

Fakultät für Physik und Astronomie

Ruprecht-Karls-Universität Heidelberg

Diplomarbeit
im Studiengang Physik

vorgelegt von
Horst Haußecker
aus Eppingen

1993

doi: 10.5281/zenodo.14550

Mehrgitter-Bewegungssegmentierung
in Bildfolgen
mit Anwendung zur
Detektion von Sedimentverlagerungen

Die Diplomarbeit wurde ausgeführt am
Institut für Umweltphysik
unter der Betreuung von
Priv.-Doz. Dr. B. Jähne

Gutachter:
Prof. Dr. U. Platt

*Leave the beaten track occasionally,
and dive into the woods.
You will be certain to find something
that you have never seen before.*

Alexander Graham Bell

Zusammenfassung

In der vorliegenden Arbeit wird ein bildanalytisches Meßverfahren zur automatischen Detektion von Bewegung in Bildfolgen vorgestellt. Diese qualitative Bewegungssegmentierung stellt einen wichtigen Verarbeitungsschritt der low-level Bildverarbeitung dar und wird so optimiert, daß die Ergebnisse denen des menschlichen visuellen Systems entsprechen. Neben klassischen Filtertechniken werden Ansätze der künstlichen Intelligenz in Form von Fuzzy-Logik verwendet, um dieses Ziel zu erreichen. Das Ergebnis ist ein robustes Verfahren, welches selbst mit bloßem Auge schwer zu erkennende Bewegung detektiert.

Eine zeitliche Faltung der Bildsequenz mit Binomialfiltern extrahiert die Strukturen des ruhenden Hintergrundes. Die Abhängigkeit der benötigten Filtergrößen von der Größe und Geschwindigkeit der Objekte wird analytisch hergeleitet. Eine neue Mehrgitterstruktur reduziert den Rechenaufwand um fast zwei Größenordnungen. Die damit erhaltenen Hintergrundbilder werden anschließend mit den Bildern der Originalsequenz verglichen. Mit Hilfe von adaptiver Fuzzy-Logik wird eine lokale Schwelle ermittelt um sich bewegende Objekte zu detektieren.

Zu Testzwecken wird der Algorithmus sowohl auf computergenerierte Testsequenzen als auch auf Realweltszenen angewendet. Dabei zeigt sich, daß die Ergebnisse gut mit den theoretisch ermittelten Werten übereinstimmen und sich das Verfahren ohne eine Veränderung der Parameter sowohl auf die Beobachtung von Sandpartikeln als auch auf Verkehrsszenen anwenden läßt.

Als eine mögliche Anwendung des Verfahrens wird die Verlagerung von Sandkörnern in Flußsedimenten untersucht. Dabei wurden Voruntersuchungen zu Messungen durchgeführt, die im Rahmen eines Projektes in Zusammenarbeit mit der Bundesanstalt für Wasserbau erfolgen. Dessen Zielsetzung besteht darin, durch eine direkte Beobachtung der Grenzschichten zwischen verschiedenen Sedimenten, aus den Bewegungen der einzelnen Sandpartikeln in Abhängigkeit von Druckschwankungen im Boden, Rückschlüsse über mögliche Instabilitäten der Uferbefestigungen zu erhalten.

Der Versuchsaufbau besteht aus einem druckdichten Stahltank, der mit verschiedenen Sedimentschichten und Wasser gefüllt wird. Der Tankinnendruck kann durch ein Pneumatiksystem beliebig verändert und damit dem hydrostatischen Druck der Wassersäule in den Sedimentschichten additiv überlagert werden. Die Bildaufnahme erfolgt durch ein starres Endoskop mit integrierter Glasfaserbeleuchtung und angekoppelter Videokamera. Das Endoskop wird durch eine spezielle Klemmvorrichtung fest mit dem Tank verbunden und liefert Bilder der gewünschten Grenzschichten zwischen verschiedenen Sedimenten. Der gesamte Bildausschnitt beträgt dabei $8 \times 6 \text{ mm}$. Dadurch wird eine Beobachtung der mikroskopischen Vorgänge in der unmittelbaren Umgebung der Grenzschichten ermöglicht.

Zur Beurteilung der Meßergebnisse wird die beobachtete Bewegungshäufigkeit mit der Bewegung verglichen, die bei einer stabilen Filterschicht zu erwarten ist, wenn nur Luftblasen komprimiert und dekomprimiert werden. Es zeigt sich, daß dieser theoretische Verlauf für stabile Filterschichten gut erfüllt wird, während instabile Filterschichten ein deutliches Anwachsen der Bewegung bei schnellen Druckänderungen aufweisen. Dies ist ein deutliches Zeichen für das Auftreten von Materialverlagerungen und damit verbundenen Instabilitäten innerhalb der Grenzschicht.

Inhaltsverzeichnis

| | | |
|----------|--|-----------|
| 1 | Einleitung | 1 |
| I | Bildfolgenanalyse | 3 |
| 2 | Problemstellung und Überblick | 5 |
| 2.1 | Spezielle Probleme des Bildmaterials | 5 |
| 2.2 | Andere Arbeiten zur Bewegungssegmentierung | 6 |
| 3 | Zeitliche Tiefpaßfilterung | 10 |
| 3.1 | Projektionsfilter | 10 |
| 3.2 | Zeitlicher Tiefpaß | 12 |
| 3.3 | Ideale periodische Struktur | 14 |
| 3.3.1 | Größe der Faltungsmaske | 15 |
| 3.4 | Kreisförmiges Objekt | 16 |
| 3.4.1 | 3D-Fouriertransformierte | 16 |
| 3.4.2 | Größe der Faltungsmaske | 17 |
| 3.5 | Verallgemeinerung für beliebige Objekte | 21 |
| 4 | Zeitliche Gaußpyramide | 23 |
| 4.1 | Allgemeine Gaußpyramiden | 23 |
| 4.2 | Aufbau durch gezieltes Unterabtasten | 24 |
| 4.3 | Effektivität der Filterkaskade | 27 |
| 4.4 | 2D-Pyramiden | 29 |
| 4.5 | Anwendung auf die zeitliche Tiefpaßfilterung | 30 |

| | | |
|-----------|--|-----------|
| 5 | Vergleich zwischen Originalbild und Hintergrund | 32 |
| 5.1 | Idealer Hochpaßfilter | 32 |
| 5.2 | Approximation | 34 |
| 5.3 | Austausch der Koordinaten | 35 |
| 5.3.1 | Ideale periodische Struktur | 35 |
| 5.3.2 | Kreisförmiges Objekt | 36 |
| 5.4 | Motion blur | 37 |
| 5.5 | Verallgemeinerung für beliebige Objekte | 38 |
| 5.6 | Größe des Ähnlichkeitsmaßes für die beiden Beispiele | 40 |
| 5.6.1 | Periodische Struktur | 40 |
| 5.6.2 | Kreisförmiges Objekt | 41 |
| 5.6.3 | Schwelle für beliebige Objekte | 41 |
| 5.7 | Überprüfung der Ergebnisse an Testbildern | 41 |
| 6 | Lokale Schwelle durch Fuzzy-Logik | 44 |
| 6.1 | Meß- und Regelgrößen der Bewegungssegmentierung | 44 |
| 6.1.1 | Berechnung der Differenz des lokalen mittleren Grauwertes (GD) | 45 |
| 6.1.2 | Berechnung der Differenz der lokalen Varianz (VD) | 45 |
| 6.2 | Fuzzy-Logik | 46 |
| 6.2.1 | Fuzzifizierung | 49 |
| 6.2.2 | Logische Verknüpfung | 50 |
| 6.2.3 | Defuzzifizierung | 51 |
| 6.2.4 | Fuzzy-LUT | 52 |
| 6.2.5 | Optimierung der Fuzzy-Logik | 53 |
| 7 | Zusammenfassung und Ergebnisse der Bildverarbeitung | 55 |
| 7.1 | Zusammenfassung der Bildfolgenanalyse | 55 |
| 7.2 | Ergebnisse der Bildfolgenanalyse | 56 |
| II | Messung & Auswertung | 61 |
| 8 | Problemstellung der Messung | 63 |
| 8.1 | Motivation | 63 |
| 8.2 | Mechanische Belastungen in Flußufern | 64 |

| | |
|--|-----------|
| 9 Experimentelle Durchführung | 68 |
| 9.1 Sedimenttank | 68 |
| 9.2 Endoskop | 70 |
| 9.3 Aufbau des Pneumatiksystems | 72 |
| 9.3.1 Druckanstieg | 72 |
| 9.3.2 Druckabfall | 73 |
| 9.3.3 Meßwerte | 74 |
| 9.4 Versuchsanordnung und Synchronisation | 75 |
| 9.5 Eichung der Geschwindigkeit | 76 |
| 10 Ergebnisse der Messung | 77 |
| 10.1 Schichtung der Sedimente | 77 |
| 10.2 Größe der Druckschwankungen | 79 |
| 10.3 Auswertung der Bewegungssegmentierung | 79 |
| 10.4 Abhängigkeit der Bewegung von der Druckschwankung | 80 |
| 10.4.1 Modellierung der Druckabhängigkeit von Bewegung | 80 |
| 10.5 Räumliche Lokalisierung der Bewegung | 84 |
| 11 Ausblick | 90 |
| A Rekursives Filter | 91 |
| B Rauschen | 93 |
| C Varianz von Faltungsmasken | 95 |
| C.1 Skalieren von Masken | 95 |
| C.2 Kaskadierung von Faltungen | 95 |

Kapitel 1

Einleitung

Bewegung mit all ihren Parametern, wie Betrag und Richtung, ist eine wichtige Meßgröße zur Interpretation dynamischer physikalischer Vorgänge. Die Rekonstruktion des dreidimensionalen Geschwindigkeitsfeldes aus einer Folge von Bildern, stellt seit mehr als 10 Jahren eine der größten Herausforderungen für die Digitale Bildfolgenanalyse dar.

Die Projektion der dreidimensionalen Welt auf die zweidimensionale Bildebene macht eine vollständige Rekonstruktion der Vorgänge in den meisten Fällen unmöglich. Dies führt zu Problemen, die in der Literatur als Korrespondenzproblem, Apertur- bzw. Blendenproblem, Okklusion und Bewegungsdiskontinuitäten bekannt sind, um nur einige wenige zu nennen.

Ein weiteres Problem folgt aus der Tatsache, daß kontinuierliche Größen digitalisiert werden und somit nur als diskrete Meßpunkte in Raum und Zeit zur Verfügung stehen. Das Abtasttheorem stellt eine wichtige Zwangsbedingung an die Auflösung der Meßpunkte dar, um sicherzustellen, daß beim Digitalisieren keine Information verloren geht. Während das räumliche Abtasttheorem in den meisten Fällen eingehalten werden kann, ist die zeitliche Abtastfrequenz durch die Bildwiederholrate der Videonorm vorgegeben und läßt sich nur in wenigen Ausnahmen erhöhen. Doch selbst in diesen Fällen geht die Verbesserung der zeitlichen Auflösung zu Lasten der räumlichen Auflösung der Bilder, da die Gesamtzahl der Meßpunkte pro Zeit eine bestimmte Grenze nicht überschreiten kann. Somit lassen sich schnelle Vorgänge nicht immer mit der benötigten zeitlichen Auflösung beobachten, was eine Geschwindigkeitsbestimmung oft unmöglich macht.

Selbst wenn das raum-zeitliche Abtasttheorem vollständig erfüllt wird, stellt die Diskretisierung der Meßdaten ein Problem dar. Die Approximation der benötigten mathematischen Operatoren auf einem diskreten Gitter ist in den seltensten Fällen ohne Einschränkung der Genauigkeit möglich. Dies trifft vor allem für alle Arten von Differentialoperatoren zu, deren mathematische Definition auf einem Grenzübergang beruht, der nur im Fall infinitesimal benachbarter Punkte das gewünschte Ergebnis liefert. Da dynamische Vorgänge immer auf zeitlichen Ableitungen der beobachtbaren Raum-Koordinaten beruhen, stellen diese Rechenoperationen immer Fehlerquellen dar.

Obwohl die Verfahren zur Geschwindigkeitsanalyse immer weiter verbessert worden sind, indem Zwangsbedingungen eingeführt und spezielle Eigenschaften von Objekten berücksichtigt wurden,

ist eine vollständige Rekonstruktion des dreidimensionalen Geschwindigkeitsfeldes nur in den seltensten Fällen möglich. Ein allgemeines Verfahren, welches auch nur annähernd die Fähigkeiten des menschlichen visuellen Systems hat, Bewegung zu erkennen und in Betrag und Richtung abzuschätzen, ist noch lange nicht in Sicht. Selbst die rasend schnelle Leistungsverbesserung der Rechnerhardware konnte daran nichts ändern.

Eine Folgerung aus diesen Überlegungen ist, daß vor der Verarbeitung der Bilddaten zunächst untersucht werden muß, welche physikalischen Informationen sich sinnvollerweise aus den Bildern entnehmen lassen. Wenn eine quantitative Analyse der Geschwindigkeit nicht möglich ist, bietet es sich an, eher qualitative Aussagen über den Inhalt der Bildsequenz zu machen. Zu dieser Schlußfolgerung kommen auch *Verri und Poggio* [Verri, Poggio, 1987]. Sie zeigen, daß der *optische Fluß*, d. h. das 2D-Verschiebungsvektorfeld, das aus den Variationen der Grauwertverläufe in Bildfolgen gewonnen werden kann und die zweidimensionale Projektion des 3D-Geschwindigkeitsfeldes auf die Bildebene generell verschieden sind. Nur unter bestimmten Voraussetzungen sind sie gleich. Als Konsequenz schlagen sie vor, nur qualitative Eigenschaften des optischen Flusses zu extrahieren und diese mit qualitativen Eigenschaften dreidimensionaler Geschwindigkeitsfelder zu vergleichen [Verri, Poggio, 1989]. Geeignete Eigenschaften sind dabei die Anzahl und Art von Singularitäten, die Aufschluß darüber geben, ob der optische Fluß topologisch äquivalent zum Geschwindigkeitsfeld ist. *François und Bouthemy* [François, Bouthemy, 1990] benutzen ein ähnliches Verfahren, um qualitative Aussagen über die Kinematik von Objekten zu machen.

In vielen Fällen ist es jedoch gar nicht notwendig, Geschwindigkeiten von Objekten in Bildern exakt zu bestimmen. Oft reicht eine rein qualitative Analyse aus, die Bildbereiche bestimmt, in denen Bewegung stattfindet, d. h. eine *Bewegungssegmentierung*.

Beispiele für solche Anwendungen sind die Überwachung von Verkehrsanlagen und Bahnsteigen [Haass, 1984] oder die Kontrolle von Fußgängerwegen [Shio, Sklansky, 1991]. Bei physikalischen Anwendungen liefert die Beobachtung von Bewegung wichtige Hinweise auf die Anwesenheit externer Kräfte, die diese Bewegung hervorrufen.

Die Bildverarbeitungsalgorithmen, die in dieser Arbeit entwickelt wurden, dienen dazu die Bewegung von Sandpartikel in Grenzschichten zwischen verschiedenen Bodenschichten zuverlässig zu segmentieren. Dieser Anwendungsfall unterscheidet sich von den meisten Beispielen der Digitalen Bildverarbeitung durch die schlechte Qualität des Bildmaterials und physikalischen Randbedingungen, welche eine Anwendung von konventionellen Algorithmen der Bildfolgenanalyse unmöglich machen. Das Ziel war, trotzdem eine möglichst gute Segmentierung derjenigen Bildbereiche durchzuführen, in denen auch vom menschlichen Betrachter Bewegung erkannt wird. Die Algorithmen sind jedoch nicht speziell auf diese Anwendung zugeschnitten und lassen sich auch in vielen anderen Bereichen der physikalischen Forschung einsetzen, in denen eine Detektion von Bewegung nötig ist. Beispiele für weitere Einsatzmöglichkeiten sind die Untersuchung von Phasentübergängen, Schmelzprozessen und Kristallbildung oder allgemein alle Prozesse, bei denen eine zeitliche Strukturumwandlung detektiert werden soll.

Teil I

Bildfolgenanalyse

Kapitel 2

Problemstellung und Überblick

Bevor ein kurzer Überblick über den aktuellen Stand der Forschung anhand einiger Arbeiten zur Bewegungssegmentierung erfolgt, wird zunächst beschrieben, worin die Schwierigkeiten des verwendeten Bildmaterials liegen.

2.1 Spezielle Probleme des Bildmaterials

Sandbewegungen in Bodenschichten weisen einige Eigenschaften auf, die es erschweren, ihre Bewegung zu detektieren und eine Verfolgung über mehrere Bilder unmöglich machen:

- Die Ablösung einzelner Teilchen oder ganzer Bodenschichten, die durch eine Strömung hervorgerufen wird, ist ein typischer Losreißeffekt. Die Teilchen bleiben bis zu einem kritischen Punkt völlig bewegungslos. Sobald sie von der Strömung mitgerissen werden, sind sie nach kurzer Zeit (2 - 3 Bilder im Videotakt) bereits so schnell, daß sie das zeitliche Abtasttheorem nicht mehr erfüllen und eine Verfolgung unmöglich wird. Im Moment des Losreißen ist ihre Bewegung kaum zu erkennen und deshalb schwer zu detektieren. Durch die hohe Geschwindigkeit kurz nach der Ablösung verteilen sie ihren Grauwert während der Belichtungszeit über eine so große Fläche des Bildsensors, daß ihre Helligkeit fast im Rauschen des ruhenden Hintergrundes untergeht. Dieser Effekt ist in der Bildverarbeitung als Bewegungsunschärfe (engl. motion blur) bekannt, welche die Ränder von sich bewegenden Objekten sowohl räumlich als auch zeitlich verschmiert. In beiden Fällen, im Moment der beginnenden Bewegung wie auch bei hohen Geschwindigkeiten, sollen die Teilchen über ihre gesamte Fläche detektiert werden.
- Die Bewegung der Sandpartikel entspricht in Betrag und Richtung der Geschwindigkeit der Wasserströmung im Untergrund. Da diese auf den Skalen der Sedimente stark turbulent ist, können keine Objektverfolgungsalgorithmen angewendet werden, die von einer langsamen und stetigen Änderung der Geschwindigkeit ausgehen.
- Da die Szene durch ein Endoskop beobachtet wird (Abschnitt 9.2), ist die Bewegung der Sedimentteilchen auf die 2D-Ebene des Beobachtungsfensters eingeschränkt. Selbst wenn

die Geschwindigkeit exakt bestimmbar wäre, könnte man dadurch keine Aussagen über die tatsächliche 3D-Bewegung ohne Anwesenheit des Endoskopes machen.

Zusätzlich zu diesen Schwierigkeiten, die durch die physikalischen Randbedingungen verursacht werden, gibt es noch Probleme, die mit der Beleuchtung der Szene zusammenhängen:

- Um überhaupt ein Bild von Querschnitten der Sedimentgrenzschichten zu erhalten, muß der Bildausschnitt beleuchtet werden. Dies geschieht durch ein Glasfaserbündel, das durch das Endoskop hindurchgeführt wird. Die Austrittsöffnung dieses Lichtleiters liegt direkt neben dem Fenster der Optik (Abb. 9.2). Da das austretende Lichtbündel stark gerichtet ist und die Sandteilchen nur einen geringen Abstand von der Optik haben, wird der Bildausschnitt nur unzureichend homogen ausgeleuchtet. Es gibt dadurch Bildbereiche, die zu dunkel sind, während andere überbelichtet werden.
- Da sich zwischen den Objekten und der Beleuchtung eine Glasscheibe befindet, spiegelt sich die Beleuchtung an dieser. Dies gibt zusätzlich Bereiche, in denen der Grauwert der Teilchen kaum zu erkennen ist.
- Die Sandteilchen selbst reflektieren das auf sie auftreffende Licht in einem Winkel der durch die Orientierung ihrer Oberflächennormale gegeben ist. Eine Rotation der Teilchen führt dazu, daß sich Spiegelreflexe mit einer Geschwindigkeit über das Bild bewegen, welche eine schnellere Bewegung der Teilchen vortäuscht. Zusätzlich beeinflußt die Reflexion des Lichts an den einzelnen Teilchen die Beleuchtung der gesamten Szene, so daß sich auch die Helligkeit ruhender Objekte ändert.

Die Detektion der sich bewegendenden Sandteilchen über ihre gesamte Fläche stellt somit aus verschiedenen Gründen Anforderungen an die Bildverarbeitung, die über die Eingangs erwähnten grundsätzlichen Schwierigkeiten der Bildfolgenanalyse hinausgehen. Die ausführliche Diskussion der Probleme zusammen mit dem nun folgenden Überblick über andere Arbeiten auf dem Gebiet der Detektion von Bewegung soll veranschaulichen, warum die Entwicklung der Bewegungssegmentierung, die in den nächsten Kapiteln beschrieben wird, nötig war.

2.2 Andere Arbeiten zur Bewegungssegmentierung

Während sich die Literatur der letzten Jahre immer mehr auf die quantitative Analyse von Bewegung konzentriert hat, wurde die reine Bewegungssegmentierung ein wenig vernachlässigt. Die folgende Übersicht stellt trotzdem nur eine kleine aber repräsentative Auswahl der Literatur zu diesem Thema dar.

Die meisten Ansätze verwenden statistische Tests zur Detektion der Bildbereiche, in denen Bewegung stattfindet.

Nagel [Nagel, 1982] benutzt den von *Yakimovsky* [Yakimovsky, 1976] zur Detektion einer Grauwertkante vorgeschlagenen Likelihood-Ratio Test (LRT), um die Differenzen zwischen zwei aufeinanderfolgenden Bildern zu bewerten. Er vergleicht die statistischen Eigenschaften von modellierten Grauwertverläufen in lokalen Umgebungen der beiden Bilder. Dabei werden die Grauwertverteilungen als konstant mit additivem Rauschen angenommen. Später wurde das Verfahren auf lineare und quadratische Grauwertverläufe erweitert [Nagel, 1982], [Hsu et al., 1984].

Wie alle Verfahren, die nur zwei Bilder vergleichen, fehlt hier die Information einer größeren zeitlichen Umgebung, um robust gegen Rauschen zu sein. Die detektierten Gebiete der sich bewegenden Objekte bleiben dabei auf die Objektränder beschränkt. Allgemein können dadurch nur Strukturen detektiert werden, die sich in der Zeit zwischen zwei Aufnahmen ausreichend weit verschoben haben.

Bouthemy [Bouthemy, 1989] erweitert diesen Ansatz auf eine 3D-raum-zeitliche Umgebung eines Pixels um die Anwesenheit einer räumlichen und zeitlichen Kante zu erkennen. Die Orientierung dieser Kante liefert die Geschwindigkeit senkrecht zur 2D-Projektion der Kante auf die Raum-Achsen.

Westberg [Westberg, 1992] benutzt einen hierarchischen Likelihood-Ratio Test um die Kanten eines bewegten Objektes auf einer räumlichen Pyramide zu detektieren. Er zeigt, daß eine Approximation des LRT auf ein Schwellwertverfahren führt, bei dem die quadrierte und über eine räumliche Umgebung gemittelte Differenz zweier Bilder mit einer bestimmten Schwelle verglichen wird. Diese Schwelle hängt von der jeweiligen Pyramidenstufe ab. Um mehr zeitliche Information zu erhalten schlägt er vor, einen Hintergrund aus den Segmentierungsergebnissen vorangegangener Bilder zu extrahieren.

Etoh und Shirai [Etoh, Shirai, 1993] benutzen zur Bewegungssegmentierung und Verfolgung von Objekten Eigenschaften von räumlichen und zeitlichen Umgebungen eines Objektes und suchen die Interpretation des Bildinhaltes, die am besten die gemessenen Größen beschreibt. Als Merkmale verwenden sie Farbe, lokale Orientierung, Größe von Pixelclustern, räumliche und zeitliche Ableitungen. Obwohl Farbe eine wichtige Rolle bei diesem Verfahren spielt und diese nur in den wenigsten Fällen zur Verfügung steht, hat dieser Ansatz die wichtige Eigenschaft, daß zeitliche Änderungen nicht unabhängig von der räumlichen Grauwertverteilung behandelt werden. Dies ist nötig um Pixelrauschen von Grauwertänderungen zu unterscheiden, die durch Bewegungen von Objekten verursacht werden.

Andere Ansätze benutzen Markov-Felder (engl. Markov Random Fields, MRF), um ein a-priori Wissen über die Eigenschaften der Objekte als Zwangsbedingung an die Ergebnisse zu stellen. Diese Methode als Gesamtergebnis von vornherein nur bestimmte Lösungen zuzulassen und somit schon bei der lokalen Berechnung dieses Ziel zu berücksichtigen, wird auch als Regularisierung bezeichnet.

Murray and Buxton [Murray, Buxton, 1987], benutzen den bereits berechneten Optischen Fluß und wenden ein räumliches Markov-Feld an. Die Tatsache, daß die Anzahl der sich bewegenden Teilchen im voraus bekannt sein muß, macht es schwierig diesen Algorithmus anzuwenden, ohne vorher eine aufwendige Vorsegmentierung durchzuführen.

Bouthemy and Lalande [Bouthemy, Lalande, 1990] berechnen die Differenz zweier aufeinanderfolgender Bilder und regularisieren dieses Ergebnis mit einem raum-zeitlichen Markov Feld.

Letang et al. [Letang et al., 1993] berechnen einen Änderungsvektor auf mehreren zeitlichen Auflösungsstufen durch einen Likelihood Ratio Test. Dieses Ergebnis zusammen mit der durch eine Kalman-Filterung geschätzte Position des Objektes, wird durch ein A Posteriori Kriterium regularisiert. Zur Berechnung des Änderungsvektors verwenden sie eine Pyramidenstruktur. Diese wird dadurch aufgebaut, daß jeweils zwei Bilder gemittelt werden und danach nur noch jedes zweite Bild zum Aufbau der nächst höheren Stufe benutzt wird. Durch diese unzureichende Glättung vor der Reduktion auf die halbe zeitliche Auflösung (siehe Kapitel 4) entsteht ein zeitliches Aliasing. Dies wird in Kauf genommen, um die Rechenzeit nicht zu erhöhen.

Filtertechniken zur Bewegungssegmentierung werden von [Duncan, Chou, 1992] und [Porat, Friedlander, 1990] angewendet. *Duncan and Chou* zeigen, daß die Positionen von sich bewegenden Kanten durch eine Faltung der Bildsequenz mit der zweiten Ableitung einer zeitlichen Gauß-Glättungsmaske detektiert werden. Die Nulldurchgänge des Faltungsergebnisses ergeben die Position der Kante. Dies entspricht im Prinzip der Tatsache, daß eine Kante immer als Nulldurchgang eines Laplaceoperators gefunden werden kann. Die extreme Rauschanfälligkeit des Laplaceoperators wird durch eine vorherige Glättung der Bilddaten berücksichtigt.

Porat und Friedlander verwenden einen Satz richtungsselektiver Filter im Fourier-Raum. Eine Entscheidungslogik vergleicht die Energieantwort der Filter, um Geschwindigkeit zu detektieren. Dieser Ansatz kann deutlich vereinfacht werden, wenn nur Bewegung in beliebiger Richtung detektiert werden soll. In diesem Fall muß nur ein einziger Filter angewendet werden, der alle zeitlichen Frequenzen ω mit $\omega \neq 0$ herausfiltert. Solch ein Filter wird in dieser Arbeit entwickelt.

Alle Arbeiten, die hier beschrieben wurden, dienen dazu, die Ränder sich bewegnender Objekte zu detektieren. Dies ist auch sinnvoll, wenn das Ganze als Vorstufe zu einer Geschwindigkeitsbestimmung dienen soll, da nur an diesen Stellen eine Verschiebung bestimmt werden kann. Soll jedoch die gesamte Fläche der Objekte detektiert werden, ist es nicht ausreichend, solange sich die Objekte nicht so schnell bewegen, daß sie sich von einem Bild zum nächsten über ihre gesamte Fläche verschieben. Wie die benötigte Filtergröße von der Größe und der Geschwindigkeit verschiedener Objekte abhängt, wird in Kapitel 3 analytisch bestimmt.

Unabhängig vom mathematischen Ansatz liefert jedes Verfahren der Digitalen Bildverarbeitung nur im Fall völlig rauschfreier Bilder direkt das theoretisch gewünschte Ergebnis. Da ein stochastisches Pixelrauschen immer zu Abweichungen zwischen zwei Bildern führt, kann ein Ergebnis erst ab einer bestimmten Schwelle als signifikant betrachtet werden. Optimale Schwellen, die eine Fehlsegmentierung von Pixelrauschen verhindern, ergeben sich entweder analytisch aus dem Verfahren selbst oder werden experimentell bestimmt. Alle Verfahren benutzen jedoch *globale Schwellen*, d. h. Schwellen, die nicht von der Position im Bild abhängen. Dies kann jedoch immer nur einen Kompromiß zwischen optimaler Segmentierung und minimaler Fehlsegmentierung darstellen. In Kapitel 6 wird eine *lokale Schwelle* berechnet, die sich den lokalen Strukturen der Bilder anpaßt. Dadurch kann das Ergebnis dem des menschlichen visuellen Systems angenähert werden.

Insgesamt gliedert sich die Bewegungssegmentierung in vier Abschnitte: In Kapitel 3 wird gezeigt,

wie bewegte Objekte durch zeitliche Glättungen der Bildsequenz mit konventionellen Filtermasken optimal aus der Bildsequenz entfernt werden können. Eine deutliche Beschleunigung der Berechnung, die einen Einsatz des Verfahrens überhaupt erst ermöglicht, wird durch die in Kapitel 4 entwickelte Mehrgitterstruktur erreicht. Das Ergebnis dieser Berechnung ist eine Bildsequenz, die nur noch ruhende Strukturen enthält. Dieser ruhende Hintergrund muß zur Detektion der bewegten Objekte mit den Originalbildern verglichen werden. In Kapitel 5 wird dazu ein Ähnlichkeitsmaß berechnet, um eine Aussage über die Größe der Abweichung zwischen Originalbild und Hintergrund machen zu können. Dieses Ähnlichkeitsmaß wird in Kapitel 6 zur endgültigen Entscheidung über die Signifikanz der Abweichung durch die bereits erwähnte lokale Schwelle bewertet und somit letztlich entschieden, ob die Abweichung durch Bewegung verursacht wurde. In Kapitel 7 erfolgt eine Zusammenfassung des gesamten Verfahrens und eine abschließende Diskussion der Ergebnisse anhand verschiedener Bildsequenzen.

Kapitel 3

Zeitliche Tiefpaßfilterung

In diesem Kapitel wird zunächst gezeigt, wie die Frequenzverteilung eines sich bewegenden Objektes im Fourierraum aussieht. Aus dieser Information ergibt sich eine Möglichkeit zur Extraktion der Objekte einer Bildfolge, die eine bestimmte Geschwindigkeit besitzen. Da alle Objekte mit beliebiger Geschwindigkeit $\vec{v} \neq 0$ detektiert werden sollen, ist es sinnvoller, zunächst die ruhenden Objekte aus der Bildfolge zu extrahieren. Ein anschließender Vergleich dieses ruhenden Hintergrundes mit dem zugehörigen Originalbild (Kapitel 5) liefert die sich bewegenden Objekte.

3.1 Projektionsfilter

Zur Analyse der zeitlichen Entwicklung des Bildinhaltes betrachtet man eine Bildfolge als eine dreidimensionale Struktur im *Bildfolgenraum* oder $\vec{x}t$ -Raum. Dort bilden bewegte Strukturen gegen die Zeitachse geneigte dreidimensionale Grauwertstrukturen. Ein solches 3D-Orts-Zeit-Bild ist in Abb. 3.1 für das Beispiel einer Scheibe mit konstantem Grauwert dargestellt, die sich mit konstanter Geschwindigkeit \vec{v} vor einem homogenen Hintergrund bewegt. Die Dicke der Scheibe entspricht dem zeitlichen Abstand zweier Bilder.

Die Neigung der Orts-Zeit-Struktur gegen die Zeitachse liefert direkt die Geschwindigkeit¹ der Scheibe:

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = - \begin{pmatrix} \tan \alpha_1 \\ \tan \alpha_2 \end{pmatrix}. \quad (3.1)$$

Das negative Vorzeichen ergibt sich aus der Definition der Winkel.

Die Betrachtung der Bildsequenz als ein Bildstapel und somit eine 3D-Struktur hat den Vorteil, daß die Eigenschaften von Objekten und Operatoren auch im zugehörigen 3D-Fourierraum untersucht werden können. Wie sich eine im Ortsraum mit der Geschwindigkeit \vec{v} bewegende beliebige Struktur im $\vec{k}\omega$ -Raum darstellt, läßt sich erkennen, wenn man die Fouriertransformierte

¹In diesem theoretischen Teil der Bildverarbeitung werden die Einheiten $[x] = 1 \text{ Pixel}$ und $[t] = 1 \text{ Bildabstand}$ verwendet. Damit ergibt sich $[\vec{v}] = 1 \text{ Pixel/Bild}$. Die genaue Eichung der Größen wird dann in Abschnitt 9.5 anhand der Auflösung des CCD-Chip und der optischen Vergrößerung durchgeführt.

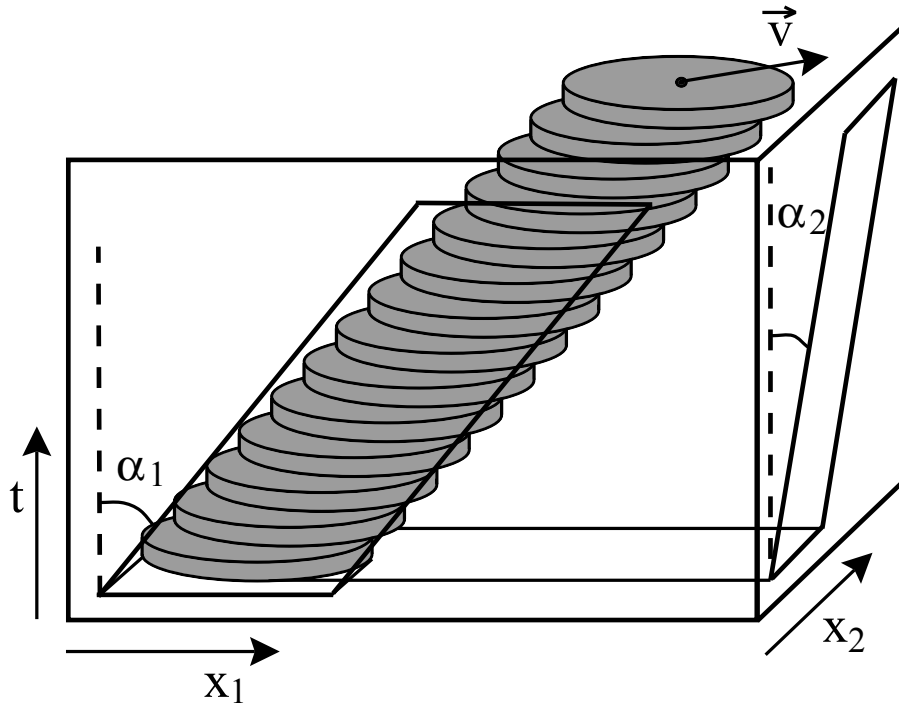


Abbildung 3.1: Darstellung der Bewegung einer Scheibe in einem Raum-Zeit-Bild.

dieser Struktur berechnet. Wird der Grauwertverlauf im Bildfolgenraum mit

$$g(\vec{x}, t) = g(\vec{x} - \vec{v}t) \quad (3.2)$$

beschrieben, so erhält man als Fouriertransformierte dieser Gleichung

$$\hat{g}(\vec{k}, \omega) = \hat{g}(\vec{k})\delta(\vec{k} \cdot \vec{v} - \omega). \quad (3.3)$$

Dies bedeutet, daß alle Objekte, die sich mit der Geschwindigkeit \vec{v} bewegen, im Fourierraum auf einer δ -Ebene liegen, deren Ebenengleichung durch $\omega = \vec{k} \cdot \vec{v}$ gegeben ist. Dabei werden die räumlichen Frequenzen der Struktur parallel zur ω -Achse auf diese Ebene projiziert. Die Ebene steht im Fourierraum senkrecht auf der Orts-Zeit-Struktur und ist unter den Winkeln α_1 und α_2 (siehe Abbildung 3.1) gegen die k_1 - bzw. k_2 -Achse geneigt.

Für ruhende Strukturen ($\vec{v} = 0$) erhält man aus (3.3)

$$\hat{g}(\vec{k}, \omega) = \hat{g}(\vec{k})\delta(\omega), \quad (3.4)$$

d. h. die k_1 - k_2 -Ebene im Fourierraum mit $\omega \equiv 0$. Wird die 3D-Fouriertransformierte einer Bildfolge mit dieser Ebene multipliziert, so erhält man nach der Rücktransformation eine Bildfolge, in der nur noch die ruhenden Strukturen zu sehen sind. Alle sich bewegendenden Objekte sind daraus verschwunden.

Wie alle Operationen der Bildverarbeitung kann auch dies direkt im Ortsraum durchgeführt werden. Die Multiplikation entspricht dort einer zeitlichen Faltung. Der dazugehörige Faltungsoperator im $\vec{x}t$ -Raum ist eine δ -Linie, die senkrecht auf der x_1 - x_2 -Ebene steht [Jähne, 1993a]

(p. 321). Durch diesen *Projektionsoperator* werden alle Bilder der Bildfolge aufaddiert. Ein Objekt, das sich bewegt, wird dadurch herausgemittelt. Das gewünschte Ergebnis wird aber nur für unendlich lange Bildsequenzen erreicht.

3.2 Zeitlicher Tiefpaß

Durch die zeitliche Begrenzung realer Bildfolgen wird die Geschwindigkeitsauflösung verschlechtert. Es können dadurch nur noch Objekte herausgefiltert werden, welche sich in dieser Zeit ausreichend bewegen. Der *zeitlich begrenzte Projektionsoperator* hat die Form:

$$H(\vec{x}, t) = \int_{t-\Delta t}^{t+\Delta t} w(t-t') g(\vec{x}, t') dt', \quad (3.5)$$

wobei $g(\vec{x}, t)$ den Grauwert der Bildfolge am Ort \vec{x} und zur Zeit t bezeichnet. Dies entspricht einer Faltung² in der Zeit mit einer geeigneten Fensterfunktion $w(t)$ als Faltungsmaske. Die Notation $H(\vec{x}, t)$ bezeichnet dabei die Tatsache, daß dies den ruhenden Hintergrund der Bildsequenz extrahiert.

Eine geeignete Fensterfunktion muß ihr Maximum bei dem betrachteten Bild zur Zeit t_0 haben und zur Grauwertterhaltung auf 1 normiert sein. Eine Mittelung über eine bestimmte Anzahl von Bildern, wobei alle Bilder gleich gewichtet werden, ist nicht ausreichend. Solch eine Faltung der Bildsequenz mit einer Rechteckfunktion erzeugt zeitliche Aliasingeffekte. Dadurch werden zeitliche Strukturen vorgetäuscht, die nicht existieren und somit die Objekte verfälschen.

Eine Faltungsmaske, welche in optimaler Weise als Glättungsfunktion geeignet ist, stellt die *Gaußmaske*

$$w(t) = B_\sigma(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \quad (3.6)$$

dar. Sie bietet eine minimale Unschärfe der Transferfunktion im Fourierraum bei minimaler Ausdehnung der Faltungsmaske im Zeitraum³.

Die diskrete Realisierung der Gaußmaske wird durch eindimensionale, *zeitliche Binomialmasken* \mathbf{B}_t^n angenähert. Eigenschaften von Binomialmasken werden in [Jähne, 1993a] (pp. 122-127) ausführlich beschrieben.

Aus (3.5) wird dann in Operatorschreibweise

$$H(\vec{x}, t) = \mathcal{B}_t^n g(\vec{x}, t). \quad (3.7)$$

²Grundlagen zur Faltung und ihre Anwendung in der Bildverarbeitung finden sich in allen Standardwerken der Bildverarbeitung, wie z. B. [Jähne, 1993a], pp. 100-116, [Gonzalez, Wintz, 1987], pp. 81-90 oder [Jain, 1989], p. 14 ff.

³Eine alternative Möglichkeit, die sich durch die Möglichkeit der direkten Implementierung in Hardware ergibt, wird in Anhang A beschrieben. Da sie jedoch nur auf einer bestimmten Bildverarbeitungskarte eingesetzt werden kann und eine zu kleine Grauwertaufflösung besitzt, wird in diesem Kapitel nicht näher darauf eingegangen.

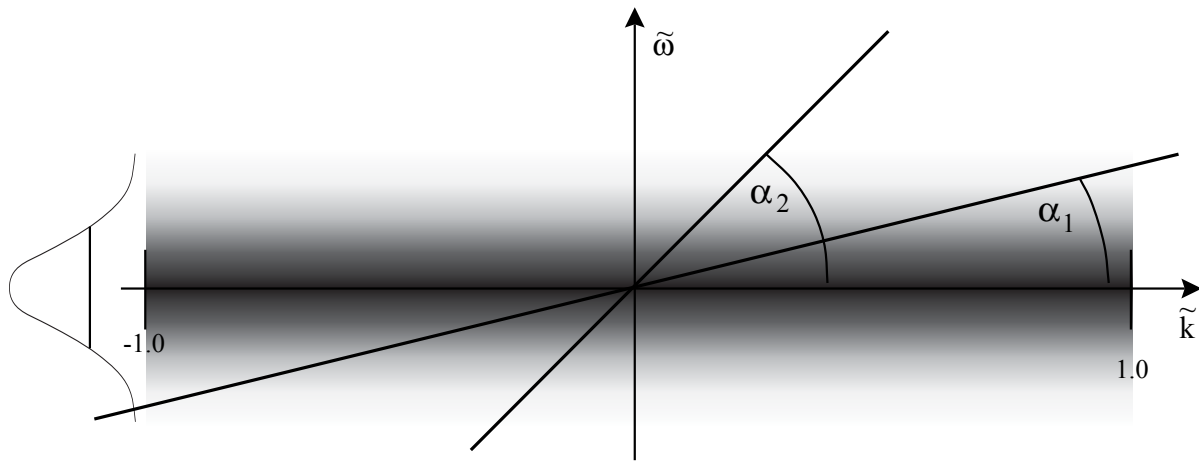


Abbildung 3.2: Wirkung einer zeitlichen Glättung im Fourierraum.

Die Breite der Faltungsmaske \mathbf{B}_t^n in der Zeit bestimmt die Anzahl der Bilder, welche bei der diskreten Realisierung der Faltung berücksichtigt werden müssen. Dies führt zu einem Rechenaufwand von $(n + 1)$ Multiplikationen und n Additionen, wobei jede der Rechenoperationen auf ein ganzes Bild anzuwenden ist. Dieser enorme Rechenaufwand läßt sich verringern, indem man die Eigenschaften von Binomialfiltern geschickt ausnutzt. Wie dies geschieht, wird in Kapitel 4 beschrieben.

Die Faltung der Bildsequenz mit einer Glättungsmaske stellt eine *zeitliche Tiefpaßfilterung* dar. Analog zu einem räumlichen Tiefpaß werden dabei hohe Frequenzen unterdrückt. Wie stark dies bei verschiedenen Filtermasken geschieht, zeigt die Transferfunktion der Filtermaske im Fourierraum ([Jähne, 1993a], S. 119).

Die zeitliche Frequenz einer räumlichen Struktur hängt sowohl von ihrer Geschwindigkeit als auch von ihrer räumlichen Frequenz ab. Bei idealen periodischen Strukturen wird dies anschaulich sofort deutlich: Eine sinusförmige Welle mit großer Wellenlänge, erzeugt bei einem ruhenden Beobachter eine kleinere zeitliche Änderung ihrer Wellenhöhe als eine Welle mit sehr kurzer Wellenlänge, die sich mit der gleichen Geschwindigkeit bewegt. Bei starren Objekten gelten diese Überlegungen für die einzelnen Komponenten der räumlichen Fouriertransformierten.

Die Wirkung einer zeitlichen Glättung auf die Frequenzen im Fourierraum wird in Abb. 3.2 veranschaulicht. Zur besseren Übersicht wird nur der Fall eindimensionaler Bewegung dargestellt. Die beiden Geraden stellen zwei Objekte mit verschiedenen Geschwindigkeiten dar. Die Geschwindigkeit folgt aus Gleichung (3.1). Das Maß der Dämpfung durch die zeitliche Glättung ist durch den Grauwertverlauf und die Transferfunktion links im Bild angedeutet. Man erkennt, daß nur noch zeitliche Frequenzen im Bild vorhanden sind, die sich um die k -Achse bei $\omega = 0$ konzentrieren. Die Ortsfrequenzen k ruhender Strukturen bleiben dadurch unbeeinflusst.

Welche Geschwindigkeiten nun bei einer zeitlichen Glättung der Bildfolge unterdrückt bzw. eliminiert werden, hängt von der Form und Größe der Objekte ab. Falls die Objekte nur sehr große räumliche Strukturen enthalten, werden sie von der zeitlichen Glättung wenig beeinflusst. Bestehen sie jedoch nur aus kleinskaligen Strukturen, so werden sie beide fast völlig verschwunden

sein. Objekte, deren räumliche Frequenzen gleichmäßig über die k -Achse verteilt sind, werden um so stärker verschmiert, je größer der Winkel, d. h. je größer ihre Geschwindigkeit ist, da die kleinen Strukturen zuerst verschwinden.

In den folgenden Abschnitten wird für zwei verschiedene Arten von Objekten eine analytische Beziehung zwischen der Größe der Objekte, ihrer Geschwindigkeit und den damit benötigten Filtergrößen hergeleitet. Dieser Zusammenhang ist wichtig, um ein optimales Ergebnis bei einem Minimum an Rechenzeit zu erreichen.

3.3 Ideale periodische Struktur

Das erste der beiden Beispiele ist eine Struktur, die einen idealen periodischen Grauwertverlauf in einem Bild aufweist und sich mit einer bestimmten Geschwindigkeit \vec{v} in beliebiger Richtung bewegt. Man könnte dies auch als *ebene Grauwertwelle* bezeichnen. Eine solche Struktur existiert normalerweise nicht in dieser Form in den Anwendungen der Bildverarbeitung. Eine Ausnahme bildet die Visualisierung von Wasseroberflächenwellen. Bei dieser speziellen Art eines *shape from shading* Verfahrens wird die Neigung der Wellen auf Grauwerte abgebildet und gibt somit der Grauwertstruktur eine physikalische Bedeutung, [Jähne, Riemer, 1990], [Beyer, 1993].

Nach dem Fourier'schen Theorem kann jedoch jede beliebige periodische Struktur nach periodischen Basisfunktionen entwickelt werden. Daher gibt eine Untersuchung der Wirkung von Filtern auf diese Grundstrukturen wichtige Aufschlüsse über deren Wirkung auf kompliziertere Objekte. Abbildung 3.3 zeigt eine solche periodische Grauwertwelle zusammen mit den benötigten Parametern. Dabei bezeichnet \vec{k}_o den Wellenvektor der Struktur:

$$\vec{k}_o = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} \frac{2\pi}{\lambda_1} \\ \frac{2\pi}{\lambda_2} \end{pmatrix}. \quad (3.8)$$

Er zeigt in die Richtung des größten Grauwertgradienten und hat die Länge der Orts-Wellenzahl in dieser Richtung.

Bewegt sich die Welle mit einer Geschwindigkeit \vec{v}_o , dann erzeugt dies eine zeitliche Frequenz ω_o . Die Beziehung zwischen den räumlichen und zeitlichen Frequenzen wird durch das Argument der δ -Funktion in Gleichung 3.3 gegeben:

$$\omega_o = \vec{k}_o \cdot \vec{v}_o. \quad (3.9)$$

Eine Bewegung senkrecht zum Vektor \vec{k}_o führt somit zu keiner zeitlichen Grauwertänderung.

Die Form der ebenen Grauwertwelle im Orts-Zeit-Raum wird damit durch

$$g(\vec{x}, t) = A \sin(\vec{k}_o \cdot \vec{x} - \omega_o t) = A \sin(\vec{k}_o \cdot (\vec{x} - \vec{v}_o t)) \quad (3.10)$$

beschrieben.

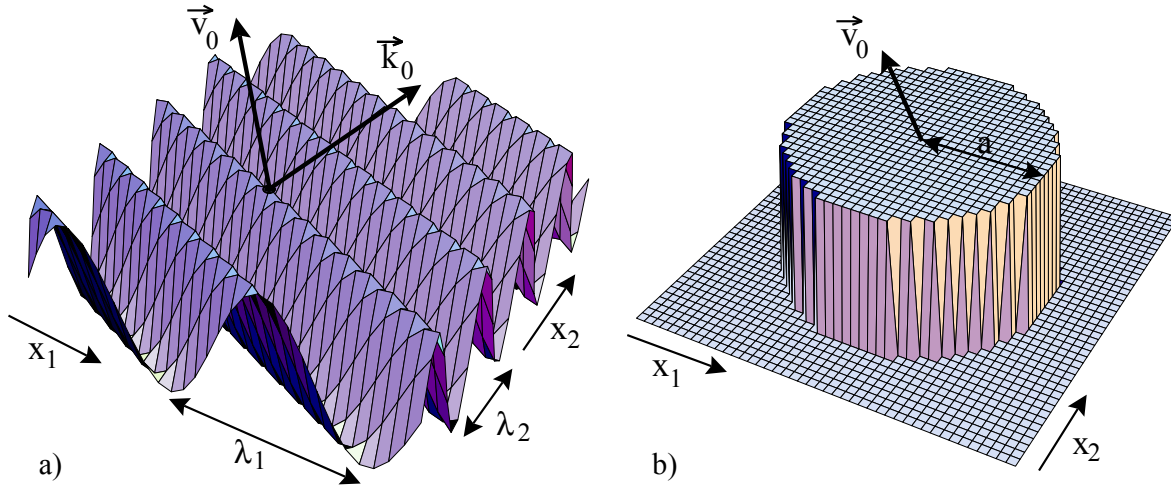


Abbildung 3.3: Beispiele für bewegte Objekte. a): Ebene Grauwertwelle. b): Bewegte Scheibe.

3.3.1 Größe der Faltungsmaske

Was passiert mit diesem Objekt, wenn es mit einer zeitlichen Binomialmaske geglättet wird? Wie aus der allgemeinen Filtertheorie bekannt ist, verursacht eine Glättung mit symmetrischen Filtern keine Phasenverschiebung des Signals. Die Amplitude einer periodischen Struktur der Frequenz ω_o wird dabei um einen Faktor gedämpft, der von der Größe des benutzten Filters abhängt. Dieser Faktor ist direkt durch den Verlauf der Transferfunktion der Faltungsmaske im Fourierraum gegeben. Sei A die Amplitude des Signals vor der Glättung, A' diejenige nach der Glättung und $\hat{\mathbf{B}}_t^n(\omega)$ die Transferfunktion des Binomialfilters \mathbf{B}_t^n , dann gilt:

$$\frac{A'}{A} = \hat{\mathbf{B}}_t^n(\omega_o). \quad (3.11)$$

Für einen gewünschten Grad der Dämpfung kann somit direkt die benötigte Faltungsmaske aus (3.11) berechnet werden. Im folgenden wird eine zeitlich periodische Struktur als *hinreichend geglättet* angenommen, wenn ihre Amplitude auf einen Faktor $e^{-1/2} \approx 0.61$ des ursprünglichen Wertes reduziert wird. Diese Größe ergibt sich aus zwei Gründen:

- Eine Verringerung der Amplitude auf diesen Wert ist ausreichend, um einen Unterschied zwischen dem geglätteten Signal und dem Originalsignal detektieren zu können. Eine weitere Glättung würde nur unnötigerweise den Rechenaufwand erhöhen.
- Eine Gaußfunktion $B_{\sigma_t}(t)$ ist nach einer Zeit $t = \sigma_t$ auf genau diesen Wert abgefallen: $B_{\sigma_t}(\sigma_t) = e^{-1/2} B_{\sigma_t}(0)$. Dies ermöglicht bei Gaußmasken eine einfache Bestimmung der zugehörigen Filtermaske einer gewünschten Transferfunktion.

Die Fouriertransformierte einer Gaußmaske $B_{\sigma_t}(t)$ ist wieder eine Gaußfunktion:

$$B_{\sigma}(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \quad \circ \text{---} \bullet \quad \hat{B}_{\sigma}(\omega) = e^{-\frac{\omega^2}{2/\sigma^2}}. \quad (3.12)$$

Da die Breiten der beiden Funktionen direkt reziprok zueinander sind, findet sich die gesuchte Breite σ_t der Faltungsmaske direkt als

$$\sigma_t = \frac{1}{\sigma_{\omega}}, \quad (3.13)$$

wenn σ_{ω} die Breite der Transferfunktion bezeichnet. Mit $\sigma_{\omega} = \omega_o$ und (3.9) ergibt sich:

$$\sigma_t = \left(\vec{v}_o \cdot \vec{k}_o \right)^{-1}. \quad (3.14)$$

Die entsprechende Binomialmaske \mathbf{B}_t^n zur diskreten Realisierung findet man durch Vergleich der Varianz der Binomialmaske mit der gewünschten Faltungsmaske. Die Varianz σ_t^2 einer Binomialmaske vom Grad n ist gegeben durch

$$\sigma_t^2 = \frac{n}{4}. \quad (3.15)$$

Mit (3.14) erhält man somit:

$$n = 4 \left(\vec{v}_o \cdot \vec{k}_o \right)^{-2}. \quad (3.16)$$

Eine Glättung der Bildsequenz mit einer zeitlichen Binomialmaske dieses Grades ergibt die gewünschte Abschwächung der Grauwertamplitude der Struktur (3.10).

3.4 Kreisförmiges Objekt

Nach der periodischen Struktur wird als zweites Beispiel ein rundes Objekt betrachtet. Diese in Abbildung 3.3b dargestellte *Grauwertscheibe* stellt ein Modell eines starren Körpers dar. Die Scheibe bewege sich mit der Geschwindigkeit \vec{v}_o . Im Gegensatz zu der periodischen Struktur bewirkt hierbei eine Bewegung in beliebiger Richtung eine zeitliche Grauwertänderung an einem Punkt. Die komplizierte räumliche Frequenzverteilung macht die Bestimmung der benötigten Faltungsmaske jedoch weniger angenehm.

3.4.1 3D-Fouriertransformierte

Die Grauwertverteilung $g(\vec{x})$ einer Scheibe mit Durchmesser $d = 2a$ und Helligkeit A (der Hintergrund habe den Grauwert 0), deren Mittelpunkt sich bei $\vec{x} = \vec{x}_o$ befindet, wird beschrieben durch

$$g(\vec{x}) = A \Pi\left(\frac{r}{2a}\right) \quad \text{mit} \quad r = \|\vec{x} - \vec{x}_o\|. \quad (3.17)$$

Ohne Beschränkung der Allgemeinheit wird im folgenden $\vec{x}_o = \vec{0}$ gesetzt. Durch die zirkulare Symmetrie der Funktion im Ortsraum reduziert sich die 2D-Fouriertransformation auf eine *1D-Hankel Transformation 0. Ordnung* [Bracewell, 1965]:

$$\begin{aligned}
\hat{g}(\vec{k}) &= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\vec{x}) e^{-i\vec{k} \cdot \vec{x}} dx_1 dx_2 \\
&= \frac{A}{(2\pi)^2} \int_{-\infty}^{\infty} \int_0^{2\pi} \Pi\left(\frac{r}{2a}\right) e^{-ikr \cos \Theta} r dr d\Theta \\
&= \frac{A}{(2\pi)^2} \int_{-\infty}^{\infty} \Pi\left(\frac{r}{2a}\right) \left[\int_0^{2\pi} e^{-ikr \cos \Theta} d\Theta \right] r dr \\
&= \frac{A}{2\pi} \int_{-\infty}^{\infty} \Pi\left(\frac{r}{2a}\right) J_0(kr) r dr \\
&= Aa \frac{J_1(ak)}{2\pi k},
\end{aligned} \tag{3.18}$$

$$\text{mit } k = \|\vec{k}\|, \quad r = \|\vec{x}\|, \quad \vec{k} \cdot \vec{x} = kr \cos \Theta \quad \text{und} \quad J_0(z) = \frac{1}{2\pi} \int_0^{2\pi} e^{-iz \cos \Phi} d\Phi.$$

Dabei enthält $\hat{g}(\vec{k})$ die *Bessel Funktion* J_1 . Wenn sich die Scheibe mit \vec{v}_o bewegt, wird die raum-zeitliche 3D-Fouriertransformation von $g(\vec{x})$ durch (3.3) bestimmt:

$$\hat{g}(\vec{k}, \omega) = Aa \frac{J_1(ak)}{2\pi k} \delta(\vec{k} \cdot \vec{v}_o - \omega). \tag{3.19}$$

Die räumliche Fouriertransformation (3.18) wird dabei parallel zur ω -Achse auf die δ -Ebene projiziert. In Abbildung 3.4 a ist dies für eine Scheibe mit Radius $a = 4 \text{ Pixel}$ illustriert. Die Scheibe bewegt sich mit der Geschwindigkeit $\|\vec{v}_o\| = 1 \text{ Pixel/Bild}$ in x_1 -Richtung. Die Verteilung der raum-zeitlichen Frequenzen auf der Ebene ist durch einen *3D surface plot* angedeutet.

3.4.2 Größe der Faltungsmaske

Abbildung 3.4 b veranschaulicht die Auswirkung einer zeitlichen Glättung auf die 3D-Fouriertransformierte der bewegten Scheibe. Die zeitliche Faltung der Bildsequenz mit einer Binomialmaske bewirkt im Fourierraum eine Multiplikation mit der Transferfunktion der Faltungsmaske. Hat die Binomialmaske die effektive zeitliche Reichweite σ_t , dann werden alle zeitlichen Frequenzen ω auf einen Faktor $\hat{B}_{\sigma_t}(\omega)$ ihres ursprünglichen Wertes reduziert. Durch die Geschwindigkeit des Objektes, d. h. der Neigung der Ebene sind die räumlichen Frequenzen direkt mit den zeitlichen verknüpft.

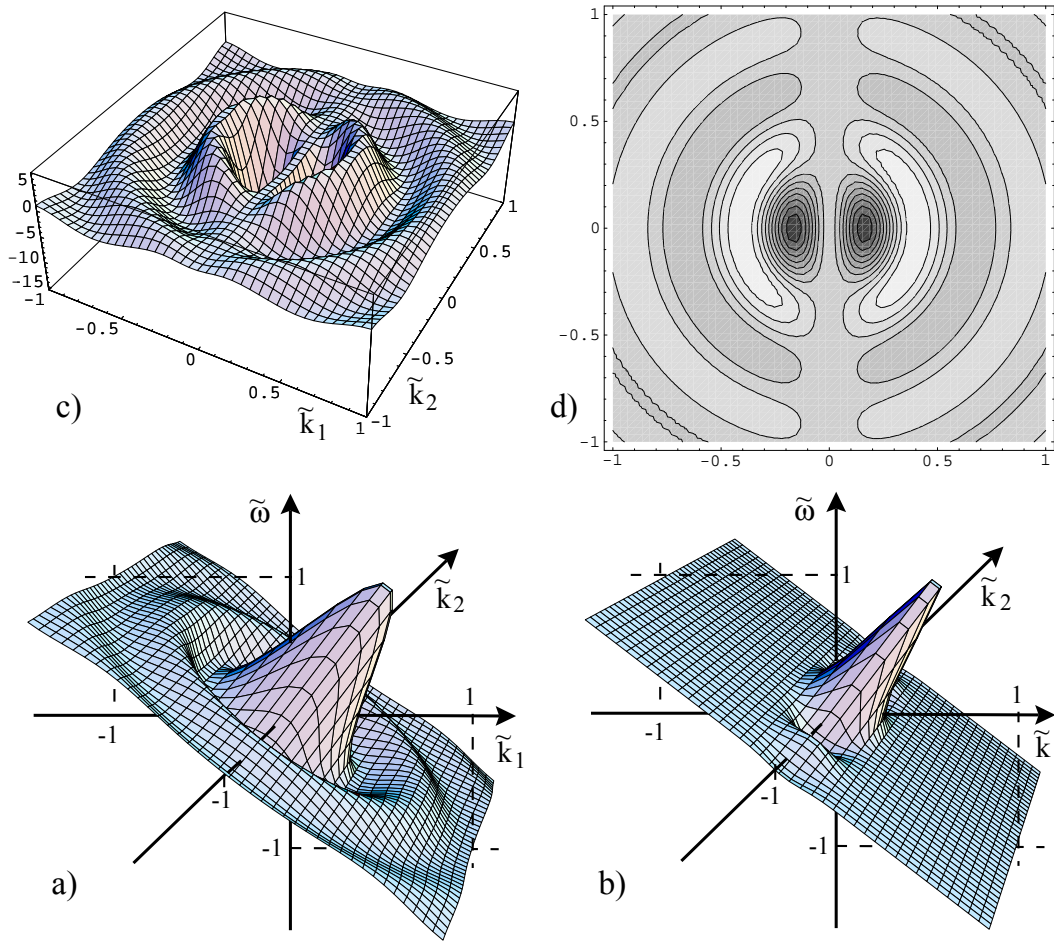


Abbildung 3.4: 3D-Fouriertransformierte einer Scheibe mit Radius $a = 4 \text{ Pixel}$, die sich mit der Geschwindigkeit $|\vec{v}_o| = 1 \text{ Pixel / Bild}$ in x_1 -Richtung bewegt: a) Originalsequenz b) zeitlich geglättete Sequenz c) Differenz zwischen Originalsequenz und zeitlicher Glättung (nach Rückprojektion in die $k_1 - k_2$ -Ebene) d) Contour Plot von c) .

Die zeitlich geglättete Frequenzverteilung \hat{g}_t wird somit beschrieben durch

$$\hat{g}_t(\vec{k}, \omega) = Aa \frac{J_1(ak)}{2\pi k} \delta(\vec{k} \cdot \vec{v}_o - \omega) \exp\left[-\frac{\omega^2}{2\sigma_\omega^2}\right]. \quad (3.20)$$

Die Auswirkung der zeitlichen Glättung hängt von der Neigung der Ebene und der Ausdehnung der Transferfunktion in σ_ω ab. Durch eine breite Transferfunktion (kleine Glättungsmaske), werden nur die großen Orts-Wellenzahlen weggeglättet. Dadurch werden einzig die Ränder des Objektes verschmiert. Je schmaler die Transferfunktion wird (große Glättungsmaske), desto größere Strukturen des Objektes werden entfernt bis es fast völlig verschwunden ist.

Umgekehrt werden bei einer festen Transferfunktion immer kleinere Orts-Wellenzahlen beeinflusst, wenn die Geschwindigkeit und damit die Neigung der Ebene zunimmt. Nur eine Ebene der Neigung Null (ruhende Objekte) bleibt unabhängig von der Größe der zeitlichen Glättung unverändert.

Analog zu Abschnitt 3.3.1 wird nun eine Scheibe als *hinreichend geglättet* definiert, wenn der Mittelpunkt der Scheibe um den gleichen Betrag abgeschwächt wird, wie die Amplitude der periodischen Struktur:

$$g_t(\vec{x} = \vec{0}) \stackrel{!}{=} e^{-1/2} g(\vec{x} = \vec{0}) = e^{-1/2} A. \quad (3.21)$$

Der Index t bezeichnet die zeitliche Glättung. Ohne Beschränkung der Allgemeinheit befinde sich der Mittelpunkt der Scheibe zu einer Zeit $t = 0$ am Ort $\vec{x} = \vec{0}$. Dieser Wert ergibt sich aus den gleichen Überlegungen wie in 3.3.1, zusammen mit der Tatsache, daß der Mittelpunkt immer am wenigsten von einer Glättung beeinflusst wird, da er sich aus den größten Strukturen des Objektes zusammensetzt. Wenn der Mittelpunkt detektiert werden kann, trifft dies auch auf alle anderen Punkte der Scheibe zu.

Um die Größe der benötigten zeitlichen Binomialmaske zu finden, wird die Rücktransformation der geglätteten Frequenzen (3.20) in den Orts-Zeit-Raum berechnet:

$$g_t(\vec{x}, t) = \frac{A a}{2\pi} \int_{-\infty}^{\infty} d^2 k \, d\omega \, \frac{J_1(ak)}{k} \delta(\vec{k} \cdot \vec{v}_o - \omega) \exp\left[-\frac{\omega^2}{2\sigma_\omega^2}\right] \exp\left[-i(\vec{k} \cdot \vec{x} - \omega t)\right].$$

Unter Ausnutzung der Eigenschaften der δ -Funktion reduziert sich dieses Volumenintegral auf ein 2D-Flächenintegral

$$g_t(\vec{x}, t) = \frac{A a}{2\pi} \int_{-\infty}^{\infty} d^2 k \, \frac{J_1(ak)}{k} \exp\left[-\frac{(\vec{v}_o \cdot \vec{k})^2}{2\sigma_\omega^2}\right] \exp\left[-i(\vec{k} \cdot \vec{x} - \omega t)\right]$$

über die Orts-Wellenzahlen. Da jedoch nur der Mittelpunkt $g(\vec{x} = \vec{0}, t = 0)$ der Scheibe interessiert, genügt es, das Integral für diesen Punkt zu berechnen. Als Ergebnis erhält man:

$$g_t(\vec{x} = 0, t = 0) = \frac{A a}{2\pi} \int_{-\infty}^{\infty} d^2 k \, \frac{J_1(ak)}{k} \exp\left[-\frac{(\vec{v}_o \cdot \vec{k})^2}{2\sigma_\omega^2}\right], \quad (3.22)$$

d. h. eine Integration über alle bei der Glättung übriggebliebenen Frequenzen. Das Skalarprodukt $\vec{v}_o \cdot \vec{k}$ im Argument der Exponentialfunktion drückt genau die oben beschriebene Tatsache aus, daß nur räumliche Strukturen, die Komponenten in Richtung der Geschwindigkeit aufweisen, gedämpft werden, während Strukturen senkrecht zu dieser Richtung unbeeinflusst bleiben. In den Abbildungen 3.4 c und 3.4 d ist dies deutlich zu erkennen. In der Anwendung werden deshalb alle Objekte, die sich bewegen, in Bewegungsrichtung verschmiert, während die Steilheit der Kanten senkrecht zur Bewegung nicht verändert wird.

Diese Tatsache, daß eine zeitliche Glättung den gleichen Effekt wie eine räumlichen Glättung hat (jedoch nur bei den Objekten mit $\|\vec{v}_o\| \neq 0$!) wird noch deutlicher, wenn man (3.22) etwas umformt:

$$g_t(\vec{x} = 0, t = 0) = \frac{A a}{2\pi} \int_{-\infty}^{\infty} d^2 k \, \frac{J_1(ak)}{k} \exp\left[-\frac{(\hat{v}_o \cdot \vec{k})^2}{2(\sigma_\omega/v_o)^2}\right] \quad (3.23)$$

mit $\vec{v}_o \cdot \vec{k} = (\hat{v}_o \cdot \vec{k}) v_o$ und $\hat{v} = \vec{v}_o/v_o$. Die Exponentialfunktion stellt die Transferfunktion einer räumlichen 1D-Glättung in Richtung \hat{v}_o dar, wobei die Breite der Transferfunktion durch $\sigma_k = \sigma_\omega/v_o$ gegeben ist und somit mit der Geschwindigkeit skaliert.

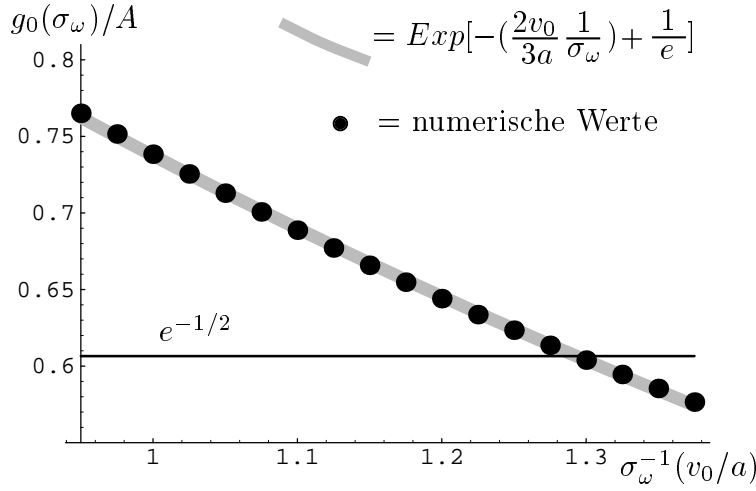


Abbildung 3.5: Numerische Lösung der Gleichung (3.23) und Fit der numerischen Daten.

Gleichung (3.23) ist eine Funktion $g_{t,o}(\sigma_\omega)$, aus der σ_ω berechnet werden kann, um (3.21) zu erfüllen. Unglücklicherweise läßt sie sich jedoch nicht analytisch lösen. Da der Integrand nicht mehr isotrop ist, ist die Rücktransformation der geglätteten Frequenzen keine Hankel-Transformation mehr.

Eine numerische Lösung von (3.23) liefert Werte, die in Abb. 3.5 als Punkte dargestellt sind. Der Verlauf dieser Punkte läßt sich in einer Umgebung von $\sigma_\omega = v_o/a$ in guter Näherung durch die Kurve

$$g_{t,o}(\sigma_\omega) = \frac{g_t(\vec{x} = 0, t = 0)}{A} = \exp \left[-\left(\frac{2v_o}{3a} \frac{1}{\sigma_\omega} \right) + \frac{1}{e} \right], \quad \text{mit } v_o = \|\vec{v}_o\| \quad (3.24)$$

darstellen. Damit erhält man direkt die Breite der Transferfunktion der benötigten Glättungsmaske aus dem Schnittpunkt der numerischen Lösung (3.24) mit der Geraden $g_{t,o}(\sigma_\omega) = e^{-1/2}$:

$$g_{t,o}(\sigma_\omega) = e^{-1/2} \iff \sigma_\omega = \frac{4}{3} \frac{e}{(e+2)} \frac{v_o}{a} \approx \frac{4v_o}{5a}. \quad (3.25)$$

Analog zu den Berechnungen für die periodische Struktur wird hieraus mit Gleichung (3.13) die effektive zeitliche Reichweite σ_t der benötigten Glättungsmaske

$$\sigma_t \approx \frac{5}{4} \frac{a}{v_o} \quad (3.26)$$

und mit (3.15) der Grad n der korrespondierenden Binomialmaske \mathbf{B}_t^n

$$n \approx \frac{25}{4} \frac{a^2}{v_o^2}. \quad (3.27)$$

Daraus wird deutlich, daß der Grad der benötigten zeitlichen Glättung und somit auch der Grad der Dämpfung bei einer festen Maskengröße nur vom Verhältnis der Geschwindigkeit des Objektes zu seiner Ausdehnung abhängt. Im Gegensatz zu der periodischen Struktur tauchen in diesen Gleichungen keine Vektoren auf. Dies zeigt, daß die Größe der zeitlichen Grauwertänderung durch die Kreissymmetrie der Scheibe nur vom Betrag der Geschwindigkeit und nicht von ihrer Richtung abhängt.

3.5 Verallgemeinerung für beliebige Objekte

In den beiden vorangegangenen Abschnitten wurde für eine periodische Struktur und ein kreisförmiges Objekt, jeweils unabhängig voneinander, untersucht wie groß eine zeitliche Glättung sein muß, um beide Objekte in gleichem Maße aus der Bildsequenz zu entfernen. In diesem Abschnitt werden die Ergebnisse verglichen und zu einer Beziehung für beliebige Objekte verallgemeinert.

Im folgenden sei λ_w die räumliche Wellenlänge der periodischen Struktur in Richtung ihrer Geschwindigkeit $\vec{v}_{o,w}$. Es zeigt sich, daß die Größe \mathfrak{x} mit

$$\mathfrak{x} = k_o^{-1} = \frac{\lambda_w}{2\pi} \quad (3.28)$$

als Maß für die Ausdehnung der Welle besser geeignet ist als λ_w . Die Komponente des Wellenvektors \vec{k}_o in Bewegungsrichtung ist dann

$$\vec{k}_o \cdot \hat{v}_{o,w} = \frac{2\pi}{\lambda_w} = \mathfrak{x}^{-1}. \quad (3.29)$$

Aus Gleichung (3.14) wird damit die skalare Gleichung

$$\sigma_{t,w} = \left(\vec{v}_o \cdot \vec{k}_o \right)^{-1} = \frac{\lambda_w}{2\pi v_{o,w}} = \frac{\mathfrak{x}}{v_{o,w}}. \quad (3.30)$$

Für eine Scheibe mit Radius a , die sich mit der Geschwindigkeit $v_{o,s}$ bewegt, erhält man nach (3.26):

$$\sigma_{t,s} = \frac{5a}{4v_{o,s}}. \quad (3.31)$$

Wird (3.30) mit (3.31) für den Fall $\mathfrak{x} = a$ und $\vec{v}_{o,w} = \vec{v}_{o,s}$ verglichen, also eine Welle und eine Scheibe mit gleicher charakteristischer Länge und gleicher Geschwindigkeit, ergibt sich:

$$\sigma_{t,s} = \frac{5}{4} \sigma_{t,w}. \quad (3.32)$$

Für den Grad n der benötigten Binomialmasken \mathbf{B}_t^n ergibt sich mit (3.16) und (3.27):

$$n_s = \frac{25}{16} n_w. \quad (3.33)$$

Dies zeigt, daß die Größen der benötigten Filtermasken bis auf einen Faktor 1.6 für beide Objekte übereinstimmen, wenn die richtigen Größen als Maß für die Ausdehnung der Objekte verwendet werden.

Bis jetzt wurde vorausgesetzt, daß sich in der Bildsequenz nur ein einzelnes Objekt mit bekannten Eigenschaften bewegt. Diese Voraussetzung wird in realen Bildszenen fast nie erfüllt. Trotzdem sind die Beziehungen, die in diesem Kapitel hergeleitet wurden, nicht auf diesen Spezialfall beschränkt. Mit einigen Modifikationen gelten die Ergebnisse auch für jede Anzahl beliebiger Objekte, welche sich dabei in alle möglichen Richtungen bewegen und zusätzlich beschleunigen können.

Die beiden betrachteten Beispiele stellen zwei Grenzfälle für reale Objekte dar. Die periodische Struktur ist das einfachste Beispiel für ein Objekt. Sie besteht im Idealfall aus nur zwei Punkten im Fourierraum und kann dazu benutzt werden, beliebige Strukturen aufzubauen. Die Scheibe mit ihrer großen, konstanten Ausdehnung und steilen Kanten stellt den anderen Extremfall dar, da sie unendlich viele Frequenzen enthält. Jedes andere Objekt kann als eine Art *Wellenpaket* mit einer charakteristischen Länge Λ_c betrachtet werden, das zwischen den beiden Extremen liegt.

Diese Betrachtung läßt einen wichtigen Schluß zu:

Jedes beliebige Objekt mit einer charakteristischen Länge Λ_c und beliebiger Geschwindigkeit kann mit diesem Verfahren optimal detektiert werden.

Die Beziehungen für die Filtergrößen können verallgemeinert werden, indem man $\tilde{\kappa}$ oder a durch Λ_c ersetzt. Werden texturierte Objekte beobachtet, so müssen nur die kleinsten Skalen der Textur betrachtet werden, da diese am meisten zu den zeitlichen Änderungen beitragen. Somit können die Gleichungen (3.16) und (3.27) vereinigt werden zum Grad n_t der benötigten zeitlichen Binomialmaske:

$$n_t = 4c^2 \frac{\Lambda_c^2}{v^2}, \quad 1 \leq c \leq \frac{5}{4}. \quad (3.34)$$

Der *Formparameter* c liegt zwischen dem Wert 1 für ideale, periodische Strukturen und dem Wert 5/4 für ein Objekt mit steilen Kanten. In der praktischen Anwendung kann folgendermaßen verfahren werden:

- Der Formparameter c muß der augenscheinlichen Form der Objekte angepaßt werden. In Realweltszenen, in denen vorwiegend starre Körper, wie z. B. Fahrzeuge oder Sandkörner beobachtet werden, liegt er normalerweise bei etwa 5/4.
- Als Wert für Λ_c wird die zu erwartende⁴ größte charakteristische Länge gewählt. Bei Objekten mit konstantem Grauwert muß die Ausdehnung der Objekte selbst, bei texturierten Objekten dagegen nur die Strukturen der Textur berücksichtigt werden.
- Für die Geschwindigkeit v in (3.34) wird die kleinste Geschwindigkeit, die man segmentieren möchte, bzw. die kleinste Geschwindigkeit, die während eines Beschleunigungsprozesses auftritt, eingesetzt. Eine Erhöhung der Geschwindigkeit verbessert nur die Wirkung der zeitlichen Glättung mit einer gegebenen Filtergröße.

⁴Unter *zu erwartend* soll hierbei entweder *durch die Kenntnis der Größe der beobachteten Objekte und der Kameraparameter berechnet* oder *anhand von Beispielbildern ermittelt* verstanden werden.

Kapitel 4

Zeitliche Gaußpyramide

In Kapitel 3 wurde gezeigt, wie der ruhende Hintergrund einer Bildszene durch eine zeitliche Tiefpaßfilterung mit Binomialmasken aus der Bildsequenz extrahiert werden kann. Die Breite der Faltungsmaske \mathcal{B}_t^n in der Zeit bestimmt die Anzahl der Bilder, die berücksichtigt werden müssen. Bei einer Binomialmaske des Grades n führt dies zu einem Rechenaufwand von $(2n+1)$ Multiplikationen und $(2n)$ Additionen, wobei jede dieser Operationen auf ein ganzes Bild anzuwenden ist. Speziell bei langsamen und großen Objekten, zu deren Detektion große Faltungsmasken verwendet werden müssen, führt dies zu einem enormen Rechenaufwand¹.

Dieser enorme Rechenaufwand läßt sich verringern, indem die Eigenschaften des Binomialfilters ausgenutzt werden: Eine große Binomialmaske läßt sich aus kleineren Binomialmasken aufbauen, wenn diese wiederholt auf sich selbst angewendet wird. Führt man diesen Schritt immer wieder aus, so nähert sich die daraus resultierende Maske immer mehr der diskreten Realisierung einer Gaußverteilung an [Jähne, 1993a]. Für die hier verwendete zeitliche Faltung bedeutet dies, daß immer nur wenige, zeitlich benachbarte Bilder gemittelt werden müssen. Dieser Vorgang wird für die dadurch entstandene neue Bildfolge mehrmals wiederholt. Verwendet man als Elementarmaske die kleinste mögliche Maske $\mathcal{B}_t^1 = 0.5 [1 \ 1]$, so reduziert sich die Berechnung auf eine einfache Addition pro Bild und eine anschließende Shift-Operation. Durch diese Zerlegung der Faltungsmasken werden alle Multiplikationen eingespart. Die Anzahl der Additionen läßt sich jedoch nicht reduzieren. Dies liegt daran, daß die effektive Reichweite der Binomialmaske \mathcal{B}_t^n nur proportional zu \sqrt{n} ansteigt oder anders ausgedrückt, daß sich bei einer sukzessiven Anwendung zweier Masken die Varianzen addieren und nicht die effektiven Reichweiten. Um große Faltungsmasken aufzubauen, muß deshalb sehr oft eine kleine Maske angewendet werden.

4.1 Allgemeine Gaußpyramiden

Den Schlüssel zum Verringern des Rechenaufwandes liefert eine konsequente Ausnutzung des Abtasttheorems ([Jähne, 1993a], pp. 45-52). Wird von einem vorgegebenen Abstand Δx_i bzw. Δt_i diskreter Grauwertstichproben in Raum und Zeit ausgegangen, so liefert das Abtasttheorem eine

¹Bei einer \mathcal{B}^{252} -Maske ca. 40 min auf einem Intel i386-Rechner und 5 min auf einem Intel i860 Risc-Prozessor.

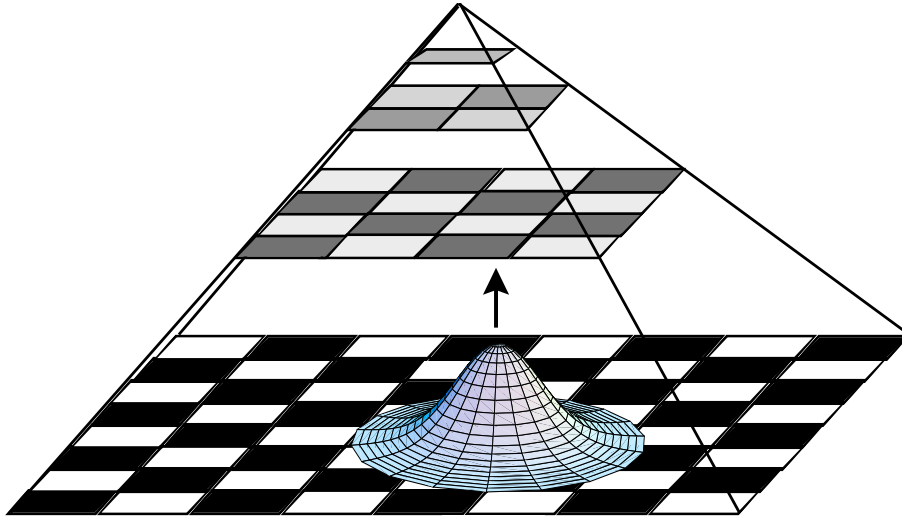


Abbildung 4.1: Schematischer Aufbau einer Gaußpyramide für 2D-Bilder. Die Quadrate entsprechen einzelnen Pixeln.

Zwangsbedingung an die maximale Frequenz (Nyquistfrequenz) einer periodischen Struktur, die von diesem diskreten Gitter unverfälscht wiedergegeben wird. Umgekehrt erhält man für eine vorgegebene Struktur einen maximalen Abstand der Abtastpunkte. Verringert sich die größte Frequenz der im Bild enthaltenen Strukturen, so vergrößert sich gleichzeitig dieser maximale Abstand. Genau dieser Fall tritt bei einer Glättung der Bildsequenz ein. Werden hohe Frequenzen durch Glättungsmasken so stark unterdrückt, daß nur noch Frequenzen unterhalb der halben Nyquistfrequenz vorhanden sind, dann wird der dazugehörige maximale Abstand der Abtastpunkte verdoppelt. Ohne Verlust an Information kann danach jeder zweite Bildpunkt bzw. jedes zweite Bild verworfen werden. Wird dieser Vorgang sukzessive fortgeführt, dann erhält man eine Bildstruktur, in der von einer Ebene zur nächsten die Anzahl der Gitterpunkte halbiert wird: eine sogenannte *Gaußpyramide* (Abb. 4.1).

Durch eine Reduktion der Bildsequenz auf die halbe zeitliche Auflösung kann die Zusammensetzung großer Binomialmasken nun viel effektiver gestaltet werden. Auf dem gröberen Gitter hat eine Glättungsmaske eine vier mal große Varianz als auf dem feineren Gitter (Anhang C.1). Ein Mitteln der reduzierten Bildfolge mit der gleichen Binomialmaske entspricht damit einer Maske mit doppelter Reichweite.

4.2 Aufbau durch gezieltes Unterabtasten

Um eine Gaußpyramide aufzubauen, muß zunächst eine geeignete Glättungsmaske gefunden werden, die alle Frequenzen oberhalb der halben Nyquistfrequenz entfernt. Die von *Burt* [Burt, 1984] zur Berechnung von Pyramiden in Ortsauflösung vorgeschlagene \mathbf{B}^4 -Glättungsmaske ist zu klein, um diese Voraussetzung zu erfüllen (Abb. 4.3a). Zur Vermeidung von Aliasingeffekten wird mindestens eine \mathbf{B}^{16} -Binomialmaske mit 17 Filterkoeffizienten benötigt. In [Jähne, 1993b] wird ein Verfahren zur effektiven Iteration großer 2D-Filtermasken durch Verwendung kleinerer Masken

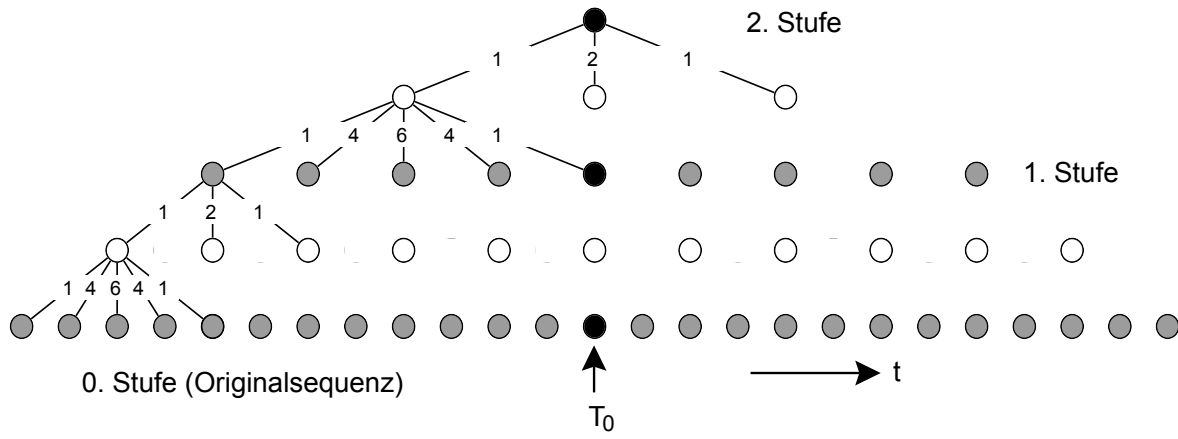


Abbildung 4.2: Zeitliche Gaußpyramide. Die Kreise entsprechen einzelnen Bildern.

beschrieben, das den Rechenaufwand noch weiter reduziert. Dabei werden kleine Faltungsmasken auf größere Gitterabstände gespreizt. Dadurch entstehende Aliasingeffekte werden verhindert, wenn durch zusätzliche Glättungen mit kleineren Faltungsmasken bei diesen Frequenzen bereits Nullstellen der Transferfunktionen vorliegen. In zwei Dimensionen stehen dazu zwei Parameter der Glättungsmasken zur Verfügung. Sie können skaliert und gleichzeitig unter verschiedenen Winkeln angewendet werden. Obwohl für die eindimensionale, zeitliche Glättung in Kapitel 3 keine Möglichkeit besteht, die Richtung der Faltung zu ändern, kann eine effektive Zerlegung der Binomialmasken durchgeführt werden. Dies geschieht durch *gezieltes Unterabtasten* unter Ausnutzung der Nullstellen der Transferfunktionen. Abbildung 4.2 zeigt schematisch den Aufbau der Gaußpyramide.

Dabei werden folgende Masken verwendet:

$$\begin{aligned} \mathbf{B}_t^4 &= \frac{1}{16} [1 \ 4 \ 6 \ 4 \ 1] \\ \mathbf{B}_t^2 &= \frac{1}{4} [1 \ 2 \ 1] \\ \mathbf{B}_{2t}^2 &= \frac{1}{4} [1 \ 0 \ 2 \ 0 \ 1]. \end{aligned}$$

Definiert man analog zu [Jähne, 1993a] den *zeitlichen Reduktionsoperator* \mathcal{R}_t , der nur jedes zweite Bild einer Bildfolge herausgreift, so entspricht die gesamte Glättungsmaske \mathcal{B}_t^{casc} zum Aufbau einer Pyramidenstufe der *Filterkaskade*

$$\mathcal{B}_t^{casc} = \mathcal{B}_{2t}^2 \mathcal{B}_t^4 = \mathcal{B}_t^2 \mathcal{R}_t \mathcal{B}_t^4. \quad (4.1)$$

Als Ausgangspunkt der *Filterkaskade* dient eine Glättung der Originalsequenz mit dem Operator \mathcal{B}_t^4 . Eine anschließende Reduktion dieser geglätteten Bildfolge auf die halbe zeitliche Auflösung mit \mathcal{R}_t verursacht unvermeidbar *zeitliche Aliasingeffekte*, da die Transferfunktion dieser kleinen Filtermaske bei der halben Nyquistfrequenz erst auf 0.25 abgefallen ist und noch weit darüber hinaus größer als Null bleibt (Abb. 4.3a). Ungeachtet dieser Tatsache, wird die Reduktion der Bildsequenz mit \mathcal{R}_t bereits in diesem Stadium durchgeführt. Im Gegensatz zu [Burt, 1984] wird dieses, in Abbildung 4.2 mit weißen Kreisen gekennzeichnete Zwischenergebnis jedoch durch eine

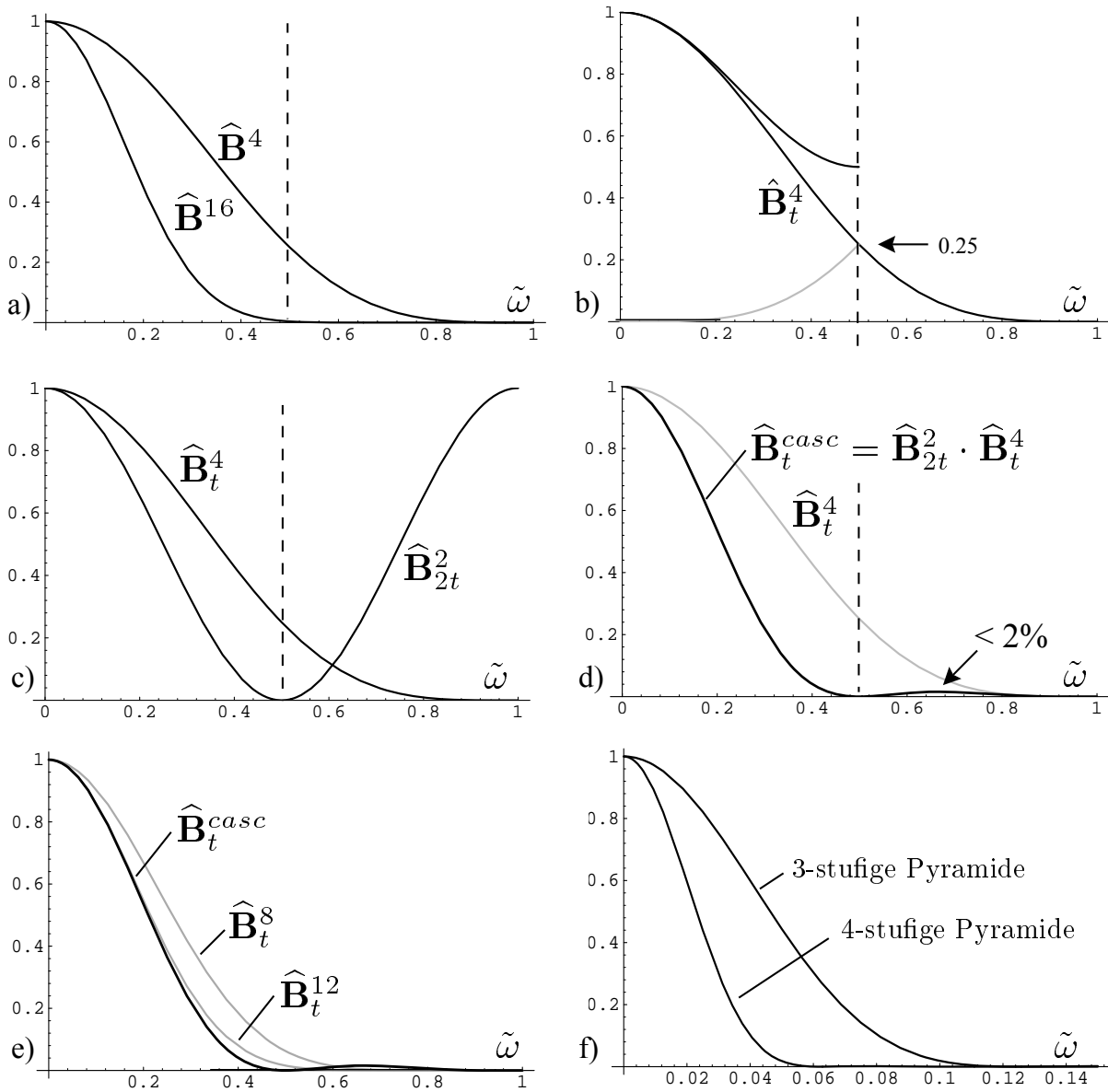


Abbildung 4.3: Transferfunktion der verwendeten Filterkaskade im Vergleich zu konventionellen Bino-mialmasken: a) Vergleich $\mathcal{B}_t^4 \leftrightarrow \mathcal{B}_t^8$. b) Lokalisierung der Aliasingeffekte im Frequenzspektrum. c) Trans-ferfunktionen der beiden Filter, die zum Aufbau der Gaußpyramide verwendet werden. d) Resultierende Transferfunktion der Filterkaskade im Vergleich zu \mathcal{B}_t^4 . e) Vergleich der Filterkaskade mit \mathcal{B}_t^8 und \mathcal{B}_t^{12} . f) Resultierende Transferfunktionen für 3- und 4-stufige Pyramiden.

weitere Glättung mit einem \mathcal{B}_t^2 -Operator nachbearbeitet. Diese anschließende Filterung dient dazu, die vorher produzierten Aliasingeffekte fast vollständig zu kompensieren. Die Wirkung der nachfolgenden Faltung \mathcal{B}_t^2 im Zusammenspiel mit der Vorglättung durch \mathcal{B}_t^4 ist in Abb. 4.3 illustriert.

Lokalisierung der Aliasingeffekte im Frequenzspektrum (Abb. 4.3b)

Wird eine Bildsequenz auf die halbe zeitliche Auflösung reduziert ohne vorher ausreichend geglättet zu werden, dann ragen die abgeschnittenen Frequenzen in die benachbarte Elementarzelle des neuen reziproken Gitters. Durch die periodische Wiederholung des Frequenzspektrums im Fourierraum erscheinen diese Frequenzen an der neuen Nyquistfrequenz bei $\tilde{\omega} = 1/2$ gespiegelt. Dort addieren sich im Spektrum die verfälschten Frequenzen zu den bereits vorhandenen korrekten Frequenzen hinzu.

Korrektur der Aliasingeffekte durch Nachglättung (Abb. 4.3c)

Eine Faltung der Bildsequenz mit \mathcal{B}_t^2 nach der Reduktion entspricht einer Faltung mit \mathcal{B}_{2t}^2 vor der Reduktion. Da der Operator auf einem Gitter mit doppeltem Gitterabstand arbeitet, sieht man die periodische Wiederholung der Transferfunktion bereits in der Elementarzelle des ursprünglichen Gitters. Der \mathcal{B}_t^2 -Operator alleine wäre daher kein guter Glättungsoperator, da er hohe Frequenzen identisch erhält. In einer Filterkaskade ergänzt er jedoch ideal den \mathcal{B}_t^4 -Operator. Die Transferfunktion von \mathcal{B}_t^2 hat bei der kritischen Grenzfrequenz $\tilde{\omega} = 1/2$ eine Nullstelle und ist in einer Umgebung dieser Frequenz, in welcher der Hauptteil der verfälschten Frequenzen liegt, kleiner als 10 %. Der Anstieg oberhalb von $\tilde{\omega} = 1/2$ spielt kaum noch eine Rolle, da die Transferfunktion des \mathcal{B}_t^4 -Operators an diesen Stellen schon fast identisch Null ist. Die Nachglättung der Bildsequenz unterdrückt somit genau die kritischen Frequenzen, bei denen Aliasingeffekte zu erwarten sind. Sie unterdrückt natürlich auch die korrekten Frequenzen, die in diesem Bereich liegen. Dies würde aber auch bei einer Glättung mit einer großen Binomialmaske geschehen, da die Transferfunktion jeder geeigneten Filtermaske bei $\tilde{\omega} = 1/2$ identisch Null sein muß und sich durch die Unschärferelation zwischen Orts- und Fourierraum diesem Wert nur langsam nähern kann. Abbildung 4.3d zeigt die resultierende Transferfunktion der Filterkaskade im Vergleich zu einem \mathcal{B}_t^4 -Operator. Es ist deutlich zu erkennen, daß die Bereiche oberhalb von $\tilde{\omega} = 1/2$ stark reduziert wurden. Das Maximum des Überschwingers ist kleiner als 2 % des Maximalwertes bei $\tilde{\omega} = 0$. Dadurch wird zeitliches Aliasing fast vollständig unterdrückt.

4.3 Effektivität der Filterkaskade

Eine weitere Eigenschaft der Filterkaskade ist bemerkenswert: Obwohl die resultierende Filtermaske bei direkter Anwendung auf nur 9 Bilder der untersten Pyramidenstufe zugreift, hat sie die gleiche Effektivität wie eine \mathbf{B}_t^{12} -Maske mit 13 Filterkoeffizienten (Abb. 4.3e). Dies liegt daran, daß die \mathbf{B}_t^{casc} -Maske weiter auseinanderliegende Bilder stärker wichtet und damit eine größere Varianz hat. In Abschnitt 4.5 wird die effektive Varianz dieser Maske berechnet.

Abbildung 4.3f zeigt die resultierende Transferfunktion für eine 3- bzw. 4-stufige Pyramide bezüglich des Gitters der Originalsequenz. Die Überschwinger wurden durch die mehrfache Anwendung der Filterkaskade noch weiter reduziert und die Transferfunktion nähert sich immer mehr

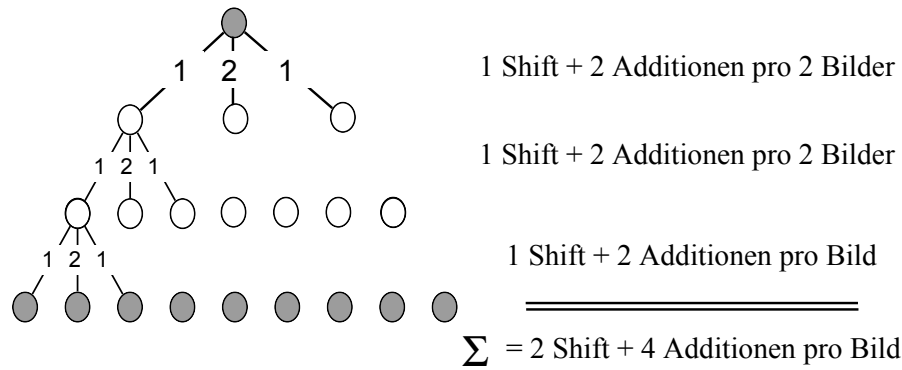


Abbildung 4.4: Detaillierte Berechnung der Gaußpyramide.

der einer idealen Gaußmaske an. Dies ist eine direkte Folge der Linearität der Faltung. Nach dem *Zentralen Grenzwertsatz* der Statistik hat eine lineare Überlagerung vieler Einzelprozesse immer die Normalverteilung als Grenzwert. Die Form der Einzelprozesse ist dabei von untergeordneter Bedeutung² [Reif, 1985]. Der Überschwinger an der Transferfunktion von \mathcal{B}_t^{casc} zeigt jedoch, daß die Filterkaskade selbst keine lineare Operation ist. Die Nichtlinearität wird durch den Reduktionsoperator \mathcal{R}_t verursacht, der eine nichtlineare Operation darstellt. Dies bedeutet, daß die Reihenfolge von Glättungen und Reduktion wichtig ist und nicht vertauscht werden darf.

| Gaußpyramide mit \mathcal{B}_t^{casc} | | | Binomialmaske gleicher Varianz | |
|---|----------------|----------------|--------------------------------|---------------|
| Stufenzahl | Shift | Additionen | Multiplikationen | Additionen |
| 1 | 2 | 4 | 13 | 12 |
| 2 | 3 | 6 | 61 | 60 |
| 3 | 3.5 | 7 | 253 | 252 |
| n | $4(1 - 0.5^n)$ | $8(1 - 0.5^n)$ | $4^{n+1} - 3$ | $4^{n+1} - 4$ |
| \vdots | \vdots | \vdots | \vdots | \vdots |
| ∞ | 4 | 8 | ∞ | ∞ |

Tabelle 4.1: Anzahl der Operationen pro Bild für Gaußpyramiden verschiedener Stufen im Vergleich zu einer direkten Faltung mit Binomialmasken gleicher Varianz.

Die gesamte Effektivität der Filterkaskade wird deutlich, wenn man die \mathbf{B}_t^4 -Maske weiter in 2 \mathbf{B}_t^2 -Masken zerlegt und die Reduktion so früh wie möglich durchführt (Abb. 4.4). Die Berechnung einer nachfolgenden Pyramidenebene benötigt nur 2 Shift Operationen und 4 Additionen pro Bild im Gegensatz zu 9 Multiplikationen und 8 Additionen pro Bild bei einer direkten Berechnung der Filterkoeffizienten! Für jede nachfolgende Pyramidenstufe reduziert sich durch die halbierte Auflösung die Anzahl der zusätzlich benötigten Operationen jeweils um die Hälfte. In Tabelle 4.1 ist für verschiedene Stufen die Anzahl der Filteroperationen pro Bild der untersten Pyramidenstufe im Vergleich zu einer direkten Berechnung der Filterkoeffizienten zusammenge-

²Dies ist auch der Grund dafür, daß Binomialmasken durch Faltung einzelner Rechteckmasken miteinander aufgebaut werden können.

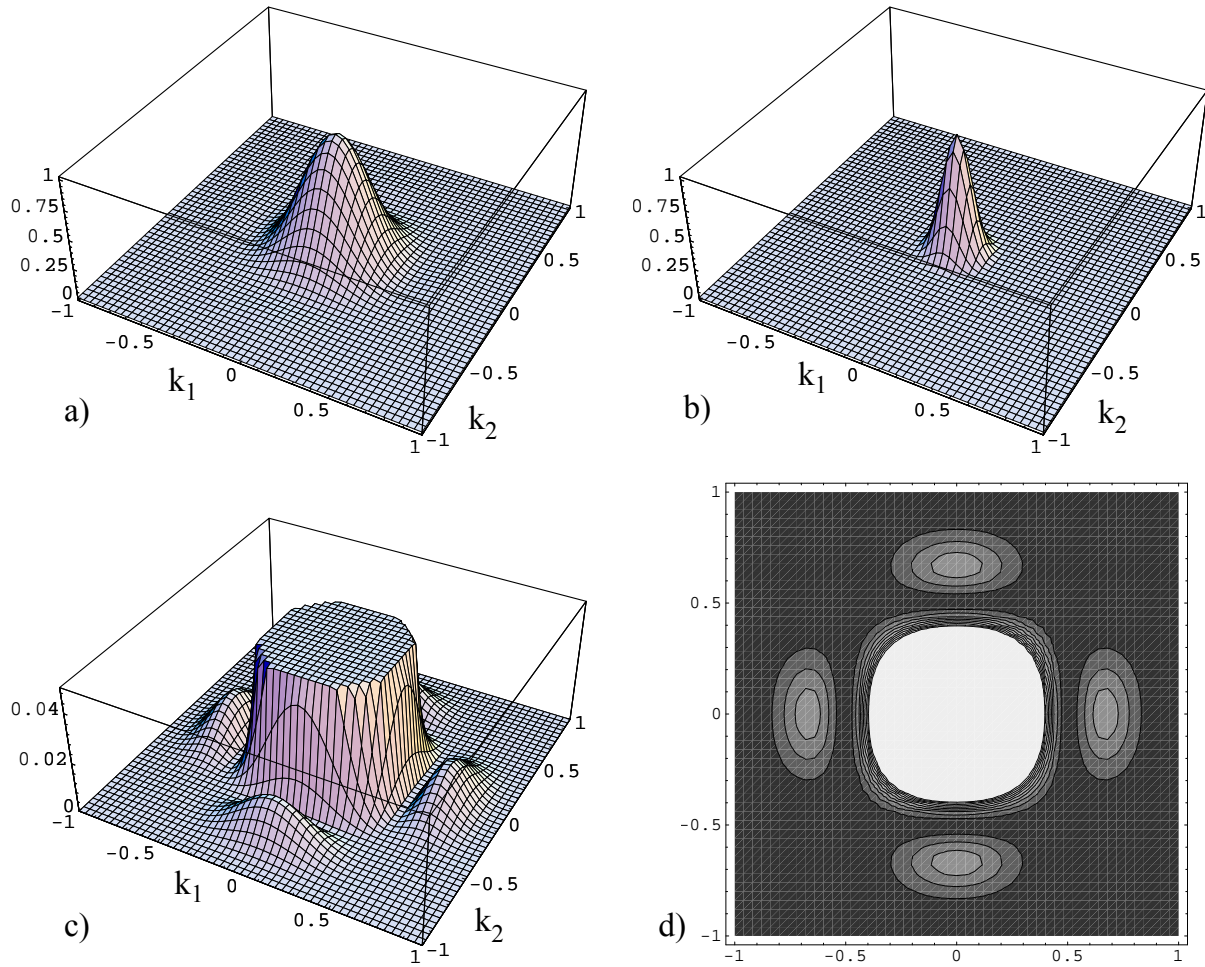


Abbildung 4.5: Transferfunktionen der Filterkaskade \mathcal{B}_l^{casc} zum Aufbau zweidimensionaler Pyramiden: a) Transferfunktion der ersten Pyramidenstufe. b) Effektive Transferfunktion einer 2-stufigen Pyramide. c) Transferfunktion der ersten Pyramidenstufe mit kleinerem Ausschnitt der z-Achse von 0 - 0.05 um die Überschwinger bei hohen Wellenzahlen hervorzuheben. d) Contourplot von b).

stellt. Während die Anzahl der Operationen und somit die Rechenzeit bei der direkten Berechnung mit der Stufenzahl stark ansteigt, konvergiert sie bei der Filterkaskade schnell und führt dazu, daß eine beliebig starke Glättung einer Bildsequenz mit maximal 4 Shift Operationen und 8 Additionen pro Bild der untersten Pyramidenstufe durchgeführt werden kann!

4.4 2D-Pyramiden

Die Filterkaskade, die in Abschnitt 4.3 zur effektiven Berechnung eindimensionaler zeitlicher Glättungen entwickelt wurde, kann auch für zweidimensionale räumliche Glättungen verwendet werden. Zum Aufbau einer 2D-Gaußpyramide kann die 1D-Filterkaskade abwechselnd auf Spalten und Zeilen des Bildes angewendet werden. Da bei der ersten Glättung bereits eine Reduktion des Bildes in dieser Richtung stattfindet, muß die zweite Glättung nur noch für die

Hälfte der Spalten bzw. Zeilen durchgeführt werden. Bei quadratischen Bildern erhöht sich der Rechenaufwand pro Pixel daher nur auf das 1.5-fache der Operationen in Tabelle 4.1. Die Transferfunktionen der 2D-Filterkaskade für ein- und zweistufige Pyramiden sind in Abbildung 4.5a und 4.5b dargestellt. Die Abbildungen 4.5c und 4.5d zeigen einen vergrößerten Ausschnitt in z -Richtung von 0 - 0.05. Man erkennt die Überschwinger bei hohen Wellenzahlen, die bereits bei der eindimensionalen Transferfunktion zu sehen waren. Durch diese lokalen Maxima wird die Transferfunktion anisotrop. Die anisotropen Teile sind jedoch kleiner als 2 % und somit vernachlässigbar klein. Der Hauptteil oberhalb von 5 % ist in guter Näherung isotrop.

4.5 Anwendung auf die zeitliche Tiefpaßfilterung

Durch ihre hohe Effektivität läßt sich die Pyramide dazu benutzen, die großen zeitlichen Glättungen in Kapitel 3 durchzuführen. Der Glättungsoperator \mathcal{B}_t^n kann durch eine bestimmte Pyramidenstufe N ersetzt werden. Um den Zusammenhang zwischen n und N zu finden, wird die effektive Varianz der Filterkaskade $\mathcal{B}_t^{casc} = \mathcal{B}_t^2 \mathcal{R}_t \mathcal{B}_t^4$ mit der Varianz von \mathcal{B}_t^n verglichen. Die effektive Varianz von \mathcal{B}_t^{casc} ergibt sich aus den folgenden beiden Eigenschaften von Binomialmasken:

- Zwei Glättungen mit Binomialmasken entsprechen einer einzigen Glättung mit einer größeren Binomialmaske, deren Varianz durch die Summe der Varianzen der beiden Glättungsmasken gegeben ist (Anhang C.2).
- Wird ein Filter auf einem größeren Gitter mit doppeltem Gitterabstand angewendet, dann hat die resultierende Transferfunktion eine vier mal so große Varianz, d. h. eine doppelt so große effektive Reichweite bezüglich des feineren Gitters (Anhang C.1).

Damit erhält man sofort für die Varianz $\sigma_{\mathcal{B}_t^{casc}}^2$ des Operators \mathcal{B}_t^{casc} :

$$\sigma_{\mathcal{B}_t^{casc}}^2 = \sigma_{\mathcal{B}_t^2 \mathcal{B}_t^4}^2 = 4 \underbrace{\sigma_{\mathcal{B}_t^2}^2}_{0.5} + \underbrace{\sigma_{\mathcal{B}_t^4}^2}_1 = 3. \quad (4.2)$$

Die Berechnung der nächst höheren Stufe entspricht somit einer Glättung mit dem Binomialoperator \mathcal{B}_t^{12} . Mit diesem Ergebnis zusammen mit den obigen Überlegungen, läßt sich für eine beliebige Pyramidenstufe N , mit effektiver Varianz σ_N^2 , die Größe einer korrespondierenden Binomialmaske \mathcal{B}_t^n der gleichen Varianz berechnen:

$$\sigma_{\mathcal{B}_t^n}^2 = \sigma_N^2 \iff n = 12 \sum_{k=1}^N 4^{(k-1)} = 12 \sum_{k=0}^{N-1} 4^k. \quad (4.3)$$

Der letzte Term in (4.3) stellt eine geometrische Reihe dar. Aufgelöst ergibt sich daraus:

$$n = 4(4^N - 1). \quad (4.4)$$

Die Umkehrfunktion von (4.4) liefert den gewünschten Zusammenhang zwischen einer vorgegebenen Binomialmaske \mathcal{B}_t^n und der dazu korrespondierenden Pyramidenstufe N :

$$N = \frac{\ln(\frac{n}{4} + 1)}{\ln 4}. \quad (4.5)$$

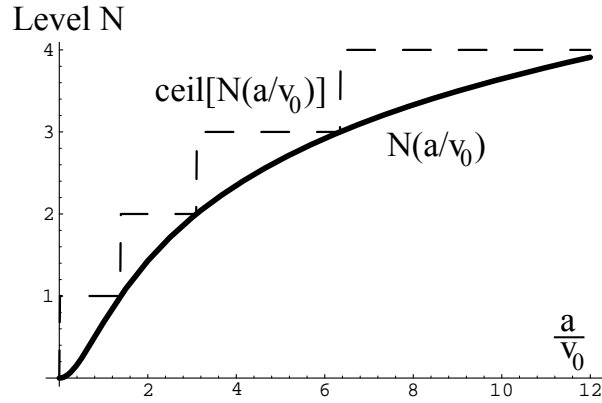


Abbildung 4.6: Stufe N der Gaußpyramide, die zur Detektion einer sich bewegenden Scheibe mit Radius a und Geschwindigkeit v_o benötigt wird.

Werden für n die Ergebnisse (3.16) und (3.27) eingesetzt, so erhält man für die periodische Struktur und die Scheibe (Kapitel 3)

$$N_{wave} = \frac{\ln \left((\vec{k}_o \cdot \vec{v}_o)^{-2} + 1 \right)}{\ln 4} \quad \text{und} \quad N_{disc} = \frac{\ln \left(\frac{25}{16} \frac{a^2}{v_o^2} + 1 \right)}{\ln 4}. \quad (4.6)$$

Dies ist die Bedingung für die minimale Pyramidenstufe, die benötigt wird um die beiden Objekte aus der Bildfolge zu entfernen. Abb. 4.6 zeigt diese Funktion für das Beispiel der Grauwertscheibe. Die Stufenfunktion $\text{ceil}(N)$ stellt die nächst größere ganze Zahl dar, die bei der Implementierung als Pyramidenstufe gewählt wird. Für beliebige Objekte der charakteristischen Länge Λ_c läßt sich (4.6) analog zu Abschnitt 3.5 verallgemeinern, indem man für n das Ergebnis (3.34) benutzt:

$$N_{gen} = \frac{\ln \left(c^2 \frac{\Lambda_c^2}{v_o^2} + 1 \right)}{\ln 4}, \quad 1 \leq c \leq \frac{5}{4}. \quad (4.7)$$

Der Formparameter c liegt zwischen den Werten für die beiden Grenzfälle: 1 für ideale periodische Strukturen und $\frac{5}{4}$ für ausgedehnte Objekte mit konstantem Grauwert.

Kapitel 5

Vergleich zwischen Originalbild und ruhendem Hintergrund

Im vorangegangenen Kapitel wurde gezeigt, wie zeitliche Binomialmasken angewendet werden müssen, um Objekte mit bestimmten Eigenschaften optimal aus einer Bildsequenz zu entfernen. Das Ergebnis ist der *ruhende Hintergrund* der Bildsequenz. Er enthält jedoch nicht nur ruhende Strukturen. Da der Grad der Dämpfung bei einer festen Maskengröße kontinuierlich von den Parametern Größe und Geschwindigkeit der Objekte abhängt, werden unterschiedliche Objekte verschieden stark unterdrückt. Die Differenz zwischen Originalbildern und den dazu korrespondierenden Hintergrundbildern liefert somit ein Signal, das von der Form, der Geschwindigkeit und dem Grauwert der Objekte in der Originalsequenz abhängt.

Dieses Kapitel befaßt sich mit der Berechnung eines Masses der Abweichung zwischen den Original- und Hintergrundbildern. Hierbei muß ebenfalls die Form der Objekte berücksichtigt werden, um darüber entscheiden zu können, ob es sich bei der Abweichung um bewegte Strukturen handelt oder andere Effekte, die einen Unterschied zwischen den Bildern verursachen.

5.1 Idealer Hochpaßfilter

Während in einem tiefpaßgefilterten Bild nur noch niedrige Frequenzen vorhanden sind, fehlen genau diese Frequenzen in der Differenz zwischen dem Originalbild und einem geglätteten Bild. Im Falle einer zeitlichen Glättung zusammen mit den Überlegungen in Kapitel 3 erhält man einen *zeitlichen Hochpaß*, der ruhende Strukturen aus der Bildsequenz entfernt.

Diese Verknüpfung eines Binomialfilters \mathbf{B}_t^n und einer punkweisen Differenz kann rein formal durch einen einzigen Filteroperator \mathcal{L} mit

$$\mathcal{L} = (\mathcal{I} - \mathcal{B}_t^n) \quad (5.1)$$

ersetzt werden, wobei \mathcal{I} den Einheitsoperator bezeichnet. Das Ergebnis $L(\vec{x}, t)$ dieser Filterung einer Bildfolge $g(\vec{x}, t)$ ist damit:

$$L(\vec{x}, t) = \mathcal{L} g(\vec{x}, t). \quad (5.2)$$

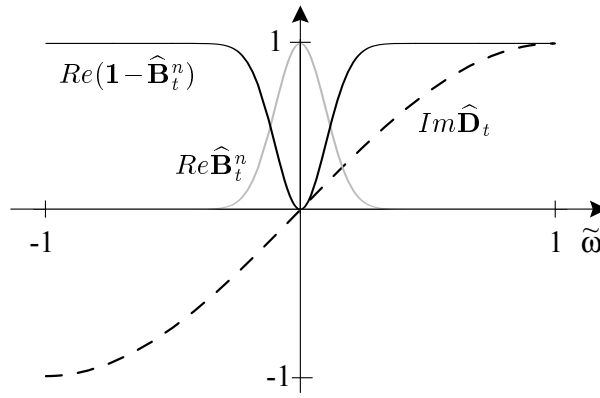


Abbildung 5.1: Transferfunktion eines zeitlichen Hochpaßfilters (rein reell) im Vergleich zu der Transferfunktion eines zeitlichen Ableitungsfilters (rein imaginär).

Abbildung 5.1 zeigt die effektive Transferfunktion dieser zusammengesetzten Operation. Verwendet man die kleinste Binomialmaske \mathbf{B}_t^1 mit

$$\mathbf{B}_t^1 = \frac{1}{2} \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad (5.3)$$

dann reduziert sich (5.1) auf

$$\mathcal{L} = \mathcal{D}_t \quad \text{mit} \quad \mathbf{D}_t = \frac{1}{2} \begin{bmatrix} 1 & -1 \end{bmatrix}, \quad (5.4)$$

d. h. die Differenz zweier aufeinanderfolgender Bilder. Die meisten der in Kapitel 2 beschriebenen Verfahren, die sich auf nur zwei Bilder beschränken, können damit nur eine Information liefern, die man im Spezialfall des einfachsten Filters der Menge der zeitlichen Hochpaßfilter dieser Arbeit erhält.

In Abb. 5.1 ist die Transferfunktion dieses *zeitlichen Differentialoperators* \mathcal{D}_t gestrichelt eingezeichnet. Da dieser Operator antisymmetrisch ist, hat er eine rein imaginäre Transferfunktion. Dies beeinflusst jedoch nur die Phase des Signals, die dadurch um $\pi/2$ verschoben wird. Der Verlauf der Transferfunktion von \mathcal{D}_t selbst gibt genau, wie bei der rein reellen Transferfunktion des Hochpaßfilters \mathcal{L}_t , die Dämpfung der Amplitude eines Signals der entsprechenden Frequenz an. Dadurch lassen sich die beiden Transferfunktionen direkt vergleichen und man erkennt, daß die des zeitlichen Differentialoperators im unteren Frequenzbereich deutlich von der des zeitlichen Hochpasses abweicht und nur Strukturen der Grenzfrequenz unverändert im Bild erhält. Dies entspricht genau dem in Kapitel 2 beschriebenen Problem, daß Verfahren, die einzig zwei aufeinanderfolgende Bilder berücksichtigen nur sehr schnelle Strukturen vollständig segmentieren können, wobei bei langsamen Objekten die Segmentierung auf einen schmalen Rand des Objektes beschränkt bleibt.

Obwohl der Filter \mathcal{L} ein fast idealer Hochpaßfilter ist, kann das Signal, das er liefert, nicht direkt zur Detektion der Objekte verwendet werden. Wie alle Filter, die spezielle periodische Strukturen aus einem Signal herausfiltern, liefert er keine Aussage über die Anwesenheit des Signals sondern das Signal selbst. Somit hat die Filterantwort die gleiche periodische Schwankung wie das Signal selbst und liefert an den Nulldurchgängen des Signals den Wert Null, unabhängig

von der Amplitude der Struktur. Bei einem zeitlichen Hochpaßfilter bedeutet dies, daß trotz Anwesenheit rein periodischer Strukturen in manchen Bildern nichts detektiert wird.

Um eine phasenunabhängige Filterantwort zu erhalten, müssen zusätzlich noch weitere Filterungen durchgeführt werden. Dabei wird mit einem *Hilbertfilter* \mathcal{H} (siehe [Jähne, 1993b], S. 182 ff) die Phase der Originalsequenz um $\pi/2$ verschoben. Eine anschließende Hochpaßfilterung dieser Bildsequenz liefert das gleiche Signal wie die erste Filterung mit einer Phasenverschiebung von $\pi/2$. Werden die quadrierten Filterergebnisse addiert, so erhält man das gewünschte phasenunabhängige Signal, d. h. die *Energieantwort*¹ \mathcal{E} des Filters \mathcal{L} :

$$\mathcal{E} = (\mathcal{L} \cdot \mathcal{L}) + (\mathcal{L} \mathcal{H} \cdot \mathcal{L} \mathcal{H}) \quad (5.5)$$

5.2 Approximation

Wie in [Jähne, 1993b] gezeigt wird, kann ein Hilbertfilter jedoch nur für Bandpaßfilter mit einem schmalen Paßband ausreichend gut realisiert werden. Für eine fast konstante Transferfunktion bei hohen Frequenzen steigt der Rechenaufwand stark an.

Für viele Anwendungen liefert eine Vereinfachung dieser Rechnung ausreichend gute Resultate. Im Fall der reinen Detektion von Objekten ist kein perfekt konstantes Signal nötig. Wenn man sicher sein kann, daß die Filterantwort bei Anwesenheit einer periodischen Struktur immer größer als Null ist und ihren minimalen Wert kennt, kann durch ein einfaches Schwellwertverfahren das gewünschte Signal vom Pixelrauschen getrennt werden.

Der erste Term der Gleichung (5.5) liefert schon ein Signal $(\mathcal{L} g)^2 \geq 0$. Es oszilliert noch mit der doppelten Frequenz der Originalstruktur um einen Mittelwert. Eine zusätzliche zeitliche Glättung reduziert diese Oszillation und verhindert, daß das Signal bis auf Null absinkt. Insgesamt erhält man dadurch ein *Ähnlichkeitsmaß* $S(\vec{x}, t)$:

$$S(\vec{x}, t) = \langle L^2(\vec{x}, t) \rangle_t = \mathcal{B}_t^m (\mathcal{L} g(\vec{x}, t))^2. \quad (5.6)$$

Wenn die gleiche Binomialmaske verwendet wird, die schon in (5.1) zur Berechnung von \mathcal{L} benutzt wurde, d. h. $m = n$, dann wird die Amplitude der Oszillation auf einen Faktor e^{-2} eingeschränkt, da die quadrierte Filterantwort die doppelte Frequenz des Originalsignals hat.

Eine zusätzliche Rechtfertigung für dieses Maß der Abweichung zwischen Originalbild und Hintergrund ergibt sich durch ein physikalisches Analogon. Betrachtet man das durch das tiefpaßgefilterte Hintergrundbild aufgespannte Grauwertgebirge als die Ruhelage einer elastischen Membran, so verursacht ein sich vor dem Hintergrund vorbeibewegendes Objekt eine Auslenkung der Membran aus ihrer Ruhelage. Diese Auslenkung ist durch die Differenz

$$L(\vec{x}, t) = (\mathcal{I} - \mathcal{B}_t^n) g(\vec{x}, t) \quad (5.7)$$

¹Im Fall periodischer Strukturen, die direkt die Amplitude eines physikalischen Signals darstellen, ergibt dies die Energie dieses Vorgangs. Im allgemeinen spricht man daher von der *Energieantwort* eines Filters, wobei nicht unbedingt eine physikalische Energie damit verknüpft sein muß. Es bezeichnet eher die Tatsache, daß von allen periodischen Strukturen des Bildes, die sich zu der beobachtbaren Grauwertdynamik des Bildes aufaddieren, dieser Anteil in den Einflußbereich des Filters fällt.

gegeben und ist mit einer *elastischen Deformationsenergie* $E(\vec{x}, t)$ der Form

$$E(\vec{x}, t) = \frac{1}{2} k L^2(\vec{x}, t) \quad (5.8)$$

verbunden, wobei k die Elastizitätskonstante der Membran bezeichnet. Eine zusätzliche Glättung der Energie mit einem Binomialfilter dient dazu, stochastische Energiefluktuationen zeitlich herauszumitteln. Überschreitet diese Deformationsenergie eine bestimmte Schwelle, ist dies auf die Anwesenheit eines bewegten Objektes zurückzuführen.

Gleichung (5.8) unterscheidet sich von (5.6) außer der fehlenden Glättung durch die Elastizitätskonstante k . Sie liefert ein Maß dafür, wie stark eine Auslenkung bei einer bestimmten Membran gewertet wird. Wenn diese Konstante von der Position im Bild abhängt, so führt dies zu lokalen Schwankungen der Empfindlichkeit der Segmentierung. Diese Überlegungen gaben Anlaß zur Entwicklung einer *lokalen Schwelle*, die in Kapitel 6 beschrieben wird und das Segmentierungsergebnis im Gegensatz zu einer *globalen Schwelle* deutlich verbessert.

5.3 Austausch der Koordinaten

Durch die Berechnung des Ähnlichkeitsmaßes S (5.6) spart man im Vergleich zu (5.5) die Bestimmung der Hilberttransformierten der Bildsequenz. Da (5.6) trotzdem immer noch zwei zeitliche Filterungen enthält, stellt sich die Frage, ob nicht eine der beiden Glättungen durch eine räumliche Glättung ersetzt werden kann. Diese kann meistens viel effizienter berechnet werden und muß im allgemeinen ohnehin angewendet werden, um das allgegenwärtige Pixelrauschen zu unterdrücken.

Den Schlüssel zur Antwort liefert die Tatsache, daß bei einer Bewegung die Position eines Objektes im Bild eindeutig mit der Zeit verknüpft ist. Im folgenden wird die Beziehung zwischen der gewünschten zeitlichen Glättung und der dazu benötigten räumlichen Glättung für die bereits in Kapitel 3 verwendeten Beispielobjekte hergeleitet.

5.3.1 Ideale periodische Struktur

Für die ebene Welle (3.10) gilt $g(\vec{x}, t) = g(\vec{k}_o(\vec{x} - \vec{v}_o t))$. Damit schreibt sich die zeitliche Glättung von $L^2(\vec{x}, t)$ in Gleichung (5.6) in Integralform als

$$\langle L^2(\vec{x}, t) \rangle_{t, \sigma_t} = \frac{1}{\sqrt{2\pi}\sigma_t} \int_{-\infty}^{\infty} dt' L^2(\vec{k}_o \cdot (\vec{x} - \vec{v}_o t')) \exp\left[-\frac{(t - t')^2}{2\sigma_t^2}\right]. \quad (5.9)$$

Sei im folgenden $x_k = \frac{\vec{k}_o \cdot \vec{x}}{k_o}$ und $v_k = \frac{\vec{k}_o \cdot \vec{v}_o}{k_o}$, mit $k_o = |\vec{k}_o|$, die Projektion des Ortsvektors und der Geschwindigkeit auf die Richtung des Wellenvektors. Ohne Beschränkung der Allgemeinheit sei weiterhin $\vec{k}_o \cdot (\vec{x} - \vec{v}_o t) = 0$, d. h. man betrachtet Phase Null. Damit ergibt sich die Beziehung zwischen der Position der Welle und der Zeit als $x_k = v_k t$ und $v_k = dx_k/dt$. Mit diesen

Transformationen ergibt sich aus Gleichung (5.9):

$$\begin{aligned}
\langle L^2(\vec{x}, t) \rangle_{t, \sigma_t} &= \frac{1}{\sqrt{2\pi}\sigma_t} \int_{-\infty}^{\infty} \frac{dx'_k}{v_k} L^2(k_o(v_k t - x'_k)) \exp \left[-\frac{(x_k - x'_k)^2}{2\sigma_t^2 v_k^2} \right] \\
&= \frac{1}{\sqrt{2\pi}(\sigma_t v_k)} \int_{-\infty}^{\infty} dx'_k L^2(k_o(x'_k - v_k t)) \exp \left[-\frac{(x_k - x'_k)^2}{2(\sigma_t v_k)^2} \right] \\
&= \langle L^2(\vec{x}, t) \rangle_{x_k, \sigma_x} \quad \text{mit} \quad \sigma_x = \sigma_t \frac{\vec{k}_o \cdot \vec{v}_o}{k_o}.
\end{aligned} \tag{5.10}$$

Gleichung (5.10) zeigt, daß eine *zeitliche Glättung* durch eine eindimensionale *räumliche* Glättung in Richtung des Vektors \vec{k}_o ersetzt werden kann, d. h. senkrecht zu den Linien konstanter Helligkeit der Welle. Die resultierende Filtermaske hat die Breite σ_x . Mit $\sigma_t = (\vec{k}_o \cdot \vec{v}_o)^{-1}$ (3.14) erhält man:

$$\sigma_x = \vec{k}_o^{-1} = \frac{\lambda_o}{2\pi} = \lambda. \tag{5.11}$$

5.3.2 Kreisförmiges Objekt

Für ein kreisförmiges Objekt, das sich mit einer Geschwindigkeit \vec{v}_o in beliebiger Richtung bewegt, gilt $g(\vec{x}, t) = g(x_v - v_o t)$, mit $x_v = \frac{\vec{v}_o \cdot \vec{x}}{v_o}$ und $v_o = \|\vec{v}_o\|$. Analog zu Gleichung (5.9) ergibt sich damit aus (5.6):

$$\langle L^2(\vec{x}, t) \rangle_{t, \sigma_t} = \frac{1}{\sqrt{2\pi}\sigma_t} \int_{-\infty}^{\infty} dt' L^2(x_v - v_o t') \exp \left[-\frac{(t - t')^2}{2\sigma_t^2} \right]. \tag{5.12}$$

Betrachtet man wieder Phase Null, so ergibt sich $x_v = v_o t$ und $v_o = dx_v/dt$. Damit erhält man aus (5.12):

$$\begin{aligned}
\langle L^2(\vec{x}, t) \rangle_{t, \sigma_t} &= \frac{1}{\sqrt{2\pi}\sigma_t} \int_{-\infty}^{\infty} \frac{dx'_v}{v_o} L^2(v_o t - x'_v) \exp \left[-\frac{(x_v - x'_v)^2}{2\sigma_t^2 v_o^2} \right] \\
&= \frac{1}{\sqrt{2\pi}(\sigma_t v_o)} \int_{-\infty}^{\infty} dx'_v L^2(x'_v - v_o t) \exp \left[-\frac{(x_v - x'_v)^2}{2(\sigma_t v_o)^2} \right] \\
&= \langle L^2(\vec{x}, t) \rangle_{x_v, \sigma_x} \quad \text{mit} \quad \sigma_x = \sigma_t v_o.
\end{aligned} \tag{5.13}$$

Analog zu Gleichung (5.10) stellt dies wieder eine Substitution der zeitlichen Glättung durch eine räumliche Glättung dar. Anders als in (5.10) muß bei der Scheibe die Glättung in Richtung der Geschwindigkeit \vec{v}_o erfolgen. Für die Breite σ_x der resultierenden Filtermaske findet man mit $\sigma_t = \frac{5}{4} \frac{a}{v_o}$ (3.26):

$$\sigma_x = \frac{5}{4} a. \tag{5.14}$$

Gleichungen (5.11) und (5.14) zeigen, daß die Größe der benötigten räumlichen Faltungsmaske nur von der Größe der Objekte abhängt und nicht von deren Geschwindigkeit. Für das runde

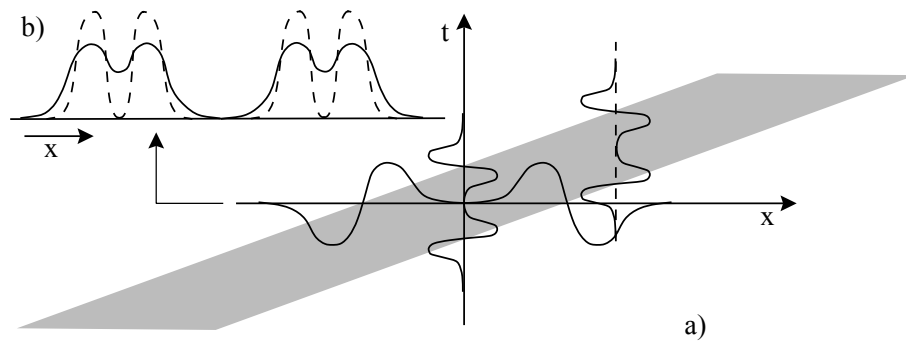


Abbildung 5.2: Eindimensionale Bewegung eines starren Körpers: a) Projektion der Filterantwort des zeitlichen Hochpaßfilters auf die Ortskoordinate. b) Räumliche Glättung des quadrierten Signals.

Objekt ist die Vertauschung von räumlicher und zeitlicher Glättung (5.13) in Abbildung 5.2 als Orts-Zeit-Bild illustriert. Zur besseren Darstellung ist einzig die Projektion auf die Bewegungsrichtung gezeigt. Man erkennt, wie durch die Verschiebung des Objektes von Bild zu Bild die Filterantwort des Hochpaßfilters eine pixelabhängige Phasenverschiebung erhält. Dadurch wird die zeitliche Form des Filtersignals im Ortsraum wiederholt, wobei das Signal in Abhängigkeit von der Geschwindigkeit skaliert wird.

Insgesamt erhält man durch die Überlegungen dieses Abschnitts die wichtigen Ergebnisse:

- Die Berechnung der Energieantwort eines Bandpaßfilters über eine Hilbertfilterung kann dadurch ersetzt werden, daß die Filterantwort quadriert und anschließend glättet wird. Dadurch spart man die rechenaufwendige Hilbertfilterung und eine zweite Anwendung des Bandpaßfilters.
- Bei der zeitlichen Filterung einer sich bewegenden Struktur kann zusätzlich die zweite zeitliche Glättung durch eine räumliche Glättung ersetzt werden. Das ist möglich, da die Hochpaßfilterung bereits die nötige zeitliche Information liefert und diese auf die Orts-Koordinaten projiziert, so daß die zweite Filteroperation vollständig in Richtung der räumlichen Koordinaten stattfinden kann.

5.4 Motion blur

Mit den Ergebnissen des Abschnitts 5.3 kann auch die bereits in Kapitel 2 beschriebene *Bewegungsunschärfe* (engl. *motion blur*) quantitativ beschrieben werden.

Die endliche Belichtungszeit eines Videobildes bei der Bildaufnahme bewirkt, daß sich ein bewegtes Objekt während der Belichtungszeit Δt an verschiedenen Positionen des Bildes befindet. Die Gesamthelligkeit des Objektes verteilt sich über die gesamte Fläche, die von ihm während dieser Zeit überstrichen wird. Ein einzelnes Bild stellt somit keine zeitliche Punktinformation mehr dar. Die auf einen bestimmten Zeitpunkt t_i projizierte Information $g(\vec{x}, t_i)$ ergibt sich aus

der integrierten Helligkeit einer zeitlichen Umgebung. Dies entspricht somit der Faltung

$$g(\vec{x}, t_i) = \int_{t_i - \Delta t/2}^{t_i + \Delta t/2} dt g(\vec{x}, t) \quad (5.15)$$

des zeitlich kontinuierlichen Vorgangs mit einer Rechteckmaske, deren Breite durch die Belichtungszeit gegeben ist und einer anschließenden diskreten, zeitlichen Abtastung. Jede Bildsequenz wird daher bereits bei der Bildaufnahme tiefpaßgefiltert. Im Gegensatz zu der in Kapitel 3 beschriebenen Glättung erfolgt hierbei die Mittelung nur über einen Bruchteil des zeitlichen Bildabstandes. Die Glättung mit einer Rechteckmaske ist das zeitliche Analogon zur sogenannten *Standardabtastung* im Ortsraum, bei der über die Fläche eines Sensorelementes des CCD Chip integriert wird, bevor diese Helligkeit einem einzelnen Pixel zugeordnet wird. Während die Größe der räumlichen Integrationsumgebung durch die Größe der Sensorelemente fest vorgegeben ist, kann die Belichtungszeit verkürzt werden, um schnelle Objekte mit ausreichender Kantenschärfe, d. h. Positionsgenauigkeit abzubilden.

Der Zusammenhang zwischen der Belichtungszeit Δt und der Stärke der räumlichen Verschmierung ergibt sich aus den gleichen Überlegungen, die zum Austausch der Koordinaten geführt haben (5.12). Ein starres Objekt, das sich mit der Geschwindigkeit v_o bewegt, erfährt daher eine Glättung in Bewegungsrichtung mit einer räumlichen Rechteckmaske. Die Größe dieser Glättungsmaske hängt sowohl von der Belichtungszeit (Varianz der zeitlichen Rechteckmaske σ_t^2) als auch von der Geschwindigkeit ab:

$$\sigma_x = \sigma_t v_o. \quad (5.16)$$

Ruhende Objekte werden dadurch nicht beeinflusst. Für kleine Geschwindigkeiten $v_o \approx 1 \text{ Pixel/Bild}$ gilt $\sigma_x \approx \sigma_t$. Damit beschränkt sich die Verschmierung auf ein Pixel am Objektrand. Bei sehr großen Geschwindigkeiten, $v_o \approx 1 \text{ Objektdurchmesser/Bild}$ entspricht die Größe der räumlichen Glättung der Objektgröße und führt auf eine deutliche Abschwächung des Signals. Dadurch ist das Objekt bereits in der Originalsequenz in gleichem Maße gedämpft wie bei einer zeitlichen Glättung zur Extraktion des Hintergrundes und deshalb schwerer zu detektieren. Eine Verkürzung der Belichtungszeit führt nach Gleichung (5.16) zu einer kleineren Verschmierung schneller Objekte. Dies ändert jedoch nichts an der Tatsache, daß die Helligkeit der Objekte geringer wird. Im Gegensatz zum Verlust an Helligkeit durch motion blur werden hier jedoch alle Objekte dunkler, unabhängig von ihrer Geschwindigkeit.

Dies zeigt insgesamt, daß sehr schnelle Teilchen, wie sie in der vorliegenden Arbeit beobachtet wurden, prinzipiell sehr schwer zu detektieren sind und keine Möglichkeit besteht, diese Schwierigkeiten bereits bei der Bildaufnahme zu beheben.

5.5 Verallgemeinerung für beliebige Objekte

Die Ergebnisse der Abschnitte 5.2 und 5.3 zeigen, daß zur Detektion der bewegten Objekte außer der zeitlichen Glättung (Kapitel 3) nur eine Punktoperation und eine weitere räumliche Glättung benötigt werden. Die Richtung der räumlichen Faltung sowie die Größe der notwendigen Filtermasken hängt sowohl von der Art der Objekte als auch von deren Bewegungsrichtung ab.

Da in realen Anwendungen beliebige Objekte in Bildsequenzen vorkommen können und vor der Bewegungssegmentierung weder ihre Bewegungsrichtung noch ihre Größe bekannt sind, müssen die Ergebnisse der Gleichungen (5.10) und (5.13) verallgemeinert werden.

Die eindimensionalen Glättungen in Richtung des Wellenvektors der Grauwertwelle (5.10) oder in Richtung der Geschwindigkeit der Scheibe (5.13) werden ersetzt durch eine zweidimensionale räumliche Glättung mit der isotropen Binomialmaske \mathcal{B}_x^n . Dies ist möglich, da Binomialmasken separabel sind ([Jähne, 1993a]). Eine 2D-Glättung kann daher zerlegt werden in eine 1D-Glättung, in jede beliebige, benötigte Richtung und eine zweite Glättung senkrecht zu dieser Richtung. Während die erste Glättung die Oszillationen der Filterantwort des zeitlichen Hochpaßfilters reduziert, unterdrückt die zweite Faltung das allgegenwärtige Pixelrauschen.

Wie aus (5.11) und (5.14) ersichtlich wird, hängt die Größe der benötigten Filtermaske ausschließlich von der Größe der Objekte ab und nicht von deren Geschwindigkeit. Damit läßt sich die Maskengröße analog zu Abschnitt 3.5 aus der charakteristischen Länge Λ_c der Objekte errechnen. Mit dem Formparameter c können die Gleichungen (5.11) und (5.14) vereinigt werden zu:

$$\sigma_{\vec{x}} = c\Lambda_c, \quad 1 \leq c \leq \frac{5}{4}. \quad (5.17)$$

Mit Gleichung (3.15) ergibt sich daraus der Grad n_x der benötigten räumlichen Binomialmaske

$$n_x = 4c^2\Lambda_c^2, \quad 1 \leq c \leq \frac{5}{4}. \quad (5.18)$$

Insgesamt berechnet sich damit das *Ähnlichkeitsmaß* S zu:

$$S(\vec{x}, t) = \mathcal{B}_x^{n_x} [(\mathcal{I} - \mathcal{B}_t^{n_t}) g(\vec{x}, t)]^2, \quad (5.19)$$

wobei sich der Grad n_t der zeitlichen Binomialmaske aus Gleichung (3.34) ergibt. Da die benötigte räumliche Glättungsmaske für ausgedehnte Objekte unter Umständen sehr groß werden kann, bietet sich auch hier zur Berechnung eine Pyramide an. Zum Aufbau dieser räumlichen 2D-Pyramide kann ebenfalls die in Kapitel 4 entwickelte Mehrgitterstruktur verwendet werden. Die gesamte Bewegungssegmentierung kann somit von vornherein auf einer Pyramidenstufe geschehen. Die effektive Reichweite der korrespondierenden Glättungsmaske muß dabei der Größe der Objekte entsprechen.

Westberg [Westberg, 1989] zeigt, daß eine Approximation eines *Likelihood Ratio Test* für den Fall sehr kleiner Kontraste der Objekte und schwer zu erkennender Bewegung auf ein ähnliches Schwellwertverfahren führt. Dabei wird auf einer räumlichen Pyramide die Größe

$$\ell(I, k) = \sum_{(x, y) \in U} I^2(x, y) \quad (5.20)$$

mit einer Schwelle $\eta(k)$ verglichen, die von der jeweiligen Pyramidenstufe k abhängt. $I(\vec{x}, t)$ bezeichnet die Differenz zweier aufeinanderfolgender Bilder. U ist eine lokale Umgebung des betrachteten Pixels. Im Gegensatz zu der räumlichen Glättung in (5.19) werden hierbei alle benachbarten Pixel gleich gewichtet. Außerdem führt die Betrachtung von nur zwei Bildern auf die eingangs genannten Probleme. Insgesamt zeigt das Verfahren von *Westberg* aber, daß sich das in dieser Arbeit benutzte Ähnlichkeitsmaß $S(\vec{x}, t)$ außer den filtertechnischen Gesichtspunkten dieses Kapitels durchaus auf wahrscheinlichkeitstheoretische Überlegungen stützen kann.

5.6 Größe des Ähnlichkeitsmaßes für die beiden Beispiele

Mit der erwarteten Dämpfung eines zeitlich variierenden Signals durch die in Kapitel 3 berechneten zeitlichen Binomialfilter kann für die beiden Beispiele die erwartete Größe des Ähnlichkeitsmaßes S berechnet werden.

5.6.1 Periodische Struktur

Wird die Amplitude A_{Welle} der Grauwertwelle (3.10) auf den Faktor $e^{-1/2}$ eingeschränkt, ergibt sich als Ergebnis dieser Glättung die periodische Struktur

$$g_t(\vec{x}, t) = \mathcal{B}_t^{n_t} g(\vec{x}, t) = A_{Welle} (1 - e^{-1/2}) \sin \left(\vec{k}_o \cdot (\vec{x} - \vec{v}_o t) \right). \quad (5.21)$$

Eingesetzt in (5.19) erhält man damit:

$$\begin{aligned} S(\vec{x}, t) &= \mathcal{B}_{\vec{x}}^{n_x} \left[A_{Welle}^2 (1 - e^{-1/2})^2 \sin^2 \left(\vec{k}_o \cdot (\vec{x} - \vec{v}_o t) \right) \right] \\ &= A_{Welle}^2 (1 - e^{-1/2})^2 \mathcal{B}_{\vec{x}}^{n_x} \left[\sin^2 \left(\vec{k}_o \cdot (\vec{x} - \vec{v}_o t) \right) \right], \end{aligned} \quad (5.22)$$

wobei in der zweiten Zeile die Linearität der Faltungsoperation ausgenutzt wurde. Mit der Näherung für hinreichend große n_x

$$\mathcal{B}_{\vec{x}}^{n_x} \left[\sin^2 \left(\vec{k}_o \cdot (\vec{x} - \vec{v}_o t) \right) \right] \cong \frac{1}{2} \quad (5.23)$$

reduziert sich (5.22) auf den konstanten Wert

$$S(\vec{x}, t) = \frac{A_{Welle}^2}{2} (1 - e^{-1/2})^2. \quad (5.24)$$

Im Idealfall gilt dies nur für die Wertepaare (\vec{x}, t) , für die in der Originalsequenz die Struktur $g(\vec{x}, t)$ vorhanden ist. Für alle anderen Bereiche gilt $g(\vec{x}, t) = g_t(\vec{x}, t) \equiv 0$ und damit $S(\vec{x}, t) = 0$.

Reale Bildszenen sind jedoch mit Rauschen behaftet. Dadurch wird das Signal S auch an Stellen von Null verschieden sein, an denen sich nichts bewegt. Um diese Abweichungen zwischen Originalbild und Hintergrund von der eigentlichen Bewegung zu trennen, muß das Signal S durch eine Schwelle binarisiert werden. Nur Werte, die diese Schwelle überschreiten, werden als Bewegung anerkannt und kleinere Werte verworfen. Der Maximalwert der Schwelle ist durch die erwartete Größe des Ähnlichkeitsmaßes in Gleichung (5.24) gegeben. Der Minimalwert der Schwelle hängt von der Stärke des Rauschens ab. Da die Berechnung von S für hohe Frequenzen eine ideale Hochpaßfilterung darstellt (Abb. 5.1), bleibt Rauschen darin vollständig erhalten. Der Erwartungswert von S für Rauschen mit der Varianz σ_n^2 (Index $n \hat{=}$ noise) ist durch die Varianz des Rauschens selbst gegeben (Anhang B). Für die Schwelle T_W zur Binarisierung von S gilt somit:

$$\sigma_n^2 \leq T_W \leq \frac{A_{Welle}^2}{2} (1 - e^{-1/2})^2. \quad (5.25)$$

5.6.2 Kreisförmiges Objekt

Analog zu Abschnitt 5.6.1 kann der erwartete Wert S für die Grauwertscheibe berechnet werden. Der Grauwert am Mittelpunkt \vec{x}_o der Scheibe wird durch die zeitliche Glättung auf den Faktor $e^{-1/2}$ seines ursprünglichen Wertes eingeschränkt. Da alle anderen Punkte der Scheibe jedoch stärker geglättet werden, ist das zu erwartende Signal S in der Scheibenmitte am kleinsten. Dort nimmt S den Wert

$$S(\vec{x} = \vec{x}_o, t) = A_{Scheibe}^2 (1 - e^{-1/2})^2 \quad (5.26)$$

an. Dieser Wert stellt also den Maximalwert der Binarisierungsschwelle dar. Der Minimalwert ist wiederum durch die Varianz σ_n^2 des Rauschens gegeben (Anhang B). Für die Binarisierungsschwelle T_S gilt daher:

$$\sigma_n^2 \leq T_S \leq A_{Scheibe}^2 (1 - e^{-1/2})^2. \quad (5.27)$$

5.6.3 Schwelle für beliebige Objekte

Wie aus den Gleichungen (5.25) und (5.27) ersichtlich wird, hängt der Erwartungswert des Ähnlichkeitsmaßes S sowohl von der *Form* als auch vom *Kontrast* der Objekte ab. Kontrastarme Objekte verursachen bei gleicher Geschwindigkeit ein wesentlich kleineres Signal als kontrastreiche Objekte. Eine *globale Schwelle*, die nicht von der Position im Bild abhängt, kann dadurch keine optimalen Ergebnisse liefern. Eine ideale Segmentierungsschwelle T sollte in der Lage sein, sich lokal den oben genannten Objektparametern anzupassen. Diese Parameter sind im allgemeinen a priori nicht bekannt. Daher muß eine Schwelle gefunden werden, die sich unter Berücksichtigung der qualitativen Aussagen der Gleichungen (5.25) und (5.27) automatisch der lokalen Bildstruktur anpaßt, ohne ein genaues Vorwissen über die Bildobjekte zu benötigen. Eine solche *lokale Schwelle* wird in Kapitel 6 entwickelt.

5.7 Überprüfung der Ergebnisse an Testbildern

Obwohl die Gleichungen (5.25) und (5.27) nicht direkt eine optimale lokalen Schwelle liefern, können sie doch dazu benutzt werden, die Ergebnisse der vorangegangenen Kapitel experimentell zu bestätigen. Wenn die Überlegungen über den Zusammenhang zwischen gewünschtem Dämpfungsgrad des Objektes und benötigter Filtergröße (4.6) stimmen und der Rauschanteil an S durch die Varianz des Rauschen gegeben ist, dann stellen sie exakte Lösungen für die minimale und maximale Schwelle dar. Eine etwas größere Schwelle sollte zu Lücken im segmentierten Objekt führen und eine zu kleine Schwelle Rauschen segmentieren. Dies wurde anhand zweier computergenerierter Testsequenzen durchgeführt. Abbildung 5.3 zeigt das erste, mittlere und letzte Bild einer Sequenz von 59 Bildern. Diese Anzahl wird benötigt um eine 3-stufige, zeitliche Gaußpyramide aufzubauen. Die dunklere der beiden Scheiben bewegt sich dabei mit der Geschwindigkeit $\vec{v}_o = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ Pixel/Bild. Den einzelnen Bildern wurde, unabhängig voneinander, normalverteiltes Rauschen der Varianz $\sigma_{noise} = 7$ Grauwerte überlagert. Abbildung 5.4 zeigt eine ähnliche Szene für eine ebene Grauwertwelle. In Tabelle 5.1 sind die verwendeten

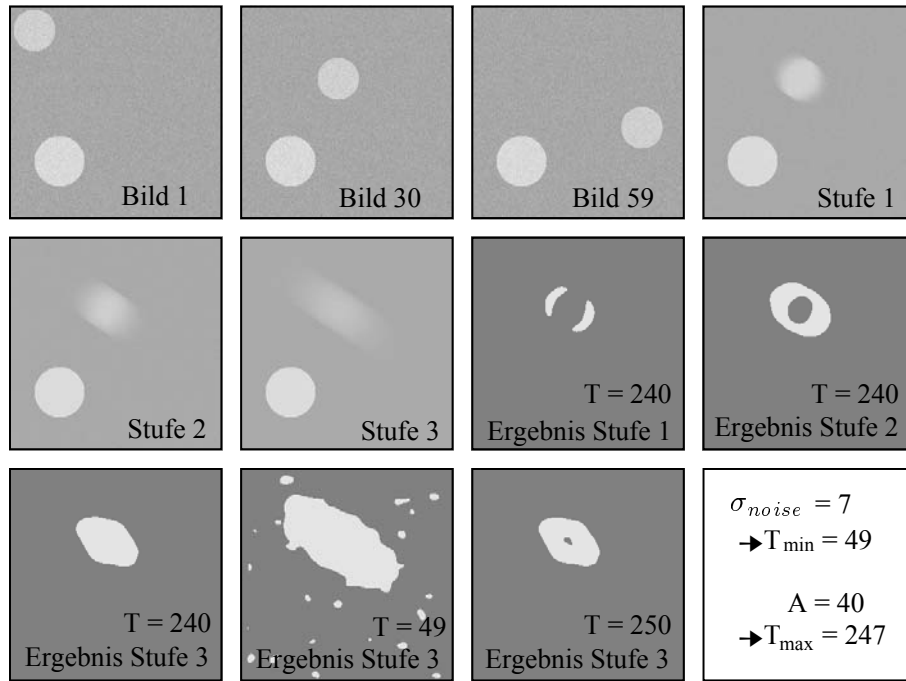


Abbildung 5.3: Computer-generierte Test-Sequenz: Die Bildnummern beziehen sich auf die Bildfolge. Stufe n bezeichnet die Stufe der zeitlichen Gaußpyramide. Die verwendete Segmentierungsschwelle wird mit T bezeichnet (Tabelle 5.1).

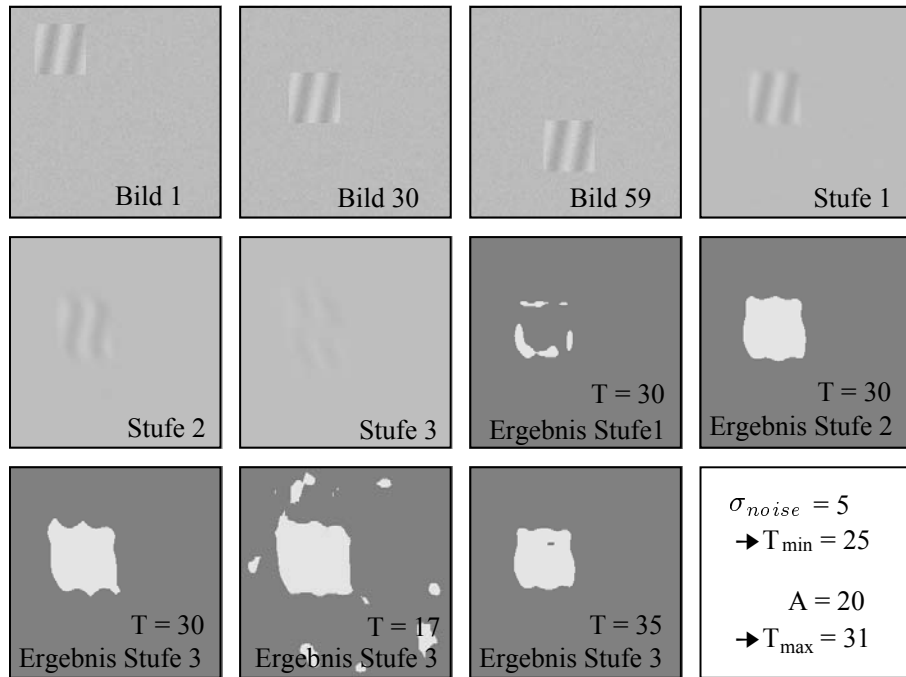


Abbildung 5.4: Computer-generierte Test-Sequenz: Die Bildnummern beziehen sich auf die Bildfolge. Stufe n bezeichnet die Stufe der zeitlichen Gaußpyramide. Die verwendete Segmentierungsschwelle wird mit T bezeichnet (Tabelle 5.1).

| | Welle | Scheibe |
|--|--|--|
| Größe $[Pixel]$ | $\Lambda = 30$ | $a = 24$ |
| Wellenvektor $\vec{k}_o [Pixel^{-1}]$ | $\frac{2\pi}{30} \begin{pmatrix} 1 \\ 0.2 \end{pmatrix}$ | — |
| Kontrast $[Grauwerte]$ | $A = 20$ | $A = 40$ |
| Geschwindigkeit $\vec{v}_o [Pixel/Bild]$ | $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ | $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ |
| Rauschen $\sigma_n [Grauwerte]$ | 5 | 7 |
| minimale Schwelle | 25 | 49 |
| maximale Schwelle | 31 | 247 |
| berechnete Pyramidenstufe n | 1.83 | 3.06 |

Tabelle 5.1: Zusammenstellung der verwendeten Parameter und der daraus folgenden Pyramidenstufen und Segmentierungsschwellen für die beiden Testsequenzen.

Parameter der beiden Testsequenzen und die damit aus (4.6), (5.25) und (5.27) berechneten Pyramidenstufen und Segmentierungsschwellen zusammengestellt.

Es zeigt sich, daß die berechneten Größen für die beiden Beispielobjekte, bis auf kleine Abweichungen durch die statistische Eigenschaft von Rauschen, exakte Lösungen darstellen. Für die optimale Segmentierungsschwelle werden die Objekte ab der (aufgerundeten) berechneten Pyramidenstufe über ihre gesamte Fläche detektiert. Während eine nur wenig größere Segmentierungsschwelle Lücken in den Objekten erzeugt, wird durch eine nur wenig kleinere Schwelle als die berechnete minimale Schwelle, deutlich Rauschen segmentiert. In Abb. 5.4 wird eine weitere Tatsache deutlich, die bereits in Abschnitt 3.5 beschrieben wurde: Zur Detektion der Objekte muß im wesentlichen die kleinste Textur berücksichtigt werden. Die Grauwertwelle selbst stellt ein quadratisches Objekt dar, dessen Textur durch die Welle gegeben ist. Die Segmentierung liefert ein optimales Ergebnis für Parameter, die nur die Form der Welle selbst berücksichtigen. Für eine zu kleine Pyramidenstufe wurde aus diesem Grund nur der Rand der quadratischen Grundfläche detektiert. Dieser besteht aus sehr kleinen Strukturen, für welche die erste Pyramidenstufe bereits ein optimales Ergebnis liefert.

Kapitel 6

Lokale Schwelle durch Fuzzy-Logik

In der Geschwindigkeit, Daten nach vorgegebenen logischen Zusammenhängen zu verknüpfen, sind Computer dem menschlichen Gehirn weit überlegen. Sie sind jedoch nicht in der Lage, aus unsicheren Daten Schlußfolgerungen zu ziehen, qualitative Zusammenhänge auf andere Anwendungen zu übertragen oder aus Erfahrung zu lernen. Diese Fähigkeiten zeichnen das menschliche Gehirn aus und werden im allgemeinen als Intelligenz bezeichnet. Ansätze der *künstlichen Intelligenz* (KI) versuchen diese Fähigkeiten auf elektronische Rechenmaschinen zu übertragen.

Regeltechnik und Qualitätskontrolle sind Anwendungen, in denen aus unsicheren Daten Schlußfolgerungen gezogen werden und auf Abweichungen von bekannten Mustern reagiert werden muß. Ein Experte, der diese Aufgabe übernimmt, macht dies oft intuitiv, ohne genau zu wissen, wie seine Handlungen von den verschiedenen Parametern beeinflußt werden. Diese Unsicherheit liegt in der Tatsache begründet, daß die meisten Systeme *nichtlinear* sind. Ein direkter funktionaler Zusammenhang zwischen den Meß- und Stellgrößen ist daher in den meisten Fällen nicht zu finden.

Die Unfähigkeit von Computern mit unsicheren Daten umzugehen, liegt in der Logik begründet, die allen Rechenmaschinen zu Grunde liegt: *Boolsche Algebra*. Diese kennt nur die Zustände *richtig* oder *falsch* bzw. *ja* oder *nein*. Obwohl das komplette Zahlensystem durch Kombinationen dieser beiden Zustände aufgebaut werden kann, bleiben alle Verknüpfungen zwischen verschiedenen Werten auf die Aussagen *richtig* oder *falsch* beschränkt. Es besteht keine Möglichkeit zu typisch menschlichen Schlußfolgerungen wie *ein wenig*, *fast* oder *jung*. Diese Art der Logik wird deshalb als *two-valued logic* bezeichnet.

Eine Möglichkeit, unsichere Daten Rechnern zugänglich zu machen, besteht darin zu einer Logik überzugehen, die in der Lage ist, qualitative Aussagen zu verknüpfen. Ein Beispiel einer solchen *multivalued logic* ist die von Lofti A. Zadeh 1965 begründete *Fuzzy-Logik* [Zadeh, 1965], [Zadeh, 1988].

6.1 Meß- und Regelgrößen der Bewegungssegmentierung

Wie in Abschnitt 5.6.3 gezeigt wurde, muß zur optimalen Segmentierung eine *lokale Schwelle* $T(\vec{x})$ gefunden werden, die sich dem Bildinhalt anpaßt. Um unabhängig von speziellen Objektfor-

men zu sein, werden die Überlegungen zur Größe der Schwelle auf statistische Eigenschaften der Grauwertverteilungen korrespondierender Umgebungen in Originalbild und Hintergrund verallgemeinert. Die minimale Schwelle wird immer noch durch die Varianz des Rauschens σ_n^2 bestimmt (Anhang B). Der größte Wert von T hängt von der Größe der zu erwartenden Grauwertänderungen ab, die durch andere Effekte als Bewegung verursacht werden, z. B. Beleuchtungsänderungen oder hardwareabhängige Digitalisierungsfehler. Die lokale Schwelle bewegt sich zwischen diesen Grenzen, um schnelle, dunkle Teilchen zu detektieren, während an anderen Stellen Beleuchtungsänderungen ignoriert werden.

Um dies zu realisieren, werden *statistische Eigenschaften* korrespondierender Umgebungen in Originalbild und Hintergrund durch Fuzzy-Logik verknüpft. Dies ist möglich, da die Texturen der Objekte am meisten zu den zeitlichen Grauwertänderungen beitragen (Abschnitt 3.5). Somit gibt die Grauwertstatistik wichtige Hinweise auf Bewegung. Die beiden einfachsten statistischen Eigenschaften sind die *lokale Varianz* und der *lokale, mittlere Grauwert* der betrachteten Grauwertstichproben. Mit höheren Momenten der Histogramme kann die exakte Grauwertstatistik iteriert werden.

Hsu *et al.* [Hsu et al., 1984] verwenden die gleichen statistischen Eigenschaften, um direkt modellierte Grauwertverteilungen in zwei aufeinanderfolgenden Bildern zu vergleichen. Da sie jedoch einzig Abweichungen von Polynomen erster oder zweiter Ordnung berücksichtigen, kann dies nur eine Approximation realer Bilder darstellen. In der vorliegenden Arbeit werden diese Differenzen der lokalen Grauwertstatistik zwischen Originalbild und Hintergrund benutzt, um das vorher berechnete *Ähnlichkeitsmaß* S zu bewerten. Die Meßgrößen sind somit die *Differenz der lokalen Varianz* (VD) und die *Differenz des lokalen, mittleren Grauwertes* (GD). Diese Werte werden für jedes Pixel berechnet und zu der Stellgröße *lokale Schwelle* $T(\vec{x})$ verknüpft.

6.1.1 Berechnung der Differenz des lokalen mittleren Grauwertes (GD)

Der lokale mittlere Grauwert $\langle G \rangle_{m,n}$ der Umgebung eines Pixels am Punkt (m, n) ist definiert durch

$$\langle G \rangle_{m,n} = \frac{1}{P} \sum_{k,l \in B} G_{m-k,n-l}, \quad (6.1)$$

wobei die Summe über alle P Punkte der lokalen Umgebung B läuft. Dies entspricht der Faltung des Bildes mit einer Rechteckmaske. Besser zur Glättung geeignet ist eine Binomialmaske \mathcal{B}_x^n , deren effektive Fläche der Größe der lokalen Umgebung entspricht, über die gemittelt wird. Die Größe GD berechnet sich damit zu

$$GD = \langle G \rangle - \langle H \rangle = \mathcal{B}_x^n G - \mathcal{B}_x^n H = \mathcal{B}_x^n (G - H), \quad (6.2)$$

wobei G das Originalbild und H den ruhenden Hintergrund bezeichnet.

6.1.2 Berechnung der Differenz der lokalen Varianz (VD)

Die lokale Varianz $V_{m,n}$ der Umgebung eines Pixels am Punkt (m, n) ist definiert durch

$$V_{m,n} = \frac{1}{P-1} \sum_{k,l \in B} (G_{m-k,n-l} - \langle G \rangle_{m,n})^2 = \langle (G - \langle G \rangle)^2 \rangle_{m,n}, \quad (6.3)$$

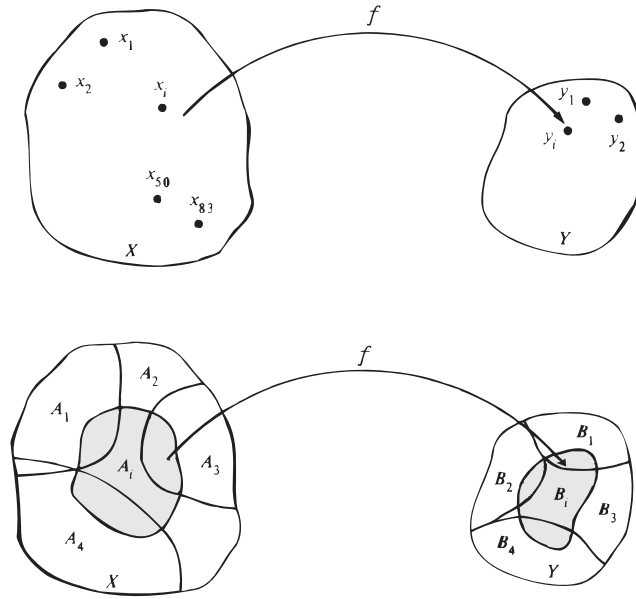


Abbildung 6.1: Die Funktion f bildet die Menge X der Meßgrößen auf die Menge Y der Stellgrößen ab. In der oberen Illustration werden verschiedene numerische Wertepaare (x_i, y_i) zur Schätzung von $f : X \rightarrow Y$ verwendet. Im zweiten Fall wird nur eine kleine Zahl von Fuzzy-Mengen A_i aus X und B_i aus Y verwendet. Die Fuzzy-Assoziation (A_i, B_i) repräsentiert die System-Struktur [Kosko, 1992].

wobei die Summe wieder über alle P Punkte der lokalen Umgebung B läuft. Der mittlere Grauwert $\langle G \rangle_{m,n}$ berechnet sich aus der gleichen lokalen Umgebung. Gleichung (6.3) läßt sich weiter vereinfachen zu

$$\langle (G - \langle G \rangle)^2 \rangle_{m,n} = \langle G^2 - 2G\langle G \rangle + \langle G \rangle^2 \rangle_{m,n} = \langle G^2 \rangle_{m,n} - \langle G \rangle_{m,n}^2. \quad (6.4)$$

Ersetzt man analog zu 6.1.1 die Mittelung $\langle \cdot \rangle$ durch eine Glättung mit einer Binomialmaske, dann ergibt sich die lokale Varianz zu

$$V_{m,n} = \mathcal{B}_x^n G^2 - (\mathcal{B}_x^n G)^2 \quad (6.5)$$

und damit VD :

$$VD = V_{m,n}(G) - V_{m,n}(H) = \mathcal{B}_x^n G^2 - (\mathcal{B}_x^n G)^2 - \mathcal{B}_x^n H^2 + (\mathcal{B}_x^n H)^2. \quad (6.6)$$

6.2 Fuzzy-Logik

Die Ansätze zur Beschreibung von Fuzzy-Logik reichen von geometrischen Interpretationen im Wahrscheinlichkeitsraum bis hin zu philosophischen Betrachtungen über die Äquivalenz von Wahrscheinlichkeit und Unsicherheit bzw. Unterschiede zwischen diesen Begriffen. Im Hinblick auf ihre Anwendung in dieser Arbeit soll die Fuzzy-Logik als eine Möglichkeit verstanden werden, den Verlauf nichtlinearer Systeme durch Verknüpfen von Teilmengen der Definitionsmenge und Wertemenge zu schätzen [Kosko, 1992]. In dieser Eigenschaft sind Fuzzy-Mengen äquivalent

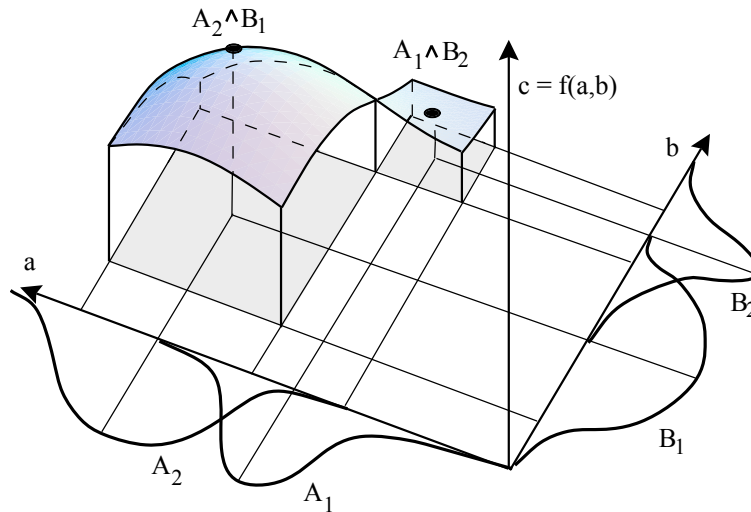


Abbildung 6.2: Schätzung einer Funktion $f(a,b)$ durch Verknüpfung des logischen Zusammenhangs von Teilmengen der Definitionsmenge zu Funktionswerten: Die Grauwertflächen im Raum entsprechen der Interpolation von Funktionswerten durch Verknüpfung der linguistischen Variablen A_1, A_2, B_1 und B_2 . Falls die linguistischen Variablen zu δ -Funktionen entarten, kann für jede Fläche nur der eingezeichnete Punkt bestimmt werden. Um den gesamten Funktionsverlauf zu schätzen, werden in diesem Fall unendlich viele Meßwerte benötigt.

zu Neuronalen Netzwerken. Beide Ansätze der künstlichen Intelligenz sind *modellunabhängige Schätzer*. Sie approximieren eine Funktion ohne eine mathematische Beschreibung des Zusammenhangs zwischen dem Output der Funktion und den Input-Parametern zu benötigen. Sie *lernen aus Beispielen!*

Neuronale Netze und Fuzzy-Systeme unterscheiden sich darin, wie sie die stückweise bekannten Funktionen schätzen (Abb. 6.1). Sie unterscheiden sich in der Art der Proben, die sie benutzen, wie sie diese repräsentieren und speichern und wie sie aus den Proben Schlußfolgerungen über den Zusammenhang zwischen Input und Output ziehen. Neuronale Netzwerke verknüpfen einzelne Meßwerte mit einzelnen Funktionswerten (x_i, y_i) . Um aus diesen einzelnen Punkten einen kontinuierlichen Verlauf der gesuchten Funktion zu bestimmen, benötigen sie eine große Zahl von Testdaten. Fuzzy-Systeme schätzen Funktionen über Kombinationen von *Fuzzy-Mengen* (A_i, B_i) . Dabei werden *Teilmengen der Input Parameter* mit *Teilmengen der Output-Parameter* verknüpft. Da sich diese Mengen in beiden Räumen überlappen, erhält man direkt einen *interpolierten Funktionsverlauf* (Abb. 6.2). In der Sprache der künstlichen Intelligenz wird ein solches System, das Mengen auf Mengen abbildet und dabei benachbarte Input-Daten auf benachbarte Output-Daten abbildet, als *assoziatives Gedächtnis* (engl. associative memory) bezeichnet. Fuzzy-Systeme werden daher im allgemeinen als *fuzzy associative memories* oder kurz FAMs bezeichnet. Während bei Neuronalen Netzwerken die Testpunkte möglichst dicht gewählt werden müssen, um eine glatte und vollständige Funktion zu schätzen, kann durch Fuzzy-Logik mit nur wenigen Verknüpfungen der komplette Funktionsverlauf beschrieben werden. Die Verknüpfungen erfolgen in einer Form, die der menschlichen Logik nachempfunden ist. Dabei werden Regeln der

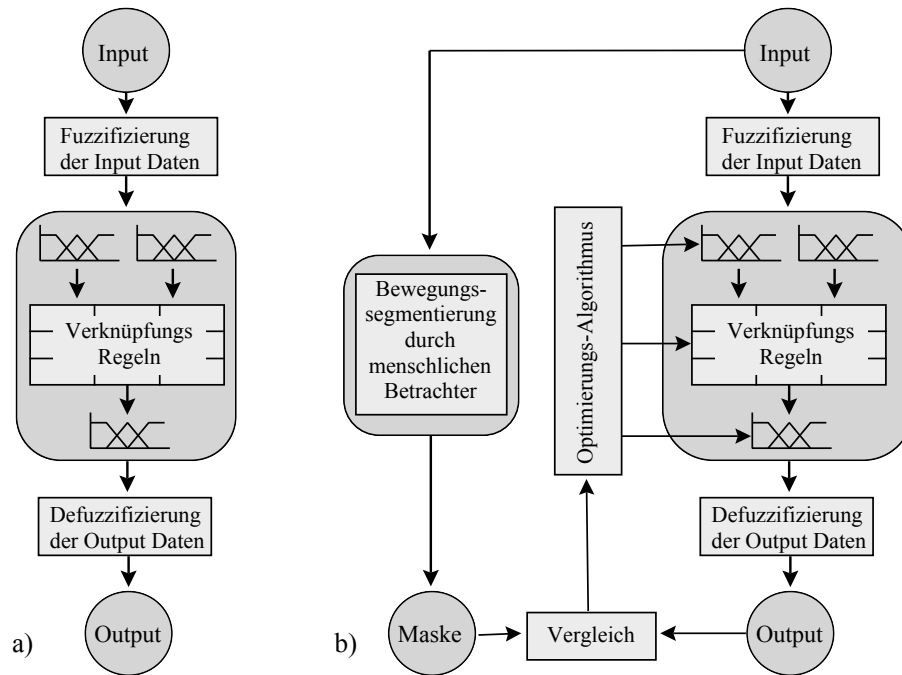


Abbildung 6.3: Schematische Darstellung der Abläufe der Fuzzy-Logik. a) Konventionelle Fuzzy-Logik: die Input-Daten werden nach festen Regeln zu Output-Daten verknüpft. b) Adaptive Fuzzy-Logik: durch Rückkopplung des Resultates auf die Form der linguistischen Variablen und die Verknüpfungsregeln kann das System optimiert werden.

Form „WENN Bedingung, DANN Konsequenz“ aufgestellt, wobei sowohl Bedingung als auch Konsequenz linguistische Variablen der Form *warm*, *klein*, *jung* usw. darstellen.

Die Verknüpfung von Eingangsdaten zu Ausgangsdaten durch Fuzzy-Logik läßt sich in 3 Schritte untergliedern (Abb. 6.3a):

- *Fuzzifizierung der Input Variablen (fuzzification)*

Die Meßgröße wird durch einen Satz von Zugehörigkeitsfunktionen beschrieben. Diese Funktionen werden als *linguistische Variablen* bezeichnet und stellen eine mathematische Realisierung von beschreibenden Eigenschaften der Größe dar (z. B. *klein*, *jung*, *viel*, *etc.*). Der numerische Wert der Meßgröße wird dabei in einen Grad der Zugehörigkeit zu den verschiedenen linguistischen Variablen umgewandelt (Abschnitt 6.2.1).

- *Logische Verknüpfung (execution of rules)*

Die linguistischen Variablen der Input-Variablen werden mit den linguistischen Variablen der Output-Variablen verknüpft. Im Vergleich zur Booleschen Algebra müssen hierbei die Operatoren der logischen Verknüpfung abgeändert werden (Abschnitt 6.2.2).

- *Defuzzifizierung der Output Variablen (defuzzification)*

Nach der Verknüpfung der Input-Parameter liegt das Ergebnis in Form von Zugehörigkeiten zu verschiedenen linguistischen Variablen der Output-Parameter vor. Diese müssen zu einem numerischen Output-Wert verknüpft werden (Abschnitt 6.2.3).

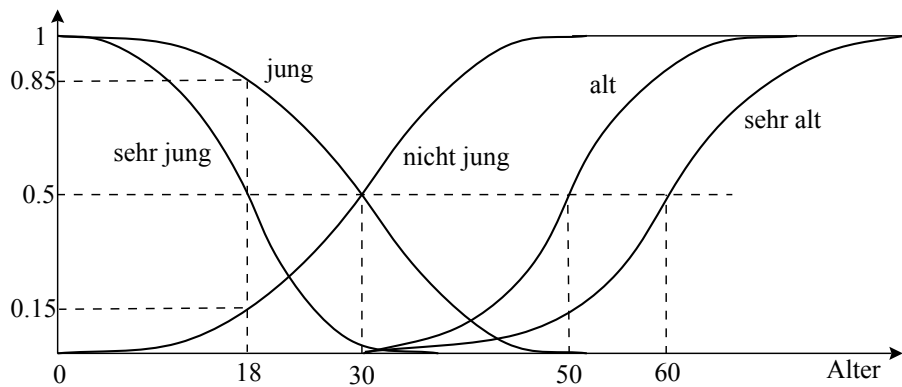


Abbildung 6.4: Linguistische Variablen für ‚Alter‘ [Zadeh, 1988].

6.2.1 Fuzzifizierung

Sei $x \in X$ eine kontinuierliche oder diskrete Meßgröße. Der Menge X wird nun eine Zahl N von *linguistischen Variablen* A_n , $0 \leq n \leq N$ zugeordnet. Jede dieser linguistischen Variablen definiert eine Zugehörigkeitsfunktion m_{A_n} , welche die Elemente x aus X auf den Grad ihrer Zugehörigkeit zu A_n abbildet:

$$m_{A_n} : X \longrightarrow [0, 1]. \quad (6.7)$$

In Abb. 6.4 sind die Zugehörigkeitsfunktionen möglicher linguistischer Variablen für die Größe „menschliches Alter“ abgebildet. In diesem Fall wäre ein 18-jähriger Mensch zu 50 % *sehr jung*, zu 85 % *jung* und damit zu 15 % *nicht jung*. *Nicht jung* stellt das Komplement der Fuzzy-Menge *jung* dar.

Jeder Input Variablen (Meßgröße) und der Output Variablen (Stellgröße) wird ein Satz solcher linguistischer Variablen zugeordnet. Eine Input Variable x wird fuzzifiziert, indem der Grad ihrer Zugehörigkeit für alle N linguistischen Variablen A_n bestimmt wird, d. h. $m_{A_n}(x) \forall n$. Der Output der Fuzzifizierung ist dann ein $N \times 1$ Vektor $\vec{v}_A = (a_0, \dots, a_{N-1})$. Die Elemente des Vektors sind die Zugehörigkeitswerte von x zu jeder einzelnen linguistischen Variablen. Abb. 6.5 zeigt die verwendete Form der linguistischen Variablen für die Größen GD und VD und T .

Die Menge der linguistischen Variablen ist zunächst beliebig und muß dem jeweiligen Problem angepaßt werden. Die Form der Zugehörigkeitsfunktionen bestimmt, wie zwischen den Werten x_n der maximalen Zugehörigkeit zu einer linguistischen Variablen A_n interpoliert wird. Für $m_{A_n}(x) = \delta(x - x_n)$ kann die zu schätzende Funktion nur für einzelne Punkte bestimmt werden (Punkte in Abb. 6.2).

Obwohl die Form der Zugehörigkeitsfunktionen einer Wahrscheinlichkeitsverteilung ähnelt und in der Literatur auch oft als Wahrscheinlichkeit der Zugehörigkeit bezeichnet wird, ist Fuzzy-Logik von der Wahrscheinlichkeitsrechnung zu unterscheiden. Fuzzy beschreibt die *Mehrdeutigkeit eines Ereignisses*. Sie beschreibt den Grad, mit dem ein Ereignis eintritt, nicht ob es eintritt. Wahrscheinlichkeit beschreibt die *Unsicherheit, mit der ein Ereignis eintritt*. Im Gegensatz zur Wahrscheinlichkeitstheorie ist Fuzzy-Logik in der Lage mit *Mehrdeutigkeiten* umzugehen.

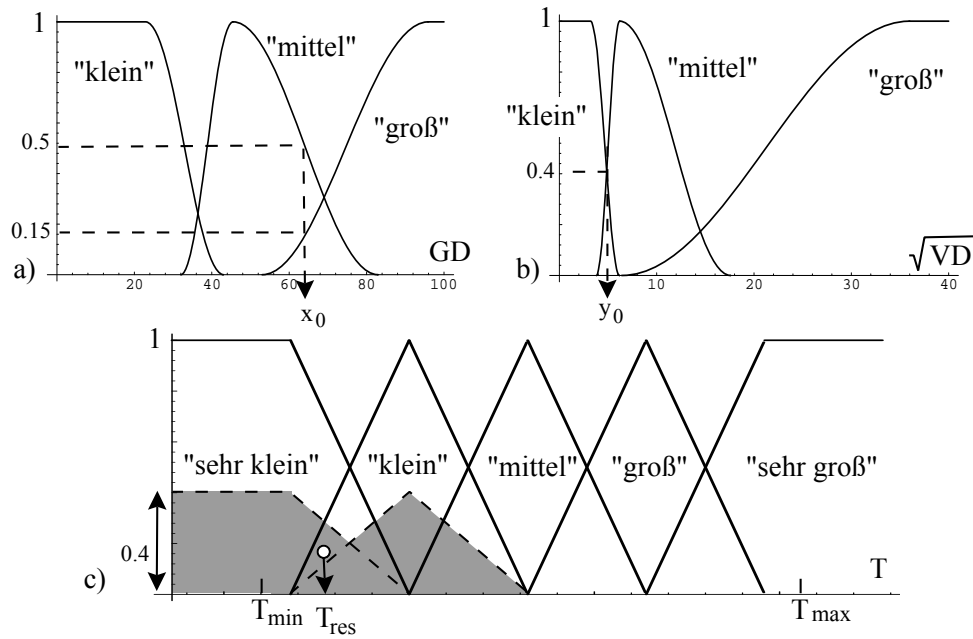


Abbildung 6.5: Verwendete optimierte linguistische Variablen für GD , VD und T . Die grau unterlegte Fläche zeigt die normierten linguistischen Variablen, die für die eingezeichneten Input-Werte x_0 und y_0 das Ergebnis T_{res} liefern (Abschnitt 6.2.3).

6.2.2 Logische Verknüpfung

Die linguistischen Variablen A_n ($0 \leq n \leq N$) und B_l ($0 \leq l \leq L$) der Input-Größen werden mit den linguistischen Variablen C_m ($0 \leq m \leq M$) der Output-Größen verknüpft. Jeder möglichen Kombination der A_n und B_l wird ein Element C_m zugeordnet, wobei ein C_m bei mehreren Kombinationen auftreten kann. Die Menge der Verknüpfungsregeln des FAM kann durch eine *logische Verknüpfungstabelle* oder *linguistische Matrix* dargestellt werden. Tabelle 6.1 zeigt die linguistische Matrix, welche in der vorliegenden Arbeit zur Verknüpfung von GD und VD zur lokalen Schwelle T verwendet wurde. Sie stellt bereits das Ergebnis der Optimierung dar, die in Abschnitt 6.2.5 beschrieben wird. Die Werte *sehr klein*, *klein*, *...*, *sehr groß* bezeichnen die linguistischen Variablen der lokalen Schwelle T , die in Abb. 6.5 dargestellt sind.

| \wedge | | GD | | |
|----------|--------|-----------|------------|------------|
| | | klein | mittel | groß |
| VD | klein | sehr groß | klein | sehr klein |
| | mittel | mittel | sehr klein | sehr klein |
| | groß | mittel | sehr klein | sehr klein |

Tabelle 6.1: Logische Verknüpfungstabelle der FAM Regeln zur Bestimmung der lokalen Schwelle (optimiert).

Die Verknüpfung geschieht dabei durch eine logische AND oder eine logische OR Operation.

Zur Verknüpfung von Fuzzy-Mengen werden die logischen Operationen der Booleschen Algebra durch Fuzzy-Operationen ersetzt:

$$AND: (a_n \wedge b_l) \longrightarrow \min\{a_n, b_l\}, \quad (6.8)$$

$$OR: (a_n \vee b_l) \longrightarrow \max\{a_n, b_l\}. \quad (6.9)$$

Für jede Kombination der Input-Variablen liefert die AND-Verknüpfung (6.8) ein Maß für das Eintreten dieser Kombination:

$$c_m = a_n \wedge b_l. \quad (6.10)$$

Da mehrere Kombinationen (n, l) zu der gleichen linguistischen Variablen C_m verknüpft werden können (z. B. $2 \times \text{mittel}$ und $5 \times \text{sehr klein}$ in Tab. 6.1), werden diese Werte als Endergebnis durch den OR-Operator (6.9) verknüpft:

$$c_m = \bigvee_{n,l \in \mathcal{M}} (a_n \vee b_l), \quad \mathcal{M} = \{(n, l) : A_n \wedge B_l = C_m\}. \quad (6.11)$$

Das Ergebnis der Verknüpfung ist ein Vektor $\vec{v}_C = (c_o, \dots, c_{M-1})$, dessen Elemente einen numerischen Wert für die Anteile der einzelnen linguistischen Variablen C_m am Gesamtergebnis darstellen, der sog. *Fit-Vektor*.

Beispiel:

In Abb. 6.5 ist dargestellt, wie für zwei Werte x_o und y_o der beiden Input-Parameter die Fit-Vektoren berechnet werden. Für x_o erhält man die Zugehörigkeitswerte $a_{\text{mittel}} = 0.5$ und $a_{\text{groß}} = 0.15$. Die Werte für y_o sind $b_{\text{klein}} = 0.4$ und $b_{\text{mittel}} = 0.4$. Die Verknüpfungstabelle (Tab.6.1) liefert

$$\begin{aligned} c_{\text{klein}} &= (a_{\text{mittel}} \wedge b_{\text{klein}}) = 0.5 \wedge 0.4 = 0.4 \\ c_{\text{sehr klein}} &= (a_{\text{groß}} \wedge b_{\text{klein}}) \vee (a_{\text{mittel}} \wedge b_{\text{mittel}}) \vee (a_{\text{groß}} \wedge b_{\text{mittel}}) \\ &= (0.15 \wedge 0.4) \vee (0.5 \wedge 0.4) \vee (0.15 \wedge 0.4) = 0.15 \vee 0.4 \vee 0.15 = 0.4 \end{aligned}$$

und damit den Fit-Vektor $\vec{v}_C = (c_{\text{sehr klein}}, c_{\text{klein}}, \dots, c_{\text{sehr groß}}) = (0.4, 0.4, 0, \dots, 0)$. Die linguistischen Variablen werden auf diesen Wert normiert (Abb. 6.5b).

6.2.3 Defuzzifizierung

Aus dem Fit-Vektor $\vec{v}_C = (c_o, \dots, c_{M-1})$ muß nun der endgültige Output-Wert y_{out} aus der Menge Y der möglichen Output-Werte errechnet werden. Das Fuzzy-System aktiviert jede linguistische Variable C_m zu einem Grad c_m und addiert diese zu einer kombinierten Fuzzy-Menge C_r

$$C_r = \sum_{i=0}^{M-1} C_i, \quad (6.12)$$

oder äquivalent dazu:

$$m_{C_r} = \sum_{i=0}^{M-1} c_i m_{C_i}. \quad (6.13)$$

Die Funktion m_{C_r} stellt somit die Summe der einzelnen Zugehörigkeitsfunktionen der linguistischen Output-Variablen dar. Aus dieser resultierenden Funktion läßt sich auf verschiedene Arten ein geeigneter numerischer Wert errechnen. Die einfachste Defuzzifizierung wird durch das Maximum von m_{C_r} realisiert:

$$y_{out} = y_{max} : \quad m_{C_r}(y_{max}) = \max_{y \in Y} m_{C_r}(y). \quad (6.14)$$

Dies wird als *maximum-membership defuzzification* bezeichnet und wird durch die populären Methoden der *maximum-likelihood*- und *maximum-a-posteriori*-Schätzer der Wahrscheinlichkeitsrechnung motiviert. Die maximum-membership defuzzification hat jedoch zwei entscheidende Nachteile. Die Funktion m_{C_r} ist in vielen Fällen bimodal und oft über einen weiten Bereich konstant. In diesen Fällen kann daher kein eindeutiger Wert y_{max} gefunden werden. Benutzt man einen großen Satz von linguistischen Variablen und sehr viele FAM-Regeln, so gilt eine Erweiterung des zentralen Grenzwertsatzes. Die Form der Funktion m_{C_r} ähnelt dann einer Gaußkurve und wird uni-modal. Dies widerspricht jedoch der Absicht der Fuzzy-Logik mit einer kleinen Anzahl von Verknüpfungen auszukommen. Außerdem berücksichtigt die maximum-membership defuzzification nicht die Information, die in der Form der Funktion m_{C_r} steckt. Da diese Funktion oft stark asymmetrisch ist, geht damit eine wichtige Information zur Interpolation verloren.

Die natürliche Alternative ist das *fuzzy centroid defuzzification* Schema. Dabei wird der Schwerpunkt der Fläche unter der Funktion berechnet:

$$y_{out} = \frac{\int_{y \in Y} y m_{C_r}(y) dy}{\int_{y \in Y} m_{C_r}(y) dy}. \quad (6.15)$$

In der praktischen Anwendung wird die Funktion m_{C_r} durch die Einhüllende \overline{m}_{C_r} der Funktionen m_{C_i} in Gleichung (6.13) ersetzt:

$$\overline{m}_{C_r}(y) = \max_{0 \leq m \leq (M-1)} m_{C_m}(y). \quad (6.16)$$

Der Schwerpunkt der beiden Funktionen weicht nur geringfügig voneinander ab. Die Funktion \overline{m}_{C_r} und der daraus erhaltene Schwerpunkt für das betrachtete Beispiel sind in Abb. 6.5 durch die grau unterlegte Fläche und den weißen Kreis gekennzeichnet.

6.2.4 Fuzzy-LUT

Da alle möglichen Werte von GD und VD für ein Bild von vornherein bekannt sind, ist es nicht notwendig, für jeden Bildpunkt die gesamte Berechnung der Fuzzy-Logik durchzuführen. Es genügt, alle Kombinationen ein mal zu berechnen und die zugehörigen Werte der lokalen Schwelle in einer Look-Up-Tabelle (LUT) zu speichern. Für die beiden Meßwerte GD und VD ist diese LUT zweidimensional, d. h. ein Bild, dessen Grauwert direkt die lokale Schwelle angibt. In der praktischen Anwendung reduziert sich die Berechnung damit auf das Auslesen eines Bildpunktes, dessen Koordinaten durch den jeweiligen Wert von GD und VD gegeben sind. Abbildung 6.6 zeigt die mit Tabelle 6.1 und den linguistischen Variablen aus Abb. 6.5 berechnete Fuzzy-LUT. Sie stellt ein Bild der gesuchten Funktion f dar und wird als *control surface* des Fuzzy-Controllers bezeichnet.

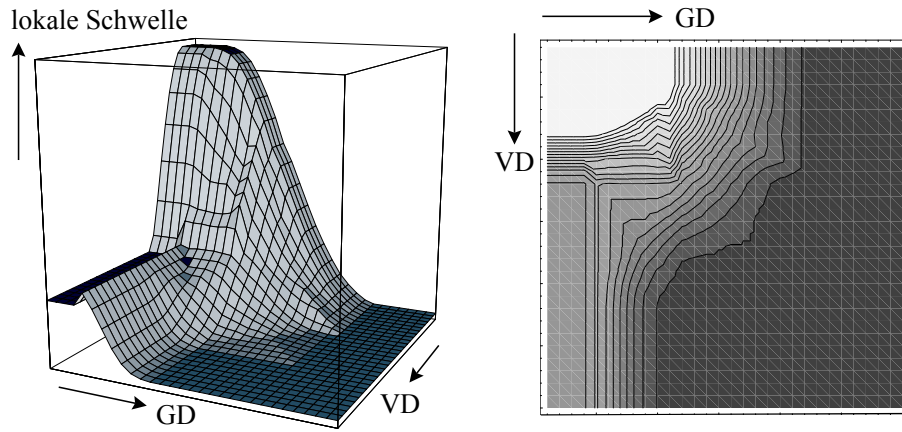


Abbildung 6.6: Look-Up-Tabelle der lokalen Schwelle in Abhängigkeit von GD und VD ($\hat{=}$ Control surface der Fuzzy-Logik).

6.2.5 Optimierung der Fuzzy-Logik

Die Form der linguistischen Variablen und die Verknüpfungsregeln werden zunächst empirisch durch Plausibilitätsüberlegungen gefunden (Abschnitt 6.2.2). Um die Parameter der Logik der Problemstellung besser anzupassen, wird *adaptive Fuzzy-Logik* eingesetzt [Cox, 1993], [Athalye et al., 1993]. Diese ist in der Lage, die Form der linguistischen Variablen und die Verknüpfungsregeln rekursiv zu ändern, um ein gewünschtes Resultat zu erreichen. Das gewünschte Ergebnis entspricht im vorliegenden Fall dem Segmentierungsergebnis eines menschlichen Beobachters. Abbildung 6.3b zeigt das modifizierte Ablaufdiagramm der adaptiven Fuzzy-Logik. Um das System zu trainieren wird das Bild jedoch parallel von einem menschlichen Beobachter segmentiert. Die beiden Ergebnisse werden verglichen und die Parameter der Logik verändert.

Vergleichsoperation

Das Segmentierungsergebnis des menschlichen Beobachters wird als Maske M benutzt. Alle Pixel, an denen Bewegung stattfindet, werden mit dem Wert 1 markiert und alle anderen Bereiche mit dem Wert -1 . Im Segmentierungsergebnis R der Fuzzy-Logik erhalten alle segmentierten Pixel den Wert 1 und alle Hintergrundpixel den Wert 0. Danach werden die beiden Bilder punktweise miteinander multipliziert und die Werte aller Pixel aufaddiert. Das Ergebnis dieser Verknüpfung ist eine skalare Größe s , deren Wert direkt die Übereinstimmung zwischen beiden Bildern angibt:

$$s = \sum_{(m,n) \in \text{Bild}} M_{m,n} \cdot R_{m,n}. \quad (6.17)$$

Für ideale Übereinstimmung erhält man als Wert die Summe der Pixel, die vom menschlichen Beobachter als Bewegung interpretiert werden. Bei maximal schlechter Übereinstimmung ergibt sich die Summe des Komplementes dieser Pixel mit negativem Vorzeichen. Zur Optimierung muß daher das Maximum dieser Funktion im Raum der Parameter der Fuzzy-Logik gefunden werden. Da der Parameterraum bei der verwendeten Logik die Dimension 23 hat (14 Formparameter der linguistischen Input-Variablen und 9 Verknüpfungen), ist dieses Maximum nicht direkt zu finden. Dies stellt ein allgemeines mathematisches *Optimierungsproblem* dar. Zur Lösung wurde

eine Vereinfachung des *Sintflutalgorithmus* [Dueck et al., 1993] verwendet. Die Parameter werden zufällig verändert. Führt die Veränderung der Parameter zu einer Verbesserung des Ergebnisses, dann werden die neuen Parameter verwendet, ansonsten die alten. Das gesamte FAM wird somit schrittweise verbessert, bis ein bestimmter Grad der Übereinstimmung erreicht wird.

Kapitel 7

Zusammenfassung und Ergebnisse der Bildverarbeitung

Nachdem in den vorangegangenen Kapiteln die einzelnen Schritte der Bildfolgenanalyse vorgestellt wurden, soll nun ein zusammenfassender Überblick über das gesamte Verfahren gegeben und die Ergebnisse anhand von Beispielbildern diskutiert werden.

7.1 Zusammenfassung der Bildfolgenanalyse

In Abbildung 7.1 ist der Ablauf des gesamten Verfahrens dargestellt. Aus einer Bildfolge wird zunächst durch eine zeitliche Faltung mit einer Gaußmaske der ruhende Hintergrund der Szene extrahiert. Eine recheneffektive Realisierung wird durch eine zeitliche Gaußpyramide erreicht. Ein Vergleich zwischen ruhendem Hintergrund und Originalbild liefert die Bildregionen, in denen mit großer Sicherheit Bewegung stattfindet. Dazu wird ein lokales Ähnlichkeitsmaß berechnet und dieses durch adaptive Fuzzy-Logik bewertet. Die Berechnungen der Fuzzy-Logik können durch Kenntnis der möglichen Meßgrößen vorab berechnet und in einer Look-Up-Tabelle gespeichert werden.

Insgesamt wird durch das Verfahren eine lokale Nachbarschaft des 3-dimensionalen Orts-Zeit-Raumes der Eigenschaft *Bewegung* zugeordnet:

$$f(V) \longrightarrow \{0, 1\}, \quad V = \{(x, y, t) : -x_u \leq x \leq x_u, -y_v \leq y \leq y_v, -t_w \leq t \leq t_w\}, \quad (7.1)$$

wobei V die lokale Umgebung und u, v, w die Ausdehnung dieser Umgebung in Richtung der verschiedenen Koordinaten bezeichnet. Einem Pixel, an dem Bewegung stattfindet, wird der Wert 1 zugeordnet und einem Hintergrund-Pixel der Wert 0. Da sich eine direkte analytische Beziehung für f nicht direkt finden läßt, werden aus V andere Merkmale berechnet, die Informationen über die gewünschte Eigenschaft enthalten:

Reduktion des Grauwertverlaufes auf lokale Eigenschaften

Zeitliche Information liefert die Berechnung des Hintergrunds. Die drei Merkmale: Ähnlichkeitsmaß S , lokale Grauwertdifferenz GD und lokale Varianzdifferenz VD enthalten Differenzen

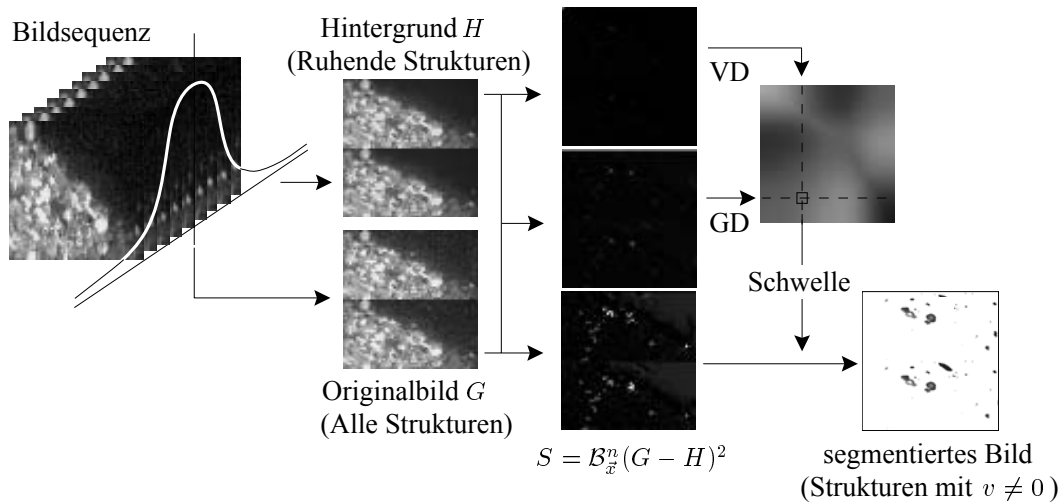


Abbildung 7.1: Ablaufdiagramm des Verfahrens der Bewegungssegmentierung.

zwischen räumlichen Eigenschaften von Original- und Hintergrundbild. Dadurch wird formal der dreidimensionale $\bar{x}t$ -Raum auf den dreidimensionalen *Merkmalsraum* reduziert, wobei die Zeitabhängigkeit eliminiert wird. Aus (7.1) wird die Funktion g :

$$g(S, GD, VD) \longrightarrow \{0, 1\}. \quad (7.2)$$

Durch diese Transformation von f auf g wird die exakte Verknüpfung einzelner Grauwerte auf die Verknüpfung von Eigenschaften dieser raum-zeitlichen Grauwertmuster reduziert.

Bewertung der Eigenschaften

Die Bestimmung einer lokalen Schwelle zur Bewertung des Ähnlichkeitsmaßes dient der effizienten Berechnung von (7.2). Die Schwelle stellt eine skalare Funktion in Abhängigkeit von zwei Parametern dar. Die Verknüpfung von S und T reduziert sich damit auf

$$h(S - T(GD, VD)) \longrightarrow \{0, 1\}, \quad \text{mit } h(S - T) = \Theta(S - T). \quad (7.3)$$

Die Berechnung von T durch Fuzzy-Logik läßt es zu, die Segmentierung durchzuführen ohne eine analytische Funktion für $T(GD, VD)$ anzugeben.

7.2 Ergebnisse der Bildfolgenanalyse

In Abbildung 7.2 sind drei verschiedene Bildszenen mit ihren einzelnen Verarbeitungsschritten dargestellt. Abbildung 7.2a zeigt einen Querschnitt einer Sedimentoberfläche in einem Wassertank. Einige Sandteilchen rollen die Böschung hinab, während sich ein sehr schnelles Teilchen mit hoher Geschwindigkeit im freien Wasserkörper durch das Bild bewegt (weißer Pfeil). Durch die hohe Geschwindigkeit geht der Grauwert dieses Teilchens fast im aufgewirbelten Schlamm unter. Man erkennt im Ergebnisbild deutlich, daß diese Struktur segmentiert wird, obwohl der Grauwertkontrast kaum größer als der des Rauschens ist. Die Bedeutung der lokalen Schwelle bei diesem speziellen Beispiel kann in Grauwertschnitten weiter verdeutlicht werden. Abbildung 7.3a

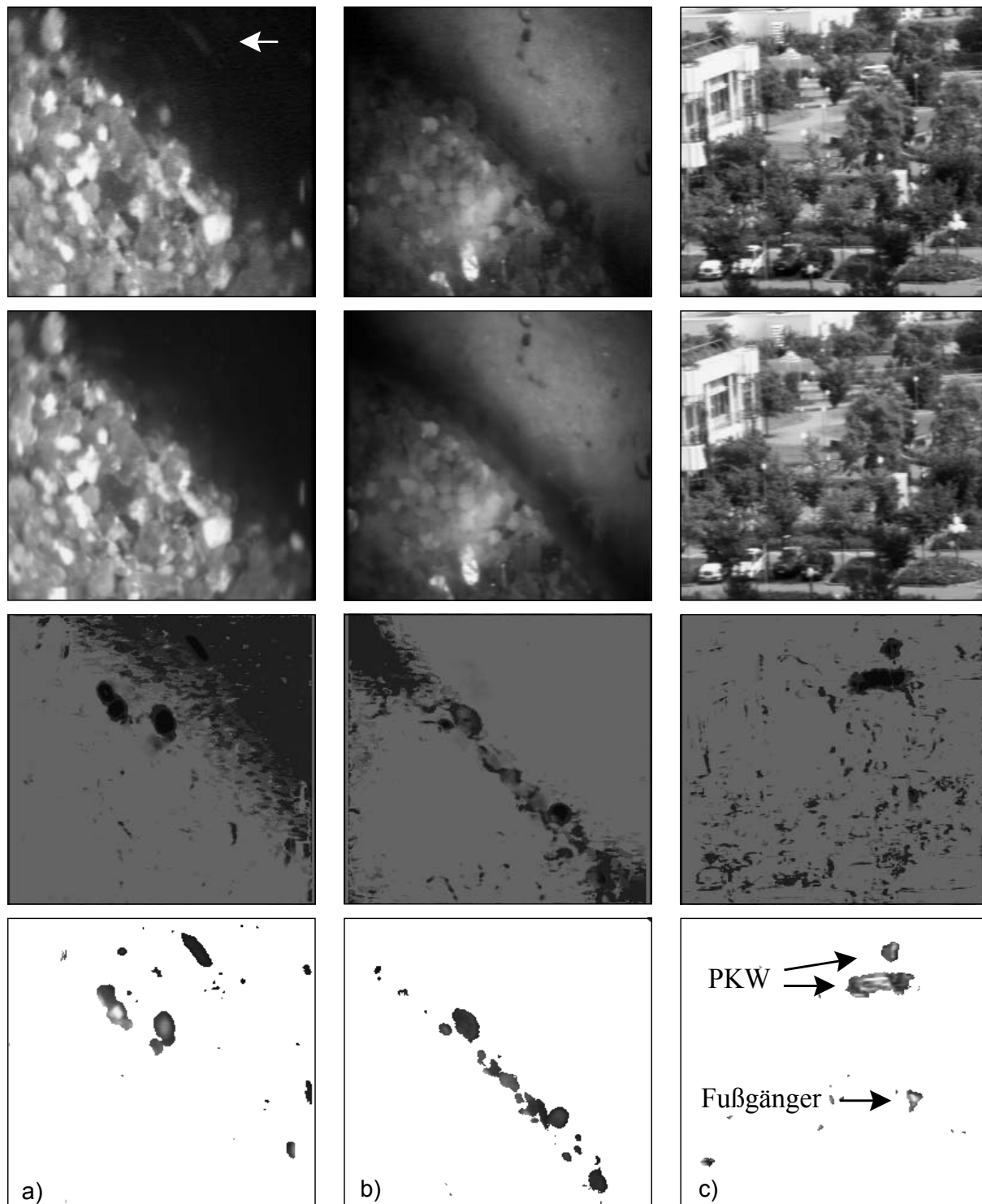


Abbildung 7.2: Beispiele zur Bewegungssegmentierung: Die drei Spalten zeigen verschiedene Beispielbilder. a) + b): Querschnitte einer Sandgrenzschicht. Der Bildausschnitt beträgt etwa 5 mm in horizontaler und vertikaler Richtung. c): Verkehrsszene. Von oben nach unten: Originalbild, ruhender Hintergrund, lokale Schwelle und Ergebnis der Segmentierung. Der Pfeil in Originalbild a) markiert den horizontalen Grauwertschnitt, der in Abb. 7.3 dargestellt ist.

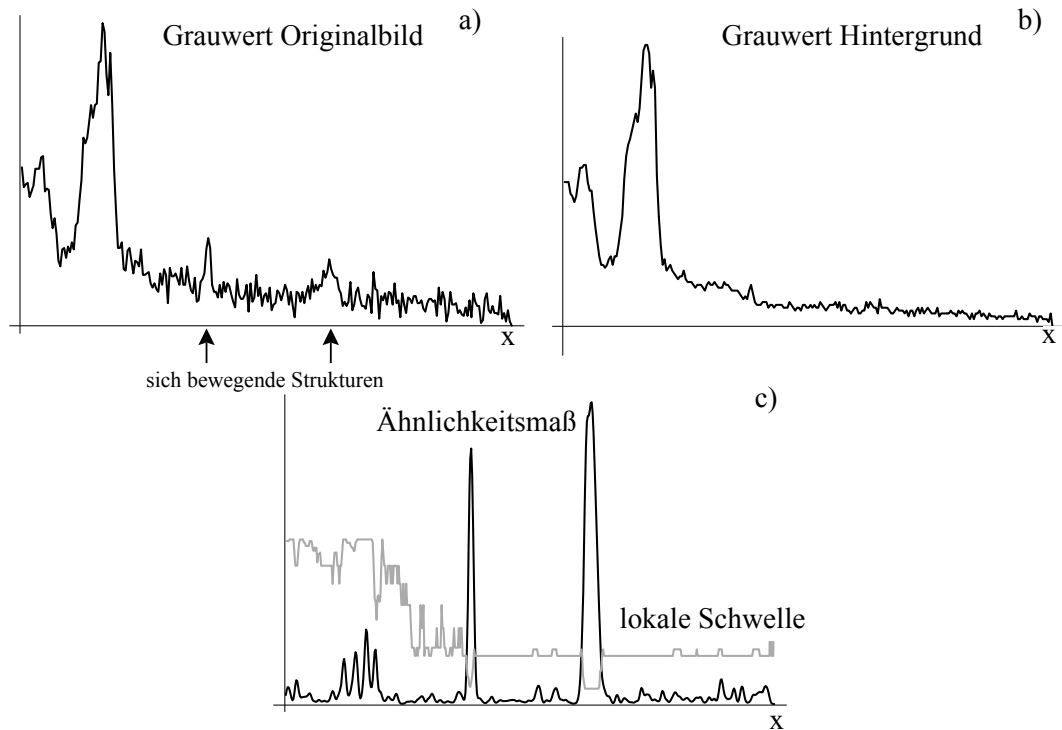


Abbildung 7.3: Grauwertschnitte, Ähnlichkeitsmaß und lokale Schwelle einer Zeile des Bildes in Abb. 7.2a.

zeigt einen horizontalen Grauwertschnitt durch das Originalbild in Abbildung 7.2a, an der durch den Pfeil markierten Stelle. Im Vergleich zu einem Schnitt durch den ruhenden Hintergrund an derselben Stelle 7.3b ist deutlich zu erkennen, daß das Rauschen durch die zeitliche Tiefpaßfilterung stark unterdrückt wird. Die beiden durch Pfeile gekennzeichneten lokalen Grauwertmaxima sind im Schnitt durch den Hintergrund, trotz des erheblich geringeren Rauschens nicht zu erkennen. Dies deutet darauf hin, daß es sich dabei um sich bewegende Strukturen handelt. Die lokale Schwelle weist an dieser Stelle zusätzlich ein deutliches Minimum auf (Abb. 7.3c). Dadurch wird die Detektion der schnellen Objekte über ihre gesamte Fläche erleichtert. Im Ergebnisbild werden an diesen Stellen auch deutlich zwei Teilchen detektiert (Abb. 7.2). An hellen Bildbereichen, die Grauwertänderungen ohne Bewegung aufweisen, bleibt die lokale Schwelle deutlich über dem Wert für dunkle Bildbereiche. Abbildung 7.4 zeigt das Segmentierungsergebnis der lokalen Schwelle im Vergleich zu einer globalen Schwelle. Eine etwas zu kleine globale Schwelle (Abb. 7.4a) führt zu Fehlsegmentierungen in den Bildbereichen unterhalb der Sandoberfläche. Außerdem beginnen die Gebiete direkt über der Oberfläche, die aufgewirbelten Schlamm enthalten, zu großen Flächen zu verschmelzen. Bei Verwendung einer lokalen Schwelle (Abb. 7.4b) wird dies verhindert. Die Teilchen an der Oberfläche und das schnelle Teilchen im freien Wasserkörper werden weiterhin vollständig segmentiert, während Fehlsegmentierungen weitgehend unterbunden werden. Eine etwas zu große Schwelle (Abb. 7.4c) weist genau wie die lokale Schwelle keine Fehlsegmentierungen auf. Die Teilchen an der Oberfläche selbst werden auch hier hinreichend gut segmentiert. Die schnellen Teilchen im freien Wasserkörper werden jedoch nur über einem Bruchteil ihrer Fläche detektiert. Eine Betrachtung des Ähnlichkeitsmaßes in Abb. 7.3c läßt vermuten, daß sich eine passende globale Schwelle trotzdem finden läßt. Dies kann in speziell-

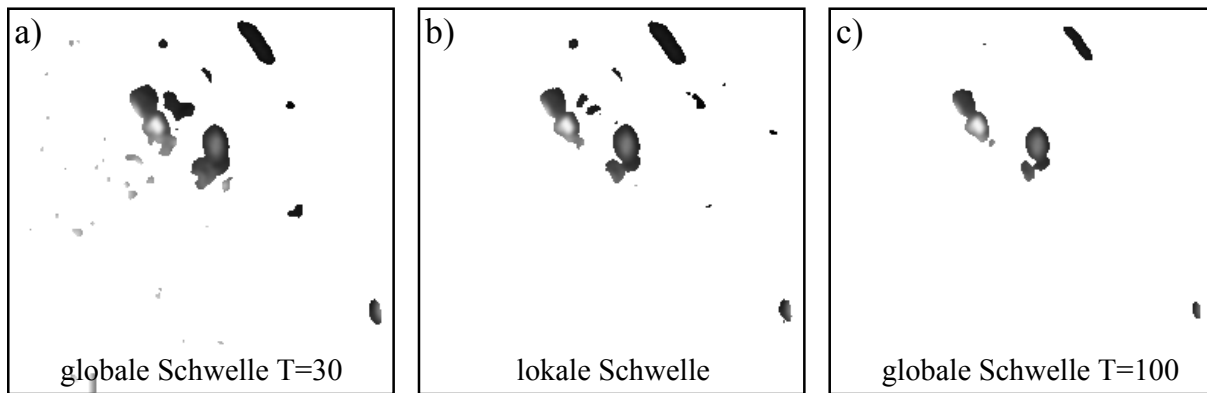


Abbildung 7.4: Ergebnis der Bewegungssegmentierung des Bildes in Abb. 7.2a zur Verdeutlichung der Wirkung der lokalen Schwelle: a): Segmentierung des Ähnlichkeitsmaßes S (Abb. refschnittc) durch eine globale Schwelle $T = 30$. b): Segmentierung des Ähnlichkeitsmaßes S durch eine lokale Schwelle (Abb. 7.2a). c): Segmentierung des Ähnlichkeitsmaßes S durch eine globale Schwelle $T = 100$.

len Fällen durchaus möglich sein. Abbildung 7.4 zeigt jedoch, daß die Suche nach einer solchen Schwelle aufwendig ist und für jedes einzelne Beispiel von neuem durchgeführt werden muß. Die lokale Schwelle wurde nicht einem speziellen Bildmaterial angepaßt und läßt sich universell einsetzen. Sie paßt sich im Gegensatz zu einer globalen Schwelle selbst dem Bildinhalt an.

Abbildung 7.2b zeigt eine Grenzschicht zwischen einer Sandoberfläche und einem aufliegenden Geotextil. Man erkennt deutlich, wie eine ganze Schicht der Sandoberfläche abrutscht.

Um die Allgemeingültigkeit des Verfahrens zu testen, wurde es zusätzlich auf eine Verkehrsszene angewendet (7.2c). Ohne eine Änderung der Parameter der Fuzzy-Logik werden dort im rechten, oberen Viertel deutlich zwei Fahrzeuge und im rechten, unteren Viertel ein Fußgänger detektiert.

Teil II

Messung & Auswertung

Kapitel 8

Problemstellung der Messung

8.1 Motivation

Natürliche Flußläufe haben sich ihr Bett im Laufe von Jahrhunderten und Jahrtausenden gegraben. In dieser Zeit veränderten sie ihren Lauf immer wieder, bis er dem entsprach, was man heute kennt: stabile Flußläufe. Dabei wurden Bodenpartikel langsam zu natürlichen Filterschichten aufgehäuft, die den Uferböschungen die nötige Stabilität gaben. Wo dies nicht der Fall war, wurden sie weggespült und gaben den Weg frei für einen neuen Lauf des Flusses.

Sollen künstliche Flußläufe modelliert werden oder natürliche Uferböschungen saniert werden, kann diese „trial and error Methode“ nicht angewendet werden. Man muß sich von vornherein über die Beschaffenheit der Sedimente und die Eigenschaft der verwendeten Baustoffe im klaren sein. Eine globale Betrachtung dieser Eigenschaften läßt nur Aussagen über die integrierte Wirkung der verwendeten Materialien zu. Einzig eine mikroskopische Betrachtung der Vorgänge im Boden kann das Verhalten der Sedimente genau beschreiben. Dies ist ein typischer Fall, bei dem eine mikroskopische Betrachtung von im einzelnen gesehen, kleinen Ereignissen, eine große Auswirkung, wie z. B. das Einstürzen ganzer Uferböschungen vorhersagen kann.

Zur Untersuchung der Verlagerung einzelner Sedimentteilchen in Abhängigkeit von Druckschwankungen in Bodenschichten von Flußläufen, sollten Ausschnitte der Flußsedimente durch ein Endoskop beobachtet werden. Eine integrierte Glasfaserbeleuchtung erlaubt eine Beobachtung tieferer Bodenschichten. Die Anwesenheit des Endoskops selbst stellt eine Einschränkung der Bewegungsfreiheit der Sandteilchen dar. Eine mögliche Bewegung der Sedimente wird dadurch zwangsweise auf die 2D-Ebene des Endoskopkopfes eingegrenzt. Zusammen mit den Überlegungen zur Qualität des Bildmaterials (Kapitel 2) ergibt sich daraus, daß eine quantitative Analyse der Geschwindigkeit nicht besonders sinnvoll ist und keine Aussagen über die tatsächliche Bewegung bei Abwesenheit eines Endoskopes zuläßt. Das Auftreten von Bewegung an sich wird jedoch nicht durch das Endoskop verhindert. Eine rein qualitative Detektion der Bildbereiche, an denen Bewegung stattfindet, kann mit dem in Teil I dieser Arbeit beschriebenen Algorithmus der Bewegungssegmentierung durchgeführt werden.

Bewegung wird immer durch externe Kräfte hervorgerufen und ist nur bei Vorhandensein von Hohlräumen und Instabilitäten bzw. Lufteinschlüssen im Sedimentverbund möglich. Daher liefert

eine Beobachtung und Lokalisierung von Partikelverschiebungen, in Abhängigkeit von hydrodynamischen und mechanischen Belastungen, wichtige Aussagen über die Stabilität von Uferbefestigungen.

8.2 Mechanische Belastungen und Instabilitäten in Flußufern

Im Rahmen des Forschungsprojektes „Materialverlagerung an Grenzschichten unter hydrodynamischer Belastung“ der Bundesanstalt für Wasserbau, sollen Kornverlagerungen in Flußsedimenten anhand von maßstabsgetreuen Uferböschungen und realistischen Druckschwankungen simuliert werden. Solche Druckschwankungen werden in realen Flüssen durch Schiffe erzeugt. Diese verdrängen große Mengen von Wasser und führen selbst bei niedriger Geschwindigkeit zu einem Heben und Senken des Wasserspiegels der Flüsse [Rath, 1992]. Während die Bugwellen der Schiffe nur oberflächliche Zerstörungen der Uferbefestigungen verursachen, indem selbst große Steine bewegt und weggerollt werden [Köhler, 1985], führen die langsameren Druckschwankungen zu Schäden der tieferen Sedimentschichten [Köhler, 1993].

Diese Schäden werden durch Änderungen der Druckpotentiale im Boden verursacht, wobei der *Porenwasserüberdruck* eine entscheidende Rolle spielt [Köhler, 1989]. Bei zeitlich konstantem Wasserspiegel herrscht im Boden ein Gleichgewicht aller angreifenden Kräfte, einschließlich Druck und Auftrieb der Sedimente im Wasser. Der Wasserdruckzustand im Boden ist dann zu jedem Zeitpunkt t gleich dem hydrostatischen Wasserdruck über dem betrachteten Bodenelement. Bei ausreichend wasserdurchlässigem Boden kann die Änderung des Wasserdruckes durch ein Absenken des Wasserspiegels ohne Verzögerung ausgeglichen werden. Alle Bodenelemente in verschiedenen Tiefen spüren diese Druckänderungen gleichzeitig, wobei die bereits bestehenden statischen Druckgradienten erhalten bleiben.

Da der mit Wasser gefüllte Porenraum des Bodens noch bis in relativ große Tiefe unter dem Grundwasserspiegel Luft in gelöster und in freier Form enthält, ist ein Sediment nicht ideal wasserdurchlässig. Dies führt zu einer Dämpfung von Druckschwankungen im Boden. Kann die Porenwasserdruckanpassung auf Grund einer nicht ausreichend wasserdurchlässigen Bodenbeschaffenheit der Absenkung nicht unmittelbar folgen, so entsteht für einen kurzen Zeitraum Δt nach Eintritt der Absenkung noch keine Porenwasserdruckänderung. Dadurch bleibt in Abhängigkeit von der Tiefe ein überhöhter Wasserdruck erhalten, der sog. Porenwasserüberdruck, der sich im Laufe der Zeit bis auf den Gleichgewichtswert abbaut. Diese instationären Druckgradienten führen zu Wasserströmungen entgegen der Richtung der Wasserspiegeländerung. Das Porenwasser wird somit durch die Sedimentschichten nach oben gedrückt. Da die Sedimentschichten und damit auch die wasserundurchlässigen Zonen in Uferböschungen parallel zur Böschung verlaufen, entsteht bei einer Wasserspiegelabsenkung ein hydraulischer Druckgradient und dadurch eine Porenwasserströmung senkrecht zur Böschungsoberfläche. Zusätzlich können, insbesondere oberhalb des abgesenkten Wasserspiegels, Fließvorgänge parallel zur Böschung entstehen (Abb. 8.1).

Um die Standsicherheit einer Uferbefestigung nicht zu gefährden, darf der im Böschungsuntergrund entstehende Porenwasserüberdruck, in der Tiefe z unterhalb des Deckwerks, den Gleichgewichtszustand aller angreifenden Kräfte nicht soweit stören, daß ein Abrutschen auf einer

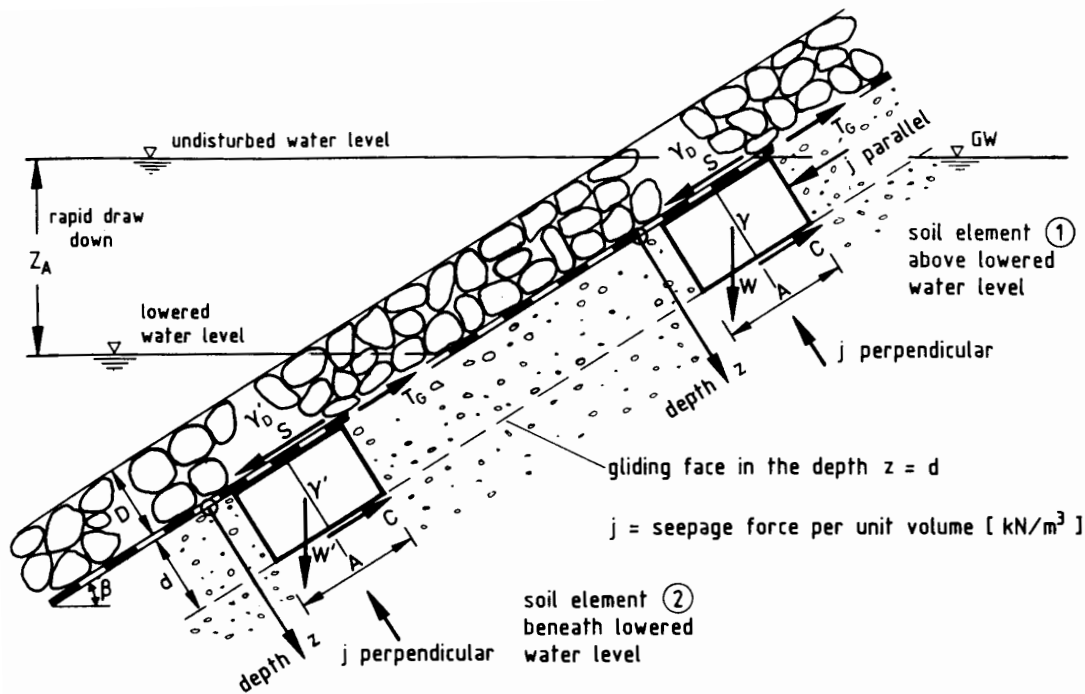


Abbildung 8.1: Gleitfläche im Untergrund einer Uferböschung und angreifende Kräfte bei einer Absenkung des Wasserspiegels. Mit j wird der Strömungsdruck bezeichnet. [Köhler, 1989].

böschungsparellen Gleitfuge ermöglicht wird. Dazu werden die verwendeten Wasserbausedimente in einer Abfolge aufeinander geschichtet, die verhindert, daß solche Gleitfugen (Abb. 8.1) in Uferböschungen entstehen. Diese speziellen Sedimentschichtungen werden als *Filter* bezeichnet. Ein solcher Filter fungiert, wenn man die bodenmechanischen Aspekte betrachtet, als eine Schutzschicht gegen Bodenerosion durch Wasserflüsse. Er sollte gegen Wasserbewegungen, durch stationäre und instationäre hydraulische Gradienten, parallel und senkrecht zur Filterschicht resistent sein [Köhler, 1993]. Ein idealer Filter hat die Eigenschaft, kleine Bodenpartikel festzuhalten und gleichzeitig als Drainage zu wirken, die Wasser ungehindert abfließen läßt, um ein Entstehen von Porenwasserüberdruck zu verhindern. Diese beiden Eigenschaften lassen sich nicht immer gleichzeitig und in optimaler Weise erfüllen.

Als klassische Filter werden *Kornfilter* verwendet. Diese bestehen aus einer Abfolge von Sandschichtungen mit nur kleinen Unterschieden der Korngrößen benachbarter Filterschichten. Die Verteilung der Korngrößen in verschiedenen Bodenarten wird im allgemeinen durch sog. *Sieblinien* (Abb. 8.2) grafisch veranschaulicht. Diese Summenkurven erhält man durch Summieren aller Fraktionsanteile eines bestimmten Sedimentes, die der Größe nach bis zu einer bestimmten Korngröße aufeinander folgen¹, [Scheffer, 1984]. Anschaulich entspricht dies dem Mengenanteil der Körner, die durch ein Sieb einer bestimmten Lochgröße hindurchfallen. Ein steiler Verlauf

¹Durch Differenzieren erhält man aus den Sieblinien die, dem Physiker wesentlich vertrautere *Häufigkeitsverteilung* der Korngrößen. Da die Sieblinien von Sand und Kies im allgemeinen einen s-förmigen Verlauf aufweisen, welcher der Error-Funktion ähnelt, haben Sedimente im wesentlichen normalverteilte Korngrößen über einer logarithmischen Skala.

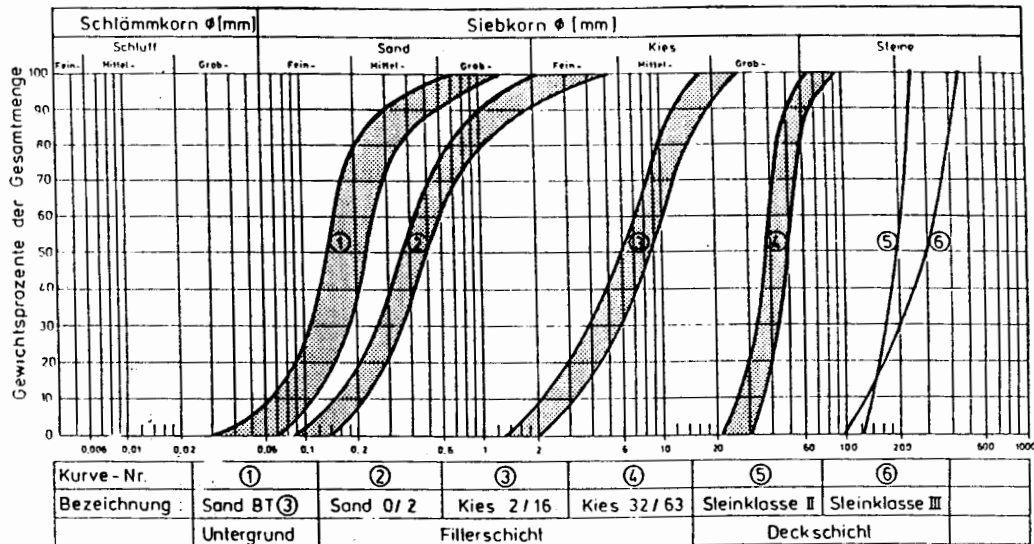


Abbildung 8.2: Sieblinien verschiedener Flußsedimente [Köhler, 1989]. In den Messungen dieser Arbeit wurden die Sedimente 1, 2 und 3 (zum Teil zusammen mit Geotextilien) verwendet.

der Kurve deutet auf einen hohen Anteil der betreffenden Fraktion hin. Ein Kornfilter ist nur dann stabil, wenn sich die Sieblinien benachbarter Sedimentschichten deutlich überlappen. Da die einzelnen Schichten der verschiedenen Sedimente in einem Kornfilter nicht beliebig dünn sein können (mindestens 15 cm Schichtdicke [Köhler, 1993]), werden Kornfilter insgesamt sehr dick um die gesamte Spannweite von feinen Sanden, bis zu großen Wasserbau-Steinen abzudecken.

Um eine kompaktere Bauweise zu ermöglichen, werden seit einiger Zeit *Geotextilien* als Filter eingesetzt. Diese werden direkt auf die feinen Sande des Untergrundes aufgelegt und von losen Steinschüttungen abgedeckt. Sie bestehen, je nach Ausführung, aus mehreren Schichten verschiedener Kunstfasern mit unterschiedlicher Struktur und haben eine Dicke von ca. 1 - 2 cm. Durch ihre verfilzte, lockere Struktur sollten sie die Eigenschaften eines idealen Filters erfüllen.

Die Stabilität eines Filters zeigt sich insbesondere an Grenzsichten zwischen verschiedenen Sedimenten oder Übergängen zwischen Sedimenten und anderen Bauwerken (z. B. unterirdische Wasserrohre). Bestehen benachbarte Filterschichten aus Sanden stark unterschiedlicher Korngrößen, können sich die kleinen Körner unter Umständen durch die Hohlräume der größeren hindurchbewegen. Dies kann zum Materialtransport über größere Strecken und zum Aushöhlen der Uferbefestigungen führen. Geotextilien scheinen auf den ersten Blick solche Materialverlagerungen zu verhindern. Es zeigt sich jedoch, daß Bodenpartikel durchaus die Textilien durchdringen können. Dies führt zu Hohlräumen unter den Geotextilien, die eine Erosion an diesen Stellen begünstigen. Durch Einlagerung von Sandpartikeln in die Filzstruktur des Geotextiles wird darüber hinaus die Wasserdurchlässigkeit vermindert. Dadurch wird der Porenwasserüberdruck unterhalb des Textiles verstärkt. Dieser führt zu einem vermehrten Transport von Wasser und Material durch die Filterschicht und wirkt der Normalkraft des Deckwerkes entgegen, was wiederum dessen Standsicherheit beeinträchtigt.

Ein Vergleich von natürlichen Filterschichten mit Geotextilien soll Aufschluß über den Nutzen dieser Materialien liefern. Während bisher integrierende Meßmethoden eingesetzt wurden, erfolgt nun eine detaillierte Beobachtung der Vorgänge in Sedimentgrenzschichten.

Ein großer Sedimenttank, der eine Simulation der Vorgänge in realen Gewässern ermöglicht, wird im Laufe des Jahres 1994 fertiggestellt werden. Im Rahmen dieser Diplomarbeit sollten, außer der Entwicklung der benötigten Bildverarbeitungsalgorithmen, Voruntersuchungen zur Kornverlagerung in Grenzschichten in einem kleinen Versuchstank erfolgen. Anhand dieser Messungen konnte die Verlässlichkeit der Bewegungssegmentierung überprüft werden. Außerdem wurden damit erstmals die mikroskopischen Vorgänge in Grenzschichten visualisiert und damit einer experimentellen Untersuchung zugänglich gemacht. Dies lieferte wichtige Aussagen über die Größe der zu erwartenden Bewegungen und die Anforderungen, die an den großen Tank gestellt werden müssen.

Ohne direkt die Vorgänge in realen Gewässern zu simulieren, sollten die Messungen Aufschluß über mögliche Bewegung innerhalb von Grenzschichten unter Druckbelastungen geben. Da die Sedimente horizontal geschichtet wurden, konnte hierbei kein Abgleiten von Filterschichten erwartet werden. Die Sedimentschichten wurden vielmehr als ein 3-Phasen-System von Luft, Wasser und Feststoffen angesehen und dessen Reaktion auf eine aufgeprägte Druckschwankung von bekanntem funktionalen Zusammenhang studiert.

Kapitel 9

Experimentelle Durchführung

9.1 Sedimenttank

Der Versuchstank (Abbildung 9.1 und Farbtafel 1) stellt ein verkleinertes Modell des geplanten großen Sedimenttanks dar. Er besteht aus Stahlplatten von 8 mm Stärke und läßt sich druckdicht verschrauben. Die Öffnungen zum Einlaß des Endoskopes können wahlweise am Boden oder Deckel des Tanks angebracht werden. Dadurch wird eine Beobachtung sowohl der unteren als auch der oberen Seite eines Geotextiles ermöglicht. Durch eine spezielle Klemmvorrichtung (Farbtafel 2) kann das Endoskop leicht in den Tank eingeführt werden. Nach dem Festschrauben wird die Einlaßöffnung druckdicht verschlossen, ohne jedoch die empfindliche Optik durch einseitige Spannungen zu quetschen. Das Endoskop wird dadurch fest mit dem Tank verbunden. Diese feste Verbindung ist notwendig, um ein Wackeln des Endoskopes zu verhindern. Eine solche Eigenbewegung der Kameraoptik würde eine Bewegung des gesamten Bildausschnittes vortäuschen und damit die Detektion der eigentlichen Bewegung verhindern.

Durch eine Wasserpumpe kann eine konstante Strömung im Sediment erzeugt werden. Da der Wasserkreislauf geschlossen ist, wird dies unabhängig vom momentanen Innendruck des Tanks ermöglicht. Der Innenraum wird durch zwei perforierte Bleche in drei Kammern aufgeteilt. In die Mitte, zwischen die beiden Bleche, wird das zu untersuchende Sediment geschichtet. Die Löcher der beiden Bleche können durch zusätzliche Gewebe¹ verkleinert werden. Dadurch kann Wasser ungehindert durch sie hindurchströmen, während die Sandkörner selbst vor dem Wegspülen geschützt werden und nicht in die Pumpe gelangen können. Durch die beiden Bleche entstehen zwei weitere Kammern direkt vor dem Ein- bzw. Ausfluß des Wasserkreislaufes. Diese dienen als *Vorfluter*, um die Strömung zu beruhigen. Durch zwei zusätzliche *Prallbleche* in der Mitte der Vorfluter wird verhindert, daß die Strömung ungehindert auf die Mitte des Sedimentes trifft. Diese kleinen Bleche sind ebenfalls mit Löchern versehen, deren Durchmesser in der Mitte am kleinsten ist und nach außen hin immer mehr zunimmt. So wie der Strömungsdruck mit zunehmendem Abstand von der Eintrittsöffnung abnimmt, wird das Durchdringen der Prallbleche erleichtert. Dadurch wird die Hauptströmung gezwungen, sich mehr oder weniger homogen über

¹ Fliegengitter für groben Sand und Siebdruckgewebe für feinen Sand.

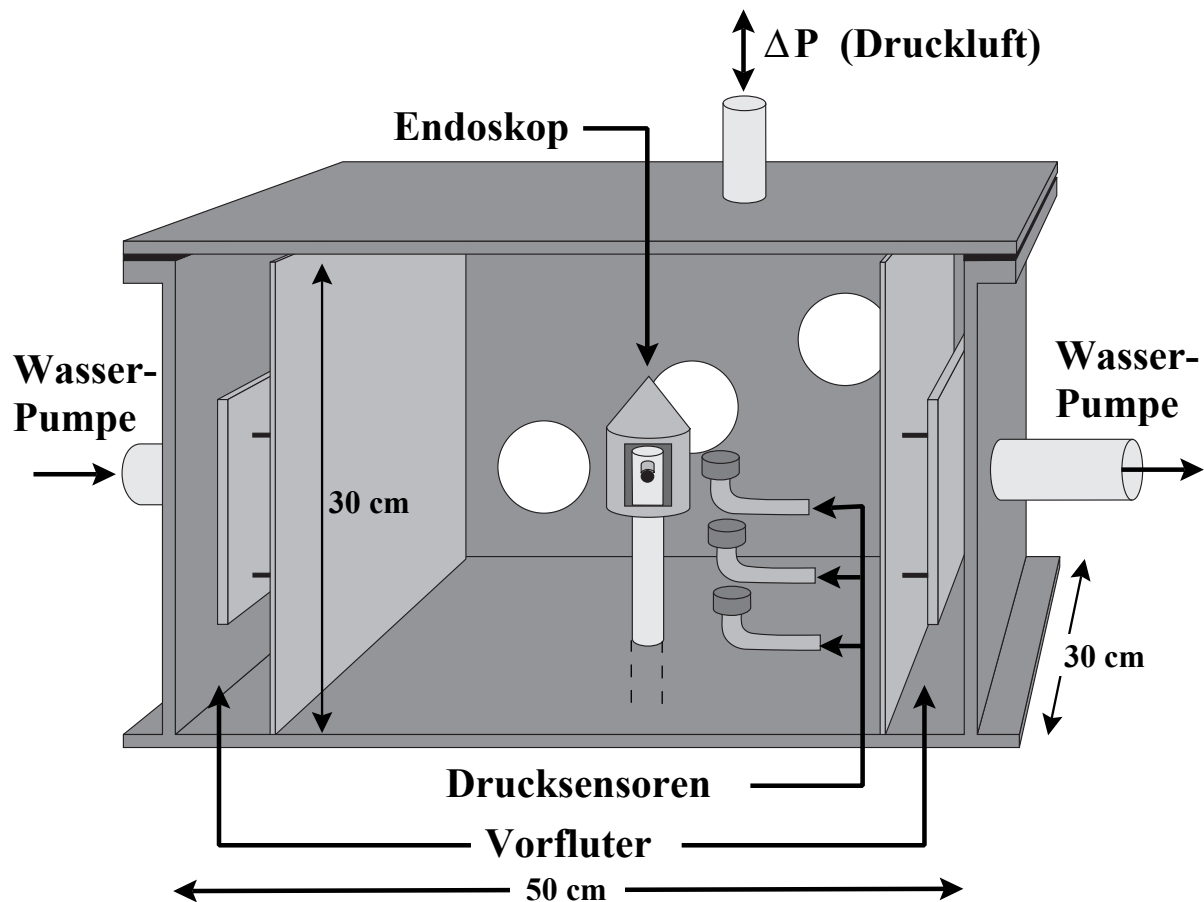


Abbildung 9.1: Schematischer Aufbau des Sedimenttanks (siehe auch Farbtafeln 1 und 3).

die gesamte Fläche des Tanks zu verteilen, bevor sie auf das Sediment trifft. Für das Absaugen des Wassers auf der gegenüberliegenden Seite gelten die gleichen Überlegungen.

Druckschwankungen im Sediment werden pneumatisch erzeugt (Abschnitt 9.3). Eine Erhöhung bzw. Absenkung des Luftdrucks der Luftschicht über dem Wasserspiegel, wird dem hydrostatischen Druck in einer bestimmten Wassertiefe additiv überlagert. Durch einen statischen Innendruck des Tanks kann eine bestimmte Wassertiefe und durch dynamische Änderungen dieses Drucks können Schwankungen des Wasserspiegels simuliert werden.

Zur Messung der Druckverteilung in verschiedenen Tiefen des Sedimentes werden drei Druckgeber² eingesetzt. Diese messen Druckunterschiede zwischen 0 und 1 bar (0–1000 cm WS) mit einer Genauigkeit von 2 cm WS gegenüber Atmosphärendruck. Damit können Wassertiefen und Druckunterschiede von bis zu 10 m Wassersäule simuliert werden. Zur groben Überwachung des Innendrucks ist im Deckel des Tanks ein Manometer eingebaut.

Bullaugen an verschiedenen Stellen der Wand und am Deckel (Farbtafel 1) ermöglichen es, die Vorgänge im Sediment und auf der Wasseroberfläche teilweise von außen zu beobachten.

²TransInstruments, Basingstoke, England, Type BHL-4291-01-50M0, Range 0-1 b G.

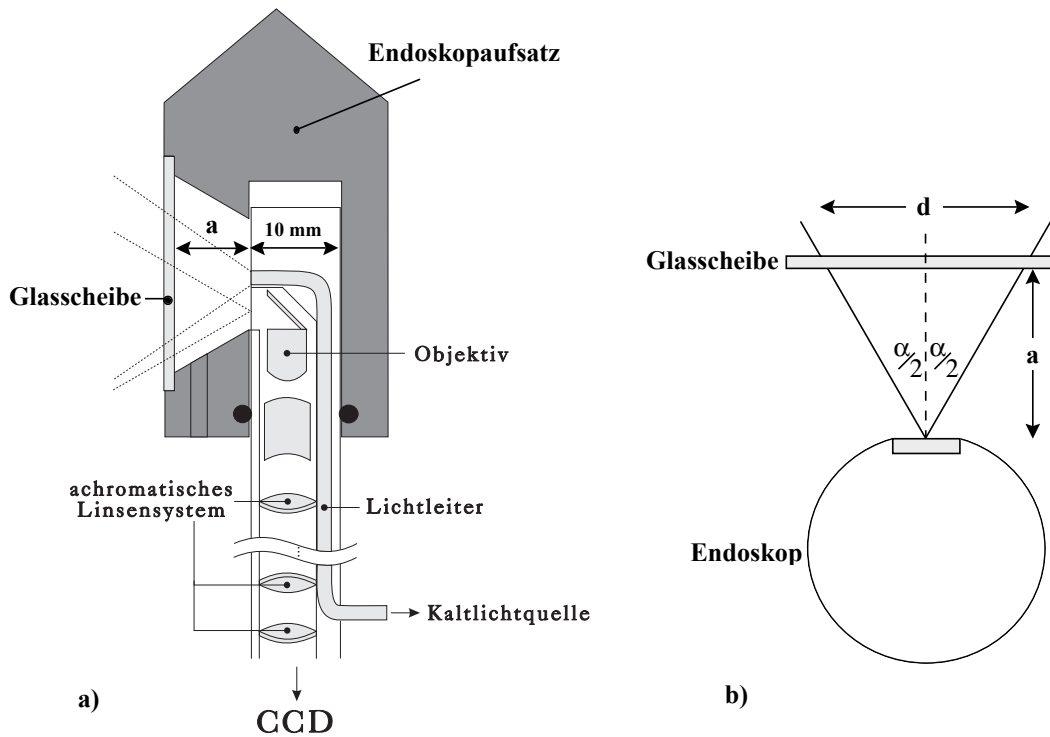


Abbildung 9.2: Schematischer Aufbau des Endoskops zusammen mit dem Endoskopaufsatz (siehe auch Farbtafeln 2 und 3). a): Längsschnitt durch das System. b): Querschnitt durch das Endoskop und die Glasscheibe des Endoskopaufsatzes an der Austrittsöffnung zur Berechnung der Größe des Bildausschnittes. Mit α wird der Öffnungswinkel der Optik bezeichnet.

9.2 Endoskop

Zur Beobachtung der Vorgänge in den Sedimentschichten wird ein *Endoskop* als Kameraobjektiv eingesetzt. Es handelt sich dabei um ein starres Endoskop³ mit einem Blickwinkel von 90° gegen die Endoskopachse. Abbildung 9.2a zeigt den schematischen Aufbau.

Das Endoskop hat einen Durchmesser von 10 mm und eine Länge von 55 cm. Die Hülle besteht, zum Schutz der Optik, aus rostfreiem Edelstahl. Sie ist druckdicht bis 3 bar und ermöglicht den Einsatz unter Wasser.

Die Beleuchtung des Bildausschnittes erfolgt mittels einer separaten Kaltlichtquelle⁴, die mit dem Endoskop durch einen Lichtleiter verbunden wird. In das Endoskop selbst sind Glasfasern eingezo-gen. Sie leiten das Licht zur Endoskopspitze. So werden die Objekte intensiv und ohne Erwärmung ausgeleuchtet.

Als Schutz der Glasscheibe der Austrittsöffnung gegen ein Zerkratzen durch Sandkörner, wird auf die Endoskopspitze ein selbstkonstruierter Endoskopaufsatz aufgesteckt. Dieser besteht aus

³Classen, 50284.

⁴Classen, Endolux 13990, 150 W.

einem Aluminiumdrehkörper, dessen Vorderseite abgeflacht ist und durch eine Glasscheibe abgedeckt wird (Abbildung 9.2a). Ist das Endoskop vollständig von Sandkörnern bedeckt, werden diese durch den Endoskopaufsatz auf einen konstanten Objektabstand von 8 mm gehalten. Ohne diesen Abstandhalter würden die Sandkörner die Austrittsöffnung der Optik vollständig von der Beleuchtung abschirmen und somit eine Bildaufnahme unmöglich machen. Die Glasscheibe des Endoskopschutzes verursacht jedoch neue Probleme: Durch die stark unterschiedlichen Brechungsindizes n von Luft und Glas (Tabelle 9.1) und dem damit verbundenen großen Reflexionskoeffizienten des Übergangs von Luft in Glas, wird ein großer Teil des Lichtes der Beleuchtungsquelle an der Glasscheibe in die Optik zurückreflektiert. Dies führt zu großflächigen Spiegelreflexen und so zu stark überbelichteten Bildbereichen. Um diese Effekte zu reduzieren, wird der Hohlraum zwischen der Glasscheibe und dem Endoskop mit Wasser aufgefüllt. Der Brechungsindex von Wasser liegt zwischen denen von Luft und Glas. Dadurch können die Reflexionen nur verringert, jedoch nicht vollständig unterdrückt werden (helle Flächen in den Bildern 10.5 - 10.9). Daraus ergeben sich die bereits erwähnten Probleme bei der Auswertung des Bildmaterials (Abschnitt 2.1). Um zu verhindern, daß Schlamm und Sandpartikel hinter die Glasscheibe gelangen, wird der Endoskopaufsatz, nach dem Auffüllen mit Wasser, durch eine Plastikschraube verschlossen.

| Substanz | n |
|-----------|----------|
| Luft | 1.000293 |
| Wasser | 1.333 |
| Quarzglas | 1.458 |

Tabelle 9.1: *Brechungsindizes verschiedener Substanzen [Hecht, 1987].*

Die Größe des beobachtbaren Bildausschnitts läßt sich aus Abb. 9.2b direkt ableiten: Der Öffnungswinkel α des verwendeten Endoskopes beträgt 60° . Bei einem Abstand $a = 8 \text{ mm}$ erhält man für den Durchmesser des Bildausschnitts:

$$d = 2a \tan(\alpha/2) \approx 9.2 \text{ mm}.$$

Die Bildübertragung erfolgt durch ein achromatisches Linsensystem. Am Endoskopende befindet sich ein Zoom-Okkular, das die Vergrößerung bis zu einem Faktor 2 stufenlos erhöht. Die Gesamtvergrößerung setzt sich aus der Vergrößerung des Objektivs und des Okkulars zusammen. Abbildung 9.3 zeigt die Abhängigkeit der Gesamtvergrößerung des verwendeten Endoskopes vom Objektabstand bei verschiedenen Einstellungen des Zoom-Okkulars. Unterhalb von 30 mm steigt die Vergrößerung stark an. Dies bedeutet, daß Partikel, die sich näher am Objektiv befinden, sehr viel stärker vergrößert werden als weiter entfernte. Bei einem Einsatz des Endoskops in Sedimenten, befinden sich alle zu beobachtenden Sandkörner unmittelbar hinter der Glasscheibe der Endoskopabdeckung. Die Gesamtvergrößerung des Endoskopes wird dadurch direkt durch den Abstand dieser Glasscheibe von der Austrittsöffnung der Optik bestimmt.

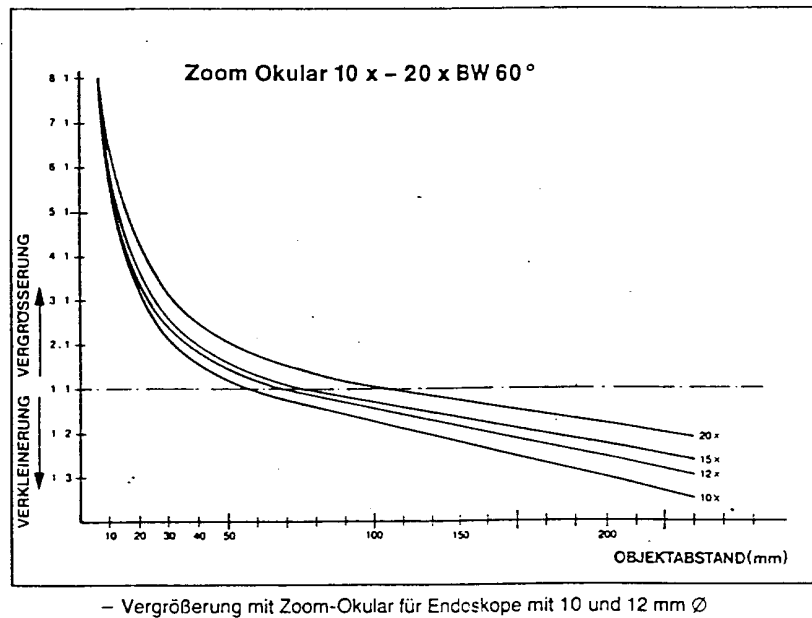


Abbildung 9.3: Vergrößerung für Endoskope in Abhängigkeit zum Objektabstand (aus techn. Prospekt, Fa. Classen).

9.3 Aufbau des Pneumatiksystems

Zur Erzeugung von Druckschwankungen im Sediment wird der Luftdruck über der Wasseroberfläche erhöht und abgesenkt. Diese Schwankungen des Luftdrucks addieren sich zu dem hydrostatischen Druck im Wasserkörper. Durch die Druckdämpfung der Sedimentschichten erreichen diese Druckschwankungen verschiedene Tiefen nur mit zeitlicher Verzögerung und verursachen den bereits beschriebenen Porenwasserüberdruck (Abschnitt 8.2).

Abbildung 9.4 zeigt den schematischen Aufbau der Pneumatikinstallation des Sedimenttanks. Als Druckversorgung wird die Preßluftleitung des Instituts angeschlossen. Sie liefert einen Druck P_o zwischen 6 und 7 bar. Die gewünschten Druckschwankungen bewegen sich zwischen 1 bar (Athmosphärendruck, 10 m WS) und 2 bar (20 m WS). Über einen Druckminderer wird der gewünschte Überdruck P_{in} eingestellt. Da die verwendeten Druckgeber direkt den relativen Überdruck gegen Athmosphärendruck anzeigen, wird im folgenden Druck immer in dieser Einheit angegeben. Dabei entspricht 0 cm WS Athmosphärendruck und 1000 cm WS dem Druck in 10 m Wassertiefe.

9.3.1 Druckanstieg

Beim Betätigen des 3-Wegeventils (Stellung 1) zu einem Zeitpunkt t_o liegt der Druck P_{in} am Tankvolumen an. Der Luftdruck des Tanks zu einem Zeitpunkt t sei $P(t)$. Für $P_{in} > P(t_o)$ strömt die Luft in den Tank. Im Tank erhöht sich der Druck im Laufe der Zeit, bis er den Gleichgewichtswert P_{in} erreicht. Die Geschwindigkeit, mit der der Druck aufgebaut wird hängt

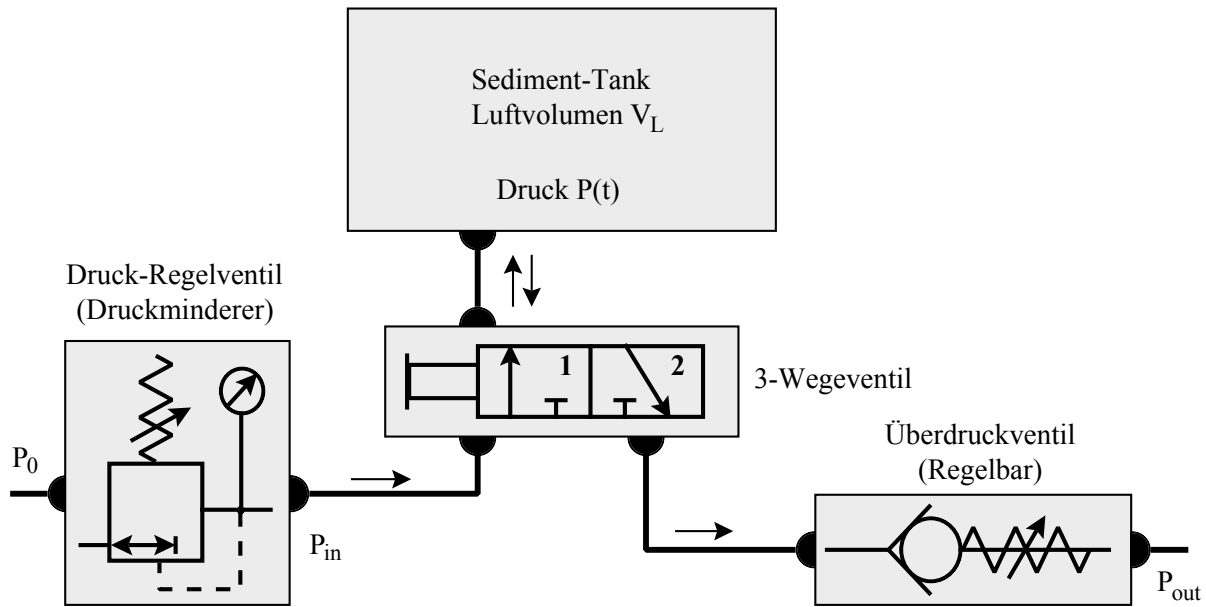


Abbildung 9.4: Pneumatiksystem des Sedimenttanks. P_o bezeichnet den Druck der Preßluftleitung des Instituts. Er beträgt ca. 7 bar. Über einen Druckminderer wird dieser Druck auf den gewünschten Arbeitsdruck P_{in} reduziert. Das 3-Wegeventil öffnet wahlweise eine Verbindung zwischen Tankvolumen und Druckminderer oder Tankvolumen und Überdruckventil. Durch das Überdruckventil wird der Innendruck des Tanks beim Ablassen des Drucks nur bis auf den Wert P_{out} abgebaut. Dieser stellt somit den Offset des Drucks dar.

von der Form der verwendeten Leitungen und deren Durchmesser ab. Geht man von der Annahme aus, daß die Änderungsgeschwindigkeit des Luftdrucks $dP(t)/dt$ direkt proportional zur anliegenden Druckdifferenz ist, dann ergibt sich für den zeitlichen Verlauf des Luftdrucks $P(t)$ die folgende Differentialgleichung:

$$\frac{dP(t)}{dt} = \kappa (P_{in} - P(t)), \quad (9.1)$$

mit der Proportionalitätskonstanten κ . Als Lösung von (9.1) erhält man einen exponentiellen Druckanstieg der Form

$$P(t) = P_{in} + (P(t_o) - P_{in}) e^{-\kappa(t-t_o)}, \quad (9.2)$$

der sich, ausgehend vom Tankinnendruck $P(t_o)$ zum Zeitpunkt t_o , asymptotisch dem Gleichgewichtswert P_{in} nähert. Der Parameter κ ergibt sich aus der Physik der Öffnung und muß experimentell bestimmt werden.

9.3.2 Druckabfall

Beim Loslassen des 3-Wegeventils (Stellung 2) zu einem Zeitpunkt t_1 wird das Luftvolumen des Tanks nach außen geöffnet. Das Überdruckventil verhindert, daß der Druck bis auf Atmosphärendruck abfällt. Eine regulierbare Feder drückt mit einer Kraft F auf eine Membran mit Querschnitt A , welche die Öffnung verschließt. Die ausströmende Luft spürt dadurch einen Gegendruck P_{out} der Größe $P_{out} = F/A$. Dieser mechanische Druck stellt für die ausströmende

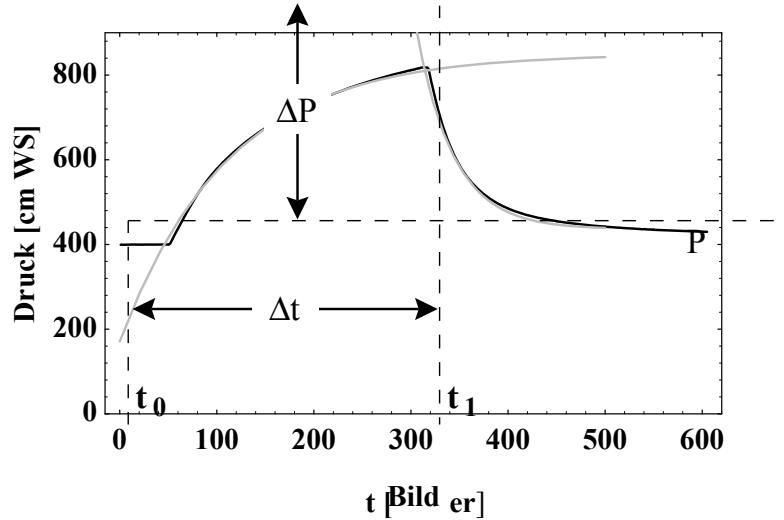


Abbildung 9.5: Druckverlauf während einer Messung. Zur Synchronisierung mit der Bildaufnahme wurden die Druckwerte im doppelten Videotakt (50 Hz) ausgelesen.

Luft den gleichen Widerstand dar, wie ein Außendruck derselben Größe. Sinkt der Innendruck des Tanks unter diesen Wert, kann keine weitere Luft ausströmen. Analog zu Abschnitt 9.3.1 ergibt sich die Differentialgleichung des Druckabfalls:

$$\frac{dP(t)}{dt} = \kappa (P_{out} - P(t)) \quad (9.3)$$

und der Druckverlauf:

$$P(t) = P_{out} + (P(t_1) - P_{out}) e^{-\kappa(t-t_1)}. \quad (9.4)$$

Da der Außendruck P_{out} kleiner ist als der Innendruck $P(t_1)$ zum Zeitpunkt t_1 , stellt dies eine abfallende Exponentialfunktion dar.

9.3.3 Meßwerte

Abbildung 9.5 zeigt den Druckverlauf des obersten der drei Druckgeber während einer Messung. Da dieser direkt unterhalb einer wasserdurchlässigen Sedimentschicht plziert war, entsprechen diese Meßwerte dem Tankinnendruck (der hydrostatische Druck in einer Wassertiefe von ca. 15 cm ist im Vergleich zu der erzeugten Druckschwankung vernachlässigbar). Das Überdruckventil war auf einen Wert $P_{out} \approx 400 \text{ cm WS}$ eingestellt, wobei sich dieser Wert nur sehr grob regulieren und nur anhand des gemessenen Druckverlaufes ablesen läßt. Der am Ventil anliegende Überdruck P_{in} betrug $P_{in} \approx 850 \text{ cm WS}$. Zu einem Zeitpunkt t_o wurde das Ventil betätigt und zu einem Zeitpunkt $t_1 = t_o + \Delta t$ wieder losgelassen.

Der gemessene Druckverlauf (schwarze Kurve) läßt sich gut durch Funktionen der Form (9.2) und (9.4) anfitten (graue Kurven). Für den Anstieg und den Abfall erhält man die beiden Funktionen

$$\begin{aligned} P(t) &= 850 - 385 e^{-0.009(t-t_o)}, & t_o \leq t \leq t_1 \\ P(t) &= 436 + 322 e^{-0.026(t-t_1)}, & t > t_1. \end{aligned} \quad (9.5)$$

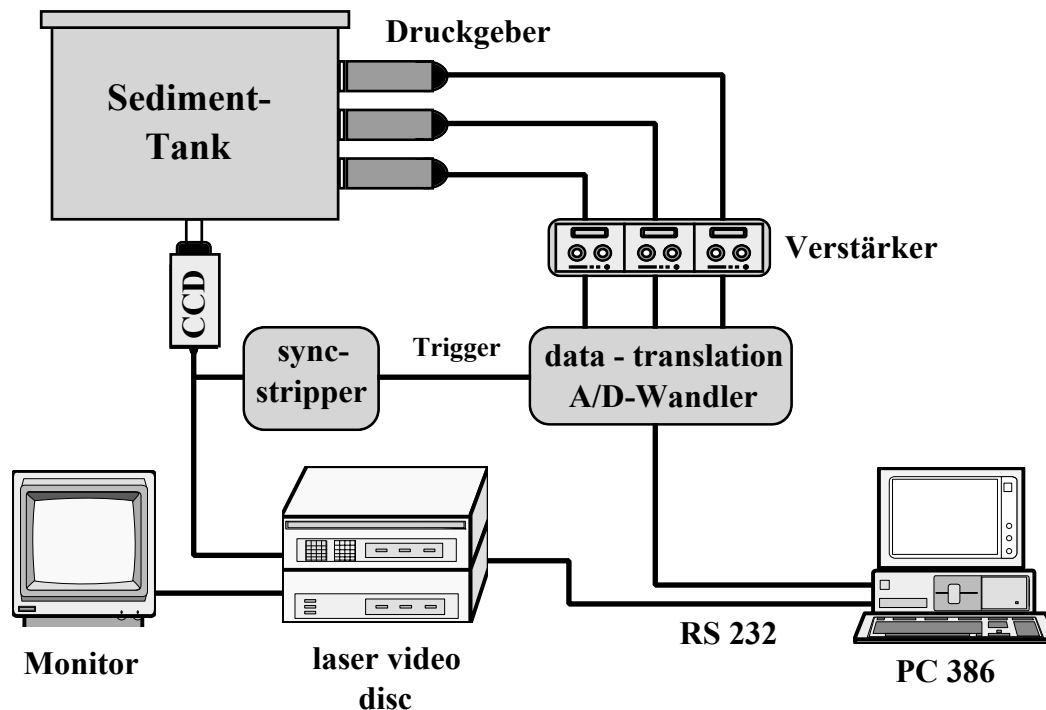


Abbildung 9.6: Versuchsaufbau: Die von der Kamera aufgenommenen Bilder werden auf einem analogen optischen Speichermedium aufgezeichnet. Die Bildaufnahme wird über die serielle Schnittstelle eines PC's gesteuert. Gleichzeitig messen drei Druckgeber im Videotakt den Druckverlauf im Sediment.

und daraus die Werte $P_{in} = 850 \text{ cm WS}$ und $P_{out} = 436 \text{ cm WS}$. Es fällt auf, daß der Druckabfall um einen Faktor drei schneller erfolgt als der Druckanstieg. Dies liegt vermutlich daran, daß innerhalb des Druckminderers, zur Reduzierung des anliegenden Drucks P_o auf den Druck P_{in} , der Querschnitt der Leitung stark verengt wird. Dadurch kann die Luft nicht schnell genug nachströmen.

Im Gegensatz zu einem linearen Druckanstieg, wie er von Schiffen in realen Gewässern erzeugt wird, enthält der exponentielle Anstieg und Abfall sich zeitlich kontinuierlich ändernde Druckgradienten. Dies ermöglicht die Untersuchung der Abhängigkeit der Bewegung von den momentanen Druckgradienten innerhalb einer einzigen Messung.

9.4 Versuchsanordnung und Synchronisation

Abbildung 9.6 zeigt den schematischen Aufbau der Versuchsanordnung. Zur Bildaufnahme wird eine Videokamera⁵ über einen Adapter an das Endoskop angekoppelt. Die Bildwiederholrate beträgt 25 Bilder/s, wobei 50 mal pro Sekunde ein Halbbild übertragen wird. Das Videosignal wird auf einer optischen Bildplatte⁶ aufgezeichnet. Eine Seite dieser Bildplatte faßt etwa 37.000 Bilder in europäischer CCIR-Norm (512×512 Pixel/Bild), was einer Speicherkapazität von etwa

⁵Sony, AVC-D7CE s/w-CCD Videokamera.

⁶Sony LVR-5000 Laser Video Recorder.

9 Gbyte entspricht. Die Steuerung des Laser Video Rekorders erfolgt über die serielle Schnittstelle eines PC's. Über diese Schnittstelle wird parallel zur Bildaufnahme die aktuelle Bildnummer der Laser Disk abgefragt. Dies geschieht in ‚Echtzeit‘, d. h. 50 mal pro Sekunde im Videotakt. Zur Kontrolle der Vorgänge im Tank kann das ankommende Videosignal an einem zusätzlichen Videomonitor verfolgt werden.

Die Druckgeber liefern einen Spannungswert, dessen Größe proportional zum gemessenen Druck ist. Dieses Signal wird von einer Verstärkerelektronik verstärkt und durch einen Analog-Digital-Wandler⁷ digitalisiert. Dieser sitzt direkt auf dem PC-Bus und verfügt über eine externe Anschlußeinheit zur Aufnahme der Datenleitungen. Um die Druckmeßwerte eindeutig einer Bewegung zuzuordnen zu können, müssen diese im gleichen Takt aufgenommen werden wie die Videobilder. Dazu wird das Videosignal zusätzlich durch einen *sync-stripper*⁸ geführt. Dieser extrahiert den *vertical blank* des ursprünglichen Signals. Als Ausgang erhält man ein TTL-Signal, dessen Amplitude mit dem doppelten Videotakt variiert und somit einen absoluten Zeittakt zum Digitalisieren der Druckwerte durch den A/D-Wandler liefert. Dieser wird so gesteuert, daß er zu Beginn jedes Halbbildes die Spannungswerte der drei Druckgeber einliest. Im PC werden die Spannungssignale in Druck umgerechnet und zusammen mit der laufenden Bildnummer abgespeichert.

Durch diese Synchronisierung ist man in der Lage, die gemessenen Druckwerte nachträglich einem bestimmten Bild zuzuordnen und damit die Abhängigkeit der segmentierten Bewegung von dem momentanen Druck exakt zu bestimmen. Die zeitliche Auflösung sowohl des Drucks als auch der gemessenen Bewegung beträgt dadurch 20 ms.

9.5 Eichung der Geschwindigkeit

Wie bereits in Abschnitt 9.2 beschrieben wurde, setzt sich die Gesamtvergrößerung des Endoskops aus der Vergrößerung des Objektivs und der des Okkulars zusammen. Die Größe des Bildes auf dem CCD-Chip und damit der Abbildungsmaßstab wird zusätzlich durch die Geometrie und Optik des Kameraadapters bestimmt. Dieser enthält eine Linse der Brennweite $f = 30\text{ mm}$, die das Bild des Endoskops auf den CCD-Chip abbildet. Der CCD-Chip besteht aus 756×581 Sensorelementen (h×v) bei einer Abmessung von $8.8 \times 6.6\text{ mm}$. Bei der verwendeten Vergrößerung (Abstand 8 mm, Zoomokkular 20×) hat der beobachtbare Bildausschnitt eine Größe von $8 \times 6\text{ mm}$ (Eichaufnahme einer Millimeterskala). Dies entspricht ungefähr einer 1 : 1-Vergrößerung des Bildausschnitts auf die Bildebene.

Bei der Auswertung wird das Videosignal von einer Bildverarbeitungskarte⁹ digitalisiert. Die Auflösung der Karte beträgt 512 Zeilen und 512 Spalten (CCIR-Norm). Insgesamt erhält man damit eine Ortsauflösung von $16\text{ }\mu\text{m}$ in horizontaler und $12\text{ }\mu\text{m}$ in vertikaler Richtung. Der zeitliche Abstand der Bilder (Halbbilder) im Videotakt beträgt 20 ms. Eine Geschwindigkeit von $\vec{v} = 1\text{ Pixel/Bild}$ entspricht damit einer realen Geschwindigkeit von $v_x = 0.8\text{ mm/s}$ in horizontaler und $v_y = 0.6\text{ mm/s}$ in vertikaler Richtung.

⁷Data Translation, DT 2821.

⁸Thomas Münsterer, Eigenbau.

⁹ITI OFG, der Firma Imaging Technologies.

Kapitel 10

Ergebnisse der Messung

In diesem Kapitel sollen die Ergebnisse der Bewegungssegmentierung zur Detektion der Kornverlagerungen im Versuchstank vorgestellt werden. Dabei wurden fünf Versuchsbedingungen mit verschiedenen Filterschichtungen ausgewertet.

10.1 Schichtung der Sedimente

Abbildung 10.1 zeigt die Abfolge der Sedimentschichten und Geotextilien bei den verschiedenen Messungen. Das Endoskop beobachtet jeweils eine Grenzschicht zwischen verschiedenen Sedimenten bzw. zwischen Sediment und aufliegendem Geotextil. Bei den einzelnen Messungen wurden die folgenden Filter untersucht:

- **Messung A: Sand BT3 - Geotextil Terrafix 600 - Kies 2/16:**
Das Endoskop beobachtet den Übergang Sand BT3 - Geotextil. Dieses Geotextil besteht aus einer ca. 0.5 cm dicken Lage weißer, verfilzter Kunstfasern.
- **Messung B: Sand BT3 - Geotextil 905 RS - Kies 2/16:**
Das Endoskop beobachtet den Übergang Sand BT3 - Geotextil. Dieses zweilagige Geotextil besteht aus einer dünnen Filzschicht (0.5 cm) und einem dickeren Geflecht schwarzer, drahtiger Kunstfasern (1.5 cm).
- **Messung C: Sand BT3 - Kies 2/16:**
Das Endoskop beobachtet den Übergang Sand BT3 - Kies 2/16. Hierbei werden zwei Sedimente stark unterschiedlicher Korngröße direkt aufeinandergeschichtet. Dies stellt das klassische Beispiel eines instabilen Filters dar.
- **Messung D: Sand BT3 - Sand 02 - Kies 2/16:**
Das Endoskop beobachtet den Übergang Sand BT3 - Sand 02. Dieses Kornfilter soll stabil sein.
- **Messung E: Sand BT3 - Sand 02 - Kies 2/16:**
Gleiches Kornfilter wie bei Messung D. Das Endoskop beobachtet den Übergang Sand 02-Kies 2/16.

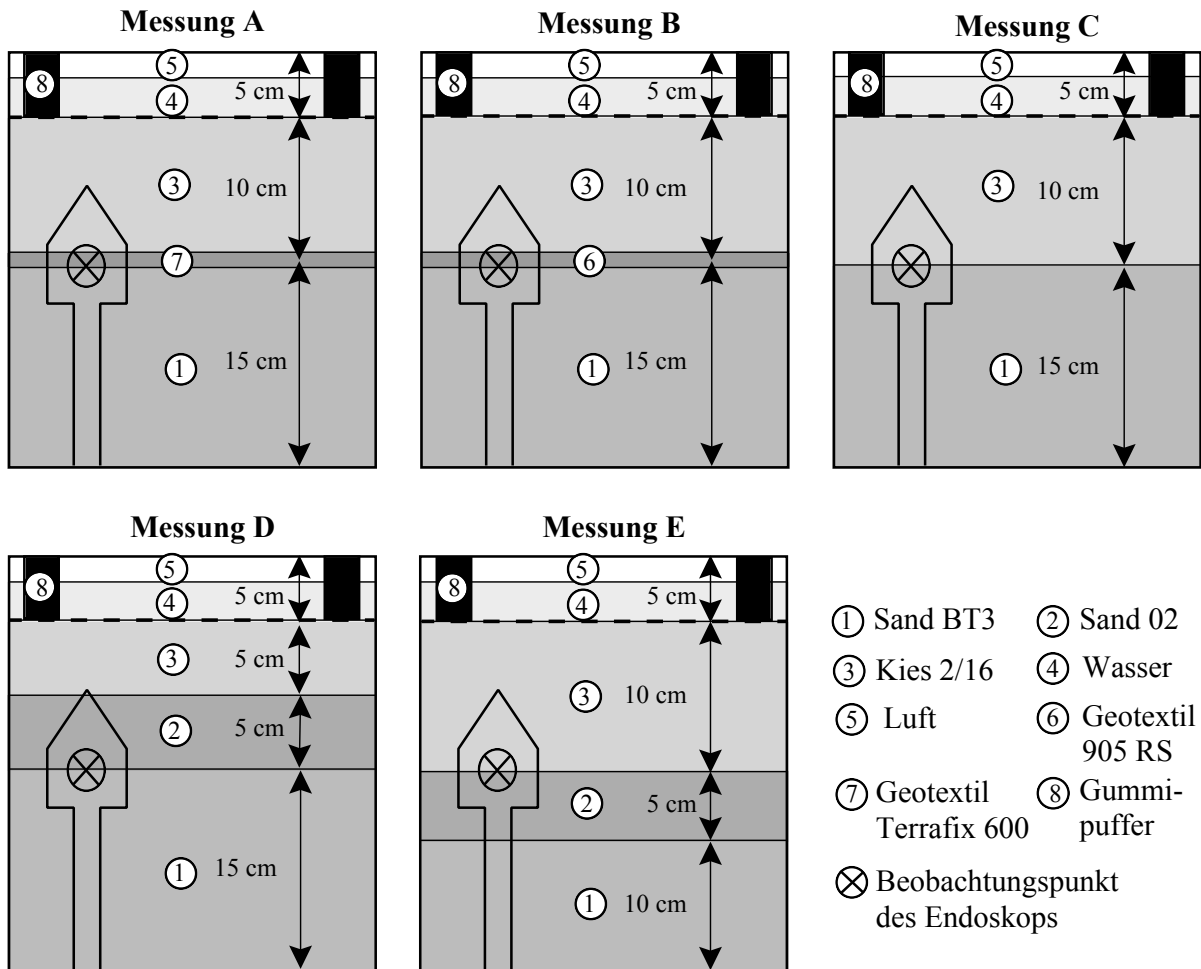


Abbildung 10.1: Schematischer Querschnitt durch die Filterschichten im Sedimenttank bei den verschiedenen Messungen. Auf der obersten Sedimentschicht liegt eine Lochplatte. Diese wird über quaderförmige Gummipuffer vom Tankdeckel auf das Sediment gepreßt. Dadurch wird ein ‚aufschwimmen‘ der gesamten Sedimentschichten bei Druckschwankungen verhindert.

Durch Druckschwankungen wird die im Boden vorhandene Luft komprimiert und dekomprimiert. Diese Volumenänderungen verursachen eine Veränderung des Auftriebs der Filterschichten im Wasser. Bei ungenügender Auflast des Sediments führt dies unweigerlich zu einem Heben und Senken der Filterschichten. Diese unnatürlichen Bewegungen werden vermieden, wenn das Sediment von oben durch eine zusätzliche Kraft belastet wird. In realen Flußufern wird diese Kraft durch die Gewichtskraft der Deckschicht erzeugt. Im Versuchstank wird eine Metallplatte mit Löchern auf die oberste Sedimentschicht aufgelegt und über vier Gummipuffer vom aufgeschraubten Deckel festgeklemmt (Abbildung 10.1). Durch diese Umklammerung des Sediments kann Bewegung nur noch durch Kornverlagerungen in vorhandenen Hohlräumen und Strukturveränderungen innerhalb des Filters verursacht werden und gibt direkt Aufschluß über das Vorhandensein solcher Instabilitäten.

10.2 Größe der Druckschwankungen

Jede der fünf Versuchsanordnungen wurde im Durchschnitt mit sechs verschiedenen Kombinationen von statischem Druck und Druckschwankung belastet. Ein typischer Verlauf dieser Druckschwankungen ist in Abbildung 9.5 dargestellt. Dabei werden jeweils der Offset P und die Amplitude der Druckschwankung ΔP verändert. Als Offset wurden die Werte 0, 400 und 800 *cm WS* eingestellt. Diesen wurden Druckschwankungen der Amplitude 100, 200 und 500 *cm WS* überlagert, wobei die Summe aus Offset und Amplitude den Berstdruck der Druckgeber (1000 *cm WS*) nicht überschreiten durften. In Tabelle 10.1 sind die verwendeten Kombinationen von statischem Druckoffset und Amplitude der dynamischen Druckänderung zusammengestellt. Diese stellen jedoch nur ungefähre Werte dar, die nicht genau eingestellt werden konnten und so bei den einzelnen Messungen leicht variierten.

| | | P | | |
|------------|-----|-----|-----|-----|
| | | 0 | 400 | 800 |
| ΔP | 100 | × | - | - |
| | 200 | × | × | × |
| | 500 | × | × | - |

Tabelle 10.1: Kombinationen von statischem Druck P und Amplitude der Druckschwankungen ΔP der Messungen. Dabei bedeutet (×) gemessen und (-) nicht gemessen.

Die Zeitdauer der Druckschwankung war bei allen Messungen gleich. Nach dem Starten der Aufnahme wurde nach ca. 1 s das Ventil betätigt und dadurch der Tankinnendruck erhöht. Nach einer Zeit $\Delta t = 5$ s wird der Druck abgelassen. Die Messung läuft danach noch ca. 5 s weiter um den Druckabfall zu verfolgen. Pro Messung wurden 300 Bilder aufgenommen und 240 davon ausgewertet. Die restlichen 60 Bilder konnten durch die Ausdehnung der zeitlichen Pyramide nicht verwendet werden. Insgesamt sind so ca. 15.000 Bilder ausgewertet worden, was einer Datenmenge von ca. 4 Gbyte entspricht.

10.3 Auswertung der Bewegungssegmentierung

Die Bewegungssegmentierung liefert für jedes einzelne Bild die Bildregionen, in denen Bewegung stattfindet. Um diese einer quantitativen Auswertung zugänglich zu machen, wird jedes segmentierte Bild *binarisiert*. Dabei wird jedem Pixel, an dem Bewegung stattfindet, der Wert 1 und jedem Pixel des Bildhintergrunds der Wert 0 zugeordnet. Die Menge der Binärbilder stellt eine dreidimensionale Orts-Zeit-Struktur dar. Um die darin enthaltene Information interpretieren zu können, werden verschiedene Größen daraus berechnet:

- **Zeitinformation:** Für jedes segmentierte Binärbild werden alle Pixel aufsummiert. Dies liefert die Anzahl der Pixel, an denen Bewegung detektiert wurde und somit die Fläche der Bildregionen, an denen Bewegung stattfindet in Abhängigkeit von der Zeit. Durch diese

Integration über die räumlichen Koordinaten geht jede Information über die räumliche Lokalisierung der Bewegung verloren. Der zeitliche Verlauf der Bewegung läßt sich jedoch direkt mit dem Verlauf der Druckschwankung vergleichen, um eine Abhängigkeit zwischen Druck und Bewegung zu erhalten.

- **Ortsinformation:** Um eine Aussage über die Lokalisierung der Bewegung machen zu können, werden zusätzlich alle zeitlich aufeinanderfolgenden Binärbilder aufsummiert. Dies entspricht einer Integration der Bewegungsinformation in Zeitrichtung und führt zu einem Verlust der zeitlichen Information zu Gunsten der räumlichen Auflösung. Man erhält dadurch ein zweidimensionales Bild, dessen Helligkeit an jedem Bildpunkt die Häufigkeit der Bewegung an diesem Punkt während des gesamten Verlaufs der Messung angibt.
- **Tiefenprofil:** Eine zusätzliche Aufaddition aller Spalten dieses Bildes liefert die Häufigkeit der Bewegung in Abhängigkeit von der Tiefe, d. h. ein über die horizontale Koordinate gemittelt *Tiefenprofil*.

Die Ergebnisse dieser Daten für die verschiedenen Messungen werden in den folgenden Abschnitten vorgestellt und interpretiert.

10.4 Abhängigkeit der Bewegung von der Druckschwankung

Abbildung 10.2 zeigt die zeitliche Abhängigkeit der Bewegung der verschiedenen Messungen (Zeit wird immer als laufende Nummer der Halbbilder ausgedrückt). Man kann deutlich erkennen, daß die Häufigkeit der Bewegung in den verschiedenen Sedimenten den gleichen qualitativen Verlauf aufweist. Zu Beginn der Druckschwankung steigt die Häufigkeit an und fällt danach deutlich ab. Beim Einsetzen des Druckabfalls steigt die Häufigkeit wieder stark an und fällt danach asymptotisch auf Null ab. Die absolute Häufigkeit ist jedoch von Sediment zu Sediment unterschiedlich. Es fällt auf, daß beim instabilen Kornfilter deutlich mehr Bewegung stattfindet als bei den stabilen Kornfiltern oder den Geotextilien. Der zeitliche Verlauf der Bewegung sowie die Tatsache, daß die absolute Häufigkeit der Bewegung mit zunehmender Amplitude der Druckschwankung deutlich zunimmt, läßt die Schlußfolgerung zu, daß die Häufigkeit der Bewegung direkt mit der Änderungsgeschwindigkeit der Druckschwankung $dP(t)/dt$ korreliert ist.

10.4.1 Modellierung der Druckabhängigkeit von Bewegung

Den Schlüssel zu einer möglichen Erklärung dieses Zusammenhangs liefert die Beobachtung von Luftblasen im Sediment. Es fällt auf, daß diese bei einer Erhöhung des Druckes komprimiert und bei einer anschließenden Druckentlastung dekomprimiert werden.

Das Volumen einer Luftblase in Wasser wird direkt durch den Umgebungsdruck der Flüssigkeit bestimmt. Der Zusammenhang zwischen dem Luftdruck P der Blase und ihrem Volumen V ergibt sich dabei in erster Näherung durch die ideale Gasgleichung:

$$PV = \nu RT = \text{const} = c, \quad (10.1)$$

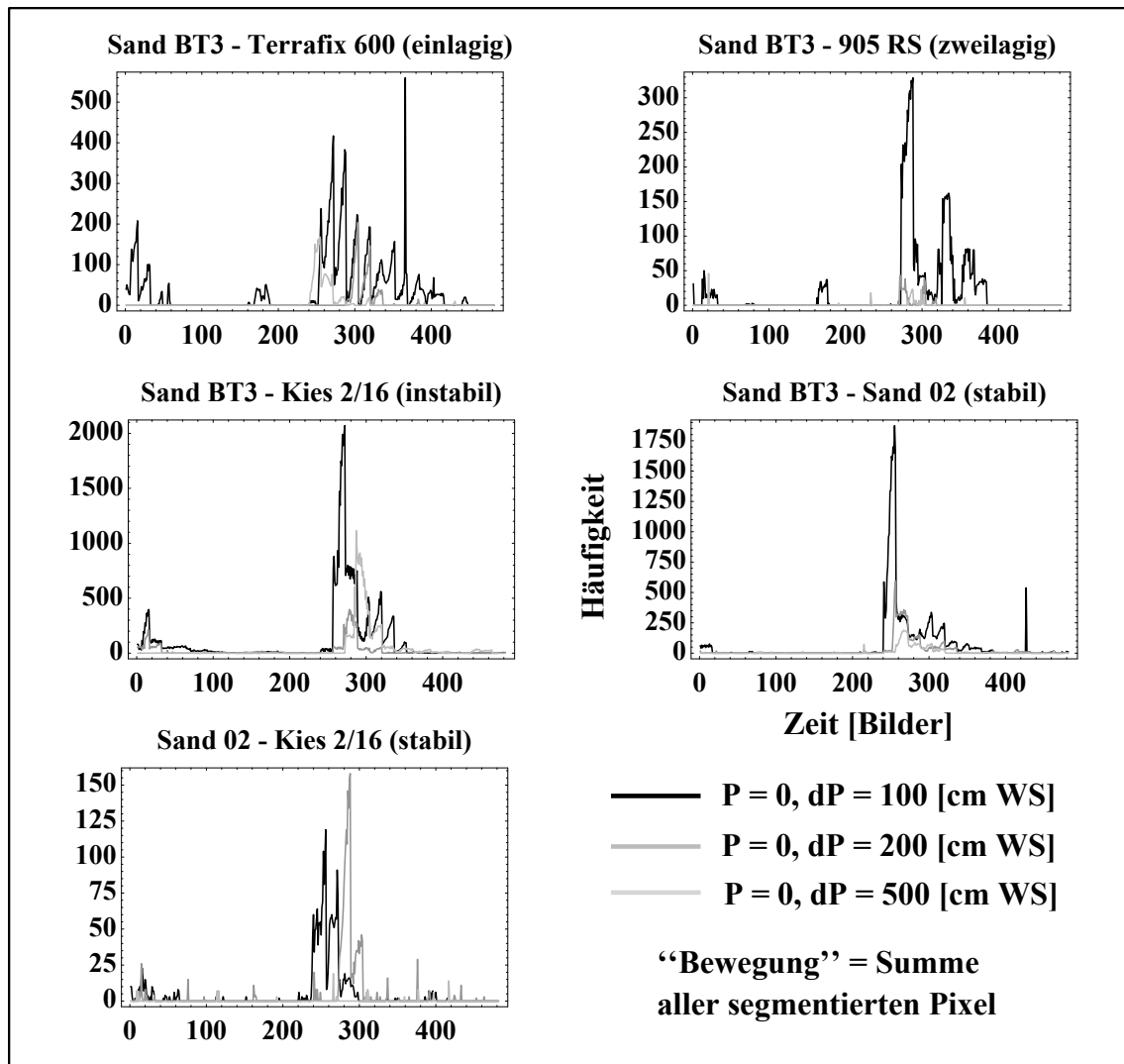


Abbildung 10.2: Zeitliche Abhängigkeit der detektierten Bewegung bei verschiedenen Filterschichten.

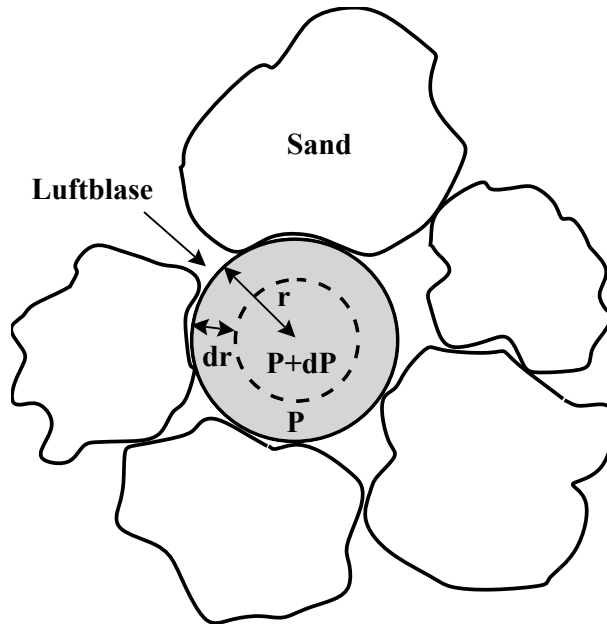


Abbildung 10.3: Veranschaulichung der Bewegung, die durch ein Komprimieren von Luftblasen erzeugt wird.

wobei die Temperatur T als konstant angenommen wird. Zur Berechnung der Größenänderung beim Komprimieren einer Luftblase durch eine infinitesimale Änderung dP des Drucks, wird das Volumen in einer Taylorreihe entwickelt:

$$V(P + dP) = V(P) - \frac{c}{P^2} dP + o(dP^2). \quad (10.2)$$

Bricht man diese Entwicklung nach der ersten Ordnung ab, so ergibt sich der Zusammenhang

$$dV = -\frac{c}{P^2} dP \quad (10.3)$$

und damit die zeitliche Änderung

$$\frac{dV(t)}{dt} = -\frac{c}{P^2(t)} \frac{dP(t)}{dt}. \quad (10.4)$$

Gleichung (10.4) stellt einen Zusammenhang zwischen der zeitlichen Änderung des Blasenvolumens und der zeitlichen Änderung des Luftdrucks her. Die Bewegungssegmentierung detektiert jedoch nicht diese Volumenänderung, sondern deren Projektion auf die zweidimensionale Bildebene, d. h. eine Flächenänderung $M(t)$ der Blasenquerschnittsfläche F , mit

$$M(t) = \frac{dF(t)}{dt}. \quad (10.5)$$

Eine Änderung des Blasenradius r um den Betrag dr (Abb. 10.3) verursacht eine Volumenänderung $dV = 4\pi r^2 dr$ und eine Änderung des Blasenquerschnitts $dF = 2\pi r dr$. Aus dem Verhältnis zwischen dV und dF ergibt sich der Zusammenhang:

$$dF = \frac{dV}{2r}. \quad (10.6)$$

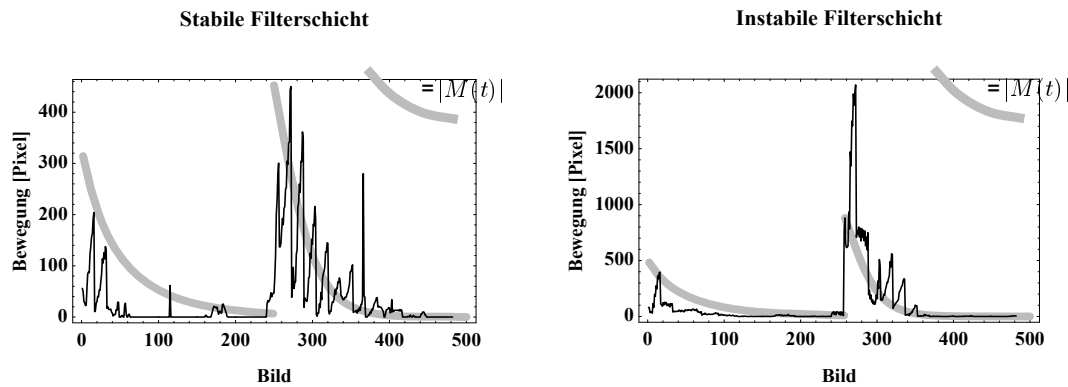


Abbildung 10.4: Vergleich des gemessenen zeitlichen Verlaufs der Bewegungshäufigkeit im Vergleich zum erwarteten Verlauf für reine Volumenänderung von Luftblasen.

Den Blasenradius r bei einem bestimmten Druck P liefert die ideale Gasgleichung (10.1):

$$r = \left(\frac{3c}{4\pi} \right)^{\left(\frac{1}{3}\right)} P^{-\left(\frac{1}{3}\right)}. \quad (10.7)$$

Zusammen mit (10.6) und (10.7) erhält man schließlich aus (10.4) die Fläche der Bewegung $|M(t)|$:

$$|M(t)| = \frac{c}{2} \left(\frac{4\pi}{3c} \right)^{\left(\frac{1}{3}\right)} P^{-\left(\frac{1}{3}\right)}(t) \frac{dP(t)}{dt}. \quad (10.8)$$

Wenn die gesamte Bewegung einzig durch die Volumenänderung der im Sediment eingeschlossenen Luftblasen verursacht würde, sollte sich der zeitliche Verlauf der Häufigkeit von Bewegung qualitativ durch den Zusammenhang (10.8) beschreiben lassen. Die Tatsache, daß über die gesamte Bildfläche integriert wird, ändert nichts an dieser Proportionalität, wenn sich die Gesamtanzahl der Blasen im Bildausschnitt nicht ändert. Selbst die Bewegung durch nachrutschende Sandteilchen, die das entstehende Volumen ausfüllen, sollte proportional zu $M(t)$ sein.

Abbildung 10.4 zeigt den zeitlichen Verlauf der Bewegung für zwei verschiedene Filter zusammen mit dem erwarteten Verlauf $M(t)$ für eine reine Volumenänderung durch Luftblasen. Für die stabile Filterschicht (Geotextil, Messung A) stimmt der erwartete Verlauf gut mit dem gemessenen Verlauf überein. Das Verhältnis der Bewegungshäufigkeit beim Druckanstieg zu der beim Druckabfall ist direkt proportional zu dem bereits in Abschnitt 9.3.3 beschriebenen Unterschied zwischen der Änderungsgeschwindigkeit des Druckanstiegs zu der des Druckabfalls. Dies deutet darauf hin, daß das Filter tatsächlich stabil ist und somit keine Partikelverlagerungen stattfinden können. Die Bildverarbeitung hat demnach die Größenänderung der Luftblasen segmentiert.

Bei der instabilen Filterschicht (Kornfilter, Messung C) gilt dies für den qualitativen Verlauf des langsamen Abfalls der Bewegungshäufigkeit. Kurz nach Beginn des Druckabfalls bei der schnellsten Druckänderung, zeigt der gemessene Verlauf jedoch eine deutliche Überhöhung gegenüber dem erwarteten Verlauf. Das Maximum des Peaks liegt doppelt so hoch wie der erwartete Wert. Dies deutet darauf hin, daß sich bei diesem Filter mehr bewegt als nur Luftblasen, die ihre Größe

ändern. Durch die schnelle Druckänderung werden Teilchen losgerissen, die sich jedoch nur dann bewegen können, wenn in der Grenzschicht Hohlräume vorhanden sind. Dies ist ein deutliches Zeichen für eine mögliche Instabilität des Filters.

10.5 Räumliche Lokalisierung der Bewegung

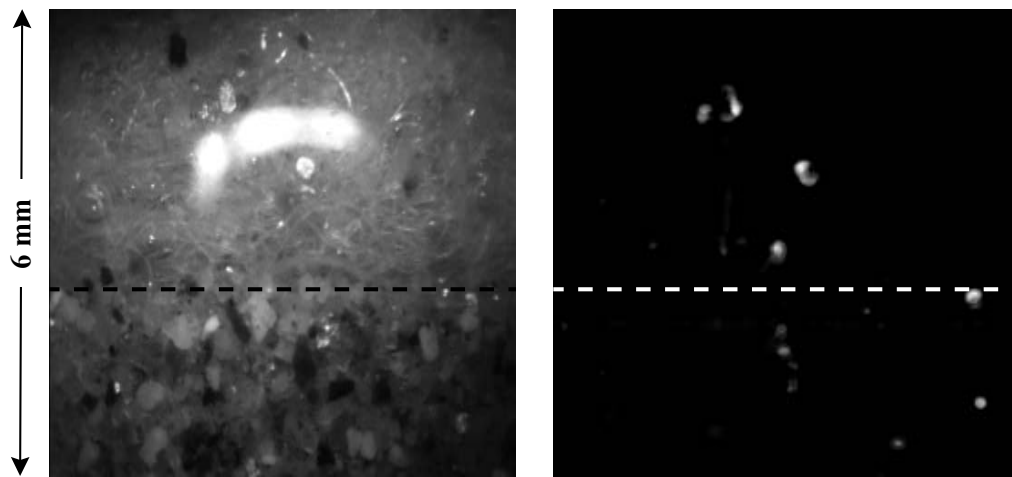
Während die Überlegungen des vorangegangenen Abschnitts Rückschlüsse über die möglichen physikalischen Vorgänge im Sediment zulassen, geben sie keinerlei Aussage über die Lokalisierung der gemessenen Bewegung innerhalb der Grenzschicht. Die Abbildungen 10.5 bis 10.9 zeigen die gemessenen räumlichen Häufigkeitsverteilungen der Bewegung und die daraus erhaltenen Tiefenprofile der Bewegungshäufigkeit für die fünf Messungen.

Das Bild in der linken oberen Ecke zeigt jeweils ein Originalbild der Grenzschicht. Die Spiegelreflexe der Beleuchtung an der Glasscheibe des Endoskopaufsatzes sind deutlich zu erkennen.

Rechts oben ist die räumliche Häufigkeitsverteilung abgebildet. Zur besseren Darstellung sind die Bilder kontrastverstärkt. Für alle Messungen wurde die Häufigkeitsverteilung für einen statischen Druck von $P = 0 \text{ cmWS}$ und eine Druckschwankung $\Delta P = 500 \text{ cmWS}$ dargestellt. Bei den Geotextilien und den stabilen Filterschichten fällt auf, daß die Bewegung nicht signifikant um die Grenzschicht herum lokalisiert ist. Insgesamt bewegen sich in diesen Filtern nur kleinere Bildbereiche. Im Gegensatz dazu tritt bei der instabilen Filterschicht (Messung C) Bewegung innerhalb eines großen, zusammenhängenden Gebiets direkt unterhalb der Grenzschicht auf.

Die Plots zeigen Tiefenprofile der Häufigkeitsverteilung für verschiedene Kombinationen von statischem Druck und Amplitude der Druckschwankung. Für die absolute Häufigkeit erkennt man folgenden Zusammenhang: Bei gleichem statischem Druck ist die Häufigkeit der Bewegung proportional zur Amplitude der Druckschwankung. Dieser Zusammenhang war schon in der Zeitabhängigkeit erkennbar. Je größer jedoch der statische Druck ist, desto weniger bewegt sich bei den gleichen Druckschwankungen. Zusammen mit den Überlegungen des vorangegangenen Kapitels ist dies ein Indiz dafür, daß die Bewegung hauptsächlich durch Luftblasen hervorgerufen wird. Diese sind bei höherem statischen Druck bereits abgesättigt, so daß sich ihr Volumen bei gleicher Druckschwankung weniger stark ändert als bei einem niedrigen statischen Druck (Vergleich zusammengedrückte Feder zu entspannter Feder).

Beim Vergleich der Tiefenprofile verschiedener Filterschichten erkennt man außerdem, daß es bei den stabilen Filtern keine signifikante Anhäufung der Bewegung im Bereich der Grenzschicht gibt. Anders bei der instabilen Filterschicht: Dort findet die Bewegung deutlich in der näheren Umgebung der Grenzschicht statt. Im Originalbild erkennt man an diesen Stellen Hohlräume, in denen Bewegung und damit das im vorangegangenen Kapitel postulierte Losreißen der Sandkörner stattfinden kann. Insgesamt scheinen die Geotextilien am stabilsten gegen Kornverlagerungen zu sein. Obwohl die Kornfilter insgesamt auch relativ stabil sind, tritt bei ihnen jedoch ein weiterer Effekt auf: Während bei kleinen Druckschwankungen die Bewegung weit unterhalb der Grenzschicht stattfindet, scheint sich die Bewegung bei großen Amplituden der Druckänderung in die Grenzschicht zu verlagern. Das könnte bedeuten, daß diese Filter gegen kleine Druckänderungen resistent sind, bei großen Druckschwankungen jedoch Umlagerungen innerhalb der Grenzschicht stattfinden.



MESSUNG A: Sand BT3 - Terrafix 600 (einlagig)

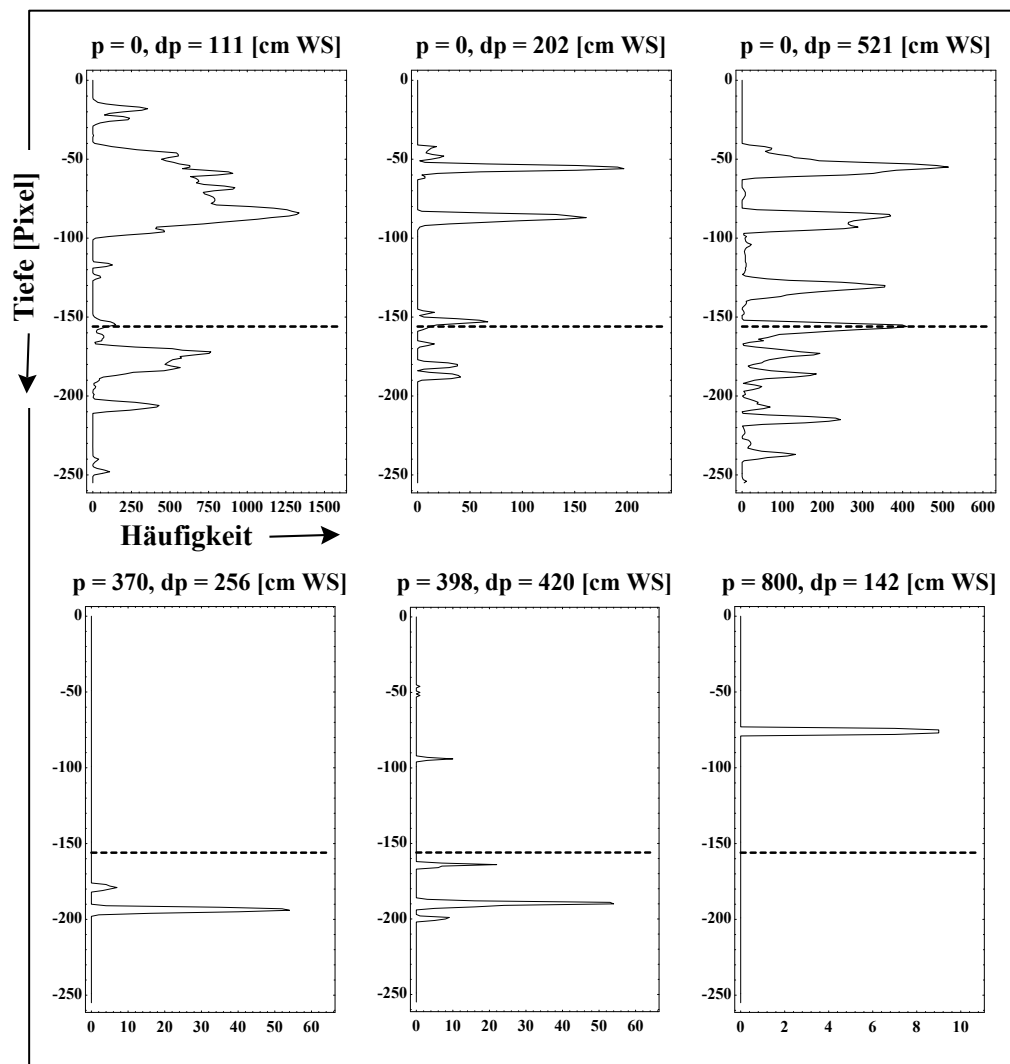
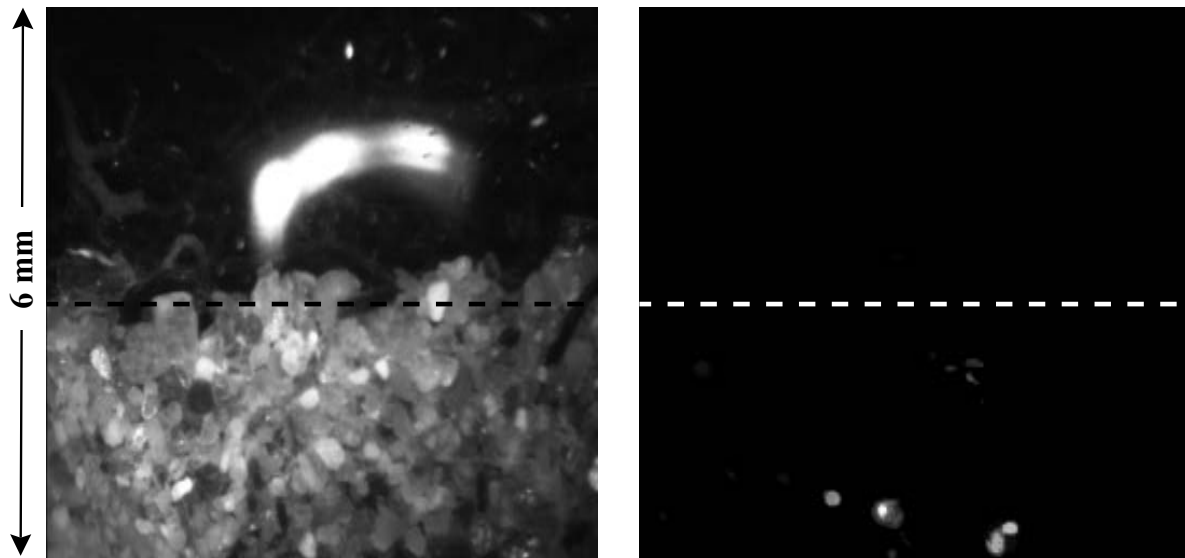


Abbildung 10.5: Räumliche Lokalisierung der Bewegung und Tiefenprofile der Bewegungshäufigkeit.



Messung B: Sand BT3 - 905 RS (zweilagig)

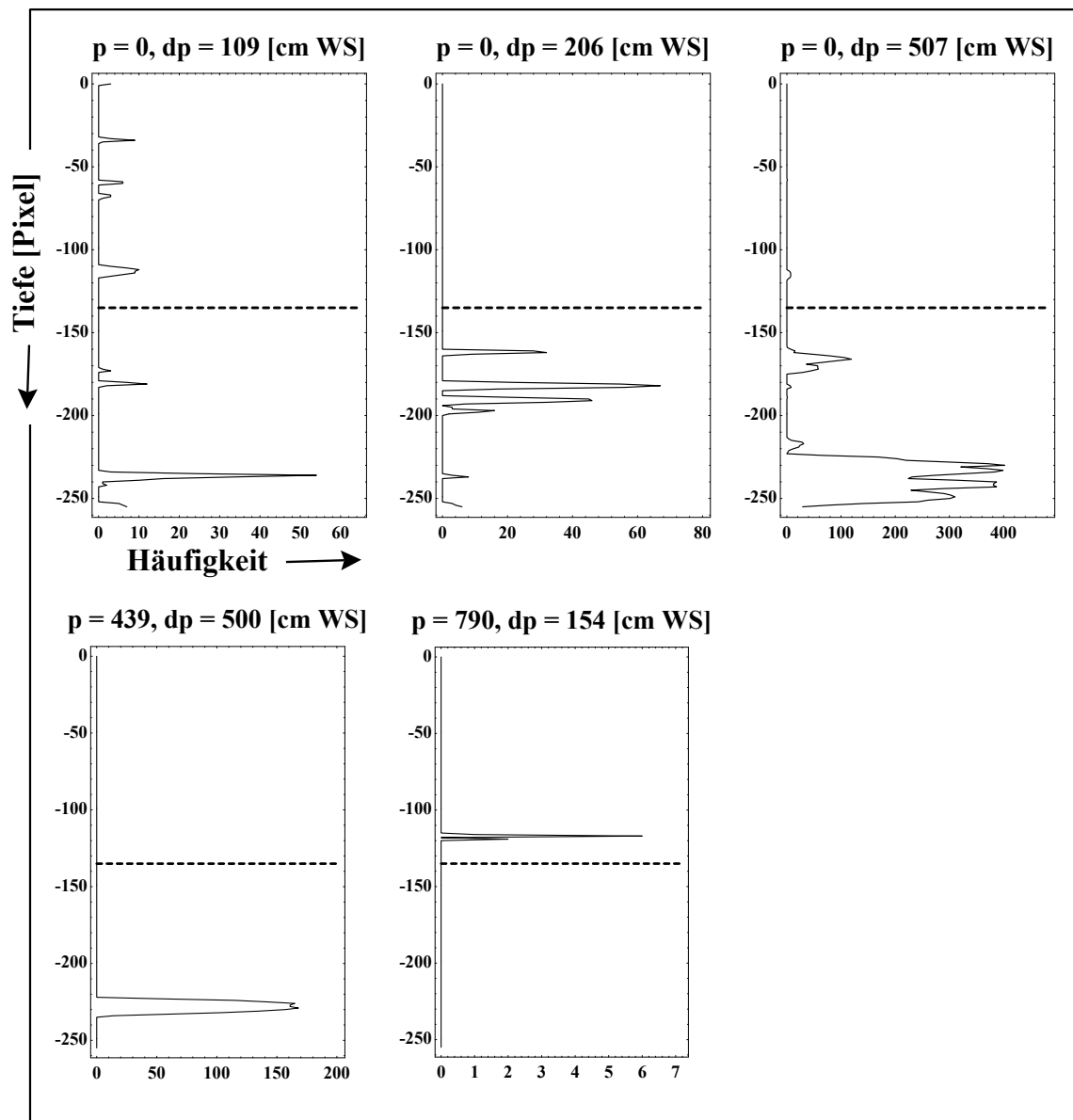
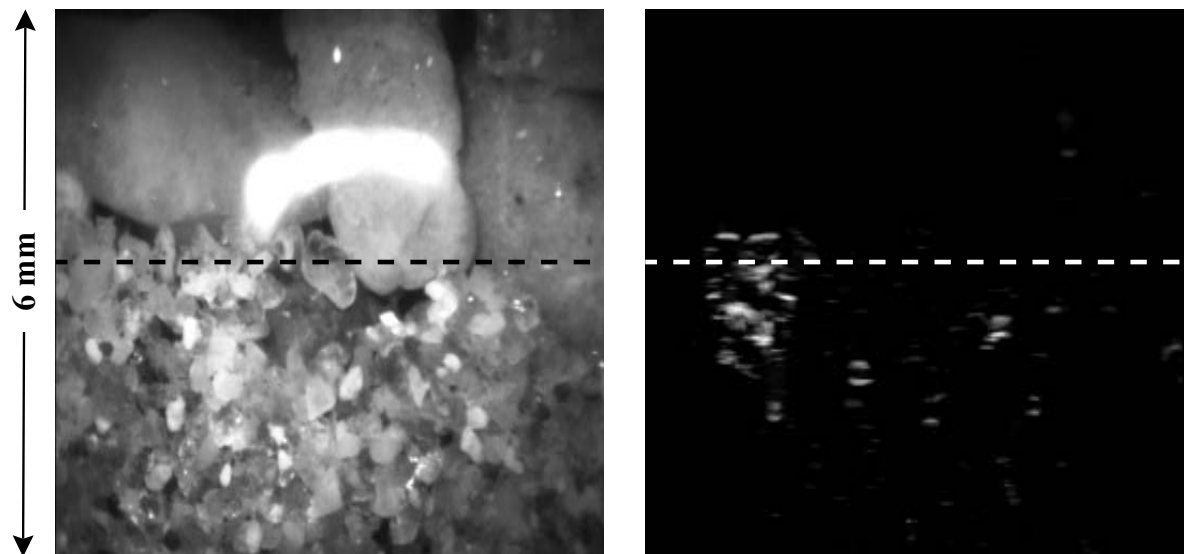


Abbildung 10.6: Räumliche Lokalisierung der Bewegung und Tiefenprofile der Bewegungshäufigkeit.



MESSUNG C: Sand BT3 - Kies 2/16 (instabil)

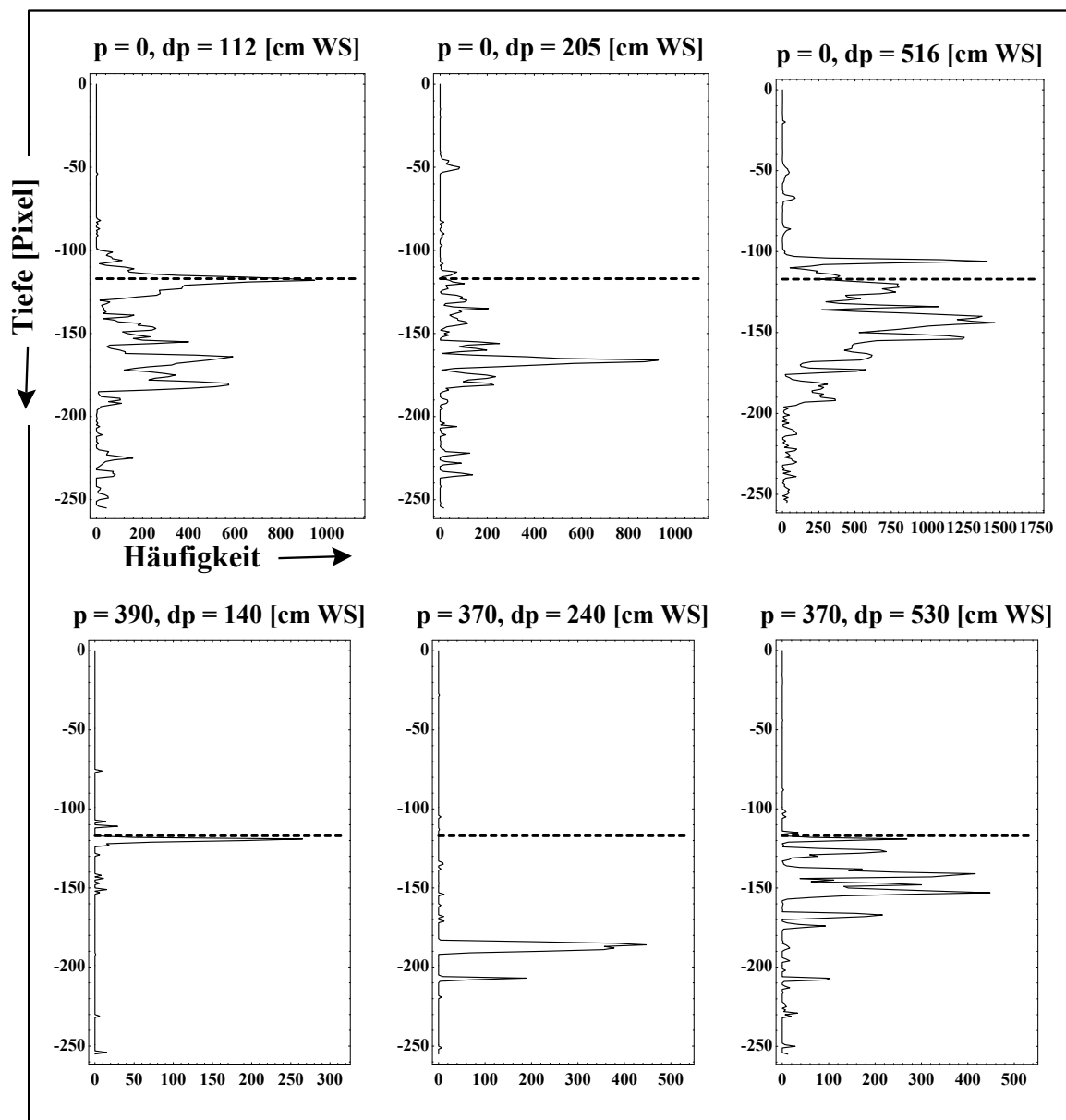
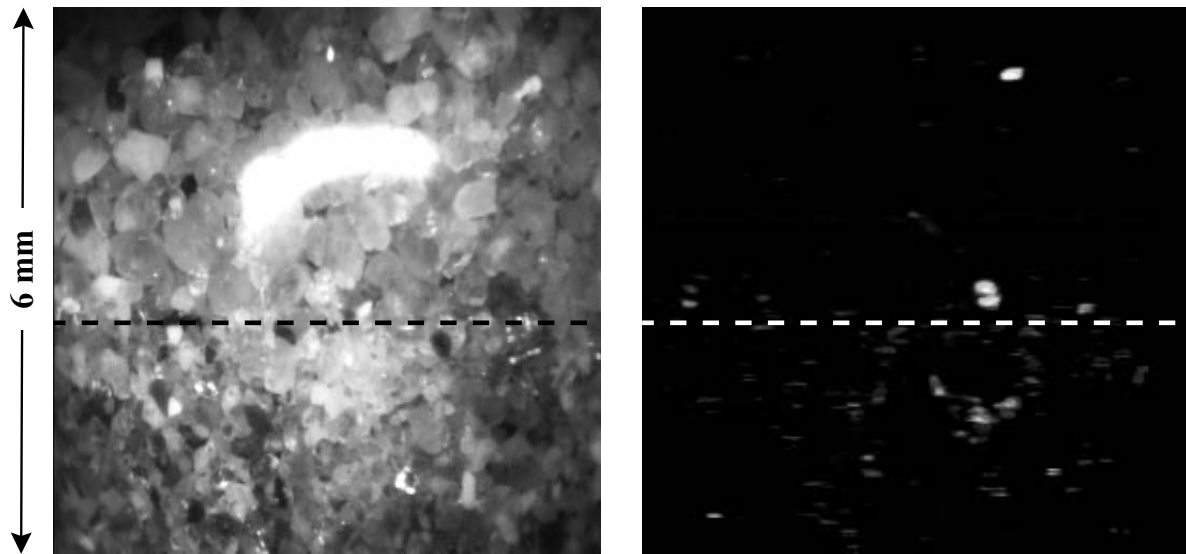


Abbildung 10.7: Räumliche Lokalisierung der Bewegung und Tiefenprofile der Bewegungshäufigkeit.



MESSUNG D: Sand BT3 - Sand 02 (stabil)

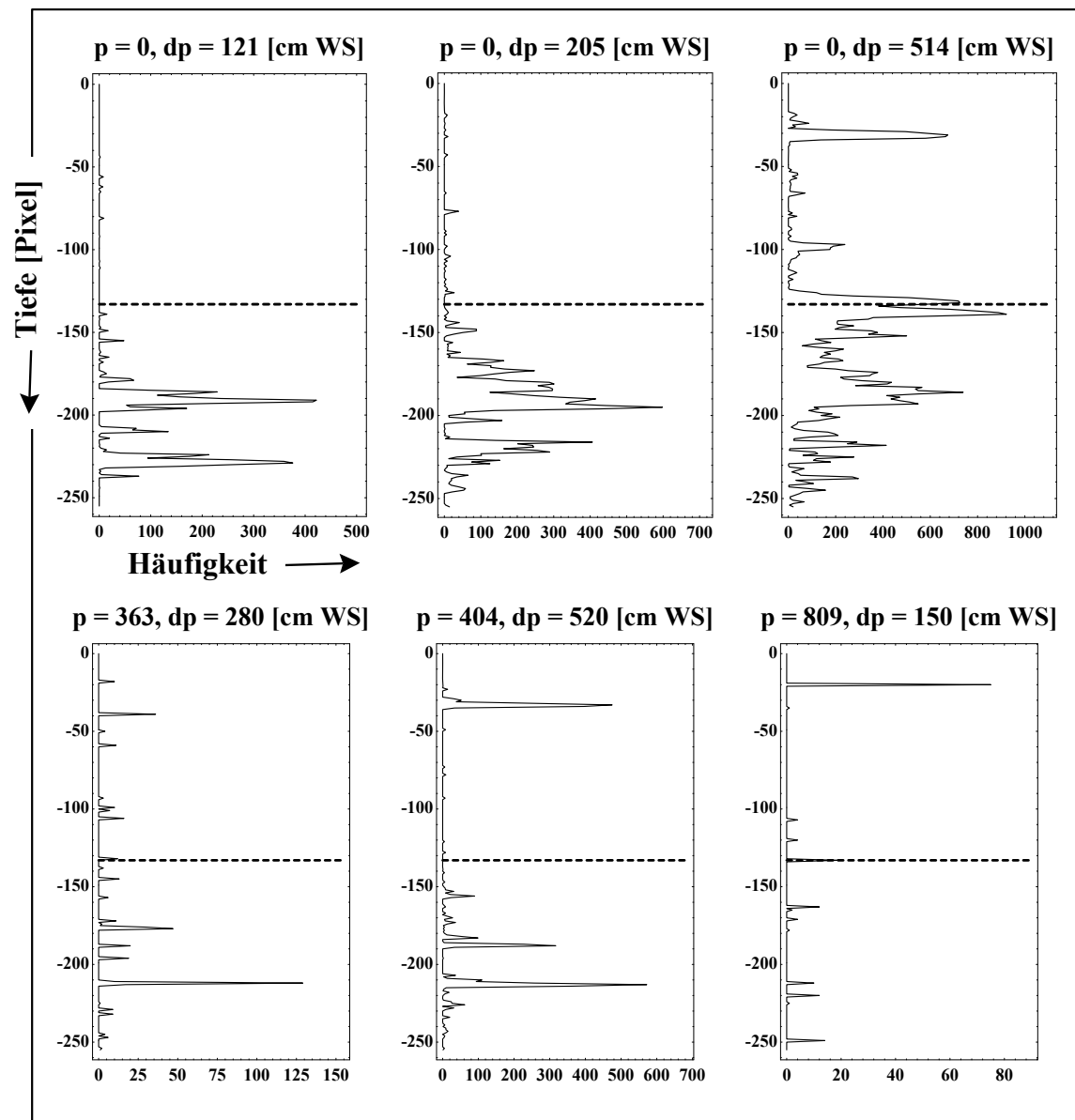
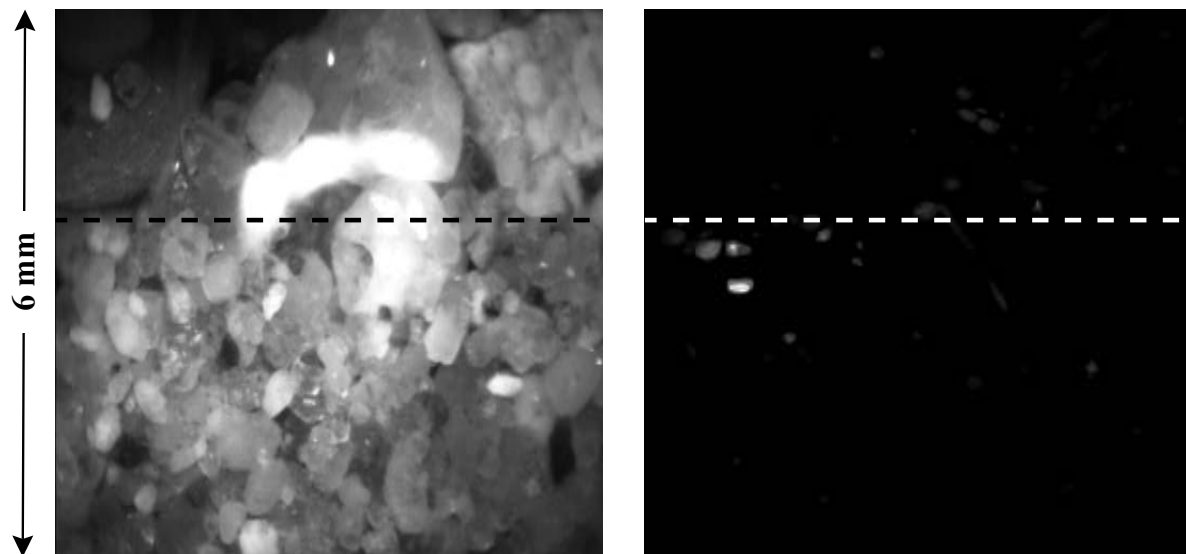


Abbildung 10.8: Räumliche Lokalisierung der Bewegung und Tiefenprofile der Bewegungshäufigkeit.



MESSUNG E: Sand 02 - Kies 2/16 (stabil)

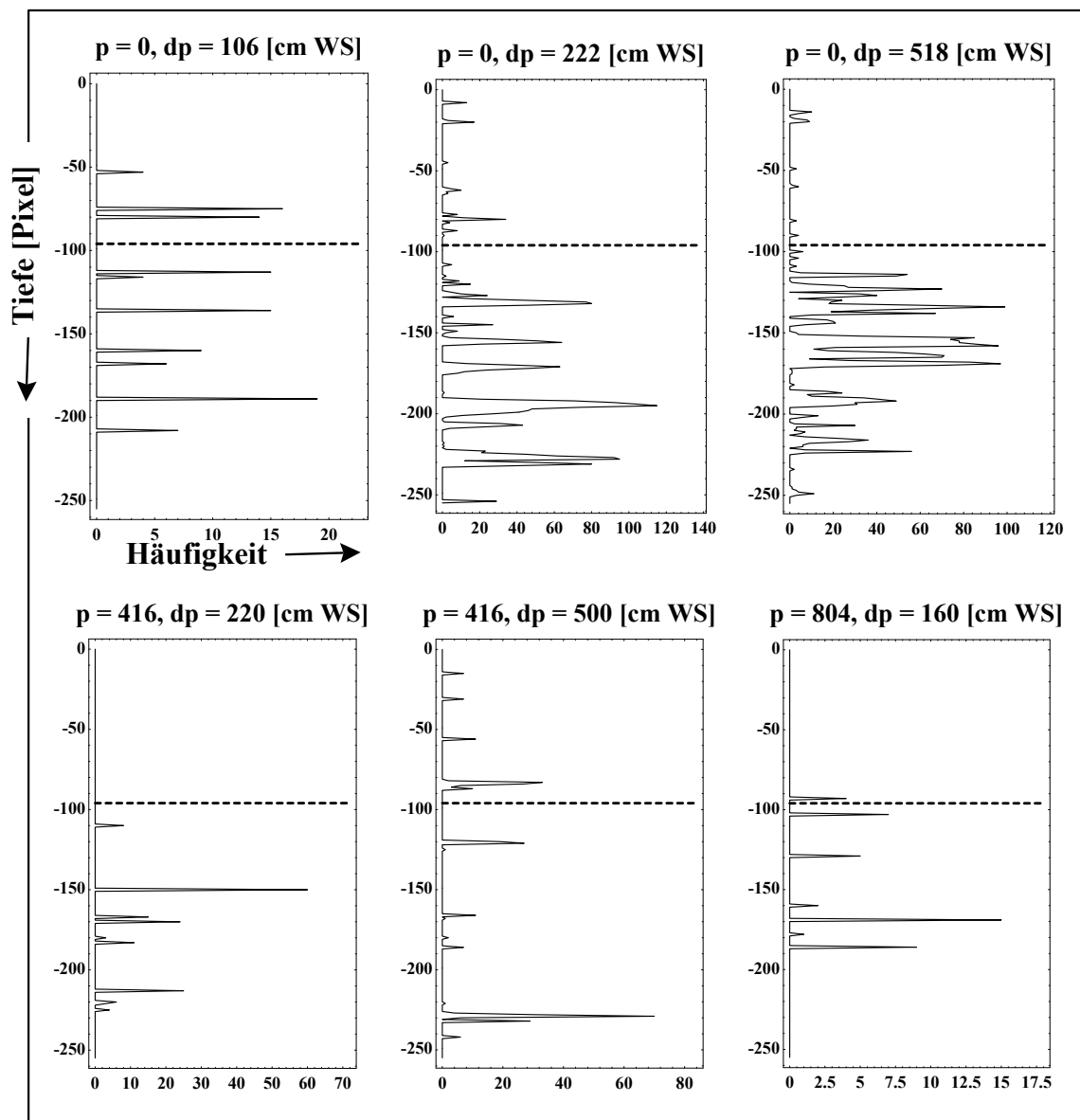


Abbildung 10.9: Räumliche Lokalisierung der Bewegung und Tiefenprofile der Bewegungshäufigkeit.

Kapitel 11

Ausblick

Die Ergebnisse der Messung zeigen, daß Kornverlagerungen in Sedimenten detektiert und Rückschlüsse über die physikalischen Vorgänge in den Grenzschichten gezogen werden können. Um eine bessere Statistik zu erhalten, sind weitere Messungen nötig, die in einem größeren Sedimenttank die Simulation realistischer Vorgänge erlauben. Zusätzlich zu den reinen Filterschichten können dann auch Grenzen zwischen Sedimenten und festen Bauteilen (z. B. Wasserrohre) untersucht werden.

In der Eigenschaft, selbst sehr langsame, zeitliche Änderungen des Bildinhaltes zu detektieren, ist das in dieser Arbeit entwickelte Verfahren nicht alleine auf die automatische Erkennung von Bewegung beschränkt. Es läßt sich gleichermaßen für alle Anwendungen einsetzen, in denen eine zeitliche Strukturveränderung des Bildinhaltes detektiert werden soll. Der betrachtete Anwendungsfall der Detektion von Sedimentverlagerungen ist dabei nur eine von vielen Möglichkeiten. Die speziellen Probleme dieser Art des Bildmaterials führten dazu, ein Verfahren zu entwickeln, das über die herkömmlichen Methoden der Detektion von Bewegung hinaus geht. Dies eröffnet die Möglichkeit auch andere Vorgänge zu untersuchen, bei denen langsame und großflächige Umwandlungen auftreten und lokalisiert werden müssen. Dazu gehören wichtige physikalische Vorgänge, wie

- Schmelzprozesse,
- Kristallisation und
- Phasenübergänge.

Bei Anwendungen, die das zweidimensionale Wachstum einer Struktur beobachten, bieten die Eigenschaften des Algorithmus noch weitere Anwendungsmöglichkeiten: Durch Verwenden sehr kleiner räumlicher und zeitlicher Faltungsmasken werden nur schnelle Veränderungen detektiert, die räumlich stark lokalisiert sind (Rand der betroffenen Region). Mit großen raum-zeitlichen Umgebungen werden langsame Änderungen detektiert und dabei über die gesamte Fläche gemittelt, an der während dieser Zeit Bewegung auftritt. Ein Vergleich der Fläche der momentanen Bewegung mit dem gesamten Gebiet könnte direkt die fraktale Struktur eines solchen Wachstumsprozesses liefern.

Anhang A

Rekursives Filter

Ein weiteres Verfahren zur Realisierung eines zeitlichen Tiefpaßfilters und somit zur Bestimmung des ruhenden Hintergrundes ergab sich durch eine effektive Ausnutzung der Bildverarbeitungs-Hardware. Die verwendete Bildverarbeitungskarte¹ bietet die Möglichkeit, das Bild, das sich bereits im Bildspeicher befindet, durch Rückkopplung zum nächsten Bild hinzuzuaddieren. Dadurch können Bildfolgen im Videotakt aufaddiert werden. Der Inhalt des Bildspeichers wird dabei mit einem Faktor α ($0 \leq \alpha \leq 1$) gewichtet und das neu ankommende Bild mit dem Faktor $(1 - \alpha)$. Dies gewährleistet, daß zu jedem Zeitpunkt t_0 die Wichtungsfaktoren aller bereits aufaddierten Bilder auf 1 normiert sind. Die dadurch resultierende Faltungsmaske $w(t)$ ist eine in der negativen Zeitachse abfallende Exponentialfunktion. Für $t > 0$ ist sie identisch 0. Um eine symmetrische Faltungsmaske ohne Unstetigkeit bei 0 zu erhalten, wird zur Faltung in der Zeit diese *rekursive Filterung* zwei mal angewendet:

Zuerst läßt man das rekursive Filter von einer Zeit $(t_0 - \Delta t)$ bis t_0 laufen. Danach wird dies von $(t_0 + \Delta t)$ bis t_0 in umgekehrter Zeitrichtung wiederholt. Anschließend werden beide Teil-Filterantworten addiert und auf 1 normiert. Das Ergebnisbild enthält den Hintergrund des Bildes zur Zeit t_0 . Man benötigt hierzu ein Speichermedium (wie z. B. die LVR²), mit dem man die Möglichkeit hat, Bilder im Videotakt rückwärts abzuspielen. Die daraus resultierende effektive Faltungsmaske hat die Form

$$w(t) = \frac{(1 - \alpha)}{2} e^{\ln \alpha \|t\|}, \quad (\text{A.1})$$

und damit als Transferfunktion die *Lorentzkurve*

$$\hat{w}(\omega) = \frac{(\alpha - 1)}{2\pi} \frac{\ln \alpha}{(\ln \alpha)^2 + \omega^2}. \quad (\text{A.2})$$

Sie fällt wie ω^{-2} ab und nähert sich für $\alpha \rightarrow 1$ einer δ -Ebene im Fourierraum an. In Abb. A.1 sind die Faltungsmaske und die dazugehörige Transferfunktion für verschiedene Rückkopplungsparameter α dargestellt.

Die Tatsache, daß die Transferfunktion eine Lorentzkurve darstellt, zeigt deutlich, daß es sich bei der Tiefpaßfilterung um ein typisches *Resonanzphänomen* handelt. Dabei bleibt nur die Frequenz $\omega = 0$ unverändert erhalten, wobei alle anderen Frequenzen stark unterdrückt werden.

¹ITI FG100 Frame grabber der Firma Imaging Technologies

²Sony LVR-5000 Laser Video Recorder.

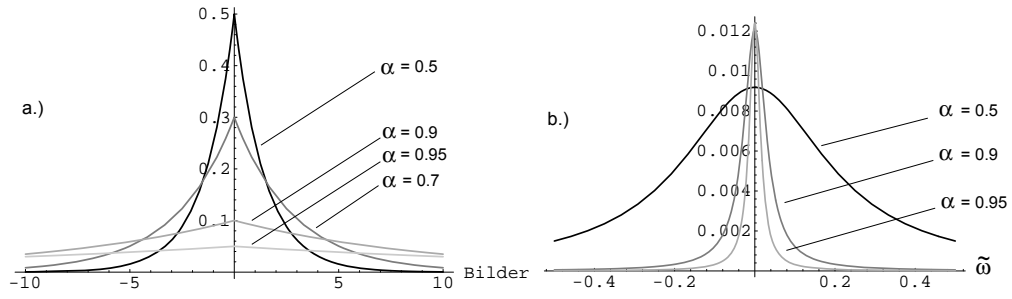


Abbildung A.1: *Effektive Faltungsmaske in der Zeit a) und Transferfunktion b) des rekursiven Filters.*

Die gesamte rekursive Filterung entspricht somit der Faltung

$$H(\vec{x}, t_0) = \frac{(1 - \alpha)}{2} \int_{t_0 - \Delta t}^{t_0 + \Delta t} e^{\ln \alpha \|t_0 - t\|} g(\vec{x}, t) dt. \quad (\text{A.3})$$

Anhang B

Rauschen

In diesem Abschnitt soll der Erwartungswert des Ähnlichkeitsmaßes S (Abschnitt 5.2) für Rauschen hergeleitet werden. Betrachtet man den Grauwert eines einzelnen Pixels als *Zufallsvariable*, dann kann Rauschen im allgemeinen als additive und mittelwertfreie Abweichung vom exakten Wert mit Standardabweichung σ_n beschrieben werden.

Sei $b(\vec{x}, t)$ der exakte Grauwertverlauf der Bildsequenz und $n(\vec{x}, t)$ die additive Störung durch Rauschen. Der gemessene Grauwert $g(\vec{x}, t)$ eines Bildpunktes ergibt sich damit zu

$$g(\vec{x}, t) = b(\vec{x}, t) + n(\vec{x}, t). \quad (\text{B.1})$$

Für normalverteiltes, mittelwertfreies Rauschen ist die Häufigkeitsverteilung $p(n)$ von $n(\vec{x}, t)$ durch die diskrete Approximation der Gaußschen Verteilungsfunktion

$$p(n) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left[-\frac{n^2}{2\sigma_n^2}\right] \quad (\text{B.2})$$

gegeben (Abb B.1). Eine Glättung der Bildsequenz mit einer hinreichend großen Maske¹ liefert den *Erwartungswert* $E(g)$

$$E(g) = \int_{-\infty}^{\infty} dg p(g) g \quad (\text{B.3})$$

des Bildsignals. Für isotropes Rauschen erhält man mit (B.2):

$$\langle n \rangle_{\vec{x}} = \langle n \rangle_t = E(n) = 0. \quad (\text{B.4})$$

Der tiefpaßgefilterte ruhende Hintergrund $g_t(\vec{x}, t)$ der Bildsequenz kann somit in guter Näherung als frei von Rauschen angenommen werden, d. h. $g_t(\vec{x}, t) = b_t(\vec{x}, t)$. Die Differenz $L(\vec{x}, t)$ zwischen der Originalsequenz und dem ruhenden Hintergrund enthält somit unverändert das Rauschen der Originalsequenz:

$$L(\vec{x}, t) = g(\vec{x}, t) - g_t(\vec{x}, t) = \underbrace{b(\vec{x}, t) - b_t(\vec{x}, t)}_{l(\vec{x}, t)} + n(\vec{x}, t). \quad (\text{B.5})$$

¹Als hinreichend groß gilt eine Glättungsmaske, wenn ihre effektive Reichweite deutlich größer als die raumzeitliche Kohärenz des Rauschens ist. Für isotropes, normalverteiltes Rauschen ist dies schon für kleine Masken erfüllt.

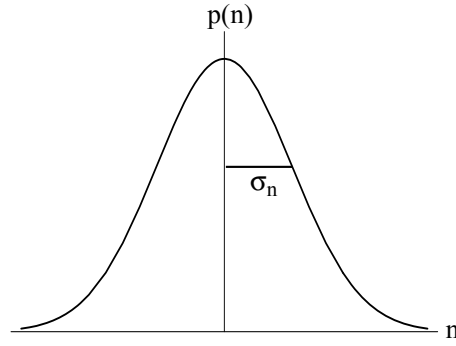


Abbildung B.1: Häufigkeitsverteilung der additiven Grauwertschwankungen für normalverteiltes, mittelwertfreies Rauschen.

Dies entspricht der Tatsache, daß die Berechnung von L eine ideale Hochpaßfilterung darstellt und deshalb Rauschen als hochfrequente Struktur unverändert erhält. Für das Ähnlichkeitsmaß S (5.19) gilt dann:

$$\begin{aligned}
 S(\vec{x}, t) &= \mathcal{B}_{\vec{x}}^m [l(\vec{x}, t) + n(\vec{x}, t)]^2 \\
 &= \langle [l(\vec{x}, t) + n(\vec{x}, t)]^2 \rangle_{\vec{x}} \\
 &= \underbrace{\langle l^2(\vec{x}, t) \rangle_{\vec{x}}}_{S_o(\vec{x}, t)} + 2 \underbrace{\langle l(\vec{x}, t) n(\vec{x}, t) \rangle_{\vec{x}}}_o + \underbrace{\langle n^2(\vec{x}, t) \rangle_{\vec{x}}}_{E(n^2)}.
 \end{aligned} \tag{B.6}$$

Der erste Term enthält keinen Rauschanteil und liefert das Ähnlichkeitsmaß $S_o(\vec{x}, t)$ für völlig rauschfreie Bildsequenzen. Der zweite Term stellt ein Integral über eine schnell oszillierende Funktion dar und ist deshalb identisch Null. Der dritte Term ist unabhängig vom Bildinhalt und hängt nur vom Rauschen ab. Für mittelwertfreies Rauschen ist dies per Definition die Varianz des Rauschens:

$$E(n^2) = \sigma_n^2. \tag{B.7}$$

Dies liefert das wichtige Ergebnis: *Mittelwertfreies, normalverteiltes Rauschen liefert einen konstanten Beitrag zum Ähnlichkeitsmaß, dessen Wert durch die Varianz des Rauschens gegeben ist:*

$$S(\vec{x}, t) = S_o(\vec{x}, t) + \sigma_n^2. \tag{B.8}$$

Sehr schnelle Teilchen, die keinen räumlichen Überlapp zwischen zwei aufeinanderfolgenden Bildern haben, stellen zeitliches Rauschen dar. Im Gegensatz zu reinem Pixelrauschen ist hierbei der räumliche Mittelwert in einem Bild (erster Term in (B.8)) ungleich Null. Dieser Offset ermöglicht es, sehr schnelle Objekte vom Pixelrauschen zu trennen.

Anhang C

Varianz von Faltungsmasken

C.1 Skalieren von Masken

Die Varianz $\sigma_{n_i}^2$ einer diskreten Verteilung von Zufallsvariablen $n_i, i \in \mathbb{Z}$ ist definiert als

$$\sigma_{n_i}^2 = \langle (n_i - \langle n_i \rangle)^2 \rangle = \sum_{i \in \mathbb{Z}} p(n_i) (n_i - \langle n_i \rangle)^2, \quad (\text{C.1})$$

wobei $p(n_i)$ die Häufigkeitsverteilung von n_i darstellt. Sie beschreibt die mittlere quadratische Abweichung der Meßwerte von ihrem Mittelwert. Bei diskreten Faltungsmasken h_i liefert die Varianz ein Maß für die effektive Reichweite der Maske. Die Zufallsvariable n_i wird dabei durch die Position i der Filterkoeffizienten innerhalb der Maske ersetzt. Die diskrete Häufigkeitsverteilung ist durch die Filterkoeffizienten h_i gegeben. Aus (C.1) wird dann:

$$\sigma_{h_i}^2 = \sum_{i \in \text{Maske}} h_i (i - \langle i \rangle)^2. \quad (\text{C.2})$$

Für symmetrische Filtermasken der Größe $(2N + 1)$ gilt:

$$\langle i \rangle = 0 \quad \text{und} \quad h_{-i} = h_i.$$

Damit reduziert sich (C.2) auf

$$\sigma_{h_i}^2 = 2 \sum_{i=1}^N h_i i^2.$$

Wird die Filtermaske h_i auf einem Gitter mit doppelter Gitterkonstante angewendet, so ergibt sich für die effektive Varianz $\sigma_{h_{2i}}^2$ des Filters h_{2i} :

$$\sigma_{h_{2i}}^2 = 2 \sum_{i=1}^N h_i (2i)^2 = 4 \sigma_{h_i}^2. \quad (\text{C.3})$$

C.2 Kaskadierung von Faltungen

Das hintereinander Ausführen von zwei Glättungen mit Binomialmasken entspricht einer Glättung mit einer größeren Faltungsmaske. Diese entsteht durch die Faltung der beiden kleinen

Masken miteinander. Die beiden Faltungsmasken G_1, G_2 und ihre zugehörigen Transferfunktionen \hat{G}_1, \hat{G}_2 haben in kontinuierlicher Darstellung einen gaußförmigen Verlauf:

$$\begin{aligned} G_1(x) &= \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{x^2}{2\sigma_1^2}\right] & \circ \text{---} \bullet & \quad \hat{G}_1(k) = \exp\left[-\frac{k^2\sigma_1^2}{2}\right] \\ G_2(x) &= \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{x^2}{2\sigma_2^2}\right] & \circ \text{---} \bullet & \quad \hat{G}_2(k) = \exp\left[-\frac{k^2\sigma_2^2}{2}\right] \end{aligned}$$

Für die Faltung der beiden Masken miteinander gilt:

$$\begin{aligned} \hat{G}_1 \cdot \hat{G}_2 &= \exp\left[-\frac{k^2}{2}\sigma_1^2\right] \exp\left[-\frac{k^2}{2}\sigma_2^2\right] = \exp\left[-\frac{k^2}{2}(\sigma_1^2 + \sigma_2^2)\right] \\ \Rightarrow G_1 * G_2 &= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} \exp\left[-\frac{x^2}{2(\sigma_1^2 + \sigma_2^2)}\right]. \end{aligned} \tag{C.4}$$

Das hintereinander Ausführen von Faltungen mit Binomialmasken entspricht somit einer Faltung mit einer einzigen Binomialmaske, wobei deren Varianz durch die Summe der Varianzen der einzelnen Glättungsmasken gegeben ist.

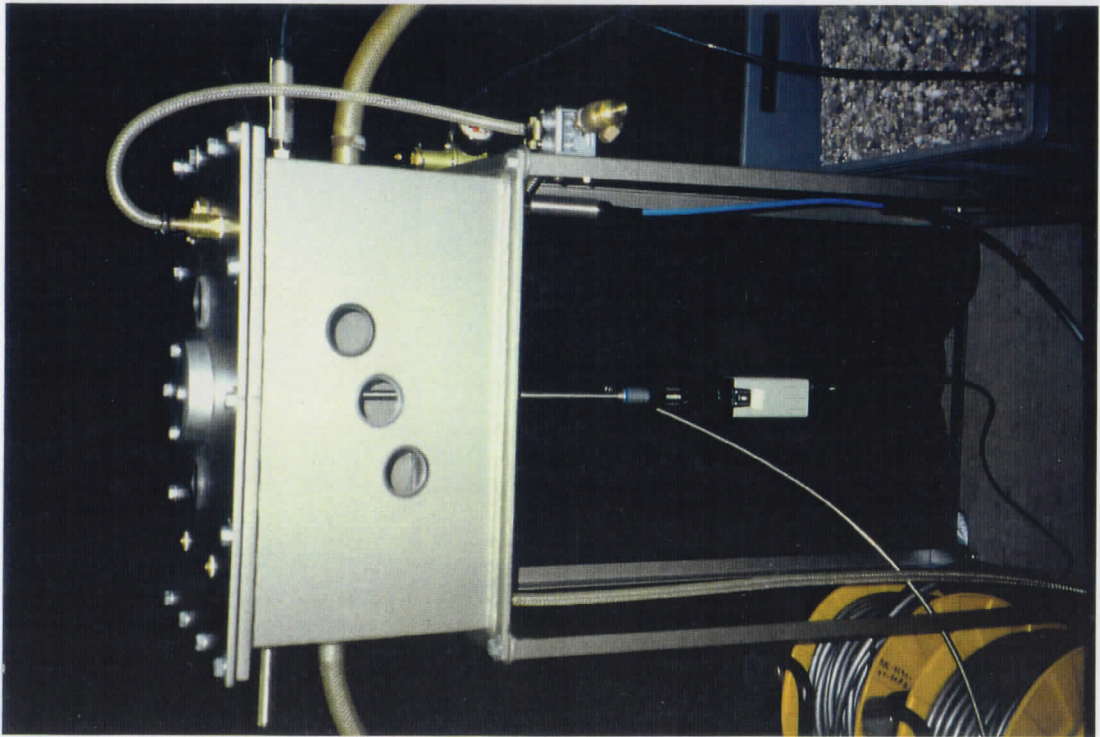
Literaturverzeichnis

- [Athalye et al., 1993] Athalye, A.; Edwards, D.; Manoranjan, V. S.; de Sam Lazaro, A.: *On designing a fuzzy control system using an optimization algorithm*. Fuzzy Sets and Systems, 56, pp. 281-290, 1993.
- [Beyer, 1993] Beyer, M.: *Untersuchung der Relaxationszeiten von winderzeugten Wasseroberflächenwellen durch periodische Variation der Windgeschwindigkeit*. Diplomarbeit, Univ. Heidelberg, 1993.
- [Bouthemy, 1989] Bouthemy, P.: *A Maximum Likelihood Framework for Determining Moving Edges*. IEEE Trans. PAMI, 11, pp. 499-511, 1989.
- [Bouthemy, Lalande, 1990] Bouthemy, P.; Lalande, P.: *Detection and tracking of moving objects based on a statistical regularization method in space and time*. Proc. Computer Vision - ECCV 90, Lecture Notes in Computer Science 427, O.Faugeras (Ed.), pp. 307-311, Springer, New York, 1990.
- [Bracewell, 1965] Bracewell, R.: *The fourier transform and its applications*. Second edition, revised, McGraw-Hill, New York, 1986.
- [Burt, 1984] Burt, P. J.: *The pyramid as a structure for efficient computation*. Multiresolution image processing and analysis, A. Rosenfeld (Ed.), Springer Series in Information Sciences, Vol. 12, Springer, New York, 1984.
- [Cox, 1993] Cox, E.: *Adaptive fuzzy systems*. IEEE Spectrum, pp. 27-31, February 1993.
- [Dueck et al., 1993] Dueck, G.; Scheurer, T.; Wallmeier, H. M.: *Toleranzschwelle und Sintflut: neue Ideen zur Optimierung*. Spektrum der Wissenschaft, pp. 42-51, März 1993.
- [Duncan, Chou, 1992] Duncan, J.; Chou, T.: *On the Detection of Motion and the Computation of Optical Flow*. IEEE Trans. PAMI, 14, pp. 346-352, 1992.
- [Etoh, Shirai, 1993] Etoh, M.; Shirai, Y.: *Segmentation and 2D Motion Estimation by Region Fragments*. Proc. Fourth International Conference on Computer Vision - ICCV 93, Berlin, Germany, pp. 192-199, 1993.
- [François, Bouthemy, 1990] François, E.; Bouthemy, P.: *The derivation of qualitative information in motion analysis*. Proc. Computer Vision - ECCV 90, Lecture Notes in Computer Science 427, O. Faugeras (Ed.), pp. 226-230, Springer, New York, 1990.

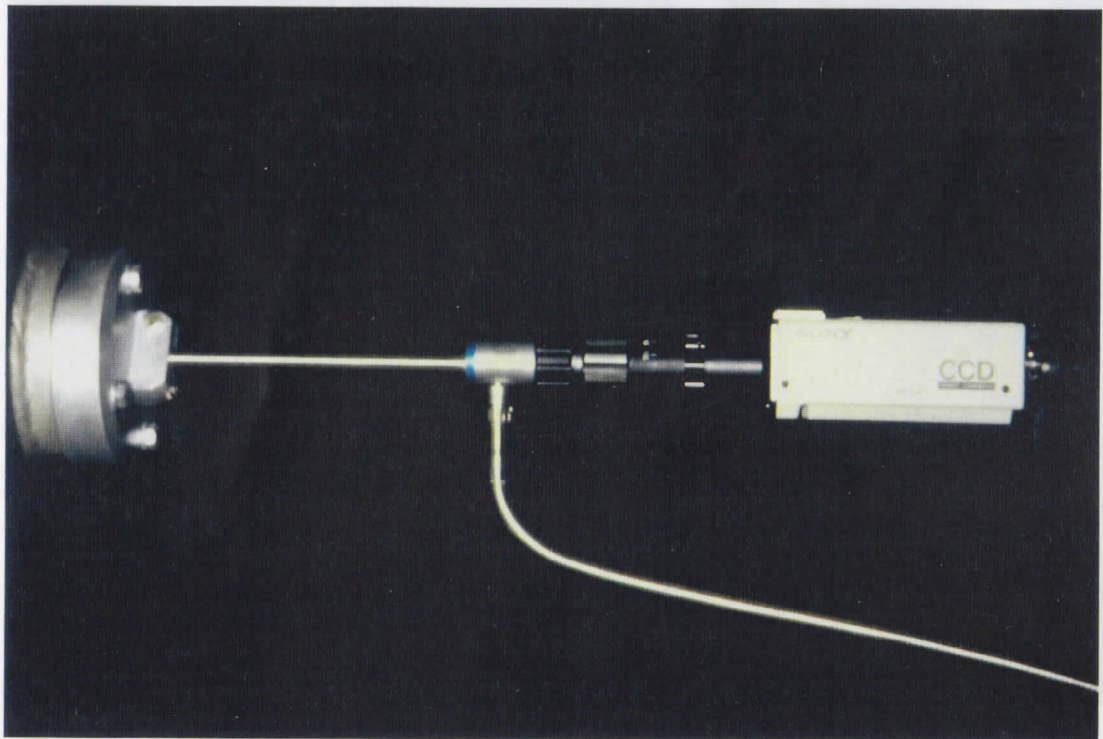
- [Gonzalez,Wintz,1987] Gonzalez, R. C.; Wintz, I.: *Digital Image Processing*. Addison-Wesley, 1987.
- [Haass,1984] Haass, U.L.: *Verfolgung dynamischer Änderungen in TV-Bildszenen zur Deutung von Bewegungen*. FhG-Berichte 2-84, 12, 1984.
- [Haußecker, Jähne,1993] Haußecker, H; Jähne, B.: *Ein Mehrgitterverfahren zur Bewegungssegmentierung in Bildfolgen*. Proc. 15. DAGM Symp. Mustererkennung 1993, S.J. Pöppel (Ed.), pp. 24-31, Springer, Berlin 1993.
- [Hecht,1987] Hecht, E.: *Optics*. Addison-Wesley, 1987.
- [Horn, Schunk,1981] Horn, B.K.P; Schunk, B.G.: *Determining optical flow*. Artificial Intelligence, 17, pp. 185-204, 1981.
- [Hsu et al.,1984] Hsu, Y.Z.; Nagel, H.-H.; Rekers, G.: *New likelihood test methods for change detection in image sequences*. Comp. Graph. Image Proc., 26, pp. 73-106, 1984.
- [Jain,1989] Jain, A.K.: *Fundamentals of digital Image Processing*. Prentice-Hall, Englewood Cliffs, 1989.
- [Jähne, Riemer,1990] Jähne, B.; Riemer K.: *Two-dimensional wave number spectra of small-scale water surface waves*. J. Geophys. Res., 95, pp. 531-546, 1990.
- [Jähne,1993a] Jähne, B.: *Digital Image Processing - Concepts, Algorithms and Scientific Applications*. 2. Auflage, Springer, Berlin, 1993.
- [Jähne,1993b] Jähne, B.: *Spatio-Temporal Image Processing, Theory and Applications*. Lecture Notes in Computer Science, Vol. 751, Springer, Berlin, 1993.
- [Köhler,1985] Köhler, H.-J.: *Modellversuche für die Dimensionierung von Deckwerken an Wasserstraßen - Stabilität loser Steinschüttungen*. Mitt.-Bl.d.BAW 1985 Nr. 56, Bundesanstalt für Wasserbau Karlsruhe 1985.
- [Köhler,1989] Köhler, H.-J.: *Messung von Porenwasserüberdrücken im Untergrund*. Mitt.-Bl.d.BAW 1989 Nr. 66, Bundesanstalt für Wasserbau Karlsruhe 1989.
- [Köhler,1993] Köhler, H.-J.: *The influence of hydraulic head and hydraulic gradient on the filtration process*. Bundesanstalt für Wasserbau Karlsruhe 1993.
- [Kosko,1992] Kosko, B.: *Neural networks and fuzzy systems: a dynamical systems approach to machine intelligence*. Prentice-Hall, 1992.
- [Letang et al.,1993] Letang, J.; Rebuffel, V.; Bouthemy, P.: *Motion detection robust to perturbations: a statistical regularization and temporal integration framework*. Proc. Fourth International Conference on Computer Vision - ICCV 93, Berlin, Germany, pp. 21-30, 1993.
- [Murray, Buxton,1987] Murray, D.; Buxton, B. F.: *Scene Segmentation from Visual Motion Using Global Optimization*. IEEE Trans. PAMI, 9, pp. 220-228, 1987.

- [Murray, Buxton, 1990] Murray, D.; Buxton, B. F.: *Experiments in the machine interpretation of visual motion*. MIT Press, Cambridge Massachusetts, 1990.
- [Nagel, 1978] Nagel, H.-H.: *Formation of an object concept by analysis of systematic time variations in the optically perceptible environment*. Comp. Graph. Image Proc. 7, 149, 1978.
- [Nagel, 1982] Nagel, H.-H.; Rekers, G.: *Moving object masks based on an improved likelihood test*. Proc. Int. Conf. Patt. Recogn., pp. 1140-1142, München 1982.
- [Porat, Friedlander, 1990] Porat, B., Friedlander, B.: *A Frequency Domain Algorithm for Multiframe Detection and Estimation of Dim Targets*. IEEE Trans. PAMI, 12 pp. 398-401, 1990.
- [Rath, 1992] Rath, R.: *Amplitudenmessung von Wasseroberflächenwellen mittels digitaler Bildanalyse*. Diplomarbeit, Univ. Heidelberg, 1992.
- [Reif, 1985] Reif, F.: *Fundamentals of statistical and thermal physics*. McGraw-Hill, Auckland, 1985.
- [Rosenfeld, Kak, 1982] Rosenfeld, A.; Kak, A.C.: *Digital picture Processing*. 2nd ed. Vol. I and II, Academic Press, Orlando, 1982.
- [Scheffer, 1984] Scheffer, F.: *Lehrbuch der Bodenkunde*. Enke, Stuttgart, 1984.
- [Shio, Sklansky, 1991] Shio, A.; Sklansky, J.: *Segmentation of People in Motion*. Proc. IEEE Workshop on Visual Motion, pp. 325-332, 1991.
- [Singh, 1992] Singh, A.: *Optical flow computation*. IEEE Computer Society Press, 1992.
- [Schwartz, Klir, 1992] Schwartz, D. G.; Klir, G. J.: *Fuzzy Logic flowers in Japan*. IEEE Spectrum, pp. 32-35, July 1992.
- [Verri, Poggio, 1987] Verri, A.; Poggio, T.: *Against quantitative optical flow*. Proc. First International Conference on Computer Vision - ICCV 87, London, England, pp. 171-180, 1987.
- [Verri, Poggio, 1989] Verri, A.; Poggio, T.: *Motion Field and Optical Flow: Qualitative Properties*. IEEE Trans. PAMI, 11, pp. 545-558, 1989.
- [Westberg, 1989] Westberg, L.: *Hierarchical object-background segmentation*. Licenciate thesis, Dept. TTT, Royal Inst. of Tech., Stockholm, Sweden, 1989.
- [Westberg, 1992] Westberg, L.: *Hierarchical Contour-Based Segmentation of Dynamic Scenes*. IEEE Trans. PAMI, 14, pp. 946-952, 1992.
- [Yakimovsky, 1976] Yakimovsky, Y.: *Boundary and object detection in real world images*. J.ACM, 23, 599, 1976.
- [Zadeh, 1965] Zadeh, L.A.: *Fuzzy Sets*. Information and Control, 8, pp. 338-353, 1965.
- [Zadeh, 1988] Zadeh, L.A.: *Fuzzy Logic*. IEEE Computer, pp. 83-92, April 1988.

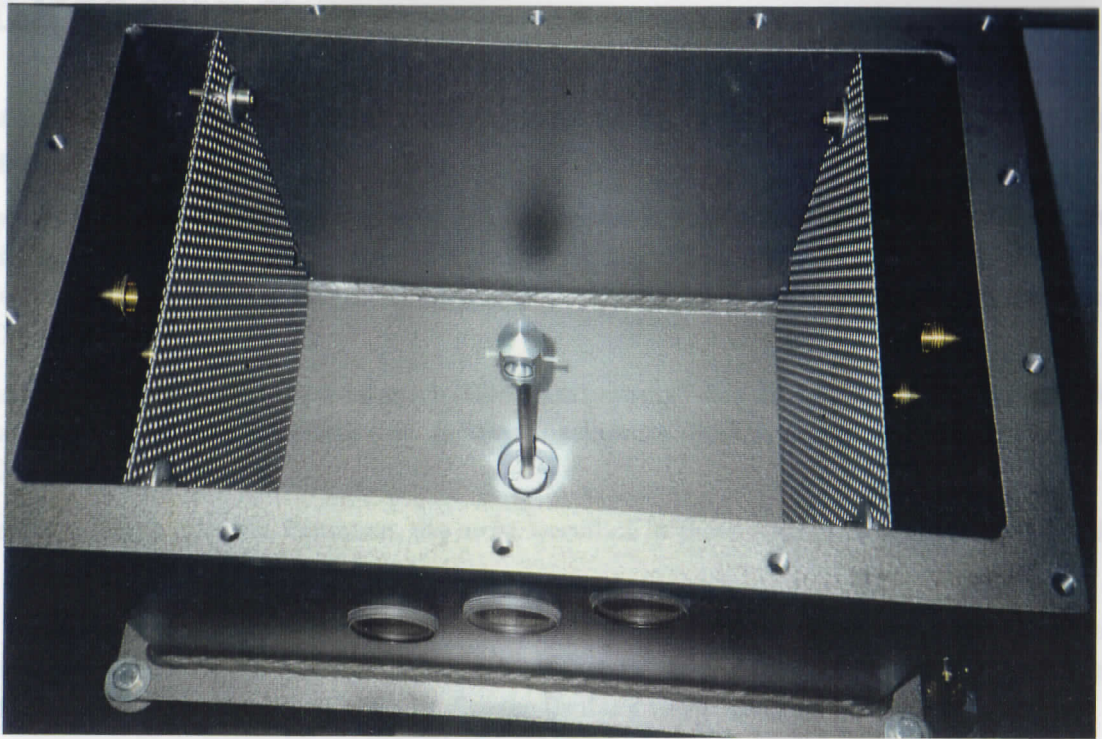
Farbtafeln



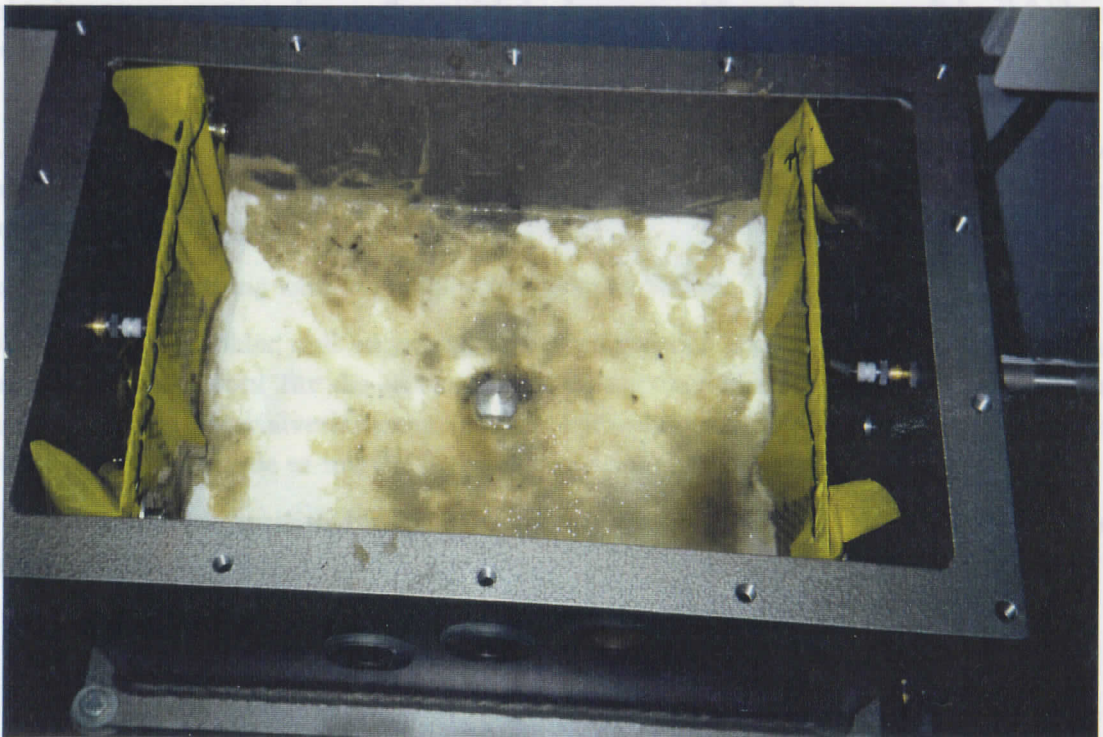
Farbtafel 1: Sedimenttank mit Anschlüssen. Auf der rechten Seite befindet sich die Pneumatikinstallation. Unter dem Tank hängt das Endoskop mit angekoppelter CCD-Kamera. Die drei Druckgeber sind jeweils rechts und links oben und rechts unter dem Tank zu erkennen.



Farbtafel 2: Endoskop mit angekoppelter CCD-Kamera und Lichtleiter. Man erkennt die Klemm-
vorrichtung am Tankboden, durch die das Endoskop druckdicht mit dem Tank verbunden wird.



Farbtafel 3: Blick in den Sedimenttank vor dem Auffüllen. Man erkennt die beiden Vorfluter und das Endoskop mit dem Endoskopaufsatz.



Farbtafel 4: Blick in den Sedimenttank während des Auffüllens. Das Endoskop ist fast vollständig von der Wasser-Sand-Mischung bedeckt. Die beiden Bleche sind mit gelbem Siebdruckgewebe bespannt, um zu verhindern, daß das Sediment herausgespült wird.

Danksagung

Abschließend möchte ich mich bei all denjenigen bedanken, die zum Gelingen dieser Arbeit beigetragen haben:

Als erstes möchte ich mich bei Dr. Bernd Jähne bedanken, der mir diese interessante Arbeit ermöglicht hat. Er half stets mit gezielten Tips zur richtigen Zeit und hatte auf jedes Problem eine Antwort. Zusätzlich gab er mir die Möglichkeit zu einer eigenen Veröffentlichung im Rahmen meiner Diplomarbeit.

Die gesamte Arbeitsgruppe „Gasaustausch/Bildverarbeitung“ des Instituts für Umweltphysik war stets zur Stelle, wenn Hilfe nötig war. Jedem einzelnen zu danken würde den Rahmen dieser kurzen Danksagung sprengen.

Mein herzlicher Dank gilt drei Personen, die nicht beruflich in diese Arbeit verwickelt waren und dennoch wesentlich dazu beigetragen haben:

Meinem Vater danke ich ganz besonders für die Arbeit, die er in die Konstruktion und den Bau des Sedimenttanks investiert hat. Ohne seine technischen Fähigkeiten und seinen Einsatz an Zeit und Material wären die Messungen in diesem Umfang nicht möglich gewesen.

Meiner Verlobten Corinna Linke gilt ein ganz herzlicher Dank. Sie hatte wohl am meisten unter dieser Arbeit zu leiden. Ihren Kenntnissen der Rechtschreibung ist es zu verdanken, daß dem Leser nicht das zugemutet werden muß, was ich in den ersten Manuskripten an Schreibfehlern fabriziert habe. Von der vielen Zeit, in der sie auf mich verzichten mußte, will ich gar nicht erst reden.

Last but not least danke ich der zweijährigen Maike Löchelt. Durch das zeitweise Ausleihen ihres „Sandschäufelchens“ (wenn auch unter Tränen) hat sie entscheidend zum Gelingen der Messungen am Sedimenttank beigetragen.

Die finanzielle Unterstützung kam von zwei verschiedenen Seiten:

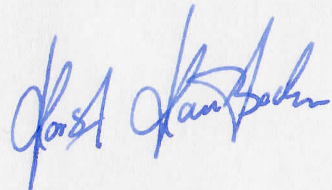
Herrn Dipl. Ing. H.-J. Köhler und seinen Mitarbeitern von der Bundesanstalt für Wasserbau, Karlsruhe danke ich, sowohl für die finanzielle Unterstützung dieser Arbeit im Rahmen des Forschungsprojektes „Materialverlagerungen unter hydrodynamischen Belastungen“, als auch für die vielen Eimer Sand, die er mir für meine Messungen zur Verfügung gestellt hat.

Meinen Eltern danke ich herzlich für die ganz private finanzielle Unterstützung, die mir das Studium erst ermöglicht hat.

Erklärung:

Ich versichere, daß ich diese Arbeit selbstständig verfaßt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, 21. Dezember 1993



(Unterschrift)