

INAUGURAL - DISSERTATION
zur
Erlangung der Doktorwürde
der
Naturwissenschaftlich-Mathematischen
Gesamtfakultät
der
Ruprecht-Karls-Universität
Heidelberg

vorgelegt von
Dipl.-Phys. Martin Brocke
aus Marburg

Tag der mündlichen Prüfung: 27.11.2002

Statistische Ereignisdetektion in Bildfolgen

Gutachter: Prof. Dr. Bernd Jähne
Prof. Dr. Ulrich Platt

Zusammenfassung

Diese Arbeit stellt eine Technik zur Detektion plötzlich auftauchender, untypischer Einzelergebnisse in verrauschten, unruhigen Bildfolgen vor. Verfahren zur Ausreißerdetektion aus der statistischen Datenanalyse wurden für die Bildverarbeitung angepaßt, geeignete Teststatistiken gefunden und die gesuchten Pixel mit Hypothesentests abgetrennt. Dafür wird die Bildfolge als Zeitreihen-Ensemble einzelner Pixel aufgefaßt, die – stationäre Szenen vorausgesetzt – Stichproben der Teststatistik sind. Zeitliche Stationarität ist typisch für viele Anwendungen (Überwachung und räumlich-zeitlicher Meßdatenaufnahme). Das Verfahren hat als einzigen Parameter das Signifikanzniveau des Hypothesentests, womit über verschiedene Datensätze und Anwendungen hinweg Vergleichbarkeit hergestellt wird. Zur binären Entscheidung über Ausreißerzugehörigkeit wird ein Gütemaß dafür mitgeliefert. Unerwünschte Effekte (Verdeckungen/Fehldetektionen bei multiplen Ausreißern) unterdrückt eine iterative Erweiterung. Der Unterschied zu bisherigen Verfahren ist die Ausnutzung der vollen zeitlichen Information anstelle räumlicher Stichproben, die nur aus zwei aufeinanderfolgenden Einzelbildern entstammen. Zwei praktische Umsetzungen des Verfahrens sind: 1) Bei einer Prozeßkontrolle eines Fügeverfahrens mit stark rauschender Sensorik aus der industriellen Qualitätssicherung konnte die Ereignisdetektion in ein automatisiertes Überwachungssystem bis zur Patentreife weiterentwickelt werden. Es ersetzt in der Massenproduktion eine 100%-Sichtprüfung und überwindet erstmals die Schwierigkeiten herkömmlicher, räumlich integrierender Sensoren. 2) Bei der Analyse von thermographischen Bilddaten konnten Bildintensitäten zweier unterschiedlicher physikalischer Ursachen voneinander getrennt werden und die Parameterschätzung der Wärmeflüsse durch frühzeitige Ausreißerdetektion verbessert werden. Das entwickelte Verfahren gibt dem Anwender ein einfach konfigurierbares, schnelles, modular einsetzbares Werkzeug zur Überwachung von Bildfolgen auf untypische Ereignisse an die Hand.

Abstract

This thesis presents a technique to detect statistically unlikely changes in noisy image sequences. Methods for outlier detection are well known in statistical data analysis. This work applies these techniques to image processing. Appropriate statistical tests are performed to identify the relevant pixels by hypothesis testing. The image sequence is represented as a separate time series for each image pixel with the assumption that at steady state the scene is static. This assumption is commonly made for many applications in surveillance and spatio-temporal measurements. The significance level related to the hypothesis test remains the only free parameter. This allows an even comparison of the algorithm's performance across different data sets. A confidence measure is calculated for each binary decision (inlier vs. outlier). Effects such as occlusion or false positives that occur for multiple outliers are controlled by an iterative extension. The algorithm was put into practice twice 1) A complete computer vision system for an industrial laser welding process control was patented. It replaces human visual inspection for mass production and improves robustness over spatially integrating sensors. 2) The algorithm has been applied to infrared image sequences in order to distinguish events caused by two separate processes. Hence heat flux parameter estimation was improved by an outlier detector module at the beginning of the estimation scheme. The technique presented has proven to be an easy-to-configure, modular, and fast tool for event detection in image sequences.

Diese Dissertation wurde an der Universität Heidelberg über den [HeiDok-Katalog](#) der Universitätsbibliothek Heidelberg und über die [Deutsche Bibliothek Frankfurt](#) als elektronischer Volltext veröffentlicht und kann unter <http://www.ub.uni-heidelberg.de/archiv/> kostenlos und ohne Zugangsbeschränkung als PDF-Dokument abgerufen werden.

Inhaltsverzeichnis

Inhaltsverzeichnis	i
1 Einführung	1
1.1 Motivation und Zielsetzung	2
1.2 Überblick zur vorliegenden Arbeit	3
2 Verfahren zur Ausreißeranalyse und Änderungsdetektion	5
2.1 Überblick	5
2.2 Methoden aus der statistischen Datenanalyse	6
2.2.1 Was sind Ausreißer?	6
2.2.2 Klassische Verfahren zur Ausreißerdetektion	7
2.2.3 Simonoff-Methode zur Ausreißerdetektion	13
2.2.4 Wilks-Methode zur Ausreißerdetektion	19
2.2.5 Methode nach Hadi mit robusten Schätzern	21
2.2.6 Zusammenfassung	23
2.3 Änderungsdetektoren der Bildverarbeitung	23
2.3.1 Einleitung und Überblick	23
2.3.2 Objektdetektion versus Änderungsdetektion	24
2.3.3 Änderungsdetektion mit lokalen Bildmodellen	25
2.3.4 Ansätze zur Glättung mit Markov-Zufalls-Feldern	31
2.3.5 Änderungsdetektion mit Hypothesentests	31
2.3.6 Glättung der Änderungskarten mit Bayes-Methoden	38
2.3.7 Ansätze zur Kompensation von Beleuchtungsschwankungen	41
2.3.8 Bewegungsschätzung versus Änderungsdetektion	42
2.3.9 Kontinuierliche Berechnung des Hintergrundes	43
2.3.10 Änderungsdetektoren für Multimediaformate	45
2.3.11 Schlußbemerkung	46

3	Statistischer Ansatz zur Detektion von Ereignissen	47
3.1	Überblick	47
3.2	Bildinhalte und Bildverstehen bei Sequenzen	47
3.3	Herleitung der varianznormierten Mittelwertabweichung	49
3.3.1	Analogie zum Inversen Kontrastverhältnis	49
3.3.2	Varianznormierte zeitliche Abweichung	50
3.3.3	Nachbarschaften	50
3.3.4	Schwellwertbildung und Bemerkungen zur Teststatistik	52
3.3.5	Mögliche Modifikationen	54
3.3.6	Iterativer Ansatz	55
3.3.7	Resultierende Größen und Weiterverarbeitung	58
3.3.8	Zusammenfassung	59
3.4	Invariante Merkmale aus der momentenbasierten Objektanalyse	60
3.4.1	Überblick	60
3.4.2	Merkmale	60
3.4.3	Reguläre und zentrale Momente	61
3.4.4	Invarianten aus Momenten	63
3.4.5	Der Momententensor	66
3.4.6	Die Invarianten nach Hu	68
3.4.7	Vollständigkeit von Momenten und charakteristische Funktion	71
3.4.8	Inverses Momentenproblem	72
3.4.9	Legendre-Momente	74
3.4.10	Weitere Invariantensätze	76
3.4.11	Momente und Rauschen	78
3.4.12	Diskrete Momente	80
3.4.13	Dreidimensionale Momente und Invarianten	81
3.4.14	Schnelle Berechnung von Momenten	81
3.4.15	Zusammenfassung	83
3.5	Polynomklassifikatoren	83
3.5.1	Übersicht	83
3.5.2	Problematik der Klassifikation	84
3.5.3	Mathematisches Gerüst der Polynomklassifikatoren	85
3.5.4	Abgrenzung zu anderen Klassifikatoren	87
3.5.5	Zusammenfassung	89

4	Ereignisdetektion zur Qualitätssicherung	91
4.1	Übersicht	91
4.2	Bildverarbeitung im industriellen Umfeld	92
4.3	Laserschweißen	93
4.4	CMOS-Kameratechnik	98
4.5	Aufgabenstellung und Prozeß	103
4.6	Industrielle Randbedingungen und Anforderungen	105
4.7	Lösungsansatz für die Qualitätssicherung	107
4.8	Aufbau und Ablauf des Überwachungssystems	109
4.9	Algorithmik zur Änderungsdetektion	111
4.10	Teststatistik und Bestimmung kritischer Werte	115
4.11	Rauschen im Vergleich zu den Variationen	120
4.12	Erkennung von Objekten und deren Beschreibung	123
4.13	Merkmale	124
4.14	Klassifikation	125
4.15	Durchgeführte Versuche	127
4.16	Ergebnisse	129
4.17	Vergleich mit anderen Systemen	135
5	Ausreißerdetektion in Thermographiedaten	137
5.1	Übersicht	137
5.2	Motivation	138
5.3	Datenaufnahme und Kameratechnik	139
5.4	Reflektionen als Ausreißer	141
5.5	Bestimmung des Wärmeflusses	142
5.6	Lösungsansatz durch Ausreißerdetektion	148
5.7	Konstruktion stationärer Szenenstücke	150
5.8	Durchführung und Ergebnisse der Ereignisdetektion	152
5.9	Grenzen des Verfahrens	155
6	Resumée	159
6.1	Wahl des Verfahrens	159
6.2	Herausragende Eigenschaften der verwendeten Teststatistik	161
6.3	Wahl der Anwendungen	162
6.4	Rückführung auf bekannte Verfahren	163
6.5	Schlußbemerkung	164

A Anhang: Rechenschema für inverses Momentenproblem	167
Abbildungsverzeichnis	170
Tabellenverzeichnis	171
Literaturverzeichnis	188

»An outlier is an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different mechanism.«

Douglas M. Hawkins [[Hawkins, 1980](#)]

Kapitel 1

Einführung

Mit dieser Arbeit soll die Brücke zwischen der multidimensionalen Bildverarbeitung und der statistischen Datenauswertung geschlagen werden, um in Bildfolgen untypische Ereignisse automatisch zu erkennen. Die dabei entwickelten Methoden werden auf zwei völlig unterschiedlichen Datensätze einmal aus der industriellen Fertigung und Qualitätssicherung, zum anderen aus der Bilddatenaufnahme im Rahmen wissenschaftlicher Versuchskampagnen angewendet und auf ihre Einsatztauglichkeit geprüft.

In der Verarbeitung von Bildfolgen, in der Regel zeitlichen Abfolgen von Grauwertbildern, ist die Detektion von Ereignissen und Änderungen in den Bildern ein vorrangiges Ziel: die Erfassung von Bildveränderung (engl. *change detection*) ist Voraussetzung und Grundlage für die Berechnung von Optischem Fluß, Bewegung und Segmentierung von bewegten Objekten. Moderne Anwendungen sind zu finden bei der Stereobild-Analyse, in der Kodierung von Videomaterial (MPEG-Standard), der Videoüberwachung (Objektschutz, Verkehrsüberwachung), Robotik und der industriellen Qualitätssicherung von Prozessen, aus der wir eine Anwendung in dieser Arbeit ausführlich diskutieren werden. Änderungsdetektion ist dabei ein oft verborgen und im Vorfeld der Verarbeitung (engl. *early vision*) stattfindender Prozeß, für den Algorithmen Ende der 70er Jahre entworfen wurde, als Datenspeicher und Rechenspeicher erstmals eine Verarbeitung von Bildfolgen zuließen. Heute wären *change detectors* wahrscheinlich vergessen, wenn nicht in den 90er Jahren ein enormer Bedarf an effizienten Kompressionsmöglichkeiten von Bildfolgen durch die Möglichkeit der Übertragung von Bilddaten im Internet und die Realisierung großer digitaler Videoarchive aufgekommen wäre.

Fassen wir die einzelnen Pixel in einer Bildfolge als Ensemble von Zeitreihen oder bei mehrkanaligen Bildern als multivariate Werte auf, so stellt uns die Disziplin der statistischen Datenanalyse unter dem Begriff der Ausreißerdetektion (engl. *outlier detection*) Werkzeuge zur Suche nach abrupten Sprüngen in den Daten zur Verfügung. Ähnlich wie die Änderungsdetektoren in der Bildverarbeitung sind auch die Ausreißerdetektoren der Statistik ein – gemessen an der jeweiligen Disziplin – altes, hier sehr zu unrecht unbeachtetes, Werkzeug. Ausreißerdetektion sollte jeder Verarbeitung experimentell ermittelter Datensätze vorausgehen und hat zum Ziel, offensichtliche Meßfehler, Ablesefehler, Zahlendreher, bei automatischer Datenaufnahme auch Geräteausfälle oder eingebrachte Störungen auszuschließen, um Resultate in Form von Meßgrößen oder validierten Modellen auf tatsächlich relevante Daten zu stützen.

Obwohl letztlich mit großer Sicherheit davon auszugehen ist, daß bei aufwendigen Experimenten und großen Datensätzen immer Ausreißer vorhanden sind, werden systematische Tests zur Suche danach selten betrieben. Gründe dafür mögen darin liegen, daß die Ausreißerdetektion fast ausschließlich in der Sprache der Statistiker formuliert ist, die zusätzlich zum Verständnis der Inhalte in den Sprachgebrauch der Anwenderdisziplinen übertragen werden muß, daß die statistische Datenanalyse nicht zum Lehrkanon naturwissenschaftlicher Grundlagen gehört, ihre Original-Literatur aufgrund des Alters der Methoden wenig zugänglich ist und die Verfahren auch mit heutigen Mitteln der Auswertung zeitraubend geblieben sind. Für eine stärkere Beachtung spricht, daß nicht wenige experimentelle Nachweise und Experimente der modernen Physik, die sich deutlich an der statistischen Nachweisgrenze bewegen, ausschließlich von Statistikern zu einem Ergebnis geführt wurden oder zu falschen Schlüssen, weil die statistische Absicherung der Resultate unsauber geführt wurde.

Während in der statistischen Datenanalyse von "Ausreißern" gesprochen wird und in der Bildverarbeitung und Überwachungstechnik von "Änderungen" die Rede ist, wird in dieser Arbeit meist nach "Ereignissen" gesucht. Auch ist in der Statistik der Begriff "Ausreißer" eher negativ besetzt, und es werden in der Regel keine weiteren Untersuchungen außer dem reinen Erkennen und Löschen an den Ausreißern vorgenommen, wogegen es in dieser Arbeit gerade um ein Verständnis der physikalischen Ursachen jener Bildelemente geht, die als "Ausreißer" erscheinen. Ebenso ist der Begriff der "Änderungen" in der Bildverarbeitung historisch bedingt heute automatisch mit den Gebieten Bewegungsanalyse und Optischem Fluß verknüpft; beide Themen werden in dieser Arbeit nur am Rande vorkommen. Die Gründe gegen die Verwendung der Bezeichnungen "Ausreißer" und "Änderungen" rechtfertigen aber nicht unbedingt, durchgängig und weitaus passender von "untypischen Ereignissen" zu sprechen: immer dort, wo es die Nähe zu bereits bekannten Detektionsverfahren oder der in den oben genannten Disziplinen übliche Sprachgebrauch nahelegt, werden wir weiterhin von Ausreißern oder Änderungen sprechen.

1.1 Motivation und Zielsetzung

Der in dieser Arbeit zentrale Ereignisdetektor wird in Aufgaben der Bildfolgenverarbeitung eingesetzt, in denen gerade die Änderungen von Bild zu Bild oder über mehrere Bilder hinweg gesucht werden: diese stellen die Datenregionen von eigentlichem Interesse dar.

Ziel ist es, einen solchen möglichst generell gehaltenen Algorithmus auf verschiedene Bilddaten anzuwenden, mit früheren Verfahren zu vergleichen und die Probleme in zwei unterschiedlichen Anwendungen zu lösen. Die gestellte Aufgabe ist aus der Praxis motiviert, in der nach einem einfachen und schnellen Verfahren gesucht wird, um plötzliche und untypische Ereignisse in Bildfolgen stationärer Prozesse automatisch zu finden.

Für die industrielle Aufgabenstellung aus der Prozeßüberwachung ist dies die sichere Detektion von Prozeßfehlern und außergewöhnlichen Ereignissen, die die Qualität der Prozesse mindern und zu fehlerhaften Bauteilen führen. Die zu entwerfende Algorithmik soll als Qualitätskontrolle des Prozesses diese schadhaften Teile schnell und sicher identifizieren und mit einem Alarm melden.

Für die umweltphysikalische Aufgabenstellung gilt es, in Bilddaten mit starken Intensitätsvariationen ungewöhnlich helle und plötzlich auftauchende Bildregionen zu finden. Grauwertintensitäten

in diesen Gebieten werden von einem anderen physikalischen Prozeß als die übrigen Werte im Bild verursacht und dürfen daher nicht in die Parameterschätzung eines Modelles einfließen, das für die reguläre Verteilung der Grauwerte entworfen wurde. Konkret soll für die gleichzeitige Berechnung von Temperaturfeldern und Bewegungsfeldern zu den bereits implementiert vorliegenden robusten Verfahren ein Ausschlußverfahren offensichtlicher Ausreißer-Effekte hinzukommen und die Parameterschätzungen stabiler und sicherer zu machen.

1.2 Überblick zur vorliegenden Arbeit

Die vorliegende Arbeit wurde am Interdisziplinären Zentrum für wissenschaftliches Rechnen der Universität Heidelberg als Industriepromotion bei der Robert Bosch GmbH (Forschung und Vorausbildung, Automatisierungstechnik) in Schwieberdingen betreut. Seitens beider Stellen flossen durch diese Arbeit neue Erkenntnisse in ein landesgefördertes Projekt *OPEN-EYE* zahlreicher Partner aus Wissenschaft und Industrie, bei dem der Einsatz neuer Sensorik (in unserem Fall der HDRC-Sensor) für die multidimensionale Bildverarbeitung (hier Sequenzbildverarbeitung) erarbeitet werden sollte. Mit Abschluß dieser Promotion wurden Ausschnitte und einzelne Ergebnisse in [Brocke, 2002] und [Jähne, Bernd und Brocke, Martin et al., 2002] (geplant ist [Brocke, 2003]) veröffentlicht sowie in [Brocke, Martin u. a., 2002] patentiert. Zu den bearbeiteten Gebieten wurden in drei Teilbereichen Diplomarbeiten angefertigt [Sax, 2001] [Ritter, 2001] [Seiffert, 2001] und kleinere Praktika betreut. Eine parallel laufende Arbeit [Schmidt, 2003] konzentrierte sich auf die physikalischen Hintergründe beim Laserschweißen sowie den Fertigungsnachweis des in Kapitel 4 vorgestellten Systems. Eine Anschlußarbeit [Hader, 2004] konzentriert sich unter dem Stichwort data-mining auf die systematische Auswertung großer Datensätze, wie sie bei der Verarbeitung von Bildfolgen in der industriellen Massenproduktion typischerweise anfallen.

In Kapitel 2 (Seite 5) werden die in der statistischen Datenanalyse bekannten Verfahren zur Ausreißerdetektion und die in der digitalen Bildverarbeitung anerkannten Verfahren zur Änderungsdetektion vorgestellt. Zahlreiche zentrale Begriffe und Effekte werden eingeführt und praxiserprobte Methoden beschrieben. In Kapitel 3 ab Seite 47 wird aus dieser theoretischen Vorbereitung ein neues Verfahren zur Ereignisdetektion in stationären und möglicherweise stark verrauschten Bildfolgen entworfen und mit einer iterativen Verschleifung um eine robusten Erweiterung ergänzt. Die beiden Kapitel 4 (Seite 91ff.) und 5 (Seite 137ff.) behandeln dann ausführlich die beiden sehr unterschiedlichen Umsetzungen des Verfahrens in die Praxis. Beide Anwendungen werden motiviert und ihre Bedeutung erläutert, Sensorik der jeweiligen Bildaufnehmer (hochdynamische CMOS-Kamera und Infrarot-Kamera), genereller Versuchsaufbau und Randbedingungen der Datenaufnahme ebenso besprochen wie die genaue Anpassung des in dieser Arbeit zentralen Verfahrens. Aus den Ergebnissen ziehen wird im Kapitel 6 auf Seite 159 ein Resumée und bewerten abschließend Vorteile, herausragende Eigenschaften und Begrenzungen des Verfahrens im Vergleich mit anderen Methoden.

Kapitel 2

Verfahren zur Ausreißeranalyse und Änderungsdetektion

2.1 Überblick

In diesem Kapitel werden Ausreißerdetektoren (engl. *outlier detectors*), so werden sie in der statistischen Datenanalyse genannt, und Detektoren für Änderungen (engl. *change detectors*), wie sie in der Bildverarbeitung bekannt sind, vorgestellt. Wir konzentrieren uns dabei auf klassische und allgemein gebräuchliche Verfahren beider Disziplinen und gehen an solchen Stellen tiefer in die Theorie, an denen Ideen und Erkenntnisse für das neue, im Rahmen dieser Dissertation gefundene Verfahren (siehe Kapitel 3) relevant sind. Zwischen den beiden Fachgebieten der statistischen Datenanalyse (Abschnitt 2.2) und der Bildfolgenverarbeitung (Abschnitt 2.3) gibt es bezüglich der Detektion ungewöhnlicher und einmaliger Ereignisse in zweidimensionalen Zeitreihen, also Bildsequenzen, nach intensiver Sichtung der Literatur bisher keine Verbindungen.

Nach einer einleitenden Begriffsklärung im Abschnitt 2.2.1 besprechen wir mit der Thompson-Teststatistik in Abschnitt 2.2.2 eines der wichtigsten klassischen Verfahren und erläutern anhand der vor allem von Grubbs, Chandra Sekar und Pearson geführten Diskussion dazu unerwünschte Effekte (Seite 10), die multiple Ausreißer auf einen Detektor haben können. Auch wird eine allgemeine Formulierung für einen Hypothesentest auf Basis einer Signifikanz α gegeben. Mit den Methoden von Simonoff und denen von Wilks in den Abschnitten 2.2.3 und 2.2.4 besprechen wir moderne Verfahren, die Fehldetektionen durch multiple Ausreißer nach den Ideen und Ansätzen von Rosner ausschließen. Die Methode von Hadi (Abschnitt 2.2.5) schließlich kombiniert aktuelle Ansätze robuster Schätzer mit Ausreißerdetektoren. Auf Übersichtsartikel zu zahlreichen weiteren Verfahren verweisen wir in der Zusammenfassung (Abschnitt 2.2.6).

Für die bereits bekannten Verfahren aus dem Bereich der Bildverarbeitung leiten wir in Abschnitt 2.3.1 mit einem historischen Rückblick, intuitiven Ansätzen und deren Probleme zu einem kurzen Einschub in Abschnitt 2.3.2 über, in dem die Änderungsdetektion von der Objektdetektion abgegrenzt wird. In Abschnitt 2.3.3 wird das inzwischen klassische und erste Verfahren zur Änderungsdetektion aus den Arbeiten von Nagel et al. vorgestellt, zu dem Bouthemy et al. (Abschnitt 2.3.4) die Bayes-Methoden von Geman et al. zur Regularisierung der Ergebnisse hinzufügte. Aach et al. benutzte diese

in zahlreichen Anwendungen der Bildverarbeitung – für den Bereich der Änderungsdetektion werden zuerst seine Hypothesentests mit unterschiedlichen Teststatistiken in Abschnitt 2.3.5 vorgestellt und dann die Glättungen (Abschnitt 2.3.6), die im wesentlichen die Entscheidungsschwellen modifizieren. Aktuelle Erweiterungen zum Problem der Beleuchtungsschwankungen werden in Abschnitt 2.3.7 genannt. Nachdem die Änderungsdetektion von der Bewegungsschätzung in Abschnitt 2.3.8 abgegrenzt wurde, werden noch kurz für diese Arbeit interessante Aspekte der kontinuierlichen Berechnung eines Hintergrundbildes und der Bildfolgenanalyse im Multimediabereich in den Abschnitten 2.3.9 und 2.3.10 angesprochen.

2.2 Methoden aus der statistischen Datenanalyse

2.2.1 Was sind Ausreißer?

Seit die Naturwissenschaften theoretische Modelle am Experiment nachweisen wollen, beschäftigen sich Datenanalytiker mit der Frage, wie mit extremen, weit ab der Mehrzahl liegenden Meßpunkten, zu verfahren sei. Schon über die Definition der Ausreißer (engl. *outliers*), also letztlich die Frage, was “extrem” und “weit abgelegen” in der Praxis bedeutet, bemerkte Bernoulli treffend

»I see no way of drawing a dividing line between those that are to be utterly rejected and those that are to be wholly retained; it may even happen that the rejected observation is the one that would have supplied the best correction to the others.«

[Bernoulli, Daniel und Allen, C. G., 1777]

Während bei univariaten Daten die Abgelegenheit einiger Werte am oberen und unteren Ende des Wertebereiches durch einen einfachen Blick auf die Meßwerte auffällt, muß bei multivariaten Daten ein Ausreißer keineswegs extreme Werte, auch nicht in einer einzigen seiner Dimensionen, annehmen, und kann dennoch ein Ausreißer sein (zum Beispiel bei Kugelschalen-förmigen Punktwolken). Ohnehin stellt sich die Frage, welche Rangfolge man multivariaten Ausreißern mit schwachen Abweichungen in allen Dimensionen gegenüber multivariaten Ausreißern starker Abweichungen in nur einer Dimension gibt, vor allem, wenn die unterschiedlichen Dimensionen völlig verschiedene Wertebereiche abdecken und unterschiedlichste physikalische Effekte abbilden. Andere beim Übergang zu multivariaten Daten auftretende Probleme sind, daß Ausreißer nicht mehr nur Parameter der mittleren Lage und deren Streuungen verzerren, sondern auch Korrelationen.

Zahlreiche mehr oder weniger ungenaue Definition von Ausreißern gibt es neben einer Vielfalt von Namensgebungen, die aber in der Regel gemeinsam haben, daß Ausreißer die Anpassung eines Modelles an die Daten erschweren oder unmöglich machen - kein Wunder, daß sie oft als *contaminants* bezeichnet werden [Davies, Laurie und Gather, Ursula, 1993].

Zwei grundlegende, aber letztlich nur im akademischen Sinn sauber zu trennende Vorgehensweisen bieten sich an. Das ältere Verfahren, das mit den grundlegenden Arbeiten von Pearson, Thompson und Grubbs in den 30er Jahren formalisiert wurde, ist die möglichst sichere Identifikation von Ausreißern zusammen mit einem Gütemaß, daß die Ausprägung der Ausreißereigenschaft beschreibt und auf die Frage “Wie weit weg?” antwortet. Wie dann mit den derart gefundenen Datenpunkten zu

verfahren ist, obliegt dem einzelnen und der konkreten Aufgabe. Auf dieser ersten Vorgehensweise liegt der Schwerpunkt der vorliegenden Arbeit. Seit Ende der 60er Jahre kam ein neuer Ansatz auf: Die anpassenden Verfahren haben zum Ziel, die Modelle, die von den Daten bestätigt oder widerlegt werden sollen, und die Verfahren, die zu dieser Überprüfung notwendig sind, so anzupassen, daß sie stabil sind gegenüber der Existenz von Ausreißern: die Aussagen des Experimentes können dann durch mehr oder weniger häufige Ausreißer nicht umgestoßen werden. Man spricht hier von *robuster Statistik* oder *robusten Verfahren* (teilweise auch aus dem Englischen von *Robustifikation*). Erste Methoden, bei denen die Identifikation auf Methoden der robusten Schätzung zurückgreift, kamen in den 90er Jahren auf und werden als Brückenschlag zwischen beiden generellen Verfahrensweisen in Abschnitt 2.2.5 und 5.6 vorgestellt.

2.2.2 Klassische Verfahren zur Ausreißerdetektion

Verfahren zur Ausreißerdetektion gehen meist von folgendem Szenario aus: eine Stichprobe X mit N Werten $\{x_1, \dots, x_N\}$ sei aus zwei Verteilungen aufgebaut: $\{x_{i_1}, \dots, x_{i_m}\}$ mit $x_{i_p} \in N(\bar{x}_m, \sigma_m)$ wobei $p = 1, \dots, m$ und $\{x_{j_1}, \dots, x_{j_m}\}$ mit $x_{j_q} \in N(\bar{x}_k, \sigma_k)$ wobei $q = 1, \dots, k$ und $m+k=N$. Ziel ist es, die $k < N/2$ Ausreißer in der Gesamtstichprobe X zu identifizieren (robuste Schätzmethoden würden sich auf die möglichst korrekte Schätzung von \bar{x}_m und σ_m konzentrieren). Allgemein wird dazu ein θ_α gesucht, so daß die Wahrscheinlichkeit $P(S(x_{i_p}, X) > \theta_\alpha \mid \text{keine Ausreißer präsent}) = \alpha/N$ für eine noch festzulegende Teststatistik $S(x_n, X)$. Ausreißer sind dann gegeben durch

$$S(x_n, X) \begin{cases} > \\ \leq \end{cases} \theta_\alpha \begin{cases} x_n \text{ ist ein Ausreißer} \\ x_n \text{ ist kein Ausreißer} \end{cases}$$

In einer der ersten Arbeiten zur Detektion eines einzelnen Ausreißers schlug Thompson [Thompson, 1935] vor, die Abweichung der zu prüfenden Werte x_n von ihrem arithmetischen Mittel (normiert auf ihre geschätzte Varianz) zu benutzen.

$$\tau_n = \frac{x_n - \langle x \rangle_N}{s} \begin{cases} > \\ \leq \end{cases} \theta_{\alpha;N} \tag{2.1}$$

$$\text{mit } s = \frac{1}{N} \sqrt{\sum_{i=r}^N (x_r - \langle x \rangle_N)^2} \quad \text{und} \quad \langle x \rangle_N = \frac{1}{N} \sum_{r=1}^N x_r$$

Thompson zeigte, daß die Größe

$$\xi_n = \sqrt{\frac{\tau_n^2 (N-2)}{N-1-\tau_n^2}}$$

nach Student t_{N-2} -verteilt ist mit $N-2$ Freiheitsgraden (die Größe $\xi \cdot (N-2)^{-1/2}$ ist nach R. A. Fisher z -verteilt). Aus der Signifikanz α und dem Freiheitsgrad $N-2$ erhält man den kritischen Wert $\xi_{\alpha;N}$ als Grenze des Integrationsintervall bei der zweiseitigen Aufintegration der t_{N-2} -Verteilung aus dem Unendlichen. Thompson tabelliert (Tabelle 2.1) diese kritischen Werte $\xi_{\alpha;N}$ und $\theta_{\alpha;N}$ für τ bei verschieden große Datensätzen N und Signifikanzen α .

krit. Werte	$\alpha = 0,2$		$\alpha = 0,1$		$\alpha = 0,05$	
	$\xi_{0,2;N}$	$\theta_{0,2;N}$	$\xi_{0,1;N}$	$\theta_{0,1;N}$	$\xi_{0,05;N}$	$\theta_{0,05;N}$
3	9,514	1,406	19,081	1,412	38,189	1,414
10	2,896	2,146	3,355	2,294	3,833	2,414
20	2,878	2,447	3,197	2,623	3,510	2,779
30	2,930	2,609	3,208	2,792	3,479	2,958
40	2,980	2,718	3,238	2,904	3,487	3,075
50	3,024	2,800	3,269	2,987	3,505	3,160
100	3,176	3,039	3,393	3,226	3,601	3,401
200	3,340	3,258	3,539	3,441	3,730	3,615
500	3,564	3,523	3,747	3,699	3,922	3,867
700	3,646	3,615	3,824	3,788	3,995	3,953
1000	3,733	3,709	3,907	3,879	4,075	4,043

Tabelle 2.1: Die kritischen Werte $\xi_{\alpha;N}$ für ξ und die entsprechenden $\theta_{\alpha;N}$ für die Teststatistik τ in (2.1) bei verschiedenen Signifikanzen α und Datensatzgrößen N . Der Wert $\xi_{\alpha;N}$ ist gegeben durch die Integrationsgrenze, bis zu der bei der Aufintegration der Student-Verteilung mit $N-2$ Freiheitsgraden aus dem Unendlichen der Integralwert α/N erreicht ist. Siehe auch Abbildung 4.18 (Seite 119) und Tabelle 5.1 (Seite 152).

Thompson verallgemeinerte diese Teststatistik, so daß k Ausreißer angenommen werden können, woraus sich dann neue kritische Werte ergeben. Die Teststatistik sei dazu

$$\tau' = \frac{\langle x \rangle_k - \langle x \rangle_N}{s} = \frac{(N-k)(\langle x \rangle_k - \langle x \rangle_m)}{N s} = \frac{(N-k)(\langle x \rangle_N - \langle x \rangle_m)}{k s}$$

$$\text{mit } \langle x \rangle_k = \frac{1}{k} \sum_{q=1}^k x_{j_q} \quad \text{und} \quad \langle x \rangle_m = \frac{1}{m} \sum_{p=1}^m x_{i_p}$$

Mit einem Beweis von R.A. Fisher kann gezeigt werden, daß die Größe

$$\zeta = \tau' \sqrt{\frac{(N-2) k}{N-k-k\tau'^2}}$$

wieder Student-verteilt ist mit $N-2$ Freiheitsgraden. Mit der Signifikanz α ergeben sich für verschiedene Datensatzgrößen N wieder kritische Werte $\zeta_{\alpha;N;k}$, aus denen die Schwellwerte $\theta'_{\alpha;N;k}$ für die Teststatistik durch Umkehrung der letzten Gleichung berechnet werden können. Für $k=1$ ergeben sich wieder die $\theta_{\alpha;N} = \theta'_{\alpha;N;1}$, so daß mit besagter Umkehrung auch folgt

$$\theta'_{\alpha;N;k} = \sqrt{\frac{N-k}{k(N-1)}} \theta'_{\alpha;N;1} = \sqrt{\frac{N-k}{k(N-1)}} \theta_{\alpha;N}$$

Bei dieser Erweiterung ist zu beachten, daß k fest gegeben sein muß und in der Teststatistik τ' die Aufteilung des Datensatzes die k Ausreißer und m Nicht-Ausreißer bereits bekannt sein müssen. Da dies bei der Anwendung des Verfahrens noch nicht der Fall ist, müssen zahlreiche Aufteilungen des Datensatzes durchgerechnet werden, bis die durch das Kriterium ausgeschlossenen k Ausreißer

sich mit der zuvor benutzten Aufteilung des Datensatzes decken und ein konsistentes Bild geben. Der enorme Rechenaufwand dafür ist aber nicht der einzige Grund, warum das Verfahren sich nicht durchsetzen konnte. Viel schwerer wiegen die von [Pearson, Erwin S. und Sekar, Chandra, 1936] bei einer Untersuchung der Leistungsstärke der verschiedenen Teststatistiken gefundenen Ergebnisse, daß die Präsenz von (mehreren) Ausreißern die Annahmen der Normalverteilung derart stören kann, daß die Verfahren nicht mehr funktionieren. Dies gilt auch für die noch einzuführende Teststatistik der ESD.

Als Teststatistiken für den einzelnen Punkt x_n gegenüber den Daten X sind beispielsweise auch noch gebräuchlich die Kurtosis, die *studentized range* [David, H. A. u. a., 1954] als varianznormierte Größe des Wertebereichs $(x_{[N]} - x_{[1]})/s$ (mit eckigen Klammern sind üblicherweise die nach Wert sortierten Datensätze indiziert, also $x_{[1]} \leq x_{[2]} \leq \dots \leq x_{[N]}$) und vor allem der Studentisierter Extrembereich (engl. *extreme studentized deviate*, ESD)

$$ESD = |\tau_n|_{[N]} = \max_{n=1, \dots, N} |\tau_n| = \max_{n=1, \dots, N} \frac{|x_n - \langle x \rangle_N|}{s}$$

der den Abstand der stärksten Abweichung vom geschätzten Mittelwert $\langle x \rangle_N$ angibt. Der Wertebereich der ESD geht bis $(N-1)/\sqrt{N}$. Bei der Anwesenheit eines Ausreißers wird in τ_n der Varianzschätzer s größer und für den Ausreißer der Zähler kleiner, für die Nicht-Ausreißer wird der Zähler aber größer: der τ_n des Ausreißers sinkt also, während die übrigen τ_n -Werte ungefähr gleich bleiben. Der ESD-Wert ist dann nicht mehr so ausgeprägt. Auf dieses von Pearson und Chandra Sekar in ihrer grundlegenden Arbeit [Pearson, Erwin S. und Sekar, Chandra, 1936] zur ESD als Fortsetzung von Thompsons Ansätzen beschriebene Problem, daß Ausreißer in zahlreichen Teststatistiken sich selbst verdecken, kommen wir auf Seite 2.2.2 nochmal zurück. Einige Autoren nehmen daher zumindest den gerade zu testenden Wert x_n aus der Berechnung der Schätzer für Mittelwert und Varianz heraus. Für die Anwesenheit von zwei oder mehr Ausreißern in Datensätzen kleiner als $N \approx 30 \dots 40$ sind die Verfahren wegen Verletzung der Normalverteilungsannahme nicht mehr brauchbar.

Grubbs [Grubbs, 1950] knüpfte in seiner Dissertation an dieser Stelle an und entwirft Tests für ein oder zwei Ausreißer. Die Teststatistik

$$S(x_{[N-1]}, x_{[N]}, X_{[1]}) = \frac{S_{(N-1;N)}^2}{S^2} = \frac{\sum_{n=1}^{N-2} (x_{[n]} - \langle x \rangle_{(N-1;N)})^2}{\sum_{n=1}^N (x_n - \langle x \rangle_N)^2}$$

$$\text{mit } \langle x \rangle_{(N-1;N)} = \frac{1}{N-2} \sum_{n=1}^{N-2} x_{[n]}$$

sucht nach zwei Ausreißern auf der rechten Seite und benutzt den nach Größen sortierten Datensatz $X_{[1]}$ und eine Notation, die in runden tiefergestellten Klammern $_{(N-1;N)}$ durch Semikolon getrennt jene Elemente aufzählt, die im sortierten Datensatz nicht berücksichtigt wurden. Ähnliche Statistiken $\frac{S_{(1;2)}^2}{S^2}$ und $\frac{S_{(N)}^2}{S^2}$ bzw. $\frac{S_{(1)}^2}{S^2}$ existieren für die Suche nach einem oder zwei Ausreißern am rechten oder linken Rand. Grubbs berechnet die Verteilungen der ESD, tabelliert kritische Werte dafür und für die von ihm vorgeschlagenen Statistiken und findet den Zusammenhang

$$\frac{S_{(N)}^2}{S^2} = 1 - \frac{ESD^2}{N-1}$$

Die Ergebnisse wurden in [Tietjen, Gary L. Moore Roger H. und Beckman, R. J., 1973] für lineare Regression und [Tietjen, 1972] für die vier Tests von Grubbs verallgemeinert auf die Suche nach k Ausreißern und die gleichzeitige Suche am rechten und linken Ende des Datensatzes $X_{[]}$.

In der Regel ist davon auszugehen, daß sich in großen Datensätzen ein multiples Ausreißerproblem stellt. Mehrfache Ausreißer erfordern eine besondere Beachtung der Konstruktion der Suchverfahren, denn sie können die verwendeten Statistiken derart verfälschen, daß sie zwar einzeln zu identifizieren wären, sich aber bei mehrfachen Auftauchen gegenseitig verdecken oder sogar “gute Daten” (engl. *inliers*) als potentielle Ausreißer identifiziert werden. Vom Maskierungseffekt (engl. *masking*) [Pearson, Erwin S. und Sekar, Chandra, 1936] spricht man, wenn ein oder mehrere sehr stark ausgeprägte Ausreißer im benutzten Detektor andere vorhandene Ausreißer verdecken. Beispielsweise ziehen sie den Mittelwert derart auf ihre Seite, daß schwächer ausgeprägte Ausreißer schließlich als optimale Daten dastehen, und somit vor dem Detektor versteckt werden. Schlimmer noch kann die Mehrheit der Daten, die die tatsächliche Ausreißer-freie Beobachtung repräsentiert, durch verschobene Mittelwerte in den Verdacht einer Abweichung zur anderen Seite kommen Dies ist als Überschwemmungseffekt (engl. *swamping*) bekannt [Hadi, Ali S. und Simonoff, Jeffrey S., 1993], der “gute” Daten als Ausreißer identifiziert, die in großer Zahl die Liste der Ausreißer überschwemmen. Die Begriffe “false positive” und “false negative” dafür sollen in dieser Arbeit vermieden werden, weil Ausreißer gerade die gesuchten Daten sind und ihr Auffinden eher positiv als negativ verstanden werden wird.

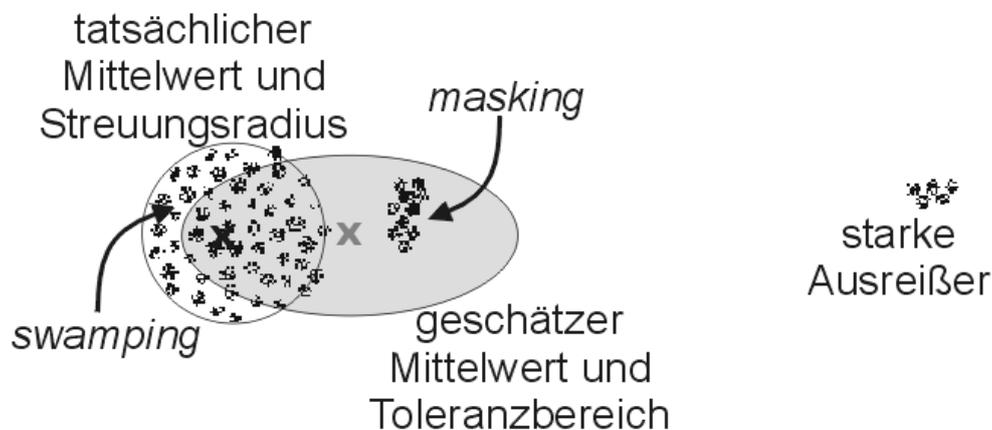


Abbildung 2.1: Ausreißer können einen einfachen Mittelwertschätzer und einen Toleranzbereich für die Ausreißeridentifikation so stark verändern, daß vorhandene Ausreißer nicht gefunden werden (*masking*) und Datenpunkte aus der Masse der Beobachtungen fälschlicherweise als Ausreißer identifiziert (*swamping*) werden.

Eine gute Erläuterung der notwendigen Kompensation dieser Effekte findet sich im Übersichtsartikel von [Beckman, R. J. und Cook, R. D., 1983]. Wichtig ist dabei auch die Feststellung, daß die Ursache für beide Effekte nicht in den Daten liegt (dort wird es immer mehr als einen Ausreißer geben, wählt man die Stichprobe groß genug), sondern einzig und allein von den Identifikationsmethoden auf der Suche nach Ausreißern (die auf jedem Iterationsschritt immer nur nach einem Ausreißer

suchen, alles andere würde zum Beispiel die mehr oder weniger unbegründete Annahme beinhalten, nur nach Tupeln oder Tripeln zu suchen) verursacht wird. Masking und Swamping wird also von den Detektoren eingebracht, nicht von den Daten verursacht.

Daher sind die Effekte unterschiedlich, wenn unterschiedliche Ausreißerdetektoren verwendet werden. Diesen Effekten tragen robuste Schätzer, mehrstufige Verfahren und vor allem Verfahren, die in unterschiedliche Richtungen suchen und verifizieren, Rechnung [Hawkins, 1985].

Unter einflußreichen Beobachtungen versteht man Teilmengen von Datensätzen, deren Streichung substantielle Veränderungen in Schätzern, Konfidenzregionen, Testverfahren, Modellparametern oder Modellpräferenzen zu Folge hat. Solche Teilmengen befinden sich in der Regel an den Rändern der Mehrheit der Datenpunkte, denn gerade dort werden Modelle und deren Parameter (Exponenten, Ausläufe von Kurven) gestützt. Hilfreich ist die Definition der Hebelwerte (engl. *leverage values*)

$$\mathbf{x} (\mathbf{x}^T \mathbf{x})^{-1} \mathbf{x}^T$$

Diese Größe ist eng mit der Mahalanobis-Distanz verbunden und kann als Maß für den Abstand vom Zentroid der Daten angesehen werden. Punkte nah am Zentrum tragen weniger Information für die Modellbildung als Punkte, die weiter außen liegen. Andererseits sind Ausreißer gerade dort besonders gefährlich.

Ausreißer und einflußreiche Größen lassen sich nicht eindeutig trennen, obwohl wir erstere gerne völlig eliminieren würden, letztere aber gerade die gesuchten Modellstützen (man denke nur an Support-Vektor-Verfahren) und eigentlichen Informationsträger einer Messung sind. Oft identifizieren Ausreißerdetektoren einflußreiche Meßwerte und anders herum werden bei der Einflußanalyse nachträglich potentielle Ausreißer, zumindest unter dem gerade angenommenen Modell, identifiziert. Dennoch stehen selbst innerhalb der Statistik beide Gebiete sehr getrennt und lassen rückgekoppelte oder parallele Verfahren vor allem für multivariate Daten vermissen.

Die frühen Verfahren zur Ausreißerdetektion wurden von Grubbs noch einmal in [Grubbs, 1969] mit zahlreichen Beispielen zusammengefaßt. Darin werden auch Arbeiten von [Dixon, 1953] diskutiert, die die Schätzung der Varianz mit s z. B. in (2.1) vermeiden. Auch werden Methoden vorgestellt, bei denen in der *ESD* die Varianz aus anderer Quelle mit verschiedenen Freiheitsgraden geschätzt wird oder tatsächlich die Varianz σ bekannt ist und anstelle von s verwendet wird. Es ergeben sich damit neue kritische Werte, die Grubbs tabelliert. Für die bisher vorgestellten Verfahren mußte die Zahl k der gesuchten Ausreißer stets bekannt sein, meist werden ein oder zwei Ausreißer gesucht. [Grubbs, 1969] verweist dazu auf Methoden, die anhand der Kurtosis von Daten die Optimalität der Detektoren für ein oder zwei Ausreißer beschreiben. Rosner [Rosner, 1975] schlug zum ersten mal Verfahren vor, die Schritt für Schritt auf $1, \dots, k$ Ausreißer prüfen und nicht nur auf genau k Ausreißer. Als eingebettete Teststatistik identifiziert Rosner die *ESD* als beste der bisher genannten. Er legte mit dieser Arbeit auch die Grundlage für die später (Abschnitt 2.2.3ff.) entwickelten Vorwärts- bzw. Rückwärtsstrategien (auch *inward* und *outward*), womit auch multivariate Datensätze erstmals sicher zu untersuchen sind. Dies rechtfertigt sicher auch den Sprachgebrauch, seit Rosner von “modernen Ausreißerdetektoren” zu sprechen. Unter *inward* versteht man von außen nach innen wandernde Methoden, die auf einen einzelnen Ausreißer testen, diesen herausnehmen und den verkleinerten Datensatz erneut testen, bis kein Ausreißer mehr gefunden wird. *outward*-Methoden suchen ähnlich wie *inward*-Methoden nach Ausreißern, prüfen aber eine Gruppe von potentiellen Ausreißern in einem

zweiten Schritt mit einer anderen Teststatistik von innen nach außen heraus: also vom um die Gruppe reduzierten Datensatz aus. Gegebenenfalls muß ein Kandidat, der als Nicht-Ausreißer identifiziert wurde, in den somit wieder anwachsenden reduzierten Datensatz zurückgelegt werden.

Davies und Gather [Davies, Laurie und Gather, Ursula, 1993] schlagen ein Hypothesentestverfahren vor, bei dem es nicht mehr darauf ankommt, eine bestimmte Anzahl (oder bis zu einer bestimmten Anzahl) von Ausreißern zu finden: alle in einem Bereich niedriger Wertewahrscheinlichkeit liegenden Daten werden als Ausreißer interpretiert (α -Ausreißer-Konzept). Der Ansatz dafür ist bereits einleitend in diesem Abschnitt genannt worden; als Teststatistik kommt der Betrag $|\tau_n|$ des Thompson-Tests (2.1) zu Einsatz. Ohne die Verteilung der Ausreißer überhaupt näher zu kennen oder kennen zu wollen, wird mit einer Signifikanz α die Menge der

$$\{x_n \in X \text{ mit } |x_n - \bar{x}| > \sigma\theta_\alpha\}$$

als Ausreißer der $N(\bar{x}, \sigma)$ -verteilten Daten bezeichnet. Die Schwelle θ_α entspricht dabei der Integrationsgrenze, bis zu der die $N(0, 1)$ -Gaußkurve beidseitig aus dem Unendlichen aufintegriert werden muß, um den Integralwert α zu erhalten:

$$\alpha \stackrel{!}{=} \frac{2}{\sqrt{2\pi}} \int_{\theta_\alpha}^{\infty} e^{-\frac{\theta^2}{2}} d\theta$$

Der Wert α wird anfangs festgelegt, aus ihm ergeben sich die Ausreißer-Regionen $[-\infty, \bar{x} - \theta_\alpha\sigma]$ und $[\bar{x} + \theta_\alpha\sigma, +\infty]$ für die Daten x_n . Man beachte, daß die Annahme, X sei normalverteilt weder vor noch nach der Ausreißerentnahme zutrifft: in der Ausreißer-Region werden zwangsläufig immer auch "gute Daten" liegen, die der üblichen in normalverteilten Daten inhärente Variabilität entsprechen, und dann fälschlicherweise als Ausreißer klassifiziert werden. Die Wahrscheinlichkeit, daß ein einzelner Wert x_n oder mehrere Werte x_1, \dots, x_o gleichzeitig nicht in der Ausreißer-Region liegen, ist

$$\begin{aligned} P(x_n \in [\bar{x} - \theta_\alpha\sigma, \bar{x} + \theta_\alpha\sigma]) &= 1 - \alpha \\ P(x_n, \dots, x_o \in [\bar{x} - \theta_{\alpha_o}\sigma, \bar{x} + \theta_{\alpha_o}\sigma]) &= 1 - \alpha \\ \text{mit } \alpha_o &= 1 - (1 - \alpha)^{1/o} \end{aligned}$$

Auch wird in [Davies, Laurie und Gather, Ursula, 1993] der Begriff des Bruchpunktes eines Verfahrens bei Maskierung angegeben und für zahlreiche Teststatistiken berechnet.

Es wird allgemein davon ausgegangen, daß Meßdaten normalverteilt sind. Der allgemeine Grenzwertsatz und die Vielzahl an zufällig und sicher annähernd normalverteilten Einflußgrößen auf eine Messung legt nahe, daß die Meßwerte selber in hohem Maße normalverteilt sind. Dies kann vor allem für univariate Daten einfach getestet werden mit Verfahren wie dem Shapiro-Wilks-Test, dem Kolmogorov-Smirnov-Test, an Histogrammen und Quantile-Quantile-Plots.

Aber nicht immer läßt sich die Verteilung der Daten durch eine Gaußkurve anpassen. Bei Daten, die zum Beispiel nicht unter oder über bestimmte Grenzen gehen können (Zeitintervalle können nicht kleiner werden als Null, Konzentrationen nicht größer als 100% u.s.w.) begegnet man mit ln-Normalverteilungen modifizierten Gaußkurven. Aber darüber hinaus gibt es auch Daten, die grundsätzlich anderen Verteilungen folgen. So sind Risiken zwar normalverteilt, Preisschwankungen und Renditeverteilungen von Finanzprodukten zeigen aber *heavy tails* und gleichzeitig *high peakedness*. Auch bei

Kommunikationslasten in Computer- oder Telefonnetzwerken treten Verteilungen mit *heavy tails* auf und wurden in den 90er Jahren zu einem aktuellen Forschungsgebiet. Solche “schweren Enden” in Meßdaten können einerseits grundsätzlich als Ausreißer interpretiert werden, was der Normalverteilung weiterhin einen gewissen Alleinvertretungsanspruch für die Verteilung von Meßdaten allgemein einräumt, oder als Hinweis auf eine grundsätzlich neuartige und von der Gaußkurve verschiedene Verteilung angesehen werden. Die Wahrheit wird sicher zwischen beiden extremen Interpretationen liegen und letztlich ist die Frage nach den tatsächlichen Ursachen von schweren Enden auch die Frage nach den Ursachen von Ausreißern.

[Beckman, R. J. und Cook, R. D., 1983], [Barnett, Vic und Lewis, Toby, 1994] und [Gather, Ursula und Becker, Claudia, 1997] geben gute Übersichten zu den Typen von Ausreißerdetektoren und [Gather, Ursula und Becker, Claudia, 1997] mißt an mehreren Kriterien deren Leistungsstärke auch bezüglich Maskierung und Überschwemmung. Bei [Anscombe, 1960] findet sich ein historischer Überblick zu Verfahren vor Thompson.

2.2.3 Simonoff-Methode zur Ausreißerdetektion

Der oben beschriebene Maskierungseffekt kann ganz erheblich reduziert und vermindert werden, indem Such- und Selektionsstrategien in verschiedene Richtungen benutzt werden. Herausragendes Verfahren zur Ausreißerdetektion ist die Simonoff-Methode, die zurück geht auf [Simonoff, 1992] und [Rosner, 1975]. Sie ist ein Vorwärts-Rückwärts-Verfahren, das pro Schritt einen einzelnen Ausreißer sucht oder bestätigt. Unter dem Überbegriff *forward-backwards* (oder auch Rückwärts-Testen, engl. *backwards-stepping*) versteht man dabei, daß potentielle Ausreißer “vorwärts” gesucht werden (die größten Abweichungen vom allgemeinen Mittel der Daten werden zuerst gefunden und dann zu kleineren Abweichungen hinabgestiegen), aber verdächtige Datenpunkte “rückwärts” getestet werden (also immer vom verdächtigen Datenpunkt mit dem geringsten Abstand ausgehend unter Einbezug aller anderen Kandidaten mit höherer Abweichung). Vorwärts-Vorwärts-Verfahren gelten als anfällig für den Maskierungseffekt; multiple Detektoren (suchen nach Paaren oder Tripeln von Ausreißern pro Iterationschritt) gelten als anfällig für Maskierung und Überschwemmungseffekte [Hawkins, 1985].

Gegeben sei eine Punktmenge $S_0 = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ aus n Messungen eines p -dimensionalen Vektors $\mathbf{x}_i = (x_{i,1}, \dots, x_{i,p})^T$, in der K Ausreißer und $(n - K)$ “gute” Messungen vermischt sind. Die Größe K ist natürlich unbekannt; oft würde es für die Versuchsauswertung bereits reichen, ihren Wert oder eine obere Schranke für ihren Wert zu kennen, auch ohne die einzelnen Ausreißer genau zu benennen. Wir bezeichnen daher im Folgenden mit k die Anzahl Ausreißer, die wir maximal in S_0 vermuten. Die Werkzeuge für die Simonoff-Methode sind:

- ein Maß zur Identifikation von potentiellen Ausreißern, das in der Vorwärtssuche den Abstand eines einzelnen Datenpunktes gegen alle anderen bestimmt,
- ein Abbruchkriterium k , das festlegt, wieviele Ausreißer maximal gesucht werden sollen,
- eine Ordnung, die alle potentiellen Ausreißer in eine Reihenfolge von “wenig extrem” bis “sehr extrem”, bringt,

- ein Testverfahren, das beim Rückwärtstesten schrittweise den verdächtigen Datenpunkt mit der geringsten Ordnung auf seine Ausreißereigenschaften prüft,
- ein Kriterium für den Rückwärtstest, das die Entscheidung “Ausreißer” oder “keine Ausreißer” fällt.

Das bemerkenswerte an der Simonoff-Methode und den darin enthaltenen unabhängigen Vorwärts-Rückwärts-Verfahren ist, daß ein im Rückwärts-Testen gefundener Ausreißer alle anderen verdächtigen Datenpunkte mit höherer Ordnung zu Ausreißern macht – unabhängig davon, ob sie einzeln getestet als echte Ausreißer identifiziert worden wären oder nicht. Hier zeigt sich, wie wichtig die Ordnung der potentiellen Ausreißer und damit ihre Reihenfolge beim Rückwärtstesten ist. Im Folgenden bezeichnet der Index i die Vorwärtsrichtung, der Index j die Rückwärtsrichtung. Die genaue Vorgehensweise der Simonoff-Methode für univariate und multivariate Daten ist:

- Man bildet die Teilmengen \mathcal{S}_i mit $i = 1, \dots, k$, indem aus der Teilmenge \mathcal{S}_{i-1} der jeweils extremste Datenpunkt $\mathbf{z}^{[i]}$ entfernt wird. Im Sinn der Vorwärtsselektion wird den Teilmengen bei jedem Schritt ein weiterer Datenpunkt dauerhaft entfernt; die Teilmengen \mathcal{S}_i hat dann $n - i$ Elemente und $\mathcal{S}_i \cup \{\mathbf{z}^{[i]}\} = \mathcal{S}_{i-1}$.
- Für jede der k Teilmengen \mathcal{S}_i wird eine Teststatistik \mathcal{T}_i berechnet, die Auskunft gibt, wie stark der zugehörige und gerade entnommene Datenpunkt $\mathbf{z}^{[i]}$ vom Rest (ohne die bereits zuvor herausgenommenen Datenpunkte) entfernt ist.
- Die Teststatistik \mathcal{T}_j wird jetzt rückwärts $j = k, \dots, 1$ mit dem kritischen Wert t_c verglichen. Wenn keine Ausreißereigenschaft nachgewiesen werden kann, wird j dekrementiert. Bestätigt der Vergleich die Ausreißereigenschaft auf der Stufe j_t , so sind alle potentiellen Ausreißer $\mathbf{z}^{[i]}$ mit $i = 1, \dots, j_t$ als Ausreißer bestätigt. Auch wenn darunter vielleicht einzelne Stufen $j < j_t$ sind, bei denen die jeweilige Teststatistik den Verdacht nicht in einen tatsächlichen Ausreißer überführen würde. Es reicht aus, daß es auf einer tieferen Stufe bereits einen Ausreißer gibt. In diesem Schritt liegt die Vermeidung des Maskierungseffektes, der ja gerade bewirken würde, daß ein Ausreißer auf einer höheren Stufe durch die Existenz von Ausreißern in niedrigeren Stufen (also in der Teilmenge, gegen die er verglichen wurde) eine solch geringe Ausprägung bekommt (eben maskiert wird durch die starken Abweichungen auf niedrigeren Stufen), daß er in der Teststatistik seiner Stufe den kritischen Wert nicht erreicht und folglich nicht erkannt wird. Eine reine Vorwärtsprüfung hätte weder den maskierten Ausreißer identifiziert, noch den auf einer tieferen Stufe liegenden Ausreißer, der diese Maskierung verursacht hat.

Für die Identifikation verdächtiger $\mathbf{z}^{[i]}$, deren Anordnung und die Teststatistiken \mathcal{T}_i gibt es zahlreiche Verfahren, die von der Dimension der Daten und der jeweiligen Anwendung abhängen. Aufgrund des Zentralen Grenzwertsatzes kann bei den meisten Datensätzen von einer Normalverteilung ausgegangen werden. Simonoff verwendet auf Normalverteilungen die extreme studentisierte Abweichung (engl. *extreme studentised deviate*, *ESD*) bzw. die Mahalanobis-Distanz als Identifikator, ein *Single-Linkage-Clustering* zur Anordnung und die Student-Verteilung und F-Verteilung für die Bestimmung der kritischen Werte. Die Verfahren sollen im einzelnen vorgestellt werden.

Als typisches Maß für die Identifikation potentieller Ausreißer in univariaten Daten im Vorwärts-Identifizieren und gleichzeitig als Teststatistik im Rückwärts-Testen hat sich die extreme studentisierte Abweichung etabliert, die die Abweichung eines Datenpunktes vom Mittelwert des Datensatzes auf dessen Standardabweichung normiert.

$$d_l = \frac{x_l - \langle x_m \rangle_m}{\left\langle (x_n - \langle x_m \rangle_m)^2 \right\rangle_n^{1/2}} \quad (2.2)$$

Die Mittelungen über die Indices m und n laufen dabei über den gerade zur Prüfung stehenden Datensatz \mathcal{S}_i von $m, n = 1, \dots, n-i$ für alle $l = 1, \dots, n-i$ Datenpunkte. Die dazugehörige Verteilung ist die Student-Verteilung mit dem Freiheitsgrad $n-1$, wodurch sie sich wegen typischerweise großen n stark einer Normalverteilung $N(0, 1)$ annähert. Für multivariate Datensätze verallgemeinert (2.2) in die Mahalanobis-Distanz (mit der Streumatrix \mathbf{S} normiertes euklidisches Skalarprodukt)

$$D_l^2 = (\mathbf{x}_l - \langle \mathbf{x}_m \rangle_m)^T \mathbf{S}^{-1} (\mathbf{x}_l - \langle \mathbf{x}_m \rangle_m) \quad \text{mit} \quad \mathbf{S} = \frac{1}{n-1} \sum_{l=1}^n (\mathbf{x}_l - \langle \mathbf{x}_m \rangle_m) (\mathbf{x}_l - \langle \mathbf{x}_m \rangle_m)^T \quad (2.3)$$

Unter der Jackknife-Mahalanobis-Distanz versteht man eine Berechnung von (2.3), wobei aus Mittelung $\langle \rangle_m$ und S der jeweilige Datenpunkt \mathbf{x}_l herausgenommen wurde.

Im Datensatz gilt es nun, einzelne Datenpunkte oder kleine Gruppen als potentielle Ausreißer zu identifizieren, die extreme Abstände d_l bzw. D_l haben.

Im univariaten Fall stellen wir eine nach den Stufen $i = 1, \dots, k$ geordnete Tabelle der Werte $|d_i|$ mit den dazu gehörenden gerade entnommenen verdächtigen Datenpunkten $z^{[i]}$ auf, wobei darauf zu achten ist, daß beim Übergang zu jeder neuen Stufe für den verbliebenen Datensatz \mathcal{S}_i nach Entnahme der potentiellen Ausreißer das Mittel und die Standardabweichung in (2.2) neu zu berechnen sind. Dabei kann für $i_2 > i_1$ durchaus $d_{i_2}^2 < d_{i_1}^2$.

Gruppen oder einzelne extreme Vektoren lassen sich bei multivariaten und in der Regel hochdimensionalen Datensätzen nicht mehr durch einfaches Anschauen der Zahlenwerte oder Abstände in einzelnen Dimensionen identifizieren. Auch gilt es, über verschiedene Dimensionen hinweg extreme Abstände im Rahmen der Vorwärts-Identifikation in eine eindeutige hierarchische Ordnung zu bringen. Bei multivariaten Daten kann im Gegensatz zu univariaten Daten bereits die Ordnung durch Maskierung beeinflusst werden. Es ist also ein Ordnungsverfahren gesucht, das nicht ausschließlich nach Mahalanobis-Distanzen sortiert und diesen Effekt bei Maskierungen unterbindet. Das Single-Linkage-Clustering-Verfahren ist ein solches exploratives Verfahren zur Aufdeckung von Gruppierungen, Konzentrationen oder Inselbildungen in Datensätzen. Rohlf [Rohlf, 1975] zeigte, daß dieses Verfahren eine multivariate Erweiterung des Lückentests (engl. *gap test*) für Zufallszahlengeneratoren ist. Als graphische Veranschaulichung der Ergebnisse dient ein sogenanntes Dendrogramm (Abbildung 2.2).

Die Vorgehensweise ist einfach: es werden ausgehend von allen Datenpunkten agglomerativ (im Unterschied zu divisiven Clustering-Methoden, die den gesamten Datensatz in immer kleinere Gruppen unterteilen) jeweils die zwei Datenpunkte oder bereits zu Gruppen zusammengefaßte Datenpunkte

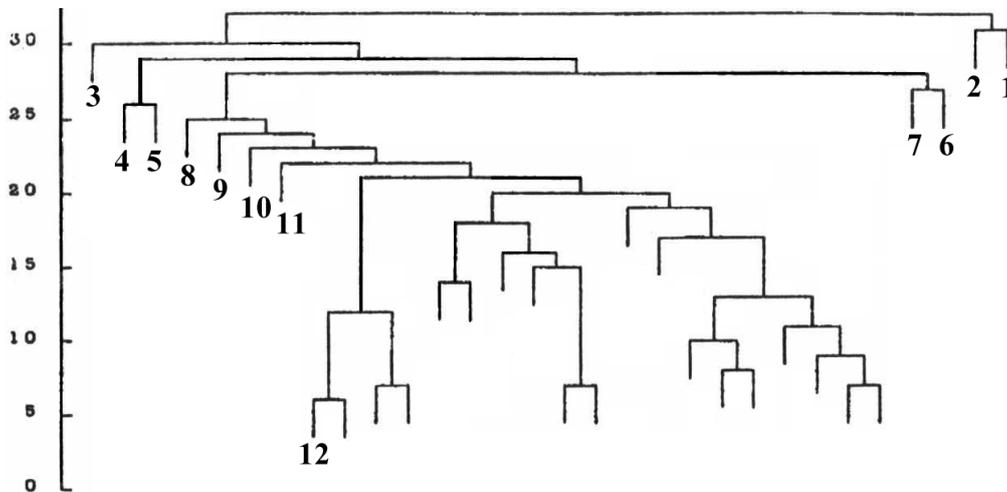


Abbildung 2.2: Dendrogramm mit markierten Ausreißerkandidaten $\{z^{[1]}, \dots, z^{[12]}\}$ aus Simonoff [Simonoff, 1992]. Die Achse mit den Werten der Metrik wurde abschließend umskaliert.

zu einem Cluster vereint, die den geringsten Abstand haben. Am Ende des Verfahrens sind aufsteigend hierarchisch alle Datenpunkte und die dabei entstandenen Cluster zu einem einzigen Baum zusammengefaßt. Vorteile des hierarchischen Verfahrens sind die Unabhängigkeit von der Anzahl der Klassen und das Sichtbarwerden innerer Strukturen in den Daten: bei reiner Zuordnung zu Clustern würde diese zusätzliche Information fehlen. Nachteile sind, daß getroffene Zuordnungsentscheidungen nicht nachträglich modifiziert werden können, ohne die gesamte oberhalb folgende Struktur neu berechnen zu müssen. Dendrogramme bieten bei sehr großen Datensätzen keine Informationsreduktion und werden unhandlich. Diese Nachteile werden im BIRCH-Verfahren teilweise behoben, dort werden statistische Größen eingeführt und durch iterative Modifikationen die Qualität des Baumes erhöht. Je später ein Datenpunkt, zu einem Cluster hinzutritt oder je später ein kleiner Untercluster mit anderen vereint wird, umso höher ist die Wahrscheinlichkeit, daß es sich dabei um Ausreißer handelt. Daher wird die gesuchte Ordnung potentieller Ausreißerkandidaten anhand des entstandenen Baumes rückwärts aufgeschlüsselt. Als Ausreißerkandidat mit höchster Priorität gilt der Datenpunkt der als letztes zu den anderen Clustern hinzugekommen ist. Wird auf der untersuchten Ebene nicht ein einzelner Datenpunkt zusammengefaßt, sondern tritt ein eigener kleiner Cluster zu dem (stets größer gewählten) Hauptcluster hinzu, so wird in dem kleinen Cluster der einzelne Datenpunkt als Ausreißerkandidat hoher Priorität notiert, der den größten Abstand zum Hauptcluster hat. Verzweigt der untersuchte Untercluster weiter in kleinere Cluster, so wird die Suche mit den genannten Prioritäten auf niedrigerer Ebene fortgesetzt. Auf diese Weise läßt sich der gesamte Baum als nach Vereinigungsabständen geordnete Liste schreiben. Die obersten k Einträge der Liste sind die für die Simonoff-Methode gesuchten verdächtigen Ausreißer, die anschließend rückwärts-getestet werden.

Es wird nun ein Abstandsmaß von Clustern und eine Metrik im Datenraum benötigt. Als Abstandsmaß wird für das Single-Linkage-Clustering-Verfahren nicht die Mahalanobis-Distanz benutzt, da in die Schätzer für Kovarianzmatrix und Mittelwert die Ausreißer eingehen und daher zu einer

vertauschten Anordnung führen können. Es wird folgende Metrik benutzt: der Abstand δ_{C_1, C_2} von zwei Clustern C_1 und C_2 (einzelne Datenpunkte sind Cluster der Größe 1) ist durch den geringsten Abstand von zwei Datenpunkten in den beiden Clustern gegeben.

$$\delta_{C_1, C_2} = \min_{u, v} \{d_{u, v} \text{ mit } \mathbf{x}_u \in C_1, \mathbf{x}_v \in C_2\} \quad (2.4)$$

Dieser Cluster-Abstand trägt auch den Namen Nächster-Nachbar-Abstand (engl. *nearest neighbor*), siehe auch Abschnitt 3.5.4. Andere Möglichkeiten sind das Complete-Linkage-Clustern mit dem entferntesten Nachbarn (*furthest-neighbor*), das Centroid-Clustern mit Abständen der Clustermittelpunkte (Mittelung oder Median), die Ward-Methode (fusioniert als hierarchisches Verfahren sukzessive diejenigen Cluster, mit deren Fusion die geringste Erhöhung der gesamten Fehlerquadratsumme einhergeht). Das Single-Linkage-Clustering neigt zu wenigen großen Gruppen und zur Kettenbildung, also der späten Hinzunahme von einzelnen Datenpunkten in bereits angewachsene Cluster, was der Ausreißeranalyse sehr entgegenkommt. Complete-Linkage-Clustering bildet eher viele kleine Gruppen, es tendiert zu einer ‘‘Verinselung’’ und versucht gleich starke Gruppen zu bilden. Centroid-Clustering beschreitet einen konservativen Mittelweg. Sein Vorteil gegenüber anderen Verfahren ist auch, daß die resultierenden Dendrogramme einfach zu interpretieren sind und das Verfahren keine Festlegung auf eine bestimmte Anzahl von Clustern verlangt.

Die für die Abstandsberechnung von zwei Datenpunkten oder Schwerpunkten im Datenraum benötigte Metrik $d_{u, v}$ ist eine Minkowski-Metrik, die meist als euklidische Metrik mit $q = 2$ benutzt wird.

$$d_{u, v} = \left(\sum_{r=1}^p |x_{u, r} - x_{v, r}|^q \right)^{\frac{1}{q}} \quad (2.5)$$

Die Verwendung dieser Metrik über alle p Dimensionen des Datenraumes verlangt eine Transformation der Datenmatrix $\mathbf{X} \rightarrow \mathbf{X} \hat{\Sigma}^{-1/2}$, wobei $\hat{\Sigma}$ ein Schätzer (sinnvollerweise ein robuster Schätzer nach [Rousseeuw, Peter J. und Leroy, Annik M., 1987]) für die Varianz $\Sigma = \text{Var}(\mathbf{X})$ ist.

Anschließend liegen die k potentiellen Ausreißer in einer Ordnung vor, auf die nun das Rückwärts-Testen angewendet wird.

Bei statistischen Tests wird eine Hypothese, die Null-Hypothese \mathcal{H}_0 , auf ihr tatsächliches Zutreffen geprüft. Dabei kann es zu Fehlentscheidungen kommen. Unter einem Fehler erster Art (engl. *error type I*) versteht man ein unberechtigtes Ablehnen der Null-Hypothese: \mathcal{H}_0 ist in Wirklichkeit wahr, der Test hat aber \mathcal{H}_0 fälschlicherweise abgelehnt. Dieser Fehlertyp wird auch α -Fehler oder Irrtumswahrscheinlichkeit genannt. Unter einem Fehler zweiter Art (engl. *error type II*), oder β -Fehler versteht man ein Beibehalten der Null-Hypothese \mathcal{H}_0 , obwohl sie in Wirklichkeit nicht zutrifft. Der Wert α für einen Hypothesen-Tests gibt also an, daß man zu $100\alpha\%$ bereit ist, einen Fehler vom Typ I zu riskieren. Gibt es zur Null-Hypothese nur das Gegenstück ‘‘Die Null-Hypothese gilt nicht’’ so spricht man statt von einer Eins-Hypothese \mathcal{H}_1 oft auch einfach von einer Alternativ-Hypothese oder Gegen-Hypothese.

Als Null-Hypothese formulieren wir, daß kein Ausreißer gefunden wurde. Dies bedeutet beim Rückwärts-Testen auf genau k mögliche Ausreißer, daß eben nicht k , sondern weniger Ausreißer

(was auch beinhaltet, daß es überhaupt keine Ausreißer gibt) vorhanden sind. Die Null-Hypothese lautet: *Es befinden sich $(k-1)$ Ausreißer im Datensatz.* Die dazu gehörende Gegen-Hypothese ist: *Es befinden sich k Ausreißer im Datensatz.*

Die ESD (2.2) bzw. Mahalanobis-Distanz (2.3) wird in der Simonoff-Methode auch als Teststatistik benutzt. Um sie gegen einen kritischen Wert t_c zu vergleichen, sind typische unter einer gewissen noch erlaubten Fehlzuordnung mögliche Quotienten von Standardabweichungen ausschlaggebend. Sie sind abhängig von der Anzahl Freiheitsgrade, die sich aus der Dimension und den Größen der Teilmengen S_i ergeben. Solche Quotientenwerte liefert für multivariate Datensätze die F-Verteilung. Diese hat als Parameter den α -Fehler und die beiden Freiheitsgrade der miteinander zu vergleichenden Stichproben p und $n-p$. Eine wichtige Eigenschaft der F-Verteilung ist $F^{-1}(1-\alpha) = F(\alpha)$.

Für den univariaten Fall ($p=1$, ein Freiheitsgrad wird zu eins) geht die F-Verteilung in das Quadrat t_{n-1}^2 der Student-Verteilung mit dem verbleibenden zweiten Freiheitsgrad $n-p=n-1$ über. Der kritische Wert t_c ist dann gegeben durch die nach α und dem zweiten Freiheitsgrad tabellierten Signifikanzschranken der Student-Verteilung zu einer gegebenen Irrtumswahrscheinlichkeit α für einen zweiseitigen Test.

Die Wahrscheinlichkeit, bei mehreren Untersuchungen auf das Zutreffen einer Hypothese aus Zufall einem Irrtum im Sinne von α (Fehler Typ I) zu erliegen, wird durch die Entwicklung des Binoms $(\alpha+(1-\alpha))^s$ gegeben, wenn s die Anzahl der Untersuchungen ist. Die Bonferroni-Ungleichung approximiert diese Entwicklung und sagt, daß die Wahrscheinlichkeit, eine oder mehrere der s Null-Hypothesen im Sinne von α fälschlicherweise zu verwerfen, nicht größer ist als die Summe der Irrtumswahrscheinlichkeiten $s\alpha$. In unserem Fall bedeutet dies, daß wir bei der Bestimmung der kritischen Werte t_c mit der F-Verteilung von α zu α/n übergehen können, und mit der Bonferroni-Ungleichung das Binom mehrfacher Untersuchungen sehr konservativ approximieren.

Bei univariaten Daten muß die mit den ESD (2.2) aufgestellte Tabelle der k potentiellen Ausreißer nun rückwärts getestet werden. Als Teststatistik dienen die d_j , die gegen $t_c = F_{t[n-1]}^{-1}(1-\alpha/n)$ von $j = k$ bis zu $j = 1$ getestet werden. Gilt für ein bestimmtes j_t die Ungleichung $t_c < d_{j_t}$, so ist die Eins-Hypothese "*Es befinden sich j_t Ausreißer im Datensatz.*" bestätigt und der Test wird mit j_t gefundenen Ausreißern $\{z^{[1]}, \dots, z^{[j_t]}\}$ abgebrochen. Andernfalls gilt die Null-Hypothese und j wird zur nächsten Stufe dekrementiert.

Wie bei den univariaten Daten werden die Mahalanobis-Distanzen der einzelnen Ausreißerkandidaten $\mathbf{z}^{[i]}$ beginnend mit dem extremsten Kandidaten $\mathbf{z}^{[1]}$ für jede Stufe mit ihrer Teilmenge S_i mit $i = 1, \dots, k$ vorwärts neu berechnet, nachdem der potentielle Ausreißer der vorangegangenen $(i-1)$ -Stufe entfernt wurde. Bei multivariaten Daten wird die Mahalanobis-Distanz (2.3) rückwärts ($j = k, \dots, 1$) gegen die $F_{p,n-p,\alpha/n}$ Verteilung getestet (auch T^2 -Test genannt).

$$D_j^2 \begin{cases} \geq \\ < \end{cases} \frac{p(n-1)}{n-p} F_{p,n-p,\alpha/n} \begin{cases} \text{Es gibt } j \text{ Ausreißer.} \\ \text{Es gibt } \leq (j-1) \text{ Ausreißer.} \end{cases} \quad (2.6)$$

Penny [Penny, 1996] hat dazu folgende Korrektur vorgeschlagen, die in der Literatur als Wilks-Wert bekannt ist und sich allgemein als besserer kritischer Wert durchgesetzt hat:

$$t_c = \frac{p(n-1)^2 F_{p,n-p-1,\alpha/n}}{n(n-p-1 + pF_{p,n-p-1,\alpha/n})}$$

Für kleine n weicht dieser stark von Simonoffs ursprünglichem kritischen Wert ab. Penny berechnet auch die korrekten Vorfaktoren für den Wilks-Wert der Jackknife-Mahalanobis-Distanzen.

So mächtig die Fähigkeit der Simonoff-Methode zur Vermeidung von Maskierungen und Überschwemmungen bei der Detektion von Ausreißern ist, ihr großer Nachteil ist die Notwendigkeit, a priori eine Anzahl der vermuteten k Ausreißer vorzugeben. Simonoff rät zu einem konservativen Wert von $k=(n-1)/2$ und nimmt dabei ein “Über-Testen” in Kauf.

2.2.4 Wilks-Methode zur Ausreißerdetektion

Die Ausreißer-Streuungs-Quotienten von Wilks [Wilks, 1963] sind eines der ersten formalisierten und heute bereits ein klassisches Verfahren zur Ausreißersuche. Caroni [Caroni, Chrys und Prescott, Philip, 1992] haben den Test in ein sequenzielles Verfahren zur Suche nach k Ausreißern in multivariaten Daten erweitert. Da die Wilks-Methode mit diesen Erweiterungen die Maskierungseffekte und Überschwemmungseffekte unterdrücken kann und eine multivariate Erweiterung vorliegt, sei diese Methode hier vorgestellt.

Die Teststatistik der Wilks-Methode ist ein Likelihood-Verhältnistest und prüft, ob ein Datenpunkt zu einer p -dimensionalen Normalverteilung $N_p(\langle \mathbf{x}_m \rangle_m, \Sigma)$ gehört oder zu einer verschobenen Normalverteilung $N_p(\langle \mathbf{x}_m \rangle_m + \mathbf{a}, \Sigma)$. Die genaue Kenntnis der Parameter Mittelwert, Verschiebung (engl. *slippage*) und Varianz ist dabei nicht nötig. In der sequenziellen Formulierung der Methode lautet die Null-Hypothese also auf Zugehörigkeit aller $\mathbf{x}_i \in \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ zu $N_p(\langle \mathbf{x}_m \rangle_m, \Sigma)$

$$\text{Null-Hypothese } \mathcal{H}_0 : \mathbf{x}_i \sim N_p(\langle \mathbf{x}_m \rangle_m, \Sigma) \quad \text{für } i = 1, \dots, n$$

Die Gegen-Hypothese besagt, daß es bis auf den Ausreißer mit dem Index l , der zu einer mittelwertverschobenen Normalverteilung gehört, alle anderen Datenpunkte in die Verteilung der Null-Hypothese fallen.

$$\text{Gegen-Hypothese } \mathcal{H}_1 : \mathbf{x}_i \sim N_p(\langle \mathbf{x}_m \rangle_m, \Sigma) \quad \text{für } i \neq j \quad \text{und} \quad \mathbf{x}_j \sim N_p(\langle \mathbf{x}_m \rangle_m + \mathbf{a}, \Sigma)$$

Wilks definiert eine Streumatrix \mathbf{A} ähnlich wie in (2.3) aus allen Quadraten und Kreuzprodukten des Datensatzes.

$$\mathbf{A} = \sum_{i=1}^n (\mathbf{x}_i - \langle \mathbf{x}_m \rangle_m) (\mathbf{x}_i - \langle \mathbf{x}_m \rangle_m)^T \quad (2.7)$$

Die Matrix $\mathbf{A}^{[l]}$ wird analog über $(n-1)$ Datensätze ohne den Ausreißer \mathbf{x}_l gebildet. Das Likelihood-Verhältnis aus den Determinanten beider Matrizen dient als Teststatistik. Der Ausreißerkandidat \mathbf{x}_l wird identifiziert als der Datenpunkt, der das Verhältnis auf den Wert D_l minimiert

$$D_l = \min_l (W_l) = \frac{\det \mathbf{A}^{[l]}}{\det \mathbf{A}} = 1 - \frac{n}{n-1} (\mathbf{x}_l - \langle \mathbf{x}_m \rangle_m)^T \mathbf{A}^{-1} (\mathbf{x}_l - \langle \mathbf{x}_m \rangle_m) \quad (2.8)$$

Eine geometrische Veranschaulichung gelang Wilks über die Simplex-Volumen. Aus den n Datensätzen lassen sich $\binom{n}{p}$ verschiedene Teilmengen mit p Elementen bilden, die jeweils einen Simplex mit einem Schwerpunkt $\langle \mathbf{x}_m^{(p)} \rangle_m$ bilden. Die Summe der quadrierten Volumina aller möglichen Simplexes ergibt $(p!)^{-2} \det \mathbf{A}$. Die Teststatistik sucht nach dem Datenpunkt, der diese Volumensumme am stärksten verkleinert – also nach dem Datenpunkt, der als einzelner bei Hinzunahme das Volumen am stärksten aufbläht. Es liegt auf der Hand, diesen Datenpunkt als Ausreißerkandidaten zu identifizieren.

Man beachte die Analogie in (2.8) zur Mahalanobis-Distanz. Wichtige Eigenschaften des Likelihood-Verhältnisses sind unter der Voraussetzung, daß die Null-Hypothese gilt

$$0 \leq W_i \leq 1 \quad \text{und} \quad \sum_{i=1}^n W_i = n \left(1 - \frac{p}{n-1}\right) \quad \text{und} \quad \langle W_i \rangle_i = \frac{n-p-1}{n-1} \quad \text{und} \quad \text{Var } W_i = \frac{2p(n-p+1)}{(n-1)(n+1)}$$

$$W_i \sim B\left(\frac{n-p-1}{2}, \frac{p}{2}\right) \quad \text{mit der Eulerschen Beta-Funktion:} \quad B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$$

Das Verhältnis W_l prüft \mathbf{x}_l gegen den Mittelwert $\langle \mathbf{x}_m \rangle_m$ ab (bekannt als T^2 -Test, zu dem die F-Verteilung gehört) und wie in (2.6) folgt dann [Caroni, Chrys und Prescott, Philip, 1992] der für spätere Berechnung der kritischen Werte t_c wichtige Zusammenhang mit der F-Verteilung (Details dazu in [Caroni, Chrys und Prescott, Philip, 1992])

$$W_l = \left(1 + \frac{T_l^2}{n-2}\right)^{-1} \sim \left(1 + \frac{p}{n-p-1} F_{p, n-p-1, \alpha/n}\right)^{-1} = t_c$$

Die Werte für $F_{p, n-p-1, \alpha/n}$ sind abhängig von der jeweils aktuellen Größe n des Datensatzes. Sie sind tabelliert in [Wilks, 1963] [Caroni, Chrys und Prescott, Philip, 1992] [Hawkins, 1980] [Barnett, Vic und Lewis, Toby, 1994]. Die sequenzielle Anwendung der Wilks-Teststatistik geschieht nun in folgenden Schritten:

- Es wird eine Anzahl k maximal im Datensatz vermuteter Ausreißer a priori festgelegt.
- Das Likelihood-Verhältnis wird über alle Datenpunkte minimiert (2.8) und somit der erste potentielle Ausreißer $\mathbf{x}^{[1]}$ und sein Likelihood-Verhältnis D_1 identifiziert.
- Der Datensatz wird um $\mathbf{z}^{[1]}$ reduziert und die Likelihood-Verhältnisse neu berechnet. Dabei ist folgende rekursive Formulierung hilfreich:

$$D_2 = \min_l W_l^{[1]} = \frac{\det \mathbf{A}^{[1,2]}}{\det \mathbf{A}^{[1]}} = 1 - \frac{n-1}{n-2} \left(\mathbf{x}_2 - \langle \mathbf{x}_m \rangle_m^{[1]}\right)^T \left(\mathbf{A}^{[1]}\right)^{-1} \left(\mathbf{x}_2 - \langle \mathbf{x}_m \rangle_m^{[1]}\right)^T$$

$$\text{mit} \quad \left(\mathbf{A}^{[1]}\right)^{-1} = \mathbf{A}^{-1} + \frac{n}{(n-1)D_1} \mathbf{A}^{-1} \left(\mathbf{x}_1 - \langle \mathbf{x}_m \rangle_m\right) \left(\mathbf{x}_1 - \langle \mathbf{x}_m \rangle_m\right)^T \mathbf{A}^{-1}$$

- Dies wird solange wiederholt, bis k potentielle Ausreißer $\{\mathbf{x}^{[1]}, \dots, \mathbf{x}^{[k]}\}$ und die dazu gehörenden Teststatistiken D_1, \dots, D_k gefunden sind.

- Beim Rückwärts-Testen der Ausreißerkandidaten werden über $j = k, \dots, 1$ die Hypothesen geprüft. Dabei ist zu beachten, daß wegen der auf den verschiedenen Stufen j unterschiedlich großen Datensätze die kritischen Werte für jede Stufe neu zu berechnen sind:

$$D_j \begin{cases} \leq \\ > \end{cases} \left(1 + \frac{p}{n-j-p-1} F_{p, n-j-p-1, \alpha/(n-j)} \right)^{-1} \begin{cases} \text{Es gibt } j \text{ Ausreißer.} \\ \text{Es gibt } \leq (j-1) \text{ Ausreißer.} \end{cases}$$

- Die Anzahl der Ausreißer beträgt also 0 oder $j_t = \operatorname{argmax}_j$ (Die Aussage $D_j \leq t_c(j)$ ist wahr).

Die stufenweise Berechnung der kritischen Werte erlaubt in der sequenziellen Wilks-Methode eine bessere Anpassung der α -Fehler. Das Verfahren unterdrückt durch seine Vorwärts-Rückwärts-Eigenschaft Maskierungseffekte und Überschwemmungen. Durch die Analogie zur Mahalanobis-Distanz in der Berechnung der Teststatistik ähnelt dieses Verfahren in den Ergebnissen, Vorteilen und Nachteilen sehr der Simonoff-Methode.

2.2.5 Methode nach Hadi mit robusten Schätzern

In Kapitel 5 werden wir unsere Ausreißerdetektion auf Daten anwenden, die bisher erst nach umfangreicher Bildverarbeitung und Filteroperationen im Rahmen von robusten Schätzern für die Parameter der ermittelten Modelle implizit gegen Ausreißer in den zugrunde liegenden Bilddaten abgesichert werden. Wir stellen dabei die Frage, ob eine Ausreißerdetektion oder robuste Verfahren sinnvoller auf der niedrigen Ebene der Meßwerte oder auf der hohen Ebene der Modellparameter angebracht ist. Andersherum können wir an dieser Stelle die Ideen der robusten Verfahren in die Ausreißerdetektion bringen, wie zahlreiche Autoren ([Davies, Laurie und Gather, Ursula, 1993] und darin enthaltene Referenzen) nahelegen. Beispielhaft sei die von Hadi [Hadi, 1992] eingeführte und mehrfach modifizierte [Hadi, Ali S. und Simonoff, Jeffrey S., 1993] [Hadi, 1994] [Hadi, Ali S. und Simonoff, Jeffrey S., 1994] Methode vorgestellt. Diese hat auch den Vorteil, die maximal vermutete Ausreißer-Anzahl k a priori nicht angeben zu müssen. Eine einführende Übersicht zu weiteren Methoden gibt [Atkinson, 1994].

In den bereits beschriebenen Verfahren wurden die einzelnen Datenpunkte gegen einen Mittelwert verglichen, wobei das Risiko besteht, daß dieser Mittelwert gerade durch die gesuchten Ausreißer verschoben ist. Anstelle des Mittelwertes ist der Median eine robuste Alternative, da er durch vorhandene Ausreißer weniger stark verzerrt wird. Er besitzt mit 50% den besten Bruchpunkt, den ein robuster Schätzer aufweisen kann: Mindestens die Hälfte der Werte müssen durch Ausreißer ersetzt werden, um die Aussagekraft des Medianwertes zu stören. Solche robusten Schätzer haben jedoch einen ungleich höheren rechentechnischen Aufwand.

Wir formulieren die Mahalanobis-Distanz (2.3) mit ihrer Streumatrix, die auch bei Wilks (2.7) auftaucht, mit dem komponentenweisen Median \mathbf{x}_M der Daten $\{\mathbf{x}\}$ um zu

$$D_l^2 = (\mathbf{x}_l - \mathbf{x}_M)^T \mathbf{S}^{-1} (\mathbf{x}_l - \mathbf{x}_M) \quad \text{mit} \quad \mathbf{S}_M = \frac{1}{n-1} \sum_{l=1}^n (\mathbf{x}_l - \mathbf{x}_M) (\mathbf{x}_l - \mathbf{x}_M)^T$$

und sortieren die Daten nach diesen Werten. Die Startdistanzen ermittelt Hadi anschließend durch

$$D_l^2 = (\mathbf{x}_l - \mathbf{x}_v)^T \mathbf{S}_v^{-1} (\mathbf{x}_l - \mathbf{x}_v)$$

wobei

$$\mathbf{S}_v = \frac{\sum_{l=1}^n v_l (\mathbf{x}_l - \mathbf{x}_v) (\mathbf{x}_l - \mathbf{x}_v)^T}{\left(\sum_{l=1}^n v_l\right) - 1} \quad \text{und} \quad \mathbf{x}_v = \frac{\sum_{l=1}^n v_l \mathbf{x}_l}{\sum_{l=1}^n v_l}$$

und der Gewichtungsfunktion

$$v_l = \begin{cases} 1 & \text{wenn } l \leq \frac{n+p+1}{2} \\ 0 & \text{sonst} \end{cases}$$

Nach diesen Startinstanzen wird erneut sortiert. Durch die Verwendung des Medians und der Gewichtungsfunktion v_l , die Teile des sortierten Datensatzes ausblendet, ist die *Basis* genannte Unterstichprobe der ersten $p+1$ Daten mit hoher Wahrscheinlichkeit ausreißerfrei und kaum durch einen Ausreißerverzerrten Mittelwert von Masking oder Swamping bedroht. Für die nun folgende Berechnung der Mittelwerte \mathbf{x}_b und Streumatrix \mathbf{S}_b in der Mahalanobis-Distanz nach (2.3) wird die Nicht-Basis der übrigen $n-p-1$ Werte nicht benutzt. Für nicht-invertierbare Streumatrizen siehe [Hadi, 1992]. Die Daten werden nach diesen neuen Distanzen sortiert und der Umfang der Basis vergrößert: die ersten $p+2$ bilden eine neue Basis. Dieses Anwachsen der Basis wird nun iterativ wiederholt, bis bei einer Basisgröße r ein Abbruchkriterium erfüllt ist. Danach werden die Distanzen

$$D_l^2 = (\mathbf{x}_l - \mathbf{x}_b)^T (q' \mathbf{S}_b)^{-1} (\mathbf{x}_l - \mathbf{x}_b)$$

mit

$$q' = \frac{\left(1 + \frac{r}{n-p}\right)^2 \cdot \left(100 \frac{n+p+1}{2n} - q\right)}{\chi_{p;0,5}^2}$$

mit dem Quantil q der n Daten berechnet, aus denen dann die Ausreißer entnommen werden. In [Hadi, Ali S. und Simonoff, Jeffrey S., 1994] wird für den Fall $n > 3p+1$ ein durch Simulationsanalysen verbesserter Korrekturfaktor

$$q' = \left(1 + \frac{2}{n-1-3p} + \frac{p+1}{n-p}\right)^2$$

für multivariate Daten der Dimension p genannt. Die Suche wird abgebrochen, wenn für eine Basisgröße r im sortierten Datensatz die Distanz des $(r+1)$ -ten Punktes

$$D_{r+1}^2 = (\mathbf{x}_{r+1} - \mathbf{x}_b)^T (q' \mathbf{S}_b)^{-1} (\mathbf{x}_{r+1} - \mathbf{x}_b) \geq \chi_{p;\alpha/n}^2$$

erfüllt. Alle Punkte, die diese Ungleichung erfüllen, sind dann als Ausreißer identifiziert.

2.2.6 Zusammenfassung

Während letztlich bei allen Datenaufnahmen und Messungen gründliche Ausreißerdetektionen einer Auswertung vorangehen sollten, sind die Verfahren im naturwissenschaftlich-technischen Bereich auch heute noch wenig bekannt und selten eingesetzt. So wären die Fragestellungen zur Ausreißeridentifikation nach ihrer ersten Blütezeit in den 30er Jahren vor allem in den Zeitschriften BIOMETRIKA und TECHNOMETRICS fast in Vergessenheit geraten, wenn nicht Jahrzehnte später mit den robusten Schätzern [Huber, 1980] [Rousseeuw, Peter J. und Leroy, Annik M., 1987] neue Akzente gesetzt worden wären. Die weit verstreute und sporadische Literatur versuchen [Hawkins, 1980] und [Barnett, Vic und Lewis, Toby, 1994] zusammenzufassen. Der Eingang selbst klassischer hier vorgestellter Verfahren in Standard-Software zur statistischen Datenauswertung steht aber noch aus; auch die im folgenden Abschnitt gesichteten Verfahren, die in der Bildverarbeitung unter dem Namen "Ausreißerdetektoren" in Teilen sogar nochmal erfunden wurden, vermitteln ein ähnliches Bild.

Ein Grund für diese Vernachlässigung mag in den komplexen und auch heute noch Rechenzeitintensiven Iterationen der Verfahren liegen, die teilweise eher graphisch orientierte Zwischenschritte haben und immer wieder auf umfangreiche Tabellenwerke zur Bestimmung kritischer Werte oder Grenzen zurückgreifen. Ein anderer Grund für die Zurückhaltung der Experimentatoren mag darin liegen, daß es mit der bloßen Identifikation der Ausreißer nicht getan ist. Was tun mit Ausreißern? Weder das rücksichtslose Streichen ohne jede Interpretation noch die unbegründete Beibehaltung offensichtlich falscher Werte sind gangbare Wege. Eine individuelle Ursachenforschung oder erneute Messung sind in der experimentellen Praxis groß angelegter Experimente aber oft schlichtweg unmöglich.

2.3 Änderungsdetektoren der Bildverarbeitung

2.3.1 Einleitung und Überblick

Aus der Bildverarbeitung sind zahlreiche Verfahren zur Änderungsdetektion bekannt. Ende der 70er Jahre, als erstmals die technischen Möglichkeiten für eine Bildfolgenverarbeitung durch Daten- und Rechenpeicher sowie Rechenleistung bereitstanden, wurde mit dem Entwurf solcher Verfahren begonnen, die binäre Änderungskarten bestimmten. Historisch ließe sich das Forschungsfeld vielleicht an der Hamburger Taxi-Szene [Hsu, Y. Z. u. a., 1984] oder vorausgehenden Arbeiten [Nagel, 1978] festmachen. Zu noch früheren Arbeiten finden sich in [Huang, 1981] und [Huang, 1982] gute Übersichten von Nagel und anderen. Von reinen Änderungsdetektoren entwickelte sich die Bildfolgenanalyse mit mächtigen Konzepten (Optischer Fluß [Horn, B. K. P. und Schunk, B., 1981]) in das Gebiet der Bewegungsschätzung (engl. *motion estimation*), also weg von der reinen Detektion und hin zur gleichzeitigen Segmentierung von Objekten und der Bestimmung ihrer Bewegung. Darin tauchen heute die Änderungsdetektoren nur noch versteckt auf und werden nicht mehr explizit so benannt. Eine gewisse Wende oder Hinwendung zu den frühen Ergebnissen der Änderungsdetektion kam in den letzten Jahren durch die Suche nach effizienten Kodierverfahren für Bildfolgen im Multimedia-Bereich auf. Von den frühen Verfahren sollen hier die wichtigsten vorgestellt werden. Es lassen sich dabei drei Forschergruppen ausmachen: Jain und Nagel (sowie Hsu und Rekers) vor allem mit [Jain,

Ramesh und Nagel, H.-H., 1979] und [Hsu, Y. Z. u. a., 1984] in den 80er Jahren und in den 90er Jahren dann Bouthemy, Lalande (auch Heitz) ausführlich in [Bouthemy, Patrick und Lalande, Patrick, 1993] (früher in [Bouthemy, Patrick und Lalande, Patrick, 1990] [Heitz, F. und Bouthemy, Patrick, 1990] [Lalande, Patrick und Bouthemy, Patrick, 1990]) sowie die Arbeiten von Aach, Kaup (auch Mester, Toth) insbesondere in [Aach, 1993], [Aach, Til u. a., 1993b]. Es wäre falsch, von unterschiedlichen Schulen oder gar gegensätzlichen Verfahren zu sprechen, da die Arbeiten der Autoren untereinander Ideen aufgreifen und Methoden vermischen.

Ein sehr intuitiver Ansatz [Jain, Ramesh u. a., 1979] zur Detektion von geänderten Bereichen in einer Bildfolge $g(\mathbf{x}, t)$ ist die Berechnung eines Differenzbildes $d(\mathbf{x}, t) = g(\mathbf{x}, t) - g(\mathbf{x}, t - \delta t)$, in dem die gesuchten Unterschiede unabhängig von den verschiedenen generellen Intensitäten zu Tage treten. Da die Bilder $g(\mathbf{x}, t)$ in der Regel räumlich wie zeitlich unabhängiges Rauschen enthalten, tritt dies in den Differenzbildern ebenso auf und konkurriert mit den tatsächlich gesuchten Änderungen, die meist von bewegten Objekten herrühren. Ein einfacher Schwellwert auf $|d(\mathbf{x}, t)|$, der eine binäre Änderungskarte mit den dort abgelegten Entscheidungen (“geändert” oder “nicht geändert”) erzeugt, beinhaltet immer die Problematik, daß bei zu niedrigen Schwellwerten die gesuchte Änderungskarte mit Rauschen überschwemmt wird, bei zu hohem Schwellwert eventuell gesuchte Regionen mit Änderungen nicht erfaßt werden. Schwellwertverfahren kranken im allgemeinen an ihrer meist empirischen Bestimmung der Schwellwerte, die stark von der jeweiligen Applikation abhängen. Eine datengetriebene Schwellwertbestimmung oder lokale Schwellwerte stellen Verbesserungsansätze dar. Sinnvoll ist eine Kopplung des eingesetzten Schwellwertes an die Varianz σ_k des Kamerarauschens. Um deren Wert sicher zu schätzen, muß jedoch über viele möglichst änderungsfreie (ausschließlich rauschbehaftete) Bilder gemittelt werden, und gerade diese änderungsfreien Daten liegen ja nicht vor. Zur sauberen Trennung von Pixelrauschen und den tatsächlich gesuchten Änderungen in den berechneten Änderungskarten sei auf den grundlegenden Unterschied beider Phänomene verwiesen: Rauschen ist pixel-gebunden und tritt gleichverteilt und als vereinzelte Pixel in Form von Löchern oder Flecken in den Änderungskarten auf. Die gesuchten Änderungen, die ursächlich an Bewegungsphänomene geknüpft sind, erzeugen räumlich wie zeitlich kompakte, glatt berandete, homogene Gebiete in den Änderungskarten. Morphologische Glättungen [Soille, 1999a] [Soille, 1999b], Medianfilter oder Makropixel (alle Pixel einer kleinen Umgebung in der Änderungskarte bekommen eine aus den Werten der Änderungskarte bestimmte Entscheidung), mit denen Änderungskarten nachträglich geglättet werden, können Rauschpixel unterdrücken. Der Vorteil solcher Zusammenfassungen mehrerer Pixel ist die gegen Rauschen robustere Entscheidung; der Nachteil ein Verlust an Auflösung und der Übergang zu einer Stichprobe von höherer Dimension als ein einzelner skalarer Entscheidungswert. Im Folgenden sollen nach einem kleinen Einschub zur Zielsetzung die klassischen in der Bildverarbeitung bekannten Verfahren zur Änderungsdetektion vorgestellt werden, die über diesen einfachen Ansatz hinausgehen.

2.3.2 Objektdetektion versus Änderungsdetektion

Unausgesprochen werden Änderungsdetektoren häufig als Detektoren für bewegte Objekte in Bildfolgen benutzt. In der Regel wird im Anschluß an eine Änderungsdetektion die erstellte Änderungskarte für die Berechnung von Tiefeninformation aus Stereobildern, Bewegungsschätzung (Richtung,

Geschwindigkeit) von bewegten Objekten oder dergleichen benutzt. Hierbei muß aber beachtet werden, daß die gefundenen Änderungsregionen weit mehr als nur die bewegten Objekte umfassen und ihre Ränder nicht die Ränder der Objekte sind: mitbewegte Schatten, durch Objektbewegung veränderte Lichtsituation im Aufnahmebereich, schnelle Beleuchtungswechsel, freiwerdender Hintergrund erzeugen in den Änderungskarten Bereiche, die der direkten Benutzung der Änderungsregionen als Objektberandungen widersprechen. Auch das aus der Theorie des Optischen Flusses bekannte Blendenproblem [Jähne, 2002] tritt bei der Änderungsdetektion auf: bewegt sich ein Objekt homogener und nicht-texturierter Intensität langsam durch die Bildfolge, so werden vom Änderungsdetektor je nach zeitlicher Tiefe der in die Berechnung einbezogenen Bildanzahlen nur die senkrecht zur Bewegungsrichtung stehenden Objektränder gefunden, das Objektinnere wegen der bis auf Rauschen unveränderten Intensitäten aber nicht. Änderungsdetektion ist also zugleich mehr (Schattendetektion, Beleuchtungsvariantions-Detektion) und auch weniger (Blendenproblem) als Objektdetektion. Zahlreiche Verfahren existieren, um aus Änderungskarten die tatsächlich interessierenden Regionen bewegter Objekte herauszuarbeiten. Dafür werden zusätzlich zu den Intensitäten des Differenzbildes und daraus abgeleiteten Prüfstatistiken (siehe Abschnitt 2.3.5) Bewegungsfelder mit der Gradientenmethode (siehe Abschnitt 2.3.8) bestimmt und Texturinformation hinzugenommen. Unterscheidungen sind beispielsweise, daß freiwerdender Hintergrund im Gegensatz zum Objekt keine Bewegung zeigt, Objekte haben mitbewegte Texturinformation und degleichen mehr. Hötter [Hötter, Michael und Thoma, Robert, 1988] [Hötter, Michael u. a., 1995] benutzt für Überwachungsaufgaben mit statischen Kameras auf Änderungskarten aus Differenzbildern (Verfahren der Abschnitte 2.3.5 und 2.3.6) die lokalen Merkmale Bilddifferenz, Textur und Bewegung und unterscheidet anhand einer statistischen Analyse dieser Merkmale über die gesamte Aufnahmezeit zwischen relevanten und nicht-relevanten Objekten, die Änderungen verursacht haben. Die Idee, für statistische Analysen (Mittelwerte, Varianzen) den gesamten zeitlichen Verlauf verschiedener lokal aufgelöster Merkmale zu nutzen, werden wir als ganz wichtigen Baustein der in Kapitel 3 entworfenen Algorithmik auch für die Änderungsdetektion selbst umsetzen. Hötter löst sich auch von der blockbasierten (Makro-pixel) Merkmalsanalyse und verbessert die Auflösung der Objekte erheblich durch eine pixelbasierte und objektorientierte Statistik der Merkmale, die in hierarchischen Verfahren solange zu den unterschiedlich klassifizierten Änderungen zugeordnet werden, bis sich Regionen homogener Merkmale herausgebildet haben. Dies erlaubt zum Beispiel die Unterscheidung zwischen freiwerdendem Hintergrund und tatsächlich bewegten Objekten im Vordergrund. Um die Objektsegmentierung besser zu stützen, benutzt [Mech, Roland und Wollborn, Michael, 1997] zusätzlich Kanteninformation aus Ableitungen auf den zweidimensionalen Bildern der Sequenz. Die lokal angepaßten Schwellwerte führen die Idee der Ortsauflösung aus den Vorarbeiten Hötters fort. Eine ebenfalls vor allem an neuen Multimediastandards orientierte Unterscheidung mitbewegter Schatten stellt [Stauder, Jürgen u. a., 1999] vor (hier findet sich auch eine gute Übersicht zu Verfahren, die andere nicht Objekt-verursachte Änderungen ausschließen). Im Abschnitt 2.3.10 werden aktuelle Arbeiten zur Objekterkennung in Bildfolgen genannt.

2.3.3 Änderungsdetektion mit lokalen Bildmodellen

Zuerst sei hier auf den von Nagel und Jain beschrittenen Weg des paarweisen Bildvergleichs (engl. *interframe differences*) aller Pixel verwiesen, bei dem für die Entscheidung, ob in einem Pixel Änderung vorliegt oder nicht, durch eine Modellierung einer kleinen räumlichen Bildumgebung um das

fragliche Pixel für beide zu untersuchenden Bilder zwei Parametersätze gefunden werden. Stimmen die Parametersätze überein, liegt keine Änderung vor; andernfalls geht man davon aus, daß eine Änderung zwischen den beiden Bildern zu einem Unterschied der Modellparameter geführt hat. Das für das Testen der kleinen Bildbereiche gegeneinander benutzte Verfahren ist ein Likelihood-Ratio-Test von Yakimovski [Yakimovsky, 1976], der ursprünglich zur Kantendetektion in Einzelbildern konzipiert war und von Nagel [Nagel, 1978] für die Analyse zeitlicher Bildfolgen übertragen wurde. Anfangs mit einer Bildfenstermodellierung über eine Konstante [Jain, Ramesh und Nagel, H.-H., 1979], erweitert in [Nagel, Hans-Hellmut und Rekers, G., 1982] und [Rekers, 1982] auf eine lineare Bildfunktion, sei an dieser Stelle die Modellierung mit bivariaten Polynomen zweiter Ordnung aus [Hsu, Y. Z. u. a., 1984] vorgestellt. Kombinationen aus Änderungsdetektion und Objektsegmentierungen finden sich in [Jain, Ramesh u. a., 1979]. Die im Folgenden vorgestellte Methode wurde von Bouthemy [Bouthemy, Patrick und Lalande, Patrick, 1990] [Bouthemy, Patrick und Lalande, Patrick, 1993] um eine Bayes-Glättung der Änderungskarten erweitert (siehe Abschnitt 2.3.4).

Seien w_1 und w_2 zwei in der Regel quadratisch gewählte kleine Ausschnitte mit n_1 bzw. n_2 Pixeln, die für die ursprünglich von Yakimovski durchgeführte Kantendetektion in einem Bild an unterschiedlichen Orten (so daß beide Fenster aneinanderstoßen) lagen, in der von Nagel auf Änderungsdetektion in Bildfolgen benutzten Methode aber aus zwei aufeinanderfolgenden Bildern $g_1(\mathbf{x}) = g(\mathbf{x}, t - \delta t)$ und $g_2(\mathbf{x}) = g(\mathbf{x}, t)$ an derselben Position \mathbf{x} stammen. Es wird angenommen, daß sich die Grauwertintensitäten aus einem wahren Wert und additiv hinzutretendem gaußverteilterm Rauschen zusammensetzen und damit $N_i(\bar{g}_i, \sigma_i)$ -verteilt sind. Mit dem Index i seien im ganzen Abschnitt die beiden Stichproben $i = 1, 2$ abgekürzt. Wenn $p(x_i, \beta)$ die Wahrscheinlichkeitsdichte der Zufallsvariablen x_1, \dots, x_n ist, die von den unbekanntem Parametern $\beta = (\beta_1, \dots, \beta_J)^T$ abhängt, dann ist die Likelihood gegeben durch $\prod_i p(x_i, \beta)$. Die beste Schätzung $\hat{\beta}$ aller Parameter β ist durch die Maximierung der Likelihood L gegeben

$$L(\hat{\beta}) = \max. \quad \Rightarrow \quad \frac{1}{L} \frac{\partial L}{\partial \beta_j} \Big|_{\hat{\beta}_j} = 0 \quad \Rightarrow \quad \sum_i \frac{1}{p(x_i, \beta)} \frac{\partial p(x_i, \beta)}{\partial \beta_j} \Big|_{\hat{\beta}_j} = 0 \quad \text{mit } j = 1, \dots, J$$

Dieses Gleichungssystem wird nach den $\hat{\beta}_j$ aufgelöst. In den beiden Stichproben w_i ist die Likelihood der Grauwertintensitäten gegeben durch

$$\frac{1}{\sqrt{2\pi\sigma_i^2}} \exp \left\{ -\frac{(g_i(\mathbf{x}) - \bar{g}_i)^2}{2\sigma_i^2} \right\}$$

Die wahren Größen für den Mittelwert \bar{g} an einem bestimmten Punkt (\mathbf{x}, t) und die in allen (\mathbf{x}, t) als gleich angenommenen Varianz σ_i^2 sind unbekannt und werden mit der Maximum-Likelihood-Methode über Stichproben aus den beiden Fenstern w_i mit $i = 1, 2$ geschätzt zu

$$\langle g_i \rangle_{w_i} = \frac{1}{n_i} \sum_{\mathbf{x} \in w_i} g_i(\mathbf{x}) \quad s_i^2 = \frac{1}{n_i} \sum_{\mathbf{x} \in w_i} (g_i(\mathbf{x}) - \langle g_i \rangle_{w_i})^2$$

Statistisch nähern sich diese Schätzungen an die wahren Werte für größere n_i an, praktisch macht es bei nicht-trivialen Bildern aber keinen Sinn, große Fenster w_i zu benutzen, weil darin die Bildinformation g_i nicht mehr als Konstante $\langle g_i \rangle_{w_i}$ angenommen werden kann. In den beiden Stichproben w_i

sind die Grauwertintensitäten annähernd verteilt nach

$$\frac{1}{\sqrt{2\pi s_i^2}} \exp \left\{ -\frac{(g_i(\mathbf{x}) - \langle g_i \rangle_{w_i})^2}{2s_i^2} \right\}$$

Die Verbundwahrscheinlichkeit für jede Stichprobe w_i beträgt dann

$$p_i(g_i(\mathbf{x}) | N_i(\bar{g}_i, \sigma_i)) = \left(\frac{1}{\sqrt{2\pi s_i^2}} \right)^{n_i} \exp \left\{ -\frac{1}{2s_i^2} \sum_{\mathbf{x} \in w_i} (g_i(\mathbf{x}) - \langle g_i \rangle_{w_i})^2 \right\} = (2\pi e)^{-\frac{n_i}{2}} \frac{1}{s_i^{n_i}}$$

Die Stichproben als Ganzes werden also ausschließlich durch die geschätzte Varianz charakterisiert. Vergleiche der Stichproben untereinander lassen sich daher über Vergleiche von diesen Varianzen anstellen. Sei $w_0 = w_1 \cup w_2$ die aus beiden Stichproben zusammengesetzte Stichprobe. Ihre Parameter lassen sich durch die Parameter der Teilstichproben ausdrücken.

$$\begin{aligned} n_0 &= n_1 + n_2 \\ \langle g_0 \rangle_{w_0} &= \frac{n_1 \langle g_1 \rangle_{w_1} + n_2 \langle g_2 \rangle_{w_2}}{n_0} \\ s_0^2 &= \frac{1}{n_0} \left(\sum_{\mathbf{x} \in w_1} (g_1(\mathbf{x}) - \langle g_0 \rangle_{w_0})^2 + \sum_{\mathbf{x} \in w_2} (g_2(\mathbf{x}) - \langle g_0 \rangle_{w_0})^2 \right) \\ &= \frac{1}{n_1 + n_2} \left[n_1 s_1^2 + n_2 s_2^2 + n_1 (\langle g_0 \rangle_{w_0} - \langle g_1 \rangle_{w_1})^2 + n_2 (\langle g_0 \rangle_{w_0} - \langle g_2 \rangle_{w_2})^2 \right] \\ p_0(g_0(\mathbf{x}) | N_0(\bar{g}_0, \sigma_0)) &= (2\pi e)^{-\frac{n_1+n_2}{2}} \frac{1}{s_0^{n_1+n_2}} \end{aligned}$$

Es wird nun geprüft, ob eine Änderung im Pixel (\mathbf{x}, t) , um das das Fenster w_2 angeordnet ist, vorliegt, indem die Hypothese H_0 , die beiden Stichproben stammen aus derselben Verteilung $N_1(\bar{g}_1, \sigma_1) = N_2(\bar{g}_2, \sigma_2)$, gegen die Hypothese H_1 , daß die beiden Teil-Verteilungen unterschiedlich sind, geprüft wird.

H_0 : $N_1(\bar{g}_1, \sigma_1) = N_2(\bar{g}_2, \sigma_2)$ Pixel in (\mathbf{x}, t) wird als geändert markiert.

H_1 : $\bar{g}_1 \neq \bar{g}_2$ und $\sigma_1 \neq \sigma_2$ Pixel in (\mathbf{x}, t) wird als ungeändert beibehalten.

Anhand der gegebenen Gesamtstichprobe w_0 wird nun gefragt, welche Wahrscheinlichkeit höher ist (Maximum-Likelihood-Prinzip): die Wahrscheinlichkeit, daß der Stichprobe eine einzige Normalverteilung zugrunde liegt, oder die Wahrscheinlichkeit, daß die Messungen der Stichprobe aus zwei verschiedenen Normalverteilungen stammen. Die Wahrscheinlichkeit für die Abstammung aus einer einzigen Normalverteilung ist durch $p_0(g_0(\mathbf{x}) | N_0(\bar{g}_0, \sigma_0))$ gegeben. Die Auftrittswahrscheinlichkeit für die Gesamtstichprobe, wenn diese aus zwei unterschiedlichen Normalverteilungen stammt, ist gegeben durch die Verbundwahrscheinlichkeit $p_1 \cdot p_2$. Aus den Wahrscheinlichkeiten für die beiden Hypothesen bildet man das *Likelihood-Ratio*

$$l = \frac{\text{Likelihood}(H_1)}{\text{Likelihood}(H_0)} = \frac{p_1(g_1(\mathbf{x}) | N_1(\bar{g}_1, \sigma_1)) \cdot p_2(g_2(\mathbf{x}) | N_2(\bar{g}_2, \sigma_2))}{p_0(g_0(\mathbf{x}) | N_0(\bar{g}_0, \sigma_0))} = \frac{s_0^{n_1+n_2}}{s_1^{n_1} \cdot s_2^{n_2}}$$

Dieser Quotient wird 1, wenn keine Entscheidung getroffen werden kann, er wird > 1 , je stärker die Hypothese H_1 , also das Vorhandensein zweier unterschiedlicher Verteilungen, naheliegt. Yakimovski verwendete das Likelihood-Ratio direkt als Maß für die Existenz von Kanten; analog kann es bei der Änderungsdetektion als Maß für Vorhandensein von Änderungen in (\mathbf{x}, t) benutzt werden. Mit den Ersetzungen für die zusammengesetzte Varianz s_0^2 , mit einem Schwellwert $\theta > 1$, mit gleich großen Teilstichproben $n = n_1 = n_2$ und mit Quadrierung zur Vermeidung von Wurzelberechnungen wird der Ausdruck umgeformt zu

$$l^{2/n} = \frac{(s_0^2)^2}{s_1^2 \cdot s_2^2} = \frac{\left(\frac{s_1^2 + s_2^2}{2} + \left(\frac{\langle g_1 \rangle_{w_2} - \langle g_2 \rangle_{w_2}}{2} \right)^2 \right)^2}{s_1^2 s_2^2} \begin{matrix} \leq \\ > \end{matrix} \theta \begin{cases} H_0 : \text{Es liegt keine Änderung vor} \\ H_1 : \text{Es liegt Änderung vor} \end{cases}$$

Mit dem Verfahren lassen sich Grauwertsprünge von $\approx 3/256$ in Bilddaten mit einem Signal-Rausch-Verhältnis von 45db gut trennen. Für die Glättung der Änderungskarten setzten Hsu, Nagel et al. an jeder Stelle \mathbf{x} verkettete AND-Operationen aus den binären Entscheidungen von drei aufeinanderfolgenden Änderungskarten ein (Verbesserungen dazu in Abschnitt 2.3.4). Sie verwendeten bei $n = 12$ für θ die Werte 2,5, 4 und 10. Geht man davon aus, daß die Mittelwerte um $s_1 + s_2$ auseinander liegen, so ergibt sich für $\theta = 4$ gerade $s_1 = s_2$, bei $\theta = 10$ ist $s_1 \approx 3,2s_2$.

Bei der Berechnung der beiden Varianzen s_1^2 und s_2^2 werden alle Elemente der Stichprobe bisher gegen eine Konstante (hier der beste Schätzer für den Mittelwert) verglichen. Alle Abweichungen, die sich gegen dieses sehr einfache Bildmodell ergeben, werden über die geschätzten Varianzen als Änderung interpretiert. Somit ist bei diesem einfachen Ansatz zu erwarten, daß vermeintliche Änderungen häufig auf den Modellfehler zurückzuführen sind. Verbesserungen des Stichprobenmodells gegenüber einer Konstanten sind lineare oder bivariate Polynome als Näherungen des Stichprobeninhaltes.

$$\begin{aligned} s_{i,\text{const.}}^2 &= \frac{1}{n_i} \sum_{\mathbf{x} \in w_i} \left(g_i(\mathbf{x}) - \underbrace{\beta_{i,1}}_{\text{konstantes Modell}} \right)^2 \\ s_{i,\text{lin.}}^2 &= \frac{1}{n_i} \sum_{\mathbf{x} \in w_i} \left(g_i(\mathbf{x}) - \underbrace{[\beta_{i,1} + \beta_{i,1}x + \beta_{i,3}y]}_{\text{lineares Modell}} \right)^2 \\ s_{i,\text{quad.}}^2 &= \frac{1}{n_i} \sum_{\mathbf{x} \in w_i} \left((g_i(\mathbf{x}) - \underbrace{[\beta_{i,1} + \beta_{i,1}x + \beta_{i,3}y + \beta_{i,4}x^2 + \beta_{i,5}y^2 + \beta_{i,6}xy]}_{\text{quadratisches Modell}}) \right)^2 \end{aligned}$$

Anstelle einer empirischen Ermittlung der Schwellwerte ist eine Bestimmung der kritischen Bereiche bei bekanntem Verteilungs-Verhalten des Likelihood-Ratio im Sinne eines Hypothesentests mit der Signifikanz α sinnvoller. Nach dem Pearson-Neyman-Lemma liegt außerhalb von $\theta^{n/2}$ die kritische Region. Mit einer vorgegebenen Signifikanz α , dem Risiko für Fehlentscheidungen bei Aufrechterhalten oder Verwerfen von H_0 (Risiko für Fehler erster Art: Zurückweisung der Null-Hypothese, trotz Zutreffen), wird der Schwellwert mit der Integration der H_0 -Likelihood über die kritische Region festgelegt.

Beide Verbesserungen – Hypothesentest mit Signifikanzmaß und genauere Bildmodelle – werden jetzt vorgestellt. Dazu bezeichnet (wieder mit $i = 1, 2$) der Vektor $\mathbf{g}^{(i)} = (g_1^{(i)}, \dots, g_{n_i}^{(i)})^T$ die

$n = n_1 = n_2$ Grauwertintensitäten $g_j^{(i)} \in \{g(\mathbf{x}, t_i) : \mathbf{x} \in w_i\}$ der Teilstichprobe w_i mit $t_1 = t - \delta t$ und $t_2 = t$. Für beide Stichproben schreiben wir die Abbildungsgleichung des quadratischen Bildmodelles $\bar{\boldsymbol{\beta}}^{(i)} = (\bar{\beta}_1^{(i)}, \dots, \bar{\beta}_6^{(i)})^T$ auf die Teilstichprobe $\mathbf{g}^{(i)}$ mit dem $N(0, \sigma\mathbb{E})$ -verteilten Rauschen $\mathbf{e} = (e_1, \dots, e_n)$ (für beide Teilstichproben gleich). Das bedeutet, daß die Grauwerte $\mathbf{g}^{(i)}$ normalverteilt mit $N(\bar{\boldsymbol{\beta}}^{(i)}, \sigma\mathbb{E})$ sind.

$$\mathbf{g}^{(i)} = \mathbf{X} \cdot \bar{\boldsymbol{\beta}}^{(i)} + \mathbf{e} \quad \text{mit} \quad \mathbf{X} = \begin{pmatrix} 1 & x_1 & y_1 & x_1^2 & y_1^2 & x_1 y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & y_n & x_n^2 & y_n^2 & x_n y_n \end{pmatrix}$$

Die $n \times 6$ Matrix \mathbf{X} enthält die Polynomterme der Koordinaten und ist für beide Teilstichproben gleich, da die Position und Größe des Fensters identisch sind. Die Koordinaten $\mathbf{x}_j = (x_j, y_j)^T$ gehören zu den beiden Grauwerten $g_j^{(i)}$. Die Null-Hypothese H_0 besagt, daß beide Teilstichproben $\mathbf{g}^{(1)}$ und $\mathbf{g}^{(2)}$ aus derselben Normalverteilung $N(\bar{\boldsymbol{\beta}}, \sigma\mathbb{E})$ stammen; die Hypothese H_1 sagt, daß die Teilstichproben aus unterschiedlichen Verteilungen $N(\bar{\boldsymbol{\beta}}^{(1)}, \sigma\mathbb{E})$ und $N(\bar{\boldsymbol{\beta}}^{(2)}, \sigma\mathbb{E})$ stammen. Die Likelihood-Funktion ist gegeben mit

$$L^{(i)} = \left(\frac{1}{2\pi\sigma} \right)^{\frac{n}{2}} \exp \left\{ - \frac{(\mathbf{g}^{(i)} - \mathbf{X}\bar{\boldsymbol{\beta}}^{(i)})^T (\mathbf{g}^{(i)} - \mathbf{X}\bar{\boldsymbol{\beta}}^{(i)})}{2\sigma^2} \right\}$$

Die Maximum-Likelihood-Methode führt auf das Gleichungssystem

$$\left. \frac{\partial \log L^{(i)}}{\partial \bar{\boldsymbol{\beta}}^{(i)}} \right|_{\substack{\bar{\boldsymbol{\beta}}^{(i)} = \boldsymbol{\beta}^{(i)} \\ \sigma^2 = s^2}} = \mathbf{0} \quad \left. \frac{\partial \log L^{(i)}}{\partial \sigma^2} \right|_{\substack{\bar{\boldsymbol{\beta}}^{(i)} = \boldsymbol{\beta}^{(i)} \\ \sigma^2 = s^2}} = 0$$

Daraus erhält man als beste Schätzer für $\bar{\boldsymbol{\beta}}^{(i)}$ und σ^2

$$\boldsymbol{\beta}^{(i)} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{g}^{(i)} \quad (s^{(i)})^2 = \frac{1}{n} \left[(\mathbf{g}^{(i)} - \mathbf{X}\boldsymbol{\beta}^{(i)})^T (\mathbf{g}^{(i)} - \mathbf{X}\boldsymbol{\beta}^{(i)}) \right]$$

Wenn H_1 zutrifft, dann ergibt sich mit den 12 Parametern $\boldsymbol{\beta}^{(1)}$ und $\boldsymbol{\beta}^{(2)}$ eine Varianz von

$$s^2(H_1) = \frac{1}{2n} \left[\sum_i (\mathbf{g}^{(i)} - \mathbf{X}\boldsymbol{\beta}^{(i)})^T (\mathbf{g}^{(i)} - \mathbf{X}\boldsymbol{\beta}^{(i)}) \right]$$

Trifft H_0 zu, so sind

$$\boldsymbol{\beta}^{(0)} = (\mathbf{X}^{(0)T} \mathbf{X}^{(0)})^{-1} \mathbf{X}^{(0)T} \mathbf{g}^{(0)} \quad s^2(H_0) = \frac{1}{2n} \left[(\mathbf{g}^{(0)} - \mathbf{X}^{(0)}\boldsymbol{\beta}^{(0)})^T (\mathbf{g}^{(0)} - \mathbf{X}^{(0)}\boldsymbol{\beta}^{(0)}) \right]$$

wobei für die Gesamtstichprobe die Matrix und der Vektor kombiniert werden zu

$$\mathbf{X}^{(0)} = (\mathbf{X}, \mathbf{X})^T \quad \mathbf{g}^{(0)} = (\mathbf{g}^{(1)}, \mathbf{g}^{(2)})^T$$

Beide Größen $\beta^{(0)}$ und $s^2(H_0)$ werden jetzt wieder durch die entsprechenden Größen der Teilstichproben ausgedrückt. Man erhält (Details der Ableitung in [Hsu, Y. Z. u. a., 1984])

$$\beta^{(0)} = \frac{\beta^{(1)} + \beta^{(2)}}{2} \quad s^2(H_0) = s^2(H_1) + \frac{1}{4n} (\beta^{(1)} - \beta^{(2)})^T \mathbf{X}^T \mathbf{X} (\beta^{(1)} - \beta^{(2)})$$

Mit

$$F_1 = \frac{1}{2} \sigma^{-2} (\beta^{(1)} - \beta^{(2)})^T \mathbf{X}^T \mathbf{X} (\beta^{(1)} - \beta^{(2)})$$

$$F_2 = \sigma^{-2} (\mathbf{g}^{(1)} - \mathbf{X}\beta^{(1)})^T (\mathbf{g}^{(1)} - \mathbf{X}\beta^{(1)}) + (\mathbf{g}^{(2)} - \mathbf{X}\beta^{(2)})^T (\mathbf{g}^{(2)} - \mathbf{X}\beta^{(2)})$$

ist das Likelihood-Ratio dann

$$l_n^{\frac{1}{2}} = \frac{s^2(H_0)}{s^2(H_1)} = 1 + \frac{(\beta^{(1)} - \beta^{(2)})^T \mathbf{X}^T \mathbf{X} (\beta^{(1)} - \beta^{(2)})}{2 (\mathbf{g}^{(1)} - \mathbf{X}\beta^{(1)})^T (\mathbf{g}^{(1)} - \mathbf{X}\beta^{(1)}) + (\mathbf{g}^{(2)} - \mathbf{X}\beta^{(2)})^T (\mathbf{g}^{(2)} - \mathbf{X}\beta^{(2)})}$$

$$= 1 + \frac{1}{2} \frac{F_1/6}{F_2/12} = 1 + \frac{F}{2}$$

Weil $\beta^{(1)} - \beta^{(2)}$ normalverteilt ist mit $N(0, 2\sigma^2(\mathbf{X}^T \mathbf{X})^{-1})$, hat F_1 6 Freiheitsgrade und F_2 hat $2n - 12$ Freiheitsgrade. Die Größe F folgt einer F -Verteilung und hat 6 Freiheitsgrade im Zähler und $2n - 12$ im Nenner. Mit den Signifikanzwerten α werden die kritischen F -Werte $F_{\alpha;6,2n-12}$ als Tabellenwerte nachgeschlagen. Der Hypothesentest lautet dann für das Bildmodell mit quadratischem Polynom

$$F = \frac{F_1/6}{F_2/12} \begin{cases} \leq & F_{\alpha;6,2n-12} = 2(\sqrt{\theta}-1) = 2(l_n^{\frac{1}{2}}-1) \\ > & \end{cases} \quad \begin{cases} H_0 : \text{Es liegt keine Änderung vor.} \\ H_1 : \text{Es liegt Änderung vor.} \end{cases}$$

Für das konstante Bildmodell erhält man ein Likelihood-Ratio

$$l = \left(1 + \frac{1}{2} \frac{(\langle g_1 \rangle_{w_1} - \langle g_2 \rangle_{w_2})^2}{s_1^2 + s_2^2} \right)^n = \left(1 + \frac{t^2}{2n-2} \right)^n \quad \text{mit} \quad t^2 = (n-1) \frac{(\langle g_1 \rangle_{w_1} - \langle g_2 \rangle_{w_2})^2}{s_1^2 + s_2^2}$$

Die Größe t folgt einer t -Verteilung nach Student mit $2n - 2$ Freiheitsgraden. Ihre kritischen Werte $t_{\alpha;2n-2}$ sind tabelliert. Man testet dann t anstelle von l gegen die $t_{\alpha;2n-2}$; für $t > t_{\alpha;2n-2}$ wird H_0 verworfen.

Das vorgestellte Verfahren des Likelihood-Vergleiches betrachtet paarweise immer nur zwei Bilder aus der Bildfolge, greift also nicht auf die darüber hinaus in vorangegangenen oder späteren Bildern vorhandene Information zurück. Darin ähnelt es den Differenzbildverfahren. Wichtige Voraussetzung ist weiterhin, daß die Bildausschnitte durch biquadratische Polynome approximiert werden: es wird also ein Bildmodell eingeführt, daß nur zu einem gewissen Grad die tatsächliche Vielfalt in seine Parameter abbilden kann. Die verbleibenden Fehler der Modellierung werden dabei ausschließlich dem Rauschen zugeordnet und gehen daher in den Hypothesentest fälschlicherweise als Änderungen ein. Differenzbildansätze vermeiden die Notwendigkeit solcher Modellannahmen für die Stichproben,

da durch die Differenzbildung der zu modellierende Bildinhalt verschwindet und nur noch Annahmen über das verbleibende Rauschen gefordert werden. Da die Stichprobe lokal gebildet wird, geht stets auch Ortsauflösung verloren. Die Berechnung der Modellparameter in allen Bildumgebungen ist je nach gewähltem Modell und Verfahren äußerst rechenaufwendig. Ein herausragender Vorteil des Verfahrens ist, daß keine Größen bekannt oder voreingestellt werden müssen, wie das bei den einfachen Schwellwertverfahren auf Differenzbildern der Fall ist. Anstelle mehr oder weniger empirisch verbesserter Werte greift dieses Verfahren auf Schwellen der Hypothesentests und zugrunde liegender Verteilungen zurück, die weitaus universeller und theoretisch fundierter sind als anwendungsangepaßte Stellgrößen.

2.3.4 Ansätze zur Glättung mit Markov-Zufalls-Feldern

Bouthemy [Bouthemy, Patrick und Lalande, Patrick, 1990] [Bouthemy, Patrick und Lalande, Patrick, 1993] benutzte wie in Abschnitt 2.3.6 die auf Geman [Geman, Stuart und Geman, Donald, 1984a] zurückgehenden (ursprünglich für Einzelbilder konzipierten) Methoden, um ein Zuordnungsproblem (engl. *labelling*) in Bildfolgen zu lösen: welche Bildpunkte gehören zum Hintergrund, welche zu einem im Vordergrund bewegten Objekt? Dafür werden zu einem Bild dessen beide zeitliche Nachbarn in Vergangenheit und Zukunft in Form der Bilddifferenz oder Ableitung benutzt und darauf ein maximum-a-posteriori-Schätzer (MAP) angewandt. Das Zuordnungsbild $l(t)$ wird dann bei gegebenen Ableitungen und binären Karten für die zeitliche Änderung durch die Maximierung von

$$l^{(*)}(t) = \operatorname{argmax}_{l(t), l(t-\delta t)} P(l(t), l(t-\delta t), \text{Ableitung}(t), \text{Änderungskarte}(t))$$

über alle Zuordnungsbilder berechnet. Für die Bestimmung der binären Änderungskarte benutzt Bouthemy das Verfahren aus [Hsu, Y. Z. u. a., 1984] (siehe Abschnitt 2.3.3). Als logarithmiertes Likelihood-Ratio wird dabei mit einem linearen Bildmodell (Rauschen unabhängig von den unterschiedlichen Hypothesen) vereinfacht

$$\frac{1}{n^2} \sum_{\Delta x, \Delta y} d(x+\Delta x, y+\Delta y, t)^2 + \frac{\left\{ \sum_{\Delta x, \Delta y} \Delta x d(x+\Delta x, y, t) \right\}^2}{\sum_{\Delta x, \Delta y} \Delta x^2} + \frac{\left\{ \sum_{\Delta x, \Delta y} \Delta y d(x, y+\Delta y, t) \right\}^2}{\sum_{\Delta x, \Delta y} \Delta y^2}$$

wobei die Δx und Δy für ein 3×3 -Fenster $-1, 0, +1$ betragen. Dieses Log-Likelihood-Ratio folgt einer χ^2 -Verteilung mit drei Freiheitsgraden. Details zu den Markov-Zufalls-Feldern, der zugehörigen Energiefunktion und dem dahinter stehenden Bayes-Ansatz sind in Abschnitt 2.3.6 beschrieben, wo eine solche Glättung (engl. *regularization*) auf Bildmodell-freien Hypothesentest (Abschnitt 2.3.5) erfolgt. In den Arbeiten der Gruppe um Bouthemy finden sich auch Multiskalen-Ansätze für Änderungsdetektionen.

2.3.5 Änderungsdetektion mit Hypothesentests

Aach untersuchte in [Aach, 1993] Bayes-Methoden für die Bildverarbeitung und setzte diese in den Bereichen Änderungsdetektion für Bildfolgen, Segmentierung von zweidimensionalen Einzelbildern, Segmentierung von Volumendaten (insbesondere Röntgentomographie und Computertomographie)

und der Disparitätsschätzung von Stereo-Aufnahmen ein. Eine ähnliche anwendungsorientierte Einführung in Bayes-Methoden mit einem Schwerpunkt auf Gibbs/Markov-Zufallsfelder geben [Winkler, 1995] oder für Wahrscheinlichkeitsansätzen allgemein [Hornegger, J. u. a., 1999]. Die reine Änderungsdetektion [Aach, Til u. a., 1993b] [Aach, Til u. a., 1993a] [Aach, Til u. a., 1997] (ausführlicher in [Aach, Til und Kaup, André, 1995]) wird als Hypothesentest auf einer Stichprobe aus Differenzbildern durchgeführt. Die Bayes-Methoden kommen zum Tragen, wenn anschließend eine Glättung der ermittelten Änderungsmaske durchgeführt wird. Diese Glättung entspricht weitgehend den Ansätzen der Segmentierung von Einzelbildern mit Bayes-Methoden. Beide Stufen sollen hier und in Abschnitt 2.3.6 kurz beschrieben werden.

Das Verfahren arbeitet, wie das in Abschnitt 2.3.3 vorgestellte, auf dem Vergleich zweier aufeinanderfolgender Bilder, in diesem Fall sogar direkt auf dem Differenzbild $d(\mathbf{x}, t)$, das aus den Bildern $g(\mathbf{x}, t)$ und $g(\mathbf{x}, t - \delta t)$ der gesamten Bildfolge durch Subtraktion gebildet wird (zeitliche Abtaste ist δt). Aach begründet diesen Ansatz mit der damit möglichen Unterdrückung von Beleuchtungsschwankungen, die in der Regel als schwache Veränderung (zum Beispiel entlang der Tageszeiten oder durch Driften bei elektrischer Beleuchtung) auf großen Zeitskalen auftritt. Ansätze zu schnellen Beleuchtungswechseln werden in Abschnitt 2.3.7 besprochen. Zur Trennung von langsamen und schnellen Veränderungen bietet sich ein Hochpaß an, der in der Form der Differenzbildberechnung besonders einfach ist. Die Größe d^2/s^2 ist t-verteilt mit dem Mittelwert 0 und $n - 1$ Freiheitsgraden. Wenn die geschätzte Varianz s^2 zur wahren Varianz σ^2 wird, dann ist d^2/σ^2 gaußverteilt: $N(0, 1)$. Die in Abschnitt 2.3.1 im Rahmen eines intuitiven Ansatzes einleitend genannten Ideen zu Fehler-typen werden jetzt fundiert ausgebaut zu einem Hypothesentest anhand von Stichproben und darauf gebildeter geeigneter Stichprobenfunktionen. Einer Stichprobe wird das Ergebnis “keine Änderung” (Null-Hypothese H_0) oder das Resultat “Es liegt Änderung vor” (Eins-Hypothese H_1) zugeordnet. Stichproben werden in einem kleinen räumlichen Fenster w_x rund um den Grauwert $d(\mathbf{x}, t)$ genommen und das Ergebnis dann in einer binären Änderungskarte $b(\mathbf{x}, t)$ abgelegt. Für die Entscheidung werden in der Stichprobe, die sich zu einem Vektor $\mathbf{d}_x(t)$ zusammenfassen läßt, die Werte des Differenzbildes im Fenster um \mathbf{x} mit dem Kamerarauschen verglichen und einem Signifianztest unterzogen. Da das Kamerarauschen in ungeänderten Bereichen leicht zu ermitteln oder abzuschätzen ist, kann ein Zutreffen oder Nicht-Zutreffen der Null-Hypothese H_0 leicht festgestellt werden. Ungleich schwerer ist es, ein Zutreffen der Eins-Hypothese H_1 festzustellen, da die genaue Verteilung und die Eigenschaften geänderter Bildpunkte nicht bekannt sind. Letztlich weiß man nur, daß sie *nicht* dem änderungsfrei bestimmten Kamerarauschen entsprechen. Der Hypothesentest ist also eher ein Befürworten oder Ausschließen der Null-Hypothese H_0 , wobei ein Ausschließen von H_0 automatisch ein Befürworten von H_1 ist. Ein direktes Befürworten von H_1 ist bei der Suche nach beliebigen, über das Kamerarauschen hinaus gehenden Änderungen, letztlich unmöglich. Für die Stichprobenfunktion wird ein Wertebereich ermittelt, der unter der Null-Hypothese H_0 sehr unwahrscheinlich ist: in diesem Bereich wird H_0 verworfen. Dies beinhaltet aber auch immer, daß die “wahre Änderungsmaske” an dieser Stelle durchaus *keine* Änderung aufweisen könnte. Das Risiko einer solchen Fehlentscheidung ist im Parameter α quantifiziert. Dieser Parameter steuert im wesentlichen die Auswahl des unwahrscheinlichen Wertebereiches der Stichprobenfunktion für H_0 . Da die Gegenhypothese nicht charakterisiert ist, läßt sich dort der analoge Parameter β nicht aufstellen; α erfaßt daher nur Fehler erster Art.

Das Kamerarauschen mit der Varianz σ_k^2 wird additiv, gaußverteilt, mittelwertfrei und örtlich wie zeitlich unabhängig angenommen. Dann sind die pixelweisen Differenzen ebenfalls gaußverteilt, mittelwertfrei und haben die Varianz $2\sigma_k^2$. Die Werte der ungeänderten Bereiche (H_0) in den Differenzbildern verteilen sich dann wie

$$p(d(\mathbf{x}, t) | H_0) = \frac{1}{\sqrt{2\pi}2\sigma_k} \exp\left\{-\frac{d^2(\mathbf{x}, t)}{4\sigma_k^2}\right\}$$

In der Regel ist nur der Betrag $|d|$ interessant. Für Werte d , die außerhalb eines Intervalls $[-t_\alpha, t_\alpha]$ liegen, kann mit der Irrtumswahrscheinlichkeit α die Null-Hypothese verworfen werden. Aus gegebenem α kann die Schwelle t_α über die Fehlerfunktion erf bestimmt werden

$$\alpha = \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{t_\alpha}{2\sigma_k}\right) \quad \text{mit:} \quad \operatorname{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-x^2} dx$$

Die Entscheidung für oder gegen H_0 mit einzelnen Pixeln anhand ihrer Werte $d(\mathbf{x}, t)$ ist wegen der Breite der Gaußfunktion und der damit schwierigen Festlegung eines Schwellwertes t wenig robust. Ein besseres Maß ist die Verbundverteilung aller N_w Werte der Stichprobe $\mathbf{d}_\mathbf{x}(t)$ im Fenster w_x um (\mathbf{x}, t)

$$p(\mathbf{d}_\mathbf{x}(t) | H_0) = \left(\frac{1}{\sqrt{4\pi}\sigma_k}\right)^{N_w} \exp\left\{-\frac{\sum_{\mathbf{x} \in w_x} d^2(\mathbf{x}, t)}{4\sigma_k^2}\right\}$$

Diese Verteilung hängt nur von der lokalen Quadratsumme der Differenzen ab, die wir als einfache Stichprobenfunktion $D_\mathbf{x}$

$$D_\mathbf{x}^2(\mathbf{d}_\mathbf{x}(t)) = \sum_{\mathbf{x} \in w_x} \frac{d^2(\mathbf{x}, t)}{2\sigma_k^2} = \frac{1}{2\sigma_k^2} \sum_{\mathbf{x} \in w_x} (g(\mathbf{x}, t) - g(\mathbf{x}, t - \delta t))^2$$

für einen Hypothesentest nutzen wollen. Eine χ_m^2 -Verteilung mit Freiheitsgrad $n-1 = m$ ist definiert als die Verteilung der Werte $\chi^2 := (n-1) s_x^2 / \sigma_x^2$, wenn $(n-1) s_x^2$ die Summe der quadrierten Mittelwertdifferenzen $\sum_i (x_i - \langle x_i \rangle)^2$ über eine $N(\bar{x}, \sigma_x)$ -normalverteilte (dabei ist $\langle x_i \rangle$ der Schätzer des wahren Mittelwertes \bar{x}) Stichprobe von n Werten ist. Ihre Verteilung ist gegeben durch

$$p(\chi^2 \equiv u | N(\bar{x}, \sigma_x), m) = \frac{u^{\frac{m}{2}-1} e^{-\frac{u}{2}}}{2^{\frac{m}{2}} \Gamma(\frac{m}{2})} \quad \text{mit} \quad \Gamma(a) = \int_0^\infty e^{-b} b^{a-1} db$$

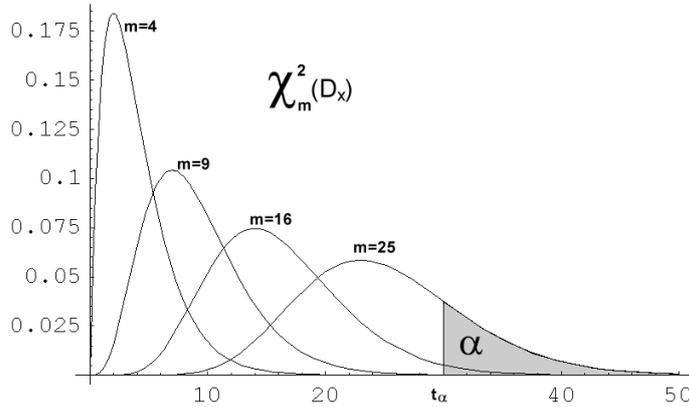
Mit $g_1 = g(\mathbf{x}, t)$ und $g_2 = g(\mathbf{x}, t - \delta t)$ und $2 \langle g \rangle_t = g_1 + g_2$ und der Umformung

$$\frac{d(\mathbf{x}, t)}{2} = \frac{1}{2} (g_1 - g_2)^2 = (g_1 - \langle g \rangle_t)^2 + (g_2 - \langle g \rangle_t)^2$$

findet man leicht, daß unter zutreffender Null-Hypothese die Verteilung $p(D_\mathbf{x} | H_0)$ der Stichprobenfunktion eine $\chi_{N_w}^2$ -Verteilung mit N_w Freiheitsgraden ist. Das Risiko α (auch Signifikanz) einer Fehlentscheidung ist die ab einer Schwelle t_α aufintegrierte und außerhalb liegende Wahrscheinlichkeit dieser Verteilung

$$P(D_\mathbf{x}^2 > t_\alpha | H_0) = \int_{t_\alpha}^\infty p(D_\mathbf{x}^2 | H_0) dD_\mathbf{x}^2 = \alpha$$

womit der Schwellwert t_α zu gegebenem Risiko von Fehlentscheidungen bestimmt ist (er muß in Tabellenwerken der kritischen Werte $\chi_{\alpha,m}^2$ nachgeschlagen werden). Wünschenswert ist ein möglichst steiler Abfall der Verteilung, um ein kompaktes Wertebereich zu erhalten. Dies erreicht man mit hohen Freiheitsgraden, Nachteil dabei ist aber, daß durch große Stichproben die Fenster w_x so groß werden, daß eine zu starke Mittelung über die Stichprobe stattfindet.



t_α	$\alpha=10^{-2}$	$\alpha=10^{-6}$
$m = 4$	13,28	36
$m = 9$	21,67	45
$m = 16$	32,00	60
$m = 25$	44,31	75
$m = 36$	58,62	94,5
$m = 50$	76,15	118,75

Abbildung 2.3: Die χ_m^2 -Verteilung mit verschiedenen Freiheitsgraden m und der Schwelle $t_\alpha = 30$ für die Fläche $\alpha=0.25$ unter der Verteilung mit 25 Freiheitsgraden

Die Entscheidung für Beibehaltung oder Verwerfen der Null-Hypothese wird dann zu

$$D_{\mathbf{x}}^2 \begin{cases} \leq & t_\alpha \\ > & t_\alpha \end{cases} \begin{cases} H_0 \text{ beibehalten: keine Änderung vorhanden} \\ H_1 : \text{es sind liegt eine Änderung vor} \end{cases}$$

Wenn die Verbundverteilungsdichte einer Stichprobe kleiner als

$$\left(\frac{1}{\sqrt{4\pi\sigma_k^2}} \right)^{N_w} \exp \left\{ -\frac{t_\alpha}{2} \right\}$$

ist, dann wird H_0 verworfen und auf Änderung entschieden. Dieser Hypothesen-Test wird für alle \mathbf{x} und t durchgeführt; die Entscheidungen bilden die Folge der Änderungskarte $b(\mathbf{x}, t)$. Beschleunigungen sind möglich, indem die Entscheidung allen im Fenster w_x enthaltenen Pixeln zugeordnet wird (Makro-Pixel).

Aach vergleicht [Aach, Til u. a., 1993b] mehrere Stichprobenfunktionen, darunter auch die lokale Betragssumme und den lokalen Betrag des Mittelwertes

$$D_{\mathbf{x}}(\mathbf{d}_{\mathbf{x}}(t)) = 2\sqrt{2} \sum_{\mathbf{x} \in w_x} \frac{|d(\mathbf{x}, t)|}{\sqrt{2}\sigma_k} \quad D_{\mathbf{x}}(\mathbf{d}_{\mathbf{x}}(t)) = \left| \sum_{\mathbf{x} \in w_x} \frac{d(\mathbf{x}, t)}{\sqrt{2}\sigma_k} \right|$$

Die lokale Betragssumme ergibt sich unter der Annahme von Laplace-verteilterm Kamerarauschen und führt zu einem χ^2 -Signifikanztest mit $2N_w$ Freiheitsgraden, der lokale Betrag des Mittelwertes (hier der Summe der Werte der Stichprobe) führt zu einem einseitig $N(0, N_w)$ -normal (wegen der nur geschätzten Varianz ist dies genauer nur eine t -Verteilung nach Student mit $N_w - 1$ Freiheitsgraden,

die für hohe N_w in eine Gaußverteilung übergeht) verteiltem Maß, dessen kritischer Wert t_α sich mit der erf-Funktion und entsprechender Signifikanz α ergibt. Motivation für diese Stichprobenfunktion ist die Tatsache, daß eine große Änderungs-behaftete Region gegebenenfalls zu schwach ausgeprägte Werte im Differenzbild hat, aber aus der Homogenität des Differenzbildes über die lokale Stichprobe unterschieden werden kann von zufällig positiven und negativen Differenzen gleicher Anzahl. Untersuchungen zeigen, daß die lokale Quadratsumme insbesondere mit der plausiblen Annahme von gaußverteilterm Rauschen die besten Ergebnisse zeigt. Für die Stichprobengröße werden typischerweise $N_w = 25$ gewählt; das Verfahren ist sehr stabil gegen den Signifikanzwert α , der $10^{-2} \dots 10^{-6}$ betragen kann.

Die sich ergebenden Änderungskarten sind in der Regel weitestgehend korrekt in ihrer Identifizierung großflächiger kompakter Regionen, enthalten aber Fehler in Form einzelner, im gesamten Bild verstreuter Pixel oder kleiner Regionen, die sich unterscheiden lassen in Fehler erster Art (α -Fehler, Fehldetektion: Es wird auf "geändert" entschieden, obwohl tatsächlich keine Änderung vorliegt; H_0 wird fälschlicherweise zurückgewiesen) und Fehler zweiter Art (β -Fehler, Versager: Es wird trotz einer tatsächlich vorhandenen Änderung auf "Keine Änderung" entschieden; H_0 wird fälschlicherweise beibehalten). Die Histogramme der Stichprobenfunktion für die Änderungs-behafteten und unveränderten Pixel (gegebenenfalls nach einer zusätzlichen morphologischen Glättung zu ermitteln) überlappen sich in der Regel sehr stark (Abbildung 2.4). Die vom Schwellwert t_α aus aufsummierte Fläche α für Fehler erster Art (analog für H_1 die Fläche β der Fehler zweiter Art) heißt auch Fehlerfläche. Durch Verschiebung von t_α kann dann der Fehler erster Art verringert werden, dies geht aber auf Kosten der Anzahl Zweiter-Art-Fehler. Wäre die Verteilung $p(D_x^2 | H_1)$ bekannt, so ergäbe sich der optimale Schwellwert t_{opt} im Schnittpunkt der Verteilungen: jede Variation des Schwellwertes würde dann jeweils eine der beiden Fehlerflächen stärker vergrößern als durch die Verkleinerung der anderen Fehlerfläche gewonnen wäre. Während es einerseits begrüßenswert ist, daß die Verteilung $p(d(\mathbf{x}, t) | H_1)$ nicht bekannt sein muß für dieses Verfahren (und damit die Methode offen hält für zahlreiche Detektionsaufgaben), so nachteilig ist es, keine Aussagen über den Fehler zweiter Art machen zu können. Die Frage, wie optimal das beschriebene Verfahren ist, konzentriert sich daher auf die Wahl der Stichprobenfunktion D_x bei gegebener Signifikanz α . Ein Testverfahren, das den Fehler zweiter Art bei gegebenem α minimiert, wird als *mächtigster Test* (engl. *most powerfull test*) bezeichnet. Man kann zeigen, daß die Stichprobenfunktion der lokalen Quadratsumme unter bestimmten Voraussetzungen ein solcher mächtigster Test ist und in der Tat bei gegebenem α gegenüber anderen Stichprobenfunktionen den Fehler zweiter Art minimiert. Das Pearson-Neyman-Kriterium besagt, daß ein mächtigster Test durch das *Likelihood ratio*

$$l(\mathbf{d}_x(t)) = \frac{p(\mathbf{d}_x(t) | H_0)}{p(\mathbf{d}_x(t) | H_1)}$$

gegeben ist, das eine Abbildung auf einen Wertebereich vornimmt, dessen kritische Regionen durch die gegebene Signifikanz α festgelegt werden. Zur Entscheidung über Beibehaltung oder Zurückweisung von H_0 wird der skalare Wert des Likelihood Ratio dann mit dem Schwellwert t_α verglichen. Kann l als eine Funktion einer anderen Funktion geschrieben werden, so ist auch diese ein mächtigster Test. Unter der Voraussetzung, daß die Differenzen $d(\mathbf{x}, t)$ der Änderungs-behafteten Bildpunkte mittelwertfrei, unabhängig und $N(0, \sqrt{2}\sigma_c)$ -gaußverteilt sind (wobei $\sigma_c \gg \sigma_k$), kann das Likelihood

Ratio für die Änderungsdetektion

$$l(\mathbf{d}_x(t)) = \frac{\frac{1}{\sigma_k^{N_w}} \exp \left\{ -\frac{1}{4\sigma_k^2} \sum_{\mathbf{x} \in w_x} d^2(\mathbf{x}, t) \right\}}{\frac{1}{\sigma_c^{N_w}} \exp \left\{ -\frac{1}{4\sigma_c^2} \sum_{\mathbf{x} \in w_x} d^2(\mathbf{x}, t) \right\}} = \left(\frac{\sigma_c}{\sigma_k} \right)^{N_w} \exp \left\{ -\frac{1}{2} \left(1 - \frac{\sigma_k^2}{\sigma_c^2} \right) D_x^2 \right\}$$

geschrieben werden. Es ist in der Tat eine Funktion der Stichprobenfunktion der lokalen Quadratsumme; damit ist diese ein mächtigster Test. Die genannten Voraussetzungen für $p(\mathbf{d}_x(t) | H_1)$ sind zwar nicht abwegig, aber auch nicht unbedingt sehr realistisch: Mittelwertfreiheit bedeutet, daß gleich viele Änderungen in die hellere wie dunklere Richtung der Intensitäten gehen. Die Gauß'sche Verteilung trifft sicher für natürliche Objekte zu, die derselben physikalischen Abbildung wie die sonstige Szene unterliegen. Die höhere Varianz ergibt sich einfach aus der Tatsache, daß Änderungen über das Rauschmaß hinaus vorliegen. Die Unabhängigkeit der Differenzwerte ist zweifelhaft, da die Ursachen für die Änderungen in ausgedehnten flächigen Objekten liegen und nicht mehr wie bei H_0 in voneinander vereinzelter Rauschprozessen der Bildpixel.

Auch bei optimaler Einstellung der Schwelle t_α verbleibt ein Fehler, der sich aus der Summe der beiden Fehlerflächen ergibt. Wünschenswert wäre ein Verfahren mit zwei Schwellwerten $t_\alpha \neq t_\beta$, die so deutlich auseinanderliegen (in dieser Anwendung $t_\beta < t_\alpha$), daß die verbleibenden Fehlerflächen klein sind und damit der Gesamtfehler gering. Für Stichprobenwerte $t_\beta < D_x^2 < t_\alpha$ setzt dies allerdings voraus, daß die Zugehörigkeit zu $p(D_x^2 | H_i)$ bereits bekannt ist oder aus anderen Ergebnissen eine Präferenz vorliegt. Es ist sicher nur annähernd richtig, im Zwischenbereich der beiden Schwellen eine bekanntlich abfallende Verteilung $p(D_x^2 | H_0)$ zu ignorieren und ausschließlich nach der Kontextinformation zu entscheiden. Eine Mischung aus der Verteilung und der Häufigkeit für Beibehaltung von H_0 , die aus der Kontextinformation nahegelegt wird, wäre ein weitaus stetigeres Verfahren. Eine solche stetige Hinzunahme weiterer Information ist die von Aach beschriebene kontextadaptive Steuerung der Entscheidungsschwelle.

Zusätzliche Information, die eine Entscheidung für Beibehaltung oder Verwerfen von H_0 stützt, kann der räumlichen und zeitlichen Nachbarschaft des untersuchten Pixels entnommen werden. Wie bereits einleitend bemerkt, nutzen Medianfilter auf den binären Änderungskarten (dies ist eine einfache Mehrheitsentscheidung) wie auch morphologische Operationen nicht die volle zur Verfügung stehende Information der Differenzbilder aus. Die Idee eines Zwei-Schwellen-Verfahrens ist im Kern: wenn Kontextinformation die Hypothese H_0 stützt, so wird anhand eines großen t_α entschieden. Legt die Zusatzinformation eine Ablehnung von H_0 nahe, so wird anhand von einer kleineren Schwelle t_β entschieden. Stichproben mit Werten $< t_\beta$ oder $> t_\alpha$ sind eindeutig den in diesen Gebieten dominanten Verteilungen zugeordnet. Im Zwischenbereich entscheidet dann ausschließlich die Kontextinformation über die Aufrechterhaltung der Null-Hypothese. Als zusätzliche Information bietet sich die bereits indizierte Änderungskarte $b(\mathbf{x}, t - \delta t)$ des vorangegangenen Differenzbildes an. Die darin zur Stichprobe (an den selben Stellen im Bild wie die $\mathbf{d}_x(t)$ im Fenster w_x rund um \mathbf{x}) korrespondierenden Entscheidungen $\mathbf{b}_x(t - \delta t)$ werden in die Verteilung als weitere Beobachtung mit aufgenommen (die

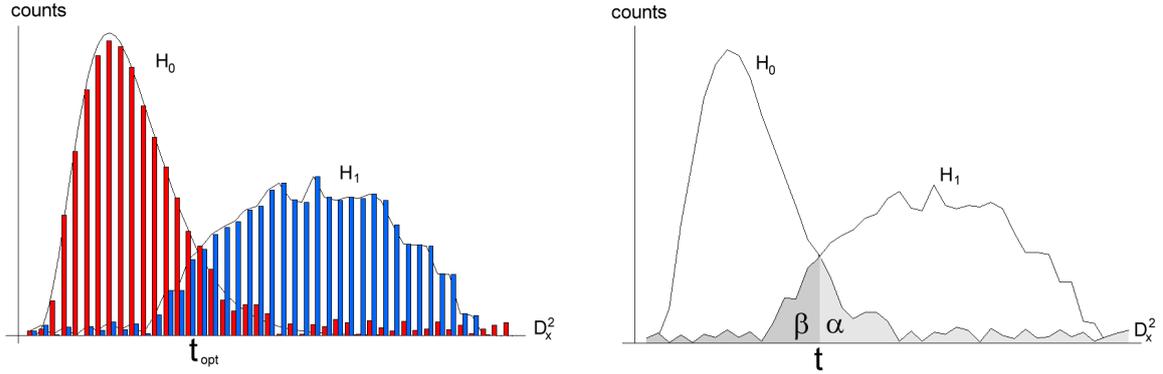


Abbildung 2.4: **Links:** Histogramme der Stichprobenwerte D_x^2 für die beiden durch H_0 und H_1 charakterisierten Bildpunkte nach Ermittlung der Änderungskarte (gegebenenfalls nach zusätzlicher morphologischer Glättung). Die Verteilung der H_0 zeigt ungefähr das erwartete $\chi_{N_w}^2$ -Verhalten, die Verteilung H_1 einen unregelmäßigen, stark von der jeweiligen Anwendung abhängigen Verlauf. Beide Histogramme überlappen stark, was zu Fehlern erster und zweiter Art führt. **Rechts:** Die optimale Schwelle ist durch den Schnittpunkt gegeben, da eine weitere Verschiebung der Schwelle dann stets zu einem Zuwachs der mit α und β charakterisierten Gesamtfehlerfläche ergeben würde.

Änderungskarte soll die Verteilungen der Stichprobenwerte nicht beeinflussen)

$$\begin{aligned}
 \frac{p(\mathbf{d}_x(t), b(\mathbf{x}, t - \delta t) \mid H_0)}{p(\mathbf{d}_x(t), b(\mathbf{x}, t - \delta t) \mid H_1)} &= \frac{p(\mathbf{d}_x(t) \mid H_0, b(\mathbf{x}, t - \delta t)) \cdot P(b(\mathbf{x}, t - \delta t) \mid H_0)}{p(\mathbf{d}_x(t) \mid H_1, b(\mathbf{x}, t - \delta t)) \cdot P(b(\mathbf{x}, t - \delta t) \mid H_1)} \\
 &= \frac{p(\mathbf{d}_x(t) \mid H_0)}{p(\mathbf{d}_x(t) \mid H_1)} \cdot \frac{P(b(\mathbf{x}, t - \delta t) \mid H_0)}{P(b(\mathbf{x}, t - \delta t) \mid H_1)} \\
 &\begin{cases} > \\ \leq \end{cases} T \quad \begin{cases} H_0 & \text{keine Änderung} \\ H_1 & \text{Änderung vorhanden} \end{cases}
 \end{aligned}$$

Der mächtigste Test als Likelihood Ratio lautet dann

$$\begin{aligned}
 l(\mathbf{d}_x(t)) &= \frac{p(\mathbf{d}_x(t) \mid H_0)}{p(\mathbf{d}_x(t) \mid H_1)} \\
 &= \left(\frac{\sigma_c}{\sigma_k} \right)^{N_w} \exp \left\{ - \left(1 - \frac{\sigma_k^2}{\sigma_c^2} \right) \frac{D_x^2}{2} \right\} \\
 &\begin{cases} > \\ \leq \end{cases} T \cdot \frac{P(b(\mathbf{x}, t - \delta t) \mid H_1)}{P(b(\mathbf{x}, t - \delta t) \mid H_0)} \quad \begin{cases} H_0 & \text{keine Änderung} \\ H_1 & \text{Änderung vorhanden} \end{cases}
 \end{aligned}$$

Zu dem Schwellwert T tritt nun eine Gewichtung durch die Verteilungen der bereits gefundenen Entscheidungen für das Vorgängerbild hinzu. Löst man nach der Stichprobenfunktion auf und vernach-

lässt dabei $\sigma_k/\sigma_c \ll 1$, dann ergibt sich das Entscheidungskriterium

$$D_{\mathbf{x}}^2 \begin{matrix} \leq \\ > \end{matrix} \underbrace{2 \ln \left[T \left(\frac{\sigma_c}{\sigma_k} \right)^{N_w} \right]}_{t_\alpha} + \underbrace{2 \ln \left[\frac{P(b(\mathbf{x}, t - \delta t) | H_0)}{P(b(\mathbf{x}, t - \delta t) | H_1)} \right]}_{t_b} \quad \begin{cases} H_0 & \text{keine Änderung} \\ H_1 & \text{Änderung vorhanden} \end{cases}$$

Wenn in der vorangegangenen Änderungskarte die Häufigkeit geänderter Bildpunkte in der Stichprobe höher ist, wird $P(b(\mathbf{x}, t - \delta t) | H_0) < P(b(\mathbf{x}, t - \delta t) | H_1)$ und damit der Summand zum Schwellwert t_α negativ. Durch die Absenkung werden Änderungen jetzt befürwortet. Für die genaue Ausgestaltung der Berechnung des Zusatztermes t_b ist zu beachten, daß in der N_w großen Stichprobe 2^{N_w} verschiedene Besetzungsmöglichkeiten der Änderungsmaske $b(\mathbf{x}, t - \delta t)$ existieren. Denkbare Vereinfachungen sind die Reduzierung der Stichprobe $\mathbf{b}_x(t - \delta t)$ auf das zentrale Bildpixel $b(\mathbf{x}, t - \delta t)$ oder eine Mehrheitsentscheidung in der Stichprobe, die in $t_q = \pm \text{const.}$ nur noch das Vorzeichen festlegt.

2.3.6 Glättung der Änderungskarten mit Bayes-Methoden

Geman [[Geman, Stuart und Geman, Donald, 1984b](#)] stellen 1984 die Analogie zwischen Bildern und statistischen mechanischen Systemen her, in denen Grauwertintensitäten und Kantenorientierungen wie Zustände von Teilchen im Festkörperlattice aufgefaßt werden. Weil die dazugehörige Energiefunktion der Gibbs-Verteilung folgt, spricht man auch von Markov-Zufallsfeldern (engl. *Markov random fields, MRF*). Die entwickelten Methoden eignen sich zur Rekonstruktion von verrauschten, unscharfen oder nichtlinear transformierten Bildern, zur Kantendetektion bis hin zur Objekterkennung und Bewegungsschätzung. Wegen des zugrunde liegenden Bayes-Ansatzes spricht man auch von Bayes-Bildverarbeitung, die sich grundsätzlich von den Methoden linearer Filter unterscheidet. Bayes-Schätzer kennen zwei Arten von Informationen: Vorwissen und empirische, gemessene Daten. Anders formuliert gibt es zwei Ursachen für Unsicherheit und Zufälligkeit: Vorwissen ist in der Regel unvollständig und gemessene Daten durch zufällig verteilte Ereignisse beeinträchtigt. Durch die Minimierung der Energiefunktion wird nach dem idealen Ausgangsbild gesucht, wenn das beobachtete, verrauschte Bild gegeben ist. Durch Konstruktion der Energiefunktion, in die das Vorwissen über den Prozeß der Beeinträchtigung eingeht, können auch ein Kantenbild oder eine Bewegungsschätzung das Ergebnis sein. Einen guten Überblick der ersten Dekade der Bayes-Bildverarbeitung gibt [[Winkler, 1995](#)]. Aach [[Aach, 1993](#)] führt typische Anwendungsfelder der Bayes-Methoden in der Bildverarbeitung vor.

Zur Glättung und damit verbesserten Segmentierung der in Abschnitt 2.3.5 gefundenen Änderungskarten benutzen [[Aach, Til u. a., 1993b](#)] [[Aach, Til u. a., 1991](#)] [[Aach, Til und Kaup, André, 1995](#)] [[Aach, Til u. a., 1997](#)] Bayes-Methoden. Die besprochenen Stichprobenfunktionen $D_{\mathbf{x}}$ lösen die genauen Nachbarschaftsbeziehungen innerhalb der Stichprobe $\mathbf{d}_x(t)$ weitgehend auf. Die Nachteile aus diesem Informationsverlust versucht die Glättung mit Bayes-Methoden auszugleichen. Kontextinformation aus der raumzeitlichen Nachbarschaft fließt in Form von Wahrscheinlichkeiten ein, die sich in das vorhandene Zuordnungsschema als zusätzliche multiplikative Gewichtung nahtlos einfügen.

Die Bayes-Theorie liefert einen statistischen Ansatz zur Regularisierung von Daten und Beobachtungen, in unserem Fall den Änderungskarten. Diese werden als Realisationen von Zufallsgrößen betrachtet und durch Wahrscheinlichkeitsverteilungen, empirisch also beobachteten Häufigkei-

ten, beschrieben. Vorhandenes Vorwissen oder zusätzlicher Kontext werden durch geeignete *a-priori*-Wahrscheinlichkeiten modelliert und fließen so in das Verfahren ein. Unter den zahlreichen Schätzverfahren der Bayes-Theorie ist der *Maximum-a-posteriori*-Schätzer (MAP) der bekannteste. Dieses Verfahren sucht bei bekannten gestörten Beobachtungen \mathcal{Y} als Schätzung für die Originaldaten \mathcal{X} genau den Datensatz $\mathcal{X}_{\text{MAP}}^*$, der die a-posteriori-Verteilung $p(\mathcal{X}|\mathcal{Y})$ maximiert. Die Bayesregel formt diese Verteilung um; die Normierung spielt bei der Maximierung keine Rolle.

$$\mathcal{X}_{\text{MAP}}^* = \operatorname{argmax}_{\mathcal{X}} p(\mathcal{X}|\mathcal{Y}) = \operatorname{argmax}_{\mathcal{X}} \frac{p(\mathcal{Y}|\mathcal{X}) \cdot P(\mathcal{X})}{P(\mathcal{Y})} = \operatorname{argmax}_{\mathcal{X}} \frac{\text{Likelihood} \cdot \text{a-priori-Verteilung}}{\text{Normierung}}$$

Es wird also intuitiv unmittelbar einsichtig genau das Originalbild vermutet, das bei gegebener Beobachtung die höchste Wahrscheinlichkeit aller denkbaren Originalbilder hat. Die Bayesregel besagt, daß $p(\mathcal{X}|\mathcal{Y})$ proportional ist zu $P(\mathcal{X}, \mathcal{Y}) = p(\mathcal{Y}|\mathcal{X}) \cdot P(\mathcal{X})$. Die Normierungskonstante ergibt sich aus der Integration über alle \mathcal{X} . Basierend auf den Markov-Zufallsfeldern für die Originaldaten wird nach dem Hammersley-Clifford-Theorem die a-priori-Verteilung $P(\mathcal{X})$ durch die aus der statistischen Physik bekannte Gibbs-Verteilung (normiert mit der Zustandssumme Z) mit der Energiefunktion $E(\mathcal{X})$

$$p(\mathcal{X}) = \frac{1}{Z} e^{-E(\mathcal{X})}$$

beschrieben, die kleine Energien bevorzugt. Die Likelihood $p(\mathcal{D}|\mathcal{B})$ für unabhängige Pixel zerfällt in ein Produkt $\prod_{\mathbf{x}} p(d(\mathbf{x}, t) | b(\mathbf{x}, t))$ über alle \mathbf{x} . Dies gilt für räumlich nicht korrelierte Pixel, wie sie unter H_0 sicher auftreten, aber unter H_1 kaum zu erwarten sind. Eine detaillierte Diskussion dieser Annahme findet sich im Anhang von [Aach, Til und Kaup, André, 1995]. Ungeänderte Pixel an der Stelle (\mathbf{x}, t) verteilen sich wieder die eingangs bekannte Gaußverteilung, für geänderte Pixel nimmt $p(d(\mathbf{x}, t) | b(\mathbf{x}, t))$ unter den bereits kritisch besprochenen Annahmen ebenfalls eine Gaußverteilung mit der Varianz $\sqrt{2}\sigma_c$ an. Die Energiefunktion soll die Änderungskarten raumzeitlich möglichst glatt machen. So kann die Energiefunktion $E_{\mathbf{x},t}(b(\mathbf{x}, t))$ der zu bestimmenden Änderungskarte eine Analyse über alle $n_p = 9$ Paare des zur Entscheidung anstehenden Pixels $b(\mathbf{x}, t)$ mit denen einer 8-er-Nachbarschaft um das Pixel $b(\mathbf{x}, t - \delta t)$ durchführen und aus den n_E unterschiedlich (mit *geändert* und *ungeändert*) besetzten Paaren einen Wert für die ‘‘Unglattheit’’ ermitteln. Die Energiefunktion $E_{\mathbf{x},t} = \gamma n_p^{-1} n_E e$ mit einer positiven Konstanten e (Potential), die als Kostenterm die Unregelmäßigkeit der n_E ungleichen Paare bestraft, ermittelt also die Kosten, die je nach Entscheidung für das zur Untersuchung anstehende Pixel durch die Anzahl der verursachten ungleichen Paare entstehen (Abbildung 2.5). Sie hat einen kleinen Wert, wenn die Entscheidung wenig ungleiche Paare erzeugt, also glatte Änderungskarten generiert.

Der MAP-Schätzer entscheidet über die Hypothesen so, daß die a-posteriori-Wahrscheinlichkeit maximal wird. Dieser Ansatz kann mit $T = 1$ und der Bayes-Regel umgeschrieben werden.

$$\begin{cases} \text{Wenn } P(\mathcal{B}(b(\mathbf{x}, t) = \textit{ungeändert}) | \mathcal{D}) \geq P(\mathcal{B}(b(\mathbf{x}, t) = \textit{geändert}) | \mathcal{D}) & \text{dann } H_0 \\ \text{Wenn } P(\mathcal{B}(b(\mathbf{x}, t) = \textit{ungeändert}) | \mathcal{D}) < P(\mathcal{B}(b(\mathbf{x}, t) = \textit{geändert}) | \mathcal{D}) & \text{dann } H_1 \end{cases}$$

$$\frac{p(\mathcal{B}(b(\mathbf{x}, t) = \textit{ungeändert}) | \mathcal{D})}{p(\mathcal{B}(b(\mathbf{x}, t) = \textit{geändert}) | \mathcal{D})} \geq T \quad \begin{cases} H_0 \text{ keine Änderung} \\ H_1 \text{ Änderung vorhanden} \end{cases}$$

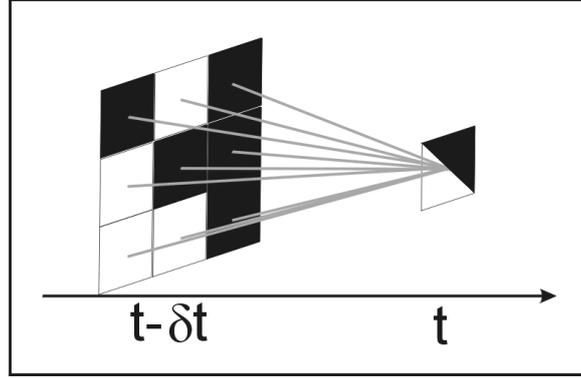


Abbildung 2.5: Je nach Entscheidung in $b(\mathbf{x}, t)$ ergeben sich zur vorangehenden Änderungskarte unterschiedliche Anzahlen von ungleichen Pixelpaaren. Die Energiefunktion bekommt je nach Entscheidung die Terme $2e_v + e_h + e_d$ oder $e_h + 3e_d + e_z$, wenn für vertikale, horizontale, diagonale oder zentrale Paare unterschiedliche Potentiale e_i aufsummiert werden.

mit
$$\frac{p(\mathcal{D} \mid \mathcal{B}(b(\mathbf{x}, t) = \text{ungeändert}))}{p(\mathcal{D} \mid \mathcal{B}(b(\mathbf{x}, t) = \text{geändert}))} = \frac{p(d(\mathbf{x}, t) \mid b(\mathbf{x}, t) = \text{ungeändert}) \cdot \prod_{\mathbf{x}' \neq \mathbf{x}} p(d(\mathbf{x}', t) \mid b(\mathbf{x}', t))}{p(d(\mathbf{x}, t) \mid b(\mathbf{x}, t) = \text{geändert}) \cdot \prod_{\mathbf{x}' \neq \mathbf{x}} p(d(\mathbf{x}', t) \mid b(\mathbf{x}', t))}$$

$$\frac{p(d(\mathbf{x}, t) \mid H_0)}{p(d(\mathbf{x}, t) \mid H_1)} \begin{cases} \geq & T \cdot \frac{P(\mathcal{B}(b(\mathbf{x}, t) = \text{geändert}))}{P(\mathcal{B}(b(\mathbf{x}, t) = \text{ungeändert}))} \\ & \left\{ \begin{array}{l} H_0 \text{ keine Änderung} \\ H_1 \text{ Änderung vorhanden} \end{array} \right. \end{cases}$$

Ohne Vorwissen müssen die a-priori-Wahrscheinlichkeiten gleichverteilt angenommen werden, so daß deren Quotient zu 1 verschwindet und wieder der aus Abschnitt 2.3.5 bekannte Hypothesentest vorliegt. Der letzte Ausdruck wird vom einzelnen Pixel $d(\mathbf{x}, t)$ zu einer sichereren Entscheidungsfindung auf die Stichprobe $\mathbf{d}_{\mathbf{x}}(t)$ ausgedehnt

$$\frac{p(\mathbf{d}_{\mathbf{x}}(t) \mid H_0)}{p(\mathbf{d}_{\mathbf{x}}(t) \mid H_1)} \begin{cases} \geq & T \cdot \frac{\exp\{-E_{\mathbf{x},t}(b(\mathbf{x}, t) = \text{geändert})\}}{\exp\{-E_{\mathbf{x},t}(b(\mathbf{x}, t) = \text{ungeändert})\}} \\ & \left\{ \begin{array}{l} H_0 \text{ keine Änderung} \\ H_1 \text{ Änderung vorhanden} \end{array} \right. \end{cases}$$

$$\left(\frac{\sigma_c}{\sigma_k}\right)^{N_w} \exp\left\{-\left(1 - \frac{\sigma_k^2}{\sigma_c^2}\right) \frac{D_{\mathbf{x}}^2}{2}\right\} \begin{cases} \geq & T \cdot \frac{\exp\{E_{\mathbf{x},t}(b(\mathbf{x}, t) = \text{ungeändert})\}}{\exp\{E_{\mathbf{x},t}(b(\mathbf{x}, t) = \text{geändert})\}} \\ & \left\{ \begin{array}{l} H_0 \\ H_1 \end{array} \right. \end{cases}$$

$$D_{\mathbf{x}}^2 \begin{cases} \leq & t_{\alpha} + 2 \underbrace{[E_{\mathbf{x},t}(b(\mathbf{x}, t) = \text{geändert}) - E_{\mathbf{x},t}(b(\mathbf{x}, t) = \text{ungeändert})]}_{t_E} \\ & \left\{ \begin{array}{l} H_0 \text{ keine Änderung} \\ H_1 \text{ Änderung vorhanden} \end{array} \right. \end{cases}$$

Durch den Zusatzterm t_E werden solche Entscheidungen unterstützt, die $b(\mathbf{x}, t)$ so ähnlich machen wie die Stichprobe $\mathbf{b}_{\mathbf{x}}(t - \delta t)$. Für t_{α} wird die Signifikanzschwelle der $\chi_{N_w}^2$ -Verteilung benutzt, auch wenn dies bei bekanntem σ_c das MAP-Kriterium $T = 1$ verletzt. Wegen der abzählbaren Möglichkeiten für den Zusatzterm t_E läßt sich dieser auch in Tabellenform ablegen. Je nach Ausgestaltung der Energiefunktion kommt es zu verschiedenen vielen Werten für t_E , was letztlich eine gesteuerte Erweiterung des Zwei-Schwellen-Algorithmus aus dem Abschnitt 2.3.5 darstellt. Über die positive Konstante γ in der Energiefunktion wird der Einfluß des Zusatzterms auf t_{α} skaliert. Insgesamt schwankt bei

symmetrischer Energiefunktion das Kriterium zwischen $t_\alpha \pm 2 \ln \gamma e$. Die Auswirkungen des Zusatzterms und das Konvergenzverhalten des Verfahrens lassen sich am sinnvollsten durch Vergleiche der Ergebnisse mit t_α und $t_\alpha + t_E$ und Auszählung der neu getroffenen Entscheidungen in $b(\mathbf{x}, t)$ erfassen. Solange nur die bereits zur Zeit $t - \delta t$ bestimmte vergangene Änderungsmaske für die Energiefunktion benutzt wird, kommt das Verfahren ohne Iteration aus. Symmetrischer wird das Verfahren mit mindestens einer Iteration, die Änderungskarten rund um den zu regularisierenden Bildpunkt (Gegenwart) und der beiden zeitlich benachbarten Bilder (Zukunft, Vergangenheit) einbaut wie in [Bouthemey, Patrick und Lalande, Patrick, 1993].

Eine Glättung der Änderungskarten mit Bayes-Methoden und Markov-Random-Fields hat sich im Vergleich zu Glättungsfiltern als eine weitaus bessere Methode etablieren können, da dabei in steuerbarer und datengetriebener Weise vorliegende Information über Vorhandensein und Berandung von Änderungsregionen einfließt. Eine solche Bayes-Erweiterung gibt es bei [Bouthemey, Patrick und Lalande, Patrick, 1990] [Bouthemey, Patrick und Lalande, Patrick, 1993] auch für die Methode aus Abschnitt 2.3.3.

2.3.7 Ansätze zur Kompensation von Beleuchtungsschwankungen

Um eine Änderungsdetektion in Richtung einer Objektdetektion (Abschnitt 2.3.2) fortzuentwickeln, gibt es Ansätze, die irrelevanten Änderungen wie zum Beispiel schnelle Beleuchtungsschwankungen zu ignorieren. Die von Toth [Toth, Daniel u. a., 2000] (kürzer auch in [Aach, Til u. a., 2000]) in das in den Abschnitten 2.3.5-2.3.6 beschriebene Konzept eingearbeiteten beleuchtungsinvarianten Verfahren trennen die Intensitätsschwankungen durch bewegte oder variierte Beleuchtung von den Grauwertschwankungen, die durch Objektbewegungen verursacht werden. Die Änderungsdetektoren mit Hypothesentests und anschließenden Bayes-Verfahren zur Glättung der Änderungsmasken werden dann nur auf die von den Beleuchtungsschwankungen befreiten Bilddaten angewendet. Die Berechnung von Differenzbildern entspricht einem zeitlichen Hochpaßfilter, der schnelle Änderungen wie Rauschen und das gesuchte Auftauchen von Bewegung passieren läßt, langsame Driften, wie sie von Beleuchtungsschwankungen eingebracht werden, aber unterdrückt. Schnelle Beleuchtungswechsel können so nicht ausgeschlossen werden. Da Beleuchtung immer ein diffuser, großflächiger Prozeß ist, wogegen Rauschen sich in Pixeln abspielt und Bewegungsänderungen zu räumlich sehr kompakten Gebieten gehört, ist zur Unterdrückung von schnellen Beleuchtungsschwankungen ein Hochpaßfilter auch in räumlicher Dimension naheliegend. Für Oberflächen, die sich wie Lambertsche Strahler verhalten, setzt sich die aufgezeichnete Grauwertintensität $g(\mathbf{x}, t) = i(\mathbf{x}, t) \cdot r(\mathbf{x}, t)$ aus einer Beleuchtungskomponente i und den Reflektionseigenschaften r zusammen. Beide sollen separiert werden, um in der Änderungsdetektion nur noch Information aus den durch die Objekte gegebenen Anteil r zu untersuchen. Wendet man auf $\log g = \log r + \log i$ einen Tiefpaßfilter in Form einer Faltung $\log g * h_{\mathbf{x}}$ mit einer räumlichen Gaußmaske an und subtrahiert anschließend, so läßt sich mit diesem homomorphen Filter

$$r \approx \exp \left\{ \log g - \underbrace{(\log g * h_{\mathbf{x}})}_{\approx \log i} \right\}$$

jedes Bild der Bildfolge annähernd von schnellen Beleuchtungsschwankungen bereinigen. Als Folge von Differenzbildern kann dann $d(\mathbf{x}, t) = r(\mathbf{x}, t) - r(\mathbf{x}, t - \delta t)$ in die in den Abschnitten 2.3.5-2.3.6

vorgestellten Verfahren eingehen.

Auch hier bleibt das Verfahren strikt in der Verarbeitung von Differenzbildern, also von Informationen aus zwei Bildern der gesamten bereitstehenden Bildfolge. Bei der Abtrennung der Beleuchtungsvariation wird sogar nur auf einem Einzelbild gearbeitet. Gerade an dieser Erweiterung zeigt sich, wie sehr die *Änderungsdetektion* mit den Einschränkungen aus Abschnitt 2.3.2 in der praktischen Umsetzung zu einer *Objektdetektion* wird. In der Tat repräsentieren schnelle, kompakte zeitliche Änderungen in der Regel die Grauwertregion eines Objektes.

2.3.8 Bewegungsschätzung versus Änderungsdetektion

Wie einleitend festgestellt, hat sich das Arbeitsgebiet der Änderungsdetektion der 80er und 90er Jahre über die Bewegungsdetektion inzwischen in die Fragestellungen der Bewegungsschätzung weiterentwickelt (einen guten Überblick zu dieser Entwicklung gibt [Mítiche, Amar und Bouthemy, Patrick, 1996]). Die Verschiebung der Forschungsschwerpunkte darf aber nicht vergessen machen, daß allen Bewegungsschätzern eine Änderungsdetektion zugrunde liegt, die nach Unterschieden in den einzelnen aufeinanderfolgenden Bildern einer Sequenz sucht.

Bewegung in einer Bildfolge kann charakterisiert werden durch Verschiebungsvektoren, die Bildelemente eines Bildes mit den korrespondierenden Elementen des darauffolgenden Bildes verbinden. Es ergeben sich dann dreidimensionale Verschiebungsvektorfelder aus 2D-Vektoren. Die daraus resultierenden Geschwindigkeitsfelder heißen *Optischer Fluß* [Horn, B. K. P. und Schunk, B., 1981]. Um Optischen Fluß zu bestimmen, gibt es zahllose Methoden [Haußecker, H. und Spies, H., 1999] [Barron, J. L. und Liptay, A., 1994] bis hin zu Bayes-Erweiterungen [Simoncelli, 1999]. Die am häufigsten verwendeten Verfahren lassen sich einteilen in zwei Gruppen: merkmalsbasierte Methoden und Gradienten-basierte Methoden. (Transformationsbasierte Methoden wie Fourierraum-Analyse sei hier ausgelassen, da dieses Konzept in der vorliegenden Arbeit nirgends weiter verfolgt wird.) In der praktischen Umsetzung werden oft Elemente aus beiden Methoden benutzt, um gute Ergebnisse zu erzeugen. Bewegungsschätzung ist, wird häufig gleichzeitig mit Objekterkennung betrieben: in einer Szene werden parallel Objekte gefunden und deren Bewegung bestimmt. Zwei Ansätze existieren dazu, die hier als Extreme gegenübergestellt werden (in der Praxis finden sich eher gemischten Realisierungen, die die Schwächen der einen Methode mit den Vorteilen der anderen in einer Nachverarbeitung abfangen). Man kann die Bildfolge als Einzelbilder einer Segmentierung aufgrund von Grauwerten, Texturmerkmalen, Kanten oder bekanntem Vorwissen über die zu suchenden Objekte unterziehen und den gefundenen Objekten dann in einem Merkmalsraum neben den genannten Merkmalen Positionen und Orientierungen zuweisen. Segmentierung bedeutet also Objektgrenzen und Kantenstücke zu finden, daraus glatte und geschlossene Regionen zu aggregieren. Bewegungsschätzung heißt dann, die Listen von Merkmalsvektoren verschiedener Bilder zu kombinieren und eine Objektverfolgung, und damit Geschwindigkeitsschätzung, auf der Ebene der Merkmale durchzuführen. Ganz anders dagegen der Gradienten-basierte Ansatz: hier werden zu allen Pixeln, deren Objektzugehörigkeiten zu diesem Zeitpunkt nicht bekannt sind, mit Strukturtensoransätzen [Jähne, 1993] [Haußecker, H. und Jähne, B., 1997] und die räumlich-zeitlichen Ableitungen berechnet, um dann aus dem sich ergebenden Verschiebungsvektorfeld für Grauwerte einheitlich (über eine Region in Raum- und Zeitkoordinaten) bewegte Bereiche als Objekte zu identifizieren. Hier wird also aus der rein wahrnehmenden Bewegungs-

detektion eine Schätzung des Bewegungsfeldes. Objektidentifikation und Bewegungsschätzung finden gerade in anderer Reihenfolge statt als beim Segmentierungs-Ansatz. Für die Gradientenmethode existieren von Heitz und Bouthemy auch Bayes-Ansätze [Winkler, 1995]. Die Änderungsdetektoren, die in beiden Ansätzen enthalten sind, suchen Änderungen jeweils in einer sehr kleinen Nachbarschaft aus der gesamten zur Verfügung stehenden Sequenz. Beim Optischen Fluß ist es offensichtlich, daß Änderung in Form der Berechnung der Ableitungen detektiert wird. Dies stellt letztlich eine (eventuell speziell gewichtete) Bildung von Differenzbildern dar. Es gehen also zwei, oder in der Praxis einige mehr, zeitlich aufeinanderfolgende Ausschnitte in die Berechnung ein. Bei der Lösung der *motion constraint equation* [Jähne, 2002] wird ein Funktional über ein kleines räumlich-zeitliches Fenster minimiert. Methoden der Bewegungsschätzung eignen sich vor allem bei kleinen Verschiebungen im Gegensatz zu den in 2.3.3 und 2.3.5 vorgestellten Verfahren, die dann oft in das Blendenproblem der Änderungsdetektion (Abschnitt 2.3.2) hineinlaufen. Schlechte zeitliche Abtastung hat den Vorteil bei den reinen Änderungsdetektoren, daß kein Blendenproblem für die überlappenden Bereiche auftritt. Bewegungsschätzer benutzen in der Regel erheblich größere Bereiche der zeitlichen Dimension, da für die Glättungen in der Berechnung des Optischen Flusses vergleichsweise große Masken benutzt werden. Diese Idee soll später aufgegriffen werden: sie stellt einen erheblichen Unterschied zu den bisher vorgestellten Änderungsdetektoren dar, die ausschließlich auf zwei zeitlich aufeinanderfolgenden Bildern arbeiteten. Bei Merkmals-basierten Verfahren werden Änderungen der Vektoren in der Regel über einige wenige zeitlich aufeinanderfolgende Vektoren bestimmt.

In der gradientenbasierten Bewegungsbestimmung lautet die Gleichung für den Optischen Fluß

$$\nabla g(\mathbf{x}, t) \cdot \mathbf{u}(\mathbf{x}, t) + \frac{\partial g(\mathbf{x}, t)}{\partial t} = 0$$

zu der die bereits von Horn und Schunck vorgeschlagenen Glättungsterme (das Bewegungsfeld $\mathbf{u}(\mathbf{x}, t) = (u_1, u_2)$) soll stückweise glatt sein) hinzu kommen und zur Minimierung von

$$\iint (u_1 g_x + u_2 g_y + g_t)^2 + \gamma^2 (u_{1x}^2 + u_{1y}^2 + u_{2y}^2 + u_{2x}^2)^2 dx dy$$

wobei γ den Einfluß des Glattheitstermes gegen die Gleichung des Optischen Flusses steuert. Nachteil der Methode ist, daß sie mit Diskontinuitäten im Geschwindigkeitsfeld nicht umgehen kann. Robuste Schätzmethoden wurden für das Minimierungsproblem ebenfalls eingesetzt: wenn $r_i = 0$ für die Gradientengleichung im Pixel \mathbf{x}_i steht, so wird nicht $\sum_i r_i^2$ minimiert, sondern $\sum_i \rho(r_i)$ wobei ρ ein M-Estimator (abgeschnittene quadratische Funktion, Tukey-Funktion, Laurentz-Funktion) sein kann, für die der Einfluß von Ausreißern gegen Null geht.

2.3.9 Kontinuierliche Berechnung des Hintergrundes

Einen anderen Ansatz der Änderungsdetektion, vor allem motiviert aus Anwendungen der Sicherheitstechnik und Verkehrs-Überwachungsaufgaben, versucht die Trennung von Vordergrund und Hintergrund. [Donohoe, G. W. u. a., 1988] führte ein Referenzbild ein, in dem der Hintergrund einer Szene möglichst vollständig erfaßt ist. Vor diesem bewegen sich in der Szene die Objekte, deren Positionsänderung erfaßt werden soll. Durch unterschiedliche Schattenwürfe und bei Objektbewegung

freiwerdendem Hintergrund ist eine regelmäßige Neuberechnung des Hintergrund-Referenzbildes nötig. Einfache Formulierungen für das Referenzbild sind

$$g_{\text{ref}}(\mathbf{x}, t) = \alpha g(\mathbf{x}, t) - (1 - \alpha) g(\mathbf{x}, t - \Delta t) \quad \text{mit } \alpha \in [0, 1]$$

Karmann [Karmann, Klaus-Peter u. a., 1990] [Karmann, Klaus-Peter und von Brandt, Achim, 1990] [Kilger, 1992] berechnet dieses Hintergrundbild mit Kalmann-Filtern [Ridder, C. u. a., 1995]. Die Berechnung von Änderungen erfolgt mit solchen Referenzbildern wie in den bereits geschilderten Verfahren zum paarweisen Bildvergleich: in diesem Fall werden nicht zwei aufeinander folgende Bilder der Sequenz verglichen, sondern ein Bild der Sequenz mit dem aktuellen Referenzbild. Letztlich werden dadurch Zuordnungsprobleme (Unterscheidung von Objekten, Schatten, freiwerdender Hintergrund, homogene Objektinnenflächen) aus der Interpretation der Änderungskarte verschoben in die korrekte Berechnung des Referenzbildes, das mehr oder weniger permanent überprüft werden muß, neue Bildinhalte unterschiedlich schnell (hart und weich) integrieren muß und an verschiedenen Positionen auch immer unterschiedliche Gütemaße für den lokal aktuellen Wert hat. Neuere Übersichten dazu geben [Toyama, K. und Wallflower, 1999] und [Makarov, 1996]. Koller [Koller, Dieter u. a., 1994a] führt den Ansatz fort und baut für Verkehrsüberwachungen mit fester Kamera ein System, das Fahrzeuge über Objekterkennung von raum-zeitlichen Kanten des Originalbildes multipliziert mit der binären Änderungsmaske identifiziert. Zur Berechnung der Änderungsmaske $b(\mathbf{x}, t)$ wird die Differenz aus aktuellem Bild und dem Hintergrundbild

$$g_{\text{ref}}(\mathbf{x}, t + \Delta t) = g_{\text{ref}}(\mathbf{x}, t) + [\alpha_1 (1 - b(\mathbf{x}, t)) + \alpha_2 b(\mathbf{x}, t)] (g(\mathbf{x}, t + \Delta t) - g(\mathbf{x}, t))$$

mit einem globalen Schwellwert verglichen (ausführlich in [Koller, Dieter u. a., 1994b]). Das Verfahren konstruiert also aus Differenzbildern fortlaufend ein Hintergrundbild, wobei Stellen mit Änderungen nicht mit in die Konstruktion übernommen werden ($\alpha_1 \gg \alpha_2$). Während die Differenzbildung $g(\mathbf{x}, t) - g_{\text{ref}}(\mathbf{x}, t)$ und Schwellwertbildung zur Ermittlung von $b(\mathbf{x}, t)$ dadurch sehr einfach werden, sind die Probleme der Änderungsdetektion vor allem in die korrekte Konstruktion des Hintergrundbildes $g_{\text{ref}}(\mathbf{x}, t)$ verlagert. Zur korrekten Konstruktion des initialen Hintergrundbildes und für Referenzen auf tiefer gehende Abhandlungen siehe [Gutchessy, D. u. a., 2001] oder [Naoya, 2001] für die Benutzung statistischer Tests zur korrekten Ermittlung des Hintergrundes.

Haußecker [Haußecker, Horst und Jähne, Bernd, 1993] [Haußecker, 1993] berechnet ein Hintergrundbild mit einer zeitlichen Gaußpyramide, bei der die Originalsequenz durch die Maske mit den Gewichten

$$1/16 [1, 4, 6, 4, 1]$$

geglättet bzw. ein Ausschnitt von $\Delta t = 5$ Bildern mit einer besonderen Gewichtung zeitlich gemittelt wird. Wie in Abbildung 2.6 dargestellt, wird aus dem Ergebnis nur jedes zweite Bild benutzt, um mit einer erneuten gewichteten Mittelung über $\Delta t' = 3$ und einer Maske $1/4 [1, 2, 1]$ die erste Pyramidenstufe aufzubauen. Die Aliasing-Effekte durch die Unterabtastung werden durch die zweite Maske reduziert. Höhere Stufen der zeitlichen Pyramide bauen sich nach dem gleichen Rechenprinzip auf. In der höchsten Stufe verbleibt ein einzelnes Bild, das bis auf Randeffekte vom Anfang und Ende der Bildfolge deren zeitlichem Mittelwert $\langle g \rangle_{\Delta T}$ entspricht. Mit der Pyramide liegt abschließend eine Extraktion des ruhenden Hintergrundes auf verschiedenen Zeitskalen (den Pyramidenstufen) vor.

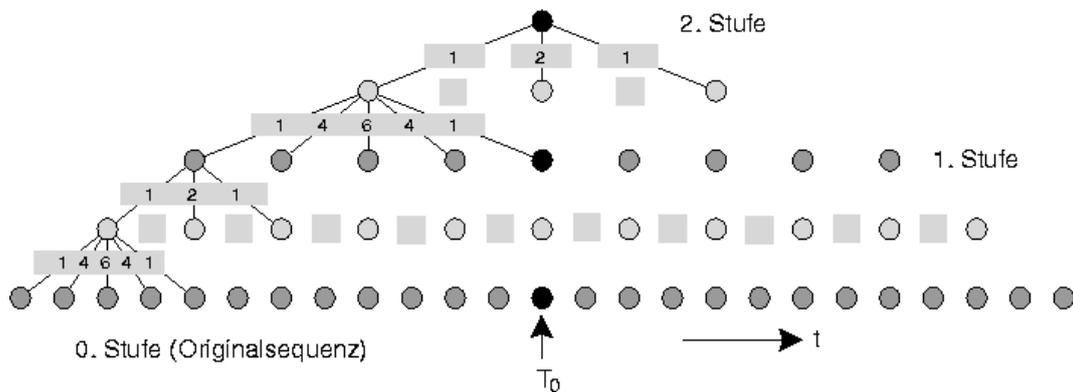


Abbildung 2.6: Konstruktion der von Haußecker [Haußecker, Horst und Jähne, Bernd, 1993] verwendeten zeitlichen Gaußpyramide zur Bestimmung des Hintergrundbildes auf verschiedenen Zeitskalen

Für die Änderungsdetektion schlägt Haußecker eine quadrierte Differenz aus Originalbild und Hintergrundbild vor, aus der eine lokale Stichprobe gezogen wird (räumliche Faltung mit Gewichtung in der Maskenmitte). Dieses Stichprobenmaß wird mit einem lokalen Schwellwert verglichen, der mit Fuzzy-Methoden aus einer lokalen Umgebung des Stichprobenmaßes eines jeden Bildes anhand von zwei Größen ermittelt wird. Diese sind die lokale Varianz des Stichprobenmaßes und der mittlere Wert des Stichprobenmaßes in einer kleinen räumlichen Umgebung um das Pixel, für das ein Schwellwert bestimmt werden soll.

2.3.10 Änderungsdetektoren für Multimediaformate

Eine Renaissance haben Änderungsdetektoren in den späten 90er Jahren durch die Videobearbeitung erfahren. Mit MPEG [Le Gall, 1991] [Sikora, 1997] werden dafür effiziente Kodier- und Kompressionsverfahren entworfen, die nicht die gesamte Bildinformation enthalten, sondern nur die Änderungen auf ein komplett übernommenes Einzelbild (engl. *key frame*) kodieren. Die Suche nach geeigneten *key frames* ist auch die Hauptaufgabe bei Anwendungen der Videoindizierung (engl. *video indexing, video retrieval*) [Ahanger, G. und Little, T. D. C., 1996] [Zhang, H. u. a., 1993]: dabei werden Videodaten automatisch nach inhaltlich interessanten Bildänderungen oder neuen Szeneinstellungen (nach Schnitten, weichen Übergängen, Schwenks, Lichtwechseln, Fahrten, Zooms, Ein/Aus/Überblendungen) durchsucht und an geeigneten Stellen Einzelbilder entnommen, die den Datenstrom inhaltlich partitionieren [Boreczky, J. S. und Rowe, L. A., 1996]. So können in einer einfach zu durchsuchenden Bilddatenbank solche Bündel von Einzelbildern bereits einen guten Überblick zu dem dazugehörigen Videomaterial geben. Beide Anwendungen bedürfen einer automatisierten Detektion von Szenengrenzen (engl. *shot boundary detection*), die der vorgestellten Änderungsdetektion sehr ähnlich ist. Da Videodaten meist bereits komprimiert vorliegen, liegt der Schwerpunkt heute auf Verfahren, in dem die Indizierung direkt auf komprimiertem Material abläuft, ohne dieses zwischenzeitlich zu entpacken [Zhang, H. J. u. a., 1994]. Bewegungsschätzung geht bei heutigen Kodierverfahren für Videos stets über eine Objektsegmentierung, die anhand der Objekte auch ein Verständnis des Inhaltes erlaubt, was ein unsegmentiertes Bewegungsfeld nicht leisten kann. Dufaux

[Dufaux, Frédéric und Moscheni, Fabrice, 1995] gibt dazu eine gute Übersicht der Verfahren.

[Bouthemy, Patrick u. a., 1999] verwendet eine sehr einfache Bewegungsschätzung von sechs Parametern der dominanten Bewegung und ermittelt daraus, ob Zoomen, Schnitte, Fahrten etc. vorliegen.

Sethi [Sethi, Ishwar und Patel, Nilesh V., 1995] identifiziert einen Histogrammvergleich zweier Ausschnitte aus aufeinanderfolgenden Bildern als geeignetes Verfahren, der für entlang einer Zeitachse ein skalares Maß für die Übereinstimmung der Histogramm-Formen berechnet

$$\sum_{\forall i} \frac{(H(i, g(x, y, t)) - H(i, g(x, y, t + 1)))^2}{(H(i, g(x, y, t)) + H(i, g(x, y, t + 1)))^2}$$

Dieses Maß ist χ^2 -verteilt. Alle von ihm untersuchten Verfahren beruhen auf dem Vergleich zweier aufeinanderfolgender Bilder.

Die in komprimierten Videos direkt zugänglichen Werte der DCT (Discrete Cosine Transform) sind stellen sich als gute Grundlage für die Änderungsdetektion dar.

Moderne Verfahren zur Detektion von Szenenwechseln direkt in MPEG-komprimierten Videodaten beschreiben [Feng, J. u. a., 1996] [Fernando, W. A. C. u. a., 1999a] [Fernando, W. A. C. u. a., 1999b].

2.3.11 Schlußbemerkung

In der Bildverarbeitung sind zahllose Verfahren zur Änderungsdetektion bekannt und gebräuchlich. Die hier vorgestellten wichtigsten und theoretisch fundierten Verfahren der expliziten Änderungsdetektion aus der Bildvorverarbeitung (engl. *early vision*) werden ergänzt durch viele implizite Verfahren, die bei der Bildfolgenanalyse eingearbeitet wurden. Aus den Erkenntnissen der Hypothesentests, Stichprobenwahl, Möglichkeiten zur Glättung und der Wahl der gegeneinander zu vergleichenden Daten (bis auf die Detektoren mit Hintergrundreferenzbildern wurden stets Differenzbilder zugrunde gelegt, also nur zwei Bilder einer ganzen Bildfolge benutzt) werden wir im folgenden Kapitel 3 eine für unsere Zwecke passende Änderungsdetektion aufbauen.

Kapitel 3

Statistischer Ansatz zur Detektion von Ereignissen

Unser Ziel ist es, stark verrauschte Prozesse oder Prozesse mit typischerweise hohen Variationen zu überwachen und Ausreißer zu detektieren, die nicht den üblichen Variationen entsprechen, sich aber in deren Größenordnungen bewegen. Bei der Verarbeitung von Bildsequenzen, also dreidimensionaler (x, y, t) Daten, sind dafür Anwendungen in der Überwachung im weitesten Sinne denkbar, so daß die hier zu entwickelnde Algorithmik als Alarm-Geber oder als ein Modul der Datenvorverarbeitung zur Überprüfung der Datenkonsistenz dienen kann.

3.1 Überblick

In diesem Kapitel wird die in dieser Arbeit verwendete Rechenmethode entwickelt. Im ersten Teil wird gezeigt, wie aus den Bilddaten Regionen extrahiert werden können, die untypische Ereignisse aufweisen. Dieser Ansatz nimmt in Details einige Ideen früherer Arbeiten auf, setzt jedoch insgesamt wegen der typischen Eigenschaften (Abschnitt 3.2) der hier zur Untersuchung anstehenden Daten neu an und beschreibt ein bisher auch in angrenzenden Gebieten nicht bekanntes Verfahren (Abschnitt 3.3). Die identifizierten Regionen werden als Objekte verstanden, und es werden Merkmale zugeordnet: klassische Verfahren (Abschnitt 3.4) werden dazu ausführlich diskutiert. Schließlich leitet ein Abschnitt zur geeigneten Merkmalsauswahl über in die vergleichende Darstellung von Klassifikatoren (Abschnitt 3.5). Damit liegen die Werkzeuge für zwei Anwendungen bereit, die in den nachfolgenden Kapiteln 4 und 5 beschrieben werden.

3.2 Bildinhalte und Bildverstehen bei Sequenzen

Im Rahmen dieser Arbeit sollen Bildsequenzen eines häufig in industriellen oder wissenschaftlichen Anwendungen vorkommenden Typs bearbeitet und eine genau dafür passende Detektion von Ereignissen eingeführt werden. Das Charakteristische an den betrachteten Sequenzen ist, daß sie stark verrauschte festen Szenen eines Prozesses zeigen, in denen überwiegend inhaltlich wenig passiert – bis

auf gelegentlich auftretende völlig untypische Ereignisse (Sensorausfälle, Zusammenbruch oder Ausbruch aus dem normalen Prozeß, hinzutretende Fremdkörper, Beleuchtungsausfälle, Verdeckungen). Diese sollen von der Algorithmik sicher gefunden werden und unter Angabe eines Maßes für die Sicherheit der Detektion und der Angaben über die Position im (x, y, t) -Bildstapel automatisch quantifiziert werden.

Kamera und Szene sollen zueinander fest sein: dies stellt eine typische Situation in der Überwachung oder Datenaufnahme dar, ist aber im Vergleich zu Filmen oder Multimedia-Daten (Schwenks, Schnitten, Überblendungen, Perspektivenwechsel, Kamerafahrten, Zoomen, ...) eine Einschränkung. Dennoch sind eine sehr große Anzahl anfallender Bilddaten von zeitlichen Abläufen in wissenschaftlichen, industriellen oder technischen Anwendungen von solchem Inhalt. Die gestellte Aufgabe ist also der Entwurf einer Überwachung für stückweise stationäre Prozesse. Der Bildinhalt der hier untersuchten Arten von Sequenz sei aufgeteilt in die folgenden Gruppen:

- \mathcal{A}) Objekte und Vordergrund vor einem gleichmäßigen Hintergrund
- \mathcal{B}) Pixel-Rauschen, das durch Quanteneffekte und die Sensorik entsteht
- \mathcal{C}) bekannte und erwartete Veränderungen im Vordergrund oder Hintergrund
- \mathcal{D}) unerwartete Änderungen (auftauchende Objekte, Sensordefekte, ...)

Durch das Rauschen \mathcal{B}) begrenzt, ist es unser Ziel, Inhalte der Gruppe \mathcal{D}) zu detektieren, ohne jedoch auf erwartete Ereignisse aus \mathcal{C}) zu reagieren. Die zu vermeidende Erfassung von Änderungen aus \mathcal{B}) und \mathcal{C}) bezeichnen wir als Fehldetektionen (engl. *false positives*).

Eine intuitive Herangehensweise wäre die Bildung von Differenzbildern aus jeweils zwei in der Sequenz aufeinanderfolgender Bilder. Darin die Veränderungen zu erkennen, läuft in zwei Probleme hinein: Veränderungen tauchen mit umgekehrten Vorzeichen jeweils doppelt, beim Erscheinen oder Verschwinden, auf. Starkes Rauschen und vom Prozeß selbst eingebrachte Änderungen (\mathcal{B} und \mathcal{C}), lassen sich in der Folge von Differenzbildern nicht mehr von den eigentlich gesuchten Veränderungen (\mathcal{D}) unterscheiden. Die reine Differenzbildung führt außerdem in das ganz praktische Problem, die geeigneten Schwellwerte zu finden: meist kommen hier mehr oder weniger unsicher bestimmte empirische Werte zum Einsatz, die auf die einzelne Anwendung hin ausgewählt wurden. Zahlreiche Gründe führen also dazu, eine einfache Differenzbildung aufeinanderfolgender Bilder zur Lösung des Detektionsproblems zu verwerfen.

Das als stark angenommene Rauschen hat die Größenordnung der gesuchten zeitlichen Änderungen (engl. *temporal change*). Was unterscheidet die Gruppen \mathcal{B}) und \mathcal{D})? Rauschen ist unkorreliert! Es verteilt sich gleichmäßig und zufällig in seinen Werten und Positionen in (x, y, t) ; die gesuchten zeitlichen Änderungen sind gleichmäßig über alle Pixel verteilt. Sie haben andere Ursachen als das Rauschen und gruppieren sich zeitlich und örtlich in der Sequenz. Die Konzentration der gesuchten Änderung gegenüber der Gleichmäßigkeit des Rauschens wird im hier benutzten Verfahren genutzt. Darüber hinaus wird die gesamte in der Sequenz vorhandene Information für die Änderungsdetektion benutzt, während die Methode mit Differenzbildern immer nur zwei Bilder der Sequenz gegeneinander vergleicht. Indem wir die gesamten Daten der Sequenz zum Abgleich gegen ein einzelnes zu

untersuchendes Bild benutzen werden, fällt auch das Rauschen \mathcal{B}) mit der Wurzel der Anzahl benutzter Bilder und kann somit effizient unterdrückt werden gegenüber den gesuchten Erscheinungen aus \mathcal{D}).

Ziel ist es also, den generellen Inhalt \mathcal{A}) und seine Variationen \mathcal{C}) durch statistische Methoden aus der gesamten Bildfolge zu rekonstruieren und dabei das Rauschen \mathcal{B}) zu unterdrücken. Beim Vergleich einzelner Bilder der Sequenz mit dem rekonstruierten generellen Aussehen und den generell erwarteten Variationen treten dann schlagartige und kurzzeitige Änderungen deutlich zutage. Diese in räumlicher und zeitlicher Ausdehnung kompakten Bereiche einer Sequenz werden als die gesuchten Ereignisse interpretiert und später als segmentierte Gebiete analysiert.

3.3 Herleitung der varianznormierten Mittelwertabweichung

3.3.1 Analogie zum Inversen Kontrastverhältnis

Sei die Grauwertintensität der n -ten Bildfolge mit $g(x, y, t, n)$ bezeichnet, wobei die räumlichen Koordinaten häufig mit $\mathbf{x} = (x, y)^T$ abgekürzt werden, und $\langle \cdot \rangle_{i_s, \Delta i}$ der Mittelwert bezüglich einer Dimension $i \in \{x, y, \mathbf{x}, t, n\}$ über den Bereich Δi (ΔI bezeichnet das gesamte zur Verfügung stehende Intervall) beginnend von i_s . Das zeitliche Mittel eines Ausschnittes Δt ist zum Beispiel

$$\langle g(x, y, t, n) \rangle_{t_s, \Delta t} = \frac{1}{\Delta t} \sum_{t'=t_s}^{t_s+\Delta t} g(x, y, t', n) \quad (3.1)$$

Die Idee für die Änderungsdetektion kommt von der Fähigkeit des menschlichen visuellen Systems, Ereignisse zu bemerken, indem ein lokales Kontrastmaß benutzt wird. In der Bildverarbeitung wird das Inverse Kontrastverhältnis definiert

$$\frac{\langle g(\mathbf{x}) \rangle_{\mathbf{x}, \Delta \mathbf{x}}}{\left\langle \left(g(\mathbf{x}) - \langle g(\mathbf{x}) \rangle_{\mathbf{x}, \Delta \mathbf{x}} \right)^2 \right\rangle_{\mathbf{x}, \Delta \mathbf{x}}^{\frac{1}{2}}} \quad (\text{Inverses Kontrastverhältnis}) \quad (3.2)$$

Dadurch werden Kanten verstärkt, weißes Gauß-Rauschen um den Faktor $(\Delta X)^{-1}$ in einer Umgebung idealer konstanter Intensität unterdrückt. Die Mittelwertbildung entspricht einer räumlichen Faltung mit einem gleich-gewichteten Rechteckfenster. Als *Statistische Skalierung* (engl. *statistical scaling*) wird (3.2) bezeichnet, wenn die Mittelung im Zähler entfällt [Wallis, 1976]. Beide Transformationen schwächen feine Strukturen, lassen aber den visuellen Eindruck eines Bildes bestehen. Die Differenz beider Ausdrücke gibt an, zu welchem Grad ein einzelnes Pixel im Vergleich zu seiner räumlichen Umgebung variiert, wobei diese Abweichung als Vielfaches der mittleren Abweichung aller Pixel in dieser Umgebung angegeben wird. Durch die Normierung treten gleich starke Abweichungen in einer ohnehin “unruhigen Umgebung” gegenüber solchen in “ruhigen” Umgebungen zurück. Dieses Konzept wird aus den Dimensionen des Raumes übertragen auf die Dimension Zeit, um Abweichungen in zeitlicher Abfolge zu finden.

3.3.2 Varianznormierte zeitliche Abweichung

Für die n -te Sequenz wird an jeder Stelle x eines Bildes die Differenz zum zeitlichen Mittelwert dieses Pixels in der gesamten Bildfolge normiert auf dessen zeitliche Standardabweichung. Dies geschieht für alle Bilder, so daß das Ergebnis $f(x, y, t, n)$ eine varianznormierte Abweichung vom zeitlichen Mittelwert darstellt. Ist die Variation der Intensitätswerte eines Pixels über alle t bereits groß, so wird die Differenz im Zähler durch einen großen Nenner zu einem kleineren Wert skaliert. An Stellen in der Sequenz, die fast keine zeitliche Variation zeigen, wird bereits eine kleine Abweichung des momentanen Grauwertes vom Mittel als ernst zu nehmende Änderung interpretiert.

$$f(x, y, t, n) = \frac{g(x, y, t, n) - \langle g(x, y, t, n) \rangle_{t_s, \Delta t}}{\sqrt{\langle [g(x, y, t, n) - \langle g(x, y, t, n) \rangle_{t_s, \Delta t}]^2 \rangle_{t_s, \Delta t}}} \begin{cases} \geq \theta & \text{zeitliche Änderungen } \mathcal{D} \\ < \theta & \text{keine Änderungen} \end{cases} \quad (3.3)$$

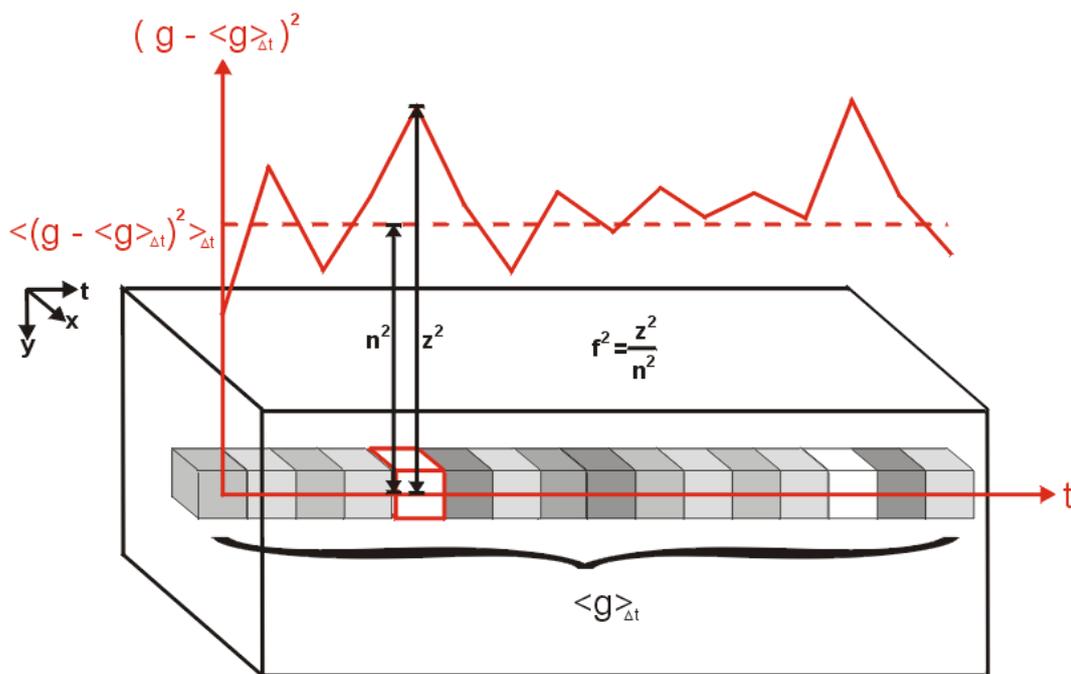


Abbildung 3.1: Graphische Erläuterung der varianznormierten zeitlichen Abweichung: ein Pixel an der Stelle x wird über die Zeitachse gemittelt und seine seitliche Varianz berechnet. Zu jedem Pixelwert wird der Quotient aus Abweichung und mittlerer Abweichung gebildet.

3.3.3 Nachbarschaften

Das vorgestellte Verfahren (3.3) ist völlig frei von jeglichen Nachbarschaftsbeziehungen: vertauschen wir die x -Positionen der zeitlichen Werte der Pixel, so erhalten wir für die Teststatistik f im vertausch-

ten Koordinatensystem dasselbe Ergebnis. Auch eine Neuordnung der einzelnen Bilder innerhalb des analysierten Intervalles Δt würde keine Änderung bringen. Vertauscht man Pixel sogar gleichzeitig in x und t , so ändert dies in den Ergebnissen in f nichts. Formal bedeutet das mit einem Vertauschungsoperator $\mathcal{V}_{x,y,t}$, daß $f(g) = \mathcal{V}_{x,y,t}^{-1} \otimes f(\mathcal{V}_{x,y,t} \otimes g)$. Die Berechnung der varianznormierten Mittelwertabweichung f verzichtet also völlig auf Informationen aus den Nachbarschaftsbeziehungen von Ausreißern. Einleitend hatten wir in Abschnitt 3.2 gerade die Tendenz der Abweichungen (\mathcal{D}) zur Bildung von kompakten Objekten als wichtiges Charakteristikum ausgemacht, um sie von eventuell gleich starkem, aber zufällig verteiltem Rauschen (\mathcal{B}) zu trennen. Während für die Berechnung von Mittelwerten und Varianzen in Zähler und Nenner von (3.3) selbstverständlich keine Beschränkung auf Ausschnitte stattfinden darf, sondern die gesamte zur Verfügung stehende Information genutzt werden muß, um das Pixelrauschen effizient zu unterdrücken, so muß bei der Identifikation der Abweichungen sehr wohl deren Nachbarschaftskorrelation genutzt werden, um sie überhaupt vom Rauschen trennen zu können.

Dafür wird die Teststatistik f einer Nachbarschaftsanalyse unterzogen, die für Pixel, die potentiell die Eigenschaften einer Abweichung zeigen (noch ist unklar, ob sie zur Gruppe \mathcal{C} oder \mathcal{D} gehören), also hohe Werte in f zeigen, deren zeitliche und räumliche Nachbarn analysiert. Zeigen auch sie hohe, vielleicht sogar ähnliche Werte in f , so wird das untersuchte Pixel zur Gruppe der gesuchten Abweichungen \mathcal{D} gerechnet, andernfalls handelt es sich um gleichverteiltes Rauschen. Diese Nachbarschaftsanalyse kann mit einem dreidimensionalen Filter geschehen, mit dem f gefaltet wird und welches den Wert f von Pixeln mit hohen f -Werten und ebensolchen Nachbarn weiter anhebt oder bei fehlender Unterstützung aus der Nachbarschaft den hohen f -Wert eines Pixels senkt. Sinnvoller ist es, eine Präferenz oder Vernachlässigung von einzelnen Pixeln je nach deren f -Werten und Nachbarschaften nicht in eine Modifizierung des Ausdruckes von f einfließen zu lassen, sondern im Rahmen einer separaten morphologischen Operation auf der Teststatistik f durchzuführen. Damit beschäftigt sich Abschnitt 3.3.4.

Bisher aus der Literatur bekannte Verfahren gehen beim Umgang mit Informationen aus der Nachbarschaft nicht zweigleisig vor, sondern beschränken sich entweder durchgehend auf Ausschnitte (Differenzbilder, Analyse von je drei aufeinanderfolgenden Bildern, Ableitungen in zeitlicher Richtung) oder untersuchen wie bei den Überwachungsalgorithmen bzw. Multimediakompressionsverfahren stets die volle Information der Bildfolge. Während erstere Schwierigkeiten mit dem vom Rauschen eingebrachten Änderungen haben (die im kleinen Ausschnitt eben nicht mehr von tatsächlich gesuchten Ereignissen zu unterscheiden sind), bedürfen letztere zusätzlicher Methoden des background-update oder key-frames, um die zuviel gesammelte Information wieder zu eliminieren. Die Störungen sind dabei nichts anderes als *masking* und *swamping* (Seite 10).

In der strengen Trennung der Behandlung der Nachbarschaften (Nachbarschafts-freies Verfahren zur Ermittlung von statistischen Schätzern und Nutzung der Nachbarschafts-Beziehungen bei der Identifikation) liegt im Vergleich zu den bisher bekannten Verfahren die eine der wesentlichen Neuerungen der hier vorgestellten Methode.

3.3.4 Schwellwertbildung und Bemerkungen zur Teststatistik

Liegt $f(x, y, t, n)$ über einer bestimmten Schwelle θ , so geht man davon aus, daß das Pixel an der Stelle (x, y, t, n) eine relevante zeitliche Änderung aufweist, andernfalls nicht.

$$b(x, y, t, n) = \mathcal{T}_\theta [f(x, y, t, n)] \Leftrightarrow b(x, y, t, n) = \begin{cases} 0 & f(x, y, t, n) \geq \theta \quad \text{zeitliche Änderungen } \mathcal{D} \\ 1 & f(x, y, t, n) < \theta \quad \text{keine Änderungen} \end{cases} \quad (3.4)$$

In der binären Maske $b(x, y, t, n)$ werden Positionen, an denen sich in $f(x, y, t, n)$ Abweichungen fanden, zu 0 gesetzt; sonst ist $b(x, y, t, n) = 1$. Das hier vorgestellte Verfahren zeichnet sich gegenüber intuitiven Ansätzen gerade dadurch aus, daß die Auswahl der Schwellwertes θ bei dieser Schwellwertoperation $\mathcal{T}_\theta [f]$ nicht durch die einzelnen Anwendungen geschieht. Im theoretischen Abschnitt 2.1 (auf Seite 8) wurde der Schwellwert θ zurückgeführt auf die beiden Größen

- Stichprobengröße N
- Signifikanzniveau α

Für die später bearbeiteten Anwendungen haben wir laut Tabelle 2.1 (Seite 8) typischerweise Werte im Bereich $2,5 \leq \theta \leq 3,5$, basierend auf Stichprobengrößen $50 \leq N \leq 500$ und nutzen Signifikanzen von 10 %.

In den Anwendungen werden wir sehen, daß – ähnlich wie in [Aach, 1993] (Seite 76) beschrieben – die genaue Wahl des Signifikanzniveaus α keine entscheidende Rolle spielt. Typischerweise benutzt man in der Statistik zur Ausreißerdetektion die Werte 0,1 oder 0,05. Eine nachträgliche Begründung kann durch einen Vergleich der Verteilungen der Teststatistiken von einerseits Ausreißer-befreiten Daten und andererseits ausschließlich der Ausreißerpunkte (dies entspricht der Abbildung 2.4 links) angestellt werden. Der Wert α ist sinnvollerweise für alle (x, y, t, n) gleich.

Die Stichprobengröße N ist anfangs gleich der Anzahl Δt der im stationären zeitlichen Ausschnitt der Bildfolge enthaltenen Einzelbilder. Im Verfahren mit mehreren Detektionsschritten, das erst in Abschnitt 3.3.6 beschrieben wird, ist dies für höhere Iterationen nicht mehr der Fall. Durch Herausnahme ortsabhängig unterschiedlicher Anzahlen von Änderungen bei der iterativen Neuberechnung der Teststatistik $f^{(i)}$ wird die Stichprobengröße dann ortsabhängig kleiner. Die Summe $\sum_{\Delta t} b(x, y, t, n)$ gibt für die Sequenz n an allen Orten \mathbf{x} an, wie groß die Stichprobe ist. Aus diesem Stichprobengrößen-Bild lassen sich für ein gegebenes α alle Schwellwerte $\theta_{\alpha; N(\mathbf{x})}$ berechnen. Bei den Anwendungen wird später festgestellt, daß sich erst Änderungen des Schwellwertes um Beiträge von 0,1 und größer auf das visuelle Ergebnis auswirken. Bei konstantem Signifikanzniveau α entsprechen solche Änderungen einer Abnahme der Stichprobengröße um rund 25% bis 35%. Eine Berechnung von $\theta_{0,1; N(\mathbf{x})}$ für jede einzelne vorkommende Stufe in N würde also viel Rechenzeit kosten und das Ergebnis nicht entscheidend genauer machen im Vergleich zu einer stufenweisen Erhöhung. Es macht daher Sinn, eine Look-up-Table, ähnlich wie bei der als “13-Schwellwerte-Verfahren” bekannten Methode in [Aach, Til u. a., 1993a], anzulegen und in höheren Iterationen, wenn die Kontrolle durch $\sum_t b$ dies anzeigt, den Schwellwert anzupassen. In der Regel ist Vorsicht geboten, wenn sich die Stichprobengröße um 30% oder mehr verringert, da dann auch iterative Verfahren am relativ hohen Anteil von Änderungen

scheitern können. Schließlich bedeutet eine Stichprobengröße, die um 50 % verkleinert wurde, daß Änderungen und “gute” Datenpunkte gleich häufig vorkommen und letztlich nicht mehr zu entscheiden ist, was nun die “guten Daten” und was deren Störungen sind.

Anstelle eines rein pixelweise angewendeten Schwellwertes θ benötigt man für die geforderte Nachbarschaftsanalyse eine Umgebung, die um jedes zu untersuchende Pixel gelegt wird, um eine Bewertung der Geschlossenheit einer abweichenden Region zu bekommen, wie dies in Abschnitt 2.3.6 mit Bayes-Glättungen realisiert ist. Es würde das Verfahren überfrachten, beliebig große Umgebungen zuzulassen und datengetrieben deren Größe und Gewichtungen der mehr oder weniger entfernten Pixel anzupassen. In der Regel wird eine Umgebung fester Größe, die nächste und übernächste Nachbarn in die Bewertung einbindet, gute Ergebnisse liefern. Die Koordinaten x und y können gleich behandelt werden, wenn nicht gezielt nach nicht-isotropen Änderungen (Ausfälle von Zeilen oder Spalten eines Sensorfeldes) gesucht wird. In Anlehnung an Bouthemy [Bouthemy, Patrick und Lalande, Patrick, 1990] [Bouthemy, Patrick und Lalande, Patrick, 1993] (hier werden in t -Richtung keine diagonalen Nachbarn, sondern nur Nachbarn ersten Grades berücksichtigt) ist es bei relativ schnell bewegten Objekten sinnvoller, in zeitlicher Ausdehnung weniger Nachbarn mit einzubeziehen. Anstelle einer echten Faltung mit einer 3D-Gewichtsmatrix mag eine Schwellwertoperation (für Nachbarn ersten Grades θ_1 , Nachbarn zweiten Grades θ_2 mit $\theta_1 < \theta_2$) und anschließende Zählung der detektierten Nachbarn eine angemessene Beschleunigung sein. Die Schwellwerte sollten keine (mehr oder weniger willkürlich) fest vorgegebenen Parameter sein, sondern sich am Wert $f(x, y, t, n)$ des untersuchten Pixels ausrichten. Haben mehr als 30% bis 50% der Nachbarn ebenfalls hohe Werte in der Teststatistik, so ordnet man das untersuchte Pixel zu den gesuchten Abweichungen vom Typ \mathcal{D} zu und setzt in der binären Maske an der Pixelposition den Wert auf 0; andernfalls bleibt er bei 1. Je stärker sich die gesuchten Abweichungen im Mittel von der durch das Rauschen gegebenen Varianz abheben, desto eher kann auf eine umfangreiche Nachbarschaftsanalyse und genaue Untersuchung des Verfahrens mit künstlichen Daten verzichtet werden.

Damit sind die in kompakten Nachbarschaften auftretenden gesuchten Änderungen von den gleichverteilten Rauschpixeln getrennt. Eine Rückkoppelung dieses Ergebnisses in die Berechnung von f wird im Abschnitt 3.3.6 eingeführt.

Die Größe (hier schon mit der erst in Abschnitt 3.3.6 eingeführten Iteration notiert)

$$\beta(x, y, n) = \frac{1}{\Delta t} \sum_{\Delta t} b(x, y, t, n) = \left\langle \prod_{j=0}^i b^{(j)}(x, y, t, n) \right\rangle_{\Delta t}$$

gibt an, welcher Anteil an Pixeln über das Zeitintervall Δt an der Stelle \mathbf{x} des n -ten Datensatzes keine starken Abweichungen aufwies. Dies kann zur Absicherung der Berechnung eine wichtige Größe sein, denn bei einem zu geringen Anteil “sauberer” Daten kommt es in Mittelwert und Standardabweichung zu Verzerrungen.

Mit den Verfahren aus (3.3) und (3.4) kann – wie in Abschnitt 3.2 entworfen – der typische Bildinhalt \mathcal{A}) durch den Mittelwert $\langle g(x, y, t, n) \rangle_{t_s, \Delta t}$ vom aktuell zu untersuchenden Bild $g(x, y, t, n)$ entfernt werden und die verbleibende Differenz unter Berücksichtigung der in der Folge enthaltenen Schwankungen \mathcal{B}), die durch die Varianz im Nenner von (3.3) berücksichtigt werden, bewertet werden. Das Rauschen \mathcal{C}) wird durch die Mittelungen über einen möglichst großen Bereich Δt unterdrückt.

Wenn die Richtung der Ausschläge (Lage deutlich oberhalb oder deutlich unterhalb des Mittelwertes) keine Rolle spielt, dann kann als Teststatistik $|f(x, y, z, t)|$ benutzt werden. Dies führt nochmal zu einer Schreibweise von (3.3), in der sehr deutlich wird, daß eine zeitlich lokalisierte Variation mit der über die Zeit gemittelten Variation verglichen wird

$$\sqrt{\frac{\left[g(x, y, t, n) - \langle g(x, y, t, n) \rangle_{t_s, \Delta t} \right]^2}{\left\langle \left[g(x, y, t, n) - \langle g(x, y, t, n) \rangle_{t_s, \Delta t} \right]^2 \right\rangle_{t_s, \Delta t}}}$$

In der Berechnung von (3.3) kann die Varianz im Nenner auch zu $\langle g^2 \rangle - \langle g \rangle^2$ aufgelöst werden. Insgesamt müssen für den ganzen Ausdruck also zwei Mittelwerte gebildet werden; zusammen mit der Differenzbildung im Zähler benötigt man rechentechnisch daher drei Durchgänge durch die Bildfolge. Interessant für die Bestimmung oberer Schranken von f sind die Summen von f , $|f|$ und f^2

$$\langle f(x, y, t, n) \rangle_{t_s, \Delta t} = 0$$

$$\langle |f(x, y, t, n)| \rangle_{t_s, \Delta t} > \frac{1}{\sqrt{T}}$$

$$\langle f(x, y, t, n)^2 \rangle_{t_s, \Delta t} = 1$$

Bei der Summation über t können Zwischenergebnisse als Schranken benutzt werden. Dadurch kann ein dritter Durchgang bei der Berechnung von (3.3) vermieden werden, wenn ausschließlich die Ungleichung $f \leq \theta$ geprüft werden soll. Der erste Grenzwert ist dafür ungeeignet, weil er durch positive und negative Beiträge um 0 schwankt, die letzten beiden Grenzwert gehen monoton auf $(\sqrt{T})^{-1}$ und 1 zu. Aus den Summen folgt natürlich auch sofort $\sum_t f \rightarrow 0$ und $\sum_t f^2 \rightarrow T$.

3.3.5 Mögliche Modifikationen

Andere Mittelungen in (3.3) wären denkbar: die Mittelung über die Raumkoordinaten wurde weiter oben bereits besprochen, die Mittelung über die n , also die verschiedenen Datensätze, ist denkbar und wird in den Anwendungskapiteln diskutiert. Bei den mehrfachen Mittelungen ist interessant, in der Teststatistik $\langle g(x, y, t, n) \rangle_{\Delta \mathbf{X}}$ anstelle von $g(x, y, t, n)$ zu benutzen und f auf eine Folge von Skalaren zu reduzieren. Dieser läßt bei Mittelung über das gesamte Bild $\Delta \mathbf{X}$ keine Aussage mehr über die räumliche Position der möglicherweise detektierten Änderung zu.

Eine weitere Möglichkeit der Modifikation besteht in der Wahl des Fensters ΔT für die Untersuchung der Bildfolge. Eine Bildfolge, die sich vom Szeneninhalte in stückweise stationäre Abschnitte einteilen läßt, untersucht man sinnvollerweise mit angepaßten Intervallen Δt . Denkbar ist auch eine weniger starre Einteilung in mehrere Intervalle, sondern eine Daten-getriebene Einteilung der Bildfolge, bei der die Intervalle verschieden groß sein können. Als Grundlage für diese Daten-getriebene Segmentierung der Bildfolge in Intervalle stationärer Szenen kann die Mittelung $\langle \langle g \rangle_{\Delta \mathbf{X}} \rangle_{\Delta T}$ anstelle der zeitlichen Mittelung in (3.3) dienen.

Als letzte hier nennenswert relevante Modifikation der Teststatistik (3.3) sei eine unterschiedliche Gewichtung der Werte, die in die Mittelung eingehen, genannt. Die Mittelungen vom Typ (3.1) gewichtet alle Beiträge gleichmäßig. Denkbar wären dabei auch Gewichtungen, die nicht konstant sind über die Größe des Fensters Δi .

Kontinuierliche Veränderungen in Bildfolgen, wie zum Beispiel eine langsame Bewegung der beobachteten Objekte, langsame Beleuchtungsänderungen und dergleichen mehr sind mit der hier vorgestellten Algorithmik auch unter der Modifikation kleiner Mittelungsintervalle schwer gegen die gesuchten abrupten Änderungen abzugrenzen.

3.3.6 Iterativer Ansatz

Die gesuchten zeitlichen Änderungen hinterlassen, auch wenn sie nur in wenigen aufeinanderfolgenden Bildern der Sequenz auftauchen, in der Mittelung und Varianzberechnung in (3.3) eine Spur. Bei der Berechnung von Mittelwert und Varianz (3.3) gehen vorhandene Abweichungen voll in die Ergebnisse ein, die daher nur ungefähre Schätzer der wahren Mittelwerte (inhaltlich die Gruppe \mathcal{A}) und Varianzen (inhaltlich die Gruppe \mathcal{C}) sind. Man könnte dem Argument von Aach [Aach, 1993] (Seite 31) folgen, daß die Schätzwerte genügend genau sind, weil sehr viele Datenpunkte aus Δt eingehen, von denen nur ein geringer Anteil (durch β charakterisiert) zu den Änderungen zählt. Diese Einschätzung beachtet jedoch nicht die aus der statistischen Datenanalyse und Ausreißerdetektion bekannten Ergebnisse. Der unerwünschte Effekt (Mittelwert und Varianz sollen ja gerade die Bildinhalte \mathcal{A} und \mathcal{B} möglichst sauber wiedergeben) ist zwar klein, wenn das Verhältnis der Pixelanzahlen mit Änderungen zur Pixelanzahl ohne Änderungen über die Zeit klein ist, doch bei abschnittweiser Berechnung von (3.3) in den Intervallen $[0, t_1], [t_1, t_1+t_2], \dots$ kann dies zu erheblichen Verfälschungen führen. Mittelwerte werden verschoben und Varianzen heraufgesetzt, was bei mehreren Änderungen zu unterschiedlichen Zeitpunkten sogar dazu führen kann, daß nur die stärkste Änderung detektiert wird und durch deren Einträge in Mittelwert und Varianz jede schwächere maskiert wird. Aus der statistischen Datenanalyse ist dies als Maskierungseffekt (engl. *masking*) bekannt. In extremen Fällen kommt es durch die Mittelwertverschiebung sogar dazu, daß Werte vom Typ \mathcal{A}) als Änderungen detektiert werden, weil sie nach der Verschiebung nicht mehr im Toleranzbereich des Mittelwertes liegen. Dies wird als Überschwemmungseffekt (engl. *swamping*) bezeichnet, da die Liste potentieller Pixel mit den gesuchten Änderungen schlagartig anwächst. Die beiden Effekte *masking* und *swamping* wurden bereits im Abschnitt 2.2.1 (Seite 10) erläutert und in Abbildung 2.1 (Seite 10) illustriert.

Ein weiterer Grund für die nach einem ersten Identifikationsschritt notwendige Neuberechnung von f unter Auslassung bestimmter Pixel in der Schätzung von Mittelwert und Standardabweichung ist die Tatsache, daß vereinzelte Pixel mit starkem Rauschen ebenso wie in kompakten Regionen vorkommende gesuchte Änderungen bei einer ersten Berechnung von f gleichberechtigt eingingen und jeweils hohe Werte in f zeigen. Für die Verbesserung der Mittelwerte und Standardabweichungen dürfen natürlich in einer erneuten Berechnung von f nur Pixel ausgelassen werden, die zu den gesuchten Änderungen gehören, nicht aber solche Pixel, deren hoher f -Wert auf Pixelrauschen zurückgeht (dieses soll gerade mit der Varianz geschätzt werden). Dazu greifen wir bei der erneuten Berechnung der Teststatistik auf die Ergebnisse der morphologischen Nachbarschaftsanalyse zurück: hier wurde bereits erfolgreich zwischen den Gruppen \mathcal{B} und \mathcal{D} getrennt.

Um den Einbezug von identifizierten Ausreißern in die Berechnung der Schätzer zu vermeiden und damit *masking* und *swamping* zu unterdrücken, wird analog zu den in den Abschnitten 2.2.3 bis 2.2.5 vorgestellten Verfahren der Ausdruck (3.3) iterativ (Iterationsindex i in hochgestellten Klammern) berechnet. Zuerst soll nur die Mittelwertberechnung verbessert werden, in einem zweiten Schritt wird die Idee der Verbesserung weiter ausgebaut und auch die Berechnung der Varianz korrigiert. Für die Mittelwertbildung (3.1) wird eine modifizierte Normalisierung benötigt, bei der die in der i -ten Iteration binäre Maske $b^{(i)}(x, y, t, n)$ hilfreich ist.

$$\langle g(x, y, t, n) \rangle_{t_s, \Delta t}^{(i)} = \frac{\sum_{\Delta t} (g(x, y, t, n) \cdot \prod_{j=0 \dots i} b^{(j)}(x, y, t, n))}{\sum_{\Delta t} \prod_{j=0 \dots i} b^{(j)}(x, y, t, n)} \quad \text{mit } b^{(0)}(x, y, t, n) = 1$$

$$f^{(i)} = \frac{g(x, y, t, n) - \langle g \rangle_{t_s, \Delta t}^{(i)}}{\sqrt{\left\langle \left[g(x, y, t, n) - \langle g \rangle_{t_s, \Delta t}^{(i)} \right]^2 \right\rangle_{t_s, \Delta t}}} \quad (3.5)$$

$$b^{(i+1)}(x, y, t, n) = \mathcal{T}_\theta [f^{(i)}]$$

Für die Anfangswerte $\langle g \rangle^{(0)}$ und $f^{(0)}$ werden (3.1) und (3.3) benutzt, denn $b^{(0)}(x, y, t, n) = 1$. Für höhere Iterationen wird die binäre Maske durch die Schwellwertoperation $\mathcal{T}_\theta [f^{(i)}]$ festgelegt. Pixel, die unter der Schwelle θ liegen, bleiben bei der Aktualisierung der Maske $b^{(i+1)}(x, y, t, n)$ auf 1 gesetzt, alle anderen - Pixel mit starken Abweichungen - werden zu 0 gesetzt. Die Produktbildung \prod sammelt die in verschiedenen Iterationen als Änderung detektierten Pixel ein. In die iterative Selektion durch Mittelwert, Varianz und deren Normierungen gehen also nur Pixel ein, die in keinem einzigen Iterationsschritt den Schwellwert θ überstiegen. Dies ist im Sinne von Abschnitt 2.2.3 eine reine Vorwärts-Methode. In (3.5) wird im Unterschied zu (3.3) im Zähler nicht mehr das aktuelle Pixel mit einem durch vorhandene Änderungen verunreinigtem Mittelwert verglichen, sondern mit einem Mittelwert, in dem Pixel, die in allen vorangegangenen Iterationsschritten als Änderung detektiert wurden, unterdrückt sind. Der verbesserte Mittelwert geht im Nenner auch in die Standardabweichung ein. Wichtig im Vergleich zu den bereits vorgestellten Methoden zur Ausreißerdetektion ist auch die Tatsache, daß in jeder Iterationsstufe die originalen Bilddaten mit verbesserten Mittelwerten und verbesserten Standardabweichungen verglichen werden, nicht aber Bilddaten, die von den schon gefundenen Abweichungen "bereinigt" sind. Doch auch die Mittelung $\left\langle (g - \text{const})^2 \right\rangle$ im Nenner von (3.3) zur Berechnung der Varianz muß noch einer solchen iterativen Bereinigung um bereits gefundene Pixel starker Ausschläge erweitert werden: in die Berechnung des für die n -te Bildfolge zweidimensionalen Varianzbildes darf an keiner Stelle auf Daten zurückgegriffen werden, in denen bereits identifizierte starke Abweichungen enthalten sind. Dazu wird die Differenz im Ausdruck für die Varianz analog zur iterativen Mittelwertberechnung mit der binären Maske multipliziert und einer gesonderten Normierung unterzogen, die berücksichtigt, daß an verschiedenen Stellen \mathbf{x} der n -ten Bildfolge unterschiedliche Anzahlen von Abweichungen zu Multiplikationen mit 0 führten. Die sich

daraus ergebende iterative Berechnung ist dann

$$\begin{aligned}
\text{Startwert: } b^{(0)}(x, y, t, n) &= 1 \quad \left(\text{daraus } \mathfrak{b}^{(0)} = \beta^{(0)} = 1 \quad \text{und} \quad \langle g \rangle^{(0)} = \langle g \rangle \right) \\
\mathfrak{b}^{(i)}(x, y, t, n) &= \prod_{j=0}^i b^{(j)}(x, y, t, n) \\
\beta^{(i)}(x, y, n) &= \frac{1}{\Delta t} \sum_{t_s}^{t_s + \Delta t} \mathfrak{b}^{(i)}(x, y, t, n) = \langle \mathfrak{b}^{(i)}(x, y, t, n) \rangle_{t_s, \Delta t} \\
\langle g(x, y, t, n) \rangle_{t_s, \Delta t}^{(i)} &= \frac{\frac{1}{\Delta t} \sum_{t_s}^{t_s + \Delta t} (g(x, y, t, n) \cdot \mathfrak{b}^{(i)}(x, y, t, n))}{\beta^{(i)}(x, y, n)} = \frac{\langle g(x, y, t, n) \cdot \mathfrak{b}^{(i)}(x, y, t, n) \rangle_{t_s, \Delta t}}{\langle \mathfrak{b}^{(i)}(x, y, t, n) \rangle_{t_s, \Delta t}} \\
f^{(i)} &= \frac{g(x, y, t, n) - \langle g \rangle_{t_s, \Delta t}^{(i)}}{\sqrt{\frac{\frac{1}{\Delta t} \sum_{t_s}^{t_s + \Delta t} \left([g(x, y, t, n) - \langle g \rangle_{t_s, \Delta t}^{(i)}]^2 \cdot \mathfrak{b}^{(i)}(x, y, t, n) \right)}{\beta^{(i)}(x, y, n)}}} = \frac{g(x, y, t, n) - \langle g \rangle_{t_s, \Delta t}^{(i)}}{\sqrt{\frac{\langle [g(x, y, t, n) - \langle g \rangle_{t_s, \Delta t}^{(i)}]^2 \mathfrak{b}^{(i)}(x, y, t, n) \rangle_{t_s, \Delta t}}{\langle \mathfrak{b}^{(i)}(x, y, t, n) \rangle_{t_s, \Delta t}}}
\end{aligned} \tag{3.6}$$

update: $b^{(i+1)}(x, y, t, n) = \mathcal{T}_\theta [f^{(i)}]$

Für die erste Stufe $i = 0$ ergibt sich wie in (3.3) $f^{(0)} = f$. Erst in höheren Iterationen werden dann durch die unterschiedlichen $b^{(j)}$, die in \mathfrak{b} aufmultipliziert sind, die besonderen Mittelungen und deren Normierungen notwendig. Die Größe β gibt an, wie sich für eine Pixelposition (x, y) Anzahlen von Pixeln mit und ohne Abweichungen verhalten. Das Integral $\sum_{\Delta x} \beta / \Delta x$ gibt an, wie stark die n -te Sequenz durch solche Abweichungen kontaminiert ist. Es reicht aus, daß ein Pixel auf einer einzigen Iterationsstufe als starke Abweichung identifiziert wurde, um es dauerhaft in allen höheren Iterationen als solches markiert zu lassen. Das Produkt \prod über die $b^{(i)}$ entspricht einer logischen AND-Operation. Als Alternative wäre denkbar, daß ein Pixel durchgängig auf allen Iterationsstufen seine Eigenschaft als starke zeitliche Änderung durch $f \geq \theta$ zeigen muß, was ein Hin- und Herwechseln von Pixeln zwischen den Gruppen \mathcal{A} und \mathcal{D} von Iteration zu Iteration verbieten würde. Dafür muß $\mathfrak{b}^{(i)}$ als OR-Verknüpfung der $b^{(j)}$ unter Auslassung von $b^{(0)}$ berechnet werden. Es lassen sich jedoch Fälle konstruieren, bei denen ohne weitere Konvergenz Pixel von Iteration zu Iteration nur noch zwischen den beiden Zuständen hin- und herspringen. Eine andere Schreibweise des Ablaufs in (3.6) definiert eine iterativ bereinigte Bildfolge (die Koordinatenangaben (x, y, t, n) wurden unterdrückt):

$$g^{(i)} = g^{(i-1)} \mathfrak{b}^{(i)} + \frac{\langle g^{(i-1)} \mathfrak{b}^{(i)} \rangle_{t_s, \Delta t}}{\langle \mathfrak{b}^{(i)} \rangle_{t_s, \Delta t}} \text{Neg} [\mathfrak{b}^{(i)}] \tag{3.7}$$

wobei die Neg-Operation eine binäre Zahl b gemäß $1 - b$ invertiert. Diese Operation wird ausgeführt, wenn $\mathfrak{b}^{(i)}$ berechnet wurde. Sie ersetzt in der Bildfolge $g^{(i+1)}$ überall dort, wo Pixel zu starke Abweichungen in $\mathfrak{b}^{(i)}$ zeigen, diese Werte durch den Mittelwert über die Zeit aller übrigen Werte. Anfangs gilt $g^{(0)} = g$. Da den Bilddaten fortan die Abweichungen aus früheren Iterationen fehlen, muß \mathfrak{b} immer als verkettete AND-Operation der $b^{(i)}$ berechnet werden. Mit den iterativ bereinigten

Bildfolgen lassen sich die komplizierten Mittelungen vermeiden und weiterhin (3.3) benutzen. Beide Rechenmethoden der Iteration sind gleichwertig. Während bei der oben ausführlich vorgestellten Methode Abweichungen durch die Multiplikation mit 0 nicht mehr eingehen, werden bei den bereinigten Bildfolgen die Werte zwar benutzt, aber zuvor durch Werte ersetzt, die bei den Mittelungen keine Beiträge liefern und damit ebenso wie im ersten Fall neutralisiert sind. Interessant bei den bereinigten Bildfolgen ist sicher, daß am Ende ein Datensatz ohne starke Änderungen vorliegt. Fraglich ist dabei jedoch, ob ein reines Ersetzen durch Mittelwerte, also ohne ein Rauschen vom Typ \mathcal{C} , eine physikalisch sinnvolle Veränderung der Daten darstellt. Es wäre zumindest besser, den eingesetzten Werten nicht nur den Mittelwert der Abweichungs-freien Werte zu geben, sondern auch deren Varianz, was jedoch aufwendige und willkürlich ist. Für die praktische Umsetzung der Formeln ist die Methode der bereinigten Bildfolge ungünstig, da die binären Objekte \mathbf{b} und β erheblich weniger Rechenbedarf und Speicherbedarf haben als die iterativen $g^{(i)}$ und deren komplizierte Berechnung.

Da bei der Iteration die Stichprobengröße verringert wird, aus der sich der Schwellwert (Tabelle 2.1) für den Hypothesentest ergibt, muß der Schwellwert in jeder Iterationsstufe für die lokal unterschiedlich großen Stichproben nachgeführt werden. Die Größe $\Delta t \cdot \beta^{(i)}(x, y)$ beschreibt diese Stichprobengrößen $N(x, y)$ und für die Implementierung bietet sich eine Look-Up-Tabelle an, die zu den einzelnen Werten die aus der Tabelle zu entnehmenden Schwellwerte $\theta_{\alpha; N}^{(i)}(x, y)$ in Form eines Bildes ablegt.

Die Iteration kann nach einer festen Anzahl von Schritten oder dann abgebrochen werden, wenn im Iterationsschritt i in der Aktualisierung der binären Maske $b^{(i+1)}(x, y, t, n)$ keine Änderungen mehr stattfinden. Dafür dient als skalares Maß der Anteil der Änderungen in der gesamten Sequenz, der bei hohen Iterationsstufen gegen 0 geht

$$\frac{1}{\Delta T \Delta \mathbf{X}} \sum_{\Delta \mathbf{X}} \sum_{\Delta T} \left| b^{(i+1)}(x, y, t, n) - b^{(i)}(x, y, t, n) \right|$$

3.3.7 Resultierende Größen und Weiterverarbeitung

Als Ergebnisse erhält man aus dem letzten Iterationsschritt dann das Produkt $\mathbf{b} = \prod_j b^{(j)}$ über alle binären Masken der unterschiedlichen Iterationsstufen. Darin auf 0 gesetzte Pixel wurden von der Schwellwertoperation auf der Teststatistik f als Änderungs-behaftet detektiert. Die Stärke dieser Abweichung ist in der zuletzt durchgeführten Iteration der Teststatistik $f^{(i)}$ hinterlegt. Das für die n -te Sequenz zweidimensionale Objekt $\beta^{(i)}$ gibt für jedes Pixel an, welcher Datenumfang (Anteil an der gesamten Sequenz) bei der Berechnung der zeitlichen Mittelwerte auf der letzten Iterationsstufe einging. Dies kann als Gütemaß für die Werte von \mathbf{b} und $f^{(i)}$ dienen. Darüber hinaus geben die eventuell zwischengespeicherten Mittelwerte $\langle g \rangle$ und Varianzen Var (Nenner in der Berechnung von f) visuell einen guten Eindruck über den aufgezeichneten Prozeß. Aus ihnen läßt sich auch das Pixelrauschen als Signal-zu-Rausch-Verhältnis bestimmen, indem man $20 \log(\langle g \rangle / \text{Var})$ für alle (x, y, n) bestimmt und gegen $\langle g \rangle$ aufträgt.

Die mit der Teststatistik, Nachbarschaftsanalyse und Schwellwertoperation \mathcal{T}_θ gefundenen Pixel mit hohen Abweichungen können zu zusammenhängenden zweidimensionalen Regionen pro Bild oder zu dreidimensionalen Regionen in der Bildfolge zusammengefaßt und indiziert (engl. *labelling*) werden. Um vereinzelte Pixel ohne markierte Nachbarn auszuschließen, bieten sich eine Schwelle für

die Größe der Regionen oder klassische morphologische Operationen aus der Binärbildverarbeitung [Soille, 1999a] [Soille, 1999b] an.

An dieser Stelle sei bereits vorweggenommen, daß die in den Kapiteln 5 und 4 zur Auswertung anstehenden Daten zeitlich sehr schlecht abgetastet sind, da sich die Objekte und Regionen mit Änderungen weitaus schneller entwickeln und bewegen, als die Bildrate der eingesetzten Kameras eine echte Verfolgung der Form der Regionen erlaubt. Vielfach decken sich in aufeinanderfolgenden Bildern der Sequenz vorhandene Regionen auch nicht in Teilen, da die Fortbewegung des Objektes zu schnell war im Vergleich zum Bildeinzug der Kamera. Eine Analyse der Regionen mit Änderungen erfolgt daher bildweise im Zweidimensionalen. Dennoch sei die Algorithmik auch für den dreidimensionalen Fall vorgestellt, da in naher Zukunft leistungsstarke Kameras für diese Zwecke zur Verfügung stehen werden.

3.3.8 Zusammenfassung

Zusammenfassend sei festgehalten, daß die varianznormierte Mittelwertabweichung (3.1) und deren iterative Berechnungen die gestellte Aufgabe löst, verschiedene Szeneninhalte zu trennen und Pixel, die schlagartige Änderungen zeigen, sicher und mit einem Gütemaß versehen zu detektieren. Die iterativen Ansätze verhindern Maskierungs- und Überschwemmungseffekte. Die vorgeschlagene Methode ist bei Einsatz einer typischen kritischen Schwelle aus Tabelle 2.1 (Seite 8) parameterfrei. Modifikationen (Betragsbildung, Anwendung auf mehrere Teilintervalle) erlauben eine datengetriebene Analyse von Bildsequenzen, die aus mehreren stückweise stationären Szenen bestehen und Ausschläge in höhere oder niedrigere Grauwertintensitäten zeigen. Die detektierten Pixel werden mit klassischen Methoden der Bildverarbeitung zu Regionen (2D oder 3D) zusammengefaßt und indiziert. Als Ergebnisse liegen dann vor:

- eine binäre Sequenz mit indizierten Regionen, deren Pixel starke Änderungen im zeitlichen Ablauf haben
- die Teststatistik f (letzte durchgeführte Iteration von $f^{(i)}$) als Sequenz mit einer Aussage, wie stark in jedem Pixel ein solcher Ausschlag war
- aus den iterativen Verfahren: β und dessen räumliches Mittel, die beide eine Aussage darüber machen, auf welchem Datenumfang die Mittelwertberechnungen der einzelnen Pixel beruhen und wie häufig Abweichungen in den Pixeln (x, y, n) und der Sequenz n insgesamt vorkommen

Im folgenden Abschnitt wird beschrieben, wie aus dieser noch pixelbasierten Repräsentation der Änderungen eine objektbasierte Merkmalsanalyse der schlagartigen Änderungen wird. Dazu werden die in f identifizierten Regionen als Objekte aufgefaßt, zu denen aus den Werten in f und den dazugehörigen Gütemaßen eine Reihe von Merkmalen extrahiert werden. Die Merkmale aller in einer Bildfolge gefundener Objekte dienen später der Klassifikation (Abschnitt 3.5) der gesamten Bildfolge in Klassen, wie zum Beispiel “starke Abweichungen detektiert” oder “keine Abweichungen detektiert”. Damit läßt sich auf den mit der Bildfolge aufgezeichneten Prozeß zurückschließen, ob dieser einen typischerweise erwarteten Verlauf hatte oder Ausnahme-Ereignisse auftraten.

3.4 Invariante Merkmale aus der momentenbasierten Objektanalyse

3.4.1 Überblick

Zu interessanten und bereits detektierten Regionen im Bildstapel muß ein einfacher Satz von Parametern gefunden werden, der diese Regionen und in unserem Fall auch deren Relevanz für die Ausreißerdetektion charakterisiert und eine Bewertung ermöglicht. Intuitiv wären Größe, umschlossene Werte, Position und äußere geometrische Form eine sinnvolle Auswahl. In diesem Abschnitt wollen wir jedoch ein mathematisch geschlossenes Gerüst für die vollständige Beschreibung von Grauwertbereichen vorstellen, das über unsere beiden konkreten Anwendungen hinaus erweiterbar ist und auch für höherdimensionale Aufgaben brauchbare Parametersätze liefert. Die Objektanalyse, die auf den als “reguläre Momente” oder “zentrale Momente” bezeichneten Integralen der Bilder oder Regionen beruht, wurde historisch gesehen von Ming-Kuei Hu 1961 [Hu, 1961] [Hu, 1962] aus der im neunzehnten Jahrhundert intensiv betriebenen Mathematik der algebraischen Formen und Invariantentheorie (vor allem Boole, Cayley, Sylvester; auch Salmon, Arnhold, Hermite, Eisenstein, Clebsch, Gordan, Lie, Klein, Capelli und Hilbert) auf die Mustererkennung übertragen. In den 70er Jahren wurden zahlreiche Sätze von Momenten vorgeschlagen, aus denen leicht Sätze invarianter Größen, die für die Mustererkennung wesentlichen Größen, abgeleitet werden können. Vergleiche zwischen verschiedenen Sets von Momenten in der Praxis der Klassifikation führten zahlreiche Autoren zum Beispiel [Belkasim, S. O. u. a., 1991] [Belkasim, S. O. u. a., 1993] und [Wood, 1996] durch; frühe Anwendungen konzentrierten sich vor allem auf *template matching* (Schrifterkennung und militärische Aufgaben). Insbesondere die Parallelisierbarkeit der Berechnung der Momente führte zu zahlreichen Publikationen zur schnellen Berechnung von Momenten. Einige wenige Ansätze zur Theorie der Momente in mehr als zwei Dimensionen liegen vor. Der heutige Trend zur multidimensionalen Bildverarbeitung und der Einzug der 3D-Bildverarbeitung mit Volumenbildern oder Zeitserien in die Standardanwendungen, die mit einfacher Hardware bereits lösbar sind, könnte zu einer Wiederentdeckung der höherdimensionalen Aspekte dieses klassischen Werkzeuges der Mustererkennung führen. Ein interessanter Aspekt sind hierbei die Bemühungen um FPGA-Realisierungen binärer oder bereits 8-bit tiefer Momentenberechnung auf Framegrabberkarten oder in intelligenten Kamerasensorboards. Einführungen in die momentenbasierte Objektanalyse finden sich in Lehrbüchern der Bildverarbeitung [Jain, 1989] [Jähne, 2002]. Die für die Bildverarbeitung relevante Mathematik der Invarianten im zweidimensionalen Raum und eine Übersicht zur wissenschaftlichen Literatur liefern Reiss [Reiss, 1993], Mundy [Mundy, Joseph L. und Zisserman, Andrew, 1992] und Prokop [Prokop, Richard J. und Reeves, Anthony P., 1992].

Im folgenden stellen wir die regulären Momente und die zentralen Momente als Merkmale vor und konstruieren daraus die ersten Hu-Invarianten. Legendre-Momente werden im Vergleich dazu vorgestellt, und einige Bemerkungen zur schnellen Berechnung von Momenten und deren Invarianten auch im dreidimensionalen Raum schließen das Kapitel ab.

3.4.2 Merkmale

Mit der Extraktion von Merkmalen aus Bilddaten verfolgt man eine Reduktion auf die in der vorliegenden Anwendung relevante Information in den Daten. Redundanz und uninteressante Signale sollen

dabei möglichst vollständig eliminiert werden. Merkmale sollen untereinander unabhängig sein (verschwindende Kovarianz) und sollten dies auch unter orthogonalen Transformationen bleiben, so daß jedes dazu genommene Merkmal tatsächlich neue Information über die Bilddaten einbringt. Merkmale sollten frei von systematischen Fehlern sein, ihre Streuung sollte bekannt oder abschätzbar sein, und sie sollten robust sein gegenüber Störungen wie Rauschen.

Von besonderem Interesse sind Invarianzen in Merkmalen. So wird in der Bildverarbeitung oft eine Unabhängigkeit der Merkmale von der Lage der Objekte im Bild gewünscht, eine andere Invarianz wäre die Konstanz von Merkmalen bei Skalierung der Bilder (Vergrößerung, Zoomen). Oft wird auch gewünscht, daß Merkmale unabhängig vom Drehwinkel der Bilder oder der darin enthaltenen Objekte sind. Gewünschte Invarianzen hängen jedoch sehr stark von der Aufgabenstellung ab: Bei der Prüfung auf Vorhandensein eines Objektes wird man sehr wohl Translationsinvarianz der Merkmale fordern, aber bereits bei der Vermessung der Lage des Objektes anhand von Merkmalen widerspräche dies der Aufgabenstellung. Andere Beispiele wären: bei fest installierten Kameras und fixierten Szenen ist Skalierungsinvarianz überflüssig; Rotationsinvarianz ist bei der Schrifterkennung zur Unterscheidung der Zeichen “6” und “9” nicht erwünscht.

3.4.3 Reguläre und zentrale Momente

Momente sind aus der Statistik für die Beurteilung von Verteilungen und deren Form, aber auch aus der Physik der starren Körper als Drehmomente bekannt. Das Drehmoment eines starren Massenkörpers der Dichte $\rho(x, y, z)$ an der Stelle $r = (x, y, z)$ ist durch das Raumintegral über $r^2 \rho(x, y, z)$ gegeben. Das Drehmoment ist invariant unter Rotation und Translation des Massenkörpers, wenn der Koordinatenursprung in den Schwerpunkt des Körpers gelegt wird. Wir übertragen diese Definition in die Bildverarbeitung und benutzen anstelle der Massendichte die Grauwerte $g(x, y)$ und bilden zweidimensionale Flächenintegrale über die Bildkoordinaten. Mit der Definition der Momente soll hier der Weg zu möglichst vollständigen Sätzen von Parametern beschriftet werden. Für die Mustererkennung definiert man die Momente als Integrale über das Produkt aus Grauwerten $g(x, y)$ an der Stelle (x, y) und den Monomen $x^p y^q$ der Koordinaten

$$\begin{aligned} m_{pq} &= \int_x \int_y x^p \cdot y^q \cdot g(x, y) dx dy && \text{(reguläre Momente)} \\ \mu_{pq} &= \int_x \int_y (x - x_s)^p \cdot (y - y_s)^q \cdot g(x, y) dx dy && \text{(zentrale Momente)} \end{aligned} \quad (3.8)$$

dabei erstrecken sich die Integrale bei Momenten von ganzen Bildern über alle vorkommenden Grauwerte $g(x, y)$. Die Koordinaten des Grauwertschwerpunktes (x_s, y_s) sind gegeben durch

$$x_s = \frac{\int_x \int_y x g(x, y) dx dy}{\int_x \int_y g(x, y) dx dy} = \frac{m_{10}}{m_{00}} \quad y_s = \frac{\int_x \int_y y g(x, y) dx dy}{\int_x \int_y g(x, y) dx dy} = \frac{m_{01}}{m_{00}} \quad \text{(Grauwertschwerpunkt)} \quad (3.9)$$

Wegen der Definition des Schwerpunktes ist stets $\mu_{01} = \mu_{10} = 0$. Torres-Méndez [Torres-Méndez, L. A. u. a., 2000] konnte aus dem Bild-Drehmoment (mit geeigneter Normierung ist auch Skaleninvarianz zu erreichen) und weiteren 15 ebenfalls sehr anschaulich konstruierten topologischen Parametern,

die invariant sind unter Rotation, Translation und Skalierung, eine robuste binäre Buchstabenerkennung konstruieren. Bei der Objektanalyse werden alle nicht zu dem ausgewählten (in der Regel geschlossenem) Objekt gehörenden Pixel unterdrückt, also zu Null gesetzt. Die Integration läuft dann nur über die zu dem Objekt gehörenden Grauwerte. Bei der reinen Erfassung der geometrischen Form werden die Grauwerte $g(x, y)$ binär zu 1 gesetzt. Für unsere Anwendungen werden die tatsächlichen Grauwerte $g(x, y)$ jedoch ausschlaggebend sein, so daß wir unter Objektanalyse im Folgenden stets Momente unter Berücksichtigung von $g(x, y)$ verstehen. Dieses Vorgehen ist analog zur Berechnung von Drehmomenten starrer Körper mit der Massendichte $\rho(x, y, z)$ an der Stelle der Grauwerte $g(x, y)$. In der Praxis der Bildverarbeitung wird die Integration durch eine Summation über alle Zeilen und Spalten eines Bildes ersetzt (siehe Abschnitt 3.4.12). Ein Übergang zu zylindersymmetrischen Koordinaten $dx dy \rightarrow r d\varphi dr$ wie für rotationssymmetrische Körper in der klassischen Mechanik bietet sich in der Bildverarbeitung wegen der bereits kartesisch vorliegenden Grauwert-Daten kaum an, da dies eine weitere Abtastung der Daten erfordert. Ausnahmen bildet sicher die Buchstabenerkennung, bei der aus den Eingangsdaten oft nicht die Grauwerte, sondern extrahierte Linienelemente (Bogensegmente, engl. *splines*) benutzt werden. Details zu dieser Koordinatentransformation finden sich bei der in [Foggia, Pasquale u. a., 1999] vorgestellten Anwendung. Reguläre Momente werden oft [Teh, Cho-Huak und Chin, Roland T., 1988] als *geometrische Momente* bezeichnet.

Zu einer Ordnung $N = p + q$ gibt es $N + 1$ Momente. Momente haben die Dimension $[x]^{n+2}$. Das Moment $m_{00} = \mu_{00}$ entspricht der Fläche (bei binären Objekten) bzw. der Grauwertsumme des Objektes (engl. *total image power*). Dies legt eine Normierung aller Momente für $N > 1$ mit m_{00} analog zu den Grauwertschwerpunktskoordinaten (3.9) nahe. Bei der Berechnung von Invarianten werden wir auf eine geeignetere Normierung stoßen. Die numerischen Werte der Momente fallen stark ab mit steigender Ordnung, was zu Vorschlägen [Boyce, J. F. und Hossak, W. J., 1983] für numerisch leichter zu behandelnde Definitionen führte oder durch einfaches Logarithmieren [Wong, Robert Y. und Hall, Ernest L., 1978] abgefangen werden kann. Eine obere Grenze der stets positiven regulären Momente für Bilder der Größe $X \times Y$ läßt sich abschätzen zu

$$0 < m_{pq} \leq \left(\frac{X^{p+1} - 1}{p + 1} \right) \left(\frac{Y^{q+1} - 1}{q + 1} \right) \max_{x,y} \{g(x, y)\}$$

Für ein anschauliches Verständnis der Momente zweiter Ordnung konstruiert Teague [Teague, 1980] [Teague, 1982] aus den zentralen Momenten bis zweiter Ordnung eines ganzen Bildes eine Bildellipse mit der großen und kleinen Halbachse a und b und dem Orientierungswinkel φ (zwischen x-Achse und großer Hauptachse). Die Parameter der Ellipse folgen aus den Rotationseigenschaften der zentralen Momente zweiter Ordnung bei Drehung des Koordinatensystems $x, y \rightarrow x', y'$.

$$\begin{bmatrix} \mu'_{20} \\ \mu'_{11} \\ \mu'_{02} \end{bmatrix} = \begin{bmatrix} \frac{1+\cos 2\varphi}{2} & -\sin 2\varphi & \frac{1-\cos 2\varphi}{2} \\ \frac{\sin 2\varphi}{2} & \cos 2\varphi & -\frac{\sin 2\varphi}{2} \\ \frac{1-\cos 2\varphi}{2} & \sin 2\varphi & \frac{1+\cos 2\varphi}{2} \end{bmatrix} \begin{bmatrix} \mu_{20} \\ \mu_{11} \\ \mu_{02} \end{bmatrix} \quad (3.10)$$

Der vollständige Ausdruck für die Transformation zentraler Momente unter Rotation um den Winkel

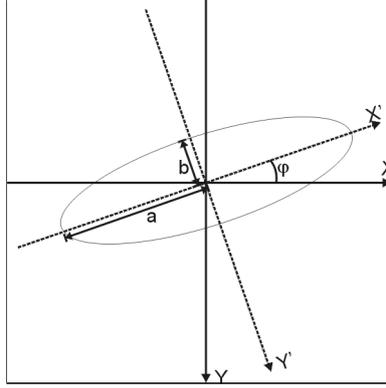


Abbildung 3.2: Repräsentation eines Bildes durch eine Ellipse aus den Momenten bis einschließlich zweiter Ordnung

φ lautet

$$\mu'_{pq} = \sum_{n=0}^q \sum_{m=0}^p (-1)^{q-n} \binom{q}{n} \binom{p}{m} (\sin \varphi)^{q+m-n} (\cos \varphi)^{p+n-m} \mu_{p+q-m-n, m+n} \quad (3.11)$$

wobei $\binom{q}{n} = \frac{q!}{n!(q-n)!}$ ist. Für eine gegebene Ordnung $N = p + q$ besteht das transformierte Moment μ'_{pq} ausschließlich aus Termen der Ordnung N . Für die Ellipsenparameter erhält man aus (3.10)

$$\begin{aligned} a &= \sqrt{\frac{\mu_{20} + \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}}{\mu_{00}/2}} && \text{(große Halbachse)} \\ b &= \sqrt{\frac{\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}}{\mu_{00}/2}} && \text{(kleine Halbachse)} \\ \varphi &= \frac{1}{2} \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) && \text{(Orientierungswinkel)} \\ G &= \frac{\mu_{00}}{\pi ab} && \text{(mittlerer Grauwert)} \end{aligned} \quad (3.12)$$

Bei der Invertierung der tan-Funktion wird der Orientierungswinkel so gewählt, daß $-\pi/2 \leq \varphi \leq \pi/2$. Für unbestimmte Werte entnehme man den Orientierungswinkel der Tabelle 3.1.

3.4.4 Invarianten aus Momenten

Für die Mustererkennung und insbesondere das *template matching* spielen Momente eine bedeutende Rolle, weil sich aus diesen Objektparametern Größen konstruieren lassen, die invariant sind unter Translation, Rotation, Achsenskalierung und gegebenenfalls auch unter Punktspiegelungen und Achsenspiegelungen. Somit kann sich die Suche nach Objekten und deren Lage innerhalb von Bildern auf den Abgleich weniger Parameter beschränken und muß nicht wie bei der mittels Kreuzkorrelationen

μ_{11}	$(\mu_{20} - \mu_{02})$	Intervall	φ
0	< 0		-90°
< 0	< 0	$] - 90^\circ \dots - 45^\circ [$	$1/2 \arctan(\xi) - 90^\circ$
< 0	0		-45°
< 0	> 0	$] - 45^\circ \dots - 0^\circ [$	$1/2 \arctan(\xi)$
0	> 0		0
0	0		0
> 0	> 0	$] 0^\circ \dots + 45^\circ [$	$1/2 \arctan(\xi)$
> 0	0		$+45^\circ$
> 0	< 0	$] + 45^\circ \dots + 90^\circ [$	$1/2 \arctan(\xi) + 90^\circ$

wobei $\xi = 2\mu_{11} / (\mu_{20} - \mu_{02})$

Tabelle 3.1: Ergebnisbereiche für den Orientierungswinkel und Werte an singulären Punkten der tan-Funktion

durchgeführten Suche eine große Zahl zweidimensionaler Integrationen über das *template* durchführen. Bei der Konstruktion eines Invariantensatzes für die praktische Anwendung sind folgende Kriterien zu beachten, die mit den Anforderungen an den Klassifikator (Abschnitt 3.5) eng verbunden sind:

- **Robustheit**
Wie robust sind die Invarianten gegen Rauschen und Rechenungenauigkeiten (diskrete Abtastung) bei den Transformationen?
- **Rechenkomplexität**
Läßt sich das Invarianten-Set durch sinnvolle Zerlegung schnell und mit wenig Speicheraufwand berechnen?
- **Diskriminanz**
Lassen sich mit den gewählten Invarianten die zu untersuchenden Grauwertobjekte klar trennen, gut beschreiben, sicher klassifizieren?
- **Generalität**
Lassen sich einmal entwickelte Invariantensätze und Randbedingungen ohne umfangreiche Neuanpassungen auf andere Probleme übertragen?
- **Flexibilität**
Lassen sich bestimmte Forderungen oder Randbedingungen auf einfache Weise ein/ausschalten (Wegfall bestimmter Transformationsinvarianzen, andere Rauschbedingungen, Bildauflösungen, ...)?

Durch die Verschiebung um (x_s, y_s) (3.9) auf ein dem Objekt angepaßtes Koordinatensystem sind die zentralen Momente (3.8) im Gegensatz zu den regulären Momenten translationsinvariant. Zur Konstruktion der Invarianz unter Achsenskalierung um den Faktor α sei $g'(x', y')$ das Objekt nach der

Achsenkalierung, $g(x, y)$ das ursprüngliche Objekt und die Transformation $g'(x', y') = g'(\alpha x, \alpha y) = g(x, y)$. Dann erhalten wir im neuen Koordinatensystem

$$\iint x'^p y'^q g'(x', y') dx' dy' = \iint (\alpha x)^p (\alpha y)^q g(x, y) \alpha dx \alpha dy = \alpha^{p+q+2} \iint x^p y^q g(x, y) dx dy \quad (3.13)$$

Zu regulären oder zentralen Momenten (3.8) tritt der Faktor α^{p+q+2} hinzu. Dies legt folgende Normierung nahe

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\frac{p+q+2}{2}}} \quad (\text{normierte zentrale Momente}) \quad (3.14)$$

Dieser Ausdruck ist im Original bei [Hu, 1962] fehlerhaft. Eine andere gebräuchliche [Belkasim, S. O. u. a., 1991] Normierung zur Erzeugung von Skaleninvarianzen ist $(\mu_{20} + \mu_{02})^\gamma$ mit $\gamma = (p+q+2)/4$. Dadurch wird der größte Radius $\sqrt{\mu_{20} + \mu_{02}}$ des Ellipsoides zu 1. Normierte zentrale Momente sind invariant unter Achsenkalierung um den Faktor α . Die Normierung bewirkt, daß alle η_{pq} einheitenlos sind und $\eta_{00} = 1$. Achsenkalierungen werden insbesondere benutzt, wenn Momente ganzer Bilder berechnet werden: eine Skalierung auf $x, y \in [0 \dots 1]$ erleichtert die numerische Verarbeitung der Momente erheblich. Normierungen von Momenten sind von der Normierung ganzer Bilder zu unterscheiden, auch wenn in (3.14) der Übergang von μ_{00} (Normierung eines Grauwertobjektes) zur Grauwertsumme des gesamten Bildes auf der Hand liegt. Normierungen von Bildern haben wir bereits in der Hauptachsentransformation in (3.10) und Abbildung 3.2 kennengelernt. Letztlich ist auch die Translation in der Definition der zentralen Momente eine Normierung, wenn Momente ganzer Bilder benutzt werden. Weitere mögliche Bildnormierungen sind die Kontrastnormierung [Maitra, 1979] (wenn sich der Kontrast eines Bildes oder eines Grauwertobjektes um den Faktor c verändert, so gilt $m_{pq} = cm_{pq}$) und Normierungen gegenüber Spiegelungen.

Wegen der fehlenden Translationsinvarianz haben die regulären Momente (3.8) in der Praxis keine Bedeutung, obschon sie sich mit Look-Up-Tabellen für die x^p und y^q schnell berechnen ließen. Interessant für den Einsatz bei der Objektsuche in Bildern sind erst die zentralen Momente, die sich

auch aus den regulären Momenten berechnen lassen. Für die ersten drei Ordnungen gilt

$$\begin{aligned}
\mu_{00} &= m_{00} & \mu_{01} &= 0 & \mu_{10} &= 0 \\
\mu_{02} &= m_{02} - \frac{m_{01}^2}{m_{00}} \\
\mu_{20} &= m_{20} - \frac{m_{10}^2}{m_{00}} \\
\mu_{11} &= m_{11} - \frac{m_{10}m_{01}}{m_{00}} \\
\mu_{30} &= m_{30} - 3\frac{m_{10}m_{20}}{m_{00}} + 2\frac{m_{10}^3}{m_{00}^2} \\
\mu_{03} &= m_{03} - 3\frac{m_{01}m_{02}}{m_{00}} + 2\frac{m_{01}^3}{m_{00}^2} \\
\mu_{21} &= m_{21} - 2\frac{m_{11}m_{10}}{m_{00}} - \frac{m_{20}m_{01}}{m_{00}} + 2m_{01}\left(\frac{m_{10}}{m_{00}}\right)^2 \\
\mu_{12} &= m_{12} - 2\frac{m_{11}m_{01}}{m_{00}} - \frac{m_{02}m_{10}}{m_{00}} + 2m_{10}\left(\frac{m_{01}}{m_{00}}\right)^2
\end{aligned} \tag{3.15}$$

Die vollständige Berechnungsvorschrift lautet

$$\mu_{pq} = \sum_{k=0}^p \sum_{l=0}^q \frac{p!}{k!(p-k)!} \frac{q!}{l!(q-l)!} (-1)^{k-l} m_{p-k,q-l} m_{10}^k m_{01}^l m_{00}^{l-k} \tag{3.16}$$

Die in (3.14) erreichte Definition der normierten zentralen Momente, die bereits invariant unter Verschiebung und Achsenskalierung sind, läßt sich für die nicht-trivialen Fälle $N > 1$ nicht mehr weiter ausbauen, um auch Rotationsinvarianz oder Invarianz unter Spiegelungen zu erhalten. Diese beiden Typen von Transformationen sind für das *template matching* jedoch unentbehrlich. Erst geeignete Linearkombinationen der η_{pq} werden diese Eigenschaft erfüllen. Eine aus der Analogie mit dem Drehmomententensor der klassischen Mechanik anschauliche Herleitung dieser Invarianten erlaubt der Momententensor.

3.4.5 Der Momententensor

Der Momententensor enthält drei unabhängige Größen, nämlich die Momente der zweiten Ordnung. Da zentrale Momente verwendet werden, ist die relevante Information der regulären Momente nullter und erster Ordnung bereits enthalten.

$$\mathbf{M} = \begin{bmatrix} \mu_{20} & -\mu_{11} \\ -\mu_{11} & \mu_{02} \end{bmatrix} \tag{3.17}$$

Daraus erhalten wir die Determinante und die Spur

$$\begin{aligned}\det \mathbf{M} &= \mu_{20}\mu_{02} - \mu_{11}^2 \\ \text{Spur } \mathbf{M} &= \mu_{20} + \mu_{02}\end{aligned}\quad (3.18)$$

Die Rotationsmatrix lautet

$$\mathbf{R} = \begin{bmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{bmatrix} \quad (3.19)$$

Bei der Hauptachsentransformation werden die Eigenwerte λ_i des Momententensors \mathbf{M}' die neuen, voneinander unabhängigen Größen. Im Momententensor \mathbf{M} waren drei Größen vorhanden, in \mathbf{M}' sind es nur noch zwei: die restliche Information ist im Drehwinkel φ enthalten. Die Hauptachsentransformation ist gegeben durch

$$\begin{aligned}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} &= \mathbf{R}^T \mathbf{M} \mathbf{R} \\ &= \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} \mu_{20} & -\mu_{11} \\ -\mu_{11} & \mu_{02} \end{bmatrix} \begin{bmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{bmatrix} \\ &= \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} \mu_{20} \cos \varphi + \mu_{11} \sin \varphi & \mu_{20} \sin \varphi - \mu_{11} \cos \varphi \\ -\mu_{02} \sin \varphi + -\mu_{11} \cos \varphi & \mu_{02} \cos \varphi - \mu_{11} \sin \varphi \end{bmatrix} \\ &= \begin{bmatrix} \mu_{20} \cos^2 \varphi + \mu_{02} \sin^2 \varphi + \mu_{11} \sin 2\varphi & \frac{1}{2} (\mu_{20} - \mu_{02}) \sin 2\varphi - \mu_{11} \cos 2\varphi \\ \frac{1}{2} (\mu_{20} - \mu_{02}) \sin 2\varphi - \mu_{11} \cos 2\varphi & \mu_{02} \cos^2 \varphi + \mu_{20} \sin^2 \varphi - \mu_{11} \sin 2\varphi \end{bmatrix}\end{aligned}$$

Wir entnehmen der Gleichung die Außerdiagonalelemente und formen um zu:

$$\begin{aligned}0 &= \frac{1}{2} (\mu_{20} - \mu_{02}) \sin 2\varphi - \mu_{11} \cos 2\varphi \\ \frac{1}{2} (\mu_{20} - \mu_{02}) \sin 2\varphi &= \mu_{11} \cos 2\varphi \\ \frac{\sin 2\varphi}{\cos 2\varphi} &= \frac{\mu_{11}}{\frac{1}{2} (\mu_{20} - \mu_{02})} \\ \tan 2\varphi &= \frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\end{aligned}$$

Daraus ergibt sich schließlich für die Eigenwerte und den Drehwinkel

$$\begin{aligned}\lambda_1 &= \mu_{20} \cos^2 \varphi + \mu_{02} \sin^2 \varphi + \mu_{11} \sin 2\varphi \\ \lambda_2 &= \mu_{02} \cos^2 \varphi + \mu_{20} \sin^2 \varphi - \mu_{11} \sin 2\varphi \\ \varphi &= \frac{1}{2} \arctan \left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right)\end{aligned}\quad (3.20)$$

Determinante und Spur sind Invarianten unter Rotation, was wir mit dem transformierten Momententensor \mathbf{M}' schnell nachrechnen können

$$\begin{aligned}
\det \mathbf{M}' &= \lambda_1 \lambda_2 \\
&= (\mu_{20}^2 + \mu_{02}^2) \sin^2 \varphi \cos^2 \varphi + \mu_{20} \mu_{02} (\cos^4 \varphi + \sin^4 \varphi) \\
&\quad + \mu_{02} \mu_{11} \sin 2\varphi (\cos^2 \varphi - \sin^2 \varphi - \cos 2\varphi) \\
&\quad + \mu_{20} \mu_{11} \sin 2\varphi (\sin^2 \varphi - \cos^2 \varphi + \cos 2\varphi) \\
&\quad - \frac{1}{4} \sin^2 2\varphi (\mu_{20}^2 + \mu_{02}^2) + \frac{1}{2} \mu_{02} \mu_{20} \sin^2 2\varphi \\
&\quad - \mu_{11}^2 \sin^2 2\varphi - \mu_{11}^2 \cos^2 2\varphi \\
&= \mu_{02} \mu_{20} - \mu_{11}^2
\end{aligned}$$

$$\text{Spur } \mathbf{M}' = \lambda_1 + \lambda_2 = \mu_{20} + \mu_{02}$$

Mit der Spur und der Determinante (3.18) des Momententensors (3.17) haben wir zusammen mit der Grauwertsumme μ_{00} drei Invarianten unter Rotation, Translation und Achsenskalierung gefunden. In der Analogie des Trägheitsellipsoides ist es anschaulicher, die Exzentrizität e einer Ellipse als Invariante zu nutzen. Diese setzt sich über die große und kleine Halbachsen aus (3.12) zusammen aus der Grauwertsumme μ_{00} , der Spur und der Determinante (3.18).

$$e = \sqrt{\frac{a^2 - b^2}{a^2}} = \sqrt{\frac{2\sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}}{\mu_{20} + \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}}} = \sqrt{\frac{2\sqrt{(\text{Spur } \mathbf{M})^2 - 4 \det \mathbf{M}}}{\text{Spur } \mathbf{M} + \sqrt{(\text{Spur } \mathbf{M})^2 - 4 \det \mathbf{M}}}} \quad (3.21)$$

In der Mustererkennung benutzt wird eher eine andere Exzentrizität ϵ des Ellipsoides der Momente zweiter Ordnung, die definiert ist über

$$\epsilon = \frac{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}{(\mu_{20} + \mu_{02})^2} = \frac{(\text{Spur } \mathbf{M})^2 - 4 \det \mathbf{M}}{(\text{Spur } \mathbf{M})^2} \quad (3.22)$$

Für beide Exzentrizitäten gilt, daß sie für den Fall des Kreises $a = b$ (was gleichbedeutend ist mit $\mu_{20} = \mu_{02}$ und $\mu_{11} = 0$) zu Null werden und für sehr langgestreckte Ellipsen gegen 1 gehen. Die Exzentrizität e nimmt den Wert $1/2$ an, wenn $a = \frac{2b}{\sqrt{3}} \approx 1,15 b$; für die Exzentrizität ϵ ist das der Fall, wenn $a = \sqrt{\frac{\sqrt{2}+1}{\sqrt{2}-1}} b \approx 2,41 b$.

3.4.6 Die Invarianten nach Hu

Eine systematische Ableitung der Invarianten gelang Ming-Kuei Hu [Hu, 1961] [Hu, 1962] aus der Theorie der algebraischen Invarianten aus binären Formen (in diesem Fall sind damit homogene Polynome gemeint und nicht binärwertige Pixelregionen) wie sie von Cayley eingeführt wurden. Zu dem von Hu benutzten fundamentalen Theorem über die Invarianzen bei Momenten hat Reiss [Reiss, 1991] [Reiss, 1993] einige relevante Ergänzungen hinzugefügt. Der Beweis führt für den Fall der Rotationsinvarianz im wesentlichen zur Transformationsgleichung (3.11), womit die momentenbasierte Objektanalyse in der Mustererkennung begründet wurde. Teague [Teague, 1980] leitet diese über die von ihm vorgeschlagenen Zernike-Momente [Khotanzad, Alireza und Hong, Yaw Hua, 1990] her, die sich auf die aus der theoretischen Optik bekannten Zernike-Polynome [Zernike, 1934] stützen und zu einer

der relativen Rotationsinvarianz bei den komplexen Momenten (3.24) analogen Gleichung für die Ableitung invarianter Größen führen. Die Mathematik der momentenbasierten Invarianten geht zurück auf die Theorie der algebraischen Invarianten (unter affinen und projektiven Transformationen), die beginnend mit Boole 1841 formuliert und von Cayley [Cayley, 1856] und Sylvester [Sylvester, 1878] ausgebaut wurde. Nach den abschließenden Beiträgen von Hilbert [Hilbert, 1890] [Hilbert, 1893] rückten die algebraischen Invarianten aus dem Fokus. Historische Überblicke zu deren Erforschung geben die Beiträge von Weiss [Weiss, 1993], Dickson [Dickson, 1914], Abhyankar [Abhyankar, 1991] oder allgemeiner von Parshall [Parshall, 1989]. Wegen der heute schwer zugänglichen Originalliteratur sei auf moderne Zusammenfassungen der Mathematik der algebraischen Invarianten in [Reiss, 1993] und [Dickson, 1914] verwiesen. Verweise auf gruppentheoretische Grundlagen und alternative Sätze von Invarianten (basierend auf der Fouriertransformierten oder mittels neuronaler Netze) finden sich im Übersichtsartikel von Wood [Wood, 1996]. Die umfangreichste Liste von zweidimensionalen und dreidimensionalen momentenbasierten Invarianten wird von Salmon [Salmon, 1885] gegeben und ist in [Reiss, 1991] [Reiss, 1993] vervollständigt. Reiss [Reiss, 1993] weist auch auf den besonders einfachen Beweis über die komplexen Momente (3.23) hin, die von Davies [Davies, 1977] und Abu-Mostafa [Abu-Mostafa, Yaser S. und Psaltis, Demitri, 1984] [Abu-Mostafa, Yaser S. und Psaltis, Demitri, 1985] eingeführt wurden.

$$c_{pq} = \iint (x + iy)^p \cdot (x - iy)^q \cdot g(x, y) \, dx dy \quad (\text{komplexe Momente}) \quad (3.23)$$

Mit Polarkoordinaten $x + iy = r e^{i\theta}$ und $\sqrt{-1} = i$ wird aus den komplexen Momenten

$$\begin{aligned} c_{pq} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x + iy)^p (x - iy)^q g(x, y) \, dx dy \\ &= \int_0^{2\pi} \int_0^{\infty} r \, r^p \cdot r^q \cdot e^{i(p-q)\theta} \cdot g(r, \theta) \, d\theta \, dr \\ &= \int_0^{\infty} r^{p+q+1} \left[\int_0^{2\pi} e^{i(p-q)\theta} g(r, \theta) \, d\theta \right] \, dr \end{aligned}$$

Wir sehen sofort, daß eine Rotation (3.19) um den Winkel φ gegeben durch $(x', y')^T = \mathbf{R} \cdot (x, y)^T$ für die komplexen Momente bedeutet:

$$c'_{pq} = c_{pq} e^{-i(p-q)\varphi} \quad (3.24)$$

In der Sprache der algebraischen Invarianten bezeichnet man diese Eigenschaft als *relative* Invarianz. Der Betrag und natürlich auch das Betragsquadrat dieses Ausdrucks sind unabhängig vom Rotationswinkel φ : die $|c_{pq}|$ werden also invariant unter Rotation, was als *absolute* Invarianz bezeichnet wird.

$$\begin{aligned} |c'_{pq}| &= |c_{pq}| \\ c'_{pq} c'_{qp} &= c_{pq} c_{qp} \end{aligned} \quad (3.25)$$

Weiterhin gilt aus der Definition (3.23) der komplexen Momente, daß wegen $c_{pq} = c_{qp}^*$ (wobei der $*$ das komplex Konjugierte bezeichnet) sich komplexe Momente einer Ordnung stets und ausschließlich

aus den zentralen Momenten der selben Ordnung zusammensetzen. Über die Bildung des Betrages lassen sich also zu einer Ordnung $N = p + q$ und den dazu gehörenden $N + 1$ komplexen Momenten insgesamt $1 + N/2$ unabhängige Invarianten konstruieren. Im einzelnen:

$$\begin{aligned}
c_{00} &= m_{00} \\
c_{10} &= m_{10} + im_{01} = c_{01}^* \\
c_{11} &= m_{20} + m_{02} \\
c_{20} &= m_{20} - m_{02} + 2im_{11} = c_{02}^* \\
c_{21} &= m_{30} - m_{12} + i(m_{21} + m_{03}) = c_{12}^* \\
c_{30} &= m_{30} - 3m_{21} + i(3m_{21} + m_{03}) = c_{03}^*
\end{aligned}$$

Abu-Mostafa [Abu-Mostafa, Yaser S. und Psaltis, Demitri, 1984] zeigt, daß $(N + 1)(N + 2)/2 - 1$ Invarianten aus den $(N + 1)(N + 2)/2$ komplexen Momenten gewonnen werden können, wenn man Invarianten vom Typ

$$c_{rs}c_{tu}^k + c_{sr}c_{ut}^k \quad \text{mit} \quad (s - r) = k(t - u)$$

hinzunimmt, bei denen der Phasenwinkel nicht allein durch die Betragsbildung verschwindet. Wir ersetzen nun in (3.23) die regulären Momente durch die translationsinvarianten zentralen Momente, so daß die Betragsquadrate $|c_{pq}|^2$ sowohl translationsinvariant als auch rotationsinvariant werden. In den Betragsquadraten dieser komplexen Momente erkennen wir sofort die Spur des Momententensors und Zähler und Nenner der Exzentrizität (3.22) wieder. In der Nomenklatur von Hu lauten die Invarianten [Hu, 1962] bis zur vierten Ordnung (weitere bis zur Ordnung $N = 9$ werden genannt bei [Belkasim, S. O. u. a., 1991] [Belkasim, S. O. u. a., 1993] und [Li, 1992]) dann explizit:

$$\begin{aligned}
\psi_0 &= c_{00} = \mu_{00} \\
\psi_1 &= c_{11} = \mu_{20} + \mu_{02} \\
\psi_2 &= |c_{20}|^2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \\
\psi_3 &= |c_{30}|^2 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} + \mu_{03})^2 \\
\psi_4 &= |c_{21}|^2 = (\mu_{30} - \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \\
\psi_5 &= (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12}) \left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2 \right] + \\
&\quad + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03}) \left[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2 \right] \\
\psi_6 &= (\mu_{20} - \mu_{02}) \left[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2 \right] + \\
&\quad + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \\
\psi_7 &= (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12}) \left[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2 \right] + \\
&\quad + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03}) \left[3(\mu_{30} - \mu_{12})^2 - (\mu_{21} + \mu_{03})^2 \right] \\
\psi_8 &= (\mu_{40} - 6\mu_{22} + \mu_{04})^2 + 16(\mu_{13} - \mu_{31})^2 \\
\psi_9 &= (\mu_{40} - \mu_{04})^2 - 2(\mu_{31} + \mu_{13})^2 \\
\psi_{10} &= \mu_{40} - 2\mu_{22} + \mu_{04}
\end{aligned} \tag{3.26}$$

Eine alternative Herleitung der ersten zwölf Rotationsinvarianten (bis $N = 4$) und eine allgemeine Ableitung des Drehwinkels (3.12) (3.20) findet sich bei Rothe [Rothe, Irene und Voss, Klaus, 1992].

Zur allgemeinen Berechnung des Rotationswinkels aus den komplexen Momenten siehe [Rothe, Irene und Voss, Klaus, 1992].

$$\begin{aligned}
\psi_1 &= \text{Spur } \mathbf{M} \\
\psi_2 &= (\text{Spur } \mathbf{M})^2 - 4 \det \mathbf{M} \\
\epsilon &= \frac{\psi_2}{\psi_1^2} \\
e &= \sqrt{\frac{2\sqrt{\psi_2}}{\psi_1 + \sqrt{\psi_2}}} \\
\varphi &= \frac{1}{2} \arctan \frac{\text{Im}\{c_{20}\}}{\text{Re}\{c_{20}\}}
\end{aligned} \tag{3.27}$$

3.4.7 Vollständigkeit von Momenten und charakteristische Funktion

Für die praktische Anwendung von zentralen Momenten (3.8) und den daraus konstruierbaren Invarianten bei der Wiedererkennung von (verschobenen, rotierten, skalierten) Objekten in einem *template matching* Problem ist die wichtigste Frage: wie viele Momente werden zu einer sauberen Trennung von Grauwertobjekten benötigt? Welche Ordnung N von Momenten reicht aus, um Objekte eindeutig zu beschreiben? Antworten auf diese Frage geben einerseits das Vollständigkeitstheorem, andererseits Untersuchungen zur Robustheit der Momente gegenüber Rauschen (Abschnitt 3.4.11). Diese leiten auch zu informations-theoretischen Fragen über, welche Strukturinformation eines Objektes sich in welchen Ordnungen wiederfindet. Weiterführende Betrachtungen dazu finden sich in [Abu-Mostafa, Yaser S. und Psaltis, Demitri, 1984]. Um ein Grauwertobjekt $g(x, y)$ durch seine regulären Momente darzustellen, schreiben wir die Fouriertransformierte $\hat{g}(x, y)$ des Grauwertobjektes mit der Reihenentwicklung der Exponentialfunktion um und transformieren zurück in den Ortsraum. Wir erhalten dann das Vollständigkeitstheorem.

$$\begin{aligned}
\hat{g}(k_1, k_2) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) e^{-i(k_1x+k_2y)} dx dy \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) \left[\sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(-ik_1)^p}{p!} x^p \frac{(-ik_2)^q}{q!} y^q \right] dx dy \\
&= \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(-i)^{p+q}}{p!q!} k_1^p k_2^q \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) x^p y^q dx dy \\
&= \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(-i)^{p+q}}{p!q!} k_1^p k_2^q m_{pq}
\end{aligned}$$

$$g(x, y) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{i(k_1x+k_2y)} \left[\sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(-i)^{p+q}}{p!q!} k_1^p k_2^q m_{pq} \right] dk_1 dk_2 \tag{3.28}$$

Die wichtigste Aussage des Vollständigkeitstheorems (3.28) ist, daß wegen der beiden Terme k_1^p und k_2^q die Summation über die Ordnungen nicht mit der Integration vertauscht werden können. Dies ist

für die numerische Praxis ein entscheidender Nachteil, da es so nicht möglich ist, eine Genauigkeitsgrenze für die Berechnung des Ausdrucks anzugeben und die Summation abzubrechen, wenn höhere Summanden keine relevanten Änderungen mehr zum Ergebnis beitragen. Nur für wenige geometrische Sonderfälle wie das binäre Rechteck oder den binären Einheitskreis (alle Momente $N \geq 2$ sind Null) läßt sich (3.28) geschlossen schreiben und berechnen.

Die inverse Darstellung der Grauwertobjekte ähnelt bis auf die komplexwertigen Anteile der *charakteristischen Funktion* M der zentralen Momente. Die Momente lassen sich durch Ableitungen aus der Charakteristischen Funktion darstellen.

$$M(k_1, k_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) e^{k_1(x-x_s)+k_2(y-y_s)} dx dy \quad (\text{Charakteristische Funktion}) \quad (3.29)$$

$$\mu_{pq} = \left. \frac{\partial^{(p+q)} M(k_1, k_2)}{\partial k_1^p \partial k_2^q} \right|_{k_1=k_2=0}$$

Mit der Taylorentwicklung der Fouriertransformierten erhalten wir durch Vergleich mit der Schreibweise der Fouriertransformierten mit Momenten folgende Berechnungsvorschrift der Momente mit der Fouriertransformierten

$$\hat{g}(k_1, k_2) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(-i)^{p+q}}{p! q!} k_1^p k_2^q m_{pq}$$

$$\hat{g}(k_1, k_2) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{k_1^p k_2^q}{p! q!} \frac{d^{p+q}}{dk_1^p dk_2^q} \hat{g}(0, 0)$$

$$(-i)^{p+q} m_{pq} = \frac{d^{p+q}}{dk_1^p dk_2^q} \hat{g}(0, 0)$$

Bevor wir in Abschnitt 3.4.9 eine für die Rekonstruktion von $g(x, y)$ erheblich geeignetere Definition der Momente kennenlernen werden, sei an dieser Stelle mit regulären Momenten bis zu einer festen Ordnung N ein Schema für die Rekonstruktion angegeben.

3.4.8 Inverses Momentenproblem

Hinter den praktischen Fragen zur Vollständigkeit einer Objektbeschreibung mit nur endlich vielen Momenten steht letztlich das inverse Momentenproblem, bei dem aus einer Reihe von Momenten bis zu einer bestimmten Ordnung N das ursprüngliche Objekt $g(x, y)$ rekonstruiert werden soll. Bei der praktischen Ausformulierung des inversen Momentenproblems, wie es erstmals von Teague [Teague, 1980] behandelt wurde, werden wir auf die Vorteile einer Legendre-Formulierung (Abschnitt 3.4.9) der Momente stoßen.

Das Grauwertobjekt $g(x, y)$, das allgemein nur durch unendlich viele reguläre Momente m_{pq} vollständig beschrieben werden kann, soll durch das Objekt $f(x, y)$ angenähert werden, wobei $f(x, y)$ aus endlich vielen Momenten m_{pq} von $g(x, y)$ bis zu einer Ordnung N zusammengesetzt ist. Formal schreiben wir $f(x, y)$ als Polynominal, dessen Summe nur über Kombinationen mit $m+n \leq N$ reicht,

was gerade die Näherung ausmacht. Das Rekonstruktionsobjekt $f(x, y)$ soll nur höchstens von der Ordnung sein, für die zu $g(x, y)$ auch Momente m_{pq} in Betracht gezogen werden.

$$f(x, y) = \sum_n \sum_m f_{n,m} x^n y^m$$

Diese formale Schreibweise macht nur Sinn, wenn die Koeffizienten $f_{n,m}$ bestimmt werden. Diese sollen durch die Momente m_{pq} bis einschließlich der Ordnung N ausgedrückt werden. Für den Fall $N=3$ bedeutet dies:

$$f(x, y) = f_{0,0} + f_{1,0}x + f_{0,1}y + f_{1,1}xy + f_{2,0}x^2 + f_{0,2}y^2 + f_{3,0}x^3 + f_{2,1}x^2y + f_{1,2}xy^2 + f_{0,3}y^3$$

Die Momente m_{pq} zu $f(x, y)$ lauten

$$m_{pq} = \iint f(x, y) x^p y^q dx dy$$

wobei das Objekt umskaliert wurde, so daß $x, y \in [-1, +1]$.

Für $N=3$ erhalten wir $(1 + N^2/2 + 3N/2)^2 = 10$ Gleichungen:

$$\begin{aligned} m_{00} &= 4f_{00} + \frac{4}{3}f_{20} + \frac{4}{3}f_{02} \\ m_{10} &= \frac{4}{3}f_{10} + \frac{4}{5}f_{30} + \frac{4}{9}f_{12} \\ m_{01} &= \frac{4}{3}f_{01} + \frac{4}{5}f_{03} + \frac{4}{9}f_{21} \\ m_{11} &= \frac{4}{9}f_{11} \\ m_{20} &= \frac{4}{3}f_{00} + \frac{4}{9}f_{02} + \frac{4}{5}f_{20} \\ m_{02} &= \frac{4}{3}f_{00} + \frac{4}{5}f_{02} + \frac{4}{9}f_{20} \\ m_{30} &= \frac{4}{5}f_{10} + \frac{4}{7}f_{30} + \frac{4}{15}f_{12} \\ m_{03} &= \frac{4}{5}f_{01} + \frac{4}{7}f_{03} + \frac{4}{15}f_{21} \\ m_{21} &= \frac{4}{5}f_{01} + \frac{4}{7}f_{03} + \frac{4}{15}f_{21} \\ m_{12} &= \frac{4}{9}f_{10} + \frac{4}{15}f_{30} + \frac{4}{15}f_{12} \end{aligned}$$

Diese Abbildung $\vec{m} = \mathbf{F}^{-1}\vec{f}$ muß invertiert werden, um aus den bekannten $m_{pq} = m_{pq}$ die Koeffizienten $f_{n,m}$ für die Rekonstruktion von $f(x, y)$ zu bestimmen. Für $N=3$ lautet die invertierte Abbildung

$$\begin{bmatrix} f_{0,0} \\ f_{1,0} \\ f_{0,1} \\ f_{1,1} \\ f_{2,0} \\ f_{0,2} \\ f_{3,0} \\ f_{0,3} \\ f_{2,1} \\ f_{1,2} \end{bmatrix} = \frac{1}{16} \begin{bmatrix} 14 & 0 & 0 & 0 & -15 & -15 & 0 & 0 & 0 & 0 \\ 0 & 90 & 0 & 0 & 0 & 0 & -105 & 0 & 0 & -45 \\ 0 & 0 & 90 & 0 & 0 & 0 & 0 & -105 & -45 & 0 \\ 0 & 0 & 0 & 36 & 0 & 0 & 0 & 0 & 0 & 0 \\ -15 & 0 & 0 & 0 & 45 & 0 & 0 & 0 & 0 & 0 \\ -15 & 0 & 0 & 0 & 0 & 45 & 0 & 0 & 0 & 0 \\ 0 & -105 & 0 & 0 & 0 & 0 & 175 & 0 & 0 & 0 \\ 0 & 0 & -105 & 0 & 0 & 0 & 0 & 175 & 0 & 0 \\ 0 & 0 & -45 & 0 & 0 & 0 & 0 & 0 & 135 & 0 \\ 0 & -45 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 135 \end{bmatrix} \cdot \begin{bmatrix} m_{00} \\ m_{10} \\ m_{01} \\ m_{11} \\ m_{20} \\ m_{02} \\ m_{30} \\ m_{03} \\ m_{21} \\ m_{12} \end{bmatrix}$$

Es stellt sich sofort die Frage, bis zu welcher Ordnung N Momente m_{pq} berücksichtigt werden müssen, um eine bestimmte Qualität der Näherung von $f(x, y)$ an $g(x, y)$ zu erhalten. Dazu wäre anhand

des Beispiels von $N = 3$ das beschriebene Rechenschema auf $N = 2$ und $N = 4$ zu erweitern. Für $N = 2$ reduziert sich die Darstellung von $f(x, y)$ durch $f_{n,m}$ auf die ersten sechs Terme, so daß die Matrix \mathbf{F}^{-1} leicht aus dem Fall $N = 3$ abgeleitet werden kann. Wegen der Außerdiagonalelemente gilt dies leider nicht mehr für die invertierte Abbildung. Für $N = 2$ erhalten wir:

$$\mathbf{F}^{-1} = 4 \begin{bmatrix} 1 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{9} & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{5} & \frac{1}{9} \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{9} & \frac{1}{5} \end{bmatrix} \quad \mathbf{F} = \frac{1}{16} \begin{bmatrix} 14 & 0 & 0 & 0 & -15 & -15 \\ 0 & 12 & 0 & 0 & 0 & 0 \\ 0 & 0 & 12 & 0 & 0 & 0 \\ 0 & 0 & 0 & 36 & 0 & 0 \\ -15 & 0 & 0 & 0 & -45 & 0 \\ -15 & 0 & 0 & 0 & 0 & -45 \end{bmatrix}$$

Die Abbildungsmatrix \mathbf{F} muß für jede Ordnung N neu berechnet werden. Für die Ordnung $N = 4$ und ein vereinfachtes Rechenschema siehe Anhang A. Die Notwendigkeit, \mathbf{F}^{-1} für jede Ordnung neu zu invertieren, liegt an der Koppelung der Momente. Sinnvolle Näherungen und Lösungsverfahren nennt Talenti [Talenti, 1987]. Gelänge eine Definition der Momente m_{pq} , so daß die Gleichungen "entmischt" sind, so wäre das Verfahren gerade für iterative Näherungen erheblich brauchbarer. Legendre-Momente besitzen eine solche Orthogonalität, die unerwünschte Redundanz geht dadurch verloren.

3.4.9 Legendre-Momente

Das Grauwertobjekt $g(x, y)$ sei dargestellt durch die orthogonalen Legendre-Polynome P_n und die Koeffizienten λ_{pq}

$$g(x, y) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \lambda_{pq} P_p(x) P_q(y) \quad \text{mit} \quad P_n(x) = \frac{1}{n!2^n} \frac{d^n}{dx^n} (x^2 - 1)^n$$

Aus $P_0(x) = 1$ und $P_1(x) = x$ lassen sich die Legendre-Polynome auch rekursiv berechnen

$$P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{n}{n+1} P_{n-1}(x)$$

Numerische Näherungen für hohe Legendre-Polynome nennt [Liao, Simon X. und Pawlak, Mirosław, 1996]. Mit der Orthogonalitätseigenschaft [Sansone, 1991] der Legendre-Polynome

$$\int_{-1}^{+1} P_n(x) \cdot P_m(x) dx = \frac{2}{2n+1} \delta_{m,n}$$

erhalten wir für die Koeffizienten λ_{pq}

$$\lambda_{pq} = \frac{(2p+1)(2q+1)}{4} \int_{-1}^{+1} \int_{-1}^{+1} g(x, y) \cdot P_p(x) \cdot P_q(y) dx dy \quad (\text{Legendre-Momente}) \quad (3.30)$$

Die Berechnung erfolgt über die Multiplikation der Legendre-Darstellung von $g(x, y)$ mit $P_p(x) \cdot P_q(y)$ und anschließender Integration. Der Monom-Kernel $x^p y^q$ der regulären Momente würde also durch den Legendre-Kernel $P_p(x) \cdot P_q(y)$ ersetzt. Die entstehenden Summanden haben den gleichen Grad wie beim Monom-Kernel. Mit der Summenschreibweise der Legendrepolynome P_m und den tabellierten Koeffizienten c_{mj}

$$P_m(x) = \sum_{j=0}^m c_{mj} x^j$$

lassen sich die Legendre-Momente durch die regulären Momente (3.8) darstellen

$$\lambda_{pq} = \frac{(2p+1)(2q+1)}{4} \sum_{j=0}^p \sum_{k=0}^q c_{pj} c_{qk} m_{jk}$$

Zum Beispiel lauten diese für die ersten Ordnungen

$$\lambda_{00} = \frac{1}{4}m_{00} \quad \lambda_{10} = \frac{3}{4}m_{10} \quad \lambda_{20} = \frac{9}{8}m_{20} - \frac{3}{8}m_{00} \quad \lambda_{11} = \frac{9}{4}m_{11}$$

Andere Legendre-Momente ergeben sich durch die Vertauschung der Indices. Wir können die Entwicklung von $g(x, y)$ aus Abschnitt 3.4.8 in Legendre-Polynome nun bei einer Ordnung $p+q = N$ abbrechen und erhalten eine Näherung für $g(x, y)$

$$g(x, y) \approx f(x, y) = \sum_{p=0}^N \sum_{q=0}^{N-p} \lambda_{pq} P_p(x) \cdot P_q(y) = \sum_{p=0}^N \sum_{q=0}^p \lambda_{p-q,q} \cdot P_{p-q}(x) \cdot P_q(y) \quad (3.31)$$

Dies ist genau dasselbe wie die zuvor benutzte Summe mit $n+m \leq N$

$$f(x, y) = \sum_n \sum_m f_{n,m} x^n y^m$$

nur daß nun die Terme umgruppiert wurden. Der Vorteil dieser Schreibweise ist, daß die Legendre-Momente λ_{pq} entkoppelt sind und wir die bereits berechneten λ_{pq} nicht neu berechnen müssen, wenn wir zu höheren Ordnungen $N+1, N+2, \dots$ übergehen. Die Korrelation der zentralen/regulären Momente im Gegensatz zu den Legendre-Momenten zeigt gerade bei Klassifikationsaufgaben, wo im Parameterraum unabhängige Größen bevorzugt werden, den Vorteil der Legendre-Formulierung.

Den Rekonstruktionsfehler $e_0(N)$ zur Ordnung N (der Index 0 bezeichnet den Fall ohne Rauschen, siehe auch Abschnitt 3.4.11), den wir mit der Näherung $f(x, y)$ an das Objekt $g(x, y)$ ist gegeben durch

$$e_0^2(N) = \iint (f(x, y) - g(x, y))^2 dx dy$$

den man sinnvollerweise mit $\iint (g(x, y))^2 dx dy$ normiert. Für die Legendre-Momente (3.30) ergibt sich damit aus (3.31)

$$e_0^2(N) = \iint_{-1}^{+1} (f(x, y) - g(x, y))^2 dx dy - \sum_{p=0}^N \sum_{q=0}^p \frac{4\lambda_{pq}^2}{(2p - 2q + 1)(2q + 1)} \quad (3.32)$$

Durch Einsetzen der Darstellung der Legendre-Momente durch reguläre Momente erhalten wir den Rekonstruktionsfehler der regulären Momente. Höhere Fehlerterme durch Diskretisierung (Abschnitt 3.4.12) nennt [Liao, Simon X. und Pawlak, Mirosław, 1996].

3.4.10 Weitere Invariantensätze

Neben den unter bestimmten Transformationen bereits invarianten Momenten oder Invariantensätzen, die bisher vorgestellt worden sind, finden sich in der Literatur unzählige weitere Definitionen. Häufig genannt werden Zernike-Momente [Teague, 1980] und verwandte Definitionen, die eine Erweiterung der Legendre-Momente mit radialen Polynomen darstellen. Einen Sonderweg - wenn auch auf der Hand liegend - bei der Suche nach Invarianten gingen Flusser und Suk in ihren Arbeiten [Flusser, Jan und Suk, Thomas, 1993] [Flusser, 1998] [Flusser und Suk, 1999] und berechneten die sich aus allen affinen Transformationen

$$\begin{array}{lll} (I) & \begin{array}{l} x' = x + x_o \\ y' = y \end{array} & (II) & \begin{array}{l} x' = x \\ y' = y + y_o \end{array} & (III) & \begin{array}{l} x' = \alpha x \\ y' = \alpha y \end{array} \\ (IV) & \begin{array}{l} x' = \alpha_0 x \\ y' = y \end{array} & (V) & \begin{array}{l} x' = x + \beta_1 y \\ y' = y \end{array} & (VI) & \begin{array}{l} x' = x \\ y' = y + \beta_2 x \end{array} \end{array}$$

ergebenden Randbedingungen für das allgemeine Polynom

$$\sum_i k_i \mu_{p_1(i), q_1(i)} \cdots \mu_{p_N(i), q_N(i)}$$

je nach einbezogener Ordnung N entstehen dadurch umfangreiche Invarianten, die (3.26) ähneln. In der aktuellen Forschung wird versucht, bekannte Invariantensätze sinnvoll mit eventuell vorhandenem Vorwissen über die Objekte zu fusionieren und daraus robustere Mustererkenner zu konstruieren. Anpassungen der Invariantensätze an typische Effekte in Bilddaten wie Beleuchtungsänderungen, Kamerabewegungen, Verdeckungen oder perspektivische Transformationen aber auch die Verwendung von Farbinformationen wurden von Alferez [Alferez, Ronald-Bryan O. und Wang, Yuan-Fang, 1999] [Wang, Yuan-Fang und Alferez, Ronald-Bryan O., 1998] vorgeschlagen, womit die algebraisch motivierte Konstruktion von Invarianten sich deutlich auf die Anwendungspraxis der Bildverarbeitung zubewegt. Ein nicht auf Momente basierender Ansatz sind Fourierdeskriptoren, die in [Trier, Oivind Due. u. a., 1996] mit den hier bereits vorgestellten momenten-basierten Parametern anhand der Buchstabenerkennung verglichen werden. Im englischen Sprachgebrauch wird zwischen Formerkennung (*shape detection*) im Vergleich [Loncaric, 1998] zu Objekterkennung (*object detection*) streng unterschieden. Einen Vergleich verschiedener mathematischer Methoden zur Konstruktion von Invarianten

stellt Weiss [Weiss, 1993] an; Rothe [Rothe, Irene und Süsse, Herbert, 1993] (ausführlicher in [Rothe, Irene u. a., 1994]) stellt mehrere Konzepte in einen verallgemeinerten Zusammenhang.

Lokale Invarianten [Rivlin, Ehud und Weiss, Isaac, 1995] beschreiben im Unterschied zu den bisher hier kennengelernten Parametern nicht das Objekt und dessen Grauwertinhalt global als Ganzes, sondern sie beschreiben die Form (engl. *shape*) eines Objektes, seine äußere Begrenzung (im zweidimensionalen Raum also die Randkurve) mit lokalen Parametern - zum Beispiel einem Plot der lokalen Krümmung über die Bogenlänge der Randkurve (man spricht von einer Signatur). Solche lokalen Invarianten auf Abschnitten von Konturen enthalten stets Ableitungen und nur Informationen über die Form und nicht das Objekt und seinen Inhalt, sie sind aber robust gegenüber partiellen Verdeckungen (im Plot fehlen dann Abschnitte, dennoch ist ein Vergleich solcher Signaturen noch möglich). Momente auf Konturstücken behandelt [Lambert, Georg u. a., 1995]. Zur allgemeinen Konstruktion von Invarianten auf Kurven siehe [Forsyth, David u. a., 1992].

Einen auf Monomen als Kernel-Funktion ähnlich wie (3.8) beruhenden Ansatz zur Konstruktion von Invarianten beschrieben Burkhardt [Burkhardt, 1979] und Schulz-Mirbach [Schulz-Mirbach, 1995] mit interessanten Anwendungen [Siggelkow, Sven u. a., 2001]. Hierbei wird über einfache Monome aus einzelnen Grauwerten $g(x, y)$ eines Bildes über die Transformationsgruppe aufintegriert, um Invarianten bezüglich dieser Transformationen zu erhalten. Für eine Transformation $\mathbf{A}_{\varphi, t_x, t_y}$ aus Rotation und Translation

$$g(x', y') = \mathbf{A} \{g(x, y)\}$$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}$$

bilden die möglichen Werte von $\varphi = 0, \dots, 2\pi$ und $t_{x,y} = 0, \dots, X, Y$ (das Bild soll außerhalb zyklisch fortgesetzt werden, daher kann man die Translation auf eine Bildgröße beschränken) eine Transformationsgruppe \mathfrak{A} . Es soll eine Invariante $I(g)$ auf dem Bild $g(x, y)$ konstruiert werden. Gefordert wird also $I(g) = I(\mathbf{A}\{g\})$ für alle Transformationen der Gruppe \mathfrak{A} . Dafür integriert man eine Funktion auf f auf dem Bild über die Transformationsgruppe, so daß die Abhängigkeit von deren Parametern φ und $t_{x,y}$ verschwindet.

$$I(g) = \frac{1}{|\mathfrak{A}|} \int_{\mathfrak{A}} f(\mathbf{A}\{g\}) d\mathbf{A}$$

$$= \frac{1}{2\pi XY} \int_{t_x=0}^X \int_{t_y=0}^Y \int_{\varphi=0}^{2\pi} f(\mathbf{A}\{g\}) d\varphi dt_x dt_y$$

Die Normierung $|\mathfrak{A}|$ erhält man aus der Annahme, daß die Funktion f bereits eine Invariante ist $f(g) = f(\mathbf{A}\{g\})$ und in der Integration zu einem Faktor wird. Benutzt man beispielsweise die Funktion $f(g) = g(0, 0)$, so ist die Invariante einfach der Mittelwert des Bildes

$$I(g) = \frac{2\pi}{2\pi XY} \int_{t_x=0}^X \int_{t_y=0}^Y g(t_x, t_y) dt_x dt_y$$

Die genannten Autoren benutzen auf diskreten Bildern $g[m, n]$ für f sehr einfache Monome der Art $g[0, 0]$ $g[0, 1]$ $g[0, 3]$ oder $g[0, 0]^2 g[0, 2]$. Die daraus entstehenden Invarianten sind robust gegenüber partiellen Verdeckungen und meistens auch die Erkennung artikulierter Objekte. Durch die Additivität $I(g_1 + g_2) = I(g_1) + I(g_2)$ der Invarianten sind vorausgehende Separationen von Objekten mit den bekannten Schwierigkeiten bei teilweisen Bedeckungen nicht nötig.

3.4.11 Momente und Rauschen

Die Trennschärfe zwischen Objekten und deren Momenten und Invarianten für Klassifikationsaufgaben hängt nicht nur vom theoretischen Auflösungsvermögen der Bilddaten und der benutzten Momente ab, sondern auch vom Einfluß des Bildrauschens auf den Originalobjekten und seinen Auswirkungen auf die Werte der unterschiedlichen Momente. Rauschuntersuchungen auf Momenten und daraus abgeleiteten Invarianten wurden von Teh [Teh, Cho-Huak und Chin, Roland T., 1988] [Teh, Cho-Huak und Chin, Roland T., 1986] und Abu-Mostafa [Abu-Mostafa, Yaser S. und Psaltis, Demitri, 1984] vor allem auf den Momenten ganzer Grauwertbilder durchgeführt. Zu Fragen des durch die Quantisierung und Abtastung eingebrachten "Rauschen" siehe [Reiss, 1993] und [Liao, Simon X. und Pawlak, Mirosław, 1996], die auch eine datengetriebene Rekonstruktion der Grauwertobjekte mit Legendre-Momenten bis zu einer optimalen Ordnung (hinzukommende Ordnungen bringen mehr Details in die Rekonstruktion als ihre Rauschfehler und Diskretisierungsfehler zerstören) vorstellen.

Für Rauschuntersuchungen von Momenten der hier eingeführten unterschiedlichen Definitionen (3.8) (3.23) (3.30) und Vergleiche von Momenten verschiedener Ordnung untereinander bietet sich ein Signal-Rausch-Verhältnis (SNR, engl. *signal noise ratio*) aus den Varianzen der Momente und des aufgebrauchten Rauschens an. Für additives, weißes, mittelwertfreies Rauschen $n(x, y)$ der Varianz σ_n^2 und ein Grauwertobjekt $g(x, y)$, das durch ein über ein Ensemble s mittelwertfreies Zufallsfeld

$$\langle g(x, y) \rangle_s = 0$$

gegeben ist, erhalten wir jeweils eine Autokorrelationsfunktion

$$A_{gg}(x, y, u, v) = \left\langle (g(x, y))^2 \right\rangle_s e^{-C_1|x-u|-C_2|y-v|}$$

$$A_{nn}(x, y, u, v) = \sigma_n^2 \delta(x - u) \delta(y - v)$$

Dabei sind $C_{1,2}$ positive Konstanten und der Erwartungswert des Grauwertquadrates die mittlere Energie des Zufallsfeldes. Mit der Verallgemeinerung der Momente über eine Kernel-Funktion $m_{pq}(x, y)$, die bei den verschiedenen bisher definierten Momenten ein unterschiedliches Aussehen annimmt,

$$\mathfrak{M}_{pq} = \iint m_{pq}(x, y) \cdot g(x, y) dx dy$$

und dem analogen Ausdruck für das Rauschen

$$\mathfrak{N}_{pq} = \iint m_{pq}(x, y) \cdot n(x, y) dx dy$$

schreiben wir für das Signal-Rausch-Verhältnis eines verallgemeinerten Momentes vom Grad p, q mit den Varianzen des Momentes und des Rauschens

$$SNR_{pq} = \frac{\text{var}[\mathfrak{M}_{pq}]_s}{\text{var}[\mathfrak{N}_{pq}]_s}$$

Die Varianzen werden dabei als Summen der Differenzenquadrate über das Ensemble gebildet. Eine sinnvolle Normierung wäre durch den Faktor $\sigma_n^2 / \langle (g(x, y))^2 \rangle$ gegeben. Für reguläre Momente und Legendre-Momente ergibt sich dabei

für reguläre Momente:

$$SNR_{pq} = \frac{(2p+1)(2q+1)}{4\sigma_n^2} \iiint (xu)^p (yv)^q A_{gg}(x, y, u, v) dx dy du dv$$

für Legendre-Momente:

$$SNR_{pq} = \frac{(2p+1)(2q+1)}{4\sigma_n^2} \iiint P_p(x) P_p(x) P_q(y) P_q(y) A_{gg}(x, y, u, v) dx dy du dv$$

Die Integralgrenzen sind dabei auf das Intervall $[-1, +1]$ skaliert. Trägt man der Einfachheit halber für $C_1 = C_2 = 1$ die normierten SNR_{pp} über $p = 0, 1, \dots$ auf, so fallen die SNR für höhere Ordnungen ab (Einflüsse des Rauschens nehmen zu) und die SNR der Legendre-Polynome liegen deutlich unter den SNR der regulären Polynome.

Während Momente höherer Ordnung erlauben, feinere Strukturen des Grauwertobjektes zu charakterisieren, schränken die Effekte des Rauschens (stärker werdender Einfluß bei höheren Ordnungen) die Hinzunahme immer höherer Ordnungen ein, weil sie gegebenenfalls ein erreichtes Ergebnis sogar wieder verschlechtern. Die optimale Ordnung für ein gegebenes SNR bestimmt [Teh, Cho-Huak und Chin, Roland T., 1988] für additives, nicht mit dem Bildinhalt korreliertem Rauschen. Aus dem rauschfreien Rekonstruktionsfehler $e_n^2(N)$ (3.32) wird mit der Fläche des Grauwertobjektes a dann

$$e_n^2(N) = \frac{e_0^2(N) + \sigma_n^2 \cdot (N+1) \cdot (N+2) / 2}{1/a \iint (f(x, y))^2 dx dy}$$

der eine neue Definition des Signal-Rausch-Verhältnisses

$$snr = \frac{1/a \iint (f(x, y))^2 dx dy}{\sigma_n^2}$$

nahelegt. Trägt man $e_n^2(N)$ in Abhängigkeit der Anzahl aller benutzten unabhängigen Momente $(N+1)(N+2)/2$ auf, so liegt wie erwartet das Minimum dieser Funktion für steigende snr bei höheren Ordnungen N .

Dieses Ergebnis wird noch deutlicher, wenn ein Signal-Rausch-Verhältnis gebildet wird, daß den Quotienten einer **Differenz** von Momenten einer Ordnung N zu höheren Ordnungen $> N$ durch die Wurzel der Varianz des Rauschen dividiert. Für die komplexen Momente zeigt [Abu-Mostafa, Yaser S. und Psaltis, Demitri, 1984], daß diese Signaldifferenz proportional zu $1/(p+q)$ wird. Die Kovarianz für komplexe Momente wird durch die δ -Distributionen zu einem sehr einfachen Ausdruck, dem wir sofort die Varianz $\text{Var}[c_{pq}]$ entnehmen.

$$\begin{aligned}
\text{Cov}[c_{pq}, c_{rs}] &= \langle c_{pq} c_{rs}^* \rangle \\
&= \iiint \int (x + iy)^p (x - iy)^q (u + iv)^r (u - iv)^s \cdot A_{gg}(x, y, u, v) dx dy du dv \\
&= \sigma_n^2 \iiint \int (x + iy)^p (x - iy)^q (u + iv)^r (u - iv)^s \cdot \delta(p - q) \cdot \delta(r - s) dx dy du dv \\
&= \frac{2\pi\sigma_n^2}{p+q+r+s} \delta(p - q) \cdot \delta(r - s) \\
\text{Var}[c_{pq}] &= \text{Cov}[c_{pq}, c_{pq}] = \frac{\pi\sigma_n^2}{p+q+1}
\end{aligned}$$

Deren Wurzel fällt bei steigender Ordnung mit $1/\sqrt{p+q}$. Das Signal-Rausch-Verhältnis verhält sich dann für additives, weißes Rauschen wie $1/\sqrt{p+q}$.

Rauschen beeinträchtigt die Verwendung von Momenten in höheren Ordnungen jedoch nicht nur durch seinen größeren Einfluß auf die Unsicherheit bei den Werten höherer Momente. Bei Invarianten aus Momenten – wie grundsätzlich bei allen Polynomen aus Momenten – tritt bei höheren Ordnungen ein immer stärker werdender Bias auf. Wird zum Beispiel das Moment $\hat{\mu}_{pq}$ mit dem mittelwertfreien und additiven Rauschanteil n_{pq} gestört, so trägt der Erwartungswert des Momentes selbst zwar keinen Bias $\langle \mu_{pq} \rangle = \langle \hat{\mu}_{pq} + n_{pq} \rangle = \hat{\mu}_{pq}$, aber die Invarianten bekommen, je höher deren Ordnung, stärkeren Bias.

$$\begin{aligned}
\langle \psi_2 \rangle &= \langle \mu_{02}\mu_{20} - \mu_{11}^2 \rangle = \langle (\hat{\mu}_{02} + n_{02})(\hat{\mu}_{20} + n_{20}) - (\hat{\mu}_{11} + n_{11})^2 \rangle \\
&= \langle \hat{\mu}_{02}\hat{\mu}_{20} + n_{02}\hat{\mu}_{20} + n_{20}\hat{\mu}_{02} + n_{02}n_{20} - \hat{\mu}_{11}^2 - 2n_{11}\hat{\mu}_{11} - n_{11}^2 \rangle \\
&= \hat{\mu}_{02}\hat{\mu}_{20} - \hat{\mu}_{11}^2 - \langle n_{11}^2 \rangle \\
&\leq \hat{\mu}_{02}\hat{\mu}_{20} - \hat{\mu}_{11}^2
\end{aligned}$$

3.4.12 Diskrete Momente

In der Praxis der Bildverarbeitung werden wegen der bereits diskret vorliegenden Bilddaten die Integrale in den Definitionen (3.8) umgeschrieben zu Summen über alle Pixel der Fläche $A = \Delta x \Delta y$ des $X \times Y$ großen Bildgitters. So zum Beispiel für reguläre Momente (Kantenlängen werden in der Regel gesetzt zu $\Delta x = \Delta y = 1$)

$$m_{pq} = \sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} i^p \cdot j^q \cdot g(i, j) \Delta x \Delta y$$

Lin [Lin, W. G. und Wang, S. S., 1994] und Liao [Liao, Simon X. und Pawlak, Mirosław, 1996] weisen darauf hin, daß dies nur eine Näherung ist und für höhere Genauigkeit folgender Ausdruck verwendet werden sollte

$$\begin{aligned}
m'_{pq} &= \sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} g(i, j) \iint_A x^p y^q dx dy = \sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} g(i, j) \int_{x_i - \Delta x/2}^{x_i + \Delta x/2} \int_{y_j - \Delta y/2}^{y_j + \Delta y/2} x^p y^q dx dy \\
&= \frac{1}{(p+1)(q+1)} \sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} g(i, j) \left[\left(i + \frac{1}{2} \right)^{p+1} - \left(i - \frac{1}{2} \right)^{p+1} \right] \left[\left(j + \frac{1}{2} \right)^{q+1} - \left(j - \frac{1}{2} \right)^{q+1} \right]
\end{aligned}$$

Bei zentralen Momenten kommt durch die in der Regel nicht ganzzahlige Verschiebung um den Schwerpunkt bei Rundung auf ganzzahlige Werte $i - i_s$ und $j - j_s$ in den Monomen ein weiterer Fehler hinzu. Die sich ergebenden Diskretisierungsfehler lassen sich auf skalierten Bildern $x, y = -1, \dots, +1$ abschätzen zu

$$\begin{aligned}
|m'_{pq} - m_{pq}| &\leq \sqrt{\frac{4V(g) \cdot \max g}{(2p+1)(2q+1)} \Delta x \Delta y} \\
|m_{pq} - m_{pq}| &\leq \frac{\max g}{2} (p\Delta x + q\Delta y) + \sqrt{\frac{4V(g) \cdot \max g}{(2p+1)(2q+1)} \Delta x \Delta y}
\end{aligned}$$

wobei $V(g)$ die totale Variation der Grauwerte über das Bild ist. Ab der Ordnung $N \geq 2$ ergibt sich ein Korrekturterm der Diskretisierungen untereinander

$$m'_{pq} - m_{pq} = \frac{q(q-1)}{24} (\Delta y)^2 m_{p,q-2} + \frac{p(p-1)}{24} (\Delta x)^2 m_{p-2,q} + O((\Delta x \Delta y)^2)$$

3.4.13 Dreidimensionale Momente und Invarianten

Die Erweiterung [Sadjadi, Firooz A. und Hall, Ernest L., 1980] auf dreidimensionale Grauwertregionen führt analog zu den bereits vorgestellten zweidimensionalen Momenten auf die

$$\mu_{pqr} = \iiint_{x y z} (x - x_s)^p \cdot (y - y_s)^q \cdot (z - z_s)^r \cdot g(x, y, z) dx dy dz \quad (\text{dreidimensionale zentrale Momente}) \quad (3.33)$$

deren Schwerpunkt (x_s, y_s, z_s) wie in (3.9) definiert ist. Skalierungsinvarianz analog zu (3.14) kann jetzt erreicht werden mit der Normierung

$$\eta_{pqr} = \frac{\mu_{pqr}}{\frac{\mu_{000}^{p+q+r+3}}{3}} \quad (\text{normierte zentrale Momente}) \quad (3.34)$$

Sadjani [Sadjadi, Firooz A. und Hall, Ernest L., 1980] leitet drei Invarianten, die aus Momenten der Ordnung $N = 2$ bestehen, ab und ermittelt deren numerischen Werte für drei binäre Volumenstücke unterschiedlicher Form in jeweils mehreren Lagen. Weitere Invarianten aus einer zur Herleitung mit komplexen Momenten (3.23) in Abschnitt 3.4.6 analogen Rechnung finden sich bei Lo [Lo, Chong-Huah und Don, Hon-Son, 1989]. Canterakis [Canterakis, 1996] verfolgt einen Ansatz mit sphärischen Polynomen und erhält Invarianten der dritten Ordnung.

$$\begin{aligned}
\Psi_1 &= \mu_{200} + \mu_{020} + \mu_{002} \\
\Psi_2 &= \mu_{200}\mu_{020} + \mu_{200}\mu_{002} + \mu_{020}\mu_{002} - \mu_{101}^2 - \mu_{110}^2 - \mu_{011}^2 \\
\Psi_3 &= \mu_{200}\mu_{020}\mu_{002} - \mu_{002}\mu_{110}^2 + 2\mu_{110}\mu_{101}\mu_{011} - \mu_{020}\mu_{101}^2 - \mu_{200}\mu_{011}^2
\end{aligned} \quad (3.35)$$

3.4.14 Schnelle Berechnung von Momenten

Im diskreten rechteckigen Bildraum werden die Integrale in (3.8) und (3.23) zu Summen über Zeilen und Spalten eines Bildes, bzw. der von der Region abgedeckten Pixel. Meist wird dabei auf binäre Bilder $g(x, y) \in \{0, 1\}$ vereinfacht. Dann eignen sich die zu berechnenden Ausdrücke in hohem Maße für eine parallelisierte Verarbeitung. Arbeiten dazu nennt [Jiang, X. Y. und Bunke, Horst, 1991b].

Zakaria [Zakaria, M. F. u. a., 1987] vergleicht die direkte und nicht-optimierte Berechnung von Momenten bis einschließlich dritter Ordnung mit der von ihm vorgeschlagenen δ -Methode. Dabei werden Beiträge zu regulären Momenten zeilenweise berechnet und diese Zwischenergebnisse für höhere Momente erneut benutzt. Darüber hinaus greift diese Methode bei der Berechnung auf binären Formen ausschließlich auf die Koordinatenposition des ersten und letzten zur Form gehörenden Pixels in den jeweiligen Zeilen zu. Dies geht letztlich auf den Greenschen Satz zurück. Philips [Philips, 1993] ergänzte Überlegungen, die sich vor allem auf die Randpixel einer Form stützen und den bei Jiang [Jiang, X. Y. und Bunke, Horst, 1991b] [Jiang, X. Y. und Bunke, Horst, 1991a] ausführlich besprochenen (von Li [Li, Bing Cheng und Shen, J., 1991] eingeführten) Anwendung des Greenschen Satzes und seiner geeigneten Umformung von Flächenintegralen zu Kurvenintegralen entlang des Randes Γ einer einfachen geschlossenen Region erweitern. Die regulären Momente binärer Bilddaten lassen sich dann für p (oder analog dazu für q) umschreiben zu

$$m_{pq} = \frac{1}{1+p} \iint_{\Gamma} x^{p+1} y^q dx dy$$

Für das worst-case Szenario (Form bedeckt das gesamte Bild) auf quadratischen binären Bilddaten der Kantenlänge X verlangt die direkte Berechnung $10X^2$ Additionen und doppelt so viele Multiplikationen. Die δ -Methode benötigt dagegen $X^2 + 6X$ Additionen und $25X$ Multiplikationen. Je nach tatsächlichen clock-cycle-Zeiten und Genauigkeit ergibt sich daraus eine Beschleunigung, die bei Zakaria für eine typische Konfiguration mit dem Faktor 35 abgeschätzt wird. Eine von Li [Li, Bing Cheng und Shen, J., 1991] [Li, 1993] aufgezeigte konsequente Umsetzung des Greenschen Satzes und Benutzung eines Pascalschen Dreieckes kann die Komplexität von $O(X^2)$ auf $O(X)$ auch bei den Additionen reduzieren und benötigt dabei nur einige initiale Multiplikationen $\approx O(1)$. Auf Grauwertobjekten kann prinzipiell keine Komplexität unter $O(X^2)$ erreicht werden. Ganz allgemein gilt für Momente der Ordnung N auf n -dimensionalen Grauwertobjekten eine direkte nicht-optimierte Rechen-Komplexität von $O(N^n X^n)$ sowohl in den Multiplikationen wie in den Additionen, was den Bedarf an optimierten Berechnungen in hohen Dimensionen (für dreidimensionale Objekte [Yang, 1995] [Yang, Luren u. a., 1997] und [Li, Bing Cheng und Ma, S. D., 1994], [Flusser und Suk, 1999] für binäre Objekte) zeigt. Bei voller Ausnutzung des Greenschen Satzes (gute Zusammenfassung in [Yang, Luren und Albrecht, Fritz, 1994]) kann auf eine bildhafte zweidimensionale Repräsentation der binären Region verzichtet und gleich eine Ketten-Code-ähnliche Repräsentation des Randes benutzt werden, deren Berechnung den Gewinn bei der Beschleunigung allerdings wieder reduziert. Bei Grauwertobjekten kann nicht auf die reine Kontur zurückgegriffen werden, dennoch lassen sich zur Beschleunigung zweidimensionale Momente aus eindimensionalen Momenten zusammensetzen [Yang, Luren und Albrecht, Fritz, 1994]. Dadurch wurde eine Komplexität von $4X^2$ Additionen und $3X^2$ Multiplikationen erreicht. Für die dreidimensionalen Momente existiert der zum Greenschen Satz analoge Gauß'sche Satz über Ströme, mit dem Volumenintegral über Divergenzen zu Integralen über die Oberfläche werden. Voraussetzung für den Einsatz in der Berechnung von Momenten ist ein schneller Algorithmus zur Detektion der Oberflächen. Yang [Yang, 1995] erreichte damit auf binären 3D-Objekten die zweite Ordnung mit $2NX^2$ Multiplikationen und $(1/2N^2 + 7/2N + 3) X^2$ Additionen. Weitere algorithmische Entwürfe zur beschleunigten Berechnung und gute Übersichten zum Stand der Entwicklungen werden bei Li [Li, 1993] [Li, Bing Cheng und Ma, S. D., 1994] und Jiang [Jiang, X. Y. und Bunke, Horst, 1991b] und Flusser [Flusser und Suk, 1999] genannt.

3.4.15 Zusammenfassung

Mit der momentenbasierten Objekterkennung mit Parametern, die invariant unter affinen Transformationen bleiben, haben wir ein Verfahren der Objekterkennung mit Invarianten (engl. *invariant object recognition*, IOF) vorgestellt. Mit den unter den affinen Transformationen Translation, Rotation und Achsenskalierung invarianten Parametern (3.26), die wir aus den normierten zentralen Momenten berechnen können, haben wir ein mächtiges Werkzeug, um detektierte geschlossene Regionen in ihrer geometrischen Form und ihrer Relevanz für unsere Aufgabe mit einem kleinen Satz skalarer Werte zu beschreiben. Diese wurden aus einem mathematisch fundierten Gerüst abgeleitet und sind erweiterbar in höhere Ordnungen wie auch in die dreidimensionale Formanalyse. Wir haben uns dabei weitgehend auf die Analyse von Grauwertobjekten beschränkt und spezielle Fälle wie binäre Regionen oder Bildinvarianten nur am Rande behandelt. Invarianten unter projektiven Transformationen haben wir dabei ebenso außer acht gelassen wie die lokalen Invarianten oder die Konstruktion von Invarianten aus Integralen über die Transformationsgruppe.

Andere Verfahren der IOF sind Fourierdeskriptoren vor allem auf Objekträndern, direkte optische Techniken, genetische Algorithmen oder der Einsatz neuronaler Netze. Unter den zahllosen Invariantensätzen der momentenbasierten Objekterkennung gelten die Zernike-Momente [Teague, 1980] als erste Wahl, wenn es um die Rekonstruktion von Grauwertobjekten geht, insbesondere bei additivem Rauschen [Khotanzad, Alireza und Hong, Yaw Hua, 1990]. Sie haben aber den entscheidenden Nachteil, im Diskreten wegen der notwendigen Normierung auf die Einheitskreis $x^2 + y^2 \leq 1$ und den zugrundeliegenden radialen Polynominalen einen ungleich höheren Rechenaufwand zu verlangen als alle anderen vorgestellten Invariantensätze. Legendre-Momente sind zwar mit ihrer Orthogonalität (nicht redundant) theoretisch attraktiv, haben aber das schlechteste Signal-Rausch-Verhältnis [Teh, Cho-Huak und Chin, Roland T., 1988]. In der in Kapitel 4 vorgestellten Anwendung werden wir daher auf normierte zentrale Momente und deren Hu-Invarianten (3.26) zurückgreifen.

Diesen umfangreichen Abschnitt zur den Invarianten in der Mustererkennung mag die Einsicht aus dem Übersichtsartikel von Weiss [Weiss, 1993] beschließen, daß die Suche nach Objekten letztlich wohl doch viel einfacher ist als die Suche nach geeigneten Invarianten.

3.5 Polynomklassifikatoren

3.5.1 Übersicht

Nach einer Motivation von Klassifikationsverfahren als solchen stellen wir die Mathematik von drei verschiedenen Klassifikationsverfahren vor, mit denen später die zu den detektierten Änderungen aus Abschnitt 3.3.7 gefundenen Parameter aus Abschnitt 3.4 eine Zuordnung in die beiden Klassen “relevante Änderung” und “keine Änderung” für einzelne Bilder oder die gesamte Bildsequenz vorgenommen werden kann. Die Klassifikatoren werden bei der industriell orientierten Anwendung in Kapitel 4 benutzt, sollen aber bereits hier in ihrer Theorie vorgestellt und von anderen Verfahren abgegrenzt werden.

3.5.2 Problematik der Klassifikation

Probleme der Mustererkennung haben in der Regel zwei Stufen. Zuerst wird aus dem Bildmaterial eine Anzahl F von Parametern (engl. *features*) generiert, die für jedes Bild oder jeden Bildstapel einen Vektor im F -dimensionalen Parameterraum bilden, womit die für den Zweck der Mustererkennung entscheidenden Daten in kompakter Form vorliegen und letztlich eine enorme Datenreduktion verglichen mit den Eingangsbilddaten stattfand. In einem zweiten Schritt werden die gefundenen Vektoren und damit die zugrunde liegenden Bilddaten im Parameterraum bestimmten Klassen zugeordnet, die im Parameterraum als möglichst einfache Gebiete dargestellt sind. Nah beinander liegende Vektoren sollten Bilddaten mit sehr ähnlichen Mustern entsprechen, die einzelnen Klassen sollten ebenfalls durch einen charakteristischen Klassenvektor repräsentiert sein, wodurch jeder neu gemessene Vektor einer Klasse, nämlich der nächsten erreichbaren, zugeordnet werden kann. Diese Klassifikation setzt eine geeignete Metrik und ein geeignetes Abstandsmaß im Parameterraum voraus.

Obwohl die Klassifikation eine eher statistisch entscheidungstheoretische Aufgabe ist, zu der zahlreiche Lösungen aus der Datenverarbeitung vorliegen, muß bei der Konstruktion des Klassifikators auch immer auf die Auswahl der Parameter geachtet werden. Diese müssen sowohl gute Beschreibende des Musters sein, als auch für die Klassifikation und ihre Mathematik geeignet sein.

Für die Konstruktion von Klassifikatoren gibt es grundsätzlich zwei Ansätze: zum einen können die Wahrscheinlichkeitsverteilungen der Merkmale und damit die Klassen im Parameterraum modelliert werden und dann neue Merkmalsvektoren anhand des Modelles zugeordnet werden. Verteilungsfreie Klassifikatoren entwerfen anhand eines Trainingsdatensatzes (engl. *learning set*) mit bekannten Klassenzugehörigkeiten Abstandsmaße und ordnen neu zu klassifizierende Daten dann zu. Dabei wird letztlich mit zu konstruierenden Funktionalen und Gewichten ein Abbildungsproblem vom Merkmalsraum in den Entscheidungsraum über Minimierung auf einer zu definierenden Metrik gelöst. Wir werden uns hier auf letztere Vorgehensweise beschränken, um beliebige Merkmale hinzunehmen und weglassen zu können, ohne über deren Verteilungsmodelle nachdenken zu müssen.

Entscheidende Fragen sind: Welches sind die geeigneten beschreibenden Parameter? Wie viele Parameter müssen benutzt werden? Die Dimension des Parameterraumes sollte dabei so klein wie möglich gehalten werden (engl. *curse of dimensionality*). Sind die gewählten Parameter voneinander unabhängig? Wie stark beeinflussen Meßfehler oder Rauschen die einzelnen Parameter? Für die Klassifikation unserer Ereignisdetektion aus Abschnitt 4.9 werden wir auf Parameter der momentenbasierten Objektanalyse aus Abschnitt 3.4 zurückgreifen, wo zahlreiche der genannten Fragen bereits diskutiert wurden. Grundlegende Untersuchungen zu unserer hier gestellten Aufgabe finden sich bei Ritter [Ritter, 2001] (unter Verwendung eines bei [Pudil, P. u. a., 1994] vorgestellten Verfahrens zur Suche geeigneter Merkmale) und der Anschlußarbeit von Hader [Hader, 2004]. Einen umfangreichen Überblick und eine allgemeine Einführung zur Auswahl von geeigneten Parametern insbesondere für die Buchstabenerkennung (engl. *optical character recognition*, OCR) geben Trier [Trier, Oivind Due. u. a., 1996]. Weitere Arbeiten liegen von Munkelt [Munkelt, 1993] vor.

3.5.3 Mathematisches Gerüst der Polynomklassifikatoren

Polynomklassifikatoren (*polynome classifiers*, PC) kombinieren die einzelnen Merkmale mit Gewichten zu einem Polynom, dessen skalarer Wert für die Klassifikation benutzt wird. Die Gewichte werden aus dem Trainingsdatensatz generiert. Aus den $Q - 1$ Merkmalen wird unter Hinzunahme der 1 ein Vektor $\mathbf{m} = (1, m_2, m_3, \dots, m_Q)^T$ gebildet, dessen Elemente dann zu Monomen des Grades P kombiniert werden. Diese insgesamt $L = (P+Q)!/P!Q!$ Monome bilden dann den zu klassifizierenden Vektor $\mathbf{H} = (H_1, \dots, H_L)^T$. Die Kombinatorik der Potenzen der einzelnen Merkmale m_q für jedes H_l steckt dem Satz der p_{lq} . Die 1 wurde in \mathbf{m} hinzugenommen, um auch Monome vom Grad $< P$ in \mathbf{H} zu haben. Der Monomenvektors (auch *Polynomstruktur* genannt) wird in der Regel nach dem Grad der Monome sortiert: beginnend mit der 1 folgen dann die $Q - 1$ Merkmale selbst, darauf die $Q(Q-1)/2$ quadratischen Terme und so weiter.

$$\mathbf{H} = \left(1, \dots, \prod_{q=1}^Q m_q^{p_{lq}}, \dots \right)^T \quad \text{mit} \quad \sum_q p_{lq} = P \quad (3.36)$$

$$\mathbf{H} = \left(1, m_2, m_3, \dots, m_Q, m_2^2, \dots, m_Q^2, m_2 m_3, \dots, m_2^2 m_3, \dots \right)^T$$

Es wird nun eine Abbildung aus dem L -dimensionalen Merkmalsraum in den \mathfrak{k} -dimensionalen Entscheidungsraum der \mathfrak{k} Klassen gesucht. Dies geschieht mit einem inneren Produkt des Monomvektors (3.36) mit einem zu jeder Klasse κ individuellem Vektor von Gewichten $\mathbf{a}^{(\kappa)} = (a_1^{(\kappa)}, \dots, a_L^{(\kappa)})^T$. Der resultierende Skalar $s^{(\kappa)}$ wird auch *Trennfunktion* der Klasse κ genannt, da $\mathbf{a}^{(\kappa)\mathbf{T}} \cdot \mathbf{H}$ letztlich nur ein mit den $a_l^{(\kappa)}$ gewichtetes Polynom der Merkmale m_l ist.

$$s^{(\kappa)}(\mathbf{H}) = \mathbf{a}^{(\kappa)\mathbf{T}} \cdot \mathbf{H} = \sum_{l=1}^{\frac{(P+Q)!}{P!Q!}} a_l^{(\kappa)} H_l \quad (3.37)$$

Die Gewichte $a_l^{(\kappa)}$ werden später so bestimmt, daß der Wert $s^{(\kappa)}(\mathbf{H})$ der Trennfunktion ein Maß für die Zugehörigkeit von \mathbf{H} zur Klasse κ ist. Sichere Zugehörigkeit bedeutet dabei den Wert 1, sichere Nicht-Zugehörigkeit den Wert 0. Für die Klassifikation brauchen wir für jede Klasse eine Trennfunktion (also einen eigenen Satz von Gewichten), die sich zusammenfassen läßt in einer $\mathfrak{k} \times L$ großen Gewichtsmatrix

$$\mathbf{A} = \left(\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \dots, \mathbf{a}^{(\mathfrak{k})} \right) \quad (3.38)$$

Die skalaren Werte der \mathfrak{k} Trennfunktionen schreiben wir ebenfalls in einen Vektor

$$\mathbf{s}(\mathbf{H}) = \left(s^{(1)}(\mathbf{H}), s^{(2)}(\mathbf{H}), \dots, s^{(\mathfrak{k})}(\mathbf{H}) \right)^T$$

so daß die Abbildung in den Entscheidungsraum der Klassifikation als Matrixabbildung geschrieben

werden kann

$$\mathbf{s}(\mathbf{H}) = \mathbf{A}^T \cdot \mathbf{H} = \left(\sum_{l=1}^L a_l^{(1)} H_l, \sum_{l=1}^L a_l^{(2)} H_l, \dots, \sum_{l=1}^L a_l^{(\mathfrak{k})} H_l \right)^T \quad (3.39)$$

Analog zu den Zugehörigkeitsvektoren $\mathbf{s}(\mathbf{H})$ gibt es für jede Klasse κ Klassenvektoren (auch *Zielvektoren*), die in der Koordinate κ den Wert 1 und sonst 0 für sichere Zugehörigkeit zur Klasse κ haben (κ -ter Einheitsvektor der Dimension \mathfrak{k}). Mit der diskreten Kroneckerfunktion $\delta_{i,j}$ lassen sich die Klassenvektoren schreiben als

$$\mathbf{k}^{(\kappa)} = (\delta_{0,\kappa}, \dots, \delta_{i,\kappa}, \dots, \delta_{\mathfrak{k},\kappa})^T = (\dots, 0, \underbrace{1}_{\text{Index } \kappa}, 0, \dots)^T$$

Der Vektor \mathbf{H} wird der Klasse κ^* zugeordnet, für die der Abstand des Zugehörigkeitsvektors zum Klassenvektor $\mathbf{k}^{(\kappa^*)}$ minimal ist über alle Klassenvektoren. Als Maß für die Sicherheit einer Zuordnung dient der Betrag des Abstandes.

$$\kappa^* = \operatorname{argmin}_{\kappa} \left| \mathbf{s}(\mathbf{H}) - \mathbf{k}^{(\kappa)} \right| = \operatorname{argmin}_{\kappa} \left[\sum_{\kappa_0=1}^{\mathfrak{k}} \left(s^{(\kappa_0)} - \delta_{\kappa, \kappa_0} \right)^2 \right]^{\frac{1}{2}} \quad (3.40)$$

Den wesentliche Schritt bei der Klassifikation stellt jedoch die Berechnung der Gewichtsmatrix (3.38) dar. Allgemein kann festgehalten werden, daß das Variationsproblem der Suche nach einer geeigneten Abbildung aus dem Merkmalsraum in den Erwartungsraum der Klassen durch die Wahl der Monome als Basisfunktionen der polynomen Trennfunktionen überführt wurde in ein Optimierungsproblem der Gewichte.

Zu jeder Klasse κ gibt es im Trainingsdatensatz V_{κ} Monomvektoren $\mathbf{h}_{\mathbf{v}}^{(\kappa)}$. Die Berechnung der Gewichtsmatrix \mathbf{A} erfolgt mit dem Trainingsdatensatz $\{\mathbf{h}\}$, für dessen Monomvektoren die Klassenzugehörigkeit $\mathbf{s}(\{\mathbf{h}\}) \in \{\mathbf{k}^{\kappa}\}$ zu einer der \mathfrak{k} verschiedenen Klassen eindeutig bekannt ist. Wir fordern, daß für die Gewichtsmatrix \mathbf{A} das Residuum (Reststreuung) auf dem ganzen Trainingsdatensatz minimiert. Mit den Vektoreigenschaften $\mathbf{v}_1^T \mathbf{v}_2 = \operatorname{Spur}(\mathbf{v}_2 \mathbf{v}_1)$ und $\operatorname{Spur}(\mathbf{M}_1 \mathbf{M}_2) = \operatorname{Spur}(\mathbf{M}_2 \mathbf{M}_1)$ erhalten wir für das Residuum

$$\begin{aligned} S^2(\mathbf{A}) &= E \left\{ \left| \mathbf{s}(\mathbf{h}_{\mathbf{v}}^{(\kappa)}) - \mathbf{k}^{(\kappa)} \right|^2 \right\} \\ &= E \left\{ \left| \mathbf{A}^T \cdot \mathbf{h}_{\mathbf{v}}^{(\kappa)} - \mathbf{k}^{(\kappa)} \right|^2 \right\} \\ &= E \left\{ \left(\mathbf{A}^T \cdot \mathbf{h}_{\mathbf{v}}^{(\kappa)} - \mathbf{k}^{(\kappa)} \right)^T \left(\mathbf{A}^T \cdot \mathbf{h}_{\mathbf{v}}^{(\kappa)} - \mathbf{k}^{(\kappa)} \right) \right\} \\ &= E \left\{ \operatorname{Spur} \left\{ \left(\mathbf{A}^T \cdot \mathbf{h}_{\mathbf{v}}^{(\kappa)} - \mathbf{k}^{(\kappa)} \right) \left(\mathbf{A}^T \cdot \mathbf{h}_{\mathbf{v}}^{(\kappa)} - \mathbf{k}^{(\kappa)} \right)^T \right\} \right\} \\ &= \operatorname{Spur} \left\{ E \left\{ \mathbf{k}^{(\kappa)} \mathbf{k}^{(\kappa)T} \right\} \right\} - 2 \operatorname{Spur} \left\{ \mathbf{A}^T E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{k}^{(\kappa)T} \right\} \right\} + \operatorname{Spur} \left\{ \mathbf{A}^T E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{h}_{\mathbf{v}}^{(\kappa)T} \right\} \mathbf{A} \right\} \\ &= \operatorname{Spur} \left\{ E \left\{ \left| \mathbf{k}^{(\kappa)} \right|^2 \right\} \right\} - 2 \operatorname{Spur} \left\{ \mathbf{A}^T E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{k}^{(\kappa)T} \right\} \right\} + \operatorname{Spur} \left\{ \mathbf{A}^T E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{h}_{\mathbf{v}}^{(\kappa)T} \right\} \mathbf{A} \right\} \\ &= 1 - 2 \operatorname{Spur} \left\{ \mathbf{A}^T E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{k}^{(\kappa)T} \right\} \right\} + \operatorname{Spur} \left\{ \mathbf{A}^T E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{h}_{\mathbf{v}}^{(\kappa)T} \right\} \mathbf{A} \right\} \end{aligned} \quad (3.41)$$

Der Erwartungswert $E\{\}$ wird dabei über alle $\sum_{\kappa} V_{\kappa}$ Tupel aus Monomvektoren und Zielvektoren des Trainingsdatensatzes gebildet. Aus einem Variationsansatz mit der optimalen Gewichtsmatrix

$$\mathbf{A}^* = \operatorname{argmin}_{\mathbf{A}} S^2(\mathbf{A}) \quad (3.42)$$

erhalten wir die Ungleichung

$$S^2(\mathbf{A}^*) \leq S^2(\mathbf{A}^* + \delta\mathbf{A}) \quad (3.43)$$

Wir lösen die Variation $S^2(\mathbf{A}^* + \delta\mathbf{A})$ analog zu (3.41) auf

$$S^2(\mathbf{A}^* + \delta\mathbf{A}) = 1 - 2 \operatorname{Spur} \left\{ (\mathbf{A}^* + \delta\mathbf{A})^T E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{k}^{(\kappa)\mathbf{T}} \right\} \right\} + \\ + \operatorname{Spur} \left\{ (\mathbf{A}^* + \delta\mathbf{A})^T E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{h}_{\mathbf{v}}^{(\kappa)\mathbf{T}} \right\} (\mathbf{A}^* + \delta\mathbf{A}) \right\}$$

multiplizieren aus und setzen mit (3.41) in (3.43) ein.

$$0 < \underbrace{\operatorname{Spur} \left\{ \delta\mathbf{A}^{\mathbf{T}} E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{h}_{\mathbf{v}}^{(\kappa)\mathbf{T}} \right\} \delta\mathbf{A} \right\}}_{\geq 0} - 2 \operatorname{Spur} \left\{ \delta\mathbf{A}^{\mathbf{T}} \left(E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{k}^{(\kappa)\mathbf{T}} \right\} - E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{h}_{\mathbf{v}}^{(\kappa)\mathbf{T}} \right\} \delta\mathbf{A}^* \right) \right\} \quad (3.44)$$

Um die Ungleichung immer zu erfüllen, muß der zweite Term verschwinden. Daraus folgt die Bestimmungsgleichung für \mathbf{A}^* , für die die beiden Matrizen $E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{h}_{\mathbf{v}}^{(\kappa)\mathbf{T}} \right\}$ und $E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{k}^{(\kappa)\mathbf{T}} \right\}$ aus dem Trainingsdatensatz zu schätzen sind und danach durch eine lineare Multiplikation von Matrizen aufzulösen ist.

$$E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{h}_{\mathbf{v}}^{(\kappa)\mathbf{T}} \right\} \mathbf{A}^* = E \left\{ \mathbf{h}_{\mathbf{v}}^{(\kappa)} \mathbf{k}^{(\kappa)\mathbf{T}} \right\} \quad (3.45)$$

Setzt man anstelle der ‘‘harten’’ Label $\mathbf{k}^{(\kappa)}$ Wahrscheinlichkeiten (Regressionsfunktions-Ansatz) ein, so führt dies ebenfalls zu (3.45). Liegen L Datensätze im Trainingsdatensatz vor, dann hat (3.45) nur eine Lösung. Bei der Inversion der $L \times L$ Matrix $E \{ \mathbf{h}\mathbf{h} \}$ müssen mögliche Abhängigkeiten der Merkmale untereinander berücksichtigt werden, die die Matrix nicht regulär machen. Details zur Inversion, Schätzung und iterativen Berechnung der Matrix \mathbf{A} finden sich bei Niemann [Niemann, 1983]. Dort wird auch der Ansatz in (3.41) auf den Bayes-Klassifikator zurückgeführt. Polynomklassifikatoren sind eingehend beschrieben bei Schürmann [Schürmann, 1977]. Anwendungen (vor allem in der Buchstabenerkennung) in der Bildverarbeitung bespricht der Artikel [Anderson, Peter G. und Gaboriski, Roger S., 1993].

3.5.4 Abgrenzung zu anderen Klassifikatoren

Es seien an dieser Stelle zwei weitere einfache Klassifikatoren eingeführt und vorgestellt, um als Vergleich zu dem in Abschnitt 4.14 realisierten Polynomklassifikator zu dienen.

Der Nächste-Nachbar-Klassifikator (engl. *nearest neighbour*, NN) ordnet einen Parametervektor zu der Klasse des Trainingsdatensatzes, aus der der Vektor mit dem geringsten Abstand zum Parametervektor stammt. Als Abstandsmaß dient die euklidische Entfernung im Parameterraum. Damit einzelne Parameter oder Gruppen von Parametern im Vektor das Abstandsmaß nicht dominieren, müssen die Parameter jeweils normiert werden. Dabei wird analog zur Änderungsdetektion in Abschnitt 3.3.2 der Mittelwert des Parameters über die entsprechende Klasse im Trainingsdatensatz subtrahiert und durch die dazugehörige Standardabweichung dividiert. Die Abstände des Parametervektors von Vektoren des Trainingsdatensatzes in den einzelnen Dimensionen bedeuten dann Vielfache der Standardabweichung. In Formeln schreiben wir: Der L -dimensionale Parametervektor \mathbf{H} soll einer der \mathfrak{k} Klassen zugeordnet werden, für die ein Trainingsdatensatz $\{\mathbf{h}\}$ vorliegt. Zu jeder Klasse κ seien dies V_κ Vektoren $\mathbf{h}_v^{(\kappa)} = [h_{v1}^{(\kappa)}, h_{v2}^{(\kappa)}, \dots, h_{vL}^{(\kappa)}]^T$ der Dimension L . Die Zuordnung zur Klasse κ^* erfolgt durch die Minimierung des Abstandsmaßes $d\{\}$ über alle Klassen $\kappa = 1, \dots, \mathfrak{k}$. Anstelle der euklidischen Entfernung können auch andere Abstandsmaße (Mahalanobis-Maß) verwendet werden.

$$\kappa^* = \operatorname{argmin}_{\kappa=1, \dots, \mathfrak{k}} d\{\mathbf{H}, \mathbf{h}_v^{(\kappa)}\} \quad \text{wobei } v = 1, \dots, V_\kappa$$

$$d\{\mathbf{H}, \mathbf{h}_v^{(\kappa)}\} = \sqrt{\sum_{l=1}^L \left(\hat{H}_l^{(\kappa)} - \hat{h}_{vl}^{(\kappa)} \right)^2}$$

$$\hat{H}_l^{(\kappa)} = \frac{H_l - \langle h_{vl}^{(\kappa)} \rangle_v}{\sigma_{h_l}^{(\kappa)}} \quad \hat{h}_{vl}^{(\kappa)} = \frac{h_{vl}^{(\kappa)} - \langle h_{vl}^{(\kappa)} \rangle_v}{\sigma_{h_l}^{(\kappa)}}$$

$$\langle h_{vl}^{(\kappa)} \rangle_v = \frac{1}{V_\kappa} \sum_{v=1}^{V_\kappa} h_{vl}^{(\kappa)} \quad \sigma_{h_l}^{(\kappa)} = \frac{1}{V_\kappa} \sqrt{\sum_{v=1}^{V_\kappa} \left(h_{vl}^{(\kappa)} - \langle h_{vl}^{(\kappa)} \rangle_v \right)^2}$$

Unter dem k -Nächste-Nachbar-Klassifikator (k -NN) versteht man ein sehr ähnliches Verfahren, bei dem die Zugehörigkeit von \mathbf{H} zur Klasse κ^* aber nicht von einem einzigen Abstandswert bzw. Repräsentanten $\mathbf{h}_{v_0}^{(\kappa^*)}$ der Klasse κ^* bestimmt wird, sondern von den k Vektoren $\mathbf{h}_v^{(\kappa)}$, die den geringsten Abstand haben. Die Entscheidung kann für eine Klasse aus den k eventuell unterschiedlichen Zugehörigkeiten κ^* “demokratisch” oder auch mit dem Abstandsmaß gewichtet erfolgen. Verweise auf Arbeiten zu schnellen Suchstrategien und effizienten Implementierungen finden sich in [Belkassim, S. O. u. a., 1991].

Ein zweiter Klassifikator ist der *minimum-mean-distance*-Klassifikator (MMD) [Jain, Anil K. u. a., 2000], der die Zugehörigkeit von \mathbf{H} zur Klasse κ^* nicht an allen $\mathbf{h}_v^{(\kappa)}$ des Trainingsdatensatzes prüft, sondern an einem aus diesen berechneten mittleren Repräsentanten $\langle \mathbf{h}_v^{(\kappa)} \rangle_v$. Dieser enthält in seinen L Dimensionen die Mittelwerte $\langle h_{vl}^{(\kappa)} \rangle_v$. Der Abstand zu den \mathfrak{k} Klassen-Repräsentanten $\langle \mathbf{h}_v^{(\kappa)} \rangle_v$ wird darüber hinaus in den einzelnen Koordinaten mit dem Quadrat von $\sigma_{h_l}^{(\kappa)}$ gewichtet. Der geringste Abstand $\delta\{\}$ bestimmt wieder die Klassenzugehörigkeit κ^* .

$$\kappa^* = \operatorname{argmin}_{\kappa=1, \dots, \mathfrak{k}} \delta\{\mathbf{H}, \kappa\} = \operatorname{argmin}_{\kappa=1, \dots, \mathfrak{k}} \sum_{l=1}^L \frac{\left(H_l - \langle h_{vl}^{(\kappa)} \rangle_v \right)^2}{\left(\sigma_{h_l}^{(\kappa)} \right)^2}$$

Der Hauptunterschied der beiden alternativ vorgestellten Klassifikatoren für die Praxis ist also, daß beim NN-Klassifikator der Vektor \mathbf{H} mit $\kappa < \sum_{\kappa} V_{\kappa}$ Vektoren des Trainingsatzes über das Abstandsmaß verglichen wird, während beim MMD-Klassifikator nur κ Vergleiche stattfinden.

In unserer Anwendung in Kapitel 4 werden wir einen Polynomklassifikator verwenden. Bei der Abgrenzung des Polynomklassifikators gegen ein Multi-Layer-Perzeptron (MLP) folgen wir der Argumentation von Kreßel [Kreßel, Ulrich u. a., 1990]. Beide Verfahren berechnen Skalarprodukte und sind daher schnell und parallelisierbar. Weil der Polynomklassifikator mit den Monomen (3.36) nicht-lineare Basisfunktionen hat, diese aber linear kombiniert (3.37), gibt es für die Gewichte eine direkte Lösung mit globalem Optimum. Den Gewichten läßt sich zudem eine Rangordnung entnehmen, die für die Auswahl von Merkmalen wertvolle Information enthält. Dieser Vorteil gilt auch im Vergleich mit den k-Nächste-Nachbar-Klassifikatoren. Ein wichtiger Nachteil bei k-Nächste-Nachbar-Klassifikatoren ist die hohe Anzahl von Vergleichen, die bei großen Trainingsdatensätzen durchzuführen sind. Dies verlangt sehr effiziente Suchstrategien und im Speicher gehaltene Merkmalsvektoren des Trainingsdatensatzes, wenn der Algorithmus schnell bleiben soll. Solche Probleme treten beim Polynomklassifikator mit einer einmaligen Berechnung eines Skalarproduktes zur Laufzeit der Klassifikation und der weniger Speicher-intensiven Vorhaltung der Gewichte nicht auf. Einen vergleichenden Überblick zu Klassifikationsverfahren geben [Schürmann, 1996] und [Niemann, 1983]. In [Toth, Daniel und Aach, Til, 2001] werden Minimum-distance und eine fuzzy-Implementierung [Keller, James M. u. a., 1985] eines kNN-Verfahrens miteinander verglichen.

3.5.5 Zusammenfassung

Zu den als potentielle Bereiche von Änderungen identifizierten Regionen innerhalb des Bildstapels liegen Merkmale vor, anhand derer eine abschließende Entscheidung getroffen werden soll, ob die identifizierte Änderung ein relevantes Ereignis ist oder nicht. Die Abbildung aus dem Merkmalsraum mit typischerweise 10^1 Merkmalen in den Entscheidungsraum mit 2 Dimensionen ist eine weitere Datenreduktion, die direkt auf die gewünschte Aussage führt. Wahl und Konstruktion eines Klassifikators muß eng verwoben mit der Auswahl und Bewertung von Merkmalen geschehen, wie die gestellten Fragen gezeigt haben. Die hier verwendeten verteilungsfreien Klassifikatoren, die eine solche Abbildung bewerkstelligen, benötigen Vorwissen (Trainingsdaten) in Form von Merkmalsvektoren und der jeweiligen Klassenzugehörigkeit. Aus den vorgestellten Klassifikatoren (Nächster-Nachbar, k-Nächster-Nachbar, Minimum-Mean-Distance und Polynomklassifikator) haben wir den Polynomklassifikator als für unsere Zwecke geeigneten identifiziert.

Kapitel 4

Ereignisdetektion zur Qualitätssicherung

Die erste Anwendung der vorgestellten statistischen Methode zur Detektion von Änderungen wird Bildfolgen aus der industriellen Qualitätskontrolle untersuchen. Aufgabe ist es, einen zu filmenden Prozeß abzusichern gegen plötzlich auftauchende Schwankungen, Objekte, die das Bildfeld durchqueren, Sensorausfälle und dergleichen mehr. Es sind also für jede einzelne Bildfolge all jene Pixel zu markieren, die zu plötzlichen, singulären und unerwarteten Ereignissen in ihren Grauwerten gehören. Die Detektion von geänderten Regionen wird hierbei zu einer Objektdetektion weiterentwickelt und erlaubt als zentraler Baustein der Auswertung schließlich eine abschließende Gut/Schlecht-Aussage über den beobachteten Prozeß und damit über das in diesem Prozeß gefertigte Bauteil.

4.1 Übersicht

Moderne Bildverarbeitungssysteme ermöglichen heute industrielle Inspektions- und Qualitätssicherungsaufgaben, die bis vor kurzem noch als unrealisierbar galten. In diesem Kapitel wird eine Anwendung der Algorithmik aus Kapitel 3 für die industrielle Prozeßüberwachung zur Qualitätssicherung vorgestellt. Qualitätskontrolle in der automatisierten Massenproduktion bedeutet traditionell die Prüfung des fertigen Werkstückes, kann aber auch die Überwachung des Herstellungsprozesses selbst sein. Ein von einer Kamera beobachteter Arbeitsschritt ist dann in der Regel auf Anomalien zu prüfen. Solche Ausreißer deuten auf Fehler am Produkt oder im Prozeß hin und müssen zum Ausschleusen der betroffenen Teile und einem Signal für Wartungsbedarf an der betroffenen Station führen.

Nach einer Einführung in die drei Themen industrielle Bildverarbeitung (Abschnitt 4.2), Laserschweißen (Abschnitt 4.3) und CMOS-Kameras (Abschnitt 4.4) werden aus der konkreten Anwendung (Abschnitt 4.5) die technischen Randbedingungen (Abschnitt 4.6) entnommen, und ein Lösungsansatz (Abschnitt 4.7) für die gestellte Aufgabe entworfen. Nach der Bereitstellung der Bilddaten und eines Ablaufkonzeptes der Auswertung (Abschnitt 4.8) passen wir die in Kapitel 3 vorgestellte Algorithmik auf dieses Problem an (Abschnitt 4.9 bis 4.11) und erläutern die Ergebnisse (Abschnitt 4.15 und 4.16) der Auswertungen. Eine Bewertung und Vergleich (Abschnitt 4.17) mit anderen Systemrunden das Kapitel ab.

4.2 Bildverarbeitung im industriellen Umfeld

Im Bereich der industriellen Fertigung hat die Bildverarbeitung (engl. *machine vision*) in den letzten zehn Jahren einen entscheidenden Anteil am Automatisierungsgrad moderner Produktionsabläufe. Bildverarbeitung wird eng verzahnt mit der Robotik eingesetzt für Positionierungs-, Zuführungs- und Lagekontrollaufgaben, zur Vermessung von Abständen, Flächen, Farben oder Oberflächen, zur Vollständigkeitsprüfung von Produktionsschritten (Bestückungen, Bündelungen), zur Identifikation (Barcodelesen), zur Orientierung von Robotern (Bahnführung), zur Sortierung (zum Beispiel von Schüttgut oder bei der Lagerhaltung von Kleinteilen). Gerade bei der Qualitätsprüfung wird sie oft als Ersatz zu einer vom Menschen durchgeführten visuellen Sichtprüfung installiert und bietet darüber hinaus neue Möglichkeiten der Sichtbarmachung und Bewertung (interferenzbasierte hochauflösende Oberflächenprüfung, thermographische Daten zur Beurteilung von Kontaktierungen, Röntgen-Computertomographie zur zerstörungsfreien Prüfung).

Ziel ist es dabei stets, Prüfsysteme als Inline-Systeme vollständig in Produktionsabläufe zu integrieren, da heutige Qualitätsanforderungen so hoch sind, daß stichprobenartige Einzelprüfungen (Offline-Systeme) nicht mehr als ernstzunehmende Qualitätssicherung verstanden werden. Der Markt für industrielle Bildverarbeitung in Europa wird auf 1,8 Milliarden € jährlich geschätzt (2001, weltweit 7 Milliarden) und wächst seit vielen Jahren mit 20% und mehr. Ausschlaggebend für den Erfolg der industriellen Bildverarbeitung sind als Motor Qualitätsstandards wie DIN ISO 9000ff und der Trend zur automatisierten 100%-Kontrolle. Der PC gab die entscheidenden Impulse für die enormen Leistungssteigerungen und legte die technische Grundlage für die heutigen kompakten Systeme, bei denen ein PC Daten aus mehreren Kameras verarbeitet und damit Roboter selbst steuert oder der vorhandenen Maschinensteuerung die notwendigen Daten bereitstellt. Industrielle Bildverarbeitung paßt darüber hinaus genau in den Trend zum durchgehenden Informationsfluß bei der Montage und Herstellung: die Weitergabe von Qualitätsdaten aus der Produktion an den Kunden wird immer mehr zum Verkaufsargument einer informationstechnisch konsequent begleiteten Herstellung.

Industrielle Bildverarbeitung ist nicht mehr als eine Datenreduktion hochdimensionaler Daten (Bildfolgen, Mehrkanalbilder, ...) auf eine binäre Gut/Schlecht-Aussage über das produzierte Teil oder einen beobachteten Prozeß. In der einfachen Logik der automatisierten Produktion entspricht die Schlecht-Aussage in der Regel einer Verschrottung des geprüften Teiles ohne weitere Nacharbeit.

Die Anforderungen an industrielle Bildverarbeitung sind: Robustheit, Schnelligkeit, geringer Wartungsaufwand, kein Justageaufwand beim Umrüsten. Ein System muß von fachfremdem Personal ohne großen Schulungsaufwand intuitiv bedienbar sein und in der oft ununterbrochenen Produktion ein Minimum an Service verlangen. Einfache Bedienelemente, geringe Anzahl frei wählbarer Parameter, Einsatz leicht tauschbarer Standardkomponenten, ein gewisses Maß an Selbstüberprüfung und Konsistenzprüfung, Unabhängigkeit von äußeren Einflüssen (Temperatur, Störstrahlung, sich ändernde Außenbeleuchtungen) sind für ein robustes System unerlässlich. Taktraten in der Elektronikindustrie oder dem Automobilbereich betragen in der Regel 1 – 10 s, in Einzelfällen < 1 s oder bis zu 30 s. Innerhalb eines solchen Taktes müssen Bildaufnahme, Bildverarbeitung, Gut/Schlecht-Entscheidung und die gesamte Kommunikation der Startsignale und Ergebnisse mit dem Transportsystem einer Produktionsstation zeitsicher abgeschlossen sein. Für die reine Rechenzeit auf einem im Speicher befindlichen Bilddatensatz verbleiben dann in der Regel nur ein Drittel Takt.

4.3 Laserschweißen

Um die hier überwachte Industrieapplikation einordnen zu können, wird im folgenden die Bedeutung des Lasers anhand der gewählten Schweißanwendung vorgestellt. Einführungen in die Thematik Laserschweißen finden sich auch in [Dilthey, 2000]. In der industriellen, automatisierten Massenproduktion hat der Laser in den letzten beiden Jahrzehnten für die Aufgabenstellungen Verbinden, Bohren, Hohnen, Entgraten, Schneiden mehr und mehr Einzug gehalten. Besonders in der Fügetechnik löst das Laserschweißen die vorher bekannten Verfahren (Lichtbogenschweißen, Gasschmelzschweißen, Löten) zunehmend ab. Entscheidend dafür sind die erreichbaren hohen Schweißtiefen und die Wirtschaftlichkeit beim Laserschweißen.

Mit einem Weltmarkt für Laserquellen von 8,8 Milliarden € (2000) hat die Lasermaterialbearbeitung in den letzten beiden Jahrzehnten für die industrielle, automatisierte Massenproduktion stark an Bedeutung gewonnen. Prognosen für die kommenden Jahre gehen weiter von einem jährlichen Wachstum von ca. 10% für die Makrobearbeitung und rund 14% für die Mikrobearbeitung aus [Mayer, 2001]. Die Anwender von Lasermaterialbearbeitungssystemen sind die Halbleiter- und Mikroelektronikindustrie (23%), Leiterplattenindustrie (10%), Elektromechanische Industrie (9%), Automobilindustrie und Zulieferer (10%), Job Shops (18%) sowie andere metallverarbeitende (18%) und nicht-metallverarbeitende Industriezweige (12%). Die Aufgabenstellungen sind entsprechend der unterschiedlichen Anwendungsfelder vielfältig: Schweißen und Schneiden (45%), Markieren und Bohren (20%) sowie andere Verfahren (35%) [Steen, 1998] [Mayer, 2001]. Anfangs im makroskopischen Bereich für Materialien wie Stahl eingesetzt, erobert das Laserschweißen inzwischen auch mikroskopische Aufgaben und die Bearbeitung von Kupfer, Aluminium oder gar Kunststoffen.

Die starke Verbreitung des Laserschweißens und das prognostizierte Wachstum verdeutlichen die Notwendigkeit, diesen Fertigungsprozeß zu beherrschen und ihn überwachen zu können. Der Laserschweißprozeß besitzt jedoch eine hohe Komplexität und eine große Mannigfaltigkeit an Einflußgrößen. Diese lassen sich in Gruppen einteilen [Hügel, 1992]: Strahleigenschaften, Prozeßgestaltung, Bauteileigenschaften, Gerätetechnik. Diese vier wichtigen Kategorien werden nun kurz besprochen. Die Strahleigenschaften werden durch die Wellenlänge, Strahlqualität, Betriebsmodus gepulst (pw = pulsed wave) oder kontinuierlich (cw = continuous wave), Energie oder Leistung sowie der Polarisation beschrieben. Anhand der Wellenlänge werden drei verschiedene Laserquellen unterschieden. CO₂-Laser, Nd:YAG-Laser und zunehmend auch Diodenlaser. Sie werden in der folgenden Tabelle mit ihren wesentlichen Eigenschaften einander gegenübergestellt (aktualisiert zu [Herrmann, 2002]).

Eigenschaft	Einheit	CO ₂ -Laser	Nd:YAG-Laser	Diodenlaser
Wellenlänge	µm	10,6	1,06	0,80 - 0,95
Gesamtwirkungsgrad	%	5 - 10	1 - 3	30 - 40
Laserkopfgröße	dm ³	10 ³	10 ²	10 ¹ - 10 ³
Ausgangsleistung	kW	20 - 40	5 - 6	2 - 6
Leistungsdichte	W/cm ²	10 ^{6...20}	10 ^{5...9}	10 ^{3...6}
Strahlparameterprodukt	mm · mrad	4 - 10	25	25 - 200
Wartungsintervall	h	1000	500	2000 - 10000
Kosten	€/w	150 - 300	200 - 600	100 - 300

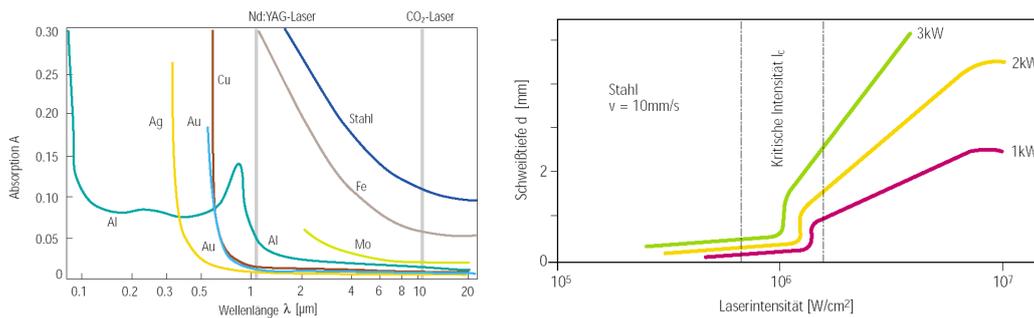


Abbildung 4.1: Links: Nd:YAG-Laser arbeiten eher im Infraroten, was die Übertragung mit Lichtwellenleitern erlaubt. Nachteilig ist, daß sie schlechter in Metalle einkoppeln. Rechts: Erst ab einer kritischen Intensität I_c bildet sich eine Dampfkapillare aus, und das Wärmeleitungsschweißen geht in ein Tiefschweißen (in Abbildung 4.2 veranschaulicht) über.

Das laseraktive Medium beim CO₂-Laser ist Kohlenstoffdioxid-Gas, das über Gasentladungen zum Leuchten gebracht wird. Beim Nd:Yag-Laser, einem Festkörperlaser, wird ein mit optisch aktiven Nd-Ionen dotierter YAG-Kristall (Yttrium-Aluminium-Granat, einige Yttrium-Ionen sind durch Neodym ersetzt) mit Anregungslampen (Halogenlampen oder Krypton-Bogenlampen) optisch gepumpt und zum Leuchten angeregt. Der CO₂-Laser zeichnet sich durch sehr hohe Leistungen bis 20 kW bei gleichzeitig hoher Strahlqualität aus. Aufgrund seiner hohen Wellenlänge wird sein Licht von Glas vollständig absorbiert und daher über Spiegelsysteme zum Werkstück geleitet, was unter Umständen justageaufwendig oder verschmutzungsanfällig sein kann. Abbildung 4.1 vergleicht die Absorptionsgrade bei verschiedenen Wellenlängen. Nicht-Eisenhaltige Metalle lassen sich aufgrund ihres Absorptionsverhaltens der Laserenergie besser mit Nd:YAG bearbeiten. Der Nd:YAG-Laserstrahl und der Diodenlaser mit einer Wellenlänge im nahen Infrarot lassen sich sehr einfach und wartungsfrei über Lichtwellenleiter (LWL), die bei $\approx 1 \mu\text{m}$ so gut wie gar nicht absorbieren, an den Prozeßort transferieren. Bei den Diodenlasern befindet sich die Weiterentwicklung noch in vollem Gange – teilweise erreichen sie aber jetzt schon die Strahlqualität bei geringerer Leistung oder die gleiche Leistung bei geringerer Strahlqualität wie ein Nd:YAG-Laser. CO₂-Laser werden vorzugsweise zum Tiefschweißen massiver Metallkörper, etwa im Schiffsbau, oder wenn es auf große Einschweißiefen ankommt, zum Beispiel im Fahrzeug- und Flugzeugbau, erfolgreich eingesetzt. Nd:YAG- und zunehmend auch Diodenlaser sind für das Fügen kleiner oder dünnwandiger Bauteile prädestiniert, wie zahlreiche Anwendungen aus der Medizintechnik und der elektrotechnischen und elektromechanischen Industrie mit Massenfertigung von Konsumprodukten und zahllosen Kleinteilen belegen. Eine vergleichende Untersuchung von Lasertypen findet sich bei [Hack, 1998]. Die Prozeßgestaltung beinhaltet alle Parameter wie das Verwenden von Prozeß- oder Schutzgasen, das Zuführen von Zusatzwerkstoffen zur Nahtverstärkung, die genaue Positionierung des Laserstrahls gegenüber der Fügestelle des Bauteils und nicht zuletzt die Vorschubgeschwindigkeit, mit der das Werkstück verschweißt wird. Wird die Geschwindigkeit verringert, so wird pro Zeiteinheit mehr Energie in den vorbeigeführten Werkstoff eingebracht, was zu einer größeren Einschweißtiefe führt. Es ist einsichtig, daß Bauteileigenschaften wie Form der Fügegeometrie, verwendeter Werkstoff oder auch die Oberflächenbeschaffenheit (Rauheit, Kratzer, Verschmutzung) die Schweißung beeinflussen. Die Oberflächenbeschaffenheit kann die Einkopplung des Lasers

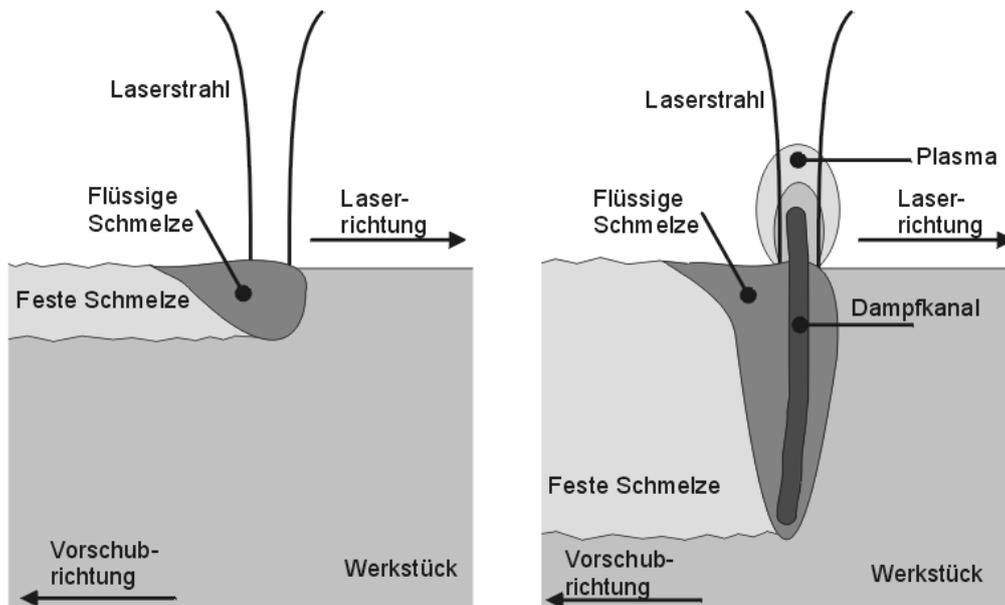


Abbildung 4.2: Schematische Darstellung von Wärmeleitungsschweißen und Lasertiefschweißen, bei dem durch höheren Energiefluß ein Metall verdampft, ein Plasma ausgebildet wird und die Laserenergie durch den Dampfkanal viel tiefer in das Material eindringen kann.

begünstigen oder im Falle der Verschmutzung Prozessinstabilitäten nach sich ziehen. Der Werkstoff selbst trägt über seine Schweißneigung bei und kann durch fertigungstechnische Inhomogenitäten ebenfalls zu Prozeßinstabilitäten wie Spritzern oder Löchern in der Naht führen. Die Bauteilgeometrie bestimmt nicht nur den Schweißprozeß in seiner Qualität, sondern beeinflusst auch in erheblichem Maße die Emissionen, die zur Überwachung herangezogen werden. Die Gerätetechnik als letzte Parametergruppe wird gerne vernachlässigt, obwohl dies nicht immer richtig ist. Das Laserstrahlschweißen läßt sich in Abhängigkeit von der an der Fügestelle erreichten Leistungsdichte (Abbildung 4.1 links) in zwei Prozeßmoden einteilen: beim Wärmeleitungsschweißen (Abbildung 4.2 links) wird die Strahlung in einer dünnen Schicht ($< 1 \mu\text{m}$) an der Materialoberfläche der Metalle absorbiert [Dausinger, 1995] und in Wärme umgewandelt. Der Energietransport in das Werkstück erfolgt über Wärmeleitung. Mit steigender Leistungsdichte steigt die Temperatur lokal am Wechselwirkungsort. Wird die Schmelztemperatur des Metalls überschritten, bildet sich entsprechend dem Einwirkungszeitraum ein Schmelzbad aus. Zusätzlich zur Wärmeleitung kommt noch die Konvektion in der Schmelze als ein weiterer Mechanismus des Energietransports hinzu. Diese Schmelzströmungen resultieren aus den in Abhängigkeit von der Temperatur variierenden Oberflächenspannungen. Durch einen kontinuierlichen Vorschub des Werkstückes wird ein Schmelzfleck, in dem das geschmolzene Material der zu fügenden Teile sich ineinander vermischt, durch das Werkstück gezogen. Bei geeigneten Parametern ist dieser Zustand (quasi-)stationär. Die so erreichten Nahtgeometrien zeigen eine Einschweißtiefe, die kleiner oder gleich der Nahtbreite ist. Die Energien beim Wärmeleitungsschweißen reichen im Unterschied zum Tiefschweißen nicht aus, um das Metall zu verdampfen. Beim Tiefschweißen (Abbildung 4.2 rechts) ist die Leistungsdichte so hoch, daß im Bereich der Wechselwirkungszone des Lasers mit

der Schmelze die Verdampfungstemperatur überschritten wird. Der abströmende Metaldampf drückt die Schmelze auseinander, so daß der Laser tiefer liegende Bereiche der Fügestelle erreichen kann. Die Folge ist die Ausbildung einer Dampfkapillare (engl. *keyhole*) innerhalb der Schmelze mit einem Durchmesser in der Größenordnung des Strahldurchmessers. Durch Mehrfachreflektionen an den Wänden in dieser Kavität wird das an den Dampfkanal angrenzende Material aufgeschmolzen und zirkuliert heftig. Über dem Werkstück steht eine leuchtende Plasmafackel. Das Prozeßleuchten des Metaldampfes hat bei der Wellenlänge des CO₂-Lasers einen erheblich höheren Absorptionsgrad als bei der Wellenlänge des Nd:YAG-Laser. Auch beim Tiefschweißen spielen diese Schmelzströmungen und der Werkstücktransport eine wichtige Rolle für den (quasi-) stationären Prozeß. Die Schmelze muß im Gegensatz zum Wärmeleitungsschweißen die Kapillare umfließen. Zusammen mit der höheren Einschweißtiefe ergeben sich komplexere Strömungsmuster in der Schmelze. Dies hat insgesamt zur Folge, daß der Tiefschweißprozeß in der Regel wesentlich unruhiger ist als ein Wärmeleitungsschweißprozeß. Die erzielten Einschweißiefen sind mindestens so groß wie der Strahldurchmesser bzw. die Nahtbreite, in der Regel betragen sie jedoch ein Vielfaches davon.

Vorteile und Nachteile des Laserschweißens gegenüber herkömmlichen Verfahren der Verbindungstechnik, wie dem Metall-Aktivgasschweißen (MAG) oder Wolfram-Inertgasschweißen (WIG), lassen sich zusammenfassen:

- Geringe thermische Belastung des Werkstücks und damit wenig Verzug. Die hohen Energiedichten gehen voll in den Fügeprozeß ein und heizen durch Verlustwärme das Bauteil kaum auf. Das ist die Grundlage zum Fügen von Präzisionsbauteilen. Lediglich das Elektronenstrahlschweißen erreicht vergleichbare Energiedichten.
- Einfache Handhabung der Gerätetechnik und hohe Prozeßgeschwindigkeit. Das Laserschweißen stellt geringe Anforderungen an die Zugänglichkeit in der Geometrie des Werkstückes, ist über den Laser leicht steuerbar, leichter automatisierbar, arbeitet berührungslos und mit hoher Geschwindigkeit.
- Großes Nahttiefe- zu Nahtbreiteverhältnis und nacharbeitsfreie Nahtqualität. Die Nähte, die beim Laserschweißen erstellt werden können, sind generell schlanker und unterstützen die Miniaturisierung der bearbeiteten Teile. Gleichzeitig erfordern Schweißnähte, die mit einem Laser erstellt wurden, in der Regel keine Nacharbeit.

Die Qualitätsanforderungen an die Laserschweißverbindung können entsprechend der Funktion des Bauteils vielfältig sein: Festigkeit, Dichtigkeit, Verzugfreiheit, Oberflächengüte und andere. Diese Funktionsanforderungen stehen in einem kausalen Zusammenhang mit einigen wenigen Schweißnahtcharakteristika, den qualitätsrelevanten Merkmalen:

- Nahtlage in Relation zur Fügestelle: z.B. Lateralversatz oder Winkelfehler
- Nahtgeometrie: z.B. Einschweißtiefe oder Anbindequerschnitt
- Nahtfehler: z.B. Löcher oder Poren

Diese genannten Merkmale lassen sich online und offline überwachen, um den Schweißprozeß abzusichern [Glumann, 1996] [Griebsch, J. u. a., 1996]. offline heißt, daß die Schweißverbindung nach dem Fertigungsprozeß einer zerstörenden oder nicht zerstörenden Werkstückprüfung unterzogen wird. Zerstörende Verfahren sind Schliff und Festigkeitstests wie Zug- oder Biegebelastung. Naturgemäß eignet sich die zerstörende Prüfung nicht für eine 100%-Kontrolle, sondern lediglich für eine mit Stichproben durchgeführte statistische Absicherung. Naheliegende zerstörungsfreie Tests sind die Sichtprüfung durch den Menschen oder durch Bildverarbeitungssysteme zur Beurteilung von Oberflächenmerkmalen. Weitere Verfahren sind aktive Thermographie, Wirbelstromprüfungen sowie Röntgen-Computertomographie oder Ultraschall für im Inneren des Bauteiles liegende Merkmale. offline-Prüfungen prüfen immer nur das fertige Ergebnis eines Prozesses, nicht den Prozeß selbst. Nachteile der Sichtprüfung oder der Techniken zur inneren Prüfung sind die enormen Kosten. Sichtprüfung ist als nicht-technisches Verfahren nicht voll reproduzierbar und hat allein durch Ermüdung und subjektiven Ermessensspielraum einen Schlupf. Bildverarbeitungssysteme kämpfen bei der Inspektion von nicht-ebenen Metalloberflächen mit Spiegelungen, die nur durch aufwendige Beleuchtung in den Griff zu bekommen sind. Schweißnähte mit ihrer schillernden und schuppigen Oberfläche lassen sich nur äußerst schwierig offline mit Bildverarbeitung prüfen. Anstelle der klassischen Abbildung unter geeigneter Beleuchtung sind daher bei breiten Nähten Lichtschnittverfahren gebräuchlich. Üblicherweise wurde bisher damit aber keine Nahtinspektion, sondern vor allem eine Nahtverfolgung, also eine Detektion der Fügestelle, vor der Schweißung mit entsprechender Bahnführung des Lasers [Horn, 1994] [Ruoff, 1989] realisiert.

online-Prüfung bedeutet, daß bereits während der Schweißung mit geeigneten Sensoren Größen erfaßt werden, die einen Rückschluß auf die qualitätsrelevanten Merkmale erlauben. Diese Größen werden eingeteilt in:

- Primärindikatoren beschreiben das Prozeß- und Schweißergebnis. Sie sind also identisch mit den qualitätsrelevanten Merkmalen, die zu den Oberbegriffen Nahtlage, Nahtgeometrie und Nahtfehler gehören.
- Sekundärindikatoren beschreiben den Schweißprozeß. Diese Indikatoren sind akustische und optische Prozeßemissionen, die sich meist einfach detektieren lassen. Dagegen gestaltet sich die Korrelation mit dem Schweißergebnis, den Primärindikatoren, in der Regel schwieriger. Wie hängt beispielsweise ein Signalverlauf im ultravioletten Spektralbereich mit der Einschweißtiefe zusammen? Die Antwort kann nur mit einem unter bestimmten Voraussetzungen gültigen Modell gegeben werden.
- Tertiärindikatoren beschreiben die Einflußgrößen auf den Prozeß und das Ergebnis. Dies sind die zuvor erläuterten Laser-, Bauteil- und Maschinenparameter.

Klarer Vorteil der online-Prüfung ist der Blick in den Prozeß und damit in die "Entstehung von Qualität". Hier liegt auch der Schlüssel zu einem physikalischen Verständnis der Ursachen von Qualitätseinbußen und Prozeßunruhen. Nachteilig ist die durch den Prozeß vorgegebene Aufnahmezeit der Meßdaten und beim Schweißprozeß speziell die Tatsache, daß dieser mit empfindlicher Sensorik kaum zu beobachten ist: gleißend helle Metallteile, Störstrahlung und -felder, Hitze, Spritzer und Oxid-Schmauch gefährden im Dauerbetrieb die Sensorik.

Heutiger Stand der Technik ist es, die akustischen und optischen Prozeßemissionen mit geeigneten Detektoren aufzuzeichnen. Diese Detektoren sind Mikrophone, räumlich integrierende Photodioden oder CCD-Kameras mit Standard-Video-Bildrate. Die Emissionen lassen sich grob den Prozeßbestandteilen zuordnen:

- Körper- oder Luftschall: Abgesehen von wenigen speziellen Applikationen, etwa dem Laserbeschriften mit gepulsten Lasern, haben die Störanfälligkeit gegenüber Umwelteinflüssen und die eher schlechten Korrelationen zu den qualitätsrelevanten Merkmalen bislang einen industriellen Einsatz verhindert.
- ultraviolettes Spektrum ($0,2 \dots 0,38 \mu\text{m}$): Das Plasma, das sich beim CO_2 -Laserschweißen in und über der Dampfkapillare bildet und in diesem Bereich strahlt. Damit lassen sich Aussagen über den direkten Wechselwirkungsbereich des Lasers mit dem Werkstoff, also dem Dampfkanal, machen.
- sichtbares Spektrum ($0,38 \dots 0,78 \mu\text{m}$): Auch in diesem Bereich strahlt das Plasma. Beim Nd:YAG-Laserschweißen entsteht kein Plasma. Stattdessen strahlt der heiße Metalldampf im sichtbaren Spektrum.
- nahes Infrarot ($0,78 \dots 3,0 \mu\text{m}$): Metalldampf, Schmelzbad und -spritzer sowie wiedererstartetes Material strahlen in diesem Wellenlängenbereich. Eine Unterscheidung ist durch eine geeignete Wahl des Meßfeldes bei Photodioden oder räumlich auflösende Sensoren wie einer Kamera möglich.
- Laserwellenlänge (Halbleiter: $0,8 \dots 0,95 \mu\text{m}$; Nd:YAG: $1,06 \mu\text{m}$; CO_2 : $10,6 \mu\text{m}$): Die reflektierte Laserleistung ist ein Maß für die dem Prozeß verloren gehende Energie.

All diese Signale und Prozeßanteile besitzen eine verschiedene zeitliche Dynamik, die nicht vergessen werden sollte. Das bedeutet, daß mit langsamen Sensoren wie einer CCD-Videokamera z.B. schnelle Vorgänge wie die Bewegung der Kapillare oder Schmelzbadwellen nicht aufgelöst werden können. Dagegen reicht die Standard-Video-Frequenz von 50 Hz durchaus zur Überwachung der Nahtlage. Um schnellere Prozesse, wie etwa die von Schmelzauswürfen in Form von Spritzern, beobachten zu können, sind Bildsensoren mit einer wesentliche höheren zeitlichen Dynamik notwendig.

4.4 CMOS-Kameratechnik

Als bildgebender Flächensensor war der CCD-Sensor (**C**harged **C**oupled **D**evice) stets das Werkzeug der ersten Wahl und ist zur Zeit am weitesten verbreitet. Entlang seiner Entwicklung entstanden aus seinen typischen technischen Randbedingungen die heute bekannten und meist nur noch historisch zu begründenden Bilderfassungssysteme, Videonormen, Bildraten und Bildgrößen. Vereinfacht läßt sich das Funktionsprinzip dieses Sensors mit der Solarzelle vergleichen: Jedes Pixel ist mit einer Photodiode ausgestattet, die jeweils parallel mit einem Kondensator verschaltet ist, der vor der Belichtung mit einer einheitlichen Spannung aufgeladen wird. Fällt Licht auf die Photodiode, wird diese je nach einwirkender Lichtmenge mehr oder weniger elektrisch leitend. Auf diese Weise entlädt

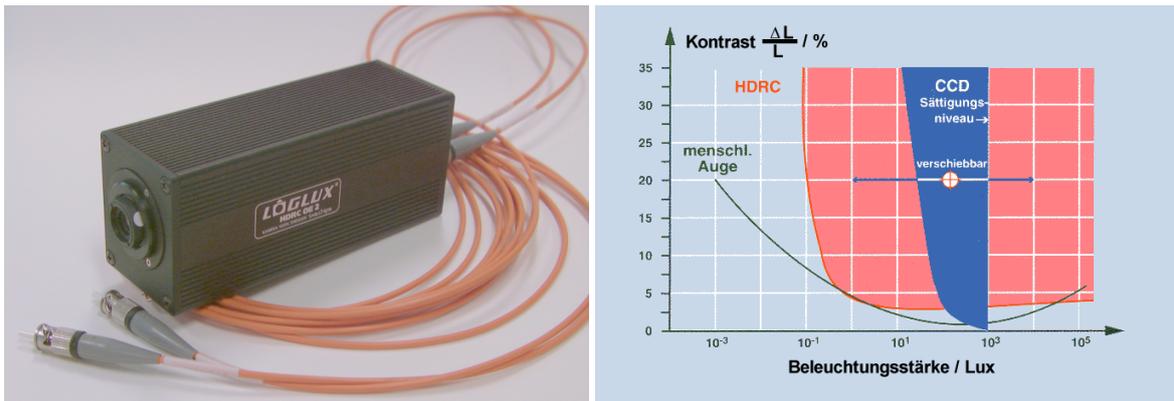


Abbildung 4.3: Links: HDRC-Kamera (High-Dynamic-Range-CMOS) mit den beiden Lichtwellenleitern und einem Normalobjektiv (Gehäuse: $10\text{ cm} \times 5\text{ cm} \times 4\text{ cm}$, später auch als 5^3 cm^3 -Gehäuse verfügbar) Rechts: Der Dynamikbereich des HDRC-Sensors initiiert mit der logarithmische Kennlinie biologische Vorbilder

sich der Kondensator proportional zur Belichtung. Die Ladungen werden nun im positiv geladenen Silizium-Kristall des Pixels in Potentialtöpfen gespeichert. Wenig Lichteinfall bewirkt eine geringe Ladung, viel Lichteinfall erzeugt eine hohe Ladung. Nach einer fest definierten Belichtungszeit werden die Potentialtöpfe mit der in ihnen enthaltenen Ladungsmenge durch Ansteuern der Kondensatorgruppen nach dem Eimerketten-Prinzip (engl. *bucket-chain*) von Sensorzelle zu Sensorzelle zum Zeilenende durchgereicht. Dort werden die Ladungen per Taktsignal “herausgeschoben” und vom Ausleseregister übernommen. Auf diese Weise wird der CCD-Sensor Zeile für Zeile ausgelesen. Aus dem Photostrom wurde also eine Ladungsmenge, die als serielles analoges Signal anschließend mit einem A/D-Wandler als weiterem erforderlichen Bauelement weiterverarbeitet werden kann.

In den letzten Jahren tauchte als bildgebender Sensor ein neuer Typ, der CMOS-Sensor, auf. Unter CMOS (Complementary Metal Oxide Semiconductor) versteht man eigentlich keinen Sensor, sondern das Aufbauprinzip für einen Baustein und dessen Herstellungstechnologie für immer kleinere integrierte Schaltkreise (IC) aus Transistoren. Die so hergestellten Chips finden sich zum Beispiel als Speicherbausteine in allen PCs. Heute ist CMOS Technologie die dominante Halbleitertechnologie (90% aller ICs) für Mikroprozessoren, Speicher (SRAM und DRAM) und anwenderspezifische integrierte Schaltungen (ASICs). Hauptvorteil von CMOS gegenüber bipolarer Technologie ist der viel kleinere Energieverbrauch: statische CMOS Gatter haben (fast) keinen statischen Energieverbrauch. Die Gatter verbrauchen nur Strom, wenn sie wirklich umschalten. Dadurch lassen sich wesentlich mehr CMOS Gatter auf einem IC integrieren als bei NMOS oder bipolarer Technologie. Die Nutzung der CMOS Technologie für Bildsensoren ist seit langem bekannt (die Zeitschrift Elektor veröffentlichte 1982 eine Bastelanleitung, um aus DRAM-Speicherbänken Bildsensoren herzustellen). Dies wurde erst in den 90ern mit dem Voranschreiten der DRAM Technologie in Produkte umgesetzt, da letztere der Motor für immer kleinere Transistoren und immer reineres Silizium ist. Erst dadurch konnten Aktiv-Pixel-Strukturen, die prinzipiell mehr als drei Transistoren brauchen, erfolgreich in die Sensoren integriert werden. Auch hier wird mit Hilfe vieler, in Form eines Rasters angeordneter Photodioden bei Lichteinwirkung ein Photostrom erzeugt. Dabei ist ein Widerstand mit der Photodi-

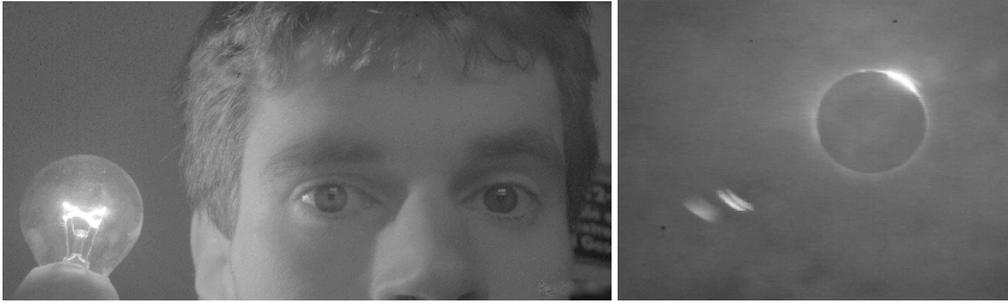


Abbildung 4.4: Links: Mit der logarithmischen Kennlinie lassen sich hell leuchtende Objekte und im Schatten liegende Bereiche gleichzeitig abbilden Rechts: Sonnenfinsternis vom 11. August 1999 über Europa mit der HDRC-Kamera aufgezeichnet

ode in Reihe geschaltet und die Spannung wird über diesem Widerstand gemessen. Diese Messung kann kontinuierlich erfolgen, braucht also keine Integrationszeiten oder Auslese-Totzeit wie die des CCD-Sensors. In diesem Sinne sammelt der CMOS-Sensor auch keine Ladungsträger an, weil die Konversion von Ladungsträgern in eine Spannung für jeden Pixel individuelle dort vor Ort stattfindet; beim CCD-Sensor findet dies jedoch in einem allen Pixeln gemeinsamen Ausgabestruktur am Ende der Eimer-Kette statt.

Für den Massenmarkt sehr wichtig sind Stückzahlen und Preis. CCD-Sensoren werden auf sehr speziellen Fertigungsstraßen, von denen es weltweit nur rund ein Dutzend gibt, hergestellt. CMOS-Sensoren können dagegen wie die nicht-bildgebenden ICs mit billigen Rohstoffen auf Standard-Fertigungsstraßen hergestellt werden. Abnehmende Strukturbreiten der Halbleiterindustrie schaffen für die bildgebenden Bausteine jedoch vielfältige Probleme durch Übersprecheffekte, Resonanzen und thermische Effekte, daß üblicherweise eher zurückhaltend auf 0,25–0,35 μm -Technologie produziert wird. Die kleinsten bekannten CMOS-Pixel sind in der Ausdehnung mit 3,3 μm (auf Basis des 0,25 μm Prozesses) rund ein Drittel größer als CCD-Pixel. CMOS-Pixel haben auch eher ungünstige Füllfaktoren (optisch sensitive Fläche im Vergleich zur Gesamtfläche des Pixel, belegt durch Verstärkertransistoren und Verschaltung) von typischerweise rund 25%-40%; dem kann ähnlich wie CCDs mit Mikrolinsensfeldern oder bei CMOS-Sensoren mit TFA (Thin Film on ASIC) [Schneider, Bernd u. a., 1999] entgegengewirkt werden. Zu den Technologien CCD und CMOS, den Unterschieden und einer Bewertung siehe auch [Litwiller, 2001] und [Braggins, Don und Schibli, Nikolaus, 2001].

CCD-Sensoren zeigen hauptsächlich zwei Störeffekte. Unter *Blooming* versteht man eine unerwünschte Überstrahlung im Bild. Muß ein Element des CCD-Sensors bedingt durch das Motiv sehr viel Licht aufnehmen, entsteht dort eine zu hohe elektrische Ladung, die dann auf benachbarte Elemente überspringt. Die ausgelesene Bildinformation ist dann immer weiß ohne weitere Zeichnung (Differenzierung). Bei CMOS-Chips tritt gar kein Blooming auf, da hier die einzelnen Zellen adressiert werden und unabhängig voneinander sind: überschüssige Ladungen können über den pixel-individuellen Reset-Transistor abfließen. Ein ähnlicher Bildfehler ist der *Smear*-Effekt. Er ist eine Bildstörung, die bei geringen Beleuchtungsstärken bedingt durch den Auslesevorgang von CCD-Chips auftritt. Sie stellt sich als senkrechter heller Streifen dar, der von hellen punktförmigen Lichtquellen ausgeht. Verursacht wird dieser Fehler auch durch überschüssige Ladungen, die sich beim

zeilenweisen Auslesen über benachbarte Elemente ausbreiten. CMOS-Sensoren zeigen im Gegenzug dazu ein höheres *fixed pattern noise*: ein latent vorhandenes Bild trotz räumlich homogener Beleuchtung. Diese Ungleichmäßigkeit in der Photonen-Spannungs-Konversion ist durch den CMOS-Prozeß bedingt. Diese Effekte verstärken sich mit weiterer Miniaturisierung der Halbleiterstrukturen. Üblicherweise wird durch Offset-Einstellungen der Umsetzer in den einzelnen Pixeln eine für jeden Chip individuelle Korrektur durchgeführt.

Während beim CCD-Sensor nur ein serielles zeilenweises Auslesen eines gesamten Bildes möglich ist, kann das Auslesen beim CMOS-Sensors wie bei Speicherbausteinen wahlfrei erfolgen, also durch direkte Adressierung der einzelnen Sensorelemente zu jedem beliebigen Zeitpunkt und in beliebiger Abfolge. Das wird durch eine Adressierung wie bei Speichern erreicht. Diese hat eine schaltungstechnisch bedingte Pixelrate, die letztlich festlegt, wie schnell Bilder ausgelesen werden können: Große Bilder haben eine kleine Bildrate, kleine Ausschnitte von Bildern werden entsprechend schneller ausgelesen. Typischerweise ist die Bildrate (in fps, *frames per second*) mit Offset-Werten der Synchronisationszeiten (umgerechnet in Pixelanzahlen) für kleine Bildausschnitte begrenzt

$$f_{\text{frame}} = \frac{f_{\text{pixel}}}{(\Delta x + n_{\text{hsync}})\Delta y + n_{\text{vsync}}}$$

Pixelrate:	$f_{\text{pixel}} = 8 \text{ M Pixel} \cdot \text{Hz}$
Zeilensynchronisation (<i>line-enable</i>):	$n_{\text{vsync}} = 32 \text{ Pixel}$
Bildsynchronisation (<i>frame-enable</i>):	$n_{\text{hsync}} = 32 \text{ Pixel}$

Für ein 64×64 Bild ergibt sich beispielsweise beim HDRC-Sensor eine Bildrate von 1295 fps (ohne Synchronisationsoffsets wären es 1953 fps; die Offsets machen bei solch kleinen Ausschnitten bereits 33% der Datenübertragung aus).

Der CMOS-Sensor ist prädestiniert für die Integration weiterer Intelligenz (Single-Chip-Kamera, Smart-Pixel) in Form von festen Schaltungen, die direkt pixelweise Vorverarbeitung der Bildverarbeitung (engl. *early vision*) analog oder digital übernehmen. Mit speziellen Grauwert-Konversionen aus Look-up-Tabellen, Gain- und Offsetanpassungen, Schwellwert-Operationen oder Binarisierungen, Lauflängenkodierungen, Berechnungen von Differenzbildern oder Ableitungen bis hin zum kompletten DSP ist damit der Weg zur "intelligenten Kamera" und der "on-chip" Bildverarbeitung technisch möglich und umgeht die für den CCD-Sensor historisch bedingte Umsetzung auf analoge Signale und deren Quantisierung im Framegrabber. Beispiele dafür in [Hauschild, 1999] und [Schanz, 1998].

Die CMOS-Kamera wird für den PC zu einem externen Speicherbaustein, so daß für Hochgeschwindigkeitsanwendungen heute vor allem die Bus-Rate des PCs der Flaschenhals ist. Moderne Übertragungsprotokolle wie CameraLink (bis 2,38 Gbit/s) und FireWire setzen hier an. Auch aus der Wahl größerer Strukturgrößen bei der Herstellung photosensitiver CMOS-Sensoren gegenüber CMOS-Chips allgemein erwächst das Problem, daß aktuelle DSP-Cores und RISC-Prozessoren (auf kleine Strukturen optimiert) ihre Vorteile nicht voll ausspielen können. Voraussetzung für aufwendige Bildverarbeitung oder gar Bildfolgenverarbeitung auf dem Chip sind ausreichend bemessene Speicherregister. Ein sehr interessantes und hochdynamisches Feld ist dabei der Einsatz von FPGA-Bausteinen (**F**ield **P**rogrammable **G**ate **A**rrays): rekonfigurierbarer Logik, die gegenüber in Hardware festverdrahteter Intelligenz (ASICs) den Vorteil hat, wiederholt modifizierbar zu sein [Noffz, K.-H. u. a., 1999]. So wurde im Rahmen des Open-Eye-Projektes auf der Industriemesse VISION2001 eine FPGA-Realisierung des Strukturensors [Jähne, 1993] vorgestellt. Während es nur eine Frage der Zeit ist, bis ausreichend große Speicherregister für höhere Bildverarbeitung standardmäßig für FPGAs

bereitstehen, verhindert das Problem der äußerst komplizierten Programmierung und Simulation von FPGAs (derzeit mit den Sprachen VHCL und DHCL), das gleichzeitig ein tiefes Verständnis für Softwaretechnik und Schaltungsdesign verlangt, eine größere Verbreitung in ingenieurtechnische Bereiche.

CMOS-Sensoren, die nicht integrieren wie CCD-Sensoren, leiden an einem hohen Rauschen: das SNR (engl. *Signal-to-Noise-Ratio*) ist erheblich höher ($56 e^-$ für Sensoren nach HDRC-Photodiodenaufbau bzw. $13 e^-$ für Active-Pixel-Sensors) als bei CCD-Sensoren, die teilweise nur 5 bis 20 Rauschelektronen bei einer Sättigung von 100000 bis 500000 Elektronen haben. CMOS-Sensoren haben einen höheren Dunkelstrom (stark temperaturabhängig) als CCD-Sensoren. Allein aus der Bitbreite bei der Digitalisierung des Signals ergibt sich ein SNR (auch Quantisierungsrauschen)

$$SNR_Q = 20 \log 2^{\text{Bittiefe}} \text{ dB}$$

das zum Beispiel für 10 bit tiefe Signale (1024 Graustufen) 60, 21 dB beträgt. Hinzu kommt der bei CMOS-Sensoren im Vergleich zu CCDs häufigere Totalausfall von einzelnen Pixeln.

CMOS-Sensoren haben einen steigenden Quantenwirkungsgrad [Seiffert, 2001] im nahen Infrarot (700–950 nm) und erreichen typischerweise bei 700–800 nm ihre maximale Empfindlichkeit. sie sind also in gewisser Weise eine Kombination aus thermographischer Kamera und visueller Kamera.

CMOS-Sensoren nehmen nur rund $1/3 \dots 1/10$ der Leistung eines CCD-Sensors auf, da sie schaltungstechnisch keine zur Eimerkette der CCD vergleichbaren Verbraucher haben. Zusammen mit dem direkt digital vorliegenden Ausgangssignal prädestiniert sie dies für mobile oder miniaturisierte Anwendungen. Durch die Fertigung auf Standardprozessen und dafür bereitstehenden großen Kapazitäten ist der CMOS-Sensor in der Regel billiger, wenn er in großer Stückzahl als Massenprodukt hergestellt wird. Diese Skaleneffekte greifen natürlich nicht für spezielle, für bildgebende ICs optimierte Jobs mit naturgemäß kleineren Stückzahlen für Abnehmer in der industriellen Bildverarbeitung, die in keiner Weise vergleichbar sind mit den Massenmärkten der Unterhaltungselektronik.

Normale CCD und CMOS Kameras erreichen eine Dynamik von 50 – 70 dB. Zum Vergleich: zwischen vollem Sonnenschein und den Lichtverhältnissen einer mondlosen Nacht liegen ≈ 170 dB, der professionelle fotografische Film kann mit Blendeneinstellungen (Blenden von 32 bis 1, 8, also 8 Stufen) und Verschußeinstellungen ($1/30$ s bis $1/4000$ s, also 7 Stufen) einen Bereich von 90 dB abbilden. Das menschliche Auge umfaßt rund 200 dB. Zur Erhöhung (Abbildung 4.3) der Dynamik (> 120 dB) gibt es verschiedene Ansätze: Verrechnung verschiedener Integrationszeiten [Robertson, Mark A. u. a., 2000] bzw. Mehrfachbelichtungen [Chen, Z. und Mu, G., 1995] oder Empfindlichkeiten (Chips vom IMS Fraunhofer Duisburg [Schrey, 2001] [Schanz, 1998]) oder logarithmisch integrierende Sensoren (Chips des IMS Stuttgart [Seiger, Ulrich u. a., 1999] [Graf, Heinz-Gerd u. a., 1995] [Höfflinger, 1999]), sowie schaltungstechnisch realisierte Mischformen linearer und logarithmischer Kennlinien (Photonfokus AG, [Seitz, 1999]). Die rein logarithmischen Sensoren wandeln dabei durch einen MOS-FET mit logarithmischer Kennlinie den Strom der Fotodiode in eine Spannung um. Prototypen einer logarithmischen CMOS-Kamera mit adaptiven Pixeln (Belichtungszeit und Kalibration gegen fixed-pattern-noise) wurden auch von [Loose, Markus u. a., 1996] [Loose, Markus u. a., 1998] [Loose, 1999] gebaut. Für den industriellen Einsatz gibt es heute eine bereits kaum mehr zu überblickende Vielfalt linearer CMOS-Kameras oder Hochgeschwindigkeitskameras (vielfach CMOS, Probleme mit

Eigenschaft	CCD-Sensor	CMOS-Sensor
Prozeß	ausgereift verlangt eigenen Herstellungsprozeß	Technologie für Bildsensoren noch nicht ausgereizt, große bereitstehende Kapazitäten aus Halbleiterherstellung
Rauschen	sehr niedrig	hoch
Taktrate	160 MPixel/s	1 GPixel/s
Zugriff	Vollbild	wahlfreie Ausschnitte (random access)
Leistungsaufnahme	3–6 W	< 0,1 W
Ausgänge	analog braucht AD-Wandler	bis zu 16 Digitalausgänge mit 10 bit
Quantenrauschen (N_{e^-} - Rauschen)	bis $100 e^-$ $3 e^-$ (Mikroskopie, Astronomie)	$10 \dots 150 e^-$
Dunkelstrom bei 25°	$2 \dots 50 \text{ pA/cm}^2$	$100 \dots 1000 \text{ pA/cm}^2$
DSNU dark signal	< 1% V_{sat}	10% V_{sat} (on Chip eliminierbar)

Tabelle 4.1: Vergleich von CCD-Sensoren und CMOS-Sensoren

der Datenrate werden durch kameranahe Zwischenspeicher gelöst, deren Größe die insgesamt aufgezeichnete Pixelzahl begrenzt). Unter den CMOS-Kameras mit (teilweise) logarithmischer Kennlinie war Ende 1999 nur eine voll einsatzfähige und für industrielle Endkunden verlässlich verfügbare Kamera in der Lage, die volle Bildratensteigerung bei kleineren Ausschnitten zu bieten.

CMOS Sensoren befinden sich erst am Anfang ihres Weges. Zusammenfassend (Tabelle 4.1) lauten die Vorteile und Nachteile der CMOS-Sensoren: integrierbar, geringe Leistungsaufnahme, erlaubt durch Ausschnittwahl hohe Bildraten, digital konfigurierbar, kompakt durch direkten digitalen Ausgang, hohes Rauschen, kein Blooming, kein Smearing. Im Rahmen dieser Arbeit kann der seit einigen Jahren entbrannte Streit um den "besten" Sensor nicht entschieden werden. Beide Technologien werden mittelfristig nebeneinander existieren und ihre Vorteile in bestimmten Anwendungen ausspielen können. Für die Wahl des in Abschnitt 4.8 verwendeten Bildsensors und mögliche Alternativen wurden intensive Studien an einem eigens dafür konstruierten Teststand bezüglich der zeitlichen und örtlichen Auflösungseigenschaften bei verschiedener Helligkeitsstufen durchgeführt. Die detaillierten Ergebnisse dazu finden sich in den beiden Diplomarbeiten [Seiffert, 2001] [Sax, 2001].

4.5 Aufgabenstellung und Prozeß

In der konkret vorliegenden Aufgabenstellung soll die Qualitätssicherung einer Schweißnaht für ein Produkt aus dem Automobilbereich verbessert werden. Es handelt sich um eine axiale verschweißte Kehlnaht am Überlappstoß, mit der eine mechanisch grob fixierte Scheibe einen Hohlraum, der im Gebrauch unter hohen Drucken stehen wird, abdichtet. Die Spritzlochscheibe soll an einem Kraftstoffeinspritzventil (Abbildung 4.5) befestigt werden. Jährlich werden weltweit 200 Millionen Stück dieses Produktes vor allem in der Erstausrüstung von Kraftfahrzeugen verbaut. Als Ersatzteil kostet

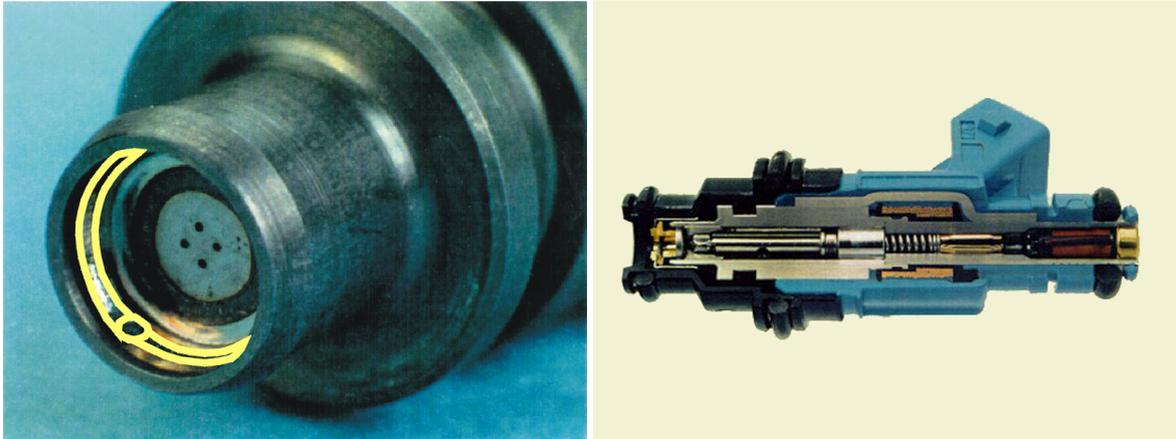
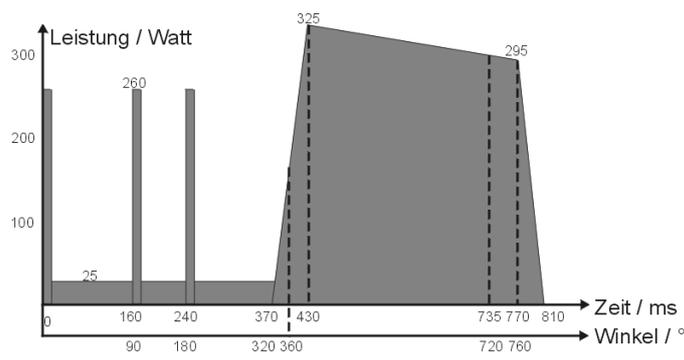


Abbildung 4.5: Links: Am unteren Ende befindet sich eine Spritzlochscheibe (typischerweise mit vier oder fünf Löchern), die mit einer Rundschweißung im hülsenförmigen Metallkörper fixiert ist. Die für die hier beschriebene Anwendung relevante Naht und ein darin vorhandenes Loch sind markiert. Die Metallhülse hat einen Innendurchmesser von ca. 6,5 mm Rechts: Das gesamte Einspritzventil fertig montiert und aufgeschnitten (Länge ca. 6 cm). Die Spritzlochscheibe befindet sich hier an der linken Seite des Bauteils.

es für einen gängigen Typ einzeln 30€ . Allein an der in dieser Arbeit involvierten Produktionslinie werden rund 150 verschiedene Teilettypen hergestellt, die sich zum Beispiel in ihrer äußeren Bauform und Geometrie, der Dicke der Spritzlochscheibe, verwendeten Metalloberflächen und dergleichen mehr unterscheiden. Insgesamt wurden an dem hier involvierten Produktionsstandort (weltweit gibt es ungefähr 10 Produktionsstandorte) seit 1992 bis 2002 rund 80 Millionen dieser Bauteile in Großserienfertigung hergestellt. Der in dieser Arbeit untersuchte Prozeß auf dem gewählten Produkt wird derzeit im Jahr rund 14 Millionen mal ausgeführt.

Die Schweißung wird als Wärmeleitungsschweißung mit einem durchgängig (*continuous wave, cw*) betriebenen Haas Laser (500 W , 0,6 mm Faser) Nd:YAG-Laser ausgeführt, unter dem das Werkstück um seine Achse um $\approx 800^\circ$ rotiert. Um ein Verziehen der Metallscheibe am Beginn der kontinuierlichen Nahtschweißung zu vermeiden, werden vor der eigentlichen Schweißnaht drei Heftpunkte im Winkelabstand von 90° gesetzt (Abbildung 4.6), um die Scheibe auch gegen große Verspannungskräfte zu fixieren. Diese drei Heftpunkte werden später beim Setzen der kontinuierlichen Naht der Länge 425° überschweißt. Auf Schutzgas oder Lotzufuhr wird verzichtet. Der Laser schaut unter 45° auf das Werkstück (Abbildung 4.9). Alle relevanten Größen der Schweißnaht faßt die Tabelle 4.6 zusammen.

Bei diesem Bauteil ist vor allem ein sauberer Prozeß gefordert, der einerseits glatte, gleichmäßige und dichte Nähte hinterläßt und außenseitig an der Spritzlochscheibe keine losen Metallrückstände hinterläßt. Trotz der Wahl geeigneter Parameter (Energie, Fokusslage, Fokusgeometrie, Oberflächenbeschaffenheit) kommt es bei den Bauteilen während des Schweißens zu eruptiv austretenden Spritzern. Dabei wird flüssiges Metall aus dem Schmelzbad herausgeschleudert. Schmelzspritzer, die beispielsweise aufgrund von Materialinhomogenitäten entstehen können, lagern sich als kleine Metallkugeln an benachbarten Oberflächen an oder hinterlassen ab einer gewissen Größe ein Loch in der Schweißnaht (Abbildung 4.8). Die Metallkugeln, die $10 \dots 100 \mu\text{m}$ Durchmesser haben, lösen sich später



Nahtbreite	0,7 mm
Schweißtiefe	0,4 mm
Gesamtenergie	117 J
Überschweißung	65°
Intensität	1830 W/cm ²
Scheibenumfang	20,5 mm
Schweißweg	24,2 mm
Umdrehungszeit	320 ms
Drehgeschwindigkeit	6,4 cm/s

Abbildung 4.6: Typischerweise benutzte Laserpulsform für 0,2 mm Materialdicke. Während der Werkstückträger anfangs auf seine konstante Rotationsgeschwindigkeit beschleunigt, werden drei 10 ms breite Laserpulse gesetzt, die die Scheibe in der Hülse gegen Verzug (vor allem zu Beginn der kontinuierlichen Schweißung) fixieren. Gleichzeitig findet ein Vormattieren der Naht bei schwacher Energie statt. Die Heftpunkte haben später einen Abstand von 90° untereinander und 180° vom Beginn der kontinuierlichen Schweißung. Diese rampt schnell auf und ab und hat eine leicht fallende Leistung, um die bereits in das Material eingebrachte Wärme zu kompensieren. Die Naht wird mit einem Überlapp von $\approx 65^\circ$ (Gesamtlänge also 425°) geschweißt.

unter der Betriebsbelastung und können zu Folgeschäden an den Dichtflächen der Ventile führen. Die Löcher und Poren in der Naht stellen unter Betriebsbelastung mit Drucken von 110 bar eine potentielle Leckstelle der Naht dar oder sind bereits von Anfang an so undicht, daß durch sie Kraftstoff unkontrolliert austreten kann. Nicht zu jedem Loch läßt sich eine Ablagerung finden. Es ist nicht ausgeschlossen, daß es auch Spritzer gibt, die kein von außen sichtbares Loch hinterlassen, da die Naht nach der Eruption noch flüssig genug sein kann, sich zumindest oberflächlich wieder zu schließen und dadurch einen in der Naht liegenden Hohlraum von einem Durchmesser von rund 0,2 mm bilden kann. Ein solcher Defekt läßt sich naturgemäß mit einer Sichtprüfung nicht entdecken. Eine zerstörungsfreie Prüfung zum Beispiel mit computertomographischen Verfahren [Eisele, 2002] bei allen Teilen ist in der geforderten Auflösung, Taktzeit und Genauigkeit in der Massenproduktion heute nicht möglich.

Spritzervorkommen für jedes individuelle Teil vorher zu sagen, ist ohne die Kenntnis der tieferen Ursachen heute noch nicht möglich und wahrscheinlich auch künftig wegen des chaotischen Verhaltens der flüssigen Schmelze nicht realisierbar. Die turbulenten Eigenschaften der flüssigen Schmelze und der an sich unruhige Prozeß des Schweißens lassen eine individuelle Vorhersage unmöglich erscheinen, zumal noch grundlegendes Verständnis für die Rahmenbedingungen fehlt, unter denen solche Ereignisse häufiger oder seltener auftreten. Saubere Oberflächen, glatte Stoßkanten, ein kleiner Fügespalt, ruhige und gleichmäßige Führung, korrekte Fokuslage sind sicher Mindestvoraussetzung. Die Vorgänge während des Schweißprozesses selbst oder gar ein Blick in die räumliche oder zeitliche Struktur des Schmelzbades und seiner Dynamik entzogen sich jedoch bisher der systematischen Beobachtung und Überwachung.

4.6 Industrielle Randbedingungen und Anforderungen

Zur Qualitätskontrolle der beschriebenen Schweißnaht wurde an den fünf Produktionslinien bis Mitte 2000 eine Sichtprüfung aller gefertigten Bauteile am Ende der Montage im Dreischicht-Betrieb

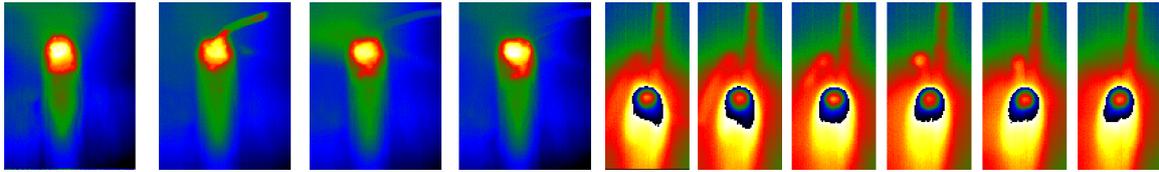


Abbildung 4.7: Falschfarbaufnahmen (10 bit Tiefe) im zeitlichen Abstand von 8 ms von einer Laserschweißung (Wärmeleitungsschweißen) auf einem glatten Stahl, der aber durch Vorbehandlung (Verschmutzung, Oxidation, Oberflächenunebenheiten) sehr anfällig für Spritzer gemacht wurde. Diese Aufnahmen entstanden im Rahmen von Voruntersuchungen zur Fragestellung, ob sich Prozeßunruhe und Spritzer beim Laserschweißen mit bildgebenden Verfahren überhaupt erfassen lassen können. **Links:** Ein Spritzer löst sich rechts oben im Bild deutlich sichtbar aus der Schmelze. In Folgebildern zieht er eine helle Spur hinter sich nach. Im dritten Bild ist auch eine eruptive, großflächige Aufhellung (links oben) durch Metaldampf sichtbar. Solche Aufhellungen kommen häufig vor und stehen nicht unbedingt im Zusammenhang mit Spritzern. **Rechts:** Bei der Verschweißung eines zu weit offenstehenden Spaltes (erscheint warm) mit eingebrachter Verschmutzung löst sich deutlich sichtbar aus der Schmelze ein Spritzer.

durchgeführt. Je nach Typ wurden bei $\approx 0,10/00$ der Teile die Fehler “Loch” oder “Metallablagerung” gefunden. Solche Teile wurden ausgeschleust, gegebenenfalls nachuntersucht und schließlich verschrottet. Eine Nacharbeit der Nähte ist nicht wirtschaftlich. Insgesamt stellen mit rund 50 % die Fehler der Laserschweißverbindungstechnik die Mehrheit aller am Produkt auftretenden Fehler dar, die zu einer Ausschleusung bei der Endsichtprüfung führen. Der Fertigungstakt beträgt 5 s, die Produktion läuft bis auf Umrüstungen auf andere Teiletypen ununterbrochen an sechs von sieben Tagen. Eine einzelne Produktionslinie mit ihren rund 20 Stationen für die verschiedenen Montageschritte wird im laufenden Betrieb und bei Umrüstungen von einer Person betreut. Ohne bisher erkennbaren oder quantifizierbaren Grund zeigen verschiedene Teiletypen, aber auch unterschiedliche Chargen, sehr ungleiche Fehlerraten. Daraus ergab sich die Anforderung, eine Überwachung zu entwerfen, die jeden Prozeß prüft, ohne dabei auf “angelerntes” Wissen aus vergangenen Prozessen zurückzugreifen. Auch wenn diese Vermeidung einer Statistik über die Dimension n viel mögliche Information verschenkt, so sind doch alle bisher eingesetzten Methoden für die Prozeßüberwachung in diesem Fall gerade daran gescheitert, daß die Teiletypen zwar innerhalb ihrer Chargen sehr homogene Fehlerraten hatten, Chargen oder Typen untereinander aber starke Sprünge zeigten.

Eine minimal undichte Naht oder eine kleine Verschmutzung an der zur Brennkammer offenen Außenseite der Spritzlochscheibe kann einen plötzlichen Totalausfall und irreparablen Motorschaden verursachen. Daher sind die Kundenanforderungen für diese produktsicherheitsrelevante Laserschweißung sehr hoch: von der von Fachleuten anfangs für realistisch gehaltenen Marke von 60 ppm (parts per million, 10^{-6}) Fehlerteilen im Feld wurde in der revolutionären Verschärfung der Qualitätsanforderungen Ende der 80er Jahre abgerückt, ein Zielwert von 15 ppm gesetzt und in den Folgejahren erreicht. Mitte 2000 lag die Fehlerrate bei nur noch 2 ppm, womit die 0 ppm-Marke statistisch nicht mehr nachweisbarer Fehler in greifbare Nähe gerückt ist. Solche Anforderungen lassen sich bei gleichzeitigem Kostendruck, just-in-time-Produktionen, weiterhin hoher Typteile-Vielfalt und fortlaufender Verbesserungen an den Herstellungsschritten nur erreichen, wenn anstelle einer reinen Nachinspektion auch die Ursachen von Fehlern konsequent angegangen werden. Die sichere Beherrschung

solcher Verfahren wird mehr und mehr Verkaufsargument gegenüber den Kunden.

Es wird also eine Ergänzung oder ein Ersatz der Sichtprüfung durch ein technisches, reproduzierbares Verfahren gesucht, das direkt den Prozeß und damit die Entstehung von möglicher Minderqualität beobachtet und Wege zu einem Ursachenverständnis öffnet. Eine solche Prozeßüberwachung und Qualitätsaussage muß in der Taktzeit abgeschlossen sein, muß voll automatisiert in die vorhandene Montagestation integriert sein, darf keinen oder nur geringsten (in jedem Fall automatisierbaren) Umrüstaufwand haben, benötigt eine Selbstüberprüfung und einfache Bedienelemente (Sensorik und Bedienoberfläche) und muß die gestellte Aufgabe in dem hier geschilderten Fall, die Schweißnaht auf Poren oder abgesetzte Metallperlen zu prüfen, mindestens so gut bewältigen wie die bisher eingesetzte Sichtprüfung.

4.7 Lösungsansatz für die Qualitätssicherung

Die vorgestellte Anwendung eignet sich ideal für den Einsatz einer Prozeßüberwachung mittels Bildfolgenauswertung mit einem Sensor, der hohe Bildraten und einen hohen Dynamikbereich bietet. Mit dem raumzeitlich aufgelöstem Blick in den Prozeß wären erstmals Teile-individuelle Daten zugänglich, die ein grundsätzliches Verständnis der im Prozeß ablaufenden Vorgänge ermöglichen und daher den Weg weisen zu einem besseren Grundlagenverständnis der Entstehung von Spritzern und Defekten beim Laserschweißen. Der Einsatz in der Fertigung hoher Stückzahlen erlaubt darüber hinaus auch den Zugang zu einer Statistik mit umfangreicher Datengrundlage über verschiedene Bauteile, Schweißparameter, Produktchargen und längere Zeiträume an mehreren unterschiedlichen Produktionslinien, die in diesem Umfang bei rund 1 ‰ Fehlern im Laborversuch niemals zu erheben gewesen wären.

Der bildgebende Sensor für diese Ausgabenstellung muß mit der enormen Helligkeitsdynamik beim Blick in die Laserschweißung zurechtkommen und Bilder in hoher Rate liefern. Die vorgestellte HDRC-Sensorik stellt erstmals einen solchen Sensor für den industriellen Einsatz zu Verfügung und erschließt damit einen neuen Aufgabenbereich der Bildsequenzverarbeitung. Die logarithmische Kennlinie nach biologischem Vorbild und der wahlfreie Zugriff auf kleine Bildausschnitte mit der damit verbundenen Erhöhung der Bildrate sind die Schlüssel zum räumlich-zeitlich aufgelöstem Zugang zum sichtbaren und nahem IR-Spektrum des Schweißprozesses.

Bei der gestellten Aufgabe wurde bisher die Qualität durch eine vollständige Sichtkontrolle der betroffenen Schweißnähte am fertigen Produkt durchgeführt. Der personelle Aufwand bei pro Produktionslinie größenordnungsmäßig 10^4 gefertigten Teilen pro Tag ist enorm, wobei allgemein davon auszugehen ist, daß bei einer von erfahrenem und geschultem Personal durchgeführten Sichtprüfung weiterhin nur 96% der tatsächlich vorhandenen Fehler gefunden werden (Schlupf).

Mit einer Bildfolgenanalyse wird nicht direkt nach den beiden gesuchten qualitätsmindernden Merkmalen der Schweißnaht ("Pore" und "Metallablagerung") gesucht, sondern nach Eruptionen und aus der Schmelze entweichenden Spritzern. Es ist dabei der Nachweis zu erbringen, daß durch die Detektion von Spritzern tatsächlich alle Poren und Metallablagerungen erfaßt werden.

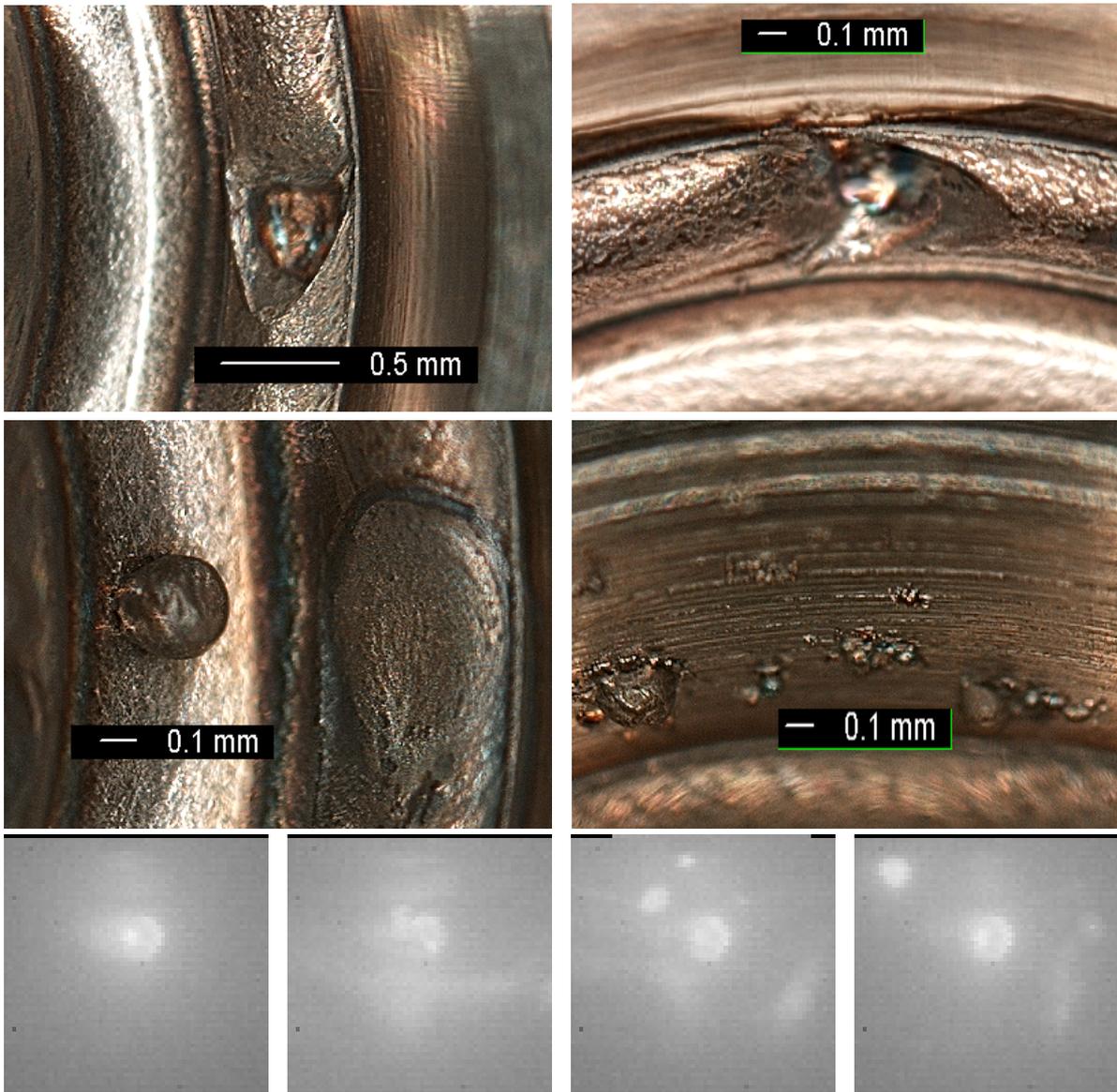


Abbildung 4.8: **Oben:** Löcher in der Schweißnaht haben meist Dreiecksform, die sich aus dem Tiefenverlauf (Abbildung 4.2) des flüssigen Metalles ergibt. Die vorderen unregelmäßigen Ränder führen dabei in der Regel sehr steil in die Naht und in einen verrundeten Hohlraum, der durch die hier verwendete Aufnahmetechnik bläulich schimmert (man beachte auch Abbildung 4.24). Der Laser koppelt nach dem eruptiven Auswurf neu in das Metall ein und bildet erneut einen Ort flüssigen Metalles, der durch das Material wandert. Davon ist unmittelbar nach dem Einkoppeln auch die Lochvorderseite betroffen: erneutes Aufschmelzen ergibt eine verrundete Kante. Die rückseitigen Kanten des Loches haben meist deutliche Kanten: hier ist die gesamte Schmelze ausgeworfen, der Rand des Schmelzbades nach ihrem Auswurf erstarrt, und nicht mehr weiter vom Laserlicht getroffen oder überschweißt worden. Löcher zeigen oft auch Nahteinschnürungen. **Links unten:** Ein Metallspritzer hat sich unmittelbar neben dem Loch wieder abgesetzt. Spritzer bilden meist dunkle Metallperlen aus, die fest im Innenraum der Hülse anhaften. Rechts neben dem Spritzer das dazugehörige Loch, das im Überschweißbereich gelegen vom Laser abermals aufgeschmolzen wurde und von einer dunkel berandeten und von der sonst üblichen Nahtoberfläche deutlich zu unterscheidenden Metallschicht bedeckt wird. Zufälligerweise fällt diese Stelle mit dem Endkrater der Schweißnaht zusammen. Es ist recht wahrscheinlich, daß unter der Nahtoberfläche Hohlräume vorhanden sind. **Rechts unten:** Auch nach längerem Flug (hier bis zur gegenüberliegenden Innenseite der Hülse) sind Spritzer noch heiß und damit flüssig genug, um sich weitgehend mit dem Metall, auf das sie auftreffen, zu vereinen und sehr fest anzuhaften. **Untere Reihe:** Anhand der zu diesem Spritzer aufgezeichneten Bilddaten (vier aufeinanderfolgende Bilder) läßt sich erkennen, daß der Spritzer als ein Objekt herausgelöst wurde und diagonal in die rechte obere Ecke des Bildfeldes flog (andere helle Punkte sind Spiegelungen des Spritzers in den Metallflächen).

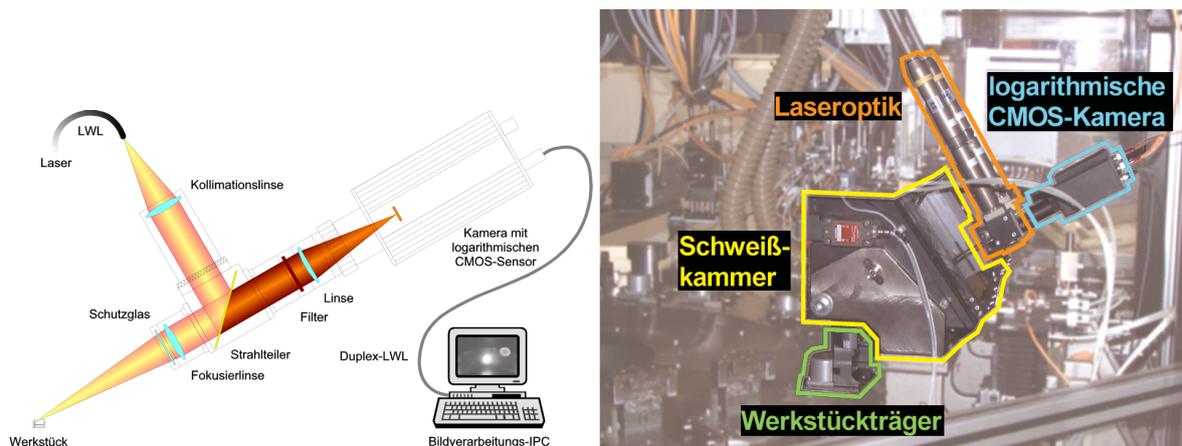


Abbildung 4.9: Links: Schematischer Aufbau: über einen Umlenkspiegel angekoppelt sieht die Kamera durch entsprechende Schutzfilter stets ein fokussiertes Bild mit dem Laserstrahl in der Mitte Rechts: Ein Rundtisch führt die bestückten Werkstückträger paarweise zu den beiden parallelen Schweißkammern (hier hintereinander verdeckt), die über die Werkstückträger abgesenkt werden. Deutlich erkennbar ist die um 90° gekippte Ankkopplung der beiden Kameras an die Laseroptik.

4.8 Aufbau und Ablauf des Überwachungssystems

In diesem Abschnitt werden eine grobe Übersicht zum Ablauf der Prozeßüberwachung und die wesentlichen Gerätekomponenten vorgestellt. Die Details zur Übertragung der Algorithmen aus Kapitel 3 auf das vorliegende Problem werden danach in Abschnitt 4.9 vorgestellt.

Das System zur Prozeßüberwachung (Abbildung 4.9) besteht pro Produktionslinie im wesentlichen aus einer Kamera, einer Anschlußoptik, ein PC mit der Software zur Auswertung und Maschinenkommunikation, entsprechender Bedieneroberfläche und Eingabemöglichkeiten sowie unterbrechungsfreier Stromversorgung. Eine mechanische Einheit zur Ausschleusung von Bauteilen nach einer NIO-Klassifizierung ist in der Regel an den einzelnen Stationen bereits vorhanden und ansprechbar.

Die eingesetzte Kamera (Abbildung 4.3 links) hat einen logarithmischen HDR-CHIP mit VGA-Auflösung (640×480 Pixel), eine Pixelgröße von $21^2 \mu\text{m}^2$ (Füllfaktor 40%), einen Dynamikbereich von 120 dB, eine Pixelrate von 8 M Pixel/s (25 Vollbilder pro Sekunde) bei einer Tiefe von 10 bit, was 1024 Graustufen entspricht. Das Setzen beliebiger rechteckiger Bildausschnitte und Gain-Offset-Verstärkerwerte erfolgt über die Software der Kameraeinbindung, kann also direkt in die ablaufende Bildverarbeitung einbezogen werden (Objektverfolgung, automatische Helligkeitsanpassung). Die Kamera benötigt eine 3,3 V-Versorgung und verbraucht weniger als 300 mW.

Die Kamera wird über einen Umlenkspiegel und durch einen Graufilter geschützt an die Optik des Lasers angeschlossen: auf diese Weise ist bei korrekt fokussiertem Laser auch immer das Bild der Kamera scharf. Das später auf 64×64 Pixel reduzierte Bildfeld zeigt aus der im Bild aufrechtstehende Hülse (die Spritzlochscheibe setzt im unteren Bildfeld an) einen ca. 1,5 mm breiten und hohen Ausschnitt. Der Brennfleck erscheint mit einem Durchmesser von 16 Pixeln. Ein Pixel sind

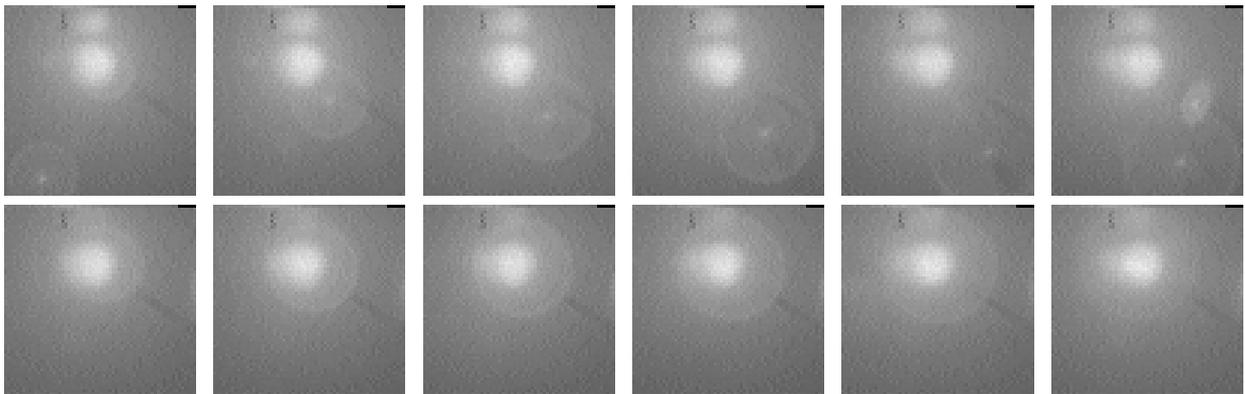


Abbildung 4.10: Spritzer, deren Bewegung eine starke Komponente in z -Richtung haben, erscheinen als Unschärfe-Scheibe mit zentraler Aufhellung. **Oben:** Für verschiedene Spritzer in verschiedenen Flughöhen erscheinen die Unschärfekreise in dieser Sequenz mit unterschiedlichen Durchmessern. Diese Durchmesser können als Maß für die Höhe des Spritzers dienen. **Unten:** Ein Spritzer fliegt auf der optischen Achse dem Betrachter entgegen. Durch seine Unschärfescheibe, die über die Ausdehnung der Schmelze anwächst, kann er detektiert werden.

ungefähr 0,02 mm. Die Tiefenschärfe des Bildes ist gering, wodurch zum Beispiel Spritzer, die vor dem Fokus liegen, ein typisches Unschärfebild bekommen.

Die Bilddaten werden über Lichtwellenleiter zur Framegrabberkarte und dem PC übertragen. Der PC ist nach außen hin des weiteren zur Fernwartung an das Firmennetzwerk und über eine digitale Ein-Ausgabe-Karte (2×16 bit, opto-entkoppelt) direkt an den Laser und die SPS-Maschinensteuerung der Laserstation angeschlossen. Vom Laser kommt das Startsignal zur Bildaufnahme, wenn das in [Abbildung 4.2](#) dargestellte Schweißprogramm startet und den ersten Heftpunkt setzt. Damit beginnt die Bildaufnahme, die nach einer festen, ausreichend großzügig bemessenen Bildfolgenlänge (1400 Bilder) endet und sofort in die Auswertung der in [Kapitel 3](#) vorgestellten Algorithmik springt. Als Ergebnisse liegen dann die Bilddaten selbst und eine hierarchisch (Daten zur Produktionslinie, Bildsequenz/Prozeß, dessen Einzelbilder, deren Objekte in Einzelbildern) organisierte XML-Datei mit den Einstellungen, Zwischenergebnissen und der IO/NIO-Aussage vor. Für IO-geprüfte Teile wird nur die XML-Ergebnisdatei gespeichert, für NIO-Teile werden Bilddaten und Ergebnisdatei gespeichert. Zusätzlich werden auf der Bedieneroberfläche in deren Betriebsmodus für die automatische Überwachung die IO/NIO-Aussage und wesentliche Parameter oder relevante Einzelbilder dargestellt und eine fortlaufende Statistik über alle geprüften Teile und Prüfergebnisse aufgebaut. Dargestellt werden zum Beispiel Anzahlen der IO-Teile und NIO-Teile, Zeitpunkte von Programmstarts, aktuelle und durchschnittliche Prüfzeiten, mittlere Grauwerte der Bilder der Sequenz, Maßzahlen und deren Schwellwerte sowie Einzelbilder und deren Änderungskarten mit starken Änderungen. NIO-Ergebnisse führen automatisch (auf Wunsch auch erst nach manueller Quittierung durch den Bediener) zu einer Ausschleusung des Bauteils, die nach Kommunikation mit der SPS-Maschinensteuerung von dieser und der von ihr kontrollierten Mechanik dazu durchgeführt wird. Dieser Ablauf ist unterfüttert mit permanenten Selbsttests und Spiegelungen der an die SPS-Maschinensteuerung kommunizierten Daten, um sicherzustellen, daß bei Ausfällen oder allen unvorhergesehenen Störungen des Prozesses

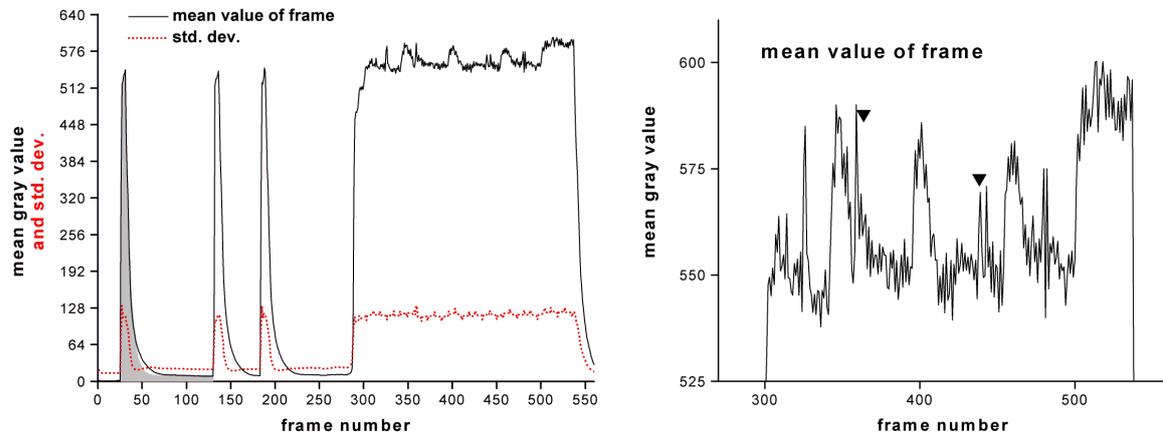


Abbildung 4.11: **Links:** Über die Bilder räumlich gemittelter Grauwertverlauf $\langle g(x, y, t) \rangle_{\Delta X, \Delta Y}$ und dessen Standardabweichung. Die Kurve folgt im wesentlichen der eingebrachten Laserenergie (Abbildung 4.6). Bei der Überschweifung der Heftpunkte und dem bereits verschweißten Teil der Naht (65° Überschweifung) ist die Bildhelligkeit erhöht. Der Grauwertverlauf ist wie gefordert stückweise stationär und wird zur Bearbeitung in mehrere zeitliche Zonen Δt_i eingeteilt. **Rechts:** Ein Ausschnitt aus der nebenstehenden Abbildung weist (markiert mit Pfeilen \blacktriangledown) auf einen Spritzer (links, unmittelbar nach einer Aufhellung) und einen weiteren Spritzer (rechts), der von einer kleineren lokalen Aufhellung gefolgt wird.

oder der Kommunikation stets Abbruchsignale oder Warnungen gesendet werden und niemals ein Bauteil als IO geprüft wird, wenn diese Prüfung durch beliebige äußere Einflüsse gestört war. Die Taktzeit beträgt 5 s, davon wird für die Bildaufnahme rund 1 s und die Kommunikation und Speicherung der Ergebnisse sowie Ausschleusung 2 s benötigt. Es verbleiben also für die reine Berechnung, die im folgenden Abschnitt vorgestellt wird, rund 2 s (Pentium III PC, 700 Mz).

4.9 Algorithmik zur Änderungsdetektion

Bei den verwendeten Aufnahmezeiten von bis zu 1000 fps ist es nicht möglich, die Bildindices von Start und Ende einzelner Abschnitte der Sequenz a priori festzulegen und für die fortlaufende Messung und Produktion unverändert beizubehalten. Mögliche Verzögerungen des Startsignals für die Bildaufnahme, unsichere Kamerasynchronisation und die Tatsache, daß bei der verwendeten Kamera kontinuierlich Daten gesendet werden und im schlimmsten Fall fast ein Bild bis zum nächsten möglichen Bildstart gewartet werden muß, machen dies unmöglich. Messungen an den Heftpunkten, die durch schlagartig eingebrachte Laserenergie gesetzt werden, haben gezeigt, daß der zeitliche Jitter im Regelfall normalverteilt ist und mit 2-3 Bildern variiert.

Die über das gesamte Bild gemittelte Grauwertintensität $\langle g(\mathbf{x}, t, n_1) \rangle_{\Delta \mathbf{X}}$ eines Prozesses (Abbildung 4.11) folgt im wesentlichen der eingebrachten Laserleistung (Abbildung 4.2). Im exponentiellen Abklingen der Heftpunkte zeigt sich nochmal das in Abschnitt 4.4 angesprochene typische Abklingverhalten von Grauwertintensitäten für CMOS-Sensoren. Die Grauwerte fallen mit einer exponentiellen Zerfallsrate von rund $\lambda \approx 5,16s^{-1}$ ab. Von einem Grauwert sind bereits im Folgebild nur noch

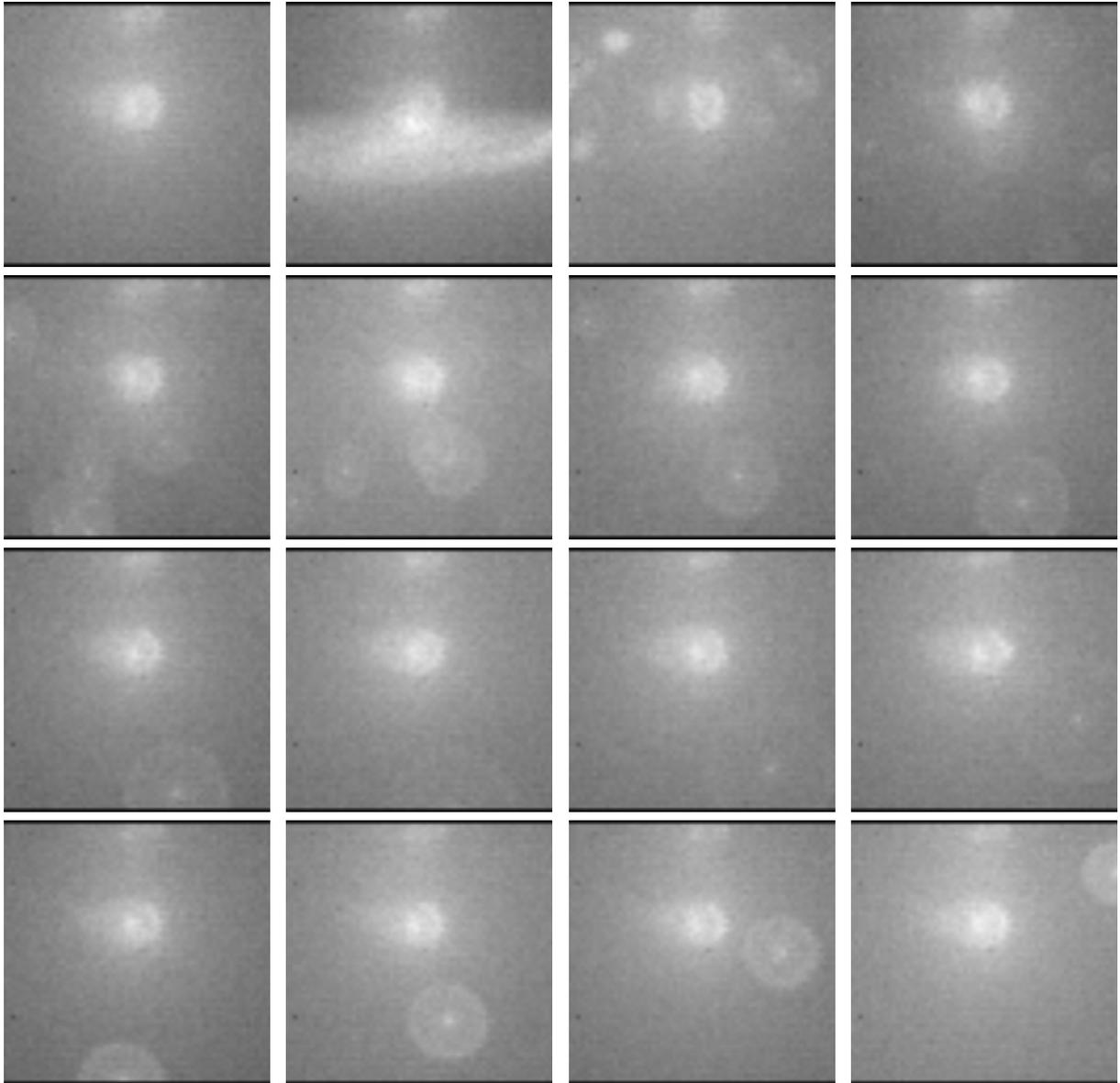


Abbildung 4.12: Unmittelbar auf ein Gut-Bild folgend kommt es zu einer Aufhellung (auch in den Mittelwerten in [Abbildung 4.11](#) gut erkennbar), danach kommt es zu einem explosiven Ablösen zahlreicher Partikel. In den mittleren Reihen gut zu verfolgen ein Spritzer, der zuerst zum unteren Bildrand fliegt, dann aber unscharf am rechten Rand aus dem Bild austritt. Rund 20 ms später durchquert ein weiterer Spritzer das Bildfeld von unten nach rechts oben. Auswertungen hierzu in den [Abbildungen 4.16](#) und [4.17](#)

33% des Wertes vorhanden. Deutlich tritt die Überschweißungszone von 65° hervor. Spritzer jedoch lassen sich in dieser Mittelung oder der Varianz dazu (in Abbildung 4.11 zwei Ereignisse bei den Bildindices 368 und 435) nicht verlässlich ausmachen. Daran erkennt man wieder die Schwierigkeiten von räumlich integrierenden Sensoren an dieser Aufgabe. Steigt das pixelweise Rauschen an, was durch einfaches Hinzufügen von multiplikativem normalverteiltem Rauschen der Standardabweichung $\sigma \cdot g$ simuliert werden kann, ändert sich im Verlauf der Mittelwerte wenig, denn die Unruhe im Verlauf der Kurve entsteht vor allem aus der Unruhe des Prozesses an sich, die als additiver, ungefähr räumlich konstanter Beitrag zu den Grauwertintensitäten hinzukommt. Das typische Pixelrauschen wird stets kleiner sein als die Unruhe der Emissionsintensitäten eines Schweißprozesses. Das Pixelrauschen beeinflusst vor allem die Standardabweichung.

Verschiedene Teiletypen haben je nach Material und Wandungsstärke der Spritzlochscheibe unterschiedliche Pulsverläufe (Abbildung 4.2): so sind Rampen des kontinuierlichen Teiles unterschiedlich steil, das Plateau der kontinuierlichen Schweißung unterschiedlich hoch und gegebenenfalls auch konstant oder sogar schwach steigend. Tests haben jedoch ergeben, daß alle verfügbaren Variationen von Typen und Laserpulsverläufen zu einem Grauwertverlauf $\langle g(\mathbf{x}, t, n_1) \rangle_{\Delta \mathbf{x}}$ führen, der sich nur um einen für den gesamten Wertebereich konstanten Multiplikationsfaktor unterscheidet. Unterschiedliche Typen oder Laserpulsverläufe haben also bei der verwendeten Algorithmik keinen Einfluß, da varianznormierte Mittelwertschätzer diese geringen multiplikativen Modifikationen (Faktoren im Bereich von $0,87 \dots 1,08$) einerseits herausrechnen, andererseits ohnehin alle Prozesse individuell geprüft werden und keine Teile-übergreifenden Mittelungen über die Dimension n stattfinden.

Für die Auswertung werden die 1400 Bilder der Sequenz in verschiedene Segmente eingeteilt. Im Bereich der Heftpunkte und Vormattierung (siehe Abbildungslegende 4.2) wird nur die Lage der Heftpunkte, genauer deren Punkte stärksten Anstiegs und Abfalls, detektiert. Im Bereich bis zum Beginn der kontinuierlichen Schweißung wird nur eine Konsistenzprüfung zur Absicherung durchgeführt

- Sind weniger als drei Heftpunkte vorhanden ?
- Finden zwischen den Heftpunkten keine starken Helligkeitsschwankungen statt ?
- Traten Kameraausfälle, fehlende Zeilen oder zerstörte Bildbereiche (Grauwert ist dann 0) auf ?

Die Änderungsdetektion findet nur für den Bereich der kontinuierlichen Schweißung inklusive der Überschweißung statt. Dafür müssen Beginn und Ende (Jitter beträgt 2-3 Bilder) der kontinuierlichen und der Beginn Überschweißung detektiert werden, um daraus die Lage von zwei zu untersuchenden Zonen zu bestimmen. Hierfür können in einem ersten Ansatz feste Schwellwerte benutzt werden, die Typenvielfalt oder auch die Unterschiedlichkeit von ADC-Gain-Offset-Einstellungen an den Kameras, verschieden stark abschwächende Optiken an den einzelnen Produktionslinien legen jedoch möglichst parameterfreie Verfahren dazu nahe. Als besonders brauchbar hat sich eine Änderungsdetektion auf der Skalarfolge $\langle g(\mathbf{x}, t, n_1) \rangle_{\Delta \mathbf{x}_c}$ herausgestellt, die mit einem Fenster von Δx_c eine zentrale (zum Beispiel 32×32) große Region aus den Grauwerten $g(\mathbf{x}, t, n)$ ermittelt. Diese eindimensionale Änderungsdetektion läuft sinnvollerweise mit einem Zeitfenster $\Delta t_w \approx 10 - 20$ rund um den gerade



Abbildung 4.13: **A:** Das initiale Mittelwertbild läßt noch keine Verunreinigungen durch noch nicht detektierte Änderungen erkennen (Wertebereich $0 \dots 1023$) **B:** Differenz der Mittelwertbilder aus der ersten und dritten Iterationsstufe (die Grau-Skala reicht von $-2,5$ (schwarz) bis $+2,5$ (weiß) bei einer Mittelung über $\Delta t = 198$ Bilder des Wertebereiches $0 \dots 1023$) **C:** initiale Standardabweichung (Skala von 0 (weiß) bis 26 (schwarz)): sie enthält neben der Unruhe der Schmelze auch die Aufhellung und Bahnen der Spritzer (vergleiche Abbildung 4.12) **D:** Standardabweichung nach der ersten Iteration: die starken Änderungen sind in der ersten Iteration bereits beseitigt, es fallen jetzt nur noch feine Strukturen der Spritzer auf. **E:** Standardabweichung nach der dritten Iteration: Änderungen sind weitestgehend beseitigt, es verbleibt die Unruhe der Schmelze mit ihrer typischen Front und Auskühlungszone sowie die obere Reflektion an der Metallhülse und eine zuvor nicht sichtbare Reflektionszone am unteren Bildrand (hier spiegelt sich das Metalldampfleuchten in der Spritzlochscheibe)

untersuchten Zeitwert. An der Größe

$$\frac{\langle g(\mathbf{x}, t, n) \rangle_{\Delta \mathbf{x}_c} - \langle g(\mathbf{x}, t, n) \rangle_{\Delta \mathbf{x}_c, \Delta t_w}}{\sqrt{\left(\langle g(\mathbf{x}, t, n) \rangle_{\Delta \mathbf{x}_c} - \langle g(\mathbf{x}, t, n) \rangle_{\Delta \mathbf{x}_c, \Delta t_w} \right)^2}}$$

werden dann Ausschläge typischerweise > 3 gesucht, an deren Zeitkoordinaten die Zonengrenzen festgemacht werden, in denen jetzt die zweidimensionale Änderungsdetektion zur Erstellung von Änderungskarten und Extraktion von Objekten stattfindet. In der Bildfolge aus Abbildung 4.11 sind an zwei Stellen Spritzer (im Ausschnitt Abbildung 4.11 von markiert mit \blacktriangledown) zu detektieren, deren Einzelbilder in Abbildung 4.12 gezeigt sind.

Nach der Segmentierung der Bildfolge in einzelne Intervalle wird nun das Verfahren zur Änderungsdetektion (3.3) wie in Abschnitt 3.3 (Seite 49) beschrieben benutzt. Dabei wird die in Abschnitt 3.3.6 vorgestellte iterative Berechnung (3.6) (Seite 57) benutzt. In Abschnitt 3.3.7 und 3.3.8 (Seiten 58 und 59) sind die Größen zusammengetragen, die man als Ergebnisse erhält. In den Bildern der Mittelwerte und Standardabweichungen, die im Zähler und Nenner der Teststatistik eingehen, erkennt man (Abbildung 4.13) deutlich, welche Spuren die vorhandenen Änderungen in den Daten hinterlassen und wie die iterative Schätzung von Mittelwert und Standardabweichung diese nach und nach eliminiert. Es hat sich gezeigt, daß diese Iteration bei Schritt $i=3$ abgebrochen werden kann, weil sich in höheren Iterationen keine Änderungen mehr ergeben. Die Grauwertverteilungen der Höhe h einer mit nur wenig Änderungen behafteten Bildfolge des Schweißprozesses lassen sich mit den beiden Verteilungen

skalierte Normalverteilung: $h \cdot e^{-\frac{(g - \langle g \rangle)^2}{2\sigma_g^2}}$

skalierte In-Normalverteilung: $\frac{h}{g} \cdot \exp \left(\langle \ln g \rangle - \sigma_l^2 + \frac{(\ln g) - \ln(e^{\langle g \rangle} - \sigma_l^2)}{2\sigma_l^2} \right)^2 - \frac{(\langle \ln g \rangle - \ln g)^2}{2\sigma_l^2} \right)$

anpassen (Abbildung 4.14 links), wobei die skalierte In-Normalverteilung gerade im linken Ausläufer etwas besser anpaßt. Dieser wurde in den Fits stärker gewichtet, da die in den Daten vermuteten Änderungen in der Regel Ausschläge zu höheren Werten sind. Dennoch kommt die angepaßte In-Normalverteilung der Gaußkurve sehr nahe. In Abbildung 4.14 (rechts) sind für einen bearbeiteten Teilbereich Δt die als Änderungs-behaftet markierten Pixel hervorgehoben. Wie zu erwarten wandert diese Unterverteilung mit zunehmender Iteration zum Zentrum der Gesamtverteilung. Anhand des prozentual ausgedrückten Anteils Änderungs-behafteter Pixel an der Gesamtzahl von Pixeln zu gegebenem Grauwert läßt sich ablesen, daß Änderungen bis hinab in den Grauwertbereich 450 vorkommen. Als Ergebnis ist festzuhalten, daß Spritzer typischerweise Grauwerte im Bereich 530 . . . 670 haben, die vor allem auch durch die Unschärfe außerhalb der Fokusebene durchaus Gauß-ähnlich verteilt sind. In höheren Bereichen ab Grauwerten 700 und aufwärts handelt es sich um starke, großflächig auftretende Aufhellungen und einzelne Reflektionen an den Metallflächen, die derart hohe Grauwerte erreichen können. Der Anteil der Änderungen steigt schließlich schlagartig auf 100% an (außerhalb des Plots), zeigt aber wegen der geringen Datenbasis starke Schwankungen.

Die Detektionsergebnisse für die in Abbildung 4.12 gezeigten Ereignisse sind in 4.16 und 4.17 dargestellt. Deutlich erkennbar ist in der ersten Iteration (vor allem bei der großflächigen Aufhellung des Gesamtbildes) die Entstehung von Artefakten an Orten, in denen zu einem späteren Zeitpunkt starke Abweichungen detektiert werden. Hier zeigt sich die Notwendigkeit des iterativen Ansatzes: nach der dritten Iteration sind die Spuren der Änderungen in den Mittelwerten und Standardabweichungen (Abbildung 4.13) eliminiert, so daß in den Teststatistiken der einzelnen Bilder keine Artefakte mehr auftreten. In der Flächendarstellung 4.19 der Grauwerte sowohl des Eingangsbildes **A** als auch der Teststatistik **D** (dritte Iteration) für ein einzelnes Bild aus Abbildung 4.16 (drittes von unten) wird noch einmal die hervorhebende Wirkung der Teststatistik klar: im Eingangsbild nur sehr schwach ausgeprägte und kaum erkennbare Erhebungen werden in der Teststatistik hervorgehoben und sind mit der Schwellwertoperation einfach zu erfassen.

4.10 Teststatistik und Bestimmung kritischer Werte

Wie im theoretischen Teil ausgeführt, geht die Student-Verteilung der Werte der ausreißerfreien Teststatistik wegen der hohen Freiheitsgrade (es werden im Beispiel 198 Einzelbilder ausgewertet) in eine Normalverteilung über. Wie in Abschnitt 3.3.4 angekündigt, muß hier überprüft werden, inwieweit sich die beiden zu den Hypothesen H_0 und H_1 gehörenden Verteilungen trennen lassen und ob H_0 tatsächlich die für die Berechnung der Schwellwerte vorausgesetzte Kurvenform annimmt. Für die verschiedenen Iterationsschritte sind die Verteilungen in Abbildung 4.15 dargestellt. Durch die Bereinigung um Ausreißer nimmt die Fläche unter den Kurven mit höheren Iterationen ab. Da prinzipiell jeder Iterationsschritt noch Änderungen enthält, insbesondere auch die das ganze Einzelbild betreffenden additiven Aufhellungen durch den wabernden Metaldampf, haben die Kurven vor allem bei den Ausreißern nach oben, in den hellen Bereich, keine exakte Gaußform. Auf der linken Seite, bei den Abweichungen ins Dunkle, kann man davon ausgehen, daß die Bilddaten Ausreißerfrei sind: es gibt keine anderen Abweichungen unter den iterativ verbesserten Mittelwert außer dem normalverteilten Pixelrauschen. Entsprechend bestätigt sich dies in der Anpassung einer Gaußfunktion an den linken Flügel der höchsten Iterationsstufe. Die Existenz eruptiver Aufhellungen des gesamten Bildes durch

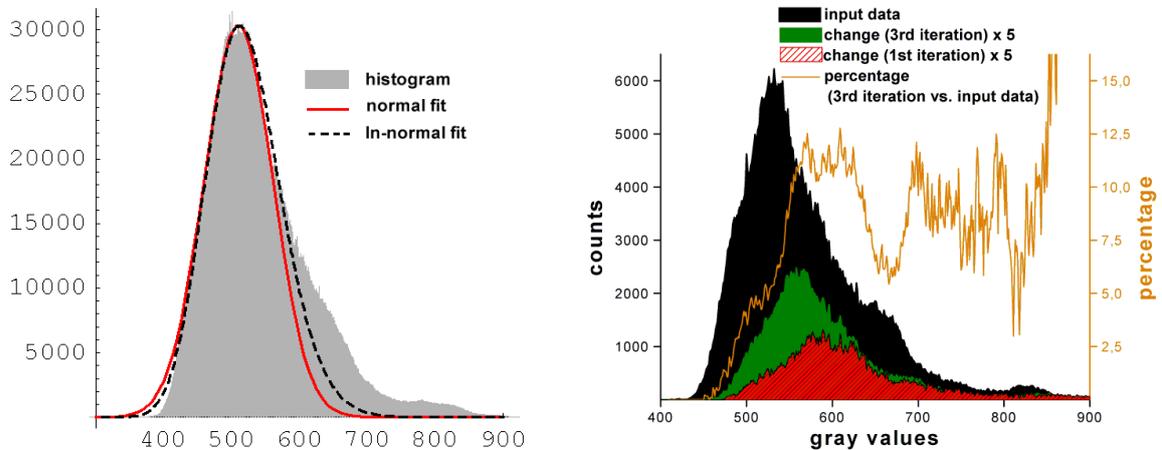


Abbildung 4.14: Links: Die Grauwerte des gesamten für die Änderungsdetektion in Frage kommenden Bereiches der kontinuierlichen Schweißung sind annähernd normalverteilt, wobei wegen der logarithmischen Kennlinie der Kamera eine ln-normale Verteilung (unterbrochene Linie) naheliegt, deren Fit der Gaußkurve (durchgezogene Linie) zwar sehr nahe kommt, aber die Daten vor allem auch im dunklen Bereich besser anpaßt. Im Grauwertbereich 600 . . . 700 und 750 . . . 850 zeigen sich die für die Szene typischen Aufhellungen der Schmelze. Vergleichbare Abweichungen in dunklere Grauwerte im Bereich 300 . . . 500 kommen nicht vor: Eruptionen und Spritzer sind typischerweise heller und nicht dunkler als das mittlere Aussehen des Prozesses. Bei den Anpassungen durch die Normalverteilung $\langle g \rangle = 510$, $\sigma_n = 55$, $h = 30300$ und die logarithmische Normalverteilung $\langle \ln g \rangle = 6,25 = \ln 518,01$, $\sigma_l = 0,105$, $h = 30300$ an die Daten wurde der Bereich 300 . . . 570 stärker gewichtet, da nur hier von einer Verteilung im Sinne der Anpassungen ausgegangen werden kann. Es gilt außerdem: $\exp(6,25 + 0,105) = 575$ und $\exp(6,25 - 0,105) = 466$ Rechts: Die Grauwertverteilung (schwarz) eines für die Änderungsdetektion ausgewählten Bereiches Δt und die Verteilungen (zur Lesbarkeit um Faktor 5 skaliert) der in den verschiedenen Iterationsstufen als Änderung detektierten Pixel.

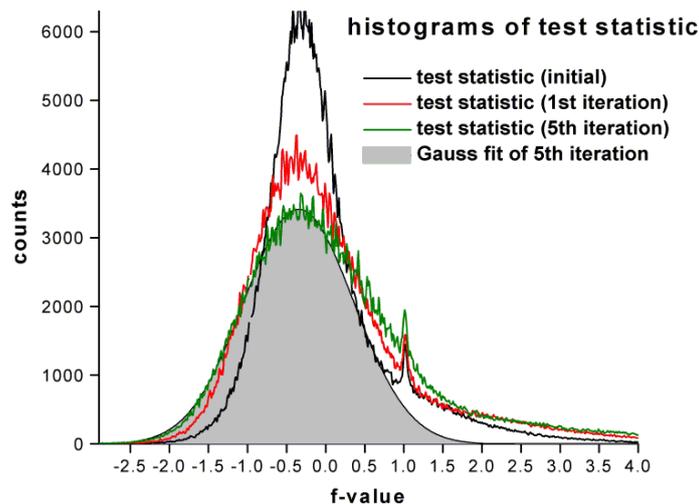


Abbildung 4.15: Werteverteilung der Teststatistik zu verschiedenen Iterationsschritten mit einem Gaußfit an den fünften Iterationsschritt (nur rechte Seite für den Fit benutzt).

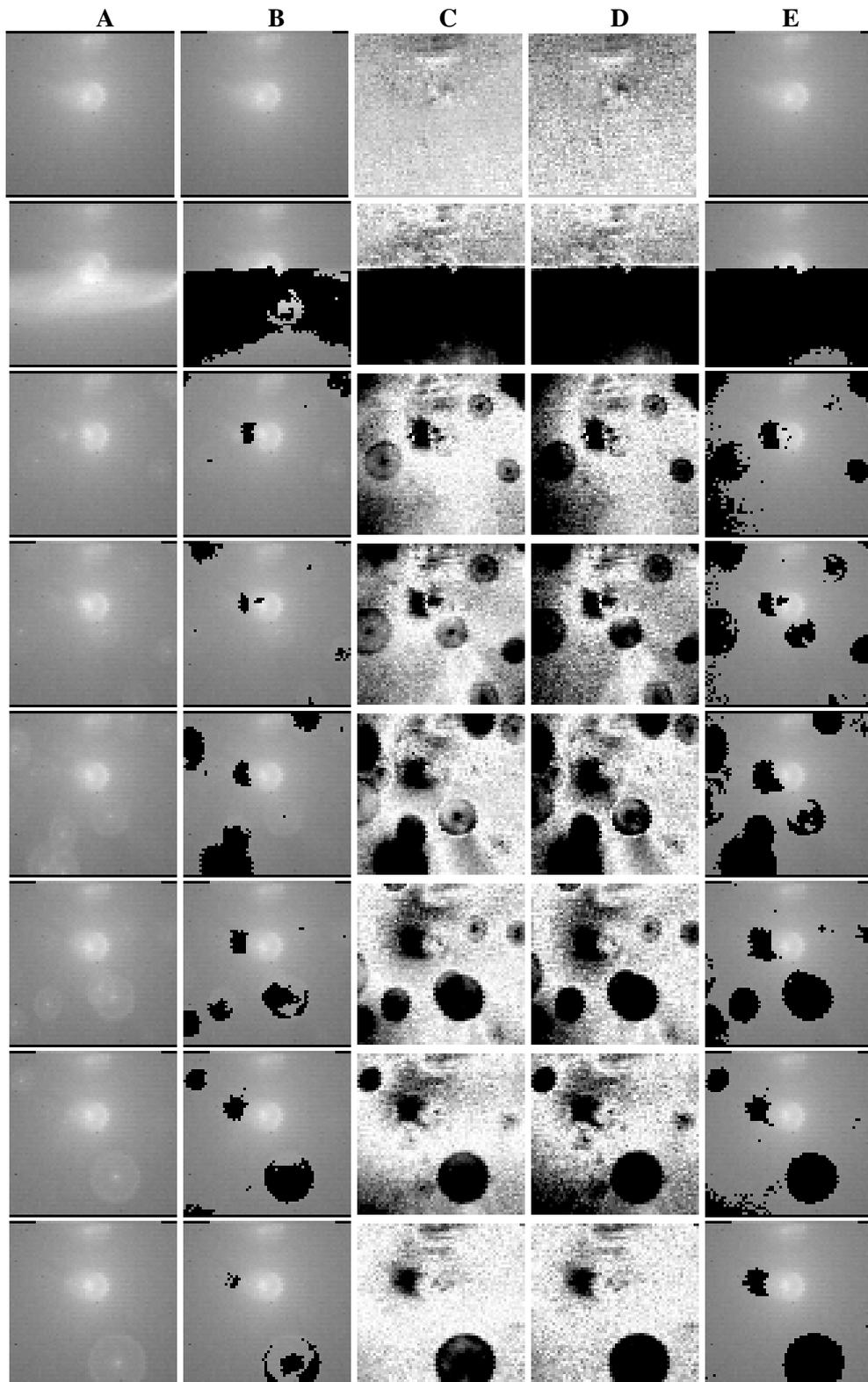


Abbildung 4.16: **A:** Eingangsbilder wie in [Abbildung 4.12](#) **B:** Eingangsbild mit überlagerter Änderung nach einer Iteration **C:** Teststatistik nach zwei Iterationen **D:** Teststatistik nach drei Iterationen **E:** Eingangsbild mit überlagerter Änderungskarte nach drei Iterationen (Fortsetzung der [Abbildung in Bild 4.17](#))

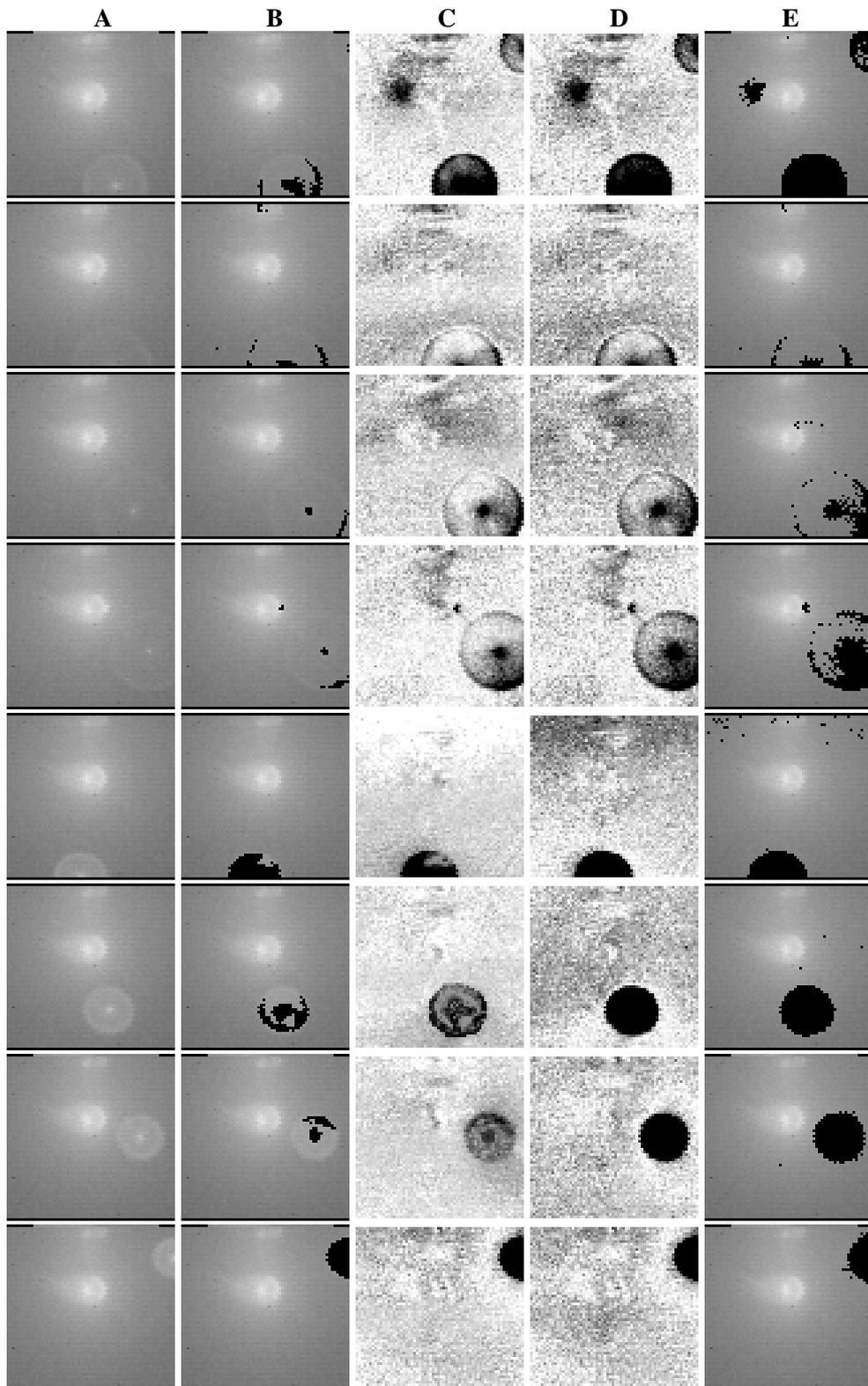
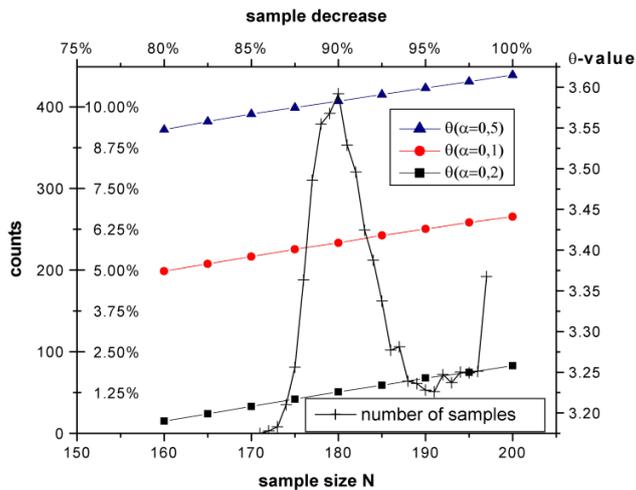


Abbildung 4.17: Fortsetzung von Abbildung 4.16 A: Eingangsbilder B: Eingangsbild mit überlagelter Änderung nach einer Iteration C: Teststatistik nach zwei Iterationen D: Teststatistik nach drei Iterationen E: Eingangsbild mit überlagelter Änderungskarte nach drei Iterationen



N	$\theta_{\alpha=0,2}$	$\theta_{\alpha=0,1}$	$\theta_{\alpha=0,05}$
160	3,190	3,374	3,548
165	3,199	3,383	3,558
170	3,208	3,392	3,567
175	3,217	3,401	3,575
180	3,226	3,409	3,583
185	3,234	3,418	3,591
190	3,243	3,426	3,599
195	3,250	3,434	3,607
200	3,258	3,441	3,615

Abbildung 4.18: Wie verringert sich bis zur letzten Iteration die Größe der 64^2 Stichproben vom initialen Wert 198? Im Mittel haben nach Abschluß der letzten Iteration die Samples eine Größe von 182,8, was einer Abnahme um 8,1 % entspricht. Die meisten Stichproben (etwa 10,4 % der 64^2 Stück) werden um 18 Einträge (entspricht 9,1 %) reduziert. Die dazugehörige Abnahme der $\theta_{\alpha,N}$ bewegt sich im Bereich 0,032, was nur minimale Auswirkungen auf die Ausreißerdetektion bei der Schwellwertbildung hat, wie Tabelle 2.1 (Seite 8).

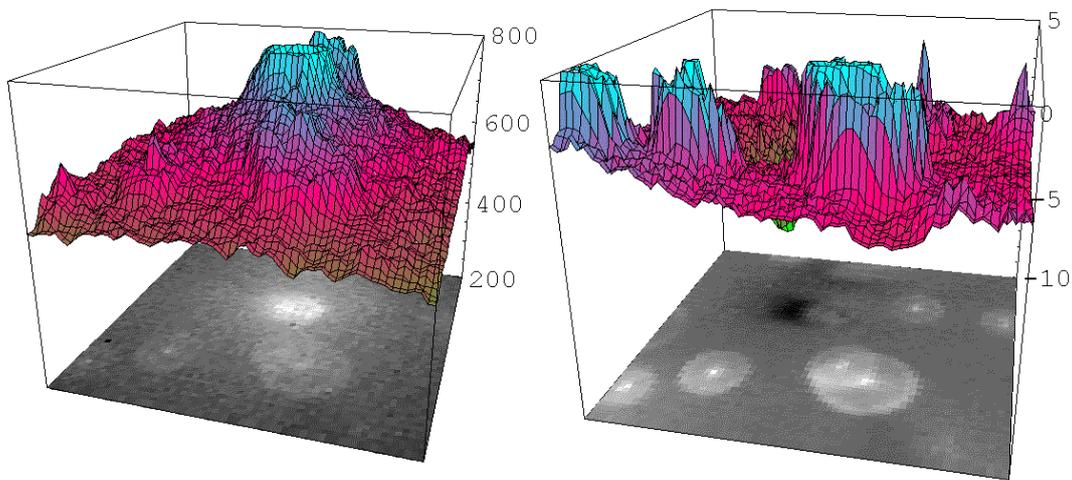


Abbildung 4.19: Links: Grauwertfläche eines einzelnen Bildes (drittes von unten aus Abbildung 4.16). Die Detektion des Spritzers mit einem einfaches Schwellwertverfahren oder einem Differenzbildverfahren ist wegen der hellen und stark variierenden Schweißung in der Bildmitte unmöglich. **Rechts:** In der zum nebenstehenden Bild gehörenden Teststatistik (dritte Iteration) erscheinen auch die im Eingangsbild nur schwach erkennbaren Partikel (zum Beispiel vorne links oder hinten rechts) deutlich hervorgehoben. Bereits eine einfache Schwellwertoperation mit einem globalen Wert $\theta_{\alpha,T}$ führt zu Änderungskarten mit kompakten Objekten und sehr wenig vereinzelt Rauschpixeln.

das über der Schweißung wabernde Metaldampfleuchten macht sich auch in der Verschiebung der Kurvenmaxima unter den Wert 0 bemerkbar. Das obere, schwere Ende der Verteilungen bleibt auch in höheren Iterationen unverändert: wie beabsichtigt werden die das ganze Bild betreffenden flackernden Aufhellungen nicht als Ausreißer detektiert. Das Artefakt bei +1 entsteht aus der möglichen Division durch 0 in der Teststatistik: wird die Standardabweichung sehr klein, explodieren die Werte der Teststatistik. Dies wird rechentechnisch abgefangen durch ein Setzen auf 1.

Der Schwellwert θ wird wie in Abschnitt 3.3.4 beschrieben und in Tabelle 2.1 dargestellt in Abhängigkeit von der Signifikanz $\alpha = 0,1$ und den Stichprobengrößen (initial 198, nach der letzten Iteration im Mittel $\langle\beta\rangle_{\Delta x, \Delta y} = 182$). Die Verteilung der Reduktionen der 64^2 Stichproben und die dazugehörigen Verläufe der Schwellwerte $\theta_{\alpha;N}$ sind in der Abbildung bei Tabelle 4.18 dargestellt. Solche Abnahmen, also Änderungsanteile von 8,1 % sind unkritisch für das vorgestellte Verfahren. Der anfängliche Wert für θ beträgt also bei der Verarbeitung der hier durchgerechneten stationären Szene 3,615 und fällt dann im Mittel ab auf 3,414. Bei der visuellen Inspektion der Änderungskarten von diesem Vorgehen gegenüber einem Vorgehen mit räumlich-zeitlich konstantem θ auf initialem Niveau stellt man zwar Unterschiede fest, es ist jedoch praktisch nicht zu entscheiden, welche Detektion “besser” ist. Das iterative Nachführen des Schwellwertes hat in diesem Fall also wenig Einfluß auf die Güte des Ergebnisses.

4.11 Rauschen im Vergleich zu den Variationen

Die in den vorliegenden Daten vorkommende Unruhe muß unterschieden werden in mehrere Beiträge. Das reine Pixelrauschen bei ruhender Szene wird vom Hersteller mit 45 dB bei maximalem Grauwert angegeben, was sich mit unseren Messungen (Abbildung 4.20 oben) deckt. Zu den laut Hersteller rund 0,7 % defekten Pixeln kommen rund 3 % Pixel hinzu, die bei einer homogenen mittleren Beleuchtung über einen längeren Zeitraum ein Vielfaches der Standardabweichung von den Werten ihrer Umgebung abweichen. Beide Effekte lassen sich zu einem Sensorrauschen zusammenfassen. Über einen sehr weiten Bereich von Intensitäten ergibt sich ein Anstieg des SNR ab rund 10 cd m^{-2} (aus [Seiffert, 2001]).

Während bei CCD Sensoren zeitliches Rauschen vor allem durch das Schrotrauschen (engl. *shot noise*) sowie das thermische Rauschen und $1/f$ Rauschen am Ausgangsverstärker eingebracht wird, sind die zusätzlichen Quellen für das generell höhere Rauschen bei CMOS-Sensoren die Anteile aus den Transistoren für die einzelnen Pixel (bei *Active-Pixel-Sensors*) und die einzelnen Zeilen (thermisches Rauschen und $1/f$ Rauschen). Die Analyse der einzelnen Bestandteile wird weiterhin verkompliziert durch Schaltkreise, die nicht-stationär arbeiten, und nicht-lineare Kennlinien in der Konversion gesammelter Ladung in die Ausgangsspannung. Arbeiten dazu wurden vorgelegt von [Tian, 2000] und [Tian, Hui u. a., 2001], der zur Analyse der Rauschanteile und insbesondere deren zeitlich nicht-stationärer Anteile und Ursachen nicht im Spektralraum, sondern im zeitlichen Ortsraum arbeitete. Insbesondere beim reset kann die konventionelle Analyse im Spektralraum nicht benutzt werden. Über das Rauschverhalten des benutzten Sensors bei zeitlich veränderlichen Intensitäten stellt [Seiffert, 2001] Ergebnisse zusammen; im Fall zeitlich konstanter Intensitäten finden sich Untersuchungen bei [Gröning, 2002].

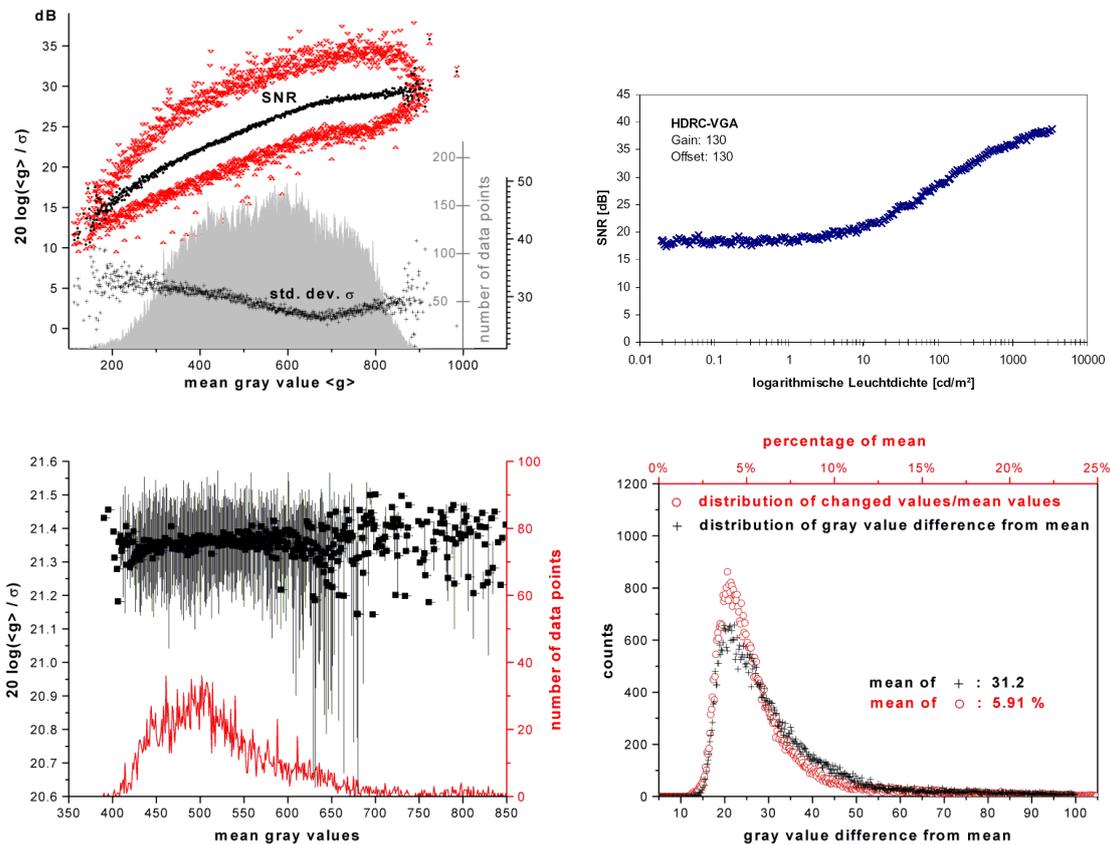


Abbildung 4.20: Oben links: Das Pixelrauschen des Sensors ist erheblich geringer als das der Variationen, die in den Schweißdaten auftreten. Der Plot wurde aus mehreren Aufnahmen unterschiedlich heller natürlicher und ereignisfreier Szenen mit ruhender Kamera gewonnen. Dargestellt ist das über die Grauwerte hin ansteigende SNR mit oberen und unteren Schranken, das Histogramm der zugrundeliegenden Sequenz und der Verlauf (+) des Rauschens in absoluten Grauwerten. Oben rechts: Jenseits des konstanten Bereiches steigt das SNR für besonders hohe Intensitäten in der logarithmischen Skala an. Unten links: Signal-Rausch-Verhältnis aus Bilddaten einer Schweißung ohne Spritzer. Für die vorkommenden Grauwerte ist das SNR und die für die Berechnung benutzte Anzahl von Datenpunkten aufgetragen. Die Länge der Fehlerbalken ergibt sich aus den minimalen und maximalen Werten für das Rauschen σ , das unter den in einem Grauwert $\langle g \rangle_t$ zusammengefaßten Datenpunkten vorkam. Unten rechts: Wie verändern die detektierten Spritzer und Aufhellungen die Grauwerte in den Daten? Die Verteilungen geben an, um welche Beiträge die Intensitäten typischerweise ansteigen.

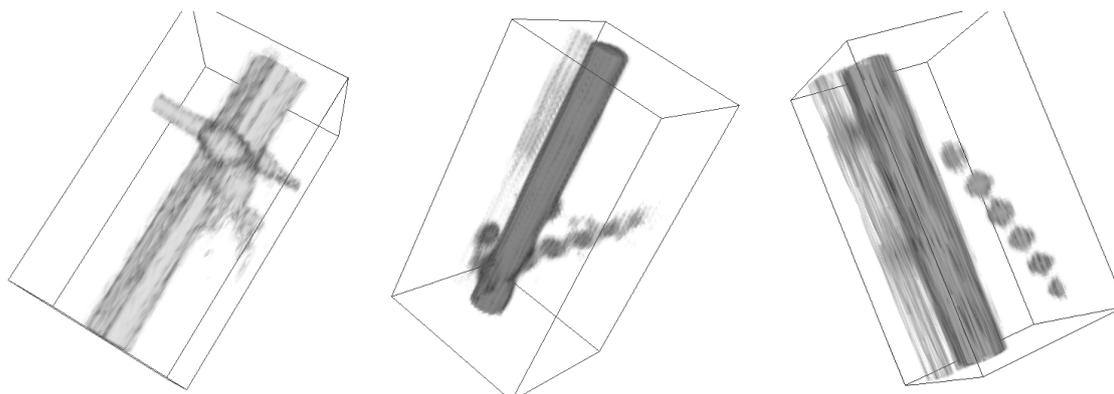


Abbildung 4.21: Plots im x - y - t -Raum für Ausschnitte aus drei Prozessen. Für die detektierten Änderungskarten wurden die Grauwerte aus den Originalbilddaten benutzt, alle anderen Pixel bleiben Null: dargestellt ist somit das pixelweise Produkt $b \hat{g}$. Eine zylinderförmige Hülse rund um die Schweißung erscheint beim Rendering als Artefakt und wurde belassen, um Ausdehnung und Position der Schmelze ungefähr zu verzeichnen. **Links:** Aufhellungen und Eruptionen des Metaldampfes erscheinen im x - y - t -Plot als Strukturen senkrecht zur Zeitachse (15 Einzelbilder). **Mitte:** Sich ablösendes Metall steht diagonal zur Zeitachse, der Winkel ist ein Maß für die Flugeschwindigkeit (15 Einzelbilder) **Rechts:** In sechs von zehn aufeinanderfolgenden Einzelbildern ist die Trajektorie eines einzelnen Spritzers zu erkennen. Das Rendering in der Visualisierungssoftware läßt die Scheibenform der detektierten Objekte als diffuse Punktwolken erscheinen.

Die direkt in den Daten der Schweißung (ohne Ereignisse wie zum Beispiel Spritzer) ermittelten Werte für das SNR (Abbildung 4.20 unten links) zeigen ein stärkeres Rauschen, was vor allem auf die Prozeßunruhe zurückzuführen ist: das über der Schweißung stehende Metaldampfleuchten bringt additive, das gesamte Bildfeld aufhellende Beiträge ein. Dieser Effekt zeigt sich auch in den schweren Ausläufern in Abbildung 4.15 zu höheren Helligkeiten. Diese Variationen haben über weite Bereiche der vorkommenden Grauwerte ein SNR von $\approx 21,35$ dB, was einem Rauschen entspricht, das rund $1/11,7$ mal schwächer ist als der dazugehörige Grauwert. Für die Analyse der von Spritzern und anderen detektierten Änderungen eingebrachten Variationen analysieren wir das pixelweise Produkt $b(x, y, t, n) \cdot [g(x, y, t, n) - \langle g^{(I)}(x, y, t, n) \rangle_{\Delta t}]$ aus Änderungsmaske (1 für Änderungen) und Mittelwertabweichungen (Mittelwertbild aus letzter Iteration). Diese Verteilung (Abbildung 4.20 unten rechts) gibt an, um wieviele Grauwerte geänderte Pixel angestiegen sind. In der Abbildung stehen “+” für Differenzen in absoluten Grauwerten und “o” für Differenzen prozentual vom Wert des Mittelwertbildes an der betreffenden Stelle (x, y) . Die von einem Spritzer eingebrachte Veränderung macht im Mittel einen Anstieg um 31 Grauwerte aus (am häufigsten ist der Anstieg um 22,3 Grauwerte), das entspricht anteilig am Mittelwert einem Anstieg um 5,9 % (am häufigsten ein Anstieg um 3,9 %). Verglichen mit dem Pixelrauschen und den Variationen durch das Metaldampfwabern sind die Anstiege aus dem gesuchten Rauschen also nur rund halb so groß.

4.12 Erkennung von Objekten und deren Beschreibung

In dreidimensionalen x - y - t -Plots 4.21 der Teststatistik oder auch den unteren vier (auch in der Originalsequenz aufeinanderfolgenden!) Einzelbildern in Abbildung 4.17 wird deutlich, daß sich die Änderungskarten in aufeinanderfolgenden Einzelbildern nur selten überlappen: die Bewegung der Partikel ist zu groß verglichen mit ihrer Ausdehnung und der zeitlichen Abtastrate der Kamera. Die im Überwachungsbereich der Bildfolge, der kontinuierlichen Schweißung inklusive der Überschweißung, als Änderungs-behaftet identifizierten Pixel werden daher für jedes Bild einzeln zu zweidimensionalen Regionen zusammengefaßt.

Morphologische Operationen, auf die in Abschnitt 3.3.7 verwiesen wurde, haben sich nicht als notwendig erwiesen. Die Anzahl der pro Bild t in der invertierten binären Maske $\text{Neg } \mathbf{b}(x, y, t, n)$ (markiert alle Pixel mit starken zeitlichen Änderungen mit 1, sonst 0) enthaltenen Einzel-Pixel ist bei einer Bildgröße von 64×64 so gering, daß eine morphologische Unterdrückung mehr Rechenzeit kostet, als durch überlange Listen im folgenden Verarbeitungsschritt verbraucht wird. Die in sich geschlossenen (wir benutzen eine 8er-Nachbarschaft) Objekte der einzelnen Bilder aus der n -ten Statistik-Folge $f^{(i)}(\mathbf{x}, t, n)$ (nicht zu verwechseln mit den Grauwertbildern $g(\mathbf{x}, t, n)$) werden mit ihren Merkmalen pro Bild in eine eine Liste gelesen. Diese Objekte werden vereinzelt und indiziert (engl. *labelling*). In der Praxis wird dafür ausgehend von $\text{Neg } \mathbf{b}(x, y, t, n)$ allen zu einem Objekt zugehörigen Pixel ein eindeutiger Index als Wert zugeordnet; die zweidimensional indizierte Maske sei mit $l(x, y, t, n)$ bezeichnet.

Es ist in der hier beschriebenen Anwendung nicht in vertretbarem Maße machbar und vom Kunden auch nicht erwünscht, die Daten aller aufgezeichneten Bildfolgen zu konservieren. Dennoch werden für die Verbesserung des Verfahrens, Fehlersuche und für allgemeine Trendaussagen pro Bildfolge und Prozeß wenn nicht die Bilddaten selbst, so doch wenigstens eine ausführliche Aufstellung der damit berechneten Ergebnisse benötigt. Zu jeder Bildfolge entsteht eine hierarchisch strukturierte Datei, die Angaben enthält über

- technische Randbedingungen:
Identifikation der Produktionslinie, eingesetzter Hardware und Softwarebausteine, globale Parameter der Software für Maschinenkommunikation und Bedienerkommunikation und Algorithmen der Bildfolgenanalyse
- zur Bildfolge:
Datum, Dateiname, Auflösung, globale Parameter und Ergebnisse der Prozeßkontrolle
- zu den Segmenten der Bildfolge:
Bildindices des Intervalls, Mittelwerte und Varianzen (gemittelt über x, y, t) in den Bilddaten g und in der Teststatistik f sowie zum Signal-Rausch-Verhältnis Angaben zur Häufigkeit starker Abweichungen aus β
- zu Einzelbildern aller Segmente:
Bildindex, Anzahl gefundener Objekte, sowie Maxima/Minima, mittlerer Grauwert/Varianz über die gesamte Bildfläche, in der binären Maske \mathbf{b} , in den Bilddaten g und in der Teststatistik f

- zu Objekten aller Einzelbilder:
deren Merkmale wie Fläche, Schwerpunkt, Grauwertsumme, Momente, Hu-Invarianten in der binären Maske \mathbf{b} , in den Bilddaten g und in der Teststatistik f

Während der erste Informationsblock nur einmal pro Datei vorkommt, treten die übrigen unterschiedlich häufig auf: je nachdem wie viele Objekte beispielsweise in einem Einzelbild gefunden oder wie viele Segmente in der Bildfolge identifiziert wurden. Diese Daten wurden zwecks sinnvoller Weiterverarbeitung [Ritter, 2001] in ein XML-fähigem Format aufgezeichnet und nach jedem Zwischenschritt der Bildverarbeitungskette gesichert, um auch bei Systemausfällen Informationen so nah wie möglich am Zeitpunkt des Ausfalls zu haben und daraus gegebenenfalls Ursachen zu rekonstruieren.

4.13 Merkmale

Als Merkmale (Abschnitt 3.4, Seite 60) für die Charakterisierung der detektierten bildweisen Änderungsregionen werden zentrale Momente (3.8) und deren Hu-Invarianten (3.26) zu den in $l(x, y, t, n)$ enthaltenen Objekten basierend auf den Grauwerten des Produktes $f^{(i)}(x, y, t, n) \text{Neg}[\mathbf{b}(x, y, t, n)]$ extrahiert. Der Merkmalsvektor für ein Objekt (mit dem Index o gekennzeichnet)

$$\mathbf{m}_o(n, t) = (1, m_{o,2}, m_{o,3}, \dots, m_{o,Q})^T$$

besteht aus Invarianten unter Rotation und Translation. Dies sind für jedes Objekt (der besseren Lesbarkeit wegen ist der Index o hier weitestgehend unterdrückt) aus den Berechnungen in den Abschnitten 3.3.6-3.3.7 (Seite 55ff.)

binäre Fläche des Objektes in $l(x, y, t, n)$	$\sum_{x,y \in o} \mathbf{b}(x, y, t, n)$
Signifikanzvolumen in $f^{(i)}(x, y, t, n)$	$v_o = \sum_{x,y \in o} f^{(i)}(x, y, t, n)$
maximaler Wert in $f^{(i)}(x, y, t, n)$	$\max_{x,y \in o} f^{(i)}(x, y, t, n)$
Exzentrizität (3.22) eines Objektes auf Basis von $f^{(i)}(x, y, t, n)$	$\epsilon_f = \frac{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}{(\mu_{20} + \mu_{02})^2}$
Grauwertvolumen in $g(x, y, t, n)$	$\sum_{x,y \in o} g(x, y, t, n)$
maximaler Wert in $g(x, y, t, n)$	$\max_{x,y \in o} g(x, y, t, n)$
Exzentrizität (3.22) eines Objektes auf Basis von $g(x, y, t, n)$	$\epsilon_g = \frac{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}{(\mu_{20} + \mu_{02})^2}$

Der mittlere Grauwert der Objekte in g oder die mittlere statistische Signifikanz in $f^{(i)}$ ergeben sich aus den Volumina dividiert durch die binäre Fläche des Objektes. In der Teststatistik $f^{(i)}$ steckt die pixelweise Information, wie stark eine Abweichung über die normale Varianz des Rauschens hinausgeht. Grauwertfläche und Grauwertvolumen eines Objektes in $f^{(i)}$ geben also Auskunft darüber, wie ausgeprägt eine detektierte Region mit schlagartiger zeitlicher Änderung der Grauwertintensitäten in ihrer Ausdehnung und Stärke der Abweichung ist. Für die tatsächlich dahinter stehenden Grauwertintensitäten muß auf g zurückgegriffen werden. Auf eine Verwendung der Gütemaße der Daten in $f^{(i)}$, wie sie in β enthalten sind, wurde verzichtet. Denkbar wäre eine geeignet normierte Gewichtung der Teststatistik $f^{(i)}$ mit β durch pixelweise Multiplikation $f^{(i)} \cdot \beta / \sum_{x,y \in o} \beta$.

Ritter hat seiner Diplomarbeit [Ritter, 2001] diese Merkmale untersucht und kam für diese Anwendung zu dem Ergebnis, daß das Signifikanzvolumen besonders geeignet ist für die Differenzierung unterschiedlicher Regionen in der Änderungsmaske und die Hervorhebung tatsächlich relevanter Ereignisse unter einer Vielzahl von Regionen ermöglicht. Dies erlaubt die binäre Fläche mit deutlich weniger Information über die tatsächliche Ausprägung der Region in der Teststatistik nicht. Der maximale Wert dagegen ist als alleiniges Merkmal völlig ungeeignet. Erst als zusätzliche Information zur binären Fläche kann er die Form und Ausprägung der Region rudimentär beschreiben, was sich durch das Signifikanzvolumen allerdings erheblich genauer erreichen läßt. Für eine detailliertere Formbeschreibung der Änderung sind die Momente auf der Teststatistik $l \cdot f^{(i)}$ in jedem Fall geeigneter als die Momente auf den ursprünglichen Objekten. Erst wenn dadurch abgesichert ist, daß es zum Beispiel aufgrund von Rundheit (Exzentrizität) um einen Spritzer handelt, macht eine Analyse in den Grauwerten, eventuell nach Subtraktion des Mittelwertbildes, Sinn, wenn es um Fragestellungen wie Formveränderung, Auskühlung, Flugbahn des Partikels geht. Für die Gut-Schlecht-Entscheidung zur Qualität der Naht ist das Signifikanzvolumen, entweder der größten Objekte in der Sequenz oder in Summe über mehrere aufeinanderfolgende Bilder zur Herstellung der zeitlichen Nachbarschaft, das richtige Maß.

4.14 Klassifikation

Nach diesen Untersuchungen zeigte sich, daß bereits mit wenigen Merkmalen eine robuste Qualitätssicherung zu realisieren ist. Für die in Dauerbetrieb genommene Realisierung wurden $Q - 1 = 4$ verschiedene invariante Merkmale des Objektes o_m eines Einzelbildes, das die größte binäre Fläche $\sum_{x,y \in o} b$ besitzt,

$$\mathbf{m}_{o_m} = (1, v, \mu_{20}, \mu_{02}, \epsilon)^T$$

in die Klassifikation übernommen. Eine in Einzelfällen noch bessere Wahl ist die Grauwertsumme μ_{00} auf Basis der Teststatistik $f^{(i)}$, da hier auch die tatsächlichen Werte von $f^{(i)}$ und nicht nur deren Überschreiten der Schwelle θ eingehen. Objekte mit der Fläche 1 werden grundsätzlich ausgeschlossen, was gegebenenfalls schon durch morphologische Operationen aus Abschnitt 3.3.7 erreicht wurde. Einzelbilder ohne Objekte der Fläche > 1 generieren keine solche Vektoren. Die Komponenten des Merkmalsvektors lassen sich zu $(P+Q)!/P!Q!$ Monomen vom Grad bis zu $P=5$ (wir prüften Werte von 4...6) für diese Anwendung kombinieren, die in den Klassifikationsvektor \mathbf{H} für den Polynomklassifikator (PC) eingehen. Eine ausführliche Beschreibung des PC, einer Abgrenzung zu anderen möglichen Klassifikatoren und eine Begründung der Bevorzugung des PC findet sich im theoretischen Teil in den Abschnitten 3.5.1-3.5.5 (Seite 83ff.).

$$\mathbf{H} = \left(1, \dots, \prod_{q=1}^Q m_q^{p_{lq}}, \dots \right)^T \quad \text{mit} \quad \sum_q p_{lq} = P = 5$$

$$\mathbf{H} = \left(1, m_2, m_3, \dots, m_Q, m_2^2, \dots, m_Q^2, m_2 m_3, \dots, m_2^2 m_3, \dots \right)^T$$

Der Merkmalsraum hat damit $L = 15120$ Dimensionen und wird auf den Klassenraum mit $\mathfrak{k} = 2$ Dimensionen (“Prozeß IO” und “Prozeß NIO”) abgebildet. Dies erfolgt über das innere Produkt (3.37)

von Klassifikationsvektor und Gewichtsvektor. Da nur zwei Klassen vorhanden sind, die sich einander ausschließend gegenüber stehen, kann man von einem Ein-Klassen-Problem sprechen und mit nur einer einzigen Trennfunktion arbeiten, die die Zugehörigkeit zu einer der beiden Klassen angibt. Ein zu geringer Zugehörigkeitswert würde automatisch die Zugehörigkeit zur anderen Klasse bedeuten. Für die rechen-technische Umsetzung ist jedoch die Arbeit mit zwei Gewichtsfunktionen $s^{(NIO)}(\mathbf{H})$ und $s^{(IO)}(\mathbf{H})$ einfacher. Bei deren Berechnung muß über die Skalarprodukte aller vorhandenen Objekte o_m aufsummiert werden. Die Gewichtsmatrix \mathbf{A} wird zu einer 2×15120 Matrix. Der Ansatz (3.37) lautet mit den genannten Anpassungen

$$s^{(\kappa)}(\mathbf{H}) = \sum_{o_m} \mathbf{a}^{(\kappa)\mathbf{T}} \cdot \mathbf{H} = \sum_{o_m} \sum_{l=1}^{L=15120} a_l^{(\kappa)} H_l \quad \text{mit } \kappa \in \{NIO, IO\}$$

$$\mathbf{s}(\mathbf{H}) = \mathbf{A}^{\mathbf{T}} \cdot \mathbf{H} = \left(\sum_{l=1}^L a_l^{(NIO)} H_l, \sum_{l=1}^L a_l^{(IO)} H_l, \right)^{\mathbf{T}} \quad \text{mit } \mathbf{A} = (\mathbf{a}^{(NIO)}, \mathbf{a}^{(IO)})$$

Die beiden Komponenten von $\mathbf{s}(\mathbf{H})$ geben an, wie stark die analysierte Bildfolge zur Klasse $\kappa = NIO$ oder $\kappa = IO$ gehört. Mit der Differenz von den idealen Klassenvektoren $\mathbf{k}^{(NIO)}$ und $\mathbf{k}^{(IO)}$ ergibt sich die Klassenzugehörigkeit einer Bildfolge, die genau der Klasse zugeordnet wird, für die der Ausdruck

$$\left| \mathbf{s}(\mathbf{H}) - \mathbf{k}^{(\kappa)} \right| = \left| \left(s^{(NIO)}(\mathbf{H}), s^{(IO)}(\mathbf{H}) \right)^{\mathbf{T}} - \mathbf{k}^{(\kappa)} \right| \quad \text{mit} \quad \begin{aligned} \mathbf{k}^{(NIO)} &= (1, 0)^{\mathbf{T}} \\ \mathbf{k}^{(IO)} &= (0, 1)^{\mathbf{T}} \end{aligned}$$

am kleinsten ist. In Abschnitt 3.5.3 ist detailliert beschrieben, wie aus einem Trainingsdatensatz (TD) mit bekannten Klassenzugehörigkeiten der Merkmalsvektoren die Gewichtsmatrix berechnet wird. Dazu wird in der oben genannten Formel für den Restabstand die in \mathbf{s} enthaltene Gewichtsmatrix \mathbf{A} solange variiert, bis der Restabstand (wobei die tatsächliche Klassenzugehörigkeit $\mathbf{k}^{(\kappa)}$ bekannt ist) für alle Datensätze des TD minimal wird. Die Berechnungsvorschrift 3.45 ist äquivalent zu dieser Variation der Gewichte, die die Differenz von der bekannten Klassenzugehörigkeit minimiert. Für unsere Zwecke stand ein Trainingsdatensatz mit klassifizierten Bildfolgen zur Verfügung. Darin enthalten waren Bildfolgen, die visuell deutlich Abweichungen zum normalen Prozeß zeigten und Bildfolgen, in denen der aufgezeichnete Prozeß keine außergewöhnlichen Ereignisse zeigte, so daß die dazugehörigen Bauteile auch von der Sichtprüfung in der Produktion als IO bewertet wurden.

Die Gewichtsmatrix wurde mehrfach bestimmt, indem aus dem Trainingsdatensatz 10 Teilmengen gebildet wurden, so daß die berechneten Gewichtsmatrizen jeweils gegeneinander validiert wurden. Dabei kam eine asymmetrischen Kostenfunktion zum Einsatz, die "false accept", also NIO-Bildfolgen zu $\kappa = IO$ klassifizierte 1000 stärker bestrafte als "false alarms", also IO-Bildfolgen, die zur Klasse NIO zugeordnet wurden. Details zur optimierten Berechnung der Gewichtsmatrix für diese Anwendung finden sich in [Hader, 2004].

In Vorversuchen, wenn es zum Beispiel darum geht, den Trainingsdatensatz zu generieren, oder um die Machbarkeit des Verfahrens zügig unter Beweis zu stellen, eignet sich die oben geschilderte Vorgehensweise weniger. Als sinnvolle Wahl eines sehr sicheren Merkmals hat sich dabei das Volumen v_o des Objektes mit der größten Fläche herausgestellt. Die binäre Fläche des Objektes selbst ist dagegen ein weniger gutes Maß, vor allem wenn es darum geht, unterschiedlich starke Abweichungen vom Normalzustand sicher in eine Rangfolge einzuordnen. Verwendet man dies direkt als skalare Größe $s(t, n)$ über die Zeitachse der n -ten Sequenz, so beobachtet man für Bildfolgen mit relevanten

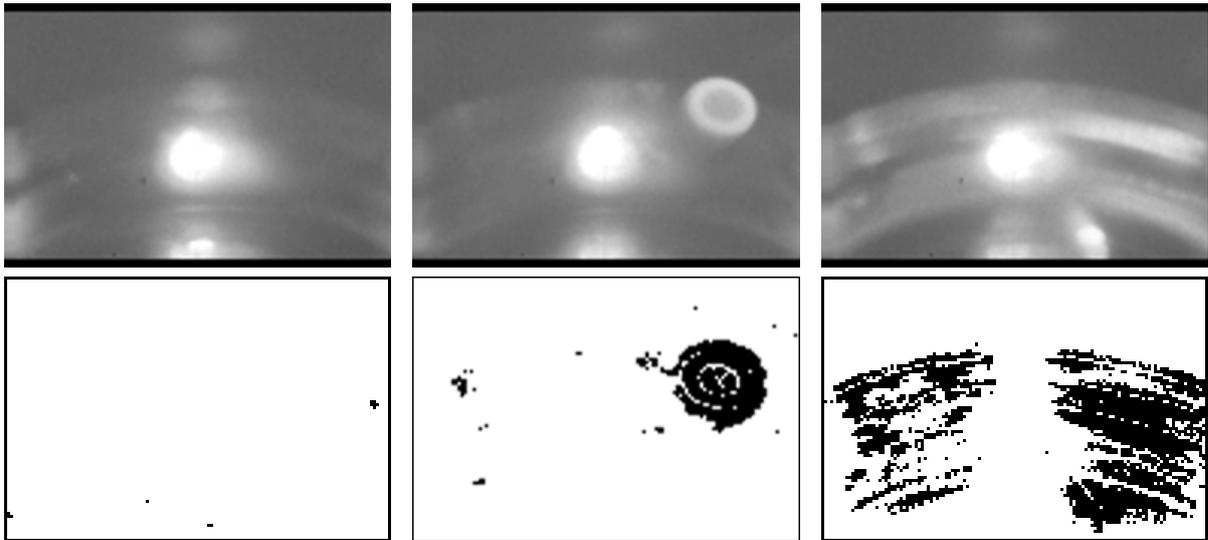


Abbildung 4.22: Bilddaten von einem im Rahmen einer Machbarkeitsstudie (Tabelle 4.2) detektierten Spritzer
Oben: Drei aufeinanderfolgende Bilder einer Rundschweißung einer Metallhülse, bei der sich ein Spritzer löst und wegfliegt. Im dritten Bild spiegelt sich das Leuchten des Spritzers noch in den Metallflächen des Bauteils.
Unten: Die korrespondierenden Änderungskarten nach Kapitel 3 dargestellt, die den Spritzer und die von ihm verursachte Aufhellung erfassen und die Schweißung selber erfolgreich ausblenden.

Abweichungen ein δ -ähnlichen Anstieg für Einzelbilder, in denen Objekte identifiziert wurden. Die Höhe der Peaks kann gleichzeitig als Maß für Relevanz der Abweichung dienen. Eine einfache Prüfung, ob $s(t, n)$ über einem (weitgehend frei wählbarem) Schwellwert liegt, gegebenenfalls in mehreren direkt aufeinanderfolgenden Einzelbildern, um bei ausreichend hoher Bildrate bzw. einer über mehrere Bilder anhaltenden Trägheit der Änderungen gegen Rauschen abgesichert zu sein, hat sich als gutes Kriterium erwiesen, um Machbarkeitsstudien voranzubringen. Für Einzelbilder ohne Auffälligkeiten ist $s(t, n)$ im Vergleich zu den Höhen der Spitzen bei solchen Ereignissen annähernd 0. Die Summe der Volumen $\sum_o v_o$ aller im Einzelbild enthaltenen Objekte konnte als ähnlich gute Größe für den Einstieg in die Verarbeitung nachgewiesen werden. Hierbei wird $s(t, n)$ sehr empfindlich, so daß leicht eine Vielzahl unterschiedlicher Typen von plötzlichen Prozeßabweichungen, letztlich eine Suche “nach allem, was ungewöhnlich ist”, sicher detektiert und automatisch gesammelt werden konnte, um einen tatsächlich aussagefähigen Testdatensatz zu generieren. Der Schwellwert für θ sollte dann auch entsprechend niedrig gewählt sein.

4.15 Durchgeführte Versuche

Für die im folgenden beschriebene Umsetzung eines Überwachungssystems wurden zuerst Grundsatzstudien zur Identifikation leistungsstarker Systemkomponenten und Probeaufnahmen von Bilddaten mit künstlich provozierten Spritzern im Labor (Abbildung 4.7 und 4.23) angefertigt.

Eine Machbarkeitsstudie in der Produktion belegte, daß das vorgeschlagene System und Verfahren auch unter Produktionsbedingungen einsatzfähig ist (Tabelle 4.2 und Abbildung 4.22). Als problema-

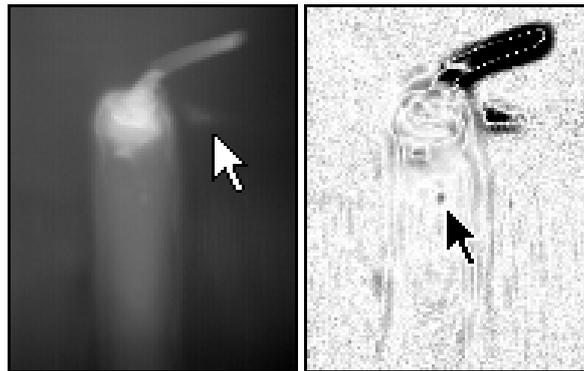


Abbildung 4.23: **Links:** Schmelzbad, Erstarrungszone und Laserfokus sind gut zu erkennen. Ein Spritzer löst sich und fliegt nach rechts aus der Schmelze fort. Ein weiterer, sehr kleiner Spritzer (Pfeil) ist unmittelbar rechts neben der Schmelze zu erkennen. **Rechts:** Die Änderungsdetektion erfaßt den großen Spritzer problemlos und unterdrückt die Schmelze korrekt. Der kleine Spritzer rechts neben dem Schmelzbad tritt in der Änderungsdetektion deutlich hervor. Darüber hinaus läßt sich im Schmelzbad und dessen Übergang in die Erstarrungszone ein Schlacketeilchen (Pfeil) ausmachen, dessen Bewegungen im Schmelzbad einen Eindruck von der Dynamik dort geben. Beide Bilder stammen aus der in Abbildung 4.7 erwähnten Grundlagenuntersuchung.

tisch erwies sich die Tatsache, daß die gesuchten Ereignisse pro Produktionslinie nur rund einmal in 24 h ununterbrochener Produktion und damit ziemlich selten auftreten. Auch war es im Rahmen dieser Studie nicht möglich, als fehlerhaft beanstandete Teile aus der Weiterverarbeitung sofort auszuschleusen und genauer zu untersuchen. Nach der Sichtprüfung, im Produktionsablauf rund 20–60 min nach der Laserschweißung, ließen sich gefundene Fehler nur schwer den Bilddaten zuordnen. Ergebnis war, daß unter Inkaufnahme einer hohen Anzahl Pseudofehler bereits ein einfacher Ansatz wahrscheinlich alle gesuchten Ereignisse finden kann. Mit den gewonnenen Bilddaten zahlreicher einwandfreier Prozesse und zwei Prozessen mit Spritzern konnte dann eine geeignete Merkmalsauswahl und Klassifikation gefunden werden; gleichzeitig wurde Hardware-seitig die Bildrate erhöht und ein robustes Startsignal für die Bildaufnahme gewählt.

Ein Dauertest (Tabelle 4.4 und Abschnitt 4.16) über 243 Tage (davon 162 ausgewertet) konnte dann die Fertigungstauglichkeit des Verfahrens und des Systems zeigen und belegen, daß das hier vorgeschlagene System die gestellte Ausgabe zuverlässig löst, so daß ein dauerhafter Einsatz an allen Produktionslinien begonnen wurde.

Aus diesem Dauertest wurden beispielhaft die Ergebnisse von sechs aufeinanderfolgenden Tagen in Tabelle 4.3 vorgestellt. Die Überwachung war bei rund 98% der produzierten Teile aktiv. Bereits in diesem Ausschnitt wird die Strategie klar, eher hohe Anzahlen an Pseudofehlern in Kauf zu nehmen, als Schlupf zu riskieren. Dies bildet sich auch in den Kostenfunktionen der Klassifikation ab. Von den 51 Prozessen, die als „fehlerhaft“ detektiert wurden ($2,1\%$) zeigen bei der genaueren Betrachtung der Bildfolgen und mit einiger Erfahrung nur rund die Hälfte tatsächlich Spritzer: die übrigen Ereignisse sind eruptives Aufleuchten des Metallampfes, der sich teilweise so ungünstig in der metallischen Innenseite des Ventils spiegelt, daß weite Bereiche des Bildfeldes schlagartig aufgehellt werden. Von den in den Bilddaten eindeutig belegten Spritzern finden wir bei der Betrachtung der dazugehörigen ausgeschleusten Bauteile bei der Hälfte (13 von 26) ein tatsächlich in der Naht sichtbares Loch oder

Ablagerungen; die andere Hälfte der Bauteile zeigt bei äußerlicher Betrachtung keinerlei Auffälligkeiten (solche Bauteile wurden in die Weitermontage zurückgestellt). Da hierzu Bilddaten vorliegen und daraus sehr genau die betreffende Stelle der Naht zu berechnen ist (Abbildungen 4.26) bietet sich an, für solche Teile einen Abschleif der Naht vorzunehmen, um gegebenenfalls verborgene Hohlräume (Abbildung 4.24) im Inneren der Naht, auf die der Spritzer ein Hinweis sein kann, nachzuweisen.

Für die Arbeitsabläufe in der Produktion muß man sich vor Augen halten, daß der scheinbar hohe Anteil von Fehlalarmen angesichts der an sich sehr seltenen Ereignisse nur dazu führt, daß pro Produktionslinie und Tag rund ein Teil fälschlicherweise ausgeschleust wird. Es kann ohne weiteres nach kurzer Draufsicht wieder in die Montage eingestellt werden, wenn es sich um eine Fehldetektion handelt. Selbst wenn die Einstellungen des Systems verlangen, daß solche Ausschleusungen vom betreuenden Personal sofort und Teile-individuell an der Steuereinheit der betreffenden Montagestation zum Beispiel durch Knopfdruck quittiert werden müssen, so ist der Arbeitsaufwand verglichen mit der im Gegenzug eingesparten 100%-Sichtprüfung aller Teile vernachlässigbar.

Die Besonderheit dieses gewählten Ausschnittes der Dauertest liegt darin, daß in diesem Zeitraum der einzige vom hier getesteten System zu detektierende Fehler (Loch in der Naht mit unmittelbar daneben angelagertem Spritzer) nicht erkannt wurde und als IO-Teil weitergeführt wurde. Bilddaten des Prozesses liegen daher leider nicht vor. Die Geometrie des Defektes deutet jedoch darauf hin, daß der Spritzer nicht erkannt wurde, weil er nur eine sehr kurze Flugstrecke zurückgelegt (eher ein Auswurf als ein tatsächlich frei fliegender Metallkörper) und sich in einem Bildbereich niedergelassen hat, in dem eine noch eine sehr hohe Varianz der Grauwerte bedingt durch den Rand der Schmelze vorliegt. Dies gab den Anstoß dazu, das *masking/swamping*-Verhalten (Erläuterungen dazu im Theorieteil Kapitel 2.2.2 Seite 10 und Abbildung 2.1) der eingesetzten Algorithmik nochmal zu untersuchen und die in Kapitel 3.3.6 (Seite 55) entworfene Iteration einzusetzen.

4.16 Ergebnisse

Hier sollen wichtige Ergebnissen, für die zahlreiche Datensätze und Bauteile vorliegen, an einzelnen Beispielen vorgestellt werden. Sie zeigen, wie hilfreich eine systematische Detektion von Spritzern und Unregelmäßigkeiten im Schweißprozeß ist, wenn anschließend sowohl das Bauteil, als auch die Bildfolge zu detaillierten Untersuchungen bereit stehen. Die gezeigten Beispiele wurden allesamt während der Langzeitstudie ausgeschleust und anschließend vergleichend mit den Bildfolgen und den Zwischenergebnissen der Algorithmik eingehend untersucht.

Es konnte gezeigt werden, daß sowohl Löcher als auch Spritzer (Abbildung 4.8) Erscheinungsformen haben, die mit dem grundsätzlichen Ablauf des Schweißprozesses (Abbildung 4.2) in Einklang sind. Durch die parallele Qualitätskontrolle während der Langzeitstudie (Tabelle 4.4) konnte unter realen Produktionsbedingungen (konventionelle Sichtprüfung und Prozeßkontrolle mit Bildfolgenanalyse beide im Einsatz) eindeutig nachgewiesen werden, daß zu Spritzern stets Löcher oder Fehlstellen gehören und diese in den meisten Fällen mit dem Bauteil und den Bilddaten vom Prozeß eindeutig zuzuordnen sind (Abbildung 4.25). Dies rechtfertigt den Ansatz, zur Qualitätskontrolle der Schweißnähte auf Fehlstellen und Löcher eine Bildfolgenanalyse zur Detektion von Spritzern und Eruptionen anzubieten. Nach Anpassungen und Verbesserungen der Algorithmik kam es nur noch zu Detektions-

Machbarkeitsstudie	Produktionslinie 1	Produktionslinie 2
Bildfenster	5,4 mm × 3,7 mm	
Auflösung	128 Pixel × 88 Pixel	
Bildrate	280 fps	
Gesamtdauer	67 h	
Überwachungszeitraum	34 h $\hat{=}$ 50%	16 h $\hat{=}$ 23%
produzierte Teile	\approx 17175	\approx 17175
Fehler lt. Sichtprüfung	4 $\hat{=}$ 100 ppm	
überwachte Schweißungen	8740	(3980)
gefundene Ereignisse	14	(wegen schlechter Bildqualität nicht auswertbar)
davon tatsächlich Spritzer	2	
davon Triggerprobleme der Kamera	3	
davon Lichtreflexe unklarer Herkunft	5	
davon eruptive Aufhellungen ohne Spritzer	4	

Tabelle 4.2: Machbarkeitsstudie unter realen Produktionsbedingungen zur Detektion von Spritzern. Vier der 34350 produzierten Teile zeigten laut Sichtprüfung die gesuchten Ereignisse. Unter den 8740 (rund 25% aller) überwachten Schweißprozessen wurden 2 davon eindeutig gefunden (Abbildung 4.22) und konnten fehlerhaften Teilen zugeordnet werden. Zwölf weitere Prozesse (darunter zum Beispiel auch drei Schweißungen, bei denen der Start der Laserschweißung von der Kamera verpaßt wurde) wurden von der Schweißüberwachung als fehlerhaft identifiziert, von der Sichtprüfung aber nicht beanstandet. Die entsprechenden Teile dazu lagen nicht vor. Der Pseudofehleranteil von \approx 85% gilt als vertretbar, da die Ereignisse insgesamt nur sehr selten vorkommen (alle 2,5 h)

fehlern zweiter Art (Teile wurden als NIO deklariert, obwohl sie die Sichtprüfung für IO einordnen würde). Fehler erster Art (tatsächliche defektes Teil wird als IO deklariert) kamen nicht mehr vor – soweit sich dies nach der mehrmonatigen Prüfung von $3 \cdot 10^6$ Teilen und darin zu erwartenden Fehleranzahlen der Größenordnung 10^3 feststellen läßt.

Dauertest (Ausschnitt)	Produktionslinie 1	Produktionslinie 2
Bildfenster	2,8 mm × 2,8 mm	
Auflösung	64 Pixel × 64 Pixel	
Bildrate	648 fps entspricht 1,543 ms pro Bild	
Gesamtdauer	≈ 42 h	
reale Taktzeit	6,18 s	
überwachte Schweißungen	12145	12162
nicht gefundene Ereignisse	0	1
gefundene Ereignisse	21	30
davon:		
- eruptive Aufhellungen ohne Spritzer	13	12
- Spritzer in Bilddaten erkennbar	8	18
davon:		
- Defekte an der Naht (lt. Sichtprüfung)	3	10
- keine Nahtfehler (lt. Sichtprüfung)	5	8

Tabelle 4.3: Auswertung eines Ausschnittes von 42 h Produktionszeit eines bestimmten Teiletyps des Dauertests zur Fertigungstauglichkeit.

Nachweis Fertigungstauglichkeit	Produktionslinie 1	Produktionslinie 2
Gesamtdauer	162 Tage ausgewertet (Laufzeit 243 Tage)	
produzierte Teile	≈ 899793	≈ 934760
reale Taktzeit	6,4 s	
überwachte Schweißungen	885644	918641
nicht gefundene Ereignisse	31	56
gefundene Ereignisse	2187	2176
davon:		
- eruptive Aufhellungen ohne Spritzer	1771	1404
- Spritzer in Bilddaten erkennbar	416	772
davon:		
- Defekte an der Naht (lt. Sichtprüfung)	168	237
- keine Nahtfehler (lt. Sichtprüfung)	248	535

Tabelle 4.4: Auswertung des Nachweises für die Fertigungstauglichkeit für die ersten zwei Drittel der gesamten Erprobung. Ein weiterer Ausschnitt der Erprobung ist in Tabelle 4.3 bereits im Detail besprochen worden.

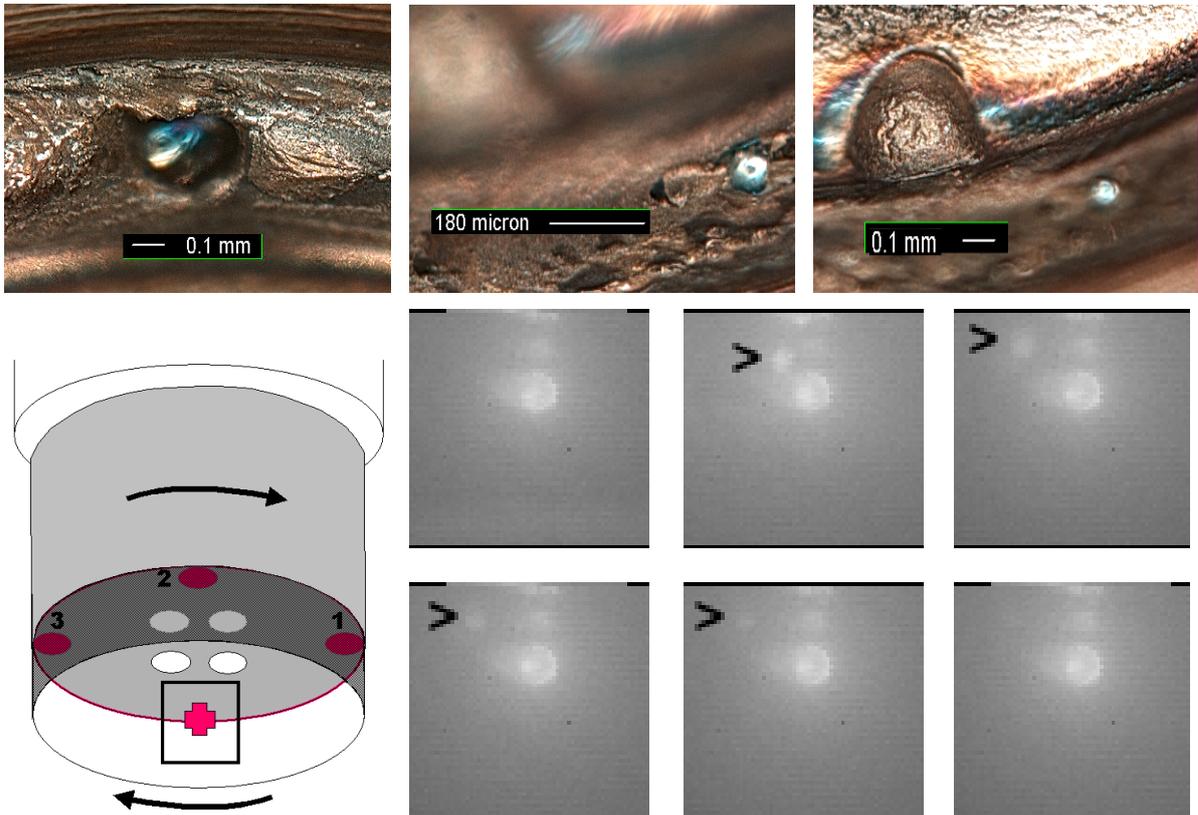


Abbildung 4.24: **Links oben:** In der hier benutzten Aufnahmetechnik erscheinen die inneren Hohlräume von Löchern in der Schweißnaht, die wie konvexe Hohlspiegel das Licht zurückwerfen, in einem bläulichen Glanzlicht **Mitte oben:** In unmittelbarer Nachbarschaft zu einem großen Spritzer befindet sich ein äußerlich sehr klein erscheinendes Loch in der Naht **Rechts oben:** Gleiche Aufnahme wie in der Mitte, jedoch mit einer anderen Fokusebene: wieder sind Loch und Spritzer zu erkennen. Angesichts der ausgeworfenen Masse muß davon ausgegangen werden, daß unter der Nahtoberfläche ein großer Hohlraum steckt, der sich beim Erstarren der Schmelze fast vollständig verschlossen hat und nur eine kleine Öffnung zurückließ. Dieses Loch stammt nicht aus dem Überschweißbereich der Naht. **Links unten:** Schematische Ansicht der Schweißung mit Bildbereich, der auf dem Kopf steht, so daß sich im Bild das Metall nach links zu bewegen scheint **Mitte und rechts unten:** Sechs aufeinanderfolgende Bilder, bei denen die Entstehung des gezeigten Auswurfes (Abbildungen oben Mitte und links) deutlich zu erkennen ist

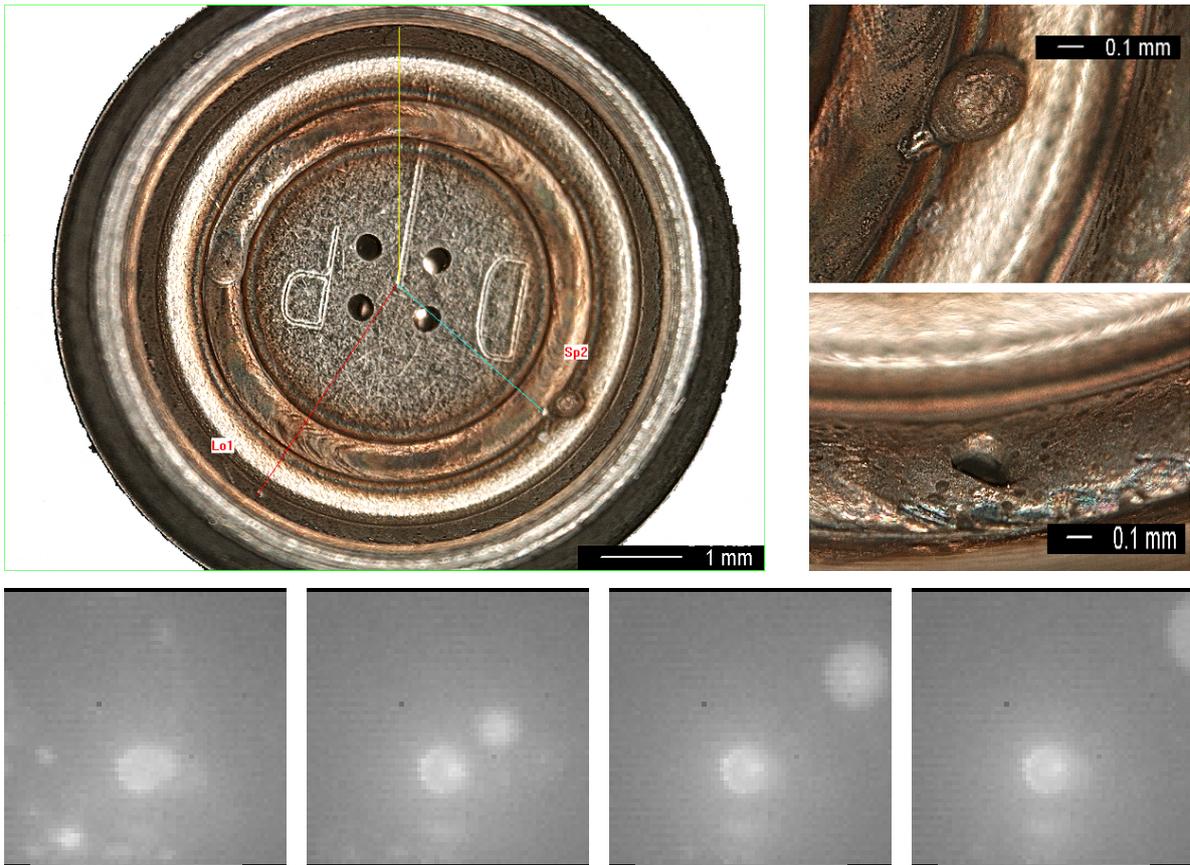


Abbildung 4.25: Links: Gesamtansicht der verschweißten Spritzlochscheibe. Gelb markiert das Ende der kontinuierlichen Schweißung, rot ein Loch in der Schweißnaht, blau einen Spritzer. Rechts: In den beiden Detailaufnahmen von Spritzer (oben) und Loch (unten) ist deutlich zu erkennen, daß das Loch zum Spritzer hin geöffnet ist und dieser offenbar auch aus der Richtung des Loches flog, bevor er an der Hülse auftraf und erstarrte. Unten: Vier aufeinanderfolgende Bilder zeigen, wie sich der Spritzer herauslöst und in die entsprechende Richtung fliegt.

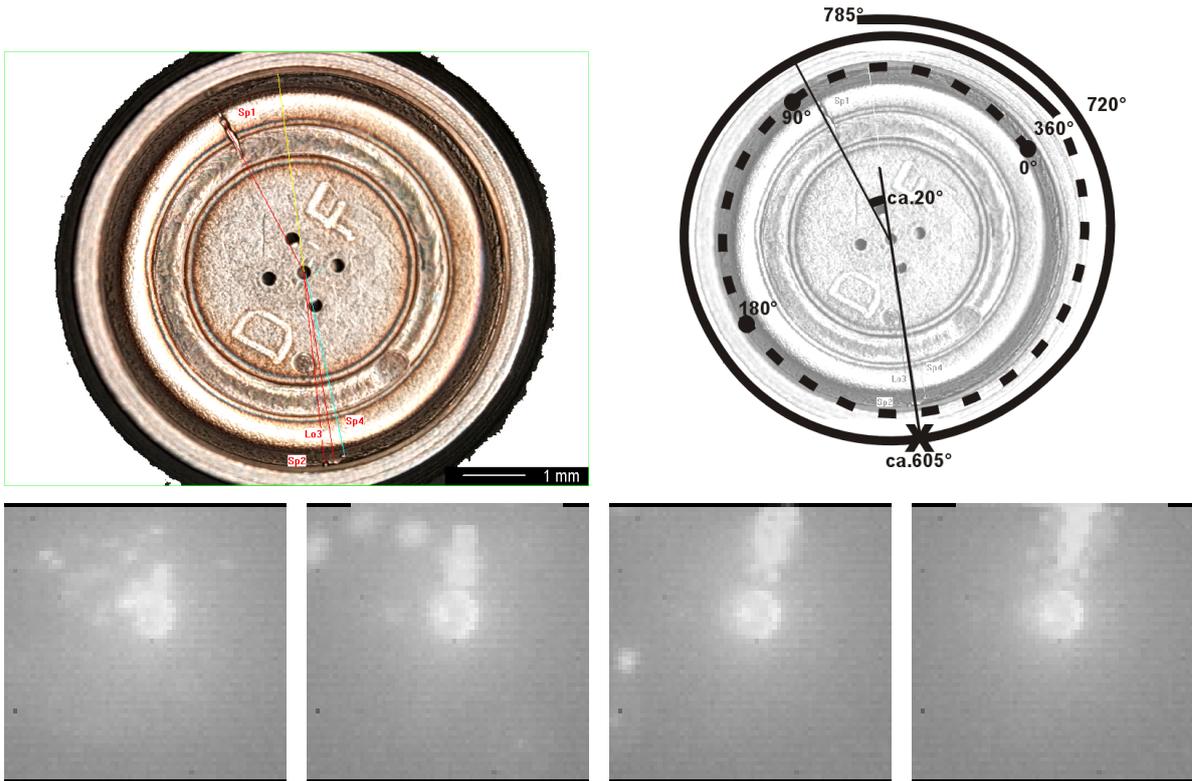


Abbildung 4.26: **Links:** Bei diesem Prozeß hat sich bei ca. 605° ein Spritzer aus der Schweißnaht herausgelöst und in insgesamt drei Fragmenten (siehe Abbildung 4.27) auf der Scheibe niedergelassen. Zwei kleine Metallreste befinden sich unmittelbar neben dem entstandenem Loch, der Hauptteil des Spritzers flog zur gegenüberliegenden Seite der Hülse und erstarrte dort als längliche Metalleruption. Das Ende der kontinuierlichen Schweißung ist gelb markiert, Löcher und Spritzer sind blau und rot markiert. **Rechts:** In der schematischen Ablaufzeichnung (Bauteil dreht sich im Uhrzeigersinn bzw. Laser dreht sich gegen Uhrzeigersinn) ist zu erkennen, daß sich der Spritzer rund 245° nach Beginn der kontinuierlichen Schweißung gelöst hat. Der Versatz in der Ausrichtung des Spritzers und der Position des Loches in der Schweißnaht beträgt 20°, dies entspricht einer Flugzeit in der sich weiter drehenden Hülse ($1,125^\circ/\text{ms}$) von 18 ms. **Unten:** In vier Bildern, die rund 6,2 ms abdecken und den Bildpositionen von 604° bis 609° entsprechen, sieht man, wie sich ein länglicher Spritzer aus der Schmelze herauslöst.

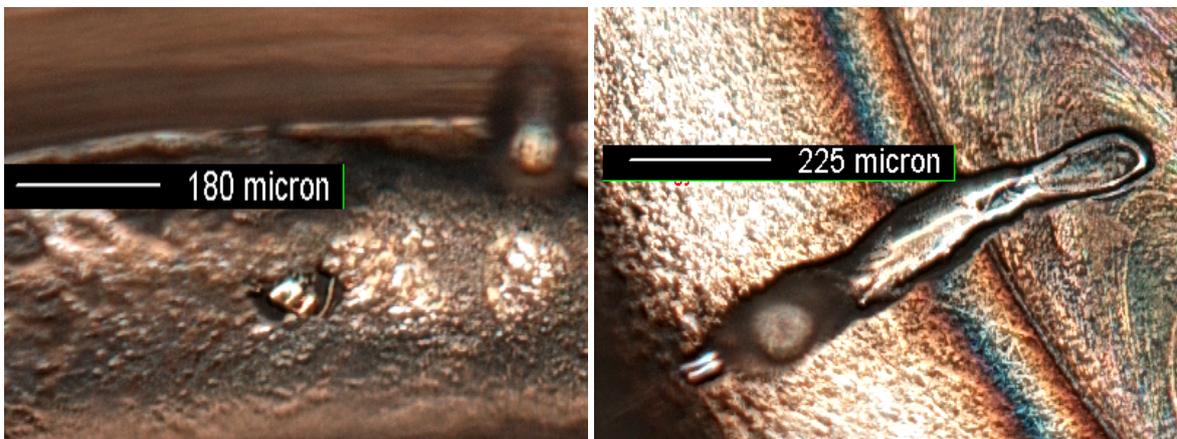


Abbildung 4.27: Die Abbildung zeigt die zu Abbildung 4.26 gehörenden Details. Weder Loch noch Spritzer wurden überschweißt. **Links:** Das Loch in der unteren Bildmitte erscheint verglichen zur ausgeworfenen Metallmenge äußerst klein. Es steht zu vermuten, daß unter der Nahtoberfläche im Bereich der Einschnürung (am oberen Nahttrand zu erkennen) Hohlräume verborgen sind. Am rechten Bildrand ist ein kleiner Metallauswurf am oberen Rand der Naht zu erkennen. **Rechts:** Der länglich ausgedehnte Spritzer erscheint gleichmäßig erkaltet.

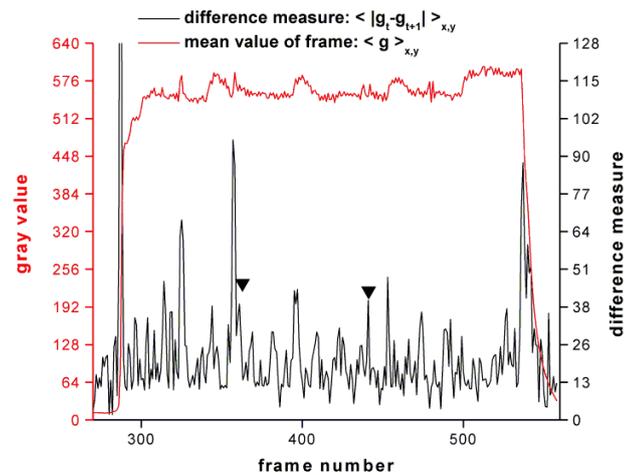


Abbildung 4.28: Vergleich räumlich integrierender Verfahren oder Differenzbildverfahren: weder mit räumlich integrierender Sensorik (simuliert mit dem räumlichen Mittelwert $\langle g \rangle_{\Delta X, \Delta Y}$) noch mit Differenzbildverfahren (simuliert anhand der Daten durch die Funktion $\langle |g(x, y, t) - g(x, y, t - 1)| \rangle_{\Delta X, \Delta Y}$) können die in diesem Kapitel detektierten Spritzer (durch ▼ markiert) sicher gegen andere Aufhellungen unterschieden werden.

Einen Teil der relativ hohen Zahl von Pseudofehlern machen wahrscheinlich Ereignisse aus, in denen es tatsächlich zu Eruptionen oder heftigen Schwankungen des Prozeßleuchtens kommt (im Bildmaterial nachgewiesen), aber keine von außen für die Sichtprüfung zu identifizierende Fehlstelle in der Schweißnaht auftritt. Es gibt zahlreiche Hinweise (Beispiele in den Abbildungen 4.24) auf die Möglichkeit, daß der Prozeß vorhandene Löcher selbst wieder als Hohlraum unter einer dünnen Metallschicht versteckt oder solche Stellen, wenn sie im Überschweißbereich liegen, schlichtweg durch erneutes Verschweißen verdeckt werden (Abbildung 4.8 rechts unten). Für eine systematische Suche, die Bestätigung zum Beispiel durch ein Abschleifen der Nähte und eine Aussage zur Häufigkeit dieses Phänomens, werden momentan ausgeschleuste Teile gesammelt und zurückbehalten.

Die Ergebnisse zeigen auch, daß mit dem Verfahren tatsächlich ein Einblick und Ursachenverständnis des Schweißprozesses an sich möglich ist: Flugzeitberechnungen wie in den Abbildungen 4.26 und 4.27 werden von [Schmidt, 2003] noch ausgebaut: anhand der in den Bilddaten und aus den Spritzergeometrien ermittelten Fluggeschwindigkeiten und Massen der kugelförmig angenommenen Spritzer lassen sich kinetische Energien und Wärmemengen ermitteln, die als Spritzer ausgeworfen der Schweißnaht und damit ihrer Qualität fehlen.

4.17 Vergleich mit anderen Systemen

Im industriellen Bereich eingesetzte zerstörungsfreie Prüfungen des Inneren von Produkten sind die Ultraschallprüfung, die Wirbelstromprüfung oder die Analyse mit Röntgenlicht (Röntgen-Computertomographie). Beide Verfahren erreichen für die in unserer Applikation auf Kleinbauteilen entstehenden Defekte in der Größenordnung 0,1–1 mm nicht das nötige Auflösungsvermögen zu wirtschaftli-

chen Kosten einer 100%-Prüfung.

Systeme zur online-Prozeßüberwachung von Laserschweißungen werden auf dem europäischen Markt von den Unternehmen Jurca, Nordmann und Prometec angeboten. Die Systeme von Jurca und Nordmann setzen dabei auf räumlich integrierende Sensoren, deren Daten mit Hüllkurven, Trends und Ausreißerdetektoren auf Abweichungen überwacht werden. Solche Systeme waren im Rahmen von ausgedehnten Machbarkeitsstudien an den hier beschriebenen Laserstationen montiert und zeigten in der industriellen Praxis vor allem Probleme mit dem hohen Justageaufwand zahlreichen Parameter angesichts der großen Typteile-Vielfalt und häufiger Umrüstungen der betroffenen Produktionslinien. In den festen Schwellwerten wird jeweils ein Erfahrungsschatz kodiert, der für andere Teiletypen oder neue Chargen nicht mehr gültig ist. Eine sichere Erkennung der Qualitätsausfälle auf der Schweißnaht konnte mit diesen Systemen bisher nicht realisiert werden. Den wesentlichen Grund dafür kann man in der Abbildung 4.28 ablesen, wo ein räumlich integrierender Sensor durch die räumliche Mittelwertbildung über alle Einzelbilder auf der in diesem Kapitel beispielhaft vorgestellten Bildfolge simuliert wurde (vergleiche auch Abbildung 4.11): eine sichere Detektion der beiden Spritzer ist gegen die zahlreichen Aufhellungen nicht möglich. Differenzbildverfahren konnten wir bereits durch genauere Betrachtungen des Rauschens, prozeßtypischer Variationen und einem Vergleich zur Stärke der tatsächlich gesuchten Änderungen in Abschnitt 4.11 ausschließen. In Abbildung 4.11 ist darüber hinaus der mittlere Verlauf eines solchen Differenzbildes aufgetragen: auch hier treten die gesuchten Ereignisse aus der Prozeßruhe und dem Pixelrauschen des Sensors nicht ausreichend hervor.

Prometec bietet seit Anfang 2001 ein kameragestützt bildverarbeitendes System zur Schweißüberwachung von allen drei gängigen Lasertypen mit einer optisch auf dem Laserstrahl eingekoppelten CMOS-Kamera (120 dB) und geht damit systemtechnisch den hier in dieser Arbeit umgesetzten Weg. Es werden laut Hersteller bis zu acht Schweißmerkmale (Einschweißtiefe, Durchschweißung, Löcher, Spritzer, Schweißposition, Fokusslage, ...) verfolgt und bewertet [Beersiek, 1999] [Beersiek, 2001]. Erfahrungen aus der Praxis oder ein Nachweis der Fertigungstauglichkeit wie in dieser Arbeit waren uns nicht zugänglich. Nordbruch [Nordbruch, Stefan u. a., 2000] stellte ein ähnliches System vor, das auch den HDRC-Sensor zur Prozeßüberwachung beim Laserschweißen in Laborversuchen ohne Taktzeitbegrenzungen einsetzt. Dabei wurde auf das Hochgeschwindigkeitsmerkmal sowie den vollen Dynamikumfang der HDRC-Kamera verzichtet und Vollbilder mit sehr geringer Bildrate und in 8 bit aufgenommen. Weitere Unterschiede zum hier gezeigten Verfahren sind die Beschränkung auf Einzelbilder in der Verarbeitung (es werden vier Merkmale extrahiert, die aus Lauflängendarstellungen von segmentierten Objekten resultieren) und die Verwendung von standardisierten Softwarebibliotheken für Einzelbildverarbeitung im industriellen Umfeld.

Das vorgestellte System (siehe auch [Schmidt, 2003]) wurde insbesondere unter dem Aspekt der Robustheit gegenüber Chargendriften (es finden keine Mittelungen über die Dimension n der Teileanzahl statt: jedes Bauteil wird unabhängig von allen anderen bewertet) in Deutschland patentiert [Brocke, Martin u. a., 2002] und ist in vielen Produktionslinien dauerhaft im Einsatz. Schon heute ist absehbar, daß in Zukunft Überwachungssysteme, die herkömmliche integrierende Sensorik und raumzeitlich auflösende Sensorik zusammenführen (engl. *sensor fusion*), eingesetzt werden, um noch mehr Information aus dem Prozeß zu sammeln. Zahlreiche Ansätze dafür [Blom, A. H. M u. a., 1999] [Bassel, 2001] werden entwickelt. Fertigungstaugliche Lösungen stehen aber auch Ende 2002 noch nicht bereit.

Kapitel 5

Ausreißerdetektion in Thermographiedaten

Für die zweite Anwendung der entwickelten Detektion von Ereignissen greifen wir auf Meßdaten zurück, die auf Forschungs Expeditionen des Institutes für Umweltphysik der Universität Heidelberg bei Feldmessungen gewonnen wurden. Sie wurden in [Garbe, 2001] ausgewertet und in den Kontext von Laborergebnissen vom Heidelberger Aeolotron (Wind-Wellenkanal) gestellt. Sowohl bei den Bildfolgen aus dem Labor wie denen aus Feldmessungen treten nach der Kalibration Regionen schlagartig erhöhter Bildintensitäten auf, die schnell und unter Veränderung ihrer unregelmäßigen, wenig kompakten Form durch das Bildfeld wandern. Diese Bereiche in einer Bildvorverarbeitung sicher zu detektieren, so daß die eigentlichen Auswertungen solche Pixel nicht mehr miteinbeziehen, das ist die hier an die Ausreißerdetektion gestellte Aufgabe.

5.1 Übersicht

Im ersten Abschnitt 5.2 führen wir in die Thematik des Gasaustausches an Meeresoberflächen ein und geben Gründe an für das Interesse an Modellen, Parametrisierungen und Transferraten, welche die aufwendigen Labor- und Felduntersuchungen rechtfertigen. Die Sensorik und Datengewinnung sind in Abschnitt 5.3 beschrieben, woraus sich das Zustandekommen der Reflektionen (Abschnitt 5.4) verstehen läßt. Eine Übersicht zu Verfahren der Auswertung dieser Daten durch zahlreiche andere Autoren wird in Abschnitt 5.5 sehr kompakt zusammengestellt – vor allem, um ein Verständnis für die Problematik der Ausreißer in den Daten zu vermitteln und die Notwendigkeit bisheriger robuster Parameterschätzer oder der hier zusätzlich anwendbaren (Abschnitt 5.6) Ausreißerdetektion anschaulich zu machen. Nachdem die Bildfolgen in Abschnitt 5.7 in stückweise stationäre Szenen eingeteilt wurden, führen wir in Abschnitt 5.8 die in dieser Arbeit vorgeschlagene Detektion durch und erläutern abschließend die Ergebnisse und Grenzen (Abschnitt 5.9) des Verfahrens.

5.2 Motivation

Zur Vorhersage globaler Klimaveränderungen spielt die genaue Abschätzung des CO₂-Haushaltes und eine Kenntnis der Quellen, Senken und zugehörigen Transferraten eine entscheidende Rolle. Die globalen mittleren Jahrestemperaturen zeigen über die letzten 200.000 Jahre eine klare Korrelation mit dem CO₂-Gehalt der Atmosphäre. Der Anstieg der Konzentration des atmosphärischen CO₂ durch Verbrennung fossiler Energieträger seit Beginn der industriellen Revolution leitet einen Treibhauseffekt ein, der zu einer globalen Erwärmung mit drastischen Konsequenzen führen kann. Ozeane stellen das größte Reservoir für gasförmiges CO₂ dar und binden rund 19¹⁸g Kohlenstoff, während Atmosphäre und Biosphäre zusammen gerade ein Zehntel dieser Größe aufnehmen. Eine ganze Reihe von Prozessen (globale Ozeanströmungen und Durchmischungen, biologische Senken, ...) ist in der Lage, einmal im Wasser gelöstes CO₂-Gas aus der Oberflächenschicht der Meere aufzunehmen und zu speichern. Dies unterstreicht, wie wichtig die Kenntnis von CO₂-Transferraten aus der Atmosphäre in das Wasser der Ozeane ist. Darüber hinaus deuten die Abschätzungen der globalen Transferraten in Weltmeere, Atmosphäre und Biosphäre auf eine fehlende Senke für CO₂ von beträchtlicher Größe hin, deren Existenz noch immer nicht zufriedenstellend belegt werden konnte. Ziel ist es daher, solche Transferraten unter unterschiedlichen, typischen Umweltbedingungen wie Oberflächenrauheit, Wasser- und Lufttemperaturen, Luftblasen durch Wellenbrechung, Wind- und Wellengeschwindigkeiten möglichst räumlich-zeitlich auf kleinen Skalen aufgelöst im Labor und unter natürlichen Bedingungen zu messen. Bisher konnten Flußraten nur mit grober zeitlicher und örtlicher Auflösung und mit sehr großer Unsicherheit bestimmt werden. Zwischen 1992 und 2000 wurden Werte $-1,7 \dots - 3,3 \text{ Gt yr}^{-1}$ mit hohen Unsicherheiten für den globalen Kohlenstoff-Transfer veröffentlicht (Übersicht in [Schimpf, 2000]), was vor allem an der jeweils unterschiedlichen Parametrisierung des Problems lag. Die Windgeschwindigkeit gilt als wichtigster Parameter, ohne daß ihr genauer Einfluß voll verstanden wäre. Zum Einfluß der Wassertemperaturen an der Oberfläche, von Wellenfeldern und Oberflächenfilmen stehen Modelle und quantitative Antworten auf die Frage nach deren Einfluß auf die Transferraten noch aus.

Wärme kann als Tracer für den Gasaustausch dienen, da sich Wärmeleitung, Diffusion und Viskosität als Transportphänomene allgemein zusammenfassen lassen und ähnlichen, rein statistischen Gesetzen gehorchen [Jähne, 1990]. Das Verhältnis von Gas-Transferraten und Wärme-Transferraten wird durch das Verhältnis der dimensionslosen Schmidtzahlen und einen durch den mittleren Neigungswinkel der Oberflächenwellen bestimmten Exponenten beschrieben. Die Schmidtzahlen sind zusammengefaßte Materialkonstanten aus kinematischer Viskosität und Diffusions-Koeffizienten für Wasser und das jeweils zur Untersuchung stehende Gas. Über diesen Zusammenhang kann die Transfergeschwindigkeit von Gasen aus dem Wärmetransfer mit einer ausreichenden Genauigkeit für bestimmte Ungenauigkeiten in der Diffusionskonstante und dem Verhältnisexponenten bestimmt werden [Haußecker, H. u. a., 1995]. Mit der Konstantflußmethode (engl. *controlled flux technique*) [Jähne, B. u. a., 1989] [Haußecker, 1996], bei der die Flußrate eines Tracers vorgegeben und der resultierende Konzentrationsunterschied über die viskose Grenzschicht gemessen wird, lassen sich auf kleinen Skalen (zeitlich und räumlich) Transferraten von CO₂ bestimmen, indem Wärme als eingesetzter Tracer und ihre räumlich-zeitliche Veränderung mit Infrarot-Kameras gemessen wird. Bei der passiven Thermographie werden die aufgrund von natürlichen Temperaturdifferenzen vorhandenen Wärme-



Abbildung 5.1: Ozeanoberfläche auf äquatorialen Breiten: die hellen Punkte auf der Wasseroberfläche sind weitgehend auf Reflektionen des Sonnenlichtes zurückzuführen (Reflektion an Oberflächenelementen unter geeignetem Winkel)

flüsse genutzt; aktive Thermographie würde zum Beispiel mit Lasern, die im Infrarot emittieren, die Wasseroberfläche kurzzeitig punktuell aufheizen und Abkühlvorgänge beobachten. Bei Austauschprozessen in der Grenzschicht ergeben sich für Wärme Einstellzeiten im Sekundenbereich, wodurch die Parametrisierung auch mit schnell fluktuierenden meteorologischen Größen (Windgeschwindigkeit, Temperatur) möglich ist.

5.3 Datenaufnahme und Kameratechnik

Die in diesem Abschnitt mit den in Kapitel 3 vorgestellten Methoden analysierten Daten stammen von der GasExII-Forschungsexpedition 2001. Dabei wurden bei unterschiedlichsten meteorologischen Verhältnissen Felddaten in Form von Infrarot-Aufnahmen der Ozeanoberfläche aus äquatorialen Breiten des Pazifik, weitab von Küsten oder Inseln, gewonnen. Dieser Meeresbereich zeichnet sich dadurch aus, daß gemäßigte Windbedingungen und ein ständiges Aufströmen kalter Wassermassen aus der Tiefsee für eine gleichmäßige Grenzschicht sorgen und eine CO₂-Quelle vorhanden sind.

Die Datenaufnahme fand von einem autonomen Katamaran aus statt, an Bord dessen sich Kamera, Datenträger, Instrumente für meteorologische Beobachtungen (Windgeschwindigkeit) und zur Bestimmung von Oberflächenfilmen, Geräte zur Messung der mittleren Wellenneigung sowie Stromversorgung, Ortsbestimmung, Steuerung und Funkverbindung befanden. Der Katamaran wurde rund ein Dutzend mal gegen Sonnenuntergang ausgesetzt und führte seine rund vierstündigen Messungen in Sichtweite vor dem Mutterschiff aus.

Die eingesetzte Kamera (Typ: Galileo, Hersteller: Amber) arbeitet auf InSb-Halbleiterbasis und nutzt wie eine CCD-Kamera den Photoeffekt aus (zu bildgebenden Infrarotsensoren siehe auch [Hau-Becker, 1996]). Durch Kühlung des Sensors mit einem Stirling-Kühler auf eine Temperatur von rund 80 K wird die Eigenstrahlung des Sensors vermindert, und es reduziert sich nach dem Stefan-Boltzmann-Gesetz die pro Zeiteinheit einfallende Photonenzahl auf ein vom Detektor leicht zu

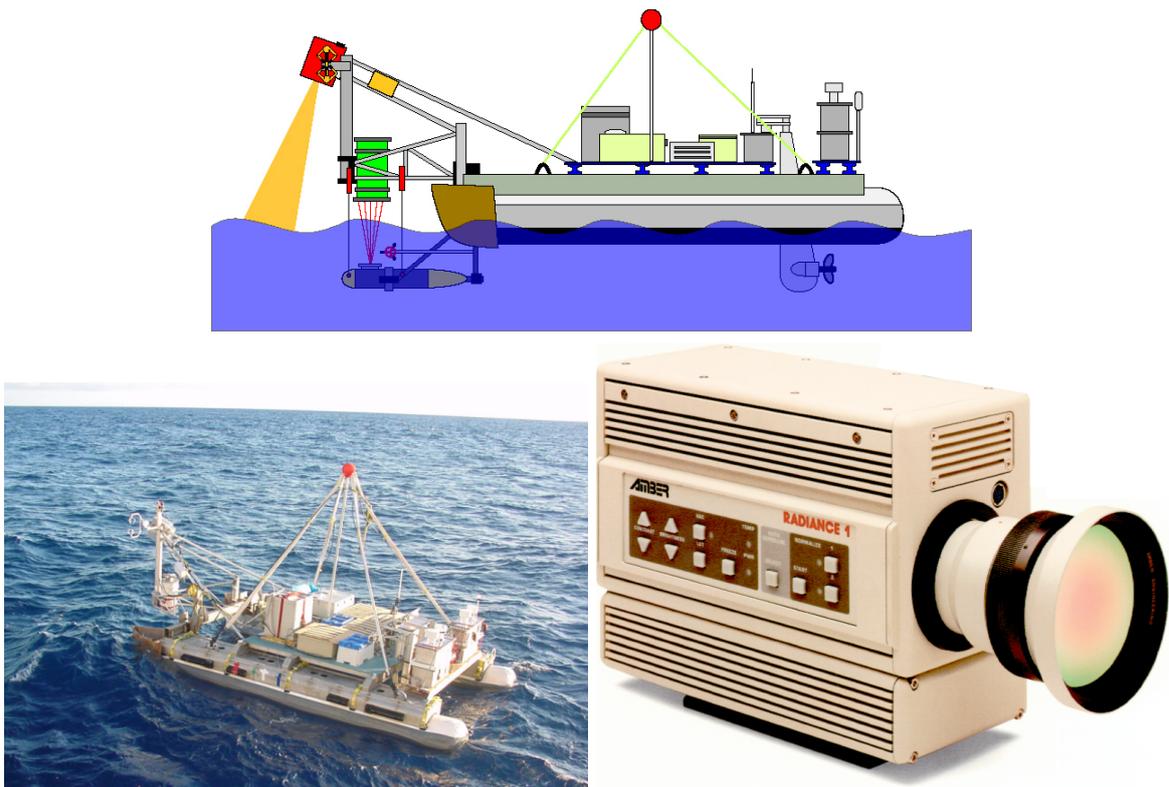


Abbildung 5.2: Versuchsaufbau für die Feldmessungen der Wärmeflüsse auf Meeresoberflächen **Oben:** schematischer Aufbau des Katamarans mit dem Ausleger für die IR-Kamera **Unten links:** Der Katamaran kurz vor Beginn einer Messung: Gut sichtbar der Ausleger mit dem Gehäuse für die Kamera, das durch einen konstanten Luftstrom von der Kamera weg die Optik vor Spritzwasser schützt. **Unten rechts:** eine der eingesetzten IR-Kameras von Amber

verarbeitendes Maß. Diese Kamera hat ein $NE\Delta T$ (engl. *noise equivalent temperature range*) von 28 mK; dies ist die Temperaturdifferenz, die von der Kamera gerade noch aufgelöst werden kann; die zu beobachtenden Fluktuationen liegen im Bereich von 0,01–0,3 K. Aus Bildfolgen homogener temperierter Flächen ermittelt man ein Signal-Rausch-Verhältnis von 30 dB. Das kalibrierte Gesamtsystem [Schimpf, Uwe u. a., 1999] kann relative Temperaturschwankungen von 0,028 K unterscheiden und absolute Temperaturen auf 0,1 K genau messen. Zur radiometrischen Kalibration der Kamera und Bildfolgen beschreiben [Schimpf, 2000] und [Garbe, 2001] [Chrzanowski, Krzysztof u. a., 2001] eine Vorgehensweise, deren Ergebnisse mit den Herstellerangaben sehr gut übereinstimmen. Der erfaßte Spektralbereich der Kamera ist 3–5 μm . Für Raumtemperatur ist der thermische Kontrast hier weitaus höher als für das 8–12 μm -Band, in dem die Atmosphäre ebenfalls durchlässig ist. Ursache dafür ist die hier noch hohe Steigung der Planck-Kurve, während im zweiten Band das Maximum der Planck-Kurve bei rund 20 °C erreicht ist. Die Dicke der wasserseitigen Grenzschicht beträgt 300 μm bis 1 mm, so daß trotz der Integration über die Eindringtiefe der Infrarot-Kamera von 1 μm bis 90 μm im 3–5 μm -Band die tatsächliche Oberflächentemperatur gemessen wird.

Die Bildfolgen liegen bereits kalibriert als 256 × 256 große Bilder mit 512 Bildern pro Folge vor. Die Bildrate beträgt 100 fps für die ursprünglich mit 12 bit (4096 Graustufen) aufgezeichneten Daten. Die Kamera blickt aus rund 2 m Höhe unter 30 ° Zenitwinkel auf die Wasseroberfläche. Die Wahl dieses Winkels begründet sich mit der geringen mittleren Reflektivität unter diesem Winkel. Horizontal polarisierte Reflektionen erreichen ein Minimum der Reflektivität $\rho_{||}$ beim Brewsterwinkel (57 °). Der Mittelwert über horizontale und vertikale Polarisation bleibt bis rund 35 ° Zenitwinkel konstant und steigt erst danach an. Mit einem 50 mm-Objektiv deckt die Kamera ein Bildfeld von circa 40 cm × 40 cm ab, was zu einer Auflösung von 1,6 mm/Pixel führt.

5.4 Reflektionen als Ausreißer

Die Reflektivität der Wasseroberfläche führt sowohl im Labor als auch bei Feldmessungen zu Reflexen, wie sie aus dem sichtbaren Spektrum (Abbildung 5.1) gut bekannt sind. Diese rühren von Reflektionen des Sonnenlichtes. Auch bei den stets nachts durchgeführten Messungen im infraroten Spektralbereich kommt es zu solchen Reflexen bei nicht homogener Wolkendecke (Abbildung 5.3). Physikalische Ursache dafür ist, daß die Wasseroberfläche kein idealer schwarzer Strahler (“*grauer Strahler*”) ist. Die auch als “*non-blackness*” bekannte Reflektivität einer Wasseroberfläche vermindert die emittierte Strahlung im Vergleich zum idealen schwarzen Strahler. Bei gleicher Temperatur wird das Verhältnis der Emission eines realen Körpers zur Emission eines idealen schwarzen Körpers durch die spektrale Emissivität $\epsilon(\lambda)$ in Abhängigkeit von der Wellenlänge λ beschrieben. Aus der Energieerhaltung ergibt sich, daß die Summe der drei Beiträge spektrale Transmissivität $\tau(\lambda)$ (Stahlungsanteil, der nicht absorbiert wird, sondern den Körper durchdringt) spektrale Emissivität und spektrale Reflektivität $\rho(\lambda)$ zusammen den Wert 1 annimmt. Für Wasser gilt im Spektralband 7–15 μm der Wert $\rho = 0,0136$. Die dazu äquivalente Schwarzkörpertemperatur des klaren Himmels liegt ungefähr 60 K unter der Lufttemperatur, bei bewölktem Himmel ist der Wert noch kleiner. Wenn man den Anteil an Himmelsstrahlung, der an der Meeresoberfläche reflektiert wird, nicht mit in die Temperaturbestimmung einbezieht, so ergibt sich bei klarem Himmel ein Fehler von 0,5 K, bei Wolken rund 0,1–0,55 K. Bei durchgehend klarem Himmel oder durchgehend homogener Bewölkung

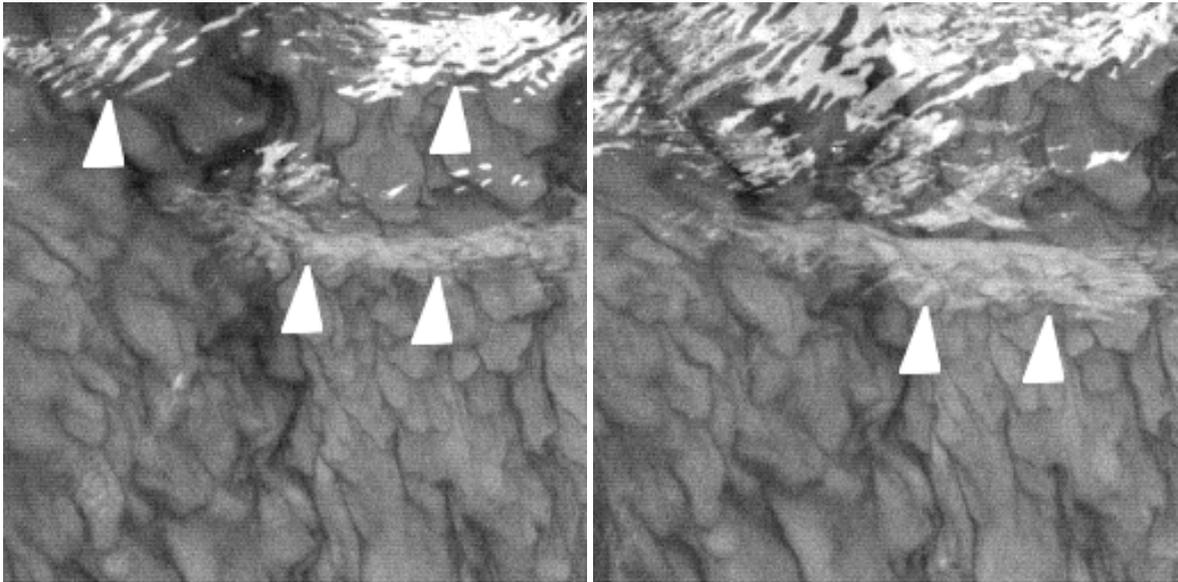


Abbildung 5.3: Pfeile markieren die Reflektionen, die sich in zwei Typen unterscheiden lassen

tritt also zu den Meßdaten ein Offset in der beschriebenen Größe hinzu. Liegt eine Mischbedeckung vor (nicht geschlossene Wolkendecke, sehr inhomogene Bewölkung), so führt dies zu additiven Offsets von Teilen des Himmels, die als Reflektionen an solchen Stellen durch das Bild laufen, deren Wasseroberflächenneigung den passenden Winkel zur Inhomogenität hat. Die genauen Temperaturverschiebungen für verschiedene Bewölkungstypen tabelliert [Haußecker, 1996].

In den Bilddaten (Abbildung 5.3) lassen sich folglich zwei Typen von Reflexen ausmachen: zum einen erkennt man in den kurz aufeinanderfolgend aufgenommenen Daten sehr helle Reflektionen (linke Abbildung, oberer Bildrand). Dabei handelt es sich wahrscheinlich um direkte Reflektionen von Wärmestrahlung zum Beispiel durch Aufbauten des Versuches oder andere warme Zonen der Umgebung und des Himmels. Ein zweiter Typ äußert sich in den kalibrierten Bilddaten nicht als “Überläufe in hohe Grauwerte”, sondern als Region, in der additiv ein wahrscheinlich homogener Grauwert zum Wärmemuster der Wasseroberfläche hinzutritt. Solche Effekte kommen bei Bilddaten aus mondlosen, klaren Nächten oder Nächten bei gleichmäßiger Wolkendecke erheblich seltener vor. Die Reflexe laufen mit der Phasengeschwindigkeit (Orte gleichen Winkels) durch das Bild, was zunächst einmal unabhängig von der generellen Bewegung des Mediums und damit seines Wärmemusters ist (Abbildung 5.4).

5.5 Bestimmung des Wärmeflusses

Die Bestimmung des Wärmeflusses in den gegebenen Daten wird ausführlich in [Garbe, 2001], [Garbe, C. S. u. a., 2001], [Garbe, C. S. und Jähne, B., 2001] und [Garbe, Christoph S. u. a., 2002] (in Planung: [Brocke, 2003]) vorgestellt und soll an dieser Stelle nur kurz zusammengefaßt werden, um die robusten Komponenten des Verfahrens vorzustellen und aufzuzeigen, wie Änderungskarten aus

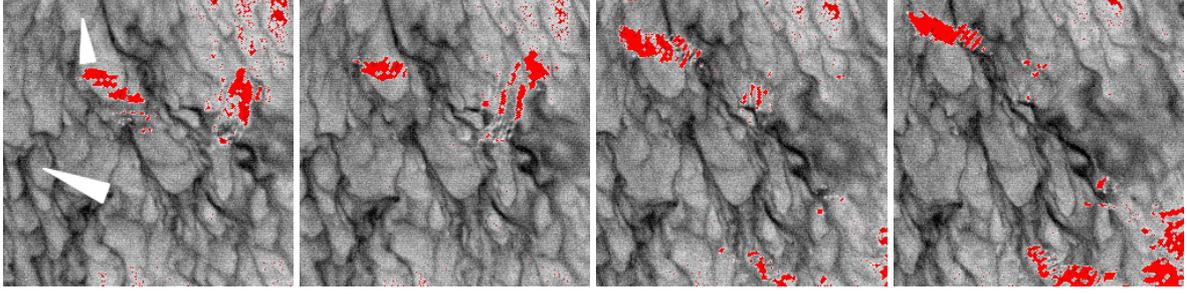


Abbildung 5.4: Zonen mit Reflektionen und das generelle Wärmemuster der Wasseroberfläche bewegen sich mit unterschiedlichen Geschwindigkeiten (Richtung und Beträge) durch das Bild (vier aufeinanderfolgende Aufnahmen): die rot markierten Zonen der Reflektionen bewegen sich nach oben, während die Flüssigkeit sich eher nach links bewegt.

einer Vorverarbeitung der Daten in dieses Schema eingebaut werden können.

Aus der Kontinuitätsgleichung und den beiden Fick'schen Gesetzen folgt für Transportphänomene (Diffusion einer Konzentration c) allgemein (erste Zeile) und für die Wärmemenge $V\rho c_p T$ bzw. Temperatur T als spezielle Größe (zweite Zeile)

$$\begin{aligned} \mathbf{j} &= -D\nabla c & \frac{d}{dt}c &= D\Delta c = -\nabla\mathbf{j} \\ \mathbf{j}_h &= -D_h\rho c_p\nabla T & c_p\rho\frac{d}{dt}T &= D_h\Delta T = -\nabla\mathbf{j}_h \end{aligned}$$

Für zwei Temperaturen $T_1(z_1)$ und $T_2(z_2)$ in unterschiedlichen Tiefen z_1 und z_2 ergibt sich ein Ausdruck, der die Definition einer Transfargeschwindigkeit k_h (engl. *piston velocity*) nahelegt.

$$\mathbf{j}_h = -D_h\rho c_p \frac{T_1(z_1) - T_2(z_2)}{z_1 - z_2} = -k_h\rho c_p\Delta T \quad \text{mit} \quad k_h = \frac{j_h}{\rho c_p\Delta T}$$

Lösungen für die Transportgleichung der Wärmemenge bzw. Temperatur können unter bestimmten Randbedingungen angegeben werden. Mit den Temperaturen des Wasserkörpers, der Wasseroberfläche und den Randbedingungen $T(z, t=0) = T_{\text{bulk}}$ und $T(z=0, t) = T_{\text{surf}}$ erhält man für den Wärmefluß \mathbf{j}_h (siehe [Crank, 1975])

$$j_h(0, t) = -D_h\rho c_p \frac{T_{\text{surf}} - T_{\text{bulk}}}{\sqrt{\pi D_h t}}$$

An der Meeresoberfläche gilt eher die Randbedingung eines konstanten Wärmeflusses, was zu

$$\frac{\partial (T_{\text{surf}} - T_{\text{bulk}})}{\partial z} = -\frac{j_h}{D_h\rho c_p}$$

führt. Mit Analogieschlüssen (siehe [Garbe, 2001], vollständige Ableitung bei [Liu, W. T. und Businger, J. A., 1985] [Soloviev, A. V. und Schlüssel, P., 1994]) folgt aus der vorigen Lösung für diese Randbedingung folgenden Zusammenhang ab:

$$j_h(0, t) = \frac{T_{\text{surf}} - T_{\text{bulk}}}{\alpha\sqrt{t}} \quad \text{mit} \quad \alpha = \frac{2}{\sqrt{\pi D_h\rho c_p}}$$

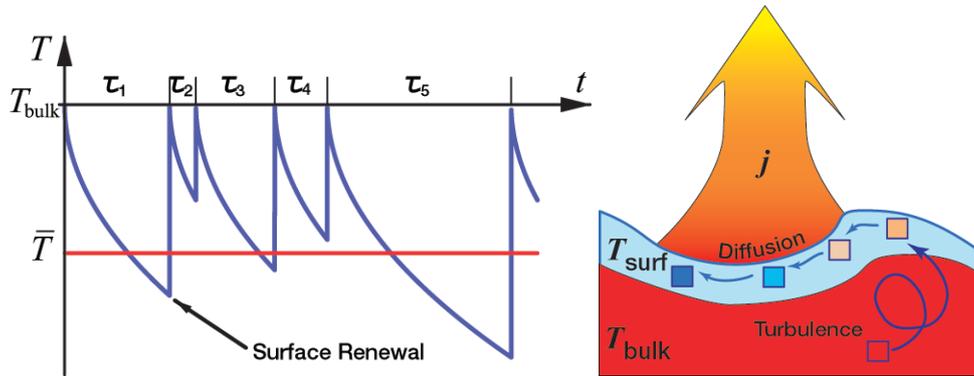


Abbildung 5.5: Links: Statistisch verteilte Erneuerungsereignisse bringen die Oberflächentemperatur in unregelmäßigen Abständen auf die Temperatur des kälteren Wasserkörpers unterhalb der Grenzschicht. Die mittlere Temperatur ergibt sich aus der aus der Häufigkeit dieser Ereignisse, der Temperatur T_{textbulk} und dem Auskühlungsverhalten an der Oberfläche. Rechts: Schematische Darstellung des Oberflächenerneuerungsmodells

oder umgeformt

$$T_{\text{surf}} - T_{\text{bulk}} = \alpha j_h \sqrt{t - t_0} \quad \text{mit: } t \geq t_0$$

Differenzieren nach der Zeit t und erneutes Einsetzen in die Ausgangsgleichung führt schließlich zu

$$j_h = \frac{\sqrt{2}}{\alpha} \sqrt{(T_{\text{surf}} - T_{\text{bulk}}) \frac{d}{dt} (T_{\text{surf}} - T_{\text{bulk}})}$$

Als Tracer für den Gasaustausch kann der Wärmefluß an der Oberfläche also bestimmt werden, indem $\Delta T := T_{\text{surf}} - T_{\text{bulk}}$ und die totale zeitliche Ableitung von ΔT bestimmt werden. Beides ist mit Bildfolgen räumlich-zeitlich aufgelöst möglich. Aus dem Wärmefluß j_h ergibt sich die Transfergeschwindigkeit k_h , die über die Schmidtzahlen der Medien und den Schmidtzahlexponenten zur Transfergeschwindigkeit für CO_2 -Gas führt. Bevor die Implementierung der Bestimmung von ΔT und der totalen zeitlichen Ableitung davon über die Verarbeitung der IR-Sequenzen und damit auch die dafür genutzte Methode der Ausreißerdetektion vorgestellt wird, soll zuvor das Modell der zeitlichen Verteilung der Oberflächentemperaturen vorgestellt werden.

Für das Zustandekommen der zeitlichen Wärmeverteilung an der Wasseroberfläche gibt es zahlreiche Modelle (Übersicht in [Jähne, 1980]) für den turbulenten Wärmetransport vom Wasserkörper zur Oberfläche, von denen das Oberflächenerneuerungsmodell [Danckwerts, 1970] mit statistisch verteilten Erneuerungen die möglichen Exponenten der Schmidtzahl korrekt vorhersagt. Dabei wird angenommen, daß kleine Wasservolumina in der Grenzschicht an der Oberfläche in zufällig verteilten Zeitabständen mit Wasser aus dem tieferen Wasserkörper durchmischt und auf dessen Temperatur gebracht werden. Zusätzlich kommt es zu einer diffusiven Auskühlung der Wasseroberfläche (siehe Abbildung 5.5). Dem Modell zugrunde liegt eine Verteilungsfunktion $p(t)$ der Zeitabstände zwischen solchen aufeinanderfolgenden Erneuerungsereignissen für ein Volumenelement. Die Verteilungsfunktion wurde von [Soloviev, A. V. und Schlüssel, P., 1994] als ln-Normalverteilung

$$p(t) = \frac{1}{\sqrt{\pi} \sigma_m t / t'} \exp \left\{ -\frac{(\ln(t/t') - m)^2}{\sigma_m^2} \right\}$$

vorgeschlagen, wobei m der Mittelwert der $\ln t/t'$ ist und σ_m^2 die zugehörige Varianz (t' ist ein Skalenfaktor, der für einheitenfreie Größen sorgt). [Haußecker, H. u. a., 2001] belegte diese Verteilung experimentell. Daraus ergibt sich auch eine mittlere Zeitdauer zwischen den Erneuerungen von

$$\langle t \rangle = \int_0^{\infty} p(t) \frac{t}{t'} dt = t' \cdot \exp \left\{ \frac{\sigma_m^2}{4} + m \right\}$$

Zu einer festgehaltenen Erneuerungszeit t berechnet [Haußecker, 1996] eine Verteilung für die Oberflächentemperatur von

$$p(T_{\text{surf}}|t) = \frac{2(T_{\text{surf}} - T_{\text{bulk}})}{t \alpha^2 j_h^2}$$

Unter der Annahme, daß die Erneuerungsereignisse In-normalverteilt sind, ergibt sich daraus eine Oberflächentemperatur-Verteilung von

$$\begin{aligned} p(T_{\text{surf}}) &= \int_0^{\infty} p(T_{\text{surf}}|t) p(t) dt \\ &= \frac{(T_{\text{surf}} - T_{\text{bulk}})^2}{(\alpha j_h)^2} \exp \left\{ \frac{\sigma_m^2}{4} - m \right\} \left\{ 1 - \operatorname{erf} \left[\frac{\sigma_m}{2} - \frac{m}{\sigma_m} + \frac{1}{\sigma_m} \ln \left[\left(\frac{T_{\text{surf}} - T_{\text{bulk}}}{\alpha j_h} \right)^2 \right] \right] \right\} \end{aligned} \quad (5.1)$$

für den Fall $T_{\text{surf}} < T_{\text{bulk}}$. Kurvenanpassungen dieser Funktion an die Histogramme von einzelnen IR-Bildern sind in Abbildung 5.6 gezeigt. Die mittlere Oberflächentemperatur \bar{T}_{surf} wird entweder durch eine einfache Mittelung über das einzelne Bild (oder mehrere aufeinanderfolgende, um eine größere Datenbasis zu haben) berechnet oder über die numerische Berechnung des Integrals

$$\bar{T}_{\text{surf}} = \frac{\int_{-\infty}^{\infty} T_{\text{surf}} p(T_{\text{surf}}) dT_{\text{surf}}}{\int_{-\infty}^{\infty} p(T_{\text{surf}}) dT_{\text{surf}}}$$

Das Modell für die Oberflächen-Erneuerung kann anhand der oben hergeleiteten Bestimmungsgleichung für den Wärmefluß überprüft werden. In

$$t - t_0 = \left(\frac{T_{\text{surf}}(t - t_0) - T_{\text{bulk}}}{\alpha j_h} \right)^2$$

können die beiden Temperaturen aus den Ergebnissen der Analyse der Verteilungen (für einen im Bild konstant angenommenen Wert T_{bulk}) und direkt aus den Daten (für $T_{\text{surf}}(t)$) bestimmt werden. Zur Berechnung der zugehörigen Lebensdauer $\tau := t - t_0$ fehlt nur der Wert des Wärmeflusses j_h . Dieser wird aus der Ableitung der Oberflächentemperatur nach der Zeit

$$j_h = \frac{2}{\alpha} \sqrt{\tau} \frac{d}{d\tau} T_{\text{surf}}(\tau)$$

entnommen, so daß sich die Lebensdauer an jedem Pixel der Sequenz zu

$$\tau = \frac{1}{2} \frac{T_{\text{surf}}(\tau) - T_{\text{bulk}}}{\dot{T}_{\text{surf}}(\tau)}$$

ergibt. Wieder wird die totale zeitliche Ableitung der Temperatur benötigt, zu deren Berechnung eine robuste Einbettung des erweiterten Strukturtenors dient. Trägt man die Häufigkeiten der in einer Sequenz vorkommenden Lebensdauern τ zur Überprüfung des Oberflächenerneuerungsmodelles auf, so muß sich die vom Modell vorhergesagte ln-Normalverteilung ergeben, was [Garbe, Christoph S. u. a., 2002] bestätigen konnte.

Die Bestimmung von ΔT aus kalibrierten IR-Sequenzen geschieht über die Histogramm-Analyse der Infrarotbilder, bei der T_{bulk} durch den Achsenschnittpunkt der Verteilung bildweise als räumlich konstanter Wert bestimmt wird und ΔT sich als Differenz daraus zu den Bilddaten als räumlich-zeitlich aufgelöster Wert ergibt. Zur Bestimmung der totalen zeitlichen Ableitung von ΔT wird eine robuste Einbettung eines Total-Least-Squares-Verfahrens (TLS) des erweiterten Strukturtenors benutzt. In der Bestimmungsgleichung (engl. *brightness change constraint equation, BCCE*) für den optischen Fluß \mathbf{f} [Horn, B. K. P. und Schunk, B., 1981] lassen sich mit der üblichen Schreibweise für partielle Ableitungen $\partial g/\partial i := g_i$ ein Datenvektor \mathbf{d} und ein Parametervektor \mathbf{p} voneinander trennen.

$$\frac{dg}{dt} = \frac{\partial g}{\partial t} + \frac{\partial g}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial g}{\partial y} \frac{\partial y}{\partial t} = g_t + (\mathbf{f}\nabla)g = (g_x, g_y, g_t)^T \cdot (f_1, f_2, 1) = \mathbf{d}^T \cdot \mathbf{p} = 0$$

Das Problem ist schlecht gestellt, weil zwei Größen $\mathbf{f} = (\partial x/\partial t, \partial y/\partial t)^T$ aus nur einer Gleichung zu bestimmen sind. In der Regel benutzt man daher aus einer Umgebung um den betreffenden Datenpunkt insgesamt n Datenpunkte, nimmt den optischen Fluß \mathbf{f} in dieser Umgebung als konstant an und löst dann ein Gleichungssystem aus n Gleichungen, das sich als Matrix $\mathbf{D} \cdot \mathbf{p} = \mathbf{0}$ mit einer $n \times 3$ -Datenmatrix \mathbf{D} schreiben läßt. Bei $n > 2$ wird für das überbestimmte Problem genau der Parametervektor \mathbf{p}^* gesucht, dessen Summe von Residuen-Quadraten (Betragsquadrat des Residuenvektors \mathbf{r})

$$\sum_{i=1}^n (\mathbf{d}_i^T \cdot \mathbf{p})^2 = \sum_{i=1}^n r_i^2 = |\mathbf{r}|^2 \rightarrow \min$$

minimal ist. Vergleicht man die totale zeitliche Ableitung

$$\frac{d\Delta T}{dt} = \frac{\partial \Delta T}{\partial t} + (\mathbf{u}\nabla) \Delta T$$

mit der BCCE, so ist evident, daß ΔT nicht die Grauwertintensität ersetzen kann, da die Bedingung $d/dt \Delta T = 0$ nicht gilt. Ganz im Gegenteil wird gerade dieser Wert ungleich Null gesucht. Unter gleichzeitiger Berücksichtigung einer durch das Geschwindigkeitsfeld \mathbf{u} gegebenen Bewegung des Wärmemusters an der Wasseroberfläche und einer in den zitierten Arbeiten für diese Gleichungen als konstant angenommenen Quell-Rate von Wärme aus dem Wasserkörper an die Oberfläche ist die totale zeitliche Ableitung von ΔT zu bestimmen. Würde die Wasseroberfläche ruhen, so wäre die totale zeitliche Änderung der Temperatur durch die lokale Änderung $\partial/\partial t \Delta T$ an einem festen Ort gegeben und aus den Daten durch Differenzbildung im Bildstapel einfach zu ermitteln. Die BCCE muß also erweitert werden, und die Annahme einer konstanten Änderung c führt zu

$$g_t + (\mathbf{f}\nabla)g - c = (g_x, g_y, g_t, -1)^T \cdot (f_1, f_2, 1, c) = \mathbf{d}^T \cdot \mathbf{p} = 0$$

Für die Änderung der Temperatur kann der Datenvektor den IR-Sequenzen und deren Ableitungen entnommen werden, im Parametervektor entspricht $c \equiv d/dt \Delta T$ der gesuchten Temperaturänderung.

In einer Umgebung mit n Pixeln mit konstant angenommenem Fluß $\mathbf{f} \equiv \mathbf{u}$ wird die Datenmatrix

$$\begin{pmatrix} g_{x,1} & g_{y,1} & g_{t,1} & -1 \\ g_{x,2} & g_{y,2} & g_{t,2} & -1 \\ \vdots & \vdots & \vdots & \vdots \\ g_{x,n} & g_{y,n} & g_{t,n} & -1 \end{pmatrix} \cdot \begin{pmatrix} \delta x \\ \delta y \\ \delta t \\ c \end{pmatrix} = \mathbf{D} \cdot \mathbf{p} = \mathbf{0} \quad \text{mit} \quad \mathbf{D} \in \mathbb{R}^{n \times 4} \quad \text{und} \quad \mathbf{p} \in \mathbb{R}^4$$

Die Minimierung (engl. *total least squares*, *TLS*) kann so umformuliert werden, daß der erweiterte Strukturtensor und sein Eigenwertproblem erkennbar werden.

$$|\mathbf{Dp}| = \int_{-\infty}^{+\infty} w(\mathbf{x} - \mathbf{x}', t - t') \mathbf{p}^T \mathbf{D}^T \mathbf{D} \mathbf{p} \, d\mathbf{x}' dt' = \mathbf{p}^T \mathbf{J} \mathbf{p} \rightarrow \min$$

$$\text{mit} \quad \mathbf{J} = \int_{-\infty}^{+\infty} w(\mathbf{x} - \mathbf{x}', t - t') \mathbf{D}^T \mathbf{D} \, d\mathbf{x}' dt' = \begin{pmatrix} \langle g_x \cdot g_x \rangle_n & \langle g_x \cdot g_y \rangle_n & \langle g_x \cdot g_t \rangle_n & \langle -g_x \rangle_n \\ \langle g_y \cdot g_x \rangle_n & \langle g_y \cdot g_y \rangle_n & \langle g_y \cdot g_t \rangle_n & \langle -g_y \rangle_n \\ \langle g_t \cdot g_x \rangle_n & \langle g_t \cdot g_y \rangle_n & \langle g_t \cdot g_t \rangle_n & \langle -g_t \rangle_n \\ \langle -g_x \rangle_n & \langle -g_y \rangle_n & \langle -g_t \rangle_n & +1 \end{pmatrix}$$

Die Randbedingung $\mathbf{p}^T \mathbf{p} = 1$ zur Vermeidung der trivialen Lösung $\mathbf{p} = \mathbf{0}$ wird mit dem Lagrange-Multiplikator λ in das Funktional

$$\mathcal{L}(\mathbf{p}, \lambda) = \mathbf{p}^T \mathbf{J} \mathbf{p} + \lambda (1 - \mathbf{p}^T \mathbf{p})$$

gekoppelt, so daß sich im Lagrange-Formalismus $\mathbf{p}^* = \text{argmin}_{\mathbf{p}} \mathcal{L}(\mathbf{p}, \lambda)$ aus den $\partial \mathcal{L} / \partial p_j \stackrel{!}{=} 0$ das Gleichungssystem

$$\mathbf{J} \mathbf{p} = \lambda \mathbf{p}$$

als Eigenwertgleichung mit dem kleinsten Eigenwert λ schreibt, der auch der minimale Wert im TLS-Verfahren ist:

$$\min |\mathbf{Dp}| = \mathbf{p}^T \mathbf{J} \mathbf{p} = \mathbf{p}^T \lambda \mathbf{p} = \lambda$$

Gesucht ist also der Eigenvektor \mathbf{p}^* zum kleinsten Eigenwert des erweiterten Strukturensors. Die totale Ableitung $d/dt \Delta T \equiv c = p_4^*/p_3^*$ ergibt sich aus den letzten beiden Komponenten dieses Eigenvektors; das Geschwindigkeitsfeld \mathbf{u} ergibt sich aus $(p_1^*/p_3^*, p_2^*/p_3^*)^T$.

Die Reflexe in den Daten führen zu Ausreißern, die mit einer robusten TLS-Methode durch *Least Median of Squares Orthogonal Distances* (LMSOD) nach [Bab-Hadiashar, Alireza und Suter, David, 1998] unterdrückt werden, so daß nicht alle Werte einer Umgebung in die Bestimmung des Parametervektors \mathbf{p}^* eingehen. Ausreißer werden dabei anhand der Residuen $r = \mathbf{d}^T \mathbf{p}$ ermittelt, für deren Berechnung eine möglichst gute Schätzung von \mathbf{p} bekannt sein muß. Generell werden dabei aus den n aus einer Umgebung zur Verfügung stehenden Daten zur Bildung der Matrix $\mathbf{D} = (\mathbf{d}_1, \dots, \mathbf{d}_n)^T$ solange Unterstichproben gebildet, bis eine Lösung für \mathbf{p} mit möglichst kleinen Residuenquadraten bereitsteht. Dann werden für alle Daten der Umgebung die Residuen erneut berechnet, danach sortiert und für die folgende Berechnung des Parametervektors nur die darin nicht als Ausreißer identifizierten Daten benutzt. Im einzelnen heißt das: es wird eine zufällig ausgewählte Unterstichprobe von $k = 4$

Datensätzen $\hat{\mathbf{D}} = (\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \mathbf{d}_4)^T$ gebildet, für die sich ein Parametervektor $\hat{\mathbf{p}}$ exakt bestimmen läßt. Für alle anderen $i = 1, \dots, n-k$ Daten werden die quadrierten Residuen $r_i^2 = (\mathbf{d}_i^T \hat{\mathbf{p}})^2$ und deren Median $\text{Med}(r_i^2)$ bestimmt. Dies geschieht für

$$m = \frac{\ln(1 - \Pi)}{\ln(1 - (1 - \epsilon)^k)}$$

Unterstichproben, wobei ϵ die geschätzte Rate an Ausreißern ist und Π die sehr nahe 1 angesetzte Wahrscheinlichkeit, daß mindestens eine Unterstichprobe den Parametervektor ausreißerfrei bestimmt hat. Aus allen m initialen Parametervektoren $\hat{\mathbf{p}}$ wird der mit dem zugehörigen kleinsten Medianwert der Residuenquadrate über den Restdatensatzes benutzt, um damit alle $j = 1, \dots, n$ Residuenquadrate $r_j^2 = (\mathbf{d}_j^T \hat{\mathbf{p}}^*)^2$ zu berechnen und die Daten nach deren Größe zu sortieren. Mit

$$\frac{|r_j|}{1,4826 \left(1 + \frac{5}{n-k}\right) \sqrt{\text{Med } r_j^2}} \begin{matrix} \leq \\ > \end{matrix} \begin{matrix} 2,5 \\ 2,5 \end{matrix} \left\{ \begin{array}{l} \text{kein Ausreißer} \\ \text{Ausreißer} \end{array} \right.$$

werden dann Ausreißer identifiziert. Über die nicht als Ausreißer gekennzeichneten n_a Daten (mit $k \leq n_a \leq n$) wird nun abschließend mit der Matrix $\mathbf{D}_a = (\mathbf{d}_1, \dots, \mathbf{d}_{n_a})^T$ das Eigenwertproblem des Strukturtenors $\mathbf{D}_a^T \mathbf{D}_a$ berechnet, dessen Elemente nun über n_a statt n gemittelt werden (es wird eine binominale Filtermaske benutzt).

Rechentechnisch findet in der Verarbeitung der IR-Sequenzen an mehreren Stellen eine Mittelung und Abtastung der Daten statt, bevor die robuste Detektion der durch Reflexe (erscheinen fälschlicherweise als “zu warme” Regionen) eingebrachten Ausreißer diese aus der Parameterschätzung herausnimmt.

Die 256×256 großen Bilder werden bei [Garbe, 2001] auf der ersten Pyramidenstufe verarbeitet und dazu bildweise geglättet mit einem Filterkernel $^{1/16}(1, 4, 6, 4, 1)^T$, um die hohen räumlichen Frequenzen so zu unterdrücken, daß auf der folgenden Pyramidenstufe der Größe 128×128 kein Aliasing-Effekt mehr auftritt [Jähne, 2002]. Auf der neu abgetasteten Pyramidenebene findet eine weitere Glättung mit einem Kernel $^{1/4}(1, 2, 1)^T$ statt. Alle folgenden Rechnungen finden auf dieser Skalenebene statt. Die Ableitungen in der erweiterten BCCE werden mit $5 \times 5 \times 5$ optimiertem Sobelfilter [Schar, 2000] berechnet. Für das TLS Verfahren werden $9 \times 9 \times 3$ Datenpunkte ausgewählt und der Ausreißerdetektion unterzogen. Für das Eigenwertproblem des erweiterten Strukturtenors wird über die verbliebenen $n_a \leq n$ Punkte mit einer binominalen Fensterfunktion w gemittelt. Die Größe der Masken und der Übergang auf eine höhere Pyramidenebene können zur Folge haben, daß eine vorhandene Ausreißerregion mit ihrem additivem Grauwertbeitrag 4 bis 5 Pixel von ihrer Kante aus noch Einfluß auf die Daten hat. Schwierigkeiten können dann entstehen, wenn die für die Ausreißeranalyse benutzte Datenprobe sich dadurch nicht mehr in “gute Daten” und Ausreißer einteilen läßt, sondern auch vermeintlich “gute Daten” durch die Nähe zu einer vorhandenen Ausreißerregion bereits gestört sind.

5.6 Lösungsansatz durch Ausreißerdetektion

Zu einer sicheren Detektion der Reflexe bietet sich das in Kapitel 3 entworfene Verfahren an. Es würde bereits auf den Originaldaten Ausreißer markieren – unabhängig von der Physik der Temperaturdif-

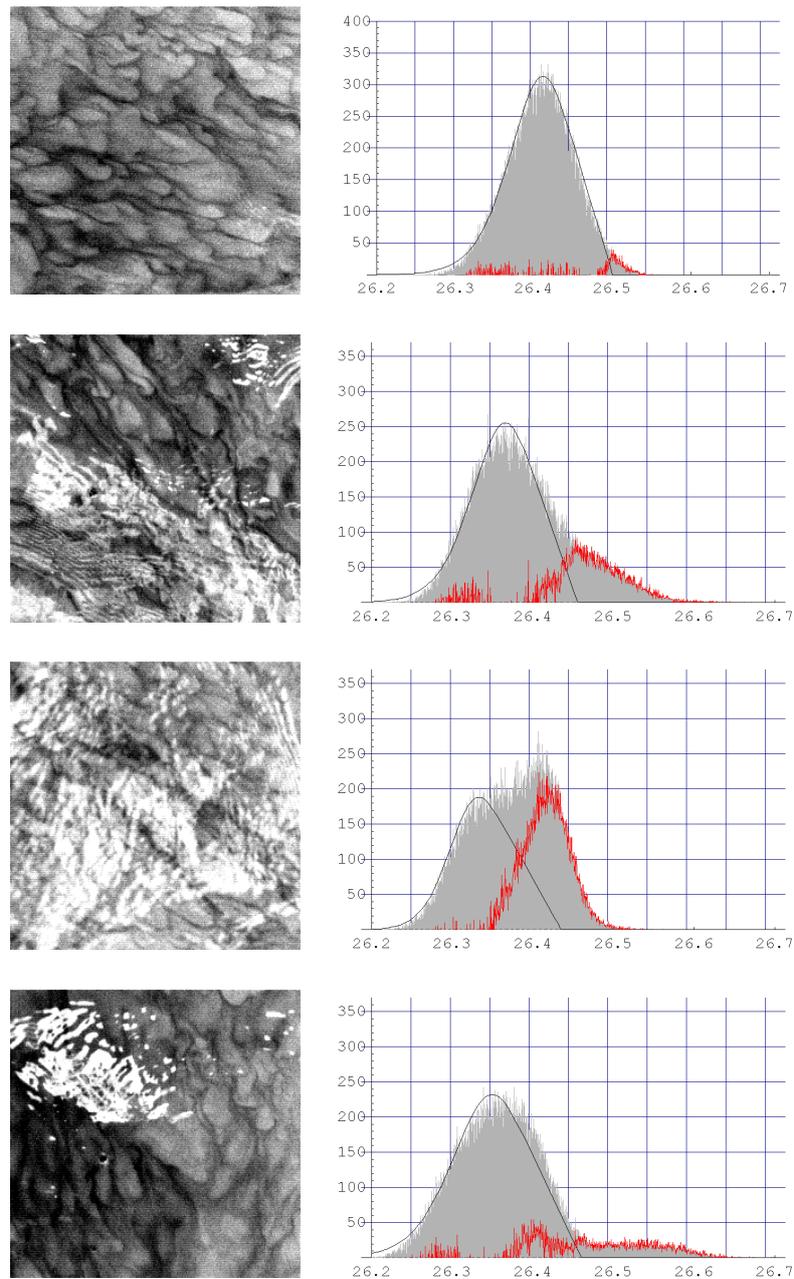


Abbildung 5.6: Einzelbilder und deren Grauwertverteilungen (grau), Kurvenfit nach (5.1) der Oberflächentemperaturen T_{surf} (schwarz) und positiver Anteil der Differenz beider Kurven (rot). **Erstes Bildpaar:** Grauwertverteilung und Histogramm eines Einzelbildes, in dem keine Reflexe auftreten. Auch bei solchen Bildern gibt es durch das hinzukommende Rauschen bei $\approx T_{bulk}$ etwas häufiger Intensitäten, als vom Modell vorhergesagt. **Zweites Bildpaar:** Reflexionen vor allem im unteren Bildbereich führen zu einem breiteren Auslaufen der Häufigkeiten im hellen Bereich. **Drittes Bildpaar:** Reflexionen und tatsächliche Wärmedaten der Oberfläche führen zu zwei dicht überlagerten Peaks im Histogramm, die sich durch einfache Schwellwertoperationen auf den Grauwerten nicht trennen lassen. **Viertes Bildpaar:** Sehr helle Reflexe ließen sich durch einen Schwellwert abtrennen, dennoch läßt das Histogramm gut erkennen, daß zwischen den Häufungen der hellen Reflexionen und den echten Wärmedaten noch weitere schwächere Aufhellungen liegen.

ferenzen und zugrunde liegenden Modellen der Oberflächentemperatur und ihres zeitlichen Verlaufes. Einmal markierte Regionen können dann bereits beim Übergang auf eine höhere Pyramidenstufe durch geeignete Gewichte in den Mittelungen weniger berücksichtigt und aus den Ableitungen zur Konstruktion des Datenvektors \mathbf{d}^T ganz herausgenommen werden. Eine sicher weiterhin empfehlenswerte Ausreißerdetektion mittels LMSOD kann dann in erheblich engerem Rahmen (von vorne herein kleineren Residuen) arbeiten und eine Analyse auf Daten durchführen, die erheblich besser auf das zugrundeliegende Modell passen.

In Abbildung 5.6 ist deutlich zu erkennen, daß im Einzelbild der Kurvenfit und damit die Bestimmung von T_{bulk} als Achsenschnittpunkt nicht gelingt, wenn Reflexe im Bild vorhanden sind. Das beschriebene additive Verhalten wird ebenfalls gut deutlich, denn die Verteilung läßt sich anpassen an eine Kurve, die aus der ursprünglichen Verteilung ohne Reflexe und einer konstant verschobenen zufällig aus der ursprünglichen Verteilung gezogenen Unterverteilung gebildet wird. Es wird auch deutlich, daß eine einfache Schwellwertbildung auf den Grauwertintensitäten der Ausgangsdaten die Ausreißer nicht separieren kann, ohne die Wärmeverteilung und damit die Bestimmung von T_{bulk} und $T_{\text{surf}} - T_{\text{bulk}}$ erheblich zu stören. Der additive, generell positive Beitrag zu einer Unterstichprobe ist gemessen am gesamten Wertebereich und verglichen mit den genannten Variationen durch Rauschen der Sensorik und lokaler Varianz der Grauwerte im Einzelbild groß genug, um von der Ausreißerdetektion gefunden zu werden.

5.7 Konstruktion stationärer Szenenstücke

Für die vorgeschlagene Änderungsdetektion ist nicht nur die lokale Verteilung der Intensitäten in einem Einzelbild von Bedeutung. Eine ganz entscheidende Voraussetzung ist, daß die zu untersuchenden Szenen stückweise stationär sind oder sich sinnvoll in solche Stücke einteilen lassen. Bereits in Abbildung 5.4 wurde deutlich, daß bei den vorliegenden Daten mit ihren beiden überlagerten Bewegungsmustern (Wärmesignatur der Wasseroberfläche und Reflektionswinkel der Wellen zu passenden Strahlungs-Inhomogenitäten der Umgebung) kaum von einem Hintergrund und einem Vordergrund zu sprechen sein kann. Um die Voraussetzung des vorgeschlagenen Verfahrens zur Detektion der Reflexe zu erfüllen, wird es also ganz entscheidend sein, stückweise stationäre Szenen bereitzustellen. Dies kann zuerst einmal nur durch eine Beschränkung auf kleine Ausschnitte Δt_i erreicht werden. Die Untersuchung in Abbildung 5.7 zeigt, wie sich die Intensitäten im Mittelwertbild $\langle g \rangle_{\Delta t}$ gegenüber unterschiedlicher Szenenlänge Δt verhalten. Als Maß dafür dient die Veränderung des über Δt gemittelten Szenenstückes. Dieses Mittelwertbild $\langle g \rangle_{\Delta t}$ sollte für stationäre Szenen gleich bleiben; seine lokalen Variationen

$$\left\langle \left(\langle g \rangle_{\Delta t} - \langle \langle g \rangle_{\Delta t} \rangle_{\Delta X, \Delta Y} \right)^2 \right\rangle_{\Delta X, \Delta Y}^{1/2}$$

sollten also für verschieden große Δt gleich sein. Es wurde über Stücke der Länge $\Delta t = 2 \dots 32$ gemittelt und die lokale Varianz der Mittelwertbilder mit der lokalen Varianz des ersten Bildes

$$\left\langle \left(g(t_0) - \langle g(t_0) \rangle_{\Delta X, \Delta Y} \right)^2 \right\rangle_{\Delta X, \Delta Y}^{1/2}$$

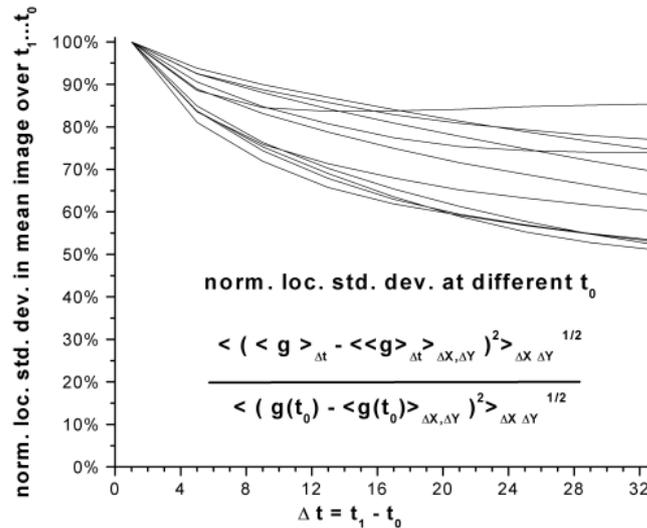


Abbildung 5.7: Die lokalen Varianzen des Mittelwertbildes $\langle g \rangle_{\Delta t}$ für verschiedene Startpunkte t_0 der Mittelung über $\Delta t := t_1 - t_0$ laufen bereits für kleine Windgeschwindigkeiten (hier 2 m/s) schnell gegen 50 % der lokalen Varianz des Bildes $g(t_0)$

verglichen und für Szenenstücke mit verschiedenen Startpunkten t_0 aufgetragen. Man sieht deutlich, daß die Standardabweichung innerhalb der Mittelwertbilder über Δt sofort nach Hinzunahme weiterer Bilder stark fällt: die Mittelwertbilder werden also bereits nach wenigen Bildern durch den sich bewegenden Hintergrund verschmiert. Dieser Effekt ist unterschiedlich stark für einzelne Abschnitte, generell aber mit der Windgeschwindigkeit korreliert. Zum Vergleich wurden synthetische Bilddaten erstellt, die bis auf das bereits quantifizierte Sensorrauschen, das Auskühlungsverhalten der Oberfläche und der Erneuerungsereignisse keine Bewegung enthielten und somit stationär im hier verwendeten Sinn waren. Das verwendete Maß fiel dann bis $\Delta t = 10$ auf rund 97% ab und blieb dort stabil.

Es stellt sich also die Frage, wie in den Bildfolgen Stücke entnommen werden können, die zumindest annähernd stationär sind. Dabei ist $\Delta t = 3$ die kleinste für die Änderungsdetektion rechen-technisch mögliche Szenenlänge. Es hat sich herausgestellt, daß eine Wahl von $\Delta t = 5 \dots 20$ je nach Windgeschwindigkeit die Bewegungsunschärfe des Hintergrundes einerseits vermeidet, aber auch eine möglichst große Stichprobe benutzt.

Diese Einteilung der Bildfolge in gleich große Stücke Δt wird erweitert durch eine gleitende Mittelung: es wird für ein Bild an der Stelle t_0 der Sequenz ein Mittelwertbild und daraus auch ein Varianzbild über den Bereich $(t_0 - \Delta t/2) \dots (t_0 + \Delta t/2)$ berechnet. Der rechentechnische Aufwand darin kann durch geschickte Ablage von Zwischenergebnissen und iterative Verfahren erheblich beschleunigt werden. Eine unterschiedliche Gewichtung der einzelnen Bilder (zum Beispiel im Zentrum des Zeitintervalls stärker als außen) brachte keine entscheidenden Veränderungen. Die Mittelungen in (3.3) (auf Seite 50) lauten dann für das Mittelwertbild und analog die Standardabweichung in der in

dieser Arbeit gebräuchlichen Schreibweise im Unterschied zu(3.1)

$$\langle g(x, y, t) \rangle_{t-\Delta t/2, \Delta t} = \frac{1}{\Delta t} \sum_{t'=t-\Delta t/2}^{t+\Delta t/2} g(x, y, t')$$

Ergebnisse für die mittlere statistische Erneuerungsrate $1/\langle \tau \rangle$ gibt [Garbe, Christoph S. u. a., 2002] für unterschiedliche Windgeschwindigkeiten. Die Werte bewegen sich bei Fehlern von 1...3% zwischen 1,47 s und 30,38 s für Windgeschwindigkeiten von 2...8 m/s. Dies entspricht bei der gegebenen Bildrate 147...3038 Bildern. Aus Plots [Garbe, Christoph S. u. a., 2002] der Häufigkeiten der Erneuerungsereignisse über deren mittlere Zeitdauer entnimmt man, daß die häufigsten Zeitdauern für den angegebenen Bereich von Windgeschwindigkeiten bei 0,2...0,8 s (also 20...800 Bildern) liegt. Unter Berücksichtigung der unterschiedlichen Windgeschwindigkeiten sieht man, daß der Effekt der Grauwertkonvergenz im Vergleich zu Erneuerungsereignissen stets schneller, also bereits in kleineren Bildstapeln, abgeschlossen ist. Die Änderungsdetektion arbeitet also mit Szenenstücken, die kürzer sind als die Erneuerungszeiten.

5.8 Durchführung und Ergebnisse der Ereignisdetektion

Das Verfahren zur Änderungsdetektion (3.3) wie in Abschnitt 3.3 (Seite 49) beschrieben wird nun ohne weitere Veränderungen benutzt. Gearbeitet wird auf den kalibrierten Bilddaten in ihrer vollen Auflösung von 256×256 mit Zeitfenstern von typischerweise $\Delta t = 5 \dots 20$ wie in Abschnitt 5.7 beschrieben. Die iterative Erweiterung (3.6) aus Abschnitt 3.3.6 (Seite 57) kommt ebenfalls zum Einsatz. Es macht dabei wenig Sinn, die initiale Änderungsmaske $b^{(i)}$ angesichts vorhandener Reflektionen mit 1 an allen Orten starten zu lassen. Wir haben jeweils über einen Ausschnitt von $\Delta t_{\text{fit}} = 10$ Bildern (gleitende Fenster) einen Fit der Verteilung (5.1) an die rechte Seite der Grauwertverteilung ausgeführt und damit T_{bulk} über den Achsenschnittpunkt aufgrund mehrerer Bilder ermittelt.

N	$\theta_{\alpha=0,2}$	$\theta_{\alpha=0,1}$	$\theta_{\alpha=0,05}$	N	$\theta_{\alpha=0,2}$	$\theta_{\alpha=0,1}$	$\theta_{\alpha=0,05}$
3	1,406	1,412	1,414	13	2,264	2,426	2,563
4	1,645	1,689	1,710	14	2,297	2,461	2,602
5	1,791	1,869	1,917	15	2,327	2,494	2,638
6	1,894	1,996	2,067	16	2,354	2,523	2,670
7	1,974	2,093	2,182	17	2,380	2,551	2,701
8	2,041	2,172	2,273	18	2,404	2,577	2,728
9	2,097	2,238	2,349	19	2,426	2,601	2,754
10	2,126	2,294	2,414	20	2,447	2,623	2,779
11	2,190	2,343	2,470	25	2,537	2,718	2,880
12	2,229	2,387	2,519	30	2,609	2,792	2,958

Tabelle 5.1: Schwellwerte $\theta_{\alpha;N}$ in Abhängigkeit der Signifikanz α und der Stichprobengröße N für die verwendete Teststatistik, wie Tabelle 2.1 (Seite 8).

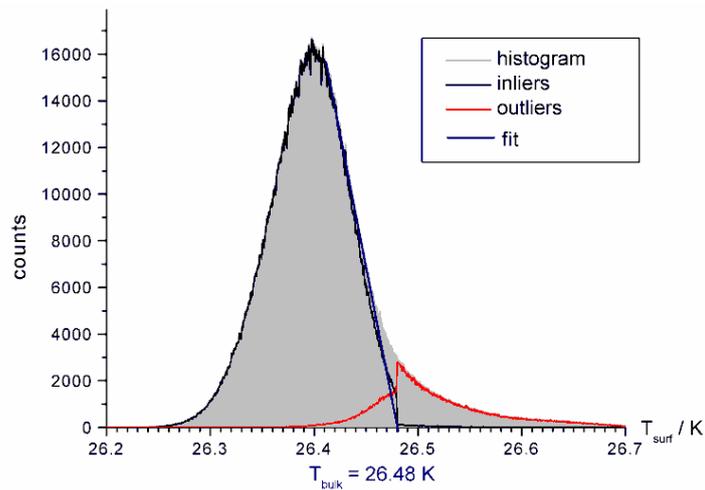
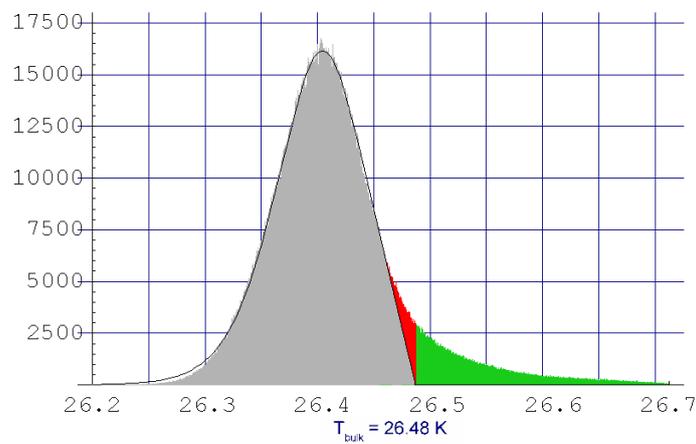
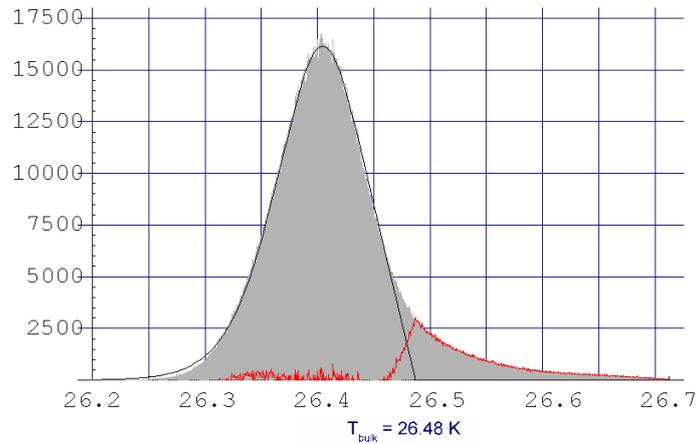
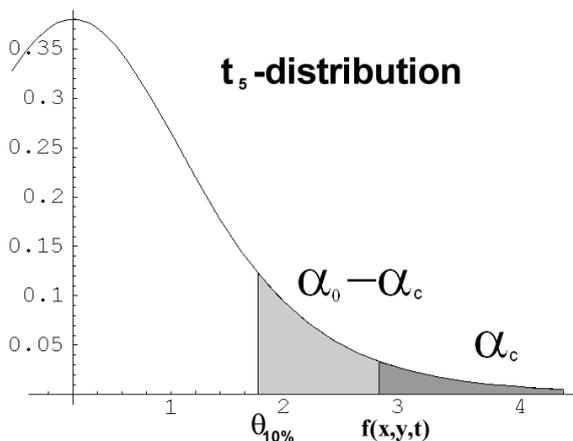


Abbildung 5.8: Detektionsergebnisse für den in Abbildung 5.9 gezeigten Ausschnitt einer Bildfolge **Oben:** Histogramm der kalibrierten Intensitäten über 55 aufeinanderfolgende Bilder (grau) mit Kurvenfit ($T_{\text{bulk}} = 26,48 \text{ K}$) und dem positiven Anteil der Differenz beider Kurven (rot) **Mitte:** Während die Ausreißer im grünen Bereich aus einer sicher bestimmten Tiefenwassertemperatur einfach anhand ihres Grauwertes in den Originaldaten zu bestimmen sind, ist dies für Ausreißer im roten Bereich nicht mehr der Fall **Unten:** Gleicher Ausschnitt der Bildfolge wie oben, die Ausreißerdetektion hat die beiden Stichproben der “inlier” (schwarze Kurve) und “outlier” (rote Kurve) getrennt. Zusätzlich ist der nahezu lineare Abfall der obere Flanke der theoretischen Verteilung zum Schittpunkt bei $T_{\text{bulk}} = 26,48 \text{ K}$ eingezeichnet (blau).

Voraussetzung dafür ist, daß die Temperatur des Tiefenwassers über den betrachteten Zeitraum konstant ist, was bei solch kurzen Zeitausschnitten sicher gegeben ist. Grauwerte $g(x, y, t) > T_{\text{bulk}}$ werden in der initialen Änderungsmaske bereits als ‘‘Ausreißer’’ markiert. Dabei ist eine Kontrolle der sich auch später in den Iterationen verringernden Stichprobengrößen $N(x, y)$ wichtig, um die Schwellwerte $\theta_{\alpha;N}^{(i)}(x, y)$ in jedem Iterationsschritt gemäß Tabelle 5.1 anzupassen und deren Absinken in der Größe $\Delta t \cdot \beta^{(x,y)}(x, y) = N(x, y)$ in allen Iterationen i zu kontrollieren. Für die Signifikanz α wählten wir den Wert 0,10. Aus der Differenz des initialen Kurvenfits von (5.1) und der tatsächlich im Einzelbild vorhandenen Grauwertverteilung läßt sich bereits ein ungefähres Maß für den Anteil an Ausreißern im jeweils untersuchten Einzelbild abschätzen. Es bietet sich an, daraus ein Abbruchkriterium für die Iteration zu machen und diese so lange laufen zu lassen, bis dieses Verhältnis erreicht ist. Bei den vorliegenden Daten ergab sich, daß durchgängig nach 3 Iterationschritten ein stabiles Ergebnis erreicht war. In Abschnitt 3.3.7 und 3.3.8 (Seiten 58 und 59) sind die Größen zusammengetragen, die man als Ergebnisse des Verfahrens erhält.

Im Anschluß an die iterative Detektion sind Glättungen der Änderungsmasken mit morphologischen Operatoren auf Einzelbildern (typischerweise mit 3×3 -Nachbarschaften) oder unter Berücksichtigung zeitlicher Nachbarn ($3 \times 3 \times 3$) denkbar. In diesem Beispiel wurde darauf verzichtet, da die resultierenden Änderungsmasken zusammen mit den Originalbilddaten für die oben beschriebene weitere Verarbeitung mit der erweiterten Strukturtenormmethode ohnehin neu abgetastet werden. Für diese weitere Verarbeitung wünscht man sich allerdings nicht nur eine binäre Zugehörigkeitsaussage (Inlier oder Outlier) zu jedem Pixel, sondern ein sinnvolles Gütemaß zu dieser Entscheidung. Als solches verwenden wir wie in der Anwendung in Kapitel 4 die Teststatistik $f^{(I)}(x, y, z)$ und das Schwellwertbild $\theta_{10\%;N(x,y)}^{(I)}(x, y, t)$ zu jedem Bild der Sequenz (die letzte Iterationsstufe ist wieder mit I indiziert). Interessant als Gütemaß ist der Abstand der Teststatistik vom Schwellwert: je größer sein Betrag, desto sicherer war die Entscheidung der Klassifikation in die beiden Gruppen Inlier und Outlier. Um solche Abstände auch untereinander für verschieden große Stichprobengrößen vergleichbar zu machen, ist eine Rückrechnung auf die zu jedem $f^{(I)}(x, y, t)$ unter Berücksichtigung der am Ort (x, y) gültigen Stichprobengröße $N^{(I)}(x, y)$ gehörenden Signifikanzen α notwendig. Es entsteht dann mit der Studentverteilung t_N (mit N Freiheitsgraden) zu jedem Pixel der Sequenz die



entscheidungskritische Signifikanz:

$$\alpha_c(x, y, t) = \int_{f^{(I)}(x,y,t)}^{\infty} t_{N(x,y)}(f') df'$$

Dieser Wert gibt an, bis zu welcher Signifikanz man für das Pixel am Ort (x, y, t) gehen müßte, um

die getroffene Entscheidung ins Gegenteil zu kehren. Es gilt wieder

$$\alpha_c(x, y, t) \begin{cases} \geq \\ < \end{cases} \alpha_o \left\{ \begin{array}{l} H_0 : \text{kein Ausreißer} \\ H_1 : \text{Ausreißer} \end{array} \right.$$

wenn α_o der für die Detektion gewählte Wert Signifikanzwert war. Mögliche Gütemaße wären also $|\alpha_c - \alpha_o|$ oder $|\alpha_c - \alpha_o / \alpha_o|$. Rechenzeitintensive Aufintegrationen können durch vorherige Tabellierung der Werte vermieden werden.

In der Abbildung 5.8 (oben) wird ein initialer Kurvenfit (5.1) an einen verhältnismäßig großen Ausschnitt Δt_{fit} gezeigt. Der rot markierte positive Anteil der Differenz stellt annähernd die Verteilung der Ausreißer dar, zu denen aus dieser Analyse der Verteilungen natürlich noch keine Positionen (x, y) vorliegen. Die in der schematischen Darstellung in der Abbildung 5.8 (Mitte) grün markierten Ausreißer lassen sich aus einer einmal durch Kurvenfit ermittelten Tiefenwassertemperatur durch eine einfache Schwellwertoperation in den Originaldaten lokalisieren. Dies ist für Ausreißer im roten Bereich nur anhand der Verteilungen und der einzelnen Grauwerte nicht mehr möglich. Nach der Änderungsdetektion mit den oben beschriebenen Einstellungen ergibt sich der in der Abbildung 5.8 unten dargestellte Plot: die schwarz markierte Kurve gibt die Verteilung der Inlier an, die rot markierte Kurve die der Outlier. Die Summe beider Kurven ergibt wieder die ursprüngliche Verteilung (grau hinterlegt). Eingezeichnet ist auch die aus dem Kurvenfit mit der Tiefenwassertemperatur resultierende Flanke der Kurve, die fast linear auf den T_{bulk} -Schnittpunkt zuläuft (blau). Deutlich zu erkennen ist, daß die Verteilung der detektierten Inlier in hohem Maß der vorhergesagten Verteilung (5.1) folgt, obwohl wir aus diesem Modell nur die Tiefenwassertemperatur in die initiale Änderungsmaske einfließen lassen. Darüber hinaus ist in der interessanten Ausreißer-Region – nämlich bei Grauwerten knapp unterhalb der Tiefenwassertemperatur – ein gutes Ergebnis erzielt worden: Ausreißer kommen in der Tat bis hinnab zu 26,5 K vor. Das initiale Setzen der Änderungsmaske läßt sich auch nach einigen Iterationen noch immer schwach an einem sprunghaftem Verlauf der Verteilung der Ausreißer bei der entsprechenden Tiefenwassertemperatur ablesen. Interessant ist auch, wie die benutzte Detektion das Rauschverhalten der Bilddaten mitberücksichtigt und in den Ergebnissen erkennbar werden läßt. Pixelzahlen aus dem Bereich knapp unter der Tiefenwassertemperatur sind eher schwach unterschätzt. Im Bereich knapp über der Tiefenwassertemperatur dagegen findet sich eine rund ebenso große Anzahl Pixel: hierbei handelt es sich um Rauscheffekte, die aus der theoretischen Vorhersage des Kurvenfits in den realen Daten auch Grauwerte über der Tiefenwassertemperatur als Nicht-Reflexe zulassen.

In Abbildung 5.9 sind 16 Änderungskarten zusammen mit den Originaldaten des als Beispiel behandelten Ausschnittes (55 Bilder) dargestellt. In den Verteilungen für jedes einzelne Bild sind die eingangs als Inlier gekennzeichneten Pixel schwarz markiert, die nach Abschluß der Iterationen gefundenen Ausreißer sind rot markiert.

5.9 Grenzen des Verfahrens

Auch bei dieser Anwendung konnte das in Kapitel 3 entworfene Verfahren die gesuchten Ereignisse, in diesem Fall paßt der Begriff “Ausreißer” besser als die Bezeichnung “Änderungen”, finden und – mit einem Gütemaß für die Sicherheit der Entscheidung Ausreißer/kein Ausreißer versehen – der weiteren Auswertung bereit stellen.

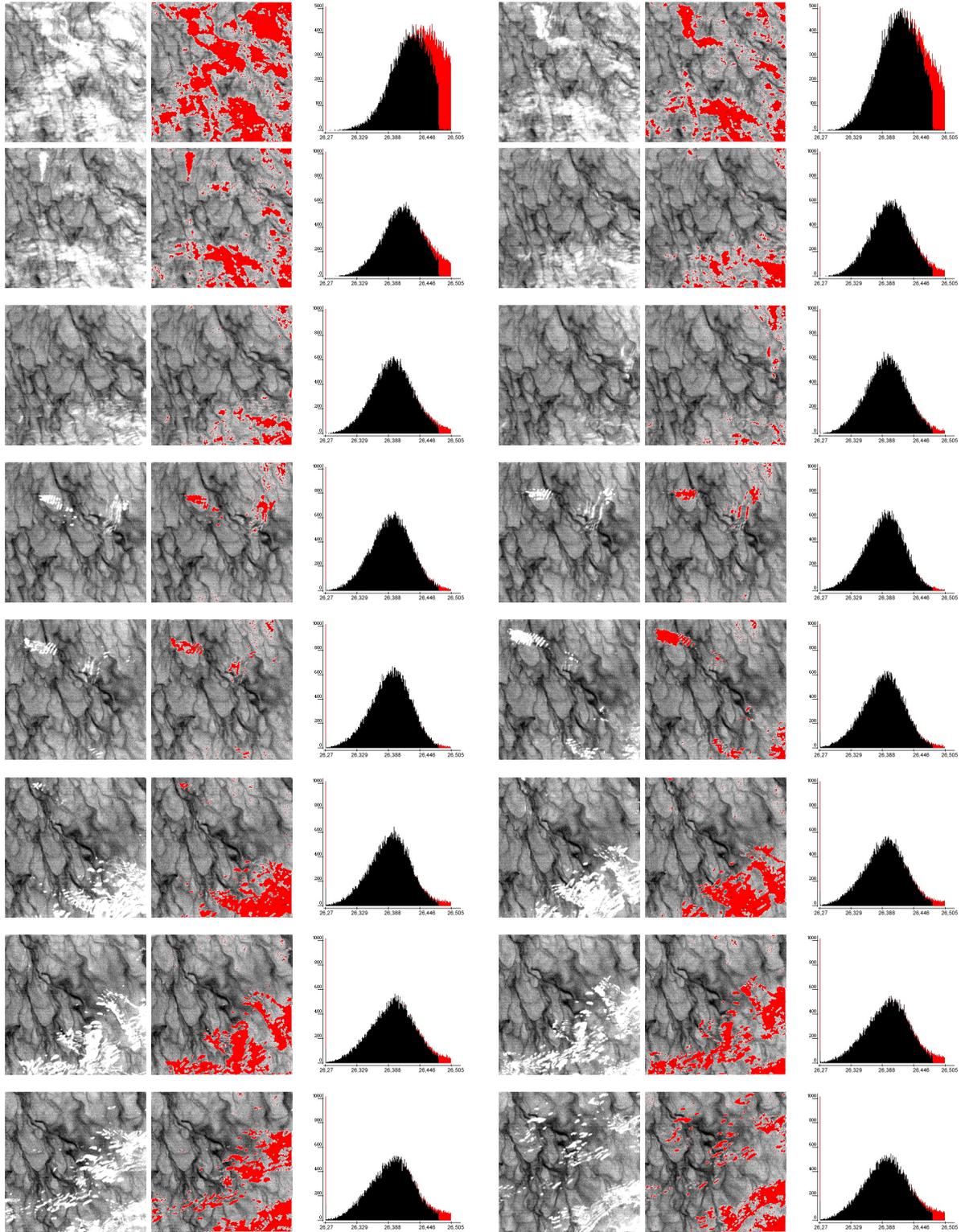


Abbildung 5.9: Ausreißerdetektion in 16 aus 55 aufeinanderfolgenden Bildern: Originalbild nach Temperaturkalibration, Originalbild mit darüber gelegter Maske der als Ausreißer detektierten Gebiete (rot) und Grauwertverteilung der initialen Inlier (schwarz) und abschließend detektierten Ausreißer (rot)

Als problematisch verbleibt dabei die Konstruktion stückweise stationärer Szenen. Grundsätzlich entsprechen die Daten nicht den in Abschnitt 3.2 vorausgesetzten Bildinhalten, da sich auch das Wärmemuster der Wasseroberfläche zeitlich durch Bewegung und Auskühlung schnell ändert und nicht mehr von einer still stehenden Szene gesprochen werden kann. Der hier gewählte Ansatz, die genannte Voraussetzung wieder herzustellen, indem die Szenenstücke geschickt ausgewählt werden und darüber hinaus gleitende Mittelwertschätzer und Varianzschätzer zur Verarbeitung der Stücke benutzt werden, hat unzweifelhaft eine Lösung des Detektionsproblems möglich gemacht. Dennoch ist festzuhalten, daß wir mit Szenenstücken bis hinab zu $\Delta t = 5$ an die Grenzen des Verfahrens gehen, wenn man gleichzeitig beachtet, daß in einer Vielzahl von Einzelbildern der Anteil der Ausreißer an die 50% beträgt. Versuche, mit den im theoretischen Abschnitt 2.3.6 behandelten Bayes-Ansätzen zur Glättung der Änderungskarten fehlende Ausreißer zu detektieren und Lücken in den Karten zu schließen, zeigten wenig Erfolg: die Regionen der über die Bilder ziehenden Reflexe sind keine kompakten Objekte, sondern haben typischerweise untereinander unverbundene "geriffelte" Formen, bei denen es keinen Sinn macht, Kompaktheit oder Glattheit zu erzwingen.

Ähnlich wie bei der in Kapitel 4 behandelten Anwendung wäre auch hier eine höhere zeitliche Abtastrate der Bilder eine Lösungsmöglichkeit, um die Länge der stationären Szenenstücke auch bei hohen Windgeschwindigkeiten und Wasserbewegungen deutlich über einer sinnvollen unteren Grenze von 4 Bildern zu halten und einen stillstehenden Hintergrund schlichtweg durch höhere Bildraten zu erzeugen. Wenn sich die Reflexe und das Wärmemuster jedoch mit der gleichen Geschwindigkeit bewegen, ist eine auch noch so hohe Bildrate keine Lösungsmöglichkeit mehr. Ansätze können dann nur noch aus der Tatsache erwachsen, daß sich das materialgebundene Wärmemuster des Wassers und die an den Neigungswinkel der Wasseroberfläche gebundenen Reflexe zwar oft mit gleichem Geschwindigkeitsbetrag, aber häufig in leicht unterschiedliche Richtungen bewegen (Abbildung 5.4). Bisher wird das aus der Eigenwertanalyse des erweiterten Strukturtenors bereitstehende Geschwindigkeitsfeld $f(x, y, t)$ nicht weiter zur Auswertung genutzt. Es ist zu erwarten, daß die Verteilung des Geschwindigkeitsfeldes in seinen Richtungen in zwei Hauptbestandteile zerfällt, wobei für die Ausreißer zusätzlich gilt, daß sie signifikant höhere Grauwerte in den Ausgangsbilddaten haben. Damit stünde zusammen mit den bisherigen Analysen der Verteilungen und Abweichungen wahrscheinlich genug Information bereit, um Ausreißer sicher zu trennen. Solche Ansätze bedeuten jedoch auch eine in der Regel iterative und damit rechenintensive Trennung zweier überlagerter Geschwindigkeitsfelder (ein Feld als Bewegung des zeitlich variablen Hintergrundmusters und ein Feld als räumlich konstanter additiver Beitrag zu diesem Muster), was die Ausreißeranalyse wieder in das Verfahren einbinden würde und nicht mehr als modularen Baustein voranstellt.

Kapitel 6

Resumée

Nach den beiden in den Kapiteln 4 und 5 beschriebenen Umsetzungen soll das in Kapitel 3 entworfene Verfahren zur Ereignisdetektion in den Kontext der im theoretischen Teil dieser Arbeit vorgestellten wichtigen Verfahren zur Ausreißerdetektion (Abschnitt 2.2) und Änderungsdetektion (Abschnitt 2.3) gestellt werden.

6.1 Wahl des Verfahrens

Die benutzte Ereignisdetektion ist eine Analogie der Ausreißerdetektion von Thompson (Abschnitt 2.2.2) für die Bildverarbeitung und damit ein klassisches Verfahren der Ausreißeranalyse mit signifikanzbasiertem Hypothesentest. Der einzige frei wählbare Parameter darin ist das Signifikanzniveau; die Stichprobengröße (Anzahl der Bilder) wird durch die Aufgabenstellung bestimmt und sollte maximal sein. Jedes Pixel wird einzeln in einem Hypothesentest auf seine Ausreißereigenschaft hin geprüft auf Basis seines eigenen zeitlichen Werteverlaufes. Die Verteilung der Teststatistik unter der Gegenhypothese H_1 muß dafür nicht bekannt sein, was ein weiterer Vorteil des Hypothesentests ist. In der Tat wurde die Teststatistik von Thompson ohne jede Nachbarschaftsinformation der Pixel untereinander benutzt: diese floß erst bei den Glättungen oder der Formanalyse (Abschnitte 3.4 und 4.13) ein, die direkt auf den Werten der Teststatistik ausgeführt wurde. Aus dieser Teststatistik entstandene Verfahren, wie ESD, die Statistiken von Grubbs oder die Maße, die der Mahalanobis-Distanz ähneln, lassen sich stets auf eine ähnliche Konstruktionsvorschrift zurückführen, bei der ein aktuell zu prüfender Wert gegen einen möglichst ausreißerfrei geschätzten Mittelwert durch Bildung eines Abstandes verglichen wird. Um die Abstände verschiedener Datensätze, Iterationsstufen oder Verfahrensschritte miteinander vergleichen zu können, wird in der Regel auf eine in der Stichprobe vorhandene Variabilität, also Größen wie die Standardabweichung, normiert. Simonoff und Wilks (Abschnitte 2.2.3 und 2.2.4) haben in der statistischen Datenanalyse aktuell häufig eingesetzte und multivariat erweiterbare Verfahren zur Ausreißerdetektion entworfen, deren Verhalten bezüglich *masking* und *swamping* (Seite 10) am besten verstanden ist. Die Konzepte des Vorwärts-Rückwärts-Testens lassen sich an diesen Methoden gut ablesen, und gleichzeitig wird deutlich, daß der Preis für die gute Absicherung gegen die unerwünschten Effekte ein enormer Rechenaufwand ist, der zudem – wie auch schon bei den Teststatistiken von Grubbs – aus häufigen Sortierschritten großer Stichproben und Unterstichproben nach

Abstandsmaßen besteht. Auch kommen graphische Zwischenschritte vor. Eine der wesentlichen Anforderungen im Rahmen dieser Arbeit war die Entwicklung ausreichend schneller Verfahren zur Ausreißerdetektion. Insbesondere für die industrielle Anwendung galt die Vorgabe, keinen akademischen Machbarkeitsnachweis für einen hochentwickelten Detektor zu implementieren, sondern eine in den Produktionstakt von nur wenigen Sekunden integrierbare Lösung bereitzustellen, die auf Standard-PCs und in der Industrie eingesetzten Betriebssystemen zeitsicher die Bildsequenzen klassifiziert und entscheidungssicher Qualitätsmängel der Produkte aussortiert. Ähnliches gilt auch für die Auswertung der IR-Bilddaten zur Wärmeflußbestimmung: ein zusätzliches Modul zur Ausreißerdetektion in der Bildvorverarbeitung würde keine Akzeptanz erfahren, wenn die ohnehin zeitintensive Parameterschätzung erheblich ausgedehnt würde. Mit ausgefeilten Vorwärts-Rückwärts-Methoden wären diese Zeitvorgaben, aber auch der Wunsch nach möglichst parameterfreien und ebenso leicht verständlichen wie leicht zu bedienenden Modulen nicht umzusetzen gewesen.

Methoden aus [Hsu, Y. Z. u. a., 1984] (Abschnitt 2.3.3) und die in [Aach, 1993] genannten Hypothesentests arbeiten alle auf Differenzbildern und fordern daher keine stückweise stationären Sequenzen wie wir sie für unsere Teststatistik $f(x, y, t)$ benötigen. Die Beschränkung auf Stichprobenwerte aus zwei aufeinanderfolgenden Bildern setzt allerdings voraus, daß sich das Sensor-Rauschen und die in den Bildfolgen vorhandenen Variationen, die nicht als Ereignisse detektiert werden sollen, in engen Grenzen halten. Dies war für unsere beiden Anwendungsgebiete nicht der Fall, da neben stark rauschender Sensorik auch Szenen von an sich hoher Unruhe zu verarbeiten waren. Eine Lösung boten die neuen Ansätze einer Teststatistik auf zeitlichen Stichproben an festgehaltenen Orten anstelle der bei Aach et al. und Nagel et al. benutzten Stichproben aus mäßig großen räumlichen Umgebungen und minimalen zeitlichen Umgebungen.

Der Einsatz von Bayes-Glättungen auf den ermittelten Änderungskarten nach Geman et al. und Bouthemy et al. (Aach et al. modifizieren damit die Schwellwerte der Hypothesenentscheidung) erwies sich für beide hier untersuchten Anwendungen als nicht nötig, was aber für die verwendete Teststatistik nicht allgemein gelten muß. Bei der Objektdetektion zeigte sich, daß auch mit ungeglätteten Änderungskarten sehr gut von Ereignissen auf Objekte geschlossen werden kann, da die volle Werteinformation (und nicht nur die binäre Entscheidung) in die mächtigen Werkzeuge der momentenbasierten Formanalyse und Polynomklassifikation eingingen. Im Fall der Detektion der Reflexe machte eine Glättung der nicht-kompakten Formen der gesuchten Regionen oder eine Formangleichung grundsätzlich keinen Sinn.

Für die Einteilung der Bilddaten in möglichst stationäre Stücke haben wir die zeitliche Mittelung als gleitend formuliert und mußten zusätzlich auf eine sehr kleine Stichprobengröße gehen. Bei den Änderungsdetektoren, die ihr Hauptaugenmerk nicht auf die Schwellwerte oder Teststatistik legen, sondern auf die saubere Konstruktion eines Hintergrundbildes (um dann mit einem einfachen Differenzbildverfahren vom aktuellen Bild zum Hintergrundbild fortzufahren), haben wir dazu im Abschnitt 2.3.9 den Multiskalenansatz mit zeitlichen Pyramiden von [Haußecker, Horst und Jähne, Bernd, 1993] als einen der wenigen uns bekannten zeit-tiefen Ansätze zur statistischen Auswertung von Bildfolgen genannt. Die dort neue Idee, zeitliche Mittelungen gleichzeitig auf mehreren, unterschiedlich abgetasteten zeitlichen Skalen (also: verschieden großen Stichprobengrößen auf der Zeitachse) durchzuführen, wurde bei der Bearbeitung der schlecht-stationären Bildsequenzen in Kapitel 5 in recht einfachen Ansätzen (Abschnitt 5.7) zur vorbereitenden Analyse und zeitlichen Seg-

mentierung der Sequenzen benutzt. Das von Haußecker et al. benutzte Verfahren geht aber nicht so weit, eine Teststatistik, einen Hypothesentest (in allgemeiner Formulierung am Ende von Abschnitt 2.2.2) und die durch die Signifikanz gegebenen Schwellwerte für die Entscheidungsfindung anzugeben. Stattdessen wird eine lokal nachgeführte, anschaulich abgeleitete Schwelle konstruiert, was eine Vergleichbarkeit des Verfahrens auf verschiedenen Daten erschwert. Wie andere Autoren auch benutzt Haußecker et al. ein Differenzbildverfahren zur Konstruktion seines Ähnlichkeitsmaßes. Es muß hinzugefügt werden, daß wie bei allen anderen Verfahren, die ein Hintergrundbild konstruieren, auch hier implizit eine Statistik über die zeitliche Dimension betrieben wird – jedoch ohne daß dies in der Teststatistik deutlich formuliert ist. Das Analogon zur zeitlichen Multiskalenmittelung in der statistischen Datenverarbeitung sind die Methoden von Wilks und Hadi (Abschnitte 2.2.4 und 2.2.5), in denen mit großem Aufwand nach einer möglichst sauberen Basis gesucht wird, aus der heraus der Test im vom Rosner eingeführten Sinne des Vorwärts-Rückwärts-Testens dann von innen heraus “anwächst”.

6.2 Herausragende Eigenschaften der verwendeten Teststatistik

Zeitliche Mittelwerte und Standardabweichungen werden von uns zu einer Teststatistik zusammengesetzt, die zur Vergleichbarkeit mit anderen Daten oder auch anderen Teststatistiken direkt eine entscheidungskritische Signifikanz (5.8) als allgemein übertragbares Gütemaß für ihre Entscheidung über Beibehaltung oder Ablehnung der H_0 -Hypothese liefert. Dies erleichtert es, das in dieser Arbeit entworfene und auf seine Brauchbarkeit hin überprüfte Verfahren als einen Baustein in bereits vorhandene Algorithmik einzubauen oder voranzustellen (so geschehen bei der Wärmeflußberechnung von Infrarotdaten in Abschnitt 5.6). Denn als weitere Ergebnisgröße (Abschnitt 3.3.7) tritt ein Maß für die Sicherheit, mit der jede Pixel-individuelle Entscheidung getroffen wurde, zu der binären Entscheidung Ereignis/kein Ereignis hinzu. Die in der Teststatistik $f(x, y, t)$ enthaltenen Schätzer für zeitliches Stichprobenmittelwert und zeitliche Standardabweichung werden durch ein iteratives Verfahren robust gemacht, so daß die unerwünschten Effekte des *masking* und *swamping* unterdrückt werden. Für die Anwendung in der industriellen Prozeßkontrolle, bei der die Ereignisdetektion zu einem bestimmten Grad auch eine Objektdetektion von Partikeln im Bildfeld ist, konnte auf der Teststatistik weitergerechnet werden und die Teststatistik selbst (Abschnitt 4.13) als im Maß für die Objekte einer momentenbasierten Formanalyse ihrer Regionen $> \theta_\alpha$ unterzogen werden. Nicht nur die erhaltene Änderungskarte, sondern auch die hier benutzte Teststatistik ist also in beiden vorgestellten Anwendungen weit über die reine Schwellwertbildung im Hypothesentest hinaus eine überaus nützliche Größe, um weitere Information über die Ausreißereigenschaft der einzelnen Pixel in nachfolgende Bildverarbeitung zu integrieren. Hinweise oder Umsetzungen einer derartigen Weiterverwendung der Teststatistik unter dem Aspekt eines Gütemaßes finden sich bei anderen Autoren nicht.

Das hier vorgestellte Verfahren benutzt im deutlichen Unterschied zu den Verfahren bei Aach et al. und Nagel et al. mehr als zwei aufeinanderfolgende Bilder aus einer Sequenz (oder die Differenz aus aktuellem Bild und Hintergrundbild). Es überträgt dagegen die Idee des Hypothesentests konsequent in die zeitliche Dimension und führt die Konstruktion einer zeitlichen Teststatistik über (möglichst) viele Bilder durch. Dies ist unter der wichtigen Voraussetzung (Abschnitt 3.2) stückweise stationärer Szenen möglich (zur Verletzung dieser Voraussetzung siehe Abschnitte 5.7 und 5.9).

Tatsächlich stellt diese Voraussetzung keine besondere Einschränkung in der Auswahl der An-

	industrielle Qualitätssicherung (Kapitel 4, Seite 91)	umweltphysikalische Wärmeflußmessungen (Kapitel 5, Seite 137)
Aufgabe	Ereignisdetektion in der Prozeßüberwachung	Ausreißerdetektion in der Datenvorverarbeitung
Sensor	HDR-CMOS	IR (Galileo, $NE\Delta t = 25$ mK)
Beobachtungsfeld	10 mm × 10 mm	40 cm × 40 cm
Auflösung	64 × 64 × 1200 frames	256 × 256 × 512 frames
Bildrate	1200 fps	100 fps
Datentiefe	10bit (log. comp.)	12 bit
ROI: $\Delta x \times \Delta y \times \Delta z$	64 × 64 × 460	251 × 251 × 5
Datenumfang	300 GB (raw, ROI)	30 GB (raw, ROI)
Häufigkeit der Ereignisse Szene	in 1 ‰ der Bildfolgen fester Hintergrund und helle bewegte Objekte im Vordergrund	bis 40 % jeder Bildfolge kontinuierliche Änderung im Bildhintergrund und helle Regionen im Vordergrund

Tabelle 6.1: Zusammenschau der wichtigsten Eigenschaften der beiden Anwendungen

wendungen dar: eine szenenfeste Kamera und ein verhältnismäßig gleichförmiger Szeneninhalt, dem wir in unseren Anwendungen mit neuartiger Sensorik (Abschnitt 4.4) oder dem bewegten Hintergrund extrem hohes Rauschen und Variation zugestehen, solche Verhältnisse sind absolut typisch für Überwachungsaufgaben und Ereignisdetektionen im weitesten Sinn, so daß sich zahlreiche Applikationen für das Verfahren finden lassen.

6.3 Wahl der Anwendungen

Die beiden in dieser Arbeit geschilderten Anwendungsgebiete waren dabei besonders attraktiv, da für die industrielle Qualitätssicherung die Ereignisdetektion nicht bloß modularer Baustein, sondern zentraler Schlüssel zur Lösung der gesamten Aufgabe war und darüber hinaus im Gebiet der Überwachung von Laserschweißprozessen ein neues Verfahren (Vergleiche zu herkömmlichen Verfahren in Abschnitt 4.17) zur Patentreife und zum Einsatz in der Massenproduktion brachte. In der Anwendung bei der Ausreißerdetektion bei der Berechnung von Wärmequellen und Bewegungsfeldern auf IR-Daten ist es interessant, den Benutzern zu dem bereits vorhandenem robusten TLS-Schätzer für Parameter (am Ende von Abschnitt 5.5) in der erweiterten Strukturtenormethode einen weiteren Ausreißerdetektor an die Hand zu geben, der zukünftig den direkten Vergleich zwischen robusten Parameterschätzern (und impliziter Ausreißerdetektion, zur Einführung siehe Abschnitt 2.2.1) und der direkten vorgeschalteten Ausreißerdetektion auf den Originaldaten erlaubt. Darüber hinaus diente die Anwendung in Kapitel 5 auch, um die Grenzen der vorgeschlagenen Ereignisdetektion aufzuzeigen.

6.4 Rückführung auf bekannte Verfahren

Zentrale Neuerung des hier vorgestellten Verfahrens ist die Abkehr von der Beschränkung auf zwei aufeinanderfolgende Bilder und von Teststatistiken auf Differenzbildern. Um das hohe Rauschen und die erlaubte starke Variabilität in den Bilddaten in den Griff zu bekommen, muß die gesamte in der Szene enthaltene Information in die Statistik, also den Schätzer für Mittelwert und Standardabweichung, eingehen. Diesen Zeitreihen-Ansatz für eine Ereignisdetektion in Bildern mittels Hypothesentests (Abschnitt 3.3.2) verfolgte nach unserer Kenntnis – trotz der zahlreichen sich dafür anbietenden Aufgaben – bisher kein anderer Autor. Den Vergleich der Verfahren soll eine erneute Zusammenchau abschließen, die gerade auf die mit dieser Arbeit beschrittene Erweiterung in die Dimension Zeit abhebt und den Weg zurück zu Differenzbildverfahren geht. Die verwendete Teststatistik (3.3)

$$\frac{g(x, y, t) - \langle g(x, y, t) \rangle_{t, \Delta t}}{\sqrt{\left\langle \left(g(x, y, t) - \langle g(x, y, t) \rangle_{t, \Delta t} \right)^2 \right\rangle_{t, \Delta t}}}$$

ist definiert für eine Stichprobengröße $\Delta t \geq 3$. Für $\Delta t = 2$ kann man keine sinnvolle Varianz mehr berechnen; formal nehmen Zähler und Nenner den Betrag $1/2 (g_1 - g_2)$ an, so daß die Teststatistik $f(x, y, t)$ auf zwei Bildern formal die Werte ± 1 annimmt. Da die Berechnung einer Standardabweichung aus zwei Werten keinen Sinn macht, wäre es bei einer Reduzierung auf zwei Einzelbilder angebracht, nicht die zeitliche Varianz aus den Daten selbst, sondern einen separat ermittelten Wert für das Sensorrauschen oder eine räumliche Standardabweichung aus einem im Mittel konstanten Bild $\langle g(x, y) \rangle_{t, \Delta t} = \text{const. } \forall (x, y)$ zu benutzen. In diesem Fall wird die Teststatistik für die beiden Bilder zu den Zeiten t_1 und t_2 zu

$$\frac{\pm \frac{1}{2} [g(x, y, t_1) - g(x, y, t_2)]}{\sigma_o}$$

Durch Quadrieren vermeidet man die sich im Vorzeichen niederschlagende Bevorzugung eines der beiden Bilder gegenüber dem anderen. Um eine statistische Verteilung angeben zu können, muß nun gemittelt werden. Da die zeitliche Mittelung nicht mehr möglich ist, verbleibt zur Konstruktion einer Stichprobe nur der Ausweg über eine Mittelung in einer räumlich angemessen kleinen Umgebung $\mathbf{x} \in w_{\mathbf{x}}$ um \mathbf{x} , die dennoch eine ausreichend große Stichprobe erzeugt. Man gelangt zu dem Ausdruck

$$\frac{1}{2\sigma_o^2} \sum_{\mathbf{x} \in w_{\mathbf{x}}} (g(\mathbf{x}, t_1) - g(\mathbf{x}, t_2))^2$$

Dies ist gerade die von Aach bei der Änderungsdetektion auf Differenzbildern (Abschnitt 2.3.5) benutzte Teststatistik! Der von Nagel aus dem Yakimovski-Maß für die Änderungsdetektion auf die Zeitachse übertragene Likelihood-Test (Abschnitt 2.3.3) lautet

$$l^{2/n} = \frac{(s_0^2)^2}{s_1^2 \cdot s_2^2} \leq \theta$$

Dabei sind die Varianzen s_i^2 in den räumlichen Umgebungen w_i der beiden Bilder der Zeitpunkte t_i (mit $i = 1, 2$) geschätzt, und die Stichprobe w_0 vereint die beiden Umgebungen w_1 und w_2 . Man erhält

durch Auflösen der Eigenschaften der Verbundstichprobe nach den Einzelstichproben daraus

$$l^{2/n} = \frac{\left(\frac{s_1^2 + s_2^2}{2} + \left(\frac{\langle g_1 \rangle_{w_2} - \langle g_2 \rangle_{w_2}}{2} \right)^2 \right)^2}{s_1^2 s_2^2}$$

und unter der Annahme sehr ähnlicher Varianzen in beiden Bildern $s_1 = s_2 = s$ folgt weiter

$$l^{1/n} = \frac{\frac{2s^2}{2} + \left(\frac{\langle g_1 \rangle_{w_2} - \langle g_2 \rangle_{w_2}}{2} \right)^2}{s \cdot s}$$

und schließlich mit

$$\sqrt{2} \left(l^{1/n} - 1 \right) = \frac{\left(\langle g_1 \rangle_{w_2} - \langle g_2 \rangle_{w_2} \right)^2}{s^2}$$

wieder ein Ausdruck, der mit dem Zusatz, daß sich die Grauwerte in den beiden aufeinanderfolgenden Bildern im Mittel nicht stark ändern, umschreibt zu

$$l^{1/n} - 1 = \left[\frac{g(x, y, t_1) - g(x, y, t_2)}{2s} \right]^2$$

für alle \mathbf{x} in der Umgebung. Dies ist aber gerade wieder die Teststatistik von Aach und die oben durchgeführte Reduzierung der Teststatistik $f(x, y, t)$ auf zwei Bilder. Die beiden zusätzlichen Forderungen bei der Umformung bedeuten aber nichts anderes als die Forderung nach stationären Szenenstücken. Damit konnte gezeigt werden, daß die von uns benutzte Teststatistik auf die Verfahren von Aach et al. und Nagel et al. zurückgeführt werden kann, wenn man einerseits die zeitliche Mittelung auf zwei Bilder reduziert und dadurch Differenzbildverfahren erreicht, andererseits unser Verfahren um eine Mittelung (denn es ist ein nachbarschaftsloses Verfahren, daß keinen räumlichen Zusammenhang der Pixel kennt) in einer räumlichen Umgebung erweitert.

6.5 Schlußbemerkung

Einleitend vorangestellt hatten wir auf Seite 1 eine der vielen möglichen und insbesondere in den wissenschaftlichen Arbeiten im Bereich der Datenanalyse kursierenden Definitionen für ungewöhnliche Ereignisse, wie sie Hawkins [Hawkins, 1980] vorschlug. Für den an der quantitativen Umsetzung interessierten Wissenschaftler stellen sich dazu sofort die beiden Fragen “*To deviate from what?*” und “*Where to put the borderline in between **much** and **too much**?*” Die in dieser Arbeit gefundenen Antworten darauf sind einerseits der möglichst robust geschätzte Mittelwert – wir gelangen in wenigen Iterationen dorthin – und zum anderen ein Vielfaches der durch den verbesserten Mittelwert bereits robusten Standardabweichung. Es gibt zu einer Stichprobengröße und einem Risiko von Fehlurteilen, das man bereit ist einzugehen, an, wo die gesuchte Grenze zu ziehen ist. Mit einer sinnvollen Auswahl der Dauer der Szenen und einem vertretbaren Risiko von 5% bis 10% Fehlurteilen (erster Ordnung) in dieser Teststatistik gelang eine sichere Detektion der gesuchten Ereignisse in zwei völlig unterschiedlichen Anwendungen und Themengebieten (Übersichtstabelle 6.1). Im ersten Fall konnte das Verfahren zu einer vollständigen und überzeugenden Lösung des bisher ungelösten Problems

der sicheren Prozeßüberwachung und Qualitätskontrolle beim Wärmeleitungsschweißen mit Lasern ausgebaut werden. Die zweite Praxisanwendung zeigt die Grenzen des Verfahrens auf, wenn nicht mehr alle notwendigen Voraussetzungen sicher eingehalten werden. Doch stellt auch in diesem Fall die mit dieser Teststatistik durchgeführte Detektion von Werten, die einer anderen Verteilung gehorchen, als daß dies vom zugrunde liegenden Modell für die bereits robustifizierte Parameterschätzung vorhergesehen wird, eine sinnvolle Ergänzung der Auswertung dar.

Anhang A

Rechenschema für inverses Momentenproblem

Für den Fall $N = 4$ lauten die Matrizen F^{-1} und F der Abbildung $\vec{m} = F^{-1}\vec{f}$ dann:

$$\begin{bmatrix} m_{00} \\ m_{10} \\ m_{01} \\ m_{11} \\ m_{20} \\ m_{02} \\ m_{30} \\ m_{03} \\ m_{21} \\ m_{12} \\ m_{40} \\ m_{04} \\ m_{31} \\ m_{13} \\ m_{22} \end{bmatrix} = 4 \begin{bmatrix} 1 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{5} & \frac{1}{5} & 0 & 0 & \frac{1}{9} \\ 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & \frac{1}{9} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{9} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{15} & \frac{1}{15} & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & \frac{1}{5} & \frac{1}{9} & 0 & 0 & 0 & \frac{1}{15} & \frac{1}{7} & 0 & 0 \\ 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{7} & 0 & 0 & \frac{1}{15} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{7} & \frac{1}{15} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{9} & 0 & 0 & 0 & 0 & \frac{1}{15} & 0 & 0 & \frac{1}{15} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{9} & \frac{1}{25} & 0 & 0 \\ \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & \frac{1}{15} & \frac{1}{7} & 0 & 0 & 0 & \frac{1}{25} & \frac{1}{9} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{21} & 0 \\ 0 & 0 & 0 & \frac{1}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{21} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{15} & \frac{1}{15} & 0 & 0 & 0 & 0 & 0 & \frac{1}{21} & \frac{1}{21} & 0 & 0 \\ \frac{1}{9} & 0 & 0 & 0 & \frac{1}{15} & \frac{1}{15} & 0 & 0 & 0 & 0 & \frac{1}{21} & \frac{1}{21} & 0 & 0 & \frac{1}{25} \end{bmatrix} \begin{bmatrix} f_{00} \\ f_{10} \\ f_{01} \\ f_{11} \\ f_{20} \\ f_{02} \\ f_{30} \\ f_{03} \\ f_{21} \\ f_{12} \\ f_{40} \\ f_{04} \\ f_{31} \\ f_{13} \\ f_{22} \end{bmatrix}$$

Nach Matrixinversion ist F dann

$$\frac{1}{16} \begin{bmatrix} \frac{243}{8} & 0 & 0 & 0 & -\frac{675}{8} & -\frac{675}{8} & 0 & 0 & 0 & 0 & \frac{945}{16} & \frac{945}{16} & 0 & 0 & \frac{225}{4} \\ 0 & 90 & 0 & 0 & 0 & 0 & -105 & 0 & 0 & -45 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 90 & 0 & 0 & 0 & 0 & -105 & -45 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 414 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -315 & -315 & 0 \\ -\frac{675}{8} & 0 & 0 & 0 & \frac{1215}{8} & \frac{225}{4} & 0 & 0 & 0 & 0 & -\frac{4725}{8} & 0 & 0 & 0 & -\frac{675}{4} \\ -\frac{675}{8} & 0 & 0 & 0 & \frac{225}{4} & \frac{1215}{8} & 0 & 0 & 0 & 0 & 0 & -\frac{4725}{8} & 0 & 0 & -\frac{675}{4} \\ 0 & -105 & 0 & 0 & 0 & 0 & 175 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -105 & 0 & 0 & 0 & 175 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -45 & 0 & 0 & 0 & 0 & 135 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -45 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 135 & 0 & 0 & 0 & 0 & 0 \\ \frac{945}{16} & 0 & 0 & 0 & -\frac{4725}{8} & 0 & 0 & 0 & 0 & 0 & \frac{11025}{16} & 0 & 0 & 0 & 0 \\ \frac{945}{16} & 0 & 0 & 0 & 0 & -\frac{4725}{8} & 0 & 0 & 0 & 0 & 0 & \frac{11025}{16} & 0 & 0 & 0 \\ 0 & 0 & 0 & -315 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 525 & 0 & 0 \\ 0 & 0 & 0 & -315 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 525 & 0 \\ \frac{225}{4} & 0 & 0 & 0 & -\frac{675}{4} & -\frac{675}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{2025}{4} \end{bmatrix}$$

Zur Erleichterung der Berechnung können wir einige Regeln ableiten: Die Summanden, aus denen sich die Momente zusammensetzen, lauten

$$m_{pq} = \dots + \int_{-1}^{+1} \int_{-1}^{+1} f_{ij} x^i y^j x^p y^q dx dy + \dots = \dots + f_{ij} \frac{x^{p+i+1}}{p+i+1} \Big|_{-1}^{+1} \frac{y^{q+j+1}}{q+j+1} \Big|_{-1}^{+1} + \dots = \dots + 4f_{ij} \frac{1}{p+i+1} \frac{1}{q+j+1} + \dots$$

Die Summanden sind nur ungleich Null, wenn $p+i$ und $q+j$ beide gerade sind. Die m_{pq} enthalten also nur Terme mit f_{ij} in denen p und j dieselbe Eigenschaft (beide gerade oder beide ungerade) haben. Das gleiche gilt für q und i . Wir haben stets den Faktor 4, die Brüche werden durch die Indices festgelegt. Deren Zähler sind stets 1, die Nenner sind ungerade Zahlen. Die Gleichung für m_{qp} erhält man aus m_{pq} durch Vertauschung der Indices in den f_{ij} . In den Gleichungen für m_{pp} sind die Elemente für f_{ij} und f_{ji} gleich.

Abbildungsverzeichnis

2.1	Masking- und swamping-Effekte	10
2.2	Beispiel für ein Dendrogramm	16
2.3	Signifikanzschwelle bei der χ^2_{25} -Verteilung	34
2.4	Hypothesentest an Stichprobenfunktionen	37
2.5	Glättung mit Bayes-Ansätzen	40
2.6	Konstruktion eines zeitlichen Skalenraumes	45
3.1	Schema der varianznormierten zeitlichen Abweichung	50
3.2	Bildrepräsentation als Ellipse durch Momente bis zu zweiter Ordnung	63
4.1	Wellenlängenbereiche und krit. Intensität beim Laserschweißen	94
4.2	Wärmeleitungsschweißen und Lasertiefschweißen	95
4.3	HDRC-Kamera und ihre Helligkeitsdynamik	99
4.4	Beispiele der log. Kennline bei der HDRC-Kamera	100
4.5	Einspritzventil	104
4.6	Laserpulsform	105
4.7	Beispielsequenz Laserschweißen (Voruntersuchungen)	106
4.8	Beispiele für Nahtfehler beim Laserschweißen	108
4.9	Aufbau Laserschweißen	109
4.10	Beispiele für entgegenkommende Partikel	110
4.11	Räumlich gemittelter Intensitätsverlauf	111
4.12	Beispiel für einen Partikelwurf	112
4.13	Mittelwerte und Standardabweichungen in verschiedenen Iterationen	114
4.14	Intensitätsverteilungen der inlier und outlier	116
4.15	Verteilung der Teststatistik	116
4.16	Detektionsergebnisse bei der Qualitätskontrolle	117
4.17	Detektionsergebnisse bei der Qualitätskontrolle (2. Teil)	118
4.18	Beispiel zur Verringerung der Stichprobe bei der Detektion	119
4.19	Detektionsergebnisse als Grauwertflächen: Originalbild und Teststatistik	119

4.20	Messungen zum Rauschen am HDRC-Sensor im Vergleich zu den Ereignissen	121
4.21	Ergebnisse der Detektion als x-y-t-Plots	122
4.22	Typischer Ablauf eines Partikelauswurfes (Machbarkeitsstudie)	127
4.23	Beispiele für Schmelzbadüberwachung (Grundlagenuntersuchung)	128
4.24	Beispiele für gefundene Nahtdefekte (Hohlräume)	132
4.25	Beispiele für gefundene Nahtdefekte (Partikelauswurf)	133
4.26	Beispiel für gefundene Nahtdefekte (Flugzeitberechnung)	134
4.27	Beispiel für gefundene Nahtdefekte (mögl. Hohlräume)	134
4.28	Abgrenzung zu anderen Verfahren der Spritzerdetektion	135
5.1	Ozeanoberfläche mit <i>sun-glitter</i>	139
5.2	Versuchsaufbau für Wärmeflußmessungen und IR-Kamera	140
5.3	Typische Bilddaten der Wärmeflußmessung	142
5.4	Bewegung der Reflexe und des Wärmemusters	143
5.5	Oberflächenerneuerungsmodell und zeitlicher Temperaturverlauf	144
5.6	Temperaturverteilungen in Einzelbildern der Wärmeflußmessungen	149
5.7	Bestimmung stationärer Szenenstücke bei Wärmefluß-Bildfolgen	151
5.8	Detektionsergebnisse anhand der Temperaturverteilungen der IR-Bildfolgen	153
5.9	Detektionsergebnisse für die IR-Sequenzen anhand ausgewählter Einzelbilder	156

Tabellenverzeichnis

2.1	Tabelle der Schwellwerte $\theta_{\alpha;N}$ für diverse Stichprobengrößen N	8
3.1	Tabelle für Orientierungswinkel aus Momenten	64
4.1	Table mit Vergleich von CCD-Sensoren und CMOS-Sensoren	103
4.2	Tabelle Erkennungsraten in der Machbarkeitsuntersuchung	130
4.3	Tabelle Erkennungsraten im Dauertest (Ausschnitt)	131
4.4	Tabelle Erkennungsraten im Dauertest (Nachweis Fertigungstauglichkeit)	131
5.1	Tabelle der Schwellwerte $\theta_{\alpha;N}$ für kleine Stichprobengrößen $N = 3 \cdots 30$	152
6.1	Tabelle zum Vergleich der beiden Anwendungen	162

Literaturverzeichnis

- [Aach 1993] AACH, Til: *Bayes-Methoden zur Bildsegmentierung, Änderungsdetektion und Verschiebungsvektorschätzung*, RWTH Aachen, Fortschrittberichte VDI Reihe 10, Nr.261, Düsseldorf, Dissertation, 1993
- [Aach, Til und Kaup, André 1995] AACH, TIL ; KAUP, ANDRÉ: Bayesian Algorithms for adaptive Change Detection in Image Sequences Using Markov Random Fields. In: *Signal Processing: Image Communication* 7 (1995), Nr. 2, S. 147–160
- [Aach, Til u. a. 1991] AACH, TIL ; KAUP, ANDRÉ ; MESTER, RUDOLF: A Statistical Framework for Change Detection in Image Sequences. In: *Proceedings 13 ième Colloque GRETSI, Juan-Les-Pins, France* (1991), S. 1149–1152
- [Aach, Til u. a. 1993a] AACH, TIL ; KAUP, ANDRÉ ; MESTER, RUDOLF: Change Detection in Image Sequences Using Gibbs Random Fields a Bayesian approach. In: *IEEE Proceedings ISPACS International Workshop on Intelligent Signal Processing and Communications Systems, Sendai, Japan, , 1993*, S. 376–381
- [Aach, Til u. a. 1993b] AACH, TIL ; KAUP, ANDRÉ ; MESTER, RUDOLF: Statistical model-based change detection in moving video. In: *Signal Processing* 31 (1993), Nr. 2, S. 165–180
- [Aach, Til u. a. 1997] AACH, TIL ; KAUP, ANDRÉ ; MESTER, RUDOLF: Statistical model-based change detection in moving video. In: *Proceedings International Conference on Acoustic, Speech and Signal Processing, ICASSP 1997, Munich, Germany* Bd. 4: Multidimensional signal processing, neural networks, 1997, S. 2657
- [Aach, Til u. a. 2000] AACH, TIL ; TOH, DANIEL ; METZLER, VOLKER: Bayesian spatio-temporal motion detection under varying illumination. In: *European Signal Processing Conference EUSIPCO 2000* (2000)
- [Abhyankar 1991] ABHYANKAR, Shreeram S.: Invariant theory and enumerative combinatorics of young tableaux. In: *DARPA-ESPRIT Workshop on Applications of Invariance in Computer Vision*, March 1991, S. 55–108
- [Abu-Mostafa, Yaser S. und Psaltis, Demitri 1984] ABU-MOSTAFA, YASER S. ; PSALTIS, DEMITRI: Recognitive aspects of moment invariants. In: *PAMI IEEE Trans. Pattern Analysis and Machine Intelligence* 6 (1984), S. 698–706
- [Abu-Mostafa, Yaser S. und Psaltis, Demitri 1985] ABU-MOSTAFA, YASER S. ; PSALTIS, DEMITRI: Image normalisation by complex moments. In: *PAMI IEEE Trans. Pattern Analysis and Machine Intelligence* 7 (1985), S. 46–55
- [Ahanger, G. und Little, T. D. C. 1996] AHANGER, G. ; LITTLE, T. D. C.: A survey of technologies for parsing and indexing digital video. In: *Journal of Visual Communication and Image Representation* 7 (1996), Nr. 1, S. 28–43

- [Alferez, Ronald-Bryan O. und Wang, Yuan-Fang 1999] ALFEREZ, RONALD-BRYAN O. ; WANG, YUAN-FANG: Geometric and Illumination Invariants for Objects Recognition. In: *PAMI IEEE Trans. Pattern Analysis and Machine Intelligence* 21 (1999), Nr. 6, S. 505–535
- [Anderson, Peter G. und Gaboriski, Roger S. 1993] ANDERSON, PETER G. ; GABORSKI, ROGER S.: A Polynomial Method Augmented by Supervised Training for hand-Printed Character Recognition. In: *Proceedings of the International Conference on Artificial Neural Networks and genetic Algorithms ANNGA*, , 1993
- [Anscombe 1960] ANSCOMBE, F. J.: Rejection of outliers. In: *Technometrics* 2 (1960), Nr. 2, S. 123–147
- [Atkinson 1994] ATKINSON, A. C.: Fast, very robust methods for detection of multiple outliers. In: *Journal of the American Statistical Association* 89 (1994), S. 1329–1339
- [Bab-Hadiashar, Alireza und Suter, David 1998] BAB-HADIASHAR, ALIREZA ; SUTER, DAVID: Robust Optic Flow Computation. In: *International Journal of Computer Vision IJCV* 29 (1998), Nr. 1, S. 59–77
- [Barnett, Vic und Lewis, Toby 1994] BARNETT, VIC ; LEWIS, TOBY: *Outliers in Statistical Data*. 3. Wiley & Sons, Chichester, New York,, 1994
- [Barron, J. L. und Liptay, A. 1994] BARRON, J. L. ; LIPTAY, A.: Optic Flow to Measure Minute Increments in Plant Growth. In: *Bioimaging* 2 (1994), Nr. 1, S. 57–61
- [Beckman, R. J. und Cook, R. D. 1983] BECKMAN, R. J. ; COOK, R. D.: Outlier...s. In: *Technometrics, silver issue 25th anniversary* 25 (1983), S. 119–149
- [Beersiek 1999] BEERSIEK, Joerg: On-line monitoring of Keyhole Instabilities during Laser Beam Welding. In: *Proceedings of International Congress on Applications of Lasers and Electro-Optics ICALEO 1999, San Diego*, , 1999
- [Beersiek 2001] BEERSIEK, Joerg: A CMOS camera as a tool for process analysis not only for laser beam welding. In: *Proceedings of International Congress on Applications of Lasers and Electro-Optics ICALEO 2001, Jacksonville*, , 2001
- [Belkasim, S. O. u. a. 1991] BELKASIM, S. O. ; SHRIDHAR, M. ; AHMADI, M.: Pattern recognition with moment invariants: A comparative study and new results. In: *Pattern Recognition* 24 (1991), Nr. 12, S. 1117–1138
- [Belkasim, S. O. u. a. 1993] BELKASIM, S. O. ; SHRIDHAR, M. ; AHMADI, M.: Corrigendum. In: *Pattern Recognition* 26 (1993), Nr. 1, S. 377
- [Bernoulli, Daniel und Allen, C. G. 1777] BERNOULLI, DANIEL ; ALLEN, C. G.: The Most Probable Choice between Several Discrepant Observations and the Formation Therefrom of the Most Likely Induction. In: *Biometrika* 48 (1777), S. 1–18. – translated reprint in 1961
- [Blom, A. H. M u. a. 1999] BLOM, A. H. M ; BRASSEL, JAN-OLIVER ; OTHERS: Quality Classification and Process Control of Micro-Spot Laser Welding. In: ASHAYERI, JALAL (Hrsg.) ; SULLIVAN, WILLIAM (Hrsg.) ; AHMAD, MUNIR (Hrsg.): *Proceedings of the Ninth International FAIM Conference - Flexible Automation and Intelligent Manufacturing, Tilburg*. New York : Begell House, 1999, S. 929–941
- [Boreczky, J. S. und Rowe, L. A. 1996] BORECZKY, J. S. ; ROWE, L. A.: Comparison of Video Shot Boundary Detection Techniques. In: *Proceedings of IS&T / SPIE International Symposium on Electronic Imaging: Science and Technology, San Jose, California* Bd. Storage and Retrieval for Image and Video Databases IV, 1996

- [Bouthemy, Patrick u. a. 1999] BOUTHEMY, PATRICK ; GELGON, M. ; GANANSIA, F.: A unified approach to shot change detection and camera motion characterization. In: *IEEE Trans.on Circuits and Systems for Video Technology* 9 (1999), Nr. 7, S. 1030–1044
- [Bouthemy, Patrick und Lalande, Patrick 1990] BOUTHEMY, PATRICK ; LALANDE, PATRICK: Detection and tracking of moving objects based on a statistical regularization method in space and time. In: FAUGERAS (Hrsg.): *Proceedings of the European Conference on Computer Vision ECCV 1990, Antibes, France*, Springer, New York, 1990, S. 307–311
- [Bouthemy, Patrick und Lalande, Patrick 1993] BOUTHEMY, PATRICK ; LALANDE, PATRICK: Recovery of moving object masks in an image sequence using local spatiotemporal contextual information. In: *Optical Engineering* 32 (1993), June, Nr. 6, S. 1205–1212
- [Boyce, J. F. und Hossak, W. J. 1983] BOYCE, J. F. ; HOSSAK, W. J.: Moment Invariants for pattern recognition. In: *Pattern Recognition Letters* 1 (1983), July, Nr. 5-6, S. 451–456
- [Braggins, Don und Schibli, Nikolaus 2001] BRAGGINS, DON ; SCHIBLI, NIKOLAUS: Image sensors for vision / Bildsensoren und ihre Zukunft. In: *Sensor Report* (2001), Nr. 5, S. 14–17
- [Brassel 2001] BRASSEL, Jan-Oliver: *Prozesskontrolle beim Laserstrahl-Mikroschweißen*, LFT Lehrstuhl für Fertigungstechnologie Universität Erlangen, Dissertation, 2001. – Verlag Meisenbach, Bamberg
- [Brocke 2002] BROCKE, Martin: Statistical Image Sequence Processing for Temporal Change Detection. In: VAN GOOL, Luc (Hrsg.): *Pattern Recognition 2002, 24. DAGM Symposium, Zürich*, Springer-Verlag, 2002 (LNCS), S. 215–223
- [Brocke 2003] BROCKE, Martin: In Planung: Statistical Image Sequence Processing for Event Detection. In: JÄHNE, Bernd (Hrsg.): *Image Sequence Analysis to Investigate Dynamic Processes*. Heidelberg : Springer-Verlag, 2003 (LNCS - Lecture Notes in Computer Science)
- [Brocke, Martin u. a. 2002] BROCKE, MARTIN ; SCHMIDT, MATTHIAS DIRK ; OTHERS: Verfahren zur automatischen Beurteilung von Laserverarbeitungsprozessen / Deutsches Patent DE 10103255 A1, Offenlegung 14.08. , 2002 (Rollnummer 39374). – Forschungsbericht
- [Burkhardt 1979] BURKHARDT, Hans: *Transformationen zur lageinvarianten Merkmalgewinnung*, University of Karlsruhe, Dissertation, 1979. – erschienen als Fortschrittbericht (Reihe 10, Nr.7) der VDI-Zeitschriften, VDI-Verlag
- [Canterakis 1996] CANTERAKIS, Nikolaos: Complete Moment Invariants and Pose Determination for Orthogonal Transformations of 3D Objects. In: *Mustererkennung 1996, 18. DAGM-Symposium, September 1996*, , 1996, S. 339–350
- [Caroni, Chrys und Prescott, Philip 1992] CARONI, CHRYS ; PRESCOTT, PHILIP: Sequential application of Wilks's multivariate outlier test. In: *Applied Statistics* 41 (1992), Nr. 2, S. 355–364
- [Cayley 1856] CAYLEY, Arthur: A second memoir upon quantics. In: *Philosophical Transactions of the Royal Society of London* 146 (1856), S. 101–126
- [Chen, Z. und Mu, G. 1995] CHEN, Z. ; MU, G.: High-dynamic-range image acquisition and display by multi-intensity imagery. In: *Journal of Imaging Science and Technology*, 39 (1995), Nr. 6, S. 559–564
- [Chrzanowski, Krzysztof u. a. 2001] CHRZANOWSKI, KRZYSZTOF ; MATYSZKIEL, ROBERT ; OTHERS: Uncertainty of temperature measurement with thermal cameras. In: *Optical Engineering* 40 (2001), Nr. 6, S. 1106–1114

- [Crank 1975] CRANK, J.: *The Mathematics of Diffusion*. 2nd. Oxford : Clarendon Press, 1975
- [Danckwerts 1970] DANCKWERTS, P. V.: *Gas-liquid reactions*. New York : MacGraw-Hill, 1970
- [Dausinger 1995] DAUSINGER, Friedrich: *Strahlwerkzeug Laser: Energieeinkopplung und Prozeßeffektivität*, Forschungsberichte des IFSW (Institut für Strahlwerkzeuge) - University of Stuttgart, Dissertation, 1995
- [David, H. A. u. a. 1954] DAVID, H. A. ; HEARTLEY, H. O. ; PEARSON, E. S.: The distribution of the ratio a single sample of range to standart deviation. In: *Biometrika* 41 (1954), S. 482–493
- [Davies 1977] DAVIES, P. J.: Plane regions determined by complex moments. In: *Journal of Approximation Theory* 19 (1977), S. 148–153
- [Davies, Laurie und Gather, Ursula 1993] DAVIES, LAURIE ; GATHER, URSULA: The identification of multiple outliers. In: *Journal of American Statistical Association* 88 (1993), S. 782–792
- [Dickson 1914] DICKSON, Leonard E.: *Algebraic Invariants*. John Wiley & Sons, 1914
- [Dilthey 2000] DILTHEY, Ulrich: *Laserstrahlschweißen: Prozesse, Werkstoffe, Fertigung und Prüfung*. DVS-Verlag, Düsseldorf, 2000
- [Dixon 1953] DIXON, W. J.: Proceeding Data with Outliers. In: *Biometrics* 9 (1953), march, Nr. 1, S. 74–89
- [Donohoe, G. W. u. a. 1988] DONOHOE, G. W. ; HUSH, D. R. ; AHMED, N.: Change detection for target detection and classification in video sequences. In: *Proceedings International Conference on Acoustic, Speech and Signal Processing, ICASSP*, 1988, S. 1084–1087
- [Dufaux, Frédéric und Moscheni, Fabrice 1995] DUFAUX, FRÉDÉRIC ; MOSCHENI, FABRICE: Segmentation-Based Motion Estimation For Second Generation Video Coding Techniques. In: TORRES, L. (Hrsg.) ; KUNT M. (Hrsg.): *Video Coding: the Second Generation Approach*. Academic Publishers, 1995, Kap. 6, S. 219–263
- [Eisele 2002] EISELE, Heiko: *Arbeitstitel: Röntgencomputertomographie im industriellen Einsatz*, Universität Heidelberg, Dissertation, 2002
- [Feng, J. u. a. 1996] FENG, J. ; LO, K. T. ; MEHRPOUR, H.: Scene change detection algorithm for MPEG video sequence. In: *Proceedings IEEE International Conference on Image Processing ICIP 1996, Lausanne, Switzerland*, , 1996
- [Fernando, W. A. C. u. a. 1999a] FERNANDO, W. A. C. ; CANAGARAJAH, C. N. ; BULL, D. R.: Fade and Dissolve Detection in Uncompressed and Compressed Video Sequences. In: *Proceedings of IEEE International Conference on Image Processing, ICIP 1999, Kobe, Japan*, 1999
- [Fernando, W. A. C. u. a. 1999b] FERNANDO, W. A. C. ; CANAGARAJAH, C. N. ; BULL, D. R.: Wipe Scene Change Detection in Video Sequences. In: *Proceedings of IEEE International Conference on Image Processing ICIP 1999, Kobe, Japan*, 1999
- [Flusser 1998] FLUSSER, Jan: Fast calculation of geometric moments of binary images. In: *Proceedings 22nd OAGM'98 Workshop Pattern Recognition Medical Computer Vision, Illmitz, Austria, 1998*, , 1998, S. 265–274
- [Flusser und Suk 1999] FLUSSER, Jan ; SUK, Thomas: *On the Calculation of Image Moments*. 1999. – unpublished

- [Flusser, Jan und Suk, Thomas 1993] FLUSSER, JAN ; SUK, THOMAS: Pattern recognition by affine moment invariants. In: *Pattern Recognition* 26 (1993), Nr. 1, S. 167–174
- [Foggia, Pasquale u. a. 1999] FOGGIA, PASQUALE ; SANSONE, CARLO ; OTHERS: Combining statistical and structural approaches for handwritten character description. In: *Image and Vision Computing* 17 (1999), Nr. 9, S. 701–711
- [Forsyth, David u. a. 1992] FORSYTH, DAVID ; MUNDY, JOSEPH L. ; ZISSERMAN, ANDREW: Transformational invariance - a primer. In: *Image and Vision Computing* 10 (1992), S. 39–45
- [Garbe 2001] GARBE, C. S.: *Measuring Heat Exchange Processes at the Air-Water Interface from Thermographic Image Sequence Analysis*. Heidelberg, Germany, University of Heidelberg, Dissertation, December 2001
- [Garbe, C. S. und Jähne, B. 2001] GARBE, C. S. ; JÄHNE, B.: Reliable estimates of the sea surface heat flux from image sequences. In: *Proc. of the 23rd DAGM Symposium*. Munich, Germany : Springer-Verlag, 2001 (Lecture Notes in Computer Science, LNCS 2191), S. 194–201
- [Garbe, C. S. u. a. 2001] GARBE, C. S. ; SPIES, H. ; JÄHNE, B.: Estimation of surface flow and net heat flux from infrared image sequences. In: *Journal of Mathematical Imaging and Vision* (2001). – Special Edition 'Analysis of Fluid Motion from Images', accepted
- [Garbe, Christoph S. u. a. 2002] GARBE, CHRISTOPH S. ; JÄHNE, BERND ; HAUSSECKER, HORST: Measuring the Sea Surface heat Flux and Probability Distribution of Surface Renewal Events. In: *Gas Transfer at Water Surfaces, Geophysical Monograph* (2002), Nr. 127, S. 109–114
- [Gather, Ursula und Becker, Claudia 1997] GATHER, URSULA ; BECKER, CLAUDIA: Outlier Identification and Robust Methods. In: MADDALA, G. S. (Hrsg.) ; RAO C.R. (Hrsg.): *Handbook of Statistics* Bd. Vol. 15: Robust Inference. Elsevier, Amsterdam, 1997, S. 123–143
- [Geman, Stuart und Geman, Donald 1984a] GEMAN, STUART ; GEMAN, DONALD: Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images. In: *IEEE Trans: PAMI* 6 (1984), Nr. 6, S. 721–741
- [Geman, Stuart und Geman, Donald 1984b] GEMAN, STUART ; GEMAN, DONALD: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. In: *PAMI IEEE Trans. Pattern Analysis and Machine Intelligence* 6 (1984), S. 721–741
- [Glumann 1996] GLUMANN, Christiane: *Verbesserte Prozeßsicherheit und Qualität durch Strahlkombination beim Laserschweißen*, University of Stuttgart - Forschungsberichte des IFSW (Institut für Strahlwerkzeuge), Dissertation, 1996
- [Graf, Heinz-Gerd u. a. 1995] GRAF, HEINZ-GERD ; HÖFFLINGER, BERND ; OTHERS: Elektronisch sehen - Alternative zur CCD-Struktur bewältigt hohe Helligkeitsdynamik. In: *Elektronik* (1995), Nr. 3, S. 52–57
- [Gribsch, J. u. a. 1996] GRIBSCH, J. ; SCHLICHTERMANN, L. ; JURCA, L. ET AL.: Quality Assurance of Industrial Spot Welding with a Pulsed Nd:YAG-Laser. In: *Proceedings of International Congress on Applications of Lasers and Electro-Optics ICALEO 1996, Michigan, USA*, 1996, S. 164–173
- [Gröning 2002] GRÖNING, H.: *Monokulares 3D-Tracking und radiometrische Kalibrierung*. Heidelberg, Germany, University of Heidelberg, Dissertation, 2002. – in preparation

- [Grubbs 1950] GRUBBS, Frank E.: Sample criteria for testing outliers observations. In: *Annals of Mathematical Statistics* 21 (1950), S. 27–58
- [Grubbs 1969] GRUBBS, Frank E.: Procedures for Detecting Outlying Observations in samples. In: *Technometrics* 11 (1969), S. 1–21
- [Gutchessy, D. u. a. 2001] GUTCHESSY, D. ; TRAJKOVICZ, M. ; OTHERS: A Background Model Initialization Algorithm for Video Surveillance. In: *Proceedings of IEEE International Conference on Computer Vision ICCV 2001*, , 2001
- [Hack 1998] HACK, Rüdiger: *System- und verfahrenstechnischer Vergleich von Nd:YAG- und CO₂ Lasern im Leistungsbereich bis 5 kW*, University of Stuttgart - Forschungsberichte des IFSW (Institut für Strahlwerkzeuge), Dissertation, 1998
- [Hader 2004] HADER, Sören: *Arbeitstitel: Klassifikation und Data Mining*, Universität Mannheim, Dissertation, 2004
- [Hadi 1992] HADI, Ali S.: Identifying Multiple Outliers in Multivariate Data. In: *Journal of the Royal Statistical Society B* 54 (1992), Nr. 3, S. 761–771
- [Hadi 1994] HADI, Ali S.: A Modification of a Method for Detection of Outliers in Multivariate Samples. In: *Journal of the Royal Statistical Society B* 56 (1994), Nr. 2, S. 393–396
- [Hadi, Ali S. und Simonoff, Jeffrey S. 1993] HADI, ALI S. ; SIMONOFF, JEFFREY S.: Procedures for the Identification of Multiple Outliers in Linear Models. In: *Journal of American Statistical Association JASA* 88 (1993), Nr. 424, S. 1264–1272
- [Hadi, Ali S. und Simonoff, Jeffrey S. 1994] HADI, ALI S. ; SIMONOFF, JEFFREY S.: Improving the Estimation and Outlier Identification Properties of the least Median of Squares and Minimum Volume Ellipsoid Estimators. In: *Parisankhyan Samikkha* 1 (1994), Nr. 1, S. 61–70
- [Haußecker 1996] HAUSSECKER, H.: *Messung und Simulation von kleinskaligen Austauschvorgängen an der Ozeanoberfläche mittels Thermographie*, University of Heidelberg, Dissertation, 1996
- [Haußecker 1993] HAUSSECKER, Horst: *Mehrgitter-Bewegungssegmentierung in Bildfolgen mit der Anwendung zur Detektion von Sedimentverlagerungen*, Universität Heidelberg, Diplomarbeit, 1993
- [Haußecker, H. und Jähne, B. 1997] HAUSSECKER, H. ; JÄHNE, B.: A Tensor Approach for Precise Computation of Dense Displacement Vector Fields. In: PAULUS, E. (Hrsg.) ; WAHL, F. M. (Hrsg.): *DAGM*. Braunschweig : Springer, September 1997, S. 199–208
- [Haußecker, H. u. a. 1995] HAUSSECKER, H. ; REINELT, S. ; JÄHNE, B.: Heat as a Proxy Tracer for Gas Exchange Measurements in the Field: Principles and Technical Realization. In: JÄHNE, B. (Hrsg.) ; MONAHAN, E. C. (Hrsg.): *Air-Water Gas Transfer - Selected Papers from the Third International Symposium on Air-Water Gas Transfer*. Heidelberg : AEON Verlag & Studio Hanau, 1995, S. 405–413
- [Haußecker, H. u. a. 2001] HAUSSECKER, H. ; SCHIMPF, U. ; OTHERS: Physics from IR image sequences: Quantitative analysis of transport models and parameters of air-sea gas transfer. In: SALTZMAN, E. (Hrsg.) ; DONELAN, M. (Hrsg.) ; OTHERS (Hrsg.): *Gas Transfer at Water Surfaces*. American Geophysical Union, 2001 (Geophysical Monograph)
- [Haußecker, H. und Spies, H. 1999] HAUSSECKER, H. ; SPIES, H.: Motion. In: JÄHNE, B. (Hrsg.) ; HAUSSECKER, H. (Hrsg.) ; GEISSLER, P. (Hrsg.): *Handbook of Computer Vision and Applications* Bd. 2. Academic Press, 1999, Kap. 13

- [Haußecker, Horst und Jähne, Bernd 1993] HAUSSECKER, HORST ; JÄHNE, BERND: Ein Mehrgitterverfahren zur bewegungssegmentierung in Bildfolgen. In: *Mustererkennung 1993, 15. DAGM Symposium*, , 1993, S. 27–31
- [Hauschild 1999] HAUSCHILD, Ralf: *Integrierte CMOS-Kamerasysteme für die zweidimensionale Bildsensorik*, xxx, Dissertation, 1999
- [Hawkins 1980] HAWKINS, Douglas M.: *Identification of Outliers*. Chapman and Hall, London, 1980
- [Hawkins 1985] HAWKINS, Douglas M.: Masking and Swamping. In: KOTZ, SAMUEL (Hrsg.) ; JOHNSON, NORMAN LLOYD (Hrsg.) ; READ, CAMPBELL B. (Hrsg.): *Encyclopedia of Statistical Sciences* Bd. 5. John Wiley & Sons, 1985
- [Heitz, F. und Bouthemy, Patrick 1990] HEITZ, F. ; BOUTHEMY, PATRICK: Motion estimation and segmentation using a global bayesian approach. In: *Proceedings International Conference on Acoustic, Speech, and Signal Processing, ICASSP*, April 1990, S. 2305–2308
- [Herrmann 2002] HERRMANN, Andreas: *Fertigungsorientierte Verfahrensentwicklung des Weichlötens mit Diodenlasern*, Universität Stuttgart, Dissertation, 2002
- [Höflinger 1999] HÖFFLINGER, Bernd: CMOS-Bildsensor, Bilder wie sie unser Auge mag. In: *Elektronik* (1999), Februar, Nr. 5, S. 30–34
- [Hügel 1992] HÜGEL, Helmut: *Strahlwerkzeug Laser*. Teubner, 1992
- [Hilbert 1890] HILBERT, David: Über die Theorie der algebraischen Formen. In: *Mathematische Annalen* 36 (1890), S. 473–534
- [Hilbert 1893] HILBERT, David: Über die vollen Invariantensysteme. In: *Mathematische Annalen* 42 (1893), S. 313–373
- [Horn 1994] HORN, Armin: *Optische Sensorik zur Bahnführung von Industrierobotern mit hohen Bahngeschwindigkeiten*, ISW Universität Stuttgart, Band 103, Dissertation, 1994
- [Horn, B. K. P. und Schunk, B. 1981] HORN, B. K. P. ; SCHUNK, B.: Determining Optical Flow. In: *Artificial Intelligence* 17 (1981), S. 185–204
- [Hornegger, J. u. a. 1999] HORNEGGER, J. ; PAULUS, D. ; NIEMANN, H.: Probabilistic Modeling in Computer Vision. In: JÄHNE, BERND (Hrsg.) ; GEISSLER, PETER (Hrsg.) ; HAUSSECKER, HORST (Hrsg.): *Handbook of Computer Vision and Applications* Bd. 2: Signal Processing and Pattern Recognition. Academic Press, 1999, Kap. 26
- [Hsu, Y. Z. u. a. 1984] HSU, Y. Z. ; NAGEL, H.-H. ; REKERS, G.: New likelihood test methods for change detection in image sequences. In: *Computer Vision, Graphics and Image Processing* 26 (1984), S. 73–106
- [Hötter, Michael u. a. 1995] HÖTTER, MICHAEL ; MESTER, RUDOLF ; MEYER, M: Detection of moving objects in natural scenes by a stochastic multi-feature analysis of video sequences. In: *29th Annual 1995 International IEEE Carnahan Conference on Security Technology, Sanderstead, Surrey, England* (1995), S. 47–52
- [Hötter, Michael und Thoma, Robert 1988] HÖTTER, MICHAEL ; THOMA, ROBERT: Image Segmentation Based on Object Oriented Mapping Parameter Estimation. In: *Signal Processing: Image Communication* 15 (1988), october, Nr. 3, S. 315–334

- [Hu 1961] HU, Ming-Kuei: Pattern Recognition by Moment Invariants. In: *Proceedings of the IRE Institute of Radio Engineers* Bd. 49, 1961, S. 1428
- [Hu 1962] HU, Ming-Kuei: Visual pattern recognition by moment invariants. In: *IRE (Institute of Radio Engineers) transactions on information theory* IT- 8 (1962), S. 179–187
- [Huang 1981] HUANG, Thomas S.: *Image Sequence Analysis*. Springer-Verlag, 1981
- [Huang 1982] HUANG, Thomas S.: *Image Sequence Processing and Dynamic Scene Analysis*. Springer-Verlag, Nato Advanced Study Institute, 1982 (NATO ASI Series F: computer and system sciences)
- [Huber 1980] HUBER, Peter J.: *Robust Statistics*. New York, Wiley, 1980
- [Jain 1989] JAIN, Anil K.: *Fundamentals of digital image processing*. Prentice Hall, New Jersey, 1989
- [Jain, Anil K. u. a. 2000] JAIN, ANIL K. ; DUIN, ROBERT P. W. ; MAO, JIANCHANG: Statistical Pattern Recognition: A Review. In: *IEEE PAMI Transactions on Pattern Analysis and Machine Intelligence* 22 (2000), Nr. 1, S. 4–37
- [Jain, Ramesh u. a. 1979] JAIN, RAMESH ; MARTIN, W. N. ; AGGARWAL, J. K.: Segmentation through the detection of changes due to motion. In: *Journal Computer Graphics and Image Processing* 11 (1979), S. 13–34
- [Jain, Ramesh und Nagel, H.-H. 1979] JAIN, RAMESH ; NAGEL, H.-H.: On the analysis of accumulative difference pictures from image sequences of real world scenes. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1 (1979), Nr. 2, S. 206–214
- [Jähne 1980] JÄHNE, B.: *Parametrisierung des Gasaustausches mit Hilfe von Laborexperimenten*. University of Heidelberg, Institut für Umweltphysik, Dissertation, 1980
- [Jähne 1993] JÄHNE, B.: *Spatio-Temporal Image Processing : Theory and Scientific Applications*. Springer-Verlag, 1993 (Lecture Notes in Computer Science)
- [Jähne 1990] JÄHNE, Bernd: From mean fluxes to a detailed experimental investigation of the gas transfer process. In: *2nd International Symposium on Gas Transfer at Water Surfaces - Air-Water Mass Transfer, Minneapolis 1990*, ASCE, 1990, S. 244–256
- [Jähne 2002] JÄHNE, Bernd: *Digital Image Processing*. 5th. Heidelberg, Germany : Springer-Verlag, 2002
- [Jähne, B. u. a. 1989] JÄHNE, B. ; LIBNER, P. ; OTHERS: Investigating the transfer process across the free aqueous boundary layer by the controlled flux method. In: *Tellus* 41B (1989), Nr. 2, S. 177–195
- [Jähne, Bernd und Brocke, Martin et al. 2002] JÄHNE, BERND ; BROCKE, MARTIN ET AL.: Für Anspruchsvolle - Multidimensionale Bildverarbeitung in der Produktion. In: *Journal Qualität und Zuverlässigkeit* 47 (2002), November, Nr. 11, S. 1154–1159
- [Jiang, X. Y. und Bunke, Horst 1991a] JIANG, X. Y. ; BUNKE, HORST: Ein konturbasierter Ansatz zur Berechnung von Momenten. In: *Mustererkennung 1991, 13. DAGM Symposium, München*. Heidelberg : Springer, 1991, S. 143–150
- [Jiang, X. Y. und Bunke, Horst 1991b] JIANG, X. Y. ; BUNKE, HORST: Simple and fast computation of moments. In: *Pattern Recognition* 24 (1991), Nr. 8, S. 801–806

- [Karmann, Klaus-Peter und von Brandt, Achim 1990] KARMANN, KLAUS-PETER ; VON BRANDT, ACHIM: Moving Object Recognition Using an Adaptive Background Memory. In: CAPPELLINI, V. E. (Hrsg.): *Time-Varying Image Processing and Moving Object Recognition* Bd. 2. Elsevier, Amsterdam, The Netherlands, 1990, S. 289–296
- [Karmann, Klaus-Peter u. a. 1990] KARMANN, KLAUS-PETER ; VON BRANDT, ACHIM ; GERL, R.: Moving object segmentation based on adaptive reference images. In: TORRES, L. et al. (Hrsg.): *Proceedings of the EUSIPCO, Barcelona, September 1990*, 1990, S. 951–954
- [Keller, James M. u. a. 1985] KELLER, JAMES M. ; GRAY, MICHAEL R. ; GIVENS, JAMES A. JR.: A Fuzzy K-Nearest Neighbor Algorithm. In: *IEEE Transactions on Systems, Man and Cybernetics* 15 (1985), Nr. 4, S. 580–585
- [Khotanzad, Alireza und Hong, Yaw Hua 1990] KHOTANZAD, ALIREZA ; HONG, YAW HUA: Invariant image recognition by Zernike moments. In: *PAMI IEEE Trans. Pattern Analysis and Machine Intelligence* 12 (1990), S. 489–497
- [Kilger 1992] KILGER, Michael: A shadow handler in a video based real time traffic monitoring system. In: *IEEE Workshop on Applications of Computer Vision, 1992*, 1992, S. 1060–1066
- [Koller, Dieter u. a. 1994a] KOLLER, DIETER ; WEBER, JOSEPH ; MALIK, JITENDRA: Robust multiple car tracking with occlusion reasoning. In: *European Conference on Computer Vision, ECCV 1994*, , 1994, S. 189–196
- [Koller, Dieter u. a. 1994b] KOLLER, DIETER ; WEBER, JOSEPH ; MALIK, JITENDRA: Robust Multiple Car Tracking with Occlusion Reasoning / University of California Berkeley. January 1994 (UCB/CSD-93-780 and UCB-ITS-PWP-94-01). – Forschungsbericht. – ISSN 1055-1417
- [Krebel, Ulrich u. a. 1990] KRESSEL, ULRICH ; FRANKE, J. ; SCHÜRMAN, JÜRGEN: Polynomklassifikator versus Multilayer-Perzeptron. In: GROSSKOPF, R. E. H. (Hrsg.): *Mustererkennung - 12. DAGM Symposium Oberkochen-Aalen*, , 1990, S. 75–81
- [Lalande, Patrick und Boutheymy, Patrick 1990] LALANDE, PATRICK ; BOUTHEMY, PATRICK: A statistical approach to the detection and tracking of moving objects in an image sequence. In: *Proc. of the Eusipco, Spain*, 1990, S. 947–950
- [Lambert, Georg u. a. 1995] LAMBERT, GEORG ; GAO, HUA ; OTHERS: Linienmomente und Invarianten zur Echtzeitverarbeitung vektorisierter Konturen. In: *Mustererkennung 1995, 17. DAGM Symposium, Bielefeld* (1995), S. 1–14
- [Le Gall 1991] LE GALL, D.: MPEG: A video compression standard for multimedia applications. In: *Communications of the ACM* 34 (1991), Nr. 4, S. 46–58
- [Li 1993] LI, Bing C.: A new computation of geometric moments. In: *Pattern Recognition* 26 (1993), S. 109–113
- [Li 1992] LI, Y.: Reforming the theory of invariant moments for pattern recognition. In: *Pattern Recognition Letters* 25 (1992), Nr. 7, S. 723–730
- [Li, Bing Cheng und Ma, S. D. 1994] LI, BING CHENG ; MA, S. D.: Efficient computation of 3D moments. In: *Proceedings 12th International Conference Pattern Recognition* Bd. 1, 1994, S. 22–26

- [Li, Bing Cheng und Shen, J. 1991] LI, BING CHENG ; SHEN, J.: Fast computation of moment invariants. In: *Pattern Recognition* 24 (1991), Nr. 8, S. 807–813
- [Liao, Simon X. und Pawlak, Mirosław 1996] LIAO, SIMON X. ; PAWLAK, MIROSLAW: On image analysis by moments. In: *PAMI IEEE Trans. Pattern Analysis and Machine Intelligence* 18 (1996), Nr. 3, S. 254–266
- [Lin, W. G. und Wang, S. S. 1994] LIN, W. G. ; WANG, S. S.: A note on the calculation of moments. In: *Pattern Recognition Letters* 15 (1994), November, Nr. 11, S. 1065–1070
- [Litwiller 2001] LITWILLER, Dave: CCD versus CMOS. In: *Photonics Spectra special issue* (2001)
- [Liu, W. T. und Businger, J. A. 1985] LIU, W. T. ; BUSINGER, J. A.: Temperature Profile in Molecular Sea Surface Temperatures. In: *Journal of Geophysical Research* 90 (1985), S. 11587–11601
- [Lo, Chong-Huah und Don, Hon-Son 1989] LO, CHONG-HUAH ; DON, HON-SON: 3D moment forms: their construction and application to object identification and positioning. In: *PAMI IEEE Trans. Pattern Analysis and Machine Intelligence* 11 (1989), Nr. 10, S. 1053–1064
- [Loncaric 1998] LONCARIC, Sven: A survey of shape analysis techniques. In: *Pattern Recognition* 31 (1998), Nr. 8, S. 983–1001
- [Loose 1999] LOOSE, Markus: *A Self-Calibrating CMOS Image Sensor with Logarithmic Response*, Universität Heidelberg, Dissertation, 1999
- [Loose, Markus u. a. 1996] LOOSE, MARKUS ; MEIER, KARLHEINZ ; SCHEMMEL, JOHANNES: Entwicklung einer Kamera mit adaptiven Photorezeptoren in analoger CMOS Technologie. In: JÄHNE, Bernd (Hrsg.): *Mustererkennung 1996, 18. DAGM Symposium, Heidelberg*, Springer Verlag, 1996, S. 301–312
- [Loose, Markus u. a. 1998] LOOSE, MARKUS ; MEIER, KARLHEINZ ; SCHEMMEL, JOHANNES: CMOS image sensor with logarithmic response and self calibrating fixed pattern noise correction. In: BERNARD, T. M. (Hrsg.): *Proceedings International Symposium on Electronic Image Capture and Publishing - Advanced Focal Plane Arrays and Electronic Cameras*, , 1998 (SPIE 3410)
- [Maitra 1979] MAITRA, S: Moment invariants. In: *Proceedings of IEEE*, 1979 (67), S. 697–699
- [Makarov 1996] MAKAROV, Aleksej: Comparison of background extraction based intrusion detection algorithms. In: *Proceedings of the IEEE International Conference on Image Processing, ICIP 1996* Bd. 1, , 1996, S. 521
- [Mayer 2001] MAYER, Arnold: Lasermarkt: Stand und Perspektiven. In: DAUSINGER, F. (Hrsg.) ; HÜGEL, HELMUT (Hrsg.) ; OTHERS (Hrsg.): *Stuttgarter Lasertage 2001, SLT 2001*, September 2001, S. 1–6
- [Mech, Roland und Wollborn, Michael 1997] MECH, ROLAND ; WOLLBORN, MICHAEL: A Noise Robust Method For Segmentation Of Moving Objects In Video Sequences. In: *Proceedings International Conference on Acoustic, Speech, and Signal Processing, ICASSP 1997, Munich, Germany* (1997), S. 2657–2660
- [Mitiche, Amar und Boutheymy, Patrick 1996] MITICHE, AMAR ; BOUTHEMY, PATRICK: Computation and analysis of image motion: a synopsis of current problems and methods. In: *International journal of computer vision, IJCV* 19 (1996), July, Nr. 1, S. 29–55
- [Mundy, Joseph L. und Zisserman, Andrew 1992] MUNDY, JOSEPH L. ; ZISSERMAN, ANDREW: *Geometric Invariance in Computer Vision*. The MIT Press, Cambridge, 1992

- [Munkelt 1993] MUNKELT, Olaf: Zur Auswahl von Merkmalen. In: PÖPPL, S. (Hrsg.) ; HANDELS, H. (Hrsg.): *Mustererkennung 1993, 15. DAGM-Symposium*, , 1993, S. 84–93
- [Nagel 1978] NAGEL, H.-H.: Formation of an object concept by analysis of systematic time variations in the optically perceptible environment. In: *Computer Graphics Image Processing 7 (1978)*, Nr. 149, S. 194
- [Nagel, Hans-Hellmut und Rekers, G. 1982] NAGEL, HANS-HELLMUT ; REKERS, G.: Moving Object Masks Based on an Improved Likelihood Ratio Test. In: *Proceedings International Conference on Pattern Recognition (ICPR) (1982)*
- [Naoya 2001] NAOYA, Ohta: A Statistical Approach to Background Subtraction for Surveillance Systems. In: *Proceedings of IEEE International Conference on Computer Vision ICCV 2001* Bd. 2, , 2001, S. 481–486
- [Niemann 1983] NIEMANN, Heinrich: *Klassifikation von Mustern*. Springer, 1983
- [Noffz, K.-H. u. a. 1999] NOFFZ, K.-H. ; LAY, R. ; OTHERS: Field Programmable Gate Array image processing. In: JÄHNE, BERND (Hrsg.) ; GEISSLER, PETER (Hrsg.) ; HAUSSECKER, HORST (Hrsg.): *Handbook of Computer Vision and Applications* Bd. 3: Systems and Applications. Academic Press, 1999, Kap. 2
- [Nordbruch, Stefan u. a. 2000] NORDBRUCH, STEFAN ; TSCHIRNER, P. ; GRÄSER, A.: Analyse von HDRC-Bildern des Werkstoffübergangs des MSG-Schweißprozesses. In: SOMMER, G. I. (Hrsg.): *Mustererkennung 2000, 22. DAGM Symposium, Kiel*, Springer Verlag, Heidelberg, 2000, S. 147–154
- [Parshall 1989] PARSHALL, Karen H.: Toward a history of nineteenth - century invariant theory. In: *The History of Modern Mathematics*. Academic Press, 1989, S. 157–208
- [Pearson, Erwin S. und Sekar, Chandra 1936] PEARSON, ERWIN S. ; SEKAR, CHANDRA: The efficiency of statistical tools and a criterion for the rejection of outlying observations. In: *Biometrika* 28 (1936), S. 308–320
- [Penny 1996] PENNY, Kay I.: Appropriate critical values when testing for a single multivariate outlier by using the mahalanobis distance. In: *Applied Statistics* 45 (1996), Nr. 1, S. 73–81
- [Philips 1993] PHILIPS, Wilfried: A new fast algorithm for moment computation. In: *Pattern Recognition* 26 (1993), S. 1619–1621
- [Prokop, Richard J. und Reeves, Anthony P. 1992] PROKOP, RICHARD J. ; REEVES, ANTHONY P.: A survey of moment-based techniques for unoccluded object representation and recognition. In: *Graphical Models and Image Processing* 54 (1992), September, Nr. 5, S. 438–460
- [Pudil, P. u. a. 1994] PUDIL, P. ; NOVOVICOVA, J. ; KITTLER, J.: Floating Search Methods in Feature Selection. In: *Pattern Recognition Letters* 15 (1994), Nr. 11, S. 1119–1125
- [Reiss 1991] REISS, Thomas H.: The revised fundamental theorem of moment invariants. In: *PAMI IEEE Trans. Pattern Analysis and Machine Intelligence* 13 (1991), S. 830–834
- [Reiss 1993] REISS, Thomas H.: *Recognizing Planar Objects Using Invariant Image Features*. Springer, 1993 (Lecture notes in computer science 676)
- [Rekers 1982] REKERS, G.: *Die Charakterisierung von Zuständen und Zustandsänderungen in Straßenszenen sowie deren Identifizierungen in TV-Bildfolgen*, Universität Hamburg, Diplomarbeit, 1982

- [Ridder, C. u. a. 1995] RIDDER, C. ; MUNKELT, OLAF ; KIRCHNER, H.: Adaptive background estimation and foreground detection using kalman-filtering. In: *In Proceedings International Conference on Recent Advances in Mechatronics*, 1995, S. 193–199
- [Ritter 2001] RITTER, Soltan: *Klassifikation von Änderungen in Bildfolgen*, Fachhochschule Esslingen und Robert Bosch GmbH FV/PLF2, Diplomarbeit, 2001
- [Rivlin, Ehud und Weiss, Isaac 1995] RIVLIN, EHUD ; WEISS, ISAAC: Local invariants for recognition. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17 (1995), Nr. 3, S. 226–238
- [Robertson, Mark A. u. a. 2000] ROBERTSON, MARK A. ; BORMAN, SEAN ; STEVENSON, ROBERT L.: Estimation-Theoretic Approach to Dynamic Range - Enhancement Using Multiple Exposures. In: *Image Processing* (2000)
- [Rohlf 1975] ROHLF, F. J.: Generalization of the gap test for the detection of multivariate outliers. In: *Biometrika* 31 (1975), S. 93–101
- [Rosner 1975] ROSNER, Bernard: On the detection of outliers. In: *Technometrics* 17 (1975), Nr. 2, S. 221–227
- [Rothe, Irene und Süsse, Herbert 1993] ROTHE, IRENE ; SÜSSE, HERBERT: Ein allgemeiner Zugang zur Berechnung von Invarianten. In: *Mustererkennung 1993, 15. DAGM-Symposium, Lübeck, , 1993, S. 67–74*
- [Rothe, Irene und Voss, Klaus 1992] ROTHE, IRENE ; VOSS, KLAUS: Orientierungsbestimmung von Objekten durch Momentinvarianten. In: *Mustererkennung 1992, 14. DAGM Symposium, Dresden, , 1992, S. 42–49*
- [Rothe, Irene u. a. 1994] ROTHE, IRENE ; VOSS, KLAUS ; OTHERS: A General Method to Determine Invariants / Fakultät für Informatik und Mathematik, Universität Jena. 1994. – Forschungsbericht. Fakultät für Informatik und Mathematik, Universität Jena
- [Rousseeuw, Peter J. und Leroy, Annik M. 1987] ROUSSEEUW, PETER J. ; LEROY, ANNIK M.: *Robust Regression and Outlier Detection*. Wiley & Sons, New York, 1987
- [Ruoff 1989] RUOFF, Wolfgang: *Optische Sensorsysteme zur On-line-Führung von Industrierobotern*. Springer Verlag, 1989
- [Sadjadi, Firooz A. und Hall, Ernest L. 1980] SADJADI, FIROOZ A. ; HALL, ERNEST L.: Three dimensional moment invariants. In: *PAMI IEEE Trans. Pattern Analysis and Machine Intelligence* 2 (1980), Nr. 2, S. 127–136
- [Salmon 1885] SALMON, George: *Lessons introductory to the Modern Higher Algebra*. 4th. Dublin : Figgis & Co., 1885
- [Sansone 1991] SANSONE, Giovanni: *Orthogonal Functions*. New York : Dover Publications, 1991
- [Sax 2001] SAX, Alexander: *Bau eines Teststandes für schnelle und hochdynamische CMOS-Bildverarbeitungssysteme*, Fachhochschule Esslingen, Diplomarbeit, 2001
- [Schanz 1998] SCHANZ, Michael: *Eindimensionale CMOS Bildsensorik mit integrierter Signalverarbeitung*, Universität Duisburg, Fachbereich Elektrotechnik, Dissertation, Jun 1998
- [Scharr 2000] SCHARR, H.: *Optimale Operatoren in der Digitalen Bildverarbeitung*. Heidelberg, Germany, University of Heidelberg, Dissertation, 2000

- [Schimpf 2000] SCHIMPF, U.: *Untersuchung des Gasaustausches und der Mikroturbulenz an der Meeresoberfläche mittels Thermographie*. Heidelberg, Germany, University of Heidelberg, Dissertation, 2000
- [Schimpf, Uwe u. a. 1999] SCHIMPF, UWE ; HAUSSECKER, HORST ; JÄHNE, BERND: Thermography for small-scale air-sea interaction. In: JÄHNE, BERND (Hrsg.) ; GEISSLER, PETER (Hrsg.) ; HAUSSECKER, HORST (Hrsg.): *Handbook of Computer Vision and Applications* Bd. 3: Systems and Applications. Academic Press, 1999, Kap. 35
- [Schmidt 2003] SCHMIDT, Matthias D.: *Prozessüberwachung beim Laserstrahlschweißen mit hochdynamischer Bildverarbeitung*, IFSW Institut für Strahlwerkzeuge Universität Stuttgart, Dissertation, 2003
- [Schneider, Bernd u. a. 1999] SCHNEIDER, BERND ; RIEVE, PETER ; BÖHM, MARKUS: Image Sensors in TFA (Thin Film on ASIC) Technology. In: JÄHNE, BERND (Hrsg.) ; GEISSLER, PETER (Hrsg.) ; HAUSSECKER, HORST (Hrsg.): *Handbook of Computer Vision and Applications* Bd. 1: Sensors and Imaging. Academic Press, 1999, Kap. 9
- [Schrey 2001] SCHREY, Olaf: *Methoden zur Dynamikerweiterung in der zweidimensionalen CMOS-Bildsensorik*, Fachbereich Elektronik, Gesamthochschule Duisburg, Dissertation, 2001
- [Schürmann 1977] SCHÜRMAN, Jürgen: *Polynomklassifikatoren für die Zeichenerkennung*. Oldenbourg Verlag, 1977
- [Schürmann 1996] SCHÜRMAN, Jürgen: *Pattern Classification - A Unified View of Statistical and Neural Approaches*. John Wiley, 1996
- [Schulz-Mirbach 1995] SCHULZ-MIRBACH, Hanns: Invariant features for gray scale images. In: *Mustererkennung 1995, 17. DAGM Symposium, Bielefeld, , 1995, S. 1–17*
- [Seger, Ulrich u. a. 1999] SEGER, ULRICH ; APEL, UWE ; HÖFFLINGER, BERND: HDRC-Imagers for Natural Visual Perception. In: JÄHNE, BERND (Hrsg.) ; GEISSLER, PETER (Hrsg.) ; HAUSSECKER, HORST (Hrsg.): *Handbook of Computer Vision and Application* Bd. 1: Sensors and Imaging. San Diego : Academic Press, 1999, Kap. II-8
- [Seiffert 2001] SEIFFERT, Thomas: *Messverfahren und Kenngrößen zur Beurteilung des dynamischen Kontrastaufklärungsvermögens elektronischer Bildaufnehmer*, Universität Karlsruhe (TH), Diplomarbeit, 2001
- [Seitz 1999] SEITZ, Peter: Solid-State Image Sensing. In: JÄHNE, BERND (Hrsg.) ; GEISSLER, PETER (Hrsg.) ; HAUSSECKER, HORST (Hrsg.): *Handbook of Computer Vision and Applications* Bd. 2: Signal Processing and Pattern Recognition. Academic Press, 1999, Kap. 14
- [Sethi, Ishwar und Patel, Nilesh V. 1995] SETHI, ISHWAR ; PATEL, NILESH V.: A Statistical Approach to Scene Change Detection. In: *SPIE Conference on Storage and Retrieval for Image and Video Database III* Bd. 2420, , 1995, S. 329–338
- [Siggelkow, Sven u. a. 2001] SIGGELKOW, SVEN ; SCHAEEL, MARC ; BURKHARDT, HANS: SIMBA - Search IMages By Appearance. In: *Mustererkennung 2001, 23. DAGM Symposium, München*, Springer, 2001, S. 7–16
- [Sikora 1997] SIKORA, Thomas: MPEG Digital Video-Coding Standards. In: *IEEE Signal Processing Magazine* 14 (1997), Nr. 5, S. 82–100
- [Simoncelli 1999] SIMONCELLI, E. P.: Bayesian Multiscale Differential Optical Flow. In: JÄHNE, BERND (Hrsg.) ; GEISSLER, PETER (Hrsg.) ; HAUSSECKER, HORST (Hrsg.): *Handbook of Computer Vision and Applications* Bd. 2: Signal Processing and Pattern Recognition. Academic Press, 1999, Kap. 14

- [Simonoff 1992] SIMONOFF, Jeffrey S.: General approaches to stepwise identification of unusual values in data analysis. In: STAHEL, WERNER. (Hrsg.) ; WEISBERG, S. (Hrsg.): *Directions in Robust Statistics and Diagnostics*. Heidelberg : Springer-Verlag, 1992, Kap. 2, S. 223–242
- [Soille 1999a] SOILLE, Pierre: *Morphological Image Analysis*. Springer-Verlag, 1999
- [Soille 1999b] SOILLE, Pierre: Morphological Operators. In: JÄHNE, BERND (Hrsg.) ; GEISSLER, PETER (Hrsg.) ; HAUSSECKER, HORST (Hrsg.): *Handbook of Computer Vision and Applications* Bd. 2: Signal Processing and Pattern Recognition. Academic Press, 1999, Kap. 21
- [Soloviev, A. V. und Schlüssel, P. 1994] SOLOVIEV, A. V. ; SCHLÜSSEL, P.: Parameterization of the Cool Skin of the Ocean and the Air-Ocean Gas Transfer on the Basis of Modeling Surface Renewal. In: *Journal of Physical Oceanography* 24 (1994), Nr. 6, S. 1339–1346
- [Stauder, Jürgen u. a. 1999] STAUDER, JÜRGEN ; MECH, ROLAND ; OSTERMANN, JÖRN: Detection of Moving Cast Shadows for Object Segmentation. In: *IEEE Transactions on Multimedia* 1 (1999)
- [Steen 1998] STEEN, William M.: *Laser material processing*. Springer, 1998
- [Sylvester 1878] SYLVESTER, James J.: Proof of the hitherto undemonstrated fundamental theorem of invariants. In: *Philosophical Magazine* 5 (1878), S. 178–188
- [Talenti 1987] TALENTI, Giorgio: Recovering a function from a finite number of moments. In: *Inverse Problems* 3 (1987), S. 5001–517
- [Teague 1980] TEAGUE, Michael R.: Image Analysis via the General Theory of Moments. In: *Journal of the Optical Society of America* 70 (1980), Nr. 8, S. 920–930
- [Teague 1982] TEAGUE, Michael R.: Calculation of irradiance moments. In: *Applied Optics* 21 (1982), S. 3292–3298
- [Teh, Cho-Huak und Chin, Roland T. 1986] TEH, CHO-HUAK ; CHIN, ROLAND T.: On digital approximation of moment invariants. In: *Computer Vision, Graphics and Image Processing* 33 (1986), S. 318–326
- [Teh, Cho-Huak und Chin, Roland T. 1988] TEH, CHO-HUAK ; CHIN, ROLAND T.: On image analysis by the methods of moments. In: *PAMI IEEE Trans. Pattern Analysis and Machine Intelligence* 10 (1988), Nr. 4, S. 496–513
- [Thompson 1935] THOMPSON, William R.: On a criterion for the rejection of observations and the distributions of the ratio of the deviation to the sample standart deviation. In: *Annals of Mathematical Statics* 6 (1935), S. 214–219
- [Tian 2000] TIAN, H.: *Noise Analysis in CMOS Image Sensors*, Ph.D Thesis, Stanford University, CA, Dissertation, 2000
- [Tian, Hui u. a. 2001] TIAN, HUI ; FOWLER, BOYD ; EL GAMAL, ABBAS: Analysis of Temporal Noise in CMOS Photodiode Active Pixel Sensor. In: *IEEE Journal of solid-state Circuits* 36 (2001), Nr. 1, S. 42
- [Tietjen 1972] TIETJEN, Gary L. Moore Roger H.: Some Grubbs-type statistics for the detection of several outliers. In: *Technometrics* 14 (1972), S. 583–597
- [Tietjen, Gary L. Moore Roger H. und Beckman, R. J. 1973] TIETJEN, GARY L. MOORE ROGER H. ; BECKMAN, R. J.: Testing for a Single Outlier in Simple Linear Regression. In: *Technometrics* 15 (1973), S. 717–721

- [Torres-Méndez, L. A. u. a. 2000] TORRES-MÉNDEZ, L. A. ; RUIZ-SUÁREZ, J. C. ; OTHERS: Translation, Rotation and Scale-Invariant Object Recognition. In: *IEEE Transactions On Systems, Man and Cybernetics, Part C: Applications And Reviews* 30 (2000), February, Nr. 1
- [Toth, Daniel und Aach, Til 2001] TOTH, DANIEL ; AACH, TIL: Improved minimum distance classification with Gaussian outlier detection for industrial inspection. In: *11th International Conference on Image Analysis and Processing Palermo, Italy* (2001)
- [Toth, Daniel u. a. 2000] TOTH, DANIEL ; AACH, TIL ; METZLER, VOLKER: Illumination Invariant Change Detection. In: *Southwest Symposium on Image Analysis and Interpretation 2000, Austin, Texas* (2000)
- [Toyama, K. und Wallflower 1999] TOYAMA, K. ; WALLFLOWER: Principles and practice of background maintenance. In: *Proceedings of IEEE International Conference on Computer Vision ICCV 1999*, 1999, S. 255–261
- [Trier, Oivind Due. u. a. 1996] TRIER, OIVIND DUE. ; JAIN, ANIL K. ; TAXT, TORFINN: Feature extraction methods for character recognition - a survey. In: *Pattern Recognition* 29 (1996), S. 641–662
- [Wallis 1976] WALLIS, R. H.: An approach to the space variant restoration and enhancement of images. In: *Proceedings of Symposium on Current Mathematical Problems in Image Science 1976*, 1976, S. 329–340
- [Wang, Yuan-Fang und Alferez, Ronald-Bryan O. 1998] WANG, YUAN-FANG ; ALFEREZ, RONALD-BRYAN O.: A Unified Framework for Image-Derived Invariants. In: *The 3rd Asian Conference on Computer Vision, Hong Kong*, , 1998, S. 542–549
- [Weiss 1993] WEISS, Isaac: Geometric Invariants and Object Recognition. In: *International Journal of Computer Vision* 10 (1993), Nr. 3, S. 207–231
- [Wilks 1963] WILKS, S. S.: Multivariate Statistical Outliers. In: *Sankhya Serie A.* 25 (1963), S. 407–426
- [Winkler 1995] WINKLER, Gerhard: Image analysis, random fields and dynamic Monte Carlo methods: A Mathematical Introduction. In: *Applications of Mathematics* (1995), Nr. 27
- [Wong, Robert Y. und Hall, Ernest L. 1978] WONG, ROBERT Y. ; HALL, ERNEST L.: Scene matching with moment Invariants. In: *Computer Graphics and Image Processing* 8 (1978), S. 16–24
- [Wood 1996] WOOD, Jeffrey: Invariant pattern recognition: a review. In: *Pattern Recognition* 29 (1996), Nr. 1, S. 1–17
- [Yakimovsky 1976] YAKIMOVSKY, Yoram: Boundary and object detection and real world images. In: *Journal of the ACM Association for Computing Machinery* 23 (1976), Nr. 4, S. 599–618
- [Yang 1995] YANG, Luren: Fast computation of 3-D geometric moments using a discrete Gauss' theorem. In: *Computer Analysis of Image and Pattern* Bd. 970. Springer, 1995, S. 649–654
- [Yang, Luren und Albreghsten, Fritz 1994] YANG, LUREN ; ALBREGSTEN, FRITZ: Fast and exact computation of moments using discrete Green's theorem. In: *Proceedings of NOBIM*, , 1994, S. 82–90
- [Yang, Luren u. a. 1997] YANG, LUREN ; ALBREGTSEN, FRITZ ; TAXT, TORFINN: Fast computation of three dimensional geometric moments using a discrete divergence theorem and a generalization to higher dimensions. In: *GMIP Graphical Models and Image Processing* 59 (1997), Nr. 2, S. 97–109
- [Zakaria, M. F. u. a. 1987] ZAKARIA, M. F. ; VROOMEN, L. J. ; OTHERS: Fast algorithm for the computation of moment invariants. In: *Pattern Recognition* 20 (1987), Nr. 6, S. 639–643

- [Zernike 1934] ZERNIKE, Fritz: Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode. In: *Physica* 1 (1934), S. 689–704
- [Zhang, H. u. a. 1993] ZHANG, H. ; KANKANHALLI, A. ; SMOLIAR S.W.: Automatic Partitioning of Full-motion Video. In: *Journal of Multimedia Systems* 1 (1993), Nr. 1, S. 10–28
- [Zhang, H. J. u. a. 1994] ZHANG, H. J. ; YOW, C. Y. ; OTHERS: Video Parsing Using Compressed Data. In: *Proceedings SPIE 1994 Symposium on Image and Video Processing, San José, California, 1994*, S. 142–149