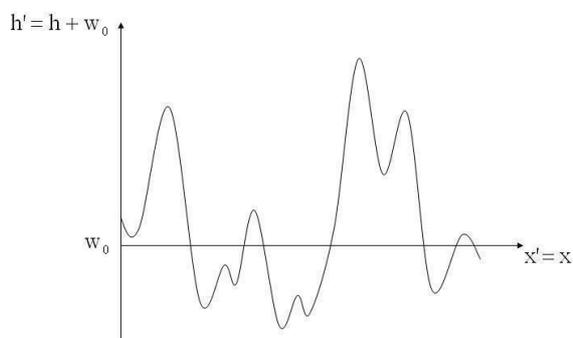
Der Wasserscheidenalgorithmus ermöglicht somit eine vollständige Partitionierung eines Bildes

$$\mathbf{b} = \bigcup_i \mathcal{R}_i \text{ mit } \mathcal{R}_i \cap \mathcal{R}_j = \emptyset, \forall i \neq j \quad (2.2)$$

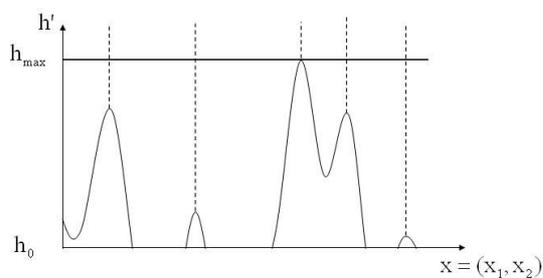
in (bezüglich der Intensität) homogene Regionen \mathcal{R}_i mit $i \geq 1$ sowie deren Wasserscheiden, die in \mathcal{R}_0 zusammengefasst werden. Ausgegeben werden Bilder mit den Labels i der Regionen. Die Vollständigkeit der Partitionierung ist entscheidend, da sie ein einfaches Auffinden der für die Klassifikation benötigten Nachbarregionen ermöglicht [Drauschke et al. 2006(2)].



(a)



(b)



(c)

Abbildung 2.3: Rauschunterdrückung bei der Bildsegmentierung

Aufgrund von Rauschen entstehen jedoch sehr viele kleine Regionen, da mehr lokale Minima auftreten als bei rauschfreien Signalen. Eine solche Übersegmentierung ist in Hinblick auf die anschließende Merkmalsextraktion zu vermeiden. Dazu kann man auf eine Rauschunterdrückung zurückgreifen, wie sie in Abbildung 2.3 schematisch dargestellt ist.

Zunächst wird ein Schwellwert w_0 festgelegt, der den Bereich des Rauschens abgrenzt (Abbildung 2.3(a)). Um diesen Schwellwert wird das Koordinatensystem (x, h) in Richtung h verschoben (Abbildung 2.3(b)). Als neues Koordinatensystem ergibt sich somit $(x', h') = (x, h + w_0)$. Setzen wir nun alle negativen Werte auf Null, dann erhalten wir regionale Minima und damit größere Einzugsgebiete (Abbildung 2.3(c)). Mit diesem Verfahren erreicht man also, dass geringe Änderungen der Intensität in einem Bild nicht als Wasserscheiden markiert werden. Der Wasserscheidenalgorithmus kann ansonsten wie oben beschrieben angewandt werden.

2.1.2 Aufbau eines Skalenraumes

Mit dem Ziel vor Augen, Strukturen im Bild zu detektieren, die sich einem Gebäude zuordnen lassen, beschäftigen wir uns nun mit deren Auflösung. Zu berücksichtigen ist, dass einzelne Bildobjekte nur in einem begrenzten Auflösungsbereich auftreten [Koenderink 1984]. Kleine Strukturen (z.B. Klinker, Dachziegel) existieren nur in Bildern mit hoher Auflösung. In diesen Bildern lassen sich große Strukturen (z.B. die gesamte Fassade) jedoch schlecht erkennen. Aus diesem Grund bietet es sich an, unterschiedliche Auflösungsstufen zu betrachten.

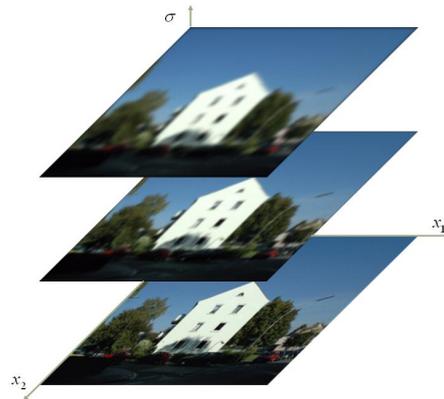


Abbildung 2.4: Schematische Darstellung eines Skalenraumes in den Bildkoordinaten x_1, x_2 und dem Skalierungsparameter σ . Bei $\sigma = 0$ liegt das Originalbild vor. Die Skalen $\sigma \geq 0$ enthalten eine Glättung des Originalbildes. Mit wachsendem σ nimmt die Größe des Glättungsfilters zu und die Auflösung ab.

Ein Skalenraum bietet eine Möglichkeit der Darstellung der Auflösungsstufen (Abbildung 2.4) in Abhängigkeit eines Parameters σ . Bilder werden dabei nach ihrem Auflösungsgrad geordnet und als Layer übereinandergelegt. Wir identifizieren das Bild mit der höchsten Auflösung (Originalbild) mit der ersten Skala ($\sigma = 0$). Für die theoretische Einführung eines Skalenraumes gehen wir zunächst von einem kontinuierlichen Stapel von Layern aus. Im praktischen Teil beschränken wir uns später auf eine diskrete Auswahl dieser Layer.

Die einzelnen Skalen ($\sigma > 0$) erhält man nach [Lindeberg 1990] durch eine einparametrische Transformation des Originalbildes $\mathbf{b}(x_1, x_2)$.

$$\mathbf{b}(x_1, x_2; \sigma) = \begin{cases} \mathbf{b}(x_1, x_2) & \text{falls } \sigma = 0 \\ \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \mathbf{b}(x_1 - i, x_2 - j) g(x_1, x_2; \sigma) & \text{falls } \sigma > 0 \end{cases} \quad (2.3)$$

Dies entspricht einer Faltung $\mathbf{b}(x_1, x_2; \sigma) = \mathbf{b}(x_1, x_2) * g(x_1, x_2; \sigma)$ des Originalbildes mit einem Filter $g(x_1, x_2; \sigma)$. Wir verwenden einen Gaußfilter, da dieser das Entstehen neuer Strukturen (insbesondere neuer Extrema) verhindert [Koenderink 1984] und damit eine Vergleichbarkeit zwischen den einzelnen Skalen ermöglicht. In der allgemeinen Form für n -dimensionale Signale hat der Gaußfilter nach [Koch 1997] die Form

$$g(\mathbf{x}; \sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \Sigma)^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})' \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})}. \quad (2.4)$$

Unter der Annahme einer ortserhaltenden Filters erhalten wir $\boldsymbol{\mu} = \mathbf{0}$ als Mittelwert. Gehen wir zusätzlich von einer in beide Achsrichtungen gleichstarken Glättung aus, so kann die Varianz zu $\Sigma = \sigma^2 I$ angenommen werden. Mit diesen Werten ergibt sich aus (2.4) folgende Größe

$$g(\mathbf{x}; \sigma) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2} \mathbf{x}' \mathbf{x}}. \quad (2.5)$$

Zudem können wir uns auf die Glättung zweidimensionaler Bildsignale ($n = 2$) beschränken. Somit erhält der Gaußfilter die Form

$$g(x_1, x_2; \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}(x_1^2 + x_2^2)}. \quad (2.6)$$

In der praktischen Umsetzung gehen wir vom kontinuierlich definierten Skalenraum zu diskreten Auflösungsstufen mit logarithmischem Abstand über.

Der von uns betrachtete Skalenraum besteht aus den Skalen $s = 1, \dots, s_{max}$, denen je ein Skalierungsparameter $\sigma \in [0, \sigma_{max}]$ zugeordnet ist. Die Skalen setzen sich aus dem ungefilterten Originalbild ($s = 1 \hat{=} \sigma = 0$), N_1 Skalen mit einem Parameter $\sigma \in]0, 1[$ und N_2 Skalen mit $\sigma \in [1, \sigma_{max}]$ zusammen. Für die Bestimmung des Skalierungsparameters σ in Abhängigkeit von s berechnen wir

$$\sigma = \sigma_0^{s-N_1-2} \quad (2.7)$$

aus einem Startwert $\sigma_0 = (\sigma_{max})^{\frac{1}{s_{max}-1}}$.

Als Filtergröße des Gaußfilters wählen wir das Doppelte der dreifachen Standardabweichung $d = \lceil 2 \cdot 3\sigma \rceil$. Für eine ungerade Pixelanzahl der Filter wird gegebenenfalls noch eine 1 zu d hinzuaddiert.

2.1.3 Stabile Regionen im Skalenraum

Greifen wir nun den Wasserscheidenalgorithmus aus Kapitel 2.1.1 wieder auf. Mit Hilfe des Wasserscheidenalgorithmus kann für jede Skala eine Segmentierung vorgenommen werden. Das Ergebnis ist nach Formel (2.2) eine vollständige Partitionierung der Skala in Regionen $\mathcal{R}(\sigma)$, $i = 1, \dots, n$.

$$\mathbf{b}(x_1, x_2; \sigma) = \bigcup_i \mathcal{R}_i(\sigma) \text{ mit } \mathcal{R}_i(\sigma) \cap \mathcal{R}_j(\sigma) = \emptyset, \forall i \neq j \quad (2.8)$$

Die dabei erfassten Regionen unterscheiden sich in Abhängigkeit der Skalierung. Zwischen den unterschiedlichen Auflösungsstufen können Regionen neu entstehen, geteilt werden, verschmelzen oder sich auflösen. Regionen gehen jedoch immer vollständig ineinander auf und werden nicht auf mehrere Regionen verteilt. Dies ist von Vorteil für die Merkmalsextraktion, da es uns ermöglicht das Verhalten von Regionen über den gesamten Skalenraum zu verfolgen. Eine besondere Rolle spielen dabei diejenigen Regionen \mathcal{R}_i , die sich über mehrere aufeinanderfolgende Skalen nur geringfügig unterscheiden. Solche Regionen werden als „stabile Regionen“ bezeichnet.

Für die Messung der Stabilität gibt es nach [Drauschke et al. 2006(1)] verschiedene Kriterien geometrischer und topologischer Art. Ein geeigneter Indikator ist der Flächeninhalt der Wasserscheidenregionen, da er bei der Verschmelzung zweier Regionen stark anwächst. Wir betrachten daher die Änderung des Flächeninhalts einer Region \mathcal{R}_i über λ aufeinanderfolgende Stufen im Skalenraum (d.h. über einem Intervall $\mathcal{I} = [\sigma_k, \sigma_{k+\lambda}]$). Dazu werden in jeder Skala $\sigma_l \in \mathcal{I}$ alle Regionen $\mathcal{R}_j(\sigma_l)$ bestimmt, die einen Schnitt mit der Region $\mathcal{R}_i(\sigma_k)$ aufweisen. Aus den $\mathcal{R}_j(\sigma_l)$ wird diejenige Region gewählt, die den größten Schnitt mit $\mathcal{R}_i(\sigma_k)$ hat. Die Differenzen der Regionen $\mathcal{R}_i(\sigma_k)$ und $\mathcal{R}_i(\sigma_l)$ dürfen zusammen einen Schwellwert t nicht überschreiten [Drauschke et al. 2006(2)]. Eine Region \mathcal{R}_i ist demnach genau dann auf dem Intervall $[\sigma_k, \sigma_{k+\lambda}]$ stabil, wenn $\forall \sigma_l \in \mathcal{I}$ gilt

$$|[\mathcal{R}_i(\sigma_k) \setminus \mathcal{R}_j(\sigma_l)] \cup [\mathcal{R}_j(\sigma_l) \setminus \mathcal{R}_i(\sigma_k)]| \leq t. \quad (2.9)$$

Ein Ablaufschema zur Bestimmung der stabilen Regionen im Skalenraum ist in Abbildung 2.5 gegeben.

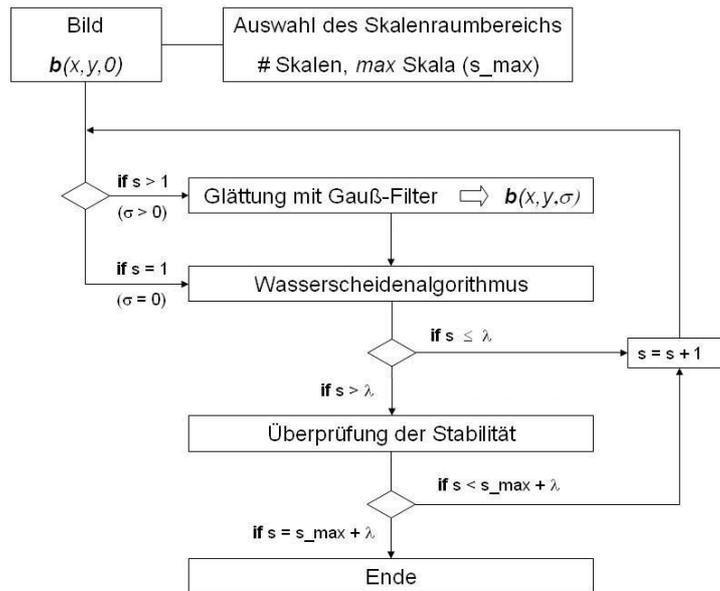


Abbildung 2.5: Ablaufschema für die Berechnung des Skalenraumes und der Bestimmung der stabilen Regionen.

Vorab ist ein diskreter Ausschnitt des Skalenraumes zu wählen. Dies beinhaltet die Festlegung des maximalen Skalenniveaus σ_{max} sowie der Anzahl N_1 der Skalen mit $\sigma \in]0, 1[$ und der Anzahl N_2 der Skalen mit $\sigma \in [1, \sigma_{max}]$ (vgl. Kapitel 2.1.2).

Da das Originalbild nicht gefiltert werden muss, wird für die erste Skala ($s = 1$) also für den Fall $\sigma = 0$ zunächst nur der Wasserscheidenalgorithmus ohne vorangestellte Filterung berechnet. In allen anderen Fällen wird erst der Filterschritt durchlaufen, in dem das Originalbild mit einem Gaußfilter (2.6) geglättet wird. Anschließend werden auch für diese Bilder die Wasserscheidenregionen bestimmt.

Für die ersten λ Schritte, inklusive dem Originalbild, liegen noch nicht genügend Informationen für die Stabilitätsanalyse vor. Hierbei bezeichnet λ die Anzahl der Skalen, über die eine Region ihren Flächeninhalt nur geringfügig ändern darf, damit sie als stabil gilt. Ab der Skala $\lambda + 1$ kann auch die Stabilität überprüft werden, indem man für jede Wasserscheidenregion in diesem Bild testet, wie viele Pixel in weiteren λ Skalen zu der jeweiligen Region gehören. Um dies für alle Skalen bis $s = s_{max}$ durchführen zu können, benötigt man λ zusätzliche Skalen, die alleine der Überprüfung der Stabilität dienen. Der Algorithmus endet somit nach $s_{max} + \lambda$ Skalen. Aussagekräftige Information über die Stabilität von Regionen erhalten wir jedoch nur für die Skalen bis s_{max} .

Die Ausgabe umfasst die skalierten Bilder sowie die Darstellung der Wasserscheidenregionen inklusive der stabilen Regionen (Abbildung 2.6).

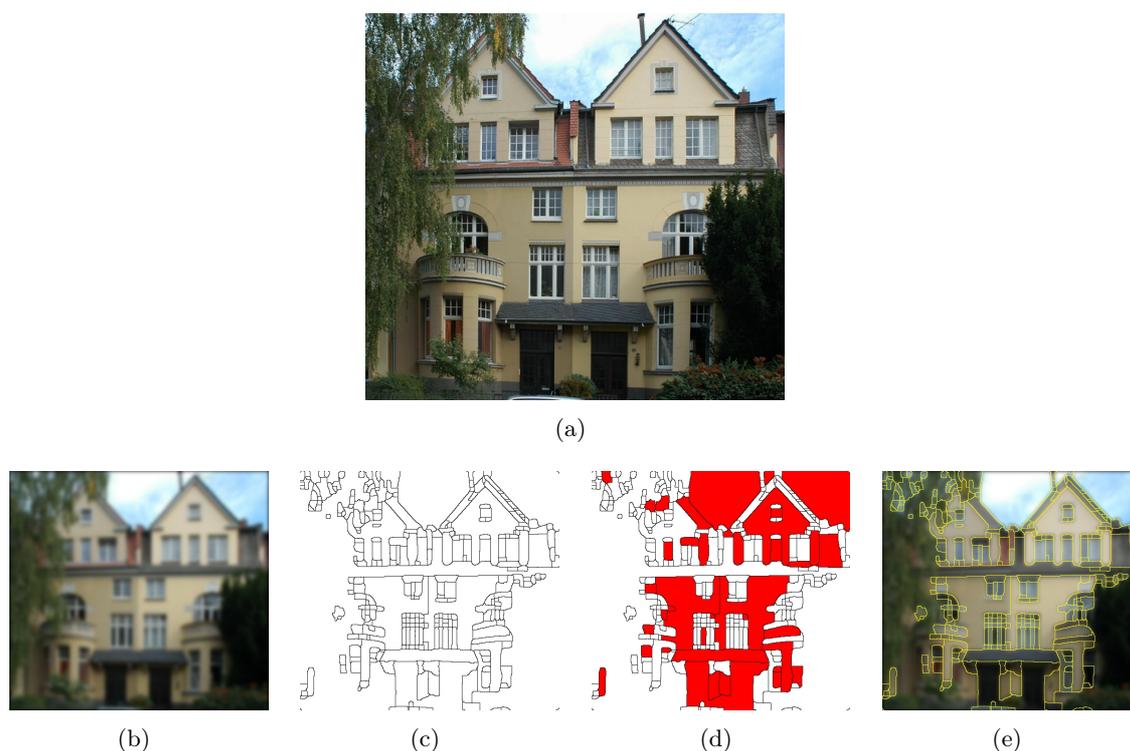


Abbildung 2.6: Ausgabe nach dem Aufbau des Skalenraumes: (a) Originalbild, (b) geglättetes Bild der Stufe $\sigma = 4$, (c) Wasserscheidenregionen der Stufe $\sigma = 4$, (d) stabile Regionen (rot) der Stufe $\sigma = 4$, (e) Wasserscheidenregionen der Stufe $\sigma = 4$ im geglätteten Bild.

Der Hauptaugenmerk dieser Arbeit richtet sich auf die stabilen Regionen des Skalenraums und deren Nachbarregionen. Es stellt sich die interessante Frage, ob sich in Verbindung mit stabilen

Regionen Merkmale finden lassen, die eine Detektion von Gebäuden in terrestrischen Farbbildern ermöglichen. Dieser Frage werden wir in den Kapiteln 3 und 4 nachgehen.

2.2 Klassifikationsmethoden

In diesem Kapitel werden drei ausgewählte Klassifikationsmethoden vorgestellt, die bei der Gebäudeextraktion eingesetzt werden: die Bayes-Klassifikation, die lineare Diskriminanzanalyse nach Fisher sowie eine Klassifikation mit Hilfe von Supportvektormaschinen. Nachstehende Formeln beziehen sich auf [Duda et al. 2001], [Franc, Hlaváč 2004] und [Christianini, Shawe-Taylor 2000]. Für alternative Klassifikationsmethoden sei ebenfalls auf diese Literatur verwiesen.

Um zwei oder mehr Klassen $y \in \{1, \dots, m\}$ unterscheiden zu können, werden Merkmale benötigt, die diese Klassen auszeichnen. Voraussetzung für eine fehlerfreie Klassifikation ist die Separierbarkeit der Daten auf der Grundlage dieser Merkmale. Diese Voraussetzung ist jedoch bei komplexen Klassifikationsproblemen wie der Gebäudeextraktion für einzelne Merkmale nicht gegeben. Daher ist ein einzelnes Merkmal in der Regel nicht ausreichend, um eine Klasse zu beschreiben. Zudem werden je nach Art der Klasse andere Merkmale benötigt. Es sind also mehrere Merkmale aufzustellen. Diese werden in einem n -dimensionalen Merkmalsvektor \mathbf{x} zusammengefasst.

Die Merkmale sollten so gewählt werden, dass sie invariant bezüglich Änderungen des Objekts und der Beobachtungen sind. Um geeignete Merkmale auswählen zu können, benötigt man Trainingsdaten, deren Klassenzugehörigkeit bekannt ist.

Im allgemeinen Fall unterscheiden wir m Klassen. Speziell für den Fall der Gebäudeextraktion, können wir uns zunächst auf ein Zweiklassenproblem beschränken, um Gebäudeteile vom Hintergrund zu trennen.

In einer Lernphase wird zunächst die Verteilung der Merkmale aus den Trainingsdaten geschätzt, deren Klassenzugehörigkeit bekannt ist. Die Zuweisung von Testdaten mit unbekannter Klassenzugehörigkeit zu einer der Klassen erfolgt anschließend in der Arbeitsphase. Entscheidungsgebend ist dabei ein Distanzmaß zwischen den Testdaten und den Klassen von Trainingsdaten.

2.2.1 Bayes-Klassifikation

Die Bayes-Klassifikation ist ein Bestandteil der statistischen Entscheidungstheorie. Sie basiert auf der Formel von Bayes

$$P(y|\mathbf{x}) = \frac{p(\mathbf{x}|y)P(y)}{p(\mathbf{x})}, \quad (2.10)$$

welche die a posteriori Wahrscheinlichkeit $P(y|\mathbf{x})$ für das Auftreten eines wahren Zustandes $y \in \{1, \dots, m\}$ unter der Beobachtung eines Merkmalsvektors \mathbf{x} beschreibt. Einfließende Größen sind die a priori Wahrscheinlichkeit $P(y)$ für das Auftreten von y ohne Vorwissen, die Likelihoodfunktion $p(\mathbf{x}|y)$ sowie die Dichte $p(\mathbf{x})$.

Ausgehend von disjunkten Klassen $y = 1, \dots, m$, können wir mit Hilfe der Formel von Bayes (2.10) die a posteriori Wahrscheinlichkeit für jede dieser Klassen bestimmen. Aus den unterschiedlichen $P(y|\mathbf{x})$ kann man eine Entscheidungsregel $q(\mathbf{x})$ ableiten, die Aufschluss darüber

gibt, welche Klassenzuordnung am besten geeignet ist.

Durch das Einführen einer Verlustfunktion $W(q|y)$ wird beschrieben, welche Konsequenz eine Entscheidung für eine bestimmte Klasse (und damit gegen die anderen Klassen) hat [Rinne 2003]. Eine kostenintensive Entscheidung wird durch einen höheren Wert in der Verlustfunktion bestraft. Wird als Verlustfunktion die Funktion

$$W_{0/1}(q(\mathbf{x}, y)) = \begin{cases} 0 & \text{falls } q(\mathbf{x}) = y \\ 1 & \text{falls } q(\mathbf{x}) \neq y \end{cases} \quad (2.11)$$

gewählt, so wird die Wahrscheinlichkeit der Missklassifikation minimiert. Das Risiko R einer Missklassifikation entspricht dem Erwartungswert des Verlustes. Unter beobachtetem \mathbf{x} entspricht das Risiko für eine einzelne Entscheidung $q(\mathbf{x}) = j$ (bezeichnet als q_j) dem Wert

$$R(q_j|\mathbf{x}) = \sum_{k=1}^m W_{1/0}(q_j|y^k) P(y^k|\mathbf{x}), \quad (2.12)$$

falls die wahre Klassenzugehörigkeit $y = k$ (bezeichnet als y^k) ist. Gesucht ist nun diejenige Entscheidung q_j , für die $R(q_j|\mathbf{x})$ minimal ist.

Betrachten wir das Zweiklassenproblem ($m = 2$). Dann lautet die Entscheidungsregel

$$q(\mathbf{x}) = \begin{cases} 1 & \text{falls } R(q_1|\mathbf{x}) \leq R(q_2|\mathbf{x}) \\ 2 & \text{falls } R(q_2|\mathbf{x}) < R(q_1|\mathbf{x}) \end{cases}. \quad (2.13)$$

Setzen wir (2.11) in (2.12) ein, so führt dies auf die Gleichungen

$$\begin{aligned} R(q_1|\mathbf{x}) &= 0 \cdot P(y_1|\mathbf{x}) + 1 \cdot P(y_2|\mathbf{x}) \\ R(q_2|\mathbf{x}) &= 1 \cdot P(y_1|\mathbf{x}) + 0 \cdot P(y_2|\mathbf{x}). \end{aligned} \quad (2.14)$$

Unter der Berücksichtigung der Formel von Bayes (2.10) erhalten wir schließlich folgende Entscheidungsregel.

$$q(\mathbf{x}) = \begin{cases} 1 & \text{falls } p(\mathbf{x}|y_1)P(y_1) \geq p(\mathbf{x}|y_2)P(y_2) \\ 2 & \text{falls } p(\mathbf{x}|y_2)P(y_2) > p(\mathbf{x}|y_1)P(y_1) \end{cases} \quad (2.15)$$

2.2.2 Lineare Diskriminanzanalyse nach Fisher

Das Ziel einer Diskriminanzanalyse ist die Zerlegung des Merkmalsraumes in disjunkte Bereiche anhand von Trainingsdaten. Gegeben sei ein Satz von Trainingsdaten $(S) = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_d, y_d)\}$, wobei für jeden Merkmalsvektor \mathbf{x}_i die Klassenzugehörigkeit $y_i \in \{1, 2\}$ bekannt ist. Weiter sei $\mu_{1,2}$ der Mittelwert der jeweiligen Klasse. Gesucht ist eine Diskriminanzfunktion $f(\mathbf{x})$, die über die Klassenzugehörigkeit einer beobachteten Größe mit Merkmalsausprägungen \mathbf{x}_i entscheidet.

Man unterscheidet zwei Fälle der Diskriminanzanalyse, in Abhängig davon, ob die Verteilung des Merkmalsvektors \mathbf{x} bekannt ist oder nicht [Rinne 2003]. Bei bekannter Verteilung kann die Diskriminanzfunktion mit Hilfe eines Maximum-Likelihood-Ansatzes aufgestellt werden. Da wir

die Kenntniss der Verteilung nicht bei allen Merkmalen für die Gebäudeextraktion voraussetzen können, beschränken wir uns auf den zweiten Fall, die Fisher-Diskriminanzanalyse.

Statt einen m -dimensionalen Merkmalsraum zu betrachten, reduziert die Fisher-Diskriminanzanalyse das Klassifikationsproblem auf den eindimensionalen Fall durch Projektion der \mathbf{x}_i auf eine Gerade durch den Ursprung.

$$\bar{x}_i = \mathbf{w}^T \mathbf{x}_i \tag{2.16}$$

beschreibt eine solche Projektion, falls $\|\mathbf{w}\| = 1$.

Gesucht ist diejenige Richtung \mathbf{w} der Projektionsgeraden, die eine optimale Separierbarkeit der \bar{x}_i gewährleistet. Abbildung (2.7) veranschaulicht das Prinzip der Diskriminanzanalyse nach Fisher. In Abbildung 2.7 (a) ist die Abszisse als Projektionsgerade gewählt. Für das gegebene Szenario ist diese Wahl der Projektionsgeraden ungeeignet, da die beiden projizierten Punktmengen nicht klar separierbar sind. Eine bessere Wahl ist in Abbildung 2.7 (b) veranschaulicht.

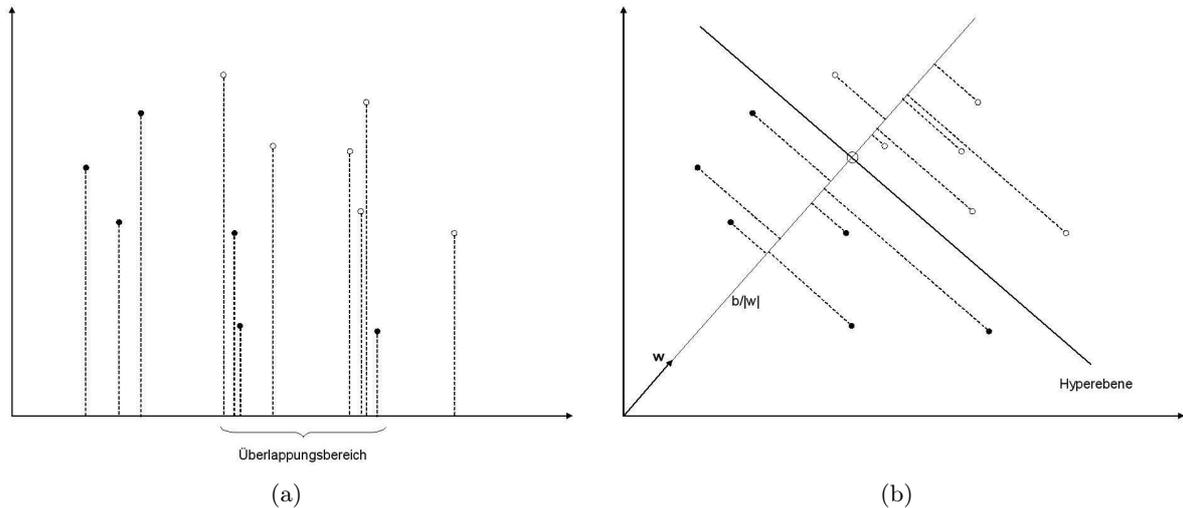


Abbildung 2.7: Schematische Darstellung der Diskriminanzanalyse nach Fisher. (a) Die Projektion auf die Abszisse (b) Die Projektionsgerade in Richtung \mathbf{w} ermöglicht eine optimale Projektion, da sie den Abstand zwischen den projizierten Punktmengen maximiert. Die auf \mathbf{w} senkrecht stehende Hyperebene $\mathbf{w}^T \mathbf{x} + b = 0$ trennt die beiden Klassen.

Die Separierbarkeit in einer Richtung \mathbf{w} kann durch den Abstand der Projektionsmittelpunkte beschrieben werden. Er kommt in der Merkmalsfunktion $F(\mathbf{w})$ zum Ausdruck.

$$F(\mathbf{w}) = \frac{|\bar{\mu}_1 - \bar{\mu}_2|}{\bar{s}_1^2 + \bar{s}_2^2} \tag{2.17}$$

Hierbei bezeichnet $\bar{\mu}_k = \frac{1}{n_k} \sum_{y_i=k} \bar{x}_i = \frac{1}{n_k} \sum_{y_i=k} \mathbf{w}^T \mathbf{x}_i = \mathbf{w}^T \mu_k$ den Mittelwert der Projektionen \bar{x}_i aus der Klasse $k = 1, 2$ und $\bar{s}_k = \sqrt{\sum_{y_i=k} (\bar{x}_i - \bar{\mu}_k)^2}$ die zugehörige Standardabweichung.

Durch die Einführung von Streumatrizen S_W zwischen den Klassen und S_B innerhalb einer Klasse, die aus der Stichprobe gewonnen werden können, erhält man eine andere Formulierung von (2.17).

Seien diese Kovarianzmatrizen gegeben durch

$$\begin{aligned} S_W &= \sum_{y_i=1} (\mathbf{x}_i - \mu_1)(\mathbf{x}_i - \mu_1)^T + \sum_{y_i=2} (\mathbf{x}_i - \mu_2)(\mathbf{x}_i - \mu_2)^T, \\ S_B &= (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T. \end{aligned}$$

Nach dem Einsetzen in (2.17), erhält man die Merkmalsfunktion

$$F(\mathbf{w}) = \frac{\mathbf{w}^T S_B \mathbf{w}}{\mathbf{w}^T S_W \mathbf{w}}. \quad (2.18)$$

Um eine bestmögliche Separierbarkeit zu gewährleisten, ist \mathbf{w} so zu wählen, dass $F(\mathbf{w})$ in (2.18) maximiert wird.

$$\mathbf{w} = S_W^{-1}(\mu_1 - \mu_2) \quad (2.19)$$

stellt die Lösung dieser Optimierungsaufgabe dar, falls S_W nicht singulär ist. Andernfalls kann das Problem mittels quadratischer Programmierung gelöst werden.

Um die Klassen voneinander trennen zu können, benötigt man neben dem Richtungsvektor \mathbf{w} noch einen Schwellwert b . Diese bestimmen gemeinsam die Diskriminantenfunktion

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b. \quad (2.20)$$

Der Schwellwert kann beispielsweise aus der Gleichung $\mathbf{w}^T \mu_1 + b = -(\mathbf{w}^T \mu_2 + b)$ gewonnen werden.

Für den Fall $f(\mathbf{x}) = 0$ wird durch die Diskriminantenfunktion eine optimale separierende Hyperebene im m -dimensionalen Merkmalsraum beschrieben. Sie ist optimal in dem Sinne, dass sie die Klassen bestmöglich trennt. Die Klassifikationsregel für zwei Klassen lautet daher

$$q(\mathbf{x}) = \begin{cases} 1 & \text{falls } f(\mathbf{x}) \geq 0 \\ 2 & \text{falls } f(\mathbf{x}) < 0. \end{cases} \quad (2.21)$$

2.2.3 Supportvektormaschinen (SVM)

Oftmals ist der bei einer Klassifikation betrachtete Merkmalsraum nicht linear separierbar. In solchen Fällen kann man jedoch durch eine (nicht-lineare) Transformation $\phi(\mathbf{x}_i)$ in einen Merkmalsraum mit höherer Dimension dort eine lineare Separierbarkeit erreichen [Vapnik 1995]. Diese Aufgabe erfüllen Supportvektormaschinen (SVM). Nachdem man die Daten im Raum höherer Dimension durch eine Hyperebene separiert hat (Abbildung 2.8), erhält man nach der Rücktransformation in den Ausgangsraum eine nicht-lineare trennende Funktion.

Gegeben sei ein Satz von Trainingsdaten $(S) = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_l, y_l)\}$ mit bekannter Klassenzugehörigkeit $y_i \in \{1, 2\}$.

Zum besseren Verständnis gehen wir zunächst vom einfachen Fall aus und nehmen linear separierbare Merkmalsvektoren an. Als Transformationsvorschrift wählen wir daher die identische Abbildung $\phi(\mathbf{x}_i) = \mathbf{x}_i$.

Analog zur linearen Diskriminante suchen wir eine trennende Hyperebene $f(\mathbf{x}) = \mathbf{w}'\mathbf{x} + b = 0$.

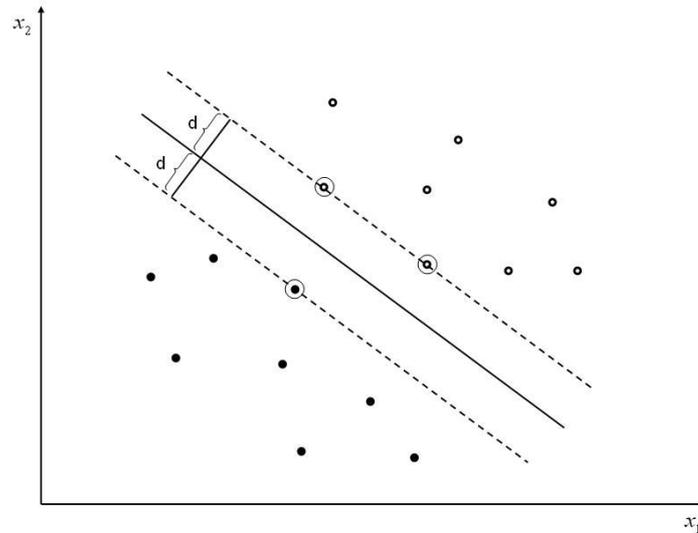


Abbildung 2.8: Hyperebene nach nicht-linearer Transformation mit einer Kernelfunktion. Vektoren auf dem gestrichelten Band sind die Supportvektoren.

In welcher Dimension diese gefunden wird, spielt zunächst keine Rolle. Die Suche nach einer solchen Hyperebene führt im strengen Fall, also für den Fall der Separierbarkeit der Daten, auf das „Maximal Margin“ Optimierungsproblem. Das bedeutet, dass der Abstand der \mathbf{x} von der Hyperebene maximiert wird.

Aus formalen Gründen führen wir für das Zweiklassenproblem die Größe $z_i \in \{-1, 1\}$ ein, mit $z_i = 1$ falls $y_i = 1$ und $z_i = -1$ falls $y_i = 2$. Somit lautet das Optimierungsproblem

$$\begin{aligned} & \text{minimiere}_{\mathbf{w}, b} && \frac{1}{2} \mathbf{w}' \mathbf{w} \\ & \text{unter der Bedingung} && z_i (\mathbf{w}' \mathbf{x} + b) \geq 1 \\ & && i = 1, \dots, l \end{aligned} \quad (2.22)$$

Die Lagrange-Funktion dieses Optimierungsproblems ist

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2} \mathbf{w}' \mathbf{w} - \sum_{i=1}^l \alpha_i (z_i (\mathbf{w}' \mathbf{x} + b) - 1), \quad (2.23)$$

mit den Lagrange-Multiplikatoren $\alpha_i \geq 0$. Mit Hilfe der Lagrange-Funktion lässt sich das Optimierungsproblem (2.22) in einer dualen Form darstellen. Dazu differenzieren wir $L(\mathbf{w}, b, \alpha)$.

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{w}} &= \mathbf{w} - \sum_{i=1}^l z_i \alpha_i \mathbf{x}_i, \\ \frac{\partial L}{\partial b} &= \sum_{i=1}^l z_i \alpha_i \end{aligned}$$

Unter der Berücksichtigung, dass

$$\frac{\partial L}{\partial \mathbf{w}} = \mathbf{0}, \quad \frac{\partial L}{\partial b} = 0 \quad (2.24)$$

gelten muss, erhalten wir nach Einsetzen der Ableitungen in (2.23)

$$L(\mathbf{w}, b, \alpha) = \sum_{i=1}^l \alpha_i = \frac{1}{2} \sum_{i=1}^l z_i z_j \alpha_i \alpha_j \mathbf{x}'_i \mathbf{x}_j. \quad (2.25)$$

Dies führt auf folgende quadratische Form des Optimierungsproblems

$$\begin{array}{ll} \text{maximiere}_{\alpha} & L(\mathbf{w}, b, \alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l z_i z_j \alpha_i \alpha_j \mathbf{x}'_i \mathbf{x}_j \\ \text{unter der Bedingung} & \sum_{i=1}^l z_i \alpha_i = 0 \\ & i = 1, \dots, l. \end{array}$$

Da die duale Form des Optimierungsproblems von b unabhängig ist, muss diese Größe auf anderem Wege bestimmt werden. Wir greifen auf die Karush-Kuhn-Tucker Bedingungen zurück. Sie besagen, dass die optimale Lösung von \mathbf{w}, b, α folgende Gleichung erfüllen muss.

$$\alpha_i (z_i (\mathbf{w}' \mathbf{x}_i + b) - 1) = 0, \quad i = 1, \dots, l \quad (2.26)$$

Diese Bedingung ist im Fall $\alpha_i \neq 0$ nur für diejenigen \mathbf{x}_i erfüllt, die am dichtesten an der Hyperebene liegen (Abbildung 2.8). Für alle anderen Vektoren sind die Lagrange-Multiplikatoren zwangsläufig gleich Null. Da nur die nahe an der Hyperebene gelegenen Merkmalsvektoren für die Optimierungsaufgabe von Bedeutung sind, werden sie als Supportvektoren bezeichnet.

Ist eine geringe Anzahl von Merkmalsvektoren nicht durch eine Hyperebene trennbar, kann man die strenge Abgrenzung der Klassen des „Maximal Margin“ Klassifikators durch Einführung von Schlupfvariablen $\xi_i, i = 1; \dots, d$ umgehen (Abbildung 2.9).

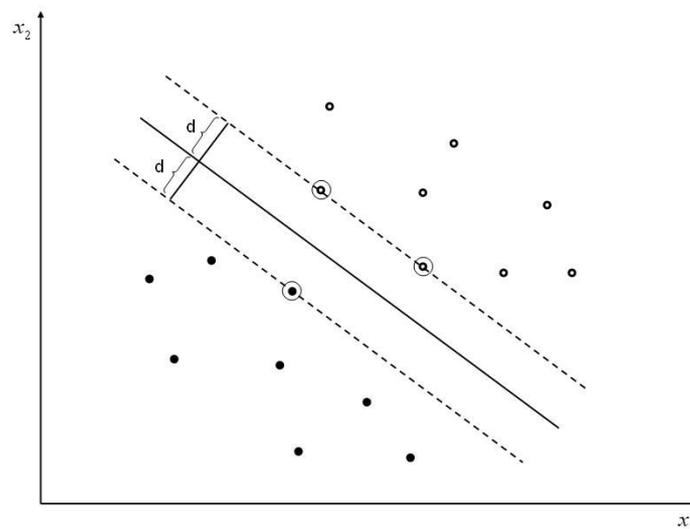


Abbildung 2.9: Darstellung von Merkmalsvektoren, die nicht linear separierbar sind. Einführung von Schlupfvariablen ξ_i für diejenigen Vektoren (x_1, x_2) , die den Grenzbereich (gestrichelt dargestellt) überschreiten.

Das führt auf das Optimierungsproblem des „Soft Margin“ Klassifikators

$$\begin{aligned} \text{minimiere}_{\xi, \mathbf{w}, b} \quad & \frac{1}{2} \mathbf{w}' \mathbf{w} + C \sum_{i=1}^l \xi_i^p \\ \text{unter der Bedingung} \quad & z_i(\mathbf{w}' \mathbf{x}_i + b) \geq 1 - \xi_i \\ & \xi_i \geq 0 \text{ für den Fall } p = 1 \\ & i = 1, \dots, l. \end{aligned}$$

Hierbei bezeichnet $C > 0$ eine Regulierungskonstante und p die zur Berechnung der Distanz verwendete Norm. Die L1-Norm ($p = 1$) liefert lineare Schlupfvariablen, bei Verwendung der L2-Norm liegen quadratische Schlupfvariablen vor. Je nach Art der Schlupfvariablen erhält man unterschiedliche Lagrange-Funktionen. In Analogie zum „Maximal Margin“ Klassifikator differenzieren wir diese und setzen die Ergebnisse für ξ , \mathbf{w} , b in die Lagrange-Funktion ein.

Wir betrachten zunächst den linearen Fall ($p = 1$). Hier ergibt sich

$$L(\mathbf{w}, b, \xi, \alpha, \mathbf{r}) = \frac{1}{2} \mathbf{w}' \mathbf{w} + C \sum_{i=1}^l \xi_i - \sum_{i=1}^l \alpha_i (z_i(\mathbf{w}' \mathbf{x}_i + b) - 1 + \xi_i) - \sum_{i=1}^l r_i \xi_i. \quad (2.27)$$

mit den Lagrange-Multiplikatoren $\alpha_i \geq 0$ und $r_i \geq 0$. Nach dem Ersetzen der ξ , \mathbf{w} , b ergibt dies

$$L(\mathbf{w}, b, \xi, \alpha, \mathbf{r}) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l z_i z_j \alpha_i \alpha_j \mathbf{x}'_i \mathbf{x}_j \quad (2.28)$$

und erhalten wir auch hier ein duales Optimierungsproblem.

$$\begin{aligned} \text{maximiere}_{\alpha} \quad & L(\mathbf{w}, b, \alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l z_i z_j \alpha_i \alpha_j \mathbf{x}'_i \mathbf{x}_j \\ \text{unter der Bedingung} \quad & \sum_{i=1}^l z_i \alpha_i = 0, \quad 0 \leq \alpha_i \leq C \\ & i = 1, \dots, l. \end{aligned}$$

Für den Fall quadratischer Schlupfvariablen ($p = 2$) ergibt sich

$$L(\mathbf{w}, b, \xi, \alpha) = \frac{1}{2} \mathbf{w}' \mathbf{w} + \frac{C}{2} \sum_{i=1}^l \xi_i^2 - \sum_{i=1}^l \alpha_i (z_i(\mathbf{w}' \mathbf{x}_i + b) - 1 + \xi_i). \quad (2.29)$$

Nach dem Ersetzen der aus der Differenzierung von L gewonnenen ξ , \mathbf{w} , b erhält man

$$\begin{aligned} L(\mathbf{w}, b, \xi, \alpha) &= \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l z_i z_j \alpha_i \alpha_j \mathbf{x}'_i \mathbf{x}_j - \frac{1}{2C} \alpha' \alpha \\ &= \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l z_i z_j \alpha_i \alpha_j \left(\mathbf{x}'_i \mathbf{x}_j + \frac{1}{C} \delta_{ij} \right). \end{aligned} \quad (2.30)$$

Hierbei bezeichnet δ_{ij} das Kronecker-Symbol, welches den Wert $\delta_{ij} = 1$ falls $i = j$ und sonst $\delta_{ij} = 0$ annimmt. Das führt schließlich auf das duale Optimierungsproblem

$$\begin{aligned} \text{maximiere}_\alpha \quad & L(\mathbf{w}, b, \xi, \alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l z_i z_j \alpha_i \alpha_j (\mathbf{x}'_i \mathbf{x}_j + \frac{1}{C} \delta_{ij}) \\ \text{unter der Bedingung} \quad & \sum_{i=1}^l z - i \alpha_i = 0, \quad \alpha_i \geq 0 \\ & i = 1, \dots, l. \end{aligned}$$

An dieser Stelle wollen wir uns wieder den nicht-linearen SVM zuwenden. Ersetzt wird die identische Abbildung durch eine nicht-lineare Abbildung $\phi(\mathbf{x}_i)$. Wir erhalten beispielsweise nach Gleichung (2.30) für den Fall quadratischer Schlupfvariablen

$$L(\mathbf{w}, b, \xi, \alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l z_i z_j \alpha_i \alpha_j \left(\phi(\mathbf{x}_i)' \phi(\mathbf{x}_j) + \frac{1}{C} \delta_{ij} \right). \quad (2.31)$$

als Lagrange-Funktion.

Das Skalarprodukt $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)' \phi(\mathbf{x}_j)$ wird als Kernel-Funktion bezeichnet. Ansätze für die Kernel-Funktion sind nach [Vapnik 1995] unter anderem

$$\begin{aligned} K(\mathbf{x}_i, \mathbf{x}_j) &= (\mathbf{x}'_i \mathbf{x}_j + 1)^d && \text{polynomial} \\ K(\mathbf{x}_i, \mathbf{x}_j) &= \exp(-\gamma |\mathbf{x}_i - \mathbf{x}_j|^2) && \text{radiale Basisfunktion} \\ K(\mathbf{x}_i, \mathbf{x}_j) &= \exp\left(-\frac{|\mathbf{x}_i - \mathbf{x}_j|^2}{2\sigma^2}\right) && \text{Gauß'sche radiale Basisfunktion} \end{aligned}$$

Damit erhalten wir das Optimierungsproblem für den transformierten Merkmalsraum höherer Dimension durch

$$\begin{aligned} \text{maximiere}_\alpha \quad & L(\mathbf{w}, b, \xi, \alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l z_i z_j \alpha_i \alpha_j (K(\mathbf{x}_i, \mathbf{x}_j) + \frac{1}{C} \delta_{ij}) \\ \text{unter der Bedingung} \quad & \sum_{i=1}^l z - i \alpha_i = 0, \quad \alpha_i \geq 0 \\ & i = 1, \dots, l. \end{aligned}$$

Das Problem der teilweise sehr hohen Dimension wird dadurch kompensiert, dass nur wenige Vektoren (die Supportvektoren) zur Beschreibung der Hyperebene erforderlich sind.

2.2.4 Kriterien zur Auswahl einer Klassifikation und Fehler

Die Auswahl einer Klassifikation hängt einerseits von der benötigten Vorinformation ab und andererseits von der Aussagekraft der Einteilung in die verschiedenen Klassen.

Für die Bayes-Klassifikation wird Vorinformation über die a priori Wahrscheinlichkeit $P(y)$ sowie die Likelihoodfunktion $p(\mathbf{x}|y)$ benötigt. Bei Trainingsdaten mit bekannter Klassenzugehörigkeit kann $P(y)$ relativ einfach aus der Häufigkeit des Auftretens der Klassen $y_j, j = 1, \dots, m$ bestimmt werden. Für jede einzelne Klasse ergibt sich

$$P(y_j) = \frac{\# \text{ Daten mit Klassenzugehörigkeit } y_j}{\# \text{ Daten gesamt}}. \quad (2.32)$$

Schwieriger zu bestimmen ist dagegen die Likelihoodfunktion. Sie kann entweder aus Histogrammen der einzelnen Klassen bestimmt werden oder mit Hilfe eines **Expectation-Maximization-Algorithmus** (EM-Algorithmus) iterativ berechnet werden.

Die Anwendung von Histogrammen kann ungeachtet der Kenntniss der zugrundeliegenden Verteilung angewandt werden. Sie hat aber den entscheidenden Nachteil, dass jeweils das gesamte Histogramm gespeichert werden muss. Probleme in Laufzeit und Speicherkapazität treten vor allem bei größeren Datenmengen auf.

Der EM-Algorithmus schätzt die Parameter einer Mischverteilung in zwei Schritten iterativ. Unter der Annahme von Näherungswerten für die Parameter der Verteilung wird zunächst der Erwartungswert berechnet. Anschließend wird dieser maximiert und das Ergebnis als neuer Näherungswert verwendet. Dieses Vorgehen setzt die Kenntnis des Verteilungstyps voraus. Nur so können die zu schätzenden Parameter angesetzt werden.

Vielfach ist die Verteilung jedoch unbekannt und aus der zu geringen Datenmenge nicht ableitbar. Die Annahme einer Normalverteilung stellt dann nur eine grobe Näherung dar. In solchen Fällen sind andere Klassifikationen meist besser geeignet.

Verwendet man die lineare Diskriminante nach Fisher oder die Supportvektormaschinen, dann benötigt man keine Vorinformation über die Verteilung. Die SVM bedürfen lediglich der Auswahl einer Kernelfunktion (vgl. Kapitel 2.2.3).

Bei der Wahl der trennenden Funktion zwischen den Klassen sollten zwei Punkte berücksichtigt werden. Zum einen sollen Fehler bei der Klassenzuordnung vermieden werden. Andererseits ist ein Modell, das die Klassen perfekt trennt, sich aber nur an einen speziellen Fall von Trainingsdaten anpasst, nicht geeignet. Ein Kompromiss zwischen Generalisierbarkeit und optimaler Klassenzuordnung ist daher erforderlich [Duda et al. 2001]. Datenmengen, die aufgrund der Merkmalsauswahl nicht separierbar sind, führen zwangsläufig zu Fehlern in der Klassifikation. Fehler treten dann auf, wenn eine Entscheidung $q_i \in \{1, \dots, m\}$ getroffen wird, die einen Vektor von Merkmalsausprägungen zur i -ten Klasse zuordnet, die richtige Klasse aber $y_j \in \{1, \dots, m\}$ wäre. Nach [Franc, Hlaváč 2004] führen wir als tatsächlichen Fehler der Klassifikation folgende Größe ein.

$$\epsilon(q_i, y_j) = \frac{\sum_j \# \text{Entscheidungen } q_i \neq y_j}{\# \text{Daten gesamt}} \quad (2.33)$$

Analog dazu können wir den Fehler für eine einzelne Klasse durch

$$\epsilon_j(q_i) = \frac{\# \text{Entscheidungen } q_i \neq y_j}{\# \text{Daten der Klasse } y_j} \quad (2.34)$$

beschreiben.

Durch die Wahl eines geeigneten Klassifikators sowie einer ausreichend großen Anzahl von Merkmalen können Fehler verringert werden. Kapitel 3 befasst sich mit der Merkmalsauswahl. In Kapitel 4 werden die Ergebnisse der drei vorgestellten Klassifikationsmethoden präsentiert und bewertet.

Kapitel 3

Analyse des Skalenraumes

Die Extraktion von Gebäuden aus terrestrischen Farbbildern erfordert die Auswahl geeigneter Bildmerkmale. Für diese Arbeit stellen die in Kapitel 2 definierten stabilen Regionen im Skalenraum die Grundlage der Bildanalyse dar. Das Vorhandensein stabiler Regionen sagt alleine jedoch noch nichts darüber aus, ob sich im Bild ein Gebäude befindet. Aus diesem Grund werden verschiedene Merkmale der stabilen Regionen sowie deren Nachbarregionen für die Gebäudeextraktion herangezogen. Nachstehend werden neben den untersuchten Merkmalen die für die Analyse benötigten Werkzeuge sowie die Ausgangsdaten vorgestellt.

3.1 Werkzeuge

Das Programm zur Gebäudeextraktion ist in MATLAB implementiert, da für die zu realisierenden Aufgaben einige nützliche Toolboxen in MATLAB vorliegen. Bei der Lösung der unterschiedlichen Aufgaben von der Vorverarbeitung bis zur Klassifikation kommen folgende Toolboxen als Hilfsmittel zum Einsatz:

1. Image Rectify Tool des Fachbereichs Photogrammetrie an der Rheinischen Friedrich-Wilhelms-Universität Bonn
(abrufbar unter <http://www.ipb.uni-bonn.de/projects/etrimis/internal/download>)
Betrachtet man Gebäude in Bildern, so sind die in der Realität parallelen Linien nicht mehr parallel, sondern laufen auf Fluchtpunkte zu. Da dies die Vergleichbarkeit der Bilder beeinträchtigt, ist eine Entzerrung der Bilder notwendig. Das Image Recognition Tool ermöglicht eine solche Entzerrung.
2. LabelMe Toolbox des Massachusetts Institute of Technology (MIT)
(abrufbar unter <http://labelme.csail.mit.edu/LAbeLMeToolbox.zip>)
Für die Klassifikation werden Trainingsdaten \mathbf{x}_i mit bekannter Klassenzugehörigkeit y_i benötigt. Die LabelMe Toolbox ist ein Werkzeug, mit dem man Bildregionen mit Labels versehen kann. In einem Vorverarbeitungsschritt können den Regionen so Klassen manuell zugewiesen werden.
3. Annotation Toolbox (Version 1-7) des Fachbereichs Photogrammetrie an der Rheinischen Friedrich-Wilhelms-Universität Bonn
(abrufbar unter <http://www.ipb.uni-bonn.de/projects/etrimis/internal/download>)

Die Annotation Toolbox baut auf der LabelMe Toolbox auf. Sie ermöglicht die Auswahl geschlossener Polygone in Bildern zur Darstellung von Klassengrenzen. Wir sprechen dabei auch von Annotation. Die Verwaltung der Polygone erfolgt mittels XML-Dokumenten. Gespeichert werden neben der Klassenzugehörigkeit der Polygone auch Informationen über deren Lage. Für jede Wasserscheidenregion kann man somit die Klassenzugehörigkeit über eine Zuordnung ihrer Lage bezüglich der Polygone ermitteln.

4. Statistical Pattern Recognition Toolbox des Center of Machine Perception an der Czech University Prag

(abrufbar unter http://cmp.felk.cvut.cz/cmp/cmp_software.html)

Die Statistical Pattern Recognition Toolbox umfasst eine Sammlung statistischer Methoden zur Mustererkennung. Daneben werden auch Werkzeuge zur Visualisierung der Klassifikationsergebnisse bereitgestellt.

3.2 Ausgangsdaten

Analysiert werden 60 terrestrisch aufgenommene Farbbilder von Gebäuden. Diese stammen sowohl aus eigenen Aufnahmen als auch aus dem Archiv des Bonner Fachbereichs für Photogrammetrie. Die Bilder umfassen unterschiedliche Aufnahmesituationen und spiegeln das breite Formen- und Farbenspektrum der Gebäude wider. Dargestellt sind sowohl einzelne Häuser als auch Häuserzeilen und Ausschnitte von Häusern. Zudem liegen die Bilder in verschiedenen Auflösungen vor. Wie die Aufnahmen dennoch vergleichbar gemacht werden können, wird in Kapitel 3.3.1 beschrieben.

Für die Klassifikation werden Trainings- und Testdaten benötigt. Dafür wird jeweils die Hälfte des Bildmaterials eingesetzt. Als Kriterium zur Auswahl der Trainingsdaten dient in erster Linie die Eigenschaft, dass die Daten möglichst allgemeingültig sind. Wegen der starken Variation der Gebäude in Form und Farbe ist die Definition eines Standardhauses jedoch problematisch. Als Testdaten verwenden wir daher gezielt vielfältig ausgeprägte Gebäude, die wir nach Gruppen sortiert auf die ihre Detektierbarkeit untersuchen.

3.3 Vorverarbeitung der Bilddaten

Bevor die Bilddaten der Merkmalsextraktion unterzogen oder klassifiziert werden können, sind einige Schritte vorab notwendig.

Zunächst muss die Vergleichbarkeit der unterschiedlichen Bilder hergestellt werden. Weiterhin ist der Skalenraum wie in Kapitel 2.1 beschrieben aufzubauen und nach Anwendung des Wasserscheidenalgorithmus die Überprüfung der Regionen auf Stabilität vorzunehmen. Schließlich benötigen wir für die Trainingsdaten noch eine Beschreibung, welche der Wasserscheidenregionen zu der Objektklasse „Gebäude“ gehören.

3.3.1 Entzerrung

Die vorliegenden Bilddaten sind zunächst nicht vergleichbar, da sie aus unterschiedlichen Aufnahmesituationen heraus entstanden sind. Durch Variation des Aufnahmewinkels oder des Abstandes der Kamera zum Objekt treten andere Auflösungen auf. Zudem treten innerhalb eines Bildes unterschiedliche Auflösungen auf, da parallele Linien im Bild in der Regel nicht parallel verlaufen, sondern verzerrt dargestellt werden.

Dies alles beeinträchtigt die Vergleichbarkeit von Wasserscheidenregionen. Eine vorangestellte Entzerrung des Bildes schafft Abhilfe. Das Image Rectify Tool ermöglicht eine solche Entzerrung durch die Angabe zweier Linienpaare, die je zu einem unterschiedlichen Fluchtpunkt laufen. Entzerrt wird jeweils diejenige Seite des Gebäudes mit der geringsten Verzerrung. Die Entzerrung der anderen Seiten würde eine zu starke Deformation des Objektes nach sich ziehen.

Nach der Entzerrung weisen die einzelnen Bilder in allen Pixeln die gleiche Auflösung auf. Zwischen den Bildern besteht jedoch weiterhin ein Unterschied in der Auflösung. Für die vorliegenden entzerrten Bilder beträgt die Auflösung zwischen 1 und 7 cm pro Pixel. Dieser Aspekt ist beim Vergleich der Skalierungsstufen zu berücksichtigen.

3.3.2 Skalenwahl und Auflösung

Im Skalenraum wird das Originalbild in verschiedenen Auflösungsstufen dargestellt, um bestimmte Strukturen auszublenden. Bei der Auswahl der Skalierung ist die Auflösung der Objekte im Originalbild zu berücksichtigen, die hier 1-7 cm pro Pixel beträgt.

Unter der Voraussetzung dieser Startwerte, wählen wir $\sigma_{max} = 4$ als maximales Skalenniveau. Das entspricht einer maximalen Filtergröße von 175 cm bezogen auf das reale Objekt.¹ Ein solcher Filter reicht aus, um Fenster von Gebäuden auszublenden. Größere Filter würden die Informationen zu stark vermischen und nicht zu aussagekräftigen Bildern führen.

Wir verwenden 20 logarithmisch angeordnete Skalen, für die der Wasserscheidenalgorithmus auch Kapitel 2.1.1 berechnet wird. Im Bereich zwischen $\sigma = 0$ und $\sigma = 1$ liefert die Segmentierung mit dem Wasserscheidenalgorithmus eine starke Zergliederung der geglätteten Bilder. Stabile Regionen sind daher in diesem Bereich nicht aussagekräftig. Aus diesem Grund werden Skalen erst ab einer Größe von $\sigma = 1$ betrachtet. Wir erhalten somit $N_1 = 0$ Skalen mit $\sigma \in]0, 1[$ und $N_2 = 20$ Skalen mit $\sigma \in [1, 4]$. Zusammen mit dem Originalbild ergibt dies $s_{max} = 21$ Stufen im Skalenraum.

Abbildung 3.1 zeigt einen Ausschnitt des Skalenraums und das Verhalten der Wasserscheidenregionen über den Skalenraum.

¹Die Filtergröße berechnet sich aus der Formel $d = \lceil 2 \cdot 3\sigma \rceil$ nach Kapitel 2.1.2. Für den Fall einer geraden Pixelanzahl d , wird noch ein Pixel hinzuaddiert. Anschließend erfolgt noch die Umrechnung in cm.



(a) Wasserscheiden, Originalbild



(b) Stabile Regionen, Originalbild

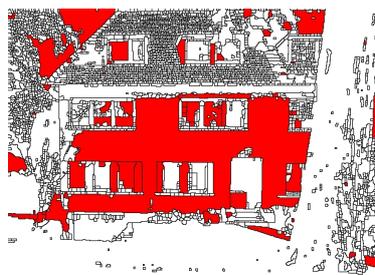
(c) Wasserscheiden, $\sigma = 1$ (d) Stabile Regionen, $\sigma = 1$ (e) Wasserscheiden, $\sigma = 2.07$ (f) Stabile Regionen, $\sigma = 2.07$ (g) Wasserscheiden, $\sigma = 2.98$ (h) Stabile Regionen, $\sigma = 2.98$ (i) Wasserscheiden, $\sigma = 4$ (j) Stabile Regionen, $\sigma = 4$

Abbildung 3.1: Darstellung der Regionen im Skalenraum. Das Originalbild (a) weist eine Auflösung von 2.08 [cm] auf. Links befinden sich die durch Filterung des Originalbildes entstandenen Bilder in unterschiedlichen, durch den Skalenparameter σ gekennzeichneten Auflösungsstufen. Die gelben Wasserscheiden partitionieren das jeweilige Bild. Auf der rechten Seite sind die Wasserscheiden schwarz dargestellt. Rote Flächen kennzeichnen stabile Regionen.

3.3.3 Beschreibung der Klassenzugehörigkeit der Trainingsdaten

Der Einsatz von Trainingsdaten setzt eine bekannte Klassenzugehörigkeit voraus.

Mit Hilfe der Annotation Toolbox können im Bild Gebiete gekennzeichnet werden, die zu einer Klasse gehören. Diese werden wie in Kapitel 3.1 beschrieben als Polygone in einem XML-Dokument gespeichert. Da wir nur die beiden Klassen Gebäude und Hintergrund betrachten, ist es ausreichend, die Gebäude durch Polygone zu kennzeichnen. Markiert werden ausschließlich die sichtbaren Teile der Gebäude, da Verdeckungen das Klassifikationsergebnis negativ beeinflussen.

Von Interesse ist nun, welche der stabilen Regionen in einem Gebäude liegen. Eine einfache „Punkt-in-Polygon“-Suche führt hier zum Ziel. Ein Punkt liegt genau dann in einem Polygon, wenn die Anzahl der Schnittpunkte einer beliebigen, an dem Punkt beginnenden Halbgeraden mit dem Polygon ungerade ist.

Diesen Test für jedes Pixel einer stabilen Region anzuwenden ist allerdings sehr rechenintensiv, bedenkt man, dass die stabilen Regionen vor allem bei geringer Auflösung sehr groß werden können. Die Idee ist daher, den Test auf die Randpixel zu beschränken. Randpixel erhält man leicht durch Erodieren des Bildes. Im Prinzip bedeutet dies, dass wir eine stabile Region auswählen und diese von Außen her schälen. Die Differenz der ursprünglichen Region und der erodierten Region ergibt den Rand.

Aufgrund von Ungenauigkeiten bei der Annotation der Bilder, die wir zu 5 Pixel annehmen, erodieren wir die stabile Region zunächst um diese 5 Pixel, bevor wir wie oben beschrieben den Rand bestimmen. Dieses Vorgehen dient dazu, Schnitte der stabilen Regionen mit dem Polygon zu vermeiden. Durch die Glättung des Bildes kann es dennoch vorkommen, dass stabile Regionen teilweise innerhalb und teilweise außerhalb des Gebäudes liegen. In solchen Fällen ordnen wir die betreffenden Regionen dem Gebäude zu, sobald mindestens 60% der Randpixel der Region im Gebäude liegen. Dabei ist zu berücksichtigen, dass in den höheren Skalen stabile Regionen durch Verschmelzung entstehen können, die im Gebäude liegen, aber über dessen Rand hinausragen. In solchen Fällen ist es möglich, dass weniger als 60% der Randpixel im Gebäude liegen. Das führt zu einer Zuordnung der stabilen Region zum Hintergrund. Wir treffen hier jedoch keine Fallunterscheidung, da wir solche Regionen nicht als aussagekräftig erachten.

3.4 Untersuchte Merkmale

Der Fokus liegt in diesem Kapitel auf der Auswahl von Merkmalen für die Gebäudeextraktion sowie der Umsetzung der Merkmalsextraktion. Eine Darstellung und Diskussion der Ergebnisse der Merkmalsextraktion erfolgt in Kapitel 4.

Einzelne Merkmale reichen alleine nicht für die Detektion von Gebäuden oder auch nur Gebäudeteilen aus. Eine Vielzahl von Merkmalen, kombiniert zu einem Merkmalsvektor verspricht bessere Ergebnisse. Wie einleitend bereits in Kapitel 1 erwähnt, führt ein großer Merkmalsvektor jedoch schnell zu einem hohen Rechen- und Speicheraufwand. Um die Datenmenge zu reduzieren, sollten zwei Aspekte beachtet werden.

Zum einen sind möglichst gut geeignete Merkmale zu verwenden und diese in geeigneter Weise zu kombinieren. Was unter geeigneten Merkmalen zu verstehen ist und wie man diese unter einer Fülle von Merkmalen auswählt, wird in [Guyon, Elisseeff 2003] beschrieben. Wir wählen im Fol-

genden einige für den menschlichen Betrachter einsichtige Merkmale. Die Auswahl der Merkmale umfasst den Umfang, den Flächeninhalt und den Formfaktor der stabilen Regionen sowie Eigenschaften benachbarter Regionen bezüglich Anzahl, Größe und Farbton. Diese Merkmale werden nachstehend näher erläutert.

Zum anderen kann die Datenmenge reduziert werden, wenn man die Zahl der betrachteten Skalen abhängig vom jeweiligen Merkmal variiert. Einige Merkmalsausprägungen ändern sich beispielsweise nicht über einen Teil oder sogar den gesamten Skalenraum. Wir sprechen in diesem Fall von einer Skalenunabhängigkeit der Merkmale, andernfalls von einer Skalenabhängigkeit. Für die Untersuchungsergebnisse zur Skalen(un)abhängigkeit sei ebenfalls auf Kapitel 4 verwiesen.

3.4.1 Merkmale der stabilen Regionen

Zunächst betrachten wir die stabilen Regionen getrennt von ihrer Umgebung. Wir stellen nun Merkmale auf, die sich nur auf die stabilen Regionen beziehen. Dazu gehören neben Umfang, Fläche und Formfaktor auch ein Maß für die Homogenität der Regionen.

3.4.1.1 Umfang, Flächeninhalt und Formfaktor

Als erste Eigenschaft der stabilen Regionen untersuchen wir deren Größe. Motiviert ist die Wahl dieses Kriteriums in der Betrachtung eines idealen Gebäudes mit homogener Fassade, die sich gut vom Hintergrund abgrenzt. Bei einem solchen Gebäude müssten Wasserscheiden des quadratischen Gradientenbildes nur am Rand und im Bereich von Türen und Fenstern auftreten. Im Vergleich zu stabilen Regionen im Bereich von Vegetation vermuten wir daher größere Regionen im Bereich von Gebäuden. Ob diese Annahme auch für nicht-ideale Gebäude zutrifft, werden die Ergebnisse in Kapitel ?? zeigen.

Die Größe einer stabilen Region kann entweder über den Umfang oder den Flächeninhalt bestimmt werden. Abhängig von der Form einer Region können Umfang und Flächeninhalt variieren. Aus diesem Grund werden beide Merkmale untersucht.

Der Flächeninhalt F kann einfach über die Summe der Pixel einer Region bestimmt werden. Um auf die jeweilige stabile Region zugreifen zu können, gehen wir vom gelabelten Bild L der stabilen Regionen aus und durchlaufen die Labels von 1 bis zum maximalen Label n . Vor der Aufsummierung der Pixel, werden die Pixel der aktuellen Region auf 1 gesetzt und jeweils alle anderen Labels ausgeschaltet (d.h. diese Pixel werden auf 0 gesetzt).

```

for i = 1:n
    R = zeros(size(L));
    R(find(L==i)) = 1;
    F(i) = sum(sum(R));
end

```

Für den Umfang betrachten wir ebenfalls Binärbilder, die nur eine der stabilen Regionen enthalten (Gebietspixel = 1, Hintergrund = 0). Der erste Schritt zur Auswahl der zu betrachtenden Region erfolgt daher analog zur Regionenauswahl für die Berechnung des Flächeninhalts.

Zur Bestimmung des Umfangs greifen wir auf eine Look-Up-Tabelle zurück (vgl. Abbildung 3.2).

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
du	0	1	1	1	1	1	2	1	1	2	1	1	1	1	1	0	

Abbildung 3.2: Look-Up Tabelle des Umfangs

Diese Tabelle gibt Aufschluss über den Zuwachs des Umfangs dU für die jeweils beobachtete Konstellation von Pixeln in einem (3×3) -Ausschnitt.

Nach Anwendung des Filters

$$f_U = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 4 & 8 \end{bmatrix} \tag{3.1}$$

auf das Binärbild erhält man über die Werte 1 bis 15 einen eindeutigen Zuwachs dU für jede Pixelposition. Die Summe $U = \sum dU$ liefert den gesamten Umfang.

Zusätzlich zu Umfang und Flächeninhalt wird ein Formfaktor betrachtet. Der Formfaktor beschreibt wie rund oder länglich eine betrachtete Region ist. Er berechnet sich direkt aus dem Umfang U und dem Flächeninhalt F zu

$$k = \frac{U^2}{4\pi F} \geq 1.$$

Aufgrund der offensichtlichen Variabilität und teilweisen Verdeckung von Gebäuden erscheint dieses Merkmal großräumig (d.h. für die gesamte Fassade) ungeeignet. Selbst bei einer rechteckigen Fassade bestehen Gebäude nur in wenigen Bildern aus einer einzigen stabilen Region. Dennoch wird der Formfaktor als Merkmal aufgenommen. Die Untersuchung zielt speziell auf den kleinräumigen Bereich von Fenstern und Türen. Als ideale rechteckige Form angenommen, mit einem Seitenverhältnis von 1 : 1 bis 1 : 2 sollte der Formfaktor einen Wert zwischen 1.27 und 1.43 annehmen (vgl. Abbildung 3.3).

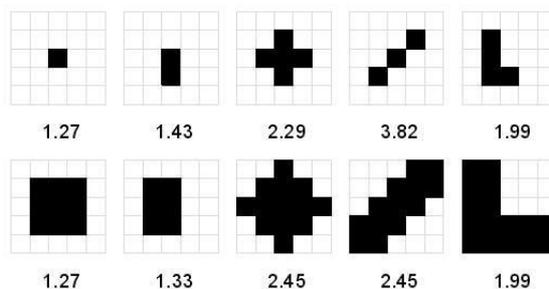


Abbildung 3.3: Darstellung des Formfaktors für unterschiedliche geometrische Figuren.

3.4.1.2 Homogenität

Für die nachstehenden Überlegungen erweist es sich als vorteilhaft, nur einen statt drei Bildlayer zu betrachten. Betrachten wir daher Grauwertbilder $b_{gr}(x, y)$. Diese kann man leicht aus Farbbildern ableiten, indem man nur deren Helligkeitswerte übernimmt.

Der Wasserscheidenalgorithmus aus Kapitel 2.1.1 unterdrückt eine Übersegmentierung durch das Setzen eines Schwellwertes. Dies führt dazu, dass innerhalb der Wasserscheidenregionen unterschiedliche Grauwerte auftreten. Die Regionen sind also nicht homogen. Das Merkmal der Homogenität ist aus der Überlegung heraus entstanden, dass Regionen im Bereich von Fassaden geringere Grauwertunterschiede aufweisen müssten als Regionen im Bereich von Vegetation. Abbildung 3.4 legt dies nahe. Im Bereich der hellen Fassade fallen die Grauwertunterschiede geringer aus als im Bereich der Vegetation rechts im Bild. Diese Darstellung zeigt zugleich aber auch eine Problematik dieses Merkmals. Stärkere Grauwertunterschiede können auch im Bereich von Schieferverkleidung, wie sie im Dachbereich vorliegt, oder bei Klinkerfassaden auftreten.



Abbildung 3.4: Darstellung eines Grauwertbildes $b_{gr}(x_1, x_2)$ mit Wasserscheidenregionen. Regionen im Bereich der hellen Fassade weisen eine größere Homogenität auf als Regionen im Bereich von Vegetation.

Als Maß für die Homogenität betrachten wir die Standardabweichung der Grauwerte $(b_{gr}^j(x_1, x_2)_i)$ mit $(x_1, x_2)_i \in \mathcal{R}_j$ einer Region.

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (b_{gr}^j(x_1, x_2)_i - \bar{b}_{gr}^j)^2} \quad (3.2)$$

Hierbei bezeichnet n die Anzahl der Pixel in einer Region \mathcal{R}_j und $\bar{b}_{gr}^j = \frac{1}{n} \sum_{i=1}^n b_{gr}^j(x_1, x_2)_i$ den Mittelwert der Region.

3.4.2 Merkmale direkter Nachbarn

Wenden wir uns nun den Nachbarschaften der stabilen Regionen zu.

Wir haben bereits gesehen, dass eine Segmentierung mit Hilfe des Wasserscheidenalgorithmus eine vollständige Partitionierung des Bildes liefert (Kapitel 2.1.1). Aus diesem Grund können wir Nachbarschaften zwischen zwei Regionen wie folgt definieren.

Als „direkt benachbart“ bezeichnen wir zwei Regionen, die eine gemeinsame Wasserscheide als Rand haben. Nachstehend betrachten wir speziell die direkten Nachbarn von stabilen Regionen. Diese Nachbarn müssen selber nicht stabil sein.

Zur Bestimmung der direkten Nachbarn können wir diese Definition nutzen. Wir müssen lediglich die Randpixel einer stabilen Region durchlaufen und deren angrenzende Regionen abgreifen. Die Identifikation der Regionen erfolgt mit Hilfe des Labelings.

Zu Steigerung der Effizienz werden nicht alle Randpixel durchlaufen. Es genügt, Kreuzungspunkte der Randpixel zu betrachten. Als Kreuzungspunkte werden diejenigen Pixel bezeichnet, an denen mehr als zwei Randstücke zusammenstoßen.

Es treten jedoch auch andere Fälle auf, an denen diese Filterung einen Wert größer oder gleich 3 ergibt, allerdings kein Kreuzungspunkt vorliegt. Um solche Fälle auszuschließen, wird als zusätzliches Kriterium das Vorhandensein eines Randpixels im Mittelpunkt des 3×3 -Fensters gefordert. Das bedeutet, dass an den Stellen, an denen im gefilterten Bild ein Wert größer oder gleich 3 auftritt, im Binärbild des Randes eine 1 stehen muss.

Zum einfachen Abgriff werden benachbarte Regionen in einer Adjazenzmatrix A_{dir} der Größe $(n \times n_r)$ gespeichert. Hierbei bezeichnet n die Anzahl der stabilen Regionen und n_r die Anzahl aller Regionen, so dass die stabilen Regionen zeilenweise und alle Regionen spaltenweise abgetragen werden. Die Adjazenzmatrix könnte folgendermaßen aussehen.

$$A_{dir (n \times n_r)} = \begin{bmatrix} 0 & 0 & 1 & 0 & \dots & 1 \\ 1 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & 1 & 0 & \dots & 0 \end{bmatrix} \quad (3.3)$$

Direkte Nachbarn sind durch eine 1 in der Adjazenzmatrix gekennzeichnet. In diesem Beispiel sind somit die Regionen mit den Labels 3 und n_r direkte Nachbarn der ersten stabilen Region.

3.4.2.1 Flächeninhalt der direkten Nachbarn

Die Motivation in der Betrachtung von Nachbarregionen liegt darin, dass Merkmale einer einzelnen Region in der Regel nicht so aussagekräftig sind. Dies sticht vor allem beim Flächeninhalt hervor. Große stabile Regionen treten nicht nur im Bereich von Gebäuden, sondern auch im Himmel und in Schattenbereichen auf. In diesen Fällen handelt es sich jedoch meist um einzelne große Regionen. Wir stellen daher die Hypothese auf, dass flächenmäßig große Nachbarregionen eher in Gebäuden zu finden sind.

Dies führt uns auf die Betrachtung des Flächeninhalts der Nachbarregionen als Klassifikationsmerkmal. Wir benötigen jedoch für jede stabile Region eine eindeutige Größe. Hier bietet sich der Median der Flächeninhalte aller Nachbarregionen an sowie ein 5%– und ein 95%–Quantil. Mit diesen Merkmalen können wir jedoch noch keine Aussage treffen, ob die stabile Region, de-

ren Nachbarn wir untersuchen, ebenfalls einen großen Flächeninhalt aufweist. Für ein weiteres Merkmal setzen wir daher den Flächeninhalt der stabilen Region zu dem Flächeninhalt ihrer Nachbarn ins Verhältnis.

3.4.2.2 Farbtöne der direkten Nachbarn

Eine Auswahl der Farbtöne der direkten Nachbarn einer stabilen Region als Merkmal ist motiviert in der Betrachtung des folgenden Beispiels.

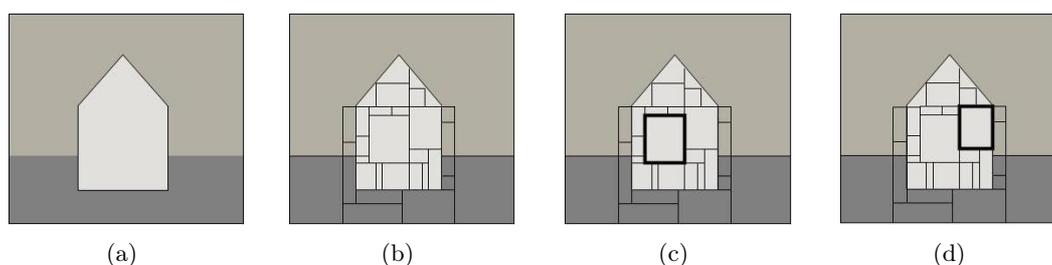


Abbildung 3.5: Idealfassade

Gegeben sei eine idealisierte Fassade, wie in Abbildung 3.5(a) dargestellt. Durch Betrachtung einiger Gebäude erkennt man, dass in vielen Fällen große Teile der Fassade (oft sogar die ganze Fassade) nur geringfügig in der Farbe variieren. Angenommen, die idealisierte Fassade besteht nach Anwendung des Wasserscheidenalgorithmus aus Kapitel 2.1.1 aus mehreren Regionen (Abbildung 3.5(b)), von denen eine stabil ist. Liegt die stabile Region im Inneren des Gebäudes (Abbildung 3.5(c)), so haben benachbarte Regionen vermutlich eine ähnliche Farbe. Befindet sich die stabile Region dagegen im Rand der Fassade, so haben die benachbarten Regionen aus dem Gebäude ebenfalls die gleiche Farbe. Bei benachbarten Regionen, die nicht zum Gebäude gehören, ist es dagegen weniger wahrscheinlich, dass die gleiche Farbe auftritt.

Als ein Klassifikationsmerkmal verwenden wir den prozentualen Anteil der Nachbarregionen einer stabilen Region, die einen ähnlichen Farbton haben.

Es bleibt an dieser Stelle noch zu klären, was wir unter einer „ähnlichen“ Farbe verstehen.

Bei den Ausgangsdaten handelt es sich um RGB-Bilder. Die Farbanteile R = Rot, G = Grün und B = Blau eignen sich jedoch nicht gut für einen Farbvergleich, da es sich um drei Merkmale handelt. Besser geeignet ist ein Vergleich in einer einzelnen Größe. Aus diesem Grund wird das RGB-Bild zunächst in den HSV-Farbraum (vgl. Abbildung 3.6) transformiert. Dieser enthält die Größen Farbton (H = hue), Sättigung (S = saturation) und Grauwert (V = value).

Eine solche Transformation behandeln u.a. [Foley et al. 1995]. Für unsere Analyse betrachten wir ausschließlich den Farbton. Dieser nimmt Werte aus dem Intervall $\mathcal{I} = [0, 1]$ an (siehe Abbildung 3.7).

Auf einen Vollkreis transformiert entspricht dies dem Farbton Rot bei einem Winkel von 0° , Grün bei 120° und Blau bei 240° .

Da Regionen im Skalenraum in der Realität keinen homogenen Farbton aufweisen, berechnen wir den Mittelwert für jede Region und vergleichen diesen.

Wir sehen Farbtöne als „ähnlich“ an, wenn sie sich im Mittel um weniger als $dh = 0.05$ unterscheiden.

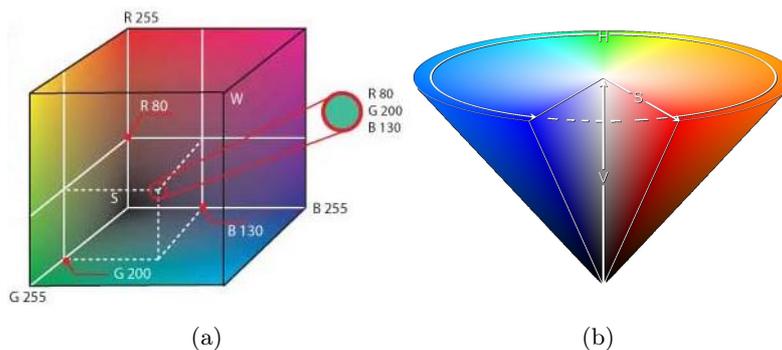


Abbildung 3.6: (a) RGB-Farbraum ($R = \text{Rot}$, $G = \text{Grün}$, $B = \text{Blau}$)
 (http://upload.wikimedia.org/wikipedia/commons/0/03/RGB_farbwuerfel.jpg)
 (b) HSV-Farbraum ($H = \text{Farbton}$, $S = \text{Sättigung}$, $V = \text{Grauwert}$)
 (http://upload.wikimedia.org/wikipedia/commons/e/ea/HSV_cone.png)

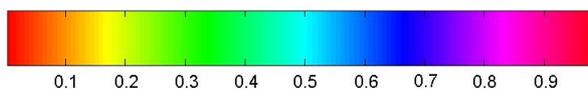


Abbildung 3.7: Dezimale Darstellung der HSV-Farbwerte

Da Rotwerte sowohl im Bereich von $h = 0$ als auch bei $h = 1$ auftreten, wird dieser Fall abgefangen und getrennt behandelt. Durch Addition einer 1 auf die kleinen Farbwerte und Verknüpfung der so neu gewonnenen Farbskala mit der alten, erhalten benachbarte Rotwerte einen ähnlichen Wert. Betrachtet wird nun das Intervall $\mathcal{I}^* = [0.5, 1.5]$. Nach der Bildung des Mittelwertes der Farbtöne erfolgt für den Vergleich der benachbarten Regionen die Rücktransformation auf das Intervall \mathcal{I} .

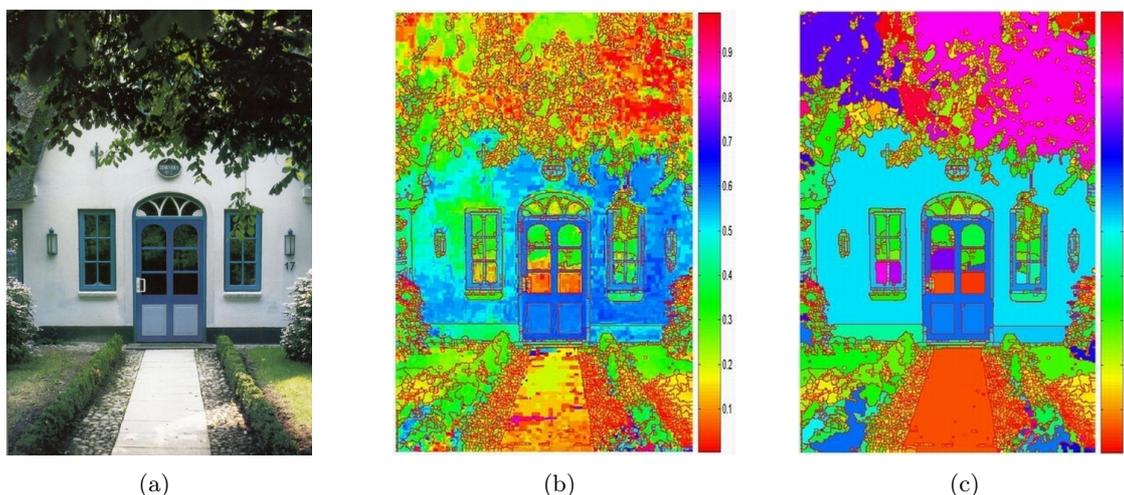


Abbildung 3.8: HSV-Farbraum (a) Originalbild, (b) Farbtöne der Wasserscheidenregionen im HSV-Farbraum, (c) Mittelwert der Farbtöne der Wasserscheidenregionen

Bislang haben wir uns nur dafür interessiert, ob Farbdifferenzen im Rahmen des Schwellwertes von $dh = 0.05$ liegen. Ebenfalls interessant sind die Beträge der Farbdifferenzen. Als weitere Merkmale werden daher der Median der Farbdifferenzen sowie deren 5%- und 95%-Quantilwert eingeführt.

3.4.3 Merkmale indirekter Nachbarn

Direkte Nachbarschaften zwischen beliebigen Regionen können einfach über deren gemeinsame Wasserscheide bestimmt werden (vgl. Kapitel 3.4.2). Um Nachbarschaften zwischen stabilen Regionen definieren zu können, benötigen wir ein anderes Kriterium als eine gemeinsame Wasserscheide.

Betrachten wir das aus einem Ausgangsbild 3.9(a) berechneten Bild der stabilen Regionen 3.9(b). Hieraus können wir ein Binärbild der stabilen Regionen ableiten und auf dessen Hintergrund das Exoskelett berechnen 3.9(c). Dieses kann als gemeinsamer Rand zweier stabiler Regionen interpretiert werden.

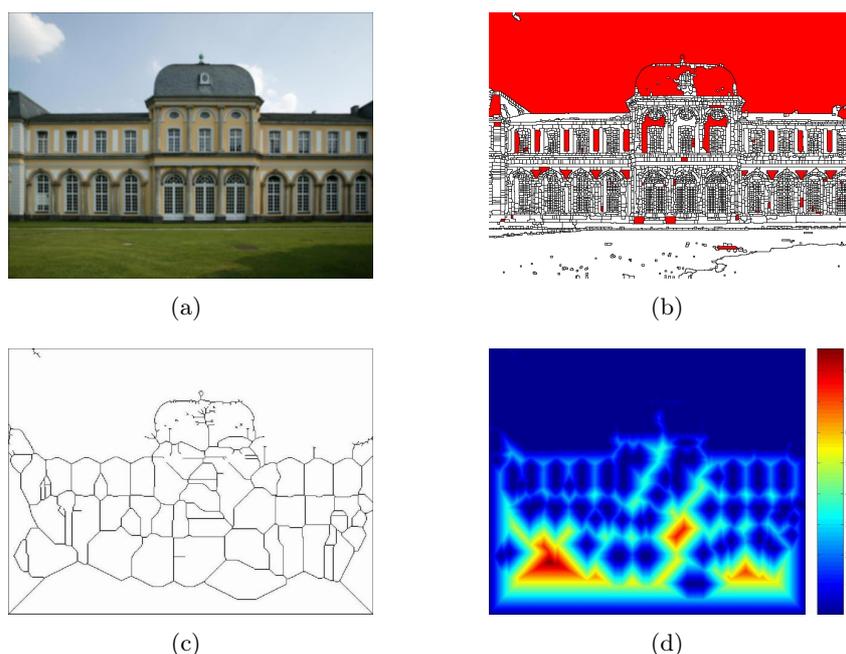


Abbildung 3.9: Dargestellt sind (a) Ausgangsbild, (b) stabile Regionen des Wasserscheidenalgorithmus, (c) Skelett des Hintergrundes auf dem Bild der stabilen Regionen, (d) Bild der Distanztransformation.

Es treten jedoch Fälle auf, in denen stabile Regionen aufgrund ihrer großen Distanz nicht als benachbart angesehen werden können, obwohl sie eine gemeinsame Skelettkante besitzen. Aus diesem Grund ist es erforderlich, die Distanz zwischen den stabilen Regionen zu bestimmen. Hierfür wird zusätzlich eine N4-Distanztransformation auf dem Hintergrund berechnet 3.9(d). Die Pixel des Hintergrundes erhalten als Wert den kürzesten Abstand zu den stabilen Regionen bzw. zum Bildrand. Die Pixel der stabilen Regionen erhalten den Wert 0.

Aus der Kombination von Skelett und Distanztransformation erhält man folgende Definition der Nachbarschaft zweier stabiler Regionen:

Zwei (stabile) Regionen heißen „indirekt benachbart“, wenn sie eine gemeinsame Skelettkante besitzen und wenn die minimale Distanz der gemeinsamen Skelettpixel zu den betrachteten Regionen kleiner einem Schwellwert d ist.

Die Variable d ist abhängig vom untersuchten Gegenstand zu wählen. Sie wurde für die weiteren Betrachtungen empirisch ermittelt und auf $d = 10$ [Pixel] festgelegt.

Der Ablauf zur Bestimmung der indirekt benachbarten Regionen ist in Abbildung 3.10 skizziert. Abbildung 3.10(a) stellt das Exoskelett der stabilen Regionen dar.

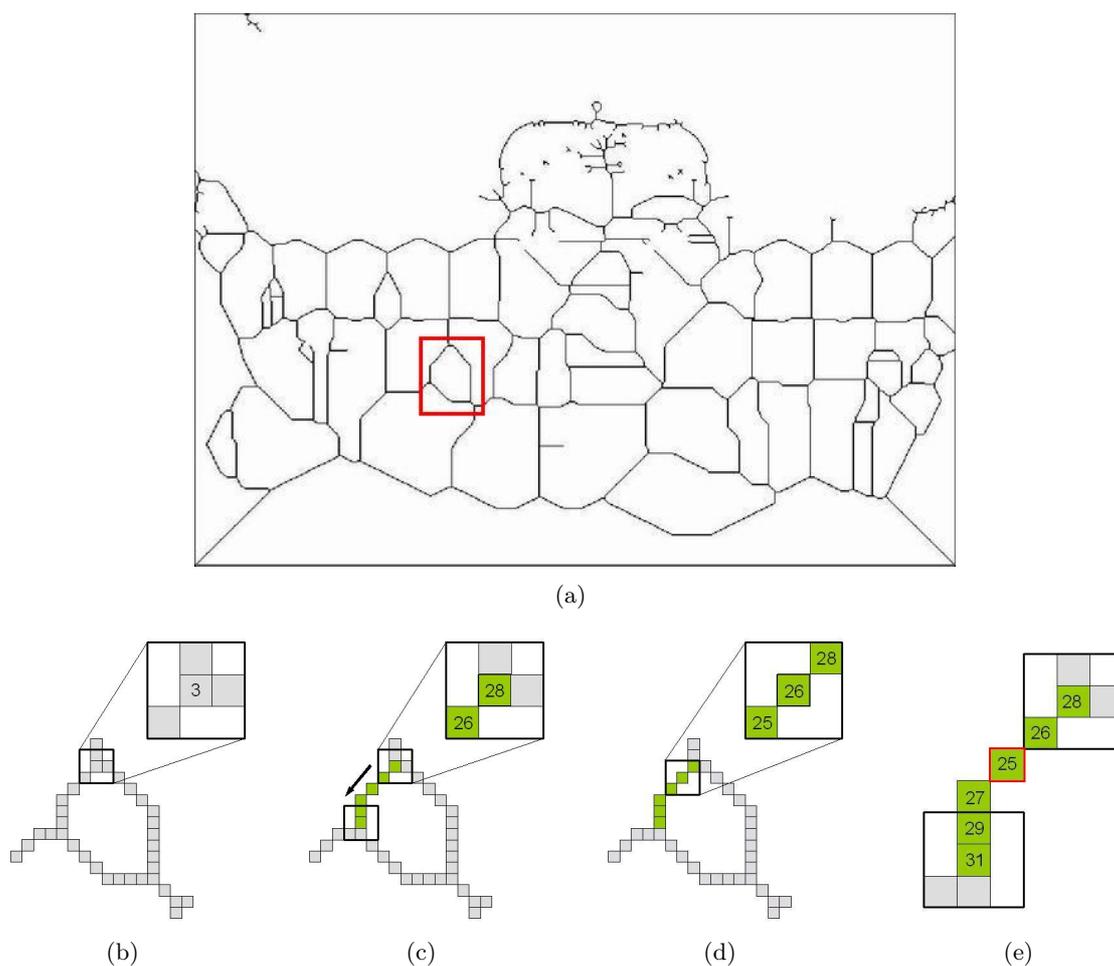


Abbildung 3.10: Bestimmung der Distanzen zwischen zwei stabilen Regionen. Ausgehend vom Bild des Skeletts (a) werden zunächst alle Kreuzungspunkte bestimmt. Dies sind Punkte mit mindestens 3 Nachbarn aus einer N8-Nachbarschaft (b). Von diesen Punkten aus wird das Skelett bis zum nächsten Kreuzungspunkt verfolgt (c), (d). Gespeichert wird jeweils die minimale Distanz (e).

Zunächst benötigt man nach der Definition eine gemeinsame Skelettkante zwischen den stabilen Regionen. Jede dieser Skelettkanten ist durch zwei Kreuzungspunkte begrenzt. In Analogie zu den Kreuzungspunkten der Randpixel aus Kapitel 3.4.2, erhalten wir die Kreuzungspunkte des Skeletts an den Stellen im Bild, an denen die Anwendung des Filters

$$f_K = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

auf das Skelett (Skelettpixel = 1, Hintergrund = 0) einen Wert größer oder gleich 3 ergibt (3.10(b)). Bereits in einem 3×3 -Fenster um den Kreuzungspunkt kann man anhand des Labelings ablesen, welche stabilen Regionen eine gemeinsame Skelettkante aufweisen. Tritt eine Skelettkante zwischen zwei Pixeln mit dem gleichen Label auf, so kann sie außer Acht gelassen werden. Somit werden Skelettkanten ausgeschlossen, die ins Leere verlaufen und nicht an einem

Kreuzungspunkt enden.

Man kennt zudem die Position der Skelettkante, die zwei verschiedenen Regionen zugeordnet ist, in dem 3×3 -Fenster. Durch wiederholtes Verschiebung des Fensters entlang der Skelettkante bis zum nächsten Kreuzungspunkt (3.10(c)-(d)) erhält man für jede Position einen Distanzwert. Gespeichert wird jeweils die minimale Distanz.

Untersucht werden soll, wieviele stabile Regionen nach der oben aufgeführten Definition benachbart sind. Eine größere Anzahl stabiler Regionen sowie ein dichteres Auftreten vermuten wir im Bereich von Gebäuden. Inwieweit sich diese Vermutung bestätigt, ist in Kapitel ?? beschrieben.

Weiterhin spielt das Merkmal des Flächeninhalts wie schon bei den direkten Nachbarn eine Rolle. Wir berechnen daher auch hier den Median der Flächeninhalte der indirekten Nachbarn sowie die 5%– und 95%–Quantile der Flächeninhalte.

Kapitel 4

Auswertung der Ergebnisse

Mit der Intention, geeignete Merkmale für die Gebäudeextraktion zu finden, haben wir im vorangehenden Kapitel 17 Kandidaten vorgestellt. Nach der Analyse von 60 terrestrischen Gebäudeaufnahmen in 21 Skalen liegen uns die Ausprägungen der Merkmale im Skalenraum vor.

Weitere Untersuchungen haben zum Ziel, aus den vorgestellten Merkmalen der stabilen Regionen sowie der direkten und indirekten Nachbarregionen eine für die Klassifikation geeignete Auswahl zu treffen. Dabei stellt sich zunächst einmal die Frage, ob wir Merkmalsextraktionen beliebiger Skalen verwenden können oder ob sich die Merkmalsausprägungen im Skalenraum ändern. Um dieser Frage nachzugehen, werden die Merkmale in Kapitel 4.1 auf Skalenabhängigkeit getestet. Anschließend betrachten wir in Kapitel 4.2 die Merkmale in Hinblick darauf, wie gut sie Daten aus einem Gebäude von sonstigen Daten unterscheiden können. Ein weiterer Aspekt, der über die Eignung eines Merkmals Aufschluss geben soll, ist das in Kapitel 4.3 behandelte Verhalten bezüglich Bildrauschen.

Für die Gebäudeextraktion geeignet erscheinende Merkmale werden in verschiedenen Formen kombiniert und fließen zusammen mit dem Trainings- und Testdatensatz, bestehend aus je 30 Bildern, in Kapitel 4.4 in die Klassifikation ein. Anhand der Ergebnisse des in Kapitel 2.2.4 beschriebenen tatsächlichen Fehlers können wir Aussagen über die Güte der Merkmalsextraktion treffen. Wir vergleichen die Ergebnisse nicht nur für unterschiedliche Merkmalskombinationen und über die Skalen des Skalenraumes, sondern auch für unterschiedliche Klassifikatoren.

Schließlich untersuchen wir ebenfalls in Kapitel 4.4 noch, wie gut sich spezielle Kategorien von Gebäudetypen im Bild detektieren lassen. Ein besonderer Augenmerk liegt dabei auf Häusern mit Klinker.

Die Ergebnisse werden an ausgewählten Beispielen vorgestellt. Für weitere Beispiele sei auf die im Anhang befindliche CD verwiesen.

4.1 Skalenabhängigkeit der Merkmale

Unter dem Begriff der Skalenabhängigkeit verstehen wir, dass sich Merkmalsausprägungen über den Skalenraum hinweg ändern. Diese Änderungen können im gesamten Skalenraum oder nur in Teilen auftreten. An den Stellen, an denen keine Skalenabhängigkeit vorliegt, müssen Merkmale nur für eine Skala erfasst werden. Die Berücksichtigung dieses Aspekts kann die Datenmenge und damit auch die Auswertzeit reduzieren.

Aus diesem Grund untersuchen wir die Merkmale im Folgenden auf ihre Abhängigkeit über

Teile des Skalenraumes. Hierfür bestimmen wir die mittleren Merkmalsausprägungen für jede Skala. Durch diese Daten werden stückweise ausgleichende Geraden approximiert. Aus einer von Null verschiedenen Geradensteigung können wir auf die Skalenabhängigkeit der Merkmale schließen. Eine horizontale ausgleichende Gerade sagt allerdings noch nichts darüber aus, ob eine Skalenunabhängigkeit vorliegt. In diesem Fall kann ein anderer Modellansatz erforderlich sein. Wir haben in einem solchen Fall die Zahl der Geradenstücke variiert.

Die Abbildungen 4.1(a)-(f) zeigen die skalenweise berechneten Mittelwerte der Merkmalsausprägungen sowie jeweils mehrere ausgleichende Geradenstücke. Ein Hypothesentest liefert für jedes Merkmal die Aussage, welche der zugehörigen Geradensteigungen sich signifikant von Null unterscheiden. Diese Geraden sind schwarz dargestellt. Geraden, deren Steigung nicht signifikant von Null abweicht, sind rot gekennzeichnet.

Skalenabhängig über den gesamten Skalenraum sind sowohl der Umfang und der Flächeninhalt der stabilen Regionen (Abbildung 4.1(a)+(b)) als auch der Flächeninhalt der direkten und indirekten Nachbarn der stabilen Regionen. Diese Erkenntnis ist nicht überraschend, da die Regionen mit zunehmender Skalierung immer stärker verschmelzen.

Das Verhältnis zwischen den Flächeninhalten stabiler Regionen und den Flächeninhalten ihrer direkten oder indirekten Nachbarregionen weist dagegen keine Skalenabhängigkeit auf (Abbildung 4.1(c)). Unterschiede im skalenabhängigen Wachstum zwischen den stabilen und nicht-stabilen Regionen scheinen daher nicht gravierend.

Ebenso skalenunabhängig sind die Merkmale, die im Zusammenhang mit der Betrachtung von Farbtönen aufgestellt wurden (Abbildung 4.1(d)), da sie sich auf eine Gegenüberstellung zu den stabilen Regionen beziehen und sich im Skalenraum verhältnismäßig ähnlich zu diesen ändern.

Lediglich drei Merkmale sind nur in bestimmten Bereichen skalenabhängig. Hierzu gehört der Formfaktor (Abbildung 4.1(e)) und die Homogenität (Abbildung 4.1(f)) der stabilen Regionen wie auch die Anzahl der indirekten Nachbarn. Während die Homogenität in den kleinen Skalen stark variiert und in hohen Skalen keine Skalenabhängigkeit aufweist, tritt bei den anderen beiden Merkmalen der umgekehrte Fall ein. Diese Aussage ist jedoch mit Vorsicht zu behandeln, da die mittleren Merkmalsausprägungen in den als skalenunabhängig erkannten Bereichen stark streuen. Das führt dazu, dass im Hypothesentest auch Steigungen steilerer Geraden als nicht signifikant von Null abweichend erkannt werden. Dennoch erscheint uns die partielle Skalenabhängigkeit plausibel, wie nachstehende Überlegungen zeigen.

Im Falle der Homogenität variieren die Grauwerte weniger stark, je mehr das Bild geglättet wird. Und ab einer gewissen Stufe ändert eine weitere Glättung wenig am Ergebnis. Dies spricht für die Skalenunabhängigkeit dieses Merkmals in hohen Skalen.

Formfaktor und Anzahl der indirekten Nachbarn orientieren sich dagegen an der Größe der stabilen Regionen, die bekanntlich in den hohen Skalen durch Verschmelzen stärker zunimmt. Je größer eine Region ist, desto variabler kann ihr Verhältnis zwischen Umfang und Flächeninhalt sein. Das führt auf eine ausgeprägtere Skalenabhängigkeit in den höheren Skalen. Außerdem kann eine Region mit großem Umfang mehr Nachbarn haben als eine Region mit kleinem Umfang. Dies muss jedoch nicht zwangsläufig der Fall sein. Es könnte auch nur eine große Nachbarregion auftreten. Dies spricht wiederum für die hohe Variabilität in den hohen Skalen.

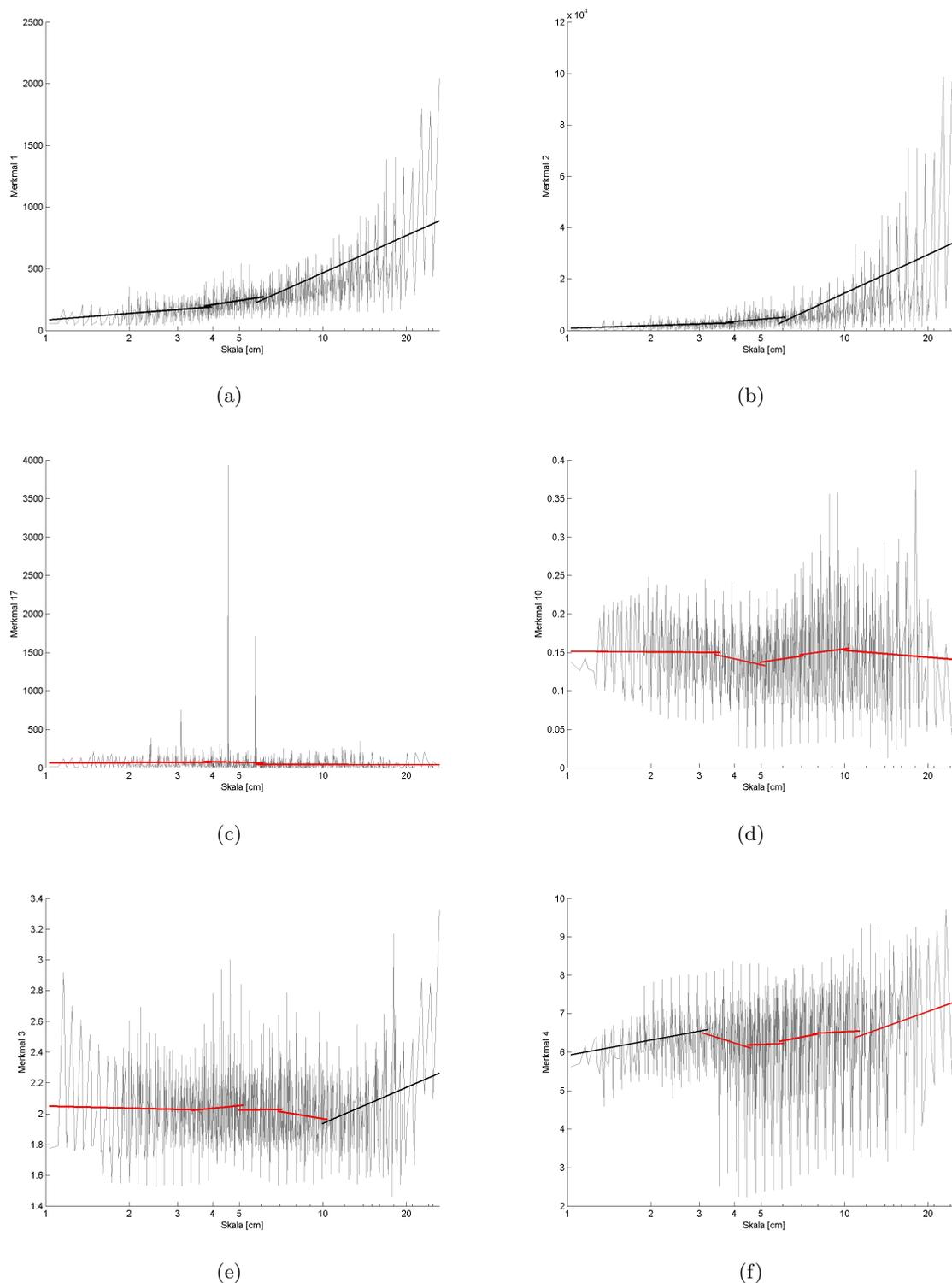


Abbildung 4.1: Die Abbildungen zeigen die Ergebnisse des Tests auf Skalenabhängigkeit für die Merkmale Umfang (a) und Flächeninhalt (b) der stabilen Regionen, Flächenverhältnis zwischen stabilen Regionen und deren indirekten Nachbarn (c), Median der Farbtondifferenzen zwischen stabilen Regionen und deren direkten Nachbarn (d), Formfaktor (e) und Homogenität (f) der stabilen Regionen. In Richtung der x-Achse sind die Skalen logarithmisch abgetragen, in Richtung der y-Achse sind die Merkmalsausprägungen jeweils für eine Skala gemittelt dargestellt. Desweiteren sind verschiedene ausgleichende Geradenstücke eingezeichnet. Schwarz dargestellte Geradenstücke haben eine signifikant von Null verschiedene Steigung und weisen daher auf eine Skalenabhängigkeit hin. Die Steigung roter Geradenstücke weicht dagegen nicht signifikant von Null ab.

Wir schließen daraus, dass eine Betrachtung des Skalenraumes für die Merkmale der stabilen Regionen neue Erkenntnisse liefert. Für Merkmale der direkten und indirekten Nachbarregionen gilt dies jedoch nur bedingt. Speziell diejenigen Merkmale, die sich aus einem Vergleich zwischen den Eigenschaften der Nachbarregionen und der stabilen Regionen ergeben, liefert in verschiedenen Skalen ähnliche Ergebnisse. Der Grund ist, dass sich die Merkmale der Nachbarn über den Skalenraum im gleichen Maße ändern wie die Merkmale der stabilen Regionen.

4.2 Ergebnisse der Analyse einzelner Merkmale

Die Analyse einzelner Merkmale soll uns zu Merkmalen führen, die Daten aus Gebäuden gut von anderen Daten unterscheiden können. Hierzu betrachten wir eindimensionale Histogramme (Abbildung 4.2) der Merkmalsausprägungen getrennt nach Regionen, die innerhalb eines Gebäudes liegen und solchen die außerhalb liegen. Das Histogramm der Gebäudedaten ist hellgrau dargestellt, das Histogramm der sonstigen Daten schwarz und der Überlappungsbereich dunkelgrau. An der Stärke der Überlappung der beiden Histogramme können wir entnehmen, welche Merkmale geeignet sind, die Daten zu separieren.

Umfang Es zeichnet sich ab, dass stabile Regionen kleinen Umfangs hauptsächlich im Gebäude zu finden sind. Es handelt sich dabei unter anderem um Regionen im Bereich von Fensterrahmen. Entgegen erster Vermutungen treten in den analysierten Daten Umfänge bis 20 cm sogar ausschließlich in Gebäuden auf. Die Abwesenheit solcher Regionen außerhalb von Gebäuden muss mit der fehlenden Stabilität der Regionen im Bereich von Vegetation begründet werden. Falls im Bereich von Vegetation stabile Regionen auftreten, dann meist in größeren Schattenregionen.

Stabile Regionen mit Umfängen bis etwa 150 cm liegen eher in einem Gebäude. Dies gilt für die Aufnahmesituation in den 60 vorliegenden Bildern. Eine größere Stichprobe könnte zu andern Ergebnissen führen.

Große Umfänge entstehen durch Verschmelzen der stabilen Regionen. Sie erweisen sich als nicht aussagekräftig, da sie gleichermaßen in Gebäuden als auch außerhalb vorkommen.

Flächeninhalt Die Zahl der stabilen Regionen, die in einem Gebäude liegen, nimmt mit wachsendem Flächeninhalt ab. Dagegen findet man für jeden Flächeninhalt etwa gleich viele stabile Regionen außerhalb von Gebäuden.

Flächen mit kleinem Inhalt können daher besser zugeordnet werden. Ab etwa 1000 cm² erscheint der Flächeninhalt als Merkmal weniger geeignet, da sich die beiden Histogramme um weniger als $\frac{2}{3}$ unterscheiden.

Formfaktor Wie in Kapitel 3.4.1.1 bereits vermutet, liegen stabile Regionen mit einem Formfaktor von 1.3–1.7 vermehrt in Gebäuden. Dieser Formfaktor umfasst Rechtecke mit einem Seitenverhältnis von 1 : 1 bis 1 : 2, wie sie bei Fenstern und Türen auftreten.

Als problematisch erweist sich hier wieder die geringe Größe der Stichprobe für stabile Regionen außerhalb von Gebäuden. Das zugehörige Histogramm zeigt ansonsten eine ähnliche Verteilung wie das Histogramm der stabilen Regionen, die innerhalb eines Gebäudes liegen. Das erschwert die Trennung der beiden Klassen. Für die vorliegenden Daten würde

die Entscheidung bis zu einem Formfaktor von 2.0 jedoch zugunsten der Lage im Gebäude ausfallen.

Homogenität Die Homogenität stabiler Regionen weist für beide Histogramme vergleichbare Verteilungen auf, die sich etwa um einen Faktor von $\frac{2}{3}$ voneinander unterscheiden. Wir können daraus schließen, dass es in einem Gebäude zu viele Regionen gibt, die genauso inhomogen sind wie ihre Umgebung. Dieses Merkmal ist für die Klassifikation daher nicht geeignet.

Flächeninhalt direkter Nachbarn Analog zum Flächeninhalt der stabilen Regionen nimmt die Zahl der stabilen Regionen, die zu einem Gebäude gehören, mit dem Flächeninhalt der Nachbarn ab. Jedoch ist die Zahl der Nachbarregionen mit kleinem Flächeninhalt gestiegen. Das erschwert das Differenzieren zwischen Gebäude und Hintergrund.

Der Flächeninhalt indirekter Nachbarn muss nicht gesondert betrachtet werden, da es sich hierbei wieder um den Flächeninhalt stabiler Regionen handelt. Der Median des Flächeninhalts der indirekten Nachbarn sowie die 5%- und 95%-Quantile eignen sich in ähnlicher Weise bedingt für die Klassifikation.

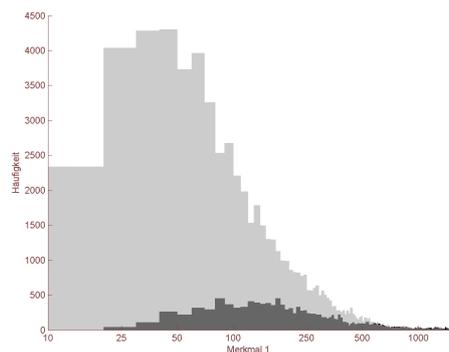
Flächenverhältnis Für beide Klassen liegt bis auf einen Faktor von $\frac{2}{3}$ die gleiche Verteilung vor. Eigentlich wäre das Flächenverhältnis damit ungeeignet. Bei der Verteilung hebt sich jedoch der Wert 1 für stabile Regionen innerhalb eines Gebäudes hervor. Wir werden auf dieses Merkmal bei der Klassifikation als Ergänzung des Merkmalsvektors zurückkommen.

Farbtöne Geringe Farbtendifferenzen treten etwas häufiger im Bereich von Gebäuden auf als außerhalb. Ansonsten sind die Farbtendifferenzen nicht sehr aussagekräftig. Besonders große Probleme beim Separieren der Daten ergeben sich jedoch für den prozentualen Anteil der Nachbarregionen, die einen ähnlichen Farbton wie die zugehörige stabile Region haben.

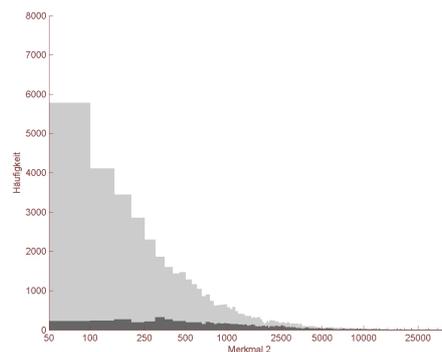
Anzahl indirekter Nachbarn Die Anzahl der indirekten Nachbarn erscheint uns ebenfalls als nicht aussagekräftig.

Keines der Merkmale ist für sich alleine geeignet, eine Klassifikation zwischen Gebäude und Hintergrund vorzunehmen. Nach [Guyon, Elisseeff 2003] können Merkmale, die alleine nutzlos sind, in Kombination mit anderen jedoch sehr nützlich sein. Wir benötigen daher eine geeignete Kombination von Merkmalen.

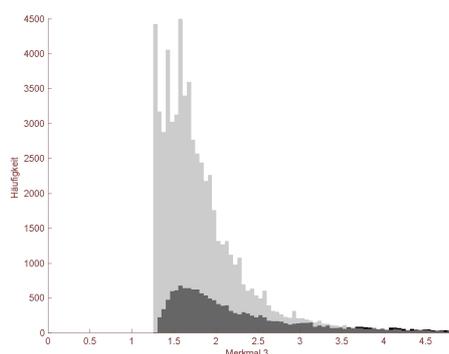
Da der Umfang als einziges Merkmal einen, wenn auch kleinen überlappungsfreien Bereich der Histogramme für Gebäudedaten und sonstige Daten aufweist, stellt er ein wichtiges Merkmal für die Klassifikation dar. Alle anderen Merkmale sind sehr kritisch zu betrachten, da sich die Histogramme über den gesamten Bereich der Merkmalsausprägungen überlappen. Nach der vorangestellten Analyse der Histogramme erachten wir jedoch auch die jenigen Merkmale für bedingt geeignet, die um den Flächeninhalt der stabilen Regionen entstanden sind. Dazu gehört auch der Formfaktor. Gegebenenfalls kann auch der Flächeninhalt der direkten Nachbarn sowie deren Flächenverhältnis mit zur Klassifikation herangezogen werden. Dies geschieht in Kapitel 4.4. Die dort vorgestellten Klassifikationsergebnisse geben mehr Aufschluss über die Güte der Merkmale.



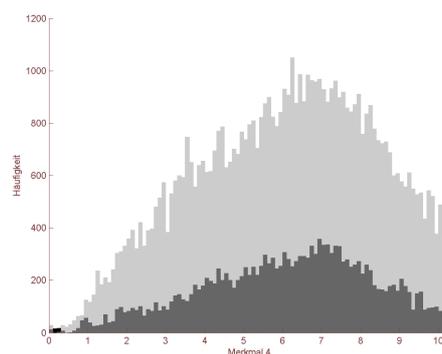
(a) Umfang



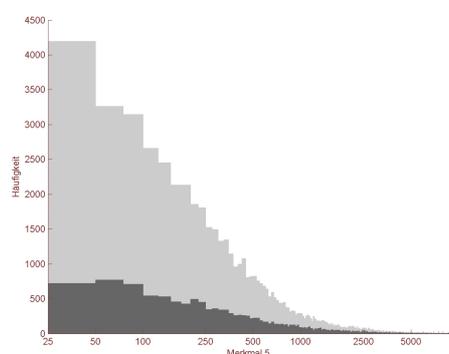
(b) Flächeninhalt



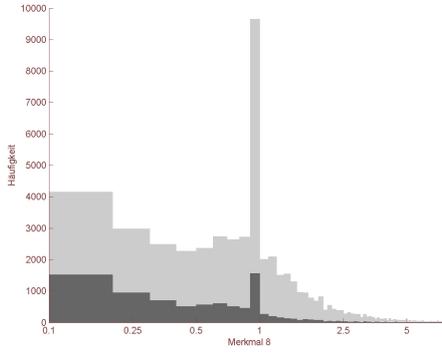
(c) Formfaktor



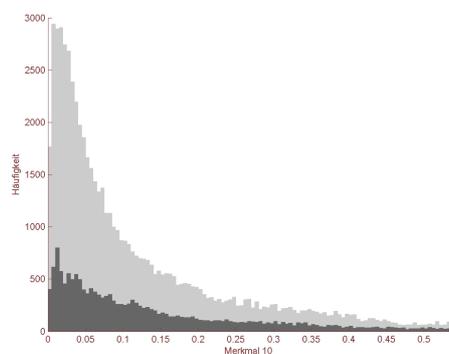
(d) Homogenität



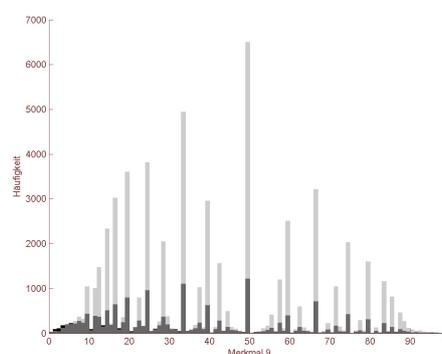
(e) Flächeninhalt der direkten Nachbarn



(f) Flächenverhältnis zwischen stabilen Regionen und deren direkten Nachbarn



(g) Farbtondifferenzen zwischen stabilen Regionen und deren direkten Nachbarn



(h) Prozentanteile der direkten Nachbarn mit ähnlichem Farbton

Abbildung 4.2: Eindimensionale Histogramme ausgewählter Merkmale.

4.3 Der Einfluss von Rauschen

Weiterhin verfolgen wir das Ziel, geeignete Merkmale für die Gebäudeextraktion auszuwählen. Ein Kriterium ist ihre Unempfindlichkeit gegenüber Bildrauschen. Inwieweit sich Bildrauschen auf die Merkmalsausprägungen auswirkt, wird nun untersucht. Wir analysieren den Einfluss von Gauß'schem Rauschen (Abbildung 4.3(b)) und „Salt and Pepper“ Rauschen (Abbildung 4.3(c)) in unterschiedlichen Stärken.



(a)



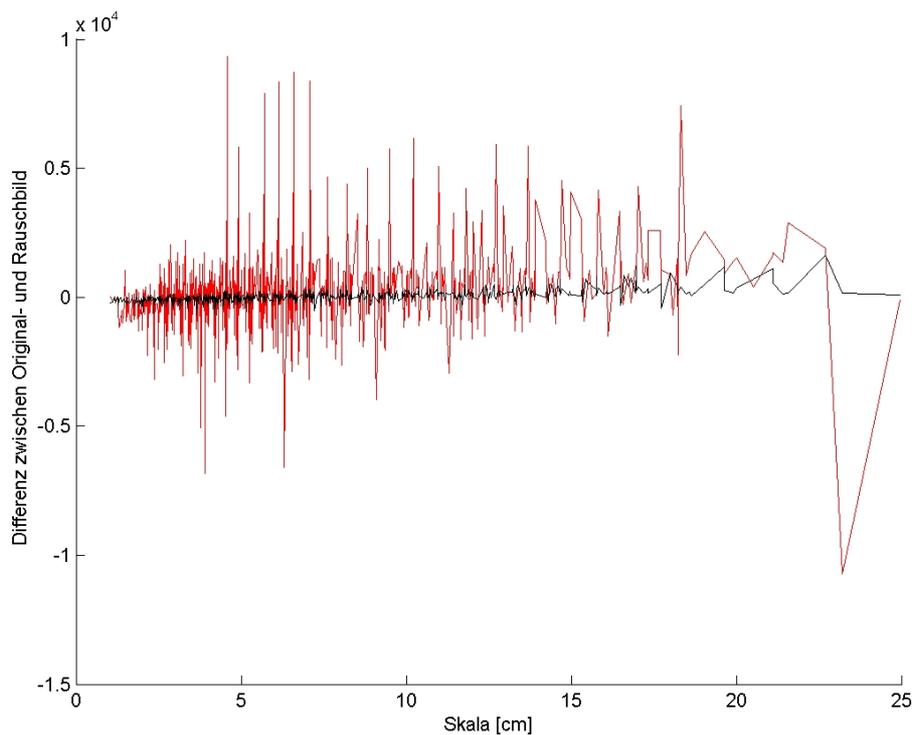
(b)



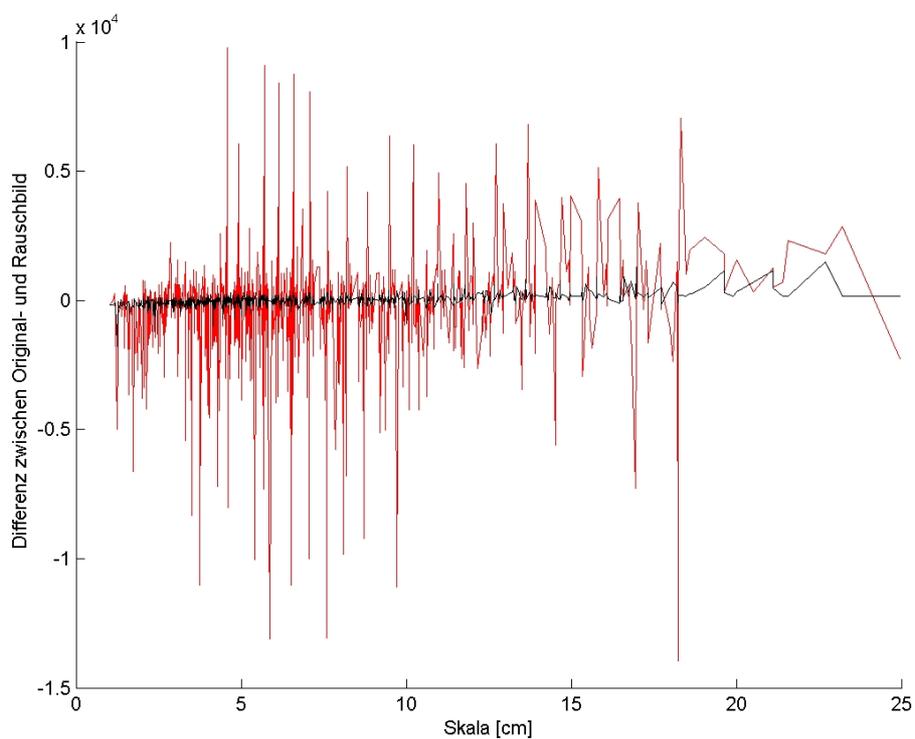
(c)

Abbildung 4.3: Das Originalbild (a) wurde einmal mit Gauß'schem Rauschen (b) mit einer Standardabweichung von 0.05 versehen und einmal mit „Salt and Pepper“ Rauschen (c) mit einer Rauschdichte von 0.05.

Da wir von normalverteilten Daten ausgehen, ist es naheliegend zunächst den Einfluss von Gauß'schem Rauschen zu untersuchen. Die Stärke des Rauschens variieren wir über die Wahl der Standardabweichung von 0.01, 0.02 und 0.05. Bereits bei kleinen Standardabweichungen treten für alle Merkmale große Differenzen zwischen den mittleren Merkmalsausprägungen von Original- und Rauschbild auf. Abbildung 4.4(a) verdeutlicht dies exemplarisch für den Umfang der stabilen Regionen. Für Regionen innerhalb eines Gebäudes treten die Differenzen besonders stark hervor. Diese Diskrepanz zwischen stabilen Regionen innerhalb und außerhalb eines Gebäudes tritt für alle Merkmale auf, die sich ausschließlich auf die stabilen Regionen beziehen. Bei Merkmalen, die eine Betrachtung der Nachbarregionen erfordern, weisen die Merkmalsausprägungen von Original- und Rauschbild unabhängig der Lage der stabilen Regionen bezüglich eines Gebäudes große Differenzen auf.



(a)



(b)

Abbildung 4.4: Differenz zwischen den skalenweise gemittelten Umfängen der stabilen Regionen aus Original- und Rauschbild, dargestellt für Gauß'sches Rauschen (a) mit einer Standardabweichung von 0.01 und „Salt and Pepper“ Rauschen (b) mit einer Rauschdichte von 0.01.

Weiter wird der Einfluss von „Salt and Pepper“ Rauschen analysiert. Diese Art von Rauschen kann durch einen Defekt in den Sensorelementen des Aufnahmesystems verursacht werden. Manche Pixel erscheinen dadurch weiß, andere erscheinen schwarz.

Durch alleiniges Betrachten des Rauschbildes (Abbildung 4.3(c)) wird deutlich, dass die stabilen Regionen durch das „Salt and Pepper“ Rauschen stark beeinflusst werden. Das bestätigen auch die Differenzen der mittleren Merkmalsausprägungen aus Original- und Rauschbild (Abbildung 4.4(b)). Sie treten nochmals stärker als beim Gauß'schen Rauschen hervor. Für beide Arten von Bildrauschen betrachten wir jetzt noch die Korrelation zwischen Original und Rauschbild. Da wir, wie wir in Kapitel 4.4 zeigen werden, nicht von normalverteilten Merkmalen ausgehen können, verwenden wir statt des bei Normalverteilung üblichen Pearson-Korrelationskoeffizienten den Rangkorrelationskoeffizienten nach Spearman, nachzulesen bei [Rinne 2003]. Tabelle 4.1 zeigt die Rangkorrelationskoeffizienten für alle Merkmale exemplarisch für Gauß'sches Rauschen mit einer Standardabweichung von 0.01 und für „Salt and Pepper“ Rauschen mit einer Rauschdichte von 0.01.

	Merkmal	Gauß		Salt and Pepper	
		in	out	in	out
stabile Region	Umfang	0.0	0.1	0.0	0.2
	Flächeninhalt	0.0	0.2	0.1	0.2
	Formfaktor	0.1	0.1	0.1	0.0
	Homogenität	0.1	0.1	0.0	0.1
direkte Nachbarn	Median des Flächeninhalts	0.1	0.1	0.1	0.0
	5%-Quantil des Flächeninhalts	0.1	0.1	0.1	0.1
	95%-Quantil des Flächeninhalts	0.1	0.1	0.1	0.2
	Flächenverhältnis	0.2	0.1	0.1	0.0
	Anteil ähnlicher Farbtöne	0.1	0.0	0.1	0.1
	Median der Farbtendifferenzen	0.1	0.0	0.1	-0.0
	5%-Quantil der Farbtendifferenzen	-0.0	0.0	-0.0	0.0
	95%-Quantil der Farbtendifferenzen	0.2	0.1	0.1	0.1
indirekte Nachbarn	Anzahl	0.1	0.1	0.1	0.1
	Median des Flächeninhalts	0.0	0.2	0.1	0.1
	5%-Quantil des Flächeninhalts	-0.1	0.1	0.1	0.1
	5%-Quantil des Flächeninhalts	0.0	0.2	0.1	0.1
	Flächenverhältnis	0.0	0.1	0.0	0.0

Tabelle 4.1: Dargestellt sind die Rangkorrelationskoeffizienten nach Spearman getrennt nach der Art des Rauschens und der Lage der stabilen Regionen innerhalb („in“) oder außerhalb („out“) von Gebäuden. Liegt der Betrag des Korrelationskoeffizienten unter 0.3, dann müssen die Ergebnisse aus Original- und Rauschbild als unkorreliert angesehen werden.

Keiner der Werte überschreitet einen Betrag von 0.3. Wir können daher bei keinem Merkmal eine Korrelation zwischen dem Original- und Rauschbild feststellen.

Aus unseren Untersuchungen können wir folgendes Fazit ziehen. Sowohl Gauß'sches Rauschen als auch „Salt and Pepper“ Rauschen haben einen starken Einfluss auf Lage und Form der Wasserscheidenregionen. Dadurch werden alle Merkmale beeinflusst, da sie aus den Wasserscheidenregionen hervorgehen. Besonders Merkmale, die sich nicht nur auf eine einzelne (stabile) Region beziehen, weisen stark variierende Ausprägungen auf. Dazu gehören sowohl die Merkmale der direkten als auch der indirekten Nachbarn. Bildrauschen ist daher möglichst zu vermeiden.

Bezüglich des Rauschverhaltens gut geeignete Merkmale können wir mit diesem Wissen nicht ausfindig machen. Wir können lediglich solche Merkmale bevorzugen, die sich direkt auf die stabilen Regionen beziehen.

4.4 Klassifikation

Bevor wir eine Klassifikation mit den in Kapitel 4.2 ausgewählten Merkmalen durchführen können, ist vorab ein geeigneter Klassifikator zu bestimmen.

4.4.1 Auswahl des Klassifikators

Die Auswahl eines für die Gebäudeextraktion geeigneten Klassifikationsverfahrens beschränken wir auf die in Kapitel 2.2 vorgestellten Verfahren der Diskriminanzanalyse nach Fisher, der Bayes-Klassifikation sowie der linearen und nicht-linearen Supportvektormaschinen. Deren Vor- und Nachteile wurden dort bereits angesprochen. Sie sollen jetzt nochmals aufgegriffen werden, um die Wahl des Klassifikator einzuschränken.

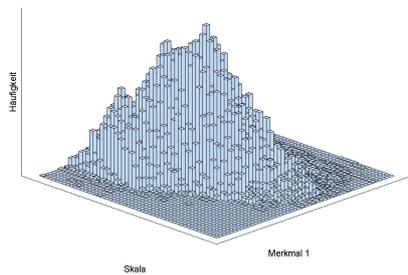
Die lineare Diskriminanzanalyse nach Fisher ist aufgrund ihrer Linearität gut generalisierbar. Sie weist allerdings mehr Fehler auf, falls die Daten nicht linear separierbar sind.

Für die Bayes-Klassifikation werden Vorinformationen in Form der Kenntniss über die zugrundeliegende Verteilung der Merkmale benötigt. Da keines der Merkmale negative Werte annehmen kann, können wir zunächst nicht von einer Normalverteilung ausgehen. Die Aussage darüber, welche Verteilung im einzelnen Fall vorliegt, kann durch die Betrachtung von Histogrammen erfolgen. Die in Abbildung 4.5 dargestellten Histogramme stellen die Häufigkeit des Auftretens von stabilen Regionen in Gebäuden für alle Ausprägungen eines Merkmals und über alle Skalen dar. Da die Histogramme für stabile Regionen außerhalb von Gebäuden größtenteils ähnlich aussehen, haben wir auf deren Darstellung verzichtet. Bei der Bestimmung einer Verteilung aus einem Histogramm ist jedoch zu beachten, dass eine gewisse Anzahl an Daten erforderlich sind. Eine kleine Datenmenge, wie sie hier vorliegt, ist nicht so aussagekräftig.

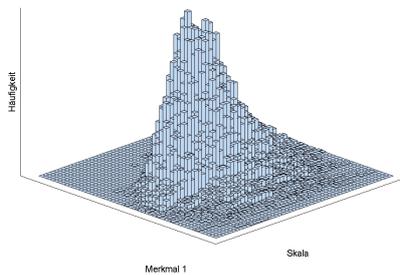
Für einige Merkmale, wie den Umfang und Flächeninhalt, können wir aus dem Histogramm (Abbildung 4.5(a)-(d)) auf eine genäherte log-Normalverteilung schließen. Für den Formfaktor wäre eine β -Verteilung geeignet (Abbildung 4.5(e)-(f)). Andere Merkmale weisen eine Mischverteilung auf.

Dieser Umstand erschwert die Bayes-Klassifikation. Alternativ könnte das gesamte Histogramm gespeichert werden, was jedoch einen hohen Speicheraufwand mit sich bringt.

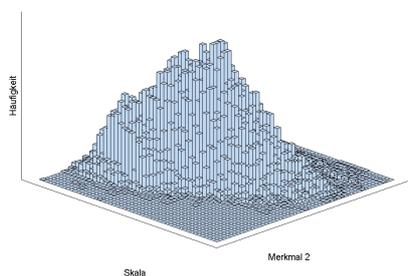
Supportvektormaschinen benötigen keine Vorinformation über die Verteilung. Sie erzeugen in hohen Dimensionen eine trennende Hyperebene und sind damit auch gut generalisierbar. Ihr einziger Nachteil ist die etwas längere Laufzeit.



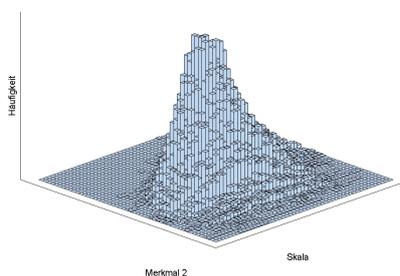
(a) Umfang



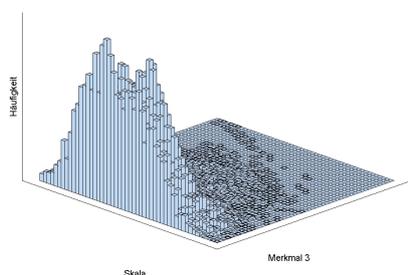
(b) Umfang



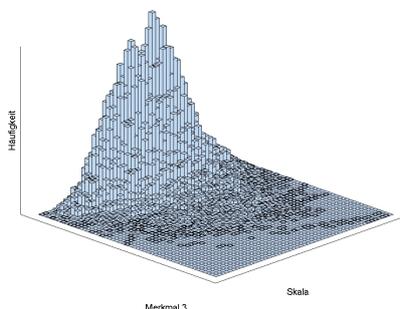
(c) Flächeninhalt



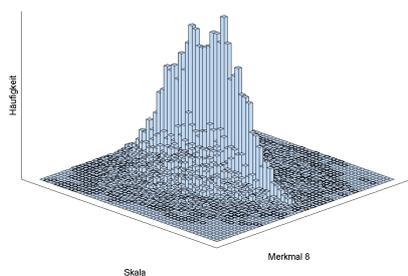
(d) Flächeninhalt



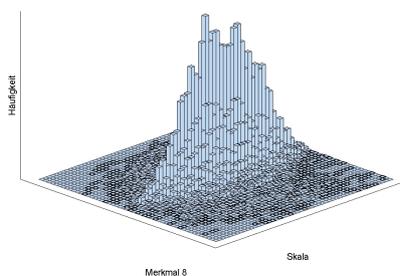
(e) Formfaktor



(f) Formfaktor



(g) Flächenverhältnis



(h) Flächenverhältnis

Abbildung 4.5: Histogramm

Des Weiteren sollen erste Klassifikationsergebnisse zur Entscheidung für ein Verfahren beitragen. Wir werden hier der Frage nachgehen, wie gut durch die einzelnen Klassifikationsverfahren die Daten separiert werden können. Zur besseren Anschauung beschränken wir uns beim Vergleich der Klassifikationsergebnisse zunächst auf die Kombination jeweils zweier Merkmale. Exemplarisch sind die Ergebnisse für die Kombination der Merkmale Flächeninhalt und Formfaktor in der Tabelle 4.2 und in den Abbildungen 4.6-4.9 festgehalten.

Klassifikationsverfahren	Skala σ	tatsächlicher Fehler			# richtige Zuordn. pro 100 Regionen		
		in (ϵ_1)	out (ϵ_2)	all (ϵ)	in	out	all
Fisher	1.0	0.3089	0.1995	0.2152	67-71	76-84	77-80
	4.0	0.5552	0.2300	0.3642	40-49	73-81	60-67
Bayes	1.0	0.7236	0.0314	0.1310	26-30	95-99	85-88
	4.0	0.5409	0.2600	0.3759	41-50	70-78	59-66
SVM (p=1)	1.0	0.7276	0.0164	0.1187	25-29	97-100	87-89
	4.0	0.5658	0.2175	0.3612	39-48	74-82	61-67
SVM (p=2)	1.0	0.6911	0.0232	0.1193	29-33	96-99	87-89
	4.0	0.5231	0.2500	0.3627	43-52	71-79	60-67

Tabelle 4.2: Fehlerverhalten bei verschiedenen Klassifikatoren unter der Anwendung von nicht-linearen Supportvektormaschinen (p=2) auf die Skalen $\sigma = 1$ und $\sigma = 4$. Die Werte resultieren aus dem Beispiel der Kombination der Merkmale Flächeninhalt der stabilen Regionen, Median des Flächeninhalts der direkten Nachbarregionen und Median des Flächeninhalts der indirekten Nachbarregionen der stabilen Regionen. Dargestellt sind der tatsächliche Fehler in der Zuordnung der stabilen Regionen sowie die daraus ermittelte Anzahl der Regionen, die in 99% der Fälle richtig zugeordnet wird. Wir unterscheiden zusätzlich, ob es sich um stabile Regionen innerhalb („in“) oder außerhalb („out“) eines Gebäudes handelt.

Die Tabelle 4.2 stellt den nach den Formeln (2.33) und (2.34) berechneten tatsächliche Fehler dar, der beim Zuordnen der Regionen zu den Klassen „Gebäude“ und „Hintergrund“ entsteht. Dieser ist getrennt nach dem allgemeinen Fehler (ϵ) und nach den klassenspezifischen Fehlern (ϵ_1 , ϵ_2) aufgetragen. Weiter gibt die Tabelle Auskunft über die Zahl der Regionen, die in 99% der Fälle von 100 Regionen richtig zugeordnet werden.

Die Supportvektormaschinen erzielen hier die besten Ergebnisse in der Klassifikation beliebiger stabiler Regionen (in der Tabelle mit „all“ gekennzeichnet), gefolgt von der Bayes-Klassifikation. Die Abbildungen 4.6-4.9 veranschaulichen diesen Sachverhalt graphisch. Sie zeigen die Merkmalsausprägungen von Flächeninhalt und Formfaktor für die stabilen Regionen, die im Gebäude liegen und für die stabilen Regionen des Hintergrundes. Für jeder der Klassifikationsmethoden ist die Trennfunktion zwischen den Klassen eingezeichnet.

Dass die lineare Diskriminanzanalyse nach Fisher für stabile Regionen in Gebäuden zumindest für kleine Skalen bessere Ergebnisse liefert, gewichten wir gegenüber dem Vorteil der Supportvektormaschinen, für alle Regionen bessere Ergebnisse zu liefern, geringer.

Aus diesem Grund und da sie keine Vorinformation benötigen, verwenden wir die Supportvektormaschinen für die Klassifikation. Unsere Wahl fällt zudem auf die nicht-lineare Form der Supportvektormaschinen, die etwas weniger Rechenzeit als die linearen Supportvektormaschinen benötigt.



Abbildung 4.6: Lineare Diskriminanzanalyse nach Fisher für die Merkmale Flächeninhalt (x-Achse) und Formfaktor (y-Achse) in den Skalen $\sigma = 1$ (a) und $\sigma = 4$ (b). Dargestellt sind die Merkmalsausprägungen von stabilen Regionen in Gebäuden als blaue Punkte und Merkmalsausprägungen von stabilen Regionen im Hintergrund als rote Kreise. Die durchgezogene Linie kennzeichnet die Grenze zwischen den Klassen.



Abbildung 4.7: Bayes-Klassifikation für die Merkmale Flächeninhalt (x-Achse) und Formfaktor (y-Achse) in den Skalen $\sigma = 1$ (a) und $\sigma = 4$ (b). Dargestellt sind die Merkmalsausprägungen von stabilen Regionen in Gebäuden als blaue Punkte und Merkmalsausprägungen von stabilen Regionen im Hintergrund als rote Kreise. Die durchgezogene Linie kennzeichnet die Grenze zwischen den Klassen. Die gestrichelte Linie stellt den Fehlersaum dar.



Abbildung 4.8: Lineare Supportvektormaschinen für die Merkmale Flächeninhalt (x-Achse) und Formfaktor (y-Achse) in den Skalen $\sigma = 1$ (a) und $\sigma = 4$ 4.8(b). Dargestellt sind die Merkmalsausprägungen von stabilen Regionen in Gebäuden als blaue Punkte und Merkmalsausprägungen von stabilen Regionen im Hintergrund als rote Kreise. Die durchgezogene Linie kennzeichnet die Grenze zwischen den Klassen.



Abbildung 4.9: Nicht-lineare Supportvektormaschinen für die Merkmale Flächeninhalt (x-Achse) und Formfaktor (y-Achse) in den Skalen $\sigma = 1$ (a) und $\sigma = 4$ 4.9(b). Dargestellt sind die Merkmalsausprägungen von stabilen Regionen in Gebäuden als blaue Punkte und Merkmalsausprägungen von stabilen Regionen im Hintergrund als rote Kreise. Die durchgezogene Linie kennzeichnet die Grenze zwischen den Klassen.

4.4.2 Klassifikationsergebnisse

Nach der Entscheidung für die nicht-linearen Supportvektormaschinen, betrachten wir nun deren Ergebnisse. Zur besseren Anschauung beginnen wir auch hier zunächst mit der Betrachtung des zweidimensionalen Falls. Wir untersuchen das Verhalten der Klassifikationsergebnisse über den Skalenraum, stellen die Kombinationen unterschiedlicher Merkmale gegenüber und erforschen deren spezielle Anwendbarkeit auf bestimmte Kategorien von Haustypen.

Exemplarisch für die Kombination von Flächeninhalt und Formfaktor der stabilen Regionen wird nun das Fehlerverhalten der Klassifikation über den Skalenraum hinweg dargestellt. Tabelle 4.3 stellt neben der Anzahl der pro Skala in allen Bildern erfassten stabilen Regionen auch deren nach den Formeln (2.33) und (2.34) berechneten tatsächliche Fehler dar. Des Weiteren ist die Zahl der gesamten richtigen Zuordnung in 99% der Fälle veranschaulicht.

Skala (σ)	# stabile Regionen			tatsächlicher Fehler			# richtige Zuordnungen		
	in	out	all	in (ϵ_1)	out (ϵ_2)	all (ϵ)	in	out	all
1.0000	3289	618	3907	0.6911	0.0232	0.1193	948-1085	593-612	3388-3492
1.0757	3751	700	4451	0.7113	0.0292	0.1280	1012-1155	667-690	3823-3938
1.1571	4127	730	4857	0.7309	0.0245	0.1240	1038-1184	701-722	4195-4313
1.2447	4391	799	5190	0.7370	0.0266	0.1283	1080-1230	765-789	4461-4586
1.3389	4580	845	5425	0.7464	0.0273	0.1324	1086-1238	809-833	4641-4770
1.4402	4746	880	5626	0.7487	0.0273	0.1363	1116-1270	843-867	4792-4925
1.5493	4845	935	5780	0.7734	0.0259	0.1441	1023-1173	897-922	4878-5015
1.6665	4944	988	5932	0.7671	0.0247	0.1438	1076-1229	950-975	5009-5148
1.7927	5008	1028	6036	0.7620	0.0328	0.1510	1115-1270	979-1008	5052-5196
1.9284	4894	1067	5961	0.7582	0.0368	0.1603	1107-1261	1011-1043	4932-5078
2.0743	3906	1041	4947	0.7934	0.0396	0.1939	742-873	983-1015	3916-4059
2.2313	3139	1021	4160	0.7578	0.0522	0.2211	699-822	949-985	3171-3308
2.4002	2584	983	3567	0.7589	0.0557	0.2388	568-680	909-946	2649-2780
2.5819	2153	947	3100	0.7089	0.0731	0.2615	573-681	856-898	2226-2352
2.7773	1792	926	2718	0.7049	0.0807	0.2791	480-579	829-872	1899-2019
2.9875	1503	874	2377	0.6640	0.1225	0.3060	458-552	741-791	1591-1707
3.2136	1272	841	2113	0.6466	0.1491	0.3290	406-494	688-742	1362-1473
3.4569	1093	799	1892	0.5921	0.2098	0.3529	404-488	601-660	1170-1277
3.7185	943	738	1681	0.5523	0.2363	0.3603	383-462	533-593	1024-1126
4.0000	805	673	1478	0.5231	0.2500	0.3627	347-420	475-533	894-989

Tabelle 4.3: Fehlerverhalten über den Skalenraum unter der Anwendung von nicht-linearen Supportvektormaschinen ($p=2$). Die Werte resultieren aus dem Beispiel der Kombination der Merkmale Flächeninhalt und Formfaktor der stabilen Regionen. Dargestellt ist die Anzahl der stabilen Regionen für jede Skala, sowie die Anzahl der Regionen, die in 99% der Fälle richtig zugeordnet wird. Diese berechnet sich aus dem tatsächlichen Fehler nach den Formeln 2.33 und 2.34. Wir unterscheiden zusätzlich, ob es sich um stabile Regionen innerhalb („in“) oder außerhalb („out“) eines Gebäudes handelt.

Die Tabelle 4.3 bestätigt unsere Vermutung, dass die Zahl der stabilen Regionen, bedingt durch Verschmelzung, in den höheren Skalen abnimmt. Weiter wird deutlich, dass sich die Güte der Zuordnung der stabilen Regionen zu den Klassen „Gebäude“ und „Hintergrund“ mit zunehmender Skalierung insgesamt verschlechtert. Auffällig ist das Verhältnis der Fehler ϵ_1 und ϵ_2 der einzelnen Klassen. Stabile Regionen, die zu einem Gebäude gehören, werden viel öfter falsch zugeordnet als solche, die zum Hintergrund gehören. In Bildern hoher Auflösung kann

die Fehlzuordnung von Gebäuderegionen bis zu 80% betragen. Das bedeutet, dass eine stabile Region aus einem Gebäude liegt zu 80% als Hintergrundregion angesehen wird. Dies ist ein sehr pessimistischer Ansatz. Jedoch kann man in Kombination mit dem kleinen Fehler von 2% für Hintergrundregionen, aus einem Klassifikationsergebnis, das die Klasse „Gebäude“ detektiert, recht zuverlässig auf das Vorhandensein eines Gebäudes schließen.

In den hohen Skalen verbessert sich das Fehlerverhältnis zugunsten derjenigen Regionen, die im Gebäude liegen, gegenüber der Hintergrundregionen.

Dies ist noch kein befriedigendes Ergebnis für die Gebäudeextraktion. Es zeigt jedoch, dass die Betrachtung des Skalenraumes sinnvoll ist, da sich die verschiedenen Klassen in verschiedenen Skalen unterschiedlich gut detektieren lassen.

Betrachten wir im Folgenden verschiedene Merkmalskombinationen. Tabelle 4.4 stellt eine Auswahl der Ergebnisse dar. Wir zeigen neben dem besten Ergebnis, das wir für eine feste Wahl von Trainings- und Testdaten erzielen konnten, auch einige negative Beispiele.

Das beste Ergebnis (Zeile 1) ergibt sich für eine Kombination aus Merkmalen, die sich ausschließlich auf die stabilen Regionen selber beziehen. Hier wurden die Merkmale Umfang, Flächeninhalt und Formfaktor herangezogen. In 99% der Fälle werden bis zu 90 von 100 stabilen Regionen richtig zugeordnet. Durch Hinzunahme weiterer Merkmale (Zeilen 2-4) verschlechtert sich dieses Ergebnis jedoch. Für die in Kapitel 4.2 aufgestellten Kandidaten für die Merkmalsextraktion (Zeile 4) werden nur 83-86 von 100 Regionen richtig erkannt. Geringfügig verbessert sich das Ergebnis, wenn wir hieraus nur die Merkmale des Flächeninhalts der stabilen Regionen und ihrer Nachbarn auswählen (Zeile 5). Andere Merkmale wie die Farbwerte (Zeile 6) oder sogar ein Kombination aus allen Möglichkeiten (Zeile 7) liefern zu schlechte Ergebnisse, insbesondere auch im Hinblick auf die Detektierbarkeit von Regionen in Gebäuden.

Merkmalskombination	tatsächlicher Fehler			# richtige Zuordn. pro 100 Regionen		
	in (ϵ_1)	out (ϵ_2)	all (ϵ)	in	out	all
1,2,3	0.6789	0.0239	0.1181	30-34	96-99	87-90
1,2,3,7	0.6707	0.0287	0.1211	31-35	95-99	87-89
1,2,3,10	0.6951	0.0239	0.1205	28-33	96-99	87-89
1,2,3,5,8,14	0.6992	0.0602	0.1521	28-32	91-96	83-86
2,5,14	0.7114	0.0396	0.1363	27-31	94-98	85-88
4,9,10	0.8984	0.0389	0.1626	9-12	94-98	82-85
alle	0.9146	0.0519	0.1761	7-10	92-97	81-84

Tabelle 4.4: Fehlerverhalten bei verschiedenen Merkmalskombinationen unter der Anwendung von nicht-linearen Supportvektormaschinen ($p=2$). Dargestellt sind der tatsächliche Fehler in der Zuordnung der stabilen Regionen sowie die daraus ermittelte Anzahl der Regionen, die in 99% der Fälle richtig zugeordnet wird. Wir unterscheiden zusätzlich, ob es sich um stabile Regionen innerhalb („in“) oder außerhalb („out“) eines Gebäudes handelt. Die Abkürzungen für die Merkmale sind folgendermaßen kodiert: 1 = Umfang, 2 = Flächeninhalt, 3 = Formfaktor, 4 = Homogenität, 5 = Median des Flächeninhalts der direkten Nachbarn, 6,7 = das zugehörige 5%- und 95%-Quantil, 8 = Flächenverhältnis der direkten Nachbarn, 9 = Prozentanteil ähnlicher Farbtöne, 10 = Median der Farbtendifferenzen, 11,12 = das zugehörige 5%- und 95%-Quantil, 13 = Anzahl indirekter Nachbarn, 14 = Median des Flächeninhalts der indirekten Nachbarn, 15,16 = das zugehörige 5%- und 95%-Quantil, 17 = Flächenverhältnis der indirekten Nachbarn.

Um eine Gebäudeextraktion durchführen zu können, die mit möglichst verschiedenen Gebäudetypen umgehen kann, empfiehlt es sich, Merkmale zu verwenden, die relativ einfach und allgemeingültig sind. Das führt uns weg von den Merkmalen der Nachbarregionen, hin zu den Merkmalen der stabilen Regionen. Wie wir bereits gesehen haben, liefern diese bei vielfältigen Trainings- und Testdaten die besten Ergebnisse.

Es stellt sich die Frage, ob wir bei allen Gebäudetypen die gleichen Ergebnisse erzielen können. Abschließend wollen wir daher die Resultate aus der Klassifikation einzelner Gebäudetypen diskutieren. Wir untersuchen anhand zweier Merkmalskombinationen, ob es Gebäudetypen gibt, die allgemein besser detektiert werden können als andere und ob es Merkmale gibt, mit denen man spezielle Gebäudetypen besonders gut extrahieren kann.

Wir verwenden nun nicht mehr die gesamten 30 Testdaten, sondern wählen daraus bestimmte Kategorien von Haustypen aus. Gegenüber gestellt werden Häuser mit Klinker und Putz, Häuser mit weißem und farbigem Anstrich sowie Häuser mit typischer und untypischer Form. Am Trainingsdatensatz ändern wir dagegen nichts.

Die Tabellen 4.5 und 4.6 stellen wiederum den tatsächlichen Fehler der Klassifikation sowie die Zahl der richtigen Zuordnungen von 100 Regionen in 99% der Fälle exemplarisch für die Skala mit $\sigma = 1$ dar.

Haustyp	tatsächlicher Fehler			# richtige Zuordn. pro 100 Regionen		
	in (ϵ_1)	out (ϵ_2)	all (ϵ)	in	out	all
beliebig	0.6789	0.0239	0.1181	30-34	96-99	87-90
Klinker	0.6765	0.0237	0.0833	30-34	96-99	91-93
Putz	0.76747	0.0263	0.1324	21-25	96-99	85-88
helle Fassade	0.7833	0.0209	0.1244	20-24	96-99	86-89
farbige Fassade	0.6915	0.0426	0.1388	29-33	94-98	85-88
typische Hausformen	0.6389	0.0213	0.1469	34-38	96-99	84-87
untypische Hausformen	0.6383	0.0348	0.1021	34-38	94-98	89-91

Tabelle 4.5: Fehlerverhalten bei verschiedenen Gebäudetypen unter der Anwendung von nicht-linearen Supportvektormaschinen ($p=2$). Die Werte resultieren aus dem Beispiel der Kombination der Merkmale Umfang, Flächeninhalt und Formfaktor der stabilen Regionen. Dargestellt sind der tatsächliche Fehler in der Zuordnung der stabilen Regionen sowie die daraus ermittelte Anzahl der Regionen, die in 99% der Fälle richtig zugeordnet wird. Wir unterscheiden zusätzlich, ob es sich um stabile Regionen innerhalb („in“) oder außerhalb („out“) eines Gebäudes handelt.

Die erste Tabelle 4.5 bezieht sich auf die Merkmalskombination von Umfang, Fläche und Formfaktor, die sich bereit für beliebige Gebäude bewährt hat (siehe Tabelle 4.4). Es zeigt sich, dass insbesondere Gebäude mit Klinker mit Hilfe dieses Merkmalsvektors gut erkannt werden. Der Formfaktor spielt hier eine entscheidende Rolle. Für die Skala mit $\sigma = 1$, sind die Bilder noch in sehr kleine Wasserscheiden unterteilt. Die Klinker werden zum Teil mit erfasst, so dass der Formfaktor ihre rechteckige Form erfassen kann.

Des weiteren zeichnet sich ab, dass helle Fassaden leichter zu detektieren sind als farbige und untypische Hausformen leichter als typische Hausformen. Als untypische Hausformen haben wir auch Gebäude mit aufgenommen, bei denen das Dach nicht zu erkennen ist. Wir vermuten einen daher Zusammenhang in der Einfachheit der Struktur, ausgenommen dem Spezialfall „Klinker“, und der besseren Detektierbarkeit mittels Umfang, Flächeninhalt und Formfaktor.

Haustyp	tatsächlicher Fehler			# richtige Zuordn. pro 100 Regionen		
	in (ϵ_1)	out (ϵ_2)	all (ϵ)	in	out	all
beliebig	0.7114	0.0396	0.1363	27-31	94-98	85-88
Klinker	0.7000	0.0958	0.1487	28-32	87-93	84-87
Putz	0.6471	0.0296	0.1183	33-37	95-99	87-89
helle Fassade	0.6167	0.0366	0.1154	36-41	94-98	87-90
farbige Fassade	0.6915	0.0686	0.1611	29-33	90-96	82-85
typische Hausformen	0.6944	0.0284	0.1638	28-33	95-99	82-85
untypische Hausformen	0.6383	0.0481	0.1140	34-38	93-97	87-90

Tabelle 4.6: Fehlerverhalten bei verschiedenen Gebäudetypen unter der Anwendung von nicht-linearen Supportvektormaschinen ($p=2$). Die Werte resultieren aus dem Beispiel der Kombination der Merkmale Flächeninhalt der stabilen Regionen, Median des Flächeninhalts der direkten Nachbarregionen und Median des Flächeninhalts der indirekten Nachbarregionen der stabilen Regionen. Dargestellt sind der tatsächliche Fehler in der Zuordnung der stabilen Regionen sowie die daraus ermittelte Anzahl der Regionen, die in 99% der Fälle richtig zugeordnet wird. Wir unterscheiden zusätzlich, ob es sich um stabile Regionen innerhalb („in“) oder außerhalb („out“) eines Gebäudes handelt.

Die zweite Tabelle stellt die Ergebnisse aus der Kombination des Flächeninhalts der stabilen Regionen, des Medians vom Flächeninhalt der direkten Nachbarregionen und des Median vom Flächeninhalt der indirekten Nachbarregionen dar. Wieder sind es die untypischen Hausformen und hellen Fassaden, die am besten erkannt werden. Mit der Begründung von oben, dass einfache Flächen besser erkannt werden, können wir nun auch Fassaden mit Putz leichter erkennen als Fassaden mit Klinker. Für die Detektion von Klinker fehlt in dieser Merkmalskombination der Formfaktor.

Wir folgern daraus, dass wir Vorinformation zum gesuchten Gebäudetyp gezielt in die Wahl der Klassifikationsmerkmale einfließen lassen können.

Kapitel 5

Zusammenfassung und Ausblick

Die automatische Detektion von Gebäuden in terrestrischen Farbbildern ist in der Wissenschaft ein ungelöstes Problem. In dieser Diplomarbeit haben wir den langen Weg zur Lösung dieses Problems ein Stück weit beschritten. Wir haben Merkmale aufgestellt, die wir als geeignet erachteten, um Gebäude zu detektieren. Die Merkmalsfindung orientierte sich an Wasserscheidenregionen, die über verschiedene Auflösungsstufen der Bilder hin stabil sind. Unterschieden wurden Merkmale dieser stabilen Regionen sowie Merkmale ihrer direkten und indirekten Nachbarregionen.

Um die Eignung der Merkmale zu prüfen, wurden diese zunächst einzeln untersucht. Eine Analyse auf Skalenabhängigkeit hat dabei ergeben, dass die Betrachtung des Skalenraumes sinnvoll ist, da einige der untersuchten Merkmale ganz oder teilweise skalenabhängig sind. In Histogrammen konnten wir mit keinem der aufgeführten Merkmale alleine die Daten in „Gebäude“ und „Hintergrund“ separieren. Dies führte letztendlich auf die Kombination von Merkmalen und der Durchführung einer Klassifikation auf der Grundlage von 60 Gebäudeaufnahmen.

Die Suche nach einem geeigneten Klassifikator führte uns zu der linearen Diskriminanzanalyse nach Fisher, der Bayes-Klassifikation und den Supportvektormaschinen. Letztere überzeugten durch einen geringen allgemeinen Fehler in der Zuweisung der stabilen Regionen zu den Klassen „Gebäude“ und „Hintergrund“. Dennoch detektieren sie Regionen, die in der Realität im Gebäude liegen, wesentlich schlechter als die Diskriminanzanalyse nach Fisher. Wenn jedoch eine Region als Gebäude detektiert wird, ist dies mit hoher Wahrscheinlichkeit richtig. Und Regionen außerhalb der Gebäude lassen sich mit den Supportvektormaschinen bis auf einen Fehler von 2% detektieren.

Das wirft die interessante Frage auf, ob sich durch Kombination der Verfahren, bessere Ergebnisse in der Gebäudedetektion erzielen lassen. Weitere Untersuchungen könnten an dieser Stelle ansetzen.

Wir konnten anhand der Klassifikationsergebnisse verschiedene Aspekte zeigen.

Zunächst einmal hängt das Klassifikationsergebnis vom Skalenraum ab. Hintergrundregionen lassen sich in niedrigen Skalen besser detektieren als in hohen Skalen. Gebäude lassen sich allgemein sehr schlecht detektieren, besser jedoch in den hohen Skalen.

Für die Gebäudeextraktion ohne Einschränkung des Gebäudetyps sind die Eigenschaften der stabilen Regionen besser geeignet als die Eigenschaften ihrer Nachbarn. Wir konnten zeigen, dass die Hinzunahme von Merkmalen der Nachbarregionen die Gebäudeextraktion negativ

beeinflusst. Die besten Ergebnisse erzielen dagegen Merkmale, die von den stabilen Regionen selber stammen. Dazu gehören ihr Umfang, Flächeninhalt und Formfaktor. Dennoch weist die Klassifikation mit diesen Merkmalen immer noch einen Fehler von etwa 12% in den kleinen Skalen auf. In den hohen Skalen steigt die Zahl der Fehlklassifikationen weiter an.

Die Gebäudeextraktion eines bestimmten Haustyps kann durch die Hinzunahme bestimmter Merkmale verbessert werden. Beispielsweise lassen sich Häuser mit Klinker besser durch Hinzunahme des Formfaktors extrahieren.

Wir wollen diese Arbeit mit einigen Ideen zur weiteren Forschung abschließen.

Zur Verbesserung der Klassifikationsergebnisse kann einmal die gezielte Hinzunahme weiterer Merkmale führen. Wie unsere Ergebnisse zeigen, sollten sich diese Merkmale im allgemeinen Fall möglichst direkt auf die stabilen Regionen beziehen. Bei der Betrachtung spezieller Haustypen können auch bestimmte Merkmale der Nachbarregionen nützlich sein.

Hier könnte auch eine Variation in der Definition der indirekten Nachbarn zu anderen Ergebnissen führen. Zur Definition der indirekten Nachbarschaften haben wir deren aus einer Distanztransformation gewonnenen Abstände herangezogen. Weitere Untersuchungen könnten an diese Stelle zeigen, ob man durch Variation des Schwellwertes zu geeigneteren Merkmalen der indirekten Nachbarn gelangt.

Die weitere Forschung könnte auch auf die Kombination von Klassifikationsergebnissen aus mehreren Skalen zielen, die einzeln sehr unterschiedliche Ergebnisse hervorbringen. Das gilt auch für die Kombination von Klassifikationsmethoden, die oben bereits vorgeschlagen wurde.

Auf die bereits vorhandenen Klassifikationsergebnisse könnte eine Post-Klassifikation angesetzt werden, die Nachbarschaften von stabilen Regionen näher untersucht. Falls alle Nachbarn die gleiche Klasse aufweisen wie die stabile Region, dann ist es wahrscheinlicher, dass die Zuordnung richtig war. Der umgekehrte Fall ergibt sich analog.

Falls die Ergebnisse durch weitere Nachforschungen verbessert werden können, wäre ein nächster Schritt, die Klassen weiter zu differenzieren.

Auch die Anwendung der Klassifikation auf beliebige Bilder wäre zu untersuchen. Welche Ergebnisse erhalten wir, wenn die Bilder aus ländlichen Regionen stammen oder gar keine Gebäude im Bild dargestellt sind?

Als letzter Punkt sei noch ein weit verbreitetes Problem der Bildanalyse aufgegriffen. Schatten bereiten immer wieder Probleme bei der Extraktion von Objekten aus Bildern. Auch bei einigen der betrachteten stabilen Regionen im Bereich von Vegetation handelt es sich um Schattenregionen. Um diese vorab zu detektieren, könnten die Untersuchungen zu den Farbtönen aus dem HSV-Farbraum weiterführen.

Anhang A

Bedienung des in Matlab implementierten Programms

Nachstehend werden die in Matlab implementierten Programmteile vorgestellt. Diese Darstellung umfasst eine kurze Beschreibung, inklusiver der Ein- und Ausgabeparameter, der übergeordneten Programme und deren Unterfunktionen.

A.1 Vorverarbeitungsschritt

Labeling Nummerierung der Wasserscheidenregionen

Eingabe:

- Bild der stabilen Regionen
- Bild aller Wasserscheidenregionen

Ausgabe:

- Labelbild der stabilen Regionen inklusive der Anzahl der stabilen Regionen
- Labelbild aller Wasserscheidenregionen inklusive der Anzahl der aller Wasserscheidenregionen

SkelettSpeichern Bestimmt wird das Skelett zwischen den stabilen Regionen.

Eingabe:

- Bild der stabilen Regionen

Ausgabe:

- Skelett
- Labelbild der Regionen, die durch das Skelettieren entstehen

Rauschen Verrauschen der Originalbilder durch Gauß'sches Rauschen oder „Salt and Pepper“

Rauschen

Eingabe:

- Originalbild

- Standardabweichung für Gauß'sches Rauschen
- Rauschdichte für „Salt and Pepper“ Rauschen

Ausgabe:

- Rauschbilder

A.2 Merkmalsextraktion

Hauptprogramm Mit dem „Hauptprogramm“ werden für vorgegebene Bilddaten die in Kapitel 3 beschriebenen Merkmale extrahiert.

Eingabe:

- Anzahl der Bilder und Skalen
- bei Bedarf: auszuschließende Bilder und Skalen

für alle Bilder

- Originalbild
- Polygone aus der Annotation

für alle Skalen

- Bild der stabilen Regionen
- Labelbild der stabilen Regionen
- Bild aller Wasserscheidenregionen
- Labelbild aller Wasserscheidenregionen
- Skelett
- Labelbild des Skeletts
- Bild mit Wasserscheiden

Ausgabe:

- Matrix (`index_in`) der Anzahl der stabilen Regionen, die in einem Gebäude liegen
(Dimension: $\#$ Skalen \times $\#$ Bilder)
- matrix (`index_out`) der Anzahl der stabilen Regionen, die nicht in einem Gebäude liegen
(Dimension: $\#$ Skalen \times $\#$ Bilder)
- 2 Merkmalsvektoren für jedes Merkmal (je einer für die stabilen Regionen innerhalb und außerhalb von Gebäuden)

Für jedes Bild und jede Skala werden die Merkmalsausprägungen nacheinander an diesen Vektor gehängt. Der Abgriff von Merkmalsausprägungen einer bestimmten Skala kann über die Anzahl der stabilen Regionen aus den Matrizen `index_in` und `index_out` erfolgen.

Funktionen, die im „Hauptprogramm“ aufgerufen werden:

StableIn Hier wird anhand der Polygone aus der Annotation geprüft, ob eine stabile Region einem Gebäude zugeordnet ist.

Umfang Die Zahl der Umfangspixel einer Region wird anhand einer Look-up-Tabelle berechnet.

sizeNeighbours Zunächst wird eine Adjazenzmatrix aufgestellt, die für direkt benachbarte Regionen einer stabilen Region den Eintrag 1 enthält. Weiter wird der Flächeninhalt für jede dieser Nachbarregionen berechnet und daraus der Median bestimmt.

colNeighbours Die Funktion transformiert ein RGB-Bild in den HSV-Farbraum. Die daraus gewonnenen Farbtonwerte werden für jede Wasserscheidenregion gemittelt. Für alle zu einer stabilen Region benachbarten Regionen wird daraus einmal der Median berechnet und zum anderen die Differenz zum Farbtonwert der stabilen region bestimmt.

Distanzh Berechnet wird die Distanztransformation auf dem Binärbild der stabilen Regionen.

anzNeighbours Um die Anzahl der indirekten Nachbarn einer stabilen Region bestimmen zu können, wird hier ebenfalls eine Adjazenzmatrix aufgestellt. Diese enthält dort eine 1 als Eintrag, wo eine indirekte Nachbarschaft vorliegt und sonst 0.

A.3 Ausgabe der Ergebnisse

Ergebnisse Das Programm „Ergebnisse“ umfasst die Auswertung der extrahierten Merkmale, wie in Kapitel 4 vorgestellt. Dies beinhaltet die Untersuchung auf Skalenabhängigkeit, ein- und zweidimensionale Histogramme der Merkmale, die Untersuchung des Rauschverhaltens sowie die Klassifikation der Merkmale.

Eingabe:

- Anzahl der Bilder und Skalen
- Auswahl eines Merkmals und Angabe, ob eine logarithmische Darstellung gewünscht ist
- Auswahl der Ausgabe (Analyse der Skalenabhängig, ein- und zweidimensionale Histogramme, Analyse des Rauschverhaltens, Klassifikation)

Fallabhängige Eingabe:

- Feinheit des Histogramms
- Art des Rauschbildes (Gauß oder „Salt and Pepper“)
- Auswahl des Klassifikators

- Auswahl mehrerer Merkmale für die Klassifikation
- Auswahl der Trainings- und Testdaten für die Klassifikation
- lineare oder quadratische Supportvektormaschinen

Ausgabe für die Analyse der Skalenabhängigkeit:

- skalenweise gemittelte Merkmalsausprägungen
- ausgleichende Geradenstücke durch die gemittelten Merkmalsausprägungen
- rote Darstellung der Geradenstücke, deren Steigung nicht signifikant von 0 abweicht
- schwarze Darstellung der Geradenstücke, deren Steigung signifikant von 0 abweicht (Skalenabhängigkeit)

Ausgabe für die Histogramme:

- eindimensionales Histogramm über eine Merkmalsausprägung
- zweidimensionales Histogramm über eine Merkmalsausprägung und alle Skalen

Ausgabe für das Rauschverhalten:

- Differenz zwischen den skalenweise gemittelten Merkmalsausprägungen aus Original- und Rauschbild
- Korrelationskoeffizient

Ausgabe für die Klassifikation:

- Graphikausgabe der Merkmalsvektoren und der Trennfunktion (bei zweidimensionalen Merkmalsvektoren)
- tatsächlicher Fehler

Funktionen, die im Programm „Ergebnisse“ aufgerufen werden:

Skalenkoord Die Funktion „Skalenkoord“ sortiert die Koordinaten der Skalenachse aufsteigend. Die Sortierreihenfolge wird ebenfalls ausgegeben, um die Merkmalsausprägungen entsprechend umsortieren zu können.

Skalenabhaengigkeit Aus den nach den Skalenkoordinaten Merkmalsausprägungen wird skalenweise der Mittelwert bestimmt. Anschließend werden die Endpunkte und die Steigung der ausgleichenden Geradenstücke durch Aufruf der Funktion „ausgleichendeGerade“ bestimmt. Die Skalenabhängigkeit wird aus der Geradensteigung bestimmt, wenn diese signifikant von 0 abweicht.

histogramm Mit Hilfe der Funktion „histogramm“ können zwei eindimensionale Histogramme überlappend gezeichnet werden.

Rauschverhalten Aus den nach den Skalenkoordinaten Merkmalsausprägungen wird jeweils für Original- und Rauschbild skalenweise der Mittelwert bestimmt. Ausgegeben wird die Differenz der Mittelwerte zwischen Original- und Rauschbild sowie Fehlergrenzen für diese Differenz.

Diskriminante, Bayes und SVM Die drei Funktionen dienen der Klassifikation der Daten aufgrund beobachteter Merkmalsausprägungen. Sie liefern ein Modell für die Trennfunktion sowie den tatsächlichen Fehler.

Literaturverzeichnis

- [Brügelmann, Förstner 1992] Brügelmann, R. and W. Förstner (1992): *Noise Estimation for Color Edge Extraction*. In: Förstner, Ruwiedel (Eds.): *Robust Computer Vision*, pp. 90-106, Wichmann, Karlsruhe.
- [Christianini, Shawe-Taylor 2000] Cristianini, N. and J. Shawe-Taylor (2000): *An Introduction to Support Vector Machines*. Cambridge University Press, Cambridge.
- [Drauschke et al. 2006(1)] Drauschke, M.; Schuster, H.-F. und W. Förstner (2006): *Detectability of Buildings in Aerial Images over Scale Space*. Symposium of ISPRS Commission III: Photogrammetric Computer Vision, Bonn, Germany, pages 7-12, 2006. Abrufbar unter: <http://www.ipb.uni-bonn.de/papers> (Datum: 07.01.2007)
- [Drauschke et al. 2006(2)] Drauschke, M.; Schuster, H.-F. und W. Förstner (2006): *Stabilität von Regionen im Skalenraum*. Publikationen der DGPF, Band 15: Geoinformation und Erdbeobachtung, pp. 29-36, Berlin, 2006. Abrufbar unter: <http://www.ipb.uni-bonn.de/papers> (Datum: 07.01.2007)
- [Duda et al. 2001] Duda, R. O.; Hart, P. E. and D. G. Stork (2001): *Pattern Classification*. John Wiley, New York.
- [Foley et al. 1995] Foley, J. D. et al.(1995): *Computer Graphics: Principles and Practice*. Addison-Wesley, Amsterdam.
- [Franc, Hlaváč 2004] Franc, V. and V. Hlaváč (2004): *Statistical Pattern Recognition Toolbox for Matlab*.
- [Guyon, Elisseeff 2003] Guyon, I. and A. Elisseeff (2003): *An Introduction to Variable and Feature Selection*. In: *Journal of Machine Learning Research* 3, pp 1157-1182, 2003.
- [Koch 1997] Koch, K.-R. (1997): *Parameterschätzung und Hypothesentests*. Dümmler Verlag, Bonn.
- [Koenderink 1984] Koenderink, J. (1984): *The Structure of Images*. In: *Biological Cybernetics* 50, pp 363-370.
- [Lindeberg 1990] Lindeberg, T. (1990): *Scale-Space for Discrete Signals*. In: *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 12(3), 1990, pp 234-254.
- [Olsen, Nielsen 1997] Olsen, O.F. and M. Nielsen (1997): *Multi-Scale Gradient Magnitude Watershed Segmentation*. In: *ICIAP'97 - 9th International Conference on Image Analysis and Processing*, Lecture Notes in Computer Science Vol. 1310, pp.9-13.

- [Platt 1998] Platt, J.C. (1998): *Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines*. Technical Report MSR-TR-98-14, Microsoft Research, Redmond.
- [Rinne 2003] Rinne, H. (2003): *Taschenbuch der Statistik*. Harri Deutsch Verlag, Frankfurt am Main.
- [Vapnik 1995] Vapnik, V. (1995): *The Nature of Statistical Learning Thoery*. Springer Verlag.
- [Vincent, Soille 1991] Vincent, L. and P. Soille (1991): *Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations*. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.13, No. 6, June 1991, pp 583-598.