

Tobias Hanning

Vektorielle Mehrniveaupassung

Anwendungen in der Bildsegmentierung

Eingereicht zur Erlangung des Grades eines Doktors der Naturwissenschaften an
der Fakultät für Mathematik und Informatik
der Universität Passau im September 2001



Lehrstuhl für Numerische Mathematik
und Analysis
Prof. Dr. Klaus Donner

Für Richard, Anneliese, Andrea, Ilona und Julia

Zusammenfassung

Die Anwendungen der vektoriellen Mehrniveaupassung in der Bildsegmentierung stehen in unmittelbarer Nachbarschaft zum verbreiteten Ansatz, Bilder mit Methoden der Variationsrechnung über einen Energieterm zu segmentieren. Beiden Verfahren ist gemeinsam, daß sie versuchen, das gegebene Bild durch stückweise stetige Funktionen zu approximieren. Die maximalen Teilmengen des Definitionsbereichs, auf denen die approximierende Funktion stetig ist, bilden dann die Segmente. Im Gegensatz zu dem in der Literatur häufig unter dem Schlagwort Mumford-Shah-Modell bekannten Energieminimierungsverfahren ist der Raum der Funktionen, mit denen das Bild approximiert wird, bei der vektoriellen Mehrniveaupassung ein endlicher Vektorraum. Ein weiterer Unterschied zu diesem Ansatz ist das System der erlaubten Mengen. Es werden nur Segmentierungen erlaubt, deren Segmente aus diesem Mengensystem sind. Die durch diese Einschränkung schlankere Theorie führt zu einer gesicherten Existenz einer optimalen Lösung des Segmentierungsproblems, die der gängigen Vorstellung einer Segmentierung genügt. Die Berechnung lokaler Optima ist algorithmisch innerhalb der Theorie umsetzbar.

Schlüsselwörter: Bildsegmentierung, vektorielle Mehrniveaupassung, Energieminimierung, stückweise stetige Approximation, Mumford-Shah-Modell

Danksagung

Nicht immer während meiner 5 Jahre am bayerischen Forschungsinstitut für wissensbasierte Systeme (FORWISS) in der Universität Passau habe ich geglaubt, daß ich eine Dissertation abschliessen werde. Bei einer Besprechung mit einem potentiellen Kunden antwortete ich auf die Frage, ob ich Mitarbeiter oder Promovend bei Prof. Donner sei, etwas verlegen mit: "Mitarbeiter". Es war Prof. Donner, der - ohne zu zögern - "und Promovend" ergänzte. Ich möchte an dieser Stelle Herrn Prof. Donner für sein Vertrauen zu mir und meiner Arbeit, die ich in dieser Form an anderer Stelle sicher nicht hätte abschliessen können, danken.

Weiter möchte ich Herrn Dr. Erich Fuchs danken, ohne dessen Engagement für das FORWISS im allgemeinen und meine Dissertation im besonderen ich nicht hätte promovieren können. Herr Dr. Fuchs, Herr Dipl. Inf. Holger Farr und Herr Dipl. Inf. Michael Kellner haben sich um meine Arbeit sicherlich besonders verdient gemacht.

Herrn Prof. Graf danke ich für die Übernahme der Zweitkorrektur und seine aufmunternden Worte zu Beginn des Rigorosums.

Herrn Dipl. Math. René Schöne danke ich für die Zusammenarbeit am Beweis von Lemma 3.6.16, Herrn Dipl. Inf. Martin Grajcar für zahlreiche Anregungen und Korrekturen.

Für Ihre Korrekturen danke ich auch Heike Engel, Dipl. Math. Simone Graf, Roman Hösl M.A., Roswitha Krieger, Marion Lang und Dr. Ralf Schuster.

Für Ihre Unterstützung während des Studiums danke ich Hedwig Emmerich.

Abschließend möchte ich mich bei meiner Freundin Roswitha Krieger für viele entgangene Stunden entschuldigen. Sie hat, zu meinem größten Bedauern, in den letzten zwei Jahren besonders unter dieser Arbeit gelitten.

Inhaltsverzeichnis

Symbolverzeichnis	iv
1 Einleitung	1
1.1 Einleitung	1
1.2 Aufbau der Arbeit	3
1.3 Grundlegende Begriffe	5
1.3.1 Bild, Pixel und diskretes Pixelbild	5
1.3.2 Bemerkungen zur Farbbildverarbeitung	9
2 Bildsegmentierungen	13
2.1 Einleitung	13
2.2 Der Begriff „Segmentierung“	14
2.2.1 Definitionen von Segmentierungen	14
2.2.2 Eigene Annäherung an den Begriff „Segmentierung“	16
2.3 Gängige Verfahren zur Segmentierung von Bildern	19
2.3.1 Schwellwertverfahren	19
2.3.2 Regionenwachstumsverfahren, „Region Splitting and Merging“	24
2.3.3 Wasserscheidentransformationen	25
2.3.4 Weitere Methoden zur Bildsegmentierung	26
2.3.5 Bemerkungen zur Segmentierung von Farbbildern	26
2.3.6 Bewertung von Segmentierungsverfahren	27
2.4 Bildsegmentierung mit Methoden der Variationsrechnung	31
2.4.1 „The weak membrane“	31
2.4.2 „The weak membrane“ im Diskreten	36
2.4.3 Eliminierung nicht-signifikanter lokaler Optima	39
2.4.4 Verwandte Ansätze	44
2.4.5 Bewertung	47
2.5 Quantisierung von Funktionen und Bildsegmentierungen	49
2.5.1 Optimale und fast optimale Quantisierung integrierbarer Funktionen	50
2.5.2 Quantisierung von stückweise konstanten Funktionen	57
2.5.3 Anwendung von Quantisierungsalgorithmen zur Bildsegmentierung	60
2.5.4 Bewertung	62
3 Theoretische Grundlagen der Mehrniveaupassung	63
3.1 Einleitung	63
3.2 Das Mengensystem	69

3.2.1	Der metrische Raum $(\mathfrak{A}, \mathfrak{d})$	69
3.2.2	Bedingungen an das Mengensystem \mathfrak{A}	74
3.3	Das System der Segmentierungen	76
3.4	Die Klasse der erlaubten Funktionen	84
3.5	Problemstellung	91
3.5.1	Die Zwei-Niveau-Situation	91
3.5.2	Die allgemeine Problemformulierung – optimale Segmentierung . . .	101
3.6	Existenz der Lösung	105
3.6.1	Der Satz von Fréchet und Kolmogorov	105
3.6.2	Der metrische Raum $\mathfrak{Z}_k(\mathfrak{A})$	110
3.6.3	Kompaktheit von $(\mathfrak{Z}_k(\mathfrak{A}), \delta_Z)$ und $(\mathfrak{S}_k(\mathfrak{A}), \mathfrak{d})$	112
3.6.4	Stetigkeit der Funktion Ψ	117
3.6.5	Existenz des Maximums	122
3.7	Bestimmung lokaler Maxima der Funktion Ψ	123
3.7.1	Reduktion auf 2-Niveaupassungsprobleme	123
3.7.2	Hinweise zur Bestimmung lokaler Minima von 2-Niveaupassungs- problemen	128
4	Algorithmen und Anwendungen der vektoriellen Mehrniveaupassung	148
4.1	Einleitung	148
4.2	Von infiniten zu finiten Systemen von Mengen	150
4.2.1	Pixelbasierte Segmentierungen	150
4.2.2	Exkurs: Approximation stückweiser konstanter Funktionen mit kon- stanten Funktionen	153
4.2.3	Pixelbasiertes lokales Maximum	158
4.3	Algorithmen zur Bestimmung eines pixelbasierten lokalen Maximums	160
4.3.1	Ein abstrakter Algorithmus zur Berechnung eines pixelbasierten lo- kalen Maximums einer k -Niveaupassung	160
4.3.2	Reduktion der k -Niveaupassung auf iterative 2-Niveaupassungen . .	160
4.3.3	2-Niveaupassungsalgorithmen	165
4.4	Eindimensionale Bilder und das Mengenmodell der zusammenhängenden Mengen	170
4.4.1	Grundsätzliches Vorgehen	170
4.4.2	Beispiele für die Approximation mit konstanten Funktionen	172
4.4.3	Beispiel für die Approximation mit quadratischen Funktionen	177
4.5	Das Mengenmodell der grauwertbenachbarten Pixelmengen	180
4.5.1	Die Mehrniveaupassung als Schwellwertverfahren	180
4.5.2	Zusammenhang zwischen 2-Niveaupassung und der Varianzanalyse von Otsu	181
4.5.3	Zusammenhang zwischen Quantisierung und 2-Niveaupassung	182
4.5.4	Ein Algorithmus zur Bestimmung von k Schwellwerten	184
4.5.5	Beispiele	188
4.5.6	Übertragung auf Farbbilder	190
4.6	Das Mengenmodell der pixelzusammenhängenden Mengen	191
4.6.1	Definition des Mengenmodells der pixelzusammenhängenden Mengen	191
4.6.2	Die pixelbasierte Umgebung einer Segmentierung mit pixelzusammen- hängenden Mengen	193

4.6.3	Algorithmische Überprüfung von Pixelzusammenhang	196
4.6.4	Anmerkungen zur Segmentierung von Farbbildern	197
4.6.5	Initialisierung	200
4.6.6	Der Algorithmus	202
4.6.7	Beispiele	208
4.6.8	Einfach pixelzusammenhängende Pixelmengen	219
5	Niveaureduktion und Schätzung der Anzahl der Segmente	221
5.1	Einleitung	221
5.2	Pareto-Optima	223
5.3	Niveaureduktion	226
5.4	Schätzung der Anzahl der Segmente	235
5.4.1	Der Approximationsfehler der Mehrniveaupassung	235
5.4.2	Der größte normalisierte Fehler	238
5.4.3	Ein Algorithmus zur Schätzung der Anzahl der Segmente über den größten normalisierten Fehler	241
5.4.4	Beispiele	242
5.4.5	Abschätzung über den Sensorfehler	249
5.4.6	Andere Fehlerfunktionen zur Abschätzung der Anzahl der Segmente	250
6	Schlußbemerkungen und Ausblick	251
6.1	Einleitung	251
6.2	Das Konzept der vektoriellen Mehrniveaupassung als Segmentierungsver- fahren	252
6.2.1	Zusammenfassung: Vektorielle Mehrniveaupassung – Anwendungen in der Bildsegmentierung	252
6.2.2	Vergleich mit den Bedingungen von Haralick und Shapiro	253
6.2.3	Vergleich mit den Forderungen von Mumford und Shah	254
6.2.4	Vergleich mit den notwendigen Bedingungen einer Segmentierung	254
6.3	Vergleich zu anderen Segmentierungsverfahren	255
6.3.1	Vergleich mit den Methoden der Variationsrechnung	255
6.3.2	Vergleich mit den Funktionsquantisierungsansätzen	260
6.3.3	Einordnung der vektoriellen Mehrniveaupassung als Segmentierungs- verfahren	260
6.4	Ausblick	262
6.4.1	Praktischer Ausblick	262
6.4.2	Theoretischer Ausblick	266
6.5	Abschließende Bemerkung	267
	Liste der Algorithmen	268
	Index	270
	Literaturverzeichnis	273

Symbolverzeichnis

$:\Leftrightarrow$	definitorische Äquivalenz
$:=$	definitorische Gleichheit
$\delta_{ij} := \begin{cases} 1 & , \text{ falls } i = j \\ 0 & , \text{ sonst} \end{cases}$	Kroneckersymbol
$\dim V$	Dimension des Vektorraums V
\mathbf{A}^t	transponierte Matrix zur Matrix \mathbf{A}
\mathbf{I}	$n \times n$ Einheitsmatrix (n sollte aus Zusammenhang bekannt sein)
$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^t = (x_1, \dots, x_n)$	transponierter Vektor
$\mathbf{0}$	Nullvektor
λ	Lebesgue-Maß
$\mathbb{R}_+ := [0, \infty[$	nicht-negative reelle Zahlen
$\mathbb{N} := \{1, 2, \dots\}$	Menge der natürlichen Zahlen
$\mathbb{N}_0 := \mathbb{N} \cup \{0\}$	Menge der natürlichen Zahlen mit 0,
$\text{round}(x) := \max\{z \in \mathbb{Z} \mid z \leq x + \frac{1}{2}\}$	Rundung für reelle Zahlen
$\text{sgn}(x) = \begin{cases} 1 & , \text{ falls } x > 0 \\ 0 & , \text{ falls } x = 0 \\ -1 & , \text{ sonst} \end{cases}$	Vorzeichen von x
$A \uplus B$	Vereinigung disjunkter Mengen, (d. h. es gilt zusätzlich $A \cap B = \emptyset$)
$\{f = y\}$	Abkürzung für die Menge $\{x \mid f(x) = y\}$
$ A $	Kardinalität der Menge A
$\wp(M)$	Potenzmenge von M
$\wp_k(M) := \{A \in \wp(M) \mid A \leq k\}$	Potenzmenge von M mit höchstens k Elementen
$\mathcal{B}(M)$	Borelsche σ -Algebra in M

$\mathcal{C}(A, B)$	Raum der stetigen Funktionen von A nach B
$\mathcal{C}^1(A, B)$	Raum der einmal stetig differenzierbaren Funktionen von A nach B
$\mathcal{L}^p(A, B)$	Raum der p -fach integrierbaren meßbaren Funktionen von A nach B
$L^p(A, B)$	Quotientenraum der p -fach integrierbaren Funktionen von A nach B (vgl. z. B. [Bau92], S.99)
$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}) \right)$	Gradient von $f : \mathbb{R}^n \rightarrow \mathbb{R}$ an der Stelle $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$
$\ x\ _2 := \sqrt{\sum_{i=1}^n x_i^2}$	euklidische Norm für $x \in \mathbb{R}^n$
$\ x\ _2 := \sqrt{\langle x x \rangle}$	euklidische Norm für $x \in E$, dabei ist E ein Raum mit Skalarprodukt $\langle \cdot \cdot \rangle$
$\ x\ _\infty := \max\{ x_i \mid i \in \{1, \dots, n\}\}$	Maximumsnorm für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$
$\ f\ _p := \left(\int \ f(x)\ _2^p d\lambda(x) \right)^{\frac{1}{p}}$	L^p -Norm auf $L^p(A, B)$
id	identische Abbildung $x \mapsto x$
\mathcal{O}	\mathcal{O} -Notation zur Beschreibung der Laufzeitkomplexität (vgl. etwa [AHU83], S. 17)

Bemerkung

Abgesetzte Formeln im Text enthalten bewußt keine Satzzeichen. Die Lesbarkeit der Formel wurde über die Lesbarkeit des Satzes gestellt.

Kapitel 1

Einleitung

1.1 Einleitung

Die Erkennung von Objekten in Bildern vollzieht das menschliche Gehirn unbewußt und somit erscheint dieser Prozeß selbstverständlich. Betrachtet man jedoch die Menge von scheinbar ungeordneten Zahlen, die das Ergebnis einer digitalen Kodierung eines Bildes darstellt, wird die hohe Komplexität von algorithmischen Bildverarbeitungsprozessen deutlich. Es erscheint unmöglich, eine Relation zwischen den Daten und den abgebildeten Objekten im Bild zu erkennen. Tatsächlich ist dies beim heutigem Stand der Technik auch nur in sehr begrenztem Umfang möglich. Das ist allerdings kein Grund, sich nicht mit der automatischen Auswertung von Bildern zur Objekterkennung zu beschäftigen.

Die Bedeutung der Segmentierung innerhalb der Bildanalyse läßt sich vielleicht am besten anhand eines Zitats aus dem Bildverarbeitungsbuch von Rafael Gonzales und Richard Woods veranschaulichen:

The first step in image analysis generally is to segment the image ([GW92], S.413)

Bei jeder Art von Bildauswertung ist implizit oder explizit ein Segmentierungsschritt, ein Schritt, der Pixel sinnvoll gruppiert, gegeben. So einfach die Bildauswertung auch sein möge, immer wird einzelnen Pixeln eine Objekt-Bedeutung zugewiesen.

Man kann an dieser Stelle zwischen Objekt- und Bildsegmentierung unterscheiden:

Bei der *Objektsegmentierung* versucht man, ein vorher klar definiertes Objekt, z. B. einen schwarzen Punkt vor einem weißen Hintergrund, im Bild zu bestimmen.

Bei der *Bildsegmentierung* versucht man, ohne detailliertes Objektwissen das gegebene Bild in sinnvolle Bereiche aufzuteilen. Ziel einer Bildsegmentierung muß es sein, von der Ebene der einzelnen Pixel ohne größeren Informationsverlust zu Objekten mit größerer Aussagekraft zu gelangen. Im Gegensatz zu den Pixeln kann man diesen Objekten mehrere Eigenschaften wie z. B. Form, Textur oder Ähnlichkeit zu anderen Objekten zuordnen. Viele High-Level Bildverarbeitungsalgorithmen basieren auf der Interpretation solcher Objekte.

Fehler bei der Bildsegmentierung wirken sich in allen nachfolgenden Schritten aus und können das Ergebnis eines Bildverarbeitungsprozesses verfälschen oder gar verhindern. Daher ist die Bildsegmentierung eines der zentralen Themen der Bildverarbeitung.

Die vorliegende Arbeit konzentriert sich auf diesen Bereich, der im folgenden einfach nur „Segmentierung“ genannt wird, und soll mit der vektoriellen Mehrniveaupassung zur Verbesserung der vorhandenen Verfahren beitragen.

Diese Arbeit entstand mit Unterstützung der Europäischen Union im Rahmen des Brite-Euram Projekts IntelliWood (IMS 29680) am Bayerischen Forschungsinstitut für wissenschaftsbasierte Systeme (FORWISS).

1.2 Aufbau der Arbeit

Im folgenden zweiten Kapitel wird der Begriff der „Segmentierung“ beleuchtet, der, obwohl in der Bildverarbeitung ständig präsent, in den meisten Arbeiten zum Thema „Segmentierung“ überhaupt nicht oder nur unzureichend definiert wird. Auch die vorliegende Arbeit kann nur Denkanstöße und den Versuch einer Definition geben. Es sollen aber zumindest notwendige Kriterien entwickelt werden, die Verfahren zur „Segmentierung“ von Bildern erfüllen sollten. Anschließend werden im zweiten Kapitel aus der großen Zahl der Segmentierungsverfahren einige gängige Methoden und Ansätze zur „Segmentierung“ von Bildern vorgestellt. Im Mittelpunkt des Kapitels werden zwei approximierende Verfahren zur Segmentierung stehen: Auf der einen Seite werden Methoden der Variationsrechnung in der Bildsegmentierung behandelt, die in der Literatur vor allem durch Mumford und Shah bekannt geworden sind, auf der anderen Seite wird der Funktionsquantisierungsansatz von Kämpke und Kober betrachtet. Diese beiden Ansätze stehen in einem direkten Zusammenhang zu der vorliegenden Arbeit und erfüllen die in dieser Arbeit aufgestellten Qualitätskriterien besonders gut. Den Ansätzen ist gemeinsam, daß sie das Bild durch stückweise stetige Funktionen, deren maximale Stetigkeitsbereiche dabei die Segmente bilden, approximieren. Die Methoden der Variationsrechnung versuchen ein Energiefunktional, bestehend unter anderem aus einem Maß der Bildtreue, durch eine allgemeine stetige Funktion zu minimieren. Während im allgemeinen eine Quantisierung über den Erwartungswert einer Zufallsvariablen definiert wird, beschäftigen sich die Arbeiten von Kämpke und Kober mit der Quantisierung von Funktionen bezüglich des euklidischen Abstands. Das heißt unter anderem, daß sie die gegebenen Funktionen mit stückweise konstanten, also insbesondere stückweise stetigen, Funktionen approximieren. Allerdings beschränken sie sich dabei auf Funktionen mit eindimensionalem Definitionsbereich und nutzen lediglich Quantisierungen des Histogramms eines Bildes zur Segmentierung desselben.

Die Quantisierung einer Funktion, die auf einem eindimensionalen Intervall definiert ist, liefert implizit eine Segmentierung dieser Funktion durch zusammenhängende Mengen. Den Schritt von ein- zu mehrdimensionalen Definitionsbereichen vollzieht die vektorielle Mehrniveaupassung, deren theoretische Grundlagen in Kapitel 3 gelegt werden. Dabei wird das Konzept um das System der erlaubten Segmentierungen erweitert. Dieses Konzept erlaubt es, Bedingungen an die jeweiligen Segmentierungen heranzutragen: So kann man z. B. fordern, daß eine Segmentierung nur aus rechteckigen oder nur aus zusammenhängenden Mengen bestehen soll. Viele Ergebnisse, die Kämpke und Kober für die Approximation durch konstante Funktionen gewonnen haben, lassen sich auch für die Approximation durch Funktionen eines endlich-dimensionalen Vektorraums unter gewissen Bedingungen aufrecht erhalten. Dabei wird auch deutlich werden, daß der Ansatz der Mehrniveaupassung den wesentlichen Kern des Ansatzes von Kämpke und Kober mit einschließt und dem Ansatz von Mumford und Shah nahe kommt, indem er unter anderem die vielen Freiheiten, die dieser Ansatz bietet, einschränkt und dadurch auf das Ziel der Bildsegmentierung hin ausrichtet. Ein wesentlicher Teil des dritten Kapitels dieser Arbeit ist auch dem Beweis der Existenz einer optimalen Segmentierung im Kontext der Mehrniveaupassung gewidmet. Der Existenzbeweis gestaltet sich dabei sowohl in der Aussage als auch in der Beweisführung einfacher als der innerhalb des Kontexts des Mumford-Shah-Modells.

Jede Theorie sollte sich auch an ihrer praktischen Umsetzung messen lassen: In Kapitel 4 werden zahlreiche Anwendungen der Mehrniveaupassung zu verschiedenen Segmentierungs-

problemen vorgestellt. Im Gegensatz zum Mumford-Shah-Modell erfolgt bei der vektoriel-
len Mehrniveaupassung die Diskretisierung innerhalb der Theorie. Dies ist ein klarer Vor-
teil des festen Mengensystems und der Einschränkung auf einen endlichen Vektorraum zur
Approximation. Gänzlich unproblematisch ist die Diskretisierung allerdings nicht: Unter
Umständen ist das globale Optimum einer Mehrniveaupassung nicht in der Diskretisierung
des Mengenmodells enthalten, wie in Kapitel 4 gezeigt wird.

Das Kapitel 5 beschäftigt sich mit einem Thema, das viele Segmentierungsansätze nur
indirekt angehen oder umgehen. Es handelt sich dabei um die Frage, wieviele Objekte
oder Segmente in einem gegebenen Bild enthalten sind. Innerhalb der Mehrniveaupass-
ung versucht diese Arbeit, einige Denkanstöße und Heuristiken zu geben, diese Frage als
Mehrzieloptimierungsproblem zu formulieren und zu lösen. In zahlreichen Fällen führen
die angegebenen Algorithmen zu zufriedenstellenden Lösungen, ein allgemeines, mathe-
matisch fundiertes Vorgehen läßt sich jedoch nicht daraus ableiten.

In Kapitel 6 wird dann noch einmal ein Resümee der Anwendungen der vektoriel-
len Mehrniveaupassung in der Bildsegmentierung gezogen. Außerdem werden die Methoden der
Mehrniveaupassung in die Welt der Segmentierungsverfahren eingeordnet. Dabei verdienen
die Methoden der Variationsrechnung aufgrund ihrer theoretischen Nähe zur vektoriel-
len Mehrniveaupassung eine besondere Aufmerksamkeit.

1.3 Grundlegende Begriffe

Der zentrale Begriff dieser Arbeit ist sicherlich der Begriff der „Segmentierung“. Deshalb wird auf diesen Begriff in einem eigenen Abschnitt, nämlich in Abschnitt 2.2.1, gesondert eingegangen.

Um jedoch überhaupt von „Bildverarbeitung“ sprechen zu können, müssen zunächst die grundlegenden Begriffe wie „Bild“, „Pixel“ oder „Nachbarschaft“ definiert werden, da sie, je nach Autor, in der Literatur unterschiedlich aufgefaßt werden. Besondere Aufmerksamkeit widmet dieses Unterkapitel dem Thema „Farbbildverarbeitung“ und den besonderen Schwierigkeiten, die das Arbeiten mit Farbbildern mit sich bringt.

1.3.1 Bild, Pixel und diskretes Pixelbild

Sei für $m_1, m_2 \in \mathbb{R}_+$ die Menge $M := [0, m_1] \times [0, m_2]$ der Definitionsbereich eines Bildes. M heißt auch *Mattscheibe*. Auf M sei ein Bild eine quadratisch integrierbare Abbildung $f : M \rightarrow \mathbb{R}$ oder $f : M \rightarrow \mathbb{R}^d$, wobei sich z. B. ein Farbbild durch einen Wertebereich im \mathbb{R}^3 auszeichnen kann, wie der Abschnitt 1.3.2 zeigen wird. Ein Computerbild dieses Bildes kann nur einen endlichen Definitions- und Wertebereich besitzen. Daher bedeutet ein Computerbild immer eine Diskretisierung eines kontinuierlichen Bildes durch eine physikalische Einheit, z. B. einem Framegrabber. Ein einfaches Modell dieser Digitalisierung ist die Integration über ein Pixel. Für eine genauere Darstellung des gesamten Aufnahmeprozesses sei auf [Haa00] oder [Lau03] verwiesen.

Ein Pixel¹ sei ein quadratischer Bereich² mit Fläche 1 auf der Mattscheibe.

Definition 1.3.1 (Pixel)

Für ein $l \in \mathbb{N}$ seien $m_1, \dots, m_l \in \mathbb{N}$. Weiter sei $M := [0, m_1] \times \dots \times [0, m_l]$ eine Mattscheibe³.

- i. Ein *Pixel* ist dann eine Region mit Kantenlänge 1, die durch ihre Koordinaten repräsentiert wird: Für natürliche Zahlen i_1, \dots, i_l mit $0 \leq i_1 < m_1, \dots, 0 \leq i_l < m_l$ sei

$$[i_1, \dots, i_l] := [i_1, i_1 + 1[\times \dots \times [i_l, i_l + 1[$$

das Pixel mit den Koordinaten i_1, \dots, i_l .

- ii. Die Menge

$$\mathbb{P}(M) := \{ [i_1, \dots, i_l] \mid 0 \leq i_1 < m_1, \dots, 0 \leq i_l < m_l \}$$

heiße die *Menge aller Pixel von M*.

¹engl. Abkürzung für *Picture Element*.

²idealisierte Annahme, die aber o. B. d. A. machbar ist.

³Der Begriff „Mattscheibe“ macht anschaulich nur für $l = 2$, vielleicht noch $l = 1$, Sinn. Aber rein mathematisch läßt sich das Konzept für endliche l kanonisch fortsetzen.

iii. Es sei

$$\mathcal{P}(M) := \left\{ \bigcup_{i=1}^n p_i \mid p_i \in \mathbb{P}(M), 1 \leq i \leq n \text{ für ein } n \in \mathbb{N} \right\} = \wp(\mathbb{P}(M))$$

die Menge der *pixelbasierten Mengen in M*.

iv. Für eine Menge $A \in \mathcal{P}(M)$ sei

$$\mathbb{P}(A) := \{p \in \mathbb{P}(M) \mid p \subseteq A\}$$

die *Menge aller Pixel in A*.

Definition 1.3.2 (Diskretes Pixelbild)

Eine Funktion $g : M \rightarrow \mathbb{R}^d$ heißt *diskretes Pixelbild*, falls für alle $p \in \mathbb{P}(M)$ die Funktion $g|_p$ konstant ist. Daher ist es legitim, wenn auch mathematisch nicht ganz korrekt, $g(p)$ statt $g|_p$ zu notieren. g ist allerdings immer noch über einen kontinuierlichen Definitionsbereich definiert, obwohl sowohl der Name als auch die Notation $g(p)$ etwas anderes andeuten.

Ist der Wertebereich eines diskreten Pixelbildes g eine Teilmenge der natürlichen Zahlen einschließlich der 0, so heie g ein *diskretes Grauwertpixelbild*. Ist der Wertebereich hingegen eine Teilmenge des \mathbb{N}_0^3 , so heie g *diskretes Farbpixelbild*.

Bemerkung 1.3.3 (Subpixelgenauigkeit)

Nicht immer besteht das Ergebnis eines Bildverarbeitungsprozesses auf einem diskreten Pixelbild aus einer Pixelmenge. Der Mittelpunkt eines Kreises etwa kann als Ergebnis einer Kreispassung auf diskreten Pixelkoordinaten durchaus nicht ganzzahlige Koordinaten haben. Da in diesem Fall das Ergebnis unterhalb der Genauigkeit der Pixelkoordinaten ermittelt wurde, spricht man von *Subpixelgenauigkeit*.

Definition 1.3.4 (Histogramm)

Sei g ein diskretes Grauwertpixelbild mit Werten in \mathbb{N}_0 und sei $G := \max_{p \in \mathbb{P}(M)} g(p)$. Dann

heißt

$$h_g := \begin{array}{l} \{0, \dots, G\} \rightarrow \mathbb{N}_0 \\ i \mapsto |g^{-1}(i)| = |\{g = i\}| \end{array}$$

das *Histogramm des diskreten Pixelbildes g*.

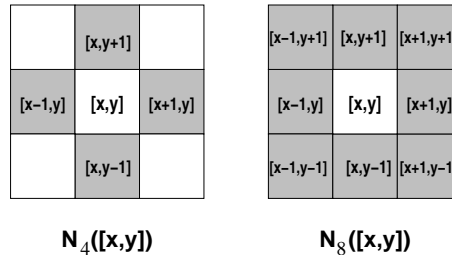


Abbildung 1.1: 4-er bzw. 8-er Nachbarschaft eines Pixels $[x, y]$ im Inneren eines Bildes.

Definition 1.3.5 (4-er und 8-er Nachbarschaft)

Sei für $m_1, m_2 \in \mathbb{R}_+$ die Mattscheibe $M := [0, m_1] \times [0, m_2]$. Für ein Pixel $[x, y] \in \mathbb{P}(M)$ heißt die Menge

$$N_4([x, y]) := \{[x', y'] \mid [x', y'] \in \mathbb{P}(M), |x' - x| + |y' - y| = 1\}$$

die *4-er Nachbarschaft* von $[x, y]$.

Die Menge

$$N_8([x, y]) := \{[x', y'] \mid [x', y'] \in \mathbb{P}(M), \max\{|x' - x|, |y' - y|\} = 1\}$$

heißt die *8-er Nachbarschaft* von $[x, y]$.

Bemerkung 1.3.6

Trivial folgt für zwei Pixel $p, q \in \mathbb{P}(M)$ mit $p \in N_4(q)$, daß auch $q \in N_4(p)$ liegt. Dies gilt natürlich auch für die 8-er Nachbarschaft.

Weiter folgt aus der Definition 1.3.5 für alle $p \in \mathbb{P}(M) : N_4(p) \subset N_8(p)$. Liegt das Pixel p nicht am Rand des Bildes, so hat die Menge $N_4(p)$ vier und die Menge $N_8(p)$ acht Elemente. Daher rührt die Namensgebung für diese Pixelnachbarschaften. Abbildung 1.1 veranschaulicht die geometrische Bedeutung der Definition 1.3.5.

Definition 1.3.7 (Pixelzusammenhang)

Eine Pixelmenge $P \in \mathcal{P}(M)$ heißt *pixelzusammenhängend* bezüglich der 4-er bzw. 8-er Nachbarschaft, falls für alle Pixel $p_1, p_2 \in P$ ein Pfad $(q_i)_{1 \leq i \leq n}$ existiert mit

- i. $q_1 = p_1$
- ii. $q_k = p_2$
- iii. $\forall i \in \{1, \dots, k-1\} : q_i \in P \quad \text{und} \quad q_{i+1} \in N_4(q_i) \text{ bzw. } N_8(q_i)$

Die leere Pixelmenge ist per Definition pixelzusammenhängend.

Bemerkung 1.3.8

Der Begriff des Pixelzusammenhangs kann zu sprachlichen Paradoxien führen: Zunächst einmal ist klar, daß, falls eine Pixelmenge P pixelzusammenhängend bezüglich der 4-er oder 8-er Nachbarschaft ist, die Menge $\bigcup_{p \in P} p$ zusammenhängend ist. Auf der anderen Seite ist es bei der 4-er Nachbarschaft durchaus möglich, daß eine zusammenhängende Menge aus Pixeln nicht pixelzusammenhängend ist. Abbildung 1.1 verdeutlicht die Problematik: Die aus den dunkel eingefärbten Pixeln bestehende Menge, die die 4-er Nachbarschaft skizziert, ist bezüglich der 4-er Nachbarschaft nicht pixelzusammenhängend, obwohl die Menge bestehend aus diesen Pixeln zusammenhängend ist.

Bei der 8-er Nachbarschaft tritt dieses Phänomen nicht auf.

Lemma 1.3.9

Sei $N_* := N_4$ bzw. $N_* := N_8$.

Sei $g : M \rightarrow \mathbb{R}^d$ ein diskretes Pixelbild. Auf $\mathbb{P}(M)$ sei eine Relation \sim definiert durch

$$p_1 \sim p_2 \Leftrightarrow g(p_1) = g(p_2) \wedge p_2 \in N_*(p_1)$$

für zwei Pixel $p_1, p_2 \in \mathbb{P}(M)$.

Dann ist die transitive Hülle der Relation \sim eine Äquivalenzrelation.

Beweis:

vgl. [Vos93], S. 13

Definition 1.3.10 (Pixelzusammenhangskomponente)

Sei $g : M \rightarrow \mathbb{R}^d$ ein diskretes Pixelbild.

Dann heißen die Äquivalenzklassen der transitiven Hülle der Relation \sim aus Lemma 1.3.9 *Pixelzusammenhangskomponenten* bezüglich der 4-er bzw. 8-er Nachbarschaft. Das heißt, eine Pixelzusammenhangskomponente in g ist eine maximal pixelzusammenhängende Menge von Pixeln gleichen Wertes.

1.3.2 Bemerkungen zur Farbbildverarbeitung

Das physio- und psychologische Phänomen des menschlichen Farbsehens ist in seiner Gesamtheit noch nicht entschlüsselt. Die physikalischen Grundlagen des Begriffes Farbe können allerdings in formalen Begriffen formuliert und durch experimentelle und theoretische Ergebnisse gestützt werden (vgl. [GW92], S. 222).

Grundsätzlich ist die Farbe eines Objektes, die das menschliche Auge wahrnimmt, dadurch definiert, welche Wellenlängen des Lichtes von dem Objekt reflektiert und welche absorbiert werden. Eine unbunte Farbe (Schwarz, Weiß und alle Graustufen) läßt sich durch einen Wert, ihre Helligkeit, beschreiben. Daher haben diskrete Grauwertbilder eine Teilmenge des \mathbb{N}_0 als Wertebereich.

Für die Beschreibung einer bunten Farbe sind jedoch drei Werte nötig. Eine Farbe läßt sich also als Vektor eines dreidimensionalen Vektorraums kodieren (vgl. [FBS90]). Eine Möglichkeit dies zu tun, ist es, alle Farben als Kombination der drei sogenannten *Primärfarben* rot, grün und blau zu kodieren. 1931 hat die CIE (Commission Internationale de l'Eclairage) die spezifischen Wellenlängen der Primärfarben standardisiert: Blau hat eine Wellenlänge von 435.8nm, grün 546.1nm und rot genau 700nm. Diese Primärfarben können nun addiert werden und ergeben die sogenannten *Sekundärfarben*: Magenta als Kombination von rot und blau, cyan als grün und blau, und gelb als Addition von rot und grün. Auf der Grundlage dieser Kodierungsmöglichkeit entstand das RGB-Farbmodell und das CMY-Farbmodell (vgl. 1.3.2.1.1). Alle anderen Farben ergeben sich in diesen Modellen als Linearkombinationen der Primärfarben, bzw. – im CMY-Farbmodell – der Sekundärfarben.

Eine andere Möglichkeit, Farben als dreidimensionale Vektoren zu kodieren, ist die Charakterisierung der Farben als Mischung aus Farbton (Hue), Sättigung (Saturation) und Helligkeit (Intensity). Dabei ist der Farbton die dominante Wellenlänge im Wellengemisch des Lichts, das der Beobachter aufnimmt. Die Sättigung ist ein Wert für den Anteil von weißem Licht, das mit dem Farbton gemischt die Farbe ergibt. Auf der Grundlage dieser Kodierungsmöglichkeit entstand das HSI-Farbmodell (vgl. 1.3.2.1.2).

1.3.2.1 Farbmodelle

Der Sinn von Farbmodellen in der Bildverarbeitung ist es, standardisierte Methoden bereitzustellen, um mit Farbbildern umgehen zu können. Allen Farbmodellen ist gemeinsam, daß sie die einzelnen Farben als Punkte in einem dreidimensionalen Koordinatensystem darstellen. Als Hauptmodelle haben sich dabei in der Geschichte der Bildverarbeitung das RGB- und das HSI-Modell herauskristallisiert.

1.3.2.1.1 Das RGB-Farbmodell

Der Vorteil des RGB-Farbmodells ist, daß ihm ein kartesisches Koordinatensystem zugrunde liegt. Jede Farbe wird als Linearkombination der Intensitäten der drei Primärfarben rot, grün und blau dargestellt. Der Farbraum ist dann ein Würfel mit Ursprung Schwarz

und rot, grün und blau als Eckfarbwerte an den Achsen. Die restlichen Ecken des Quadrats sind dann magenta (= rot + blau), cyan (= grün + blau), gelb (= rot + grün) und weiß (= rot + grün + blau). Die unbunten Farben, das heißt die Grautöne, liegen in diesem Farbmodell auf der Geraden durch die Punkte Schwarz und Weiß. Abbildung 1.2 veranschaulicht dieses Farbmodell. Zahlreiche in der Bildverarbeitung bekannte Farb-

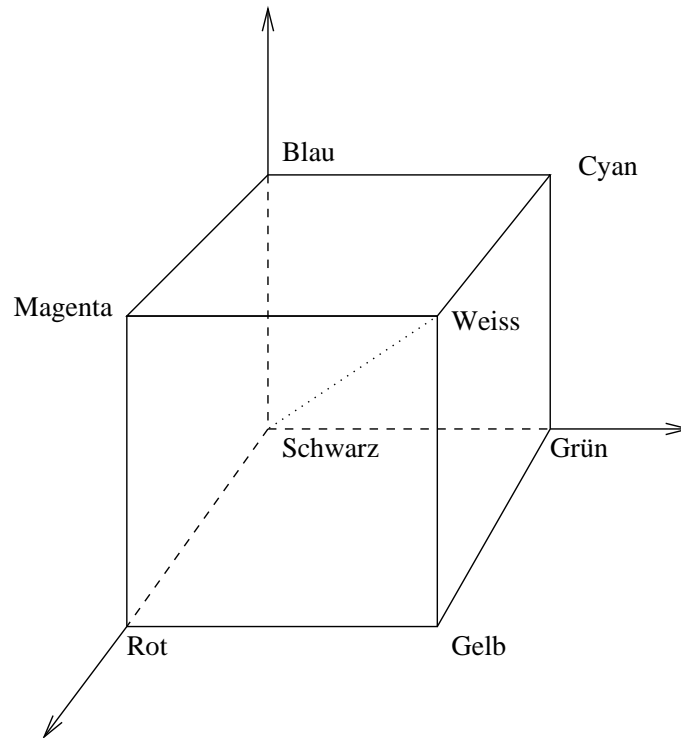


Abbildung 1.2: Skizze des RGB-Farbwürfels mit den Ecken rot, grün, blau, magenta, cyan, gelb sowie schwarz und weiß. Auf der gepunkteten Linie zwischen Schwarz und Weiß sind die Grauwerte angeordnet.

modelle lassen sich durch eine lineare Abbildung auf dieses Farbmodell zurückführen. Als Beispiele seien die Modelle CMY (= Cyan, Magenta, Yellow), YIQ oder YUV genannt (vgl. [GW92] oder [PB99]).

1.3.2.1.2 Das HSI-Farbmodell

Das HSI-(= Hue, Saturation, Intensity) Farbmodell, oft auch HSV-Farbmodell genannt, wobei das V für „Value“ steht, wurde 1978 von Alvy Ray Smith entwickelt. Es spaltet die Farbinformation in zwei Teile: Einen unbunten (achromatischen) Anteil, nämlich die Helligkeit oder Intensität, und einen bunten (chromatischen) Anteil, nämlich Sättigung und Farbton. Diese Aufteilung kommt dem menschlichen Farbsehen näher als die RGB-Darstellung. Farbton und Sättigung (engl. Hue and Saturation) spannen ein Dreieck auf, wie es in Abbildung 1.3 skizziert wird.

Alle Farben (= Farbton + Sättigung) liegen in diesem Dreieck, das von Primärfarben

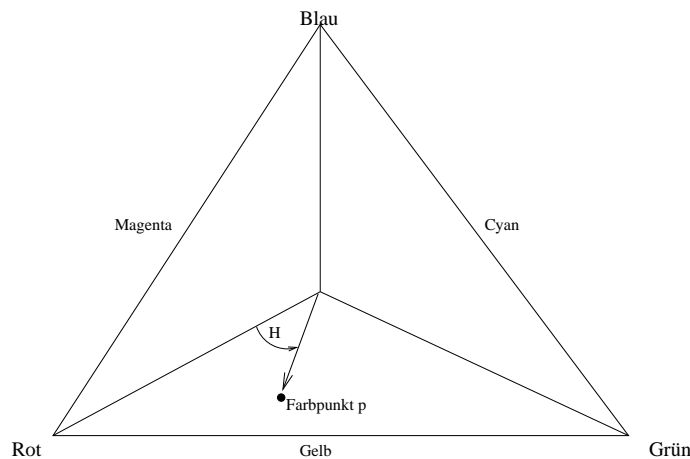


Abbildung 1.3: Skizze der Farbton und Sättigungsebene („Hue“ und „Saturation“). Der Farbton ist der Winkel zur Rotachse, die Sättigung ist der Abstand des Farbpunkts P zum Zentrum des Dreiecks.

rot, grün und blau aufgespannt wird. Jede Farbe P in diesem Dreieck ist durch zwei Werte gegeben: dem Farbwinkel, relativ zur Rotachse und dem Abstand des Farbwerts zum Mittelpunkt des Farbdreiecks. Dieser Abstand wird relativ zur Größe des Dreiecks angegeben, da die Größe des Farbdreiecks abhängig ist von der Intensität (vgl. Abbildung 1.4). Dieser Abstand entspricht dann der Sättigung. Der Farbwinkel kodiert den Farbton. So entspricht ein Winkel von 0° rot, ein Winkel von 60° gelb und ein Winkel von 120° grün.

Ein Umrechnungsalgorithmus einer Darstellung bezüglich des RGB-Farbmodells zu einer Darstellung im HSI-Farbmodell und umgekehrt findet man zum Beispiel in [GW92], S. 229 – 238. Weitere Farbmodelle, die in analoger Weise der Trennung der chromatischen und achromatischen Farbanteile folgen sind z. B. das CIEL^{*} u^*v^* - oder das CIEL^{*} a^*b^* -Farbmodell (vgl. [WS82]).

1.3.2.2 Probleme einer Metrik in Farbmodellen

Das Hauptproblem beim Vergleich von Farben ist es, die Perzeption der Farben durch das menschliche Auge nachzuahmen. Dieses Problem kann grundsätzlich wohl nie gelöst werden, da, wie eingangs erwähnt, das Farbsehen des Menschen nicht gänzlich erforscht und sicher auch individuell verschieden ist. Trotzdem wäre es schön, in den jeweiligen Farbmodellen eine Metrik definieren zu können, die die folgende Eigenschaft hat:

Seien f_1, f_2, f_3 drei Werte innerhalb eines Farbmodells und d eine Metrik des Farbmodells. Dann sollte aus $d(f_1, f_2) < d(f_1, f_3)$ folgen, daß die Mehrheit der Menschen empfindet, die Farbe, die f_1 zugeordnet ist, ist der Farbe von f_2 ähnlicher, als der von f_3 .

Diese Forderung ist natürlich nur ein nicht zu erreichendes Ideal, an dem aber mögliche Metriken der Farbmodelle gemessen werden sollten.

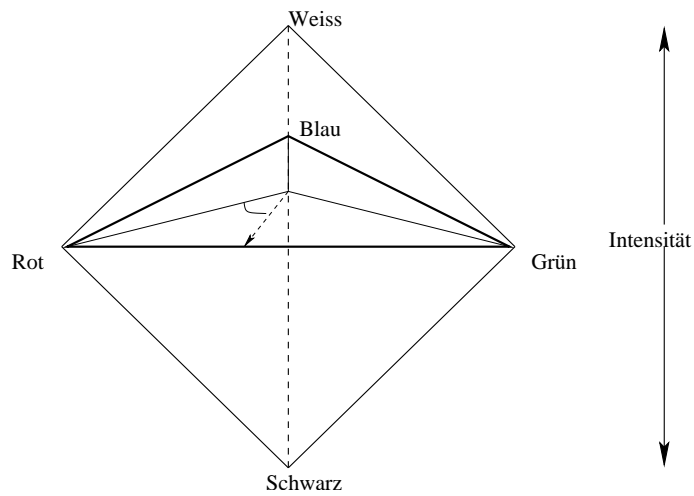


Abbildung 1.4: Skizze der Darstellung von Farben im HSI-Farbmodell. Für jede Intensität existiert ein Farbdreieck aufgespannt von Farbton und Sättigung.

Die naheliegendste Metrik für das RGB-Farbmodell ist sicherlich die euklidische. Die Nachteile dieser Metrik werden u. a. in [PB99] dargelegt. Dort wird zunächst die Koordinatentransformation des RGB-Farbwürfels in den YUV-Farbwürfel mittels der Vorschrift

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.1 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

vorgeschlagen. In diesem Koordinatensystem ist dann wieder die euklidische Metrik anzuwenden, was speziell im Zusammenhang mit der vektoriellen Mehrniveaupassung von Farbbildern (vgl. speziell Abschnitt 4.6) wünschenswert erscheint. Eine Übersicht von verschiedenen Ansätzen einer Metrik, die zumindest versucht, die oben aufgestellte Forderung zu erfüllen, findet man in [Haf98], Abschnitt 4.3.2, oder in [WS82], Chapter 3.

Kapitel 2

Bildsegmentierungen

2.1 Einleitung

Das Ziel dieses Kapitels ist es, sich zunächst einmal mit der Welt der Bildsegmentierung auseinanderzusetzen. Dazu ist es nötig, sich dem Begriff der „Segmentierung“ zu nähern. Dies erscheint, da der Begriff in der Bildverarbeitung sehr häufig anzutreffen ist, einfacher, als es wirklich ist. Das erste Unterkapitel dieses Kapitels wird die anerkanntesten Definitionen kritisch vorstellen und auch einen eigenen Ansatz präsentieren, an dem sich die Anwendungen der vektoriellen Mehrniveaupassung in der Bildverarbeitung messen sollen.

Im zweiten Unterkapitel werden die gängigsten Methoden zur Segmentierung von Bildern präsentiert. Aufgrund der Fülle an Algorithmen und Ansätzen kann im Unterkapitel allerdings nur der Versuch einer einigermaßen vollständigen Übersicht und Einordnung der verbreitetsten Methoden gemacht werden.

Der Schwerpunkt dieses Kapitels liegt allerdings bei den Methoden der Variationsrechnung zur Bildsegmentierung, sowie bei den Ansätzen der Segmentierung über die Quantisierung von Funktionen. Die vektorielle Mehrniveaupassung wird Anregungen dieser beiden Segmentierungsansätze aufnehmen. Aufgrund dieser Nähe ist beiden Ansätzen jeweils ein eigenes Unterkapitel gewidmet.

2.2 Der Begriff „Segmentierung“

Der Begriff „Segmentierung“ bzw. „Bildsegmentierung“ ist sicherlich einer der zentralen Begriffe der Bildverarbeitung überhaupt. Es gibt jedoch keine einheitliche Definition des Begriffes.

Im folgenden werden zunächst einige bekanntere Versuche in der Literatur, den Begriff „Segmentierung“ zu definieren, vorgestellt. Besonders hervorzuheben sind dabei die Ansätze von Haralick und Shapiro und von Mumford und Shah, wobei letztere ihre Definition des Begriffes Segmentierung zur Grundlage ihrer Theorie der Segmentierung mit Methoden der Variationsrechnung machen.

2.2.1 Definitionen von Segmentierungen

Viele Autoren in der Bildverarbeitung umgehen es, den Begriff „Segmentierung“ zu definieren. Denn schon Azriel Rosenfeld und Avinash Kak schreiben 1982:

It should be emphasized that there is no single standard approach to segmentation. ([RK82], S.56)

In der Tat kann man das, was dem menschlichen Auge und Gehirn so leicht zu fallen scheint, nicht so einfach in mathematischen Termen ausdrücken.

In [HS92] gehen Robert Haralick und Linda Shapiro dem Problem der Definition des Begriffes „Segmentierung“ elegant aus dem Weg, indem sie beschreiben, welche Bedingungen die meisten Algorithmen, die sich „Segmentierungsalgorithmen“ nennen, erfüllen („*Instead, general segmentation procedures tend to obey the following rules*“ [HS92], S. 509). Als diese Bedingungen listen sie die folgenden auf:

- i. Segmente einer Segmentierung sind uniform oder homogen bezüglich einer Charakteristik wie zum Beispiel dem Grau- oder Farbwert oder der Textur.*
- ii. Adjazente Segmente haben signifikant unterschiedliche Werte in der Charakteristik, in der sie uniform sind.*
- iii. Das Innere der Segmente ist einfach und hat, wenn überhaupt, nur wenige kleine „Löcher“.*
- iv. Die Ränder von Segmenten sind ebenfalls einfach und nicht zerfranst.*

(2.1)

Es herrscht Einigkeit darüber, daß ein Verfahren, das ein Bild „segmentiert“, dieses in disjunkte Bereiche, deren Vereinigung wieder das Bild ergibt, aufteilt. Aber welche Eigenschaften diese Bereiche haben sollen, läßt sich nicht so leicht festlegen. Der Begriff „homogen“ in Punkt *i*. läßt einigen Interpretationsspielraum zu. Ebenso offen bleibt die Frage, in wie *viele* Bereiche ein Bild aufgeteilt werden soll.

In einigen wissenschaftlichen Veröffentlichungen findet man folgende Definition für ein „Segment“:

Ein Segment ist eine Menge von miteinander verbundenen Pixeln eines Bildes, in der alle Pixel zueinander ähnlich sein müssen (vgl. etwa [Wol00], S. 66).

Eine solche Definition wird allerdings Bildflächen nicht gerecht, die, weil sie vielleicht beim Aufnahmeprozeß schräg beleuchtet wurden, einen Farbverlauf haben, der sich kontinuierlich in einer Richtung ändert. Deshalb fordern David Mumford und Jayant Shah von einer Segmentierung eines Bildes f in n disjunkte Bereiche R_1, \dots, R_n ,

daß sich das Bild f innerhalb eines Bereichs $R_i, (i \in \{1, \dots, n\})$ nur sanft („smoothly“) verändert, an den Grenzen der Bereiche aber sprunghaft. (vgl. [MS89], S. 578).

(2.2)

Diese Definition – man vergleiche sie mit den Punkten *i.* und *ii.* der Bedingungen von Haralick und Shapiro – führt die Autoren dann unweigerlich zu den nach ihnen benannten Verfahren zur Bildsegmentierung, bei denen Bilder durch stückweise stetige Funktionen, die gewissen Glattheitsbedingungen gehorchen, approximiert werden. Die Bereiche, auf denen die Approximation stetig ist, bilden dann die Segmente. Die Segmentgrenzen sind durch die Unstetigkeitsstellen charakterisiert. Auf diese Verfahren wird in Abschnitt 2.4 eingegangen. Ein Nachteil dieser Definition ist das Fehlen jeglicher Charakterisierung der Form der Segmente. Man findet auch in dieser Definition keinen direkten Hinweis auf die Bestimmung der Anzahl der Segmente. Sie ist – abhängig von einer geeigneten Parametrisierung – gegeben durch die Anzahl der stetigen Teilstücke der Approximation.

Einen weiteren Nachteil der Definition von Mumford und Shah zeigt Bild 2.1¹: Der Hintergrund dieses Bildes enthält Pixel, deren Farbwerte denen der Hand näher sind, als denen der dunkleren Pixel des Hintergrunds. Außerdem sind die „Sprünge“ zwischen den Farb-



Abbildung 2.1: Bild mit „unähnlichen“ und „unglatten“ Segmenten.

werten des Hintergrundes gewaltig. Trotzdem unterscheidet man intuitiv zwischen den drei

¹Das Bild wurde http://www.caip.rutgers.edu/~comanici/segm_images.html entnommen.

Bereichen „Hand“, „Ring“ und „Hintergrund“. In der Praxis der Variationsrechnung in der Bildverarbeitung kann man dieses Problem jedoch umgehen, indem man ein Featurebild, statt eines Grauwert- bzw. Farbbilds, segmentiert (vgl. etwa [KLM94]). Das heißt, es wird der Segmentierung ein Schritt – nämlich der der Featureextraktion – vorgeschaltet. Dieser Schritt wird jedoch nicht in die Definition des Begriffes „Segmentierung“ aufgenommen.

2.2.2 Eigene Annäherung an den Begriff „Segmentierung“

Die eigene Annäherung an den Begriff „Segmentierung“ folgt zwei Maximen:

- Eine Segmentierung ist immer problemabhängig. Selbst wenn man ein Bild einfach „nur“ in mehrere Segmente unterteilen will, muß man sich schon Gedanken über das Aussehen dieser Segmente und die Sensoreigenschaften des Bildaufnahmeprinzips gemacht haben.
- Eine Segmentierung ist ein Approximationsproblem. Zahlreiche Filme glänzen heutzutage mit computergenerierten Bildern, die für das menschliche Auge kaum vom Original zu unterscheiden sind. Daher liegt es nahe, Bilder zu segmentieren, indem sie durch künstlich generierte Bilder approximiert werden. Die Segmente zeichnen sich dabei durch eine einheitliche Repräsentation der Approximation aus. Insbesondere ist daher das Segmentierungsproblem mathematisch formulierbar.



Abbildung 2.2: Bild einer hellen Granittranche vor gelbem Hintergrund.

Betrachtet man Bild 2.2 mit der Aufgabe, das Bild in die Bereiche „Granitplatte“ und „Hintergrund“ zu segmentieren, so erhält man ein Problem, das sich klar als mathematisches Approximationsproblem formulieren läßt:

Sei $f : M \rightarrow \{0, \dots, G\}^3$ ein Farbwertbild. Gesucht ist die Zerlegung des Definitionsbereichs M des Bildes in zwei disjunkte, einfach zusammenhängende² Mengen A, B , so daß

²Eine Teilmenge des \mathbb{R}^2 heißt *einfach zusammenhängend*, wenn ihr Rand zusammenhängend ist (vgl. [BS91], S. 244). Vergleiche auch Abschnitt 4.6.8

der Abstand $\text{dist}(f, g\mathbf{1}_A + h\mathbf{1}_B)$ minimal wird, wobei g und h Funktionen eines Funktionenraums sind, die f auf A bzw. B approximieren.

Dabei spielt der Funktionenraum in der Modellierung eine große Rolle. Bei der oben gestellten Aufgabe wird man z. B. g und h aus dem Raum der konstanten Funktionen wählen.

Im Beispiel der Abbildung 2.2 war schon vor dem Segmentierungsprozeß bekannt, wieviele Bereiche zu segmentieren sind. Wie kann man den Segmentierungsprozeß aber definieren, wenn man vorher nicht weiß, wieviele Objekte sich im Bild befinden? Bild 2.3 zeigt mehrere Stücke von Granittranchen vor gelbem Hintergrund. Die Anzahl der zu segmen-

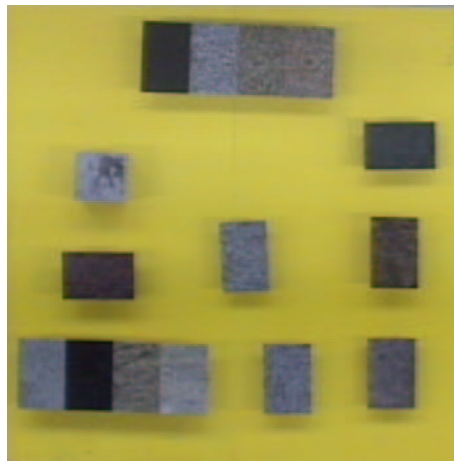


Abbildung 2.3: Bild von mehreren Granittranchenstücken vor gelbem Hintergrund.

tierenden Bereiche ist hier ein weiterer Parameter, den es zu ermitteln gilt. Dabei wird die Ermittlung dieses Parameters von dem Bestreben bestimmt, ihn möglichst klein zu halten. Auf der anderen Seite wird die Minimierung der zu segmentierenden Bereiche den Fehler der Approximation erhöhen. Bildet jedes einzelne Pixel ein Segment, so ist zwar der Approximationsfehler gleich 0, aber das Ergebnis wenig aussagekräftig. Das heißt, die Minimierung des Approximationsfehlers läuft der Minimierung der Anzahl der Bereiche entgegen. Man befindet sich also im Kontext einer Mehrzieloptimierung: Approximationsfehler versus Anzahl der zu segmentierenden Bereiche.

Dabei spielt unter anderem die Sensorcharakteristik eine entscheidende Rolle: Hat man zum Beispiel einen idealen Sensor, d. h. das Signal ist gänzlich unverrauscht, so muß man, wenn man von einfarbigen, ideal beleuchteten Objekten ausgeht, für jeden unterschiedlichen Sensorwert einen unterschiedlichen Bereich segmentieren. Ist das Signal stark verrauscht, so wird man großzügiger mit der Vereinigung von unterschiedlichen Signalen umgehen.

Fazit

Da, wie schon eingangs erwähnt, eine Segmentierung immer problemabhängig ist, muß für jedes praktische Problem eine eigene Definition formuliert werden. Daher sollte ein

gutes Segmentierungsverfahren immer eine Auswahl von Möglichkeiten und Werkzeugen bereitstellen, um das jeweilige Segmentierungsproblem formulieren und lösen zu können. Trotzdem gibt es einige Bedingungen, denen alle Segmentierungsverfahren unterworfen sein sollten. Für ein Bild $f : M \rightarrow \mathbb{R}^d$ sollte ein Segmentierungsverfahren den folgenden Bedingungen genügen:

Definition 2.2.1 (Notwendige Bedingungen einer Segmentierung)

Sind für ein $n \in \mathbb{N}$ die Mengen $A_1, \dots, A_n \subset M$ und die Funktionen g_1, \dots, g_n das Ergebnis eines Segmentierungsverfahrens, so sollte gelten:

- i. $A_1 \uplus \dots \uplus A_n = M$,
- ii. Die approximierenden Funktionen g_1, \dots, g_n sind aus einem Funktionenraum V und $\text{dist}(f, \sum_{i=1}^n g_i \mathbf{1}_{A_i})$ ist ein (lokales) Minimum unter allen Ergebnissen mit gleichem n ,
- iii. die Segmente A_1, \dots, A_n gehorchen gewissen Bedingungen an ihr Aussehen, und
- iv. n ist das Pareto-Optimum einer Mehrzieloptimierung.

Insbesondere sollte das Segmentierungsverfahren Mittel, d. h. Funktionen bereitstellen, die das Minimierungsproblem in ii. im Sinne des jeweiligen Segmentierungsprozesses gestalten.

Dabei sollte die Distanz dist in Punkt ii. der Definition 2.2.1 eine geeignete Metrik für den Abstand von Originalbild und Approximation sein. Das Aussehen der Segmente in Punkt iii. kann z. B. die Forderung nach Zusammenhang oder Konvexität für die Segmente bedeuten.

Die Forderung in Punkt iv. nach Pareto-Optimalität ist wohl der größte Unterschied zu anderen Ansätzen der Definition des Begriffes „Segmentierung“. Eine Situation einer Mehrzieloptimierung heißt *pareto-optimal*, wenn sich das Optimierungsproblem in keiner Komponente verbessern läßt, ohne es in einer anderen zu verschlechtern (vgl. auch Abschnitt 5.2).

Wie man deutlich sieht, steht bei der Definition 2.2.1 die Zahl n der Segmente im Vordergrund. Sie ergibt sich nicht, wie bei den Definitionen von Haralick und Shapiro bzw. Mumford und Shah indirekt aus den Eigenschaften der Segmente bzw. Segmentgrenzen, sondern ist ein Ziel einer Mehrzieloptimierung.

An diesen Kriterien sollte ein Segmentierungsverfahren gemessen werden. Ein nicht zu vernachlässigender Faktor bei der Beurteilung von Segmentierungsverfahren sollte auch die algorithmische Umsetzbarkeit bzw. die Laufzeit der jeweiligen Algorithmen auf Rechnersystemen sein.

2.3 Gängige Verfahren zur Segmentierung von Bildern

In diesem Unterkapitel sollen die gängigsten Methoden außerhalb des Ansatzes der Variationsrechnung zur Segmentierung von Bildern vorgestellt werden. Natürlich kann eine solche Übersicht nur einen Teil sowohl der unzähligen Verfahren und Algorithmen als auch der Theorien zur Segmentierung von Bildern wiedergeben. Daher werden nur solche Verfahren vorgestellt, die sich in der Literatur durchgesetzt haben. Besondere Aufmerksamkeit wird dabei dem Schwellwertverfahren von Nobuyuki Otsu in Abschnitt 2.3.1.1 gewidmet werden, da sich in Abschnitt 4.5 herausstellen wird, daß es sich auch als Mehrniveaupassungsproblem modellieren läßt.

Bemerkungen zur Segmentierung von Farbbildern und zur allgemeinen Bewertung von Segmentierungsverfahren schließen dieses Unterkapitel ab.

2.3.1 Schwellwertverfahren

Sei $\{0, \dots, G\}$ der Wertebereich eines diskreten Grauwertpixelbildes g . Bei einem Schwellwertverfahren werden zunächst algorithmisch Werte $\{t_2, \dots, t_k\} \subset \{0, \dots, G\}$, sogenannte Schwellwerte, bestimmt. Mit diesen Schwellwerten wird dann die Segmentierung von g berechnet, indem man alle Pixel, die zwischen zwei aufeinanderfolgenden Schwellwerten liegen, zusammenfaßt. Definiert man $S_i := \{p \in \mathbb{P}(M) : t_i \leq g(p) < t_{i+1}\}$ für $i = 1, \dots, k$, wobei $t_0 := 0$ und $t_{k+1} := G + 1$ definiert sei, so bildet S_1, \dots, S_k eine Segmentierung des Bildes g bei denen alle Pixel eines Segments einen ähnlichen Grauwert haben. Der Nachteil eines solches Vorgehens ist, daß man auf die Form der Segmente keinen Einfluß hat. Es befindet sich damit im Widerspruch zu Punkt iii. der Definition 2.2.1 der notwendigen Bedingungen an einer Segmentierung, bzw. zu Punkt iii. der Bedingungen von Haralick und Shapiro (vgl. (2.1), S. 14). Die Segmente S_1, \dots, S_k sind im allgemeinen nicht einmal zusammenhängend.³

Die zentrale Aufgabe bei Schwellwertverfahren ist es, ein geeignetes Verfahren für die automatische Bestimmung der Schwellwerte anzugeben. Das wichtigste Hilfsmittel dazu ist in den meisten Anwendungen das Histogramm des Bildes. Wenn sich zwei Objekte (meist ist eines davon der Hintergrund) im Bild klar anhand ihres Grauwertes unterscheiden, so werden diese Grauwerte im Histogramm des Bildes lokale Maxima sein. Genauer gesagt wird, da die Grauwerte in einem realen Bild immer auch innerhalb des Objektes ein wenig variieren, eine Art „Hügel“ im Graphen des Histogramms rund um den mittleren Grauwert des Objektes zu sehen sein. Einige Autoren (vgl. etwa [GW92], S. 447ff) gehen davon aus, daß dieser „Hügel“ im Idealfall einer Gaußschen Glockenkurve entspricht (vgl. z. B. Abbildung 2.4).

Der große Vorteil eines histogrammbasierten Schwellwertverfahrens ist die Geschwindigkeit, da die Algorithmen nur auf dem Definitionsbereich des Histogramms, statt auf dem des Bildes, arbeiten.

³Diesen Nachteil kann man umgehen, indem man, wie in Abschnitt 4.6, die Pixelzusammenhangskomponenten (vgl. Definition 1.3.10) der Mengen S_1, \dots, S_k als Segmente definiert.

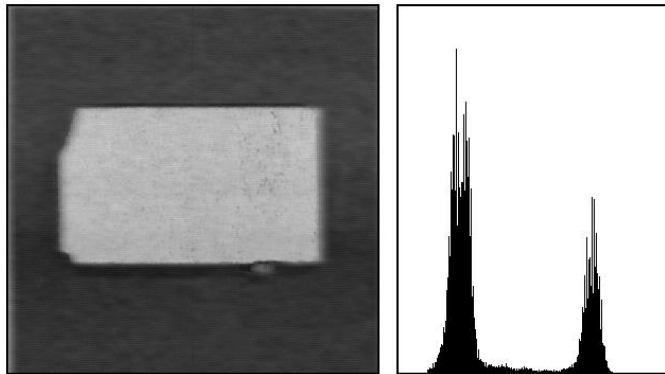


Abbildung 2.4: Beispiel eines Histogramms mit einem Definitionsbereich von 256 Grauwerten. Bei dem Bild handelt es sich um den Blaukanal der Aufnahme von Bild 2.2.

2.3.1.1 Varianzanalyse, Otsu 1979

Obwohl aus dem Jahre 1979, und damit vergleichsweise alt, findet man das Verfahren von Nobuyuki Otsu ([Ots79]) als Referenz in zahlreichen Veröffentlichungen über histogrammbasierte Schwellwertverfahren.

Otsu nimmt die Einträge des Histogramms als Wahrscheinlichkeiten einer diskreten Zufallsvariablen und bestimmt denjenigen Grauwert, für den sich, nach einer Varianzanalyse, das Histogramm am besten in zwei zusammenhängende Teile separieren läßt.

Sei $\{0, \dots, G\}$ die Menge der Grauwerte eines diskreten Grauwertpixelbildes g und h_g das Histogramm des Bildes (vgl. Definition 1.3.2 und 1.3.4). Dann kann man für ein beliebiges Pixel p des Bildes aus $h_g(i)$ die Wahrscheinlichkeit schätzen, daß $g(p)$ den Wert i hat. Diese Überlegung rechtfertigt das Vorgehen, das Histogramm h_g als Grundlage eines Wahrscheinlichkeitsmaßes zu sehen.

Sei $X : \Omega \rightarrow \{0, \dots, G\}$ eine Zufallsvariable, deren Wahrscheinlichkeit durch das Histogramm h_g bestimmt ist. Definiert man $N := \sum_{i=0}^G h_g(i)$, so sei die Wahrscheinlichkeit $P[X = i]$, die Wahrscheinlichkeit, daß ein Pixel den Grauwert i trägt, $p_i := \frac{h_g(i)}{N}$. Es gilt dann $\sum_{i=0}^G p_i = 1$ und alle p_i sind größer oder gleich 0, wobei im folgenden alle p_i mit $0 \leq i \leq G$ größer als 0 angenommen werden, um Fallunterscheidungen zu vermeiden.

Ziel ist es nun, einen Grauwert $k^* \in \{0, \dots, G-1\}$ zu bestimmen, so daß sich für $k = k^*$ die beiden Klassen $C_0(k) := \{0, \dots, k\}$ und $C_1(k) := \{k+1, \dots, G\}$ am besten trennen lassen. Dieser Grauwert k^* ist dann ein geeigneter Schwellwert, um die beiden Klassen zu segmentieren.

Die Wahrscheinlichkeit der Klasse $C_0(k)$ ist $\omega_0(k) := P[X \in C_0(k)] = \sum_{i=0}^k p_i$. Analog ist $\omega_1(k) := P[X \in C_1(k)] = \sum_{i=k+1}^G p_i = 1 - \omega_0(k)$ die Wahrscheinlichkeit der Klasse $C_1(k)$. Definiere $\mu(k) := \sum_{i=0}^k i p_i$. Dann ist $E[X] = \mu(G)$ der Erwartungswert von X .

Der bedingte Erwartungswert der Klasse $C_0(k)$ ist

$$\mu_0(k) := E[X|X \in C_0(k)] = \sum_{i=0}^k i \frac{p_i}{\omega_0(k)} = \frac{\mu(k)}{\omega_0(k)}$$

und analog sei

$$\mu_1(k) := E[X|X \in C_1(k)] = \frac{\mu(G) - \mu(k)}{\omega_1(k)} = \frac{\mu(G) - \mu(k)}{1 - \omega_0(k)}$$

Die bedingte Varianz der Klassen C_0 bzw. C_1 ist

$$\sigma_0(k)^2 = \text{VAR}[X|X \in C_0(k)] = E[(X - \mu_0(k))^2|X \in C_0(k)] = \sum_{i=0}^k (i - \mu_0(k))^2 \frac{p_i}{\omega_0(k)} \quad (2.3)$$

$$\text{bzw. } \sigma_1(k)^2 = \sum_{i=k+1}^G (i - \mu_1(k))^2 \frac{p_i}{\omega_1(k)}.$$

Die Gesamtvarianz der Zufallsvariable X ist $\sigma^2 := \sum_{i=0}^G (i - \mu(G))^2 p_i$.

Die Varianzanalyse unterscheidet nun zwischen der Varianz innerhalb der Klassen (im folgenden $\sigma_W(k)^2$) und der Varianz zwischen den Klassen (im folgenden $\sigma_B(k)^2$).⁴ Sie sind wie folgt definiert:

$$\sigma_W(k)^2 := \omega_0(k)\sigma_0(k)^2 + \omega_1(k)\sigma_1(k)^2$$

und

$$\sigma_B(k)^2 := \omega_0(k)(\mu_0(k) - \mu(G))^2 + \omega_1(k)(\mu_1(k) - \mu(G))^2$$

Es gilt

$$\sigma_W(k)^2 + \sigma_B(k)^2 = \sigma^2 \quad (2.4)$$

Die Summe der Varianz innerhalb der Klassen und der Varianz zwischen den Klassen entspricht also der Gesamtvarianz.

Zwei Klassen sind „gut“ zu trennen, wenn die Varianz innerhalb der Klassen gering und die Varianz zwischen den Klassen hoch ist. Otsu betrachtet in seiner Arbeit nun den Wert $\chi(k) := \frac{\sigma_B(k)^2}{\sigma_W(k)^2}$. Dieser sollte maximal sein unter allen möglichen Schwellwerten k , d. h. $\sigma_B(k)^2$ soll ein hoher Wert und $\sigma_W(k)^2$ gleichzeitig ein niedriger Wert sein. Ist $\sigma_W(k)^2 = 0$, so sei $\chi(k) := \infty$ definiert. Dieser Fall ist allerdings ein Idealfall, der in den meisten Bildern nicht auftritt.

Für eine effektive Bestimmung des Wertes k , für den $\chi(k)$ maximal wird, wird nicht $\chi(k)$, sondern $\frac{\chi(k)}{\chi(k)+1}$ maximiert. Da für alle k gilt $\chi(k) \geq 0$, ist dies zulässig. Nun ist aber, wie man durch leichte Umformungen mit Hilfe von Gleichung (2.4) zeigen kann, $\frac{\chi(k)}{\chi(k)+1} = \frac{\sigma_B(k)^2}{\sigma^2}$. Daher genügt es, nur $\sigma_B(k)^2$, die Varianz zwischen den Klassen, zu maximieren, um gleichzeitig auch die Varianz innerhalb der Klassen zu minimieren. Das heißt als Schwellwert liefert dieses Verfahren einen Wert k^* mit

$$\sigma_B(k^*)^2 = \max_{k \in \{0, \dots, G-1\}} \sigma_B(k)^2$$

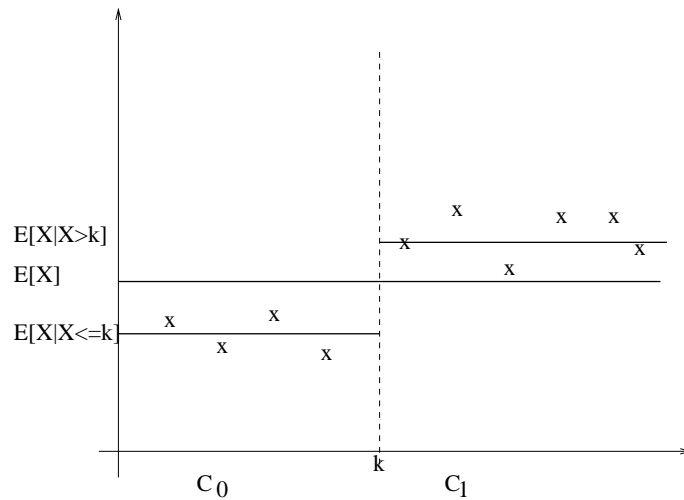


Abbildung 2.5: Veranschaulichung: Wähle k so, daß die Abweichung innerhalb der Klassen gering und zwischen den Klassen groß ist.

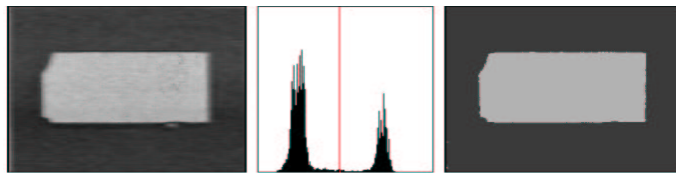


Abbildung 2.6: Beispiel der Schwellwertbestimmung nach dem Verfahren von Otsu (Originalbild 2.4, Histogramm mit eingetragem Schwellwert, Ergebnis nach Anwendung des Schwellwertes).

Bemerkung 2.3.1

Es ist

$$\begin{aligned}
 \sigma_B(k)^2 &= \omega_0(k)(\mu_0(k) - \mu(G))^2 + \omega_1(k)(\mu_1(k) - \mu(G))^2 \\
 &= \omega_0(k)\mu_0(k)^2 - 2\omega_0(k)\mu_0(k)\mu(G) + \omega_0(k)\mu(G)^2 \\
 &\quad + \omega_1(k)\mu_1(k)^2 - 2\omega_1(k)\mu_1(k)\mu(G) + \omega_1(k)\mu(G)^2 \\
 &= \omega_0(k)\mu_0(k)^2 + \omega_1(k)\mu_1(k)^2 \\
 &\quad - 2\mu(G) \underbrace{(\omega_0(k)\mu_0(k) + \omega_1(k)\mu_1(k))}_{=\mu(G)} + \mu(G)^2 \underbrace{(\omega_0(k) + \omega_1(k))}_{=1} \\
 &= \omega_0(k)\mu_0(k)^2 + \omega_1(k)\mu_1(k)^2 - \mu(G)^2
 \end{aligned}$$

Da $\mu(G)^2$ konstant ist, genügt es also, um $\sigma_B(k)^2$ zu maximieren, lediglich den Term

⁴Die Wahl der Indizes „W“ bzw. „B“ folgt dem Originalartikel, d. h. sie sind dem Englischen entlehnt: „W“ steht für „within“ und „B“ für „between“.



Abbildung 2.7: Beispiel der Schwellwertbestimmung nach dem Verfahren von Otsu. Zu sehen ist von links nach rechts das Originalbild, das Histogramm mit eingetragenen Schwellwert und das Ergebnis nach der Anwendung des Schwellwertes zur Segmentierung. Das Originalbild ist die Abbildung einer Granittranche und ihres Schattens vor einem weißen Hintergrund. Der weiße Hintergrund hat die höchste Helligkeit, die ein Pixel bei dieser Aufnahme annehmen kann. Daher entsteht, aufgrund der großen Fläche des Hintergrunds, ein unverhältnismäßig hoher Histogrammwert für G . Wie man sieht, wird der Schatten, der heller ist als das Objekt, noch zum Hintergrund gerechnet.

$\omega_0(k)\mu_0(k)^2 + \omega_1(k)\mu_1(k)^2$ zu maximieren. Es ist

$$\omega_0(k)\mu_0(k)^2 + \omega_1(k)\mu_1(k)^2 = \frac{(\sum_{i=0}^k ip_i)^2}{\sum_{i=0}^k p_i} + \frac{(\sum_{i=k+1}^G ip_i)^2}{\sum_{i=k+1}^G p_i} \quad (2.5)$$

Auf diese Gleichheit wird an späterer Stelle, nämlich in Abschnitt 4.5, noch einmal verwiesen werden.

2.3.1.2 Erweiterung der Varianzanalyse auf $n > 1$ Schwellwerte

Mit Hilfe des oben beschriebenen Ansatzes läßt sich leicht auch mehr als ein Schwellwert bestimmen. In dem Originalartikel [Ots79] schlägt Otsu vor, die Varianzanalyse über $n+1$ Klassen, statt über 2 Klassen, zu machen. Dann hängt allerdings σ_B^2 von n statt nur von einem Parameter ab. Man sucht dann Schwellwerte k_1^*, \dots, k_n^* mit

$$\sigma_B(k_1^*, \dots, k_n^*)^2 = \max_{0 \leq k_1 < \dots < k_n < G} \sigma_B(k_1, \dots, k_n)^2$$

Auf die Berechnung des globalen Maximums von $\sigma_B(k_1^*, \dots, k_n^*)^2$ sollte jedoch verzichtet werden, da $\binom{G+1}{n}$ Fälle zu unterscheiden sind. Bei $G = 255$ wären daher schon für nur 3 Schwellwerte 2731135 Fälle zu untersuchen. Daher sollte man stattdessen das Histogramm, u. U. mit Hilfe des Verfahrens von Otsu, in sinnvolle Teile unterteilen und auf diese den Algorithmus zur Bestimmung eines Schwellwertes anwenden. Genauso könnte man die einzelnen Parameter sukzessive optimieren, bis sich kein Parameter mehr einzeln verbessern läßt. Dieses Verfahren sichert zwar kein globales Maximum, ist jedoch in der Laufzeit viel effektiver. Ein Algorithmus der genau dies leistet, ist der Algorithmus 11 (**Histogrammanalyse**), der in Abschnitt 4.5 als Anwendung der Mehrniveaupassung präsentiert wird.

2.3.1.3 Weitere Schwellwertverfahren

Schon 1972 versuchte man, das Histogramm durch eine Summe von Normalverteilungen zu approximieren ([CK72]). Allerdings zeigen die Beispielhistogramme der Abbildungen 2.6, 2.7 oder auch 4.18 in Abschnitt 4.5 die Grenzen eines solchen Ansatzes, insbesondere bei ungeglätteten Histogrammen: Zum einen fallen viele nicht-signifikante lokale Optima der nichtlinearen Approximation an, zum anderen verteilen sich nicht alle Grauwerte von Objekten gemäß einer Gaußverteilung auf das Histogramm. Es kommt oft vor, daß die ansteigende und absteigende Flanke einer Spitze im Histogramm unterschiedlich schnell an- bzw. absteigt. Außerdem stellt sich die Frage, wie bzw. an welcher Stelle man zwei Gaußkurven mit unterschiedlicher Varianz, die sich überlagern, trennen soll.

Die meisten histogrammbasierten Schwellwertbestimmungsverfahren versuchen „Täler“ zwischen den markantesten „Hügeln“ im Histogramm zu detektieren. In [Mat96] wird ein Verfahren vorgeschlagen, das der Autor „Graph-theoretical clustering“ nennt. Bei diesem Verfahren wird ein Histogramm mit Zeigern benutzt. Zusätzlich zum Wert des Histogramms an der Stelle i , befindet sich an dieser Stelle noch ein Zeiger, der auf eine Stelle zwischen $i - \delta$ und $i + \delta$ (für ein $\delta \in \mathbb{N}$) des Histogramms zeigt. Die Grundidee bei diesem Verfahren ist, daß jeder Stelle i im Histogramm auf den Nachbarn zeigt, dessen Histogrammwert in einer vorgegebenen Nachbarschaft ($i - \delta$ bis $i + \delta$) am größten ist. Lokale Maxima sind dann Werte, die auf sich selbst zeigen. Alle Grauwerte des Bildes werden nun in der Segmentierung mit dem lokalen Maximum, auf das sie direkt oder indirekt (nach Bildung der transitiven Hülle der Zeiger-Relation) zeigen, identifiziert. Man erhält so ein segmentiertes Bild.

Auch die Verwendung neuronaler Netze zur Detektion der oben genannten „Täler“ im Histogramm findet in der Literatur Beachtung (vgl. z. B. [PSA00]).

Einen interessanten Ansatz bietet auch die Arbeit [KK98]. Dort wird eine Reihe von Funktionsquantisierungen des Histogramms mit abnehmender Anzahl von Quantisierungsstufen genutzt, um die signifikanten „Täler“ im Histogramm zu bestimmen. Im Abschnitt 2.5.3.2 wird auf dieses Verfahren noch näher eingegangen.

Weitere interessante Schwellwertverfahren findet man in Kapitel 7.3 des Buches [GW92] oder auf den Seiten 520 - 524 inbzw. [HS92]. Eine sehr umfassende Übersicht von Binarisierungsalgorithmen (das heißt Algorithmen, die einen Schwellwert berechnen und mit diesem ein gegebenes Grauwertbild in ein Binärbild umwandeln) findet man in [TJ95]. Dieser Artikel beschränkt sich allerdings nicht nur auf reine Schwellwertverfahren.

2.3.2 Regionenwachstumsverfahren, „Region Splitting and Merging“

Unter dem Begriff „Regionenwachstumsverfahren“ werden in der Regel Verfahren zusammengefaßt, die Pixel oder kleinere Pixelregionen, also zusammenhängende Mengen von Pixeln, zu größeren pixelzusammenhängenden Regionen wachsen lassen. Das grundsätzliche Vorgehen ist ein in der Bildverarbeitung immer noch aktuelles Verfahren, das schon über 30 Jahre alt ist. In [BF70] wurde es vermutlich das erste Mal erwähnt.

Ein gängiger Algorithmus ist folgender:

- Wähle eine initiale Menge von Pixeln, die einen gewissen Abstand voneinander haben (im allgemeinen wird diese Menge als „seed points“ bezeichnet). Jedes solche Startpixel bildet eine eigene Startregion.
- Lasse alle Regionen wachsen, indem ihnen nach gewissen Ähnlichkeitskriterien (zum Beispiel der Grau- oder Farbwert, die Textur oder ähnliches) benachbarte Pixel hinzugefügt werden.

Ein solches Verfahren hängt sehr stark von der Auswahl der Startregionen ab. In [AB94] und in [MJ97] finden sich Hinweise für eine gute Auswahl von Startpixeln. Eine Einführung in Regionenwachstumsverfahren findet man in Kapitel 7.4 in [GW92] oder auch in Kapitel 10.3 in [HS92].

Eine Alternative zu den doch recht einfachen Regionenwachstumsverfahren sind die sogenannten „Split and Merge“ Verfahren. Dabei wird ein Bild solange in disjunkte Segmente weiter aufgeteilt, bis diese innerhalb ihrer Fläche ein vorher festgelegtes Ähnlichkeitskriterium erfüllen. Diese Segmente werden dann anschließend nach Ähnlichkeitskriterien wieder zusammengefügt: Sind zwei Segmente einander ähnlich (zum Beispiel im durchschnittlichen Grau- bzw. Farbwert oder in der Textur), so werden sie zusammengefügt. Dies wird solange gemacht, bis vorher festgelegte Nebenbedingungen, wie z. B. die Anzahl der resultierenden Segmente oder der mittlere Fehler pro Segment, erfüllt sind. Ein sehr bekanntes Verfahren in diesem Bereich ist die Quadtree-Zerlegung, die auch in der „Image Processing Toolbox“ der mathematisch-technischen Entwicklungsplattform „Matlab“ zu finden ist.⁵ Sie geht zurück auf Horowitz und Pawlidis [HP76]. Eine nähere Beschreibung dieses Verfahrens findet man unter anderem in [GW92] und in [HS92].

2.3.3 Wasserscheidentransformationen

Eine wichtige Gruppe der Segmentierungsverfahren sind die sogenannten Wasserscheidentransformationen, die gerade in den letzten Jahren besondere Aufmerksamkeit erhielten. Die diesen Verfahren zugrunde liegende Idee läßt sich am besten verstehen, wenn man die Grauwerte eines diskreten Grauwertpixelbildes als Höhenangaben eines topographischen Reliefs betrachtet. Trifft ein „Regentropfen“ auf ein solches Relief, so wird er in die Richtung des nächsten lokalen Minimums fließen. Die Einflußbereiche lokaler Minima nennt man „Wasserbecken“. Unter einer „Wasserscheide“ versteht man dann Bereiche, bei denen man nicht entscheiden kann, in welche Richtung der Regentropfen abfließen wird, da mehrere Abflußrichtungen möglich sind. Die Wasserscheiden sind darüber hinaus Trennlinien zwischen den Wasserbecken. Nicht-signifikante lokale Minima werden eliminiert, indem man die zwischen ihnen liegende Wasserscheide durch „Überfluten“ entfernt. Ein Vorteil der Wasserscheidentransformation ist die Eigenschaft, Details und Feinheiten zu erhalten. Eine genaue Einführung mit zahlreichen Beispielen findet sich in [Wol00] ab Seite 74. In jüngerer Zeit gehen immer mehr Autoren dazu über, die Wasserscheidentransformation in einen morphologischen Zusammenhang einzuführen (vgl. etwa [Soi98]).

⁵Vergleiche <http://www.mathworks.com>.

2.3.4 Weitere Methoden zur Bildsegmentierung

Die Liste der heuristischen Verfahren zur Segmentierung von Bildern ist außerordentlich groß. Eine Aufzählung zahlreicher gängiger Verfahren, wie z. B. kantenbasierte Verfahren, findet sich unter anderem in [Wol00], [GW92] und [HS92]. Gerade im Bereich der Regionenwachstumsverfahren aus Abschnitt 2.3.2 sind bei der Auswahl der Ähnlichkeitskriterien der Phantasie keine Grenzen gesetzt. In der Literatur findet man Ansätze von Fuzzy-Methoden (vgl. z. B. [MIN⁺00]) über neuronale Netze (vgl. z. B. [WZX92]) bis hin zu vollkommen heuristischen Methoden.

Ein ebenso großer Bereich sind die statistischen Methoden zur Bildsegmentierung, die vor allem aus der Spracherkennung, d. h. aus der Verarbeitung eindimensionaler Signale, ihren Weg in die Bildsegmentierung – und damit in die Verarbeitung zweidimensionaler Signale – gefunden haben. Besonders Markov-Felder zur Segmentierung von Texturen sind sehr weit verbreitet. Eine Einführung zu diesem Thema findet man in [Gim99].

Eine umfassende Übersicht über (Farb-)Bildsegmentierung mit neuronalen Netzen, Fuzzy-Methoden und Markov-Ketten findet man auch in [PP93].

Bemerkung 2.3.2

Eine eingehende Betrachtung der zahlreichen heuristischen und statistischen Verfahren ist an dieser Stelle nicht angebracht. Denn bei den meisten heuristischen Ansätzen, zu denen letztlich auch die Ansätze über neuronale Netze zu zählen sind, ist mathematisch oft nicht klar, wonach sie eigentlich suchen bzw. welchen Fehler sie minimieren. Ein Vergleich solcher Verfahren ist also im allgemeinen recht schwierig, auch wenn es in der Literatur Ansätze gibt, dieses doch zu tun (vgl. Abschnitt 2.3.6). Trotzdem liefern viele dieser Verfahren gute Ergebnisse. Dabei wird der Begriff „gut“ allerdings mit dem Auge gemessen, er unterliegt also keiner absoluten Norm.

Bei den statistischen Methoden steht meist die Segmentierung von Bildmerkmalen („Features“) des Bildes im Vordergrund. Es werden dabei zunächst lokale Bildmerkmale gewonnen, deren Ähnlichkeit gemessen wird. Es wird also auf einem Metabild, nämlich einem Merkmalsbild, gerechnet. Auf diesem Metabild werden dann wieder herkömmliche Verfahren angewandt, um Pixel mit ähnlichem Merkmalsvektor zusammenzufassen. Deshalb liegen diese Verfahren in ihrem Approximationsverhalten bezüglich des Originalbildes außerhalb des Punktes ii. der Definition 2.2.1 und damit nicht im Fokus dieser Arbeit.

2.3.5 Bemerkungen zur Segmentierung von Farbbildern

Viele Verfahren der Grauwertbildsegmentierung lassen sich leicht auf Farbbilder übertragen, indem man Farbbilder als drei-kanalige Bilder betrachtet, die sich in drei Grauwertbilder aufspalten lassen (z. B. in ein Rot-, Grün- und Blaubild). Es lassen sich etwa Schwellwertverfahren auf den Bildern einzeln berechnen und die schwellwertsegmentierten Bilder wieder zusammenfügen (vgl. z. B. [OPR79] oder Algorithmus 13 (**Initiales Farbschwellwertverfahren**), S. 201). Auch Ähnlichkeitsmaße für Regionenwachstums- oder

„Split and Merge“-Strategien lassen sich übertragen, indem man statt einer Norm auf \mathbb{R} zu einer Norm auf \mathbb{R}^3 wechselt, mit der die Ähnlichkeit von Pixeln bzw. Regionen gemessen wird (vgl. z. B. [PB99]).

Für die Segmentierung von Farbbildern wurden allerdings auch eigene, nicht von den Grauwertmethoden übernommene Verfahren entwickelt, wie z. B. das sogenannte „Color Indexing“ über Farbhistogramme nach Swain und Ballard [SB91]. Dabei wird eine Funktion bin definiert, die jeder Farbe einen „Index“ zugeordnet. Ein 3D-Histogramm wird dabei in sogenannte „bins“ aufgeteilt, beispielsweise in $4 \times 4 \times 4$ „bins“, und in ein 1D-Histogramm überführt.

Sei $g : M \rightarrow \{0, \dots, 255\}^3$ ein diskretes Farbpixelbild. Dann ist die Abbildung bin definiert durch

$$bin : \{0, \dots, 255\}^3 \rightarrow \{0, \dots, 84\} \\ (r, g, b) \mapsto 16 \text{round}(\frac{r}{64}) + 4 \text{round}(\frac{g}{64}) + \text{round}(\frac{b}{64})$$

Das Farbhistogramm $F_g : \{0, \dots, 84\} \rightarrow \mathbb{N}_0$ von g ist dann definiert durch

$$F_g(i) := |\{p \mid bin(g(p)) = i, p \in \mathbb{P}(M)\}|$$

Auf diesem Histogramm kann man dann wieder Schwellwerte definieren, um Pixel mit ähnlichem Index zu Farbclustern zusammenzufassen (vgl. z. B. [BR99]), die das Bild segmentieren.

Ein weiterer Ansatz der Farbbildverarbeitung, den die Segmentierung von Grauwertbildern nur äußerst eingeschränkt bieten kann, ist die Segmentierung durch Farbklassen oder Farbnamen. Die Grundidee ist einfach: Man faßt einfach alle benachbarten Pixel mit der gleichen Farbe zu einem Segment zusammen. Dabei ist ein Farbmodell wie das HSI-Farbmodell eher geeignet als das RGB-Farbmodell, da hier Helligkeit und Farbton getrennt sind (vgl. Abschnitt 2.3.5). Bei der reinen Klassifikation von Farbpixeln wird für jedes Pixel ein Merkmalsvektor bestimmt, der auch aus mehr als nur der Farbe des Pixels (z. B. den Farben der Nachbarn) bestehen kann. Auf der Basis dieses Merkmalsvektors wird das Pixel dann einer Klasse zugeordnet (vgl. z. B. [OKS80] und [CM97]).

2.3.6 Bewertung von Segmentierungsverfahren

Eine Grundlage zur Bewertung von Segmentierungsverfahren bieten natürlich die Definitionsversuche des Abschnitts 2.2. Y. J. Zhang stellt in dem Artikel [Zha96] weitere gängige Bewertungsmethoden von Segmentierungsansätzen vor. Hier soll eine kleine Übersicht gegeben werden. Für eine genauere Einsicht sei auf den Originalartikel verwiesen.

Die beiden ersten Qualitätskriterien lassen sich direkt aus der Definition von Haralick und Shapiro (vgl. S. 14, (2.1)), bzw. von Mumford und Shah (vgl. S. 15, (2.2)) ableiten: Eine Segmentierung sollte den Definitionsbereich eines Bildes so zerlegen, daß sich das Bild innerhalb der Segmente nur sanft und zwischen den Segmenten stark verändert.

2.3.6.1 Bewertung der Uniformität der Regionen

Ein Bewertungskriterium ist die Uniformität innerhalb der Regionen. In [Zha96] wird zwar nur ein „gray-level uniformity measure“, abgekürzt **GU**, vorgestellt, aber der Übergang zu

Farbbildern durch den Normwechsel von der Betragsnorm zur euklidischen Norm ist kanonisch. In Anlehnung an den Originalartikel soll jedoch die Bezeichnung **GU** beibehalten werden.

Sei R_1, \dots, R_n eine disjunkte Zerlegung eines Bildes $f : M \rightarrow \mathbb{R}^d$ in n Regionen. Man betrachte den Fehler

$$\mathbf{GU} := \sum_{i=1}^n \int_{R_i} \|f - m_{R_i}\|_2^2 d\lambda$$

Dabei sei m_{R_i} der durchschnittliche Farb- oder Grauwert des Bildes f auf der Region R_i .

Dieses Qualitätsmaß mißt den Abstand der Segmente zum Mittel über dem Segment. Ein Verfahren, das genau diesen Fehler minimiert, ist die 2-Niveaupassung nach Donner [Don91] (vgl. 3.5.11 i)). Für die Approximation des gegebenen Bildes mit stückweise konstanten Funktionen bezüglich des euklidischen Abstands entspricht die Forderung nach Minimierung des Fehlers **GU** genau der Bedingung ii. in Definition 2.2.1. Allerdings ist diese Definition weder auf die Approximation mit stückweise konstanten Funktionen noch auf einen euklidischen Abstand festgelegt. Die Minimierung des Fehler **GU** hingegen führt z. B. bei einem affinen Grauwertverlauf zu keinem befriedigenden Ergebnis (vgl. Abbildung 3.4, S. 67).

Man kann innerhalb einer stochastischen Modellierung diesen Fehler auch über die Summe bedingter Varianzen beschreiben (vgl. z. B. [LN85]). So eine Modellierung würde es auch erlauben, einen Ansatz, der Texturen segmentiert, d. h. Regionen, deren Grau- bzw. Farbwert eben nicht durch einen konstanten Wert gegeben ist, zu beurteilen.

2.3.6.2 Bewertung des Kontrastes zwischen den Regionen

Schon Nobuyuki Otsu versucht in [Ots79], den Kontrast zwischen zwei Regionen zu maximieren (vgl. auch Abschnitt 2.3.1.1). Auch die Definition des Begriffes „Segmentierung“ durch Mumford und Shah (S. 15, (2.2)) veranlaßt, wie bereits eingangs erwähnt, den Kontrast zwischen zwei adjazenten Regionen zu maximieren.

In [Zha96] wird ein „gray-level contrast measure“ abgekürzt **GC** angegeben, das sich zwar wieder nur auf Grauwertbilder bezieht, aber ebenso wie das „gray-level uniformity measure“ kanonisch auf Farbbilder übertragen werden kann. Für eine Objekt/Hintergrund-Situation bezeichne f_o den durchschnittlichen Grau- bzw. Farbwert der Pixel des Objektes und f_b den des Hintergrundes. Dann sei der Wert des Kontrastes zwischen den beiden Regionen:

$$\mathbf{GC} := \frac{\|f_o - f_b\|_2}{\|f_o + f_b\|_2}$$

2.3.6.3 Andere Bewertungskriterien

- Bewertung der Form der Regionen

In [SSWC88] vergleichen Sahoo, Soltani, Wong und Chen verschiedene Schwellwertverfahren, indem sie die Uniformität der helleren Region bewerten.

Sei $g : \mathbb{P}(M) \rightarrow \{0, \dots, G\}$ ein diskretes Grauwertbild. Weiter sei $g_{N(x,y)}$ der mittlere Grauwert von g über alle Nachbarpixel⁶ von $[x, y]$, $C \in \mathbb{R}_+$ ein Normalisierungsfaktor und T ein Schwellwert, den ein Schwellwertverfahren als Segmentierung liefert. Dann schlagen Sahoo, Soltani, Wong und Chen in [SSWC88] folgenden Wert vor, den sie als „shape measure“, abgekürzt **SM**, bezeichnen.

$$\mathbf{SM} := \sum_{p \in \mathbb{P}(M)} U(g(p) - g_{N(p)}) \|\text{grad}h(p)\|_2^2 U(g(p) - T)$$

Dabei sei $U : \mathbb{R} \rightarrow \{0, 1\}$, $x \mapsto \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$ und $\text{grad}h(p)$ der Gradient des Bildes g im Pixel p . Der Gradient kann über eine kontinuierliche Approximation h des diskreten Bildes, zum Beispiel durch Polynome, in der Umgebung des Pixels p berechnet werden.

Der Wert **SM** ist umso kleiner, je uniformer die Region ist, über der der Wert ermittelt wurde.

- Empirische Diskrepanzmethoden

Die empirischen Diskrepanzmethoden zur Bewertung von Segmentierungsalgorithmen können helfen, etwas zu messen, was sich nicht in mathematischen Termen ausdrücken läßt. Sie messen einen Fehler, indem sie das vom Segmentierungsalgorithmus gelieferte Ergebnisbild mit einem Idealergebnis einer Segmentierung vergleichen. Dieses Idealbild kann dann auch mit dem Auge gemessen sein. Es muß also keiner mathematischen Optimalitätsbedingung genügen, sondern kann einer ästhetischen Optimalität Rechnung tragen.

- Bewertung über die Anzahl der Objekte im Bild:

Ist das Idealergebnis einer Segmentierung bekannt und liegt das Ergebnis eines Segmentierungsalgorithmus vor, so ist der Vergleich der Anzahl der segmentierten Objekte ein Kriterium zur Gütebestimmung. Man wird ein Verfahren, das weniger oder mehr Regionen segmentiert als die ideale Segmentierung, geringer bewerten als ein Verfahren, das genau diese Anzahl segmentiert. Man sollte bei einer Bewertung noch die Größe der nicht oder zuviel berechneten Segmente berücksichtigen.

- Bewertung der falsch segmentierten Pixel:

Gegeben sei ein Objekt O und der Hintergrund B . Für ein Pixel p seien zwei Zufallsvariablen $Y_{ideal}(p), Y_{seg}(p) : \Omega \rightarrow \{O, B\}$ gegeben. Dabei habe Y_{ideal} die Verteilung der Idealsegmentierung und Y_{seg} die Verteilung des algorithmisch segmentierten Bildes. In [LCP90] wird nun vorgeschlagen, die Wahrscheinlichkeit einer Fehlklassifikation („Probability of error“, abgekürzt PE) eines Pixels p zu betrachten. Es ist

$$\begin{aligned} PE(p) &:= P[Y_{seg}(p) = B | Y_{ideal}(p) = O] P[Y_{seg}(p) = O] \\ &\quad + P[Y_{seg}(p) = O | Y_{ideal}(p) = B] P[Y_{seg}(p) = B] \\ &\qquad\qquad\qquad Y_{ideal}\text{-fast sicher} \end{aligned}$$

⁶Zur eingehenderen Betrachtung des Begriffes „Nachbarpixel“ vgl. Definition 1.3.5.

Dabei bezeichnet $P[Y_{seg}(p) = B | Y_{ideal}(p) = O]$ die Wahrscheinlichkeit, daß ein Pixel des Objekts fälschlicherweise als Pixel des Hintergrunds klassifiziert wird. $P[Y_{seg}(p) = O | Y_{ideal}(p) = B]$ bezeichnet den umgekehrten Fall. Im Idealfall, also wenn die ideale Segmentierung und algorithmische Segmentierung übereinstimmen, gilt $PE(p) = 0$.

- Bewertung der Position der falsch segmentierten Pixel:

Auch die Positionen der falsch zugeordneten Pixel können bei der Bewertung von Segmentierungsalgorithmen eine Rolle spielen. So wird schon 1977 in [YMB77] vorgeschlagen, einen Fehler zu betrachten, der den Abstand fehlklassifizierter Pixel zu ihrem eigentlichen Segment mißt. Sei N die Anzahl der falsch zugeordneten Pixel und $d(i)$ der Abstand des i -ten falsch segmentierten Pixel zu seinem eigentlichen Segment. Dann sei

$$D = \sum_{i=1}^N d^2(i)$$

das Maß der räumlichen Fehlsegmentierung.

- Andere empirische Diskrepanzmethoden:

In [Zha96] werden noch eine Reihe weiterer Bewertungskriterien vorgestellt, die die Diskrepanz zu einer idealen Segmentierung als Grundlage ihrer Bewertung haben. So kann man z. B. Bildmerkmale, die sich aus den idealen und berechneten Segmenten ergeben, vergleichen oder die Fläche der Differenz zwischen diesen Segmenten messen.

2.4 Bildsegmentierung mit Methoden der Variationsrechnung

Der Begriff „Variation“ wurde 1755 von dem italienisch-französischen Mathematiker Joseph-Louis Lagrange in einem Brief an Leonard Euler eingeführt, in dem der erst 20jährige einige Verbesserungen eines Beweises vorschlägt. Die Variationsrechnung befaßt sich mit der Ermittlung von Extrema von Funktionalen, die auf einer Klasse von Funktionen definiert sind, und Werte in den reellen Zahlen haben. Analog zur Extremwertbestimmung bei reellen Funktionen kann man auch in diesem Kontext relative Extrema angeben. In der Variationsrechnung wird bewiesen, daß eine Funktion, die einem Funktional ein lokales Extremum erteilt, unter gewissen Voraussetzungen, die hier nicht weiter erörtert werden sollen, einer Differentialgleichung genügt. In gleicher Weise, wie die Extremwertrechnung bei reellen Funktionen den Begriff der ersten und zweiten Ableitung zur Bestimmung relativer Extrema benutzt, führt die Variationsrechnung den Begriff der ersten und zweiten Variation ein, die im wesentlichen den Ableitungen entsprechen und bei der Bestimmung von Extrema analog genutzt werden.

In der Bildverarbeitung dient die Variationsrechnung dazu, ein gegebenes Bild unter Nebenbedingungen möglichst genau durch stückweise stetige Funktionen zu approximieren. Im Jahr 1984 veröffentlichten Stuart und Donald Geman einen Artikel mit dem Titel „Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images“, in dem sie (in einem stochastischen Zusammenhang) den Begriff der Segmentierungsenergie (segmentation energy) als Funktional einführen [GG84]. Dieser Artikel war Bezugspunkt für eine Reihe von Veröffentlichungen in diesem Bereich, von denen [MS85], [MS89] und [BZ87] zu den am häufigsten zitierten Werken gehören.

Im wesentlichen soll hier der Ansatz in der Sichtweise von Andrew Blake und Andrew Zisserman in [BZ87] vorgestellt werden. Die Ausführungen bleiben aber nicht auf dieses Buch beschränkt, sondern versuchen, einen allgemeinen Einblick in das große Feld der Bildsegmentierung mit Methoden der Variationsrechnung zu geben.

2.4.1 „The weak membrane“

Die Bezeichnung „weak membrane“ geht zurück auf Blake und Zisserman ([BZ87]), die damit den Approximationsprozeß durch stückweise stetige Funktionen charakterisieren. Der Begriff soll eine Approximation durch den Vergleich mit einer Membran beschreiben: Sie ist biegsam, so daß sie sich der gegebenen Funktion anpaßt. Wird sie allerdings zu stark gebogen, so bricht sie, weil sie schwach („weak“) ist. Die Bruchstellen der „weak membrane“ entsprechen dann den Unstetigkeitsstellen der approximierenden Funktion.

Sei $f \in L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}))$ das Bild und $\mathcal{U} \subset \mathcal{L}^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}))$ ein geeigneter Raum stückweise differenzierbarer Funktionen von M nach \mathbb{R} .⁷

⁷Die genaue Definition von \mathcal{U} unterscheidet sich von Autor zu Autor. Eine mathematisch korrekte Version findet man in [MS95], S. 151ff. Da an dieser Stelle nur die Prinzipien der Methoden der Variationsrechnung zur Bildsegmentierung vorgestellt werden sollen, wurde aus Übersichtsgründen auf die aufwendige korrekte mathematische Formulierung verzichtet.

Gesucht ist die Funktion $u \in \mathcal{U}$, die das Funktional

$$E(u) = D(u) + S(u) + P(u) \quad (2.6)$$

minimiert. Dabei ist

- das Funktional D definiert durch

$$D(u) := \int_M (u - f)^2 d\lambda \quad (2.7)$$

ein Maß für die Bildtreue der gesuchten Funktion,

- das Funktional S definiert durch

$$S(u) := \beta^2 \int_M \|\nabla u\|_2^2 d\lambda \quad (2.8)$$

ein Maß für die Steigungen der gesuchten Funktion und

- das Funktional P ein „Strafmaß“ für jede Diskontinuität der gesuchten Funktion.

Die Namen der Terme sind englisch motiviert: D steht für *data*, S für *spine* und P für *penalty*. β ist ein zu wählender Parameter, der den Einfluß der Steigungen der zu suchenden Funktion auf das Optimum festlegt.

Besondere Aufmerksamkeit verlangt das „Strafmaß“ P . In den meisten Fällen wird P in Abhängigkeit von Menge der Unstetigkeitsstellen der gesuchten Funktion formuliert. Da die Menge der Unstetigkeitsstellen gleichzeitig auch die Segmentgrenzen bildet, sollte der gesamte Konturzug aller Segmentgrenzen „bestraft“ werden.

Blake und Zisserman schlagen mehrere Varianten vor:

- P als Maß der gesamten Konturlänge. Eine Funktion u , die den Energieterm E minimiert, minimiert dann auch die Konturlänge der Segmentgrenzen.
- P kann zusätzlich einen Term enthalten, der die Krümmung der Kontur mißt.
- Ecken in den Konturverläufen können bewertet werden.
- Topologische Eigenschaften, wie etwa Zusammenhang oder Konvexität der kontinuierlichen Anteile der gesuchten Funktion, können bewertet werden.

Die Autoren verfolgen nur den ersten Vorschlag und setzen

$$P(u) = \alpha \text{length}(K)$$

wobei K die Menge aller Unstetigkeitsstellen ist. α ist dabei ein weiterer vom Benutzer zu wählender Parameter.

Mumford und Shah sind in [MS85] mathematisch etwas präziser und definieren die zu minimierende Energie als

$$E(u, K) = \int_M (u - f)^2 d\lambda + \beta^2 \int_{M \setminus K} \|\nabla u\|_2^2 d\lambda + \alpha \mathcal{H}^1(K) \quad (2.9)$$

wobei \mathcal{H}^1 das eindimensionale Hausdorffmaß bezeichne. Weiter ist $K \subset M$ eine abgeschlossene Menge mit Hausdorffdimension 1, die die Menge der Konturränder, also der Unstetigkeitsstellen, bestimmt. Das Energiefunktional (2.9) heißt in der Literatur auch *Mumford-Shah-Modell*, wie überhaupt der gesamte Ansatz der Energieminimierung zur Bildsegmentierung durch ein Funktional der Form $E = D + S + P$ meist als *Mumford-Shah-Ansatz* bezeichnet wird, da diese beiden Autoren in ihren Veröffentlichungen Pionierarbeit zu diversen theoretischen Problemen, insbesondere zur Existenz und zum Aussehen der energieminimierenden Funktion, geleistet haben.

Die Sätze 2.4.4 und 2.4.5 sind zwei der wichtigsten Existenzaussagen zu dieser Form des Energiefunktionals. Sie charakterisieren auch die Form der Menge der Unstetigkeitsstellen K .

Definition 2.4.1 (Segmentierung, Segment)

Im Zusammenhang der Bildsegmentierung mit Ansätzen aus der Variationsrechnung heißt eine Funktion $u \in \mathcal{U}$ auch *Segmentierung*. Die Mengen

$$\text{seg}(u) := \left\{ A \mid \begin{array}{l} A \text{ ist zusammenhängend,} \\ u|_A \text{ ist stetig} \\ \text{und } A \text{ ist maximal mit dieser Eigenschaft} \end{array} \right\}$$

heißen die *Segmente* der Segmentierung u . Minimiert u das Energiefunktional (2.6) bzw. (2.9), so heißt u eine *optimale Segmentierung*.

Bemerkung 2.4.2

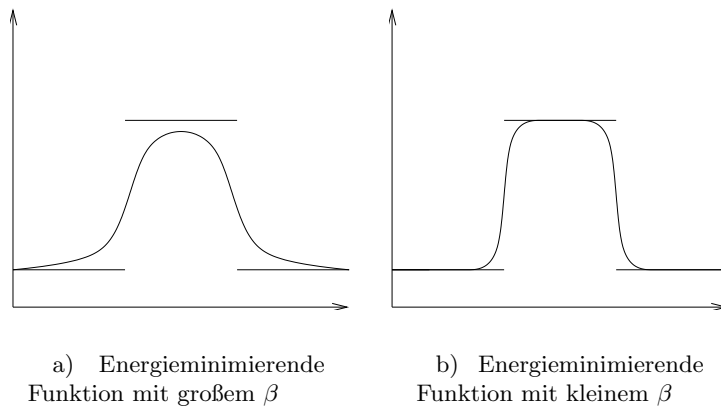
Die Abbildung 2.8 verdeutlicht die Wechselwirkung der Parameter α und β am Beispiel der Energieminimierung auf einer Hutfunktion: In Abbildung 2.8 i. wurde α so gewählt, daß das Strafmaß P keine Unstetigkeitsstellen der energieminimierenden Funktion zuläßt. Während in Abbildung 2.8 i. a) β sehr groß gewählt wurde und so starke Steigungen vermieden werden, erlaubt ein kleineres β eine höhere Bildtreue der energieminimierenden Segmentierung (Abbildung 2.8 i. b)). Absolute Bildtreue erhält man, wenn α so gewählt wurde, daß P zwei Unstetigkeitsstellen zuläßt, wie in Abbildung 2.8 ii. dargestellt.

Definition 2.4.3 (Rektifizierbar)

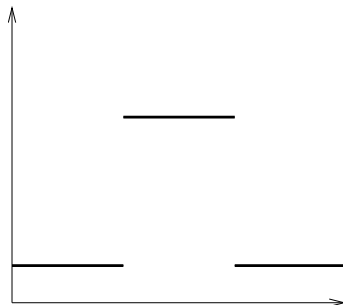
Eine Kurve $c : [0, 1] \rightarrow \mathbb{R}^2$ heißt *rektifizierbar* genau dann, wenn

$$\sup\{\|c(a_1) - c(a_2)\|_2 + \cdots + \|c(a_{n-1}) - c(a_n)\|_2 \mid n \in \mathbb{N}, 0 \leq a_1 < \cdots < a_n \leq 1\} < \infty$$

gilt.



i. Approximation mit kontinuierlichen Funktionen



ii. Approximation mit stückweise kontinuierlichen Funktionen

Abbildung 2.8: Energieminimierende Funktionen der Hutfunktion.

Satz 2.4.4

Es existiert ein Tupel (u_0, K_0) , das den Energieterm $E(u, K)$ in Gleichung (2.9) minimiert, wobei

- i. $u_0 \in \mathcal{C}^1(M \setminus K)$ und
- ii. Für jedes $\varepsilon > 0$ existiert eine endliche Anzahl rektifizierbarer Kurven c_1, \dots, c_n mit $\mathcal{H}^1(K_0 \setminus \bigcup_{i=1}^n c_i) < \varepsilon$.

Beweis:

siehe [MS95], Chapter 15

Satz 2.4.5

Für den Energieterm $E(u, K)$ in Gleichung (2.9) existiert eine Folge $(u_n, K_n)_{n \in \mathbb{N}}$, die gegen das optimale Tupel (u_0, K_0) konvergiert, wobei für alle n die Menge K_n eine endliche Menge von Lipschitz-stetigen Bögen ist.

Die Konvergenz von K_n gegen K erfolgt dabei bezüglich des Hausdorff-Abstands.

Beweis:

siehe [MMS92]

Bemerkung 2.4.6

Mumford und Shah äußerten in ihrer Arbeit [MS89] die Vermutung, daß die Menge der Segmentgrenzen der optimalen Segmentierung K_0 aus einer endlichen Menge von \mathcal{C}^1 -Bögen besteht. Diese Behauptung, die in der Literatur auch unter dem Namen „Mumford-Shah conjecture“ bekannt ist, konnte jedoch bislang noch nicht mathematisch bewiesen werden. Lediglich für die Approximation mit stückweise konstanten Funktionen wurde die Richtigkeit der Behauptung von mehreren Autoren nachgewiesen (unter anderem auch in [MS95]). Dennoch zeigen diese Annahme und Satz 2.4.4, daß die Menge K der Unstetigkeitsstellen einer Segmentierung, d. h. die Menge der Segmentgrenzen, eine regelmäßige Struktur aufweist, was eine gewisse Nähe zu Charakterisierung iv. des Begriffes „Segmentierung“ von Haralick und Shapiro (S. 14, (2.1)) anzeigt.

2.4.2 „The weak membrane“ im Diskreten

Um den Prozeß der Funktionalminimierung auf diskreten Daten mit dem Computer berechnen zu können, muß der kontinuierliche Ansatz durch ein diskretes Minimierungsproblem approximiert und dabei der eigentliche Bereich der Variationsrechnung verlassen werden. Dies geschieht in Anlehnung an die Methode der *finiten Elemente*. Diskretisiert wird im folgenden der Energieterm (2.6).

2.4.2.1 Der eindimensionale Fall: „line process“

Sei im folgenden das (eindimensionale) Bild gegeben durch einen Vektor $d = (d_0, \dots, d_N)$. Eine Funktion $\tilde{u} : [0, N] \rightarrow \mathbb{R}$, $\tilde{u} \in \mathcal{U}$ wird durch einen Polygonzug mit Knoten $(i, u_i = \tilde{u}(i))_{0 \leq i \leq N}$ interpoliert. Zur besseren Veranschaulichung kann man sich vorstellen, daß die Funktionen aus \mathcal{U} nur an endlich vielen äquidistanten Stellen ausgewertet werden.

Die interpolierende Funktion u ist dann auf jedem Intervall $[i-1, i]$, mit $i \in \{1, \dots, N\}$, affin-linear. Diese Funktion u läßt sich nun als endlicher Vektor (u_1, \dots, u_N) darstellen und läßt sich im folgender einfacher handhaben.

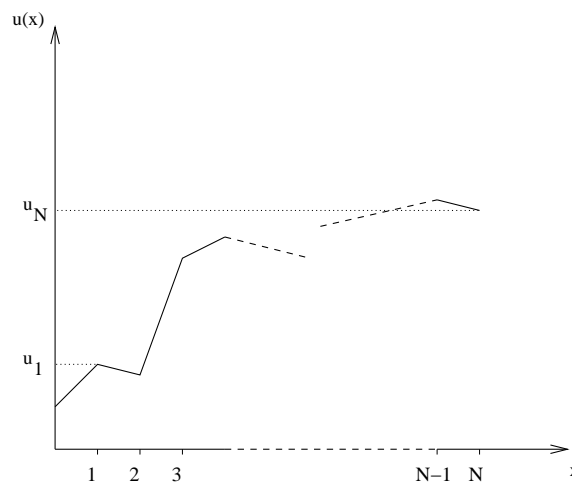


Abbildung 2.9: Diskrete Darstellung von u als Polygonzug.

Die Energien D , S und P der Terme (2.7) und (2.8) werden dann folgendermaßen umformuliert:

$$D(u) := \sum_{i=0}^N (u_i - d_i)^2$$

$$S(u, l) := \beta^2 \sum_{i=1}^N (u_i - u_{i-1})^2 (1 - l_i)$$

$$P(u, l) := \alpha \sum_{i=1}^N l_i$$

Dabei ist $l = (l_1, \dots, l_N)$ ein Binärvektor, wobei $l_i = 1$ anzeigt, daß im Intervall $]i-1, i[$ eine Unstetigkeit der Funktion \tilde{u} vorliegt, und $l_i = 0$ anzeigt, daß die Funktion \tilde{u} im Intervall $]i-1, i[$ stetig ist. Ist $l_i = 1$, so wird der Energieterm des Polygonzuges $(i, u_i)_{0 \leq i \leq N}$ für das Intervall $]i-1, i[$ nicht addiert. Stattdessen wird die Strafe α vergeben.

Die Bestimmung der optimalen Belegung des Binärvektors l nennen die Autoren „line process“. Ziel ist es, den „line process“ möglichst zu isolieren.

Anmerkung:

An dieser Stelle ist es eigentlich nicht mehr gerechtfertigt, von einem Funktional $E = D + S + P$ zu reden. Vielmehr ist E an dieser Stelle eine Funktion von $\mathbb{R}^N \times \{0, 1\}^N$ nach \mathbb{R}_+ definiert durch $E(u, l) := D(u) + S(u, l) + P(u, l)$.

Sei die Funktion $h_{\alpha, \beta} : \mathbb{R} \times \{0, 1\} \rightarrow \mathbb{R}$ definiert durch $h(t, k) := \beta^2 t^2 (1 - k) + \alpha k$, dann ist

$$S(u, l) + P(u, l) = \sum_{i=1}^N h_{\alpha, \beta}(u_i - u_{i-1}, l_i)$$

Diese Darstellung erlaubt es, die Optimierung der Energiefunktion in zwei Optimierungen zu spalten, da der Term $D(u)$ nicht von l abhängt.

$$E_0 = \min_{u, l} \left\{ D(u) + \sum_{i=1}^N h_{\alpha, \beta}(u_i - u_{i-1}, l_i) \right\} \quad (2.10)$$

$$E_0 = \min_u \left\{ D(u) + \min_l \left\{ \sum_{i=1}^N h_{\alpha, \beta}(u_i - u_{i-1}, l_i) \right\} \right\} \quad (2.11)$$

Statt nun E_0 über eine gemeinsame Minimierung von $u \in \mathbb{R}^{N+1}$ und $l \in \{0, 1\}^N$ zu berechnen (Gleichung (2.10)), kann man den Wert E_0 bestimmen, indem man die Unstetigkeitsstellen isoliert betrachtet (Gleichung (2.11)).

Definiert man $g_{\alpha, \beta}(t) := \min_{k \in \{0, 1\}} h_{\alpha, \beta}(t, k)$, so ist $E_0 = \min_{u \in \mathbb{R}^{N+1}} D(u) + \sum_{i=1}^N g_{\alpha, \beta}(u_i - u_{i-1})$.

Die Funktion $g_{\alpha, \beta}$ ist das Minimum zweier Funktionen:

$$g_{\alpha, \beta}(t) = \begin{cases} \beta^2 t^2 & , \text{ falls } |t| < \frac{\sqrt{\alpha}}{\beta} \\ \alpha & , \text{ sonst} \end{cases}$$

Das bedeutet, es ist

$$l = \begin{cases} 1 & , \text{ falls } |t| > \frac{\sqrt{\alpha}}{\beta} \\ 0 & , \text{ sonst} \end{cases} \quad (2.12)$$

Damit befindet sich im Intervall $]i-1, i[$ eine Unstetigkeitsstelle, falls $|u_{i-1} - u_i| > \frac{\sqrt{\alpha}}{\beta}$.

Der Wert $\frac{\sqrt{\alpha}}{\beta}$ ist also ein Schwellwert für Unstetigkeitsstellen.

Die Suche nach der minimalen Energie E_0 erfolgt mit Methoden der nichtlinearen Optimierung, wie z. B. Gradientenabstiegsverfahren mit Schrittweitensteuerung (vgl. z. B. [Bra86] oder [Spe93]). Weitere Ansätze zur Optimierung finden sich in Abschnitt 2.4.3.

2.4.2.2 Erweiterung auf den zweidimensionalen Fall

Eine mögliche Erweiterung des Polygonansatzes auf den zweidimensionalen Fall ist die Darstellung durch Dreieckspaare, auf deren Dreiecken die möglichen Funktionen affin-linear sind.

Gegeben sei ein diskretes Bild⁸ $d = (d_{ij})_{\substack{0 \leq i \leq N \\ 0 \leq j \leq N}}$ und die Funktionenklasse

$$\mathfrak{U} := \{u \in C^1([0, N]^2, \mathbb{R}_+) \mid u|_{n(i,j)} \text{ linear}, u|_{s(i,j)} \text{ linear}, 1 \leq i, j \leq N\}$$

Dabei ist $n(i, j) := \{(x, y) \in \mathbb{R}^2 \mid i \leq x, j \leq y, x + y \leq (i + 1) + (j + 1)\}$ und $s(i, j) := \{(x, y) \in \mathbb{R}^2 \mid x \leq i, y \leq j, (i - 1) + (j - 1) \leq x + y\}$ (vgl. Abbildung 2.10).

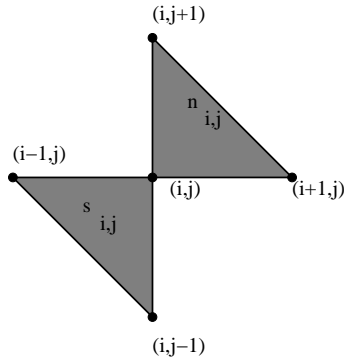


Abbildung 2.10: Lineare Dreieckselemente $n(i, j)$ (oben) und $s(i, j)$ (unten).

$n_{i,j}$ sei die Boolesche Variable, die anzeigt, ob in $n(i, j)$ eine Unstetigkeitsstelle vorliegt, $s_{i,j}$ der Boolesche Ausdruck gleicher Aussage für die Menge $s(i, j)$. Liegt eine solche Unstetigkeitsstelle vor, so wird die Energie des entsprechenden Dreiecks nicht aufsummiert, sondern durch die Strafe α ersetzt.

Eine Unstetigkeitskontur besteht dann aus einer Kette von Dreiecken, die als Unstetigkeiten enthaltend markiert wurden. Die „Strafe“ P für jede Unstetigkeit in der zu minimierenden Energie ist dann α multipliziert mit der Anzahl der unstetigen Dreiecke

$$P(u, (l, m)) := \alpha \sum_{(i,j) \in \{1, \dots, N-1\}^2} l_{i,j} + m_{i,j}.$$

Analog zum eindimensionalen Fall ergibt sich damit folgende Energiefunktion:

$$\begin{aligned} E(u, (l, m)) = & \\ D(u) + & \sum_{(i,j) \in \{1, \dots, N-1\}^2} h_{\alpha, \beta} \left(\sqrt{(u_{i,j} - u_{i-1,j})^2 + (u_{i,j} - u_{i,j-1})^2}, s_{i,j} \right) \\ & + \sum_{(i,j) \in \{1, \dots, N-1\}^2} h_{\alpha, \beta} \left(\sqrt{(u_{i,j} - u_{i+1,j})^2 + (u_{i,j} - u_{i,j+1})^2}, n_{i,j} \right) \end{aligned}$$

⁸Man beachte den Unterschied zum diskreten Pixelbild gemäß Definition 1.3.2.

Auch im zweidimensionalen Fall kann man nun den „line process“ von der Minimierung des Abstands zum Bild trennen.

2.4.3 Eliminierung nicht-signifikanter lokaler Optima

Bei der Approximation eines Bildes mit stückweise stetigen Funktionen ergeben sich zwangsläufig viele lokale Optima, die nur geringen Aussagewert haben und weit vom globalen Optimum entfernt sind. Ein großer Teil der Veröffentlichungen im Bereich der Variationsrechnung zur Bildsegmentierung beschäftigt sich mit der Elimination solcher nicht-signifikanter lokaler Optima. Einige weit verbreitete Ansätze sollen an dieser Stelle vorgestellt werden.

2.4.3.1 Der GNC-Algorithmus

Blake und Zisserman schlagen einen zweistufigen Algorithmus zur Eliminierung nicht-signifikanter lokaler Optima bei der Minimierung der Energiefunktion vor. Sie nennen ihn „Graduated Non-Convexity Algorithm“. Das Prinzip dieses Algorithmus ist einfach: Im ersten Schritt approximiert man die gegebene nicht-konvexe Kostenfunktion durch eine konvexe, deren Minimum man durch ein Gradientenabstiegsverfahren bestimmt.

Im zweiten Schritt definiert man eine Folge von Funktionen, die absteigend gegen die gegebene Energiefunktion konvergiert. Dabei sollte jede dieser Funktionen in einer Umgebung des Minimums der Vorgängerfunktion konvex sein. Ausgehend vom Minimum des vorhergehenden Energieterms in der Folge bestimmt man das (lokale) Minimum der aktuellen Funktion.

Ein solches Vorgehen führt nicht immer zwangsläufig zum globalen Optimum einer Kostenfunktion. Im Zusammenhang mit der Minimierung der Energiefunktion E kann man jedoch Aussagen über Effektivität und Konvergenz machen.

2.4.3.1.1 Konvexe Approximation der Energiefunktion im eindimensionalen Fall

Die Zielfunktion $E(u) = D(u) + \sum_{i=1}^N g_{\alpha,\beta}(u_i - u_{i-1})$ wird approximiert durch eine konvexe Funktion $E^*(u) = D(u) + \sum_{i=1}^N g_{\alpha,\beta}^*(u_i - u_{i-1})$.

Eine Funktion $f : \mathbb{R}^N \rightarrow \mathbb{R}$ ist jedenfalls dann konvex, wenn die Hesse-Matrix der Funktion positiv semidefinit ist. Für die Hesse-Matrix $(\mathbf{H}_{i,j})_{1 \leq i,j \leq N}$ von E^* gilt

$$\mathbf{H}_{i,j} = \frac{\partial^2 E^*}{\partial u_i \partial u_j}(u) = 2\delta_{ij} + \sum_{k=1}^N g_{\alpha,\beta}^{*''}(u_k - u_{k-1}) \mathbf{Q}_{k,i} \mathbf{Q}_{k,j}$$

wobei

$$\mathbf{Q}_{k,i} := \frac{\partial(u_k - u_{k-1})}{\partial u_i} = \begin{cases} 1 & , \text{ falls } i = k \\ -1 & , \text{ falls } i = k - 1 \\ 0 & , \text{ sonst} \end{cases}$$

definiert sei.

Sei für ein $c^* > 0$

$$\mathbf{H}' := 2\mathbf{I} - c^* \mathbf{Q}^t \mathbf{Q}$$

Nun gilt:

\mathbf{H} ist positiv definit, wenn \mathbf{H}' positiv definit ist und $g''_{\alpha,\beta}(t) \geq -c^*$ für alle t gilt (vgl. [BZ87], Appendix D.1.).

\mathbf{H}' ist positiv definit, falls für den größten Eigenwert v_{\max} von $\mathbf{Q}^t \mathbf{Q}$ gilt:

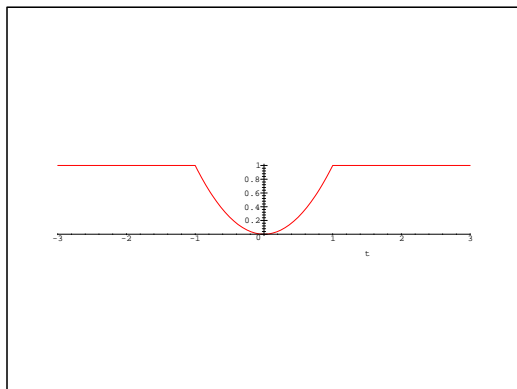
$$v_{\max} \leq \frac{2}{c^*} \quad (2.13)$$

Definiere

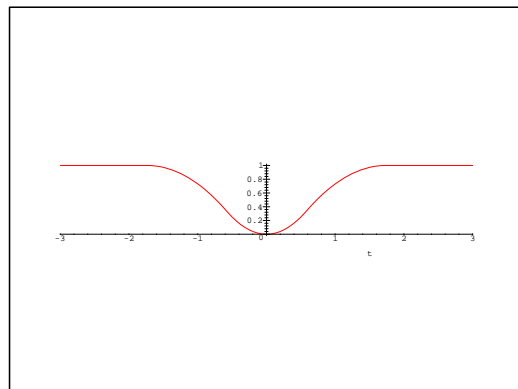
$$g_{\alpha,\beta}^*(t) := \begin{cases} \beta^2 t^2 & , \text{ falls } |t| < q \\ \alpha - c^* \frac{(|t|-r)^2}{2} & , \text{ falls } q \leq |t| < r \\ \alpha & , |t| \geq r \end{cases}$$

wobei $r^2 := \alpha \left(\frac{2}{c^*} + \frac{1}{\beta^2} \right)$ und $q := \frac{\alpha}{\beta^2 r}$ definiert sei.

Der Hintergrund dieser Transformation ist eine Approximation von unten: $g_{\alpha,\beta}^*$ erhält man, indem man $g_{\alpha,\beta}$ mit einer Parabel der Form $t \mapsto \frac{1}{2}c^*t^2 + bt + a$ approximiert (Abbildung 2.11 skizziert diese Approximation).



Originalfunktion $g_{\alpha,\beta}$



Approximation von $g_{\alpha,\beta}$ durch $g_{\alpha,\beta}^*$

Abbildung 2.11: Skizze zur Definition von $g_{\alpha,\beta}^*$.

Nun gilt:

- i. $\forall t : g_{\alpha,\beta}^*(t) \leq g_{\alpha,\beta}(t)$
- ii. $g''_{\alpha,\beta} \geq -c^*$

Wählt man $c^* = \frac{2}{v_{\max}}$, um g^* „so nah wie möglich“ bei g zu erhalten, so ist die Gleichung (2.13) erfüllt und \mathbf{H}' ist positiv definit. Man erhält daher E^* als konvexe Approximation der Energiefunktion E .

2.4.3.1.2 Konvexe Approximation der Energiefunktion im zweidimensionalen Fall

Um im zweidimensionalen Fall analog vorgehen zu können, müssen die Wurzeln innerhalb der Funktion $h_{\alpha,\beta}$ beseitigt werden. Statt E wie in Abschnitt 2.4.2.2 zu formulieren, schlagen Blake und Zisserman daher folgenden Energieterm vor:

$$E(u, (l, m)) = D(u) + \sum_{(i,j) \in \{1, \dots, N-1\}^2} h_{\alpha,\beta}(u_{i,j} - u_{i-1,j}, l_{i,j}) + \sum_{(i,j) \in \{1, \dots, N-1\}^2} h_{\alpha,\beta}(u_{i,j} - u_{i,j+1}, m_{i,j})$$

Bei dieser Energie bekommen die Booleschen Variablen $l_{i,j}$ und $m_{i,j}$ neue Bedeutungen:

$l_{i,j} = 1$ zeigt eine Unstetigkeit im adjazenten linken Rechteck an.

$m_{i,j} = 1$ zeigt eine Unstetigkeit im adjazenten oberen Rechteck an.

Dieses Vorgehen stimmt zwar mit der intuitiven Vorstellung wenig überein, ist aber, da die Wurzeln wegfallen, einfacher zu rechnen. Mit dieser neuen Energiefunktion ist es nun möglich, ähnlich wie im Eindimensionalen eine Funktion $g_{\alpha,\beta}^*$ und damit eine modifizierte Energiefunktion E^* , die konvex ist, zu bestimmen. Für eine genauere Darstellung und für eine Abschätzung der Güte der konvexen Approximation sei auf den Originaltext verwiesen ([BZ87], S. 134 – 141).

2.4.3.1.3 Stufenweise Nicht-Konvexität

Um im Fall unzureichender Ähnlichkeit der konvexen Funktion E^* mit der Originalenergiefunktion E trotzdem einen Algorithmus zu haben, der das globale Minimum von E liefert, schlagen Blake und Zisserman ein Verfahren vor, das stufenweise die nicht-konvexe Funktion E durch lokal konvexe Funktionen $E^{(p)}$ annähert (**Graduated Non Convexity**).

Sei

$$g_{\alpha,\beta}^{(p)}(t) := \begin{cases} \beta^2 t^2 & , \text{ falls } |t| < q \\ \alpha - c \frac{(|t|-r)^2}{2} & , \text{ falls } q \leq |t| < r \\ \alpha & , \text{ falls } |t| \geq r \end{cases}$$

wobei $c := \frac{c^*}{p}$, $r := \alpha \left(\frac{2}{c} + \frac{1}{\beta^2} \right)$ und $q := \frac{\alpha}{\beta^2 r}$ definiert sei.

Setze

$$E^{(p)}(u) := D(u) + \sum_{i=1}^N g_{\alpha,\beta}^{(p)}(u_i - u_{i-1})$$

dann ist $E^{(1)} = E^*$, da $c = c^*$ ist und daher $g_{\alpha,\beta}^{(1)} = g_{\alpha,\beta}^*$ gilt. Weiter gilt, daß für p gegen 0 fallend, die Funktion $g_{\alpha,\beta}^{(p)}$ gegen $g_{\alpha,\beta}$ strebt.

Der GNC-Algorithmus

Berechne die Segmentierung $u^* = u^{(1)}$, das globale Minimum der konvexen Funktion $E^{(1)}$. Sei für ein $k \in \mathbb{N}, k > 0$ die Folge $(p_i)_{i=1}^k$ streng monoton fallend mit $p_k = 0$ und $p_1 = 1$. Für $i = 2$ wird nun mit $u^{(1)}$ als Startwert ein lokales Minimum $u^{(2)}$ von $E^{(p_2)}$ berechnet und so weiter:

$u^{(i+1)}$ ist das lokale Minimum der Funktion $E^{(p_{i+1})}$, berechnet mit $u^{(i)}$ als Startwert der nichtlinearen Optimierung ($1 \leq i \leq k - 1$).

Bei diesem Vorgehen ist es zwar nicht ausgeschlossen, daß der Algorithmus ein falsches – nämlich nur ein lokales – Minimum der Energiefunktion findet. Blake und Zisserman zeigen jedoch für eine ausreichend große Klasse von Fällen, daß der Algorithmus tatsächlich das globale Minimum der Energiefunktion liefert (vgl. [BZ87], S. 142 – 151). Wichtig dabei ist auch die geeignete Wahl der Folge $(p_i)_{i=1}^k$.

2.4.3.2 Der Multiskalen-Ansatz

In [Mar82] (S. 46) bemerkt David Marr:

In other words, the primitives of our representation must work at a number of different scales.

Die in der Literatur unter dem Namen „Multiskalen-Algorithmen“ (engl. „multiscale algorithms“) bekannten Verfahren beziehen sich meist auf diese Aussage Marrs. Die meisten Algorithmen in diesem Bereich verschmieren das gegebene Bild und versuchen, in den verschmierten Bildern stabile Strukturen, d. h. Strukturen, die auch in weniger verschmierten Bildern vorkommen, zu entdecken.⁹

Bei den Methoden der Variationsrechnung zur Bildverarbeitung wird in der Regel der Schwellwert für die Unstetigkeitsstellen variiert. Eine gute Beschreibung für ein solches Verfahren findet man in [KLM94]. Der Raum \mathcal{U} , der Definitionsbereich des Energiefunktionals, wird dort auf die Menge der stückweise konstanten Funktionen beschränkt. Daher ist der Strafterm S für die Steigungen immer 0 und kann somit ignoriert werden. Damit kann als einziger Parameter zur Einstellung des Schwellwertes der Sprunghöhen nur der Parameter α des Strafmaßes P der Unstetigkeitsstellen variiert werden. Das Strafmaß P wird in [KLM94] über die Gesamtlänge der Segmentgrenzen gemessen. Man beachte, daß bei dieser Wahl des Definitionsbereichs die auf einem Segment A approximierende Funktion durch $\frac{1}{\lambda(A)} \int_A f d\lambda$, d. h. dem durchschnittlichen Grauwert des Segments, gegeben ist.

Begonnen wird mit einer Energieminimierung mit einer geringen Strafe α_0 für die Länge der Segmentgrenzen. Das Ergebnis u_0 dieser Energieminimierung ist dann der Startwert eines Energieminimierungsalgorithmus mit einer höheren Strafe α_1 . Dieses Verfahren wird

⁹Der Begriff „Multiskalen-Algorithmus“ wird an dieser Stelle sehr pragmatisch verwendet. Für eine detailliertere Einführung des Begriffes, der eigentlich aus dem Modell der Diffusionsrechnung in der Bildverarbeitung stammt, vgl. etwa [MS95], S. 7 oder [Wei98].

mit sukzessiver Erhöhung der Strafe α wiederholt, bis ein gewünschtes Ergebnis erreicht wird.¹⁰ In [KLM94] wird dazu noch vorgeschlagen, daß, bevor eine Segmentierung als Startwert verwendet wird, zwei adjazente Segmente, die sich bezüglich eines Ähnlichkeitsmaßes unter allen adjazenten Segmenten am ähnlichsten sind, vereinigt werden. Man erhält so eine Segmentierung u_1 , deren Energiewert $E(u_1)$ sehr schnell aus den Energietermen der einzelnen Segmente berechnet werden kann. Diese Vereinigung wird für ein vorgegebenes $\alpha_1 > \alpha_0$ solange wiederholt, bis sich keine Segmente mehr finden, deren Vereinigung den Wert des Energieterms (mit α_1 als Strafe für die Länge der Segmentgrenzen) reduziert.

Sei nun $g : M \rightarrow \mathbb{R}$ ein diskretes Pixelbild, und sei u_0 die energieminimierende Segmentierung eines Energiefunktional E definiert über g mit Strafmaß α_0 für die Länge der Segmentgrenzen. Ist dann u_1 aus u_0 durch die Vereinigung zweier adjazenter Segmente entstanden, so ist das Ähnlichkeitsmaß so gewählt, daß die Differenz $E(u_1) - E(u_0)$ die geringste unter allen möglichen Vereinigungen adjazenter Segmente ist.¹¹ Als Ähnlichkeitsmaß δ_E ergibt sich dann für zwei pixelbasierte Segmente $A, B \in \text{seg}(u)$:

$$\delta_E(A, B) := \int_A (\tilde{\gamma}_g(A) - g)^2 d\lambda + \int_B (\tilde{\gamma}_g(B) - g)^2 d\lambda + \alpha \lambda(\text{Rand}(A \cup B)) - \int_{A \cup B} (\tilde{\gamma}_g(A \cup B) - g)^2$$

Dabei bezeichne $\tilde{\gamma}_g(A)$ den durchschnittlichen Grauwert der Funktion g auf der Menge A .¹² Die Menge $\text{Rand}(A)$ ist die Menge aller Randpixel des Segments A und daher Lebesguemeßbar. Durch Umformungen kann man zeigen, daß

$$\delta_E(A, B) = \alpha \lambda(\text{Rand}(A \cup B)) - \lambda(A) \tilde{\gamma}_g(A)^2 - \lambda(B) \tilde{\gamma}_g(B)^2 + \lambda(A \cup B) \tilde{\gamma}_g(A \cup B)^2 \quad (2.14)$$

gilt. Auf diese Gleichheit wird an späterer Stelle, nämlich in Abschnitt 5.3, eingegangen.

2.4.3.3 Andere Methoden

Im Prinzip können sämtliche Methoden der nichtlinearen Optimierung nicht-konvexer Funktionen auch bei der Minimierung des Energiefunktional angewandt werden. Schlagwörter wie „simulated annealing“, „genetische Algorithmen“ usw. finden jedoch, da die Zielfunktion genau bekannt ist, in der Literatur nur wenig Beachtung. Weitere Methoden, das Energiefunktional zu minimieren, finden sich z. B. in [Ric90] und [MS94].

¹⁰Der Begriff „gewünscht“ heißt hier, daß ein optisch zufriedenstellendes Segmentierungsergebnis erzielt werden soll. Dieses Ergebnis läßt sich also keineswegs mathematisch fassen.

¹¹Da u_0 das Funktional E minimiert, muß $E(u_1) - E(u_0) \geq 0$ sein.

¹²Die Bezeichnung $\tilde{\gamma}_g(A)$ wurde in Anlehnung an die Definition 3.5.9 des Kapitels 3 gewählt. Das Beispiel 3.5.11 i) rechtfertigt diese Wahl.

2.4.4 Verwandte Ansätze

Das Prinzip der Minimierung eines Energiefunktional zur Segmentierung von Bildern ist nicht auf die Minimierung der Energie des Mumford-Shah-Modells beschränkt. Von den zahlreichen anderen Energietermen sollen an dieser Stelle exemplarisch vier weitere erwähnt werden.

2.4.4.1 Der „2.1-D Sketch“

Der Name „2.1-D Sketch“ soll andeuten, daß zusätzlich zur zweidimensionalen Information des gegebenen Bildes noch eine Überdeckungsrelation genutzt wird. Diese Überdeckungsrelation ist Träger dreidimensionaler Informationen wie „Region A liegt vor Region B “, liefert jedoch kein komplettes dreidimensionales Bild.

Die Überdeckungsrelation, genauer gesagt handelt es sich dabei um eine Präordnung, wird im „2.1-D Sketch“ mit dem klassischen Mumford-Shah-Modell verbunden. Die Energieminimierung bezüglich des Energiefunktional nach Mumford und Shah bzw. Blake und Zisserman liefert Segmentierungen durch disjunkte zusammenhängende Mengen. In [NM90] erweitern Mark Nitzberg und David Mumford das klassische Mumford-Shah-Modell um das Konzept, auch sich überlappende Segmente zuzulassen. Sie modellieren damit Überdeckungseffekte.

Ist \mathcal{R} die Menge der Regionen, die sich auch überlappen können, so sei \prec eine Präordnung auf \mathcal{R} . Für zwei Regionen $R_1, R_2 \in \mathcal{R}$ gelte $R_1 \prec R_2$ genau dann, wenn R_1 durch R_2 verdeckt wird.

Ist Θ die Menge aller solcher Präordnungen, so wird der Energieterm

$$E : (\wp P(\mathcal{R}), \Theta) \rightarrow \mathbb{R}_+$$

$$E : (\{R_1, \dots, R_n\}, \prec) \mapsto \sum_{i=1}^n \left(\mu^2 \int_{R'_i} (g - m_i)^2 + \varepsilon \lambda(R_i) + \alpha P(\text{Rand}(R_i)) \right)$$

mit $\bigcup_{i=1}^n R_i = M$ minimiert.

Dabei ist für $1 \leq i \leq n$ der Wert $m_i := \frac{1}{\lambda(R'_i)} \int_{R'_i} g d\lambda$ der Mittelwert von g auf der Menge R'_i . Die Approximation erfolgt also über stückweise konstante Funktionen.

Für eine Menge $\{R_1, \dots, R_n\} \in \mathcal{P}(\mathcal{R})$ und eine Präordnung \prec auf \mathcal{R} bezeichnet die Menge R'_i den unverdeckten Anteil von $R_i \in \{R_1, \dots, R_n\}$ bezüglich der Präordnung. Es ist also $R'_i := R_i \setminus \bigcup_{R \in \{R_1, \dots, R_n\}, R_i \prec R} R$. P bezeichnet den Strafterm, der die Länge des Randes von

R_i , also inklusive des verdeckten Anteils, bestraft. Die Formulierung dieses Strafterms fällt beim „2.1-Sketch“ wesentlich differenzierter aus als in Abschnitt 2.4.1. Die Abbildung 2.12 soll andeuten, was P leisten sollte.

Die Parameter μ , ε und α sind analog zum Abschnitt 2.4.1 freie Parameter, die abhängig von der jeweiligen Anwendung gesetzt werden können.

Für eine genauere Darstellung dieses Ansatzes sei auf den Originalartikel [NM90] verwiesen.

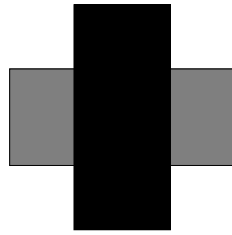


Abbildung 2.12: Die graue Region wird von der schwarzen Region überdeckt. Der Strafterm P sollte derart sein, daß der Rand der beiden grauen Flächen stärker bestraft wird, als Rand der gemeinsamen Fläche (inklusive des verdeckten Anteils).

2.4.4.2 Segmentierung über die minimale Beschreibungslänge

In [Lec89] formuliert Yvan Leclerc das Segmentierungsproblem über die Suche nach der minimalen Länge der Beschreibung eines Bildes. Dazu ist zunächst einmal eine formale Sprache zu definieren, mit der man Bilder kodieren kann. Ein Bild könnte etwa durch stückweise konstante Bildbereiche und einem gleichverteilten Rauschen beschrieben werden.

Um ein Bild in diesem Modell zu kodieren, kann man etwa die einzelnen Regionen durch ihren Rand im sogenannten „Chain-Code“ (vgl. [GW92], S. 484) beschreiben. Zusätzlich muß noch der Mittelwert der Region notiert werden. Das Rauschen wird dadurch kodiert, daß für jedes Pixel der Abstand des Mittelwertes seiner Region zum Bild notiert wird. Dabei sollte die Kodierung eines größeren Abstands länger sein als die eines kleineren Abstands. Je weniger Regionen man kodiert bzw. je kürzer die Kodierung der Regionen ist, desto größer wird dann die Kodierung der Abstände und umgekehrt.

In [Lec89] werden Energietermine für die Approximation sowohl durch stückweise konstante als auch durch stückweise polynomiale Funktionen vorgestellt.

2.4.4.3 Affin invariante Segmentierung

Beim Ansatz von Ballester, Caselles und Gonzales in [BCG96] wird der Bildentstehungsprozeß in die Überlegungen einbezogen. Die Aufnahme eines Bildes entsteht in ihrem Modell durch eine projektive Abbildung von Objekten. Ein Energiefunktional sollte also gegenüber projektiven Abbildungen invariant sein. Eine Näherung der projektiven Invarianz ist die affine Invarianz, was einer Abbildung mit Parallelprojektion entspricht. Eine Parallelprojektion wird näherungsweise bei Aufnahmen von Objekten aus großer Entfernung erreicht.

Als Szenario verwenden die Autoren Luftaufnahmen von Straßen und Feldern. Das Bild, das von der Linse auf die Bildebene projiziert wird, liegt daher im wesentlichen in einer Ebene. Sei g ein solches Bild einer Ebene, das durch Parallelprojektion entstanden ist. Dann sollte ein Energiefunktional E folgende Eigenschaft haben:

Für alle affinen Abbildungen \mathbf{A} und optimale Minimierer u_0 von E gilt: $E(u_0 \circ \mathbf{A})$ ist ebenfalls ein optimaler Minimierer von E auf dem Bild $g \circ A$.

Die Autoren schlagen ein Energiefunktional vor, das diese Vorgabe für alle affinen Abbildungen mit orthogonalem linearen Anteil erfüllt (vgl. [BCG96]). Eine interessante Eigenschaft der affin invarianten Segmentierung von Ballester, Caselles und Gonzales ist, daß lange und dünne Regionen kaum bestraft werden. Diese Eigenschaften der affin invarianten Segmentierung sind z. B. bei der Auswertung von Satellitenbildern nützlich.

2.4.4.4 Snakes

Auch der inzwischen in der Bildverarbeitung sehr verbreitete Ansatz von Kass, Witkin und Terzopoulos zur Detektion geschlossener Kantenkurven in Bildern wird in Termen der Variationsrechnung formuliert. Dieses unter dem Namen „Snakes“ bekannte Verfahren wurde 1987 in [KWT87] vorgestellt. Das Energiefunktional dient hier jedoch nicht dazu, eine stückweise stetige Approximation des Bildes zu finden, sondern eine Energie zu minimieren, die lediglich auf einer geschlossenen Konturlinie definiert ist.

Ein Snake v ist eine parametrisierte Kurve auf der Bildfläche

$$v : \begin{array}{l} [0, 1] \rightarrow M \\ s \mapsto (x(s), y(s)) \end{array}$$

Das Energiefunktional kann in drei Teilenergien aufgeteilt werden:

$$E(v) = \int_{[0,1]} E_{snake}(v(s))d\lambda(s) + \int_{[0,1]} E_{image}(v(s))d\lambda(s) + \int_{[0,1]} E_{con}(v(s))d\lambda(s)$$

Dabei dient die interne Energie E_{snake} dazu, die Krümmung der geschlossenen Kurve zu steuern. Hier spielen also wieder Ableitungen der Argumentfunktion eine Rolle.

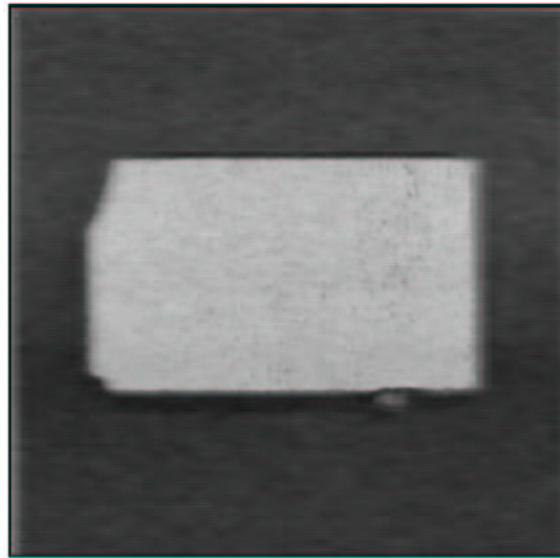
Die Bildenergie E_{image} dient dazu, diejenigen Merkmale im Bild als besonders günstig für die Energieminimierung zu setzen, die für die Problemstellung als wichtig gesehen werden. Im Falle einer Kantendetektion wäre etwa $-\|\text{grad}(f)(v(s))\|_2$ eine gute Wahl. Im Funktional E_{con} werden äußere Zwänge (engl. *constraints*), die z. B. der Benutzer vorgeben kann, kodiert.

2.4.5 Bewertung

Ausgehend von der Definition des Begriffes „Segmentierung“ durch Mumford und Shah als eine Zerlegung eines Bildes in disjunkte Bereiche, auf denen sich das Bild nur schwach und zwischen denen sich das Bild stark ändert (vgl. auch Abschnitt 2.2.1, S. 15, (2.2)), erscheint der Ansatz, das Bild durch eine stückweise stetige Funktion zu approximieren, deren Stetigkeitsbereiche die Segmente bilden, als sehr sinnvoll. Dabei muß die approximierende Funktion gewissen Glattheitsbedingungen im Inneren der Segmente und Sprungbedingungen am Rand der Segmente gehorchen. Daher ergibt sich die Modellierung als Energieminimierungsverfahren mit dem Energiefunktional $E = D + S + P$, wobei D die Approximationseigenschaften, S die Glattheitseigenschaften und P die Sprungeigenschaften der Segmentierung steuert, fast zwangsläufig. Die Parameter α und β im Energiefunktional (Abschnitt 2.4.1) bieten eine gute Möglichkeit, die Glattheit und die Sprunghöhen des Bildes, bei deren Überschreitung eine Segmentgrenze gezogen werden muß, zu steuern. Allerdings bestehen zwischen diesen Parametern Wechselwirkungen, die der ungeübte Benutzer nur langsam erschließen kann (vgl. Abbildung 2.8 oder Gleichung (2.12)).

Die Bedingungen iii. und iv. von Haralick und Shapiro (vgl. S. 14, (2.1)), nämlich, daß das Innere der Segmente nur wenige kleine Löcher enthalten soll, und daß die Ränder der Segmente nicht ausgefranst sein sollen, werden allerdings nur unzureichend erfüllt. Die Sätze 2.4.4 und 2.4.5 stellen zwar fest, daß das optimale Ergebnis gewissen Bedingungen an die Form der Segmentgrenzen genügt, für die Praxis sind diese Aussagen jedoch viel zu vage. Es ist lediglich gesichert, daß Löcher und Zerfransungen durch das Energiefunktional bestraft werden. Die Existenz dieser Phänomene kann man jedoch selbst bei optimalen Segmentierungen nicht ausschließen.

Weiter ist der Übergang von der Theorie zur Praxis bei den Methoden der Variationsrechnung zur Bildsegmentierung problematisch, wie insbesondere der Abschnitt 2.4.2.1 und die Schwierigkeiten mit zweidimensionalen Definitionsbereichen in 2.4.2.2 zeigen. Auch die Existenzaussagesätze 2.4.4 und 2.4.5 lassen sich nicht diskret formulieren. Ein Grund für diese Schwierigkeiten liegt darin, daß die Menge der Unstetigkeitsstellen K in der Theorie die Hausdorff-Dimension 1 hat (vgl. Gleichung (2.9)) und daher keine Vereinigung von Pixeln, die ja Hausdorff-Dimension 2 haben, sein kann.



Eingabebild



Ergebnissegmentierung

Abbildung 2.13: Beispiel eines Ergebnisses des GNC-Algorithmus mit $\alpha = 5000$ und $\beta = 10$.

2.5 Quantisierung von Funktionen und Bildsegmentierungen

Als Quantisierung bezeichnet man die Approximation einer Funktion durch eine stückweise konstante Funktion. Im Zusammenhang mit Bildsegmentierungen ist die zu approximierende Funktion meist auf einem kompakten Intervall aus \mathbb{R} oder einer kompakten Menge $M \subset \mathbb{R}^2$ definiert. Eine Quantisierung des Histogramms kann als Grundlage einer Segmentierung herangezogen werden (vgl. Abschnitt 2.5.3.2 bzw. Abschnitt 4.5). Der Wertebereich der zur approximierenden Funktion ist entweder diskret (z. B. bei Graustufen- oder Farbkodierungen) oder positiv reell.

Sei $\mathcal{S}_n := \left\{ \sum_{i=1}^n y_i \mathbf{1}_{[x_i, x_{i+1}[} \mid x_1 < \dots < x_{n+1} \in \mathbb{R}, y_1, \dots, y_n \in \mathbb{R} \right\}$ die Menge der stückweise konstanten Funktionen mit höchstens n Stufen. Die meisten Autoren (vgl. z. B. [GG92]) definieren die skalare Quantisierung bezüglich einer Dichte f einer Zufallsvariablen X . Sie suchen eine Funktion $s_n \in \mathcal{S}_n$, mit $s_n = \sum_{i=1}^n y_i \mathbf{1}_{[x_i, x_{i+1}[}$, die den Fehler

$$\|f - s_n\|_{\text{quant}}^2 := E[(X - s_n(X))^2] = \sum_{i=1}^n \int_{[x_i, x_{i+1}[} (x - y_i)^2 f(x) d\lambda \quad (2.15)$$

minimiert.¹³

Der Ansatz von Thomas Kämpke und Rudolf Kober in den Artikeln [KK98], [KK99] und [Käm00] minimiert nicht den erwarteten Fehler einer Approximation einer Dichte. Vielmehr wird dort im Sinne der L^p -Räume approximiert. Dieser Typ der Quantisierung, der im folgenden *Funktionsquantisierung* heißen soll, wird in diesem Abschnitt vorgestellt.¹⁴

Vereinfachend kann man sagen, daß eine Quantisierung bezüglich des Fehlers $\|\cdot\|_{\text{quant}}$ dazu tendiert, die Breakpoints x_i dorthin zu setzen, wo die Funktion f große Werte hat, während eine Quantisierung bezüglich des Fehlers $\|\cdot\|_2$ die Breakpoints eher dorthin setzen wird, wo sich f stark verändert.

Einige Ergebnisse der Theorie der Funktionsquantisierung bezüglich der L^2 -Norm werden als Spezialfälle von grundlegenden Aussagen der Mehrniveaupassung in Kapitel 3 wieder auftauchen. Die inhaltliche Nähe zu [Don91] und zur Mehrniveaupassung (vgl. Kapitel 3) soll durch die Wahl der Notation, abweichend von den Originalartikeln, unterstrichen werden.

Im folgenden Abschnitt 2.5.1 werden die wichtigsten theoretischen Ergebnisse zur Funktionsquantisierung vorgestellt. Im zweiten Abschnitt dieses Unterkapitels wird auf die algorithmische Umsetzung von Funktionsquantisierungen stückweiser konstanter Funktionen eingegangen. Abschließend wird im vorletzten Abschnitt der Bezug zur Segmentierung von Bildern hergestellt. Abschließend wird die Funktionsquantisierung zur Bildsegmentierung nach den Maßstäben des Unterkapitels 2.2 bewertet.

¹³Zur Quantisierung bezüglich $\|f - \cdot\|_{\text{quant}}$ vgl. auch Abschnitt 4.5.

¹⁴Beide Probleme, das der Quantisierung bezüglich einer Dichte und das der Funktionsquantisierung, lassen sich unter bestimmten Voraussetzungen durch Substitution ineinander überführen (vgl. etwa [Käm00]).

2.5.1 Optimale und fast optimale Quantisierung integrierbarer Funktionen

Alle Resultate dieses Abschnitts sind [Käm00] entnommen. Dort sind auch alle Beweise der vorgestellten Sätze und Lemmata zu finden. Viele Ergebnisse werden im Originalartikel für L^p -Räume mit $p \in \{1, 2, \dots, \infty\}$ angegeben. Für die vorliegende Arbeit – wie auch für die meisten Anwendungen der Funktionsquantisierung in der Bildsegmentierung – ist allerdings nur $p = 2$ von Relevanz. Deswegen wird an dieser Stelle nur dieser Fall vorgestellt. Für die Fälle $p \neq 2$ sei auf den Originalartikel verwiesen.

In diesem Abschnitt gelte immer $f \in \mathcal{C}([a, b], \mathbb{R})$.

Definition 2.5.1 (Menge der Funktionsquantisierungen)

\mathcal{S}_n definiert durch

$$\mathcal{S}_n := \left\{ \sum_{i=1}^n y_i 1_{[x_i, x_{i+1}[} \mid y_1, \dots, y_n \in \mathbb{R}, x_1 \leq x_2 \leq \dots \leq x_{n+1} \in \mathbb{R} \right\}$$

heißt *die Menge der Funktionsquantisierungen durch n Stufen*. Ist $s_n \in \mathcal{S}_n$ der Form $s_n = \sum_{i=1}^n y_i 1_{[x_i, x_{i+1}[}$, so gilt insbesondere $s_n(x) = 0$ für $x < x_1$ oder $x > x_{n+1}$. Die Punkte x_1, \dots, x_{n+1} heißen *Breakpoints* oder *Knotenpunkte* der Funktionsquantisierung s_n .

Bemerkung 2.5.2

Sei $s_n \in \mathcal{S}_n$ mit $s_n = \sum_{i=1}^n y_i 1_{[x_i, x_{i+1}[}$. Dann heißen die Mengen $[x_i, x_{i+1}[$ auch *Niveaumengen* der Funktionsquantisierung s_n und y_i die *Niveauhöhen* der Niveaumenge $[x_i, x_{i+1}[$ für $1 \leq i \leq n$.

Man vergleiche zu diesen Bezeichnungen die Einleitung zu den theoretischen Grundlagen der Mehrniveaupassung in Abschnitt 3.1 und die Definition 3.3.1 (S. 76).

Definition 2.5.3 (Optimale Funktionsquantisierung)

Eine Funktion $s_n^0 \in \mathcal{S}_n$ heißt *n -optimale Funktionsquantisierung für f bezüglich der L^2 -Norm*, falls

$$\|f - s_n^0\|_2 = \min_{s_n \in \mathcal{S}_n} \|f - s_n\|_2$$

gilt.

Satz 2.5.4 (Existenz)

Eine n -optimale Funktionsquantisierung existiert unter den Voraussetzungen dieses Abschnitts immer.

Lemma 2.5.5 (Massenerhaltungslemma)

Ist s_n^0 die n -optimale Funktionsquantisierung von f , so gilt:

$$\int_{[a, b]} s_n^0(x) d\lambda = \int_{[a, b]} f d\lambda$$

2.5.1.1 Funktionsquantisierung mit Nebenbedingungen

Bei der Funktionsquantisierung bezüglich der L^2 -Norm sind zwei Parameter zu bestimmen. Ein Parameter ist die Sequenz der Breakpoints x_1, \dots, x_{n+1} , der andere die Sequenz der Funktionswerte y_1, \dots, y_n der Funktionsquantisierung. In den beiden folgenden Abschnitten werden einige Ergebnisse vorgestellt, die sich ergeben, wenn man einen dieser Parameter festhält und nach dem anderen optimiert. Das Resultat des ersten Abschnitts ist, daß die Partition, d. h. die Sequenz der Breakpoints, die Funktionswerte eindeutig bestimmt.

Eine entsprechende Aussage für feste Funktionswerte und variable Breakpoints im zweiten Abschnitt fällt etwas differenzierter aus.

2.5.1.1.1 Feste Partitionen

In diesem Abschnitt soll untersucht werden, wie sich bei fest vorgegebener Partition des Definitionsbereichs die Funktionswerte der optimalen Funktionsquantisierung von f verhalten.

Satz 2.5.6

Sei s_n^0 eine n -optimale Funktionsquantisierung von f und $[x_i, x_{i+1}[$ ein Intervall dieser Funktionsquantisierung mit $x_i < x_{i+1}$. Dann ist der Wert $y_{i,0}$ von s_n^0 auf diesem Intervall eindeutig bestimmt.

Genauer gilt

$$y_{i,0} = \frac{1}{x_{i+1} - x_i} \int_{[x_i, x_{i+1}[} f d\lambda$$

Definition 2.5.7

(i) Definiere $\tilde{\gamma}_f : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}$ durch

$$\tilde{\gamma}_f : A \mapsto \begin{cases} \frac{1}{\lambda(A)} \int_A f d\lambda & , \text{ falls } \lambda(A) > 0 \\ 0 & , \text{ sonst} \end{cases}$$

(ii) Sei $n \in \mathbb{N}$ mit $n \geq 2$. Dann sei die Menge $\mathcal{Z}_n(a, b)$ definiert durch

$\mathcal{Z}_n(a, b) := \{(x_2, \dots, x_n) \mid a < x_2 < \dots < x_n < b\}$. Die Funktion $\Phi_f : \mathcal{Z}_n(a, b) \rightarrow \mathbb{R}$ sei definiert durch

$$\Phi_f((x_2, \dots, x_n)) := \left\| f - \sum_{i=1}^n \tilde{\gamma}_f([x_i, x_{i+1}[) \mathbf{1}_{[x_i, x_{i+1}[} \right\|_2^2$$

mit $x_1 := a$ und $x_{n+1} := b$.

Bemerkung 2.5.8

Mit der obigen Definition lassen sich die Funktionswerte $y_{i,0}$ der n -optimalen Funktionsquantisierung aus Satz 2.5.6 als $\tilde{\gamma}_f([x_i, x_{i+1}[)$ schreiben. Weiter zeigt Satz 2.5.6, daß man das Problem der optimalen Funktionsquantisierung auf ein Problem zurückführen kann, das lediglich von den Partitionen abhängt. Es genügt nämlich Φ_f zu minimieren.

Man vergleiche hierzu die entsprechenden Lemmata 3.5.3 (S. 93) und 3.5.21 (S. 102) der vektoriellen Mehrniveaupassung.

2.5.1.1.2 Feste Niveauhöhen

Seien nun die Niveauhöhen fest gewählt. Wünschenswert wäre es, ein Pendant zu Satz 2.5.6 zu haben, das das Problem der Funktionsquantisierung auf das Problem, die richtigen Niveauhöhen zu finden, zurückführt. Dies läßt sich allerdings nicht in so allgemeiner Weise formulieren wie im vorangegangenen Abschnitt.

Lemma 2.5.9 (Notwendige Bedingung der Partition)

Sei $f \in \mathcal{C}([a, b], \mathbb{R}) \cap \mathcal{L}^2(\mathcal{B}([a, b]), \mathcal{B}(\mathbb{R}))$ zusätzlich monoton steigend und differenzierbar auf $]a, b[$. Seien weiter die Niveauhöhen y_1, \dots, y_n fest gewählt mit $f(a) < y_1 < \dots < y_n < f(b)$.

Dann erfüllt die optimale Partition $a = x_{1,0} < x_{2,0} < \dots < x_{n+1,0} = b$ die Gleichung $f(x_{i,0}) = \frac{y_{i-1} + y_i}{2}$, für $i = 2, \dots, n$.

Partitionen, die diese Bedingung erfüllen, sind lokale Minima der Funktion Φ_f aus Definition 2.5.7.

Bemerkung 2.5.10

Die Lösungen $x_{i,0}$ im Lemma 2.5.9 müssen nicht eindeutig bestimmt sein. Daher liefert das Lemma 2.5.9 nur eine notwendige Bedingung an die Partition und keine algorithmisch so einfach umsetzbare Anweisung, wie es bei Lemma 2.5.6 im Falle der festen Partition für die Niveauhöhe der Fall ist.

Das folgende Lemma entfernt die Bedingung der Monotonie.

Lemma 2.5.11

Sei f differenzierbar auf $]a, b[$. Weiter seien aufeinander folgende Niveaus von unterschiedlicher Höhe (d. h. es gilt $y_i \neq y_{i+1}$ für $i = 1, \dots, n-1$).

Dann genügen die Breakpoints $x_{i,0}$ der optimalen Segmentierung der Bedingung

$$x_{1,0} = a, \quad x_{n+1,0} = b$$

und

$$f(x_{i,0}) = \frac{y_{i-1} + y_i}{2} \wedge \operatorname{sgn}(f'(x_{i,0})) = \operatorname{sgn}(y_i - y_{i-1})$$

für $i = 2, \dots, n$.

Definition 2.5.12 (Kritische Punkte)

Punkte x_i , die der Breakpointbedingung in Lemma 2.5.11 genügen, heißen *kritische Punkte* für die Funktionsquantisierung von f .

2.5.1.2 Variable Partitionen und variable Niveauhöhen: Algorithmen zur Funktionsquantisierung ohne Nebenbedingungen

Die Ergebnisse der vorhergehenden Abschnitte werden nun verwendet, um Algorithmen zur Berechnung der n -optimalen Funktionsquantisierung von f auf einem Intervall $[a, b]$ bezüglich der $\|\cdot\|_2$ -Norm bereitzustellen.

2.5.1.2.1 Ein generischer Algorithmus zur Funktionsquantisierung

Der Algorithmus 1 (**GenP**) nutzt die Ergebnisse der beiden vorangegangenen Abschnitte, um eine möglichst optimale Partition zu berechnen. Abwechselnd werden zunächst für eine feste Partition die optimalen Niveauhöhen bestimmt und dann für feste Niveauhöhen eine optimale, d. h. eine Partition, die der Breakpointbedingung in Lemma 2.5.11 genügt, gewählt, bis ein zu wählendes Stopkriterium erfüllt ist.

Algorithmus 1 (GenP) Berechnung einer Funktionsquantisierung

Input: Funktion $f \in \mathcal{L}^2(\mathcal{B}([a, b], \mathcal{B}(\mathbb{R})))$,
Anzahl der Niveaumengen $n \in \mathbb{N}$

Output: Funktionsquantisierung von f mit n Quantisierungsstufen

- 1: Initialisierung: Wähle eine beliebige Partition $a = x_1^{(1)} < x_2^{(1)} < \dots < x_{n+1}^{(1)} = b$
- 2: $k \leftarrow 1$
- 3: **while** Ein Stopkriterium ist nicht erfüllt **do**
- 4: **rem** Berechnung der Niveauhöhen $y_i^{(k)}$:
- 5: $y_i^{(k)} := \tilde{\gamma}_f([x_i^{(k)}, x_{i+1}^{(k)}])$, für $i = 1, \dots, n$.
- 6: **rem** Berechnung der Partition:
- 7: Berechne eine Partition $x_1^{(k+1)} < x_2^{(k+1)} < \dots < x_{n+1}^{(k+1)}$ derart, daß
 - i. $x_1^{(k+1)} := a$ und $x_{n+1}^{(k+1)} := b$ und
 - ii. für $i = 2, \dots, n$ $f(x_i^{(k+1)}) = \frac{y_{i-1}^{(k)} + y_i^{(k)}}{2}$ und $\text{sgn}(f'(x_i^{(k+1)})) = \text{sgn}(y_i^{(k)} - y_{i-1}^{(k)})$, gilt, falls ein solches $x_i^{(k+1)}$ unterschiedlich von $x_i^{(k)}$ existiert.
Ansonsten sei $x_i^{(k+1)} := x_i^{(k)}$.
- 8: $k \leftarrow k + 1$
- 9: **end while**
- 10: **return** $\sum_{i=1}^n y_i^{(k+1)} \mathbf{1}_{[x_i^{(k+1)}, x_{i+1}^{(k+1)}]}$

Anmerkungen zum Algorithmus:

Gilt im Algorithmus $x_i^{(k)} = x_{i+1}^{(k)}$ für ein i , so wird $y_i^{(k)}$ auf $f(x_i^{(k)})$ gesetzt.

Ist f nicht differenzierbar in $x_i^{(k+1)}$ für ein i , so wird der Ausdruck $\text{sgn}(f'(x_i^{(k+1)}))$ ersetzt durch -1 bzw. 1 , abhängig davon, ob f in einer Umgebung von $x_i^{(k+1)}$ steigt oder sinkt.

Während die Berechnung der Niveauhöhen in Zeile 5 der Schleife eindeutig ist, trifft dies für die Partition in Zeile 7 im allgemeinen nicht zu. Ein Stopkriterium der **while**-Schleife kann z. B. über einen Schwellwert der Differenzen $|x_i^{(k+1)} - x_i^{(k)}|$ implementiert werden.

Aus den beiden vorangegangenen Abschnitten folgt der Satz 2.5.13.

Satz 2.5.13

Ist $((x_1^{(k)}, \dots, x_{n+1}^{(k)}))_{k \in \mathbb{N}}$ eine Folge von Partitionen, generiert durch den Algorithmus 1 (**GenP**), so gilt:

Die Folge $(\Phi_f(x_2^{(k)}, \dots, x_n^{(k)}))_{k \in \mathbb{N}}$ ist monoton fallend.

2.5.1.2.2 Vereinfachung des Algorithmus für monotone Funktionen

Für monoton steigende Funktionen f vereinfacht sich der Algorithmus 1(**GenP**) auf triviale Weise im Schritt 7 gemäß Lemma 2.5.9. O. B. d. A. sei $f([a, b]) \subseteq \mathbb{R}_+$.

Dann ergibt sich aus dem Algorithmus 1(**GenP**) mit Lemma 2.5.9 der vereinfachte Algorithmus 2 (**FP**).

Algorithmus 2 (FP) Berechnung einer Funktionsquantisierung einer monotonen Funktion

Input: Monoton steigende Funktion $f \in \mathcal{L}^2(\mathcal{B}([a, b], \mathcal{B}(\mathbb{R})))$,
Anzahl der Niveaumengen $n \in \mathbb{N}$

Output: Funktionsquantisierung $\sum_{i=1}^n y_i^{(k+1)} \mathbf{1}_{[x_i^{(k+1)}, x_{i+1}^{(k+1)}[}$ von f mit n Quantisierungs-

stufen

1: Initialisierung: Wähle eine beliebige Partition $a = x_1^{(1)} < x_2^{(1)} < \dots < x_{n+1}^{(1)} = b$

2: $k \leftarrow 1$

3: **while** Ein Stopkriterium ist nicht erfüllt **do**

4: **rem** Berechnung der Niveaus $y_i^{(k)}$:

5: $y_i^{(k)} := \tilde{\gamma}_f([x_i^{(k)}, x_{i+1}^{(k)}])$, für $i = 1, \dots, n$.

6: **rem** Berechnung der Partition

7: Berechne eine Partition $x_1^{(k+1)} < x_2^{(k+1)} < \dots < x_{n+1}^{(k+1)}$, die der Bedingung

$$f(x_i^{(k+1)}) = \frac{y_{i-1}^{(k)} + y_i^{(k)}}{2}$$

genügt

8: $k \leftarrow k + 1$

9: **end while**

10: **return** $\sum_{i=1}^n y_i^{(k+1)} \mathbf{1}_{[x_i^{(k+1)}, x_{i+1}^{(k+1)}[}$

2.5.1.2.3 Monotone Funktionen und zwei Niveaus

Sei wieder $f \in \mathcal{C}([a, b], \mathbb{R})$ monoton steigend und $f([a, b]) \subseteq \mathbb{R}_+$.

Betrachtet wird nun der Fall $n = 2$, d. h. man erhält 2 Niveauhöhen. Außerdem sei f invertierbar. In diesem Kontext ist x_2 der einzige zu bestimmende Parameter im Algorithmus 2 (FP). Aus dem Algorithmus ergibt sich eine Folge von Partitionen mit $x_2^{(k+1)} = g_f(x_2^{(k)})$, wobei $g_f : \mathbb{R} \rightarrow \mathbb{R}$ definiert sei durch

$$g_f(x) := f^{-1} \left(\frac{\tilde{\gamma}_f([a, x]) + \tilde{\gamma}_f([x, b])}{2} \right)$$

Weiter sei dabei $\tilde{\gamma}_f(\{a\}) := a$ und $\tilde{\gamma}_f(\{b\}) := b$ definiert.

Lemma 2.5.14

Sei f zweimal differenzierbar auf $]a, b[$ und nicht notwendig monoton. Dann sind die Funktionen $y_1(x) := \tilde{\gamma}_f([a, x])$ und $y_2(x) := \tilde{\gamma}_f([x, b])$ konvex (bzw. konkav), falls f konvex (bzw. konkav) ist.

Bemerkung 2.5.15

Aus der Konvexität von f folgt im allgemeinen *nicht* die Konvexität von Φ_f .

Satz 2.5.16

Sei f zweimal differenzierbar und streng monoton steigend. Dann ist die oben definierte Funktion g_f konvex (bzw. konkav), wenn f konvex (bzw. konkav) ist.

Lemma 2.5.17 (Spezialisierung des Fixpunktsatzes von Brouwer)

Sei $h : [a, b] \rightarrow [a, b]$ stetig und streng monoton steigend mit $h(a) \geq a$ und $h(b) \leq b$. Weiter sei h entweder konvex oder konkav. Dann hat h genau einen Fixpunkt $\tilde{x} \in [a, b]$, und es gilt:

$$\forall x \in [a, b] : \lim_{n \rightarrow \infty} h^n(x) = \tilde{x}$$

Dabei ist $h^k := \underbrace{h \circ \dots \circ h}_{k\text{-mal}}$.

Lemma 2.5.18

Sei f monoton auf $[a, b]$. Dann ist jeder kritische Punkt $x_{2,0}$ von f bezüglich einer Niveauhöhe y ein Fixpunkt von g_f und umgekehrt.

Korollar 2.5.19

Für konvexe (bzw. konkave), streng monoton steigende und zweimal differenzierbare Funktionen f konvergiert der Algorithmus **FP** gegen eine 2-optimale Funktionsquantisierung.

Ist f wie im Korollar angegeben, aber weder konvex noch konkav, so kann g_f mehrere Fixpunkte haben. Diese kann man in abstoßende und anziehende Fixpunkte unterteilen.

Definition 2.5.20

Sei $h : [a, b] \rightarrow [a, b]$. Ein Fixpunkt \tilde{x} von h heißt *anziehend*, falls es eine Umgebung U von x gibt mit:

$$\forall x \in U : \lim_{n \rightarrow \infty} h^n(x) = \tilde{x}$$

Existiert keine solche Umgebung, so heißt \tilde{x} *abstoßend*.

Satz 2.5.21

Ist f monoton, so liefert jeder anziehende Fixpunkt von g_f ein lokales Minimum von $\|f - \cdot\|_2$. Jeder abstoßende Fixpunkt liefert ein lokales Maximum.

2.5.2 Quantisierung von stückweise konstanten Funktionen

Im Gegensatz zur zu quantisierenden Funktion f des vorangegangenen Kapitels sind die Funktionen, die in der Bildverarbeitung betrachtet werden, meist diskret. Speziell Computerbilder sind auf einem Pixelgitter definiert und haben einen endlichen Wertebereich. Aufgrund dieser Überlegung ist es sinnvoll, die Funktionsquantisierung von stückweise konstanten Funktionen mit endlichem Wertebereich noch einmal gesondert zu betrachten. In [KK99] werden Algorithmen vorgeschlagen, die speziell für dieses Problem formuliert wurden. Dabei unterscheiden die Autoren zwischen Algorithmen, die die optimale Funktionsquantisierung exakt berechnen, und Algorithmen, die die optimale Funktionsquantisierung nur approximieren, dafür aber in der Laufzeit schneller sind.

Sei im folgenden

$$g : \begin{array}{l} [a, b[\rightarrow \mathbb{R} \\ x \mapsto \sum_{i=1}^N v_i \mathbf{1}_{[w_i, w_{i+1}[}(x) \end{array}$$

eine stückweise konstante Funktion mit N Werten $\{v_1, \dots, v_N\}$ und Breakpoints $a = w_1 < \dots < w_{N+1} = b$.

Um die folgenden Funktionsquantisierungsalgorithmen zu verstehen, muß man wissen, daß im Falle einer Funktionsquantisierung von g die Menge der Breakpoints der n -optimalen Funktionsquantisierung s_n eine Teilmenge der Breakpoints von g ist (vgl. Satz 4.2.9 in Abschnitt 4.2.2, (S. 155)).

2.5.2.1 Exakte Funktionsquantisierung von stückweise konstanten Funktionen

Das Funktionsquantisierungsproblem stückweiser konstanter Funktionen läßt sich über die Suche nach dem billigsten Pfad in einem endlichen gerichteten Graphen exakt lösen. Die Knoten des Graphen sind dabei die Breakpoints $K := \{w_1, \dots, w_{N+1}\}$ der zu quantisierenden Funktion g . Es existiert eine gerichtete Kante $\text{arc}(w_i, w_j)$ zwischen w_i und w_j , falls $i < j$ gilt. Die Kosten dieser Kante werden mit $\text{cost}_{w_i, w_j} := \int_{[w_i, w_j[} (g - \tilde{\gamma}_f([w_i, w_j]))^2 d\lambda$ beziffert. Der Graph (K, A, cost) mit $A := \{\text{arc}(w_i, w_j) \mid w_i, w_j \in K, w_i < w_j\}$ läßt sich dann in $\mathcal{O}(N^2)$ Schritten berechnen.

Sei $C_i(k)$ der Preis des billigsten Pfades vom Knoten w_i zum Endknoten $w_{N+1} = b$ mit exakt k Knoten. Dann können alle Werte $C_i(k)$ mit dem Algorithmus 3 (**ExaktQ**), der eine Implementierung des Dijkstra Algorithmus (vgl. z. B. [AHU83]) ist, berechnet werden.

Algorithmus 3 (ExaktQ) Hilfsalgorithmus zur Berechnung der exakten Funktionsquantisierung einer stückweise konstanten Funktion

Input: Stückweise konstante Funktion $f : [a, b] \rightarrow R$ mit N Stufen,
Anzahl der Quantisierungsstufen n

Output: Hilfsstruktur p zur Berechnung der optimalen Funktionsquantisierung

```

1: rem Initialisierung:
2: for  $i \in \{2, \dots, N + 1\}$  do
3:    $C_i(0) := \text{cost}_{w_i, w_{N+1}}$ 
4: end for
5:  $l(0) := \text{cost}_{w_1, w_{N+1}}$ 
6: for  $k = 1$  to  $n - 2$  do
7:   rem Berechnung des billigsten Pfades von  $w_1$  nach  $w_{N+1}$  mit exakt  $k$  Knoten
8:    $l(k) := \min_{i \in \{2, \dots, N+1-k\}} \text{cost}_{w_1, w_i} + C_i(k-1)$ 
9:   Wähle  $p(k, 1) \in \{i \mid \text{cost}_{w_1, w_i} + C_i(k-1) = l(k)\}$ 
10:  rem Update der Werte, um einen weiteren Breakpoints einfügen zu können:
11:  for  $i \in \{2, \dots, N + 1 - k + 1\}$  do
12:     $C_i(k) := \min_{j \in \{i+1, \dots, N+1-k\}} \text{cost}_{w_i, w_j} + C_j(k-1)$ 
13:    Wähle  $p(i, k) \in \{j \mid \text{cost}_{w_i, w_j} + C_j(k-1) = C_i(k)\}$ 
14:  end for
15: end for
16: return  $p$ 

```

Satz 2.5.22

Bei Termination des Algorithmus 3 (**ExaktQ**) gilt in der Notation des Algorithmus:
Ist nun $p^*(n-1) := p(n-1, 1)$ und für $k \in \{1, \dots, n-2\}$ $P^*k := p(p^*(k+1), k)$ definiert,
so ist das folgende richtig:

- i. Setzt man $w_{1,0} := a, w_{2,0} := p^*(n-1), \dots, w_{n,0} := p^*(1), w_{n+1,0} := b$ und

$$s_{n,0} := \sum_{i=1}^n \tilde{\gamma}_f([w_{i,0}, w_{i+1,0}]) \mathbf{1}_{[w_{i,0}, w_{i+1,0}[}$$

so ist $s_{n,0}$ eine n -optimale Funktionsquantisierung von g .

- ii. Es ist $\sqrt{l(n-1)} = \|g - s_n\|_2$.

Beweis:

vgl. [KK99], S. 621

Satz 2.5.23

Der Algorithmus 3 **ExaktQ** hat eine Laufzeitbeschränkung von $\mathcal{O}(N^2)$.

Die Berechnung der n -optimalen Funktionsquantisierung aus dem Ergebnis **ExaktQ** gemäß Satz 2.5.22 erfolgt in $\mathcal{O}(n)$.

Beweis:

vgl. [KK99], S. 621

2.5.2.2 Approximation der Funktionsquantisierung von stückweise konstanten Funktionen

Eine Laufzeit in der Komplexitätsklasse $\mathcal{O}(N^2)$ ist in vielen Anwendungen der Bildverarbeitung, speziell bei der Echtzeitverarbeitung von Videobildern, zu aufwendig. In [KK99] werden mehrere Algorithmen angegeben, die zwar heuristisch sind, aber zum einen eine schnellere worst-case-Laufzeit garantieren und zum anderen eine akzeptable – wenn auch nicht immer die optimale – Lösung liefern.

Eine einfache Idee zur Laufzeitverbesserung ist das Unterteilen des Definitionsbereichs $[a, b]$ in $n - 1$ Intervalle I_1, \dots, I_{n-1} , die jeweils etwa $\frac{N}{n-1}$ Breakpoints der Funktion g enthalten. In diesen Intervallen wird dann eine 2-optimale Funktionsquantisierung von g , eingeschränkt auf das jeweilige Intervall, gesucht. Dies erfolgt mittels des Algorithmus 3 **ExaktQ** in $\mathcal{O}(|I_i \cap \{w_1, \dots, w_N\}|^2)$ für das Intervall I_i . Anschließend werden die auf den Intervallen bestimmten Breakpoints als Breakpoints der Funktionsquantisierung herangezogen. Dieses Vorgehen hat einen Aufwand von $\mathcal{O}(N^2/n)$ (vgl. [KK99], S. 624).

Ein noch schnellerer Algorithmus nimmt n (im allgemeinen nicht adjazente) Segmente der Funktion g als Ausgangsbasis und erweitert diese zu einer Funktionsquantisierung s_n von g (vgl. [KK99], S.624).

2.5.3 Anwendung von Quantisierungsalgorithmen zur Bildsegmentierung

Bisher wurden die Verfahren zur Funktionsquantisierung von Kämpke und Kober noch völlig ohne Bezug zum übergeordneten Thema dieses Kapitels, „Bildsegmentierungen“, vorgestellt. Dieser Abschnitt soll nun die Verbindung zum Thema herstellen.

2.5.3.1 Funktionsquantisierung als direkte Segmentierung

Eine Funktionsquantisierung einer Funktion $f : [a, b] \rightarrow \mathbb{R}$ bezüglich der L^2 -Norm $\|\cdot\|_2$ kann als Approximation dieser Funktion durch eine stückweise konstante Funktion, die daher auch stückweise stetig ist, gesehen werden. Bei eindimensionalen Bildern kann man die Bereiche, auf denen die Funktionsquantisierung des Bildes konstant ist, als Segmente einer Segmentierung definieren. An dieser Stelle wird die Verbindung zu Abschnitt 2.4 deutlich. In [KLM94] geben Koepfler, Lopez und Morel Algorithmen im Rahmen der Methoden der Variationsrechnung zur Bildsegmentierung an, bei denen zur stückweisen stetigen Approximation nur stückweise konstante Funktionen herangezogen werden. Allerdings steht eine Übertragung der wichtigsten Ergebnisse des Abschnitts 2.5.1 auf zweidimensionale Bilder noch aus.

Als Erweiterung dieses Verfahrens auf zwei- bis beliebig dimensionale Bilder kann auch die 2-Niveaupassung gesehen werden, die erstmals in [Don91] erwähnt wurde. In Kapitel 3 wird auf die Analogien der Mehrniveaupassung zur Funktionsquantisierung bezüglich der Norm $\|\cdot\|_2$ eingegangen.

Obwohl eine derartige Ausweitung der Funktionsquantisierung zu deutlich besseren Ergebnissen führt, hat sich in der Literatur nur die folgende Anwendung durchgesetzt.

2.5.3.2 Funktionsquantisierung des Histogramms als Grundlage von Schwellwertverfahren

Bei der Suche nach geeigneten Schwellwerten zur Segmentierung eines Grauwertbildes bietet es sich an, das Histogramm als Hilfe zu nutzen, um Schwellwerte bildabhängig festlegen zu können. Einige Verfahren, die sich auf das Histogramm stützen, wurden bereits in Abschnitt 2.3.1 vorgestellt.

In [KK98] und [KK99] wird – mit leichten Modifikationen – ein Verfahren vorgestellt, das nicht nur das Ergebnis einer Funktionsquantisierung des Histogramms mit n Stufen, sondern auch die „Historie“, d. h. die Zwischenergebnisse der Funktionsquantisierungen mit $m > n$ Stufen, nutzt. Dabei werden die sogenannten „inneren lokalen Minima“ der Funktionsquantisierungen betrachtet.

Definition 2.5.24 (Inneres lokales Minimum)

Ein Segment $[x_i, x_{i+1}[$ heißt *inneres lokales Minimum* einer Treppenfunktion $s_n = \sum_{i=1}^n \alpha_i \mathbf{1}_{[x_i, x_{i+1}[}$, falls $i \notin \{1, n\}$ und $\alpha_i < \alpha_{i+1}$ und $\alpha_i < \alpha_{i-1}$.

$n_{min}(s_n)$ sei die Anzahl der inneren lokalen Minima von s_n .

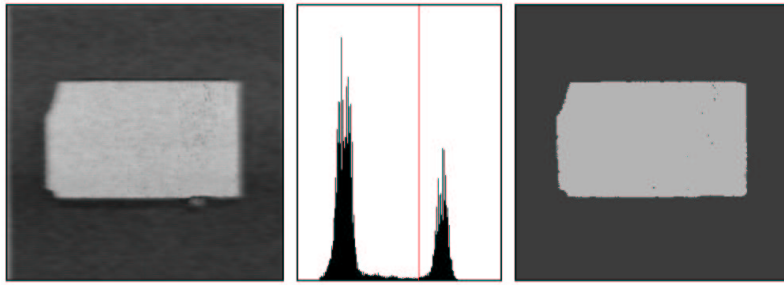


Abbildung 2.14: Beispiel der Bestimmung eines Schwellwertes mit Hilfe des Algorithmus 4 (**SEG**). Im mittleren Bild ist der Schwellwert in das Histogramm rot eingetragen. Das rechte Bild zeigt das Ergebnis der Schwellwertsegmentierung mit dem ermittelten Schwellwert. Man vergleiche dieses Ergebnis mit Bild 2.6 (S. 22).

Der Algorithmus 4 (**SEG**) führt, bei aufsteigendem n , solange n -optimale Funktionsquantisierungen aus, bis eine vorher fest gewählte Anzahl ($n_{cluster} - 1$) von inneren lokalen Minima erreicht wird. Der Name $n_{cluster}$ soll andeuten, daß man nach Beendigung des Algorithmus $n_{cluster}$ Teilstücke des Histogramms erhält.

Algorithmus 4 (SEG) Berechnung von Schwellwerten über die Funktionsquantisierung des Histogramms eines diskreten Grauwertbildes g

Input: Diskretes Grauwertpixelbild g ,
Anzahl der Schwellwerte $n_{cluster}$

Output: $n_{cluster}$ Schwellwerte zur Schwellwertsegmentierung von g

- 1: **rem** Initialisierung
 - 2: $n \leftarrow 2$
 - 3: Setze s_2^0 als die 2-optimale Funktionsquantisierung von h_g , dem Histogramm von g
 - 4: **repeat**
 - 5: $n \leftarrow n + 1$
 - 6: Berechne s_n^0 die optimale Funktionsquantisierung von g
 - 7: **until** $n_{min}(s_n^0) \geq n_{cluster} - 1$
 - 8: **return** $n_{cluster} - 1$ innere lokale Minima von s_n^0
-

Bei der Ausführung dieses Algorithmus ist zu beachten, daß $n_{cluster} < n_{min}(g)$ gelten sollte. Bei einer praktischen Implementierung würde man die Algorithmen 4 (**SEG**) und 3 (**ExaktQ**) verbinden.

Als Schwellwerte für die Segmentierung eines Bildes werden nun Werte in den $n_{cluster} - 1$ inneren lokalen Minima bestimmt. Dies geschieht nach Kämpke und Kober wieder mit Hilfe von inneren lokalen Minima.

Mit Hilfe der $n_{cluster} - 1$ Schwellwerte kann nun das Bild in $n_{cluster}$ Segmente aufgeteilt werden, wobei die Segmente nicht notwendig pixelzusammenhängend sind.

2.5.4 Bewertung

Da die Anwendungen der Funktionsquantisierung zur Segmentierung von Bildern mit zweidimensionalem Definitionsbereich im wesentlichen auf histogrammbasierte Schwellwertverfahren (vgl. Abschnitt 2.3.1) hinauslaufen, sind zur Bewertung zunächst die gleichen Kriterien heranzuziehen. Ein Schwellwertverfahren bietet keinerlei Kontrolle über die Form der Segmente und steht damit im Gegensatz zu den Punkten iii. und iv. der Forderungen von Haralick und Shapiro (vgl. Abschnitt 2.2.1, S. 14, (2.1)) bzw. Punkt iii. der notwendigen Bedingungen einer Segmentierung nach Definition 2.2.1.

Beschränkt man sich jedoch auf Bilder mit eindimensionalem Definitionsbereich, so bietet die Funktionsquantisierung einen Ansatz, dessen Erweiterung direkt zur vektoriellen Mehrniveaupassung führt. Für diese Bilder bietet die Funktionsquantisierung nach Kämpke und Kober nämlich eine Approximation durch stückweise konstante Funktionen. Dabei ist, wie später auch bei der Mehrniveaupassung, die Anzahl der Segmente fest vorgegeben. In diesem Kontext ist die Funktionsquantisierung durchaus als Segmentierung zu bezeichnen.

Das Aussehen der Segmente gehorcht der Zusammenhangsbedingung, genügt also den Punkten iii. und iv. der Beobachtung von Haralick und Shapiro bzw. Punkt iii. der notwendigen Bedingungen einer Segmentierung nach Definition 2.2.1. Weiter sind die Segmente uniform oder homogen bezüglich einer Charakteristik, nämlich dem Funktionswert, (vgl. (2.1) i., S. 14), da eine Funktionsquantisierung den Abstand zur Funktion minimiert. Automatisch ist dadurch auch die Forderung ii. von Haralick und Shapiro erfüllt, da sich adjazente Segmente durch einen unterschiedlichen Durchschnittsfunktionswert auszeichnen. Daher genügt die Funktionsquantisierung in diesem Zusammenhang sowohl den Forderungen von Haralick und Shapiro als auch denen von Mumford und Shah (vgl. (2.2), S. 15) an eine Segmentierung. Auch die Punkte i. – iii. der notwendigen Bedingungen einer Segmentierung nach Definition 2.2.1 (S. 18) sind bei der Funktionsquantisierung von Bildern mit eindimensionalem Definitionsbereich erfüllt. Lediglich die Bedingung iv., die die Anzahl der Segmente als Pareto-Optimum einer Mehrzieloptimierung verlangt, findet keinen Niederschlag in den Arbeiten von Kämpke und Kober.

Viele Ideen dieses Unterkapitels fließen in die vektorielle Mehrniveaupassung mit ein. Zum Beispiel läßt sich die Zusammenhangsbedingung der Segmente auch auf Bilder mit zweidimensionalem Definitionsbereich übertragen. Es sind bei der Übertragung gewisse Forderungen an das System der betrachteten Segmente zu stellen, damit diese den Punkt iii. der Definition 2.2.1 auch für diese Anwendung erfüllen. Es ergibt sich ein weiterer Parameter, nämlich der des Mengenmodells, dessen Konzept es erlaubt, mit Bildern mit ein- als auch mit mehrdimensionalen Definitionsbereich umzugehen.

Ein wesentlicher Unterschied zur vektoriellen Mehrniveaupassung besteht darin, daß sich die Mehrniveaupassung nicht auf die Approximation durch konstante Funktionen auf den Segmenten beschränkt. Um z. B. auch affinen Grau- oder Farbwertverläufen auf den Segmenten gerecht werden zu können, wird mit Funktionen eines endlichen Vektorraums approximiert. Es wird sich aber zeigen, daß einige grundlegende Konzepte der Arbeiten von Kämpke und Kober auch mit diesen Erweiterungen ihre Gültigkeit behalten.

Kapitel 3

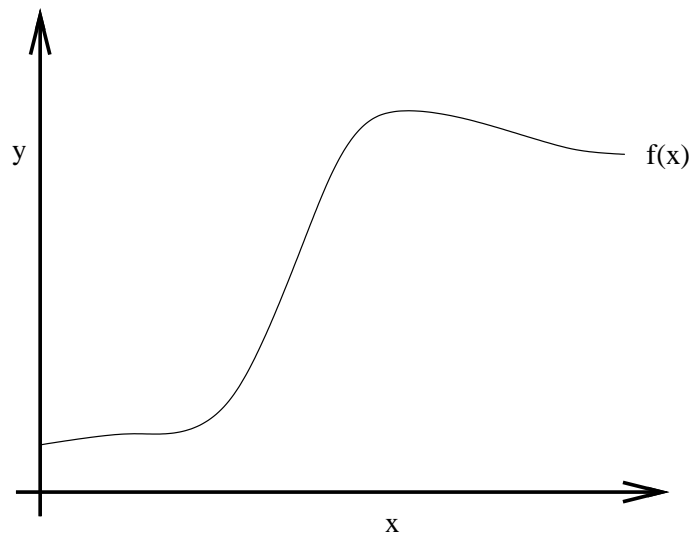
Theoretische Grundlagen der Mehrniveaupassung

3.1 Einleitung

In Abschnitt 2.5.4 wurde die Notwendigkeit, Bedingungen an die Segmente einer Segmentierung heranzutragen zu können, angedeutet. Auch die Annäherungen an den Begriff „Segmentierung“ im Abschnitt 2.2, sowohl von Haralick und Shapiro (vgl. Abschnitt 2.2.1, S. 14, (2.1)), als auch die der vorliegenden Arbeit (vgl. Definition 2.2.1), fordern – indirekt oder direkt – diese Eigenschaft. Im Kontext der Funktionsquantisierung nach Kämpke und Kober in Abschnitt 2.5 ist diese Möglichkeit allerdings nicht vorgesehen. Das Fehlen dieser Möglichkeit hat auch zur Folge, daß die Autoren das Potential, das die Funktionsquantisierung zur direkten Segmentierung von Bildern, d. h. ohne den Umweg über das Histogramm, nicht ausschöpfen (vgl. Abschnitt 2.5.3.1). Die 2-Niveaupassung, erstmals vorgestellt von Klaus Donner in [Don91], schließt diese Lücke und eröffnet so einen neuen Zugang zum Thema Segmentierung. Der Ausgangspunkt zur Entwicklung der 2-Niveaupassung war das Projekt ANGIO (vgl. [Kol94]), in dessen Rahmen ein Verfahren entwickelt wurde, das aus MR-Schnittbildern des Kopfes mit Hilfe eines Gefäßatlases die 3D-Struktur der Kopfadern rekonstruiert. Ein wichtiger Schritt in diesem Verfahren war die Segmentierung der Blutgefäße. Daraus ergab sich eine zwei-Niveau-Situation, bei der das hellere Blutgefäß vom dunkleren Hintergrund getrennt werden sollte.

Gegeben sei eine quadratisch integrierbare Funktion $f : M \rightarrow \mathbb{R}$, wobei $M \subset \mathbb{R}^d$ eine kompakte Menge, im allgemeinen ein rechteckiger Definitionsbereich eines Bildes, sei. In [Don91] wurde ein Verfahren vorgestellt, das ein Bild f mit einer 2-wertigen Funktion $s_2 = \alpha \mathbf{1}_A + \beta \mathbf{1}_{M \setminus A}$ approximiert (vgl. auch Abbildung 3.2). Dabei heißen die Mengen $A = \{s_2 = \alpha\}$ und $M \setminus A = \{s_2 = \beta\}$ *Niveaumengen* und die Werte α und β *Niveauhöhen*. Das Verfahren heißt *2-Niveaupassung*.

Um nun einen Segmentierungsprozeß auf einem Bild, d. h. auf einem zweidimensionalen Definitionsbereich, modellieren zu können, ist es wichtig, daß die Mengen A und B gewissen Bedingungen gehorchen. Man braucht also ein System erlaubter Mengen für den Segmentierungsprozeß. So wurde zum Beispiel im Projekt ANGIO verlangt, daß die Menge

Abbildung 3.1: Funktion $f : M \rightarrow \mathbb{R}$.

mit der höheren Niveauhöhe, also im Bild das hellere Blutgefäß, *kegelsternförmig* sei. Es handelt sich dabei um eine Übertragung des Begriffes „sternförmig“ auf Pixelmengen (vgl. hierzu [Kus92]).

Da die 2-Niveaupassung, insbesondere mit dem Mengenmodell der kegelsternförmigen Mengen, für eine spezielle Anwendung entwickelt wurde, stößt das Konzept bei allgemeinen Segmentierungsproblemen schnell an Grenzen: Will man mehr als ein Objekt segmentieren, mit Farbbildern oder kontinuierlichen Grauwertänderungen im zu segmentierenden Objekt arbeiten, so bietet das klassische Konzept der 2-Niveaupassung kaum Gestaltungsmöglichkeiten.

Um die gegebene Funktion noch besser approximieren zu können, d. h. um den in der Norm des L^2 -Vektorraums gemessenen Abstand von f und der Näherungsfunktion noch mehr zu verringern, kann man die Anzahl der Niveaumengen erhöhen und den Funktionenraum erweitern, mit dem auf den jeweiligen Niveaumengen gepaßt wird, erweitern. Man hat also zwei Ansatzpunkte, um die Theorie der klassischen 2-Niveaupassung nach Donner zu erweitern:

- Erhöhung der Anzahl der Niveaumengen
Dies hat zur Folge, daß man sich tiefere Gedanken über die Bedingungen, denen die Niveaumengen gehorchen sollen, machen muß. Bei zwei Niveaus macht es oft Sinn, ein Vordergrund- und ein Hintergrundniveau zu wählen, um dann nur Forderungen – wie z. B. die Kegelsternförmigkeit – an das Vordergrundniveau zu stellen. Aus dem Vordergrundniveau ergibt sich dann automatisch das Hintergrundniveau. Dies ist bei mehr als zwei Niveaumengen nicht mehr möglich.
- Erhöhung der Dimension des Funktionenraums
Die klassische 2-Niveaupassung ist eine Approximation mit stückweise konstanten Funktionen. Auf einer Niveaumenge wird also die auf diese Menge eingeschränkte

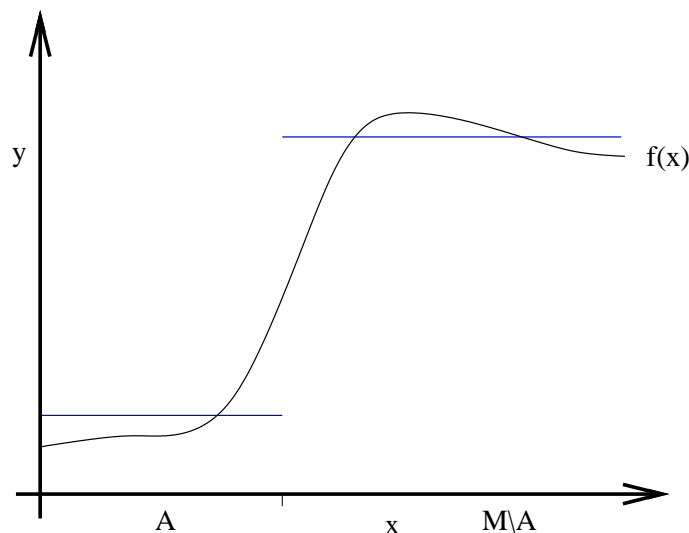


Abbildung 3.2: Approximation von f mit einer 2-wertigen Funktion.

Funktion mit einer konstanten Funktion gepaßt. Nimmt man statt diesem eindimensionalen Raum einen höherdimensionalen Funktionenraum, z. B. einen Polynomraum, so erhöht sich die Exaktheit der Passung, wie man in z. B. Abbildung 3.3 deutlich sieht. Ein erfreulicher Nebeneffekt dieser erweiterten Modellierung ist die Ausdehnung der Niveaupassungsstrategien auch auf mehrkanalige Bilder (etwa Bilder im RGB-Farbmodell). Allerdings muß man in Kauf nehmen, daß einige einfache und algorithmisch leicht umsetzbare Konzepte der 2-Niveaupassung, insbesondere bei der Suche nach lokalen Optima, verallgemeinert werden müssen und dabei in ihrer Darstellung komplizierter erscheinen.

Beziehung zu den Methoden der Variationsrechnung

Auch die in Kapitel 2.4 vorgestellten Konzepte und Problemstellungen motivieren dazu, die bisherigen Ansätze der 2-Niveaupassung zu erweitern, um fließenden Farbübergängen, wie z. B. in Abbildung 3.4, gerecht werden zu können. Allerdings führt die Erhöhung der Anzahl der Niveaumengen und der Dimension des approximierenden Funktionenraums aus mehreren Gründen nicht von der 2-Niveaupassung zu einer Art Mumford-Shah-Modell: Die Anzahl der Niveaumengen wird ein Parameter der Segmentierung von Bildern mit vektorieller Mehrniveaupassung. Dies ist ein deutlicher Unterschied zum Ansatz der Segmentierung mit Methoden der Variationsrechnung, wo sich die Anzahl der Segmente aus der Energieminimierung ergibt und nur indirekt über die Parameter α und β steuerbar ist (vgl. Abschnitt 2.4.1, insbesondere S. 31).

Auch die Erhöhung der Dimension des approximierenden Funktionenraums wird nicht zu einer derart allgemeinen Formulierung eines Approximationsproblem führen, wie es im Falle der Variationsmethoden der Fall ist. Statt mit beliebigen differenzierbaren Funktionen, wie im Falle des Ansatzes von Mumford und Shah, wird der Raum der Funktionen,

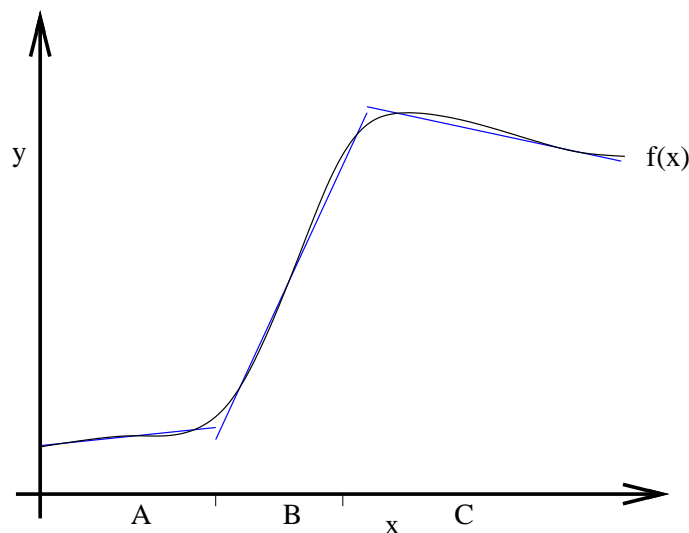


Abbildung 3.3: Approximation von f mit einer stückweise affinen Funktion auf drei Niveaumengen.

mit dem die vektorielle Mehrniveaupassung approximiert, immer ein endlich dimensionaler Vektorraum bleiben.

Schließlich wird auch das Konzept der 2-Niveaupassung, Bedingungen an die Niveaumengen heranzutragen, nicht aufgegeben. Auch dies ist ein deutlicher Unterschied zum Mumford-Shah-Modell. Der Vorteil dieses Konzepts ergibt sich jedoch direkt aus der Definition 2.2.1 der notwendigen Bedingungen einer Segmentierung auf Seite 18 dieser Arbeit, die im Punkt iii. verlangt, daß das Aussehen der Segmente reglementierbar sein sollte.

Übersicht des Kapitels

Zunächst einmal ist im folgenden zu klären, wie Bedingungen an das Aussehen der Segmente zu handhaben sind und wie man die Dimension des Raums der approximierenden Funktionen sinnvoll erhöht, um das Segmentierungsproblem als Mehrniveaupassungsproblem zu formulieren. Daher gliedert sich das Kapitel wie folgt:

Im zweiten Unterkapitel wird ein System erlaubter Mengen vorgestellt, aus dem später die Segmente einer Segmentierung gewählt werden müssen.

Im dritten Unterkapitel wird das System der Segmentierungen durch k erlaubte Mengen eingeführt, das, wie der Name schon sagt, k Mengen aus dem System der erlaubten Mengen als Segmente einer Segmentierung zusammenfaßt. Dabei ist zu beachten, daß – im Gegensatz zur 2-Niveaupassung nach Donner ([Don91]) – alle Segmente, d. h. nicht nur das Vordergrundniveau, die Bedingungen, die das System der erlaubten Mengen fordert, erfüllen. Ein wichtiges Ergebnis dieses Unterkapitels wird sein, daß aus der Abgeschlossenheit des Systems der erlaubten Mengen die Abgeschlossenheit des Systems der Segmentierungen durch k erlaubte Mengen innerhalb des System aller Segmentierungen folgt.



Originalbild



Approximation mit stückweise konstanten Funktionen

Abbildung 3.4: Beispiel eines Segmentierungsproblems, bei dem eine affin lineare Approximation auf mehreren Niveaumengen sinnvoll ist. Insbesondere die Oberfläche des Buchstabens zeigt einen affinen Farbverlauf. Die Approximation mit stückweise konstanten Funktionen führt zu keinem zufriedenstellenden Ergebnis (vgl. auch Abbildung 5.8, S. 234).

Das vierte Unterkapitel widmet sich dem Vektorraum der approximierenden Funktionen. Es handelt sich dabei um einen endlich dimensionalen Vektorraum mit der Eigenschaft, daß die Einschränkung der Funktionen des Raums auf eine erlaubte Menge mit positivem Lebesgue-Maß nicht zu einem Dimensionsverlust des Vektorraums führt. Über die Definition einer erlaubten Familie von Skalarprodukten, derart, daß zu jeder erlaubten Menge ein Skalarprodukt existiert, wird die Grundlage des Approximationsproblems der vektoriellen Mehrniveaupassung gelegt.

Dieses Approximationsproblem wird dann im fünften Unterkapitel formuliert. Dabei wird – auch aus historischen Gründen – zunächst einmal das Problem für die 2-Niveaupassung neu formuliert. Es wird sich allerdings zeigen, daß die Übertragung auf die Mehrniveausituation völlig unproblematisch ist.

Im sechsten Unterkapitel wird die Existenz einer optimalen Segmentierung innerhalb der Theorie der vektoriellen Mehrniveaupassung zur Bildsegmentierung nachgewiesen. Die Theorie der Mehrniveaupassung ist also in sich schlüssig.

Das siebte und letzte Unterkapitel der theoretischen Grundlagen der Mehrniveaupassung zur Bildsegmentierung versucht, einige Ergebnisse des Funktionsquantisierungsansatzes von Kämpke und Kober (vgl. Abschnitt 2.5) zur Berechnung lokaler Optima auf die vektorielle Mehrniveaupassung zu übertragen. Dies wird zwar nur für die 2-Niveausituation durchgeführt, ist jedoch für die praktische Umsetzung, die fast immer auf iterierte 2-Niveaupassungen hinausläuft, von zentralem Interesse.

Für die Praxistauglichkeit des Konzepts sei auf das Kapitel 4 verwiesen.

3.2 Das Mengensystem

Das Konzept eines Systems von erlaubten Mengen ist einer der wesentlichen Unterschiede der vektoriiellen Mehrniveaupassung zum Ansatz von Mumford und Shah, mit Methoden der Variationsrechnung Bilder zu segmentieren (vgl. Abschnitt 2.4). Das System der erlaubten Mengen bestimmt dabei, welche Segmentierungen erlaubt sind und welche nicht. Bei einer Segmentierung eines Bildes in zwei „Bereiche“ wird man z. B. nur zusammenhängende Mengen verwenden wollen. Viele konkrete Anwendungen benötigen noch speziellere Mengensysteme (vgl. z. B. [Kus92]). Die Auswahl eines geeigneten Mengensystems ist ein problemabhängiger Parameter der Mehrniveaupassung.

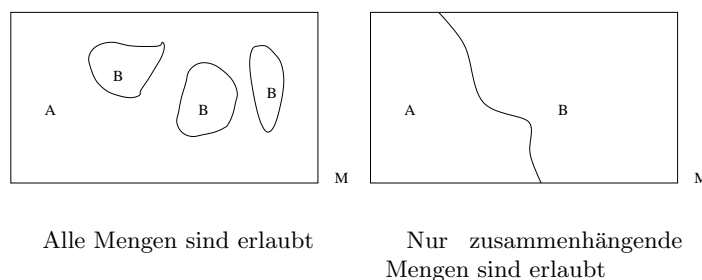


Abbildung 3.5: Segmentierungen mit jeweils zwei Segmenten aus unterschiedlichen Systemen erlaubter Mengen.

Das zentrale Anliegen dieses Abschnitts ist es, eine geeignete Metrik auf dem eingeschränkten Mengensystem zu definieren, die in Abschnitt 3.3 zu einer Metrik auf dem System der Segmentierungen durch k erlaubte Mengen erweiterbar ist. Eine Metrik ist nötig, um Stetigkeiten von Funktionen, die auf Mengen definiert sind, formulieren zu können.

3.2.1 Der metrische Raum (\mathcal{A}, δ)

Im folgenden werden Funktionen betrachtet, die auf Mengen definiert sind. Um Begriffe wie „lokales Minimum“ oder „Stetigkeit“ festzulegen, ist ein topologischer Mengenraum zu definieren. Klassische Ansätze, z. B. über den Hausdorff-Abstand, wie er etwa in der Theorie der Fraktale verwandt wird (vgl. z. B. [Edg98]), scheiden aus, da sie das Problem der λ -Nullmengen nicht beheben. Es ist nämlich für alle Lebesgue-integrierbaren Funktionen f : $\int_A f d\lambda = \int_B f d\lambda = 0$, falls $\lambda(A) = \lambda(B) = 0$ ist, obwohl die Hausdorffdistanz dieser Mengen beliebig sein kann. Bei der Segmentierung von Bildern ist eine solche Topologie ungeeignet, sollte es doch für zwei Mengen, deren Abstand größer als 0 ist, Funktionen (d. h. Bilder) geben, auf denen sich ihre Integrale unterscheiden.

In [Don91] führt Klaus Donner eine solche Topologie über einen metrischen Raum ein. Sie basiert darauf, daß Mengen A_1, A_2 mit $\lambda(A_1 \setminus A_2) = 0 = \lambda(A_2 \setminus A_1)$ als Vertreter derselben Äquivalenzklasse identifiziert werden. Es wird nicht eine Metrik auf einem Mengenraum eingeführt, sondern eine Halbmetrik, die durch den Übergang zum Quotientenraum bezüglich der Äquivalenzrelation, die durch diese Halbmetrik definiert ist, zur Metrik wird.

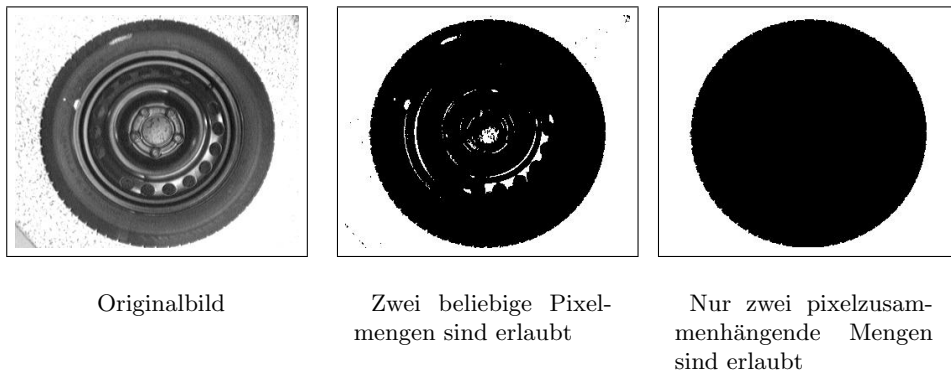


Abbildung 3.6: Das Ergebnis einer 2-Niveaupassung ist abhängig von der Wahl des Systems der erlaubten Mengen: Will man etwa das Rad im Originalbild vom Hintergrund trennen, so bietet es sich an, nur Segmentierungen durch zwei pixelzusammenhängende Mengen zu erlauben.

Mit leichten Modifikationen, besonders bei der Definition 3.2.10 in Abschnitt 3.2.2, wird der Ansatz, der in [Don91] beschrieben wird, übernommen.

Betrachtet wird im folgenden eine Teilmenge \mathcal{A} der Borelschen σ -Algebra $\mathcal{B}(M)$ von M .

Definition 3.2.1

Sei $d : \mathcal{B}(M) \times \mathcal{B}(M) \rightarrow \mathbb{R}_+$ definiert durch

$$d(A_1, A_2) := \lambda((A_1 \setminus A_2) \uplus (A_2 \setminus A_1))$$

für Mengen $A_1, A_2 \in \mathcal{B}(M)$.

Weiter sei \sim eine Relation auf $\mathcal{B}(M)$ definiert durch

$$A_1 \sim A_2 \Leftrightarrow d(A_1, A_2) = 0$$

Bemerkung 3.2.2

Seien $A, B \in \mathcal{B}(M)$. Einige nützliche äquivalente Umformungen von d , die durch einfache Mengenlogik und Eigenschaften des Lebesguemaßes gefolgert werden können, sind:

- i. $d(A, B) = \lambda(A \setminus B) \uplus \lambda(B \setminus A) = \lambda(A \setminus B) + \lambda(B \setminus A)$
- ii. $d(A, B) = \lambda((A \cup B) \setminus (A \cap B)) = \lambda(A \cup B) - \lambda(A \cap B)$

Lemma 3.2.3

d ist eine Halbmetrik¹ und \sim ist eine Äquivalenzrelation auf $\mathcal{B}(M)$.

Beweis:

d ist Halbmetrik:

Die Symmetrie ist trivial zu zeigen. Die Dreiecksungleichung folgt aus der Isotonie des Lebesguemaßes. Seien $A, B, C \in \mathcal{A}$. Dann gilt:

$$\begin{aligned}
d(A, B) + d(B, C) &= \lambda(A \setminus B \cup B \setminus A) + \lambda(B \setminus C \cup C \setminus B) \\
&\geq \lambda(A \setminus B \cup B \setminus A \cup B \setminus C \cup C \setminus B) \\
&\geq \lambda(A \setminus (B \cup C) \cup (B \cap C) \setminus A \cup (A \cap B) \setminus C \cup C \setminus (A \cup B)) \\
&= \lambda((A \setminus (B \cup C) \cup (A \cap B) \setminus C) \cup ((B \cap C) \setminus A \cup C \setminus (A \cup B))) \\
&= \lambda(((A \setminus B) \setminus C \cup (A \cap B) \setminus C) \cup ((B \setminus A \cap C \setminus A) \cup (C \setminus A \cap C \setminus B))) \\
&= \lambda((A \setminus C) \cup (C \setminus A \cap (C \setminus B \cup B \setminus A))) \\
&= \lambda((A \setminus C) \cup (C \setminus A)) \\
&= d(A, C)
\end{aligned}$$

\sim ist Äquivalenzrelation:

Reflexivität und Symmetrie sind trivial zu zeigen. Zum Beweis der Transitivität gelte $A \sim B$ und $B \sim C$. Dann ist $d(A, B) = 0$ und $d(B, C) = 0$. Dann folgt aber mit der Dreiecksungleichung der Halbmetrik d , daß $d(A, C) = 0$ gelten muß. Daher ist nach Definition $A \sim C$. □

Bemerkung 3.2.4

Man beachte, daß $d : \mathcal{A} \times \mathcal{A} \rightarrow \mathbb{R}_+$ keine Metrik ist, sondern nur eine Halbmetrik. Denn beispielsweise gilt $d([0, 1]^n,]0, 1[^n) = 0$, aber nicht $]0, 1[^n =]0, 1[^n$. Um dennoch einen metrischen Raum zu erhalten, kann man nun die Äquivalenzrelation \sim aus Definition 3.2.1 nutzen, um einen Quotientenraum zu definieren, auf dem d eine Metrik ist.

Definition 3.2.5

Für eine Teilmenge $\mathcal{A} \subset \mathcal{B}(M)$ ist mit Lemma 3.2.3 die Relation \sim eine Äquivalenzrelation auf \mathcal{A} . Daher ist die Definition

$$\mathfrak{A} := \mathcal{A} / \sim$$

wohldefiniert. O sei im Quotientenraum \mathfrak{A} die Äquivalenzklasse der leeren Menge. Es ist also $O = \{A \in \mathcal{A} \mid \lambda(A) = 0\}$.

Weiter sei $\mathfrak{d} : \mathfrak{A} \times \mathfrak{A} \rightarrow \mathbb{R}_+$ die durch d induzierte Metrik auf \mathfrak{A} . Dadurch ist der Quotientenraum $(\mathfrak{A}, \mathfrak{d})$ ein metrischer Raum, da nun $\mathfrak{d}(A, B) = 0 \Rightarrow A = B$ gilt. Im folgenden sei \mathfrak{A} stets auf diese Weise aus einem Mengensystem $\mathcal{A} \subseteq \mathcal{B}(M)$ entstanden.

¹Eine Halbmetrik erfüllt die Eigenschaften einer Metrik, bis auf die Bedingung $\forall A, B \in \mathcal{B}(M) : d(A, B) = 0 \Leftrightarrow A = B$

Bemerkung 3.2.6 (Quotientenraum \mathfrak{A})

Sei $\mathcal{K}(M)$ die Menge aller kompakten Teilmengen von M . Dann ist $\mathcal{K}(M)$ bezüglich der Hausdorff-Metrik ein vollständiger metrischer Raum (vgl. z. B. [Bar88], S. 37). Diese Tatsache kann jedoch im Zusammenhang mit Segmentierungen nur schlecht genutzt werden. Denn es existiert keine paarweise disjunkte Zerlegung von M durch Mengen aus $\mathcal{K}(M)$, außer der trivialen Zerlegung von M durch eine Menge, nämlich M selbst.

Der Übergang von \mathcal{A} zum Äquivalenzklassenmengensystem \mathfrak{A} und die Tatsache, daß Mengen, die sich nur um Nullmengen unterscheiden, für alle integrierbaren Funktion das gleiche Integral haben, erleichtert die Notation von Segmentierungen wesentlich. Denn eine Unterscheidung zwischen $]0, 1[$, $[0, 1[$, $]0, 1]$ und $[0, 1]$ ist im Quotientenraum \mathfrak{A} obsolet. Der Ausdruck $f = \mathbf{1}_{]0,1[}1 + \mathbf{1}_{[1,2]}2$ definiert daher eine Funktion aus $L^2(\mathcal{B}([0, 2]), \mathcal{B}(\mathbb{R}))$, die an der Stelle 1 nicht ausgewertet werden kann (und im folgenden auch nicht muß).

Bemerkung 3.2.7 (Übertragung der Mengenoperationen)

Da alle Mengen einer Äquivalenzklasse das gleiche Lebesguemaß haben, kann man für eine Äquivalenzklasse $A_0 \in \mathfrak{A}$ das Maß der Äquivalenzklasse als das Maß eines Repräsentanten, also $\lambda(A_0) := \lambda(A)$ für ein $A \in A_0$, definieren.

Weiter sei für Äquivalenzklassen $A_0, B_0 \in \mathfrak{A}$:

$$A_0 \sqsubset B_0 \quad :\Leftrightarrow \quad (A_0 \neq B_0) \wedge \exists A, B \in \mathcal{A} : (A \in A_0, B \in B_0, A \subset B)$$

Für den Schnitt von Mengen gilt

$$A_0 \sqcap B_0 = C_0 \quad :\Leftrightarrow \quad \exists A, B, C \in \mathcal{A} : (A \in A_0, B \in B_0, C \in C_0, A \cap B = C)$$

Man beachte, daß es im Fall $A_0 \sqcap B_0 = O$ durchaus Repräsentanten $A \in A_0, B \in B_0$ geben kann, für die $A \cap B = C \neq \emptyset$ ist. Dann gilt aber $\lambda(C) = 0$. Damit ist C in O , der Äquivalenzklasse der leeren Menge, enthalten. Ähnlich werden die Mengenoperationen \cup, \cap und \setminus auf \mathfrak{A} übertragen.

Eine Äquivalenzklasse heiße *zusammenhängend*, falls alle Elemente A der Äquivalenzklasse, die aus mehreren Zusammenhangskomponenten bestehen, nur eine Zusammenhangskomponente mit positivem Lebesgue-Maß haben. Dabei heiße eine Menge $Z \subset A$ *Zusammenhangskomponente* von A , falls Z zusammenhängend ist und keine zusammenhängende Menge $B \subset A$ existiert mit $Z \subsetneq B \subset A$.

Im folgenden wird – eigentlich falsch – von „Mengen“ A aus \mathfrak{A} gesprochen werden. Man vergleiche zu dieser verkürzenden, legeren Sprechweise die allgemein in der Mathematik übliche Sprechweise von „Funktionen“ $f \in L^p(\mathcal{B}(M), \mathcal{B}(\mathbb{R}))$, obwohl f in diesem Fall eine Äquivalenzklasse bezeichnet und gar nicht punktweise ausgewertet werden kann. Dementsprechend wird auch die Notation der Mengenoperationen \cap, \cup, \subset für die Äquivalenzklassen beibehalten. Im folgenden wird \mathfrak{A} auch als Mengensystem bezeichnet, obwohl Quotientenraum eines Mengensystems die korrekte Bezeichnung wäre.

Bemerkung 3.2.8 (Translationseigenschaften der Metrik \mathfrak{d})

Offensichtlich ist die Metrik \mathfrak{d} translationsinvariant. Denn es gilt für alle $A, B \in \mathfrak{A}$ und $t \in M$ mit $A + t \subset M$ und $B + t \subset M$: $\mathfrak{d}(A, B) = \mathfrak{d}(A + t, B + t)$.

Zusätzlich zu dieser Translationsinvarianz besitzt \mathfrak{d} noch eine bedingte einseitige Translationsinvarianz:

Es sei $M \subset \mathbb{R}^m$ kompakt. Weiter seien vier Mengen $A_1, A_2, B_1, B_2 \in \mathfrak{A}$ gegeben, wobei $\lambda(A_1) = \lambda(A_2)$ und $\lambda(B_1) = \lambda(B_2)$ sei. Dann gilt:

$$\lambda(A_1 \cap B_1) > \lambda(A_2 \cap B_2) \Rightarrow \mathfrak{d}(A_1, B_1) < \mathfrak{d}(A_2, B_2)$$

Gilt jedoch $\lambda(A_1 \cap B_1) = 0$ und $\lambda(A_2 \cap B_2) = 0$, so ist $\mathfrak{d}(A_1, B_1) = \mathfrak{d}(A_2, B_2)$ unabhängig von der Entfernung der Mengen im Sinne eines Hausdorff-Abstands oder der Lage der Mengen.

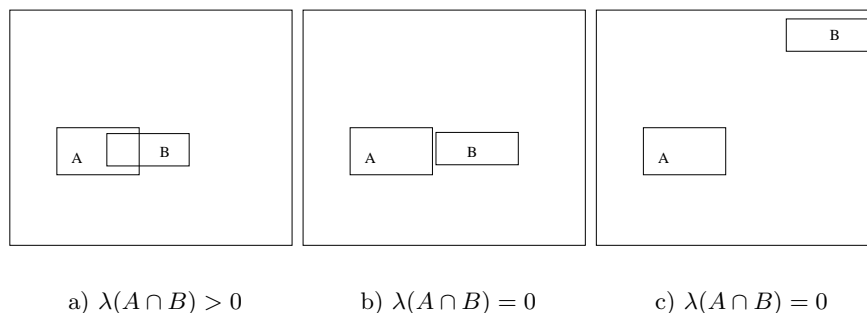


Abbildung 3.7: Bedingte einseitige Translationsinvarianz der Metrik \mathfrak{d} : Sobald zwei Mengen einen Schnitt haben, der einer λ -Nullmenge entspricht, hat die Lage dieser beiden Mengen keine Auswirkung auf ihren Abstand in der Metrik \mathfrak{d} . Der Abstand der Mengen A und B in Skizze b) und c) ist gleich und größer als der in Skizze a).

Es gilt also für alle $A, B \in \mathfrak{A}$ mit $\lambda(A) > 0, \lambda(B) > 0$ und $\lambda(A \cap B) = 0$:

$$\forall t \in M, B + t \in \mathfrak{A} : \lambda(A \cap (B + t)) = 0 \Rightarrow \mathfrak{d}(A, (B + t)) = \mathfrak{d}(A, B)$$

Man vergleiche dazu auch Abbildung 3.7.

Bemerkung 3.2.9 (Größenabhängigkeit der Metrik \mathfrak{d})

Der Abstand, gemessen in der Metrik \mathfrak{d} , ist abhängig von der Größe der beteiligten Mengen, da diese Metrik nur über das Lebesguemaß der Differenz definiert ist. Abbildung 3.8 veranschaulicht den Effekt, der daraus entsteht. Die Tatsache, daß Abstände von der Größe der beteiligten Mengen abhängig sind, wird insbesondere bei der Diskretisierung der Mengen in Abschnitt 4.2 Schwierigkeiten bereiten (vgl. Bemerkung 4.2.16, S. 159).

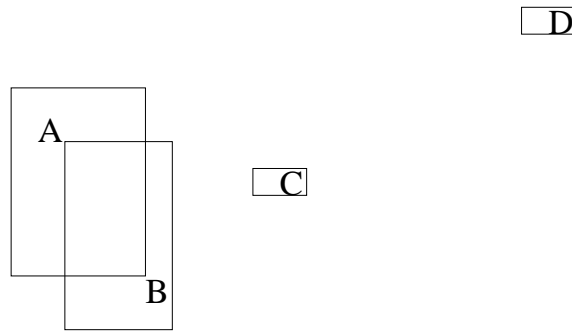


Abbildung 3.8: Der Abstand gemessen in der Metrik \mathfrak{d} ist abhängig von der Größe der verglichenen Mengen: $\mathfrak{d}(A, B)$ ist größer als $\mathfrak{d}(C, D)$.

3.2.2 Bedingungen an das Mengensystem \mathfrak{A}

Für die späteren Beweise ist es nötig, daß \mathfrak{A} die Bedingungen der Definition 3.2.10 erfüllt.

Definition 3.2.10

Erfüllt ein Mengensystem $\mathfrak{A} = \mathcal{A}/\sim$ die Eigenschaften

(OM) $O \in \mathfrak{A}$ und $M \in \mathfrak{A}$, und

(A) \mathfrak{A} ist abgeschlossen in $(\mathcal{B}(M)/\sim, \mathfrak{d})$,

so heie \mathfrak{A} ein *erlaubtes Mengensystem*.

Bemerkung 3.2.11

Aus der Eigenschaft (OM) folgt, da das erlaubte Mengensystem \mathfrak{A} nicht leer ist. Zustzlich kann man noch eine weitere Eigenschaft fr das erlaubte Mengensystem fordern:

(DE) fr alle $A, B \in \mathfrak{A}$ mit $A \subsetneq B$ gilt:

$$\lambda(B) = \sup_{C \in \mathfrak{A}, A \subsetneq C \subsetneq B} \lambda(C) \quad \text{und} \quad \lambda(A) = \inf_{C \in \mathfrak{A}, A \subsetneq C \subsetneq B} \lambda(C)$$

Dabei heie die Eigenschaft (DE) *Dichteigenschaft*. Fr ein erlaubtes Mengensystem mit Dichteigenschaft gilt: Ist $M \neq O$, so ist $|\mathfrak{A}| = \infty$. Denn es gilt: $\forall A, B \in \mathfrak{A} : \exists C \in \mathfrak{A} \setminus \{A, B\} : A \subsetneq C \wedge C \subsetneq B$.

Die Dichteigenschaft spielt fr die spteren Beweise keine Rolle. Dennoch ist es im Sinne eines Segmentierungsproblems, wenn die erlaubten Segmente die Dichteigenschaft erfllen. Ist \mathfrak{A} beispielsweise ein Menge von Rechtecken, so sollte zwischen zwei Rechtecken A, B mit $A \subset B$ immer noch ein Rechteck C liegen, da auch im erlaubten Mengensystem liegt.

Beispiel 3.2.12

Sei $\mathcal{A} := \{[0, t] \mid t \in [0, 1]\}$. Beim Übergang zur Quotientenmenge stellt man fest, daß jede Äquivalenzklasse nur einen Repräsentanten hat. Daher ist es legitim, $\mathfrak{A} = \{[0, t] \mid t \in [0, 1]\}$ zu schreiben. Die Bedingungen (0M), (A) und (DE) der Definition 3.2.10 sind, wie man leicht nachweist, erfüllt. Man kann sogar zeigen, daß \mathfrak{A} zu $[0, 1]$ homöomorph ist. Auch für $\mathcal{A}' := \{[0, t[,]0, t],]0, t[\mid t \in [0, 1]\}$ gilt: $\mathfrak{A}' := \mathcal{A}' / \sim = \mathfrak{A}$.

3.3 Das System der Segmentierungen

Das Mengensystem der erlaubten Mengen nach Definition 3.2.10 ermöglicht es, Bedingungen an die einzelnen Segmente einer Segmentierung heranzutragen. Es bietet sich z. B. an für Segmentierungsprobleme ein erlaubtes Mengensystem von zusammenhängenden Mengen zu benutzen (vgl. auch Abschnitt 2.2, insbesondere (2.1)). Vorsicht ist allerdings beim Zusammensetzen der einzelnen Mengen aus \mathfrak{A} zu einer disjunkten Zerlegung von M geboten: Ist A, B mit $A \cup B = M, A \cap B = O$ eine Zerlegung der Menge M , so ist $B = M \setminus A$, aber aus dem Zusammenhang von A folgt nicht der Zusammenhang von $M \setminus A = B$. Daher ist eine genauere Definition des Begriffes „Segmentierung durch erlaubte Mengen“ erforderlich.

Definition 3.3.1 (System der Segmentierungen)

Sei \mathfrak{A} ein erlaubtes Mengensystem, $A \in \mathfrak{A}$ und $k \in \mathbb{N}$. Dann sei $\mathfrak{S}_k(\mathfrak{A}, A)$ definiert durch

$$\mathfrak{S}_k(\mathfrak{A}, A) := \left\{ \mathbf{B} \in \wp(\mathfrak{A}) \mid \biguplus_{B \in \mathbf{B}} B = A, |\mathbf{B}| \leq k, |\mathbf{B} \setminus \{O\}| < k \Rightarrow O \in \mathbf{B} \right\}$$

$\mathfrak{S}_k(\mathfrak{A}, A)$ heie das *System der Segmentierungen von A durch k erlaubte Mengen*. Eine Menge heie dabei *erlaubt*, wenn sie im Mengensystem \mathfrak{A} enthalten ist.

Ist $\mathbf{B} \in \mathfrak{S}_k(\mathfrak{A}, A)$, so heien die Mengen $B \in \mathbf{B}$ die *Segmente* oder *Niveaumengen* der Segmentierung \mathbf{B} . Ist $B \neq O$, so heie B auch *echtes Segment*. O heie *das leere Segment*.

Meist ist $A = M$. In diesem Fall wird $\mathfrak{S}_k(\mathfrak{A}) := \mathfrak{S}_k(\mathfrak{A}, M)$ notiert. Bis auf den Abschnitt 3.7.1 wird dies im folgenden fast immer der Fall sein.

Bemerkung 3.3.2

- a) Bei $\mathfrak{S}_k(\mathfrak{A})$ handelt es sich um die Menge der disjunkten Zerlegungen von M durch hchstens k Mengen aus \mathfrak{A} . Enthlt eine Segmentierung weniger als k echte Segmente, so ist das leere Segment O in der Segmentierung enthalten. Diese Vorschrift mute gemacht werden, um die Abgeschlossenheit des Systems der Segmentierungen garantieren zu knnen (vgl. Satz 3.3.13): Konvergiert in einer konvergenten Folge von Segmentierungen eine Folge von Segmenten gegen das leere Segment, so mu dieses auch in der Zielsegmentierung enthalten sein, wenn das System abgeschlossen sein soll. Da aber O nicht die leere Menge \emptyset ist, sondern eine quivalenzklasse, mute die Definition 3.3.1 diese etwas technische Vorschrift enthalten. Man beachte, da aus der Definition folgt: $\forall \mathbf{A} \in \mathfrak{S}_k(\mathfrak{A}) : O \notin \mathbf{A} \Rightarrow |\mathbf{A}| = k$.

Statt einer Tupelschreibweise wurde die Mengenschreibweise der Zerlegungen gewhlt, um Identittsprobleme bei gleichen Zerlegungen in verschiedener Reihenfolge zu vermeiden.

- b) Gelegentlich wird in dieser Arbeit auch von einer „Segmentierung einer Funktion $f : M \rightarrow \mathbb{R}^d$ “ die Rede sein. Gemeint ist dann immer die Segmentierung des Definitionsbereichs M der Funktion f .

Beispiel 3.3.3

Für $\mathfrak{A} := \{[0, t] \mid t \in [0, 1]\}$ – wie in Beispiel 3.2.12 – ist $\mathfrak{S}_2(\mathfrak{A}) = \{\{[0, 1], O\}\}$. Dieses System der Segmentierungen ist natürlich akademisch.

Ist $\mathfrak{A}' := \{[x_1, x_2] \mid x_1, x_2 \in [0, 1]\}$, so ergibt sich

$$\mathfrak{S}_2(\mathfrak{A}') = \{\{[0, t], [t, 1]\} \mid t \in [0, 1]\} = \{\{A, [0, 1] \setminus A\} \mid A \in \mathfrak{A}\}$$

Man beachte bei diesem Beispiel, daß $[0, t]$ und $[t, 1]$ nur Repräsentanten ihrer Äquivalenzklasse sind und im Quotientenraum \mathfrak{A}' durchaus $[0, t] \cap [t, 1] = O$ gilt.

Bemerkung 3.3.4

Um Abstände zwischen Elementen aus $\mathfrak{S}_k(\mathfrak{A})$ beschreiben zu können, wird eine von \mathfrak{d} induzierte Metrik auf $\mathfrak{S}_k(\mathfrak{A})$ definiert.

Definition 3.3.5 (Metrik der Segmentierungen)

Auf $\wp_k(\mathcal{B}(M)/\sim)$ sei τ wie folgt definiert:

$$\tau : (\mathbf{A}, \mathbf{B}) \mapsto \sum_{A \in \mathbf{A}} \min_{B \in \mathbf{B}} \mathfrak{d}(A, B)$$

Mit dieser Definition sei

$$\mathfrak{d}' : \wp_k(\mathcal{B}(M)/\sim) \times \wp_k(\mathcal{B}(M)/\sim) \rightarrow \mathbb{R}_+$$

$$(\mathbf{A}, \mathbf{B}) \mapsto \max\{\tau(\mathbf{A}, \mathbf{B}), \tau(\mathbf{B}, \mathbf{A})\}$$

definiert.

Die Einschränkung von \mathfrak{d}' auf $\mathfrak{S}_k(\mathfrak{A}) \times \mathfrak{S}_k(\mathfrak{A})$ heißt die *Metrik der Segmentierungen* $\mathfrak{S}_k(\mathfrak{A})$.

Lemma 3.3.6

\mathfrak{d}' ist eine Metrik auf $\wp_k(\mathcal{B}(M)/\sim)$

Beweis:

Die Symmetrie ist offensichtlich und die Dreiecksungleichung ergibt sich aus den Metrikeigenschaften von \mathfrak{d} . Auch ist offensichtlich, daß \mathfrak{d}' zwei identischen Segmentierungen den Abstand null zuordnet. Seien also $\mathbf{A}, \mathbf{B} \in \wp_k(\mathcal{B}(M)/\sim)$ mit $\mathfrak{d}(\mathbf{A}, \mathbf{B}) = 0$ gegeben. Dann gilt jedoch mit Definition 3.3.5:

$$\forall A \in \mathbf{A} : \exists B \in \mathbf{B} : d(A, B) = 0 \Rightarrow A = B$$

und

$$\forall B \in \mathbf{B} : \exists A \in \mathbf{A} : B = A$$

Daher folgt $\mathbf{A} = \mathbf{B}$.

□

Bemerkung 3.3.7

Für $k = 1$ ist \mathfrak{d}' auf $\mathfrak{S}_1(\mathfrak{A})$ genau die Metrik \mathfrak{d} . Natürlich ist das System der Segmentierungen von M durch eine Menge trivial: Es ist $\mathfrak{S}_1(\mathfrak{A}) = \{\{M\}\}$. Allerdings kann man im Fall $k = 1$ die Metrik \mathfrak{d}' kanonisch zur Metrik \mathfrak{d} auf \mathfrak{A} fortsetzen. Daher wird im folgenden \mathfrak{d}' mit \mathfrak{d} notiert. Die genaue Definition sollte immer aus dem Zusammenhang bekannt sein.

Beispiel 3.3.8

Abbildung 3.9 zeigt einige Beispiele für Distanzen gemessen in der Metrik \mathfrak{d} auf $\mathfrak{S}_5(\mathfrak{A})$. Es ist dabei zu beachten, daß, wenn sich nur zwei Segmente zweier Segmentierungen unterscheiden, dieselbe Veränderung zweimal gezählt wird. Denn wenn ein Segment von einer Segmentierung zur nächsten kleiner geworden ist, muß ein anderes Segment oder mehrere andere Segmente aufgrund der disjunkten Zerlegungseigenschaft größer geworden sein:

In der Beispielskizze von Abbildung 3.9 ist $\mathfrak{d}(\mathbf{A}, \mathbf{A}') = \mathfrak{d}(A_1, A'_1) + \mathfrak{d}(A_2, A'_2) = 5 + 5 = 10$. Man vergleiche dazu auch Lemma 3.3.9.

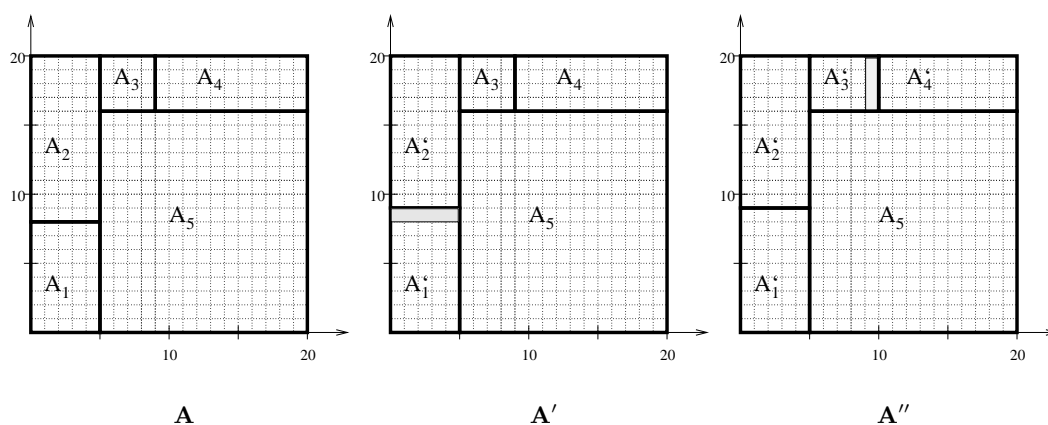


Abbildung 3.9: In diesem Beispiel ist $\mathfrak{d}(\mathbf{A}, \mathbf{A}') = 10$, $\mathfrak{d}(\mathbf{A}', \mathbf{A}'') = 8$ und $\mathfrak{d}(\mathbf{A}, \mathbf{A}'') = 18$.

Ebenso wie die Metrik auf \mathfrak{A} , verhält sich auch die Metrik \mathfrak{d} der Segmentierungen ungewohnt, wenn sich Segmente, die sich minimal unterscheiden, nicht überlappen (vgl. Abbildung 3.10).

Lemma 3.3.9

Seien $A, B \in \mathfrak{A}$. Dann gilt:

$$\mathfrak{d}(A, B) = \mathfrak{d}(M \setminus A, M \setminus B)$$

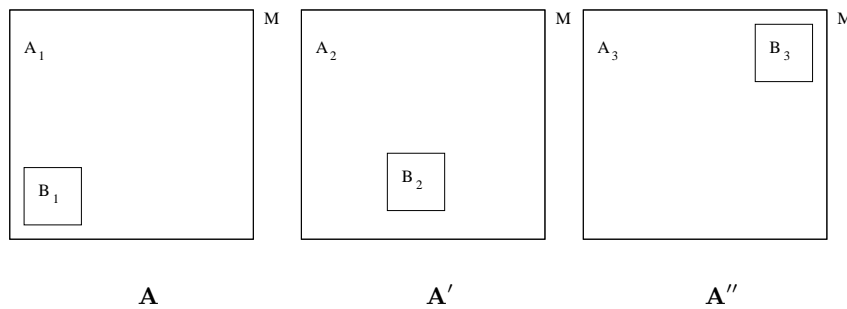


Abbildung 3.10: Es gilt: $\mathfrak{d}(\mathbf{A}, \mathbf{A}') = \mathfrak{d}(\mathbf{A}, \mathbf{A}'') = \mathfrak{d}(\mathbf{A}', \mathbf{A}'')$.

Beweis:

Es ist

$$\begin{aligned}
 \mathfrak{d}(M \setminus A, M \setminus B) &= \lambda(M \setminus A \cup M \setminus B) - \lambda(M \setminus A \cap M \setminus B) \\
 &= \lambda(M \setminus (A \cap B)) - \lambda(M \setminus (A \cup B)) \\
 &= \lambda(M) - \lambda(A \cap B) - \lambda(M) + \lambda(A \cup B) \\
 &= \lambda(A \cup B) - \lambda(A \cap B) \\
 &= \mathfrak{d}(A, B)
 \end{aligned}$$

□

Satz 3.3.10

Genügt \mathfrak{A} den Bedingungen in Definition 3.2.10, so ist $\mathfrak{S}_2(\mathfrak{A})$ abgeschlossen in $(\wp_k(\mathcal{B}(M)/\sim), \mathfrak{d})$.

Beweis:

Sei $(\mathbf{A}_n)_{n \in \mathbb{N}}$ eine Folge in $\mathfrak{S}_2(\mathfrak{A})$ mit $\lim_{n \rightarrow \infty} \mathbf{A}_n = \mathbf{A}_\infty$. Zu zeigen ist, daß $\mathbf{A}_\infty \in \mathfrak{S}_2(\mathfrak{A})$ gilt. Nach Definition 3.3.1 muß für alle $n \in \mathbb{N}$ $|\mathbf{A}_n| = 2$ gelten. Sei also $\mathbf{A}_n = \{A_n, B_n\}$ für $n \in \mathbb{N}$.

O. B. d. A. gelte für alle $n \in \mathbb{N}$ der Abstand $\mathfrak{d}(A_n, A_{n+1}) \leq \mathfrak{d}(A_n, B_{n+1})$. Dies ist durch Umbezeichnung immer möglich. Dann folgt mit Lemma 3.3.9 $\mathfrak{d}(B_n, B_{n+1}) = \mathfrak{d}(A_n, A_{n+1})$ und $\mathfrak{d}(B_n, A_{n+1}) = \mathfrak{d}(A_n, B_{n+1})$. Da aber nach Voraussetzung $\mathfrak{d}(A_n, B_{n+1}) \geq \mathfrak{d}(A_n, A_{n+1})$ gilt, folgt aus der Definition 3.3.5 der Metrik \mathfrak{d} :

$$\mathfrak{d}(\{A_n, B_n\}, \{A_{n+1}, B_{n+1}\}) = \mathfrak{d}(A_n, A_{n+1}) + \mathfrak{d}(B_n, B_{n+1})$$

Aus der Konvergenz von $(\{A_n, B_n\})_{n \in \mathbb{N}}$ folgt dann die Existenz von Mengen $A_\infty, B_\infty \in \mathbf{A}_\infty$ mit

$$\forall \varepsilon > 0 : \exists n_0 \in \mathbb{N} : \forall n > n_0 : \mathfrak{d}(A_n, A_\infty) \leq \frac{1}{2}\varepsilon \wedge \mathfrak{d}(B_n, B_\infty) \leq \frac{1}{2}\varepsilon \quad (3.1)$$

Daher ist $(\{A_n\})_{n \in \mathbb{N}}$ eine Folge in \mathfrak{A} , die gegen A_∞ konvergiert. Da \mathfrak{A} abgeschlossen ist, folgt $A_\infty \in \mathfrak{A}$. Analog ist $B_\infty \in \mathfrak{A}$. Daher bleibt nur noch $\lambda(A_\infty \cap B_\infty) = 0$ und $\lambda(A_\infty \cup B_\infty) = \lambda(M)$ zu zeigen, um $\{A_\infty, B_\infty\} \in \mathfrak{G}_2(\mathfrak{A})$ nachzuweisen.

Aber mit der Aussage (3.1) und der Definition 3.3.5 von \mathfrak{d} folgt für alle $n > n_0$:

$$\lambda(A_n \setminus A_\infty) < \frac{1}{2}\varepsilon, \lambda(A_\infty \setminus A_n) < \frac{1}{2}\varepsilon, \lambda(B_n \setminus B_\infty) < \frac{1}{2}\varepsilon, \lambda(B_\infty \setminus B_n) < \frac{1}{2}\varepsilon$$

Für $A_\infty \cap B_\infty$ gilt

$$\begin{aligned} A_\infty \cap B_\infty &\subseteq ((A_\infty \cap B_\infty) \setminus (A_n \cap B_n)) \cup (A_n \cap B_n) \\ &= (A_n \cap B_n) \cup ((A_\infty \setminus (A_n \cap B_n)) \cup (B_\infty \setminus (A_n \cap B_n))) \\ &\subseteq (A_n \cap B_n) \cup (A_\infty \setminus A_n) \cup (B_\infty \setminus B_n) \end{aligned} \quad (3.2)$$

und daher

$$\begin{aligned} \lambda(A_\infty \cap B_\infty) &\leq \lambda(A_n \cap B_n) + \lambda(A_\infty \setminus A_n) + \lambda(B_\infty \setminus B_n) \\ &= 0 + \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon \end{aligned}$$

Da $A_n \uplus B_n = M \supseteq A_\infty \cup B_\infty$ gelten muß, ist

$$\begin{aligned} A_\infty \cup B_\infty &= (A_n \uplus B_n) \setminus (A_n \setminus (A_\infty \cup B_\infty) \uplus B_n \setminus (A_\infty \cup B_\infty)) \\ &\supseteq (A_n \uplus B_n) \setminus (A_n \setminus A_\infty \uplus B_n \setminus B_\infty) \end{aligned} \quad (3.3)$$

und somit gilt

$$\begin{aligned} \lambda(A_\infty \cup B_\infty) &\geq \lambda((A_n \uplus B_n) \setminus (A_n \setminus A_\infty \uplus B_n \setminus B_\infty)) \\ &\geq \lambda(A_n \uplus B_n) - (\lambda(A_n \setminus A_\infty) + \lambda(B_n \setminus B_\infty)) \\ &= \lambda(M) - \left(\frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon\right) = \lambda(M) - \varepsilon \end{aligned}$$

Da ε beliebig war, folgt die Behauptung. □

Lemma 3.3.11

Für $A, B \in \mathfrak{A}$ gilt:

$$\mathfrak{d}(A, B) < \frac{\lambda(A)}{2} \Rightarrow \lambda(A \cap B) > \frac{\lambda(A)}{2}$$

Beweis:

Es ist

$$\begin{aligned}
\mathfrak{d}(A, B) < \frac{\lambda(A)}{2} &\Rightarrow \lambda(A \setminus B) + \lambda(B \setminus A) < \frac{\lambda(A)}{2} \\
&\Rightarrow \lambda(A \setminus B) < \frac{\lambda(A)}{2} \\
[A \setminus B = A \setminus (A \cap B)] &\Rightarrow \lambda(A \setminus (A \cap B)) < \frac{\lambda(A)}{2} \\
&\Rightarrow \lambda(A) - \lambda(A \cap B) < \frac{\lambda(A)}{2} \\
&\Rightarrow \lambda(A) - \frac{\lambda(A)}{2} < \lambda(A \cap B) \\
&\Rightarrow \frac{\lambda(A)}{2} < \lambda(A \cap B)
\end{aligned}$$

□

Lemma 3.3.12

Sei $(\mathbf{A}^{(n)})_{n \in \mathbb{N}}$ eine Folge in $\mathfrak{S}_k(\mathfrak{A})$, die gegen ein $\mathbf{A} = \{A_1, \dots, A_l\}$ konvergiert. Dann ist $\{A_1, \dots, A_l\}$ paarweise disjunkt und es existiert ein $n_0 \in \mathbb{N}$, so daß gilt:

Für alle $i \in \{1, \dots, l\}$ mit $A_i \neq O$ läßt sich eine Folge $(A_i^{(n)})_{\substack{n \in \mathbb{N} \\ n > n_0}}$ konstruieren, so daß gilt:

- i. $\lim_{\substack{n \rightarrow \infty \\ n > n_0}} A_i^{(n)} = A_i$
- ii. $\forall n \in \mathbb{N}, n > n_0 : A_i^{(n)} \in \mathbf{A}^{(n)}$ ist eindeutig bestimmt.
- iii. $\forall i \in \{1, \dots, l\} : A_i \in \mathfrak{A}$
- iv. Für alle $n \in \mathbb{N}$ mit $n > n_0$ sind die $A_i^{(n)}$ für $i \in \{1, \dots, l\}$ mit $A_i \neq O$ paarweise disjunkt.

Beweis:

Sei $\mathbf{A} = \lim_{n \rightarrow \infty} \mathbf{A}^{(n)}$ und $\{A_1, \dots, A_l\} = \mathbf{A} \setminus \{O\}$. Sei $\varepsilon > 0$ derart, daß für alle A_i mit $1 \leq i \leq l$ gilt: $\varepsilon < \frac{1}{2}\lambda(A_i)$.

Dann folgt aus der Konvergenz der Folge $\{A_1^{(n)}, \dots, A_k^{(n)}\}_{n \in \mathbb{N}}$, daß ein n_0 existiert mit

$$\forall n \geq n_0 : \mathfrak{d}(\mathbf{A}^{(n)}, \{A_1, \dots, A_l\}) \leq \mathfrak{d}(\mathbf{A}^{(n)}, \mathbf{A}) < \varepsilon$$

Sei $i \in \{1, \dots, l\}$ beliebig. Aus der Definition 3.3.5 von \mathfrak{d} folgt, daß für alle $n > n_0$ ein Segment $A_i^{(n)} \in \mathbf{A}^{(n)}$ existiert mit $\mathfrak{d}(A_i, A_i^{(n)}) < \varepsilon$. Insbesondere gilt damit nach

Voraussetzung $\vartheta(A_i, A_i^{(n)}) < \frac{1}{2}\lambda(A_i)$. Damit ist aber nach Lemma 3.3.11 $\lambda(A_i^{(n)} \cap A_i) > \frac{\lambda(A_i)}{2}$. Dies kann aber, da die Segmente in $\mathbf{A}^{(n)}$ paarweise disjunkt sind, nur für genau ein Segment in $\mathbf{A}^{(n)}$ gelten. Daher folgen mit der Konvergenz von $(\mathbf{A}^{(n)})_{n \in \mathbb{N}}$ die Behauptungen i. und ii.

Aus der Abgeschlossenheit von \mathfrak{A} folgt die Behauptung iii.

Weiter gilt für $A_i, A_j \in \{A_1, \dots, A_l\}$: $A_i \neq A_j \Rightarrow A_i \cap A_j = O$ Denn: Es ist $A_i \cap A_j \supseteq (A_i^{(n)} \cap A_j^{(n)}) \cup (A_i \setminus A_i^{(n)}) \cup (A_j \setminus A_j^{(n)})$. Daher folgt die Behauptung iv. analog zur Beweisführung von Satz 3.3.10 (vgl. dort Gleichung (3.2)). \square

Satz 3.3.13

$\mathfrak{S}_k(\mathfrak{A})$ ist abgeschlossen in $(\wp_k(\mathcal{B}(M)/\sim), \vartheta)$.

Beweis:

Sei $(\mathbf{A}^{(n)})_{n \in \mathbb{N}}$ eine Folge in $\mathfrak{S}_k(\mathfrak{A})$, die gegen ein $\{A_1, \dots, A_l\}$ konvergiert. Mit Lemma 3.3.12 sind dann alle Elemente von $\{A_1, \dots, A_l\}$ paarweise disjunkt und aus \mathfrak{A} . Daher bleibt nur noch

$$\lim_{\substack{n \rightarrow \infty \\ n > n_0}} \lambda(M \setminus \bigcup_{i=1}^l A_i^{(n)}) = 0$$

zu zeigen.

Mit Lemma 3.3.12 existiert ein n_0 derart, daß für jedes $i \in \{1, \dots, l\}$ eine eindeutig bestimmte Folge $(A_i^{(n)})_{n \in \mathbb{N}}$ existiert, mit $\forall n > n_0 : A_i^{(n)} \in \mathbf{A}^{(n)}$ und $\lim_{\substack{n \rightarrow \infty \\ n > n_0}} A_i^{(n)} = A_i$. Außerdem sind diese Folgen paarweise verschieden.

Da $\biguplus_{A \in \mathbf{A}^{(n)}} A = M \supseteq \biguplus_{i=1}^l A_i$ gelten muß, ist

$$\begin{aligned} \biguplus_{i=1}^l A_i &= \biguplus_{A \in \mathbf{A}^{(n)}} A \setminus \left(\biguplus_{A \in \mathbf{A}^{(n)}} (A \setminus \biguplus_{i=1}^l A_i) \right) \\ &\supseteq \biguplus_{A \in \mathbf{A}^{(n)}} A \setminus \left(\biguplus_{i=1}^l (A_i^{(n)} \setminus \biguplus_{i=1}^l A_i) \right) \\ &\supseteq \biguplus_{A \in \mathbf{A}^{(n)}} A \setminus \left(\biguplus_{i=1}^l (A_i^{(n)} \setminus A_i) \right) \end{aligned}$$

Daher läßt sich die Behauptung analog zur Beweisführung von Satz 3.3.10 (vgl. dort Gleichung (3.3)) zeigen.

Anmerkung: Für alle Folgen $(B^{(n)})_{n \in \mathbb{N}}$ mit $\forall n > n_0 : B^{(n)} \in \mathbf{A}^{(n)}$ und $\forall n > n_0 : B^{(n)} \notin \{A_1^{(n)}, \dots, A_l^{(n)}\}$ gilt: $\lim_{\substack{n \in \mathbb{N} \\ n > n_0}} B^{(n)} = O$. Daher muß in Definition 3.3.1 gefordert werden, daß alle Segmentierungen mit weniger als k echten Segmenten das leere Segment O enthalten. \square

Bemerkung 3.3.14

In Bemerkung 3.3.7 wird deutlich, daß das System der Segmentierungen $\mathfrak{S}_1(\mathfrak{A}) = \{\{M\}\}$ höchstens von akademischen Interesse ist. Es sei daher im Rest der Arbeit für $\mathfrak{S}_k(\mathfrak{A})$ immer $k \geq 2$ gefordert.

3.4 Die Klasse der erlaubten Funktionen

Die Klasse der erlaubten Funktionen bestimmt in erheblicher Weise das Approximationsproblem. In dieser Arbeit wird das Approximationsproblem auf einer Menge $A \in \mathfrak{A}$ durch orthogonale Projektion gelöst. Dafür benötigt man allerdings ein Skalarprodukt und eine orthonormale Basis des Funktionenraums. Dieses Skalarprodukt – und damit die Orthonormalbasis – sollte abhängig von der Menge A sein. Das Mengensystem bestimmt also die Basen des Funktionenraums, mit dem stückweise approximiert wird. Im Mittelpunkt dieses Abschnitts steht daher weniger der Funktionenraum der erlaubten Funktionen als vielmehr seine Darstellung bezüglich der Orthonormalbasis des jeweiligen durch ein Segment bestimmten Skalarprodukts.

Ist E ein vollständiger normierter reeller Vektorraum mit Norm $\|\cdot\|$, so sei im folgenden $f \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ das Bild, wobei $M \subset \mathbb{R}^k$ mit $\lambda(M) > 0$ kompakt sei.

Definition 3.4.1

Für ein System erlaubter Mengen \mathfrak{A} sei $\sigma(\mathfrak{A}) \subset \mathcal{B}(M) / \sim$ die kleinste σ -Algebra, die \mathfrak{A} enthält.

Definition 3.4.2 (Erlaubte Funktionen)

Sei $V \subset L^2(\mathcal{B}(M), \mathcal{B}(E))$ ein Vektorraum stetiger Funktionen mit endlicher Basis $\mathcal{E} = (e_1, \dots, e_n)$, so daß für alle $i \in \{1, \dots, n\}$ und $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) > 0$ gilt:

$V|_A := \{f|_A \mid f \in V\}$ ist ein Vektorraum mit $\dim V|_A = n$.

Erfüllt V diese Eigenschaft, so heiße V eine *Klasse von erlaubten Funktionen*.

Es existiert dann zu jedem $f \in V|_A$ genau ein $g \in V$ mit $g|_A = f$. Außerdem gilt $\forall A \in \sigma(\mathfrak{A}) : \forall i \in \{1, \dots, n\} : \lambda(A) > 0 \Rightarrow \int_A \|e_i\|_2 d\lambda > 0$

Bemerkung 3.4.3

Die Klasse der erlaubten Funktionen ist also nichts anderes als ein endlich dimensionaler Vektorraum mit gutartigem Verhalten bei Einschränkungen auf Mengen aus dem Mengensystem \mathfrak{A} . Die Fragen, die sich ergeben, wenn ein solcher Funktionenraum bei Einschränkungen auf Mengen aus \mathfrak{A} nicht gutartig, d. h. mit Dimensionsverlust, reagiert, sind von zentraler Bedeutung in der Arbeit von Verena Lauren ([Lau03]). Deshalb soll an dieser Stelle nicht näher darauf eingegangen werden.

Der Rest dieses Abschnitts wird sich mit der Frage einer geeigneten Darstellung dieses Vektorraums beschäftigen.

Die folgende Definition 3.4.4 stellt den Zusammenhang zwischen dem Mengensystem und einer Familie von Skalarprodukten her. Grundsätzlich gilt: „Die Menge bestimmt das Skalarprodukt“. Das heißt, zu jeder Menge $A \in \sigma(\mathfrak{A})$ existiert genau ein Skalarprodukt, bezüglich dessen man eine Orthonormalbasis des Funktionenraums angeben kann. Dabei sollten die einzelnen Skalarprodukte der einzelnen Mengen nicht beziehungslos nebeneinander stehen.

Die folgende Definition führt die Familie von Skalarprodukten, die im folgenden das Approximationsproblem beschreibt, ein.

Definition 3.4.4 (Skalarprodukte auf $L^2(\mathcal{B}(M), \mathcal{B}(E))$)

Auf $L^2(\mathcal{B}(M), \mathcal{B}(E))$ sei eine Menge $\{\langle \cdot | \cdot \rangle_A \mid A \in \sigma(\mathfrak{A}), \lambda(A) > 0\}$ von symmetrischen Bilinearformen definiert, derart, daß für alle $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) > 0$ genau eine symmetrische Bilinearform $\langle \cdot | \cdot \rangle_A$ existiert mit

- i. $\langle \cdot | \cdot \rangle_A|_A$ ist ein Skalarprodukt auf $L^2(\mathcal{B}(A), \mathcal{B}(E))$.²
- ii. Für alle $f, g \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ ist $|\langle f | g \rangle_M| < \infty$.
- iii. Für alle $f, g \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ gilt:
 $\forall \varepsilon > 0 : \exists \delta > 0 : \forall A \in \sigma(\mathfrak{A}) : \lambda(A) < \delta \Rightarrow |\langle f | g \rangle_A| < \varepsilon$.
- iv. Für alle $A, B \in \sigma(\mathfrak{A})$ mit $\lambda(A \cap B) = 0$ gilt:
 $\forall f, g \in L^2(\mathcal{B}(M), \mathcal{B}(E)) : \langle f | g \rangle_A + \langle f | g \rangle_B = \langle f | g \rangle_{A \uplus B}$.
- v. Für alle $A \in \sigma(\mathfrak{A})$ und alle $f, g \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ ist $\langle f \mathbf{1}_A | g \rangle_A = \langle f | g \rangle_A$.

Eine Familie von symmetrischen Bilinearformen, die diese Eigenschaften erfüllt, heie *erlaubt*.

Bemerkung 3.4.5

- a) Die Familie der erlaubten Bilinearformen wurde auf $\sigma(\mathfrak{A})$ statt auf \mathfrak{A} definiert, um einen Maraum zu erhalten. Diese Eigenschaft ist in vielen Beweisen ntzlich. Die zentralen spteren Aussagen werden jedoch auf \mathfrak{A} formuliert.
- b) Die Bedingung iii. formuliert die Stetigkeit von Maen.

Fr jede Wahl von f und g aus $L^2(\mathcal{B}(M), \mathcal{B}(E))$ ist $\nu : \sigma(\mathfrak{A}) \rightarrow \mathbb{R}_+$ mit $\nu(A) = |\langle f | g \rangle_A|$ ein Ma. Denn:

Offensichtlich ist ν ein Inhalt. Aus der Bedingung iii. folgt die \emptyset -Stetigkeit von ν . Daher ist ν Prma auf einer σ -Algebra, also ein Ma (vgl. [Bau92], S. 12f).

Damit ist die Bedingung iii. nichts anderes als die Forderung der Stetigkeit des Maes ν bezglich des Lebesgue-Maes λ . Die angegebene Bedingung ist dann quivalent dazu, da jede λ -Nullmenge auch eine ν -Nullmenge ist (vgl. [Bau92], S. 113).

² $\langle \cdot | \cdot \rangle_A$ ist eigentlich nur auf $L^2(\mathcal{B}(M), \mathcal{B}(E))$ definiert. Funktionen aus $L^2(\mathcal{B}(A), \mathcal{B}(E))$ men daher auf $M \setminus A$ fortgesetzt werden. In diesem Sinne sei $\langle \cdot | \cdot \rangle_A|_A$ die Bilinearform $\langle \cdot | \cdot \rangle_A$ betrachtet nur auf Funktionen aus $L^2(\mathcal{B}(A), \mathcal{B}(E))$, die durch 0 zu Funktionen auf $L^2(\mathcal{B}(M), \mathcal{B}(E))$ fortgesetzt werden.

- c) Die Bedingung v. in 3.4.4 ist für viele der späteren Beweise eigentlich nicht nötig. Sie soll jedoch die Thematik dieser Arbeit einschränken und die Beweisführung von zahlreichen Fallunterscheidungen befreien. In der Arbeit von Verena Lauren ([Lau03]) wird verstärkt auf die Probleme, die aus dem Auslassen dieser Bedingung folgen, eingegangen.
- d) Da nach der Definition der Klasse der erlaubten Funktionen V jede Funktion aus V auf einem beliebigen Träger $A \in \sigma(\mathfrak{A})$ mit positivem Lebesgue-Maß eindeutig bestimmt ist, ist mit Bedingung i. bei $\lambda(A) > 0$ die Bilinearform $\langle \cdot | \cdot \rangle_A$ ein Skalarprodukt auf V . Daher wird im folgenden immer vom Skalarprodukt $\langle \cdot | \cdot \rangle_A$ die Rede sein.

Beim ersten Lesen der Bedingung i. könnte man sich fragen, warum $\langle \cdot | \cdot \rangle_A$ nicht gleich als Skalarprodukt auf $L^2(\mathcal{B}(M), \mathcal{B}(E))$ definiert wird. Das folgende Beispiel zeigt unter anderem den Sinn dieser Definition.

Beispiel 3.4.6

Ist auf E ein Skalarprodukt $\langle \cdot | \cdot \rangle$ definiert und sei für $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) > 0$ das Skalarprodukt $\langle \cdot | \cdot \rangle_A : L^2(\mathcal{B}(M), \mathcal{B}(E)) \times L^2(\mathcal{B}(M), \mathcal{B}(E)) \rightarrow \mathbb{R}$ definiert durch

$$\langle f | g \rangle_A := \int_A \langle f | g \rangle d\lambda$$

Dann erfüllt die Familie $\{\langle \cdot | \cdot \rangle_A \mid A \in \sigma(\mathfrak{A}), \lambda(A) > 0\}$ die Bedingungen aus Definition 3.4.4.

Denn:

- i. Für alle $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) > 0$ ist $\langle \cdot | \cdot \rangle_A$ ein Skalarprodukt auf $L^2(\mathcal{B}(A), \mathcal{B}(E))$. Distributivität, Kommutativität und Homogenität ergeben sich aus den Eigenschaften des Skalarprodukts auf E . Die positive Definitheit folgt aus der Tatsache, daß nur Funktionen aus $L^2(\mathcal{B}(A), \mathcal{B}(E))$ betrachtet werden: Eine Funktion f , für die $\int_A \langle f | f \rangle d\lambda = \int_A \|f\|_2^2 d\lambda = 0$ gilt, ist auf A fast überall 0, liegt also in derselben Äquivalenzklasse wie die Nullfunktion. Allerdings ist $\langle \cdot | \cdot \rangle_A$ kein Skalarprodukt auf $L^2(\mathcal{B}(M), \mathcal{B}(E))$. Denn z. B. für die Funktion $f := 0\mathbf{1}_A + \mathbf{1}_{M \setminus A}$ gilt $\langle f | f \rangle_A = 0$, aber f ist nicht in der Äquivalenzklasse der 0-Funktion in $L^2(\mathcal{B}(M), \mathcal{B}(E))$, falls $\lambda(M \setminus A) > 0$ ist.
- ii. Da $\langle \cdot | \cdot \rangle_A$ auf dem Raum der quadratisch integrierbaren Funktionen definiert ist, folgt die Bedingung ii. Denn mit der Ungleichung von Cauchy und Schwarz gilt für alle $f, g \in L^2(\mathcal{B}(A), \mathcal{B}(E))$:

$$\infty > \int_M \|f\|_2 \|g\|_2 d\lambda \geq \int |\langle f | g \rangle| d\lambda$$

- iii. Der Satz von Radon-Nikodym (vgl. z. B. [Bau92], S. 116) liefert die Erfüllung von Bedingung iii.: Definiert man $\nu(A) := \int_A |\langle f|g \rangle| d\lambda$, so beschreibt $|\langle f|g \rangle|$ die Dichte des Maßes ν bezüglich des Lebesgue-Maßes λ . Daher existiert nach [Bau92], 17.8 Satz, S. 113, zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit

$$\lambda(A) < \delta \Rightarrow \nu(A) < \varepsilon$$

und es gilt $|\int_A \langle f|g \rangle d\lambda| \leq \int_A |\langle f|g \rangle| d\lambda < \varepsilon$.

- iv. Für alle A und B aus \mathfrak{A} mit $\lambda(A \cap B) = 0$ sowie für alle $f, g \in V$ gilt:

$$\langle f|g \rangle_A + \langle f|g \rangle_B = \int_A \langle f|g \rangle d\lambda + \int_B \langle f|g \rangle d\lambda = \int_{A \cup B} \langle f|g \rangle d\lambda$$

- v. Die Bedingung v. ist ebenso leicht ersichtlich.

Bemerkung 3.4.7

Das Beispiel 3.4.6 ist die hauptsächliche Anwendung der Definition 3.4.4 im praktischen Teil dieser Arbeit. Zum Beweis aller Sätze der theoretischen Grundlagen der vektoriellen Mehrniveaupassung reicht jedoch die abstrakte Definition 3.4.4 vollkommen aus. Damit ist auch Nicht-Standardanwendungen eine theoretische Grundlage gegeben: So könnte man z. B. über das Skalarprodukt die Approximation gewichten (vgl. auch Abschnitt 4.6.4 und Abschnitt 6.4.1.5).

Wenn nicht anders erwähnt, sei jedoch in den folgenden Beispielen dieser Arbeit die jeweilige Familie von erlaubten Skalarprodukten immer vom Skalarprodukt des Raums E induziert, wie es in Beispiel 3.4.6 in der Fall ist.

Bemerkung 3.4.8

Mit $\|\cdot\|_A := \sqrt{\langle \cdot | \cdot \rangle_A}$ ist zu jedem $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) > 0$ eine Norm auf V definiert. Mittels des Skalarprodukts und der Norm auf V ist es möglich, zu jedem $A \in \sigma(\mathfrak{A})$ eine Orthonormalbasis (e_1^A, \dots, e_n^A) bezüglich $\langle \cdot | \cdot \rangle_A$ anzugeben. Diese ist so allerdings nicht eindeutig bestimmt. Im Rest der Arbeit sei jedoch bei gegebener Basis (e_1, \dots, e_n) von V die Orthonormalbasis (e_1^A, \dots, e_n^A) mit Hilfe des Orthogonalisierungsverfahren von Gram und Schmidt (vgl. Algorithmus 5 (**Gram-Schmidt**)) entstanden.³

³Zur Verwendung dieses Verfahrens vgl. auch Lemma 3.6.21 und Bemerkung 3.6.23.

Algorithmus 5 (Gram-Schmidt) Berechnung einer Orthonormalbasis (e_1^A, \dots, e_n^A) aus einer Basis $\mathcal{E} = (e_1, \dots, e_n)$ bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_A$

Input: Basis $\mathcal{E} = (e_1, \dots, e_n)$ des Vektorraums V ,
Skalarprodukt $\langle \cdot | \cdot \rangle_A$ auf V

Output: Orthonormalbasis (e_1^A, \dots, e_n^A) von V bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_A$

```

1:  $e_1^A := \frac{e_1}{\|e_1\|_A}$ 
2:  $k := 1$ 
3: while  $k < n$  do
4:    $\tilde{e}_{k+1}^A := e_{k+1} - \sum_{i=1}^k \langle e_i^A | e_{k+1} \rangle_A e_i^A$ 
5:    $e_{k+1}^A := \frac{\tilde{e}_{k+1}^A}{\|\tilde{e}_{k+1}^A\|_A}$ 
6:    $k \leftarrow k + 1$ 
7: end while
8: return  $(e_1^A, \dots, e_n^A)$ 

```

Definition 3.4.9 (Ausgezeichnete Orthonormalbasis)

Sei $\mathcal{E} = (e_1, \dots, e_n)$ eine Basis des Vektorraums der erlaubten Funktionen. Für $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) > 0$ sei (e_1^A, \dots, e_n^A) die eindeutig bestimmte Orthonormalbasis, die durch die Orthonormalisierung von \mathcal{E} gemäß dem Orthogonalisierungsverfahren nach Gram und Schmidt (Algorithmus 5 (Gram-Schmidt)) entstanden ist.

$\mathcal{E}^A := (e_1^A, \dots, e_n^A)$ sei der Bezeichner dieser Orthonormalbasis.

Für $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) = 0$, d. h. $A = O$, sei $\mathcal{E}^A := (0, \dots, 0)$.⁴

(e_1^A, \dots, e_n^A) heie die ausgezeichnete Orthonormalbasis bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_A$ und der Basis \mathcal{E} .

Definition 3.4.10 (Koordinatendarstellung)

Fr eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V heit die Abbildung $\mathcal{K}_{\mathcal{B}} : V \rightarrow \mathbb{R}^n$ definiert durch

$$\mathcal{K}_{\mathcal{B}}(f) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \text{ mit } \sum_{i=1}^n x_i v_i = f$$

die Koordinatendarstellung von f bezüglich der Basis \mathcal{B} .

Bemerkung 3.4.11

Sei $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) > 0$. Ist dann $\mathbf{A} \in \mathbb{R}^{n \times n}$ definiert durch $\mathbf{A} := (\langle e_i | e_j \rangle_A)_{1 \leq i, j \leq n}$, so gilt

$$\forall f, g \in V : \langle f | g \rangle_A = \mathcal{K}_{\mathcal{E}}(f)^t \mathbf{A} \mathcal{K}_{\mathcal{E}}(g)$$

⁴ \mathcal{E}^O ist im strengen Sinne keine Basis!

und insbesondere ist für $(e_1, \dots, e_n) = (e_1^A, \dots, e_n^A)$

$$\begin{aligned}\forall f, g \in V : \langle f | g \rangle_A &= \mathcal{K}_{\mathcal{E}^A}(f)^t ((\langle e_i^A | e_j^A \rangle_A)_{1 \leq i, j \leq n}) \mathcal{K}_{\mathcal{E}^A}(g) \\ &= \mathcal{K}_{\mathcal{E}^A}(f)^t \mathcal{K}_{\mathcal{E}^A}(g)\end{aligned}$$

Da \mathcal{E}^A eine Orthonormalbasis bildet, gilt

$$\mathcal{K}_{\mathcal{E}^A}(f) = \begin{pmatrix} \langle f | e_1^A \rangle_A \\ \vdots \\ \langle f | e_n^A \rangle_A \end{pmatrix}$$

Beispiel 3.4.12

Alle hier aufgeführten Beispiele sind Konkretisierungen von Beispiel 3.4.6

- i) Sei $E := \mathbb{R}$ mit der Multiplikation \cdot als Skalarprodukt und $V := \{\alpha \mid \alpha \in \mathbb{R}\}$ der Vektorraum der konstanten Funktionen von \mathbb{R} nach \mathbb{R} . Die Familie der erlaubten Skalarprodukte sei dann definiert wie in Beispiel 3.4.6. Dann ist (1) eine Basis dieses Raums und für $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) > 0$ ist $e_1^A = \frac{1}{\sqrt{\lambda(A)}}$ die ausgezeichnete Orthonormalbasis bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_A$ und der Basis (1) gemäß Definition 3.4.9.
- ii) Sei $E := \mathbb{R}^d$ mit dem Standardskalarprodukt $\langle \cdot | \cdot \rangle$. Die Familie der erlaubten Skalarprodukte sei wieder definiert wie in Beispiel 3.4.6. $V := \{(\alpha, \dots, \alpha) \mid \alpha \in \mathbb{R}\}$ sei ein Vektorraum konstanter Funktionen in \mathbb{R}^d mit Basis $(\underbrace{(1, \dots, 1)}_{d\text{-mal}})$. Dann ist für $A \in \sigma(\mathfrak{A})$ mit positivem Lebesgue-Maß

$$\|(1, \dots, 1)\|_A^2 = \int_A \langle (1, \dots, 1) | (1, \dots, 1) \rangle d\lambda = \int_A d\lambda = d\lambda(A)$$

Daher ist $e_1^A = \frac{1}{\sqrt{d\lambda(A)}}(1, \dots, 1)$ die ausgezeichnete Orthonormalbasis bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_A$ und der Basis $((1, \dots, 1))$ gemäß Definition 3.4.9.

- iii) Sei $E := \mathbb{R}^3$ mit dem Standardskalarprodukt $\langle \cdot | \cdot \rangle$ wie in ii). Mit Hilfe dieses Zielraums lassen sich auch Farbbilder kodieren, indem man sie z. B. in einen Rot-, Grün und Blaukanal aufspaltet (vgl. Abschnitt 1.3.2). V sei der Vektorraum aller konstanten Funktionen von M nach \mathbb{R}^3 . Eine Basis dieses Raums ist (e_1, e_2, e_3) mit $e_1 := (1, 0, 0)$, $e_2 := (0, 1, 0)$ und $e_3 := (0, 0, 1)$. Dann ist für $A \in \sigma(\mathfrak{A}) \setminus \{O\}$ mit $e_1^A = \frac{1}{\sqrt{\lambda(A)}}e_1$, $e_2^A = \frac{1}{\sqrt{\lambda(A)}}e_2$, $e_3^A = \frac{1}{\sqrt{\lambda(A)}}e_3$ die Basis (e_1^A, e_2^A, e_3^A) die ausgezeichnete Orthonormalbasis bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_A$ und der Basis (e_1, e_2, e_3) gemäß Definition 3.4.9, wie man leicht nachrechnet.⁵
- iv) Sei $M := [0, 1]$ und $E := \mathbb{R}$ mit der Multiplikation reeller Zahlen als Skalarprodukt. Die Klasse der erlaubten Funktionen sei $V = \{f \mid f(x) = ax + t, a, t \in \mathbb{R}\}$. Dann ist (e_1, e_2) mit $e_1(x) := 1$ und $e_2(x) := x$ eine Basis von V .

Weiter sei $\mathfrak{A} := \{[0, t] \mid t \in [0, \dots, 1]\}$. Daraus ergibt sich für $[0, t] \in \mathfrak{A}$ mit $t > 0$:

$$e_1^{[0,t]}(x) = \frac{1}{\sqrt{t}} \quad \text{und} \quad e_2^{[0,t]}(x) = 2\sqrt{\frac{3}{t^3}}\left(x - \frac{1}{2}t\right)$$

⁵Natürlich läßt sich diese Basis auch für beliebige \mathbb{R}^d mit $d \neq 3$ vollkommen analog definieren.

3.5 Problemstellung

Dieses Unterkapitel dient dazu, das Approximationsproblem, das die vektorielle k -Niveaupassung darstellt, zu formulieren.

Dabei werden die Konzepte der beiden vorangegangenen Abschnitte zusammenfließen: Auf der einen Seite das System der Segmentierungen von M durch k erlaubte Mengen und auf der anderen Seite die Klasse der erlaubten Funktionen. Ziel dabei ist es, ein Approximationsproblem zu formulieren, das ein gegebenes Bild f durch eine stückweise definierte Funktion approximiert. Diese Funktion wird auf den einzelnen Segmenten einer Segmentierung des Definitionsbereichs von f durch k erlaubte Mengen durch eine Funktion aus der Klasse der erlaubten Funktionen beschrieben.

Das wichtigste Ergebnis dieses Abschnitts wird sein, daß sich das Problem, Mengen und Funktionen zu finden, die den Abstand zum Bild minimieren, auf ein Problem reduzieren läßt, das nur über die Segmentierungen parametrisiert ist. Die Segmentierung wird also die zu verwendenden Funktionen bestimmen.

Weiter werden verschiedene Formulierungen, die zueinander äquivalent sind, vorgestellt. Je nach Problem – ob Beweis oder Algorithmus – erlauben es diese unterschiedlichen Darstellungen des Problems, die jeweils günstigste zu wählen.

Im ersten Abschnitt dieses Unterkapitels wird die Segmentierung mit zwei Segmenten, also die 2-Niveaupassung, betrachtet, weil dieser Fall eine einfache Parametrisierung erlaubt. Die Ergebnisse lassen sich dann im zweiten Abschnitt leicht auf den Fall der Segmentierung mit einer höheren Anzahl von Niveaumengen übertragen.

3.5.1 Die Zwei-Niveau-Situation

Im Falle von zwei Niveaus ist $\mathfrak{S}_2(\mathfrak{A}) = \{\{A, M \setminus A\} \mid A \in \mathfrak{A}, M \setminus A \in \mathfrak{A}\}$, d. h. bei der Auswahl von $A \in \mathfrak{A}$ ist darauf zu achten, daß auch $M \setminus A$ in \mathfrak{A} enthalten ist, um mit $\{A, M \setminus A\}$ eine Segmentierung von M durch zwei erlaubte Mengen zu erhalten.

In diesem Abschnitt soll jedoch zunächst auf die Überprüfung $M \setminus A \in \mathfrak{A}$ verzichtet werden. Es werden also Segmentierungen der Form $\{\{A, M \setminus A\} \mid A \in \mathfrak{A}\}$ betrachtet. Diese Menge ist eine Obermenge von $\mathfrak{S}_2(\mathfrak{A})$, die eine einfachere Parametrisierung – mit nur einer Menge als Parameter – erlaubt, da mit A die zweite Menge der Segmentierung $M \setminus A$ automatisch gegeben ist. Aus Gründen der Übersicht wird an dieser Stelle also nur dieses System der Segmentierungen verwendet. Man beachte daher im folgenden, daß $M \setminus A$ nicht in \mathfrak{A} sein muß. Allerdings muß $M \setminus A$ in $\sigma(\mathfrak{A})$, der kleinsten σ -Algebra, die \mathfrak{A} enthält, sein.

Ziel der 2-Niveaupassung ist es, eine Menge $A \in \mathfrak{A}$ und zwei Funktionen $g_1, g_2 \in V$ so zu bestimmen, daß sie die Funktion

$$\tilde{\Phi} : \mathfrak{A} \times V \times V \rightarrow \mathbb{R} \\ (A, g_1, g_2) \mapsto \|f - g_1 \mathbf{1}_A - g_2 \mathbf{1}_{M \setminus A}\|_M^2$$

minimieren. $\|\cdot\|_M$ sei dabei definiert wie in Bemerkung 3.4.8, also induziert von der Familie der erlaubten Skalarprodukte (vgl. Definition 3.4.4). Die Existenz eines Minimums der Funktion $\tilde{\Phi}$ muß allerdings noch gesichert werden. Dies geschieht im anschließenden Unterkapitel 3.6.

Bemerkung 3.5.1

- a) Aus den Eigenschaften der Familie der erlaubten Skalarprodukte (vgl. Definition 3.4.2) folgt:

$$\begin{aligned}
& \tilde{\Phi}(A, g_1, g_2) \\
&= \|f - g_1 \mathbf{1}_A - g_2 \mathbf{1}_{M \setminus A}\|_M^2 \\
&= \langle f - g_1 \mathbf{1}_A - g_2 \mathbf{1}_{M \setminus A} | f - g_1 \mathbf{1}_A - g_2 \mathbf{1}_{M \setminus A} \rangle_M \\
&= \langle f - g_1 \mathbf{1}_A - g_2 \mathbf{1}_{M \setminus A} | f - g_1 \mathbf{1}_A - g_2 \mathbf{1}_{M \setminus A} \rangle_A \\
&\quad + \langle f - g_1 \mathbf{1}_A - g_2 \mathbf{1}_{M \setminus A} | f - g_1 \mathbf{1}_A - g_2 \mathbf{1}_{M \setminus A} \rangle_{M \setminus A} \\
&= \langle f - g_1 | f - g_1 \rangle_A + \langle f - g_2 | f - g_2 \rangle_{M \setminus A} \\
&= \langle f | f \rangle_A - 2 \langle f | g_1 \rangle_A + \langle g_1 | g_1 \rangle_A \\
&\quad + \langle f | f \rangle_{M \setminus A} - 2 \langle f | g_2 \rangle_{M \setminus A} + \langle g_2 | g_2 \rangle_{M \setminus A} \\
&= \|f\|_M^2 - 2 \langle f | g_1 \rangle_A + \langle g_1 | g_1 \rangle_A - 2 \langle f | g_2 \rangle_{M \setminus A} + \langle g_2 | g_2 \rangle_{M \setminus A}
\end{aligned}$$

Da $\|f\|_M^2$ für jede Wahl von $A \in \mathfrak{A}$ und $g_1, g_2 \in V$ konstant ist, genügt es also, im folgenden den Term

$$\begin{aligned}
\Phi(A, g_1, g_2) &:= \tilde{\Phi}(A, g_1, g_2) - \|f\|_M^2 \\
&= \langle g_1 | g_1 \rangle_A - 2 \langle f | g_1 \rangle_A + \langle g_2 | g_2 \rangle_{M \setminus A} - 2 \langle f | g_2 \rangle_{M \setminus A}
\end{aligned}$$

zu betrachten.

- b) Weil V eine endliche Basis besitzt, existiert für jedes $g \in V$ eine Darstellung bezüglich dieser Basis. Da mit Definition 3.4.9 für jede erlaubte Menge $A \in \mathfrak{A} \setminus \{O\}$ bzw. $A \in \sigma(\mathfrak{A}) \setminus \{O\}$ eine ausgezeichnete Orthonormalbasis $\mathcal{E}^A = (e_1^A, \dots, e_n^A)$ existiert, bietet es sich an, für jede Funktion $g \in V$, die nur auf A ausgewertet werden soll, die Darstellung von g bezüglich dieser Basis zu wählen: $g = \sum_{i=1}^n \alpha_i e_i^A$. Daher kann man die Funktion Φ über die Koeffizienten der Darstellung zur jeweiligen Orthonormalbasis parametrisieren:

$$\Phi'(A, \alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n) := \Phi(A, \sum_{i=1}^n \alpha_i e_i^A, \sum_{i=1}^n \beta_i e_i^{M \setminus A}),$$

(Dabei gilt nach Definition 3.4.9 $e_i^A := 0$ für $i \in \{1, \dots, n\}$ für Mengen $A \in \sigma(\mathfrak{A})$ mit leerem Lebesgue-Maß.)

Φ' hat als Definitionsbereich die Menge $\mathfrak{A} \times \mathbb{R}^n \times \mathbb{R}^n$. Da eine Unterscheidung von Φ und Φ' aus der Parametrisierung eindeutig hervorgeht, wird Φ' wieder mit Φ bezeichnet. Ebenso bezeichnet in legerer Schreibweise für $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ und $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{R}^n$ der Ausdruck $\Phi(A, \alpha, \beta)$ die Funktion $\Phi(A, \alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n)$.

Definition 3.5.2

Für $f \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ sei $\gamma_f : \sigma(\mathfrak{A}) \rightarrow \mathbb{R}^n$ definiert durch

$$\gamma_f(A) := \begin{cases} \begin{pmatrix} \langle f | e_1^A \rangle_A \\ \vdots \\ \langle f | e_n^A \rangle_A \end{pmatrix}, & \text{falls } \lambda(A) > 0 \\ \mathbf{0} & \text{sonst} \end{cases}$$

Für $1 \leq i \leq n$ bezeichne $\gamma_f(A)_i$ den i -ten Eintrag des Vektors $\gamma_f(A)$.

Satz 3.5.3

Für alle $A \in \mathfrak{A}$ und $\alpha, \beta \in \mathbb{R}^n$ gilt

$$\Phi(A, \alpha, \beta) \geq \Phi(A, \gamma_f(A)^t, \gamma_f(M \setminus A)^t)$$

Beweis:

Sei $\lambda(A) = 0$, also $A = O$. Dann ist $\Phi(A, \alpha, \beta) = \left\langle \sum_{i=1}^n \beta_i e_i^M \middle| \sum_{j=1}^n \beta_j e_j^M \right\rangle_M - 2 \left\langle \sum_{i=1}^n \beta_i e_i^M \middle| f \right\rangle_M = \Phi(A, x, \beta)$ für beliebige $x \in \mathbb{R}^n$, also insbesondere auch für $x = \gamma_f(A)^t = \gamma_f(O)^t = 0$. Weil die Behauptung für den Fall $\lambda(A) = \lambda(M)$, also $A = M$, vollkommen analog gezeigt werden kann, reicht es nun aus, die Behauptung für alle $A \in \mathfrak{A}$ mit $\lambda(A) > 0$ und $\lambda(M \setminus A) > 0$ zu zeigen.

Sei also $A \in \mathfrak{A}$ mit $\lambda(A) > 0$ und $\lambda(M \setminus A) > 0$. Dann gilt mit Bemerkung 3.5.1 a)

$$\Phi(A, \alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n) \geq \Phi(A, \gamma_1, \dots, \gamma_n, \eta_1, \dots, \eta_n)$$

\Leftrightarrow

$$\left\| f - \sum_{i=1}^n \alpha_i e_i^A \mathbf{1}_A - \sum_{j=1}^n \beta_j e_j^{M \setminus A} \mathbf{1}_{M \setminus A} \right\|_M \leq \left\| f - \sum_{i=1}^n \gamma_i e_i^A \mathbf{1}_A - \sum_{j=1}^n \eta_j e_j^{M \setminus A} \mathbf{1}_{M \setminus A} \right\|_M$$

Daher genügt es zu zeigen, daß das Tupel $(\gamma_f(A)^t, \gamma_f(M \setminus A)^t)$ die Funktion

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R}^n &\rightarrow \mathbb{R}_+ \\ (\alpha, \beta) &\mapsto \left\| f - \sum_{i=1}^n \alpha_i e_i^A \mathbf{1}_A - \sum_{j=1}^n \beta_j e_j^{M \setminus A} \mathbf{1}_{M \setminus A} \right\|_M \end{aligned}$$

minimiert.

Da nach Definition 3.4.4 iv.

$$\left\| f - \sum_{i=1}^n \alpha_i e_i^A \mathbf{1}_A - \sum_{j=1}^n \beta_j e_j^{M \setminus A} \mathbf{1}_{M \setminus A} \right\|_M^2 = \left\| f - \sum_{i=1}^n \alpha_i e_i^A \right\|_A^2 + \left\| f - \sum_{i=1}^n \beta_i e_i^{M \setminus A} \right\|_{M \setminus A}^2$$

gilt, kann man das Minimierungsproblem für α und β getrennt betrachten. Nach [Die85] (6.3.1, S. 118) ist aber $\left\| f - \sum_{i=1}^n \alpha_i e_i^A \right\|_A^2$ genau dann minimal, wenn $\sum_{i=1}^n \alpha_i e_i^A$ die orthogonale Projektion von f auf den Vektorraum V bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_A$ ist.

Ist jedoch $\sum_{i=1}^n \alpha_i e_i^A$ diese orthogonale Projektion, so muß, da (e_1^A, \dots, e_n^A) eine Orthonormalbasis von V bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_A$ ist, für alle $1 \leq j \leq n$

$$0 = \left\langle f - \sum_{i=1}^n \alpha_i e_i^A \middle| e_j^A \right\rangle_A = \langle f | e_j^A \rangle_A - \alpha_j$$

gelten und daher muß $\alpha_j = \langle f | e_j^A \rangle_A$ sein. Analog zeigt man, daß $\beta_j = \langle f | e_j^{M \setminus A} \rangle_{M \setminus A}$ für $1 \leq j \leq n$ gelten muß. Damit folgt die Behauptung mit Definition 3.5.2. \square

Bemerkung 3.5.4

Bei der Beweisführung des Satzes 3.5.3 handelt es sich um ein Beispiel der Optimalitätssuche über orthogonale Projektion auf den Raum $V|_A$ mit $\lambda(A) > 0$ beschrieben durch $\sum_{i=1}^n \gamma_f(A)_i e_i^A = \sum_{i=1}^n \langle f | e_i^A \rangle_A e_i^A$.

$\gamma_f(A)$ ist also der Vektor der orthogonalen Projektion der Einschränkung von f auf A bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_A$. Man kann – streng genommen – nur innerhalb des Vektorraums $L^2(\mathcal{B}(A), \mathcal{B}(E))$ von orthogonaler Projektion auf den Raum $V|_A$ sprechen. (Man siehe auch Fußnote zu 3.4.4 i.). Da aber $\dim V|_A = \dim V$ gilt und die Funktion f nur auf A betrachtet wird, ist das Vorgehen im Beweis mit 3.4.4 v. legitim, wenn auch mathematisch vereinfacht dargestellt.

Lemma 3.5.5

Für alle $A \in \sigma(\mathfrak{A})$ gilt:

$$\left\| f - \sum_{i=1}^n \gamma_f(A)_i e_i^A \right\|_A^2 = \|f\|_A^2 - \|\gamma_f(A)\|_2^2$$

Beweis:

Weil $\sum_{i=1}^n \gamma_f(A)_i e_i^A$ die Orthogonalprojektion von $f|_A$ auf V_A bezüglich $\langle \cdot | \cdot \rangle_A$ ist, folgt:

$$\begin{aligned} & \left\| f - \sum_{i=1}^n \gamma_f(A)_i e_i^A \right\|_A^2 \\ &= \left\langle f - \sum_{i=1}^n \gamma_f(A)_i e_i^A \middle| f - \sum_{i=1}^n \gamma_f(A)_i e_i^A \right\rangle_A \\ &= \left\langle f \middle| f - \sum_{i=1}^n \gamma_f(A)_i e_i^A \right\rangle_A - \underbrace{\left\langle \sum_{i=1}^n \gamma_f(A)_i e_i^A \middle| f - \sum_{i=1}^n \gamma_f(A)_i e_i^A \right\rangle_A}_{=0 \text{ (Orthogonalität)}} \end{aligned}$$

$$\begin{aligned}
&= \left\langle f \left| f - \sum_{i=1}^n \gamma_f(A)_i e_i^A \right. \right\rangle_A \\
&= \langle f|f \rangle_A - \left\langle f \left| \sum_{i=1}^n \gamma_f(A)_i e_i^A \right. \right\rangle_A \\
&= \langle f|f \rangle_A - \sum_{i=1}^n \gamma_f(A)_i \langle f|e_i^A \rangle_A \\
&= \langle f|f \rangle_A - \sum_{i=1}^n \gamma_f(A)_i \gamma_f(A)_i
\end{aligned}$$

□

Definition 3.5.6

Die Funktion $\Psi : \mathfrak{A} \rightarrow \mathbb{R}_+$ sei definiert durch

$$\Psi(A) := \|\gamma_f(A)\|_2^2 + \|\gamma_f(M \setminus A)\|_2^2$$

Bemerkung 3.5.7

Mit Hilfe von Satz 3.5.3 kann man nun die Anzahl der Parameter der Funktion Φ zur Suche relativer Minima reduzieren. Es reicht aus, statt nach Minima der Funktion $\Phi : \mathfrak{A} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$, nach Maxima der Funktion $\Psi : \mathfrak{A} \rightarrow \mathbb{R}_+$ zu suchen, denn für $A \in \mathfrak{A}$ mit $\lambda(A) > 0, \lambda(M \setminus A) > 0$ gilt Lemma 3.5.5

$$\begin{aligned}
&\Phi(A, \gamma_f(A)^t, \gamma_f(M \setminus A)^t) \\
&= \left\| f - \sum_{i=1}^n \gamma_f(A) e_i^A \mathbf{1}_A - \sum_{i=1}^n \gamma_f(M \setminus A) e_i^{M \setminus A} \mathbf{1}_{M \setminus A} \right\|_M^2 \\
&= \left\| f - \sum_{i=1}^n \gamma_f(A) e_i^A \right\|_A^2 + \left\| f - \sum_{i=1}^n \gamma_f(M \setminus A) e_i^{M \setminus A} \right\|_{M \setminus A}^2 \\
&= \|f\|_A^2 - \|\gamma_f(A)\|_2^2 + \|f\|_{M \setminus A}^2 - \|\gamma_f(M \setminus A)\|_2^2 = \|f\|_M^2 - \Psi(A)
\end{aligned}$$

Für $\lambda(A) = 0$ ist, wie man leicht nachrechnet,

$$\Phi(A, \mathbf{0}^t, \gamma_f(M)^t) = \sum_{i=1}^n \gamma_f(M)_i^2 = -\Psi(A).$$

Analog gilt $\Phi(A, \gamma_f(A)^t, \mathbf{0}^t) = -\Psi(A)$ für $\lambda(M \setminus A) = 0$.

Bemerkung 3.5.8

Der Vektor $\gamma_f(A)$ ist gegeben in der Koordinatendarstellung bezüglich der ausgezeichneten Basis (e_1^A, \dots, e_n^A) . Diese Basis ist für jede Menge $A \in \mathfrak{A}$ unterschiedlich. Um dennoch eine einheitliche Darstellung der Vektoren $\gamma_f(A)$ zu haben, sollte der Vektor der orthogonalen Projektion in der Koordinatendarstellung bezüglich der Ausgangsbasis $\mathcal{E} = (e_1, \dots, e_n)$ von V betrachtet werden, die unabhängig vom Mengensystem ist.

Definition 3.5.9

Sei $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) > 0$ die Matrix $\mathbf{E}_A^{A, \mathcal{E}} := (\langle e_i^A | e_j \rangle_A)_{1 \leq i, j \leq n}$ definiert.⁶ Dann sei der Vektor $\tilde{\gamma}_f(A) \in \mathbb{R}^n$ definiert durch $\tilde{\gamma}_f(A) := (\mathbf{E}_A^{A, \mathcal{E}})^{-1} \gamma_f(A)$.

Für $A \in \mathfrak{A}$ mit $\lambda(A) = 0$, also $A = O$, sei $\tilde{\gamma}_f(O) := \mathbf{0}$ definiert.

Bemerkung 3.5.10

Für die Matrix $\mathbf{E}_A^{A, \mathcal{E}}$ definiert wie in Definition 3.5.9 gilt:

$$\mathbf{E}_A^{A, \mathcal{E}} \mathcal{K}_{\mathcal{E}}(e_i) = \begin{pmatrix} \langle e_i | e_1^A \rangle_A \\ \vdots \\ \langle e_i | e_n^A \rangle_A \end{pmatrix} = \mathcal{K}_{\mathcal{E}^A}(e_i)$$

Die Matrix $\mathbf{E}_A^{A, \mathcal{E}}$ ist also die Matrix des Basiswechsels von \mathcal{E} nach \mathcal{E}^A . Daher ist $\tilde{\gamma}_f(A)$ für $A \neq O$ der Vektor $\gamma_f(A)$ in der Koordinatendarstellung von \mathcal{E} .

Dieser, von dem jeweiligen Segment A unabhängigen, Darstellung des Vektors $\gamma_f(A)$ wird bei der algorithmischen Bestimmung lokaler Maxima der Funktion Ψ , insbesondere in den Algorithmen des Kapitels 4, eine große Rolle zukommen.

Beispiel 3.5.11

Alle hier aufgeführten Beispiele sind Konkretisierungen von Beispiel 3.4.6. (Vergleiche auch Beispiel 3.4.12)

⁶vgl. auch Definition 3.7.27.

- i) Sei $E := \mathbb{R}$ mit der Multiplikation \cdot als Skalarprodukt und $V := \{\alpha \mid \alpha \in \mathbb{R}\}$ der Vektorraum der konstanten Funktionen. Dann ist (1) eine Basis dieses Raums und für $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) > 0$ ist $e_1^A = \frac{1}{\sqrt{\lambda(A)}}$ die ausgezeichnete Orthonormalbasis bezüglich $\langle \cdot | \cdot \rangle_A$ und (1) nach Definition 3.4.9 (vgl. Beispiel 3.4.12 i)).

Damit ist $\gamma_f(A) = \frac{1}{\sqrt{\lambda(A)}} \int_A f d\lambda$ und nach Bemerkung 3.5.10 $\tilde{\gamma}_f(A) = \frac{1}{\lambda(A)} \int_A f d\lambda$.

Also ergibt sich für die Funktion Ψ :

$$\begin{aligned} \Psi(A) &= \|\gamma_f(A)\|_2^2 + \|\gamma_f(M \setminus A)\|_2^2 \\ &= \frac{1}{\lambda(A)} \left(\int_A f d\lambda \right)^2 + \frac{1}{\lambda(M \setminus A)} \left(\int_{M \setminus A} f d\lambda \right)^2 \\ &= \tilde{\gamma}_f(A) \int_A f d\lambda + \tilde{\gamma}_f(M \setminus A) \int_{M \setminus A} f d\lambda \end{aligned}$$

Die Maximierung dieser Funktion ist die bereits erwähnte 2-Niveaupassung, die Klaus Donner in [Don91] einführt.

Bei genauerer Betrachtung von $\tilde{\gamma}_f(\cdot)$ erkennt man, daß die 2-Niveaupassung genau den Fehler **GU**, der in Abschnitt 2.3.6.1 (S. 27) skizziert wurde, minimiert. Sie ist also bezüglich dieses Qualitätsmaßes optimal.

- ii) Sei $E := \mathbb{R}^d$ mit dem Standardskalarprodukt $\langle \cdot | \cdot \rangle$. $V := \{(\alpha, \dots, \alpha) \mid \alpha \in \mathbb{R}\}$ sei der Vektorraum der konstanten Funktionen mit Basis $((1, \dots, 1))$ (vgl. Beispiel 3.4.12 ii)). Es ist $e_1^A = \frac{1}{\sqrt{d\lambda(A)}}(1, \dots, 1)$ die ausgezeichnete Orthonormalbasis bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_A$ und der Basis $((1, \dots, 1))$ gemäß Definition 3.4.9.

Daher ist für $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) > 0$

$$\gamma_f(A) = \frac{1}{\sqrt{d\lambda(A)}} \int_A \langle f | (1, \dots, 1) \rangle d\lambda = \frac{1}{\sqrt{d\lambda(A)}} \int_A \sum_{i=1}^d f_i d\lambda$$

und es ergibt sich für Ψ :

$$\begin{aligned} \Psi(A) &= \|\gamma_f(A)\|_2^2 + \|\gamma_f(M \setminus A)\|_2^2 \\ &= \frac{1}{d\lambda(A)} \left(\sum_{i=1}^d \int_A f_i d\lambda \right)^2 + \frac{1}{d\lambda(M \setminus A)} \left(\sum_{i=1}^d \int_{M \setminus A} f_i d\lambda \right)^2 \end{aligned}$$

- iii) Sei $E := \mathbb{R}^3$ mit dem Standardskalarprodukt $\langle \cdot | \cdot \rangle$. Es sei $f \in L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^3))$ mit $f = (f_1, f_2, f_3)$. Weiter sei $V := \mathbb{R}^3$ der Vektorraum der konstanten Funktionen mit der Standardbasis $\mathcal{E} := ((1, 0, 0), (0, 1, 0), (0, 0, 1))$. Für $A \in \sigma(\mathfrak{A}) \setminus \{O\}$ ist mit $e_1^A = \frac{1}{\lambda(A)}(1, 0, 0)$, $e_2^A = \frac{1}{\lambda(A)}(0, 1, 0)$, $e_3^A = \frac{1}{\lambda(A)}(0, 0, 1)$ die Basis (e_1^A, e_2^A, e_3^A) die ausgezeichnete Orthonormalbasis bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_A$ und der Basis \mathcal{E} nach Definition 3.4.9 (vgl. Beispiel 3.4.12 iii)).

Für $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) > 0$ ist daher $\gamma_f(A)_1 = \frac{1}{\sqrt{\lambda(A)}} \int_A f_1 d\lambda$, $\gamma_f(A)_2 =$

$$\frac{1}{\sqrt{\lambda(A)}} \int_A f_2 d\lambda, \gamma_f(A)_3 = \frac{1}{\sqrt{\lambda(A)}} \int_A f_3 d\lambda \text{ und damit } \gamma_f(A) = \begin{pmatrix} \gamma_f(A)_1 \\ \gamma_f(A)_2 \\ \gamma_f(A)_3 \end{pmatrix}.$$

Also ergibt sich für Ψ :

$$\begin{aligned}\Psi(A) &= \|\gamma_f(A)\|_2^2 + \|\gamma_f(M \setminus A)\|_2^2 \\ &= \frac{1}{\lambda(A)} \sum_{i=1}^3 \left(\int_A f_i d\lambda \right)^2 + \frac{1}{\lambda(M \setminus A)} \sum_{i=1}^3 \left(\int_{M \setminus A} f_i d\lambda \right)^2\end{aligned}$$

iv) Sei $M := [0, 1]$, $E := \mathbb{R}$ mit der Multiplikation als Skalarprodukt und $V := \{f \mid f(x) = ax + t \text{ mit } a, t \in \mathbb{R}\}$, dann ist (e_1, e_2) mit $e_1(x) := 1$ und $e_2(x) := x$ eine Basis von V .

Weiter sei $\mathfrak{A} := \{[t_1, t_2] \mid t_1, t_2 \in [0, 1], t_1 < t_2\}$ (vgl. Beispiel 3.3.3). Dann ist für $[0, t] \in \mathfrak{A}$ mit $t > 0$ die ausgezeichnete Orthonormalbasis $(e_1^{[0,t]}, e_2^{[0,t]})$ gemäß Definition 3.4.9 gegeben durch

$$e_1^{[0,t]}(x) = \frac{1}{\sqrt{t}} \quad \text{und} \quad e_2^{[0,t]}(x) = \sqrt{\frac{3}{t^3}}(2x - t).$$

Weiter ist $(e_1^{M \setminus [0,t]}, e_2^{M \setminus [0,t]})$ gemäß Definition 3.4.9 gegeben durch

$$e_1^{M \setminus [0,t]}(x) = e_1^{[t,1]}(x) = \frac{1}{\sqrt{1-t}} \quad \text{und} \quad e_2^{M \setminus [0,t]}(x) = e_2^{[t,1]}(x) = \sqrt{\frac{3}{(1-t)^3}}(2x - t - 1).$$

Daraus ergeben sich

$$\begin{aligned}\gamma_f([0, t])_1 &= \frac{1}{\sqrt{t}} \int_{[0,t]} f(x) d\lambda(x) \\ \gamma_f([0, t])_2 &= 2\sqrt{\frac{3}{t^3}} \int_{[0,t]} f(x) \left(x - \frac{1}{2}t\right) d\lambda(x) \\ \gamma_f([t, 1])_1 &= \frac{1}{\sqrt{1-t}} \int_{[t,1]} f(x) d\lambda(x) \\ \gamma_f([t, 1])_2 &= \sqrt{\frac{3}{(1-t)^3}} \int_{[t,1]} f(x) (2x - t - 1) d\lambda(x)\end{aligned}$$

und damit als zu optimierende Funktion Ψ :

$$\begin{aligned}\Psi([0, t]) &= \\ &= \frac{1}{t} \left(\int_{[0,t]} f(x) d\lambda(x) \right)^2 + \frac{12}{t^3} \left(\int_{[0,t]} f(x) \left(x - \frac{1}{2}t\right) d\lambda(x) \right)^2 \\ &+ \frac{1}{1-t} \left(\int_{[t,1]} f(x) d\lambda(x) \right)^2 \\ &+ \frac{3}{(1-t)^3} \left(\int_{[t,1]} f(x) (2x - t - 1) d\lambda(x) \right)^2\end{aligned}$$

Die Matrix $\mathbf{E}_{[0,t]}^{[0,t],\varepsilon}$ ist in diesem Fall gegeben durch

$$\begin{aligned} \begin{pmatrix} \langle e_1^{[0,t]} | e_1 \rangle_{[0,t]} & \langle e_2^{[0,t]} | e_1 \rangle_{[0,t]} \\ \langle e_1^{[0,t]} | e_2 \rangle_{[0,t]} & \langle e_2^{[0,t]} | e_2 \rangle_{[0,t]} \end{pmatrix} &= \begin{pmatrix} \int_{[0,t]} \frac{1}{\sqrt{t}} d\lambda(x) & \int_{[0,t]} \frac{1}{\sqrt{t}} x d\lambda(x) \\ \int_{[0,t]} \sqrt{\frac{3}{t^3}} (2x-t) d\lambda(x) & \int_{[0,t]} \sqrt{\frac{3}{t^3}} (2x-t) x d\lambda(x) \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{t} & \frac{\sqrt{t^3}}{2} \\ 0 & \frac{\sqrt{3t^3}}{6\sqrt{t^3}} \end{pmatrix} \end{aligned}$$

Für die Inverse von $\mathbf{E}_{[0,t]}^{[0,t],\varepsilon}$ gilt

$$\left(\mathbf{E}_{[0,t]}^{[0,t],\varepsilon} \right)^{-1} = \begin{pmatrix} \frac{1}{\sqrt{t}} & -\sqrt{\frac{3}{t}} \\ 0 & 2\sqrt{\frac{3}{t^3}} \end{pmatrix}$$

Daher ergibt sich für $\tilde{\gamma}_f([0,t])$:

$$\tilde{\gamma}_f([0,t]) = \begin{pmatrix} \frac{1}{t} \int_{[0,1]} f(x) d\lambda(x) - \frac{6}{t^2} \int_{[0,1]} f(x) (x - \frac{1}{2}t) d\lambda(x) \\ \frac{12}{t^3} \int_{[0,t]} f(x) (x - \frac{1}{2}t) d\lambda(x) \end{pmatrix}$$

Bemerkung 3.5.12

Im vorangegangenen Beispiel 3.4.12 iii) wurde die Funktion f aus dem Raum $L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^3))$ gewählt. Offensichtlich läßt sich das Beispiel auf beliebige endlich dimensionale Wertebereiche \mathbb{R}^d erweitern. Die Einschränkung auf den \mathbb{R}^3 erfolgt, weil dieser Fall in der Praxis der Segmentierung von Farbbildern häufig auftaucht. Denn Farbbilder lassen sich als Funktionen von M in den \mathbb{R}_+^3 kodieren, indem man sie als 3-kanalige Bilder modelliert (vgl. Abschnitt 1.3.2).

Man könnte auch im Beispiel ii) die Dimension des Wertebereichs von f auf $d = 3$ setzen. In diesem Fall würde man ein Farbbild durch ein stückweise konstantes Grauwertbild approximieren, da die Werte $(\alpha, \alpha, \alpha) \in \mathbb{R}_+^3$ im RGB-Farbwürfel die Grauwerte kodieren.

Definition 3.5.13 (f_A)

Mit den Bezeichnungen von oben und $A \in \sigma(\mathfrak{A})$ sei die Funktion $f_A : M \rightarrow E$ definiert durch

$$f_A := \begin{cases} \sum_{i=1}^n \gamma_f(A)_i e_i^A \mathbf{1}_A & , \lambda(A) > 0 \\ 0 & , \text{sonst} \end{cases}$$

Bemerkung 3.5.14

Offensichtlich gilt mit Bemerkung 3.5.10 $f_A = \sum_{i=1}^n \tilde{\gamma}_f(A)_i e_i \mathbf{1}_A$.

Beispiel 3.5.15

Im Kontext des Beispiels 3.5.11 iv) ist für $t > 0$:

$$\begin{aligned}
f_{[0,t]}(x) &= \gamma_f([0,t])_1 e_1^{[0,t]}(x) + \gamma_f([0,t])_2 e_2^{[0,t]}(x) \\
&= \frac{1}{\sqrt{t}} \int_{[0,t]} f(\chi) d\lambda(\chi) \frac{1}{\sqrt{t}} + 2\sqrt{\frac{3}{t^3}} \int_{[0,t]} f(\chi) \left(\chi - \frac{1}{2}t\right) d\lambda(\chi) \sqrt{\frac{3}{t^3}} (2x - t) \\
&= \frac{1}{t} \int_{[0,t]} f(\chi) d\lambda(\chi) - \frac{6}{t^2} \int_{[0,t]} f(\chi) \left(\chi - \frac{1}{2}t\right) d\lambda(\chi) \\
&\quad + \frac{12}{t^3} \left(\int_{[0,t]} f(\chi) \left(\chi - \frac{1}{2}t\right) d\lambda(\chi) \right) x \\
&= \tilde{\gamma}_f([0,t])_1 e_1 + \tilde{\gamma}_f([0,t])_2 e_2
\end{aligned}$$

wie man leicht mit Beispiel 3.5.11 iv) zeigt.

Satz 3.5.16 (Neuformulierung des Problems für zwei Niveaus)

(A, g_1, g_2) ist ein (lokales) Minimum der Funktion

$$\begin{aligned}
\tilde{\Phi} : \mathfrak{A} \times V \times V &\rightarrow \mathbb{R} \\
(B, g_1, g_2) &\mapsto \|f - g_1 \mathbf{1}_B - g_2 \mathbf{1}_{M \setminus B}\|_M^2
\end{aligned}$$

genau dann, wenn $g_1 \mathbf{1}_A = f_A$ und $g_2 \mathbf{1}_{M \setminus A} = f_{M \setminus A}$ gilt, und $A \in \mathfrak{A}$ ein (lokales) Minimum der Funktion

$$\hat{\Phi} : \mathfrak{A} \rightarrow \mathbb{R}_+ \\
B \mapsto \|f - f_B - f_{M \setminus B}\|_M^2$$

ist,

genau dann, wenn A ein (lokales) Maximum der Funktion

$$\Psi : \mathfrak{A} \rightarrow \mathbb{R}_+ \\
B \mapsto \|\gamma_f(B)\|_2^2 + \|\gamma_f(M \setminus B)\|_2^2$$

ist.

Beweis:

Folgt aus Bemerkung 3.5.7, Satz 3.5.3 und der Definition von f_A in Definition 3.5.13.

3.5.2 Die allgemeine Problemformulierung – optimale Segmentierung

Die im vorangegangenen Abschnitt aufgestellten Konzepte für die Zwei-Niveau-Situation lassen sich leicht auf die k -Niveau-Situation übertragen. Allerdings kann man in diesem allgemeineren Kontext nicht mehr zwischen Vordergrundniveau $A \in \mathfrak{A}$, das aus einem erlaubten Mengensystem \mathfrak{A} stammt, und Hintergrundniveau $M \setminus A$, das nicht notwendig in dem erlaubten Mengensystem \mathfrak{A} enthalten ist, unterscheiden. Daher wird das Optimierungsproblem der k -Niveaupassung nicht auf erlaubten Mengensystemen, sondern auf dem System der erlaubten Segmentierungen gemäß Definition 3.3.1 definiert.

Definition 3.5.17 (Menge der Segmentierungen auf M)

Sei $\mathfrak{S}_k(\mathfrak{A})$ ein System von Segmentierungen von M durch k Mengen aus \mathfrak{A} nach Definition 3.3.1 und V eine Klasse von erlaubten Funktionen gemäß 3.4.2. Dann heie die Menge

$$\mathcal{F}(\mathfrak{S}_k(\mathfrak{A}), V) := \{h \in L^2(\mathcal{B}(M), \mathcal{B}(E)) \mid \exists \mathbf{A} \in \mathfrak{S}_k(\mathfrak{A}) : \forall A \in \mathbf{A}, A \neq O : h|_A \in V|_A\}$$

die Menge der Segmentierungen auf M durch k erlaubte Segmente und Funktionen aus V .

Bemerkung 3.5.18

Im folgenden soll nun der Begriff „optimale Segmentierung“ definiert werden. An dieser Stelle spielt auch die Wahl der Familie der erlaubten Skalarprodukte (vgl. Definition 3.4.4) eine entscheidende Rolle, da der Abstand, bezuglich dessen approximiert wird, von der Norm $\|\cdot\|_M = \sqrt{\langle \cdot | \cdot \rangle_M}$ induziert wird. Mit der Bedingung iv. aus der Definition der erlaubten Skalarprodukte 3.4.4 ist es dann mglich, diesen Abstand auf jedem Segment einer Segmentierung einzeln zu berechnen bzw. die Approximation auf jedem Segment einzeln auszufhren. Denn fr jede Segmentierung $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ gilt $\|\cdot\|_M^2 = \sum_{A \in \mathbf{A}} \|\cdot\|_A^2$.⁷

Definition 3.5.19 (Optimale Segmentierung)

Sei $\{\langle \cdot | \cdot \rangle_A \mid A \in \sigma(\mathfrak{A}), \lambda(A) > 0\}$ eine Familie von erlaubten Skalarprodukten gem Definition 3.4.4. Ist dann $s_{k,0} \in \mathcal{F}(\mathfrak{S}_k(\mathfrak{A}), V)$ mit

$$\|s_{k,0} - f\|_M = \min_{s_k \in \mathcal{F}(\mathfrak{S}_k(\mathfrak{A}), V)} \|s_k - f\|_M$$

so heie $s_{k,0}$ eine *optimale Segmentierung von f durch k erlaubte Mengen des Mengensystem \mathfrak{A} und Funktionen aus der Klasse der erlaubten Funktionen V* .

⁷Dabei sei $\|\cdot\|_O := 0$ definiert.

Bemerkung 3.5.20

Oft wird in dieser Arbeit vereinfachend von einer „optimalen Segmentierung“ $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ einer Funktion f gesprochen. Man mache sich aber an dieser Stelle klar, daß es sich dabei nur um eine legere, abkürzende Schreibweise handeln kann, da der zugehörige Vektorraum der erlaubten Funktionen und die Familie von erlaubten Skalarprodukten weggelassen wurden. Diese fließen jedoch immer in den Begriff „optimale Segmentierung“ mit ein, wie in Definition 3.5.19 eindeutig zu sehen ist.

Die Existenz der optimalen Segmentierung wird in Unterkapitel 3.6 gezeigt.

Satz 3.5.21

Sei $s_{k,0} \in \mathcal{F}(\mathfrak{S}_k(\mathfrak{A}), V)$ mit $s_{k,0} = \sum_{A \in \mathbf{A}} h_A \mathbf{1}_{A_j}$ eine optimale Segmentierung von f mit Funktionen aus V bezüglich des Mengensystems \mathfrak{A} .

Dann gilt :

$$\forall A \in \mathbf{A} : A \neq O \Rightarrow h_A = \sum_{i=1}^n \gamma_f(A)_i e_i^A$$

Für $A \in \mathbf{A}$ mit $A = O$ ist die Funktion h_A ohne Bedeutung, da ihr Träger Lebesgue-Maß null hat.

Es ist also $s_{k,0} = \sum_{A \in \mathbf{A}} f_A$.

Beweis:

Der Beweis erfolgt vollkommen analog zum Beweis des Satzes 3.5.3 über orthogonale Projektion bezüglich der Skalarprodukte $\langle \cdot | \cdot \rangle_A$ (vgl. auch Bemerkung 3.5.4 und Definition 3.5.13).

□

Bemerkung 3.5.22

Nicht nur Satz 3.5.3, sondern auch die Bemerkungen 3.5.1 und 3.5.7 lassen sich kanonisch auf die k -Niveau-Situation übertragen. Die k -Niveau-Situation erfordert allerdings, daß alle Segmente aus dem erlaubten Mengenmodell sind. Daher wird die k -Niveaupassung auf etwas andere Art parametrisiert als die 2-Niveaupassung: Gesucht ist eine Funktion $s_{k,0} \in \mathcal{F}(\mathfrak{S}_k(\mathfrak{A}), V)$, die die Funktion

$$\tilde{\Phi} : \begin{array}{l} \mathcal{F}(\mathfrak{S}_k(\mathfrak{A}), V) \rightarrow \mathbb{R}_+ \\ h \mapsto \|f - h\|_M^2 \end{array}$$

minimiert. Dabei ist eine Segmentierung $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ und für jedes $A \in \mathbf{A}$ eine Funktion aus V zu bestimmen, die diese Minimalfunktion $s_{k,0}$ beschreiben. Dieses Problem läßt sich nun, völlig analog zur 2-Niveau-Situation, auf ein Problem reduzieren, das nur über die Segmentierungen parametrisiert ist. Die Existenz des Minimums wird dabei im Unterkapitel 3.6 nachgewiesen.

Satz 3.5.23 (Neuformulierung des Problems für k -Niveaus)

Sei $g \in \mathcal{F}(\mathfrak{S}_k(\mathfrak{A}), V)$. Dann existiert eine Segmentierung $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ mit $g = \sum_{A \in \mathbf{A}} g_A \mathbf{1}_A$, wobei für alle $A \in \mathbf{A} \setminus \{O\}$ die Funktion g_A aus der Klasse der erlaubten Funktionen V ist.

Es gilt:

g ist ein (lokales) Minimum der Funktion

$$\tilde{\Phi} : \mathcal{F}(\mathfrak{S}_k(\mathfrak{A}), V) \rightarrow \mathbb{R}_+ \\ h \mapsto \|f - h\|_M^2$$

genau dann, wenn $\forall A \in \mathbf{A} \setminus \{O\} : g \mathbf{1}_A = f_A$ gilt und $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ (lokales) Minimum der Funktion

$$\hat{\Phi} : \mathfrak{S}_k(\mathfrak{A}) \rightarrow \mathbb{R}_+ \\ \mathbf{B} \mapsto \|f - \sum_{B \in \mathbf{B}} f_B\|_M^2$$

ist,

genau dann, wenn \mathbf{A} ein (lokales) Maximum der Funktion

$$\Psi : \mathfrak{S}_k(\mathfrak{A}) \rightarrow \mathbb{R}_+ \\ \mathbf{B} \mapsto \sum_{B \in \mathbf{B}} \|\gamma_f(B)\|_2^2$$

ist.

Beweis:

Eine Durchsicht der Beweise aus dem vorangegangenen Abschnitt 3.5.1, die zum Beweis des Satzes 3.5.16 führen, zeigt, daß sich nicht nur der Beweis des Satzes 3.5.3, sondern auch alle anderen Beweise und Bemerkungen leicht auf die Situation mit k Niveaumengen anstelle von zwei Niveaumengen übertragen lassen. □

Definition 3.5.24 (Zielfunktion einer k -Niveaupassung)

Sei $f \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ ein Bild.

Sei $\{\langle \cdot | \cdot \rangle_A \mid A \in \sigma(\mathfrak{A}), \lambda(A) > 0\}$ eine Familie von erlaubten Skalarprodukten gemäß Definition 3.4.4.

Sei V eine Klasse von erlaubten Funktionen gemäß Definition 3.4.2.

Sei \mathfrak{A} ein Mengensystem, das die Bedingungen von Definition 3.2.10 erfüllt.

Dann heie die Funktion

$$\Psi : \mathfrak{S}_k(\mathfrak{A}) \rightarrow \mathbb{R}_+ \\ \mathbf{B} \mapsto \sum_{B \in \mathbf{B}} \|\gamma_f(B)\|_2^2$$

die Zielfunktion einer k -Niveaupassung der Funktion f .

Das Problem, ein (lokales) Maximum der Funktion Ψ zu finden, heie k -Niveaupassungsproblem.⁸

⁸Die Familie der erlaubten Skalarprodukte ist dabei in der Definition von γ_f (Definition 3.5.2) versteckt.

Bemerkung 3.5.25

Die in Definition 3.5.24 definierte Funktion Ψ ist das Kernstück dieser Arbeit. Nach Satz 3.5.23 ist das Maximum dieser Funktion die optimale Segmentierung eines Bildes bezüglich eines gegebenen Mengen- und Funktionensystems und einer gegebenen Familie von erlaubten Skalarprodukten. Daher ist die Bezeichnung k -Niveaupassungsproblem für die Maximierung der Funktion Ψ gerechtfertigt.

Das Ziel des nächsten Abschnitts 3.6 ist es, die Existenz dieses Maximums zu zeigen. Im darauf folgenden Abschnitt wird eine Charakterisierung von lokalen Maxima von Ψ gegeben. In Kapitel 4 werden die Ergebnisse dieses Kapitels dazu genutzt, Algorithmen zur Segmentierung von diskreten Pixelbildern zu formulieren.

3.6 Existenz der Lösung

Um die Existenz einer Lösung nachzuweisen, wird gezeigt, daß die Zielfunktion der k -Niveaupassung Ψ aus Definition 3.5.24 eine stetige Funktion auf einem kompakten Definitionsbereich ist. Daher nimmt die Funktion auf ihrem Definitionsbereich sowohl ihr Maximum als auch ihr Minimum an.

Es sind also zwei Dinge zu zeigen:

- i. $\mathfrak{S}_k(\mathfrak{A})$, der Definitionsbereich von Ψ , ist kompakt. Bei dem Beweis dieser Behauptung fällt dem Satz von Fréchet und Kolmogorov aus der Funktionalanalysis eine entscheidende Rolle zu.
- ii. Die Funktion Ψ ist stetig bezüglich \mathfrak{d} , der Metrik auf dem System der Segmentierungen.

3.6.1 Der Satz von Fréchet und Kolmogorov

Der Satz von Fréchet und Kolmogorov gibt Bedingungen an, unter denen eine Funktionenfamilie relativ kompakt ist. Im Abschnitt 3.6.2 werden einzelne Segmentierungen durch Funktionen kodiert. Die Menge aller Funktionen, die nach dieser Kodierungsvorschrift eine Segmentierung bezeichnen, erfüllt dann die Voraussetzungen des Satzes von Fréchet und Kolmogorov und ist somit relativ kompakt.

Definition 3.6.1 (Relativ kompakt)

Eine Teilmenge A eines topologischen Raums heißt *relativ kompakt*, wenn der Abschluß der Menge A kompakt ist.

Lemma 3.6.2

Sei (M, δ) ein kompakter metrischer Raum. Eine Teilmenge $A \subset M$ ist genau dann relativ kompakt, wenn für jedes $\varepsilon > 0$ eine endliche Menge $\{a_1, \dots, a_n\}$ in A existiert, so daß für jedes $a \in A$ ein $a_i \in \{a_1, \dots, a_n\}$ existiert mit $\delta(a, a_i) < \varepsilon$.

A kann also durch eine endliche Menge von offenen Kugeln mit Radius ε und Zentrum in A überdeckt werden.

Beweis:

vgl. [Yos78], S. 13

Satz 3.6.3 (Ascoli-Arzelá)

Sei (M, δ) ein kompakter metrischer Raum und $\mathcal{C}(M, \mathbb{R}^d)$ der Banachraum der stetigen Funktionen auf M mit Werten in \mathbb{R}^d , versehen mit der Supremumsnorm.

Eine Teilmenge $\mathcal{F} \subset \mathcal{C}(M, \mathbb{R}^d)$ ist relativ kompakt in $\mathcal{C}(M, \mathbb{R}^d)$, falls

- i. \mathcal{F} gleichmäßig beschränkt ist, und
- ii. \mathcal{F} gleichgradig stetig ist, d. h.

$$\forall \varepsilon > 0 : \exists \zeta > 0 : \forall f \in \mathcal{F} : \forall s, t \in M : \delta(s, t) \leq \zeta \Rightarrow \|f(s) - f(t)\|_2 \leq \varepsilon$$

Beweis:

vgl. z. B. [Wer95], S. 68f. Die Übertragung auf den vektorwertigen Fall erfolgt kanonisch.

Satz 3.6.4 (Fréchet und Kolmogorov)

Sei $M \in \mathcal{B}(\mathbb{R}^k)$ kompakt mit $\lambda(M) < \infty$. Sei $1 \leq p < \infty$ und sei $\mathcal{F} \subset L^p(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$ eine Menge von Funktionen, wobei für alle $f \in \mathcal{F}$ gilt:

$$\sup_{f \in \mathcal{F}} \|f\|_p < \infty \tag{3.4}$$

$$\lim_{\|t\|_\infty \rightarrow 0} \int_M \|f(x+t) - f(x)\|_2^p d\lambda(x) = 0, \quad \text{gleichmäßig in allen } f \in \mathcal{F} \tag{3.5}$$

$$\lim_{\alpha \rightarrow \infty} \int_{\|x\|_2 > \alpha} \|f(x)\|_2^p d\lambda(x) = 0, \quad \text{gleichmäßig in allen } f \in \mathcal{F} \tag{3.6}$$

Dann ist \mathcal{F} relativ kompakt in $L^p(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$

Beweis:

Der Beweis erfolgt vollkommen analog zum Beweis des Satzes in [Yos78], S. 275f. Dort werden allerdings nicht der vektorwertige Fall und der mehrdimensionale Definitionsbereich untersucht, was insbesondere den Beweis von Behauptung 1 (s.u.) erschwert.

Zunächst wird ein Translationsoperator T_t definiert: Für $t \in \mathbb{R}^d$ und $f = (f_1, \dots, f_d) \in \mathcal{F}$ sei $(T_t f)(x) = f(x+t)$. Da f nur auf M definiert ist, wird, um die Arbeit zu vereinfachen, $f(x) = 0$ für alle $x \in \mathbb{R}^k \setminus M$ gesetzt. Mit der Voraussetzung (3.5) folgt dann

$$\lim_{\|t\|_2 \rightarrow 0} T_t f = f \text{ gleichmäßig in } f \in \mathcal{F} \tag{3.7}$$

Weiter wird eine Mittelwertfunktion $M_a : \mathcal{F} \rightarrow L^p(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$ definiert durch

$$(M_a f)_i(x) := \frac{1}{(2a)^k} \int_{R_a} \|(T_t f)_i(x)\|_2 d\lambda(t) \quad (\text{mit } R_a := [-a, a]^k)$$

Behauptung 1:

$$\|(M_a f)(x) - f(x)\|_2 \leq \int_{R_a} \frac{1}{(2a)^k} \|f(x+t) - f(x)\|_2 d\lambda(t) \quad (3.8)$$

Bemerkung: Da f aus $L^p(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$ ist, kann man f eigentlich nicht eindeutig punktweise auswerten. Daher sei für den Beweis dieser Behauptung \hat{f} eine Funktion aus der Äquivalenzklasse von f .

Beweis von Behauptung 1:

$$\begin{aligned} & \| (M_a f)(x) - f(x) \|_2 \\ &= \sqrt{\sum_{i=1}^d \left(\int_{R_a} (M_a f)_i(x) d\lambda(t) - \hat{f}_i(x) \right)^2} \\ &= \sqrt{\sum_{i=1}^d \left(\int_{R_a} \frac{1}{(2a)^k} \hat{f}_i(x+t) d\lambda(t) - \frac{1}{(2a)^k} \int_{R_a} \hat{f}_i(x) d\lambda(t) \right)^2} \\ &= \sqrt{\sum_{i=1}^d \left(\frac{1}{(2a)^k} \int_{R_a} \hat{f}_i(x+t) d\lambda(t) - \hat{f}_i(x) \right)^2} \\ (M) \quad &\leq \int_{R_a} \sqrt{\sum_{i=1}^d \left(\frac{1}{(2a)^k} (\hat{f}_i(x+t) - \hat{f}_i(x)) \right)^2} d\lambda(t) \\ &= \int_{R_a} \frac{1}{(2a)^k} \left\| \hat{f}(x+t) - \hat{f}(x) \right\|_2 d\lambda(t) \end{aligned}$$

Wobei die Ungleichung (M) mit der kontinuierlichen Version der Ungleichung von Minkowski (vgl. [Flo81], S. 225) folgt.

Damit ist Behauptung 1 bewiesen.

Bemerkung: Man beachte, daß es durchaus Funktionen \hat{f} aus der Äquivalenzklasse f geben kann, mit $\hat{f}_i(x) = \infty$ für ein $x \in M$ und ein $i \in \{1, \dots, d\}$.

Behauptung 2:

$$\|M_a f - f\|_p \leq \left(\frac{1}{(2a)^k} \int_{R_a} \int_M \|f(x+t) - f(x)\|_2^p d\lambda(x) d\lambda(t) \right)^{\frac{1}{p}} \quad (3.9)$$

Beweis von Behauptung 2: Sei q derart, daß $\frac{1}{p} + \frac{1}{q} = 1$. Mit Hilfe der Hölderschen Ungleichung (H) und des Satzes von Fubini (F) gilt dann:

$$\begin{aligned}
& \|M_a f - f\|_p \\
&= \left(\int_M \|(M_a f)(x) - f(x)\|_2^p d\lambda(x) \right)^{\frac{1}{p}} \\
(3.8) \quad &\leq \left(\int_M \left(\frac{1}{(2a)^k} \int_{R_a} \|f(x+t) - f(x)\|_2^p d\lambda(t) \right)^p d\lambda(x) \right)^{\frac{1}{p}} \\
(H) \quad &\leq \frac{1}{(2a)^k} \left(\int_M \left(\left(\int_{R_a} \|f(x+t) - f(x)\|_2^p d\lambda(t) \right)^{\frac{1}{p}} \left(\int_{R_a} 1^q d\lambda \right)^{\frac{1}{q}} d\lambda(s) \right)^p \right)^{\frac{1}{p}} \\
&= \frac{1}{(2a)^k} \left(\int_M \left(\int_{R_a} \|f(x+t) - f(x)\|_2^p d\lambda(t) \right) \left((2a)^k \right)^{\frac{p}{q}} d\lambda(s) \right)^{\frac{1}{p}} \\
&\leq \left(\frac{1}{(2a)^k} \int_M \int_{R_a} \|f(x+t) - f(x)\|_2^p d\lambda(t) d\lambda(s) \right)^{\frac{1}{p}} \\
(F) \quad &= \left(\frac{1}{(2a)^k} \int_{R_a} \int_M \|f(x+t) - f(x)\|_2^p d\lambda(s) d\lambda(t) \right)^{\frac{1}{p}}
\end{aligned}$$

Daher folgt Behauptung 2.

Mit Gleichung (3.9) folgt

$$\|M_a f - f\|_p \leq \sup_{\|t\|_\infty \leq a} \|T_t f - f\|_p$$

und damit gilt mit Ergebnis (3.7):

$$\lim_{a \rightarrow 0} M_a f = f \quad \text{gleichm\u00e4\u00dfig in } f \in \mathcal{F}. \quad (3.10)$$

Behauptung 3:

$$\{M_a f \mid f \in \mathcal{F}\} \quad \text{ist gleichgradig stetig.}$$

Beweis von Behauptung 3:

Analog zum Beweis der Gleichung (3.9) zeigt man:

$$\begin{aligned}
\|(M_a f)(s_1) - (M_a f)(s_2)\|_2 &\leq \frac{1}{(2a)^k} \int_{R_a} \|f(s_1+t) - f(s_2+t)\|_2 d\lambda(t) \\
&\leq \left(\frac{1}{(2a)^k} \int_{R_a} \|f(s_1+t) - f(s_2+t)\|_2^p d\lambda(t) \right)^{\frac{1}{p}}
\end{aligned}$$

Daher folgt die Behauptung 3 mit der Voraussetzung (3.5).

Behauptung 4:

$$\{M_a f \mid f \in \mathcal{F}\} \quad \text{ist gleichmäßig beschränkt.}$$

Behauptung 4 folgt aus der Definition von $M_a f$.

Aus Behauptung 3 und 4 folgt mit dem Satz von Ascoli-Arzelá 3.6.3 die relative Kompaktheit von $\{M_a f \mid f \in \mathcal{F}\}$ in $\mathcal{C}(M, \mathbb{R}^d)$. Es bleibt noch zeigen, daß $\{M_a f \mid f \in \mathcal{F}\}$ relativ kompakt in $L^p(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$ ist. Dazu wird Lemma 3.6.2 genutzt: Mit diesem Lemma folgt nämlich, daß für jedes $a > 0$ eine endliche Anzahl von Funktionen $M_a f_1, M_a f_2, \dots, M_a f_n$ mit $f_i \in \mathcal{F}$ existiert, so daß

$$\exists \alpha > 0 : \forall f \in \mathcal{F} : \exists j \in \{1, \dots, n\} : \sup_{\|s\|_\infty \leq \alpha} \|(M_a f)(s) - (M_a f_j)(s)\|_2 < \varepsilon$$

gilt. Um zu zeigen, daß $\{M_a f \mid f \in \mathcal{F}\}$ relativ kompakt in $L^p(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$ ist, muß die Supremumsnorm des Raums $\mathcal{C}(M, \mathbb{R}^d)$ durch die Norm $\|\cdot\|_p$ des Raums $L^p(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$ ersetzt werden.

Wähle für $f \in \mathcal{F}$ beliebig ein j , so daß $\sup_{\|s\|_\infty \leq \alpha} \|(M_a f)(s) - (M_a f_j)(s)\|_2 < \varepsilon$ gilt.

Damit folgt:

$$\begin{aligned} \|M_a f - M_a f_j\|_p^p &= \int_M \|M_a f(s) - M_a f_j(s)\|_2^p d\lambda(s) \\ &= \int_{\|s\|_\infty \leq \alpha} \|M_a f(s) - M_a f_j(s)\|_2^p d\lambda(s) \\ &\quad + \int_{\|s\|_\infty > \alpha} \|M_a f(s) - M_a f_j(s)\|_2^p d\lambda(s) \\ &\leq (2\alpha)^k \varepsilon^p + \int_{\|s\|_\infty > \alpha} \|M_a f(s) - M_a f_j(s)\|_2^p d\lambda(s) \end{aligned} \quad (3.11)$$

Weiter gilt

$$\forall \varepsilon > 0 : \exists (a, \alpha) \in \mathbb{R} \times \mathbb{R}^k : \int_{\|s\|_\infty > \alpha} \|M_a f(s) - M_a f_j(s)\|_2^p d\lambda(s) \leq \varepsilon \quad (3.12)$$

Denn mit der Ungleichung von Minkowski folgt:

$$\begin{aligned} &\int_{\|s\|_\infty > \alpha} \|M_a f(s) - M_a f_j(s)\|_2^p d\lambda(s) \\ &< \left(\|M_a f - f\|_p \right. \end{aligned} \quad (3.13)$$

$$\left. + \left(\int_{\|s\|_2 > \alpha} \|f(s) - f_j(s)\|_2^p d\lambda(s) \right)^{\frac{1}{p}} \right)^p \quad (3.14)$$

$$\left. + \left(\int_{\|s\|_2 > \alpha} \|f_j(s) - (M_a f_j)(s)\|_2^p d\lambda(s) \right)^{\frac{1}{p}} \right)^p \quad (3.15)$$

Für den Term (3.13) gilt mit Voraussetzung (3.6) und Gleichung (3.10): $\forall \varepsilon > 0 : \exists a > 0 : \|M_a f - f\|_p < \frac{\varepsilon}{3}$. Für festes a existiert für jedes ε ein $\alpha > 0$, so daß die Terme (3.14) und (3.15) kleiner $\frac{\varepsilon}{3}$ sind. Daher folgt die Gleichung (3.12). Und aus Gleichung (3.11) ergibt sich:

$$\forall \varepsilon > 0 : \exists a > 0 : \|M_a f - M_a f_j\|_p^p < \varepsilon$$

Alle Abschätzungen gelten gleichmäßig in $f \in \mathcal{F}$, woraus die relative Kompaktheit von $\{M_a f \mid f \in \mathcal{F}\}$ in $L^p(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$ für ein ausreichend kleines $a > 0$ folgt.

Da mit der Gleichung (3.10) für $a \rightarrow 0$ die Menge $\{M_a f \mid f \in \mathcal{F}\}$ gleichmäßig gegen \mathcal{F} konvergiert, folgt die Behauptung des Satzes. \square

3.6.2 Der metrische Raum $\mathfrak{Z}_k(\mathfrak{A})$

Für den Beweis der Existenz einer optimalen Segmentierung ist die Kompaktheit von $\mathfrak{S}_k(\mathfrak{A})$ von zentraler Bedeutung. Um jedoch die Kompaktheit des Systems der Segmentierungen zu zeigen, wird an dieser Stelle das System der geordneten Segmentierungen eingeführt. Im nächsten Abschnitt wird dann gezeigt, daß das System der geordneten Segmentierungen kompakt ist. Aus der Kompaktheit dieses Systems kann man dann die Kompaktheit von $\mathfrak{S}_k(\mathfrak{A})$ folgern.

Definition 3.6.5 (System der geordneten Segmentierungen)

Die Menge $\mathfrak{Z}_k(\mathfrak{A}) \subset \mathfrak{A}^k$ definiert durch

$$\mathfrak{Z}_k(\mathfrak{A}) := \{(A_1, \dots, A_k) \mid \biguplus_{i=1}^k A_i = M, A_i \in \mathfrak{A}, 1 \leq i \leq k\}$$

heißt das System der geordneten Segmentierungen von M durch k erlaubte Mengen aus \mathfrak{A} .

Bemerkung 3.6.6

Man könnte an dieser Stelle fragen, warum dieses System der geordneten Segmentierungen von M durch k erlaubte Mengen aus \mathfrak{A} nicht zur Definition des Begriffes optimale Segmentierung 3.5.19 eingeführt wurde. Sicher spricht für dieses System, daß sich Dinge wie Kompaktheit und Stetigkeit der Funktion $\mathfrak{Z}_k(\mathfrak{A}) \rightarrow \mathbb{R}_+, (A_1, \dots, A_k) \mapsto \sum_{i=1}^k \|\gamma_f(A_i)\|_2^2$ leichter zeigen lassen. Der Grund für die vereinfachte Beweisführung liegt in der Ordnung der Segmente, die durch die Tupel-Schreibweise induziert wird. Diese Ordnung hat allerdings auch einen entscheidenden Nachteil: Es gilt $([0, 1], [1, 2]) \neq ([1, 2], [0, 1])$, obwohl diese beiden Elemente eines Systems geordneter Segmentierung die gleiche Zerlegung von $[0, 2]$ beschreiben. Bei diesem speziellen Mengenmodell erscheint es noch einfach, solche Phänomene durch die Einführung einer Ordnung auf den Segmenten auszuschließen. Bei komplexeren Mengenmodellen ist dies jedoch nicht möglich.

Definition 3.6.7 (Metrik des Systems der geordneten Segmentierungen)

Die Metrik δ_Z sei durch

$$\delta_Z : (\mathcal{B}(M)/\sim)^k \rightarrow \mathbb{R}_+ \\ ((A_1, \dots, A_k), (B_1, \dots, B_k)) \mapsto \sum_{i=1}^k \mathfrak{d}(A_i, B_i)$$

definiert. Schränkt man den Definitionsbereich von δ_Z auf ein System geordneter Segmentierungen ein, so heie

$$\delta_Z : \mathfrak{Z}_k(\mathfrak{A}) \times \mathfrak{Z}_k(\mathfrak{A}) \rightarrow \mathbb{R}_+ \\ ((A_1, \dots, A_k), (B_1, \dots, B_k)) \mapsto \sum_{i=1}^k \mathfrak{d}(A_i, B_i)$$

die Metrik des Systems der geordneten Segmentierungen von M durch k erlaubte Mengen aus \mathfrak{A} .

Lemma 3.6.8

δ_Z ist eine Metrik auf $(\mathcal{B}(M)/\sim)^k$.

Beweis:

folgt aus den Metrikeigenschaften von \mathfrak{d} auf $(\mathcal{B}(M)/\sim)$.

Beispiel 3.6.9

Um das Beispiel aus der Bemerkung 3.6.6 aufzugreifen: Mit der Metrik δ_Z gilt $\delta_Z(([0, 1], [1, 2]), ([1, 2], [0, 1])) = 4$.

Satz 3.6.10

Die Abbildung

$$((\mathcal{B}(M)/\sim)^k, \delta_Z) \rightarrow (\mathcal{P}_k(\mathcal{B}(M)/\sim), \mathfrak{d}) \\ (A_1, \dots, A_k) \mapsto \{A_1, \dots, A_k\}$$

ist stetig und surjektiv, aber nicht injektiv.

Beweis:

Die Surjektivitt ist trivial zu zeigen. Da die Abbildung nicht injektiv ist, zeigt Bemerkung 3.6.6. Zum Beweis der Stetigkeit seien (A_1, \dots, A_k) und (B_1, \dots, B_k) aus $\mathfrak{Z}_k(\mathfrak{A})$ mit $\delta_Z((A_1, \dots, A_k), (B_1, \dots, B_k)) < \varepsilon$. Dann gilt

$$\begin{aligned} \sum_{i=1}^k \mathfrak{d}(A_i, B_i) < \varepsilon &\Rightarrow \sum_{i=1}^k \mathfrak{d}(A_i, B_i) + \sum_{i=1}^k \mathfrak{d}(B_i, A_i) < 2\varepsilon \\ &\Rightarrow \sum_{i=1}^k \min_{1 \leq j \leq n} \mathfrak{d}(A_i, B_j) + \sum_{i=1}^k \min_{1 \leq j \leq n} \mathfrak{d}(B_i, A_j) < 2\varepsilon \\ \text{[Definition 3.3.5]} &\Rightarrow \mathfrak{d}(\{A_1, \dots, A_k\}, \{B_1, \dots, B_k\}) < 2\varepsilon \end{aligned}$$

und damit die Behauptung. □

Korollar 3.6.11

Ist \mathfrak{A} abgeschlossen in $\mathcal{B}(M)/\sim$, so ist $\mathfrak{Z}_k(\mathfrak{A})$ abgeschlossen in $(\mathcal{B}(M)/\sim)^k, \delta_Z$.

Beweis:

Mit Satz 3.6.10 und Satz 3.3.13 ist $\mathfrak{Z}_k(\mathfrak{A})$ das Urbild einer abgeschlossenen Menge unter einer stetigen Funktion. Daher ist auch $\mathfrak{Z}_k(\mathfrak{A})$ abgeschlossen nach [Die85], 3.11.4. c), S. 53.

□

3.6.3 Kompaktheit von $(\mathfrak{Z}_k(\mathfrak{A}), \delta_Z)$ und $(\mathfrak{S}_k(\mathfrak{A}), \mathfrak{d})$

Ziel dieses Abschnitts ist es, die Kompaktheit des Systems der geordneten Segmentierungen von M durch k erlaubte Mengen aus \mathfrak{A} zu zeigen. Dabei wird jedes Element des Systems durch eine Abbildung kodiert. Die Menge aller dieser Abbildungen wird dann die Voraussetzungen des Satzes 3.6.4 von Fréchet und Kolmogorov erfüllen. Mit der Abgeschlossenheit des Systems der Segmentierungen folgt dann die Kompaktheit von $(\mathfrak{Z}_k(\mathfrak{A}), \delta_Z)$ und daher auch die Kompaktheit von $(\mathfrak{S}_k(\mathfrak{A}), \mathfrak{d})$ als Bild von $(\mathfrak{Z}_k(\mathfrak{A}), \delta_Z)$ unter einer stetigen Abbildung.

Bemerkung 3.6.12

Auf die Analogie der Definition des Mengensystems \mathfrak{A} zur Konstruktion der L^p -Räume wurde schon in Abschnitt 3.2.1 hingewiesen.

Sei $\mathcal{A} \subset \mathcal{B}(M)$ ein Mengensystem und sei $\mathcal{I} := \{\mathbf{1}_A \mid A \in \mathcal{A}\} \subset \mathcal{L}^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}))$. Sei weiter \sim_L eine Äquivalenzrelation auf \mathcal{I} mit $\mathbf{1}_A \sim_L \mathbf{1}_B \Leftrightarrow \int_M |\mathbf{1}_A - \mathbf{1}_B|^2 d\lambda = 0$ für $\mathbf{1}_A$ und $\mathbf{1}_B$ aus \mathcal{I} . Dann ist

$$\mathbf{1}_A \sim_L \mathbf{1}_B \Leftrightarrow \int_M |\mathbf{1}_A - \mathbf{1}_B|^2 d\lambda = 0 \Leftrightarrow \lambda(A \setminus B) = 0 \wedge \lambda(B \setminus A) = 0$$

Daher entspricht jede Äquivalenzklasse in \mathcal{I}/\sim_L einer Äquivalenzklasse in \mathfrak{A} (vgl. Definition 3.2.1).

Anmerkung: Das Quadrat des Absolutbetragabstands im Inneren des Integrals ist eigentlich nicht nötig, entspricht aber der Äquivalenzbedingung zur Bildung des L^2 -Quotientenraums aus dem Raum der quadratisch integrierbaren Funktionen \mathcal{L}^2 (vgl. z. B. [Bau92], S.99).

Die Abbildungen

$$\begin{array}{l} \mathcal{A} \rightarrow \mathcal{I} \\ A \mapsto \mathbf{1}_A \end{array} \quad \text{bzw.} \quad \begin{array}{l} \mathfrak{A} \rightarrow \mathcal{I}/\sim_L \\ A \mapsto \mathbf{1}_A \end{array}$$

sind also bijektiv.

Definition 3.6.13

Sei \mathfrak{A} ein erlaubtes Mengensystem, dann sei $\mathcal{G}_k(\mathfrak{A}) \subset L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}))$ definiert durch

$$\mathcal{G}_k(\mathfrak{A}) := \left\{ \sum_{i=1}^k i \mathbf{1}_{A_i} \mid (A_1, \dots, A_k) \in \mathfrak{Z}_k(\mathfrak{A}) \right\}$$

Lemma 3.6.14

Die Abbildung

$$j_k : (\mathcal{G}_k(\mathfrak{A}), \|\cdot\|_2) \rightarrow (\mathfrak{Z}_k(\mathfrak{A}), \delta_Z)$$

$$\sum_{i=1}^k i \mathbf{1}_{A_i} \mapsto (A_1, \dots, A_k)$$

ist stetig.

Beweis:

Seien (A_1, \dots, A_k) und (B_1, \dots, B_k) beliebige Elemente aus $\mathfrak{Z}_k(\mathfrak{A})$. Für $f, g \in \mathcal{G}_k(\mathfrak{A})$ gilt dann:

$$\begin{aligned} \int_M (f - g)^2 d\lambda &= \sum_{i=1}^k \int_{A_i} (f - g)^2 d\lambda \\ &= \sum_{i=1}^k \left(\int_{A_i \setminus B_i} (f - g)^2 d\lambda + \int_{A_i \cap B_i} (f - g)^2 d\lambda \right) \\ &\geq \sum_{i=1}^k \int_{A_i \setminus B_i} (f - g)^2 d\lambda \end{aligned} \quad (3.16)$$

Analog ist $\int_M (f - g)^2 d\lambda \geq \sum_{i=1}^k \int_{B_i \setminus A_i} (f - g)^2 d\lambda$. Sei $\varepsilon > 0$ beliebig. Weiter seien $\sum_{i=1}^k i \mathbf{1}_{A_i}$ und $\sum_{i=1}^k i \mathbf{1}_{B_i}$ zwei Elemente aus $\mathcal{G}_k(\mathfrak{A})$ mit $\left\| \sum_{i=1}^k i \mathbf{1}_{A_i} - \sum_{i=1}^k i \mathbf{1}_{B_i} \right\|_2^2 < \frac{\varepsilon}{2}$. Dann gilt:

$$\begin{aligned} \varepsilon &> \left\| \sum_{i=1}^k i \mathbf{1}_{A_i} - \sum_{i=1}^k i \mathbf{1}_{B_i} \right\|_2^2 + \left\| \sum_{i=1}^k i \mathbf{1}_{A_i} - \sum_{i=1}^k i \mathbf{1}_{B_i} \right\|_2^2 \\ \text{[Ungleichung (3.16)]} &\geq \sum_{i=1}^k \int_{A_i \setminus B_i} \left(\sum_{i=1}^k i \mathbf{1}_{A_i} - \sum_{i=1}^k i \mathbf{1}_{B_i} \right)^2 d\lambda \\ &\quad + \sum_{i=1}^k \int_{B_i \setminus A_i} \left(\sum_{i=1}^k i \mathbf{1}_{A_i} - \sum_{i=1}^k i \mathbf{1}_{B_i} \right)^2 d\lambda \\ &= \sum_{i=1}^k i^2 \lambda(A_i \setminus B_i) + \sum_{i=1}^k i^2 \lambda(B_i \setminus A_i) = \sum_{i=1}^k i^2 \mathfrak{d}(A_i, B_i) \\ &\geq \sum_{i=1}^k \mathfrak{d}(A_i, B_i) \end{aligned}$$

Damit folgt die Behauptung. □

Bemerkung 3.6.15

Die Abbildung j_k aus Lemma 3.6.14 definiert eine eindeutige Kodierung von Funktionen aus $\mathcal{G}_k(\mathfrak{A})$ durch Elemente aus $\mathfrak{Z}_k(\mathfrak{A})$, die zudem stetig ist. Es ist dieses Kodierungslemma, das die Definition 3.6.5 der geordneten Segmentierungen nötig macht: Ohne eine Ordnung der Segmente wäre eine solche stetige Kodierung nicht möglich.

Die Menge $\mathcal{G}_k(\mathfrak{A})$ wird benötigt, um mit Hilfe des Satzes von Fréchet und Kolmogorov (Satz 3.6.4) die Kompaktheit von $\mathfrak{Z}_k(\mathfrak{A})$ zu zeigen.

Lemma 3.6.16

Sei

$$S_k := \left\{ f \in L^2(\mathcal{B}(M), \mathcal{B}(E)) \mid \exists A_1, \dots, A_k \in \mathcal{B}(M), f = \bigsqcup_{i=1}^k A_i \right\}$$

eine Menge von Treppenfunktion mit höchstens k Stufen. Dann ist S_k abgeschlossen in $L^2(\mathcal{B}(M), \mathcal{B}(E))$.

Beweis:

Sei $(f_k)_{k \in \mathbb{N}}$ eine Cauchy-Folge in S_k mit $\lim_{k \rightarrow \infty} f_k = f$. Da der Raum $L^2(\mathcal{B}(M), \mathcal{B}(E))$ vollständig ist, gilt $f \in L^2(\mathcal{B}(M), \mathcal{B}(E))$. Zu zeigen ist $f \in S_k$. Nach Voraussetzung existiert für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ mit

$$\forall n > n_0 : \|f - f_n\|_2^2 = \int_M \|f - f_n\|_2^2 d\lambda < \varepsilon \quad (3.17)$$

Daher existiert eine Folge $(\hat{f}_n)_{n \in \mathbb{N}}$ in $\mathcal{L}^2(\mathcal{B}(M), \mathbb{R})$ derart, daß für alle $n \in \mathbb{N}$ gilt: $\hat{f}_n \in f_n$. Nach Voraussetzung existiert eine Funktion $\hat{f} \in \mathcal{L}^2(\mathcal{B}(M), \mathbb{R})$ und $\hat{f} \in f$ mit $\lim_{n \rightarrow \infty} \|\hat{f}_n - \hat{f}\|^2 = 0$.

Wähle ein δ mit $\frac{1}{2} > \delta > 0$ beliebig.

Annahme: Es existiert eine Menge $A \in \mathcal{B}(M)$ mit $\lambda(A) > 0$ derart, daß für alle $x \in A$ gilt: $\hat{f}(x) \notin [i - \delta, i + \delta]$ für alle $i \in \{1, \dots, k\}$.

Dann ist aber für alle $x \in A$ $|\hat{f}(x) - \hat{f}_n| > \delta$. Nun gilt aber für alle $n \geq n_0$

$$\begin{aligned} \varepsilon \geq \|f_n - f\|^2 &= \int_M |\hat{f}_n(x) - \hat{f}(x)|^2 d\lambda(x) \\ &\geq \int_A |\hat{f}_n(x) - \hat{f}(x)|^2 d\lambda(x) \\ &\geq \delta \lambda(A) \end{aligned}$$

Da ε jedoch beliebig war, folgt ein Widerspruch zur Annahme, daß heißt, es existiert keine Menge $A \in \mathcal{B}(M)$ mit $\lambda(A) > 0$ derart, daß für alle $x \in A$ gilt: $\hat{f}(x) \notin [i - \delta, i + \delta]$ für alle $i \in \{1, \dots, k\}$. Vielmehr muß in diesem Fall $\lambda(A) = 0$ gelten.

Da $\delta \in]0, \frac{1}{2}[$ beliebig gewählt war, muß λ -fast sicher $\hat{f}(x) \in \{1, \dots, k\}$ für $x \in M$ gelten. Also ist \hat{f} bis auf ein λ -Nullmenge eine Funktion mit höchstens k Stufen. Daher gilt $f \in S_k$.

□

Satz 3.6.17

Sei \mathfrak{A} ein erlaubtes Mengensystem. Dann gilt:

- i. $\mathcal{G}_k(\mathfrak{A})$ ist abgeschlossen in $L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}))$.
- ii. $\mathcal{G}_k(\mathfrak{A})$ ist kompakt in $L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}))$.

Beweis:

i. Offensichtlich gilt $\mathcal{G}_k(\mathcal{B}(M)/\sim) = S_k$. Daher ist $\mathcal{G}_k(\mathcal{B}(M)/\sim)$ nach Lemma 3.6.16 abgeschlossen in $L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}))$. Für \mathfrak{A} ist $\mathcal{G}_k(\mathfrak{A})$ nach Korollar 3.6.11 und Lemma 3.6.14 das Urbild einer abgeschlossenen Menge unter einer stetigen Bijektion. Daher ist in diesem Fall $\mathcal{G}_k(\mathfrak{A})$ abgeschlossen in $(S_k, \|\cdot\|_2)$ nach [Die85], 3.11.4 c), S. 53. Mit [Die85] 3.10.3, S. 51, folgt daher die Behauptung.

ii. Da $\mathcal{G}_k(\mathfrak{A})$ abgeschlossen ist, genügt es zum Beweis von ii. die relative Kompaktheit von $\mathcal{G}_k(\mathfrak{A})$ nachzuweisen. Es reicht also zu zeigen, daß $\mathcal{G}_k(\mathfrak{A})$ den Voraussetzungen des Satzes von Fréchet und Kolmogorov (Satz 3.6.4) genügt. Die Voraussetzungen (3.4) und (3.6) des Satzes sind trivial erfüllt. Der Nachweis der Gültigkeit der Bedingung (3.5), nämlich

$$\lim_{\|t\|_\infty \rightarrow 0} \int_M |g(x+t) - g(x)|^2 d\lambda(x) = 0 \quad \text{gleichmäßig in allen } g \in \mathcal{G}_k(\mathfrak{A})$$

fällt auch nicht besonders schwer:

Sei $g := \sum_{i=1}^k i \mathbf{1}_{A_i} \in \mathcal{G}_k(\mathfrak{A})$ beliebig. Für $t \in \mathbb{R}^k$ und $B \in \mathcal{B}(M)$ bezeichne $B+t$ die Menge $\{b+t \mid b \in B, b+t \in M\}$. Dann gilt:

$$\begin{aligned} & \int_M (g(x) - g(x+t))^2 d\lambda(x) \\ &= \int_M \left(\sum_{i=1}^k i \mathbf{1}_{A_i} - \sum_{i=1}^k i \mathbf{1}_{A_i-t} \right)^2 d\lambda \\ &= \int_M \left(\sum_{i=1}^k i \mathbf{1}_{A_i \setminus (A_i-t)} - \sum_{i=1}^k i \mathbf{1}_{(A_i-t) \setminus A_i} \right)^2 d\lambda \\ &\leq \int_M \left(\sum_{i=1}^k i \mathbf{1}_{A_i \setminus (A_i-t)} \right)^2 + \left(\sum_{i=1}^k i \mathbf{1}_{(A_i-t) \setminus A_i} \right)^2 d\lambda \\ &= \int_M \left(\sum_{i=1}^k i \mathbf{1}_{A_i \setminus (A_i-t)} \right)^2 d\lambda + \int_M \left(\sum_{i=1}^k i \mathbf{1}_{(A_i-t) \setminus A_i} \right)^2 d\lambda \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^k \int_{A_i \setminus (A_i - t)} i^2 d\lambda + \sum_{i=1}^k \int_{(A_i - t) \setminus A_i} i^2 d\lambda \\
&= \sum_{i=1}^k \lambda(A_i \setminus (A_i - t)) i^2 + \sum_{i=1}^k \lambda((A_i - t) \setminus A_i) i^2 \\
&\leq k^2 \sum_{i=1}^k (\lambda(A_i \setminus (A_i - t)) + \lambda((A_i - t) \setminus A_i))
\end{aligned}$$

Nun gilt aber für $1 \leq i \leq k$, daß $\lim_{\|t\|_\infty \rightarrow 0} \lambda((A_i - t) \setminus A_i) = 0$ ist. Damit folgt

$$\begin{aligned}
&\lim_{\|t\|_\infty \rightarrow 0} \int_M |g(x+t) - g(x)|^2 d\lambda(x) \\
&\leq \lim_{\|t\|_\infty \rightarrow 0} k^2 \sum_{i=1}^k (\lambda(A_i \setminus (A_i - t)) + \lambda((A_i - t) \setminus A_i)) \\
&= 0
\end{aligned}$$

Also folgt die Voraussetzung (3.5) des Satzes 3.6.4.

Somit sind also sämtliche Voraussetzungen des Satzes von Fréchet und Kolmogorov erfüllt und es folgt die Kompaktheit der Menge $\mathcal{G}_k(\mathfrak{A})$. □

Satz 3.6.18

Die Menge $\mathfrak{Z}_k(\mathfrak{A})$ ist kompakt bezüglich δ_Z .

Beweis:

Mit Bemerkung 3.6.12, Satz 3.6.17 und dem Lemma 3.6.14 ist $\mathfrak{Z}_k(\mathfrak{A})$ das Bild einer kompakten Menge unter einer stetigen Abbildung. Daher folgt die Behauptung mit [Die85], 3.17.9., S. 68. □

Satz 3.6.19

$\mathfrak{S}_k(\mathfrak{A})$ ist bezüglich der Metrik \mathfrak{d} kompakt.

Beweis:

Mit den Sätzen 3.6.18 und 3.6.10 ist $\mathfrak{S}_k(\mathfrak{A})$ das stetige Bild einer kompakten Menge und daher selbst kompakt nach [Die85], 3.17.9., S. 68. □

3.6.4 Stetigkeit der Funktion Ψ

Ziel dieses Abschnitts ist es, die Stetigkeit der Zielfunktion der k -Niveaupassung bezüglich der Metrik \mathfrak{d} des Systems der Segmentierungen $\mathfrak{S}_k(\mathfrak{A})$ zu zeigen.

Für $A \in \mathfrak{A} \setminus \{O\}$ sei (e_1^A, \dots, e_n^A) die ausgezeichnete Basis bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_A$ und der Basis (e_1, \dots, e_n) von V gemäß Definition 3.4.9.

Lemma 3.6.20

Für alle Funktionen $f \in L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$ ist die Funktion $\mathfrak{A} \setminus \{O\} \rightarrow \mathbb{R}_+, A \mapsto \|f\|_A$ gleichmäßig stetig bezüglich \mathfrak{d} und der Betragsmetrik auf \mathbb{R} .

Beweis:

Mit Definition 3.4.4 iii. existiert für jedes $\varepsilon > 0$ ein $\delta > 0$ mit $\lambda(A) < \delta \Rightarrow |\langle f | f \rangle_A| < \varepsilon$ für $A \in \sigma(\mathfrak{A})$. Daher folgt für alle Mengen $A, A' \in \sigma(\mathfrak{A})$ mit $\lambda(A \setminus A') + \lambda(A' \setminus A) = \delta(A, A') < \delta$, daß $|\langle f | f \rangle_A - \langle f | f \rangle_{A'}| = |\langle f | f \rangle_{A \setminus A'} - \langle f | f \rangle_{A' \setminus A}| \leq |\langle f | f \rangle_{A \setminus A'}| + |\langle f | f \rangle_{A' \setminus A}| = 2\varepsilon$ gelten muß. Daher folgt die Behauptung. \square

Lemma 3.6.21

Für $1 \leq i \leq n$ und $f \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ ist die Abbildung

$$\mathfrak{A} \setminus \{O\} \rightarrow \mathbb{R}, A \mapsto \langle f | e_i^A \rangle_A$$

gleichmäßig stetig bezüglich \mathfrak{d} und Betragsnorm auf \mathbb{R}

Beweis:

Der Beweis erfolgt über die Konstruktion der ausgezeichneten Basis (e_1^A, \dots, e_n^A) gemäß Algorithmus 5 (**Gram-Schmidt**). Sei dazu $f \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ und $A \in \mathfrak{A} \setminus \{O\}$ beliebig.

Induktion über i :

$i = 1$: Die Abbildung $A \mapsto \langle e_1 | f \rangle_A$ ist gleichmäßig stetig. Denn:

Sei $\varepsilon > 0$ beliebig. Nach Definition 3.4.4 iii. (S. 85) der Familie der erlaubten Skalarprodukte gibt es ein $\delta > 0$, so daß für alle $A' \in \mathfrak{A}$ mit $\mathfrak{d}(A, A') < \delta$ gilt:

$$|\langle e_1 | f \rangle_{A \setminus A'}| < \frac{\varepsilon}{2} \text{ und } |\langle e_1 | f \rangle_{A' \setminus A}| < \frac{\varepsilon}{2}. \text{ Damit ist aber}$$

$$\begin{aligned} |\langle e_1 | f \rangle_A - \langle e_1 | f \rangle_{A'}| &= |\langle e_1 | f \rangle_{A \setminus A'} - \langle e_1 | f \rangle_{A' \setminus A}| \\ &\leq |\langle e_1 | f \rangle_{A \setminus A'}| + |\langle e_1 | f \rangle_{A' \setminus A}| \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

Da nach Lemma 3.6.20 die Abbildung $A \mapsto \|e_1\|_A$ gleichmäßig stetig ist, folgt mit $\langle e_1^A | f \rangle_A = \frac{1}{\|e_1\|_A} \langle e_1 | f \rangle_A$ die Behauptung für $i = 1$.

$i > 1$: Nach Induktionsvoraussetzung sind für $j < i$ die Abbildungen $A \mapsto \langle e_j^A | f \rangle_A$ gleichmäßig stetig. Dies gilt insbesondere auch für $f = e_i$. Damit ist die Abbildung

$$A \mapsto \sum_{j=1}^{i-1} \langle e_j^A | f \rangle_A \langle e_j^A | e_i \rangle_A$$

gleichmäßig stetig. Somit existiert mit Definition 3.4.4 iii. ein $\delta > 0$, so daß für alle $A' \in \mathfrak{A}$ mit $\mathfrak{d}(A, A') < \delta$ gilt:

$$\begin{aligned} & \left| \langle e_i | f \rangle_{A \setminus A'} \right| < \frac{\varepsilon}{3}, \quad \left| \langle e_i | f \rangle_{A' \setminus A} \right| < \frac{\varepsilon}{3} \\ \text{und} \quad & \left| \sum_{j=1}^{i-1} \langle e_j^A | f \rangle_A \langle e_j^A | e_i \rangle_A - \sum_{j=1}^{i-1} \langle e_j^{A'} | f \rangle_{A'} \langle e_j^{A'} | e_i \rangle_{A'} \right| < \frac{\varepsilon}{3} \end{aligned}$$

Damit gilt mit der Konstruktion von \tilde{e}_i^A

$$\begin{aligned} & \left| \langle \tilde{e}_i^A | f \rangle_A - \langle \tilde{e}_i^{A'} | f \rangle_{A'} \right| \\ &= \left| \left\langle e_i - \sum_{j=1}^{i-1} \langle e_i | e_j^A \rangle_A e_j^A \middle| f \right\rangle_A - \left\langle e_i - \sum_{j=1}^{i-1} \langle e_i | e_j^{A'} \rangle_{A'} e_j^{A'} \middle| f \right\rangle_{A'} \right| \\ &= \left| \langle e_i | f \rangle_A - \sum_{j=1}^{i-1} \langle e_i | e_j^A \rangle_A \langle e_j^A | f \rangle_A \right. \\ & \quad \left. - \langle e_i | f \rangle_{A'} + \sum_{j=1}^{i-1} \langle e_i | e_j^{A'} \rangle_{A'} \langle e_j^{A'} | f \rangle_{A'} \right| \\ &= \left| \langle e_i | f \rangle_{A \setminus A'} - \langle e_i | f \rangle_{A' \setminus A} \right. \\ & \quad \left. - \sum_{j=1}^{i-1} \langle e_i | e_j^A \rangle_A \langle e_j^A | f \rangle_A + \sum_{j=1}^{i-1} \langle e_i | e_j^{A'} \rangle_{A'} \langle e_j^{A'} | f \rangle_{A'} \right| \\ &\leq \left| \langle e_i | f \rangle_{A \setminus A'} \right| + \left| \langle e_i | f \rangle_{A' \setminus A} \right| \\ & \quad + \left| \sum_{j=1}^{i-1} \langle e_i | e_j^{A'} \rangle_{A'} \langle e_j^{A'} | f \rangle_{A'} - \sum_{j=1}^{i-1} \langle e_i | e_j^A \rangle_A \langle e_j^A | f \rangle_A \right| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

Daher folgt die gleichmäßige Stetigkeit der Abbildung $A \mapsto \langle \tilde{e}_i^A | f \rangle_A$. Mit Lemma 3.6.20 folgt dann die gleichmäßige Stetigkeit von $A \mapsto \langle e_i^A | f \rangle_A$ für $A \in \mathfrak{A} \setminus \{O\}$.

□

Korollar 3.6.22

Für alle $f \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ ist die Abbildung

$$\mathfrak{A} \rightarrow \mathbb{R}, A \mapsto \|\gamma_f(A)\|_2$$

gleichmäßig stetig bezüglich \mathfrak{d} und Betragsnorm auf \mathbb{R}

Beweis:

Mit Lemma 3.6.21 gilt für beliebige $f \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ und $i \in \{1, \dots, n\}$, daß die Funktion

$$A \mapsto \langle f | e_i^A \rangle_A$$

gleichmäßig stetig ist auf $\mathfrak{A} \setminus \{O\}$. Daher läßt sie sich nach [Die85] 3.15.6, S. 62, eindeutig gleichmäßig stetig auf \mathfrak{A} fortsetzen. □

Bemerkung 3.6.23

Der Beweis des vorangegangenen Lemmas zeigt, daß es wichtig ist, auf welche Art und Weise die einzelnen ausgezeichneten Orthonormalbasen (e_1^A, \dots, e_n^A) entstehen. Die gleichmäßige Stetigkeit der Abbildung $A \mapsto \langle f | e_i^A \rangle_A$ muß garantiert sein, wenn man, etwa aus numerischen Gründen, ein anderes Verfahren als den in Bemerkung 3.4.8 und Definition 3.4.9 festgelegten Algorithmus nach Gram-Schmidt zur Bestimmung der einzelnen Basen verwenden will.

Satz 3.6.24

Sei $\mathcal{N} := \{(A_1, \dots, A_k) \in \mathfrak{Z}_k(\mathfrak{A}) \mid \exists i \in \{1, \dots, k\} : A_i = O\}$. Dann ist die Funktion

$$\tilde{\Psi} : \begin{array}{l} (\mathfrak{Z}_k(\mathfrak{A}) \setminus \mathcal{N}, \mathfrak{d}) \rightarrow (\mathbb{R}_+, |\cdot|) \\ (A_1, \dots, A_k) \mapsto \sum_{i=1}^k \|\gamma_f(A_i)\|_2^2 \end{array}$$

gleichmäßig stetig.

Beweis:

Da $\|\gamma_f(A)\|_2^2 = \sum_{i=1}^n \langle f | e_i^A \rangle_A^2$ ist, folgt aus Korollar 3.6.22 die Behauptung. □

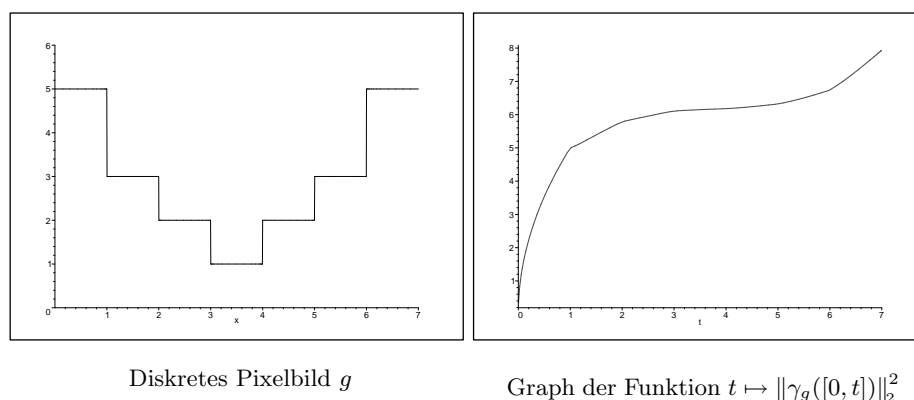


Abbildung 3.11: Diskretes eindimensionales Pixelbild g und Verlauf der Funktion $t \mapsto \|\gamma_g([0, t])\|_2^2$ aus Beispiel 3.6.25.

Beispiel 3.6.25

Es sei $M := [0, 7]$ und $E := \mathbb{R}$ mit der Multiplikation als Skalarprodukt. \mathfrak{A} sei das Mengensystem der zusammenhängenden Mengen in M . Die Familie der Skalarprodukte sei gegeben wie in Beispiel 3.4.6. Es ist $\mathfrak{S}_2(\mathfrak{A}) = \{\{[0, t], [t, 7]\} \mid t \in [0, 7]\}$ das System der Segmentierungen von M durch zwei zusammenhängende Mengen. Die Klasse der erlaubten Funktionen habe die Basis $(1, \text{id})$.

Gegeben sei gemäß Definition 1.3.2 ein diskretes Pixelbild $g : M \rightarrow \mathbb{R}$ mit $g([0]) := 5, g([1]) := 3, g([2]) := 2, g([3]) := 1, g([4]) := 2, g([5]) := 3, g([6]) := 5$. Dann kann gezeigt werden, daß die Funktion

$$\begin{aligned}]0, 7] &\rightarrow \mathbb{R}_+ \\ t &\mapsto \|\gamma_g([0, t])\|_2^2 \end{aligned}$$

einen Verlauf hat, wie ihn Abbildung 3.11 skizziert. Insbesondere ist $\lim_{t \rightarrow 0} \|\gamma_g([0, t])\|_2^2 = 0$.

Lemma 3.6.26

Es gilt

$$\lim_{\substack{A \rightarrow O \\ A \in \mathfrak{A}}} \|\gamma_f(A)\|_2 = 0$$

Beweis:

Nach Lemma 3.5.5 gilt für alle $A \in \sigma(\mathfrak{A})$ mit $\lambda(A) > 0$:

$$0 \leq \|f - f_A\|_A^2 = \|f\|_A^2 - \|\gamma_f(A)\|_2^2$$

Aus der Eigenschaft iii. der erlaubten Skalarprodukte (vgl. Definition 3.4.4, S. 85) folgt

$$\lim_{\substack{A \rightarrow O \\ A \in \mathfrak{A}}} \|f\|_A^2 = 0$$

Damit folgt die Behauptung. □

Satz 3.6.27

Die Funktion

$$\Psi : \mathbf{A} \mapsto \sum_{A \in \mathbf{A}} \|\gamma_f(A)\|_2^2$$

ist stetig.

Beweis:

Sei $\mathbf{A}_{n \in \mathbb{N}}^{(n)}$ eine Folge in $\mathfrak{S}_k(\mathfrak{A})$, die gegen ein $\{A_1, \dots, A_l\}$ mit $l \leq k$ konvergiert. Mit Satz 3.3.13 ist $\{A_1, \dots, A_l\} \in \mathfrak{S}_k(\mathfrak{A})$. Weiter existiert nach Lemma 3.3.12 ein n_0 derart, daß für jedes $i \in \{1, \dots, l\}$ eine eindeutig bestimmte Folge $(A_i^{(n)})_{n \in \mathbb{N}, n > n_0}$ existiert, so daß für alle $n > n_0$ das Folgenglied $A_i^{(n)} \in \mathbf{A}^{(n)}$ ist und $\lim_{n \rightarrow \infty} A_i^{(n)} = A_i$ gilt. Außerdem sind diese Folgen paarweise verschieden.

Daher folgt aus Korollar 3.6.22, Lemma 3.6.26 und der Anmerkung zum Beweis des Satzes 3.3.13 (S. 82):

$$\begin{aligned} \lim_{n \rightarrow \infty} \Psi(\mathbf{A}^{(n)}) &= \lim_{n \rightarrow \infty} \sum_{A \in \mathbf{A}^{(n)}} \|\gamma_f(A)\|_2^2 \\ &= \lim_{n \rightarrow \infty} \left(\sum_{i=1}^l \|\gamma_f(A_i^{(n)})\|_2^2 + \sum_{A \in \mathbf{A} \setminus \{A_1^{(n)}, \dots, A_l^{(n)}\}} \|\gamma_f(A)\|_2^2 \right) \\ &= \sum_{i=1}^l \|\gamma_f(A_i)\|_2^2 + 0 \\ &= \Psi(\lim_{n \rightarrow \infty} \mathbf{A}^{(n)}) \end{aligned}$$

Daher folgt die Behauptung. □

3.6.5 Existenz des Maximums

Aufgrund der Kompaktheit von $\mathfrak{Z}_k(\mathfrak{A})$ und der Stetigkeit der Funktion Ψ ist die Existenz eines globalen Optimums der k -Niveaupassung gesichert.

Satz 3.6.28

Die Funktion

$$\tilde{\Psi} : \begin{array}{l} \mathfrak{Z}_k(\mathfrak{A}) \rightarrow \mathbb{R}_+ \\ (A_1, \dots, A_k) \mapsto \sum_{i=1}^k \|\gamma_f(A_i)\|_2^2 \end{array}$$

nimmt auf $\mathfrak{Z}_k(\mathfrak{A})$ ihr Maximum an.

Beweis:

Nach Satz 3.6.24 ist die Funktion $\tilde{\Psi}$ gleichmäßig stetig auf $\mathfrak{Z}_k(\mathfrak{A}) \setminus \mathcal{N}$. Daher existiert genau eine stetige Fortsetzung von $\tilde{\Psi}$ auf den Abschluß von $\mathfrak{Z}_k(\mathfrak{A}) \setminus \mathcal{N}$. Diese ist selbst gleichmäßig stetig (vgl. z. B. [Die85], 3.15.6., S. 62). Dieser Abschluß ist jedoch die Menge $\mathfrak{Z}_k(\mathfrak{A})$. Da $\mathfrak{Z}_k(\mathfrak{A})$ kompakt ist (Satz 3.6.18), folgt die Behauptung, denn stetige Funktionen auf kompakten Mengen nehmen ihr Maximum an. (vgl. [Die85], 3.17.10., S. 68)

□

Satz 3.6.29

Die Funktion

$$\Psi : \begin{array}{l} \mathfrak{S}_k(\mathfrak{A}) \rightarrow \mathbb{R}_+ \\ \mathbf{A} \mapsto \sum_{A \in \mathbf{A}} \|\gamma_f(A)\|_2^2 \end{array}$$

nimmt auf $\mathfrak{S}_k(\mathfrak{A})$ ihr Maximum an.

Beweis:

Da die Funktion Ψ stetig ist auf $\mathfrak{S}_k(\mathfrak{A})$ (Satz 3.6.27) und $\mathfrak{S}_k(\mathfrak{A})$ nach Satz 3.6.19 kompakt ist, folgt die Behauptung mit [Die85], 3.17.10., S. 68.

□

Bemerkung 3.6.30

Um die Existenz einer optimalen Lösung einer 2-Niveaupassung im Sinne von Abschnitt 3.5.1 zu zeigen, ist die Stetigkeit der Funktion

$$\Psi : \begin{array}{l} \mathfrak{A} \rightarrow \mathbb{R}_+ \\ A \mapsto \|\gamma_f(A)\|_2^2 + \|\gamma_f(M \setminus A)\|_2^2 \end{array}$$

und die Kompaktheit von \mathfrak{A} zu zeigen, da diese Funktion nicht über das System der Segmentierungen $\mathfrak{S}_2(\mathfrak{A})$ parametrisiert wurde. Beide Eigenschaften sind jedoch mit den Mitteln dieses Abschnitts leicht zu zeigen.

3.7 Bestimmung lokaler Maxima der Funktion Ψ

Obwohl das vorangegangene Unterkapitel die Existenz eines globalen Maximums der Zielfunktion Ψ sichert, werden sich die meisten Algorithmen aus Gründen der Laufzeiteffizienz auf die Suche nach einem lokalen Maximum beschränken. Denn in der Regel handelt es sich bei dem Optimierungsproblem der k -Niveaupassung um ein nichtlineares Optimierungsproblem. In Kapitel 4 werden Algorithmen für diskrete Pixelbilder angegeben, die lokale Maxima bestimmen. Die theoretischen Grundlagen der dort zu findenden Algorithmen sind allerdings auch für beliebige Bilder $f \in L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$ gültig. Aus diesem Gründen werden diese Grundlagen schon in diesem Kapitel vorgestellt.

Im ersten Abschnitt wird gezeigt, daß es unter bestimmten Voraussetzungen möglich ist, das Problem der k -Niveaupassung von Bildern auf iterative 2-Niveaupassungen zu beschränken. Diese Voraussetzungen sind so allgemein, daß sich der zweite Abschnitt nur mit der Suche von lokalen Maxima von 2-Niveaupassungen beschäftigt.

Bemerkung 3.7.1 (Strenges lokales Maximum)

Sei (E, d) ein metrischer Raum. Ein Punkt $x_0 \in E$ heißt dann strenges lokales Maximum einer Funktion $f : E \rightarrow \mathbb{R}$, falls eine Umgebung U von x_0 existiert mit

$$\forall x \in U : f(x) < f(x_0)$$

Daher folgt aus dem Satz 3.6.29, also der Annahme des Maximums durch die Zielfunktion Ψ der k -Niveaupassung, nicht die Existenz eines strengen lokalen Maximums.

3.7.1 Reduktion auf 2-Niveaupassungsprobleme

Dieser Abschnitt definiert zunächst die Menge der Segmentierungen durch zwei Segmente innerhalb einer schon existierenden Segmentierung. Kann man alle Segmentierungen innerhalb einer Umgebung einer lokal optimalen Segmentierung durch iterative Veränderungen von jeweils zwei Segmenten beschreiben, so läßt sich das k -Niveaupassungsproblem auf iterative 2-Niveaupassungsprobleme reduzieren.

Bemerkung 3.7.2

Sei \mathfrak{A} ein Mengensystem erlaubter Mengen. Sei weiter $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ und seien $A_i, A_j \in \mathbf{A}$ mit $A_i \neq A_j$ und $A_i \cup A_j \in \mathfrak{A}$. Dann gilt:

$$\forall \{A'_i, A'_j\} \in \mathfrak{S}_2(\mathfrak{A}, A_i \cup A_j) : (\mathbf{A} \setminus \{A_i, A_j\}) \cup \{A'_i, A'_j\} \in \mathfrak{S}_k(\mathfrak{A})$$

Das heißt, solange die Mengen A'_i und A'_j innerhalb der eingeschränkten Zerlegung $\mathfrak{S}_2(\mathfrak{A}, A_i \cup A_j)$ möglich sind, sind sie es auch im ursprünglichen System der Segmentierungen.

Definition 3.7.3 (Segmentierungen innerhalb einer Segmentierung)

Sei \mathfrak{A} ein System erlaubter Mengen. Dann sei für alle $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ und $A_i, A_j \in \mathbf{A}$ mit $A_i \cup A_j \in \mathfrak{A}$ die Menge $\mathfrak{S}_{A_i, A_j}(\mathbf{A})$ definiert durch

$$\mathfrak{S}_{A_i, A_j}(\mathbf{A}) := \{(\mathbf{A} \setminus \{A_i, A_j\}) \cup \{A'_i, A'_j\} \mid A'_i, A'_j \in \mathfrak{S}_2(\mathfrak{A}, A_i \cup A_j)\}$$

$\mathfrak{S}_{A_i, A_j}(\mathbf{A})$ heie die Menge der Segmentierungen der Menge $A_i \cup A_j$ durch zwei Segmente innerhalb der Segmentierung \mathbf{A} .

Mit Bemerkung 3.7.2 gilt $\mathfrak{S}_{A_i, A_j}(\mathbf{A}) \subseteq \mathfrak{S}_k(\mathfrak{A})$.

Beispiel 3.7.4

Sei $M := [0, 10]$ und $\mathfrak{A} := \{[t_1, t_2] \mid t_1, t_2 \in M\}$. Dann ist $\{[0, 1], [1, 3], [3, 10]\} \in \mathfrak{S}_3(\mathfrak{A})$. Und mit Definition 3.7.3 ist die Menge $\mathfrak{S}_{[1, 3], [3, 10]}(\mathfrak{A})$ gegeben durch

$$\mathfrak{S}_{[1, 3], [3, 10]}(\mathfrak{A}) = \{\{[0, 1], [1, t], [t, 10]\} \mid t \in [1, 10]\}$$

Bemerkung 3.7.5

Sei $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ und $A_i, A_j \in \mathbf{A}$. Dann sind alle $\mathbf{A}' \in \mathfrak{S}_{A_i, A_j}(\mathbf{A})$ der Form $\mathbf{A} \setminus \{A_i, A_j\} \cup \{A'_i, A'_j\}$ für ein $\{A'_i, A'_j\} \in \mathfrak{S}_2(\mathfrak{A}, A_i \cup A_j)$. Für die Zielfunktion der k -Niveaupassung Ψ gilt

$$\begin{aligned} \Psi(\mathbf{A}') &= \|\gamma_f(A'_i)\|_2^2 + \|\gamma_f(A'_j)\|_2^2 \\ &+ \sum_{A \in \mathbf{A} \setminus \{A_i, A_j\}} \|\gamma_f(A)\|_2^2 \end{aligned} \quad (3.18)$$

Dabei ist für alle $\mathbf{A}' \in \mathfrak{S}_{A_i, A_j}(\mathbf{A})$ der Wert $\sum_{A \in \mathbf{A} \setminus \{A_i, A_j\}} \|\gamma_f(A)\|_2^2$ in Term (3.18) gleich. Um ein lokales Maximum in $\mathfrak{S}_{A_i, A_j}(\mathbf{A})$ zu berechnen, ist also lediglich ein lokales Maximum in $\mathfrak{S}_2(\mathfrak{A}, A_i \cup A_j)$ zu bestimmen.

Satz 3.7.6

Sei (E, d) ein metrischer Raum. Dann ist für jedes $a \in E$ und jedes $\varepsilon > 0$ die Menge $K_\varepsilon(a) := \{x \in E \mid d(x, a) < \varepsilon\}$ eine offene Umgebung von a .

Beweis:

vgl. [Die85], 3.6.1., S. 44

Bemerkung 3.7.7

Sei $M := [0, 1]$ und $\mathfrak{A} := \{[x_1, x_2] \mid x_1, x_2 \in [0, 1], x_1 \leq x_2\}$ das Mengensystem der erlaubten Mengen. Betrachtet wird nun $\mathfrak{S}_3(\mathfrak{A})$, das System der Segmentierungen von $[0, 1]$ durch 3 Intervalle. $\{[0, x_1], [x_1, x_2], [x_2, 1]\}$ sei ein Element aus $\mathfrak{S}_3(\mathfrak{A})$ mit $x_1 > 0$. Sei weiter ein $\varepsilon \in \mathbb{R}_+$ mit $x_1 > \varepsilon > 0$ gegeben. Wählt man dann $x'_1 := x_1 - \frac{\varepsilon}{2}$, so gilt $\mathfrak{d}(\{[0, x_1], [x_1, x_2], [x_2, 1]\}, \{[0, x'_1], [x'_1, x_2], [x_2, 1]\}) \leq \varepsilon$.

Anhand dieses Beispiels kann man sich leicht überlegen, daß es für jedes Element aus $\mathbf{A} \in \mathfrak{S}_3(\mathfrak{A})$ und jedem $\varepsilon > 0$ ein Element $\mathbf{A}' \in \mathfrak{S}_3(\mathfrak{A})$ mit $\mathfrak{d}(\mathbf{A}, \mathbf{A}') < \varepsilon$ gibt, das aus \mathbf{A} nur durch die Veränderung zweier adjazenter Segmente entstanden ist. Für jedes Element $\mathbf{A} = \{A_1, A_2, A_3\} \in \mathfrak{S}_3(\mathfrak{A})$ mit $O \notin \mathbf{A}$ und jeder Umgebung U von \mathbf{A} sind also die Mengen $\{\mathbf{A}' \in \mathfrak{S}_{A_1, A_2}(\mathbf{A})\} \cap U$, $\{\mathbf{A}' \in \mathfrak{S}_{A_2, A_3}(\mathbf{A})\} \cap U$ und $\{\mathbf{A}' \in \mathfrak{S}_{A_1, A_3}(\mathbf{A})\} \cap U$ nicht leer.

Definition 3.7.8

Für $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ sei

$$\mathcal{U}_2(\mathbf{A}) := \{\mathbf{A}' \mid \exists A_i, A_j \in \mathbf{A} : A_i \cup A_j \in \mathfrak{A}, \mathbf{A}' \in \mathfrak{S}_{A_i, A_j}(\mathbf{A})\}$$

Weiter sei für $\varepsilon > 0$ beliebig die Menge $\mathcal{U}_2^\varepsilon(\mathbf{A})$ definiert durch

$$\mathcal{U}_2^\varepsilon(\mathbf{A}) := \{\mathbf{B} \in \mathcal{U}_2(\mathbf{A}) \mid \mathfrak{d}(\mathbf{B}, \mathbf{A}) < \varepsilon\}$$

Bemerkung 3.7.9

Im Kontext der Bemerkung 3.7.7 gilt für alle $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$, wie man sich leicht überlegen kann, $\mathcal{U}_2^\varepsilon(\mathbf{A}) \subseteq K_\varepsilon(\mathbf{A})$. Diese Inklusion kann jedoch garantiert werden, denn unter Umständen erlaubt das gewählte Mengensystem keine Veränderungen in nur zwei Segmenten (vgl. auch Bemerkung 4.2.16).

Satz 3.7.10

Sei \mathfrak{A} derart, daß für jedes $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ und jedes $\varepsilon > 0$ die Inklusion $\mathcal{U}_2^\varepsilon(\mathbf{A}) \subseteq K_\varepsilon(\mathbf{A})$ gilt. Dann ist die folgende Aussage richtig:

Ist $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ ein lokales Maximum einer k -Niveaupassung Ψ , so gilt:

Für alle $A_i, A_j \in \mathbf{A}$ mit $A_i \cup A_j \in \mathfrak{A}$ ist $\{A_i, A_j\}$ ein lokales Maximum der 2-Niveaupassung

$$\Psi_2 : \begin{array}{l} \mathfrak{S}_2(\mathfrak{A}, A_i \cup A_j) \rightarrow \mathbb{R}_+ \\ \{B, (A_i \cup A_j) \setminus B\} \mapsto \|\gamma_f(B)\|_2^2 + \|\gamma_f((A_i \cup A_j) \setminus B)\|_2^2 \end{array}$$

Beweis:

Es ist

$$\mathcal{U}_2(\mathbf{A}) = \{\mathbf{A} \setminus \{A_i, A_j\} \cup \{A'_i, A'_j\} \mid A_i, A_j \in \mathbf{A}, A_i \cup A_j \in \mathfrak{A}, \{A'_i, A'_j\} \in \mathfrak{S}_2(\mathfrak{A}, A_i \cup A_j)\}$$

Daher folgt die Behauptung aus der Definition 3.7.8 von $\mathcal{U}_2^\varepsilon(\mathbf{A})$ mit Bemerkung 3.7.5.

Bemerkung 3.7.11

Sei $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$. Weiter sei für alle $A_i, A_j \in \mathbf{A}$ mit $A_i \cup A_j \in \mathfrak{A}$ die Segmentierung $\{A_i, A_j\} \in \mathfrak{S}_2(\mathfrak{A}, A_i \cup A_j)$ ein lokales Maximum der 2-Niveaupassung

$$\Psi_2 : \begin{array}{l} \mathfrak{S}_2(\mathfrak{A}, A_i \cup A_j) \rightarrow \mathbb{R}_+ \\ \{B, (A_i \cup A_j) \setminus B\} \mapsto \|\gamma_f(B)\|_2^2 + \|\gamma_f((A_i \cup A_j) \setminus B)\|_2^2 \end{array}$$

Sei nun $\mathbf{A} = \{A_1, \dots, A_k\}$. Nach Voraussetzung existiert dann ein $\varepsilon > 0$, so daß für alle $\mathbf{B} \in \mathcal{U}_2^\varepsilon(\mathbf{A})$ gilt: $\Psi(\mathbf{B}) < \Psi(\mathbf{A})$. Seien nun $\mathbf{B}, \mathbf{C} \in \mathcal{U}_2^\varepsilon(\mathbf{A})$ der Form $\mathbf{B} = \{B_1, B_2, A_3, \dots, A_k\}$ und $\mathbf{C} = \{A_1, \dots, A_{k-2}, C_{k-1}, C_k\}$. Da $\Psi(\mathbf{B}) < \Psi(\mathbf{A})$ ist, muß $\|\gamma_f(B_1)\|_2^2 + \|\gamma_f(B_2)\|_2^2 < \|\gamma_f(A_1)\|_2^2 + \|\gamma_f(A_2)\|_2^2$ gelten. Analog muß $\|\gamma_f(C_{k-1})\|_2^2 + \|\gamma_f(C_k)\|_2^2 < \|\gamma_f(A_{k-1})\|_2^2 + \|\gamma_f(A_k)\|_2^2$ sein. Daher ist auch $\Psi(\{B_1, B_2, A_3, \dots, A_{k-2}, C_{k-1}, C_k\}) < \Psi(\mathbf{A})$. Aus dieser Überlegung kann man folgern, daß für alle $\mathbf{D} \in S_k(\mathfrak{A})$, die durch Verschmelzung zweier Elemente aus $\mathcal{U}_2^\varepsilon(\mathbf{A})$ entstanden sind, $\Psi(\mathbf{D}) < \Psi(\mathbf{A})$ gelten muß.

Definition 3.7.12

Für $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ sei die Menge $\widehat{\mathcal{U}}_2^\varepsilon(\mathbf{A})$ eine Mengen von Segmentierungen definiert durch

$$\widehat{\mathcal{U}}_2^\varepsilon(\mathbf{A}) := \{\mathbf{C} \in \mathfrak{S}_k(\mathfrak{A}) \mid \forall C \in \mathbf{C} : \exists \mathbf{B} \in \mathcal{U}_2^\varepsilon(\mathbf{A}) : C \in \mathbf{B}\}$$

Nach Bemerkung 3.7.11 ist $\widehat{\mathcal{U}}_2^\varepsilon(\mathbf{A})$ die Menge aller Segmentierungen, die durch Verschmelzen (im Sinne der Bemerkung) endlich vieler Segmentierungen, die sich in nur zwei Komponenten von \mathbf{A} unterscheiden, entstanden sind.

Beispiel 3.7.13

Wie in Bemerkung 3.7.7 sei $M := [0, 1]$ und $\mathfrak{A} := \{[x_1, x_2] \mid x_1, x_2 \in [0, 1], x_1 \leq x_2\}$ das Mengensystem der erlaubten Mengen. Betrachtet wird wieder $\mathfrak{S}_3(\mathfrak{A})$, das System der Segmentierungen von $[0, 1]$ durch 3 Intervalle. Sei $\mathbf{A} = \{[0, a_1], [a_1, a_2], [a_2, 1]\}$ ein Element aus $\mathfrak{S}_3(\mathfrak{A})$ mit $0 < a_1 < a_2 < 1$. Weiter sei $\varepsilon > 0$ und $\varepsilon < \min\{\frac{a_1}{2}, \frac{a_2 - a_1}{2}, \frac{1 - a_2}{2}\}$. Die Segmentierung $\mathbf{B} = \{[0, b_1], [b_1, b_2], [b_2, 1]\}$ sei ein beliebiges Element aus $K_\varepsilon(\mathbf{A})$. Dann ist $0 < b_1 < a_2$ und $a_1 < b_2 < 1$. Definiert man nun

$$\begin{aligned} \mathbf{B}_1 &:= \{[0, b_1], [b_1, a_2], [a_2, 1]\} \quad \text{und} \\ \mathbf{B}_2 &:= \{[0, a_1], [a_1, b_2], [b_2, 1]\} \end{aligned}$$

so gilt $\mathbf{B}_1 \in \mathcal{U}_2^\varepsilon(\mathbf{A})$ und $\mathbf{B}_2 \in \mathcal{U}_2^\varepsilon(\mathbf{A})$. Daher muß $\mathbf{B} \in \widehat{\mathcal{U}}_2^\varepsilon(\mathbf{A})$ gelten. Damit folgt:
 $K_\varepsilon(\mathbf{A}) \subseteq \widehat{\mathcal{U}}_2^\varepsilon(\mathbf{A})$.

Es ist also für jedes solche \mathbf{A} die Menge $\widehat{\mathcal{U}}_2^\varepsilon(\mathbf{A})$ eine Umgebung von \mathbf{A} , da sie mit Satz 3.7.6 eine offene Umgebung von \mathbf{A} umschließt.

Ist $\mathbf{A} \in \mathfrak{S}_3(\mathfrak{A})$ der Form $\mathbf{A} = \{[0, 1], O\}$, so kann man sich leicht überlegen, daß $\widehat{\mathcal{U}}_2^\varepsilon(\mathbf{A})$ keine Umgebung von \mathbf{A} ist, da alle Elemente in $\widehat{\mathcal{U}}_2^\varepsilon(\mathbf{A})$ der Form $\{[0, t], [t, 1], O\}$ sein müssen.

Daher kann für kein Mengensystem erlaubter Mengen \mathfrak{A} die Menge $\widehat{\mathcal{U}}_2^\varepsilon(\mathbf{A})$ eine Umgebung von $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ sein, wenn $|\mathbf{A} \setminus \{O\}| = k - 2$ gilt.

Satz 3.7.14

Sei \mathfrak{A} derart, daß für alle $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ mit $O \notin \mathbf{A}$ die Menge $\widehat{\mathcal{U}}_2^\varepsilon(\mathbf{A})$ eine Umgebung von \mathbf{A} ist. Dann gilt für alle $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ mit $O \notin \mathbf{A}$:

\mathbf{A} ist ein lokales Maximum der Zielfunktion der k -Niveaupassung Ψ genau dann, wenn für alle $A_i, A_j \in \mathbf{A}$ mit $A_i \cup A_j \in \mathfrak{A}$ die Segmentierung $\{A_i, A_j\}$ ein lokales Maximum der 2-Niveaupassung

$$\Psi_2 : \begin{array}{l} \mathfrak{S}_2(\mathfrak{A}, A_i \cup A_j) \rightarrow \mathbb{R}_+ \\ \{B, (A_i \cup A_j) \setminus B\} \mapsto \|\gamma_f(B)\|_2^2 + \|\gamma_f((A_i \cup A_j) \setminus B)\|_2^2 \end{array}$$

ist.

Beweis:

folgt mit der Definition 3.7.12 von $\widehat{\mathcal{U}}_2^\varepsilon(\mathbf{A})$ und Bemerkung 3.7.11 sowie Satz 3.7.10

Bemerkung 3.7.15

Die Bemerkung 3.7.13 zeigt, daß es Mengensysteme gibt, in denen die Voraussetzungen des Satzes 3.7.14 erfüllt sind. Die Einschränkung $O \notin \mathbf{A}$ in den Voraussetzungen des Satzes hat für die meisten praktischen Anwendungen keine Bedeutung, weil ein lokales Maximum der Zielfunktion der k -Niveaupassung in den meisten Fällen aus k Segmenten bestehen muß. Lediglich wenn das zu segmentierende Bild $f \in L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$ selbst der Form $f = \sum_{i=1}^l f_{A_i}$ mit $A_i \in \mathfrak{A}$ für $1 \leq i \leq l$ und $l < k$ ist, kommen lokale Maxima mit weniger als k Segmenten vor.

Bemerkung 3.7.16

Wäre die Menge der Segmentierungen der Menge $A_i \cup A_j$ durch zwei Segmente innerhalb der Segmentierung \mathbf{A} definiert worden durch

$$\mathfrak{S}_{A_i, A_j}(\mathbf{A})' := \{\mathbf{A}' \in \mathfrak{S}_k(\mathfrak{A}) \mid \forall A \in \mathbf{A}, A \neq A_i, A \neq A_j : A \in \mathbf{A}'\}$$

so wäre die Einschränkung auf Segmentierungen \mathbf{A} mit $O \notin \mathbf{A}$ in der Voraussetzung des Satzes 3.7.14 nicht nötig. Denn bei dieser Definition werden nicht automatisch zwei Segmente durch zwei, sondern unter Umständen auch durch mehr Segmente ersetzt. Dies war bei der Definition 3.7.3 nicht der Fall.

Da in dieser Arbeit jedoch auch auf die algorithmische Umsetzung des Segmentierungsansatzes geachtet wird, wurde im Hinblick auf das Kapitel 4 (insbesondere Abschnitt 4.3.2) der Ansatz der Definition 3.7.3 gewählt.

3.7.2 Hinweise zur Bestimmung lokaler Minima von 2-Niveaupassungsproblemen

Der vorangegangene Abschnitt 3.7.1 rechtfertigt es, sich bei der Bestimmung lokaler Maxima der Funktion Ψ auf die 2-Niveau-Situation einzuschränken.

Die zentrale Frage dieses Abschnitts kann man wie folgt formulieren:

Zwei Niveaumengen $A \cup B$ und C „streiten“ sich um die Menge B . Soll das Segment $A \cup B$ seine Teilmenge B behalten oder soll es B an die Niveaumenge C abgeben?

Das heißt, es soll die Frage beantwortet werden, wann

$$\|f - f_{A \cup B} - f_C\|_M > \|f - f_A - f_{B \cup C}\|_M$$

gilt.

Die Antwort auf diese Frage ist für die zu entwickelnden Algorithmen von zentraler Bedeutung, da diese immer nur Teilmengen B mit positivem Lebesgue-Maß behandeln können, sich aber nur in der Theorie mit infinitesimal kleinen Mengen auseinandersetzen können. In Anbetracht der in Abschnitt 2.2.2 (S. 18) geforderten algorithmischen Umsetzbarkeit einer Segmentierungstheorie, gewinnt diese Frage zusätzlich an Gewicht. Dabei wird bei der Beantwortung die Ausgangsfunktion f so modifiziert, daß die Frage leichter zu beantworten ist, die Ergebnisse sich aber auf die Originalfunktion f übertragen lassen.

Während dieses Abschnitts sei stets $A, B, C \in \mathfrak{A}$ eine Zerlegung von M , d. h. $A \uplus B \uplus C = M$, in drei erlaubte Mengen mit positivem Lebesgue-Maß. Desweiteren sei im gesamten Abschnitt $f \in L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$ und sowohl $A \cup B \in \mathfrak{A}$ als auch $B \cup C \in \mathfrak{A}$. Weiter sei $\mathcal{E} = (e_1, \dots, e_n)$ eine Basis von V , der Klasse der erlaubten Funktionen und für $D \in \mathfrak{A} \setminus \{O\}$ sei $\mathcal{E}^D = (e_1^D, \dots, e_n^D)$ die ausgezeichnete Basis bezüglich der Basis \mathcal{E} und dem Skalarprodukt $\langle \cdot | \cdot \rangle_D$. Die Familie der Skalarprodukte gemäß Definition 3.4.4 sei gegeben durch $\{\langle \cdot | \cdot \rangle_D \mid D \in \sigma(\mathfrak{A}), \lambda(D) > 0\}$.

Definition 3.7.17

Um eine einfache und intuitive Schreibweise in den Beweisen dieses Abschnitts zu haben, seien die Matrizen $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{n \times n}$ definiert durch

$$\begin{aligned}\mathbf{A} &:= (\langle e_i^A | e_j \rangle_A)_{1 \leq i, j \leq n} \\ \mathbf{B} &:= (\langle e_i^B | e_j \rangle_B)_{1 \leq i, j \leq n} \\ \mathbf{C} &:= (\langle e_i^C | e_j \rangle_C)_{1 \leq i, j \leq n}\end{aligned}$$

Es gilt also mit Definition 3.5.9: $\tilde{\gamma}_f(A) = \mathbf{A}^{-1} \gamma_f(A)$, sowie $\tilde{\gamma}_f(B) = \mathbf{B}^{-1} \gamma_f(B)$ und $\tilde{\gamma}_f(C) = \mathbf{C}^{-1} \gamma_f(C)$.

Definition 3.7.18

Für $f \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ sei $\bar{f} := f_{A,B,C} := f_A + f_B + f_C$.
Offensichtlich ist $f_{A,B,C} \in \mathcal{L}^2(\mathcal{B}(M), \mathcal{B}(E))$.

Lemma 3.7.19

Es gilt $\langle f_A | f_A \rangle_A = \|\gamma_f(A)\|_2^2$ und damit

$$\langle \bar{f} | \bar{f} \rangle_M = \langle f_A | f_A \rangle_A + \langle f_B | f_B \rangle_B + \langle f_C | f_C \rangle_C = \|\gamma_f(A)\|_2^2 + \|\gamma_f(B)\|_2^2 + \|\gamma_f(C)\|_2^2$$

Beweis:

Es ist

$$\begin{aligned}\langle \bar{f} | \bar{f} \rangle_M &= \langle \bar{f} | \bar{f} \rangle_A + \langle \bar{f} | \bar{f} \rangle_B + \langle \bar{f} | \bar{f} \rangle_C \\ &= \langle f_A | f_A \rangle_A + \langle f_B | f_B \rangle_B + \langle f_C | f_C \rangle_C\end{aligned}$$

Für die einzelnen Terme gilt nun:

$$\begin{aligned}\langle f_A | f_A \rangle_A &= \left\langle \sum_{i=1}^n \gamma_f(A)_i e_i^A \left| \sum_{j=1}^n \gamma_f(A)_j e_j^A \right. \right\rangle_A \\ &= \sum_{i=1}^n \gamma_f(A)_i \sum_{j=1}^n \gamma_f(A)_j \langle e_i^A | e_j^A \rangle_A \\ [(e_1^A, \dots, e_n^A) \text{ ist Orthonormalbasis}] &= \sum_{i=1}^n \gamma_f(A)_i^2\end{aligned}$$

Analog ist $\langle f_B | f_B \rangle_B = \sum_{i=1}^n \gamma_f(B)_i^2$ und $\langle f_C | f_C \rangle_C = \sum_{i=1}^n \gamma_f(C)_i^2$. Daher folgt die Behauptung. □

Satz 3.7.20 (Verallgemeinertes Massenerhaltungslemma)

Für g aus V , der Klasse der erlaubten Funktionen, gilt:

- i. $\forall D \in \{A, B, C\} : \langle f|g \rangle_D = \langle f_D|g \rangle_D$
- ii. $\langle f|g \rangle_M = \langle f_{A,B,C}|g \rangle_M$

Beweis:

Da f_D die Orthogonalprojektion von $f|_D$ bezüglich $\langle \cdot | \cdot \rangle_D$ auf V ist (vgl. Bemerkung 3.5.4), gilt

$$\langle f - f_D | g \rangle_D = 0 \Rightarrow \langle f | g \rangle_D - \langle f_D | g \rangle_D = 0 \Rightarrow \langle f | g \rangle_D = \langle f_D | g \rangle_D$$

Damit folgt i.

Aus der Additivität des Skalarprodukt folgt

$$\langle f | g \rangle_M = \langle f | g \rangle_A + \langle f | g \rangle_B + \langle f | g \rangle_C = \langle f_A | g \rangle_A + \langle f_B | g \rangle_B + \langle f_C | g \rangle_C = \langle f_{A,B,C} | g \rangle_M$$

und damit ii. □

Bemerkung 3.7.21

Der Satz 3.7.20 ist eine verallgemeinerte Version des Massenerhaltungslemma 2.5.5 von Kämpke und Kober (vgl. [Käm00]), das in Abschnitt 2.5.1 präsentiert wurde.

Ist $1 \in V$, so gilt mit Satz 3.7.20 $\langle f_{A,B,C} | 1 \rangle_M = \langle f | 1 \rangle_M$. Ist dann $\langle \cdot | \cdot \rangle_M$ definiert wie in Beispiel 3.4.12 i., d. h. es ist $E := \mathbb{R}$ mit der Multiplikation als Skalarprodukt und $\langle f | g \rangle_A := \int_A f g d\lambda$, so gilt

$$\int_M f_{A,B,C} d\lambda = \langle f_{A,B,C} | 1 \rangle_M = \langle f | 1 \rangle_M = \int_M f d\lambda$$

Damit wird die Analogie zum Lemma 2.5.5 deutlich.

Lemma 3.7.22

Es gilt

$$\langle f - f_{A \cup B} - f_C | f - f_{A \cup B} - f_C \rangle_M = \langle f | f \rangle_M - \langle f_{A \cup B} | f_{A \cup B} \rangle_M - \langle f_C | f_C \rangle_M$$

Beweis:

Zunächst einmal gilt für $D \in \{A, B, A \cup B\}$

$$\begin{aligned} \langle f | f_D \rangle_M &= \left\langle f \left| \sum_{i=1}^n \gamma_f(D)_i e_i^D \right. \right\rangle_D \\ &= \sum_{i=1}^n \gamma_f(D)_i \langle f | e_i^D \rangle_D \\ \text{[Def. 3.5.2]} &= \sum_{i=1}^n \gamma_f(D)_i^2 \\ &= \langle f_D | f_D \rangle_D = \langle f_D | f_D \rangle_M \end{aligned} \tag{3.19}$$

Damit kann man nun

$$\begin{aligned} &\langle f - f_{A \cup B} - f_C | f - f_{A \cup B} - f_C \rangle_M \\ &= \langle f | f \rangle_M + \langle f_{A \cup B} | f_{A \cup B} \rangle_M + \langle f_C | f_C \rangle_M \\ &\quad - 2 \langle f | f_{A \cup B} \rangle_M - 2 \langle f | f_C \rangle_M + 2 \langle f_{A \cup B} | f_C \rangle_M \\ \text{[Gleichung (3.19)]} &= \langle f | f \rangle_M + \langle f_{A \cup B} | f_{A \cup B} \rangle_M + \langle f_C | f_C \rangle_M \\ &\quad - 2 \langle f_{A \cup B} | f_{A \cup B} \rangle_M - 2 \langle f_C | f_C \rangle_M + 2 \langle f_{A \cup B} | f_C \rangle_M \\ &= \langle f | f \rangle_M - \langle f_{A \cup B} | f_{A \cup B} \rangle_M + 2 \langle f_{A \cup B} | f_C \rangle_M - \langle f_C | f_C \rangle_M \end{aligned}$$

folgern. Aber mit der Definition 3.5.13 von f_C bzw. $f_{A \cup B}$ folgt:

$$\langle f_{A \cup B} | f_C \rangle_M = \langle f_{A \cup B} | 0 \rangle_{A \cup B} + \langle 0 | f_C \rangle_C = 0 \text{ und damit die Behauptung.}$$

□

Lemma 3.7.23

Es gilt

$$\gamma_f(A \cup B) = \gamma_{\bar{f}}(A \cup B) \quad \text{und} \quad \gamma_f(B \cup C) = \gamma_{\bar{f}}(B \cup C)$$

Beweis:

Für $i \in \{1, \dots, n\}$ ist

$$\begin{aligned} \gamma_{\bar{f}}(A \cup B)_i &= \langle f_{A, B, C} | e_i^{A \cup B} \rangle_{A \cup B} = \langle f_A + f_B | e_i^{A \cup B} \rangle_{A \cup B} \\ &= \langle f_A | e_i^{A \cup B} \rangle_A + \langle f_B | e_i^{A \cup B} \rangle_B \\ \text{[Satz 3.7.20]} &= \langle f | e_i^{A \cup B} \rangle_A + \langle f | e_i^{A \cup B} \rangle_B \\ &= \gamma_f(A \cup B)_i \end{aligned}$$

Damit folgt die Behauptung für $A \cup B$. Der Beweis für $B \cup C$ erfolgt analog.

□

Satz 3.7.24

Sei $A, B, C \in \mathfrak{A}$ eine Zerlegung von M in drei erlaubte Mengen mit jeweils positivem Lebesgue-Maß, dann gilt

$$\begin{aligned} \|f - f_A - f_{B \cup C}\|_M &< \|f - f_{A \cup B} - f_C\|_M \\ &\Leftrightarrow \|f_{B,C} - f_{B \cup C}\|_{B \cup C} < \|f_{A,B} - f_{A \cup B}\|_{A \cup B} \end{aligned}$$

Beweis:

Mit Lemma 3.7.23 und Satz 3.5.16 (S. 100) gilt:

$$\begin{aligned} \|f - f_A - f_{B \cup C}\|_M &< \|f - f_{A \cup B} - f_C\|_M \\ &\Leftrightarrow \|\gamma_f(A)\|_2^2 + \|\gamma_f(B \cup C)\|_2^2 < \|\gamma_f(A \cup B)\|_2^2 + \|\gamma_f(C)\|_2^2 \\ &\Leftrightarrow \|\gamma_{\bar{f}}(A)\|_2^2 + \|\gamma_{\bar{f}}(B \cup C)\|_2^2 < \|\gamma_{\bar{f}}(A \cup B)\|_2^2 + \|\gamma_{\bar{f}}(C)\|_2^2 \\ &\Leftrightarrow \|\bar{f} - \bar{f}_A - \bar{f}_{B \cup C}\|_M < \|\bar{f} - \bar{f}_{A \cup B} - \bar{f}_C\|_M \end{aligned}$$

Aus der Gleichheit $\|\bar{f} - \bar{f}_A - \bar{f}_{B \cup C}\|_M = \|\bar{f} - \bar{f}_{B \cup C}\|_{B \cup C}$ bzw. $\|\bar{f} - \bar{f}_{A \cup B} - \bar{f}_C\|_M = \|\bar{f} - \bar{f}_{A \cup B}\|_{A \cup B}$ folgt dann die Behauptung. \square

Bemerkung 3.7.25

Der Satz 3.7.24 erlaubt es, sich bei der Beantwortung der Eingangs dieses Abschnitts gestellten Frage auf $f_{A,B,C}$, anstelle der Originalfunktion f , zurückzuziehen. Da $f_{A,B,C}$ eingeschränkt auf die Menge A bzw. B oder C aus dem Vektorraum der erlaubten Funktionen ist, wird die Beantwortung der Eingangs des Abschnitts gestellten Frage durch diesen Satz erheblich vereinfacht.

Bemerkung 3.7.26

Es läßt sich sogar ein genauer Zusammenhang zwischen der Approximation von f und der von \bar{f} angeben. Es ist nämlich

$$\|f - f_A - f_{B \cup C}\|_M^2 = \|f\|_M^2 - \|\bar{f}\|_M^2 + \|f_{B,C} - f_{B \cup C}\|_{B \cup C}^2$$

Denn:

Mit Lemma 3.7.22 gilt

$$\begin{aligned} \|f - f_A - f_{B \cup C}\|_M^2 &= \langle f - f_A - f_{B \cup C} | f - f_A - f_{B \cup C} \rangle_M \\ &= \langle f | f \rangle_M - \langle f_A | f_A \rangle_M - \langle f_{B \cup C} | f_{B \cup C} \rangle_M \\ &= \langle f | f \rangle_M - \langle f_{A,B,C} | f_{A,B,C} \rangle_M + \langle f_{A,B,C} | f_{A,B,C} \rangle_M \\ &\quad - \langle f_A | f_A \rangle_M - \langle f_{B \cup C} | f_{B \cup C} \rangle_M \\ &= \|f\|_M^2 - \|f_{A,B,C}\|_M^2 \\ &\quad + \langle f_{A,B,C} | f_{A,B,C} \rangle_M - \langle f_A | f_A \rangle_M - \langle f_{B \cup C} | f_{B \cup C} \rangle_M \\ \text{[Lemma 3.7.22]} &= \|f\|_M^2 - \|f_{A,B,C}\|_M^2 + \|f_{A,B,C} - f_A - f_{B \cup C}\|_M^2 \\ &= \|f\|_M^2 - \|f_{A,B,C}\|_M^2 + \|f_{B,C} - f_{B \cup C}\|_M^2 \end{aligned}$$

Definition 3.7.27

Seien $A, B, C \in \mathfrak{A}$ mit $\lambda(A) > 0, \lambda(B) > 0$ und $\lambda(C) > 0$. Dann sei $\mathbf{E}_C^{A,B} \in \mathbb{R}^{n \times n}$ definiert durch

$$\mathbf{E}_C^{A,B} := (\langle e_i^A | e_j^B \rangle_C)_{1 \leq i, j \leq n}$$

$\mathbf{E}_{C^{i,j}}^{A,B} := \langle e_i^A | e_j^B \rangle_C$ bezeichne das j -te Element der i -ten Zeile der Matrix $\mathbf{E}_C^{A,B}$.

Weiter sei $\mathbf{E}_B^{A,\varepsilon} := (\langle e_i^A | e_j \rangle_B)_{1 \leq i, j \leq n}$.

Lemma 3.7.28

Sind $f, g \in V$, so gilt für alle $A \in \mathfrak{A}$

$$\langle f | g \rangle_A = \sum_{i=1}^n \langle f | e_i^A \rangle_A \langle g | e_i^A \rangle_A$$

Beweis:

vgl. [Die85], 6.5.2., S.125.

Lemma 3.7.29

Mit \mathbf{B} und \mathbf{C} wie in Definition 3.7.17 gilt

$$\mathbf{E}_B^{B \cup C, B} \mathbf{B} = \mathbf{E}_B^{B \cup C, \varepsilon} \quad \text{und} \quad \mathbf{E}_C^{B \cup C, C} \mathbf{C} = \mathbf{E}_C^{B \cup C, \varepsilon}$$

Insbesondere sind die Matrizen $\mathbf{E}_B^{B \cup C, \varepsilon}$ und $\mathbf{E}_C^{B \cup C, \varepsilon}$ invertierbar.

Beweis:

Zunächst ist

$$\begin{aligned} \left(\mathbf{E}_B^{B \cup C, B} \mathbf{B} \right)_{i,j} &= \sum_{r=1}^n \langle e_i^{B \cup C} | e_r^B \rangle_B \langle e_j | e_r^B \rangle_B \\ \text{[Lemma 3.7.28]} &= \langle e_i^{B \cup C} | e_j \rangle_B \\ &= \mathbf{E}_B^{B \cup C, \varepsilon} \end{aligned}$$

Daher ist $\mathbf{E}_B^{B \cup C, B} \mathbf{B} = \mathbf{E}_B^{B \cup C, \varepsilon}$. Die Matrix $\mathbf{E}_B^{B \cup C, B}$ ist die Matrix der Basis $(e_1^{B \cup C}, \dots, e_n^{B \cup C})$ in der Koordinatendarstellung bezüglich der Basis (e_1^B, \dots, e_n^B) . Die Matrix \mathbf{B} ist ebenfalls eine Basisdarstellung, daher ist $\mathbf{E}_B^{B \cup C, \varepsilon}$ das Produkt zweier invertierbarer Matrizen und damit selbst invertierbar. \square

Bemerkung 3.7.30

Der folgende Satz wird in diesem Kapitel nur im Sinne eines Lemmas, d. h. als Zwischenergebnis, genutzt. Er wird jedoch bei den praktischen Anwendungen der Mehrniveaupassung eine herausragende Rolle spielen. Insbesondere wird er innerhalb des Kontexts der Mehrniveaupassung das kontrollierte Zusammenfügen („Merging“) von Niveaumengen erlauben (vgl. Abschnitt 5.3). Daher hat er auch in diesem Kapitel den Namen „Satz“ verdient.

Satz 3.7.31

Es gilt

$$\gamma_f(B \cup C) = \mathbf{E}_B^{B \cup C, B} \gamma_f(B) + \mathbf{E}_C^{B \cup C, C} \gamma_f(C)$$

Beweis:

Zunächst einmal ist nach Lemma 3.7.23 $\gamma_{\bar{f}}(B \cup C) = \gamma_f(B \cup C)$. Weiter ist für alle $i \in \{1, \dots, n\}$:

$$\begin{aligned} \gamma_{\bar{f}}(B \cup C)_i &= \langle \bar{f} | e_i^{B \cup C} \rangle_{B \cup C} \\ &= \langle f_B + f_C | e_i^{B \cup C} \rangle_{B \cup C} \\ &= \langle f_B | e_i^{B \cup C} \rangle_B + \langle f_C | e_i^{B \cup C} \rangle_C \end{aligned} \quad (3.20)$$

Ferner ist

$$\begin{aligned} \langle f_B | e_i^{B \cup C} \rangle_B &= \left\langle \sum_{j=1}^n \gamma_f(B)_j e_j^B \middle| e_i^{B \cup C} \right\rangle_B \\ &= \sum_{j=1}^n \gamma_f(B)_j \langle e_j^B | e_i^{B \cup C} \rangle_B \end{aligned}$$

und analog $\langle f_C | e_i^{B \cup C} \rangle_C = \sum_{j=1}^n \gamma_f(C)_j \langle e_j^B | e_i^{B \cup C} \rangle_C$. Daher gilt mit Gleichung (3.20)

$$\begin{aligned} \gamma_{\bar{f}}(B \cup C)_i &= \sum_{j=1}^n \gamma_f(B)_j \langle e_j^B | e_i^{B \cup C} \rangle_B + \sum_{l=1}^n \gamma_f(C)_l \langle e_l^B | e_i^{B \cup C} \rangle_C \\ &= \sum_{j=1}^n \left(\gamma_f(B)_j \mathbf{E}_B^{B \cup C, B} + \gamma_f(C)_j \mathbf{E}_C^{B \cup C, C} \right) \end{aligned} \quad (3.21)$$

und damit insgesamt:

$$\gamma_f(B \cup C) = \gamma_{\bar{f}}(B \cup C) = \mathbf{E}_B^{B \cup C, B} \gamma_f(B) + \mathbf{E}_C^{B \cup C, C} \gamma_f(C)$$

□

Lemma 3.7.32

Es ist

$$\left\| \bar{f} - \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C)_i e_i^{B \cup C} \right\|_{B \cup C}^2 = \|\bar{f}\|_{B \cup C}^2 - \left\langle \bar{f} \left| \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C)_i e_i^{B \cup C} \right. \right\rangle_{B \cup C}$$

Beweis:

Es ist $\sum_{i=1}^n \gamma_{\bar{f}}(B \cup C)_i e_i^{B \cup C} = f_{B \cup C}$ nach Bemerkung 3.5.4 und Lemma 3.7.23 die orthogonale Projektion von \bar{f} auf V bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_{B \cup C}$. Daher ist analog zu Lemma 3.7.22 $\|\bar{f} - f_{B \cup C}\|_{B \cup C}^2 = \|\bar{f}\|_{B \cup C}^2 - \|f_{B \cup C}\|_{B \cup C}^2$. Weiter gilt

$$\begin{aligned} \|f_{B \cup C}\|_{B \cup C}^2 &= \langle f_{B \cup C} | f_{B \cup C} \rangle_{B \cup C} \\ &= \left\langle \sum_{i=1}^n \gamma_f(B \cup C)_i e_i^{B \cup C} \left| f_{B \cup C} \right. \right\rangle_{B \cup C} \\ &= \sum_{i=1}^n \gamma_f(B \cup C)_i \langle e_i^{B \cup C} | f_{B \cup C} \rangle_{B \cup C} \\ [\text{Lemma 3.7.23 + Gleichung (3.20)}] &= \sum_{i=1}^n \langle f_B + f_C | e_i^{B \cup C} \rangle_{B \cup C} \langle e_i^{B \cup C} | f_{B \cup C} \rangle_{B \cup C} \\ [\text{Lemma 3.7.28}] &= \langle f_{B,C} | f_{B \cup C} \rangle_{B \cup C} \\ &= \left\langle \bar{f} \left| \sum_{i=1}^n \gamma_f(B \cup C)_i e_i^{B \cup C} \right. \right\rangle_{B \cup C} \end{aligned}$$

□

Lemma 3.7.33

Für $1 \leq i \leq n$ sei $\mathbf{B}^{(i)} \in \mathbb{R}^{n \times n}$ durch $\mathbf{B}_{r,l}^{(i)} := \mathbf{E}_{B^{i,l}}^{B \cup C, B} \mathbf{E}_{B^{i,r}}^{B \cup C, B}$ und $\mathbf{C}^{(i)} \in \mathbb{R}^{n \times n}$ durch $\mathbf{C}_{r,l}^{(i)} := \mathbf{E}_{C^{i,l}}^{B \cup C, C} \mathbf{E}_{C^{i,r}}^{B \cup C, C}$ sowie $\mathbf{G}^{(i)} \in \mathbb{R}^{n \times n}$ durch $\mathbf{G}_{r,l}^{(i)} := \mathbf{E}_{C^{i,l}}^{B \cup C, C} \mathbf{E}_{B^{i,r}}^{B \cup C, B}$ definiert (für $r, l \in \{1, \dots, n\}$).

Dann gilt:

$$\begin{aligned} &\left\langle \bar{f} \left| \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C)_i e_i^{B \cup C} \right. \right\rangle_{B \cup C} \\ &= \sum_{i=1}^n \gamma_f(B)^t \mathbf{B}^{(i)} \gamma_f(B) + 2 \sum_{i=1}^n \gamma_f(B)^t \mathbf{G}^{(i)} \gamma_f(C) + \sum_{i=1}^n \gamma_f(C)^t \mathbf{C}^{(i)} \gamma_f(C) \end{aligned}$$

Beweis:

Es ist

$$\begin{aligned}
& \left\langle f \middle| \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C) e_i^{B \cup C} \right\rangle_{B \cup C} \\
&= \left\langle f_B \middle| \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C) e_i^{B \cup C} \right\rangle_{B \cup C} + \left\langle f_C \middle| \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C) e_i^{B \cup C} \right\rangle_{B \cup C} \\
&= \left\langle f_B \middle| \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C) e_i^{B \cup C} \right\rangle_B + \left\langle f_C \middle| \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C) e_i^{B \cup C} \right\rangle_C \tag{3.22}
\end{aligned}$$

Nun ist aber

$$\begin{aligned}
\langle f_B | e_i^{B \cup C} \rangle_B &= \left\langle \sum_{r=1}^n \gamma_f(B)_r e_r^B \middle| e_i^{B \cup C} \right\rangle_B = \sum_{r=1}^n \gamma_f(B)_r \langle e_r^B | e_i^{B \cup C} \rangle_B \\
&= \sum_{r=1}^n \gamma_f(B)_r \mathbf{E}_{B^{i,r}}^{B \cup C, B} \tag{3.23}
\end{aligned}$$

Für den Term $\langle f_B | \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C) e_i^{B \cup C} \rangle_B$ gilt damit

$$\begin{aligned}
& \left\langle f_B \middle| \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C) e_i^{B \cup C} \right\rangle_B \\
&= \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C)_i \langle f_B | e_i^{B \cup C} \rangle_B \\
\text{[Gleichung (3.21)]} &= \sum_{i=1}^n \left(\sum_{l=1}^n \mathbf{E}_{B^{i,l}}^{B \cup C, B} \gamma_f(B)_l + \mathbf{E}_{B^{i,l}}^{B \cup C, C} \gamma_f(C)_l \right) \langle f_B | e_i^{B \cup C, B} \rangle_B \\
&= \sum_{i=1}^n \sum_{l=1}^n \mathbf{E}_{B^{i,l}}^{B \cup C, B} \gamma_f(B)_l \langle f_B | e_i^{B \cup C, B} \rangle_B \\
&\quad + \sum_{i=1}^n \sum_{l=1}^n \mathbf{E}_{C^{i,l}}^{B \cup C, C} \gamma_f(C)_l \langle f_B | e_i^{B \cup C, B} \rangle_B \\
\text{[Gleichung (3.23)]} &= \sum_{i=1}^n \sum_{l=1}^n \mathbf{E}_{B^{i,l}}^{B \cup C, B} \gamma_f(B)_l \left(\sum_{r=1}^n \gamma_f(B)_r \mathbf{E}_{B^{i,r}}^{B \cup C, B} \right) \\
&\quad + \sum_{i=1}^n \sum_{l=1}^n \mathbf{E}_{C^{i,l}}^{B \cup C, C} \gamma_f(C)_l \left(\sum_{r=1}^n \gamma_f(B)_r \mathbf{E}_{B^{i,r}}^{B \cup C, C} \right) \\
&= \sum_{i=1}^n \sum_{l=1}^n \sum_{r=1}^n \gamma_f(B)_l \mathbf{E}_{B^{i,l}}^{B \cup C, B} \mathbf{E}_{B^{i,r}}^{B \cup C, B} \gamma_f(B)_r \\
&\quad + \sum_{i=1}^n \sum_{l=1}^n \sum_{r=1}^n \gamma_f(B)_r \mathbf{E}_{B^{i,r}}^{B \cup C, B} \mathbf{E}_{C^{i,l}}^{B \cup C, C} \gamma_f(C)_l
\end{aligned}$$

Analog folgt

$$\begin{aligned} \left\langle f_C \left| \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C) e_i^{B \cup C} \right. \right\rangle_C &= \sum_{i=1}^n \sum_{l=1}^n \sum_{r=1}^n \gamma_f(C)_l \mathbf{E}_C^{B \cup C, C}{}_{i,l} \mathbf{E}_C^{B \cup C, C}{}_{i,r} \gamma_f(C)_r \\ &\quad + \sum_{i=1}^n \sum_{r=1}^n \sum_{l=1}^n \gamma_f(B)_r \mathbf{E}_B^{B \cup C, B}{}_{i,r} \mathbf{E}_C^{B \cup C, C}{}_{i,l} \gamma_f(C)_l \end{aligned}$$

und damit die Behauptung durch Einsetzen in die Gleichung (3.22). \square

Lemma 3.7.34

Es gilt

$$\begin{aligned} \left\langle \bar{f} \left| \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C) e_i^{B \cup C} \right. \right\rangle_{B \cup C} &= \gamma_f(B)^t \left(\mathbf{E}_B^{B \cup C, B} \right)^t \mathbf{E}_B^{B \cup C, B} \gamma_f(B) \\ &\quad + 2\gamma_f(B)^t \left(\mathbf{E}_B^{B \cup C, B} \right)^t \mathbf{E}_C^{B \cup C, C} \gamma_f(C) \\ &\quad + \gamma_f(C)^t \left(\mathbf{E}_C^{B \cup C, C} \right)^t \mathbf{E}_C^{B \cup C, C} \gamma_f(C) \end{aligned}$$

Beweis:

Mit den Bezeichnungen aus Lemma 3.7.33 gilt:

$$\begin{aligned} \sum_{i=1}^n \gamma_f(B)^t \mathbf{B}^{(i)} \gamma_f(B) &= \sum_{i=1}^n \sum_{l=1}^n \sum_{r=1}^n \gamma_f(B)_r \mathbf{B}_{r,l}^{(i)} \gamma_f(B)_l \\ &= \sum_{r=1}^n \sum_{l=1}^n \gamma_f(B)_r \gamma_f(B)_l \left(\sum_{i=1}^n \mathbf{B}_{r,l}^{(i)} \right) \end{aligned}$$

Aber es ist

$$\sum_{i=1}^n \mathbf{B}_{r,l}^{(i)} = \sum_{i=1}^n \mathbf{E}_B^{B \cup C, B}{}_{i,l} \mathbf{E}_B^{B \cup C, B}{}_{i,r} = \left(\left(\mathbf{E}_B^{B \cup C, B} \right)^t \mathbf{E}_B^{B \cup C, B} \right)_{r,l}$$

Der Rest folgt analog. \square

Satz 3.7.35

Sei $A, B, C \in \mathfrak{A}$ eine Zerlegung von M .

Dann gilt:

i. Die Matrix $\left(\mathbf{E}_B^{B \cup C, \varepsilon}\right)^t \mathbf{E}_C^{B \cup C, \varepsilon}$ ist symmetrisch.

ii. Es gilt

$$\|\bar{f} - \bar{f}_A - \bar{f}_{B \cup C}\|_M^2 = (\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C))^t \left(\mathbf{E}_B^{B \cup C, \varepsilon}\right)^t \mathbf{E}_C^{B \cup C, \varepsilon} (\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C))$$

Beweis:

Zunächst wird die Behauptung ii. bewiesen: Nach Lemma 3.7.32 folgt mit den Lemmata 3.7.19 und 3.7.34

$$\begin{aligned} \|\bar{f} - \bar{f}_A - \bar{f}_{B \cup C}\|_M^2 &= \left\| \bar{f} - \sum_{i=1}^n \gamma_{\bar{f}}(B \cup C)_i e_i^{B \cup C} \right\|_{B \cup C}^2 \\ &= \|\gamma_f(B)\|_2^2 + \|\gamma_f(C)\|_2^2 \\ &\quad - \gamma_f(B)^t \left(\mathbf{E}_B^{B \cup C, B}\right)^t \mathbf{E}_B^{B \cup C, B} \gamma_f(B) \\ &\quad - 2\gamma_f(B)^t \left(\mathbf{E}_B^{B \cup C, B}\right)^t \mathbf{E}_C^{B \cup C, C} \gamma_f(C) \\ &\quad - \gamma_f(C)^t \left(\mathbf{E}_C^{B \cup C, C}\right)^t \mathbf{E}_C^{B \cup C, C} \gamma_f(C) \\ &= \gamma_f(B)^t \mathbf{I} \gamma_f(B) + \gamma_f(C)^t \mathbf{I} \gamma_f(C) \\ &\quad - \gamma_f(B)^t \left(\mathbf{E}_B^{B \cup C, B}\right)^t \mathbf{E}_B^{B \cup C, B} \gamma_f(B) \\ &\quad - 2\gamma_f(B)^t \left(\mathbf{E}_B^{B \cup C, B}\right)^t \mathbf{E}_C^{B \cup C, C} \gamma_f(C) \\ &\quad - \gamma_f(C)^t \left(\mathbf{E}_C^{B \cup C, C}\right)^t \mathbf{E}_C^{B \cup C, C} \gamma_f(C) \\ &= \gamma_f(B)^t \left(\mathbf{I} - \left(\mathbf{E}_B^{B \cup C, B}\right)^t \mathbf{E}_B^{B \cup C, B}\right) \gamma_f(B) \\ &\quad - 2\gamma_f(B)^t \left(\mathbf{E}_B^{B \cup C, B}\right)^t \mathbf{E}_C^{B \cup C, C} \gamma_f(C) \\ &\quad + \gamma_f(C)^t \left(\mathbf{I} - \left(\mathbf{E}_C^{B \cup C, C}\right)^t \mathbf{E}_C^{B \cup C, C}\right) \gamma_f(C) \\ &= (\mathbf{B} \tilde{\gamma}_f(B))^t \left(\mathbf{I} - \left(\mathbf{E}_B^{B \cup C, B}\right)^t \mathbf{E}_B^{B \cup C, B}\right) (\mathbf{B} \tilde{\gamma}_f(B)) \\ &\quad - 2(\mathbf{B} \tilde{\gamma}_f(B))^t \left(\mathbf{E}_B^{B \cup C, B}\right)^t \mathbf{E}_C^{B \cup C, C} (\mathbf{C} \tilde{\gamma}_f(C)) \\ &\quad + (\mathbf{C} \tilde{\gamma}_f(C))^t \left(\mathbf{I} - \left(\mathbf{E}_C^{B \cup C, C}\right)^t \mathbf{E}_C^{B \cup C, C}\right) (\mathbf{C} \tilde{\gamma}_f(C)) \end{aligned}$$

$$= \tilde{\gamma}_f(B)^t \mathbf{B}^t \left(\mathbf{I} - \left(\mathbf{E}_B^{B \cup C, B} \right)^t \mathbf{E}_B^{B \cup C, B} \right) \mathbf{B} \tilde{\gamma}_f(B) \quad (3.24)$$

$$- 2\tilde{\gamma}_f(B)^t \mathbf{B}^t \left(\mathbf{E}_B^{B \cup C, B} \right)^t \mathbf{E}_C^{B \cup C, C} \mathbf{C} \tilde{\gamma}_f(C) \quad (3.25)$$

$$+ \tilde{\gamma}_f(C)^t \mathbf{C}^t \left(\mathbf{I} - \left(\mathbf{E}_C^{B \cup C, C} \right)^t \mathbf{E}_C^{B \cup C, C} \right) \mathbf{C} \tilde{\gamma}_f(C) \quad (3.26)$$

Dabei bezeichnet \mathbf{I} die $n \times n$ Einheitsmatrix. Gezeigt wird nun, daß sich die Terme (3.24)-(3.26) auf die Form

$$\begin{aligned} \tilde{\gamma}_f(B)^t \mathbf{H} \tilde{\gamma}_f(B) - 2\tilde{\gamma}_f(B)^t \mathbf{H} \tilde{\gamma}_f(C) + \tilde{\gamma}_f(C)^t \mathbf{H} \tilde{\gamma}_f(C) \\ = (\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C))^t \mathbf{H} (\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C)) \end{aligned}$$

mit $\mathbf{H} = \left(\mathbf{E}_B^{B \cup C, \varepsilon} \right)^t \mathbf{E}_C^{B \cup C, \varepsilon}$ bringen lassen.

Mit Lemma 3.7.29 folgt

$$\begin{aligned} \mathbf{B}^t \left(\mathbf{E}_B^{B \cup C, B} \right)^t \mathbf{E}_C^{B \cup C, C} \mathbf{C} &= \left(\mathbf{E}_B^{B \cup C, B} \mathbf{B} \right)^t \mathbf{E}_C^{B \cup C, C} \mathbf{C} \\ &= \left(\mathbf{E}_B^{B \cup C, \varepsilon} \right)^t \mathbf{E}_C^{B \cup C, \varepsilon} \end{aligned} \quad (3.27)$$

Behauptung 1:

Für $i, j \in \{1, \dots, n\}$ gilt

$$\langle e_i | e_j \rangle_B = \sum_{r=1}^n \langle e_r^{B \cup C} | e_i \rangle_B (\langle e_r^{B \cup C} | e_j \rangle_B + \langle e_r^{B \cup C} | e_j \rangle_C) \quad (3.28)$$

Denn aus der Definition des Skalarprodukts $\langle \cdot | \cdot \rangle_{B \cup C}$ folgt

$$\langle e_r^{B \cup C} | e_j \rangle_{B \cup C} = \langle e_r^{B \cup C} | e_j \rangle_B + \langle e_r^{B \cup C} | e_j \rangle_C$$

Damit gilt

$$\begin{aligned} \langle e_i | e_j \rangle_B &= \left\langle e_i \left| \sum_{r=1}^n \langle e_j | e_r^{B \cup C} \rangle_{B \cup C} e_r^{B \cup C} \right. \right\rangle_B \\ &= \sum_{r=1}^n \langle e_r^{B \cup C} | e_i \rangle_B \langle e_r^{B \cup C} | e_j \rangle_{B \cup C} \\ &= \sum_{r=1}^n \langle e_r^{B \cup C} | e_i \rangle_B (\langle e_r^{B \cup C} | e_j \rangle_B + \langle e_r^{B \cup C} | e_j \rangle_C) \end{aligned}$$

Behauptung 2:

Es gilt

$$\mathbf{B}^t \left(\mathbf{I} - \left(\mathbf{E}_B^{B \cup C, B} \right)^t \mathbf{E}_B^{B \cup C, B} \right) \mathbf{B} = \left(\mathbf{E}_B^{B \cup C, \varepsilon} \right)^t \mathbf{E}_C^{B \cup C, \varepsilon} \quad (3.29)$$

Denn Lemma 3.7.29 gilt:

$$\begin{aligned} & \mathbf{B}^t \left(\mathbf{I} - \left(\mathbf{E}_B^{B \cup C, B} \right)^t \mathbf{E}_B^{B \cup C, B} \right) \mathbf{B} \\ &= \mathbf{B}^t \mathbf{B} - \mathbf{B}^t \left(\mathbf{E}_B^{B \cup C, B} \right)^t \mathbf{E}_B^{B \cup C, B} \mathbf{B} \\ &= \mathbf{B}^t \mathbf{B} - \left(\mathbf{E}_B^{B \cup C, \varepsilon} \right)^t \mathbf{E}_B^{B \cup C, \varepsilon} \end{aligned}$$

Setzt man $(T_{i,j})_{1 \leq i, j \leq n} := \mathbf{B}^t \mathbf{B} - \left(\mathbf{E}_B^{B \cup C, \varepsilon} \right)^t \mathbf{E}_B^{B \cup C, \varepsilon}$ so ist für $i, j \in \{1, \dots, n\}$

$$\begin{aligned} T_{i,j} &= \sum_{r=1}^n \langle e_i | e_r^B \rangle_B \langle e_j | e_r^B \rangle_B - \sum_{r=1}^n \langle e_i | e_r^{B \cup C} \rangle_B \langle e_j | e_r^{B \cup C} \rangle_B \\ [\text{Lemma 3.7.28}] &= \langle e_i | e_j \rangle_B - \sum_{r=1}^n \langle e_i | e_r^{B \cup C} d \rangle_B \langle e_j | e_r^{B \cup C} \rangle_B \\ [\text{Gleichung (3.28)}] &= \sum_{r=1}^n \langle e_r^{B \cup C} | e_i \rangle_B (\langle e_r^{B \cup C} | e_j \rangle_B + \langle e_r^{B \cup C} | e_j \rangle_C) \\ &\quad - \sum_{r=1}^n \langle e_i | e_r^{B \cup C} d \rangle_B \langle e_j | e_r^{B \cup C} \rangle_B \\ &= \sum_{r=1}^n \langle e_r^{B \cup C} | e_i \rangle_B \langle e_r^{B \cup C} | e_j \rangle_C \\ &\quad + \sum_{r=1}^n \langle e_r^{B \cup C} | e_i \rangle_B \langle e_r^{B \cup C} | e_j \rangle_B \\ &\quad - \sum_{r=1}^n \langle e_r^{B \cup C} | e_i \rangle_B \langle e_r^{B \cup C} | e_j \rangle_B \\ &= \sum_{r=1}^n \langle e_r^{B \cup C} | e_i \rangle_B \langle e_r^{B \cup C} | e_j \rangle_C \\ &= \left[\left(\mathbf{E}_B^{B \cup C, \varepsilon} \right)^t \mathbf{E}_C^{B \cup C, \varepsilon} \right]_{i,j} \end{aligned}$$

Daher folgt die Behauptung 2.

Analog zeigt man

$$\mathbf{C}^t \left(\mathbf{I} - \left(\mathbf{E}_C^{B \cup C, C} \right)^t \mathbf{E}_C^{B \cup C, C} \right) \mathbf{C} = \left(\mathbf{E}_B^{B \cup C, \varepsilon} \right)^t \mathbf{E}_C^{B \cup C, \varepsilon} \quad (3.30)$$

Setzt man nun die Gleichungen (3.29), (3.27) und (3.30) in die Terme (3.24), (3.25) und (3.26) ein, so folgt die Behauptung ii. des Satzes

Mit Gleichung (3.29) folgt die Symmetrie der Matrix $\left(\mathbf{E}_B^{B \cup C, \varepsilon} \right)^t \mathbf{E}_C^{B \cup C, \varepsilon}$ und damit die Behauptung i. des Satzes. □

Korollar 3.7.36

Die Matrix $\begin{pmatrix} \mathbf{E}_A^{A \cup B, \mathcal{E}} \\ \mathbf{E}_B^{A \cup B, \mathcal{E}} \end{pmatrix}^t \mathbf{E}_B^{A \cup B, \mathcal{E}}$ ist positiv definit.

Beweis:

Eine Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ ist genau dann positiv definit, wenn für alle $x \in \mathbb{R}^n$ mit $x \neq \mathbf{0}$ gilt: $x^t \mathbf{A} x > 0$. Für die Matrix $\begin{pmatrix} \mathbf{E}_A^{A \cup B, \mathcal{E}} \\ \mathbf{E}_B^{A \cup B, \mathcal{E}} \end{pmatrix}^t \mathbf{E}_B^{A \cup B, \mathcal{E}}$ gilt aber mit Satz 3.7.35 ii. daß für alle f

$$(\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B))^t \begin{pmatrix} \mathbf{E}_A^{A \cup B, \mathcal{E}} \\ \mathbf{E}_B^{A \cup B, \mathcal{E}} \end{pmatrix}^t \mathbf{E}_B^{A \cup B, \mathcal{E}} (\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B)) \geq 0$$

gelten muß, da der Abstand $\|f - f_{A \cup B} - f_C\|_M^2$ niemals negativ sein kann. Es ist

$$\|f - f_{A \cup B} - f_C\|_M^2 = \|f - f_{A \cup B}\|_A^2 + \|f - f_{A \cup B}\|_B^2 + \underbrace{\|f - f_C\|_C^2}_{=0}$$

Daher ist $\|f - f_{A \cup B} - f_C\|_M^2$ genau dann 0, wenn $\|f - f_{A \cup B}\|_A^2 = 0$ und $\|f - f_{A \cup B}\|_B^2 = 0$ gilt. In dem Fall muß also $f|_{A \cup B} = f_{A \cup B}$ gelten. Also hat $f|_{A \cup B}$ eine eindeutig bestimmte Darstellung bezüglich der Basis \mathcal{E} . Es sei diese $f|_{A \cup B} = \sum_{i=1}^n \alpha_i e_i|_{A \cup B}$. Mit dieser Darstellung ergibt sich $\gamma_f(A)_i = \left\langle \sum_{j=1}^n \alpha_j e_j|_{A \cup B} \middle| e_i^A \right\rangle_A = \sum_{j=1}^n \alpha_j \langle e_j | e_i^A \rangle_A$ und mit $\alpha = (\alpha_1, \dots, \alpha_n)$. Daher gilt mit Definition 3.7.17 $\gamma_f(A) = \mathbf{A}\alpha$. Analog ist $\gamma_f(B) = \mathbf{B}\alpha$. Nach den Definitionen 3.5.9 und 3.7.17 gilt daher

$$\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B) = \mathbf{A}^{-1}\gamma_f(A) - \mathbf{B}^{-1}\gamma_f(B) = \alpha - \alpha = 0$$

Also gilt $x^t \begin{pmatrix} \mathbf{E}_A^{A \cup B, \mathcal{E}} \\ \mathbf{E}_B^{A \cup B, \mathcal{E}} \end{pmatrix}^t \mathbf{E}_B^{A \cup B, \mathcal{E}} x = 0 \Leftrightarrow x = \mathbf{0}$.

Daher ist nur noch zu zeigen, daß sich alle $x \in \mathbb{R}^n$ durch eine Funktion $f \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ derart kodieren lassen, daß $x = (\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B))$ gilt. Zunächst einmal wird gezeigt, daß sich für jedes $y \in \mathbb{R}^n$ eine Funktion $f_y \in V$ konstruieren läßt, mit $\gamma_{f_y}(A) = y$:

Da f_y aus V ist, hat f_y eine Darstellung der Form $f_y = \sum_{i=1}^n \alpha_i e_i^A$, wobei für alle $i \in \{1, \dots, n\}$ die Skalare $\alpha_i \in \mathbb{R}$ eindeutig bestimmt sind. Setzt man $\alpha_i = y_i$, so gilt nach Definition 3.5.2

$$\gamma_{f_y}(A)_i = \langle f_y | e_i^A \rangle_A = \left\langle \sum_{j=1}^n y_j e_j^A \middle| e_i^A \right\rangle_A = \sum_{j=1}^n y_j \langle e_j^A | e_i^A \rangle_A = y_i$$

da (e_1^A, \dots, e_n^A) eine Orthonormalbasis bezüglich $\langle \cdot | \cdot \rangle_A$ ist.

Für ein beliebiges $x \in \mathbb{R}^n$ setzt man $y := \mathbf{A}x$ und konstruiert ein f_y mit $\gamma_{f_y}(A) = y$ wie oben. Definiert man dann $f \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ durch $f|_A = f_y$ und $f|_{B \cup C} = 0$, so folgt mit Punkt v. der Definition 3.4.4 $\gamma_f(A) = y$ und $\gamma_f(B) = 0$. Daher ist

$$\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B) = \mathbf{A}^{-1}\gamma_f(A) - \mathbf{B}^{-1}\gamma_f(B) = \mathbf{A}^{-1}\mathbf{A}x - \mathbf{B}^{-1}0 = x$$

für beliebige $x \in \mathbb{R}^n$. □

Korollar 3.7.37

Sei $A, B, C \in \mathfrak{A}$ eine Zerlegung von M . Die Matrizen $\mathbf{E}_B^{B \cup C, \varepsilon}$ und $\mathbf{E}_C^{B \cup C, \varepsilon}$ sowie $\mathbf{E}_A^{A \cup B, \varepsilon}$ und $\mathbf{E}_B^{A \cup B, \varepsilon}$ seien definiert wie in Definition 3.7.27.

Dann gilt

$$\begin{aligned} \|f - f_A - f_{B \cup C}\|_M &< \|f - f_{A \cup B} - f_C\|_M \Leftrightarrow \\ &(\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C))^t \left(\mathbf{E}_B^{B \cup C, \varepsilon} \right)^t \mathbf{E}_C^{B \cup C, \varepsilon} (\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C)) \\ &< (\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B))^t \left(\mathbf{E}_A^{A \cup B, \varepsilon} \right)^t \mathbf{E}_B^{A \cup B, \varepsilon} (\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B)) \end{aligned}$$

Beweis:

Der Beweis folgt mit Satz 3.7.24 und Satz 3.7.35. □

Beispiel 3.7.38

Alle hier aufgeführten Beispiele sind Konkretisierungen von Beispiel 3.4.6. Die Auflistung der einzelnen Anwendungsbeispiele folgt den Beispielen 3.4.12 und 3.5.11 auf den Seiten 90f und 96ff.

- i) Sei $E := \mathbb{R}$ mit der Multiplikation \cdot als Skalarprodukt und $V := \{\alpha \mid \alpha \in \mathbb{R}\}$ der Vektorraum der konstanten Funktionen. Dann ist (1) eine Basis dieses Raums und für $A \in \mathfrak{A}$ mit $\lambda(A) > 0$ ist $e_1^A = \frac{1}{\sqrt{\lambda(A)}}$ die ausgezeichnete Orthonormalbasis bezüglich der Basis (1) und $\langle \cdot | \cdot \rangle_A$. Damit ist $\gamma_f(A) = \frac{1}{\sqrt{\lambda(A)}} \int_A f d\lambda$ und $\tilde{\gamma}_f(A) = \frac{1}{\lambda(A)} \int_A f d\lambda$. Die Matrix $\mathbf{E}_B^{B \cup C, \varepsilon} \in \mathbb{R}^{1 \times 1}$ ist dann durch $\langle e_1^{B \cup C} | e_1 \rangle_B = \int_B \frac{1}{\sqrt{\lambda(B \cup C)}} d\lambda = \frac{\lambda(B)}{\sqrt{\lambda(B \cup C)}}$. Analog ist $\mathbf{E}_C^{B \cup C, \varepsilon} = \frac{\lambda(C)}{\sqrt{\lambda(B \cup C)}}$. Dann gilt mit Korollar 3.7.37:

$$\begin{aligned} \|f - f_A - f_{B \cup C}\|_M &< \|f - f_{A \cup B} - f_C\|_M \\ \Leftrightarrow (\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C))^t \left(\mathbf{E}_B^{B \cup C, \varepsilon} \right)^t \mathbf{E}_C^{B \cup C, \varepsilon} (\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C)) \\ &< (\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B))^t \left(\mathbf{E}_A^{A \cup B, \varepsilon} \right)^t \mathbf{E}_B^{A \cup B, \varepsilon} (\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B)) \\ \Leftrightarrow \frac{\lambda(B)\lambda(C)}{\lambda(B \cup C)} (\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C))^2 &< \frac{\lambda(A)\lambda(B)}{\lambda(A \cup B)} (\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B))^2 \\ \Leftrightarrow \frac{\lambda(A \cup B)\lambda(C)}{\lambda(A)\lambda(B \cup C)} (\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C))^2 &< (\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B))^2 \end{aligned}$$

Sei $\tilde{\gamma}_f(A) > \tilde{\gamma}_f(B) > \tilde{\gamma}_f(C)$. Dann ist

$$\begin{aligned}
& \|f - f_A - f_{B \cup C}\|_M < \|f - f_{A \cup B} - f_C\|_M \\
& \Leftrightarrow \sqrt{\frac{\lambda(A \cup B)\lambda(C)}{\lambda(A)\lambda(B \cup C)}}(\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C)) < \tilde{\gamma}_f(A) - \tilde{\gamma}_f(B) \\
& \Leftrightarrow \sqrt{\frac{\lambda(A \cup B)\lambda(C)}{\lambda(A)\lambda(B \cup C)}}\tilde{\gamma}_f(B) + \tilde{\gamma}_f(B) < \sqrt{\frac{\lambda(A \cup B)\lambda(C)}{\lambda(A)\lambda(B \cup C)}}\tilde{\gamma}_f(C) + \tilde{\gamma}_f(A) \\
& \Leftrightarrow \left(\sqrt{\frac{\lambda(A \cup B)\lambda(C)}{\lambda(A)\lambda(B \cup C)}} + 1 \right) \tilde{\gamma}_f(B) < \sqrt{\frac{\lambda(A \cup B)\lambda(C)}{\lambda(A)\lambda(B \cup C)}}\tilde{\gamma}_f(C) + \tilde{\gamma}_f(A) \\
& \Leftrightarrow \tilde{\gamma}_f(B) < \frac{\sqrt{\frac{\lambda(A \cup B)\lambda(C)}{\lambda(A)\lambda(B \cup C)}}\tilde{\gamma}_f(C) + \tilde{\gamma}_f(A)}{\left(\sqrt{\frac{\lambda(A \cup B)\lambda(C)}{\lambda(A)\lambda(B \cup C)}} + 1 \right)}
\end{aligned}$$

Für Mengen $B \in \mathfrak{A}$ mit $\lambda(B)$ hinreichend klein, muß dann $\tilde{\gamma}_f(B) \leq \frac{1}{2}(\tilde{\gamma}_f(A) + \tilde{\gamma}_f(C))$ gelten. Man vergleiche zu diesem Ergebnis [Don91], Corollary 2.4, und [Kus92], S. 23f.

- ii) Sei $E := \mathbb{R}^d$ mit dem Standardskalarprodukt $\langle \cdot | \cdot \rangle$ und $V := \{(\alpha, \dots, \alpha) \mid \alpha \in \mathbb{R}\}$ der Vektorraum der konstanten Funktionen mit Basis $((1, \dots, 1))$. Für $A \in \mathfrak{A}$ mit $\lambda(A) > 0$ ist $e_1^A = \frac{1}{\sqrt{d\lambda(A)}}(1, \dots, 1)$ die ausgezeichnete Orthonormalbasis bezüglich der gegebenen Basis und $\langle \cdot | \cdot \rangle_A$. Es ist dann $\gamma_f(A) = \left(\frac{1}{\sqrt{d\lambda(A)}} \int_A \langle f | (1, \dots, 1) \rangle d\lambda \right)$ (vgl. Beispiel 3.4.12 ii)). Leicht zeigt man, daß $\tilde{\gamma}_f(A) = \left(\frac{1}{d\lambda(A)} \int_A \langle f | (1, \dots, 1) \rangle d\lambda \right)$ ist. Für $\mathbf{E}_B^{B \cup C, \varepsilon}$ ergibt sich damit

$$\begin{aligned}
\mathbf{E}_B^{B \cup C, \varepsilon} &= \int_B \left\langle \frac{1}{\sqrt{d\lambda(B \cup C)}}(1, \dots, 1) \middle| (1, \dots, 1) \right\rangle d\lambda \\
&= \frac{1}{\sqrt{d\lambda(B \cup C)}} \int_B \langle (1, \dots, 1) | (1, \dots, 1) \rangle d\lambda \\
&= \frac{1}{\sqrt{d\lambda(B \cup C)}} \int_B d\lambda = \frac{d\lambda(B)}{\sqrt{d\lambda(B \cup C)}}
\end{aligned}$$

Mit Korollar 3.7.37 gilt dann

$$\begin{aligned}
& \|f - f_A - f_{B \cup C}\|_M < \|f - f_{A \cup B} - f_C\|_M \\
& \Leftrightarrow \frac{d\lambda(B)d\lambda(C)}{d\lambda(B \cup C)}(\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C))^2 < \frac{d\lambda(A)d\lambda(B)}{d\lambda(A \cup B)}(\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B))^2 \\
& \Leftrightarrow \frac{\lambda(A \cup B)\lambda(C)}{\lambda(A)\lambda(B \cup C)}(\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C))^2 < (\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B))^2
\end{aligned}$$

- iii) Sei $E := \mathbb{R}^3$ mit dem Standardskalarprodukt $\langle \cdot | \cdot \rangle$. Es sei $f \in L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^3))$ mit $f = (f_1, f_2, f_3)$. Weiter sei $V := \mathbb{R}^3$ der Vektorraum konstanter Funktionen mit der Standardbasis $\mathcal{E} := ((1, 0, 0), (0, 1, 0), (0, 0, 1))$.

Dann ist mit $e_1^A = \frac{1}{\lambda(A)}(1, 0, 0)$, $e_2^A = \frac{1}{\lambda(A)}(0, 1, 0)$, $e_3^A = \frac{1}{\lambda(A)}(0, 0, 1)$ die Basis (e_1^A, e_2^A, e_3^A) die ausgezeichnete Orthonormalbasis bezüglich des Skalarprodukts $\langle \cdot | \cdot \rangle_A$ und der Basis \mathcal{E} (vgl. Beispiel 3.4.12 iii)).

Dann ist für $A \in \mathfrak{A}$ mit $\lambda(A) > 0$ der Vektor $\tilde{\gamma}_f(A)$ gegeben durch $\tilde{\gamma}_f(A)_1 = \frac{1}{\lambda(A)} \int_A f_1 d\lambda$, $\tilde{\gamma}_f(A)_2 = \frac{1}{\lambda(A)} \int_A f_2 d\lambda$, $\tilde{\gamma}_f(A)_3 = \frac{1}{\lambda(A)} \int_A f_3 d\lambda$ und damit $\tilde{\gamma}_f(A) = \frac{1}{\lambda(A)} (\int_A f_1 d\lambda, \int_A f_2 d\lambda, \int_A f_3 d\lambda)$.

Für die Matrix $\mathbf{E}_B^{B \cup C, \mathcal{E}} \in \mathbb{R}^{3 \times 3}$ ergibt sich

$$\begin{aligned} \mathbf{E}_B^{B \cup C, \mathcal{E}} &= \begin{pmatrix} \langle e_1^{B \cup C} | e_1 \rangle_B & \langle e_1^{B \cup C} | e_2 \rangle_B & \langle e_1^{B \cup C} | e_3 \rangle_B \\ \langle e_2^{B \cup C} | e_1 \rangle_B & \langle e_2^{B \cup C} | e_2 \rangle_B & \langle e_2^{B \cup C} | e_3 \rangle_B \\ \langle e_3^{B \cup C} | e_1 \rangle_B & \langle e_3^{B \cup C} | e_2 \rangle_B & \langle e_3^{B \cup C} | e_3 \rangle_B \end{pmatrix} \\ &= \begin{pmatrix} \frac{\lambda(B)}{\sqrt{\lambda(B \cup C)}} & 0 & 0 \\ 0 & \frac{\lambda(B)}{\sqrt{\lambda(B \cup C)}} & 0 \\ 0 & 0 & \frac{\lambda(B)}{\sqrt{\lambda(B \cup C)}} \end{pmatrix} \end{aligned}$$

Mit Korollar 3.7.37 ergibt sich dann

$$\begin{aligned} \|f - f_A - f_{B \cup C}\|_M &< \|f - f_{A \cup B} - f_C\|_M \\ \Leftrightarrow (\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C))^t (\mathbf{E}_B^{B \cup C, \mathcal{E}})^t \mathbf{E}_C^{B \cup C, \mathcal{E}} (\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C)) \\ &< (\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B))^t (\mathbf{E}_A^{A \cup B, \mathcal{E}})^t \mathbf{E}_B^{A \cup B, \mathcal{E}} (\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B)) \\ \Leftrightarrow \sum_{i=1}^3 \frac{\lambda(B)\lambda(C)}{\lambda(B \cup C)} (\tilde{\gamma}_f(B)_i - \tilde{\gamma}_f(C)_i)^2 &< \sum_{i=1}^3 \frac{\lambda(A)\lambda(B)}{\lambda(A \cup B)} (\tilde{\gamma}_f(A)_i - \tilde{\gamma}_f(B)_i)^2 \\ \Leftrightarrow \frac{\lambda(A \cup B)\lambda(C)}{\lambda(A)\lambda(B \cup C)} \|\tilde{\gamma}_f(B) - \tilde{\gamma}_f(C)\|_2^2 &< \|\tilde{\gamma}_f(A) - \tilde{\gamma}_f(B)\|_2^2 \end{aligned}$$

- iv) Sei $M := [0, 1]$, $E := \mathbb{R}$ mit der Multiplikation als Skalarprodukt und $\mathfrak{A} := \{[t_1, t_2] \mid t_1, t_2 \in [0, 1]\}$. $V := \{f \mid f(x) = ax + t \text{ mit } a, t \in \mathbb{R}\}$, dann ist (e_1, e_2) mit $e_1(x) := 1$ und $e_2(x) := x$ eine Basis von V .

Dann ist für $t \in]0, 1[$ die Menge $[0, t] \in \mathfrak{A}$. Es ist dann $e_1^{[0, t]}(x) = \frac{1}{\sqrt{t}}$ und $e_2^{[0, t]}(x) = 2\sqrt{\frac{3}{t^3}}(x - \frac{1}{2}t)$. Ebenso ist die Menge $[t, 1] \in \mathfrak{A}$ und es gilt $e_1^{[t, 1]} = \frac{1}{\sqrt{1-t}}$ und $e_2^{[t, 1]}(x) = \sqrt{\frac{3}{(t-1)^3}}(2x - t - 1)$ (vgl. Beispiel 3.4.12 iv)).

Sei für $0 < t_1 < t_2 < 1$ die Menge $A := [0, t_1]$, $B := [t_1, t_2]$ und $C := [t_2, 1]$. Dann

ergibt sich

$$\begin{aligned}
\langle e_1^{A \cup B} | e_1 \rangle_A &= \langle e_1^{[0, t_2]} | 1 \rangle_{[0, t_1]} = \frac{t_1}{\sqrt{t_2}} \\
\langle e_1^{A \cup B} | e_2 \rangle_A &= \langle e_1^{[0, t_2]} | \text{id} \rangle_{[0, t_1]} = \frac{1}{\sqrt{t_2}} \int_{[0, t_1]} x d\lambda(x) \\
&= \frac{t_1^2}{2\sqrt{t_2}} \\
\langle e_2^{A \cup B} | e_1 \rangle_A &= \langle e_2^{[0, t_2]} | 1 \rangle_{[0, t_1]} = 2\sqrt{\frac{3}{t_2^3}} \int_{[0, t_1]} x - \frac{1}{2}t_2 d\lambda(x) \\
&= 2\sqrt{\frac{3}{t_2^3}} \left(\frac{1}{2}t_1^2 - t_1 \frac{1}{2}t_2 \right) \\
&= 2\sqrt{\frac{3}{t_2^3}} \frac{1}{2}t_1(t_1 - t_2) \\
&= \sqrt{\frac{3}{t_2^3}} t_1(t_1 - t_2)
\end{aligned}$$

und

$$\begin{aligned}
\langle e_2^{A \cup B} | e_2 \rangle_A &= \langle e_2^{[0, t_2]} | \text{id} \rangle_{[0, t_1]} = 2\sqrt{\frac{3}{t_2^3}} \int_{[0, t_1]} \left(x - \frac{t_2}{2} \right) x d\lambda(x) \\
&= 2\sqrt{\frac{3}{t_2^3}} \int_{[0, t_1]} x^2 - \frac{t_2}{2} x d\lambda(x) \\
&= 2\sqrt{\frac{3}{t_2^3}} \left(\frac{1}{3}t_1^3 - \frac{t_2 t_1^2}{4} \right) \\
&= \frac{\sqrt{3}}{6\sqrt{t_2^3}} t_1^2 (4t_1 - 3t_2)
\end{aligned}$$

Damit ergibt sich als Matrix $\mathbf{E}_A^{A \cup B, \mathcal{E}}$:

$$\mathbf{E}_{[0, t_1]}^{[0, t_2], \mathcal{E}} = \begin{pmatrix} \frac{t_1}{\sqrt{t_2}} & \frac{t_1^2}{2\sqrt{t_2}} \\ \sqrt{\frac{3}{t_2^3}} t_1(t_1 - t_2) & \frac{\sqrt{3}}{6\sqrt{t_2^3}} t_1^2 (4t_1 - 3t_2) \end{pmatrix}$$

Analog zeigt man:

$$\mathbf{E}_B^{A \cup B, \mathcal{E}} = \mathbf{E}_{[t_1, t_2]}^{[0, t_2], \mathcal{E}} = \begin{pmatrix} \frac{t_2 - t_1}{\sqrt{t_2}} & \frac{t_2^2 - t_1^2}{2\sqrt{t_2}} \\ \sqrt{\frac{3}{t_2^3}} t_1(t_2 - t_1) & \frac{\sqrt{3}}{6\sqrt{t_2^3}} (t_2^3 + 3t_2 t_1 - 4t_1^3) \end{pmatrix}$$

und

$$\mathbf{E}_B^{B \cup C, \mathcal{E}} = \mathbf{E}_{[t_1, t_2]}^{[t_1, 1], \mathcal{E}} = \begin{pmatrix} \frac{t_2 - t_1}{\sqrt{1 - t_1}} & \frac{1}{2\sqrt{1 - t_1}} t_2^2 - t_1^2 \\ \sqrt{\frac{3}{(1 - t_1)^3}} (t_2^2 + t_1 t_2 - t_2 + t_1) & \frac{\sqrt{3}}{6\sqrt{(1 - t_1)^3}} (4t_2^3 - t_1^3 - 3t_1 t_2^2 + 3t_1^2 - 3t_2^2) \end{pmatrix}$$

sowie

$$\mathbf{E}_C^{B \cup C, \varepsilon} = \mathbf{E}_{\begin{smallmatrix} [t_1, 1], \varepsilon \\ [t_2, 1] \end{smallmatrix}} = \begin{pmatrix} \frac{1-t_2}{\sqrt{1-t_1}} & \frac{1-t_2^2}{2\sqrt{1-t_1}} \\ \sqrt{\frac{3}{(1-t_1)^3}}(t_1 t_2 + t_2 - t_1 - t_2^2) & \frac{\sqrt{3}}{6\sqrt{(1-t_1)^3}}(1 - 4t_2^3 - 3t_1 + 3t_1 t_2^2 + 3t_2^2) \end{pmatrix}$$

Bemerkung 3.7.39

Mit Satz 3.7.35 und Lemma 3.7.29 ist die Matrix $(\mathbf{E}_A^{A \cup B, \varepsilon})^t \mathbf{E}_B^{A \cup B, \varepsilon}$ symmetrisch und invertierbar, und nach Korollar 3.7.36 ist sie positiv definit. Daher ist die Abbildung

$$s : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \\ (x, y) \mapsto x^t \left(\mathbf{E}_B^{B \cup C, \varepsilon} \right)^t \mathbf{E}_C^{B \cup C, \varepsilon} y$$

ein Skalarprodukt (vgl. z. B. [Fis86], 6.1.8. Bemerkung, S. 186) und die Abbildung

$$d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+ \\ (x, y) \mapsto \sqrt{s(x-y, x-y)}$$

definiert eine Metrik auf dem \mathbb{R}^n .

Daher ist die folgende Definition 3.7.40 wohldefiniert.

Definition 3.7.40 (Metrik des Segments A bezüglich des Segments B)

Für zwei Mengen $A, B \in \mathfrak{A}$ mit $\lambda(A) > 0$, $\lambda(B) > 0$ und $\lambda(A \cap B) = 0$ heie die Metrik

$$d_{A,B} : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+ \\ (x, y) \mapsto \sqrt{s(x-y, x-y)}$$

die Metrik des Segments A bezüglich des Segments B. Dabei sei s das Skalarprodukt aus der vorangegangenen Bemerkung 3.7.39.

Bemerkung 3.7.41

Die Flle i), ii) und iii) des Beispiel 3.7.38 zeigen, da bei Approximation mit konstanten Funktionen auf den Segmenten $d_{A,B}$ lediglich von den Groen der Mengen A und B im Lebesgue-Ma abhngt. Diese Tatsache kann algorithmisch ausgenutzt werden, wie man in Abschnitt 4.6.6.4 des Kapitels 4 auf Seite 206 sehen kann.

Korollar 3.7.42

Seien $A, B_1, C_1 \in \mathfrak{A}$ sowie $A, B_2, C_2 \in \mathfrak{A}$ Zerlegungen von M . Sei weiter $\Psi : \mathfrak{A} \rightarrow \mathbb{R}_+$ mit

$$\Psi(A) = \|\gamma_f(A)\|_2^2 + \|\gamma_f(M \setminus A)\|_2^2.$$

die Zielfunktion einer 2-Niveaupassung. Dann gilt:

$$\Psi(A \cup B_1) > \Psi(A \cup B_2) \Leftrightarrow d_{A, B_1}(\tilde{\gamma}_f(A), \tilde{\gamma}_f(B_1)) < d_{A, B_2}(\tilde{\gamma}_f(A), \tilde{\gamma}_f(B_2))$$

Beweis:

Mit Satz 3.5.16 gilt $\Psi(A \cup B_1) > \Psi(A \cup B_2) \Leftrightarrow \|f - f_{A \cup B_1} - f_{C_1}\|_M^2 < \|f - f_{A \cup B_2} - f_{C_2}\|_M^2$. Daher folgt die Behauptung mit Satz 3.7.35 und Bemerkung 3.7.39.

Bemerkung 3.7.43

Das Korollar 3.7.42 kann als Grundlage zur algorithmischen Suche nach lokalen Maxima der Zielfunktion der 2-Niveaupassung dienen. Wie bei vielen nichtlinearen Optimierungsalgorithmen, hat sich auch in den Anwendungen der vektoriellen Mehrniveaupassung die Verbesserung in die Richtung des stärksten Anstiegs bzw. Wertzuwachses als zufriedenstellende Strategie zur algorithmischen Bestimmung signifikanter lokaler Maxima herausgestellt. Hat ein Algorithmus ausgehend von einer Segmentierung $\{A, M \setminus A\}$ nun mehrere Mengen $B \subset M \setminus A$ zur Wahl, um sie A hinzuzufügen, so kann man mit Hilfe des Korollars die Menge B wählen, die den stärksten Wertzuwachs verursacht. Dabei muß die Menge $M \setminus (A \cup B)$ nicht betrachtet werden.

Erst in Abschnitt 4.6 wird sich jedoch die Bedeutung des Korollars für die Algorithmen der vektoriellen Mehrniveaupassung zeigen.

Kapitel 4

Algorithmen und Anwendungen der Mehrniveaupassung im Bereich der Bildsegmentierung

4.1 Einleitung

Obwohl die theoretischen Grundlagen der vektoriellen Mehrniveaupassung auch Anwendungen über den Bereich der Bildsegmentierung hinaus erlauben, soll in dieser Arbeit in erster Linie auf die vielfältigen Möglichkeiten, die die Mehrniveaupassung zur Segmentierung von Bildern bietet, eingegangen werden. Es ist selbstverständlich, daß dabei der Aspekt der algorithmischen Umsetzung der vorgestellten Verfahren im Vordergrund stehen muß. Gerade diese Umsetzung von theoretischen Resultaten – gewonnen bei der Betrachtung von kontinuierlichen Bildern – bereitet dem Ansatz, Bilder über Energieminimierung zu segmentieren, erhebliche Probleme, die letztlich algorithmisch nur über die Methode der finiten Elemente lösbar sind (vgl. Abschnitt 2.4.2).

Die Approximationsprobleme, die die vektorielle Mehrniveaupassung auf den Segmenten zu lösen hat, sind nach Bemerkung 3.5.4 (S. 94) orthogonale Projektionen auf endlich dimensionale Vektorräume. Da die Menge der diskreten Pixelbilder mit festem Definitionsbereich M ebenfalls einen endlich dimensionalen Vektorraum definieren, sind die Approximationsprobleme auf Pixelmengen algorithmisch lösbar. Eine Annäherung der optimalen approximierenden Funktion auf einem Segment über die Methode der finiten Elemente (d. h. durch ein Polygon), wie es in Abschnitt 2.4.2.1 beim Mumford-Shah-Modell gemacht wird, ist an dieser Stelle also unnötig.

Auch die Behandlung der Menge der Unstetigkeitsstellen auf einem zweidimensionalen Definitionsbereich bereitet den Ansätzen der Energieminimierung große Schwierigkeiten (vgl. Abschnitt 2.4.5, S. 47). Hier zeigt sich ein nicht zu unterschätzender Aspekt des Ansatzes der vektoriellen Mehrniveaupassung: Das Konzept des Systems der erlaubten Segmentierung, das im Abschnitt 3.3 eingeführt wurde, ermöglicht eine einfache Übertragung auf diskrete Pixelbilder, indem einfach unter allen möglichen Segmentierungen nur diejenigen betrachtet werden, die sich auch als Vereinigungen von Pixeln darstellen lassen.

Natürlich ist der Übergang von infiniten zu finiten Mengensystemen auch bei der vektoriellen Mehrniveaupassung nicht ganz unproblematisch. So wird sich insbesondere herausstellen, daß ein Maximum der Zielfunktion einer k -Niveaupassung über einem diskreten Pixelbild nicht notwendig selbst pixelbasiert sein muß.

Übersicht des Kapitels

Im Gegensatz zum vorherigen Kapitel sind die im folgenden betrachteten Systeme von Segmentierungen endlich. Bevor die einzelnen Anwendungsbeispiele vorgestellt werden, zeigt das zweite Unterkapitel zunächst einmal das grundsätzliche Vorgehen und die Schwierigkeiten des Übergang von infiniten zu finiten Mengensystemen auf. Im Mittelpunkt dabei steht der Begriff der „pixelbasierten Segmentierung“ und des „pixelbasierten lokalen Maximums“. Alle später vorgestellten Algorithmen arbeiten mit diesen Begriffen.

Im dritten Unterkapitel werden zunächst einmal Algorithmen für beliebige pixelbasierte Mengensysteme präsentiert. Dabei wird sich für pixelbasierte lokale Maxima zeigen, daß man unter bestimmten Voraussetzungen das Problem, ein pixelbasiertes lokales Maximum einer k -Niveaupassung zu finden, auf iterative 2-Niveaupassungen zurückführen kann, wie es bei allgemeinen lokalen Maxima in Abschnitt 3.7.1 gezeigt wurde.

In den nachfolgenden Unterkapiteln werden spezielle Mengensysteme und konkrete Instanzen der abstrakten Algorithmen für diese speziellen Mengensysteme vorgestellt. Als erste einfache Anwendung der vektoriellen Mehrniveaupassung zur Bildsegmentierung wird im vierten Unterkapitel die Segmentierung von diskreten Pixelbildern mit eindimensionalem Definitionsbereich durch Intervalle betrachtet.

Im fünften Unterkapitel wird das Mengenmodell der „grauwertbenachbarten“ Pixelmengen definiert. Es wird gezeigt, daß eine Mehrniveaupassung bezüglich dieses Mengenmodells nichts anderes liefert als das Schwellwertverfahren von Otsu, das in Abschnitt 2.3.1.1 vorgestellt wurde.

Im sechsten Unterkapitel wird das Mengenmodell der pixelzusammenhängenden Mengen betrachtet. Dieses Mengenmodell ist wohl das naheliegendste Mengensystem zur Segmentierung von Bildern mit zweidimensionalem Definitionsbereich. Leider erfüllt jedoch das infinite Mengensystem der zusammenhängenden Mengen, dessen Diskretisierung das Mengensystem der pixelzusammenhängenden Mengen ergeben sollte, nicht die Voraussetzungen eines Systems erlaubter Mengen (vgl. Definition 3.2.10, S. 74). Der Algorithmus zur Bestimmung eines pixelbasierten Maximums bezüglich dieses Mengenmodells bleibt jedoch gültig.

4.2 Von infiniten zu finiten Systemen von Mengen

Die Definition des Systems der Segmentierungen in Definition 3.3.1 (S. 76) geht von einem Mengensystem mit unendlich vielen Elementen aus, was sich direkt aus der Dichte-eigenschaft (DE) der Definition 3.2.10 der erlaubten Mengensysteme ableiten läßt. Dem gegenüber steht bei diskreten Pixelbildern eine endliche Anzahl von Pixeln und daher eine endliche Anzahl von pixelbasierten Mengen, d. h. Mengen, die Vereinigungen von Pixeln sind. Prinzipiell gibt es zwei Möglichkeiten dieses Problem zu lösen:

- a) Einbettung pixelbasierter Mengen innerhalb eines infiniten Mengensystems
Verfolgt man diesen Ansatz, so sind Optima von Funktionen wie zum Beispiel in Beispiel 3.5.11 iv) mit algebraischen oder numerischen Methoden zu bestimmen. Ein Beispiel für einen solchen Ansatz ist auch die Funktion in Beispiel 4.2.4. Sicher würde es für viele Fragestellungen nichtlineare Optimierungsalgorithmen geben, die gegen ein signifikantes Optimum konvergieren. Allerdings sind solche Verfahren bei der Segmentierung von Bildern in ihrer Laufzeit oft zu lang.
- b) Einschränkung auf pixelbasierte Mengen innerhalb eines infiniten Mengensystems
Dieser Ansatz, bei einem infiniten Mengensystem nur endlich viele Elemente des Systems zu betrachten, erleichtert die zu entwickelnde Algorithmik. Allerdings ist hier bei der Definition des Begriffes „Optimum“ Vorsicht geboten, wie Beispiel 4.2.4 und Bemerkung 4.2.5 zeigen.

Da sich diese Arbeit in erster Linie mit den Anwendungen der vektoriellen Mehrniveaupassung in der Bildsegmentierung auseinandersetzt, soll hier der Weg b) beschrrieben werden. Der Weg, der in a) beschrieben wird, ist jedoch keinesfalls uninteressant. Selbst im Bereich der Segmentierung von Bildern hat er durchaus seine Berechtigung, etwa bei subpixelgenauen Segmentierungen. Allerdings ist für diesen Ansatz eine Parametrisierung der Segmentierungen wie z. B. in Beispiel 4.2.4 nötig. Für weitere Anwendungsbeispiele sei auf [Lau03] verwiesen.

4.2.1 Pixelbasierte Segmentierungen

Ziel dieses Abschnitts ist es, den Begriff der „pixelbasierten Segmentierung“, d. h. der Diskretisierung des Segmentierungsbegriffs aus Abschnitt 3.3 einzuführen. Die Menge der pixelbasierten Segmentierungen ist eine Untermenge des Systems der Segmentierungen durch erlaubte Mengen. Es kann allerdings vorkommen, daß selbst bei diskreten Pixelbildern das globale Maximum der Zielfunktion einer k -Niveaupassung nicht auf einer pixelbasierten Segmentierung angenommen wird.

Definition 4.2.1 (Pixelbasierte Segmentierungen)

Sei \mathfrak{A} ein System erlaubter Mengen und $A \in \mathfrak{A}$. Dann heiße die Menge $S_k(\mathfrak{A}, A)$ definiert durch

$$S_k(\mathfrak{A}, A) := \mathfrak{S}_k(\mathfrak{A} \cap \mathcal{P}(M), A)$$

die Menge der pixelbasierten Segmentierungen von A in $\mathfrak{S}_k(\mathfrak{A})$. Ist $A = M$, so werde $S_k(\mathfrak{A}) := S_k(\mathfrak{A}, M)$ notiert. Dies ist, bis auf Abschnitt 4.3, im folgenden immer der Fall.

Bemerkung 4.2.2

Vergleicht man die Definitionen von \mathfrak{A} und $\mathcal{P}(M)$ so kann – streng mathematisch – für kein Mengensystem \mathfrak{A} der Schnitt $\mathfrak{A} \cap \mathcal{P}(M) \neq \emptyset$ sein. Denn \mathfrak{A} besteht aus Äquivalenzklassen (vgl. Definition 3.2.5, S. 71) und ist somit von einem anderen Typ als $\mathcal{P}(M)$. Allerdings kann man die Elemente aus $\mathcal{P}(M)$ auch als Repräsentanten ihrer Äquivalenzklasse sehen, denn ein Übergang $\mathcal{P}(M) \rightarrow \mathcal{P}(M)/\sim$, wobei \sim wie in Definition 3.2.1 (S. 70) definiert sei, ist ein einfacher Isomorphismus, der jedes Element auf die Äquivalenzklasse abbildet, die nur aus diesem Element besteht. Daher ist die Notation $\mathfrak{A} \cap \mathcal{P}(M)$ durchaus legitim. Im folgenden sei also $\mathfrak{A} \cap \mathcal{P}(M)$ nur die legere Notation der Menge $\mathfrak{A} \cap \mathcal{P}(M)/\sim$.

Bemerkung 4.2.3

Es kann vorkommen, daß das Mengensystem \mathfrak{A} keine pixelbasierten Mengen enthält (etwa bei Kreisen). Mit den Fragestellungen, die sich aus einem solchen Mengensystem ergeben, beschäftigt sich die Arbeit von Verena Lauren ([Lau03]). Deshalb soll an dieser Stelle nicht weiter auf solche Mengensysteme eingegangen werden. Aber auch Mengensysteme, die pixelbasierte Mengen enthalten, können zu optimalen Segmentierungen führen, die nicht pixelbasiert sind, wie das folgende Beispiel zeigt.

Beispiel 4.2.4

Es sei $M := [0, 7]$ und $E := \mathbb{R}$ mit der Multiplikation als Skalarprodukt. \mathfrak{A} sei das Mengensystem der zusammenhängenden Mengen in M . Die Familie der Skalarprodukte sei gegeben wie in Beispiel 3.4.6.

Es ist $\mathfrak{S}_2(\mathfrak{A}) = \{\{[0, t], [t, 7]\} \mid t \in [0, 7]\}$ das System der Segmentierungen von M durch zwei zusammenhängende Mengen. Die Klasse der erlaubten Funktionen sei der Vektorraum, der durch die Funktionen e_1, e_2 mit $e_1(x) := 1$ und $e_2(x) := x$ aufgespannt wird.

Gegeben sei gemäß Definition 1.3.2 ein diskretes Pixelbild $g : M \rightarrow \mathbb{R}$ mit $g([0]) := 5, g([1]) := 3, g([2]) := 2, g([3]) := 1, g([4]) := 2, g([5]) := 3, g([6]) := 5$. Abbildung 4.1 zeigt den Graphen dieser Funktion.

Betrachtet wird nun die Zielfunktion der 2-Niveaupassung

$$\Psi : \begin{array}{l} [0, 7] \rightarrow \mathbb{R}_+ \\ t \mapsto \|\gamma_f([0, t])\|_2^2 + \|\gamma_f([t, 7])\|_2^2 \end{array}$$

Wie man sieht, läßt sich in diesem Beispiel die Zielfunktion leicht über den Parameter t parametrisieren, so daß eine analytische Untersuchung der Zielfunktion möglich ist. Es kann gezeigt werden, daß die Zielfunktion ihr Maximum bei $t = 3 + \frac{1}{2}$ annimmt. Die Abbildung 4.2 zeigt den Graphen der Zielfunktion. Abbildung 4.3 zeigt die optimale Segmentierung des diskreten Pixelbildes g durch zwei zusammenhängende Segmente. Man beachte zudem, daß die Zielfunktion Ψ bei $t = 1$ und $t = 6$ lokale Maxima annimmt.

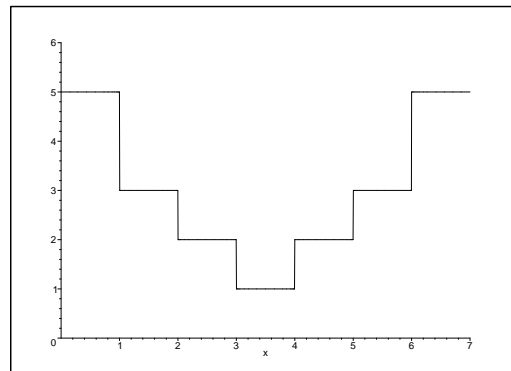


Abbildung 4.1: Diskretes eindimensionales Pixelbild aus Beispiel 4.2.4.

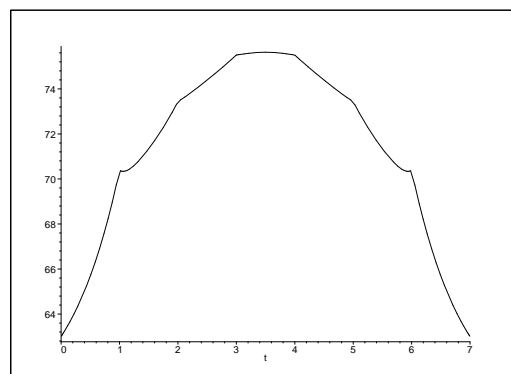
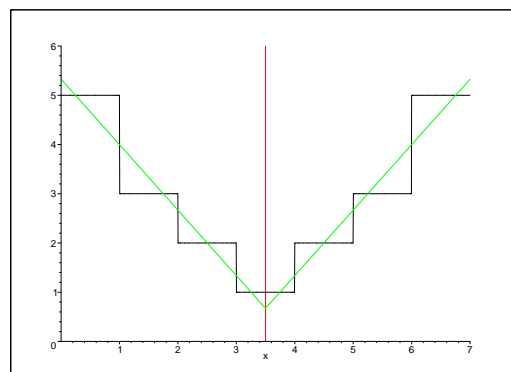
Abbildung 4.2: Graph der Zielfunktion Ψ aus Beispiel 4.2.4.

Abbildung 4.3: Optimales Ergebnis der 2-Niveaupassung aus Beispiel 4.2.4: Segmentweise Approximation mit linearen Funktionen (grün). Die Segmentgrenzen sind rot eingezeichnet.

Bemerkung 4.2.5

Das Beispiel 4.2.4 zeigt, daß eine optimale Segmentierung von diskreten Pixelbildern nicht pixelbasiert sein muß. Es verdeutlicht zwei Probleme:

1. Das absolute Maximum der Zielfunktion der k -Niveaupassung muß nicht pixelbasiert sein.
2. Unter Umständen existieren zwar lokale Maxima die pixelbasiert sind, aber bei einer Auswertung der Zielfunktion nur an Pixelgrenzen keine „pixelbasierten“ lokalen Maxima sind. (Man vergleiche dazu die Definition 4.2.15 dieses Begriffes, die in Abschnitt 4.2.3 vorgestellt wird.)

Würde die Zielfunktion nur an Pixelgrenzen ausgewertet, was bei den später vorgestellten Algorithmen immer der Fall ist, so würde man die Werte 3 und 4 als Maximalstellen erhalten. Es ist daher eine diskrete Version des Begriffes „lokales Maximum“ notwendig, die die Werte 3 und 4 als Stellen, an denen das diskrete lokale Maximum angenommen wird, liefert.

4.2.2 Exkurs: Approximation stückweiser konstanter Funktionen mit konstanten Funktionen

Für eine kleine Untermenge der Mehrniveaupassungsprobleme reicht es aus, nur die pixelbasierten Segmentierungen zu betrachten, um ein globales Optimum der Zielfunktion einer k -Niveaupassung zu bestimmen. Sind alle Funktionen des Vektorraums V konstant, so kann man sich, unter bestimmten Voraussetzungen, auf die Menge der pixelbasierten Segmentierungen $S_k(\mathfrak{A})$ beschränken, wenn man eine optimale Segmentierung aus $\mathfrak{S}_k(\mathfrak{A})$ sucht.

In diesem Abschnitt werden nur Funktionen aus $L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$ betrachtet. Als Klasse der erlaubten Funktionen wird nur der Vektorraum der komponentenweise konstanten Funktionen zugelassen.

Definition 4.2.6

Für $E := \mathbb{R}^d$ mit dem Standardskalarprodukt sei V_c der Vektorraum der komponentenweise konstanten Funktionen. Das heißt, es ist

$$V_c := \{(c_1, \dots, c_d) \mid c_1, \dots, c_d \in \mathbb{R}\}$$

Bemerkung 4.2.7

Für V_c gilt:

- i. $e_1 := (1, 0, \dots, 0), \dots, e_d := (0, \dots, 0, 1)$ ist eine Basis von V_c .
- ii. $\forall A \in \sigma(A), \lambda(A) > 0 : \dim V_c|_A = d$.

Man beachte, daß in diesem Kontext für alle $A \in \sigma(\mathfrak{A}) \setminus \{O\}$ die Funktion $f_A|$ auf dem Segment A konstant ist.

Lemma 4.2.8

Sei $\{\langle \cdot | \cdot \rangle_A \mid A \in \sigma(\mathfrak{A}), \lambda(A) > 0\}$ eine Familie erlaubter Skalarprodukte definiert wie in Beispiel 3.4.6 (S. 86). V_c sei definiert wie in Definition 4.2.6. Außerdem sei $A \uplus B_1, B_2 \uplus C$ eine Zerlegung von M durch Mengen aus $\sigma(\mathfrak{A})$. Ebenso seien die Mengen A und C aus $\sigma(\mathfrak{A})$. Es sei $B := B_1 \cup B_2$.

Gilt dann $f|_{B_1} = f|_{B_2} = c$ für eine Konstante $c \in \mathbb{R}^d$, so folgt:

$$\begin{aligned} \text{Entweder es gilt } \|f - f_A - f_{B \cup C}\|_M &\leq \|f - f_{A \cup B_1} - f_{B_2 \cup C}\|_M \\ \text{oder es gilt } \|f - f_{A \cup B} - f_C\|_M &\leq \|f - f_{A \cup B_1} - f_{B_2 \cup C}\|_M \end{aligned}$$

Beweis:

Zunächst einmal gilt $B_1 \in \sigma(\mathfrak{A})$ und $B_2 \in \sigma(\mathfrak{A})$ und daher auch $B \in \sigma(\mathfrak{A})$ (vgl. [Bau92], S. 3f).

Nun ist

$$\begin{aligned} &\|f - f_{A \cup B_1} - f_{B_2 \cup C}\|_M^2 \\ &= \|f - f_{A \cup B_1}\|_A^2 + \|f - f_{A \cup B_1}\|_{B_1}^2 + \|f - f_{B_2 \cup C}\|_{B_2}^2 + \|f - f_{B_2 \cup C}\|_C^2 \\ &= \|f - f_{A \cup B_1}\|_A^2 + \|c - f_{A \cup B_1}\|_{B_1}^2 + \|c - f_{B_2 \cup C}\|_{B_2}^2 + \|f - f_{B_2 \cup C}\|_C^2 \\ &= \int_A \|f - f_{A \cup B_1}\|_2^2 d\lambda \\ &\quad + \int_{B_1} \|c - f_{A \cup B_1}\|_2^2 d\lambda + \int_{B_2} \|c - f_{B_2 \cup C}\|_2^2 d\lambda \\ &\quad + \int_C \|f - f_{B_2 \cup C}\|_2^2 d\lambda \end{aligned}$$

Nun kann man zwei Fälle unterscheiden: Entweder gilt $\|c - f_{A \cup B_1}\|_2 \leq \|c - f_{B_2 \cup C}\|_2$ oder es ist $\|c - f_{A \cup B_1}\|_2 > \|c - f_{B_2 \cup C}\|_2$. Betrachtet wird zunächst der erste Fall:

$$\|c - f_{A \cup B_1}\|_2 \leq \|c - f_{B_2 \cup C}\|_2$$

Dann gilt aber:

$$\begin{aligned} & \int_A \|f - f_{A \cup B_1}\|_2^2 d\lambda + \int_{B_1} \|c - f_{A \cup B_1}\|_2^2 d\lambda \\ & + \int_{B_2} \|c - f_{B_2 \cup C}\|_2^2 d\lambda + \int_C \|f - f_{B_2 \cup C}\|_2^2 d\lambda \\ & \geq \int_A \|f - f_{A \cup B_1}\|_2^2 d\lambda \\ & \quad + \int_{B_1} \|c - f_{A \cup B_1}\|_2^2 d\lambda + \int_{B_2} \|c - f_{A \cup B_1}\|_2^2 d\lambda \\ & \quad + \int_C \|f - f_{B_2 \cup C}\|_2^2 d\lambda \\ & \geq \int_{A \cup B_1 \cup B_2} \|f - f_{A \cup B_1}\|_2^2 d\lambda + \int_C \|f - f_{B_2 \cup C}\|_2^2 d\lambda \\ & \geq \int_{A \cup B} \|f - f_{A \cup B}\|_2^2 d\lambda + \int_C \|f - f_C\|_2^2 d\lambda \\ & = \|f - f_{A \cup B} - f_C\|_M^2 \end{aligned}$$

Dabei folgt die letzte Ungleichung mit Satz 3.5.3 zusammen mit Bemerkung 3.5.1. Daher gilt in diesem Fall: $\|f - f_{A \cup B_1} - f_{B_2 \cup C}\|_M \geq \|f - f_{A \cup B} - f_C\|_M$. Vollkommen analog kann man für den Fall $\|c - f_{A \cup B_1}\|_2 > \|c - f_{B_2 \cup C}\|_2$ auf $\|f - f_{A \cup B_1} - f_{B_2 \cup C}\|_M > \|f - f_A - f_{B \cup C}\|_M$ schließen. Damit folgt die Behauptung. \square

Satz 4.2.9

Sei $M := [0, m]$ mit $m \in \mathbb{N}$ und \mathfrak{A} die Menge der Intervalle in M , $E := \mathbb{R}^d$ mit dem Standardskalarprodukt. Weiter sei die Familie erlaubter Skalarprodukte wie in Beispiel 3.4.6. Außerdem sei V_c definiert wie in Definition 4.2.6 und $g : M \rightarrow \mathbb{R}^d$ ein diskretes Pixelbild und Ψ die Zielfunktion der k -Niveaupassung von g . Dann gilt:

Für alle $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ existiert eine pixelbasierte Segmentierung $\mathbf{A}' \in S_k(\mathfrak{A})$ mit $\Psi(\mathbf{A}') \geq \Psi(\mathbf{A})$.

Bei der Bestimmung lokaler Maxima von k -Niveaupassungen eindimensionaler diskreter Pixelbilder mit konstanter Approximation auf den Segmenten genügt es also, sich auf pixelbasierte Segmentierungen zurückzuziehen.

Beweis:

Sei $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ und $[x_1, x_2] \in \mathbf{A}$ ein Segment mit $x_2 \notin \mathbb{N}$. Weiter sei $p_{x_2} \in \mathbb{P}(M)$ das Pixel mit $\{x_2\} \cap p_{x_2} = \{x_2\}$. Dann ist sowohl $[x_1, x_2] \setminus p_{x_2} \in \mathfrak{A}$ als auch $[x_1, x_2] \cup p_{x_2} \in \mathfrak{A}$. Nach Voraussetzung ist g auf p_{x_2} konstant. Daher existiert nach Lemma 4.2.8 eine Segmentierung

$\mathbf{A}' \in \mathfrak{S}_k(\mathfrak{A})$ mit $[x_1, x_2] \setminus p_{x_2} \in \mathbf{A}'$ oder $[x_1, x_2] \cup p_{x_2} \in \mathbf{A}'$, die der Funktion Ψ einen höheren oder gleichen Wert zuteilt.

Auf diese Art lassen sich alle Segmente, die nicht pixelbasiert sind, in pixelbasierte Segmente umwandeln, ohne daß der Wert der Funktion Ψ der Segmentierung \mathbf{A} sinkt. Damit folgt die Behauptung. \square

Satz 4.2.10

Sei $\mathfrak{E} \subset \mathcal{B}(M) / \sim$ ein System meßbarer Mengen. Weiter sei \mathfrak{A} sei ein Mengensystem erlaubter Mengen gemäß Definition 3.2.10, so daß für alle Mengen $A \in \mathfrak{A}$ und $B \in \mathfrak{E}$ mit $A \cap B \neq \emptyset$ gilt: $A \cup B \in \mathfrak{A}$ und $A \setminus B \in \mathfrak{A}$.

Sei $\{\langle \cdot | \cdot \rangle_A \mid A \in \sigma(\mathfrak{A} \cup \mathfrak{E}), \lambda(A) > 0\}$, eine Familie der erlaubten Skalarprodukte, definiert wie in Beispiel 3.4.6. V_c sei definiert wie in Definition 4.2.6. Existiert dann eine disjunkte Zerlegung $\bigsqcup_{i=1}^l B_i$ von M durch Mengen aus \mathfrak{E} , so daß f auf den einzelnen Segmenten B_i der Zerlegung konstant ist ($1 \leq i \leq l$), so gilt:

Es existiert eine global optimale Segmentierung $\{A_1, \dots, A_k\} \in \mathfrak{S}_k(\mathfrak{A})$ von f mit

$$\forall i \in \{1, \dots, k\} : \exists i_1, \dots, i_n \in \{1, \dots, l\} : A_i = \bigcup_{j \in \{i_1, \dots, i_n\}} B_j \quad (4.1)$$

Beweis:

Sei $\mathbf{A} = \{A_1, \dots, A_k\}$ eine optimale Segmentierung von f . Gezeigt wird nun, daß man, wenn \mathbf{A} die Bedingung (4.1) nicht erfüllt, aus \mathbf{A} eine Segmentierung \mathbf{A}' konstruieren kann, die ebenfalls optimal ist.

$A \in \mathbf{A}$ verletze die Bedingung (4.1). Trotzdem existieren Indizes (o. B. d. A. seien es $1, \dots, m$) mit $A \subsetneq \bigsqcup_{i=1}^m B_i$. Da A die Bedingung (4.1) verletzt, existieren noch andere Segmente der Segmentierung \mathbf{A} , deren Schnitt mit $\bigsqcup_{i=1}^m B_i$ nicht leer ist. Betrachtet wird nun zunächst die Menge B_1 . Es existieren nun Segmente $C_1, \dots, C_p \in \mathbf{A}$ mit $C_i \neq A$ und $B_1 = (\bigsqcup_{i=1}^p B_1 \cap C_i) \uplus A \cap B_1$.

Nun wird zunächst das Segment C_1 betrachtet. Genauer gesagt die Mengen $A \setminus B_1 \uplus A \cap B_1$ und $C_1 \cap B_1 \uplus C_1 \setminus B_1$. Nun ist $f_{C_1 \cap B_1} = f_{A \cap B_1} = c$ für eine Konstante c . Dann sind nach Lemma 4.2.8 für $M_{lokal} := A \cup C_1$ zwei Fälle zu unterscheiden:

a) Es gilt

$$\|f - f_{A \setminus B_1 \uplus A \cap B_1} - f_{C_1 \cap B_1 \uplus C_1 \setminus B_1}\|_{M_{lokal}} \geq \|f - f_{A \setminus B_1} - f_{A \cap B_1 \uplus C_1 \cap B_1 \uplus C_1 \setminus B_1}\|_{M_{lokal}}$$

oder es gilt

$$\|f - f_{A \setminus B_1 \uplus A \cap B_1} - f_{C_1 \cap B_1 \uplus C_1 \setminus B_1}\|_{M_{lokal}} \geq \|f - f_{A \setminus B_1 \uplus A \cap B_1 \uplus C_1 \cap B_1} - f_{C_1 \setminus B_1}\|_{M_{lokal}}$$

Da \mathbf{A} allerdings schon eine optimale Segmentierung war, muß in dem Fall (a) oder b)), der gilt, sogar Gleichheit gelten.

Im Fall a) kann man dann $A' = A \setminus B_1$ und $C'_1 = C_1 \cup (A \cap B_1)$ und damit $\mathbf{A}' = (\mathbf{A} \setminus \{A, C_1\}) \uplus \{A', C'_1\}$ setzen und erhält so eine optimale Segmentierung mit $A' \cap B_1 = O$.

Im Fall b) setzt man $A' := A \setminus B_1 \uplus A \cap B_1 \uplus C_1 \cap B_1$ und $C'_1 := C_1 \setminus B_1$. Dann ist $\mathbf{A}' := \mathbf{A} \setminus \{A, C_1\} \cup \{A', C'_1\}$ eine optimale Segmentierung mit $C'_1 \cap B_1 = O$.¹ Nun wird statt C_1 die Menge C_2 betrachtet. Wieder kann man nun die Fälle a) und b) unterscheiden. Wendet man die Ersetzungen nach der Fallunterscheidung sukzessive für alle C_2, \dots, C_p an, so existiert entweder ein $C \in \{C_2, \dots, C_p\}$ für das Fall a) eintritt, oder man kann, falls für alle $C \in$

$\{C_2, \dots, C_p\}$ Fall b) eintritt, $A' := A \cup \left(\biguplus_{C \in \{C_2, \dots, C_p\}} C \cap B_1 \right)$ und für $C \in \{C_2, \dots, C_p\}$

$C' := C \setminus B_1$ setzen. Setzt man dann $\mathbf{A}' := \mathbf{A} \setminus \{A, C_2, \dots, C_p\} \cup \{A', C'_2, \dots, C'_p\}$, so ist \mathbf{A}' eine optimale Segmentierung mit $A' \cap B_1 = B_1$. Nach den Voraussetzungen an das Mengensystem ist \mathbf{A}' wieder eine gültige Segmentierung aus $\mathfrak{S}_k(\mathfrak{A})$.

Da man dieses Verfahren für alle $B \in \{B_1, \dots, B_m\}$ anwenden kann, kann man also aus \mathbf{A} eine optimale Segmentierung \mathbf{A}' erzeugen, in der das Segment A so verändert wurde, daß es die Bedingung (4.1) nicht mehr verletzt. Da $A \in \mathbf{A}$ beliebig war folgt die Behauptung. \square

Korollar 4.2.11

Sei $\{\langle \cdot | \cdot \rangle_A \mid A \in \sigma(\mathfrak{A}), \lambda(A) > 0\}$ eine Familie der erlaubten Skalarprodukte definiert wie in Beispiel 3.4.6 und V_c wie in Definition 4.2.6. Weiter sei \mathfrak{A} ein Mengensystem erlaubter Mengen derart, daß für alle Mengen $A \in \mathfrak{A}$ und $p \in \mathbb{P}(M)$ mit $A \cap p \neq O$ gilt

i. $A \cup p \in \mathfrak{A}$

ii. $A \setminus p \in \mathfrak{A}$

Dann gilt für diskrete Pixelbilder:

Ist $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ ein lokales Maximum der Zielfunktion Ψ der k -Niveaupassung eines diskreten Pixelbildes, so existiert ein $\mathbf{A}' \in S_k(\mathfrak{A})$ mit $\Psi(\mathbf{A}') \geq \Psi(\mathbf{A})$.

Beweis:

folgt aus Satz 4.2.10, Definition 4.2.1 und der Definition 1.3.2 (S. 6)

Bemerkung 4.2.12

Unter den Voraussetzungen des Korollars 4.2.11, reicht es also aus, zur Bestimmung eines lokalen Maximums der Zielfunktion der k -Niveaupassung nur pixelbasierte Mengen zu betrachten. Ein Beispiel eines Segmentierungsproblems, das diese Voraussetzungen erfüllt findet sich in Abschnitt 4.5.

Leider erfüllt das Mengensystem der zusammenhängenden Mengen im zweidimensionalen weder die Voraussetzungen des Satzes 4.2.9 noch die des Korollars 4.2.11.

¹Man beachte, daß es an dieser Stelle des Beweises sein kann, daß die konstruierte Segmentierung kein Element aus $\mathfrak{S}_k(\mathfrak{A})$ ist.

4.2.3 Pixelbasiertes lokales Maximum

In diesem Abschnitt soll der Begriff des pixelbasierten Maximums definiert werden. Dies geschieht über die Definition einer pixelbasierten Umgebung, die nicht mit dem in der Analysis üblichen Umgebungsbegriff verwechselt werden sollte, da sie über einen festen Abstand zum Element definiert wird.

Im folgenden ist $g : M \rightarrow \mathbb{R}^d$ immer ein diskretes Pixelbild. Und $\Psi : \mathfrak{S}_k(\mathfrak{A}) \rightarrow \mathbb{R}$ die Zielfunktion einer k -Niveaupassung von g bezüglich des Mengensystems \mathfrak{A} nach Definition 3.5.24.

Definition 4.2.13 (Pixelbasierte Umgebung)

Sei $\mathbf{A} \in S_k(\mathfrak{A})$. Weiter sei $\epsilon := \min\{\vartheta(\mathbf{A}, \mathbf{B}) \mid \mathbf{B} \in S_k(\mathfrak{A}) \setminus \{\mathbf{A}\}\}$. Dann heie die Menge

$$\mathfrak{U}_p(\mathbf{A}) := \{\mathbf{B} \in S_k(\mathfrak{A}) \mid \vartheta(\mathbf{A}, \mathbf{B}) = \epsilon\}$$

die *pixelbasierte Umgebung* von \mathbf{A} in $S_k(\mathfrak{A})$.

Bemerkung 4.2.14

Der Begriff der pixelbasierten Umgebung tuscht darber hinweg, da es sich hier nicht um eine Umgebung im eigentlichen mathematischen Sinne handelt. $\mathfrak{U}_p(\mathbf{A})$ mu *keine* offene Menge in $S_k(\mathfrak{A})$, die \mathbf{A} enthlt, umfassen, hat also mit dem klassischen Umgebungsbegriff der Analysis nichts gemeinsam. Auch ist mit der pixelbasierten Umgebung von \mathbf{A} genau eine Teilmenge von $S_k(\mathfrak{A})$ festgelegt.

Definition 4.2.15 (Pixelbasiertes lokales Maximum)

Eine pixelbasierte Segmentierung $\mathbf{A} \in S_k(\mathfrak{A})$ heie *pixelbasiertes lokales Maximum* der k -Niveaupassung Ψ genau dann, wenn

$$\forall \mathbf{B} \in \mathfrak{U}_p(\mathbf{A}) : \Psi(\mathbf{A}) \geq \Psi(\mathbf{B})$$

gilt.

Bemerkung 4.2.16 (Probleme des diskreten Umgebungsbegriffs)

Sei $M := [0, 20]^2$ und $\mathfrak{A} := \{[t_1, t_2] \times [t_3, t_4] \mid t_1, t_2, t_3, t_4 \in [0, 20], t_1 \leq t_2, t_3 \leq t_4\}$. Das Mengensystem \mathfrak{A} ist also die Menge aller Rechtecke in M . Man zeigt leicht, daß \mathfrak{A} die Bedingungen aus Definition 3.2.10 (S. 74) erfüllt. \mathfrak{A} ist also ein Mengensystem erlaubter Mengen.

\mathbf{A} sei eine pixelbasierte Segmentierung von M durch 5 Segmente. Die Segmente seien dabei $A_1 := [0, 5] \times [0, 8]$, $A_2 := [0, 5] \times [8, 20]$, $A_3 := [5, 9] \times [16, 20]$, $A_4 := [9, 20] \times [16, 20]$ und $A_5 := [5, 20] \times [0, 16]$. Dann ist $\mathbf{A} := \{A_1, A_2, A_3, A_4, A_5\}$ aus $S_5(\mathfrak{A})$. Abbildung 4.4 a) zeigt eine Skizze der pixelbasierten Segmentierung \mathbf{A} .

Sei nun eine weitere pixelbasierte Segmentierung \mathbf{A}' gegeben durch die Segmente A_3, A_4, A_5 und $A'_1 := [0, 5] \times [0, 9]$ sowie $A'_2 := [0, 5] \times [9, 20]$. Es sei also $\mathbf{A}' := \{A'_1, A'_2, A_3, A_4, A_5\}$ definiert. Abbildung 4.4 b) zeigt eine Skizze von \mathbf{A}' .

Ein dritte Segmentierung \mathbf{A}'' sei gegeben durch A_1, A_2, A_5 und $A'_3 := [5, 10] \times [16, 20]$ sowie $A'_4 := [10, 20] \times [16, 20]$. Es sei also $\mathbf{A}'' := \{A_1, A_2, A'_3, A'_4, A_5\}$. Abbildung 4.4 c) zeigt eine Skizze von \mathbf{A}'' .

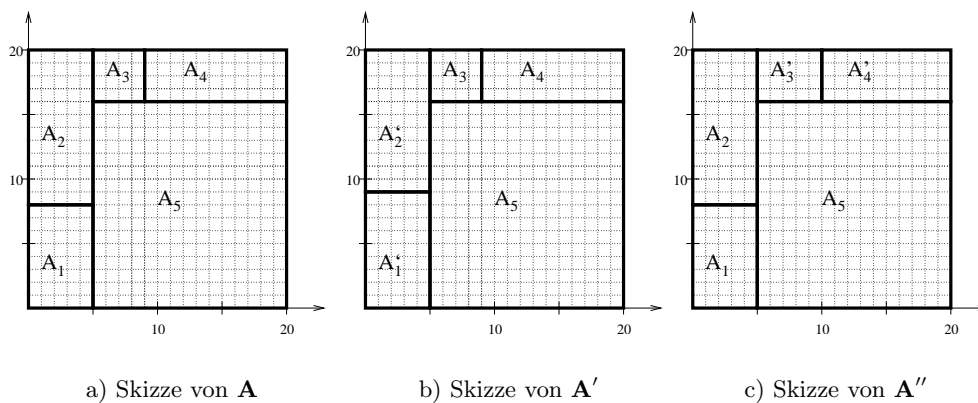


Abbildung 4.4: Skizzen der Elemente \mathbf{A} , \mathbf{A}' und \mathbf{A}'' aus $S_5(\mathfrak{A})$ aus Beispiel 4.2.16.

Es ist nun $\vartheta(\mathbf{A}, \mathbf{A}') = 10$ und $\vartheta(\mathbf{A}, \mathbf{A}'') = 8$. Man kann zeigen, daß $\mathbf{A}'' \in \mathfrak{U}_p(\mathbf{A})$ gilt. Dann kann aber \mathbf{A}' nach der Definition 4.2.13 nicht in $\mathfrak{U}_p(\mathbf{A})$ sein. Überprüft man also die pixelbasierte lokale Maximalität von \mathbf{A} , so spielt die Segmentierung \mathbf{A}' , wie alle Segmentierungen die aus Veränderungen der Segmente A_1 und A_2 entstanden sind, keine Rolle, obwohl die Veränderung $A_1, A_2 \mapsto A'_1, A'_2$ zu den kleinsten Veränderungen zählt, die diese beiden Segmente in $S_5(\mathfrak{A})$ machen können. Daher ist der Begriff der pixelbasierten Umgebung in diesem Fall nicht zufriedenstellend. Vielmehr wäre bei einer Definition eines Umgebungsbegriffs, die solche Phänomene ausschließen soll, auf die minimalen Veränderung jedes Segments zu achten, statt nur die minimalen Veränderungen der gesamten Segmentierung in die Definition des Umgebungsbegriffs aufzunehmen. An dieser Stelle wird nicht weiter auf dieses Problem eingegangen, da sich bei den meisten praktischen Anwendungen der vektoriellen Mehrniveaupassung die Definition 4.2.13 als ausreichend erwiesen hat.

4.3 Algorithmen zur Bestimmung eines pixelbasierten lokalen Maximums

Schon in Abschnitt 2.2.2 („Eigene Annäherung an den Begriff Segmentierung“) wurde gefordert, zur Beurteilung von Segmentierungsverfahren auch die algorithmische Umsetzbarkeit heranzuziehen. Dieser Abschnitt kann nun, dank den Vorarbeiten des vorangegangenen Abschnitts 4.2, Algorithmen zur Segmentierung von Bildern über die vektorielle Mehrniveaupassung formulieren. Zunächst wird ein abstrakter allgemeingültiger Algorithmus zur k -Niveaupassung präsentiert. Dieser läßt sich jedoch unter bestimmten Voraussetzungen analog zum Abschnitt 3.7.1 auf iterative 2-Niveaupassungen reduzieren, so daß sich der Hauptteil dieses Abschnitts mit der Berechnung lokaler pixelbasierter Maxima der Zielfunktion der 2-Niveaupassung auseinandersetzt. Zum Abschluß des Abschnitts werden noch Hinweise zu effektiven Update-Strategien aufgezeigt, die die Modellierung des Funktionenraums V als endlich dimensionaler Vektorraum mit einer Orthonormalbasis bezüglich jedes Segments bietet.

In diesem Abschnitt sei $g : M \rightarrow \mathbb{R}^d$ ein diskretes Pixelbild. $\Psi : \mathfrak{S}_k(\mathfrak{A}) \rightarrow \mathbb{R}$ sei die Zielfunktion einer k -Niveaupassung auf g bezüglich des Mengensystems \mathfrak{A} nach Definition 3.5.24 (S. 103).

4.3.1 Ein abstrakter Algorithmus zur Berechnung eines pixelbasierten lokalen Maximums einer k -Niveaupassung

Der Algorithmus 6 (**k-Niveaupassung**) (161) bietet ein simples allgemeingültiges Verfahren zur Berechnung eines pixelbasierten lokalen Maximums einer k -Niveaupassung, indem einfach für eine bestehende Segmentierung die pixelbasierte Umgebung der Segmentierung nach einer besseren Segmentierung abgesucht wird. Wird keine solche Segmentierung gefunden, so ist die bestehende Segmentierung nach Definition 4.2.15 ein pixelbasiertes lokales Maximum. Andernfalls wird die bessere Segmentierung als die aktuelle Segmentierung gewählt und man sucht in der pixelbasierten Umgebung dieser neuen Segmentierung nach einer besseren Segmentierung. Dieses Vorgehen wird solange wiederholt, bis ein pixelbasiertes lokales Maximum gefunden wurde.

Da $S_k(\mathfrak{A})$ endlich ist, terminiert der Algorithmus 6 (**k-Niveaupassung**) in einem pixelbasierten lokalen Maximum.

Für eine große Klasse von Mehrniveaupassungsproblemen läßt sich dieser abstrakte Algorithmus vereinfachen, insbesondere was die Wahl $\mathbf{B}' \in \mathfrak{U}_p(\mathbf{B})$ im Zeile 5 des Algorithmus angeht.

4.3.2 Reduktion der k -Niveaupassung auf iterative 2-Niveaupassungen

Eine k -Niveaupassung ist ein globales Optimierungsproblem. „Global“ heißt dabei, daß die Zielfunktion Ψ auf ganz M definiert ist, also eine Funktion auf dem ganzen Bild zu optimieren ist. Um dieses Problem zu vereinfachen, bietet es sich an, immer nur zwei Segmente

Algorithmus 6 (k-Niveaupassung) Berechnung eines pixelbasierten lokalen Maximums einer k -Niveaupassung Ψ

Input: Diskretes Pixelbild $g : M \rightarrow \mathbb{R}^d$,
beliebige Segmentierung $\mathbf{A} \in S_k(\mathfrak{A})$

Output: pixelbasiertes lokales Maximum \mathbf{B}

```

1:  $\mathbf{B} \leftarrow \mathbf{A}$ 
2: continue  $\leftarrow$  TRUE
3: while continue = TRUE do
4:   continue  $\leftarrow$  FALSE
5:   for  $\mathbf{B}' \in \mathcal{U}_p(\mathbf{B})$  do
6:     if  $\Psi(\mathbf{B}') > \Psi(\mathbf{B})$  then
7:       continue  $\leftarrow$  TRUE
8:        $\mathbf{B} \leftarrow \mathbf{B}'$ 
9:     end if
10:  end for
11: end while
12: return  $\mathbf{B}$ 

```

zu betrachten und ein lokales Maximum einer 2-Niveaupassung, definiert auf diesen zwei Segmenten, zu berechnen. Es stellt sich die Frage, ob man ein globales Maximum einer k -Niveaupassung durch iteratives Ausführen lokaler 2-Niveaupassungen erhält.

Diese Frage muß, wie die Bemerkung 4.3.3 zeigt, im allgemeinen negativ beantwortet werden. Auch die Ergebnisse des Abschnitts 3.7.1 zeigen, daß eine Reduktion auf iterative 2-Niveaupassungen, nur durch starke Bedingungen an das Mengensystem \mathfrak{A} erreicht werden kann (vgl. Satz 3.7.10 und Satz 3.7.14, 125ff). Ziel dieses Abschnitts ist es, vergleichbare Bedingungen für die Menge der pixelbasierten Segmentierungen innerhalb der Menge der Segmentierungen zu konstruieren. Diese Bedingungen werden für eine, für die meisten Bedürfnisse in den Anwendungen der Mehrniveaupassungen zur Bildsegmentierung ausreichende, Unterklasse von Mengensystemen erfüllt sein.

Definition 4.3.1

Ist \mathfrak{A} ein Mengensystem erlaubter Mengen und $\mathbf{A} \in S_k(\mathfrak{A})$. Dann sei für alle $A_i, A_j \in \mathbf{A}$ mit $A_i \cup A_j \in \mathfrak{A}$ die Menge $S_{A_i, A_j}(\mathbf{A})$ definiert durch

$$S_{A_i, A_j}(\mathbf{A}) := \{\mathbf{A} \setminus \{A_i, A_j\} \cup \{A'_i, A'_j\} \mid A'_i, A'_j \in S_2(\mathfrak{A}, A_i \cup A_j)\}$$

$S_{A_i, A_j}(\mathbf{A})$ heie die Menge der pixelbasierten Segmentierungen der Menge $A_i \cup A_j$ durch zwei Segmente innerhalb der pixelbasierten Segmentierung \mathbf{A} .

Beispiel 4.3.2

Sei $M := [0, 10]$ und $\mathfrak{A} := \{[t_1, t_2] \mid t_1, t_2 \in M\}$. Dann ist $\{[0, 1], [1, 3], [3, 10]\} \in S_3(\mathfrak{A})$. Mit Definition 4.3.1 ist dann $S_{[1, 3], [3, 10]}(\mathfrak{A}) = \{\{[0, 1], [1, n], [n, 10]\} \mid n \in \{1, \dots, 10\}\}$. (Vergleiche auch Beispiel 3.7.4, S. 124.)

Bemerkung 4.3.3 (Probleme bei der Reduktion auf die 2-Niveau-Situation)

Ist \mathfrak{A} , ähnlich wie in Bemerkung 4.2.16, eine Menge von Rechtecken in $[0, 10] \times [0, 20]$, so läßt sich das von diesem Mengensystem induzierte System von Segmentierungen nicht auf iterative 2-Niveaupassungen reduzieren, um allgemein lokale pixelbasierte Maxima der Zielfunktion der k -Niveaupassung zu bestimmen. Man kann sich leicht an Abbildung 4.5 vergegenwärtigen, daß sich bei minimalen Veränderungen einer pixelbasierten Segmentierung unter Umständen mehr als ein Segment verändern muß, um innerhalb des erlaubten Mengensystems zu bleiben.

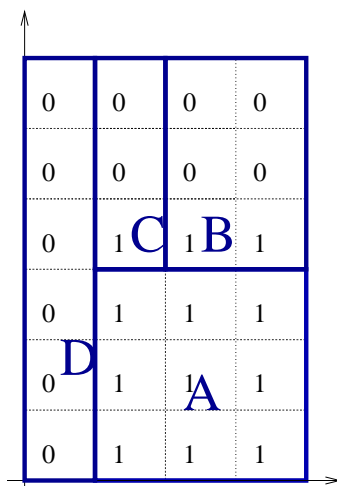


Abbildung 4.5: Beispielskizze eines Falls, in dem sich eine 4-Niveaupassung nicht auf iterative 2-Niveaupassungen zurückführen läßt: Um eine optimale Segmentierung der gegebenen Funktion durch vier Rechtecke gemäß Beispiel 4.3.3 zu erhalten, müßte das Segment A Pixel von C und B erhalten. Dies kann jedoch, ohne das Mengensystem zu verlassen, nur in einem Schritt geschehen: Denn es ist A nicht möglich nur eine Teilmenge von C aufzunehmen, da es sonst kein Rechteck mehr sein würde. Man beachte, daß die optimale Segmentierung in der pixelbasierten Umgebung der dargestellten Segmentierung liegt.

Die folgende Definition 4.3.4 formuliert eine Bedingung an das System von Segmentierungen, die hinreichend ist, um das Problem der Mehrniveaupassung auf die iterative Ausführung von 2-Niveaupassungen zurückzuführen,

Definition 4.3.4 (Hinreichende Bedingung für die Niveaureduktion)

$S_k(\mathfrak{A})$ heiße auf die Zwei-Niveau-Situation reduzierbar, falls für alle $\mathbf{A} \in S_k(\mathfrak{A})$ mit $O \notin \mathbf{A}$ gilt:

$$\forall \mathbf{A}' \in \mathfrak{U}_p(\mathbf{A}) : \exists A_i, A_j \in \mathbf{A} : A_i \cup A_j \in \mathfrak{A} \wedge \mathbf{A}' \in S_{A_i, A_j}(\mathbf{A})$$

Bemerkung 4.3.5

Man beachte die Analogie der hinreichenden Bedingung 4.3.4 zur Voraussetzung des Satzes 3.7.14 (S. 127).

Die Menge der auf die Zwei-Niveau-Situation reduzierbaren Segmentierungen ist nicht leer, wie Beispiel 4.3.2 zeigt:

In diesem Beispiel ist für alle Elemente \mathbf{A} aus $S_3(\mathfrak{A})$ mit $O \notin \mathbf{A}$ der Wert $\epsilon = \min\{\vartheta(\mathbf{A}, \mathbf{B}) \mid \mathbf{B} \in S_3(\mathfrak{A}) \setminus \{\mathbf{A}\}\} = 2$. Dieser kann aber nur durch die Veränderung eines Pixels in der Segmentierung \mathbf{A} erreicht werden: Also können nur zwei Segmente der Segmentierung \mathbf{A} verändert werden: Eines, das das Pixel abgibt, und eines, das das Pixel aufnimmt.

Satz 4.3.6

Ist $S_k(\mathfrak{A})$ auf die Zwei-Niveau-Situation reduzierbar und ist $\mathbf{A} = \{A_1, \dots, A_k\} \in S_k(\mathfrak{A})$ mit $O \notin \mathbf{A}$, so gilt:

Gilt für alle $i \neq j$ mit $A_i \cup A_j \in \mathfrak{A}$, daß $\{A_i, A_j\}$ ein pixelbasiertes lokales Maximum der 2-Niveaupassung

$$\Psi_{A_i, A_j}^{(2)} : S_2(\mathfrak{A}, A_i \cup A_j) \rightarrow \mathbb{R} \\ \{B, (A_i \cup A_j) \setminus B\} \mapsto \|\gamma_g(B)\|_2^2 + \|\gamma_g((A_i \cup A_j) \setminus B)\|_2^2$$

ist ($i, j \in \{1, \dots, k\}$),

so ist \mathbf{A} ein pixelbasiertes lokales Maximum der k -Niveaupassung Ψ .

Beweis:

Folgt direkt mit Definition 4.3.4 und der Definition 4.2.15 des pixelbasierten lokalen Maximums (vgl. auch Bemerkung 3.7.5). □

Bemerkung 4.3.7

Die Einschränkung auf Mengen $\mathbf{A} \in S_k(\mathfrak{A})$ mit $O \notin \mathbf{A}$ in der Voraussetzung des Satzes 4.3.6 ist in der Praxis ohne große Bedeutung. Ist nämlich eine Segmentierung \mathbf{A} mit $O \in \mathbf{A}$ ein pixelbasiertes lokales Optimum im Sinne der Definition 4.2.15, so kann man sich leicht überlegen, daß in der pixelbasierten Umgebung eine Segmentierung \mathbf{A}' existieren muß, für die $\Psi(\mathbf{A}') = \Psi(\mathbf{A})$ gilt (vgl. auch Bemerkung 3.7.15, S. 127).

Algorithmus 7 (Iteration) Berechnung eines pixelbasierten lokalen Maximums einer k -Niveaupassung Ψ

Input: Diskretes Pixelbild $g : M \rightarrow \mathbb{R}^d$,
beliebige Segmentierung $\mathbf{A} = \{A_1, \dots, A_k\} \in S_k(\mathfrak{A})$ mit $O \notin \mathbf{A}$

Output: pixelbasiertes lokales Maximum $\{B_1, \dots, B_k\}$

```

1:  $\mathbf{B} = \{B_1, \dots, B_k\} \leftarrow \{A_1, \dots, A_k\}$ 
2: continue  $\leftarrow$  TRUE
3: while continue = TRUE do
4:   continue  $\leftarrow$  FALSE
5:   rem Schleife über alle möglichen 2-Niveaueombinationen
6:   for  $i = 1$  bis  $i = k - 1$  do
7:     for  $j = i$  bis  $j = k$  do
8:       if  $B_i \cup B_j \in \mathfrak{A}$  then
9:         Berechne  $B'_i, B'_j$  mit  $\{B'_i, B'_j\}$  ist ein pixelbasiertes lokales Maximum von

```

$$\Psi_{B_i, B_j}^{(2)} : S_2(B_i \cup B_j) \rightarrow \mathbb{R}$$

$$\{C, (B_i \cup B_j) \setminus C\} \mapsto \|\gamma_g(C)\|_2^2 + \|\gamma_g((B_i \cup B_j) \setminus C)\|_2^2$$

```

10:        if  $B'_i \neq B_i$  und  $B'_j \neq B_j$  then
11:          continue  $\leftarrow$  TRUE
12:           $B_i \leftarrow B'_i, B_j \leftarrow B'_j$ 
13:        end if
14:      end if
15:    end for
16:  end for
17: end while
18: return  $\{B_1, \dots, B_k\}$ 

```

Satz 4.3.8

Ist $S_k(\mathfrak{A})$ ein auf die Zwei-Niveau-Situation reduzierbares pixelbasiertes System von Segmentierungen, so berechnet der Algorithmus 7 (**Iteration**) ein pixelbasiertes lokales Maximum der Zielfunktion Ψ .

Beweis:

Mit Satz 4.3.6 folgt, daß bei Termination des Algorithmus $\forall \mathbf{B}' \in \mathfrak{U}_p(\mathbf{B}) : \Psi(\mathbf{B}) \geq \Psi(\mathbf{B}')$ gilt.

Zu zeigen ist also nur noch, daß der Algorithmus terminiert. Dies folgt jedoch aus der Tatsache, daß nur ein endliches Mengensystem als Definitionsmenge genommen wird und der Wert $\Psi(\mathbf{B})$ in der Doppelschleife des Algorithmus streng monoton steigt. Dies ist jedoch nur für endliche viele Zustände möglich.

Bemerkung: Streng genommen müßte in Zeile 9 gefordert werden, daß B'_i und B'_j nicht leer sind. Mit Bemerkung 4.3.7 kann diese Forderung immer erfüllt werden.

□

4.3.3 2-Niveaupassungsalgorithmen

Der vorangegangene Abschnitt motiviert dazu, sich noch einmal gezielt mit Zwei-Niveau-Situation auseinanderzusetzen. Der Algorithmus zur 2-Niveaupassung, der an dieser Stelle vorgeschlagen wird, sucht solange in der pixelbasierten Umgebung einer aktuellen Segmentierung, bis eine bessere Segmentierung gefunden wurde. Diese wird dann als aktuelle Segmentierung gesetzt. Der Algorithmus endet, wenn keine Verbesserung in der pixelbasierten Umgebung der aktuellen Segmentierung gefunden wird.

Im Mittelpunkt dieses Abschnitts stehen unter anderem Hinweise zur Auswahl einer geeigneten Segmentierung aus der pixelbasierten Umgebung der aktuellen Segmentierung und Anmerkungen zu geeigneten Update-Strategien für eine effektive Implementierung der Algorithmen.

Algorithmus 8 (2-Niveaupassung) Berechnung eines pixelbasierten lokalen Maximums einer 2-Niveaupassung Ψ bezüglich eines pixelbasierten Systems von Segmentierungen $S_2(\mathfrak{A})$

Input: diskretes Pixelbild $g : M \rightarrow \mathbb{R}^d$,
Startsegmentierung $\mathbf{A} = \{A, M \setminus A\} \in S_2(\mathfrak{A})$

Output: pixelbasiertes lokales Maximum \mathbf{B}

```

1:  $\mathbf{B} = \{B, M \setminus B\} \leftarrow \{A, M \setminus A\}$ 
2:  $max \leftarrow \|\gamma_g(B)\|_2^2 + \|\gamma_g(M \setminus B)\|_2^2$ 
3:  $continue \leftarrow \text{TRUE}$ 
4: while  $continue = \text{TRUE}$  do
5:   for  $\{B', M \setminus B'\} \in \mathfrak{U}_p(\mathbf{B})$  do
6:      $t \leftarrow \|\gamma_g(B')\|_2^2 + \|\gamma_g(M \setminus B')\|_2^2$ 
7:     if  $t > max$  then
8:        $continue \leftarrow \text{TRUE}$ 
9:        $max \leftarrow t$ 
10:     $\mathbf{B} = \{B, M \setminus B\} \leftarrow \{B', M \setminus B'\}$ 
11:   end if
12: end for
13: end while
14: return  $\mathbf{B}$ 

```

Satz 4.3.9

Der Algorithmus 8 (2-Niveaupassung) terminiert und der Rückgabewert ist ein pixelbasiertes lokales Maximum einer 2-Niveaupassung.

Beweis:

Sei $\mathbf{B} = \{B, M \setminus B\}$ das Ergebnis des Algorithmus. Sicher gilt, nach Termination für alle $\{B', M \setminus B'\} \in \mathfrak{U}_p(\mathbf{B})$: $\|\gamma_g(B)\|_2^2 + \|\gamma_g(M \setminus B)\|_2^2 \geq \|\gamma_g(B')\|_2^2 + \|\gamma_g(M \setminus B')\|_2^2$, da sonst

die **for**-Schleife nicht terminiert. Auf der anderen Seite muß die **for**-Schleife terminieren, da das System der pixelbasierten Segmentierungen endlich ist, und die Variable *max* bei jedem Durchlauf der Schleife wächst.

□

4.3.3.1 Hinweise zur Auswahl geeigneter Segmentierungen $\{B, M \setminus B\} \in \mathfrak{U}_p(\mathbf{A})$

Gegeben sei eine Segmentierung $\mathbf{A} = \{A, M \setminus A\} \in S_2(\mathfrak{A})$. In vielen Fällen, man vergleiche zum Beispiel das Mengensystem \mathfrak{A}' des Beispiels 3.2.12 (S. 75), ist $\mathfrak{U}_p(\mathbf{A})$ einelementig. Daher ist in solchen Fällen, die Wahl $\{B, M \setminus B\} \in \mathfrak{U}_p(\mathbf{A})$ determiniert.

Aber Beispiele mit zweidimensionalem Definitionsbereich M , wie zum Beispiel in Abschnitt 4.6, erzeugen eine weitaus komplexere pixelbasierte Umgebung $\mathfrak{U}_p(\mathbf{A})$. Wie schon an anderer Stelle angedeutet, befindet man sich selbst im Falle der vektoriellen 2-Niveaupassung im Kontext eines nichtlinearen Optimierungsproblems. Eine einfache Methode zur Bestimmung eines signifikanten lokalen Maximums einer pixelbasierten 2-Niveaupassung ist es, immer diejenige Segmentierung aus $\mathfrak{U}_p(\mathbf{A})$ zu wählen, die den stärksten Anstieg in der Zielfunktion verspricht. Das Korollar 3.7.42 (S. 147) kann dabei helfen, diejenige Segmentierung aus $\mathfrak{U}_p(\mathbf{A})$ zu wählen, die eine größtmögliche Wertsteigerung der Funktion $\Psi : S_2(\mathfrak{A}) \rightarrow \mathbb{R}$ verspricht.

4.3.3.2 Hinweise zur algorithmischen Bestimmung des Wertes $\|\gamma_g(A)\|_2^2$

Von zentraler Bedeutung für eine effiziente algorithmische Umsetzung einer 2- bzw. k -Niveaupassung ist die schnelle algorithmische Bestimmung des Wertes $\|\gamma_g(A)\|_2^2$. Dabei ist zu beachten, daß sich die Segmente der gerade aktuellen Segmentierung innerhalb der Algorithmen (wie z. B. Algorithmus 8 (**2-Niveaupassung**)) nur langsam ändern. In den meisten Fällen wird nur ein Pixel hinzugefügt oder entfernt. Eine einfache Realisierung der Berechnung des Wertes $\|\gamma_g(A)\|_2^2$ wäre die Umsetzung der Definition 3.5.2 für nicht-leere pixelbasierte Segmente A : $\gamma_g(A) := (\langle g | e_1^A \rangle_A, \dots, \langle g | e_n^A \rangle_A)$. Dafür ist allerdings bei jeder Veränderung des Segments das Skalarprodukt $\langle \cdot | \cdot \rangle_A$ und die Orthonormalbasis (e_1^A, \dots, e_n^A) bezüglich dieses Skalarprodukts neu zu bestimmen. Bei einem unüberlegten Vorgehen müßte man für jede Veränderung eines Segments alle Pixel dieses Segments mehrmals durchlaufen, was zu einem erheblichen Rechenaufwand führen würde. Dieser ist leicht zu vermeiden, indem man die iterative Rechenvorschrift zur Bestimmung der Orthonormalbasis gemäß Algorithmus 5 (**Gram-Schmidt**) abrollt und ausnützt, daß für das Skalarprodukt $\langle \cdot | \cdot \rangle_{A \uplus B}$ gilt: $\forall f, g \in L^2(\mathcal{B}(M), \mathcal{B}(E)) : \langle f | g \rangle_{A \uplus B} = \langle f | g \rangle_A + \langle f | g \rangle_B$ (vgl. Definition 3.4.4 iv., S. 85). Diese Eigenschaft in der Definition der erlaubten Skalarprodukte erlaubt es, Werte, die schon einmal für das Segment A berechnet wurden, zur Berechnung des Wertes $\|\gamma_g(A \uplus p)\|_2^2$ für das Segment $A \uplus p$ heranzuziehen.

Für einen Spezialfall soll an dieser Stelle vorgeführt werden, welche Werte ein Algorithmus für ein effizientes Updateverfahren zur Bestimmung des Wertes $\|\gamma_g(A)\|_2^2$ bereitstellen muß, und wie diese zu nutzen sind. Das Vorgehen kann dann entsprechend für andere Fälle vollkommen analog nachvollzogen werden.

Sei $E := \mathbb{R}$ mit der Multiplikation als Skalarprodukt. Für alle Mengen $A \in \mathfrak{A} \setminus \{O\}$ sei $\langle \cdot | \cdot \rangle_A$ definiert wie Beispiel 3.4.6. $\mathcal{E} := (1, \text{id}, \text{id}^2)$ sei eine Basis der Klasse der erlaubten

Funktionen V . Dann sind die durch das Skalarprodukt $\langle \cdot | \cdot \rangle_A$ bestimmten Basisfunktionen (e_1^A, e_2^A, e_3^A) auf Linearkombinationen von Funktionen aus \mathcal{E} zurückföhrbar, indem man die durch das Gram-Schmidtsche Orthogonalisierungsverfahren gegebene rekursive Definition der Orthonormalbasis ausformuliert.

Es ergibt sich gemäÙ Algorithmus 5 (**Gram-Schmidt**) für $A \in \mathfrak{A} \cap \mathcal{P}(M)$ mit $\lambda(A) > 0$:

$$\begin{aligned} \tilde{e}_1^A &= 1 \\ \text{und damit ist } e_1^A &= \frac{1}{\|\tilde{e}_1^A\|_A} \tilde{e}_1^A = \frac{1}{\sqrt{\lambda(A)}} \\ \text{Weiter ist} \\ \tilde{e}_2^A &= \text{id} - \langle \text{id} | e_1^A \rangle_A e_1^A \\ &= \text{id} - \frac{1}{\lambda(A)} \langle \text{id} | 1 \rangle_A 1 \\ \text{also ergibt sich } e_2^A &= \frac{1}{\|\tilde{e}_2^A\|_A} \tilde{e}_2^A \end{aligned}$$

Für \tilde{e}_3^A gilt:

$$\begin{aligned} \tilde{e}_3^A &= \text{id}^2 - \langle \text{id}^2 | e_2^A \rangle_A e_2^A - \langle \text{id}^2 | e_1^A \rangle_A e_1^A \\ &= \text{id}^2 - \frac{1}{\|\tilde{e}_2^A\|_A^2} \langle \text{id}^2 | \tilde{e}_2^A \rangle_A \tilde{e}_2^A - \frac{1}{\lambda(A)} \langle \text{id}^2 | 1 \rangle_A 1 \\ &= \text{id}^2 - \frac{1}{\|\tilde{e}_2^A\|_A^2} \left\langle \text{id}^2 \left| \text{id} - \frac{1}{\lambda(A)} \langle \text{id} | 1 \rangle_A 1 \right. \right\rangle_A \left(\text{id} - \frac{1}{\lambda(A)} \langle \text{id} | 1 \rangle_A 1 \right) \\ &\quad - \frac{1}{\lambda(A)} \langle \text{id}^2 | 1 \rangle_A 1 \\ &= \text{id}^2 - \frac{1}{\|\tilde{e}_2^A\|_A^2} \left(\langle \text{id}^2 | \text{id} \rangle_A - \frac{1}{\lambda(A)} \langle \text{id}^2 | 1 \rangle_A \langle \text{id} | 1 \rangle_A \right) \left(\text{id} - \frac{1}{\lambda(A)} \langle \text{id} | 1 \rangle_A 1 \right) \\ &\quad - \frac{1}{\lambda(A)} \langle \text{id}^2 | 1 \rangle_A 1 \end{aligned}$$

Daher ist

$$e_3^A = \frac{1}{\|\tilde{e}_3^A\|_A} \tilde{e}_3^A$$

Setzt man nun

$$\begin{aligned} a &:= -\frac{1}{\|\tilde{e}_2^A\|_A^2} \left(\langle \text{id}^2 | \text{id} \rangle_A - \frac{1}{\lambda(A)} \langle \text{id}^2 | 1 \rangle_A \langle \text{id} | 1 \rangle_A \right) \\ \text{und} \\ b &:= \frac{\langle \text{id} | 1 \rangle_A}{\lambda(A) \|\tilde{e}_2^A\|_A^2} \left(\langle \text{id}^2 | \text{id} \rangle_A - \frac{1}{\lambda(A)} \langle \text{id}^2 | 1 \rangle_A \langle \text{id} | 1 \rangle_A \right) - \frac{\langle \text{id}^2 | 1 \rangle_A}{\lambda(A)} \end{aligned}$$

so ergibt sich eine einfachere Darstellung für e_3^A :

$$e_3^A = \frac{1}{\|\tilde{e}_3^A\|_A} (\text{id}^2 + a \text{id} + b 1)$$

Für die Normen von \tilde{e}_2^A bzw. \tilde{e}_3^A gilt:

$$\begin{aligned}
\|\tilde{e}_2^A\|_A^2 &= \langle \tilde{e}_2^A | \tilde{e}_2^A \rangle_A \\
&= \left\langle \text{id} - \frac{1}{\lambda(A)} \langle \text{id} | 1 \rangle_A 1 \left| \text{id} - \frac{1}{\lambda(A)} \langle \text{id} | 1 \rangle_A 1 \right\rangle_A \\
&= \langle \text{id} | \text{id} \rangle_A - 2 \left\langle \text{id} \left| \frac{1}{\lambda(A)} \langle \text{id} | 1 \rangle_A 1 \right\rangle_A + \frac{1}{\lambda(A)^2} \langle \langle \text{id} | 1 \rangle_A 1 | \langle \text{id} | 1 \rangle_A 1 \rangle_A \\
&= \langle \text{id} | \text{id} \rangle_A - \frac{2}{\lambda(A)} \langle \text{id} | 1 \rangle_A^2 + \frac{1}{\lambda(A)^2} \langle \text{id} | 1 \rangle_A^2 \langle 1 | 1 \rangle_A \\
&= \langle \text{id} | \text{id} \rangle_A - \frac{2}{\lambda(A)} \langle \text{id} | 1 \rangle_A^2 + \frac{1}{\lambda(A)^2} \langle \text{id} | 1 \rangle_A^2 \lambda(A) \\
&= \langle \text{id} | \text{id} \rangle_A - \frac{1}{\lambda(A)} \langle \text{id} | 1 \rangle_A^2
\end{aligned}$$

und

$$\begin{aligned}
\|\tilde{e}_3^A\|_A^2 &= \langle \tilde{e}_3^A | \tilde{e}_3^A \rangle_A = \langle \text{id}^2 + \text{aid} + \text{b1} | \text{id}^2 + \text{aid} + \text{b1} \rangle_A \\
&= \langle \text{id}^2 | \text{id}^2 \rangle_A + a^2 \langle \text{id} | \text{id} \rangle_A + b^2 \langle 1 | 1 \rangle_A \\
&\quad + 2a \langle \text{id}^2 | \text{id} \rangle_A + 2b \langle \text{id}^2 | 1 \rangle_A + 2ab \langle \text{id} | 1 \rangle_A \\
&= \langle \text{id}^2 | \text{id}^2 \rangle_A + 2a \langle \text{id}^3 | 1 \rangle_A + (a^2 + 2b) \langle \text{id}^2 | 1 \rangle_A + 2ab \langle \text{id} | 1 \rangle_A + b^2 \lambda(A)
\end{aligned}$$

Nachdem man eine Darstellung der Orthonormalbasis (e_1^A, e_2^A, e_3^A) als Linearkombination der Basisvektoren aus \mathcal{E} hat, ist es möglich, den Vektor

$$\gamma_g(A) = \begin{pmatrix} \gamma_g(A)_1 \\ \gamma_g(A)_2 \\ \gamma_g(A)_3 \end{pmatrix} = \begin{pmatrix} \langle g | e_1^A \rangle_A \\ \langle g | e_2^A \rangle_A \\ \langle g | e_3^A \rangle_A \end{pmatrix}$$

ebenfalls als Linearkombination von Skalarprodukten mit den Basisvektoren aus \mathcal{E} darzustellen. Es ist nämlich

$$\begin{aligned}
\langle g | e_1^A \rangle_A &= \left\langle g \left| \frac{1}{\sqrt{\lambda(A)}} 1 \right\rangle_A \\
&= \frac{1}{\sqrt{\lambda(A)}} \langle g | 1 \rangle_A
\end{aligned}$$

und

$$\begin{aligned}
\langle g | e_2^A \rangle_A &= \frac{1}{\|\tilde{e}_2^A\|_A} \left\langle g \left| \text{id} - \frac{1}{\lambda(A)} \langle \text{id} | 1 \rangle_A 1 \right\rangle_A \\
&= \frac{1}{\|\tilde{e}_2^A\|_A} \left(\langle g | \text{id} \rangle_A - \frac{\langle \text{id} | 1 \rangle_A}{\lambda(A)} \langle g | 1 \rangle_A \right)
\end{aligned}$$

sowie

$$\begin{aligned}\langle g|e_3^A \rangle_A &= \frac{1}{\|\tilde{e}_3^A\|_A} \langle g|\text{id}^2 + a\text{id} + b1_A \rangle_A \\ &= \frac{1}{\|\tilde{e}_3^A\|_A} (\langle g|\text{id}^2 \rangle_A + a \langle g|\text{id} \rangle_A + b \langle g|1 \rangle_A)\end{aligned}$$

Da für den Algorithmus nur der Wert $\|\gamma_g(A)\|_2^2 = \sum_{i=1}^3 \gamma_g(A)_i^2$ berechnet werden muß, ist es an keiner Stelle in einer Implementierung notwendig, die Wurzel aus einem Term zu ziehen. Damit wird diese, auf den meisten Rechnerarchitekturen in der Laufzeit teure und numerisch instabile, Operation vermieden.

Ein Algorithmus zur effizienten Bestimmung des Wertes $\|\gamma_g(A)\|_2^2$ bezüglich der Basis \mathcal{E} sollte also bei jeder Veränderung des Segments A die folgenden Werte aktualisieren:

$$\begin{aligned}\lambda(A) &= \sum_{p \in A} 1, \\ \langle g|1 \rangle_A &= \sum_{p \in A} \langle g(p)|1 \rangle_p, \\ \langle g|\text{id} \rangle_A &= \sum_{p \in A} \langle g(p)|\text{id} \rangle_p, & \langle 1|\text{id} \rangle_A &= \sum_{p \in A} \langle 1|\text{id} \rangle_p, \\ \langle g|\text{id}^2 \rangle_A &= \sum_{p \in A} \langle g(p)|\text{id}^2 \rangle_p, & \langle 1|\text{id}^2 \rangle_A &= \sum_{p \in A} \langle 1|\text{id}^2 \rangle_p, \\ \langle g|\text{id}^3 \rangle_A &= \sum_{p \in A} \langle g(p)|\text{id}^3 \rangle_p, & \langle 1|\text{id}^3 \rangle_A &= \sum_{p \in A} \langle 1|\text{id}^3 \rangle_p, \\ & & \langle 1|\text{id}^4 \rangle_A &= \sum_{p \in A} \langle 1|\text{id}^4 \rangle_p\end{aligned}$$

Dabei sollten die alten berechneten Werte benutzt werden: Nimmt das Segment A ein Pixel p auf, so ist beispielsweise $\langle g|1 \rangle_{A \cup p} = \langle g|1 \rangle_A + g(p)$. Vereinigt sich das Segment A mit einem ganzen Segment B , so können ebenfalls die Werte, die für das Segment B gespeichert sind, genutzt werden: $\langle g|1 \rangle_{A \cup B} = \langle g|1 \rangle_A + \langle g|1 \rangle_B$. Verliert ein Segment A ein Pixel p , so ergibt sich $\langle g|1 \rangle_{A \setminus p} = \langle g|1 \rangle_A - g(p)$, falls $p \subset A$ ist.

4.4 Eindimensionale Bilder und das Mengenmodell der zusammenhängenden Mengen

Um die Algorithmen des vorangegangenen Abschnitts 4.3 zu veranschaulichen, soll an dieser Stelle zunächst einmal auf eine einfache Anwendung der Mehrniveaupassung eingegangen werden: Gegeben seien diskrete Grauwertpixelbilder von einem Intervall in \mathbb{R} nach \mathbb{R} , also Bilder mit eindimensionalem Definitions- und Wertebereich. Diese sollen im folgenden durch zusammenhängende Pixelmengen, also Pixelintervalle, segmentiert werden.

4.4.1 Grundsätzliches Vorgehen

Sei $M = [0, m] \subset \mathbb{R}$ mit $m \in \mathbb{N}$. Das System erlaubter Mengen \mathfrak{A} sei das Mengensystem der zusammenhängenden Mengen in $[0, m]$. Ähnlich wie in Beispiel 3.5.11 iv) ist dann $\mathfrak{A} := \{[t_1, t_2] \mid t_1 \leq t_2, t_1, t_2 \in [0, m]\}$. Daher ist $\mathfrak{A} \cap \mathbb{P}([0, m]) = \{[t_1, t_2] \mid t_1 \leq t_2, t_1, t_2 \in \{0, \dots, m\}\}$. Und damit ist für $k \in \mathbb{N}$ mit $k < m$ das System der pixelbasierten Segmentierungen von M durch k Mengen gegeben durch

$$S_k(\mathfrak{A}) = \{\{[0, t_1], \dots, [t_{k-1}, m]\} \mid 0 \leq t_1 \leq \dots \leq t_{k-1} \leq m, t_1, \dots, t_{k-1} \in \mathbb{N}_0\}$$

Man sieht leicht, daß sich jedes Element aus $S_k(\mathfrak{A})$ auch durch seine Breakpoints t_1, \dots, t_{k-1} kodieren läßt. Daher sei im Rest dieses Abschnitts mit (t_1, \dots, t_{k-1}) das Element $\{[0, t_1], \dots, [t_{k-1}, m]\}$ aus $S_k(\mathfrak{A})$ kodiert. Vereinfachend wird $(t_1, \dots, t_{k-1}) \in S_k(\mathfrak{A})$ notiert und damit die Kodierungsabbildung, die implizit hinter dieser Aussage steht, weggelassen.

Für (t_1, \dots, t_{k-1}) aus $S_k(\mathfrak{A})$ läßt sich nun leicht zeigen, daß

$$\mathfrak{U}_p((t_1, \dots, t_{k-1})) = \{(t'_1, \dots, t'_{k-1}) \in S_k(\mathfrak{A}) \mid \exists i \in \{1, \dots, k-1\} : t'_i \in \{t_i - 1, t_i + 1\}, \forall j \neq i : t'_j = t_j\}$$

gilt. Das heißt, jedes Element aus der Umgebung der Segmentierung, die durch den Vektor (t_1, \dots, t_{k-1}) kodiert ist, ist durch das Verschieben eines einzelnen Breakpoints t_i um ± 1 entstanden.²

Dies zeigt, daß dieses System der Segmentierungen auf die Zwei-Niveau-Situation reduzierbar ist. Daher folgt die Korrektheit des Algorithmus 7 (**Iteration**) für dieses Mengensystem. Der Algorithmus 9 (**1D-k-Niveaupassung**) auf der nächsten Seite ist eine auf dieses Mengensystem spezialisierte Version des Algorithmus 7 (**Iteration**), der insbesondere die einfache Auswahl einer Segmentierung aus der pixelbasierten Umgebung einer bestehenden Segmentierung durch das Verschieben eines Schwellwertes ausnützt.

In Zeile 7 des Algorithmus 9 (**1D-k-Niveaupassung**) wird ein lokales (oder globales) pixelbasiertes Maximum der 2-Niveaupassung

$$\Psi_{[t_n, t_{n+1}][t_{n+1}, t_{n+2}]}^{(2)} : S_{[t_n, t_{n+1}][t_{n+1}, t_{n+2}]}((t_1, \dots, t_{k-1})) \rightarrow \mathbb{R} \\ \{[t_n, x], [x, t_{n+2}]\} \mapsto \|\gamma_g([t_n, x])\|_2^2 + \|\gamma_g([x, t_{n+2}])\|_2^2$$

²In [Käm01] wird diese Umgebung ebenfalls über die Breakpoints definiert und als U_1 bezeichnet.

Algorithmus 9 (1D-k-Niveaupassung) k -Niveaupassung bezüglich der pixelbasierten Segmentierung eines Intervalls durch k zusammenhängende Segmente

Input: Diskretes Pixelbild $g : [0, m] \rightarrow \mathbb{R}$,
 $k \in \mathbb{N}$ mit $k \leq m$

Output: pixelbasierte Segmentierung $\{[0, t_1], \dots, [t_{k-1}, m]\}$

- 1: Wähle Startsegmentierung $0 =: t_0 < t_1 < \dots < t_{k-1} < t_k := m$
- 2: continue \leftarrow TRUE
- 3: **while** continue = TRUE **do**
- 4: **for** $n = 0$ bis $n = k - 2$ **do**
- 5: continue \leftarrow FALSE
- 6: old $\leftarrow t_{n+1}$
- 7: new $\leftarrow n_0$, wobei n_0 ein lokales (oder globales) Maximum der Funktion

$$\Psi_{[t_n, t_{n+1}][t_{n+1}, t_{n+2}]}^{(2)} : \begin{array}{l} \{t_n, \dots, t_{n+2}\} \rightarrow \mathbb{R}_+ \\ x \mapsto \|\gamma_g([t_n, x])\|_2^2 + \|\gamma_g([x, t_{n+2}])\|_2^2 \end{array}$$

ist

- 8: **if** (old \neq new) **then**
 - 9: continue \leftarrow TRUE
 - 10: $t_{n+1} \leftarrow$ new
 - 11: **end if**
 - 12: **end for**
 - 13: **end while**
 - 14: **return** $\{[0, t_1], \dots, [t_{k-1}, m]\}$
-

berechnet. Dies kann man einfach durch Ausprobieren aller möglichen pixelbasierten Segmentierungen der Menge $[t_n, t_{n+2}]$ durch zwei zusammenhängende pixelbasierte Mengen berechnen. Dabei sollte man allerdings bei der Berechnung der Funktionswerte nicht auf Update-Strategien, wie in Abschnitt 4.3.3.2 vorgeschlagen, verzichten.

Um das Mehrniveaupassungsproblem komplett zu formulieren, ist es noch nötig, sowohl die Klasse der erlaubten Funktionen V , als auch die Familie der erlaubten Skalarprodukte zu spezifizieren.

Analog zum Beispiel 3.5.11 iv) (S. 98) sei in diesem Abschnitt $E := \mathbb{R}$ mit der Multiplikation reeller Zahlen als Skalarprodukt. Dann ist für $f, g \in L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}))$ und $A \in \mathfrak{A}$ mit $\lambda(A) > 0$: $\langle f|g \rangle_A = \int_A f(x)g(x)d\lambda(x)$ und damit $\|f - g\|_M = \sqrt{\int_M (f(x) - g(x))^2 d\lambda(x)}$.

Als mögliche Klassen erlaubter Funktionen sollen an dieser Stelle zwei vorgestellt werden: Zuerst einmal der Raum der konstanten Funktionen und dann der Raum der quadratischen Funktionen. Es ist also im ersten Beispiel $\mathcal{E} := (1)$ und im zweiten Fall $\mathcal{E} := (1, \text{id}, \text{id}^2)$.

Anmerkung: Streng genommen müßte in Zeile 7 des Algorithmus 9 (**1D-k-Niveaupassung**) gefordert werden, daß das lokale (oder globale) Maximum der Funktion $\Psi_{[t_n, t_{n+1}][t_{n+1}, t_{n+2}]}^{(2)}$ in der Menge $\{t_n + 1, \dots, t_{n+2} - 1\}$ liegt, um als Ergebnis eine Segmentierung mit k echten Segmenten zu erhalten. Mit Bemerkung 4.3.7 kann diese Forderung jedoch immer erfüllt werden.

4.4.2 Beispiele für die Approximation mit konstanten Funktionen

In der ersten Beispielfamilie ist $\mathcal{E} = (1)$. Daher ist V der Raum der konstanten Funktionen von M nach \mathbb{R} , d. h. auf den einzelnen Segmenten wird mit konstanten Funktionen approximiert.

Beispiel 4.4.1

Das diskrete Pixelbild $g : [0, 500] \rightarrow \mathbb{R}_+$ sei gegeben durch

$$g(x) := \begin{cases} 0 & , 0 \leq x < 100 \\ 40 & , 100 \leq x < 200 \\ 60 & , 200 \leq x < 300 \\ 120 & , 300 \leq x \leq 500 \end{cases}$$

Abbildung 4.6 skizziert das diskrete Pixelbild g . Da das diskrete Pixelbild g selbst auf

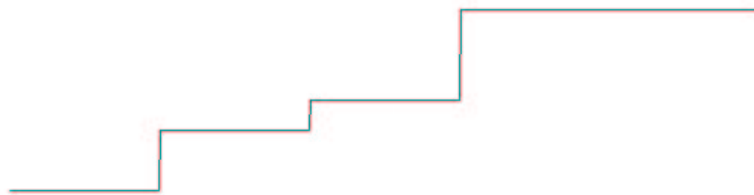


Abbildung 4.6: Skizze des diskreten Pixelbildes g .

4 Intervallen konstant ist, sollte man von einer 4-Niveaupassung erwarten, daß sie als Segmentgrenzen genau die Sprungstellen der Funktion g liefert.

Ist nun $(100, 300, 400)$ die Kodierung der Startsegmentierung, d. h. das Intervall $[0, 500]$ wird segmentiert in die Intervalle $[0, 100]$, $[100, 300]$, $[300, 400]$ und $[400, 500]$, wie es Abbildung 4.7 skizziert, so befindet sich der Algorithmus 9 (**1D-k-Niveaupassung**) gleich zu Beginn in einen pixelbasierten lokalen Optimum, aus dem er nicht mehr herauskommen kann, auch nicht wenn immer das globale Maximum der einzelnen 2-Niveaupassungen berechnet wird. Dieses pixelbasierte lokale Optimum der 4-Niveaupassung ist jedoch nicht das globale Maximum der Zielfunktion. Abbildung 4.8 visualisiert das Ergebnis des Algorithmus mit der errechneten konstanten Approximation auf den Segmenten.

Bei einer besseren Wahl der Startsegmentierung, nämlich indem man die Breakpoints zwischen die Sprünge des diskreten Grauwertbildes g legt, wie in Abbildung 4.9 skizziert, führt der Algorithmus 9 (**1D-k-Niveaupassung**) zu dem gewünschten Ergebnis, nämlich dem pixelbasierten globalen Maximum der Zielfunktion der 4-Niveaupassung (vgl. Abbildung 4.10).

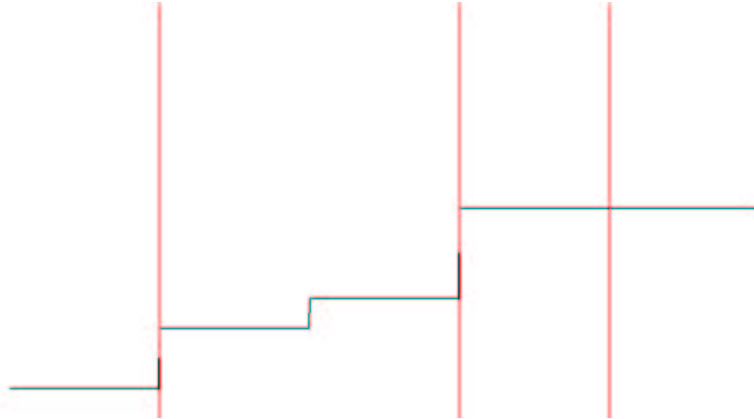


Abbildung 4.7: Breakpoints der Startsegmentierung (rot) von g , die schon ein pixelbasiertes lokales Maximum ist.

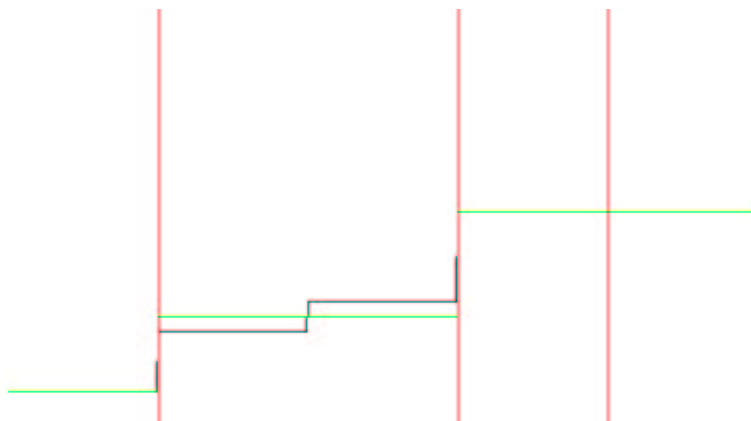


Abbildung 4.8: Breakpoints der Ergebnissegmentierung (rot) mit stückweiser Approximation (grün) auf den Segmenten.

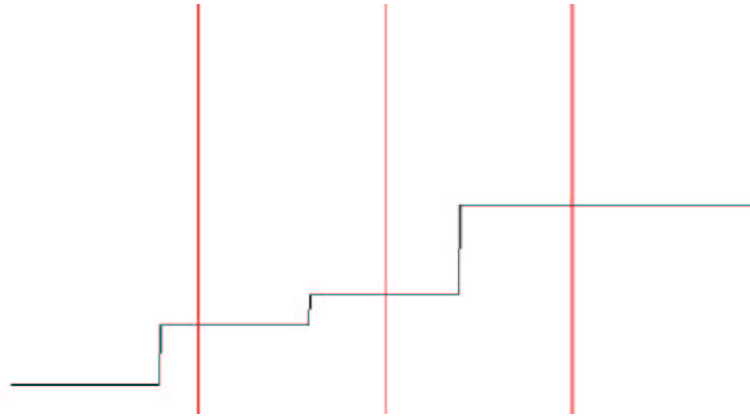


Abbildung 4.9: Alternative Breakpoints (rot) einer initialen Segmentierung einer 4-Niveaupassung von g .

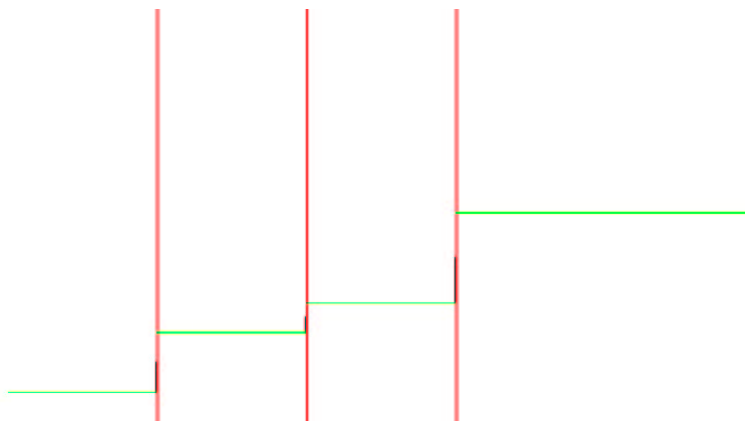


Abbildung 4.10: Ergebnis des Algorithmus 9 (**1D-k-Niveaupassung**) ausgehend von der Startsegmentierung in Abbildung 4.9.

Bemerkung 4.4.2 (Auswahl einer guten Startsegmentierung)

Das vorangegangene Beispiel 4.4.1 zeigt, wie wichtig die Auswahl einer guten Startsegmentierung ist. Es stellt sich die Frage, wie man algorithmisch eine gute Startsegmentierung generiert. Im Abschnitt 5.3 wird der Algorithmus 16 (**Iterative Niveaoverschmelzung**) (S. 226) vorgestellt, der ausgehend von einer gegebenen Segmentierung mit k_1 Segmenten, diese solange iterativ verschmilzt bis k_2 Segmente ($k_2 < k_1$) übrig bleiben. Dabei wird bei der Auswahl der zu verschmelzenden Segmente darauf geachtet, daß ihre Verschmelzung einen möglichst kleinen Fehler in der Funktion Ψ verursacht. Diesen Algorithmus kann man dann wie folgt anwenden, um eine gute Startsegmentierung für eine Mehrniveaupassung mit k Niveaus zu erhalten:

- i. Berechne mit dem Algorithmus 9 (**1D-k-Niveaupassung**) ein lokales pixelbasiertes Maximum einer l -Niveaupassung. Dabei sei $l > k$.
- ii. Verschmelze mit dem Algorithmus 16 (**Iterative Niveaoverschmelzung**) solange Segmente, bis noch k Segmente übrig bleiben.
- iii. Nutze das Ergebnis von ii. als Startsegmentierung einer k -Niveaupassung im Algorithmus 9 (**1D-k-Niveaupassung**)

Beispiel 4.4.3 (Verrauschtes Eingangssignal)

Gegeben sei wieder das diskrete Grauwertpixelbild g aus Beispiel 4.4.1. Nun jedoch wird das Signal durch ein Rauschen, das uniform auf $[-20, 20]$ verteilt ist, gestört. Ein typisches Ergebnis einer solchen Rauschaddition ist in Abbildung 4.11 skizziert.

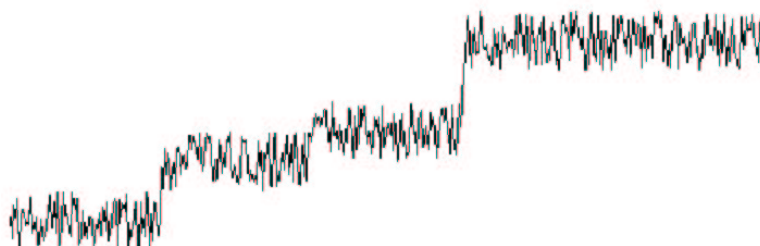


Abbildung 4.11: Funktion g aus Beispiel 4.4.1 mit additiven Rauschen, uniform auf $[-20, 20]$.

Das Ergebnis einer 4-Niveaupassung mit pixelbasierten zusammenhängenden Mengen ist in Abbildung 4.12 zu sehen. Wie man sieht, ist das Ergebnis mit dem Ergebnis der 4-Niveaupassung des unverrauschten Originals aus Beispiel 4.4.1 durchaus vergleichbar. Ein Grund dafür ist, daß bei der Berechnung der einzelnen Maxima von 2-Niveaupassungen

in Algorithmus 9 (**1D-k-Niveaupassung**) immer das globale Maximum berechnet wurde, und so nicht-signifikante pixelbasierte lokale Maxima der einzelnen 2-Niveaupassungen unterdrückt wurden.

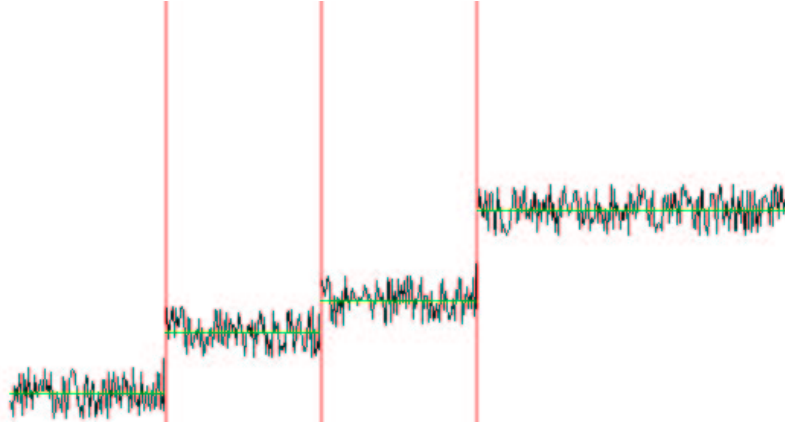


Abbildung 4.12: Ergebnis des Algorithmus 9 (**1D-k-Niveaupassung**) angewandt auf das verrauschte diskrete Pixelbild g aus Abbildung 4.11. Die Segmentgrenzen sind rot und die Approximationen mit konstanten Funktionen auf den Segmenten grün dargestellt.

4.4.3 Beispiel für die Approximation mit quadratischen Funktionen

In der zweiten Beispielreihe der Segmentierung eindimensionaler Bilder mit zusammenhängenden Mengen sei die Basis der Klasse der erlaubten Funktionen $\mathcal{E} := (1, \text{id}, \text{id}^2)$. Das heißt, auf den einzelnen Segmenten wird mit quadratischen Funktionen approximiert.

Beispiel 4.4.4

Die Funktion $\text{pix} : \mathbb{R} \rightarrow \mathbb{R}$ sei gegeben durch $\text{pix}(x) := \text{round}(x - \frac{1}{2})$. Sie ist daher nach Definition 1.3.1 auf (eindimensionalen) Pixeln konstant. Daher handelt es sich bei der Funktion $g : [0, 500] \rightarrow \mathbb{R}_+$ gegeben durch

$$g(x) = \begin{cases} -\frac{1}{200}\text{pix}(x)^2 + 2\text{pix}(x) & , 0 \leq x < 200 \\ 320 - \frac{3}{5}\text{pix}(x) & , 200 \leq x \leq 500 \end{cases}$$

um ein diskretes Pixelbild.

Dieses diskrete Pixelbild ist auf $[0, 200]$ im wesentlichen (d. h. bis auf die round-Operation) quadratisch und auf $[200, 500]$ im wesentlichen linear. Vom Algorithmus 9 (**1D-k-Niveaupassung**) wird man nun verlangen, daß er die Segmentgrenze einer 2-Niveaupassung bezüglich des Mengensystems der zusammenhängenden Mengen genau an die Stelle des Wechsels vom quadratischen zum linearen Teil der Funktion legt. Abbildung 4.13 skizziert das diskrete Pixelbild g . Man beachte, daß die Skizze eine stetige Funktion suggeriert, obwohl g durch die round-Operation nicht stetig sein kann.



Abbildung 4.13: Skizze des diskreten Pixelbildes g aus Beispiel 4.4.4.

Die Abbildung 4.14 zeigt das Ergebnis einer 2-Niveaupassung mit quadratischen Funktionen bezüglich des Mengensystems der pixelbasierten zusammenhängenden Mengen. Im Gegensatz zum Beispiel 4.4.1 spielt an dieser Stelle die Wahl der Startsegmentierung keine Rolle, weil es nur ein lokales Maximum der Zielfunktion der 2-Niveaupassung gibt.

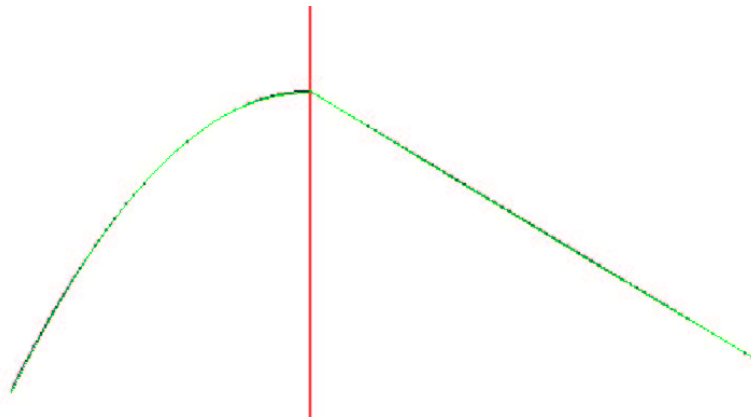


Abbildung 4.14: Ergebnis einer 2-Niveaupassung mit Algorithmus 9 (**1D-k-Niveaupassung**): Die Segmentgrenze (rot) wird genau dorthin gelegt, wo der quadratische Teil der Funktion in den linearen umbricht. Die Approximation auf den einzelnen Segmenten (grün) ist genau die Funktion (bis auf Rundungsfehler, die zum Teil auf die round-Operation zurückzuführen sind).

Die Abbildung 4.15 zeigt die gleiche Pixelfunktion g überlagert mit einem starken uniformen Rauschen.

Auch in diesem Fall verhält sich der Algorithmus 9 (**1D-k-Niveaupassung**) gutartig und setzt die Grenze der zwei zu ermittelnden Segmente an fast die gleiche Stelle wie im unverrauschten Fall (vgl. Abbildung 4.16).

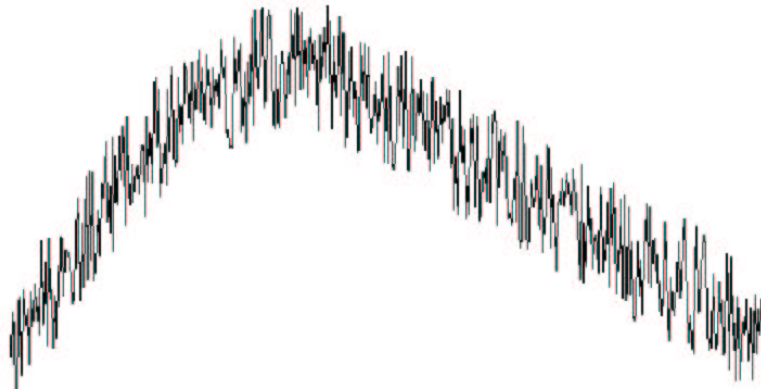


Abbildung 4.15: Skizze der mit einem starkem uniformen Rauschen überlagerten Pixelfunktion g .

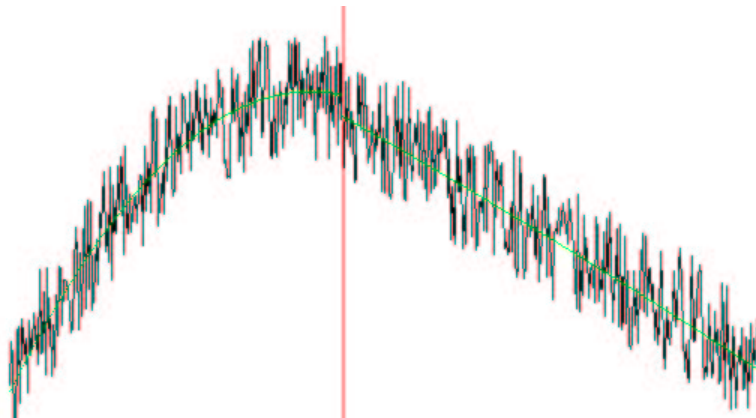


Abbildung 4.16: Ergebnis einer 2-Niveaupassung mit Algorithmus 9 (**1D-k-Niveaupassung**): Die Segmentgrenze (rot) wird fast genau dort hin gelegt, wo der quadratische Teil der unverrauschten Funktion in den linearen umbricht.

4.5 Das Mengenmodell der grauwertbenachbarten Pixelmengen

Dieser und der folgende Abschnitt widmet sich Bildern im herkömmlichen Sinne, also Bildern mit zweidimensionalem Definitionsbereich. Im Mittelpunkt dieses Abschnitts soll nun das einfachste aller Mengenmodelle stehen: Es sei im folgenden $\mathfrak{A} := \mathcal{B}(M)/\sim$. Dann ergibt sich als $S_k(\mathfrak{A})$ die Menge $\{\{A_1, \dots, A_l\} \mid A_i \subset \mathcal{P}(M), 1 \leq i \leq l \leq k, \bigcup_{i=1}^l A_i = M\}$. Es sei $E := \mathbb{R}$ mit der Multiplikation als Skalarprodukt. Weiter sei V der Vektorraum der konstanten Funktionen von M nach \mathbb{R} mit Basis $\mathcal{E} = (1)$. Die Familie der erlaubten Skalarprodukte sei definiert wie in Beispiel 3.4.6 (S. 86).

Eine k -Niveaupassung eines diskreten Grauwertpixelbildes bezüglich $S_k(\mathfrak{A})$ und dem Raum der konstanten Funktionen ist nicht anderes als ein Schwellwertverfahren, wie der Abschnitt 4.5.1 zeigt. Der zum globalen Maximum einer 2-Niveaupassung bezüglich $S_k(\mathfrak{A})$ und V korrespondierende Schwellwert entspricht dabei genau dem Schwellwert, den auch die Varianzanalyse von Otsu (vgl. Abschnitt 2.3.1.1) als Schwellwert liefert. Dies wird in Abschnitt 4.5.2 gezeigt. Auf der anderen Seite kann man aber auch zeigen, daß die k -Niveaupassung über ein diskretes Grauwertpixelbild g in diesem Kontext nichts anderes ist als die Quantisierung des Histogramms des Bildes, wie in Abschnitt 4.5.3 dargelegt wird.

Sei nun g ein diskretes Grauwertpixelbild mit Wertebereich $\{0, \dots, G\}$.

4.5.1 Die Mehrniveaupassung als Schwellwertverfahren

Will man auf g bezüglich der Potenzmenge der Pixel eine k -Niveaupassung mit konstanten Funktionen durchführen, so bietet es sich an, daß man, wenn man ein Pixel zu einer Menge hinzunimmt, sofort alle Pixel mit dem gleichen Grauwert hinzufügt: Denn auf der einen Seite existiert mit Korollar 4.2.11 eine optimale Segmentierung, die Pixel mit gleichem Grauwert nicht in verschiedene Segmente verteilt, und auf der anderen Seite gibt es bei diesem Mengenmodell keine Begründung, warum ein Pixel mit Grauwert i zu einem Segment A gehören soll, ein anderes Pixel mit dem gleichen Grauwert hingegen nicht. Das heißt in diesem Fall genügt es, Vereinigungen von Mengen der Form $\{g = i\} \subset \mathfrak{A} \cap \mathcal{P}(M)$ zu betrachten.

Das Beispiel 3.7.38 i) auf Seite 142 zeigt darüberhinaus, daß ein Segment einer optimalen Segmentierung aus *grauwertbenachbarten* Elementen aus \mathfrak{G}' bestehen muß. Zwei Mengen $\{g = i\}$ und $\{g = j\}$ heißen dabei grauwertverbunden genau dann, wenn $|i - j| = 1$ gilt. Denn:

Sei $P_i = \{g = i\} \in \mathfrak{A} \cap \mathcal{P}(M)$ ein Segment einer Segmentierung für ein $i \in \{0, \dots, G\}$. Soll nun ein Pixel hinzugefügt werden, so gilt mit Korollar 3.7.37 und Bemerkung 3.7.41, daß zwei Konstanten $c_1 > 0, c_2 > 0$ existieren, derart, daß für alle $p \in \mathbb{P}(M \setminus P_i)$ gilt:

$$\begin{aligned} \|g - g_{P_i \cup p} - g_{M \setminus (P_i \cup p)}\|_M &< \|g - g_{P_i} - g_{M \setminus P_i}\|_M \\ &\Leftrightarrow c_1(\tilde{\gamma}_g(p) - \tilde{\gamma}_g(P_i))^2 < c_2(\tilde{\gamma}_g(p) - \tilde{\gamma}_g(M \setminus P_i)) \end{aligned}$$

Nun ist aber mit Beispiel 3.5.11 i) (S. 96) $\tilde{\gamma}_g(p) = g(p)$ und $\tilde{\gamma}_g(P_i) = i$. Wenn also ein Pixel $p \in M \setminus P_i$ mit $\|g - g_{P_i \cup p} - g_{M \setminus (P_i \cup p)}\|_M < \|g - g_{P_i} - g_{M \setminus P_i}\|_M$ existiert, so muß

entweder ein Pixel mit dem Grauwert $i + 1$ oder ein Pixel mit Grauwert $i - 1$ diese Approximationsverbesserung ebenfalls erfüllen.

Zusammenfassend gilt also:

- Es sind nur Vereinigungen von Mengen der Form $\{g = i\}$ zu betrachten.
- Es sind nur grauwertbenachbarte Pixel auszutauschen.

Daher genügt es, im folgenden nur das Mengensystem

$$\mathfrak{G} := \left\{ \bigcup_{n_1 \leq i \leq n_2} P_i \mid 0 \leq n_1 \leq n_2 \leq G, P_i = \{g = i\} \right\} \subset \mathfrak{A} \cap \mathcal{P}(M)$$

zu betrachten. Dies bedeutet aber, daß man eigentlich gleich auf dem Histogramm h_g des Bildes g arbeiten kann (vgl. Definition 1.3.4, S. 6).

Definition 4.5.1 (grauwertbenachbarte Pixelmengen)

Das Mengensystem \mathfrak{G} , definiert wie oben, heie *das Mengenmodell der grauwertbenachbarten Pixelmengen*.

Bemerkung 4.5.2

- Beim Mengenmodell \mathfrak{G} der grauwertbenachbarten Pixelmengen handelt es sich eigentlich nur um ein Metamengenmodell: Es ist eine Menge sinnvoller Pixelmengen innerhalb des Mengensystems $\mathcal{P}(M)$.
- Das Mengenmodell \mathfrak{G} erlaubt ein Arbeiten auf dem Histogramm des Bildes. Da es fur die k -Niveaupassung bezuglich dieses Mengenmodells unerheblich ist, wo sich die Pixel im Bild befinden, handelt es sich bei diesem Ansatz eigentlich um keinen echten zweidimensionalen Ansatz zur Bildsegmentierung.

4.5.2 Zusammenhang zwischen 2-Niveaupassung und der Varianzanalyse von Otsu

Sei h_g das Histogramm des diskreten Grauwertpixelbildes $g : M \rightarrow \{0, \dots, G\}$ gema Definition 1.3.4 (S. 6). Betrachtet wird nun die Situation fur zwei Niveaus. Dann ergibt

sich für die Funktion $\Psi_g : \mathfrak{G} \rightarrow \mathbb{R}, A \mapsto \|\gamma_g(A)\|_2^2 + \|\gamma_g(M \setminus A)\|_2^2$ nach Beispiel 3.4.12 i) (S. 90):

$$\begin{aligned} \Psi_g\left(\bigsqcup_{0 \leq i \leq n} P_i\right) &= \tilde{\gamma}_g\left(\bigsqcup_{0 \leq i \leq n} P_i\right) \int_{\bigsqcup_{0 \leq i \leq n} P_i} g d\lambda + \tilde{\gamma}_g(M \setminus \bigsqcup_{0 \leq i \leq n} P_i) \int_{M \setminus \bigsqcup_{0 \leq i \leq n} P_i} g d\lambda \\ &= \tilde{\gamma}_g\left(\bigsqcup_{0 \leq i \leq n} P_i\right) \sum_{j=0}^n \lambda(P_j)j + \tilde{\gamma}_g(M \setminus \bigsqcup_{0 \leq i \leq n} P_i) \sum_{j=n+1}^G \lambda(P_j)j \\ &= \tilde{\gamma}_g\left(\bigsqcup_{0 \leq i \leq n} P_i\right) \sum_{j=0}^n h_g(j)j + \tilde{\gamma}_g(M \setminus \bigsqcup_{0 \leq i \leq n} P_i) \sum_{j=n+1}^G h_g(j)j \end{aligned}$$

Weiter ist

$$\tilde{\gamma}_g\left(\bigsqcup_{0 \leq i \leq n} P_i\right) = \begin{cases} \frac{\int_{\bigsqcup_{0 \leq i \leq n} P_i} g d\lambda}{\lambda(\bigsqcup_{0 \leq i \leq n} P_i)} = \frac{\sum_{j=0}^n h_g(j)j}{\sum_{j=0}^n h_g(j)} & , \text{ falls } \sum_{j=0}^n h_g(j) > 0 \\ 0 & , \text{ sonst} \end{cases}$$

Daher ergibt sich

$$\begin{aligned} \Psi_g\left(\bigsqcup_{0 \leq i \leq n} P_i\right) & \tag{4.2} \\ &= \begin{cases} \frac{\left(\sum_{j=0}^n h_g(j)j\right)^2}{\sum_{j=0}^n h_g(j)} + \frac{\left(\sum_{j=n+1}^G h_g(j)j\right)^2}{\sum_{j=n+1}^G h_g(j)} & , \text{ falls } \sum_{j=0}^n h_g(j) > 0 \text{ und } \sum_{j=n+1}^G h_g(j) > 0 \\ \frac{\left(\sum_{j=0}^n h_g(j)j\right)^2}{\sum_{j=0}^n h_g(j)} & , \text{ falls } \sum_{j=0}^n h_g(j) > 0 \text{ und } \sum_{j=n+1}^G h_g(j) = 0 \\ \frac{\left(\sum_{j=n+1}^G h_g(j)j\right)^2}{\sum_{j=n+1}^G h_g(j)} & , \text{ falls } \sum_{j=0}^n h_g(j) = 0 \text{ und } \sum_{j=n+1}^G h_g(j) > 0 \\ 0 & , \text{ sonst} \end{cases} \end{aligned}$$

Definiert man nun $\sigma : \{0, \dots, G\} \rightarrow \mathbb{R} : \sigma(n) := \Psi_g(\bigsqcup_{0 \leq i \leq n} P_i)$, so erhält man als Optimierungsproblem genau den gleichen Term, der auch in der Varianzanalyse von Nobuyuki Otsu (vgl. Abschnitt 2.3.1.1) maximiert wird. Den Beweis dieser Behauptung liefert ein Vergleich mit Gleichung (2.5) in der Bemerkung 2.3.1 auf Seite 22.

4.5.3 Zusammenhang zwischen Quantisierung und 2-Niveaupassung

Untersucht werden soll zunächst einmal die allgemeine Quantisierung bezüglich des Fehlers $\|f - \cdot\|_{\text{quant}}$ (vgl. Abschnitt 2.5, S. 49) durch zwei Mengen.

Gegeben sei ein Mengensystem $\mathfrak{A}' \subset \mathcal{B}([0, G])$ erlaubter Mengen. Für eine Funktion $f \in L^2(\mathcal{B}([0, G]), \mathcal{B}(\mathbb{R}_+))$ sei die 2-optimale Quantisierung bezüglich des Fehlers

$\|f - \cdot\|_{\text{quant}}$ gesucht. Das heißt, gesucht ist eine 2-Niveaufunktion $s_2 \in \mathcal{S}_2$ mit $\mathcal{S}_2 := \{\alpha \mathbf{1}_A + \beta \mathbf{1}_{M \setminus A} \mid A \in \mathfrak{A}'\}$ für die $E[(f - s_2)^2]$ minimal wird. Also sind α, β und $A \in \mathfrak{A}'$ gesucht für die $Q^* : \mathfrak{A}' \times \mathbb{R} \times \mathbb{R}$ mit

$$Q^*(A, \alpha, \beta) := \int f(x) (x - \alpha \mathbf{1}_A(x) + \beta \mathbf{1}_{M \setminus A}(x))^2 d\lambda(x)$$

minimal wird. Zunächst einmal gilt:

$$\begin{aligned} Q^*(A, \alpha, \beta) &= \int f(x) (x^2 - 2\alpha x \mathbf{1}_A - 2\beta x \mathbf{1}_{M \setminus A} + \alpha^2 \mathbf{1}_A(x) + \beta^2 \mathbf{1}_{M \setminus A}(x)) d\lambda(x) \\ &= \int x^2 f(x) d\lambda(x) - 2\alpha \int_A x f(x) d\lambda(x) - 2\beta \int_{M \setminus A} x f(x) d\lambda(x) \\ &\quad + \alpha^2 \int_A f(x) d\lambda(x) + \beta^2 \int_{M \setminus A} f(x) d\lambda(x) \end{aligned}$$

Durch ein Ableitungsargument kann man zeigen, daß die Werte α und β allein durch die Wahl der Menge A bestimmt sind, wenn man das gewünschte Minimum erhalten will.

Es ist nämlich, bei festem A für

$$h : (\alpha, \beta) \mapsto \alpha^2 \int_A f(x) d\lambda(x) - 2\alpha \int_A x f(x) d\lambda(x) + \beta^2 \int_{M \setminus A} f(x) d\lambda(x) - 2\beta \int_{M \setminus A} x f(x) d\lambda(x)$$

(α_0, β_0) mit

$$\alpha_0 = \frac{\int_A x f(x) d\lambda(x)}{\int_A f(x) d\lambda(x)} \quad \text{und} \quad \beta_0 = \frac{\int_{M \setminus A} x f(x) d\lambda(x)}{\int_{M \setminus A} f(x) d\lambda(x)}$$

das globale Minimum von h , falls es existiert. Es ist nämlich (α_0, β_0) die einzige Nullstelle der Ableitung von h , bei positiv definiter Hesse-Matrix. Daher ist h konvex und (α_0, β_0) das globale Minimum von h .

Definiert man nun für α_0, β_0 wie oben die Funktion $Q : \mathfrak{A}' \rightarrow \mathbb{R}_+$ durch

$$Q(A) := \begin{cases} Q(A, \alpha_0, \beta_0) & , \text{ falls } \int_A f d\lambda > 0 \text{ und } \int_{M \setminus A} f d\lambda > 0 \\ Q(A, 0, \beta_0) & , \text{ falls } \int_A f d\lambda = 0 \text{ und } \int_{M \setminus A} f d\lambda > 0 \\ Q(A, \alpha_0, 0) & , \text{ falls } \int_A f d\lambda > 0 \text{ und } \int_{M \setminus A} f d\lambda = 0 \\ Q(A, 0, 0) & , \text{ falls } \int_A f d\lambda = 0 \text{ und } \int_{M \setminus A} f d\lambda = 0 \end{cases}$$

so ergibt sich für alle $A \in \mathfrak{A}'$ mit $\int_A f d\lambda > 0$ und $\int_{M \setminus A} f d\lambda > 0$:

$$\begin{aligned}
Q(A) &= \int x^2 f(x) d\lambda(x) + \left(\frac{\int_A x f(x)}{\int_A f(x) d\lambda(x)} \right)^2 \int_A f(x) d\lambda(x) \\
&\quad - 2 \left(\frac{\int_A x f(x)}{\int_A f(x) d\lambda(x)} \right) \int_A x f(x) d\lambda(x) \\
&\quad + \left(\frac{\int_{M \setminus A} x f(x)}{\int_{M \setminus A} f(x) d\lambda(x)} \right)^2 \int_{M \setminus A} f(x) d\lambda(x) \\
&\quad - 2 \left(\frac{\int_{M \setminus A} x f(x)}{\int_{M \setminus A} f(x) d\lambda(x)} \right) \int_{M \setminus A} x f(x) d\lambda(x) \\
&= \int x^2 f(x) d\lambda(x) - \left(\frac{(\int_A x f(x) d\lambda(x))^2}{\int_A f(x) d\lambda(x)} + \frac{(\int_{M \setminus A} x f(x) d\lambda(x))^2}{\int_{M \setminus A} f(x) d\lambda(x)} \right)
\end{aligned}$$

Statt $A \mapsto Q(A)$ zu minimieren, kann man auch, völlig analog zur 2-Niveaupassung (vgl. Bemerkung 3.5.1, S. 92), die Funktion

$$\Psi_f^{\text{quant}}(A) := -Q(A) + \int_M x^2 f(x) d\lambda(x)$$

maximieren. Für alle $A \in \mathfrak{A}'$ mit $\int_A f d\lambda > 0$ und $\int_{M \setminus A} f d\lambda > 0$ ist dann $\Psi_f^{\text{quant}}(A) = \left(\frac{(\int_A x f(x) d\lambda(x))^2}{\int_A f(x) d\lambda(x)} + \frac{(\int_{M \setminus A} x f(x) d\lambda(x))^2}{\int_{M \setminus A} f(x) d\lambda(x)} \right)$. Ist nun $\mathfrak{A}' := \{[0, t] \mid 0 \leq t \leq G\}$, so gilt mit Gleichung (4.2) für alle $n \in \{0, \dots, G\}$:

$$\Psi_{h_g}^{\text{quant}}([0, n]) = \Psi_g \left(\biguplus_{0 \leq i \leq n} P_i \right)$$

Also ist eine 2-Niveaupassung eines Grauwertbildes bezüglich des Mengenmodell der grauwertbenachbarten Pixelmengen nichts anderes als eine Quantisierung mit 2 Stufen bezüglich einer Zufallsvariablen, deren Dichte durch das Histogramm h_g des Bildes gegeben ist.

4.5.4 Ein Algorithmus zur Bestimmung von k Schwellwerten

Obwohl eine 2-Niveaupassung mit konstanten Funktionen bezüglich des Mengenmodells der grauwertbenachbarten Pixelmengen keine anderen Schwellwerte liefert als die Varianzanalyse von Otsu, es also bereits eine Anzahl von Algorithmen in diesem Bereich gibt, soll an dieser Stelle ein Algorithmus angegeben werden, der ein pixelbasiertes lokales Maximum einer k -Niveaupassung bezüglich des Mengenmodells der grauwertbenachbarten Pixelmengen liefert. Es soll nicht zuletzt eine Lücke geschlossen werden, die beim Verfahren von Otsu im Originalartikel [Ots79] offen gelassen wird: die Erweiterung des Verfahrens

auf mehr als einen Schwellwert (vgl. Abschnitt 2.3.1.2). Im Originalartikel schlägt Otsu zwar eine Erweiterung des Verfahrens vor, bleibt allerdings bei der Forderung das globale Optimum der Varianzanalyse mit $k - 1$ Schwellwerten zu berechnen. Dies ist allerdings mit einem erheblichen Laufzeitaufwand verbunden, der durch die Beschränkung auf die Ermittlung lokaler Optima in der Praxis ohne größere Einbußen deutlich gedrückt werden kann.

Definition 4.5.3

Mit $F_{h_g} : \{0, \dots, G\}^2 \rightarrow \mathbb{R}$ definiert durch

$$F_{h_g}(a, b) := \begin{cases} \frac{\left(\sum_{j=a}^b h_g(j)j\right)^2}{\sum_{j=a}^b h_g(j)} & , \text{ falls } a < b \text{ und } \sum_{j=a}^b h_g(j) > 0 \\ 0 & , \text{ sonst} \end{cases}$$

sei $F_{h_g}^{(2)} : \{0, \dots, G\}^3 \rightarrow \mathbb{R}$ definiert durch

$$F_{h_g}^{(2)}(l, u, t) := \begin{cases} F_{h_g}(l, t) + F_{h_g}(t + 1, u) & , \text{ falls } l \leq t < u \\ -\infty & , \text{ sonst} \end{cases}$$

Algorithmus 10 (2-Niveauquantisierung) Berechnung einer optimalen Quantisierung für einen Ausschnitt eines Histogramms h_g

Input: Histogramm h_g eines diskreten Pixelbildes g ,
 l, u mit $l < u$, $l, u \in \{0, \dots, G + 1\}$

Output: Schwellwert t_0 mit $F_{h_g}^{(2)}(l, u, t_0) = \max_{n \in \{u, \dots, l\}} F_{h_g}^{(2)}(l, u, t)$

```

1: max ← -∞
2: t0 ← -1
3: for t = l + 1 bis t = u - 1 do
4:   if F2(l, u, t) > max then
5:     max ← Fhg(2)(l, u, t)
6:     t0 ← n
7:   end if
8: end for
9: return t0

```

Satz 4.5.4

Sei für ein gegebenes diskretes Pixelbild g definiert auf M mit Wertebereich in $\{0, \dots, G\}$ die Menge P_i definiert als $P_i := \{p \in \mathbb{P}(M) \mid g(p) = i\}$. Ist nun t_0 das Ergebnis des Algorithmus 10 (2-Niveauquantisierung), so gilt:

$\{\biguplus_{i=l}^{t_0} P_i, \biguplus_{i=t_0+1}^u P_i\}$ ist das pixelbasierte globale Maximum einer 2-Niveaupassung mit konstanten Funktionen des diskreten Pixelbildes $g|_{\bigcup_{i=l}^u P_i}$ bezüglich des Mengensystems der grauwertbenachbarten Pixelmengen \mathfrak{G} .

Beweis:

folgt mit Gleichung (4.2) auf Seite 182, Definition 4.5.3 und Bemerkung 4.3.7

Bemerkung 4.5.5

Man beachte, daß im Algorithmus 10 (**2-Niveauquantisierung**) das globale Maximum einer eingeschränkten 2-Niveaupassung berechnet wird. An dieser Stelle könnte man sich auch auf die Berechnung eines lokalen Maximums beschränken, wie es im allgemeinen 2-Niveaupassungsalgorithmus 8 (**2-Niveaupassung**) gemacht wird. Dies führt allerdings bei ungeglätteten Histogrammen oft zu nicht-signifikanten lokalen Maxima der Zielfunktion der 2-Niveaupassung.

Bemerkung 4.5.6

Ist $\mathbf{A} \in S_k(\mathfrak{G})$, so ist \mathbf{A} der Form $\{\biguplus_{0 \leq i < t_1} P_i, \biguplus_{t_1 \leq i < t_2} P_i, \dots, \biguplus_{t_{k-1} \leq i < t_k} P_i\}$. Wobei $t_0 := 0$ und $t_k := G + 1$ festgelegt ist. Daher läßt sich jedes $\mathbf{A} \in S_k(\mathfrak{G})$ analog zum Abschnitt 4.4.1 durch einen Vektor (t_0, \dots, t_k) mit $0 = t_0 \leq t_1 \leq \dots \leq t_{k-1} \leq t_k = G$ kodieren.

Leider kodiert jedoch nicht jede Verschiebung eines Schwellwertes $t_i \in \{t_1, \dots, t_{k-1}\}$ eine Segmentierung in $\mathfrak{U}_p(\mathbf{A})$, da die Fläche des minimalen Zuwachses eines jeden Segments $A \in \mathbf{A}$ abhängig davon ist, wieviele Pixel einen gleichen, benachbarten Grauwert haben. Umgekehrt ist es jedoch offensichtlich, daß jedes Element aus $\mathfrak{U}_p(\mathbf{A})$ nur durch die Verschiebung eines einzigen Schwellwertes kodiert sein muß, da man nur so die minimale Veränderung einer bestehenden Segmentierung erhält. Daher ist $S_k(\mathfrak{G})$ nach Definition 4.3.4 auf die Zwei-Niveau-Situation reduzierbar.

Satz 4.5.7

Mit der Kodierung der Segmente nach Bemerkung 4.5.6 gilt:

Der Algorithmus 11 (**Histogrammanalyse**) (S. 187) berechnet ein pixelbasiertes lokales Maximum einer k -Niveaupassung mit konstanten Funktionen eines diskreten Pixelbildes $g : M \rightarrow \{0, \dots, G\}$ bezüglich des Mengensystems der grauwertbenachbarten Pixelmengen.

Beweis:

folgt mit Bemerkung 4.5.6 und Satz 4.3.8.

Man beachte, daß im gesamten Algorithmus $0 = t_0 < t_1 < \dots < t_k := G + 1$ gilt. Daher wird für alle Histogramme h_g mit $\forall i \in \{0, \dots, G\} : h_g(i) > 0$ stets eine Segmentierung mit k echten Segmenten kodiert.

□

Algorithmus 11 (Histogrammanalyse) Zerlegung des Histogramms in k Segmente

Input: Histogramm h_g eines diskreten Pixelbildes g ,
 $k \in \mathbb{N}$ mit $k \leq G$

Output: Schwellwerte t_0, \dots, t_k

```
1: Wähle Startzerlegung  $0 = t_0 < \dots < t_k = G + 1$ 
2: continue  $\leftarrow$  TRUE
3: while continue = TRUE do
4:   for  $i = 0$  bis  $i = k - 2$  do
5:     continue  $\leftarrow$  FALSE
6:     old  $\leftarrow t_{i+1}$ 
7:     new  $\leftarrow$  2-Niveauquantisierung( $h_g, t_i, t_{i+2}$ )
8:     if (old  $\neq$  new) then
9:       continue  $\leftarrow$  TRUE
10:       $t_{i+1} \leftarrow$  new
11:    end if
12:  end for
13: end while
14: return  $t_0, \dots, t_k$ 
```

4.5.5 Beispiele

Abschließend sollen an dieser Stelle noch einige Beispielergebnisse des Algorithmus 11 (**Histogrammanalyse**) präsentiert werden. Die Abbildungen 4.17 bis 4.21 haben alle den gleichen Aufbau: Auf der linken Seite ist jeweils das Originalbild zu sehen, in der Mitte ist das Histogramm, das der Algorithmus als Eingabe verwendet, graphisch dargestellt. In diese Darstellung des Histogramms werden dann die durch den Algorithmus ermittelten Schwellwerte als rote Linien eingetragen. Auf der rechten Seite ist das Ergebnisbild, nach Anwendung der errechneten Schwellwerte und Mittelung der Grauwerte über alle Pixel zwischen jeweils zwei adjazenten Schwellwerten, dargestellt. Das heißt, auf der rechten Seite ist für k Schwellwerte t_0, \dots, t_k das Bild

$$\sum_{i=0}^{k-1} g_{\bigcup_{j=t_i}^{t_{i+1}-1} P_j} = \sum_{i=0}^{k-1} \tilde{\gamma}_g \left(\bigcup_{j=t_i}^{t_{i+1}-1} P_j \right) \mathbf{1}_{\bigcup_{j=t_i}^{t_{i+1}-1} P_j}$$

(wobei $t_0 = 0$ und $t_k = G + 1$ gilt) zu sehen (vgl. auch Beispiel 3.5.11 i), S. 96).

In Abbildung 4.17 wird noch einmal auf das Beispiel in Abbildung 2.7 (S. 23) aus dem Abschnitt 2.3.1.1 Bezug genommen. Allerdings werden an dieser Stelle zwei Schwellwerte statt einem verwandt. Das Ergebnis teilt das Bild in drei Bereiche, nämlich in das Bild der Granittranche, das Bild des Schattens und den weißen Hintergrund. Deutlich zu sehen ist in diesem Beispiel, wie auch in dem der Abbildung 4.21, daß bei der Aufnahme das Bild überbelichtet wurde. Daher überragt der höchste Grauwert, der der Farbe Weiß entspricht, im Histogramm alle übrigen Werte um ein vielfaches.

Die Beispiele 4.18 und 4.19 sind Bilder einer Granittranche, die eine unregelmäßige Textur aufweist. Hier wird die Quantisierungseigenschaft des Algorithmus 11 (**Histogrammanalyse**) deutlich: Auch mit nur drei verschiedenen Grauwerten der optimalen Segmentierung, d. h. nach Anwendung von zwei errechneten Schwellwerten, ist die Struktur der Oberfläche noch relativ gut zu erkennen.

Das Beispiel 4.20 zeigt, daß der Algorithmus auch mit bereits quantisierten Bildern umgehen kann: Das Eingabebild selbst besitzt nur 20 Graustufen.

Beispiel 4.21 ist ein Beispiel aus einer industriellen Anwendung: In einem ersten Schritt eines Bildverarbeitungsprozesses soll hier zunächst einmal der Reifen von dem helleren Hintergrund getrennt werden. Die Abbildung 4.21 zeigt das Ergebnis, wenn man dieses Problem mit dem Algorithmus 11 (**Histogrammanalyse**) lösen will: Einige Pixel, die nicht zum Bild des Reifens, sondern zum Bild des Hintergrunds gehören, werden ebenfalls als „dunkler Vordergrund“ eingestuft, während Pixel im Inneren des Bildes des Reifens als „heller Vordergrund“ eingestuft werden.

Dieses Beispiel zeigt, daß an dieser Stelle das Mengenmodell der grauwertbenachbarten Pixelmengen nicht ausreicht.

Der Algorithmus 10 (**2-Niveauquantisierung**) wurde in keinem der Beispielbilder mehr als 15mal aufgerufen. Der Funktionswert der Funktion F_{h_g} mußte nie öfter als 2000mal berechnet werden.



Abbildung 4.17: Beispiel der Schwellwertbestimmung für zwei Schwellwerte mittels des Algorithmus 11 (**Histogrammanalyse**) (vgl. auch Abbildung 2.7, S. 23).

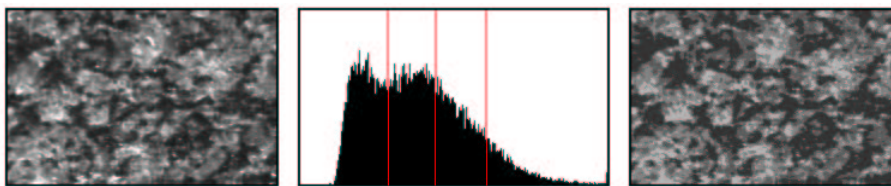


Abbildung 4.18: Beispiel der Schwellwertbestimmung für drei Schwellwerte mittels des Algorithmus 11 (**Histogrammanalyse**). Bei dem Bild handelt es sich um die Oberfläche einer Granittranche.

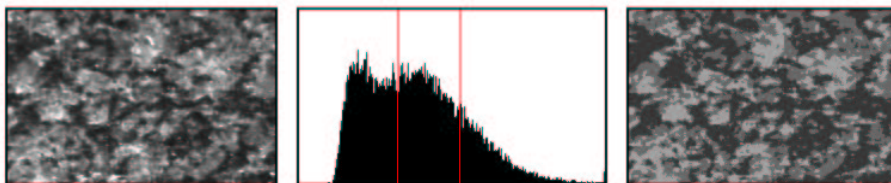


Abbildung 4.19: Schwellwertbestimmung für dasselbe Ursprungsbild wie in Abbildung 4.18 mittels des Algorithmus 11 (**Histogrammanalyse**) für zwei statt für drei Schwellwerte.

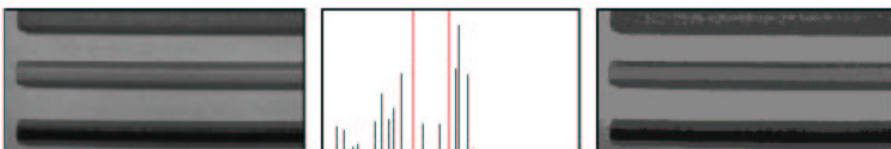


Abbildung 4.20: Beispiel der Schwellwertbestimmung für zwei Schwellwerte eines Bildes mit nur 20 Graustufen mittels des Algorithmus 11 (**Histogrammanalyse**).

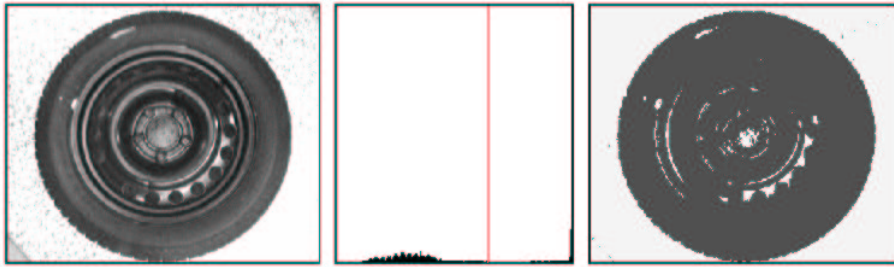


Abbildung 4.21: Beispiel der Schwellwertbestimmung für einen Schwellwert mittels des Algorithmus 11 (**Histogrammanalyse**).

4.5.6 Übertragung auf Farbbilder

Sei nun $E := \mathbb{R}^3$ und $g : M \rightarrow \{0, \dots, G\}^3 \subset E$ ein Farbbild im RGB-Farbmodell. V sei dabei wie in Beispiel 3.4.12 iii), also der Vektorraum der kanalweise konstanten Funktionen von M nach \mathbb{R}^3 mit Basis (e_1, e_2, e_3) , wobei $e_1 := (1, 0, 0)$, $e_2 := (0, 1, 0)$ und $e_3 := (0, 0, 1)$ sei. Das bedeutet, daß ein diskretes Farbpixelbild durch ein Farbbild mit k Farben approximiert werden soll. Betrachtet wird nun die k -Niveaupassung bezüglich des pixelbasierten Mengensystems $S_k(\mathfrak{A}) = \{\{A_1, \dots, A_l\} \mid A_i \subset \mathcal{P}(M), 1 \leq i \leq l \leq k, \biguplus_{i=1}^l A_i = M\}$ aller Pixelmengen. Dabei sei die Familie der Skalarprodukte von Standardskalarprodukt des \mathbb{R}^3 induziert wie in Beispiel 3.4.6.

Analog zum Abschnitt 4.5.1 kann man folgern, daß es genügt, nur Vereinigungen von Mengen der Form $P_c := \{g = c\}$ für alle Farbwerte c im Wertebereich von g zu betrachten, und daß die Segmente einer optimalen Segmentierung in diesem Kontext *farbwertverbunden* sein müssen. Zwei Mengen P_{c_1}, P_{c_2} heißen dabei farbwertverbunden, wenn kein diskreter Farbwert $c \in g(M) \setminus \{c_1, c_2\}$ existiert mit $\|c_1 - c\|_2 < \|c_1 - c_2\|_2$ oder $\|c_2 - c\|_2 < \|c_1 - c_2\|_2$.

Allerdings lassen sich die Farbwerte in $\{0, \dots, G\}^3$ nicht so ordnen, daß man eine zu \mathfrak{G} vergleichbare effektive Kodierung der Segmentierungen bezüglich dieses Mengensystems durch aufsteigende Folgen angeben kann.

Daher läßt sich im farbigen Fall der abstrakte Algorithmus 8 (**2-Niveaupassung**) nicht so einfach darstellen, wie der Algorithmus 10 (**2-Niveauquantisierung**) im grauwertigen Fall. Analog zum Abschnitt 4.5.3 kann man jedoch zeigen, daß eine Segmentierung bezüglich des Systems der farbwertverbundenen Mengen einer Vektorquantisierung mit dem Farbhistogramm als Dichte entspricht. Dabei ist ein Farbhistogramm in diesem Zusammenhang die Funktion $\text{ch}_g : \mathbb{N}_0^3 \rightarrow \mathbb{N}_0$, $\text{ch}_g(r, g, b) = |g^{-1}((r, g, b))|$.

4.6 Das Mengenmodell der pixelzusammenhängenden Mengen

In den meisten Anwendungen der vektoriellen Mehrniveaupassung in der Bildsegmentierung ist der Definitionsbereich des zu segmentierenden Bildes zweidimensional. Das Mengensystem der zusammenhängenden Mengen in einer zwei- oder höherdimensionalen Obermenge ist jedoch nicht abgeschlossen bezüglich der Metrik \mathfrak{d} , wie Bemerkung 4.6.1 zeigt. Das heißt, daß bei diesem Mengensystem die Voraussetzungen für die Existenz eines globalen Maximums einer k -Niveaupassung nicht gegeben sind, weil die Eigenschaften eines System erlaubter Mengen gemäß Definition 3.2.10 (S. 74) nicht erfüllt sind.

Beschränkt man sich jedoch auf pixelzusammenhängenden Mengen, so erhält man ein endliches Mengensystem, auf dem die Zielfunktion Ψ der vektoriellen Mehrniveaupassung ihr globales Maximum annimmt.

Dieses Unterkapitel ist wie folgt aufgebaut:

Zunächst einmal wird das Mengenmodell der pixelzusammenhängenden Mengen definiert. Im zweiten Abschnitt wird dann die pixelbasierte Umgebung einer Segmentierung durch pixelzusammenhängende Mengen genauer betrachtet. Es wird sich dabei herausstellen, daß jedes Element der pixelbasierten Umgebung durch den Austausch eines einzelnen Pixels zwischen zwei Segmenten der aktuellen Segmentierung charakterisiert ist. Der dritte Abschnitt beschäftigt sich mit der Frage der algorithmischen Überprüfung des Pixelzusammenhangs. Der vierte Abschnitt liefert Anmerkungen zur Segmentierung von Farbbildern innerhalb der vektoriellen Mehrniveaupassung. Im fünften Abschnitt wird auf das wichtige Thema der effizienten Initialisierung eines Algorithmus zur Segmentierung bezüglich des Mengenmodells der pixelzusammenhängenden Mengen eingegangen. Im Mittelpunkt dieses Unterkapitels steht allerdings der im sechsten Abschnitt präsentierte Algorithmus zur Segmentierung von Farbbildern bezüglich des Mengenmodells der pixelzusammenhängenden Mengen. Ein Abschnitt mit zahlreichen Beispielen, einige davon aus industriellen Anwendungen, und ein Abschnitt über einfach pixelzusammenhängende Mengen schließen dieses Unterkapitel ab.

Im folgenden sei stets $\mathbb{P}(M) \neq \emptyset$.

4.6.1 Definition des Mengenmodells der pixelzusammenhängenden Mengen

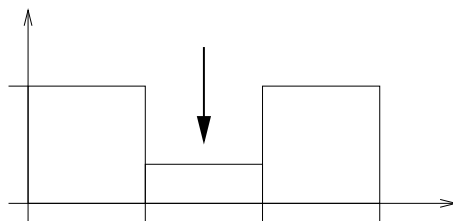
Bemerkung 4.6.1

Sei $M := [0, 3] \times [0, 1]$. Die Äquivalenzrelation \sim sei definiert wie in Definition 3.2.1 (S. 70). Das Mengensystem \mathfrak{A} sei definiert durch

$$\mathfrak{A} := \{A \in \mathcal{B}(M) / \sim \mid A \text{ ist zusammenhängend}\}$$

Dann ist \mathfrak{A} nicht abgeschlossen.

Denn: Sei $A_n := ([0, 1] \times [0, 1]) \cup ([1, 2] \times [0, 1/n]) \cup ([2, 3] \times [0, 1])$. Dann ist für alle $n \in \mathbb{N}$ die Menge $A_n \in \mathfrak{A}$. Aber es ist $\lim_{n \rightarrow \infty} A_n = ([0, 1] \times [0, 1]) \cup ([2, 3] \times [0, 1]) \notin \mathfrak{A}$.

Abbildung 4.22: Skizze der Mengen A_n .**Bemerkung 4.6.2** (N_*)

Trotz der ernüchternden Erkenntnis der Bemerkung 4.6.1 ist das Mengensystem der pixelzusammenhängenden Mengen für viele Anwendungen sehr wichtig. Allerdings ist es nicht möglich, die pixelzusammenhängenden Mengen als endliche Untermenge der zusammenhängenden Mengen zu sehen. Vielmehr muß man sich ganz auf die pixelbasierten Mengen und die pixelbasierte Definition des Begriffes „Pixelzusammenhang“ aus Definition 1.3.7 (S. 7) zurückziehen.

An dieser Stelle muß man sich für eine Nachbarschaftsdefinition entscheiden: Im folgenden stehe N_* entweder immer für N_4 oder immer für N_8 . Die Vor- und Nachteile der 8-er gegenüber der 4-er Nachbarschaft sowie die Schwierigkeiten der Übertragung des Begriffes „zusammenhängend“ auf Pixelmengen soll an dieser Stelle nicht behandelt werden. Vielmehr sei dazu z. B. auf [Vos93] verwiesen. Offensichtlich ist, daß die 8-er Nachbarschaft eine größere Flexibilität bereitstellt als die 4-er Nachbarschaft, was je nach Problem Vor- und Nachteile haben kann. Auch existieren bei der 4-er Nachbarschaft zusammenhängende Mengen von Pixeln, die nicht pixelzusammenhängend sind (vgl. Beispiel 4.6.26).

Es ist keineswegs so, daß man sich beim Nachbarschaftsbegriff auf die 4-er und 8-er Nachbarschaft beschränken muß. Ebenfalls in [Vos93] findet man einen abstrakten Ansatz zum Nachbarschaftsbegriff. Im Rahmen dieser Arbeit soll jedoch nur auf die 4-er und 8-er Nachbarschaft eingegangen werden. Es ist allerdings im folgenden durchaus möglich N_* durch eine andere Nachbarschaftsrelation zu ersetzen. Es hat auf die Definitionen und Sätze, bis auf Satz 4.6.6, keine Auswirkungen. Bei den Beispielen wurde allerdings aus implementierungstechnischen Gründen auf eine weitere Differenzierung verzichtet.

Definition 4.6.3 (System der pixelzusammenhängenden Mengen)

$\mathfrak{Z} \subset \mathcal{B}(M)$ sei ein System von Mengen, das die Bedingungen der Definition 3.2.10 (S. 74) für ein erlaubtes Mengensystem erfüllt, und für das gilt:

$$\mathfrak{Z} \cap \mathcal{P}(M) = \{A \in \mathcal{P}(M) \mid A \text{ ist pixelzusammenhängend bezüglich der } N_*\text{-Nachbarschaft}\}$$

$S_k(\mathfrak{Z})$ ist durch diese Vorschrift eindeutig bestimmt und heiße *das System der pixelzusammenhängenden Mengen*

4.6.2 Die pixelbasierte Umgebung einer Segmentierung mit pixelzusammenhängenden Mengen

Sei eine Segmentierung $\mathbf{A} \in S_k(\mathfrak{Z})$ gegeben. Für die Algorithmen der vektoriiellen Mehrniveaupassung ist es nötig, möglichst schnell Segmentierungen aus der pixelbasierten Umgebung einer bestehenden pixelbasierten Segmentierung zu berechnen (vgl. Algorithmus 6 (**k-Niveaupassung**), Zeile 5). Im Falle der Segmentierung bezüglich des Mengensystems der pixelzusammenhängenden Mengen unterscheidet sich eine Segmentierung \mathbf{A}' aus $\mathfrak{U}_p(\mathbf{A})$ von \mathbf{A} dadurch, daß genau ein Pixel p eines Segments $A \in \mathbf{A}$ aus dem Segment entfernt und einem anderem Segment hinzugefügt wurde:

$$\forall \mathbf{A}' \in \mathfrak{U}_p(\mathbf{A}) : \exists A_i, A_j \in \mathbf{A}, p \subseteq A_i : (\mathbf{A} \setminus \{A_i, A_j\}) \cup \{A_i \setminus p, A_j \cup p\} = \mathbf{A}'$$

Denn offensichtlich ist die Verschiebung eines einzelnen Pixels die Operation, die die kleinste Veränderung von \mathbf{A} in der Metrik der Segmentierungen \mathfrak{d} ausmacht. Daher müssen nach der Definition der pixelbasierten Umgebung 4.2.13 alle Segmentierungen aus der pixelbasierten Umgebung einer Segmentierung aus pixelzusammenhängenden Mengen so zu erhalten sein. Damit ist $S_k(\mathfrak{Z})$ insbesondere auf die Zwei-Niveau-Situation reduzierbar.

Es ist nur ein solcher Austausch von Pixeln erlaubt, der pixelzusammenhängende Pixelmengen in pixelzusammenhängende Pixelmengen überführt. Wird ein Pixel p von Segment A_i zum Segment A_j einer Segmentierung durch pixelzusammenhängende Mengen verschoben, so muß $N_*(p) \cap \mathbb{P}(A_j) \neq \emptyset$ gelten. Daher kann es sich bei dem zu verschiebenden Pixel p nur um ein Randpixel von A_i handeln.

Definition 4.6.4 (Randpixel)

Sei $\mathbf{A} \in S_k(\mathfrak{Z})$ eine pixelbasierte Segmentierung. Dann heie ein Pixel p *Randpixel des Segments* $A \in \mathbf{A}$ *bezüglich der 4-er Nachbarschaft bzw. 8-er Nachbarschaft*, falls $p \subseteq A$ gilt und

$\exists p' \in N_4(p) : p' \not\subseteq A$ (4-er Nachbarschaft) bzw.

$\exists p' \in N_8(p) : p' \not\subseteq A$ (8-er Nachbarschaft).

$Rn(A)$ sei Menge aller Randpixel des Segments A .

Bemerkung 4.6.5

Man beachte, daß aus der Definition 4.6.4 und der Definition 1.3.5 (S. 7) der 4-er bzw. 8-er Nachbarschaft folgt, daß Pixel, die am Rand des Bildes liegen, nicht automatisch Randpixel eines Segments sein müssen. Vielmehr müssen sie an der Berührungslinie eines Segments zu einem anderem Segment liegen. Daher existiert für alle Randpixel p eines pixelbasierten Segments $A \in \mathbf{A}$ ein weiteres Segment $A' \in \mathbf{A}$ mit $N_*(p) \cap \mathbb{P}(A') \neq \emptyset$. Es ist daher die Menge $A' \cup p$ pixelzusammenhängend.

Lemma 4.6.6

Sei $\mathbf{A} \in S_k(\mathfrak{Z})$. Dann existiert für alle $A \in \mathbf{A}$ mit $\lambda(A) > 0$ ein Pixel $p_A \in Rn(A)$ derart, daß $A \setminus p_A$ pixelzusammenhängend ist.

Dies gilt sowohl für die 4-er als auch für die 8-er Nachbarschaft.

Beweis:

vgl. [Ros70], Theorem 3 und Corollary, S. 152

Lemma 4.6.7

Sei $\mathbf{A} \in S_k(\mathfrak{Z})$ mit $O \notin \mathbf{A}$. Dann gilt: Ist \mathbf{A}' eine Segmentierung aus $\mathfrak{U}_p(\mathbf{A})$, so existieren zwei Segmente A_i und A_j aus \mathbf{A} , sowie ein Pixel aus $Rn(A_i)$, mit

$$(\mathbf{A} \setminus \{A_i, A_j\}) \cup \{A_i \setminus p, A_j \cup p\} = \mathbf{A}'$$

Daher ist insbesondere $S_k(\mathfrak{Z})$ auf die Zwei-Niveau-Situation reduzierbar.

Beweis:

Nach Lemma 4.6.6 existiert für alle $A_i \in \mathbf{A}$ mit $A_i \neq O$ ein Pixel $p_A \in Rn(A_i)$ derart, daß $A_i \setminus p_A$ ist pixelzusammenhängend ist.

Nach Definition 1.3.7 (vgl. auch Bemerkung 4.6.5) folgt aber, daß ein Segment $A_j \neq A_i$ in \mathbf{A} mit $N_*(p_A) \cap \mathbb{P}(A_j) \neq \emptyset$ existieren muß. Daher ist $A_j \cup p_A$ ebenfalls pixelzusammenhängend. Da der Austausch eines Pixels die kleinste Veränderung einer pixelbasierten Segmentierung sein muß, folgt die Behauptung. □

Bemerkung 4.6.8

Wie man sich leicht überlegen kann, lassen sich nicht alle Randpixel eines pixelbasierten Segments entfernen, ohne den Pixelzusammenhang des Segments zu verletzen. Daher muß man zwischen Randpixeln und solchen Pixeln unterscheiden, deren Entfernung aus dem Segment den Pixelzusammenhang nicht zerstört. Aus diesem Grunde wird in der folgenden Definition 4.6.9 die Menge der „Kandidatenpixel“ eingeführt.

Definition 4.6.9 (Kandidatenpixel)

Für eine pixelzusammenhängende Menge A heiße die Menge

$$\text{Kand}(A) := \{p \in \text{Rn}(A) \mid A \setminus p \text{ ist pixelzusammenhängend}\}$$

die Menge der *Kandidatenpixel der pixelbasierten Menge* A .

Nach Lemma 4.6.6 ist $\text{Kand}(A)$ für alle nichtleeren pixelzusammenhängenden Mengen nicht leer. Für eine pixelbasierte Segmentierung $\mathbf{A} \in S_k(\mathfrak{Z})$ sei

$$\text{Kand}(\mathbf{A}) := \bigcup_{A \in \mathbf{A}} \text{Kand}(A)$$

die Menge der *Kandidatenpixel der Segmentierung* \mathbf{A} . Da im gesamten Unterkapitel $\mathbb{P}(M) \neq \emptyset$ gilt, ist $\text{Kand}(\mathbf{A})$ nicht leer.

Bemerkung 4.6.10

Sei $\mathbf{A} \in S_k(\mathfrak{Z})$ mit $O \notin \mathbf{A}$. Nach Lemma 4.6.7 ist jede pixelbasierte Segmentierung aus der pixelbasierten Umgebung durch den Austausch eines Pixels p von einem Segment zu einem benachbarten Segment in der Segmentierung kodiert. Für jedes Kandidatenpixel $p \in \text{Kand}(\mathbf{A})$ gilt:

- i. $\exists A \in \mathbf{A}, p \subset A, A \setminus p$ ist pixelzusammenhängend, und
- ii. $\exists A' \in \mathbf{A} : p \not\subset A', \mathbb{P}(A') \cap N_*(p) \neq \emptyset$, d. h. $A' \cup p$ ist pixelzusammenhängend.

Es gilt sogar, daß die Menge der Kandidatenpixel maximal in dieser Eigenschaft ist. Daher ist $\text{Kand}(\mathbf{A})$ genau die Menge von Pixeln, deren Austausch zwischen zwei Segmenten die pixelbasierte Umgebung der Segmentierung \mathbf{A} bezüglich des Mengenmodells der pixelzusammenhängenden Mengen kodiert.

4.6.3 Algorithmische Überprüfung von Pixelzusammenhang

Um zu entscheiden, ob ein Pixel p des Randes eines pixelzusammenhängenden Segments ein Kandidatenpixel nach Definition 4.6.9 ist, ist es nötig, zu prüfen, ob die Menge der Pixel des Segments nach Entfernen des Pixels p immer noch pixelzusammenhängend ist. Die Beantwortung der Frage, ob eine Pixelmenge $P \setminus \{p\}$ pixelzusammenhängend ist, fällt für Pixel $p \in P$ algorithmisch nicht leicht aus. Denn wie die Definition 1.3.7 der pixelzusammenhängenden Pixelmengen zeigt, ist die Zusammenhangseigenschaft keine lokale Eigenschaft der Pixelmenge. Es sind vielmehr alle minimalen Pfade von Pixel zu Pixel in $P \setminus \{p\}$ zu überprüfen, in denen p enthalten war. Sei $(q_i)_{1 \leq i \leq n}$ ein Pfad von $p_1 \in P$ nach $p_2 \in P$ in P mit $q_j = p$ für ein j mit $1 < j < n$. Dann gilt aber $q_{j-1} \in N_*(p) \cap P$ und $q_{j+1} \in N_*(p) \cap P$. Daher läßt sich das Problem darauf reduzieren, für alle $p_1, p_2 \in (N_*(p) \cap P)$ einen Pfad in $P \setminus \{p\}$ zu finden, der p_1 und p_2 verbindet. Obwohl das Problem darauf eingeschränkt wurde, Pfade zu suchen, die Pixel einer lokalen Pixelmenge verbinden, ist es immer noch nicht lokal lösbar, da die Pfade nicht lokal begrenzt sein müssen, wie die Abbildung 4.23 zeigt. Die Abbildung skizziert eine Pixelmenge, aus der man ein beliebiges Pixel p entfernen kann, ohne den Pixelzusammenhang der Menge zu verletzen. Bei der Überprüfung des Pixelzusammenhangs der Restmenge muß in diesem Beispiel (bei 4-er Nachbarschaft) für jedes Pixel aus der Nachbarschaft des entfernten Pixels ein Pfad über alle anderen Pixel der Restmenge konstruiert werden.

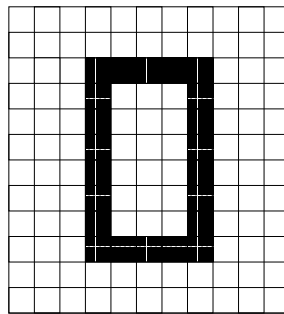


Abbildung 4.23: Beispielpixelmenge, bei der die Überprüfung des Pixelzusammenhangs nicht lokal geschehen kann.

Für praktische Anwendungen bedeutet dies einen erheblichen Laufzeitaufwand, da jede Überprüfung des Pixelzusammenhangs einer Pixelmenge $P \setminus \{p\}$ im schlechtesten Fall mit $O(|P|)$ abgeschätzt werden muß. In den meisten Fällen reicht es jedoch aus, nur Pfade bestehend aus Pixeln der Nachbarschaft des Pixels, das entfernt werden soll, zu konstruieren, wie die Abbildung 4.24 zeigt.

Bei den realen Implementierungen von Mehrniveaupassungen bezüglich des Mengenmodells der pixelzusammenhängenden Mengen wurde daher im Rahmen dieser Arbeit aus Laufzeitgründen auf die globale Überprüfung des Pixelzusammenhangs der Pixelmenge $P \setminus \{p\}$ verzichtet. Das hat zur Folge, daß die betrachteten Segmentierungen zwar aus pixelzusammenhängenden Segmenten bestehen, aber eben nicht alle möglichen pixelzusammenhängenden Segmentierungen betrachtet werden.

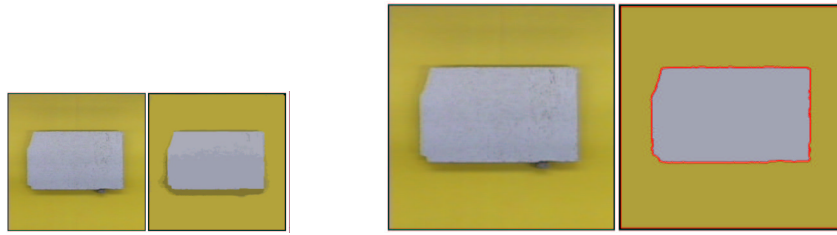


Abbildung 4.24: Beispielpixelmengen, bei der die Überprüfung des Pixelzusammenhangs lokal geschehen kann. Im ersten Fall kann ein Pfad durch alle Randpixel des zu entfernenden Pixels gelegt werden, im zweiten Fall nicht. Das heller eingefärbte Pixel p in der rechten Pixelmenge kann nicht entfernt werden, ohne den Pixelzusammenhang zu verletzen. Im Beispiel gilt die 4-er Nachbarschaft.

Definition 4.6.11 ((Erweiterte) lokale Pixelzusammenhangsüberprüfung)

Sei $P \in \mathcal{P}(M)$ eine pixelzusammenhängende Pixelmenge aus der das Pixel p entfernt werden soll.

Die oben beschriebene Methode zur Überprüfung des Pixelzusammenhangs von P nur den Pixelzusammenhang der Menge $(N_*(p) \cap P)$ zu überprüfen heie *lokale Pixelzusammenhangsüberprüfung*. Die Überprüfung des Pixelzusammenhangs der Menge

$$\bigcup_{p' \in N_*(p)} N_*(p') \cap P \setminus \{p\}$$

heie *erweiterte lokale Pixelzusammenhangsüberprüfung*.

Bemerkung 4.6.12

Ist die (erweiterte) lokale Pixelzusammenhangsüberprüfung erfolgreich, so ist $P \setminus \{p\}$ pixelzusammenhngend. Die Umkehrung gilt nicht.

In praktischen Anwendungen hat sich im Falle der 8-er Nachbarschaft die lokale Pixelzusammenhangsüberprüfung als ausreichend herausgestellt. Im Falle der 4-er Nachbarschaft sollte auf die erweiterte Pixelzusammenhangsüberprüfung zurckgegriffen werden.

4.6.4 Anmerkungen zur Segmentierung von Farbbildern

Eine groe Klasse von Anwendungen der vektoriiellen Mehrniveaupassung in der Bildsegmentierung ist die Segmentierung von Farbbildern. Natrlich lassen sich alle Ergebnisse innerhalb der Segmentierung von Farbbildern auch auf Grauwertbilder bertragen. Alle Beispiele des Abschnitts 4.6 bestehen allerdings aus Farbbildern. Bei Farbbildern ist der Wertebereich eines Bildes f eine Teilmenge des \mathbb{R}^3 (vgl. Abschnitt 1.3.2). Die bertragung der Ergebnisse und Algorithmen auf Bilder mit einem Wertebereich im \mathbb{R}^d , mit $d \in \mathbb{N}$ beliebig, erfolgt kanonisch, wird aber an dieser Stelle nicht durchgefhrt.

4.6.4.1 Anmerkungen zum Approximationsproblem

Die vektorielle Mehrniveaupassung stellt Methoden und Algorithmen zur Verfügung, um die Funktion

$$\hat{\Phi} : S_k(\mathfrak{Z}) \times V^k \rightarrow \mathbb{R}_+$$

$$(\{B_1, \dots, B_k\}, v_1, \dots, v_k) \mapsto \left\| g - \sum_{i=1}^k v_i \mathbf{1}_{B_i} \right\|_M^2$$

zu minimieren. Der Fehler wird dabei in der vom Skalarprodukt induzierten Norm $\|\cdot\|_M = \sqrt{\langle \cdot, \cdot \rangle_M}$ gemessen, die in allen Anwendungen dieser Arbeit durch das Skalarprodukt des Zielraums der Funktion g bestimmt ist (vgl. Beispiel 3.4.6, S. 86). In diesem Abschnitt ist der Zielraum der \mathbb{R}^3 und das Skalarprodukt das Standardskalarprodukt des \mathbb{R}^3 . Das heißt, daß Abstände hier in der euklidischen Metrik gemessen werden.

Im RGB-Farbmodell ist die euklidische Metrik als Diskriminanzmaß von Farben nicht geeignet (vgl. Abschnitt 1.3.2.2). Es ist jedoch möglich, den RGB-Farbwürfel so linear zu transformieren, daß die euklidische Metrik auf dem Ergebnis der Transformation bessere Ergebnisse zur Unterscheidung von Farben liefert (vgl. z. B. [PB99], [Haf98] oder Abschnitt 1.3.2.2).

Auf der anderen Seite kann man auch über das Skalarprodukt einzelne Kanäle stärker gewichten. Die Idee zur Approximation von mehrkanaligen Bildern nicht unbedingt die euklidische Norm heranzuziehen findet sich schon in [KLM94] mit der Definition einer gewichteten Norm $\|\cdot\|_{pond}$. Diese kann man nutzen, um die Wichtigkeit der Kanäle des Bildes zu berücksichtigen. Die abstrakt gehaltene Definition 3.4.4 (S. 85) der Familie der erlaubten Skalarprodukte zur Mehrniveaupassung erlaubt eine solche Gewichtung auch in der Theorie der vektoriellen Mehrniveaupassung.

4.6.4.2 Anmerkungen zur Wahl der Basis \mathcal{E} auf mehrkanaligen Bildern

Ein Farbbild wird vielfach auch als ein mehrkanaliges Bild gesehen, das aus 3 Kanälen, z. B. dem Rot-, Grün- und Blaukanal, besteht (vgl. Abschnitt 1.3.2.1.1). Da die Pixelwerte des Bildes auf den einzelnen Kanälen unterschiedlich sein müssen, um Farbwerte, die nicht Grau sind, kodieren zu können, bietet es sich an, die Basis \mathcal{E} wie in Beispiel 3.5.11 iii) so zu wählen, daß jede Basisfunktion ihren Wertebereich nur in einem Kanal des Bildes hat.

Eine typische Basis für eine Mehrniveaupassung eines RGB-Bildes $f : M \rightarrow \mathbb{R}_+^3$ wäre $\mathcal{E} = (e_1, e_2, e_3, e_4, e_5, e_6)$ mit

$$e_1(x) := (1, 0, 0), \quad e_2(x) := (0, 1, 0), \quad e_3(x) := (0, 0, 1),$$

$$e_4(x) := (x, 0, 0), \quad e_5(x) := (0, x, 0), \quad e_6(x) := (0, 0, x)$$

Allgemein sollte die Basis \mathcal{E} der Klasse der erlaubten Funktionen auf RGB-Bildern also die folgende Definition erfüllen:

Definition 4.6.13 (Erlaubte Basis zur Farbbildsegmentierung)

Erfüllt eine Basis \mathcal{E} eines Vektorraums $V \subset L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^3))$ von erlaubten Funktionen die Bedingungen

- i. \mathcal{E} ist der Form $(e_{1_r}, e_{1_g}, e_{1_b}, \dots, e_{n_r}, e_{n_g}, e_{n_b})$,
- ii. $\forall i \in \{1, \dots, n\} : e_{i_r}(x)_1 = e_{i_g}(x)_2 = e_{i_b}(x)_3$ für alle $x \in M$, und
- iii. $\forall i \in \{1, \dots, n\} : e_{i_r}(x)_2 = e_{i_r}(x)_3 = 0, e_{i_g}(x)_1 = e_{i_g}(x)_3 = 0, e_{i_b}(x)_1 = e_{i_b}(x)_2 = 0$
für alle $x \in M$,

so heie \mathcal{E} eine *erlaubte Basis zur Farbbildsegmentierung*.

4.6.5 Initialisierung

Wie das Beispiel 4.4.1 (S. 172) zeigt, ist eine gute Initialisierung, d. h. eine gute Auswahl einer Startsegmentierung, im Algorithmus 6 (**k-Niveaupassung**) von zentraler Bedeutung. Eine Möglichkeit diese zu erhalten, bietet das folgende Vorgehen:

- i. Berechne ein vorsegmentiertes Bild g_q , entweder als vektorquantisiertes Bild (vgl. Abschnitt 4.5.6 oder [GG92]) oder als Ergebnis des Algorithmus 13 (**Initiales Farbschwellwertverfahren**).
- ii. Die Menge $\{A_1, \dots, A_l\}$ sei die Menge der Pixelzusammenhangskomponenten in g_q . Mit Lemma 1.3.9 ist dann $\{A_1, \dots, A_l\}$ eine Zerlegung von M in l pixelzusammenhängende Mengen. Ist $l < k$, so ist Schritt i. mit mehr Schwellwerten anzuwenden, oder sich auf l Niveaus zu beschränken.
- iii. Verschmelze mit Algorithmus 16 (**Iterative Niveaverschmelzung**) (S. 226) iterativ jeweils zwei Segmente bis k Segmente übrig bleiben.

Dabei hat sich für Schritt i. das zweite Verfahren (Algorithmus 13 (**Initiales Farbschwellwertverfahren**)) als sehr effizient herausgestellt. Es liefert zwar kein optimales Ergebnis im Sinne einer Vektorquantisierung oder einer k -Niveaupassung bezüglich des Mengenmodells aller Pixelmengen (vgl. Abschnitt 4.5.6), hat sich aber in der Praxis als ausreichend erwiesen. Es hat gegenüber den Ansätzen über eine Vektorquantisierung einen nicht zu unterschätzenden Geschwindigkeitsvorteil. Man kann an dieser Stelle aber auch jedes andere Segmentierungsverfahren als Initialisierung der vektoriellen Mehrniveaupassung bezüglich des Mengensystems der pixelzusammenhängenden Mengen nutzen.

Der Algorithmus 13 (**Initiales Farbschwellwertverfahren**) implementiert separate Histogrammanalysen nach Abschnitt 4.5, wendet die für jeden der drei Farbkanäle einzeln berechneten Schwellwerte auf den jeweiligen Kanal an, und fügt anschließend die segmentierten Kanäle wieder zu einem drei-kanaligen Bild zusammen.

Um den Algorithmus 13 (**Initiales Farbschwellwertverfahren**) effizient formulieren zu können, wird zunächst der Algorithmus 12 (**Schwellwertsegmentierung**) vorgestellt, der ein Bild mit vorgegebenen Schwellwerten segmentiert.

Algorithmus 12 (Schwellwertsegmentierung) Segmentierung eines diskreten Pixelbildes über gegebene Schwellwerte

Input: Diskretes Grauwertpixelbild g ,

Schwellwerte t_0, \dots, t_k mit $0 = t_0 \leq \dots \leq t_k = G + 1$

Output: Schwellwertsegmentiertes diskretes Grauwertpixelbild g_k mit k Graustufen

1: Setze für alle $p \in \mathbb{P}(M)$: $g_k(p) \leftarrow \max\{t \in \{t_0, \dots, t_k\} \mid g(p) < t\}$

2: **return** g_k

Algorithmus 13 (Initiales Farbschwelwertverfahren) Berechnung einer Startsegmentierung durch ein kanalweises Schwellwertverfahren

Input: Diskretes Farbpixelbild $g = (g_{\text{rot}}, g_{\text{grün}}, g_{\text{blau}})$,
 $k \in N$

Output: Diskretes Farbpixelbild g_q mit in k Stufen quantisierten Farbkanälen

- 1: $h_{g_{\text{rot}}}$ sei das Histogramm von g_{rot}
 - 2: $h_{g_{\text{grün}}}$ sei das Histogramm von $g_{\text{grün}}$
 - 3: $h_{g_{\text{blau}}}$ sei das Histogramm von g_{blau}
 - 4: $t_0^{\text{rot}}, \dots, t_k^{\text{rot}} \leftarrow \text{Histogrammanalyse}(h_{g_{\text{rot}}}, k)$
 - 5: $t_0^{\text{grün}}, \dots, t_k^{\text{grün}} \leftarrow \text{Histogrammanalyse}(h_{g_{\text{grün}}}, k)$
 - 6: $t_0^{\text{blau}}, \dots, t_k^{\text{blau}} \leftarrow \text{Histogrammanalyse}(h_{g_{\text{blau}}}, k)$
 - 7: $g_{q_{\text{rot}}} \leftarrow \text{Schwellwertsegmentierung}(g_{\text{rot}}, t_0^{\text{rot}}, \dots, t_k^{\text{rot}})$
 - 8: $g_{q_{\text{grün}}} \leftarrow \text{Schwellwertsegmentierung}(g_{\text{grün}}, t_0^{\text{grün}}, \dots, t_k^{\text{grün}})$
 - 9: $g_{q_{\text{blau}}} \leftarrow \text{Schwellwertsegmentierung}(g_{\text{blau}}, t_0^{\text{blau}}, \dots, t_k^{\text{blau}})$
 - 10: **return** $g_q = (g_{q_{\text{rot}}}, g_{q_{\text{grün}}}, g_{q_{\text{blau}}})$
-

Beispiel 4.6.14

Die Abbildung 4.25 zeigt zwei kleine Granittranchen vor einem gelben Hintergrund. Ebenfalls abgebildet ist für jeden Kanal R, G und B das Histogramm, für das jeweils mittels des Algorithmus 11 (**Histogrammanalyse**) zwei Schwellwerte bzw. ein Schwellwert berechnet werden bzw. wird.

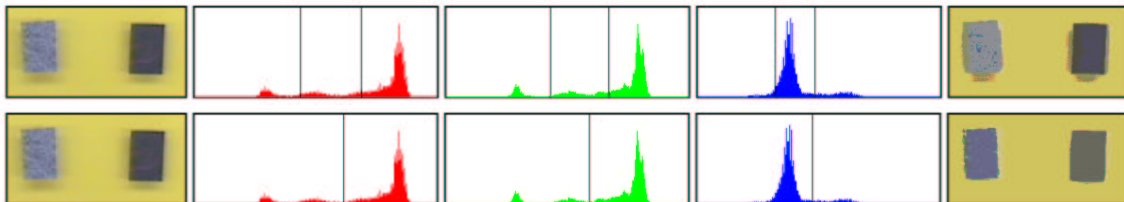


Abbildung 4.25: Beispiel eines Ergebnisses des Algorithmus 13 (**Initiales Farbschwelwertverfahren**). Links ist das Eingabebild g zu sehen, in der Mitte die drei Histogramme $h_{g_{\text{rot}}}$, $h_{g_{\text{grün}}}$ und $h_{g_{\text{blau}}}$ mit den jeweils algorithmisch ermittelten Schwellwerten. Rechts ist dann das Bild g_q nach Schwellwertsegmentierung der einzelnen Farbkanäle zu sehen. In der oberen Reihe wurde jeweils mit zwei Schwellwerten, in der unteren mit einem Schwellwert pro Kanal, segmentiert.

Das Ergebnis der Verschmelzung der einzeln quantisierten Kanäle ist jeweils auf der rechten Seite der Abbildung zu sehen. Das untere Ergebnis mit einem Schwellwert pro Kanal sollte als Initialisierung einer 3-Niveaupassung genügen.

4.6.6 Der Algorithmus

Die Segmentierung bezüglich des Mengenmodells der pixelzusammenhängenden Mengen hat sich in den praktischen Anwendungen der Methoden der Mehrniveaupassung als eine der wichtigsten Applikationen in diesem Kontext erwiesen. Deshalb soll der abstrakte Algorithmus 6 (**k-Niveaupassung**) für diese Anwendung optimal spezialisiert werden. Im Mittelpunkt steht dabei der Schritt 5 des abstrakten Algorithmus: Die effiziente Auswahl einer Segmentierung aus der pixelbasierten Umgebung einer aktuell gültigen Segmentierung.

4.6.6.1 Bemerkungen zur Auswahl einer pixelbasierten Segmentierung aus $\mathfrak{U}_p(\mathbf{A})$

Wie bereits in Abschnitt 4.6.2 festgestellt, ist die Menge $\mathfrak{U}_p(\mathbf{A})$, die Menge aller pixelzusammenhängenden Segmentierungen aus der pixelbasierten Umgebung der pixelzusammenhängenden Segmentierung \mathbf{A} , durch die Menge aller Wechsel von Kandidatenpixeln zu einem benachbarten Segment in \mathbf{A} gegeben. Um daher eine laufzeit-effiziente Spezialisierung des abstrakten Algorithmus 6 (**k-Niveaupassung**) zu formulieren, ist es nötig die Liste der Kandidatenpixel effizient zu verwalten. Das Korollar 3.7.42 motiviert dazu, die Liste der Kandidatenpixel eines jeden pixelbasierten Segments A nach ihrem Abstand in der Metrik $d_{A \setminus p, p}$ für alle $p \in \text{Kand}(A)$ zu sortieren, um auf die besten Kandidatenpixel in konstanter Zeit, d. h. in $\mathcal{O}(1)$, zugreifen zu können.

Für jedes Pixel p aus der Menge der Kandidatenpixel kann man ein Nachbarsegment auszeichnen: das Segment, das nach Bemerkung 3.7.41 die höchste Wertsteigerung beim Wechsel des Pixels zu dem Segment verspricht. Dieses Segment wird in Definition 4.6.16 als „bester Nachbar“ des Pixels p bezeichnet.

Für eine effiziente Verwaltung der Pixel des Bildes sollte für jedes Pixel bekannt sein, zu welchem Segment es gehört.

Definition 4.6.15 ($seg_{\mathbf{A}}$)

Für $\mathbf{A} \in S_k(\mathfrak{Z})$ sei die Abbildung $seg_{\mathbf{A}} : \mathbb{P}(M) \rightarrow \mathfrak{Z} \cap \mathcal{P}(M)$ definiert durch $seg_{\mathbf{A}}(p) := A_p$ mit $A_p \in \mathbf{A}$ und $p \subset A_p$. Da es sich bei \mathbf{A} um eine pixelbasierte Segmentierung handelt, ist diese Definition wohldefiniert.

Definition 4.6.16 (bester Nachbar)

Sei $\mathbf{A} \in S_k(\mathfrak{Z})$. Für $p \in \text{Kand}(\mathbf{A})$ sei die Menge $\text{Neighbor}(p) \subset \mathfrak{Z}$ definiert durch

$$\text{Neighbor}(p) := \{seg_{\mathbf{A}}(q) \mid q \in N_*(p), seg_{\mathbf{A}}(q) \neq seg_{\mathbf{A}}(p)\}$$

Die Menge $\text{BestNeighbor}(p) \subset \mathfrak{Z}$ sei definiert durch

$$\text{BestNeighbor}(p) := \left\{ A \in \text{Neighbor}(p) \mid d_{p,A}(p, A) = \min_{B \in \text{Neighbor}(p)} \{d_{p,B}(p, B)\} \right\}$$

$\text{BestNeighbor}(p)$ heie die Menge der *besten Nachbarn des Kandidatenpixel* p .

Bemerkung 4.6.17

Für ein Pixel p erfüllt die Menge der besten Nachbarn zwei Eigenschaften: Zunächst einmal ist Verbindung von p mit jedem Segment aus dieser Menge erlaubt, d. h. sie führt zu einer erlaubten Segmentierung. Weiter gilt mit Korollar 3.7.37 und Definition 3.7.40, daß diese Verbindung am ehesten zu einer Wertverbesserung der Zielfunktion der k -Niveaupassung $\Psi : S_k(\mathfrak{Z}) \rightarrow \mathbb{R}_+$ führt. Daher genügt es, wenn – wie im abstrakten Algorithmus 6 (**k-Niveaupassung**) – alle möglichen Segmentierungen in der Umgebung einer Segmentierung \mathbf{A} zu prüfen sind, sich für jedes Pixel aus der Menge der Kandidatenpixel auf die besten Nachbarn zu beschränken.

Definition 4.6.18 (kritischer Wert)

Sei $\mathbf{A} \in S_k(\mathfrak{Z})$ und $p \in \text{Kand}(\mathbf{A})$. Da nach Definition 4.6.16 für alle $B \in \text{BestNeighbor}(p)$ der Wert $d_{p,B}(p, B)$ gleich ist, ist die Definition

$$\text{crit}(p) := d_{p, \text{seg}(p) \setminus p}(p, \text{seg}(p) \setminus p) - d_{p,B}(p, B)$$

mit $B \in \text{BestNeighbor}(p)$ wohldefiniert. $\text{crit}(p)$ heie *der kritische Wert* von p .

Bemerkung 4.6.19

Sortiert man alle Kandidatenpixel nach ihrem kritischen Wert $\text{crit}(p)$, so gengt es, lediglich das Pixel mit dem hochsten kritischen Wert darauf zu berprfen, ob $\text{crit}(p) > 0$ gilt. Denn, wenn die aktuelle Segmentierung $\mathbf{A} \in S_k(\mathfrak{Z})$ kein pixelbasiertes lokales Maximum ist, so mu ein Pixel p existieren, derart, da $\mathbf{A} \setminus \{\text{seg}_{\mathbf{A}}(p), B\} \uplus \{\text{seg}_{\mathbf{A}}(p) \setminus p, B \cup b\}$, mit $B \in \text{BestNeighbor}(p)$, eine bessere Segmentierung ist. Nach Korollar 3.7.37 (S. 142) und Definition 4.6.18 mu dann aber $\text{crit}(p) > 0$ sein. Umgekehrt gilt auch, da, wenn ein Pixel $p \in \text{Kand}(\mathbf{A})$ existiert mit $\text{crit}(p) > 0$, die pixelbasierte Segmentierung \mathbf{A} kein lokales pixelbasiertes Maximum sein kann.

Eine Strategie vieler Optimierungsverfahren ist es, die Richtung des strksten Wertzuwachses whlen. Das Korollar 3.7.42 motiviert daher, die Menge der Kandidatenpixel nach ihrem kritischen Wert $\text{crit}(p)$ zu sortieren und immer das Pixel $p \in \text{Kand}(p)$ zu whlen, dessen Austausch den strksten Wertzuwachs der k -Niveaupassung verspricht.

4.6.6.2 Der Algorithmus **k-NiveaupassungZ**

Die vorangegangenen Überlegungen werden nun im Algorithmus 14 (**k-NiveaupassungZ**) zusammengefaßt. Bei diesem Algorithmus handelt es sich um eine Konkretisierung des abstrakten pixelbasierten k -Niveaupassungsalgorithmus 6 (**k-Niveaupassung**), bei dem lediglich die Auswahl einer pixelbasierten Segmentierung aus der pixelbasierten Umgebung der aktuellen Segmentierung zielgerichteter angegangen wird.

Algorithmus 14 (k-NiveaupassungZ) Berechnung eines pixelbasierten lokalen Maximums einer k -Niveaupassung Ψ bezüglich des Mengensystems $S_k(\mathfrak{Z})$ und der Klasse der erlaubten Funktionen V

Input: Diskretes Pixelbild $g : M \rightarrow \mathbb{R}^d$,
beliebige Segmentierung $\{A_1, \dots, A_k\} \in S_k(\mathfrak{Z})$,
Basis (e_1, \dots, e_n)

Output: pixelbasiertes lokales Maximum $\{B_1, \dots, B_k\}$ bezüglich der Basis (e_1, \dots, e_n)

```

1:  $\mathbf{B} = \{B_1, \dots, B_k\} \leftarrow \{A_1, \dots, A_k\}$ 
2: Berechne  $seg = seg_{\mathbf{B}}$  und bestimme  $\text{Kand}(\mathbf{B})$ 
3: continue  $\leftarrow$  TRUE
4: while continue = TRUE do
5:   continue  $\leftarrow$  FALSE
6:   Wähle  $p \in \text{Kand}(\mathbf{B})$  mit  $crit(p)$  maximal in  $\text{Kand}(\mathbf{B})$ 
7:   if  $crit(p) > 0$  then
8:     Wähle  $B \in \text{BestNeighbor}(B)$ 
9:     continue  $\leftarrow$  TRUE
10:     $seg(p) \leftarrow seg(p) \setminus p$ ,  $B \leftarrow B \cup p$ 
11:    rem Update der Funktion  $seg$ 
12:     $seg(p) \leftarrow B$ 
13:    rem Update der Kandidatenpixel
14:    Alle Nachbarpixel sind zu überprüfen, ob sie Kandidatenpixel sind oder werden können
15:    Alle Nachbarpixel der Nachbarpixel sind zu überprüfen, ob sie eventuell aus der Kandidatenliste zu streichen sind.
16:  end if
17: end while
18: return  $\{B_1, \dots, B_k\}$ 

```

Satz 4.6.20

Der Algorithmus 14 (**k-NiveaupassungZ**) terminiert und liefert ein pixelbasiertes lokales Maximum einer k -Niveaupassung bezüglich der Klasse der erlaubten Funktionen V mit der Basis (e_1, \dots, e_n) .

Beweis:

Zunächst gilt während des gesamten Algorithmus (bis auf die Zeilen 10 - 12) $seg = seg_{\mathbf{B}}$. Denn die Verschiebung des Pixels p in Zeile 10 des Algorithmus wird in der Zeile 12 für die Abbildung seg nachvollzogen.

Bei dem Algorithmus 14 (**k-NiveaupassungZ**) handelt es sich um eine Abwandlung des abstrakten Algorithmus 6 (**k-Niveaupassung**).

Nach Lemma 4.6.7 besteht die pixelbasierte Umgebung einer bestehenden Segmentierung $\mathbf{B} \in S_k(\mathfrak{Z})$ aus Segmentierungen, die sich von \mathbf{B} nur durch den Austausch eines Pixels von einem Segment zu einem benachbarten Segment unterscheiden. Insbesondere läßt sich $S_k(\mathfrak{Z})$ auf die Zwei-Niveau-Situation reduzieren. Daher genügt es, sich bei der Überprüfung der Wertverbesserung in Zeile 6 des Algorithmus 6 (**k-Niveaupassung**) auf diese zwei beteiligten Segmente zu beschränken. Nach Korollar 3.7.37 (S. 142) und der Definition 4.6.18 ist die Abfrage $crit(p) > 0$ in Zeile 6 des Algorithmus 14 (**k-NiveaupassungZ**) zu dieser Überprüfung äquivalent.

Die Wahl $p \in \text{Kand}(\mathbf{B})$ mit $crit(p)$ maximal in $\text{Kand}(\mathbf{B})$ in Zeile 5 kodiert lediglich eine gezielte Auswahl der zwei Segmente $seg_{\mathbf{B}}(p)$ und $B \in \text{BestNeighbor}(p)$, deren Austausch des Pixels p eine Segmentierung aus der pixelbasierten Umgebung von \mathbf{B} kodiert (vgl. Bemerkung 4.6.19). Existiert kein $p \in \text{Kand}(\mathbf{B})$ mit $crit(p) > 0$, so ist keine Wertsteigerung der Zielfunktion der k -Niveaupassung in der Umgebung von \mathbf{B} mehr möglich (vgl. auch Bemerkung 4.6.19). Daher muß in diesem Fall \mathbf{B} ein pixelbasiertes lokales Maximum der k -Niveaupassung sein. Aus der Endlichkeit des pixelbasierten Mengensystems folgt die Termination.

□

Bemerkung 4.6.21

Da sich nach Lemma 4.6.7 das System der Segmentierungen durch pixelzusammenhängende Mengen auf die Zwei-Niveau-Situation reduzieren läßt, kann man für dieses Mengensystem auch den Algorithmus 7 (**Iteration**) anwenden. Iterativ kann man dann für jeweils zwei Segmente ein lokales pixelbasiertes Maximum einer 2-Niveaupassung berechnen, bis für keine Kombination von zwei Segmenten eine Wertverbesserung möglich ist. Nach Satz 4.3.8 terminiert ein solches Vorgehen in einem pixelbasierten lokalen Maximum der Zielfunktion der k -Niveaupassung.

In der Praxis hat sich allerdings herausgestellt, daß der Ansatz, den der Algorithmus 14 (**k-NiveaupassungZ**) verfolgt, effektiver ist. Die Auswahlstrategie des stärksten Wertzuwachses, die durch die Ordnung der Kandidatenpixel nach ihrem kritischem Wert kodiert ist, führt zu signifikanteren Ergebnissen.

4.6.6.3 Verwaltung der Kandidatenliste

Eine effiziente Verwaltung der Menge $\text{Kand}(\mathbf{B})$ kann im Schritt 5 des Algorithmus 14 (**k-NiveaupassungZ**) zu deutlichen Laufzeitverbesserungen führen.

Liegt die Menge der Kandidatenpixel an dieser Stelle nämlich als nach ihrem kritischem Wert geordnete Liste vor, so erfolgt die Auswahl in Schritt 5 in $\mathcal{O}(1)$, indem man einfach das erste Element der Liste nimmt.

Allerdings verändert jeder Durchlauf der **while**-Schleife (Zeile 4 – 17) die Menge $\text{Kand}(\mathbf{B})$: Zum einen werden Elemente aus ihr entfernt und hinzugefügt (Schritt 13 und 14), zum anderen verändert sich auch die Ordnung der Elemente bezüglich $\text{crit}(p)$, wenn sich die Segmente verändern. Ist $\tilde{\gamma}_f(B) \neq \tilde{\gamma}_f(B \cup p)$, so ändert sich für jedes Pixel q eines Segments B , dem gerade ein Pixel p hinzugefügt wurde, der Wert $d_{B,q}(B, q)$ in $d_{B \cup p,q}(B \cup p, q)$. Analoges gilt für die Pixel des Segments, das ein Pixel verliert. Daher ist unter Umständen die vorherige Ordnung der Liste zerstört.

Allerdings verändert sich nur der Abstand derjenigen Pixel der Menge $\text{Kand}(\mathbf{B})$, deren Segment verändert wurde. Für alle anderen Pixel ändert sich der Wert $\text{crit}(p)$ nicht. Deshalb sollte für jedes Segment $B \in \mathbf{B}$ eine eigene geordnete Liste der Kandidatenpixel $\text{Kand}(B)$, geordnet nach $\text{crit}(p)$ aufgebaut werden. Damit ist teure Operation des Ordnen einer Liste, die in jedem Schleifendurchlauf gemacht werden muß, auf die Listen jener Segmente beschränkt, die in diesem Durchlauf geändert wurden. Für die Auswahl in Schritt 6 des Algorithmus müssen dann nur noch k Werte miteinander verglichen werden.

In einer realen Implementierung kann man zudem noch berücksichtigen, daß bei großen Segmenten die Hinzunahme oder Abgabe eines einzelnen Pixels die Ordnung bezüglich der Metrik $d_{B,p}$ nur unwesentlich ändern wird und es auch nicht sehr wahrscheinlich ist, daß ein soeben hinzugefügtes Pixel wieder am Anfang der Kandidatenliste des Segments stehen wird. Daher muß eine Aktualisierung der Ordnung der Liste nur sporadisch, d. h. nicht bei jeder Änderung, sondern nur bei einer gewissen Anzahl von Änderungen, durchgeführt werden.

4.6.6.4 Relaxation des Funktionenraums

Je kleiner die Basis ist, desto leichter fällt die Berechnung der für den Algorithmus 14 (**k-NiveaupassungZ**) entscheidenden Werte $\text{crit}(p)$. Denn zur Berechnung dieser Werte benötigt man die Orthonormalbasen der jeweils an der Berechnung dieser Werte beteiligten Segmente (vgl. Definition 4.6.19 und die Abschnitte 3.7.2 und 4.3.3.2), die für jede Veränderung der Segmente neu berechnet werden müssen.

Eine besondere Erleichterung der Berechnung der Werte $d_{p,B}(p, B)$ (vgl. Definition 4.6.18) erhält man, wenn die Basis eine erlaubte Basis für die Farbbildsegmentierung ist, die, wie in Beispiel 3.5.11 iii) (S. 96), aus komponentenweise konstanten Funktionen besteht. Denn dann ist, wie Beispiel 3.7.38 iii) (S. 142) zeigt, die Berechnung des Werte $d_{p,B}(p, B)$ unabhängig von der Position des Pixels p im Bild g .

Diese Überlegungen motivieren dazu, sich, bei einer Mehrniveaupassung mit einer für die Farbbildsegmentierung erlaubten Basis mit mehr als 3 Basisfunktionen, zunächst einmal

auf die Berechnung eines pixelbasierten lokalen Maximums einer Mehrniveaupassung mit den ersten drei Basisfunktionen, die im günstigsten Fall konstant sind, zu beschränken, um mit diesem Ergebnis als Startsegmentierung die Berechnung eines pixelbasierten lokalen Maximums einer Mehrniveaupassung bezüglich einer um die nächsten drei Basisfunktionen erweiterten Basis zu starten.

Der Algorithmus 15 (**Funktionenraum-Relaxation**) skizziert ein solches Vorgehen.

Algorithmus 15 (Funktionenraum-Relaxation) Berechnung eines pixelbasierten lokalen Maximums einer k -Niveaupassung Ψ bezüglich des Mengensystems $S_k(\mathfrak{Z})$ und der Klasse der erlaubten Funktionen V , wobei V eine erlaubte Basis zur Farbbildsegmentierung besitzt

Input: Diskretes Pixelbild $g : M \rightarrow \mathbb{R}^d$,
beliebige Segmentierung $\{A_1, \dots, A_k\} \in S_k(\mathfrak{Z})$,
für die Farbbildsegmentierung erlaubte Basis
 $(e_{1_r}, e_{1_g}, e_{1_b}, \dots, e_{n_r}, e_{n_g}, e_{n_b})$

Output: pixelbasiertes lokales Maximum \mathbf{B}

- 1: $\mathbf{B} = \{B_1, \dots, B_k\} \leftarrow \{A_1, \dots, A_k\}$
 - 2: $i \leftarrow 1$
 - 3: **while** $i < n$ **do**
 - 4: $\mathbf{B} \leftarrow k$ -Niveaupassung $Z(g, \mathbf{B}, (e_{1_r}, e_{1_g}, e_{1_b}, \dots, e_{i_r}, e_{i_g}, e_{i_b}))$
 - 5: $i \leftarrow i + 1$
 - 6: **end while**
 - 7: **return** \mathbf{B}
-

4.6.7 Beispiele

In allen Beispielen dieses Abschnitts wurde die 4-er Nachbarschaft gewählt.

Beispiel 4.6.22 (Granitplatten)

Bei der ersten Beispielfamilie handelt es sich um eine industrielle Anwendung der vektoriellen Mehrniveaupassung. Im Rahmen eines Projektes erstellte das FORWISS für eine granitverarbeitende Firma ein Programm zur Schnittoptimierung von Granittranchen. Grundvoraussetzung für eine solche Optimierung ist die Bereitstellung der Geometrie der jeweiligen Tranchen. Dazu werden die zu verarbeitenden Tranchen senkrecht vor einem gelben Hintergrund in einem Abstand von circa 5 Metern mit einer CCD-Kamera aufgenommen. In den so gewonnenen Bildern müssen dann die Granittranchen vom gelben Hintergrund getrennt werden. Die Farbbilder sind dabei bezüglich des RGB-Farbmodells gegeben. Die Schwierigkeit bei dieser Anwendung liegt darin, daß die Farben der Granittranchen stark variieren und die industrielle Umgebung teilweise wechselnden Lichtverhältnissen unterworfen ist.

In diesem Szenario wurde der Vektorraum der kanalweise konstanten Funktionen als Raum der erlaubten Funktionen gewählt. Dabei ist $E := \mathbb{R}^3$ mit dem Standardskalarprodukt versehen. Als Basis wurde $\mathcal{E} := \{e_r, e_g, e_b\}$ mit $e_r := (1, 0, 0)$, $e_g := (0, 1, 0)$ und $e_b := (0, 0, 1)$ gewählt. Bis auf Abbildung 4.26, die das Ergebnis einer 3-Niveaupassung des vorsegmentierten Bildes aus Abbildung 4.25 zeigt, sind alle Beispielbilder direkt aus der Industrieumgebung. Die Ergebnisse sind nach folgendem Verfahren gewonnen:

1. Berechnung eines vorsegmentierten Bildes mit Hilfe des Algorithmus 13 (**Initiales Farbschwellwertverfahren**) mit einem Schwellwert pro Kanal. Dabei sollten die initialen Segmente eine gewisse Mindestgröße haben. Einzelne Pixel werden dabei den Nachbarsegmenten, denen sie von Farbwert am ähnlichsten sind, zugeordnet.
2. Verschmelzung der initialen Segmente über den Algorithmus 16 (**Iterative Niveaverschmelzung**) bis zwei Segmente übrig bleiben.
3. 2-Niveaupassung gemäß Algorithmus 6 (**k-NiveaupassungZ**)

Man sieht, daß mit Hilfe dieser zwei Parameter – der Anzahl der Schwellwerte in der Vorsegmentierung und der Anzahl der Ergebnissegmente – die vektorielle Mehrniveaupassung ein sehr robustes Verfahren zur Trennung von Tranche und Hintergrund liefert.

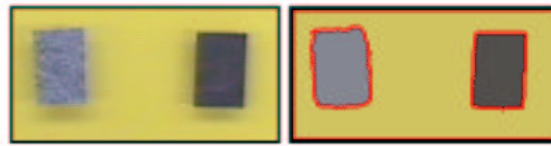


Abbildung 4.26: Ergebnis einer 3-Niveaupassung des vorsegmentierten Bildes aus Abbildung 4.25.

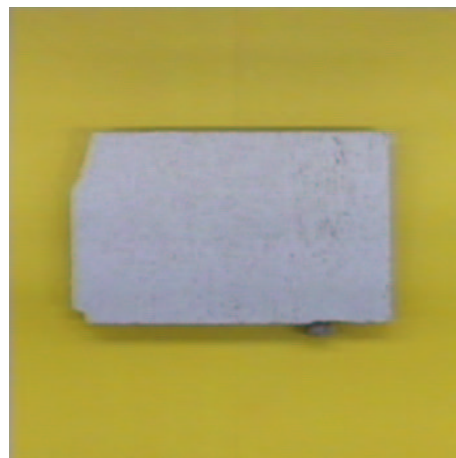


Abbildung 4.27: Typisches Eingabebild im Kontext des Beispiels 4.6.22.

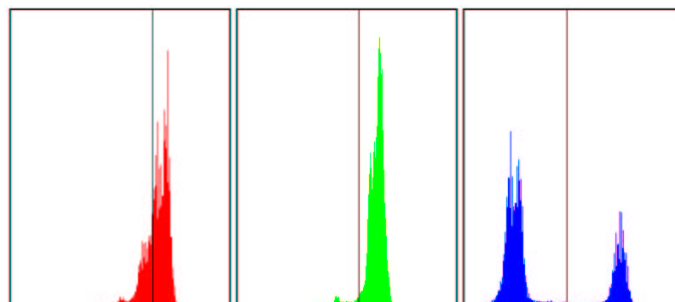


Abbildung 4.28: Kanalweise Histogramme des Eingabebildes aus Abbildung 4.27. Die Schwellwerte, die der Algorithmus 10 (**Histogramm-analyse**) für einen Schwellwert pro Kanal liefert, sind eingezeichnet.

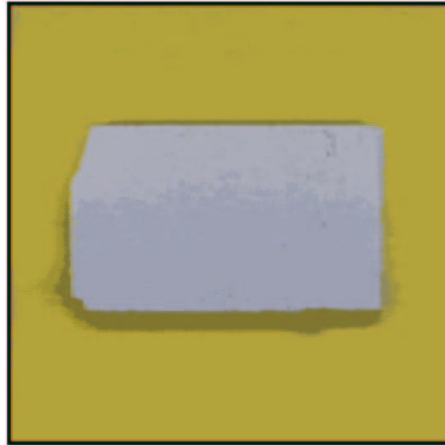


Abbildung 4.29: Ergebnis der Vorsegmentierung des Eingabebildes aus Abbildung 4.27 mit dem Algorithmus 13 (**Initiales Farbschwellwertverfahren**) mit einem Schwellwert.

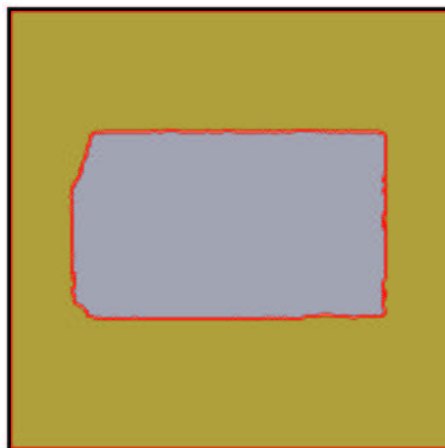


Abbildung 4.30: Ergebnis des im Beispiel 4.6.22 beschriebenen Verfahrens für das Eingabebild aus Abbildung 4.27.

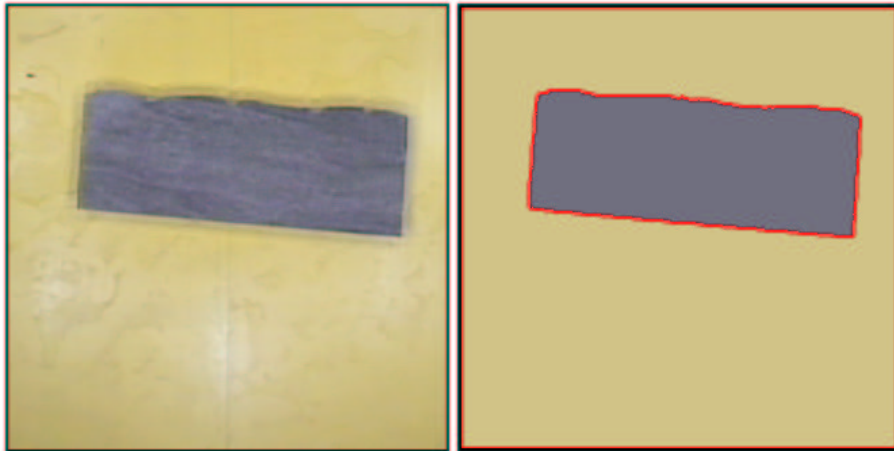


Abbildung 4.31: Typisches Eingabe- und Ergebnisbild des in Beispiel 4.6.22 beschriebenen Verfahrens.

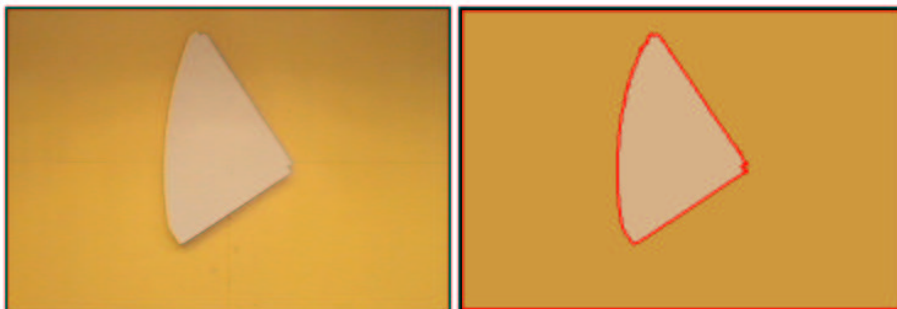


Abbildung 4.32: Typisches Eingabe- und Ergebnisbild des in Beispiel 4.6.22 beschriebenen Verfahrens.

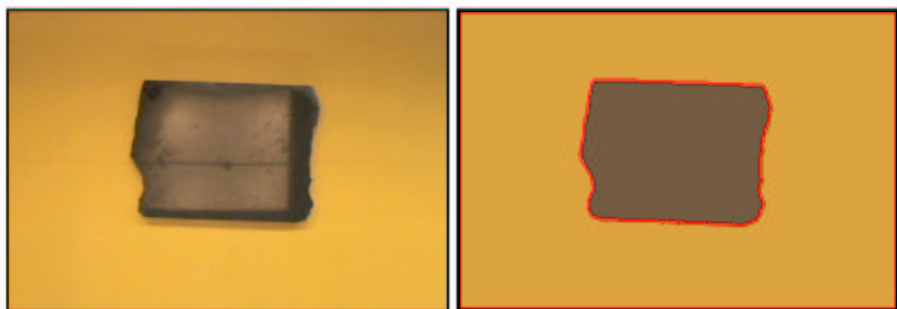


Abbildung 4.33: Typisches Eingabe- und Ergebnisbild des in Beispiel 4.6.22 beschriebenen Verfahrens.

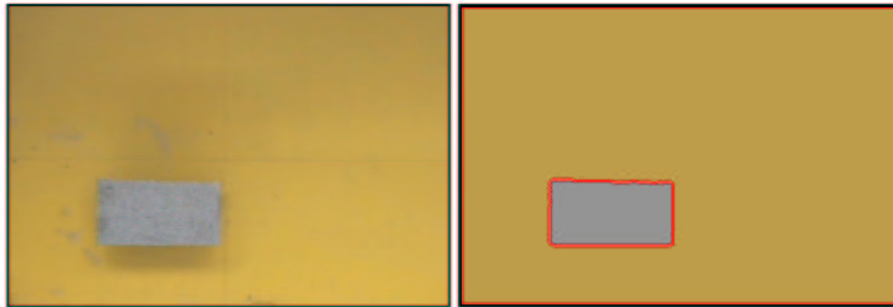


Abbildung 4.34: Typisches Eingabe- und Ergebnisbild des in Beispiel 4.6.22 beschriebenen Verfahrens.

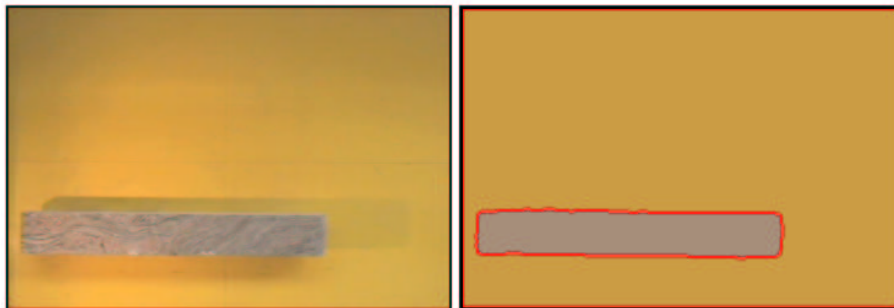


Abbildung 4.35: Typisches Eingabe- und Ergebnisbild des in Beispiel 4.6.22 beschriebenen Verfahrens.



Abbildung 4.36: Typisches Eingabe- und Ergebnisbild des in Beispiel 4.6.22 beschriebenen Verfahrens. Man beachte, daß aufgrund der Reduktion auf zwei Segmente nach Algorithmus 16 (**Iterative Niveaverschmelzung**) das innere gelbe Segment verschwunden ist. Auch der untere Schatten der Granittranche wird zum dunkleren Segment gezählt.

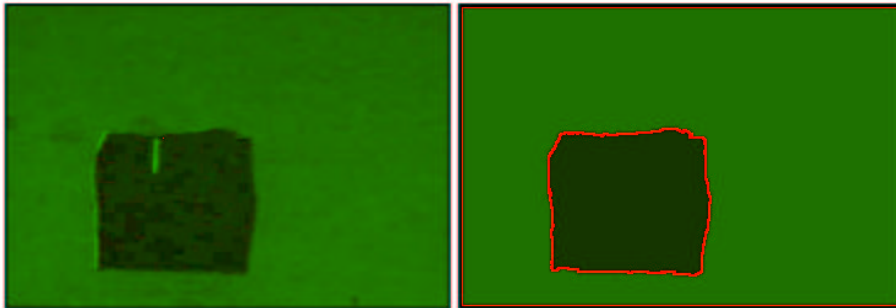


Abbildung 4.37: Transformiert man das RGB-Bild aus Abbildung 4.36 in den HSI-Farbraum (vgl. Abschnitt 1.3.2.1.2), so erhält man ein besseres Ergebnis. Das linke Bild ist die normalisierte RGB-Darstellung der HSI-Werte des Eingabebildes aus Abbildung 4.36. Dabei wurde der I-Kanal, also die Intensität, weggelassen. Man beachte, daß der Schatten, der das Ergebnis in Abbildung 4.36 erheblich störte, nun nicht mehr zu sehen ist. Nimmt man dieses Bild als Eingabebild des im Beispiel 4.6.22 beschriebenen Algorithmus, so erhält man das gewünschte Ergebnis (rechts).

Bemerkung 4.6.23 (Laufzeiten)

Die Laufzeit des Algorithmus hängt natürlich stark von Parametern, wie z. B. die Anzahl der Schwellwerte im Algorithmus 13 (**Initiales Farbschwellwertverfahren**) oder die Mindestgröße für initiale Segmente, ab. Für die Beispielanwendung aus Beispiel 4.6.22 ergeben sich im Schnitt (gemittelt über alle Bilder von Abbildung 4.32 bis Abbildung 4.36) folgende Werte, die auf einem 500MHz Alpha-Prozessor (EV56) unter RedHat-Linux 2.2.18 gemessen wurden:

Mindestgröße	Initialisierung	Niveaureduktion	Niveaupassung
1 Schwellwert pro Kanal			
1 Pixel	0.22s	0.22s - 16s	0.19s
10 Pixel	0.22s	0.10s - 0.20s	0.23s
20 Pixel	0.22s	0.09s - 0.13s	0.24s
2 Schwellwerte pro Kanal			
1 Pixel	0.22s	0.22s - 22s	0.23s
10 Pixel	0.22s	0.09s - 0.21s	0.24s
20 Pixel	0.22s	0.08s - 0.15s	0.24s
3 Schwellwerte pro Kanal			
1 Pixel	0.22s	10s - 80s	0.25s
10 Pixel	0.22s	0.22s	0.24s
20 Pixel	0.22s	0.14s - 0.16s	0.24s

Dabei bestehen die Bilder aus 360×245 Pixeln. Die teilweise hohen Laufzeiten bei der Niveaureduktion kommen zustande, weil der Algorithmus 16 (**Iterative Niveaoverschmelzung**) eine Laufzeit hat, die man quadratisch in Abhängigkeit von der Anzahl der Startsegmente abschätzen muß (vgl. Bemerkung 5.3.10). Die Anzahl der Startsegmente sollte also nicht zu groß gewählt werden.

Die Ergebnisbilder sind trotz der Parameter „Mindestgröße der initialen Segmente“ und „Anzahl der Schwellwerte“ in gedruckter Form nicht zu unterscheiden.

Beispiel 4.6.24

Das Bild (vgl. Abbildung 4.38) dieses Beispiels ist der Internet-Seite³ der Autoren des Artikels [CM97] entnommen. In diesem Artikel wird ein Verfahren zur Segmentierung von Farbbildern über eine Farbklassifikation von Pixeln präsentiert. In Abbildung 4.39 ist das Ergebnis des dort vorgestellten Algorithmus zu sehen. Die Abbildung 4.40 zeigt das gemäß Algorithmus 13 (**Initiales Farbschwellwertverfahren**) vorsegmentierte Bild. Die Abbildung 4.41 zeigt, daß das Ergebnis einer 7-Niveaupassung mit dem Ergebnis des Originalartikels vergleichbar ist. Dabei kommt die 7-Niveaupassung ohne eine Farbraumtransformation aus, während der in [CM97] vorgeschlagene Algorithmus auf dem CIEL^{*}u^{*}v^{*}-Farbraum arbeitet. Der CIEL^{*}u^{*}v^{*}-Farbraum ist ein Farbraum, der ähnlich dem HSI-Farbraum den chromatischen vom unchromatischen Anteil in der Farbdarstellung trennt (vgl. [WS82], 3.3.9).

³http://www.caip.rutgers.edu/~comanici/segm_images.html.



Abbildung 4.38: Eingabebild.



Abbildung 4.39: Ergebnisbild des in [CM97] vorgeschlagenen Algorithmus.



Abbildung 4.40: Mit einem Schwellwert pro Kanal gemäß Algorithmus 13 (**Initiales Farbschwellwertverfahren**) vorsegmentiertes Bild.

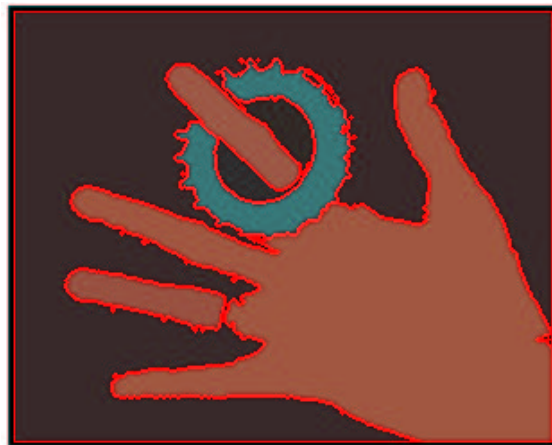


Abbildung 4.41: Ergebnis einer 7-Niveaupassung des Bildes aus Abbildung 4.38.

Beispiel 4.6.25

Das Beispiel 4.43 zeigt ein Problem des Algorithmus 14 (**k-NiveaupassungZ**): Angenommen es existieren zwei deutlich trennbare Segmente A und B mit durchschnittlichen Farbwerten c_A und c_B . Durch Rauschen existiere in B ein isoliertes Pixel p mit dem Farbwert c_A . Ist nun eine Startsegmentierung $\{A', B'\}$ mit $p \in A'$, so ist das Pixel p in dieser Startsegmentierung dem falschen Segment zugeordnet. Der Algorithmus 14 (**k-NiveaupassungZ**) kann dieses Pixel jedoch oft nicht vom Segment entfernen, ohne die Zusammenhangsbedingung zu verletzen (vgl. Abbildung 4.42). Daher kommt es in sol-

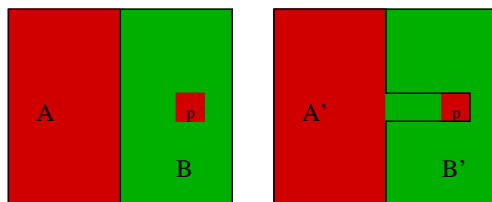


Abbildung 4.42: Illustration zum Beispiel 4.6.25.

chen Fällen unter Umständen zu einem fransigen Segmentrand. Fransige Segmentgrenzen lassen sich durch nachfolgende Algorithmen beheben. In [Jai89] wird z. B. ein Verfahren präsentiert, das die Segmentgrenzen über eine Approximation durch B-Splines glättet.

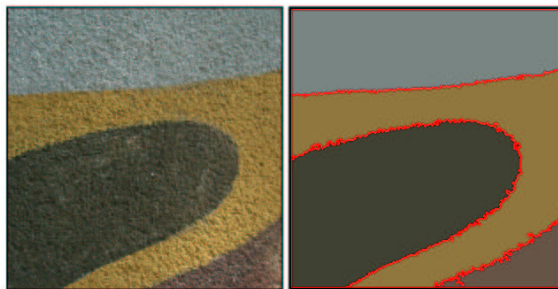


Abbildung 4.43: Eingabebild (Graffiti an einer Wand) und Ergebnis einer 4-Niveaupassung des Eingabebildes.

Laufzeiten für das Beispielbild der Abbildung 4.43 (309×317 Pixel)

Mindestgröße	Initialisierung	Niveaureduktion	Niveaupassung
1 Schwellwert pro Kanal			
1 Pixel	0.33s	46s	0.33s
10 Pixel	0.33s	0.20s	0.36s
20 Pixel	0.33s	0.10s	0.36s
2 Schwellwerte pro Kanal			
1 Pixel	0.33s	226s	0.52s
10 Pixel	0.33s	0.87s	0.39s
20 Pixel	0.33s	0.30s	0.37s

(Gemessen auf einem 500MHz Alpha-Prozessor (EV56) unter RedHat-Linux 2.2.18.)

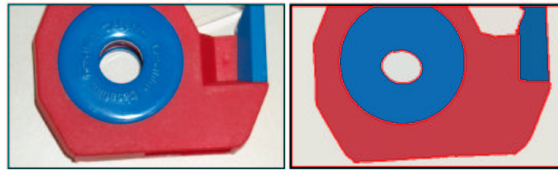


Abbildung 4.44: Eingabebild und Ergebnis einer 6-Niveaupassung des Eingabebildes.

Laufzeiten für das Beispielbild der Abbildung 4.44 (375×221 Pixel)

Mindestgröße	Initialisierung	Niveaureduktion	Niveaupassung
1 Schwellwert pro Kanal			
1 Pixel	0.20s	0.30s	0.51s
10 Pixel	0.20s	0.08s	0.54s
20 Pixel	0.20s	0.05s	0.58s
2 Schwellwerte pro Kanal			
1 Pixel	0.20s	2s	0.23s
10 Pixel	0.20s	0.12s	0.57s
20 Pixel	0.20s	0.11s	0.52s

(Gemessen auf einem 500MHz Alpha-Prozessor (EV56) unter RedHat-Linux 2.2.18.)

4.6.8 Einfach pixelzusammenhängende Pixelmengen

Eine Teilmenge des \mathbb{R}^2 heißt *einfach zusammenhängend* (vgl. [BS91], S. 244)⁴, wenn ihr Rand zusammenhängend ist. Anschaulich heißt das, daß es keine „Löcher“ in der zusammenhängenden Menge gibt. Ein Mengensystem einfach zusammenhängender Pixelmengen ist daher für viele Segmentierungsprobleme interessant. Allerdings führt der Begriff des Pixelzusammenhangs in diesem Kontext zu vermeintlichen Paradoxien, wie Beispiel 4.6.26 zeigt.

Beispiel 4.6.26

Gegeben sei die Pixelmenge $P = \{[0, 1], [1, 1], [2, 1], [1, 0], [1, 2]\}$. Die Abbildung 4.45 skizziert diese Pixelmenge. Diese Menge ist dann pixelzusammenhängend bezüglich der 4-er Nachbarschaft. Weiter kann man sich leicht vergegenwärtigen, daß die Menge $A = \bigcup_{p \in P} p$ nicht nur zusammenhängend, sondern auch einfach zusammenhängend ist. Nun ist $R := P \setminus \{[1, 1]\}$ die Menge der Randpixel von P nach Definition 4.6.4. Wieder gilt $A' = \bigcup_{p \in R} p$ ist zusammenhängend. Allerdings ist R bezüglich der 4-er Nachbarschaft *nicht* pixelzusammenhängend. Daher reicht es nicht aus, für eine Definition einfach

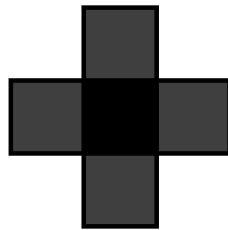


Abbildung 4.45: Skizze der in Beispiel 4.6.26 beschriebenen Pixelmenge.

pixelzusammenhängender Mengen zu fordern, daß der Rand pixelzusammenhängend ist. Vielmehr muß man sich auf die ursprüngliche Definition des Begriffes „einfach zusammenhängend“ zurückziehen.

Definition 4.6.27 (Einfach pixelzusammenhängende Menge)

Eine Menge von Pixeln P heie *einfach pixelzusammenhängend* bezüglich der 4-er bzw. 8-er Nachbarschaft, falls sie pixelzusammenhängend ist bezüglich der 4-er bzw. 8-er Nachbarschaft, und falls die Menge $\bigcup_{p \in P} p$ einfach zusammenhängend ist.

⁴Eine allgemeinere Definition des Begriffes findet man in [Die85], S. 208

Bemerkung 4.6.28

Man kann sich leicht überlegen, daß – eine korrekte Initialisierung vorausgesetzt – eine Überprüfung gemäß Definition 4.6.11 Verletzungen des einfachen Pixelzusammenhangs durch Pixelaustausch erkennt. Alle in diesem Unterkapitel vorgestellten Algorithmen und Verfahren, bis auf die Anmerkungen zur Initialisierung in Abschnitt 4.6.5, behalten also auch für ein Mengenmodell der einfach pixelzusammenhängenden Mengen ihre Gültigkeit. Eine initiale Segmentierung durch einfach pixelzusammenhängende Mengen kann man z. B. erhalten, indem man in einer Segmentierung durch pixelzusammenhängende Menge alle Segmente, die nur ein Nachbarsegment haben und nicht am Rand liegen, entfernt und die Pixel des Segments dem Nachbarsegment hinzufügt.

Kapitel 5

Niveaureduktion und Schätzung der Anzahl der Segmente

5.1 Einleitung

Bis zu dieser Stelle wurde die Forderung iv. der notwendigen Bedingungen einer Segmentierung aus Definition 2.2.1 (S. 18), nämlich, daß die Anzahl der Segmente einer Segmentierung ein Pareto-Optimum einer Mehrzieloptimierung sein sollte, nicht beachtet. Dieses Kapitel soll nun dieses Versäumnis nachholen. Es soll allerdings gleich an dieser Stelle darauf hingewiesen werden, daß sich die Ergebnisse dieses Kapitel keineswegs mathematisch so begründen lassen wie die Ergebnisse der vorangegangenen Kapitel. Letztlich ist es, nach der Meinung des Autors dieser Arbeit, nicht möglich, ein allgemeingültiges Verfahren, das die Anzahl der Objekte ohne Nebenwissen in einem Bild berechnen kann, anzugeben. Daher kann auch dieses Kapitel nur Ansätze und Hinweise geben, wie man unter Umständen die Anzahl der Objekte im Bild schätzen kann.

Das grundsätzliche Vorgehen dabei ist einfach: Man startet mit einer lokal optimalen Segmentierung durch eine große Anzahl von Segmenten und entfernt solange Segmente, bis ein auftretender Approximationsfehler zu groß ist. Vereinfachend heißt das, daß man versucht mit so wenig Segmenten wie nötig auszukommen. Dabei ist genau zu spezifizieren, was „nötig“ eigentlich heißen könnte.

Übersicht des Kapitels

Im anschließenden zweiten Unterkapitel wird der aus der Ökonomie bekannte Begriff des „Pareto-Optimums“ eingeführt. Ein Zustand einer Mehrzieloptimierung heißt dabei „Pareto-Optimal“, falls man in keiner Komponente eine Verbesserung erreichen kann ohne eine andere zu verschlechtern. Dabei sind die zwei Komponenten im Falle der Segmentierung klar: Auf der einen Seite ist die Anzahl der Segmente möglichst klein zu halten, auf der anderen Seite muß man einen Approximationsfehler minimieren, der mit Abnahme der Anzahl der Segmente steigt.

Im dritten Unterkapitel wird ein Algorithmus vorgestellt, der, obwohl er vom theoretischen Standpunkt nicht sonderlich interessant erscheint, in den praktischen Anwendungen der Mehrniveaupassung zur Bildsegmentierung einen sehr großen Raum einnimmt. Es ist dies ein Algorithmus zur sukzessiven Reduktion der Anzahl der Niveaus innerhalb einer Mehrniveaupassung.

Das vierte Unterkapitel widmet sich dann dem eigentlichen Problem dieses Kapitels, nämlich der Schätzung der Anzahl der Objekte oder besser gesagt der Schätzung der Anzahl der deutlich trennbaren Segmente im gegebenen Bild. Im zentralen Algorithmus dieses Unterkapitels wird die Anzahl der Niveaus als Pareto-Optimum einer Zweizieloptimierung unter gewissen Nebenbedingungen betrachtet. Eine Teilfunktion der Optimierung ist die Approximationsgüte bei variierender Anzahl von Segmenten. Dabei wird sich herausstellen, daß sich die $\|\cdot\|_M$ -Norm als Approximationsmaß der Zweizieloptimierung schlechter geeignet ist, als das Maximum der gemittelten Abstände auf den einzelnen Segmenten. Dieser Fehler ermöglicht es, daß unter Umständen bei Abnahme der Anzahl der Niveaus der Approximationsfehler sinkt, wie die Beispiele in Unterkapitel 5.4 belegen.

5.2 Pareto-Optima

Der Begriff des *Pareto-Optimums* ist in der Ökonomie weit verbreitet. Man versteht im allgemeinen darunter eine (gesellschaftliche) Situation, in der es nicht möglich ist, die Wohlfahrt eines Individuums durch eine Re-Allokation der Ressourcen zu erhöhen ohne gleichzeitig die eines anderen Individuums zu verringern. Anders formuliert: Eine Situation, in der A besser gestellt werden kann, ohne daß dabei B gleichzeitig schlechter gestellt werden muß. Das Gesamtsystem muß sich in diesem Zustand nicht im Optimum befinden.

Der Begriff *Pareto-Optimum* geht zurück auf den italienischen Ökonomen Vilfredo Pareto (1848 - 1923). Er lehrte ab 1893 an der Universität von Lausanne und veröffentlichte 1896 - 1897 die Reihe „Cours d'Économie Politique“, in der er versucht, in mathematischen Modellen zu beweisen, daß die Verteilung von Einkommen und Wohlfahrt in einer Gesellschaft nicht zufällig sind, sondern gewissen Regeln unterliegen. Aus dieser Arbeit leitet sich der Begriff des *Pareto-Optimums* ab.

Die mathematische Definition des *Pareto-Optimums* innerhalb einer Mehrzieloptimierung folgt in dieser Arbeit der Dissertation von Eckard Zitzler [Zit99], ist aber in seiner Notation allgemein verbreitet. Obwohl allgemein in der Mehrzieloptimierung ein Mehrzielmaximierungsproblem gelöst werden soll, wird an dieser Stelle ein Mehrzielminimierungsproblem definiert, das der späteren Verwendung, nämlich der Minimierung des Approximationsfehlers bei gleichzeitiger Minimierung der Anzahl der Segmente, eher entgegenkommt.

Definition 5.2.1 (Mehrzieloptimierung)

Eine Mehrzieloptimierung (U, N) besteht aus einer Menge von Zielfunktionen $U = (u_1, \dots, u_k)$ und einer Menge von Nebenbedingungen $N = (e_1, \dots, e_m)$. Ziel ist es, den Vektor

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_k(\mathbf{x}))$$

unter den Nebenbedingungen

$$e_1(\mathbf{x}) \leq 0, \dots, e_m(\mathbf{x}) \leq 0$$

zu minimieren.

Den Vektor \mathbf{y} zu minimieren, heißt im Zusammenhang der Mehrzieloptimierung, die einzelnen Einträge des Vektors zu minimieren. Die Minimierung eines einzelnen Vektoreintrages kann jedoch unter Umständen der Minimierung eines anderen Vektoreintrags entgegenstehen.

Dabei ist $\mathbf{x} = (x_1, \dots, x_n) \in X$ und $\mathbf{y} = (y_1, \dots, y_k) \in Y$. X heißt der *Entscheidungsraum* und Y der *Zielraum* der Mehrzieloptimierung.

Definition 5.2.2

Für zwei Vektoren \mathbf{u}, \mathbf{v} des Entscheidungsraums X mit $\mathbf{u} = (u_1, \dots, u_k)$ und $\mathbf{v} = (v_1, \dots, v_k)$ definiere

$$\begin{aligned}\mathbf{u} = \mathbf{v} & :\Leftrightarrow \forall i \in \{1, \dots, k\} : u_i = v_i \\ \mathbf{u} \leq \mathbf{v} & :\Leftrightarrow \forall i \in \{1, \dots, k\} : u_i \leq v_i \\ \mathbf{u} < \mathbf{v} & :\Leftrightarrow \mathbf{u} \leq \mathbf{v} \wedge \mathbf{u} \neq \mathbf{v}\end{aligned}$$

Definition 5.2.3 (Pareto-Dominanz)

Für zwei Vektoren \mathbf{u}, \mathbf{v} des Entscheidungsraums X mit $\mathbf{u} = (u_1, \dots, u_k)$ und $\mathbf{v} = (v_1, \dots, v_k)$ definiere

$$\begin{aligned}\mathbf{u} \prec \mathbf{v} \text{ (u dominiert v)} & :\Leftrightarrow \mathbf{f}(\mathbf{u}) < \mathbf{f}(\mathbf{v}) \\ \mathbf{u} \preceq \mathbf{v} \text{ (u dominiert v schwach)} & :\Leftrightarrow \mathbf{f}(\mathbf{u}) \leq \mathbf{f}(\mathbf{v}) \\ \mathbf{u} \sim \mathbf{v} \text{ (u und v sind indifferent zueinander)} & :\Leftrightarrow \mathbf{f}(\mathbf{u}) \not\leq \mathbf{f}(\mathbf{v}) \wedge \mathbf{f}(\mathbf{v}) \not\leq \mathbf{f}(\mathbf{u})\end{aligned}$$

Definition 5.2.4 (Pareto-Optimum)

Es sei $N(X) := \{\mathbf{x} \in X \mid n_1(\mathbf{x}) \leq 0, \dots, n_m(\mathbf{x}) \leq 0\}$. Sei $A \subset N(X)$. Ein Vektor \mathbf{x} aus dem Entscheidungsraum X heißt *nicht dominiert von A* genau dann, wenn $\nexists \mathbf{a} \in A : \mathbf{a} \prec \mathbf{x}$. \mathbf{x} heißt ein *Pareto-Optimum* genau dann, wenn \mathbf{x} nicht von $N(X)$ dominiert wird.

Beispiel 5.2.5

Um die Begriffe *Dominanz* und *Pareto-Optimum* zu verdeutlichen, sei als Beispiel auf die Skizze einer Zweizieloptimierung in Abbildung 5.1 verwiesen. Das Beispiel kommt ohne Nebenbedingungen N aus. In der Skizze dominiert e den Wert f im Zielraum. Weiter dominiert b die Werte c, d, e, f des Zielraums. Die Werte a und b sind die einzigen Werte, die von keinem anderen Wert dominiert werden. Sie sind daher Pareto-Optima. Innerhalb der Menge $\{c, d, e, f\}$ sind die Werte c und e Pareto-optimal.

Bemerkung 5.2.6 (Formulierung der Mehrzieloptimierung)

Im Kontext der Segmentierung von Bildern durch Approximation des Bildes auf Segmenten ergibt sich fast kanonisch ein Mehrzieloptimierungsproblem: Die Anzahl der Segmente versus den Approximationsfehler. Im Rahmen der Definitionen dieses Abschnitts kann man dies wie folgt formulieren:

Als Entscheidungsraum dient $\mathfrak{S}_k(\mathfrak{A})$ der Raum der Segmentierungen von M durch k erlaubte Mengen. Die Zielfunktionen $U = (u_{no}, u_{approx})$ seien auf der einen Seite der Approximationsfehler u_{approx} und auf der anderen Seite die Anzahl der echten Segmente u_{no} . Die Menge der Nebenbedingungen erlaubt es beispielsweise, Sensoreigenschaften zu berücksichtigen: Man kann verlangen, daß der Approximationsfehler auf einem Segment eine vorher festgelegte Höchstgrenze nicht überschreiten darf. Bei dieser Höchstgrenze würde es Sinn machen, den Approximationsfehler unter Umständen in der Maximumsnorm zu messen.

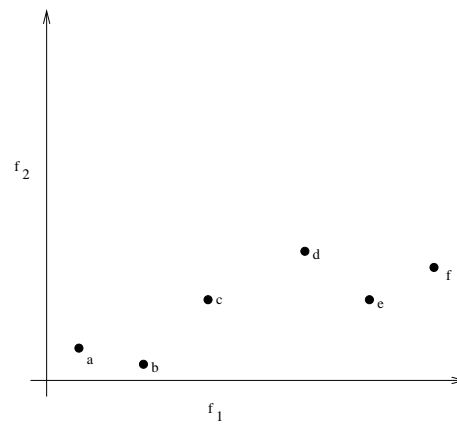


Abbildung 5.1: Beispielskizze einer Mehrzieloptimierung mit zwei Zielfunktionen.

5.3 Niveaureduktion

An mehreren Stellen dieser Arbeit wurde bereits auf den Algorithmus 16 (**Iterative Niveaoverschmelzung**) hingewiesen. Sowohl in der Bemerkung 4.4.2 in Abschnitt 4.4.2, als auch in Abschnitt 4.6.5, wird der unten stehende Algorithmus dazu verwendet, eine geeignete Startsegmentierung für die jeweiligen Algorithmen zu konstruieren.

Der Algorithmus 16 (**Iterative Niveaoverschmelzung**) kann als „Merging“-verfahren (vgl. Abschnitt 2.3.2) gesehen werden. Denn sukzessive werden jeweils zwei Segmente A und B vereinigt, die sich bezüglich des Ähnlichkeitsmaßes $\|\gamma_g(A)\|_2^2 + \|\gamma_g(B)\|_2^2 - \|\gamma_g(A \cup B)\|_2^2$ unter allen möglichen Vereinigungen von Segmenten am ähnlichsten sind. Natürlich ist bei einer solchen Vereinigung zu überprüfen, ob das Ergebnis eine Segmentierung durch erlaubte Mengen ergibt.

Algorithmus 16 (Iterative Niveaoverschmelzung) Iterative Verschmelzung von einzelnen Segmenten

Input: Diskretes Pixelbild $g : M \rightarrow \mathbb{R}^d$,
 pixelbasiertes Maximum einer k -Niveaupassung $\{A_1, \dots, A_k\} \in S_k(\mathfrak{A})$,
 $l \in \mathbb{N}$ mit $l < k$

Output: Segmentierung $\{B_1, \dots, B_l\} \in S_l(\mathfrak{A})$

- 1: $i \leftarrow k$
 - 2: $\{B_1, \dots, B_i\} \leftarrow \{A_1, \dots, A_k\}$
 - 3: **while** $i > l$ **do**
 - 4: Wähle Indizes i_0, i_1 mit $1 \leq i_0 < i_1 \leq i$ derart, daß $B_{i_0} \cup B_{i_1} \in \mathfrak{A}$ und

$$\|\gamma_g(B_{i_0})\|_2^2 + \|\gamma_g(B_{i_1})\|_2^2 - \|\gamma_g(B_{i_0} \cup B_{i_1})\|_2^2$$

$$= \min \left\{ \|\gamma_g(B_{j_0})\|_2^2 + \|\gamma_g(B_{j_1})\|_2^2 - \|\gamma_g(B_{j_0} \cup B_{j_1})\|_2^2 \mid \right.$$

$$\left. B_{j_0} \cup B_{j_1} \in \mathfrak{A}, 1 \leq j_0 < j_1 \leq i \right\}$$
 - 5: $B_{i_0} \leftarrow B_{i_0} \cup B_{i_1}, B_{i_1} \leftarrow B_{i_1}, B_i \leftarrow O$
 - 6: $i \leftarrow i - 1$
 - 7: **end while**
 - 8: **return** $\{B_1, \dots, B_i\}$
-

Bemerkung 5.3.1

Der Algorithmus 16 (**Iterative Niveaoverschmelzung**) überführt streng genommen Elemente aus $S_k(\mathfrak{A})$ in Elemente aus $S_l(\mathfrak{A})$. Das Ergebnis des Algorithmus kann also nicht aus $S_l(\mathfrak{A})$ mit $l < k$ sein. Allerdings wird in jedem Schleifendurchlauf ein Element der gerade gültigen Segmentierung auf O gesetzt. Dieses ist in der Nummerierung der Segmentierung im Algorithmus das letzte Segment. Daher wird es ab Zeile 6 des Algorithmus einfach ignoriert. Da es ein leeres Segment war, ist die in Zeile 7 gültige Menge $\{B_1, \dots, B_i\}$ eine Segmentierung in $S_i(\mathfrak{A})$.

Mit dieser Bemerkung ergibt sich das folgende Lemma:

Lemma 5.3.2

Nach Beendigung des Algorithmus 16 (**Iterative Niveaoverschmelzung**) gilt:

- i. $i = l$
- ii. $\{B_1, \dots, B_l\} \in S_l(\mathfrak{A})$

Beweis:

folgt mit Bemerkung 5.3.1

Bemerkung 5.3.3

Man beachte, daß man für beliebige erlaubte Mengensysteme \mathfrak{A} die Termination des Algorithmus 16 (**Iterative Niveaoverschmelzung**) nicht sichern kann. Die Bedingung $B_{i_0} \cup B_{i_1} \in \mathfrak{A}$ in Zeile 4 des Algorithmus ist unter Umständen für kein Paar B_{i_0}, B_{i_1} von Segmenten aus $\{B_1, \dots, B_k\}$ erfüllt (vgl. etwa Beispiel 4.3.3).

Man kann jedoch zeigen, daß der Algorithmus für pixelbasierte Systeme von Segmentierungen, die auf die Zwei-Niveau-Situation reduzierbar sind, immer terminiert.

Bemerkung 5.3.4

Für die Zielfunktion der pixelbasierten Mehrniveaupassung

$$\Psi : \mathbf{B} \mapsto \sum_{B \in \mathbf{B}} \|\gamma_g(B)\|_2^2$$

bedeutet die Auswahl der zu vereinigenden Segmente in Zeile 4 des Algorithmus 16 (**Iterative Niveaoverschmelzung**), daß das Ergebnis der Vereinigung den geringsten Wertverlust von Ψ unter allen erlaubten Vereinigungen darstellt. Hier findet man eine starke Analogie zum Multiskalenansatz im Kontext der Methoden der Variationsrechnung, wo die zur Vereinigung bestimmten Segmente so ausgesucht werden, daß ihre Vereinigung den geringsten Fehlerzuwachs für das Energiefunktional bedeutet (vgl. Abschnitt 2.4.3.2).

Sei g ein diskretes Grauwertpixelbild. Ist $\mathcal{E} = (1)$, so ist im Kontext von Beispiel 3.5.11 i) (S. 96) für zwei Segmente A, B das „Ähnlichkeitsmaß“ gegeben durch

$$\begin{aligned} \|\gamma_g(A)\|_2^2 + \|\gamma_g(B)\|_2^2 - \|\gamma_g(A \cup B)\|_2^2 = \\ \tilde{\gamma}_g(A) \int_A g d\lambda + \tilde{\gamma}_g(M \setminus A) \int_{M \setminus A} g d\lambda - \tilde{\gamma}_g(A \cup B) \int_{A \cup B} g d\lambda \end{aligned}$$

Dabei ist $\tilde{\gamma}_g(A) = \frac{1}{\lambda(A)} \int_A g d\lambda$ der durchschnittliche Grauwert des Bildes g auf dem Segment A mit $\lambda(A) > 0$. Es ist daher

$$\begin{aligned} \tilde{\gamma}_g(A) \int_A g d\lambda &= \tilde{\gamma}_g(A) \left(\frac{\lambda(A)}{\lambda(A)} \int_A g d\lambda \right) \\ &= \tilde{\gamma}_g(A) \lambda(A) \tilde{\gamma}_g(A) \\ &= \lambda(A) \tilde{\gamma}_g(A)^2 \end{aligned}$$

Damit folgt also, daß in diesem Fall gilt:

$$\begin{aligned} \|\gamma_g(A)\|_2^2 + \|\gamma_g(B)\|_2^2 - \|\gamma_g(A \cup B)\|_2^2 = \\ \lambda(A) \tilde{\gamma}_g(A)^2 + \lambda(B) \tilde{\gamma}_g(B)^2 - \lambda(A \cup B) \tilde{\gamma}_g(A \cup B)^2 \end{aligned}$$

Dieses Ähnlichkeitsmaß hat eine starke Analogie zu dem in [KLM94] propagierten Ähnlichkeitsmaß. Man vergleiche es mit Gleichung (2.14) in Abschnitt 2.4.3.2 (S. 43). Bis auf den Term, der den Rand des neuen Segments $A \cup B$ bewertet sind sie, bis auf das Vorzeichen, gleich. Das umgekehrte Vorzeichen ist leicht zu erklären: Im Kontext des Abschnittes 2.4 wird ein Energieterm minimiert und im Kontext der Mehrniveaupassung die Zielfunktion Ψ maximiert. Der Energieterm des Abschnittes 2.4.3.2 minimiert, bis eben auf die Bewertung des Randes, den Abstand einer stückweise konstanten Funktion zum Bild. Dies ist aber genau das, was die Mehrniveaupassung bezüglich der Basis \mathcal{E} tut (vgl. auch Abschnitt 6.3.1).

Bemerkung 5.3.5 (Hinweise zur Implementierung)

Mit Hilfe der Hinweise aus Abschnitt 4.3.3.2 (S. 166ff) und dem Satz 3.7.31 (S. 134) ist es möglich, ein Updateverfahren zu entwickeln, das es erlaubt, für zwei Segmente A, B einer pixelbasierten Segmentierung den Wert $\|\gamma_g(A \uplus B)\|_2^2$ in einer Laufzeit in $\mathcal{O}(1)$ aus den Werten $\gamma_g(A)$ und $\gamma_g(B)$ zu berechnen. Nach Satz 3.7.31 ist nämlich

$$\gamma_f(A \cup B) = \mathbf{E}_A^{A \cup B, A} \gamma_f(A) + \mathbf{E}_B^{A \cup B, B} \gamma_f(B)$$

Die Matrix $\mathbf{E}_A^{A \cup B, A}$ ist dabei gegeben durch $\mathbf{E}_{A^{i,j}}^{A \cup B, A} = \left\langle e_i^{A \cup B} \middle| e_j^A \right\rangle_A$ (für $1 \leq i, j, \leq n$). Da nach Definition 3.4.4 iv) $\langle \cdot | \cdot \rangle_{A \cup B} = \langle \cdot | \cdot \rangle_A + \langle \cdot | \cdot \rangle_B$ gilt, kann man die Matrix $\mathbf{E}_A^{A \cup B, A}$ aus Werten, die für die Segmente A und B abgespeichert wurden, in konstanter Laufzeit algorithmisch berechnen. Dies kann man erreichen, indem man – analog zum Abschnitt 4.3.3.2 – die Basisfunktionen e_i^A der ausgezeichneten Basis \mathcal{E}^A nach ihrer iterativen Definition gemäß dem Algorithmus 5 „abrollt“. Das heißt, wenn die für diese Berechnung relevanten Werte mit der Startsegmentierung $\{A_1, \dots, A_k\}$ gegeben sind, müssen in einer effizienten Implementierung des Algorithmus 16 (**Iterative Niveaueschmelzung**) keine Berechnungen mehr auf Pixelebene durchgeführt werden. Insbesondere ist die Laufzeit lediglich davon abhängig, wie schnell sich die Auswahl der zu vereinigenden Segmente in Schritt 4 durchführen läßt. Die Laufzeit dieses Schritts hängt stark von der Wahl des Systems der erlaubten Mengen ab.

Bemerkung 5.3.6

Die Auswahl

$$\begin{aligned} & \|\gamma_g(B_{i_0})\|_2^2 + \|\gamma_g(B_{i_1})\|_2^2 - \|\gamma_g(B_{i_0} \cup B_{i_1})\|_2^2 \\ &= \min\{\|\gamma_g(B_{j_0})\|_2^2 + \|\gamma_g(B_{j_1})\|_2^2 - \|\gamma_g(B_{j_0} \cup B_{j_1})\|_2^2 \mid B_{j_0} \cup B_{j_1} \in \mathfrak{A}, 1 \leq j_0 < j_1 \leq i\} \end{aligned}$$

in Schritt 4 des Algorithmus 16 (**Iterative Niveaoverschmelzung**) und die anschließende Vereinigung der Segmente B_{i_0} und B_{i_1} führt ein pixelbasiertes lokales Maximum einer k -Niveaupassung nicht immer in ein pixelbasiertes lokales Maximum einer $(k-1)$ -Niveaupassung über, wie das folgende Beispiel 5.3.7 zeigt.

Beispiel 5.3.7

Sei $M := [0, 8]$, $E := \mathbb{R}$ mit der Multiplikation als Skalarprodukt und V der Vektorraum der konstanten Funktionen von M nach \mathbb{R} mit Basis (1). Weiter sei $\mathfrak{A} := \{[t_1, t_2] \mid t_1, t_2 \in M\}$ und damit

$$S_3(\mathfrak{A}) = \{\{[0, n_1], [n_1, n_2], [n_3, 8]\} \mid 0 \leq n_1 \leq n_2 \leq n_3 \leq 8, n_1, n_2, n_3, \in \mathbb{N}_8 \cup \{0\}\}.$$

Sei nun $g : M \rightarrow \mathbb{N}$ ein diskretes Pixelbild mit

$$g(x) := \begin{cases} 1, & 0 \leq x < 3 \\ 2, & 3 \leq x < 4 \\ 4, & 4 \leq x < 5 \\ 2, & 5 \leq x < 8 \end{cases}$$

Abbildung 5.2 veranschaulicht g .

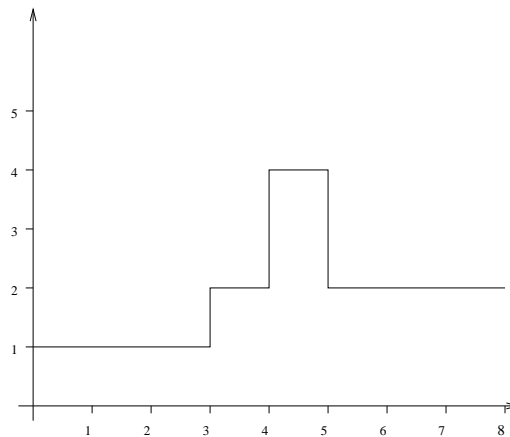


Abbildung 5.2: Veranschaulichung des diskreten eindimensionalen Pixelbildes g aus Beispiel 5.3.7.

Es ist dann offensichtlich $\{[0, 4], [4, 5], [5, 8]\}$ nicht nur ein pixelbasiertes lokales Maximum, sondern auch das globale Maximum einer 3-Niveaupassung.

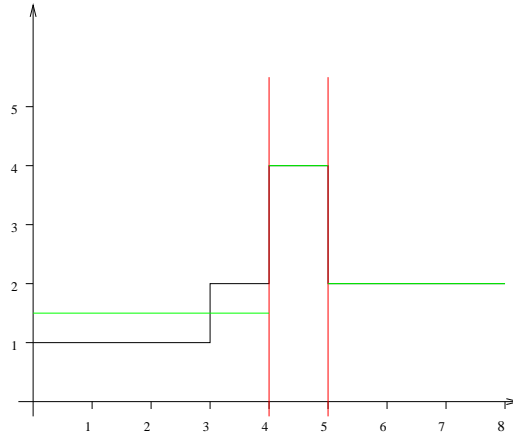


Abbildung 5.3: Globales pixelbasiertes Maximum einer 3-Niveaupassung von g gemäß Beispiel 5.3.7. Die grünen Linien skizzieren die approximierenden stückweise konstanten Funktionen $g_{[0,4]}$, $g_{[4,5]}$ bzw. $g_{[5,8]}$ (vgl. Definition 3.5.13).

Wie man leicht nachrechnet, ist

$$\begin{aligned} \|\gamma_g([0, 4])\|_2^2 &= \frac{1}{4}(3 \cdot 1 + 2)^2 = 6 + \frac{1}{4}, \\ \|\gamma_g([4, 5])\|_2^2 &= (1 \cdot 4)^2 = 16, \\ \text{und } \|\gamma_g([5, 8])\|_2^2 &= \frac{1}{3}(3 \cdot 2)^2 = 12 \end{aligned}$$

Die Abbildung 5.3 skizziert diese optimale Segmentierung.

Weiter gilt

$$\|\gamma_g([0, 5])\|_2^2 = \frac{1}{5}(3 \cdot 1 + 1 \cdot 2 + 1 \cdot 4)^2 = 16 + \frac{1}{5} \quad \text{und} \quad \|\gamma_g([4, 8])\|_2^2 = \frac{1}{4}(1 \cdot 4 + 3 \cdot 2)^2 = 25$$

und daher ist

$$\begin{aligned} &\|\gamma_g([4, 5])\|_2^2 + \|\gamma_g([5, 8])\|_2^2 - \|\gamma_g([4, 8])\|_2^2 = 28 - 25 = 3 \\ &< 6 + \frac{1}{20} = (6 + \frac{1}{4}) + 16 - (16 + \frac{1}{5}) \\ &= \|\gamma_g([0, 4])\|_2^2 + \|\gamma_g([4, 5])\|_2^2 - \|\gamma_g([0, 5])\|_2^2 \end{aligned}$$

Es ist daher

$$\begin{aligned} &\|\gamma_g([4, 5])\|_2^2 + \|\gamma_g([5, 8])\|_2^2 - \|\gamma_g([4, 8])\|_2^2 \\ &= \min\{\|\gamma_g(B_{j_0})\|_2^2 + \|\gamma_g(B_{j_1})\|_2^2 - \|\gamma_g(B_{j_0} \cup B_{j_1})\|_2^2 \mid B_{j_0} \cup B_{j_1} \in \mathfrak{A}, 1 \leq j_0 < j_1 \leq 3\} \end{aligned}$$

Das heißt, der Algorithmus 16 (**Iterative Niveauerschmelzung**) würde in der Zeile 4 die Segmente $[4, 5]$ und $[5, 8]$ zur Vereinigung auswählen, um so die Segmentierung $\mathbf{B} = \{[0, 4], [4, 8]\}$ aus $S_2(\mathfrak{A})$ zu erhalten (vgl. Abbildung 5.4).

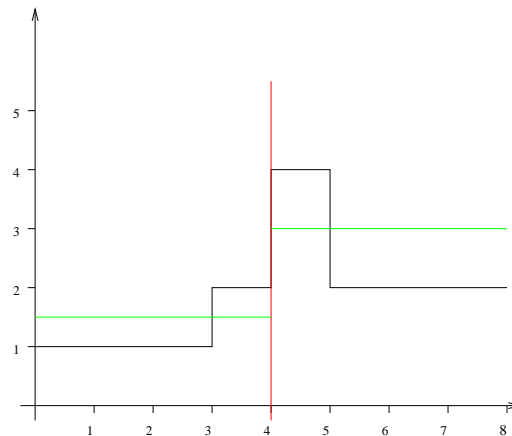


Abbildung 5.4: Ergebnis der Verschmelzung der Segmente $[4, 5]$ und $[5, 8]$: Eine Segmentierung mit zwei Segmenten, die kein lokales pixelbasiertes Maximum ist.

Es ist aber $\Psi(\mathbf{B}) = \|\gamma_g([0, 4])\|_2^2 + \|\gamma_g([4, 8])\|_2^2 = 31 + \frac{1}{4}$ kein pixelbasiertes lokales Maximum der Funktion $\Psi : S_2(\mathfrak{A}) \rightarrow \mathbb{R}, \{B_1, B_2\} \mapsto \|\gamma_g(B_1)\|_2^2 + \|\gamma_g(B_2)\|_2^2$. Denn es gilt $\|\gamma_g([0, 3])\|_2^2 + \|\gamma_g([3, 8])\|_2^2 = 3 + \frac{1}{5}(2 + 4 + 6)^2 = 31 + \frac{4}{5} > 31 + \frac{1}{4} = \Psi(\mathbf{B})$ und $\{[0, 3], [3, 8]\} \in \mathfrak{U}_p(\mathbf{B})$. Damit existiert eine pixelbasierte Segmentierung aus der pixelbasierten Umgebung von \mathbf{B} , die der Funktion Ψ , ein höheren Wert zuteilt. Also ist \mathbf{B} kein pixelbasiertes lokales Maximum der Funktion Ψ .

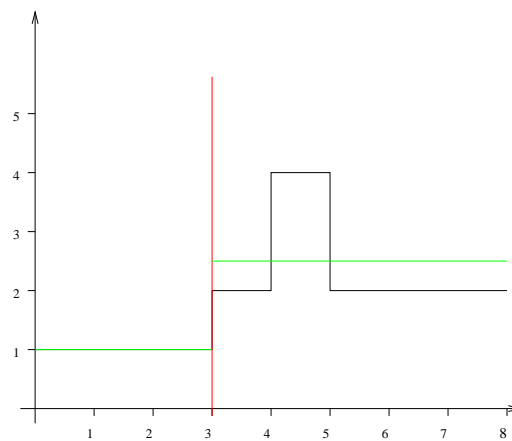


Abbildung 5.5: Pixelbasiertes lokales Maximum einer 2-Niveaupassung von g gemäß Beispiel 5.3.7.

Beispiel 5.3.8 (Konstante Niveaureduktion)

Betrachtet wird ein diskretes Farbpixel $g : M \rightarrow \{0, \dots, 255\}^3$ im RGB-Farbmodell. Weiter sei $E := \mathbb{R}^3$ mit dem Standardskalarprodukt und \mathcal{E} sei eine erlaubte Basis zur Farbbildsegmentierung mit

$$e_1(x) := (1, 0, 0), e_2(x) := (0, 1, 0), e_3(x) := (0, 0, 1)$$

Die Abbildung 5.6 zeigt ein Ergebnis des Algorithmus 16 (**Iterative Niveaoverschmelzung**) für das Mengensystem der pixelzusammenhängenden Mengen. Das Eingabebild, das nach der Vorsegmentierung mit dem Algorithmus 13 (**Initiales Farbschwellwertverfahren**) mit einem Schwellwert pro Farbkanal aus 38 Segmenten besteht, wird auf drei Segmente reduziert. Dabei werden die roten Buchstaben – trotz des deutlichen Farbunterschieds – wegen ihrer geringen Fläche mit dem Segment des gelben Schildhintergrunds vereinigt.



Abbildung 5.6: Eingabebild und Ergebnis des Algorithmus 16 (**Iterative Niveaoverschmelzung**) gemäß Beispiel 5.3.8.

Beispiel 5.3.9 (Affine Niveaureduktion)

Es sei \mathcal{E} eine Basis zur Farbbildsegmentierung mit

$$\begin{aligned} e_1(x) &:= (1, 0, 0), e_2(x) := (0, 1, 0), e_3(x) := (0, 0, 1), \\ e_4(x) &:= (x, 0, 0), e_5(x) := (0, x, 0), e_6(x) := (0, 0, x) \end{aligned}$$

Wie man leicht sieht, ist \mathcal{E} eine erlaubte Basis zur Farbbildsegmentierung gemäß Definition 4.6.13 (S. 198). Betrachtet wird wieder das System der pixelzusammenhängenden Mengen.

Weiter sei $g : M \rightarrow \{0, \dots, 255\}^3$ das diskrete Farbpixelbild im RGB-Farbmodell, das durch Abbildung 5.7 skizziert wird. Die Abbildung zeigt den Buchstaben „W“, auf dessen Oberfläche sich der Farbwert affin von links unten nach rechts oben zu ändern scheint. Das Bild wurde mit einem Textverarbeitungsprogramm erstellt, so daß man davon ausgehen kann, daß der Farbverlauf tatsächlich mit Hilfe einer affinen Funktion generiert wurde.

In der Abbildung 5.8 werden die in Zeile 6 des Algorithmus 16 (**Iterative Niveaoverschmelzung**) gültigen Zwischenergebnisse des Algorithmus dargestellt. Man könnte die Abbildungen auch als Endergebnisse der Algorithmenaufrufe

$$\text{Iterative Niveaoverschmelzung}(g, \{A_1, \dots, A_{17}\}, i)$$

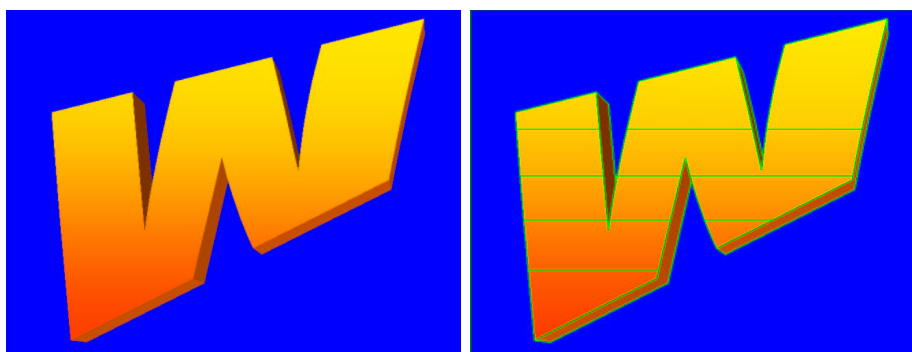


Abbildung 5.7: Eingabebild für den Algorithmus 16 (**Iterative Niveaverschmelzung**). Auf der linken Seite ist das Originalbild zu sehen, auf der rechten Seite die Startsegmentierung für den Algorithmus. Dabei sind die Segmentgrenzen grün eingezeichnet. Dargestellt wird die approximierende Funktion $\sum_{i=1}^{17} g_{A_i}$.

mit $i = 17, \dots, 6$ sehen. Das Resultat mit 4 Segmenten an den Seiten des Buchstabens, einem Hintergrundniveau und einem Oberflächensegment auf dem „W“ erscheint als Endergebnis eines Segmentierungsverfahrens sinnvoll.

Bemerkung 5.3.10

Das vorherige Beispiel 5.3.9 relativiert die Aussage von Bemerkung 5.3.6. Es zeigt, daß der Algorithmus für praktische Fälle oft auch ohne eine korrekte Nachbehandlung des Ergebnis z. B. durch den Algorithmus 6 (**k-Niveaupassung**) eine optisch ausreichend gute Segmentierung mit der gewünschten Anzahl von Segmenten liefert. Man kann zwar nicht garantieren, daß das Ergebnis ein pixelbasiertes lokales Maximum einer k -Niveaupassung ist, aber in der Praxis hat sich gezeigt, daß der Verzicht auf den Algorithmus 6 (**k-Niveaupassung**) zur Nachbesserung des Ergebnisses einen erheblichen Laufzeitgewinn unter der Duldung optisch zu verschmerzender Fehler bedeutet (vgl. auch [HFKL01]). Der Laufzeitgewinn bei diesem System der pixelbasierten Segmentierungen geht darauf zurück, daß man nur einmal algorithmisch feststellen muß, welches Segment mit welchem benachbart ist. Danach sind keine Berechnungen mehr auf Pixelebene erforderlich (vgl. Bemerkung 5.3.5).

Unter der Annahme, daß alle Nachbarn eines jeden Segments schon bestimmt sind, kann man die Laufzeit einer Schleife für die Reduktion von k auf $k - 1$ Niveaus mit $\mathcal{O}(k^2)$ abschätzen. Da die Berechnung der Ähnlichkeit zweier Segmente mit Bemerkung 5.3.5 in $\mathcal{O}(1)$ geschehen kann und jedes der k Segmente höchstens k Nachbarsegmente haben kann, ist die Ähnlichkeit höchstens k^2 -mal zu überprüfen. Daß k im Vergleich zur Anzahl der Pixel in $\mathbb{P}(M)$ in der Regel sehr klein ist, sollte den Geschwindigkeitsvorteil noch deutlicher machen.



Abbildung 5.8: Verlauf des Algorithmus 16 (**Iterative Niveaueschmelzung**). Dargestellt ist jeweils die in Zeile 6 des Algorithmus gültige Segmentierung. Dabei sind die Segmentgrenzen grün eingezeichnet.

5.4 Schätzung der Anzahl der Segmente

Wie zuvor schon mehrmals erwähnt, ist die Schätzung der Anzahl der Segmente in einem Bild eine Mehrzieloptimierung: Auf der einen Seite ist die Segmentanzahl, auf der anderen Seite ein Approximationsfehler zu minimieren. Setzt man den zu minimierenden Approximationsfehler als $\|\cdot\|_M$, so wird der Fehler betrachtet, der bei der Mehrniveaupassung minimiert wird. Daher wird dieser Approximationsfehler bei sinkender Anzahl von Segmenten monoton steigen. Im ersten Abschnitt dieses Unterkapitels wird ein Algorithmus vorgestellt, der den größten Anstieg dieses Fehlers als Schätzgrundlage hat.

Dabei sollte man an dieser Stelle eigentlich, statt von der Schätzung der Anzahl der Segmente, von der Schätzung der Anzahl der optimal optisch zu trennenden Segmente im Bild sprechen. Die Trennbarkeit kann man, wie es z. B. Nobuyuki Otsu (vgl. [Ots79] oder Abschnitt 2.3.1.1) tut, mit Hilfe der Varianz messen. Ein an diesem Ansatz entlehnter Fehler wird im zweiten Abschnitt vorgestellt.

Im darauffolgenden Abschnitt wird ein Schätzalgorithmus vorgestellt, der auf der Basis dieser Fehlerfunktion die Anzahl der zu trennenden Objekte als Pareto-Optimum einer Zweizieloptimierung schätzt.

5.4.1 Der Approximationsfehler der Mehrniveaupassung

Der naheliegendste Approximationsfehler, den man betrachten kann, ist derjenige bezüglich dessen optimiert wird. Betrachtet werden soll daher in diesem Abschnitt der Approximationsfehler $\|\cdot\|_M$. Dabei ist nach Lemma 3.5.5 (S. 94) für eine Segmentierung $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ der Fehler $\|f - \sum_{A \in \mathbf{A}} f_A\|_M^2 = \|f\|_M^2 - \sum_{A \in \mathbf{A}} \|\gamma_f(A)\|_2^2$. Daher genügt es, den Wert $\sum_{A \in \mathbf{A}} \|\gamma_f(A)\|_2^2$ zu betrachten, der sowohl bei der k -Niveaupassung als auch bei der Niveaureduktion gemäß Algorithmus 16 (**Iterative Niveauerschmelzung**) ohnehin berechnet werden muß. Es ist daher nicht nötig, zur Berechnung des Approximationsfehlers weitere Algorithmik heranzuziehen. Insbesondere ist es also nicht nötig, zur Berechnung des Approximationsfehlers $\|\cdot\|_M$ die Pixel des gesamten Bildes zu durchlaufen.

Da bezüglich des Approximationsfehlers $\|\cdot\|_M$ in der Zielfunktion der k -Niveaupassung optimiert wird, ist es offensichtlich, daß der Approximationsfehler für optimale Segmentierungen bei sinkender Anzahl monoton steigt. Daher ist fast jeder Punkt der Mehrzieloptimierung (u_{no}, u_{approx}) ein Pareto-Optimum. Unter allen diesen Punkten zeichnet sich jedoch derjenige des stärksten Fehleranstiegs bei Reduktion um ein Segment aus. Der Algorithmus 17 (**Stärkster Anstieg**) aus Seite 236 berechnet diesen Punkt. Das anschließende Beispiel 5.4.1 zeigt jedoch die Grenzen eines solchen Ansatzes: Bei diesem Approximationsfehler werden große Segmente besonders stark gewichtet.

Beispiel 5.4.1

Gegeben sei der Kontext des Abschnitts 4.5. Das heißt, es wird ein diskretes Grauwertpixelbild betrachtet und es ist $E := \mathbb{R}$ mit dem Standardskalarprodukt. Die Approximation

Algorithmus 17 (Stärkster Anstieg) Schätzung der Anzahl der Segmente über den stärksten Anstieg des Approximationsfehler

Input: Diskretes Pixelbild $g : M \rightarrow \mathbb{R}^d$,
Startanzahl k von Segmenten

Output: geschätzte Anzahl von Objekten im diskreten Pixelbild

```

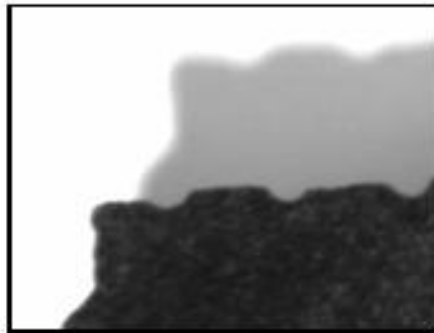
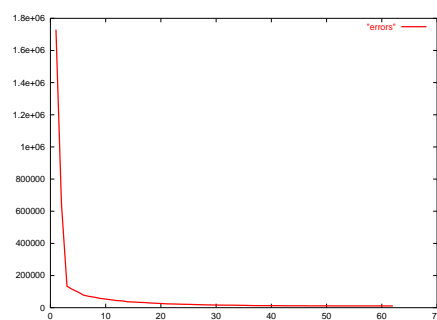
1: for  $i = k$  abwärts bis  $i = 1$  do
2:   Berechne ein lokales pixelbasiertes Maximum  $\mathbf{A}_i \in S_i(\mathfrak{A})$  einer  $i$ -Niveaupassung des
   Bildes  $g$ 
3:    $\text{error}_i := \left\| g - \sum_{A \in \mathbf{A}_i} g_A \right\|_M^2$ 
4: end for
5:  $\max \leftarrow -1$ 
6:  $\text{estimatedSegments} \leftarrow k$ 
7: for  $i = k - 1$  abwärts bis  $i = 1$  do
8:   if  $\text{error}_{i-1} - \text{error}_i \geq \max$  then
9:      $\text{estimatedSegments} \leftarrow i$ 
10:     $\max \leftarrow \text{error}_{i-1} - \text{error}_i$ 
11:   end if
12: end for
13: return  $\text{estimatedSegments}$ 

```

erfolgt mit konstanten Funktionen bezüglich des Mengensystems der grauwertverbundenen Mengen.

Das zu segmentierende Bild wurde schon im Abschnitt 2.3.1.1 als Beispiel einer Segmentierung über einen Schwellwert mit Hilfe der Varianzanalyse präsentiert (vgl. Abbildung 2.7). Die Abbildung 5.9 (oben) zeigt noch einmal das Eingabebild g , das Bild einer Granittranche vor weißem Hintergrund, die einen Schatten auf diesen Hintergrund wirft. Abbildung 5.9 (Mitte) zeigt den Verlauf der durch den Algorithmus 17 (**Stärkster Anstieg**) definierten Folge $(\text{error}_i)_{i=1}^6$. Es ist deutlich zu sehen, daß sich der stärkste Anstieg (bei fallender Segmentanzahl) beim Übergang von zwei zu einem Segmenten befindet. Daher liefert der Algorithmus zwei zu trennende Segmente als Ergebnis der Schätzung. Abbildung 5.9 (unten) zeigt ein pixelbasiertes lokales Maximum einer 2-Niveaupassung.

Bei genauerer Betrachtung fällt auf, daß auch der Anstieg des Fehlers bei der Reduktion von drei auf zwei Segmente im Verhältnis zu den übrigen Anstiegen sehr stark ist. Generell ist zu bemerken, daß der Approximationsfehler $\| \cdot \|_M$ stark von der Größe der Objekte abhängt. Daher ist die Überlegung, diesen Fehler relativ zur Größe der jeweiligen Segmente zu berechnen, wie es im nächsten Abschnitt geschieht, berechtigt.

Eingabebild g 

Entwicklung des Fehlers error_i im Algorithmus 17 (**Stärkster Anstieg**):
Das Ergebnis des Algorithmus ist 2.



Das Ergebnis einer pixelbasierten 2-Niveaupassung mit konstanten Funktionen bezüglich des Mengensystems der grauwertverbundenen Mengen gemäß Beispiel 5.4.11

Abbildung 5.9: Ergebnis des Algorithmus 17 (**Stärkster Anstieg**) auf einem Bild gemäß Beispiel 5.4.1.

5.4.2 Der größte normalisierte Fehler

Der Ansatz des größten normalisierten Fehlers verbindet zwei grundlegende Ideen:

- i. Der Approximationsfehler sollte unabhängig von der Größe der Segmente sein.
- ii. Es sollte der Fehler des fehlerhaftesten Segments gezählt werden.

Die Motivation der zweiten Idee ergibt sich aus folgender Überlegung: Ist \mathbf{A} ein lokales Optimum einer k -Niveaupassung, so wird die Mehrniveaupassung für jedes Segment $A \in \mathbf{A}$ den Fehler $\|f - f_A\|_A$ minimal halten. Angenommen, es befinden sich l durch konstante Niveaus deutlich trennbare Segmente im Bild. Ist nun $k > l$, so ist der Fehler $\|f - f_A\|_A$ für jedes $A \in \mathbf{A}$ relativ gering. Eine Niveaureduktion über den Algorithmus 16 (**Iterative Niveaverschmelzung**) wird dann iterativ Segmente verschmelzen, die sich sehr ähnlich sind, d. h. die einen ähnlichen Niveauverlauf haben. Dies ist nur möglich, wenn mindestens genauso viele Segmente zur Segmentierung zur Verfügung stehen, wie trennbare Objekte im Bild sind. Ist dies nicht der Fall, so müssen zwei Segmente, die sich nicht ähnlich sind, miteinander verbunden werden. Steht dann ein Segment weniger als optisch trennbare Objekte im Bild zur Verfügung, so zeichnen sich die mit Hilfe des Algorithmus 16 (**Iterative Niveaverschmelzung**) als Initialisierung berechneten lokalen Minima einer k -Niveaupassung in der Regel dadurch aus, daß sich $k - 1$ Segmente gebildet haben, auf denen der Approximationsfehler $\|\cdot\|_A$ gering ist, und das letzte Segment den Rest des Definitionsbereichs abdeckt. Daher genügt es, den Fehler des fehlerhaftesten Segments zu betrachten.

Salopp formuliert sollen also beginnend mit einer ausreichenden Anzahl von Segmenten, diese solange reduziert werden, bis die Anzahl der Segmente nicht mehr ausreicht und ein Segment zwei deutlich im Bild trennbare Objekte überdecken muß.

Diese Überlegungen führen zu dem folgendem Fehlermaß.

Definition 5.4.2 (Größter normalisierter Fehler)

Für eine Mengen $A \in \mathfrak{A}$ und $f, g \in L^2(\mathcal{B}(M), \mathcal{B}(E))$ heiße der Fehler

$$n_A(f, g) := \begin{cases} \frac{1}{\lambda(A)} \|f - g\|_A^2 & , \lambda(A) > 0 \\ 0 & , \text{sonst} \end{cases}$$

der *normalisierte Fehler auf A zwischen f und g* . Für eine Segmentierung $\mathbf{A} \in \mathfrak{S}_k(\mathfrak{A})$ heiße der Fehler

$$n_{\mathbf{A}}(f, g) := \max_{A \in \mathbf{A}} n_A(f, g)$$

der *größte normalisierte Fehler der Segmentierung \mathbf{A} zwischen f und g* .

Bemerkung 5.4.3

Betrachtet man das Mengensystem der grauwertverbundenen Mengen und die Klasse der konstanten Funktionen, wie dies in Abschnitt 4.5 geschieht, so verbirgt sich hinter dem Begriff des normalisierten Fehlers nichts anderes als die bedingte Varianz. Ersetzt man in der Gleichung (2.3) des Abschnitts 2.3.1.1 (S. 21) die durch die Varianzanalyse definierten Begriffe durch die entsprechenden Definitionen der vektoriellen Mehrniveaupassung, so gilt mit $A_k := \biguplus_{i=0}^k P_i$ (vgl. Abschnitt 4.5):

$$\sigma_0(k)^2 = n_{A_k}(f, f_{A_k})$$

Bemerkung 5.4.4

Im Gegensatz zu dem in Abschnitt 5.4.1 betrachteten Approximationsfehler $\|\cdot\|$ ist es beim größten normalisierten Fehler durchaus möglich, daß der Approximationsfehler bei abnehmender Anzahl von Segmenten sinkt. Die Begründung liegt in der Normalisierung des Fehlers: Kleine Segmente mit starken Störungen werden stärker bewertet als beim Approximationsfehler $\|\cdot\|_M$. Wird so ein kleines Segment zu einem größeren Segment hinzugefügt, so kann der normalisierte Approximationsfehler des zusammengeführten Segments kleiner sein als der des gestörten kleineren Ausgangssegments. Daher kann der größte normalisierte Approximationsfehler bei einer Niveaureduktion kleiner werden, wie die Abbildung 5.10 am Beispiel eines Bildes mit eindimensionalem Definitionsbereich zeigt.

Lemma 5.4.5

Es gilt

$$n_A(f, f_A) = \frac{1}{\lambda(A)} \|f\|_A^2 - \|\gamma_f(A)\|_2^2$$

Beweis:

folgt mit Lemma 3.5.5 (S. 94) und Definition 3.5.13 (S. 99)

Bemerkung 5.4.6

Das Lemma 5.4.5 ist die Entsprechung des aus der Wahrscheinlichkeitsrechnung bekannten Zusammenhangs $\text{VAR}[X] = \text{E}[(X - \text{E}[X])^2] = \text{E}[X^2] - \text{E}[X]^2$.

Für diskrete Pixelbilder g erlaubt das Lemma 5.4.5 eine Berechnung des Wertes $n_A(g, g_A)$, ohne alle Pixel des Segments A zu durchlaufen. Vielmehr sollte ein Algorithmus, der eine optimale Segmentierung $\mathbf{A} = \{A_1, \dots, A_k\}$ berechnet, bei jedem Pixelaustausch der Segmente die Werte $\|g\|_A^2 = \sum_{p \in \mathbb{P}(A)} \|g\|_p^2$ und $\|\gamma_g(A)\|_2^2$ aus zuvor berechneten Werten aktualisieren. Zur Aktualisierung des Wertes $\|\gamma_g(A)\|_2^2$ in einer Update-Strategie sei auf Abschnitt 4.3.3.2 verwiesen.

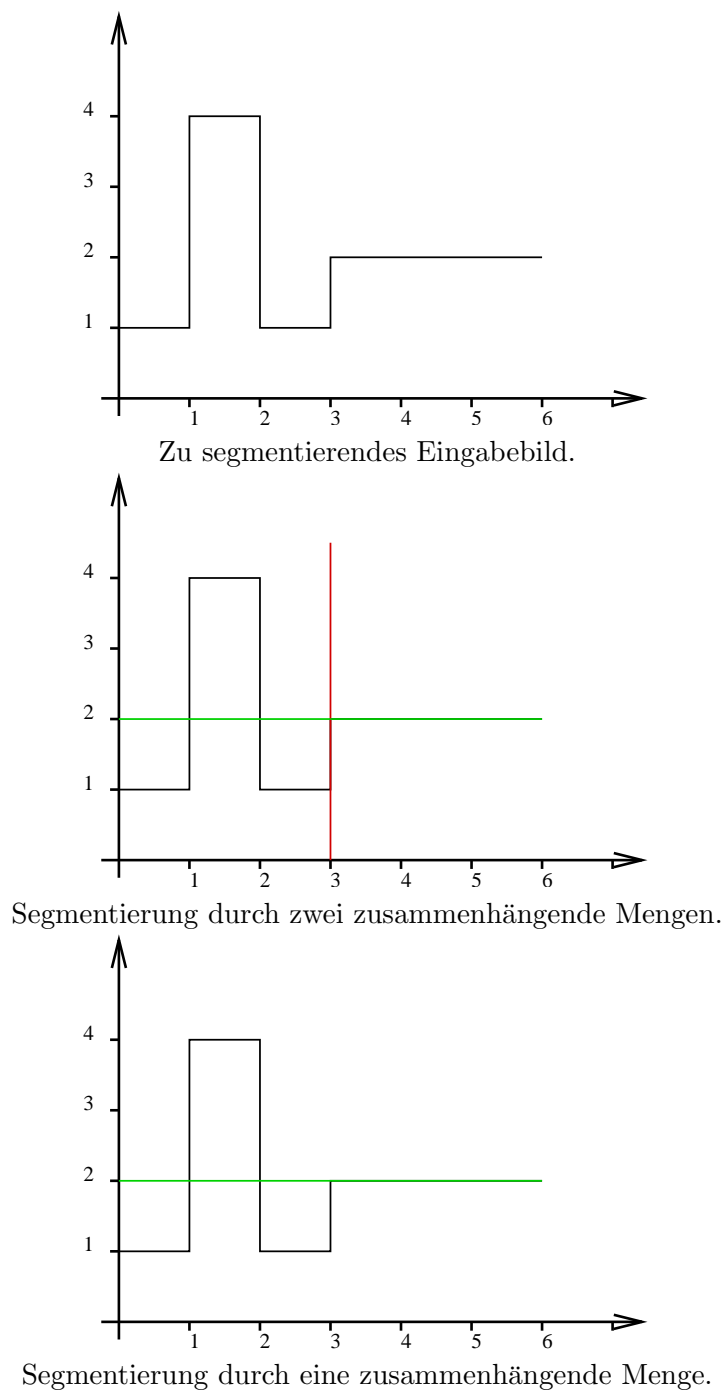


Abbildung 5.10: Der größte normalisierte Fehler der Segmentierung des Bildes (oben) durch zwei zusammenhängende Mengen mit konstanten Funktionen ist 2 (Mitte), der größte normalisierte Fehler der Segmentierung durch eine zusammenhängende Menge ist 1 (unten). Damit ist der größte normalisierte Fehler der Segmentierung mit weniger Segmenten in diesem Fall kleiner.

5.4.3 Ein Algorithmus zur Schätzung der Anzahl der Segmente über den größten normalisierten Fehler

Mit Hilfe des im vorangegangenen Abschnitts 5.4.2 eingeführten größten normalisierten Fehlers ist es nun möglich, ein Pareto-Optimum einer Mehrzieloptimierung zu berechnen, das einer etwas differenzierteren Bedingung gehorcht als das des Algorithmus 17 (**Stärkster Anstieg**). Die Heuristik, die der folgende Algorithmus 18 (**Anzahlschätzer**) implementiert, liefert in vielen Fällen ein optisch schöneres Ergebnis, als die des Algorithmus 17 (**Stärkster Anstieg**).

Algorithmus 18 (Anzahlschätzer) Schätzung der Anzahl der Segmente über den größten normalisierten Fehler

Input: Diskretes Pixelbild $g : M \rightarrow \mathbb{R}^d$,
Startanzahl k von Segmenten

Output: geschätzte Anzahl von Objekten im diskreten Pixelbild

```

1: for  $i = k$  abwärts bis  $i = 1$  do
2:   Berechne ein lokales pixelbasiertes Maximum  $\mathbf{A} \in S_i(\mathfrak{A})$  einer  $i$ -Niveaupassung des
   Bildes  $g$ 
3:    $g_{\mathbf{A}} \leftarrow \sum_{A \in \mathbf{A}} g_A$ 
4:    $\text{error}_i := n_{\mathbf{A}}(g, g_{\mathbf{A}})$ 
5: end for
6:  $\text{estimatedSegments} \leftarrow k$ 
7:  $\text{criticalValue} \leftarrow \text{error}_k$ 
8: for  $i = k - 1$  abwärts bis  $i = 1$  do
9:   if  $\text{error}_i \leq \text{criticalValue}$  then
10:     $\text{estimatedSegments} \leftarrow i$ 
11:   end if
12:   if  $\text{error}_{i-1} < \text{error}_i$  und  $\text{error}_i > \text{criticalValue}$  then
13:     $\text{criticalValue} \leftarrow \text{error}_i$ 
14:   end if
15: end for
16: return  $\text{estimatedSegments}$ 

```

Bemerkung 5.4.7

Der Algorithmus 18 (**Anzahlschätzer**) wurde für Mehrniveaupassungen mit konstanten Funktionen bereits in [Han00] vorgeschlagen.

Sind für $i \in \{1, \dots, k\}$ die Werte $error_i$ und $criticalValue(i) = \max\{error_j \mid j > i\}$ wie im Algorithmus definiert, so berechnet der Algorithmus 18 (**Anzahlschätzer**) den Wert

$$\min\{i \in \{1, \dots, k\} \mid error_i \leq criticalValue(i)\}$$

Satz 5.4.8

Mit den Bezeichnungen aus Bemerkung 5.2.6 und des Algorithmus 18 (**Anzahlschätzer**) seien die Funktionen $u_{no}, u_{approx} : \{1, \dots, k\} \rightarrow \mathbb{R}_+$ definiert durch $u_{no}(i) := i$ und $u_{approx}(i) := error_i$. Weiter sei N , die Menge der Nebenbedingungen, leer und $U = (u_{no}, u_{approx})$ analog zu Definition 5.2.1. In diesem Kontext gilt:

Ist i_0 das Ergebnis des Algorithmus 18 (**Anzahlschätzer**), so ist $(i_0, error_{i_0})$ ein Pareto-Optimum der Zweizieloptimierung (U, E) .

Beweis:

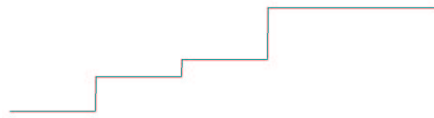
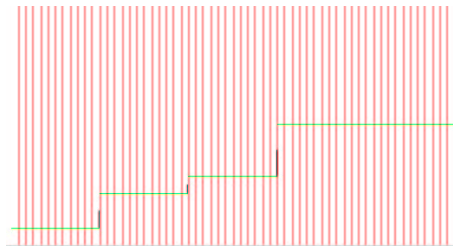
Der Entscheidungsraum X der Definition 5.2.1 ist $X = \{(i, error_i) \mid i \in \{1, \dots, k\}\}$. Ein Tupel $(i_0, error_{i_0})$ ist genau dann kein Pareto-Optimum, wenn ein Tupel $(i, error_i)$ existiert, mit $i < i_0$ und $error_i \leq error_{i_0}$. Nun betrachtet der Algorithmus der Reihe nach die Tupel $(k, error_k)$ bis $(1, error_1)$. Ist $(i_0, error_{i_0})$ kein Pareto-Optimum, so ist aber der Wert $error_i$ des Tupels $(i, error_i)$, das die Pareto-Optimalität verletzt, auch kleiner als der Wert $criticalValue$ zum Zeitpunkt, an dem das Ergebnis auf i_0 gesetzt wird. Da der Algorithmus 18 (**Anzahlschätzer**) nach Bemerkung 5.4.7 jedoch den Wert $\min\{i \in \{1, \dots, k\} \mid error_i \leq criticalValue(i)\}$ als Ergebnis liefert, folgt die Behauptung. \square

5.4.4 Beispiele**Beispiel 5.4.9**

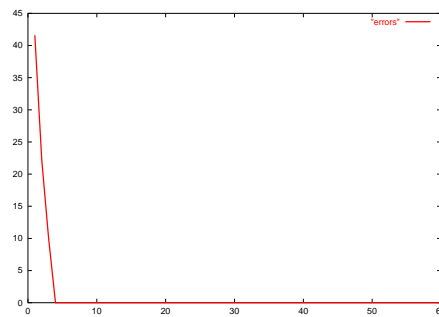
Das erste Beispiel, das das Vorgehen des Algorithmus 18 (**Anzahlschätzer**) illustrieren soll, greift das Beispiel 4.4.3 aus Abschnitt 4.4.2 auf. Dort war $M = [0, 500]$, $E := \mathbb{R}$ mit der Multiplikation als Skalarprodukt, $\mathcal{E} := (1)$ und $g : [0, 500] \rightarrow \mathbb{R}_+$ war gegeben durch

$$g(x) = \begin{cases} 0 & , 0 \leq x < 100 \\ 40 & , 100 \leq x < 200 \\ 60 & , 200 \leq x < 300 \\ 120 & , 300 \leq x \leq 500 \end{cases}$$

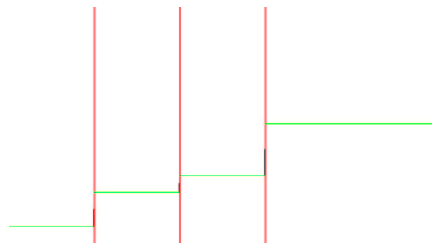
Abbildung 5.11 zeigt das diskrete Pixelbild g , die Startsegmentierung bezüglich des Mengensystems der pixelzusammenhängenden Intervalle mit 60 Segmenten, die Entwicklung des Fehler $error_i$ im Algorithmus 18 (**Anzahlschätzer**) und das Ergebnis einer pixelbasierten Mehrniveaupassung in diesem Kontext mit dem Ergebnis des Algorithmus als vorgegebene Anzahl von Niveaus. Die Abbildung 5.12 zeigt das diskrete Pixelbild \tilde{g} das aus g durch Überlagerung eines auf $[-20, 20]$ uniformen Rauschens entstand. Dort sind ebenfalls die Startsegmentierung, die Entwicklung des Fehler $error_i$ und das Ergebnis abgebildet.

Skizze von g .

Startsegmentierung mit 60 Segmenten.

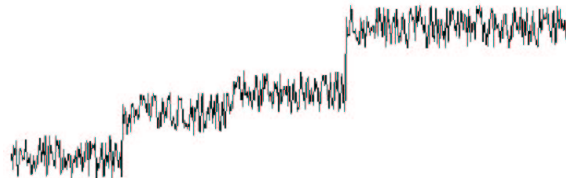
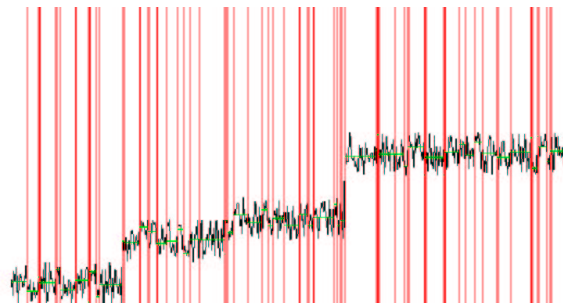


Entwicklung des Fehlers error_i im Algorithmus 18 (**Anzahlschätzer**):
Das Ergebnis des Algorithmus ist 4.

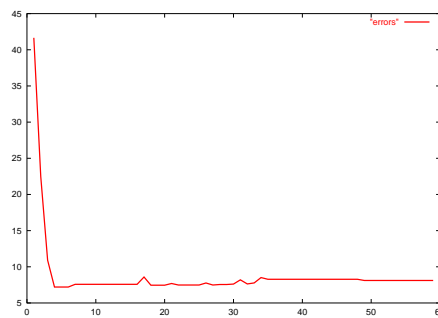


Das Ergebnis einer pixelbasierten 4-Niveaupassung gemäß Beispiel 5.4.9.

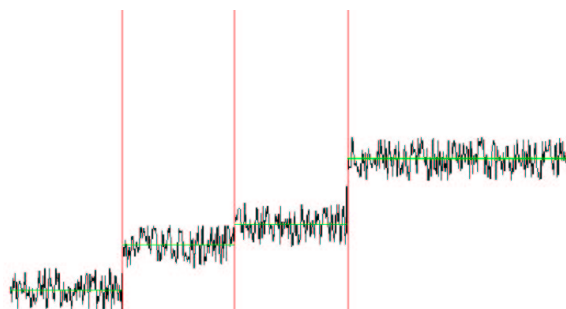
Abbildung 5.11: Ergebnis des Algorithmus 18 (**Anzahlschätzer**) auf einem diskreten Pixelbild gemäß Beispiel 5.4.9.

Skizze von \tilde{g} .

Startsegmentierung mit 60 Segmenten.



Entwicklung des Fehlers error_i im Algorithmus 18 (**Anzahlschätzer**):
Das Ergebnis des Algorithmus ist 4, wie im unverrauschten Fall.



Das Ergebnis einer pixelbasierten 4-Niveaupassung gemäß Beispiel 5.4.9.

Abbildung 5.12: Ergebnis des Algorithmus 18 (**Anzahlschätzer**) auf einem verrauschten Bild gemäß Beispiel 5.4.9.

Beispiel 5.4.10

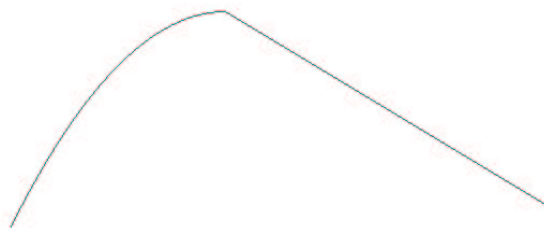
Das zweite Beispiel greift das Beispiel 4.4.4 des Abschnitts 4.4.3 auf. Dort war $M := [0, 500]$, $E := \mathbb{R}$ mit der Multiplikation als Skalarprodukt, $\mathcal{E} := (1, \text{id}, \text{id}^2)$ und das diskrete Pixelbild $g : [0, 500] \rightarrow \mathbb{R}_+$ war gegeben durch

$$g(x) = \begin{cases} -\frac{1}{200}\text{pix}(x)^2 + 2\text{pix}(x) & , 0 \leq x < 200 \\ 320 - \frac{3}{5}\text{pix}(x) & , 200 \leq x \leq 500 \end{cases}$$

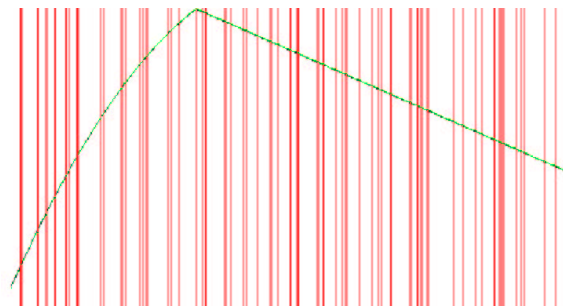
wobei $\text{pix} : \mathbb{R} \rightarrow \mathbb{R}$ durch $\text{pix}(x) := \text{round}(x - \frac{1}{2})$ gegeben ist. Betrachtet wird das Mengensystem der pixelzusammenhängenden Intervalle in $[0, 500]$.

In der Abbildung 5.13 (oben) wird dieses eindimensionale Bild skizziert. Die Entwicklung des Fehlers error_i für $i = 1, \dots, 60$ ist in der Mitte der Abbildung dargestellt. Man beachte, daß für fallende i mit $i \geq 2$ der Fehler $n_{\mathbf{A}}(g, g_{\mathbf{A}})$ ebenfalls fällt. Daß der Fehler für mehr als 2 Segmente nicht null ist, kann man auf die round-Operation und numerische Instabilitäten zurückführen, die sich bei größeren Segmenten aufheben. Der Algorithmus 18 (**Anzahlschätzer**) liefert als geschätzte Anzahl von Objekten im Bild dem Wert 2. Unten in der Abbildung 5.13 ist dann die optimale Segmentierung mit zwei Segmenten skizziert.

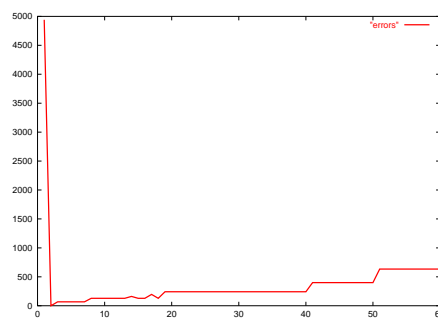
In der Abbildung 5.14 wurde die Funktion g mit einem starken uniformen Rauschen überlagert. Trotzdem kommt der Algorithmus 18 (**Anzahlschätzer**) bei der Analyse der Werte $(\text{error}_i)_{i=1}^{60}$, die in der Mitte der Abbildung skizziert sind, zum selben Ergebnis wie im unverrauschten Fall. Der Fehler $n_{\mathbf{A}}(g, g_{\mathbf{A}})$ hat sich allerdings für alle lokal-optimalen Segmentierungen drastisch erhöht.



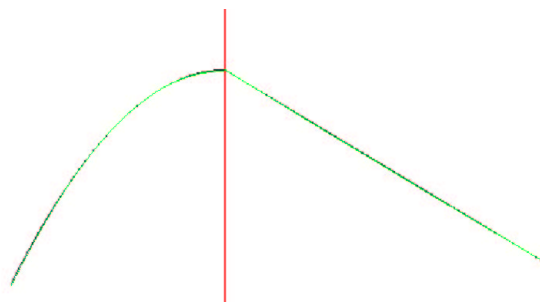
Skizze des diskreten Pixelbildes g aus Beispiel 5.4.10.



Startsegmentierung mit 60 Segmenten.

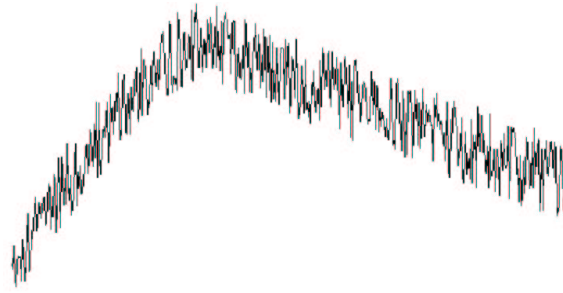
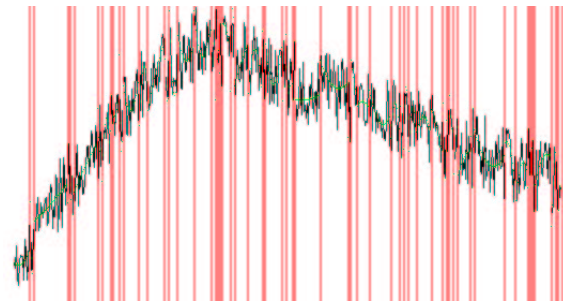


Entwicklung des Fehlers $error_i$ im Algorithmus 18 (**Anzahlschätzer**):
Das Ergebnis des Algorithmus ist 2.

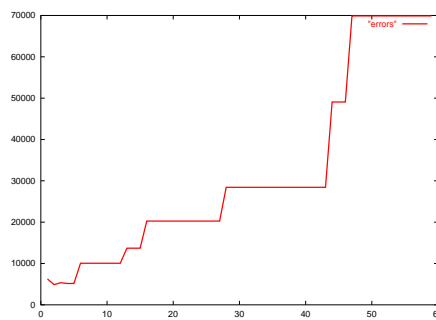


Das Ergebnis einer pixelbasierten 2-Niveaupassung gemäß Beispiel 5.4.10.

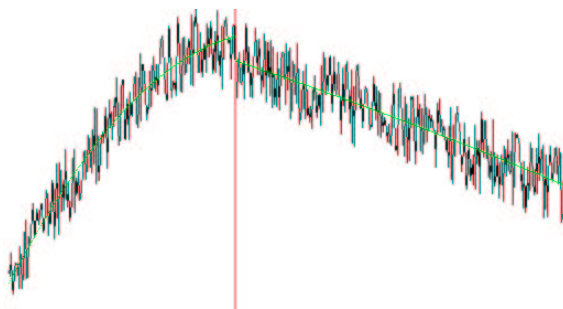
Abbildung 5.13: Ergebnis des Algorithmus 18 (**Anzahlschätzer**) auf dem diskreten Pixelbild g gemäß Beispiel 5.4.10.

Skizze des verrauschten Bildes \tilde{g} .

Startsegmentierung mit 60 Segmenten.



Entwicklung des Fehlers error_i im Algorithmus 18 (**Anzahlschätzer**):
Das Ergebnis des Algorithmus ist 2.

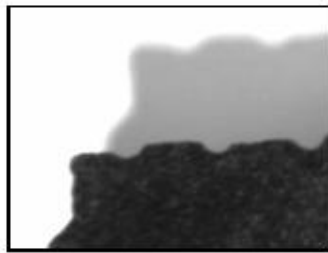
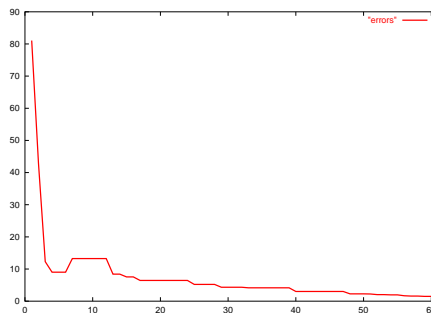


Das Ergebnis einer pixelbasierten 2-Niveaupassung gemäß Beispiel 5.4.10.

Abbildung 5.14: Ergebnis des Algorithmus 18 (**Anzahlschätzer**) auf dem uniform verrauschten Bild gemäß Beispiel 5.4.10.

Beispiel 5.4.11

Das dritte Beispiel nimmt Bezug auf das Beispiel 5.4.1, das die Performanz des Algorithmus 17 (**Stärkster Anstieg**) verdeutlichen sollte. Der Algorithmus schätzte – bei Segmentierungen durch grauwertverbundene Mengen und konstanter Approximation auf den Segmenten – die Anzahl der Objekte im Bild auf zwei. Der Algorithmus 18 (**Anzahlsschätzer**), der auf der Grundlage des größten normalisierten Fehlers, statt auf der des Fehlers $\|f - \cdot\|_M$ arbeitet, liefert ein optisch ansprechenderes Ergebnis, nämlich drei zu trennende Segmente: den weißen Hintergrund, den grauen Schatten und die dunklere Granittranche.

Eingabebild g .

Entwicklung des Fehlers error_i im Algorithmus 18 (**Anzahlsschätzer**):
Das Ergebnis des Algorithmus ist 3.



Das Ergebnis einer pixelbasierten 3-Niveaupassung mit konstanten Funktionen bezüglich des Mengensystems der grauwertverbundenen Mengen gemäß Beispiel 5.4.11.

Abbildung 5.15: Ergebnis des Algorithmus 18 (**Anzahlsschätzer**) auf einem Bild gemäß Beispiel 5.4.11.

5.4.5 Abschätzung über den Sensorfehler

Ist der bildgebende Sensor genau bekannt, so kann man die Anzahl der im Bild trennbaren Segmente über die Sensoreigenschaften ermitteln. Allerdings sind jedoch in den seltensten Fällen alle Eigenschaften des bildgebenden Sensors (bzw. im Falle einer CCD-Kamera des bildgebenden Sensor-Arrays) bekannt. Im Falle einer CCD-Kamera etwa genügt es nicht, das Rauschverhalten des CCD-Arrays zu kennen. Vielmehr müssen auch die Digital/Analog-Wandlung in der Kamera, die Leitungseigenschaften des Kabels, die Analog/Digital-Wandlung im Framegrabber, die vielen Beleuchtungsparameter und zahlreiche andere äußere Einflüsse bekannt sein. Nur in einer sterilen Laborumgebung kann man letztlich die Sensoreigenschaften einer CCD-Kamera halbwegs genau bestimmen. Trotzdem soll – der Vollständigkeit halber – an dieser Stelle ein Algorithmus zur Bestimmung der Anzahl der trennbaren Segmente im Bild über einen vorgegebenen Sensorfehler präsentiert werden.

Definition 5.4.12

Sei $A \in \mathfrak{A} \cap \mathcal{P}(M)$ und $g : M \rightarrow \mathbb{R}_+^d$ ein diskretes Pixelbild. Der Fehler

$$p(g)_A := \max_{p \in \mathbb{P}(A)} \{\|g(p) - g_A(p)\|_2^2\}$$

heiße der *maximale Pixelfehler von g auf dem Segment A* . Für $\mathbf{A} \in S_k(\mathfrak{A})$ sei

$$p(g)_{\mathbf{A}} := \max\{p(g)_A \mid A \in \mathbf{A}\}$$

definiert. $p(g)_{\mathbf{A}}$ heiße der *größte maximale Pixelfehler von g auf der Segmentierung \mathbf{A}* .

Bemerkung 5.4.13

- a) Erfolgt die Berechnung des lokalen pixelbasierten Maximums in Algorithmus 19 (**Epsilon**) über die Niveaureduktion des Abschnitts 5.3, so wird in den meisten Fällen der größte maximale Pixelfehler bei sinkender Segmentanzahl steigen. Daher würde es in der Praxis im Algorithmus genügen, solange Segmente zu verschmelzen, bis der größte maximale Pixelfehler die vorgegebene ε -Schranke überschreitet.
- b) Obwohl in den seltensten Fällen eine genaue ε -Schranke als Sensoreigenschaft für den Algorithmus 19 (**Epsilon**) angegeben werden kann, ist der Algorithmus in vielen praktischen Anwendungen einsetzbar. Ein Benutzer kann dann solange die ε -Schranke verändern, bis ihm das Ergebnis optisch gefällt.

Algorithmus 19 (Epsilon) Schätzung der Anzahl der Segmente über den Sensorfehler

Input: Diskretes Pixelbild $g : M \rightarrow \mathbb{R}^d$,
 Startanzahl k von Segmenten,
 $\varepsilon > 0$

Output: geschätzte Anzahl von Objekten im diskreten Pixelbild

```

1: for  $i = k$  abwärts bis  $i = 1$  do
2:   Berechne ein lokales pixelbasiertes Maximum  $\mathbf{A} \in S_i(\mathfrak{A})$  einer  $i$ -Niveaupassung des
   Bildes  $g$ 
3:    $\text{error}_i := p(g)_{\mathbf{A}}$ 
4: end for
5:  $\text{estimatedSegments} \leftarrow k$ 
6: for  $i = k$  abwärts bis  $i = 1$  do
7:   if  $\text{error}_i \leq \varepsilon$  then
8:      $\text{estimatedSegments} \leftarrow i$ 
9:   end if
10: end for
11: return  $\text{estimatedSegments}$ 

```

5.4.6 Andere Fehlerfunktionen zur Abschätzung der Anzahl der Segmente

Die gleichen Kriterien, die bei der Minimierung des Energiefunktional nach Mumford und Shah (vgl. Abschnitt 2.4) die Anzahl der Unstetigkeitsstellen direkt oder indirekt bestimmen, können auch bei Segmentierung mit Methoden der vektoriiellen Mehrniveaupassung angewendet werden.

Im Bereich der Mehrniveaupassung erscheinen bezüglich des Mengenmodells der pixelzusammenhängenden Mengen auch die folgenden Abschätzungen über vom Benutzer definierte Schwellwerte sinnvoll:

- Abschätzung über die Sprunghöhe im Bild
 Eine Niveaureduktion erfolgt solange, bis der Abstand $\|g(p_1) - g(p_2)\|_2$ zweier benachbarter Randpixel p_1, p_2 adjazenter Segmente einen vorher festgelegten Schwellwert überschreitet.
- Abschätzung über die Sprunghöhe in der Segmentierung
 Eine Niveaureduktion erfolgt solange, bis der Abstand $\|g_A(p_A) - g_B(p_B)\|_2$ zweier benachbarter Randpixel p_A, p_B adjazenter Segmente A, B einen vorher festgelegten Schwellwert überschreitet.
- Abschätzung über die Steigungen in der Segmentierung
 Eine Niveaureduktion erfolgt solange, bis das Integral der Steigungen der approximierenden Funktion auf einem Segment einen vorher festgelegten Schwellwert überschreitet.

Natürlich ist auch eine Linearkombination aller in diesem Unterkapitel vorgestellten Fehlerfunktionen als neue Fehlerfunktion denkbar

Kapitel 6

Schlußbemerkungen und Ausblick

6.1 Einleitung

Abschließend sollen die Anwendungen der vektoriellen Mehrniveaupassung in der Bildsegmentierung sowohl mit den Definitionen von Segmentierung als auch mit den gängigsten Segmentierungsverfahren verglichen werden.

Im folgenden Unterkapitel werden noch einmal die wichtigsten Konzepte des Ansatzes der Bildsegmentierung über eine vektorielle Mehrniveaupassung zusammengefaßt. Anschließend wird geprüft, ob diese Konzepte den verschiedenen Definitionen des Begriffes „Segmentierung“, die in Abschnitt 2.2 präsentiert wurden, genügen.

Das dritte Unterkapitel orientiert sich an der Praxis der Bildsegmentierung. Die vektorielle Mehrniveaupassung wird mit den gängigen Segmentierungsverfahren verglichen. Es ist offensichtlich, daß dabei der Vergleich zum Mumford-Shah-Modell und zum Ansatz von Kämpke und Kober im Vordergrund steht.

Der Ausblick im vierten Unterkapitel zeigt, daß, sowohl in praktischer als auch in theoretischer Hinsicht, die vektorielle Mehrniveaupassung nicht auf den Bereich der Bildsegmentierung beschränkt werden muß.

6.2 Das Konzept der vektoriellen Mehrniveaupassung als Segmentierungsverfahren

Die beiden vorangegangenen Kapitel 4 („Algorithmen und Anwendungen der Mehrniveaupassung im Bereich der Bildsegmentierung“) und 5 („Niveaureduktion und Schätzung der Anzahl der Segmente“) zeigen insbesondere in ihren Beispielen die Rechtfertigung für den Einsatz der vektoriellen Mehrniveaupassung zur Bildsegmentierung. Die folgenden Abschnitte werden zeigen, daß dieser Ansatz der Segmentierung dem Vergleich mit den notwendigen Bedingungen einer Segmentierung (Definition 2.2.1) oder den Forderungen, die Haralick und Shapiro bzw. Mumford und Shah an eine Segmentierung stellen (vgl. Abschnitt 2.2), standhält. Insgesamt bietet die vektorielle Mehrniveaupassung einen sehr allgemeinen Ansatz, den man – durch die Wahl eines geeigneten Systems erlaubter Mengen bzw. Funktionen – so spezialisieren kann, daß er der jeweiligen Definition des Begriffes „Segmentierung“ aus Abschnitt 2.2 genügt.

6.2.1 Zusammenfassung: Vektorielle Mehrniveaupassung – Anwendungen in der Bildsegmentierung

Die vektorielle Mehrniveaupassung zur Bildsegmentierung ist ein Verfahren, das eine Segmentierung eines Bildes durch die Minimierung des Abstands einer segmentweise stetigen Funktion zum Bild bestimmt. Jede Segmentierung im Kontext der vektoriellen Mehrniveaupassung unterliegt dabei starken Nebenbedingungen, was sowohl das Aussehen der Segmente als auch die Approximation auf den Segmenten angeht. So müssen die Segmente aus einem System erlaubter Mengen sein und die approximierende Funktion auf jedem Segment aus einer Klasse erlaubter Funktionen. Beide Mengen, die der erlaubten Mengen und die der erlaubten Funktionen, sind vor dem Segmentierungsprozeß vom Benutzer je nach Problemklasse wählbar. Selbst das Approximationsverhalten läßt sich durch die Wahl einer erlaubten Familie von Skalarprodukten steuern.

Der Anwender der vektoriellen Mehrniveaupassung kann also an drei Stellen den Segmentierungsprozeß modellieren:

- i. Das System der erlaubten Segmentierungen
Das System der erlaubten Segmentierungen gestattet es dem Anwender, problem-spezifisch ein adäquates Aussehen der Segmente der optimalen Segmentierung fest-zulegen. Der Abschnitt 3.3 stellt allerdings Forderungen an die Eigenschaften des Systems der Segmentierungen, die z. B. das System der zusammenhängenden Men-gen im Zweidimensionalen nicht erfüllt.
- ii. Die Klasse der erlaubten Funktionen
Die Klasse der erlaubten Funktionen ist nach Abschnitt 3.4 ein endlich dimensionaler Vektorraum stetiger Funktionen, der sich ohne Dimensionsverlust auf jedes nichtleere Segment beschränken läßt. Für jedes Segment kann das Approximationsproblem als orthogonale Projektion auf den Raum der erlaubten Funktionen formuliert werden, was der algorithmischen Umsetzung der vektoriellen Mehrniveaupassung sehr ent-gegenkommt. Je nach Problem kann ein geeigneter Funktionenraum vom Anwender gewählt werden.

iii. Die Familie der erlaubten Skalarprodukte

Die Familie der erlaubten Skalarprodukte formuliert das Approximationsverhalten der segmentweise definierten Funktion zum Bild (vgl. Definition 3.4.4). Der zu minimierende Abstand wird als Norm über die Familie der erlaubten Skalarprodukte induziert. In den meisten Fällen der Praxis ist die durch die Norm induzierte Metrik die des $L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}^d))$, so daß die Approximation des Bildes in gewohnter Art und Weise statt findet. Es ist allerdings auch möglich, über das Skalarprodukt Einfluß auf das Approximationsverhalten zu nehmen. Man könnte z. B. bei mehrkanaligen Bildern die einzelnen Kanäle unterschiedlich gewichten.

Ein weiterer Parameter, der dem Anwender der vektoriellen Mehrniveaupassung in die Hand gegeben wird, ist die Anzahl der Segmente einer Segmentierung. Eine Schätzung dieses Parameters kann allerdings auch algorithmisch erfolgen, wie das Kapitel 5 („Niveaureduktion und Schätzung der Anzahl der Segmente“) zeigt. Der Anwender sollte sich bei der Verwendung von Algorithmen des fünften Kapitels allerdings immer bewußt sein, daß es sich dabei lediglich um eine heuristische Schätzung einer plausiblen Anzahl von Segmenten handelt.

6.2.2 Vergleich mit den Bedingungen von Haralick und Shapiro

Segmentierungen über die vektorielle Mehrniveaupassung erfüllen die Bedingungen, die Haralick und Shapiro für Segmentierungen beobachtet haben (vgl. (2.1), S. 14):

i.: Segmente einer Segmentierung sind uniform oder homogen bezüglich einer Charakteristik wie zum Beispiel dem Grau- oder Farbwert oder der Textur.

ii.: Adjazente Segmente haben signifikant unterschiedliche Werte in der Charakteristik, in der sie uniform sind.

Die beiden ersten Bedingungen von Haralick und Shapiro lassen sich über die Klasse der erlaubten Funktionen beantworten. Erlaubt man nur konstante Funktionen, so sind die beiden Bedingungen sicherlich trivial erfüllt. Aber auch für höherdimensionale Räume erlaubter Funktionen verspricht die Definition der erlaubten Funktionen 3.4.2 (S. 84) Uniformität auf den Segmenten über die Bedingung der Stetigkeit. Die erlaubten Unstetigkeitsstellen der approximierenden Funktion an den Segmentgrenzen führen in dem überwiegenden Teil der Anwendungen zu den geforderten signifikant unterschiedlichen Werten in der Approximation auf adjazenten Segmenten.

iii.: Das Innere der Segmente ist einfach und hat, wenn überhaupt, nur wenige kleine „Löcher“.

iv.: Die Ränder von Segmenten sind ebenfalls einfach und nicht zerfranst.

Die dritte und vierte Charakterisierung einer Segmentierung kann die vektorielle Mehrniveaupassung über das System der erlaubten Segmentierungen erfüllen. Es ist offensichtlich, daß nicht jedes System von Segmentierungen durch erlaubte Mengen nach Definition 3.3.1 diese beiden Punkte erfüllt. Das System der grauwertbenachbarten Mengen in Abschnitt 4.5, das zu einem Schwellwertverfahren führt, erfüllt sie beispielsweise nicht. Aber über die Definition der erlaubten Mengen nach Definition 3.2.10 kann man Systeme von Segmentierungen definieren, die die Punkte iii. und iv. der Definition einer Segmentierung von Haralick und Shapiro erfüllen.

6.2.3 Vergleich mit den Forderungen von Mumford und Shah

Mit den obigen Ausführungen zu den Punkten i. und ii. der Bedingungen, denen nach Haralick und Shapiro der Begriff „Segmentierung“ unterliegt, sollte klar sein, daß die vektorielle Mehrniveaupassung durchaus den Forderungen von Mumford und Shah an eine Segmentierung (vgl. (2.2), S. 15) entspricht: Denn die vektorielle Mehrniveaupassung teilt den Definitionsbereich eines Bildes in disjunkte Bereiche, auf denen sich das Bild nur stetig, zwischen denen sich das Bild im allgemeinen jedoch unstetig ändert. Da die Definition von Mumford und Shah untrennbar mit der praktischen Umsetzung durch das Mumford-Shah-Modell (vgl. Abschnitt 2.4) verbunden ist, sei für den Vergleich auch auf den Abschnitt 6.3.1 verwiesen.

6.2.4 Vergleich mit den notwendigen Bedingungen einer Segmentierung

Der eigene Ansatz dieser Arbeit, notwendige Bedingungen einer Segmentierung zu definieren (vgl. Definition 2.2.1, S. 18), ist vor dem Hintergrund der vektoriellen Mehrniveaupassung entstanden. Daher werden alle Forderungen dieser Definition durch die Anwendungen der vektoriellen Mehrniveaupassung zur Bildsegmentierung erfüllt, wie man an den Punkten i. bis iv. der Definition überprüfen kann. Die vektorielle Mehrniveaupassung stellt darüber hinaus auch jene Flexibilität bereit, die implizit in der Definition 2.2.1 gefordert wird, insbesondere in den Punkten ii. (Bereitstellung von Funktionen zur Lösung eines segmentweisen Approximationsproblems), iii. (Bedingungen an das Aussehen der Segmente sind formulierbar) und iv. (Die Anzahl der Segmente ist ein Pareto-Optimum einer Mehrzieloptimierung).

6.3 Vergleich zu anderen Segmentierungsverfahren

In diesem Abschnitt sollen die Segmentierungsalgorithmen über die vektorielle Mehrniveaupassung mit anderen Segmentierungsansätzen verglichen werden. Dabei bieten sich aufgrund der inhaltlichen Nähe die Methoden der Variationsrechnung zur Bildsegmentierung (vgl. Abschnitt 2.4) und der Funktionsquantisierungsansatz von Kämpke und Kober (vgl. Abschnitt 2.5) besonders an. Aber auch zur „Split and Merge“-Technik läßt sich ein Bezug finden.

6.3.1 Die vektorielle Mehrniveaupassung im Vergleich mit den Methoden der Variationsrechnung zur Bildsegmentierung

Die Ähnlichkeit dieser beiden Ansätze zur Segmentierung von Bildern ist offensichtlich: Beide, sowohl die bildenergieminimierenden Verfahren über die Methoden der Variationsrechnung als auch die vektorielle Mehrniveaupassung, bestimmen die optimale Segmentierung eines Bildes über eine Approximation mit stückweise stetigen Funktionen, deren stetige Bereiche die Segmente bilden. Betrachtet man das Energiefunktional $E = D + S + P$, so fällt auf, daß das Funktional D , das den Abstand der stückweise stetigen Approximation zum Bild „bestraft“, auch die Grundlage der vektoriellen Mehrniveaupassung ist (vgl. Abschnitt 3.1).

Die Funktionale S und P findet man bei der Formulierung der vektoriellen Mehrniveaupassung nicht. Ihre Aufgaben übernehmen die Klasse der erlaubten Funktionen und das System der erlaubten Mengen.

6.3.1.1 Vergleich der Segmentgrenzen

Bei der vektoriellen Mehrniveaupassung ist die Gestalt der Segmente direkt über die Definition des System der erlaubten Mengen steuerbar. Im Gegensatz dazu kann der Anwender des Mumford-Shah-Modells nur wenig Einfluß auf das Aussehen der Segmentgrenzen der Segmentierung nehmen. Selbst die Theorie dieses Ansatzes bietet nur vage Informationen über die Gestalt dieser Grenzen. Die Sätze 2.4.4 (S. 35) und 2.4.5 (S. 35) und die Bemerkung 2.4.6 (S. 35) zeigen zwar, daß die Grenzlinien einer Energieminimierung des Funktionals $E = D + S + P$ mathematischen Glattheitsbedingungen genügen, und der Strafterm P erlaubt dem Anwender, eine zu lange Grenzlinie zu „bestrafen“, aber Aussagen über das optische Aussehen liefert dieser Ansatz nicht. Dies wäre sicherlich auch nicht im Sinne der Idee der Energieminimierung, die eben auch das Aussehen der Segmentgrenzen als Energieminimierung im Strafterm P , d. h. als freies Spiel der vorgegebenen „Kräfte“, modelliert.

Die vektorielle Mehrniveaupassung bietet kein Äquivalent zu diesem Strafterm P , daher muß zwangsläufig ein System erlaubter Mengen eingeführt werden, um optisch ansehnliche Ergebnisse zu erhalten.

Der Vorteil dieses Mengensystems ist es, daß das Ergebnis einer Mehrniveaupassung genau den Zielvorstellungen des Benutzers entsprechen muß. Dies kann allerdings auch ein

Nachteil sein: Ein Mengensystem kann sich unter Umständen als zu streng gewählt herausstellen. In diesem Fall bietet es keine Spielräume mehr: Segmentiert man mit rechteckigen Mengen, kann man keine rundlichen Segmente erwarten. Andererseits kann man auch ein zu wenig restriktives Mengensystem wählen, in dem vielleicht zu starke Zerfransungen des Randes der Segmente bei einigen Segmentierungen auftreten (vgl. etwa Abbildung 4.43, S. 217).

6.3.1.2 Vergleich der Approximation auf den Segmenten

Eine große Gemeinsamkeit der Segmentierung mit vektorieller Mehrniveaupassung und der Segmentierung mit Methoden der Variationsrechnung ist die Approximation des Bildes durch stetige Funktionen auf den Segmenten. Um dieser Approximation einen sinnvollen Rahmen zu geben, d. h. um das Approximationsproblem so zu formulieren, daß die Lösung des Problems auch in den Rahmen des Begriffes „Segmentierung“ fällt, gehen beide Segmentierungsverfahren jedoch unterschiedlich vor.

Während die vektorielle Mehrniveaupassung den Raum der Funktionen, mit denen approximiert wird, einschränkt, lassen die Ansätze der Variationsrechnung diesen Raum weitgehend unspezifiziert und „bestrafen“ große Funktionsänderungen der approximierenden Funktion durch das Funktional S des Energiefunktionals $E = D + S + P$ (vgl. Abschnitt 2.4.1, Gleichung (2.8), S. 32). Dieser Ansatz erlaubt es der approximierenden Funktion, dem Verlauf des Bildes zu folgen, wenigstens so weit es das Funktional S ermöglicht. Es wird nicht der Abstand zum Bild minimiert, sondern eben ein Energiefunktional, in dem der Abstand zum Bild nur ein Teil einer Zielfunktion ist. Bei der Diskretisierung dieses Ansatzes durch die Methode der finiten Elemente nach Blake und Zisserman (vgl. Abschnitt 2.4.2.1) führt dies jedoch dazu, daß es einen durch die Parametrisierung der Strafterme S und P determinierten Schwellwert für eine Sprungstelle des Bildes gibt, bei dem die Methoden der Variationsrechnung eine Unstetigkeitsstelle haben *müssen* (vgl. Gleichung (2.12), S. 37).

Die vektorielle Mehrniveaupassung geht einen anderen Weg um das Approximationsproblem auf den Segmenten segmentierungsgerecht einzuschränken: Bei diesem Verfahren darf nur ein endlich dimensionaler Vektorraum von Funktionen zur Approximation herangezogen werden. Im Gegensatz zur Energieminimierung nach Mumford und Shah wird nur der Abstand zum Bild minimiert. Bei einer Approximation durch Funktionen eines endlichen Vektorraums ist die Sprunghaftigkeit der approximierenden Funktion, auf Pixeln durch die Einschränkung auf den Vektorraum stark begrenzt. Denn Pixel haben eine positive Fläche und daher ist nach Definition 3.4.2 das Approximationsproblem auf einem Pixel eindeutig lösbar. Dadurch ist die Sprunghaftigkeit der approximierenden Funktion nicht mehr veränderbar. Das Sprungverhalten hängt also direkt von der Klasse der erlaubten Funktionen ab. Wählt man etwa einen Polynomraum mit niedrigem Grad, so ist die Sprunghöhe, die eine approximierende Funktion zwischen zwei Pixeln überwinden kann, in der Regel sehr gering. Auf der anderen Seite kann es auch sein, daß – etwa bei periodischen Texturen – Sprünge zwischen Pixeln gewollt sind. Hier bietet die vektorielle Mehrniveaupassung bei richtiger Auswahl der Klasse der erlaubten Funktionen einen größeren Spielraum als das Mumford-Shah-Modell (vgl. auch Abschnitt 6.3.1.5).

Da das Approximationsproblem auf dem ganzen Segment eindeutig lösbar ist, kann die

Lösung nicht so flexibel auf lokale Bildveränderungen reagieren, wie beim Ansatz nach Mumford und Shah. Allerdings müssen lokale Ausreißer der Bildfunktion, die einen gewissen Schwellwert überschreiten, nicht zwangsläufig zu einer Segmentgrenze führen. Die Abbildung 6.1 (S. 258) verdeutlicht diesen Unterschied im Approximationsverhalten.

Die Bewertung des Verhaltens gegenüber lokalen Ausreißern der Bildfunktion muß problemabhängig sein: Ist der Benutzer nicht gewillt, einen lokalen Wertsprung einer gewissen Höhe als Rauschen abzutun, so ist das Verhalten der Energieminimierung zur Bildsegmentierung vorzuziehen.¹ Will man andererseits einzelne Pixel, die inmitten eines gleichförmigen Segments plötzlich einen herausragenden Wert haben, nicht überbewerten, so bietet sich die vektorielle Mehrniveaupassung an.

Ein nicht zu unterschätzender Vorteil der Einschränkung auf einen endlich dimensional Vektorraum ist auch die einfachere algorithmische Umsetzung der Approximation: Hat man eine Orthonormalbasis des Vektorraums, so ist die optimal approximierende Funktion aus dem Vektorraum eine orthogonale Projektion (vgl. Bemerkung 3.5.4, S. 94). Dies ist bei diskreten Pixelbildern eine leicht zu berechnende Operation (vgl. Abschnitt 4.1). Der algorithmische Aufwand ist daher ungleich geringer als bei der Energieminimierung, bei der man schnell zu nichtlinearen Optimierungsproblemen kommt, die wesentlich aufwendiger zu lösen sind (vgl. Abschnitte 2.4.2.1, 2.4.3.1 und 2.4.3.2).

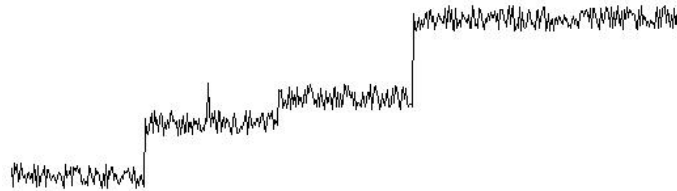
6.3.1.3 Vergleich der Anzahl der Segmente

Ein weiterer wesentlicher Unterschied zwischen dem Ansatz von Mumford und Shah und dem der vektoriellen Mehrniveaupassung ist der Umgang mit der Anzahl der durch die Segmentierung zu erwartenden Segmente. Während die Anzahl der Segmente bei der Energieminimierung des Mumford-Shah-Funktional nur indirekt über die Parameter α und β der Strafterme S und P gesteuert werden kann, ist sie bei der vektoriellen Mehrniveaupassung ein freier Parameter, der entweder vom Benutzer oder von einem Schätzalgorithmus (vgl. Kapitel 5) vorgegeben wird.

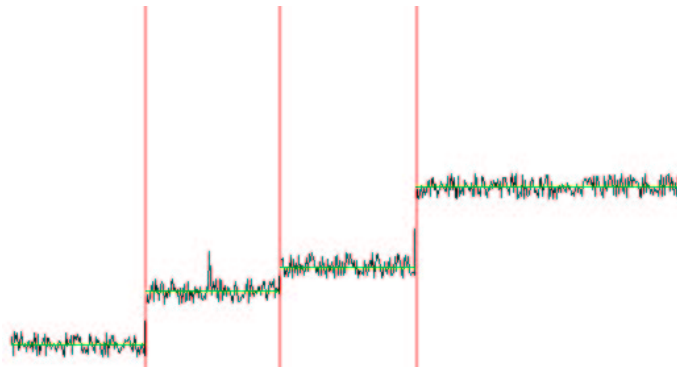
Der Ansatz der Variationsrechnung schafft immer dann ein neues Segment, wenn das Energiefunktional es fordert, d. h. wenn das Bild einen Sprung macht, der zu einer hohen „Strafe“ durch den Term S führt. Der Vorteil dieses Vorgehens liegt auf der Hand: Ist das bildgebende Verfahren in seinen Sensoreigenschaften gut bekannt, und sind daher die Parameter α und β fest vorgegeben, so kann der Variationsansatz mit Bildern, die eine unbekannter Anzahl von Objekten enthalten, umgehen. Auf der anderen Seite kann in einem solchen Fall für die vektorielle Mehrniveaupassung jederzeit der Algorithmus 19 (**Epsilon**) des Abschnitts 5.4.5 zur Schätzung der Anzahl der Segmente herangezogen werden. Liegt jedoch, wie meist in der Praxis, eine solch tiefgehende Kenntnis über das bildgebende Verfahren nicht vor, so muß ein Anwender der energieminimierenden Verfahren die Parameter α und β solange variieren, bis er ein ihm vernünftig erscheinendes Segmentierungsergebnis erhält, obwohl unter Umständen die Anzahl der zu segmentierenden Objekte in der Problemstellung klar definiert ist.

Ein weiterer Unterschied bei der Bestimmung der Anzahl der Segmente findet man in der Bewertung der Form der Segmente. Während bei den Segmentierungsalgorithmen über

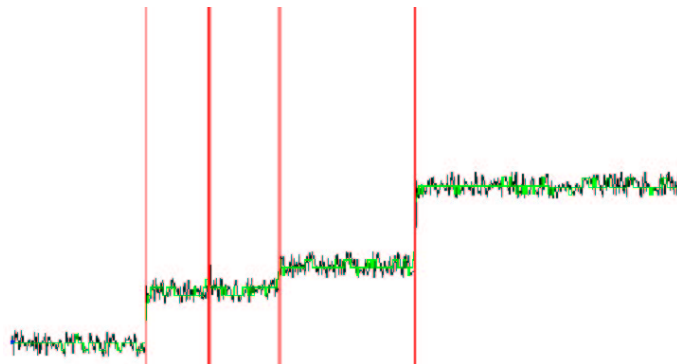
¹Abschnitt 5.4.6 zeigt jedoch, daß man ein solches Verhalten auch innerhalb der vektoriellen Mehrniveaupassung modellieren kann.



Eindimensionales Originalbild: Man beachte den hohen Funktionswert ungefähr in der Mitte des 2. „Plateaus“.



4-Niveaupassung mit konstanten Funktionen bezüglich des Mengensystems der zusammenhängenden Mengen.



Variationsansatz: Es ist nicht möglich, die Parameter α und β so zu wählen, daß sich genau die vier Segmente der 4-Niveaupassung ergeben.

Abbildung 6.1: Vergleichendes Beispiel der Approximationseigenschaften der vektoriiellen Mehrniveaupassung und des Energieminimierungsansatzes nach Mumford und Shah.

die Energieminimierung immer die Länge des Randes der Segmente als Parameter in die Ermittlung der Anzahl der Segmente miteinfließt, ist dies bei den Schätzverfahren der vektoriellen Mehrniveaupassung in Kapitel 5 nicht der Fall. Vielmehr wird bei diesen Schätzalgorithmen immer ein Approximationsfehler, der unter Umständen über die Größe der Segmente gemittelt ist (vgl. Abschnitt 5.4.2), als Parameter der Schätzung herangezogen. Bei keinem der Verfahren spielt die Form der Segmente eine Rolle, wie dies bei einer Bewertung der Randlänge implizit der Fall ist.

6.3.1.4 Vergleich der Sprunghöhen zwischen den Segmenten

In der Bewertung der Sprunghöhen zwischen den Segmenten liegt ein weiterer fundamentaler Unterschied zwischen der Bildsegmentierung über die vektorielle Mehrniveaupassung und der über die Energieminimierung des Mumford-Shah-Funktional. Die Parametrisierung der „Strafen“ für Unstetigkeitsstellen und Steigungen über die Parameter α und β liefert – wie bereits erwähnt – einen festen Schwellwert für die Sprunghöhen des Bildes an den Segmentgrenzen. Über den Umweg der Niveaureduktion und der Abschätzung der Anzahl der Segmente bietet die vektorielle Mehrniveaupassung indirekt auch die Möglichkeit, die Sprunghöhen zwischen den Segmenten zu kontrollieren, wie Abschnitt 5.4.6 zeigt.

6.3.1.5 Segmentierung von Texturen

Die Segmentierung von Texturen mit Hilfe von energieminimierenden Verfahren, gemäß dem Mumford-Shah-Modell, kann nur indirekt erfolgen. Der Term S des Energiefunktional „bestraft“ Schwankungen der approximierenden Funktion. Deshalb erzeugen Texturen mit regelmäßigen Schwankungen im Grau- oder Farbwertverlauf einen großen Wert im Energiefunktional. Es ist dies schließlich auch im Sinne der Definition des Begriffes „Segmentierung“ von Mumford und Shah (vgl. Abschnitt 2.2.1), in der gefordert wird, daß sich das Bild innerhalb eines Segments nur schwach ändert. Also eignet sich das Mumford-Shah-Modell nicht zur direkten Segmentierung von solchen Texturen. Stattdessen wird anstelle des normalen Bildes ein Featurebild, d. h. ein Metabild mit aus dem Originalbild extrahierten Features, segmentiert. Die aus dem Originalbild gewonnenen Featurevektoren sollten dabei derart sein, daß sie sich innerhalb eines Segments nur wenig und zwischen den Segmenten stark unterscheiden. Ein solches Vorgehen findet man z. B. in [KLM94].

Die vektorielle Mehrniveaupassung kennt, wie in Abschnitt 6.3.1.2 bereits erwähnt, keinen Strafterm für Steigungen der Approximation, da lediglich der Abstand der approximierenden Funktion zum Bild gemessen wird. Daher muß, bei entsprechender Wahl der Klasse der erlaubten Funktionen, ein regelmäßig schwankender Grau- oder Farbwertverlauf nicht zu einem hohen Approximationsfehler auf der Texturfläche und damit unter Umständen zur Bildung mehrerer Niveaus auf dieser Fläche führen. Als Klasse der erlaubten Funktionen könnte man zur Segmentierung von Texturen mit periodischem Grauwertverlauf einen Funktionenvektorraum periodischer Funktionen wählen. $\mathcal{E} := (1, \cos, \sin, \dots, \cos 2\pi n, \sin 2\pi n)$ für ein $n \in \mathbb{N}$ wäre beispielsweise ein solcher Vektorraum.

6.3.2 Die vektorielle Mehrniveaupassung im Vergleich zu den Quantisierungsansätzen zur Bildsegmentierung

Bereits in der Bewertung der Ansätze von Kämpke und Kober zur Bildsegmentierung über Funktionsquantisierung (Abschnitt 2.5.4) wurde darauf hingewiesen, daß man die vektorielle Mehrniveaupassung als Erweiterung dieses Ansatzes sehen kann. Übernommen wurde die stückweise Approximation des Bildes. Bei der Quantisierung von Funktionen mit eindimensionalem Definitionsbereich war implizit das System von Segmentierungen durch Intervalle gegeben. Dieses wurde für die vektorielle Mehrniveaupassung im Abschnitt 3.3 benannt und abstrahiert. Dadurch war es möglich, viele Ergebnisse von Kämpke und Kober auch für Funktionen mit höherdimensionalem Definitionsbereich zu formulieren.

Während bei der Funktionsquantisierung nach Kämpke und Kober die Funktion (oder das Bild) auf den Segmenten nur durch konstante Funktionen approximiert wird, erweitert die vektorielle Mehrniveaupassung den Raum der approximierenden Funktionen auf endlich dimensionale Vektorräume von Funktionen, die gewissen Bedingungen unterworfen sind (vgl. Definition 3.4.2, S. 84). Trotzdem behalten wichtige Ergebnisse des Unterkapitels 2.5 auch in dem erweiterten Kontext der vektoriellen Mehrniveaupassung ihre Gültigkeit. Der wichtigste Satz in diesem Zusammenhang ist, daß die Partition (oder Segmentierung) die approximierende Funktion eindeutig bestimmt. Im Fall der Funktionsquantisierung ist dies der Satz 2.5.6 (S. 51), der im Satz 3.5.21 (S. 102) seine Entsprechung im Kontext der vektoriellen Mehrniveaupassung findet. Auch das sogenannte „Massenerhaltungslemma“ 2.5.5 (S. 50) hat ein Pendant in Satz 3.7.20 (S. 130). Dies gilt allerdings nicht für einige für die Algorithmik der Funktionsquantisierung wichtige Sätze. Das Lemma über die notwendige Bedingung der Partition 2.5.9 (S. 52) ließe sich sicherlich in einer sehr abstrakten Version auf die Mehrniveaupassung übertragen, hätte aber algorithmisch wohl keine Bedeutung. Das Lemma 2.5.21 (S. 56), über das sich lokal optimale Partitionen als Fixpunkte der gegebenen Funktion berechnen lassen, gilt nur für monotone Funktionen und ist daher für die vektorielle Mehrniveaupassung im allgemeinen uninteressant.

6.3.3 Einordnung der vektoriellen Mehrniveaupassung als Segmentierungsverfahren

Auch abseits des Mumford-Shah-Modells und der Funktionsquantisierung gibt es verwandtschaftliche Beziehungen von anderen Segmentierungsverfahren zur vektoriellen Mehrniveaupassung. Schon der Abschnitt 4.5 zeigt, daß es durchaus möglich ist, mit dem Instrumentarium der Mehrniveaupassung ein Schwellwertverfahren zu formulieren. Man könnte den Algorithmus 14 (**k-NiveaupassungZ**) in Abschnitt 4.6 auch als Regionwachstumsverfahren sehen. Allerdings wachsen die Regionen in diesem Algorithmus nicht nur, sondern sie schrumpfen unter Umständen auch.

Der Algorithmus zur Niveaureduktion, Algorithmus 16 (**Iterative Niveaverschmelzung**) aus Unterkapitel 5.3, ist auch Grundlage eines klassischen „Split and Merge“-Verfahrens, das in den praktischen Anwendungen der vektoriellen Mehrniveaupassung eine große Rolle spielt:

i. „Split“

Der Algorithmus 13 (**Initiales Farbschwellwertverfahren**) bietet ein schnelles Verfahren, ein gegebenes Farbbild in pixelzusammenhängende Mengen, die Pixel ähnlichen Farbwertes zusammenfassen, zu zerlegen.

ii. „Merge“

Der Algorithmus 16 (**Iterative Niveaueschmelzung**) vereinigt iterativ jeweils die beiden adjazenten pixelzusammenhängenden Segmente, deren Vereinigung den geringsten Fehler in der Zielfunktion der k -Niveaupassung bedeutet.

6.4 Ausblick

Natürlich kann in dieser Arbeit nur einer kleiner Ausschnitt der Möglichkeiten und Probleme der vektoriellen Mehrniveaupassung präsentiert werden. Sowohl bei den praktischen Anwendungen der Resultate des Kapitels 3 („Theoretische Grundlagen der vektoriellen Mehrniveaupassung“) als auch bei der Theorie der Mehrniveaupassung sind noch Lücken zu schließen.

6.4.1 Praktischer Ausblick

Um den Umfang dieser Arbeit zu begrenzen, wurde sie auf die Anwendungen der vektoriellen Mehrniveaupassung in der Bildsegmentierung beschränkt. Dabei wurden in den praktischen Beispielen auch nur Basen zur polynomialen Approximation des gegebenen Bildes betrachtet. Interessant wäre sicherlich in diesem Zusammenhang ebenfalls die Segmentierung von Texturen mit periodischen Funktionen oder Wavelets. Aber auch Anwendungen, die auf den ersten Blick wenig mit Segmentierung zu tun haben, sind mit den in Kapitel 3 bereitgestellten Grundlagen möglich.

6.4.1.1 Erweiterung auf Subpixelgenauigkeit

Analog zum Abschnitt 4.2 gibt es zwei Möglichkeiten, den pixelbasierten Ansatz der praktischen Anwendungen der Mehrniveaupassung aus Kapitel 4 zu einem subpixelgenauen Ansatz zu erweitern. Eine einfache Möglichkeit dies zu tun, ist es, die Größe der Pixel zu verkleinern. Statt das Pixelgitter als disjunkte Zerlegung des Bildes durch quadratische Regionen mit der Fläche 1 zu definieren, kann man die Fläche der Pixel auch auf l^2 festlegen. Die Seitenlänge der Pixel als quadratische Region ist dann l . Diesen Wert kann man als Parameter einer subpixelgenauen Segmentierung nutzen, indem man ihn sukzessiv verkleinert.

Dieser Weg, den pixelbasierten Ansatz zu erweitern, geht allerdings von einem Bild aus, das nur auf einem Pixelgitter mit Pixelseitenlänge 1 definiert ist. Für einen Ansatz mit parametrisierten Mengenmodellen, der auch den Bildaufnahme-prozeß mit in die Modellierung aufnimmt, sei auf [Lau03] verwiesen.

6.4.1.2 Die Mehrniveaupassung zur Kantendetektion

Kantendetektions- und Segmentierungsverfahren sind naturgemäß eng miteinander verbunden. Auf der einen Seite können Kanten im Bild dazu genutzt werden, Segmente zu bestimmen, auf der anderen Seite kann man aber auch Segmentgrenzen als Kanten im Bild ansehen. Gerade die energieminimierenden Verfahren sind auch als Kantendetektoren anwendbar (vgl. Abschnitte 2.4.2.1, 2.12 und 2.4.4.4). In [MS95] werden einige klassische Kantendetektionsverfahren, wie z. B. der Ansatz von Hildreth und Marr (vgl. [Mar82]), in einem Variationsmodell beschrieben.

Auch die Segmentgrenzen einer Segmentierung über eine vektorielle Mehrniveaupassung lassen sich als Grundlage einer Kantendetektion nutzen, wie Beispiel 6.4.1 zeigt.

Beispiel 6.4.1

Im EU-Projekt INTELLIWOOD soll die Maserung von Hölzern untersucht werden. Bei eingescannten Bildern war der Übergang von der dunkleren Maserung zum helleren Hintergrund fließend über einer Breite von bis zu acht Pixeln. Daher hatten klassische Kantenoperatoren, wie z. B. der Sobeloperator (vgl. [Sob70]), Schwierigkeiten die Kanten zu finden, da sie meist auf kleinen Bildausschnitten (in der Regel 3×3 oder 5×5 Pixel) arbeiten. An dieser Stelle erwies sich eine vorherige Segmentierung über die vektorielle Mehrniveaupassung bezüglich des Mengenmodells der pixelzusammenhängenden Mengen (vgl. Abschnitt 4.6) als vorteilhaft. Abbildung 6.2 zeigt ein Beispiel dieser Anwendung.



Abbildung 6.2: Mit Hilfe des Algorithmus 6 (**k-Niveaupassung**) berechnete Kanten einer Holzoberfläche. Auf der linken Seite ist das Eingabebild zu sehen, auf der rechten Seite die ermittelten Kanten.

6.4.1.3 Vordergrundniveaus und Hintergrundniveau

In vielen Segmentierungen werden Objekte vom Hintergrund getrennt. In diesem Fall ist es oft nicht sinnvoll zu verlangen, daß die Menge der Pixel, die dem Hintergrund zugeordnet wird, auch den Einschränkungen des Mengensystems unterworfen ist (vgl. Abbildung 6.3). Gerade im Kontext der Segmentierungen durch pixelzusammenhängende Mengen erscheint es oft hinderlich, auch Pixelzusammenhang vom Hintergrundsegment zu fordern. Daher bietet es sich an, die Überprüfung des Zusammenhangs im Algorithmus 14 (**k-NiveaupassungZ**) für das Segment, das den Hintergrund repräsentiert, auszusetzen. An dieser Stelle ist allerdings Vorsicht geboten: Wurde ein Segment einmal als Hintergrund festgelegt, so muß es während der gesamten Laufzeit des Algorithmus Hintergrundsegment bleiben. Daher ist im Vorfeld genau zu prüfen, welche Pixel dem Hintergrund zuzuordnen sind und welche nicht.

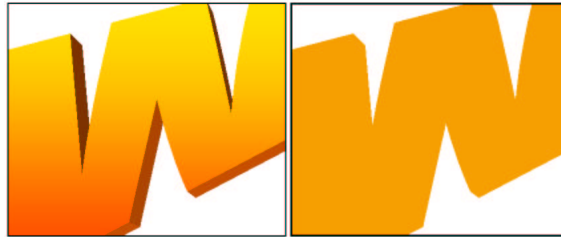


Abbildung 6.3: Deutlich ist im linkem Bild zu sehen, daß der weiße Hintergrund nicht pixelzusammenhängend ist. Das rechte Bild zeigt eine 2-Niveaupassung, bei der die Zusammenhangsbedingung für den Hintergrund ausgesetzt wurde.

6.4.1.4 Mehrniveaupassung von Transformationen

In vielen Bereichen der digitalen Bildverarbeitung ist es nötig, Transformationen von Pixeln zu behandeln. Für ein optisches Fahrassistenzsystem kann es z. B. sinnvoll sein, (zusammenhängende) Bereiche in Bildern einer Sequenz zu gruppieren, die sich in dieselbe Richtung bewegen, um etwa den Hintergrund von anderen sich ebenfalls bewegenden Verkehrsteilnehmern zu unterscheiden. Statt eines Bildes $g : \mathbb{P}(M) \rightarrow \mathbb{R}$ könnte man dazu ein Bild von Transformationsvektoren, die für bestimmte Gebiete berechnet wurden, betrachten. Für Transformationsvektoren, die sich durch Vektoren des \mathbb{R}^2 kodieren lassen, könnte man etwa Bilder der Form $t : \mathcal{P}(M) \rightarrow \mathbb{R}^2$ untersuchen.

Eine ähnliche Situation liegt beim Entzerren von optisch verzerrten Bildern vor.

Die Arbeit von Georg Pisinger ([Pis02]) wird sich unter anderem mit Problemen dieser Art auseinandersetzen.

Beispiel 6.4.2 (Entzerrung)

Abbildung 6.4 (links) zeigt das Bild eines Punktmusters, das auf ein regelmäßiges Punktgitter transformiert werden soll. Abbildung 6.4 (Mitte) zeigt die Transformationsvektoren der einzelnen im Bild gefundenen schwarzen Punkte.

Berechnet wird nun zunächst ein Zwischenbild h . Dafür wird zunächst die Voronoi-Zerlegung bezüglich der schwarzen Bildpunkte ermittelt. Dann wird im Zwischenbild h jedem Pixel eines Voronoi-Gebietes der Transformationsvektor des schwarzen Bildpunktes innerhalb des Gebietes zugeordnet. Beginnend mit der Voronoi-Zerlegung als Startsegmentierung werden nun mit dem Algorithmus 16 (**Iterative Niveaueschmelzung**) solange adjazente Gebiete miteinander verschmolzen, bis die gewünschte Anzahl an Transformationsgebieten erreicht ist. Als Raum der erlaubten Funktionen wurde in diesem Fall der Raum der quadratischen Polynome genommen. In der Beispielabbildung wurde die gewünschte Anzahl über den Algorithmus 18 (**Anzahlschätzer**) berechnet. In Abbildung 6.4 (rechts) sind die ermittelten Ergebnissegmente farblich unterschiedlich markiert. Die

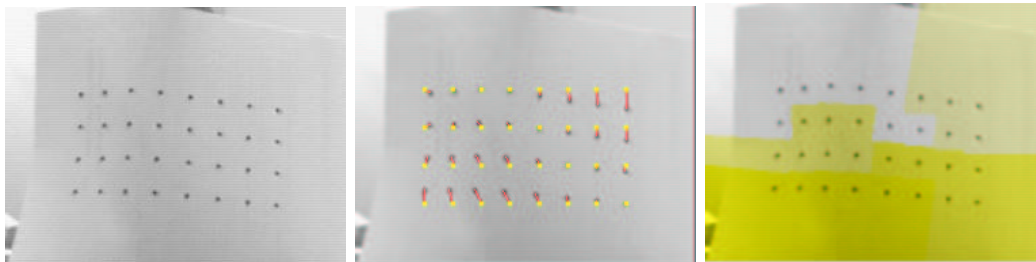


Abbildung 6.4: Beispiel einer Entzerrung mit vektorieller Mehrniveaupassung: Links ist das Ausgangsmuster zu sehen. In der Mitte sieht man die Transformationen auf das vorgegebene Punktegitter (gelb). Rechts ist das Ergebnis des in Beispiel 6.4.2 beschriebenen Verfahrens zu sehen.

dazugehörigen Abbildungen bilden alle schwarzen Bildpunkte auf die dazugehörigen gelben Rasterpunkte mit einer Genauigkeit im Pixelbereich ab.

In einer industriellen Anwendung hat diese Art der Entzerrung von Mustern ansprechende Ergebnisse geliefert. Es ist dabei allerdings zu bedenken, daß die Übergänge an den Segmentgrenzen unstetig sind. Eine Voraussetzung, die bei den meisten Bildverzerrungen, auch in diesem Beispiel, nicht gegeben ist.

6.4.1.5 Verwendung anderer Orthonormalbasen von Polynomräumen

Ein großer Bereich der Numerik beschäftigt sich mit der Approximation über orthogonale Polynombasen. Gerade unter dem Gesichtspunkt der numerischen Stabilität, die einige der in dieser Arbeit vorgestellten Anwendungen der Mehrniveaupassung vermissen lassen (vgl. etwa Abbildung 5.13, S. 246), bietet sich eine genauere Untersuchung des Approximationsverhaltens an.

Die Beispiele 3.4.12 iv) und 3.5.11 iv) zeigen die Verwandtschaft der vektoriellen Mehrniveaupassung zur Passung über orthogonale Polynombasen. Viele Konzepte der Approximation über orthogonale Polynombasen, gerade im algorithmischen und numerischen Bereich, können auch für die Approximation der vektoriellen Mehrniveaupassung auf den einzelnen Segmenten verwendet werden. Ansätze wie die 3-Term-Rekursion oder Update-Verfahren in gleitenden Zeitfenstern (vgl. [Fuc99]) lassen sich sicher leicht innerhalb der Theorie der vektoriellen Mehrniveaupassung umsetzen. Auch Orthonormalbasen zu anderen Skalarprodukten als dem Standardskalarprodukt des $L^2(\mathcal{B}(M), \mathcal{B}(\mathbb{R}))$ stehen nicht unbedingt im Widerspruch zur Definition 3.4.4 der Familie der erlaubten Skalarprodukte.

Auch im Bereich der Approximation über orthogonalen Polynome gibt es Ansätze für Update-Algorithmen, die die Zunahme bzw. Abnahme der zu approximierenden Daten berücksichtigen (vgl. z. B. [EGK91]). Diese Ansätze könnte man, zumindest für Bilder mit eindimensionalem Definitionsbereich, in die Anwendungen der vektoriellen Mehrniveaupassung einbringen, um numerische Instabilitäten zu reduzieren.

6.4.2 Theoretischer Ausblick

Eine Möglichkeit, den theoretischen Ansatz im Rahmen der Bildsegmentierung zu erweitern, besteht darin, wie beim Mumford-Shah-Modell zusätzlich die Länge der Segmentgrenzen der Segmentierungen zu minimieren. Das heißt, bei einem zweidimensionalen Definitionsbereich wäre die Funktion

$$\mathfrak{S}_k(\mathfrak{A}) \rightarrow \mathbb{R} : \mathbf{A} \mapsto \left\| f - \sum_{A \in \mathbf{A}} f_A \right\|_M^2 + \sum_{A \in \mathbf{A}} \mathcal{H}^1(\partial A) \quad (6.1)$$

zu minimieren, wobei ∂A den Rand der Menge A bezeichne. Dabei sollten die Ränder der Mengen des Mengensystem \mathfrak{A} Hausdorff-Dimension 1 haben. Die Bemerkungen 3.5.1 und 3.5.7 zeigen, daß die Minimierung der Funktion (6.1) zur Maximierung der Funktion

$$\mathfrak{S}_k(\mathfrak{A}) \rightarrow \mathbb{R} : \mathbf{A} \mapsto \sum_{A \in \mathbf{A}} \|\gamma_f(A)\|_2^2 - \sum_{A \in \mathbf{A}} \mathcal{H}^1(\partial A)$$

äquivalent ist. Grundsätzlich sollte sich bei dieser neuen Problemstellung in Kapitel 3 nicht viel ändern. Lediglich das Unterkapitel 3.7.2 müßte sicher noch einmal grundlegend überarbeitet werden. Die Algorithmen des Unterkapitels 4.6 („Das Mengemodell der pixelzusammenhängenden Mengen“) bleiben grundsätzlich auch für dieses modifizierte Problem gültig. Lediglich der Wert $\text{crit}(\cdot)$ der Kandidatenpixel müßte neu definiert werden.

Die Einschränkung auf Anwendungen der vektoriiellen Mehrniveaupassung auf Bildsegmentierungen bestimmt auch die Zielrichtung des Kapitels 3, das die theoretischen Grundlagen der Mehrniveaupassung legt. Insbesondere die Forderungen der Definition der erlaubten Skalarprodukte 3.4.4 (S. 85) sind eindeutig auf diese Anwendung hin ausgelegt. Die Arbeit von Verena Lauren ([Lau03]) zeigt hier einen viel allgemeineren Ansatz. Man sollte dabei allerdings bedenken, daß sowohl die Existenzaussage als auch die Algorithmik wesentlich von dieser Definition abhängen. Beides gestaltet sich für einen allgemeineren Ansatz ungleich schwieriger.

Eine weitere Herausforderung im Bereich der Mehrniveaupassung bedeutet die Erweiterung auf Bewegtbilder. Ist $(f_i)_{i=1}^n$ eine Sequenz von Bildern, so könnte eine Aufgabe darin bestehen, eine Segmentierung $\{A_1, \dots, A_k\} \in \mathfrak{S}_k(\mathfrak{A})$ und eine Folge von Transformationsvektoren $(T^{(i)})_{i=1}^n$ zu bestimmen, die den Abstand

$$\sum_{i=1}^n \left\| f_i - \sum_{j=1}^k g_j \mathbf{1}_{T_j^{(i)}(A_j)} \right\|_M^2$$

minimiert.

Außerdem könnte man stetige Anschlüsse an Segmentgrenzen fordern, was zu einer Splinpassung mit freien Breakpoints führt.

In vielen Applikationen der Informatik tauchen Probleme auf, bei denen zwei oder mehr „Parteien“ verhandeln müssen, um insgesamt ein Optimum zu erhalten. Dabei könnte es sich z. B. um mehrere Prozessoren handeln, die Prozesse unter sich aufteilen. Für alle diese Probleme muß der Ansatz, der in Kapitel 3 („Theoretischen Grundlagen der Mehrniveaupassung“) beschrieben wird, auf einer noch abstrakteren Stufe nachvollzogen werden.

6.5 Abschließende Bemerkung

Die in dieser Arbeit vorgestellten Ergebnisse machen deutlich, daß es sich bei dem Ansatz der Bildsegmentierung mit den Methoden der vektoriellen Mehrniveaupassung um einen wertvollen Ansatz handelt, der sich mit bereits etablierten Ansätzen wie dem der Energieminimierung des Mumford-Shah-Modells messen kann. Sicher besitzt die Theorie der vektoriellen Mehrniveaupassung nicht die mathematische Tiefe der Energieminimierung, die ja letztlich auf über 200 Jahre Variationsrechnung zurückgreifen kann, aber das entscheidende Kriterium zur Bewertung sollte die Praxistauglichkeit sein: Wird das Verfahren den gegebenen Problemen gerecht? Was nützt eine Modellierung, wenn die daraus entstehende Theorie an der Anwendung vorbeigeht?

Die Modellierung des Segmentierungsprozesses als Energieminimierung im Sinne von Mumford und Shah wird dem Problem sicher gerecht. Beim Übergang von der Theorie zur Algorithmik hat dieser Ansatz allerdings Probleme durch die sehr allgemein formulierte Modellierung und muß daher Einbußen in der Reichhaltigkeit in Kauf nehmen. Letztlich befinden sich streng genommen die Algorithmen, die innerhalb der Energieminimierung präsentiert werden, außerhalb der eigenen Theorie.

Bei der vektoriellen Mehrniveaupassung ist dies nicht der Fall. Die Einschränkungen werden hier auf der Theorieseite vollzogen: Es wird nur ein eingeschränktes, erlaubtes System von Segmentierungen betrachtet. Dieses ist so konzipiert, daß sich Segmentierungen auf Pixelebene als Untermenge von Segmentierungen mit kontinuierlichem Definitionsbereich modellieren lassen. Die Theorie der vektoriellen Mehrniveaupassung ist daher näher an der Anwendung und der algorithmischen Umsetzung. Dennoch ist diese Theorie so reichhaltig, daß Segmentierungsalgorithmen von Schwellwertverfahren bis hin zur „Split and Merge“-Techniken und letztlich auch energieminimierende mit eingeschränktem Funktionenraum innerhalb der Theorie modelliert werden können.

Liste der Algorithmen

1	(GenP) Berechnung einer Funktionsquantisierung	53
2	(FP) Berechnung einer Funktionsquantisierung einer monotonen Funktion .	55
3	(ExaktQ) Hilfsalgorithmus zur Berechnung der exakten Funktionsquantisierung einer stückweise konstanten Funktion	58
4	(SEG) Berechnung von Schwellwerten über die Funktionsquantisierung des Histogramms eines diskreten Grauwertbildes g	61
5	(Gram-Schmidt) Berechnung einer Orthonormalbasis (e_1^A, \dots, e_n^A) aus einer Basis $\mathcal{E} = (e_1, \dots, e_n)$ bezüglich des Skalarprodukts $\langle \cdot \cdot \rangle_A$	88
6	(k-Niveaupassung) Berechnung eines pixelbasierten lokalen Maximums einer k -Niveaupassung Ψ	161
7	(Iteration) Berechnung eines pixelbasierten lokalen Maximums einer k -Niveaupassung Ψ	164
8	(2-Niveaupassung) Berechnung eines pixelbasierten lokalen Maximums einer 2-Niveaupassung Ψ bezüglich eines pixelbasierten Systems von Segmentierungen $S_2(\mathcal{A})$	165
9	(1D-k-Niveaupassung) k -Niveaupassung bezüglich der pixelbasierten Segmentierung eines Intervalls durch k zusammenhängende Segmente	171
10	(2-Niveauquantisierung) Berechnung einer optimalen Quantisierung für einen Ausschnitt eines Histogramms h_g	185
11	(Histogrammanalyse) Zerlegung des Histogramms in k Segmente	187
12	(Schwellwertsegmentierung) Segmentierung eines diskreten Pixelbildes über gegebene Schwellwerte	200
13	(Initiales Farbschwellwertverfahren) Berechnung einer Startsegmentierung durch ein kanalweises Schwellwertverfahren	201
14	(k-NiveaupassungZ) Berechnung eines pixelbasierten lokalen Maximums einer k -Niveaupassung Ψ bezüglich des Mengensystems $S_k(\mathfrak{Z})$ und der Klasse der erlaubten Funktionen V	204
15	(Funktionsraum-Relaxation) Berechnung eines pixelbasierten lokalen Maximums einer k -Niveaupassung Ψ bezüglich des Mengensystems $S_k(\mathfrak{Z})$ und der Klasse der erlaubten Funktionen V , wobei V eine erlaubte Basis zur Farbbildsegmentierung besitzt	207

-
- 16 (**Iterative Niveauverschmelzung**) Iterative Verschmelzung von einzelnen Segmenten 226
- 17 (**Stärkster Anstieg**) Schätzung der Anzahl der Segmente über den stärksten Anstieg des Approximationsfehler 236
- 18 (**Anzahlschätzer**) Schätzung der Anzahl der Segmente über den größten normalisierten Fehler 241
- 19 (**Epsilon**) Schätzung der Anzahl der Segmente über den Sensorfehler . . . 250

Index

- (e_1^A, \dots, e_n^A) , 88
- Φ , 92
- $\hat{\Phi}$, 100
- $\mathbb{P}(A)$, 5
- $\mathbb{P}(M)$, 5
- $\mathcal{P}(M)$, 5
- Ψ , 95
- $\mathfrak{S}_k(\mathfrak{A}, A)$, 76
- $\mathfrak{S}_{A_i, A_j}(\mathbf{A})$, 124
- $\mathfrak{S}_k(\mathfrak{A})$, 76
- $\mathfrak{Z}_k(\mathfrak{A})$, 110
- δ_Z , 111
- \mathfrak{d} , 71, 77
- γ_f , 92
- $\tilde{\gamma}_f$, 96
- $\|f - \cdot\|_{\text{quant}}$, 49
- \prec, \preceq , 224
- $\langle \cdot | \cdot \rangle_A$, 85
- $\sigma(\mathfrak{A})$, 84
- \sqcap , 72
- \sqsubset , 72
- \mathfrak{G} , 181
- 2-Niveaupassung, 63
- 2-Niveaupassung mit konstanten Funktionen, 97, 142
- 4-er Nachbarschaft, 7
- 8-er Nachbarschaft, 7
- O , 71
- \mathfrak{A} , 71
- \mathfrak{Z} , 192
- \mathbf{A} , 129
- Ascoli und Arzelá, Satz von, 106
- auf die Zwei-Niveau-Situation reduzierbar, 162
- ausgezeichnete Orthonormalbasis, 88
- \mathbf{B} , 129
- bester Nachbar, 202
- BestNeighbor(p), 202
- Breakpoints, 50
- \mathbf{C} , 129
- CIEL* a^*b^* -Farbmodell, 11
- CIEL* u^*v^* -Farbmodell, 11
- CMY-Farbmodell, 9, 10
- Color Indexing, 27
- $\text{crit}(p)$, 203
- $d_{A,B}$, 146
- Dichteeigenschaft, 74
- Dichteeigenschaft (DE), 74
- diskretes Farbpixelbild, 6
- diskretes Grauwertpixelbild, 6
- diskretes Pixelbild, 6
- E^* (konvexe Approximation der Energiefunktion), 39
- \mathcal{E}^A , 88
- $\mathbf{E}_C^{A,B}$, 133
- $\mathbf{E}_B^{A,\varepsilon}$, 133
- echtes Segment, 76
- einfach pixelzusammenhängend, 219
- einfach zusammenhängend, 219
- Energiefunktional, 33
- Energieterm, 32
- erlaubte Basis zur Farbbildsegmentierung, 198
- erlaubtes Mengensystem, 74
- erweiterte lokale Pixelzusammenhangsüberprüfung, 196
- $\mathcal{F}(\mathfrak{S}_k(\mathfrak{A}), V)$, 101
- \bar{f} , 129
- f_A , 99
- $f_{A,B,C}$, 129
- Farbhistogramm, 27, 190
- farbwertverbunden, 190
- finite Elemente, 36

- Fréchet und Kolmogorov, Satz von, 106
 Funktional, 31
 Funktionsquantisierung, 49
- $\mathcal{G}_k(\mathfrak{A})$, 113
 GNC-Algorithmus, 39, 42
 größter normalisierter Fehler, 238
 grauwertbenachbarte Pixelmengen, 181
 grauwertverbunden, 180
 gray-level contrast measure (**GV**), 28
 gray-level uniformity measure (**GU**), 27
- Histogramm, 6
 HSI-Farbmodell, 10
 HSV-Farbmodell, 10
- inneres lokales Minimum, 60
- \mathcal{K}_B , 88
 $K_\varepsilon(a)$, 124
 Kand(A), 195
 Kand(\mathbf{A}), 195
 Kandidatenpixel, 195
 Kantendetektion, 262
 Knotenpunkte, 50
 kritische Punkte, 53
 kritischer Wert, 203
- leeres Segment, 76
 line process, 37
 lokale Pixelzusammenhangsüberprüfung,
 196
- Mattscheibe, 5
 Mehrzieloptimierung, 223
 Menge der Segmentierungen auf M durch
 Funktionen aus V , 101
 Metrik der geordneten Segmentierungen,
 111
 Metrik der Segmentierungen, 77
 Metrik des Segments A bezüglich des Seg-
 ments B , 146
 Multiskalen-Algorithmen, 42
 Mumford-Shah-Ansatz, 33
 Mumford-Shah-Modell, 32
- $N_4([x, y])$, 7
 $N_8([x, y])$, 7
 N_* , 192
- $n_A(f, g), n_{\mathbf{A}}(f, g)$, 238
 Nachbarschaft, 7
 Niveau, 50
 Niveaumenge, 76
 normalisierter Fehler, 238
- optimale Funktionsquantisierung, 50
 optimale Segmentierung, 101
 optimale Segmentierung, Variationsansatz,
 33
 Otsu-Verfahren, 20
- Pareto-Dominanz, 224
 Pareto-Optimum, 224
 Pfad, 7
 $p(g)_A, p(g)_{\mathbf{A}}$, 249
 Pixel, 5
 pixelbasierte Mengen, 5
 pixelbasierte Umgebung, 158
 pixelbasiertes lokales Maximum, 158
 pixelzusammenhängend, 7
 Pixelzusammenhangskomponente, 8
 Primärfarben, 9
 probability of error, 29
- Quadtree-Zerlegung, 25
- Randpixel, 193
 Regionenwachstumsverfahren, 24
 rektifizierbar, 33
 relativ kompakt, 105
 RGB-Farbmodell, 9
 $Rn(A)$, 193
- $S_{A_i, A_j}(\mathbf{A})$, 161
 $S_k(\mathfrak{A})$, 150
 $S_k(\mathfrak{Z})$, 193
 \mathcal{S}_n , 50
 Satz von Ascoli und Arzelá, 106
 Satz von Fréchet und Kolmogorov, 106
 Schwellwertverfahren, 19
 $seg_{\mathbf{A}}$, 202
 Segment, 76
 Segment, Variationsansatz, 33
 segmentation energy, 31
 Segmentierung, 14
 Segmentierung durch Farbklassen, 27
 Segmentierung, Variationsansatz, 33

-
- Segmentierungsenergie, 31
Sekundärfarben, 9
Snakes, 46
Split and Merge, 24
strenges lokales Maximum, 123
Subpixelgenauigkeit, 6
System der geordneten Segmentierungen,
110
- $\mathcal{U}_p(\mathbf{A})$, 158
 $\mathcal{U}_2(\mathbf{A})$, 125
 $\mathcal{U}_2^\varepsilon(\mathbf{A})$, 125
 $\hat{\mathcal{U}}_2^\varepsilon(\mathbf{A})$, 126
- V , 84
 $V|_A$, 84
Varianzanalyse, 20
Variationsrechnung, 31
- YIQ-Farbmodell, 10
YUV-Farbmodell, 10
- Zielfunktion einer k -Niveaupassung, 103
Zusammenhangskomponente, 72

Literaturverzeichnis

- [AB94] ADAMS, R.; BISCHOF, L.: *Seeded Region Growing*; IEEE Transactions on Pattern Analysis and Machine Intelligence; Bd. 16 (6), S. 641–647, 1994
- [AHU83] AHO, A. V.; HOPCROFT, J. E.; ULLMAN, J. D.: *Data Structures and Algorithms*; Addison-Wesley, 1983
- [Bar88] BARNSELY, M. F.: *Fractals Everywhere*; Academic Press, Boston, 1988
- [Bau92] BAUER, H.: *Maß- und Integrationstheorie*; 2. Aufl.; DeGruyter, 1992
- [BCG96] BALLESTER, C.; CASELLES, V.; GONZALES, M.: *Affine Invariant Segmentation by Variational Method*; SIAM Journal on Applied Mathematics; Bd. 56 (1), S. 294–325, 1996
- [BF70] BRICE, C. R.; FENNEMA, C. L.: *Scene Analysis Using Regions*; Artificial Intelligence; Bd. 1, S. 205–226, 1970
- [BR99] BALTHASAR, D.; REHRMANN, V.: *Robustes histogrammbasiertes Farbmatching*; in: *FarbBV' 99, 5. Workshop Farbbildverarbeitung*; S. 51–58; Technische Universität Illmenau, 1999
- [Bra86] BRAESS, D.: *Nonlinear Approximation Theory*; Springer, 1986
- [BS91] BRONSTEIN, I. N.; SEMENDJAJEW, K. A.: *Taschenbuch der Mathematik*; 25. Aufl.; Teubner, Stuttgart, Leipzig, 1991
- [BZ87] BLAKE, A.; ZISSERMAN, A.: *Visual Reconstruction*; MIT Press, Cambridge MA, 1987
- [CK72] CHOW, C. K.; KANEKO, T.: *Automatic Boundary Detection of the Left Ventricle from Cineangiograms*; Computers and Biomedical Research; Bd. 5, S. 388–410, 1972
- [CM97] COMANICIU, D.; MEER, P.: *Robust Analysis of Feature Spaces: Color Image Segmentation*; in: *IEEE Conference on Computer Vision and Pattern Recognition*; S. 750–755; San Juan, Puerto Rico, June 1997
- [Die85] DIEUDONNE, J. A.: *Grundzüge der modernen Analysis*; Bd. 1; 3. Aufl.; Vieweg, Braunschweig, 1985

- [Don91] DONNER, K.: *Hierarchical Approximation for Pattern Recognition*; in: *Applications of Artificial Intelligence*; S. 213–226; Prague, 1991
- [Edg98] EDGAR, G. A.: *Integral, Probability, and Fractal Measures*; Springer, New York, 1998
- [EGK91] ELHAY, S.; GOLUB, G.; KAUTSKY, J.: *Updating and Datedating of Orthogonal Polynomials with Data Fitting Applications*; SIAM Journal on Matrix Analysis and Applications; Bd. 12, S. 327–353, 1991
- [FBS90] FALK, D. S.; BRILL, D. R.; STORK, D. G.: *Ein Blick ins Licht: ein Einblick in die Natur des Lichts und des Sehens*; Birkhäuser, Springer, Basel, Berlin, 1990
- [Fis86] FISCHER, G.: *Lineare Algebra*; 9. Aufl.; Vieweg, Braunschweig, 1986
- [Flo81] FLORET, K.: *Maß- und Integrationstheorie*; Teubner, Stuttgart, 1981
- [Fuc99] FUCHS, E.: *Schnelle Quadratmittelapproximation in gleitenden Zeitfenstern mit diskreten orthogonalen Polynomen*; Dissertation; Universität Passau, 1999
- [GG84] GEMAN, S.; GEMAN, D.: *Stochastic relaxation, Gibbs distributions, and the Bayesian Restoration of Images*; IEEE Transactions on Pattern Analysis and Machine Intelligence; Bd. 6, S. 721–724, 1984
- [GG92] GERSHO, A.; GRAY, R. M.: *Vector Quantization and Signal Compression*; Kluwer, Boston, 1992
- [Gim99] GIMEL'FARB, G. L.: *Image Textures and Gibbs Markov Random fields*; Nr. 16 in *Computational Imaging and Vision*; Kluwer, 1999
- [GW92] GONZALES, R. C.; WOODS, R. E.: *Digital Image Processing*; Addison-Wesley, 1992
- [Haa00] HAAS, J.: *Echtzeit-Korrespondenzprobleme in Bildsequenzen und Subpixelgenauigkeit*; Dissertation; Universität Passau, 2000
- [Haf98] HAFNER, W.: *Segmentierung von Video-Bildfolgen durch Adaptive Farbklassifikation*; Dissertation; Technische Universität München, 1998
- [Han00] HANNING, T.: *Nonparametric Segmentation of Grayvalue Images by n-level-set-fitting*; in: *World Multiconference on Systemics, Cybernetics and Informatics 2000*; Bd. V; S. 90–95; Orlando, July 2000
- [HFKL01] HANNING, T.; FARR, H.; KELLNER, M.; LAUREN, V.: *Segmentation of vector images by n-level-set-fitting*; in: *IEEE International Conference of Image Processing 2001*; Bd. II; S. 793 – 796; Thessaloniki, October 2001
- [HP76] HOROWITZ, S. L.; PAVLIDIS, T.: *Picture Segmentation by a Tree Traversal Algorithm*; Journal of the ACM; Bd. 23, S. 368–388, 1976
- [HS92] HARALICK, R. M.; SHAPIRO, L. G.: *Computer and Robot Vision*; Bd. 1; Addison-Wesley, 1992

- [Jai89] JAIN, A. K.: *Fundamentals of Digital Image Processing*; Prentice-Hall, 1989
- [Käm00] KÄMPKE, T.: *Optimal and Near Optimal Quantization of Integrable Functions*; Computers and Mathematics with Applications; Bd. 40, 2000
- [Käm01] KÄMPKE, T.: *Efficient Quantization Algorithms for Discrete Functions*; eingereicht, 2001
- [KK98] KÄMPKE, T.; KOBER, R.: *Nonparametric Optimal Binarization*; in: *Proceedings of the 14th International Conference on Pattern Recognition ICPR*; Brisbane, 1998
- [KK99] KÄMPKE, T.; KOBER, R.: *Discrete Signal Quantization*; Pattern Recognition; Bd. 32 (4), S. 619–634, April 1999
- [KLM94] KOEPFLER, G.; LOPEZ, C.; MOREL, J.: *A Multiscale Algorithm for Image Segmentation by Variational Method*; SIAM Journal on Numerical Analysis; Bd. 31 (1), S. 294–325, 1994
- [Kol94] KOLLER, W.: *Segmentierung und Diagnose von Gefäßen mit Atlasunterstützung*; Dissertation; Universität Passau, 1994
- [Kus92] KUSTERER, S.: *Querschnittspassung von Blutgefäßen in MR-Angiographien*; Diplomarbeit; Universität Passau, 1992
- [KWT87] KASS, M.; WITKIN, A.; TERZOPOULOS, D.: *Snakes: Active Contour Models*; in: *1st International Computer Vision Conference.*; S. 259–264, 1987
- [Lau03] LAUREN, V.: *Semilineare Approximation in der Bildrekonstruktion*; Dissertation; Universität Passau, voraussichtlich 2003
- [LCP90] LEE, S. U.; CHUNG, S. Y.; PARK, R. H.: *A Comparative Performance Study of Several Global Thresholding Techniques*; Computer Vision Graphics and Image Processing; Bd. 52, S. 171–190, 1990
- [Lec89] LECLERC, Y.: *Constructing Stable Descriptions for Image Partitioning*; International Journal of Computer Vision; Bd. 3, S. 73–102, 1989
- [LN85] LEVINE, M. D.; NAZIF, A.: *Dynamic Generated Image Segmentations*; IEEE Transactions on Pattern Analysis and Machine Intelligence; Bd. 7, S. 155–164, 1985
- [Mar82] MARR, D.: *Vision*; Freeman, 1982
- [Mat96] MATAS, J.: *Colour Based Object Recognition*; Dissertation; University of Surrey, 1996
- [MIN⁺00] MAEDA, J.; ISHIKAWA, C.; NOVIANTO, S.; TADEHARA, N.; SUZUKI, Y.: *Rough and Accurate Segmentation of Natural Color Images Using Fuzzy Region-Growing Algorithm*; in: *15th International Conference on Pattern Recognition*; Bd. 3; S. 642–645; Barcelona, 2000

- [MJ97] MEHNERT, A.; JACKWAY, P.: *An Improved Region Growing Algorithm*; Pattern Recognition Letters; Bd. 18, S. 1065–1071, 1997
- [MMS92] MASO, G. D.; MOREL, J. M.; SOLIMINI, S.: *A Variational Method in Image Segmentation: Existence and Approximation Results*; Acta Mathematica; Bd. 168, S. 89–151, 1992
- [MS85] MUMFORD, D.; SHAH, J.: *Boundary Detection by Minimizing Functionals*; in: *Proceedings of the IEEE Conference on Computer vision and pattern recognition*; S. 22–26; San Francisco, 1985
- [MS89] MUMFORD, D.; SHAH, J.: *Optimal Approximation by Piecewise Smooth Functions and Associated Variational Problems*; Communications Pure and Applied Mathematics; Bd. 42, S. 577–684, 1989
- [MS94] MUMFORD, D.; SHAH, J.: *Pattern Theory: a Unifying Perspective*; in: *First European Congress of Mathematics*; S. 187–224; Springer, San Francisco, 1994
- [MS95] MOREL, J.-M.; SOLIMINI, S.: *Variational Methods for Image Segmentation*; Birkhäuser, Boston, 1995
- [NM90] NITZBERG, M.; MUMFORD, D.: *The 2.1-D Sketch*; in: *3. International Conference on Computer Vision*; S. 138–144, 1990
- [OKS80] OHTA, Y.; KANDAHE, T.; SAKAI, T.: *Color Information for Region Segmentation*; Computer Graphics Image Processing; Bd. 13, S. 222–241, 1980
- [OPR79] OHLANDER, R.; PRICE, K.; REDDY, D. R.: *Picture Segmentation Using a Recursive Region Splitting Method*; Computer Graphics and Image Processing; Bd. 8 (4), S. 313–333, 1979
- [Ots79] OTSU, N.: *A Threshold Selection Method from Gray-Level Histograms*; IEEE Transactions of Systems Man and Cybernetics; Bd. 9, S. 62–66, 1979
- [PB99] PALUS, H.; BERESKA, D.: *Region-based Colour Image Segmentation*; in: *FarbBV' 99, 5. Workshop Farbbildverarbeitung*; S. 67–71; Technische Universität Illmenau, 1999
- [Pis02] PISINGER, G.: *Lokale Stützstrukturen zur Transformationspassung von Bildern*; Dissertation; Universität Passau, voraussichtlich 2002
- [PP93] PAL, N. R.; PAL, S. K.: *A Review on Image Segmentation Techniques*; Pattern Recognition; Bd. 26 (9), S. 1277–1294, 1993
- [PSA00] PAPAMARKOS, N.; STROUTHOPOULUS, C.; ANDREADIS, I.: *Multithresholding of Color and Gray-Level Images Through a Neural Network Technique*; Image and Vision Computing; Bd. 18, S. 213–222, 2000
- [Ric90] RICHARDSON, T.: *Scale Independent Piecewise Smooth Segmentation of Images via Variational Methods*; Dissertation; MIT, 1990; MA 02139
- [RK82] ROSENFELD, A.; KAK, A.: *Digital Picture Processing Vol. 2*; 2. Aufl.; Academic Press, 1982

- [Ros70] ROSENFELD, A.: *Connectivity in Digital Pictures*; Journal of the ACM; Bd. 17, S. 146–160, 1970
- [SB91] SWAIN, M. J.; BALLARD, D. H.: *Color Indexing*; International Journal of Computer Vision; Bd. 7 (1), S. 11–32, 1991
- [Sob70] SOBEL, I. E.: *Camera Models and Machine Perception*; Dissertation; Stanford University; Stanford, CA, 1970; Electrical Engineering Department
- [Soi98] SOILLE, P.: *Morphologische Bildverarbeitung: Grundlagen, Methoden, Anwendungen*; Springer, 1998
- [Spe93] SPELLUCCI, P.: *Numerische Verfahren der nichtlinearen Optimierung*; Birkhäuser, 1993
- [SSWC88] SAHOO, P. K.; SOLTANI, S.; WONG, A.; CHEN, Y.: *A Survey of Thresholding Technologies*; Computer Vision Graphics and Image Processing; Bd. 41, S. 233–260, 1988
- [TJ95] TRIER, O. D.; JAIN, A. K.: *Goal-directed Evaluation of Binarization Methods*; IEEE Transactions of Pattern Analysis and Machine Intelligence; Bd. 17, S. 1191–1201, 1995
- [Vos93] VOSS, K.: *Discrete Images, Objects, and Functions in Z^n* ; Algorithms and Combinatorics; Springer, Berlin, 1993
- [Wei98] WEICKERT, J.: *Anisotropic Diffusion in Image Processing*; Teubner, Stuttgart, 1998
- [Wer95] WERNER, D.: *Funktionalanalysis*; Springer, Berlin, 1995
- [Wol00] WOLF, M.: *Posturerkennung starrer und nichtstarrer Objekte*; Dissertation; Universität Erlangen-Nürnberg, 2000
- [WS82] WYSZECKI, G.; STILES, W. S.: *Color Science: Concepts and Methods, Quantitative Data and Formulae*; 2. Aufl.; Wiley, New York, 1982
- [WZX92] WANG, T.; ZHUANG, X.; XING, X.: *Robust Segmentation of Noisy Images Using a Neural Network Model*; Image and Vision Computing; Bd. 10, S. 233–240, 1992
- [YMB77] YANSOFF, W. A.; MUI, K. K.; BACUS, J. W.: *Error Measures for Scene Segmentation*; Pattern Recognition; Bd. 9, S. 217–231, 1977
- [Yos78] YOSIDA, K.: *Functional Analysis*; 5. Aufl.; Springer, Berlin, 1978
- [Zha96] ZHANG, Y. J.: *A Survey on Evaluation Methods for Image Segmentation*; Pattern Recognition; Bd. 29 (8), S. 1335–1346, August 1996
- [Zit99] ZITZLER, E.: *Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications*; Dissertation; ETH Zürich, 1999; TIK-Schriftenreihe Nr. 30