



Westfälische Wilhelms-Universität Münster

Anwendung optimaler Steuerungsprobleme mit L^∞ -Steuerbeschränkung auf ein Modellproblem der Bildverarbeitung

Diplomarbeit

Institut für Numerische und Angewandte Mathematik

vorgelegt von
Lucas Franek

Betreuer:
Prof. Dr. H. Maurer
PD Dr. M. Wagner

Inhaltsverzeichnis

1	Einleitung	1
2	Problemstellung und mathematische Grundlagen	4
2.1	Die Modellierung des Problems – Bildstörungen	4
2.2	Schlecht gestellte Probleme	6
2.3	Funktionenräume	6
2.4	Repräsentation von Bildern – Diskretes und kontinuierliches Bildmodell . .	8
2.4.1	Diskretes Bildmodell	8
2.4.2	Kontinuierliches Bildmodell - geeignete Funktionenräume	11
2.4.3	Übergänge zwischen den Modellen	12
3	Variationsprobleme und Probleme der optimalen Steuerung	14
3.1	Variationsprobleme	14
3.1.1	Existenzsätze	15
3.1.2	Notwendige Optimalitätsbedingungen	16
3.2	Probleme der optimalen Steuerung	18
3.2.1	Das Pontrjaginsche Maximumprinzip für mehrdimensionale Steuerungsprobleme	19
4	Variationsmethoden in der Bildverarbeitung	21
4.1	Der Variationsansatz zur Lösung des image-restoration-Problems	21
4.1.1	Herleitung des Variationsansatzes	21
4.1.2	Die Wahl der Potentialfunktion f	23
4.1.3	Existenz und Eindeutigkeit einer Lösung	25
4.1.4	Anwendung der Euler-Lagrange-Gleichungen	26
4.2	Das Kantenerkennungsfunktional	27
4.3	Fallstudie: Impulsrauschen	29
4.3.1	Impulsrauschen in der Forschung	29
4.3.2	Entzerrung und Impulsrauschen – L^1 -Datenterm und Mumford-Shah-Regularisierungsterm	32
4.4	Steuerung und Steuerbeschränkung im Variationsansatz	34
4.5	L^∞ -Steuerbeschränkung und „rotating mask“	36
4.5.1	Modellierung für kontinuierliche Bilder	36
4.5.2	Modellierung für diskrete Bilder	37
5	Numerische Verfahren – Minimierungsmethoden	42
5.1	Diskretisierungsschemata	42
5.2	Indirekte Verfahren	43
5.2.1	Gradientenabstiegsverfahren	44
5.2.2	Fixpunktiteration	45

5.2.3	Herleitung eines iterativen Schemas für das Kantenerkennungsfunktional	46
5.3	Direkte Verfahren	48
5.4	Nichtlineare Optimierungsprobleme – Notwendige Optimalitätsbedingungen	50
5.5	Innere-Punkt-Verfahren – IPOPT und LOQO	51
6	Numerische Resultate	54
6.1	Bewertungsmaßstäbe für die Ergebnisse	54
6.2	Kantenerkennung	56
6.3	Quadratische und TV-Regularisierung	63
6.4	Steuerungsproblem mit L^∞ -Steuerbeschränkung	63
6.5	Steuerungsprobleme mit rotating mask	76
6.5.1	Der eindimensionale Fall: Anwendung auf Signale	76
6.5.2	Der zweidimensionale Fall	77
6.5.3	Fallbeispiel: Struktur mit kleinen Variationen und Salt-and-Pepper-Rauschen	80
6.5.4	Lena-Testbild	83
6.6	Diskussion der Resultate	87
7	Zusammenfassung und Ausblick	88
A	Beispiel eines AMPL-Programms	90
B	Tabellen	93
	Danksagung	101
	Eidesstattliche Erklärung	102
	Abbildungsverzeichnis	103
	Literaturverzeichnis	105

1 Einleitung

Motivation

In dieser Arbeit wollen wir optimale Steuerungsprobleme mit L^∞ -Steuerbeschränkung untersuchen und mit ihrer Hilfe zwei Grundprobleme der Bildverarbeitung behandeln. Wir setzen uns zum Ziel, ein verrauschtes Bild zu entrauschen und außerdem die Kanten oder Objektgrenzen der Bildobjekte zu extrahieren. Diese beiden Probleme sind in der Bildverarbeitung unter dem Namen *image-restoration-Problem* bzw. *image-segmentation-Problem* bekannt.

Das erste Problem, d.h. die Bildwiederherstellung oder -entrauschung, dient in der Praxis oft als Vorverarbeitungsschritt, um die Qualität eines Bildes zu verbessern. Dieser ist notwendig, da Bilder oft in einer gestörten Form vorkommen, wobei Bildstörungen auf vielfältige Weise entstehen können. Zum Beispiel entstehen gestörte Satellitenbilder durch atmosphärisches Rauschen, verzerrte Bilder, wenn sich Kamera und Objekt relativ zu einander bewegen. Selbst moderne Digitalkameras können bei schlechten Lichtverhältnissen verrauschte Bilder produzieren, sodass auf die Bildentrauschung trotz technischem Fortschritt nicht immer verzichtet werden kann.

Es gibt sehr viele unterschiedliche Ansätze, um das Bild zu entrauschen und seine Qualität somit in einem gewissen Sinne zu verbessern (vgl. [7]). In dieser Arbeit beschäftigen wir uns vor allem mit dem Variationsansatz zur Lösung dieses Problems. Der Variationsansatz basiert auf der Minimierung eines jeweils geeigneten Energiefunktional und hat die Verbesserung der Bildqualität zum Ziel. Bezeichnet $I : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ das gestörte und $x : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ das gesuchte Bild, so hat das Energiefunktional die folgende Gestalt:

$$F(x) = \int_{\Omega} (I(s) - S(x(s)))^2 ds + \mu \int_{\Omega} f(|\nabla x(s)|) ds, \quad x \in W_0^{1,p}(\Omega). \quad (1.1)$$

Dabei ist S eine geeignete Transformation der Funktion x , und die Funktion f wird so gewählt, dass unerwünschte Bildeigenschaften bestraft werden. Funktionale dieser Gestalt sind im Bereich der Bildverarbeitung schon seit etwa 20 Jahren Gegenstand der Forschung und somit bereits gut untersucht. Der Ansatz in dieser Arbeit besteht nun darin, eine Steuerung zu erklären und im Zusammenhang mit diesen Funktionalen eine Steuerbeschränkung zu betrachten. Wir erhalten also ein Problem der optimalen Steuerung, was laut Kenntnisstand des Verfassers in Verbindung mit dem *image-restoration-* und *image-segmentation-Problem* noch nicht untersucht wurde. Wir führen also zunächst eine künstliche Steuerung $u : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ein, indem wir sie durch

$$u(s) = (u_1(s), u_2(s))^T := \nabla x(s) \quad (1.2)$$

erklären. Wir betrachten dann Beschränkungen der $\|\cdot\|_q$ -Norm von u für $1 \leq q < \infty$ und

erhalten somit insgesamt das folgende Steuerungsproblem:

$$\begin{aligned}
 &\text{Minimiere } F(x, u) = \int_{\Omega} (I(s) - S(x(s)))^2 ds + \mu \int_{\Omega} f(|u(s)|) ds, \\
 &(x, u) \in W_0^{1,p}(\Omega) \times L^p(\Omega, \mathbb{R}^2), \\
 &\text{unter } u(s) = (u_1(s), u_2(s))^T = \nabla x(s), \\
 &u(s) \in K := \{(v_1, v_2) \in \mathbb{R}^2 \mid |v_1|^q + |v_2|^q \leq R^q\} \quad (\forall) s \in \mathbb{R}.
 \end{aligned} \tag{1.3}$$

Dieses Steuerungsproblem ist ein Problem vom Dieudonné-Rashevsky-Typ. Notwendige Optimalitätsbedingungen für Probleme dieser Gestalt wurden von Wagner [31, 32] entwickelt und liefern somit das theoretische Fundament für diese Probleme.

Motiviert wird die Definition der Steuerung (1.2) durch das zweite Grundproblem der Bildverarbeitung, das in dieser Arbeit behandelt wird. Wir setzen uns zum Ziel, neben der Bildwiederherstellung die Kanten des Bildes zu bestimmen, behandeln also das image-segmentation-Problem. Um dieses Ziel zu erreichen, reicht es nicht aus, das Funktional (1.1) zu betrachten. Um ein Kantenbild $k : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ zu erhalten, wurden aus diesem Grund komplexere Energiefunktionale entwickelt. Ein Ansatz, der von Ambrosio/Tortorelli [2] vorgeschlagen wurde, besteht darin, für ein festes $\epsilon > 0$ und Parameter c_1, c_2, c_3, c_4 das folgende Funktional zu minimieren:

$$\begin{aligned}
 F_{\epsilon}(x, k) = &c_1 \int_{\Omega} (x(s) - I(s))^2 ds + c_2 \int_{\Omega} |\nabla x(s)|^2 \cdot (k(s)^2 + c_4) ds \\
 &+ c_3 \int_{\Omega} \left(\epsilon |\nabla k(s)|^2 + \frac{1}{4\epsilon} (k(s) - 1)^2 \right) ds.
 \end{aligned}$$

Dieses liefert nach Minimierung und geeigneter Parameterwahl ein wiederhergestelltes Bild x und ein Kantenbild k . Wir werden dieses Funktional nutzen, um die dort erhaltenen Kantenbilder mit den Kantenbildern zu vergleichen, die wir durch das Lösen des optimalen Steuerungsproblems (1.3) erhalten. Dabei besteht die Idee grundsätzlich darin, Kantenbilder mit Hilfe der optimalen Steuerung zu definieren. Da wir gerade die Steuerung durch den Bildgradienten definieren und dieser eine große Steigung an Kanten aufweist, enthält er Kanteninformationen, die wir nutzen wollen. Wir erhalten somit durch die Lösung des Steuerungsproblems ein Verfahren, das ein wiederhergestelltes Bild und ein Kantenbild liefert. Je nachdem, welche Art von Steuerbeschränkung wir wählen, erhalten wir auch unterschiedliche Kantenbilder. Insbesondere untersuchen wir in den Versuchsreihen den Fall $q = \infty$, also die Beschränkung der Maximumsnorm von u .

Bei der Untersuchung des Steuerungsproblems und der Durchführung vielfältiger Versuchsreihen werden wir darauf stoßen, dass Steuerungsprobleme insbesondere bei Salt-and-Pepper-Rauschen sinnvoll eingesetzt werden können. Wir werden ein neues Verfahren entwickeln, das es ermöglicht, synthetische Bilder mit einer geringen Rauschintensität zu entrauschen, ohne dabei Kanten zu zerstören. Dabei werden wir unser Verfahren mit einigen Verfahren aus der aktuellen Forschung vergleichen, die in den letzten Jahren unter anderem von Nikolova [22] und Bar/Kiryati/Sochen [5] entwickelt wurden.

In dieser Arbeit lösen wir die Variations- und Steuerungsprobleme mit Hilfe von direkten Verfahren, d.h. wir diskretisieren die Probleme zuerst und wenden anschließend einen Optimierungssolver an. Dieses Verfahren wurde von Theissen [26] für optimale Steuerprozesse mit partiellen Differentialgleichungs-Restriktionen erfolgreich angewandt. Dabei wurden in

[26] semilineare elliptische und parabolische Differentialgleichungen in den Nebenbedingungen untersucht, während die Nebenbedingungen in dieser Arbeit partielle Differentialgleichungen erster Ordnung sind.

In der Literatur werden im Gegensatz zu den direkten Verfahren oft die indirekten Verfahren zur Lösung der Probleme beschrieben. Wir wollen untersuchen, ob wir auch mit Hilfe der direkten Verfahren gute Ergebnisse bei Problemen der Bildverarbeitung erzielen, und unser Vorgehen teilweise auch mit den indirekten Verfahren vergleichen.

Aufbau

In Kapitel 2 werden wir einige Grundlagen der mathematischen Bildverarbeitung, die wir in den darauf folgenden Kapiteln benötigen werden, zusammenstellen. In Kapitel 3 fassen wir die Grundlagen der Theorie der Variationsrechnung sowie der optimalen Steuerungsprobleme zusammen. Anschließend werden wir in Kapitel 4 diese Theorie verwenden, um den Variationsansatz für das image-restoration-Problem herzuleiten. Weiterhin modellieren wir eine Lösung des Problems mit Hilfe von optimalen Steuerungsproblemen. Wir entwickeln ein neues Verfahren – das „Steuerungsproblem mit rotating mask“ – mit dessen Hilfe wir kontrolliertes Salt-and-Pepper-Rauschen eliminieren können. Wir stellen die Methoden vor, die wir als Vergleichsverfahren nutzen werden. In Kapitel 5 behandeln wir die numerische Lösung der Variations- und Steuerungsprobleme. Anschließend stellen wir in Kapitel 6 die Versuchsreihen zusammen, die wir zu den unterschiedlichen Methoden durchgeführt haben. Wir diskutieren die Vor- und Nachteile der unterschiedlichen Ansätze. Zuletzt werden wir in Kapitel 7 das Vorgehen dieser Arbeit kurz zusammenfassen und die wichtigsten Resultate dieser Arbeit wiederholen, sowie einen Ausblick auf offene Probleme liefern.

Literatur zur modernen mathematischen Bildverarbeitung ist vor allem in englischer Sprache verfügbar, sodass einige englische Begriffe des spezifischen Vokabulars nur inadäquat ins Deutsche übersetzt werden können. Häufig werden in dieser Arbeit die englischen Bezeichnungen daher in Klammern mit angegeben oder gar ausschließlich benutzt.

Bis auf das in der Bildverarbeitung bekannte Lena-Testbild wurden sämtliche Bilder und Abbildungen vom Verfasser mit Hilfe von MATLAB erstellt.

2 Problemstellung und mathematische Grundlagen

In diesem Kapitel führen wir einige Begriffe und Grundlagen der mathematischen Bildverarbeitung ein, die wir in dieser Arbeit benötigen werden. Wir erklären, wie Bildstörungen entstehen, und geben dafür ein mathematisches Modell an. Anschließend wollen wir die Frage behandeln, in welchen Funktionenräumen unsere wiederhergestellten Bilder sinnvollerweise aufgesucht werden sollen. Es ist offensichtlich, dass Bilder normalerweise Sprünge in der Farbintensität enthalten, sodass es nicht sinnvoll ist, Bilder als stetig differenzierbar anzunehmen. Wir werden deshalb eine größere Klasse von Räumen betrachten, die sogenannten Sobolevräume.

In diesem Zusammenhang werden wir die unterschiedlichen Repräsentationen von Bildern untersuchen und zwischen dem diskreten und dem kontinuierlichen Bildmodell unterscheiden.

2.1 Die Modellierung des Problems – Bildstörungen

Das Problem der Bildwiederherstellung (engl.: image restoration) ist eines der ältesten und am häufigsten untersuchten Probleme der Bildverarbeitung. Image restoration dient oft als Vorverarbeitungsschritt für Bilder, deren Bildqualität beeinträchtigt wurde. Diese Verminderung der Bildqualität geschieht oft bei der Bildaufzeichnung. Jede Digitalkamera erzeugt zum Beispiel ein sogenanntes Grundrauschen. Dieses entsteht, da die Photodioden bereits durch Wärme oder eine Grundspannung Elektronen abgeben. Bei Tageslicht wird dieses Grundrauschen überdeckt und fällt daher nicht auf. Das Rauschen wird hingegen sichtbar, wenn die Lichtverhältnisse zu schlecht sind. Verrauschte und verzerrte Bilder entstehen auch, wenn sie etwa von einem Satelliten aus dem Weltall aufgezeichnet werden.

Bildstörungen werden in zwei Klassen eingeteilt. Man unterscheidet dazu den systematischen und den zufälligen Fehler. Der **systematische Fehler** ist auf das Verfahren der Bildfassung zurückzuführen und bewirkt ein Verschmieren und Verzerren des Bildes. Dieser Fehler entsteht, wenn sich die Kamera zum Beispiel relativ zu der Szene bewegt. Der **zufällige Fehler** wird im Bild im Rauschen sichtbar und kann meistens durch eine Gauss'sche Verteilung modelliert werden. Um das Problem mathematisch zu formulieren, betrachten wir Bilder als Funktionen einer Menge $\Omega \subset \mathbb{R}^2$ in die reellen Zahlen \mathbb{R} , also

$$x : \Omega \rightarrow \mathbb{R}.$$

Anzumerken ist, dass wir für den eindimensionalen Fall $\Omega \subset \mathbb{R}$ eine Vereinfachung erhalten, und x in diesem Fall eindimensionales Bild oder **Signal** nennen. Wenn es sinnvoll erscheint, werden wir auf diesen Fall zurückgreifen.

Der Störvorgang lässt sich mathematisch ausdrücken. Bezeichnet $\tilde{x} : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ das tatsächliche ungestörte und $I : \Omega \subset \mathbb{R}^2 \rightarrow [a, b] \subset \mathbb{R}$ das aufgezeichnete gestörte Bild, so

lässt sich lineares Rauschen durch den Zusammenhang

$$I(s) = S(\tilde{x}(s)) + T(\tilde{x}(s))N(s), \quad \forall s \in \Omega \quad (2.1)$$

modellieren. Dabei bezeichnen S und T geeignete Transformationen eines Bildes x und $N(s)$ den zufälligen Fehler in einem Punkt $s \in \Omega$. S und T repräsentieren den systematischen Fehler. Zum Beispiel erhält man:

1. **Additives Rauschen** für $S(\tilde{x}) = \tilde{x}$ und $T(\tilde{x}) = 1$; oft ist der Erwartungswert \mathbb{E} des Rauschens Null: $\mathbb{E}(N) = 0$.
2. **Multiplikatives Rauschen** (engl. speckle noise) für $S(\tilde{x}) = 0$ und $T(\tilde{x}) = \tilde{x}$; hierbei gilt oft $\mathbb{E}(N) = 1$. Diese Art Rauschen tritt z.B. bei Radar- und Ultraschallbildern auf.
3. **Verzerrung** (engl. blur) erreicht man, wenn $N = 0$ und S ein Faltungsoperator mit geeignetem Kern ist: $I = S * \tilde{x}$, wobei wir mit $*$ die Faltung einer Funktion bezeichnen.

Den ersten Fall schreiben wir noch einmal explizit auf, da wir meistens ein additives Rauschen mit Erwartungswert 0 annehmen werden:

$$I(s) = S(\tilde{x}(s)) + N(s), \quad \forall s \in \Omega \quad (2.2)$$

Eine weitere wichtige Art des Rauschens ist das **Impulsrauschen** (engl.: impulsive noise). Es gehört zum **nichtlinearen Rauschen** und kann unterteilt werden in das **Salt-and-Pepper-Rauschen** und das **Random-Valued Impulsrauschen**. Impulsrauschen kann durch fehlerhafte Pixel in Kamera-Sensoren, fehlerhafte Speicherplätze in der Hardware oder Transmission in verrauschten Kanälen verursacht werden. Kennzeichnend für Impulsrauschen ist, dass es ungestörte Pixel geben kann. Weiterhin nehmen gestörte Pixel einen Wert an, ganz unabhängig davon, welchen Wert sie vorher hatten. Dies steht im Gegensatz zum additiven Rauschen, wo der Farbwert des ursprünglichen Pixels mit in die Berechnung des verrauschten Pixels eingeht. Beim Salt-and-Pepper-Rauschen nimmt ein verrauschtes Pixel entweder den Maximal- oder den Minimalwert in gegebenem Wertebereich $[a, b] \in \mathbb{R}$ an. Der Minimalwert entspricht dann z.B. schwarzem „Pepper“ und der Maximalwert weißem „Salt“. Das Salt-and-Pepper-Rauschen kann wie folgt mathematisch modelliert werden. Bei einer gegebenen binären Zufallsverteilung $\omega(s), \forall s \in \Omega$, d.h. $\omega(s) = 0$ oder 1 , soll gelten:

$$I(s) = \begin{cases} \tilde{x}(s) & | \text{ wenn } \omega(s) = 0, \\ N(s) & | \text{ wenn } \omega(s) = 1. \end{cases}$$

Dabei ist $N(s)$ ebenfalls eine binäre Zufallsvariable, die die Werte a bzw. b annehmen kann. Den Erwartungswert von ω wollen wir **Rauschintensität** nennen. Im Folgenden betrachten wir nur noch Graustufenbilder mit Wertebereich $[0, 1]$, also $x : \Omega \rightarrow [0, 1]$. Dabei soll der Wert 0 Schwarz und der Wert 1 Weiß darstellen.

Um das image-restoration-Problem zu lösen, d.h. die Originalbilder so gut es geht wiederherzustellen, wurden unterschiedliche Zugänge entwickelt und mathematische Techniken verwendet, wie z.B. Wavelets, statistisch-basierte Filter, partielle Differentialgleichungen und der Variationsansatz (vgl. [7] für einen groben Überblick). In dieser Arbeit wird insbesondere der Variationsansatz betrachtet, jedoch wird auch die enge Verknüpfung zum Ansatz mit partiellen Differentialgleichungen deutlich werden.

2.2 Schlecht gestellte Probleme

Bei dem image-restoration-Problem handelt es sich um ein inverses Problem, d.h. man möchte den Vorgang, der durch das Fehlermodell (2.2) gegeben ist, in gewissen Sinne wieder rückgängig machen, um so wieder das ungestörte Bild zu erhalten. Für uns von Bedeutung ist, dass es sich dabei um ein schlecht gestelltes Problem handelt. Genauer betrachten wir dazu normierte Räume X, Y und einen linearen Operator $A : X \rightarrow Y$. Für gegebenes $g \in Y$ suchen wir dann eine Lösung $f \in X$ der Gleichung

$$Af = g. \tag{2.3}$$

Definition 2.1 (Schlecht gestelltes Problem nach Hadamard):

Das Problem (2.3) ist gut gestellt, wenn die folgenden Eigenschaften gelten

1. Existenz: für alle $g \in Y$ existiert eine Lösung von (2.3),
2. Eindeutigkeit: die Lösung ist eindeutig,
3. f hängt stetig von g ab.

Das Problem heißt schlecht gestellt, wenn es nicht gut gestellt ist.

Nun kann man zeigen, dass – wie bereits in Abschnitt 2.1 angedeutet wurde – die Verzerrung eines Bildes, die in diesem Fall dem linearen Operator A in (2.3) entspricht, durch eine Faltung mit Gauss'schem Kern modelliert werden kann und dieser Faltungsoperator kompakt ist. Das hat zur Folge, dass das Problem schlecht gestellt ist. Dies soll hier jedoch nicht genauer erläutert werden, sondern stattdessen auf [8], pp. 11 verwiesen werden. Informell kann an dieser Stelle noch die folgende Begründung geliefert werden: Die Verzerrung eines Bildes entspricht einem Diffusionsprozess, d.h. der Anwendung einiger Zeitschritte der Wärmeleitungsgleichung $\frac{d}{dt}x = \Delta x$. Die Inversion dieses Prozesses wäre somit ein rückwärtiger Diffusionsprozess $\frac{d}{dt}x = -\Delta x$, der jedoch instabil ist (vgl. [15], pp. 216).

2.3 Funktionenräume

Wir führen einige funktionalanalytische Begriffe und Funktionenräume ein, die wir benötigen werden. Es bezeichne Ω eine offene Teilmenge des \mathbb{R}^n , wobei wir im Folgenden nur solche Gebiete betrachten wollen, die einen hinreichend glatten Rand – einen sogenannten Lipschitzrand – besitzen. Dazu definieren wir:

Definition 2.2 (Lipschitzgebiet):

Es sei $\Omega \subset \mathbb{R}^n$ offen, beschränkt und zusammenhängend. Ω heißt Lipschitzgebiet (im strengen Sinne), wenn zu jedem Randpunkt $s \in \partial\Omega$ eine Umgebung U_s und eine bijektive Abbildung $\varphi_s : U_s \rightarrow Q$ auf einen offenen Quader $Q : (a_1, b_1) \times \dots \times (a_n, b_n) \subset \mathbb{R}^n$ existieren, und die folgenden Bedingungen erfüllen:

1. φ_s ist aus einer Drehung und einer Verschiebung zusammengesetzt
2. φ_s bildet ab:

$$\begin{aligned} U_s \cap \Omega & \text{ auf } Q_s \cap \{t \in \mathbb{R}^n | a_n < t_n < f_s(t_1, \dots, t_{n-1})\} \\ U_s \cap \partial\Omega & \text{ auf } Q_s \cap \{t \in \mathbb{R}^n | a_n < t_n = f_s(t_1, \dots, t_{n-1}) < b_n\} \\ U_s \setminus (\Omega \cap \partial\Omega) & \text{ auf } Q_s \cap \{t \in \mathbb{R}^n | f_s(t_1, \dots, t_{n-1}) < t_n < b_n\} \end{aligned}$$

wobei $f_s : (a_1, b_1) \times \dots \times (a_{n-1}, b_{n-1}) \rightarrow \mathbb{R}$ Lipschitz-stetig sei.

Für $1 \leq p < \infty$ definieren wir den Lebesgueraum $L^p(\Omega)$ durch

$$L^p(\Omega) := \left\{ x : \Omega \rightarrow \mathbb{R} \text{ messbar} \mid \int_{\Omega} |x(s)|^p ds < \infty \right\}. \quad (2.4)$$

Dieser ist zusammen mit der L^p -Norm

$$\|x\|_p := \left(\int_{\Omega} |x(s)|^p ds \right)^{\frac{1}{p}}$$

ein Banachraum. Wenn nicht anders angegeben, so bezeichnen wir mit $\|\cdot\|$ immer die L^2 -Norm. Der Raum der wesentlich beschränkten, messbaren Funktionen $L^\infty(\Omega)$ ist wie folgt definiert:

$$L^\infty(\Omega) := \{x : \Omega \rightarrow \mathbb{R} \text{ messbar} \mid \|x(s)\|_\infty < \infty\}, \text{ wobei}$$

$$\|x\|_\infty := \operatorname{ess\,sup}_{s \in \Omega} |x(s)|.$$

$\operatorname{ess\,sup}(x)$ bezeichnet dabei das Supremum der Funktion x bis auf Mengen vom Maß Null.

Wir führen nun den Begriff der schwachen Ableitung ein, mit dessen Hilfe wir anschließend die Sobolevräume definieren können. Dazu definieren wir zunächst den Raum

$$L^1_{loc}(\Omega) := \left\{ x : \Omega \rightarrow \mathbb{R} \text{ messbar} \mid \int_K |x(s)| ds < \infty, \forall K \subset \Omega, K \text{ kompakt} \right\}.$$

Für $\alpha = (\alpha_1, \dots, \alpha_n)^T \in \mathbb{R}^n$ bezeichne $|\alpha|$ den Multiindex:

$$|\alpha| := \sum_{i=1}^n \alpha_i, \quad \partial^\alpha := \frac{\partial^{|\alpha|}}{\partial s_1^{\alpha_1} \dots \partial s_n^{\alpha_n}}.$$

Es bezeichne weiterhin $C_0^\infty(\Omega) := \{v \in C^\infty(\Omega) \mid \operatorname{supp}(v) \text{ kompakt, } \operatorname{supp}(v) \subset \Omega\}$ den Vektorraum der C^∞ -Funktionen mit kompakten Träger. Er wird auch Raum der Testfunktionen genannt. Mit $(C_0^\infty(\Omega))'$ wollen wir den Raum aller stetigen linearen Funktionale $x : C_0^\infty(\Omega) \rightarrow \mathbb{R}$ bezeichnen. Dieser Raum wird auch Raum der Distributionen genannt. Jetzt können wir den Begriff der schwachen Ableitung einer Funktion $x \in L^1_{loc}(\Omega)$ einführen. Die Funktion $\omega \in L^1_{loc}(\Omega)$ heißt schwache Ableitung der Ordnung $|\alpha|$, wenn gilt:

$$\int_{\Omega} \omega(s) \varphi(s) ds = (-1)^{|\alpha|} \int_{\Omega} x(s) \partial^\alpha \varphi(s) ds, \quad \forall \varphi \in C_0^\infty(\Omega).$$

Wir können ω mit $\partial^\alpha x$ identifizieren, da die schwache Ableitung für $x \in C^{|\alpha|}(\Omega)$ mit der klassischen Ableitung übereinstimmt.

Nun können wir die Sobolevräume definieren, indem wir verlangen, dass die schwachen Ableitungen in den Lebesgueräumen enthalten sind:

Definition 2.3 (Sobolevraum):

Für $1 \leq p \leq \infty$, $k \in \mathbb{N}_0$, definieren wir den Sobolevraum $W^{k,p}(\Omega)$ durch

$$W^{k,p}(\Omega) := \{x \in L^p(\Omega) \mid \exists \partial^\alpha x \in L^p(\Omega), \forall |\alpha| \leq k\}.$$

Ist Ω beschränkt, so ist $W^{1,\infty}(\Omega)$ der Raum der Lipschitzfunktionen, d.h. der Funktionen, die Lipschitz-stetig sind.

Wenn wir einen allgemeineren Wertebereich \mathbb{R}^n der Funktionen aus $L^p(\Omega)$ betrachten wollen, schreiben wir $L^p(\Omega, \mathbb{R}^n)$, und definieren diesen Raum ganz entsprechend zu der Definition des Lebesgueraums (2.4). Dieselbe Schreibweise soll auch für die Sobolevräume und den Raum der Testfunktionen gelten, also $W^{1,p}(\Omega, \mathbb{R}^n)$ bzw. $(C_0^\infty(\Omega, \mathbb{R}^n))'$.

Wir führen nun noch einen weiteren Raum ein, der in der Bildverarbeitung oft verwendet wird. Ist Ω eine beschränkte offene Menge in \mathbb{R}^n , so definieren wir zunächst für $x \in L^1(\Omega)$ die Totalvariation

$$|x|_{TV} := \sup \left\{ \int_{\Omega} x \operatorname{div} \varphi \, ds \mid \varphi \in C_0^\infty(\Omega, \mathbb{R}^n), \|\varphi\|_\infty \leq 1 \right\},$$

wobei $\|\varphi\|_\infty \leq 1$ bedeutet, dass alle Komponenten von φ eine Maximumsnorm kleiner als 1 haben. Ist die schwache Ableitung von x in $L^1(\Omega)$ enthalten, so kann man zeigen, dass in diesem Fall

$$|x|_{TV} = \int_{\Omega} |\nabla x(s)| \, ds \tag{2.5}$$

gilt (vgl. [15], pp. 48). Wir definieren den Raum der Funktionen beschränkter Totalvariation $BV(\Omega)$ als den Raum aller L^1 -messbaren Funktionen mit beschränkter Totalvariation:

Definition 2.4 (Raum der Funktionen beschränkter Totalvariation):

Der Raum

$$BV(\Omega) := \{x \in L^1(\Omega) \mid |x|_{TV} < \infty\} \tag{2.6}$$

ist mit der Norm

$$\|x\|_{BV} := \|x\|_1 + |x|_{TV}$$

ein Banachraum.

2.4 Repräsentation von Bildern – Diskretes und kontinuierliches Bildmodell

In der Bildverarbeitung unterscheidet man das kontinuierliche und das diskrete Bildmodell. Werden Bilder mit Hilfe des Computers bearbeitet, so geschieht diese Verarbeitung letztlich mit diskreten Bildern. Die Problemmodellierung und Analyse des Modells basiert jedoch meistens auf dem kontinuierlichen Bildmodell. Dabei ist es normalerweise nicht von Interesse, Bilder als stetig voranzusetzen, da sie oft Unstetigkeiten enthalten, wie zum Beispiel Kanten. Hingegen wird auf natürliche Weise oft die Beschränktheit von Bildern vorausgesetzt.

2.4.1 Diskretes Bildmodell

Ein diskretes Bild erhält man, indem man von einem äquidistanten $(N_1 \times N_2)$ -Gitter ausgeht und jedem Gitterpunkt einen Grau- bzw. Farbwert zuweist. Somit kann ein Bild also durch eine Matrix $x \in \mathbb{R}^{N_1 \times N_2}$ beschrieben werden. Wir wollen einige Bezeichnungsweisen

eingeführen, die in den folgenden Kapiteln immer dann nützlich sein werden, wenn wir mit dem diskreten Bildmodell arbeiten. Die Gitterpunkte eines zweidimensionalen Bildes sollen sowohl in der Breite als auch in der Länge einen Abstand h zueinander haben, den wir bei Bildern immer gleich 1 setzen. Diese sind also gegeben durch:

$$s_{i,j} := (ih, jh) \quad \text{mit } (i, j) \in \{1, \dots, N_1\} \times \{1, \dots, N_2\}.$$

Für den Wert des Bildes x in einem Gitterpunkt $s_{i,j}$ führen wir folgende Schreibweise ein:

$$x_{i,j} := x(s_{i,j}) \quad \text{mit } (i, j) \in \{1, \dots, N_1\} \times \{1, \dots, N_2\}.$$

Abkürzend schreiben wir auch für das Gitter der Größe $N_1 \times N_2$:

$$\Omega_{N_1, N_2} := \{1, \dots, N_1\} \times \{1, \dots, N_2\}.$$

Für jeden Punkt $s_{i,j}$ im Innern des Gitters definieren wir seine Nachbarschaft durch:

$$\mathcal{N}_{i,j} := \{s_{i,j-1}, s_{i,j+1}, s_{i-1,j}, s_{i+1,j}\}. \quad (2.7)$$

Die Nachbarschaft eines Punktes ist also die Menge der vier direkt anliegenden Punkte im Gitter. Für einen Randpunkt $s_{i,j}$ mit $(i, j) \in \{1\} \times \{2, \dots, N_2 - 1\}$ erklären wir seine Nachbarschaft durch:

$$\mathcal{N}_{i,j} := \{s_{i,j-1}, s_{i,j+1}, s_{i+1,j}\}.$$

Analog definieren wir die Nachbarschaft für alle anderen Randpunkte, wobei die Nachbarschaft eines Eckpunktes dementsprechend nur aus zwei Punkten besteht.

Wir wollen in dieser Arbeit sehr genau zwischen **Gitterpunkten** und **Pixeln** unterscheiden. Für einen gegebenen Gitterpunkt $s_{i,j} \in \Omega_{N_1, N_2}$ definieren wir den zugehörigen Pixel durch:

$$P_{i,j} := [s_{i,j} - \frac{h}{2}, s_{i,j} + \frac{h}{2}] \times [s_{i,j} - \frac{h}{2}, s_{i,j} + \frac{h}{2}].$$

Während der Gitterpunkt wirklich nur ein Punkt ist, verstehen wir in dieser Arbeit unter einem Pixel eine Menge von Punkten, wobei alle Punkte in dem Pixel den gleichen Farbwert besitzen sollen. Wenn wir also von einem verrauschten Pixel sprechen, so ist nicht nur ein Punkt in diesem Pixel verrauscht, sondern der gesamte zugehörige Bereich erhält denselben Farbwert. Statt Gitterpunkt sagen wir oft – wenn dies aus dem Zusammenhang klar ist – auch einfach nur Punkt.

Für einen Gitterpunkt im Innern des Gitters können wir die **Vorwärts-** und **Rückwärtsdifferenzenquotienten** definieren. Wir führen für einen Punkt $s_{i,j}$ im Innern des Gitters die folgenden Bezeichnungen ein:

$$\begin{aligned} \delta_{s_1}^+ x_{i,j} &:= \frac{x_{i+1,j} - x_{i,j}}{h}, & \delta_{s_2}^+ x_{i,j} &:= \frac{x_{i,j+1} - x_{i,j}}{h}, \\ \delta_{s_1}^- x_{i,j} &:= \frac{x_{i,j} - x_{i-1,j}}{h}, & \delta_{s_2}^- x_{i,j} &:= \frac{x_{i,j} - x_{i,j-1}}{h}. \end{aligned}$$

Für den eindimensionalen Fall $\Omega \subset \mathbb{R}$ gelten entsprechende Bezeichnungen, zum Beispiel bezeichne $x_i := x(s_i)$ den Wert im den Gitterpunkt s_i und $\delta_s^+ x_i$ den Vorwärtsdifferenzenquotienten in diesem Punkt.

Wir wollen nun kurz die Güte der Approximation der Ableitungen durch finite Differenzen diskutieren, und betrachten dazu die Funktion $x : \Omega \subset \mathbb{R} \rightarrow \mathbb{R}$. Wenn die Funktion

hinreichend glatt ist, so gilt für kleines h die Approximation $x'(s_i) \approx \delta_s^+ x(s_i)$. Genauer folgt zunächst nach dem Mittelwertsatz:

$$x(s_i + h) - x(s_i) = x'(s_i)h + \frac{1}{2}x''(\xi)h^2, \quad \xi \in [s_i, s_i + h].$$

Und somit erhalten wir die Abschätzung

$$\|\delta_s^+ x - x'\|_\infty = \sup_{s \in \Omega} |\delta_s^+ x(s) - x'(s)| \leq \frac{h}{2} \sup_{\xi \in \Omega} |x''(\xi)| = \frac{h}{2} \|x''\|_\infty.$$

In glatten Regionen, in denen x zweimal stetig differenzierbar ist, ist $\|x''\|_\infty$ unabhängig von h , und wir erhalten eine gute Näherung der Ableitung für $h \rightarrow 0$. Hingegen ist bei einer Kante $x'(s) \approx \frac{1}{h}$ und $x''(s) \approx \frac{1}{h^2}$, und der Fehler, der bei der Finite-Differenzen-Approximation entsteht, kann daher nur noch durch den Wert $> \frac{1}{2}$ (für $h < 1$) abgeschätzt werden, sodass wir den Approximationsfehler nicht hinreichend gut abschätzen können. Wie wir sehen werden, liefert die Diskretisierung in der Praxis jedoch gute Ergebnisse und ist leicht anwendbar.

Ist ein diskretes Bild x gegeben, so lässt sich diesem wiederum ein kontinuierliches Bild $\hat{x} : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ zuordnen, indem man zum Beispiel zwischen den einzelnen Gitterpunkten linear interpoliert oder es als stückweise konstantes Bild auffasst. Der Einfachheit halber betrachten wir ein eindimensionales Bild, das auf N Gitterpunkten $s_1 < \dots < s_N$ gegeben ist, wobei die Gitterpunkte jeweils den äquidistanten Abstand h haben sollen. Wenn wir zwischen den Gitterpunkten linear interpolieren, erhalten wir das kontinuierliche Bild

$$\hat{x}(s) = x_i + \frac{x_{i+1} - x_i}{h}(s - s_i) \quad \text{für } s \in [s_i, s_{i+1}[, \quad i \in 1, \dots, N - 1.$$

Ein stückweise konstantes Bild lässt sich bei gegebenen Konstanten c_n mit der charakteristischen Funktion χ schreiben als

$$\hat{x}(s) = \sum_{n=1}^N c_n \chi_{[s_n - \frac{h}{2}, s_n + \frac{h}{2}[}(s). \quad (2.8)$$

Dabei nimmt χ auf dem (eindimensionalen) Pixel $P_n := [s_n - \frac{h}{2}, s_n + \frac{h}{2}[$ den Wert 1 an und sonst 0. Informell können wir sagen, dass wir jedem Pixel P_n den Farbwert des zugehörigen Gitterpunktes s_n zuordnen.

Nun bestimmen wir den Gesamtfehler eines Bildes, das durch das additive Gauss'sche Rauschen gestört wurde. Wir betrachten dazu das Fehlermodell (2.2) für ein diskretes gestörtes Bild I und das ungestörte Bild \tilde{x} mit $\tilde{x}, I \in \mathbb{R}^{n \times n}$. Das Fehlermodell hat im diskreten Fall dann die Gestalt

$$I_{i,j} = S\tilde{x}_{i,j} + N_{i,j}, \quad \forall (i,j) \in \Omega_{n,n}, \quad (2.9)$$

wobei S eine $n \times n$ -Matrix und die $N_{i,j}$ unabhängige normalverteilte Zufallsvariable mit Varianz σ^2 und Erwartungswert 0 sind. Für den normierten quadratischen Gesamtfehler gilt

$$\frac{1}{n^2} \sum_{i,j=1}^n (N_{i,j})^2 = \frac{1}{n^2} \sum_{i,j=1}^n \mathbb{E}((N_{i,j})^2) = \frac{1}{n^2} \sum_{i,j=1}^n \mathbb{E}((N_{i,j} - \mathbb{E}(N_{i,j}))^2) \quad (2.10)$$

$$= \frac{1}{n^2} \sum_{i,j=1}^n \mathbb{V}(N_{i,j}) = \sigma^2. \quad (2.11)$$

Weiterhin erhalten wir durch Umstellung von (2.9):

$$I_{i,j} - S\tilde{x}_{i,j} = N_{i,j}, \quad \forall (i,j) \in \Omega_{n,n},$$

und somit gilt

$$\frac{1}{n^2} \sum_{i,j=1}^n (I_{i,j} - S\tilde{x}_{i,j})^2 = \frac{1}{n^2} \sum_{i,j=1}^n (N_{i,j})^2 = \sigma^2.$$

Für $n \rightarrow \infty$ erhalten wir den Übergang zum kontinuierlichen Bild, und es folgt:

$$\int_{\Omega} (I(s) - S(\tilde{x}(s)))^2 ds = \int_{\Omega} (N(s))^2 ds = \sigma^2. \quad (2.12)$$

2.4.2 Kontinuierliches Bildmodell - geeignete Funktionenräume

In diesem Abschnitt soll die Frage diskutiert werden, welche mathematischen Funktionenräume für ein kontinuierliches Bild $x : \Omega \rightarrow \mathbb{R}$ im Zusammenhang mit unserer Problemstellung aus Abschnitt 2.1 geeignet sind. Wir unterscheiden dabei in der Regel zwischen einem Funktionenraum \mathcal{X} , der unsere beobachteten Bilder enthält, und einem Funktionenraum \mathcal{Y} , in dem wir unser gewünschtes Bild suchen. Unsere beobachteten Bilder sind in der Regel verrauscht. Besitzt das Rauschen die Varianz σ^2 und Mittelwert 0, so gilt (vgl. (2.12))

$$\int_{\Omega} (N(s))^2 ds = \sigma^2.$$

Somit ist zumindest $N \in L^2(\Omega)$. Da das ungestörte Bild in der Regel viel regulärer ist und nicht so große Oszillationen aufweist wie das Gauss'sche Rauschen N , folgern wir insgesamt, dass sich das beobachtete Bild $I : \Omega \rightarrow \mathbb{R}$ in $L^2(\Omega)$ befindet. Der Bildraum \mathcal{Y} soll in der Regel Bilder enthalten, die nicht so stark oszillieren. Also ist es sinnvoll, mehr Regularität zu fordern. Dies erreichen wir, indem wir verlangen, dass die schwache Ableitung unseres Bildes beschränkt ist, d.h. wir verlangen, dass das gesuchte Bild im Sobolevraum $W^{1,p}(\Omega)$, $p \geq 1$, enthalten ist. Dies deckt sich gewissermaßen auch mit der intuitiven Definition von Glattheit: Ein Bild ist glatt, wenn der Gradient nicht zu stark oszilliert und somit zumindest beschränkt ist. Diese Tatsache werden wir auch bei der Modellierung der Methoden zur Lösung unseres Problems verwenden. Bilder in Sobolevräumen zu vermuten ist aus Sicht der Bildverarbeitung jedoch nicht ideal, da man in diesen Räumen keine Bilder mit Kanten modellieren kann. Dass stückweise stetige Funktionen nicht im Sobolevraum $W^{1,2}(\Omega)$ enthalten sind, wollen wir kurz darlegen. Wir betrachten dazu die Heavisidefunktion auf dem Intervall $[0, 1]$, die im eindimensionalen offensichtlich eine Kante aufweist:

$$H(s) = \begin{cases} 0 & | s \in [0, \frac{1}{2}), \\ 1 & | s \in [\frac{1}{2}, 1]. \end{cases}$$

Wir bestimmen die schwache Ableitung von H . Für die Testfunktionen $\varphi \in C^\infty([0, 1])$ mit $\varphi(0) = \varphi(1) = 0$ gilt mittels partieller Integration:

$$\begin{aligned} \int_0^1 H'(s)\varphi(s)ds &= H(s)\varphi(s)|_0^1 - \int_0^1 H(s)\varphi'(s)ds \\ &= - \int_0^1 H(s)\varphi'(s)ds = - \int_{\frac{1}{2}}^1 \varphi'(s)ds = \varphi(\frac{1}{2}). \end{aligned}$$

Die schwache Ableitung der Heavisidefunktion ist also die Dirac-Delta-Distribution $\delta : \varphi \mapsto \varphi(\frac{1}{2})$. Diese liegt jedoch nicht in $L^2([0, 1])$. Dies erkennen wir, indem wir eine Folge von Testfunktionen φ_n betrachten, für die gilt:

$$\int_{\Omega} \varphi_n(s) ds = 1, \quad \varphi_n(\frac{1}{2}) \rightarrow \infty \text{ für } n \rightarrow \infty.$$

Nach dem Rieszschen Darstellungssatz [4], pp. 76, ist jedes stetige lineare Funktional auf $L^2([0, 1])$ von der Gestalt

$$\varphi \mapsto \int_{\Omega} \delta(s) \varphi_n(s) ds.$$

Nun ist jedoch mit obiger Folge φ_n für $n \rightarrow \infty$

$$\int_{\Omega} \delta(s) \varphi_n(s) ds \rightarrow \infty,$$

und somit $\delta \notin L^2([0, 1])$. Daher ist die Heavisidefunktion nicht im Sobolevraum $W^{1,2}([0, 1])$ enthalten. Es kann sogar noch mehr gezeigt werden. Wir zitieren das folgende Lemma aus [9], pp. 40 und verzichten auf den Beweis:

Lemma 2.1:

Sei $D \subset \Omega$ ein Gebiet mit C^1 -Rand. Dann ist die Funktion

$$x(s) = \begin{cases} 1 & | \text{ wenn } s \in D, \\ 0 & | \text{ sonst.} \end{cases}$$

nicht in $W^{1,p}(\Omega)$ für $p \geq 1$.

Dies bedeutet nun, dass selbst der größte Sobolevraum $W^{1,1}(\Omega)$ keine stückweise konstanten Funktionen, also insbesondere keine Kanten enthält. Dies ist jedoch der Fall im Raum $BV(\Omega)$ (vgl. (2.6) und [9], pp. 50). Aus diesem Grund und weil Existenzsätze einfacher zu formulieren sind als für Lebesgue-Räume, findet er in der Bildverarbeitung oft Verwendung.

2.4.3 Übergänge zwischen den Modellen

Wie bereits angemerkt wurde, kann ein diskretes Bild etwa durch Interpolation in ein kontinuierliches Bild übergehen. Umgekehrt erhalten wir aus einem kontinuierlichen Bild ein diskretes Bild, wenn wir jedem Gitterpunkt jeweils einen Mittelwert aus dem zugehörigen Bereich des kontinuierlichen Bildes zuordnen. Der Wechsel zwischen den unterschiedlichen Bild-Repräsentationen in dieser Arbeit kann durch das folgende vierstufige Schema zusammengefasst werden:

Wir wollen dieses Schema kurz erläutern. In Kapitel 4 werden wir unsere Methoden zur Problemlösung modellieren und die Bilder meistens in Sobolevräumen aufsuchen. Anschließend werden wir in Kapitel 5 diese Modelle diskretisieren und dabei diskrete Bilder erhalten. Wir lösen das erhaltene Problem in Kapitel 6, und erhalten wiederum ein diskretes Bild als Lösung. Meistens werden diese Bilder nun bewertet. Es besteht jedoch noch die Möglichkeit, das Bild zunächst noch linear zu interpolieren und anschließend zu bewerten, was in der gängigen Literatur meistens nicht angemerkt wird. Meistens werden wir jedoch die zugehörigen kontinuierlichen Bilder als stückweise konstant auffassen, also eine Definition wie in (2.8) verwenden.

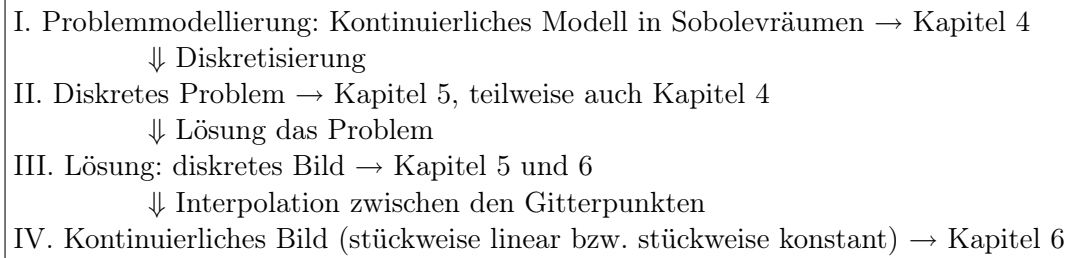


Abbildung 2.1: Überblick über unterschiedliche Bildrepräsentationen in dieser Arbeit

Die Frage, inwiefern die kontinuierlichen Bilder bzw. die diskreten Bilder, die wir letztendlich mit unseren Verfahren erhalten, mit unseren Bildern aus der Problemmodellierung zusammenhängen, erweist sich aus mehreren Gründen als schwierig. Lösen wir das image-restoration-Problem etwa mit Hilfe von partiellen Differentialgleichungen, so könnten wir für gewisse Fälle mit Konsistenz- und Stabilitätsaussagen eine Konvergenz der diskreten Bilder gegen das gesuchte Bild im Sobolevraum beweisen, etwa nach dem Vorbild [3], pp. 233 ff. Da wir in dieser Arbeit das Problem oft als Steuerungsproblem modellieren, können entsprechende Aussagen nicht behandelt werden, da die zugehörigen Konvergenzbeweise noch nicht ausgearbeitet wurden.

3 Variationsprobleme und Probleme der Optimalen Steuerung – notwendige Optimalitätsbedingungen

In der Bildverarbeitung ist man oft bestrebt, bestimmte Aufgaben als Probleme der Minimierung von Energiefunktionalen zu modellieren. Das theoretische Fundament für solche Aufgaben liefern die Variationsrechnung und die Theorie der optimalen Steuerung. In diesem Kapitel stellen wir die Grundlagen dieser Theorie zusammen und formulieren insbesondere notwendige Optimalitätsbedingungen, die eine Minimalstelle der entsprechenden Energiefunktionale erfüllen muss.

3.1 Variationsprobleme

Wir betrachten Funktionen $x : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, wobei wir annehmen, dass x im Sobolevraum $W^{1,p}(\Omega)$ enthalten ist. In der Variationsrechnung werden dann Energiefunktionale $F : W^{1,p}(\Omega) \rightarrow \mathbb{R}$ der folgenden Gestalt betrachtet:

$$F(x) = \int_{\Omega} f(s, x(s), Jx(s)) ds, \tag{3.1}$$

wobei

$$Jx(s) = \begin{pmatrix} \frac{\partial x_1}{\partial s_1}(s) & \dots & \frac{\partial x_1}{\partial s_n}(s) \\ \vdots & & \vdots \\ \frac{\partial x_m}{\partial s_1}(s) & \dots & \frac{\partial x_m}{\partial s_n}(s) \end{pmatrix}$$

die Jacobi-Matrix von x im Punkt s bezeichne. Bei einer gegebenen Funktion f lautet nun das Ziel, die minimale Energie zu bestimmen, also eine Funktion x zu finden, sodass das Funktional F minimal wird. Dabei stellen sich die Fragen nach der Eindeutigkeit und Existenz einer solchen Minimalstelle x . Wir werden dabei meistens fordern, dass x außerdem die Dirichletrandbedingung

$$x = y \text{ auf } \partial\Omega \text{ für gegebenes } y \in W^{1,p}(\Omega, \mathbb{R}^n) \tag{3.2}$$

erfüllt. Wenn wir mit $W_0^{1,p}(\Omega, \mathbb{R}^n)$ den Raum der Funktionen aus $W^{1,p}(\Omega, \mathbb{R}^n)$ bezeichnen, die auf dem Rand $\partial\Omega$ den Wert 0 annehmen, ist die Schreibweise $x - y \in W_0^{1,p}(\Omega, \mathbb{R}^n)$ gleichbedeutend mit (3.2). Analog dazu schreiben wir auch $x \in y + W_0^{1,p}(\Omega, \mathbb{R}^n)$, und damit ist die Randbedingung also implizit im Lösungsraum enthalten.

Wir klären zunächst die Frage, wann eine Lösung existiert, und geben anschließend notwendige Optimalitätsbedingungen an.

3.1.1 Existenzsätze

In der Variationsrechnung unterscheidet man den klassischen Ansatz und die direkte Methode. Mit Hilfe des klassischen Ansatzes werden wir im nächsten Abschnitt Bedingungen herleiten, die eine Minimalstelle erfüllen muss. Unter der direkten Methode der Variationsrechnung versteht man, für das Problem

$$\text{Minimiere } F(x) \tag{3.3}$$

eine Minimalfolge x_k zu finden, die gegen die Minimalstelle \hat{x} konvergiert. Nach dem Satz von Weierstrass nimmt eine stetige Funktion auf einer kompakten Menge Maximal- und Minimalstelle an. Die Stetigkeit für F zu fordern, wäre in einigen Fällen zu einschränkend. Stattdessen fordern wir allgemeiner, dass F folgenuntermittig ist, und erhalten den folgenden Satz, der die Existenz einer Minimalstelle garantiert:

Satz 3.1 (Fundamentaler Existenzsatz der Optimierung):

Sei \mathcal{X} ein topologischer Raum mit Topologie τ und $F : (\mathcal{X}, \tau) \rightarrow \mathbb{R} \cup \infty$ ein nach unten beschränktes Funktional. Es gelte

1. Folgenuntermittigkeit: für jede Folge $x_k \rightarrow x$ in der Topologie τ gelte

$$F(x) \leq \liminf_k F(x_k).$$

2. Kompaktheit: es existiert ein $\alpha \in \mathbb{R}$, sodass die Niveaumenge

$$S_\alpha := \{x \in \mathcal{X} \mid F(x) \leq \alpha\}$$

nichtleer und folgenkompakt in der Topologie τ ist.

Dann existiert eine Minimalstelle \hat{x} von (3.3).

Beweis. Es sei x_n eine Minimalfolge in \mathcal{X} , d.h. $\lim_{n \rightarrow \infty} F(x_n) = \inf_{x \in \mathcal{X}} F(x) > -\infty$. Sei α so gegeben, dass S_α nichtleer ist. Dann existiert $N \in \mathbb{N}$, sodass für alle $n \geq N : F(x_n) \leq \alpha$. Somit ist x_n für $n \geq N$ in der folgenkompakten Menge S_α und hat daher eine konvergente Teilfolge $x_{n_k} \rightarrow \hat{x}$. Nun folgt aus der Folgenuntermittigkeit von F :

$$F(\hat{x}) \leq \liminf_{k \rightarrow \infty} F(x_{n_k}) = \lim_{k \rightarrow \infty} F(x_{n_k}) = \inf_{x \in \mathcal{X}} F(x) \leq F(\hat{x})$$

und damit $\inf_{x \in \mathcal{X}} F(x) = F(\hat{x})$. □

Ist das Funktional F von der Gestalt (3.1), so lässt sich ein Existenzsatz angeben, der einfacher zu verifizieren ist und für unsere Zwecke ausreicht. In diesem Fall lässt sich unter geeigneten Voraussetzungen die Folgenuntermittigkeit aus der Konvexität von f folgern. Die Beschränktheit der Folge erhalten wir aus der Koerzivität des Funktionals F . Dabei heißt F koerziv, wenn $\lim F(x) = \infty$ für $|x| \rightarrow \infty$. Wir präzisieren diese Überlegungen, indem wir den folgenden Satz aus [17], pp. 84, zitieren:

Satz 3.2 (Existenzsatz für Funktionale mit konvexem Integranden):

Sei $\Omega \subset \mathbb{R}^n$ offen, beschränkt und habe einen Lipschitzrand. Es sei der Integrand $f \in C^0(\bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n)$, $f = f(s, \omega, \xi)$. Weiterhin gelten die folgenden beiden Bedingungen:

- (i) $\xi \rightarrow f(s, \omega, \xi)$ sei konvex für jedes $(s, \omega) \in \bar{\Omega} \times \mathbb{R}$,

(ii) $\exists a \in L^1(\Omega), b > 0, c > 0, p > 1$, sodass für jedes $(s, \omega, \xi) \in \bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n$ gilt:

$$f(s, \omega, \xi) \geq a(s) + b|\omega|^p + c|\xi|^p.$$

Dann hat das Problem

$$\text{Minimiere } \left\{ F(x) = \int_{\Omega} f(s, x(s), Jx(s)) ds \mid x \in y + W_0^{1,p}(\Omega) \right\}$$

eine Lösung. Ist zudem $(\omega, \xi) \rightarrow f(s, \omega, \xi)$ strikt konvex, so ist die Lösung eindeutig.

3.1.2 Notwendige Optimalitätsbedingungen

Der klassische Ansatz der Variationsrechnung beruht auf der Aufstellung notwendiger Optimalitätsbedingungen, die wir nun formulieren. Dazu ist es zunächst notwendig, den Differenzierbarkeitsbegriff in Banachräumen zu erklären. Fundamental dafür ist der Begriff der Gâteaux-Ableitung:

Definition 3.1 (Gâteaux-Ableitung):

Seien \mathcal{X} und \mathcal{Y} Banachräume. und $F : \mathcal{X} \rightarrow \mathcal{Y}$. Wir definieren die Richtungsableitung von F in einem Punkt x in Richtung y durch

$$F'(x, y) := \lim_{t \rightarrow 0, t > 0} \frac{F(x + ty) - F(x)}{t},$$

falls der Grenzwert auf der rechten Seite existiert. F heiÙe Gâteaux-differenzierbar in x , wenn der Grenzwert für alle $y \in \mathcal{X}$ existiert und zudem ein linearer stetiger Operator $A : \mathcal{X} \rightarrow \mathcal{Y}$ existiert, sodass

$$F'(x, y) = Ay \quad \forall x \in \mathcal{X}.$$

A heißt auch Gâteaux-Ableitung in x , und wir bezeichnen sie mit $F'(x)$.

Die Gâteaux-Ableitung ist also ein lineares stetiges Funktional auf \mathcal{X} , also ein Element des Dualraums \mathcal{X}^* . Wenn \mathcal{X} ein Hilbertraum ist, kann nach dem Satz von Riesz (vgl. [4], pp. 76) jedes Element aus \mathcal{X}^* mit einem Element aus dem Dualraum \mathcal{X} identifiziert werden. Den Ausdruck aus \mathcal{X} , den wir in einem solchen Fall mit der Gâteaux-Ableitung identifizieren, nennen wir Gradient von F . Wir verdeutlichen dies an einem Beispiel.

Beispiel 1. Sei H ein Hilbertraum mit zugehörigem Skalarprodukt $\langle \cdot, \cdot \rangle$ und $f : H \rightarrow H$ eine Funktion, definiert durch

$$f(x) = \langle x, x \rangle$$

Für die Gâteaux-Ableitung von f gilt

$$\begin{aligned} f'(x)y &= \lim_{t \rightarrow 0, t > 0} \frac{f(x + ty) - f(x)}{t} \\ &= \lim_{t \rightarrow 0, t > 0} \frac{\langle x + ty, x + ty \rangle - \langle x, x \rangle}{t} \\ &= \lim_{t \rightarrow 0, t > 0} \frac{2t \langle x, y \rangle + t^2 \langle x, x \rangle}{t} \\ &= \langle 2x, y \rangle. \end{aligned}$$

Die Ableitung ist also in diesem Fall durch die Vorschrift $f'(x)y = 2 \langle x, y \rangle$ gegeben, während wir mit dem Ausdruck $f'(x) = 2x$ den Gradienten identifizieren.

Der folgende Satz ist nun fundamental für die Herleitung notwendiger Optimalitätsbedingungen für das Standardfunktional der Variationsrechnung (3.1):

Satz 3.3 (Notwendige Optimalitätsbedingung):

$F : \mathcal{X} \rightarrow \mathbb{R}$ sei Gâteaux-differenzierbar und \bar{x} sei eine globale Minimalstelle von F , d.h.

$$F(\bar{x}) = \inf \{F(x) | x \in \mathcal{X}\}.$$

Dann gilt

$$F'(\bar{x})y = 0 \quad \forall y \in \mathcal{X}. \tag{3.4}$$

Beweis. Sei \bar{x} die Minimalstelle von F . Für $y \in \mathcal{X}, t \in \mathbb{R}$ definieren wir die reelle Funktion $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ durch

$$\varphi(t) := F(\bar{x} + ty).$$

Somit nimmt φ in $t = 0$ sein Minimum an, und da φ differenzierbar ist, gilt für jedes $y \in \mathcal{X}$

$$\begin{aligned} 0 = \varphi'(0) &= \frac{d}{dt} F(\bar{x} + ty)|_{t=0} \\ &= \lim_{h \rightarrow 0} \frac{F(\bar{x} + (t+h)y) - F(\bar{x} + ty)}{h} \Big|_{t=0} \\ &= \lim_{h \rightarrow 0} \frac{F(\bar{x} + hy) - F(\bar{x})}{h} \\ &= F'(\bar{x})y. \end{aligned}$$

□

Gilt die Beziehung (3.4) für ein \bar{x} , so ist dies noch nicht hinreichend für eine Minimalstelle, jedoch gilt der folgende Satz:

Satz 3.4 (Hinreichende Optimalitätsbedingung):

Sei F Gâteaux-differenzierbar und zusätzlich konvex. Für $\bar{x} \in \mathcal{X}$ gelte

$$F'(\bar{x})y = 0 \quad \forall y \in \mathcal{X}.$$

Dann ist \bar{x} eine Minimalstelle von F .

Beweis. Der Beweis findet sich in [16], pp. 51. □

Wir geben nun noch die Optimalitätsbedingungen für Funktionale F der Gestalt (3.1) an. Wir fordern dann, dass die gesuchte Minimalstelle x zweimal stetig differenzierbar ist. Dann erhalten wir Optimalitätsbedingungen für eine globale Minimalstelle, die sich leicht auf unsere Probleme anwenden lassen. Die folgenden Optimalitätsbedingungen lassen sich aus Satz 3.3 herleiten und werden auch als Euler-Lagrange-Gleichungen bezeichnet:

Satz 3.5 (Euler-Lagrange-Gleichung):

Sei $m, n \geq 1$, $\Omega \in \mathbb{R}^n$ ein Lipschitzgebiet. Weiterhin sei der Integrand $f(s, x, \xi)$ mit $f : \Omega \times \mathbb{R}^m \times \mathbb{R}^{nm} \rightarrow \mathbb{R}$ eine C^2 -Funktion. Ist $x^* \in C^2(\Omega, \mathbb{R}^m)$ eine Minimalstelle von

$$F(x) = \int_{\Omega} f(s, x(s), Jx(s)) ds \quad \text{mit } x \in x_0 + W_0^{1,p}(\Omega, \mathbb{R}^m),$$

dann gilt für jedes $s \in \Omega$ und $1 \leq i \leq m$ die notwendige Bedingung

$$F'(x^*) = -\operatorname{div} \left(\frac{\partial f}{\partial \xi^i}(s, x^*, Jx^*) \right) + \frac{\partial}{\partial x_i} f(s, x^*, Jx^*) = 0.$$

Dabei bezeichne div den Divergenz-Operator, der für eine Funktion $x : \mathbb{R}^n \rightarrow \mathbb{R}$ definiert ist als:

$$\operatorname{div} x = \sum_{i=1}^n \frac{\partial x}{\partial s_i}.$$

Beweis. Ein Beweis dieses Satzes findet sich in [16], pp.87. □

Im Fall $n = 1$ sind die Euler-Lagrange-Gleichungen ein System von m gewöhnlichen Differentialgleichungen. Ist $m = 1$, so ist die Euler-Lagrange-Gleichung eine partielle Differentialgleichung.

3.2 Probleme der optimalen Steuerung

Um Probleme der optimalen Steuerung zu erhalten, nehmen wir zunächst zu dem Variationsproblem

$$\text{Minimiere } \left\{ F(x) = \int_{\Omega} f(s, x(s), Jx(s)) ds \quad \text{mit } x \in W_0^{1,p}(\Omega, \mathbb{R}^m), \Omega \subset \mathbb{R}^n \right\} \quad (3.5)$$

Beschränkungen der partiellen Ableitungen von x hinzu:

$$Jx(s) \in K \subset \mathbb{R}^{nm} \quad (\forall) s \in \Omega, K \text{ konvex.}$$

Für eine Menge $M \subset \mathbb{R}^n$ bedeute die Bezeichnung $(\forall) s \in M$, dass eine Aussage „für fast alle $s \in M$ “ gilt, d.h. wenn es eine Nullmenge N gibt, sodass die Aussage für alle $s \in M \setminus N$ gilt.

Durch Einführung einer Steuerung u erhalten wir aus diesem Variationsproblem mit Nebenbedingung ein Problem der optimalen Steuerung:

$$(P) \quad \text{Minimiere } F(x, u) = \int_{\Omega} f(s, x(s), u(s)) ds \quad (3.6)$$

$$\text{unter } u(s) = Jx(s), \quad (\forall) s \in \Omega, \quad (3.7)$$

$$u(s) \in K \subset \mathbb{R}^{nm}, \quad (\forall) s \in \Omega. \quad (3.8)$$

Dabei suchen wir wie in dem Variationsproblem (3.5) den Zustand x in $W_0^{1,p}(\Omega, \mathbb{R}^m)$ auf, und erhalten so auf natürliche Weise für die Steuerung den Funktionenraum $L^p(\Omega, \mathbb{R}^{nm})$. Wir definieren die Menge der zulässigen Steuerungen (oder auch Steuerbereich genannt) durch

$$U = \{u \in L^p(\Omega, \mathbb{R}^{nm}) \mid u(s) \in K \subset \mathbb{R}^{nm} \quad (\forall) s \in \Omega\}.$$

Die Nebenbedingung $u \in U$ nennen wir dann Steuerbeschränkung, während wir Nebenbedingungen der Gestalt

$$u(s) = Jx(s), \quad (\forall) s \in \Omega$$

Zustandsgleichungen nennen wollen. Die Menge aller Paare $(x, u) \in W_0^{1,p}(\Omega, \mathbb{R}^m) \times L^p(\Omega, \mathbb{R}^{nm})$, die die Nebenbedingungen (3.7) und (3.8) erfüllen, nennen wir **zulässige Menge**. Jedes solche Paar heißt **zulässig**. Es folgt dann aus der Steuerbeschränkung, dass $x \in W_0^{1,p}(\Omega, \mathbb{R}^m) \cap$

$W_0^{1,\infty}(\Omega, \mathbb{R}^m)$ ist. Im Fall $1 < p \leq n$ besitzt x einen Lipschitz-stetigen Repräsentanten. Die Aufgabe der optimalen Steuerung besteht darin, ein zulässiges Paar (x^*, u^*) zu finden, für welches gilt:

$$F(x^*, u^*) \leq F(x, u) \text{ für alle zulässigen Paare } (x, u)$$

Ein mehrdimensionales Steuerungsproblem der Gestalt (P) ist ein Problem vom Dieudonné-Rashevsky-Typ. Steuerungsprobleme dieser Gestalt finden bei Torsionsproblemen für prismatische Stäbe mit elastischem Materialverhalten Anwendung. In [31], pp. 108 ff. wird vorgeschlagen, Probleme der Bildverarbeitung als Steuerungsprobleme dieser Art aufzufassen. So ergibt sich dies in natürlicher Weise für das in der Bildverarbeitung bekannte shape-from-shading-Problem oder das Problem des optischen Flusses. In dieser Arbeit werden wir das image-restoration-Problem als Steuerungsproblem der Gestalt (P) modellieren.

Bemerkung 3.1. Mehrdimensionale Steuerungsprobleme der Gestalt (P) wurden bisher numerisch kaum behandelt. Die Arbeit von Dewess/Helbig [18] kann dabei als Vorläufer angesehen werden, wobei in dieser Arbeit nicht das Problem (P) unmittelbar numerisch gelöst wurde, sondern das duale Problem.

Im Folgenden werden notwendige Optimalitätsbedingungen formuliert. Dabei hängt das weitere Vorgehen davon ab, ob der Integrand f konvex oder nichtkonvex ist. Wir werden uns auf den konvexen Fall beschränken.

3.2.1 Das Pontrjaginsche Maximumprinzip für mehrdimensionale Steuerungsprobleme

Es gelte $n \geq 1, m \geq 2, 1 < p < \infty$ und K sei ein konvexer Körper mit $0 \in \text{int}(K)$. Weiterhin sei Ω Abschluß eines beschränkten Lipschitzgebietes. Der Integrand $f(s, \omega, \xi)$ mit $f: \Omega \times \mathbb{R}^n \times \mathbb{R}^{mn} \rightarrow \mathbb{R}$ sei in s stetig und nach allen x_i und ξ_{ij} stetig differenzierbar. Wenn wir zudem nur konvexe Integranden betrachten, so können wir die folgenden notwendigen Optimalitätsbedingungen für das Dieudonné-Rashevsky-Problem der Gestalt (P) anwenden, die in der folgenden Form erstmals von Wagner [31], pp. 92 ff. formuliert wurden:

Satz 3.6 (Pontrjaginsches Maximumprinzip für das Dieudonné-Rashevsky-Problem mit konvexem Integranden):

Das Dieudonné-Rashevsky-Problem (P) habe eine globale Minimalstelle (x^, u^*) und es gelten die obigen Voraussetzungen. Der Integrand $f(s, \omega, \xi)$ sei konvex bezüglich ξ für alle $(s, \omega) \in \Omega \times \mathbb{R}^n$. Dann existieren Multiplikatoren $\lambda_0 > 0$ und $y \in L^q(\Omega, \mathbb{R}^{nm})$, wobei q die zu p Konjugierte sei, d.h. $\frac{1}{p} + \frac{1}{q} = 1$, die nicht gleichzeitig verschwinden und die folgenden beiden Bedingungen erfüllen:*

$$(M) : \quad \lambda_0 \int_{\Omega} (f(s, x^*(s), u(s)) - f(s, x^*(s), u^*(s))) ds \\ - \sum_{i,j} \int_{\Omega} (u_{ij}(s) - u_{ij}^*(s)) y_{ij}(s) ds \geq 0 \quad \forall u \in U,$$

$$(K) : \quad \lambda_0 \sum_i \int_{\Omega} \frac{\partial f}{\partial \xi_i}(s, x^*(s), u^*(s)) \cdot (x_i(s) - x_i^*(s)) ds \\ + \sum_{i,j} \int_{\Omega} \left(\frac{\partial x_i}{\partial s_j}(s) - \frac{\partial x_i^*}{\partial s_j}(s) \right) y_{ij}(s) ds = 0 \quad \forall x \in W_0^{1,p}(\Omega, \mathbb{R}^n).$$

Aus der Bedingung (M) kann eine punktweise Formulierung hergeleitet werden. Es gilt

$$(MP) : \quad \lambda_0 (f(s, x^*(s), \xi) - f(s, x^*(s), u^*(s))) \\ - \sum_{i,j} (\xi_{ij} - u_{ij}^*(s)) y_{ij}(s) ds \geq 0 \quad \forall \xi \in K \quad (\forall) s \in \Omega.$$

Beweis. Der Beweis dieses Satzes findet sich in [31], pp. 93 ff. □

Die punktweise Formulierung des Maximumsprinzips (MP) eignet sich zur Auswertung, wenn ein Problem der Gestalt (P) vorliegt.

Wir schließen nun noch einige Bemerkungen zu den notwendigen Bedingungen an.

Bemerkung 3.2. Wie eine Neuformulierung dieses Satzes in [32], pp. 6 zeigt, ist es nicht nötig, die Stetigkeit des Integranden f bezüglich s zu fordern. Diese Bedingung kann abgeschwächt werden, indem wir nur fordern, dass f messbar und wesentlich beschränkt bezüglich s ist. Der Beweis bleibt unverändert.

Bemerkung 3.3. Die Frage nach der Existenz einer Lösung für Probleme vom Dieudonné-Rashevsky-Typ wird in [24], pp. 217 ff. beantwortet. Dabei muss für nichtkonvexe Integranden das relaxierte Problem betrachtet werden, das die nötigen Konvexitätseigenschaften erfüllt. Unter gewissen Voraussetzungen kann dann die Existenz einer globalen Minimalstelle bewiesen werden.

4 Variationsmethoden in der Bildverarbeitung

4.1 Der Variationsansatz zur Lösung des image-restoration-Problems

4.1.1 Herleitung des Variationsansatzes

Variationsmethoden zur Bildwiederherstellung beruhen auf der Minimierung von Energiefunktionalen. Dieser Ansatz besitzt den Vorteil leicht verständlicher und intuitiver Problemmodellierung. Weiterhin lassen sich entwickelte Modelle leicht auf neue Probleme erweitern.

Wir gehen nun von dem Fehlermodell (2.2) aus, d.h. unser beobachtetes Bild sei verzerrt und mit additivem Gauss'schem Rauschen gestört. Nun wollen wir daraus einen Variationsansatz zur Bildwiederherstellung herleiten. Sei also I das beobachtete Bild, \tilde{x} das gesuchte Originalbild, S eine gegebene Transformation und N ein stochastisches Rauschen mit Varianz σ^2 und Erwartungswert Null. Im Folgenden nehmen wir weiterhin an, dass $I, \tilde{x} \in W^{1,2}(\Omega)$ und S ein linearer Operator auf $L^2(\Omega)$ ist. Nach (2.12) gilt dann die Gleichung

$$\int_{\Omega} (I(s) - S(\tilde{x}(s)))^2 ds = \sigma^2.$$

Wir betrachten diese Gleichung als Nebenbedingung, die erfüllt sein soll, und erhalten nun mit einem noch näher zu spezifizierenden Energiefunktional \mathcal{F} das folgende Modell, dessen Lösung das wiederhergestellte Bild ist:

$$\begin{aligned} \min \mathcal{F}(x), \\ \text{unter } \|I - S(x)\|^2 = \sigma^2. \end{aligned} \tag{4.1}$$

Dabei haben wir der Übersichtlichkeit halber das Argument s weggelassen und lassen es auch im Folgenden meistens weg. Das Funktional \mathcal{F} wählen wir so, dass unerwünschte Zustände des Bildes stark bestraft werden. So können wir zum Beispiel zum Glätten von Strukturen im Bild ihre Fläche minimieren, also \mathcal{F} so wählen, dass glatte Strukturen bevorzugt werden. Die Nebenbedingung stellt dabei sicher, dass die Differenz des beobachteten Bildes I und des gesuchten Bildes x die gleiche Varianz haben, wie das Rauschen N . Für das image-restoration-Problem bedeutet dies zum Beispiel, dass das gesuchte Bild nicht zu sehr vom beobachteten Bild abweicht. Dabei sei angemerkt, dass hier a priori-Informationen über die Varianz des Rauschens verwendet werden. Da diese in der Praxis im Voraus nicht bekannt ist, wird in der Literatur meistens die Minimierung der zu (4.1) gehörigen Lagrange-Funktion betrachtet, also in diesem Fall das Minimierungsproblem

$$\text{Minimiere } \left\{ \mathcal{F}(x) + \lambda \|I - S(x)\|^2 \mid x \in W_0^{1,2}(\Omega) \right\}, \tag{4.2}$$

wobei der Lagrange-Multiplikator $\lambda \geq 0$ als Bestrafungsparameter zu interpretieren ist. Ist λ nämlich groß, so wird der Wert $\|I - S(x)\|^2$ stark gewichtet, und folglich werden große Abweichungen vom beobachteten Bild stärker bestraft, als bei kleinem λ . Der Parameter λ kontrolliert also die Gewichtung zwischen Datentreue einerseits und glatten Bildern andererseits, wobei die Glattheit des Bildes durch \mathcal{F} bestimmt wird. In der Praxis wählen wir λ a priori und versuchen durch Testen ein „optimales“ λ zu finden. Optimal ist hier nun so zu verstehen, dass das durch die Minimierung erhaltene Bild \hat{x} die gewünschten Eigenschaften erhält, die wir von ihm erwarten. Dass das Problem mit Nebenbedingung (4.1) und das Problem ohne Nebenbedingungen (4.2) auf die gleiche Lösung führen, wurde zum Beispiel für ein spezielles Funktional \mathcal{F} in [12] untersucht.

Die Regularisierung \mathcal{F}

Entscheidend für die Güte des Bildes, das man durch die Minimierung von (4.2) erhält, ist vor allem das Energiefunktional \mathcal{F} , das wir **Regularisierungsterm** (oder **Regularisierung**) nennen. Da der Betrag des Gradienten in einem Punkt ein Maß für die Steigung in diesem Punkt ist, bedeutet dies, dass ein verrauschtes Bild viele Punkte hat, die einen großen Betragsgradienten aufweisen. Um also das Rauschen zu eliminieren, liegt es nahe, große Betragsgradienten zu bestrafen und das Bild somit zu glätten. Dies motiviert die Abhängigkeit des Funktionals \mathcal{F} von $|\nabla x|$. Da es sinnvoll ist, nicht nur lokale, sondern auch globale Unterschiede des Betragsgradienten zu betrachten, untersuchen wir das Integral über diese, also Funktionale der Gestalt

$$\mathcal{F}(x) = \int_{\Omega} f(|\nabla x|) ds.$$

Die Eigenschaften der sogenannten **Potentialfunktion** f entscheiden darüber, welche Glättungseigenschaften das Funktional \mathcal{F} hat. Wenn wir also von dem wiederhergestellten Bild erwarten, dass es keine großen Oszillationen aufweisen soll, können wir die Potentialfunktion f nun so wählen, dass f sein Minimum in 0 annimmt und für große Argumente ebenfalls große Werte annimmt. Geeignete Funktionen für f sind dann also zum Beispiel $f(t) = t^2, |t|, |t|^\alpha$ mit $1 < \alpha$ oder auch die nichtkonvexen Funktionen $\sqrt{\alpha + t^2}, \log(\alpha t^2 + 1)$. Insbesondere betrachten wir vor allem die folgenden beiden Funktionale:

$$\text{Dirichlet-Integral } \mathcal{F}(x) = \int_{\Omega} |\nabla x|^2 ds, \quad (4.3)$$

$$\text{Totalvariation } \mathcal{F}(x) = \int_{\Omega} |\nabla x| ds \text{ bzw. } \mathcal{F}(x) = |x|_{TV}. \quad (4.4)$$

Nach Multiplikation von (4.2) mit $\mu = \frac{1}{\lambda}$, die nur eine Normierung des zweiten Summanden bewirkt und an der globalen Minimalstelle des Funktionals nichts ändert, erhalten wir also das folgende Funktional, das wir von nun an betrachten:

$$F(x) = \int_{\Omega} (I(s) - S(x(s)))^2 ds + \mu \int_{\Omega} f(|\nabla x(s)|) ds. \quad (4.5)$$

Der erste Summand des Funktionals gibt die Abweichung vom Ausgangsdatensatz an, und sorgt bei der Minimierung des Funktionals somit dafür, dass das wiederhergestellte Bild dem beobachteten Bild ähnlich sieht. Wir wollen ihn **Datenterm** nennen. Der zweite Summand ist der **Regularisierungsterm** und wirkt glättend. Der Parameter μ heißt auch

Regularisierungsparameter. Wir sehen auch an diesem Funktional, dass der Datenterm notwendig ist. Würde er fehlen, so würden wir für die genannten Regularisierungsterme (4.3) und (4.4) ein Bild mit konstantem Grauwert als Minimum erhalten, da der Gradient nach Minimierung überall gleich 0 wäre.

Die Notwendigkeit der Regularisierung

Wir wollen nun einige Aspekte erwähnen, die die Notwendigkeit des Regularisierungsterms betreffen. Wie bereits in Abschnitt 2.2 angedeutet wurde, ist die Entzerrung eines Bildes ein schlecht gestelltes Problem. Wir wollen an dieser Stelle noch einmal eine genauere Begründung angeben und nehmen für einen Moment an, dass $f = 0$ ist, wir also keine Regularisierung haben. Weiterhin seien x sowie I diskrete Variablen des \mathbb{R}^m , und S werde in diesem diskreten Ansatz als $m \times m$ -Matrix aufgefasst. So gilt für eine hinreichende glatte globale Minimalstelle x des Funktionals (4.5) – falls dieses existiert – notwendig die Euler-Lagrange-Gleichung (vgl. Satz 3.5)

$$S^*Sx - S^*I = 0, \tag{4.6}$$

wobei S^* die Transponierte von S bezeichne. Zur Motivation des Regularisierungsterms lassen sich also nun zwei Gründe angeben:

1. Ist $S \neq E$, wobei E die Einheitsmatrix bezeichne, so ist in typischen Anwendungen S^*S nicht invertierbar oder schlecht konditioniert. Somit ist der Regularisierungsterm in diesem Fall notwendig, um eine eindeutige Lösung und somit ein gut gestelltes Problem zu gewährleisten.
2. Ist $S = E$, der systematische Fehler also nicht vorhanden, so wäre wegen (4.6) $x = I$ und somit würden wir als Ergebnis das verrauschte Bild erhalten. Im Fall der kontinuierlichen Bilddarstellung wäre also $x \in L^2(\Omega)$, was unerwünscht ist, da der Raum $L^2(\Omega)$ Funktionen enthält, die zu große Oszillationen aufweisen, vgl. Abschnitt 2.4. Der Regularisierungsterm ist in diesem Fall sinnvoll, weil er zu große Beträge des Gradienten von x bestraft und garantiert, dass das erhaltene Bild ausreichend regulär ist, d.h. $x \in W^{1,p}(\Omega)$.

4.1.2 Die Wahl der Potentialfunktion f

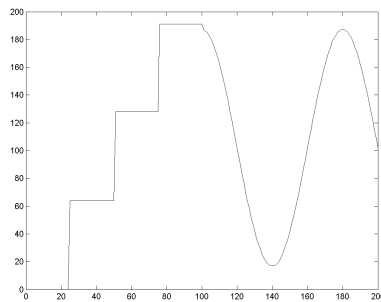
Für die Wahl des Integranden $f(r)$ in (4.5) gibt es, wie wir bereits erwähnt haben, mehrere Möglichkeiten, die unterschiedliche Auswirkungen auf das Bild haben. Wir wollen an dieser Stelle einige gut untersuchte Resultate zusammenfassen:

1. Die *quadratische Regularisierung* $f(r) = r^2$ – wie die Regularisierung mit Hilfe des Dirichlet-Integrals (4.3) auch genannt wird – bewirkt eine *isotrope Diffusion*, d.h. es wird in alle Richtungen geglättet und somit auch in Normalenrichtung der Kanten. Folglich werden Kanten zu stark geglättet (vgl. [3], pp. 63 ff.), was natürlich unerwünscht ist. Dieser Effekt wird in der Fachsprache „*oversmoothing*“ genannt und wird in Abbildung 4.1c an einem Signal deutlich.
2. Die *Totalvariation* (oder TV-Regularisierung) $|r|_{TV}$ wurde als Regularisierungsterm zuerst von Rudin/Osher/Fatemi in [25] untersucht und ist aufgrund wünschenswerter Eigenschaften sehr populär in der Forschung. Das zugehörige Modell lautet

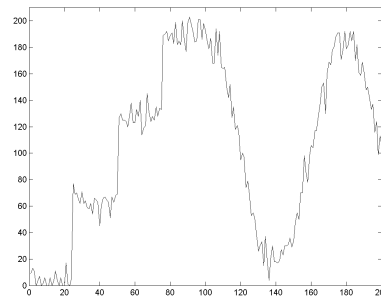
$$F(x) = \int_{\Omega} (I(s) - x(s))^2 ds + \mu |x(s)|_{TV}, \tag{4.7}$$

und wird auch als ROF-Modell bezeichnet. Es verringert Rauschen und erhält gleichzeitig scharfe Kanten. Das Glättungsverhalten nennen wir **anisotrop**, d.h. dass bei Kanten bevorzugt in Richtung der Tangente der Isophote geglättet wird. Ihr Nachteil ist, dass Kanten auch dort entstehen, wo vorher keine waren. Dieser Effekt wird „staircasing“ genannt und wird auch in Abbildung 4.1d deutlich.

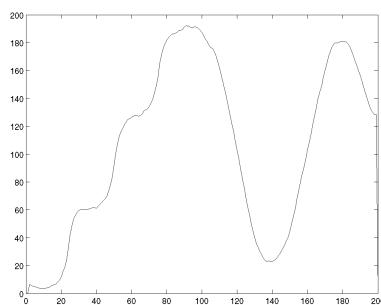
3. Die *konvexe, nichtquadratische Regularisierung* $f(r) = \sqrt{1 + r^2}$ bewirkt eine isotrope Glättung in Punkten mit kleinem Betragsgradienten, sowie eine anisotrope Glättung, wenn der Betragsgradient groß ist.
4. Die von *Perona und Malik* vorgeschlagene nichtkonvexe Funktion $f(r) = \frac{r^2}{1+r^2}$ führt in den Anwendungen zu den besten numerischen Resultaten, da sie eine anisotrope Glättung bewirkt und scharfe Kanten erzielt. Da die Funktion nichtkonvex ist, können wir Satz 3.2 nicht mehr anwenden. Es kann sogar gezeigt werden, dass eine Lösung des Minimierungsproblems im Raum $W^{1,p}(\Omega)$ nicht existiert, vgl. dazu [3] pp. 81. An dieser Stelle weisen wir auf die Diskrepanz zwischen Theorie und Praxis hin, die in der Bildverarbeitung manchmal vorkommt. Ein Verfahren kann in der Theorie Mängel aufweisen, wie z.B. fehlende Existenz in geeigneten Räumen oder fehlende Eindeutigkeit der Lösung, und dennoch in der Praxis zu guten Ergebnissen führen.



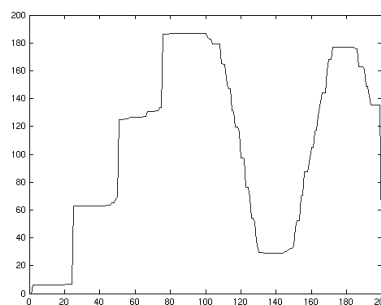
(a) Ungestörtes Signal



(b) Signal mit Gauss'schem Rauschen



(c) Quadratische Regularisierung



(d) TV-Regularisierung

Abbildung 4.1: Vergleich von quadratischer und TV-Regularisierung an einem verrauschten Signal: Während die quadratische Regularisierung aus scharfen Kanten Kurven macht, neigt die TV-Regularisierung zu „staircasing“.

4.1.3 Existenz und Eindeutigkeit einer Lösung

Wir wollen hier nur exemplarisch die Existenz eines Minimums für das Variationsproblem mit quadratischer Regularisierung zeigen. Dazu werden wir das folgende Lemma benötigen:

Lemma 4.1 (Youngsche Ungleichung):

Seien a, b reelle Zahlen. Dann gilt die Youngsche Ungleichung

$$ab \leq \frac{a^2}{2} + \frac{b^2}{2}. \quad (4.8)$$

Beweis. Es ist offensichtlich $a = b$ genau dann, wenn (4.8) mit Gleichheit gilt. Ist $a \neq b$, so folgern wir:

$$\begin{aligned} 0 < (a - b)^2 &= a^2 - 2ab + b^2 \\ \Rightarrow 2ab &< a^2 + b^2, \end{aligned}$$

und damit ist die Aussage gezeigt. □

Nun können wir den folgenden Satz beweisen:

Satz 4.1 (Existenz und Eindeutigkeit bei quadratischer Regularisierung):

Sei $\Omega \subset \mathbb{R}^n$ offen, beschränkt und habe einen Lipschitzrand. Ist $\mu > 0$, dann hat das Problem

$$\text{Minimiere } \left\{ F(x) = \int_{\Omega} (I(s) - x(s))^2 ds + \mu \int_{\Omega} |\nabla x(s)|^2 ds \mid x \in W_0^{1,p}(\Omega) \right\}$$

eine eindeutig bestimmte globale Minimalstelle.

Beweis. In diesem Fall können wir auf Satz 3.2 zurückgreifen, da der Integrand durch $f(s, \omega, \xi) = (I(s) - \omega)^2 + \mu |\xi|^2$ gegeben ist. Damit ist offensichtlich $\xi \mapsto f(s, \omega, \xi)$ konvex. Wir zeigen nun die Bedingung (ii) aus Satz 3.2: Es ist

$$(I - \omega)^2 = I^2 - 2I\omega + \omega^2 \geq \frac{1}{2}\omega^2 - I^2,$$

wobei wir ausgenutzt haben, dass für $I, \omega \geq 0$ die Youngsche Ungleichung (4.8) gilt:

$$\omega I \leq \frac{\omega^2}{2} + \frac{I^2}{2} \Leftrightarrow -\left(2\left(\frac{\omega}{\sqrt{2}}\right)\left(I\sqrt{2}\right)\right) \geq -\left(\frac{\omega^2}{2} + 2I^2\right).$$

Deshalb folgt nun

$$f(s, x, \nabla x) \geq \frac{1}{2}x^2 - I^2 + \mu |\nabla x|^2,$$

und damit ist Bedingung (ii) aus Satz 3.2 erfüllt. Nun zeigen wir noch, dass der Datenterm

$$F_{DT}(x) := \int_{\Omega} (I(s) - x(s))^2 ds$$

strikt konvex ist. Es ist zu zeigen:

$$F_{DT}(\alpha x + (1 - \alpha)y) < \alpha F_{DT}(x) + (1 - \alpha)F_{DT}(y), \quad \forall \alpha \in (0, 1), \forall x \neq y.$$

Dies sehen wir so:

$$\begin{aligned}
 F_{DT}(\alpha x + (1 - \alpha)y) &= \int_{\Omega} (I - (\alpha x + (1 - \alpha)y))^2 ds \\
 &= \int_{\Omega} (I - \alpha I + \alpha I - (\alpha x + (1 - \alpha)y))^2 ds \\
 &= \int_{\Omega} (\alpha^2(I - x)^2 + (1 - \alpha)^2(I - x)^2 + 2\alpha(1 - \alpha)(I - x)(I - y)) ds \\
 &\leq \int_{\Omega} (\alpha^2(I - x)^2 + (1 - \alpha)^2(I - x)^2 \\
 &\quad + \alpha(1 - \alpha) [(I - x)^2 + (I - y)^2]) ds \\
 &= \int_{\Omega} \alpha(I - x)^2 + (1 - \alpha)(I - y)^2 ds \\
 &= \alpha F_{DT}(x) + (1 - \alpha)F_{DT}(y).
 \end{aligned}$$

Dabei haben wir im dritten Schritt wieder die Youngsche Ungleichung (4.8) in der folgenden Form verwendet:

$$(I - x)(I - y) \leq \frac{1}{2} ((I - x)^2 + (I - y)^2),$$

wobei „ \leq “ genau dann gilt, wenn $x = y$ ist. Da die Funktion $\xi \rightarrow |\xi|^2$ strikt konvex ist, ist der quadratische Regularisierungsterm offensichtlich ebenfalls strikt konvex, und somit auch das gesamte Funktional F . Aus Satz 3.2 folgt nun die Existenz und Eindeutigkeit der Lösung. \square

Was die Existenz und Eindeutigkeit von Lösungen allgemeinerer Fälle angeht, merken wir an:

1. Die Existenz einer Lösung für das Problem der TV-Regularisierung wird zum Beispiel in [9], pp. 52, oder [12], pp. 169 ff. gezeigt. Dabei reicht es nicht aus, die Lösung in Sobolevräumen zu vermuten, stattdessen ist es erforderlich, den größeren Raum $BV(\Omega)$ zu betrachten. In diesem Raum existiert tatsächlich eine Lösung. Da das Funktional strikt konvex ist, ist diese sogar eindeutig (vgl. [9], pp. 52).
2. Die Existenzfrage für allgemeine konvexe Potentialfunktionen wird in [3], pp. 66 ff., behandelt. Es ist notwendig, das relaxierte Funktional zu betrachten. Dann kann unter gewissen Voraussetzungen ebenfalls die Existenz und Eindeutigkeit einer Lösung im Raum $BV(\Omega)$ gefolgert werden.

4.1.4 Anwendung der Euler-Lagrange-Gleichungen

Um die enge Verknüpfung des Variationsansatzes mit dem Lösungsansatz der partiellen Differentialgleichungen aufzuzeigen, wenden wir die Euler-Lagrange-Gleichungen aus Satz 3.5 auf den in Abschnitt 4.1.1 hergeleiteten Variationsansatz an. Wir nehmen unsere Lösungen x in dem Raum $C^2(\Omega)$ an und wollen uns im Folgenden auf den Fall der quadratischen Regularisierung und der Totalvariation beschränken.

Quadratische Regularisierung

Für die quadratische Regularisierung (4.3) betrachten wir wegen (2.5) das Funktional

$$F(x) = \int_{\Omega} (S(x(s)) - I(s))^2 ds + \mu \int_{\Omega} |\nabla x(s)|^2 ds.$$

Anwendung von Satz 3.5 liefert die notwendige Optimalitätsbedingung

$$\begin{aligned} F'(x) &= -\operatorname{div}(2\mu\nabla x) + 2S^*Sx - 2S^*I = 0 \\ &\Rightarrow -\mu\Delta x + S^*Sx - S^*I = 0 \end{aligned}$$

mit der Randbedingung $x = x_0$ auf $\partial\Omega$ für gegebenes $x_0 \in W^{1,p}(\Omega)$. Es bezeichne dabei Δ den Laplace-Operator

$$\Delta x = \frac{\partial^2 x}{\partial s_1^2} + \frac{\partial^2 x}{\partial s_2^2}.$$

TV-Regularisierung

Für die TV-Regularisierung (4.4) betrachten wir für $x \in C^2(\Omega)$ wegen das Funktional

$$F(x) = \int_{\Omega} (S(x(s)) - I(s))^2 ds + \mu \int_{\Omega} \sqrt{\left(\frac{\partial x}{\partial s_1}\right)^2 + \left(\frac{\partial x}{\partial s_2}\right)^2} ds.$$

Wir erhalten wieder durch Anwendung des Satzes 3.5

$$F'(x) = -\operatorname{div} \left(\mu \frac{2\frac{\partial x}{\partial s_1}}{\sqrt{\left(\frac{\partial x}{\partial s_1}\right)^2 + \left(\frac{\partial x}{\partial s_2}\right)^2}}, \mu \frac{2\frac{\partial x}{\partial s_2}}{\sqrt{\left(\frac{\partial x}{\partial s_1}\right)^2 + \left(\frac{\partial x}{\partial s_2}\right)^2}} \right) + S^*Sx - S^*I = 0 \quad (4.9)$$

$$\Rightarrow -\mu \operatorname{div} \left(\frac{\nabla x}{|\nabla x|} \right) + S^*Sx - S^*I = 0 \quad (4.10)$$

mit der Randbedingung $x = x_0$ auf $\partial\Omega$ für gegebenes $x_0 \in W^{1,p}(\Omega)$.

4.2 Das Kantenerkennungsfunktional

Ein bekanntes und gut untersuchtes Problem in der Bildverarbeitung ist neben der Bildentrauschung das Problem der Segmentierung. In einem Bild sind oft unterschiedliche Objekte zu sehen. Das Problem der Segmentierung besteht darin, die Ränder oder Kanten dieser Objekte zu finden. Um das Problem zu lösen, müssen zwei fundamentale Schritte ausgeführt werden. Zunächst muss definiert werden, was eine Kante ist, und anschließend muss ein Algorithmus gefunden werden, der diese Kanten extrahiert.

Intuitiv definieren wir eine Kante als die Menge der Punkte, in denen der Betrag des Gradienten groß ist, also formal durch

$$\Gamma := \{s \in \Omega \mid |\nabla x(s)| \geq p\},$$

wobei der Grenzparameter p angibt, wann ein Punkt zu einer Kante gehört. Wir sehen jedoch schnell, dass eine allein auf dieser Definition basierende Kantenerkennung aus zwei Gründen zu naiv ist:

1. Die Kantenerkennung kann nicht mit Rauschen umgehen, da sie zwischen einem betragsgroßen Gradienten, der durch Rauschen, und einem betragsgroßen Gradienten, der durch echte Kanten hervorgerufen wird, nicht unterscheiden kann.
2. Die Definition einer Kante ist abhängig von dem Parameter p , der a priori festgelegt werden muss. Insbesondere ist die Definition nicht invariant unter einer linearen Abbildung $x \rightarrow \lambda x$ mit $\lambda \in (0, \infty)$, während für das menschliche Auge die Kantensmengen der Bilder x und des Vielfachen λx gleich wären.

Diese Probleme motivierten neue Verfahren (vgl. [15], pp. 318 ff.), die versuchten diese Probleme zu lösen, jedoch auf dem Vergleich benachbarter Pixel basierten. Ein globaler Ansatz zur Kantenbestimmung, der auf der Minimierung eines Energiefunktional basierte, wurde 1989 von Mumford und Shah entwickelt. Die Idee besteht darin, ausgehend von einem aufgezeichneten Bild I eine nahezu stückweise konstante Approximation x und eine optimale Kantenmenge K zu finden, indem ein geeignetes Energiefunktional E_{MS} minimiert wird. Das Problem wollen wir abstrakt wie folgt formulieren:

Minimiere $E_{MS}(x, K)$.

Das Energiefunktional, das wir betrachten wollen, besteht dabei aus drei Termen:

1. Einem Glattheitsterm, der garantiert, dass die Funktion x in homogenen Regionen, also solchen die nicht zur Kantenmenge gehören, glatt genug ist:

$$\int_{\Omega \setminus K} |\nabla x(s)|^2 ds.$$

2. Einem Datenterm:

$$\int_{\Omega \setminus K} |x(s) - I(s)|^2 ds.$$

3. Einem Term, der die „Gesamtlänge“ der Kanten im Bild misst und sicherstellt, dass nicht das ganze Bild mit Kanten gefüllt wird. Wir bezeichnen diese Kantenlänge mit $\text{length}(K)$.

Der erste und dritte Term aus dieser Auflistung sorgen dafür, dass Rauschen nicht als Kante erkannt wird. Definiert man nun die Länge von K mit Hilfe des eindimensionalen Hausdorffmaßes \mathcal{H} , so erhält man das berühmte Mumford-Shah-Funktional:

$$E_{MS}(x, K) = \int_{\Omega \setminus K} \left(\alpha |x(s) - I(s)|^2 + |\nabla x(s)|^2 \right) ds + \beta \mathcal{H}(K). \quad (4.11)$$

Die Parameter α und β entscheiden über die Gewichtung der einzelnen Summanden. Das Hausdorffmaß ist eine Verallgemeinerung der Fläche einer Untermannigfaltigkeit und soll hier nicht näher spezifiziert werden (vgl. z.B. [3], pp. 41). Der Erfolg des Mumford-Shah-Funktional erklärt sich damit, dass bei geeigneter Parameterwahl nur zusammenhängende Kanten dargestellt werden, d.h. Rauschen nicht als Kante erkannt wird. Zur Frage der Existenz und Eindeutigkeit des zugehörigen Minimierungsproblems vergleiche man etwa [15], pp. 355 ff.

Das Funktional (4.11) selbst zu minimieren ist kein triviales Problem, da gleichzeitig sowohl eine Funktion x als auch eine Menge K gesucht werden. Es wird auch „freies Randwertproblem“ genannt, da die Randwerte der Funktion x auf dem Rand ∂K nicht bekannt sind.

Aus diesem Grund wird in der Literatur (vgl. [3], pp. 135 und [6], pp. 231) das folgende äquivalente Problem betrachtet, das nur die Funktion x als Optimierungsvariable enthält:

$$E(x) = \int_{\Omega} \left(\alpha |x(s) - I(s)|^2 + |\nabla x(s)|^2 \right) ds + \beta \mathcal{H}(S_x).$$

Dabei bezeichne S_x die Menge der Unstetigkeiten von x (vgl. [3], pp. 135 ff., [6], pp. 231 ff., für genauere Ausführungen). Ambrosio/Tortorelli [2] haben vorgeschlagen, das Funktional E durch eine Folge von Funktionalen F_ϵ zu approximieren. Für festes $\epsilon > 0$ haben diese Funktionale $F_\epsilon(x, k)$ mit einer zusätzlichen Funktion $k : \Omega \rightarrow [0, 1]$ die folgende Gestalt, die wir aus [6], pp. 232 übernehmen:

$$\begin{aligned} F_\epsilon(x, k) = & c_1 \int_{\Omega} (x(s) - I(s))^2 ds + c_2 \int_{\Omega} |\nabla x(s)|^2 \cdot (k(s)^2 + c_4) ds \\ & + c_3 \int_{\Omega} \left(\epsilon |\nabla k(s)|^2 + \frac{1}{4\epsilon} (k(s) - 1)^2 \right) ds. \end{aligned} \quad (4.12)$$

Für festes ϵ bewirkt die Minimierung des Funktionals F_ϵ sowohl eine Glättung des Bildes x als auch eine Kantenerkennung. Dabei nimmt die **Kantenfunktion** k in einem Punkt s , der auf einer Kante liegt, einen Wert ≈ 0 an, während sie in einem Punkt s , der auf keiner Kante liegt, einen Wert ≈ 1 annimmt. Somit werden Kanten schwarz und Punkte innerhalb homogener Regionen weiß dargestellt. Der erste Term gewährleistet wieder, dass das gesuchte Bild x nicht zu stark von dem aufgezeichneten Bild I abweicht. Der zweite Term ist ein Regularisierungsterm, bestehend aus der gesuchten Funktion x , gekoppelt mit dem Quadrat der Kantenfunktion k . Dies führt dazu, dass in Punkten mit betragsgroßem Gradienten $|\nabla x|$ die Kantenfunktion k dazu neigt, einen Wert ≈ 0 anzunehmen, also eine Kante identifiziert wird. Das dritte Integral besteht aus einem Regularisierungsterm für k sowie einem Term, der $k(s) \approx 1$ für fast alle Punkte s erzwingt. Je kleiner $\epsilon > 0$ ist, umso mehr neigt k dazu, den Wert 1 anzunehmen, und umso öfter kann $|\nabla k|$ groß werden. Dies bedeutet, dass dünne Kanten bevorzugt werden. Die Parameter $c_1, c_2, c_3, c_4 \geq 0$ sind Konstanten und bestimmen das Ergebnis, indem sie die Terme unterschiedlich stark gewichten.

Bourdin zeigte in [6], dass im Sinne von Γ -Konvergenz $F_\epsilon(x, k) \rightarrow E(x)$ für $\epsilon \rightarrow 0$ gilt. Insbesondere sind die Minima im Falle der Konvergenz gleich.

Die Funktionale der Gestalt (4.12) nennen wir **Kantenerkennungsfunktional** von Ambrosio-Tortorelli und werden sie in Kapitel 6 als Vergleichsverfahren nutzen, um die Qualität von Kantenbildern vergleichen zu können.

Für die numerische Behandlung der Folge (4.12) verweisen wir auf Kapitel 5, Seite 46 ff.

4.3 Fallstudie: Impulsrauschen

4.3.1 Impulsrauschen in der Forschung

Der einfachste Ansatz, um Impulsrauschen zu reduzieren, besteht darin, den Median-Filter (vgl. [15], pp. 200) zu verwenden. Dies ist zwar mit wenig Rechenaufwand verbunden, jedoch hat dieser Filter Schwierigkeiten bei zu großer Rauschintensität ([15], pp. 200). Es wurden deshalb basierend auf diesem Filter viele Verbesserungen, wie zum Beispiel der adaptive Median-Filter oder der Multistate Median-Filter entwickelt. Ihr größter Nachteil

ist, dass sie Kanten nicht hinreichend gut erhalten. Die zugehörigen Algorithmen für den Median-Filter bzw. den adaptiven Median-Filter finden sich in [14], pp. 1480.

Das Problem der mit Impulsrauschen gestörten Bilder wurde insbesondere in den letzten Jahren mit Hilfe des Variationsansatzes unter anderem von Nikolova [22, 23], Chan/Shen/Ho [15, 14] und Bar/Kiryati/Sochen [5] behandelt. Verwenden wir denselben Ansatz wie für das additive Gauss'sche Rauschen (vgl. (4.5)), so stoßen wir bald auf folgendes Problem: Der Datenterm bestraft Abweichungen vom beobachteten Bild quadratisch und gewichtet damit die Eliminierung eines Ausreißers sehr stark. Wenn man also einen starken Ausreißer (engl.: outlier) in einer homogenen Region hat, so würde die Eliminierung dieses Ausreißers eine hohe Energie verursachen, da die Differenz $x - I$ quadratisch einginge. Dies hätte zur Folge, dass ein großer Regularisierungsparameter gewählt werden müsste, um den Ausreißer vollständig zu eliminieren, was wiederum eine starke Glättung des gesamten Bildes bedeutet. Aus diesem Grund wurden von Nikolova [22, 23] nichtglatte Datenterme (insbesondere die L^1 -Norm) in Verbindung mit kantenerhaltenden Regularisierungstermen untersucht.

Konvexe Zielfunktionen mit nichtglattem Datenterm

Wir betrachten den Ansatz von Nikolova [22], pp. 99 ff., der für diskrete Bilder modelliert wurde, nun genauer. Es sei dazu $I \in \mathbb{R}^{n \times n}$ das beobachtete Bild, das auf dem Gitter $\Omega_{n,n}$ gegeben sei. Dabei ist $\Omega_{n,n}$ wie in Abschnitt 2.4.1 definiert. Außerdem bezeichne $\mathcal{N}_{i,j}$ wie in Abschnitt 2.4.1 die Nachbarschaft eines Gitterpunktes $s_{i,j}$. Da wir Bilder in ihrer diskreten Repräsentation betrachten, erhalten wir folglich kein Variationsproblem, sondern ein endlichdimensionales Optimierungsproblem. Nikolova schlägt zur Rauschreduzierung die Minimierung der folgenden Zielfunktion vor:

$$F(x) = \sum_{(i,j) \in \Omega_{n,n}} |x_{i,j} - I_{i,j}| + \frac{\beta}{2} \sum_{s_{i,j} \in \Omega_{n,n}} \sum_{s_{m,n} \in \mathcal{N}_{i,j}} \varphi(x_{i,j} - x_{m,n}). \quad (4.13)$$

Die erste Summe in (4.13) entspricht unserem bisherigen Datenterm. Nun geht die Differenz $x - I$ jedoch nicht mehr quadratisch ein, und insbesondere ist der Integrand im Nullpunkt nicht mehr differenzierbar. Die zweite Summe entspricht unserem bisherigen Regularisierungsfunktional. Für jeden Gitterpunkt $s_{i,j} \in \Omega_{n,n}$ werden die Differenzen mit den Punkten aus der Nachbarschaft betrachtet. Wenn wir annehmen, dass der Abstand zwischen den Gitterpunkten $h = 1$ beträgt, handelt es sich um die jeweiligen Differenzenquotienten mit den Punkten aus der Nachbarschaft, also um eine Approximation des Gradienten. Für die Potentialfunktion φ werden in [22], pp. 99, die konvexen Funktionen

$$\begin{aligned} \varphi(t) &= \sqrt{\alpha + t^2}, \alpha > 0, \\ \varphi(t) &= |t|^\alpha, 1 < \alpha \leq 2 \end{aligned}$$

vorgeschlagen. In diesen Fällen ist die Zielfunktion insgesamt konvex, da die einzelnen Summen konvex sind. Somit ist jede lokale Minimalstelle auch global, aber die Eindeutigkeit einer Minimalstelle ist nicht mehr gewährleistet.

Für diesen diskreten Ansatz wird in [23] nun gezeigt, dass er sich besser zur Eliminierung von Ausreißern eignet als vergleichbare Methoden mit L^2 -Datentermen.

Wir werden diesen Ansatz als Vergleichsverfahren verwenden, wenn es darum geht, Bilder mit Impulsrauschen zu entrauschen. Um einfacher auf ihn verweisen zu können, bezeichnen wir die Minimierung des Funktionals (4.13) kurz als **Methode von Nikolova**, da die Zielfunktionale in diesem Zusammenhang als erste von Nikolova eingehend untersucht wurden.

Bemerkung 4.1. Wie in [14] festgestellt wird, hat die Methode von Nikolova Schwierigkeiten bei sogenannten **Patches**, d.h. zusammenhängenden Pixelketten, die durch Rauschen entstehen. In [14] wird deshalb eine Verbesserung der Methode von Nikolova vorgeschlagen, die das Funktional mit einem auf dem Median-Filter basierenden Rauschdetektor verknüpft. Die Autoren schlagen ein Zwei-Phasen-Verfahren vor, das in der ersten Phase die verrauschten Pixel bestimmt und in der zweiten Phase einen Variationsansatz durchführt. In der ersten Phase wird mit Hilfe des adaptiven Median-Filters die Menge der Pixel bestimmt, die durch Rauschen entstanden sind. Es bezeichne \hat{x} das Bild, das durch Anwendung des adaptiven Median-Filters auf I entsteht. Dann ist die Menge der verrauschten Gitterpunkte gegeben durch:

$$\mathcal{M} := \{s_{i,j} \in \Omega_{n,n} \mid \hat{x}_{i,j} \neq I_{i,j} \text{ und } I_{i,j} \in \{0, 1\}\}.$$

Nun brauchen die Punkte aus $\Omega_{n,n} \setminus \mathcal{M}$ nicht verändert zu werden, da sie gerade die ungestörten Punkte sind. In der zweiten Phase wird der Variationsansatz (4.13) verwendet, um die neuen Werte an den Punkten der Menge \mathcal{M} zu bestimmen. Nach Angaben der Autoren ist diese Methode geeignet, um Impulsrauschen mit einer Intensität von bis zu 90 % zu reduzieren.

L^1 -Datenterm und TV-Regularisierung – kontinuierliches Modell

Das zu (4.13) gehörende kontinuierliche Modell wurde von Chan/Esodoglu [13] untersucht. Für ein gegebene Daten $I \in L^1(\Omega)$ betrachten die Autoren ein Energiefunktional, bestehend aus L^1 -Datenterm und TV-Regularisierungsterm, wobei sie $\Omega = \mathbb{R}^n$ annehmen:

$$F(x) = \int_{\mathbb{R}^n} |I(s) - x(s)| ds + \lambda |x(s)|_{TV} \quad (4.14)$$

Für dieses Modell geben die Autoren Eigenschaften an, die für das Standard-ROF-Modell (4.7), d.h. L^2 -Datenterm mit TV-Regularisierungsterm, nicht gelten. Wir fassen die wichtigsten Eigenschaften kurz zusammen:

1. Nicht-Eindeutigkeit der Lösung des zugehörigen Minimierungsproblems: Das Funktional F ist nicht mehr strikt konvex und somit die Eindeutigkeit der Lösung nicht mehr gegeben. Für eine weiße Kreisscheibe auf schwarzem Grund lässt sich die Lösung der beiden Probleme explizit ausrechnen (vgl. [13]). Es bezeichne $B_r(0)$ diese Kreisscheibe um den Nullpunkt mit Radius r und $\chi_{B_r(0)}$ die charakteristische Funktion, die auf der Kreisscheibe den Wert 1 und sonst den Wert 0 annimmt. Dann ist für einen gegebenen Regularisierungsparameter λ die Lösung des Standard-ROF-Modells

$$x_\lambda(s) \equiv \begin{cases} 0 & | \text{ wenn } \lambda \geq r, \\ (1 - \lambda) \chi_{B_r(0)}(s) & | \text{ wenn } 0 \leq \lambda < r. \end{cases}$$

Die Lösung des Modells (4.14) ist hingegen gegeben durch

$$x_\lambda(s) \equiv \begin{cases} 0 & | \text{ wenn } \lambda > \frac{r}{2}, \\ \chi_{B_r(0)}(s) & | \text{ wenn } 0 \leq \lambda < \frac{r}{2}. \end{cases}$$

Wenn $\lambda = \frac{r}{2}$, so ist die Lösung nicht eindeutig, sondern gegeben durch die Lösungsmenge:

$$\{c \chi_{B_r(0)}(s) : c \in [0, 1]\}.$$

Dabei tritt diese Lösungsmenge in der numerischen Praxis so nicht auf. Numerisch gesehen werden wir Approximationen verwenden, die auch bei dem Modell (4.14) auf eine eindeutige Lösung führen werden. Wir verweisen auf Seite 50.

2. Kontrastinvarianz: Im Gegensatz zu dem quadratischen Datenterm ist das Funktional (4.14) kontrastinvariant im Folgenden Sinne: Ist $x(s)$ Lösung des Minimierungsproblems bei gegebenen Daten $I(s)$, so ist $cx(s)$ auch Lösung des Minimierungsproblems, wenn das Bild $cI(s)$ beobachtet wurde, wobei $c > 0$ konstant sei. Abbildung 4.2 zeigt die praktischen Konsequenzen dieser Eigenschaft. Wir betrachten das Originalbild aus Abbildung 4.2a. Es enthält zwei groß- und zwei kleinskalige Rechtecke, die jeweils unterschiedlichen Kontrast zum Hintergrund aufweisen. Die beiden kleinskaligen Rechtecke können wir als Rauschen ansehen und uns das Ziel setzen, diese zu eliminieren, ohne dabei die großskaligen Rechtecke zu zerstören. In der obigen Zeile der Abbildung 4.2 verwenden wir dazu den L^2 -Datenterm, während wir in der unteren Zeile den L^1 -Datenterm verwenden. Während beim L^2 -Datenterm die Rechtecke mit gleicher Grauwertintensität beim gleichen Regularisierungsparameter λ verschwinden, erhalten wir durch Verwendung des L^1 -Datenterms den gewünschten Effekt. Die beiden kleinskaligen Rechtecke verschwinden zuerst, während die größeren Rechtecke erhalten bleiben. Wählen wir λ groß genug, so verschwinden auch die großskaligen Rechtecke, und zwar für dieselbe Wahl von λ .
3. Ingesamt lässt sich somit auch sagen, dass das Modell (4.14) die „Geometrie“ der Bildobjekte besser respektiert und erhält. Damit meinen wir, dass der Regularisierungsprozess mehr von der geometrischen Form der Objekte als von ihrem Kontrast abhängt. Kleine Objekte bleiben im Kontrast erhalten, auch wenn λ größer wird, bis sie abrupt verschwinden.

4.3.2 Entzerrung und Impulsrauschen – L^1 -Datenterm und Mumford-Shah-Regularisierungsterm

Wird ein Bild gleichzeitig verzerrt und mit Impulsrauschen versehen, so stoßen wir bei der Wiederherstellung des Bildes auf Schwierigkeiten, die von Bar/Kiryati/Sochen in [5], pp. 279 ff. untersucht wurden. Die Autoren zeigen, dass eine Filterung mit einem 3×3 Median-Filter, gefolgt von einer Entzerrung mit Hilfe eines Variationansatzes zu unbefriedigenden Ergebnissen führt. Von Bar/Kiryati/Sochen wird deshalb eine Methode vorgeschlagen, die sich sowohl dazu eignet, Bilder zu entzerren als auch gleichzeitig Impulsrauschen zu eliminieren. Es bezeichne, wie im Fehlermodell (2.2), S den systematischen Fehler, der das Bild verzerrt. Die Autoren schlagen vor, für den Datenterm die modifizierte L^1 -Norm zu verwenden (vgl. [5], pp. 282):

$$\|S(x) - I\|_\eta := \sqrt{(S(x) - I)^2 + \eta}. \quad (4.15)$$

Dabei wird aus numerischen Gründen η positiv und fest gewählt, was zur Folge hat, dass der Datenterm auch im Nullpunkt differenzierbar wird.

Bemerkung 4.2. Für $(S(x) - I)^2 \ll \eta$ tendiert die Norm (4.15) dazu, die L^2 -Norm zu

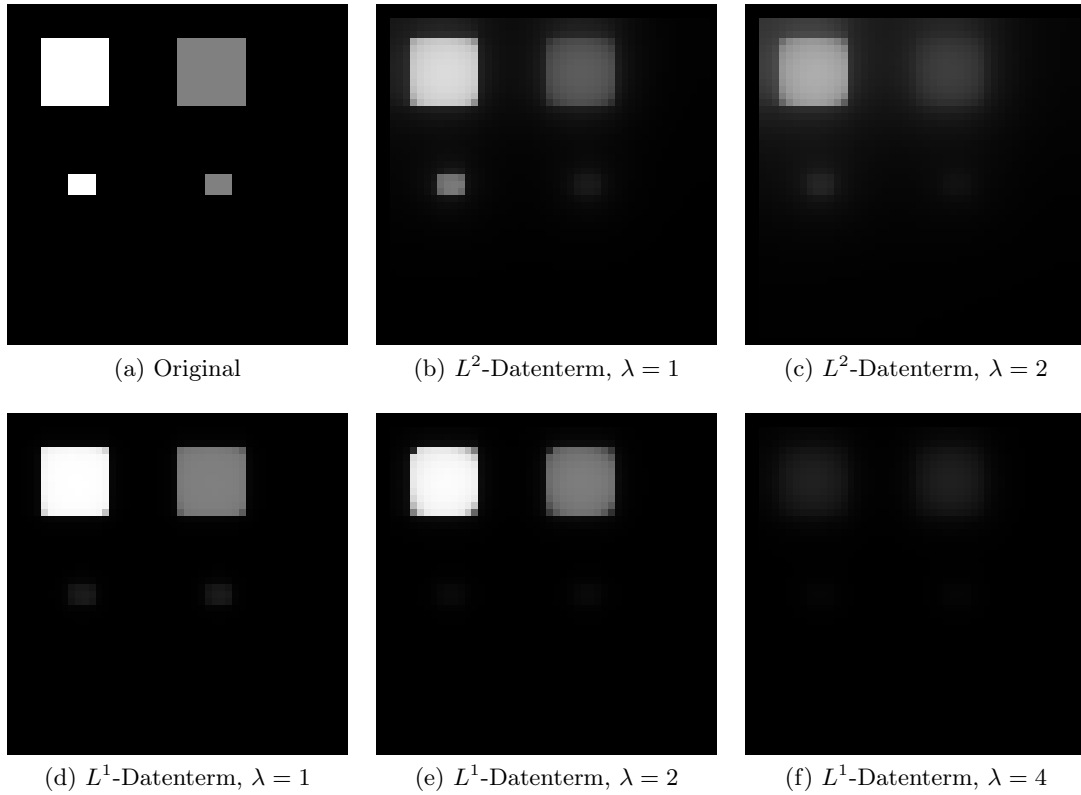


Abbildung 4.2: L^1 -Datenterm garantiert im Gegensatz zum L^2 -Datenterm die Kontrastinvarianz

approximieren, was wir an folgender Abschätzung sehen:

$$\begin{aligned} \sqrt{(S(x) - I)^2 + \eta} &= \sqrt{\eta} \sqrt{1 + \frac{(S(x) - I)^2}{\eta}} \\ &\approx \sqrt{\eta} \sqrt{1 + \frac{(S(x) - I)^2}{\eta} + \frac{(S(x) - I)^4}{4\eta^2}} \\ &= \sqrt{\eta} \left(1 + \frac{(S(x) - I)^2}{2\eta} \right). \end{aligned}$$

Als Regularisierungsterm wählen die Autoren den Regularisierungsterm aus der Approximation des Mumford-Shah-Funktional (4.12) mit $c_4 = 0$ (vgl. [5], pp. 283), d.h.

$$\mathcal{F}_\epsilon(x, k) = c_2 \int_{\Omega} |\nabla x(s)|^2 \cdot k(s)^2 ds + c_3 \int_{\Omega} \left(\epsilon |\nabla k(s)|^2 + \frac{1}{4\epsilon} (k(s) - 1)^2 \right) ds, \quad (4.16)$$

und erhalten damit das Funktional

$$F_\epsilon(x, k) = \int_{\Omega} \sqrt{(S(x) - I)^2 + \eta} ds + \mathcal{F}_\epsilon(x, k). \quad (4.17)$$

Eine Minimierung dieses Funktional sollte also nun eine entrauschte Version unseres veräuschten Bildes liefern. Gleichzeitig erhalten wir, da wir den „Kantenerkennungsterm“

aus dem Ambrosio-Tortorelli-Funktional verwenden, ein Kantenbild k . Auch diesen Ansatz werden wir als Vergleichsverfahren nutzen, wenn wir mit Impulsrauschen verrauschte Bilder entrauschen wollen. Um diesem Ansatz einen Namen zu geben, auf den wir leicht verweisen können, nennen wir ihn **Methode von Bar/Kiryati/Sochen**.

Nach [5], pp. 293 ist diese Methode in dem Sinne robust, dass bei unterschiedlichen Bildern mit gleicher Rauschintensität ähnliche Parameter gewählt werden können, um ähnlich gute Ergebnisse zu erzielen. Nach dem einmaligen Auffinden von geeigneten Parameterpaaren, müssen diese also nur noch gering variiert werden, um auch bei anderen Bildern gute Ergebnisse zu erzielen.

Bemerkung 4.3. Cai/Chan/Nikolova schlagen in [11] eine weitere Verbesserung dieses Modells vor. Dabei handelt es sich wie in [14] um ein Zwei-Phasen-Verfahren, das in der ersten Phase den Median-Filter als Rauschdetektor verwendet und in der zweiten Phase den Variationsansatz (4.17), vgl. Bemerkung 4.1.

4.4 Steuerung und Steuerbeschränkung im Variationsansatz

Nach dem Kenntnisstand des Verfassers wurden Steuerungsprobleme in Verbindung mit dem image-restoration-Problem noch nicht untersucht. Der Steuerungsansatz in dieser Arbeit beruht auf der Idee, den Bildgradienten als Steuerung aufzufassen und anschließend den Einfluß einer Steuerbeschränkung zu untersuchen. Wir betrachten also den Variationsansatz (4.5) und führen für $x \in W_0^{1,p}(\Omega)$ eine künstliche Steuervariable $u \in L^p(\Omega, \mathbb{R}^2)$ ein, indem wir sie durch

$$u(s) = (u_1(s), u_2(s))^T := \nabla x(s), (\forall) s \in \Omega$$

erklären. Somit erhalten wir das Steuerungsproblem

$$\begin{aligned} \text{Minimiere } F(x, u) &= \int_{\Omega} (I(s) - S(x(s)))^2 ds + \mu \int_{\Omega} f(|u(s)|) ds \\ \text{unter } u(s) &= (u_1(s), u_2(s))^T = \nabla x(s), (\forall) s \in \Omega. \end{aligned} \quad (4.18)$$

Uns interessiert nun die Frage, wie sich Steuerbeschränkungen der Gestalt

$$|u_1(s)|^q + |u_2(s)|^q \leq R^q, \quad u \in L^p(\Omega, \mathbb{R}^2) \quad (\forall) s \in \Omega, 1 \leq q < \infty, R \in \mathbb{R} \quad (4.19)$$

auswirken. Hier wird also die $\|\cdot\|_q$ -Norm von u beschränkt. Der Steuerbereich U lautet also in diesem Fall:

$$U = \{u \in L^p(\Omega, \mathbb{R}^2) \mid |u_1(s)|^q + |u_2(s)|^q \leq R^q, (\forall) s \in \Omega.\}$$

Die Motivation für unser Vorgehen ist die folgende: Der Gradient – und somit unsere Steuerung – drückt die Oszillationen im Bild aus. Beschränken wir also unsere Steuerung, so werden Oszillationen beschränkt und das Bild damit glatter. Zu klären ist dann auch die Frage, in welchen Fällen diese Beschränkung effektiv ist, und ob und in welchem Sinne die Steuerbeschränkung selbst regularisierend wirkt, d.h. zu untersuchen ist insbesondere der Fall $\mu = 0$. Eine weitere Motivation für unser Vorgehen ist die Kantenerkennung. Da große Gradienten Kanten im Bild anzeigen, stellt sich für uns die Frage, inwiefern eine Kantenerkennung mit Hilfe der Steuerung sinnvoll ist. Die Idee besteht darin, Kanten dort zu definieren, wo die Steuerung aktiv wird, d.h. wo $|u_1(s)|^q + |u_2(s)|^q = R^q$ gilt, oder auch

dort wo die Schranke R „fast“ angenommen wird, d.h. wo der Wert $|u_1(s)|^q + |u_2(s)|^q$ nahe R liegt.

Neben Beschränkungen der $\|\cdot\|_q$ -Norm von u für $1 \leq q < \infty$ ist es auch von besonderem Interesse, diese Beschränkung für $q = \infty$ zu untersuchen. Dies bedeutet, dass wir die Maximumsnorm von u beschränken:

$$\begin{aligned} -R_1 &\leq u_1(s) \leq R_1, \quad (\forall) s \in \Omega, \\ -R_2 &\leq u_2(s) \leq R_2, \quad (\forall) s \in \Omega. \end{aligned} \tag{4.20}$$

Dabei ist nicht notwendig $R_1 = R_2$, sodass wir die Komponenten von u unterschiedlich stark beschränken können. Wie wir sehen werden, gibt es durchaus Fälle, bei denen es sinnvoll ist die eine Komponente des Gradienten stärker zu beschränken als die andere.

Wir schreiben das Steuerungsproblem (4.18) mit der Steuerbeschränkung (4.19) bzw. (4.20) noch einmal vollständig auf:

$$\begin{aligned} (P_1) \quad \text{Minimiere } F(x, u) &= \int_{\Omega} (I(s) - S(x(s)))^2 ds + \mu \int_{\Omega} f(|u(s)|) ds, \\ (x, u) &\in W_0^{1,p}(\Omega) \times L^p(\Omega, \mathbb{R}^2) \\ u(s) &= (u_1(s), u_2(s))^T = \nabla x(s), \quad (\forall) s \in \Omega, \\ u(s) &\in K := \{(v_1, v_2) \in \mathbb{R}^2 \mid |v_1|^q + |v_2|^q \leq R^q, 1 \leq q \leq \infty\} \quad (\forall) s \in \Omega. \end{aligned}$$

(P_1) ist ein Problem vom Dieudonné-Rashevsky-Typ. Dementsprechend läßt sich das Pontrjaginsche Maximumprinzip aus Abschnitt 3.2.1 (Satz 3.6) anwenden.

Satz 4.2 (Pontrjaginsches Maximumprinzip für das Steuerungsproblem (P_1) mit Steuerbeschränkung):

Sei $\Omega \subset \mathbb{R}^2$ Abschluß eines Lipschitzgebietes im strengen Sinne. Es sei $1 < p < \infty$ und $\mu > 0$, $R > 0$. Die beobachtete Funktion $I : \Omega \rightarrow [0, 1]$ sei stetig, und die Potentialfunktion f sei konvex und zweimal stetig differenzierbar. Dann existieren für jede globale Minimalstelle $(x^*, u^*) \in W_0^{1,p}(\Omega) \times L^p(\Omega, \mathbb{R}^2)$ von (P_1) Multiplikatoren $\lambda_0 > 0$ und $y \in L^{\bar{p}}(\Omega, \mathbb{R}^2)$, $\frac{1}{\bar{p}} + \frac{1}{p} = 1$, die nicht gleichzeitig verschwinden, so dass die folgenden Bedingungen erfüllt sind:

$$\begin{aligned} (M) : \quad \lambda_0 \mu \int_{\Omega} (f(u_1(s), u_2(s)) - f(u_1^*(s), u_2^*(s))) ds \\ - \int_{\Omega} (u_1(s) - u_1^*(s)) y_1(s) ds - \int_{\Omega} (u_2(s) - u_2^*(s)) y_2(s) ds \geq 0 \quad \forall u \in U, \end{aligned}$$

$$\begin{aligned} (K) : \quad 2\lambda_0 \int_{\Omega} (x^*(s) - I(s)) (x(s) - x^*(s)) ds \\ + \int_{\Omega} \left(\frac{\partial x}{\partial s_1}(s) - u_1^*(s) \right) y_1(s) ds + \int_{\Omega} \left(\frac{\partial x}{\partial s_2}(s) - u_2^*(s) \right) y_2(s) ds = 0 \\ \forall x \in W_0^{1,p}(\Omega, \mathbb{R}^n), \end{aligned}$$

$$\begin{aligned} (MP) : \quad \lambda_0 \mu f(u_1^*(s), u_2^*(s)) + u_1^*(s) y_1(s) + u_2^*(s) y_2(s) \\ = \max_{\xi \in \mathbb{R}^2, |\xi| \leq R} (-\lambda_0 \mu f(\xi_1, \xi_2) + \xi_1 y_1(s) + \xi_2 y_2(s)) \quad \forall \xi \in K \quad (\forall) s \in \Omega. \end{aligned}$$

Nachdem wir nun das grundsätzliche Steuerungsproblem (P_1) eingeführt haben, wollen wir im nächsten Abschnitt insbesondere noch einmal die Beschränkung der Maximumsnorm von u studieren und das Modell erweitern.

4.5 L^∞ -Steuerbeschränkung und „rotating mask“

Der im letzten Abschnitt eingeführte Steuerungsansatz (P_1) läßt sich nun erweitern. Im Allgemeinen ist im Voraus nicht bekannt, welche Richtung in einem Punkt gerade beschränkt werden soll. Ziel eines guten Verfahrens ist es, *anisotrop* zu glätten, d.h. entlang der Kanten (d.h. der Tangentialrichtung der Isophoten) und nicht über sie hinweg (d.h. in Normalenrichtung der Isophoten). Nun ist dies allein durch das Steuerungsproblem (4.18) mit der Beschränkung der Maximumsnorm der Steuerung (4.20) noch nicht gewährleistet.

Im Grunde muss in jedem Punkt unterschieden werden, ob der Punkt zu einer homogenen Region, zu einer Kante oder zum Rauschen gehört. Dies erreichen wir, indem wir die Nachbarschaft eines Punktes betrachten. Wir stellen in den nächsten beiden Unterabschnitten einen kontinuierlichen und einen diskreten Zugang vor, bei denen dieser Ansatz entwickelt wird.

4.5.1 Modellierung für kontinuierliche Bilder

Wenn wir voraussetzen, dass unsere Bilder hinreichend glatt sind, ist die Richtungsableitung eines Bildes x in einem Punkt s in Richtung v gegeben durch $\nabla x(s)^T v$. Wenn wir alle Richtungsableitungen in einem Punkt betrachten wollen, können wir $v(\alpha) := (\cos(\alpha), \sin(\alpha))^T$ setzen und erhalten dann abhängig von dem Winkel $0 \leq \alpha \leq \pi$ eine Richtungsableitung in Richtung $v(\alpha)$. Wir wollen nun in jedem Punkt die Richtungsableitung $\nabla x(s)^T v(\alpha)$ betrachten und daraus ein Modell entwickeln, das anisotrop glättet. Dies wird durch folgende Beobachtung motiviert: In einem Punkt, der zu einer homogenen Region oder zu einer Kante gehört, gibt es mindestens eine Richtung, in der die Richtungsableitung klein, sogar nahe 0 ist. Es gibt dann nämlich eine Richtung – zumindest die Tangentenrichtung der Isophote – und damit einen Winkel α , so dass $\nabla x^T v(\alpha)$ klein ist. In einem Ausreißer sind hingegen in der Regel alle Richtungsableitungen groß. Diese Unterscheidung zwischen Kante und homogener Region einerseits und Ausreißern andererseits genügt, um ein vernünftiges Modell zu erhalten. Wir erklären die Steuerung wieder durch $u(s) = (u_1(s), u_2(s))^T := \nabla x(s)$ und führen nun eine Steuerbeschränkung der folgenden Gestalt ein:

$$-R \leq (\cos \alpha(s), \sin \alpha(s))(u_1(s), u_2(s))^T \leq R.$$

Dabei ist der Winkel $0 \leq \alpha \leq \pi$ eine neue Optimierungsvariable, die von $s \in \Omega$ abhängt. Diese Steuerbeschränkung nutzt nun die oben geschilderte Unterscheidung zwischen Ausreißer und Kante aus, wenn wir die Schranke R klein wählen. Gehört der Punkt s nämlich zu einer Kante oder homogenen Region, so wird der Winkel α so gefunden, dass die Richtungsableitung in Richtung $v(\alpha)$ klein ist. Handelt es sich andererseits um einen Ausreißer, so existiert eine derartige Richtung nicht, und die Steuerung muss zwingend klein gewählt werden, wenn die Steuerbeschränkung eingehalten werden soll. Folglich wird der Ausreißer gedämpft, ohne dass Kanten zerstört werden. Die Steuerbeschränkung können wir auch als rotierende Maske (engl.: rotating mask) interpretieren: In jedem Punkt wird eine Rotation durchgeführt, bis die optimale Richtung und damit ein optimaler Rotationswinkel α gefunden wird. Eine (Filter-)Maske wird in dem Sinne angewandt, dass Ausreißer gedämpft werden und Kanten ungestört bleiben.

Wir erhalten also das folgende Steuerungsproblem mit „rotating mask“:

$$\begin{aligned} \text{Minimiere } F(x, u, \alpha) &= \int_{\Omega} (I(s) - S(x(s)))^2 ds + \mu \int_{\Omega} |u(s)|^p ds \\ (x, u, \alpha) &\in W_0^{1,\infty}(\Omega) \times L^\infty(\Omega, \mathbb{R}^2) \times L^\infty(\Omega), \\ u(s) &= (u_1(s), u_2(s))^T = \nabla x(s), \quad (\forall) s \in \Omega, \\ -R &\leq (\cos \alpha(s), \sin \alpha(s))(u_1(s), u_2(s))^T \leq R, \quad (\forall) s \in \Omega, \\ \alpha(s) &\in [0, a] \subset \mathbb{R}. \end{aligned}$$

Wir merken an, dass eventuell eine zusätzliche Regularisierung des Winkels im Zielfunktional sinnvoll wäre und für ein stabileres Verfahren sorgen würde. Jedoch schien diese zusätzliche Regularisierung in den Versuchsreihen des Verfassers keine Vorteile zu bringen und wird daher im Folgenden nicht betrachtet.

4.5.2 Modellierung für diskrete Bilder

Wir wollen nun aufbauend auf den Ideen des letzten Abschnitts eine Problemlösung basierend auf dem diskreten Bildmodell modellieren. Der Vorteil dieser Herangehensweise ist, dass wir anschließend ein Verfahren erhalten, das sich leicht implementieren läßt. Wir betrachten zunächst den Fall eindimensionaler Signale und anschließend den Fall zweidimensionaler Bilder.

Der eindimensionale Fall

Es sei zunächst $x \in \mathbb{R}^N$ ein diskretes Signal, das auf einem Gitter Ω_N gegeben sei. In jedem Punkt s_i im Innern des Gitters ist nun der Vorwärts- und Rückwärtsdifferenzenquotient $\delta_s^+ x_i$ bzw. $\delta_s^- x_i$ wie in Abschnitt 2.4.1 definiert. Wir können dieselbe Idee verwenden wie im kontinuierlichen Fall, indem wir den Unterschied zwischen einer Kante und einem Ausreißer feststellen. In einem Ausreißer sind sowohl der Vorwärts- als auch der Rückwärtsdifferenzenquotient groß, während bei einer Kante mindestens ein Differenzenquotient klein ist, sogar einen Wert nahe 0 annimmt. Wir betrachten nun in jedem Punkt den Vorwärts- und den Rückwärtsdifferenzenquotient gleichzeitig, indem wir die Steuerung in einem Punkt s_i wie folgt erklären:

$$u_i = \left(u_i^{(1)}, u_i^{(2)} \right)^T = \left(\delta_s^+ x_i, \delta_s^- x_i \right)^T.$$

Zur Verdeutlichung betrachten wir die Abbildung 4.3 als einfaches Beispiel für ein Signal. An einer Kante in einem Punkt s_i erhalten wir für die Steuerung $u_i = (0, 0.5)^T$ oder $u_i = (-0.5, 0)^T$, während wir an einem Ausreißer $u_i = (-0.5, 0.5)$ oder $u_i = (0.5, -0.5)$ erhalten. Wir betrachten u_i als einen Vektor im Koordinatensystem und können allgemeiner folgende Beobachtung aufstellen: In jedem Fall liegt u_i im Fall einer Kante entweder auf der negativen x -Achse des Koordinatensystems oder auf der positiven y -Achse, vgl. Abbildung 4.4a. Ein Ausreißer hingegen liegt in jedem Fall im Innern des zweiten oder vierten Quadranten. Wir wollen den Bereich $[-R_1, R_1] \times [-R_2, R_2]$ zulässigen Rotationsbereich nennen und wählen $R_2 \approx 0$. Wenn wir weiterhin R_1 groß genug wählen – in unserem Beispiel etwa $R_1 = 0.5$ – so kann jeder Steuervektor, der eine Kante repräsentiert, in den

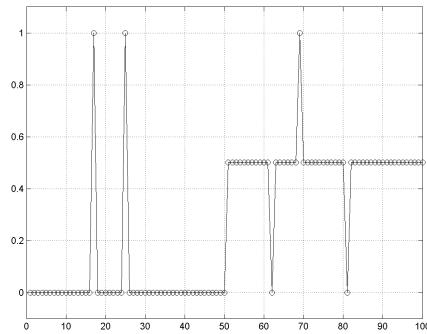
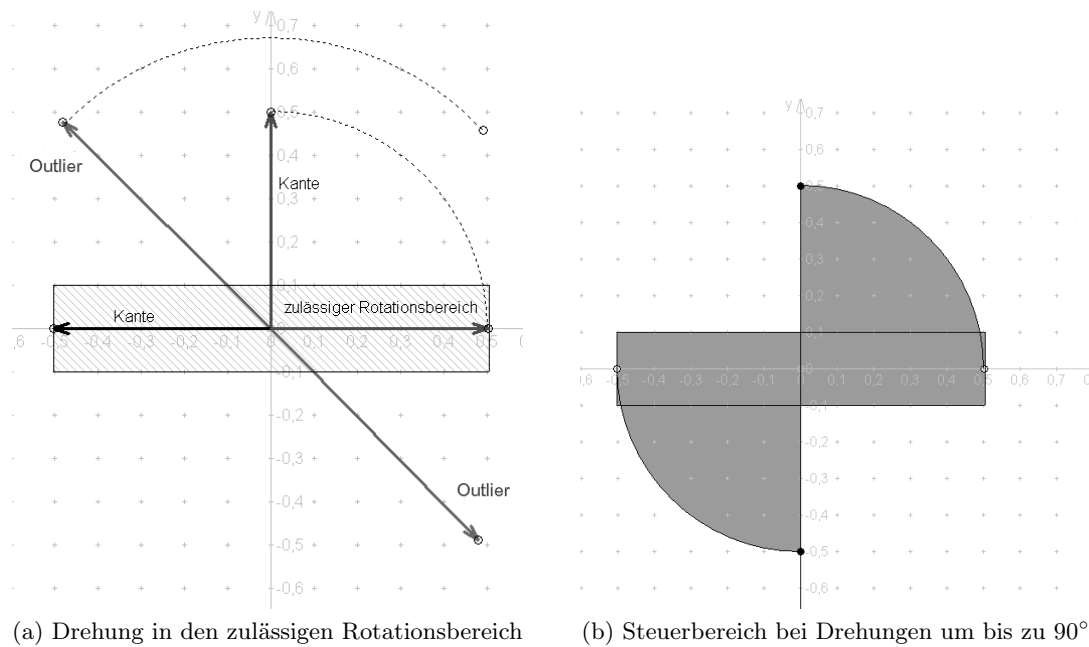


Abbildung 4.3: Signal



(a) Drehung in den zulässigen Rotationsbereich

(b) Steuerbereich bei Drehungen um bis zu 90°

Abbildung 4.4: Anwendung der Rotationsmatrix

zulässigen Rotationsbereich gedreht werden. Dazu wenden wir eine Rotationsmatrix der Gestalt

$$\begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

auf den Steuervektor an und lassen geeignete Drehwinkel α zu. Wir führen nun eine neue Optimierungsvariable $\alpha \in \mathbb{R}^N$ ein und betrachten in einem Punkt s_i die folgende Steuerbeschränkung des gedrehten Steuervektors:

$$\begin{aligned} -R_1 &\leq \cos(\alpha_i)u_i^{(1)} + \sin(\alpha_i)u_i^{(2)} \leq R_1, \\ -R_2 &\leq -\sin(\alpha_i)u_i^{(1)} + \cos(\alpha_i)u_i^{(2)} \leq R_2, \\ 0 &\leq \alpha_i \leq \frac{\pi}{2}. \end{aligned}$$

Damit diese Beschränkungen erfüllt sind, muss das Verfahren (bzw. die rotating mask) also in jedem Punkt s_i den Winkel α_i so bestimmen, dass die Steuervektoren im zulässigen Rotationsbereich liegen. Um zu vermeiden, dass Ausreißer ebenfalls in den zulässigen Rotationsbereich gedreht werden können, lassen wir nur Drehungen der Vektoren im Uhrzeigersinn um den Winkel $\leq 90^\circ$ zu. Ausreißer werden also in diesem Fall zwangsläufig gedämpft, da die Steuerung klein werden muss, um die Beschränkung noch erfüllen zu können.

Wir merken an, dass der zulässige Rotationsbereich nicht mit der Menge der zulässigen Steuerungen verwechselt werden darf. Wenn wir Drehungen um 90° zulassen, so liegen zulässige Steuervektoren nur im ersten und dritten Quadranten und in dem zulässigen Rotationsbereich selbst, vgl. Abbildung 4.4a. Wie in Abbildung 4.4b deutlich wird, ist insbesondere der Steuerbereich damit nicht konvex.

Wir schreiben das vollständige Steuerungsproblem in seiner diskreten Form nun nochmal auf. Dabei ist nun zu beachten, dass wir das Zielfunktional mittels einer Summe approximieren, sodass wir das folgende Steuerungsproblem mit rotating mask betrachten:

$$\begin{aligned}
 (P_2) \quad \text{Minimiere } F(x) &= \sum_{i \in \Omega_N} (x_i - I_i)^2 \\
 (x, u, \alpha) &\in \mathbb{R}^N \times \mathbb{R}^{(N-2) \times 2} \times \mathbb{R}^{N-2}, \\
 u_i &= \begin{pmatrix} u_i^{(1)} \\ u_i^{(2)} \end{pmatrix} = (\delta_s^+ x_i, \delta_s^- x_i), \\
 -R_1 &\leq \cos(\alpha_i) u_i^{(1)} + \sin(\alpha_i) u_i^{(2)} \leq R_1, \\
 -R_2 &\leq -\sin(\alpha_i) u_i^{(1)} + \cos(\alpha_i) u_i^{(2)} \leq R_2, \\
 0 &\leq \alpha_i \leq \frac{\pi}{2}.
 \end{aligned}$$

Eine weitere Bemerkung betrifft eine Vereinfachung dieses Modells. Da im Grunde nur Drehungen um 90° eine Rolle spielen, müssen Winkel $0 < \alpha < \frac{\pi}{2}$ nicht betrachtet werden. Deshalb lassen sich die Nebenbedingungen mit Hilfe von zwei neuen Optimierungsvariablen $y^{(1)}, y^{(2)} \in \mathbb{R}^N$ auch wie folgt formulieren:

$$\begin{aligned}
 -R_1 &\leq y_i^{(1)} u_i^{(1)} + y_i^{(2)} u_i^{(2)} \leq R_1, \\
 -R_2 &\leq y_i^{(2)} u_i^{(1)} + y_i^{(1)} u_i^{(2)} \leq R_2, \\
 0 &= y_i^{(1)} y_i^{(2)}, \\
 1 &= y_i^{(1)} + y_i^{(2)}.
 \end{aligned} \tag{4.21}$$

Es ist zu beachten, dass wir durch die beiden letzten beiden Gleichungen die Definition von zwei binären Variablen simuliert haben. $y^{(1)}, y^{(2)}$ können nämlich in jedem Punkt jeweils nur die Werte $\{0, 1\}$ annehmen.

Der zweidimensionale Fall

Wir betrachten nun Bilder $x \in \mathbb{R}^{N_1 \times N_2}$ auf dem zugehörigen Gitter Ω_{N_1, N_2} und betrachten wie im eindimensionalen Fall für jeden Punkt $s_{i,j}$ des Gitters seine Nachbarschaft, um daraus ein Kriterium für Ausreißer zu entwickeln. Wir beschränken uns dabei auf die Nachbarschaft $\mathcal{N}_{i,j}$, wie sie in Abschnitt 2.4.1 (2.7) erklärt wurde. Nun erklären wir einen

Gitterpunkt zu einem Ausreißer, wenn die Differenzen mit allen vier Punkten aus der Nachbarschaft groß sind. In diesem Fall soll der Ausreißer gedämpft werden. Hingegen gehört ein Gitterpunkt zu einer homogenen Region beziehungsweise zu einer Kante, wenn es mindestens einen Punkt in der Nachbarschaft gibt, mit dem die Differenz klein ist. Mathematisch gesehen betrachten wir bezüglich der Komponente s_1 als auch bezüglich der Komponente s_2 den Vorwärts- und Rückwärtsdifferenzenquotienten in jedem Punkt gleichzeitig. Wir führen also die folgenden vier Steuervariablen ein, die in jedem Punkt (i, j) im Innern des Gitters Ω_{N_1, N_2} wohldefiniert sind:

$$u_{i,j}^{(1)} = \delta_{s_1}^+ x_{i,j}, \quad u_{i,j}^{(2)} = \delta_{s_2}^+ x_{i,j}, \quad u_{i,j}^{(3)} = \delta_{s_1}^- x_{i,j}, \quad u_{i,j}^{(4)} = \delta_{s_2}^- x_{i,j}.$$

Dass Ausreißer gedämpft werden und Kanten ungestört bleiben, erreichen wir nun, indem wir mit einer zusätzlichen Optimierungsvariablen $y = (y^{(1)}, y^{(2)}, y^{(3)}, y^{(4)}) \in \mathbb{R}^{N_1 \times N_2 \times 4}$ die Steuerbeschränkungen (bzw. die rotating mask) für jeden Punkt (i, j) im Innern des Gitters wie folgt modellieren:

$$\begin{aligned} -R &\leq y_{i,j}^{(1)} u_{i,j}^{(1)} \leq R, \\ -R &\leq y_{i,j}^{(2)} u_{i,j}^{(2)} \leq R, \\ -R &\leq y_{i,j}^{(3)} u_{i,j}^{(3)} \leq R, \\ -R &\leq y_{i,j}^{(4)} u_{i,j}^{(4)} \leq R, \\ \sum_{k=1}^4 y_{i,j}^{(k)} &\geq 1. \end{aligned}$$

Wir wollen kurz deutlich machen, wieso diese Steuerbeschränkungen eine Dämpfung von Ausreißern bewirken. Die letzte Nebenbedingung garantiert, dass zumindest eine der Steuervariable $u_{i,j}^{(k)}, 1 \leq k \leq 4$ nicht mit zu kleinen Zahlen ≈ 0 multipliziert wird. Wir wählen nun die Schranke R nahe 0 und zwingen – wegen der letzten Nebenbedingung – mindestens eine Steuerungskomponente $u_{i,j}^{(k)}$, einen kleinen Wert anzunehmen. Ist der Punkt $s_{i,j}$ nun etwa Teil einer horizontal oder vertikal verlaufenden Kante, so wird für mindestens ein k die zugehörige Steuervariable $u_{i,j}^{(k)}$ klein, und das zugehörige $y_{i,j}^{(k)}$ kann so groß gewählt werden, dass die letzte Nebenbedingung erfüllt ist. Damit bleiben also Kanten unbeeinflusst. Ist umgekehrt $s_{i,j}$ ein Ausreißer, so wird dieser zwangsweise gedämpft, da es ja mindestens eine Steuerungskomponente $u_{i,j}^{(k)}$ geben muss, die nahe 0 ist.

Wir schreiben das vollständige Steuerungsproblem mit rotating mask in seiner diskreten Form auf:

$$\begin{aligned} (P_3) \quad \text{Minimiere } F(x) &= \sum_{(i,j) \in \Omega_{N_1, N_2}} (x_{i,j} - I_{i,j})^2 \\ (x, u, y) &\in \mathbb{R}^{N_1 \times N_2} \times \mathbb{R}^{(N_1-2) \times (N_2-2) \times 4} \times \mathbb{R}^{(N_1-2) \times (N_2-2) \times 4}, \\ u_{i,j}^{(1)} &= \delta_{s_1}^+ x_{i,j}, \quad u_{i,j}^{(2)} = \delta_{s_2}^+ x_{i,j}, \\ u_{i,j}^{(3)} &= \delta_{s_1}^- x_{i,j}, \quad u_{i,j}^{(4)} = \delta_{s_2}^- x_{i,j}, \end{aligned}$$

$$\begin{aligned}
 -R &\leq y_{i,j}^{(1)} u_{i,j}^{(1)} \leq R, \\
 -R &\leq y_{i,j}^{(2)} u_{i,j}^{(2)} \leq R, \\
 -R &\leq y_{i,j}^{(3)} u_{i,j}^{(3)} \leq R, \\
 -R &\leq y_{i,j}^{(4)} u_{i,j}^{(4)} \leq R, \\
 \sum_{k=1}^4 y_{i,j}^{(k)} &\geq 1.
 \end{aligned}$$

Wir merken an dieser Stelle an, dass diagonale Kanten mit der Kantenbreite von einem Pixel nach dieser Modellierung als Ausreißer erkannt werden würden und damit verändert werden. Dies ließe sich beheben, indem man etwa alle acht anliegenden Gitterpunkte zu der Nachbarschaft zählen würde. Eine weitere Einführung von Steuervariablen scheint jedoch nicht sinnvoll, da zu viele Steuervariablen das Verfahren zu langsam machen würden. In diesem Fall müßte man einen Vorverarbeitungsschritt durchführen, der aus den acht anliegenden Gitterpunkten eine bessere Approximation des Gradienten bestimmen würde. Es ist möglich, das Steuerungsproblem (P_2) für eindimensionale Signale auf zweidimensionale Bilder zu übertragen. Dies können wir erreichen, indem wir die vier Steuervariablen jeweils miteinander kombinieren. Insbesondere betrachten wir die folgenden Nebenbedingungen:

$$\begin{aligned}
 -R_1 &\leq \cos(\alpha_{i,j})(u_{i,j}^{(1)} + u_{i,j}^{(3)}) + \sin(\alpha_{i,j})(u_{i,j}^{(2)} + u_{i,j}^{(4)}) \leq R_1, \\
 -R_2 &\leq -\sin(\alpha_{i,j})(u_{i,j}^{(1)} + u_{i,j}^{(3)}) + \cos(\alpha_{i,j})(u_{i,j}^{(2)} + u_{i,j}^{(4)}) \leq R_2.
 \end{aligned}$$

Diese Formulierung führt zu Schwierigkeiten, wenn das Bild ein Quadrat enthält. Dann werden unabhängig von der Kombination der Steuervariablen in zwei Eckpunkten jeweils zwei Steuervariablen groß, und man kann sich überlegen, dass die Steuerung in diesen Eckpunkten zwangsläufig gedämpft wird. Der Eckpunkt wird also als Ausreißer behandelt. In unseren Versuchsreihen konnten wir bei Bildern, die keine expliziten Quadrate enthalten, mit diesen Nebenbedingungen jedoch gute Ergebnisse erzielen.

5 Numerische Verfahren – Minimierungsmethoden

Zur numerischen Lösung von Variationsproblemen und Problemen der optimalen Steuerung unterscheidet man direkte und indirekte Verfahren:

- Bei den direkten Verfahren diskretisiert man das Problem zunächst und erhält so ein endlichdimensionales Optimierungsproblem. Der Vorteil dieser Methode besteht darin, dass man zur Lösung dieses endlichdimensionalen Problems nun einen Solver heranziehen kann. Mit Hilfe dieses Ansatzes lassen sich unsere Optimierungsprobleme also meist schnell und leicht implementieren.
- Bei den indirekten Verfahren wendet man im ersten Schritt die notwendigen Optimalitätsbedingungen für eine Minimalstelle an. Für ein entsprechendes Problem sind dies zum Beispiel die Euler-Lagrange-Gleichungen aus Kapitel 3, Satz 3.5. Auf diese Weise erhält man ein System von partiellen Differentialgleichungen und diskretisiert dieses erst jetzt. Hängt das Zielfunktional quadratisch von der Optimierungsvariable ab, so erfüllt das Minimum ein System von linearen Gleichungen. Das diskrete Problem wird nun mit Hilfe eines iterativen Verfahrens gelöst.

In diesem Kapitel nutzen wir indirekte und direkte Verfahren, um für einige Optimierungsprobleme, die im letzten Kapitel vorgestellt wurden, explizit zu zeigen, wie sie numerisch gelöst werden. So leiten wir für das Kantenerkennungsfunktional mit Hilfe des indirekten Verfahrens ein iteratives Schema her. Wir geben anschließend an, wie Steuerungsprobleme geeignet diskretisiert werden, und erläutern dann, wie sie mit Hilfe der Solver IPOPT und LOQO, die auf sogenannten Innere-Punkt-Verfahren basieren, gelöst werden.

5.1 Diskretisierungsschemata

In diesem Abschnitt wollen wir einige grundlegende Bezeichnungen einführen, die wir benötigen werden, wenn wir unsere diskreten Verfahren herleiten.

Sowohl bei den direkten als auch bei den indirekten Verfahren arbeiten wir mit diskreten Bildern $x \in \mathbb{R}^{N_1 \times N_2}$. Der Definitionsbereich Ω der Bilder wird also als äquidistantes $(N_1 \times N_2)$ -Gitter Ω_{N_1, N_2} aufgefasst. Die Gitterpunkte sollen dabei sowohl in der Breite als auch in der Länge einen Abstand h haben, den wir bei Bildern immer gleich 1 setzen. Wir definieren zunächst Indextmengen, die wir in diesem Kapitel benötigen werden:

$$\begin{aligned} D(\Omega) &:= \{1, \dots, N_1 - 1\} \times \{1, \dots, N_2 - 1\}, \\ D(\Gamma) &:= (\{1\} \times \{1, \dots, N_2\}) \cup (\{N_1\} \times \{1, \dots, N_2\}) \cup \\ &\quad (\{1, \dots, N_1\} \times \{1\}) \cup (\{1, \dots, N_1\} \times \{N_2\}), \\ DIN(\Omega) &:= \Omega_{N_1, N_2} / D(\Gamma). \end{aligned}$$

Da der Definitionsbereich somit eine einfache Gestalt hat, ist es möglich, die auftretenden Differentialgleichungen mittels finiter Differenzen zu approximieren.

Diskretisierung des Laplace-Operators

In Abschnitt 2.4.1 haben wir bereits die Diskretisierung der ersten Ableitung mittels finiter Differenzen angegeben. Nun wollen wir noch die zweite Ableitung einer Funktion approximieren, um anschließend eine Diskretisierung des Laplace-Operators zu erreichen. Ist $f : \mathbb{R} \rightarrow \mathbb{R}$ eine hinreichend glatte Funktion, so gilt mittels Taylorentwicklung in den Punkten $x_{j+1}, x_{j-1} \in \mathbb{R}$

$$\begin{aligned} f(x_{j+1}) &= f(x_j) + f'(x_j)h + \frac{1}{2}f''(x_j)h^2 + \frac{1}{6}f'''(x_j)h^3 + \mathcal{O}(h^4), \\ f(x_{j-1}) &= f(x_j) - f'(x_j)h + \frac{1}{2}f''(x_j)h^2 - \frac{1}{6}f'''(x_j)h^3 + \mathcal{O}(h^4). \end{aligned}$$

Addition der beiden Gleichungen ergibt

$$\begin{aligned} f(x_{j+1}) + f(x_{j-1}) &= 2f(x_j) + f''(x_j)h^2 + \mathcal{O}(h^4) \\ \Rightarrow f''(x_j) &= \frac{f(x_{j+1}) - 2f(x_j) + f(x_{j-1}))}{h^2} + \mathcal{O}(h^4). \end{aligned}$$

Wenn wir dies auf den zweidimensionalen Fall übertragen, erhalten wir für $h = 1$ und die Funktion $x : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ eine Approximation des Laplace-Operators im Gitterpunkt $s_{i,j} \in \Omega_{N_1, N_2}$:

$$\Delta x|_{i,j} \approx x_{i+1,j} + x_{i-1,j} + x_{i,j+1} + x_{i,j-1} - 4x_{i,j}.$$

Diese Diskretisierung nennen wir auch Fünf-Punkte-Stern. Wir führen nun die folgende Bezeichnung ein:

$$\bar{x}_{i,j} = x_{i+1,j} + x_{i-1,j} + x_{i,j+1} + x_{i,j-1}.$$

Dann läßt sich der Laplace-Operator schreiben als:

$$\Delta x|_{i,j} \approx \bar{x}_{i,j} - 4x_{i,j}. \quad (5.1)$$

Eine genauere Approximation des Laplace-Operators stellt der Neun-Punkte-Stern dar (vgl. [3], pp. 254), der zudem im Gegensatz zum Fünf-Punkte-Stern invariant unter Rotationen ist. Allerdings haben wir bereits mit dem Fünf-Punkte-Stern in dieser Arbeit gute Ergebnisse erzielen können, sodass eine Erweiterung als nicht notwendig angesehen wurde.

5.2 Indirekte Verfahren

In der gängigen Literatur zur Bildverarbeitung werden die in Kapitel 4 vorgestellten Energiefunktionale numerisch häufig mit Hilfe des indirekten Verfahrens minimiert. Bei diesem „first-optimize-then-discretize“-Ansatz wendet man zuerst notwendige Optimalitätsbedingungen an und diskretisiert anschließend das erhaltene System von partiellen Differentialgleichungen.

Wir betrachten dazu für $x : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit $x \in C^2(\Omega)$ das Standard-Energiefunktional der Variationsrechnung:

$$F(x) = \int_{\Omega} f(s, x(s), Jx(s)) ds \quad (5.2)$$

Durch Anwendung der Euler-Lagrange-Gleichungen erhalten wir (vgl. Satz (3.5)) im Fall $m = 1$ die partielle Differentialgleichung

$$F'(x) = 0. \tag{5.3}$$

Um diese Gleichung numerisch zu lösen, können wir ein Gradientenabstiegsverfahren oder eine Fixpunktiteration anwenden. Wir betrachten im Folgenden diese beiden numerischen Ansätze genauer.

5.2.1 Gradientenabstiegsverfahren

Motiviert wird dieser erste Ansatz durch das aus der endlichdimensionalen Optimierung bekannte Gradientenabstiegsverfahren. Um eine Lösung der Gleichung (5.3) zu erhalten, betrachten wir für eine Funktion $F \in C^1(\mathbb{R}^n, \mathbb{R})$ die folgende Differentialgleichung, die wir durch künstliche Einführung einer Zeitvariable t erhalten und **Gradientenfluss** nennen wollen:

$$\begin{aligned} \frac{d}{dt}x(t) &= -\nabla F(x(t)), \quad t > 0, \\ x(0) &= x_0, \quad t = 0. \end{aligned}$$

Die Optimierungsvariable $x(t)$ hängt also von der Zeit ab, und wir können zeigen, dass mit fortschreitender Zeit die Funktionswerte von F abnehmen, bis die Funktion an einem kritischen Punkt ankommt, d.h. bis $\nabla F(x(t)) = 0$ gilt. Für Zeitpunkte $t_1 > t_2 \geq 0$ gilt nämlich

$$F(x(t_1)) \geq F(x(t_2)).$$

Das sieht man so:

$$\begin{aligned} F(x(t_2)) - F(x(t_1)) &= (F \circ x)(t_2) - (F \circ x)(t_1) = \int_{t_1}^{t_2} (F \circ x)'(t) dt \\ &= \int_{t_1}^{t_2} \nabla F(x(t)) \frac{d}{dt}x(t) dt = - \int_{t_1}^{t_2} \nabla F(x(t)) \nabla F(x(t)) dt \\ &= - \int_{t_1}^{t_2} \|\nabla F(x(t))\|^2 dt \leq 0. \end{aligned}$$

Für $\nabla F(x(t_1)) \neq 0$ gilt die Ungleichung sogar strikt, und F fällt streng monoton. Wir betrachten nun für das Standard-Energiefunktional (5.2) analog Funktionen $x(s, t)$, die nun zusätzlich von der Zeit abhängen, und somit den Gradientenfluss

$$\begin{aligned} \frac{\partial}{\partial t}x(s, t) &= -F'(x(s, t)), \\ x(s, 0) &= x_0(s), \quad t = 0. \end{aligned}$$

Dann gilt für das Energiefunktional

$$\frac{\partial}{\partial t} (F(x(s, t))) = F'(x(s, t)) \frac{\partial}{\partial t}x(s, t) = - \left\| \frac{\partial}{\partial t}x(s, t) \right\|^2 \leq 0,$$

und es ist deshalb zu erwarten, dass F mit der Zeit abnimmt. Außerdem erhalten wir $\frac{\partial}{\partial t} (F(x(s, t))) = 0$, wenn $\frac{\partial}{\partial t}x(s, t) = 0$, sowie $\frac{\partial}{\partial t}x(s, t) = 0$, wenn $F'(x(s, t)) = 0$, d.h.

wenn wir an einem stationären Punkt angekommen sind. Nun können wir ein explizites Schema verwenden, um eine Minimierung von F numerisch zu erreichen:

$$\frac{x^{n+1} - x^n}{\tau^n} = -F'(x^n) \quad \Rightarrow \quad x^{n+1} = x^n - \tau^n F'(x^n), \quad (5.4)$$

wobei die **Schrittweite** $\tau^n > 0$ einem Zeitschritt entspricht. Dabei ist zu beachten, dass die τ^n in jedem Iterationsschritt geeignet gewählt werden müssen, um in jedem Schritt wirklich einen Abstieg zu erreichen. Dies erreichen wir etwa mit der Schrittweitensteuerung von Armijo-Goldstein (vgl. [1], pp. 89). Das Verfahren (5.4) heißt **Gradientenabstiegsverfahren**, da es in jedem Iterationsschritt einen Abstieg in Richtung des Gradienten vornimmt.

Beispiel 2. Wir leiten das Gradientenabstiegsverfahren für das Zielfunktional bestehend aus L^2 -Datenterm und quadratischer Regularisierung her. Wir nehmen an, dass S die Identität ist und betrachten also das folgende Funktional, das wir minimieren wollen:

$$F(x) = \int_{\Omega} (x(s) - I(s))^2 ds + \mu \int_{\Omega} |\nabla x(s)|^2 ds.$$

Wie wir in Abschnitt 4.1.4 gesehen haben, liefert die Anwendung der Euler-Lagrange-Gleichung in diesem Fall die elliptische Differentialgleichung:

$$0 = -\mu \Delta x(s) + x(s) - I(s).$$

Wir führen nun eine künstliche Zeit t ein und erhalten den Gradientenfluss, welcher hier folglich eine parabolische Differentialgleichung darstellt:

$$\frac{\partial}{\partial t} x(s, t) = \mu \Delta x(s, t) - (x(s, t) - I).$$

Durch Diskretisierung, und indem wir für den Laplace-Operator den Fünf-Punkte-Stern (5.1) verwenden, erhalten wir folgendes explizites Schema:

$$x_{i,j}^{n+1} = x_{i,j}^n - \tau^n (\mu (\bar{x}_{i,j}^n - 4x_{i,j}^n) - (x_{i,j}^n - I_{i,j})) \quad \text{für } (i, j) \in DIN(\Omega).$$

Wenn wir die Dirichletrandbedingungen betrachten, können wir zum Beispiel annehmen, dass das Bild einen schwarzen Rand hat, indem wir

$$x_{i,j} = 0 \quad \text{für } (i, j) \in D(\Gamma)$$

fordern.

5.2.2 Fixpunktiteration

Da es sich beim Gradientenabstiegsverfahren um ein explizites Verfahren handelt, muss die CFL-Bedingung (Courant-Friedrichs-Lewy-Bedingung) eingehalten werden, um ein stabiles Verfahren zu gewährleisten (vgl. [10], pp. 12). Dies kann jedoch insgesamt zu einem langsamen Verfahren führen. Eine Alternative zum Gradientenabstiegsverfahren wurde deshalb von Vogel und Oman in [29] vorgeschlagen.

Wir betrachten dazu den Variationsansatz mit TV-Regularisierung und gehen von der Euler-Lagrange-Gleichung aus, die in diesem Fall wie folgt lautet (vgl. (4.10), wobei S nun die Identität sei):

$$0 = -\mu \operatorname{div} \left(\frac{\nabla x(s)}{|\nabla x(s)|} \right) + (x(s) - I(s)) \quad \text{mit } s \in \Omega. \quad (5.5)$$

Das Energiefunktional ist in Regionen mit $|\nabla x| = 0$ nicht differenzierbar, und wir verwenden die modifizierte L^1 -Norm als Approximation, um numerische Schwierigkeiten für den Fall $|\nabla x| \approx 0$ zu vermeiden:

$$|\nabla x|_\eta := \sqrt{|\nabla x|^2 + \eta}, \quad \eta > 0. \quad (5.6)$$

Für $|\nabla x| \ll \sqrt{\eta}$ erhalten wir $\sqrt{|\nabla x|^2 + \eta} \approx \sqrt{\eta} \left(1 + \frac{|\nabla x|^2}{2\eta}\right)$ und für $|\nabla x| \gg \sqrt{\eta}$ gilt $\sqrt{|\nabla x|^2 + \eta} \approx |\nabla x|$ (vgl. Bemerkung 4.2). Das bedeutet, dass wir in glatten Regionen eine quadratische Regularisierung erhalten und bei Kanten eine lineare Funktion. Dabei spielt $\sqrt{\eta}$ die Rolle eines Grenzparameters.

Mit (5.6) läßt sich (5.5) in der Gestalt

$$x + \mu L(x)x = I \quad (5.7)$$

darstellen, wobei der elliptische Differentialoperator $L(x)$ durch

$$L(x)y = -\operatorname{div} \left(\frac{\nabla y}{|\nabla x|_\eta} \right)$$

definiert ist. Aus der nichtlinearen Gleichung (5.7) erhalten wir nun eine Fixpunktiteration, indem wir die Gleichung wie folgt linearisieren:

$$(1 + \mu L(x^n))x^{n+1} = I, \quad x^0 \text{ gegeben.} \quad (5.8)$$

In jeder Iteration ist so nur noch eine lineare Diffusionsgleichung zu lösen. Wenn wir diese diskretisieren, erhalten wir ein hochdimensionales lineares Gleichungssystem, welches mit dem vorkonditioniertem konjugierten Gradientenverfahren („PCG-Verfahren“) gelöst werden kann (vgl. [28], pp. 103). Im Gegensatz zum Gradientenabstiegsverfahren besitzt dieses Verfahren nun folgende Vorteile: Eine Schrittweitenbestimmung in jedem Iterationsschritt ist nicht nötig, und das Verfahren ist unabhängig von der CFL-Bedingung. In [19], pp. 7 ff. wurde gezeigt, dass die Diskretisierung der Iteration (5.8) für jeden Startwert x^0 global konvergent ist.

5.2.3 Herleitung eines iterativen Schemas für das Kantenerkennungsfunktional

Wir wollen für das Kantenerkennungsfunktional von Ambrosio-Tortorelli (4.12) das Minimum numerisch mit Hilfe einer indirekten Methode bestimmen. Dies bedeutet, dass wir zunächst die Euler-Lagrange-Gleichung anwenden und in diesem Fall nun ein System bestehend aus zwei partiellen Differentialgleichungen erhalten. Durch Diskretisierung leiten wir dann ein iteratives Schema zur Lösung des Problems her. Wir betrachten also das Kantenerkennungsfunktional (4.12):

$$F_\epsilon(x, k) = \int_\Omega c_1 (x(s) - I(s))^2 + c_2 |\nabla x(s)|^2 \cdot (k(s)^2 + c_4) + c_3 \left(\epsilon |\nabla k(s)|^2 + \frac{1}{4\epsilon} (1 - k(s))^2 \right) ds. \quad (5.9)$$

Dieses Funktional soll bzgl. x und k minimiert werden. Aufgrund des Produkts $|\nabla x|^2 k^2$ ist das Funktional nicht konvex und somit nicht mehr gewährleistet, dass ein numerisches Verfahren gegen ein Minimum konvergiert. Wenn wir hingegen k fest lassen, so hängt F_ϵ quadratisch von x ab und wie in Satz 4.1 folgt die Konvexität von F_ϵ bzgl. x . Lassen wir wiederum x fest, so ist das Funktional quadratisch und konvex bzgl. k . Um nun ein numerisches Verfahren zu erhalten, lassen wir also jeweils x und k fest und bestimmen zunächst die Euler-Lagrange-Gleichung von F_ϵ bzgl. x und anschließend bzgl. k . Um die folgenden Gleichungen abzukürzen, bezeichnen wir den Integranden von (5.9) mit f . Wir lassen zunächst k fest und nehmen an, dass $x \in C^2(\Omega)$. Dann können wir Satz 3.5 aus Kapitel 3 anwenden und erhalten

$$\begin{aligned}\frac{\partial f}{\partial x} &= c_1 2(x - I), \\ f_{x_{s_1}} &= c_2 2 \frac{\partial x}{\partial s_1} (k^2 + c_4), \\ f_{x_{s_2}} &= c_2 2 \frac{\partial x}{\partial s_2} (k^2 + c_4), \\ \Rightarrow \operatorname{div} (f_{x_{s_1}}, f_{x_{s_2}}) &= c_2 2 \left(\left(\frac{\partial^2 x}{\partial s_1^2} \right) + \left(\frac{\partial^2 x}{\partial s_2^2} \right) \right) (k^2 + c_4).\end{aligned}$$

Also ist die Euler-Lagrange-Gleichung eine elliptische Differentialgleichung:

$$c_1(x - I) - c_2 \Delta x (k^2 + c_4) = 0. \quad (5.10)$$

Wir bestimmen die Euler-Lagrange-Gleichungen für $k \in C^2(\Omega)$ analog, indem wir x fest lassen:

$$\begin{aligned}\frac{\partial f}{\partial k} &= 2c_2 k |\nabla x|^2 - c_3 \frac{1}{4\epsilon} 2(1 - k), \\ f_{k_{s_1}} &= c_3 \epsilon 2 \left(\frac{\partial k}{\partial s_1} \right), \\ f_{k_{s_2}} &= c_3 \epsilon 2 \left(\frac{\partial k}{\partial s_2} \right) \\ \Rightarrow \operatorname{div} (f_{k_{s_1}}, f_{k_{s_2}}) &= 2 \epsilon c_3 \left(\left(\frac{\partial^2 k}{\partial s_1^2} \right) + \left(\frac{\partial^2 k}{\partial s_2^2} \right) \right).\end{aligned}$$

Somit erhalten wir

$$c_2 k |\nabla x|^2 - c_3 \frac{1}{4\epsilon} (1 - k) - \epsilon c_3 \Delta k = 0. \quad (5.11)$$

Mit der Definition der beiden elliptischen Differentialoperatoren

$$L(k) = -\frac{c_2}{c_1} (k^2 + c_4) \Delta \quad \text{und} \quad L(x) = \frac{c_2 4\epsilon}{c_3} |\nabla x|^2 - 4\epsilon^2 \Delta,$$

lassen sich die Gleichungen (5.10) und (5.11) darstellen als

$$(1 + L(k))x = I \quad \text{und} \quad (1 + L(x))k = 1,$$

wobei 1 die Identität bezeichnet. Es liegt also ein gekoppeltes System zweier elliptischer Differentialgleichungen vor, das wir iterativ durch das folgende alternierende Schema lösen können: $x^n \rightarrow k^n \rightarrow x^{n+1}$, d.h. durch die beiden Gleichungen

$$(1 + L(x^n))k^n = 1 \quad \text{und} \quad (1 + L(k^n))x^{n+1} = I. \quad (5.12)$$

Wir können alternativ auch ein paralleles Schema $x^n \rightarrow k^{n+1}, k^n \rightarrow x^{n+1}$ verwenden:

$$(1 + L(x^n))k^{n+1} = 1 \quad \text{und} \quad (1 + L(k^n))x^{n+1} = I.$$

Nun müsste man – ähnlich wie bei der Fixpunktiteration (5.8) – in jedem Iterationsschritt nach der Diskretisierung zwei lineare Gleichungssysteme lösen. Wir leiten jedoch ein explizites Verfahren her und betrachten dazu die partiellen Differentialgleichungen aus (5.12) separat. Zunächst betrachten wir

$$\left(1 + \frac{c_2 4\epsilon}{c_3} |\nabla x|^2 - 4\epsilon^2 \Delta\right) k^n = 1.$$

Indem wir die Approximation des Laplace-Operators (5.1) in der Gestalt $\Delta k^n \approx \bar{k}^n - 4k^{n-1}$ verwenden, erhalten wir in jedem Punkt $(i, j) \in DIN(\Omega)$ das folgende explizite Schema:

$$\begin{aligned} & \left(1 + \frac{c_2 4\epsilon}{c_3} \left((\delta_{s_1}^+ x_{i,j}^n)^2 + (\delta_{s_2}^+ x_{i,j}^n)^2 \right) + 16\epsilon^2\right) k_{i,j}^n = 1 + 4\epsilon^2 \bar{k}_{i,j}^{n-1} \\ \Leftrightarrow & k_{i,j}^n = \frac{1 + 4\epsilon^2 \bar{k}_{i,j}^{n-1}}{\left(1 + \frac{c_2 4\epsilon}{c_3} \left((\delta_{s_1}^+ x_{i,j}^n)^2 + (\delta_{s_2}^+ x_{i,j}^n)^2 \right) + 16\epsilon^2\right)}. \end{aligned}$$

Analog leiten wir für

$$\left(1 - \frac{c_2}{c_1} (k^2 + c_4) \Delta\right) x^{n+1} = I$$

mit der Approximation $\Delta x^{n+1} \approx \bar{x}^n - 4x^{n+1}$ das folgende explizite Schema in jedem Punkt $(i, j) \in DIN(\Omega)$ her:

$$\begin{aligned} & \left(1 + \frac{4c_2}{c_1} ((k_{i,j}^n)^2 + c_4)\right) x_{i,j}^{n+1} = I + \frac{c_2}{c_1} ((k_{i,j}^n)^2 + c_4) \bar{x}_{i,j}^n \\ \Leftrightarrow & x_{i,j}^{n+1} = \frac{I + \frac{c_2}{c_1} ((k_{i,j}^n)^2 + c_4) \bar{x}_{i,j}^n}{1 + \frac{4c_2}{c_1} ((k_{i,j}^n)^2 + c_4)}. \end{aligned}$$

Auch hier müssen noch die Randwerte der Gitterpunkte auf dem Rand sowohl für x als auch für k angegeben werden. Der Einfachheit halber können wir auch hier die Dirichlet-randbedingung fordern, also

$$x_{i,j} = 0 \quad \text{für} \quad (i, j) \in D(\Gamma) \quad \text{und} \quad k_{i,j} = 0 \quad \text{für} \quad (i, j) \in D(\Gamma)$$

Ein theoretisches Konvergenzresultat ist nach Kenntnisstand des Verfassers nicht vorhanden, jedoch liefert das Verfahren in der Praxis und bei den in dieser Arbeit durchgeführten Versuchen bereits nach wenigen Iterationen gute Ergebnisse.

5.3 Direkte Verfahren

Die Idee der direkten Verfahren besteht darin, ein Steuerungsproblem durch Diskretisierung in ein endlichdimensionales nichtlineares Optimierungsproblem zu transformieren und auf dieses Problem anschließend ein numerisches Verfahren (Innere-Punkt-Methode) anzuwenden, die in Abschnitt 5.5 behandelt wird.

Wir betrachten als Beispiel das Steuerungsproblem (P_1) aus Abschnitt 4.4, wobei S nun die Identität sei und wir außerdem $f(r) = r^2$ annehmen:

$$\begin{aligned} \text{Minimiere } F(x, u) &= \int_{\Omega} (I(s) - x(s))^2 ds + \mu \int_{\Omega} |u(s)|^2 ds; \quad (x, u) \in W_0^{1,p}(\Omega) \times L^p(\Omega, \mathbb{R}^2) \\ \text{unter } u(s) &= (u_1(s), u_2(s))^T = \nabla x(s), \quad (\forall) s \in \Omega, \\ u(s) &\in K := \{(v_1, v_2) \in \mathbb{R}^2 \mid |v_1|^q + |v_2|^q \leq R^q\} \quad (\forall) s \in \Omega. \end{aligned} \tag{5.13}$$

Wie bei den indirekten Verfahren diskretisieren wir das Gebiet Ω durch ein Gitter Ω_{N_1, N_2} . Haben die Gitterpunkte jeweils äquidistanten Abstand h , so lässt sich das Zielfunktional mittels einer einfachen Kubaturformel approximieren. Eine solche ließe sich zum Beispiel mit der Simpson-Formel herleiten, doch verwenden wir hier der Einfachheit halber für eine Funktion $g : \Omega \rightarrow \mathbb{R}$ die zweidimensionale Trapez-Kubaturformel:

$$\int_{\Omega} g(s) ds \approx \sum_{(i,j) \in D(\Omega)} \frac{h^2}{4} (g(s_{i,j}) + g(s_{i,j+1}) + g(s_{i+1,j}) + g(s_{i+1,j+1})).$$

Nachdem wir das Zielfunktional mit Hilfe dieser Formel diskretisiert haben, diskretisieren wir die Nebenbedingungen, indem wir die einzelnen partiellen Ableitungen durch Finite-Differenzen approximieren. Für das Problem (5.13) erhalten wir also insgesamt mit den Optimierungsvariablen $(x, u) \in \mathbb{R}^{N_1 \times N_2} \times \mathbb{R}^{(N_1-1) \times (N_2-1) \times 2}$ das folgende endlichdimensionale Optimierungsproblem mit Nebenbedingungen:

$$\begin{aligned} \text{Minimiere } F(x, u) &\approx \sum_{(i,j) \in \Omega_{N_1, N_2}} h^2 (I_{i,j} - x_{i,j})^2 + \mu \sum_{(i,j) \in D(\Omega)} h^2 \left((u_{i,j}^{(1)})^2 + (u_{i,j}^{(2)})^2 \right) \\ u_{i,j}^{(1)} &= \delta_{s_1}^+ x_{i,j}, \quad (i, j) \in D(\Omega) \\ u_{i,j}^{(2)} &= \delta_{s_2}^+ x_{i,j}, \quad (i, j) \in D(\Omega) \\ -R^q &\leq |u_{i,j}^{(1)}|^q + |u_{i,j}^{(2)}|^q \leq R^q, \quad (i, j) \in D(\Omega) \end{aligned}$$

Dieses Optimierungsproblem können wir nun mit Hilfe eines geeigneten Solvers lösen. Wir verwenden dabei Solver, die intern sogenannte Innere-Punkt-Verfahren benutzen (vgl. Abschnitt 5.5).

An dieser Stelle wollen wir noch exemplarisch darauf eingehen, wie einige der in Kapitel 4 vorgestellten Optimierungsprobleme geeignet diskretisiert werden.

Sind allgemeinere Regularisierungsterme als die quadratische Regularisierung gegeben, wie zum Beispiel die TV-Regularisierung, oder das Regularisierungsfunktional (4.16) aus der Approximation des Mumford-Shah-Funktionals, so lässt sich auf analoge Weise das Zielfunktional mittels der Quadraturformel diskretisieren. Zu beachten ist nur, dass die Regularisierung jeweils geeignet durch Finite-Differenzen approximiert wird.

Als Beispiel betrachten wir dazu das Zielfunktional (4.14) bestehend aus L^1 -Datenterm und TV-Regularisierungsterm genauer. Wir nehmen dabei an, dass $x \in W^{1,2}(\Omega)$, und betrachten wegen (2.5) also das folgende Modell:

$$F(x) = \int_{\Omega} |I(s) - x(s)| ds + \lambda \int_{\Omega} |\nabla x(s)| ds$$

Nun ist weder der Datenterm, noch der Regularisierungsterm im Nullpunkt differenzierbar. Die einfachste Möglichkeit besteht nun darin, beide Terme mit Hilfe der modifizierten L^1 -Norm zu approximieren. Mit zwei Parametern $\beta > 0, \delta > 0$ erhalten wir also

$$F(x) \approx \int_{\Omega} \sqrt{(I(s) - x(s))^2 + \beta} ds + \lambda \int_{\Omega} \sqrt{|\nabla x(s)|^2 + \delta} ds.$$

Auf diese Weise erhalten wir sogar ein strikt konvexes Funktional, das dann – im Gegensatz zu dem Funktional (4.14) – nun eine eindeutige Minimalstelle besitzt (vgl. auch [13] pp.21). Durch Anwendung der Trapez-Kubaturformel erhalten wir nun die Approximation:

$$F(x) \approx \sum_{(i,j) \in \Omega_{N_1, N_2}} h^2 \sqrt{(I_{i,j} - x_{i,j})^2 + \beta} + \lambda \sum_{(i,j) \in D(\Omega)} h^2 \sqrt{((\delta_{s_1}^+ x_{i,j})^2 + (\delta_{s_2}^+ x_{i,j})^2)^2 + \delta}.$$

Die in Abschnitt 4.5.2 modellierten Steuerungsprobleme mit rotating mask (vgl. die Probleme (P_2) und (P_3)) sind bereits in diskreter Form gegeben, sodass eine direkte Anwendung des Solvers möglich ist.

5.4 Nichtlineare Optimierungsprobleme – Notwendige Optimalitätsbedingungen

Um Innere-Punkt-Verfahren herleiten zu können, wiederholen wir in diesem Abschnitt einige Grundlagen der nichtlinearen Optimierung und formulieren insbesondere notwendige Optimalitätsbedingungen für ein nichtlineares Optimierungsproblem mit Nebenbedingungen. Seien $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine differenzierbare Funktion und $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ stetig differenzierbar. Dann hat das Standardproblem der nichtlinearen Optimierung die Gestalt

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{unter} \quad & g_i(x) \leq 0 \quad i = 1, \dots, k, \\ & g_i(x) = 0 \quad i = k + 1, \dots, m. \end{aligned} \tag{5.14}$$

Ein Punkt $\bar{x} \in \mathbb{R}^n$ heißt **zulässig**, wenn er die Nebenbedingungen aus (5.14) erfüllt, d.h. wenn gilt:

$$\begin{aligned} g_i(\bar{x}) &\leq 0 \quad i = 1, \dots, k, \\ g_i(\bar{x}) &= 0 \quad i = k + 1, \dots, m. \end{aligned}$$

Definition 5.1 (Indexmengen):

Für einen zulässigen Punkt \bar{x} definieren wir die Menge der aktiven Indizes der Ungleichungen $I(\bar{x})$ und die Menge der Indizes der Gleichungen $J(\bar{x})$ wie folgt:

$$\begin{aligned} I(\bar{x}) &:= \{i \in \{1, \dots, k\} \mid g_i(\bar{x}) = 0\}, \\ J(\bar{x}) &:= I(\bar{x}) \cup \{k + 1, \dots, m\}. \end{aligned}$$

Definition 5.2 (Lagrange-Funktion):

Die Lagrange-Funktion $\mathcal{L} : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$ für das Standardproblem (5.14) ist definiert durch

$$\mathcal{L}(x, \lambda_0, \lambda) := \lambda_0 f(x) + \lambda g(x),$$

wobei $(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m$ Lagrange-Multiplikator genannt wird.

Nun können wir notwendige Optimalitätsbedingungen erster Ordnung für das Standardproblem der nichtlinearen Optimierung wie folgt formulieren:

Satz 5.1 (Notwendige Optimalitätsbedingungen von Karush-Kuhn-Tucker):

Sei x^* lokale Minimalstelle des Standardproblems der nichtlinearen Optimierung (5.14). Dann gilt:

1. Es gibt einen Zeilenvektor $(\lambda_0, \lambda) \in \mathbb{R} \times \mathbb{R}^m \neq (0, 0_m)$ mit

$$L_x(x^*, \lambda_0, \lambda) = \lambda_0 f'(x^*) + \sum_{j \in J} \lambda_j g'_j(x^*) = 0,$$

$$\lambda_0 \geq 0, \lambda_j \geq 0 \quad \forall i \in I(x^*), \lambda_i = 0 \text{ für } i \notin J(x^*).$$

2. Ist x^* regulär, d.h.

- a) die Gradienten $(g_{k+1})_x(x^*), \dots, (g_m)_x(x^*)$ sind linear unabhängig,
- b) $\exists v \in \mathbb{R}^n$ sodass $(g_i)_x(x^*)v < 0$ für $i \in J(x^*)$ und $(g_i)_x(x^*)v = 0$ für $i \in \{k+1, \dots, m\}$,

dann kann $\lambda_0 = 1$ gesetzt werden.

3. Ist x^* normal, d.h., sind die Gradienten $(g_i)_x(x^*)$, $i \in J(x^*)$ linear unabhängig, so ist λ mit $\lambda_0 = 1$ eindeutig bestimmt

Beweis. Der Beweis findet sich in [21], Kapitel 2, Satz (4.16). □

5.5 Innere-Punkt-Verfahren – IPOPT und LOQO

Eine Möglichkeit der Lösung nichtlinearer Optimierungsprobleme mit Nebenbedingungen bieten Innere-Punkt-Verfahren. Die Idee besteht darin, zur Zielfunktion einen Barriere-term zu addieren, so dass Verletzungen der Nebenbedingung stark bestraft werden. Wir betrachten dazu das Optimierungsproblem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{unter} \quad & g(x) \leq 0. \end{aligned}$$

Dabei seien die Zielfunktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ und die Nebenbedingung $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ zweimal stetig differenzierbar. Wir führen nun Schlupfvariablen $y \in \mathbb{R}^m$ ein und erhalten ein äquivalentes Optimierungsproblem:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{unter} \quad & g(x) - y = 0, \\ & y \leq 0. \end{aligned} \tag{5.15}$$

Um die Ungleichungsrestriktionen zu eliminieren, subtrahieren wir im Zielfunktional einen logarithmischen Barriereterm und erhalten:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) - \mu \sum_{i=1}^m \ln(y_i) \\ \text{unter} \quad & g(x) - y = 0. \end{aligned} \tag{5.16}$$

Dabei ist $\mu > 0$ vorausgesetzt. Nun sieht man leicht, dass für $y \rightarrow 0$ die Zielfunktion in (5.16) beliebig große Werte annimmt. Deswegen muss ein Minimum im Innern der zulässigen Menge des Problems (5.15) liegen. Das Verfahren heißt deshalb Innere-Punkt-Verfahren. Der Einfluß des Barriereterms $-\mu \sum_{i=1}^n \ln(y_i)$ auf die Lösung hängt vom Parameter μ ab, so dass wir für $\mu \rightarrow 0$ erwarten, dass die Lösung des Problems (5.16) unter gewissen Voraussetzungen gegen die Lösung des ursprünglichen Problem (5.15) konvergiert. Nun betrachtet man die Lagrange-Funktion von (5.16):

$$\mathcal{L}(x, y, \lambda) = f(x) - \mu \sum_{i=1}^m \ln(y_i) - \sum_{i=1}^m \lambda_i (g_i(x) - y_i(x)).$$

Dabei ist $\lambda = (\lambda_1, \dots, \lambda_m)$ der Lagrange-Multiplikator für die Ungleichungs-Nebenbedingung. Anschließend kann man die notwendigen Optimalitätsbedingungen von Karush-Kuhn-Tucker aus Satz 5.1 auswerten. Es bezeichne $Y = \text{diag}(y_1, \dots, y_m)$ die Diagonalmatrix mit den Schlupfvariablen y_i auf der Hauptdiagonale und $e = (1, \dots, 1)^T \in \mathbb{R}^m$, dann gilt:

$$\begin{aligned} \mathcal{L}(x, y, \lambda)_x &= \nabla f(x) + g'(x)\lambda = 0, \\ \mathcal{L}(x, y, \lambda)_y &= \mu Y^{-1}e + \lambda = 0, \\ \mathcal{L}(x, y, \lambda)_\lambda &= -g(x) + y = 0. \end{aligned}$$

Wir multiplizieren die zweite Gleichung nun mit Y und erhalten:

$$\begin{aligned} \nabla f(x) + g'(x)\lambda &= 0, \\ \mu e + Y\lambda &= 0, \\ -g(x) + y &= 0. \end{aligned}$$

Wir fassen die linke Seite als Funktion $F(x, y, \lambda)$ auf und können das Barriereproblem für einen festen Parameter μ_j nun mit Hilfe des Newton-Verfahrens lösen:

$$F'(x, y, \lambda) \Delta(x, y, \lambda) = -F(x, y, \lambda).$$

Wir führen dazu die folgenden Bezeichnungen ein:

$$\begin{aligned} H^k &:= \mathcal{L}_{xx} = f''(x^k) - \sum_{i=1}^m g_i''(x^k) \lambda_i^k, \\ \Lambda^k &:= \text{diag}(\lambda_1^k, \dots, \lambda_m^k), \quad Y^k := \text{diag}(y_1^k, \dots, y_m^k). \end{aligned}$$

In jedem Iterationsschritt k ist dann ein nichtlineares Gleichungssystem zu lösen:

$$\begin{bmatrix} H_k & 0 & -g'(x^k)^T \\ 0 & \Lambda^k & Y^k \\ -g'(x^k) & E & 0 \end{bmatrix} \begin{pmatrix} d_k^x \\ d_k^\lambda \\ d_k^z \end{pmatrix} = - \begin{pmatrix} \nabla f(x^k) + g'(x^k)\lambda^k \\ \mu e + Y\lambda^k \\ -g(x^k) + w^k \end{pmatrix}$$

Hat man die Suchrichtung durch das Lösen dieses Gleichungssystems berechnet und geeignete Schrittweiten $\alpha_k \in (0, 1]$ bestimmt, kann man den nächsten Iterationspunkt mittels

$$\begin{aligned} x_{k+1} &:= x_k + \alpha_k d_k^x, \\ y_{k+1} &:= y_k + \alpha_k d_k^y, \\ \lambda_{k+1} &:= \lambda_k + \alpha_k d_k^\lambda \end{aligned}$$

berechnen.

Bei der Durchführung der Versuchsreihen in dieser Arbeit haben wir den Solver IPOPT verwendet und manchmal zum Vergleich LOQO herangezogen. Beide Solver basieren auf dem Innere-Punkt-Verfahren und wurden entwickelt, um hochdimensionale nichtlineare Optimierungsprobleme mit Nebenbedingungen zu lösen. Für Ausführungen und Details zu dem von A. Wächter entwickelten Solver IPOPT verweisen wir auf [30]. Für Details zum Solver LOQO, der von R.J. Vanderbei entwickelt wurde, sei auf [27] verwiesen.

Wir haben das AMPL-Interface (A Modelling Language for Programming) benutzt, um unsere Probleme in eine geeignete Modellierungssprache zu übersetzen und anschließend an die Solver übergeben zu können. Mit Hilfe von AMPL konnten wir so unsere Optimierungsprobleme in eine leicht verständliche Programmiersprache übersetzen. Eine Einführung in AMPL leistet [20].

Die Programmiersprache MATLAB haben wir verwendet, um die Matrizen, mit denen die Solver intern rechnen, als Bilder geeignet darzustellen. Auf der MATLAB-Homepage <http://www.mathworks.de> findet sich die Dokumentation der verwendeten MATLAB-Anweisungen. Außerdem nutzen wir MATLAB, um die indirekten Verfahren zu programmieren.

6 Numerische Resultate

In diesem Kapitel sollen die meisten Methoden, die in Kapitel 4 vorgestellt und entwickelt wurden, anhand von Beispielen und Testbildern untersucht und verglichen werden. Um den Überblick über die unterschiedlichen Methoden und Ziele nicht zu verlieren, fassen wir diese in dem folgenden Schema kurz zusammen und widmen uns in den jeweiligen Abschnitten den entsprechenden Details.

- I. Kantenerkennungsfunktional von Ambrosio-Tortorelli
Ziel: Bestimmung der optimalen Parameter für ein synthetisches Testbild mit einfachen geometrischen Formen, Anwendung auf verrauschtes Testbild, gute Entrauschung und gute Kantenbilder → Abschnitt 6.2

- II. Energiefunktionale mit quadratischer und TV-Regularisierung
Ziel: Entrauschen des verrauschten Testbildes → Abschnitt 6.3

- III. Steuerungsproblem mit L^p - und insbesondere L^∞ -Steuerbeschränkung
Ziel: Entrauschen des verrauschten Testbildes und Kantenerkennung mit Hilfe der Steuerung, Anwendung mit und ohne zusätzliche Regularisierung im Zielfunktional, Vergleich mit der Methode aus I. → Abschnitt 6.4

- IV. Steuerungsproblem mit rotating mask
Ziel: Entrauschen von Bildern mit Salt-and-Pepper-Rauschen, Vergleich mit aktuellen Methoden aus der Forschung, Anwendung auf synthetische Bilder und auf das „Lena“-Testbild
→ Abschnitt 6.5

Abbildung 6.1: Überblick über die Durchführung der Versuchsreihen

Im letzten Abschnitt versuchen wir dann, die Ergebnisse dieses Kapitels zusammenzufassen und die unterschiedlichen Methoden zu bewerten.

6.1 Bewertungsmaßstäbe für die Ergebnisse

Neben der subjektiven Bewertung eines Bildes ist es oft wichtig, eine quantitative Maßzahl zu definieren, um die Güte eines wiederhergestellten Bildes bewerten zu können. Wir definieren dazu:

Definition 6.1 (SNR und PSNR):

Sei $x \in \mathbb{R}^{n \times n}$ das Originalbild und $I \in \mathbb{R}^{n \times n}$ das gestörte bzw. mit einer Methode wiederhergestellte Bild. Die **Signal-Noise-Ratio SNR** (deutsch: *Signal-Rausch-Verhältnis*) und die

Peak-Signal-Noise-Ratio PSNR sind dann definiert als

$$SNR := 10 \log_{10} \left(\frac{\sum_{i=1}^n \sum_{j=1}^n x_{i,j}^2}{\sum_{i=1}^n \sum_{j=1}^n (x_{i,j} - I_{i,j})^2} \right), \quad (6.1)$$

$$PSNR := 10 \log_{10} \left(\frac{\max_{i,j} x_{i,j}^2}{\sum_{i=1}^n \sum_{j=1}^n (x_{i,j} - I_{i,j})^2} \right). \quad (6.2)$$

Für kontinuierliche Bilder $x : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ beschreibt $\|x - I\|_2$ den quadratischen Fehler der Differenz $x - I$. Das SNR bzw. das PSNR sind für diesen Fall definiert durch:

$$SNR := 10 \log_{10} \left(\frac{\|x\|_2}{\|x - I\|_2} \right), \quad PSNR := 10 \log_{10} \left(\frac{\|x\|_\infty}{\|x - I\|_2} \right).$$

Das SNR wird in der Einheit dB (deciBel) angegeben. Die Zusammenhänge zwischen Signal-Noise-Ratio und subjektiver Wahrnehmung sollen hier nicht weiter erläutert werden. Generell kann gesagt werden, dass die Bilder umso besser aussehen, je größer das PSNR bzw. das SNR ist. Wenn das gestörte Bild I dem ursprünglichen Bild x nämlich „nahe“ kommt, so werden die Nenner in (6.1) bzw. (6.2) klein und folglich das SNR bzw. PSNR groß.

Bei der Bewertung der wiederhergestellten Bilder bestimmen wir das SNR oder das PSNR des wiederhergestellten Bildes im Verhältnis zum ursprünglichen Bild.

Bei der Bewertung der Kantenbilder, die wir erhalten, gehen wir wie folgt vor:

1. Für ein einfaches synthetisches Bild, das nur Rechtecke oder Kreise enthält, lässt sich ein ideales Kantenbild analytisch bestimmen. Anschließend müssen diese idealen Kanten „verpixelt“ werden, d.h. wir projizieren sie auf ein diskretes Gitter, um sie quantitativ vergleichen zu können. Wir erstellen dieses ideale diskrete Kantenbild \bar{k} und nutzen es zum Vergleich mit den Kantenbildern, die wir durch unsere Verfahren erhalten.
2. Die Kantenbilder k , die wir durch unsere numerischen Verfahren erhalten, sind in diskreter Darstellung gegeben. Wir bestimmen die Güte des Kantenbildes, indem wir die L^2 -Norm der Differenz $k - \bar{k}$ bestimmen und damit eine quantitative Maßzahl für seine Abweichung zum idealen Kantenbild \bar{k} erhalten. Das Kantenbild k ist dann also umso besser, je kleiner dieser Wert ist.
3. Dabei vergleichen wir das ideale Kantenbild \bar{k} sowohl mit den Kantenbildern, die wir durch das Kantenerkennungsfunktional erhalten, als auch mit den Kantenbildern aus der Auswertung der optimalen Steuerung.
4. Außerdem vergleichen wir die Kantenbilder, die wir aus der optimalen Steuerung gewinnen, mit dem besten Kantenbild aus der Auswertung des Kantenerkennungsfunktional.

In dieser Arbeit kann nur eine gewisse Anzahl von Bildern aufgenommen werden. Bei der Auswahl der Bilder war es dem Verfasser wichtig, dass zum einen die quantitativen Zahlen, wie PSNR oder die L^2 -Normdifferenz, gute Werte liefern. Andererseits wurden die Bilder

so ausgewählt, dass sie die Aussagen, die in dieser Arbeit gemacht werden, belegen können. Manchmal kommt es dabei zu einer Diskrepanz zwischen der qualitativen Aussage und dem quantitativen Wert. Zum Beispiel ist dies etwa bei der Kantenbewertung der Fall. Die quantitative Bewertungszahl hängt nämlich immer davon ab, wie wir gerade die idealen Kanten „verpixeln“. Auf diese Weise kann es zu der folgenden Diskrepanz kommen: Die Werte einer Methode können schlecht sein, obwohl das Kantenbild subjektiv gesehen gut ist. Dabei ist natürlich immer die Frage, was man unter „gut“ versteht. Zum Beispiel kann es wünschenswert sein, dickere Kanten zu erhalten, um diese in einem Bild hervorzuheben. Dies würde jedoch bei geeigneter Verpixelung der Idealkanten zu einem schlechten quantitativen Wert führen. Die quantitativen Bewertungen sind also immer abhängig von der zugrunde liegenden Art der Projektion der Idealkanten auf ein diskretes Gitter. Der Verfasser hat deshalb bei der Bilderauswahl versucht, einen Ausgleich dieser Diskrepanz zu erreichen, und gleichzeitig zu erklären, welche Faktoren bei der Bilderauswahl eine Rolle gespielt haben.

Um den Lesefluss nicht zu stören, werden die Tabellen zu den Versuchsreihen mit den jeweiligen quantitativen Bewertungen im Anhang dargestellt. Es wird jedoch in den Abschnitten jeweils auf die Tabellen verwiesen.

6.2 Kantenerkennung

Wir wollen die Approximation des Mumford-Shah-Funktional nach Ambrosio/Tortorelli und die damit verbundene Kantenerkennung testen. Wir schreiben das Funktional aus Abschnitt 4.2 der Übersichtlichkeit halber noch einmal auf:

$$\begin{aligned}
 F_\epsilon(x, k) = & c_1 \int_{\Omega} (x(s) - I(s))^2 ds + c_2 \int_{\Omega} |\nabla x(s)|^2 \cdot (k(s)^2 + c_4) ds \\
 & + c_3 \int_{\Omega} \left(\epsilon |\nabla k(s)|^2 + \frac{1}{4\epsilon} (k(s) - 1)^2 \right) ds.
 \end{aligned} \tag{6.3}$$

Wir benutzen das iterative Schema, das wir in Abschnitt 5.2.3 hergeleitet haben, zur numerischen Lösung. Um unsere Verfahren angemessen vergleichen und bewerten zu können, betrachten wir an dieser Stelle nur synthetische Bilder. Bei synthetischen Bildern können wir die Kanten des Bildes leicht selbst definieren und so am besten beurteilen, wie gut die Kanten tatsächlich erkannt werden. Wir betrachten also das synthetische Bild der Größe 100×100 aus Abbildung 6.2a und eine davon mit Gauss'schem Rauschen verrauschte Version (vgl. Abb. 6.2b). Da die Grauwertübergänge in diesem Bild sehr abrupt sind, ist das ideale analytische Kantenbild leicht zu bestimmen. Wir erklären die idealen Kanten in diesem Fall einfach als die Objektgrenzen. Diese sind, da die Grauwertübergänge eindeutig sind, durch diese Vereinbarung eindeutig bestimmt. Dies ist bei natürlichen Bildern nicht unbedingt der Fall. Dort stellt sich die Frage, wann ein Sprung der Grauwertintensität groß genug ist, um ihn als Kante zu identifizieren (vgl. die Argumentation in Abschnitt 4.2). Um diese idealen Kanten unseres synthetischen Bildes jetzt zu verpixeln, gibt es mehrere Möglichkeiten. Wir entscheiden uns dafür, jedem Pixel oberhalb und unterhalb einer idealen Kante den Wert 0, also schwarz, zuzuweisen. Alle anderen Pixel erhalten den Wert 1, d.h. weiß. Auf diese Weise erhalten wir das ideale Kantenbild \bar{k} aus Abbildung 6.2c, das wir als Referenzbild nutzen. Wir machen an dieser Stelle nochmal auf die Diskrepanz zur subjektiven Bewertung aufmerksam. Bei der Verpixelung wäre es auch möglich, jeweils nur ein anstatt zwei Pixel „schwarz zu malen“, zum Beispiel jeweils das Pixel oberhalb bzw.

rechts von der idealen Kante. Dies würde im Übrigen vielleicht den numerischen Verfahren besser entsprechen, die ja Finite-Differenzen benutzen und sich jeweils bei geeigneter Parameterwahl für ein und nicht zwei Pixel „entscheiden“. Auf diese Weise würden wir im Vergleich zum Kantenbild aus Abbildung 6.2c dünnere Kanten erhalten, was gegebenenfalls auch wünschenswert sein kann. Ein Nachteil der gewählten Verpixelung ist nämlich, dass aus dem Kantenbild für das menschliche Auge nicht unbedingt hervorgeht, dass es sich bei den Kanten rechts unten im Bild 6.2c um Kanten von dünnen schwarzen Linien handelt. Hier wären dünnere Kanten wünschenswert.

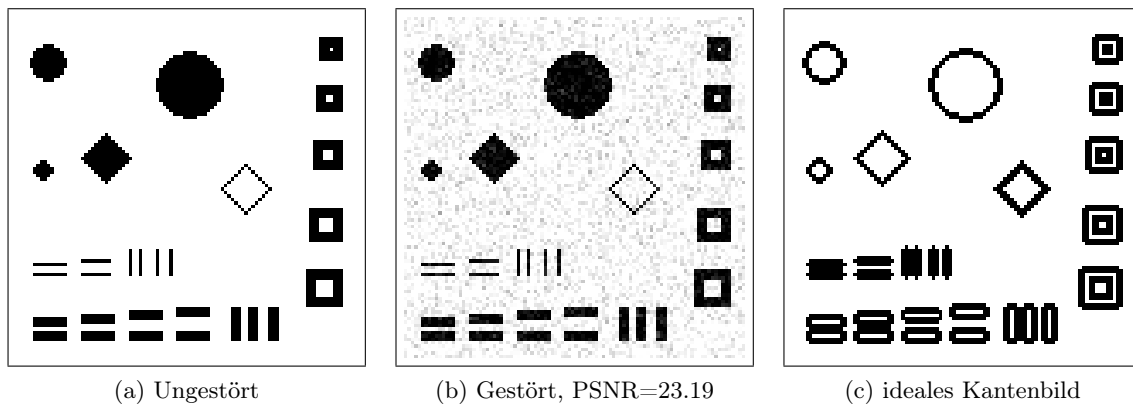


Abbildung 6.2: Synthetisches Testbild

Die Werte der Versuchsreihe finden sich in Tabelle B.1. Wir beschreiben nun das Vorgehen in unserer ersten Versuchsreihe:

Zunächst normieren wir in der ganzen Versuchsreihe $c_1 = 1$ und variieren nur die anderen Parameter, was an dem Ergebnis nichts ändert. Wir lassen nun in den Versuchen jeweils alle bis auf einen Parameter fest, und variieren diesen Parameter, um herauszufinden, welche Wirkung eine Variation hat.

Zuerst variieren wir c_2 . Die Ergebnisse finden sich in Abbildung 6.3. In der linken Spalte ist das Ergebnisbild zu sehen, rechts das zugehörige Kantenbild. Das Verfahren verhält sich so, wie wir es erwarten. Wird c_2 größer, so wird der Gradient des Bildes x mehr gewichtet, das Bild also geglättet. Damit ist gleichzeitig verbunden, dass das Kantenbild k auch eher dazu neigt, Werte nahe 0 anzunehmen. Die Kanten werden also breiter. Besonders gut ist die Diffusion des Bildes an der Raute im rechten Teil des Bildes zu beobachten. Mit der Diffusion der Objekte im Bild geht natürlich eine Verschlechterung der Bildqualität einher, und deshalb ist es nicht verwunderlich, wenn die PSNR-Werte sinken (vgl. Tabelle B.1).

In einem zweiten Schritt untersuchen wir die Auswirkungen des Parameters c_3 . Die Ergebnisse finden sich in Abbildung 6.4. Wie wir vermuten, wirkt sich die Variation von c_3 auf die Kanten aus, während sie auf das Bild selbst kaum Auswirkungen hat. Die Kanten werden dabei für kleinere c_3 dicker und intensiver. Die Begründung ist auch hier einfach: Für große c_3 wird der dritte Summand in (6.3) stärker bestraft und das Kantenbild k neigt somit dazu, Werte nahe 1 anzunehmen. Dasselbe Argument gilt bei der Variation von ϵ , wie in Abbildung 6.6 zu sehen ist. Dabei sei daran erinnert, dass ϵ ursprünglich die Konvergenz der Approximation von Ambrosio-Tortorelli bestimmt. Dass die Kanten bei kleiner werdendem ϵ dünner werden, deckt sich insofern mit der Theorie, als ja die Kantenmenge gerade gegen die Menge der Unstetigkeiten von x konvergiert.

Zuletzt betrachten wir den Parameter c_4 . Bei den bisherigen Versuchen haben wir c_4 Null gelassen. Die Betrachtung von (6.3) gibt Anlass zu der Vermutung, dass c_4 lediglich eine zusätzliche Glättung unseres Bildes x bewirkt. Dies wird in Abbildung 6.5 bestätigt. Eine derartige zusätzliche Glättung kann von Vorteil sein, wenn es sich um ein verrauschtes Bild handelt, das wir in unserem nächsten Schritt betrachten wollen.

Wir testen das Kantenerkennungsfunktional nun im Zusammenhang mit einem verrauschten Bild. Das Ziel ist dabei, sowohl gute Kanten als auch ein wiederhergestelltes Bild ohne Rauschen zu erhalten. Die benutzten Parameterpaare und Bewertungen finden sich in Tabelle B.2. Da die Auswirkungen der Parameter detailliert erklärt wurden, reicht es an dieser Stelle aus, wenn wir nur die „besten“ Ergebnisse in Abbildung 6.7 angeben. Die Parameterpaare haben wir dabei so gewählt, dass sie gleichzeitig sowohl ein gutes wiederhergestelltes, als auch ein gutes Kantenbild geliefert haben. Zu bemerken ist, dass der Parameter c_2 nicht zu klein gewählt werden darf, da das Rauschen sonst nicht geglättet wird. c_3 kann anschließend so gewählt werden, dass wir ein möglichst gutes Kantenbild erhalten.

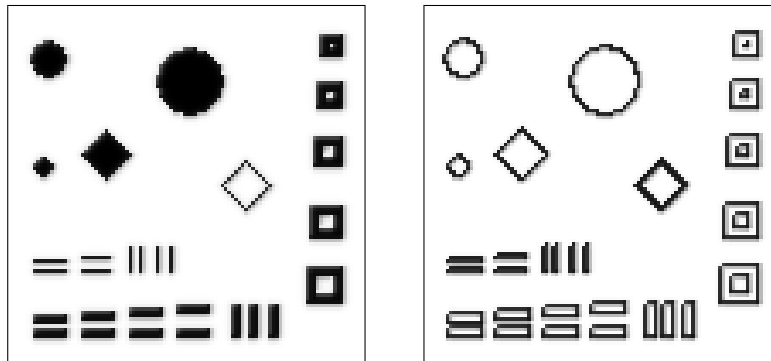
In Abbildung 6.7(b) ist das Rauschen zufriedenstellend entfernt worden, wobei sich der PSNR-Wert des Bildes nicht verbessert hat (vgl. Tabelle B.2). Dies hängt damit zusammen, dass zwar die Beseitigung des Rauschens die Qualität des Bildes verbessert, eine geringe Diffusion der Objekte aber bereits so sehr ins Gewicht fällt, dass die quantitativen PSNR-Werte insgesamt schlechter werden. Wir können jedoch festhalten, dass wir mit Hilfe des Ambrosio-Tortorelli-Funktional Bilder entrauschen können und gleichzeitig gute Kantenbilder erhalten.

Als optimales Kantenbild mit Hilfe dieses Ansatzes legen wir das Kantenbild aus Abbildung 6.3b fest, da dieses quantitativ (im Vergleich zum festgelegten Idealbild) den besten Wert geliefert hat.

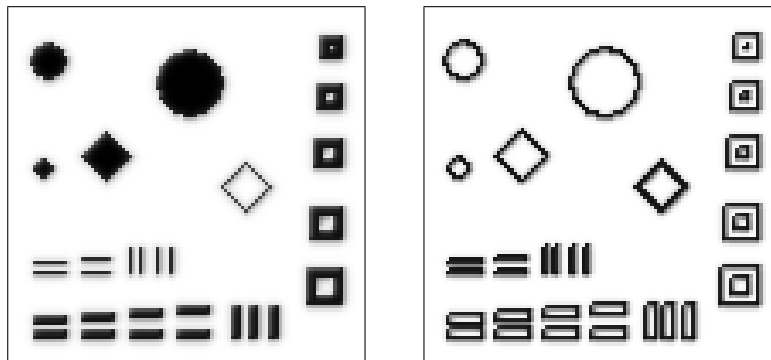
Es ist anzumerken, dass die Minimierung des Ambrosio-Tortorelli-Funktional in den Experimenten auch mit Hilfe der direkten Verfahren vorgenommen wurde und so auch sehr ähnliche Ergebnisse erzielt wurden. Auch der Nachteil der direkten Verfahren wurde jedoch deutlich: in der Regel war die Laufzeit länger. Die Ergebnisse finden sich auf der beigefügten CD.

Abbildung 6.3: Variiere c_2 : c_2 wird von oben nach unten größer. Andere Parameter bleiben fest. $c_1 = 1$, $c_3 = 0,001$, $c_4 = 0$, $\epsilon = 0,05$

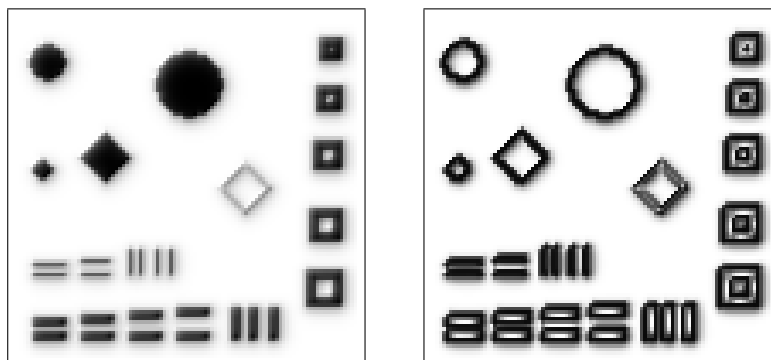
(a) $c_2 = 0.05$



(b) $c_2 = 0.1$



(c) $c_2 = 0.5$



(d) $c_2 = 1$

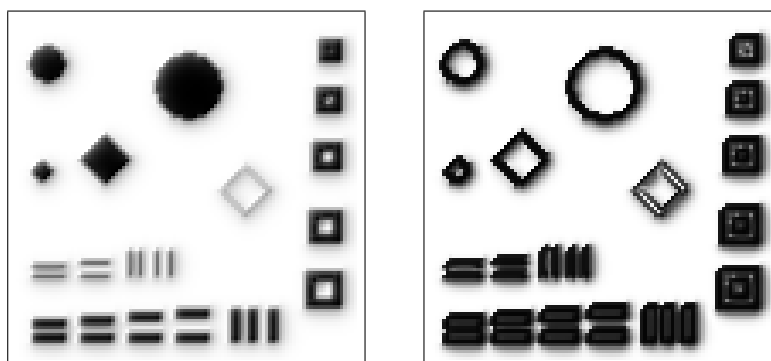


Abbildung 6.4: Variiere c_3 : c_3 wird von oben nach unten kleiner. Andere Parameter bleiben fest. $c_1 = 1$, $c_2 = 0,001$, $c_4 = 0$, $\epsilon = 0,05$

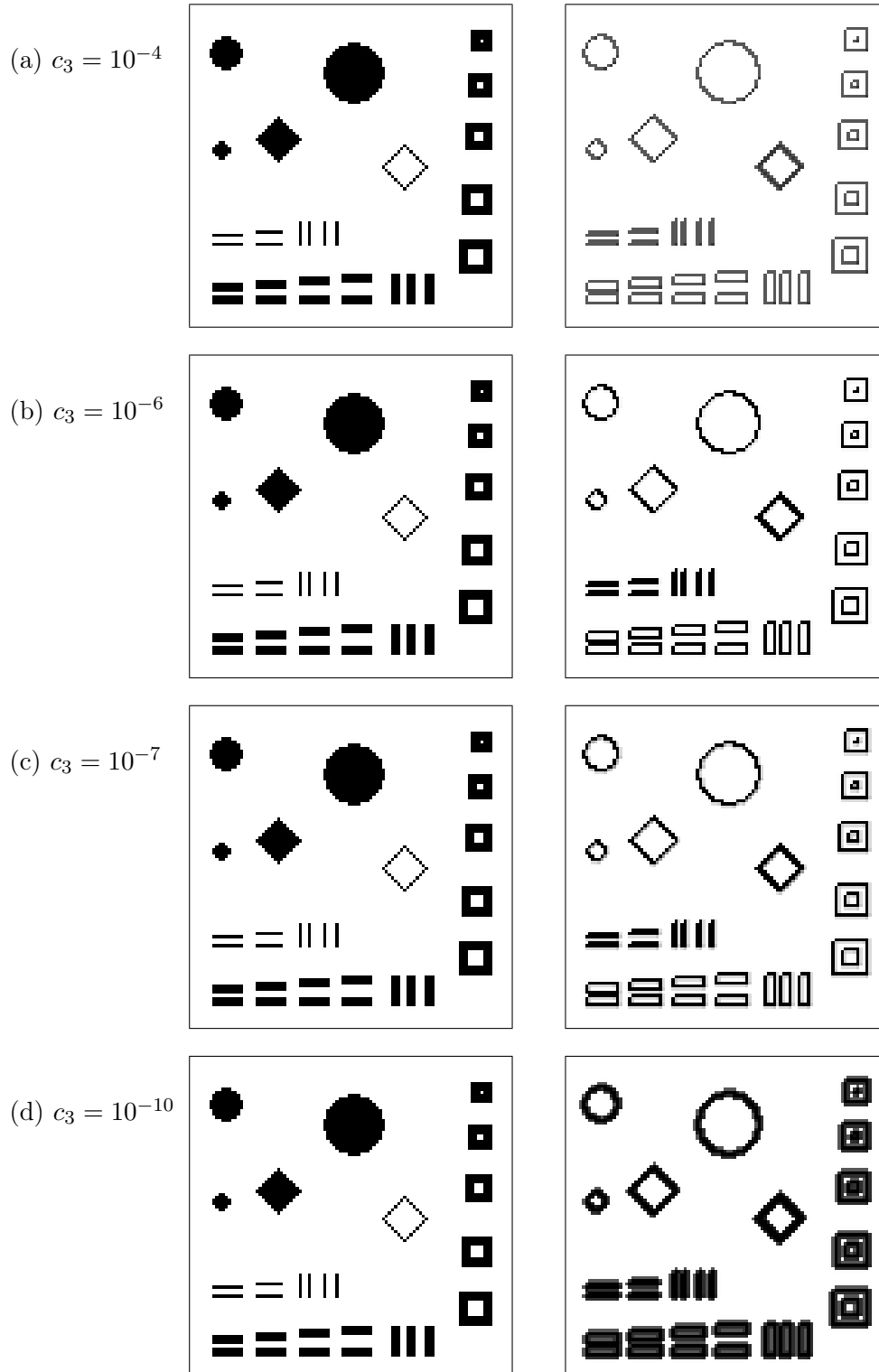
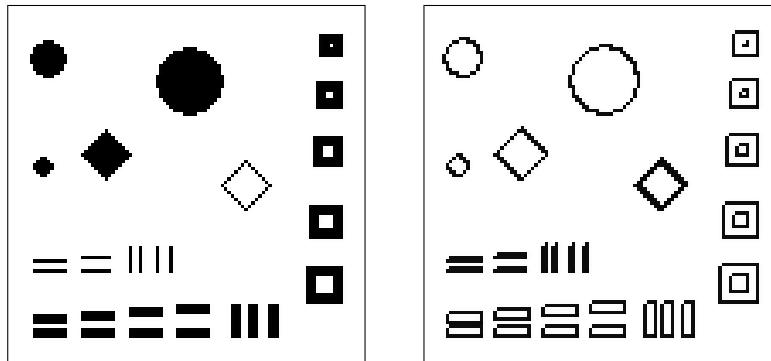
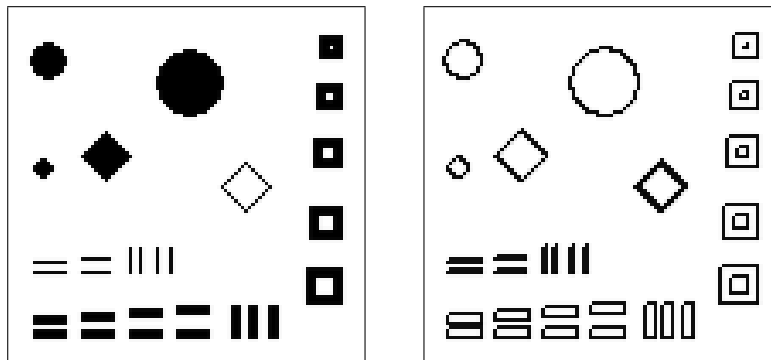


Abbildung 6.5: Variiere c_4 : c_4 wird von oben nach unten größer. Andere Parameter bleiben fest. $c_1 = 1$, $c_2 = 0,001$, $c_3 = 0,001$, $\epsilon = 0,05$

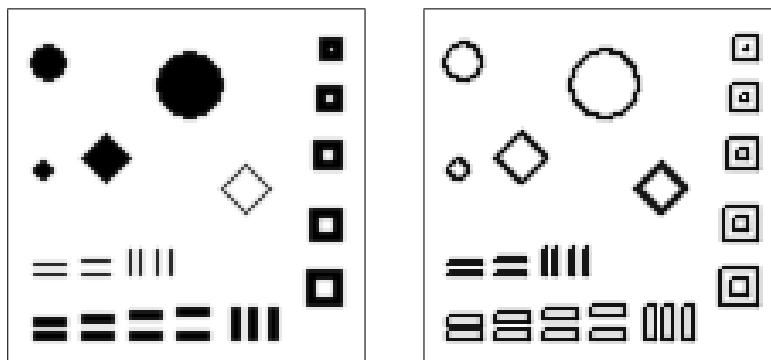
(a) $c_4 = 0.1$



(b) $c_4 = 2$



(c) $c_4 = 10$



(d) $c_4 = 20$

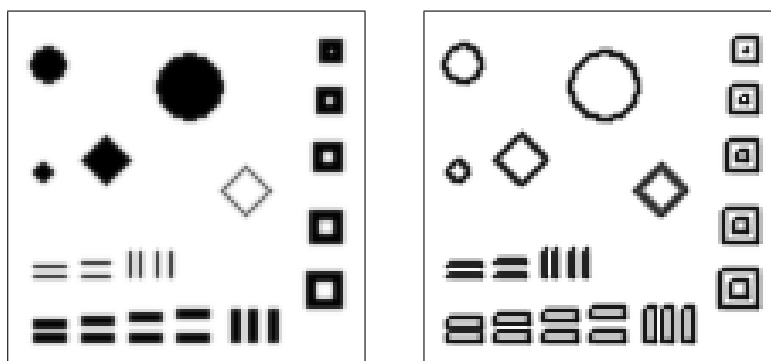


Abbildung 6.6: Variiere ϵ : ϵ ist im linken Bild größer als im rechten. Andere Parameter bleiben fest. $c_1 = 1$, $c_2 = 0,001$, $c_3 = 0,00001$, $c_4 = 0$

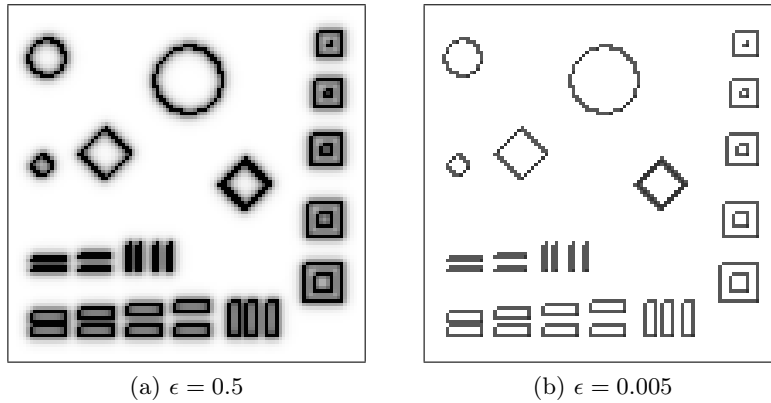
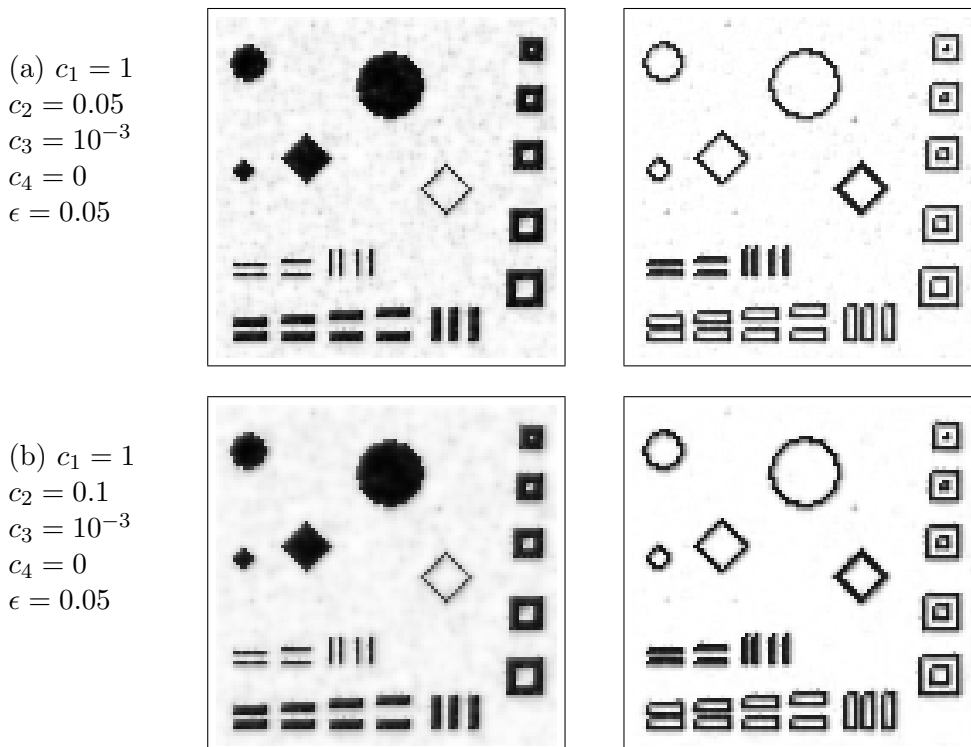


Abbildung 6.7: Entrauschen und gleichzeitige Kantenerkennung des verrauschten Testbildes.



6.3 Quadratische und TV-Regularisierung

Wir betrachten dieselben synthetischen Bilder wie im letzten Abschnitt und wollen das Rauschen im Bild aus Abbildung 6.2b mit Hilfe des Variationsansatzes aus Abschnitt 4.1.1 entfernen, d.h. wir betrachten das Problem

$$\text{Minimiere } F(x) = \int_{\Omega} (I(s) - x(s))^2 ds + \mu \int_{\Omega} f(|\nabla x(s)|) ds; (x, u) \in W_0^{1,p}(\Omega) \times L^p(\Omega, \mathbb{R}^2). \quad (6.4)$$

Als Regularisierungsterme verwenden wir dabei die quadratische als auch die TV-Regularisierung. Um das Problem numerisch zu lösen verwenden wir an dieser Stelle das direkte Verfahren aus Abschnitt 5.3 und den Solver IPOPT. Dabei ist es beim direkten Verfahren notwendig, für die TV-Regularisierung die folgende Approximation zu verwenden:

$$|\nabla x|_{\beta} := \sqrt{|\nabla x|^2 + \beta}. \quad (6.5)$$

Wir zeigen in Abbildung 6.8 exemplarisch nur die besten Bilder, an denen auch der Unterschied zwischen quadratischer und TV-Regularisierung noch einmal deutlich wird. Die quadratische Regularisierung bewirkt nicht nur eine Glättung des Rauschens, sondern auch eine zu starke Glättung der Kanten. Dieser negative Effekt hat zur Folge, dass der PSNR-Wert insgesamt deutlich schlechter wird, während wir mit Hilfe der TV-Regularisierung eine Verbesserung erreichen (vgl. Tabelle B.3). Auch im Vergleich zum Ambrosio-Tortorelli-Funktional aus dem letzten Abschnitt erreichen wir mit Hilfe der TV-Regularisierung bessere PSNR-Werte für das wiederhergestellte Bild.

In Abbildung 6.9 zeigen wir, was geschieht, wenn der Regularisierungsparameter zu groß gewählt wird. Bei der TV-Regularisierung wird derselbe Effekt deutlich wie in Abbildung 4.2. Die kleinskaligen Objekte fangen an zu verschwinden. Bei zu klein gewähltem Regularisierungsparameter würde das Rauschen wiederum nicht entfernt werden.

In diesem Abschnitt kann insgesamt festgehalten werden, dass die TV-Regularisierung auch bei unseren Versuchen mit den direkten Verfahren keine oder nur eine geringfügige Diffusion der Kanten in einem Bild bewirkt. Ziel des nächsten Abschnitts wird es sein, diese Eigenschaft zu nutzen und gleichzeitig eine Kantenerkennung zu erreichen.

6.4 Steuerungsproblem mit L^{∞} -Steuerbeschränkung – Auswirkungen der Steuerbeschränkung

In diesem Abschnitt wollen wir die Auswirkungen einer Steuerbeschränkung untersuchen und nehmen wieder das Testbild aus Abbildung 6.2a, um unseren Steuerungsansatz mit den bisherigen Methoden vergleichen zu können.

Wir betrachten also das Steuerungsproblem (P_1) aus Abschnitt 4.4, wobei hier S die Iden-

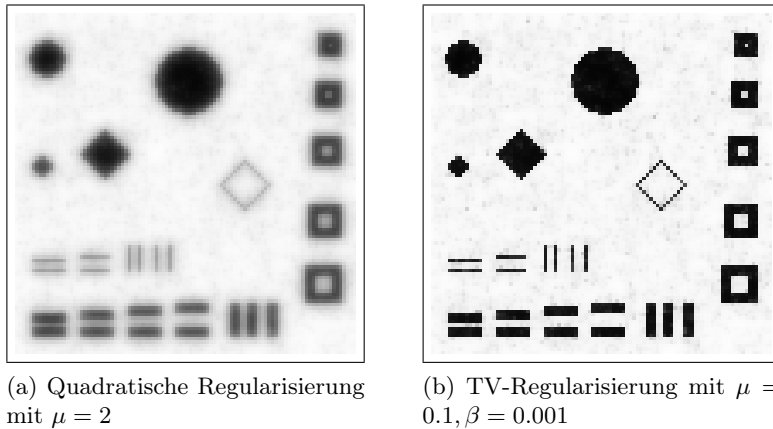


Abbildung 6.8: Glättung basierend auf dem Variationsansatz

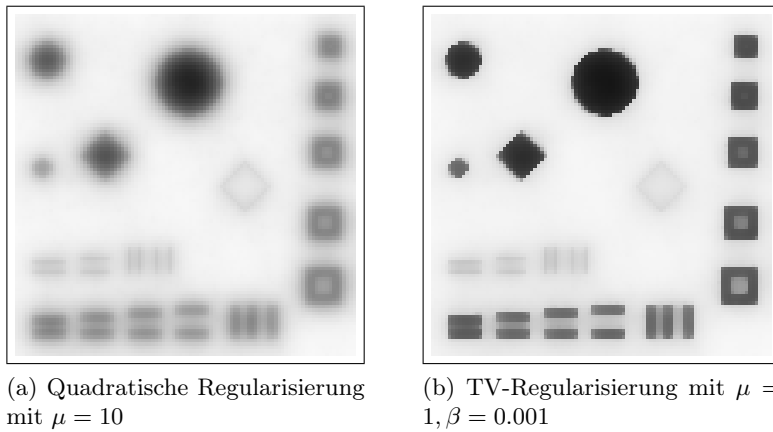


Abbildung 6.9: Glättung basierend auf dem Variationsansatz mit größerem Regularisierungsparameter

tität sei:

$$(P_1) \quad \text{Minimiere } F(x, u) = \int_{\Omega} (I(s) - x(s))^2 ds + \mu \int_{\Omega} f(|u(s)|) ds$$

$$(x, u) \in W_0^{1,p}(\Omega) \times L^p(\Omega, \mathbb{R}^2),$$

$$u(s) = (u_1(s), u_2(s))^T = \nabla x(s), (\forall) s \in \Omega,$$

$$u(s) \in K := \{(v_1, v_2) \in \mathbb{R}^2 \mid |v_1|^q + |v_2|^q \leq R^q\} \quad (\forall) s \in \Omega.$$

Wir haben bereits in Abschnitt 4.4 angemerkt, dass es mit Hilfe der optimalen Steuerung möglich ist, ein Kantenbild zu erhalten. Dies wollen wir an dieser Stelle genauer diskutieren, bevor wir dazu übergehen, die Versuchreihe für das Problem (P_1) auszuwerten. Da die Steuerung gerade durch den Gradienten definiert ist, erhalten wir bei gewissen Parametern ein Kantenbild, indem wir dem Kantenbild an den Stellen, wo die Steuerung aktiv wird, einen schwarzen Pixel zuweisen. An Stellen, an denen die Steuerung nahe Null ist, es sich also um eine homogene Region handelt, werden die Pixel des Kantenbildes weiß gesetzt. Bei Zwischenwerten der Steuerung wird dem Kantenbild ein entsprechender Wert der Grauwertskala zugewiesen, der etwa durch lineare Interpolation der Grauwertskala gewonnen werden kann. Indem wir die Steuerung also auf diese Weise auswerten, erhalten wir ein Kantenbild, da die Steuerung ja gerade dort aktiv wird, wo der Gradient groß ist,

also an den Kanten. Aus diesem Grund ist es auch nicht sinnvoll, die Steuerung im allgemeineren Fall der verrauschten Bilder als Kantendetektor zu verwenden, da Rauschen auch mit großen Gradienten einhergeht und also auch als Kante erkannt werden würde. Bei verrauschten Bildern wird eine Glättung erforderlich sein, um das Rauschen aus den Kantenbildern zu entfernen.

Um aus der optimalen Steuerung ein Kantenbild zu erhalten, gibt es mehrere Möglichkeiten. Die optimale Steuerung $u = (u^{(1)}, u^{(2)})$ ist nach Anwendung des numerischen Verfahrens zunächst in diskreter Form gegeben. Dieser diskreten Steuerung kann eine kontinuierliche Steuerung $\hat{u} := (\hat{u}^{(1)}, \hat{u}^{(2)})$ zugewiesen werden, indem wir (wie in Abschnitt 2.4.1 erwähnt) zwischen den einzelnen Gitterpunkten linear interpolieren oder die einzelnen Komponenten auf jedem Pixel als konstant auffassen. Nun kann das Kantenbild mit Hilfe dieser kontinuierlichen Steuerung auf zwei unterschiedliche Weisen definiert werden. Es bezeichnen im Folgenden $P_{i,j}$ die einzelnen Pixel des gegebenen Gitters $\Omega_{n,n}$. Sie besitzen die Länge 1 und die Breite 1. Wir definieren das Kantenbild $k \in \mathbb{R}^{n,n}$ nun durch eine der folgenden Möglichkeiten:

1. Motiviert durch die Beschränkung der q -Norm von u für $1 \leq q < \infty$ setzen wir für jeden Gitterpunkt $s_{i,j} \in \Omega_{n,n}$:

$$k_{i,j} := \int_{P_{i,j}} ((\hat{u}_1(s))^q + (\hat{u}_2(s))^q)^{\frac{1}{q}} ds. \quad (6.6)$$

2. Motiviert durch die Beschränkung der Maximumnorm von u setzen wir für jeden Gitterpunkt $s_{i,j} \in \Omega_{n,n}$:

$$k_{i,j} := \max_{s \in P_{i,j}} \max(\hat{u}_1(s), \hat{u}_2(s)). \quad (6.7)$$

3. Wir merken an, dass bei der Betrachtung von stückweise konstantem \hat{u} , sich (6.6) bzw. (6.7) vereinfacht schreiben lassen als:

$$k_{i,j} := |(\hat{u}_1(s_{i,j}))^q + (\hat{u}_2(s_{i,j}))^q|, \quad \text{bzw.} \quad (6.8)$$

$$k_{i,j} := \max(\hat{u}_1(s_{i,j}), \hat{u}_2(s_{i,j})). \quad (6.9)$$

Vor allem diese beiden Definitionen werden wir zur Auswertung verwenden, unsere Bilder und die Steuerung also als stückweise konstant auffassen. Wir könnten \hat{u} auch als lineare Interpolation von u auffassen, d.h. zwischen den Gittermittelpunkten linear interpolieren. Jedoch ist ein Vorteil dieser Darstellung nicht direkt ersichtlich und würde bei den in dieser Arbeit benutzten Bildern zu viel Rechenzeit bei der Bildbewertung erfordern.

4. Wir können auch nur einen gewissen Anteil der Steuerung nutzen, nämlich den Anteil, wo sie „nahezu“ aktiv wird, zum Beispiel die obersten 10% der Steuerung. Dies erreichen wir, indem wir das Kantenbild auf die folgende Weise definieren: Für jeden Gitterpunkt $s_{i,j} \in \Omega_{n,n}$ setzen wir

$$k_{i,j} := \begin{cases} \int_{P_{i,j}} (\hat{u}_1(s))^q + (\hat{u}_2(s))^q ds & | \text{wenn } R^q - \frac{1}{10}R^q \leq |(\hat{u}_1(s))^q + (\hat{u}_2(s))^q| \leq R^q, \\ 0 & | \text{sonst.} \end{cases} \quad (6.10)$$

Wenn wir die Kanten auf diese Weise definieren, erhalten wir tatsächlich Kantenbilder, wobei nach der in dieser Arbeit verwendeten Notation 0 schwarz und 1 weiß entspricht, die Kanten in diesen Bildern also weiß wären, während die homogenen Regionen schwarz dargestellt würden. Dies ändert jedoch an der Qualität der Kantenbilder nichts, da wir die Farbwertskala einfach invertieren können.

Zur weiteren Erläuterung betrachten wir zum Beispiel die Definition (6.8) genauer. Befinden wir uns in einer homogenen Region, so ist die Steuerung 0 und somit $k_{i,j} = 0$, und der Punkt (nach Inversion der Farbskala) damit weiß. Umgekehrt, wenn wir uns an einem Punkt $s_{i,j}$ befinden, in dem die Steuerung aktiv wird, bedeutet dies gerade, dass $k_{i,j} = |(\hat{u}_1(s_{i,j}))^q + (\hat{u}_2(s_{i,j}))^q| = R^q$ und damit das Kantenbild im Punkt $s_{i,j}$ den höchsten Wert der Grauwertskala $[0, R^q]$ annimmt, d.h. nach geeigneter Skalierung die Farbe schwarz annimmt.

Mit diesen unterschiedlichen Möglichkeiten Kanten zu definieren, werten wir jetzt schrittweise die Ergebnisse aus, die wir durch numerisches Lösen von (P_2) erhalten haben. Dabei haben wir das direkte Verfahren aus Abschnitt 5.3 verwendet und untersuchen nun zunächst Steuerbeschränkungen mit $1 \leq q < \infty$. In Abbildung 6.10 (und in Tabelle B.4) finden sich die Ergebnisse für den Fall $q = 2$. Wir haben hier keine Regularisierung vorgenommen, also die Potentialfunktion $f = 0$ gesetzt. In der linken Spalte finden sich die wiederhergestellten Bilder. In der rechten Spalte werden in dieser und in den folgenden Abbildungen die zugehörigen Kantenbilder dargestellt, die wir hier durch (6.10) definiert haben. In Abbildung 6.10 haben wir von oben nach unten die Steuerbeschränkung R immer kleiner gewählt. Die Auswirkungen auf die Bilder sind gut zu erkennen. Es findet eine Diffusion bzw. Glättung der Objekte im Bild statt und zwar umso mehr, je kleiner wir die Schranke R wählen. Außerdem werden die Kanten in den Kantenbildern breiter. Es kann gesagt werden, dass der Glättungseffekt, der durch die Steuerbeschränkung bewirkt wird, stärker ist, als die Anwendung der quadratischen Regularisierung (vgl. etwa Abb. 6.9a). Um eine ähnliche Glättung zu erreichen, müsste man bei der quadratischen Regularisierung den Regularisierungsparameter sehr groß wählen.

In Tabelle B.4 wird außerdem deutlich, dass es bei ungestörten Bildern kaum einen Unterschied macht, ob wir die gesamte Steuerung nutzen, um die Kantenbilder zu definieren, oder wie in der Definition (6.10) nur die obersten 10 % der Steuerung. Es kann gesagt werden, dass quantitativ und qualitativ das Kantenbild aus Abbildung 6.10 (b) sehr gut ist. Weiterhin kann festgehalten werden, dass die Kantenbilder bei zu kleiner Schranke nicht mehr brauchbar sind, da die Kanten dann zu sehr das Bild füllen.

Im nächsten Schritt untersuchen wir die Beschränkung der Maximumnorm von u , d.h. den Fall $q = \infty$. Die Steuerbeschränkung hat also die Gestalt

$$\begin{aligned} -R_1 &\leq u_1(s) \leq R_1, \\ -R_2 &\leq u_2(s) \leq R_2, \end{aligned} \tag{6.11}$$

wobei zunächst für die Schranken $R := R_1 = R_2$ gelten soll. Wir lassen nun in den Versuchen jeweils alle bis auf einen Parameter fest, und variieren diesen Parameter, um herauszufinden, welche Wirkung eine Variation hat. Wir wenden das Verfahren nun auf verrauschte synthetische Bild an. Die Ergebnisse finden sich in den Abbildungen 6.11 und 6.12 sowie in Tabelle B.5. Dabei stellen wir in diesem Fall die gesamte Steuerung dar, d.h. wir definieren unsere Kantenbilder durch (6.7). Auch hier variieren wir die Schranke R und wählen sie (in der Abbildung von oben nach unten) immer kleiner. Es ist derselbe Effekt wie beim ungestörten Bild zu beobachten. Um das Rauschen zu entfernen (vgl. Abb. 6.12), ist es

notwendig die Schranke sehr klein zu wählen, was gleichzeitig bewirkt, dass die Objekte im Bild stark diffundieren. Auch die Kantenbilder enthalten entweder zu viel Rauschen oder haben zu breite Kanten.

Aus diesen Gründen gehen wir dazu über, eine Regularisierung einzuführen und im Zusammenhang mit der Steuerbeschränkung zu betrachten. Wir betrachten hier nur die TV-Regularisierung, da diese in Abschnitt 6.3 bereits ein gutes Ergebnis geliefert hat. Zunächst wählen wir einen festen Regularisierungsparameter und variieren in Abbildung 6.13 wiederum die Schranke R . Je kleiner wir R wählen, um so breiter und deutlicher werden die Kanten in den Kantenbildern, was durchaus wünschenswert sein kann, um Kanten hervorzuheben (vgl. insbesondere Abb. 6.14h). Die PSNR-Werte der wiederhergestellten Bilder werden dabei zwar gleichzeitig schlechter (vgl. Tabelle B.5), was wir jedoch in Kauf nehmen, um breitere Kanten zu erhalten. Dass die PSNR-Werte schlechter werden, ist einfach wieder dadurch zu erklären, dass die Objekte im Bild diffundieren und dies eine große Abweichung vom Originalbild bewirkt. Ein weiterer negativer Effekt, der die Kantenbilder betrifft, ist auch hier gut zu beobachten. In den Kantenbildern werden bei kleiner werdender Schranke einige verrauschte Pixel sichtbar (vgl. insbesondere Abb. 6.14h). Jedoch geschieht dies nicht in dem Maße wie in den Abbildungen 6.12 und 6.12, also ohne zusätzliche Regularisierung. Dies belegt unsere Vermutung, dass im Steuerungsansatz eine Regularisierung im Zielfunktional notwendig ist.

Um diese Nachteile zu beheben, können wir nun noch einen Schritt weiter gehen und die verwendeten Methoden nacheinander in einem Zwei-Phasen-Verfahren anwenden. In der ersten Phase bestimmen wir mit Hilfe des Steuerungsproblems (P_1) ein Kantenbild mit der gewünschten Kantenbreite, die wir durch die Schranke R steuern können. Anschließend wenden wir die TV-Regularisierung auf das erhaltene Kantenbild an, um die restlichen verrauschten Pixel aus diesem Kantenbild zu entfernen. Wir haben die Bilder aus Phase 1 und Phase 2 in Abbildung 6.14 gegenübergestellt.

Wenn wir die quantifizierte Kantengüte als Bewertungsmaßstab wählen, so liefert die Schranke $R = 0.5$ die besten Ergebnisse (vgl. die Abbildungen 6.13 (b) und 6.14(c) sowie die Tabellen B.6 und B.7). Allerdings betonen wir noch einmal, dass subjektiv gesehen das Kantenbild aus Abbildung 6.14(d) als besser gewertet werden könnte, da dort die Kanten breiter und damit deutlicher hervortreten. Wir betonen, dass, wenn wir ein Kantenbild mit breiteren Kanten zugrunde legen würden, also unsere idealen Kanten jeweils auf drei oder vier Pixel des diskreten Gitters projizieren würden, wir auch quantitativ bessere Werte erhalten würden.

In Abbildung 6.15 sind die zu den Zustandsgleichungen (6.11) adjungierten Variablen dargestellt, die von IPOPT ebenfalls berechnet werden und somit leicht erhalten werden können. Indem wir die Adjungierten darstellen, erhalten wir (wie man in der Abbildung sieht) auch Kantenbilder. Bei den Adjungierten ist in geringerem Ausmaß derselbe Effekt zu beobachten wie bei der Auswirkung der Steuerung auf das Bild. Je kleiner die Schranke gewählt wird, umso intensiver sind die Kanten. Gleichzeitig ist das Rauschen im Gegensatz zu den Bildern aus Abbildung 6.14h in den Kantenbildern nicht zu sehen. Somit erhalten wir bei gering verrauschten Bildern gute Kantenbilder ohne Rauschen, indem wir die Adjungierten darstellen. Für stärkeres Rauschen konnte vergleichbar gute Ergebnisse nicht mehr erzielt werden. Dann tritt das Rauschen auch wieder in den Kantenbildern auf.

In diesem Abschnitt können wir also insgesamt festhalten, dass sowohl eine L^2 -Steuerbeschränkung als auch eine L^∞ -Steuerbeschränkung breitere Kanten bewirken. Die Steuerbeschränkungen bewirken eine Diffusion der Objekte im Bild und sind nur in Verbindung

mit einer zusätzlichen TV-Regularisierung effektiv. Durch anschließende Anwendung der TV-Regularisierung in einer zweiten Phase erhalten wir gute Kantenbilder.

Abbildung 6.10: Steuerbeschränkung im Fall $q = 2$. Anwendung auf ein ungestörtes Testbild aus Abb. 6.2a

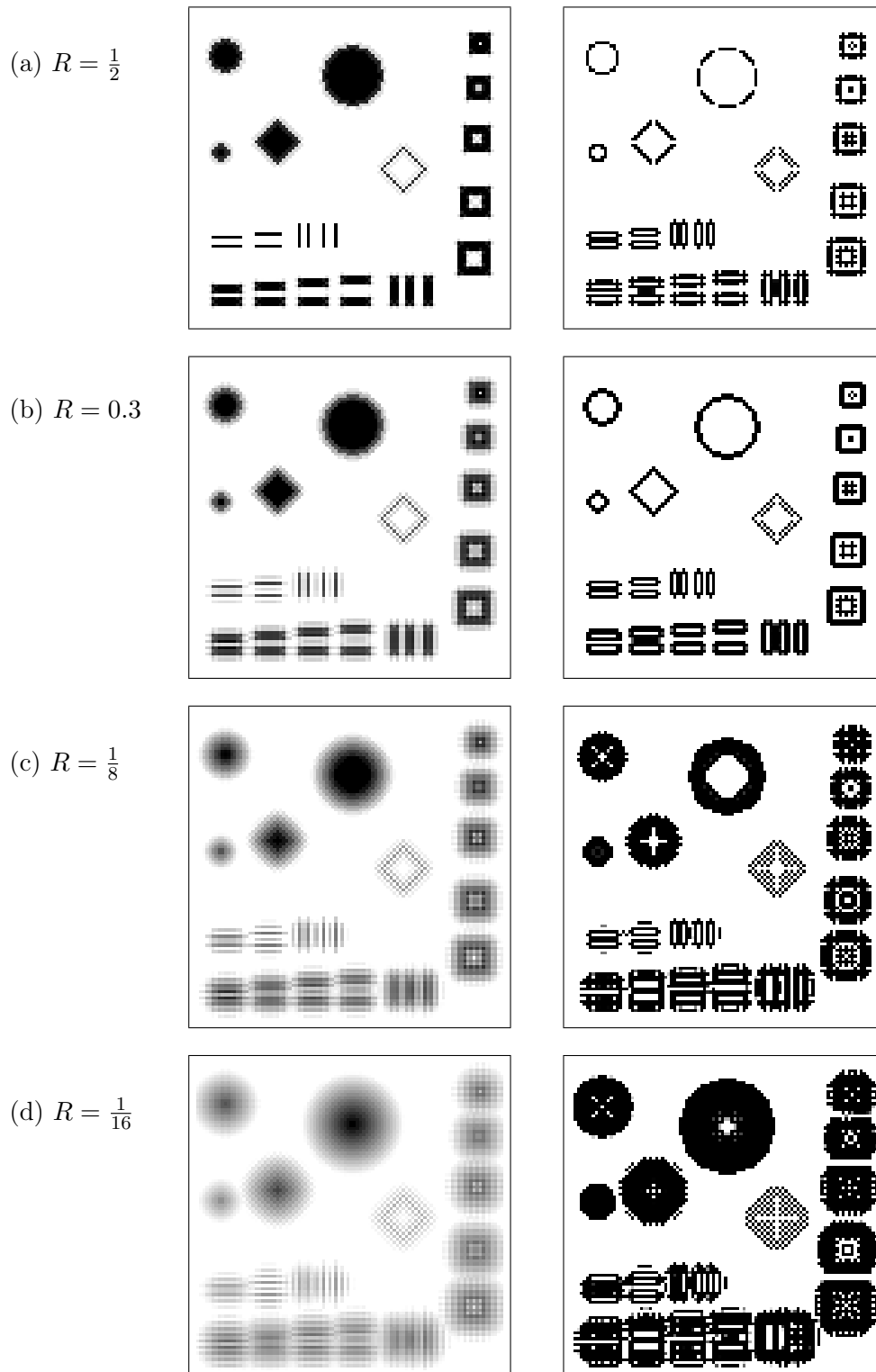


Abbildung 6.11: Anwendung auf verrauschtes Testbild aus Abb. 6.2b mit Maximumnorm und Schranke $R := R_1 = R_2$

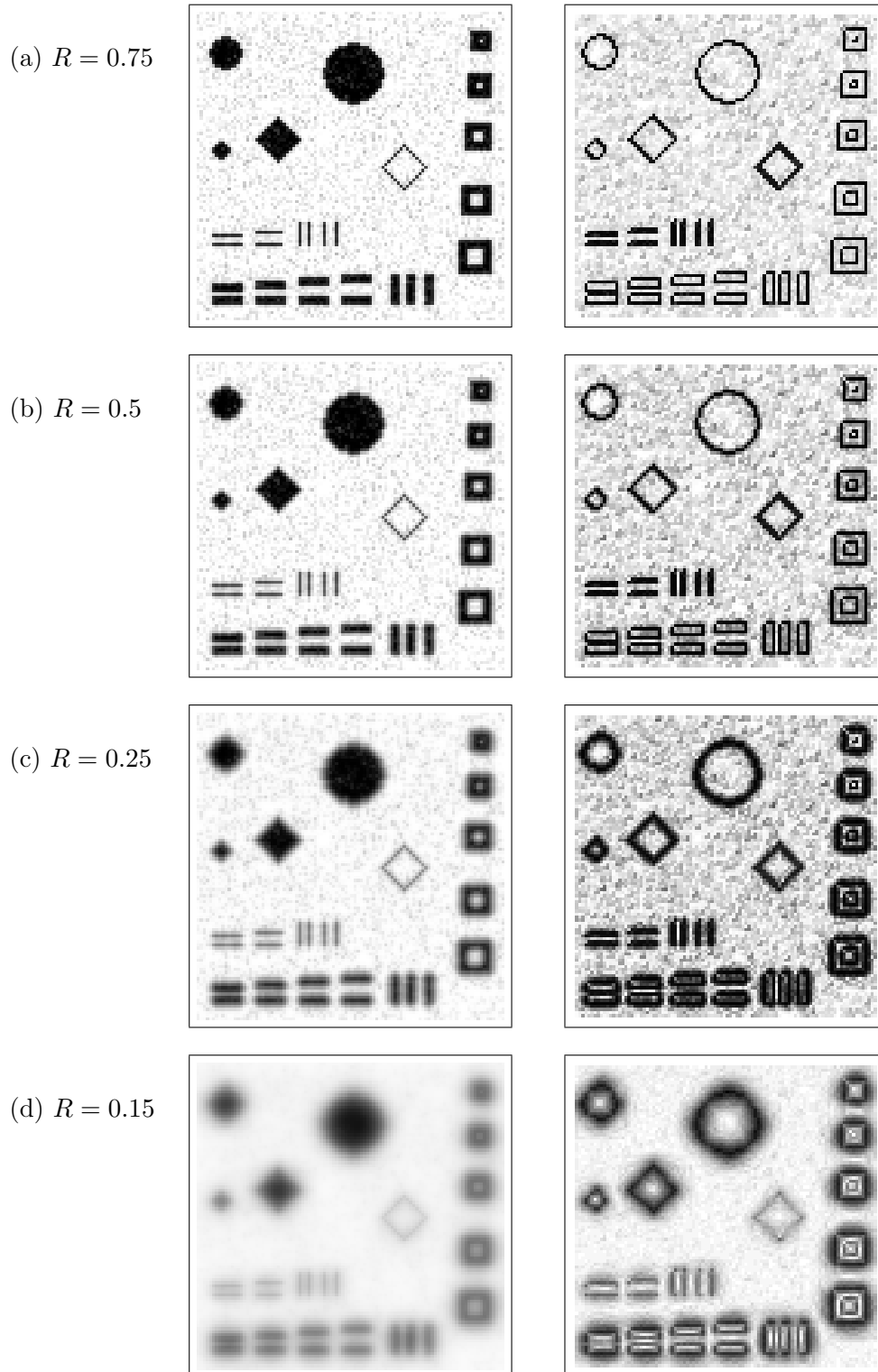


Abbildung 6.12: Anwendung auf verrauschtes Testbild aus Abb. 6.2b mit Maximumnorm

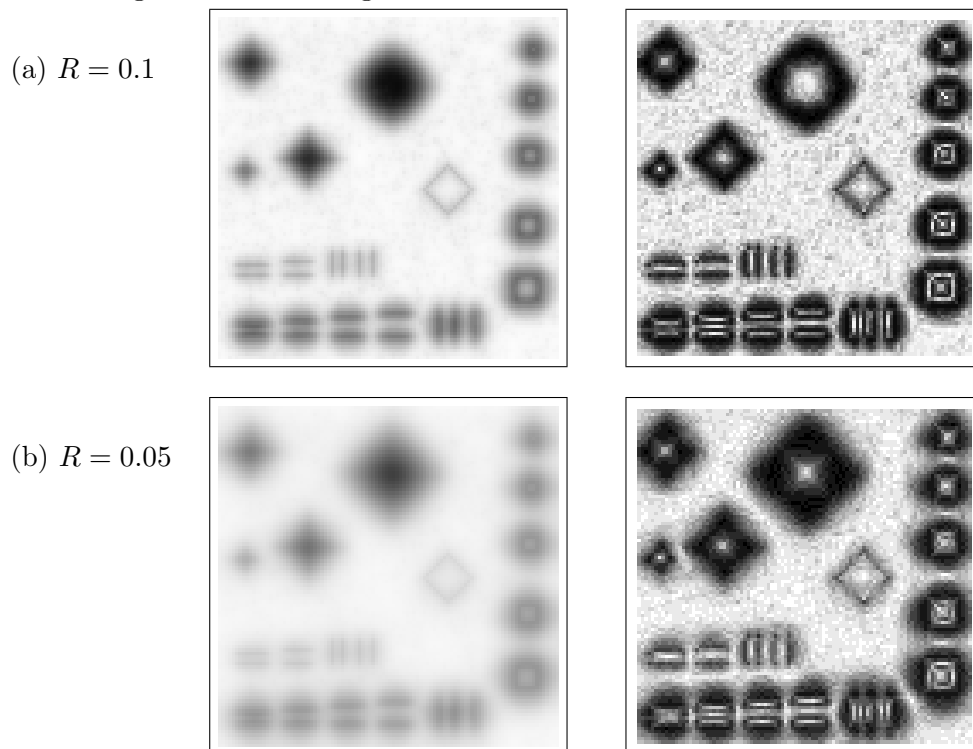


Abbildung 6.13: TV-Regularisierung mit $\beta = 0.001, \mu = 0.1$ und L^∞ -Steuerbeschränkung mit $R := R_1 = R_2$

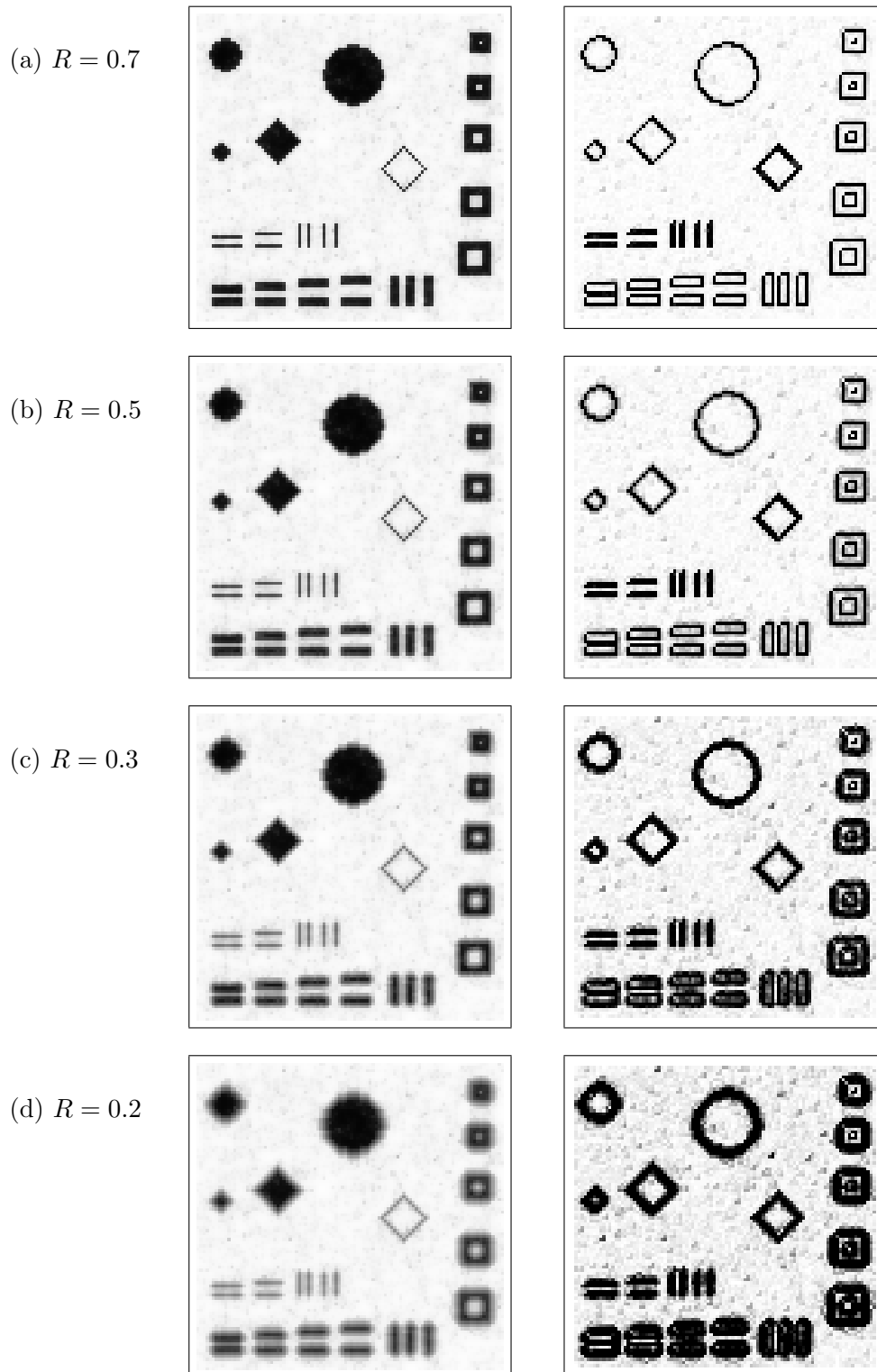
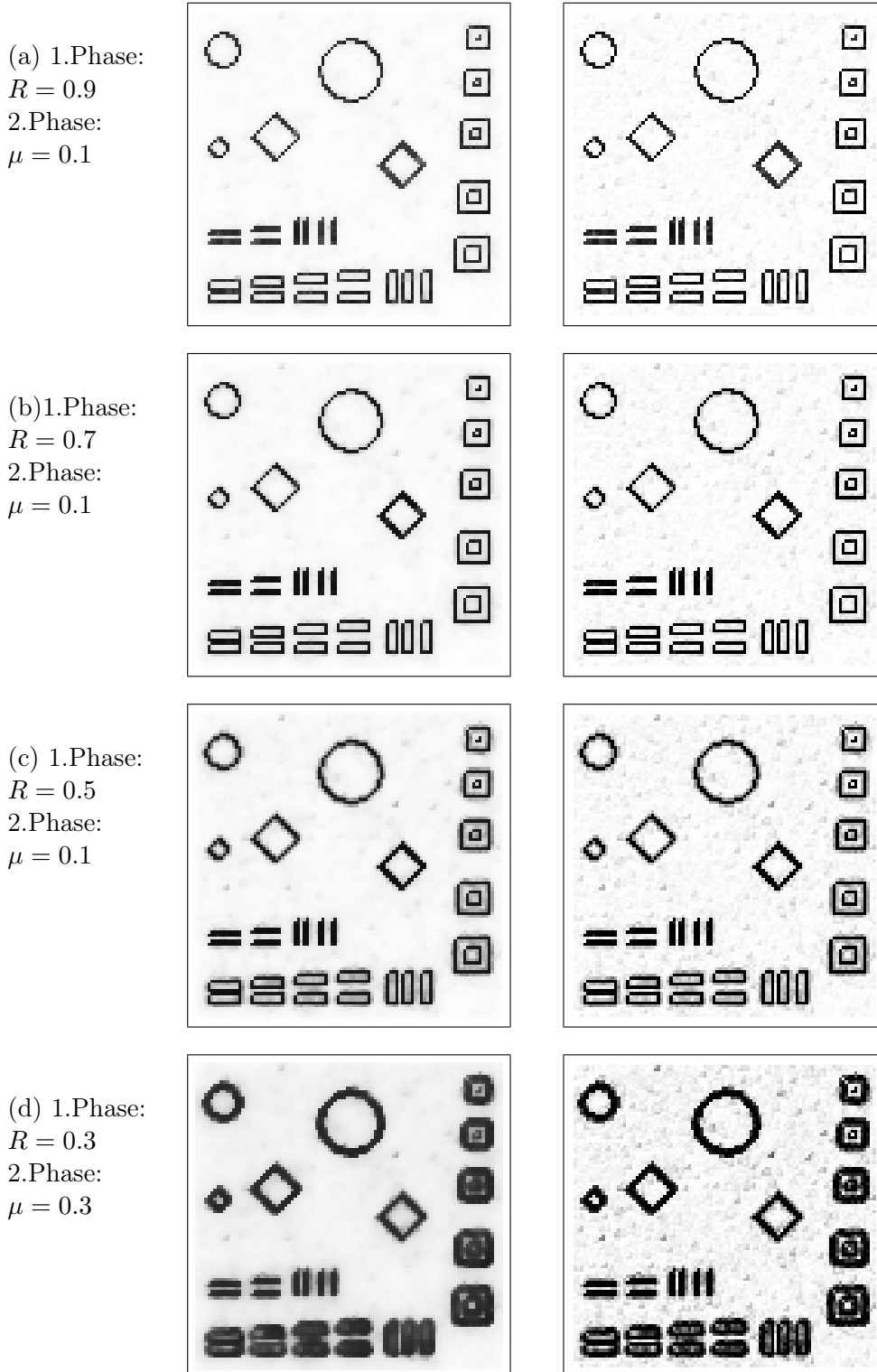
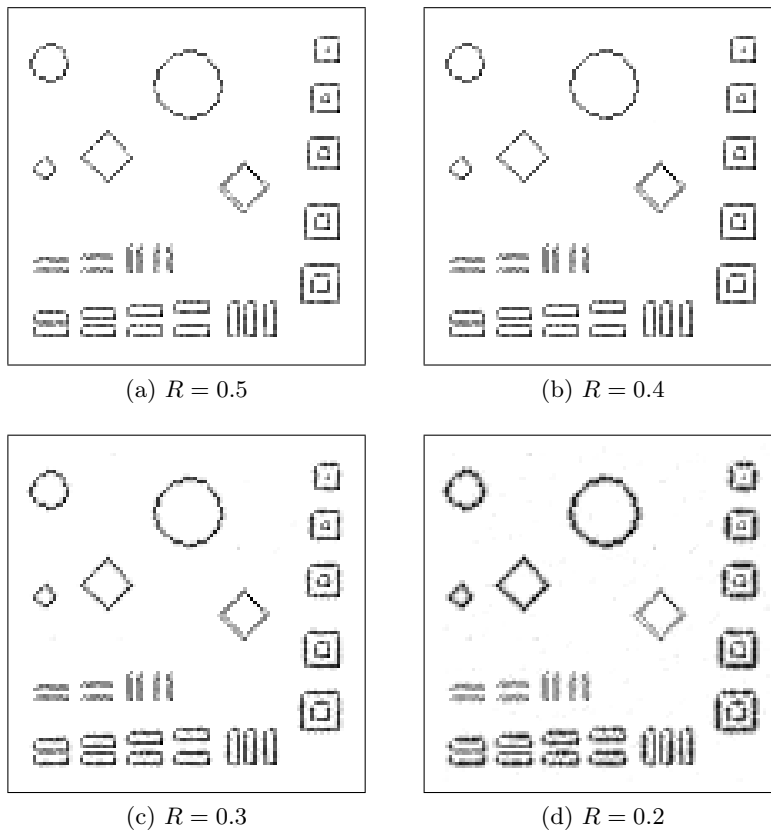


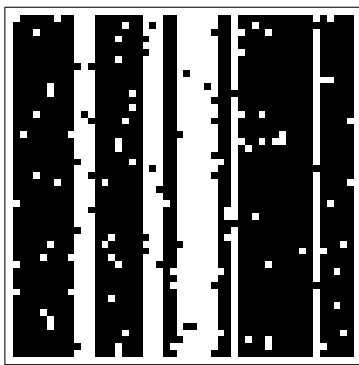
Abbildung 6.14: Zwei-Phasen-Verfahren: In der rechten Spalte die Kantenbilder nach der ersten Phase, in der linken Spalte die Kantenbilder nach der zweiten Phase, d.h. nach Anwendung der TV-Regularisierung mit Regularisierungsparameter μ



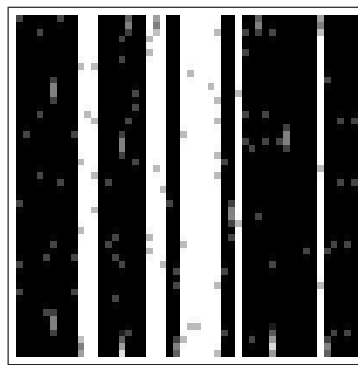
Abbildung 6.15: Adjungierte des Zustands bei L^∞ -Steuerbeschränkung

Streifenbild

Wie bereits in Abschnitt 4.4 angedeutet wurde, kann es Fälle geben, in denen es sinnvoll ist die eine Komponente der Steuerung stärker zu beschränken als die andere. Wir betrachten Steuerbeschränkungen der Gestalt (6.11) und können R_1 und R_2 unterschiedlich wählen. Dies ist zum Beispiel sinnvoll, wenn wir das verrauschte Streifenbild aus Abbildung 6.16a betrachten und wissen, dass das Bild nur vertikale Streifen aufweist. In der Praxis ist dies etwa bei Strichcodes der Fall. Mit diesem a priori-Wissen können wir vertikale Glättung zulassen, während wir Glättung in der horizontalen Richtung verhindern, da so die Streifen sonst diffundieren würden. Für die horizontale Richtung wählen wir also eine große Schranke R_2 , während wir für die vertikale Richtung die Schranke R_1 klein wählen. Die Ausreißer werden somit gezwungen in vertikaler Richtung zu diffundieren. Die Abbildungen 6.16b-6.16d zeigen diesen Effekt: Die Steuerung wird in der ersten Komponente immer stärker beschränkt und die Ausreißer werden ausgeglättet, bis wir schließlich das Idealbild erhalten. Dabei werden Kanten nicht zerstört. Weitere Versuche zeigen, dass wir auch noch bei einer Rauschintensität von 40 % sehr gute und bei 60 % Rauschintensität akzeptable Ergebnisse erhalten, vgl. Abbildung 6.17.



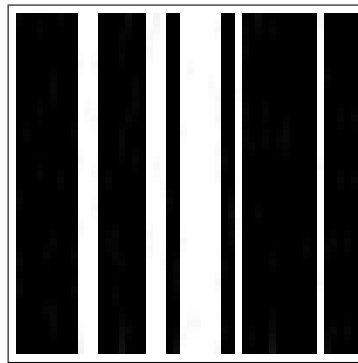
(a) Salt-and-Pepper-Rauschen mit 10% Rauschintensität



(b) $R_1 = 0.25$



(c) $R_1 = 0.1$

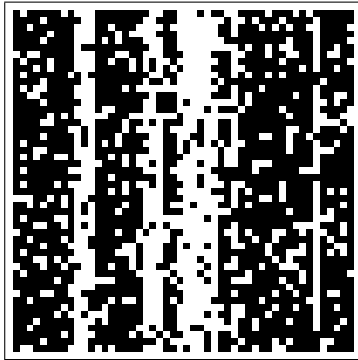


(d) $R_1 = 0.001$

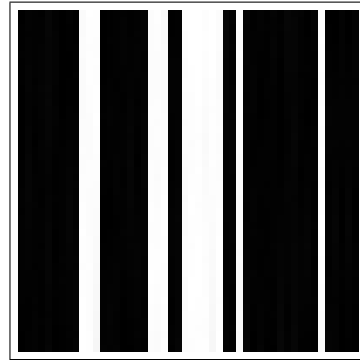
Abbildung 6.16: Streifenbild mit L^∞ -Steuerbeschränkung und $R_2 = 1$

Abbildung 6.17: Streifenbild mit L^∞ -Steuerbeschränkung und $R_1 = 10^{-3}$, $R_2 = 1$. In der linken Spalte jeweils das verrauschte Bild. In der rechten Spalte das wiederhergestellte Bild.

(a) 40%
Rausch-
intensität

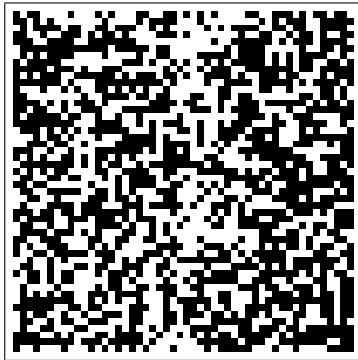


(a) 40% Rauschintensität

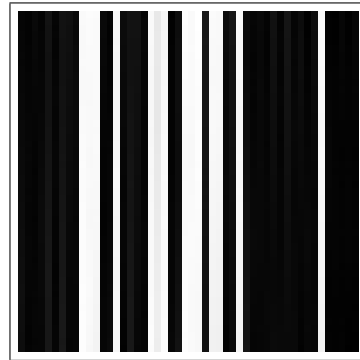


(b) wiederhergestelltes Bild

(b) 60%
Rausch-
intensität



(c) 60% Rauschintensität



(d) wiederhergestelltes Bild

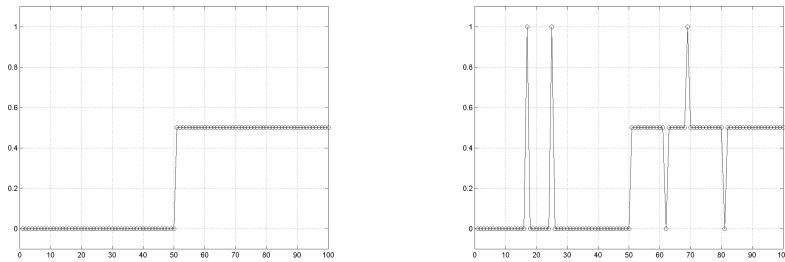
6.5 Steuerungsprobleme mit rotating mask

Im letzten Abschnitt haben wir insbesondere beim Streifenbild den Nutzen der L^∞ -Steuerbeschränkung kennengelernt und wollen nun die allgemeineren Steuerungsprobleme mit rotating mask studieren, wie wir sie in Abschnitt 4.5 formuliert haben. Wir wollen herausfinden, wo diese Modellierung Probleme löst, und wo sie an ihre Grenzen stößt. Dabei verwenden wir direkt die Modelle, die wir bereits in Abschnitt 4.5.2 für diskrete Bilder hergeleitet haben, und lösen diese auch hier mit Hilfe des direkten Verfahrens und dem Solver IPOPT. Wir untersuchen insbesondere den Fall der mit Salt-and-Pepper verrauschten Bilder, da dort die Wirkung der Rotationsmatrix besonders gut zur Geltung kommt. Wie im Abschnitt 4.3 über Impulsrauschen festgestellt wurde, ist es wichtig bei Salt-and-Pepper-Rauschen einen L^1 -Datenterm zu verwenden, was wir im Folgenden anhand von Beispielen auch für unsere Steuerungsprobleme nachweisen wollen.

6.5.1 Der eindimensionale Fall: Anwendung auf Signale

Wir betrachten zunächst Signale, da dort die Wirkung der rotating mask besonders deutlich wird. Gegeben sei also das Signal aus Abbildung 6.18a, und eine mit einigen wenigen Ausreißern verrauschte Version, vgl. Abbildung 6.18b.

Ziel ist es dabei die Ausreißer zu dämpfen oder gar ganz zu eliminieren, während die Kante



(a) Originalsignal mit Kante

(b) Signal mit einigen Ausreißern

Abbildung 6.18: Signal, Original und verrauscht

erhalten bleiben soll. Um das Problem zu lösen verwenden wir das Steuerungsproblem (P_2) aus Abschnitt 4.5.2:

$$(P_2) \quad \text{Minimiere } F(x) = \sum_{i \in \Omega_N} (x_i - I_i)^2$$

$$(x, u) \in \mathbb{R}^N \times \mathbb{R}^{N \times 2},$$

$$u_i = \begin{pmatrix} u_i^{(1)} \\ u_i^{(2)} \end{pmatrix} = (\delta_s^+ x_i, \delta_s^- x_i),$$

$$-R_1 \leq \cos(\alpha_i) u_i^{(1)} + \sin(\alpha_i) u_i^{(2)} \leq R_1,$$

$$-R_2 \leq -\sin(\alpha_i) u_i^{(1)} + \cos(\alpha_i) u_i^{(2)} \leq R_2,$$

$$0 \leq \alpha_i \leq \frac{\pi}{2}.$$

Die Ergebnisse finden sich in Abbildung 6.19. In der linken Spalte haben wir im Datenterm die modifizierte L^1 -Norm mit einer kleinen Konstante $\beta > 0$ verwendet, während wir in

der rechten Spalte im Datenterm die L^2 -Norm verwendet haben. Wir lassen $R_2 = 0.5$ konstant und variieren R_1 . Wie erwartet – und in der linken Spalte gut zu sehen ist – werden beim Senken der Schranke R_1 auch die Ausreißer gedämpft, bis sie schließlich fast vollständig verschwinden. Dies wird mit dem L^2 -Datenterm nur mangelhaft erreicht, da es dort für das Verfahren günstiger scheint eine neue Kante entstehen zu lassen, als die Ausreißer vollständig zu dämpfen. Bei sehr kleinen Schranken $R_1 \leq 0.1$ neigt der Solver Ipopt dazu, Schwierigkeiten zu haben, und das Verfahren konvergiert nur langsam oder gar nicht. In Abbildung 6.20 haben wir aus diesem Grund die Nebenbedingungen der Gestalt (4.21) verwendet, also nur Drehungen um genau 90° zugelassen. Das Verfahren mit diesen Nebenbedingungen führt zu einer noch besseren Dämpfung der Ausreißer, und der Solver IPOPT scheint für sehr kleine Schranken stabilere Ergebnisse zu liefern.

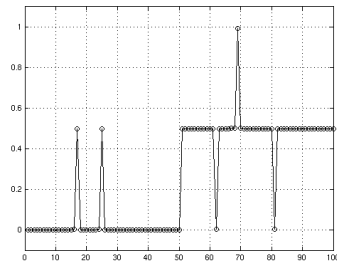
6.5.2 Der zweidimensionale Fall

Wir betrachten nun ein Testbild bestehend aus übereinandergelegten Quadraten unterschiedlicher Intensität und versehen dieses mit Salt-and-Pepper-Rauschen, siehe Abbildung 6.21a. Ziel ist es wieder, das Rauschen zu entfernen, die Kanten der Quadrate jedoch zu erhalten. Um dies zu erreichen verwenden wir das Steuerungsproblem (P_3) aus Abschnitt 4.5.2:

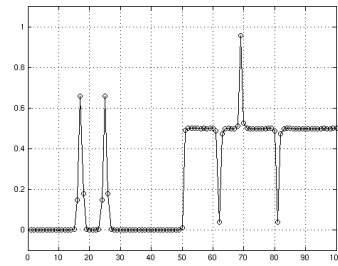
$$\begin{aligned}
 (P_3) \quad \text{Minimiere } F(x) &= \sum_{(i,j) \in \Omega_{N_1, N_2}} (x_{i,j} - I_{i,j})^2 \\
 (x, u, y) &\in \mathbb{R}^{N_1 \times N_2} \times \mathbb{R}^{(N_1-2) \times (N_2-2) \times 4} \times \mathbb{R}^{(N_1-2) \times (N_2-2) \times 4}, \\
 u_{i,j}^{(1)} &= \delta_{s_1}^+ x_{i,j} \quad u_{i,j}^{(2)} = \delta_{s_2}^+ x_{i,j} \\
 u_{i,j}^{(3)} &= \delta_{s_1}^- x_{i,j} \quad u_{i,j}^{(4)} = \delta_{s_2}^- x_{i,j} \\
 -R &\leq y_{ij}^{(1)} u_{ij}^{(1)} \leq R \quad \forall (i, j) \in \text{DIN}(\Omega) \\
 -R &\leq y_{ij}^{(2)} u_{ij}^{(2)} \leq R \quad \forall (i, j) \in \text{DIN}(\Omega) \\
 -R &\leq y_{ij}^{(3)} u_{ij}^{(3)} \leq R \quad \forall (i, j) \in \text{DIN}(\Omega) \\
 -R &\leq y_{ij}^{(4)} u_{ij}^{(4)} \leq R \quad \forall (i, j) \in \text{DIN}(\Omega) \\
 \sum_{k=1}^4 y_{ij}^{(k)} &\geq 1 \quad \forall (i, j) \in \text{DIN}(\Omega)
 \end{aligned}$$

Die Ergebnisse finden sich in Abbildung 6.21. Die Ausreißer werden bei kleiner werdender Schranke R immer mehr gedämpft, bis wir schließlich das Idealbild erhalten. Wir weisen nun noch auf Schwierigkeiten des Verfahrens hin, die in Abbildung 6.22 deutlich werden. Ist die Rauschintensität zu groß, so bilden sich Patches, d.h. Pixelketten oder Pixelgruppen, die durch nebeneinander liegende Ausreißer entstehen. Diese werden durch unser Verfahren als Kante identifiziert, da wir unser Verfahren gerade so modelliert haben, dass es nebeneinanderliegende Pixel gleicher Intensität als Kante erkennt. Aus diesem Grund bleiben diese Patches durch unser Verfahren unbehandelt.

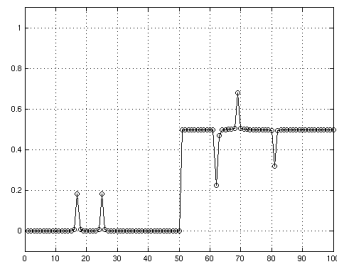
Anzumerken ist außerdem, dass es bei diesem Testbild ausreicht in jedem Punkt nur drei Punkte der Nachbarschaft zu betrachten, wie man sich leicht überlegen kann. Es kommen nämlich keine Linien mit Pixelbreite 1 vor, so dass zur Unterscheidung von Ausreißern und



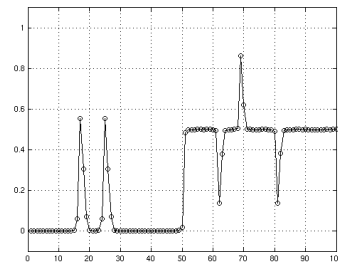
(a) $R_1 = 0.5, R_2 = 0.5$



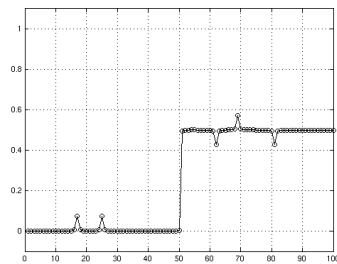
(b) $R_1 = 0.5, R_2 = 0.5$



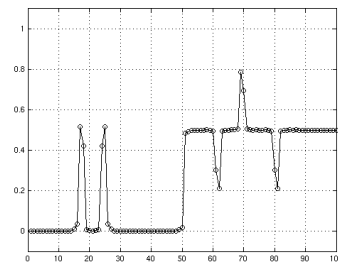
(c) $R_1 = 0.25, R_2 = 0.5$



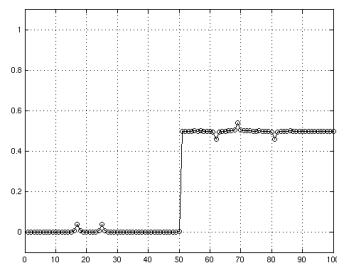
(d) $R_1 = 0.25, R_2 = 0.5$



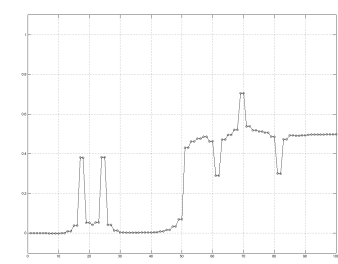
(e) $R_1 = 0.1, R_2 = 0.5$



(f) $R_1 = 0.1, R_2 = 0.5$



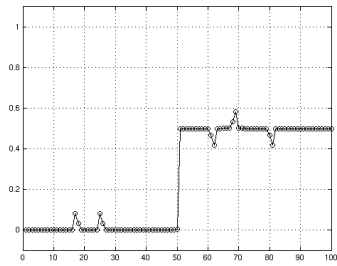
(g) $R_1 = 0.05, R_2 = 0.5$



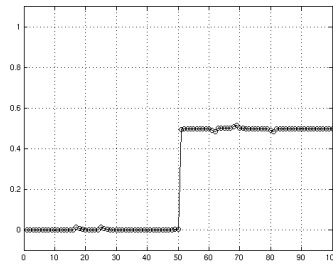
(h) $R_1 = 0.05, R_2 = 0.5$

Abbildung 6.19: In der linken Spalte wurde die modifizierte L^1 -Norm mit $\beta = 0.0001$ für den Datenterm verwendet. In der rechten Spalte wurde die L^2 -Norm verwendet. Die Ergebnisse wurden mit IPOPT erzielt. Der Winkel θ wurde in $[0, \frac{\pi}{2}]$ gewählt.

Bildobjekten in jedem Punkt bereits die Betrachtung von drei Richtungen ausreicht. Aus diesem Grund reicht es, nur 3 statt 4 Steuervariablen in (P_3) zu betrachten, ohne dass wir dadurch die Idee der Methode in diesem Fall selbst verändern. Dadurch wird lediglich die Anzahl der Steuerungsvariablen im endlichdimensionalen Optimierungsproblem kleiner,



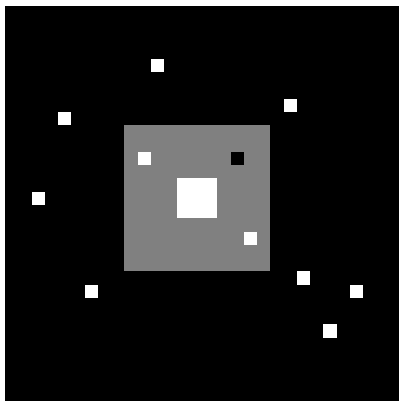
(a) $R_1 = 0.05, R_2 = 0.5$



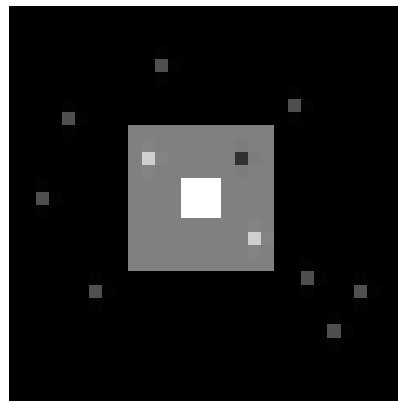
(b) $R_1 = 0.01, R_2 = 0.5$

Abbildung 6.20: $\beta = 0.001$ Ergebnisse mit IPOPT binäre Variable

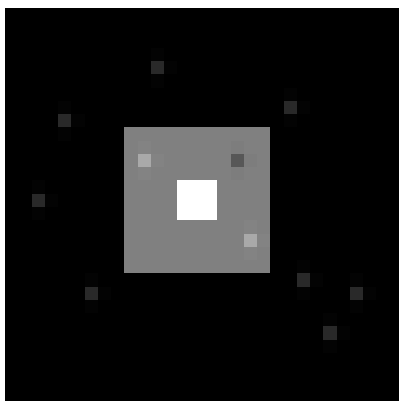
und das Verfahren konvergiert damit schneller.



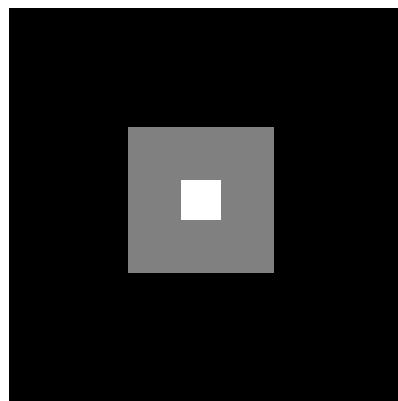
(a) Testbild mit Salt-and-Pepper-Rauschen



(b) $R_1 = 0.1$



(c) $R_1 = 0.05$



(d) $R_1 = 0.01$

Abbildung 6.21: Anwendung des Steuerungsproblems mit rotating mask auf einfaches Testbild.

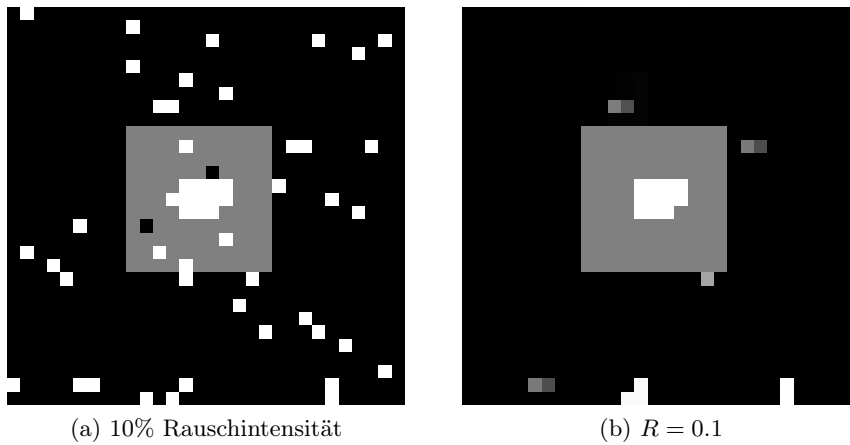


Abbildung 6.22: Anwendung des Steuerungsproblems mit rotating mask auf einfaches Testbild mit höherer Rauschintensität.

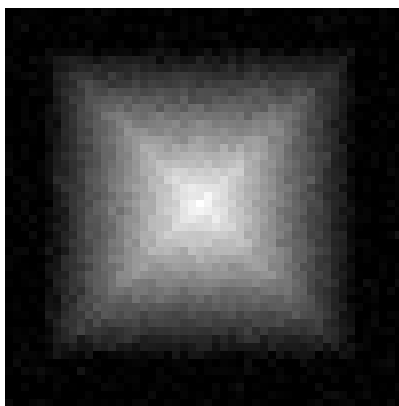
6.5.3 Fallbeispiel: Struktur mit kleinen Variationen und Salt-and-Pepper-Rauschen

Wir wollen unsere neue Methode noch an einem weiteren Testbild ausprobieren und diese mit anderen Verfahren vergleichen. Gegeben sei ein Bild eines Objektes, das kleine Variationen aufweist, die für dieses Objekt völlig natürlich sein sollen. Das Bild werde nun mit Salt-and-Pepper-Rauschen verrauscht und das Ziel sei nun, das Rauschen zu reduzieren oder zu eliminieren, ohne dass die anderen kleinen Variationen verändert werden. Wir vergleichen unsere Methode mit den von Nikolova und Bar/Kiryati/Sochen vorgeschlagenen Methoden (vgl. Abschnitt 4.3). Als Testbild betrachten wir eine Struktur, die sich mit einem „Gebirge“ vergleichen lässt. In Abbildung 6.26 sind jeweils die besten Ergebnisse zusammengestellt, die wir mit Hilfe der unterschiedlichen Verfahren erhalten haben. In der linken Spalte sind die wiederhergestellten Bilder zu sehen, in der mittleren Spalte jeweils deren Höhenbilder, während wir in der rechten Spalte jeweils die Differenz von wiederhergestelltem Bild und Originalbild bestimmt haben und anschließend das Höhenbild dieser Differenz dargestellt haben. An den PSNR-Werten wird ersichtlich, dass unsere Methode in diesem Fallbeispiel bessere Ergebnisse liefert als die beiden anderen Methoden (vgl. auch die Tabellen B.10, B.11, B.12). Auch das Höhenbild scheint dem ursprünglichen Höhenbild am nächsten zu kommen.

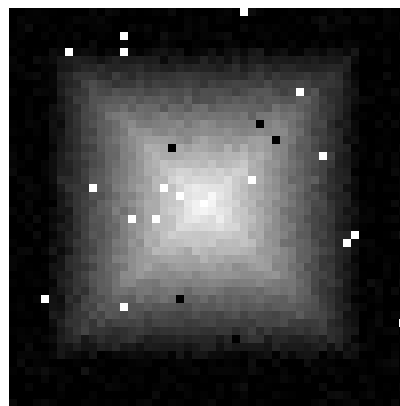
Wir haben dabei wie im letzten Abschnitt das Steuerungsproblem (P_3) benutzt, wobei wir die Nebenbedingungen diesmal wie folgt definiert haben:

$$\begin{aligned}
 -R_1 &\leq \cos(\alpha_{i,j})(u_{i,j}^{(1)} + u_{i,j}^{(3)}) + \sin(\alpha_{i,j})(u_{i,j}^{(2)} + u_{i,j}^{(4)}) \leq R_1 \\
 -R_2 &\leq -\sin(\alpha_{i,j})(u_{i,j}^{(1)} + u_{i,j}^{(3)}) + \cos(\alpha_{i,j})(u_{i,j}^{(2)} + u_{i,j}^{(4)}) \leq R_2
 \end{aligned}$$

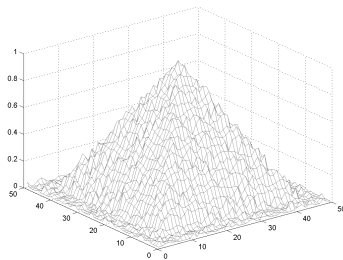
Wie in Abschnitt 4.5.2 angemerkt wurde, eignen sich diese Nebenbedingungen nicht, wenn wir scharfe Ecken im Bild haben. In diesem Fallbeispiel führten diese Nebenbedingungen jedoch zu guten Ergebnissen.



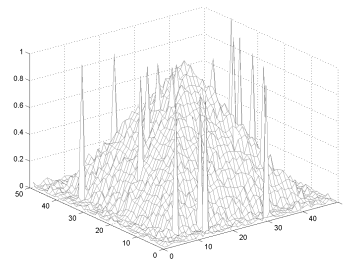
(a) Original



(b) Original mit Salt-and-Pepper



(c) Höhenbild des Originals



(d) Höhenbild des verrauschten Bildes

Abbildung 6.23: Eine Struktur mit kleinen Variationen, verrauscht mit Impulsrauschen

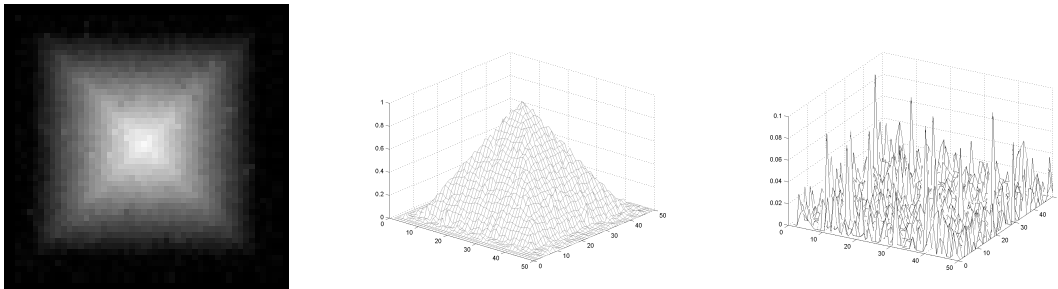


Abbildung 6.24: Methode von Bar/Kiryati/Sochen, PSNR=38,83

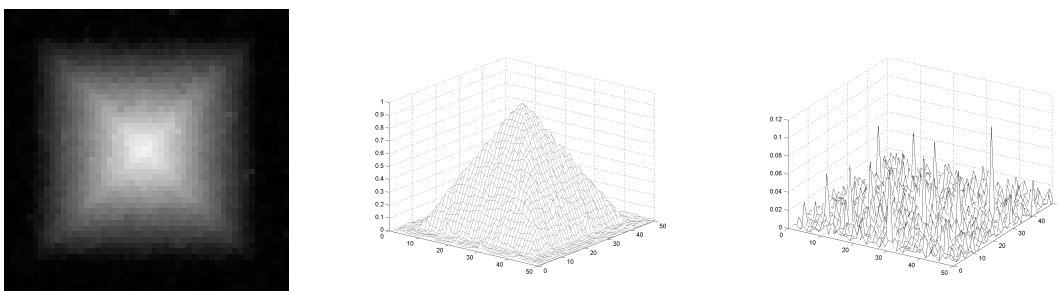


Abbildung 6.25: Methode von Nikolova, PSNR=39,09

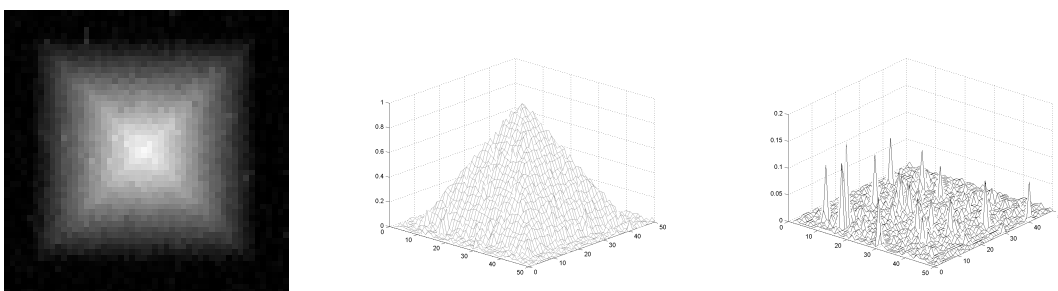


Abbildung 6.26: Steuerungsproblem mit rotating mask, PSNR=40,33

6.5.4 Lena-Testbild

Zuletzt wenden wir das Steuerungsproblem (P_3) auf das Lena-Bild an, das wir ebenfalls mit Salt-and-Pepper-Rauschen versehen, man vergleiche Abbildung 6.27a. Zum Vergleich ziehen wir wieder die Verfahren von Nikolova und Bar/Kiryati/Sochen heran. Wir sehen in Abbildung 6.27, dass bei genügend kleiner Schranke R unser Verfahren es zwar schafft, die Ausreißer zu entfernen, das Bild selbst jedoch zu sehr verändert wird, was insbesondere in Abbildung 6.27d gut zu erkennen ist. Dies ist etwa an der Hutkante oder an den Augenbrauen gut zu beobachten. Unter anderem können wir dies auf die Schwierigkeit der diagonalen Kanten zurückführen, die wir in Abschnitt 4.5 erwähnt haben. Zum anderen aber auch darauf, dass bei Bildern, die in der Natur vorkommen, auch innerhalb von Kanten kleine Intensitätsunterschiede auftreten können. Ist die Schranke R zu klein, so versucht unser Verfahren diese Intensitätsunterschiede zu beseitigen, wodurch Gruppen von Pixeln gleicher Intensität entstehen, also beinahe ein „Cartoon“. Angemerkt sei auch, dass bei zu kleinen Schranken R sowohl der Solver IPOPT als auch LOQO Schwierigkeiten haben und nur langsam oder gar nicht konvergieren. Dies mag damit zusammenhängen, dass die Nebenbedingungen nicht konvex sind. Insgesamt war es deshalb auch schwierig, Zwischenergebnisse zwischen den Bildern aus den Abbildungen 6.27c und 6.27d zu erhalten.

In Abbildung 6.28 zeigen wir jeweils zwei der besten Ergebnisse, die wir mit den Verfahren von Nikolova bzw. Bar/Kiryati/Sochen erhalten haben, sowie in Abbildung 6.29 Ergebnisse bei einer höheren Rauschintensität. Hier leidet unser Verfahren unter demselben Problem, wie in dem Testbild mit den Quadraten: Gruppen von Ausreißern werden als Kante angesehen, wie in Abbildung 6.29e deutlich wird. Stattdessen liefern die Vergleichsverfahren gute Ergebnisse und schaffen es die Ausreißer zu eliminieren. Die zugehörigen Werte finden sich auch in den Tabellen B.13, B.14 und B.15.



(a) Lena mit Salt-and-Pepper-Rauschen



(b) $R = 0.1$



(c) $R = 0.05$



(d) $R = 0.01$

Abbildung 6.27: Anwendung des Steuerungsproblems mit rotating mask auf verrauschtes Lena-Bild



(a) Bar/Kiryati/Sochen, $c_1 = 1, c_2 = 3, c_3 = 1, \epsilon = 0.05$



(b) Bar/Kiryati/Sochen, $c_1 = 1, c_2 = 2, c_3 = 0.5, \epsilon = 0.05$

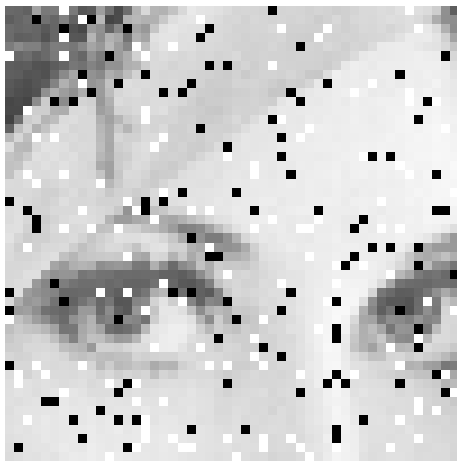


(c) Nikolova, $\mu = 0.5, \beta = \delta = 0.001$



(d) Nikolova, $\mu = 0.4, \beta = \delta = 0.001$

Abbildung 6.28: Anwendung der Verfahren von Nikolova bzw. Bar/Kiryati/Sochen auf verraushtes Lena-Bild



(a) 10 % Rauschintensität



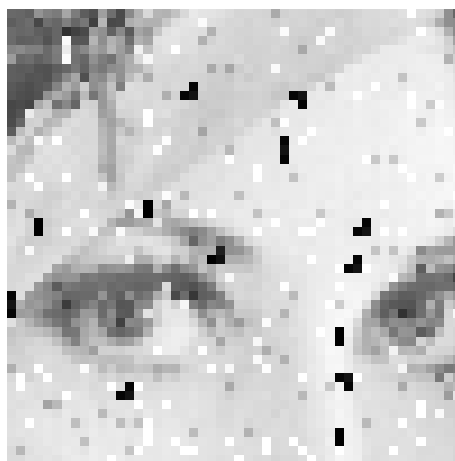
(b) Nikolova, $\mu = 1.5$



(c) Bar/Kiryati/Sochen, $c_1 = 1, c_2 = 5, c_3 = 0.5, \epsilon = 0.05$



(d) Kantenbild aus dem Verfahren von Bar/Kiryati/Sochen



(e) Anwendung der rotating mask

Abbildung 6.29: Anwendung der Verfahren von Nikolova bzw. Bar/Kiryati/Sochen auf Lena-Bild mit 10 % Rauschintensität

6.6 Diskussion der Resultate

In Abschnitt 6.2 haben wir gesehen, dass das Kantenerkennungsfunktional von Ambrosio-Tortorelli sowohl ein gutes wiederhergestelltes Bild als auch gute Kantenbilder liefert. Die Implementierung mit Hilfe eines iterativen Verfahrens führt zu einer kurzen Laufzeit. Ein Nachteil des Kantenerkennungsfunktional ist, dass es viele Parameter (insgesamt fünf) enthält, die zu wählen sind. Um das optimale Parameterpaar, das die besten Ergebnisse liefert, zu finden, kann deshalb eine Reihe von Versuchen notwendig sein.

In Abschnitt 6.4 haben wir Steuerungsprobleme insbesondere mit der L^∞ -Steuerbeschränkung untersucht. Wir haben versucht, mit Hilfe der optimalen Steuerung Kantenbilder zu erhalten, die sich mit den Kantenbildern aus Abschnitt 6.2 messen lassen. Generell läßt sich sagen, dass sich die Steuerbeschränkung glättend auf das Bild auswirkt, also einen ähnlichen Effekt bewirkt wie die quadratische Regularisierung im Variationsansatz. Aus diesem Grund ist es schwierig, Rauschen mit Hilfe der Steuerbeschränkung zu entfernen, ohne dabei das Bild selbst zu sehr zu zerstören. Wie wir gesehen haben, diffundieren die Objekte im Bild nämlich dabei zu sehr, und wir erhalten sehr breite Kanten. Nehmen wir im Zielfunktional die TV-Regularisierung hinzu, so erhalten wir sowohl gute Ergebnisse bei den wiederhergestellten Bildern als auch gute Kantenbilder durch Auswertung der optimalen Steuerung. An dieser Stelle kann die Steuerung sinnvoll eingesetzt werden. Wir können mit Hilfe der Schranke die Kantenbreite der Kanten steuern. Zwar wird bei zu kleiner Schranke auch das Rauschen im Kantenbild sichtbar, doch können wir diesen Nachteil mit einem Zwei-Phasen-Verfahren beseitigen. Wir wenden zunächst das Steuerungsproblem mit Steuerbeschränkung und anschließend den Variationsansatz mit TV-Regularisierung an. Der Nachteil dieses Zwei-Phasen-Verfahrens liegt in seiner Laufzeit. Der Vorteil liegt jedoch darin, dass wir hierbei nur zwei Parameter variieren müssen. Die Parameter werden dabei unabhängig voneinander gewählt. Zuerst bestimmen wir die Schranke so, dass wir die gewünschte Kantenbreite erhalten, und anschließend beseitigen wir in der zweiten Phase die verbleibenden gestörten Pixel.

Zuletzt haben wir in Abschnitt 6.5 Steuerungsprobleme mit rotating mask untersucht und mit den Methoden von Nikolova bzw. Bar/Kiryati/Sochen verglichen. Wir haben mit diesen Methoden versucht, Salt-and-Pepper-Rauschen bei einer Reihe von unterschiedlichen Bildern bzw. Signalen zu entfernen. Bei den synthetischen Bildern mit geringem Salt-and-Pepper-Rauschen haben wir durch das Lösen unseres Steuerungsproblems fast ideale wiederhergestellte Bilder erhalten. Diese Ergebnisse ließen sich durchaus mit den anderen Verfahren aus der Forschung vergleichen und lieferten bei den verwendeten Testbildern sogar etwas bessere Ergebnisse. Bei zu großer Rauschintensität ist hingegen eine Schwäche unseres Steuerungsansatzes deutlich geworden. Es kann Gruppen von verrauschten Pixeln nicht handhaben. Auch bei dem Lena-Bild schnitten die anderen Verfahren etwas besser ab. Hier gäbe es jedoch Ansätze, etwa mit einem Vorverarbeitungsschritt, die unser Verfahren optimieren würden. Bei den Steuerungsproblemen mit rotating mask gibt es, wie wir gesehen haben, unterschiedliche Möglichkeiten, die Nebenbedingungen zu formulieren bzw. die rotating mask zu implementieren. Abhängig von den zugrundeliegenden Bildern haben wir einige gute Lösungen gefunden. An dieser Stelle wäre noch Forschungsaufwand notwendig, um eine allgemeine Implementation zu finden, die wirklich in jedem Fall auch ausreichend schnell konvergiert.

Zuletzt sei erwähnt, dass wir im Gegensatz zur Literatur hauptsächlich direkte Verfahren benutzt haben. Diese sind anwenderfreundlicher, da sie leicht zu implementieren sind, dafür aber zu längeren Laufzeiten neigen. Unsere numerischen Ergebnisse zeigen jedenfalls, dass sich direkte Verfahren – bei nicht zu großen Bildern – dazu eignen, um Variations- und Steuerungsprobleme in der Bildverarbeitung zu lösen. Dies kann insbesondere dann sinnvoll sein, wenn wir eine neue Methode mit einem einfachen Bild testen wollen.

7 Zusammenfassung und Ausblick

Wir haben in dieser Arbeit Variations- und Steuerungsprobleme mit Anwendungen in der Bildverarbeitung behandelt. Neben den bekannten Energiefunktionalen mit quadratischer und TV-Regularisierung sowie dem Funktional von Ambrosio-Tortorelli haben wir Steuerungsprobleme zur Lösung des image-restoration- und image-segmentation-Problems entwickelt.

Der Lösungsansatz über optimale Steuerungsprobleme war bisher nicht Gegenstand der Forschung und wurde in dieser Arbeit erstmals behandelt. In umfangreichen Versuchsreihen haben wir die Wirkung einer Steuerbeschränkung untersucht und uns dabei insbesondere auf die L^∞ -Steuerbeschränkung konzentriert. Das Ergebnis dieser Versuchsreihen war, dass wir zufriedenstellende Kantenbilder erhalten, wenn wir zusätzlich im Zielfunktional eine Regularisierung verwenden. Wir sind zu dem Erkenntnis gelangt, dass die Steuerbeschränkung einerseits glättend wirkt und wir außerdem mit Hilfe der Steuerung die Breite der Kanten steuern können. Dabei können etwas breitere Kanten durchaus wünschenswert sein, um diese im Bild hervorzuheben. Dies kann als positiv gewertet werden, auch wenn unsere quantitativen Bewertungszahlen diesen Effekt nicht wiedergeben konnten. Der einzige wirklich negative Aspekt war, dass diese Kantenbilder noch einige verrauschte Pixel enthielten. Diesen Effekt konnten wir jedoch beseitigen, indem wir ein Zwei-Phasen-Verfahren vorgeschlagen haben und auf die erhaltenen Kantenbilder nochmal einen Variationsansatz mit TV-Regularisierung angewendet haben. Auf diese Weise erhalten wir einen Ansatz, dessen Kantenbilder sich qualitativ mit den Kantenbildern vom Ambrosio-Tortorelli-Funktional vergleichen lassen. Dabei sind die Vorteile unserer Methode, dass wir insgesamt weniger Parameter bestimmen müssen und leicht breitere Kanten erreichen können.

Ein weiterer Schwerpunkt dieser Arbeit war die Untersuchung von Salt-and-Pepper-Rauschen. Wir haben dazu in Kapitel 4 die Steuerungsprobleme mit rotating mask entwickelt und in Kapitel 6 mit Methoden, die von Nikolova und Bar/Kiryati/Sochen entwickelt wurden, verglichen. Dabei zeigte sich, dass unsere neue Methode bei synthetischen Bildern mit kontrolliertem Salt-and-Pepper-Rauschen zu durchaus guten Ergebnissen führte. Sie beseitigt das Rauschen, ohne dabei Kanten zu zerstören. Hier wäre weitere Forschungsarbeit möglich und notwendig, um die Steuerungsprobleme mit rotating mask auch auf allgemeinere Fälle effektiv anwenden zu können. Um diagonal verlaufende Kanten nicht zu zerstören, wäre hier noch eine andere oder erweiterte Implementation notwendig. Außerdem ließe sich unser Verfahren, ähnlich wie bei den Methoden von Nikolova und Bar/Kiryati/Sochen, durch ein Zwei-Phasen-Verfahren, wie es in [11] und [14] vorgeschlagen wurde, verbessern. In der ersten Phase würden wir die verrauschten Pixel mit Hilfe eines adaptiven Median-Filters feststellen und anschließend geeignet unser Steuerungsproblem mit rotating mask anwenden, um die neuen Werte nur an den verrauschten Pixeln zu erhalten.

In dieser Arbeit haben wir im Gegensatz zur gängigen Literatur fast ausschließlich direkte Verfahren benutzt und gezeigt, dass wir mit diesem Ansatz Aufgaben der Bildverarbeitung bearbeiten können. Mit direkten Verfahren können wir unsere Optimierungsprobleme leicht und schnell implementieren, und können insbesondere leicht Probleme mit Nebenbedingungen lösen. Die Verfahren konvergieren bei nicht allzu großen Bildern ausreichend

schnell. Ein weiterer Schritt würde nun darin bestehen, weitere Probleme der Bildverarbeitung – vor allem solche, die Nebenbedingungen enthalten – mit Hilfe dieses Ansatzes zu lösen.

Alle Ansätze in dieser Arbeit ließen sich auf verzerrte Bilder erweitern und noch einmal untersuchen. Die Modellierung wäre dabei sehr einfach, da nur der systematische Fehler – der ja ein Faltungsoperator ist – jeweils in den Datenterm integriert werden müsste. Die Ansätze dazu liegen in dieser Arbeit meistens bereits vor.

Zuletzt wäre es auch noch notwendig, die Konvergenz der Verfahren zu erforschen, also insbesondere den Nachweis der Konvergenz der diskreten Bilder gegen die Lösungen zu erbringen, die durch die Existenzsätze garantiert werden. In diesem Zusammenhang wäre die Untersuchung der Methoden mit Hilfe von finiten Elementen von Interesse. Die Diskussion der Vor- und Nachteile einer Diskretisierung mit Hilfe der Finite-Elemente-Methode anstatt der Finite-Differenzen-Methode wäre erforderlich.

A Beispiel eines AMPL-Programms

Es soll nun eines der AMPL-Programme vorgestellt werden, die wir in dieser Arbeit verwendet haben. Für die Syntax von AMPL verweisen wir nochmals auf [20], doch wurde das Programm gut auskommentiert, so dass sich viele Erklärungen bereits im Programm befinden. Die Kommentare in AMPL sind durch das Symbol `#` gekennzeichnet.

Das beobachtete oder gestörte Bild wird in Matrixform an das Programm übergeben. Diese Matrix ist im Programm mit `A` bezeichnet und hat m Zeilen und n Spalten. Das Ergebnisbild ist die Optimierungsvariable x .

```
param m; #Anzahl der Zeilen
param n; #Anzahl der Spalten

# Definition der Indextmengen, die benötigt werden
set I:=1..m;
set J:=1..n;
set I1:=2..m-1;
set J1:=2..n-1;
set Rand:=1..1;
set Randm:=m..m;
set Randn:=n..n;

# Erklärung der Parameter
param A {I,J}; # in der Matrix A wird das gestörte Bild gespeichert

# Deklaration der Optimierungsvariablen
var x{I,J};
var u1{I,J};
var u2{I,J};
var u3{I,J};
var u4{I,J};
var y1{I,J}>=0;
var y2{I,J}>=0;
var y3{I,J}>=0;
var y4{I,J}>=0;
var theta{I,J};

#Parameter für die Steuerbeschränkung
param q:=2; #beschreibt die Art der Steuerbeschränkung
param R1:=0.5; #Erste Schranke
```

```

param R2:=0.5; #Zweite Schranke

#weitere Parameter
param pi:=3.1417;
param delta:=0.001; # wird bei der modifizierten  $L^1$ -Norm im Datenterm verwendet
param beta:=0.001; # modifizierten  $L^1$ -Norm in der TV-Regularisierung
param mu:=0.1; # dies ist der Regularisierungsparameter

# um numerische Schwierigkeiten zu vermeiden wird die Totalvariation durch die
# modifizierte  $L^1$ -Norm approximiert
var tv{i in I1,j in J1}=sqrt((abs(x[i+1,j]-x[i,j]))^2+(abs(x[i,j+1]-x[i,j]))^2+beta);

#-----
# nun folgen die Zielfunktionale
# sie bestehen aus einem Datenterm und der Regularisierung
# in diesem Fall verwenden wir die Totalvariation als Regularisierung
#-----

minimize ZF: sum {i in I1} sum {j in J1} ((A[i,j]-x[i,j])^2+mu*tv[i,j]);

#liegt Salt-and-Pepper-Rauschen vor, so müssen wir für den Datenterm die
# modifizierte  $L^1$ -Norm verwenden, betrachten also das folgende Funktional
#minimize ZF: sum {i in I1} sum {j in J1} sqrt((A[i,j]-x[i,j])^2+delta)+mu*tv[i,j];

#-----
#Nebenbedingungen
#-----
# Definition der Steuerung
# bei einer  $L^q$ -Steuerbeschränkung benötigen wir zwei Steuerungsvariablen
# bei den Problemen mit rotating mask benötigen wir vier Steuerungsvariablen
s.t. Nb1 {i in I1,j in J1}: u1[i,j]=abs(x[i+1,j]-x[i,j]);
s.t. Nb2 {i in I1,j in J1}: u2[i,j]=abs(x[i,j+1]-x[i,j]);

s.t. Nb3 {i in I1,j in J1}: u3[i,j]=abs(x[i,j]-x[i-1,j]);
s.t. Nb4 {i in I1,j in J1}: u4[i,j]=abs(x[i,j]-x[i,j-1]);

#-----

# Nebenbedingungen bei einer  $L^q$ -Steuerbeschränkung

s.t. Nb5 {i in I1,j in J1}: -R1^q<=u1[i,j]^q+u2[i,j]^q<=R1^q;

# Nebenbedingungen bei einer  $L^\infty$ -Steuerbeschränkung

```

```

# s.t. Nb6 {i in I1,j in J1}: -R1<=u1[i,j]<=R1;
# s.t. Nb7 {i in I1,j in J1}: -R2<=u2[i,j]<=R2;

# Nebenbedingungen bei der rotating mask
# s.t. Nb8 {i in I1,j in J1}: -R1<=y1[i,j]*u1[i,j]<=R1;
# s.t. Nb9 {i in I1,j in J1}: -R1<=y2[i,j]*u2[i,j]<=R1;
# s.t. Nb10 {i in I1,j in J1}: -R1<=y3[i,j]*u3[i,j]<=R1;
#s.t. Nb11 {i in I1,j in J1}: -R1<=y4[i,j]*u4[i,j]<=R1;
# s.t. Nb12 {i in I1,j in J1}: y1[i,j]+y2[i,j]+y3[i,j]+y4[i,j]>=1;

#-----

# Zweite Möglichkeit, die Nebenbedingungen der rotating mask zu formulieren

#s.t. Nb13 {i in I,j in J}: -R1<=cos(theta[i,j])*(u1[i,j]+u3[i,j])
#                               +sin(theta[i,j])*(u2[i,j]+u4[i,j])<=R1;
#s.t. Nb14 {i in I,j in J}: -R2<=-sin(theta[i,j])*(u1[i,j]+u3[i,j])
#                               +cos(theta[i,j])*(u2[i,j]+u4[i,j])<=R2;
#
#s.t. Nb15 {i in I,j in J}: 0<=theta[i,j]<=pi/2;

#-----
#Dirichletrandbedingungen: Bild habe zum Beispiel einen weißen Rand
s.t. Nb16 {i in Rand, j in J}: x[i,j]=1;
s.t. Nb17 {i in Randm, j in J}: x[i,j]=1;
s.t. Nb18 {i in I, j in Rand}: x[i,j]=1;
s.t. Nb19 {i in I, j in Randn}: x[i,j]=1;

```

B Tabellen

In diesem Abschnitt sollen die wichtigsten Tabellen zu den in dieser Arbeit durchgeführten Versuchsreihen zusammengestellt werden. Es werden dabei insbesondere die Tabellen ausgewählt, die sich auf Aussagen in dieser Arbeit beziehen. Für weitere Tabellen sei auf die beiliegende CD verwiesen, wo auch die zugehörigen Bilder verfügbar sind.

Es folgen nun einige allgemeine Bemerkungen zu den Tabellen und den Bezeichnungen, die in den Tabellen verwendet werden. Der Index in jeder Tabelle bezeichnet den Bildindex, unter dem das Bild im jeweiligen Ordner auf der CD zu finden ist. Gegebenenfalls, wenn wir zu einem Bild das Kantenbild erstellt haben, gehören zu einem Index also zwei Bilder, das wiederhergestellte Bild selber (z.B. mit 'bild1.png' bezeichnet) und das Kantenbild (z.B. 'kantenbild1.png'). Alle Bilder befinden sich in den jeweiligen Unterordnern des Ordners „Bilder“ auf der CD. In der Spalte „Abbildung“ ist jeweils die Nummer der Abbildung in dieser Arbeit angegeben, falls das jeweilige Bild in der Arbeit abgedruckt wurde.

Wie in Kapitel 6 angemerkt wurde, haben wir Bilder mit Hilfe des PSNR bewertet und die Kanten mit Hilfe der L^2 -Normdifferenz. Die Kantengüte für erhaltene Kantenbilder wurde auf unterschiedliche Weisen berechnet. In den Tabellen benutzen wir die folgenden Abkürzungen:

1. **KG1**: L^2 -Normdifferenz des erhaltenen Kantenbildes mit dem idealen Kantenbild aus Abbildung 6.2c. Das erhaltene Kantenbild wurde dabei abhängig von der L^q -Steuerbeschränkung durch (6.8) bzw. (6.9) definiert.
2. **KG2**: L^2 -Normdifferenz des erhaltenen Kantenbildes mit dem idealen Kantenbild aus Abbildung 6.2c. Das erhaltene Kantenbild wurde dabei abhängig von der L^q -Steuerbeschränkung durch (6.8) bzw. (6.9) in Verbindung mit (6.10) definiert, d.h. nur die obersten 10 % der Steuerung wurden benutzt.
3. **KG3**: L^2 -Normdifferenz des erhaltenen Kantenbildes mit dem idealen Kantenbild aus der Versuchsreihe von Ambrosio-Tortorelli (vgl. Abbildung 6.3b).

Es stellt sich dabei heraus, dass es sinnvoll ist, vor allem die Kantengüte KG1 als Bewertungskriterium zu nutzen. Dies liegt daran, dass (wie bereits in Abschnitt 6.4 begründet wurde) es durchaus Sinn macht, die ganze Steuerung bei der Kantendefinition zu benutzen. Zum anderen waren die Abweichungen zwischen den Kantenbildern aus den Steuerungsproblemen und dem optimalen Kantenbild aus dem Ambrosio-Tortorelli-Funktional größer, als die Abweichungen zu dem von uns definierten idealen Kantenbild (vgl. Tabelle B.4), d.h. KG3 lieferte schlechtere Werte.

Die besten Werte in den Tabellen sind jeweils fett hervorgehoben. Bei den gestörten Bildern wurde dabei gleichzeitig darauf geachtet, dass neben der quantitativen Kantenbewertung der PSNR-Wert des wiederhergestellten Bildes gut ist. Es war dem Verfasser also wichtig ein Parameterpaar zu finden, das sowohl einen guten PSNR-Wert als auch gute Kanten lieferte. Dieses wurde dann als das beste Paar angesehen.

Beim Lena-Bild zeigte sich die Diskrepanz zwischen quantitativer Bewertung, d.h. dem PSNR-Wert, und dem subjektiven Eindruck. Der PSNR-Wert war bei Bildern, bei denen

das Salt-and-Pepper-Rauschen noch deutlich zu sehen war teilweise höher als bei Bildern ohne sichtbares Rauschen. Dies ist dadurch zu erklären, dass das Ausglätten der Ausreißer gleichzeitig immer ein Ausglätten der Bildobjekte bewirkt. Bei der Bilderauswahl war dem Verfasser wichtig, solche Bilder auszuwählen, bei denen das Rauschen nicht mehr oder nur noch kaum zu erkennen ist.

Tabelle B.1: Anwendung des Kantenerkennungsfunktionals von Ambrosio-Tortorelli auf das ungestörte Testbild aus Abbildung 6.2a.

Index	c_1	c_2	c_3	c_4	ϵ	PSNR	KG1	Abbildung
1	1	0,001	0,001	0	0,05	47,33	35,61	
2	1	0,01	0,001	0	0,05	31,8	27,45	
3	1	0,05	0,001	0	0,05	23,2	19,51	6.3(a)
4	1	0,1	0,001	0	0,05	20,98	16,1	6.3(b)
5	1	0,5	0,001	0	0,05	17,87	21,76	6.3(c)
6	1	1	0,001	0	0,05	17,54	28,05	6.3(d)
7	1	0,001	0,0001	0	0,05	50,63	27,28	6.4(a)
8	1	0,001	0,00001	0	0,05	50,98	25,37	
9	1	0,001	0,000001	0	0,05	51,16	24,75	6.4(b)
10	1	0,001	0,0000001	0	0,05	52,19	21,44	6.4(c)
11	1	0,001	10^{-10}	0	0,05	63,93	20,71	6.4(d)
12	1	0,001	0,00001	0,1	0,05	49,87	25,37	6.5(a)
13	1	0,001	0,00001	0,5	0,05	45,49	25,33	
14	1	0,001	0,00001	2	0,05	38,01	24,96	6.5(b)
15	1	0,001	0,00001	10	0,05	26,73	22,55	6.5(c)
16	1	0,001	0,00001	20	0,05	22,48	21,16	6.5(d)
17	1	0,001	0,00001	0	0,5	58,15	19,6	6.6(a)
18	1	0,001	0,00001	0	0,005	50,5	27,46	6.6(b)

Tabelle B.2: Anwendung des Ambrosio-Tortorelli-Funktional auf das gestörte Testbild aus Abbildung 6.2b.

Index	c_1	c_2	c_3	c_4	ϵ	PSNR	Kanten	Abbildung
2	1	0,001	0,001	0	0,05	23,29	36,03	
3	1	0,01	0,001	0	0,05	23,41	27,87	
4	1	0,05	0,001	0	0,05	20,85	20,32	6.7(a)
5	1	0,1	0,001	0	0,05	19,29	16,91	6.7(b)
6	1	0,2	0,001	0	0,05	17,85	16,15	
7	1	0,5	0,001	0	0,05	16,71	21,33	
8	1	0,01	0,001	0,5	0,05	22,38	28,7	
9	1	0,01	0,001	1	0,05	21,27	29,61	
10	1	0,01	0,001	5	0,05	16,99	34,63	
11	1	0,1	0,0001	0	0,05	21,51	28,46	
12	1	0,1	0,002	0	0,05	18,52	20,71	

Tabelle B.3: Anwendung der quadratischen bzw. TV-Regularisierung (modifizierte L^1 -Norm mit $\beta = 0,001$) auf das gestörte Testbild aus Abbildung 6.2b.

Index	Regularisierung	μ	PSNR	Abbildung
1	Quadratisch	0	23,19	
2	Quadratisch	0,001	23,19	
3	Quadratisch	0,01	23,26	
4	Quadratisch	0,1	22,38	
5	Quadratisch	0,3	19,9	
6	Quadratisch	0,5	18,46	
7	Quadratisch	0,7	17,54	
8	Quadratisch	1	16,62	
9	Quadratisch	1,5	15,68	
10	Quadratisch	2	15,07	6.8(a)
11	Quadratisch	2,5	14,63	
12	Quadratisch	5	13,43	
13	Quadratisch	10	12,43	6.9(a)
14	TV	0,001	23,19	
15	TV	0,01	23,78	
16	TV	0,1	25,05	6.8(b)
18	TV	0,3	21,4	
19	TV	0,5	18,62	
20	TV	0,7	16,89	
21	TV	1	15,26	6.9(b)
22	TV	10	10,05	
23	TV	2	12,59	

Tabelle B.4: Anwendung der L^2 -Steuerbeschränkung auf ungestörtes Testbild 6.2a.

Bildindex	R	PSNR	KG1	KG2	KG3	Abbildung
1	2^{-5}	11,31	61,14	69,94	67,21	
2	2^{-3}	14,52	39,91	38,39	38,36	6.10(c)
3	2^{-1}	28,06	11,88	18,97	19,55	6.10(a)
4	1,00	∞	20,25	39,37	29,02	
5	2,00	∞	20,25	39,37	29,02	
6	0,25	17,82	18,28	12,63	18,01	
7	2^{-4}	12,58	52,64	55,66	53,73	6.10(d)
31	0,40	23,19	11,60	14,89	18,33	
32	0,30	19,34	13,65	12,54	17,83	6.10(b)
33	0,20	16,47	25,68	20,12	22,57	
34	0,10	13,80	43,91	43,67	43,55	
35	0,05	12,10	55,58	60,61	60,39	
36	0,35	21,11	12,06	13,84	18,19	
37	0,38	22,31	11,71	14,49	18,40	

Tabelle B.5: Anwendung der L^∞ -Steuerbeschränkung auf gestörtes Testbild aus Abbildung 6.2b.

Index	R_1	R_2	PSNR	KG1	Abbildung
1	1	1	22,76	25,35	
2	0,75	0,75	21,26	24,77	6.11(a)
3	0,5	0,5	18,61	24,23	6.11(b)
4	0,25	0,25	15,81	29,21	6.11(c)
5	0,15	0,15	12,94	29,63	6.11(d)
6	0,1	0,1	12,98	36,88	6.12(a)
7	0,05	0,05	11,54	42,32	6.12(b)
13	1	0,5	20,02		
14	1	0,25	17,48		
15	1	0,1	13,81		
16	1	0,05	13,16		
17	1	0,01	11,07		

Tabelle B.6: Anwendung von TV-Regularisierung (mit modifizierter L^1 -Norm und $\beta = 0.001$ und Regularisierungsparameter $\mu = 0.1$) und L^∞ -Steuerbeschränkung auf gestörtes Testbild aus Abbildung 6.2b.

Index	R_1	R_2	PSNR	KG1	Abbildung
18	1	1	25,05	25,35	
19	0,9	0,9	25,04	25,26	
20	0,7	0,7	22,56	23,9	6.13(a)
21	0,5	0,5	19,39	21,38	6.13(b)
22	0,3	0,3	16,91	23,68	6.13(c)
23	0,2	0,2	15,48	30,58	6.13(d)

Tabelle B.7: Anwendung des Zwei-Phasen-Verfahrens auf gestörtes Testbild aus Abbildung 6.2b: Phase 1 mit L^∞ -Steuerbeschränkung und TV-Regularisierung, wie in Tabelle B.6. Anschließend Phase 2: Anwendung von TV-Regularisierung (mit modifizierter L^1 -Norm und $\beta = 0.001$ und Regularisierungsparameter μ).

Index	Phase 1: R	Phase 2: μ	Kanten	Abbildung
3	0,9	0,2	24,73	
4	0,9	0,3	25,19	
5	0,9	0,1	24,71	6.14(a)
6	0,7	0,1	23,29	6.14(b)
7	0,7	0,2	23,22	
8	0,7	0,3	23,58	
9	0,7	0,5	25,38	
10	0,5	0,1	20,95	6.14(c)
11	0,5	0,2	21,18	
12	0,5	0,3	21,18	
13	0,5	0,5	23,96	
14	0,3	0,1	22,93	
15	0,3	0,2	22,74	
16	0,3	0,3	22,79	6.14(d)
17	0,3	0,4	23,01	
18	0,3	0,5	23,32	
19	0,3	1	25,49	

Tabelle B.8: Anwendung der L^∞ -Steuerbeschränkung auf das Streifenbild aus Abbildung 6.16a mit Salt-and-Pepper-Noise unterschiedlicher Rauschintensität

Rauschintensität	Index	R_1	R_2	PSNR	Abbildung	Datentreueterm
0,1	1	1	1	13,4		L^2
0,1	2	0,5	1	16,16		L^2
0,1	3	0,25	1	17,34		L^2
0,1	4	0,1	1	19,26		L^2
0,1	5	0,01	1	22,89		L^2
0,1	6	0,001	1	24,59		L^2
0,1	7	0,5	1	18,02		L^1
0,1	8	0,25	1	21,94	6.16b(b)	L^1
0,1	9	0,1	1	28,01	6.16c(c)	L^1
0,1	10	0,001	1	41,86	6.16d(d)	L^1
0,2	12	0,001	1	44,68		L^1
0,4	13	0,001	1	39,47	6.17(a)	L^1
0,4	14	0,0001	1	41,32		L^1
0,6	15	0,001	1	14,32	6.17(b)	L^1
0,9	16	0,001	1	7,28		L^1

Tabelle B.9: Anwendung des Steuerungsproblems mit rotating mask auf das Quadratbild aus Abbildung 6.21a

Index	R_1	PSNR	Abbildung
1	0,01	68,68	6.21(d)
2	0,05	34,92	6.21(c)
3	0,1	29,25	6.21(b)
4	0,2	23,81	
5	0,5	20,12	

Tabelle B.10: Anwendung der Methode von Bar/Kiryati/Sochen auf eine Struktur mit kleinen Variationen und Impulsrauschen.

Index	c_1	c_2	c_3	ϵ	PSNR	Abbildung
1	1	1	1	0,05	37,11	
2	1	2	1	0,05	38,83	6.24
3	1	5	1	0,05	36,87	
4	1	0,5	1	0,05	32,75	
5	1	1	0,1	0,05	30,62	
6	1	1	0,5	0,05	37,05	
7	1	2	0,5	0,05	38,82	
8	1	2	1,5	0,05	38,83	
9	1	2	2	0,05	38,83	
10	1	2	1	0,05	38,8	
11	1	2	1	1	38,77	
12	1	2	1	0,01	38,83	
13	1	2	1	0,001	38,83	
14	1	1,5	1	0,05	38,55	
15	1	2,2	1	0,05	38,79	
16	1	2,5	1	0,05	38,65	
17	1	2	1,2	0,05	38,83	

Tabelle B.11: Anwendung der Methode von Nikolova auf eine Struktur mit kleinen Variationen und Impulsrauschen.

Index	μ	PSNR	Abbildung
1	0,1	10,02	
2	0,2	23,38	
3	0,5	38,89	
4	0,4	33,68	
5	0,6	38,32	
6	0,45	39,09	6.25
7	0,47	39,03	
8	0,43	39,06	
9	1	26,37	
10	5	26,55	

Tabelle B.12: Anwendung der rotating mask auf eine Struktur mit kleinen Variationen und Impulsrauschen.

Index	R_1	R2	PSNR	Abbildung
1	0,5	2	32,58	
2	0,25	2	40,33	6.26
3	0,2	2	34,23	
4	0,3	2	39,66	
5	0,27	2	37,92	
6	0,23	2	40,25	
7	0,25	2	40,25	

Tabelle B.13: Anwendung der Methode von Bar/Kiryati/Sochen auf das verrauschte Lena-Bild aus Abbildung 6.27a bzw. Abbildung 6.29a.

Index	Rauschintensität	c_1	c_2	c_3	ϵ	PSNR	Abbildung
1	0,01	1	1	1	0,05	35,33	
2	0,01	1	0,5	1	0,05	31,98	
3	0,01	1	2	1	0,05	33,67	
4	0,01	1	1,5	1	0,05	34,98	
5	0,01	1	1,25	1	0,05	35,41	
6	0,01	1	3	1	0,05	31,43	6.28(a)
7	0,01	1	5	1	0,05	28,76	
8	0,01	1	5	1	0,05	28,76	
9	0,01	1	10	1	0,05	25	
10	0,01	1	2	0,01	0,05	24,15	6.29(d)
11	0,01	1	2	0,5	0,05	33,84	6.28(b)
12	0,01	1	2	0,5	0,5	35,35	
13	0,01	1	1,5	0,5	0,5	35,86	
14	0,01	1	1,1	0,5	0,5	35,26	
30	0,1	1	1	0,5	0,05	27,51	
31	0,1	1	2	0,5	0,05	30,05	
32	0,1	1	5	0,5	0,5	25,29	
33	0,1	1	5	0,5	0,05	28,24	6.29(c)
34	0,1	1	2	0,5	0,01	30	
35	0,1	1	3	0,5	0,05	29,82	

Tabelle B.14: Anwendung der Methode von Nikolova auf das verrauschte Lena-Bild aus Abbildung 6.27a bzw. Abbildung 6.29a.

Index	Rauschintensität	μ	PSNR	Abbildung
15	0,01	1	29,37	
16	0,01	0,1	3,93	
17	0,01	0,5	33,93	6.28(c)
18	0,01	0,4	35,03	6.28(d)
19	0,01	0,3	10,74	
20	0,01	0,75	31,23	
21	0,01	0,3	28,27	
22	0,01	1	31,59	
23	0,01	0,75	33,56	
26	0,1	0,5	29,78	
27	0,1	1	30,15	
28	0,1	1,5	28,24	6.29(b)
29	0,1	1	4,12	

Tabelle B.15: Anwendung des Steuerungsproblems mit rotating mask auf das Lena-Bild aus Abbildung 6.27a bzw. Abbildung 6.29a

Index	Rauschintensität	R	PSNR	Abbildung
9	0,01	0,5	24,13	
10	0,01	0,2	25,96	
11	0,01	0,1	30,38	6.27d(a)
12	0,01	0,01	29,62	6.27d(d)
13	0,01	0,05	33,46	6.27d(b)
14	0,01	0,03	32,66	
15	0,1	0,05	18,74	
16	0,1	0,02	18,3	

Danksagung

An dieser Stelle möchte ich allen danken, die zum Gelingen dieser Arbeit beigetragen haben. Ich danke Herrn Prof. Dr. Maurer und Herrn Dr. Wagner für die gute Betreuung und die interessante Themenstellung, die ein sehr modernes und interessantes Gebiet der Mathematik behandelt. Ich danke Herrn Dr. Wagner weiterhin für die häufigen Besuche in Münster, die trotz der großen Entfernung eine gute Zusammenarbeit förderten. Ich danke meiner Schwester Marzena Franek und Christof Brune für die gute Zusammenarbeit und zahlreiche mathematische Diskussionen, die mir sehr weitergeholfen haben. Schließlich gilt mein besonderer Dank meiner Familie, die mir das Studium ermöglicht und mich während der Diplomarbeit unterstützt hat.

Eidesstattliche Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst habe und neben den Programmen AMPL, MATLAB und LOQO/IPOPT keine weiteren als die im Literaturverzeichnis angegebenen Hilfsmittel verwendet habe. Zudem sind alle Programme zur Benutzung von LOQO/IPOPT von mir selbst erstellt worden und auf einer CD beigelegt.

Münster, 22. November 2007

Unterschrift

Abbildungsverzeichnis

2.1	Überblick über unterschiedliche Bildrepräsentationen in dieser Arbeit	13
4.1	Vergleich von quadratischer und TV-Regularisierung an einem verrauschten Signal: Während die quadratische Regularisierung aus scharfen Kanten Kurven macht, neigt die TV-Regularisierung zu „staircaising“.	24
4.2	L^1 -Datenterm garantiert im Gegensatz zum L^2 -Datenterm die Kontrastinvarianz	33
4.3	Signal	38
4.4	Anwendung der Rotationsmatrix	38
6.1	Überblick über die Durchführung der Versuchsreihen	54
6.2	Synthetisches Testbild	57
6.3	Variiere c_2 : c_2 wird von oben nach unten größer. Andere Parameter bleiben fest. $c_1 = 1, c_3 = 0,001, c_4 = 0, \epsilon = 0,05$	59
6.4	Variiere c_3 : c_3 wird von oben nach unten kleiner. Andere Parameter bleiben fest. $c_1 = 1, c_2 = 0,001, c_4 = 0, \epsilon = 0,05$	60
6.5	Variiere c_4 : c_4 wird von oben nach unten größer. Andere Parameter bleiben fest. $c_1 = 1, c_2 = 0,001, c_3 = 0,001, \epsilon = 0,05$	61
6.6	Variiere ϵ : ϵ ist im linken Bild größer als im rechten. Andere Parameter bleiben fest. $c_1 = 1, c_2 = 0,001, c_3 = 0,00001, c_4 = 0$	62
6.7	Entrauschen und gleichzeitige Kantenerkennung des verrauschten Testbildes.	62
6.8	Glättung basierend auf dem Variationsansatz	64
6.9	Glättung basierend auf dem Variationsansatz mit größerem Regularisierungsparameter	64
6.10	Steuerbeschränkung im Fall $q = 2$. Anwendung auf ein ungestörtes Testbild aus Abb. 6.2a	68
6.11	Anwendung auf verrauschtes Testbild aus Abb. 6.2b mit Maximumnorm und Schranke $R := R_1 = R_2$	69
6.12	Anwendung auf verrauschtes Testbild aus Abb. 6.2b mit Maximumnorm	70
6.13	TV-Regularisierung mit $\beta = 0.001, \mu = 0.1$ und L^∞ -Steuerbeschränkung mit $R := R_1 = R_2$	71
6.14	Zwei-Phasen-Verfahren: In der rechten Spalte die Kantenbilder nach der ersten Phase, in der linken Spalte die Kantenbilder nach der zweiten Phase, d.h. nach Anwendung der TV-Regularisierung mit Regularisierungsparameter μ	72
6.15	Adjungierte des Zustands bei L^∞ -Steuerbeschränkung	73
6.16	Streifenbild mit L^∞ -Steuerbeschränkung und $R_2 = 1$	74
6.17	Streifenbild mit L^∞ -Steuerbeschränkung und $R_1 = 10^{-3}, R_2 = 1$. In der linken Spalte jeweils das verrauschte Bild. In der rechten Spalte das wiederhergestellte Bild.	75
6.18	Signal, Original und verrauscht	76
6.19	In der linken Spalte wurde die modifizierte L^1 -Norm mit $\beta = 0.0001$ für den Datenterm verwendet. In der rechten Spalte wurde die L^2 -Norm verwendet. Die Ergebnisse wurden mit IPOPT erzielt. Der Winkel θ wurde in $[0, \frac{\pi}{2}]$ gewählt.	78
6.20	$\beta = 0.001$ Ergebnisse mit IPOPT binäre Variable	79
6.21	Anwendung des Steuerungsproblems mit rotating mask auf einfaches Testbild.	79
6.22	Anwendung des Steuerungsproblems mit rotating mask auf einfaches Testbild mit höherer Rauschintensität.	80
6.23	Eine Struktur mit kleinen Variationen, verrauscht mit Impulsrauschen	81
6.24	Methode von Bar/Kiryati/Sochen, PSNR=38,83	82

Abbildungsverzeichnis

6.25 Methode von Nikolova, PSNR=39,09	82
6.26 Steuerungsproblem mit rotating mask, PSNR=40,33	82
6.27 Anwendung des Steuerungsproblems mit rotating mask auf verrauschtes Lena-Bild	84
6.28 Anwendung der Verfahren von Nikolova bzw. Bar/Kiryati/Sochen auf verrauschtes Lena-Bild	85
6.29 Anwendung der Verfahren von Nikolova bzw. Bar/Kiryati/Sochen auf Lena-Bild mit 10 % Rauschintensität	86

Literaturverzeichnis

- [1] W. Alt: *Nichtlineare Optimierung* Vieweg, Braunschweig, (2002)
- [2] L. Ambrosio, V.M. Tortorelli: *Approximation of functionals depending on jumps by elliptic functionals via Γ -convergence*. Communications on Pure and Applied Mathematics **43**, (1990) 999–1036
- [3] G. Aubert, P. Kornprobst: *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*. Springer; New York (2002)
- [4] A. Arnold: *Numerik partieller Differentialgleichungen*. Vorlesungen an der Westfälischen Wilhelms-Universität Münster, FS 2002
- [5] L. Bar, N. Kiryati, N. Sochen: *Image Deblurring in the Presence of Impulsive Noise* International Journal of Computer Vision **70** (2006), 279-298
- [6] B. Bourdin: *Image segmentation with finite element method*. M2AN Mathematical Modelling and Numerical Analysis **33**(1999),229-244
- [7] A. Buades, B. Coll, J.-M. Morel: *A review of image denoising algorithms, with a new one*. Multiscale Model Simul. **4**(2005),490-530
- [8] M. Burger: *Inverse problems*. Vorlesungen an der Universität Linz, FS 2005
- [9] M. Burger: *Mathematische Bildverarbeitung*. Vorlesungen an der Westfälischen Wilhelms-Universität Münster, FS 2007
- [10] M. Burger: *Numerik partieller Differentialgleichungen*. Vorlesungen an der Westfälischen Wilhelms-Universität Münster, FS 2006
- [11] J.F. Cai, R.H. Chan, M. Nikolova: *Two-Phase Methods for Deblurring Images Corrupted by Impulse Plus Gaussian Noise*. Elektronisch zugänglich unter <http://www.cmla.ens-cachan.fr/Utilisateurs/nikolova/publisM.html> (Zugriff am 20.9.2007)
- [12] A. Chambolle, P.L. Lions: *Image recovery via total variation minimization and related problems* Numerische Mathematik **76** (1997), 167-188
- [13] T.F. Chan, S. Esedoğlu: *Aspects of Total Variation Regularized L^1 -Function Approximation*. CAM-Report **(04-07)** (2004), elektronisch zugänglich unter <http://www.math.ucla.edu/applied/cam/> (Zugriff am 20.9.2007)
- [14] R.H. Chan, C.-W. Ho, M. Nikolova *Salt-and-Pepper Noise Removal by Median-type Noise Detectors and Detail-preserving Regularization* IEEE Transactions on Image Processing **14** (2005), 1479-1485.
- [15] T.F. Chan, J. Shen: *Image Processing and Analysis*. SIAM; Philadelphia (2005)
- [16] B. Dacorogna: *Direct Methods in the Calculus of Variations*. Springer, New York, (1989)
- [17] B. Dacorogna: *Introduction to the Calculus of Variations*. Imperial College Press, London, (2004)
- [18] G. Dewess, P. Helbig: *Einschätzen und Optimieren von Verkehrsflüssen*. In: A. Bachem, M. Jünger, R. Schrader (Hg.): *Mathematik in der Praxis*. Springer, Berlin (1995) 473-492
- [19] D.C. Dobsen, C.R. Vogel: *Convergence of an Iterative Method for Total Variation Denoising*. SIAM Journal on Numerical Analysis **34** (1997), 1779-1791
- [20] R. Fourer, D.M. Gay, B.W. Kernighan: *AMPL: A Modelling Language for Programming*. Duxbury Resource Center (2003)
- [21] H. Maurer: *Optimierung*. Vorlesungen an der Westfälischen Wilhelms-Universität Münster, FS 2005

- [22] M. Nikolova: *A Variational Approach to Remove Outliers and Impulsive Noise*. Journal of Mathematical Imaging and Vision **20** (2004), 99-120
- [23] M. Nikolova: *Minimizers of cost-functions involving non-smooth data-fidelity terms. Application to the processing of outliers*. SIAM Journ. on Numerical Analysis **40** (2001), 965-994
- [24] S. Pickenhain, M. Wagner: *Critical points in relaxed deposit problems*. In: A. Ioffe, S. Reich, I. Shafirir (Eds.): *Calculus of Variations and Optimal Control*, Technion 98, Vol. II (Research Notes in Mathematics, Vol. 411). Chapman & Hall/CRC Press; Boca Raton etc. 2000, 217 – 236.
- [25] L.I. Rudin, S.J. Osher, E. Fatemi: *Nonlinear total variation based noise removal algorithms*. Physica D **60** (1992), 259-268.
- [26] K. Theissen: *Optimale Steuerprozesse unter partiellen Differentialgleichungs-Restriktionen mit linear eingehender Steuerfunktion*. Dissertation, Westfälische Wilhelms-Universität Münster, Institut für Numerische und Angewandte Mathematik (2006)
- [27] R. J. Vanderbei: *LOQO: An interior point code for quadratic programming*. Statistics and Operations Research. Princeton University (1998)
- [28] C.R. Vogel, M. E. Oman: *A fast robust algorithm for total variation based reconstruction of noisy, blurred images* IEEE Trans. Image Processing to appear (1998)
- [29] C.R. Vogel, M. E. Oman: *Iterative Methods for Total Variation Denoising*. SIAM Journal of Scientific Computing **17** (1996), 227-238
- [30] A. Wächter, LT. Biegler: *On the implementation of a interior-point filter line-search algorithm for large-scale non linear programming*. Mathematical Programming **106** (2006) 25-57
- [31] M. Wagner: *Mehrdimensionale Steuerungsprobleme mit quasikonvexen Integranden*. Habilitationsschrift. BTU Cottbus 2006
- [32] M. Wagner: *Pontryagin's maximum principle for multidimensional control problems in image processing. Updated version. BTU Cottbus, Preprint*. Reihe Mathematik M-10/2007