

Diss. ETH Nr. 10196

**Experimenteller Vergleich von Korrelationskriterien in der
Bildanalyse**

ABHANDLUNG

Zur Erlangung des Titels

DOKTOR DER TECHNISCHEN WISSENSCHAFTEN

der

EIDGENÖSSISCHEN TECHNISCHEN HOCHSCHULE ZÜRICH

vorgelegt von

PETER FRANZ ASCHWANDEN

Dipl. El.-Ing. ETH
geboren am 1. Oktober 1960
von Altdorf und Zürich

Angenommen auf Antrag von:

Prof. Dr. W. Guggenbühl, Referent

Prof. Dr. O. Kübler, Korreferent

1993

Leer - Vide - Empty

meinen Eltern

Leer - Vide - Empty

DANK

Eine ganze Reihe von Verwandten und Bekannten haben direkt oder indirekt zum guten Gelingen dieser Arbeit beigetragen. Dazu gehören sicher meine Eltern, meine Schwestern und ihre Familien sowie verschiedene, gute Freunde aus Beruf, Freizeit und Studium.

Sehr wichtig war für mich auch die Unterstützung meines mehrjährigen Bürokollegen Bruno Schneuwly. Er war bei vielerlei Problemen stets ein kompetenter Gesprächspartner. Von ihm stammen im übrigen auch die Filter zur Generierung der unscharfen Bildern. Besonders geschätzt habe ich auch den Kontakt zu Guy Morbach. Das Einbinden der Bilder in dieses Buch wäre ohne seine Programmierkünste noch erheblich zeitaufwendiger gewesen.

Mein kleiner Erfolg beruht sicher zum Teil auch auf den früheren und jetzigen Mitgliedern der Gruppe für Videosignalverarbeitung und des ganzen Instituts für Elektronik. Ausserhalb dieses Kreises möchte ich besonders Herrn Dr. Federico Bonzanigo danken, der mir bei der Verfassung meiner Publikationen, die die eigentliche Basis dieser Dissertation darstellen, mit unvergleichlichem Elan und viel Erfahrung ausserordentlich geholfen hat.

Erwähnen möchte ich auch das Flugzeugwerk Emmen, das mich über mehrere Jahre hinweg im Rahmen eines Partnerprojektes unterstützte. Ganz speziell geht hier mein Dank an die Herren Dr. Utz, Dr. Monkewitz, David Parnitzky und Andreas Dürr.

Natürlich darf auch der Dank an meinen Referenten, Prof. Dr. Guggenbühl und meinen Korreferenten, Prof. Dr. Kübler nicht fehlen. Beide haben es verstanden einerseits eine sehr angenehme Atmosphäre zu schaffen und andererseits die Arbeit in verschiedenen Punkten zu lenken.

Zürich, 23. Juni 1993

Peter Aschwanden

Leer - Vide - Empty

Inhaltsverzeichnis

Inhaltsverzeichnis	1
Kurzfassung	3
Abstract	5
1. Einleitung	7
2. Terminologie	15
3. Definition der Korrelationskriterien	19
3.1. Korrelationskriterien ohne Bild-Vorverarbeitung	20
3.2. Korrelationskriterien mit Richtungsfiltern	25
3.3. Korrelationskriterien mit Bild-Hochpassfilterung	26
3.4. Binäre Korrelationskriterien	27
4. Versuche und Versuchsbedingungen	31
4.1. Allgemeines	31
4.2. Generierung der Bildsequenzen	37
4.2.1. Helligkeitsreduktion, IRIS	37
4.2.1.1. Signal-zu-Störleistungsverhältnis kontra Signal-zu- Störvarianzenverhältnis	39
4.2.2. Additives, Gauss'sches Rauschen, NOISE	41
4.2.3. Salt and Pepper Rauschen, SALTY	42
4.2.4. Unschärfe Bilder, FOCUS	43
4.2.5. Geometrische Expansion, ZOOM	45
4.3. Skala der Bildstörungsstärken	46
4.4. Praktische Relevanz der Bildsequenzen	46
4.5. Bemerkungen zu den Untersuchungen	47
5. Resultate	51
5.1. Beurteilungskriterien	51
5.2. Erwartete Detektionsratenverläufe	59
5.3. Diskussion der Detektionsratenverläufe	62
5.3.1. Korrelationskriterien ohne Bild-Vorverarbeitung	62
5.3.2. Korrelationskriterien mit Richtungsfiltern	68
5.3.3. Korrelationskriterien mit Bild-Hochpassfilterung	70
5.3.4. Binäre Korrelationskriterien	71
5.3.5. Zusammenfassung der Resultate	75

6. Rechenaufwand	77
6.1. Allgemeines	77
6.2. Verminderung des Rechenaufwandes	80
6.2.1. Extremwertsuche	80
6.2.2. Mittelwerte, Energieterme und Laplacefilterung	82
6.2.3. Verwendung von Frequenztransformationen	85
6.2.4. Richtungsfilter	87
6.2.5. Laplace-Gauss-Operator und Operator von Nack	89
6.2.6. Zusatzaufwand	90
6.3. Rechenaufwand der Korrelationskriterien	90
6.3.1. Gesamter Rechenaufwand für die KKF	90
6.3.2. Diskussion des Rechenaufwandes der einzelnen Korrelationskriterien	92
7. Diskussion von Beiträgen aus der Literatur	99
8. Schlussbemerkungen	123
9. Literaturverzeichnis	127
Anhang A.1: Nomenklatur	141
Anhang A.2: Begriffe	146
Anhang B: Bildmaterial	155
Anhang C: Objekt- und Suchfensterpositionen	175
Anhang D: Wirkung des Laplace-Gauss-Operators	179
Anhang E.1: Resultate, Grafiken	181
Anhang E.2: Resultate, Tabellen	207
Anhang E.3: Resultate, Korrelationsfunktionen	214
Anhang F: Liste der Abkürzungen	219
Lebenslauf	223

Kurzfassung

Verschiedene, in der Bildanalyse gebräuchliche Ähnlichkeitsmasse, die direkt auf (evtl. vorverarbeiteten) Intensitätsdaten arbeiten, werden bezüglich ihrer Leistungsfähigkeit als Musterlokalisatoren bei qualitativ schlechtem Bildmaterial und unterschiedlich grossen Mustern untersucht. Fünf wesentliche Störungsarten, repräsentiert durch Bildstörungsreihen, werden getestet. Die Störungsarten umfassen: Kontrast/Helligkeitsänderung, additives Gauss'sches Rauschen, 'Salt and Pepper' Rauschen, Unschärfe sowie geometrische Expansion. Jedes Korrelationskriterium wurde ca. 60'000 verschiedenen Versuchen mit Bildern unterschiedlichster Störungsstärke unterworfen. Die Auswertung der Tests erfolgt hauptsächlich mittels Analyse der Euklid'schen Distanz zwischen korrekter und gerechneter Objektposition. Im wesentlichen wird der Anteil detektierter Positionen, die innerhalb eines Toleranzkreises rund um die korrekte Objektposition liegen, als Gütemass für die Leistungsfähigkeit der Korrelationsverfahren benutzt. Auf diese Weise kann die Leistungsfähigkeit der Algorithmen bezüglich Störungsart, Störungsstärke, Mustergrösse und Mustertypus sehr einfach analysiert werden. Es zeigt sich z.B., dass bei den in der Stereo-Vision bzw. Bildkomprimierung (mittels block matching) am häufigsten verwendeten Korrelationskriterien (Summe der Differenzenbeträge: **SAD**, Summe der Differenzenquadrate: **SSD**) schon bei geringen, z.B. durch variierende Beleuchtung verursachte Helligkeitsschwankungen im Bild, starke Einbussen gegenüber der Leistung unter optimalen Bedingungen zu befürchten sind.

Neben diesen und anderen klassischen Verfahren werden auch verschiedene binäre und auf Hochpass-gefilterten oder richtungscodierten Bildern arbeitende Korrelationskriterien untersucht. Die damit verbundene Vorverarbeitung lässt die Anwendung von einfacheren und damit weniger rechenintensiven (Un-) Ähnlichkeitsmasse zu. Die Nachteile dieser Methoden treten besonders bei unscharfen oder geometrisch verzerrten Bildern zutage. In diesen Fällen wird die Detektionsrate durch die Vorverarbeitung meist wesentlich reduziert.

Zur Abschätzung des in der Regel recht grossen und damit bedeutungsvollen Rechenaufwandes wird dieser für alle Algorithmen im Detail diskutiert. Hierbei wird speziell auf die Charakteristik moderner Signalprozessoren Rücksicht genommen. Durch geschickte Implementation wie auch kon-

sequentes Ausnützen der besonderen Fähigkeiten des Zielsystems kann die Rechenzeit oft drastisch reduziert werden.

Einerseits soll die vorliegende Arbeit verschiedene aus der Literatur bekannte, teilweise aber auch neu definierte Korrelationsverfahren aufzeigen. Andererseits soll sie aber auch den potentiellen Anwender bei der Wahl und der Implementation eines entsprechenden Algorithmus' unterstützen (Bereich Bildverarbeitung, bei Videobildern mit 'normalen' Eigenschaften). Die Resultate zeigen, dass für maximale Erkennungsleistung, je nach zu erwartender Störungsart ein anderes Korrelationskriterium gewählt werden muss. Eine einzige, für alle Problemstellungen gleichermaßen optimal arbeitende Lösung existiert, wie die Versuche zeigen, nicht!

Da die zur Verfügung stehende Rechenleistung meist beschränkt ist, muss zudem oftmals ein Kompromiss zwischen der Menge der handhabbaren Situationen (Detektionsleistung des Korrelationsmasses) und dem Rechenaufwand gefunden werden. Die erarbeiteten Zahlen und Grafiken erlauben, trotz beschränkter Zahl an Experimenten, eine dem Problem angepasste Korrelator-Wahl zu treffen und die Leistungsfähigkeit vor der eigentlichen Implementation abzuschätzen. Zudem wird ein Versuchskonzept aufgezeigt, das einen einfachen und gut interpretierbaren Vergleich zwischen verschiedenen Korrelationsmassen zulässt.

Abstract

Several frequently used area correlation measures in image analysis are tested with respect to their performance in locating templates of different sizes within perturbed images. Five important perturbation types have been selected and applied to the test images: combined contrast and intensity decrease, additive zero mean Gaussian noise, 'salt and pepper' noise, image blur, and geometric image expansion. With every correlation criterion approximately 60,000 experiments have been performed. Perturbation strength has been varied within a large range, from no distortion to extremely heavy deformation. The performance of an algorithm is measured by the number of matches found within a tolerance area around the true match position. In this way the influence of perturbation type, perturbation strength, template type, and template size on performance can be analyzed easily. One interesting result is that the correlation algorithms commonly used in stereo vision and image sequence compression, the sum of squared differences (**SSD**) and the sum of absolute differences (**SAD**) respectively, perform poorly in the presence of barely visible, global changes in scene illumination or sensor sensitivity.

Beside these and other classical similarity and dissimilarity measures, binary correlation measures and correlation algorithms relying on high-pass filtered or orientation-coded images are considered. The preprocessing of the intensity data allows the use of simpler (dis-) similarity measures. The cost of this reduction of an otherwise huge computation load is a performance decrease, especially for blurred or geometrically transformed images.

Along with the ability to find correct matches the above-mentioned computation load is the most important feature of a correlation algorithm. The computational loads are analyzed theoretically and several proposals for efficient implementation are given (assuming a target system with a signal processor).

This work presents different, partly modified correlation criteria, gives a simple test setup for performance measurements of these algorithms, and establishes a performance scale which supplies the crucial information for algorithm selection to the systems engineer (for image analysis applications with 'normal' video images). The results indicate that the optimal selection has to be accomplished according to the expected perturbation type of the

envisioned application. No single correlation measure performs best under all conditions.

Simpler algorithms often have to be selected because of lack of sufficient computation power. The combined occurrence of different perturbations could also necessitate a compromise for best overall performance (the problem of combined perturbations is not considered here). Although the given numbers and figures rely on a limited number of experiments, they allow for a well-founded selection of correlation criteria in the field of image analysis.

1. Einleitung

In der Robotik wie auch in anderen Gebieten der Technik werden zur Steigerung der Autonomie von Maschinen in zunehmendem Masse Komponenten eingesetzt, die die Umwelt selbständig erfassen und in beschränktem Ausmass auch verstehen können. Neben den herkömmlichen Sensoren wie z.B. Schalter, Drucksensoren oder Temperaturfühlern kommen immer häufiger auch bildgebende Sensoren zum Einsatz. Mit den heute handelsüblichen CCD-Videokameras ist z.B. ein guter Sensor zur zweidimensionalen Erfassung des Helligkeitseindrucks einer Szene vorhanden. Neben diesen Sensoren ([18]), die Datenraten bis teilweise über 10MByte pro Sekunde unterstützen, sind aber auch die Digitalisierung der analogen Information, die Rechnerarchitektur und die Algorithmen die auf dieser Hardware implementiert werden von entscheidender Bedeutung zur Lösung obiger Aufgabenstellung – des Maschinen-Sehens. Da – trotz vielerlei Anstrengungen von diversen Forschungslaboratorien – keine allgemeine Lösung zu diesem Problem bekannt ist und der Rechenaufwand sehr stark ansteigt, wenn eine vielen Ansprüchen genügende Lösung gesucht wird, werden heute bei konkreten Implementationen, je nach Bedarf, verschiedene Systemfähigkeiten einprogrammiert. Nicht unbedingt benötigte Eigenschaften werden zur Vereinfachung des Problems und damit zur Verringerung des Aufwandes ausser Acht gelassen. Häufig können auch durch die Wahl eines geeigneten mechanischen und optischen Aufbaus die Anforderungen an das Bildverarbeitungssystem erheblich reduziert werden.

Im allgemeinsten Sinn geht es beim Maschinen-Sehen immer um die Detektion, die Klassifizierung, die Erkennung oder das Lokalisieren von Objekten im Raum. Informationen dieser Art können z.B. mit Hilfe von Auswertelgorithmen dazu benutzt werden, Roboter in einer nur schwer vorhersehbaren Umgebung selbständig arbeiten zu lassen. Der Weg vom Bild bis zur Aktion des Roboters kann je nach Aufgabenstellung sehr komplex sein. Einfache Algorithmen zur geeigneten Vorverarbeitung der Bildinformation bis hin zu komplexen Auswerteverfahren, die das Verständnis der Szene zum Ziel haben, müssen eingesetzt werden. In konventionellen Systemen zur vierdimensionalen Erfassung einer Szene (Raum und Zeit) kann die algorithmische Struktur grob wie folgt dargestellt werden:

1. Datenerfassung (Bildaufnahme, d.h. im Normalfall: Erfassung der Szene mit einem bildgebenden Sensor, Digitalisierung und Speicherung der Information).
2. Vorverarbeitung (z.B. Filterung der Bilddaten zur Rauschunterdrückung oder zur Kantenextraktion, sowie Bildtransformationen).
3. Extraktion von komprimierten Daten aus der vorverarbeiteten Bildinformation, normalerweise mit wesentlichem Informationsverlust (Konzentration auf interessierende Daten; z.B. Position oder Ausrichtung eines Helligkeitsmusters).
4. Auswertung der komprimierten Daten (z.B. erstellen einer dreidimensionalen Datenbasis, 'Verstehen der Szene').
5. Bestimmung einer geeigneten Reaktion des Systems auf die Aussenwelt (z.B. Roboterbewegung).

Die aufgeführte Liste könnte noch stark verfeinert werden, z.B. kann die Auswertung der komprimierten Daten auf vielfältigste Weise erfolgen. Je nach Anwendung werden heutzutage zur Vereinfachung des Problems und zur Geschwindigkeitssteigerung einer konkreten Applikation einzelne der oben aufgeführten Schritte ganz weggelassen oder mit Hilfe von Vorkenntnissen über das Problem vereinfacht. Z.B. ist in verschiedenen Anwendungen eine zweidimensionale Auswertung völlig genügend (z.B. Chipfabrikation, einfache Roboteranwendungen mit definierter Tiefe der Objekte). Die hier präsentierte Arbeit ist zwar in diesem riesigen Umfeld zu sehen, im Detail werden aber nur ausgewählte Lösungsansätze zum dritten Punkt der obigen Liste behandelt (und allenfalls Punkt 2).

Neuere Lösungsansätze in der Bildanalyse gehen auch weg von den klassischen Methoden in Richtung besonderer, auf das spezielle Problem angepasster Sensoren oder aber in Richtung einer umfassenden, verteilten Bildanalyse, z.B. mit Hilfe von neuronalen Netzen, die das menschliche bzw. tierische Sehen nachempfinden sollen ([28], [88], [106], [123] und [157]) oder Fuzzy Logic ([143]). Trotz der Bedeutung, die diese neuen Verfahren zweifellos haben, sind die konventionellen Methoden, zu denen auch die an dieser Stelle diskutierten Korrelationsverfahren zählen, z.B. wegen dem momentan (noch) besseren Verständnis sowie der normalerweise grösseren, erreichbaren Präzision nicht aus der heutigen und zukünftigen Bildverarbeitung wegzudenken.

Zum oben angesprochenen Kreis der Algorithmen, die den Übergang von bildhaften auf abstrakte Daten bewerkstelligen, können eine Vielzahl von Rechenverfahren gezählt werden. Als Beispiel soll nur das hier im Detail diskutierte Lokalisieren von Helligkeitsmustern in Bildern aufgeführt werden.

Die Bestimmung der Position eines Musters kann auf verschiedene Weise erfolgen. Symbolische Verfahren, die die Suche nicht direkt auf den Bilddaten, sondern mit Hilfe einer symbolischen Beschreibung des Bildinhaltes vornehmen, sind normalerweise wegen der geringeren zu bearbeitenden Datenmenge beim eigentlichen 'Matching-Prozess' relativ schnell (je nach Art der symbolischen Beschreibung). Nachteile dieser Methoden sind aber das meist recht aufwendige Erstellen der symbolischen Beschreibung, welche z.B. Elemente wie 'Linie', 'Kreuzung', 'Kreisbogen', 'mittlere, lokale Helligkeit' etc. enthalten kann, sowie die Abhängigkeit vom Helligkeitsrauschen bei der Verwendung von Elementen, die auf räumlichen Ableitungen des Bildes beruhen (z.B. Kanten). Es soll aber nicht verschwiegen werden, dass diese Verfahren ein grosses Leistungspotential haben, wenn z.B. nach Mustern gesucht wird, von denen nur grobe Beschreibungen der Form, Textur, Ausdehnung oder Ausrichtung vorliegen ([161]).

Die symbolischen Verfahren werden hier aber nicht weiter diskutiert. Vielmehr soll das Augenmerk auf die algorithmisch erheblich einfacheren, z.T. aber sehr rechenintensiven, korrelativen Ähnlichkeitsmasse gerichtet werden (area correlation). Sie basieren alle darauf, dass ein Ähnlichkeits- bzw. Unähnlichkeitsmass auf alle möglichen Gebiete (Korrelationsfenster) in einem Suchbereich angewendet wird. Zum Vergleich werden die (evtl. vorverarbeiteten) Pixelwerte des Musters und des Suchfensters herangezogen. Die Position mit maximaler Übereinstimmung zwischen Muster und Suchfensterausschnitt wird im einfachsten Fall durch Detektion des Extremums in der Menge aller in Frage kommenden (Un-) Ähnlichkeitswerte bestimmt.

Anwendungen solcher Algorithmen sind z.B. bei Positionsbestimmungen von mehr oder weniger genau bekannten Helligkeitsmustern (z.B. Vermessung von Markierungen), bei der Bildkompression (block matching), bei Trackingproblemen (Analyse eines dynamischen Vorgangs, [11] und [12]) oder auch in ähnlicher Weise bei der Tiefenbildgewinnung durch das 'depth from motion'-Verfahren ([140]) zu finden. Ein wesentlicher Nachteil dieser Korrelationskriterien besteht sicher darin, dass nicht die Korrespondenz eines einzelnen Punktes, sondern die Zuordnung einer ganzen Fläche gesucht wird. Gerade bei Vermessungsaufgaben kann dies z.B. wegen variierender Tiefe innerhalb des gesuchten Musters zu Problemen führen. Dieses Handicap, wie auch die Schwierigkeiten, die bei spiegelnd reflektierenden Objek-

ten im Zusammenhang mit Korrelationskriterien auftreten, werden hier nur angetönt, aber nicht diskutiert.

Viele Korrelationsalgorithmen erweisen sich als ausserordentlich robust gegenüber Helligkeitsrauschen. Auch gegenüber linearen Helligkeitstransformationen (exklusive Kontrastumkehr) sind die Algorithmen in der Regel äusserst insensitiv. Dagegen ist eine gute Robustheit gegenüber geometrischen Bildtransformationen¹ nur mittels spezieller Vorkehrungen, z.B. anhand des verfügbaren Vorwissens, mit vertretbarem Rechenaufwand zu erreichen (Auflösungspyramide, Verwenden eines geeigneten Koordinatensystems, Lernfähigkeit zur Adaption des Musters, [12]).

Die vorliegende Arbeit hat das Ziel die unterschiedlichen Eigenschaften der verschiedenen Korrelationskriterien aufzuzeigen und so dem Anwender die Algorithmen-Wahl (in diesem kleinen Bereich) zu vereinfachen. Die weiter unten definierten Korrelationsverfahren werden bezüglich folgender Charakteristiken experimentell untersucht:

- Robustheit gegenüber Helligkeitsänderungen
- Robustheit gegenüber additivem, mittelwertfreiem Gauss'schen Rauschen
- Robustheit gegenüber 'Salt and Pepper' Rauschen
- Robustheit gegenüber Schärfeänderungen im Bild
- Robustheit gegenüber geometrischen Transformationen (Vergrösserung)
- Rechenaufwand

Diese Eigenschaften wurden mit Hilfe von speziell generierten Bildsequenzen in einer grossen Anzahl von Versuchen getestet. Neben der Robustheit der Korrelationskriterien gegenüber verschiedenen Störungsarten wurde insbesondere auch der Einfluss der Mustergrösse im Bereich von 2x2 bis 25x25 Pixel genauer untersucht.

Der Rechenaufwand der verschiedenen Korrelationskriterien wurde rein theoretisch überprüft, aber zum Teil mit experimentellen Werten verglichen. Bei der Bewertung der Operationen wurden speziell die Fähigkeiten von modernen Signalprozessoren berücksichtigt.

Die betrachteten Störungen werden in der Praxis einerseits durch den Sensor und dessen Elektronik (Gauss'sches Intensitätsrauschen, Helligkeits-

¹ Z.B. verursacht durch Veränderung der Kameraposition relativ zur Szene oder Änderung der Kameraobjektivbrennweite.

änderungen, siehe [18]) und andererseits durch die Aufnahmebedingungen bzw. die Szene selber verursacht (Helligkeitsänderungen, geometrische Transformationen, Unschärfe). Das in der Regel selten auftretende Salt and Pepper Rauschen kann durch Übertragungs- oder Elektronikfehler erzeugt werden. Diese Störungsart wird aber auch zur Nachbildung von Muster-Verdeckungen verwendet. Dabei darf nicht unberücksichtigt bleiben, dass im praktischen Fall die Störpixel gewöhnlich in Clustern auftreten, was durch (gleichverteiltes) Salt and Pepper Rauschen nicht nachgebildet wird. Es wird also nur der Informationsausfall, der aus Verdeckungen resultiert wirklichkeitsgetreu wiedergegeben. Die geometrische Verteilung dieses Datenverlustes entspricht hingegen meist nicht der Praxis.

Die Wirkung der untersuchten Störungen kann teilweise durch entsprechenden Betrieb der Kamera bzw. Einflussnahme auf die Szene reduziert werden. Wichtige Elemente sind dabei z.B. die Beleuchtung der Szene, die automatische Verstärkungsregelung der Videokamera (AGC), die automatische Blendenregelung, die Fokussierung und die Brennweite des Objektivs, die Belichtungszeit des Sensors (shutter speed), der Sensortyp der Kamera und die spektrale Empfindlichkeit des Sensors (Filter). In der Regel lassen sich aber die Aufnahmebedingungen und die Kameraeinstellungen nicht völlig frei wählen. Deshalb muss immer auch algorithmisch versucht werden die geforderten Robustheitsspezifikationen zu erfüllen.

Das den Untersuchungen zugrunde liegende Bildmaterial ist mittels einer konventionellen CCD-Videokamera erfasst und (meist) rechnerisch verändert worden. Die Bilder sind, wie in der Mehrzahl der Anwendungen der Bildanalyse, intensitätscodiert (mittelwertbehaftete Codierung) und weisen, im Gegensatz zu synthetischen Grafikbildern, eine hohe Korrelation zwischen benachbarten Pixeln auf. Bilder von Grafikqualität sind nicht untersucht worden.

Die hier vorgestellten Versuche wurden auf IBM-kompatiblen Personal Computern durchgeführt². Zur Beschleunigung der Voruntersuchungen wurde auch auf Signalprozessorsysteme zurückgegriffen, die die Rechenzeit wesentlich verkürzten. Um Unstimmigkeiten, verursacht durch unterschiedliche Datenformate und Rechengenauigkeiten, zu vermeiden, sind die dokumentierten Versuche aber allesamt auf PCs mit der 64Bit Fließkomma-Darstellung des Co-Prozessors durchgeführt worden.

Experimentelle Untersuchungen haben dort ihre Daseinsberechtigung, wo ein beschränkter Ereignisraum, ungenügend mächtige Mathematik oder

² 33MHz Intel 80486DX

schlechte Modelle vorliegen. Im hier diskutierten Fall ist der quantitative, theoretische Vergleich der verschiedenen Verfahren schwierig, da die unterschiedlichen, z.T. nichtlinearen (Vorverarbeitungs-) Operationen zusammen mit den durch die digitale Verarbeitung bedingten Quantisierungseffekten mathematisch nur sehr schwer einheitlich zu fassen sind. Die Störmodelle sind zwar meist relativ einfach (z.B. [58], [93], [138] und [139]), deren mathematische Analyse, besonders zusammen mit einer gültigen mathematischen Beschreibung von typischen Szenen, in der Regel aber äusserst schwierig und häufig nur unter Zuhilfenahme von (zu) stark vereinfachenden Annahmen möglich³. Einige Untersuchungen, z.B. die Analyse des Verhaltens der Algorithmen bei Helligkeits- oder Kontrastschwankungen, könnten aber für einzelne Methoden sehr wohl theoretisch angegangen werden ([13]). Der Vergleich solcher Resultate bleibt allerdings trotzdem problematisch.

Obwohl die Ereignismenge nie auch nur annähernd vollständig durch Experimente abgedeckt werden kann, ist der *Versuch* als Basis der vorliegenden Untersuchungen gewählt worden. Es ist darum immer zu beachten, dass die Ergebnisse nicht ohne Relativierung auf andere Fälle übertragen werden dürfen. Die Resultate der verschiedenen Testreihen weisen aber auf eine gute Aussagekraft bezüglich der Leistungsdaten der Algorithmen hin.

Der nachfolgende Text ist folgendermassen gegliedert: Auf die Einleitung folgt im Kapitel 2: *Terminologie* ein kurzer Einblick in die verwendeten Begriffe. Ausführlichere Definitionen erscheinen entweder direkt im Text oder aber im Anhang A.1: *Nomenklatur* oder A.2: *Begriffe*. In Kapitel 3: *Definition der Korrelationskriterien* werden die untersuchten Korrelationskriterien vorgestellt. Kapitel 4: *Versuche und Versuchsbedingungen* befasst sich mit dem verwendeten Bildmaterial, dem Versuchsablauf und dem Auswertekonzept. Zur Illustration sind im Anhang B: *Bildmaterial* und Anhang D: *Wirkung des Laplace-Gauss-Operators* ein Teil der verwendeten Bilder festgehalten. Im Anhang C: *Objekt- und Suchfensterpositionen* werden zudem die interessierenden Bildausschnitte definiert.

Die Versuchsergebnisse werden im Abschnitt 5: *Resultate* diskutiert. Ausführlichere Grafiken und Tabellen zu diesem Thema sind auch im Anhang E.1: *Resultate, Grafiken* und E.2: *Resultate, Tabellen* zu finden. Kapitel 6: *Rechenaufwand* liefert einen Einblick in den Rechenaufwand der verschiedenen Korrelationskriterien; daneben werden auch verschiedene Methoden zur effizienten Implementation der besprochenen Verfahren präsentiert.

³ Z.B. vorausgesetzte Statistik (Mittelwert und Varianz) muss auch in kleinen Fenstern erfüllt sein!

Da die einzelnen Begriffe und Verfahren erst nach Kapitel 6 vollständig bekannt sind, ist die Diskussion der Literaturreferenzen an diesen späten Platz gelegt worden. Neben Artikeln, die die hier ebenfalls behandelten Korrelationskriterien erörtern, werden im Kapitel 7: *Diskussion von Beiträgen aus der Literatur* auch eine Reihe von Arbeiten, die weitere interessante Algorithmen, Ideen und Untersuchungen vorstellen, besprochen. Eine kurze Wertung der Versuche und Resultate wird in Kapitel 8: *Schlussbemerkungen* gegeben. Im anschliessenden Literaturverzeichnis sind die referenzierten Artikel aufgeführt.

2. Terminologie

Zum Verständnis der folgenden Kapitel muss zunächst eine gemeinsame Sprache definiert werden. Darum sollen nun kurz einige der wichtigsten Begriffe eingeführt werden. Weitere Definitionen sind im Anhang A.1: *Nomenklatur* und A.2, *Begriffe* zu finden. Im Text werden einzelne der hier aufgeführten Begriffsbestimmungen, falls sie von grosser Wichtigkeit sind, wiederholt. Die weiter unten angegebenen Grössenangaben sind selbstverständlich nicht allgemein gültig, sondern nur für die hier vorgestellten Untersuchungen repräsentativ.

Das Bild $b(x,y,Nr)$ enthält die mit sechs Bit aufgelöste Helligkeitsinformation einer Szene. Die Kantenlänge der quadratischen Bilder beträgt je 256 Pixel. Jedes dieser Bilder ist Teil einer *Bildsequenz*. Die Reihenfolge der Bilder innerhalb dieser Gruppe ist durch den Parameter Nr gegeben. Bilder mit hohem Wert des Parameters Nr sind immer stärker gestört als solche mit kleinem Wert.

Das gesuchte *Muster* oder *Referenzfenster* wird durch $r(u,v)$ repräsentiert. Es enthält ebenfalls helligkeitscodierte Pixel und wird aus dem ungestörten Bild der Bildsequenz extrahiert ($Nr=0$). Die Grösse des Musters ist durch $xlr=umax+1$ und $ylr=vmax+1$ gegeben (2x2 bis 25x25 Pixel).

Innerhalb der Bilder $b(x,y,Nr)$ werden rechteckige *Suchfenster* der Grösse $xls=xlr+39$ bzw. $yls=ylr+39$ platziert. Somit werden immer, unabhängig von der Referenzfenstergrösse, insgesamt 40x40 im Suchfenster vollständig enthaltene, rechteckige Regionen sog. *Korrelationsfenster* ($k(x,y)$, xlr mal ylr) auf ihre Ähnlichkeit mit dem Muster überprüft. Bei der Korrelation werden ausschliesslich die im Suchfenster enthaltenen Helligkeitswerte verwendet. Alle dem Suchfenster angehörenden Pixel tragen auch zur Musterlokalisierung bei.

Das *Objektfenster* bezeichnet dasjenige Korrelationsfenster, das gemäss dem gewählten Korrelationsverfahren am besten mit dem Muster übereinstimmt (*Objektposition*). Am Ort der **objektiv** besten Übereinstimmung zwischen Muster und Korrelationsfenster befindet sich das *wahre Objektfenster*. Bei einer Anwendung ist im Normalfall die Position dieses Fensters unbekannt. Durch die kontrollierte Generierung der in dieser Untersuchung benutzten Bilder ist die *wahre Objektposition* aber jederzeit genau bestimmt und liegt immer innerhalb des Suchfensters.

Das Auffinden der Objektposition wird *Positionierung* genannt und umfasst die Vorverarbeitung der Intensitätswerte, die Berechnung aller Korrelationswerte bzw. (*Un-*) *Ähnlichkeitsmasse* in der *Korrelationsfläche* sowie die Lokalisierung des entsprechenden Extremums mit Hilfe eines einfachen Maximum- oder Minimumdetektors. Die Korrelationsfläche enthält alle potentiellen Objektpositionen und nimmt in den vorgestellten Versuchen eine Fläche von 40x40 Pixeln ein.

Eine Überprüfung des Positionierungsergebnisses anhand von Plausibilitätsbetrachtungen wurde bei den Versuchen nicht vorgenommen. Der detektierte Extremwert wurde z.B. ungeachtet der Grösse des Korrelationswerts mit der Position des gesuchten Helligkeitsmusters identifiziert, eine Rückweisung des Ergebnisses wegen zu geringer Güte also nicht in Betracht gezogen.

Falls die Vektordistanz zwischen errechneter und wahrer Objektposition einen Betrag aufweist der kleiner oder gleich einem vorgegebenen Radius r_t ist, so wird von einer *tolerierbaren Positionierung* gesprochen. Im anderen Fall muss dieselbe Positionierung als untolerierbar zurückgewiesen werden.

Anhand der *Detektionsrate* kann die Leistungsfähigkeit der verschiedenen Korrelationsalgorithmen beurteilt werden. Sie gibt den Anteil tolerierbarer Positionierungen gemessen an der momentan betrachteten (Unter-) Menge der Positionierungen an. Die Menge der betrachteten Positionierungen kann sich z.B. durch eine bestimmte Mustergrösse, einen speziellen Bildstörungstyp oder eine genau definierte Bildstörungsstärke auszeichnen. Natürlich könnte man die Leistungsfähigkeit der Algorithmen z.B. auch anhand des Signal-zu-Rauschleistungsverhältnisses in der Korrelationsebene diskutieren. Obwohl dieses Gütemass – speziell bei theoretischen Untersuchungen – öfters verwendet wird und entsprechende Resultate publiziert werden (Übersicht über derartige Leistungskriterien in [72]), wird hier die Detektionsrate als Vergleichsgrösse verwendet. Prinzipiell ist der Ingenieur ja meist nicht am Signal-zu-Rauschleistungsverhältnis, sondern vielmehr an der Wahrscheinlichkeit für eine erfolgreiche Positionierung, also an der Detektionsrate interessiert. Zudem ist auch die Unterscheidung zwischen Signal- und Rauschleistung in der Korrelationsebene problematisch. –Das Signal-zu-Rauschleistungsverhältnis in der Korrelationsebene ist ein Mass für die Detektierbarkeit einer Korrelationsspitze und hat darum bei analogen oder optischen Implementationen aber eine gewisse Bedeutung.

Wichtige Begriffe beim Vergleich der Korrelationskriterien werden die *Robustheit* und die *Sensitivität* (gegenüber genau zu spezifizierenden Bildstörungen) sein. Sie werden als Antonyme verwendet und sind dementspre-

chend austauschbar. Ferner bezeichnen diese Begriffe, wenn sie nicht vergleichend (relativ), sondern nur auf ein einzelnes Korrelationskriterium bezogen, als absolute Wertung verwendet werden, eine spezielle Eigenschaft eines Algorithmus: Ein Korrelationskriterium wird dann als robust gegenüber einer Bildstörung charakterisiert, falls sich gemäss theoretischen Überlegungen die Reihenfolge der einzelnen Korrelationswerte, d.h. die prinzipielle Struktur der Korrelationsfunktion durch die Bildstörung nicht ändern lässt. So bleibt also bei einem gegenüber einer Störung (theoretisch) robusten Korrelationskriterium die Position des Extremwertes auch bei stärkster Bildstörung fix am selben Ort. Die in digitalen Systemen unvermeidlichen Quantisierungs- und Sättigungseffekte bleiben dabei unbeachtet! Im praktischen Experiment kann der Einfluss dieser beiden Phänomene aber einfach nachgewiesen werden.

Die *Invarianz* gegenüber einer Bildstörung verlangt im Vergleich zur Robustheit zusätzlich, dass die einzelnen Korrelationswerte durch die Störung nicht verändert werden. Auch dieser Begriff stellt eine theoretische Idealisierung dar, die in der Praxis nicht vollständig erfüllt werden kann.

Bei der Diskussion des Bildmaterials und der Resultate wird immer wieder auf die Bild- oder Störsignal-Leistung bzw. -Energie (oder entsprechende Verhältnisse davon) eingegangen. Beide Begriffe, *Leistung* und *Energie*, werden verwendet, obwohl sie physikalisch zwar eng verknüpft, aber nicht identisch sind. Jeder einzelne Bildsensor misst die während einer gewissen Zeitspanne auf ihn treffende 'Licht'-Energie. Diese 'Messung pro Zeiteinheit' entspricht natürlich genau einer (abgetasteten) Leistungsbestimmung.

Grundsätzlich wird in dieser Arbeit von Signalenergie gesprochen. Bei Energie-Verhältnissen wird von dieser Regel aber abgewichen, da der Term *Signal-zu-Rauschleistungsverhältnis* ein stehender Begriff in der Technik darstellt.

3. Definition der Korrelationskriterien

Für die weiter unten beschriebenen Tests wurden nur Korrelations-ähnliche Verfahren ausgewählt. Der Begriff 'Korrelation' wird im folgenden als Oberbegriff für alle hier besprochenen (Un-) Ähnlichkeitsmasse verwendet.

Die nachstehenden Tabellen (Tab.1-4) zeigen die formelmässigen Definitionen der untersuchten Algorithmen. Die verwendete Nomenklatur sowie die Details der Vorverarbeitung gehen aus Anhang A.1: *Nomenklatur* hervor. Die Korrelationskriterien sind anhand der involvierten Vorverarbeitungsmethoden in vier Hauptklassen eingeteilt:

- Algorithmen ohne Bild-Vorverarbeitung, die also direkt auf den Intensitätsdaten arbeiten (Tab.1)
- Algorithmen, die die lokale Orientierung des Helligkeitsgradienten verwenden (Tab.2)
- Algorithmen, die auf Hochpass-gefilterten Daten arbeiten (Tab.3)
- Algorithmen, die auf zweistufigen, binären Bildern arbeiten (Tab.4)

Vier der unten aufgeführten Verfahren sind Variationen von Original-Algorithmen, die Nack bzw. Seitz vorschlugen und vom Autor neu definiert wurden, um evtl. eine Verbesserung der Detektionsrate zu erreichen ([122], [156] bzw. [145]).

Es ist zu beachten, dass die Auswahl der (Un-) Ähnlichkeitsmasse und der Vorverarbeitungsmethoden in keiner Weise vollständig ist. Die Wahl wurde durch verschiedene Faktoren beeinflusst: Benötigte Rechenleistung, anvisierte Anwendungen, Vorversuche sowie verschiedene Überlegungen bezüglich der Vergleichbarkeit der Korrelationsverfahren.

Beim Vergleich dieser Arbeit mit anderen Untersuchungen ist darauf zu achten, dass die Bezeichnungen der Korrelationskriterien nicht in jedem Fall übereinstimmen und z.T. sogar widersprüchlich sind. Z.B. wird unter dem Begriff *normierter Korrelation* vielfach die lokal mittelwertfreie, normierte Kreuzkorrelation verstanden. Aus Konsistenzgründen werden hier eigene, leicht zu interpretierende und teilweise auch in anderen Artikeln verwendete Bezeichnungen benutzt. -Alle Korrelationsalgorithmen werden in ihrer diskreten, endlichen Form definiert. Im Anhang E.3: *Resultate, Korrelationsfunktionen* sind grafische Darstellungen von Korrelationsfunktionen aller untersuchten (Un-) Ähnlichkeitsmasse beigeheftet.

3.1. Korrelationskriterien ohne Bild-Vorverarbeitung

Das von der Mathematik vorgeschlagene Instrument zur Bestimmung der Ähnlichkeit zweier Signale, die *unnormierte Kreuzkorrelation*, im englischen als *direct correlation* oder einfach *cross-correlation* bezeichnet, kann bei unbearbeiteten, endlich ausgedehnten und mittelwertbehafteten Bildern nicht erfolgreich eingesetzt werden, da die vermeintlich grösste Übereinstimmung am Ort mit lokal maximaler Helligkeitsaussteuerung erreicht wird. Gute Abhilfe schafft hier die Normierung dieser Funktion mit dem geometrischen Mittelwert der (lokalen) Energien im Suchfenster und im Muster (**KKF**, *normierte Kreuzkorrelationsfunktion*, *correlation coefficient*, *bezogener Korrelationskoeffizient*; Herleitung: [137] oder [104]; Diskussion und Anwendung: [10], [14], [17], [48], [49], [52], [59], [64], [75], [92], [127], [134], [144], [151], [154], [164] und [174], Tab.1). Diese Funktion ist invariant gegenüber multiplikativen Helligkeitsänderungen im Suchfenster ([13]). Dagegen ist derselbe Algorithmus sensitiv bezüglich additiven Helligkeitsänderungen. Die Invarianz gegenüber multiplikativen wie auch additiven Helligkeitsänderungen kann z.B. durch die Anwendung der *lokal mittelwertfreien, normierten Kreuzkorrelationsfunktion* gewonnen werden (**KKFMF**, Herleitung: [139]; weitere Artikel: [14], [33], [52], [67], [69], [71], [75], [76], [96], [115], [117], [137], [147], [148], [156]). Von allen Intensitätswerten wird dabei der lokale, im entsprechenden Korrelationsfenster beobachtete, Helligkeitsmittelwert abgezogen. Dies resultiert schliesslich in der Normierung mit dem geometrischen Mittelwert der lokalen Signalvarianzen. Deshalb wird diese Korrelationsfunktion auch häufig *variance normalized correlation* oder *bezogener Kovarianzkoeffizient* genannt. Muster, die sich nur durch additive oder multiplikative, lokale (über das gesamte Korrelationsfenster identische) Helligkeitsdifferenzen unterscheiden, sind nicht mehr diskriminierbar, da sich dieses Korrelationskriterium gegenüber diesen Helligkeitstransformationen als invariant erweist. Die Subtraktion des lokalen Mittelwerts entspricht eigentlich einer Hochpassfilterung und hat damit natürlich einen ähnlichen, die Korrelationsspitze verstärkenden Einfluss. Die Helligkeitswerte werden in 'oberhalb' und 'unterhalb' des Mittelwerts eingeteilt. Je nachdem tragen sie darum, ganz im Gegensatz zur **KKF**, positiv oder negativ zum Zähler der **KKFMF** bei. D.h. Pixel, die nicht zur gleichen Helligkeitsklasse gehören, wie dies häufig bei geringer Ähnlichkeit zwischen Muster und Objektfenster vorkommt, reduzieren tatsächlich den Korrelationswert. -In [139] wird von Ryan und Hunt nachgewiesen, dass die **KKFMF** den Maximum Likelihood Korrelator realisiert, falls die Pixel nur durch additive und multiplikative Helligkeitsänderungen sowie mittelwertfreies Gauss'sches Rauschen verändert werden.

Moravec stellt ein weiteres interessantes Ähnlichkeitsmass vor (MOR, [119]). Der Vorteil dieser Funktion gegenüber den bisher definierten Verfahren liegt hauptsächlich beim Rechenaufwand bzw. bei der geringeren Komplexität der notwendigen Operationen. Gegenüber der **KKF** muss z.B. keine

Name	Definition
normierte Kreuzkorrelationsfunktion (KKF(x,y))	$\frac{\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} r(u,v) \cdot s(x+u,y+v)}{\sqrt{\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} r^2(u,v) \cdot \sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} s^2(x+u,y+v)}}$
lokal mittelwertfreie, normierte Kreuzkorrelationsfunktion (KKFMF(x,y))	$\frac{\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{r(u,v) - \bar{r}\} \cdot \{s(x+u,y+v) - \overline{s(x,y)}\}}{\sqrt{\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{r(u,v) - \bar{r}\}^2 \cdot \sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{s(x+u,y+v) - \overline{s(x,y)}\}^2}}$
Korrelationsfunktion von Moravec (MOR(x,y))	$\frac{2 \cdot \sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{r(u,v) - \bar{r}\} \cdot \{s(x+u,y+v) - \overline{s(x,y)}\}}{\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{r(u,v) - \bar{r}\}^2 + \sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{s(x+u,y+v) - \overline{s(x,y)}\}^2}}$
Normierte, lokal mittelwertfreie Summe der Differenzenquadrate (NSSDMF(x,y))	$\frac{\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{\{r(u,v) - \bar{r}\} - \{s(x+u,y+v) - \overline{s(x,y)}\}\}^2}{\sqrt{\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{r(u,v) - \bar{r}\}^2 \cdot \sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{s(x+u,y+v) - \overline{s(x,y)}\}^2}}$
Summe der Differenzenquadrate (SSD(x,y))	$\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{r(u,v) - s(x+u,y+v)\}^2$
Summe der Differenzenbeträge (SAD(x,y))	$\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} r(u,v) - s(x+u,y+v) $
lokal mittelwertfreie Summe der Differenzenquadrate (SSDMF(x,y))	$\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{\{r(u,v) - \bar{r}\} - \{s(x+u,y+v) - \overline{s(x,y)}\}\}^2$
lokal mittelwertfreie Summe der Differenzenbeträge (SADMF(x,y))	$\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{r(u,v) - \bar{r}\} - \{s(x+u,y+v) - \overline{s(x,y)}\} $
lokal mittelwertskalierte Summe der Differenzenquadrate (SSDMS(x,y))	$\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \left\{ r(u,v) - \frac{\bar{r}}{\overline{s(x,y)}} \cdot s(x+u,y+v) \right\}^2$
lokal mittelwertskalierte Summe der Differenzenbeträge (SADMS(x,y))	$\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \left r(u,v) - \frac{\bar{r}}{\overline{s(x,y)}} s(x+u,y+v) \right $

Tab.1: Definition der Korrelationskriterien, die keine Bild-Vorverarbeitung benötigen, also direkt auf den Grauwerten arbeiten.

Wurzel berechnet werden und die Addition im Nenner bringt besonders bei der Rechnung mit beschränkter Bitzahl Vorteile mit sich. Nenner und Zähler dieser Funktion sind ferner bei geeigneter Umformung sehr einfach über Lookup-Tabellen mit anschliessenden Akkumulatoren – ohne jeglichen Multiplikator – in Hardware zu realisieren. Dieses Ähnlichkeitsmass wird gelegentlich auch *pseudo-normierte Kreuzkorrelation* genannt. Es ist zwar invariant gegenüber additiven Helligkeitsänderungen, reagiert aber sensitiv auf multiplikative Intensitätsschwankungen. Moravec argumentiert, in realer Umgebung seien multiplikative Helligkeitsänderungen im Normalfall gering⁴. Diese Aussage stimmt nur teilweise, da z.B. Beleuchtungsschwankungen besser als multiplikative denn als additive Helligkeitsverschiebungen modelliert werden (siehe auch Kapitel 4: *Versuche und Versuchsbedingungen*). In kontrollierter Umgebung ist aber der multiplikative (wie auch der additive) Intensitätsunterschied aufeinanderfolgender, mit einer Videokamera aufgenommener Bilder in der Regel tatsächlich gering. Ähnliche Kontinuitätsaussagen lassen sich aber natürlich mit der gleichen Berechtigung für beinahe alle denkbaren Bildstörungen machen. Wie auch die Resultate der lokal mittelwertfreien Differenzmasse (ohne **NSSDMF**) zeigen, sind die Auswirkungen der erwähnten Sensitivität aber vergleichsweise klein.

Der Wertebereich der **KKF**, der **KKFMF** und des Korrelationsalgorithmus' von Moravec ist auf den Zahlenbereich zwischen +1 und -1 beschränkt. Da diese Algorithmen allesamt Ähnlichkeitsmasse darstellen, ist bei guter Übereinstimmung zwischen Referenz- und Korrelationsfenster mit Werten nahe +1 zu rechnen. Bei der **KKF** kommen negative Zahlen sogar nur vor, wenn das Such- oder Referenzfenster, z.B. als Folge einer vorherigen Hochpassfilterung, negative Pixelwerte enthält. Ein Vorteil der **KKFMF** gegenüber der **KKF** ist, dass der Dynamikbereich viel besser ausgenützt wird (siehe Anhang E.3: *Korrelationsfunktionen*). Bei der **KKF** können erfahrungsgemäss, selbst bei relativ schlechter Übereinstimmung zwischen Muster und Korrelationsfenster, Werte von über 0.9 auftreten. Die verfügbare Wertedynamik wird also bei der **KKF**, im Gegensatz zur **KKFMF**, nur schlecht ausgenützt.

Die **NSSDMF** wurde an der INRIA in Frankreich im Zusammenhang mit Stereosichtsystemen entwickelt ([52] und [75]). Da es sich dabei um die Verknüpfung des Normierungsterms der **KKFMF** mit der **SSDMF** (siehe unten) handelt, muss hier von einem Unähnlichkeitsmass gesprochen werden. Grosse Werte entsprechen also einer grossen Diskrepanz zwischen Korrelations- und Referenzfenster. Diese rechnerisch relativ aufwendig zu realisierende Mischung von zwei Algorithmen führt, wie wir später sehen

⁴ Zudem sei in vielen Fällen eine Robustheit in grossem Ausmass gar nicht gefragt.

werden, auch zu einer Mischung der Eigenschaften. Der Wertebereich dieses, wie auch jedes anderen besprochenen Unähnlichkeitsmasses, ist nur gegen unten beschränkt. Bei perfekter Übereinstimmung zwischen Muster und Korrelationsfenster wird ein Wert von Null ausgegeben.

Häufig wird in Applikationen die *Summe der Differenzenquadrate* (**SSD**: *sum of squared differences*, *Kreuzquadratsubtraktionsfunktion*, [6], [10], [11], [14], [16], [32], [48], [59], [67], [70], [91], [96], [100], [105], [116], [128], [134] und [144]) bzw. die *Summe der Differenzenbeträge* (**SAD**: *sum of absolute differences*, *Kreuzbetragsubtraktionsfunktion*, [3], [10], [11], [14], [16], [30], [41], [57], [59], [64], [67], [70], [76], [85]-[87], [89], [100], [107], [110], [112], [144], [146], [156], [159], [163], [164], [174], [176] und [177]) verwendet. Sie stellen heuristisch sehr einfach verständliche und ebenso einfach realisierbare Unähnlichkeitsmasse dar. Da sie schon bei geringen Helligkeitsänderungen im Suchfenster eine markante Leistungseinbuße erfahren, werden gelegentlich auch modifizierte Differenzmasse vorgeschlagen und eingesetzt. Zum einen sind dies die lokal mittelwertfreien (**SSDMF**, **SADMF**, [27]) und zum anderen die lokal mittelwertskalierten Varianten (**SSDMS**, **SADMS**, in der Literatur unbekannt⁵) der entsprechenden Korrelationskriterien. Sie bewirken eine erhebliche Verbesserung im oben genannten Fall. Die Resultate der lokal mittelwertfreien Algorithmen sind völlig unabhängig von additiven Helligkeitsänderungen⁶, diejenigen der lokal mittelwertskalierten Algorithmen sind invariant gegenüber

⁵ Die **SSDMS** kann analog zur **KKFMF** hergeleitet werden (Ryan bzw. Ryan und Hunt: [138], [139]; keine additive Helligkeitsverschiebung erlaubt!). Dabei wird angenommen, dass die multiplikative Helligkeitsänderung durch das Verhältnis der Helligkeitsmittelwerte ($\bar{r}/s(x, y)$) bestimmt werden kann (und nicht das Verhältnis der Standardabweichungen).

Wie Lücke in [104] zeigt, kann auch die **KKF** als quadratisches Differenzmass mit lokal skalierten Objektfensterintensitäten aufgefasst werden (siehe unten). Die in Kap. 5, *Resultate* dokumentierten kleinen Unterschiede der Robustheiten zugunsten der **KKF** zeigen, dass die Skalierung mit dem Mittelwertverhältnis (**SSDMS**) in der Regel eine relativ gute Annäherung an die entsprechende Skalierung bei der **KKF** darstellt.

$$\text{KKF}(x, y) = 1 - \frac{\overbrace{1}^{\text{konstant in } x \text{ und } y}}{2 \cdot \sum_{v=0}^{\text{MINDET}} \sum_{u=0}^{\text{MINDET}} r^2(u, v)} \cdot \sum_{v=0}^{\text{MINDET}} \sum_{u=0}^{\text{MINDET}} \left(r(u, v) - \sqrt{\frac{\sum_{v=0}^{\text{MINDET}} \sum_{u=0}^{\text{MINDET}} r^2(u, v)}{\sum_{v=0}^{\text{MINDET}} \sum_{u=0}^{\text{MINDET}} s^2(x+u, y+v)}} \cdot s(x+u, y+v) \right)^2$$

⁶ Die **SAD** und die **SSD** sind hingegen nur robust; d.h. die Reihenfolge der Korrelationswerte bleibt zwar unter idealen Verhältnissen, ohne Berücksichtigung von Quantisierungs- und Sättigungseffekten, erhalten, aber alle Korrelationswerte werden um denselben Betrag angehoben (d.h. nicht invariant).

multiplikativen Veränderungen. Ryan und Hunt zeigen in [139], dass die **SSDMF** den Maximum Likelihood Korrelator realisiert, falls die Pixel nur durch additive Helligkeitsänderungen sowie mittelwertfreies Gauss'sches Rauschen verändert werden (keine multiplikative Helligkeitsänderung!). Mittels Division durch die im Muster enthaltene Anzahl Pixel und durch die maximal mögliche Intensitätsdifferenz, bzw. deren Quadrat, könnten alle bisher definierten Unähnlichkeitsmasse auf den Bereich 0..1 skaliert werden. Im Zusammenhang mit der Überprüfung der Plausibilität einer berechneten Objektposition, anhand des entsprechenden Korrelationswertes, ist dies für Anwendungen eine interessante Eigenschaft (siehe z.B. [22], [78], [67] und [122]). Hier wird aber trotzdem die einfachere, für die betrachteten Versuche und Auswertungen indes äquivalente Form, ohne Skalierung benutzt. Damit ist der Wertebereich gegen oben durch die Anzahl Pixel im Muster multipliziert mit der maximalen Helligkeitsdifferenz (im Quadrat für die **SSD**, die **SSDMF** und die **SSDMS**) und gegen unten – bei maximaler Ähnlichkeit – durch den Wert Null begrenzt.

Da die **SAD** – insbesondere auf Hardwareplattformen mit im Vergleich zur Addition langsamer Multiplikation – schneller berechnet werden kann als die **SSD**⁷, die **SSD** aber unter gewissen Umständen vorteilhafter erscheint, sind auch Möglichkeiten untersucht worden die Leistungsfähigkeit der **SAD** durch geschickte, stückweise lineare Approximation an die **SSD** anzugleichen ([111]). Dieses Verfahren liefert aber nur in Ausnahmefällen bessere Resultate als die **SSD** oder die **SAD** selber. Darum werden im folgenden nur die reine **SSD** und **SAD** – sozusagen als Grenzfälle dieser Annäherung – untersucht.

Die Verwandtschaft unter den bislang vorgestellten Korrelationskriterien beruht auf unterschiedlichen Annahmen über den Bildinhalt bzw. dessen Veränderung und kann durch einfache Rechnung festgestellt werden. Die **KKF** entspricht z.B. der **KKFMF**, wenn gegenüber dem Muster keine additive Helligkeitsänderung im Suchfenster vorliegt ([139]). Falls die Energie in allen Korrelationsfenstern gleich gross ist, so kann die **KKF** durch die **SSD** oder gar die unnormierte Kreuzkorrelationsfunktion ersetzt werden ([96]). Die Position des Extremums wird durch einen solchen Ersatz, der je nach

⁷ Diese Aussage gilt nicht für alle Prozessortypen. Z.B ist die **SSD** auf gewissen modernen (general purpose) Signalprozessoren wegen der Möglichkeit der parallelen Ausführung von Multiplikation und Addition schneller als eine entsprechende Implementation der **SAD**. Die Absolutwert-Bildung und die Akkumulation kann nicht parallel ausgeführt werden, da diese Operationen zur gleichen Zeit die gleichen Prozessorressourcen beanspruchen würden.

Name	Definition	10
Korrelationsalgorithmus von Seitz (original; SES1(x,y))	$\sqrt{\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{r_{Sobel}(u,v) - s_{Sobel}(x+u, y+v)\}^2}$	
Korrelationsalgorithmus von Seitz (modifiziert; SEK1(x,y))	$\sqrt{\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{r_{Kirsch}(u,v) - s_{Kirsch}(x+u, y+v)\}^2}$	
Korrelationsalgorithmus von Seitz (modifiziert; SES2(x,y))	$\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} r_{Sobel}(u,v) - s_{Sobel}(x+u, y+v) $	
Korrelationsalgorithmus von Seitz (modifiziert; SEK2(x,y))	$\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} r_{Kirsch}(u,v) - s_{Kirsch}(x+u, y+v) $	

Tab.2: Definitionen der Algorithmen, die richtungscodierte Bilder benutzen. Die Differenzen sind als (kleinste) Winkelunterschiede zu verstehen.

Zielplattform die Implementierung vereinfacht, nicht beeinflusst⁸. Als gleichermassen äquivalent erweisen sich die **SSDMF**, die **NSSDMF**, die **KKFMF** und die Korrelation nach Moravec, falls die Signalvarianz (eigentlich: Mittlere quadratische Abweichung vom Mittelwert) in den Korrelationsfenstern konstant bleibt. Unter der Annahme, dass der lokale Helligkeitsmittelwert im Suchfenster nur additiv und für alle Korrelationsfenster identisch verändert wird, kann die **SSDMF** auch durch die **SSD** substituiert werden⁹. Diese Energie-, Varianzen- oder Mittelwertkonstanz ist aber in der Regel, wie auch die unterschiedlichen Ergebnisse zeigen werden – speziell für kleinere Muster oder für Bilder mit ausgeprägten Kanten – nicht erfüllt.

3.2. Korrelationskriterien mit Richtungsfiltern

Der von Seitz vorgeschlagene Algorithmus (**SES1**) eröffnet eine ganz neue Klasse von interessanten Korrelationskriterien (Tab.2). Da mit richtungscodierten Bildern gearbeitet wird, ist insbesondere eine gute Robustheit gegenüber Helligkeitsschwankungen in der Szene zu erwarten. Das in [145]

⁸ Die Korrelationswerte sind aber nicht notwendigerweise identisch!

⁹ Dasselbe gilt auch für die **SAD** und die **SADMf**. Falls zusätzlich gilt, dass die mittlere Helligkeit im Muster der mittleren Helligkeit in allen Korrelationsfenstern entspricht, so kann die **SSDMF** auch durch die **SSD** ersetzt werden (gilt auch für die **SADMf** und die **SAD**).

¹⁰ Um die Versuche zu vereinfachen, wurde bei der Implementation der **SES1** und der **SEK1** die Wurzel in den Formeln weggelassen, da sie die Rangierung der Werte und damit den Ort des Minimums des Unähnlichkeitsmasses nicht verschiebt.

definierte Verfahren, das den Sobeloperator zur Richtungscodierung zu Hilfe nimmt, unterscheidet zwischen 32 verschiedenen Richtungen (siehe Anhang A.1: *Nomenklatur*). Die Richtungsauflösung des als Alternative zum Sobeloperator eingeführten Kirschoperators (**SEK1**, **SEK2**) ist mit nur acht unterscheidbaren Richtungen recht gering. Dies wirkt sich besonders bei kleinen Mustern negativ aus¹¹. Filter mit grösseren Masken wurden nicht untersucht, obwohl sie einerseits eine bessere Richtungsauflösung ergäben und andererseits robuster gegenüber gewissen Bildstörungen sein sollten. Der Nachteil von grösseren Masken besteht darin, dass die Filterresultate einen weniger lokalen Charakter haben und deswegen, je nach Bildinhalt, Information verloren gehen kann.

Der Gebrauch der (Winkel-) **SAD** anstelle der (Winkel-) **SSD** in Seitz' Original-Algorithmus liefert, wie meine Messungen zeigen, in vielen Fällen etwas bessere Resultate. -Im Unterschied zur normalen **SSD** und **SAD** sind bei der *Winkel-SSD* und bei der *Winkel-SAD* die Differenzen als die kleinsten, absoluten Richtungsunterschiede zwischen den Richtungen der Muster- und der entsprechenden Korrelationsfensterpixel zu verstehen.

3.3. Korrelationskriterien mit Bild-Hochpassfilterung

Aus den zum Teil recht weit verbreiteten Algorithmen, die auf Hochpassgefilterten Daten arbeiten, wurden nur zwei Verfahren ausgewählt (Tab.3): Die *unnormierte Kreuzkorrelationsfunktion (direct correlation, UKKF*, [11], [33], [50], [67], [68], [118], [137], [144], [156] und [164]) und die *lokal mittelwertfreie, unnormierte Kreuzkorrelationsfunktion (Kovarianzfunktion, UKKFMF*, [147] und [141]). Beide Algorithmen erbringen ungenügende Leistungen, wenn sie direkt auf den Helligkeitsdaten arbeiten. Durch vorherige Hochpassfilterung der Bildinformation kann dieses Problem aber weitgehend behoben werden. Als vorverarbeitendes Hochpassfilter wurde der mittelwertfreie Laplaceoperator ausgewählt (siehe Anhang A.1: *Nomenklatur*). Gründe für diese Wahl waren neben häufigem Einsatz in diversen Applikationen auch die simple Implementation und die gute Leistung im Vergleich mit anderen, ähnlich einfachen Filtern bei den Vorversuchen (Sobel-, Robertsoperator etc.). Die Differenzierung, die jedes Hochpassfilter beinhaltet, lässt speziell bei verrauschten Bildern ein relativ schlechtes Verhalten erwarten. Die Versuchsergebnisse sind aber überraschenderweise auch bei Rauschstörungen recht gut.

¹¹ Die Auflösung der Sobel-Richtungsinformation ist im Gegensatz zu den Resultaten des Kirschoperators von der lokalen Helligkeit abhängig, da das benötigte Gradientenverhältnis bei geringer Intensität eine schlechte Ausnutzung der Werteskala ergibt.

Name	Definition
unnormierte Kreuzkorrelationsfunktion (UKKF(x,y))	$\sum_{v=0}^{w-1} \sum_{u=0}^{w-1} r'(u,v) \cdot s'(x+u, y+v)$
lokal mittelwertfreie, unnormierte Kreuzkorrelationsfunktion (UKKFMF(x,y))	$\sum_{v=0}^{w-1} \sum_{u=0}^{w-1} \{r'(u,v) - \bar{r}'\} \cdot \{s'(x+u, y+v) - \bar{s}'(x,y)\}$

Tab.3: Definitionen der Korrelationskriterien, die Hochpass-gefilterte Bilder benutzen (ohne Binarisierung).

Natürlich könnte bei jedem anderen (Un-) Ähnlichkeitsmass ebenfalls eine Vorverarbeitung eingeführt werden. Eine vorgängige Hochpassfilterung kann z.B. die Detektierbarkeit einer Korrelationsspitze verbessern, da damit die Aufmerksamkeit (bzw. das Gewicht) weg vom eher einförmigen Hintergrund auf die stark informationshaltigen Bildteile gelenkt wird (siehe Anhang E.3: *Resultate, Korrelationsfunktionen* sowie [6], [8], [9], [21], [33], [65], [95], [107], [113], [114], [122], [133], [134], [137], [153], [156], [171] und [172]). Lapidus wendet z.B. vor der KKF einen Sobeloperator auf die Bilder an ([95]). Dieses Hochpassfilter soll den Einfluss von Reflexionen und Schattenwurf reduzieren. Problematisch wird die Hochpassfilterung aber bei Bildstörungen, die sich im hochfrequenten Bildanteil auswirken. Wie wir später sehen werden, ist dieses Vorgehen deshalb für kleine Muster und starke Bildstörungen in der Regel mit geringeren Detektionsraten verbunden.

3.4. Binäre Korrelationskriterien

Nishihara schlägt die Verwendung der unnormierten ein-Bit-Korrelation vor (Tab.4), wobei er das Bild und das Muster vorher mit einem Laplace-Gauss-Filter vorverarbeitet (siehe Anhang A.1: *Nomenklatur*) und nachher anhand des Resultatvorzeichens binarisiert ([124], Anhang D: *Wirkung des Laplace-Gauss-Operators*; die binarisierten Filterresultate werden in der verwendeten Implementation als +1 respektive -1 interpretiert!). Diese Vorverarbeitung kann sehr rechenintensiv sein, da grosse Filterkerne eingesetzt werden sollten¹². Für kleine w darf der Filterkern kleiner gewählt werden, da die Filterkoeffizienten ausserhalb des Operator kernels einen bestimmten Minimalwert nicht mehr übersteigen und somit nicht mehr berücksichtigt werden müssen. Wenn man die recht aufwendige Filterung ausser Acht lässt, ist

¹² Filterkerngrösse bei $w=8$: 29x29 Pixel. Der Parameter w bestimmt die Filterkennlinie.

Name	Definition
Korrelationsalgorithmus nach Nishihara ($w=2, 3, \dots, 8$; NIS(x,y))	$\sum_{v=0}^{w/2-1} \sum_{u=0}^{w/2-1} r_{b1}(u, v) \cdot s_{b1}(x+u, y+v)$
Korrelationsalgorithmus nach Nack (original; NA1(x,y))	$1 + \frac{\sum_{v=0}^{w/2-1} \sum_{u=0}^{w/2-1} r_{b2}(u, v) \cdot s_{b2}(x+u, y+v)}{1 + \sum_{v=0}^{w/2-1} \sum_{u=0}^{w/2-1} s_{b2}(x+u, y+v)}$
Korrelationsalgorithmus nach Nack (modifiziert; NA2(x,y))	$\frac{1 + \sum_{v=0}^{w/2-1} \sum_{u=0}^{w/2-1} r_{a1}(u, v) \cdot s_{a1}(x+u, y+v)}{\left\{ 1 + \sum_{v=0}^{w/2-1} \sum_{u=0}^{w/2-1} s_{a1}(x+u, y+v) \right\} \cdot \left\{ 1 + \sum_{v=0}^{w/2-1} \sum_{u=0}^{w/2-1} r_{a2}(u, v) - \sum_{v=0}^{w/2-1} \sum_{u=0}^{w/2-1} s_{a2}(u, v) \cdot s_{a2}(x+u, y+v) \right\}}$

Tab.4: Definitionen der Algorithmen, die binäre, mit Filtern vorverarbeitete Bilder verwenden.

der Aufwand (Binäroperationen!) in Relation zur Leistungsfähigkeit dieser Methode beachtlich.

Auch Nack schlägt einen binäres Korrelationskriterium vor ([122] und [156]). Das originale Verfahren (**NA1**) ist derart optimiert, dass diejenige Position resultiert, deren Objektfenster die minimale Anzahl fälschlich auf eins gesetzter Pixel enthält. Dies ist nicht befriedigend und führt zu messbaren Fehlern, da die Korrelation im Fall einer einzigen Übereinstimmung von 'Eins-Pixeln' grösser ist als die Korrelation bei totaler Übereinstimmung der 'Eins-Pixel' (alle Pixel, die im Referenzfenster auf eins gesetzt sind, haben ein entsprechendes Pendant im Objektfenster) und einer einzigen Diskrepanz. Die hier vorgeschlagene Modifikation umgeht dieses Problem, indem zusätzlich auf die maximale Übereinstimmung der gesetzten Pixel optimiert wird.

Die **NA1**-Korrelation wird im Prinzip fälschlicherweise als Original-Algorithmus von Nack bezeichnet. Die Addition einer Konstanten (Eins) im Zähler wie auch im Nenner wurde ursprünglich nicht verwendet. Die Änderung des Nenners wurde hinzugefügt, um degenerierte, in den Untersuchungen aber auftretende Fälle, in denen dieser Term Null wird, beherrschen zu können. Die Modifikation des Zählers legt die Ergebniswerte wieder in den Bereich zwischen Null und Eins.

Da diese kleinen Modifikationen in Anwendungen die Leistung sowie den Rechenaufwand nicht wesentlich (negativ) beeinflussen, dafür aber das separate Behandeln von Spezialfällen ersparen, wird die **NA1** trotz dieses Vorbehalts weiterhin als Original-Algorithmus bezeichnet.

Beim originalen wie auch beim modifizierten Korrelations-Algorithmus von Nack erfolgt die Binarisierung nach einer Hochpassfilterung, die mit

einem dem Robertsoperator ähnlichen Filter durchgeführt wird (siehe Anhang A.1: *Nomenklatur*). Die zugehörige adaptive Binarisierungsschwelle wird derart angepasst, dass eine, im Rahmen der Quantisierung, mehr oder weniger konstante (wählbare) Anzahl von 'gesetzten' Kantenpixeln im Suchfenster resultiert (15%).

4. Versuche und Versuchsbedingungen

4.1. Allgemeines

Damit die Versuchsergebnisse richtig gewertet werden können, ist es notwendig die Testbedingungen – Bildmaterial, Versuchsablauf, Auswerteprozess und Rechengenauigkeit – präzise zu definieren. Die Bilder wurden derart zusammengestellt, dass sie einerseits eine gute, repräsentative Auswahl von Mustern bieten und andererseits mit einer besonderen geometrischen Verteilung potentieller Muster den Versuchsablauf vereinfachen (siehe Bildstörung **ZOOM** und **FOCUS**).

Die Tests wurden auf drei unabhängigen, aber identisch gewonnenen Sets von Bildsequenzen durchgeführt (Fig.1). Die Bildsequenzen eines Sets enthalten Bilder der gleichen Szene mit sequenz-spezifischen und von Bild zu Bild zunehmenden Bildstörungen (siehe Anhang B: *Bildmaterial*). Dies erlaubt die getrennte Beobachtung der Auswirkungen verschiedener Bildstörungsarten.

Fünf Typen von Bildstörungen wurden bei den Tests berücksichtigt. Sie entsprechen den in der Regel wichtigsten, real auftretenden und zumeist unerwünschten Bildveränderungen. Die Namen der jeweiligen Bildsequenzen

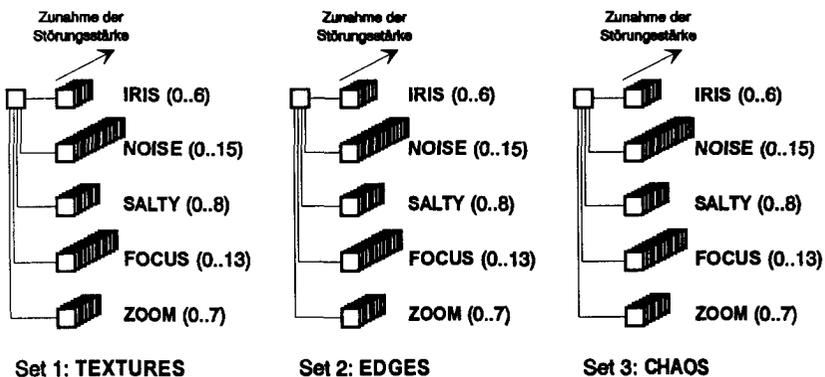


Fig.1: Bilder- 'Stammbaum' für die Bildreihen aller Sequenzsets. Alle Bildsequenzen eines einzelnen Sets haben das gleiche Anfangsbild (siehe auch Anhang B: *Bildmaterial*).

bzw. die im weiteren verwendeten Kürzel der entsprechenden Transformationen sind wie folgt definiert:

- Helligkeitsreduktion: **IRIS**
- Additives, mittelwertfreies Gauss'sches Rauschen: **NOISE**
- 'Salt and Pepper' Rauschen: **SALTY**
- Unschärfe: **FOCUS**
- Grössenänderung: **ZOOM**

Für jedes Bildsequenzenset existiert genau ein einziges ungestörtes Bild, das als Startbild Teil jeder Sequenz ist. Die verwendeten Bilder enthalten 256x256 Pixel mit jeweils sechs Bit Helligkeitsauflösung, d.h. 64 unterscheidbaren Graustufen¹³. Im Anfangsbild jedes Sequenzensets wurden manuell 18 verschiedene (wahre) Objektpositionen ausgesucht; an diesen Orten wurde die Musterinformation gewonnen (→ total 54 Objektpositionen: 3 mal 18 Positionen bei 3 Sequenzensets; Anhang C: *Objekt- und Suchfensterpositionen*).

Um nicht nur optimal gute, d.h. einfach auffindbare Muster zu untersuchen, wurde in Set 1 und Set 2 (**TEXTURES** und **EDGES**) eine mehr oder weniger zufällige (aber identische) Wahl der wahren Objektpositionen getroffen. Bei Set 1, deren Bilder hauptsächlich texturhafte Informationen enthalten, ist diese Wahl unkritisch. Die Texturen bedecken die gesamte Bildfläche und stellen somit, wie das für Texturen genügend feiner 'Körnung' üblich ist, an jeder Stelle im Bild eindeutige Informationen zur Verfügung. Etwas schwieriger ist die Situation bei Set 2, deren Bilder hauptsächlich durch mehr oder weniger einförmige Flächen und markante Kanten bestimmt sind. Die für die Korrelationen wichtige Kanteninformation kommt teilweise erst bei mittelgrossen Mustern zum Tragen. Bei Verfahren, die sich beinahe ausschliesslich auf die Kanteninformation konzentrieren, wie z.B. die **UKKF**, und die Restinformation der relativ einförmigen Flächen weitgehend unberücksichtigt lassen, ist darum teilweise mit einer gewissen Leistungseinbusse für kleine Muster zu rechnen (siehe Anhang E.2: *Tabellen*). Das gleiche gilt auch für die binären Korrelationen (**NA1**, **NA2** und **NIS**), deren Vorverarbeitungsfiler nur an speziell geeigneten Orten überhaupt eine Struktur im Muster detektieren können. Zur Reduktion dieses Effekts wurde die zufäl-

¹³ Der Einfluss der Helligkeitsquantisierung wird z.B. in [38], [68] und [92] angesprochen. Es ist aber klar, dass in den meisten Fällen (bei kontrastreichen Bildern, die die Helligkeitsdynamik gut ausnützen) eine relativ grobe Helligkeitsquantisierung vollausreicht genügt.

lige Wahl der wahren Objektpositionen für die beiden Sequenzensets in einzelnen Fällen anhand der Szene von Set 2 geringfügig manuell korrigiert.

Bei Sequenzenset 3, wird ein Ausschnitt eines Labors gezeigt, der am ehesten mit einer realen Szene z.B. aus dem Robotikbereich vergleichbar ist. Hier sind die wahren Objektpositionen, wie meist in einer Anwendung auch, sorgfältig ausgesucht worden. Die manuelle Wahl wurde hauptsächlich durch den lokalen Kontrast beeinflusst. Auch kleine Muster sollten hier genügend Bildinformation enthalten.

Es ist zu beachten, dass in keinem der drei Sequenzensets eigentliche, räumlich z.B. durch deutliche Intensitätsgrenzen abgegrenzte Objekte als Muster verwendet werden, wie das in einigen Applikationen üblich ist. Vielmehr ist, besonders bei Set 1 und 2, eine bestimmte Bildstruktur (z.B. Texturtyp) im ganzen Suchfenster zu finden und ein kleiner Ausschnitt daraus, mit gleichen Eigenschaften, wird effektiv gesucht.

Um die Zahl der geometrisch transformierten Bilder klein zu halten (**ZOOM**), wurde die Wahl der Objektpositionen bei allen Sequenzensets zusätzlich dadurch eingeschränkt, dass in jedem neuntel der Bildfläche (Objektpositionsfläche, siehe Fig.2) genau zwei wahre Objektpositionen bestimmt werden mussten. Der Grund hierfür liegt im Wunsch, für alle Muster gleiche Bedingungen zu schaffen. Die zu jedem Muster zugehörigen Suchfenster sollten darum zentrisch bezüglich jeder einzelnen Objektposition expandiert werden. Dies garantiert auch, dass die Sollpositionen, in den Objektschwerpunkten angenommen, durch diese Bildtransformation nicht verschoben werden. Pro Sequenzenset wären also wegen der 18 verschiedenen Expansionszentren (Objektpositionen) ebenso viele 'ZOOM-Sequenzen' notwendig. Dies ergäbe eine Unmenge von Bildern, die zum grössten

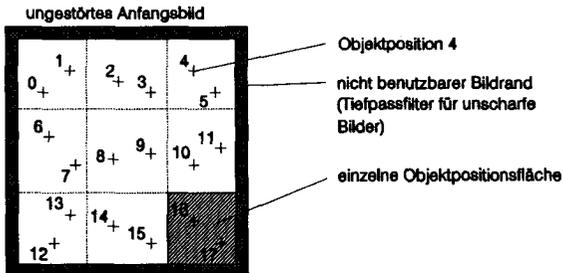


Fig.2: Wahl der wahren Objektpositionen. In jedem neuntel der Bildfläche (Objektpositionsfläche) werden zwei Positionen ausgesucht. Auch bei maximaler Mustergrösse muss das wahre Objektfenster vollständig innerhalb der nicht doppelt schraffierten Fläche Platz finden.

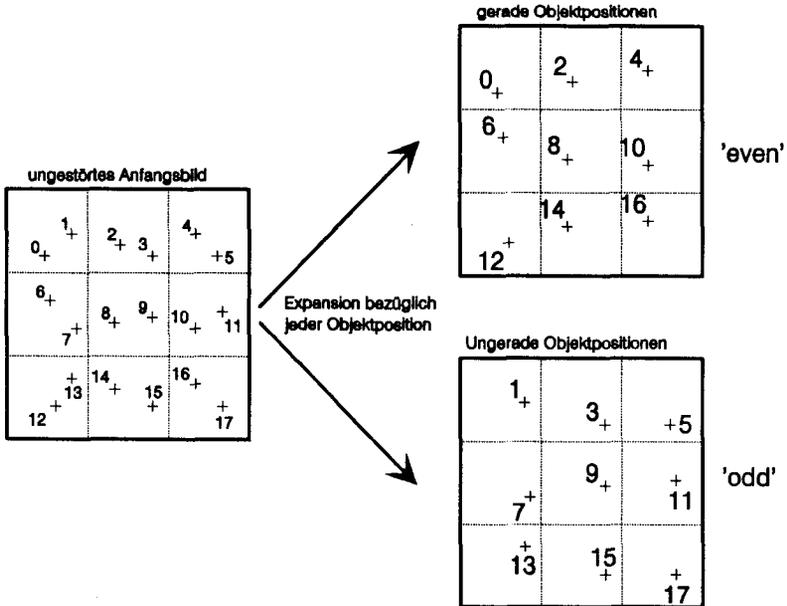


Fig.3: Zusammensetzen der einzelnen, bezüglich der wahren Objektpositionen expandierten Bilder der 'Doppel'-Sequenzen **ZOOM**. Die Objektpositionen (Kreuze) bleiben unabhängig vom Expansionsgrad fix.

Teil unnütze Information enthielten, da die Suchfenster eine maximale Grösse von nur 64x64 Pixel (rund um die Objektposition) erreichen und der Rest der Pixel, ausserhalb der Suchfenster, nie benutzt würde. Es liegt also auf der Hand aus den oben erwähnten 18 Sequenzen eine einzige 'zusammenzuschneiden'. Aus geometrischen Gründen ist dies aber nicht möglich. Als Kompromiss wurden deshalb schliesslich aus den 18 potentiellen Bildsequenzen zwei Sequenzen, 'even' und 'odd', generiert, die jeweils die Information für die Hälfte aller Objektpositionen enthalten. Die unterschiedlich expandierten Suchfenster gleicher Vergrösserung der gerade nummerierten Objektpositionen wurden somit zur Sequenz '**ZOOMeven**' und die Suchfenster der entsprechenden ungerade nummerierten Objektpositionen zur Sequenz '**ZOOModd**' zusammengefügt (Fig.3). Obwohl also jeweils zwei Sequenzen **ZOOM** pro Sequenzenset existieren, wird diese Tatsache im weiteren Verlauf dieser Darlegungen nicht mehr erwähnt. Wenn von der Sequenz **ZOOM** gesprochen wird, sind also implizit immer beide Ausprägungen dieser Sequenz gemeint.

Eine zusätzliche Einschränkung bei der Wahl der Objektposition wird durch die Sequenz **FOCUS** (unscharfe Bilder) erzwungen. Wegen der nicht verschwindenden Operatorgrösse des Tiefpassfilters (maximal 31x31 Pixel) ist der Randbereich der unscharfen Bilder nicht definiert. Die wahren Objektpositionen müssen also derart gewählt werden, dass auch die grössten wahren Objektfenster das Bild-Randgebiet sowie die Grenzen der Objektpositionsflächen nicht schneiden (Fig.2).

Die untersuchten Mustergrössen liegen im Bereich zwischen 2x2 und 25x25 Pixeln. Daraus resultieren alles in allem 1296 verschiedene Muster (24 Mustergrössen an je 18 Positionen in den 3 Sequenzensets). Für jedes (Un-) Ähnlichkeitsmass wurden also insgesamt ca. 60'000 Korrelationsfunktionen ausgewertet (alle 1296 Muster in allen 50 Bildern der entsprechenden Sequenzensets; bei **ZOOM**-Störung nur die Muster mit ungeraden Kantenlängen, also 648 Muster!).

Die Musterdaten wurden während den Versuchen mit einem bestimmten Sequenzenset unverändert belassen. Das Muster (maximale Bildqualität) wurde also mit Korrelationsfenstern gleicher oder geringerer Qualität verglichen. Natürlich wären auch Versuche interessant in denen sowohl das Suchfenster als auch das Muster gestört würden. Um alle Kombinationen durchzutesten müsste aber die Anzahl der Versuche quadriert werden und die Versuchszeit nähme dementsprechend unakzeptable Werte an.

Die Suchfenstergrösse wurde über die Korrelationsfläche definiert, die in jedem Fall 40x40 potentielle Objektpositionen enthielt. Das Suchfenster veränderte also seine Grösse in Abhängigkeit der Mustergrösse ($xls = xlr + 39$, xls , xlr : Kantenlänge des Suchfensters bzw. Musters).

Die Suchfensterposition wurde leicht asymmetrisch bezüglich der Objektposition gewählt (siehe Anhang C: *Objekt- und Suchfensterpositionen*). Diese Massnahme wurde getroffen, um auch Versuche mit Verfahren, die auf Frequenztransformationen beruhen, und damit ein 'Windowing' der Daten nötig machen, unter den selben Voraussetzungen durchführen zu können (phase only matched filter, Cepstrum-Methode, siehe Kapitel 7: *Diskussion von Beiträgen aus der Literatur*); sie verhindert nämlich, dass diese neue Klasse von Algorithmen durch das Windowing nicht bevorteilt werden. Es wird also vorgesorgt, dass tatsächlich nach der Musterinformation und nicht nach der eigentlich uninteressanten Windowing-Funktion gesucht wird! Obwohl Versuche mit einigen der Algorithmen aus dieser Klasse gemacht wurden, sollen sie hier nicht weiter erwähnt werden. Es zeigt sich aber durch dieses Beispiel, dass neben den hier präsentierten Ergebnissen über die Leistungsfähigkeit einer Anzahl von Korrelationsverfahren auch ein Versuchs-

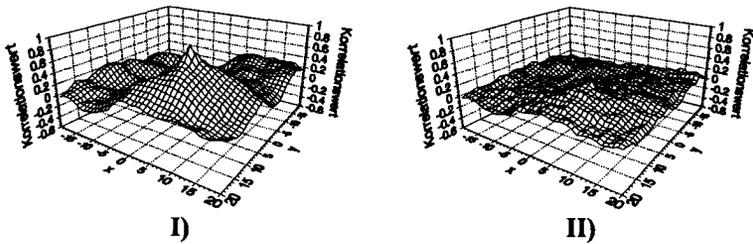


Fig.4; Beispiel einer Kreuzkorrelationsfunktion: **KKFMF** im Fall I) bei idealen Bildverhältnissen und im Fall II) bei maximal verrauschtem Bild (Set 1: **TEXTURES**, Sequenz **NOISE**, Muster 5).

ablauf vorgestellt wird, der ebenso auf andere 'Korrelations'-Algorithmen anwendbar ist.

Mit Hilfe der Korrelationskriterien wurden für alle Bilder die Positionen der Objektfenster innerhalb der jeweiligen Suchfenster bestimmt. Zur Lokalisierung der entsprechenden Korrelationsextremwerte wurde ein einfacher Maximum- bzw. Minimum-Detektor auf die Ausschnitte der Korrelationsfunktionen angewendet (Fig.4, siehe auch Anhang E.3: *Resultate, Korrelationsfunktionen*). Die so berechneten Positionen wurden anschliessend mit den wahren Objektpositionen verglichen. Einzig diese Abweichungen wurden, abhängig von den verschiedenen Testparametern wie Bildsequenz, Bildnummer, Mustergrösse und Musternummer, weiter ausgewertet. In der Literatur wird die Leistungsfähigkeit von Korrelationsalgorithmen häufig mit Hilfe des Signal-zu-Rauschleistungsverhältnisses in der Korrelationsebene beschrieben. Dies ist aber aus zweierlei Gründen problematisch. Erstens ist der Signal- bzw. Rauschanteil der Korrelationswerte in der Praxis schwierig zu bestimmen und zweitens zeigt ein Blick auf die Figuren im Anhang E.3: *Resultate, Korrelationsfunktionen*, dass z.B. die **NA2** im betrachteten, mittelmässig stark gestörten Fall – beispielsweise im Vergleich zur **KKFMF** – ein hervorragendes Signal-zu-Rauschleistungsverhältnis aufweisen muss, die Zahl erfolgreicher (tolerierbarer) Positionierungen aber in der Regel erheblich geringer ausfällt als beim Vergleichsverfahren.

Alle Berechnungen wurden auf einem IBM-kompatiblen Personal Computer durchgeführt. Die Programmierung erfolgte in TurboPascal 6.0, wobei auf schnelle Ausführungszeit optimiert wurde. Als Zahlenformat wurde der *DOUBLE* verwendet, ein 64 Bit Fließkommaformat mit 11 Bit Exponent und 53 Bit Mantisse. Dies sollte garantieren, dass ausser minimen, unau-

weichlichen Genauigkeitsverlusten bei Divisionen und Wurzelberechnungen keine Rechenfehler durch ungenügende Auflösung des Zahlenformates eingeführt wurden. Bei Implementationen – speziell auf Fixkommarechenwerken – ist gerade dieser Aspekt ebenfalls von grosser Bedeutung. In [138] werden von Ryan einige dieser Probleme zur Sprache gebracht.

4.2. Generierung der Bildsequenzen

Neben den eigentlichen Versuchen gestaltete sich die Generierung der Bilder am aufwendigsten. Um reproduzierbare Resultate von kontrolliert veränderten Bildern zu erhalten, ist eine sorgfältige Planung mit Vorausblick auf die Tests unerlässlich. Die geometrischen Randbedingungen der Objektpositionen wie auch der Dynamikbereich der Störungen müssen frühzeitig festgelegt werden, um die aufwendigen Versuche schliesslich ungestört durchführen zu können.

Erste Versuche, alle Bildstörungen experimentell zu erzeugen, zeigten, dass dadurch zwar realistische Bilder gewonnen werden könnten, die Kontrolle über die Störungen aber sehr schwierig ist (Ausmass der Störungen sowie Nebeneffekte). Darum wurde schliesslich entschieden, die Mehrzahl der Bilder rechnerisch aus Bildern realer Szenen zu generieren.

4.2.1. Helligkeitsreduktion, IRIS

Die Sequenzen **IRIS** wurden als einzige experimentell gewonnen. Dabei wurde die automatische Verstärkungsregelung (AGC), die Aperturkorrektur sowie die elektronische Blende (electronic shutter) der K235 schwarz/weiss-Kamera von Siemens ausgeschaltet und das Videosignal mit dem elektrischen Energienetz synchronisiert¹⁴. Das erste Bild der Sequenz wurde jeweils derart festgehalten, dass es sowohl sehr dunkle als auch sehr helle Pixel enthielt (siehe Histogramme in Fig.5). Die mechanische Blende des verwendeten Objektivs wurde von Bildaufnahme zu Bildaufnahme um eine Stufe geschlossen. Dies entspricht einer theoretischen, optischen Leistungsreduktion um einen Faktor zwei zwischen aufeinanderfolgenden Bildern (Fig.6 für experimentellen Leistungs- bzw. Energieverlauf). Das letzte Bild der Sequenz besitzt damit nur noch einen äusserst schwachen Kontrast (siehe Anhang B: *Bildmaterial*).

¹⁴ Um den Einfluss der mit doppelter Netzfrequenz schwankenden Beleuchtungsstärke zu minimieren.

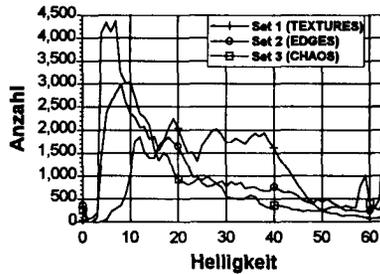


Fig.5: Histogramme der unveränderten Anfangsbilder der einzelnen Sequenzensets (Set 1: **TEXTURES**, Set 2: **EDGES**, Set 3: **CHAOS**). Die Kurven zeigen, dass die Helligkeitsdynamik, trotz gewissen Sättigungseffekten, gut ausgenutzt wird.

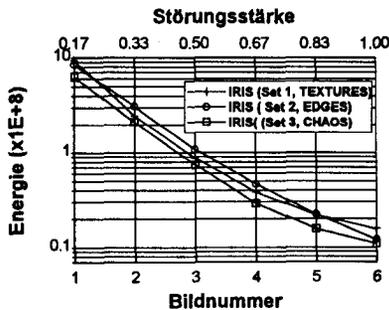


Fig.6: Energieabfall bei den Sequenzen **IRIS** für alle drei Sequenzensets.

Sinnvollerweise wird die Änderung der Blendenöffnung hauptsächlich als multiplikative, d.h. von der ungestörten Helligkeit der einzelnen Pixel abhängige Modulation der Intensität modelliert (Gl.1). Unter der Annahme eines idealen optischen Systems, das der Strahlenoptik gehorcht, und konstanter Szenenbeleuchtung verhält sich die Leistung, die auf ein Sensorpixel fällt, proportional zur Fläche der Blendenöffnung. Die gemessene Intensität ändert sich also proportional zum Quadrat des Blendendurchmessers. -Ähnlich kann auch bei einer Beleuchtungsänderung argumentiert werden. Es ergibt sich also ebenfalls ein multiplikativer Zusammenhang zwischen der Beleuchtungsleistung und dem gemessenen Sensorsignal. Dieser Effekt entspricht einer gleichzeitigen Kontrast- (Intensitätsdynamik) und Helligkeitsreduktion (mittlere Helligkeit).

Modell der IRIS-Helligkeitstransformation (ohne Quantisierungs- und Sättigungseffekte):

$$b_{\text{IRIS}}(x, y, Nr) = b(x, y, 0) \cdot m(Nr) + n_{\text{Kamera}}(x, y, Nr) \quad (1)$$

$$0 \leq m(Nr + 1) < m(Nr) \leq 1 \quad \text{für alle } Nr > 0$$

Die Intensität eines jeden Pixels $b(x, y, \theta)$ wird proportional zum Term $m(Nr)$ reduziert. $n_{\text{Kamera}}(x, y, Nr)$ bezeichnet das Kamera- und Quantisierungsrauschen, das bei der experimentellen Erfassung der Bilder nicht vollständig zu vermeiden ist. Der ganzzahlige Parameter Nr beschreibt die Nummer des Bildes innerhalb der Sequenz. Mit zunehmender Bildnummer Nr nimmt $m(Nr)$ ab.

Wegen der experimentellen Aufnahme dieser Bilder mussten natürlich spezielle Vorkehrungen getroffen werden, um geometrische Verschiebungen oder Verzerrungen zwischen den Aufnahmen klein zu halten. Dies wurde durch einen stabilen mechanischen Aufbau der Szene und der Kamera erreicht.

4.2.1.1. Signal-zu-Störleistungsverhältnis kontra Signal-zu-Störvarianzenverhältnis

Der Verlauf des Signal-zu-Störleistungsverhältnisses über die ganze Bildsequenz gibt einen Qualitätseindruck der Bilder, also die Relation zwischen der Bildnummer und der messbaren Bildqualität ([72])¹⁵. Dieser Zusammenhang wurde unter der Annahme eines völlig unverrauschten Startbildes bestimmt (Fig.7.I). Dies ist zwar nicht ganz korrekt, da selbst bei idealsten Aufnahmebedingungen ohne Sensor- und Elektronikrauschen zumindest (gleichverteiltes) Quantisierungsrauschen auftreten sollte; die Helligkeitsreduktion, ist aber auf jeden Fall stark dominant gegenüber diesem Effekt. Das eigentliche 'Störsignal' ergibt sich also im wesentlichen durch die veränderte Helligkeit im Bild.

Erheblich interessanter als das Leistungsverhältnis ist eigentlich das Verhältnis von Signal- zu Störsignalvarianz (Fig.7.II). Im Grunde ist diese Zahl für die Detektionsrate eines Algorithmus' bestimmend ([128]). Dies ist sehr einfach einzusehen. Die Signalleistung in einem Bildausschnitt kann sehr gross, die Signalvarianz dagegen gleichzeitig äusserst klein sein. Ein derartiges Muster ist aber schwierig zu lokalisieren, da unstrukturierte Flächen –

¹⁵ Nach Plunk ([132]) ist für eine 90-prozentige Sicherheit beim Auffinden des Objekts ein Signal-zu-Rauschleistungsverhältnis von 4 erforderlich (KKF?), was in etwa mit den weiter unten präsentierten Ergebnissen übereinstimmt.

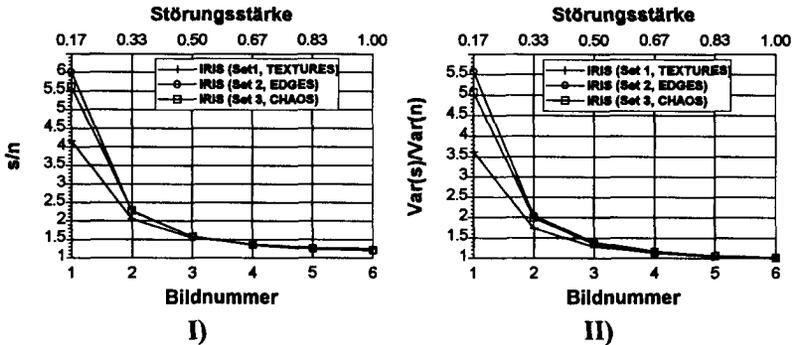


Fig. 7: Signal-zu-Störleistungsverhältnis (I) bzw. Signal-zu-Störvarianzenverhältnis (II) in Abhängigkeit der Bildnummer für die Sequenzen IRIS in allen Sequenzensets.

mit ihrer kleinen Helligkeitsvarianz – keine Detektionsmerkmale bietet. Ein Mustersignal mit kleiner Varianz lässt sich wegen fehlender positionsspezifischer Information viel schlechter auffinden als eines mit grosser Signalvarianz. In umgekehrter Weise kann für das Störsignal argumentiert werden. Eine hohe Störenergie mit kleiner Störvarianz bietet in der Regel nur kleine Probleme. Durch das Hinzufügen von hochfrequenter Störinformation (hohe Störvarianz) wird die Detektionsrate aber bei gleicher Störenergie viel stärker beeinflusst.

Diese Argumentation gilt aber nicht unbedingt bei allen Bildstörungstypen und für alle Korrelationsverfahren. So sind z.B. die Differenzmasse **SSD** oder **SAD**, wie weiter unten dokumentiert, sehr robust gegenüber Rauschen (grosse σ_n), sie versagen aber schon bei sehr kleinen Helligkeitsverschiebungen (kleine Störsignalvarianzen). Die Leistungs- oder Varianzenverhältnisse alleine, ohne Kenntnis des Bildstörungstyps (und natürlich des gewählten Algorithmus) erlauben darum ebenfalls keine zuverlässige Prognose über die Leistungsfähigkeit der Korrelationskriterien. Deswegen werden sowohl die Leistungs- als auch die Varianzenverhältnisse dargestellt. Diese Grafiken verbinden die abstrakten Bildnummern als Bildqualitätsmass mit der physikalischen Welt. Korrekterweise müssten allerdings die Leistungs- bzw. Varianzenverhältnisse der einzelnen Suchfenster und nicht der ganzen Bilder dokumentiert werden. Um den Platzbedarf für diese Charakterisierung des Bildmaterials klein zu halten und weil die Bildinformation ausserhalb der Suchfenster ja prinzipiell ähnlich strukturiert ist wie innerhalb, scheint das gewählte Vorgehen trotzdem dem Problem angepasst zu sein.

4.2.2. Additives, Gauss'sches Rauschen, NOISE

Zur rechnerischen Generierung der weiteren Bildsequenzen wurde jeweils das erste ($Nr=0$) und damit qualitativ beste Bild der Sequenz **IRIS** als Vorlage benutzt. Die Anfangsbilder aller Sequenzen eines Sets sind also identisch (Fig.1).

Die Sequenz **NOISE** wurde durch Überlagerung der ungestörten Helligkeitsinformation mit von Bild zu Bild zunehmendem, additivem und mittelwertfreiem Gauss'schem Rauschen $n_{Gauss}(x,y,Nr)$ aus dem Startbild des entsprechenden Bildsequenzensets gewonnen.

Modell der **NOISE**-Helligkeitstransformation (ohne Sättigungseffekte):

$$b_{NOISE}(x, y, Nr) = b(x, y, 0) + n_{Gauss}(x, y, Nr) \tag{2}$$

$$\sqrt{\sigma_n} = \frac{\sqrt{2}^{(Nr-1)}}{4}, \quad Nr > 0$$

Die theoretische Rauschsignal-Standardabweichung des ersten veränderten Bildes ($Nr = 1$) wurde auf eine viertel Helligkeitsstufe gesetzt (minimale Helligkeit: 0, maximale Helligkeit: 63). Die Standardabweichung des Rauschens wurde anschliessend pro Bild jeweils um den Faktor $\sqrt{2}$ erhöht (Gl.2). Weil die maximal bzw. minimal darstellbaren Helligkeiten zahlenmässig nicht überschritten werden können, musste die resultierende Information auf diesen Bereich beschränkt werden. Dies hat zur Folge, dass die experimentell bestimmte Standardabweichung geringer ist als der entsprechende theoretische Wert. Mit sechzehn Bildern ist diese Sequenz die umfang-

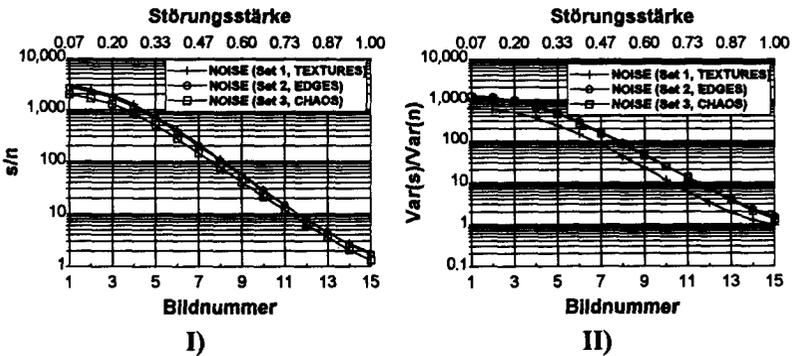


Fig.8: Signal-zu-Rauschleistungsverhältnis (I) bzw. Signal-zu-Störvarianzenverhältnis (II) in Abhängigkeit der Bildnummer für die Sequenzen **NOISE** in allen Sequenzensets.

reichste jedes Sequenzensets. Diese Länge war nötig, da einige der Algorithmen auch bei sehr stark verrauschten Suchfenstern noch gute Resultate lieferten. Um die Leistungsfähigkeit dieser Verfahren besser unterscheiden zu können, wurde die maximale (theoretische) Standardabweichung des Rauschens auf 32 Helligkeitsstufen gesetzt, was der halben Helligkeitsdynamik entspricht. Der für die Praxis interessante Zusammenhang zwischen Bildnummer und gemessenem Signal-zu-Rauschleistungsverhältnis ist der Figur 8.I zu entnehmen. Wiederum wird davon ausgegangen, dass das Anfangsbild vollkommen unverrauscht ist. In Figur 8.II sind die Varianzverhältnisse für den Bildstörungstyp **NOISE** gegeben. Da mittelwertfreies Rauschen verwendet wurde, entspricht die Rauschvarianz hier gerade der Rauschenergie¹⁶.

4.2.3. Salt and Pepper Rauschen, SALTY

‘Salt and Pepper’ Rauschen repräsentiert einen weiteren relevanten Rauschtyp. Die Sequenzen vom Typ **SALTY** wurden durch zufällige Auswahl von Pixeln, welche auf schwarz (Pepper) bzw. weiss (Salt) gesetzt wurden, erzeugt. Die Wahrscheinlichkeiten für diese Ereignisse wurden im ersten veränderten Bild auf 0.002 gesetzt und danach von Bild zu Bild um den Faktor 2 erhöht. Im letzten, d.h. neunten Bild sind somit theoretisch 51.2% aller Pixel verändert. Die Information dieser Pixel ist völlig unkorreliert mit der ursprünglichen Intensität.

Modell der **SALTY**-Helligkeitstransformation (ohne Sättigungseffekte):

$$b_{\text{SALTY}}(x, y, Nr) = \begin{cases} \text{MinIntens} & \text{falls } Z(x, y, Nr) < Schw1(Nr) \\ \text{MaxIntens} & \text{falls } Z(x, y, Nr) \geq Schw1(Nr) \\ & \text{und } Z(x, y, Nr) < Schw2(Nr) \\ b(x, y, 0) & \text{sonst} \end{cases} \quad (3)$$

$$\text{mit } Schw2(Nr) = 2 \cdot Schw1(Nr) = 2 \cdot 2^{(Nr-1)} \cdot 0.002 \quad \text{für } Nr > 0$$

$Z(x, y, Nr)$ bezeichnet ein Bild von gleichverteilten Zufallszahlen im Wertebereich von 0 bis 1. $Schw1$ und $Schw2$ kennzeichnen zwei Schwellen, die die ‘Kornrate’ für weisse ($MaxIntens$) und schwarze ($MinIntens$) Pixel bestimmen.

Im Unterschied zu Gauss'schem Rauschen, das alle Pixel betrifft, wird bei Salt and Pepper Rauschen die Information einzelner, ausgewählter Pixel

¹⁶ Die Signalvarianz entspricht aber nicht der Signalenergie!

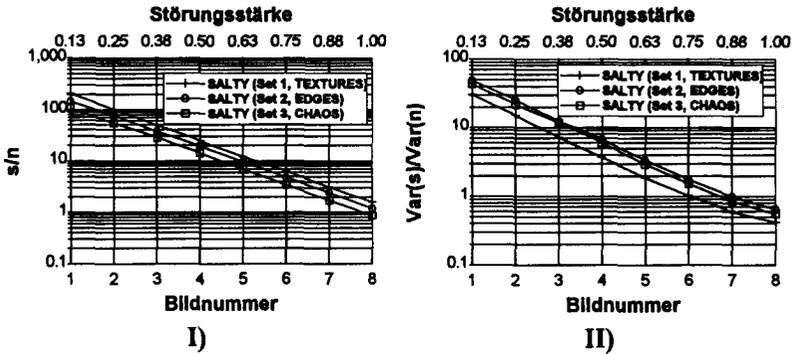


Fig.9: Signal-zu-Rauschleistungsverhältnis (I) bzw. Signal-zu-Strörvarianzenverhältnis (II) in Abhängigkeit der Bildnummer für die Sequenzen **SALTY** in allen Sequenzensets.

vollständig eliminiert. D.h. die Auswirkungen sind gleichartig wie bei einer vollständigen Abdeckung der Musterinformation (occlusion). Im Gegensatz zur echten Verdeckung, bei der entsprechende Pixel in der Regel clusterweise vorkommen, resultiert hier allerdings eine Gleichverteilung.

Durch die vollständige Beseitigung der Bildinformation an einzelnen Stützstellen des Bildes muss befürchtet werden, dass einerseits das Abtasttheorem nicht mehr eingehalten werden kann und andererseits hohe Frequenzkomponenten verändert werden. Neben den Auswirkungen der Störampplitude, ähnlich denjenigen des Gauss'schen Rauschens, muss also auch mit aliasing-Problemen gerechnet werden.

4.2.4. Unscharfe Bilder, FOCUS

Die Bildsequenzen vom Typ **FOCUS** wurden durch Filterung mit zwei-dimensionalen, rotationssymmetrischen Äquirippel-Tiefpässen (im Sinn von Tschebyscheff, siehe [142]) mit abnehmender Bandbreite aus den Anfangsbildern der Sequenzen **IRIS** erzeugt. Die Filteroperatorgröße wurde auf 31x31 Pixel festgesetzt, um alle Filterkoeffizienten, die im gewählten 16 Bit Fixkommazahlenformat ungleich Null sind zu berücksichtigen. Der Anfang des Stoppbandes, das durch eine minimale Dämpfung von 40dB gekennzeichnet ist, liegt bei 1/3, 1/4, 1/5,...1/15 Pixel⁻¹ (Fig.10).

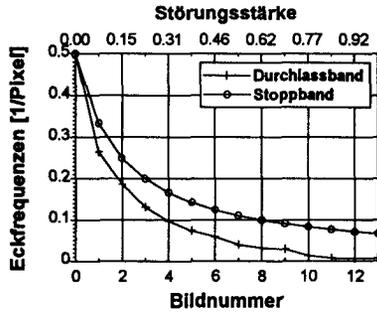


Fig.10: Start des Stoppbandes (mehr als 40dB Dämpfung) und Ende des Durchlassbereichs des Tschebyscheff-ähnlichen Äquirippel-Tiefpassfilters zur Berechnung der Sequenzen **FOCUS**.

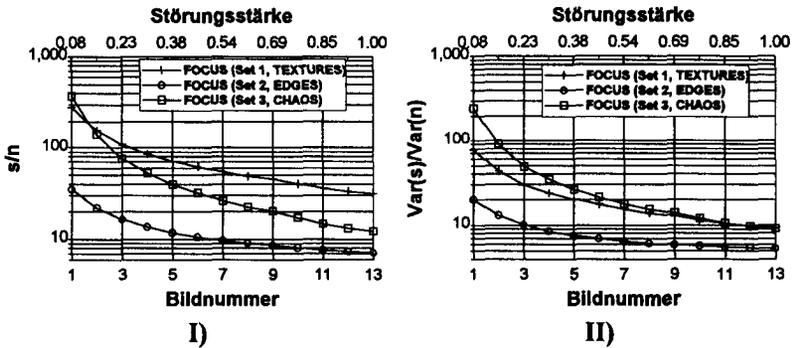


Fig.11: Signal-zu-Störleistungsverhältnis (I) bzw. Signal-zu-Störvarianzenverhältnis (II) in Abhängigkeit der Bildnummer für die Sequenzen **FOCUS** in allen Sequenzensets.

Modell der **FOCUS**-Helligkeitstransformation (im Frequenzbereich):

$$B_{\text{FOCUS}}(\omega_x, \omega_y, Nr) = B_0(\omega_x, \omega_y) * F(\omega_x, \omega_y, Nr) \quad (4)$$

*: Faltung der Operanden

$F(\omega_x, \omega_y, Nr)$: Frequenzgang des Filters

4.2.5. Geometrische Expansion, ZOOM

Die Bildsequenzen mit Namen **ZOOM** enthalten im Gegensatz zu den anderen Bildstörungssequenzen *geometrisch* transformierte Bilder. Der Bildinhalt wurde in 10%-Schritten von 100% auf 170% vergrößert (Fig.12) Zur Berechnung dieser Bilder wurde eine zweidimensionale, bilineare Interpolation verwendet. Damit die wahre Objektposition über alle Bilder konstant am gleichen Ort bleibt und nicht durch die Expansion immer weiter vom Expansionszentrum entfernt wird, wurde für jedes Muster ein eigenes Ausdehnungszentrum – die entsprechende wahre Objektposition – verwendet.

Modell der **ZOOM**-Transformation (das Ursprungsbild wird bilinear interpoliert):

$$SF_{\text{ZOOM}}(x, y, Nr) = SF_0 \left(x_0 + \frac{1}{v(Nr)} \cdot (x - x_0), y_0 + \frac{1}{v(Nr)} \cdot (y - y_0) \right) \quad (5)$$

$$v(Nr) = 1 + 0.1 \cdot Nr \quad \text{für } Nr > 0$$

Diese Expansion wird für jedes Suchfenster $SF(x, y, Nr)$ separat, relativ zum Objektzentrum (x_0, y_0) durchgeführt. Um eine Vergrößerung des Bildinhalts zu erreichen, muss der Expansionsfaktor $v(Nr)$ grösser als eins gewählt werden.

Alternativ zur Expansion hätte auch eine Bildverkleinerung realisiert werden können. Allerdings wäre dann ein zusätzliches anti-Aliasing-Filter zur Einhaltung des Abtasttheorems nötig gewesen ([33]). Beim oben beschriebenen Verfahren wird kein Filter benötigt, da man sich bei der Expansion von der Grenze des Abtasttheorems entfernt. Allerdings wird zugleich mit der Bildvergrößerung die relative Schärfe reduziert, da keine neue, vorher nicht sichtbare Information wahrnehmbar wird (Anhang B: *Bildmaterial*).

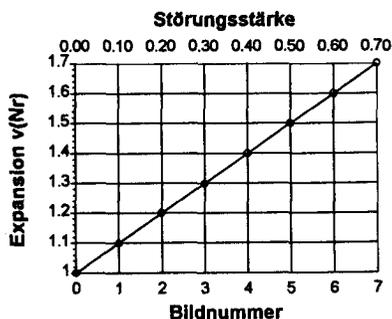


Fig.12: Expansion der Bilder in den Sequenzen **ZOOM**.

Das Aufzeichnen der Signal-zu-Rauschleistungs- oder der Varianzenverhältnisse scheint im Fall von geometrischer Transformation nicht sinnvoll zu sein, da diese Kurven sehr stark vom Bildinhalt abhängen und der Zusammenhang mit der Störung sehr komplex sein kann.

Bei den Versuchen mit den Bildsequenzen des Bildstörungstyps **ZOOM** wurden im Gegensatz zu den anderen Störungsarten nur die Resultate der Muster mit ungerader Kantenlänge ausgewertet. Dies hat den Vorteil, dass bei der Mustergewinnung die Daten symmetrisch bezüglich der wahren Objektposition anfallen. Bei geometrischen Bildtransformationen ist dies ein wichtiges Detail, da dann der Mittelpunkt des wahren Objektfensters geometrisch stabil bleibt und auf das Pixelraster zu liegen kommt.

4.3. Skala der Bildstörungsstärken

Der Einfachheit halber werden bei allen Auswertungen, unabhängig vom Bildstörungstyp, die Bildstörungsstärken durch die Bildnummern innerhalb der Sequenz charakterisiert. Diese Zahlen werden für alle Sequenzen, mit Ausnahme der Sequenzen **ZOOM**, einzeln derart auf das Intervall 0..1 abgebildet, dass eine einheitliche, qualitative – quantitativ aber nicht vergleichbare – Skala entsteht. Ein Bild mit Bildstörungsstärke Null entspricht dem unveränderten Originalbild mit maximaler Güte. Die minimale getestete Bildqualität wird bei Bildstörungsstärke Eins erreicht. Der Zusammenhang zwischen dieser Grösse und den in den Bildern messbaren Werten wie Expansion, Bandbreite des Bildsignals, Varianzen- und Signal-zu-Rauschleistungsverhältnissen wird durch die oben aufgeführten Grafiken gegeben (siehe auch Anhang E.1: *Resultate, Grafiken*).

4.4. Praktische Relevanz der Bildsequenzen

Die verschiedenen Bildsequenzen repräsentieren eine Auswahl von Nicht-idealitäten bei der Bildaufnahme, wie sie einzeln oder kombiniert in realen Anwendungen vorkommen. Die gewählte Stärke der maximalen Bildveränderungen ist aber bei Anwendungen im allgemeinen nicht üblich. Vielmehr geben diese Fälle eine Art 'worst case' wieder. Die Stärke der maximalen Bildveränderungen wurde derart gewählt, dass die verschiedenen Algorithmen gut in ihrer Leistungsfähigkeit unterschieden werden können.

Die ausgewählten Bildtransformationen haben eine grosse praktische Bedeutung und führen immer wieder zu Problemen bei Anwendungen:

IRIS: Diese Sequenz simuliert z.B. Beleuchtungsschwankungen.

- NOISE:** Vielfach kann das (Video-) Sensorrauschen als additives, mittelwertfreies Gauss'sches Rauschen modelliert werden ([18], [26], [101] und [169]).
- SALTY:** Neben der Modellierung von fatalen, in der Regel sehr seltenen Sensor- und Elektronikfehlern kann mit dieser Sequenz auch die teilweise Verdeckung eines Helligkeitsmusters durch Vordergrundinformation kontrolliert nachgebildet werden (allerdings fallen dort die Störpixel meist clusterförmig an).
- FOCUS:** Nachbildung von Unschärfe verursacht durch Defokussierung oder Distanzänderung zwischen Kamera und Objekt.
- ZOOM:** Diese Sequenz simuliert geometrische Bildtransformationen, wie sie z.B. durch Objektiv-Brennweitenänderung oder aber ändernde Kamera-Objekt-Distanz eingeführt werden. Da eine praktische (nicht rechnerische) Realisierung speziell bei kleinen Kamera-Objekt-Distanzen mit einer Defokussierung einhergeht, erscheint die gewählte rechnerische Implementation auch aus dieser Sicht sinnvoll. Natürlich wären bei den geometrischen Transformationen z.B. auch perspektivische Verzerrungen von Interesse. Aus Rechenzeitgründen sowie auch aus der Schwierigkeit heraus in einer solchen Situation die objektiv korrekte Position eines flächenhaften Musters zu bestimmen, wurde auf die Implementation dieses Verzerrungstyps verzichtet.

In Anhang B: *Bildmaterial* sind Abzüge eines Teils der Bilder beigeheftet. Die Rasterung der Bilder bringt aber leider einen Bildqualitätsverlust bezüglich der Geometrie- und Helligkeitsauflösung mit sich.

4.5. Bemerkungen zu den Untersuchungen

Um die Zahl der Experimente und damit die totale Rechenzeit auf ein vernünftiges Mass zu beschränken, mussten einige Kompromisse bei der Anzahl variiertes Parameter gemacht werden. Z.B. konnten für Anwendungen wichtige Parameter einzelner Algorithmen nicht in vollem Umfang untersucht werden (z.B. Vorverarbeitungsfiler von Nack für **NA1** und **NA2**). So wurde meist auf die in der Literatur vorgeschlagenen Parameterwerte zurückgegriffen, obwohl beim verwendeten Bildmaterial und Versuchskonzept mit anderen Einstellungen möglicherweise ein optimaleres Verhalten der Algorithmen hätte erreicht werden können. Trotz diesen Einschränkungen sollte es möglich sein, vergleichende, qualitative Erkenntnisse aus den hier vorliegenden Untersuchungen in echte Applikationen zu übertragen.

Im folgenden sind einige nicht untersuchte Störungsarten, Vor- und Nachverarbeitungsmethoden sowie Rechenstrategien aufgeführt, die die Leistungsfähigkeit der Algorithmen und/oder den Rechenbedarf beeinflussen können:

- Multiplikatives Rauschen, z.B. Dunst in Satellitenbildern kann in dieser Art modelliert werden ([39], [44]).
- Rotation und perspektivische Bildverzerrungen¹⁷.
- Nicht uniformes Rauschen ([110]).
- Simultanes Auftreten von verschiedenen Störungsarten sowie gleichzeitige Störung sowohl des Suchfensters als auch des Musters.
- Vorverarbeitung der Bilder durch Filter (z.B. [134]).
- Positionsbestimmung nicht nur mit einem einfachen Minimum- bzw. Maximumdetektor, sondern z.B. über Vorkenntnisse der statistischen Eigenschaften der Werte der entsprechenden Korrelationsfunktion ([122]).
- Rückweisung von unwahrscheinlichen Objektpositionen z.B. durch Analyse des Korrelationswertes und andere, das Resultat beurteilende Plausibilitätsbetrachtungen (z.B. [22], [40], [67], [78], [122] und [131]).
- Bestimmung der optimalen Suchfenstergröße ([65], [128]).
- Bestimmung der optimalen Mustergröße/Musterform ([33], [65], [126], [128], [148] und [150]). Dem räumlichen Objekt evtl. adaptiv angepasste, nicht rechteckige Musterformen können die Leistung eines Algorithmus positiv beeinflussen. Allerdings nimmt dabei die Komplexität der Adressrechnung bei der Implementation des Algorithmus' zu.
- Muster grösser als 25x25 Pixel.
- Verschiebung zwischen Muster und Hinter- bzw. Vordergrund.
- *Sequential similarity detection algorithm* ([2], [20], [59], [64], [67], [137], [163], [174] und [177]); wird oft mit *SSD* oder auch *SSDA* abgekürzt, dieses Akronym wird hier aber bereits für ein Unähnlichkeitsmass verwendet). Die **SSD** und die **SAD** (sowie die verwandten Verfahren) können derart implementiert werden, dass die Berechnung der entsprechenden Endsumme abgebrochen wird, sobald das Zwischenresultat eine gewisse, möglicherweise adaptive Unähnlichkeitsschwelle überschreitet. Dies kann zu einer erheblichen Reduktion des Rechenbedarfs führen.

¹⁷ Nach Silver ([147]) ist bei Bildrotationen ein prinzipiell ähnliches Verhalten zu erwarten wie bei Grössenänderungen (**ZOOM**).

- Auflösungspyramiden ([1], [2], [4], [8], [19], [23], [31], [60], [67], [83], [105], [119], [136], [137], [166] und [172]) werden einerseits zur Reduktion des Rechenaufwandes bei grossen Fenstern und andererseits zur Steigerung der erlaubten Grössendynamik des Bildinhaltes eingesetzt.
- Verminderung des Rechenaufwandes durch Unterabtastung der Bildinformation und ähnliche Verfahren ([176]).
- Durch intelligente Suchstrategien kann unter gewissen Annahmen über die Stetigkeit der Korrelationsfunktion, ohne Veränderung des Resultats, der zu bearbeitende Suchraum (Anzahl Positionen, die verifiziert werden müssen) und damit der Rechenaufwand drastisch verringert werden ([3], [80], [100], [108], [137], [152], [159] und [172]).
- Koordinatentransformationen (bzw. die Wahl geeigneter Koordinatensysteme) können zur Reduktion der Sensitivität bezüglich geometrischen Bildtransformationen bzw. zur Genauigkeitssteigerung eingesetzt werden ([15], [18], [34] und [160]).
- Helligkeitstransformationen (z.B. logarithmische Helligkeitsskala, [45]; Histogrammausgleich) können die Aufgabe für die Korrelationsalgorithmen bei schwierigen Beleuchtungsverhältnissen vereinfachen.
- In vielen Anwendungen kann durch konsequentes Ausnützen von Vorkenntnissen über die zu erwartenden Szenen bei gleicher Detektionsleistung ein einfacherer Algorithmus ausgewählt werden, der weniger grosse Ansprüche an die benötigte Rechenleistung stellt als ein Algorithmus, der ohne Vorkenntnisse auskommt (z.B. ternäre Korrelation, [92]).
- Subpixelgenauigkeit von Positionierungen kann z.B. durch Interpolation der Korrelationsfunktionen erreicht werden ([12], [67], [129], [158] und [165]).
Eigene Interpolationsversuche lassen je nach Bildinhalt und Korrelationsalgorithmus eine maximale Genauigkeitssteigerung um den Faktor 10 erwarten ([12]). Eine Untersuchung über die Eignung der verschiedenen Korrelationsfunktionen für ein derartiges Subpixel-Verfahren, müsste die Funktionen auf ihre Stetigkeit und auf Formveränderung, z.B. verursacht durch Intensitätsrauschen, untersuchen. Es ist zu erwarten, dass Korrelationen, die Hochpass gefilterte Daten verwenden eher weniger für solche genauigkeitssteigernden Interpolationen geeignet sind (schlanke Korrelationsspitzen).
- Die Datenformate am Eingang (Pixelwerte) und am Ausgang (Korrelationswerte) beeinflussen sowohl das Positionierungsverhalten wie auch die Rechenlast (bei entsprechender Hardware).

- Durch Wahl der Szenenbeleuchtung, evtl. sogar mit strukturier-tem/unstrukturiertem Licht ([124]), kann die Leistung der Algorithmen in Applikationen ebenfalls gesteigert werden.
- Die Form der Korrelationsflächen kann unter Umständen bei der Detek-tion des Extremwerts sehr wichtig sein ([40], [7] und [73]).

Eine schöne Zusammenstellung einiger der angesprochenen Verfahren findet sich z.B. in [67].

Neben diesen, in den vorliegenden Untersuchungen nicht diskutierten Verfahren und Ideen, muss auch erwähnt werden, dass die Anzahl der Ver-suche und die Auswahl des Bildmaterials nicht ausreicht, um in jedem Fall zuverlässige und präzise, quantitative Vorhersagen über die Leistung der einzelnen Algorithmen zu machen. Eine vergleichende Abschätzung kann aber sehr wohl durchgeführt werden.

Die Auswahl der untersuchten Algorithmen ist nicht vollständig. Interes-sante Algorithmen wie z.B. *(binary) phase only matched filter* und das *cepstrum matching* sind nicht verglichen worden, weil insbesondere die Wahl der Muster- und Suchfenstergrößen gewissen Einschränkungen unter-liegen (siehe Kapitel 7: *Diskussion von Beiträgen aus der Literatur*). Zudem ist die Windowing-Funktion (z.B. bei der Fourier-Transformation) entschei-dend für das Verhalten dieser Verfahren ([103]) und müsste demnach (dem Frequenzgehalt des Suchfensters entsprechend) optimiert werden.

5. Resultate

In diesem Kapitel sollen die Korrelationskriterien bezüglich ihrem Verhalten bei unterschiedlich grossen Mustern sowie ihrer Robustheit gegenüber verschiedenen Störungen untersucht werden. Die Störungsarten, die Generierung des entsprechenden Bildmaterials wie auch die Versuchsabläufe wurden in Kapitel 4: *Versuche und Versuchsbedingungen* im Detail diskutiert.

5.1. Beurteilungskriterien

Die Leistungsfähigkeit der Korrelationskriterien wird hier anhand einfacher Statistiken über die Abweichungen der berechneten Objektpositionen von den Sollpositionen beurteilt. -Bei der weiteren Besprechung wird öfters von guter oder schlechter Leistung eines Algorithmus' gesprochen. Es darf aber nicht vergessen werden, dass diese Wertung je nach Anwendung auch genau entgegengesetzt sein kann. So kann z.B. die Sensitivität eines Algorithmus gegenüber geometrischen Verzerrungen in einer Applikation wichtig sein, um nur korrekt geformte Fabrikationsteile zu detektieren. Andererseits ist beim *depth from motion* Verfahren zur Gewinnung eines Tiefeneindrucks mit Hilfe einer einzigen, kooperativ bewegten Videokamera ([140]) eine gewisse Robustheit gegenüber perspektivisch verzerrten Bildern zwingende Voraussetzung, um die Zeit während der ein bestimmtes Muster zuverlässig verfolgt werden kann und damit die insgesamt resultierende Disparität, zu maximieren. Mit 'guter' Leistung soll in diesem Zusammenhang also die hohe Robustheit gegenüber Störungen gemeint sein.

Ein wichtiger Begriff bei der Diskussion der Leistungsfähigkeit der einzelnen Korrelationsalgorithmen wird die *Detektionsrate*¹⁸, der Anteil tolerierbarer Positionierungen an der Gesamtmenge der betrachteten Experimente, sein. Sie gibt – zumindest für die untersuchten Fälle – die Wahrscheinlichkeit an, bei spezifizierter Mustergrosse und Bildstörungsstärke eine tolerierbare Positionierung zu erreichen. Unter einer tolerierbaren Position wird – unter der Annahme quadratischer Pixel – eine Position verstanden, die innerhalb eines durch den Radius r_t spezifizierten Toleranzkreises (TK_t) um die wahre Objektposition (x_0, y_0) liegt (Fig.13).

¹⁸ Diese Namengebung lehnt sich an den Begriff Falschalarmrate an, die den mittleren Anteil von fehlerhaften Positionierungen angibt.

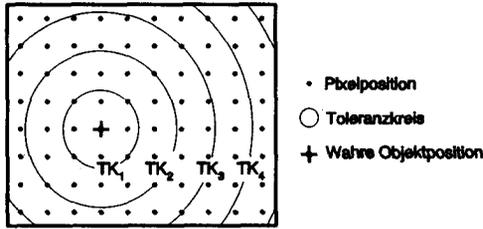


Fig.13: Beispiel einer Korrelationsfläche mit eingetragenen Toleranzkreisen TK_i .

Die Korrelationsfläche kann zur detaillierteren Fehleranalyse gleichermaßen in verschiedene Toleranzringe unterteilt werden. Die inneren und äusseren Radien dieser Teilgebiete sind durch die folgenden Beziehungen gegeben:

$$\begin{aligned} r_i(m) &= (m-1)\sqrt{2} & r_i(0) &= -0 \\ r_a(m) &= m\sqrt{2}, & \text{mit } m \in \text{ganzen Zahlen, } m \geq 1, r_a(0) &= +0 \end{aligned} \quad (6)$$

Ein Pixel innerhalb eines Toleranzrings mit Index m erfüllt somit die folgende Ungleichung:

$$TK_m: r_i(m) < \sqrt{(x-x_0)^2 + (y-y_0)^2} \leq r_a(m) \quad (7)$$

Die Diskussion der (experimentellen) Robustheiten der Algorithmen gegenüber Bildstörungen wird anhand von Grafiken und daraus abgeleiteten Zahlenwerten vorgenommen:

- *Detektionsratenverlauf* in Abhängigkeit von Mustergrosse und Bildstörungsstärke (als dreidimensionale Grafik für Toleranzkreise mit Radius $r_i = \sqrt{2}$ Pixel realisiert, siehe Fig.14 und 15).
- Aus den Detektionsratenverläufen abgeleitete *Kennzahlen* (Fig.14 und 15).
- *Fehlerhistogramm*, das die Häufigkeit der verschiedenen Positionierungsabweichungen angibt; die Toleranzringe definieren dabei die Fehlerklassen (Fig.16).

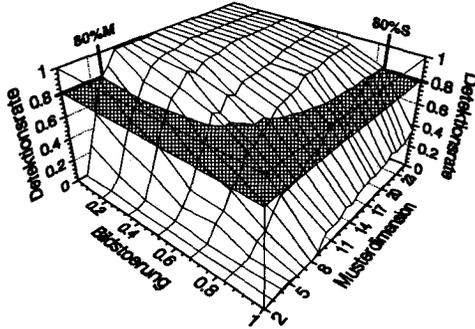


Fig.14: Erläuterung der 80%-Störungsstärke (**80%S**) und der 80%-Mustergrösse (**80%M**) anhand des Detektionsratenverlaufs von Nishiharas Korrelationsmass (**NIS**, $w=7$) beim Störungstyp **IRIS**.

Der Detektionsratenverlauf¹⁹ bezüglich Mustergrösse und Bildstörung erlaubt einen raschen Überblick über die Robustheit eines Korrelationskriteriums. Die Frage 'Unter welchen Bedingungen (Mustergrösse, Bildstörungsstärke) darf mit einer bestimmten Detektionsrate gerechnet werden?' kann damit sehr einfach beantwortet werden²⁰. Der Vergleich der Algorithmen anhand dieser Darstellung ist aber äusserst aufwendig. Darum wird die Gegenüberstellung hauptsächlich anhand von Zahlenwerten, die aus den Detektionsratenverläufen eruiert werden können, durchgeführt.

Die (*mittlere*) *Detektionsrate* stellt dabei die wichtigste Kennzahl dar. Sie bezeichnet für jeden Algorithmus und jede Störungsart die von der Mustergrösse und der Bildstörungsstärke unabhängige Wahrscheinlichkeit, eine tolerierbare Positionierung zu erreichen. Dieser Wert ist also proportional zum Integral des Detektionsratenverlaufs (über die Mustergrösse und die Störungsstärke).

Daneben muss aber auch die Form der Detektionsratenverläufe bzw. die Abhängigkeit der Detektionsrate von den Experimentierbedingungen charakterisiert werden. Hier würde sich z.B. eine Approximation einer Modellfläche an die experimentell ermittelten Stützstellen im Sinne der kleinsten Quadrate anbieten. Ein für alle Korrelationskriterien gleichsam gültiges, mathematisch fundiertes Modell, dessen Parameter einfach zu interpretieren

¹⁹ Wegen dem grossen Umfang dieser Grafiken, sind sie im Anhang E.1: *Resultate, Grafiken* beigelegt.

²⁰ Die dokumentierten Werte gelten streng natürlich nur für das verwendete Bildmaterial unter den oben definierten Bedingungen.

sind, ist aber nur schwer zu finden. Es erscheint deshalb sinnvoll, direkter zu bestimmende, der Praxis nahestehende Werte zu berechnen.

Der Applikationsingenieur, der die Wahl eines Korrelationskriteriums vornehmen muss, ist in der Regel nicht am ganzen Detektionsratenverlauf interessiert. Von Interesse ist für ihn vielmehr die Schnittfläche dieser Fläche mit einer Ebene, die durch die minimal zu garantierende Detektionsrate beschrieben wird (Fig.14). Als gute Detektionsratenschwelle hat sich z.B. die 80%-Ebene erwiesen (Detektionsrate mindestens 0.8). Sie ist einerseits genügend tief, dass praktisch alle Detektionsratenverläufe die entsprechende Ebene schneiden und andererseits genügend hoch, dass die unterschiedlich starke Abhängigkeit der Leistungsfähigkeit der verschiedenen Algorithmen von der Mustergrösse und Bildstörungsstärke gut sichtbar wird.

Versuche mit Hilfe einer Approximation im Sinne der kleinsten Quadrate die so erhaltenen Schnittkurven durch verschiedene Polynom- und Exponential-Modelle anzunähern, ergaben speziell für kleine Mustergrössen eine relativ schlechte Übereinstimmung und auch wenig aussagekräftige Parameter²¹. Deswegen wurden schliesslich direkt die Durchstosspunkte der Schnittkurve durch die Koordinatenebenen als charakteristische Werte bestimmt (Fig.14). Die 80%-Mustergrösse (**80%G**) bezeichnet die linear interpolierte (und damit real in der Regel nicht realisierbare) Mustergrösse, bei der im ungestörten Fall eine Detektionsrate von genau 80% erreicht wird. Die 80%-Störungsstärke (**80%S**) wird analog dazu als diejenige linear interpolierte Störungsstärke definiert, die bei maximal grossem Muster (25x25 Pixel) gerade noch eine 80%-Detektionsrate erlaubt. Da die 80%-Mustergrösse bei Störungsstärke Null bestimmt wird, ist sie unabhängig von der Störungsart. Somit muss nur die 80%-Störungsstärke, nicht aber die 80%-Mustergrösse, für jeden Störungstyp angegeben werden.

Für geometrische Verzerrungen, wie sie z.B. durch den Störungstyp **ZOOM** definiert sind, reicht diese Charakterisierung nicht aus, da, wie später noch gezeigt wird, die Detektionsrate der Korrelationskriterien ab einer bestimmten Grösse mit wachsender Musterabmessung abnimmt (Fig.15). -Die 80%-Mustergrösse wurde für diese Störungsart nicht neu berechnet, obwohl in diesem Fall nur Muster mit ungeraden Kantenlängen untersucht wurden.

²¹ Für die Störungen **IRIS**, **NOISE** und **FOCUS** (und teilweise **SALTY**) kann die Schnittkurve, ausser für geringe Bildstörungen, z.B. gut durch ein Polynom vierter Ordnung angenähert werden (Mustergrösse=f(Bildstörung)). Dieses Modell geht davon aus, dass die Robustheit eines Korrelationskriteriums mit wachsender Mustergrösse zunimmt und kleine Störungen einen geringen Einfluss auf die Detektionsraten haben (horizontale Tangente im ungestörten Fall, siehe Fig.14).

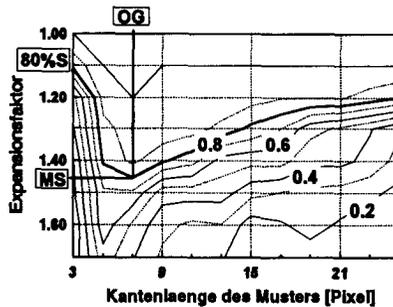


Fig.15: Konturplot des Detektionsratenverlaufs der *KKF* beim Störungs-*typ ZOOM* zur Erläuterung der optimalen Mustergrösse (*OG*) und der dazugehörigen Maximumstörungsstärke (*MS*).

Zusätzlich zum 80%-Störstärkenwert wird noch die *optimale Mustergrösse* (*OG*, nur ungerade, nicht interpolierte Werte) und die *Maximumstörungsstärke* (*MS*, linear interpolierter Wert) bestimmt (Fig.15). Die optimale Mustergrösse bezeichnet diejenige Mustergrösse, die über den grössten Verzerrungsbereich eine 80 prozentige Detektionsrate garantiert (gilt nur für *ZOOM*-Störung!). Die zugehörige Maximumstörungsstärke (*MS*) gibt das Ausmass dieser Robustheit an. Eine Vergrösserung oder Verkleinerung der Musterfläche gegenüber der optimalen Mustergrösse geht immer mit einer Reduktion der maximalen 80%-*ZOOM*-Störungsstärke einher.

Wie aus der Figur 15 hervorgeht ist die optimale Mustergrösse abhängig von der gewählten Detektionsratenschwelle. Für kleinere Schwellen darf in der Regel mit einer geringeren optimalen Mustergrösse gerechnet werden.

Die nunmehr definierten, charakteristischen Werte geben nur einen groben Anhaltspunkt über die Form des Detektionsratenverlaufs (bzw. der Form des Schnittbildes mit der 80%-Ebene). Deshalb sollten sie immer zusammen mit der mittleren Detektionsrate und im Zweifelsfall mit den entsprechenden Detektionsratenverläufen betrachtet werden.

Die für jede Korrelationsfunktion experimentell ermittelten Kurven und Zahlen stellen nur Richtwerte und vergleichende Grössen, also weder als obere noch als untere Grenze zu verstehende Detektionsraten dar. Auch die Definition einer tolerierbaren Positionierung ist natürlich von Anwendung zu Anwendung verschieden. Bei Präzisionsvermessungen wird nur eine perfekte d.h. pixelgenaue Positionierung (evtl. zusammen mit einem Verfahren zur subpixelgenauen Positionsbestimmung) akzeptabel sein, in anderen Fällen darf die tolerierbare Abweichung bei einigen – im Normalfall wenigen – Pixeln liegen.

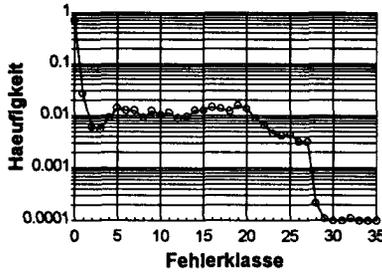


Fig. 16: Fehlerhistogramm für die **KKF** und Störungstyp **IRIS** (siehe Anhang E.1: Resultate, Grafiken). Der logarithmische Massstab erleichtert die Beurteilung der geringen Häufigkeiten von grossen Fehlern.

$$\text{Fehlerklasse} = \text{Abstand}/\sqrt{2}.$$

$$\text{Fehlerhistogramm}(i) = \frac{\sum_{\forall Pos. \in TR_i} (Pos. \in TR_i)}{\sum_{\forall Pos.} (Pos. \in SF)} \quad (8)$$

TR_i : Gebiet des Toleranzrings (Gl.6 und 7)

$Pos.$: Positionierungen

SF : Suchfenster

Die Fehlerhistogramme (siehe Gl.8 und Fig.16) stellen die Häufigkeiten der verschiedenen Fehlerklassen dar und sollen eine Aussage über die Präzision der einzelnen Korrelationsverfahren machen. Die Fehlerklassen sind wie oben bereits erwähnt durch die Toleranzringe definiert. Günstig sind Fehlerhistogramme mit einer hohen Spitze bei Fehlerdistanz Null und stark abfallendem Verlauf für zunehmende Fehlerbeträge. Unpräzisere Korrelationsalgorithmen müssten demnach eine breite Spitze aufweisen. Die allgemeine Tendenz einer verbreiterten Spitze im Fehlerhistogramm ist für die Bildstörungstypen **FOCUS** und **ZOOM** einfach einzusehen. Die Tiefpassfilterung auf den Bildern (**FOCUS**) führt, wegen der Dämpfung der präzisionsfördernden, hohen Signalanteilen, zu einer breiteren, entsprechend unpräziser lokalisierten Korrelationsspitze. Für die kontinuierliche, eindimensionale, unendlich ausgedehnte und unnormierte Kreuzkorrelationsfunktion (ohne Vorverarbeitungsfiler), die in diskreter und endlicher Form einen wesentlichen Term in verschiedenen Korrelationsalgorithmen darstellt (**KKF**, **KKFMF**, **MOR** sowie nach entsprechender Umformung auch **NSSDMF**, **SSD**, **SSDMF** und **SSDMS**), kann sehr einfach gezeigt werden, dass eine

lineare Tiefpassfilterung auf dem Bild einer entsprechenden Filterungen der Korrelationsfunktion gleichkommt:

$f(x)$: Filterfunktion

$k(x)$: Korrelationsfunktion

$$\begin{aligned}
 k_{\text{unscharf}}(x) &= \int_{-\infty}^{\infty} \underbrace{r(u) \cdot s_{\text{unscharf}}(u+x)}_{\text{Korrelation mit unscharfem (Tiefpass-gefiltertem) Bild}} du = \int_{-\infty}^{\infty} r(u) \int_{-\infty}^{\infty} \underbrace{f(\xi) \cdot s(u+x-\xi)}_{\text{Tiefpassfilterung (Faltung mit } f(\xi))} d\xi du = \\
 &= \int_{-\infty}^{\infty} f(\xi) \int_{-\infty}^{\infty} \underbrace{r(u) \cdot s(u+x-\xi)}_{\text{Korrelation mit ungefiltertem Bild verschoben um } -\xi} du d\xi = \int_{-\infty}^{\infty} f(\xi) \cdot k(x-\xi) d\xi = \\
 &= \int_{-\infty}^{\infty} \underbrace{k(\xi) \cdot f(x-\xi)}_{\text{Tiefpassfilterung der Korrelationsfunktion } k(x) \text{ (Faltung)}} d\xi \quad (9)
 \end{aligned}$$

Die geometrische Transformation des Bildinhalts (**ZOOM**) verschiebt im Muster enthaltene Bildteile gegeneinander. Wenn dadurch wichtige, lokale Helligkeitswechsel im Bild verschoben werden, kann sich die Position des Korrelationsextremwerts ebenfalls entsprechend ändern. Dies führt wie bei unscharfen Bildern zu einer verbreiterten Spitze im Fehlerhistogramm.

Bei der Beurteilung der Fehlerhistogramme ist zu beachten, dass die Anzahl der möglichen Positionen (Pixel) in den einzelnen Fehlerklassen (Toleranzringen) nicht konstant ist. Falls die Positionierungen rein zufällig und über die Korrelationsfläche gleichverteilt wären, so wären in der Regel Fehlerklassen mit kleinem Toleranzradius wegen der ihnen eigenen, kleineren Fläche nur schwach belegt (siehe Fig. 17)²².

Das Fehlerhistogramm wurde, wie schon erwähnt, ursprünglich zur Beurteilung der Positioniergenauigkeit der einzelnen Verfahren eingeführt. Die Präzision eines Algorithmus¹ sei hier anhand der Abweichung der Positionierungen vom richtigen Ort gemessen. Dabei werden aber nur tolerierbare

²² Dies gilt natürlich nur solange die Toleranzringe nicht die Begrenzungslinien der Korrelationsfläche schneiden (siehe Fig. 13).

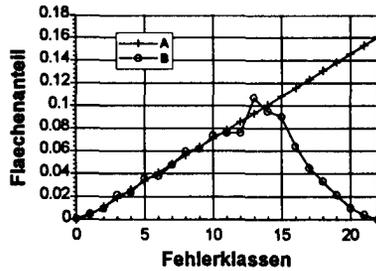


Fig.17: Flächenanteil der einzelnen Fehlerklassen bzw. Toleranzringe. Die Korrelationsfläche enthält in diesem Beispiel 41×41 mögliche Positionen. Die wahre Objektposition wird exakt in der Mitte dieses Fensters angenommen. Der Abfall der Klassenbesetzung für grosse Toleranzradien rührt von der geraden Begrenzung der Korrelationsfläche her (Kurve B, Fig.13). Kurve A zeigt den Flächenanteil ohne Pixelquantisierung und ohne Suchfenstereinschränkung.

Positionierungen in Betracht gezogen²³. Schlechte Präzision müsste sich durch eine breitere Spitze bei kleinen Fehlern bemerkbar machen. Wie aber in den Darstellungen der Fehlerhistogramme im Anhang E.1: *Resultate, Grafiken* oder auch als Beispiel in Fig.16 zu ersehen ist, kann höchstens beim Korrelationsalgorithmus von Nishihara eine leichte Verbreiterung der Spitze des Fehlerhistogramms mit Zunahme des Tiefpasscharakters des Vorverarbeitungsfilters beobachtet werden. Bei keinem anderen Verfahren kann schlüssig ein ähnliches Verhalten festgestellt werden. Vielmehr scheint der Bildinhalt selbst, also z.B. die Korrelation benachbarter Pixelwerte, die Form der Spitze im Fehlerhistogramm zu definieren (vergleiche die Fehlerhistogramme für die Bildstörungstypen **FOCUS** und **ZOOM** mit denjenigen der anderen Bildstörungstypen).

Der Einfluss der räumlichen Korrelation der Pixel auf die Form des Fehlerhistogramms ist nach Gleichung 9 zu erwarten. Die weitgehende Unabhängigkeit der Resultate vom gewählten Korrelationsalgorithmus scheint aber nicht von vornherein klar. Dieses Phänomen hängt offenbar auch mit der hier verwendeten Definition der Präzision ab: Nur nicht tolerierbare Positionierungen werden für die Beurteilung der Präzision herangezogen;

²³ Untolerierbare Positionierungen hingegen tragen zur Fehlerrate, dem Pendant zur Detektionsrate, bei. –Bei diesen, hier nicht weiter diskutierten Untersuchungen wurde die Toleranzgrenze auf Werte zwischen $\sqrt{2}$ und $5\sqrt{2}$ gesetzt.

diese Fälle werden aber durch die Fehlerrate bzw. die Detektionsrate abgedeckt.

Die Korrelations-Präzision könnte z.B. über die Steigung des Fehlerhistogramms bei kleinen Fehlern auch quantifiziert werden. Ein Blick auf die Histogramme zeigt aber, dass das Verhalten in diesem Fehlerklassenbereich weitgehend von der Häufigkeit der perfekten oder tolerierbaren Positionierungen abhängt und somit eher mit der mittleren Detektionsrate als mit der Genauigkeit zusammenhängt²⁴.

Natürlich gibt es dennoch erkennbare Unterschiede in den Fehlerhistogrammen der verschiedenen Korrelationskriterien. Diese rühren aber nicht von der entsprechenden Präzision, sondern vielmehr von 'fehlerhaft interpretierter Bildinformation' her (nicht die richtige Position sondern die Position eines für den entsprechenden Korrelationsalgorithmus ähnlichen Helligkeitsmusters gefunden, siehe z.B. Nebenmaximum im Fehlerhistogramm von **NA1** beim Bildstörungstyp **IRIS**, Anhang E.1: *Resultate, Grafiken*).

5.2. Erwartete Detektionsratenverläufe

Für die Störungen vom Typ **IRIS**, **NOISE** und **FOCUS** kann tendenziell dieselbe Kurvenform erwartet werden: Die Detektionsrate nimmt mit zunehmender Mustergrösse und abnehmender Störung stetig zu, bis sie den Wert Eins annimmt und damit ihr Maximum erreicht.

Die Tiefpassfilterung des Bildinhalts (**FOCUS**) – und damit die Dämpfung der für eine genaue Positionierung notwendigen Signalanteile höherer Raumfrequenz – erhöht die Positionierungsunsicherheit. Je unschärfer das Bild, desto unzuverlässiger also der Ort des Extremwerts. Beim Störungstyp **SALTY** ist für kleine Muster bei zunehmender Grösse mit einem steilen Anstieg der Detektionsrate zu rechnen. Dieser Anstieg wird sich in der Regel für mittlere Mustergrössen verflachen (evtl. sogar umkehren) und bei grösseren Referenzfenstern wieder anwachsen. Der Grund für dieses Verhalten ist in zwei gegenläufigen Einflüssen zu suchen; einerseits in der für kleine Muster geringen Wahrscheinlichkeit eines störenden 'Salz- oder Pfefferkorns' im wahren Objektfenster und andererseits dem reduzierten Informationsgehalt für sehr kleine Muster.

Der Störungstyp **ZOOM** erzeugt in der Regel Detektionsratenverläufe, die zuerst mit der Mustergrösse ansteigen, da der Informationsgehalt zunimmt und damit die Wahrscheinlichkeit für ähnliche Helligkeitsverteilungen (an

²⁴ Ausser bei den Bildstörungstypen **ZOOM** und **FOCUS** nimmt, bei einem Toleranzkreisradius von $\sqrt{2}$, die Anzahl perfekter Positionierungen (keine Abweichung) in der Regel etwa 90-95% der tolerierbaren Positionen ein.

der falschen Position) im selben Suchfenster abnimmt. Schon bei relativ kleinen Mustern – in der Grössenordnung von unter 10×10 Pixeln – tritt aber ein gegenläufiger Trend ein. Die Detektionsrate nimmt mit zunehmender Mustergrosse ab. Dies kann durch eine einfache Überlegung erklärt werden. Alle hier untersuchten Korrelationsverfahren basieren auf dem pixelweisen Vergleich zwischen Muster und Korrelationsfenster. Unter der (meist nicht ganz korrekten) Annahme von wenig korrelierten Pixeln muss also grundsätzlich davon ausgegangen werden, dass die Verfahren nur über eine beschränkte Toleranz gegenüber geometrischen Bildverzerrungen verfügen²⁵.

Im Fall der Vergrösserung des Bildinhalts, wie dies beim Störungstyp **ZOOM** auftritt, kann aufgrund plausibler Argumente eine vernünftige, wenn auch Algorithmus-unspezifische Expansionsgrenze abgeleitet werden, die das Verhalten bei grösseren Mustern erklärt: Eine Detektionsratenreduktion ist dann zu erwarten, wenn die Intensitäten einzelner Pixel des Objektfensters wegen der Bildtransformation wesentlich von den Nachbarn des Ursprungsbildes beeinflusst werden. Bei zentrischer Expansion, bezüglich des Objektfensters, kann somit für den eindimensionalen Fall die folgende Ungleichung aufgestellt werden (nur für Muster grösser als 4×4 Pixel und ungerade Muster-Kantenlängen xlr gültig):

$$v \cdot \left\{ \frac{xlr - 1}{2} - 1 \right\} \leq \left\{ \frac{xlr - 1}{2} - 1 \right\} + \delta \quad (10)$$

Die Grenze der erlaubten Vergrösserung v wird demnach dann erreicht, wenn die Information der zweitäussersten Objektfensterpixel um δ Pixel in Richtung der äussersten Pixel verschoben wird. Die Auflösung nach dem Expansionsfaktor v ergibt dann die nachstehende Ungleichung:

$$v \leq 1 + \frac{2\delta}{xlr - 3} \quad (11)$$

δ kann nun über die im Muster auftretende maximale Frequenzkomponente f_{\max} [Pixel⁻¹] – unabhängig von der Richtung²⁶ – abgeschätzt werden. Unter der Annahme, dass eine Intensitätsverschiebung um eine achteil Periode

²⁵ Bei der Verwendung von herkömmlichen Videokameras zur Bildaufnahme ([18]) und 'normalen' Szenen muss in der Regel auch bei guter Bildqualität von einer starken Korrelation der benachbarten Pixel ausgegangen werden.

²⁶ Die eindimensionale Grenzbedingung (Gl.11) wird nun als gute Annäherung für zweidimensionale Fälle betrachtet. Schrägdistanzen (xlr) bleiben also unberücksichtigt und führen somit eigentlich zu einer zu wenig harten Bedingung.

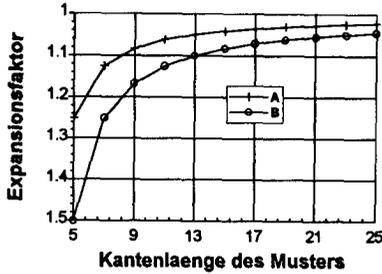


Fig. 18: Prinzipieller Verlauf der Grenze der erlaubten Expansion (Expansionsfaktor v) unter der Annahme einer maximalen Verschiebung δ um ein viertel ($f_{\max}=0.5\text{Pixel}^{-1}$, Kurve A) bzw. halbes Pixel ($f_{\max}=0.25\text{Pixel}^{-1}$, Kurve B) bei zentrischer Expansion (siehe auch Fig. 15).

(45°) von f_{\max} von den Korrelationsalgorithmen noch zu bewältigen ist²⁷, ergibt sich somit folgende Bedingung für die erlaubte Vergrößerung:

$$\delta = \frac{1}{8f_{\max}}$$

$$v \leq 1 + \frac{1}{4f_{\max} \cdot (xlr - 3)} \quad (12)$$

In Fig. 18 ist dieser Sachverhalt für $f_{\max}=0.5\text{Pixel}^{-1}$, die gemäss Abtasttheorem maximal mögliche Raumfrequenz, und $f_{\max}=0.25\text{Pixel}^{-1}$, eine etwas realistischere Grösse, grafisch in Abhängigkeit der Mustergösse xlr dargestellt. Wenn wir nun einen Blick vorauswerfen und die Detektionsratenverläufe im Anhang E.1: *Resultate, Grafiken* oder die Fig. 15 betrachten, so sehen wir, dass die Höhenlinien (Linien gleicher Detektionsraten) dieser Darstellungen für mittlere bis grosse Muster in ihrer Form sehr gut mit den Kurven in Fig. 18 übereinstimmen. Für kleine Muster gilt dieser Trend nicht, da der Informationsgehalt dann nur gering ist und die Detektionsrate

²⁷ Helava ([68]) zeigt, dass bei der unnormierten Kreuzkorrelation mit einer sinusförmigen Helligkeitsverteilung und einer 45° Verschiebung zwischen Muster- und Suchfenstersignal die entsprechende Korrelationsamplitude auf $\sqrt{2}/2$ des Ausgangswertes (0 Grad Verschiebung) fällt. Im oben besprochenen Fall wird aber erstens nur die Randinformation derart verschoben und zweitens liegen noch andere Frequenzanteile vor, die weniger stark betroffen sind. Darum ist ein entsprechend kleinerer Abfall zu erwarten.

wesentlich (negativ) beeinflusst. In [12], [14], [19], [102], [137]-[139], [150] und [163] wird dieselbe Problematik ebenfalls diskutiert.

Die Unterschiede der Detektionsratenverläufe zwischen den verschiedenen Korrelationskriterien werden durch die unterschiedliche Gewichtung der durch die geometrischen Verzerrungen entstandenen Pixelfehler erzeugt.

5.3. Diskussion der Detektionsratenverläufe

Die Diskussion der Leistungsfähigkeit der Algorithmen erfolgt hauptsächlich mit Hilfe der – im Vergleich der Verfahren übersichtlicheren, oben definierten – mittleren Detektionsraten, 80%-Mustergrößen (**80%G**), 80%-Bildstörungsstärken (**80%S**) sowie der optimalen Mustergrößen (**OG**) und Maximumstörungsstärken (**MS**). Die Aussagen können aber anhand der Fehlerhistogramme und besonders der Detektionsratenverläufe verifiziert und verfeinert werden.

Eine gute Robustheit gegenüber einer Störung macht sich durch grosse mittlere Detektionsraten, 80%- und Maximum-Störungsstärken bemerkbar. Kleine 80%-Mustergrößen deuten darauf hin, dass schon mit kleinen Referenzfenstern gute Resultate erreicht werden.

Tabelle 5 enthält die mittleren Detektionsraten und Tabelle 6 die oben besprochenen Kennzahlen, die die Schnittfläche des Detektionsratenverlaufs mit der 80%-Detektionsraten-Ebene charakterisieren (über alle Sequenzensets gemittelt; siehe Fig.14 und 15). Anhand dieser Parameter wird die Abhängigkeit der Detektionsrate von der Mustergrösse bzw. Störungsstärke diskutiert. -Im Anhang E.2: *Resultate, Tabellen* sind die entsprechenden Tabellen mit den Teilresultaten der einzelnen Sequenzensets beigeheftet.

In Tabelle 7 wird zusätzlich anhand der mittleren Detektionsraten eine störungsspezifische Rangierung der Verfahren aufgestellt. Zur Verdeutlichung der Situation sind die Felder mit den Rangierungen entsprechend schattiert worden. Insgesamt werden dadurch fünf, in der Regel gleich grosse, Rangierungskategorien unterschieden. Helle Felder kennzeichnen gute Rangierungen (kleine Rangwerte).

5.3.1. Korrelationskriterien ohne Bild-Vorverarbeitung

Das auffälligste Resultat der Versuche mit diesen Korrelationskriterien ist das äusserst schlechte Abschneiden der **SSD** und der **SAD** bei der **IRIS**-Störung! Bei allen anderen Störungsarten sind mit diesen Unähnlichkeitsmassen dagegen gute bis sehr gute Resultate zu verzeichnen. Die besten Rangierungen liefert die **KKFMF** (Tab.7). Sie zeigt zwar nur für den Bild-

Störung Alg.	IRIS	NOISE	SALTY	FOCUS	ZOOM
KKF	0.743	0.747	0.645	0.555	0.625
KKFMF	0.756	0.798	0.735	0.517	0.620
MOR	0.678	0.810	0.735	0.502	0.600
NSSDMF	0.422	0.794	0.701	0.434	0.586
SSD	0.263	0.779	0.668	0.616	0.642
SAD	0.187	0.773	0.878	0.551	0.658
SSDMF	0.703	0.754	0.647	0.567	0.620
SADMF	0.663	0.742	0.788	0.543	0.630
SSDMS	0.742	0.728	0.625	0.554	0.622
SADMS	0.721	0.723	0.792	0.525	0.633
SES1	0.724	0.785	0.794	0.415	0.552
SES2	0.739	0.793	0.822	0.436	0.600
SEK1	0.694	0.754	0.822	0.398	0.516
SEK2	0.708	0.766	0.839	0.420	0.559
UKKF	0.605	0.661	0.514	0.243	0.338
UKKFMF	0.618	0.677	0.532	0.259	0.375
NIS (w=2)	0.601	0.661	0.766	0.271	0.372
NIS (w=3)	0.654	0.692	0.670	0.331	0.413
NIS (w=4)	0.651	0.699	0.636	0.365	0.443
NIS (w=5)	0.644	0.693	0.622	0.390	0.444
NIS (w=6)	0.623	0.673	0.604	0.413	0.429
NIS (w=7)	0.616	0.662	0.588	0.440	0.402
NIS (w=8)	0.591	0.653	0.569	0.464	0.389
NA1	0.227	0.260	0.237	0.063	0.097
NA2	0.479	0.533	0.434	0.186	0.302

Tab.5: Übersicht über die mittleren Detektionsraten bei einem Toleranzkreisradius von $\sqrt{2}$ Pixeln.

störungstyp **IRIS** die absolut grösste mittlere Detektionsrate, eine schlechteste Platzierung von Rang neun, was nur noch von **MOR** annähernd erreicht wird, sowie die Tatsache, dass bei keinem Bildstörungstyp eine wirklich schlechte mittlere Detektionsrate vorliegt, erklärt aber das gute Gesamtergebnis dieses Ähnlichkeitsmasses.

Ohne Berücksichtigung von Quantisierungseffekten kann sehr einfach nachgewiesen werden, dass additive und multiplikative Helligkeitsverschiebungen keinen Einfluss auf die Korrelationswerte der **KKFMF** haben, die Funktion also invariant gegenüber diesen Transformationen ist ([13]). Die normale **KKF** erweist sich hingegen als sensitiv bezüglich additiven Helligkeitsänderungen. Dieser Einfluss ist aber beim Bildstörungstyp **IRIS** sehr klein, da das Schliessen einer Objektivblende als multiplikative Helligkeits-

Störung Alg.	80%G	IRIS 80%S	NOISE 80%S	SALTY 80%S	FOCUS 80%S	ZOOM		
						OG	MS	80%S
KKF	2.0	0.83	0.80	0.66	0.76	7	45.2	19.8
KKFMF	2.4	0.87	0.94	0.89	0.73	7	41.3	19.0
MOR	2.2	0.79	0.94	0.89	0.62	9	36.5	18.3
NSSDMF	2.2	0.36	0.89	0.75	0.44	11	35.7	18.3
SSD	2.0	0.08	0.84	0.66	0.83	5	51.1	19.8
SAD	2.0	0.06	0.83	1.00	0.71	5	49.5	19.8
SSDMF	2.2	0.81	0.82	0.65	0.77	7	43.2	19.8
SADMF	2.2	0.75	0.81	0.87	0.73	11	39.8	18.3
SSDMS	2.0	0.83	0.78	0.58	0.76	7	44.9	18.3
SADMS	2.0	0.79	0.76	0.87	0.72	7	44.0	19.0
SES1	2.0	0.86	0.86	0.90	0.44	7	29.7	17.2
SES2	2.0	0.85	0.88	0.92	0.47	5	41.1	18.3
SEK1	2.7	0.77	0.85	0.93	0.41	9	30.5	15.2
SEK2	2.7	0.80	0.84	0.93	0.44	9	32.7	18.3
UKKF	7.3	0.80	0.84	0.75	0.26	19	13.1	12.1
UKKFMF	6.8	0.81	0.86	0.75	0.31	15	15.3	12.3
NIS (w=2)	3.9	0.68	0.77	0.91	0.30	11	17.2	12.4
NIS (w=3)	4.5	0.77	0.83	0.81	0.39	11	24.2	15.1
NIS (w=4)	5.2	0.79	0.87	0.78	0.45	11	27.1	19.8
NIS (w=5)	5.3	0.84	0.88	0.75	0.49	13	24.8	20.3
NIS (w=6)	5.9	0.84	0.86	0.77	0.54	19	23.6	21.8
NIS (w=7)	6.2	0.84	0.86	0.76	0.59	17	21.6	20.3
NIS (w=8)	6.4	0.79	0.89	0.77	0.61	23	22.0	20.7
NA1	23.1	0.41	0.56	0.43	0.09	25	3.5	3.5
NA2	10.1	0.59	0.70	0.46	0.21	21	14.9	13.4

Tab.6: Übersicht über die 80%-Mustergrößen (80%G), die 80%-Bildstörungsstärken (80%S), die optimalen Mustergrößen (OG) und die Maximumstörungsstärken (MS) bei einem Toleranzkreisradius von $\sqrt{2}$ Pixeln. Sowohl MS als auch 80%S sind für den Störungstyp ZOOM in Prozenten angegeben (Größenveränderung).

reduktion modelliert werden muss. Diese Aussage wird auch durch den Vergleich der Resultate der mittelwertskalierten mit denjenigen der mittelwertbefreiten Differenzmasse gestützt. Die mittelwertskalierten Verfahren weisen beim Störungstyp IRIS in der Regel eine bessere mittlere Detektionsrate auf. Dieser Effekt manifestiert sich aber erst bei kleinen Mustergrößen mit gut strukturiertem Bildinhalt, wo ohnehin eine reduzierte Detektionsrate auftritt. Er wirkt sich daher kaum auf die 80%S-Werte aus.

Die höhere mittlere Detektionsrate der KKF gegenüber der KKFMF beim Bildstörungstyp FOCUS weist auf ein Problem hin, das besonders bei klei-

Störung Alg.	IRIS	NOISE	SALTY	FOCUS	ZOOM
KKF	2	11	15	3	5
KKFMF	1	2	9	8	7
MOR	10	1	9	9	9
NSSDMF	2	3	11	13	11
SSD	3	5	13	1	2
SAD	5	6	1	5	1
SSDMF	8	9	14	2	7
SADMF	11	12	7	6	4
SSDMS	3	13	17	4	6
SADMS	6	14	6	7	3
SES1	5	8	5	15	13
SES2	4	4	3	12	9
SEK1	9	9	3	17	14
SEK2	7	7	2	14	12
UKKF	18	21	23	23	23
UKKFMF	16	18	22	22	21
NIS (w=2)	19	21	8	21	22
NIS (w=3)	12	17	12	20	18
NIS (w=4)	13	15	16	19	16
NIS (w=5)	14	16	18	18	15
NIS (w=6)	15	19	19	16	17
NIS (w=7)	17	20	20	11	19
NIS (w=8)	20	21	21	10	20
NA1	24	25	25	23	25
NA2	21	24	24	24	24

Tab.7: Übersicht über die Rangierungen anhand der mittleren Detektionsraten bei einem Toleranzkreisradius von $\sqrt{2}$ Pixeln.

nen Referenzfenstern im Zusammenhang mit der lokalen Mittelwertbefreiung auftreten kann. Muster, die nur aufgrund der unterschiedlichen lokalen mittleren Helligkeit auseinanderzuhalten sind, können nicht mehr unterschieden werden. Das gleiche Phänomen kann auch bei der **SSDMF** sowie, in geringerem Ausmass, bei der **SADMF** beobachtet werden. Der Hochpasseffekt der Mittelwertbefreiung reduziert die verfügbare Informationsbandbreite speziell bei unscharfen Bildern merklich und vermindert so die entsprechenden Detektionsraten.

Bei rauschartigen Störungen liefert die **KKF** zwar immer noch gute, verglichen mit den ähnlich komplexen, jedoch lokal mittelwertbefreiten Verfahren wie **KKFMF**, **MOR** und **NSSDMF** aber dennoch deutlich schlech-

tere Resultate. Wie auch der 80%-Störungswert zeigt, ist dies hauptsächlich eine Folge der geringeren Robustheit gegenüber diesen Störungstypen. Die Mustergrösse scheint auf das Ergebnis keinen Einfluss zu haben.

Die Korrelation von Moravec erreicht Detektionsraten (für die verschiedenen Bildstörungstypen), die nahe den Werten der **KKFMF** liegen. Nur beim Bildstörungstyp **IRIS** ist die entsprechende Differenz etwas grösser, da dieser Algorithmus nicht robust gegenüber multiplikativen Helligkeitsänderungen ist.

Die schlechten Resultate der **NSSDMF** bei abgedunkelten und unscharfen Bildern (**IRIS** und **FOCUS**) müssen vom Nenner dieser Funktion verursacht werden, da der Zähler der einfachen **SSDMF** entspricht, die ja in der gleichen Situation recht gute Ergebnisse liefert. Weil der Musterterm für alle Objektfenster konstant bleibt, kann im Nenner der **NSSDMF** wiederum nur der Suchfensterterm einen Einfluss auf die Position des Extremas haben. Ein falsches Korrelationsminimum (Unähnlichkeitsmass!) ergibt sich also durch einen zu grossen Suchfensterterm, der eng verwandt mit der Objektfenstervarianz ist. Eine grosse Varianz ergibt sich genau dann, wenn kontrastreiche Objektfenster im Suchfenster vorhanden sind.

Aus diesen Betrachtungen folgt, dass die **NSSDMF** nicht als Korrelationsmass geeignet ist, falls ein schwach strukturiertes Muster in einem Suchfenster gesucht wird, das kontrastreichere Objektfenster enthält! -Dieser negative Effekt tritt nun hauptsächlich bei Störungsarten, die die Objektfenstervarianzen verringern, wie **IRIS**, **FOCUS** und **ZOOM** in Erscheinung, da dann die zusätzlich reduzierte Varianz am Ort des wahren Objektfensters den Unähnlichkeitswert nochmals erhöht, während die (ebenfalls etwas reduzierte) Varianz eines Objektfensters mit kräftigem Kontrast den **NSSDMF**-Korrelationswert gegenüber dem **SSDMF**-Wert sehr stark mindert.

Bei den Störungen **NOISE** und **SALTY** sind die mittleren Detektionsraten der **NSSDMF** grösser als diejenige der **SSDMF**, weil dann genau der gegenteilige Trend eintritt. Die Helligkeitsvarianzen der Objektfenster mit, gegenüber der Störung geringem Kontrast werden in Richtung der Störvarianzen verschoben (erhöht, die **NSSDMF**-Unähnlichkeitswerte werden also reduziert). Bei Objektfenstern mit grossem Kontrast werden die Helligkeitsvarianzen reduziert und damit die Unähnlichkeitswerte vergrössert. Die stark kontrasthaltigen Objektfenster werden also nicht mehr bevorteilt.

Schliesslich bleibt nur noch die Frage zu beantworten: 'Wieso ist der gleiche Effekt nicht auch bei der **KKFMF**, die ja den gleichen Nenner wie die **NSSDMF** besitzt, zu beobachten?'. Der wesentliche Unterschied besteht darin, dass die **KKFMF**, im Gegensatz zur **NSSDMF**, ein Ähnlichkeitsmass

darstellt und damit der Nenner genau umgekehrt wirkt. Kontrastreiche Objektfenster verursachen also einen grossen Nenner und reduzieren damit den Korrelationswert (reduzieren die Ähnlichkeit). Diese Reduktion wird aber durch den Zähler wieder kompensiert, der bei guter Übereinstimmung zwischen Muster und Objektfenster, einen Wert annimmt, der wegen der lokalen Mittelwertbefreiung ebenfalls mit der Signalvarianz ansteigt.

Die vorangehenden Überlegungen zeigen, dass es sich bei der **NSSDMF** um ein ungünstiges Korrelationsmass handelt, das für gute Leistungen zusätzliche Bedingungen an die Szene stellt. Da der Rechenaufwand, wie wir in Kapitel 6: *Rechenaufwand* sehen werden, ebenfalls nicht tiefer liegt als bei den Vergleichsverfahren, kann dieses Korrelationskriterium nicht empfohlen werden!

Die **SSD** und die **SAD** liefern unabhängig von der Mustergrösse selbst bei geringen Bildstörungen vom Typ **IRIS** völlig ungenügende Resultate. Globale, multiplikative Helligkeitsverschiebungen beeinflussen die Korrelationswerte in unterschiedlicher Weise und können damit die Position des Extremums verschieben ([13]). Bei allen anderen Bildstörungstypen erreichen die beiden Algorithmen sehr gute Detektionsraten! Die **SAD** besitzt beim Bildstörungstyp **SALTY** sogar eine Detektionsrate die nur von der **SEK2** annähernd erreicht wird. Das bessere Abschneiden der **SAD** verglichen mit der **SSD** rührt offenbar daher, dass die durch die 'Salz- und Pfefferkörner' verursachten grossen Differenzen zwischen Muster- und Korrelationsfensterpixel nicht quadratisch gewichtet werden und somit die Gesamtsumme nicht dramatisch beeinflussen. Die Resultatreihenfolge dieser Differenzmasse bei Rauschstörungen wird z.B. auch durch die Untersuchungen von Hinrichs ([70]) gestützt.

Eine Verbesserung der Situation beim Bildstörungstyp **IRIS** kann durch lokale Mittelwertbefreiung bzw. Mittelwertskalierung der ursprünglichen Differenzmasse erreicht werden. Dies geht in der Regel allerdings mit einer moderaten Verschlechterung – hinsichtlich des Detektionsratenverlaufs bezüglich der Störungsstärke – bei den anderen Bildstörungstypen einher (siehe 80%-Störungswerte).

Von der Theorie her ([138] und [139]) müssten die **SSDMF** und die **KKFMF** bei verrauschten Bildern die gleichen Resultate erzeugen, da beide in diesem Fall den Maximum Likelihood Korrelator realisieren. Die Unterschiede zu Gunsten der **KKFMF** sind nur durch unterschiedliches Verhalten bei Sättigungs- und Quantisierungseffekten sowie durch die Verwendung von kleinen, die statistischen Annahmen nicht erfüllenden Mustern zu erklären. Dieses Ergebnis zeigt sehr schön die beschränkte Aussagekraft von entsprechenden theoretischen Untersuchungen. Vielfach wird, damit die Theo-

rie handhabbar wird, von Stör- und Signalcharakteristiken ausgegangen, die realitätsfremd sind.

Bei der Betrachtung der Resultate stellt sich natürlich eine weitere Frage: 'Wieso schneiden die lokal mittelwertbefreiten Differenzmasse bei der **IRIS**-Störung erheblich besser ab, als die entsprechenden ursprünglichen Verfahren?'. Die Mittelwertkompensation kann ja nur additive Helligkeitsverschiebungen ausgleichen und der Störungstyp **IRIS** stellt zur Hauptsache eine multiplikative Intensitätsveränderung dar (siehe Kapitel 4.2: *Generierung der Bildsequenzen*). Zudem besteht der wesentlichste Unterschied der **SSD** bzw. der **SAD** im Vergleich zur **SSDMF** respektive **SADMF** darin, dass die unkompensierten Algorithmen nur robust und die lokal mittelwertbefreiten Differenzmasse invariant gegenüber additiven Helligkeitsänderungen sind. Beide haben also die Eigenschaft die Position des Minimalwertes bei idealer additiver Störung nicht zu verschieben. Um nun die Diskrepanz der Resultate bei multiplikativen Helligkeitsänderungen zu verstehen, muss man sich vor Augen führen, dass die **SSDMF** wie auch die **SADMF** den lokalen Helligkeitsmittelwert in der Rechnung berücksichtigen. Dieser nimmt bei einer multiplikativen Störung, je nach ursprünglicher, lokaler Helligkeit, an jeder Korrelationsfensterposition in der Regel einen anderen Wert an. Die lokale Mittelwertkompensation stellt darum eine bessere Annäherung an die tatsächlichen Verhältnisse bei multiplikativen Intensitätsverschiebungen dar, als die Annahme, dass die Mittelwertänderung über das ganze Suchfenster konstant ist. Genau diese Forderung besteht aber für die Robustheit der **SSD** und der **SAD** und wird nur sehr schlecht erfüllt.

Die **SSD** und die **SAD** zeigen besonders auch bei geometrischen Verzerrungen ein ausgezeichnetes Verhalten. Die optimale Mustergrösse (**OG**) ist mit 5x5 Pixeln sehr klein und garantiert noch bei einer maximalen Verzerrung (**MS**) von ca. 50 Prozent (Expansionsfaktor $\cong 1.5$) eine 80 prozentige Detektionsrate!

5.3.2. Korrelationskriterien mit Richtungsfiltern

Besonders interessant ist natürlich der Vergleich der Sobeloperator-basierten Verfahren (**SES1**, **SES2**) mit den neu eingeführten Verfahren, die den Kirschoperator verwenden (**SEK1**, **SEK2**). Der Sobeloperator liefert meist die besseren Resultate. Selbst eine Reduktion der Auflösung des Sobeloperators von 32 auf 8 Richtungen, wie beim Kirschoperator, lässt diese Reihenfolge, wenn auch in abgeschwächter Form, bestehen! Wie beim Kirschoperator wird durch diese Auflösungsverminderung aber einerseits die Robustheit gegenüber den Störungen leicht verringert und andererseits die

Detektionsraten für sehr kleine Muster auf unakzeptable Werte gesetzt. Dies hängt damit zusammen, dass kleine Muster bei einer derartig geringen Richtungsauflösung nur noch sehr wenig Information tragen. Ein Referenzfenster von 2×2 Pixel Grösse kann z.B. nur noch $8^4 = 4096$ unterschiedliche Richtungsmuster enthalten!

Im Gegensatz zu allen anderen untersuchten Störungsarten liefert der Kirschoperator bei 'Salt and Pepper' Rauschen durchwegs bessere mittlere Detektionsraten als der Sobeloperator. Obwohl auch die Sobeloperator-basierten Verfahren ausserordentlich robust gegenüber dieser Störung sind, erweisen sich die **SEK1**- und die **SEK2**-Korrelation, besonders wenn man noch die oben erwähnten Probleme bei kleinen Mustern berücksichtigt, als noch erfolgreicher! Ausser bei maximaler Störungsstärke bzw. minimaler Mustergösse egalisieren oder übertreffen sie sogar die **SAD**, die ja in diesem Störfall die beste mittlere Detektionsrate aufweist (nicht aus den **80%S**-Werten ablesbar!).

Bei unscharfen Bildern muss bei allen Verfahren, die auf richtungscodierten Bildern arbeiten eine deutliche Reduktion der mittleren Detektionsrate in Kauf genommen werden. Durch die Reduktion der Informationsbandbreite wird das Augenmerk von der lokalen auf die eher 'regionale' Orientierung verlegt. Im Extremfall wird überhaupt keine lokale Richtung mehr bestimmbar sein, da alle Pixel der betrachteten 3×3 Pixel grossen Umgebung dieselbe Helligkeit aufweisen. Bei den untersuchten Implementationen wird dieser Fall durch Zuweisung einer Standardrichtung gelöst, welche natürlich keinerlei Relation zur tatsächlichen, ungestörten Richtung besitzt. Eine bessere Lösung wäre z.B., bei Helligkeitsgradienten mit Beträgen unter einer gewissen Schwelle eine Richtung senkrecht zur Bildebene zuzuweisen; diese schlösse mit allen anderen Richtungen den gleichen kleinsten Winkel ein (90°)²⁸. Ein Fehler, der durch schlecht abgesicherte Richtungen entsteht, würde somit das Korrelationsmass in genau definierter Weise beeinflussen. Bei der Winkel-**SSD** würde sich diese Massnahme, unter der Annahme einer Gleichverteilung der durch die unsicheren Richtungswerte verursachten Richtungsdifferenzen, in einer Reduktion des Unähnlichkeitswertes niederschlagen. Bei der Winkel-**SAD** hingegen blieben die Korrelationswerte im Rahmen der statistischen Annahme konstant²⁹.

²⁸ Selbstverständlich könnte in diesem Fall die (unsichere) Richtungsdifferenz auch einfach unberücksichtigt belassen werden.

²⁹ Winkel-**SSD**: Die Summe der Quadrate der gleichverteilten (unsicheren) Winkeldifferenzen ($0..180^\circ$) ist grösser als die Summe der Quadrate der neu eingeführten 90° -Winkeldifferenzen.

Der Bildstörungstyp **ZOOM** verursacht ebenfalls eine gewisse, allerdings erstaunlich geringe Reduktion der mittleren Detektionsraten im Vergleich zu den robustesten Korrelationskriterien. Besonders entlang von Helligkeitskanten können entsprechende, zur erfolgreichen Detektion wichtige Richtungselemente von Muster und Objektfenster nicht mehr zur Deckung gebracht werden. Bei direkter Verwendung der Helligkeitsinformation erweist sich dieselbe Situation als weniger kritisch, da auch relativ unstrukturierte Flächen zur Korrelation beitragen.

Die Verwendung der Winkel-**SAD** anstelle der Winkel-**SSD** führt, unabhängig vom Störungstyp, zu besseren mittleren Detektionsraten. Die quadratische Gewichtung der Richtungsfehler hat in allen untersuchten Fällen also eine Verminderung der Detektionsrate zur Folge. Bei der normalen Intensitäts-**SSD** und -**SAD** kann kein entsprechender Trend beobachtet werden.

5.3.3. Korrelationskriterien mit Bild-Hochpassfilterung

Hochpass-gefilterte Bilder ergeben bei der Korrelation ähnliche Schwierigkeiten wie richtungscodierte³⁰. Bei unscharfen oder geometrisch verzerrten Bildern ergeben darum die **UKKF** und die **UKKFMF** ebenfalls schlechte Resultate. Bei den anderen Störungstypen fällt der Unterschied zu den erfolgreichereren Korrelationskriterien nicht ganz so krass aus. Der 80%-Störungswert weist darauf hin, dass die betrachteten Korrelationsalgorithmen über eine ähnliche Robustheit verfügen, wie die Vergleichsverfahren. Die 80%-Mustergrößen zeigen allerdings, dass dies nur für mittlere bis grosse Muster zutrifft, wo eine genügende Anzahl von Kantenpixeln vorliegen.

Bei unscharfen Bildern (**FOCUS**) wird die Leistung der **UKKF** und der **UKKFMF** dadurch reduziert, dass die Bildinformation in den tieferen Raumfrequenzen durch das Hochpassfilter in ihrer Bedeutung weiter gemindert wird (die hohen Frequenzen sind bei unscharfen Bildern ja schon reduziert worden; Grenzfrequenzen: $1/3$, $1/4$, .. $1/15$ Pixel⁻¹). -Geometrisch verzerrte Bilder (**ZOOM**) werden nur schlecht verarbeitet, da die Kanteninformation von Muster und Objektfenster nicht zur Deckung gebracht werden kann.

Die **UKKF** als auch die **UKKFMF** ist, wie sehr einfach einzusehen ist falls Quantisierungseffekte unberücksichtigt bleiben, in der Theorie robust

Winkel-**SAD**: Die Betragssumme der gleichverteilten (unsicheren) Winkeldifferenzen (0..180°) ist im Mittel gleich der Summe der neu eingeführten 90°-Winkeldifferenzen.

³⁰ Die Richtungsbestimmung stellt ja im Prinzip auch eine Hochpassfilterung dar.

gegenüber additiven und multiplikativen Helligkeitsänderungen³¹. Multiplikative Änderungen der Intensitäten rufen nur eine ebensolche multiplikative Änderung der Korrelationswerte hervor. Die Position des Extremwerts wird also nicht verschoben. Trotz diesen theoretisch herleitbaren Eigenschaften weisen beide Korrelationskriterien im Gegensatz zu den besten Verfahren eine geringere mittlere Detektionsrate auf. Dies ist hauptsächlich durch den starken Abfall der Detektionsrate mit abnehmender Mustergrösse zu erklären (**80%G**). Durch die Hochpassfilterung wird dem Muster Information entzogen, die besonders bei kleinen Mustern für eine eindeutige Zuordnung zum richtigen Objektfenster fehlt. Für grössere Muster ist die experimentell ermittelte, gute Robustheit gegenüber der **IRIS**-Störung aber in Übereinstimmung mit der Theorie³².

Erstaunlich gut schneiden die beiden Korrelationskriterien bei rauschartigen Störungen ab (**NOISE** und **SALTY**). Besonders für Gauss'sches Helligkeitsrauschen ist nur wegen der generell geringeren Leistungsfähigkeit bei kleinen Mustern eine gegenüber den besten Korrelationskriterien reduzierte mittlere Detektionsrate zu beobachten. Für grössere Muster (ca. 16x16 Pixel) darf bei Störungen vom Typ **IRIS**, **NOISE** und **SALTY** mit einer guten Robustheit gerechnet werden!

Vorversuche mit der **KKF** haben gezeigt, dass die Laplacefilterung des Bildmaterials, bei geringen bis mittelstarken Störungen vom Typ **NOISE** und **SALTY**, auch in diesem Fall die Detektionsraten verbessert! Die befürchtete Verschlechterung der Detektionsrate, wegen der Ableitung des verrauschten Signals, bleibt also weitgehend aus bzw. wird erst für grosse Störungsstärken relevant. Der Informationsgehalt der hohen Frequenzanteile ist demnach im erwähnten Störstärkenbereich noch genügend gross und lässt, zur Verbesserung der Detektionsleistung, sogar noch eine Hervorhebung dieser verrauschten Komponenten zu.

5.3.4. Binäre Korrelationskriterien

Bei den binären Korrelationskriterien erscheint der Algorithmus von Nishihara in jeder Beziehung deutlich leistungsfähiger als derjenige von Nack. Auch die oben diskutierten Verfahren, die auf Hochpass-gefilterten Daten arbeiten, weisen, wegen den schlechten Detektionsraten bei kleinen Mustern (**80%-Mustergrösse**), teilweise schlechtere Resultate als Nishiharas binäres Verfahren auf. In Tabelle 6 wird aber ein Problem von Nishiharas

³¹ Da ein mittelwertfreier Laplaceoperator verwendet wird, ist die **UKKF** und die **UKKFMF** sogar invariant gegenüber einer additiven Helligkeitsänderung.

³² Bei maximaler Störung ist die Detektionsrate aber ebenfalls vergleichsweise gering.

Korrelationskriterium besonders deutlich: Je nach Bildstörungstyp muss der w -Parameter des Laplace-Gauss-Filters angepasst werden, um optimale Leistung zu erreichen. Dies bedeutet, dass bei Anwendungen, die auf bestmögliche Detektionsraten bei unterschiedlichen Bedingungen angewiesen sind, eine adaptive Nachführung des w -Parameters nötig sein kann.

Der Laplace-Gauss-Operator in Nishiharas Korrelationsalgorithmus erweist sich in der Theorie als robust gegenüber multiplikativen Helligkeitsänderungen. Das Vorzeichen des Filterausgangs, das ja zur Binarisierung benutzt wird, wird durch eine derartige Störung nicht geändert, d.h. das Binärbild bleibt unbeeinflusst. Der Korrelationsalgorithmus ist also schlussendlich invariant gegenüber multiplikativen Helligkeitsänderungen.

Nishiharas Korrelationsalgorithmus zeigt bei Gauss'schem wie auch 'Salt and Pepper' Rauschen gute mittlere Detektionsraten. Das Verhalten bezüglich unterschiedlichen Vorverarbeitungsfiltren (w -Parameter des Laplace-Gauss-Operators) fällt aber bei den beiden Störungsarten nicht genau gleich aus. Bei Gauss'schem Rauschen (**NOISE**) nimmt die mittlere Detektionsrate, im Gegensatz zum Verhalten bei 'Salt and Pepper' Rauschen, zuerst mit verstärkter Tiefpasscharakteristik des Laplace-Gauss-Filters zu (w wird grösser); hochfrequente Signalanteile mit ungünstigerem Signal-zu-Rauschleistungsverhältnis werden etwas besser unterdrückt. Später, ungefähr bei $w=5$, tritt ein, bei 'Salt and Pepper' Rauschen überall absolut dominanter, gegenläufiger Trend ein. Die stärkere Tiefpasswirkung hinterlässt bei kleinen Mustern keine wieder auffindbare, lokale Struktur (**80%G** nimmt zusammen mit w zu). Die Robustheit gegenüber der Rauschstörung (**80%S**) bleibt hingegen für alle grossen Werte von w ungefähr konstant.

Die Robustheit von Nishiharas Korrelationsalgorithmus wird bei 'Salt and Pepper' Rauschen (**SALTY**) ausserordentlich stark erhöht, wenn der Filterparameter w auf zwei gesetzt wird (ca. 14% grössere mittlere Detektionsrate als bei $w=3!$). Dieses Verhalten ist erstaunlich, kann doch gerade diese Störung zum Verstoß gegen das Abtasttheorem führen³³. Das Allpassfilter dämpft, im Gegensatz zu den sonst üblichen Tiefpässen ($w > 2$), die hohen und darum bezüglich des angesprochenen Aliasing-Problems besonders gefährdeten Frequenzanteile nur wenig. Diese leichte Reduktion zusammen mit der etwas stärkeren Gewichtung der zur genauen Lokalisierung wichtigen

³³ Die Elimination der Bildinformation an einzelnen Stützstellen reduziert die Bandbreite des beobachtbaren Spektrums. Dies ist anhand des Beispiels mit einer Gleichverteilung der informationslosen Stützstellen leicht einzusehen. Die reale Abtastung erfolgt dann nur noch mit entsprechend reduzierter Rate. -Bei einer Störrate von 50% beträgt der mittlere Abstand zwischen zwei gestörten Pixeln $\sqrt{2}$ Pixel. Dies ergibt theoretisch eine Grenzfrequenz von $1/(2\sqrt{2}) \approx 0.35 \text{ Pixel}^{-1}$ für das Abtasttheorem.

und vom Aliasing nicht betroffenen Spektralanteile um 0.25 Pixel^{-1} genügt offenbar, um eine deutliche Verbesserung der Robustheit zu erreichen.

Der sonst übliche Effekt, dass bei kleinen Mustern mit 'Salt and Pepper' Rauschen ein Zwischenmaxima im Detektionsratenverlauf auftritt, kann bei Nishiharas Korrelationsalgorithmus nicht festgestellt werden, da die Störinformation vor der eigentlichen Korrelation durch das Laplace-Gauss-Filter auf eine weite Nachbarschaft des veränderten Pixels verteilt wird. Ähnliches gilt im übrigen auch für alle anderen Korrelationsverfahren, die Vorverarbeitungsfilter verwenden.

Für die Störung **IRIS** kann eine zum Verhalten bei Gauss'schem Rauschen ähnliche Tendenz bezüglich der verschiedenen Vorverarbeitungsfilter beobachtet werden. Auch hier muss die anfängliche Zunahme der mittleren Detektionsrate mit wachsendem w -Parameter durch die bessere Robustheit des gesamten Korrelationsalgorithmus' gegenüber der Störung erklärt werden (**80%S**). Die Abhängigkeit von der 80%-Mustergrösse ist genau entgegengesetzt und dominiert schliesslich für grössere w -Werte.

Die Resultate beim Bildstörungstyp **FOCUS** zeigen, dass dieses Verfahren auf scharfe Bilder angewiesen ist. Mit abnehmendem w -Wert wird die Robustheit des Korrelationsalgorithmus' gegenüber dieser Störung reduziert. Für grosse Muster ist ein markanter Abfall der Detektionsrate ungefähr dann zu erwarten, wenn die Grenzfrequenz des Bildinhalts unter die Frequenz minimaler Dämpfung des Laplace-Gauss-Operators fällt (siehe auch Detektionsratenverläufe im Anhang E.1: *Resultate, Grafiken*)³⁴.

Beim Bildstörungstyp **ZOOM** hat dieses Verfahren gegenüber dem verwandten, im Bereich der Stereo Vision klassischen, hier aber nicht untersuchten Marr-Poggio-Korrelator ([109]) den Vorteil, dass durch die Vorzeichen-Binarisierung des Filterresultats (anstelle der Bestimmung der Null-durchgangsposition) nicht Kanten sondern vielmehr, vereinfacht dargestellt, von den Kanten umschlossene Flächen hervorgehoben werden (Anhang D: *Wirkung des Laplace-Gauss-Operators*). Eine kleine Verzerrung des Bildes resultiert in einer grossen Diskrepanz im Kantenbild (Kanten kommen im Muster und im Objektfenster nicht mehr aufeinander zu liegen), während die entsprechenden Flächen noch mehrheitlich übereinstimmen. Für kleine w -Werte, bei denen das Filter die hohen Signalfrequenzen nur schwach dämpft, sind darum die Resultate schlechter als bei mittelgrosser w -Zahl. Für grosse

³⁴ Ungefähre Frequenz minimaler Dämpfung beim Laplace-Gauss-Operator für $w=2..8$: 0.25, 0.17, 0.11, 0.09, 0.08, 0.07, 0.06 Pixel^{-1} . Grenzfrequenz des Bildinhalts für die **FOCUS**-Bildsequenzen (Bildnummer 0.8): 0.5, 0.33, 0.25, 0.2, 0.17, 0.14, 0.13, 0.11, 0.1 Pixel^{-1} .

w -Parameter ist ebenfalls eine Abnahme der Detektionsraten zu beobachten, die der fehlenden Struktur in kleinen Mustern zuzuschreiben ist³⁵.

Beide von Nack inspirierten binären Korrelationen (**NA1** und **NA2**) sind bezüglich der Detektionsrate regelmässig auf den hintersten Plätzen rangiert. Die Modifikation des Original-Algorithmus' (**NA2**) ergibt aber doch eine deutliche Verbesserung der Situation. Bei scharfen und geometrisch unverzerrten Bildern ist die Detektionsrate für mittlere bis grosse Muster wenn nicht gut, so doch wenigstens akzeptabel. Bei sehr grossen Mustern kann mit einer guten Leistung gerechnet werden ([122] und [156]). Für mittlere Mustergrössen könnte eventuell durch Erhöhung des Sollwerts der Anzahl Kantenpixel, die von der Vorverarbeitung detektiert werden sollen, eine zusätzliche Verbesserung erreicht werden.

Bei unscharfen Bildern versagt der Kantendetektionsalgorithmus während bei verzerrten Bildern die Position der Kantenpixel derart verschoben wird, dass keine Deckungsähnlichkeit zwischen Muster und Objektfenster in den binären Kantenbildern erreicht werden kann. Da die Kantendetektion adaptiv arbeitet – einen fixen Anteil von Kantenpixeln garantiert – sind diese zwei Verfahren, falls von Quantisierungsfehlern abgesehen wird, bei grossen Mustern robust gegenüber allen linearen Helligkeitstransformationen.

In den Detektionsratenverläufen von **NA1** ist im übrigen bemerkenswert, dass die Detektionsraten bei grossen Mustern und Bildstörungstyp **SALTY** wie auch **NOISE** für die unveränderten Bilder schlechter sind als für die leicht verrauschten. Dies rührt von der bereits bei der Definition dieser Korrelationsfunktion beschriebenen Tatsache her, dass das Ähnlichkeitsmass Objektfenster mit wenigen, aber allesamt richtig gesetzten Pixeln gegenüber solchen mit vielen, aber teilweise falsch gesetzten Pixeln bevorzugt. D.h. ein Korrelationsfenster mit nur einem einzigen an der richtigen Stelle gesetzten Kantenpixel (Muster enthält an der gleichen Stelle ebenfalls ein gesetztes Pixel) erreicht eine Korrelation mit dem Wert Eins. Diese Objektposition ist aber aufgrund des Korrelationswerts nicht mehr von der wahren Objektposition, die im unverrauschten Fall ebenfalls den Korrelationswert Eins liefert, unterscheidbar.

Beide Verfahren, **NA1** und **NA2**, sollten aufgrund dieser Ergebnisse wenn möglich nur mit Mustergrössen von über 25x25 Pixeln verwendet werden! In diesem Bereich wird aber natürlich auch Rechenaufwand wieder

³⁵ Das Phänomen, dass für kleine Muster keine verwertbare Struktur mehr auftritt, würde auch eine abnehmende Leistung beim Bildstörungstyp **FOCUS** erwarten lassen. Dort wird aber die stetige Zunahme der mittleren Detektionsrate mit wachsenden w -Werten hauptsächlich durch mittlere und grosse Muster, die zuverlässiger aufgefunden werden, realisiert. Kleinste Muster werden also ebenfalls nicht mehr erkannt.

erheblich grösser. Der Vorteil der einfachen, binären Verarbeitung wird also mindestens zum Teil wieder eliminiert!

5.3.5. Zusammenfassung der Resultate

Die ausgeglichene Leistung über alle untersuchten Bildstörungstypen erreicht, wie z.B. aus den Rangierungen in Tabelle 7 hervorgeht, die **KKFMF** gefolgt vom Ähnlichkeitsmass von Moravec, der **KKF** und den mittelwertbefreiten bzw. mittelwertskalierten Differenzmassen.

Die **SSD** und die **SAD**, die beide in sehr vielen Applikationen recht erfolgreich verwendet werden, sind ungeeignet, falls zwischen dem Muster und dem Suchfenster, selbst in kleinem Ausmass, multiplikative Helligkeitsverschiebungen, wie sie z.B. bei Beleuchtungsänderungen auftreten, vorliegen. Dasselbe, in leicht abgeschwächter Form, gilt auch für die **NSSDMF**, die zusätzlich den Nachteil hat, dass sie auch bei unscharfen Bildern schlechte Resultate liefert.

Bei den Korrelationsverfahren, die mit richtungscodierten Bildern arbeiten, ist der Sobeloperator dem Kirschoperator, ausser bei 'Salt and Pepper' Rauschen, vorzuziehen. Die Reduktion der Richtungsauflösung des Sobeloperators (z.B. von 32 auf 8) geht mit einer leichten Verminderung der Robustheit gegenüber den Störungen und, wie beim Kirschfilter, mit unakzeptablen Detektionsraten bei sehr kleinen Mustern einher³⁶.

Mit diesen Verfahren dürfen, besonders bei verrauschten (**NOISE** und **SALTY**) oder in ihrer Helligkeit multiplikativ veränderten Bildern (**IRIS**), schon mit verhältnismässig kleinen Mustern gute bis sehr gute Resultate erwartet werden.

Die **UKKF** und die **UKKFMF** können nur solange empfohlen werden als scharfe und geometrisch unverzerrte Bilder vorliegen. Zudem muss unbedingt mit grösseren Mustern gearbeitet werden (grösser als ca. 16x16 Pixel). Die **UKKFMF** weist immer eine etwas bessere mittlere Detektionsrate auf. Dies ist sowohl durch die geringfügig bessere Robustheit gegenüber den Störungen, als auch durch die etwas kleineren Ansprüche an die Mustergrösse zu erklären.

Nishiharas Korrelationsalgorithmus weist für mittlere bis grosse Muster eine gute Robustheit auf, die nur bei unscharfen oder geometrisch verzerrten Bildern deutlich unter derjenigen der besten Korrelationskriterien liegt.

³⁶ Ausser bei 'Salt and Pepper' Rauschen ist der Sobeloperator aber auch bei dieser Auflösungsreduktion dem Kirschoperator vorzuziehen.

Die schlechtesten Resultate liefern die beiden Implementationsvarianten von Nacks Verfahren. Die Modifikation des Originalalgorithmus' verbessert das Verhalten zwar wesentlich, trotzdem sind aber immer grosse Muster notwendig, um befriedigende Resultate zu erreichen. Wie bei allen Algorithmen, die auf gefilterten Bildern arbeiten, muss auch hier mit besonders grossen Leistungseinbussen bei unscharfen oder geometrisch verzerrten Bildern gerechnet werden.

Die Auswertung der Resultate getrennt nach Sequenzenset, und damit unterschieden nach Mustertypus (texturhafte, kantenhafte bzw. ausgewählte, kontrastreiche Muster), ergibt bezüglich der mittleren Detektionsraten kein einheitliches Bild. Die entsprechenden Rangierungen weisen aber auf ein prinzipiell sehr ähnliches Verhalten bei unterschiedlich gearteter Bildinformation hin.

Allgemein kann festgestellt werden, dass bei Sequenzenset 2 und 3 (**EDGES** und **CHAOS**) bessere mittlere Detektionsraten und charakteristische Werte (**80%S** etc.) auftreten als bei Sequenzenset 1 (**TEXTURES**). Dieses Phänomen ist aber eher auf die in Set 1 z.T. sehr schwache Textur als auf den Bildtyp zurückzuführen. Der Schluss, dass teksturhafte Muster schlechter zu detektieren sind, ist darum nicht zulässig und in seiner Tendenz sogar falsch ([52])! Die Korrelationskriterien, wie sie hier untersucht werden, sind speziell für diesen Fall sehr gut geeignet, da die Eindeutigkeit des Resultats mit Texturinformation besser gewährleistet werden kann als mit (nicht allzu dichter) Kanteninformation.

Aus diesen Beobachtungen wird ersichtlich, dass der Bildinhalt die absolute Leistungsfähigkeit der Korrelationskriterien ganz wesentlich beeinflusst (eine Binsenweisheit). Die relative Leistungsfähigkeit, unabhängig von der Art des Bildinhalts, bleibt aber in gewissen Grenzen erhalten. Dies führt zum Schluss, dass eine absolute Leistungsvorhersage nur mit genauer Kenntnis der zu erwartenden Szenen möglich ist. Eine relative Voraussage scheint aber mit guter Zuverlässigkeit, selbst mit geringen Vorkenntnissen, gangbar.

Die Resultate, aufgespalten in die Teilresultate der verschiedenen Sequenzensets, sind im Anhang E.2: *Resultate, Tabellen* beigeheftet.

6. Rechenaufwand

6.1. Allgemeines

Der Rechenbedarf eines (Un-) Ähnlichkeitsmasses kann für eine Anwendung ein ebenso wichtiges Kriterium sein wie die Robustheit gegenüber Bildstörungen. Ein perfekt arbeitender Algorithmus, der eine vorgegebene Systemtaktrate nicht einhalten kann, ist ebenso unnütz wie ein Algorithmus mit ungenügender Detektionsrate, aber kleinem Rechenaufwand. Die rasante Entwicklung der hochintegrierten Schaltungen verschiebt zwar diese Grenze von Jahr zu Jahr in Richtung aufwendiger und leistungsfähiger Verfahren, die dadurch neu nutzbare Rechenleistung kann aber in der Regel nicht voll für die auf algorithmisch tiefem Niveau arbeitenden Rechenverfahren – wie z.B. die Korrelationskriterien – eingesetzt werden. Vielmehr steigt bei klassischen Realisationen in der Regel mit zunehmender Rechengeschwindigkeit das Bedürfnis komplexere, übergeordnete Algorithmen zu verwenden und wesentliche Teile der verfügbaren Rechenleistung für diese einzusetzen. Die grundlegenden Verfahren, werden häufig 'Toolbox-artig' weiterbenutzt, deren Resultate aber intelligenter weiterverarbeitet. So ist leicht vorauszu-sehen, dass auch in absehbarer Zukunft die Rechenleistung bzw. der Rechenaufwand noch von Belang sein wird.

Der heutige Stand der allgemeinen, digitalen Signalprozessoren, wie auch einiger RISC-Prozessoren (reduced instruction set computer), erlaubt ihren Einsatz in Echtzeit-Bildanalysesystemen (z.B. 50Hz-Systemtaktrate). Zumindest Teilaufgaben können von ihnen selbständig und genügend schnell gelöst werden. Obwohl spezielle Hardwarelösungen (z.B. elektronische Lösungen: [32] und [170]; analoge Implementation: [63]; optische Realisation: [35]) grundsätzlich immer für höhere Geschwindigkeiten als allgemeine Prozessorsysteme ausgelegt werden können, bietet der moderne Signalprozessor, einzeln oder im Verbund, eine attraktive, einfach auf ändernde Anforderungen anpassbare Alternative.

In der Regel enthalten die Signalprozessoren, die zur Entlastung des Ein-/Ausgangs-Buses häufig nicht in der von Neumann sondern in der Harvard-Struktur mit getrenntem Daten- und Adressbus aufgebaut sind, mehrere parallel arbeitende Funktionsblöcke. Diese erlauben z.B. gleichzeitig eine Multiplikation, Addition und Adressrechnung durchzuführen. Bei optimaler

Programmierung können also unter Berücksichtigung der Pipeline-Nebenbedingungen, wie der Latenzzeit und dem Füllgrad, und ohne verlangsamenden externen Datenzugriff pro Instruktionszyklus (etwa 40 bis 100ns, je nach Prozessortyp) mehrere Operationen parallel bewältigt werden. Als typische Funktionsblöcke sind in der Regel eine ALU (arithmetic logic unit), ein Multiplizierwerk, eine Adressrecheneinheit sowie interner Speicher (teilweise als Programm-/Daten-Cache organisiert) realisiert. Eine der wichtigsten Eigenschaften für die Signalverarbeitung ist aber die Fähigkeit die meisten Operationen, die Multiplikation eingeschlossen, in einem einzigen Instruktionszyklus³⁷ bewältigen zu können

So leistungsfähig die heutigen Prozessoren auch sind, so wichtig sind auch gute Compiler, die Hochsprachenprogramme effizient in Assemblerbefehle übersetzen. Leider ist heute gerade in diesem Bereich bei den Signalprozessoren noch ein Defizit an Leistung zu spüren. Obwohl gewisse Fortschritte zu beobachten sind – z.B. werden vermehrt auch Parallelinstruktionen unterstützt – sind doch immer noch wesentliche Geschwindigkeitsunterschiede zwischen Hochsprachen- und optimierten Assemblerprogrammen im Bereich von vier bis im Extremfall zwanzig zu beobachten. Diese Situation ist natürlich unbefriedigend, da der Hardwarevorteil der Signalprozessoren gegenüber allgemeinen Mikroprozessoren damit mindestens zum Teil durch ungenügende Softwarewerkzeuge wieder eliminiert wird. Deshalb wird heute häufig das Programmgerüst sowie Zeit-unkritische Teile in einer Hochsprache und der rechenintensivste Teil des Programms (meist innerste Schleife) in handoptimiertem Assembler geschrieben (z.B. [46]).

Da heute in der Bildanalyse in zunehmendem auch Masse Signalprozessoren zum Einsatz kommen, soll der hier diskutierte Rechenaufwand der Korrelationskriterien zwar theoretisch, aber unter Berücksichtigung der Fähigkeiten dieser Spezialprozessoren analysiert werden³⁸. Die wichtigste Annahme ist dabei, dass Multiplikationen und Additionen bezüglich der Ausführungszeit gleichwertige Operationen sind, die teilweise sogar parallel ausgeführt werden können. Dies steht im krassen Gegensatz zu der heute teilweise überholten, gelegentlich aber immer noch benutzten Regel, dass Multiplikationen, wegen ihrer vermeintlich grösseren Ausführungszeit im Vergleich zu Additionen, die Rechenlast stark dominieren. Bei der Verwendung von Fließkommazahlen (, wo die Addition wegen der notwendigen Normalisierung und De-Normalisierung auch einen gewissen Aufwand bedeutet), wie

³⁷ Dies gilt wegen der Pipelinestruktur der Signalprozessoren nicht für Einzeloperationen, sondern nur für Befehlsfolgen, die die Funktionsblöcke optimal ausnützen und die Pipeline dauernd gefüllt halten.

³⁸ Andersartige Zielsysteme führen zu anderen Resultaten.

auch bei einer wachsenden Zahl von Spezial- und general purpose Prozessoren mit schnellem Multiplizierwerk, ist dies aber oft nicht mehr richtig. Praktische Messungen an Signalprozessorimplementationen einzelner Korrelationskriterien bestätigen die Gültigkeit dieser Bemerkungen.

Des weiteren werden bei der Bewertung des Rechenaufwandes die folgenden Annahmen gemacht: Der Adressrechenaufwand wie auch der Aufwand, der durch die Schleifenkontrolle, die Programm-Ablaufsteuerung, die Initialisierung und andere Nebenaufgaben entsteht, wird vernachlässigt. Signalprozessoren können durch ihre spezielle Architektur diese Aufgaben häufig parallel zur eigentlichen Signalverarbeitung ausführen. Bei normalen CPUs verursachen sie jedoch in der Regel einen Zusatzaufwand.

Verschiedene Methoden zur Umformulierung eines Korrelationsalgorithmus³⁹ können den Rechenaufwand drastisch reduzieren, ohne die Qualität des Resultats zu beeinflussen; andere wiederum stellen nur Approximationen an die korrekte Rechnung dar. Ein Verfahren den Rechenaufwand zu reduzieren ohne das Resultat – die Position des detektierten des Extremwerts der **KKF** – zu verändern ist z.B. das vorzeichenbehaftete Quadrieren der **KKF** zur Elimination der aufwendigen Wurzelberechnung³⁹ (eindimensionaler Fall: Gl.13). Jede Wurzeloperation wird damit durch eine Multiplikation und eine Betragsbildung ersetzt⁴⁰. -Demgegenüber kann die Reduktion des Rechenbedarfs durch eine spezielle Suchstrategie zur Bestimmung der Position des Extremwerts, bei der das entsprechende Ähnlichkeitsmass nur noch an speziell ausgewählten Positionen ausgewertet wird, zu einem völlig anderen Resultat führen als der voll durchgerechnete Korrelationsalgorithmus. Solche, das Resultat möglicherweise modifizierende Verfahren werden hier nicht diskutiert (siehe z.B. [67]).

$$\text{KKF}(x) = \frac{\sum_{u=0}^{N-1} r(u) \cdot s(x+u)}{\sqrt{\sum_{u=0}^{N-1} r^2(u) \cdot \sum_{u=0}^{N-1} s^2(x+u)}} \quad (13a)$$

³⁹ Beim Signalprozessor Texas Instruments TMS320C30: Für maximale Präzision → 32 Bit Wurzelberechnung: 49 Instruktionszyklen ([177]); eine Multiplikation dauert 1 Instruktionszyklus und lässt gleichzeitig noch eine ALU-Operation zu; siehe auch [117] für ein ähnliches Vorgehen.

⁴⁰ Die Betragsbildung ist nur notwendig, falls im Suchfenster oder im Muster negative Werte auftreten.

$$\rightarrow \text{KKF}'(x) = \frac{\text{abs} \left\{ \sum_{u=0}^{\text{NMAX}} r(u) \cdot s(x+u) \right\} \left\{ \sum_{u=0}^{\text{NMAX}} r(u) \cdot s(x+u) \right\}}{\sum_{u=0}^{\text{NMAX}} r^2(u) \cdot \sum_{u=0}^{\text{NMAX}} s^2(x+u)} = \frac{\text{Zähler}(x)}{\text{Nenner}(x)} \quad (13b)$$

Weitere Betrachtungen zu den Rechenlasten finden sich auch in [10], [17], [20], [33], [41], [46], [67], [68], [71], [147], [148] und [172]. Besonders interessant ist dabei der Beitrag von Siegel et al. ([147]). Die Übereinstimmung mit den hier vorgestellten Resultaten ist aufgrund ähnlicher Annahmen sehr gut.

6.2. Verminderung des Rechenaufwandes

Im folgenden werden einige, bei der Bestimmung des Rechenaufwands berücksichtigte Verfahren diskutiert, die ohne Resultatverfälschung zur Verminderung der Rechenlast der Korrelationskriterien beitragen können.

6.2.1. Extremwertsuche

Bisweilen ist einzig die Position des Extremwerts von Interesse. Die Korrelationswerte selber werden nicht benötigt. Mit Hilfe dieser Reduktion der Ansprüche an die Qualität der Resultate ist es beispielsweise möglich, die Rechenlast bei allen (Un-) Ähnlichkeitsmassen, die in einen Nenner und Zähler zerlegbar sind, durch Beseitigung der Division zu reduzieren. Dies gelingt aber nur solange die Division aufwendiger zu berechnen ist als zwei Multiplikationen (sowie einige Vergleiche und Zuweisungen, siehe Fig.19), was heute noch bei den meisten Prozessoren zutrifft. Die Bestimmung der Position maximaler Ähnlichkeit z.B. mit Hilfe der **KKF** kann – zur Eliminierung der normalerweise aufwendigeren Division⁴¹ – demnach im ein-dimensionalen Fall folgendermassen durchgeführt werden (Fig.19, PASCAL-ähnliche Notation; siehe auch Gl.13, 14 und 15):

$$\text{KKF}(x) = \frac{\text{Zähler}(x)}{\text{Nenner}(x)} \stackrel{?}{>} \frac{\text{MaxZählerBisher}}{\text{MaxNennerBisher}} \quad (14)$$

$$\stackrel{?}{\Rightarrow} \text{Zähler}(x) \cdot \text{MaxNennerBisher} > \text{Nenner}(x) \cdot \text{MaxZählerBisher} \quad (15)$$

⁴¹ 43 Instruktionszyklen beim TMS320C30 für 32 Bit Fließkommazahl und maximale Genauigkeit ([177]).

```

PROCEDURE GetMaxKKFPosition; {Bestimmung der Position}
                               {des Extremwerts (Max)}
VAR x: INTEGER;                {lokale Definitionen}
    MaxZählerBisher, MaxNennerBisher: REAL;
BEGIN {GetMaxKKFPosition}
    MaxZählerBisher:= Zähler[0];    {Initialisierung mit ersten Werten}
    MaxNennerBisher:= Nenner[0];    {Zähler, Nenner, MaxPosition, MaxX}
    MaxPosition:= 0;                {sind global definiert}
    FOR x:= 1 TO MaxX DO            {Suche nach Maximumposition}
    BEGIN
        IF (Zähler[x]*MaxNennerBisher) >
           (Nenner[x]*MaxZählerBisher) THEN
        BEGIN                       {neues Maximum gefunden}
            MaxZählerBisher:= Zähler[x];
            MaxNennerBisher:= Nenner[x];
            MaxPosition:= x;
        END;
    END;
END; {GetMaxKKFPosition}

```

Fig.19; Bestimmung der eindimensionalen Position des maximalen Korrelationswerts (MaxPosition) ohne Division (siehe auch Gl.15, anwendbar auf **KKF**, **KKFMF**, **MOR**, **NSSDMF**, **NA1** und **NA2**). Eine Division (inkl. Extremwertsuche) wird damit durch zwei Multiplikationen sowie je einen Vergleich, einen Entscheid und 1.5 Zuweisungen⁴² ersetzt.

Die Suche der Minimumposition bei der **SSDMS** kann in ähnlicher Weise abgeändert werden (Gl.16). Somit entfällt die aufwendige, für jedes einzelne Korrelationsfenster notwendige Division. Im Prinzip wird bei der Detektion des Extremwerts getestet, ob folgende noch nicht vereinfachte Ungleichung gilt⁴³ (eindimensionaler Fall):

$$\sum_{u=0}^{\text{MAXX}} \left\{ r(u) - \frac{\sum_{\xi=0}^{\text{MAXX}} r(\xi)}{\sum_{\xi=0}^{\text{MAXX}} s(x_i + \xi)} \cdot s(x_i + u) \right\}^2 < \sum_{u=0}^{\text{MAXX}} \left\{ r(u) - \frac{\sum_{\xi=0}^{\text{MAXX}} r(\xi)}{\sum_{\xi=0}^{\text{MAXX}} s(x_{\text{bisher}} + \xi)} \cdot s(x_{\text{bisher}} + u) \right\}^2 \quad (16)$$

⁴² Dies gilt unter der Annahme, dass die Ungleichung 15 im Mittel in der Hälfte der Fälle erfüllt ist.

⁴³ Um die Mittelwertbildung, d.h. die Division der Helligkeitssummen durch die Anzahl Musterpixel zu umgehen, können in diesem Fall alle Mittelwerte durch die in den entsprechenden Fenstern enthaltenen Summen der Intensitätswerte ersetzt werden.

Falls Ungleichung 16 gilt, so ist x_i als neues Temporär-Minimum x_{bisher} zu betrachten und für den Vergleich mit dem Korrelationswert der nächsten Position $(i+1)$ zu benutzen. Die Umformung – erreicht durch Ausmultiplizieren der Ungleichung und Erweiterung mit dem Term $\left\{ \sum_{\xi=0}^{u_{max}} s(x_1 + \xi) \cdot \sum_{\xi=0}^{u_{max}} s(x_2 + \xi) \right\}^2$ – ergibt eine Formel, die anstelle der Division elf Multiplikationen enthält.

Das Ausmultiplizieren des Quadrats der Differenzen lohnt sich übrigens in der Regel genauso, wenn die Division nicht eliminiert werden soll. Die pixelweise Skalierung (Gl.16) bzw. Mittelwertbefreiung muss nach dem Ausmultiplizieren nur noch einmal pro Korrelationsfenster durchgeführt werden, da die Skalierungs- bzw. Mittelwertbefreiungs-Terme nun ausserhalb der Summation stehen (Gl.17 und 18):

$$\sum_{u=0}^{u_{max}} \left\{ \{r(u) - \bar{r}\} - \{s(x+u) - \bar{s}(x)\} \right\}^2 = \underbrace{\sum_{u=0}^{u_{max}} \{r(u) - s(x+u)\}^2}_{SSD} - \left\{ \frac{\bar{r} - \bar{s}(x)}{u_{max} + 1} \right\}^2 \quad (17)$$

bzw.

$$\sum_{u=0}^{u_{max}} \left\{ r(u) - \frac{\bar{r}}{\bar{s}(x)} \cdot s(x+u) \right\}^2 = \sum_{u=0}^{u_{max}} r^2(u) - 2 \frac{\bar{r}}{\bar{s}(x)} \sum_{u=0}^{u_{max}} r(u)s(x+u) + \left\{ \frac{\bar{r}}{\bar{s}(x)} \right\}^2 \sum_{u=0}^{u_{max}} s^2(x+u) \quad (18)$$

Bei der **SSD** – und damit auch bei der **SSDMF** (siehe Gl.17) – kann sich das Ausmultiplizieren ebenfalls auszahlen, falls die Berechnung nicht über Look-up-Tabellen erfolgt und verglichen mit der Suchfenstergrösse nicht sehr grosse oder kleine Muster benutzt werden.

6.2.2. Mittelwerte, Energieterme und Laplacefilterung

Die in der **KKF** und anderen (Un-) Ähnlichkeitsmassen erscheinende Energie $\left\{ \sum_{v=0}^{v_{max}} \sum_{u=0}^{u_{max}} r^2(u, v) \right\}$ bzw. die Varianz des Musters muss eigentlich nicht berechnet werden, da sie für alle Positionen im Suchfenster konstant ist. Dieser Term hat darum – rein additive bzw. multiplikative Verknüpfung vorausgesetzt (z.B. nicht gegeben bei **MOR**) – keinen Einfluss auf die Position des Extremwerts.

Die Energieberechnung im Suchfenster $\left\{ \sum_{v=0}^{v_{\max}} \sum_{u=0}^{u_{\max}} s^2(x+u, y+v) \right\}$, z.B. bei der **KKF**, der **MOR**, oder der **SSDMS**, muss hingegen für jede einzelne Position im Suchfenster durchgeführt werden⁴⁴. Sie kann jedoch derart modifiziert werden, dass – vom Initialisierungsaufwand abgesehen – pro Position nur zwei Additionen und zwei Subtraktionen ausgeführt werden müssen. Dabei wird jeweils vom horizontal um ein Pixel verschobenen Resultat der vorherigen Energieberechnung Gebrauch gemacht (Fig.20). Von diesem Resultat muss nur noch die Energie der letzten Pixelkolonne weg- und diejenige der neuen Pixelkolonne dazugezählt werden. Die Energien der um ein Pixel vertikal verschobenen Pixelkolonnen werden dann in ähnlicher Weise aktualisiert: Die Energie des obersten Pixels wird abgezählt und diejenige des neuen Pixels (unten) dazugezählt. Natürlich wird bei der Anwendung dieser Methode mehr Speicherplatz, nämlich zur Speicherung der Energien aller Pixelkolonnen, benötigt.

Das gleiche Verfahren kann auch bei der Filterung mit dem Laplaceoperator (Operator in Mittelwertbildung und Subtraktion des zentralen Helligkeitswerts auftrennen, $5xIs^2 + 6xIs + 2$ Additionen und xIs^2 Multiplikationen) oder bei der Bestimmung des Intensitätsmittelwerts benutzt werden. Dies ist z.B. bei der Berechnung der **KKFMF** möglich, deren Nenner zudem vorteilhafterweise ausmultipliziert wird (nur Korrelationsfensterteil) und die einzelnen Terme separat berechnet werden (Gl.23, [117], [148], [156]⁴⁵). Zusätzlich kann die resultatmässig völlig unnötige Mittelwertbefreiung des Korrelationsfensters im Zähler, wie bei der **UKKFMF**, ausgelassen werden (Gl.19):

$$\begin{aligned} \text{UKKFMF}(x, y) &= \sum_{v=0}^{v_{\max}} \sum_{u=0}^{u_{\max}} \{r(u, v) - \bar{r}\} \cdot \{s(x+u, y+v) - \overline{s(x, y)}\} \\ &= \sum_{v=0}^{v_{\max}} \sum_{u=0}^{u_{\max}} \underbrace{\{r(u, v) - \bar{r}\}}_{\text{konstant in } x \text{ und } y} \cdot s(x+u, y+v) \end{aligned} \quad (19)$$

⁴⁴ Sonst kann anstelle der **KKF** auch die **SSD**, die **UKKF** oder die Korrelation von Moravec ausgewertet werden, da im Fall konstanter lokaler Energie im Suchfenster diese Ähnlichkeitsmasse äquivalent sind und darum ihre Extrema an derselben Position aufweisen.

⁴⁵ In dieser Literaturreferenz ist bei der entsprechenden Formel ein kleiner Fehler zu korrigieren. Das Quadrieren bei der Berechnung der Suchfenstervarianz ist vergessen gegangen.

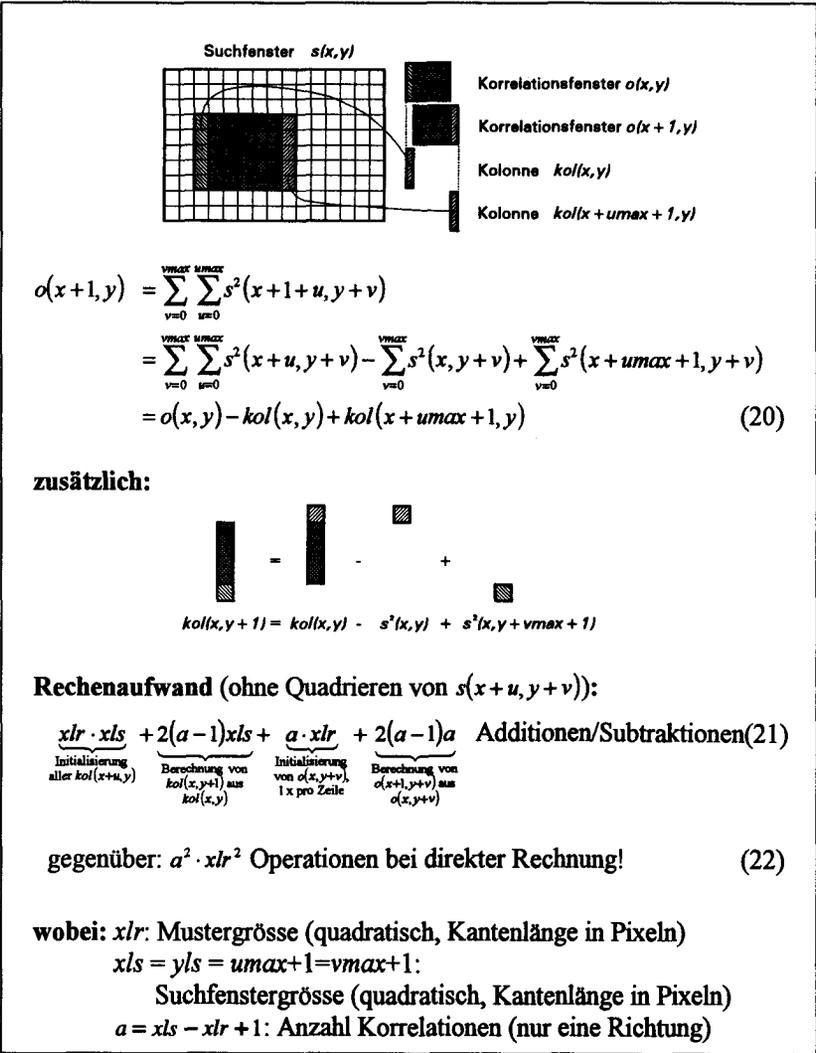


Fig.20; Reduktion des Aufwandes bei der Bestimmung der lokalen Suchfensterenergie durch Verwendung benachbarter Energieresultate (siehe auch "fast normalization" in [20]).

$$\text{KKFMF}(x, y) = \frac{\sum_{v=0}^{\text{VMAX}} \sum_{u=0}^{\text{UMAX}} \{r(u, v) - \bar{r}\} \cdot s(x+u, y+v)}{\sqrt{\left(\sum_{v=0}^{\text{VMAX}} \sum_{u=0}^{\text{UMAX}} \{r(u, v) - \bar{r}\}^2 \right) \left(\sum_{v=0}^{\text{VMAX}} \sum_{u=0}^{\text{UMAX}} s^2(x+u, y+v) - xls \cdot yls \cdot \overline{s(x, y)^2} \right)}}
 \quad (23)$$

Falls genügend Speicher vorhanden ist, kann der Rechenaufwand bei verschiedenen Verfahren weiter reduziert werden, indem die Quadrate aller Suchfenster-Pixelwerte für die Energie- oder Varianzenberechnung vorausbestimmt und zwischengespeichert werden ([27]). Dies fällt bei einem Signalprozessor aber weniger ins Gewicht, da er ja in der Lage ist gleichzeitig eine Multiplikation und eine Addition auszuführen. Beim Vorausquadrieren kann von dieser Eigenschaft meist kein Gebrauch gemacht werden.

6.2.3. Verwendung von Frequenztransformationen

Die Terme der **KKF** sowie die **UKKF** lassen sich natürlich auch über Frequenztransformationen bestimmen. Dazu können z.B. die Fourier-Transformation (Fast Fourier Transform, FFT; [20], [68], [102], [144], [151]), die Hartley-Transformation oder auch eine separierbare Version der Hartley-Transformation zusammen mit den entsprechenden Korrelationstheoremen ([24]) verwendet werden. Dieses alternative Vorgehen bringt aber erst bei grösseren Suchfenstern (ab etwa 64x64 Pixel) eine Verringerung der Rechenlast, da der Rechenaufwand dieser indirekten Berechnungsart ungefähr proportional zu $xls \cdot \{xls \cdot \ln(xls)\}$ steigt und damit bei kleinen Suchfenstern noch grösser ist als bei der direkten Rechnung mit einem Aufwand, der sich im wesentlichen proportional zu $a^2 x l r^2$ verhält (der Einfachheit halber sei hier angenommen, dass das Suchfenster quadratisch sei und der Wert der Kantenlänge als ganzzahlige Zweierpotenz darstellbar ist)⁴⁶. Selbst bei Suchfenstern von 64x64 Pixeln ist im Experiment noch kein Vorteil der FFT-basierten Rechnung auszumachen, da dann, zumindest bei den heutigen Signalprozessoren, ein Teil der Daten ausserhalb des Prozessorchips abgelegt werden müssen und damit das Rechenwerk durch die limitierende Speicher-Ein-/Ausgangsbandbreite zusätzlich gebremst wird. Die Berechnung der **KKF** z.B. über die echte bzw. die separierbare, zweidimensionale

⁴⁶ In der Dissertation von M. J. Hannah [67] wird eine ähnliche Aussage gemacht. Die Berechnung der **KKF** über die FFT soll (auf einer PDP 8) erst ab Suchfenstergrössen von 100x100 und Mustergrössen von 15x15 Pixel schneller gegenüber der direkten Rechnung sein.

Hartley-Transformation ist trotz den nur reellwertigen Operationen etwa gleich aufwendig wie die Rechnung mit Hilfe der FFT, da im ersten Fall die beiden Richtungen nicht separierbar sind und damit diese Transformation am schnellsten wiederum über die FFT berechnet wird und im zweiten Fall das Korrelationstheorem relativ kompliziert ist ([24]).

Der Rechenaufwand für die Berechnung der **KKF** mit Hilfe der FFT stellt sich folgendermassen dar: (nach Brigham, [30], S.185 und 205; nur der Zähler wird über die Transformation berechnet; ohne Windowing; nur für Suchfensterdimensionen xls , die ganzen Zweierpotenzen entsprechen und quadratische Such- und Referenzfenster; reelle Operationen):

$$\begin{aligned} \text{Additionen: } & (3 \cdot xls + xlr) \cdot \text{FFTAdd} + 2a \cdot \text{IFFTAdd} + \\ & 2xls^2 + xlr^2 + xlr \cdot xls + 2(xls + a)(a - 1) + a \cdot xlr \end{aligned} \quad (24)$$

$$\text{Multiplikationen: } (3 \cdot xls + xlr) \cdot \text{FFTMul} + 5xls^2 + 2a \cdot \text{IFFTMul} + xlr^2 + a^2 \quad (25)$$

$$\text{Divisionen: } a^2 \quad (26)$$

$$\text{Wurzeln: } a^2 \quad (27)$$

wobei:

xlr : Muster-Kantenlänge (Annahme: quadratisches Referenzfenster, $xlr = ylr = umax + 1 = vmax + 1$)

xls : Suchfenster-Kantenlänge (Annahme: quadratisches Suchfenster, $xls = yls$)

$a = xls - xlr + 1$: Anzahl Korrelationspositionen auf einer Zeile bzw. Kolonne

$\text{FFTAdd} = xls \cdot \text{ld}(xls)$: Anzahl Additionen für eine eindimensionale FFT, komplexe Eingangswerte (ld: Logarithmus zur Basis 2)

$\text{FFTMul} = 2 \cdot xls \cdot \text{ld}(xls)$: Anzahl Multiplikationen für eine eindimensionale FFT, komplexe Eingangswerte

$\text{IFFTAdd} = 2 \cdot xls \cdot \text{ld}(xls)$: Anzahl Additionen für eine eindimensionale, inverse FFT, komplexe Eingangswerte

$\text{IFFTMul} = 2 \cdot xls \cdot \text{ld}(xls)$: Anzahl Multiplikationen für eine eindimensionale, inverse FFT, komplexe Eingangswerte

Der in den Formeln 24 bis 27 festgehaltene Rechenaufwand beinhaltet die folgenden Teilaufgaben:

Vorwärtstransformation Muster:

$(xls + xlr) \cdot FFTAdd$ Additionen

$(xls + xlr) \cdot FFTMul$ Multiplikationen

Vorwärtstransformation Suchfenster:

$2 \cdot xls \cdot FFTAdd$ Additionen

$2 \cdot xls \cdot FFTMul$ Multiplikationen

(komplexe) Multiplikation der Spektren:

$2xls^2$ Additionen

$4xls^2$ Multiplikationen

Musterenergie:

xlr^2 Additionen

xlr^2 Multiplikationen (parallelisierbar)

Suchfensterenergie (siehe Fig.20):

$xlr \cdot xls + 2(a-1)xls + a \cdot xlr + 2(a-1)a$ Additionen

xls^2 Multiplikationen

Multiplikation der Muster- mit der Suchfensterenergie (Nenner **KKF**):

a^2 Multiplikationen

Für das nachstehende Beispiel werden die aufwendigen Divisionen und Wurzeloperationen durch insgesamt $3a^2$ Multiplikationen und $3.5a^2$ Additionen ersetzt (Additionen: inkl. Subtraktionen, Vergleiche, Entscheide und Zuweisungen. Siehe Gl.13 und Fig.19, Annahme: Nur positive Pixelwerte).

Beispiel: $xls = 32$, $xlr = 16 \Rightarrow$ ca. 87'700 Instruktionen⁴⁷. Bei direkter Rechnung: ca. 79'800 Instruktionen (siehe Tab.10 und Fig.22.II, keine parallelen Operationen!).

6.2.4. RichtungsfILTER

Für die Bestimmung des Rechenaufwands des Sobeloperators (**SES1**, **SES2**; inklusive Division, *arctan*-Berechnung und Quantisierung) wird eine Look-up-Tabelle der Grösse $2^{2(n+3)} \times q$ Bit vorausgesetzt, wobei 2^n die Anzahl unterscheidbarer Intensitäten im Originalbild und 2^q die Richtungsquantisie-

⁴⁷ Die Divisionen und Wurzeloperationen wurden gemäss Fig.19 und Gl.13 ersetzt, um für beide Realisationen gleiche Bedingungen zu schaffen.

rung des Helligkeitsgradienten bezeichnet⁴⁸. Aus der Berechnung aller benötigten horizontalen und vertikalen 3-er Summen (siehe Anhang A.1: *Nomenklatur*) und Verwendung dieser Look-up-Tabelle resultiert ein Aufwand⁴⁹ – wiederum ohne Overhead und Adressierung – von $9xls^2 + 12xls$ Additionen und $2xls^2 + 4xls$ Multiplikationen.

Das Kirschfilter mit seinen 8 Masken – je eine pro Richtung – wird mit einem Aufwand von $22xls^2$ Additionen, $9xls^2$ Multiplikationen, $8xls^2$ Betragsbildungen sowie je $7xls^2$ Vergleichen und Entscheidungen veranschlagt. Bei jedem Pixel im Suchfenster wird die erste Kirschmaske direkt gerechnet, indem zuerst alle Pixel mit Gewicht 3 und alle Pixel mit Gewicht -5 getrennt aufsummiert und dann entsprechend gewichtet werden. Jede weitere Maske – am selben Ort (x,y) – wird durch Korrektur des vorherigen Resultats berechnet. Ein Pixel muss dabei, gewichtet mit dem Multiplikator 8, dazu und ein zweites Pixel, mit dem selben Multiplikator, abgezählt werden. D.h. an jedem Ort im Suchfenster kommt neben den schon gerechneten acht Additionen und zwei Multiplikationen für die Berechnung der ersten Maske noch je eine weitere Addition, Subtraktion und Multiplikation (gleichzeitige Gewichtung beider Korrekturwerte) pro zusätzliche Maske hinzu. 8 Betragsbildungen und je 7 Vergleiche und Entscheidungen werden schliesslich zur eigentlichen Bestimmung der Intensitätsrichtung benötigt. Diese Betrachtungen führen zu einem, verglichen mit dem Sobeloperator, grossen Rechenaufwand. Durch eine etwas andere Betrachtungsweise liesse sich der Kirschoperator aber wie der Sobeloperator relativ einfach auf Hardware abbilden ([25]).

Bei der **SES1** und der **SEK1** kann zudem die Wurzeloperation beim Unähnlichkeitsmass weggelassen werden, da sie die Position des Extremwerts nicht verschiebt.

⁴⁸ Für $n=6$ und 32 unterscheidbare Richtungen ($32=2^5 \Rightarrow q=5$) muss somit ein Speicher von $2^{18} \times 5$ Bit bereitstehen. Der Term '+3' in der Grössenberechnung der Look-up-Tabelle entsteht durch die maximal mögliche Bitzahl des Resultats des horizontalen bzw. vertikalen Gradienten ($n+3$ Bit). Die Multiplikation mit 2 ist durch die Tatsache gegeben, dass sowohl der horizontale wie auch der vertikale Gradientenwert als Eingänge für die Look-up-Tabelle dienen. D.h. neben der *arctan*-Berechnung wird mit diesem Speicher zusätzlich auch die Division durchgeführt.

⁴⁹ Die Vorbereitung des Look-up-Tabellen-Eingangswertes (Verknüpfung des horizontalen und vertikalen Gradientenwertes) benötigt je xls^2 Multiplikationen und Additionen.

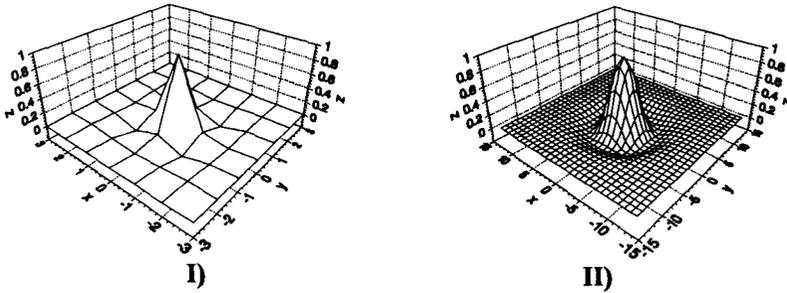


Fig.21: Maske des Laplace-Gauss-Operators für $w=2$ (I) und $w=8$ (II).

6.2.5. Laplace-Gauss-Operator und Operator von Nack

Die Rechenlast des Laplace-Gauss-Filters (NIS) ist ohne besondere Vorkehrungen zur Rechenzeitverkürzung sehr gross. Es ergeben sich $FD^2 \cdot xls^2$ Additionen und gleich viele Multiplikationen (FD : Kantenabmessung des Filterkerns). Die notwendige Binarisierung des Resultats ist dabei nicht berücksichtigt. Nishihara ([124]) verwendet eine Approximation dieses Filters durch die Differenz zweier Gauss-Operatoren (oft mit *DoG* abgekürzt, *Difference of Gaussians*; siehe Laplace-Gauss-Operator, Fig.21), was wegen der Separierbarkeit des Gauss-Operators den Rechenaufwand auf $(2 \cdot FD + 1) \cdot xls^2$ Additionen/Subtraktionen und $2 \cdot FD \cdot xls^2$ Multiplikationen reduziert⁵⁰ (Multiplikationen und Additionen sind parallelisierbar). Der Bestimmung des Rechenbedarfs wurde diese Annäherung zugrunde gelegt. Die Versuche wurden allerdings mit dem abgetasteten Originalfilter durchgeführt.

Wie beim Laplace-Gauss-Operator wird auch beim Filter, das von Nack vorgeschlagen wird der Rechenaufwand für die Binarisierung des Resultats und für die Bestimmung der adaptiven Schwelle nicht in die Rechenlast miteinbezogen. Deswegen werden nur $11 \cdot xls^2$ Operationen berücksichtigt (pro Pixel im Suchfenster 4 Subtraktionen, 4 Absolutwertbestimmungen und 3 Additionen).

⁵⁰ Diese Formel geht von zwei gleich grossen Gauss-Operatoren aus, deren Resultate voneinander subtrahiert werden. Eigentlich könnte der eine Gauss-Filterkern aber kleiner gewählt werden, so dass real ein geringerer Aufwand entstünde. In den weiteren Betrachtungen zum Rechenaufwand wird aber von zwei gleich grossen Gauss-Operatoren ausgegangen (Grösse: Siehe Anhang A.1: *Nomenklatur*).

6.2.6. Zusatzaufwand

Durch das Ignorieren des Zusatzaufwandes (Rechenlast für Adressrechnung, Schleifenkontrolle etc.) ergibt sich ein optimistischeres Bild der Rechenlast als tatsächlich im besten Fall zu erreichen ist. Trotzdem haben Messungen der Rechenzeit an einem TMS320C30-System mit teilweise handoptimiertem Assemblercode eine relativ gute Übereinstimmung mit den unten aufgetragenen Rechenlasten gezeigt⁵¹.

6.3. Rechenaufwand der Korrelationskriterien

Die oben diskutierten Rechenlasten der Teilschritte werden nun zum Gesamtrechenaufwand der einzelnen Korrelationsalgorithmen zusammengefügt.

6.3.1. Gesamter Rechenaufwand für die KKF

Für die KKF soll der gesamte Rechenaufwand, sowohl bei bezüglich Rechenaufwand optimierter, als auch nicht optimierter Vorgehensweise, exemplarisch bestimmt werden (Aufwand II respektive I, siehe Tab. 8). Die entsprechenden Angaben für die weiteren, hier ebenfalls untersuchten Korrelationskriterien werden in den Tabellen 9 und 10 gegeben, aber nicht mehr im Detail erläutert. Alle notwendigen Teil-Rechenschritte wurden oben diskutiert.

Zur Erinnerung:

$$\text{KKF}(x, y) = \frac{\sum_{v=0}^{\text{VMAX}} \sum_{u=0}^{\text{UMAX}} r(u, v) \cdot s(x+u, y+v)}{\sqrt{\sum_{v=0}^{\text{VMAX}} \sum_{u=0}^{\text{UMAX}} r^2(u, v) \cdot \sum_{v=0}^{\text{VMAX}} \sum_{u=0}^{\text{UMAX}} s^2(x+u, y+v)}} \quad (28)$$

⁵¹ Ca. 75% Mehraufwand gegenüber der theoretischen Rechenaufwandanalyse wurde gemessen (Zeitmessung, nicht Messung der Anzahl Instruktionszyklen!). Das scheint eine relativ grosse Differenz zu sein, berücksichtigt man aber die optimistischen Annahmen und die Tatsache, dass beim Experiment nur der innerste Teil der Schleife handoptimiert wurde, der verwendete C-Compiler einen relativ schlecht optimierten Code erzeugt und insbesondere prozessorexterne Daten- und Programmmzugriffe den Ablauf wesentlich verlangsamten, ist die Übereinstimmung zwischen Theorie und Experiment dennoch bemerkenswert.

Berechnung von:	Aufwand I	Gewicht I	Aufwand II	Gewicht II
A	a^2xlr^2 Add. a^2xlr^2 Mul.	1 1	a^2xlr^2 MAC	1
B	xlr^2 Add. ⁵² xlr^2 Mul.	1 1	xlr^2 MAC	1
C	xls^2 Mul.	1	xls^2 Mul.	1
D (Fig. 20, Gl. 21)	$xlr \cdot xls + 2(a-1)xls +$ $a \cdot xlr + 2(a-1)a$ Add.	1	$xlr \cdot xls + 2(a-1)xls +$ $a \cdot xlr + 2(a-1)a$ Add.	1
E	a^2 Mul.	1	a^2 Mul.	1
F	a^2 Wur.	49 ([178])	wird nicht berechnet!	
G	a^2 Div.	43 ([178])	wird nicht berechnet!	
H	$3a^2$ Add.	1	wird nicht berechnet!	
I (Gl. 13 ⁵³)	wird nicht berechnet!		a^2 Mul.	1
J (Gl. 15, Fig. 19)	wird nicht berechnet!		$5.5a^2$ Mul./Add.	1

Tab.8: Rechenaufwand bei der Bestimmung der Maximumposition der **KKF** aufgespalten in die verschiedenen Teilaufgaben. Fall II: Aufwand, der bei Signalprocessorimplementationen resultiert, wenn sowohl die 'MAC'-Instruktionen (multiply and accumulate) verwendet, als auch die aufwendigen Quadratwurzeln und Divisionen ersetzt werden (Rechen-schritte F, G und H werden durch I und J ersetzt, siehe unten)⁵⁴.

Folgende Teilschritte zur Bestimmung der Maximumposition der **KKF** können unterschieden werden:

$$A: \sum_{v=0}^{\text{MAX}} \sum_{u=0}^{\text{MAX}} r(u,v) \cdot s(x+u, y+v)$$

$$B: \sum_{v=0}^{\text{MAX}} \sum_{u=0}^{\text{MAX}} r^2(u,v)$$

$$C: s^2(x,y)$$

$$D: \sum_{v=0}^{\text{MAX}} \sum_{u=0}^{\text{MAX}} s^2(x+u, y+v)$$

$$E: \sum_{v=0}^{\text{MAX}} \sum_{u=0}^{\text{MAX}} r^2(u,v) \cdot \sum_{v=0}^{\text{MAX}} \sum_{u=0}^{\text{MAX}} s^2(x+u, y+v) \quad F: \text{Nenner}(x,y) = \sqrt{\quad}$$

$$G: \text{KKF}(x,y) = \frac{\text{Zähler}(x,y)}{\text{Nenner}(x,y)}$$

H: Extremwertsuche (normal)

⁵² Subtraktionen, Vergleiche, Entscheide und Zuweisungen werden zu den Additionsoperationen gezählt!

⁵³ Es wird angenommen, dass der Zähler, da keine negativen Pixelintensitäten auftreten sollen, nur positive Werte annehmen kann und somit die Betragsbildung aus Gl.13 überflüssig ist!

⁵⁴ $a = xls - xlr + 1$: Anzahl Korrelationspositionen (nur eine Richtung); xls : Suchfenster-Kantenlänge; xlr : Muster-Kantenlänge.

I: $Zähler^*(x,y) = Zähler^2(x,y)$

J: Extremwertsuche (Fig.19)

Die Teilschritte zur Bestimmung der Maximumposition der **KKF** verursachen den in Tabelle 8 festgehaltenen Rechenaufwand.

Die Zusammenfassung von Tabelle 8 ergibt insgesamt, unter Berücksichtigung der jeweiligen Gewichtung der Operationen, den folgenden Aufwand (gemessen in äquivalenten, 'ein-Zyklus-fähigen' Instruktionen wie z.B. MAC: Addition parallel zu Multiplikation):

Fall I ('brute force' Rechnung):

$$2a^2xlr^2 + xls^2 + 98a^2 + 2xlr^2 + xlr \cdot xls + 2xls \cdot a + xlr \cdot a - 2xls - 2a \quad (29)$$

Fall II (für den Signalprozessor optimierte Implementation):

$$a^2xlr^2 + xls^2 + 9.5a^2 + xlr^2 + xlr \cdot xls + 2xls \cdot a + xlr \cdot a - 2xls - 2a \quad (30)$$

6.3.2. Diskussion des Rechenaufwandes der einzelnen Korrelationskriterien

In den Tabellen 9 und 10 wird der Rechenaufwand in Anzahl Instruktionen für jeden Korrelationsalgorithmus einerseits für die 'brute force' Methode und andererseits für eine optimierte Signalprozessorimplementation spezifiziert. Bei der Optimierung wurden alle der oben diskutierten Methoden zur Verkleinerung der Rechenlast ausgenützt. Als Ausnahme muss hier die Berechnung der Musterenergie bzw. Mustervarianz bei der **KKF**, der **KKFMF** und der **NSSDMF** aufgeführt werden. Diese relativ wenigen, zur Berechnung der korrekten Extremwertposition eigentlich unnötigen Operationen (weil betragsmässig unabhängig von x und y) werden mitgezählt.

Additionen und Multiplikationen werden bei der Betrachtung der totalen Rechenlast gleich gewichtet, während die Divisionen mit dem Gewicht 43 und die Wurzeloperationen mit dem Gewicht 49 belegt werden (Anzahl Instruktionen im Vergleich zur Addition oder Multiplikation beim Signalprozessor TMS320C30, siehe [178]). Diese Gewichtung repräsentiert nur die Verhältnisse beim Signalprozessor TMS320C30! -Neben den oben erwähnten, Rechenzeit vermindernnden Methoden könnte auch durch Abstriche bei der Präzision der Rechnungen (bei der iterativen Bestimmung von Division und Wurzel) die Zahl der Instruktionen reduziert werden. Diese Möglichkeit wird aber nicht berücksichtigt.

Alg.	$a^2 \cdot xlr^2$	xls^2	a^2	xlr^2	$xls \cdot xlr$	$xls \cdot a$	$xlr \cdot a$	xls	a
KKF	2	1	98	2	1	2	1	-2	-2
KKFMF	2	1	103	4	2	4	2	-4	-4
MOR	2	1	54	4	2	4	2	-4	-4
NSSDMF	2	2	109	3	2	4	2	-4	-4
SSD	2	2	6	2	0	0	0	0	0
SAD	3	0	3	0	0	0	0	0	0
SSDMF	2	2	12	3	1	2	1	-2	-2
SADMf	4	0	6	2	1	2	1	-2	-2
SSDMS	2	1	56	3	2	4	2	-4	-4
SADMS	4	0	48	1	1	2	1	-2	-2
SES1	4.5	11	3	0	0	0	0	16	0
SES2	4.5	11	3	0	0	0	0	16	0
SEK1	4.5	53	3	0	0	0	0	0	0
SEK2	4.5	53	3	0	0	0	0	0	0
UKKF	2	6	3	0	0	0	0	6	0
UKKFMF	2	6	3	2	0	0	0	6	0
NIS	2	$4FD+1$	3	0	0	0	0	0	0
NA1	2	12	50	0	1	2	1	-2	-2
NA2	2	12	53	1	1	2	1	-2	-2

Tab.9: Anzahl notwendiger Instruktionen der Korrelationsverfahren für die 'brute force' Implementationen aufgespalten in die verschiedenen Terme⁵⁵. Die Suche nach der Position des Extremwerts ist mit $3 \cdot a^2$ Operationen in diesen Zahlen eingeschlossen. Wurzeln sind mit dem Multiplikator 49, Divisionen mit dem Multiplikator 43 gewichtet worden. FD: Kantenlänge des Laplace-Gauss-Filters.

Bei den folgenden Analysen ist auch zu berücksichtigen, dass z.B. binäre Operationen gleich gewichtet werden wie Fließkommaoperationen. Dies ist natürlich keine faire Betrachtungsweise. Jede andere Gewichtung erscheint aber ebenfalls willkürlich und von der betrachteten Hardware abhängig. Es gilt auch zu bedenken, dass gewisse Operationen durch Einsatz von Lookup-Tabellen gegenüber den betrachteten Implementationen wesentlich vereinfacht und beschleunigt werden könnten (z.B. **SSD**, **SAD**).

⁵⁵ $a = xls \cdot xlr + 1$: Anzahl Korrelationspositionen (nur eine Richtung); xls : Suchfenster-Kantenlänge; xlr : Muster-Kantenlänge.

Alg.	$a^2 \cdot xlr^2$	xls^2	a^2	xlr^2	$xls \cdot xlr$	$xls \cdot a$	$xlr \cdot a$	xls	a
KKF	1	1	9.5	1	1	2	1	-2	-2
KKFMF	1	1	14.5	3	2	4	2	-4	-4
MOR	1	1	13.5	3	2	4	2	-4	-4
NSSDMF	1	2	20.5	2	2	4	2	-4	-4
SSD	1	2	6	1	0	0	0	0	0
SAD	3	0	3	0	0	0	0	0	0
SSDMF	1	2	12	2	1	2	1	-2	-2
SADMf	4	0	6	2	1	2	1	-2	-2
SSDMS	1	1	18	1	2	4	2	-4	-4
SADMS	4	0	48	1	1	2	1	-2	-2
SES1	3.5	11	3	0	0	0	0	16	0
SES2	4.5	11	3	0	0	0	0	16	0
SEK1	3.5	53	3	0	0	0	0	0	0
SEK2	4.5	53	3	0	0	0	0	0	0
UKKF	1	6	3	0	0	0	0	6	0
UKKFMF	1	6	3	2	0	0	0	6	0
NIS	1	$2FD+1$	3	0	0	0	0	0	0
NA1	1	12	9.5	0	1	2	1	-2	-2
NA2	1	12	12.5	1	1	2	1	-2	-2

Tab.10: Anzahl notwendiger Instruktionen der Korrelationskriterien für die optimierten Implementationen aufgespalten in verschiedene Terme⁵⁶. Diese Zahlen gelten für Signalprozessoren mit der Möglichkeit Multiplikationen und Additionen/Subtraktionen parallel auszuführen. Die Implementationen sind auf Geschwindigkeit und nicht auf minimalen Speicherbedarf optimiert. Bei der **KKF**, **KKFMF**, **MOR**, **NSSDMF**, **NA1** und **NA2** wird einzig die Position des Extremwerts, nicht aber dessen Betrag bestimmt. Nicht eliminierbare Wurzeln werden mit dem Multiplikator 49, Divisionen mit 43 gewichtet.

Die halben Wertungen des Terms $a^2 \cdot xlr^2$ bei den Verfahren nach Seitz entstehen durch die Annahme, dass im Mittel in der Hälfte der Fälle zuerst nicht die kleine, sondern die grosse Winkeldifferenz zwischen Muster- und Korrelationsfenstergradient berechnet wird (2π -kleine_Winkeldifferenz) und darum eine Korrektur nötig wird. Die Detektion dieses Fehlers schlägt ebenfalls mit $2a^2 \cdot xlr^2$ zu Buche (Subtraktion und Vergleich mit Null). Ähnliches kann über die halben Wertungen des a^2 -Terms gesagt werden. Sie entstehen durch die Annahme, dass bei der Extremwertdetektion im Mittel in der

⁵⁶ $a = xls \cdot xlr + 1$: Anzahl Korrelationspositionen (nur eine Richtung); xls : Suchfenster-Kantenlänge; xlr : Muster-Kantenlänge.

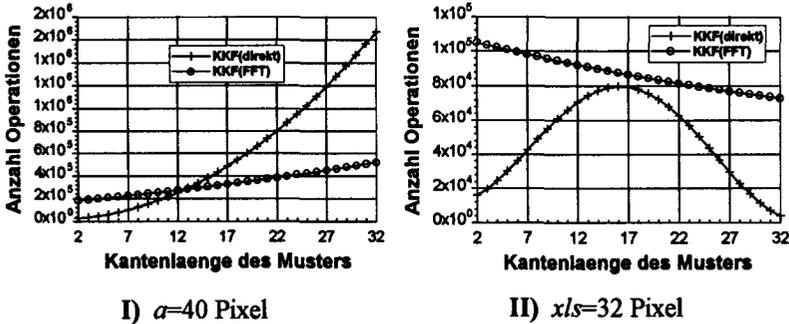
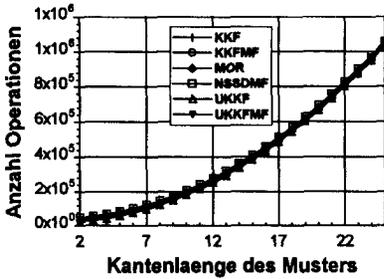
I) $a=40$ PixelII) $xls=32$ Pixel

Fig.22: Rechenaufwand gemessen in Anzahl Instruktionen in Abhängigkeit der Mustergrösse bei konstanter Korrelationsfläche (Fall I) bzw. konstanter Suchfensterfläche (Fall II). *KKF(direkt)*: Optimierte, direkte Berechnung der *KKF* unter Zuhilfenahme aller beschriebenen Massnahmen. *KKF(FFT)*: Berechnung der *KKF* über die Fourier-Transformation. Es ist zu beachten, dass der Rechenaufwand für die Berechnung über die FFT eigentlich nur für Suchfenster mit Kantenlängen, die ganzzahligen Zweierpotenzen entsprechen, definiert ist!

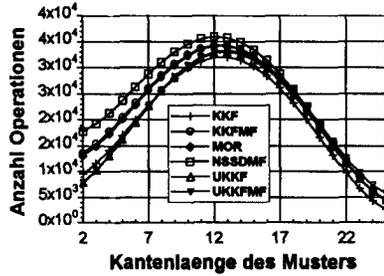
Hälfte der erforderlichen Korrelationswertvergleiche der Entscheid zugunsten eines neuen, temporären Extremums ausfällt und somit eine Zwischenspeicherung erforderlich wird (siehe Fig.19, gilt nur für: **KKF**, **KKFMF**, **MOR**, **NSSDMF**, **NA1** und **NA2**).

In den Figuren 22.I bis 25.I sind die theoretisch bestimmten Rechenlasten aller untersuchten Korrelationskriterien bei fixer Korrelationsfläche (40x40 Pixel) festgehalten. Die Grafiken zeigen deutlich, dass der wesentliche Verlauf durch den Term $a^2 \cdot x \cdot r^2$ bestimmt ist. Dank der Verwendung der parallelen Instruktionen kann gerade der Beitrag dieses Terms reduziert werden.

Bei der **SAD**, der **SADMF**, der **SADMS** sowie den entsprechenden Verfahren, die auf richtungscodierten Bildern arbeiten, kann diese Möglichkeit nicht ausgenützt werden, da die entsprechenden Operationen (z.B. Aufsummieren, Differenz und Absolutbetrag bilden) nicht parallelisierbar sind. Sie schneiden darum in diesem Vergleich sehr schlecht ab. Diese Algorithmen können aber z.B. mit Hilfe von Look-up-Tabellen schnell und einfach in Hardware realisiert werden ([32]). Die Figuren 22.II bis 25.II zeigen schliesslich die Rechenlasten in Abhängigkeit der Mustergrösse bei fixer Suchfenstergrösse. Hier wird deutlich, dass sehr kleine und sehr grosse Muster – verglichen mit der Suchfenstergrösse – einen wesentlich kleineren Aufwand bedeuten als Muster mit Kantenlängen in der Gegend der halben

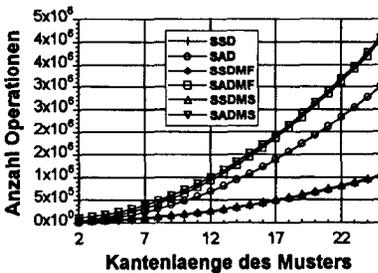


I) $\alpha=40$ Pixel

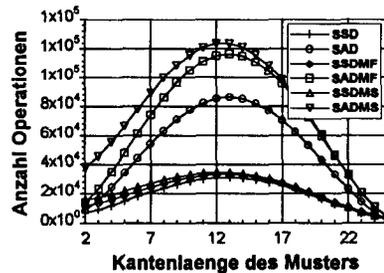


II) $xls=32$ Pixel

Fig.23: Rechenaufwand gemessen in Anzahl Instruktionen in Abhängigkeit der Mustergröße bei konstanter Korrelationsfläche (Fall I) bzw. konstanter Suchfensterfläche (Fall II) für die optimierten Implementierungen der **KKF, KKFMF, MOR, NSSDMF, UKKF und UKKFMF**.



I) $\alpha=40$ Pixel



II) $xls=32$ Pixel

Fig.24: Rechenaufwand gemessen in Anzahl Instruktionen in Abhängigkeit der Mustergröße bei konstanter Korrelationsfläche (Fall I) bzw. konstanter Suchfensterfläche (Fall II) für die optimierten Implementierungen von **SSD, SAD, SSDMF, SADMf, SSDMS und SADMS**.

Suchfensterseite. Der Initialisierungsaufwand fällt allerdings bei kleinen Mustern mehr ins Gewicht als bei grossen.

Diese Untersuchungen zeigen, dass die meisten Korrelationsverfahren sehr effizient auf Signalprozessorsystemen implementiert werden können. Die Betrags-Differenzmasse sind allerdings für diese Architektur weniger geeignet, da keine Parallelisierung der erforderlichen Operationen vorgenommen werden kann. Die binären Korrelationsverfahren schneiden eben-

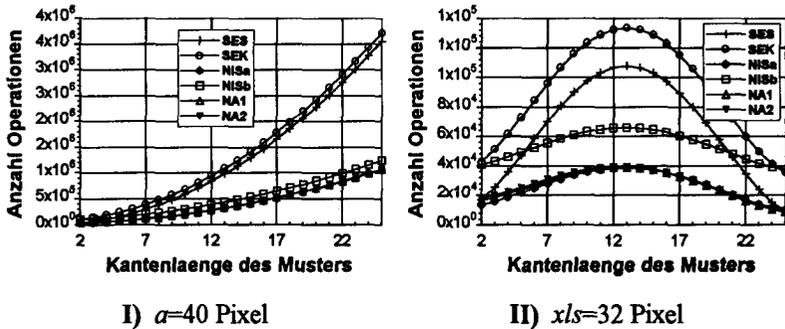


Fig.25: Rechenaufwand in Instruktionen in Abhängigkeit der Mustergröße bei konstanter Korrelationsfläche (Fall I) bzw. konstanter Suchfensterfläche (Fall II) für die optimierten Implementierungen von **SES1** (bzw. **SES2**), **SEK1**, (bzw. **SEK2**), **NIS** ($w=2$, **NISa**), **NIS** ($w=8$, **NISb**), **NA1** und **NA2**.

falls relativ schlecht ab. Der Vorteil der binären Daten kann nur von speziell geeigneter Hardware genutzt werden! Bei den Verfahren, die auf Richtungsbildern arbeiten, ist besonders die Implementation des Richtungsfilters von entscheidender Bedeutung für den Rechenaufwand. Eine relativ einfach zu realisierende, hardwaremässige Lösung für die beiden betrachteten Operatoren kann darum unter Umständen von Vorteil sein ([25]).

7. Diskussion von Beiträgen aus der Literatur

Eine ganze Reihe von Autoren beschäftigt sich mit dem Vergleich von verschiedenen Ähnlichkeitsmassen in der Bildanalyse. Die Berichte sind aber meist relativ kurz, decken häufig nur das Verhalten bei Gauss'schem Rauschen ab, testen nur eine sehr beschränkte Auswahl von Algorithmen oder geben keine Zahlen über den Rechenaufwand der betrachteten Verfahren. Trotz dieser Mängel erlaubt das Studium der Literatur aber den Vergleich von Resultaten aus fremden Untersuchungen mit den hier vorliegenden Ergebnissen. Neben einigen ebenfalls vergleichenden Arbeiten sollen auch Artikel, die hier nicht eingehend diskutierte Korrelationskriterien vorstellen, kurz erwähnt werden. Um die Verständlichkeit der Literaturdiskussion zu verbessern, wurden alle Aussagen in die hier definierte Begriffswelt übersetzt.

Gelegentlich werden in diesem Kapitel auch Aussagen zitiert, die nicht direkt mit der hier vorliegenden Untersuchung in Verbindung stehen. Diese Bemerkungen erlauben aber einen besseren Überblick über weitere Möglichkeiten der flächenbasierten Ähnlichkeitsmasse zu erhalten. Zudem beeinflussen teilweise nur heuristisch begründete 'Details' die Leistungsfähigkeit eines Korrelatorsystems vielfach derart stark, dass sie nicht unerwähnt bleiben dürfen.

Ryan ([138]) und in gekürzter Fassung auch Ryan und Hunt ([139]) geben eine theoretische Herleitung der **KKFMF**, der **SSDMF** und der **UKKFMF** (ohne Filterung). Alle diese (Un-) Ähnlichkeitsmasse realisieren unter der Annahme verschiedener Bildstörungsmodelle⁵⁷ einen Maximum Likelihood Korrelator. Als relevante Bildstörungen werden sowohl lineare Helligkeitstransformationen (additive und multiplikative Intensitätsänderungen), Gauss'sches Rauschen als auch geometrische Bildtransformationen, wie sie in Stereo-Applikationen vorkommen, berücksichtigt.

Zwei Arten von Positionierungsfehlern werden im Beitrag von Ryan und Hunt unterschieden: Der *local registration error* und der *false acquisition*

⁵⁷ Im Fall **KKFMF**: Additive und multiplikative Helligkeitsänderungen, mittelwertfreies, Gauss'sches Helligkeitsrauschen.

Im Fall **SSDMF**: Additive Helligkeitsänderungen, Gauss'sches Helligkeitsrauschen.

Im Fall **UKKFMF**: Additive und multiplikative Helligkeitsänderungen, mittelwertfreies Gauss'sches Helligkeitsrauschen. Korrelationsfensterenergie konstant!

error. Der local registration error ist ein kleiner Fehler, der durch das Rauschen der Korrelationsfunktion hervorgerufen wird. Weil im Prinzip die richtige Korrelationsspitze gefunden wurde, ergibt sich nur eine geringe Diskrepanz zwischen Soll- und Ist-Position. Der false acquisition error hingegen stellt einen größeren Fehler dar, der durch die Detektion einer völlig falschen Korrelationsspitze erfolgt.

Die von den Autoren kurz rapportierte Differenz zwischen den theoretischen und den experimentellen Resultaten ist relativ gross, da bei der theoretischen Untersuchung erstens von konstanter Korrelationsfensterenergie ausgegangen wird⁵⁸ und zweitens kleine Muster (7x3, 15x3 und 23x3 Pixel) die statistischen Annahmen in der Praxis nur schlecht erfüllen können. Gute, experimentell ermittelte Resultate liefern nur die **KKFMF** und die **SSDMF**⁵⁹.

Das zentrale Ziel dieser Arbeit war aber nicht der Vergleich von Korrelationskriterien sondern die Definition eines Gütemasses, das eine zuverlässige Voraussage über die Eignung von Korrelationsmustern zulässt und zudem die Wahl des bestgeeigneten, der gegenwärtigen Szene optimal angepassten Korrelationskriteriums unterstützt. Verschiedene zum grössten Teil kontrastabhängige Masse wurden mit wechselndem Erfolg geprüft.

Auch der Rechenaufwand der einzelnen Verfahren wird geschätzt. Es werden zwar verschiedene Umformungen zur Reduktion der Rechenlast vorgenommen, da aber die Rechenlast pro Korrelationsfenster und nicht pro Suchfenster bestimmt wird und verschiedene Vereinfachungsmöglichkeiten ausser Acht gelassen werden, können diese Resultate nicht direkt mit den hier vorliegenden Ergebnissen verglichen werden.

Einzig die Kovarianzfunktion (**UKKFMF** ohne Filterung) wird bezüglich des Verhaltens bei Rauschen und reduzierter Rechengenauigkeit detailliert theoretisch untersucht. Daneben wird aber auch das Verhalten bei geometrischen Suchfensterverzerrungen betrachtet. Als Vorbeugung im Hinblick auf entsprechende Probleme wird die Verwendung kleiner Muster oder die geometrische Rücktransformation des Musters (bei bekannter Verzerrung) vorgeschlagen.

Burt, Yen und Xu ([33]) berichten über experimentelle Untersuchungen verschiedener Korrelationsalgorithmen. Getestet wurden die **KKFMF**, die **UKKF**, die **UKKF** auf binarisierten Daten, die unnormierte Kreuzkorrelati-

58 $\left\{ \sum_{v=0}^{M-1-u} \sum_{u=0}^{M-1-u} s^2(x+u, y+v) \right\} = \text{konstant!}$ Dies trifft insbesondere an Bild-Diskontinuitäten nicht zu.

59 Die Rangierung dieser beiden Verfahren ist allerdings im Falle von verrauschten Bildern gerade umgekehrt zu den eigenen Resultaten!

onsfunktion und die lokal mittelwertfreie, unnormierte Kreuzkorrelationsfunktion (beide ohne Filterung). Die Testbilder enthalten eindimensionale, mit einem Zufallszahlengenerator erzeugte (und damit eher unrealistische) Information. Als Störungen werden Rauschen, Mittelwert- und Kontrastverschiebungen sowie die Bildskalierung eingeführt. Sowohl das Muster wie auch das Korrelationsfenster werden mit Fensterfunktionen gewichtet, die die Information am Rand abschwächen und das Korrelationskriterium auf das Musterzentrum konzentrieren sollen. Leider wird mit sehr kleinen (natürlich ebenfalls eindimensionalen) Suchfenstern, die nur drei unterscheidbare Positionen enthalten, gearbeitet. Dies, sowie die Verwendung von eindimensionalen Zufallszahlen, macht die Versuche nur schwer mit den hier vorliegenden Ergebnissen vergleichbar.

Die Gauss'sche Gewichtung der Muster- und Korrelationsfenster-Pixel zeigt gegenüber der üblichen Gleichgewichtung der Pixelwerte besonders bei der Bildgrössenskalierung (Reduktion der Grösse bis zu 75% der Ursprungsgrösse) Vorteile. Dies ist einleuchtend, da dann vom Transformationszentrum weit entfernte und damit stark veränderte Pixel wegen ihrer geringen Gewichtung nur einen kleinen Einfluss auf das Ergebnis haben. Den gleichen Effekt kann man auch bei der Reduktion der Mustergrösse feststellen.

Die **KKFMF** schneidet bei allen Versuchen am Besten ab. Die Versuche von Burt zeigen aber, dass z.B. die **UKKF** oder die lokal mittelwertfreie, unnormierte Kreuzkorrelationsfunktion nur geringfügig schlechter sind. Dies stimmt mit den hier vorliegenden Ergebnissen überein, solange nur grössere Muster betrachtet werden.

Hannah ([67]) gibt einen sehr guten und empfehlenswerten Überblick über verschiedene Korrelationsalgorithmen (**KKF**, **KKFMF**, **SSD**, **SSDMF**, **UKKF**, **SAD** und **SADMF**) und damit zusammenhängende Verfahren. Dabei werden der Rechenbedarf wie auch verschiedene Methoden zur Reduktion des Rechenbedarfs und zur Steigerung der Genauigkeit diskutiert. Der Vergleich der Rechenlasten (z.B. **SSD** und **KKF**) ist jedoch nicht mehr unbedingt repräsentativ für heutige Rechnerarchitekturen. Die Schlussfolgerungen bezüglich der Anwendung der FFT sind aber trotzdem in guter Übereinstimmung mit den eigenen Ergebnissen: Die Anwendung der FFT ist erst bei grösseren Fenstern lohnenswert!

Helava ([68]) berichtet über verschiedene, auch in der vorliegenden Arbeit besprochene Korrelationsverfahren. Anhand der unnormierten Autokorrelation zeigt er, dass die Breite der Korrelationsspitze mit Hilfe des Leistungsspektrums des Suchfensters abgeschätzt werden kann. Das Wie-

ner-Khinchin-Theorem besagt, dass die Autokorrelationsfunktion über die Fourier-Transformation mit dem Leistungsdichtespektrum verknüpft ist (unendliche Fensterausdehnung vorausgesetzt). Bei guter Bildqualität, d.h. geringen Abweichungen zwischen Muster- und Suchfensterbild, ist die Kreuzkorrelationsfunktion der Autokorrelationsfunktion sehr ähnlich und damit die Voraussage zuverlässig. Dies gilt natürlich nur für genügend grosse Bildausschnitte.

Der Einfangbereich⁶⁰ (pull in range) ist somit einerseits mit der Breite der Korrelationsspitze und andererseits mit der höchsten im Bild auftretenden Raumfrequenz verbunden. Für einen grossen Einfangbereich werden unscharfe Bilder mit geringem Detailgehalt benötigt. Für grosse Genauigkeit, d.h. schmale Korrelationsspitze, sollte hingegen mit maximaler geometrischer Auflösung gearbeitet werden. Bei grossem Suchbereich drängt sich also ein hierarchisches Vorgehen auf, das zuerst auf unscharfen Bildern die grobe Position bestimmt und dieses Ergebnis in zunehmend besser aufgelösten Bildern verfeinert. Anhand von verschiedenen Spektren wird die Auswirkung auf die Form der Korrelationsfunktion gezeigt.

Daneben werden auch unterschiedlich feine Helligkeitsauflösungen betrachtet (siehe auch [92]). Im Extremfall von nur zwei Helligkeitsstufen muss, für gleiche Detektionsraten, eine viel grössere Musterfläche berücksichtigt werden als bei feinerer Intensitätsauflösung (bis zu 72 Mal grössere Fläche, siehe dazu auch Kuhnert [92], weiter unten). Da meist eine Punktvermessung durchgeführt werden soll, ist die Mustergrösse sehr klein zu wählen und diese Bedingung somit nur schwer zu erfüllen. -Mit sechzehn Helligkeitsstufen ist die Leistungsfähigkeit sehr nahe dem Optimum. Diese Zahlenwerte sind aber mit Vorsicht zu betrachten. Sie gelten nur, falls die vorausgesetzte Helligkeitsstatistik im untersuchten Bildbereich erfüllt wird.

Ein kurzer Vergleich der Rechenlast bei direkter Berechnung der unnormierten Kreuzkorrelationsfunktion mit der Berechnung über die FFT zeigt, dass der Aufwand bei direkter Rechnung, einem Muster von 8x8 Pixeln und einem Suchfenster von 16x16 Pixeln etwa acht Mal kleiner ist.

Silver ([148]) beschreibt die in Hardware implementierte **KKFMF**. Er gibt einen knappen Überblick über die Eigenschaften dieser Funktion. Dabei wird nicht unbedingt von rechteckigen oder zusammenhängenden Mustern ausgegangen. Ein solch ungewöhnliches Vorgehen hat unter Umständen Vorteile, da bei richtiger Anwendung nicht-informationstragende Bildteile

⁶⁰ Einfangbereich: In dieser Region rund um das Korrelations-Maximum (bzw. -Minimum) steigen die Korrelationswerte bei Annäherung an die Position des Extremwerts stetig an.

keinen Beitrag zum Rechenbedarf leisten und störende Vorder- und Hintergrundinformation eliminiert werden kann. Dafür kann aber auch die Genauigkeit der Resultate reduziert werden, falls gut strukturierte Musterregionen ausgeblendet werden.

Die Invarianz der **KKFMF** bezüglich additiven und multiplikativen Helligkeitsänderungen im Bild wird besonders hervorgehoben. Die Genauigkeit der **KKFMF**, zusammen mit einer nicht spezifizierten Interpolationsmethode, wird mit 0.1 bis 0.25 Pixel angegeben. Die Abhängigkeit der Resultate vom Signal-zu-Rauschleistungsverhältnis scheint im betrachteten Gebiet (20..4) relativ gering zu sein. Allerdings wird aus dem Artikel nicht ganz klar um welche Art Störung es sich dabei handelt (und ob das Leistungsverhältnis in Dezibel gemessen wurde); für die Bildaufnahme wurde sowohl die Kameraverstärkung als auch die Beleuchtung variiert (entspricht vermutlich am ehesten der Bildstörung **IRIS**). Auch Verdeckungen des Musters sollen keinen wesentlichen Einfluss auf die Leistung dieses Korrelationsalgorithmus' haben.

In diesem Artikel wird ausserdem kurz auf das Lokalisieren von verdrehten Mustern eingegangen. Bei den beschriebenen Versuchen wurde das Muster bis zu einer Verdrehung von 25° noch zuverlässig detektiert. Es wird aber darauf hingewiesen, dass keine feinen Details im Bild vorhanden sein dürfen und symmetrische Muster in diesem Fall robuster sind. Als Abhilfe bei Problemen mit geometrisch transformierten Bildern wird unter anderem empfohlen kleinere Muster zu verwenden. Ein grosses Muster kann z.B. in mehrere kleine unterteilt werden (siehe auch [2] und [12]).

Als weitere gute Eigenschaften der **KKFMF** werden die ausgezeichnete Diskriminierungsfähigkeit und die Wiederholbarkeit erwähnt. Leider wird aber, wie bei einigen anderen Aussagen, keine Begründung theoretischer oder experimenteller Art gegeben.

Wie Helava ([68]) benutzt auch Silver die Autokorrelationsfunktion zur Schätzung der erwarteten Kreuzkorrelationsfunktion.

Zur Reduktion des Rechenaufwands wird die **KKFMF** ausmultipliziert (siehe Gl.23 auf Seite 85), vorzeichenbehaftet quadriert und der negative Wertebereich – weil in der Regel uninteressant – abgeschnitten, d.h. auf Null gesetzt. Durch das Quadrieren wird der interessierende Wertebereich nahe Eins auseinandergezogen und damit die Unterscheidung nahe beieinanderliegender Korrelationswerte vereinfacht (siehe auch [117]). Da die Extremwertsuche gemäss einem der *hill climbing* Methode ähnlichen Verfahren gemacht wird, können keine weiteren Vereinfachungen bei der Berech-

nung der Korrelation vorgenommen werden. Dafür muss aber nur ein Teil der Korrelationswerte berechnet werden⁶¹.

Insgesamt werden in diesem Artikel einige interessante Themen angesprochen und praktische Tips für die Verwendung der **KKFMF** gegeben. Die von Silver dokumentierten Resultate stehen in guter Übereinstimmung mit den Ergebnissen der vorliegenden Arbeit.

Im Artikel von Anandan ([6]) wird die **KKFMF** mit der **SSD** verglichen. Beide Korrelationskriterien werden auf Bilder angewandt, die mit dem Laplaceoperator vorverarbeitet wurden. Bei diesen Tests schnitt die **SSD** (mit Filterung!), im Gegensatz zu den vorliegenden Ergebnissen, besser ab als die **KKFMF** (ebenfalls mit Filterung!). Den Bildern wurden bei den Versuchen Gauss'sches Rauschen⁶² überlagert. Die Mustergrösse wurde auf maximal 8x8 Pixel beschränkt. Da diese Untersuchungen nur auf einem einzigen Bild durchgeführt wurden (MANDRILL), ist schwierig abzuschätzen, ob diese Resultate auch auf andere Szenen übertragbar sind. Sicher ist aber eine Hochpassfilterung, bei günstiger Bildqualität ohne allzu starkes Intensitätsrauschen und ohne geometrische Bildtransformationen, für die Positionierung von Vorteil, da sie das Hauptaugenmerk von gleichförmigen, nur ungenau lokalisierbaren Flächen weg auf dominante Kanten lenkt.

Faugeras, Fua, Hotz, Ma, Robert, Thonnat und Zhang ([52], siehe auch [75]) vergleichen die **KKF**, die **KKFMF**, die **NSSDMF** und ein damit verwandtes, mittelwertbehaftetes Unähnlichkeitsmass, die (mittelwertbehaftete) normierte Summe der Differenzenquadrate (Gl.31, **NSSD**), im Zusammenhang mit Stereoanwendungen⁶³.

$$\text{NSSD}(x, y) = \frac{\sum_{v=0}^{\text{VMAX}} \sum_{u=0}^{\text{HMAX}} \{r(u, v) - s(x+u, y+v)\}^2}{\sqrt{\sum_{v=0}^{\text{VMAX}} \sum_{u=0}^{\text{HMAX}} r^2(u, v) \cdot \sum_{v=0}^{\text{VMAX}} \sum_{u=0}^{\text{HMAX}} s^2(x+u, y+v)}} \quad (31)$$

⁶¹ Dieses Verfahren geht von einer gewissen 'Weichheit' der Korrelationsfunktion aus und findet deshalb nur bei Einhaltung dieser Bedingung den richtigen Extremwert.

⁶² Standardabweichung: 0.5 und 10% der maximalen Aussteuerung.

⁶³ Beim Vergleich mit den Resultaten von Faugeras ist darauf zu achten, dass die Nomenklatur ähnlich derjenigen von Hannah ([67]) ist und nicht mit der hier benutzten Nomenklatur übereinstimmt. Z.B. wird die normierte Kreuzkorrelationsfunktion (hier mit dem Akronym **KKF** bezeichnet) in [52] unnormierte Kreuzkorrelationsfunktion genannt. Dieser Name wird aber in der vorliegenden Arbeit für die **UKKF** benutzt. Unter *Normierung* versteht Faugeras et al. die lokale Mittelwertbefreiung.

Allgemein, d.h. unabhängig vom Korrelationsalgorithmus, kann gesagt werden, dass die Wahrscheinlichkeit für eine Fehlpositionierung mit steigendem Texturgehalt der Szene und wachsender Mustergrösse abnimmt (Signal-zu-Störvarianzenverhältnis! Siehe auch [126]). Grosse Muster sollen allerdings nur eine geringe Genauigkeit ergeben. Diese Aussage kann, nach meinen Erfahrungen, höchstens auf geometrisch transformierte Bilder oder in der Tiefe stark gestaffelte Szenen zutreffen. In allen anderen Fällen ist bei zunehmender Mustergrösse mit besserer Leistung zu rechnen.

Die mittelwertbefreiten Korrelationskriterien schneiden etwas besser ab als die beiden Vergleichsalgorithmen, wobei aber nur die **NSSD** wesentlich von der Leistung der anderen Verfahren abweicht. Diese Korrelationsfunktion erreicht bei weitem die schlechtesten Resultate. Mit der **NSSDMF** werden bei den Versuchen mit verschiedenen Mustergrössen (3x3 bis 19x19 Pixel) die besten Resultate erzielt. Dies steht in deutlichem Widerspruch zu den eigenen Untersuchungen und Ergebnissen. In der detaillierteren Beschreibung von Hotz ([75]) erscheinen die Gegensätze sogar noch grösser! Die **NSSDMF** schneidet beispielsweise bei krassen Helligkeitsveränderungen besser ab als die **KKFMF**. Dies ist auch vom theoretischen Standpunkt her nicht einzusehen. Die **KKFMF** zeichnet sich durch Störungsinvarianz sowohl bezüglich additiven wie auch multiplikativen Helligkeitsänderungen aus. Die **NSSDMF** ist aber mit Sicherheit sensitiv gegenüber multiplikativen Intensitätsverschiebungen. Die von Hotz eingeführten Bildstörungen sind darum, entgegen seinen Angaben, vermutlich nur durch Veränderung des Offset- und nicht auch des Verstärkungswerts am Analog/Digital-Videoeingangswandler des Bildverarbeitungssystems realisiert. Diese für alle Pixel gleichermaßen geltende (analoge) Helligkeitstransformation muss natürlich als additive Helligkeitsveränderung modelliert werden, die durch die lokale Mittelwertbefreiung gut kompensiert werden kann.

Korrelationsalgorithmen der betrachteten Art eignen sich nach Faugeras' Meinung gut bei texturhaften Szenen, während merkmalsbasierte Verfahren eher bei Bildern, die durch Flächen und Kanten charakterisiert werden können, einzusetzen sind. Dies ist einleuchtend, da sich texturhafte Bildteile nur schwer durch (wenige) Merkmale eindeutig von gleichartigen Texturregionen unterscheiden lassen. Bei Bildern, die vorwiegend durch Kanten charakterisiert werden können (Grafiken, CAD-Modellansichten, Bilder mit prägnanten Kanten und relativ uniformen Flächen) sind bei konventionellen Korrelationskriterien und kleinen Mustern Uneindeutigkeiten zu erwarten, weil alle Korrelationsfenster mit Kanten gleicher Ausrichtung sehr ähnliche Korrelationswerte ergeben.

Als interessante Erweiterung des Extremwertdetektors wird zudem die Vor-Rückwärtskorrelation zur robusten Lösung des Korrespondenzproblems bei Stereosichtsystemen vorgestellt: Zuerst wird in einem von zwei Stereobildern ein Muster gewonnen. Dieses Muster wird mit dem zweiten Bild entlang der *epipolar line*⁶⁴ korreliert. Am Ort maximaler Übereinstimmung wird anschliessend im zweiten Bild ein weiteres Muster extrahiert, das schliesslich wieder mit dem ersten Bild korreliert wird. Falls die Extremwertposition dieser zweiten Korrelation nicht mit der Position des ersten Musters übereinstimmt, wird das Positionierungsergebnis als unzuverlässig eingestuft und für die weiteren Betrachtungen verworfen. -Dieses Verfahren ist allerdings schon wesentlich früher von Nishihara bzw. Grimson erwähnt worden ([125]).

Aggarwal, Davis und Martin ([2]) untersuchen die effiziente Implementation von Ähnlichkeitsmassen für Tracking-Systeme. Die **KKF**, die **SSD**, die **SAD** und die **UKKF** werden betrachtet. Bei konstanter lokaler Energie im Suchfenster sind diese Korrelationsmasse, abgesehen von der **SAD**, in Bezug auf die Position des Extremums identisch. Allfällig verwendete Vorverarbeitungsalgorithmen haben beim Verfolgen von Objekten vor strukturiertem Hintergrund die Aufgabe den Einfluss des Hintergrundes zu minimieren.

Zur Beschleunigung der Korrelationsrechnung kann spezielle Hardware (z.B. SIMD- und MIMD-Rechner) eingesetzt werden. Die Autoren stellen ein der *sequential similarity detection* angelehntes, verallgemeinertes Verfahren zur Rechenzeitreduktion vor (siehe auch [174]). Dieser Algorithmus bestimmt einzig die Position und den Wert des Korrelationsextremums, nicht aber die anderen Korrelationswerte. Ohne Berücksichtigung des der Methode eigenen, erheblichen Overheads verursacht sie den minimal notwendigen Rechenaufwand zur direkten Bestimmung der Extremwertposition, da nur das Korrelationsmass an der Extremwertstelle fertig berechnet wird. Dieses Rechenverfahren kann aber nur bei rein akkumulativen Korrelationsmassen wie der **SSD**, der **SAD** und der **UKKF** verwendet werden (z.B. nicht bei der **KKF**). Die Bestimmung der Gesamtsumme wird in mehreren Schritten vorgenommen. Zuerst wird das Suchfenster nur mit einem kleinen Teil des Musters korreliert. Nach dieser Initialisierung wird die Summe pro Iterationsschritt jeweils nur noch für das günstigste Zwischenresultat, durch Einbezug weiterer Musterpixel, erweitert. Sobald an einem Ort alle Musterpixel in der Korrelation verwendet worden sind, wird die Rechnung abge-

⁶⁴ Epipolar line: In die zweite Sensorebene transformierter Abbildungsstrahl des ersten 2D-Sensors.

brochen, da an diesem Ort das gesuchte Extremum auftritt. Dieses Vorgehen garantiert die fehlerlose Detektion des Extremwertes. Das zweite Verfahren, das diskutiert wird, teilt das Muster in mehrere Unter-Regionen auf. Das jeweils nächste Unter-Muster wird nur noch an den, durch das vorherige Teil-Muster bestimmten Orten grösster Ähnlichkeit korreliert. Das Ergebnis dieser Rechnung, d.h. die Position des gefundenen Gesamt-Extremums, stimmt nicht notwendigerweise mit demjenigen der normalen Rechnung überein. Die Aufteilung des Musters in verschiedene Teilgebiete muss deswegen optimiert werden. Kleine Flächen reduzieren den Rechenaufwand; die Wahrscheinlichkeit, dass die berechnete Extremwertposition falsch ist, wird aber erhöht. Ein weiteres, ähnliches Verfahren ist das sog. *coarse-to-fine matching* das auf Bildpyramiden beruht. Pro Iterationsschritt wird die Position des Objekts nur noch in der Nähe des vorherigen Resultats, das mit geringerer geometrischer Auflösung erzielt wurde, gesucht.

Venot, Lebrouche und Roucayrol ([164]) beschreiben ein originelles Korrelationsverfahren, das die Vorzeichenwechsel von Differenzbildern (Korrelationsfenster minus Muster) auswertet (*stochastic* bzw. *deterministic sign change* Verfahren, **SSC**, **DSC**). Bei genügend verrauschten Bildern⁶⁵ sind, unter der Annahme von Gauss'schem Rauschen, $\frac{N}{2} - 1 \pm 1.96 \sqrt{\frac{N}{4}}$ Vorzeichenwechsel zu erwarten (N : Anzahl Pixel im Referenzfenster). Die Objektposition wird also an derjenigen Stelle detektiert, die diesem Wertebereich am nächsten kommt. Die Abweichung zwischen theoretischem und gemessenem Wert kann zusätzlich als Gütemass verwendet werden. Dieses Verfahren wird mit der **SAD**, der **KKF** und der unnormierten Kreuzkorrelationsfunktion im Zusammenhang mit medizinischen Röntgenaufnahmen verglichen. Alle Vergleichsalgorithmen schneiden in den drei Experimenten (einmal mit deutlicher Verdeckung) schlechter oder bestenfalls gleich gut wie die neu eingeführten Rechenverfahren ab. Eine besonders schlechte Leistung weist natürlich die unnormierte Kreuzkorrelation auf (ohne Vorverarbeitung!). Die Mustergrösse ist mit 6000 bis 9000 Pixeln vergleichsweise gross. Die Autoren bemerken, dass dieses Ähnlichkeitsmass sehr robust sei. Leider wird nicht klar, gegenüber welchen Einflüssen dieser Algorithmus robust ist. Einfache Überlegungen lassen aber eigentlich eine relativ grosse

⁶⁵ Bei zu geringem natürlichen Rauschen (unterhalb der Quantisierungsschwelle des Analog-Digital-Wandlers) kann durch Addition eines einfachen pseudo-Rauschmusters eine entsprechende Situation geschaffen werden.

Sensitivität gegenüber geometrischen Transformationen⁶⁶ und Helligkeitsverschiebungen⁶⁷ erwarten. Ganz sicher wird dieses Verfahren nur bei sehr grossen Referenzfenstern erfolgreich einsetzbar sein, da eine genügend grosse Zahl Ereignisse (Vorzeichenwechsel) in einem Korrelationsfenster beobachtet werden muss.

Chou und Chen ([39]) berichten über ein neues Matching-Verfahren mit geringem Rechenaufwand: Das sog. *moment-preserving pattern matching* (**MPPM**). Grundlage dieses binären Korrelationsalgorithmus', dessen Spezialität die aussergewöhnliche Binarisierung darstellt, ist die Annahme, dass die betrachtete Szene ein bimodales Histogramm aufweist. D.h. es sind prinzipiell zwei Helligkeitsklassen, Objekt und Hintergrund, unterscheidbar. Durch Anwendung geeigneter, von Intensitäts-Mittelwert und -Standardabweichung abhängigen (unteren wie auch oberen) Schwellwerten können so jeweils zwei Binärbilder mit gleicher Helligkeitsstatistik wie das Originalbild generiert werden. Gesetzte Pixel zeigen in den entsprechenden Binärbildern die Zugehörigkeit zur jeweiligen Helligkeitsklasse an⁶⁸. Die Summe der Resultate der beiden Korrelationsfunktionen, die mit den zwei binären 'Exklusiv-'Mustern berechnet werden, ergibt schliesslich das Ähnlichkeitsmass, das weiter ausgewertet wird. Als Korrelationsfunktion wird die unnormierte Kreuzkorrelationsfunktion verwendet, deren Operationen im binären Fall ja zu einfachen logischen UND-Operationen und Inkrementen degenerieren.

Das **MPPM**-Verfahren wird mit der **KKF**, der **SAD** dem *stochastic sign change Verfahren* (Vorzeichenwechsel-Verfahren, **SSC**, siehe Venot et al. [164], oben), der unnormierten Kreuzkorrelationsfunktion (ohne Filterung) sowie einem nicht näher spezifizierten Ähnlichkeitsmass in Verbindung mit einem binarisierenden Hochpassfilter bei additivem wie auch multiplikativem Gauss'schen Rauschen verglichen. Die **SSC** wird nicht weiter verfolgt, da ihre Rauschsensitivität sehr gross sei!

⁶⁶ Da das Korrelationsfenster und das Muster nur noch schlecht zur Deckung gebracht werden kann, enthält das Differenzbild nicht nur die Rauschkomponente des Bildsignals, sondern auch eigentliche Signal-Intensitätsdifferenzen.

⁶⁷ Eine additive Helligkeitsveränderung in der Grössenordnung der Standardabweichung des Rauschens – was bei kleinem Rauschen sehr schnell z.B. durch Beleuchtungsänderungen auftreten kann – vermag die Rauschstatistik und damit die Zahl der Vorzeichenwechsel wesentlich zu beeinflussen und kann dadurch die Positionierung erheblich stören.

⁶⁸ Ungesetzte Pixel zeigen nur an, dass sie nicht zur betrachteten Klasse gehören, die Zugehörigkeit zur anderen Klasse wird damit aber nicht impliziert.

Im Unterschied zu den hier vorgestellten Versuchen wird nicht das Suchfenster, sondern das Muster verwechselt. Als Ergebnis zeigt sich, dass die **KKF** bei additivem Rauschen die beste Leistung (mittlere Abweichung von der korrekten Position) ergibt. Dicht gefolgt wird sie von der **SAD**, die wiederum einen grösseren Vorsprung auf den **MPPM**-Algorithmus aufweist. Das Verfahren, das Hochpass-gefilterte und anschliessend binarisierte Bilder verwendet, zeigt eine genügende, aber zusätzlich reduzierte Leistung. Die unnormierte Kreuzkorrelation weist hingegen eine völlig ungenügende Leistung auf.

Die Leistung der **KKF** und der **SAD** sind genau umgekehrt zu den in diesem Text dokumentierten Ergebnissen. Dies kann einerseits von den gewählten Mustern⁶⁹ oder andererseits von der unklaren Definition der **KKF** herühren (**KKFMF?**). Zudem gilt es zu beachten, dass die Versuche von Chou und Chen zwar rangierungsmässig ähnliche (aber nicht für alle Tests identische) Resultate liefern und somit nur für diese einzelnen Szenen mit vorteilhaften Histogrammen repräsentativ sind.

Bei multiplikativem Rauschen kann ungefähr die gleiche Situation beobachtet werden wie bei additivem Rauschen. Allerdings scheint dort die **SAD**, die einzelne, grosse Fehler nur schwach gewichtet, besser als die **KKF** abzuschneiden.

Der Vergleich des **MPPM**-Algorithmus' mit der **SAD** in Kombination mit einer Rechenaufwand-reduzierenden Suchstrategie (, die den Suchraum natürlich nur unvollständig abdeckt; Muster, 8x8 Pixel, für block matching Anwendung) erscheint nicht ganz fair und ergibt für den **MPPM**-Algorithmus natürlich die besten Resultate, da nur er jeweils den ganzen Suchraum berücksichtigt.

Interessant sind noch die relativen Rechenzeiten, die erreicht wurden: Die **KKF** benötigt etwa 25 Prozent mehr Rechenleistung als die **SAD** (Sun 3/110) während das **MPPM**-Verfahren ca. 20 Prozent des **KKF**-Aufwandes beansprucht.

Secilla, Garcia und Carrascosa ([144]) beschreiben Versuche mit additiv und multiplikativ verwechselt Bildern (Gauss'sches Rauschen). Als Anwendungsgebiet steht die Elektronenmikroskopie im Vordergrund, die Bilder entsprechend schlechter Qualität liefert. Als Ähnlichkeits- bzw. Distanzmass wurden die folgenden Algorithmen untersucht: Die **SSD**, die **SAD**, die

⁶⁹ Es wurden nur zwei verschiedene Muster der Grösse 24x24 Pixel in einer Korrelationsfläche von 64x64 Positionen mit 100 unterschiedlichen Rauschstörungen betrachtet. Zudem wurde, wie oben schon erwähnt, im Gegensatz zu den hier vorgestellten Versuchen das Muster verwechselt und nicht das Suchfenster.

KKF und die unnormierte Kreuzkorrelationsfunktion (ohne vorherige Hochpassfilterung). Daneben werden noch drei weitere, oben z.T. kurz erwähnte Korrelationskriterien getestet: Das Phasendifferenzverfahren (siehe auch [5], [82] und [84]), ein Algorithmus der die Anzahl der Vorzeichenwechsel analysiert (**SSC** und **DSC**), die Maximum a posteriori Estimation (**MPE**) sowie die Methode der invarianten Momente. Die Phasendifferenzmethode basiert auf Projektionen binärer Repräsentationen von Muster und Suchfenster. Die Phasen entsprechender Spektren weisen bei idealen Verhältnissen eine lineare Differenz, proportional zur Verschiebung des Objekts, auf. Daraus kann die folgende Beziehung abgeleitet und damit die Verschiebung bestimmt werden:

$$F^{-1} \left[\frac{R(\omega_x, \omega_y)}{S(\omega_x, \omega_y)} \right] = \delta(x - \Delta x, y - \Delta y) \quad (32)$$

Das Verfahren, das die Anzahl der Vorzeichenwechsel der Differenzen zwischen Muster und Objektfenster analysiert ([164], siehe oben) liefert schlechte Resultate, da es eine lokal symmetrische Verteilung des Rauschens annimmt, Sättigungseffekte diese Verteilung üblicherweise aber verändern.

Das Verfahren der invarianten Momente gehört eigentlich nicht zu den Korrelationskriterien, ist aber eine weit verbreitete Methode der Bildanalyse. Es liefert als Musterdetektor völlig unbrauchbare Resultate bei verrauschten Bildern. Der Vorteil dieser Methode wäre natürlich die Robustheit bezüglich Skalierung und Rotation.

Bessere Resultate liefert hingegen die interessante Maximum a posteriori Estimation der Objektposition (**MPE**). Sie lehnt sich stark an die **SAD** an, ist aber noch erheblich einfacher mit spezieller Hardware zu realisieren:

$$\text{MPE}(x, y) = \sum_{v=0}^{\text{vmax}} \sum_{u=0}^{\text{umax}} m(r(u, v), s(x+u, y+v)) \quad (33)$$

mit $m(a, b) = \begin{cases} 0 & \text{falls } |a - b| < k \\ 1 & \text{falls } |a - b| \geq k \end{cases}$

Die Konstante k ist eine vom Benutzer einstellbare, den Einfluss der Störung minimierende Konstante.

Die Versuche von Secilla wurden mit künstlich verrauschten Bildern vorgenommen, wobei die Mustergrösse auf 32x32 Pixel und das Suchfenster auf 64x64 Pixel gesetzt wurde. Der Zusammenhang zwischen mittlerer Fehlerdistanz (gerechnete Position \leftrightarrow wahre Position) und Signal-zu-Rausch-

Leistungsverhältnis zeigt, dass die **KKF**, die bei diesen Versuchen über die Fourier-Transformation berechnet wird (siehe auch Kapitel 6.2.3: *Verwendung von Frequenztransformationen*), am besten abschneidet. Besonders betont wird, dass bei der **SAD**, im Gegensatz zur **KKF**, ausschliesslich ganzzahlige Ergebnisse resultieren, was bei entsprechender Hardware natürlich einen wesentlichen Vorteil darstellt. Die unnormierte Kreuzkorrelation ohne Vorverarbeitung ist unbrauchbar, da dieses Ähnlichkeitsmass in Gebieten der höchsten Energie (hell) unabhängig von der Bildstruktur maximal wird. Brauchbar bis gut erscheinen die **SSD** und die **SAD**; als akzeptabel können die Maximum a posteriori Estimation und die Methode der Phasen Differenzen bezeichnet werden. Bei zusätzlichen Versuchen zeigt es sich zudem, dass die Methode der Phasen Differenzen geometrisch sehr genau übereinstimmende Bilder benötigt!

Die gute Bewertung der **KKF**, der **SSD** und der **SAD** stimmen mit den eigenen Versuchen überein, allerdings ist die Reihenfolge der Resultate unterschiedlich.

Ghaffary ([59]) gibt einen knappen Überblick über verschiedene Verfahren aus dem Bereich der merkmalsbasierten Positionssuche und dem Korrelations-Matching. Er beschreibt das Prinzip der *sequential similarity detection*⁷⁰, die bei der **SAD**, der **SSD** und verwandten Verfahren zur Rechenzeitverkürzung eingesetzt werden kann. Neben der unnormierten Kreuzkorrelation (ohne Vorverarbeitung) wird auch kurz die **KKF** angesprochen. Erstaunlich scheint hier die Bemerkung, dass die **KKF** besonders sensitiv gegenüber Rauschen sei! Die hier dokumentierten Versuche, wie auch die Theorie ([138] und [139]), deuten eigentlich nicht auf ein derartiges Verhalten hin.

Seitz ([145]) untersucht fünf verschiedene Richtungsdetektoren im Zusammenhang mit dem RMS-Distanzmass (root mean square). Die Korrelationswerte der **SSD** entsprechen bis auf die Wurzel diesem Distanzmass. Als zusätzliche, kleine Änderung wird hier jeweils für jedes Pixel die kleinste Winkel- anstelle der Intensitätsdifferenz berechnet. Zu den untersuchten Richtungsdetektoren gehören der Roberts- und der Sobel-Operator sowie ein Verfahren, das auf den horizontalen und vertikalen Helligkeitsdifferenzen beruht (ähnlich dem Sobeloperator) ferner eine Methode, die die Richtung einer approximierenden Ebene bestimmt und zu guter letzt ein Verfahren, welches die Orientierung der am besten passenden, lokalen Spiegelachse

⁷⁰ Vorzeitiger Abbruch der Berechnung, falls das Zwischenresultat eine bestimmte Schwelle überschreitet.

berechnet. Alle Verfahren schneiden bei unverrauschten und leicht ver-
rauschten Bildern ungefähr gleich gut ab. Die absolut besten Resultate wer-
den aber vom Sobeloperator erreicht

Als Vorteil der Verwendung von Richtungsinformation gegenüber den
konventionellen Ansätzen mit den Intensitätswerten wird die bessere Dis-
kriminierungsfähigkeit hervorgehoben. Gegenüber den Verfahren, die auf
Kanteninformation arbeiten, sind einerseits wegen der unsicheren Lokalisa-
tion von Kanten und andererseits wegen der relativ problemlosen Rich-
tungsbestimmung ebenfalls bessere Resultate zu erwarten.

Ein weiterer, auch theoretisch abgeklärter Punkt ist die gegenüber der
KKF bessere Korrelationswerte-Verteilung. Unter der Annahme von genü-
gend grossen Mustern kann von einer Gleichverteilung der Richtungen inner-
halb dieses Fensters ausgegangen werden. Daraus folgt, dass die Korre-
lationswerte bei Verwendung der Richtungsinformation Gauss-verteilt sind.
Bei Verwendung der **KKF** (Intensitäten!) resultiert hingegen eine erheblich
schlechtere Korrelationswerteverteilung.

Nishihara berichtet in seinen Artikeln ([124]-[126]) von einem schnellen
und einfachen binären (Stereobild-) Korrelator (**NIS**). Im Gegensatz zu Marr
und Poggio ([109]), die in ihrem klassischen Korrelator die Nulldurchgänge
der Ableitungen nach x und y verwenden, korreliert Nishihara Vorzeichen-
binarisierte Bilder, die er aus einer Filterung mit dem Laplace-Gauss-Opera-
tor gewinnt⁷¹. Dadurch erreicht er neben einem geringeren Rechenaufwand
z.B. auch eine grössere Robustheit gegenüber Helligkeitsrauschen. Durch
die (Vorzeichen-) Binarisierung des Filterresultats wird in der Korrelations-
ebene eine scharfe Spitze an der Objektposition erzeugt (siehe Anhang E.3:
Resultate, Korrelationsfunktionen). Die Breite dieser Erhebung kann durch
den w -Parameter des Laplace-Gauss-Operators gewählt werden⁷². Das Filter
garantiert zudem eine gewisse Unabhängigkeit des Resultats von der lokalen
Helligkeit und dem lokalen Kontrast. Dieser Korrelator wird hauptsächlich
in der Robotik eingesetzt, wo schnelle Taktraten im Bereich von einem
Hertz gefragt sind. In seinen Versuchen benutzt Nishihara Muster der Grös-
se 32×32 Pixel und Filter, die durch $w=4..16$ charakterisiert sind. Für $w=4$
benutzt er dabei einen Operator von 10×10 Pixeln Grösse (in unseren Versu-
chen wurde im gleichen Fall eine Operatorgrösse von 15×15 Pixeln gewählt,
was eine genauere Approximation an das ideale Filter mit sich bringt).

⁷¹ Diese Vorzeicheninformation ist dual zu der von Marr und Poggio vorgeschlagenen
Null-Durchgangs-Kontur.

⁷² Die Breite der Korrelationsspitze in Pixeln entspricht ungefähr dem Wert des w -Para-
meters.

Damit genügend eindeutige Strukturen im Binärbild enthalten sind, sollte der w -Parameter im Normalfall gegenüber der Musterdimension klein gewählt werden. Diese Aussage wird auch durch die Resultate in der vorliegenden Arbeit untermauert.

Das Laplace-Gauss-Filter kann übrigens auch zusammen mit der **KKF** oder jedem anderen (Un-) Ähnlichkeitsmass verwendet werden. Es verstärkt bei günstiger Bildqualität den Extremwert und vermindert den Einfluss der als mehr oder weniger uniform angenommenen Umgebung.

Nack, dessen Algorithmus auch von Svedlov diskutiert wird (siehe unten, [155] und [156]), beschreibt in [122] einige Details seines Algorithmus' sowie die Untersuchungen bezüglich Kantendetektoren. Insgesamt betrachtet er elf verschiedene, einfache Kantendetektoren im Zusammenspiel mit seinem Ähnlichkeitsmass. Das von ihm definierte, binäre Ähnlichkeitsmass (**NA1**) benutzt er zusammen mit grossen Mustern, was zu relativ breiten Korrelationsspitzen führen soll. Den Entscheid, ob eine Positionierung ein zuverlässiges Resultat darstellt, fällt er aufgrund einer simplen, statistischen Analyse der Korrelationswerte (Mittelwert und Standardabweichung) sowie zweier aufeinanderfolgender Clustering-Methoden, die in der Umgebung des gefundenen Extremwerts weitere hohe Korrelationswerte verlangen (breite, unscharfe Korrelationsspitze! Siehe Anhang E.3: *Resultate, Korrelationsfunktionen*). Durch die adaptive Detektion von relativ wenigen Kantenpixeln ($\leq 25\%$) werden Nebenextrema weitgehend unterdrückt und Verdeckungen des Musters gut überbrückt. Die ermutigenden Resultate, die hier dokumentiert werden und teilweise im Gegensatz zu den eigenen Versuchen stehen, werden sicherlich durch die Wahl verhältnismässig grosser Muster und die kurz erläuterten Plausibilitätsbetrachtungen (Korrelationswert, Clustering) verursacht. Dies stimmt in etwa wieder mit den eigenen Erfahrungen überein, die ein sehr starkes Anwachsen der Detektionsrate mit zunehmender Mustergrösse aufzeigen. Allerdings läuft die Hochpassfilterung immer der ursprünglichen Forderung nach breiter Korrelationsspitze zuwider.

Svedlov, McGillem und Anuta ([155] und [156]) untersuchen mit Hilfe von Experimenten verschiedene Ähnlichkeits- und Differenzmasse sowie Vorverarbeitungsmethoden. Getestet werden die **SAD**, die **UKKF** sowie die **KKFMF**. Ein grösserer Teil des Artikels ([156]) beschäftigt sich auch mit der binären Korrelationsmethode von Nack (**NA1**). Das verwendete Bildmaterial stammt von Satelliten mit multispektralen Kameras. Die Bilder sind stark strukturiert, wobei die Kontrastunterschiede in den verschiedenen Spektralbereichen den Matching-Prozess erschweren. Die **KKFMF** schneidet wiederum – konsistent mit den hier vorliegenden Ergebnissen – am

besten ab, dicht gefolgt von der **NA1**-Korrelation und den anderen Verfahren, die auf Hochpass gefilterten Bildern arbeiten. Dies stimmt nicht unbedingt mit den eigenen Versuchen überein. Das überaus gute Abschneiden von Nacks Korrelationsalgorithmus kann aber auf die mit 51x51 Pixeln sehr gross gewählten Muster und die grossen geometrischen Toleranzmargen zurückgeführt werden. Durch die Technik der einfachen Oder-Verknüpfung von verschiedenen (spektralen) Binärbildern wird die Leistung dieses Algorithmus natürlich ebenfalls gesteigert.

Casasent, Schaefer und Sturgill ([36]) führen ein neues Binärbildverfahren ein (**HMT**, hit-miss morphological transform), das auf morphologischen Operationen basiert. Die erforderlichen Erosionen werden optisch realisiert und garantieren damit schnelle Ausführungszeiten. Im wesentlichen kann dieser Algorithmus als Summe der gesetzten Pixel der *UND*-Verbindung zweier Erosions-Ergebnisse verstanden werden. Zur Unterdrückung von Fehldetektionen wird die Erosion, die ja als binarisierte (maximal mögliche Binarisierungsschwelle) unnormierte Kreuzkorrelation mit einer Maske (structuring element) interpretiert werden kann, nicht nur mit der Objektmaske, sondern auch mit einer entsprechend gewählten Gegenmaske auf dem Gegenbild durchgeführt. Die Kombination dieser beiden Resultate ergibt eine deutlich reduzierte Wahrscheinlichkeit für Fehldetektionen.

Die Verallgemeinerung der Erosionen zu Rangordnungsoperationen ergibt bei günstiger Parameterwahl zusätzlich eine sehr gute Robustheit gegenüber Rauschen (bis 50% Rauschpixel) und geometrischen Verzerrungen (bis 25%). Diese Modifikation des Algorithmus⁷³ ist trivial, da einfach die Erosions- (bzw. Korrelations-) Schwelle tiefer angesetzt werden muss und somit auch schlechter definierte Regionen einen Korrelationsbeitrag leisten. Als Implementationsvorschlag wird ein einfacher optischer, auf schnelle Ausführungszeit getrimmter Aufbau vorgestellt.

Der wesentliche Beitrag dieser Arbeit besteht nicht in der Einführung morphologischer Operationen für die binäre Objektdetektion⁷³. Vielmehr zeigt diese Arbeit, dass durch konsequentes Ausnützen der vorhandenen Information, z.B. mit einem Muster und einem Gegenmuster sowie Kombination der jeweiligen Korrelations- (bzw. Erosions-) Resultate, auch mit binären, einfach implementierbaren Verfahren unter Umständen erstaunliche Ergebnisse erreicht werden können.

⁷³ Diese Nomenklatur könnte unschwer durch konventionelle Begriffe wie *unnormierte Kreuzkorrelation* und *Schwellwertbildung* ersetzt werden.

Kumar und Carroll ([93]) analysieren theoretisch die Reduktion des Signal-zu-Rauschleistungsverhältnisses in der Korrelationsebene für die unnormierte Kreuzkorrelationsfunktion, wenn dem Bild nicht weisses Rauschen überlagert wird⁷⁴. Die teilweise beträchtliche Leistungsreduktion wird für verschiedene Fälle theoretisch diskutiert.

Die vergleichende Studie von **Gülch** ([61]) beschreibt die momentan aktuellen Matching-Verfahren im Bereich der *3D-Computer Vision*. Sie beschäftigt sich aber nicht nur mit den notwendigen Ähnlichkeitsmassen sondern z.B. auch mit mathematischen Modellen der Szene oder der Strategie beim Einsatz der verschiedenen Algorithmen. Ziel der Versuche ist es Antworten auf die folgenden Fragen zu erhalten: 'Wie reagieren die verschiedenen Algorithmen auf unterschiedlich komplexe Szenen?', 'Wieviel a priori Wissen wird benötigt?', 'Wie gross ist die Präzision der Algorithmen?' und 'In welcher Art und wie gut überprüfen die Algorithmen ihre eigenen Resultate?'. Die bislang recht knappen Ausführungen über 15 verschiedene Algorithmen aus verschiedenen Anwendungsgebieten fassen die Resultate von flächenbasierten (area based: Ähnlichkeitsmasse, die direkt auf bildhaften Daten arbeiten) und merkmalsbasierten Verfahren zusammen. Den flächenbasierten Verfahren, wie z.B. **KKF** und **SSD**, wird eine eher bessere Genauigkeit als den merkmalsbasierten Verfahren zugestanden.

In [66] beschreibt **Hannah** die Ergebnisse, die beim Test von Gülch ([61], siehe oben) erreicht wurden. Die flächenbasierten Korrelationsverfahren werden als besser geeignet bei texturhaften Szenen taxiert, während die merkmalsbasierten Methoden bei Bildern, die hauptsächlich durch Kanten (relativ geringer Dichte) beschrieben werden können den Vorrang haben. Ein kombiniertes System könnte die Leistung der einzelnen Verfahren erheblich verbessern.

Kuhnert ([92]) befasst sich mit der ternären Korrelation (das Muster enthält nur die Werte 0, +1, -1). Er argumentiert, dass der Mensch die beste, bekannte Mustererkennung besitze und begründet darum die Wahl dieses Verfahrens mit entsprechenden physiologischen Experimenten. Bei Bildern mit vielen Kanten – die konkrete Anwendung von Kuhnert ist die Detektion von Strassenmarkierungen zur autonomen Steuerung eines Fahrzeugs – zeigt sich, dass die Detektionsrate durch die Verwendung von vielen Helligkeitsstufen oder grösseren Mustern nur geringfügig verbessert werden kann.

⁷⁴ Bei weissem Rauschen ergibt die unnormierte Kreuzkorrelationsfunktion das beste Signal-zu-Rauschleistungsverhältnis in der Korrelationsebene.

Schon sehr grob quantisierte, kleine Muster können bei entsprechendem Bildinhalt sehr gute Resultate liefern (siehe dazu auch Helava [68], oben).

Als Ähnlichkeitsmasse werden die **KKF** und die unnormierte Kreuzkorrelationsfunktion (ohne Filterung) untersucht. Die **KKF** wird verworfen, da sie wegen der Normierung zu aufwendig sei. Die ternäre Codierung, zusammen mit der Mittelwertfreiheit der synthetischen Muster (\rightarrow mittelwertfreie, unnormierte Kreuzkorrelationsfunktion) scheint die Notwendigkeit der Normierung sowieso in Frage zu stellen.⁷⁵

Besonders interessant ist die ternäre Korrelation natürlich bezüglich des Rechenaufwandes. Nur triviale Multiplikationen mit ± 1 sind notwendig. Ein Nachteil der unnormierten Kreuzkorrelation besteht aber darin, dass der Wertebereich nicht mehr einfach definiert ist. Die Festlegung von Schwellwerten ist also erheblich schwieriger als z.B. bei der **KKF**, deren Werte im fixen Intervall -1 bis $+1$ liegen.

Die eigentliche Musterdetektion kann durch vier unterschiedliche Maximumdetektoren durchgeführt werden:

- Detektion des absoluten Maximums in der Korrelationsfläche
- Detektion des ersten Maximums bzw. Umkehrpunktes, der eine bestimmte Schwelle überschreitet
- Local hill climber: In einer bestimmten, festlegbaren Umgebung (kleiner als Korrelationsfläche) darf kein höherer Korrelationswert auftreten. Dieses sequentielle Vorgehen zeichnet sich durch einen relativ geringen Rechenaufwand aus.
- Kennzahloperator: Über zwei Parameter wird die erforderliche Breite und Höhe eines Maximums bestimmt.

Fukushima ([56]) berichtet über Versuche mit der **KKFMF** im Bereich der Kartographie (Stereo-Matching). Der Einsatz eines Median-Vorverarbeitungsfilters auf den Bilddaten ergibt 30% bessere Resultate. Bei seinen Tests mit Mustern der Grösse 5×5 bis 33×33 Pixeln erscheinen die kleinen Muster mit 5×5 und 7×7 Pixeln am leistungsfähigsten. Diese Aussagen sind natürlich sehr stark vom Bildmaterial abhängig. Beim Stereo-Matching ist z.B. auch klar, dass kleine Muster bevorzugt werden müssen, da die Tiefendynamik innerhalb des Musters einerseits und die geometrische Verzerrung des Such-

⁷⁵ In Bildern mit allgemeinem, mehrstufigem Inhalt genügt die Mittelwertfreiheit aber nicht! Eine Normierung ist dann unumgänglich, um genügende Zuverlässigkeit zu erreichen.

fensterbildes andererseits das Resultat sonst zu stark beeinflussen würden (siehe z.B. Resultate bei **ZOOM**-Verzerrung).

Mostafavi und Smith berichten in zwei gekoppeltem Artikeln ([120] und [121]) von ihren theoretischen Untersuchungen bezüglich dem Verhalten der unnormierten Kreuzkorrelationsfunktion bei geometrischen Verzerrungen (affine Abbildungen, wie Translation und Rotation). Die Leistungsfähigkeit dieses Ähnlichkeitsmasses hängt vom Mass des Intensitätsrauschen, von der Autokorrelationsfunktion des Bildes und dem Ausmass der geometrischen Verzerrung ab. Die geometrische Verzerrung wirkt sich in der Korrelationsebene als Mittelung über eine kleine Umgebung aus. Somit wird die Korrelationspitze einerseits reduziert und andererseits 'weicher' gemacht.

Für jede Bildverzerrung existiert eine Mustergrösse, die die Detektionsrate oder die Präzision optimiert. Die bezüglich der Detektionsrate optimale Mustergrösse ist bei gegebener Verzerrung grösser als die günstigste Mustergrösse hinsichtlich der mittleren Präzision. Tendenziell sind kleine Muster bei Verzerrungen zu bevorzugen. Die bezüglich der mittleren Genauigkeit optimale Mustergrösse verhält sich proportional zur Breite der Spitze der Autokorrelationsfunktion.

Die Untersuchungen sind bezüglich des Bildmaterials in dem Sinne unrealistisch, als die Bildinformation gleich wie das Intensitätsrauschen modelliert wird (Gauss'sches Rauschen) und kleine Muster die statistischen Annahmen nur schlecht erfüllen.

Die Untersuchung von **Mendelsohn und Englund** ([114]) ist vorab interessant, weil im Gegensatz zu den meisten anderen Arbeiten geometrische Mustertransformationen im Detail betrachtet werden. Das Ziel dieser Arbeit war es in Binärbildern zufällig rotierte Werkstücke aus einem Haufen verschiedener Teile zu detektieren. Verwendet wurde die **UKKF** (allerdings über die Fourier-Transformation gerechnet). Dies hat den Vorteil, dass das Ähnlichkeitsmass relativ einfach auf optischem Wege implementierbar ist und damit auch bei grossen Fenstern äusserst günstige Rechenzeiten garantiert. Das Hauptaugenmerk der Untersuchung lag bei der Bestimmung der optimalen geometrischen Auflösung und der Anzahl (verdrehter) Muster, die für die sichere Detektion benötigt werden. Die besten Resultate wurden bei einer Auflösung von 256x256 Pixeln und 64 verdrehten Mustern zur Detektion eines einzigen Werkstücks erreicht. Diese Resultate sind natürlich stark von den gesuchten Werkstückformen und Ausdehnungen abhängig. Die Robustheit gegenüber rotierten Mustern wird im hier vorliegenden Bericht nicht analysiert. In bestimmten Anwendungen kann aber eine gewisse Robustheit gegenüber solchen Transformationen durchaus notwendig sein.

Andere Lösungen als die von Mendelsohn und Englund vorgestellte, benutzen ein polares Koordinatensystem, um die Anzahl Muster pro Objekt gering zu halten. Allerdings geht dabei die Translationsinvarianz der Korrelationsfunktion verloren (ein kombiniertes, iteratives Verfahren, das sowohl die Muster-Rotation wie auch -Translation verarbeiten kann, wird z.B. in [15] vorgestellt).

Jaffey und Millman ([79]) konzentrieren sich ebenfalls auf das Verhalten von Korrelationskriterien bei verschiedenen Bildtransformationen (Translation, Verkleinerung und Rotation). Neben der **KKFMF** wird auch ein mit der **SAD** verwandtes Differenzmass, die normierte Summe der Differenzenbeträge (**NSAD**, Gl.34), untersucht. Dieses Korrelationskriterium muss trotz der guten Resultate skeptisch beurteilt werden. Der Nenner, bzw. die Summe der Suchfensterintensitäten im Nenner, bewirkt nämlich, dass helle Korrelationsfenster, bei gleicher Diskrepanz zwischen Muster und Korrelationsfenster, einen geringeren Unähnlichkeitswert verursachen als entsprechende dunkle Korrelationsfenster.

$$\text{NSAD}(x, y) = \frac{\sum_{v=0}^{\text{MIPKX}} \sum_{u=0}^{\text{MIPKX}} |r(u, v) - s(x+u, y+v)|}{\sum_{v=0}^{\text{MIPKX}} \sum_{u=0}^{\text{MIPKX}} r(u, v) + \sum_{v=0}^{\text{MIPKX}} \sum_{u=0}^{\text{MIPKX}} s(x+u, y+v)} \quad (34)$$

Mit den eigenen Überlegungen und Resultaten übereinstimmend wird festgestellt, dass bei geometrischen Bildtransformationen kleine Mustergrossen gewählt werden sollten. Um bei Mustern von 8x8 Pixeln Grösse eine Positionierungsgenauigkeit von 0.15 Pixeln zu gewährleisten, dürfen die Suchfenster um maximal 30 Prozent verkleinert werden. Dies stimmt sehr schön mit der in Kapitel 5.2: *Erwartete Detektionsratenverläufe* hergeleiteten Formel 11 auf der Seite 60 für die maximal zugelassene Vergrösserung überein ($\delta=1 \Rightarrow$ 40 Prozent Vergrösserung erlaubt, dies entspricht von den Auswirkungen her in etwa einer Verkleinerung um 30 Prozent).

Marshall ([110]) präsentiert einen Vergleich zwischen einem Grauwert-Centroid-Verfahren, einem merkmalsbasierten Algorithmus und der **SAD**. Bei seinen Versuchen verwendet er synthetische Bilder. Der Hintergrund wird als (bewegtes) korreliertes Rauschen⁷⁶ modelliert. Zusätzlich wird auch Sensorrauschen im Szenenmodell berücksichtigt. Es zeigt sich, dass die **SAD** im Gegensatz zu den beiden nicht-korrelativen Vergleichsverfahren

⁷⁶ Korrelationsweite: 10..20 Pixel

praktisch unabhängig vom Signal-zu-Hintergrund-Leistungsverhältnis ist. Dies ist einleuchtend, da bei den Tests das Muster keine Hintergrundpixel enthielt! Das Sensorrauschen bewirkt, dass 'erst' ab einem Signal-zu-Rauschleistungsverhältnis von ca. Eins gute Resultate erzielt werden.

Das nicht näher beschriebene, mekmalsbasierte Verfahren schneidet aufgrund dieser Ergebnisse, besonders bei nicht-uniformem Rauschen sowie Auflösungsänderungen, Verdrehungen und Verdeckungen, besser als die SAD ab.

Wong ([171] berichtet unter anderem von einem weiteren interessanten, hier nicht untersuchten Korrelationsverfahren (**KOM**). Für jedes Muster-Korrelationsfenster-Paar wird die Häufigkeit von Helligkeitskombinationen entsprechender Pixel bestimmt (*Co-occurrence-Matrix*). $h_{ij}(x,y)$ bezeichnet für das Korrelationsfenster an der Stelle (x,y) die Häufigkeit mit der im Muster eine Helligkeit i und im entsprechenden Pixel des Korrelationsfensters die Intensität j zu beobachten ist. Die eigentliche Korrelationsfunktion wird dann wie folgt definiert:

$$\text{KOM}(x,y) = \prod_{i=0}^{\text{MaxIntens}} \frac{h_i(x,y)}{\sum_{j=0}^{\text{MaxIntens}} h_j(x,y)} \quad (35)$$

Im Fall von binären Daten (über nicht näher spezifizierte Hochpassfilterung und anschließende Binarisierung) vereinfacht sich diese Gleichung noch weiter:

$$\text{KOM}_{\text{binär}}(x,y) = \left[\frac{h_{00}(x,y)}{h_{00}(x,y) + h_{01}(x,y)} \right] \left[\frac{h_{11}(x,y)}{h_{11}(x,y) + h_{10}(x,y)} \right] \quad (36)$$

Schon bei kleineren Mustern (8x8 Pixel, Suchfenster 32x32 Pixel) soll dieses Verfahren sehr gute Resultate liefern. Wie beim binären Algorithmus von Nack ist die Vorverarbeitung und Binarisierung von entscheidender Bedeutung. Da aber exakte Intensitätsübereinstimmung gefordert wird damit der Korrelationswert zunimmt, ist in der Regel nur bei idealen Bild-Aufnahmebedingungen und grösseren Mustern eine gute Detektionsrate zu erwarten.

Ein weiterer, alternativer Korrelationsalgorithmus (im weiteren Sinne), der *cluster reward Algorithmus* (**CRA**), wird von Kuglin und Eppler in [91] vorgestellt. Er basiert, wie auch das gerade eben beschriebene Verfahren, auf der Co-occurrence-Matrix $h_{ij}(x,y)$. Dieses zweidimensionale Histo-

gramm enthält wiederum die Auftretenswahrscheinlichkeit für ein Helligkeitswertepaar (i,j) entsprechender Muster- und Korrelationsfensterpixel (Korrelationsfenster an der Position (x,y)). Zudem werden auch die (eindimensionalen, im Prinzip aus der Co-occurrence-Matrix ableitbaren) Histogramme hr_i und $hs_j(x,y)$ des Musters und des Korrelationsfensters benötigt. Das Korrelationskriterium ergibt sich dann wie folgt:

$$\text{CAR}(x,y) = \frac{\left[\sum_{j=0}^{jmax} \sum_{i=0}^{imax} h_y^2(x,y) \right] / T(x,y) - T(x,y)}{1 - T(x,y)} \quad (37)$$

$$\text{mit } T(x,y) = \sqrt{\sum_{j=0}^{jmax} hs_j^2(x,y) \cdot \sum_{i=0}^{imax} hr_i^2}, \quad imax, jmax: \text{ Maximale Intensitäten}$$

Der Wertebereich dieses Ähnlichkeitsmasses liegt zwischen 0 und 1. Für grosse Muster (z.B. 64x64 Pixel) liefert es selbst bei Kontrastumkehr, die von Korrelationskriterien in der Regel nicht mehr erfolgreich bewältigt werden kann, oder anderen, beliebigen Helligkeitspermutationen, offenbar ausgezeichnete Resultate. Deshalb ist dieses Verfahren besonders für den Einsatz mit Bildern von Sensoren unterschiedlicher spektraler Empfindlichkeit, wie sie z.B. bei Satellitenbildern vorkommen, geeignet.

Um zuverlässige Positionierungen zu erhalten, ist es nötig, die Anzahl der berücksichtigten Helligkeitsstufen klein zu halten (damit die Co-occurrence-Matrix mit genügend grossen Werten und relativ dicht besetzt ist). Kontrast-schwache Bilder, die wegen der geringen Zahl unterscheidbarer Helligkeiten keine Struktur mehr aufweisen, setzen also auch diesem interessanten Verfahren enge Grenzen.

Zum Lokalisieren eines Objekts kann anstelle eines klassischen Korrelationsverfahrens auch der Algorithmus von **Yeshurun und Schwartz** verwendet werden ([175]). Er zeichnet sich durch extreme Rauschfestigkeit aus (siehe auch [37], [97], [103] und [129]) und basiert auf dem aus der Unterwasser-Akkustik und der Sprachverarbeitung bekannten *Cepstrum* (Gl.38).

$$\text{Cepstrum}\{a(x,y)\} = \left\| F \left\{ \log \left(\| F \{ a(x,y) \} \|^2 \right) \right\} \right\|^2 \quad (38)$$

$F\{\cdot\}$ bezeichnet die Fourier-Transformation

Unter der Annahme, dass das Suchfenster eine genaue, um Δx und Δy verschobene Kopie des Musters enthält, und durch Annäherung des Log-

arithmus' ($\log(1+x)$) mittels einer Reihenentwicklung, lässt sich die Translation demnach wie folgt bestimmen:

$$a(x, y) = r(x, y) + s(x, y) = r(x, y) + r(x + \Delta x, y + \Delta y) \quad (39)$$

$$\begin{aligned} \|A(\omega_x, \omega_y)\|^2 &= \|R(\omega_x, \omega_y) \cdot (1 + e^{-j(\Delta x \omega_x + \Delta y \omega_y)})\|^2 \\ &= \|R(\omega_x, \omega_y)\|^2 \cdot (1 + \cos(\Delta x \omega_x + \Delta y \omega_y)) \end{aligned}$$

$$\text{Cepstrum}\{a(x, y)\} = \text{Cepstrum}\{r(x, y)\} + \sum_{n=-\infty}^{\infty} \frac{\delta(x - n \cdot \Delta x, y - n \cdot \Delta y)}{n} \quad (40)$$

Durch Einführung einer genügend grossen, vorgewählten Disparität beim Zusammensetzen von Muster und Suchfenster (Gl.39) wird garantiert, dass diese Dirac-Pulse vom restlichen Cepstrum getrennt werden können. Die gesuchte translatorische Verschiebung kann anschliessend durch Detektion der Diracstösse ermittelt werden.

Wie bei den (binary) phase only matched Filter (siehe weiter unten) besteht auch hier das Problem des 'Windowing', das eine wichtige Voraussetzung für die Funktion des Algorithmus' ist. Um Randeffekte bei der Fourier-Transformation zu vermindern, muss eine Windowing-Funktion gewählt werden, die die Randbereiche des Fensters genügend dämpft. Andererseits sollte die Information auch nicht zu stark beeinträchtigt werden. Dieser Kompromiss kann nur bei vergleichsweise grossen Fenstern befriedigend getroffen werden.

Von der Idee her sehr ähnlich zum Cepstrum-Algorithmus arbeitet der *joint-transform correlator*, wie er z.B. von Kuo in [94] beschrieben wird. Analog dem Cepstrum-Verfahren das Spektrum einer Kombination von verschobenem Referenz- und Suchfenster bestimmt (Gl.41).

$$I(\omega_x, \omega_y) = R(\omega_x, \omega_y) e^{j\omega_x x_0} + S(\omega_x, \omega_y) e^{-j\omega_x x_0} e^{-j\omega_y y_0} \quad (41)$$

mit x_0, y_0 als charakterisierende Grössen der Verschiebung der Einzel-fenster im Kombinationsbild

Die Rücktransformation von $|I(\omega_x, \omega_y)|^2$ ergibt ein Signal, das einerseits die entsprechenden Autokorrelationsfunktionen und andererseits die Kreuzkorrelationsfunktionen des Musters mit dem Suchfenster und des Suchfensters mit dem Muster enthält (beide Korrelationsrichtungen; jeweils unnor-

mierte Kreuzkorrelationsfunktion). Dieses Verfahren lässt sich sehr gut mit optischen Komponenten (inkl. Lichtmodulator) in einem kohärenten System realisieren. Dementsprechend können sehr schnelle Ausführungszeiten erreicht werden. Eine Schwierigkeit stellt jedoch die z.T. ungenügende Lichteffizienz dar, die die Detektierbarkeit des Extremwertes beeinträchtigt. Durch Reduktion des Kombinations-Spektrums auf die reine Phaseninformation lässt sich dieses Problem aber elegant lösen (siehe auch *binary phase only matched filter*, weiter unten). Damit wird ein nichtlinearer oder genauer gesagt ein *bipolarer joint-transform correlator* realisiert. Neben der Steigerung der Lichteffizienz scheint die Detektierbarkeit des Signals damit in gleicher Weise verbessert zu werden.

Auch die (*binary*) *phase only matched filter* (**POF** und **BPOF**) können im weitesten Sinne zu den flächenbasierten Ähnlichkeitsmassen gezählt werden. Sie realisieren im Prinzip eine unnormierte Kreuzkorrelationsfunktion mit einem modifizierten Muster. Das gesamte Amplitudenspektrum des Referenzfensters wird ungeachtet der Werte auf eins gesetzt. Der Phasenverlauf wird hingegen unverändert belassen⁷⁷. Bei binary *phase only matched* Filtern wird zusätzlich die Phaseninformation binarisiert. Zu diesen Verfahren sind eine ganze Reihe von Artikeln erschienen, die aber hier nicht alle besprochen werden können ([15], [24], [42], [43], [51], [53], [71], [73], [74], [81], [90], [98], [99] und [149]). Da diese Filter relativ einfach auf optischem Wege realisierbar sind, ist das Interesse speziell für Anwendungen mit grossen Suchfenstern und Mustern gross. Eine vergleichbare Geschwindigkeit ist auf elektronischem Weg nur mit sehr grossem Hardwareaufwand erreichbar. Der Vorteil der *phase only matched Filter* gegenüber den klassischen *matched Filtern* ist insbesondere ihre bessere Filter-Effizienz. Dies ist bei optischen Realisationen wichtig, um das Intensitätsmaximum und damit die Objektposition ohne Problem detektieren zu können (siehe Horner und Gianino [73]). Eigene Versuche wie auch andere Untersuchungen ([173]) deuten aber auf gewisse Schwierigkeiten dieses Verfahrens selbst bei optimalen Bildverhältnissen hin.

⁷⁷ Bekannterweise enthält das Phasenspektrum verglichen mit dem Amplitudenspektrum die wichtigere Information für die Erkennbarkeit eines Bildes ([77] und [167]).

8. Schlussbemerkungen

Die hier vorgestellten Ähnlichkeitsmasse wurden bezüglich ihres Rechenaufwandes und ihrer Robustheit gegenüber verschiedenartigen Bildstörungen untersucht. Für alle Korrelationsalgorithmen zeigten sich Vor- wie auch Nachteile. Der eine Algorithmus arbeitet selbst bei stärksten Bildstörungen noch zuverlässig, während ein anderer Algorithmus unter gewissen Umständen gewisse Mängel in der Robustheit aufweist, dafür aber vielleicht etwas weniger Rechenleistung beansprucht oder bei einer anderen Störungsart bessere Resultate ergibt. Diese Aussage illustriert, dass die Wahl eines Korrelationsalgorithmus' eine Optimierungsaufgabe mit verschiedenen, beeinflussenden Parametern ist. Eine einzig gültige Lösung gibt es im allgemeinen nicht! Die robusteste Lösung bestünde in einem 'hybriden' Korrelatorsystem ([40] und [167]), das je nach Situation die Resultate der geeignetsten Verfahrens stärker gewichtet. Es könnten z.B. auch komplett unterschiedliche Verfahren (z.B. merkmalsbasierte Verfahren) gleichzeitig benutzt werden, um die teilweise komplementären Eigenschaften auszunutzen. Dies ist aber wegen der hohen notwendigen Rechenleistung wie auch wegen der Behandlung der vielen evtl. widersprüchlichen Resultate sehr aufwendig oder zumindest schwierig. Darum muss der Ingenieur in der Regel beim Aufbau eines Korrelators eine einzige Wahl treffen. Hierzu sollte er sich die folgenden Fragen stellen:

- Welche Störungen können auftreten?
- Wie stark sind diese Störungen?
- Wie häufig treten die Störungen auf?
- Unter welchen Bedingungen muss der Korrelator noch zuverlässig arbeiten?
- Muss ein Fehlverhalten vom Korrelator selber detektiert werden?
- Welche Rechenzeit muss eingehalten werden?
- Welche Rechenleistung steht zur Verfügung?

Ein weiterer wichtiger Aspekt, um mit einem Korrelatorsystem optimale Leistung (Detektionsraten) zu erreichen, ist auch die Einbettung des Korrelationsalgorithmus' in das gesamte System. Ein, der jeweiligen Aufgabenstellung angepasster, übergeordneter, die Korrelation steuernder Algorithmus

kann die Leistung erheblich beeinflussen. Z.B. können durch Plausibilitätsbetrachtungen die Korrelatorresultate verifiziert werden. Nack ([122]) wendet unter anderem beispielsweise ein Korrelationswert-Clustering-Verfahren an, das Fehlpositionierungen detektiert. Als weitere Plausibilitätsbetrachtungen kommen folgende Möglichkeiten in Frage:

- Maximum/Minimum-Korrelationswert muss innerhalb eines bestimmten, durch Vorwissen definierten Bereichs bleiben.
- Statistik der Korrelationswerte in der Umgebung des Extremwerts erfüllen gewisse Bedingungen (Stetigkeit der Korrelationsfunktion! Clustering).
- mittlere Helligkeit/Varianz/Textur etc. im Objektfenster ändert von Bild zu Bild nur wenig.
- Kontinuitätsbedingungen (z.B. beim Tracking: Übereinstimmen des Korrelationsresultats mit der bereits bekannten Bewegung aus der Vergangenheit (Kalmanfilter, Prädiktor) oder sonstige Modellvorstellungen).
- Signale von anderen Sensoren und Ereignissen auswerten (z.B. Kamerabewegung zwischen zwei Bildaufnahmen).
- weiteres Vorwissen verwenden.

Einzelne, oder eine Kombination dieser und ähnlicher Betrachtungen sollten als Überwachung eines jeden Korrelators eingesetzt werden!

Am Institut für Elektronik wurden im Verlaufe der Zeit verschiedene, auf Signalprozessoren basierende Echtzeit-Trackingsysteme realisiert. Die Hardware besteht jeweils aus einer PC-basierten Signalprozessorkarte und einer frame-grabber-Karte zur Digitalisierung der Kamerainformation. Im wesentlichen kann ein solches System in die folgenden logischen Blöcke unterteilt werden: Musterdetektion (Korrelator), Prädiktor zur Schätzung der Musterbewegung und zur Steuerung des Suchfensters, Musteradaptation zur Nachführung der Information im Referenzfenster. Durch die Verwendung eines Bewegungsprädiktors (mit vergleichsweise geringem Rechenaufwand) lässt sich die Suchfenstergröße reduzieren, die Anforderungen an die erforderliche Rechenleistung also vermindern. Kleine Suchfenster haben zudem den Vorteil, dass die Wahrscheinlichkeit für Fehldetektionen geringer ist. Eine Korrelatorleistung von 100 bis 1000 Korrelationswerten pro 40msec reicht somit meist selbst zur Verfolgung von schnell beschleunigten

Mustern⁷⁸ aus. Die entsprechende Rechenleistung, z.B. für ein Muster von 16x16 Pixeln Grösse, ist heute in einem einzigen modernen Signalprozessor erhältlich. Allerdings muss der rechenaufwendigste Teil erfahrungsgemäss in Assembler geschrieben werden.

Bei der Realisierung des Korrelatorteils auf einem 16 Bit-Fixkomma-Rechenwerk wurde, trotz der üblicherweise geringeren Robustheit, die **KKF** als Detektionsalgorithmus verwendet. Der Preis (Ausführungs- und Implementationszeit) für eine ausgedehnte Mehrfachpräzisionsrechnung erschien zu hoch, um die favorisierte **KKFMF** zu implementieren. Auf einem System mit Fließkomma-Signalprozessor wurde aber später mit relativ geringem Aufwand die **KKFMF** realisiert, die auch im praktischen Betrieb einen sehr robusten Eindruck hinterlässt.

Sehr wichtig ist bei einer Trackinganwendung natürlich, dass ein Muster über lange Zeit, d.h. hunderte oder tausende von Bildern fehlerlos detektiert werden kann. Unter schwierigen Verhältnissen mit niedrigem Kontrast zwischen verfolgtem Muster und Hinter-/Vordergrund ist es darum unumgänglich, konstante Aufnahmebedingungen zu gewährleisten (AGC und Blendeneinstellung optimieren, evtl. auch abschalten) und aussagekräftige Plausibilitätstests durchzuführen. Damit werden einerseits Fehldetektionen vermieden und andererseits unabwendbare Fehler, z.B. verursacht durch die Verdeckung des verfolgten Helligkeitsmusters aufgrund von Vordergrundinformation, detektiert. Eine kurzzeitige Überbrückung derartiger Situationen ohne Positionsinformation ist dann vielfach ohne grosse Probleme, z.B. mit Hilfe eines Bewegungsmodells möglich.

Die vorliegende Arbeit gibt Anhaltspunkte unter welchen Umständen ein bestimmter Korrelationsalgorithmus erfolgreich eingesetzt werden kann. Die Betrachtung des Rechenaufwandes zeigt ausserdem wie und evtl. auch auf welchen Systemen er am günstigsten implementiert werden kann. Die letztendliche Wahl des Algorithmus' kann aber niemandem abgenommen werden. Das Ziel dieser und weiterer Arbeiten muss es sein, dem Ingenieur eine Grundlage zu geben, die Wahl aufgrund der spezifischen Umgebung, des erforderlichen Ein-/Ausgangsverhaltens ([84]) sowie der hier vorliegenden und anderweitig erhältlicher Daten selbständig auszuarbeiten. Erstrebenswert wäre eine umfassende tabellarische Zusammenstellung der Eigenschaften der Korrelationskriterien – wie auch aller anderen Rechenvorschriften. Mit Hilfe dieser Tabellen könnten Algorithmen blockschaltbildartig, wie dies teilweise beim Hardwareentwurf schon einige Zeit möglich ist, aufge-

⁷⁸ Maximal erlaubte Beschleunigung: Ca. 5..50+ Pixel pro Bild, je nach Vorkenntnissen über die Bewegung.

baut und ihr Verhalten im voraus zuverlässig abgeschätzt werden. Selbstverständlich ist die vorliegende Arbeit nur ein kleiner Schritt in diese Richtung. Neben den verwertbaren Resultaten und Erfahrungen soll sie aber auch andere Ingenieure dazu anregen, durch entsprechende Untersuchungen zu diesem Ziel beizutragen. Nicht nur die Entwicklung und die Implementation neuer Algorithmen ist wichtig, auch der Vergleich mit anderen Verfahren ist von grosser Bedeutung! Dies erlaubt ein kontrolliertes und modulares Vorgehen beim Design von grösseren Systemen.

Natürlich kann bei einer wissenschaftlichen Arbeit das hier angewendete, beinahe ausschliesslich experimentelle Vorgehen kritisiert werden. Theoretische Betrachtungen wurden nur am Rande gemacht. Der Grund hierfür liegt hauptsächlich in der Schwierigkeit eine einheitliche theoretische Betrachtungsweise, die für alle betrachteten Algorithmen Gültigkeit hat, zu entwickeln. Eine gesonderte Betrachtung der einzelnen Algorithmen erbringt in der Regel nur sehr schlecht vergleichbare Resultate. Zudem sind die üblichen statistischen Annahmen über die Bildinformation und die Bildstörung, wie sie z.B. von Kumar ([93]) und Ryan ([138]) gemacht werden, nur für grössere Bildausschnitte mehr oder weniger genau erfüllt. Darum müssen theoretische Überlegungen dieser Art – wie auch experimentelle Resultate – immer, entsprechend den Abweichungen der anwendungsspezifischen Randbedingungen von den Voraussetzungen der zu Rate gezogenen Untersuchungen, relativiert werden.

9. Literaturverzeichnis

- [1]: E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, J. M. Ogden; *pyramid methods in image processing*; RCA Engineer, 29-6, Nov./Dec. 1984; 33-41.
- [2]: J. K. Aggarwal, L. S. Davis, W. N. Martin; *correspondence processes in dynamic scene analysis*; Proc. of the IEEE, Vol. 69, No. 5, May 1981.
- [3]: Ali N. Akansu, Jung Hui Chien, M. S. Kadur; *lossless compression of block motion information in motion compensated video coding*; SPIE Proc., Vol. 1199, Visual Communication and Image Processing IV, 1989; 30-38.
- [4]: V. V. Alexandrov, N. D. Gorsky, S. N. Mysko; *a fast technique for recursive scene matching using pyramids*; Pattern Recognition Letters 3, Dec. 1985; 413-419.
- [5]: Stefano Alliney, Carlo Morandi; *digital image registration using projections*; IEEE Trans. on PAMI, Vol. 8, No. 2, March 1986; 222-233.
- [6]: P. Anandan; *computing dense displacement fields with confidence measures in scenes containing occlusion*; Proc. of the Image Understanding Workshop, Oct. 1984, New Orleans, Louisiana; 236-246.
- [7]: P. Anandan; *a computational framework and an algorithm for the measurement of motion*; Int. Jou. of Computer Vision; 2(3); 1989; 283-310.
- [8]: P. Anandan; *a unified perspective on computational techniques for the measurement of visual motion*; Proc. of the Image Understanding Workshop, Feb. 23-25 1987, Los Angeles, California, Vol. I; 719-732.
- [9]: Paul E. Anuta, Farzin Davalou; *resolution matching for registration of dissimilar images*; Proc. IEEE 1982, Pattern Recognition and Image Processing, June 14-17 1982, Las Vegas, Nevada; 327-332.
- [10]: P. Aschwanden; *Tracking-Algorithmen*, Teil A, Bericht 2; Jahresbericht 1987 zum Projekt: Hardwarenahe Verarbeitung von Video-Bildsequenzen.
- [11]: P. Aschwanden; *real-time tracker with signal processor*; Proc. of the European Signal Processing Conf. (EUSIPCO), 1988, Grenoble, France, Vol. 1; 43-46.
- [12]: P. Aschwanden; *methods for real-time tracking of uncooperative targets*; Int. Conf. on Acoustics, Speech and Signal Processing, IEEE, May 1989, Vol. 3; 1560-1563.

- [13]: P. Aschwanden; *Bericht 6*; Jahresbericht 1991 zum Projekt: Hardwarenahe Verarbeitung von Video-Bildsequenzen.
- [14]: P. Aschwanden, W. Guggenbühl; *experimental results from a comparative study on correlation-type registration algorithms*; Proc. of the 2nd Int. Workshop on Robust Computer Vision, 1992, ISPRS, Bonn, Quality of Vision Algorithms, Förstner/Ruwiedel (eds.); 268-289.
- [15]: Joseph G. Bailey, Richard A. Messner; *log-polar mapping as a preprocessing stage for an image tracking system*; SPIE Vol. 1002, Intelligent Robots and Computer Vision, Nov. 1988; 15-22.
- [16]: R. Ballmer, M. Nägelin; *Tracking mit Signalprozessor*; Studienarbeit am Inst. für Elektronik der ETH Zürich, Sommersemester 1987.
- [17]: R. Ballmer, M. Nägelin; *Steuerung/Regelung einer Videokamera*; Diplomarbeit am Inst. für Elektronik der ETH Zürich, Wintersemester 1987/88.
- [18]: Emmanuel P. Baltasvias; *multiphoto geometrically constrained matching*; Dissertation, Inst. für Geodäsie und Photogrammetrie, ETH Zürich, Dez. 1991.
- [19]: Stephen T. Barnard; *stochastic stereo matching over scale*; Proc. of the Image Understanding Workshop, April 6-8 1988, Cambridge, Massachusetts, Vol. II; 769-778.
- [20]: Daniel I. Barnea, Harvey F. Silverman; *a class of algorithms for fast digital image registration*; IEEE Trans. on Comp., Vol. 21, No. 2, Feb. 1972; 179-186.
- [21]: Yair Barniv, Hassan Mostafavi; *correlation of images with random contrast reversals*; SPIE Vol. 238, Image Processing for Missile Guidance, 1980; 156-165.
- [22]: K.-H. Bers, M. Bohner, P. Fritsche; *image sequence analysis for target tracking*; NATO ASI Series, Vol. F2, Image Sequence Processing and Dynamic Scene Analysis, 1983, edited by T. S. Huang; 493-501.
- [23]: Martin Bichsel, Peter Seitz; *Der elektronische Pförtner: Automatisches Erkennen und Identifizieren von menschlichen Gesichtern*; Proc. zum 12. DAGM-Symp., Oberkochen-Aalen, 24.-26. Sep. 1990; 106-113.
- [24]: Leo Bivetti, Andreas Wyss; *Schnelle Objektlokalisierung mit Signalprozessor*; Studienarbeit am Inst. für Elektronik der ETH Zürich, Sommersemester 1991.
- [25]: Thomas Blum, Dominik Amman; *Kompass-Filter zu SYDAMA-2*; Studienarbeit am Inst. für Elektronik der ETH Zürich, Wintersemester 1990/91.
- [26]: Robert A. Boie, Ingemar J. Cox; *an analysis of camera noise*; IEEE Trans. PAMI, Vol. 14, No. 6, June 1992; 671-674.

- [27]: J. S. Boland, III, H. S. Ranganath, W. W. Malcolm; *computational efficiency of multiple image registration algorithms*; SPIE Vol. 252, Smart Sensors II, July/Aug. 1980; 105-112.
- [28]: Abdesselam Bouzerdoum, Robert B. Pinter; *image motion processing in biological and computer vision systems*; SPIE Proc., Vol. 1199, Visual Communication and Image Processing IV, 1989; 1229-1240.
- [29]: Achim v. Brandt, Stefan Lanser; *hierarchical block matching using edge preserving smoothness for optical flow field estimation*; Proc. zum 12. DAGM-Symp., Oberkochen-Aalen, 24.-26. Sep. 1990; 634-641.
- [30]: E. Oran Brigham; *FFT, Schnelle Fourier-Transformation* (4. Auflage).
- [31]: R. F. Browne, R. M. Hodgson; *pyramidal vision applied to automatic inspection*; Third Int. Conf. on Image Processing and its Applications, Conf. Publication 307, IEE, 18-29 July 1989, Warwick, UK; 59-62.
- [32]: Christof Bühler; *Schneller Korrelator*; Diplomarbeit am Inst. für Elektronik der ETH Zürich, Wintersemester 1990/91.
- [33]: Peter J. Burt, Chihsung Yen, Xinping Xu; *local correlation measures for motion analysis, a comparative study*; Conf. Proc. IEEE 1982, Pattern Recognition and Image Processing, June 14-17 1982, Las Vegas, Nevada; 269-274.
- [34]: Terry Caelli, Shyamala Nagendran; *fast edge-only matching techniques for robot pattern recognition*; Int. Jou. on Computer Vision, Graphics, and Image Processing, Vol. 39, No. 2, Aug., 1987; 131-143.
- [35]: David Casasent; *is what you see what you get?*; BYTE, Sep. 1992; McGraw-Hill publication; 189-194.
- [36]: David Casasent, Roland Schaefer, Robert Sturgill; *optical hit-miss morphological transform*; Applied Optics, 10 Oct. 1992, Vol. 31, No. 29; 6255-6263.
- [37]: Donald G. Childers, David P. Skinner, Robert C. Kemerait; *the cepstrum: a guide to processing*; Proc. of the IEEE, Vol. 65, No. 10, Oct. 1977; 1428-1443.
- [38]: W. C. Choate, W. W. Boyd; *correlation tracking concepts for THASSID*; NTC, (record), also: National Telemetry Conf. Proc., spon. by IEEE; 1979; 79ff.
- [39]: Chun-Hsien Chou, Yung-Chang Chen; *moment-preserving pattern matching*; Pattern Recognition, 1990, Vol. 23, No. 5; 461-474.
- [40]: E. H. Conrow, J. A. Ratkovic; *almost everything one needs to know about image matching systems*; SPIE Vol. 238, Image Processing for Missile Guidance (1980); 426-453.

- [41]: V. Considine, A. S. Bhandal, J. N. Gooding; *single chip motion estimator for video CODEC applications*; Third Int. Conf. on Image Processing and its Applications, Conf. Publication 307, IEE, 18-29 July 1989, Warwick, UK; 285-289.
- [42]: Don M. Cottrell, Roger A. Lilly, Jeffrey A. Davis, Timothy Day; *optical correlator performance of binary phase-only filters using Fourier and Hartley transforms*; Applied Optics, Vol. 26 No. 18, 15 Sep. 1987; 3755-3761.
- [43]: Jeffrey A. Davis, Don M. Cottrell, Glenn W. Bach, Roger A. Lilly; *phase-encoded binary filters for optical pattern recognition*; Applied Optics, Vol. 28, No. 2, 15. Jan. 1989; 258-261.
- [44]: T. Day, J.-P. Muller; *quality assessment of digital elevation models produced by automatic stereo matchers from SPOT image pairs*; 16th Congr. of the Int. Soc. for Photogrammetry and Remote Sensing; 148-159.
- [45]: Joachim Dengler, Markus Schmidt; *the pseudo-logarithmic transformation for robust displacement estimation*; Proc. zum 12. DAGM-Symp., Oberkochen-Aalen, 24.-26. Sep. 1990; 275-281.
- [46]: Luca Denti, Marco Derighetti; *Tracking mit modernem Signalprozessor*; Studienarbeit am Inst. für Elektronik der ETH Zürich, Wintersemester 1989/90.
- [47]: Edward R. Dougherty, Charles R. Giardana; *matrix structured image processing*; Prentice Hall, 1987.
- [48]: V. N. Dvornychenko; *generalized correlation measures for use in signal and image processing*; Comp. Soc. on Pattern Recognition and Image Processing, 3-5 Aug. 1981.
- [49]: Sue Ellen Englert, Zhaoxin Sheng, R. Lynn Kirlin; *the evaluation of a cross-correlation pattern recognition technique for flow visualization*; Pattern Recognition, 1990, Vol. 23, No. 3/4; 237-243.
- [50]: Zhixi Fang, Xiaobo Li, Lionel M. Ni; *parallel algorithms for image template matching on hypercube SIMD computers*; IEEE Trans. PAMI, Vol. 9, No. 6, Nov. 1987; 835-841.
- [51]: Michael W. Farn, Joseph W. Goodman; *optimal binary phase-only matched filters*; Applied Optics, Vol. 27 No. 21, 15 Nov. 1988; 4431-4437.
- [52]: O. Faugeras, P. Fua, B. Hotz, R. Ma, L. Robert, M. Thonnat, Z. Zhang; *quantitative and qualitative comparison of some area and feature-based stereo algorithms*; Proc. of the 2nd Int. Workshop on Robust Computer Vision, 1992, ISPRS, Bonn, Quality of Vision Algorithms, Förstner/Ruwiedel (eds.); 1-26.

- [53]: Kenneth H. Fielding, Joseph L. Horner; *clutter effects on optical correlators*; SPIE Vol. 1151, Optical Information Processing Systems and Architectures, 1989; 130-137.
- [54]: James M. Florence; *design considerations for phase-only correlation filters*; SPIE Vol. 1151, Optical Information Processing Systems and Architectures, 1989; 195-202.
- [55]: M. Fondelli; *a method for three-dimensional models generation from digital image correlation*; 16th Congr. of the Int. Soc. for Photogrammetry and Remote Sensing; 189-195.
- [56]: Yoshikazu Fukushima; *generation of DTM using SPOT image near Mt. Fuji by digital image correlation*; 16th Congr. of the Int. Soc. for Photogrammetry and Remote Sensing; 225-234.
- [57]: P. Gerken, D. Adolph; *improved motion-compensating prediction with rotational block matching*; Third Int. Conf. on Image Processing and its Applications, Conf. Publication 307, IEE, 18-29 July 1989, Warwick, UK; 290-294.
- [58]: James M. Geros, A. J. Witsmeer; *performance evaluation of image correlation techniques*; SPIE Vol. 186, Digital Processing of Aerial Images, 1979; 197-205.
- [59]: Bayesteh K. Ghaffary; *a review of image matching techniques*; SPIE Vol. 596, Architectures and Algorithms for Digital Image Processing, Dec. 1984.
- [60]: Frank Glazer, George Reynolds, P. Anandan; *scene matching by hierarchical correlation*; IEEE Conf. Proc., 1983, Computer Vision and Pattern Recognition; 432-441.
- [61]: Eberhard Gülch; *results of test on image matching of ISPRS WG III/4*; ISPRS Jour. of Photogrammetry and Remote Sensing, 46, 1991; 1-18.
- [62]: Rafael C. Gonzalez, Paul Wintz; *digital image processing*; Addison-Wesley Publishing Company, 1977.
- [63]: Frank Goodenough; *IC signal processor runs 45 billion MACs/s*; Electronic Design, July 23, 1992; 51-56.
- [64]: A. Goshtasby, S. H. Gage, J. F. Bartholic; *a two-stage cross correlation approach to template matching*; IEEE Trans. PAMI, Vol. 6, No. 3, May 1984; 374-378.
- [65]: Ernest L. Hall, David L. Davies, Michael E. Casey; *the selection of critical subsets for signal, image and scene matching*; IEEE Trans. PAMI, Vol. 2, No. 4, July 1980; 313ff.

- [66]: Marsha Jo Hannah; *digital stereo image matching techniques*; 16th Congr. of the Int. Soc. for Photogrammetry and Remote Sensing, 1988; 280-293.
- [67]: Marsha Jo Hannah; *computer matching of areas in stereo images*; Dissertation, Stanford University, 1974.
- [68]: U. V. Helava; *digital correlation in photogrammetric instruments*; *Photogrammetria*, 34 (1978); 19-41.
- [69]: Edwin R. Hiller; *pulse doppler map matching*; SPIE Vol. 238, *Image Processing for Missile Guidance*, 1980; 50-58.
- [70]: Paul R. Hinrichs; *advanced terrain correlation techniques*; Proc. of the IEEE 1976 Position Location and Navigation Symp. (PLANS); 88-96.
- [71]: Claude Hoffmann, Markus Kohler; *Bewegungsdetektion mit Signalprozessor*; Diplomarbeit am Inst. für Elektronik der ETH Zürich, Wintersemester 1990/91.
- [72]: Joseph L. Horner; *metrics for assessing pattern-recognition performance*; *Applied Optics*, 10 Jan. 1992, Vol. 31, No. 2; 165-166.
- [73]: Joseph L. Horner, Peter D. Gianino; *phase-only matched filtering*; *Applied Optics*, Vol. 23, No. 6, 15. March 1984; 812-816.
- [74]: Joseph L. Horner, James R. Leger; *pattern recognition with binary phase-only filters*; *Applied Optics*, Vol. 24 No. 5, 1 March 1985; 609-611.
- [75]: Bernard Hotz; *étude de techniques de stéréovision par corrélation, application au programme véhicule autonome planétaire (V. A. P.)*; Rapport de stage de 3ème année de l'ESSI et du DEA de Robotique et Vision Artificielle, Ecole Supérieure en Science Informatique INRIA de Sophia Antipolis, Sep. 1991.
- [76]: Jun S. Huang; *criteria of template matching and image sensor noise analysis*; Proc. IEEE Comp. Soc., Workshop on Computer Vision, 1987; 315-317.
- [77]: Thomas S. Huang, James W. Burnett, Andrew G. Deczky; *the importance of phase in image processing filters*; *IEEE Trans. ASSP*, Vol. 23, No. 6, Dec. 1975; 529-542.
- [78]: R. M. B. Jackson; *target tracking using area correlation*; Int. Conf. on Advanced IR Detectors and Systems, 29-30 Oct. 1981, London, IEEE.
- [79]: S. M. Jaffey, M. W. Millman; *effects of rectification on image correlation*; SPIE Vol. 238, *Image Processing for Missile Guidance*, July/Aug. 1980; 178-183.
- [80]: Jaswant R. Jain, Anil K. Jain; *displacement measurement and its application in interframe image coding*; *IEEE Trans. on Communication Technology*, Dec. 1981, Vol. 29, No. 12; 1799-1808.

- [81]: Bahram Javidi; *nonlinear matched filter based optical correlation*; SPIE Vol. 1151, Optical Information Processing Systems and Architectures, 1989; 262-267.
- [82]: Allan D. Jepson, Michael R. M. Jenkin; *the fast computation of disparity from phase differences*; Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition, June 1989; 398-403.
- [83]: Zhixin Jiang, Weixue Lu, Dake Hu; *a new digital image registration algorithm based on the double spatial intensity gradients using pyramids*; Pattern Recognition Letters, 8, 1988; 335-340.
- [84]: Mansur R. Kabuka, E. S. McVey; *input-output characteristics for imaged transducers*; IEEE Jour. of Robotics and Automation, Vol. RA-2, No. 2, June 1986; 106-110.
- [85]: Hideo Kawai, Shinichi Tamura, Kazutaka Kani, Komyo Kariya; *eye movement analysis system using fundus images*; Pattern Recognition, 1986, Vol. 19, No. 1; 77-84.
- [86]: Ali E. Kayaalp; *high speed machine perception using passive sensing technology*; SPIE Vol. 1195, Mobile Robots IV, 1989; 33-43.
- [87]: Surender K. Kenue; *LANELOCK: detection of lane boundaries and vehicle tracking using image-processing techniques - Part II: Template matching algorithms*; SPIE Vol. 1195, Mobile Robots IV, 1989; 234-245.
- [88]: V. Z. Kepuska, S. Mason; *design and analysis of spatial image sequences*; NFP 23, 2. Jahresbericht, 5. Juni 1991.
- [89]: Harald Kirchner; *blockmatching with column and row oriented optimization*; Time-Varying Image Processing and Moving Object Recognition, 2, 1990, V. Cappellini (ed.); 280-288.
- [90]: Samuel P. Kozaitis, Atikom Roeksabutr; *image preprocessing for phase-only filtering*; SPIE Vol. 1151, Optical Information Processing Systems and Architectures, 1989; 154-164.
- [91]: C. D. Kuglin, W. G. Eppler; *map-matching techniques for use with multispectral/multitemporal data*; SPIE Vol. 238, Image Processing for Missile Guidance, 1980; 146-155.
- [92]: Klaus-Dieter Kuhnert; *Zur Echtzeit-Bildfolgenanalyse mit Vorwissen*; Dissertation, Universität der Bundeswehr, München, 1988.
- [93]: B. V. K. Vijaya Kumar, Christopher Carroll; *loss of optimality in cross correlators*; J. Opt. Soc. Am. A, Vol. 1, No. 4, April 1984, Optics and Image Science; 392-397.

- [94]: Chung J. Kuo; *theoretical expression for the correlation signal of nonlinear joint-transform correlators*; Applied Optics, 10 Oct. 1992, Vol. 31, No. 29; 6264-6271.
- [95]: Stanley N. Lapidus; *robust intelligent vision for the factory floor*; The Int. Electronic Imaging Exposition and Conf., Sep. 1984; 281-291.
- [96]: David Lavine, Lavean N. Kanal, Carlos A. Berenstein, Eric Slud, Charles Herman; *analysis of subpixel registration accuracy*; Proc. of the NASA Symp. on Mathematical Pattern Recognition and Image Analysis, 1982; 327ff.
- [97]: Dah-Jye Lee, Thomas F. Krile, Sunanda Mitra; *power cepstrum and spectrum techniques applied to image registration*; Applied Optics, 15 March 1988, Vol. 27, No. 6; 1099-1106.
- [98]: Kenneth G. Leib, Robert W. Brandstetter, Marvin D. Drake, Glen B. Franks, Ronald O. Siewert; *experimental comparison of optical binary phase-only filter and high pass matched filter correlation*; SPIE Vol. 1483, Signal and Image Processing Systems Performance Evaluation, Simulation, and Modelling, April 1991; 140-154.
- [99]: Scott D. Lindell, David L. Flannery; *ternary phase-amplitude filters for character recognition*; SPIE Vol. 1151, Optical Information Processing Systems and Architectures, 1989; 174-182.
- [100]: Xiaobo Li, Mahbuba Ferdousi, Marie Chen, Thanh Thuy Nguyen; *image matching with multiple templates*; Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition, June 1986; 610-613.
- [101]: J. M. Lloyd; *Thermal Imaging Systems*; Plenum Press, 1975.
- [102]: Thomas K. Lo, Gordon Gerson; *guidance system position update by multiple subarea correlation*; SPIE Vol. 186, Digital Processing of Aerial Images, 1979; 30-40.
- [103]: Kai-Oliver Ludwig, Heiko Neumann, Bernd Neumann; *robust estimation of local stereoscopic depth*; Proc. of the 2nd Int. Workshop on Robust Computer Vision, 1992, ISPRS, Bonn, Quality of Vision Algorithms, Förstner/Ruwiedel (eds.); 290-312.
- [104]: Hans Dieter Lüke; *Korrelationssignale; Korrelationsfolgen und Korrelationsarrays in Nachrichten- und Informationstechnik, Messtechnik und Optik*; Springer-Verlag, 1992.
- [105]: S. R. Maeng, K. Wohn; *real-time estimation of 2-D Motion for object tracking*; SPIE Vol. 1193, Intelligent Robots and Computer Vision VIII, Systems and Applications, 1989; 131-140.

- [106]: Misha A. Mahoswald, Curver Mead; *the silicon retina*; Scientific America, May 1991; 40-46.
- [107]: A. Mandanayake, W. K. Taylor; *specialised parallel hardware for flash recognition of window images in a television frame*; Third Int. Conf. on Image Processing and its Applications, Conf. Publication 307, IEE, 18-29 July 1989, Warwick, UK; 501-505.
- [108]: Avraham Margalit, Azriel Rosenfeld; *using feature probabilities to reduce the expected computational cost of template matching*; Int. Jou. on Computer Vision, Graphics, and Image Processing, Vol. 52, No. 1, Oct., 1990; 110-123.
- [109]: D. Marr, T. Poggio; *a computational theory of human stereo vision*; Proc. of the Royal Soc. of London, B. Biological Sciences, Vol. 204, 1979; 301-328.
- [110]: William C. Marshall; *evaluation of image tracker algorithms*; SPIE Vol. 1483, Signal and Image Processing Systems Performance Evaluation, Simulation, and Modelling, April 1991; 207-218.
- [111]: V. John Mathews; *multiplication free vector quantization using L1 distortion measure and its variants*; IEEE Trans. on Image Processing, Vol. 1, No. 1, Jan. 1992; 11-17.
- [112]: K. Matsuda, T. Tauda, T. Ito, S. Maki; *a new motion compensation coding scheme for video conferences*; IEEE Int. Conf. on Communications, May 14-17 1984, Amsterdam, The Netherlands; 234-237.
- [113]: Pat H. McIngvale, Robert D. Guyton; *experimental evaluation of video pre-processing algorithms for automatic target hand-off*; Southeast Conf. 1981, Proc., 5-8 April, Huntsville, Al, USA.
- [114]: J. Mendelsohn, David C. Englund; *matched filtering for robotic vision systems*; SPIE Proc., Vol. 579, Intelligent Robots and Computer Vision, 1985.
- [115]: John Merchant; *full-frame image registration by address modification*; SPIE Vol. 186, Digital Processing of Aerial Images, 1979; 49-52.
- [116]: Sunanda Mitra, Brian S. Nutter, Thomas F. Krile, Reay H. Brown; *automated method for fundus image registration and analysis*; Applied Optics, 15 March 1988, Vol. 27, No. 6; 1107-1112.
- [117]: Omid Mohtadi, Felix Safar, Jorge L. C. Sanz; *AGUILA: an automatic tube detection system*; SPIE Conf. on Applications of Artificial Intelligence X: Machine Vision and Robotics, 1992; 482-494.
- [118]: Perry A. Molley, K. Terry Stalker; *performance of the difference-squared error algorithm on an acousto-optic architecture*; SPIE Vol. 1151, Optical Information Processing Systems and Architectures, 1989; 436-444.

- [119]: Hans P. Moravec; *robot rover visual navigation*; UMI Research Press 1980/1981, Comp. Science: Artificial Intelligence.
- [120]: Hassan Mostafavi, Fred W. Smith; *image correlation with geometric distortion, Part I: Acquisition performance*; IEEE Trans. AES, Vol. 14, No. 3, May 1978; 487-493.
- [121]: Hassan Mostafavi, Fred W. Smith; *image correlation with geometric distortion, Part II: Effect on local accuracy*; IEEE Trans. AES, Vol. 14, No. 3, May 1978; 494-500.
- [122]: M. L. Nack; *temporal registration of multispectral digital satellite images using their edge images*; Proc. of the AAS/AIAA/Astrodynamics Specialist Conf., Nassau, Bahamas, July 28-30 1975, paper AAS75-104.
- [123]: Nasser M. Nasrabadi, Chang Y. Choo; *hopfield network for stereo correspondence*; IEEE Trans. on Neural Networks, Jan. 1992, Vol. 3, No. 1; 5-13.
- [124]: H. K. Nishihara; *practical real-time imaging stereo matcher*; Optical Engineering, 23(5), Sep. 1984; 536-545.
- [125]: H. K. Nishihara, N. G. Larson; *towards a real time implementation of the Marr and Poggio stereo matcher*; Proc. of the Image Understanding Workshop, April 1981, Washington, DC; 114-120.
- [126]: H. K. Nishihara, T. Poggio; *stereo vision for robotics*; ISRR 83 Conf., Bretton Woods, New Hampshire, 1983.
- [127]: Leslie Novak; *correlation algorithms for radar map matching*; IEEE Trans. AES, Vol. 14, No. 4, Jul. 1978; 641-648.
- [128]: Masatoshi Okutomi, Takeo Kanade; *a locally adaptive window for signal matching*; Int. Jou. of Computer Vision, 7:2, 1991; 143-162.
- [129]: Thomas J. Olson, David J. Coombs; *real-time vergence control for binocular robots*; Int. Jou. of Computer Vision, 7:1, 1991; 67-89.
- [130]: Rae-Hong Park, Woo Young Choi; *a new interpretation of the compass gradient edge operators*; Int. Jou. on Computer Vision, Graphics, and Image Processing, Vol. 47, No. 1, July, 1989; 259-265.
- [131]: Li Peng; *utilization of existing elevation data base to improve correlation on Kern DSR-11 analytical plotter*; 16th Congr. of the Int. Soc. for Photogrammetry and Remote Sensing; 466-475.
- [132]: Gerald W. Plunk; *simulation of multiple automatic video trackers*; SPIE Vol. 1111, Acquisition, Tracking, and Pointing III, March 1989; 172-184.

- [133]: William K. Pratt; *correlation techniques of image registration*; IEEE Trans. AES, Vol. 10, No. 3, May 1974; 353-358.
- [134]: William K. Pratt; *Digital Image Processing*; John Wiley & Sons, 1981; 551-567.
- [135]: H. S. Ranganath; *analysis of bilevel quantizers used in binary image correlators*; Image and Vision Computing, Vol. 6, No. 3, Aug. 1988 ; 193-197.
- [136]: Edward M. Riseman, Allan Hanson; *image understanding research at the University of Massachusetts*; Proc. of the Image Understanding Workshop, June 1983, Arlington, Virginia; 37-42.
- [137]: A. Rosenfeld, A. C. Kak; *Digital Picture Processing*, 2nd ed., Vol. 2, Comp. Science and Appl. Math.; 30-55.
- [138]: T. W. Ryan; *the prediction of cross-correlation accuracy in digital stereo-pair images*; Dissertation, University of Arizona, Tucson, 1980.
- [139]: T. W. Ryan, B. R. Hunt; *recognition of stereo-image cross-correlation errors*; in L. N. Kanal, A. Rosenfeld (eds.), Progress in Pattern Recognition, Vol. 1, 1981; 265-322.
- [140]: P. Schaeren, B. Schneuwly, W. Guggenbühl; *structure from motion in real time*; SPIE Conf.: Applications of AI 1993: Machine Vision and Robotics.
- [141]: D. Schmid, Ch. Umbricht; *Kamerasteuerung mit Signalprozessor*; Studienarbeit am Inst. für Elektronik der ETH Zürich, Wintersemester 1988/89.
- [142]: Bruno Schneuwly, *Ein Qualitätsmassstab für die Farbbildwiedergabe*, Dissertation ETH Zürich, 1993.
- [143]: Daniel G. Schwartz, George J. Klir; *fuzzy logic flowers in Japan*; IEEE Spectrum, July 1992; 32-35.
- [144]: Juan P. Secilla, Narciso Garcia, Jose L. Carrascosa; *locating templates in noisy environments*; Proc. of the 5th Scandinavian Conf. on Image Analysis, Stockholm, June 1987, Vol. 1; 101-108.
- [145]: Peter Seitz; *using local orientational information as image primitive for robust object recognition*; SPIE Proc., Vol. 1199, Visual Communication and Image Processing IV, 1989; 1630-1639.
- [146]: Jun Shen, Serge Castan; *fast binary image correlation algorithms for stereo vision*; Proc. of the 5th Scandinavian Conf. on Image Analysis, Stockholm, June 1987, Vol. 1; 317-323.

- [147]: Leah J. Siegel, Howard Jay Siegel, Arthur E. Feather; *parallel processing approaches to image correlation*; IEEE Trans. on Comp., Vol. 31, No. 1, Jan. 1982; 208-218.
- [148]: William M. Silver; *normalized correlation search in alignment, gauging, and inspection*; SPIE Vol. 755, Image Pattern Recognition: Algorithm Implementations, Techniques, and Technology, Jan. 1987; 23-34.
- [149]: Jeffrey A. Sloan, Lewis A. Holloway Jr.; *a self-serving optical correlator for tracking*; SPIE Vol. 1151, Optical Information Processing Systems and Architectures, 1989; 307-317.
- [150]: F. W. Smith, H. Mostafavi, T. L. Steding, R. S. Poulsen; *optimum windows for image registration*; SPIE Vol. 186, Digital Processing of Aerial Images, 1979; 98-107.
- [151]: L. D. R. Smith, G. P. Robinson, D. J. Stevenson, H. Sanghera, P. Quarendon; *a semi automatic computerised system for the production of quantitative moving three dimensional images of the heart and coronary arteries*; Second Int. Conf. on Image Processing and its Applications, IEE Conf. publ. 265, 24-26 June 1986; 72-76.
- [152]: R. Srinivasan, K. R. Rao; *predictive coding based on efficient motion estimation*; IEEE Int. Conf. on Communications, May 14-17 1984, Amsterdam, The Netherlands; 521-526.
- [153]: Thomas L. Steding, Fred W. Smith; *optimum filters for image registration*; SPIE Vol. 186, Digital Processing of Aerial Images, 1979; 88-97.
- [154]: Robin N. Strickland, Donald W. Sweeney; *optical flow computation in combustion image sequences*; Applied Optics, Vol. 27 No. 22, 15 Dec. 1988; 5213-5220.
- [155]: M. Svedlow, C. D. McGillem, P. E. Anuta; *experimental examination of similarity measures and preprocessing methods used for image registration*; Proc. IEEE 1976, Machine Processing of Remotely Sensed Data; 4A-9 bis 4A-17.
- [156]: M. Svedlow, C. D. McGillem, P. E. Anuta; *image registration: similarity measure and preprocessing method comparisons*; IEEE Trans. AES, Vol. 14, No. 1, Jan. 1978; 141-149.
- [157]: Larry N. Thibos; *image processing by the human eye*; SPIE Proc., Vol. 1199, Visual Communication and Image Processing IV, 1989; 1148-1153.
- [158]: Qi Tian, Michael N. Huhns; *algorithms for subpixel registration*; Computer Graphics and Image Processing, Vol. 35, No. 2, Aug. 1986; 220-233.

- [159]: Zhang Tianxu, Peng Jiaxiong, Lue Weixue; *an optimized searching algorithm for image matching*; 9th Int. Conf. on Pattern Recognition, 14-17 Nov. 1988, Rome Italy, Vol. 2; 1265-1267.
- [160]: M. Tistarelli, G. Sandini; *direct estimation of time-to-impact from optical flow*; Proc. of the IEEE Workshop on Visual Motion, Oct. 1991; 226-233.
- [161]: D. Y. Tseng, D. K. Conti, W. O. Eckhardt, R. Nevatia, K. E. Olin; *model-based scene matching*; SPIE Vol. 238, Image Processing for Missile Guidance, July/Aug. 1980; 225-231.
- [162]: J. D. Tubbs; *a note on binary template matching*; Pattern Recognition, 1989, Vol. 22, No. 4; 359-365.
- [163]: Gordon J. Vanderburg, Azriel Rosenfeld; *two-stage template matching*; IEEE Trans. on Comp., Vol. 26, No. 4, April 1977; 384-393.
- [164]: A. Venot, J. F. Lebruchec, J. C. Roucayrol; *a new class of similarity measures for robust image registration*; Computer Vision, Graphics, and Image Processing, 28, (1984), ; 176-184.
- [165]: R. Voles; *interpolating sampled cross-correlation surfaces of images for fine registration*; IEE Proc., Vol. 127, Part F, No. 5, Oct. 1980, Communications, Radar and signal processing; 401-404; 134.
- [166]: Gooitzen S. Van der Wal, Joseph O. Sinniger; *real time pyramid transform architecture*; SPIE Proc., Vol. 579, Intelligent Robots and Computer Vision, 1985.
- [167]: Masaki Watanabe, Yuichi Ohta; *cooperative integration of multiple stereo algorithms*; Proc. of the Third Int. Conf. on Computer Vision, Osaka, Japan, Dec. 1990; 476-480.
- [168]: Juyang Weng; *a theory of image matching*; Proc. of the Third Int. Conf. on Computer Vision, Osaka, Japan, Dec. 1990; 200-209.
- [169]: Hubert M. Wey; *Beiträge zur rauscharmen CCD-Signaldetektion*; Dissertation ETH Zürich, 1988.
- [170]: Urs Wirth; *Schneller Warper/Korrelator für SYDAMA 2*; interner Bericht des Inst. für Elektronik der ETH Zürich (Sommer 1993).
- [171]: Robert Y. Wong; *radar to optical scene matching*; SPIE Vol. 186, Digital Processing of Aerial Images, 1979; 108-114.
- [172]: Robert Y. Wong; *sequential scene matching using edge features*; IEEE Trans. AES, Vol. 14, No. 1, Jan. 1978; 128-140.

- [173]: Donald C. Wunsch, Robert J. Marks, Thomas P. Caudell, C. David Capps; *limitations of a class of binary phase-only filters*; Applied Optics, 10 Sep. 1992, Vol. 31, No. 26; 5681-5687.
- [174]: Zhong Chun Xiang, Liang Zhi Hui; *a heuristic search algorithm for image registration*; Proc. 1988 IEEE, 9th Int. Conf. on Pattern Recognition, Vol. 2; 731-735.
- [175]: Yehezkel Yeshurun, Eric L. Schwartz; *cepstral filtering on a columnar image architecture: a fast algorithm for binocular stereo segmentation*; IEEE Trans. PAMI, Vol. 11, No. 7, July 1989; 759-767.
- [176]: Andre Zaccarin, Bede Liu; *fast algorithms for block motion estimation*; IEEE Proc. ICASSP 1992, Vol. 3, M; III-449-452.
- [177]: Hong Zhenhua, Zhu Peiyong; *an improved SSDA applied in target tracking*; 9th Int. Conf. on Pattern Recognition, 14-17 Nov. 1988, Rome Italy, Vol. 2; 767-769.
- [178]: *Digital Signal Processing Applications with the TMS320 Family*. Theory, Algorithms, and Implementations. Vol. 3, 1990, Texas Instruments.

Anhang A.1: Nomenklatur

$b(x, y; Nr)$: *Bild* der Grösse 256x256 Pixel mit sechs Bit Helligkeitsauflösung (nichtlineare Auflösung der Kamera, Gammakorrektur). Der Parameter Nr bezeichnet die Position des Bildes innerhalb einer Bildsequenz.

$s(x, y)$: *Suchfenster*, in der Literatur oft als *region of interest* bezeichnet (ROI, Fig.26). Das rechteckige Suchfenster enthält alle Pixel, die für die Korrelation verwendet werden (ohne Vorverarbeitung!). In den Versuchen wird die Suchfenstergrösse jeweils derart angepasst, dass bei gegebener Mustergrösse 40x40 mögliche Objektpositionen resultieren. Dimension: xls mal yls ($xls=xlr+39$, $yls=yldr+39$).

$r(u, v)$: *Referenzfenster, Muster, template* (Fig.26). Das Referenzfenster repräsentiert das Grauwertmuster, dessen ähnlichstes Abbild im Suchfenster gesucht wird. In den Versuchen wurden quadratische Muster der Grösse 2x2 bis 25x25 Pixel verwendet. Dimension: $xlr=umax+1$, $yldr=vmax+1$.

$S(\omega_x, \omega_y), R(\omega_x, \omega_y)$: Spektrum des Suchfensters bzw. des Musters. Im Normalfall wird hiermit das Fourier-Spektrum gemeint. Denkbar ist z.B. aber auch das reellwertige Hartley-Spektrum.

$\overline{s(x, y)}, \bar{r}$: Lokaler Mittelwert von $s(x, y)$ und $r(u, v)$.

$$\overline{s(x, y)} = \sum_{v=0}^{vmax} \sum_{u=0}^{umax} s(x+u, y+v) \quad (42)$$

$$\bar{r} = \sum_{v=0}^{vmax} \sum_{u=0}^{umax} r(u, v) \quad (43)$$

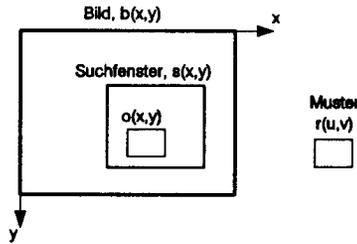


Fig.26: Definition des Bildes $b(x,y)$, Suchfensters $s(x,y)$, Musters $r(u,v)$ und Objektfensters $o(x,y)$. Das Objektfenster besitzt die gleichen Dimensionen wie das Muster; es ist Teil des Suchfensters und befindet sich am Ort des dem Muster ähnlichsten Abbildes (gemäß entsprechendem Ähnlichkeitsmass). Das Objektfenster ist ein Spezialfall eines Korrelationsfensters (Korrelationsfenster: Potentielles Objektfenster).

$s^*(x,y), r^*(u,v)$: Hochpass-gefilterte Version von $s(x,y)$ bzw. $r(u,v)$ (mittelwertfreier Laplaceoperator, siehe **UKKF** und **UKKFMF**):

$$s^*(x,y) = 8 \cdot s(x,y) - s(x-1,y-1) - s(x,y-1) - s(x+1,y-1) - s(x+1,y) - s(x+1,y+1) - s(x,y+1) - s(x-1,y+1) - s(x-1,y) \quad (44)$$

$s_{Sobel}(x,y), r_{Sobel}(u,v)$: Diese Bildausschnitte enthalten die Richtungen der lokalen Helligkeitsgradienten, welche mit dem Sobeloperator bestimmt wurden (siehe **SES1** und **SES2**):

$$v_{Sobel}(x,y) = \left\{ s(x+1,y+1) + 2 \cdot s(x,y+1) + s(x-1,y+1) \right\} - \left\{ s(x+1,y-1) + 2 \cdot s(x,y-1) + s(x-1,y-1) \right\} \quad (45)$$

$$h_{Sobel}(x,y) = \left\{ s(x+1,y+1) + 2 \cdot s(x+1,y) + s(x+1,y-1) \right\} - \left\{ s(x-1,y+1) + 2 \cdot s(x-1,y) + s(x-1,y-1) \right\} \quad (46)$$

$$s_{Sobel}(x,y) = \arctan \left\{ \frac{v_{Sobel}(x,y)}{h_{Sobel}(x,y)} \right\} \quad (47)$$

In den dokumentierten Versuchen wurde die Richtung in 32 Stufen quantisiert.

$s_{Kirsch}(x,y)$, $r_{Kirsch}(u,v)$: Diese Bildausschnitte enthalten die Richtungen der lokalen Helligkeitsgradienten, die mit Hilfe des *Kirschoperators* bestimmt wurden. Der Kirschoperator kann zwischen acht verschiedenen Richtungen unterscheiden, indem er alle seine Filtermasken gleichzeitig auf das Bild anwendet ([47]). Jede der 3x3 Pixel grossen Masken entspricht einer ausgeprägten Richtung. Diejenige Filtermaske – an einer Stelle (x,y) – deren Berechnung den grössten Absolutwert liefert, bestimmt, zusammen mit dem Vorzeichen des Ergebnisses, die lokale Helligkeitsorientierung (siehe Fig.27). Die Verwandtschaft von Sobel- und Kirsch- (und Prewitt-) Operator wird in [130] beschrieben.

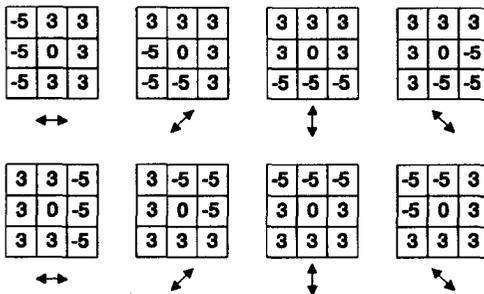


Fig.27: Definition der acht Filtermasken für den Kirschoperator. Er kann sowohl als Kanten- (Amplituden-) als auch als Richtungsdetektor verwendet werden. Die Maske mit dem betragsmässig grössten Resultat bestimmt, zusammen mit dem entsprechenden Vorzeichen, die lokale Orientierung des Helligkeitsgradienten.

$s_{bl}(x,y)$, $r_{bl}(u,v)$: Diese beiden Bildausschnitte stellen binäre Versionen von $s(x,y)$ und $r(u,v)$ dar. Die Binarisierung erfolgt durch die Vorzeichendetektion nach einer Fil-

terung mit dem *Laplace-Gauss-Operator* ([124]). Die binären Werte werden vom Korrelationskriterium als +1 und -1 interpretiert. Der Wert Null tritt also bei dieser Implementation nicht auf (die Leistungsunterschiede zwischen den beiden möglichen Implementationen – symmetrische und unsymmetrische Binarisierung – sind relativ gering)!

Laplace-Gauss-Operator:

$$\nabla^2 g = \left\{ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right\} \cdot g_\sigma(x, y) = \left\{ 1 - \frac{4r^2}{w^2} \right\} \cdot e^{\left(-\frac{4r^2}{w^2} \right)} \quad (48)$$

r^2 (nicht zu verwechseln mit $r(u, v)$, dem Muster!) entspricht dem Quadrat der Distanz zum Filtermaskenzentrum. Der Parameter w definiert die Bandgrenze des Filters und wurde bei den Versuchen auf alle ganzen Werte zwischen zwei und acht gesetzt. Im Prinzip müsste die Operatordimension unendlich gross gewählt werden. In den Versuchen wurde die Filtermaskengrösse aber jeweils derart gewählt, dass ausserhalb der Maske nur Filterkoeffizienten mit einem Absolutwert kleiner als 2^{-15} des Maximalwertes vorkommen. Daraus resultiert eine Maskengrösse von 7×7 , 11×11 , 15×15 , 19×19 , 25×25 und 29×29 Pixel für die verschiedenen w -Parameter ($w=2..8$, dieselbe Reihenfolge).

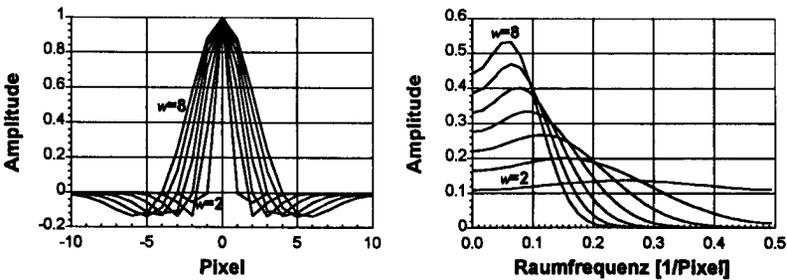


Fig.28: Filterkennlinie in der Bildebene sowie Amplitudengang des eindimensionalen Laplace-Gauss-Operators für die Werte 2..8 des w -Parameters.

$s_{b2}(x, y)$, $r_{b2}(u, v)$: Diese Bildausschnitte werden im Zusammenhang mit der binären Korrelation von Nack benötigt. Es handelt sich dabei um Hochpass-gefilterte und anschließend binarisierte Formen des Such- und Referenzfensters. Die adaptive Binarisierungsschwelle wurde jeweils für jedes einzelne Suchfenster derart gewählt, dass mindestens 15% der Pixel diese Schwelle überschritten (siehe [122] und [156]). $r_{b2}(u, v)$ ist nichts anderes als eine Kopie eines Bereichs des gefilterten Suchfensters (maximale Bildqualität, keine Störung!). Das vorgeschlagene Hochpassfilter – das beste der von Nack untersuchten Filter – kann als eine Art erweiterter Robertsoperator betrachtet werden:

$$\begin{aligned} s_{b2}(x, y) = & |s(x+1, y) - s(x-1, y)| + \\ & |s(x, y+1) - s(x, y-1)| + \\ & |s(x+1, y-1) - s(x-1, y+1)| + \\ & |s(x+1, y+1) - s(x-1, y-1)| \end{aligned} \quad (49)$$

Anhang A.2: Begriffe

Die folgende Liste enthält Begriffe, die im Verlauf des vorliegenden Textes verwendet wurden. Sie ist alphabetisch geordnet. Die quantitativen Aussagen gelten nur für die hier vorliegende Arbeit!

Ähnlichkeitsmass: Ein Ähnlichkeitsmass ist eine Rechenvorschrift, die zwei Datensätze (z.B. zwei Bildausschnitte: Muster und Korrelationsfenster) miteinander vergleicht. Der Zahlenbereich der Resultate kann bei den verschiedenen Algorithmen sehr unterschiedlich sein. Generell deuten grosse Zahlenwerte auf geringe Unterschiede zwischen Muster und Korrelationsfenster hin. Ähnlichkeitsmasse, genau wie auch Unähnlichkeitsmasse, gehören zur Klasse der Korrelationskriterien.

Auflösungspyramide: Diese Technik wird im Zusammenhang mit korrelativen Verfahren benutzt, wenn z.B. in sehr grossen Suchfenstern nach einem Muster gesucht werden muss. Das Bild wird in verschiedenen (geometrischen) Auflösungsstufen bearbeitet. Mit der Suche wird meist auf einer sehr groben Auflösungsstufe begonnen, die eine relativ schlechte Genauigkeit liefert. Das grobe Resultat wird benutzt, um die Suche in der nächst feineren Auflösungsstufe einzuschränken. Die Bilder der einzelnen Auflösungsstufen müssen vor der Unterabtastung mit einem Tiefpassfilter bearbeitet werden, um Aliasing-Probleme zu vermeiden (meist Gauss- oder Laplace-Gauss-Filter).

Bild: Einzelne Momentaufnahme einer Szene (kann auch rechnerisch generiert sein). Eine Reihe von Bildern $b(x,y;Nr)$ ergeben eine Bildsequenz. Bilddimensionen: 256x256 Pixel. Helligkeitsauflösung: 64 Stufen von schwarz bis weiss.

Bildsequenz: (Bildsequenz) Darunter wird hier eine geordnete Reihe von Bildern verstanden, die ein und dieselbe Szene mit von Bild zu Bild abnehmender Qualität beinhaltet.

Bildtransformation: Änderung der Bildgeometrie oder der Pixel-Intensitätswerte.

Depth from motion: Bei diesem Algorithmus wird, im Gegensatz zum üblichen Stereosicht-Verfahren, nur eine einzige Videokamera verwendet.

Dieser zweidimensionale Sensor wird aber derart bewegt, dass die im einfachsten Fall unbewegliche Szene von einer ganzen Reihe von Aufnahmepositionen her betrachtet werden kann. Während der Bewegung werden möglichst viele, interessante Regionen der Szene beobachtet und deren (2D-) Positionen ausgewertet. Die Daten dieser Verfolgung (Tracking), sowie die Kenntnis der Kamerabewegung lässt die Tiefenbestimmung der beobachtbaren 3D-Oberflächen zu ([140]). Der Vorteil dieses Verfahrens besteht in der Vereinfachung des Korrespondenzproblems. Aufeinanderfolgende Bilder unterscheiden sich in der Regel nur geringfügig und machen darum die Zuordnung unkritisch.

Detektionsrate: Die Detektionsrate beschreibt den Anteil erreichter tolerierbarer Positionierungen gemessen an der momentan betrachteten Menge der Positionierungen. Die (Unter-) Menge der betrachteten Positionierungen kann sich z.B. durch den Bildstörungstyp, die Bildstörungsstärke oder die Mustergröße auszeichnen.

Einfangbereich: Der Einfangbereich ist derjenige Teil der Korrelationsfläche, der in stetigem Anstieg bzw. Abfall zum Extremwert der Korrelationsfunktion führt.

Epipolar line: Dieser Begriff stammt aus der Stereobilderfassung. Eine epipolar line bezeichnet eine Kurve, die durch Transformation eines Abbildungsstrahls des ersten Sensors in die zweite (2D-) Sensorebene projiziert wurde. Nur entlang dieser Kurve muss bei Stereoanwendungen tatsächlich korreliert werden. Zur Bestimmung der Transformationsparameter ist die Kenntnis der relativen Ausrichtungen der beteiligten Kameras notwendig.

Fehlerhistogramm: Das Fehlerhistogramm zeigt die Häufigkeit der verschiedenen Abweichungen – aufgeteilt in die Fehlerklassen – zwischen berechneter und wahrer Objektposition.

Fehlerklasse: Ein Korrelationsresultat wird anhand der Toleranzringe in die entsprechende Fehlerklasse eingeteilt.

Fehlerrate: Anteil untolerierbarer Positionierungen (siehe *Detektionsrate*).

Flächen-Korrelationskriterien: Unter Flächen-Korrelationskriterien (*area correlation*) sollen Algorithmen verstanden werden, die auf (Un-)Ähnlichkeitsmassen basieren und als Informationsgrundlage direkt die (evtl. vorverarbeiteten) Intensitätswerte der Pixel verwenden.

Invarianz: Die Invarianz eines Korrelationskriteriums gegenüber einer Störung bedeutet, dass sich weder die Reihenfolge der Ähnlichkeitswerte noch die Korrelationswerte selber durch die betreffende Bild-Transformation ändern. Wegen der Intensitätsquantisierung und anderen Nichtidealitäten kann die vollständige Invarianz in der Praxis kaum je beobachtet werden!

Koordinatensystem: Für das Bild wie auch für das Such- und Referenzfenster wird ein in der Bildverarbeitung weit verbreitetes, lokales Koordinatensystem verwendet, das den Ursprung in der linken oberen Ecke des jeweiligen Datenfeldes besitzt (Fig.26 auf Seite 142). Das Zentrum aller Pixel liegt auf ganzzahligen Koordinatenschnittpunkten.

Koordinatentransformation: Die Transformation von der kartesischen $\{x,y\}$ - zur polaren $\{\rho,\varphi\}$ -Darstellung der Bilder vereinfacht z.B. die Suche nach gegenüber dem Referenzfenster verdrehten Helligkeitsmustern. Natürlich wird dafür die Detektion von translatorischen verschobenen Objekten erschwert.

Korrelationsfenster: Das Korrelationsfenster ist ein Ausschnitt aus dem Suchfenster mit derselben Grösse wie das Referenzfenster (Muster). Für jedes mögliche Korrelationsfenster im Suchfenster wird ein (Un-) Ähnlichkeitswert bezüglich des gesuchten Musters ermittelt. Das Suchfenster enthält total $\{x_{ls}-u_{max}\} \{y_{ls}-v_{max}\} = \{x_{ls}-x_{lr}+1\} \{y_{ls}-y_{lr}+1\}$ Korrelationsfenster (40x40).

Korrelationsfensterenergie: Sie entspricht der Energie im jeweiligen Korrelationsfenster an der Position (x,y) : $\sum_{v=0}^{u_{max}} \sum_{u=0}^{v_{max}} s^2(x+u, y+v)$ (siehe auch: *lokale Suchfensterenergie*).

Korrelationsfläche: Die Korrelationsfläche umfasst alle im Suchfenster enthaltenen Positionen (x,y) der Korrelationsfenster.

Korrelationskriterium: Rechenvorschrift, die auf ein Muster und ein Korrelationsfenster angewendet wird und eine Aussage über deren (Un-) Ähnlichkeit macht. Der Algorithmus schliesst hier evtl. auch eine Vorverarbeitung der Helligkeitswerte mit ein.

Korrespondenzproblem: Dieser Begriff bezeichnet das Problem der Zuordnung korrespondierender Regionen in zwei oder mehr Bildern, die

von der selben Szene, aber unter verschiedenen Bedingungen (z.B. anderer Kamerastandort) aufgenommen wurden.

Lokale Suchfensterenergie: Sie entspricht der Energie im jeweiligen Korrelationsfenster an der Position (x,y) : $\sum_{v=0}^{vmax} \sum_{u=0}^{umax} s^2(x+u, y+v)$ (siehe auch *Korrelationsfensterenergie*).

Matching: Mit Matching, Positionierung oder Musterdetektion wird hier der gesamte Prozess der Bestimmung der (abgetasteten und geometrisch beschränkten) Ähnlichkeitsfunktion an jeder Position in der Korrelationsfläche sowie die Bestimmung des Ortes maximaler Übereinstimmung zwischen Muster und Korrelationsfenster – im Sinne des angewendeten Algorithmus' – bezeichnet.

Maximumstörungsstärke: Dieser Zahlenwert (**MS**) bezeichnet für die **ZOOM**-Störung (geometrische Expansion) die maximale, zwischen den Datenstützstellen linear interpolierte Expansion, bei der mit der optimalen Mustergrösse eine 80 prozentige Detektionsrate erreicht werden kann (siehe auch *optimale Mustergrösse*, **OG**).

MIMD-Rechner: Mehrprozessor-Rechner, dessen Prozessoren zur selben Zeit verschiedene Instruktionen auf unterschiedlichen Daten ausführen (MIMD: multiple instruction multiple data).

Mittlere Detektionsrate: Die mittlere Detektionsrate bezeichnet den über alle Mustergrössen und Bildstörungsstärken gemittelten Anteil tolerierbarer Positionierungen (nicht gemittelt über die Störungstypen!).

Muster: Das Muster oder Referenzfenster entspricht bis auf mögliche Bildstörungen im Suchfenster dem Helligkeitsmuster der gesuchten Bildregion. Es enthält in den Versuchen keine Störungen. Dimension: $xlr=\{umax+1\}$, $ylr=\{vmax+1\}$, 2x2 bis 25x25 Pixel.

Objektfenster: Das Objektfenster ist eines der Korrelationsfenster im Suchfenster und befindet sich am Ort maximaler Übereinstimmung zwischen Muster und Korrelationsfenster (gemäss dem gewählten (Un-) Ähnlichkeitsmass). Diese Position wird *Objektposition* genannt.

Objektposition: Die Objektposition bezeichnet den Ort des Objektfensters.

Optimale Mustergrösse: Die optimale Mustergrösse (**OG**) kennzeichnet diejenige Mustergrösse, die die grösste Robustheit gegenüber der

ZOOM-Störung aufweist (Beurteilung gemäss Durchstosspunkt der 80%-Detektionsratenschwelle, siehe auch *Maximumstörungsstärke*, **MS**).

Pixel: Aus dem Begriff *picture element* entstanden. Einzelner Punkt des Bildes mit einer ihm eigenen Helligkeit (Helligkeitsdynamik: 0..63 in unseren Tests).

Plausibilitätsbetrachtungen: Plausibilitätsbetrachtungen werden zur Detektion von fehlerhaften oder unzuverlässigen Positionierungsergebnissen benutzt. Unkorrekte Resultate rühren z.B. von teilweise verdeckten Objekten her. Plausibilitätsbetrachtungen beinhalten beispielsweise die Kontrolle des Korrelationswerts oder gewisser Merkmale des Musters (z.B. mittlere Helligkeit) oder bei Tracking-Anwendungen die Kontrolle der Übereinstimmung der neu gefundenen Position mit der bisherigen Objektbewegung. Diese Tests haben, bei erschwerten Bedingungen, entscheidenden Einfluss auf die Leistungsfähigkeit eines Korrelatorsystems.

Positionierung: Mit Positionierung wird die Ermittlung derjenigen Position im Suchfenster bezeichnet, deren Umgebung, gemäss dem gewählten Korrelationsalgorithmus, am besten mit dem Muster übereinstimmt (siehe auch *Matching*).

Präzision: Positionsgenauigkeit mit der ein Muster detektiert wird, wobei nur tolerierbare Positionierungen zur Präzisionsbestimmung beitragen. Die untolerierbaren Positionierungen bestimmen dagegen die Fehler-rate.

Referenzfenster: Das Referenzfenster oder Muster entspricht bis auf mögliche Bildstörungen im Suchfenster dem Helligkeitsmuster der gesuchten Bildregion. Es enthält in den Versuchen keine Störungen. Dimension: $xlr=\{umax+1\}$, $ylr=\{vmax+1\}$, 2x2 bis 25x25 Pixel.

Robustheit: Die Robustheit eines Korrelationskriteriums gegenüber einer Bildtransformation signalisiert, dass sich durch die Bildstörung zwar die Zahlenwerte des betreffenden Ähnlichkeitsmasses ändern, nicht aber deren grössenmässige Reihenfolge (z.B. der Extremwert des Ähnlichkeitsmasses ändert seine Position nicht durch die Bildtransformation; siehe auch *Invarianz*).

Sequential similarity detection: Dieses Verfahren kann zur Reduktion des Rechenaufwandes bei allen rein akkumulativen Korrelationsverfahren, wie z.B. der **SSD** und der **SAD**, eingesetzt werden: Die Ermittlung eines einzelnen Korrelationswerts wird abgebrochen, sobald das Zwischenresultat eine bestimmte Korrelationsgüte unterschreitet. Es ist zu beachten, dass das Positionierungsergebnis bei korrekter Wahl der Schwelle mit demjenigen der vollständigen Rechnung übereinstimmen muss. Bei ungünstiger Wahl der Schwelle kann evtl. keine oder keine eindeutige Position gefunden werden. In der Literatur wird dieser Begriff oft durch SSD (oder SSDA) abgekürzt. Dieses Akronym ist hier aber schon durch das Unähnlichkeitsmass '**SSD**' (sum of squared differences, Summe der Differenzenquadrate) belegt.

Sequenzenset: (Set von Bildsequenzen) Ein Set von Bildsequenzen ist eine Gruppe von Bildsequenzen mit gemeinsamem Anfangsbild (siehe Fig.1 auf Seite 31). Ein Set umfasst die Sequenzen mit allen Störungstypen.

Signal-zu-Rauschleistungsverhältnis: Darunter wird hier das Leistungsverhältnis des *unverrauschten* Signals zum Rauschanteil verstanden (dies steht im Gegensatz zur Definition in [62], wo das Verhältnis der Leistungen des verrauschten Signals zum Rauschanteil gebildet wird).

SIMD-Rechner: Mehrprozessor-Rechner, dessen Prozessoren zur selben Zeit die gleichen Instruktionen (evtl. maskiert) auf unterschiedlichen Daten ausführen (SIMD: single instruction multiple data).

Störungstyp: Ausprägung (nicht Stärke!) einer Bildstörung: Z.B. Intensitätsschwankungen, Gauss'sches Rauschen, 'Salt and Pepper' Rauschen, Schärfe- und Grössenveränderungen.

Suchfenster: Um Rechenzeit zu sparen und die Wahrscheinlichkeit von Fehlpositionierungen zu reduzieren, wird die Objektsuche häufig auf einen kleinen Bildausschnitt – das Suchfenster – beschränkt. Für optimale Resultate muss in Anwendungen die Position und die Grösse (und evtl. die Form) des Suchfensters durch Vorwissen derart gewählt werden, damit die Wahrscheinlichkeit, dass das Objekt im Suchfenster enthalten ist, *maximiert* wird. Ausserhalb des Suchfensters liegende Pixel beeinflussen das Korrelationsresultat in keiner Weise. Dimensionen des (evtl. gefilterten) Suchfensters: $xls=40+umax$, $yls=40+vmax$.

Suchstrategie: Vielfach – und auch in den hier dokumentierten Versuchen – wird bei den korrelativen Verfahren das Ähnlichkeitsmass an allen

möglichen Positionen im Suchfenster ausgewertet. Zur Reduktion des Rechenaufwands kann unter der Annahme von genügend 'weichen' Korrelationsfunktionen (was bei den behandelten Beispielen nicht generell zutrifft) z.B. nach der sog. 'hill climbing' Methode oder nach einer beliebig anderen Suchstrategie vorgegangen werden. Suchstrategien haben aber meist den Nachteil, dass benachbarte Zwischenresultate der Ähnlichkeitsmasse nicht zur Reduktion des Rechenaufwandes verwendet werden können, da im allgemeinen nicht angrenzende Korrelationswerte bestimmt werden.

Symbolische Verfahren: Symbolische Korrelationsverfahren extrahieren aus der Matrix von Helligkeitswerten zuerst eine (nicht mehr an die Matrixform gebundene) symbolische Merkmals-Beschreibung des Suchfensters bzw. Musters. Im Suchfenster wird aufgrund dieser Beschreibung nach derjenigen Stelle gesucht, die mit der Musterbeschreibung am besten übereinstimmt.

Toleranzkreis/Toleranzgebiet: Ein Toleranzkreis ist, in diesem Zusammenhang, ein kreisförmiges Gebiet innerhalb der Korrelationsfläche, in dessen Zentrum sich die wahre Objektposition befindet (die Ränder der Korrelationsfläche können das ansonsten runde Gebiet durch gerade Teilstücke begrenzen, Fig.13 auf Seite 52). Mit dem Toleranzkreisradius r_t kann die Ausdehnung des Toleranzgebiets definiert werden. Die Radien der hier untersuchten Toleranzgebiete genügen der Formel:

$$r_t(i) = i \cdot \sqrt{2}, \quad i = 0, 1, \dots, n \quad (50)$$

Toleranzring: Ein Toleranzring ist ein ringförmiges Gebiet, das sich gemäss Mengenlehre folgendermassen beschreiben lässt:

$$TR_i = KF \cap TK_i \cap \overline{TK_{i-1}}, \quad i > 0 \quad (51)$$

KF : Menge der Pixel im Suchfenster;

TK_i : Menge der Pixel im Toleranzkreis mit Grenzradius $r_t(i) \leq i\sqrt{2}$

TR_i : Menge der Pixel im Toleranzring i

Tolerierbare Position: Unter einer tolerierbaren Position soll eine Position verstanden werden, die innerhalb eines spezifizierten Toleranzgebiets liegt. Der Betrag der Vektordistanz zwischen wahrer und berechneter Objektposition muss für kreisförmige Toleranzgebiete also kleiner oder gleich dem Toleranzkreisradius r_t sein. Für diesen Radius kann kein

allgemeingültiger Wert angegeben werden. Vielmehr muss er anwendungsbezogen bestimmt werden (z.B. $r_t = \sqrt{2}$ Pixel).

Tracking: Unter Tracking soll hier das Verfolgen eines Helligkeitsmusters in einer Videobildsequenz (d.h. über die Zeit) verstanden werden.

Unähnlichkeitsmass: Ein Unähnlichkeitsmass – in diesem Zusammenhang – gibt die zahlenmässige Diskrepanz zwischen Muster und Objektfenster an. Der Zahlenbereich kann bei den verschiedenen Algorithmen sehr unterschiedlich sein. Prinzipiell deuten grosse Zahlenwerte auf erhebliche Unterschiede zwischen Muster und Korrelationsfenster hin. Ein Unähnlichkeitsmass kann z.B. durch einfache Negation des Zahlenwerts in ein Ähnlichkeitsmass verwandelt werden. Unähnlichkeitsmasse, genau wie auch Ähnlichkeitsmasse, gehören zur Klasse der Korrelationskriterien.

Untolerierbare Position: Unter einer untolerierbaren Position soll eine Position verstanden werden, die ausserhalb des Toleranzgebiets liegt. Die Grenze zwischen tolerierbaren und untolerierbaren Positionen muss anwendungsspezifisch gewählt werden (z.B. $r_t = \sqrt{2}$ Pixel).

Varianzenverhältnis: Verhältnis der (unverrauschten) Signalvarianz zur Rauschvarianz. Dies ist ein besseres Bildqualitätsmass im Zusammenhang mit Matching-Algorithmen als das Signal-zu-Rauschleistungsverhältnis, da eigentlich alle korrelativen Matching-Verfahren eine bestimmte Signatur/Signal-Variation suchen ([128]).

Wahres Objektfenster: Dasjenige Korrelationsfenster, das objektiv (unabhängig vom Korrelationskriterium) die dem Muster ähnlichste Bildregion enthält.

Wahre Objektposition: Ort des wahren Objektfensters. In der Regel ist, im Gegensatz zu den hier diskutierten Untersuchungen, diese Position unbekannt!

Anhang B: Bildmaterial

Alle Bilder wurden im Format 256x256 Pixel mit sechs Bit Grauwertauf-
lösung aufgenommen. Aus Platzgründen werden im folgenden nur Stich-
proben aus dem Bildmaterial gezeigt. Um keine unerwünschten Nebeneffek-
te zu verursachen (z.B. Betonung der Kanten), wurden die Bilder für den
Druck nur mit Hilfe einer zero order hold Interpolation vergrößert. Das
verwendete Raster entspricht einem *ordered dithering*, das zwar eine regu-
läre Struktur auf das Bild prägt, sich aber relativ gut kopieren lässt. Da kei-
nerlei Kontrast- oder Helligkeits-ändernde Operationen bei der Wiedergabe
vorgenommen wurden, teilweise aber auch wegen dem nicht optimalen Wie-
dergabeprozess, sind die Bildinhalte der beiden letzten Bilder der Sequenzen
IRIS nicht mehr erkennbar. Insbesondere das jeweils letzte, maximal verän-
derte Bild dieser Sequenzen enthält aber tatsächlich nur noch sehr wenig
Information (siehe auch Kapitel 4.2: *Generierung der Bildsequenzen*).

Bei den Unschärfesequenzen (**FOCUS**) ist zu beachten, dass durch die
nicht verschwindende Tiefpassfilterkerngrösse ein nicht definierter, 16 Pixel
breiter, willkürlich auf schwarz gesetzter Rand im Bild entsteht. Allen Bil-
dern wurde zudem ein schmaler, umschreibender, schwarzer Rand bei-
gefügt, der keinerlei Information enthält und rein ästhetisch begründet ist.

Organisation der Darstellungen:

Pro Sequenzenset werden je sechs Seiten reserviert. Die Sequenzen **IRIS**,
NOISE, **SALTY** und **FOCUS** nehmen jeweils eine Seite in Anspruch. Wie
in Kapitel 4: *Versuche und Versuchsbedingungen* beschrieben, werden für
jede logische Sequenz **ZOOM** zwei Sequenzen gebraucht: Eine für die un-
gerade und eine für die gerade nummerierten Muster. Obwohl die Unterschie-
de auf dem Papier gering sind, werden der Vollständigkeit halber beide Se-
quenzen dokumentiert.

Reihenfolge der Bildsequenzen:

• Set 1, **TEXTURES**

IRIS (Bilder 0, 1, 2, 3, 5, 6)

NOISE (Bilder 0, 3, 6, 9, 12, 15)

SALTY (Bilder 0, 2, 4, 6, 7, 8)

FOCUS (Bilder 0, 2, 5, 7, 10, 13)

ZOOM (gerade Musternummern, Bilder 0, 1, 3, 5, 6, 7)

ZOOM (ungerade Musternummern, Bilder 0, 1, 3, 5, 6, 7)

• Set 2, **EDGES**

IRIS (Bilder 0, 1, 2, 3, 5, 6)

NOISE (Bilder 0, 3, 6, 9, 12, 15)

SALTY (Bilder 0, 2, 4, 6, 7, 8)

FOCUS (Bilder 0, 2, 5, 7, 10, 13)

ZOOM (gerade Musternummern, Bilder 0, 1, 3, 5, 6, 7)

ZOOM (ungerade Musternummern, Bilder 0, 1, 3, 5, 6, 7)

• Set 3, **CHAOS**

IRIS (Bilder 0, 1, 2, 3, 5, 6)

NOISE (Bilder 0, 3, 6, 9, 12, 15)

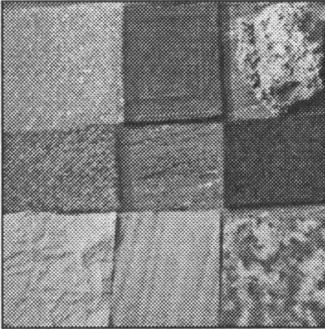
SALTY (Bilder 0, 2, 4, 6, 7, 8)

FOCUS (Bilder 0, 2, 5, 7, 10, 13)

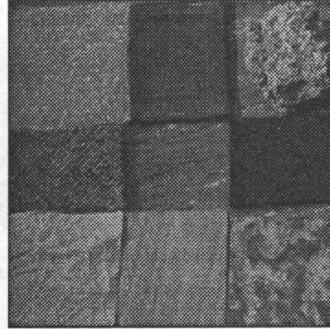
ZOOM (gerade Musternummern, Bilder 0, 1, 3, 5, 6, 7)

ZOOM (ungerade Musternummern, Bilder 0, 1, 3, 5, 6, 7)

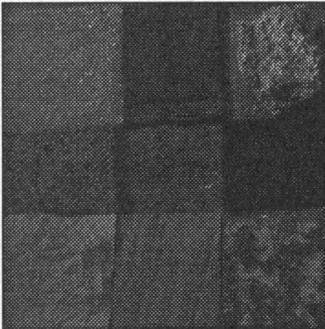
Sequenz IRIS, Set 1 (TEXTURES):



Bildnummer: 0



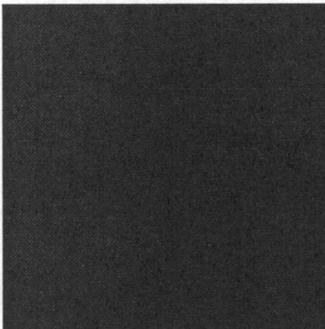
Bildnummer: 1



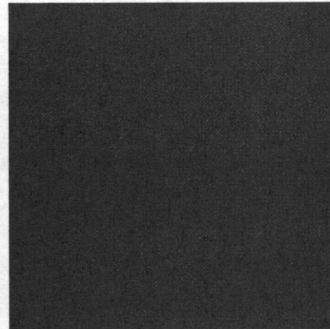
Bildnummer: 2



Bildnummer: 3

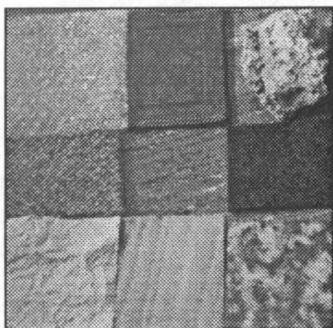


Bildnummer: 5

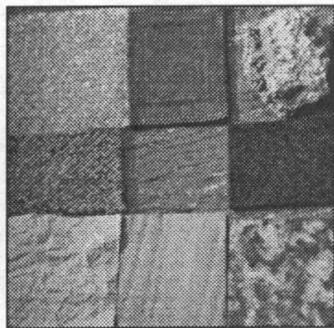


Bildnummer: 6

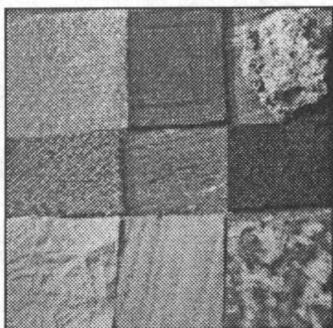
Sequenz **NOISE**, Set 1 (**TEXTURES**):



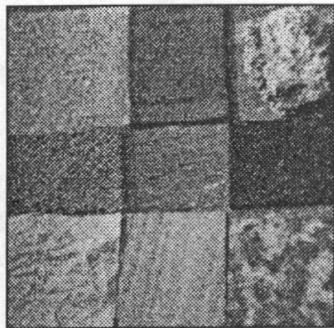
Bildnummer: 0



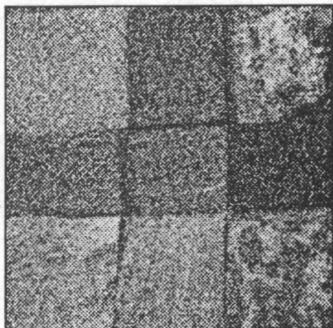
Bildnummer: 3



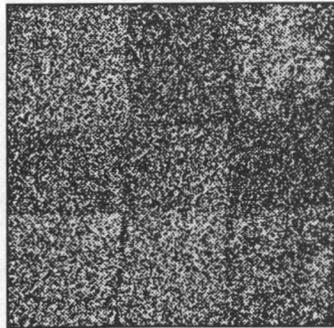
Bildnummer: 6



Bildnummer: 9

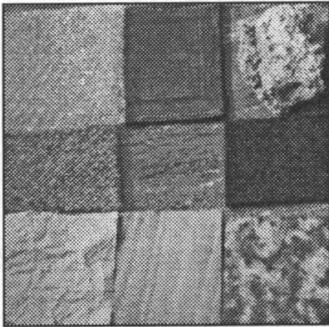


Bildnummer: 12

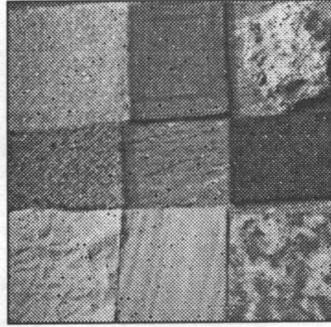


Bildnummer: 15

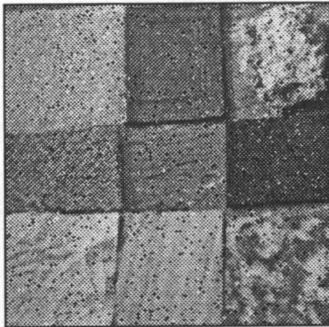
Sequenz **SALTY**, Set 1 (**TEXTURES**):



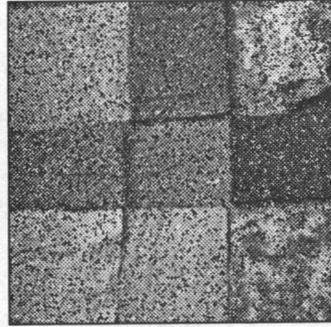
Bildnummer: 0



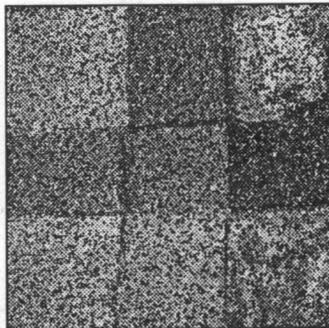
Bildnummer: 2



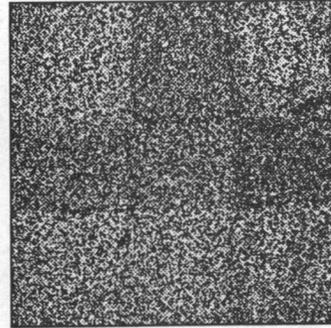
Bildnummer: 4



Bildnummer: 6

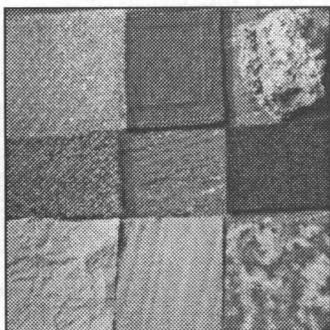


Bildnummer: 7

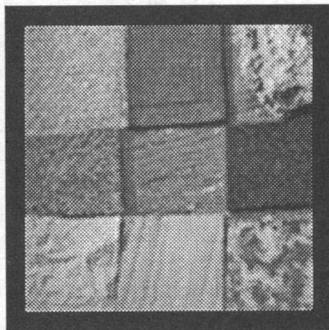


Bildnummer: 8

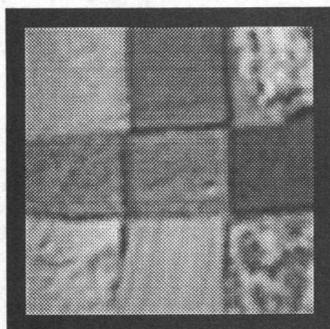
Sequenz **FOCUS**, Set 1 (**TEXTURES**):



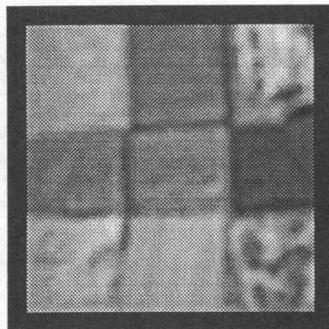
Bildnummer: 0



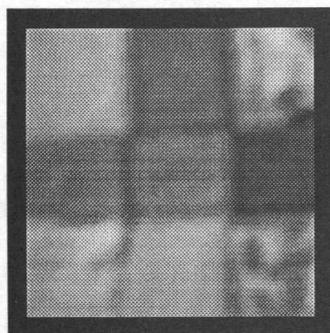
Bildnummer: 2



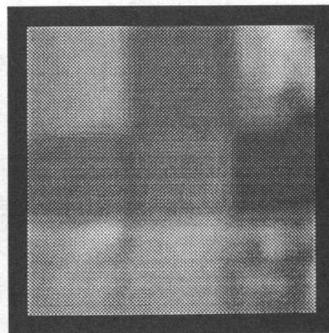
Bildnummer: 5



Bildnummer: 7

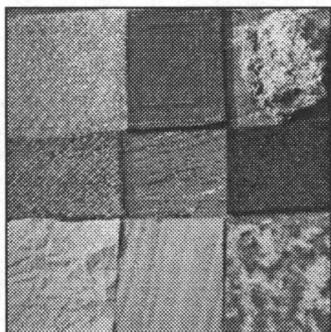


Bildnummer: 10

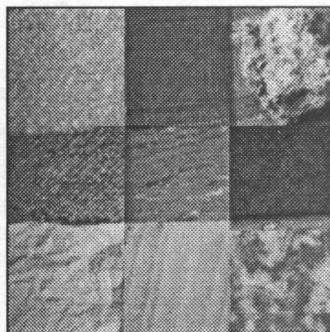


Bildnummer: 13

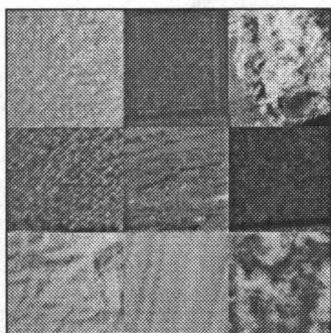
Sequenz **ZOOM** (gerade Musternummern), Set 1 (**TEXTURES**):



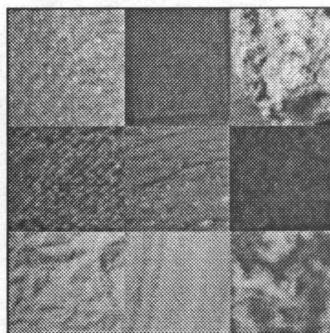
Bildnummer: 0



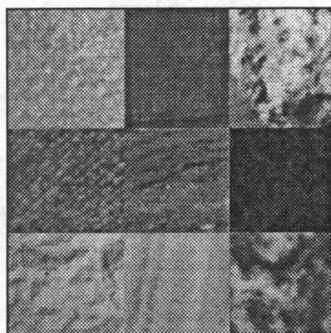
Bildnummer: 1



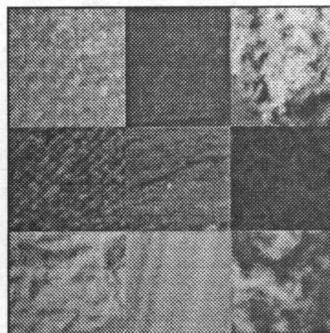
Bildnummer: 3



Bildnummer: 5

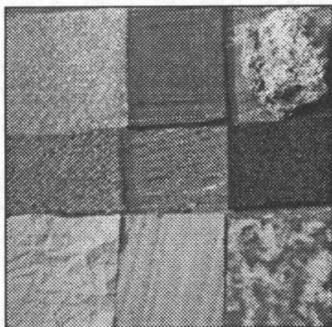


Bildnummer: 6

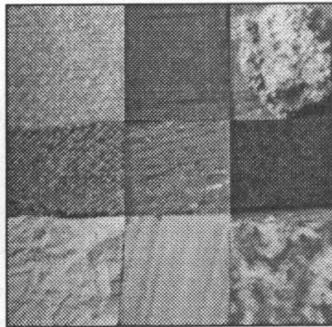


Bildnummer: 7

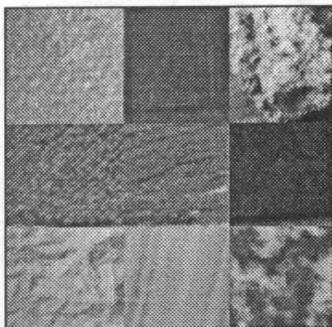
Sequenz **ZOOM** (ungerade Musternummern), Set 1 (**TEXTURES**):



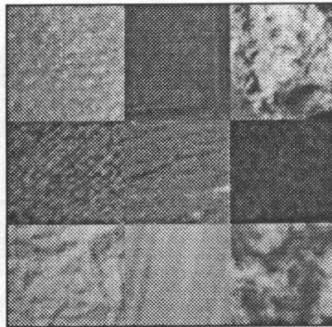
Bildnummer: 0



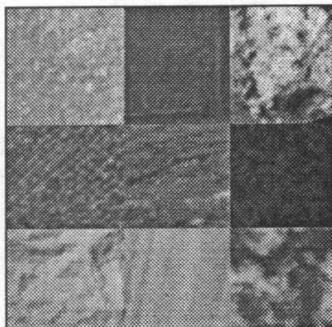
Bildnummer: 1



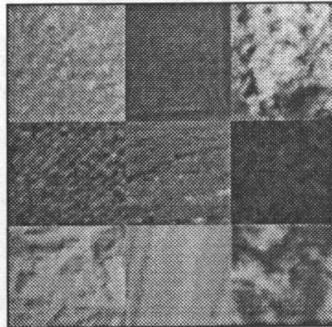
Bildnummer: 3



Bildnummer: 5

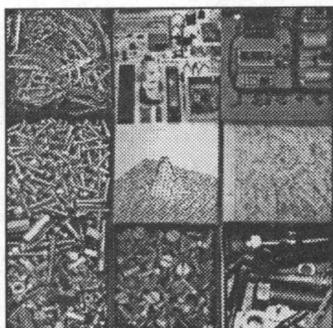


Bildnummer: 6

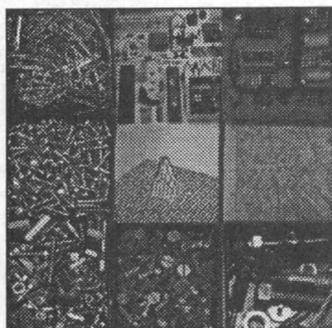


Bildnummer: 7

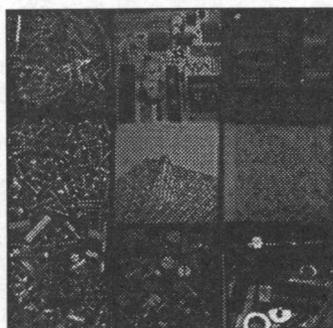
Sequenz IRIS, Set 2 (EDGES):



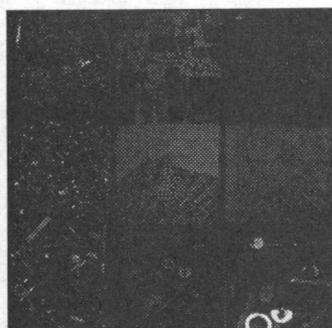
Bildnummer: 0



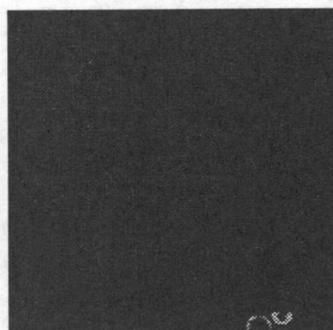
Bildnummer: 1



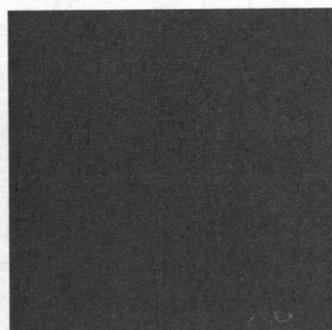
Bildnummer: 2



Bildnummer: 3

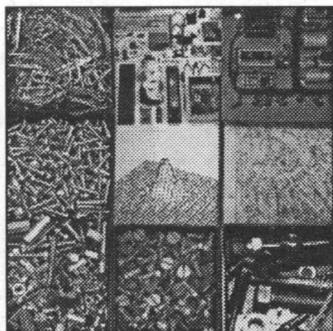


Bildnummer: 5

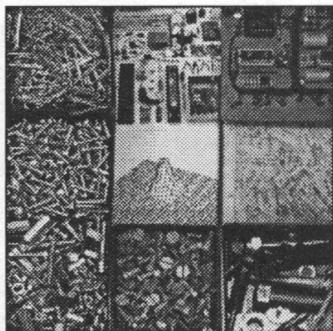


Bildnummer: 6

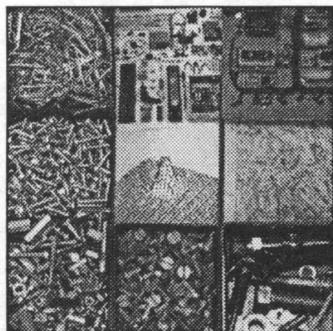
Sequenz **NOISE**, Set 2 (**EDGES**):



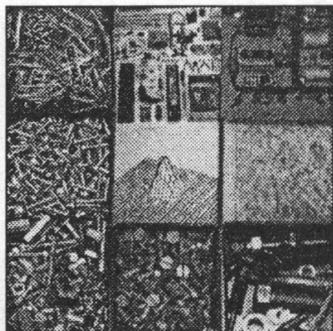
Bildnummer: 0



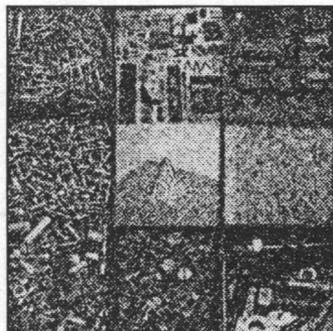
Bildnummer: 3



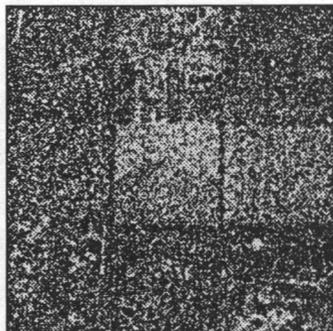
Bildnummer: 6



Bildnummer: 9

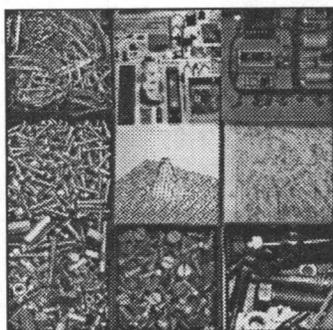


Bildnummer: 12

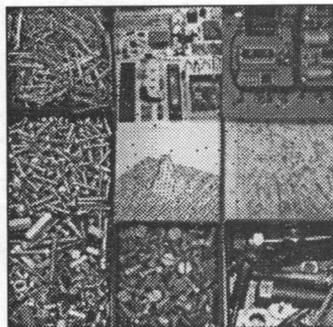


Bildnummer: 15

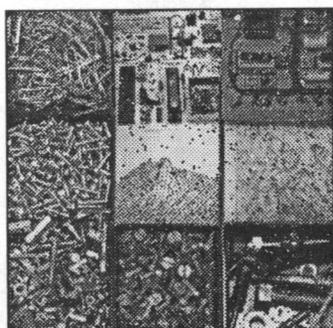
Sequenz **SALTY**, Set 2 (**EDGES**):



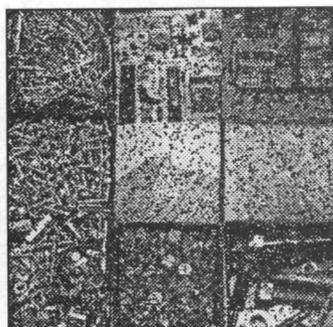
Bildnummer: 0



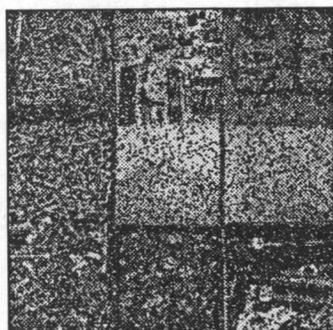
Bildnummer: 2



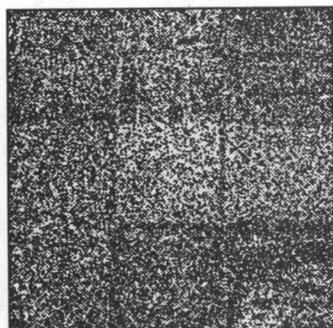
Bildnummer: 4



Bildnummer: 6

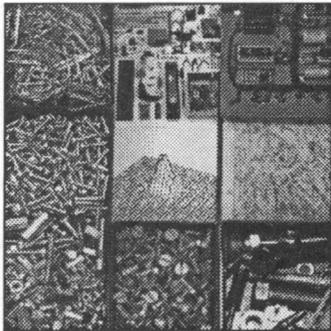


Bildnummer: 7

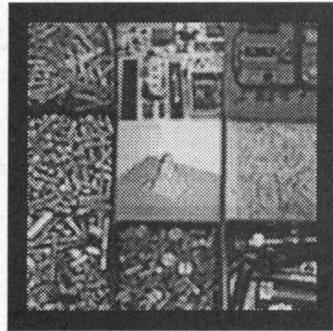


Bildnummer: 8

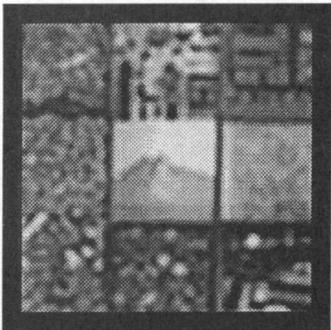
Sequenz **FOCUS**, Set 2 (**EDGES**):



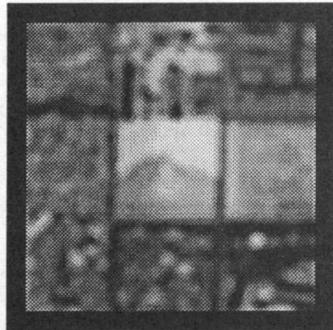
Bildnummer: 0



Bildnummer: 2



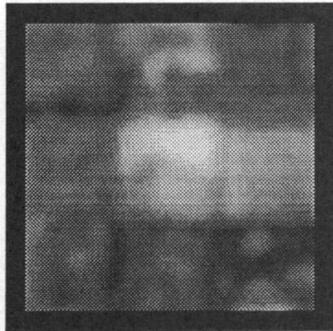
Bildnummer: 5



Bildnummer: 7

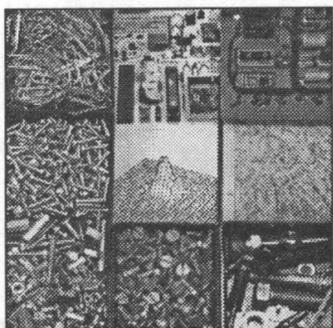


Bildnummer: 10

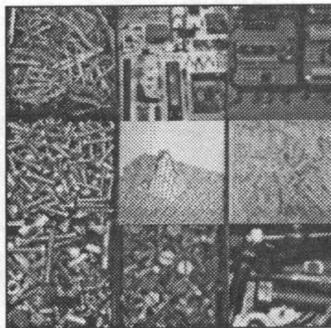


Bildnummer: 13

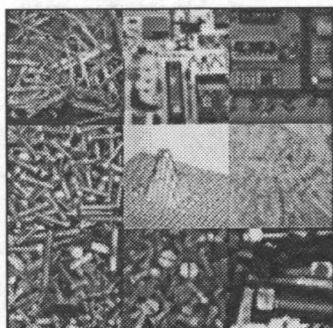
Sequenz **ZOOM** (gerade Musternummern), Set 2 (**EDGES**):



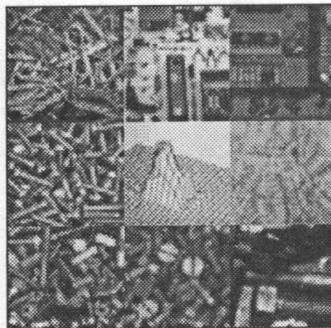
Bildnummer: 0



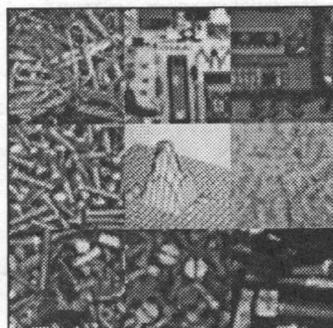
Bildnummer: 1



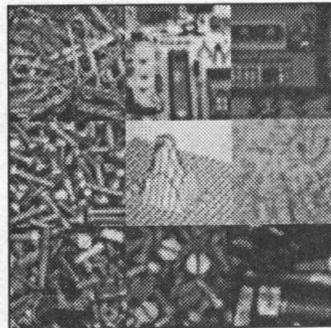
Bildnummer: 3



Bildnummer: 5

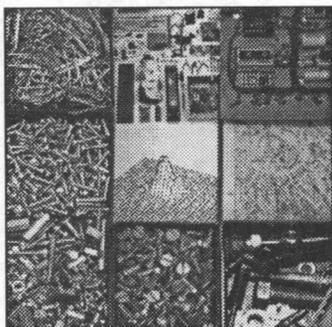


Bildnummer: 6

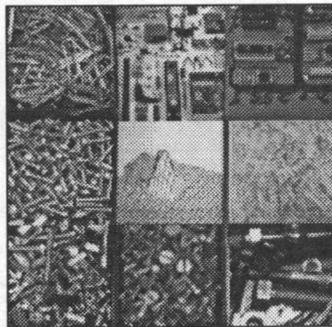


Bildnummer: 7

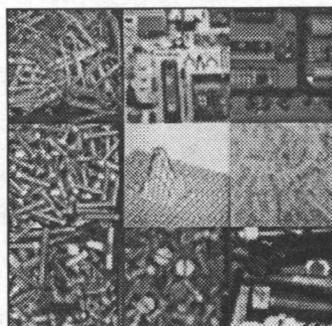
Sequenz **ZOOM** (ungerade Musternummern), Set 2 (**EDGES**):



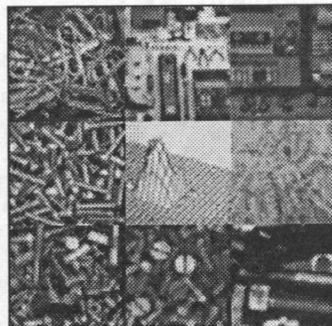
Bildnummer: 0



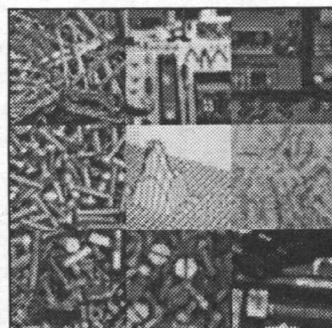
Bildnummer: 1



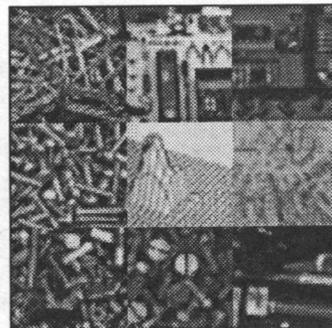
Bildnummer: 3



Bildnummer: 5

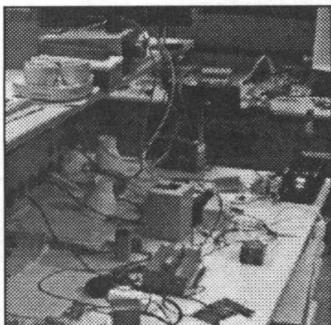


Bildnummer: 6

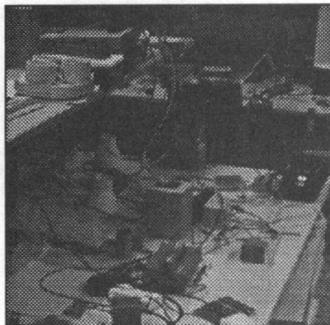


Bildnummer: 7

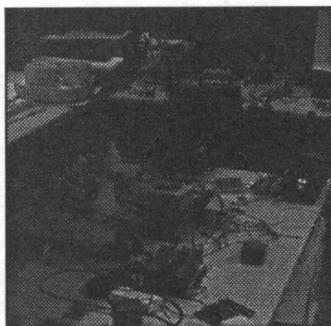
Sequenz IRIS, Set 3 (CHAOS):



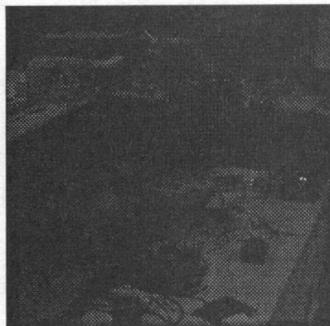
Bildnummer: 0



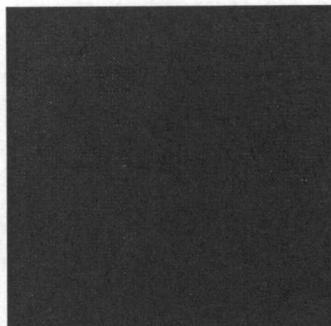
Bildnummer: 1



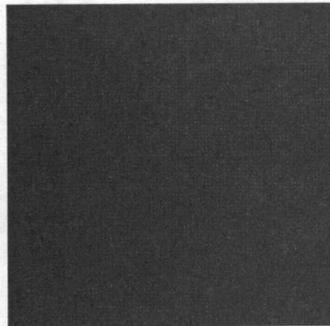
Bildnummer: 2



Bildnummer: 3

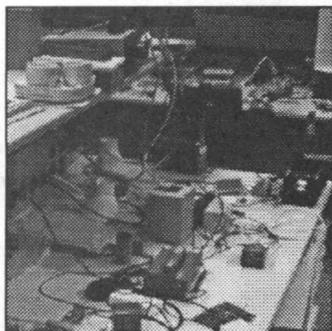


Bildnummer: 5

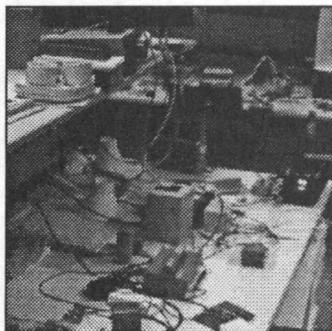


Bildnummer: 6

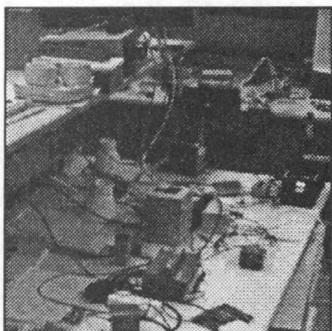
Sequenz **NOISE**, Set 3 (**CHAOS**):



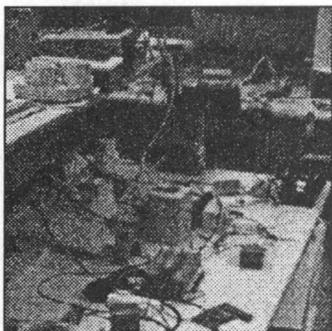
Bildnummer: 0



Bildnummer: 3



Bildnummer: 6



Bildnummer: 9

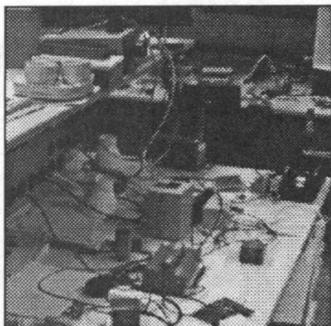


Bildnummer: 12

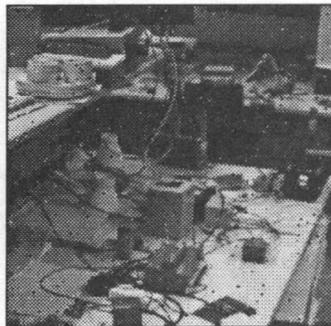


Bildnummer: 15

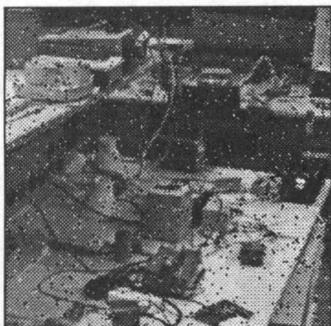
Sequenz **SALTY**, Set 3 (**CHAOS**):



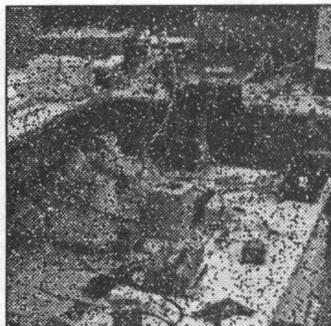
Bildnummer: 0



Bildnummer: 2



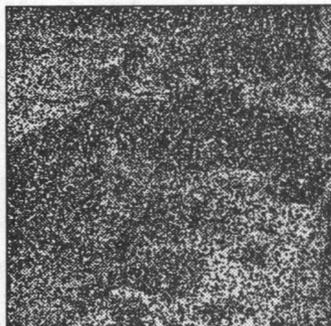
Bildnummer: 4



Bildnummer: 6

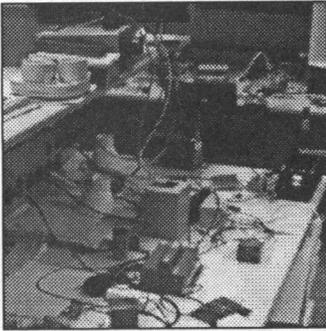


Bildnummer: 7

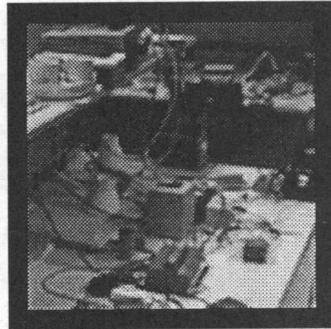


Bildnummer: 8

Sequenz **FOCUS**, Set 3 (**CHAOS**):



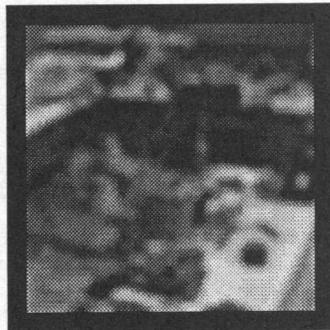
Bildnummer: 0



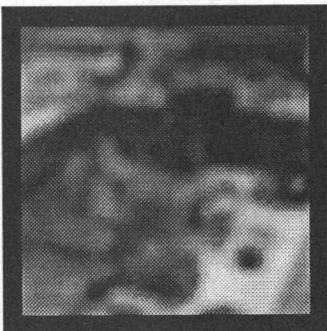
Bildnummer: 2



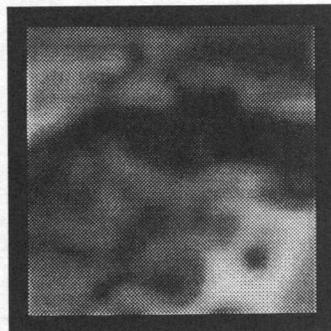
Bildnummer: 5



Bildnummer: 7

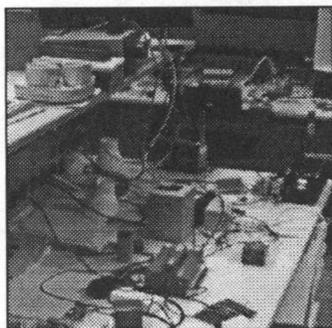


Bildnummer: 10

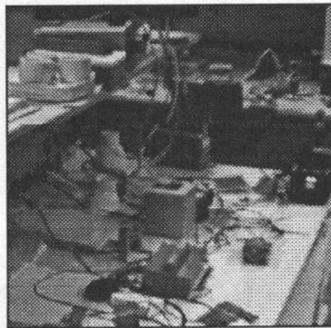


Bildnummer: 13

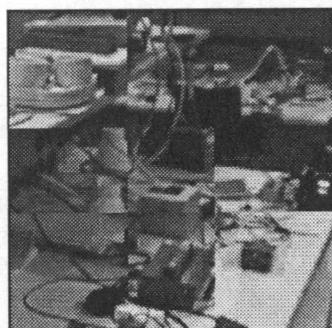
Sequenz ZOOM (gerade Musternummern), Set 3 (CHAOS):



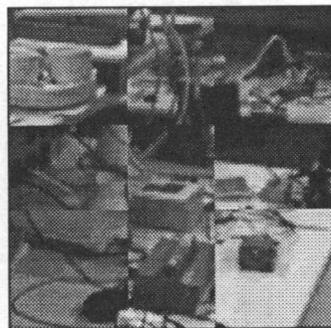
Bildnummer: 0



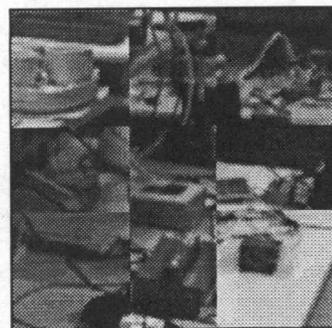
Bildnummer: 1



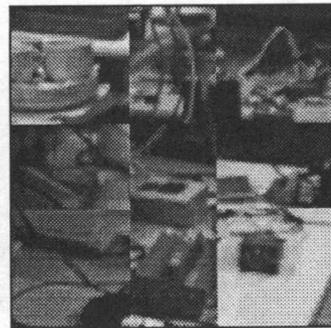
Bildnummer: 3



Bildnummer: 5

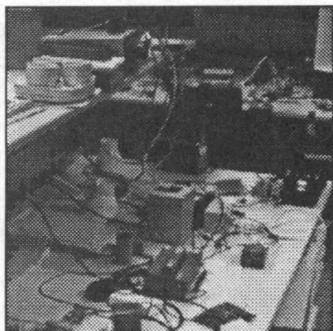


Bildnummer: 6

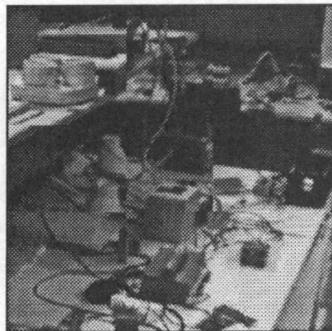


Bildnummer: 7

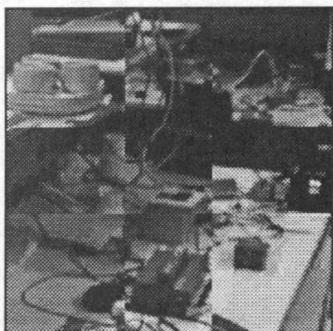
Sequenz **ZOOM** (ungerade Musternummern), Set 3 (**CHAOS**):



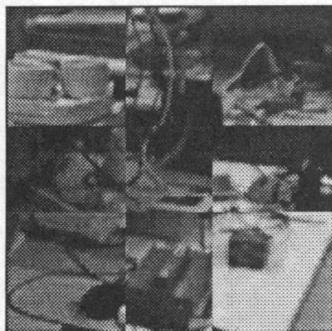
Bildnummer: 0



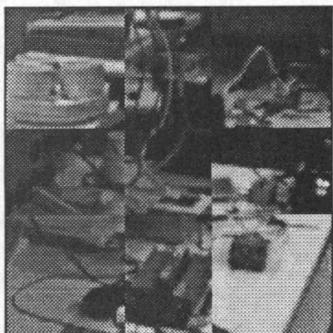
Bildnummer: 1



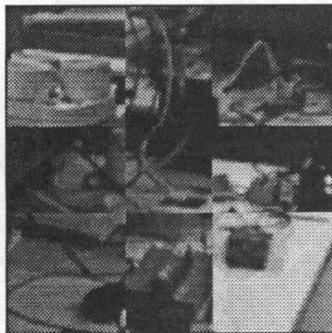
Bildnummer: 3



Bildnummer: 5



Bildnummer: 6



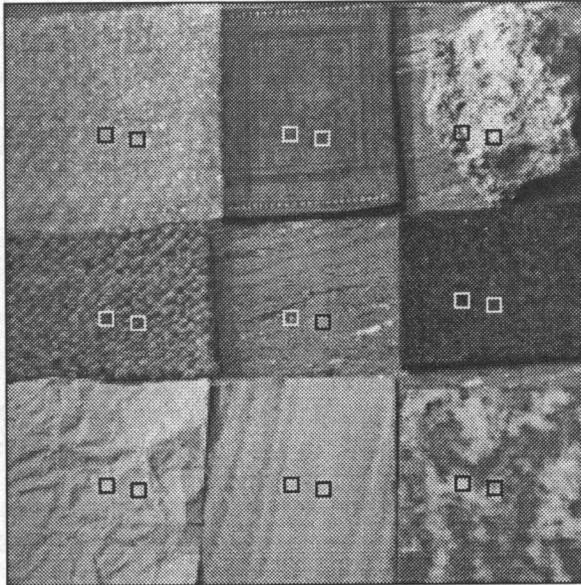
Bildnummer: 7

Anhang C: Objekt- und Suchfensterpositionen

In den folgenden Bildern sind die Objektpositionen – je nach lokaler Helligkeit – durch schwarze bzw. weisse Quadrate in den Startbildern (ungestört) der einzelnen Sequenzensets eingetragen. Die Grösse der Quadrate (7x7 Pixel) ist willkürlich gewählt. Die Muster wurden zentriert bezüglich diesen Quadraten gewonnen (Grösse 2x2 bis 25x25 Pixel, ausschliesslich quadratisch). Zur Wiedergabe der Bilder wurde eine zero order hold Interpolation verwendet. Trotz der relativ grossen Abmessungen der Bilder sind kleinste Bilddetails in Geometrie und Helligkeit wegen der Rasterung nicht mehr erkennbar. Aus ästhetischen Gründen wurde allen Bildern ein dünner, schwarzer Rand beigelegt.

Die zahlenmässigen Positionsangaben gehen von einem Koordinatensystem mit Ursprung in der linken oberen Ecke des Bildes aus. Es gilt zu beachten, dass die Objektpositionen für Bildsequenzenset 1 (**TEXTURES**) und 2 (**EDGES**) identisch gewählt wurden. Die zu den Objektfensterpositionen unsymmetrisch gewählten Suchfensterpositionen werden ebenfalls angegeben; sie sind aber nicht in den Bildern eingetragen. Es wird jeweils zuerst die Objektposition und danach, getrennt durch einen Schrägstrich, die zugehörige Suchfensterposition angegeben. Die Suchfenstergrösse lässt sich anhand der Mustergrösse berechnen: $xls=xlr+39$, $yls=yll+39$.

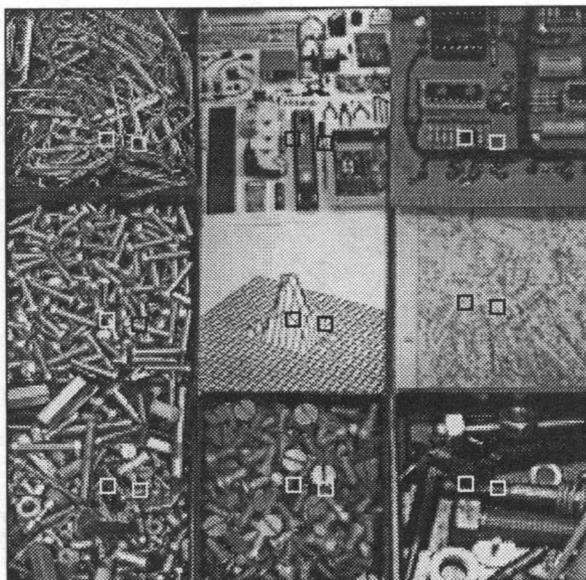
Bildsequenzenset 1 (TEXTURES):



Positionen:

- | | |
|------------------------|------------------------|
| 0: (44,58)/(52,52) | 9: (139,140)/(133,132) |
| 1: (58,60)/(52,52) | 10:(200,131)/(208,125) |
| 2: (125,58)/(133,52) | 11:(214,133)/(208,125) |
| 3: (139,60)/(133,52) | 12: (44,211)/(52,205) |
| 4: (200,58)/(208,52) | 13: (58,213)/(52,205) |
| 5: (214,60)/(208,52) | 14:(125,211)/(133,205) |
| 6: (44,138)/(52,132) | 15:(139,213)/(133,205) |
| 7: (58,140)/(52,132) | 16:(200,211)/(208,205) |
| 8: (125,138)/(133,132) | 17:(214,213)/(208,205) |

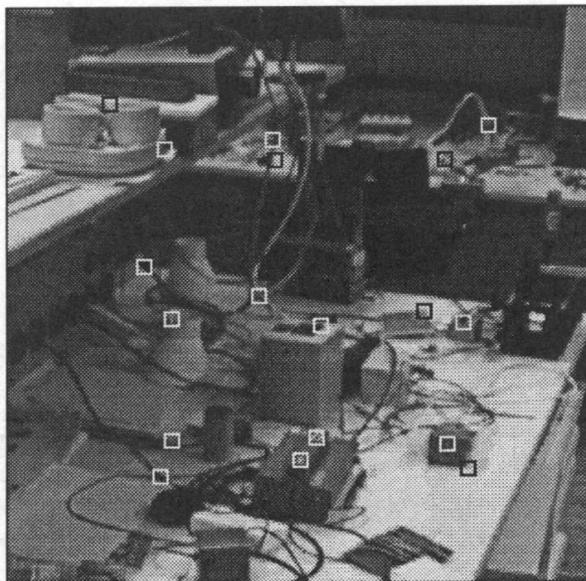
Bildsequenzenset 2 (EDGES):



Positionen:

20: (44,58)/(52,52)	29:(139,140)/(133,132)
21: (58,60)/(52,52)	30:(200,131)/(208,125)
22: (125,58)/(133,52)	31:(214,133)/(208,125)
23: (139,60)/(133,52)	32: (44,211)/(52,205)
24: (200,58)/(208,52)	33: (58,213)/(52,205)
25: (214,60)/(208,52)	34:(125,211)/(133,205)
26: (44,138)/(52,132)	35:(139,213)/(133,205)
27: (58,140)/(52,132)	36:(200,211)/(208,205)
28:(125,138)/(133,132)	37:(214,213)/(208,205)

Bildsequenzenset 3 (CHAOS):



Positionen:

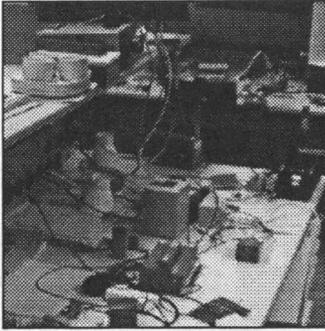
- | | |
|------------------------|------------------------|
| 40: (69,63)/(60,60) | 49:(110,128)/(127,127) |
| 41: (45,43)/(60,60) | 50:(182,135)/(194,127) |
| 42: (116,59)/(125,60) | 51:(199,140)/(194,127) |
| 43: (137,68)/(127,60) | 52: (72,192)/(60,194) |
| 44: (191,68)/(194,60) | 53: (67,208)/(60,194) |
| 45: (210,53)/(194,60) | 54:(128,201)/(127,194) |
| 46: (72,138)/(60,127) | 55:(135,191)/(127,194) |
| 47: (60,115)/(60,127) | 56:(192,194)/(194,194) |
| 48:(137,141)/(127,127) | 57:(201,205)/(194,194) |

Anhang D: Wirkung des Laplace-Gauss-Operators

Die unten aufgeführten Bilder zeigen die Wirkung des Laplace-Gauss-Operators inklusive Vorzeichenbinarisierung (Binarisierungsschwelle: 0). Der gefilterte und binarisierte Wert -1 wird schwarz der Wert +1 weiss dargestellt. Es wird deutlich, dass für grosse w -Werte der Tiefpasscharakter des Filters verstärkt wird, Details in den Bildern verschwinden und somit Muster, welche ursprünglich nur feine, lokale Merkmale enthielten, nicht mehr erfolgreich aufgefunden werden können.

Aus Platzgründen werden nur fünf Beispiele der ungestörten Szene des Sequenzensets 3, **CHAOS** gezeigt ($w = 2, 4, 6, 7$ und 8). Da die Implementation des Laplace-Gauss-Operators eine nicht verschwindende, vom Parameter w abhängige Maskengrösse aufweist, tritt am Bildrand eine undefinierte und hier schwarz markierte Zone auf.

Wirkung des Laplace-Gauss-Operators:



Original (Set 3, CHAOS, Nr.0)



$w=2$



$w=4$



$w=6$



$w=7$



$w=8$

Anhang E.1: Resultate, Grafiken

Auf den folgenden Seiten wird für jedes Korrelationskriterium das Fehlerhistogramm sowie alle Detektionsratenverläufe dargestellt. Die Detektionsratenverläufe wurden unter der Annahme eines Toleranzkreisradius' von $\sqrt{2}$ Pixel bestimmt. Die Bildstörungenstärke wurde – ausser bei der Bildstörung **ZOOM** – mit den Bildnummern (innerhalb der Bildsequenz) festgehalten und willkürlich auf den Bereich 0 bis 1 skaliert:

$$\text{Störungsstärke} = \frac{\text{Bildnummer}}{\text{MaxBildnummer}}$$

0 bezeichnet den ungestörten und 1 den Fall maximaler Störungsstärke. Der Zusammenhang dieser Grösse mit messbaren Störungsgrössen wird in Kapitel 4.2: *Generierung der Bildsequenzen* gegeben.

Bei der geometrischen Bildverzerrung (**ZOOM**) sind aus den vorne beschriebenen Gründen nur die Resultate der Muster mit ungeraden Kantenlängen berücksichtigt worden. Da der Zusammenhang zwischen Bildnummer und Störungsstärke in diesem Fall trivial ist, enthält die Störungsskala hier direkt die relative Expansion.

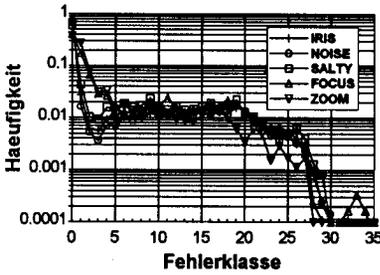
Beim Fehlerhistogramm wurde eine logarithmisch unterteilte Ordinate verwendet, um bei grösseren Abweichungen (Nummer der Fehlerklasse) die meist geringe Fehlerhäufigkeit besser beurteilen zu können. Im allgemeinen ist zu erkennen, dass sich die Fehlerhistogramme praktisch unabhängig von der Störungsart für jedes Korrelationskriterium sehr ähnlich verhalten. Wie weiter vorne bereits diskutiert, ist nur für die Störungsarten **FOCUS** und **ZOOM** regelmässig eine etwas breitere Spitze in der Nähe der Null-Abweichungsposition zu bemerken.

Die Nummer der Fehlerklasse multipliziert mit $\sqrt{2}$ gibt den äusseren Radius r_a des $\sqrt{2}$ Pixel breiten Toleranzrings an.

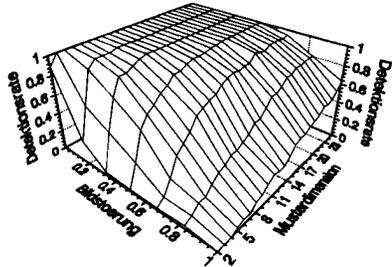
Die nun folgenden Grafiken halten die kumulierten Ergebnisse aller Sequenzensets fest (**TEXTURES, EDGES, CHAOS**).

Normierte Kreuzkorrelationsfunktion (KKF)

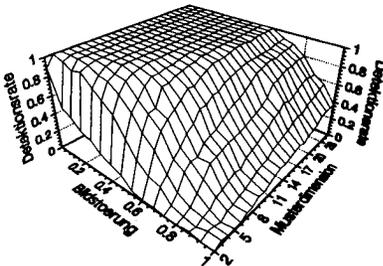
$$KKF(x, y) = \frac{\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} r(u, v) \cdot s(x+u, y+v)}{\sqrt{\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} r^2(u, v) \cdot \sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} s^2(x+u, y+v)}}$$



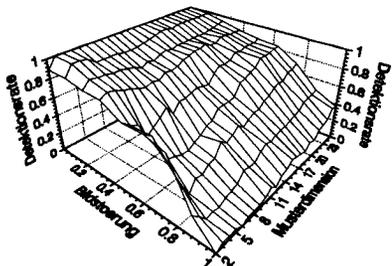
Fehlerhistogramm



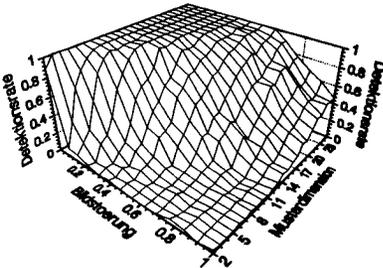
Störung: IRIS



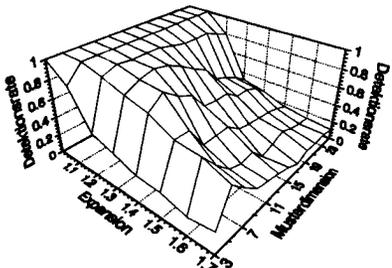
Störung: NOISE



Störung: SALTY



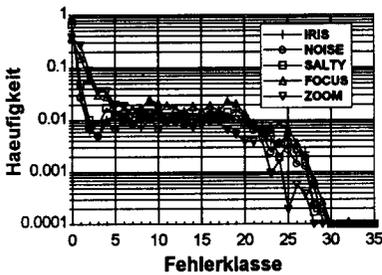
Störung: FOCUS



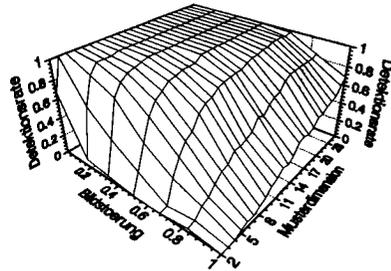
Störung: ZOOM

Lokal mittelwertfreie, normierte Kreuzkorrelationsfunktion (KKFMF)

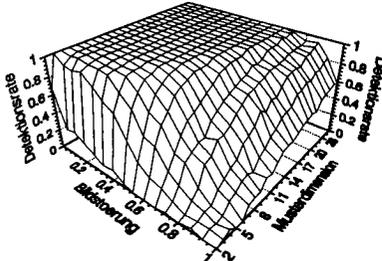
$$KKFMF(x, y) = \frac{\sum_{v=0}^{v_{\max}} \sum_{u=0}^{u_{\max}} \{r(u, v) - \bar{r}\} \cdot \{s(x+u, y+v) - \bar{s}(x, y)\}}{\sqrt{\sum_{v=0}^{v_{\max}} \sum_{u=0}^{u_{\max}} \{r(u, v) - \bar{r}\}^2 \cdot \sum_{v=0}^{v_{\max}} \sum_{u=0}^{u_{\max}} \{s(x+u, y+v) - \bar{s}(x, y)\}^2}}$$



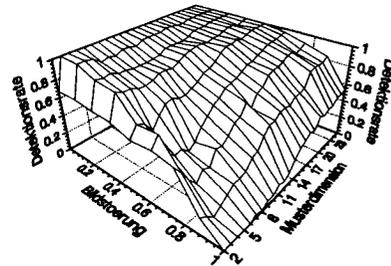
Fehlerhistogramm



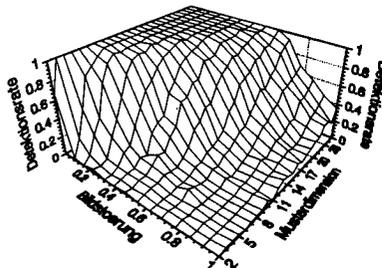
Störung: IRIS



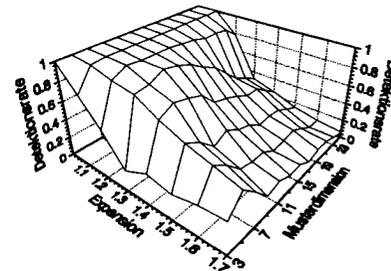
Störung: NOISE



Störung: SALTY



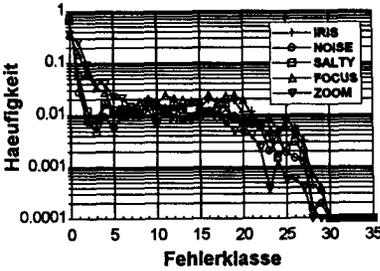
Störung: FOCUS



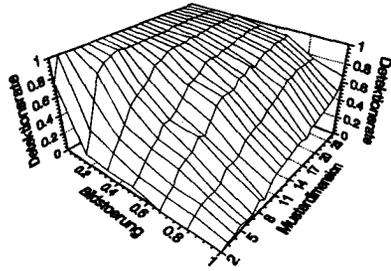
Störung: ZOOM

Korrelationskriterium nach Moravec (MOR)

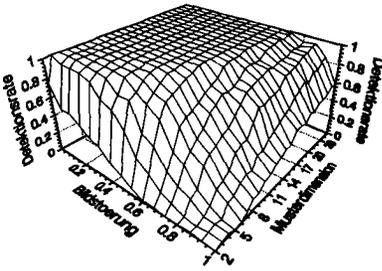
$$\text{MOR}(x, y) = \frac{2 \cdot \sum_{v=0}^{\text{MORMAX}} \sum_{u=0}^{\text{MORMAX}} \{r(u, v) - \bar{r}\} \cdot \{s(x+u, y+v) - \bar{s}(x, y)\}}{\sum_{v=0}^{\text{MORMAX}} \sum_{u=0}^{\text{MORMAX}} \{r(u, v) - \bar{r}\}^2 + \sum_{v=0}^{\text{MORMAX}} \sum_{u=0}^{\text{MORMAX}} \{s(x+u, y+v) - \bar{s}(x, y)\}^2}$$



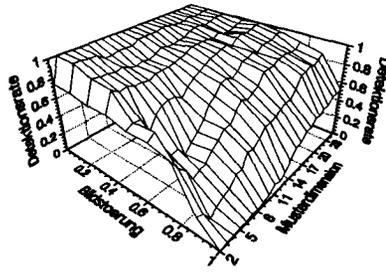
Fehlerhistogramm



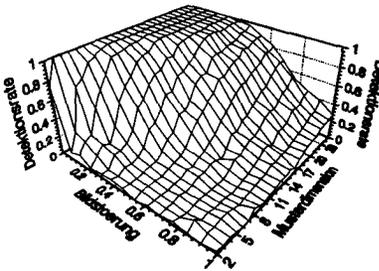
Störung: IRIS



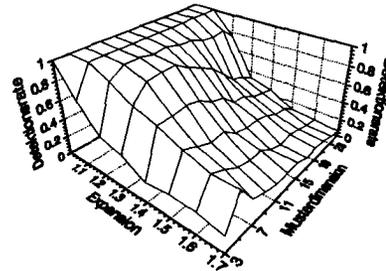
Störung: NOISE



Störung: SALTY



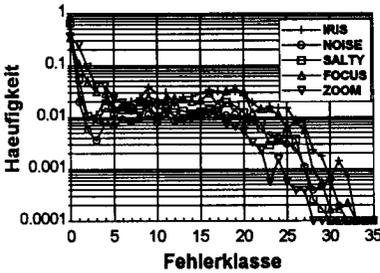
Störung: FOCUS



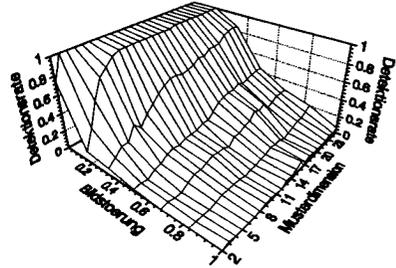
Störung: ZOOM

Normierte, lokal mittelwertfreie Summe der Differenzenquadrate
(NSSDMF)

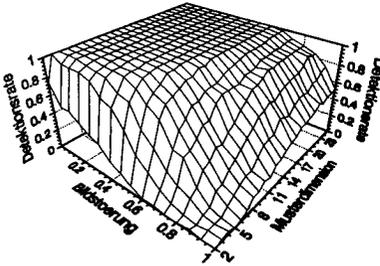
$$NSSDMF(x,y) = \frac{\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{r(u,v) - \bar{r}\} - \{s(x+u, y+v) - \bar{s}(x,y)\}}{\sqrt{\sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{r(u,v) - \bar{r}\}^2 \cdot \sum_{v=0}^{VMAX} \sum_{u=0}^{UMAX} \{s(x+u, y+v) - \bar{s}(x,y)\}^2}}$$



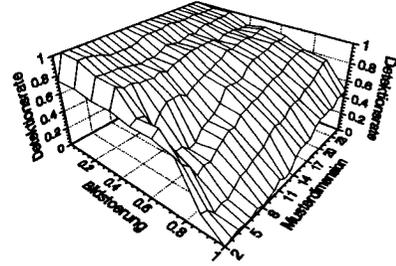
Fehlerhistogramm



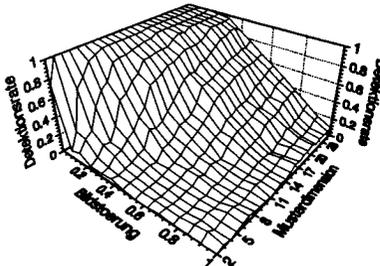
Störung: IRIS



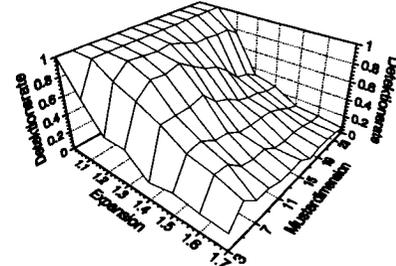
Störung: NOISE



Störung: SALTY



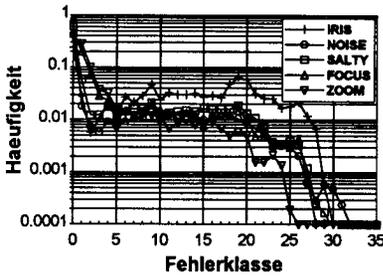
Störung: FOCUS



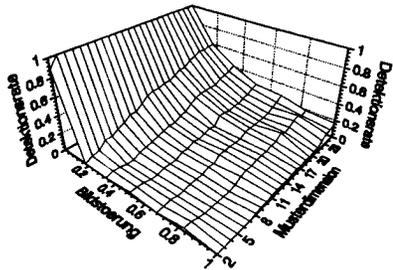
Störung: ZOOM

Summe der Differenzenquadrate (SSD)

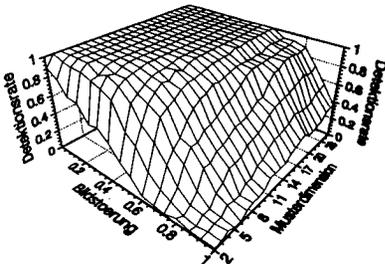
$$SSD(x, y) = \sum_{v=0}^{vmax} \sum_{u=0}^{umax} \{r(u, v) - s(x+u, y+v)\}^2$$



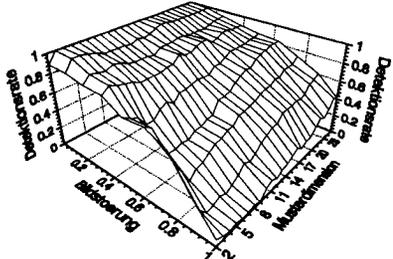
Fehlerhistogramm



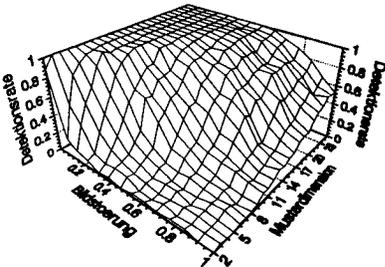
Störung: IRIS



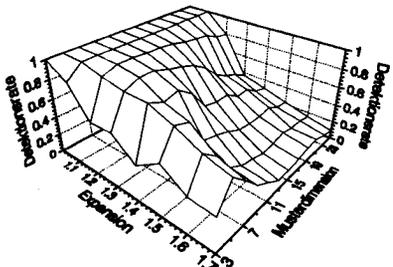
Störung: NOISE



Störung: SALTY



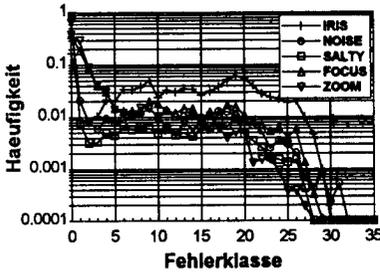
Störung: FOCUS



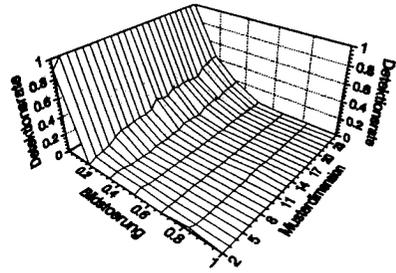
Störung: ZOOM

Summe der Differenzbeträge (SAD)

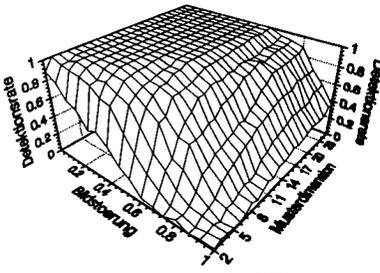
$$SAD(x, y) = \sum_{v=0}^{MIN(x, y)-1} \sum_{u=0}^{MIN(x, y)-1} |r(u, v) - s(x+u, y+v)|$$



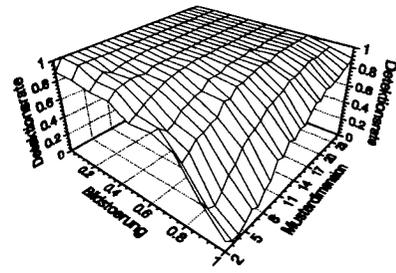
Fehlerhistogramm



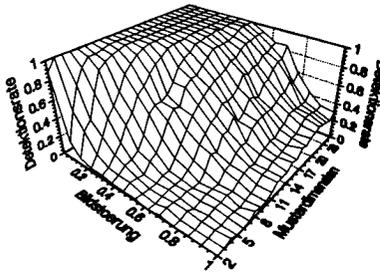
Störung: IRIS



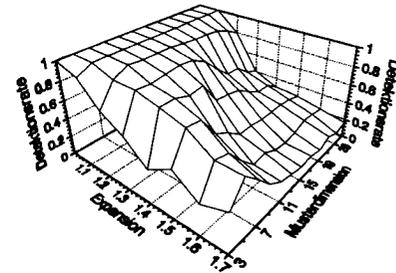
Störung: NOISE



Störung: SALTY



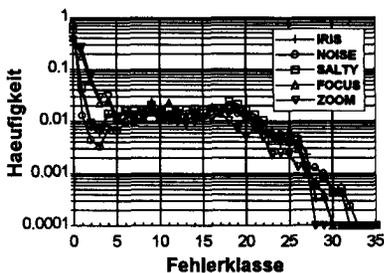
Störung: FOCUS



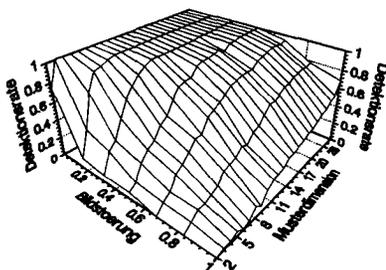
Störung: ZOOM

Lokal mittelwertfreie Summe der Differenzenquadrate (SSDMF)

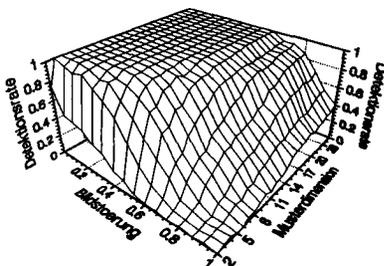
$$SSDMF(x, y) = \sum_{v=0}^{v_{max}} \sum_{u=0}^{u_{max}} \left\{ r(u, v) - \bar{r} \right\} - \left\{ s(x+u, y+v) - \bar{s}(x, y) \right\}^2$$



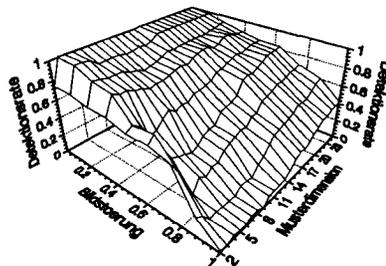
Fehlerhistogramm



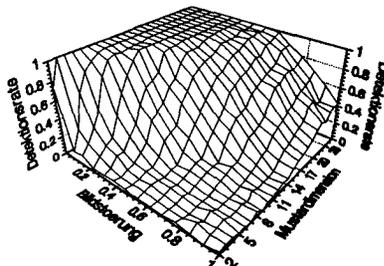
Störung: IRIS



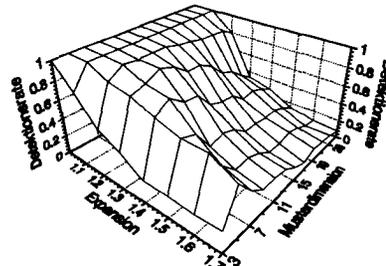
Störung: NOISE



Störung: SALTY



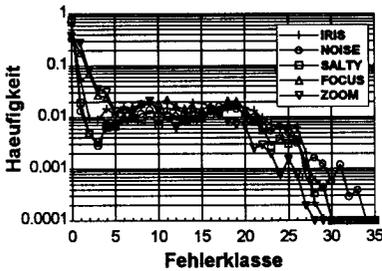
Störung: FOCUS



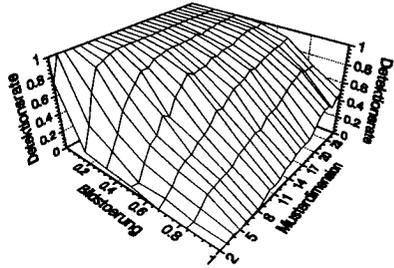
Störung: ZOOM

Lokal mittelwertfreie Summe der Differenzbeträge (SADMF)

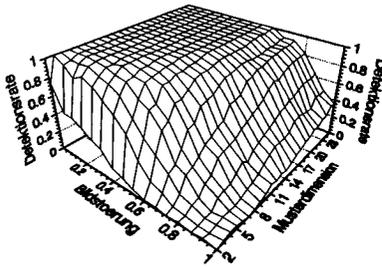
$$\text{SADMF}(x, y) = \sum_{v=0}^{\text{YMAX}} \sum_{u=0}^{\text{XMAX}} \left\{ |r(u, v) - \bar{r}| - \left| s(x+u, y+v) - \overline{s(x, y)} \right| \right\}$$



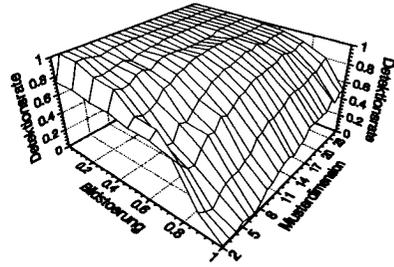
Fehlerhistogramm



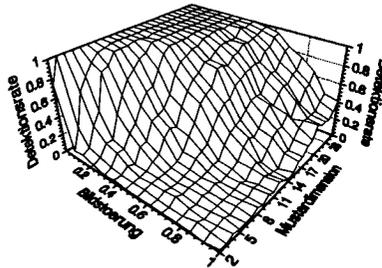
Störung: IRIS



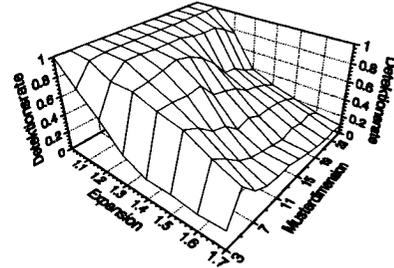
Störung: NOISE



Störung: SALTY



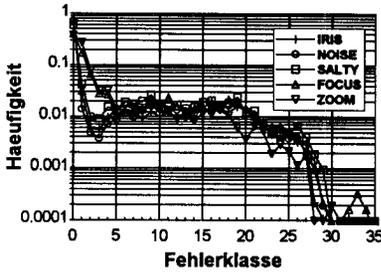
Störung: FOCUS



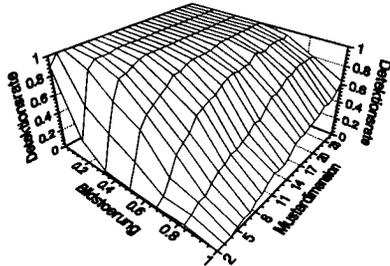
Störung: ZOOM

Lokal mittelwertskalierte Summe der Differenzenquadrate (SSDMS)

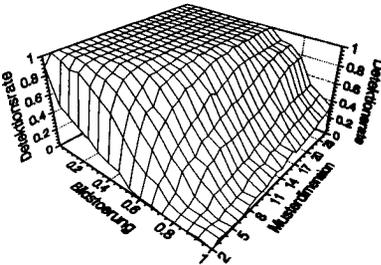
$$SSDMS(x, y) = \sum_{v=0}^{vmax} \sum_{u=0}^{umax} \left\{ r(u, v) - \frac{\bar{r}}{s(x, y)} \cdot s(x+u, y+v) \right\}^2$$



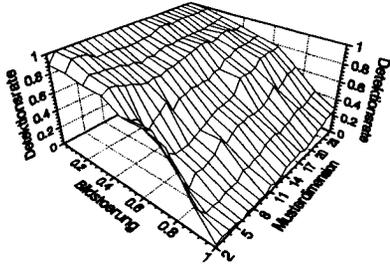
Fehlerhistogramm



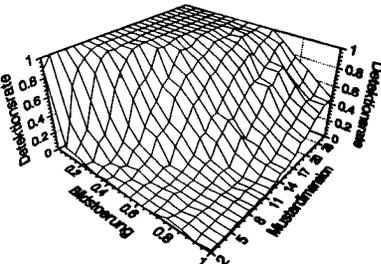
Störung: IRIS



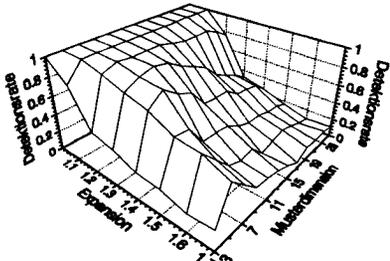
Störung: NOISE



Störung: SALTY



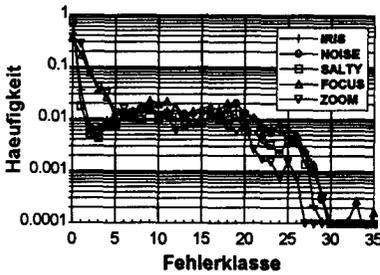
Störung: FOCUS



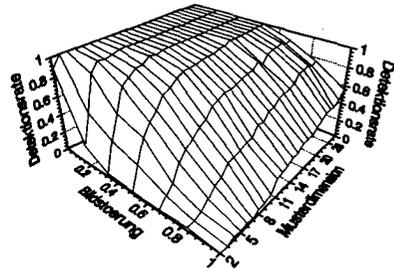
Störung: ZOOM

Lokal mittelwertskalierte Summe der Differenzenbeträge (**SADMS**)

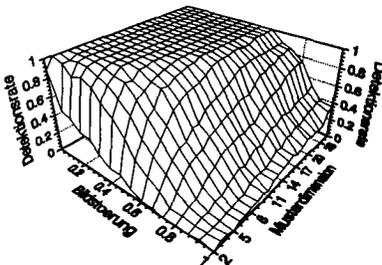
$$SADMS(x, y) = \sum_{v=0}^{M/2} \sum_{u=0}^{M/2} \left| r(u, v) - \frac{\bar{r}}{s(x, y)} s(x+u, y+v) \right|$$



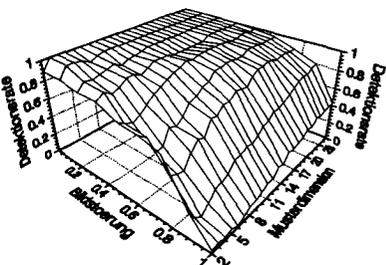
Fehlerhistogramm



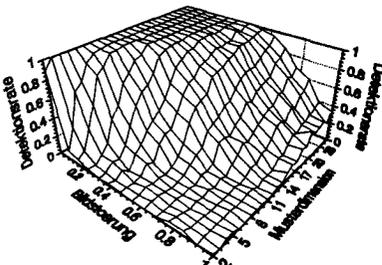
Störung: IRIS



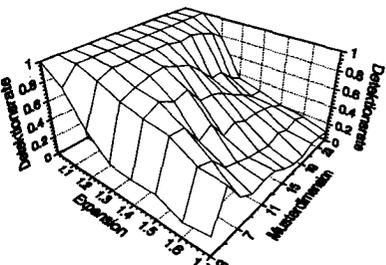
Störung: NOISE



Störung: SALTY



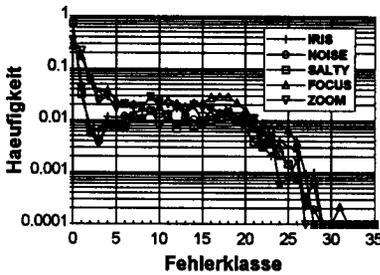
Störung: FOCUS



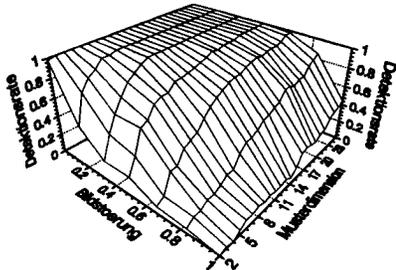
Störung: ZOOM

Korrelationskriterium nach Seitz (original, mit Sobeloperator; **SES1**)

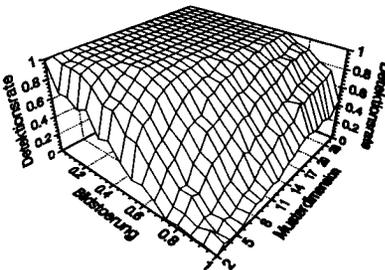
$$SES1(x, y) = \sqrt{\sum_{v=0}^{V_{max}} \sum_{u=0}^{U_{max}} \{r_{Sobel}(u, v) - s_{Sobel}(x+u, y+v)\}^2}$$



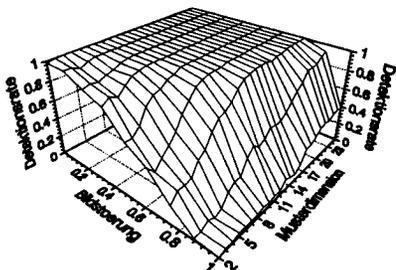
Fehlerhistogramm



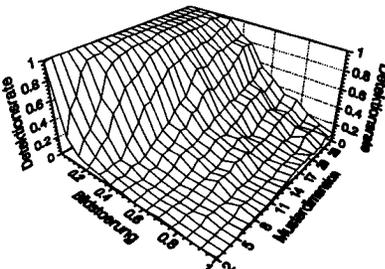
Störung: IRIS



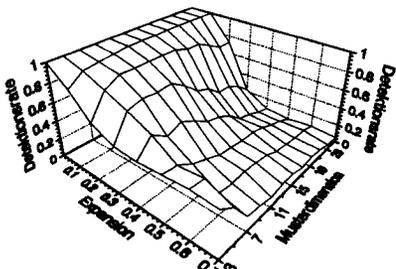
Störung: NOISE



Störung: SALTY



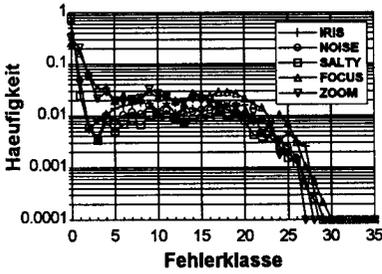
Störung: FOCUS



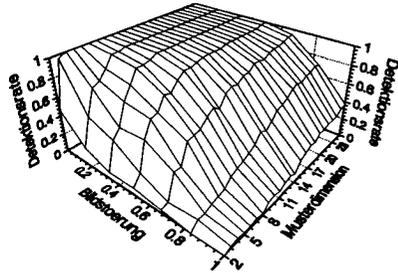
Störung: ZOOM

Korrelationskriterium nach Seitz (mit Kirschoperator; **SEK1**)

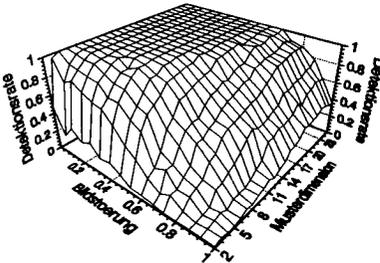
$$SEK1(x, y) = \sqrt{\sum_{v=0}^{vmax} \sum_{u=0}^{umax} \{r_{Kirsch}(u, v) - s_{Kirsch}(x+u, y+v)\}^2}$$



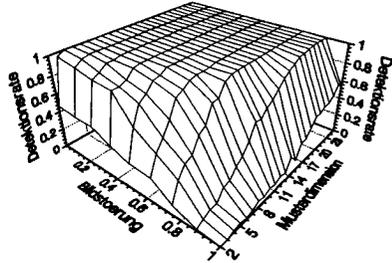
Fehlerhistogramm



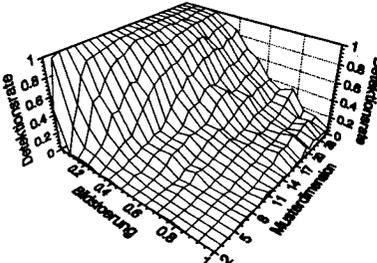
Störung: IRIS



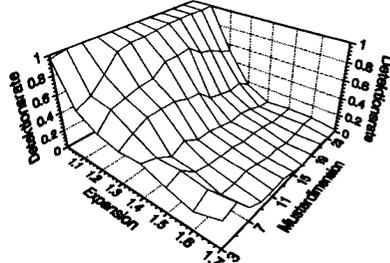
Störung: NOISE



Störung: SALTY



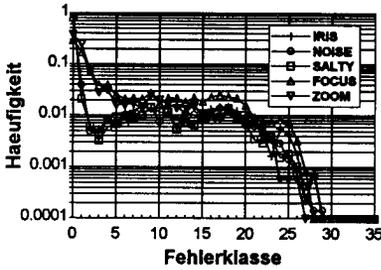
Störung: FOCUS



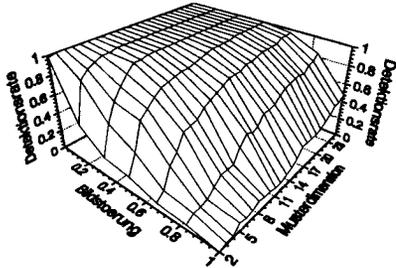
Störung: ZOOM

Korrelationskriterium nach Seitz (mit Sobeloperator; **SES2**)

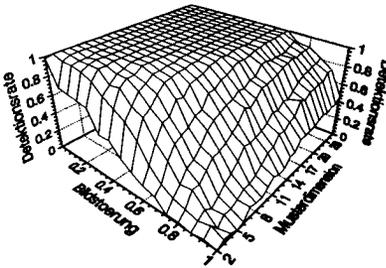
$$SES2(x, y) = \sum_{v=0}^{vmax} \sum_{u=0}^{umax} |r_{Sobel}(u, v) - s_{Sobel}(x+u, y+v)|$$



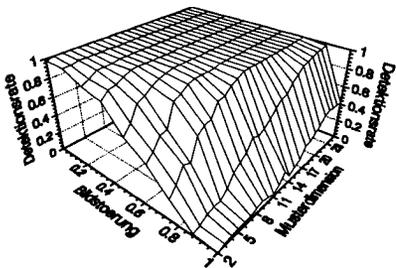
Fehlerhistogramm



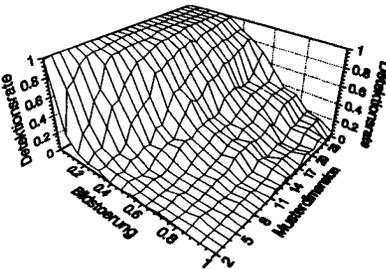
Störung: IRIS



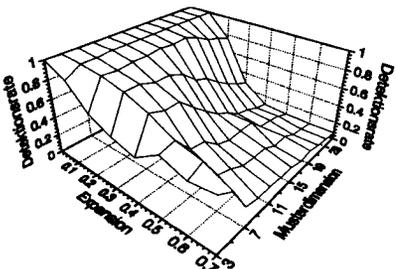
Störung: NOISE



Störung: SALTY



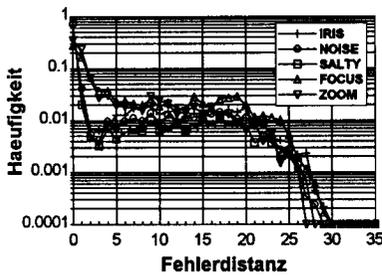
Störung: FOCUS



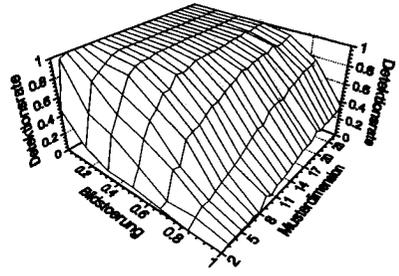
Störung: ZOOM

Korrelationskriterium nach Seitz (mit Kirschoperator; **SEK2)**

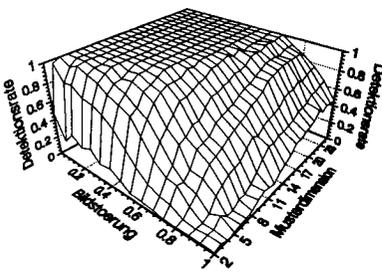
$$SEK2(x, y) = \sum_{v=0}^{v_{max}} \sum_{u=0}^{u_{max}} |r_{Kirsch}(u, v) - s_{Kirsch}(x+u, y+v)|$$



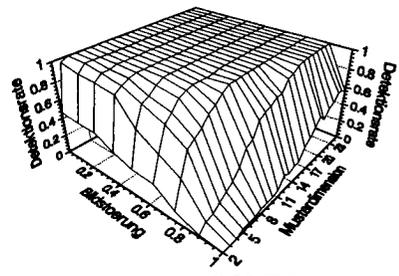
Fehlerhistogramm



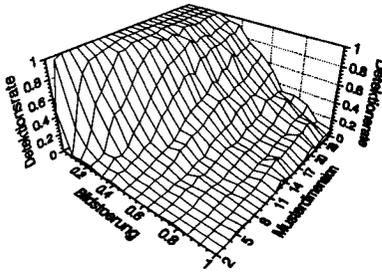
Störung: IRIS



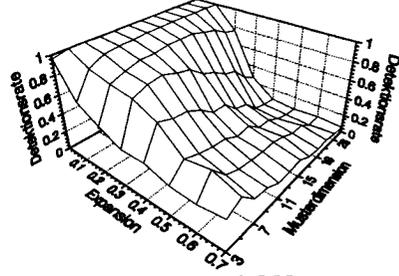
Störung: NOISE



Störung: SALTY



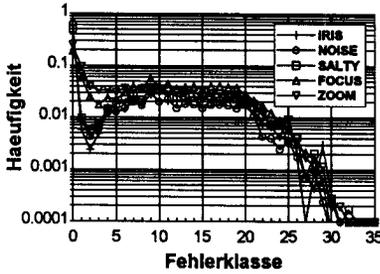
Störung: FOCUS



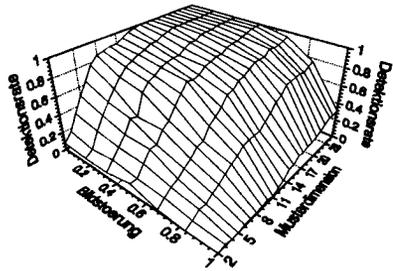
Störung: ZOOM

Unnormierte Kreuzkorrelationsfunktion (UKKF)

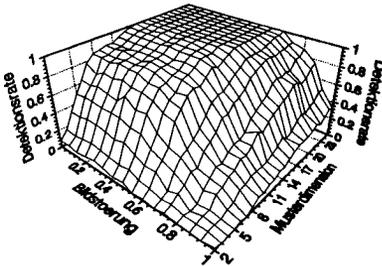
$$UKKF(x, y) = \sum_{v=0}^{v_{max}} \sum_{u=0}^{u_{max}} r'(u, v) \cdot s'(x+u, y+v)$$



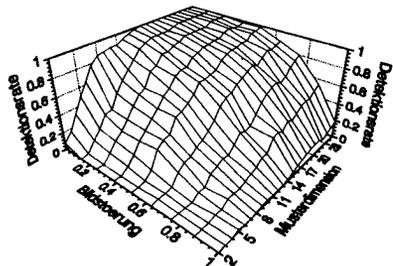
Fehlerhistogramm



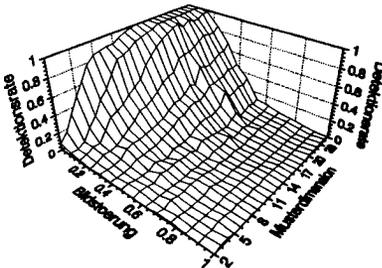
Störung: IRIS



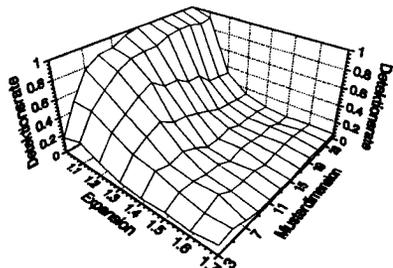
Störung: NOISE



Störung: SALTY



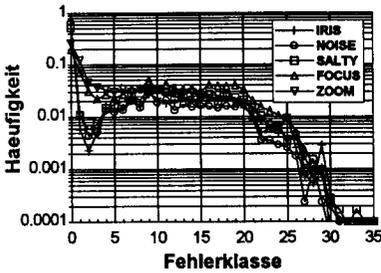
Störung: FOCUS



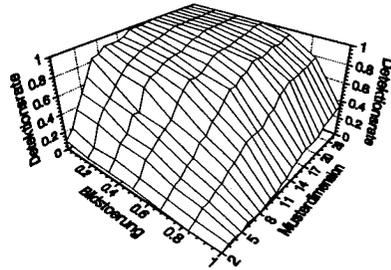
Störung: ZOOM

Lokal mittelwertfreie, unnormierte Kreuzkorrelationsfunktion (UKKFMF)

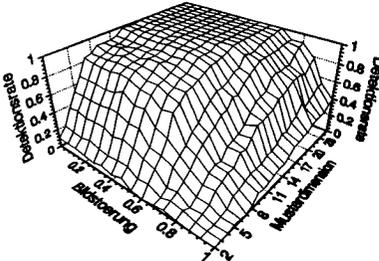
$$UKKFMF(x, y) = \sum_{v=0}^{vmax} \sum_{u=0}^{umax} \{r'(u, v) - \bar{r}'\} \cdot \{s'(x+u, y+v) - \bar{s}'(x, y)\}$$



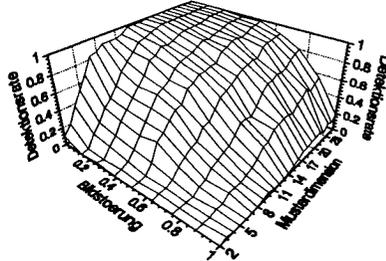
Fehlerhistogramm



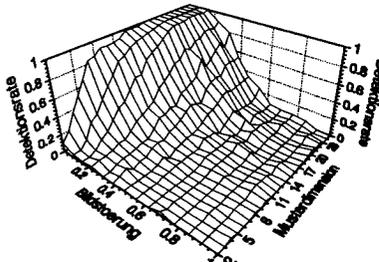
Störung: IRIS



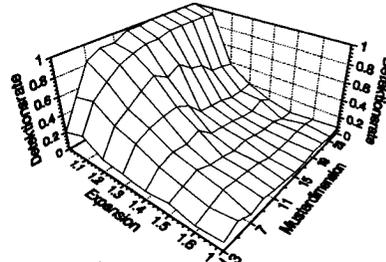
Störung: NOISE



Störung: SALTY



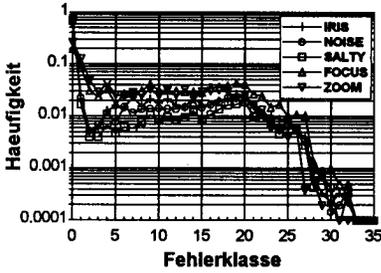
Störung: FOCUS



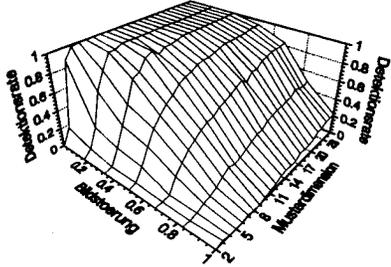
Störung: ZOOM

Korrelationskriterium nach Nishihara ($w=2$; NIS)

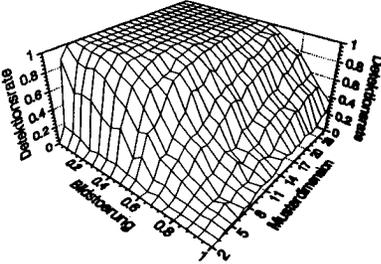
$$\text{NIS}(x, y) = \sum_{v=0}^{\text{MIN}(M, N)} \sum_{u=0}^{\text{MIN}(M, N)-v} r_{b_1}(u, v) \cdot s_{b_1}(x+u, y+v)$$



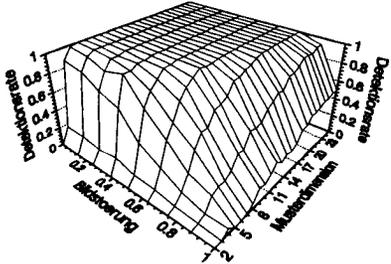
Fehlerhistogramm



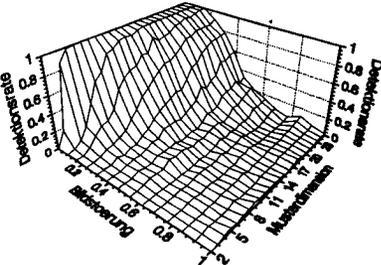
Störung: IRIS



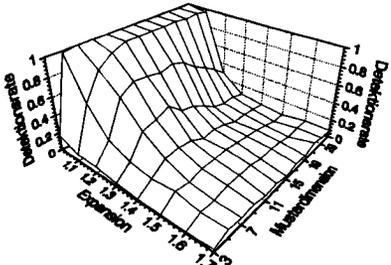
Störung: NOISE



Störung: SALTY



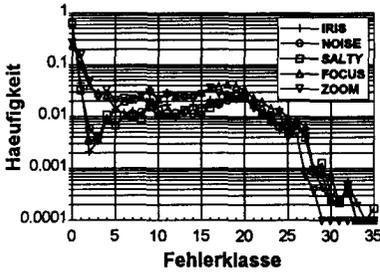
Störung: FOCUS



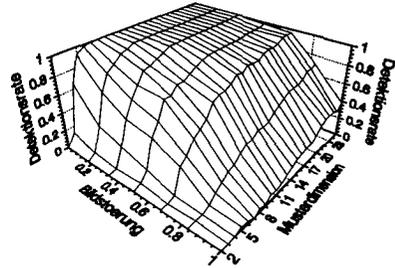
Störung: ZOOM

Korrelationskriterium nach Nishihara ($w=3$; NIS)

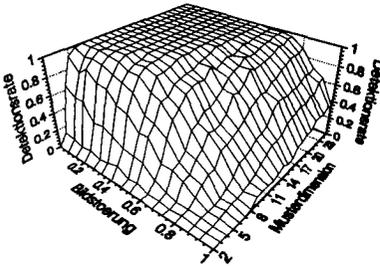
$$NIS(x, y) = \sum_{u=0}^{MIN(x, y)} \sum_{v=0}^{MIN(x, y)} r_{21}(u, v) \cdot s_{21}(x+u, y+v)$$



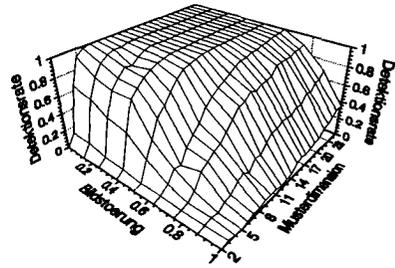
Fehlerhistogramm



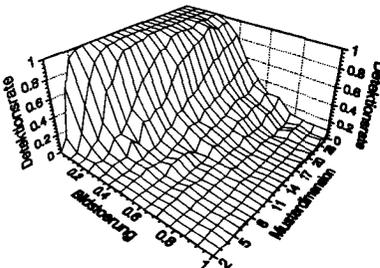
Störung: IRIS



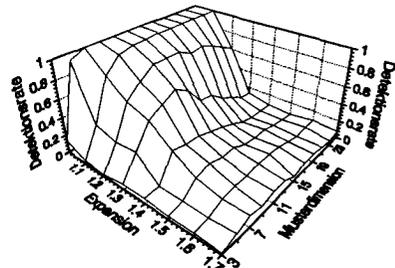
Störung: NOISE



Störung: SALTY



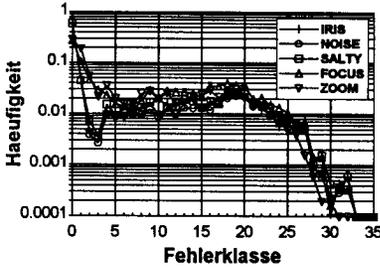
Störung: FOCUS



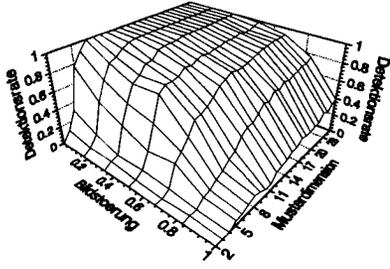
Störung: ZOOM

Korrelationskriterium nach Nishihara ($w=4$; NIS)

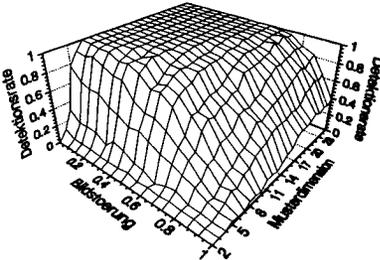
$$NIS(x, y) = \sum_{v=0}^{vmax} \sum_{u=0}^{umax} r_{b1}(u, v) \cdot s_{b1}(x+u, y+v)$$



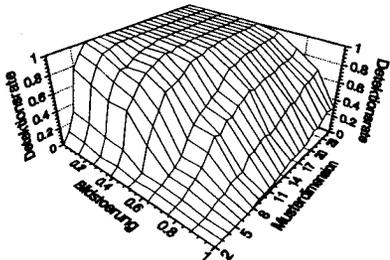
Fehlerhistogramm



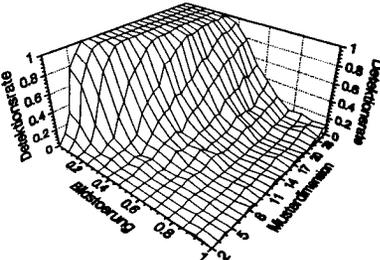
Störung: IRIS



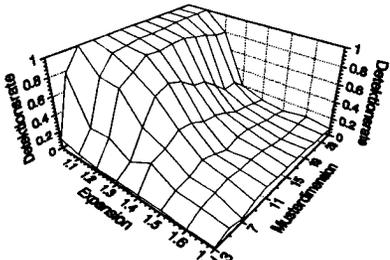
Störung: NOISE



Störung: SALTY



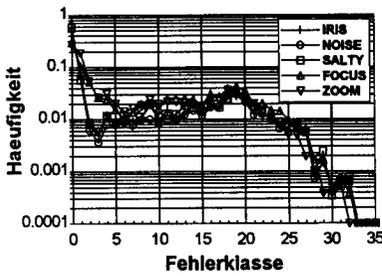
Störung: FOCUS



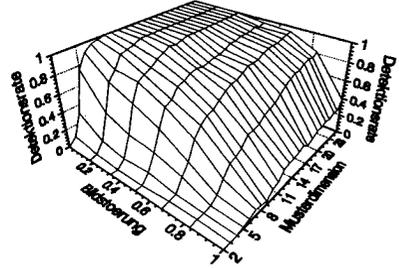
Störung: ZOOM

Korrelationskriterium nach Nishihara ($w=5$; NIS)

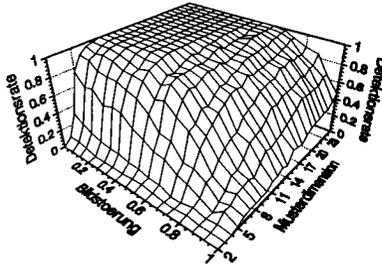
$$\text{NIS}(x, y) = \sum_{v=0}^{\text{MIN}(x, y)} \sum_{u=0}^{\text{MIN}(x, y)-v} r_{b_1}(u, v) \cdot s_{b_1}(x+u, y+v)$$



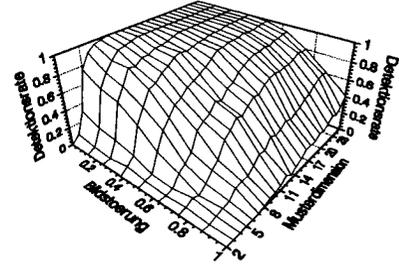
Fehlerhistogramm



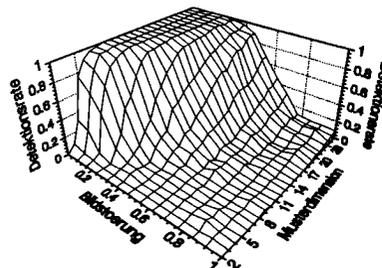
Störung: IRIS



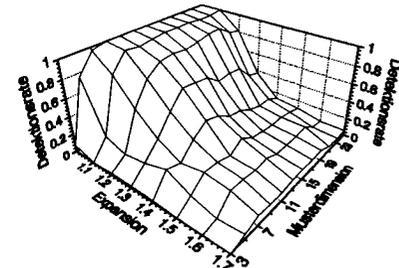
Störung: NOISE



Störung: SALTY



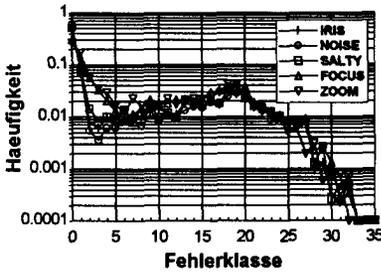
Störung: FOCUS



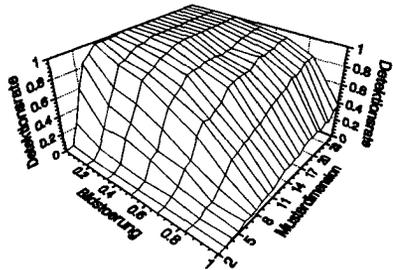
Störung: ZOOM

Korrelationskriterium nach Nishihara ($w=6$; NIS)

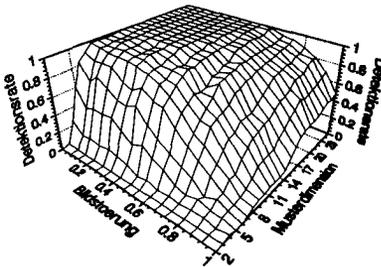
$$\text{NIS}(x, y) = \sum_{v=0}^{\text{MIN}(x, y)} \sum_{u=0}^{\text{MIN}(x, y)} r_{b1}(u, v) \cdot s_{b1}(x+u, y+v)$$



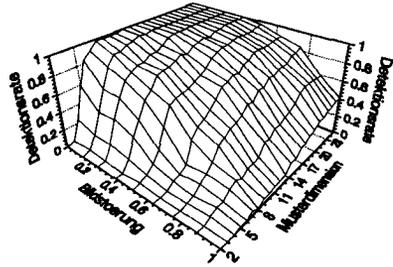
Fehlerhistogramm



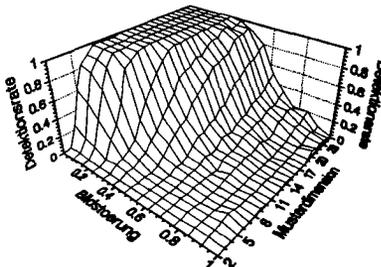
Störung: IRIS



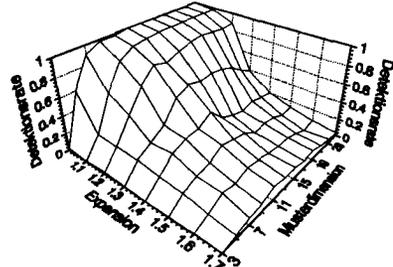
Störung: NOISE



Störung: SALTY



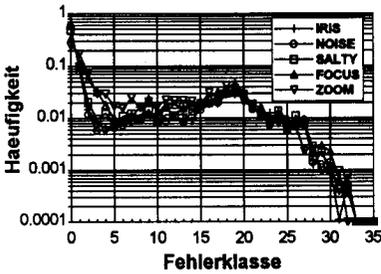
Störung: FOCUS



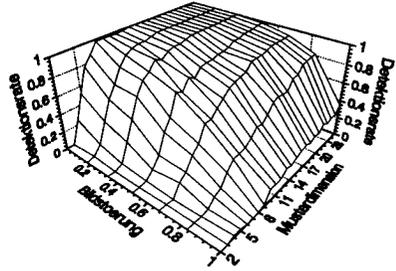
Störung: ZOOM

Korrelationskriterium nach Nishihara ($w=7$; NIS)

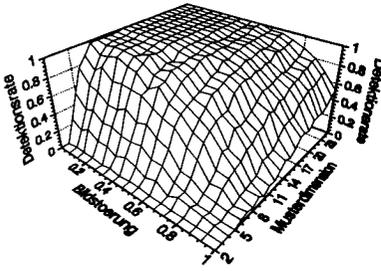
$$NIS(x, y) = \sum_{u=0}^{w/2} \sum_{v=0}^{w/2} r_{bl}(u, v) \cdot s_{bl}(x+u, y+v)$$



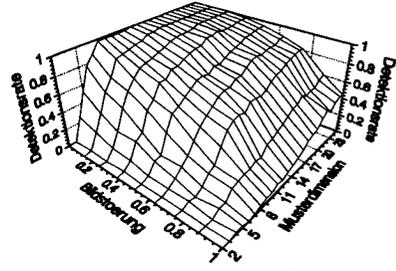
Fehlerhistogramm



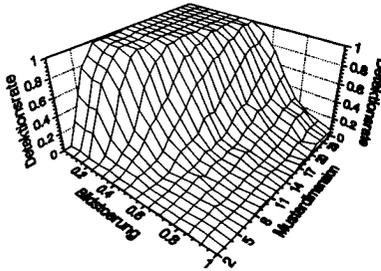
Störung: IRIS



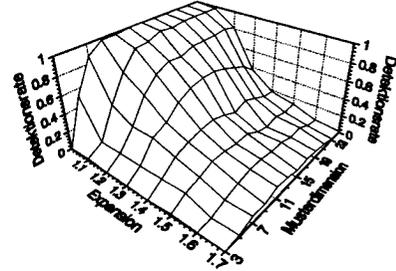
Störung: NOISE



Störung: SALTY



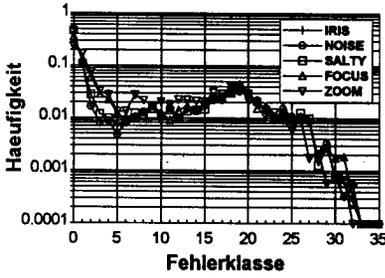
Störung: FOCUS



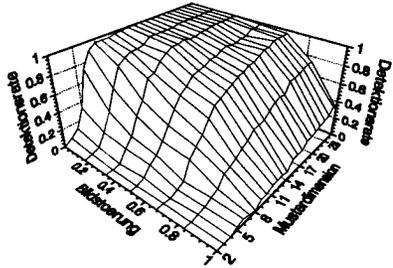
Störung: ZOOM

Korrelationskriterium nach Nishihara ($w=8$; NIS)

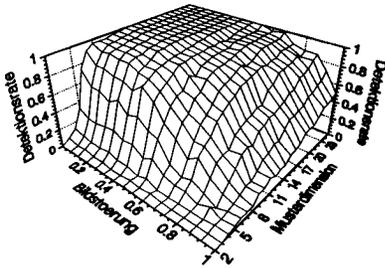
$$\text{NIS}(x, y) = \sum_{v=0}^{\text{MIN}(x, y)} \sum_{u=0}^{\text{MIN}(x, y)} r_{b1}(u, v) \cdot s_{b1}(x+u, y+v)$$



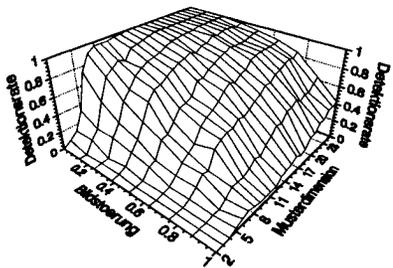
Fehlerhistogramm



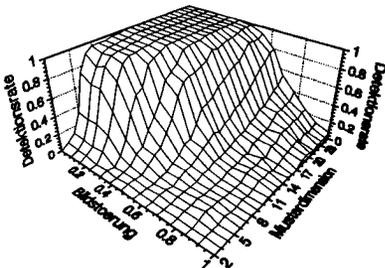
Störung: IRIS



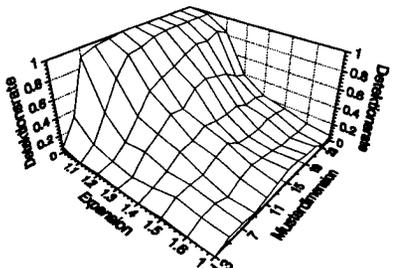
Störung: NOISE



Störung: SALTY



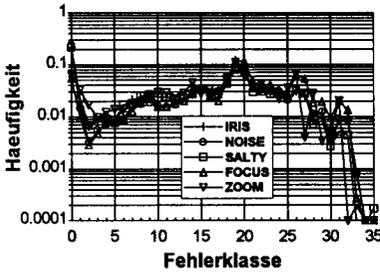
Störung: FOCUS



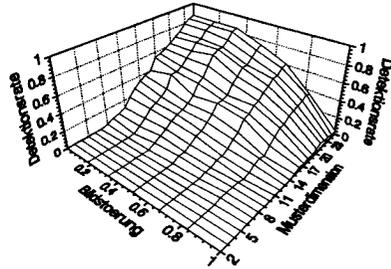
Störung: ZOOM

Korrelationskriterium nach Nack (original, NA1)

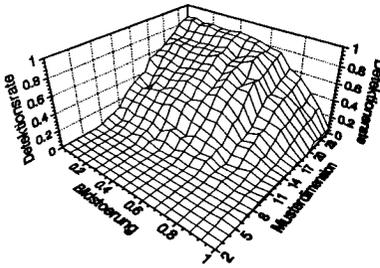
$$NA1(x, y) = \frac{1 + \sum_{v=0}^{vmax} \sum_{u=0}^{umax} r_{b2}(u, v) \cdot s_{b2}(x+u, y+v)}{1 + \sum_{v=0}^{vmax} \sum_{u=0}^{umax} s_{b2}(x+u, y+v)}$$



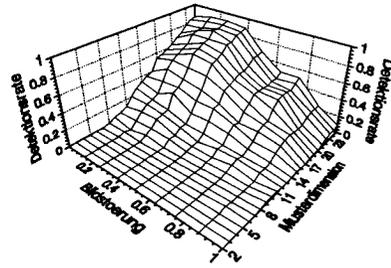
Fehlerhistogramm



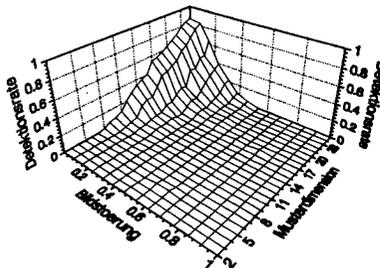
Störung: IRIS



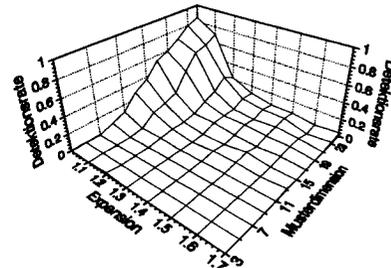
Störung: NOISE



Störung: SALTY



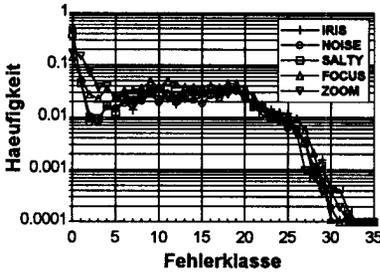
Störung: FOCUS



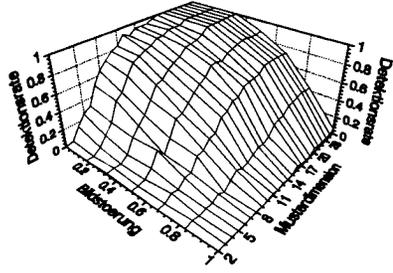
Störung: ZOOM

Korrelationskriterium nach Nack (modifiziert, NA2)

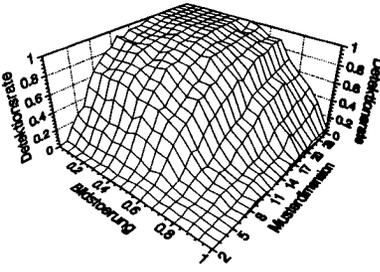
$$NA2(x,y) = \frac{1 + \sum_{v=0}^{VIRAGE} \sum_{u=0}^{KIRAGE} r_{b2}(u,v) \cdot s_{b2}(x+u,y+v)}{\left\{ 1 + \sum_{v=0}^{VIRAGE} \sum_{u=0}^{KIRAGE} s_{b2}(x+u,y+v) \right\} \left\{ 1 + \sum_{v=0}^{VIRAGE} \sum_{u=0}^{KIRAGE} r_{b2}(u,v) - \sum_{v=0}^{VIRAGE} \sum_{u=0}^{KIRAGE} r_{b2}(u,v) \cdot s_{b2}(x+u,y+v) \right\}}$$



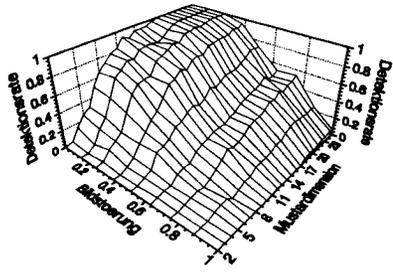
Fehlerhistogramm



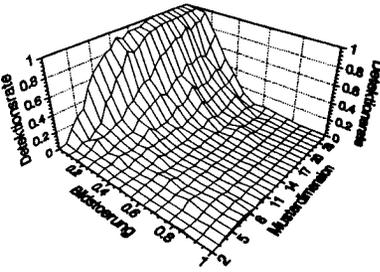
Störung: IRIS



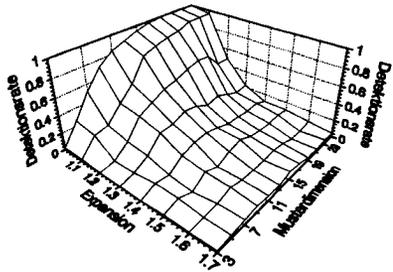
Störung: NOISE



Störung: SALTY



Störung: FOCUS



Störung: ZOOM

Anhang E.2: Resultate, Tabellen

Im folgenden werden die zahlenmässigen Resultate der Korrelationskriterien für die einzelnen Sequenzensets getrennt gegeben.

Störung Alg.	IRIS	NOISE	SALTY	FOCUS	ZOOM
KKF	0.668	0.651	0.514	0.550	0.626
KKFMF	0.689	0.732	0.653	0.505	0.616
MOR	0.658	0.734	0.640	0.468	0.592
NSSDMF	0.430	0.708	0.580	0.384	0.579
SSD	0.158	0.690	0.541	0.605	0.635
SAD	0.143	0.679	0.814	0.563	0.646
SSDMF	0.674	0.671	0.512	0.545	0.627
SADMF	0.621	0.659	0.722	0.524	0.624
SSDMS	0.669	0.642	0.505	0.550	0.625
SADMS	0.640	0.636	0.727	0.524	0.623
SES1	0.657	0.709	0.762	0.410	0.520
SES2	0.666	0.713	0.796	0.422	0.564
SEK1	0.628	0.676	0.809	0.408	0.499
SEK2	0.641	0.687	0.824	0.418	0.533
UKKF	0.600	0.637	0.448	0.216	0.339
UKKFMF	0.609	0.643	0.458	0.224	0.366
NIS (w=2)	0.526	0.586	0.725	0.238	0.323
NIS (w=3)	0.570	0.616	0.624	0.310	0.389
NIS (w=4)	0.569	0.623	0.572	0.341	0.423
NIS (w=5)	0.538	0.598	0.527	0.360	0.410
NIS (w=6)	0.514	0.582	0.500	0.387	0.402
NIS (w=7)	0.514	0.574	0.485	0.416	0.387
NIS (w=8)	0.489	0.560	0.470	0.440	0.373
NA1	0.277	0.317	0.233	0.090	0.149
NA2	0.418	0.475	0.362	0.173	0.306

Tab.11: Übersicht über die mittleren Detektionsraten für Set 1 (TEXTURES) bei einem Toleranzkreisradius von $\sqrt{2}$ Pixeln.

Störung Alg.	80%G	IRIS 80%S	NOISE 80%S	SALTY 80%S	FOCUS 80%S	ZOOM		
						OG	MS	80%S
KKF	2.1	0.70	0.66	0.28	0.73	7	51.5	17.2
KKFMF	2.7	0.75	0.79	0.73	0.72	7	44.0	17.2
MOR	2.6	0.73	0.79	0.73	0.64	7	38.7	16.0
NSSDMF	2.6	0.38	0.72	0.32	0.41	7	38.7	16.0
SSD	2.1	0.04	0.66	0.28	0.86	7	53.7	19.0
SAD	2.1	0.03	0.71	0.90	0.72	7	46.0	19.0
SSDMF	2.6	0.75	0.66	0.23	0.73	7	51.2	19.0
SADMF	2.6	0.69	0.66	0.69	0.70	11	41.2	17.2
SSDMS	2.1	0.70	0.64	0.28	0.73	7	51.5	17.2
SADMS	2.1	0.65	0.64	0.68	0.72	7	41.5	17.2
SES1	2.0	0.81	0.75	0.86	0.45	11	32.0	16.0
SES2	2.0	0.80	0.73	0.90	0.48	9	41.2	15.1
SEK1	2.6	0.69	0.72	0.91	0.43	9	32.7	15.1
SEK2	2.6	0.70	0.73	0.91	0.44	7	36.5	16.0
UKKF	6.3	0.70	0.72	0.61	0.19	9	21.2	12.0
UKKFMF	6.1	0.70	0.72	0.61	0.21	9	21.5	12.2
NIS (w=2)	3.9	0.53	0.68	0.86	0.22	9	17.2	12.3
NIS (w=3)	4.4	0.70	0.76	0.74	0.33	11	28.0	14.0
NIS (w=4)	4.9	0.64	0.77	0.68	0.42	11	28.7	20.5
NIS (w=5)	5.7	0.69	0.72	0.61	0.45	15	26.0	19.0
NIS (w=6)	6.4	0.63	0.71	0.57	0.45	15	26.0	20.9
NIS (w=7)	7.1	0.64	0.72	0.54	0.56	15	25.3	20.7
NIS (w=8)	6.6	0.70	0.70	0.58	0.58	21	23.7	21.6
NA1	16.6	0.43	0.55	0.36	0.16	21	12.4	12.2
NA2	8.4	0.43	0.56	0.39	0.21	19	15.2	13.6

Tab.12: Übersicht über die 80%-Mustergrößen (80%G), die 80%-Bildstörungsstärken (80%S), die optimalen Mustergrößen (OG) und die Maximumstörungsstärken (MS) für Set 1 (TEXTURES) bei einem Toleranzkreisradius von $\sqrt{2}$ Pixeln. Sowohl MS als auch 80%S sind für den Störungstyp ZOOM in Prozenten angegeben (Größenveränderung, relative Expansion).

Störung Alg.	IRIS	NOISE	SALTY	FOCUS	ZOOM
KKF	0.803	0.821	0.747	0.530	0.536
KKFMF	0.809	0.854	0.804	0.470	0.536
MOR	0.710	0.869	0.809	0.462	0.514
NSSDMF	0.444	0.863	0.788	0.375	0.492
SSD	0.352	0.846	0.751	0.579	0.567
SAD	0.238	0.842	0.908	0.471	0.586
SSDMF	0.739	0.832	0.749	0.543	0.527
SADMF	0.709	0.823	0.839	0.493	0.552
SSDMS	0.802	0.802	0.722	0.531	0.532
SADMS	0.783	0.800	0.843	0.486	0.559
SES1	0.779	0.842	0.816	0.365	0.450
SES2	0.802	0.852	0.838	0.391	0.513
SEK1	0.746	0.804	0.830	0.333	0.402
SEK2	0.766	0.821	0.847	0.369	0.465
UKKF	0.620	0.707	0.626	0.212	0.293
UKKFMF	0.625	0.714	0.631	0.231	0.303
NIS (w=2)	0.677	0.750	0.810	0.249	0.330
NIS (w=3)	0.735	0.754	0.714	0.323	0.363
NIS (w=4)	0.724	0.755	0.684	0.368	0.380
NIS (w=5)	0.715	0.755	0.682	0.404	0.400
NIS (w=6)	0.694	0.720	0.665	0.423	0.377
NIS (w=7)	0.671	0.709	0.645	0.459	0.377
NIS (w=8)	0.639	0.699	0.606	0.483	0.369
NA1	0.269	0.330	0.306	0.064	0.092
NA2	0.522	0.593	0.504	0.159	0.244

Tab.13: Übersicht über die mittleren Detektionsraten für Set 2 (EDGES) bei einem Toleranzkreisradius von $\sqrt{2}$ Pixeln.

Störung Alg.	80%G					ZOOM		
	80%G	IRIS 80%S	NOISE 80%S	SALTY 80%S	FOCUS 80%S	OG	MS	80%S
KKF	2.0	1.00	0.89	0.85	0.79	7	42.3	16.0
KKFMF	2.3	1.00	1.00	0.97	0.71	5	35.3	15.1
MOR	2.0	1.00	1.00	0.97	0.53	9	34.3	15.1
NSSDMF	2.0	0.37	1.00	0.94	0.37	9	32.7	15.1
SSD	2.0	0.15	0.96	0.89	0.83	5	45.3	16.0
SAD	2.0	0.15	0.96	1.00	0.56	5	54.0	16.0
SSDMF	2.0	1.00	0.95	0.90	0.79	9	36.5	15.1
SADMF	2.0	0.97	0.93	0.98	0.78	7	43.0	16.0
SSDMS	2.0	1.00	0.85	0.79	0.79	7	42.0	15.1
SADMS	2.0	1.00	0.86	0.94	0.68	7	48.0	16.0
SES1	2.0	0.92	1.00	0.92	0.37	9	21.5	15.1
SES2	2.0	0.98	1.00	0.93	0.41	5	36.0	17.2
SEK1	2.8	0.98	0.97	0.97	0.28	13	18.7	12.9
SEK2	2.8	0.98	1.00	0.99	0.42	17	22.3	17.2
UKKF	9.4	0.90	0.99	0.90	0.31	19	12.0	10.6
UKKFMF	8.7	0.90	0.99	0.90	0.33	19	12.0	10.6
NIS (w=2)	3.9	0.80	0.95	1.00	0.32	15	16.0	11.3
NIS (w=3)	4.8	0.95	0.95	0.93	0.41	15	21.6	15.1
NIS (w=4)	5.5	0.93	0.97	0.92	0.44	11	22.7	17.2
NIS (w=5)	4.7	0.91	0.98	0.91	0.53	13	24.3	20.5
NIS (w=6)	5.9	0.91	0.99	0.90	0.55	19	23.7	21.9
NIS (w=7)	5.9	0.89	0.97	0.90	0.61	17	21.5	19.0
NIS (w=8)	6.4	0.89	0.97	0.90	0.62	19	24.0	21.0
NA1	20.4	0.69	0.77	0.55	0.10	23	5.3	4.0
NA2	11.5	0.74	0.79	0.58	0.17	21	12.7	11.6

Tab.14: Übersicht über die 80%-Mustergrößen (**80%G**), die 80%-Bildstörungsstärken (**80%S**), die optimalen Mustergrößen (**OG**) und die Maximumstörungsstärken (**MS**) für Set 2 (**EDGES**) bei einem Toleranzkreisradius von $\sqrt{2}$ Pixeln. Sowohl **MS** als auch **80%S** sind für den Störungstyp **ZOOM** in Prozenten angegeben (Größenveränderung, relative Expansion).

Störung Alg.	IRIS	NOISE	SALTY	FOCUS	ZOOM
KKF	0.757	0.768	0.674	0.583	0.712
KKFMF	0.771	0.808	0.747	0.576	0.708
MOR	0.667	0.825	0.756	0.575	0.695
NSSDMF	0.392	0.809	0.736	0.543	0.687
SSD	0.278	0.800	0.711	0.665	0.723
SAD	0.179	0.798	0.911	0.620	0.742
SSDMF	0.695	0.758	0.680	0.613	0.707
SADMF	0.658	0.745	0.802	0.612	0.714
SSDMS	0.755	0.740	0.649	0.582	0.708
SADMS	0.741	0.734	0.807	0.565	0.718
SES1	0.735	0.803	0.803	0.470	0.686
SES2	0.748	0.813	0.832	0.496	0.722
SEK1	0.708	0.782	0.828	0.453	0.646
SEK2	0.718	0.791	0.846	0.472	0.678
UKKF	0.594	0.641	0.468	0.303	0.380
UKKFMF	0.620	0.674	0.509	0.323	0.455
NIS (w=2)	0.600	0.648	0.762	0.327	0.464
NIS (w=3)	0.655	0.706	0.672	0.360	0.487
NIS (w=4)	0.660	0.718	0.652	0.387	0.525
NIS (w=5)	0.680	0.725	0.657	0.406	0.522
NIS (w=6)	0.660	0.718	0.646	0.430	0.509
NIS (w=7)	0.662	0.704	0.635	0.446	0.442
NIS (w=8)	0.646	0.701	0.630	0.470	0.424
NA1	0.134	0.133	0.172	0.034	0.049
NA2	0.497	0.531	0.437	0.228	0.357

Tab.15: Übersicht über die mittleren Detektionsraten für Set 3 (CHAOS) bei einem Toleranzkreisradius von $\sqrt{2}$ Pixeln.

Störung Alg.	ZOOM							
	80%G	IRIS 80%S	NOISE 80%S	SALTY 80%S	FOCUS 80%S	OG	MS	80%S
KKF	2.0	0.82	0.85	0.76	0.76	7	46.5	25.1
KKFMF	2.0	0.90	0.96	0.91	0.83	9	43.2	25.1
MOR	2.0	0.79	0.97	0.92	0.68	5	42.0	25.1
NSSDMF	2.0	0.29	0.91	0.90	0.67	5	41.5	25.1
SSD	2.0	0.18	0.83	0.82	0.79	5	55.3	25.2
SAD	2.0	0.07	0.84	1.00	0.73	5	52.0	23.7
SSDMF	2.0	0.82	0.82	0.78	0.78	5	45.3	26.0
SADMF	2.0	0.75	0.82	0.89	0.78	5	48.0	22.7
SSDMS	2.0	0.82	0.79	0.61	0.76	7	46.5	24.3
SADMS	2.0	0.82	0.79	0.86	0.72	7	46.0	23.7
SES1	2.0	0.84	0.86	0.92	0.49	5	41.5	22.3
SES2	2.0	0.82	0.86	0.92	0.52	5	48.7	23.7
SEK1	2.6	0.74	0.86	0.93	0.45	9	41.2	20.9
SEK2	2.6	0.75	0.85	0.93	0.50	5	44.0	22.7
UKKF	7.8	0.79	0.84	0.76	0.35	17	15.2	13.6
UKKFMF	7.4	0.82	0.84	0.76	0.40	15	22.0	14.5
NIS (w=2)	3.9	0.71	0.78	0.92	0.39	11	25.2	14.0
NIS (w=3)	4.6	0.77	0.86	0.79	0.44	13	32.3	17.2
NIS (w=4)	5.1	0.77	0.89	0.80	0.51	11	33.2	20.9
NIS (w=5)	5.6	0.86	0.91	0.82	0.53	11	31.5	21.2
NIS (w=6)	5.8	0.84	0.91	0.80	0.56	15	23.3	22.7
NIS (w=7)	5.9	0.87	0.92	0.80	0.62	23	22.3	20.7
NIS (w=8)	5.9	0.86	0.94	0.83	0.63	23	20.7	19.0
NA1	25.0	0.00	0.00	0.00	0.00	25	0.0	0.00
NA2	11.2	0.71	0.73	0.51	0.26	15	22.7	16.0

Tab.16: Übersicht über die 80%-Mustergrößen (**80%G**), die 80%-Bildstörungsstärken (**80%S**), die optimalen Mustergrößen (**OG**) und die Maximumstörungsstärken (**MS**) für Set 3 (**CHAOS**) bei einem Toleranzkreisradius von $\sqrt{2}$ Pixeln. Da die Detektionsraten von **NA1** bei keiner Mustergrößen-Störstärkenkombination den Schwellwert von 0.8 erreichte, wurden alle **80%S**-Werte auf null gesetzt! Sowohl **MS** als auch **80%S** sind für den Störungstyp **ZOOM** in Prozenten angegeben (Größenveränderung, relative Expansion).

Störung	IRIS			NOISE			SALTY			FOCUS			ZOOM		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Alg.	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
KKF	4	2	2	12	10	10	16	14	14	3	4	5	4	5	6
KKFMF	1	1	1	2	3	4	9	10	10	8	9	7	8	5	7
MOR	6	13	11	1	1	1	10	9	9	9	10	8	9	9	10
NSSDMF	21	22	22	5	2	3	12	11	11	17	15	10	10	11	11
SSD	14	13	13	6	5	6	14	12	12	1	1	1	2	2	2
SAD	23	23	24	8	6	7	2	1	1	2	8	2	1	1	1
SSDMF	2	9	9	10	8	11	17	13	13	5	2	3	3	8	9
SADMF	11	14	15	11	9	12	8	4	7	6	5	4	6	4	5
SSDMS	3	3	3	14	13	13	18	15	13	3	3	6	5	7	7
SADMS	9	5	5	16	14	14	6	3	5	6	6	9	7	3	4
SES1	7	6	6	4	6	5	5	7	6	14	18	13	13	13	12
SES2	5	3	4	3	4	2	4	5	3	11	14	11	11	10	3
SEK1	10	8	8	9	12	9	3	6	4	15	19	15	14	14	14
SEK2	8	7	7	7	10	8	1	2	2	12	16	12	12	12	13
UKKF	13	20	20	15	22	22	22	22	22	23	23	22	22	22	22
UKKFMF	12	19	18	13	20	20	20	21	20	22	23	23	21	22	20
NIS (w=2)	17	16	19	20	18	18	7	8	8	11	11	21	21	21	19
NIS (w=3)	14	10	16	18	17	18	11	16	15	20	20	20	18	20	18
NIS (w=4)	15	11	13	17	15	16	13	13	17	19	17	19	15	16	15
NIS (w=5)	16	12	10	19	15	15	15	18	16	18	13	18	16	15	16
NIS (w=6)	18	15	13	19	19	16	19	19	19	16	12	17	17	17	17
NIS (w=7)	18	17	12	19	19	19	20	20	20	13	11	16	19	17	17
NIS (w=8)	20	18	17	20	20	20	21	21	21	10	7	13	20	19	17
NA1	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22
NA2	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22

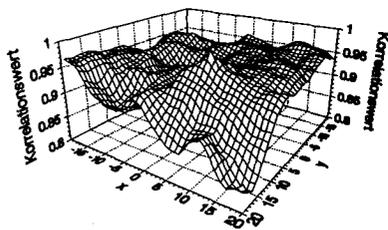
Tab.17: Übersicht über die Rangierungen für die einzelnen Bildsequenzensets (1: TEXTURES, 2: EDGES, 3: CHAOS) anhand der mittleren Detektionsraten bei einem Toleranzkreisradius von $\sqrt{2}$ Pixeln.

Anhang E.3: Resultate, Korrelationsfunktionen

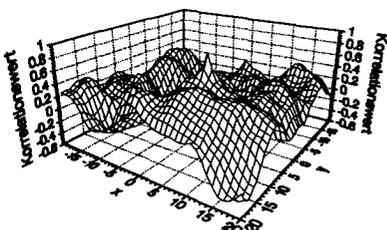
In diesem Teil des Anhangs werden für alle untersuchten Korrelationskriterien die entsprechenden Kreuzkorrelationsfunktionen für die Objektposition 5 mit einem 15x15 Pixel grossen Muster und einer 40x40 Pixel grossen Korrelationsfläche wiedergegeben (mit Ausnahme der **NIS** ($w=4$), die aus Platzgründen weggelassen wurde). Die Suchfenster wurden dem Bild 6 der Sequenz **RAUSCH** von Sequenzset 1 (**TEXTURES**) entnommen. Die (geringe) Störung kann also durch eine theoretische Rauschstandardabweichung von ca. 1.4 Helligkeitseinheiten charakterisiert werden (siehe Kapitel 4: *Versuche und Versuchsbedingungen*).

Die Wahl eines relativ grossen Musters garantiert, dass alle Korrelationskriterien ein vernünftiges Verhalten zeigen. Die sehr schöne Korrelationsfunktion von **NA2** mit ihrer sehr deutlichen Korrelationsspitze darf beispielsweise nicht über die Tatsache hinwegtäuschen, dass bei kleinen Mustern eine wesentlich schlechtere, in der Regel sogar unakzeptable Leistung erwartet werden muss. Dieser Sachverhalt zeigt im übrigen gleichzeitig, dass das Signal-zu-Rauschleistungsverhältnis in der Korrelationsebene ein ungeeignetes Merkmal zur Charakterisierung der Leistung eines Korrelationskriteriums ist. Einzig die Detektierbarkeit des Extremwertes kann dadurch (unter den gegebenen Umständen) zuverlässig beschrieben werden.

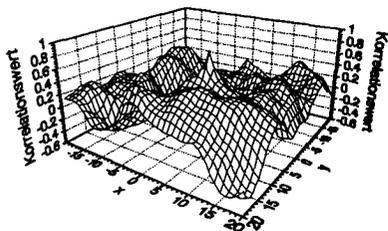
Zum besseren Vergleich und auch wegen der besseren Sichtbarkeit des Extremwerts wurden die Korrelationswerte der Unähnlichkeitsmasse mit einem negativen (konstanten) Wert multipliziert. Damit resultiert eine Darstellung wie bei den Ähnlichkeitsmassen. Hohe Werte deuten in diesen Grafiken also immer auf gute Übereinstimmung zwischen Muster und Korrelationsfenster hin.



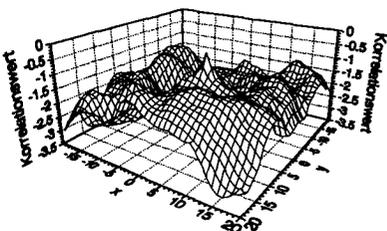
KKF



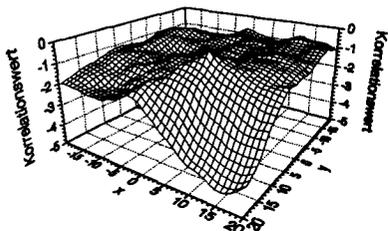
KKFMF



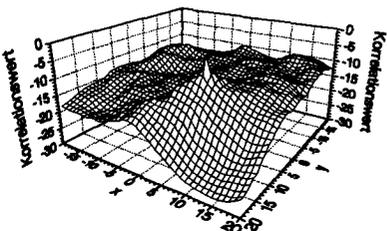
-MOR



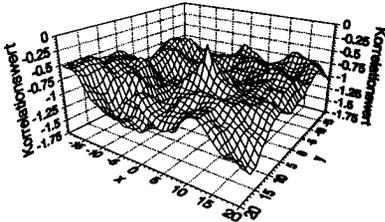
-NSSDMF



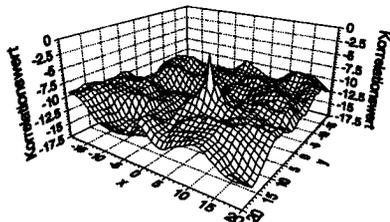
$-10^{-6} \cdot \text{SSD}$



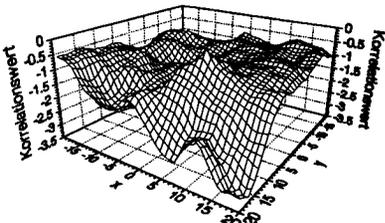
$-10^{-3} \cdot \text{SAD}$



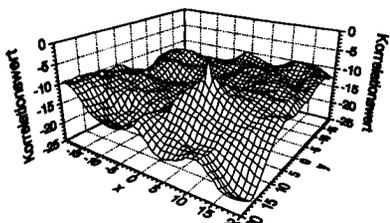
$-10^{-6} * \text{SSDMF}$



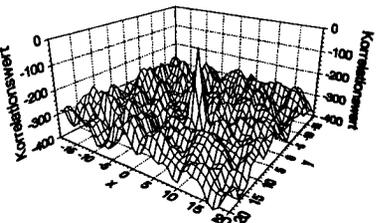
$-10^{-3} * \text{SADMF}$



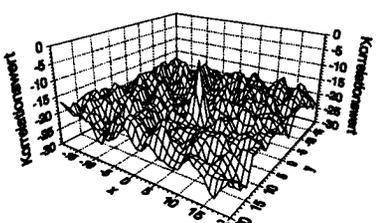
$-10^{-6} * \text{SSDMS}$



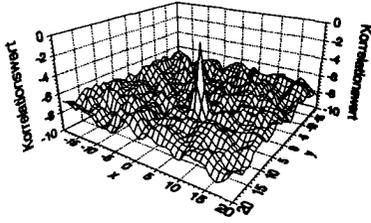
$-10^{-3} * \text{SADMS}$



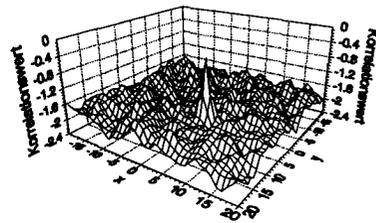
$-10^{-3} * \text{SES1}$



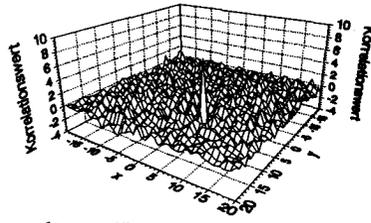
$-10^{-3} * \text{SEK1}$



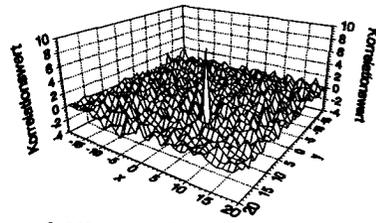
$-10^{-3} * SES2$



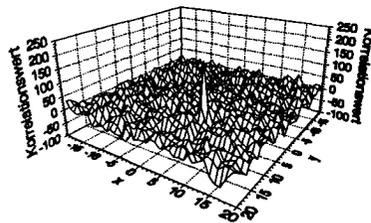
$-10^{-3} * SEK2$



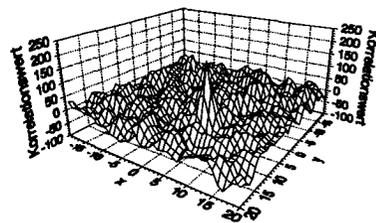
$10^{-6} * UKKF$



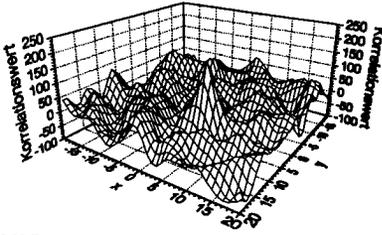
$10^{-6} * UKKFMF$



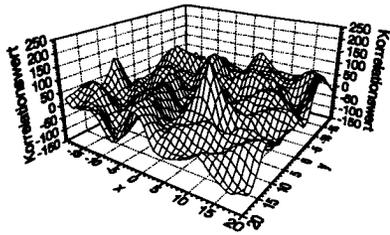
NIS ($w=2$)



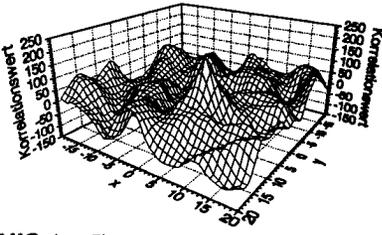
NIS ($w=3$)



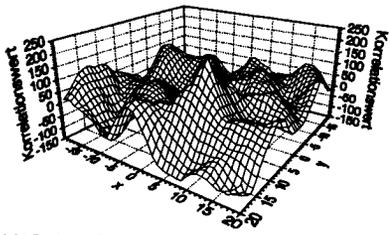
NIS ($w=5$)



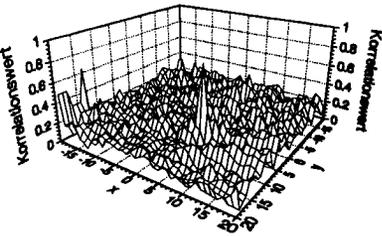
NIS ($w=6$)



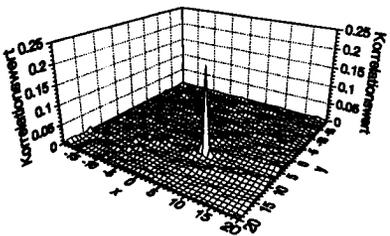
NIS ($w=7$)



NIS ($w=8$)



NA1



NA2

Anhang F: Liste der Abkürzungen

Die im Text verwendeten Abkürzungen sind im folgenden alphabetisch geordnet aufgeführt.

80%G	80%-Mustergrösse; diejenige (linear interpolierte) Mustergrösse, bei welcher mit ungestörtem Bildmaterial noch 80% tolerierbare Positionierungen resultieren.
80%S	80%-Störungsstärke; diejenige (linear interpolierte) Störung, bei welcher mit maximaler Mustergrösse (25x25 Pixel) noch 80% tolerierbare Positionierungen resultieren.
AGC	Automatic gain control, automatische Verstärkungsregelung für das Helligkeitssignal einer Videokamera.
BPOF	Binary phase-only matched filter.
CPU	Central processing unit.
CRA	Cluster reward Algorithmus, basiert auf der co-occurrence-Matrix.
DSC	Deterministic sign change Korrelationsalgorithmus.
HMT	Hit-miss morphological transform, Korrelationsalgorithmus.
KKF	Normierte Kreuzkorrelationsfunktion.
KKFMF	Lokal mittelwertfreie, normierte Kreuzkorrelationsfunktion.
KOM	Binäres Korrelationsverfahren, das auf der co-occurrence-Matrix basiert.
MIMD	Multiple instruction multiple data (Computer).
MOR	Korrelationsfunktion von Moravec.
MPE	Maximum a posteriori Estimation; an die SAD angelehntes Differenzmass.

MPPM	Moment-preserving pattern matching; binäres Korrelationsverfahren.
MS	Maximumstörungsstärke; linear interpolierte Störungsstärke vom Störungstyp ZOOM , die bei der Anwendung der <i>optimalen Mustergrösse</i> (OG) gerade noch eine Detektionsrate von 80% gewährleistet.
NA1	(Beinahe) Originales, binäres Korrelationskriterium nach Nack (mit vorgängiger Hochpassfilterung).
NA2	Modifiziertes, binäres Korrelationskriterium nach Nack (mit vorgängiger Hochpassfilterung).
NIS	Korrelationskriterium nach Nishihara; unnormierte, binäre Kreuzkorrelationsfunktion nach einer Filterung mit dem Laplace-Gauss-Operator.
NSAD	Normierte Summe der Differenzenbeträge ([79]).
NSSD	Normierte Summe der Differenzenquadrate, siehe NSSDMF .
NSSDMF	Normierte, lokal mittelwertfreie Summe der Differenzenquadrate.
OG	Optimale Mustergrösse; (nicht interpolierte) Mustergrösse, bei der die Robustheit gegenüber geometrischer Expansion (ZOOM) maximal wird (beurteilt aufgrund der 80%-Detektionsratenschwelle, siehe auch MS).
POF	Phase-only matched filter.
RMS	Root mean square.
SAD	Summe der Differenzenbeträge (sum of absolute differences)
SADMF	Lokal mittelwertfreie Summe der Differenzenbeträge
SADMS	Lokal mittelwertskalierte Summe der Differenzenbeträge
SEK1	Modifizierter Korrelationsalgorithmus nach Seitz; RMS-Wert der Richtungsabweichungen. Die Richtungen werden mit Hilfe des Kirschoperators bestimmt.

SEK2	Modifizierter Korrelationsalgorithmus nach Seitz; Summe der (kleinsten) Richtungs-differenzbeträge. Die Richtungen werden mit Hilfe des Kirschoperators bestimmt.
SES1	Originaler Korrelationsalgorithmus nach Seitz; RMS-Wert der Richtungsabweichungen. Die Richtungen werden mit Hilfe des Sobeloperators bestimmt (32 Richtungen).
SES2	Modifizierter Korrelationsalgorithmus nach Seitz; Summe der (kleinsten) Richtungs-differenzbeträge. Die Richtungen werden mit Hilfe des Sobeloperators bestimmt (32 Richtungen).
SIMD	Single instruction multiple data (Computer).
SSC	Stochastic sign change Verfahren (Korrelation).
SSD	Summe der Differenzenquadrate (sum of squared differences)
SSDMF	Lokal mittelwertfreie Summe der Differenzenquadrate
SSDMS	Lokal mittelwertskalierte Summe der Differenzenquadrate
UKKF	Unnormierte Kreuzkorrelationsfunktion (inkl. Laplacefilterung)
UKKFMF	Lokal mittelwertfreie, unnormierte Kreuzkorrelationsfunktion (inkl. Laplacefilterung)

Lebenslauf

- 1960 Geboren am 1. Oktober in Zürich
- 1967 - 1973 Primarschule in Zürich
- 1973 - 1975 Sekundarschule in Zürich
- 1975 - 1979 Mittelschule und Matura (Typus C) in Zürich
- 1980 - 1985 Studium der Elektrotechnik an der ETH Zürich
 - 1985 Diplom als Elektroingenieur (Dipl. El.-Ing. ETH)
 - 1985 Gewinn eines SAP-Förderpreises (5. Platz) mit der Diplomarbeit 'Bildererkennungssystem mit μP '.
- 1985 - 1993 Assistent und wissenschaftlicher Mitarbeiter in der Gruppe für Videosignalverarbeitung am Institut für Elektronik der ETH Zürich.
Bis Mitte 1991 Mitarbeit an verschiedenen Industrieprojekten: Vermessungsroboter (Wild Heerbrugg AG), Laserintensitätsprofil-Messgerät (Contraves AG) und Video-Verfolgungssystem (Flugzeugwerk Emmen).