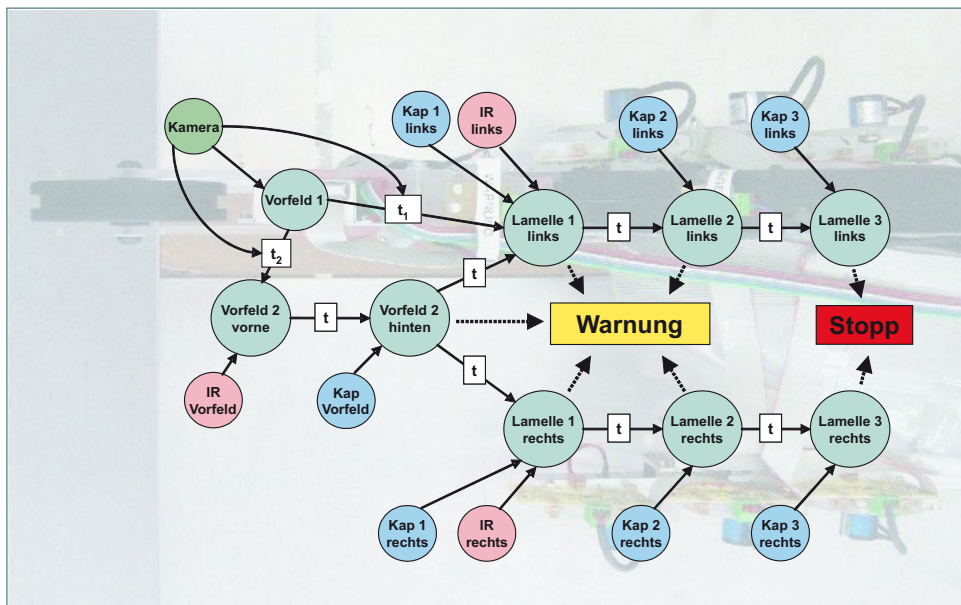


Jörg Barrho

Sensor- und bildverarbeitungs-gestützte Erkennung von Gefahrensituationen



Jörg Barrho

Sensor- und bildverarbeitungsgestützte Erkennung von Gefahrensituationen

Sensor- und bildverarbeitungsgestützte Erkennung von Gefahrensituationen

von
Jörg Barrho



universitätsverlag karlsruhe

Dissertation, Universität Karlsruhe (TH)
Fakultät für Elektrotechnik und Informationstechnik, 2007

Impressum

Universitätsverlag Karlsruhe
c/o Universitätsbibliothek
Straße am Forum 2
D-76131 Karlsruhe
www.uvka.de



Dieses Werk ist unter folgender Creative Commons-Lizenz
lizenziert: <http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

Universitätsverlag Karlsruhe 2007
Print on Demand

ISBN: 978-3-86644-156-9

Sensor- und bildverarbeitungsgestützte Erkennung von Gefahrensituationen

Zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS

von der Fakultät für

Elektrotechnik und Informationstechnik

der Universität Karlsruhe (TH)

genehmigte

DISSERTATION

von

Dipl.-Ing. Jörg Barrho

aus Villingen-Schwenningen

Tag der mündl. Prüfung: 26.06.2007

Hauptreferent: Prof. Dr.-Ing. U. Kiencke, Universität Karlsruhe (TH)

Korreferent: Prof. Dr.-Ing. G. Trommer, Universität Karlsruhe (TH)

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Industrielle Informationstechnik der Fakultät für Elektrotechnik und Informationstechnik an der Universität Karlsruhe (TH). Mein besonderer Dank gilt Herrn Prof. Dr.-Ing. U. Kiencke für die Initiierung dieser Arbeit sowie für seine stetige Unterstützung bei der Durchführung. Herrn Prof. Dr.-Ing. G. Trommer danke ich für die Übernahme des Korreferats und für das entgegengebrachte Interesse an der Arbeit.

Die vorliegende Arbeit ist im Kooperationsprojekt *Entwicklung eines Sicherheitssystems für spanende Werkzeugmaschinen „Cut-Stop“* entstanden, das durch das VDI/VDE-InnoNet-Programm gefördert wurde. Für die Begleitung des Projektes danke ich Herrn Dr. Jens Schulz von der VDI/VDE Innovation + Technik GmbH. Mein Dank gilt auch den Beteiligten des Projektes, insbesondere Herrn Kai Schumacher, Prof. Dr.-Ing. habil. Johannes Tröger, Rolf Tweer und Peter Martin.

Ohne das gute Arbeitsklima am Institut wäre vieles nicht möglich gewesen, weshalb ich mich bei meinen Kolleginnen und Kollegen ebenfalls bedanke. Viele organisatorische Probleme wurden von Frau Manuela Koffler gelöst, für die EDV Administration zeigte sich Dieter Brandt verantwortlich. Herr Hermann Dilger und seine Mitarbeiter unterstützten mich tatkräftig beim Aufbau des Versuchstandes. Herr Stefan Seelinger hat mir bei vielen praktischen Problemen geholfen. Bei Ihnen allen bedanke ich mich herzlich.

Ohne das Engagement der Studierenden, die bei mir ihre Diplom- und Studienarbeit absolviert haben, wäre die vorliegende Arbeit sicher nicht in dieser Form entstanden. Auch ihnen gilt mein Dank.

Für die sorgfältige und kritische Durchsicht des Manuskripts danke ich besonders Dr.-Ing. Julian Baumann, Dr.-Ing. Thorsten Huck, Dr.-Ing. Marcus Hiemer, Thomas Weickert, Michael Schwarz, Michael Bauer, Dr.-Ing. Thomas Barrho sowie Benjamin Kehrer.

Ferner möchte ich mich bei allen Menschen in meinem persönlichen Umfeld bedanken, die mir während meiner Promotion ein ständiger Rückhalt waren und mich begleitet haben. Allen voran sei hier meine Frau Maximiliane genannt, die durch ihre Unterstützung mir stets die Kraft und die nötigen Freiräume für diesen Lebensabschnitt gab.

Zu guter Letzt danke ich meinen Eltern dafür, dass sie mich während meines Studiums auf vielfältige Art und Weise unterstützt haben und somit die Promotion erst ermöglichten.

Karlsruhe, im Juli 2007

Jörg Barrho

*Nicht weil es schwierig ist, wagen wir es nicht,
sondern weil wir es nicht wagen, ist es schwierig.*

Sokrates

Inhaltsverzeichnis

1	Einleitung	1
1.1	Gefahrenbereiche bei Formatkreissägen	3
1.2	Problemstellung	3
1.3	Stand der Technik	4
1.3.1	Passive Sicherheitssysteme	5
1.3.2	Aktive Sicherheitssysteme	6
1.3.3	Gestenerkennung in der Bilddatenverarbeitung	10
1.3.4	Defizite der vorgestellten Systeme	11
1.4	Ziel und Aufbau der Arbeit	12
1.4.1	Ziel	12
1.4.2	Gliederung	13
2	Sensorsysteme zur berührungslosen Handerkennung	15
2.1	Sensorsysteme zur berührungslosen Objektidentifikation	15
2.1.1	Ultraschall-Sensoren	16
2.1.2	Optische Systeme	18
2.1.3	Feldbasierende Sensoren	19
2.2	Handerkennung mit thermopilen Infrarot-Sensoren	19
2.2.1	Sensorprinzip	19
2.2.2	Sensoraufbau	21
2.2.3	Messergebnisse	22
2.2.4	Verwendung bei der Handerkennung	23
2.3	Objektidentifikation mit kapazitiven Sensoren	24
2.3.1	Sensorprinzip	24
2.3.2	Auswertung von Hauptfeldänderungen	26
2.3.3	Auswertung von Streufeldänderungen	28
2.3.4	Materialidentifikation mit induktiven Sensoren	36
2.3.5	Verwendung bei der Handerkennung	38
2.4	Fazit	40
3	Bildverarbeitung zur Handerkennung	43
3.1	Verfahren zur Handidentifikation und -lokalisierung	44
3.2	Bildgebendes System	46
3.2.1	Kamera und Objektive	46
3.2.2	Beleuchtung	48

3.3	Bewegungserkennung in Bildsequenzen	49
3.3.1	Voraussetzungen und Einschränkungen	49
3.3.2	Methoden im Überblick	51
3.3.3	Alpha-Segmentierer	54
3.4	Bildvorverarbeitung	56
3.4.1	Bildsegmentierung	56
3.4.2	Reduktion von Bildrauschen	59
3.4.3	Kantenerkennung	60
3.4.4	Ablauf der Bildvorverarbeitung	62
3.4.5	Ergebnis der Vorverarbeitung	63
3.5	Merkmalsextraktion	64
3.5.1	Verfahren zur Merkmalsextraktion	65
3.5.2	Kreis-Hough-Transformation	67
3.6	Lokalisation und Klassifikation mit probabilistischen Modellen	72
3.6.1	Aufbau und Erstellen der Modelle	73
3.6.2	Lokalisation und Klassifikation	80
3.6.3	Ergebnis	84
3.7	Handrepräsentation mit Aktiven Konturen	84
3.7.1	Aktive Konturmodelle und B-Spline-Kurven	86
3.7.2	Shape-Space	88
3.7.3	Anpassung von Aktiven Konturen	89
3.7.4	Fazit	91
3.8	Verfolgung von Objekten	92
3.8.1	Tracking	92
3.8.2	Probabilistisches Tracking	93
3.8.3	Condensation Algorithmus	97
3.8.4	Fazit	101
3.9	Analyse von Bildsequenzen zur Gefahrenerkennung	102
3.10	Fazit und Ergebnisse	103
4	Informationsfusion zur Gefahrenanalyse	117
4.1	Fusionsstrukturen zur Gefahrenanalyse	118
4.1.1	Fusionsmodelle	118
4.1.2	Sensorintegration	120
4.1.3	Abstraktionsebenen	121
4.1.4	Fusionsarchitektur	121
4.1.5	Methoden zur Informationsfusion	121
4.2	Bayes'sche Netze	125
4.2.1	Grundlagen	125
4.2.2	Definition eines Bayes'schen Netzes	128
4.2.3	Entwurf von Bayes'schen Netzen	132
4.3	Vorverarbeitung der Sensordaten	134

4.3.1	Thermopile Infrarot-Sensoren	135
4.3.2	Kapazitive Sensoren	136
4.3.3	Diskretisierung der Sensordaten	142
4.4	Informationsfusion mittels Bayes'schem Netz	144
4.4.1	Fusionskonzept zur Gefahrenanalyse	145
4.4.2	Entwurf der Bayes'schen Netze	147
4.4.3	Verknüpfung der Teilnetze	151
4.4.4	Einbindung der Bildverarbeitung	153
4.4.5	Bayes'sches Netz zur Gefahrenanalyse	155
5	Funktionsstruktur des Gefahrenerkennungssystems	159
5.1	Merkmale zur Gefahrenanalyse	159
5.2	Gefahrenzonen und Überwachungsbereiche	160
5.3	Realisierung und Implementierung	162
5.3.1	Schutzlamellen	162
5.3.2	Sensorsysteme	164
5.3.3	Datenverarbeitung und Gefahrenanalyse	166
5.3.4	Aktorik	167
5.3.5	Modularer Aufbau des Systems	167
5.4	Validierung und Ergebnisse	168
5.4.1	Beispiel: Detektion einer Hand	168
5.4.2	Unterschiedliche Handgeschwindigkeiten	174
5.4.3	Verifikation der Gefahrenerkennung	175
5.4.4	Fazit	177
6	Zusammenfassung	179
A	Nomenklatur	183
B	Parameterschätzung mit dem EM-Algorithmus	187
C	Algorithmus zur Lösung des Fitting-Problems	189
D	Condensation Algorithmus	191
E	Beispiel eines Bayes'schen Netzes	193
F	Bedingte Wahrscheinlichkeitstabellen für das Bayes'sche Netz	199
	Literatur	205
	Index	217

1 Einleitung

Formatkreissägen sind stationäre Kreissägen und werden als Werkzeug für die professionelle Holzbearbeitung eingesetzt. Diese Sägen sind vier bis fünffach größer als die aus dem Heimwerkerbereich bekannten Tischkreissägen und besitzen im Gegensatz zu diesen einen Schiebeschlitten. Pro Jahr werden ca. 3000-5000 Formatkreissägen in Deutschland für den heimischen Markt und den weltweiten Export hergestellt.

Ein zentraler Bestandteil von holzbearbeitenden Betrieben wie Schreinereien oder Tischlereien sind Formatkreissägen. Mit diesen Maschinen (s. Abbildung 1.1) werden unter anderem Möbel, Treppen, Boote, etc. hergestellt. Vielfältige Arbeitsgänge werden mit Formatkreissägen durchgeführt, z.B. Längsschneiden, Besäumen, Ablängen, Nuten usw. [32].

Diese Art von Holzbearbeitungsmaschinen gehört mit zu den für menschliche Hände gefährlichsten Arbeitsgeräten, da oftmals in unmittelbarer Nähe zur Gefahrenquelle (hier: Sägeblatt) gearbeitet wird. Allein in Deutschland ereignen sich pro Jahr mehrere hundert Unfälle mit zum Teil schweren Amputationsverletzungen. Ursachen hierfür sind unter anderem eine unzureichende Ausbildung des Bedienpersonals, Missachten von Sicherheitsvorschriften und -verordnungen sowie das Demontieren von Sicherheitsvorrichtungen und die Nichtverwendung von Hilfsgeräten wie einem Schiebestock.

Die Holz-Berufsgenossenschaft hat eine Verletzungsrisiko-Analyse durchgeführt [31], die die Gefährlichkeit von Formatkreissägen belegt. Qualitativ sind die Gefahren bekannt und werden z. B. in Arbeitsvorschriften beschreiben. Um quantitativ die Gefährlichkeit von Formatkreissägen zu belegen, wird ein Maß für die Gefährlichkeit definiert. Dazu wird als so genanntes Risikomaß das relative Risi-



Abbildung 1.1: Formatkreissäge

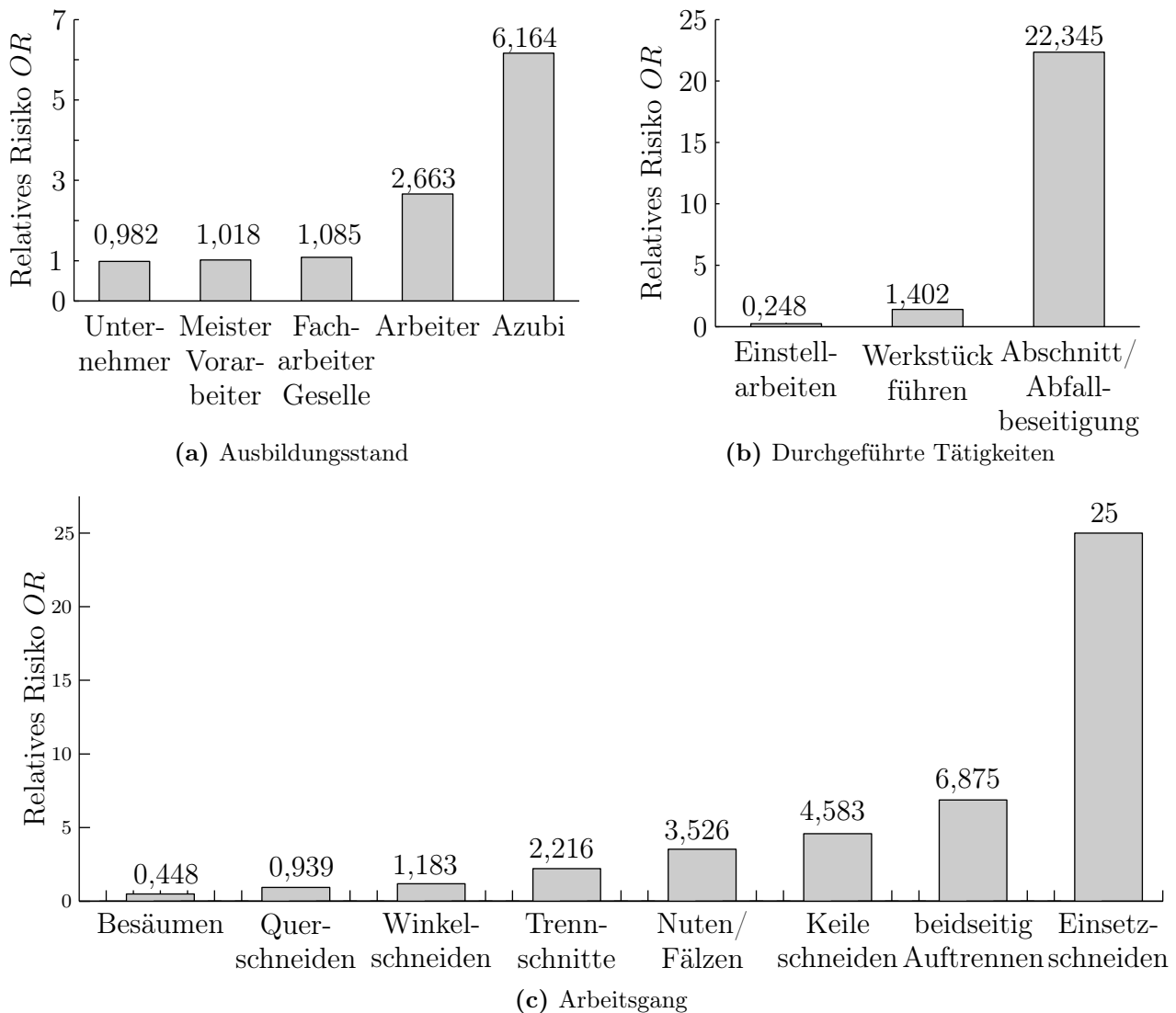


Abbildung 1.2: Betrachtung des Unfallrisikos OR aus [31]

ko OR eingeführt, welches das Risiko angibt, in einem Arbeitsbereich, bei einem Arbeitsablauf verletzt zu werden. Ein $OR > 1$ bedeutet erhöhtes, $OR < 1$ geringeres Verletzungsrisiko.

In Abbildung 1.2 (a) ist der Zusammenhang zwischen Verletzungsrisiko und dem Ausbildungsstand des Maschinenbedieners zu erkennen. Bei Beschäftigten ohne spezielle Ausbildung steigt das relative Risiko eines Unfalls stark an. Das höchste Verletzungsrisiko tragen Ungelernte bzw. Auszubildende.

Ebenso ist in Abbildung 1.2 (b) ersichtlich, dass bei der Abschnitt- bzw. Abfallbeseitigung in der Nähe des Sägeblattes ein sehr hohes Verletzungsrisiko besteht. Hierbei sind besonders unachtsame durchgeführte Wischbewegungen mit hoher Geschwindigkeit zu nennen, die in unmittelbarer Sägeblattumgebung durchgeführt werden, um kleine Sägereste zu entfernen.

Des Weiteren bergen Arbeitsgänge, bei denen in direkter Sägeblattumgebung bzw. ohne zusätzliche Hilfsmittel gearbeitet werden muss, ein besonders hohes

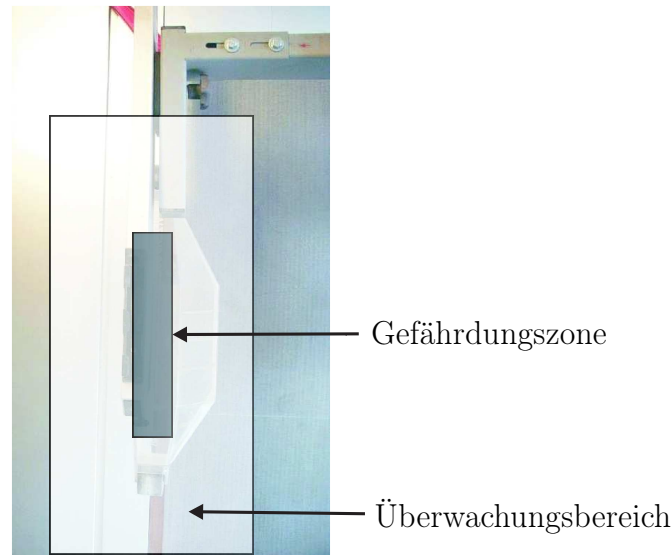


Abbildung 1.3: Gefahrenbereiche einer Formatkreissäge in Draufsicht

relatives Risiko einer Verletzung. Beispiele dazu sind Keile schneiden oder das Einsetzschneiden [32], siehe Abbildung 1.2 (c).

1.1 Gefahrenbereiche bei Formatkreissägen

In Abbildung 1.3 sind die Bereiche einer Formatkreissäge markiert, in denen bei sämtlichen Arbeitsgängen ein erhöhtes Verletzungsrisiko des Maschinenbenutzers zu erwarten ist. Die *Gefährdungszone* entspricht ungefähr der Ausdehnung der Schutzhaube der Maschine. Innerhalb dieses Bereiches ist das Unfallrisiko für den Maschinenbediener aufgrund der räumlichen Nähe zum Sägeblatt überproportional hoch. Im so genannten *Überwachungsbereich* ist eine potentielle Verletzungsgefahr für eine Hand gegeben, jedoch wesentlich geringer als in der Gefährdungszone. Außerhalb dieser beiden Bereiche ist das Verletzungsrisiko gering.

Die Abmessungen der einzelnen Gefahrenbereiche, wie sie in Abbildung 1.4 dargestellt sind, ergeben sich aus der Größe der Schutzhaube und den örtlichen Gegebenheiten des Maschinentisches.

1.2 Problemstellung

Formkreissägen stellen, wie bereits erwähnt und oben dargelegt, eine Gefährdung der körperlichen Integrität eines Menschen dar. Nach DIN EN 61508-4 [23] ist eine *Gefährdung* eine potentielle Schadensquelle. *Schaden* ist in diesem Zusammenhang die physische Verletzung oder Schädigung der Gesundheit eines Menschen. Der Umstand, durch den eine Person einer Gefährdung ausgesetzt ist, wird als *Gefährdungssituation* bezeichnet [23]. Des Weiteren ist die Kombination aus der Wahrscheinlichkeit, mit der ein Schaden auftritt, und dem Ausmaß dieses

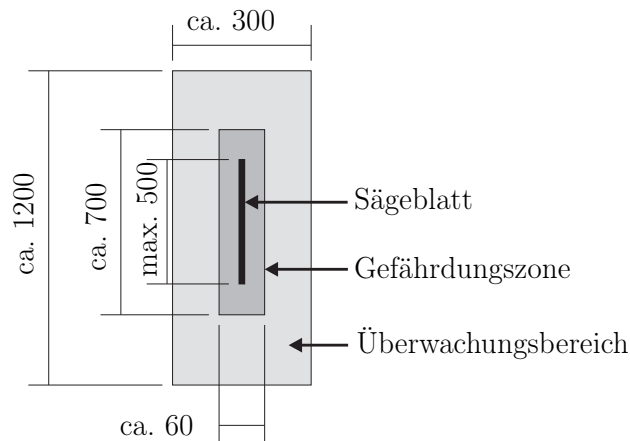


Abbildung 1.4: Abmessungen der Gefahrenbereiche einer durchschnittlichen Formatkreissäge in mm

Schadens das *Risiko* eines Schadens. Demzufolge hat eine Schutzmaßnahme die Aufgabe, das Risiko einer Verletzung zu reduzieren und die Auftrittshäufigkeit einer Gefährdungssituation zu verringern. Ein Schadens-*Restrisiko* wird trotz vorhandener Schutzeinrichtung oder Schutzmaßnahme bleiben. Dieses Restrisiko muss unterhalb des *tolerierbaren Risikos* [23] bleiben, damit die eingesetzte Schutzeinrichtung/-maßnahme wirkungsvoll die Maschinensicherheit gewährleistet.

Das tolerierbare Risiko ist abhängig von den gesellschaftlichen Wertvorstellungen [23] und hat in den letzten Jahrzehnten stark abgenommen, d. h. eine hohe Arbeitssicherheit wird erwartet, gefordert und auch durch den Gesetzgeber im Arbeitsschutzgesetz vorgeschrieben. Aus technischer Sicht bedeutet dies eine Reduktion des Restrisikos.

1.3 Stand der Technik

Um das Restrisiko eines körperlichen Schadens zu reduzieren, sind einige Lösungsvorschläge und -verfahren in den letzten Jahrzehnten entwickelt worden. Als Ergebnis sind die passiven und die aktiven Sicherheitssysteme zu nennen. Einige dieser Systeme werden in den folgenden beiden Abschnitten vorgestellt. Des Weiteren werden einige Verfahren präsentiert, die unter Ausnutzung einer Kamera und der Bildverarbeitung eine menschliche Hand unter bestimmten Voraussetzungen erkennen können. Auf konstruktive Maschinenverbesserungen, Erstellung von Sicherheitsvorschriften, Verfeinerung von Anwendungsregularien oder Erweiterung und Verbesserung der Ausbildung des Maschinenbenutzers wird in dieser Arbeit nicht eingegangen.



(a) Schiebestock



(b) Schutzhaube

Abbildung 1.5: Passive Sicherheitssysteme

1.3.1 Passive Sicherheitssysteme

Aktuell vorgeschriebene und verfügbare Sicherheitsvorkehrung werden im Folgenden beschrieben. Diese sind durchgehend passiver Natur, d. h. der Maschinenbenutzer ist für die Verwendung dieser Sicherheitssysteme selbst verantwortlich.

Schiebestock

Für schmale Werkstücke mit einer Breite von 120 mm und weniger muss ein Schiebestock (siehe Abbildung 1.5 (a)) verwendet werden. Er dient zur Führung dieser schmalen Werkstücke, so dass die Hand des Maschinenbenutzers nicht in die Gefährdungszone der Sägemaschine gelangen kann. Bei sehr schmalen Werkstücken (weniger als 30 mm Breite) sollte ein Schiebeh Holz verwendet werden [32].

Schutzhaube

Innerhalb der Schutzhaube (siehe Abbildung 1.5 (b)) befindet sich das Sägeblatt sowie der Spaltkeil. Dadurch ist das Sägeblatt abgedeckt, jedoch nur wenn die Schutzhaube auf die richtige Arbeitshöhe eingestellt wurde. Zusätzlich werden über die Schutzhaube Sägespäne und kleine Holzreste abgesaugt, so dass die Verletzungsgefahr für die Augen des Maschinenbenutzers durch herumfliegende Holzreste reduziert wird.

Spaltkeil

Der Spaltkeil (siehe Abbildung 1.6) verhindert ein versehentliches Berühren des Sägeblattes von der Maschinenrückseite her und verhindert außerdem ein Einklemmen des Sägeblattes in der Schnittfuge, was mit einem Zurückschlagen des Werkstücks einhergehen würde. Der Spaltkeil muss richtig montiert sein und die korrekte Dicke im Vergleich zum Kreissägeblatt haben [32].

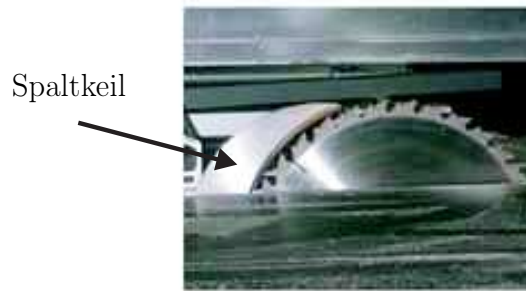


Abbildung 1.6: Spaltkeil

Weitere Hilfssysteme

Durch eine elektrische Motorbremse kann das Sägeblatt in wenigen Sekunden abgebremst werden. Mit einem Parallelanschlag bzw. einem Quer-/ Gehrungsanschlag kann das Werkstück sicher und präzise geführt werden. Besäumhilfen verhindern außerdem das „Verlaufen“ eines Werkstückes, das besäumt werden muss. Unter *Besäumen* versteht man das Abtrennen der Baumkante, z. B. an einer Holzleiste.

1.3.2 Aktive Sicherheitssysteme

Um die Sicherheit der Kreissägemaschinen weiter zu erhöhen, wurden in den letzten Jahren aktive Sicherheitssysteme entwickelt. Die Aufgabe dieser Systemen ist die automatisch Erkennung einer Gefahr für den Maschinenbenutzer und die Auslösung einer Reaktion darauf. Die Gefahrenquelle muss dabei auf eine geeignete Weise neutralisiert werden. Die Gefahrenerkennung soll außerdem vom Benutzer unbemerkt im Hintergrund arbeiten, demontagesicher sein und den Maschinenbenutzer bei seinen gewohnten Arbeitsabläufen nicht behindern. Im Folgenden werden die bisher bekannten Systeme vorgestellt. Im Fokus dabei steht die Gefahrenerkennung der einzelnen Systeme.

Allen bisher vorhandenen aktiven Sicherheitssystemen ist gemeinsam, dass diese nur für Tischkreissägen konzipiert und entwickelt wurden. Tischkreissägen sind jedoch wesentlich kleiner, kostengünstiger und werden auch im Heimwerkerbereich eingesetzt. Für Tischkreissägen werden Sägeblätter bis maximal 250 mm verwendet. Für Formatkreissäge ist dies die kleinste eingesetzte Sägeblattgröße.

Sicherheitssystem für Tischkreissägen

Zwei Schüler stellten beim Bundeswettbewerb „Jugend forscht“ im Jahr 1998 ein Sicherheitssystem für Tischkreissägen vor [68]. Die Erkennung der Hand erfolgt kapazitiv über ein elektrisches Feld. Durch den dielektrischen Einfluss einer Hand oder von Holz verändert sich dieses Feld. Diese Änderung wird durch eine Verstärkung eines Schwingkreises nachgewiesen. Da die unterschiedlichen Holzarten

unterschiedliche dielektrische Eigenschaften aufweisen, ist ein Abgleich mittels eines Potentiometers vor dem Sägen erforderlich. Im Falle einer Handdetektion wird das Sägeblatt durch einen Pneumatikzylinder unter den Arbeitstisch gezogen.

Die Nachteile dieses Systems sind zum einen eine unzuverlässige Unterscheidung zwischen Hand und verschiedenen Hölzern, da vor jedem Sägevorgang das zu sägende Holz dem System eingelernt werden muss. Zum anderen ist die Totzeit von pneumatischen Zylindern für diese Art der Anwendung zu lange.

Sägeblatt unter Hochspannung

Bei diesem patentierten Verfahren wird das Sägeblatt unter Hochspannung von ungefähr 15 kV gesetzt [25]. Durch die elektrische Entladung beim Berühren des Sägeblattes wird dann eine schnelle Ausweich- bzw. Reflexbewegung der Hand des Maschinenbenutzers weg von der Gefahrenquelle angenommen.

Durch die auftretende hohe Entladungsenergie ist unter Umständen mit ernsthaften kardialen Problemen und ggf. leichten Verbrennungen an der Kontaktstelle der Hand zu rechnen. Außerdem ist nicht gewährleistet, dass die Reflexbewegung immer weg von der Gefahrenquelle erfolgt.

Berührungslos wirkende Schutzeinrichtung zur Fingererkennung an Kreissägen

Das in [8] vorgeschlagene System arbeitet mehrschichtig. Ein Bildverarbeitungssystem erfasst den Arbeitsbereich der Tischkreissäge frontal. Tritt ein Maschinenbenutzer an die Maschine, wird mit einer Grob-zu-Fein-Strategie, beginnend am Kopf-Schulter-Bereich bis hin zu den Fingern, versucht, die Silhouette des Benutzers und den möglichen Abstand der Finger zum Sägeblatt zu erkennen. Als weiteres Sicherheitssystem ist ein Infrarot-Sensor an der Schutzhaube montiert, um in einem abgegrenzten Bereich Temperaturunterschiede zu erkennen, die entstehen, wenn eine menschliche Hand in den überwachten Bereich eindringt. Als zusätzliches Sensorsystem wird eine Feldmessung durchgeführt. Hierbei ist es erforderlich, dass der Maschinenbenutzer auf einer Trittmatte steht. Dabei wird ein hochfrequentes Signal in den menschlichen Körper eingekoppelt. Die Hände des menschlichen Körpers wirken dabei als „Sender“. An der Vorderseite der Schutzhaube sind Antennen angebracht, die das von den Händen emittierte Signal empfangen. Die Auswertung des empfangenen Signals erfolgt mittels Schwellwertverfahren. Im Falle, dass eine Gefahrensituation vom System erkannt wurde, wird eine Schutzklappe innerhalb von 60 ms vor das Sägeblatt abgesenkt und der Antriebsmotor der Säge abgeschaltet.

Das hier erwähnte Feldmessungsprinzip funktioniert nicht mehr, sobald der Maschinenbenutzer die Tischplatte berührt und damit geerdet ist. Der gleiche Effekt

tritt bei der Verarbeitung von elektrisch-leitenden Werkstoffen auf. Durch den Sensoraufbau an der Schutzhaube ist die Sicht auf das Sägeblatt stark eingeschränkt. Das Infrarot-System reagiert ebenfalls auf angewärmte Werkstückpartien, die bei der Führung des Holzes entstehen, so dass es zu Fehlauflösungen kommen kann. Die Bildverarbeitung beruht auf einer Hintergrundsubtraktion mit statischen Bildern, die zu Betriebsbeginn aufgenommen werden. Auftretende Beleuchtungsschwankungen können damit nicht ausgeglichen werden. Das Sicherheitssystem ist so konfiguriert, dass bei einer Annäherung der Hand an das Sägeblatt auf einen Abstand von weniger als 12 cm die Schutzklappe aktiviert wird. Dies ist jedoch für einige diffizile Bearbeitungsschritte kritisch. Außerdem besteht durch die schnell schließende Schutzklappe die Gefahr, dass sich der Maschinenbenutzer die Finger einklemmen kann.

Sicherheitskonzept für Tischkreissägen

Das Patent [73] beschreibt ein Sicherheitssystem, das Gefahrensituationen an einer Tischkreissäge erkennt und reagieren kann. Die Reaktion ist zweigeteilt, entweder das Sägeblatt wird abgesenkt oder das Antriebsmoment wird deaktiviert, abhängig vom erkannten Gefährdungsrisiko.

Die Gefahrensituation wird durch eine Näherungssensorik erkannt. Hierbei wird registriert, wann und wo sich eine menschliche Hand im Arbeitsbereich der Maschine befindet. Damit lässt sich die aktuelle Position relativ zum Sägeblatt sowie die Geschwindigkeit und Bewegungsrichtung einer Hand errechnen und die Gefahrensituation in sehr bzw. weniger gefährlich einteilen. Jedoch soll es dem Benutzer der Maschine immer noch möglich sein, seine Hände langsam an dem Sägeblatt vorbei zu bewegen, um filigrane Sägearbeiten durchzuführen. Die Auswerteeinheit muss dies erkennen können.

Die Näherungssensorik beruht auf der Messung der dielektrischen Konstanten in einem Raumelement und über der Zeit. Hierzu ist ein kapazitiver Aufbau notwendig. Im Patent wird dazu Folgendes vorgeschlagen. Eine Elektrode des Kondensatoraufbaus stellt das isoliert aufgehängte Sägeblatt dar. Die Gegenelektroden befinden sich in der Arbeitsplatte. Diese muss deshalb in Multilayer- bzw. Sandwichbauweise hergestellt werden. Die Gegenelektroden sind dadurch isoliert in die Arbeitsplatte eingebettet. Die Form und die Anzahl der Gegenelektroden kann nun so gewählt werden, dass eine gute räumliche Sensorauflösung möglich ist. Jedoch erhöht sich mit der Zahl der Gegenelektroden auch der Aufwand der Messdatenverarbeitung.

Die Informationsverarbeitung beruht im ersten Schritt auf einer Art Differenzverfahren, das die dielektrischen Eigenschaften der Luft bzw. Umgebung kompensieren kann. Dadurch kann die menschliche Hand robuster erkannt werden. Der Referenzkreis besteht hauptsächlich aus der Arbeitsplatte, der Messkreis aus den Gegenelektroden. Die Reaktionszeit des Systems wird mit 100-500 μs angegeben.

Nachteilig wirkt sich bei diesem System das elektrisch isoliert angebrachte Sägeblatt aus. Dadurch wird die Qualität der Schnittoberfläche verschlechtert und die Wiederholgenauigkeit eines Sägevorganges reduziert. Des Weiteren sind Arbeitsplatten in Schichtbauweise sehr aufwendig und teuer zu fertigen. Außerdem ist zu erwarten, dass die Gefahrenerkennung bei feuchtem Holz und Verbundwerkstoffen nicht mehr zuverlässig funktioniert. Da diese Erfindungsmeldung bisher nicht als Prototyp verfügbar ist, können keine weiteren Aussagen über deren Leistungsfähigkeit gemacht werden.

SawStop-Sicherheitseinrichtung

Bei diesem kommerziell verfügbaren Sicherheitssystem für Tischkreissägen wird nach einer Berührung des Sägeblattes mit der Hand das Sägeblatt in äußerst kurzer Zeit (4-24 ms) unter den Maschinentisch abgesenkt (siehe [83] oder Homepage des Herstellers: www.sawstop.com).

Die Detektion einer Hand erfolgt elektrisch durch Berührung des Kreissägeblattes. Hierzu wird über zwei Kopplungsplatten eine Rechteckspannung mit einer Frequenz von 250 kHz in den Detektionsstromkreis eingespeist. Zwischen den Kopplungsplatten ist das Sägeblatt isoliert aufgehängt. Dadurch wird die Rechteckspannung auf das Sägeblatt eingekoppelt. Es entsteht eine Kondensatorserienschaltung. Mögliche Amplitudendämpfungen im normalen Betrieb (z.B. Bearbeiten von trockenem Holz) werden von der Kontrolleinheit ausgeglet. Berührt nun ein Mensch das Sägeblatt, so wird dieses geerdet. Infolge dessen wird das Rechtecksignal signifikant gedämpft. Um die Zuverlässigkeit der Detektion eines menschlichen Gliedmaßes zu erhöhen und eine Fehlauflösung zu vermeiden, wird auch die Geschwindigkeit der Amplitudenänderung betrachtet. So werden nur starke Amplitudendämpfungen innerhalb einer Zeitspanne von ca. 100 μ s als gefährliche Berührung interpretiert. Eine Auslösung der Bremseinheit erfolgt jedoch erst, wenn die Geschwindigkeit der Amplitudenänderung durch den Kontakt Mensch/Sägeblatt größer ist als die Regelgeschwindigkeit der Kontrolleinheit. Die Regelgeschwindigkeit ist ein variabler Parameter des Regelalgorithmus, so dass die jeweiligen Umgebungsverhältnisse berücksichtigt werden können, z. B. Abstand der Kopplungsplatten zum Sägeblatt und Luftfeuchtigkeit.

Wird der Kontakt Sägeblatt/menschliche Hand erkannt, so wird die Bremseinheit aktiviert. Dabei wird ein Schaltimpuls ausgelöst, der einen Draht zum Schmelzen bringt, der zuvor eine Druckfeder auf Spannung hielt. Die nunmehr entspannte Feder schiebt ein Kunststoffelement sehr schnell in das rotierende Sägeblatt, so dass sich dessen Zähne darin verhaken, ein Rückschlag wird ausgelöst, d. h. das Sägeblatt wird schlagartig abgebremst. Das daraus resultierende Moment wird für die Bewegung des Sägeblattes bis unterhalb des Arbeitstisches genutzt. Dies ist möglich, weil Sägeblatt und Antriebsmotor sich auf einer Wippe befinden, die sich um einen Drehpunkt bewegen kann.

Ein Nachteil dieses Verfahrens liegt darin, dass das Sägeblatt zur Aktivierung des Auslösemechanismus berührt werden muss und mindestens ein Finger durch das Sägeblatt angeritzt wird. Des Weiteren kommt es zu Fehlauslösungen, wenn feuchtes Holz oder Verbundwerkstoffe gesägt werden. Konstruktiv ist dieses System ebenfalls kritisch zu bewerten, da zum einen das Sägeblatt isoliert angebracht werden muss. Dies bewirkt eine Verschlechterung der Schnittqualität. Zum anderen ist die Bremswirkung unterschiedlicher Werkzeuge mit verschiedenen Spanwinkeln problematisch, so dass der Bremsvorgang unterschiedlich lange dauern kann [85].

1.3.3 Gestenerkennung in der Bilddatenverarbeitung

In Abschnitt 1.3.2 wird ein System beschrieben, das zur Gefahrenanalyse eine Kamera mit einem Bildverarbeitungssystem einsetzt. Der Einsatz des Systems beschränkt sich auf eine Überprüfung, ob sich ein Mensch vor der Maschine befindet [8].

Die Bildverarbeitung kann jedoch auch für eine genauere Gefahrenanalyse eingesetzt werden, indem ein Bereich direkt überwacht und in diesem nach Fingern gesucht wird. Mit der Zunahme der Leistungsfähigkeit von Datenverarbeitungssystemen und Kameras in den 90er Jahren des letzten Jahrhunderts wurde begonnen, Bildverarbeitungssysteme u. a. zur Zeichensprachenerkennung oder als Mensch-Maschine-Schnittstelle (MMS) einzusetzen. Erste Systeme verwendeten dazu farbkodierte Handschuhe. Basierend auf der Konstellation der erkannten Farbmerkmale wurde dann auf eine Handgeste geschlossen [33]. Diese Methode wird auch *Marker-basierte* Merkmalsextraktion genannt, da hier speziell nach vorher bekannten Objektmarkierungen gesucht wird. Dieses Verfahren setzt allerdings immer voraus, dass der Benutzer ein entsprechendes markiertes Objekt verwendet.

Ein weiteres Anwendungsgebiet ist die automatische Gesichtserkennung z. B. zur Personenidentifikation oder Zugangskontrolle [22]. Hierbei wird Wert auf eine stabile, reproduzierbare Merkmalsextraktion gelegt, um Schlüsselmerkmale für die Klassifikation bzw. Personenidentifikation zu erhalten. Damit kann ein zuverlässiges und robustes System aufgebaut werden.

Mittlerweile ist die Leistungsfähigkeit von Standard-Computern hoch genug, um mit allgemeinen Merkmalsextraktionsverfahren arbeiten zu können. Dadurch kann der Nachteile von Marker-basierten Verfahren, der erforderlichen Verwendung von markierten Objekten, beseitigt werden. Die Herausforderungen für ein Handdetektionssysteme steigen dadurch enorm, da z. B. das zu findende Objekt eine Variation von ähnlichen Farben enthalten, der Hintergrund eine hautähnliche Farbgebung aufweisen oder sich die Beleuchtungssituation stark ändern kann. Dadurch entstehende Kontrastreduktionen oder Bildaufnahmestörungen, die nicht korrigiert oder kompensiert werden können. Je nach Anwendungsgebiet

müssen daher die bestehenden Bildverarbeitungsmethoden optimiert bzw. neue entwickelt werden, um die gestellte Aufgabe lösen zu können.

In [30] werden verschiedene Anwendungsgebiete und Lösungsansätze zur Fingerlokalisierung in Bildern aufgeführt, sowie deren Einschränkungen. Schnelle Verfahren zur Handlokalisierung und -erkennung beruhen hauptsächlich auf der Hautfarbensegmentierung und setzen einen konstanten bzw. bekannten Bildhintergrund voraus. Diese einfachen Verfahren eignen sich als MMS bei bekannter Beleuchtungssituation [21, 59].

Ein aufwendigeres Verfahren, das zusätzlich zur Farbsegmentierung ein Handkonturen-Modell besitzt, kann verschiedene flache Handstellung erkennen [61], wobei gewährleistet sein muss, dass die Hand eine bestimmte Ausdehnung haben und damit im Vordergrund sein muss. Eine andere Arbeit beschäftigt sich damit, über eine Fingerspitzen- und Handkonturerkennung Personen zu identifizieren [94]. Allerdings ist dazu ein definierter, konstanter Hintergrund erforderlich.

Eine weitere Möglichkeit der Handerkennung besteht darin, zusätzlich zur Farbsegmentierung und Konturerkennung, ein Tracking-System einzusetzen, um gezielt sich bewegende Objekte zu verfolgen [10, 16, 79]. Tracking steht hier für die Verfolgung einer Objektbewegung. Zusätzlich kann ein Handmodell bestehend aus Zylindern und Ellipsoiden verwendet werden, um zuverlässig eine Hand zu lokalisieren und zu klassifizieren [82]. Allerdings ist dieses System sehr rechenzeitaufwendig, da das verwendete Handmodell sehr umfangreich ist und das Handtracking auf die räumliche Verfolgung einer Handgeste ausgelegt ist.

1.3.4 Defizite der vorgestellten Systeme

Formatkreissägen zählen, wie bereits erwähnt, zu den gefährlichen Arbeitsgeräten. Selbst in professioneller industrieller Umgebung kommt es regelmäßig zu Unfällen und Verletzungen. Aus diesem Grund wurden im Laufe der letzten Jahrzehnte konstruktive Verbesserungen an den Maschinen vorgenommen, zudem passive Sicherheitseinrichtungen (siehe Abschnitt 1.3.1) entwickelt, Vorschriften und Sicherheitsanweisungen erstellt, sowie die Ausbildung der Maschinenbenutzer verbessert. Dadurch sind die Unfallzahlen nach Auskunft¹ von Maschinenherstellern und Holz-Berufsgenossenschaft deutlich reduziert worden. Jedoch werden passive Sicherheitseinrichtungen oftmals demontiert oder aus Kostengründen schlecht ausgebildete Arbeitskräfte mit der Bedienung der Maschinen betraut.

Aufgrund dieser Tatsachen und den Erfahrungen, wie sie z. B. in der Automobilindustrie mit der Airbag-Technologie gemacht wurden [57], liegt es nahe, aktive Sicherheitssysteme für Kreissägemaschinen zu entwickeln (siehe Abschnitt 1.3.2). In den letzten Jahren sind einige Anstrengungen in diese Richtung hin unternommen worden. Jedoch bestehen nach wie vor entweder Akzeptanzhürden seitens

¹Interne Statistik der HBG (Holz-Berufsgenossenschaft).

der Maschinenbenutzer oder die vorgestellten Lösungen sind nicht für Formatkreissägen geeignet, da diese ausschließlich für Tischkreissägen konzipiert wurden. Außerdem weisen die Systeme Schwächen hinsichtlich der Erkennungssicherheit von Gefahren und der Anwendbarkeit auf.

Die Bildverarbeitung als System zur Handerkennung wird bisher hauptsächlich als MMS eingesetzt, die zur alternativen Bedienung von Geräten dienen kann. Eine weitere Anwendung ist die Gebärdensprachenerkennung, bei der der Fokus auf einer exakten Gestenerkennung liegt. Die Systeme setzen entweder einen homogenen Hintergrund oder nur ein zu beobachtendes Objekt im überwachten Bereich voraus. Außerdem sind die eingesetzten Handmodelle von hoher Komplexität, was einen hohen Berechnungsaufwand bedeutet, oder müssen manuell initialisiert werden.

1.4 Ziel und Aufbau der Arbeit

Zur Erhöhung der Arbeitssicherheit an einer Werkzeugmaschine, im Speziellen einer Formatkreissäge, wird ein neuartiges aktives Sicherheitssystem vorgeschlagen. Im Folgenden werden die Ziele dieser Arbeit und deren Aufbau vorgestellt.

1.4.1 Ziel

Aktive Sicherheitssysteme beinhalten zwei Aufgabenteile. Im ersten Teil wird die Umgebung des Prozesses mithilfe von Sensoren überwacht. Durch die Auswertung der erhaltenen Informationen wird auf eine potentielle Gefahrensituation geschlossen. Der zweiten Aufgabenteil beinhaltet die Reaktion auf die Gefahr. Die vorliegende Arbeit befasst sich mit der Gefahrenanalyse für eine menschliche Hand im Bereich von Formatkreissägen.

Damit eine eindeutige und reproduzierbare Entscheidung einer Gefahrenerkennung möglich ist, muss die Analyse auf der Berechnung und Extraktion von Situationsmerkmalen basieren, die ebenfalls die Handdynamik beschreiben können.

Einsetzbare Sensorsysteme müssen diese Merkmale entweder selbst oder in Kombination mit anderen erfassen können. Durch die Sensorik dürfen die gewohnten Arbeitsabläufe der Maschinenbenutzer nicht beeinträchtigt werden, damit diese das Sicherheitssystem überhaupt akzeptieren. Deshalb werden nur berührungslos arbeitende Sensoren verwendet.

Nach dem Erfassen einer Situation durch die Sensoren wird eine Gefahrenanalyse durchgeführt. Die Reaktion auf eine Gefahr muss erfolgen, ohne dass eine Verletzung des Maschinenbenutzers eintritt.

Nach geltenden Sicherheitsvorschriften muss eine Berührung der Gefahrenquelle durch das Sicherheitssystem ausgeschlossen und verhindert werden. Neben akti-

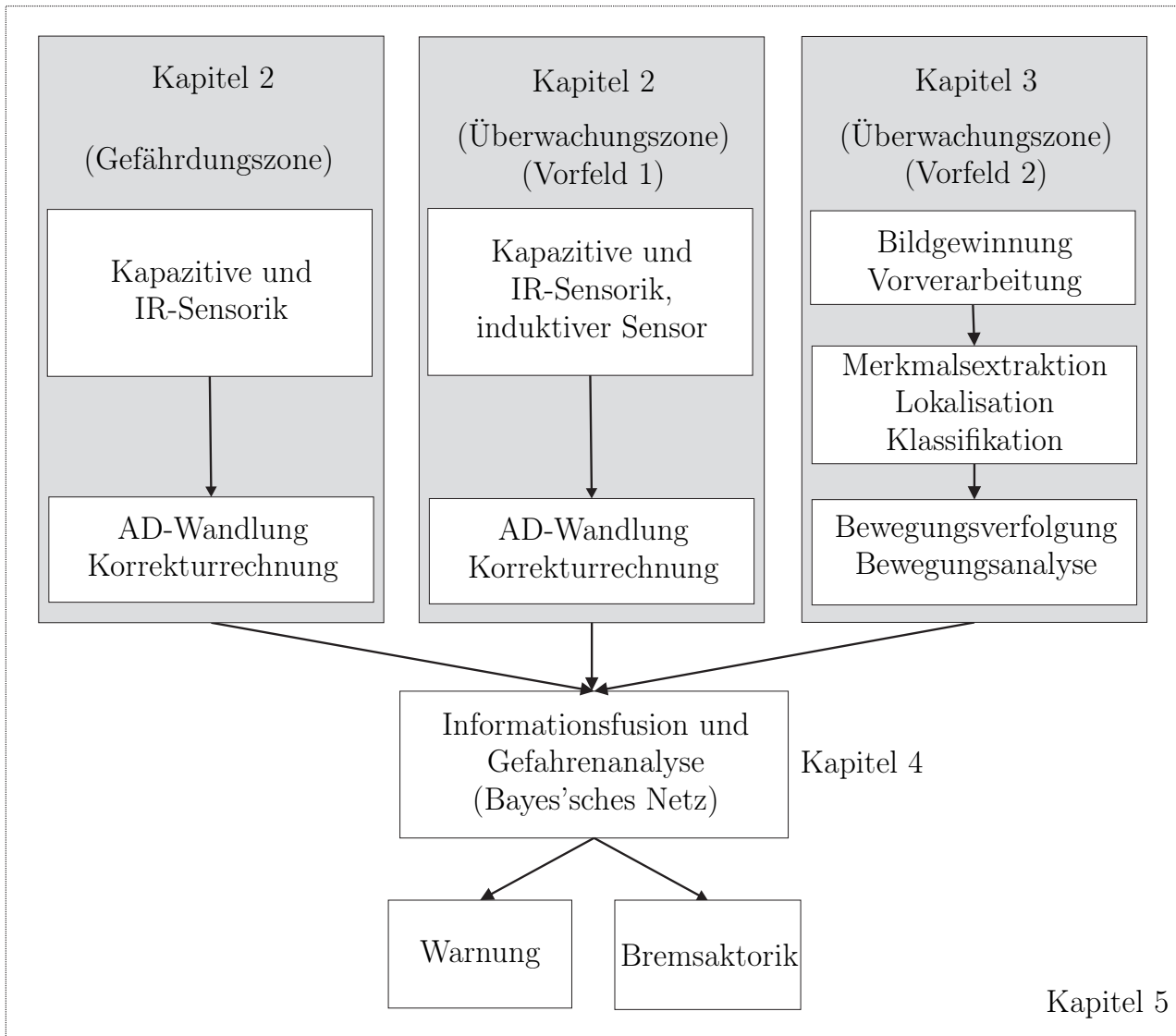


Abbildung 1.7: Struktur des Systems zur Gefahrenanalyse

ven Maßnahmen sind auch passive Schutzvorrichtungen notwendig, um die Überwachungssysteme zu unterstützen. Diese Forderung leitet sich aus den Vorschriften zum *Betreiben von Arbeitsmitteln* der Berufsgenossenschaftlichen Zentrale für Sicherheit und Gesundheit her [7].

1.4.2 Gliederung

In Abbildung 1.7 ist die Struktur des in dieser Arbeit entwickelten Systems zur Gefahrenanalyse aufgezeichnet.

In Kapitel 2 werden berührungslos arbeitende Sensorsysteme zur Situationserfassung im Überwachungsbereich und in der Gefährdungszone vorgestellt. Dies sind kapazitive Sensoren zur Messung von Impedanzänderungen im Erfassungsbereich dieser Sensoren und Infrarot-Sensoren zur Registrierung von sich schnell ändernden Wärmeprofilen. Diese beiden Effekte können das Eindringen einer Hand in

die überwachte Region beschreiben. Für die Eliminierung von Störeinflüssen auf die kapazitiven Sensoren sind weitere Sensorsysteme sowie Korrekturrechnungen der Signale notwendig.

Ein weiteres sehr leistungsstarkes, berührungslos arbeitendes Sensorsystem stellt eine Kamera mit einem Bildverarbeitungssystem dar, welches in Kapitel 3 präsentiert wird. Es ist speziell für die Überwachung einer Gefahrenstelle, bei der menschliche Hände involviert sind, entwickelt. Die Datenverarbeitungskette für die Bildverarbeitung beginnt mit einer Bewegungsdetektion, gefolgt von einer Farbsegmentierung und der Lokalisation sowie der Klassifikation einer Hand mit probabilistischen Modellen. Anschließend werden mittels Aktiver Konturen und eines Partikelfilters die Handbewegungen verfolgt.

Nach dem Erfassen der aktuellen Situation an einer Maschine mittels der genannten Sensoren ist eine Gefahrenanalyse durchzuführen. Diese erfolgt mithilfe einer auf einem Bayes'schen Netz basierenden Informationsfusion. In Kapitel 4 werden die Zusammenhänge näher erläutert und ein Bayes'sches Netz entwickelt. Damit wird parallel eine Gefahrenanalyse durchgeführt.

Das Gesamtsystem für die automatische Gefahrenanalyse im Rahmen eines aktiven Sicherheitssystems für eine Formatkreissäge wird in Kapitel 5 vorgestellt. Die Merkmale, welche die aktuelle Situation an einer Maschine beschreiben, werden definiert. Die verwendeten Sensorsysteme und die Informationsfusion müssen diese gewinnen und auswerten können, damit eine eindeutige und reproduzierbare Gefahrenanalyse möglich ist. Die im Rahmen des Gesamtprojektes (siehe [39]) entwickelten Reaktionen auf die erkannten Gefahren werden in Abschnitt 5.3.4 vorgestellt. Ein ausführliches Beispiel und die Validierung der Systeme schließen Kapitel 5 ab.

Die wesentlichen Ergebnisse dieser Arbeit sind in Kapitel 6 noch einmal zusammengefasst.

2 Sensorsysteme zur berührungslosen Handerkennung

Eine *berührungslos* arbeitende Handerkennung ist im Umfeld einer Formatkreissäge zur Gefahrenerkennung notwendig. In Abschnitt 1.4.1 wird dies von einem aktiven Sicherheitssystem gefordert, um den Maschinenbenutzer bei seiner Arbeit nicht zu beeinträchtigen.

In diesem Kapitel werden mögliche Sensorprinzipien dazu vorgestellt. Ein allgemeiner Überblick wird im folgenden Abschnitt gegeben. Infrarot- und kapazitive Sensoren eignen sich besonders für eine berührungslose Handerkennung, weshalb diese genauer in den Abschnitten 2.2 und 2.3 erklärt werden. Zur Materialidentifikation bzw. der Erkennung von Verbundmaterialien können induktive Sensoren eingesetzt werden, s. Abschnitt 2.3.4. Verbundwerkstoffe sind z. B. auf Holz basierende Materialien mit einer Oberschicht aus Aluminium als Blende. Eine Kamera ist ein Sonderfall der nicht-taktilen Sensoren, da diese eigentlich eine Gruppierung von tausenden von Helligkeitssensoren darstellt. Dies erfordert eine andere Betrachtungsweise der Sensorsignale, s. Kapitel 3.

2.1 Sensorsysteme zur berührungslosen Objektidentifikation

Eine Vielzahl von unterschiedlichen Sensorsystemen sind für die berührungslose Objekterkennung einsetzbar. Die physikalischen Grundlagen dazu bilden zum einen Schall und zum anderen elektromagnetische Wellen unterschiedlicher Wellenlängen. Andere Sensorprinzipien, wie Gamma-Radiometrie oder Trägheitssensoren, sind in diesem Kontext nicht von Bedeutung.

In Abbildung 2.1 sind die wichtigsten Sensorsysteme für berührungslose Detektionssysteme dargestellt und grau hinterlegt. Zwei Sensorprinzipien sind dabei dominant. Zum einen sind dies Systeme, denen Schall zugrunde liegt. Jedoch ist hier nur der Ultraschall-Bereich aus technischen Gründen von Interesse. Weit wichtiger und leistungsfähiger sind Systeme, die elektromagnetische Wellen zur Detektion verwenden. Dabei sind Wellen mit einer Wellenlänge unterhalb 380 nm bzw. oberhalb 10 μm zur Handerkennung nicht sinnvoll. So kann z. B. UV-Strahlung mit einer Wellenlänge von ca. 100 nm für die menschliche Haut schädlich sein und widerspricht damit dem Anspruch eines den Benutzer nicht beeinträchtigenden Sicherheitssystems.

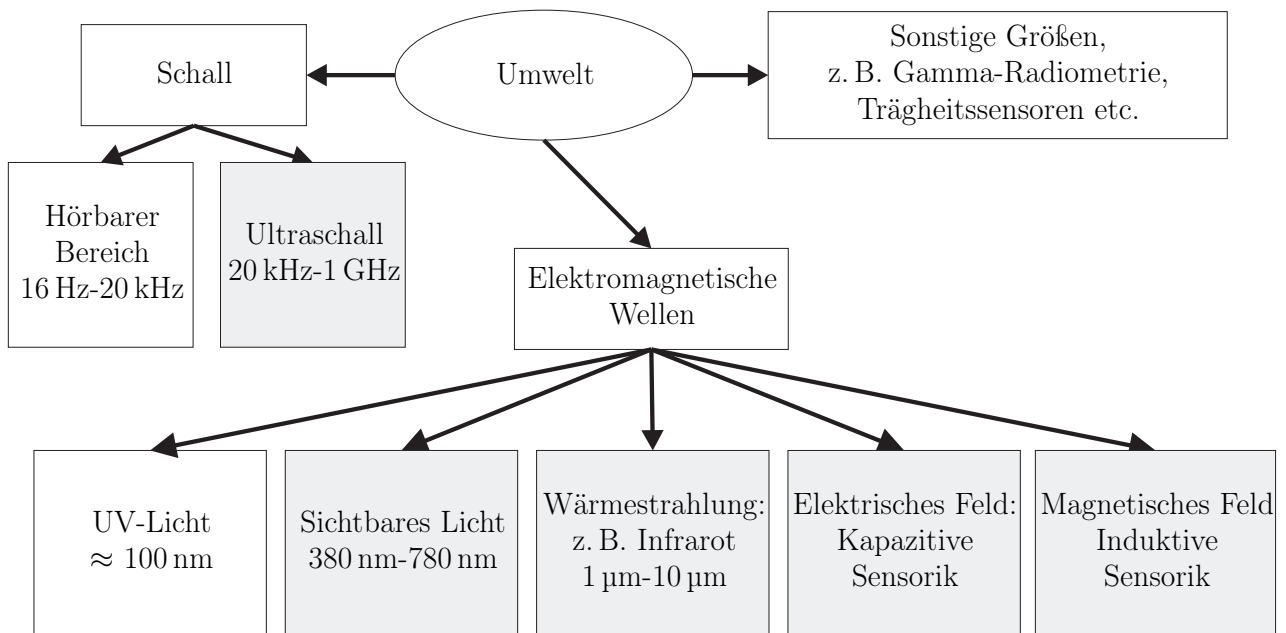


Abbildung 2.1: Übersicht physikalischer Größen für berührungslos arbeitende Sensoren

2.1.1 Ultraschall-Sensoren

Ultraschall-Sensoren erfüllen im industriellen Umfeld eine Vielzahl von Aufgaben wie z. B. Füllstandsmessungen, Materialdiagnostik, Abstandswarner bei Fahrzeugen etc. Die meisten kommerziell verfügbaren Ultraschall-Wandler arbeiten nach dem piezoelektrischen Prinzip [86]. Der Arbeitsbereich dieser Wandler ist sehr schmalbandig, dadurch ist der nutzbare Sendebzw. Empfangsbereich eines Wandlers auf wenige kHz beschränkt.

Mit Ultraschall-Sensoren ist die Messung von Laufzeiten zur Abstandsbzw. Entfernungbestimmung, Auswertung des Dopplereffekts zur Bestimmung von Objektgeschwindigkeiten und schließlich Korrelationsverfahren zur Objekterkennung möglich. Alle drei Ansätze sind für eine Handerkennung von Interesse.

Rahmenbedingungen beim Einsatz an Formatkreissägen: Die zu erwartenden Störgeräusche beim Betrieb von Formatkreissägen liegen unter 20 kHz. Dies wurde durch mehrere Messungen am Institut für Werkzeugmaschinen der Universität Stuttgart bestätigt. Eine Störung der Ultraschall-Sensoren kann daher ausgeschlossen werden. Des Weiteren kann eine Störung durch Holzstaubpartikel weitestgehend ausgeschlossen werden, da an jeder Formatkreissäge eine leistungsfähige Absaugvorrichtung vorhanden ist.

Objekte bewegen sich auf dem Tisch horizontal. Für über dem Tisch angebrachte Wandler entspricht dies einer Lateralbewegung. Die laterale Auflösung der Ultraschall-Wandler ist jedoch deutlich schlechter als die axiale Auflösung, ferner benötigt man für die laterale Lagebestimmung auch die zusätzliche Information über die axiale Entfernung.

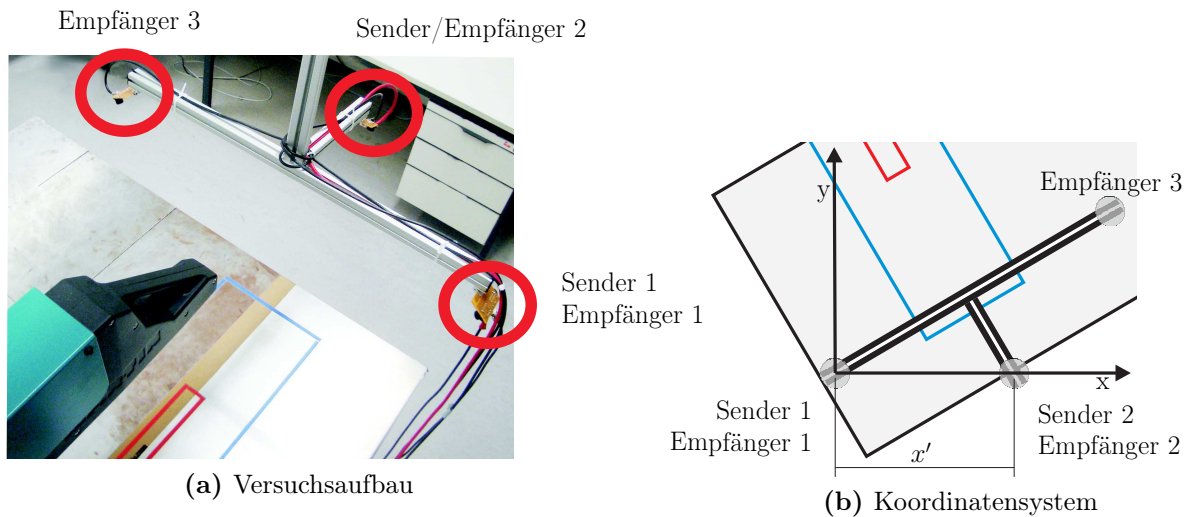


Abbildung 2.2: Ultraschall-Wandler zur Positionsbestimmung

Daneben ergibt sich ein noch zu berücksichtigender Faktor durch die zu detektierenden Objekte. Diese sind in Form und Material sehr vielfältig und reichen von Reliefhölzern über Aluminiumblöcke bis zur menschlichen Hand. Somit besitzen sie auch völlig unterschiedliche Reflexionseigenschaften, sowohl die Intensität des Echos als auch die Richtungscharakteristik betreffend. Objekte mit rauer Oberfläche fächern den Schall auf, während glatte, ebene Objekte ein scharf begrenztes Echo erzeugen.

System zur Positionsbestimmung: Hierzu wird eine Laufzeitmessung des Ultraschalls eingesetzt, s. Abbildung 2.2. Zur genauen Lokalisation eines reflektierenden Objektes sind die Laufzeitinformationen von mindestens drei Sensoren nötig. Ein Array aus mehreren Ultraschall-Wandlern wird dazu verwendet [65, 86]. Die Mehrzahl der Sender-/ Empfängerpaare führt zu zusätzlichen Redundanzen, die sich in der Störungsreduktion positiv bemerkbar machen. Bei der Ansteuerung der Sender muss darauf geachtet werden, dass die Empfängersignale immer eindeutig einem Sender zugeordnet werden können, um Mehrdeutigkeiten zu vermeiden. Dies kann am einfachsten durch die zeitlich versetzte Ansteuerung der Sender erreicht werden.

Die Schwierigkeit der geringen lateralen Auflösung kann durch Zusatzinformationen von mehreren Sensoren gelöst werden. Die mögliche große Streuung durch Objekte auf dem Tisch kann ebenso kompensiert werden. Ein Ultraschallpuls kann zum Beispiel durch starke Streuung aufgefächert werden und im besten Fall von mehreren Empfängern detektiert werden. Dadurch erhält man durch einen Sendeimpuls gleichzeitig drei Laufzeitinformationen. Ein weiterer Vorteil eines Arrays mit mehreren Sendern ist die Möglichkeit, die Sperrung eines Senders ausgleichen zu können, da immer noch das Sendesignal der anderen Sender

genutzt werden kann. Wird also ein Sender von einem Menschen oder einem Objekt verdeckt und lässt somit keine Rückschlüsse über das Gefahrenszenario zu, so ist das System zwar beeinträchtigt, aber immer noch funktionstüchtig.

Bewertung: Eine Positionsbestimmung von Objekten ist mit dem vorgestellten Aufbau im Bereich von Formatkreissägen möglich, für eine genaue Bestimmung der Objektgeschwindigkeit kann eine Dopplerfrequenz-Auswertung der Signale erfolgen. Das Auflösungsvermögen und der Einsatzbereich der verwendeten 40 kHz Ultraschallwandler ist begrenzt aufgrund der relativ breiten Schallkeule dieser Wandler. Zur Steigerung der Genauigkeit sowie der Erweiterung des Überwachungsbereiches können 200 kHz-Wandler eingesetzt werden. Die Handhabung sowie Ansteuerung dieser Wandler ist sehr umfangreich, da diese mit bis zu 200 V angesteuert werden müssen. Dazu sind keine Standardbauelemente für die Auswerte- bzw. Ansterelektronik mehr verfügbar. Dies würde den Aufbau verteuern und wesentlich komplizierter gestalten. Des Weiteren müssen die Sender eines Arrays entweder unterschiedlichen Frequenzen emittieren oder nacheinander angesteuert werden um Mehrdeutigkeiten zu verhindern. Beide Maßnahmen sind schaltungstechnisch aufwendig, wobei letztere die Reaktionszeit des System reduziert.

2.1.2 Optische Systeme

Herkömmliche optische Systeme arbeiten im elektromagnetischen Spektralbereich von 380 nm bis 780 nm. Das entspricht genau dem für das menschliche Auge sichtbaren Wellenbereich. Geeignete Sensorsysteme sind Kameras und Farbsensoren. Kameras als Sensoren zur Handerkennung und die dazugehörige Datenverarbeitung werden in Kapitel 3 genauer betrachtet.

Farbsensoren bestehen aus jeweils einer Fotodiode pro Farbkanal. Als Lichtquelle muss eine definierte, konstante Weißlichtquelle zur Verfügung stehen. Durch diese Konstellation kann die Farbe eines Objektes festgestellt werden. Diese Art von Sensoren wird hauptsächlich in der Automatisierungstechnik eingesetzt. Zur Handerkennung könnte dieses System eingesetzt werden, um mit der Hautfarbe als Merkmal eine Hand zu erkennen. Diese Art von Sensoren ist jedoch auf eine konstante Lichtquelle und eine beleuchtungsinvariante Umgebung angewiesen, um Farbverschiebungen aufgrund von Fremdlicht zu verhindern. Ansonsten würde der Sensor ungenaue oder sogar falsche Messergebnisse liefern. Im Umfeld einer Formatkreissäge kann dies jedoch nicht gewährleistet werden. Es wäre zwar möglich weitere Farbsensoren für Korrekturen zu verwenden, jedoch ist der Aufwand in Bezug auf den Nutzen beschränkt.

Der größte Teil des elektromagnetischen Spektrums liegt im nicht sichtbaren Bereich. Hierbei spielt der Infrarot-Bereich (1 μm -10 μm) für die berührungslo-

se Handerkennung eine wichtige Rolle, da sich die von der menschlichen Hand emittierte Wärmestrahlung genau in diesem Bereich befindet. In Abschnitt 2.2 wird ein auf Infrarot-Strahlung basierendes Sensorsystem vorgestellt.

2.1.3 Feldbasierende Sensoren

Kapazitive und induktive Sensoren werden hauptsächlich in der Automatisierungstechnik als Näherungsschalter oder zur Materialidentifikation eingesetzt, besitzen jedoch in den kommerziell verfügbaren Versionen eine geringe Reichweite. So sind kapazitive Sensoren bis zu einer Reichweite von 10 mm verwendbar. Induktive Sensoren werden je nach Anforderung bis zu einer Reichweite von 10 cm eingesetzt. Die Bauform dieser Sensoren ist allerdings dann ebenfalls groß.

Kapazitive Sensoren eignen sich sehr gut für eine Objektidentifikation basierend auf unterschiedlichen Materialien. Der dafür verantwortliche Sensoreffekt beruht auf der Änderung einer Kapazität, die von der dielektrischen Zahl eines Materials beeinflusst wird. So hat eine menschliche Hand mit ihrem hohen Wasseranteil eine im Vergleich zu anderen Materialien große dielektrische Zahl. Die Umsetzung dieses Prinzips in ein geeignetes Sensorsystem findet sich in Abschnitt 2.3.

Da ein kapazitives Sensorsystem leitende Werkstoffe unter Umständen von Objekten mit einer hohen dielektrischen Zahl, wie z. B. einer Hand, nicht unterscheiden kann, ist ein induktiver Sensor einzusetzen. Zwei Ansätze dazu sind in Abschnitt 2.3.4 beschrieben.

2.2 Handerkennung mit thermopilen Infrarot-Sensoren

Die thermopilen Sensoren gehören, wie auch die pyroelektrische Sensoren, zu der Familie der thermischen Detektoren. Diese Art von Sensoren sind als Bewegungsmelder aus dem Alltag bekannt, wobei fast ausschließlich pyroelektrische Sensoren zum Einsatz kommen, da diese günstiger und schaltungstechnisch einfacher sind. Dagegen werden thermopile Sensoren z. B. zur berührungslosen Temperaturmessung eingesetzt. Diese Sensoren nutzen zur Messung den Seebeck-Effekt aus. Dieser beschreibt die Entstehung einer elektrischen Spannung an den Enden eines stromlosen Leiters, der einen Temperaturgradienten aufweist. In diesem Abschnitt wird das Sensorprinzip der thermopilen Sensorik erläutert, anschließend wird die Eignung zur Handerkennung untersucht.

2.2.1 Sensorprinzip

Der Seebeck-Effekt beschreibt, dass sich zwischen den Enden eines stromlosen Leiters, der einen Temperaturgradienten aufweist, eine Spannung aufbaut (Ab-

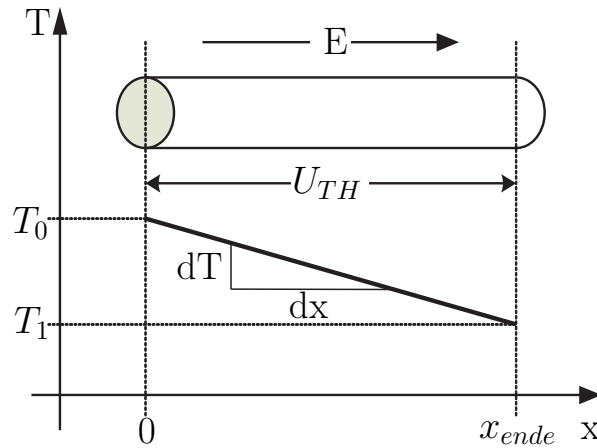


Abbildung 2.3: Entstehung der Thermospannung U_{TH} unter Einfluss eines Temperaturgradienten

bildung 2.3). Durch den Temperaturgradienten besitzt der Leiter an den Enden unterschiedliche Ladungsträgerdichten, da die freien Elektronen am wärmeren Ende eine höhere Bewegungsenergie aufweisen als am kälteren Ende. Dadurch entsteht ein Diffusionsstrom zum kälteren Ende hin, und es bildet sich ein inneres Raumladungsfeld aus, welches den Ladungsträgerfluss im stromlosen Fall kompensiert [86]. Diese entstehende Thermospannung $U_{TH}(T)$ lässt sich durch folgende Gleichung beschreiben

$$U_{TH}(T) = \int_0^{x_{\text{ende}}} E(x, T) dx .$$

Dabei ist das elektrische Feld $E(x, T)$ vom Ort x und der Temperatur T abhängig. Die Integrationsgrenzen ergeben sich aus der Leiterlänge zu $x = 0$ und $x = x_{\text{ende}}$. Wird die Temperaturabhängigkeit des Seebeck-Koeffizienten α_s vernachlässigt, ist bei homogenen Materialien die Thermospannung

$$U_{TH}(T_1, T_0) = \alpha_s \cdot \int_{T_0}^{T_1} dT = \alpha_s \cdot (T_1 - T_0) . \quad (2.1)$$

Der Seebeck-Koeffizient α_s ist außerdem materialabhängig, weiteres dazu findet sich in [26, 75].

Für technische Anwendungen ist die direkte Umsetzung des physikalischen Zusammenhangs in Gleichung (2.1) ungeeignet, da die Abgriffpunkte für die Thermospannung auf gleichem Temperaturniveau liegen würden und sich somit ein weiterer gegenläufiger Temperaturgradient bilden kann. Aus diesem Grund wird

ein zweiter elektrischer Leiter aus einem Material mit großem, dem ersten Material entgegengesetzten, Seebeck-Koeffizient verwendet, so dass sich die einzelnen Absolutbeträge der Thermospannungen addieren [75, 86]. Der dazu notwendige Messaufbau ist in Abbildung 2.4 zu sehen. Durch Addition der einzelnen Teilspannungen und durch Einsetzen von Gleichung 2.1 ist die Thermospannung

$$\begin{aligned}
 U_{TH}(T_1, T_0) &= U_1(T_1, T_0) + U_2(T_0, T_1) \\
 &= \alpha_{s,A} \cdot \int_{T_0}^{T_1} dT + \alpha_{s,B} \cdot \int_{T_1}^{T_0} dT \\
 &= \alpha_{s,A} \cdot (T_1 - T_0) + \alpha_{s,B} \cdot (T_0 - T_1) \\
 &= (\alpha_{s,A} - \alpha_{s,B}) \cdot (T_1 - T_0) .
 \end{aligned}$$

Mit einer gemessenen Thermospannung erhält man die Temperaturdifferenz ΔT zwischen Umgebungstemperatur und der Messstellentemperatur

$$\Delta T = T_1 - T_0 = \frac{U_{TH}(T_1, T_0)}{\alpha_{s,A} - \alpha_{s,B}} . \quad (2.2)$$

2.2.2 Sensoraufbau

In dieser Arbeit wird ein Infrarot-Sensor vom Typ ST60R verwendet [20]. Dieser ist mit einer diffraktiven Linse ausgestattet, so dass ein Detektionswinkel von 10° vorhanden ist und das Infrarot-Spektrum im Bereich von 3-10 μm liegt. In diesem Spektrum liegt gerade das Maximum der von einer Hand emittierten Infrarot-Strahlung. Des Weiteren ist dieser thermopile Infrarot-Sensor mit einer Reaktionszeit von ungefähr 18 ms schnell im Vergleich zu pyroelektrischen Sensoren mit durchschnittlich 150 ms.

Die aktive Sensorfläche eines ST60R-Sensors besteht aus mit Phosphor dotiertem Poly-Silizium und Gold. Diese Materialkombination für das Thermopaar besitzt

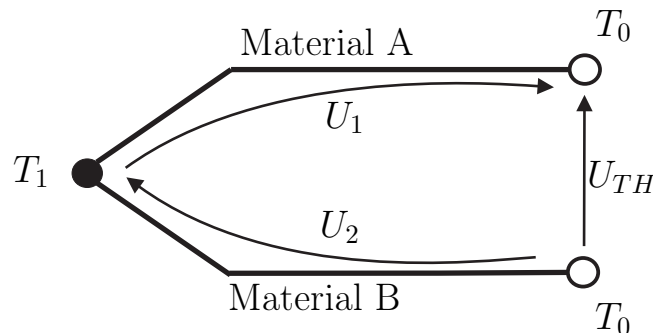


Abbildung 2.4: Prinzipieller Aufbau eines Thermoelements



Abbildung 2.5: Verstärkerschaltung für einen thermopilen Infrarot-Sensor

eine hohe Thermokraft und damit eine hohe Empfindlichkeit. Die Temperatur T_0 aus Gleichung (2.2) entspricht der Umgebungstemperatur. Die Messstelle, an der die Temperatur T_1 herrscht, befindet sich in einem mit Stickstoffoxid gefüllten Raum. Aufgrund dieser Konstruktion ist der Sensoraufbau relativ unempfindlich gegenüber Schwankungen der Umgebungstemperatur.

Die Messstelle reagiert mit einer Temperaturänderung, sobald Infrarot-Strahlung auf diese trifft. Dies bewirkt wiederum eine Änderung der Thermospannung, was gemessen und ausgewertet werden kann.

Ein einzelnes Thermopaar erzeugt selbst unter normalen Bedingungen nur eine sehr kleine Thermospannung. Aus diesem Grund werden für diesen Sensortyp 80 einzelne Thermoelemente zu einem Sensorarray zusammengefasst und in Reihe zusammengeschaltet, so dass eine Ausgangsspannung von ca. $60 \mu\text{V}$ bei 23°C entsteht. Diese immer noch sehr kleine Spannung wird nun mithilfe eines Instrumentenverstärker [84] und einer normalen Verstärkerschaltung um den Faktor 10^6 verstärkt, s. Abbildung 2.5. Der Vollweggleichrichter sorgt hauptsächlich dafür, dass die Ausgangsspannung auf den Bereich $0\text{-}10\text{ V}$ begrenzt und die Restwelligkeit reduziert ist, so dass das Ausgangssignal einfach digitalisiert und weiter verarbeitet werden kann. Um Störeinflüsse, die über die Messleitung eingekoppelt werden könnten, so gering wie möglich zu halten, muss die gesamte Messschaltung in einem Metallgehäuse geschirmt und zusätzlich geerdet werden.

2.2.3 Messergebnisse

In Abbildung 2.6 (a) ist die Ausgangsspannung eines thermopilen Infrarot-Sensors dargestellt, während eine auf einem Holz liegende Hand durch den Sensorbereich gezogen wird. Der Sensor befindet sich in einer Höhe von 7 cm . Eine Hand bzw. ein warmer Gegenstand ist durch den Sensor gut von der Umgebung zu unterscheiden.

Der Messwertverlauf in Abbildung 2.6 (b) entsteht, wenn ein durch eine Hand erwärmtes Holz durch den Erfassungsbereich des Infrarot-Sensors geschoben wird. Der Infrarot-Sensor registriert nur die Infrarot-Strahlung in dessen Detektionsbereich. Über die Art der Strahlungsquelle liefert der Sensor keine Information.

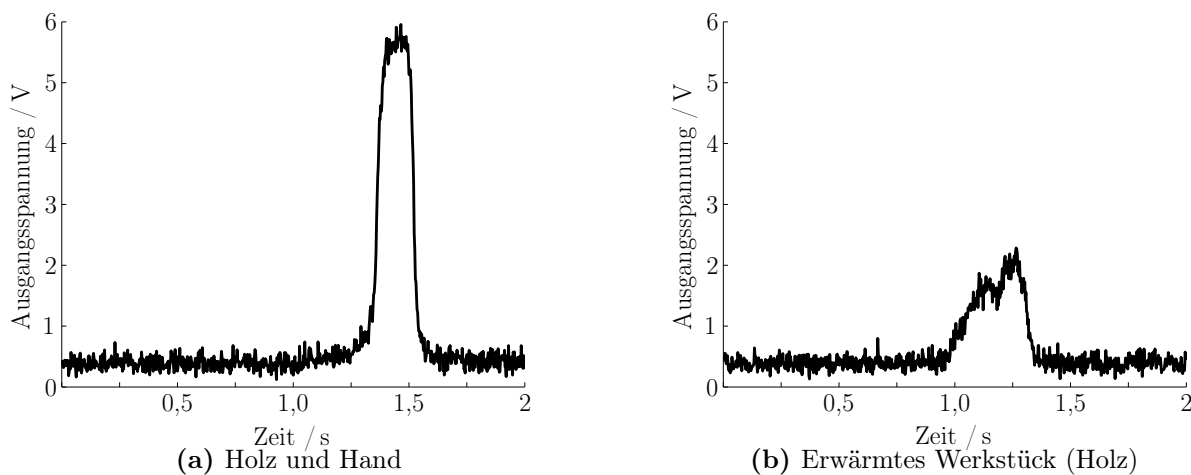


Abbildung 2.6: Messwertverlauf thermopiler Infrarot-Sensor

2.2.4 Verwendung bei der Handerkennung

Das vorgestellte Infrarot-Sensorsystem ist in der Lage, Gegenstände mit von der Umgebungstemperatur abweichender Oberflächentemperatur in dessen Detektionsbereich zu erkennen. Unterschiedliche Objekte mit der gleichen Temperatur können bei der Signalanalyse nicht einzeln aufgelöst werden. Eine exakte Objektidentifikation kann mit dieser Art von Sensor nicht erfolgen.

Die Infrarot-Sensoren erlauben jedoch eine zuverlässige und schnelle Bewegungserkennung einer Hand, solange die Umgebungstemperatur unter ca. 30°C liegt. Dies entspricht einer durchschnittlichen Handtemperatur. Über dieser Grenze funktioniert der Sensor als reiner Bewegungsmelder, ohne auf ein spezielles Objekt zu schließen.

Ist die Umgebungstemperatur unter 30°C und der Infrarot-Sensor registriert eine Temperaturänderung über einer bestimmten Schwelle, kann dies mehrere Ursachen haben:

1. Eine Hand befindet sich im Detektionsbereich des Sensors.
2. Ein Werkstück mit einer Oberflächentemperatur über 30°C wird durch den Sensorbereich bewegt.
3. Ein Werkstück und eine Hand liegen im Erkennungsbereich des Infrarot-Sensors.
4. Starke Schwankungen der Lufttemperatur innerhalb des Detektionsbereiches.

Ein einzelner Infrarot-Sensor liefert für diese Fälle ähnliche, kaum unterscheidbare Ausgangssignale. Jedoch ist dadurch die Menge an zu erkennenden Ursachen bzw. Situationen begrenzt. Mit einer nachfolgenden Datenverarbeitung ist es

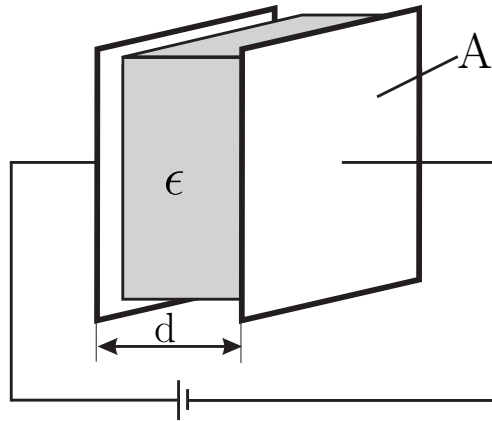


Abbildung 2.7: Aufbau eines Plattenkondensators

nun möglich, in Kombination mit Informationen anderer Sensorarten eine Objektidentifikation zuverlässiger durchzuführen, da mehr unterschiedliches Wissen über ein Szenario zur Verfügung steht (siehe auch Kapitel 4).

2.3 Objektidentifikation mit kapazitiven Sensoren

Kapazitive Sensoren finden im industriellen Alltag ein breites Anwendungsspektrum. So werden diese Sensoren unter anderem als Näherungsschalter, Füllstandssensoren, Tastschalter und als Elemente bei Touch-Screens verwendet. Eine Weiterentwicklung stellt die Materialidentifikation mittels kapazitiver Sensorik dar [38].

Der den Sensoren zugrunde liegenden physikalischen Effekt ist eine Kapazitätsänderung eines Kondensators, der entsprechend der Anwendung optimiert wird. Die Auswertung dieser Änderung lässt Rückschlüsse auf das sich im Sensorbereich befindende Material und ggf. auch auf die relative Objektposition zu.

Im Folgenden wird nun das Funktionsprinzip eines kapazitiven Sensors erklärt. Darauf aufbauend werden die beiden entwickelten Messprinzipien zur Kapazitätsänderung vorgestellt und auf deren Nutzen zur Handerkennung im Arbeitsbereich einer Formatkreissäge näher eingegangen.

2.3.1 Sensorprinzip

Ein Kondensator besteht aus zwei leitenden Elementen, den so genannten Elektroden, und einem dazwischen liegenden Isolator. Die Elektroden können in beliebiger Geometrie angeordnet sein. Der entstandene Zwischenraum wird durch den Isolator bzw. das Dielektrikum ausgefüllt.

Der einfachste Kondensator ist ein Plattenkondensator mit Luft als Dielektrikum. In Abbildung 2.7 ist dies schematisch dargestellt. Ein Plattenkondensator besteht aus zwei Platten als Elektroden mit der Fläche A , die den Abstand d

aufweisen. Dazwischen befindet sich das Dielektrikum mit der Permittivitätszahl ϵ . Die Kapazität ergibt sich dann zu [26]

$$C = \epsilon_0 \cdot \epsilon_r \cdot \frac{A}{d} . \quad (2.3)$$

Die beiden Größen A und d beschreiben den Einfluss der Geometrie auf die Kapazität. Die elektrische Feldkonstante ϵ_0 und die relative Permittivitätszahl ϵ_r ergeben die Permittivitätszahl oder Dielektrizitätszahl $\epsilon = \epsilon_0 \cdot \epsilon_r$.

Die relative Permittivitätszahl ϵ_r ist eine materialspezifische Größe, mit der die elektrische Beeinflussung von äußeren Feldern quantifiziert wird. In der Feldtheorie ist dieser Effekt als Polarisierung bekannt [58]. Bei der Polarisierung wirkt ein äußeres elektrisches Feld auf das Dielektrikum. Dabei wirken Kräfte auf die Ladungen im Material und deformieren die Atome bzw. Moleküle des Objektes derart, dass Dipolmomente entstehen, welche wiederum das angelegte Feld schwächen. Dieser Effekt ist neben den Materialeigenschaften auch von der Frequenz f des angelegten Feldes abhängig. Die komplexe relative Permittivität $\underline{\epsilon}_r$ berücksichtigt dies

$$\underline{\epsilon}_r(f) = \epsilon_r - j\epsilon_r''(f) = \epsilon_r - j \frac{\kappa}{\epsilon_0 \cdot 2\pi \cdot f} . \quad (2.4)$$

Der Realteil von $\underline{\epsilon}_r$ ist ein Maß für die Durchlässigkeit von elektrischen Feldern. Im statischen Fall ($f = 0$ Hz) ist die Permittivität nur von ϵ_r' abhängig. Der Imaginärteil in Gleichung (2.4) verkörpert die frequenzabhängige elektrische Leitfähigkeit κ des Materials [81].

Material	Permittivitätszahl ϵ_r
Vakuum	1
Luft	1,00059
Holz, trocken	2 – 3,5
Wasser	80,1

Tabelle 2.1: Permittivitätszahl ϵ_r verschiedener Materialien [26]

In Tabelle 2.1 sind einige Werte von $\underline{\epsilon}_r$ für $f = 0$ Hz aufgeführt. Der Wert von ϵ_r für Wasser ist ca. 40-mal größer als der für Holz. Dies ist entscheidend für die Nutzung eines kapazitiven Sensors zur Handerkennung, da menschliches Gewebe zu über 60 – 70 % aus Wasser besteht [76]. Des Weiteren wird menschliches Gewebe mit Blut durchströmt, das ein Elektrolyt ist und eine frequenzabhängige Leitfähigkeit aufweist [42]. Für die elektrische Beschreibung von menschlichem Gewebe wird deshalb die relative Permittivität ϵ_r'' (s. Gleichung (2.4)) verwendet.

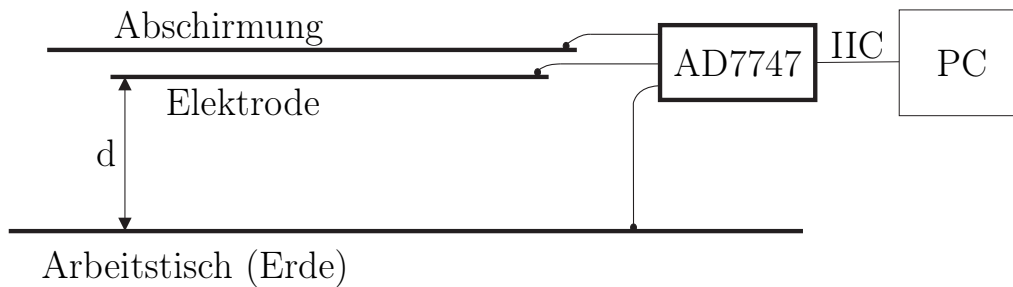


Abbildung 2.8: Aufbau zur Kapazitätsmessung

Der Wert der relativen Permittivitätszahl für menschliches Gewebe ist außerdem von der Gewebetemperatur, der Hautoberfläche und dem Messort abhängig. Um den exakten Wert ermitteln zu können, wären zusätzlich noch Leitfähigkeitsmessungen nötig [38, 60], die nicht berührungslos ausgeführt werden können und daher hier nicht weiter betrachtet werden. In der Anwendung zur Handerkennung kann jedoch angenommen werden, dass der Betrag von ϵ_r ausreichend groß gegenüber trockenem Holz und anderen Werkstoffen ist. Dies wurde durch eine Vielzahl durchgeführter Messungen nachgewiesen, s. [42] und Abbildung 2.15.

Eine kapazitive Materialerkennung beruht auf der charakteristischen Permittivitätszahl des zu untersuchenden Mediums. Der Sensoreffekt kann also auf die Beeinflussung des Hauptfeldes eines Kondensators oder dessen Streufeld durch die elektrischen Materialeigenschaften eines Objektes zurückgeführt werden. Das Hauptfeld ist vom Betrag her dominant. Streufelder entstehen an den Enden der Kondensatorelektroden, da diese nicht unendlich ausgedehnt sein können. Die Streufelder sind im Allgemeinen vernachlässigbar, können jedoch unter bestimmten Umständen für Sensoranwendungen ausgenutzt werden. Im Folgenden werden drei mögliche Arten von kapazitiven Sensoren sowie die passenden Messschaltungen vorgestellt. Im Vordergrund steht hierbei die Erkennung einer Hand bzw. eines Objektes innerhalb eines vorgegebenen Gebietes, basierend auf der Modifikation eines elektrischen Feldes durch die Änderung der Permittivitätszahl.

2.3.2 Auswertung von Hauptfeldänderungen

Ein nahe liegender Ansatz zur Objektidentifikation ist nun, das Objekt zwischen zwei Platten eines ausreichend großen dimensionierten Kondensators hindurch zu führen. Hierbei ist jedoch das Verhältnis Elektrodenfläche A zu Plattenabstand d ungünstig, da ein im Vergleich zu Standardkondensatoren sehr großer notwendiger Plattenabstand nötig ist und die Oberfläche der Platten nicht beliebig vergrößert werden kann. Betrachtet man die Gegebenheiten an einer Formatkreissäge (s. Abschnitt 5.3) so kann es zu einem Plattenabstand d von bis zu $h_S = 10$ cm kommen.

Mit einem Sensor vom Typ AD7747 von Analog Devices [2] ist es möglich, kleine Kapazitätswerte kontinuierlich zu messen. In Abbildung 2.8 ist der dazu notwendige Aufbau dargestellt. Dieser Sensor ermöglicht es, direkt eine Kapazität zwischen der aktive-geschirmten Elektrode, die unterhalb der Erweiterung der Schutzhaube angebracht ist, und dem Arbeitstisch zu messen. Die Messdaten können direkt über den IIC-Bus (Inter-Integrated Circuit-Bus) an eine Auswerteeinheit gesendet werden. Die Verarbeitungsgeschwindigkeit des AD7747 beträgt aufgrund des notwendigen 24 bit Sigma-Delta-Wandlers ca. 10 ms.

Die Fläche einer Elektrode ist bekannt und wird durch die Abmessung der Schutzhaubenerweiterung vorgegeben. Als zweite Elektrode kann der Arbeitstisch benutzt werden. So ergibt sich mithilfe der Gleichung (2.3) eine Kapazität von ungefähr 17 pF bei einem Plattenabstand von 2 cm und einer rechteckigen Elektrodenfläche von ca. $13 \text{ cm} \cdot 3 \text{ cm} = 39 \text{ cm}^2$.

Wird der Plattenabstand auf 10 cm vergrößert, so müsste nach Gleichung (2.3) die Kapazität auf ungefähr 3,5 pF absinken. Messungen mit obigem Aufbau ergaben jedoch einen Wert von ca. 14 pF. Der Grund dafür ist, dass die für eine Kapazitätsbestimmung relevante Ausdehnung der Arbeitstisch-Elektrode abhängig vom Plattenabstand im Allgemeinen nicht begrenzt ist, da der Arbeitstisch im Verhältnis zur ersten Elektrode eine große Fläche hat.

Messergebnisse

In Abbildung 2.9 (a) ist die Kapazitätsänderung in Abhängigkeit vom Plattenabstand d bzw. Schutzhaubenhöhe h_S aufgetragen. Wie gut zu sehen ist, ändert sich der Kapazitätswert in Abhängigkeit vom Plattenabstand. Diese Kurve wird auch zur Kompensation der Höhenänderung der Schutzhaube benutzt, s. Abschnitt 4.3.

In den drei Abbildungen 2.9 (b), (c), (d) sind die Messverläufe bei unterschiedlichen Szenarien dargestellt. Der Plattenabstand beträgt ca. 4,5 cm.

Für die Messverläufe in den Abbildungen 2.9 (b) und (c) sind eine Hand auf einem Holzbrett und ein Holzbrett orthogonal zur Schutzhaube durch den Detektionsbereich geschoben worden. Bei dieser Bewegung wird nur die kurze, 3 cm breite Seite des Sensorbereichs angesprochen. Deutlich zu erkennen ist, dass eine Hand gut von Holz unterschieden wird.

In Abbildungen 2.9 (d) ist der Messverlauf zu sehen, der entsteht, wenn eine Hand durch die 13 cm lange Seite des Sensorbereichs bewegt wird. Zu Beginn (0,4-0,8 s) befinden sich nur einzelne Finger im Detektionsbereich. Anschließend die ganze Hand.

Ein Objekt kann also mit einer hohen Permittivitätszahl und bei fast vollständiger Ausfüllung des Sensorraums sicher erfasst werden. Für eine eindeutige Identifikation eines Objektes ist diese Messung jedoch zu ungenau, da der Kapazitätswert stark von der Ausfüllung des Erfassungsbereiches abhängt.

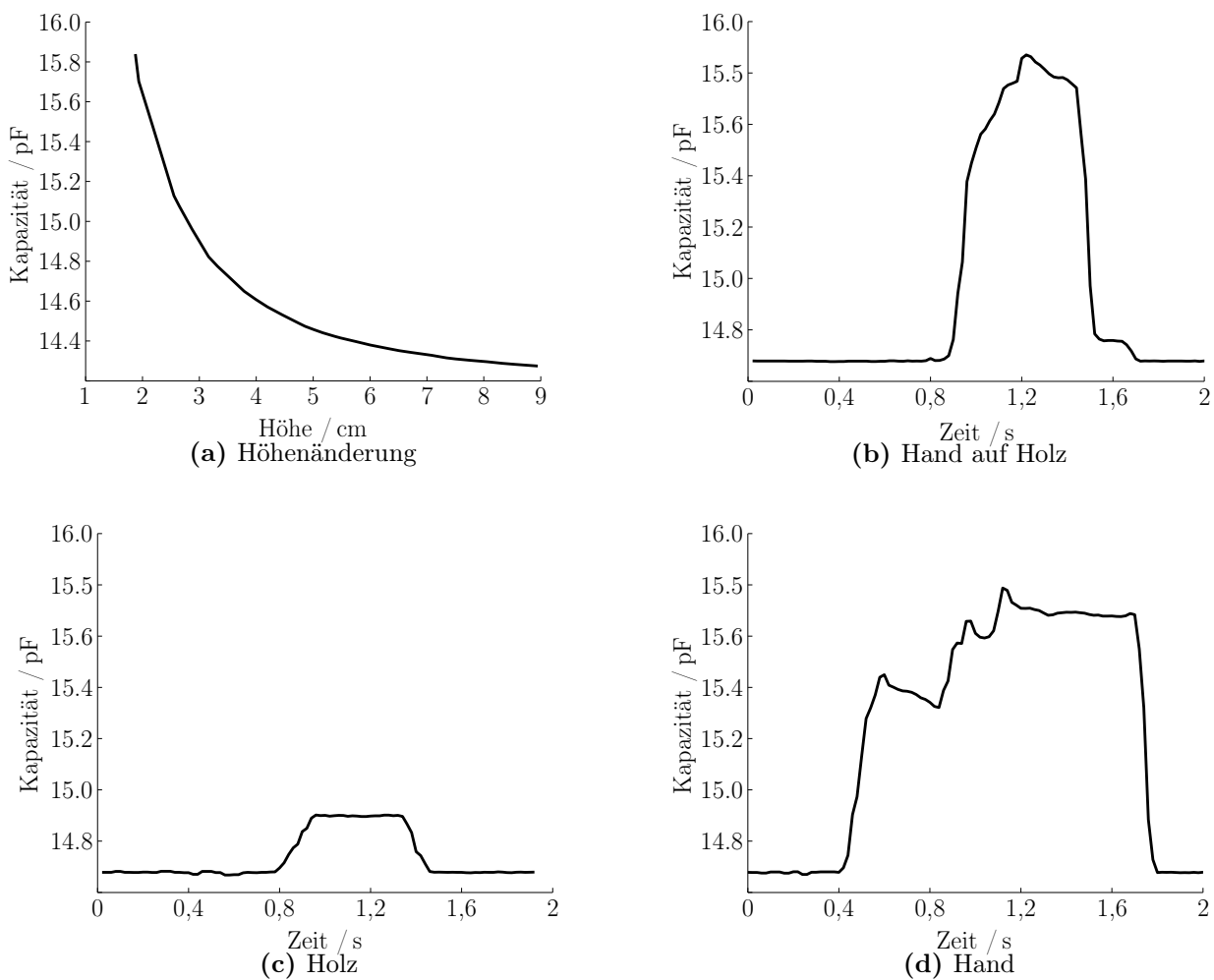


Abbildung 2.9: Messwertverlauf kapazitiver Sensor (AD7747)

2.3.3 Auswertung von Streufeldänderungen

Für einige Messanordnungen, z. B. Näherungsschalter, ist ein Plattenkondensator (siehe Abbildung 2.7) und damit die Hauptfeldmessung ungeeignet, da der Plattenabstand nicht beliebig groß gewählt werden kann. Die Streufelder, die an den Enden der Elektroden des Plattenkondensators entstehen, sind zu klein und deren Reichweite zu gering, um eine Annäherung eines Objektes zuverlässig und rechtzeitig zu erkennen. Durch eine geschickte Anordnung der Elektroden kann jedoch eine Streufeldmessung nützlich und hilfreich sein.

Eine Möglichkeit der Elektrodenanordnung ist in Abbildung 2.10 dargestellt. Hierbei sind die Elektroden eines Plattenkondensators aufgeklappt. Ein Hauptfeld existiert beim Anlegen einer Wechselspannung direkt zwischen den Elektroden. Die aktive Zone dieser Sensorik wird durch das Streufeld gebildet. Ein in die aktive Zone eingebrachtes Objekt verändert das Streufeld, abhängig von der Permittivitätszahl des Objektes und vom Abstand des Objektes von den Elektroden.

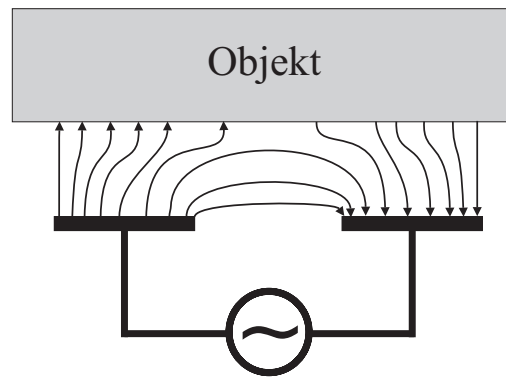
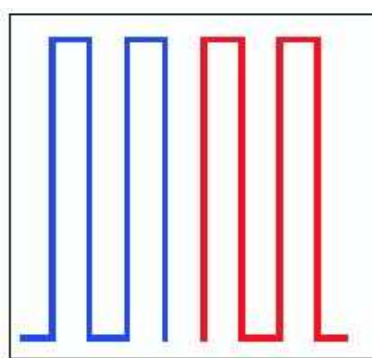
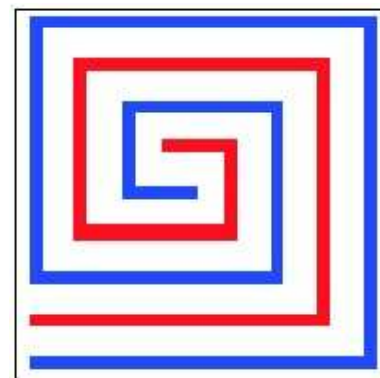


Abbildung 2.10: Feldverlauf eines aufgeklappten Plattenkondensators



(a) Erhöhte Reichweite



(b) Vergrößerte aktive Zone

Abbildung 2.11: Anordnung von Elektroden kapazitiver Sensoren

Dieser Aufbau wird in kommerziell verfügbaren, kapazitiven Näherungsschaltern verwendet [37, 86]. Die maximale Reichweite dieser Anordnung ist aufgrund des schwachen Streufeldes stark eingeschränkt und beträgt ungefähr 10 mm.

Durch Variation der aktiven Sensorflächen können auch andere kapazitive Sensoren mit unterschiedlichen Eigenschaften erzeugt werden.

Abbildung 2.11 (a) zeigt eine Anordnung von Elektroden, die auf eine Reichweite von ca. 60 mm optimiert ist und Objekte mit großer Permittivitätszahl erkennt. Die Anordnung in Abbildung 2.11 (b) hat eine vergrößerte aktive Sensorfläche von ca. 25 cm², jedoch liegt die Reichweite dieses Sensors unter 20 mm.

Eine weitere Möglichkeit der Elektrodenanordnung ist in Abbildung 2.12 dargestellt. Dabei werden zwei Drähte im Gegensatz zu zwei Platten (s. Abbildung 2.10) als Elektroden verwendet. Diese zwei parallelen Drähte sind einfach in einer Schutzlamelle zu integrieren. Zum Schutz vor mechanischer Belastung ist in der Unterseite der Lamelle eine längliche Vertiefung eingefräst, in der die Drähte verlaufen. Diese Elektrodenanordnung hat eine Reichweite von ca. 5 mm. Das Konzept der Schutzlamellen ist in Abschnitt 5.3 ausführlich erklärt.

Eine analytische Form für die Kapazität lässt sich für diese Anordnung nicht mehr angeben. Eine Simulation wäre aufgrund der veränderlichen Geometrie des

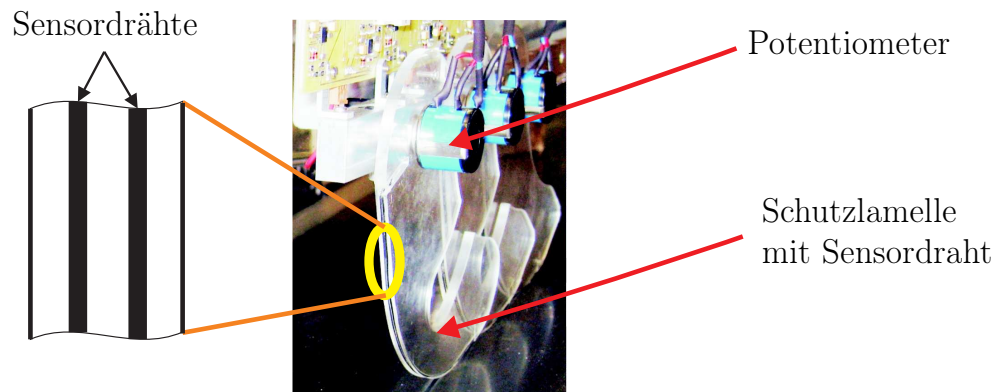


Abbildung 2.12: Schutzlamelle mit integrierter kapazitiven Sensorik

Kondensators nicht hilfreich. Außerdem können die Umgebungsparameter und -einflüsse nicht ohne weiteres vernachlässigt werden, da die Kapazität des Kondensators davon beeinflusst wird. Eine Modellierung der Kapazität ist daher nur unvollständig möglich. Durchgeführte Kapazitätsmessungen lieferten für diese Anordnung Werte im pF-Bereich.

Messung der Kapazität: Eine Alternative zur Erfassung der Kapazitätsänderung, wie diese bei der Hauptfeldmessung vorgestellt wird, bietet ein Schwingkreis als Auswerteschaltung [84]. Dabei ist der Kondensator das den Schwingkreis verstimmende Element. Diese Frequenzmodulation ist im Vergleich zu einem AD7747 wesentlich schneller. Jedoch ist ein Abgleich der Schaltung wegen parasitärer kapazitiver und induktiver Einflüsse der Bauteile und des Sensors schwierig und nicht stabil. Ursachen sind unter anderem die geringe Kapazität des Sensors, die im pF-Bereich liegt, und Schwankungen der Versorgungsspannung, die sich bei diesem empfindlichen System auf das Ausgangssignal auswirken.

Die Anforderungen an die Messschaltung sind demnach:

- Unabhängigkeit von den Schwankungen der Versorgungsspannung.
- Gewährleistung einer ausreichend hohen Empfindlichkeit.
- Schnell zu erfolgende Erfassung der Kapazitätsänderung (unter 1 ms, s. Kapitel 5).

Die oben genannten Forderungen lassen sich einfach und schnell mit einer analogen Differenzverstärkerschaltung [84], wie sie in Abbildung 2.13 zu sehen ist, erfüllen. Als Eingangssignal wird eine Sinusspannung U_e mit 40 kHz verwendet. Die komplexe Impedanz \underline{Z}_C des Sensors beträgt allgemein

$$\underline{Z}_C = \frac{1}{j2\pi fC} \propto \frac{1}{\epsilon_r} \quad (2.5)$$

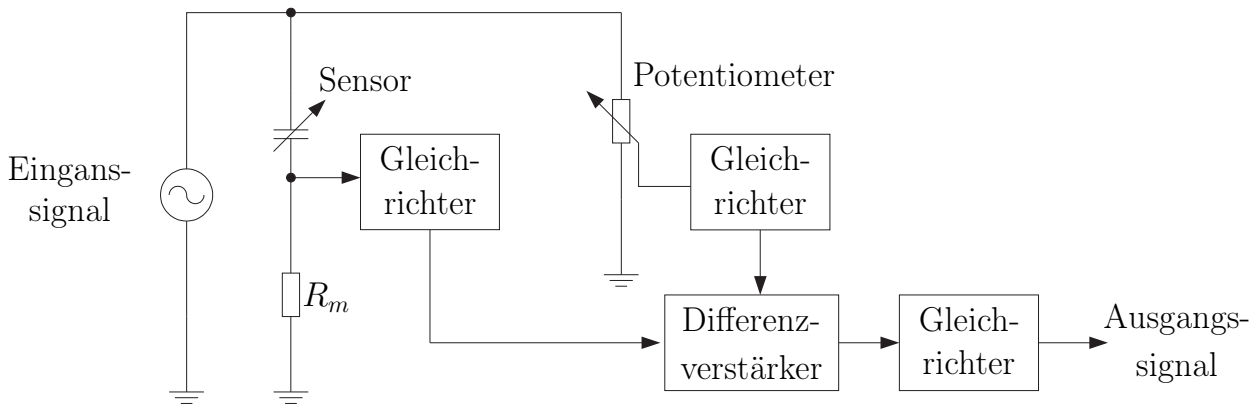


Abbildung 2.13: Messschaltung für kapazitiven Sensor

und ist damit durch eine Veränderung der Kapazität bzw. der Permittivität von außen beeinflussbar (s. Gleichung (2.3)).

Abhängig von der Art der Beeinflussung ist eine genauere Beschreibung der Sensorimpedanz \underline{Z}_C notwendig. In Abbildung 2.14 sind dazu Sensorersatzschaltbilder dargestellt. C_0 stellt die konstruktionsbedingte, nicht beeinflussbare Kapazität des Sensors dar. Im Leerlauf, d. h. wenn sich kein Objekt in Sensorreichweite befindet, und beim Einbringen von Werkstoffen wie Holz und Kunststoff, kann die Impedanz des Sensors wie in Abbildung 2.14 (a) modelliert werden. Der Kondensator C_1 ist entweder mit Luft oder mit einem Dielektrikum gefüllt. Das Dielektrikum vergrößert nun die Kapazität des Kondensators und bewirkt dadurch eine Verringerung der Sensorimpedanz, s. Gleichung (2.5).

Wird statt eines Dielektrikums ein leitender Werkstoff in Detektionsreichweite des Sensors gebracht, so entsteht dadurch eine neue Kondensatorelektrode, was einer Serienschaltung von zwei Kondensatoren gleichkommt. Dies ist in Abbildung 2.14 (b) zu sehen. Dadurch reduziert sich die Gesamtkapazität und bewirkt eine Erhöhung der Impedanz \underline{Z}_C .

In Abbildung 2.14 (c) ist ein vereinfachtes Ersatzschaltbild zu sehen, das ein Einbringen von menschlichem Gewebe in die Sensorzone darstellt. Die beiden Kondensatoren C_1 und C_2 stellen das Einkoppeln des Messsignals in das Gewebe dar. Die Impedanz \underline{Z}_{gew} repräsentiert die Leitfähigkeit und Reaktanz von menschlichem Gewebe [60] und ist reziprok von $\underline{\epsilon}_r$ abhängig, s. Gleichungen (2.4) und (2.5). Durch die Serienschaltung der Kondensatoren und das Einbringen der Gewebeampeanz erhöht sich die Gesamtimpedanz des Sensors. Diese qualitative Beschreibung der Zusammenhänge ist für das weitere Verständnis der Sensorfunktion ausreichend. Weitere Information finden sich z. B. in [60].

Wird ein konstanter Messwiderstand R_m in Reihe zum kapazitiven Sensor mit \underline{Z}_C geschaltet, so kann die Impedanzänderung des Sensors mittels des entstandenen

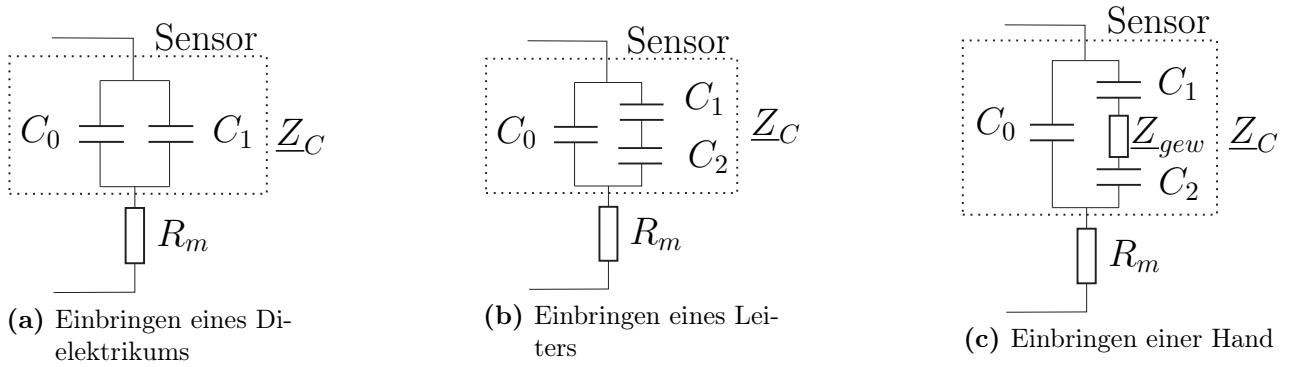


Abbildung 2.14: Ersatzschaltbild des kapazitiven Sensors

Spannungsteilers ermittelt werden. Dazu wird die Spannung am Messwiderstand U_{R_m} ausgewertet.

$$U_{R_m} = \frac{R_m}{R_m + \underline{Z}_C} \cdot U_e = \frac{1}{1 + \frac{\underline{Z}_C}{R_m}} \cdot U_e \quad (2.6)$$

Somit kann indirekt über die Spannung U_{R_m} auf die komplexe Impedanz eines Objektes geschlossen werden. Die Permittivität ϵ_r kann nicht exakt bestimmt werden, da die geometrischen Parameter des Sensors nicht analytisch ermittelt werden können und zusätzlich eine Leitwertmessung notwendig wäre. Jedoch kann qualitativ über die Impedanz- bzw. Permittivitätsänderung auf das vom Sensor erfasste Objekt geschlossen werden. Unter der Annahme einer stabilen Eingangsspannung U_e und eines festen Widerstandes R_m steigt die Spannung U_{R_m} an, wenn der Betrag von \underline{Z}_C kleiner wird, s. Gleichung (2.6). Eine mögliche Ursache hierfür ist eine Vergrößerung der Gesamtkapazität des Sensors, was auf eine Vergrößerung der Permittivität z. B. durch Holz hindeuten kann.

Die Spannung U_{R_m} fällt, wenn der Betrag von \underline{Z}_C größer wird. Dies hat z. B. eine Verringerung der Sensorkapazität oder menschliches Gewebe im Sensorbereich als Ursache.

Eine weitere Ursache für die Veränderung der Spannung U_{R_m} ist die Annäherung eines Objektes an den Sensor. Durch die Annäherung steigt der Anteil des Objektes, welcher sich im Feld des Sensors befindet. Dadurch wird ein größerer Teil des Objektes polarisiert und bewirkt dadurch eine Feldschwächung. Dies führt wiederum zu einer Impedanzänderung.

Den Effekten ist gemein, dass diese eine Amplitudenmodulation der Spannung U_{R_m} hervor rufen. Eine einfache Variante zur Demodulation ist nun eine Effektivwertberechnung mit einem Gleichrichter. Dabei geht die Information über die Phase der Spannung an R_m bezüglich der Eingangsspannung verloren. Diese ist jedoch für die Bestimmung der Impedanzänderung nicht notwendig. Des Weiteren ist eine Phasenmessung schaltungstechnisch aufwendig und nur aussagekräftig, wenn zusätzlich eine Leitwertmessung durchgeführt wird.

Die Amplitudenmodulation von U_{R_m} kann auch von Spannungsschwankungen der Versorgungsspannung oder von äußeren Einflüssen verursacht werden. Aus diesem Grund ist ein zweiter Spannungsteiler in die Schaltung integriert. Durch einmaligen Abgleich der Schaltung mittels des Potentiometers können durch den Differenzverstärker diese Fehler eliminiert werden. Außerdem werden die Spannungsänderungen an R_m , die im mV-Bereich liegen, verstärkt. Die Restwelligkeit des Signals wird durch den nachgeschalteten Gleichrichter reduziert.

Eine miniaturisierte Auswerteschaltung kann sehr nahe am Sensor angebracht werden, so dass äußere Einflüsse minimal sind. Mit diesem realisierten Schaltungstyp lassen sich keine absoluten Kapazitätswerte messen. Es können nur Impedanzänderungen festgestellt werden. Die Ursachen für diese Änderungen können folgende sein:

1. Eine Hand befindet sich in einer Entfernung von bis zu 5 mm vom Sensor. Dies bewirkt eine Impedanzänderung von über 30 % aufgrund der elektrischen Eigenschaften und der hohen Permittivitätszahl von menschlichem Gewebe.
2. Die Schutzlamelle (s. Abbildung 2.12) wurde ausgelenkt. Dies hat eine Änderung der Feldkonfiguration und der Kondensatorgeometrie zur Folge. Dadurch entstehen Kapazitätsänderungen abhängig von der Stellung der Lamelle (Positionsbestimmung der Lamelle s. unten) im Bereich von ca. 10 – 20 %.
3. Verbundwerkstoffe, die leitfähige Materialien enthalten, oder aluminiumbeschichtetes Holz werden in den Sensorbereich geschoben. Diese Impedanzänderung ist nicht von 1. zu unterscheiden. Deshalb ist eine Metallerkennung erforderlich.
4. Trockenes Holz und Kunststoffe stellen die Werkstoffe der zu sägenden Werkstücke dar. Die Kapazitätsänderung dieser Stoffe liegt bei ca. 10 bis 15 %.
5. Sehr feuchtes oder nasses Holz verursacht aufgrund des hohen Wassergehaltes ähnliche Kapazitätsänderungen, wie sie bei 1. auftreten. Holz mit hohem Restfeuchtegehalt wird jedoch nicht mit Formatkreissägemaschinen bearbeitet.

Die Fälle 2 und 3 können durch zusätzliche Maßnahmen kompensiert werden, die im Folgenden beschrieben werden. Der 5. Fall könnte nur durch eine Leitwertmessung des Objektes richtig erkannt werden. Das dafür notwendige System benötigt jedoch direkten elektrischen Kontakt zum Objekt [38], um die Leitfähigkeit des Objektes zu bestimmen. Dies lässt sich in realen Maschinenumgebungen nicht bewerkstelligen.

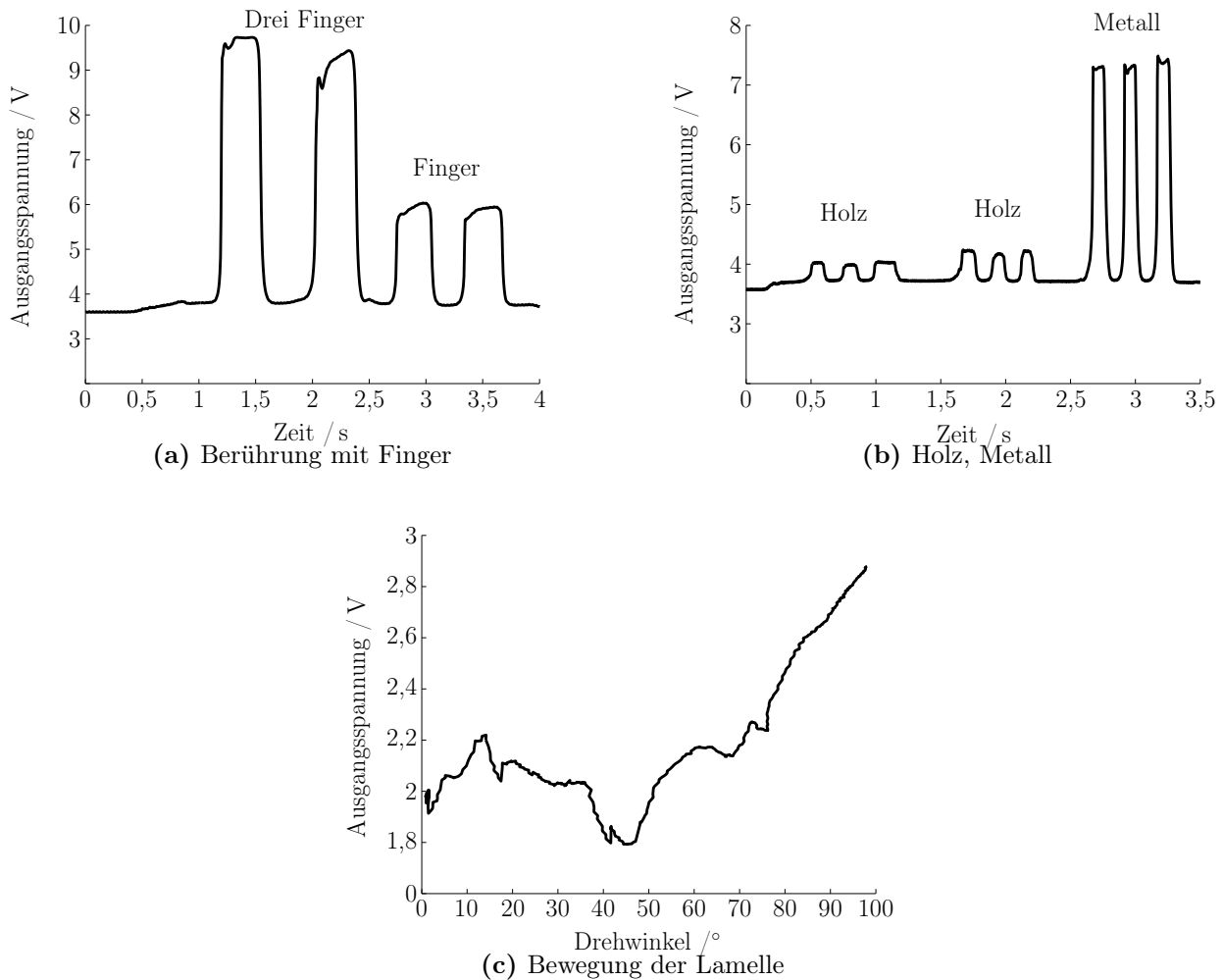


Abbildung 2.15: Messwertverlauf kapazitiver Sensor in einer Schutzlamelle

Messergebnisse

In Abbildung 2.15 (a) ist das Ausgangssignal eines kapazitiven Sensors einer Schutzlamelle dargestellt. Hierbei wurde eine fixierte Schutzlamelle zuerst zwei Mal mit drei Fingern einer Hand, danach zwei Mal mit einem Finger berührt. Gut zu erkennen ist, dass, je mehr Finger einer Hand sich im Sensorbereich befinden, die Impedanzänderung stärker und damit das Ausgangssignal höher ist. Vergleicht man das Ausgangssignal des Sensors mit Abbildung 2.15 (b), so ist gut zu erkennen, dass zwischen Metall und einem Finger nicht mehr unterschieden werden kann. Holz und Hand dagegen können gut voneinander unterschieden werden. Betrachtet man die Änderung des Ausgangssignals in Abbildung 2.15 (c), die durch eine reine Drehbewegung der Lamelle erzeugt wird, so ist ersichtlich, dass dadurch Fehler im Ausgangssignal entstehen.

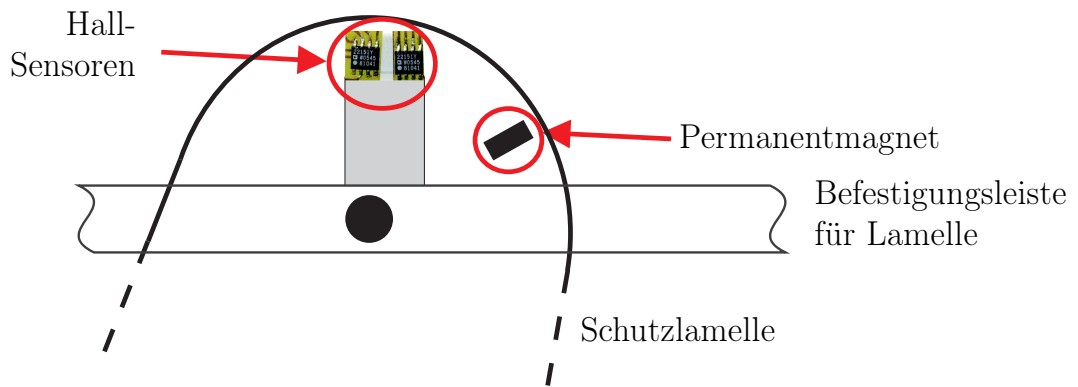


Abbildung 2.16: Messsystem für eine berührungslos arbeitende Positionsbestimmung einer Schutzlamellen

Positionsbestimmung der Schutzlamellen

Die oben vorgestellte kapazitive Sensorik liefert auch bei einer Bewegung der Schutzlamelle eine Kapazitätsänderung, die ähnlich der von Werkstoffen sein kann, s. Abbildung 2.15 (c). Wird nun eine Lamelle durch ein Holzwerkstück ausgelenkt, so ergibt sich eine deutliche Kapazitätsänderung, welche zu einer fehlerhaften Gefahrenerkennung führen könnte. Ebenso wird das richtige Sensorergebnis verfälscht, wenn eine Hand die Lamelle auslenkt. Aus diesem Grund sind die zu überwachenden Schutzlamellen über ein Potentiometer an der Schutzhaube befestigt, s. Abbildung 2.12. In Abschnitt 5.3 finden sich weitere Informationen zu den Schutzlamellen.

Die Potentiometer liefern über eine Differenzverstärkerschaltung abhängig von der Auslenkung der Lamelle einen absoluten, eindeutigen Positionswert r_{SL} . Durch eine im System hinterlegte Referenzkennlinie kann eine Korrektur der Kapazitätsänderung des Sensors erfolgen, so dass diese unabhängig von der Bewegung und der Position der Lamelle ist, s. Abschnitt 4.3.

Potentiometer sind einfach in der Handhabung, jedoch sind sie verschleißbehaftet und haben eine für die Anwendung bei Formatkreissägen ungünstige Bauform. Abhilfe schafft ein berührungslos arbeitendes Sensorsystem, das auf Hall-Sensoren zur Positionsbestimmung beruht. Hall-Sensoren basieren auf dem Hall-Effekt. Unter diesem ist das Auftreten einer quergerichteten Lorentz-Kraft bei Vorliegen einer senkrecht zu einem stromführendem Leiter gerichteten magnetischen Induktion zu verstehen. Durch die Lorentz-Kraft entsteht abhängig von der Leitergeometrie die Hall-Spannung [86], die für Sensoranwendungen ausgewertet werden kann.

In Abbildung 2.16 ist eine Skizze für ein System zur Lamellenpositionsbestimmung abgebildet. Grundlage bilden zwei an der Schutzhaube angebrachte Hall-Sensoren (Typ AD22151 von Analog Device, s. [1]) und ein auf der Lamelle befestigter Permanentmagnet. Je nach Lamellenbewegung werden die Hall-Sensoren durch den Permanentmagneten unterschiedlich angeregt. Zur Bestimmung der

Lamellenposition werden der erste, der zweite oder beide Hall-Sensor verwendet, abhängig von der Auslenkung des Lamelle.

Im Vergleich zu einem Potentiometer zur Positionsbestimmung ist der Verarbeitungsaufwand mehr als doppelt so hoch, da zwei Sensorwerte pro Lamelle digitalisiert und verarbeitet werden müssen. Der relative Fehler des Hall-Sensoren-Systems in Bezug auf ein Potentiometer liegt unter 5 % und ist damit ausreichend genau.

2.3.4 Materialidentifikation mit induktiven Sensoren

Bei der Verarbeitung von nicht-magnetischen leitenden Werkstoffen, wie z. B. Aluminiumprofilen oder aluminiumbeschichteten Holzplatten, wird die kapazitive Sensorik gestört, s. Abbildung 2.15 (b). Das leitende Material wirkt im kapazitiven Sensorfeld wie eine Mittelelektrode. Durch diese Reihenschaltung von Kondensatoren verändert sich die Gesamtkapazität der Anordnung und kann mit der oben vorgestellten Schaltung zur Kapazitätsmessung nicht mehr von der Kapazitätsänderung, die eine Hand bewirkt, unterschieden werden. Aus diesem Grund muss eine induktive Materialerkennung eingesetzt werden.

Die Erkennung leitender, nicht-magnetischer Werkstoffe wird mit Wirbelstromsensoren durchgeführt [86]. Kommerziell verfügbare Sensoren nutzen dazu z. B. eine Hartley Oszillatorschaltung [84] aus. Die Reichweite solcher Sensoren ist im Verhältnis ihrer Baugröße aufgrund der notwendigen Spulen klein.

Sensorprinzipien: Wird ein ausgedehnter elektrischer Leiter in ein Magnetfeld bewegt oder befindet sich dieser in einem sich zeitlich ändernden Magnetfeld, so ändert sich der magnetische Fluss durch den Leiter. Dieser Vorgang verursacht eine Induktionsspannung im Leiter. Da bei Raumtemperatur kein Leiter widerstandslos ist, entstehen durch diese Induktionsspannung Ströme, die so genannten Wirbelströme. Diese Ströme erzeugen nach der Lenz'schen Regel wiederum ein Magnetfeld, das dem ursprünglichen entgegengesetzt ist. Dadurch wird das ursprüngliche Feld geschwächt [26]. Je nach Leiterdicke und der Frequenz des äußeren Magnetfeldes werden die Wirbelströme an die Leiteroberfläche gedrängt. Dieser Effekt ist als Haut- oder Skin-Effekt bekannt.

Der magnetische Fluss ϕ ist die Multiplikation der magnetischen Flussdichte \mathbf{B} mit der von ihr durchsetzten Fläche \mathbf{A}

$$\phi = \mathbf{B} \cdot \mathbf{A} .$$

Die zeitliche Änderung der Flussdichte ergibt die Induktionsspannung U_{ind}

$$U_{ind} = -w \cdot \frac{d\phi}{dt}$$

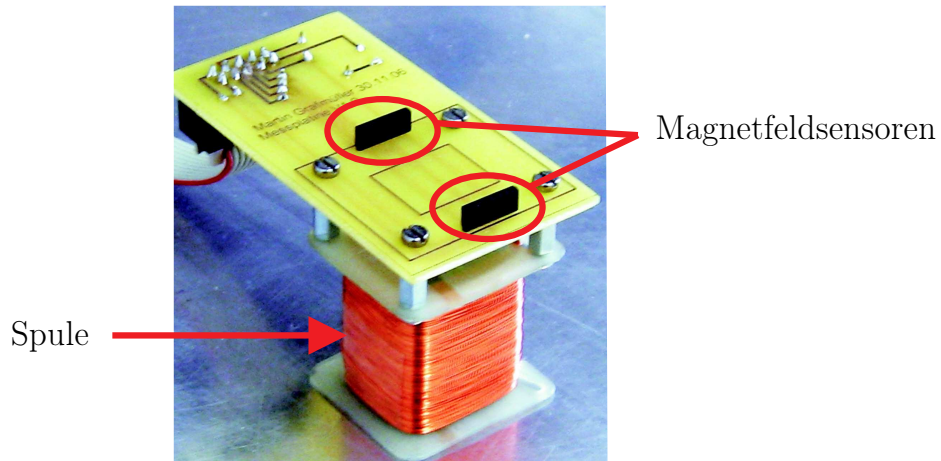


Abbildung 2.17: Versuchsaufbau des induktiven Sensors

wobei w die Anzahl der Leiterschleifen darstellt und $w = 1$ für einen ausgedehnten Leiter gilt. Bei homogenen Materialien lässt sich die Größe der Wirbelströme mit dem Ohm'schen Gesetz bestimmen.

Die Eindringtiefe δ des magnetischen Feldes in einen Leiter ist nach [58]

$$\delta = \sqrt{\frac{1}{\pi f \kappa \mu}}.$$

Dabei ist f die Frequenz des Feldes, κ die spezifische Leitfähigkeit und μ die Permeabilität des Materials. So ist z. B. für Aluminium $\kappa = 33 \cdot 10^2 \text{ S/m}$ und $\mu \approx \mu_0 = 4 \cdot 10^{-7} \frac{\text{Vs}}{\text{Am}}$. Bei einer Frequenz von $f = 1 \text{ kHz}$ beträgt die Eindringtiefe $\delta \approx 2,78 \text{ mm}$.

Direkte Messung der Magnetfeldänderung: Die entwickelte induktive Sensorik sollte möglichst kompakt sein und eine hohe Reichweite haben. Dazu wurde ein Aufbau mit einer Spule und zwei Honeywell Magnetfeldsensoren vom Typ HMC1001 [34] erstellt, s. Abbildung 2.17. Die Spule wird mit einer Rechteckspannung mit 1 kHz betrieben. Dies erlaubt eine Detektion von ca. 1 mm starkem Aluminium in einer Entfernung von ca. 6,5 cm. Die Schaltung muss vor Inbetriebnahme kalibriert werden, da diese unter anderem auch vom Erdmagnetfeld beeinflusst wird.

Die Magnetfeldsensoren sind magnetoresistive Sensoren (s. [4]) und haben eine Empfindlichkeit von 3 nT. Betrieben werden sie in einer Differenzschaltung, um äußere Störeinflüsse zu reduzieren. Die Sensoren werden zusätzlich alle 0,1 s zurückgesetzt, um eine Sättigung der Sensoren durch äußere Magnetfelder zu vermeiden.

Messergebnisse: In Abbildung 2.18 (a) ist das Ausgangssignal der Magnetfeldsensorik zu sehen. Bei 0,7 s wurde eine ca. 1 mm dicke Aluminiumplatte in

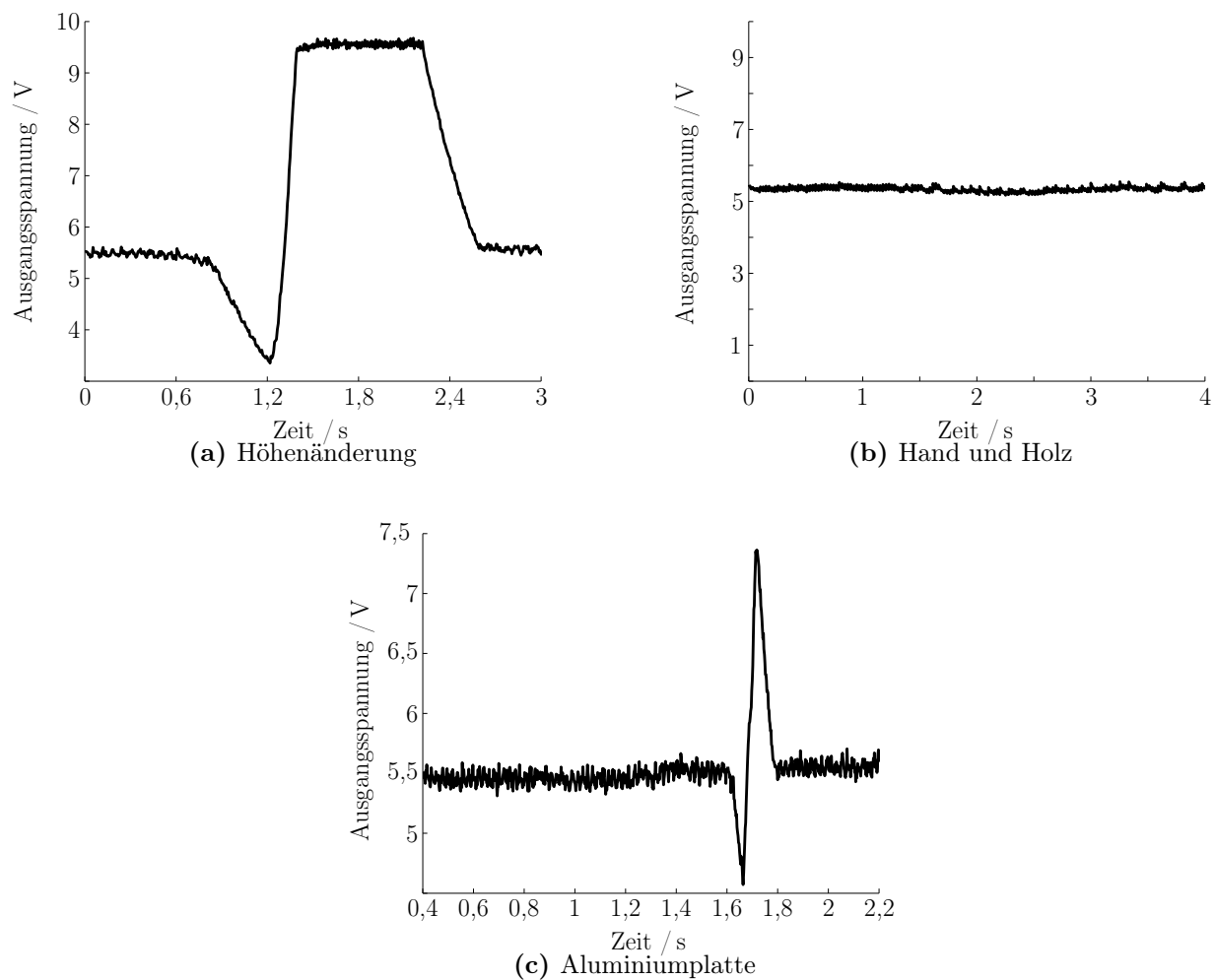


Abbildung 2.18: Messwertverlauf induktiver Sensor

einer Höhe von 6,5 cm in den Erfassungsbereich des Sensors eingebracht und anschließend auf den Sensor zu bewegt. Deutlich zu erkennen ist, dass bei einer Abnahme der Entfernung das Ausgangssignal stark ansteigt. Aufgrund der Differenzschaltung der beiden Magnetfeldsensoren fällt die Ausgangsspannung ab, bevor sie anschließend wieder ansteigt. Dieser Effekt ist auch in Abbildung 2.18 (c) zu erkennen. Hier wurde dieselbe Aluminiumplatte in einer Höhe von 4,5 cm schnell durch das Detektionsgebiet der Sensorik geführt. Werden hingegen eine Hand, Holz oder nicht-leitenden bzw. nicht-magnetische Werkstoffe in das Erfassungsgebiet gebracht, so ist der in Abbildung 2.18 (b) dargestellte Messwertverlauf zu beobachten.

2.3.5 Verwendung bei der Handerkennung

Der kapazitive Sensor zur Auswertung von Hauptfeldänderungen misst permanent die Kapazität zwischen der Verlängerung der Schutzhaube und dem Arbeitstisch, s. Abschnitt 5.3. Ändert sich diese Kapazität, so kann daraus ge-

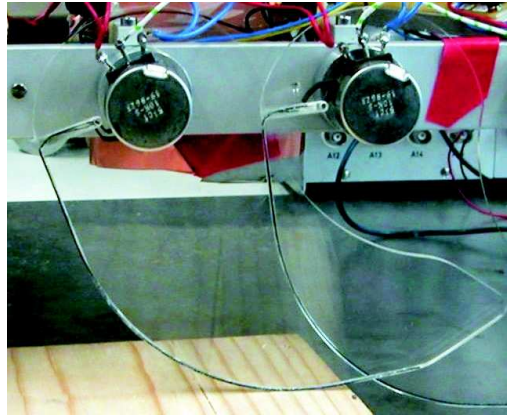


Abbildung 2.19: Holzstück hebt eine Schutzlamelle an

geschlossen werden, dass ein Objekt in das Vorfeld der Schutzhaube eingetreten ist. Aus dem absoluten Wert kann die durchschnittliche Permittivitätszahl ϵ_r für den Zwischenraum berechnet werden. Bei einer hohen Permittivitätszahl liegt die Vermutung nahe, dass es sich um eine Hand handeln könnte. Eine geringe Permittivitätszahl deutet auf Holz hin. Ähnlich arbeiten auch die kapazitiven Sensoren, die in den Lamellen eingebaut sind.

Allerdings erreicht die Permittivitätszahl bei Eintritt einer Hand ins Vorfeld nie auch nur annähernd so große Werte wie sie in Tabelle 2.1 für Wasser genannt wurden. Zwar besteht die Hand zu über 70 % aus Wasser, was zu einer großen Permittivitätszahl ϵ_r führt, allerdings füllt eine Hand nicht den gesamten Raum zwischen dem Sensor und der Arbeitsplatte aus. Der Bereich, den die Hand nicht ausfüllt, ist weiterhin mit Luft ($\epsilon_r = 1,00059$) gefüllt, und die durch den Sensorausgangswert bestimmte Materialkonstante ist nur der Durchschnittswert über den ganzen Zwischenraum.

Eine weitere Schwierigkeit für den kapazitiven Sensor in der Erweiterung der Schutzhaube ist die verstellbare Höhe der Schutzhaube. Mit ihr ändert sich der Abstand der zwei Kondensatorplatten, dem Sensor und der Arbeitsplatte. Der Abstand der beiden Kondensatorplatten fließt jedoch nach Gleichung (2.3) reziprok proportional in die Kapazität mit ein, so dass die Messergebnisse des Sensors dadurch verfälscht werden. Um den Einfluss der Höhe kompensieren zu können, wird die Höhe der Schutzhaube von einem weiteren Sensor gemessen. Der Ablauf der Kompensation wird in Abschnitt 4.3 detailliert beschrieben.

Die kapazitiven Sensoren in den Lamellen haben den Vorteil, dass sie wenige Millimeter vom Objekt entfernt sind. Dadurch ist die Durchschnittsbildung der Permittivitätszahl mit Luft vernachlässigbar.

Wie in Abbildung 2.19 zu sehen ist, dreht das Holz die Lamelle nach hinten weg. Den gleichen Effekt hätte auch eine Höhenänderung der Schutzhaube. Durch diese Drehung der Lamelle ändert sich die Feldkonfiguration des kapazitiven Sensors in den Lamellen. Dies hat wiederum eine Änderung des Streufelds zur Folge. Um nur die Änderung der Kapazität, die durch die Permittivitätszahlen ϵ_r des Hol-

zes hervorgerufen wurde, zu erhalten, muss der Einfluss der Streufeldänderung durch die Drehung der Lamelle korrigiert werden. Hierfür wird mithilfe eines Potentiometers die Stellung der Lamelle gemessen. Eine nähere Untersuchung zur Kompensation der Kapazitätsänderung durch die Lamellendrehung wird in Abschnitt 4.3 vorgenommen.

Eine Schwierigkeit für beide kapazitive Sensoren stellt das Sägen von Aluminium oder Verbundwerkstoffen dar. Wie bereits erwähnt, erhöht sich durch den metallischen Leiter im Sensorbereich die Kapazität in ähnlichem Maße, wie bei Eintritt einer Hand in den überwachten Bereich. Da eine Unterscheidung aufgrund der Bauweise der kapazitiven Sensoren nicht möglich ist, wird ein zusätzlicher, induktiver Sensor zur Metalldetektion verwendet.

2.4 Fazit

In diesem Kapitel sind einige berührungslos arbeitende Sensorsysteme vorgestellt und auf deren Eignung für Sicherheitsanwendungen hin diskutiert worden. Ausgehend von der Übersicht in Abbildung 2.1 finden folgende Systeme Verwendung zur berührungslosen Überwachung:

- **Thermopile Infrarot-Sensoren** registrieren Temperaturänderungen im Sensorbereich. Da eine Hand im Normalfall wärmer als $30\text{ }^{\circ}\text{C}$ ist, kann diese bei geringeren Umgebungstemperaturen als $30\text{ }^{\circ}\text{C}$ von einem Infrarot-Sensor erfasst werden. Jedoch ist eine eindeutige Objektidentifikation damit nicht möglich, s. Abschnitt 2.2.
- **Kapazitive Sensoren:** Basierend auf den dielektrischen Eigenschaften von Objekten lassen sich diese mit einer Hauptfeld- oder Streufeldmessung identifizieren, s. Abschnitt 2.3. Eindeutig ist dies jedoch nur mit einer parallel erfolgenden Leitwertmessung, die allerdings kontaktbehaftet arbeitet und damit ungeeignet für eine berührunglose arbeitende Gefahrenerkennung ist. Die begrenzte Reichweite der Sensoren erfordert eine dem zu untersuchenden Objekt möglichst nahe Position. Des Weiteren müssen die auf die kapazitiven Sensoren wirkenden parasitäre Umgebungseinflüsse korrigiert werden, s. Abschnitt 4.3.
- **Induktive Sensoren** ermöglichen eine eindeutige Erkennung von leitenden und magnetischen Materialien. Menschliches Gewebe kann hiermit nicht detektiert werden. Jedoch ist dieser Sensor für eine Registrierung von Verbundmaterialien aus Holz und Aluminium geeignet. Diese Materialien erzeugen bei den kapazitiven Sensoren Ausgangswerte, die von einer Hand nicht mehr zu unterscheiden sind.

- **Optischer Sensor (Kamera):** Eine Kamera besteht aus einer Matrix von Helligkeitssensoren. Durch eine geeignete Datenverarbeitung (s. Kapitel 3) ist dieses System in der Lage, eine Hand als solche visuell wahrzunehmen und damit zu identifizieren und zu lokalisieren. Der Verarbeitungsaufwand der Daten ist im Vergleich zu den anderen Sensoren um Größenordnungen höher.

Auf diesen Sensoren aufbauend wird mittels einer Informationsfusion die Gefahrenanalyse in Kapitel 4 entwickelt. Damit lassen sich die Schwächen der Sensoren mit den Stärken der anderen Sensoren beheben, um ein zuverlässiges und robustes System zur Gefahrenerkennung zu erhalten. Das Gesamtsystem ist in Kapitel 5 beschrieben und in Abbildung 1.7 dargestellt.

3 Bildverarbeitung zur Handerkennung

Systeme zur Bildverarbeitung sind hierarchisch aufgebaut [45]. Von einer Kamera ausgehend werden die Bilder bzw. Bilddaten sequentiell mit verschiedenen Bildverarbeitungsoperationen bearbeitet, um am Ende der Verarbeitungskette die gewünschten Informationen zu erhalten. In Abbildung 3.1 sind die wichtigsten Schritte der Bildverarbeitung dargestellt.

Die Bildaufnahme und -digitalisierung werden unter dem Begriff Bildgewinnung zusammengefasst. Diese erfolgt mit einer Kamera, in der beispielsweise CCD- oder CMOS-Sensorarrays zur analogen Bildaufnahme eingesetzt werden. Die analogen Bilddaten werden anschließend in der Kamera digitalisiert und über einen Datenbus, z. B. USB, FireWire oder CameraLink, an eine Datenverarbeitungseinheit übermittelt [18].

Unter Bildvorverarbeitung versteht man die Bildsegmentierung, Merkmalsextraktion sowie Glättungsfiler und Bildformatkonvertierungen. Bei einer Bildsegmentierung wird das Bild in einzelne Regionen unterteilt, um für die Weiterverarbeitung interessante Merkmale voneinander zu trennen. Eine einfache aber effektive Methode stellt die Farbsegmentierung dar. Die Merkmalsextraktion kann als Datenreduktion verstanden werden. Aus den vorhandenen Bilddaten werden nur die interessanten Merkmale *extrahiert* und an die nächsten Verarbeitungsschritte weitergeleitet. Dadurch reduziert sich der nachfolgende Verarbeitungsaufwand, da nicht mehr jedes Bildelement, ein so genanntes Pixel, weiterverarbeitet werden muss, sondern nur noch die relevanten Merkmale. Dies sind z. B. Kanten, Kreise, Texturen, etc.

Nach der Merkmalsextraktion erfolgt das Bildverstehen oder maschinelles Sehen [45]. Hierbei werden basierend auf den Merkmalen u. a. Objekte klassifiziert oder Bilder analysiert. Ähnlich dem menschlichen Sehen wird hier versucht, ein von der Kamera erfasstes Szenario zu interpretieren und zu bewerten. Dies erfolgt z. B. durch Auswerten und Vergleichen von Objektformen oder einer Objektbewegung und dessen Verfolgung.

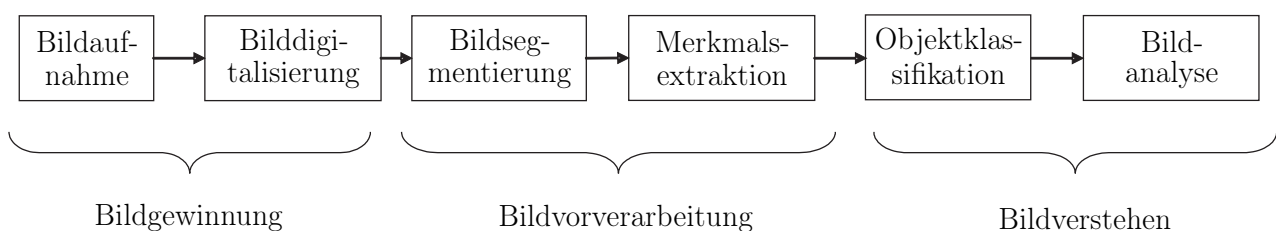


Abbildung 3.1: Schritte der Bildverarbeitung

Eine Kamera stellt einen berührungslosen Sensor dar (s. Abschnitt 2.1) und ist, in Kombination mit einer Bilddatenverarbeitung, in der Lage Gefahrenbereiche zu überwachen. Dieses Einsatzgebiet der Bildverarbeitung wird in diesem Kapitel näher betrachtet und ein System zur Handidentifikation und deren Lokalisation in einem überwachten Bereich vorgestellt. Im folgenden Abschnitt wird der entwickelte Ablauf der erforderlichen Bilddatenverarbeitung präsentiert. Anschließend werden die einzelnen Verarbeitungsschritte näher erklärt. Die erzielten Ergebnisse finden sich am Ende des Kapitels.

3.1 Verfahren zur Handidentifikation und -lokalisierung

Die hier vorgestellte Bilddatenverarbeitung ist Teil eines Gesamtsystems zur Gefahrenerkennung, s. Abbildung 1.7 und Kapitel 5. Die Ergebnisse dieses Teilsystems können in der Informationsfusion zur Gefahrenanalyse weiter verwendet werden, s. Abschnitt 4.4.4. Die Resultate sind hier eine erkannte und lokalisierte Hand sowie deren Bewegungsparameter.

Der Ablauf der allgemeinen Bildverarbeitungskette aus Abbildung 3.1 bildet die Basis der Bildverarbeitungsschritte, die zu einer Handerkennung mittels Bildverarbeitung notwendig sind. In Abbildung 3.2 sind die einzelnen Schritte dargestellt, deren Abfolge auch die Reihenfolge der nachfolgenden Abschnitte wiedergibt.

Am Anfang der Kette steht die Bildgewinnung mit einer Kamera. Abweichend von Abbildung 3.1 folgt als nächstes die Bewegungserkennung, die eigentlich als Bildsegmentierung aufzufassen ist. Der Vorteil hierbei ist, dass nur Bildausschnitte, in denen eine Bewegung erkannt wird, für die weitere Bilddatenverarbeitung betrachtet werden müssen. Dadurch lassen sich schnelle Objektbewegungen oder Reflexbewegungen einer Hand zuverlässiger erkennen und analysieren. Des Weiteren entsteht eine Gefahrensituation für die Hand nur aufgrund von Bewegung, weshalb hier eine Bewegungserkennung an erster Stelle der Bildverarbeitung stehen muss. Schlimmstenfalls liefert die Bewegungssegmentierung das gesamte Bild als *ROI (Region of Interest)*. Dies ist der Fall, wenn z. B. ein den kompletten Überwachungsbereich abdeckendes Werkstück durch diesen hindurch geschoben wird.

Die Bildvorverarbeitung gliedert sich in eine Farbraumtransformation, Farbsegmentierung, Glättungs- und Rauschfilter sowie einer Kantendetektion. Diese Verarbeitungsschritte dienen dazu, das System robust gegen Helligkeitsschwankung zu machen, Hände von Holz zu unterscheiden sowie die Datenmenge und Bildrauschen zu reduzieren.

Nach der vorangegangenen Bildsegmentierung mit einem Kantenbild als Ausgangsdatum folgt die Merkmalsextraktion. Eine modifizierte Kreis-Hough-Trans-

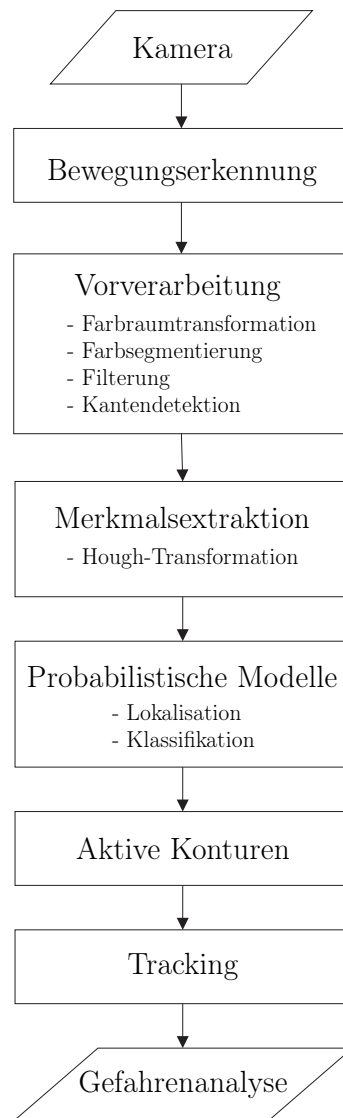


Abbildung 3.2: Ablaufdiagramm der Handerkennung

formation liefert, basierend auf dem Kantenbild, die Koordinaten von potentiellen Fingerspitzen als Merkmale. Die Fingerspitzen werden dabei als Halbkreise betrachtet.

Ausgehend von den Koordinaten der potentiellen Fingerspitzen wird mit verschiedenen probabilistischen Modellen eine Identifikation und Lokalisation einer evtl. vorhandenen Hand versucht. Probabilistische Modelle erlauben eine flexible Betrachtung der Handgeometrie, d. h. die Abstände und Beziehungen der einzelnen Modellparameter zueinander sind nicht starr, sondern erlauben Variationen innerhalb bestimmter gültiger Grenzen. Daher sind diese Modelle unabhängig von Schwankungen der Handgeometrie, die durch unterschiedliche Anwender hervorgerufen werden.

Wird eine Hand erfolgreich erkannt und lokalisiert, so erfolgt im nächsten Verarbeitungsschritt die Initialisierung einer Aktiven Kontur. Diese bildet die Umrisse einer Hand nach. Die Aktiven Konturen erlauben eine robuste Objektverfolgung

bzw. Handverfolgung. Der Rechenaufwand ist im Vergleich zu einer Merkmalsverfolgung geringer. Für das *Tracking* (engl. verfolgen) von Objekten bzw. einer Hand wird ein Partikelfilter eingesetzt.

Mithilfe der aus den Verarbeitungsschritten erzeugten Informationen wird abschließend eine Gefahrenanalyse des aktuell aufgenommenen Szenarios durchgeführt. Dabei werden die Handidentifikation und -lokalisierung wie auch die Ergebnisse der Handverfolgung ausgewertet.

3.2 Bildgebendes System

Die Bildaufnahme ist der erste Schritt der Bilddatenverarbeitung und erfolgt mittels einer Kamera. Diese ist wiederum mit einer Datenverarbeitungseinheit verbunden. Für eine zuverlässige Bildaufnahme ist eine gute Umgebungsbeleuchtung erforderlich, da es sonst zu einer Über- bzw. Unterbelichtung kommt und die Bildaufnahme damit nicht mehr zu gebrauchen ist.

3.2.1 Kamera und Objektive

In Abbildung 3.3 ist ein Kamerasystem mit allen notwendigen Komponenten dargestellt. Eine gute Beleuchtung eines Szenarios ist für eine zuverlässige Bildverarbeitung unerlässlich. Darauf wird im folgenden Abschnitt 3.2.2 eingegangen. Das Objektiv einer Kamera hat die Aufgabe, das vom aufzunehmenden Objekt reflektierte Licht zu bündeln und auf den Bildsensor der Kamera zu projizieren. Des Weiteren dient ein Objektiv zur Fokussierung eines Bildes. Über die Objektivblende wird die Lichteinstrahlung auf den Bildsensor eingestellt, um eine Überbelichtung zu vermeiden.

Die Kamera selbst besteht hauptsächlich aus einem Bildsensor, Signalverstärker, Analog-Digital-Wandler (AD-Wandler) und einer Schnittstelle zu einer Datenverarbeitung. Der Bildsensor ist bei Standard-Industriekameras ein Halbleitersensor basierend auf CCD- oder CMOS-Technologie. Für Farbkameras wird eine so genannte Bayer-Matrix verwendet. Dabei ist der Bildsensor mit einem schachbrettartigen Farbfilter überzogen, der aus Sensorelementen für blaue, grüne und rote Lichtanteile besteht, um diese Spektralfarben einzeln zu erfassen [46]. Als Schnittstellen zur Bilddatenübertragung von einer Kamera zu einem Rechner werden z. B. USB, FireWire (IEEE 1394) oder CameraLink verwendet.

Bei der Wahl der Kamera ist auf das Anwendungsgebiet wie auch auf die Anforderungen der Anwendung zu achten. So werden z. B. in der Qualitätskontrolle von elektrischen Schaltkreisen Monochromkameras mit einer Bildfrequenz von mehreren hundert Bildern pro Sekunde eingesetzt, da nur auf fehlende Bauteile hin geprüft wird. Die hohe Bildfrequenz ist erforderlich, damit der Prüfprozess dem Fertigungsprozess folgen kann. Die Bildverarbeitung ist hierbei auf die Erkennung

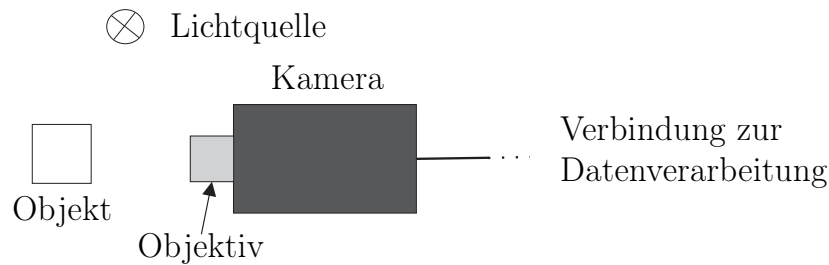


Abbildung 3.3: Kamerasystem

einfacher Muster optimiert und der Raum zur Überprüfung der Schaltkreise von äußeren Lichteinflüssen abgeschirmt und zusätzlich optimal beleuchtet. Dadurch lassen sich hohe Bildfrequenzen bei einer gleichzeitig zuverlässigen Datenverarbeitung erreichen.

Ein Kamerasystem für eine MMS (Abschnitt 1.3.3) benötigt keine derart hohen Bildfrequenzen, jedoch sind bei dieser Anwendung die Reduktion von Störeinflüssen sowie eine zuverlässige Handerkennung und Handstellungserkennung wichtig. Dies erfordert eine höhere Rechenleistung der Datenverarbeitung als im vorherigen Beispiel.

Für ein Handerkennungssystem zur Gefahrenanalyse ist ein Kompromiss zwischen hoher Bildfrequenz und zuverlässiger Handidentifikation und -lokalisierung erforderlich. Die hohe Bildfrequenz ist notwendig, da eine Hand in einem begrenzten Gebiet auch bei Handgeschwindigkeiten bis 2 m/s erfasst werden muss. Die zuverlässige Handidentifikation erfordert jedoch eine hohe Rechenkapazität des Computers zur Bilddatenverarbeitung.

In Abschnitt 5.2 wird der Bereich (*Vorfeld 1*), den eine Kamera überwacht, mit den Abmessungen 38 cm x 28,5 cm festgelegt. Um eine Hand sicher zu erkennen, muss diese mindestens 10 cm in den Überwachungsbereich hineinreichen, d. h. im ungünstigsten Fall verbleiben 18,5 cm bis zur ersten Schutzlamelle. Bewegt sich eine Hand mit einer konstanten Geschwindigkeit von 2 m/s durch diesen Bereich, so ergibt sich eine maximale Verarbeitungszeit von ca. 90 ms. Da für eine Bewegungserkennung mindestens zwei Bilder und für eine sichere Bestimmung der Bewegungsrichtung mindestens drei Bilder benötigt werden, ist eine Bildfrequenz von mindestens $f_B = 30$ fps notwendig. In diesem Fall berührt die Hand im dritten Bild die erste Lamelle. Da in diesem Beispiel die Verarbeitungszeit der Bilddaten noch nicht berücksichtigt ist, wäre unter diesen Voraussetzungen eine erfolgreiche Handerkennung und -verfolgung nicht gewährleistet. Abhilfe schafft eine höhere Bildfrequenz von ca. 50-80 fps, damit die erforderlichen Bilddaten früher zur Verfügung stehen und diese dann auch noch verarbeitet werden können. Die Anforderungen an die Verarbeitungseinheit hinsichtlich Rechenleistung und an die Verarbeitungsalgorithmen hinsichtlich Laufzeit und Zuverlässigkeit steigen dadurch an.

Um Hände von Holz und anderen Objekten zu unterscheiden, wird eine Farbsegmentierung durchgeführt, weshalb eine Farbkamera verwendet werden muss. Unter den hier angesprochenen Bedingungen kann die Kamera *A602fc* der Firma Basler [3] eingesetzt werden. Diese Farbkamera arbeitet mit einer Bildfrequenz von bis zu 100 fps und überträgt die Bilddaten über den FireWire-Bus an einen Computer. Die Auflösung dieser Kamera beträgt 640 x 480 Pixel.

3.2.2 Beleuchtung

Die korrekte Beleuchtung einer Szene ist entscheidend für die Qualität und den Erfolg der Bildverarbeitung. Nach einer Richtlinie der Berufsgenossenschaft [24] muss die Beleuchtungsstärke eines Arbeitsplatzes mit potentieller Gefährdung einer Hand mindestens 500 lx betragen und so gestaltet sein, dass kein Schattenwurf entsteht. Diese Forderungen sind durch geeignet angeordnete Leuchtröhren gut zu erfüllen. Dabei ist ebenfalls zu beachten, dass die durch den metallischen Arbeitstisch entstehenden Reflexionen minimal sind und nicht auf die Kamera treffen. Hierdurch und aufgrund von starken anderen Lichtquellen, z. B. Sonneneinstrahlung, kann es zur Überbelichtung der Kamera kommen. Dabei ist die Lichtstärke so groß, dass der Bildsensor der Kamera übersteuert ist. Die Kamera liefert dann ein homogenes einfarbiges Bild, das keine nützliche Information mehr enthält. Abhilfe bietet eine korrekte bzw. adaptierte Beleuchtung eines Szenarios, eine automatische Objektivblende oder die Anpassung der Belichtungszeit des Bildsensors. Die qualitativ beste Maßnahme ist eine korrekte Beleuchtung der Szene [45]. Dies ist auch das Mittel der Wahl, wenn eine Unterbelichtung der Szene vorliegt. Hierbei ist die Lichtstärke so gering, dass der Bildsensor keine Messwerte und damit kein Bild liefert.

Durch den Einsatz von Leuchtröhren entsteht so genanntes Flimmern, das aufgrund der 50 Hz Wechselspannung des öffentlichen Versorgungsnetzes und der dabei hervorgerufenen Helligkeitsschwankungen der Leuchtröhren entsteht. Die Helligkeitsänderungen sind dabei asynchron zur Aufnahme- bzw. Triggerfrequenz der Kamera. Eine netzsynchrone Triggerung der Bildaufnahme verhindert Flimmern komplett. Eine Kamera wird dazu über eine externe Triggerschaltung angesteuert. Diese Schaltung besteht aus einer Nulldurchgangserkennung der Netzspannung, einer Zeitverzögerung und einer Erzeugung des Triggersignals. Die Zeitverzögerung von 5 ms ist notwendig, damit die Kamera das Bild zu dem Zeitpunkt aufnimmt, bei dem die Spannung ein Maximum besitzt und die Leuchtröhre die größte Lichtabstrahlung besitzt.

Beim Einsatz einer Farbkamera ist durch Tageslichtschwankungen verursachte wechselnde Umgebungsbeleuchtung ein regelmäßiger Weißabgleich der Kamera notwendig, da durch diese Schwankungen die von der Kamera gelieferten Farbwerte variieren. Der technische Weißabgleich einer Kamera entspricht der menschlichen chromatischen Adaption. Das Ziel hierbei ist, den Wechsel der Farb-

temperatur auszugleichen, so dass weiß auch immer als weiß wahrgenommen bzw. registriert wird [46].

3.3 Bewegungserkennung in Bildsequenzen

Die Bewegungserkennung in Bildfolgen kann allgemein entweder als Segmentierung oder Merkmalsextraktion aufgefasst werden und dient der Analyse von dynamischen Vorgängen. Für die Erkennung einer Hand nach Abbildung 3.2 wird die Bewegungserkennung zu Beginn als Bildsegmentierung eingesetzt, um nur die Gebiete weiter untersuchen zu müssen, in denen tatsächlich eine aktive Handlung erfolgt. Der passive Hintergrund wird ausgeblendet und für die weitere Verarbeitung nicht mehr ausgewertet, um den gesamten Rechenaufwand der Handerkennung zu reduzieren. Eine Bewegung wird auch als Objektmerkmal angesehen, weshalb in Abschnitt 3.8 die Bewegungserkennung als Merkmalsextraktion betrachtet und zur Verbesserung der Objekt- bzw. Handverfolgung eingesetzt wird. Nachfolgend werden die Voraussetzungen und Einschränkungen der Bewegungserkennung erklärt. Aus der Literatur ist eine Vielzahl von Verfahren zur Bewegungsdetektion in Bildfolgen bekannt. In Abschnitt 3.3.2 werden aus diesem Grund einige wichtige Methoden vorgestellt und miteinander verglichen. Der in dieser Arbeit eingesetzte so genannte *Alpha-Segmentierer* zur Bewegungserkennung wird am Schluss dieses Abschnittes vorgestellt.

3.3.1 Voraussetzungen und Einschränkungen

Dynamische Veränderungen können in der Bildverarbeitung nur in Bildfolgen erkannt werden und basieren auf der Veränderung von Helligkeitswerten eines Bildelementes in aufeinander folgenden Aufnahmezeitpunkten. Diese Veränderungen werden entweder durch eine Objektbewegung oder durch Änderung der Umgebungsbeleuchtung erreicht.

Die Bewegungserkennung in der Bildverarbeitung unterliegt bestimmten Grenzen, die durch das Korrespondenz- und das Blendenproblem beschrieben werden. Es ist deshalb nicht ausreichend einzelnen Bildelemente (Pixel) zu überwachen, sondern es muss das gesamte Bild betrachtet werden.

Grauwertänderung

In Abbildung 3.4 (a) ist eine Bildfolge von acht zeitlich aufeinander folgenden Einzelbildern dargestellt. Für den menschlichen Betrachter ist eine Bewegung des weißen Punktes von der linken auf die rechte Seite zu sehen. Betrachtet man den qualitativen Verlauf eines Bildelementes in Abbildung 3.4 (b), so ist zu erkennen, dass sich der Grauwert nur zu den Zeitpunkten t_1 und t_3 ändert, ansonsten aber

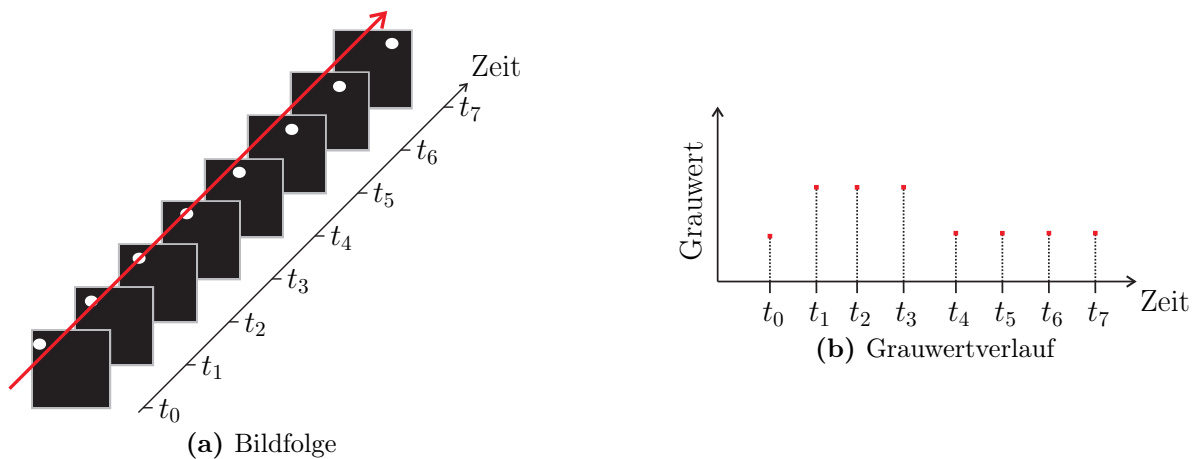


Abbildung 3.4: Grauwertverlauf eines Bildelementes über der Zeit

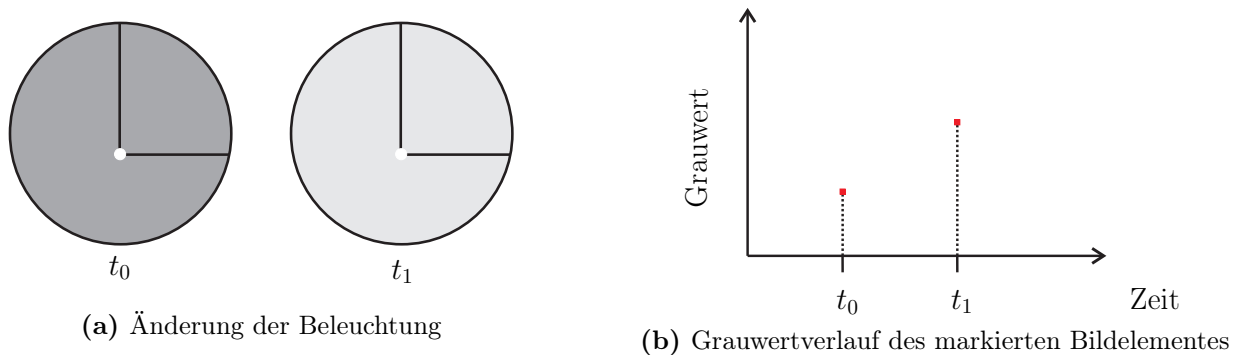


Abbildung 3.5: Grauwertänderung aufgrund von Beleuchtungsänderung

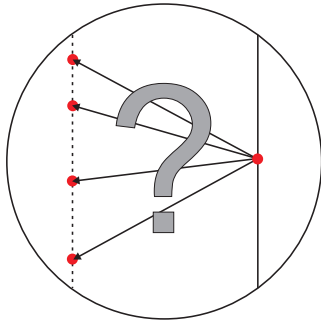
konstant bleibt. Für eine Bewegungserkennung müssen also alle Bildelemente betrachtet werden.

Beleuchtungsänderung

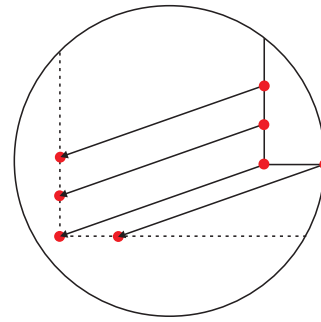
Eine Grauwertänderung eines Bildelementes tritt auch auf, wenn sich die Beleuchtung der Szene verändert. Fälschlicherweise kann hier auf eine Bewegung in der Bildfolge geschlossen werden, obwohl keine wahre Objektbewegung vorliegt. Wie in Abbildung 3.5 (a) dargestellt ist, findet keine Verschiebung des markierten Punktes statt, sondern es erfolgt eine Aufhellung des Hintergrundes. Der Grauwertverlauf dieses Bildelementes ist in Abbildung 3.5 (b) zu sehen.

Korrespondenzproblem

Das Korrespondenzproblem tritt auf, wenn eine eindeutige Zuordnung korrelierter Bildpunkte in zwei zeitlich aufeinander folgenden Bildern nicht mehr möglich ist. Somit ist eine Bewegung nicht mehr zu erkennen. In Abbildung 3.6 (a) ist ein Bildausschnitt mit einer vertikalen Kante zu sehen, die im darauf folgenden Bild verschoben worden ist. Dabei stellt die durchgezogene bzw. gepunktete Li-



(a) Mehrdeutigkeiten an einer Kante



(b) Eindeutig bestimmbar an einer Ecke

Abbildung 3.6: Blendenproblem [45]

nie die Kante im ersten bzw. im zweiten Bild dar. Es ist nicht möglich, eindeutig zu bestimmen, wohin das markierte Bildelement sich bewegt hat. Dieses spezielle Problem wird als *Blendenproblem* bezeichnet. Beinhaltet der Bildausschnitt eine Ecke, so kann der Verschiebungsvektor eindeutig bestimmt werden, s. Abbildung 3.6 (b). Damit ist es möglich, auch für alle anderen Punkte im Bild die neue Position zu bestimmen.

Ein weiteres Korrespondenzproblem tritt auf, wenn sich eine große Anzahl ähnlicher Objekte in einem Bild befindet und diese sich bewegen. Ein Beispiel hierfür ist die Geschwindigkeitsschätzung einer mit kleinen Partikeln durchsetzten strömenden Flüssigkeit. Sind dabei die Partikelabstände kleiner als die Verschiebungsvektoren, so treten Mehrdeutigkeiten auf. Eine mögliche Abhilfe ist die Verkürzung des Zeitintervalls zwischen zwei Aufnahmezeitpunkten [45].

3.3.2 Methoden im Überblick

Zur Bestimmung der Objektbewegung in Bildfolgen ist eine Vielzahl von Ansätzen und Techniken bekannt. Die im Folgenden beschriebenen Methoden bieten einen Überblick.

Differenzbildverfahren

Eine schnelle und einfache Methode zur Bestimmung von Bewegung ist das Differenzbildverfahren [45]. Dabei werden die Grauwerte von zeitlich nacheinander aufgenommenen Bildern subtrahiert und der Absolutbetrag davon gebildet. Als Ergebnis erhält man Grauwertdifferenzen. Ist die Differenz größer null bzw. größer als ein bestimmter Schwellwert, so wird dies als Bewegung interpretiert.

Diese Verfahren sind in dieser Form störanfällig, da bereits bei leichten Vibrationen der Kamera falsche Pixelpaare voneinander subtrahiert werden oder bei Schwankungen der Umgebungsbeleuchtung es zu Grauwertvariationen kommen kann. In beiden Fällen werden die Grauwertänderungen als Bewegung falsch interpretiert.

In Abschnitt 3.3.3 wird ein Verfahren vorgestellt, das auf dem Differenzbildverfahren beruht, aber durch eine Hintergrundmodellierung wesentlich robuster ist.

Optischer Fluss

Als Maß für Grauwertänderung kann der optische Fluss herangezogen werden. Dieser wird als Geschwindigkeitsvektorfeld aufgefasst. Voraussetzung für den zuverlässigen Einsatz des optischen Flusses ist eine gleichbleibende konstante Intensität der Beleuchtung [45]. Aus dieser Annahme heraus wird der optische Fluss definiert zu

$$\frac{\partial b}{\partial t} + \mathbf{v} \nabla b = 0, \quad (3.1)$$

mit b als Bild einer Sequenz und dem Geschwindigkeitsvektor $\mathbf{v} = \left(\frac{dx}{dt}, \frac{dy}{dt} \right)$.

Für die Bewegungserkennung muss Gleichung (3.1) für jeden Bildpunkt berechnet werden. Da ein Bild zweidimensional ist, kann diese Gleichung nicht eindeutig nach \mathbf{v} aufgelöst werden. Um dieses unterbestimmte Gleichungssystem trotzdem lösen zu können, wird z. B. in [62] angenommen, dass der optische Fluss innerhalb eines Gebietes konstant ist. Aus der unterbestimmten Differentialgleichung, wird ein überbestimmtes System, das mittels eines Least-Squares Schätzers gelöst werden kann.

Wavelet-Transformation

Bei einer Approximationsstufe von j sind für die Durchführung der Wavelet-Transformation [12] mit dem Haar-Wavelet 2^j zeitlich aufeinander folgende Bilder nötig, da sich die Anzahl der Koeffizienten von Stufe zu Stufe halbiert. Bei einer Bildfolge mit 2^j Bildern der Breite x und Höhe y ergeben sich $x \cdot y$ Grauwertfunktionen der Länge 2^j . Für jede dieser Grauwertfunktionen wird eine Wavelet-Transformation bis zur gewählten Stufe durchgeführt (Multiraten-Filterbank). Das Ergebnis der Hochpassfilterung sind j Graustufenbilder, die im Folgenden als Detailbilder bezeichnet werden.

Um in einem ersten Schritt eine Aussage zu erhalten, ob in dem jeweiligen Detailbild Bewegung stattgefunden hat, wird der Mittelwert μ sowie die Standardabweichung σ des jeweiligen Detailbildes bestimmt. Ausgehend von dem Grauwertmittelwert jedes Detailbildes wird ein Pixel nur dann als bewegt eingestuft, wenn es außerhalb des Intervalls von $\mu \pm 3\sigma$ liegt. Ist die Abweichung kleiner, wird der Verlauf des Pixels als verrauscht eingestuft.

Auf diese Weise werden die Bilder binarisiert, so dass bewegte Pixel als solche markiert werden. Für jedes Pixel der drei Detailbilder kann ein dreidimensionaler

Entscheidungsvektor gebildet werden, in dem festgehalten wird, ob ein Pixel in einem Detailbild markiert ist oder nicht. Bei keiner Markierung in dem Entscheidungsvektor wird das Pixel als unbewegt eingestuft, bei einer oder zwei Markierungen wird es als verrauscht angenommen und nur wenn drei Markierungen vorhanden sind gilt das Pixel an dieser Stelle als bewegt [66].

Dieser Ansatz erlaubt eine rauscharme und zuverlässige Bewegungserkennung auch bei Beleuchtungsschwankungen. Der Rechenaufwand und Speicherbedarf ist jedoch enorm. Der benötigte Speicher kann durch eine Auflösungsreduktion mittels Gauß-Pyramide verringert werden [45]. Wird vor der Wavelet-Transformation eine Kantendetektion durchgeführt, reduziert sich der Rechenbedarf der Multi-raten-Filterbank abhängig von der Anzahl der gefundenen Kanten. Durch diese beiden Maßnahmen kann die Rechenzeit um etwa zwei Drittel gesenkt werden.

Phasenmethode

Durch Fourier-Transformation ist es möglich ein Bild in den Frequenzraum zu transformieren. Schwankungen in der Beleuchtung spiegeln sich hauptsächlich in der Amplitude und geringfügig in der Phase wider. Deshalb ist diese Methode sehr robust gegenüber Beleuchtungsänderungen, da die Bildinformationen in der Phase und nicht in der Amplitude enthalten sind [45].

Ausgehend von der Phase einer 1D-Welle

$$\Phi(x, t) = kx - \omega t$$

erhält man durch räumliche und zeitliche Ableitung die Wellenzahl k und die Kreisfrequenz ω . Aus dem Verhältnis dieser beiden Größen kann die Geschwindigkeit $v = \frac{\omega}{k}$, ähnlich der Gleichung für den optischen Fluss in Gleichung (3.1), bestimmt werden.

Korrelationsmethode

Dieses Verfahren bestimmt mithilfe der Kreuzkorrelationsfunktion von Grauwertfunktionen die Ähnlichkeit von zwei Bildern in einem bestimmten Suchbereich. Bei völlig ungleichen Mustern ist der Kreuzkorrelationskoeffizient null, bei totaler Ähnlichkeit ist er eins. Ziel ist es, diesen Koeffizienten zu maximieren, wodurch auf die Verschiebung der beiden Bilder zueinander geschlossen werden kann. Aus der Verschiebung und dem zeitlichen Abstand der beiden Aufnahmen kann jetzt die Geschwindigkeit bestimmt werden, mit der sich das Objekt bewegt. Das Hauptproblem der hier vorgestellten Methode ist der hohe Rechenaufwand, der zur Maximierung des Kreuzkorrelationskoeffizienten für jeden Bildausschnitt nötig ist [47].

Fazit

Das Differenzbildverfahren eignet sich für eine schnelle Erkennung von Grauwertänderungen. So werden hierbei auch fälschlicherweise Beleuchtungsschwankungen als Bewegung interpretiert. Ein genaueres Verfahren zur Bewegungsbestimmung ist das Prinzip des optischen Flusses, jedoch sind hier konstante Beleuchtungsverhältnisse erforderlich.

Die Wavelet-Transformation und die Phasenmethode sind robust gegenüber Beleuchtungsschwankungen. Der Berechnungsaufwand und Speicherbedarf ist bei beiden Verfahren hoch, weshalb diese sich für eine schnelle Bewegungserkennung nicht eignen. Ähnlich verhält es sich mit der Korrelationsmethode, die zusätzlich auf Variationen der Beleuchtung reagiert.

Für die Handerkennung (s. Abbildung 3.2) wird im ersten Schritt eine Bewegungserkennung mit dem Ziel eingesetzt, nur die Gebiete, in denen Bewegung auftritt, weiter untersuchen zu müssen. Eine schnelle Bewegungserkennung ist deshalb notwendig, dazu bietet sich als modifiziertes Differenzbildverfahren der Alpha-Segmentierer an.

Wird in einem Gebiet aufgrund von Helligkeitsschwankungen oder eines sich langsam ändernden Hintergrundes ein Gebiet als bewegt interpretiert, bedeutet dies keine Einschränkung bzw. Zuverlässigkeitsreduktion des Systems. Im nächsten Schritt der Datenverarbeitung wird dies sofort erkannt und der Fehler behoben.

3.3.3 Alpha-Segmentierer

Die Verwendung eines beim Differenzbildverfahren üblichen statischen Hintergrundmodells ist die Ursache für die geringe Robustheit bei Beleuchtungsänderungen. Noch deutlicher wird dies bei Objekten, die im Bild bewegt wurden und dann zur Ruhe kommen. Diese Objekte werden immer als bewegt angenommen, da sie nicht an den Hintergrund angepasst werden und sich so bei der Differenzbildung stark vom Hintergrund unterscheiden.

Mit einem Differenzbildverfahren, bei dem immer nur zwei aufeinander folgende Bilder miteinander verglichen werden, ist eine zuverlässige Bewegungserkennung deshalb nicht erreichbar. Aus diesem Grund wird ein einfaches Verfahren verwendet, das in der Lage ist, den Hintergrund mit einem Hintergrundmodell anzupassen und ein Kompromiss zwischen den beiden beschriebenen Extremen darstellt [93]. Die Extremsituationen sind zum einen ein reines Differenzbildverfahren, d. h. das vorangegangene Bild wird zur Bewegungsbestimmung herangezogen. Zum anderen wird nur das initiale Hintergrundbild verwendet, so dass z. B. sich ändernde Hintergründe nicht berücksichtigt werden.

Das neue Hintergrundmodell $bg(t)$ ist

$$bg(t) = \alpha \cdot bg(t - 1) + (1 - \alpha) \cdot b(t), \quad 0 \leq \alpha \leq 1. \quad (3.2)$$

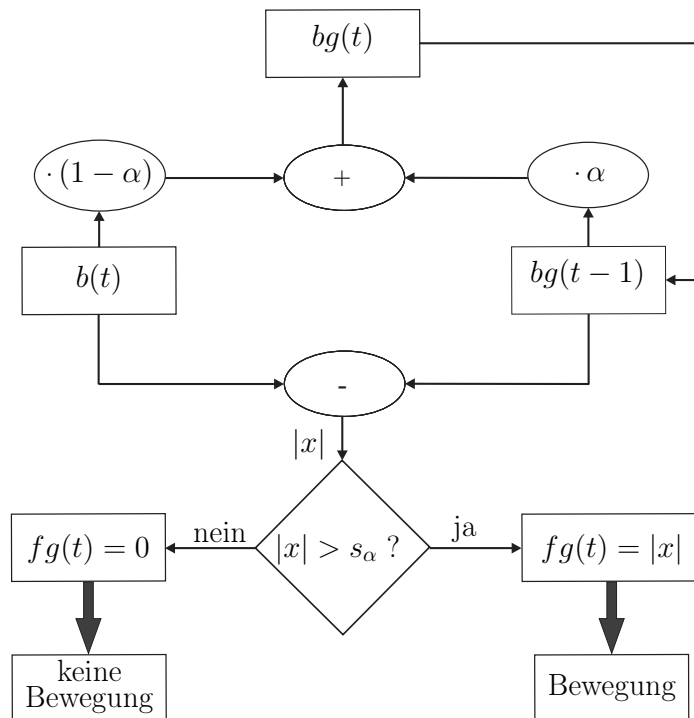


Abbildung 3.7: Ablauf des Alpha-Segmentierers

Dabei bezeichnet $bg(t)$ das Hintergrundbild zum Zeitpunkt t , $bg(t - 1)$ das Hintergrundbild zum Zeitpunkt $(t - 1)$, $b(t)$ das aktuell eingelesene Bild und α den Vergessensfaktor. Durch Gewichtung des aktuellen Bildes mit $(1 - \alpha)$ und des vorigen Hintergrundbildes mit α , werden Teile des neuen Bildes über der Zeit in das Hintergrundmodell übernommen.

Für $\alpha = 0$ wird im nächsten Schritt das zuletzt eingelesene Bild verwendet, was einem reinen Differenzbildverfahren entspricht. Im Gegensatz dazu wird für $\alpha = 1$ ein statisches Hintergrundmodell verwendet, das aus dem ersten eingelesenen Bild entsteht und über der Zeit nicht adaptiert wird.

Für die Initialisierung des Hintergrunds wird das erste eingelesene Bild $b(0)$ auch gleichzeitig als Hintergrundbild $bg(0)$ verwendet. Ab dem zweiten Bild wird nach Gleichung (3.2) für die Hintergrundmodellierung verwendet.

Der gesamte Ablauf des Alpha-Segmentierers mit Hintergrundmodellierung und Vordergrundbestimmung $fg(t)$ ist in Abbildung 3.7 zu sehen.

Bestimmung der Bewegung: Mit obigem Modell für den Hintergrund und dem Vordergrund $fg(t)$ kann auf eine potentielle Bewegung im Bildbereich geschlossen werden. Dazu wird die Differenz zwischen dem aktuell eingelesenen Bild $b(t)$ und dem Hintergrundmodell $bg(t - 1)$ berechnet und mit einem Schwellwert s_α verglichen. Das nicht aktualisierte Hintergrundmodell $bg(t - 1)$ wird zur Differenzbildung verwendet, da dieses die Veränderungen aus dem aktuell eingelesenen Bild noch nicht enthält. Der Parameter s_α wird empirisch bestimmt und ist von der Beleuchtung abhängig.

$$fg(t) = \begin{cases} 0 & |b(t) - bg(t-1)| \leq s_\alpha \\ |b(t) - bg(t-1)| & |b(t) - bg(t-1)| > s_\alpha \end{cases}$$

Ist die Differenz größer als die Schwelle s_α , wird als Grauwert der Absolutwert der Differenz an die Stelle des betrachteten Pixels geschrieben. Dadurch ist die Intensität ein Maß dafür, wie stark sich der Vordergrund vom Hintergrund unterscheidet. Der als bewegt erkannte Bereich wird mit einem Rechteck in den Bildern gekennzeichnet, dem so genannten ROI. Dessen Koordinaten stehen dann den weiteren Verarbeitungen zur Verfügung.

Für $\alpha = 0,8$ erhält man einen Alpha-Segmentierer mit einer langsamen, kontinuierlichen Adaption des Hintergrundes. Dies ist für eine robuste Bewegungserkennung ausreichend. Der Schwellwert s_α wird automatisch angepasst und ist abhängig von der Bildhelligkeit.

3.4 Bildvorverarbeitung

Die sich an die Bewegungserkennung anschließende Bildvorverarbeitung führt weitere grundlegende Bildverarbeitungsoperationen durch, s. Abbildung 3.2. Diese sind die Bildsegmentierung mit der Farbsegmentierung, die Reduktion von Bildrauschen und die Kantendetektion mit einem gradientenbasierten Verfahren. Die einzelnen Verfahren werden in diesem Abschnitt vorgestellt, wobei der Schwerpunkt auf der Eignung der Verfahren für eine robuste Handerkennung liegt.

3.4.1 Bildsegmentierung

Unter einer Bildsegmentierung versteht man allgemein die Zusammenfassung gleichartiger Pixel bezüglich eines bestimmten Merkmals. Dementsprechend existiert eine Vielzahl von unterschiedlichen Segmentierungsalgorithmen, die in vier Kategorien eingeteilt werden können. Dies sind die pixel-, kanten-, regionen-, modellbasierten Verfahren.

Nach [91] gelten für eine Segmentierung vier Regeln, die besagen, dass jedes Pixel genau zu einer Region gehört. Für jede der entstandenen Regionen gilt ein Einheitlichkeitskriterium, das vom Anwendungsfall abhängt. Die Regionen sind maximal, d. h. die Segmentierung ist vollständig. Die letzte Regel besagt, dass jedes Segment ein zusammenhängendes Gebiet bildet.

Als Merkmale für die Segmentierung eignen sich z. B. die Farben der Bildpunkte, statistische Parameter des Gesamtbildes, Texturparameter, Objektkanten und Bewegungsvektorfelder. Die Bewegungserkennung im Abschnitt 3.3 kann ebenfalls als Segmentierung aufgefasst werden.

Die Qualität der Segmentierungsergebnisse ist abhängig von der Beleuchtung der Szene. So führt eine inhomogene oder eine im Bild wechselnde Beleuchtung zu Segmentierungsfehlern, die durch weitere Verfahren bzw. durch nachfolgende Bildverarbeitungsschritte reduziert bzw. ausgeglichen werden müssen. Auch kann es zur Über- bzw. Untersegmentierung kommen und damit ein Segmentierungsergebnis unbrauchbar werden lassen.

Methoden im Überblick

Im Folgenden werden die vier Kategorien der Bildsegmentierungsalgorithmen vorgestellt. In [45, 27] finden sich weitere Informationen zu den einzeln angesprochenen Verfahren.

Pixelorientierte Segmentierung: Bei diesem Verfahren wird für jeden einzelnen Bildpunkt eine Entscheidung getroffen, ob er zu einem bestimmten Segment gehört oder nicht. Dazu sind bildpunktweise definierbare Merkmale wie Helligkeit, Farbe oder Bildpunktdifferenzen notwendig. Die Klassentrennung bzw. Entscheidung wird mit statistischen Analysemethoden durchgeführt. Der Vorteil dieses Verfahrens ist der einfache Ansatz und die schnelle Berechenbarkeit. Bekannte Verfahren sind die Farbsegmentierung, histogrammbasierte Segmentierung und clusterbasierte Segmentierung.

Kantenorientierte Segmentierung: Die kantenorientierte Segmentierung nutzt Kanten oder genauer die Übergänge zwischen Objekten oder Objekt und Hintergrund als Merkmal. Bei stark strukturierten Objekten, d. h. Objekte mit vielen Texturen, wie Linien, innerhalb der Objektgrenzen, liefert dieses Verfahren ein übersegmentiertes Bild. Für die Handerkennung ist dieser Ansatz daher nicht sinnvoll, da eine Hand wie auch Holz starke Texturen aufweisen.

Regionenorientierte Segmentierung: Bei dieser Art der Segmentierung wird die Homogenität im Inneren einer Region als Merkmal verwendet. Die Betrachtung bei der Segmentierung liegt hier nicht ausschließlich bei einem Bildelement sondern auch auf dessen Nachbarpixel. Somit wird die Zusammenhängigkeit als ein entscheidendes Kriterium der Segmentierung implizit mitbetrachtet. Die Verfahren zur regionenorientierter Segmentierung werden in agglomerative und divisive Verfahren unterteilt.

Agglomerative Verfahren verwenden Wachstumsprozesse zur Bildung von Regionen. Hierbei wird ein Startpixel gewählt, von dem aus die Nachbarpixel überprüft werden. Erfüllen diese bestimmte Kriterien, so gehören diese Nachbarpixel zur Region des Startpixels. Treffen die Kriterien nicht zu, muss ein neuer Startpunkt festgelegt werden. Die Schwierigkeit dieser Verfahren ist, dass Regionen *auslaufen*

können, d. h. bei nicht geschlossenen Objekträndern kommt es zu einer Untersegmentierung. Beispielverfahren sind das *Region-Growing-Verfahren* oder das *Saatpunktverfahren*.

Bei divisiven Verfahren erfolgt die Regionenbildung durch Zerteilungsprozesse. Als Ausgangspunkt wird das ganze Bild betrachtet und solange in Partitionen zerlegt, bis die Partitionen innerhalb homogen sind. Danach werden Partitionen, die gleiche Kriterien erfüllen, als zu einer Region gehörend betrachtet. Das *Split-and-Merge-Verfahren* ist ein Beispiel hierfür.

Die hier genannten Verfahren eignen sich für eine schnelle Bildsegmentierung zur Handerkennung nicht, da das Risiko der Über- bzw. Untersegmentierung aufgrund der schwankenden Beleuchtung hoch ist. Ebenfalls steigt der Rechenaufwand, wenn das Verfahren zur Übersegmentierung neigt, was vor allem bei inhomogenen Objekten und Hintergründen der Fall ist.

Modellbasierte Segmentierung: Die bisher beschriebenen Verfahren verwenden zur Segmentierung lokale Informationen. Durch die Hinzunahme von a priori Wissen z. B. über die mögliche Form eines zu segmentierenden Gebietes, kann die Zuverlässigkeit gesteigert werden. Als Beispiel dieser modellbasierten Segmentierung sei für gerade Objekte die Geraden-Hough-Transformation erwähnt.

In diesem Zusammenhang wird deutlich, dass die Grenzen zwischen einer Bildsegmentierung und einer Merkmalsextraktion fließend sind. In Abschnitt 3.5 wird die Hough-Transformation zur Merkmalsgewinnung eingesetzt, dabei aber spezifischer angewandt als dies bei der Segmentierung der Fall ist.

Farbsegmentierung

Die Hautfarbe wird als Standardmerkmal für optische MMS verwendet (s. Abschnitt 1.3.3) und dient dort zur einfachen Lokalisation der Hand bei homogenen, nicht-hautfarbenen Hintergründen. Um die Zuverlässigkeit dieser Handerkennung zu erhöhen, sind eine Vielzahl von verschiedenen Hautfarbenmodellen in Kombination mit dafür geeigneten Klassifikationsalgorithmen entwickelt worden [49].

Einfache Systeme bedienen sich einer Farbraumtransformation und einer Komparatorfunktion. Diese besitzen dadurch eine geringe Berechnungszeit, sind aber nicht robust gegenüber Beleuchtungsschwankungen. Die Wahl des Farbraumes und der Komparatorfunktion bzw. des Klassifikationsverfahrens ist ebenfalls abhängig vom geographischen Einsatzort. Die unterschiedlichen Ethnien dieser Erde besitzen genetisch bedingt unterschiedliche Hautfarben.

Die Farbsegmentierung ist ein pixelorientiertes Verfahren, bei dem für jedes Bildelement eine Entscheidung hinsichtlich der Farbe erfolgt. Für die Handerkennung hat sich eine Farbsegmentierung im *HSV-Farbraum* (**H**ue-**S**aturation-**V**alue) als geeignet herausgestellt, da damit eine gute Vorder-/ Hintergrundseparation mög-

lich ist und die Farben können unabhängig von der Umgebungsbeleuchtung dargestellt werden, d. h. die Farbwerte sind kaum von der Beleuchtung abhängig. Die verwendete Kamera stellt die Bilder im RGB -Farbraum dar. Eine Farbraumtransformation vom RGB - in den HSV -Farbraum ist deshalb nach folgender Vorschrift erforderlich [27]

$$\begin{aligned}
 V &= \max(R, G, B) \\
 S &= \begin{cases} \frac{V - \min(R, G, B)}{V} & , V \neq 0 \\ 0 & , \text{sonst} \end{cases} \\
 H &= \begin{cases} 60^\circ \cdot \frac{G - B}{S} & , V = R \\ 60^\circ \cdot \frac{B - R + 2}{S} & , V = G \\ 60^\circ \cdot \frac{R - G + 4}{S} & , V = B \end{cases} .
 \end{aligned}$$

Die Komparatorfunktion bzw. die Entscheidungsfunktion für ein Bildelement erzeugt aus dem Hue- und Saturation-Kanal des HSV -Bildes ein Binärbild \mathbf{F} , in dem jedes als Hautfarbe klassifizierte Pixel eine logische eins repräsentiert. Die Helligkeitsinformation (Value) wird dabei vernachlässigt. Die Berechnung erfolgt pixelweise.

$$F_{yx} = \begin{cases} 1 & , H_{yx} \in C_H \wedge S_{yx} \in C_S \\ 0 & , \text{sonst} \end{cases}$$

Die Mengen C_H und C_S enthalten die als Hautfarbe zu klassifizierende Farbwerte für den H- bzw. S-Kanal. Diese Grenzwerte sind unter anderem von der verwendeten Kamera sowie von den Farbtönen der Umgebungsbeleuchtung abhängig. Die Intervallgrenzen werden empirisch aus einer Reihe von Musteraufnahmen ermittelt, so dass schwankende Lichtsituationen die weitere Verarbeitung nicht signifikant beeinflussen.

3.4.2 Reduktion von Bildrauschen

Bei der Aufnahme von Bildern mit einer digitalen Kamera werden dem Originalbild verschiedene Rauschsignale überlagert, die die Bildqualität verschlechtern und damit die weitere Bildverarbeitung beeinflussen. Das Rauschen entsteht durch die Halbleiter-Bildsensoren (Dunkelrauschen, thermisches Rauschen), die Auslese-Verstärker (Ausleserauschen), die Digitalisierung der analogen Bildsignale (Digitalisierungsrauschen) und die Zufallsverteilung der auf ein Pixel treffenden Photonen (binäres oder Impulsrauschen) [46]. Der Einfluss der Rauschsignale kann mithilfe von Glättungsfiltern reduziert werden.

Glättungsfilter sind Tiefpassfilter, mit denen mittelwertfreies, hochfrequentes Rauschen reduziert werden kann. Nachteilig wirkt sich das Weichzeichnen durch

diese Filter aus, da scharfkantige Objektränder ebenfalls hochfrequente Spektralanteile besitzen. Der dazu notwendige Filterkern besteht aus einer abgetasteten Gaußschen Glocke mit einer von der Filtergröße abhängigen Varianz. Dieser Filtertyp wird Gaußsches oder Binominalfilter genannt.

Die Tiefpassfilterung versagt bei binärem Rauschen. Für diesen Fall werden Medianfilter eingesetzt. Hierbei wird ein Fenster definierter Größe über die Bildmatrix geschoben. Die Pixelwerte innerhalb der Filtermaske werden in eine Liste geschrieben und der Wertigkeit nach geordnet. Das Element, welches in der Mitte dieser Reihenfolge steht, wird als Filterwert an die entsprechende Stelle im Ursprungsbild zurückgeschrieben. Da sich bei fehlerhaften Pixeln deren Wert stark von den Werten seiner Umgebung unterscheidet, ist es höchst unwahrscheinlich, dass dieser als Filterwert zurückgegeben wird. Somit wird der fehlerhafte Pixelwert durch den mittleren Wert seiner Umgebung ersetzt. Die Bildschärfe bleibt bei dieser nichtlinearen Filterung erhalten [45].

3.4.3 Kantenerkennung

Lokale Helligkeitsschwankungen, lokale Helligkeitsübergänge sowie lokale Farbänderungen werden in der Bildverarbeitung als *Kanten* bezeichnet. Diese Änderungen der lokalen Helligkeit markieren Objektbegrenzungen und Texturen, so dass Informationen über die Form und die Lage der Objekte aus dem Bild gewonnen werden können. Zu einer Kanten-Übersegmentierung kann es kommen, wenn stark texturierte Objekte oder Hintergründe im Bild vorkommen.

Grundlagen

Das Ziel der Kantenerkennung ist allgemein die Erfassung von Grauwertänderungen bei gleichzeitiger Unterdrückung von konstanten Werten. Dazu eignen sich folgende Verfahren, die allesamt Nachbarschaftsoperatoren sind.

Lokale Varianz: Die Varianz ist ein Maß für die Abweichung von Messwerten untereinander. Zur Kantendetektion mittels lokaler Varianz wird nacheinander für jedes Pixel eine Schätzung der Varianz in einer Umgebung um das jeweilige Zielpixel berechnet. Die entstehende Varianzmatrix ist auch ein Maß für die Kantestärke. Der Nachteil dieses Verfahrens besteht darin, dass der Rechenaufwand hoch ist. Für jedes Pixel muss der Mittelwert und die Varianz im betrachteten Gebiet bestimmt werden.

Morphologische Kantendetektion: Dieses Verfahren eignet sich für Binärbilder, wie sie z. B. von einer Farbsegmentierung erzeugt werden. Durch Erosion oder Dilatation und anschließender Subtraktion vom Quellbild wird ein Kantenbild erzeugt [80]. Die Dicke der hierbei entstehenden Konturen hängt von Größe

und Form des Strukturelementes sowie der Anzahl der Erosions- bzw. Dilatationsdurchläufe ab. Ein wesentlicher Vorteil dieser Methode gegenüber gradientenbasierten Kantendetektoren liegt im relativ geringen Rechenaufwand. Nachteil ist hier allerdings, dass keine Information über die Stärke und Richtung der Kanten geliefert wird, sondern nur ein binäres Kantenbild.

Differenzielle Kantendetektion: Kanten werden allgemein durch Grauwertänderungen charakterisiert. Diese Änderungen können von Ableitungsfiltern ermittelt werden. Dabei werden Ableitungen erster und zweiter Ordnung betrachtet. Bei Ableitungen erster Ordnung stellt eine Kante einen Extremwert dar, d.h. die Kantendetektion entspricht einer Suche nach dem maximalen Betrag des Gradientenvektors. Hierfür ist es erforderlich, dass partielle Ableitungen in alle Richtungen berechnet werden müssen. Aus diesem Grund kann hierbei auch die Gradientenrichtung bestimmt werden.

Bei Ableitungsfilter zweiter Ordnung entspricht eine Kante einem Nulldurchgang des Grauwertverlaufes. Ein bekannter Vertreter ist das Laplace-Filter [45]. Diese Art der Kantendetektion ist aufgrund der zu berechnenden zweiten Ableitung rauschanfälliger als die gradientenbasierte.

Operatoren zur Kantendetektion

Für die der Vorverarbeitung nachfolgende Merkmalsextraktion und für eine deutlichere Unterscheidung zwischen einer Hand und Holz wird eine gradientenbasierte Kantendetektion verwendet.

Der Gradientenvektor einer Funktion zweier Parameter $b(x,y)$ kann aus den partiellen Ableitungen berechnet werden

$$\nabla b(x,y) = \begin{bmatrix} \frac{\partial b}{\partial x}(x,y) \\ \frac{\partial b}{\partial y}(x,y) \end{bmatrix} =: \begin{bmatrix} b_x(x,y) \\ b_y(x,y) \end{bmatrix} .$$

Die Ableitung kann bei den hier vorliegenden diskreten Bildern jedoch nur durch die Differentiation approximiert werden. Dies kann in x - und y -Richtung beispielsweise mit folgenden Filterkernen erfolgen

$$\mathbf{D}_x = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \quad \text{bzw.} \quad \mathbf{D}_y = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

Um Störungen durch Rauscheinflüsse zu verringern, wird senkrecht zur Richtung der Differentiation üblicherweise eine Gauß-Glättung vorgenommen. Dies kann in einem Schritt mithilfe des *Sobel*-Operators geschehen [45]

$$\mathbf{S}_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \mathbf{S}_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

Um eine verbesserte Rauschunterdrückung und Kantenverstärkung zu erhalten kann der *Scharr*-Operator eingesetzt werden [74]. Hiermit ist zusätzlich eine genaue und zuverlässige Berechnung der Gradientenrichtung möglich. Der Filterkern lautet für die Differentiation in horizontaler bzw. in vertikaler Richtung

$$\mathbf{R}_x = \begin{bmatrix} -3 & 0 & 3 \\ -10 & 0 & 10 \\ -3 & 0 & 3 \end{bmatrix} \quad \mathbf{R}_y = \begin{bmatrix} 3 & 10 & 3 \\ 0 & 0 & 0 \\ -3 & -10 & -3 \end{bmatrix}$$

Der Gradientenvektor ergibt sich damit zu (* bezeichnet die Faltung)

$$\nabla b(x, y) = \begin{bmatrix} \mathbf{R}_x * \mathbf{B} \\ \mathbf{R}_y * \mathbf{B} \end{bmatrix}.$$

Für die weitere Verarbeitung wird dieser Vektor schließlich noch in die Polarkoordinaten-Darstellung transformiert, um die Kantenstärke getrennt von ihrer Orientierung betrachten zu können

$$|\nabla b(x, y)| = \sqrt{b_x^2(x, y) + b_y^2(x, y)}$$

$$\vartheta(x, y) = \arctan \frac{b_y(x, y)}{b_x(x, y)}.$$

3.4.4 Ablauf der Bildvorverarbeitung

In Abbildung 3.8 ist das Ablaufdiagramm der gesamten Bildvorverarbeitung für die Handerkennung zu sehen. Unmittelbar nach der Bildaufnahme erfolgt die Bestimmung des ROI mithilfe der Bewegungserkennung. Hierbei wird der Alpha-Segmentierer eingesetzt. Aus dem so erhaltenen Bildausschnitt wird zum einen ein Grauwertbild erzeugt. Zum anderen wird das RGB-Bild in den HSV-Farbraum transformiert und anschließend mit einem Gauß-Filter geglättet. Das hierbei erzeugte Bild dient als Grundlage für die Farbsegmentierung, die ein Binärbild mit hautfarbenen Gebieten als logisch eins markiert. Das Binärbild und das Grauwertbild werden im folgenden Schritt mit dem logischen *UND*-Operator miteinander verknüpft. Durch diese Schritte erhält man ein Grauwertbild, das nur bewegte, hautfarbene Gebiete enthält. Diese Gebiete können sowohl Hände als auch der Hautfarbe ähnliches Holz beinhalten.

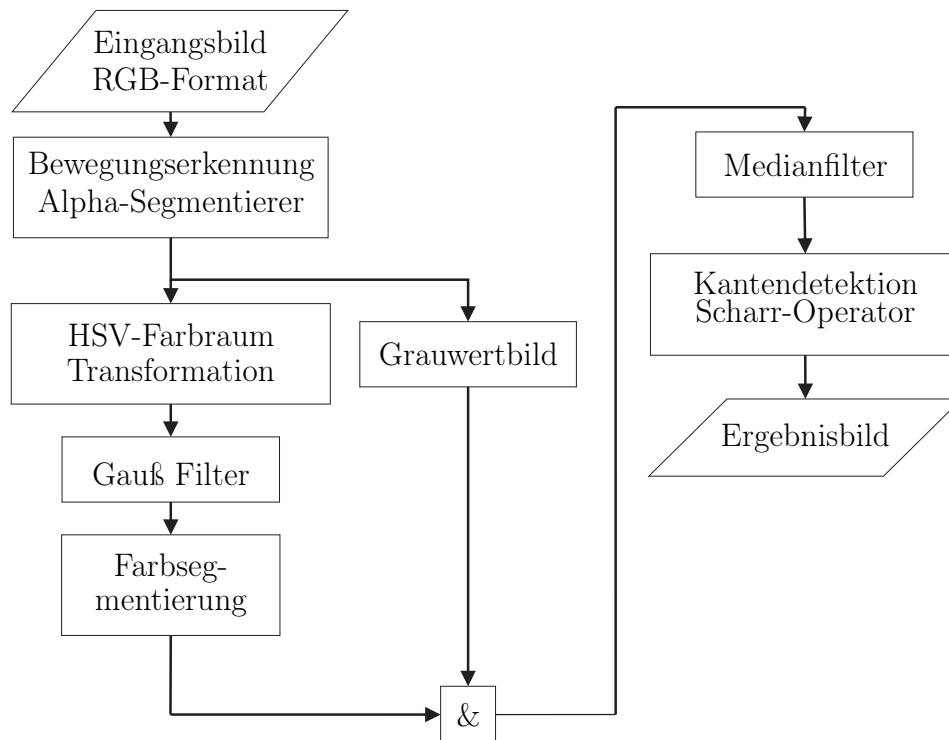


Abbildung 3.8: Ablauf der Bildvorverarbeitung

Durch die Farbsegmentierung und der anschließenden Verknüpfung mit dem Grauwertbild entstehen inhomogene Zonen innerhalb des segmentierten Gebietes, z. B. durch Fingernägel oder Hautfalten. Diese Gebiete können die Kantendetektion stören. Der Einsatz eines Medianfilters vor der Kantendetektion bewirkt, dass diese Inhomogenitäten herausgefiltert werden. Anschließend daran wird die Kantendetektion mittels Scharr-Operator durchgeführt. Als Ergebnis erhält man den Betrag der Grauwertkanten (Maß für die Kantenstärke) und die Richtung der Gradienten.

3.4.5 Ergebnis der Vorverarbeitung

Das Ergebnis der Bildvorverarbeitung ist in Abbildung 3.9 dargestellt. Zu sehen sind auch die Ergebnisse der einzelnen Verarbeitungsschritte. Der Bildhintergrund ist in diesem Szenario zweigeteilt. Die Hand liegt auf einem dunklen, gemaserten Holz (Hintergrund 1) und wird über den Maschinentisch (Hintergrund 2) durch den Kameraüberwachungsbereich geschoben. Das Holz hat in diesem Fall eine ähnliche Farbe wie eine Hand, weshalb diese nach der Farbsegmentierung nicht deutlich zu erkennen ist. Durch das Hinzufügen (entspricht der *UND*-Verknüpfung) der Informationen aus dem Grauwertbild ist es dennoch möglich, ein für die Weiterverarbeitung gutes Kantenbild zu erzeugen.

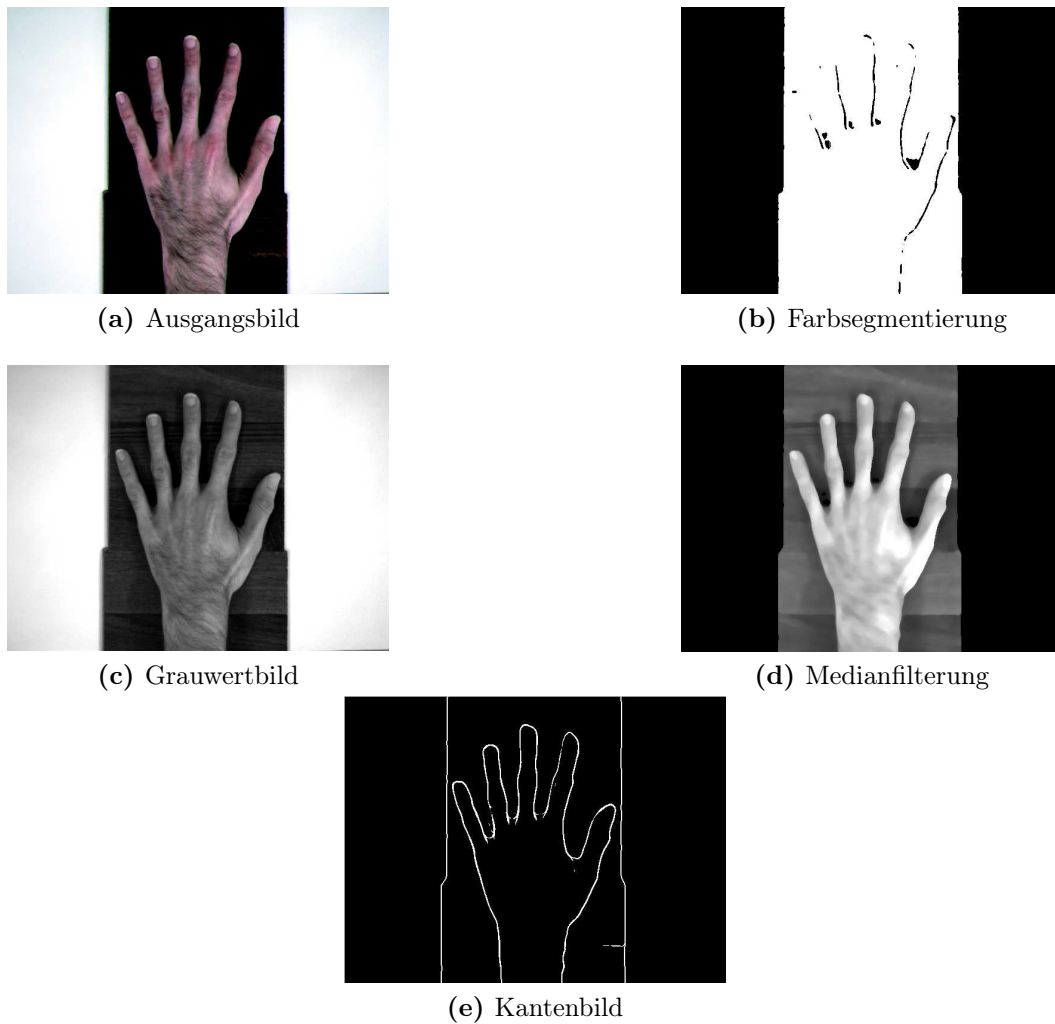


Abbildung 3.9: Ergebnisbilder der einzelnen Schritte der Bildvorverarbeitung

3.5 Merkmalsextraktion

Die Bildvorverarbeitung liefert ein segmentiertes Bild, in dem bewegte, hautfarbene Objekte enthalten sind, s. Abschnitt 3.4.1. Alle anderen Objekte bzw. Bereiche werden ausgeblendet. Dadurch erfolgt eine erste Bilddatenreduktion. Diese ist jedoch für eine Bildanalyse, wie sie für die Handerkennung notwendig ist, nicht ausreichend. Daher ist es erforderlich, aus dem vorverarbeiteten Bild die für eine Analyse entscheidenden Informationen zu extrahieren, um damit die Bilddaten weiter zu reduzieren. Diesen Verarbeitungsschritt nennt man Merkmalsextraktion.

Es gibt eine Vielzahl von Verfahren zur Merkmalsgenerierung. Einige davon sind im Folgenden aufgeführt und werden hinsichtlich ihrer Eignung für die Handerkennung betrachtet. Die Kreis-Hough-Transformation *KHT* ist ein leistungsstarkes Merkmalsextraktionsverfahren für die Detektion von Fingerspitzen, weshalb die KHT im Abschnitt 3.5.2 detailliert vorgestellt wird.

3.5.1 Verfahren zur Merkmalsextraktion

Das Ziel einer Merkmalsextraktion ist die effiziente Bereitstellung von geeigneten Merkmalen für die Bildanalyse. Im ersten Schritt sind daher die geeigneten Merkmale, die ein Objekt eindeutig charakterisieren, festzulegen. Im Anschluss daran ist ein Verfahren zur Gewinnung der festgelegten Merkmale so zu wählen, dass diese zuverlässig und schnell aus dem vorhandenen, segmentierten Bildmaterial gewonnen werden können.

Für die Bildsegmentierung werden bereits die Merkmale *Hautfarbe* und *Bewegung* benutzt, um den Suchbereich zu verringern. In Abschnitt 3.4.1 ist dies ausführlich erklärt.

Komplexe Muster werden in der Bildverarbeitung Texturen genannt. Diese stellen periodische, gleichmäßige oder einem anderen Muster folgende Oberflächeneigenschaften im Bild dar [47]. Für die Texturerkennung benötigt man verschiedene Operatoren bzw. Parameter, die diese abgebildeten Oberflächeneigenschaften beschreiben können. Dazu werden z. B. statistische Operatoren, die Musterorientierung oder die Gabor-Wavelets eingesetzt. Die Verfahren zur Texturanalyse sind jedoch rechenintensiv und benötigen außerdem eine hierarchische Auswertung. Des Weiteren arbeiten diese Verfahren nur zuverlässig, wenn sich im Bildhintergrund kein weiteres komplexes Muster befindet.

Im Folgenden werden weitere Verfahren zur Merkmalsextraktion und mögliche Merkmale vorgestellt.

Template Matching: Beim Template Matching wird das gewünschte Merkmal als Bild gespeichert. Zur Merkmalsextraktion wird dann das gespeicherte Bild (Musterbild) mit dem zu untersuchenden korreliert. Das dabei entstehende Ähnlichkeitsmaß kann mittels Schwellwertoperation ausgewertet werden. Dieses Vorgehen hat den Vorteil, dass genau das gewünschte Objekt gefunden und dessen Position ermittelt wird. Jedoch ist der Berechnungsaufwand und Speicherbedarf für eine Korrelation enorm. Des Weiteren darf das zu suchende Objekt nicht gedreht im Bild vorhanden sein. Dies würde einem anderen Musterbild entsprechen und somit würde das gewünschte Objekt nicht gefunden werden.

Um diese Nachteile zu beheben, sind weitere Verfahren in der Literatur beschrieben. Zur Beschleunigung der Berechnung lässt sich die Korrelationsfunktion als Faltung deuten, die im Frequenzbereich einer Multiplikation der fouriertransformierten Bilder entspricht. Eine weitere Methode verwendet unterschiedliche Bildauflösungsstufen und eine Grob-zu-Fein-Strategie zur genauen Lokalisation des Musterbildes [15]. Um auch das gedrehte Template zu finden, kann eine Karhunen-Loève-Transformation für das Template Matching eingesetzt werden [87].

Diese verbesserten Verfahren erlauben ein schnelleres und zuverlässigeres Lokalisieren des gesuchten Merkmals als dies mit dem ursprünglichen Verfahren mög-

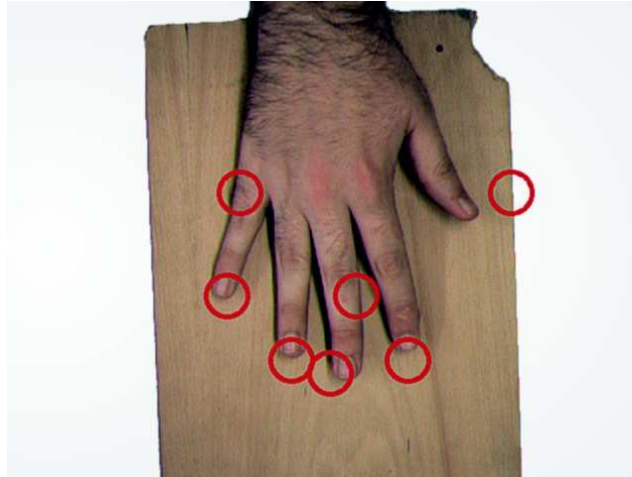


Abbildung 3.10: Ergebnis einer Gabor-Wavelet-Filterung

lich wäre. Template Matching eignet sich jedoch nicht zur zuverlässigen Handerkennung, da eine Verallgemeinerung des Templates notwendig ist, um sämtliche Finger aller möglichen Maschinenbenutzer zu erfassen und nicht nur die gespeicherten Musterbilder.

Gabor-Wavelet-Transformation: Die zweidimensionale Gabor-Wavelet-Transformation ist eine isotrope Filterung und robust gegen geringe Störungen und Beleuchtungsvariationen [22, 70]. Für die Merkmalsextraktion werden hierbei bestimmte Wellenzahlbereiche und Orientierungen betrachtet, so dass eine schnelle und robuste Wavelet-Filterung möglich ist [47]. Somit lässt sich je nach Objektkomplexität eine ausreichende Anzahl von Merkmalen erzeugen. Durch dieses Verfahren können unter anderem auch Merkmale für Objekttexturen oder Teile der Handkontur extrahiert werden.

Die Umsetzung der Merkmalsextraktion mittels Gabor-Wavelet ist umfangreich und die Berechnungen trotz Einschränkungen der betrachteten Bereiche und Orientierungen zeitintensiv. Außerdem ist durch die Einschränkungen die Genauigkeit einer Gabor-Wavelet-Transformation gering, s. Abbildung 3.10.

Generalized-Hough-Transformation: Die Hough-Transformation benötigt zur Beschreibung der zu lokalisierenden Formen parametrische Gleichungen [27], s. auch folgender Unterabschnitt 3.5.2. In vielen Fällen liegen die Formen jedoch als Menge von Bildpunkten, d. h. als Musterbild, vor. Es ist i. A. je nach Form gar nicht möglich, eine parametrische Gleichung für ein solches Musterbild aufzustellen. Hier setzt die *Verallgemeinerte-Hough-Transformation (Generalized Hough-Transform, GHT)* an [88].

Beim Entwurf einer auf der GHT aufbauenden Merkmalsextraktion wird zunächst in einer Lernphase aus einem gegebenen Musterbild der zu suchenden Form ein Modell erzeugt, die so genannten *R-Tabellen*. Die R-Tabellen erset-

zen zur Laufzeit des Systems die bei der gewöhnlichen Hough-Transformation zur Beschreibung der Form nötige analytische Gleichung. Der Parameterraum der GHT umfasst unabhängig von der Objektform die Koordinaten eines während der Lernphase definierten Objekt-Ursprungs sowie einen Drehwinkel, der die Orientierung des Objekts gegenüber seiner ebenfalls während der Lernphase definierten Normallage angibt. Dieser Parameterraum wird genau wie bei der gewöhnlichen Hough-Transformation durch ein Akkumulator-Array dargestellt, in welchem die lokalen Maxima die gefundenen Objekte repräsentieren [69].

Die wesentliche Schwierigkeit beim Einsatz der GHT, wie auch beim Template Matching oder der Gabor-Wavelet-Transformation, besteht in der Wahl der geeigneten Musterbilder und dass die zu extrahierenden Merkmale eine große Ähnlichkeit mit den verwendeten Musterbildern aufweisen müssen. Dies ist vor allem bei der Handerkennung mit den vielen verschiedenen und möglichen Hand- und Fingerformen und -größen nicht gewährleistet. Diese Verfahren sind somit besser für die Lokalisation starrer Objekte geeignet, jedoch nicht für formvariable Objekte.

3.5.2 Kreis-Hough-Transformation

Die Hough-Transformation eignet sich für die Extraktion von geometrischen Merkmalen und Objekten aus Bildern. Beispiele sind Geraden, Kreise oder Ellipsen. Diese Formen lassen sich durch analytische Gleichungen beschreiben, die als Vorwissen der Hough-Transformation zur Verfügung gestellt werden müssen. Hierdurch ist diese Art der Transformation robust gegen verrauschte Bilder, da durch Rauschen verursachte, einzelne Lücken in der Objektkontur keinen Einfluss auf das Extraktionsergebnis haben [27].

Die Hough-Transformation wurde für die Geradenerkennung entwickelt [36], ist aber für alle analytisch beschreibbaren Objekte geeignet. Im voran gegangenen Abschnitt ist die verallgemeinerte Hough-Transformation beschrieben, mit der auch komplexe Objekte durch die Hough-Transformation extrahiert werden können. Allerdings eignet sich diese Art der Transformation hauptsächlich für starre Objektformen.

Im Folgenden wird die Kreis-Hough-Transformation (KHT) näher erläutert, da diese sich für die Extraktion von Fingerspitzen als Merkmale für eine Hand eignet. Basierend auf den so gewonnenen Merkmalen kann bei der nachfolgenden Bildanalyse eine Hand in Bildsequenzen gefunden werden.

Prinzip

Für die KHT wird das Kantenbild (s. Abschnitt 3.4.3) in einen dreidimensionalen Parameterraum transformiert [69]. Dieser umfasst die Parameter Mittel-

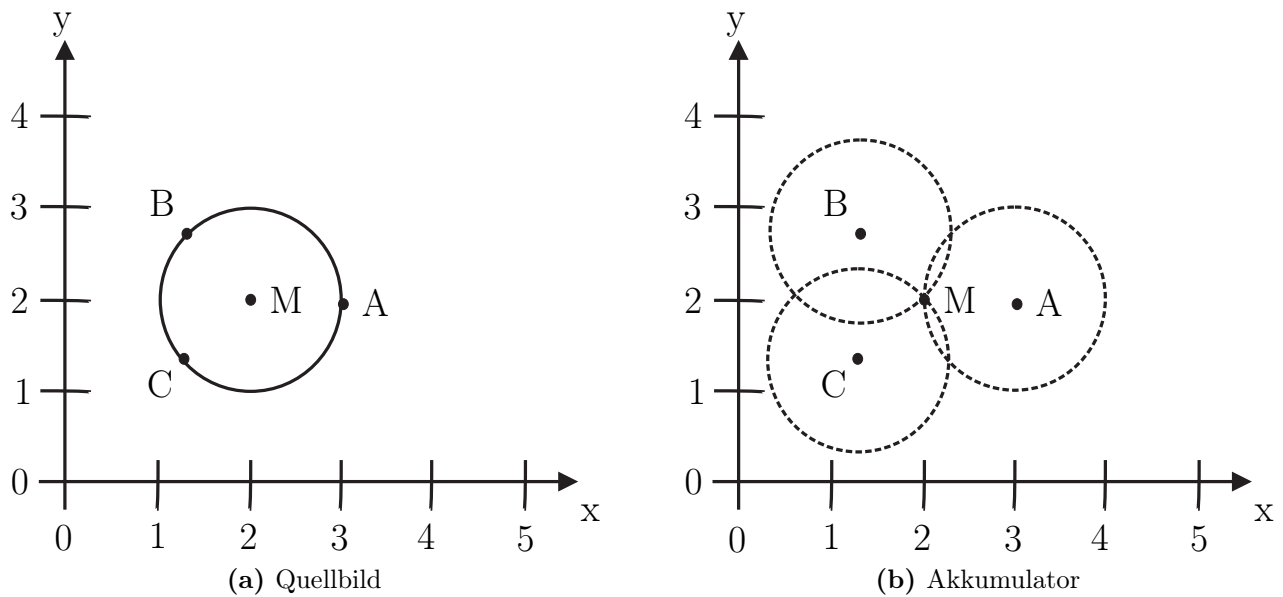


Abbildung 3.11: Hough-Transformation für Kreise

punkt (x_0, y_0) und Radius r . Die dazu notwendige analytische Beschreibung ist die Kreisgleichung

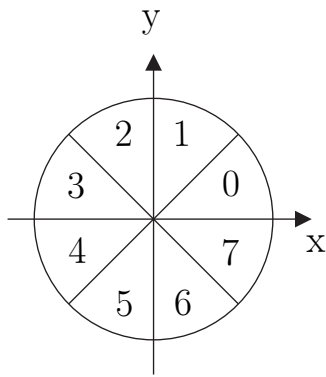
$$r^2 = (x - x_0)^2 + (y - y_0)^2 .$$

Der Parameterraum wird als Akkumulator-Array aufgefasst, wobei jede Akkumulatorzelle einem Kreis mit bestimmten Radius und Mittelpunkt entspricht.

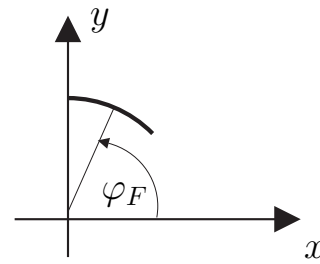
Im Verlauf der Transformation werden nacheinander sämtliche Kantenpixel des Quellbildes durchlaufen. Die Orientierung der Kanten wird hier nicht berücksichtigt. Ausgehend von jedem einzelnen Kantenpixel werden anschließend im Akkumulator-Array sämtliche Zellen inkrementiert, welche ein Objekt beschreiben, dessen Kontur dieses Kantenpixel im Quellbild treffen würde.

Zur Verdeutlichung der KHT wird zunächst nur ein bestimmter Radius betrachtet. In Abbildung 3.11 (a) ist ein zu untersuchendes Beispielbild dargestellt, das einen Kreis mit Radius $r = 1$ enthält.

Abbildung 3.11 (b) zeigt die zu $r = 1$ gehörende x - y -Ebene des Akkumulators nach einer KHT. Für die drei markierten Kantenpixel A , B und C aus dem Originalbild sind die zugehörigen Kreisbahnen gestrichelt eingezeichnet. Da sich sämtliche Kreisbahnen in M schneiden, wird die Akkumulator-Zelle mit den Koordinaten von M den größten Wert aller Akkumulatorzellen annehmen. Kreise im Originalbild werden somit durch lokale Maxima im Akkumulator repräsentiert. Nach Abschluss dieser Transformation in den Parameterraum werden im Akkumulator-Array sämtliche lokalen Maxima bestimmt, aus denen dann die Parameter aller gefundenen Kreise abgelesen werden können.



(a) Zerlegung in Oktanten



(b) Mittlerer Winkel eines Achtelkreises

Abbildung 3.12: Betrachtung der Kreisgeometrie bei der modifizierten KHT

$\varphi_F \in$	$\varphi_{F,D}$
$[0, \pi/4)$	0
$[\pi/4, \pi/2)$	1
$[\pi/2, 3\pi/4)$	2
$[3\pi/4, \pi)$	3
$[\pi, 5\pi/4)$	4
$[5\pi/4, 3\pi/2)$	5
$[3\pi/2, 7\pi/4)$	6
$[7\pi/4, 2\pi)$	7

Tabelle 3.1: Diskretisierungsstufen des Orientierungswinkels φ_F

Umsetzung

Geeignete Merkmale für eine Handerkennung stellen die Fingerspitzen dar, weshalb diese aus einem Bild zu extrahieren sind. Die Fingerkuppen können näherungsweise als Halbkreise betrachtet werden. Bei einer genaueren Betrachtungsweise stellt man jedoch fest, dass Fingerkuppen eher durch drei bis fünf Achtelkreise mit variierenden Radien beschrieben werden können. Dies bedeutet nun, dass die KHT entsprechend angepasst werden muss, d. h. der Vollkreis wird in acht Kreissegmente zerlegt, s. Abbildung 3.12 (a). Diese Zerlegung bewirkt, dass der Parameterraum auf vier Dimensionen (Mittelpunkt, Radius und Orientierung) ansteigt.

Die Orientierung ist der mittlere Winkel des Kreissegmentes im Intervall $[0, 2\pi)$ und direkt als Orientierung der Fingerspitze φ_F zu interpretieren, s. Abbildung 3.12 (b).

Für die acht möglichen Achtelkreise ergibt sich die in Tabelle 3.1 dargestellte Winkeleinteilung.

Die eigentliche Transformation kann wiederum nach dem Muster der KHT erfolgen, indem für jede der insgesamt acht diskreten Orientierungen ein eigenes, dreidimensionales Akkumulator-Array angelegt wird. Die darin zu inkrementie-

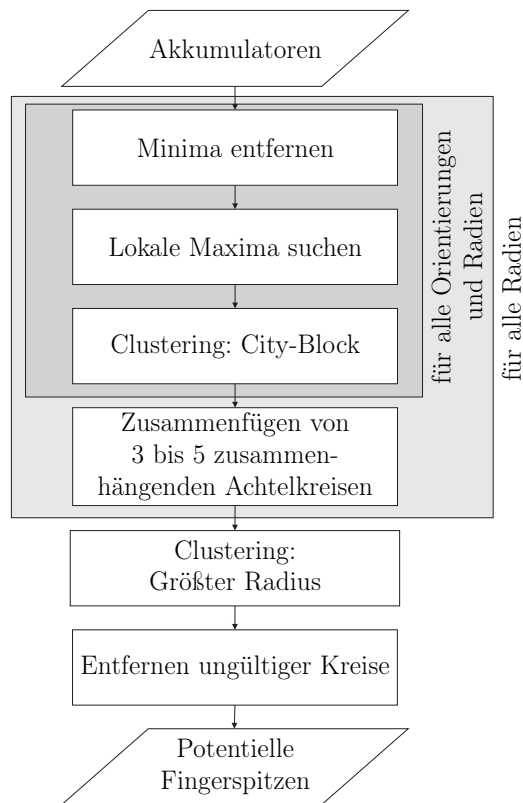


Abbildung 3.13: Ablauf der Clustering-Routinen

renden, möglichen Mittelpunkte liegen nun nicht mehr auf einem Vollkreis, sondern nur noch auf einem Kreissegment der Länge $r \cdot \frac{\pi}{4}$, einem Achtelkreis. Durch Auswertung der insgesamt acht dreidimensionalen Akkumulatoren erhält man dann Mittelpunkt, Radius und Orientierung der im Bild enthaltenen Achtelkreise.

Um die Komplexität der Berechnung weiter zu reduzieren, werden bestimmte Kreisradien festgelegt. Betrachtet man menschliche Finger, so liegt der durchschnittliche Radius einer Fingerkuppe im Bereich von $6 \leq r \leq 8$ mm, was bei gegebener Konfiguration des Kamerasystems einem Pixelradius von $13 \leq r \leq 16$ entspricht. Dies bedeutet, dass die modifizierte KHT zusätzlich noch viermal durchlaufen werden muss bzw. es werden 32 zweidimensionale Akkumulatoren benötigt.

Clustering

Die Ergebnisse der beschriebenen Hough-Transformation für Achtelkreise können noch nicht direkt zur Weiterverarbeitung eingesetzt werden, da in realen Aufnahmen viele achtelkreis-ähnliche Strukturen vorkommen, die jedoch keine Fingerspitzen darstellen. Daher ist es erforderlich, aus der Menge der lokalisierten Kreissegmente diejenigen auszuwählen, die bei geeigneter Kombination die gewünschte Näherung eines Halbkreises mit ggf. schwankendem Radius bilden. Die Bedingungen für eine solche Kombination sind zum Einen, dass die Mittel-

punkte der einzelnen Kreissegmente näherungsweise zusammenfallen und zum Anderen, dass eine passende Anzahl *nebeneinander liegender* Teilstücke vorliegt. Für einen idealen Halbkreis sind dies offensichtlich vier Achtelkreise. Es hat sich gezeigt, dass mit der Vorgabe einer Mindestzahl von drei aneinander grenzenden Kreissegmenten eine gute Extraktion der Fingerspitzen erreicht werden kann.

Ausgehend vom Kantenbild werden mit obiger modifizierter KHT Achtelkreise im Bild markiert. Durch eine Schwellwertoperation über alle Akkumulatoren werden Minima entfernt. Anschließend werden die lokalen Maxima gesucht, die den potentiellen Achtelkreisen entsprechen.

Zur Suche derjenigen Achtelkreise, deren Mittelpunkte näherungsweise übereinstimmen, bietet sich der Einsatz eines Clustering-Algorithmus zur Bewertung mehrfach detektierter Kreissegmente an. Dabei werden diejenigen Kreissegmente, für die der Abstand der Mittelpunkte einen gewissen Schwellwert unterschreitet, als *Nachbarn* bezeichnet. Die Berechnung des Abstandes erfolgt aufgrund der im Vergleich zu anderen Abstandsmaßen wesentlich effizienteren Berechnung mittels der *Cityblock-Norm* [45]

$$d(\mathbf{a}, \mathbf{b}) = |a_x - b_x| + |a_y - b_y| .$$

Diese Berechnung muss für jede Orientierung und jeden Radius durchgeführt werden. Anschließend werden die benachbarten Achtelkreise mit unterschiedlichen Orientierungen zusammengefügt, so dass näherungsweise die Fingerspitze abgebildet ist. Dies ist für jeden Radius erforderlich. Danach ist ein weiteres Clustering über die verschiedenen Radien durchzuführen, so dass diejenigen Kreissegmente mit dem größten Radius ausgewählt werden. Abschließend werden die Mittelpunkte der bisher gültigen Kreissegmente, die außerhalb des hautfarbenen Bildbereichs liegen, noch entfernt. In Abbildung 3.13 ist der Ablauf nochmals dargestellt.

Ergebnis

In Abbildung 3.14 (a) ist ein Ergebnisbild eines Akkumulators für einen Radius ($r = 13$) und eine Orientierung ($\varphi_{F,D} = 0$) dargestellt. Die Maxima in der Nähe der Fingerspitzen sind deutlich zu erkennen.

Nach dem Clustering-Prozess erhält man das in Abbildung 3.14 (b) dargestellte Bild und für die weitere Verarbeitung die Mittelpunkte der potentiellen, gefundenen Fingerspitzen.

Das hier vorgestellte Verfahren ist umfangreich. Beginnend vom Kantenbild wird die modifizierte KHT durchgeführt und in 32 verschiedene Akkumulatoren das Ergebnis für die gefundenen Achtelkreise abgespeichert. Anschließend ist noch der Clustering-Prozess abzuarbeiten. Der Rechenaufwand des gesamten Verfahrens ist abhängig von der Anzahl der potentiell gefundenen Achtelkreise. Dass dieser Aufwand gerechtfertigt ist, zeigen die erreichten, zuverlässigen Ergebnisse.

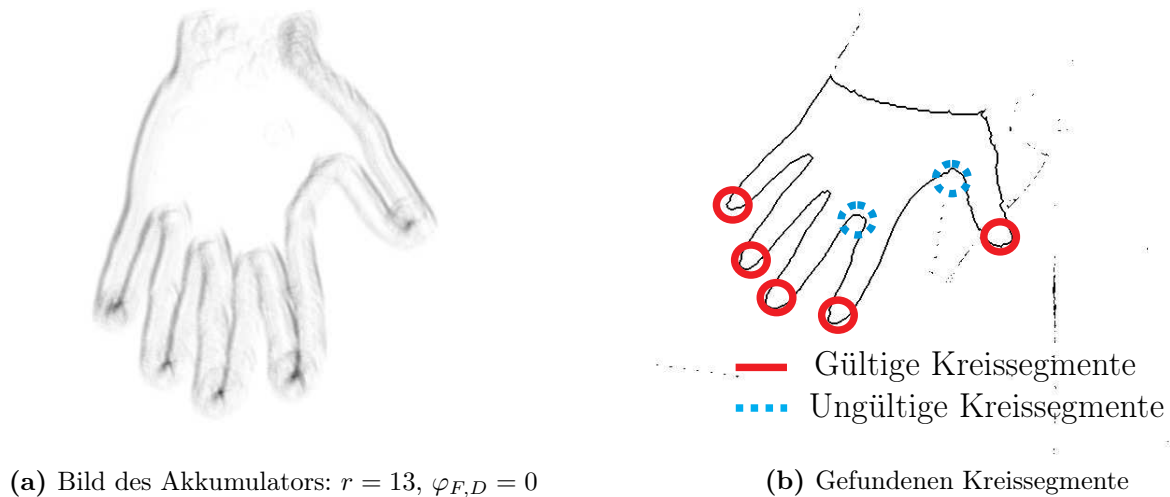


Abbildung 3.14: Ergebnis der modifizierten KHT

Basierend auf diesen ist es erst möglich, eine weitere Verarbeitung durchzuführen und potentielle Hände zu erkennen.

3.6 Lokalisation und Klassifikation mit probabilistischen Modellen

Die vorgestellte Merkmalsextraktion liefert die Position von potentiellen Fingerspitzen, s. Abschnitt 3.5. Die nun folgenden Verarbeitungsschritte haben die Aufgabe die extrahierten Merkmale auszuwerten und die aufgenommenen Szenen zu verstehen (*Bildverstehen*, Abbildung 3.1). Diese Schritte gliedern sich in drei Teile, s. Abbildung 3.2. Der erste Teil dient zur Lokalisation und Klassifikation einer Hand und zur Initialisierung einer Aktiven Kontur, s. Abschnitt 3.7. Anschließend erfolgt die Berechnung der Aktiven Kontur, die zur Präsentation von zusätzlichem Wissen über die Handform und zur Optimierung der Handverfolgung eingesetzt wird. Der letzte Teil des Bildverstehens bzw. der Bildanalyse ist das Objekt-Tracking bzw. die Handverfolgung, s. Abschnitt 3.8.

In diesem Abschnitt wird die Lokalisation und Klassifikation einer Hand, basierend auf dem Ansatz der stochastischen Modellierung nach Hornegger [35], vorgestellt. Das zugrunde liegende Modell wird durch eine parametrische Wahrscheinlichkeitsdichtefunktion beschrieben, die für eine beobachtete Merkmalskonfiguration die Wahrscheinlichkeit angibt, mit der das zu lokalisierende Objekt vorhanden ist. Bei den dabei beschriebenen Verfahren ist die explizite Lösung der Zuordnung von Modellmerkmalen zu den Bildmerkmalen nicht notwendig. Die Modellparameter werden dabei berechnet, ohne dass die Zuordnung bekannt ist. Die Parameterschätzung erfolgt also *unüberwacht*. Die Eingangsdaten für die Schätzverfahren entsprechen den extrahierten potentiellen Fingerspitzen.

Modellmerkmale

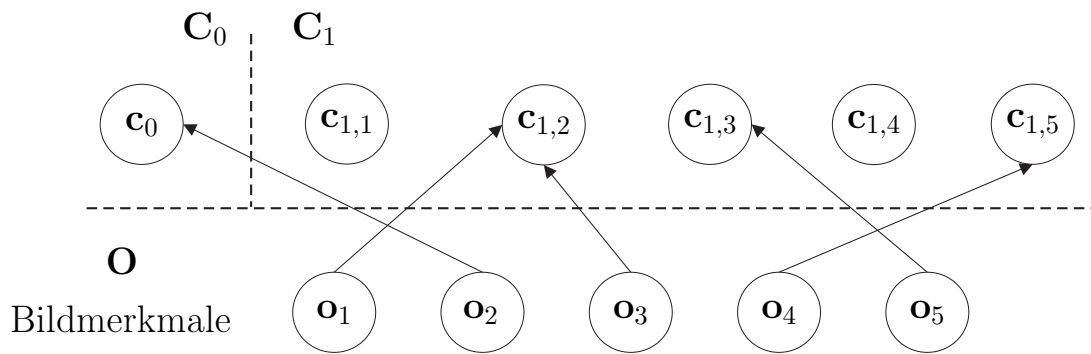


Abbildung 3.15: Beispiel für die Zuordnung von beobachteten Merkmalen

Des Weiteren ist es grundsätzlich möglich, einen Modellraum mit einer höheren Dimension als derjenigen der beobachtbaren Merkmalsvektoren zu verwenden. So kann z. B. ein 3D-Modell basierend auf zweidimensionalen Aufnahmen des Objektes trainiert werden und anschließend zur Lokalisation und Klassifikation von 2D-Aufnahmen herangezogen werden. Hierbei ist neben der kameraabhängigen Projektion der Vektoren des Modellraums in den Bildraum die Rotation um zwei weitere Achsen sowie die Translation mit drei Freiheitsgraden zu berücksichtigen. Dies hat einen fünfdimensionalen Suchraum bei der Lokalisation zur Folge. Für die hier betrachtete Handerkennung wird jedoch nur ein 2D-Modell, basierend auf zweidimensionalen Merkmalsvektoren, implementiert, da dies den für Lokalisation und Klassifikation erforderlichen Rechenaufwand deutlich reduziert. Der Grund ist die geringere Komplexität der Dichtefunktion sowie die insgesamt nur drei Freiheitsgrade der Lageparameter. Unterschiedliche Ansichten des Objektes können auch durch Modellierung als eigenständige Objektklassen berücksichtigt werden.

Im Folgenden wird der Aufbau und die Erstellung der Handmodelle nach Hornegger [35] vorgestellt, die zur Lokalisation und Klassifikation einer Hand im Bild genutzt werden.

3.6.1 Aufbau und Erstellen der Modelle

Das Objektmodell nach Hornegger [35] basiert auf einer Menge von n_κ Modellmerkmalen für eine Objektklasse Ω_κ , die durch jeweils einen Zufallsvektor $\mathbf{c}_{\kappa,l}$ ($1 \leq l \leq n_\kappa$) beschrieben werden. Diese Menge wird im Folgenden mit \mathbf{C}_κ bezeichnet

$$\mathbf{C}_\kappa = \{\mathbf{c}_{\kappa,1}, \mathbf{c}_{\kappa,2}, \dots, \mathbf{c}_{\kappa,n_\kappa}\} .$$

Jedes Modellmerkmal $\mathbf{c}_{\kappa,l}$ repräsentiert genau ein beobachtbares Merkmal des Objekts und wird durch eine parametrische Dichtefunktion charakterisiert. Nach der Merkmalsextraktion erfolgt zunächst eine Zuordnung jedes Elementes der

Menge der Beobachtungen \mathbf{O} zu einem Element von \mathbf{C}_κ . Hierbei können mehrere beobachtete Merkmale zu einem einzigen Modellmerkmal zugeordnet werden (Mehrfachdetektion). Weiter muss nicht zwangsläufig jedem Modellmerkmal auch eine Beobachtung zugeordnet werden. Die Modellierung der Zuordnungen wird im nächsten Unterabschnitt beschrieben.

Beim Training der Modellparameter während der Entwurfsphase wird davon ausgegangen, dass sich ausschließlich das zu lernende Objekt in den Trainingsaufnahmen befindet und Fehler der Merkmalsextraktion vernachlässigt werden können. Während der Laufzeit des Systems ist für die Lokalisation und Klassifikation darüber hinaus auch der Fall zu berücksichtigen, dass Merkmale erfasst werden, die keine Entsprechung im Objektmodell haben.

Diese Detektionsfehler können zum einen durch weitere in der Szene enthaltene Objekte (*Hintergrund*) hervorgerufen werden, zum anderen durch allgemeine Bildstörungen wie z. B. Rauschen. Hierzu wird im Objektmodell zusätzlich eine *Hintergrundzuordnung* eingeführt. Dabei werden diese Merkmale einer zusätzlichen Objektklasse Ω_0 , welche den Hintergrund repräsentiert, zugeordnet. In Abbildung 3.15 ist ein Beispiel für eine Zuordnung von fünf beobachteten Merkmalen \mathbf{o}_k zu den Modellmerkmalen der Objektklasse Ω_1 sowie dem Hintergrund dargestellt. Die Bildmerkmale \mathbf{o}_k enthalten die Koordinaten des Ursprungs (x_0, y_0) der Fingerspitzen und deren Rotationswinkel φ_F gegenüber der Normallage.

Vor dem Training des Modells ist zunächst für jede Objektklasse die Zahl der Modellmerkmale n_κ festzulegen. Im Fall der Handdetektion mit Fingerspitzen als Merkmale gilt $3 \leq n_\kappa \leq 5$, da in manchen Ansichten u. U. einige Fingerspitzen verdeckt sein können. Für jedes Modellmerkmal ist eine parametrische Dichtefunktion zu wählen, die die statistische Verteilung der Werte des jeweiligen Merkmals darstellen kann. Dies wird mit einer Normalverteilung erzielt. Die Dichtefunktion ist anschließend in einem Trainingslauf anhand von Musteraufnahmen zu parametrieren. Sie nimmt für normalverteilte Merkmale folgende Form an

$$p(\mathbf{c}_{\kappa,l} | \mathbf{a}_{\kappa,l}) = \mathcal{N}(\mathbf{c}_{\kappa,l} | \mu_{\kappa,l}, \Sigma_{\kappa,l}) .$$

Die Menge $\mathbf{a}_{\kappa,l}$ bezeichnet im Folgenden allgemein die Parametermenge der Dichtefunktion eines Modellmerkmals. Das Modell enthält somit keine konkreten Werte $\mathbf{c}_{\kappa,l}$ für die Modellmerkmale, sondern deren durch $\mathbf{a}_{\kappa,l}$ charakterisierte Wahrscheinlichkeitsdichtefunktion. Die $\mathbf{c}_{\kappa,l}$ werden als Zufallsvektoren aufgefasst. Bei normalverteilten Merkmalen ist damit

$$\mathbf{a}_{\kappa,l} = \{\mu_{\kappa,l}, \Sigma_{\kappa,l}\} \quad \text{mit} \quad \mu_{\kappa,l} \in \mathbb{R}^{D_m} \quad \text{und} \quad \Sigma_{\kappa,l} \in \mathbb{R}^{D_m \times D_m}$$

wobei $\mu_{\kappa,l}$ der Mittelwertvektor der Normalverteilung und $\Sigma_{\kappa,l}$ deren Kovarianzmatrix ist. D_m bezeichnet hierbei die Dimension des Modellraums. Diese ist stets größer oder gleich der Dimension der Merkmalsvektoren D_o . Beim hier betrachteten Verfahren zur Handerkennung gilt $D_m = D_o = 2$.

Zuordnung der Merkmale innerhalb einer Klasse

Im Allgemeinen ist die Zuordnung der einzelnen Beobachtungen (Bildmerkmal) $\mathbf{o}_k \in \mathbf{O}$ zu den Modellmerkmalen $\mathbf{c}_{\kappa,l}$ nicht beobachtbar, d. h. sie kann nicht aus \mathbf{O} abgeleitet werden. Um dieses Problem zu lösen, wird die Zuordnung ebenfalls stochastisch modelliert und während der Bildanalyse implizit mitgeschätzt. Hierzu wird ein m -dimensionaler Zufallsvektor eingeführt

$$\zeta_{\kappa} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_m \end{bmatrix} \in \{1, 2, \dots, n_{\kappa}\}^m \quad .$$

Jedes Element $\zeta_{\kappa,k}$ dieses Vektors nimmt den Index l desjenigen Modellmerkmals auf, welchem die jeweilige Beobachtung \mathbf{o}_k zugeordnet wird. Insgesamt gibt es damit n_{κ}^m mögliche Konfigurationen für ζ_{κ} , die jeweils mit einer eigenen diskreten Wahrscheinlichkeit auftreten

$$p(\zeta_{\kappa}) = p(\zeta_{\kappa,1} = l_1, \zeta_{\kappa,2} = l_2, \dots, \zeta_{\kappa,m} = l_m) \quad .$$

Für die Zuordnung sind somit n_{κ}^m Parameter zu schätzen.

Falls die einzelnen Merkmale in der Beobachtungsmenge unabhängig voneinander detektiert werden, sind auch die einzelnen Zuordnungen paarweise unabhängig, d. h. die Wahrscheinlichkeitsdichte der Zuordnung eines bestimmten Merkmals hängt nicht von den restlichen Zuordnungen ab. Bei der eingesetzten Hough-Transformation zur Merkmalsextraktion ist dies der Fall.

Des Weiteren ist die Reihenfolge der einzelnen Beobachtungen in \mathbf{O} aufgrund der Eigenschaften der Merkmalsextraktion i. A. undefiniert. Daher sind die Einzelwahrscheinlichkeiten ${}^k p_{\kappa,l} = p(\zeta_{\kappa,k} = l_k)$ unabhängig vom Index k .

Dann entspricht die Verbundwahrscheinlichkeit aller Zuordnungen dem Produkt der Einzelwahrscheinlichkeiten [35]

$$p(\zeta_{\kappa}) = \prod_{k=1}^m \underbrace{p(\zeta_{\kappa,k} = l_k)}_{{}^k p_{\kappa,l}} = \prod_{k=1}^m p_{\kappa,l} \quad .$$

Somit reduziert sich die Zahl der zu schätzenden Zuordnungsparameter je Klasse auf n_{κ} (nämlich $p_{\kappa,1}, \dots, p_{\kappa,n_{\kappa}}$).

Lageabhängigkeit der Dichtefunktion

Bildkoordinaten werden als Merkmalsvektoren (Bildmerkmal) der weiteren Bildverarbeitung von der Merkmalsextraktion zur Verfügung gestellt. Die Bildmerkmale sind jedoch von der Lage des Objektes im Bild abhängig und können nicht direkt mit den parametrischen Dichtefunktionen $p(\mathbf{c}_{\kappa,l} | \mathbf{a}_{\kappa,l})$ der Modellmerkmale

in Verbindung gebracht werden. Die Wahrscheinlichkeitsverteilung für die Modellmerkmale wird zunächst nur in der Objektreferenzlage angeben. Diese Referenzlage wird während des Trainings des Objektmodells beliebig festgelegt. Um Objekte in beliebigen Lagen im Bild lokalisieren zu können, müssen diese Dichtefunktionen somit noch um weitere Parameter erweitert werden.

Im Fall $D_m > D_o$ ist die Abbildung vom Modell- in den Merkmalsraum nicht invertierbar. Daher wird hier eine Transformation der Dichtefunktionen des Modells vorgenommen, anstatt umgekehrt die Merkmalsvektoren in den Modellraum zu transformieren.

Allgemein kann die Referenzlage im Modellraum durch eine Kombination aus Projektion, Rotation und Translation in jede beliebige beobachtbare Objektlage im Merkmalsraum überführt werden. Die Projektion kann für den Fall $D_m = D_o$ entfallen. Diese affine Abbildung wird durch eine Rotationsmatrix $\mathbf{R} \in \mathbb{R}^{D_o \times D_m}$ und einen Translationsvektor $\mathbf{t} \in \mathbb{R}^{D_o}$ vollständig beschrieben. Bei $D_m = D_o$ ist \mathbf{R} quadratisch, im anderen Fall nimmt sie eine rechteckige, nicht-quadratische Form an und enthält damit zusätzlich zur Rotation die Projektion. Die Transformation eines Vektors $\mathbf{c}_{\kappa,l}$ aus dem Modell- in den Merkmalsraum kann damit allgemein durch folgende Gleichung beschrieben werden

$$\mathbf{o}_k = \mathbf{R} \cdot \mathbf{c}_{\kappa,l} + \mathbf{t} \quad (3.3)$$

Da hier sowohl von 2D-Bildern als auch von einem 2D-Modell ausgegangen wird, ist die Rotationsmatrix durch einen einzigen Parameter, den Drehwinkel (im Folgenden mit φ_M bezeichnet), vollständig festgelegt

$$\mathbf{R} = \begin{bmatrix} \cos \varphi_M & \sin \varphi_M \\ -\sin \varphi_M & \cos \varphi_M \end{bmatrix} \quad (3.4)$$

Die Abbildungsmatrix 3.4 bewirkt mit Gleichung (3.3) im kartesischen Koordinatensystem eine Rotation des Vektors $\mathbf{c}_{\kappa,l}$ um den Winkel φ_M in mathematisch positiver Richtung um den Koordinatenursprung.

Damit lautet die Dichtefunktion mit integrierter Merkmalstransformation für ein beobachtetes Merkmal \mathbf{o}_k bei Annahme der Normalverteilung

$$p(\mathbf{o}_k | \mathbf{a}_{\kappa,l}, \mathbf{R}, \mathbf{t}) = \mathcal{N}(\mathbf{o}_k | \mathbf{R} \cdot \mu_{\kappa,l} + \mathbf{t}, \mathbf{R} \cdot \Sigma_{\kappa,l} \cdot \mathbf{R}^T)$$

und gibt die Wahrscheinlichkeit dafür an, dass das beobachtete Merkmal \mathbf{o}_k das Modellmerkmal $\mathbf{c}_{\kappa,l}$ der Objektklasse Ω_κ an der durch \mathbf{R} und \mathbf{t} festgelegten Position darstellt.

Parametrische Modelldichte

Die Dichtefunktion für eine beobachtete Merkmalsmenge \mathbf{O} bei statistisch unabhängigen Zuordnungen unter Berücksichtigung der Rotation, Translation und ggf. Projektion ist

$$p(\mathbf{O}|\mathbf{B}_\kappa, \mathbf{R}, \mathbf{t}) = \prod_{k=1}^m p(\mathbf{o}_k|\mathbf{B}_\kappa, \mathbf{R}, \mathbf{t}) = \prod_{k=1}^m \sum_{l=1}^{n_\kappa} p_{\kappa,l} p(\mathbf{o}_k|\mathbf{a}_{\kappa,l}, \mathbf{R}, \mathbf{t}) . \quad (3.5)$$

Sie entspricht der Verbundwahrscheinlichkeit aller Beobachtungen $\mathbf{o}_k \in \mathbf{O}$ und gibt die Wahrscheinlichkeit dafür an, dass die beobachtete Merkmalkonfiguration \mathbf{O} durch ein an der durch \mathbf{R} und \mathbf{t} festgelegten Position befindliches Objekt der Klasse Ω_κ erzeugt wird. Diese Art von Dichtefunktionen werden auch als transformierte Mischverteilungen oder Gaussian Mixture Modells genannt.

Die Menge \mathbf{B}_κ umfasst sämtliche Parameter des Objektmodells für eine Klasse Ω_κ , d. h. die Zuordnungs- und Dichteparameter jedes Modellmerkmals. Sie lautet für normalverteilte Merkmalsvektoren

$$\mathbf{B}_\kappa = \{ \{p_{\kappa,l}, \mathbf{a}_{\kappa,l}\} \mid 1 \leq l \leq n_\kappa \} = \{ \{p_{\kappa,l}, \{\mu_{\kappa,l}, \Sigma_{\kappa,l}\}\} \mid 1 \leq l \leq n_\kappa \} .$$

Diese Parameter müssen vor dem Einsatz im Gesamtsystem aus einer Sequenz von Trainingsaufnahmen ermittelt werden. Die Dichtefunktion in Gleichung (3.5) beschreibt das statistische Verhalten der Modellmerkmale einer Klasse, jedoch ohne Berücksichtigung des Hintergrundes.

Im Allgemeinen sind jedoch aufgrund von Fehlern bei der Merkmalsextraktion sowie durch zusätzliche im Bild dargestellte Objekte einige Merkmale in der Menge \mathbf{O} enthalten, denen kein Modellmerkmal zugrunde liegt. Diese Merkmale werden im Folgenden unter dem Begriff *Hintergrundmerkmale* zusammengefasst. Für diese Merkmale nimmt die gesamte Modelldichte in Gleichung (3.5), die aus dem Produkt von Einzelwahrscheinlichkeiten berechnet wird, Werte an, die nahe bei null liegen.

Daher ist vor der Lokalisation noch eine Hintergrundzuordnung durchzuführen. Dabei ist für jedes beobachtete Merkmal \mathbf{o}_k zu entscheiden, ob es nun mit größerer Wahrscheinlichkeit dem gerade lokalisierten Objekt oder dem Hintergrund zuzuordnen ist. Zur Berücksichtigung des Hintergrundes wird eine zusätzliche Objektklasse Ω_0 eingeführt, der diejenigen Merkmale zugeordnet werden, die keinem Modellmerkmal zugeordnet werden können. Diese Zuordnung wird ebenfalls wie bereits die Zuordnung einzelner Beobachtungen zu Modellmerkmalen als Zufallsvektor modelliert

$$\zeta_{H,\kappa} = \begin{bmatrix} \zeta_{H,\kappa}(\mathbf{o}_1) \\ \zeta_{H,\kappa}(\mathbf{o}_2) \\ \vdots \\ \zeta_{H,\kappa}(\mathbf{o}_m) \end{bmatrix} \in \{0, \kappa\}^m .$$

Dabei nimmt jedes Element dieses Vektors entweder den Index κ der gerade betrachteten Objektklasse an, falls das Merkmal dem Objekt zugeordnet wird. Bei einer Zuordnung des beobachteten Merkmals \mathbf{o}_k zum Hintergrund nimmt das entsprechende Element $\zeta_{H,\kappa,k}$ den Wert null an.

Ein beobachtetes Merkmal \mathbf{o}_k wird genau dann dem Hintergrund zugeordnet, wenn die Dichtefunktion für eine Beobachtung \mathbf{o}_k einen Wert unterhalb eines Schwellwertes Θ_H liefert

$$\zeta_{H,\kappa}(\mathbf{o}_k) = \begin{cases} 0 & , p(\mathbf{o}_k | \mathbf{B}_\kappa, \mathbf{R}, \mathbf{t}) \leq \Theta_H \quad \forall \kappa \\ \kappa & , \text{sonst: Zuordnung von } \mathbf{o}_k \text{ zur Klasse } \Omega_\kappa \end{cases}$$

Die Hintergrundzuordnung ist somit abhängig von \mathbf{R} und \mathbf{t} . Bei der Berechnung der Mischverteilungsdichte für die gesamte Beobachtung in Gleichung (3.5) werden nur noch diejenigen Merkmale berücksichtigt, die nicht dem Hintergrund oder anderen Objektklassen zugeordnet wurden, d. h. für die $\zeta_{H,\kappa}(\mathbf{o}_k) = \kappa$ gilt. Die Menge dieser Merkmale wird im Folgenden mit \mathbf{O}_κ bezeichnet und wird wie folgt bestimmt

$$\mathbf{O}_\kappa = \{ \mathbf{o}_k \in \mathbf{O} | \zeta_{H,\kappa}(\mathbf{o}_k) = \kappa \} .$$

Die um die Hintergrundmodellierung erweiterte Dichtefunktion 3.5 lautet somit

$$p(\mathbf{O}_\kappa | \Theta_H, \mathbf{B}_\kappa, \mathbf{R}, \mathbf{t}) = \prod_{\mathbf{o}_k \in \mathbf{O}_\kappa} p(\mathbf{o}_k | \mathbf{B}_\kappa, \mathbf{R}, \mathbf{t}) = \prod_{\mathbf{o}_k \in \mathbf{O}_\kappa} \sum_{l=1}^{n_\kappa} p_{\kappa,l} p(\mathbf{o}_k | \mathbf{a}_{\kappa,l}, \mathbf{R}, \mathbf{t}) . \quad (3.6)$$

und bildet die Grundlage für die Lokalisation und Klassifikation während der Laufzeit des Systems.

Training eines Modells

Für das Erstellen der Modelle ist die Bestimmung einer Schätzung $\hat{\mathbf{B}}_\kappa$ der modellspezifischen Parameter aus Gleichung (3.5) erforderlich. Hierfür stehen insgesamt N Trainingsaufnahmen mit je ${}^\rho m$ ($1 \leq \rho \leq N$) Merkmalen und bekannten Lageparametern ${}^\rho \mathbf{R}$ und ${}^\rho \mathbf{t}$ zur Verfügung. Allgemein ist folgendes Parameterschätzproblem zu lösen

$$\hat{\mathbf{B}}_\kappa = \operatorname{argmax}_{\mathbf{B}_\kappa} \sum_{\rho=1}^N \log p({}^\rho \mathbf{O} | \mathbf{B}_\kappa, {}^\rho \mathbf{R}, {}^\rho \mathbf{t}) .$$

Der Logarithmus wirkt sich nicht auf das Endergebnis aus, sondern wird eingesetzt, damit sehr kleine Werte der Dichtefunktion p bei der numerischen Berechnung nur geringe Rundungsfehler verursachen.

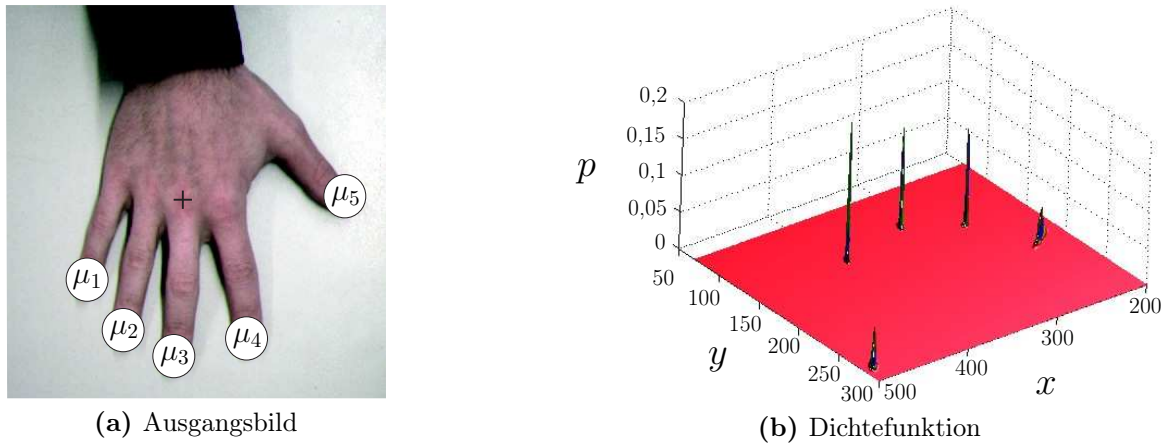


Abbildung 3.16: Modelldichtefunktion für eine rechte Hand

Da die konkrete Zuordnung von Beobachtungen zu Modellmerkmalen weder beobachtbar ist, noch im weiteren Verlauf explizit bekannt sein muss, bietet sich für die Schätzung der Modellparameter der Einsatz des Expectation-Maximization-Algorithmus (EM-Algorithmus) an. Dabei wird, ausgehend von geeigneten Startwerten, ein lokales Optimum für \mathbf{B}_κ bestimmt, während die Zuordnung implizit mitgeschätzt wird. In [19] und [17] wird der EM-Algorithmus ausführlich erklärt. Die Gleichungen der einzelnen Schätzung für \mathbf{B}_κ mithilfe des EM-Algorithmus sind im Anhang B zu finden.

In Abbildung 3.16 (a) ist als Beispiel eine rechte Hand in Normallage dargestellt. Basierend auf diesem Bild mit den eingezeichneten Positionen der Fingerspitzen wird mit dem EM-Algorithmus die Mischdichte aus Abbildung 3.16 (b) erzeugt. Zu erkennen sind die Wahrscheinlichkeitswerte für die Position der Fingerspitzen. Dieses Modell wird für die Lokalisation und Klassifikation benutzt.

Die für die Handerkennung bei Formatkreissägen benötigten Modelle, die im System abgelegt werden müssen, sind eine Hand mit gespreizten und geschlossenen Fingern sowie eine Faust. Wesentlich für die Unterscheidung einer rechten von einer linken Hand sind die Modellparameter des Daumens und des kleinen Fingers. Da der Unterschied dieser Werte nicht signifikant ist, wird auf eine explizite Berücksichtigung von rechter und linker Hand verzichtet. Die Modelle für rechte und linke Faust sind sogar nahezu identisch. Mit diesen drei Modellen wird der größte Teil der möglichen Handstellungen abgedeckt. Weitere Modelle lassen sich beliebig in das System einfügen, jedoch steigt dadurch der Speicherbedarf des Systems. Bei einer Faust werden zwar keine potentiellen Fingerspitzen mehr erkannt, jedoch bilden die Fingergelenke ebenfalls Halbkreise, die ungefähr den der Fingerspitzen entsprechen. Daher wird für die angesprochenen Handstellungen kein weiteres Merkmalsextraktionsverfahren mehr benötigt.

3.6.2 Lokalisation und Klassifikation

Durch das im vorangegangenen Abschnitt beschriebene Modelltraining sind die Schätzwerte $\hat{\mathbf{B}}_\kappa$ für die klassenspezifischen Parameter der Modelldichte in Gleichung (3.5) bekannt. Diese Modelldichte gibt die Wahrscheinlichkeit dafür an, dass der beobachteten Merkmalsmenge \mathbf{O} ein Objekt der Klasse Ω_κ an der durch $\{\mathbf{R}, \mathbf{t}\}$ beschriebenen Position zugrunde liegt. Aufgabe des Lokalisationsschrittes ist es nun, die Lageparameter \mathbf{R} und \mathbf{t} zu schätzen. Falls mehrere unterschiedliche Objektklassen betrachtet werden, und somit mehrere Parametermengen $\hat{\mathbf{B}}_\kappa$ zur Verfügung stehen, wird die Lokalisation für jede Klasse Ω_κ getrennt durchgeführt. Dabei entsteht für jede Modellklasse ein eigener Satz von Lageparametern $\{\hat{\mathbf{R}}_\kappa, \hat{\mathbf{t}}_\kappa\}$. Anschließend wird im Klassifikationsschritt entschieden, welches Objekt mit der größten Wahrscheinlichkeit lokalisiert wurde.

Die zur Lokalisation eingesetzte ML-Schätzung wird im Folgenden vorgestellt. Anschließend wird dann die Lokalisation und Klassifikation von Händen behandelt.

Maximum-Likelihood-Schätzung (ML-Schätzung): Ausgangspunkt der Schätzung ist eine parametrische Dichtefunktion der Form $p(\mathbf{O}|\mathbf{B})$. Die Menge \mathbf{O} bezeichnet hier allgemein die Menge der beobachtbaren Daten und \mathbf{B} die Parametermenge. Ziel des ML-Verfahrens ist es, Schätzwerte $\hat{\mathbf{B}}$ für die Elemente der Parametermenge so zu bestimmen, dass die Dichtefunktion p bei gegebenen Beobachtungen \mathbf{O} maximal wird

$$\hat{\mathbf{B}} = \operatorname{argmax}_{\mathbf{B}} \{ \log p(\mathbf{O}|\mathbf{B}) \} .$$

Der Logarithmus wird hier lediglich wegen Vorteilen bei der numerischen Lösung im Rechner eingefügt. Dieser wirkt sich aufgrund seiner Monotonieeigenschaft nicht auf das Ergebnis der Schätzung aus [35].

Lokalisation

Die Lokalisation erfolgt durch Maximieren der Mischverteilungsdichte in Gleichung (3.5) über dem gesamten Parameterraum $\{\mathbf{R}, \mathbf{t}\}$ mittels Maximum-Likelihood-Schätzung

$$\{\hat{\mathbf{R}}_\kappa, \hat{\mathbf{t}}_\kappa\} = \operatorname{argmax}_{\mathbf{R}, \mathbf{t}} \left\{ p(\mathbf{O}|\hat{\mathbf{B}}_\kappa, \mathbf{R}, \mathbf{t}) \right\} \quad (3.7)$$

Hierbei wird jedoch noch nicht die Zuordnung der Hintergrundmerkmale berücksichtigt. Dazu wird im Folgenden von der erweiterten Mischverteilungsdichte in Gleichung (3.6) ausgegangen. Diese kann nicht unmittelbar in Gleichung (3.7) eingesetzt werden, da die einzelnen Faktoren der Dichte in Gleichung (3.6) selbst

Wahrscheinlichkeiten darstellen und deshalb in keinem Fall größer als eins sein können. Darum wird das Produkt um so kleiner, je mehr Merkmale dem Objekt zugeordnet werden. Darüber hinaus ist der Fehler 1. Art (falsch negativ, [52]) um so geringer, je mehr Bildmerkmale erkannt und einem Modell zugeordnet werden können. Aus diesem Grund wird bei der Lokalisation die zusätzliche Nebenbedingung eingeführt, dass die Zahl der dem Objekt zugeordneten Merkmale $|\mathbf{O}_\kappa(\Theta_H, \mathbf{R}, \mathbf{t})|$ möglichst groß ist

$$\{\widehat{\mathbf{R}}_\kappa, \widehat{\mathbf{t}}_\kappa\} = \operatorname{argmax}_{\mathbf{R}, \mathbf{t}} \left\{ p(\mathbf{O}_\kappa(\Theta_H, \mathbf{R}, \mathbf{t}) | \widehat{\mathbf{B}}_\kappa, \mathbf{R}, \mathbf{t}) \mid |\mathbf{O}_\kappa(\Theta_H, \mathbf{R}, \mathbf{t})| = \max_{\mathbf{R}, \mathbf{t}} \right\}$$

Diese Nebenbedingung gilt vorrangig, d. h. bei der ML-Schätzung werden nur noch diejenigen $\{\mathbf{R}, \mathbf{t}\}$ berücksichtigt, für die $|\mathbf{O}_\kappa(\Theta_H, \mathbf{R}, \mathbf{t})|$ ein globales Maximum annimmt.

Klassifikation und Modellauswahl

Während der Lokalisation wird entschieden, welche Lageparameter für Objekte der einzelnen Klassen mit der größten Wahrscheinlichkeit die tatsächliche Objektlage beschreiben. Hierbei wird jedoch nicht berücksichtigt, ob der Beobachtung überhaupt ein entsprechendes Objekt zugrunde liegt. Um eine solche Aussage treffen zu können, wird erneut die Mischverteilungsdichte für eine Objektklasse in Gleichung (3.6) betrachtet. Diese nimmt per Definition um so größere Werte an, je mehr die beobachtete Merkmalkonfiguration mit dem trainierten Objektmodell übereinstimmt.

Für die drei modellierten Klassen kann nach deren getrennter Lokalisation durch folgende Abschätzung die am wahrscheinlichsten beobachtete Klasse Ω_κ ermittelt werden

$$\hat{\kappa} = \operatorname{argmax}_{\kappa} p(\Omega_\kappa) p(\mathbf{O} | \widehat{\mathbf{B}}_\kappa, \widehat{\mathbf{R}}_\kappa, \widehat{\mathbf{t}}_\kappa) .$$

Damit ein gesuchtes Objekt als *vorhanden* erkannt wird, muss zusätzlich Gleichung (3.6) eine Mindestwahrscheinlichkeit liefern.

Da die drei betrachteten Modelle bzw. Handstellungen ähnliche Wahrscheinlichkeitswerte liefern können, ist eine weitere auf obiger Auswahl basierende Modellauswahl notwendig. Diese erfolgt durch die geometrische Auswertung der gefundenen Merkmale \mathbf{o}_κ , welche zum gefundenen und lokalisierten Modell gehören. Dadurch kann weiteres Wissen über die Handgeometrie in den Verarbeitungsprozess einfließen und diesen dadurch zuverlässiger und robuster gestalten.

Varianzmethode zur Modellauswahl: Der hier vorgestellte Ablauf zur Modellauswahl verwendet charakteristische Varianzen der geometrischen Verteilung von Modellmerkmalen. Der gesamte Ablauf ist in Abbildung 3.18 zu sehen.

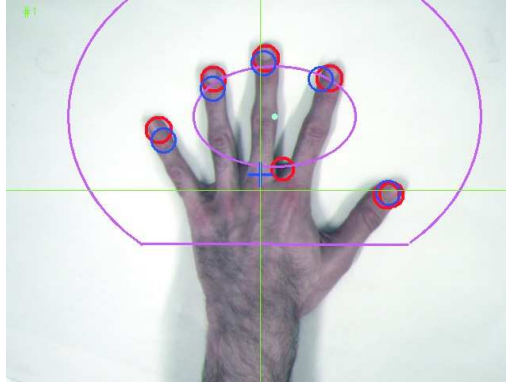


Abbildung 3.17: Gültigkeitsbereich einer flachen, gespreizten Hand

Der erste Schritt der Verarbeitung besteht darin, mithilfe der geschätzten Handorientierung φ_M die gefundenen Merkmale in die Modellnormallage zu drehen. Dies ist aufgrund der Rotationsabhängigkeit der zu berechnenden Merkmalspositionsvarianzen notwendig. Nachfolgend wird der Mittelwert $\mu_{\hat{\kappa}}$ aller dem gefundenen Modell zugeordneten Merkmale berechnet

$$\mu_{\hat{\kappa}} = \frac{1}{n_{\hat{\kappa}}} \sum_{l=1}^{n_{\hat{\kappa}}} \mathbf{o}_{\hat{\kappa},l}.$$

Anschließend wird überprüft, ob die gefundenen Merkmale in einem für jedes Modell gültigen Bereich liegen, s. Abbildung 3.17. Dazu wird der Abstand der Merkmale zu dem berechneten Mittelwert $\mu_{\hat{\kappa}}$ bestimmt. Liegt ein Merkmal außerhalb, wird es nicht berücksichtigt und der Mittelwert korrigiert. Dadurch werden falsch zugewiesene Merkmale entfernt.

Der nächste Schritt ist die Berechnung der Varianz in x- und y-Richtung

$$\sigma_{\hat{\kappa},x}^2 = \frac{1}{n_{\hat{\kappa}} - 1} \sum_{l=1}^{n_{\hat{\kappa}}} (\mathbf{o}_{\hat{\kappa},l,x} - \mu_{\hat{\kappa},x})^2$$

$$\sigma_{\hat{\kappa},y}^2 = \frac{1}{n_{\hat{\kappa}} - 1} \sum_{l=1}^{n_{\hat{\kappa}}} (\mathbf{o}_{\hat{\kappa},l,y} - \mu_{\hat{\kappa},y})^2$$

Jedes Modell besitzt charakteristische Varianzwerte, so dass durch die Auswertung der berechneten Varianzen auf das entsprechende Modell geschlossen werden kann. Für die Auswertung wird zuerst ein der Exzentrizität ϵ_V angelehntes Bewertungskriterium bestimmt. Dabei werden die berechneten Varianzwerte als Halbachsen einer Ellipse interpretiert und so die Ellipsenform beurteilt.

$$\epsilon_V = \begin{cases} \frac{\sqrt{\sigma_{\hat{\kappa},x}^4 - \sigma_{\hat{\kappa},y}^4}}{\sigma_{\hat{\kappa},x}^2}, & \text{falls } \sigma_{\hat{\kappa},x}^2 \geq \sigma_{\hat{\kappa},y}^2 \\ \frac{\sqrt{\sigma_{\hat{\kappa},y}^4 - \sigma_{\hat{\kappa},x}^4}}{\sigma_{\hat{\kappa},y}^2}, & \text{sonst} \end{cases}$$

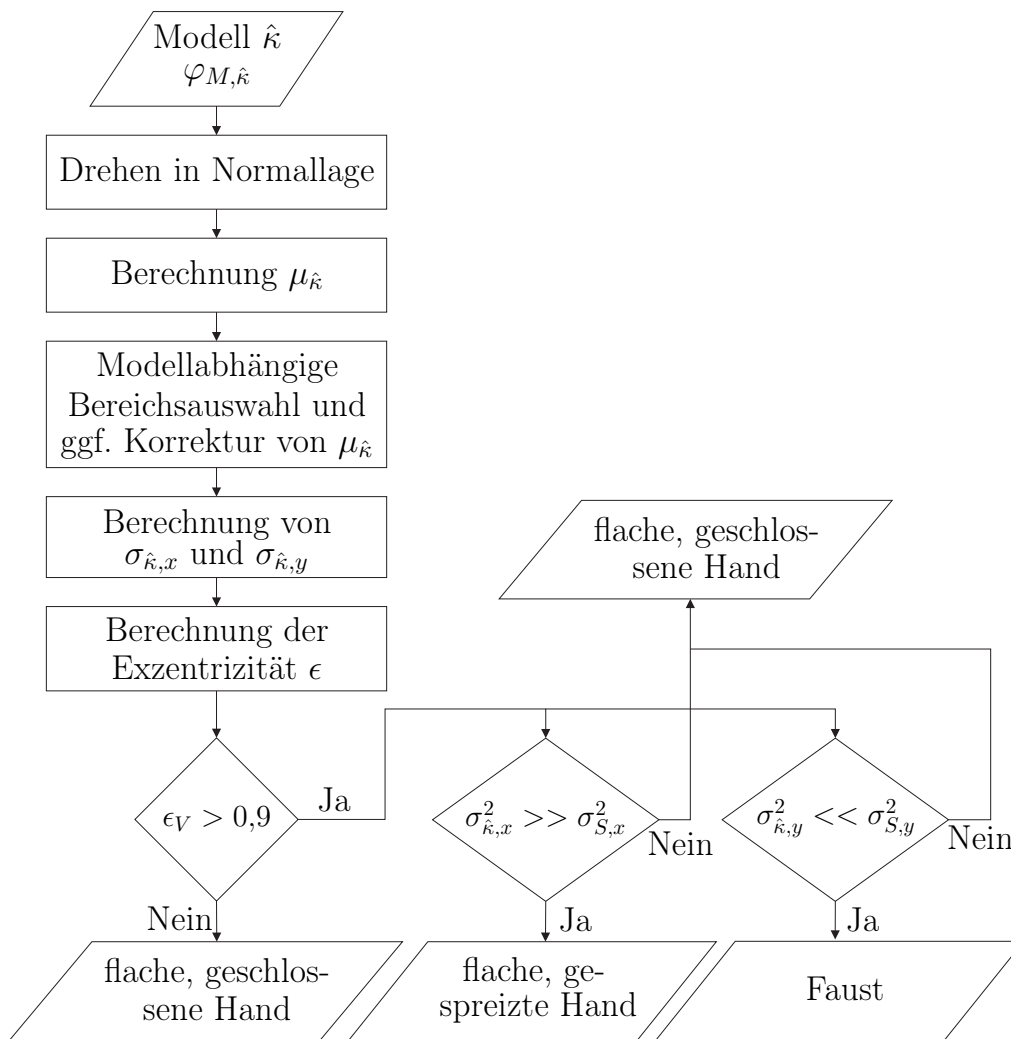


Abbildung 3.18: Ablauf der Modellauswahl mittels der Varianzmethode

Für die geometrische Interpretation des Ergebnisses werden die Standardabweichungen $\sigma_{\hat{k},x}$ und $\sigma_{\hat{k},y}$ betrachtet und eine Umrechnung der Exzentrizität ϵ_V in ϵ_S durchgeführt

$$\epsilon_S = \sqrt{1 - \sqrt{1 - \epsilon_V^2}}.$$

Liegt ϵ_S im Bereich $[0, 0,75]$ bedeutet dies, dass die Halbachsen in der gleichen Größenordnung sind und damit eine Ellipse beschrieben wird, die nicht langgestreckt und schmal, sondern kreisähnlich ist. Die Merkmalsverteilung für die flache, geschlossene Hand besitzt genau diese Eigenschaft.

Für eine schnelle Berechnung ist es allerdings von Vorteil, wenn die Wurzelfunktionen für die Bestimmung der Standardabweichungen nicht verwendet werden müssen. Außerdem sind die Werte der charakteristischen Varianzen besser unterscheidbar als die der charakteristischen Standardabweichung, weshalb die Varianzen betrachtet werden. Ein Grund dafür ist, dass die in der Varianz enthaltene

Quadrierung die Verschiedenheit der charakteristischen Modellwerte deutlicher hervorhebt, was sich bei Variationen der Handgröße vorteilhaft auswirkt.

Für den gültigen Wertebereich der flachen, geschlossenen Hand ergibt sich daher ϵ_V zu $[0, 0,9]$. Ist $\epsilon_V > 0,9$ so werden die beiden berechneten Varianzwerte mit den aus den Modellen bestimmten charakteristischen Grenzwerten verglichen. Diese Grenzwerte sind zusätzlich noch abhängig von der Kameraauflösung und Konfiguration des optischen Systems.

Eine Faust liegt demnach vor, wenn $\sigma_{\hat{\kappa},y}^2$ kleiner als $\sigma_{S,y}^2 \approx (2,2)^2 \text{ cm}^2$ ist. Dagegen erhält man für $\sigma_{\hat{\kappa},x}^2$ größer als $\sigma_{S,x}^2 \approx (4,2)^2 \text{ cm}^2$ das Modell für die flache, gespreizte Hand. Für weitere Modelle bzw. deren Auswahl kann auch ein auf Fuzzy-Logik basierendes System eingesetzt werden [53].

Durch diese Auswertung ist eine zuverlässige Auswahl der Modelle möglich, da die geometrischen Beziehungen direkt ausgewertet werden und keine parametrischen Wahrscheinlichkeiten ausgewertet werden müssen, d. h. die Varianzmethode besitzt eine höhere Evidenz als die Berechnung der parametrischen Wahrscheinlichkeiten.

3.6.3 Ergebnis

Die Ergebnisse der Lokalisation und Klassifikation sind in Abbildung 3.19 zu sehen. Ebenso ist die charakteristische Ellipse der Varianzmethode eingezeichnet. Die vorgestellten Methoden und Verfahren zur Lokalisation, Klassifikation und Modellverifikation funktionieren zuverlässig und robust. Selbst Variationen der Handgeometrie werden toleriert, d. h. das Verfahren funktioniert auch bei unterschiedlichen Maschinenbenutzern.

Es stehen nun die Position und die Rotation einer Hand im Bild zur Verfügung. Die Varianzmethode erlaubt darüber hinaus eine schnelle und zuverlässige Modellauswahl. Dies ist vor allem für das nachfolgende Tracking mit Aktiven Konturen für deren Initialisierung notwendig. Die Genauigkeit der Verfahren liegt bei 18° für die Orientierung und $5 \text{ Pixel} \approx 3 \text{ mm}$ bei der Position.

3.7 Handrepräsentation mit Aktiven Konturen

Im vorangegangenen Kapitel ist die Lokalisation und Klassifikation mithilfe von probabilistischen Modellen dargestellt. In die verwendeten Modelle fließt das Wissen über die geometrischen Zusammenhänge der Fingerspitzen, die bei der Merkmalsextraktion gefunden werden, ein. Die Komplexität und der Rechenaufwand dieses Vorgehens sind hoch. Aus diesem Grund ist ein Verfahren notwendig, das nach einer erfolgreichen Lokalisation und Klassifikation einer Hand, diese repräsentieren und weiterhin erfassen kann. Jedoch muss dieses mit geringem Aufwand

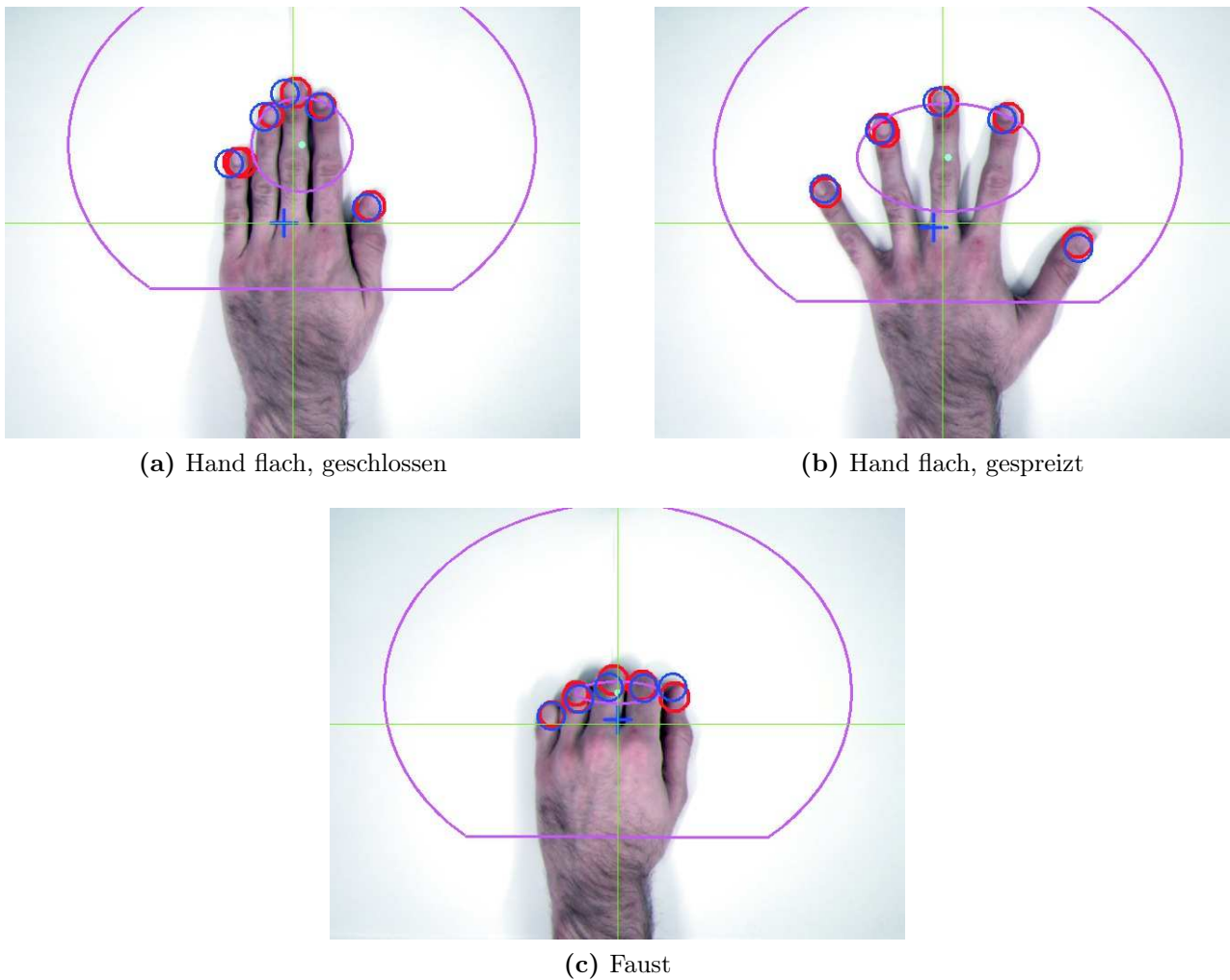


Abbildung 3.19: Ergebnis der Lokalisation und Klassifikation

erfolgen, damit die nachfolgende Objektverfolgung schnell und zuverlässig arbeiten kann.

Die so genannten Aktiven Konturen können *high-level* Merkmale wie z. B. Objektformen repräsentieren [11]. Weitere Eigenschaften, die a priori in den Aktiven Konturen integriert sind, sind beispielsweise Geschlossenheit, Glattheit oder Steigtigkeit einer Objektkontur. Der Einsatz dieser Aktiven Konturmodelle ermöglicht die Erfüllung obiger Forderungen.

Wie in Abbildung 3.2 zu sehen ist, werden die Aktiven Konturen nach der Lokalisation und Klassifikation eingesetzt. Der Grund für dieses Vorgehen ist, dass dadurch diese Konturen einfach und zuverlässig initialisiert werden können.

Die folgenden Ausführungen fassen die wichtigsten Informationen zur Anwendung der Aktiven Konturen zusammen und beschreiben das prinzipielle Vorgehen. Weitere und detaillierte Informationen sind in [11] zu finden.

3.7.1 Aktive Konturmodelle und B-Spline-Kurven

Es existieren verschiedene Möglichkeiten, um die Aktiven Konturen zu beschreiben. Der hier verwendete Ansatz nutzt deformierbare, dynamische Konturen. Dabei wird eine Musterkurve definiert und über charakteristische Größen parametrisiert. Dadurch ist die Möglichkeit der Verformung der Kontur bereits im Voraus auf bestimmte gewünschte Formen beschränkt, damit nicht beliebig deformierte Konturen entstehen können. Voraussetzung für diese deformierbaren Templates ist, dass sich die zu beschreibenden Objekte durch eine endliche Anzahl an dafür geeigneten Größen parametrisieren lassen.

Da in der Praxis oftmals Objekte von Interesse sind, deren eigentliche Kontur leicht variiert, diese aber dennoch zur gleichen Klasse gehören sollen, ist es sinnvoll zusätzlich noch eine dynamische Deformierung zuzulassen. Dabei werden die Vorteile von Snakes und der deformierbaren Templates kombiniert, so dass eine robuste, mächtige Klasse an Aktiven Konturen entsteht. *Snakes* sind parametrische Modelle, die energieminimierende Konturen darstellen [50].

Diese Aktiven Konturen werden durch B-Spline-Kurven repräsentiert, welche in ihrer Verformbarkeit beschränkt werden, so dass sie eng an einer zuvor definierten mittleren Objektform bleiben.

B-Spline-Funktion und B-Spline-Kurven: Eine *B-Spline-Funktion* wird aus einer gewichteten Summe von N_B *Basisfunktionen* $B_n(s)$ mit $n = 0, \dots, N_B - 1$ konstruiert. Im einfachsten Fall besteht jede Basisfunktion aus d Polynomen der Ordnung d mit dem Grad $d - 1$, welche über einem *Segment* der s -Achse definiert sind. Ohne Beschränkung der Allgemeinheit kann das Segment zu $s \in [0,1]$ angenommen werden. Die einzelnen Segmente sind an Knotenpunkten zusammengesetzt. In diesem Fall sind die Knotenpunkte äquidistant verteilt und *regulär*, d. h. die resultierende Spline-Funktion ist maximal glatt, sie ist also $(d - 2)$ -mal stetig differenzierbar. Mathematisch lässt sich die Spline-Funktion darstellen als

$$x(s) = \sum_{n=0}^{N_B-1} x_n B_n(s)$$

wobei x_n die jeweils zur Basisfunktion $B_n(s)$ gehörenden Gewichte sind. In Matrix-Notation lässt sich die Spline-Funktion kompakt schreiben als

$$x(s) = \mathbf{B}(s)^T \mathbf{Q}^x$$

was einem Matrixprodukt aus dem Vektor der Basisfunktionen $\mathbf{B}(s)$ und dem Vektor der Gewichte \mathbf{Q}^x entspricht. Die erste B-Spline-Funktion für $d = 3$ ist

$$B_0(s) = \begin{cases} s^2/2 & \text{für } 0 \leq s < 1 \\ \frac{3}{4} - (s - \frac{3}{2})^2 & \text{für } 1 \leq s < 2 \\ (s - 3)^2/2 & \text{für } 2 \leq s < 3 \\ 0 & \text{sonst} \end{cases}$$

während die weiteren Basisfunktionen verschobene Kopien davon sind $B_n(s) = B_0(s - n)$.

Die so definierten Spline-Funktionen können verwendet werden, um B-Spline-Kurven in der Ebene zu konstruieren. Eine parametrische Spline-Kurve

$$\mathbf{r}(s) = (x(s), y(s))^T$$

hat die Koordinaten $x(s)$, $y(s)$, die jeweils eine Spline-Funktion mit dem Parameter s darstellen. Zunächst wird ein Intervall $0 \leq s \leq L$, das aus L Segmenten besteht, und eine geeignete Basis $B_0, B_1, \dots, B_{N_B-1}$ von N_B B-Spline-Basisfunktionen gewählt.

Zu jeder Basisfunktion B_n wird ein *Kontrollpunkt* $\mathbf{q}_n = (q_n^x, q_n^y)^T$ definiert, so dass sich die Spline Kurve als eine gewichtete Vektorsumme von Kontrollpunkten ergibt

$$\mathbf{r}(s) = \sum_{n=0}^{N_B-1} B_n(s) \mathbf{q}_n \quad \text{für } 0 \leq s \leq L .$$

Zur Vereinfachung werden die einzelnen Kontrollpunkte in *Kontrollvektoren* zusammengefasst. Dazu wird ein Spline-Space \mathcal{S}_Q von Kontrollvektoren \mathbf{Q} definiert. Diese entstehen dabei aus den Kontrollpunktkoordinaten

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}^x \\ \mathbf{Q}^y \end{pmatrix} \quad \text{mit} \quad \mathbf{Q}^x = (q_0^x, \dots, q_{N_B-1}^x)^T, \quad \mathbf{Q}^y = (q_0^y, \dots, q_{N_B-1}^y)^T .$$

Damit können die Koordinatenfunktionen einer beliebigen Spline Kurve wie folgt geschrieben werden

$$x(s) = \mathbf{B}(s)^T \mathbf{Q}^x \quad \text{und} \quad y(s) = \mathbf{B}(s)^T \mathbf{Q}^y .$$

Für einen Kontrollvektor ergibt sich folgende Darstellung für parametrische Spline Kurven in der Ebene

$$\mathbf{r}(s) = \mathbf{U}(s) \mathbf{Q} \quad \text{für } 0 \leq s \leq L \quad \text{mit} \quad \mathbf{U}(s) = \begin{bmatrix} \mathbf{B}(s)^T & \mathbf{0} \\ \mathbf{0} & \mathbf{B}(s)^T \end{bmatrix} .$$

3.7.2 Shape-Space

Mit einem Spline Vektor $\mathbf{Q} \in \mathcal{S}_Q$ werden beliebige Kurvenverläufe beschrieben. Durch dessen Veränderung kann prinzipiell eine beliebige Formvariation erzielt werden. In der Praxis führen beliebige Verformungen jedoch häufig zu unerwünschten Ergebnissen. So könnte beispielsweise aus einer Handkontur durch eine beliebige Veränderung des Spline-Vektors eine Form entstehen, die nur noch wenig bis gar nichts mehr mit der ursprünglichen Handkontur gemein hat.

Demzufolge ist es sinnvoll die Art der Verformungen einzuschränken und nur noch bestimmte allgemeine Variationen des Spline Vektors zuzulassen, so dass beispielsweise die Ergebnisform letztlich eine verschobene, rotierte oder skalierte Version der Ursprungsform darstellt. Dies kann erreicht werden, indem die Spline Kurve anstatt durch den Spline Vektor \mathbf{Q} im Spline-Space \mathcal{S}_Q durch einen *Shape-Vektor* \mathbf{X} in einem *Shape-Space* \mathcal{S} (Formen-Raum) dargestellt wird.

Dazu wird im Folgenden ein allgemeiner Shape-Space definiert, der alle erlaubten Variationen einer Musterkurve umfasst. Anschließend wird dargestellt, wie die Repräsentation einer Kurve durch ihren Shape-Vektor \mathbf{X} in die bereits bekannte Darstellung im Spline-Space \mathcal{S}_Q , also die Repräsentation durch den Spline-Vektor \mathbf{Q} , überführt werden kann. Danach wird der Raum der euklidischen Ähnlichkeiten als Shape-Space eingeführt, sowie ein Überblick über die darin erlaubten Transformationen und eine mathematische Beschreibung gegeben.

Definition des Shape-Space: Allgemein ist ein Shape-Space eine lineare Parametrisierung einer Menge von erlaubten Transformationen einer *Musterkurve*. Im mathematischen Sinn ist ein Shape-Space \mathcal{S} damit eine lineare Abbildung eines Shape-Vektors $\mathbf{X} \in \mathbb{R}^{D_X}$ auf einen Spline Vektor $\mathbf{Q} \in \mathbb{R}^{D_Q}$

$$\mathbf{Q} = \mathbf{W}\mathbf{X} + \mathbf{Q}_0$$

wobei \mathbf{W} eine $D_Q \times D_X$ *Shape-Matrix* ist. Die Musterkurve \mathbf{Q}_0 entspricht dabei einer Art konstantem Offset, gegen welchen Variationen der Form gemessen werden.

Transformationen im Shape-Space: Im Raum der euklidischen Ähnlichkeiten sind vier grundlegende Variationen darstellbar: Translation in x-Richtung, Translation in y-Richtung, Rotation und isotrope Skalierung. Diese erfüllen die Spezifikationen eines Shape-Spaces. Demgegenüber wäre aber auch vorstellbar, dass ein Objekt sich in beliebiger Weise im dreidimensionalen Raum bewegen kann, d. h. zusätzlich zur Translation in x- und y-Richtung sowie der Rotation müsste anstatt der isotropen Skalierung noch eine horizontale, vertikale und diagonale Skalierung treten. Dies sind die so genannten affinen Transformationen, weswegen der Raum auch *affiner Raum* genannt wird. Im Weiteren wird der Raum der euklidischen Ähnlichkeiten als Shape-Space verwendet, da die zu verfolgende

Hand in erster Linie planare Bewegungen vollführt und volle dreidimensionale Bewegungen eine untergeordnete Rolle spielen.

Gegeben sei eine Musterkurve $\mathbf{r}_0(s)$, repräsentiert durch den Kontrollvektor \mathbf{Q}_0 . Dann ergibt der Raum der euklidischen Ähnlichkeiten bezüglich dieser Kurve einen vierdimensionalen Shape-Space \mathcal{S} mit der Shape-Matrix

$$\mathbf{W} = \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{Q}_0^x & -\mathbf{Q}_0^y \\ \mathbf{0} & \mathbf{1} & \mathbf{Q}_0^y & \mathbf{Q}_0^x \end{pmatrix}$$

wobei $\mathbf{0}$ und $\mathbf{1}$ jeweils Vektoren der Länge $D_Q = 2 \cdot N_B$ sind für die $\mathbf{0} = (0, 0, \dots, 0)^T$ und $\mathbf{1} = (1, 1, \dots, 1)^T$ gilt.

Dabei geben die ersten zwei Spalten von \mathbf{W} die horizontale bzw. vertikale Translation wieder. Während die aus Komponenten der Kurve \mathbf{Q}_0 bestehende dritte und vierte Spalte die Skalierung sowie die Rotation bestimmen. Es wird vorausgesetzt, dass der Schwerpunkt der Musterkurve \mathbf{Q}_0 im Ursprung liegt, so dass durch die dritte und vierte Spalte ausschließlich Rotation und Skalierung festgelegt werden, jedoch keine zusätzliche Translation erzeugt wird.

3.7.3 Anpassung von Aktiven Konturen

Im Folgenden wird beschrieben, wie eine zielgerichtete Deformierung der Aktiven Kontur und damit eine *Anpassung (Fitting)* an im Bild vorhandene Objekte erzielt werden kann. Im mathematischen Sinn ist das Fitting-Problem eine Optimierungsaufgabe. Aus der Menge der möglichen und erlaubten Formen $\mathbf{Q} \in \mathcal{S}_Q$ muss diejenige Form $\hat{\mathbf{Q}}$ oder $\hat{\mathbf{X}}$ ermittelt werden, die optimal auf die Zielkontur passt. Im Folgenden sei die gesuchte Zielkontur (*Merkmalskurve*) als Spline Kurve \mathbf{r}_f mit $\mathbf{r}_f(s) = \mathbf{U}(s)\mathbf{Q}_f$ dargestellt.

Dazu wird eine allgemeine Regularisierung eingesetzt, mit der das Fitting-Problem wie folgt aussieht

$$\min_{\mathbf{X}} (\mathbf{X} - \bar{\mathbf{X}})^T \bar{\mathbf{S}} (\mathbf{X} - \bar{\mathbf{X}}) + \|\mathbf{Q} - \mathbf{Q}_f\|^2 \quad \text{mit} \quad \mathbf{Q} = \mathbf{W}\mathbf{X} + \mathbf{Q}_0$$

Die Diagonalelemente der Regularisierungsmatrix $\bar{\mathbf{S}}$ werden für jeden einzelnen Parameter festgelegt. $\bar{\mathbf{X}}$ entspricht dem mittleren Shape-Vektor einer Musterkurve. Damit erfolgt die Lösung $\hat{\mathbf{X}}$ des Fitting-Problems in zwei Schritten:

1. Projektion in den Shape-Space

$$\mathbf{X}_f = \mathbf{W}^+ (\mathbf{Q}_f - \mathbf{Q}_0)$$

mit \mathbf{W}^+ als Pseudoinverse von \mathbf{W}

2. gefolgt von einer gewichteten Summation

$$\hat{\mathbf{X}} = (\bar{\mathbf{S}} + \mathcal{H})^{-1} (\bar{\mathbf{S}} \bar{\mathbf{X}} + \mathcal{H} \mathbf{X}_f) \quad \text{mit} \quad \mathcal{H} = \mathbf{W}^T \mathcal{U} \mathbf{W}$$

wobei $\mathcal{U} = \begin{bmatrix} \mathcal{B} & 0 \\ 0 & \mathcal{B} \end{bmatrix}$ gilt und \mathcal{B} eine metrische Matrix für die B-Spline-Funktionen ist.

Die bei diesem Vorgehen entstehende Schwierigkeit ist, dass die Anpassung der aus Messungen entstandenen Zielkontur (Merkmalskurve) und der Musterkurve (Aktive Kontur) nicht exakt gelingt. Ursache hierfür sind verrauschte Bilddaten, dadurch verbleibt ein geometrischer Abstand zwischen Zielkurve und Aktiver Kontur $\mathbf{r}(s)$. Diese beiden Kurven besitzen eine unterschiedliche Parametrisierung, was sich in ihrer Norm widerspiegelt. Die Differenz dieser Normen kann als Bewertungsmaß für das Fitting verwendet werden. Daraus leitet sich das *Normal Displacement* ab [11].

Normal Displacement: Anschaulich gesehen wird die gesamte Verschiebung $\mathbf{r}_f(s) - \mathbf{r}(s)$ in eine vektorielle Summe aus der Verschiebungskomponente in tangentialer Richtung und der Verschiebungskomponente in Richtung der Kurvennormalen aufgeteilt. Die tangentiale Komponente entspricht dabei der Variation der Kurvenparametrisierung und wird deshalb vernachlässigt. Somit bleibt nur noch die normale Komponente übrig, welche den reinen Kurvenabstand unter Vernachlässigung einer ggf. unterschiedlichen Parametrisierung der Kurven angibt.

Da in einem Bild keine kontinuierlichen Messungen durchgeführt werden können, muss das invariante Abstandsmaß in diskreter Form ausgedrückt werden. Dazu wird angenommen, dass die Normalenvektoren an äquidistanten Punkten $s = s_i$, $i = 1, \dots, N$ im Abstand h entlang der gesamten Kurve $\mathbf{r}(s)$ abgetastet werden und es ergibt sich die Norm-Differenz zu

$$\|\mathbf{r} - \mathbf{r}_f\|_{\bar{\mathbf{n}}} \approx \frac{1}{N} \sum_{i=1}^N ((\mathbf{r}_f(s_i) - \mathbf{r}(s_i)) \cdot \bar{\mathbf{n}}(s_i))^2 .$$

Über $\mathbf{r}(s_i) = \mathbf{U}(s_i) (\mathbf{W} \mathbf{X} + \mathbf{Q}_0)$ kann die Norm-Differenz explizit in Abhängigkeit des Shape-Vektors \mathbf{X} ausgedrückt werden

$$\|\mathbf{r} - \mathbf{r}_f\|_{\bar{\mathbf{n}}} \approx \frac{1}{N} \sum_{i=1}^N (\nu_i - \mathbf{h}(s_i))^T [\mathbf{X} - \bar{\mathbf{X}}]^2$$

mit $\nu_i = (\mathbf{r}_f(s_i) - \bar{\mathbf{r}}(s_i)) \cdot \bar{\mathbf{n}}(s_i)$ und $\mathbf{h}(s)^T = \bar{\mathbf{n}}(s)^T \mathbf{U}(s) \mathbf{W}$.

Der Algorithmus zur Lösung des gesamten Fitting-Problems ist im Anhang C zu finden.

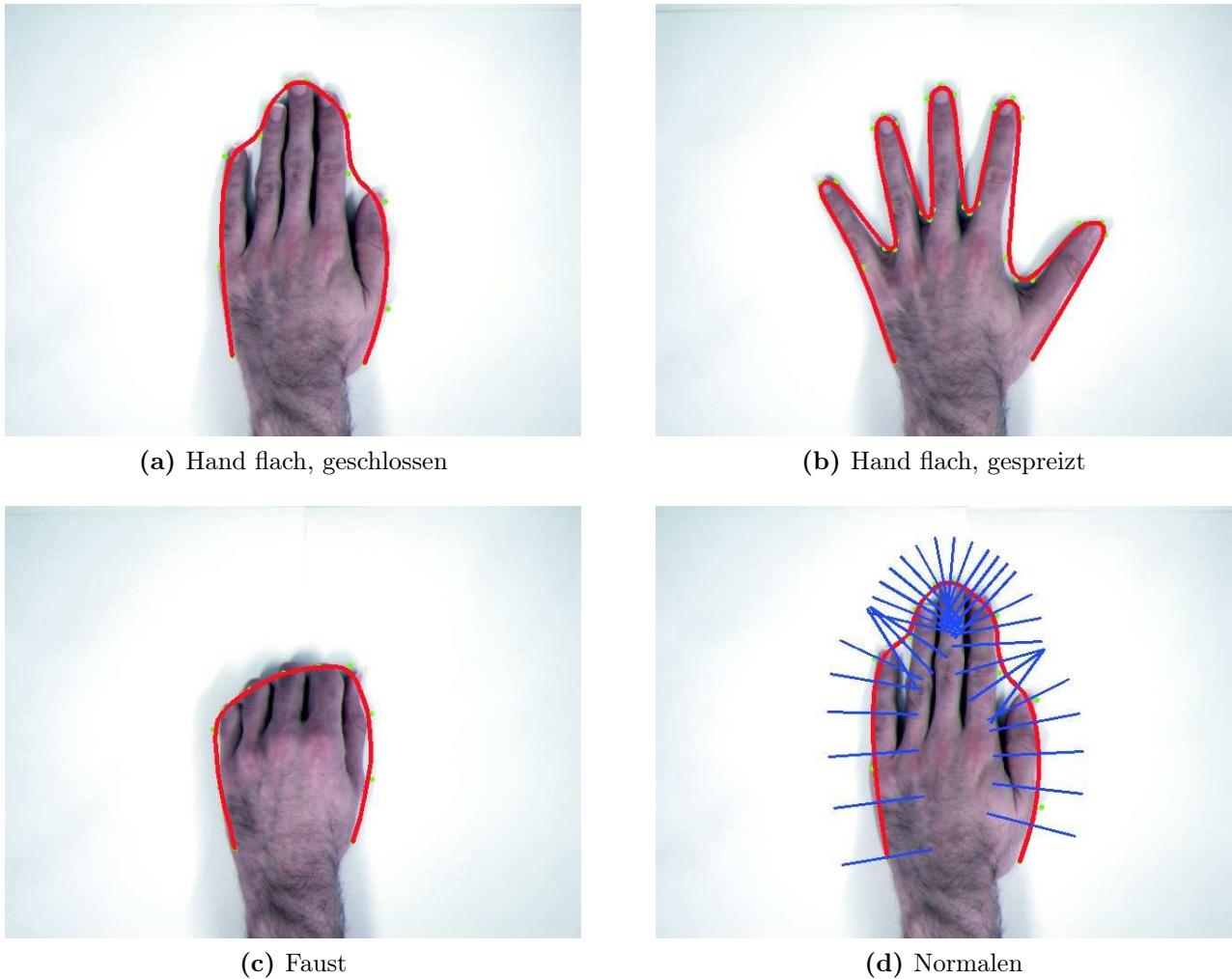


Abbildung 3.20: Aktive Konturen

3.7.4 Fazit

In den Abbildungen 3.20 (a), (b) und (c) sind die zur Handerkennung verwendeten Aktiven Konturen dargestellt. Dabei werden für die flache, geschlossene Hand 13, für die flache, gespreizte Hand 31 und für die Faust 11 Kontrollpunkte verwendet.

In Abbildung 3.20 (d) sind beispielhaft für eine Kontur die Normalen, die zur Lösung des Fitting-Problems eingesetzt werden, zu sehen.

Die hier vorgestellten Aktiven Konturen dienen dazu die bei der Lokalisation und Klassifikation gefundene Handposition der nachfolgenden Objektverfolgung zuverlässig zur Verfügung zu stellen, s. Abschnitt 3.8.3.

Die Initialisierung der Aktiven Konturen, d. h. Ermittlung der Position und Lage eines bestimmten Handmodells, wird mithilfe der Lokalisation und Klassifikation durchgeführt. Anschließend werden die Kontrollpunkte bestimmt und die Normalen festgelegt. Als nächstes erfolgt die einmalige Durchführung des Fitting-Algorithmus zur Anpassung der Musterkurve an die aktuelle Handkontur. Die

so gewonnen Parameter werden für das Tracking verwendet, das im folgenden Abschnitt erklärt wird.

3.8 Verfolgung von Objekten

Für den Menschen ist das Erkennen eines Objektes im Bild und das anschließende Verfolgen dessen Bewegung eine einfache Aufgabe. Bereits durch einen kurzen Blick werden genügend Informationen aus der Szene gewonnen. Ein Mensch kann durch die Kombination von Vorwissen über die Szene und das Objekt mit großer Exaktheit ableiten, wie schnell und in welcher Richtung sich ein Objekt bewegt. Auch kurzzeitige Verdeckungen, unruhige Hintergründe oder gar Verformungen und perspektivische Verzerrungen beeinträchtigen die menschliche Wahrnehmung kaum.

Einem Computer steht meist nur ein Bruchteil dieses Wissens zur Verfügung, da es nicht möglich ist, jede Szene in ihrer Gesamtheit in computergerechter Weise zu modellieren. Deshalb muss ein Rechner in der Lage sein, über ein ihm zur Verfügung stehendes Objektmodell, Bewegungsmodell und Sensormodell die Objektverfolgung durchzuführen.

Im Folgenden werden grundlegende Aspekte zur probabilistischen Objektverfolgung (*Tracking*) betrachtet. Anschließend wird der zur Handverfolgung verwendete Partikelfilter vorgestellt. Verwendung findet hier der Condensation Algorithmus nach Isard und Blake [41].

3.8.1 Tracking

Aus einem initialen Objektzustand versucht man beim Tracking mithilfe des Bewegungsmodells den Folgezustand im nächsten Bild zu schätzen. Es wird also eine Prädiktion durchgeführt. Aus dem Folgebild, für das der Objektzustand prädiziert wurde, werden nun mittels Objektmodell und der Merkmalsextraktion Messergebnisse aus der lokalen Umgebung der Vorhersage generiert. Da die Merkmalsextraktion ein aufwändiger Prozess ist, wird versucht den Messbereich möglichst klein zu halten. Anhand der Messungen kann nun der tatsächliche Objektzustand ermittelt werden und die ganze Prozedur mit Prädiktion, Messung und Korrektur beginnt wieder von vorn.

Bei einem einfachen Ansatz für das Tracking wird die Bildsequenz als Einzelbildfolge interpretiert und auf jedes einzelne Bild der im vorangegangenen vorgestellte Fitting-Algorithmus angewendet. Dabei entfällt der Prädiktionsschritt, da als Prädiktionwert einfach die zuletzt gefundene Objektposition verwendet wird. Ein Nachteil bei diesem einfachen Verfahren besteht in der ungenauen Vorhersage der Anfangsposition in jedem Anpassungsschritt. Sobald sich das Objekt sehr schnell bewegt, resultiert eine große Verschiebung im Bild. Damit kann im un-

günstigsten Fall die Anfangsschätzung durch Verwendung der zuletzt gefundenen tatsächlichen Objektposition derart schlecht sein, dass der Fitting-Algorithmus versagt. Zudem kann aufgrund von starken Störungen im Bildhintergrund eine unverhältnismäßig hohe Anzahl an falschen Übereinstimmungen beim Fitting-Algorithmus entstehen, so dass ebenfalls keine erfolgreiche Anpassung gelingt. Eine deutliche Verbesserung der Objektverfolgung kann dadurch erreicht werden, dass der Objektzustand (Position, Orientierung und ggf. Geschwindigkeit und Richtung) probabilistisch interpretiert wird. Dieser wird folglich durch eine Wahrscheinlichkeitsdichte beschrieben. Damit wird nicht nur ein zukünftiger Objektzustand geschätzt, sondern es können eine ganze Reihe von möglichen Objektzuständen erfasst werden.

Das Optimierungsproblem (Fitting-Algorithmus) kann in ein Schätzproblem umgewandelt werden, wenn der Objektzustand auch die Parameter der Aktiven Konturen berücksichtigt. Die dabei entstehenden Ungenauigkeiten sind vernachlässigbar, da der Vorteil einer wesentlich geringeren Komplexität überwiegt.

3.8.2 Probabilistisches Tracking

Tracking Verfahren, die mit den Grundsätzen der Wahrscheinlichkeitstheorie arbeiten, werden allgemein auch als Bayes-Tracker bezeichnet, da diese auf der Bayes'sche Formel (s. Gleichung (4.1)) bzw. deren Grundsätzen beruhen. Der Kalman Filter ist ein bekannter Vertreter der Bayes-Tracker, weshalb er in diesem Zusammenhang vorgestellt wird. Um die Nachteile des Kalman Filters zu umgehen, werden hier jedoch Partikelfilter verwendet.

Kalman Filter

Für die Herleitung und weitere Betrachtungen sei auf [92] verwiesen.

Der Kalman Filter ist laut Voraussetzungen nur für lineare und gaußverteilte Prozesse geeignet. Dies bereitet in der Bildverarbeitung Schwierigkeiten, da sich Objekte selten gemäß linearer Gleichungen bewegen. Besonders Bewegungen der menschlichen Hand sind aufgrund der vielen Freiheitsgrade meist nichtlinear. Ebenso führen extrahierte Merkmale in einem Bild oft nicht zu gaußverteilten Dichtefunktionen. Des Weiteren sind in der Regel mehrere Merkmale bzw. Objekte zu verfolgen, so dass die zu betrachtenden Dichtefunktionen multimodal werden.

Um jedoch auch Nichtlinearitäten verarbeiten zu können, wurden das linearisierte und das Extended Kalman Filter entwickelt. Dabei ist während der Laufzeit eine nichtlineare Matrix-Riccati-Differenzialgleichung zu lösen. Dies ist einerseits sehr rechenintensiv, andererseits führen Rechenzeit minimierende Vereinfachungen zu Ungenauigkeiten. Die Anforderung, nicht gaußförmige Wahrscheinlichkeitsdichten bewältigen zu können, bleibt bei diesen beiden Ansätzen unberücksichtigt.

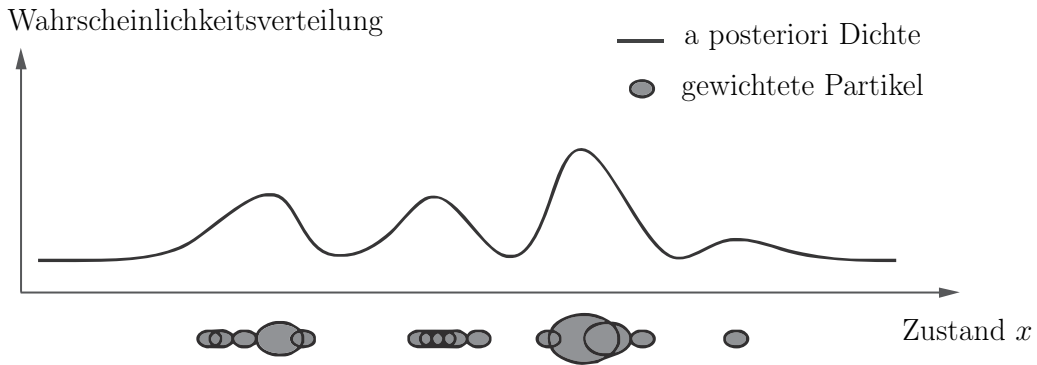


Abbildung 3.21: Factored Sampling Prozess [41]

Dieser Sachverhalt ist ausschlaggebend dafür, dass der Kalman Filter und dessen Varianten hier nicht zur Anwendung kommen.

Partikelfilter

Bei der Bildverarbeitung werden zweidimensionale Signale als Bild interpretiert. Wird das Bild nun als Dichtefunktion betrachtet, so handelt es sich dabei um eine Funktion mit mehreren Veränderlichen. Zusätzlich ist diese Funktion i. A. nicht gaußförmig sondern multimodal. Um diese Dichtefunktion und die Nichtlinearität einer Bewegung verarbeiten zu können, kann ein Partikelfilter eingesetzt werden. Die Wahrscheinlichkeitsdichten werden hierbei nicht funktional beschrieben, sondern werden in einem *Factored Sampling* Prozess erzeugt. Die dabei entstehenden gewichteten *Partikel* oder *Samples* repräsentieren die multimodale Wahrscheinlichkeitsdichte.

Beim Factored Sampling wird eine Wahrscheinlichkeitsverteilung $P(\mathbf{Z}|\mathbf{X})$ zu einem diskreten Zeitschritt t durch eine endliche Menge gewichteter Partikel dargestellt

$$\left\{ \left(\mathbf{s}_t^{(n)}, \pi_t^{(n)} \right), n = 1, \dots, N \right\} .$$

Dabei stellt $\mathbf{s}_t^{(n)}$ zum Zeitschritt t das n -te Partikel einer Menge von insgesamt N Partikeln dar und $\pi_t^{(n)}$ ist die diesem Partikel zugeordnete Gewichtung. Die einzelnen Gewichte werden anhand der folgenden Gleichung bestimmt. Als Beispiel dient hier die Beobachtungsdichte $p(\mathbf{z}_t|\mathbf{x}_t)$.

$$\pi_t^{(n)} = \frac{p(\mathbf{z}_t|\mathbf{x}_t = \mathbf{s}_t^{(n)})}{\sum_{k=1}^N p(\mathbf{z}_t|\mathbf{x}_t = \mathbf{s}_t^{(k)})} .$$

Nach der Bestimmung aller Partikelgewichte werden diese so normiert, dass sie in der Summe eins ergeben. Die Wahrscheinlichkeitsdichte wird also ohne ein funktional ausreichend beschrieben. Dieser Prozess wird auch Monte Carlo Methode

genannt. Eine graphische Veranschaulichung einer derart approximierten Wahrscheinlichkeitsverteilung für den eindimensionalen Fall zeigt Abbildung 3.21. Die in der Abbildung gezeigte eindimensionale Verteilung wird durch die als Ovale dargestellten Partikel approximiert. Dabei entspricht jeweils der Mittelpunkt eines Ovals einem Partikel, während die Fläche die Gewichtung angibt.

Bei Partikelfiltern werden statt Wahrscheinlichkeitsdichten Samples durch das System propagiert. Durch das so genannte *Resampling* wird die Degeneration der Samples verhindert. Dies tritt auf, wenn nur eine endliche Anzahl Samples (bei realen System ist das der Fall) vorhanden sind und einzelnen dieser Samples im Messschritt ein großes Gewicht zugewiesen wird. Nach wenigen Iterationsschritten wären ohne das Resampling alle Partikel bis auf wenige verschwunden. Der betrachtete Filter würde divergieren [92].

Allgemein wird bei Bayes-Trackern die Markov-Eigenschaft erster Ordnung vorausgesetzt, d. h. der aktuelle Zustand des Systems ist nur vom unmittelbar vorgegangen abhängig. Da nur der jeweils letzte Zustand des Systems für die Berechnung eines neuen benötigt wird, vereinfachen sich die zu betrachtenden Wahrscheinlichkeitsdichten und das System kann als zeitdiskrete Markov-Kette modelliert werden. Diese Art der Modellierung der Systemdynamik erlaubt eine allgemeine Systembeschreibung. Die dabei entstehende Zustandsdichte stellt den Prädiktionsschritt oder *predict step* eines Partikelfilters dar und wird durch ein Bewegungsmodell beschrieben

$$p(\mathbf{x}_t | \mathbf{X}_{t-1}) = p(\mathbf{x}_t | \mathbf{x}_{t-1}) .$$

Im Messschritt bzw. *measure step*, der dem Prädiktionsschritt folgt, werden die einzelnen Messungen \mathbf{z}_i als von einander unabhängig vorausgesetzt. Das bedeutet für die Beobachtungsdichte

$$p(\mathbf{Z}_t | \mathbf{X}_t) = \prod_{i=1}^t p(\mathbf{z}_i | \mathbf{x}_i) .$$

Die Messungen seien auch unabhängig von der Systemdynamik, so dass folgendes gilt

$$p(\mathbf{Z}_{t-1}, \mathbf{x}_t | \mathbf{X}_{t-1}) = p(\mathbf{x}_t | \mathbf{X}_{t-1}) \prod_{i=1}^{t-1} p(\mathbf{z}_i | \mathbf{x}_i) .$$

Eine zeitdiskrete Markov-Kette mit unabhängigen Beobachtungen als System wird vorausgesetzt. Damit gilt für die a priori Zustandsdichte

$$p(\mathbf{x}_t) \equiv p(\mathbf{x}_t | \mathbf{Z}_t) .$$

Die a priori Dichte des Systems ist somit identisch mit der a posteriori Zustandsdichte, was typisch für Bayes-Tracker ist. Die a posteriori Dichtefunktion wird daher nach jedem Iterationsschritt gleich der a priori Dichte für den Folgeschritt.

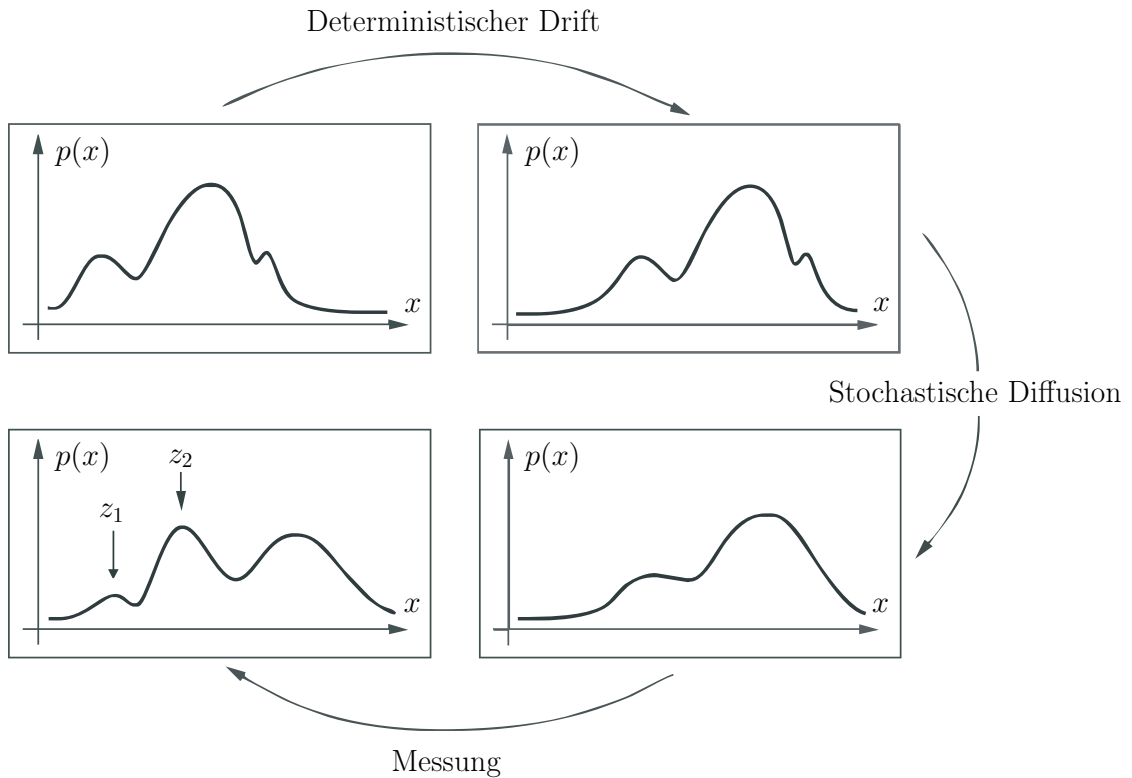


Abbildung 3.22: Propagieren der bedingten Wahrscheinlichkeitsdichte aus [41]

Für das Propagieren der Wahrscheinlichkeitsdichten durch das System wird die Formel von Bayes betrachtet

$$p(\mathbf{x}_t | \mathbf{Z}_t) = \frac{p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t)}{p(\mathbf{z}_t)}.$$

Bei $p(\mathbf{x}_t | \mathbf{Z}_t)$ handelt es sich um die a posteriori Dichte nach der Messung und bei $p(\mathbf{z}_t | \mathbf{x}_t)$ um die Beobachtungsdichte. $p(\mathbf{z}_t)$ hängt nicht vom Systemzustand ab und wird deshalb als eine Normierungskonstante $1/k_t$ weitergeführt. Die a posteriori Dichte $p(\mathbf{x}_t)$ ist aus dem vorherigen Iterationsschritt bekannt. Mit diesem Wissen verallgemeinert sich die Bayes'sche Formel zu

$$p(\mathbf{x}_t | \mathbf{Z}_t) = k_t \cdot p(\mathbf{z}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{Z}_{t-1}) \tag{3.8}$$

mit

$$p(\mathbf{x}_t | \mathbf{Z}_{t-1}) = \int_{\mathbf{x}_{t-1}} p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \mathbf{Z}_{t-1}) d\mathbf{x}_{t-1} \tag{3.9}$$

Diese Gleichung modelliert, wie aus der a posteriori Dichte $p(\mathbf{x}_{t-1} | \mathbf{Z}_{t-1})$ des vorherigen Zustands durch die Systemdynamik die a priori Dichte $p(\mathbf{x}_t | \mathbf{Z}_{t-1})$ des folgenden Systemzustands vor der Messung \mathbf{z}_t ermittelt wird.

In Abbildung 3.22 ist der Verlauf der Dichtefunktionen für den allgemeinen Fall (multimodale, nicht gaußförmige Dichte und nichtlineares System) grafisch dargestellt. Das Propagieren der Dichten durch den Partikelfilter ist darin zu erkennen.

Zu Beginn steht eine a priori Dichte $p(x)$ (*im Bild oben links*). Sie unterliegt einem stochastischen Prozess, der sich aus einem deterministischen Drift (*deterministic drift*), verursacht vom Bewegungsmodell, und einer stochastischen Diffusion (*stochastic diffusion*) zusammensetzt. Die verschiedenen lokalen Maxima der Funktionen $p(x)$ sind auf den komplexen Hintergrund, der durch Beleuchtungseffekte oder Gegenstände mit ähnlichen Eigenschaften entsteht, zurückzuführen. Der stochastische Prozess kann mit den Gleichungen (3.8) und (3.9) erklärt werden.

Die a priori Dichte $p(\mathbf{x}_{t-1}|\mathbf{Z}_{t-1})$ ist anfangs bekannt. Mithilfe der Zustandsdichte des Systems und Gleichung (3.9) lässt sich zum einen $p(\mathbf{x}_t|\mathbf{Z}_{t-1})$ die Dichte des Systems im Folgezustands vor der Messung \mathbf{z}_t berechnen. Zum anderen wird damit der deterministische Drift modelliert oder anders ausgedrückt, der folgende Zustand wird mit einem Bewegungsmodell vorhergesagt (Prädiktion). Der Effekt der stochastischen Diffusion wird durch die Beobachtungsdichte $p(\mathbf{z}_t|\mathbf{x}_t)$ kompensiert (Messung). Mit dem neuen Messwert z_t sorgt die Gleichung (3.8) dafür, dass aus $p(\mathbf{x}_t|\mathbf{Z}_{t-1})$ die a posteriori Dichte $p(\mathbf{x}_t|\mathbf{Z}_t)$ (*im Bild unten links*) ermittelt werden kann. Aus der a posteriori Wahrscheinlichkeitsfunktion $p(\mathbf{x}_t|\mathbf{Z}_t)$ wird im nächsten Iterationsschritt die a priori Dichte $p(\mathbf{x}_{t-1}|\mathbf{Z}_{t-1})$. Damit kann der Ablauf wieder von vorne beginnen.

3.8.3 Condensation Algorithmus

In [41] beschreibt Isard den Condensation Algorithmus, welcher einen Partikelfilter implementiert. Trotz des vergleichsweise einfachen Prinzips des Ansatzes, können Objekte vor stark ausgeprägten Störmerkmalen im Bildhintergrund zuverlässig erkannt und verfolgt werden. Ein weiterer Vorteil des Condensation Algorithmus besteht darin, dass dieser problemunabhängig ist, da er letztlich nur die *Verwaltung* der Partikel übernimmt. Das Bewegungsmodell beschreibt die deterministischen, zeitabhängigen Veränderung der durch die Partikel dargestellten Objektzustände. Das Beobachtungsmodell schließlich sorgt für die Gewichtung der Hypothesen, die ebenfalls durch die Partikel repräsentiert werden.

Funktionsweise

Beginnend mit einer a priori Dichte, welche entweder die Initialisierungsdichte $p(\mathbf{x}_0)$ ist oder der a posteriori Dichte $p(\mathbf{x}_{t-1}|\mathbf{Z}_{t-1})$ des vorhergegangenen Schrittes des Condensation Algorithmus entspricht, wird eine neue Partikelmenge bestimmt. Dabei werden in einem stochastischen Auswahlprozess niedrig gewichtete Partikel verworfen, während hoch gewichtete sogar mehrfach gewählt werden können. Dieser Auswahlprozess entspricht also einem *Ziehen mit Zurücklegen* und beinhaltet das Resampling. Dadurch wird die Gefahr einer Partikeldegeneration verhindert.

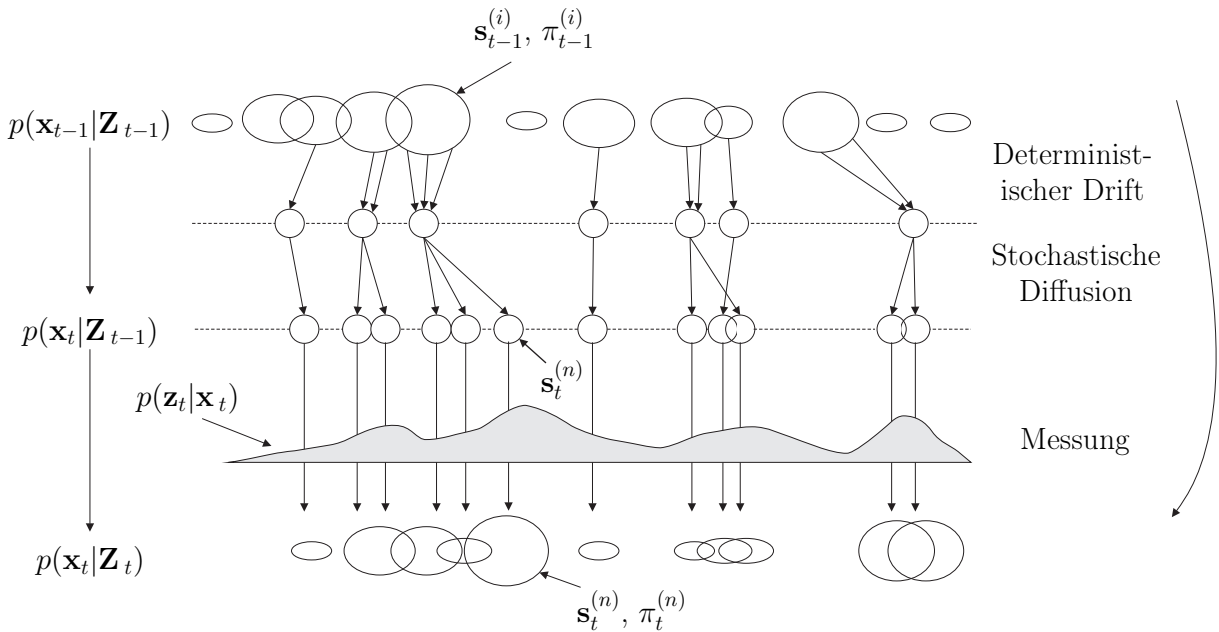


Abbildung 3.23: Ein Iterationsschritt des Condensation Algorithmus [41]

Anschließend werden diese neuen, noch ungewichteten Partikel einem deterministischen Drift und einer stochastischen Diffusion unterzogen. Diese Bewegung wird durch das Bewegungsmodell realisiert. Danach erfolgt eine Neugewichtung der einzelnen Partikel entsprechend der im Bild vorgenommenen Messung (z. B. über eine Merkmalsextraktion). Diese Neugewichtung wird durch ein Beobachtungsmodell dargestellt. Abschließend werden dann die Gewichte derart normiert, dass sie sich zu eins summieren. Damit wird die neue a posteriori Verteilung durch die neue Partikelmenge approximiert und kann zur Prädiktion des nächsten Objektzustandes verwendet werden.

Um den jeweils ersten Schritt der Partikelauswahl aus der alten Partikelmenge effizient durchführen zu können, wird jedem Partikel ein zusätzlicher Parameter $c_t^{(n)}$ zugewiesen. Dieser Parameter entspricht einem kumulativen Gewicht als Summe aller bis zum n -ten Partikel aufsummierten Gewichte. Dabei gilt

$$\begin{aligned}
 c_t^{(0)} &= 0 \\
 c_t^{(n)} &= c_t^{(n-1)} + \pi_t^{(n)} \quad (n = 1, \dots, N) .
 \end{aligned}$$

Der Auswahlprozess kann dadurch mit einem reduzierten Aufwand durchgeführt werden, da die Partikel über $c_t^{(n)}$ der Größe nach sortiert sind und demzufolge eine effiziente Binärsuche implementiert werden kann. Eine Partikelmenge stellt sich somit wie folgt dar

$$\left\{ \left(\mathbf{s}_t^{(n)}, \pi_t^{(n)}, c_t^{(n)} \right), n = 1, \dots, N \right\}$$

Eine anschauliche graphische Darstellung eines solchen Verfolgungsschrittes ist in Abbildung 3.22 und 3.23 zu sehen, während der Condensation Algorithmus im Detail im Anhang D zu finden ist.

ICondensation Algorithmus

Der *ICondensation* Algorithmus ist eine Erweiterung des vorgestellten Condensation Algorithmus. Dabei wird ein so genanntes *importance sampling* zur Verbesserung der Qualität des *factored samplings* durchgeführt. Grundlage dazu ist eine *Importance Funktion*, mit der zusätzlich verfügbares Wissen in den Partikelfilter eingebracht wird. Die Idee dabei ist, dass die Samples entsprechend dem neuen verfügbaren Wissen gewichtet werden [41].

In Abschnitt 3.3.3 ist die Bewegungserkennung als Beginn der Bildverarbeitungskette erklärt. Der dabei verwendete Alpha-Segmentierer erkennt Gebiete, in denen eine Objektbewegung stattfindet. Dies wird durch den ROI an den nächsten Verarbeitungsschritt weitergegeben. Da durch den Condensation Algorithmus in erster Linie die Aktive Kontur einer Hand verfolgt wird, würde die ROI Information neues und zusätzliches Wissen darstellen. Dies lässt sich anhand einer Importance Funktion in den Condensation Algorithmus im Auswahlschritt integrieren.

Die Importance Funktion wird durch eine Gleichverteilung dargestellt, die den ROI repräsentiert. Liegen Partikel außerhalb des ROIs, werden diese mit null gewichtet, d. h. gelöscht. Durch die Löschung ist wiederum eine erneute Normierung der Gewichte auf Eins notwendig.

Das so erweiterte System integriert zusätzlich das high-level Merkmal Bewegung in den Verfolgungsprozess. Dadurch wird die Objektverfolgung robuster und zuverlässiger, da Partikel, welche einen falschen Objektort repräsentieren würden, von vornherein nicht weiter beachtet bzw. verarbeitet werden. Des Weiteren werden neue auf der Bewegungserkennung gestützte Hypothesen generiert.

Bewegungsmodell

Der Condensation Algorithmus benötigt für den Prädiktionsschritt ein Modell der Systemdynamik. Damit werden neue mögliche Positionen der Samples berechnet bzw. Hypothesen gebildet, wohin das zu verfolgende Objekt sich bewegen wird. Dieses Bewegungsmodell stützt den Tracking Prozess, da eine weitere Informationsquelle neben der Messung herangezogen wird.

Für den hier vorgestellten Partikelfilter ist ein autoregressiver Prozess (ARP) 2. Ordnung als einfaches Bewegungsmodell ausreichend, da im Messschritt die neuen Gewichte der Samples ermittelt werden und diese ausschlaggebend für das notwendige Resampling sind. Das Dynamikmodell dient nur dazu, kurzfristige Hypothesen zu bilden, wohin sich das Objekt in den nächsten Schritten bewegen könnte. Ein Messwert bestimmt letztendlich über die Gewichtung der Partikel, welche Hypothesen überleben.

Ein ARP 1. Ordnung, auch *Random Walk* oder *Brownsche Bewegung* genannt, beschreibt eine Handbewegung nur unzureichend, da Bewegungs- bzw. Geschwin-

digkeitsänderungen nicht in die Modellierung einfließen. Aus diesem Grund wird hier ein ARP 2. Ordnung verwendet. Diese Beschreibung deckt sämtliche Handbewegungen ab, z. B. geradlinige, kreisförmige oder beschleunigende.

Der ARP 2. Ordnung kann wie folgt dargestellt werden

$$\mathbf{x}_t = A_2 \mathbf{x}_{t-2} + A_1 \mathbf{x}_{t-1} + \mathbf{D}_0 + B_0 \mathbf{w}_t \quad (3.10)$$

mit $\mathbf{D}_0 = (\mathbf{I} - A_2 \mathbf{I} - A_1 \mathbf{I})\bar{\mathbf{x}}$, wobei $\bar{\mathbf{x}}$ ein konstanter Mittelwert ist.

In Gleichung (3.10) erkennt man, dass für die Berechnung der neuen Position sowohl die Geschwindigkeit über \mathbf{x}_{t-1} gewichtet mit A_1 , sowie die Beschleunigung mit $A_2 \mathbf{x}_{t-2}$ einfließt. Hinzu kommt die konstante Driftbewegung, wiedergegeben mit dem Offset \mathbf{D}_0 , und der stochastische Anteil, verursacht durch $B_0 \mathbf{w}_t$. Die Parameter des Modells können durch empirische Tests oder durch einen von Isard und Blake vorgeschlagenen modifizierten ML-Schätzer identifiziert werden [11].

Beobachtungsmodell

Der Messschritt des Condensation Algorithmus dient dazu, die Beobachtungsdichte $p(\mathbf{z}|\mathbf{x})$ zu generieren und daraus die zu den einzelnen Samples gehörenden Gewichte zu bestimmen. Hierfür ist ein Beobachtungsmodell für eine Messung bei komplexem Bildhintergrund aufzustellen, das als modifizierter Poissonprozess interpretiert werden kann [41].

Eine zweidimensionale Beobachtungsdichte $p(\mathbf{z}|\mathbf{x})$ bei gegebener parametrisierter Merkmalskurve $\mathbf{z}(s)$ im Bild (Messung) und bei einer gegebenen geschätzten Kurve $\mathbf{r}(s)$ (Aktive Kontur) ist daher

$$p(\mathbf{z}|\mathbf{x}) = Z(\mathbf{x}) \cdot \exp \left(- \sum_{m=1}^M \frac{1}{2rM} \min \left[(\mathbf{z}_1(s_m) - \mathbf{r}(s_m))^2, v^2 \right] \right)$$

wobei die Menge der Messwerte $Z(\mathbf{x})$ als konstant zu $Z(\mathbf{x}) = Z$ gesetzt werden, da angenommen wird, dass die Variationen von Z über \mathbf{x} klein gegenüber dem anderen Term sind.

Mit $s_m = m/M$ wird $\mathbf{z}_1(s)$ definiert als das am nächsten an $\mathbf{r}(s)$ liegende Merkmal

$$\mathbf{z}_1(s) = \mathbf{z}(s') \quad \text{mit} \quad s' = \underset{s'}{\operatorname{argmin}} |\mathbf{r}(s) - \mathbf{z}(s')| .$$

Zusätzlich ist die örtliche Skalierungskonstante v definiert als

$$v = \sqrt{2}\sigma \log \left(\frac{1}{\sqrt{2\pi\tau\sigma}M} \right) .$$

wobei $\tau = q\lambda$ das Produkt aus der Wahrscheinlichkeit q für die Nicht-Detektion von Objekten und dem Poissonprozessparameter λ ist, welcher das Hintergrundrauschen modelliert. Zudem ist σ eine konstante Standardabweichung und M die Anzahl an Normalen entlang derer die Merkmalspunkte $\mathbf{z}(s_m)$ gesucht werden. Die unbekannt Parameter können z. B. aus vorhandenen Beispielbildern geschätzt werden.

Das eingeführte Beobachtungsmodell kann weiter vereinfacht werden. Dabei werden der konstante Faktor Z sowie alle konstanten Faktoren im Exponenten vernachlässigt, da diese allen Messungen gemein sind und keine Information für das Beobachtungsmodell enthalten

$$p(\mathbf{z}|\mathbf{x}) \approx \exp \left(- \sum_{m=1}^M \min \left[(\mathbf{z}_1(s_m) - \mathbf{r}(s_m))^2, v^2 \right] \right) .$$

Des Weiteren wird v dahingehend vereinfacht, dass es als Konstante frei gewählt werden kann. Anschaulich gibt μ an, in welchem Bereich entlang jeder Normalen ein Merkmal gefunden werden muss. Wenn innerhalb dieses Bereichs kein Merkmal gefunden wird, so wird ein fiktives gefundenes Merkmal im Abstand v angenommen, um zu starke Abweichungen der Form zu verhindern, falls der eigentliche Suchbereich entlang der Normalen sehr groß ist. Durch Abschätzung der Genauigkeit des Bewegungsmodells kann v gewählt werden.

3.8.4 Fazit

Der hier vorgestellte Partikelfilter (Condensation Algorithmus) wird zur Handverfolgung eingesetzt. Das zu verfolgende Objekt wird durch Aktive Konturen repräsentiert, s. Abschnitt 3.7. Es werden 50 Partikel eingesetzt, welche die Parameter für eine Aktive Kontur repräsentieren

$$\mathbf{s}_t^{(n)} = (x, y, \varphi_{AK}, v_x, v_y, \omega)^T .$$

In einem Partikel ist also die Objektposition (x, y) und -rotation φ_{AK} enthalten. Dazu kommen als Parameter noch die Geschwindigkeit der Translation in x - und y -Richtung sowie die Rotationsgeschwindigkeit ω . Zur Reduzierung der Anzahl der Freiheitsgrade der Partikel wird hier darauf verzichtet, die Skalierung als zusätzlichen Parameter einzuführen. Bei der konkreten Anwendung zur Handverfolgung ist keine Veränderung der relativen Größe der Hand zu erwarten, da diese sich näherungsweise in einer Ebene parallel zum Bildebene bewegt und somit keine affinen Verzerrungen auftreten können.

In Abbildung 3.24 ist ein Ergebnisbild als Kantenbild der Verfolgung einer flachen, geschlossenen Hand zu sehen. Darin ist die Aktive Kontur der Hand als

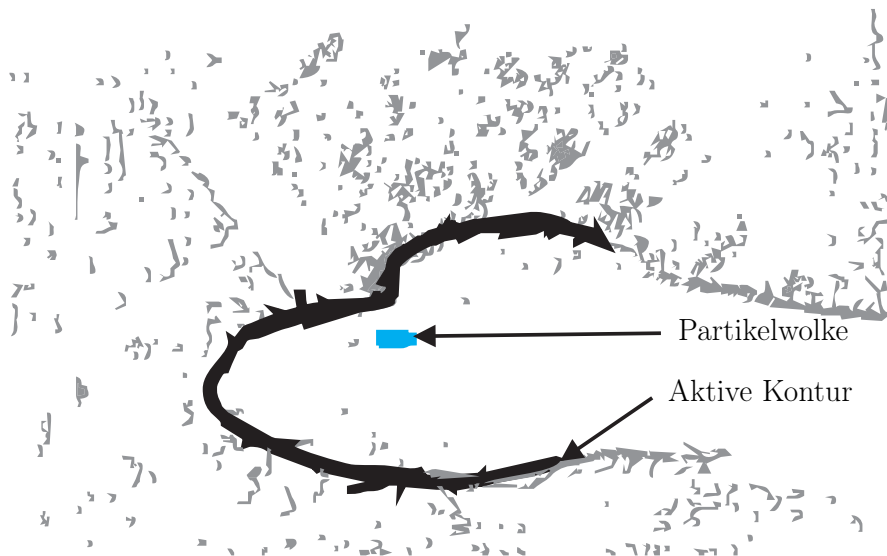


Abbildung 3.24: Ergebnisbild der Handverfolgung

verstärkte Linie abgebildet. Die Partikelwolke ist ebenfalls dargestellt. Durch den Einsatz von Aktiven Konturen und deren Verfolgung mit einem Partikelfilter konnte die Komplexität des Gesamtsystems reduziert sowie dessen Zuverlässigkeit gesteigert werden. Eine Handverfolgung ist damit möglich und bildet die Grundlage für eine anschließende Gefahrenanalyse der Handbewegung.

3.9 Analyse von Bildsequenzen zur Gefahrenerkennung

Der in Abbildung 3.2 gezeigte Ablauf zur automatischen Handerkennung endet mit der Gefahrenanalyse für die Hand. Die Zwischenschritte sind in den vorangegangenen Abschnitten aufgezeigt worden. Mit dem letzten Teil der Bildverarbeitung, dem Tracking, endet bei der Gefahrenerkennung die Bildanalyse. Als Information stehen die identifizierte Hand, deren Position, Geschwindigkeit und Bewegungsrichtung zur Verfügung. Diese Ergebnisse der Bildverarbeitung müssen nun für die nachfolgende Informationsfusion (s. Abschnitt 4.4.4) aufbereitet werden.

Für die Gefahrenanalyse ist von Interesse, wann und wo die Hand in einen für diese gefährlichen Bereich eintritt. Dazu ist neben der Handidentifikation die Auswertung der Bewegungsgeschwindigkeit und -richtung ausschlaggebend. In Abbildung 3.25 ist das *Vorfeld 1*, welches von der Kamera überwacht wird, zu sehen. Die Einteilung der Gefahrenbereiche ist in Abschnitt 5.2 erklärt.

Wird eine Hand in einem Bild erkannt, erfolgt die Auswertung ihrer Geschwindigkeit. Außerdem werden fortlaufend die Abstände $\mathbf{d}_{r,1}$ und $\mathbf{d}_{r,2}$ vom Handmittelpunkt zum *Vorfeld 1 vorne* bzw. *Lamelle 1 links* berechnet. Anschließend vergleicht man die Bewegungsrichtung mit der Orientierung der Vektoren $\mathbf{d}_{r,1}$

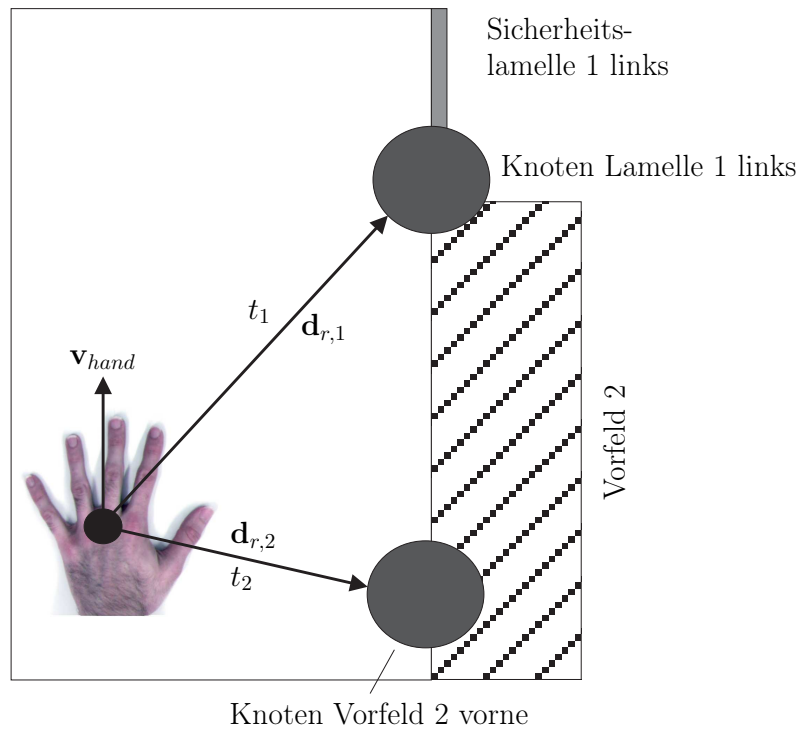


Abbildung 3.25: Beobachteter Bereich des Überwachungsbereiches

und $\mathbf{d}_{r,2}$. Bewegt sich die Hand nun in Richtung Vorfeld oder Lamelle wird entweder die Zeit t_1 oder t_2 berechnet und der Zustand im Knoten *Kamera* auf *hand* gesetzt. Wenn keine Hand im *Vorfeld 2* erkannt wird oder diese sich nicht Richtung Gefahrenquelle bewegt, bleibt der Zustand von Knoten *Kamera* auf *keine hand*. Dieses binäre Ergebnis garantiert die zuverlässige Entscheidungsfindung bei der Informationsfusion. Die Zustandsvariablen und Knoten werden in Abschnitt 4.4.4 erklärt.

3.10 Fazit und Ergebnisse

In Abbildung 3.2 ist der Ablauf der entwickelten Handerkennung dargestellt. Im Folgenden wird dieser Ablauf nun detaillierter beschrieben. Die einzelnen Teilschritte sind sukzessive in diesem Kapitel erklärt worden.

In Abbildung 3.26 ist die gesamte Bildvorverarbeitung dargestellt. Diese ist im Vergleich zu Abbildung 3.8 um einen Block erweitert. In diesem Abfrageblock wird überprüft, ob bereits das 10. Bild der Kamera zur Verarbeitung anliegt. Ist dies der Fall, wird die Bewegungserkennung überbrückt und das Bild ohne weitere Einschränkungen weiterverarbeitet. Damit lässt sich eine evtl. vorhandene aber unbewegte Hand erkennen. Bewegt sich diese Hand, kann sie durch das System u. U. schneller erkannt und verfolgt werden.

Nach der Bewegungserkennung werden die Merkmale zur Handerkennung extrahiert. Dies erfolgt mittels einer modifizierten KHT, s. Abschnitt 3.5.2. Die anschließende Lokalisation und Klassifikation beruht auf probabilistischen Model-

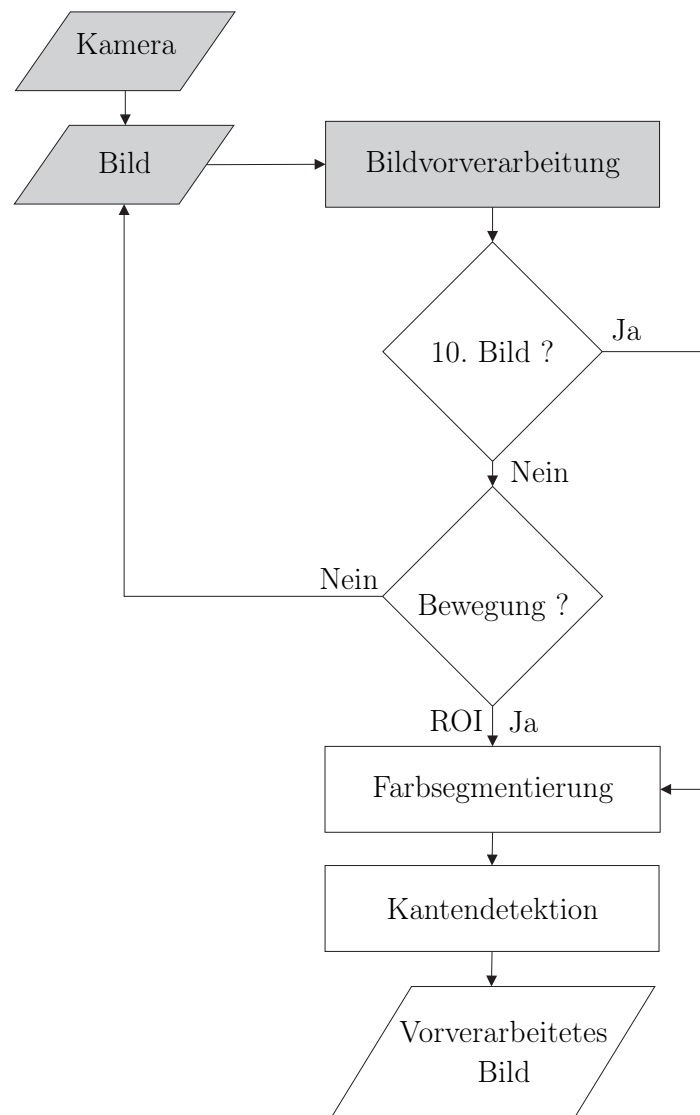


Abbildung 3.26: Ablaufdiagramm des Gesamtsystems: Vorverarbeitung

len, s. Abschnitt 3.6. Zur Auswahl des richtigen Modells wird die hier entwickelte Varianzmethode eingesetzt. Als Ergebnis dieses Bildverarbeitungsschrittes erhält man neben einer klassifizierten Hand auch deren Position und Orientierung. Der gesamte dazu notwendige Ablauf ist in Abbildung 3.27 dargestellt.

In Abbildung 3.28 ist das weitere Vorgehen nach der Lokalisation und Klassifikation einer Hand zu sehen. Durch die nunmehr vorhandenen Parameter (Handmodell, Position, Klassifikation) werden Aktive Konturen initialisiert und für die Handverfolgung mit einem Partikelfilter verwendet, s. Abschnitt 3.8. Als Partikelfilter wird der ICondensation Algorithmus eingesetzt. Nachdem eine Hand 25 Bilder lang verfolgt wurde, wird der Lokalisations- und Klassifikationsprozess erneut aufgerufen. Hiermit wird ein evtl. auftretender Handmodellwechsel erkannt, d. h. bevor das Tracking das Objekt „verliert“ wird es sicherheitshalber alle 25. Bilder neu initialisiert. Die 25 Bilder entsprechen bei einer zurzeit möglichen Verarbeitungsgeschwindigkeit von 6 fps ungefähr vier Sekunden. In dieser Zeit bewegt sich

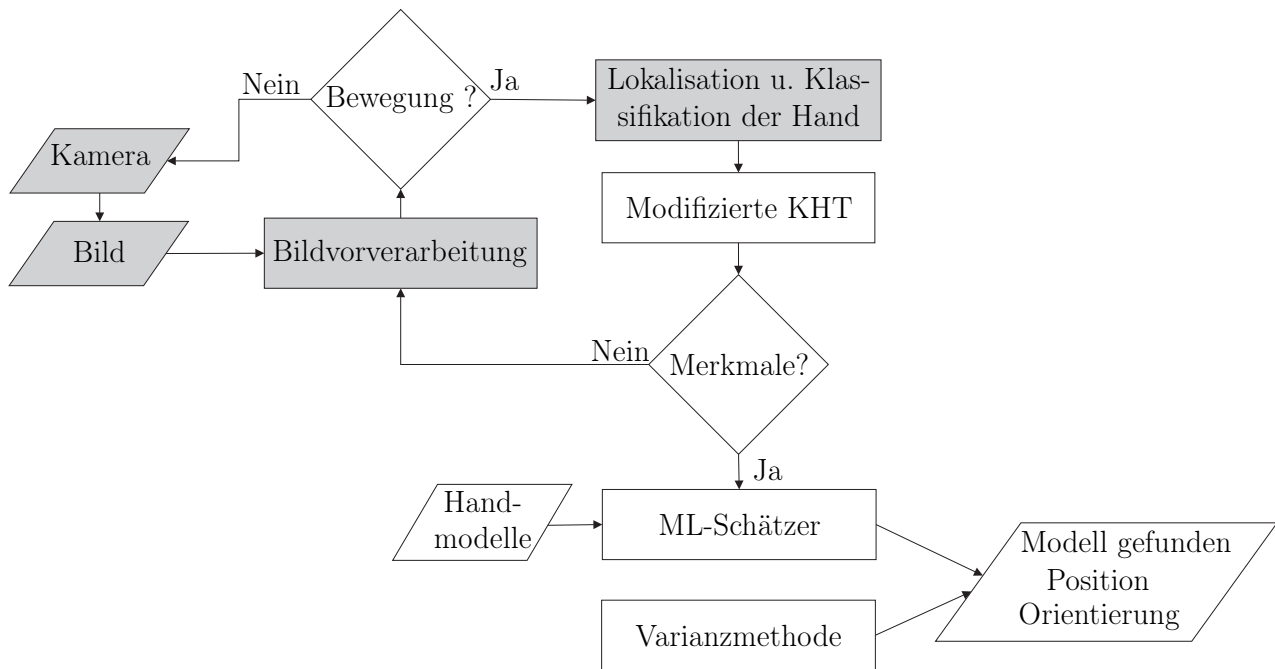


Abbildung 3.27: Ablaufdiagramm des Gesamtsystems: Lokalisation und Klassifikation

im Normalfall die Hand wieder aus dem Überwachungsbereich heraus. Des Weiteren werden die im Auswahlschritt *überlebenden* Partikel untersucht, ob sie sich alle innerhalb eines $|\mu_P \pm 3\sigma_P| < \text{Bildfläche-Bereiches}$ befinden. Ist dies nicht der Fall, bedeutet dies, dass der Partikelfilter divergiert und damit die Aktive Kontur verloren hat. Eine neue Lokalisation und Klassifikation ist notwendig. μ_P ist der Mittelwert der Partikelwolke und σ_P die Standardabweichung der einzelnen Partikel zum Mittelwert.

Der in Abbildung 3.29 dargestellte Ablauf zeigt die gesamten Bildverarbeitungsschritte zur Handerkennung. Während der Objektverfolgung wird fortlaufend eine Gefahrenanalyse erstellt. Deren Ergebnisse stehen der Informationsfusion zur Bewertung zur Verfügung.

Ergebnisse

Das oben vorgestellte System zur Gefahrenerkennung mittels Bildverarbeitung ist auf einem PC¹ in der Programmiersprache C/C++ mithilfe der OpenCV Bibliothek [40] implementiert. Als Kamera dient, wie bereits eingangs erwähnt, eine Basler A602fc [3]. Zur Bildvorverarbeitung und Merkmalsextraktion werden jeweils 70 ms pro Bild benötigt. Die Lokalisation und Klassifikation dauert abhängig von der Anzahl der gefundenen Merkmale durchschnittlich 30 ms. Das Tracking und die Gefahrenanalyse werden in 90 ms berechnet. Summiert man die jeweiligen Schritte der beiden möglichen Ablaufwege, so benötigt der Algorithmus für jeden Weg ca. $70 \text{ ms} + 70 \text{ ms} + 30 \text{ ms} = 170 \text{ ms}$ bzw. $70 \text{ ms} + 90 \text{ ms} = 160 \text{ ms}$ pro Bild.

¹Intel Pentium 4 Dual-Core Prozessor, 3 GHz, 2 GB Hauptspeicher

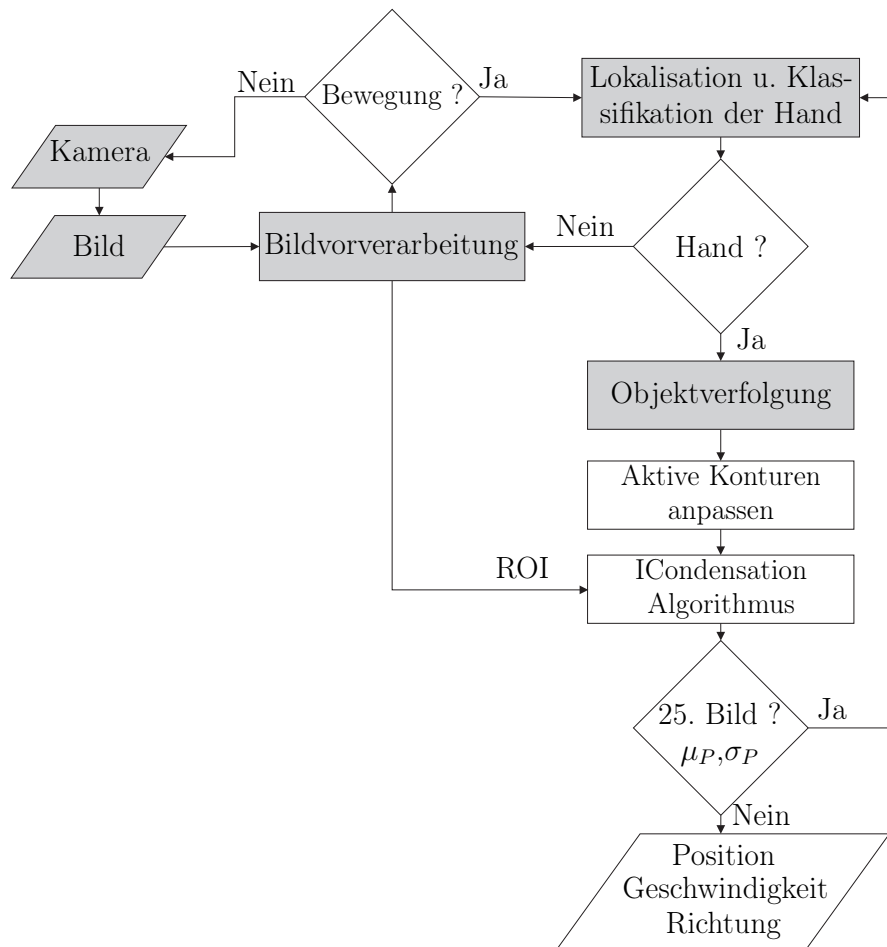


Abbildung 3.28: Ablaufdiagramm des Gesamtsystems: Objektverfolgung (Tracking)

Daraus resultiert einer Bildfrequenz von ca. 6 fps. Dies ist für die Funktionsüberprüfung und Validierung dieses Laboraufbaus ausreichend schnell. Des Weiteren wurde das Verfahren nur für die Erkennung einer Hand im Überwachungsbereich implementiert. Eine Erweiterung auf mehrere zu erkennende Hände erhöht die Komplexität des Programms und erschwert die Validierung der Einzelverfahren, die in der hier vorgestellten Methodendefinitionsphase im Vordergrund stehen. Das vorgestellte Verfahren kann jedoch ohne weitere Einschränkungen erweitert werden. So ist beispielsweise eine Erweiterung der Bewegungserkennung auf die Erfassung von mehreren, nicht zusammenhängenden Gebieten möglich ebenso wie die Bestimmung von mehreren Objekten durch die Lokalisation und Klassifikation.

In Abbildung 3.30 ist das Original- und Kantenbild einer erfolgreich lokalisierten und klassifizierten flachen Hand zu sehen. Als Hintergrund ist ein dunkles, furniertes Holz verwendet worden. Eingezeichnet sind extrahierte Merkmale, gefundenes Handmodell sowie der Bereich der Varianzmethode.

Das in Abbildung 3.31 dargestellte Szenario enthält ein helles, gemasertes Holz. In der oberen, rechten Ecke befindet sich ein Astloch, das fälschlich als Merkmal extrahiert wird. Aufgrund der Varianzmethode und den Eigenschaften des

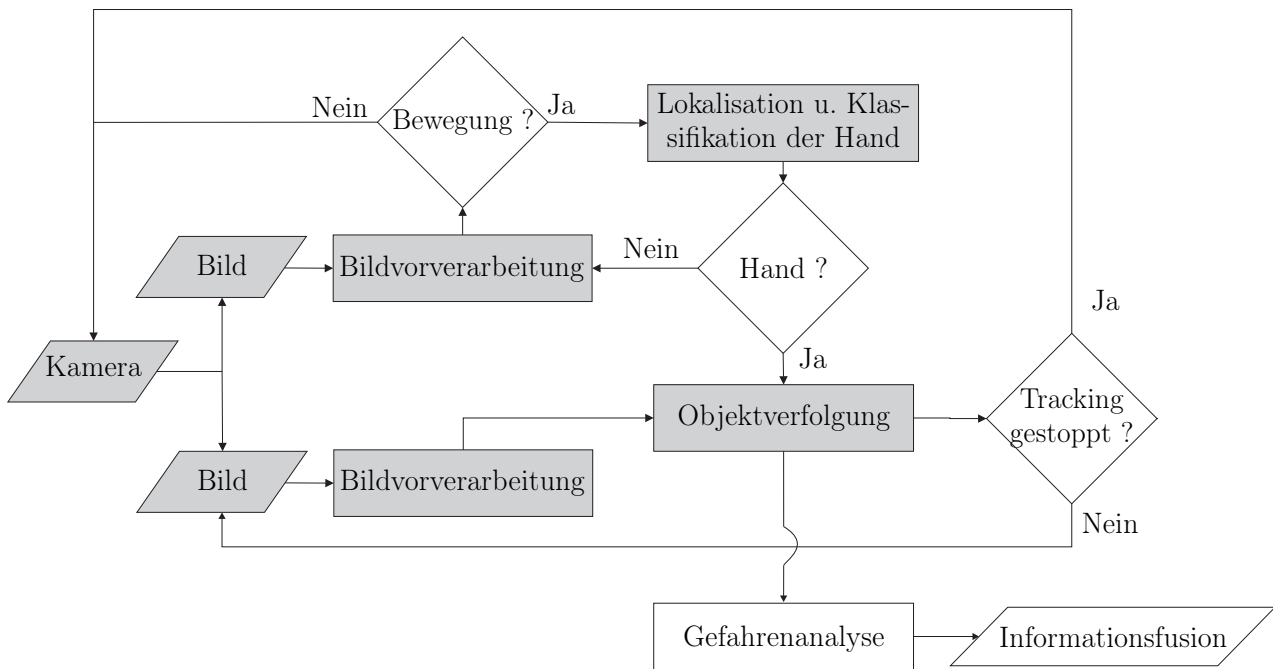
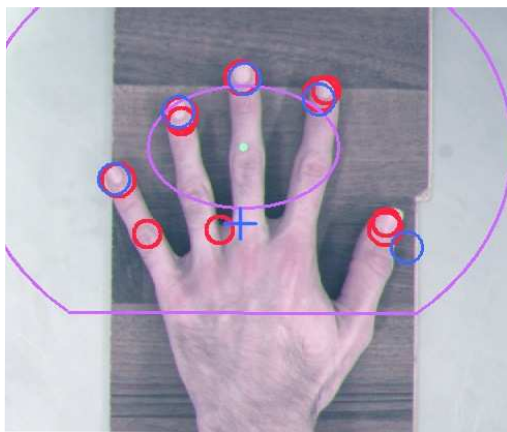


Abbildung 3.29: Ablaufdiagramm des Gesamtsystems: Gefahrenanalyse

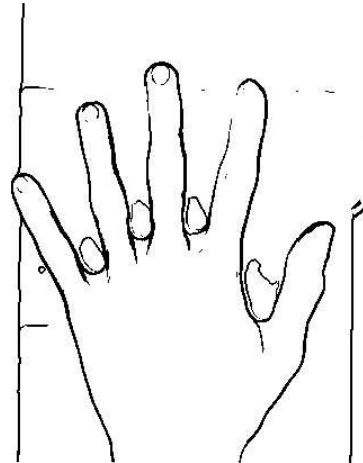
probabilistischen Modells wird dieses Merkmal bei der Lokalisation und Klassifikation ignoriert. Weiter fällt auf, dass für eine erfolgreiche Lokalisation nicht alle Fingerspitzen gefunden werden müssen. Für den Zeigefinger wurde kein Merkmal aufgrund von Beleuchtungsschwankungen extrahiert.

Befindet sich neben einer Hand auch der Hautfarbe ähnliches Holz im Erfassungsbereich der Kamera, so kann diese u. U. nicht mehr erkannt werden, s. Abbildung 3.32. Ursache ist hier die physikalische Grenze der Farbsegmentierung. Eine Trennung der Hautfarbe von der Holzfarbe ist hier nicht mehr möglich. Dadurch schlagen sowohl die Merkmalsextraktion wie auch die nachfolgende Lokalisation fehl und das System erkennt fälschlicherweise keine Hand. Abhilfe könnte hier ein Stereo-Kamerasystem zur Auswertung der Tiefeninformation oder eine Infrarot-Kamera zur Einbeziehung der Wärmeabstrahlung einer menschlichen Hand schaffen. Beide Systeme erhöhen die Systemkomplexität und den Systempreis enorm. Wird jedoch eine Hand von der Bildverarbeitung nicht erkannt, so bedeutet dies keine Einschränkung für das gesamte Handerkennungssystem, da aufgrund der Informationsfusion und den weiteren Sensoren eine Hand trotzdem detektiert werden kann, s. Abschnitt 4.4.5.

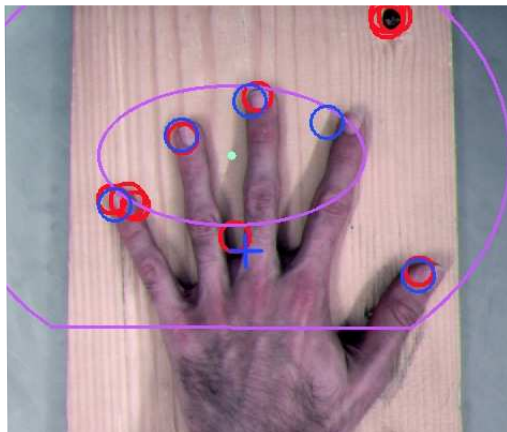
Die Leistungsfähigkeit des Systems wird anhand von ausgewählten Bildern mehrerer Bildsequenzen gezeigt. In Abbildung 3.33 sind die ersten sechs Bilder einer Sequenz dargestellt. Im ersten Bild findet die Bildvorverarbeitung, die Merkmalsextraktion (rot eingezeichnete Merkmale sowie die Lokalisation (blaue Kreise entsprechen dem Modell) und Klassifikation (lila ist der Bereich der Varianzmethode) statt. In Bild 2 ist die Initialisierung der Aktiven Kontur (Handkontur



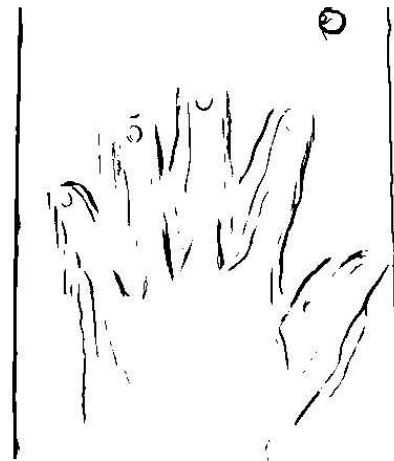
(a) Ergebnisbild



(b) Kantenbild

Abbildung 3.30: Flache Hand auf dunklem, funiertem Holz

(a) Ergebnisbild



(b) Kantenbild

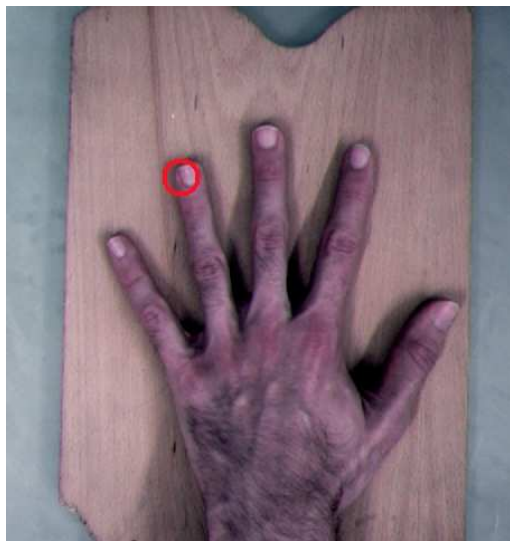
Abbildung 3.31: Flache Hand auf gemasertem Holz mit Astloch

ist rot) sowie die (blaue) Partikelwolke zu sehen. Ab Bild 3 findet die Verfolgung der erkannten Hand statt.

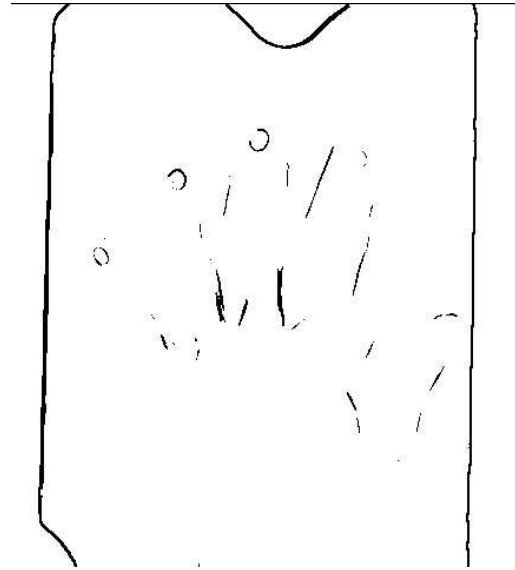
Ein Modellwechsel von einer flachen, gespreizten zu einer flachen, geschlossenen Hand ist in Abbildung 3.34 zu sehen. Das Tracking ist damit nach dem Bild 2 beendet und wird wieder nach Bild 4 gestartet. Der Modellwechsel bzw. der Handstellungswechsel wird vom System richtig durchgeführt.

In Abbildung 3.36 ist der Abbruch des Trackings dargestellt. Die Ursache ist, dass der Bereich $|\mu_p \pm \sigma_p|$ der Partikelwolke (grün) größer als $|\mu_p \pm \frac{\text{Bildbereich}}{3}|$ (lila) geworden ist.

Eine längere Handverfolgungssequenz ist in Abbildung 3.35 dargestellt. Dabei wird eine sich im Kreis bewegende Hand verfolgt. Abgebildet ist nur jeder zweite Bild, da die Positionsunterschiede der Hand ansonsten kaum darzustellen sind. Die Handverfolgung verläuft einwandfrei.



(a) Ergebnisbild



(b) Kantenbild

Abbildung 3.32: Flache Hand auf Buchenholz

In Abbildung 3.32 ist ein Beispiel aufgezeigt worden, bei dem aufgrund der Grenzen der Farbsegmentierung der Verarbeitungsalgorithmus nicht funktioniert. Wird die Hand jedoch einmal erfolgreich erkannt und lokalisiert, ist deren Verfolgung auch auf diesem problematischen Hintergrund möglich. Dies ist in Abbildung 3.37 zu sehen. Grund hierfür sind die Eigenschaften des Condensation Algorithmus in Kombination mit den Aktiven Konturen. In den Aktiven Konturen ist das a priori vorhandene Wissen über die Handform hinterlegt. Des Weiteren werden mit dem Partikelfilter ständig 50 mögliche Konturpositionen geschätzt und durch Messung (Merkmalsextraktion) bestätigt. Selbst durch einen komplexen Hintergrund verursachte Messstörungen, werden innerhalb eines bestimmten Rahmens durch die multimodale Wahrscheinlichkeitsdichte des Condensation Algorithmus kompensiert. Dadurch ist die Objektverfolgung auch bei diesem problematischen Szenario weiterhin möglich.

Die zu Abbildung 3.37 gehörende Kantensequenz ist in Abbildung 3.38 dargestellt. Die Merkmalsextraktion bzw. Fingerspitzenenerkennung würde bei dieser Sequenz ab Bild 9 versagen. Da der Messschritt des Condensation Algorithmus nur das Kantenbild in Kombination mit der Aktiven Kontur auswertet, sind weiterhin genügend Referenzpunkte zur Bestätigung der Handposition vorhanden. Bild 9 und 19 zeigen den gesamten Bildbereich aufgrund des Überspringens von jedem zehnten Bild, s. Abbildung 3.26. In den restlichen Bildern ist zusätzlich der ROI zu sehen.

In Tabelle 3.2 sind die Validierungsergebnisse der Lokalisation und Klassifikation nach dem Ablauf in Abbildung 3.27 aufgezeigt. Die Tests wurden unter den Standardbeleuchtungsbedingungen am Versuchsstand durchgeführt (Tageslicht und künstlicher Beleuchtung). Bei *Holz 1* handelt es sich um eine helle Spanplatte,

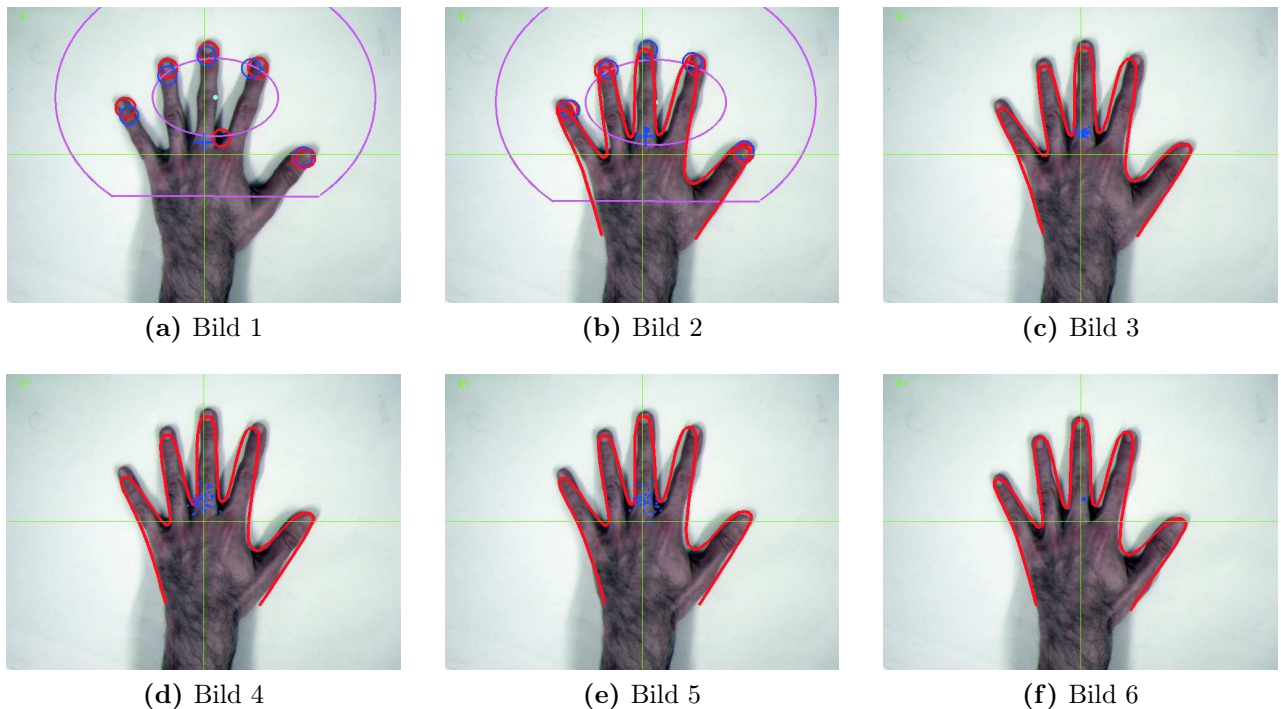


Abbildung 3.33: Initialisierung des Trackings und Verfolgung einer flachen Hand

Hintergrund	Hand: flach, gespreizt	Hand: flach, geschlossen	Faust	Gesamt
Holz 1	88 %	63 %	88 %	79 %
Holz 2	98 %	94 %	57 %	83 %
Holz 3	50 %	44 %	75 %	56 %
Metall	98 %	82 %	94 %	91 %
Gesamt	84 %	71 %	78 %	-

Tabelle 3.2: Ergebnis der Lokalisation und Klassifikation

Holz 2 ist ein Holz mit dunklem Furnier und *Holz 3* ist Buchenholz, das eine der Haut ähnliche Farbe besitzt. In den Beobachtungsbereich der Kamera wurden jeweils rechte und linke Hände, den drei Handmodellen entsprechend, nacheinander eingebracht. Die Hände wurden aus acht unterschiedlichen Richtungen in den Bereich bewegt sowie lateral verschoben. Mit diesen Tests ist eine ausreichende Validierung der Lokalisation und Klassifikation möglich, die eine zentrale Rolle in der Handidentifikation und deren Positionsbestimmung einnimmt. Die in Tabelle 3.2 dargestellten Ergebnisse zeigen, dass der entwickelte Ablauf zuverlässig eine Hand lokalisieren und einem Modell zuweisen kann.

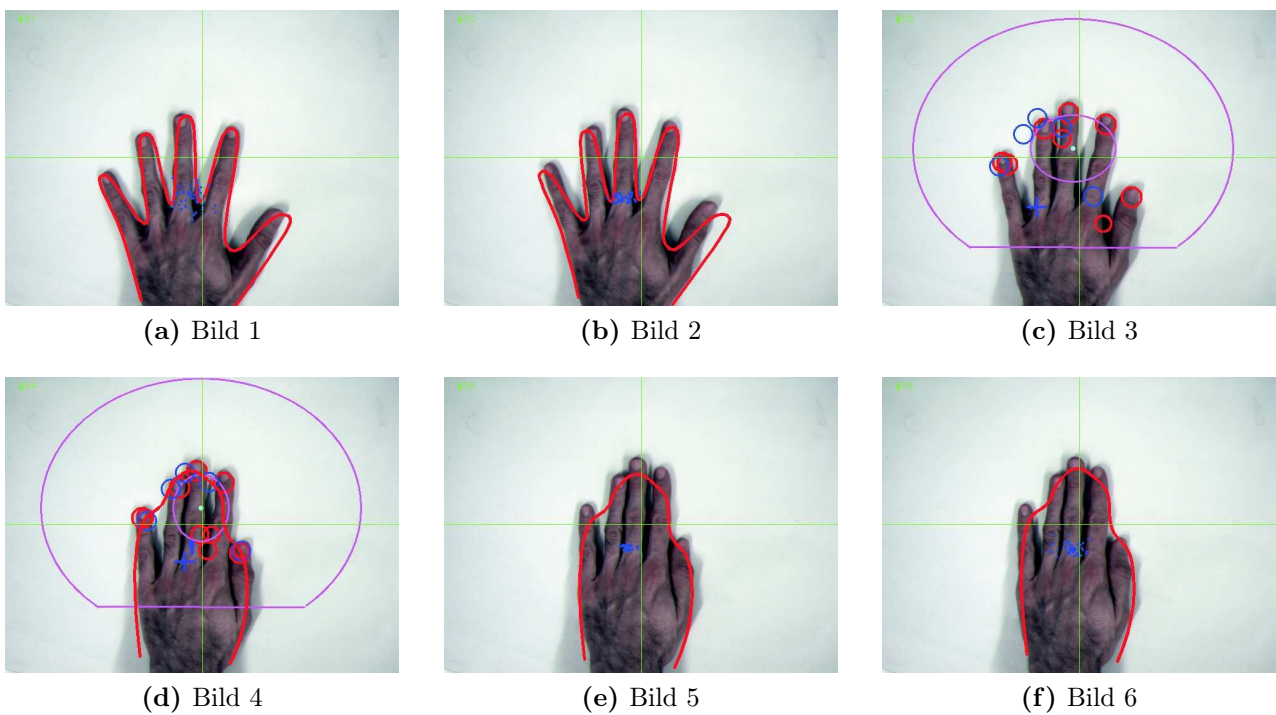
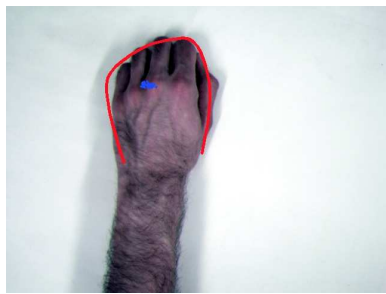


Abbildung 3.34: Modellwechsel von flacher, gespreizter zu flacher, geschlossener Hand



(a) Bild 1



(b) Bild 3



(c) Bild 5



(d) Bild 7



(e) Bild 9



(f) Bild 11



(g) Bild 13



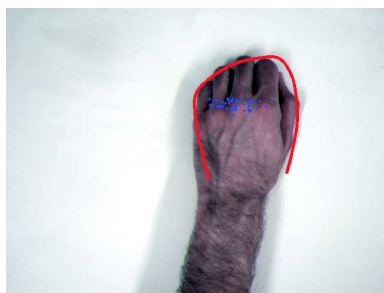
(h) Bild 15



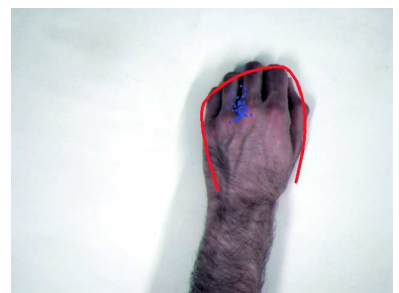
(i) Bild 17



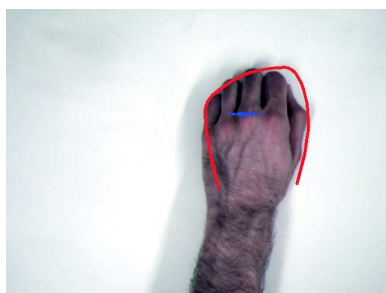
(j) Bild 19



(k) Bild 21



(l) Bild 23



(m) Bild 25



(n) Bild 27



(o) Bild 29

Abbildung 3.35: Verfolgung einer sich im Kreis bewegendes Faust

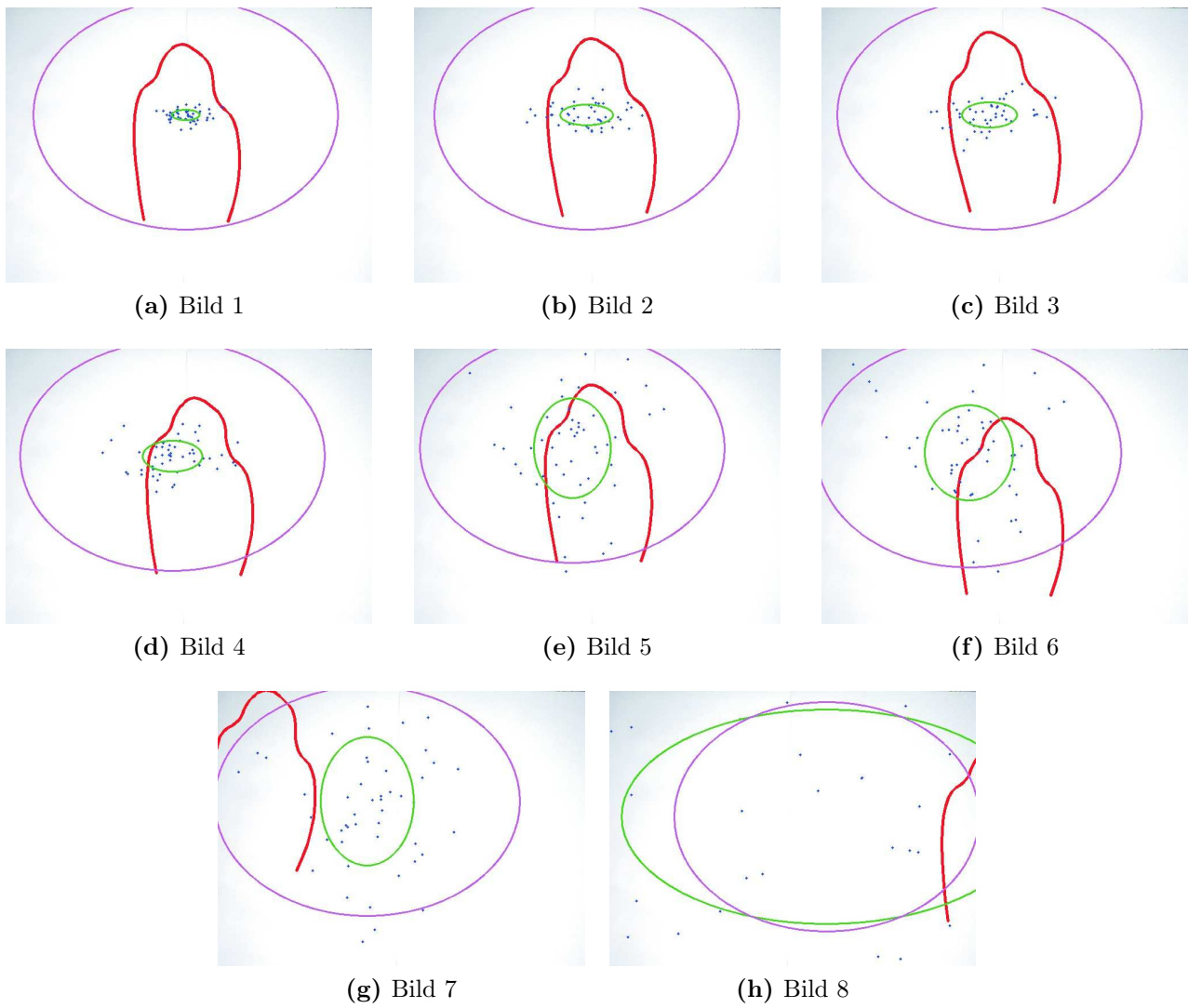


Abbildung 3.36: Verlust der Hand und Abbruch des Trackings in Bild 8

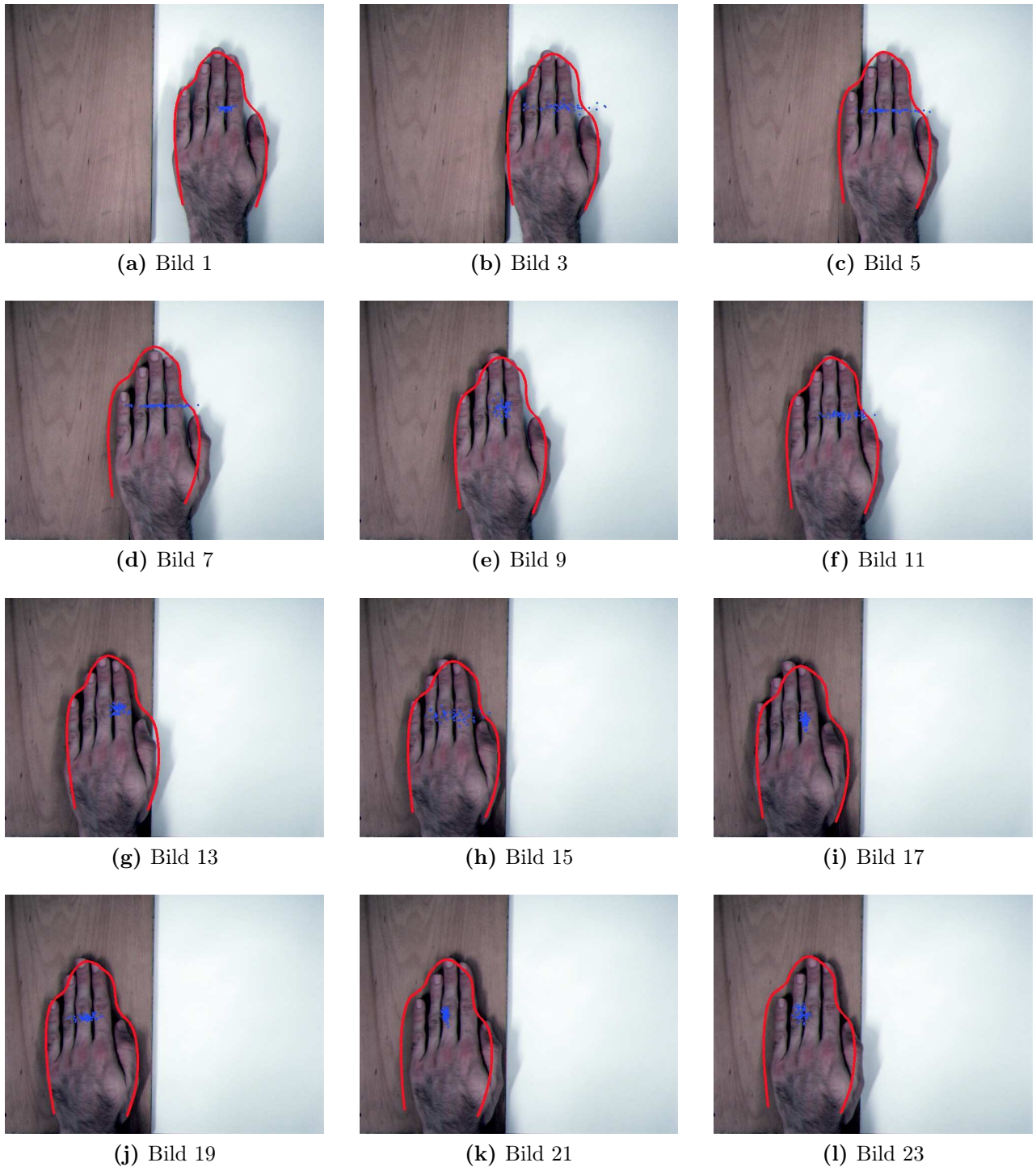


Abbildung 3.37: Verfolgung einer sich geradlinig bewogender Hand mit zweigeteiltem Hintergrund

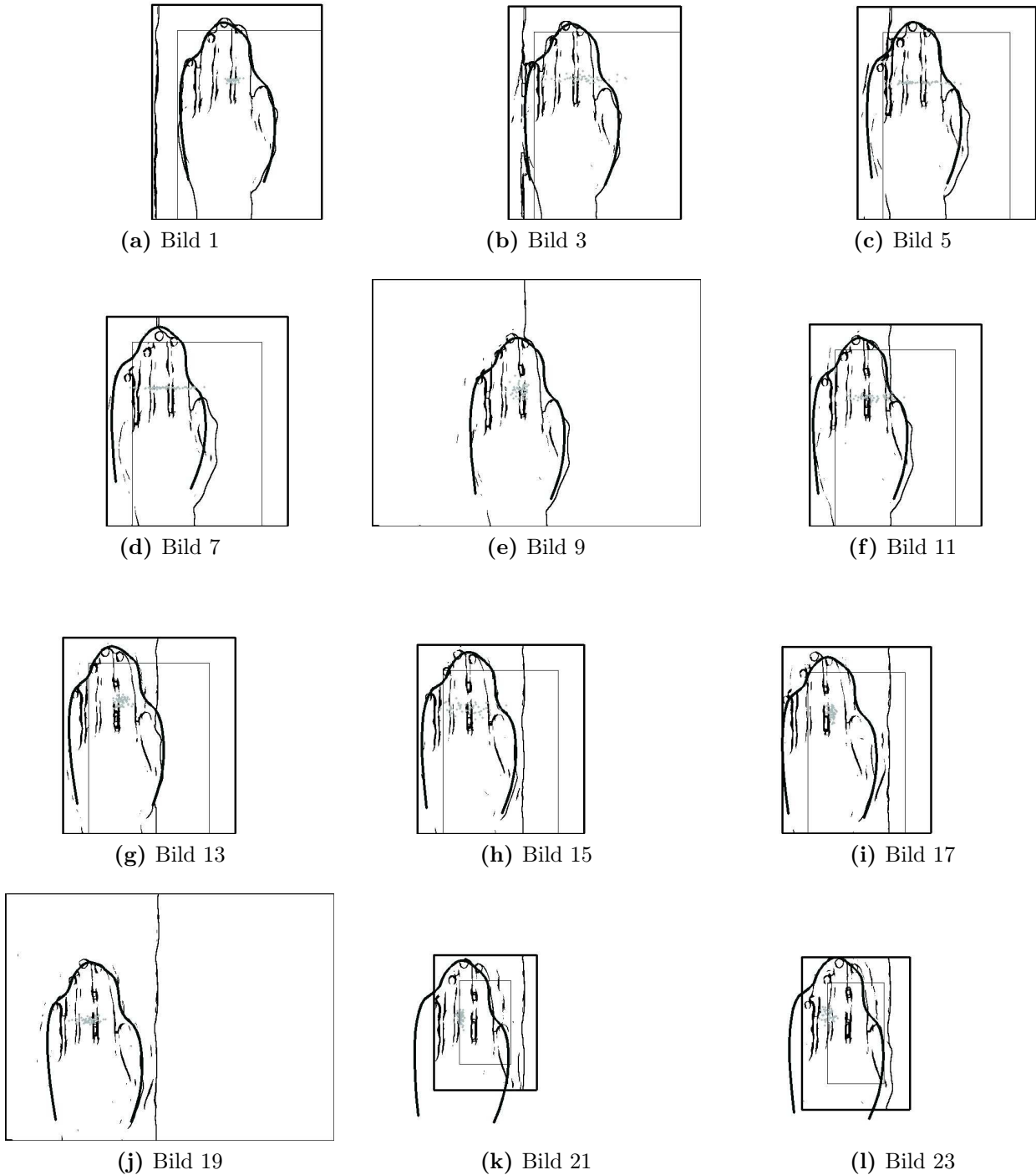


Abbildung 3.38: Kantenbilder: Verfolgung einer sich geradlinig bewogender Hand mit zweigeteiltem Hintergrund

4 Informationsfusion zur Gefahrenanalyse

Das menschliche Gehirn ist in der Lage, die Informationen und Wahrnehmungen der Sinnesorgane parallel zu verarbeiten und darüber hinaus auch die einzelnen Informationen miteinander zu verknüpfen. Dadurch ist ein Mensch fähig auf unvorhersehbare Ereignisse entsprechend zu reagieren, ein Beispiel hierfür ist Treppensteigen. Dazu sind u. a. die Augen, der Gleichgewichtssinn sowie die taktile Wahrnehmung notwendig. Darunter ist die Vibrations-, Beschleunigungs- und Druckwahrnehmung im Körper zu verstehen. Bei ausreichender Beleuchtung sind die Augen die primäre Informationsquelle, um den Bewegungsablauf zu steuern. Fällt die Beleuchtung plötzlich aus, so kann ein Mensch ohne größere Schwierigkeiten die anderen Sinne zur Bewegungskoordination nutzen.

Technische Systeme können nur bedingt auf plötzliche Situationsänderungen reagieren bzw. Sensorinformationen situationsbedingt richtig auswerten. Sie müssen dazu konstruiert und programmiert sein. Mit den Grundlagen und Methoden hierzu beschäftigen sich die Kybernetik und die künstliche Intelligenz. Ein Teilgebiet davon ist die Informationsfusion.

Die Informationsfusion hat zum Ziel, aus unterschiedlichen Sensoren und Informationsquellen, neues und genaueres Wissen über physikalische Parameter, Ereignisse und Situationen zu gewinnen [9]. Dies erfolgt durch Integration von Informationen aus verschiedenen Datenquellen, um die inhärente Unsicherheit der Sensorsignale zu beheben. Die inhärente Unsicherheit bei der Signalinterpretation entsteht durch den prinzipiellen Aufbau eines Sensors, da entscheidende Merkmale bauartbedingt nicht erfasst werden können und daher ein Informationsdefizit bleibt.

Ein Beispiel für die Entstehung von Unsicherheit bei der Deutung eines Signals ist die Auswertung der Sensordaten eines thermopilen Sensors, s. Abschnitt 2.2. Basierend auf diesen Daten wird auf eine Hand geschlossen. Dies ist nur mit einer großen Unsicherheit möglich, da nicht bestimmt werden kann, ob die gemessene Temperaturänderung durch eine Hand oder eine andere Wärmequelle verursacht wird. Durch die Hinzunahme der Information aus einem kapazitiven Sensor kann diese Unsicherheit verringert werden.

Das Prinzip der Informationsfusion zur Gefahrenanalyse, also der Hinzunahme zusätzlicher Information in obigem Beispiel, wird in diesem Kapitel erläutert. Dazu werden im Folgenden mögliche Fusionsstrukturen und -methoden präsentiert. Eine mögliche Fusionsmethode verwendet Bayes'sche Netze zur Informationsverarbeitung (Abschnitt 4.2). Die dazu notwendige Vorverarbeitung der Sensordaten

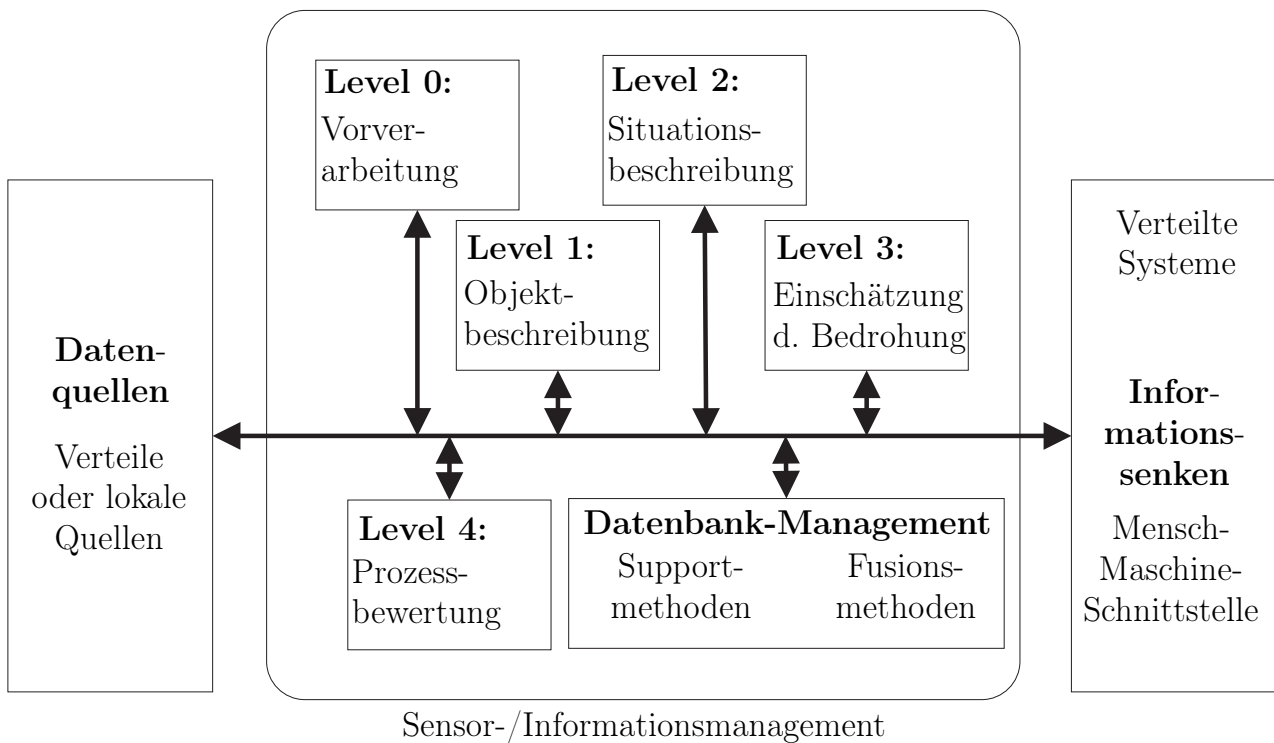


Abbildung 4.1: JDL-Modell [29]

wird in Abschnitt 4.3 dargestellt. Abschließend steht die entwickelte Informationsfusion zur Gefahrenanalyse mittels eines Bayes'schen Netzes im Zentrum der Betrachtung. Einen Überblick über das gesamte System bietet Abbildung 1.7 auf Seite 13.

4.1 Fusionsstrukturen zur Gefahrenanalyse

Die Informationsfusion wird in vielen Gebieten zur Datenverarbeitung eingesetzt, u. a. in der Verteidigungstechnik, Medizintechnik, Robotik und Fahrerassistenz im Auto. Die Anforderungen an die Methoden und die Struktur der Informationsverarbeitung sind je nach Anwendung vielfältig. Eine Übersicht findet sich in den folgenden Abschnitten. Es werden verschiedene Fusionsmodelle, -strukturen und -methoden vorgestellt.

4.1.1 Fusionsmodelle

Das aus der Militärtechnik stammende JDL-Modell (Joint Directors of Laboratories-Modell) in Abbildung 4.1 bildet die Grundlage für eine systematische Einteilung von Fusionstechniken [29, 9].

Die *Datenquellen* stehen am Anfang der Informationsverarbeitung und sind entweder lokal verfügbar oder räumlich verteilt.

Level 0 Vorverarbeitung: Die Rauschunterdrückung, Digitalisierung und Synchronisation der Daten, stellt den ersten Schritt einer Informationsfusion dar.

Level 1 Objekterkennung und Zustandsschätzung durch Sensordatenfusion. In diesem Schritt erfolgt die Objektdetektion und -klassifikation, ebenso wie die Verfolgung der Objekttrajektorien. Dadurch erhält man ein Situationsbild, das auch eine Prädiktion des Objektverhaltens ermöglicht.

Level 2 Situationsanalyse: Die erkannten Objekte werden durch eine zeitliche Analyse einem Szenario zugeordnet. Dies lässt eine Interpretation der Bewegungsmuster und der Objektdynamik zu.

Level 3 Bedrohungsanalyse: Nachdem die „Absichten“ der detektierten Objekte ermittelt wurden, werden mögliche Handlungsoptionen im Sinne einer Risikominimierung ermittelt.

Level 4 Prozessoptimierung: Diese erfolgt durch eine Evaluation der Fusion und Interaktion zwischen den Ebenen.

Mensch-Maschine-Schnittstelle: Diese stellt die Kommunikation des Systems mit einem Anwender dar. Hierbei werden zum einen Fusionsergebnisse präsentiert und zum anderen kann der Nutzer selbst in die Entscheidungsfindung eingreifen bzw. das System bei der Entscheidungsfindung unterstützen.

Omnibus-Modell

Das JDL-Modell erklärt die Informationsverarbeitung als hierarchische Verarbeitungsschritte. Eine allgemeine Darstellung der Informationsfusion bietet das Omnibus-Modell [6]. Hierbei wird die Informationsverarbeitung als Kreislauf dargestellt.

Die vier Schritte der Informationsverarbeitung sind in Abbildung 4.2 dargestellt und sind im Einzelnen:

- **Beobachten:** Messen, Signalerfassung und -verarbeitung
- **Ausrichten:** Merkmalsfusion mit vorangegangener Merkmalsextraktion und Mustererkennung
- **Entscheiden:** Kontextverarbeitung und Entscheidungsfindung
- **Handeln:** Steuern/Regeln und Ressourceneinsatz

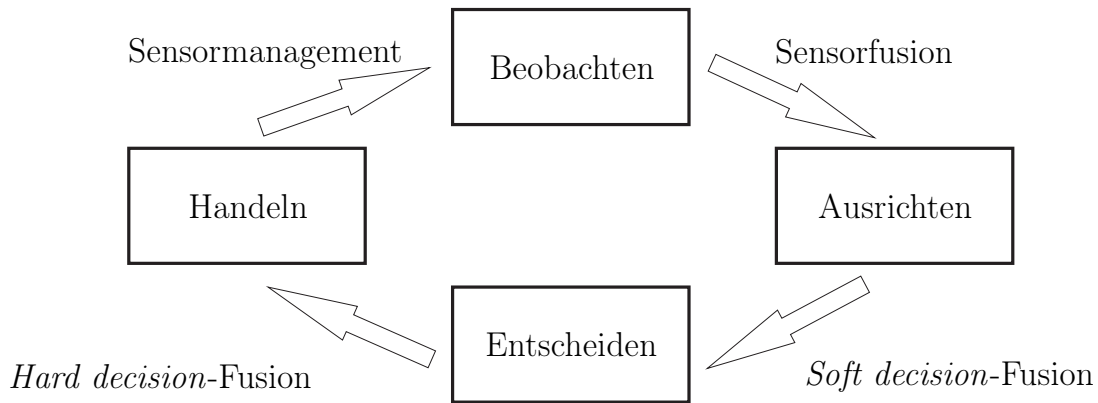


Abbildung 4.2: Omnibus-Modell [6]

Diese Schritte werden kontinuierlich durchlaufen und abgearbeitet. Die Übergänge stellen die einzelnen Fusionschritte dar: Sensordatenfusion, *soft* und *hard decision* und Sensorverwaltung.

Der *soft decision* Schritt liefert basierend auf der Sensorrohdatenfusion z. B. eine gewichtete Abfolge von Konfidenzfaktoren. Eine *hard decision* hingegen entscheidet sich für eine Aussage oder Identität.

4.1.2 Sensorintegration

Sensorinformationen in einem Multisensorsystem liegen entweder komplementär oder redundant vor [9]. Durch redundante Sensoren wird z. B. die Unsicherheit einzelner Sensorsignale reduziert bzw. das Signal-zu-Rauschverhältnis gesteigert. Dadurch sind Aussagen über die Plausibilität des Fusionsergebnisses möglich, was einer Steigerung der Zuverlässigkeit des Sensorsystems gleichkommt. Bei komplementären Informationen können die einzelnen Sensorinformationen dazu genutzt werden, um die unvollständigen Informationen einzelner oder redundanter Sensoren zu ergänzen. Hierdurch werden Mehrdeutigkeiten verhindert und der Reaktionsbereich des Systems wird vergrößert [63]. Der *Reaktionsbereich* bedeutet in diesem Zusammenhang die Fähigkeit zur Erfassung von unterschiedlichen Objektmerkmalen.

Die Kombination von einzelnen Sensorinformationen ist abhängig vom Einsatzort und der Art der verfügbaren Datenquellen. Drei mögliche Konzepte zu deren Integration sind im Folgenden aufgeführt:

- *Konkurrierende Integration*: Kombination gleichartiger, redundanter Sensoren mit gleicher Information und dem Ziel, Unsicherheiten zu reduzieren oder Rauschen zu unterdrücken.
- *Komplementäre Integration*: Kombination von unterschiedlicher Information von gleichartigen Sensoren, damit fehlende Information ausgeglichen werden kann und eine vollständige Beschreibung des Szenarios vorliegt.

- *Kooperative Integration*: Sensorinformationen liegen verteilt vor und können auch unterschiedlicher Art sein. Die Daten aller Sensoren werden ausgewertet, damit auf eine Messgröße geschlossen werden kann oder eine Entscheidung erzielt werden kann, die ein einzelner Sensor nicht zur Verfügung stellen kann.

4.1.3 Abstraktionsebenen

Eine Informationsfusion kann in unterschiedlichen Abstraktionsebenen durchgeführt werden. Dabei ist die Fusion auf höheren Ebenen meist effizienter, da jedoch durch die davor erfolgte Datenreduktion auch Informationen verloren gehen können, ist das erzielte Ergebnis nicht unbedingt besser oder zuverlässiger [9].

Die unterste Abstraktionsebene (*Signalebene*) ist die direkte Verarbeitung der Sensorsignale miteinander. Dazu muss eine Vergleichbarkeit und Synchronisation der Sensorsignale gewährleistet sein. In der nächst höheren Ebene, der so genannten *Merkmalebene*, werden Signaldiskriptoren oder Merkmale miteinander fusioniert. In der höchsten Abstraktionsebene (*Symbolebene*) werden bereits erzielte Detektions- oder Klassifikationsergebnisse vereinigt, um Entscheidungen treffen zu können.

4.1.4 Fusionsarchitektur

Eine Datenfusion erfolgt entweder mit einem zentralen Entscheidungszentrum oder verteilt mit lokalen Entscheidungszentren. Auch Mischformen von beiden Varianten treten auf [29, 89]. Bei einer zentralen Datenfusion werden alle erfassten Daten an das Entscheidungszentrum übertragen. Dadurch gehen keine Informationen verloren, jedoch ist eine hohe Bandbreite der Kommunikationseinrichtung zwischen Sensoren und Zentrum erforderlich. Eine Reduktion der Komplexität und des Übertragungsaufwandes wird erreicht, indem lokale Entscheidungszentren die Sensorwerte verarbeiten und die Endergebnisse an eine zentrale Einheit übermitteln.

4.1.5 Methoden zur Informationsfusion

Für die Sensordatenfusion wird zwischen gitterbasierten und parametrischen Ansätzen unterschieden. Gitterbasierte Verfahren werden hauptsächlich in der Robotik zur Fahrwegeplanung und Navigation verwendet. Dabei werden Belegungswahrscheinlichkeiten einer Rasterstruktur ermittelt und ausgewertet. Diese Methode wird hier nicht weiterverfolgt. Sie ist in [9] ausführlich erklärt.

Die parameterbasierten oder numerischen Ansätze zur Sensordatenfusion lassen sich in folgende Kategorien einteilen [29]:

- *Merkmalsbasierte Ansätze*: z. B. gewichtete Mittelwerte, Kalman-Filter.
- *Probabilistische Ansätze*: z. B. klassische und Bayes Statistik, Dempster-Shafer Theorie.
- *Heuristische Ansätze*: z. B. Voting Logik, Scoring- und Ordinal-Ranking-Techniken.
- *Neuronale Ansätze*: z. B. Neuronale Netze, Neuro-Fuzzy-Methoden
- *Wissensbasierte Ansätze*: z. B. Fuzzy Logik.

Im Folgenden werden einige der oben genannten, repräsentativen Methoden genauer vorgestellt. Weitere und detaillierte Informationen finden sich z. B. in [54] und [55].

Kalman-Filter

Das *Kalman-Filter* wird zur Fusion auf Signalebene bzw. von redundanten Rohdaten eingesetzt. Das rekursive Schätzverfahren bestimmt aus verrauschten Signalen die unbekannt Zustände eines dynamischen Prozesses. Zur Anwendung des Kalman-Filters wird ein Sensor- und Umgebungsmodell benötigt. Diese sind oftmals jedoch nicht vorhanden oder können erst gar nicht erstellt werden. Des Weiteren sind geeignete Anfangswerte nötig, deren Bestimmung unter Umständen schwierig ist [13], s. auch Abschnitt 3.8.

Klassische Statistik

Die klassische Statistik beruht auf der Auswertung von empirisch bestimmten Häufigkeitsverteilungen. Hierbei werden jedoch nur die Sensordaten x_{sen} als Zufallsgrößen interpretiert. Die eigentliche Messgröße y dagegen nicht. Zur Schätzung der Messgröße durch die Sensordaten wird eine so genannte Likelihoodfunktion $p_{\mathbf{x}_{sen}|y}(\mathbf{x}_{sen}|y)$ gebildet. Die Schätzung \hat{y}_{MLS} erfolgt dann mittels einer Maximum-Likelihood-Schätzung [48]

$$\hat{y}_{MLS} = \arg \max_y \{p_{\mathbf{x}_{sen}|y}(\mathbf{x}_{sen}|y)\} .$$

Bayes Statistik

Bei der Bayes Statistik wird die Messgröße ebenfalls als Realisierung einer Zufallsgröße betrachtet, so dass die a priori Wahrscheinlichkeitsdichte $p_y(y)$ zur Bestimmung der a posteriori Wahrscheinlichkeitsdichte $p_{y|\mathbf{x}_{sen}}(y|\mathbf{x}_{sen})$ verwendet wird [56, 5]

$$p_{y|\mathbf{x}_{sen}}(y|\mathbf{x}_{sen}) = \frac{p_{\mathbf{x}_{sen}|y}(\mathbf{x}_{sen}|y) \cdot p_y(y)}{p_{\mathbf{x}_{sen}}(\mathbf{x}_{sen})} . \quad (4.1)$$

Eine *Maximum a Posteriori*-Schätzung des zu ermittelnden Parameters y ergibt

$$\hat{y}_{MAP} = \arg \max_y \{p_{y|\mathbf{x}_{sen}}(y|\mathbf{x}_{sen})\} .$$

Dieses Vorgehen hat den Vorteil, dass bei gegebenen Sensordaten x_{sen} eine Wahrscheinlichkeitsverteilung für die zu schätzende Größe y bestimmbar ist. Bei der klassischen Statistik kann lediglich eine Wahrscheinlichkeitsverteilung der Sensordaten x_{sen} bei gegebenen Parameter y erfolgen.

Bei der klassischen wie auch bei der Bayes Statistik werden nur disjunkte Klassen bzw. sich gegenseitig ausschließende Hypothesen oder Zustände betrachtet. Ungewissheit kann mit dieser Art der Informationsverarbeitung nicht betrachtet werden.

Dempster-Shafer-Evidenztheorie

Die *Dempster-Shafer-Evidenztheorie* [78] wird oftmals als eine Verallgemeinerung der Bayes Statistik betrachtet. Sie bietet die Möglichkeit Ungewissheit exakter auszudrücken. Dazu wird ein Grad des Dafürhaltens (englisch: *degree of belief*), eine Plausibilität und Wahrscheinlichkeitsmassen eingeführt. Dadurch ist es möglich redundante Behauptungen oder Informationen zu integrieren, was bei der Bayes Statistik nicht möglich ist. Oftmals ist jedoch detailliertes Wissen über die Zuverlässigkeit der Sensoren oder eine genaue Sensorbeschreibung und deren Interaktionen nicht verfügbar.

Voting Logik

Die Voting Logik zählt zu den heuristischen Methoden und vereinigt die Vorteile einer parallelen und seriellen Informationsverarbeitung, so dass eine hohe Detektionsrate bei gleichzeitig hoher Zuverlässigkeit erreichbar ist [55].

Die aus den Sensordaten erzeugten Merkmale werden Vertrauensebenen zugeordnet. Diese Vertrauensebenen werden in bestimmte Detektionsmodi zusammengefasst, entsprechende Hypothesen gebildet und daraus eine Detektions- bzw. Identitätswahrscheinlichkeit berechnet. Die Voting Logik arbeitet demokratisch, d. h. je mehr Sensoren eine Hypothese unterstützen, desto wahrscheinlicher ist es, dass diese Hypothese als Ergebnis ausgegeben wird.

Zur Bildung der Vertrauensebene werden nicht-adaptierbare Schwell- und Grenzwerte eingesetzt. Bei sich ändernden Umgebungsparametern oder einer Erweiterung des Systems wirkt sich dies nachteilig aus, da Teile des Systems erneut berechnet und abgeändert werden müssen [105].

Neuronale Netze

Die Nutzung von *Neuronalen Netzen* [28] zur Informationsfusion ist besonders dann sinnvoll, wenn es nicht, oder nur schwer möglich ist, einen Algorithmus zur Verarbeitung der Sensordaten zu finden. Neuronale Netze werden mithilfe einer Trainingsphase unter Verwendung von Testdaten erstellt. Ein großer Nachteil dieses Verfahrens ist, dass a priori Wissen über die Sensordaten nicht mit in die Fusion eingebracht werden kann. Des Weiteren kann es bei redundanten Testdaten oder einseitig gewählten Lerndaten zu einer Überanpassung des Netzes kommen, so dass die Reaktion auf unbekannte Eingangsdaten nicht vorhersehbar ist.

Fuzzy Logik

Die Fuzzy Logik basiert auf einer Verallgemeinerung der klassischen Mengenlehre [53]. Hierbei wird mittels einer Zugehörigkeitsfunktion für jedes Element ein Zugehörigkeitsgrad zu einer Menge bestimmt. Die verschiedenen Mengen werden über Inferenzen miteinander kombiniert.

Die Fuzzy Logik erlaubt durch die beliebige Wahl der Zugehörigkeitsfunktionen und der Kombinationsregeln eine subjektive Repräsentation von Wissen, z. B. das menschliche Wissen über einen Zusammenhang. Dadurch können Ungewissheit über die Sensordaten und Abschätzungen von Sensorinformationen für eine Informationsfusion berücksichtigt werden.

Die Repräsentation von Wissen erfolgt oftmals mithilfe von Datenbanken, da bei umfangreich vorhandenem Expertenwissen sich eine Vielzahl von Kombinationsregeln ergeben. Eine Datenbanksuche oder die Auswertung einer großen Regelbasis mit anschließender Defuzzifizierung erfordert erheblichen Berechnungsaufwand und Speicherbedarf bei der Verarbeitung.

Fazit

Die hier angesprochenen Verfahren eignen sich für unterschiedliche Anwendungsfälle, wie z. B. Trajektorien-schätzung mit Kalman-Filter oder Mustererkennung mit Neuronalen Netzen. Eine Datenintegration zur Gefahrenanalyse ist mit einem einzelnen der obigen Verfahren nicht durchführbar. Die Gründe hierfür wurden bei den Verfahren aufgeführt. Die Kombination einiger Verfahren oder die Erstellung eines auf den Verfahren beruhenden Expertensystems liefert ebenfalls keine zielführende Lösung, aufgrund von dafür erforderlichen Rechen- und Speicherkapazität eines solchen Systems.

Die im Folgenden vorgestellten Bayes'schen Netze bieten einen Lösungsansatz zur effizienten, zuverlässigen und robusten Informationsfusion. Weitere Folgerungen und Anforderungen werden in Abschnitt 4.4.1 behandelt.

4.2 Bayes'sche Netze

Bayes'sche Netze bieten die Möglichkeit, Zusammenhänge und Systemvorwissen mithilfe von probabilistischen, graphischen Modellen darzustellen. Dies wird z. B. bei Systemen mit künstlicher Intelligenz und Expertensystemen ausgenutzt. Des Weiteren können effizient Rückschlüsse auf Ursachen von Beobachtungen gezogen werden [72, 71] und damit unter anderem Diagnosesysteme und Systeme zur Entscheidungsfindung aufgebaut werden. Bayes'sche Netze haben außerdem den Vorteil, dass deduktive und abduktive Schlussfolgerungen im Gegensatz zu klassischen Wahrscheinlichkeitsmodell, wie der klassischen oder Bayes Statistik, möglich sind. Dazu werden geeignete Inferenzalgorithmen genutzt. Bayes'sche Netze sind im Allgemeinen modular, so dass eine Netzerweiterung aufgrund von z. B. zusätzlichen Sensoren einfach durch Hinzufügen eines neuen Knotens erfolgen kann ohne dass das gesamte Netz neu aufgebaut und berechnet werden muss.

Mit einem Bayes'sches Netz kann ebenfalls eine Informationsfusion durchgeführt werden. Bei Betrachtung des Omnibus-Modell (s. Abbildung 4.2) als Darstellung einer Informationsfusion ist zu erkennen, dass alle Schritte außer dem Sensormanagement von einem Bayes'schen Netz durchgeführt werden können, s. auch Abschnitt 4.4.1.

In diesem Abschnitt werden einige notwendige Grundlagen für Bayes'sche Netze vorgestellt sowie deren Definition und Entwurf. Im Anhang E ist ein einfaches, erklärendes Beispiel zu einem Bayes'schen Netz aufgeführt.

4.2.1 Grundlagen

Im Folgenden werden einige für das Verständnis von Bayes'schen Netze notwendige Grundlagen aus der Wahrscheinlichkeits- und Graphentheorie vorgestellt. Weitere und detaillierte Informationen sind in [48] und [51] enthalten.

Wahrscheinlichkeitstheorie

In Bayes'schen Netzen spielen bedingte Wahrscheinlichkeiten und die Bayes'sche Statistik die grundlegende Rolle, weshalb diese hier vorgestellt werden.

Bedingte Wahrscheinlichkeit: Damit lässt sich ausdrücken, wie wahrscheinlich das Eintreten von Zustand A unter der Bedingung ist, dass Zustand B wahr ist.

Definition 4.1 (Bedingte Wahrscheinlichkeit) A und B seien Ereignisse, wobei $P(A) \neq 0$ ist. Die bedingte Wahrscheinlichkeit von A bei gegebenem B , ist

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A, B)}{P(B)} \quad (4.2)$$

bei gegebener Verbundwahrscheinlichkeit $P(A \cap B) = P(A, B)$.

Bayes'sches Theorem: Zur Bestimmung von bedingten Wahrscheinlichkeiten aus bereits bekannten bedingten Wahrscheinlichkeiten wird das Bayes'sche Theorem herangezogen [5].

Definition 4.2 (Bayes'sches Theorem) Falls A und B Ereignisse mit $P(A) \neq 0$ und $P(B) \neq 0$ sind, gilt

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$

Für die disjunkten Ereignisse A_1, A_2, \dots, A_n mit $P(A_i) \neq 0$ für alle i , gilt mit $1 \leq i \leq n$:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^n P(B|A_j)P(A_j)}.$$

Bedingte Unabhängigkeit: Die Zustände A und C sind von einander unabhängig, wenn bei gegebenen B folgender Zusammenhang erfüllt ist

$$P(A|B) = P(A|B, C),$$

s. auch d-Separation in Abschnitt 4.2.2.

Marginalisierung: Die Marginalisierung entspricht der Berechnung der Randwahrscheinlichkeit $P(A)$ einer Verbundwahrscheinlichkeit $P(A, B)$, dass auch der Einzelwahrscheinlichkeit $P(A)$ aller Zustände von A entspricht

$$P(A) = \sum_{i=0}^N P(A, B) = \sum_{i=0}^N P(A, b_i). \quad (4.3)$$

Dieser Zusammenhang ist der zentrale Punkt bei der Berechnung von Inferenzen in Bayes'schen Netzen.

Graphentheorie

Detaillierte Grundlagen der Graphentheorie sind in [51] zu finden. Im Folgenden werden einige Begriffe der Graphentheorie vorgestellt, die für das Verständnis von Bayes'schen Netzen hilfreich sind.

Definition 4.3 (Graph) Abhängig von der Verbindungsart der Knoten, wird zwischen zwei Arten von Graphen unterschieden

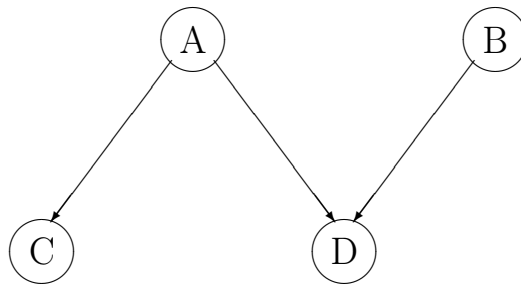


Abbildung 4.3: Beispiel für einen azyklischen gerichteten Graph

- a) Ein **ungerichteter Graph** ist durch ein Paar $U = (V, E)$ gegeben, wobei V eine endliche, nicht leere Menge von Knoten ist. Die Menge der Kanten ist durch $E \subseteq \{\{v_1, v_2\} | v_1, v_2 \in V\}$ dargestellt.
- b) Ein **gerichteter Graph** ist durch ein Paar $G = (V, E)$ gegeben, wobei $V = \{v_1, v_2, \dots, v_n\}$ eine endliche, nicht leere Menge von Knoten ist und $E = \{e_1, e_2, \dots, e_m\}$, mit $e_{ij} = (v_i, v_j)$, $v_i, v_j \in V$ die Menge der Kanten.

Definition 4.4 (Pfad) Ein Pfad ist eine Sequenz von Knoten (v_1, v_2, \dots, v_k) , $k \geq 2$ mit $(v_{i-1}, v_i) \in E$ für $2 \leq i \leq k$.

Definition 4.5 (Zyklus) Ein Zyklus ist ein Pfad mit demselben Anfangs- und Endknoten $v_1 = v_k$ und verschiedenen „Zwischenknoten“ v_i mit $2 \leq i \leq k - 1$.

Eine besondere Form des gerichteten Graphen ist der *azyklische gerichtete Graph* (engl. directed acyclic graph, DAG).

Definition 4.6 (Azyklischer gerichteter Graph) Dieser Graph ist ein gerichteter Graph ohne einen Zyklus, siehe Abbildung 4.3.

Um die Abhängigkeiten der Knoten untereinander besser darzustellen, werden die Knoten entsprechen eingeteilt:

Definition 4.7 (Elternteil) Ein Knoten v_i eines gerichteten Graphen ist ein Elternteil von v_j genau dann, wenn $v_i, v_j \in V$ und es eine Kante $(v_i, v_j) \in E$ gibt.

Definition 4.8 (Kind) Ein Knoten v_j eines gerichteten Graphen ist ein Kind von v_i genau dann, wenn $v_i, v_j \in V$ und es eine Kante $(v_i, v_j) \in E$ gibt.

4.2.2 Definition eines Bayes'schen Netzes

Ein Bayes'sches Netz ist ein azyklischer gerichteter Graph (DAG), dessen Knoten Variablen repräsentieren und dessen Kanten die Abhängigkeit der Variablen untereinander darstellen. Die einzelnen Variablen sind entweder kontinuierlich oder haben eine endliche Zahl an sich gegenseitig ausschließenden diskreten Zuständen [67]. Die Variablen können auch als Hypothesen, Wahrscheinlichkeitsfunktionen oder allgemeine Zustandsvariablen aufgefasst werden. Für jede Variable A mit den Eltern B_1, B_2, \dots, B_n existiert eine bedingte Wahrscheinlichkeitstabelle $P(A|B_1, B_2, \dots, B_n)$ [44].

Besitzt ein Knoten A keine Eltern, reduziert sich die bedingte Wahrscheinlichkeitstabelle auf eine einfache Wahrscheinlichkeitsverteilung $P(A)$. Diese a priori Wahrscheinlichkeiten müssen genau wie die Wahrscheinlichkeitstabellen vor der Verwendung eines Bayes'schen Netzes festgelegt werden. Durch sie erhält man einen unerwünschten Drift in das Modell. Dieser Drift wird aber in Kauf genommen, um ein Schließen möglichst ähnlich dem menschlichen Schlussfolgern zu erreichen.

Allgemein werden nur diskrete Bayes'sche Netze betrachtet, da kontinuierliche Bayes'sche Netze einen wesentlich größeren Rechenaufwand vor allem aufgrund der dazu notwendigen approximative Inferenzen benötigen, und damit für ein echtzeitfähiges System nicht geeignet sind [67].

d-Separation

Mithilfe der d-Separation wird die Abhängigkeit von benachbarten Knoten bestimmt [43]. Während im Grundzustand die jeweiligen Zustände der Knoten eines Netzes von ihren Vorgängern bzw. Nachfolgern abhängen, können Knoten durch Hinzunahme von Information über einen anderen Knoten voneinander unabhängig werden. Eine Aussage über die Gewissheit eines Zustandes einer Variablen nennt man *Evidenz*. Hierbei wird weiter unterschieden zwischen *harter Evidenz* und *weicher Evidenz*. Harte Evidenz bezeichnet hierbei die absolute Sicherheit über den Zustand einer Variablen, während weiche Evidenz nur mit einer gewissen Unsicherheit eine Aussage über den Zustand einer Variablen erlaubt.

Es gibt drei Arten der d-Separation, die im Folgenden einzeln vorgestellt werden.

Serielle Verbindung: Die einfachste Art der d-Separation tritt bei seriellen Verbindungen auf. Wie in Abbildung 4.4 zu sehen ist, hat Knoten A Einfluss auf Knoten B , welcher wiederum Knoten C beeinflusst. Erhöht man die Wahrscheinlichkeit für einen Zustand des Knoten A , so ändern sich leicht ersichtlich auch die Wahrscheinlichkeiten der Zustände des Knoten B . Diese Veränderung wiederum hat Einfluss auf die Wahrscheinlichkeiten des Knoten C . Ebenso wirkt sich Wissen über den Knoten C durch Knoten B auf den Knoten A aus.

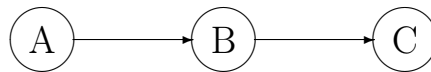


Abbildung 4.4: Serielle Verbindung

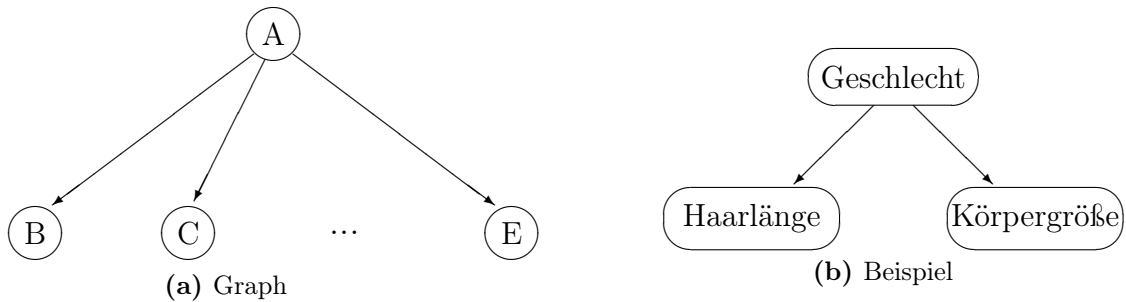


Abbildung 4.5: Divergente Verbindung

Die Einflüsse der Knoten untereinander ändern sich allerdings stark, falls der Zustand des Knoten B bekannt ist, also harte Evidenz für ihn vorliegt. Ist dies der Fall, haben die Knoten A und C keinen Einfluss mehr aufeinander, sie sind unabhängig geworden. A und C sind d -separiert durch B .

Divergente Verbindung: Der Graph in Abbildung 4.5 (a) enthält *divergente* Verbindungen. Alle Kinder von A beeinflussen sich gegenseitig, es sei denn der Zustand von A ist bekannt. Im Falle harter Evidenz für A sind B , C , ..., E d -separiert durch A .

Beispiel 1 In Abbildung 4.5 (b) ist der Zusammenhang zwischen Geschlecht (männlich, weiblich), der Haarlänge (lang, kurz) und der Körpergröße (>168 cm, <168 cm) dargestellt [44].

Kennt man eine Person nicht, hat die Haarlänge einen Einfluss auf die Wahrscheinlichkeit über das Geschlecht der Person. Und diese wiederum auf die geschätzte Körpergröße. Ist allerdings bekannt, dass es sich bei der Person um einen Mann handelt, kann man aus der Länge der Haare keine Rückschlüsse auf die Größe der Person ziehen.

Konvergente Verbindung: Die Netzstruktur einer *konvergenten* Verbindung ist in Abbildung 4.6 (a) abgebildet. Ist nichts über den Knoten A bekannt, außer dem was von seinen Eltern weitergegeben wurde, sind die Eltern (B , C , ..., E) unabhängig. Evidenz von einem von ihnen hat also keinerlei Einfluss auf die Wahrscheinlichkeiten der anderen. Ist jedoch die Wahrscheinlichkeit der Zustände des Knoten A bekannt, kann Evidenz über einen Elternknoten Auswirkungen auf die anderen Elternknoten haben. Dieser Effekt wird auch *explaining away* genannt, siehe auch Abbildung 4.7 (d).

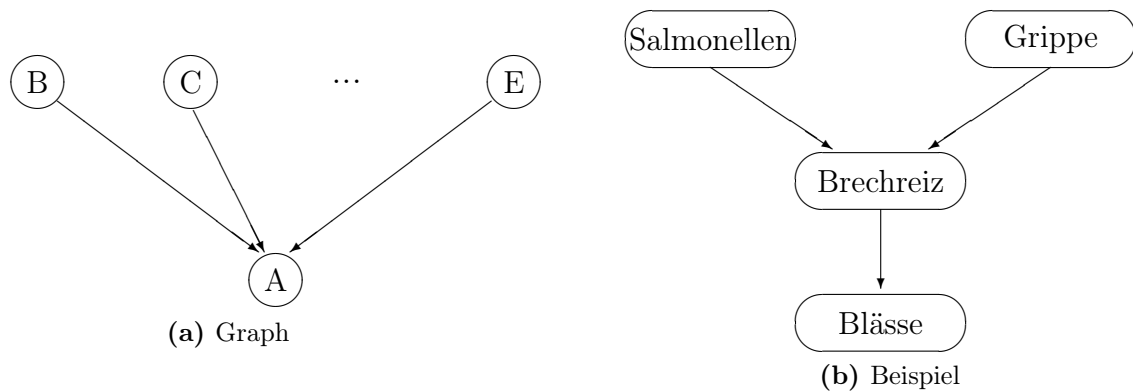


Abbildung 4.6: Konvergente Verbindung

Zu beachten ist, dass bei divergenten und seriellen Verbindungen für die *d*-Separation *harte Evidenz* für den Zwischenknoten benötigt wird, während bei konvergenter Verbindung bereits *weiche Evidenz* ausreicht.

Beispiel 2 In Abbildung 4.6 (b) ist der Zusammenhang zwischen den Krankheiten Salmonellen-Infektion und Grippe, sowie ihren Symptomen Brechreiz und Blässe visualisiert.

Ist keine Kenntnis über Symptome (Brechreiz und Blässe) vorhanden, sagt die Information, dass der Patient eine Salmonellen-Infektion hat, nichts über die Möglichkeit einer Grippe aus. Ist der Patient hingegen blass, lässt uns die Information, dass der Patient keine Salmonellen-Infektion hat, vermuten, dass er an einer Grippe leidet.

Diese drei Verbindungsarten decken die Möglichkeiten ab, durch die Information in einem Bayes'schen Netz übertragen werden kann. Um die Unabhängigkeit von Knoten eines Netzes bei gegebener Evidenz zu bestimmen, müssen folgende Regeln überprüft werden.

Definition 4.9 (d-Separation) Zwei Variablen A und B eines kausalen Netzwerks sind *d*-separiert, falls für alle Pfade zwischen A und B eine dazwischenliegende Variable V mit $V \neq A$ und $V \neq B$ existiert, so dass die Verbindung entweder

- *seriell oder divergent ist und für die Variable V harte Evidenz vorliegt.*

oder

- *konvergent ist und weder für die Variable V noch für ihre Nachkommen Evidenz vorliegt.*

	b_1	b_2	b_3
a_1	0,4	0,3	0,6
a_2	0,6	0,7	0,4

(a) $P(A|B)$

	b_1	b_2	b_3
a_1	0,16	0,12	0,12
a_2	0,24	0,28	0,08

(b) $P(A,B)$ mit $P(B) = (0,4; 0,4; 0,2)$

Tabelle 4.1: Beispiele

Berechnung der Wahrscheinlichkeiten in Bayes'schen Netzen

Bei den Variablen eines diskreten Bayes'schen Netzes handelt es sich um eine endliche Zahl sich gegenseitig ausschließender diskreter Zustände.

Eine Variable A besitzt die Zustände a_1, \dots, a_n mit der Wahrscheinlichkeitsverteilung über alle Zustände

$$P(A) = (x_1, \dots, x_n); \quad x_i \geq 0; \quad \sum_{i=1}^n x_i = 1,$$

wobei x_i die Wahrscheinlichkeit darstellt, mit der A sich im Zustand a_i befindet. Die Wahrscheinlichkeit, dass die Variable A sich in Zustand a_i befindet, wird durch $P(A = a_i)$ ausgedrückt. Ist die Variable A aus dem Kontext ersichtlich, kann auch kurz $P(a_i)$ geschrieben werden.

Die bedingte Wahrscheinlichkeit $P(A|B)$ der Variablen A und B mit den Zuständen a_1, \dots, a_n und b_1, \dots, b_m wird als $n \times m$ Tabelle dargestellt, siehe Tabelle 4.1 (a).

Die Summe aller Wahrscheinlichkeiten einer Spalte muss immer eins ergeben, damit es sich um eine gültige Wahrscheinlichkeit handelt.

Die Verbundwahrscheinlichkeit $P(A, B)$ der Variablen A und B wird ebenfalls in einer $n \times m$ Tabelle abgebildet, siehe Tabelle 4.1 (b). Für jede Konfiguration aus (a_i, b_j) existiert eine Wahrscheinlichkeit. Die Wahrscheinlichkeiten aller Einträge aufsummiert ergibt immer eins. Die Berechnung der Verbundwahrscheinlichkeit eines Elements $P(a_i, b_j)$ von $P(A, B)$ ergibt sich aus Gleichung (4.2) zu

$$P(a_i, b_j) = P(a_i|b_j)P(b_j)$$

Zur Bestimmung aller Elemente von $P(A, B)$ muss jede Spalte j der Tabelle 4.1 (a) mit $P(b_j)$ multipliziert werden.

Angewandt auf Variablen ergibt sich eine grundlegende Gleichung zur Berechnung der Verbundwahrscheinlichkeit zweier Variablen bei Bayes'schen Netzen

$$P(A, B) = P(A|B)P(B).$$

Um die Einzelwahrscheinlichkeit $P(A)$ der Variable A zu bestimmen, ist ein weiterer Rechenschritt notwendig. Die so genannte *Marginalisierung*.

Für den Zustand a_i der Variablen A gibt es genau m verschiedene Ereignisse für die A den Zustand a_i annimmt. Da diese Ereignisse sich gegenseitig ausschließen, lässt sich die Wahrscheinlichkeit $P(a_i)$ mit Gleichung (4.3) berechnen.

$$P(a_i) = \sum_{j=1}^m P(a_i, b_j) = (a_i, b_1) + (a_i, b_2) + \dots + (a_i, b_m) .$$

Um die Einzelwahrscheinlichkeiten $P(A)$ für alle Zustände von A zu bekommen, wird obige Gleichung umgeschrieben zu

$$P(A) = \sum_B P(A, B) .$$

Durch Herausbildung von B aus $P(A, B)$ in Tabelle 4.1 (b) erhält man $P(A) = (0,4; 0,6)$.

4.2.3 Entwurf von Bayes'schen Netzen

Ein Ziel beim Entwurf von Bayes'schen Netzen ist es, die Wahrscheinlichkeiten von nicht beobachtbaren Ereignissen zu schätzen. Aus diesem Grunde geht man beim Entwurf der Graphenstruktur eines Bayes'schen Netzes im Allgemeinen so vor, dass zuerst die nicht beobachtbaren Ereignisse (*Hypothesenergebnisse*) definiert werden. Die Hypothesenergebnisse werden dann zu *Hypothesenvariablen* zusammengefasst. Dabei ist darauf zu achten, dass alle Zustände einer Variablen sich gegenseitig ausschließen. Anschließend werden die beobachtbaren Variablen (*Informationsvariablen* oder auch *Evidenzvariablen*) definiert und festgelegt, wie diese mit den Hypothesenvariablen (auch *Hypothesenknoten*) in Verbindung stehen. Für die Informationsvariablen (auch *Informationsknoten*) werden die exakten Zustände nicht benötigt, sondern es ist ausreichend, Schätzwerte zu besitzen. Sind alle Knoten bzw. Variablen des Netzes definiert und ihre Abhängigkeiten durch Verbindungen zwischen den einzelnen Knoten repräsentiert, werden als nächstes die Pfeilrichtungen von den Hypothesenvariablen zu den Informationsvariablen festgelegt. Dabei ist zu beachten, dass kein zyklischer Graph entsteht und für reale Anwendungen ein kausales Netz aus den Knoten entsteht [44].

Die Struktur des Graphen kann auch automatisch mithilfe eines Strukturlernalgorithmus für Bayes'sche Netze erstellt werden. Um diese Algorithmen jedoch anwenden zu können, werden Lerndaten mit einer umfassenden Abdeckung aller möglichen Ereignisse und Zustände benötigt, was im Allgemeinen nicht gewährleistet werden kann. Bei einem manuellen Entwurf des Graphen ergibt sich eine leichter nachvollziehbare Struktur, die auch im Nachhinein einfach modular erweiterbar ist. Aus diesen Gründen wird die Möglichkeit des automatischen Strukturlernens hier nicht weiter untersucht. Eine Übersicht verschiedener Verfahren wird in [67] und [71] gegeben.

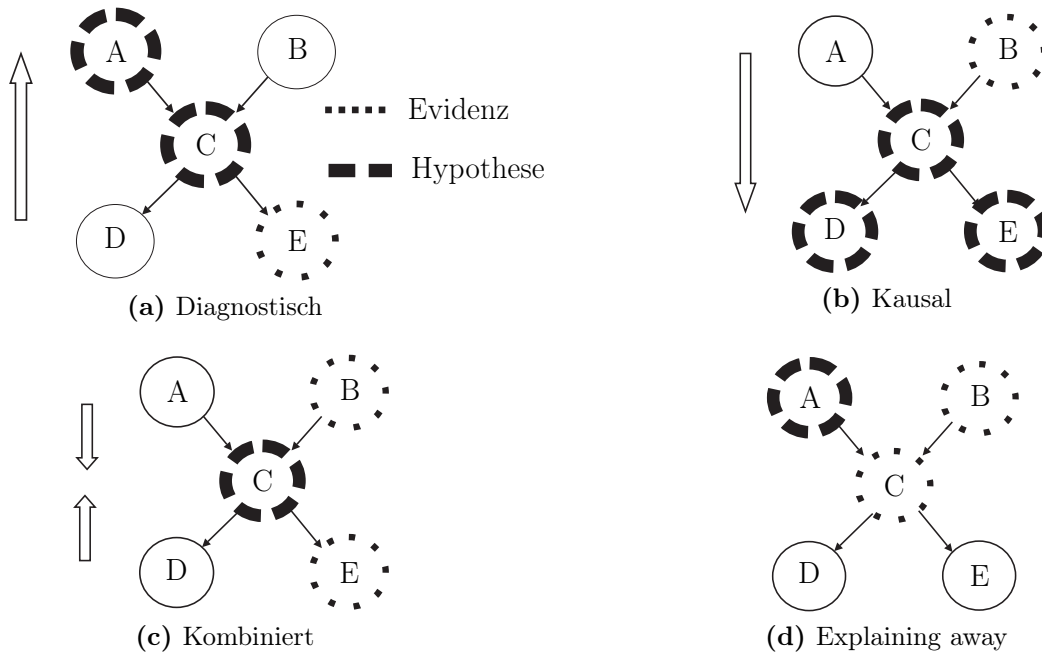


Abbildung 4.7: Mögliche Inferenzen bei Bayes'schen Netzen

Für den gerichteten azyklischen Graph (DAG), der dem Bayes'schen Netz als Graphenstruktur zu Grunde liegt, müssen nun noch die Wahrscheinlichkeitsverteilungen für die verschiedenen Variablen festgelegt werden. Dazu gibt es ebenfalls die Möglichkeit, dies mithilfe von automatisierten Lernverfahren durchzuführen. Allerdings benötigt man auch hierzu wiederum eine große Anzahl an Lerndaten, die den Ereignisraum komplett abdecken. Unterschiedliche Methoden für automatisches Parameterlernen bei Bayes'schen Netzen sind ebenfalls in [67] beschrieben.

Eine andere Möglichkeit, die Initialisierungswahrscheinlichkeiten und bedingten Wahrscheinlichkeitstabellen für das Bayes'sche Netz festzulegen, ist diese empirisch oder manuell zu bestimmen. Die einzelnen Parameter der bedingten Wahrscheinlichkeitstabellen beruhen dann meist auf Erfahrungswerten oder Statistiken. So kann zum Beispiel die Zuverlässigkeit eines Sensors mit in die Parameterwahl einfließen. Die Initialisierungswahrscheinlichkeiten werden in der Regel mit Durchschnittswerten belegt.

Das Schließen oder die Inferenz in Bayes'schen Netzen dient dazu die a posteriori Wahrscheinlichkeiten der Hypothesenknoten oder eines Zielknotens bei gegebener Evidenz bzw. Information in einem Knoten zu berechnen. Hierzu wird zwischen exakten und approximativen Verfahren unterschieden. Die exakten benötigen eine höhere Rechenzeit, sind aber in der Berechnung genauer. Bei den approximativen ist es genau umgekehrt. Bei kleinen Bayes'schen Netzen und ausreichender Rechenkapazität ist die exakte Inferenz zu bevorzugen [67].

In Bayes'schen Netzen sind je nach Anforderung unterschiedliche Schlussrichtungen (Inferenzweg) möglich. Dies ist in Abbildung 4.7 mittels der breiten Pfeile

links vom jeweiligen Graph angedeutet. Beim diagnostischen Schließen in Abbildung 4.7 (a) werden Nachrichten bzw. Informationen entgegen der Pfeilrichtung durch das Netz propagiert. Dieser Ansatz wird auch *Bottom-up Propagation* genannt.

Das Gegenteil davon ist die *Top-down Propagation* oder kausales Schließen. Dabei werden die Nachrichten bzw. Informationen in Pfeilrichtung durch das Netz gesendet, siehe Abbildung 4.7 (b).

Beide Ansätze können auch kombiniert werden. Dies ist in Abbildung 4.7 (c) dargestellt. Dabei wird sowohl kausales als auch diagnostisches Schließen verwendet. Dieser Ansatz ist enorm rechenaufwendig, da die Wahrscheinlichkeiten bzw. das *Belief-update* mehrfach berechnet bzw. durchgeführt werden müssen. Unter *Belief-update* wird die Aktivierung eines Knotens und das Aktualisieren dessen a posteriori Wahrscheinlichkeit verstanden. Die Aktualisierung wird mittels einer Marginalisierung der *Knotenverbundswahrscheinlichkeit* oder der Knotenwahrscheinlichkeitstabelle erreicht.

Der letzte Fall, der *explaining away*, ist in Abbildung 4.7 (d) dargestellt. Dabei ändert sich die a posteriori Wahrscheinlichkeit von Knoten *A* bei Vorliegen von Evidenz im gemeinsamen Knoten *C* und im Knoten *B*.

4.3 Vorverarbeitung der Sensordaten

Die Ausgangssignale von Sensoren können im Allgemeinen nicht ohne geeignete Verarbeitung für eine Informationsfusion verwendet werden. Aus diesem Grund ist eine Vorverarbeitung der Sensorsignale notwendig. Dies ist in Abbildung 4.1 als *Level 0* der Informationsverarbeitung bzw. in Abbildung 4.2 als *Beobachten* dargestellt. Dazu gehört neben einer Datenverarbeitung auch die Anti-Aliasing Filterung und Digitalisierung der Sensorsignale. Weitere Informationen dazu sind in [52] zu finden.

Die Datenverarbeitung auf dieser Ebene hat die Aufgabe, die von einem Sensor gewonnenen Informationen dem Fusionsprozess geeignet zur Verfügung zu stellen. Der Fusionsprozess in dem hier betrachteten Fall sei ein Bayes'sches Netz, s. Abschnitt 4.2 und 4.4. Die von den Sensoren bereitgestellten Informationen für den Fusionsprozess müssen nun dahingehend vorverarbeitet werden, dass diese einen Wertebereich von null bis eins besitzen und damit eine Eingangswahrscheinlichkeit für das Netz repräsentieren. Eine null bedeutet dabei, dass sich kein Objekt im Sensorbereich befindet. Hingegen soll eine eins ausgegeben werden, sobald der Sensor einen typischen Wert für das zu erkennende Objekt, wie z. B. eine Hand, erreicht oder übersteigt. Werte dazwischen zeigen an, mit welcher Wahrscheinlichkeit es sich um das zu erkennende Objekt handeln könnte.

In diesem Abschnitt werden nun die notwendigen Datenverarbeitungsschritte für die von den kapazitiven (s. Abschnitt 2.3) und Infrarot-Sensoren (s. Ab-

schnitt 2.2) gelieferten Werten vorgestellt. Des Weiteren werden im Zuge der Vorverarbeitung ebenfalls bekannte Störeinflüsse weitestgehend korrigiert. Abschließend werden auf Basis der normierten Sensorwerte Zustände für das Bayes'sche Netz bestimmt, um eine effektive Informationsfusion zu ermöglichen.

4.3.1 Thermopile Infrarot-Sensoren

Der thermopile Infrarot-Sensor (s. Abschnitt 2.2) liefert eine konstante Ausgangsspannung $u_{IR}(t)$ so lang keine Temperaturänderung im Detektionsbereich erkannt wird. Die Höhe dieser Ausgangsspannung ist von der Umgebungstemperatur abhängig und wird deshalb bei der Systeminitialisierung als $u_{IR}(0)$ abgespeichert. Ändert sich während des Betriebs die Umgebungstemperatur um mehr als 10-15 °C muss eine automatische Reinitialisierung des Systems erfolgen, um wieder zuverlässige Detektionsergebnisse zu erhalten.

Solange sich keine Wärmequelle im Detektionsbereich befindet, soll eine 0 ausgegeben werden. Durch Subtraktion der Spannung $u_{IR}(0)$ von der aktuellen Spannung $u_{IR}(t)$ wird dies erreicht. Ist das Ergebnis hiervon negativ, so wird es auf 0 gesetzt. Dies ist bis zu einer Umgebungstemperatur von ca. 27-30 °C für den hier vorgestellten Anwendungsfall sinnvoll. Für eine Temperatur über 30 °C ist eine zuverlässige Handerkennung mit Infrarot-Sensoren nur eingeschränkt möglich, da eine Hand selbst ca. 30 °C warm sein kann.

Der Referenzwert $u_{IR,ref}$ wird empirisch für jeden Infrarot-Sensor bestimmt und gibt den Wert an, mit dem auf eine Hand geschlossen werden kann. Durch Division der Subtraktion mit $u_{IR,ref}$ wird eine Normierung auf eins erreicht. Alle über eins liegenden Werte liefern keine neue Information und werden deshalb auf eins gesetzt. Mit Gleichung (4.4) wird die Bestimmung des normierten Ausgangssignals festgelegt.

$$u_{IR,norm}(t) = \begin{cases} 0 & \text{für } u_{IR}(t) < u_{IR}(0) \\ \frac{u_{IR}(t) - u_{IR}(0)}{u_{IR,ref}} & \text{für } u_{IR}(0) \leq u_{IR}(t) \leq (u_{IR}(0) + u_{IR,ref}) \\ 1 & \text{für } u_{IR}(t) > (u_{IR}(0) + u_{IR,ref}) \end{cases} \quad (4.4)$$

Diese Gleichung und der Referenzwert $u_{IR,ref}$ sind für den Bereich der Umgebungstemperatur bis ca. 30 °C gültig. Bei Umgebungstemperaturen über 30 °C würde der Sensorwert bei Einführen einer Hand in den Detektionsbereich den Ausgangswert eins nicht mehr erreichen, da die Temperaturänderung durch eine Hand zu gering ist. Dies ist jedoch ein erwünschter Effekt, da die Detektionszuverlässigkeit einer Hand bei hohen Temperaturen ohnehin abnimmt und sich dies so automatisch im Ausgangswert widerspiegelt.

4.3.2 Kapazitive Sensoren

Für jedes der zwei kapazitiven Sensorsysteme (s. Abschnitt 2.3) ist eine separate Vorverarbeitung und Normierung notwendig. Beide Systeme messen Kapazitäten bzw. Kapazitätsänderungen und reagieren daher auf die Höhenänderung der Schutzhaube. Diese Änderung kann z. B. durch Ultraschall-Sensoren oder optische Triangulationssensoren bestimmt werden und dient dann zur Korrektur der Sensorsignale. Im Folgenden werden für den kapazitiven Sensor zur Vorfeldüberwachung und für den in den Schutzlamellen integrierten Sensor die Normierung und entsprechende Korrektur vorgestellt. Weitere Informationen zur Sensorintegration finden sich in Kapitel 5.

Sensor im Vorfeld

Der kapazitive Sensor AD7747 dient zur Kapazitätsmessung für den Raum zwischen der Verlängerung der Schutzhaube und dem Arbeitstisch. Die gemessene Kapazität $C_{vor}(t)$ ist dabei von der Höhe der Schutzhaube abhängig, s. Abschnitt 2.3.2.

Der Einfluss der Schutzhaubenhöhe kann mit einer vorher bestimmten Referenzkennlinie nahezu eliminiert werden. Zur Erstellung der Referenzkennlinie wird die Kapazität C_{vor} bei bestimmten Schutzhaubenhöhen h_S gemessen. Daraus kann mithilfe einer kubischen Spline-Interpolation (s. [52]) die Referenzkennlinie erstellt werden. Der so interpolierte Kapazitätsverlauf $C_{ref}(h_S)$ ist in Abbildung 4.8 zu sehen. Solange sich kein Objekt im Detektionsbereich befindet, kann mittels Subtraktion der Referenzkennlinie $C_{ref}(h_S)$ von der aktuellen Kapazität C_{vor} eine 0 als Ausgangswert erzeugt werden.

Die Normierung des Sensorwertes kann nicht mit einem konstanten Divisor bestimmt werden, da sich durch Veränderung der Höhe der Schutzhaube auch der Zwischenraum des kapazitiven Sensors zum Arbeitstisch ändert. Ist dieser Zwischenraum klein, füllt eine Hand nur einen kleinen Teil des Zwischenraums aus. Wohingegen bei geringer Höhe der Schutzhaube fast der gesamte Zwischenraum durch die Hand ausgefüllt wird. Dies hat zur Folge, dass die Kapazitätsänderung bei großer Höhe der Schutzhaube wesentlich kleiner ist als bei einer geringen Höhe.

Damit diese kleine Kapazitätsänderung durch die Normierung nicht verschwindet, wird ein linearer, höhenabhängiger Normierungsfaktor $N_{vor}(h_S)$ eingeführt

$$N_{vor}(h_S) = \begin{cases} 1 & \text{für } h_S < 3,5 \\ h - 3 & \text{für } h_S \geq 3,5 . \end{cases}$$

Bei einer Haubenhöhe unter 3,5 cm kann konstruktionsbedingt keine Hand mehr in den Detektionsbereich des Sensors gelangen.

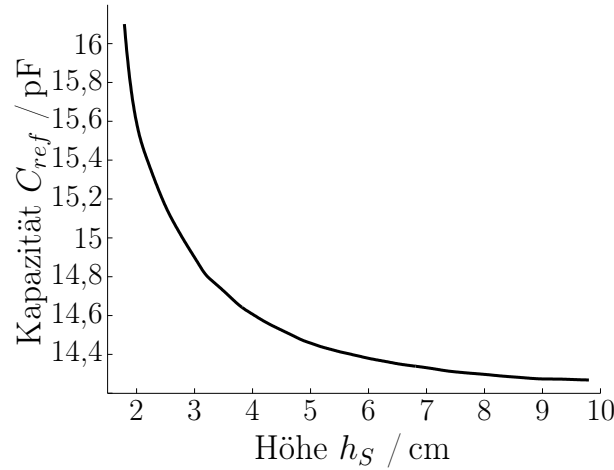


Abbildung 4.8: Kubische Spline-Interpolierte als Referenzkennlinie

Der normierte und kompensierte Sensorwert wird zusätzlich noch auf eins beschränkt und es ergibt sich folgender Zusammenhang für den Ausgabewert des Sensors

$$C_{vor,norm}(t, h_S) = \begin{cases} 0 & \text{für } C_{vor}(t) < C_{ref}(h_S) \\ C_{vor}(t) - C_{ref}(h_S) & \text{für Bedingung 1} \\ (C_{vor}(t) - C_{ref}(h_S)) \cdot (h_S - 3) & \text{für Bedingung 2} \\ 1 & \text{für Bedingung 3} \end{cases}$$

mit

$$\text{Bedingung 1} = C_{ref}(h_S) \leq C_{vor}(t) \leq (C_{ref}(h_S) + 1) \text{ und } h_S < 3,5$$

$$\text{Bedingung 2} = C_{ref}(h_S) \leq C_{vor}(t) \leq \left(C_{ref}(h_S) + \frac{1}{h_S - 3} \right) \text{ und } h_S \geq 3,5$$

$$\text{Bedingung 3} = \begin{cases} C_{vor}(t) > (C_{ref}(h_S) + 1) \text{ und } h_S < 3,5 \\ C_{vor}(t) > \left(C_{ref}(h_S) + \frac{1}{h_S - 3} \right) \text{ und } h_S \geq 3,5 \end{cases} \text{ oder}$$

Sensoren in den Schutzlamellen

Die in den Schutzlamellen integrierten kapazitiven Sensoren liefern als Ausgangssignal eine Spannung u_{lam} , die proportional zur Kapazitätsänderung bzw. Impedanzänderung im Detektionsbereich des Sensors ist, s. Abschnitt 2.3.3. Diese

werden derart korrigiert und normiert, dass der Sensor eine 0 liefert, entweder solange die Lamelle auf dem Arbeitstisch aufliegt oder sich Luft im Sensorbereich befindet. Ist ein Gegenstand zwischen Tisch und Lamelle sollen Werte größer 0 ausgegeben werden, wobei der Wertebereich wieder auf $[0,1]$ begrenzt ist und der Wert eins für eine potentielle Handerkennung stehen soll.

Wenn sich kein Gegenstand im Sensorbereich befindet, muss für die Korrektur des Sensorwertes auf 0 die Stellung der Schutzlamelle beachtet werden, da die Konfiguration des elektrischen Feldes und damit die Impedanz des Sensorsystems davon abhängig ist. Um die Stellung der Lamelle zu bestimmen, werden Potentiometer zur Winkelmessung in den Lamellendrehpunkten verwendet. Der Wert des Potentiometers ist r_{SL} . Als weitere Einflussgröße ist die aktuelle Höhe h_S der Schutzhaube zu beachten, da bei derselben Stellung einer Lamelle sich die Sensorwerte stark unterscheiden, je nachdem ob die Lamelle auf dem Arbeitstisch aufliegt oder nicht.

Zur Korrektur dieser beiden Einflüsse wird ein zweidimensionales Kennfeld verwendet. Durch Messung von Wertetripeln bestehend aus dem Wert des kapazitiven Sensors u_{lam} , dem Wert des Potentiometers r_{SL} und der Höhe h_S kann ein zweidimensionales Kennfeld erzeugt werden. Damit kann der Ausgabewert des kapazitiven Sensors bei bekannter Haubenhöhe und Lamellenposition interpoliert werden. Durch Subtraktion dieses Interpolationswertes $u_{kl}(h_S, r_{SL})$ vom gemessenen Wert des Sensors wird die Korrektur auf 0 durchgeführt.

Korrekturrechnung mit zweidimensionaler Kennfeldinterpolation: Ein zweidimensionales Kennfeld kann einen Funktionswert interpolieren, der von zwei Variablen abhängt. Zur Aufstellung des Kennfeldes müssen ein Wertepaar als Stützstellen und dessen Funktionswert zur Verfügung stehen. Normalerweise werden die Stützstellen äquidistant bzw. rechteckförmig auf dem Wertebereich angeordnet. Es ist aber auch möglich, die Stützstellen, wie in Abbildung 4.9 zu sehen ist, lediglich in einer Dimension auf eine Gerade zu legen und in der anderen Dimension beliebig auf der Geraden anzuordnen.

Zur Berechnung des Funktionswertes $u_{kl}(h_S, r_{SL})$ müssen vier Punkte gefunden werden, welche den zu interpolierenden Punkt umgeben. Hierzu ist es notwendig, zuerst die beiden Geraden in h_S -Richtung zu suchen. Anschließend sucht man jeweils zwei Punkte auf den benachbarten Geraden in r_{SL} -Richtung, die den zu interpolierenden Punkt umschließen.

Zur Erstellung des Kennfeldes werden die Wertepaare und der dazugehörige Funktionswert aufgenommen. Dazu wird die Schutzhaube von der kleinsten einstellbaren Höhe in kleinen Schritten bis zur maximal vorgesehen Höhe bewegt. Auf jeder Höhe werden die Potentiometerwerte r_{SL} und die gemessenen Werte der kapazitiven Sensoren u_{lam} bei verschiedenen Stellungen der Lamellen aufgenommen und zusammen mit der jeweiligen Höhe h_S abgespeichert. Die Stellung

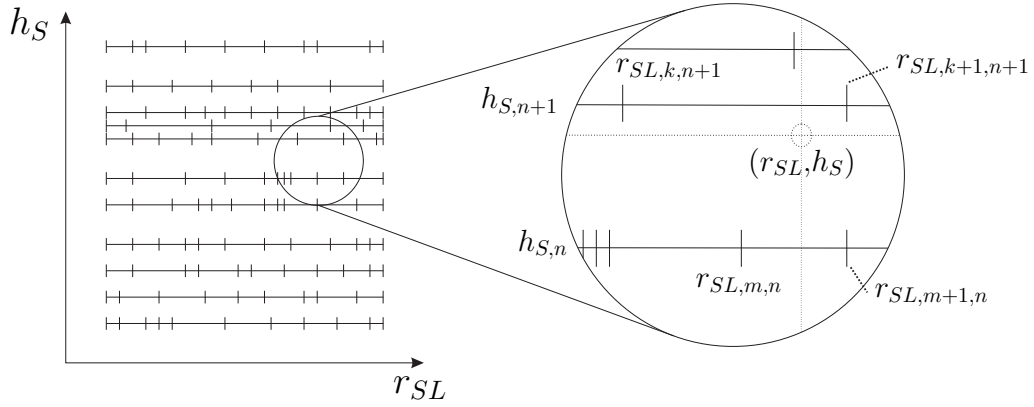


Abbildung 4.9: Anordnung der Stützstellen auf Geraden einer Dimension für die zweidimensionale Interpolation

der Lamellen wird so variiert, dass sie erst auf der Tischplatte aufliegt und dann in kleinen Schritten bis zum maximalen Öffnungswinkel nach oben angehoben wird.

Durch diese Art der Aufnahme des Wertetripels entsteht eine Anordnung der Stützstellen, die bei den verschiedenen Höhen h_S gleichmäßig auf Geraden liegen und in Richtung der Potentiometerwerte r_{SL} beliebig auf diesen Geraden verteilt sind. Zur Bestimmung der Zwischenwerte werden die Gleichungen (4.5) bis (4.7) herangezogen.

$$\begin{aligned}
 h_{S,n} &\leq h_S < h_{S,n+1} \\
 r_{SL,m,n} &\leq r_{SL} < r_{SL,m+1,n} \\
 r_{SL,k,n+1} &\leq r_{SL} < r_{SL,k+1,n+1}
 \end{aligned} \tag{4.5}$$

Die Interpolation des Korrekturwertes $\hat{u}_{kl}(h_S, r_{SL})$ ist dann

$$\begin{aligned}
 \hat{u}_{kl}(h_S, r_{SL}) &= (1 - g_1) \cdot (1 - e) \cdot u_{kl}(h_{S,n}, r_{SL,m,n}) \\
 &\quad + (1 - g_2) \cdot e \cdot u_{kl}(h_{S,n+1}, r_{SL,k,n+1}) \\
 &\quad + g_1 \cdot (1 - e) \cdot u_{kl}(h_{S,n}, r_{SL,m+1,n}) \\
 &\quad + g_2 \cdot e \cdot u_{kl}(h_{S,n+1}, r_{SL,k+1,n+1}),
 \end{aligned} \tag{4.6}$$

mit den für die Berechnung benötigten Abständen

$$\begin{aligned}
 g_1 &= \frac{r_{SL} - r_{SL,m,n}}{r_{SL,m+1,n} - r_{SL,m,n}}, \quad g_2 = \frac{r_{SL} - r_{SL,k,n+1}}{r_{SL,k+1,n+1} - r_{SL,k,n+1}}, \\
 e &= \frac{h_S - h_{S,n}}{h_{S,n+1} - h_{S,n}}.
 \end{aligned} \tag{4.7}$$

Das sich für die zweite Lamelle auf der linken Seite der Schutzhaube resultierende zweidimensionale Kennfeld ist in Abbildung 4.10 dargestellt.

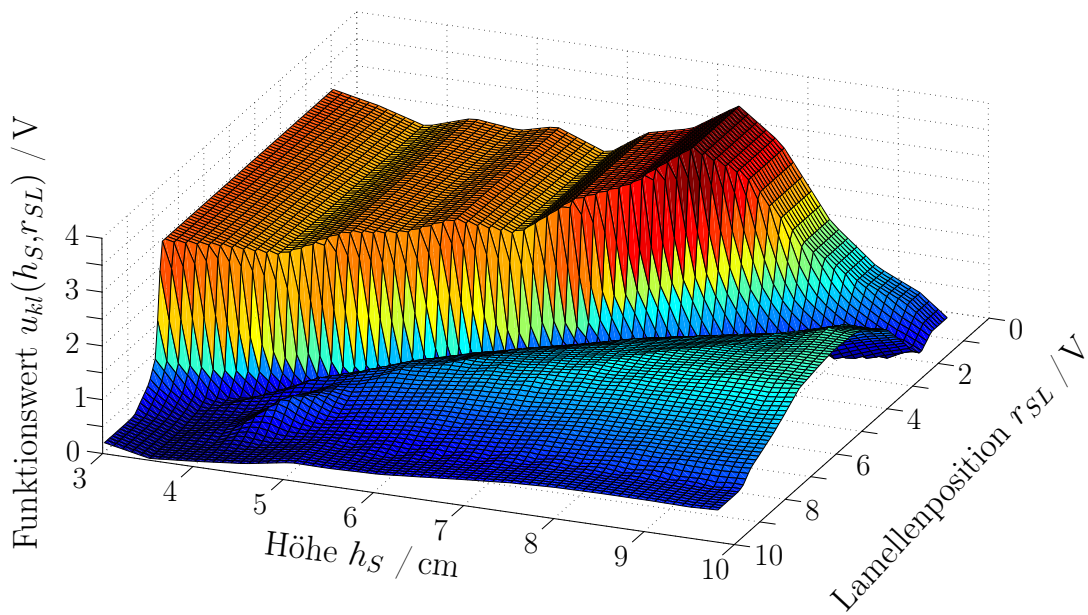


Abbildung 4.10: Zweidimensionales Kennfeld für die 2. Lamelle auf der linken Seite

Durch die Subtraktion des interpolierten Korrekturwertes $\hat{u}_{kl}(h_S, r_{SL})$ vom gemessenen Sensorwert $u_{lam}(t)$ ergibt sich der durch das Kennfeld korrigierter Sensorwert $\hat{u}_{kap,lam,i}(t)$.

$$\hat{u}_{kap,lam}(t) = u_{lam}(t) - \hat{u}_{kl}(h_S, r_{SL})$$

Das so erzielte Resultat ist aber für die Kompensation der Werte der kapazitiven Sensoren in den Lamellen noch nicht ausreichend. Der Grund hierfür ist, dass die Stellung der Lamellen und damit der Wert des Potentiometers r_{SL} bei Auflage auf den Arbeitstisch für unterschiedliche Höhen stark variiert. Der Wert eines kapazitiven Sensors $u_{lam}(t)$ wiederum ändert sich stark, wenn die Lamelle den Kontakt zum Arbeitstisch verliert. Folglich besitzen zwei benachbarte Referenzhöhen völlig unterschiedliche Funktionswerte $\hat{u}_{kl}(h_S, r_{SL})$ für die gleichen Potentiometerwerte.

Zur Erläuterung wird eine Lamelle mit einem festen Drehwinkel (konstanter Potentiometerwert $r_{SL} = const.$) betrachtet. Liegt die Lamelle auf dem Arbeitstisch auf, misst der kapazitive Sensor eine geringe Impedanz (großes u_{kl}). Hebt man nun die Schutzhaube um nur ca. 0,3 cm an, liegt die Lamelle nicht mehr auf dem Tisch auf. Die gemessene Impedanz nimmt durch den Abstand der Lamelle vom Arbeitstisch von 0,3 cm stark zu (kleines u_{kl}). Verwendet man nun einen der beiden Werte zur Interpolation des anderen, so ist das Ergebnis stark verfälscht. Zur Behebung dieses Fehlers muss die vorgestellte zweidimensionale Kennfeldinterpolation modifiziert werden.

Erweiterung des zweidimensionalen Kennfeldes: Zur Behebung des Fehlers wird der gemessene Potentiometerwert an die Potentiometerwerte der zur In-

terpolation verwendeten Referenzhöhen angepasst. Hierzu werden die Gleichungen (4.5) bis (4.7) um den Einfluss des Potentiometerwerts erweitert.

Die Bestimmung der beiden zur Interpolation verwendeten Referenzhöhen $h_{S,n}$ und $h_{S,n+1}$ bleiben unverändert, ebenso der Abstand e .

Bevor die Bestimmung der benachbarten Referenzpunkte für die Potentiometerwerte r_{SL} durchgeführt wird, werden diese an die Potentiometerwerte der beiden bereits gewählten Referenzhöhen angepasst.

Dazu werden in einem ersten Schritt für die beiden Referenzhöhen $h_{S,n}$ und $h_{S,n+1}$ die Potentiometerreferenzwerte $r_{SL,ref,n}$ und $r_{SL,ref,n+1}$ gesucht, für die die Funktionswerte $u_{kl}(h_{S,n}, r_{SL})$ und $u_{kl}(h_{S,n+1}, r_{SL})$ ihr Maximum annehmen. Dadurch erhält man die Potentiometerreferenzwerte, für die die Lamelle bei der jeweiligen Höhe auf dem Tisch aufliegt.

$$r_{SL,ref,n} = \operatorname{argmax}_{r_{SL}} \{u_{kl}(h_{S,n}, r_{SL})\}$$

$$r_{SL,ref,n+1} = \operatorname{argmax}_{r_{SL}} \{u_{kl}(h_{S,n+1}, r_{SL})\}$$

Anschließend wird die Differenz $\Delta r_{SL,ref}$ der beiden maximalen Referenzwerte der Potentiometer gebildet. Damit ergibt sich der Differenzwert für die auf dem Arbeitstisch aufliegende Lamelle bei zwei benachbarten Referenzhöhen

$$\Delta r_{SL,ref} = r_{SL,ref,n} - r_{SL,ref,n+1} \cdot$$

Um den Potentiometerwert r_{SL} nun an die Referenzwerte der beiden Höhen anzugleichen, wird für die jeweilige Höhe ein angepasster Potentiometerwert berechnet. Dies geschieht, indem die Differenz $\Delta r_{SL,ref}$ anteilig vom Potentiometerwert r_{SL} abgezogen oder dazu addiert wird

$$r_{h_{S,n}} = r_{SL} + \Delta r_{SL,ref} \cdot e$$

$$r_{h_{S,n+1}} = r_{SL} - \Delta r_{SL,ref} \cdot (1 - e)$$

Mit den so modifizierten Werten, werden nun die benachbarten Referenzwerte bestimmt.

$$r_{SL,m,n} \leq r_{h_{S,n}} < r_{SL,m+1,n}$$

$$r_{SL,k,n+1} \leq r_{h_{S,n+1}} < r_{SL,k+1,n+1}$$

Anschließend werden noch die Abstände g_1 und g_2 berechnet

$$\hat{g}_1 = \frac{r_{h_{S,n}} - r_{SL,m,n}}{r_{SL,m+1,n} - r_{SL,m,n}}, \quad \hat{g}_2 = \frac{r_{h_{S,n+1}} - r_{SL,k,n+1}}{r_{SL,k+1,n+1} - r_{SL,k,n+1}} \cdot$$

Die Gleichung zur Berechnung des Korrekturwertes unter Verwendung der angepassten zweidimensionalen Kennfeldinterpolation $\hat{u}_{kl}^*(h_S, r_{SL})$ ist damit

$$\begin{aligned}\hat{u}_{kl}^*(h_S, r_{SL}) = & (1 - \hat{g}_1) \cdot (1 - e) \cdot u_{kl}(h_{S,n}, r_{SL,m,n}) \\ & + (1 - \hat{g}_2) \cdot e \cdot u_{kl}(h_{S,n+1}, r_{SL,k,n+1}) \\ & + \hat{g}_1 \cdot (1 - e) \cdot u_{kl}(h_{S,n}, r_{SL,m+1,n}) \\ & + \hat{g}_2 \cdot e \cdot u_{kl}(h_{S,n+1}, h_{SL,k+1,n+1}) .\end{aligned}$$

Durch Subtraktion des Korrekturwertes aus der zweidimensionalen Kennfeldinterpolation vom Sensorausgangswert erhält man eine Korrektur der Ausgangswerte des kapazitiven Sensors einer Schutzlamelle

$$\hat{u}_{lam}^*(t) = u_{lam}(t) - \hat{u}_{kl}^*(h_S, r_{SL}) .$$

Korrektur und Normierung: Um negative Werte für $\hat{u}_{lam}^*(t)$ zu verhindern, wird eine Fallunterscheidung eingeführt, die negative Werte zu 0 setzt. Die Normierung auf den Maximalwert eins wird durch die Division eines konstanten Wertes erreicht. Dazu wurde empirisch der Grenzwert $u_{lam,max}$ bestimmt, der erreicht werden muss, damit auf eine Hand geschlossen werden kann. Teilt man den korrigierten Sensorwert durch diesen Grenzwert und schneidet zusätzlich noch die Werte größer eins ab, so ergibt sich folgende Formel zur Korrektur und Normierung der Ausgangssignale eines kapazitiven Sensors in einer Schutzlamellen

$$u_{lam,norm}(t) = \begin{cases} 0 & \text{für } u_{lam}(t) < \hat{u}_{kl}^*(h_S, r_{SL}) \\ \frac{u_{lam}(t) - \hat{u}_{kl}^*(h_S, r_{SL})}{u_{lam,max}} & \text{für } \hat{u}_{kl}^*(h_S, r_{SL}) \leq u_{lam}(t) \leq (\hat{u}_{kl}^*(h_S, r_{SL}) + u_{lam,max}) \\ 1 & \text{für } u_{lam}(t) > (\hat{u}_{kl}^*(h_S, r_{SL}) + u_{lam,max}) . \end{cases}$$

4.3.3 Diskretisierung der Sensordaten

Diskrete Bayes'sche Netze sind weniger rechenintensiver als welche mit kontinuierlichen Variablen. Aus diesem Grund werden die auf den $[0,1]$ -Wertebereich normierten und korrigierten Sensordaten zusätzlich noch in entsprechende Zustände z_i diskretisiert. Die Zustände für den thermopilen Sensor lauten: *kein Objekt*, *kaltes Objekt* und *warmes Objekt*. Für die kapazitiven Sensoren werden folgende Zustände definiert: *kein Objekt*, *Objekt mit kleinem ϵ_r* und *Objekt mit großem ϵ_r* .

Bei der klassischen Diskretisierung werden Schwellwertverfahren eingesetzt. Dabei können Zustandssprünge, aufgrund der eindeutig festgelegten und nicht überlappenden Zustandsübergänge, auftreten. Dies bedeutet, dass z. B. das System bei einem fallenden Sensorwert plötzlich keine Hand mehr im Detektionsbereich erkennt. Richtig wäre jedoch, dass die Wahrscheinlichkeit für das Vorhandensein einer Hand abgenommen hat, da letztendlich nicht sicher ist, ob die Hand wirklich aus dem Detektionsbereich verschwunden ist oder sich am Rand des Bereiches befindet.

Die Fuzzifizierung der Fuzzy-Logik [53] erlaubt es, mithilfe der so genannten Zugehörigkeitsfunktionen, diese Zustandssprünge zu vermeiden. Jedoch liefert dieser Bearbeitungsschritt Zugehörigkeitsgrade eines Wertes zu verschiedenen Fuzzy-Mengen. Für ein Bayes'sches Netz sind jedoch Wahrscheinlichkeiten zur Repräsentation der Eingangsgrößen notwendig. Durch Modifikation der Zugehörigkeitsfunktionen können Wahrscheinlichkeitsfunktionen erzeugt werden. Der Unterschied zwischen einer Zugehörigkeitsfunktion und einer Wahrscheinlichkeitsfunktion liegt darin, dass eine Wahrscheinlichkeitsfunktion auf eins normiert ist und daher die Summe der Zustandswahrscheinlichkeiten für einen Sensorwert eins ergeben muss.

Als Basis für die Wahrscheinlichkeitsfunktion dient die Gauß Glockenkurve [14]

$$\Psi(u) = e^{-\frac{(u-m)^2}{2\sigma^2}}, \quad (4.8)$$

mit m als Mittelwert und σ^2 der Varianz.

Mit der Varianz wird der Wendepunkt der Zugehörigkeitsfunktion festgelegt. Mit $\sigma^2 = 0,01$ ist dieser nahe am jeweiligen Mittelwert, so dass gut separierbare Wahrscheinlichkeitsfunktionen entstehen. An den experimentell ermittelten Grenzen $u_{g,1}$ und $u_{g,2}$ sollen die Glockenkurven der benachbarten Zustände jeweils den Funktionswert $\mu_{z_i}(u_{g,k}) = 0,5$ besitzen. Mit diesen Werten können die Mittelwerte $m_{g,k,j}$ aus Gleichung (4.8) mit

$$m_{g,k,j} = u_{g,k} \pm \sqrt{-2\sigma^2 \cdot \ln(\Psi(u_{g,k}))}$$

berechnet werden.

Damit können nun die Zugehörigkeitsfunktionen bestimmt werden

$$\mu_{z_1}(u) = \begin{cases} 1 & \text{für } u \leq m_{g,1,1} \\ e^{-\frac{(u-m_{g,1,1})^2}{2\sigma^2}} & \text{für } u > m_{g,1,1} \end{cases}$$

$$\mu_{z_2}(u) = \begin{cases} e^{-\frac{(u-m_{g,1,2})^2}{2\sigma^2}} & \text{für } u < m_{g,1,2} \\ 1 & \text{für } m_{g,1,2} \leq u \leq m_{g,2,1} \\ e^{-\frac{(u-m_{g,2,1})^2}{2\sigma^2}} & \text{für } u > m_{g,2,1} \end{cases}$$

$$\mu_{z_3}(u) = \begin{cases} e^{-\frac{(u-m_{g,2,2})^2}{2\sigma^2}} & \text{für } u < m_{g,2,2} \\ 1 & \text{für } u \geq m_{g,2,2} . \end{cases}$$

Um aus den Zugehörigkeitsfunktionen Wahrscheinlichkeitsfunktionen für einen Zustand $\Psi_{z_i}(u)$ zu erhalten, wird zuerst die Summe der Wahrscheinlichkeitsfunktionen $S_{z_i}(u)$ über die entsprechenden Zustände $\mu_{z_i}(u)$ gebildet. Die Werte, welche größer als 1 sind, werden durch die Differenz $\Delta S_{z_i}(u)$

$$\Delta S_{z_i}(u) = S_{z_i}(u) - 1 = \mu_{z_1}(u) + \mu_{z_2}(u) + \mu_{z_3}(u) - 1$$

dargestellt.

Der Anteil der Summe $S_{z_i}(u)$, der größer als eins ist, wird von der Zugehörigkeitsfunktion mit dem größten Wert abgezogen. Damit erhält man die Zustandswahrscheinlichkeitsfunktionen zur Diskretisierung der Sensorwerte

$$\psi_{z_1}(u) = \begin{cases} \mu_{z_1}(u) - \Delta S_{z_i}(u) & \text{für } u < u_{g,1} \\ \mu_{z_1}(u) & \text{für } u \geq u_{g,1} \end{cases} \quad (4.9)$$

$$\psi_{z_2}(u) = \begin{cases} \mu_{z_2}(u) & \text{für } u \leq u_{g,1} \\ \mu_{z_2}(u) - \Delta S_{z_i}(u) & \text{für } u_{g,1} < u < u_{g,2} \\ \mu_{z_2}(u) & \text{für } u \geq u_{g,2} \end{cases} \quad (4.10)$$

$$\psi_{z_3}(u) = \begin{cases} \mu_{z_1}(u) & \text{für } u \leq u_{g,2} \\ \mu_{z_1}(u) - \Delta S_{z_i}(u) & \text{für } u > u_{g,2} . \end{cases} \quad (4.11)$$

Der Verlauf der Wahrscheinlichkeitsfunktionen $\psi_{z_i}(u)$ ist in Abbildung 4.11 dargestellt. Durch obige Gleichungen wird der durch die Vorverarbeitung erreichte Wertebereich $[0,1]$ auf die jeweiligen drei Zustände abgebildet. Dabei werden für jeden Sensorwert Wahrscheinlichkeiten für die drei Zustände bestimmt, wobei die Summe der Wahrscheinlichkeiten wieder eins ergibt.

4.4 Informationsfusion mittels Bayes'schem Netz

Informationen aus Datenquellen und vorhandenes Wissen lassen sich auf unterschiedliche Art und Weise verknüpfen. In diesem Abschnitt wird eine Informationsfusion vorgestellt, die die Daten der Sensorsysteme aus den Kapitel 2 und

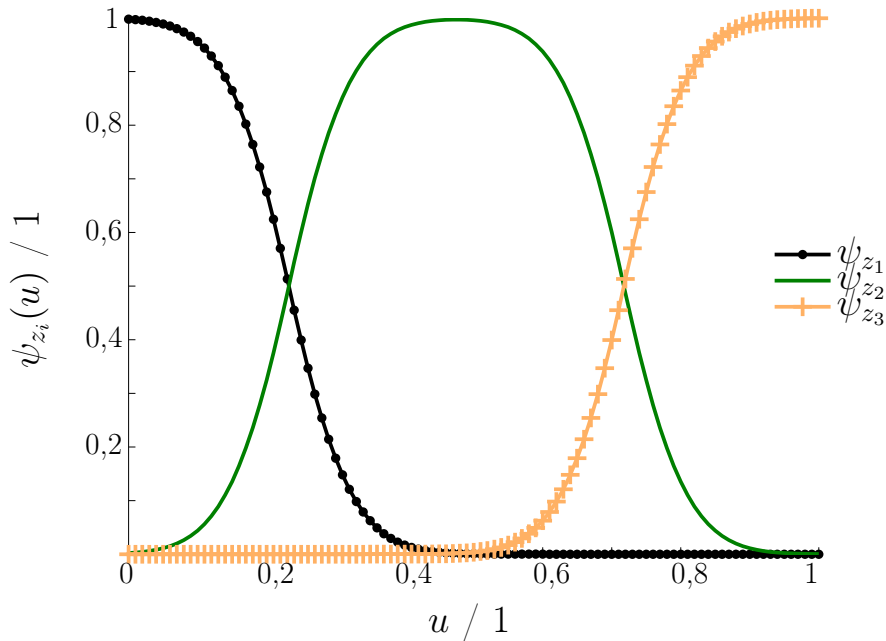


Abbildung 4.11: Wahrscheinlichkeitsfunktionen zur Diskretisierung der Sensorwerte

3 mithilfe eines Bayes'schen Netzes verknüpft. Die notwendige Vorverarbeitung und Diskretisierung wurde im vorangegangenen Abschnitt vorgestellt, ebenso wie die Grundlagen zu den Bayes'schen Netzen.

Im Folgenden wird das neu entwickelte Fusionskonzept für eine Gefahrenerkennung genauer erläutert. Im Anschluss daran werden Teilnetze zur Fusion entworfen und in Abschnitt 4.4.5 das komplette Bayes'sche Netz zur Informationsfusion vorgestellt.

4.4.1 Fusionskonzept zur Gefahrenanalyse

In Abschnitt 4.1 wurden verschiedene Fusionsstrukturen und -arten vorgestellt. Diese werden nun hinsichtlich einer Gefahrenanalyse mittels Sensordaten bewertet und verknüpft. Des Weiteren ist zu beachten, dass die Datenverarbeitung bei Sicherheitssystemen ein Kompromiss zwischen Zuverlässigkeit, Robustheit und Schnelligkeit zugrunde liegt. Deshalb scheiden komplizierte Datenbanksuchen bei Expertensystemen oder eine iterative Entscheidungsfindung aus. Außerdem sollte vorhandenes Vorwissen in das System einfließen, aber ohne eine implizit subjektive Entscheidungsfindung zu fokussieren, wie z. B. Fuzzy Logik oder Neuronale Netze.

Aus diesen Gründen werden für die Fusion zur Gefahrenanalyse beim JDL-Modell aus Abbildung 4.1 nur die *Level 0* bis *Level 3* betrachtet bzw. beim Omnibus-Modell in Abbildung 4.2 würde das *Sensormanagement* entfallen. Durch dieses Vorgehen ist eine fortlaufende Optimierung des Verarbeitungsprozesses hinsichtlich der Risikominimierung bzw. Zuverlässigkeitssteigerung nicht mehr möglich,

jedoch ist die Reaktionszeit wesentlich höher. Um trotzdem eine Risikominierung bzw. Zuverlässigkeitssteigerung zu ermöglichen, muss bei weiteren Verarbeitungsschritten Vorwissen in das System eingebracht werden.

Die Sensordaten der in Kapitel 2 und 3 vorgestellten Systeme zur Handerkennung werden auf der Symbolebene miteinander fusioniert. Durch die dabei erfolgte Datenreduktion ist ein schnelles Abarbeiten der anfallenden Daten möglich. Der Systemaufbau, s. Kapitel 5, ist so entwickelt, dass die Sensoren in einer zeitlichen Abfolge eine Hand registrieren. Aufgrund der sequenziell verfügbaren Sensorinformationen eignet sich eine kooperative Datenauswertung.

Eine lokale Fusion benötigt für jeden Sensor eine eigene Auswerteeinheit, welche die Kosten für das Gesamtsystem erhöht. Die Kommunikationslast ist allerdings stark reduziert. Eine zentrale Fusionseinheit muss hingegen in einem gewissen Maß robust gegen Kommunikationsfehler und Fehler aufgrund von Störeinkopplungen sein. Eine zentrale Einheit ist aber hinsichtlich der Kosten günstiger, da nicht für jeden Sensor eine eigene Auswerteeinheit benötigt wird.

Aus all diesen Betrachtungen ergeben sich folgende Anforderungen an das Fusionssystem:

1. Hohe Zuverlässigkeit, ausreichende Robustheit und Schnelligkeit des Gesamtsystems.
2. Um dem System Vorwissen zur Verfügung zu stellen, müssen geeignete Strukturen und Wissensrepräsentationsmethoden vorhanden sein, ohne dass die vorangegangene Forderung verletzt wird. Dadurch scheidet u. a. iterative Fusionsprozesse, Neuronale Netze und Datenbanksysteme zur Gefahrenerkennung aus.
3. Aufgrund des Systemaufbaus eignet sich eine kooperative Datenauswertung.
4. Die Sensorsysteme sind derart gestaltet, dass eine kostengünstige zentrale Fusion erfolgt, auch hinsichtlich der raumoptimalen Integration der Sensoren in ein System, s. Kapitel 5.
5. Die resultierende Entscheidung aus dem Fusionsprozess muss eindeutig sein.

Mit keiner der in Abschnitt 4.1.5 beschriebenen Fusionsmethoden lassen sich alle der obigen Forderungen erfüllen. Es ist jedoch möglich, mit einem auf Voting Logik basierenden Fusionssystem eine Gefahrenerkennung zu realisieren. Jedoch verhindern nicht-adaptierbare Schwell- und Grenzwerte eine ausreichende Robustheit des Systems.

Die Bayes Statistik ist ein mächtiges Werkzeug, um Rückschlüsse auf Ursachen zu ziehen oder um kausales oder prädiktives Schließen zu ermöglichen. Umfangreiches Vorwissen und zeitliche Zusammenhänge können jedoch nur begrenzt dargestellt werden.

Eine Abhilfe hierzu sind die Bayes'schen Netze, mit denen Vorwissen einfach graphisch und in Form von Wahrscheinlichkeitsfunktionen eingebracht werden kann. Ebenso können zeitliche Zusammenhänge durch dynamische Bayes'sche Netze dargestellt werden.

Die Struktur des Netzes hat noch weitere Funktionen. So wird implizit eine kooperative Datenauswertung durchgeführt und eine zentrale Fusionsstruktur ermöglicht. In Kombination mit einer Vorverarbeitung und dem Entwurf der Zustandswahrscheinlichkeiten kann eine hohe Robustheit des Systems erreicht werden. Ein weiterer Vorteil eines Bayes'schen Netzes ist dessen Modularität beim Entwurf. So können mehrere übersichtliche Teilnetze zu einem Gesamtnetz zusammen gefügt werden, wodurch die Zuverlässigkeit des Systems beim Entwurf ständig überprüft und optimiert werden kann.

Bayes'sche Netze können auch als Modelle von Prozessen aufgefasst werden, wodurch kausale Zusammenhänge darstellbar sind. Dies kann nun für eine automatische, computerbasierte Entscheidungsfindung ausgenutzt werden [44], d. h. die vom Netz schlussendlich erzeugten Wahrscheinlichkeiten können zur Hypothesenauswahl verwendet werden. Da die Bayes Statistik und damit auch das Bayes'sche Netz nur disjunkte Hypothesen bzw. Zustände zulässt, ist die getroffene Hypothesenwahl eindeutig. Die unterschiedlichen Gefahrenstufen lassen sich als Hypothesen darstellen und somit ist ein Bayes'sches Netz zur Gefahrenanalyse geeignet.

4.4.2 Entwurf der Bayes'schen Netze

In Abschnitt 4.2.3 werden die unterschiedlichen Entwurfsmöglichkeiten eines Bayes'schen Netzes aufgeführt. Abhängig vom Einsatzgebiet des Netzes und der Verfügbarkeit von umfassenden Lerndaten werden die Struktur, die Parameter (Wahrscheinlichkeitstabellen) und die Inferenzmethode festgelegt. Lerndaten, die eine umfassende Abdeckung sämtlicher Ereignisse und Zustände besitzen, sind im Allgemeinen nicht verfügbar. Außerdem besteht das Risiko, dass beim Entwurf durch fehlende Lerndaten das System auf entsprechende Fälle unerwartet reagiert oder dass eine *Überanpassung* (engl. *overfitting*) des zu lernenden Systems durch einseitige Lerndatenrepräsentation erfolgt. Aus diesen Gründen ist es vorteilhaft das Bayes'sche Netz zur Gefahrenerkennung manuell zu entwerfen. Der Erfolg des Netzentwurfes hängt damit vom theoretischen Expertenwissen und den zur Verfügung stehenden Messdaten ab.

Die Struktur, die Parameter und die Inferenzmethode für das zu entwerfende Netz zur Gefahrenerkennung sollen ein exaktes, kausales Schließen ermöglichen,

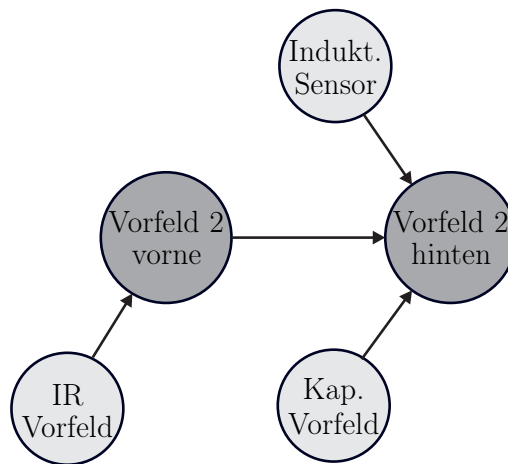


Abbildung 4.12: Bayes'sches Netz für *Vorfeld 2*

so dass eine zuverlässige und robuste Informationsfusion basierend auf den vorverarbeiteten Sensordaten entsteht. Ausgehend von diesen Anforderungen werden im Folgenden drei Teilnetze entwickelt, die sich an den lokalen Gegebenheiten einer Formatkreissäge orientieren. Eine genaue Beschreibung dazu erfolgt in Abschnitt 5.3.2.

Der hier durchgeführte manuelle Netzentwurf ermöglicht die Entwicklung und das anschließende Zusammenfügen der Teilnetze. Das Netz wird durch die Bildverarbeitung und Ausgabeknoten erweitert. Des Weiteren muss die zeitliche Abfolge der eintretenden Ereignisse beachtet werden. Dazu ist es ausreichend, eine Zeitverzögerung in das Netz einzubauen, so dass die Ergebnisse des Hypothesenknotens für eine bestimmte Zeit gespeichert werden. Das Schließen erfolgt mit exakter Inferenz, um eine hohe Genauigkeit zu erhalten. Die dazu benötigte Rechenkapazität ist im Vergleich zu den ungenaueren, approximativen Verfahren höher. Bei kleinen Netzen ist dies jedoch vernachlässigbar [67].

Überwachungszone

Die Überwachungszone ist in zwei Bereiche unterteilt, s. Abschnitt 5.3. Der größere Teil (*Vorfeld 1*) wird von einer Kamera überwacht. Deren Integration ins Bayes'sche Netz wird in Abschnitt 4.4.4 betrachtet. Der zweite Teil der Überwachungszone (*Vorfeld 2*) befindet sich unmittelbar vor der Schutzhaube, unter einer Erweiterung der Haube.

Dieser Bereich wird von einem thermopilen Infrarot-Sensor und einem kapazitiven Sensor überwacht. Zusätzlich ist in diesem Bereich noch ein induktiver Sensor zur Detektion von Verbundwerkstoffen angebracht. Diese drei Sensoren stellen die Informationsknoten für ein Netz dar. Aufgrund der räumlichen Trennung der Sensoren sind zwei Hypothesenknoten notwendig, so dass der überwachte Bereich in ein vorderes und ein hinteres Gebiet geteilt ist.

Knoten und Variablenname	Zustand	Zustandsbezeichnung
IR Vorfeld (IR_V), Vorfeld 2 vorne ($V2V$)	kein Objekt	<i>nichts</i>
	kaltes Objekt	<i>kalt</i>
	warmes Objekt	<i>warm</i>
Kap Vorfeld (KAP_V)	kein Objekt	<i>nichts</i>
	Objekt mit kleinem ϵ_r	<i>kleines ϵ_r</i>
	Objekt mit großem ϵ_r	<i>großes ϵ_r</i>
Indukt. Sensor (MAT)	Metall	<i>metall</i>
	kein Metall	<i>kein metall</i>
Vorfeld 2 hinten ($V2H$)	kein Objekt	<i>nichts</i>
	Holz	<i>holz</i>
	Metall	<i>metall</i>
	Hand	<i>hand</i>

Tabelle 4.2: Bezeichnungen der Zustände und Variablen sowie deren Abkürzungen

In Abbildung 4.12 ist das entworfene Netz dargestellt, das die räumliche Anordnung der Sensoren wie auch deren Detektionsort automatisch als Vorwissen enthält. Die hellen Knoten sind die Informationsknoten, die dunklen die Hypothesenknoten.

Die Zustände der beiden Knoten *IR Vorfeld* und *Vorfeld 2 vorne* sind identisch. In diesem Teilgebiet kann zwischen keinem, warmem oder kaltem Objekt unterschieden werden. Durch die zusätzlich einfließenden Evidenzen aus den Informationsknoten *Induktiver Sensor* und *Kap. Vorfeld* kann der Knoten *Vorfeld 2 hinten* zwischen keinem Objekt, Holz, Metall und Hand unterscheiden. In Tabelle 4.2 sind die im Netz verwendeten Variablen und mögliche Zustände aufgeführt.

Das bisher entworfene Netz besitzt drei Knoten ohne Elternknoten. Die Wahrscheinlichkeitsverteilung dieser Knoten besitzt während des normalen Betriebes keinen Einfluss auf die Informationsfusion, da für diese Knoten immer Evidenz in Form von vorverarbeiteten Sensorwerten vorliegt. Beim Ausfall eines Sensors kann auf dessen Initialisierungswahrscheinlichkeit zurückgegriffen werden. Da in diesem Fall keine Informationen über die Zustände des entsprechenden Knoten vorliegen, werden diese als gleichverteilt angenommen.

$$\begin{aligned}
 P(IR_V) &= P(\textit{nichts}, \textit{kalt}, \textit{warm}) = \left(\frac{1}{3}; \frac{1}{3}; \frac{1}{3}\right) \\
 P(KAP_V) &= P(\textit{nichts}, \textit{kleines } \epsilon_r, \textit{großes } \epsilon_r) = \left(\frac{1}{3}; \frac{1}{3}; \frac{1}{3}\right) \\
 P(MAT) &= P(\textit{kein metall}, \textit{metall}) = (1; 0)
 \end{aligned}$$

Die Übergangswahrscheinlichkeiten repräsentieren, ebenso wie die Netzstruktur, in die Fusion eingebrachtes Vorwissen. Die Tabelle für die bedingte Übergangswahrscheinlichkeiten $P(V2V|IR_V)$ und $P(V2H|V2V, KAP_V, MAT)$

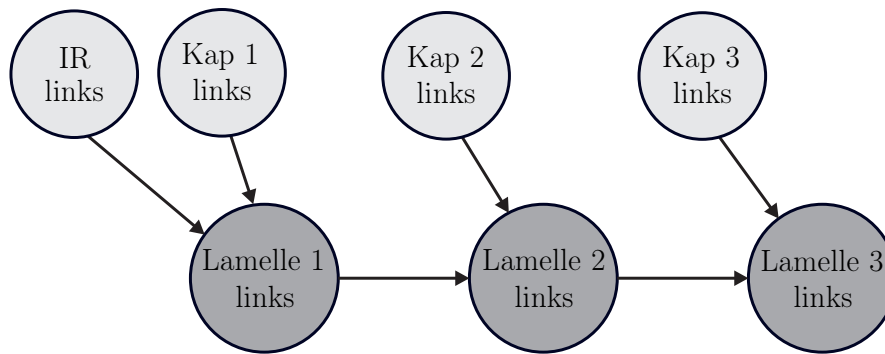


Abbildung 4.13: Bayes'sches Netz für die linke Seite der *Gefährdungszone*

der beiden Hypothesenknoten werden in Abschnitt 4.4.4 aufgrund der Einbindung der Bildverarbeitung noch modifiziert. Damit sind alle notwendigen Parameter und die Struktur des Netzes für eine Informationsfusion definiert.

Gefährdungszone

Die Gefährdungszone wird auf der rechten und linken Seite der Schutzhaube von jeweils drei kapazitiven Sensoren und einem Infrarot-Sensor überwacht. Die Gefahrenanalyse erfolgt für jede Seite separat, da die Erkennung einer Hand auf der einen Seite keinen Einfluss auf die andere Seite hat. Die Zustandswahrscheinlichkeiten der beiden Seiten beeinflussen sich gegenseitig nicht. Dadurch muss nur für eine Seite ein Teilnetz entworfen werden, das jedoch für beiden Seiten zur Fusion verwendet werden kann.

Es stehen vier Informationsquellen, die durch Informationsknoten repräsentiert werden, zur Verfügung. Die Sensoren überwachen unterschiedliche Bereiche (s. Abschnitt 5.3), weshalb für jeden Teilbereich ein Hypothesenknoten verwendet wird. Die einzelnen Bereiche decken dabei die Ausdehnung einer Schutzlamelle ab, so dass drei Hypothesenknoten notwendig sind, um die drei Lamellen pro Seite zu repräsentieren.

Aufgrund der Zuordnung der Teilbereiche zu den Lamellen, besitzt jeder kapazitive Sensor eine Kante zu dem Knoten, dessen Bereich die Lamelle begrenzt, in der dieser integriert ist. Der Infrarot-Sensor überwacht ungefähr den gleichen Bereich wie der erste kapazitive Sensor und wird deshalb ebenfalls mit dem Hypothesenknoten *Lamelle 1 links* verbunden. Letztendlich werden noch der erste und zweite und der zweite und dritte Hypothesenknoten verbunden, da eine Handdetektion im ersten Teilbereich die Wahrscheinlichkeit einer Hand im zweiten Teilbereich erhöht und eine Handdetektion im zweiten Teilbereich die Wahrscheinlichkeit des dritten Teilbereichs ebenfalls beeinflusst. In Abbildung 4.13 ist das Bayes'sche Netz zur Informationsfusion der linken Seite des Gefährdungsbereiches dargestellt.

Die Variablen der Informationsknoten (*IR links*, *Kap i links*) besitzen die gleichen drei Zustände, wie die des *Vorfeld 2*. Die Hypothesenvariablen verwenden die vier

Knoten, Variablenname	Zustand	Zustands- bezeichnung
IR links (IR_L)	kein Objekt	<i>nichts</i>
	kalter Gegenstand	<i>kalt</i>
	warmes Objekt	<i>warm</i>
Kap 1 links (KAP_L1), Kap 2 links (KAP_L2), Kap 3 links (KAP_L3)	kein Objekt	<i>nichts</i>
	Objekt mit kleinem ϵ_r	<i>kleines ϵ_r</i>
	Objekt mit großem ϵ_r	<i>großes ϵ_r</i>
Lamelle 1 links (LAM_L1), Lamelle 2 links (LAM_L2), Lamelle 3 links (LAM_L3)	kein Objekt	<i>nichts</i>
	Holz	<i>holz</i>
	Metall	<i>metall</i>
	Hand	<i>hand</i>

Tabelle 4.3: Bezeichnungen der Zustände und Variablen sowie deren Abkürzungen

Zustände: *kein Objekt, Holz, Metall, Hand*. Die Bezeichnung der Variablen und die Zustände sind in Tabelle 4.3 aufgeführt.

Die Initialisierungswahrscheinlichkeiten für die Informationsknoten, werden wiederum mit einer Gleichverteilung über alle möglichen Zustände festgelegt.

$$P(IR_L) = P(nichts, kalt, warm) = \left(\frac{1}{3}; \frac{1}{3}; \frac{1}{3}\right)$$

$$P(KAP_L1) = P(nichts, kleines \epsilon_r, großes \epsilon_r) = \left(\frac{1}{3}; \frac{1}{3}; \frac{1}{3}\right)$$

$$P(KAP_L2) = P(nichts, kleines \epsilon_r, großes \epsilon_r) = \left(\frac{1}{3}; \frac{1}{3}; \frac{1}{3}\right)$$

$$P(KAP_L3) = P(nichts, kleines \epsilon_r, großes \epsilon_r) = \left(\frac{1}{3}; \frac{1}{3}; \frac{1}{3}\right)$$

Sämtliche Tabellen der bedingten Übergangswahrscheinlichkeiten werden nach der Einbindung der Bildverarbeitung in Abschnitt 4.4.4 aufgeführt. Die Struktur, sowie die Wahrscheinlichkeitstabellen und die Initialisierungswahrscheinlichkeiten können unverändert für die rechte Seite der Gefährdungszone übernommen werden, siehe Abbildung 4.14.

4.4.3 Verknüpfung der Teilnetze

Die drei bisher entworfenen Teilnetze werden nun zu einem Netz zusammengefügt. Dazu wird der Knoten *Vorfeld 2 hinten* mit den Knoten *Lamelle 1 links* und *Lamelle 1 rechts* verbunden. Die beiden Zweige des Netzes, welche die Gefährdungszone darstellen, bleiben durch die Verknüpfung unabhängig und damit d-separiert, da aufgrund der Netzstruktur eine divergente Verbindung vorliegt und harte Evidenz für Knoten *Vorfeld 2 hinten* gilt.

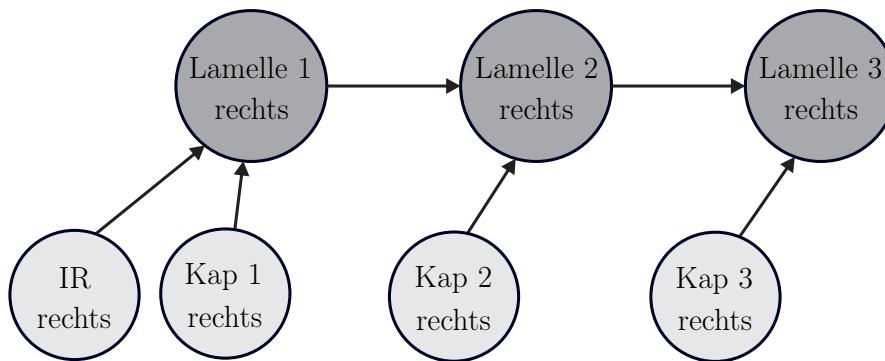


Abbildung 4.14: Bayes'sches Netz für die rechte Seite der *Gefährdungszone*

Aufgrund der Verknüpfung müssen die bedingten Wahrscheinlichkeitstabellen der beiden Knoten *Lamelle 1 links/rechts* erweitert werden, da beide jeweils einen neuen Elternknoten besitzen. Die bisherigen Werte werden beibehalten solange der Knoten *Vorfeld 2 hinten* den Zustand *nichts* liefert. Für die anderen Zustände *holz*, *metall* und *hand* müssen die bedingten Wahrscheinlichkeiten noch bestimmt werden. Dabei betrachtet man die Zustände der Knoten *IR links* bzw. *IR rechts* und *Kap 1 links* bzw. *Kap 1 rechts* unter der Voraussetzung, dass im hinteren Teil des *Vorfeld 2* bereits entweder Holz, Metall oder eine Hand detektiert wurde.

Bei dem betrachteten Fusionssystem werden die verteilten Sensordaten kooperativ integriert, d. h. die aktuellen Daten aller Sensoren werden gleichzeitig verarbeitet. Eine Hand wird aber nicht von allen Sensoren gleichzeitig registriert. Erkennt ein Sensorsystem eine Hand, so wird diese Information vernachlässigt, da die anderen Systeme die Hand zu diesem Zeitpunkt nicht erkennen.

Durch eine Wärmezunahme wird z. B. im *Vorfeld 2* die Wahrscheinlichkeit im Knoten *Vorfeld 2 hinten* für das Vorhandensein einer Hand aufgrund der Sensorwerte des Infrarot-Sensors erhöht. Ein kapazitiver Sensor in den Lamellen erkennt die Hand aber erst, wenn diese den vom thermopilen Sensor überwachten Bereich wieder verlassen hat. Die Information des Knoten *Vorfeld 2 hinten* ist für die weitere Fusion verloren.

Abhilfe könnte ein dynamisches Bayes'sches Netz schaffen [67]. Diese eignen sich jedoch nicht für einen kontinuierlichen Betrieb, da sich dadurch der Umfang des Netzes um Größenordnungen steigert und die Inferenz nicht mehr schnell und zuverlässig möglich ist. Des Weiteren müssten die Transitionen vom aktuellen zum nachfolgenden Zustand variabel sein, da sich eine Hand mit unterschiedlichen Geschwindigkeiten durch das System bewegt.

Eine einfach und effiziente Lösung ist die Speicherung der maximalen Wahrscheinlichkeit für den Zustand *hand* in einem Hypothesenknoten, um diese für die weitere Verarbeitung zu verwenden. In einem Bayes'schen Netz wird dies durch Vergleich des aktuellen Wertes eines Knotens $P_{t_n}(A)$ mit dem vorangegangenen Wert $P_{t_n-k}(A)$ erreicht. Ist die Wahrscheinlichkeit für den Zustand *hand*

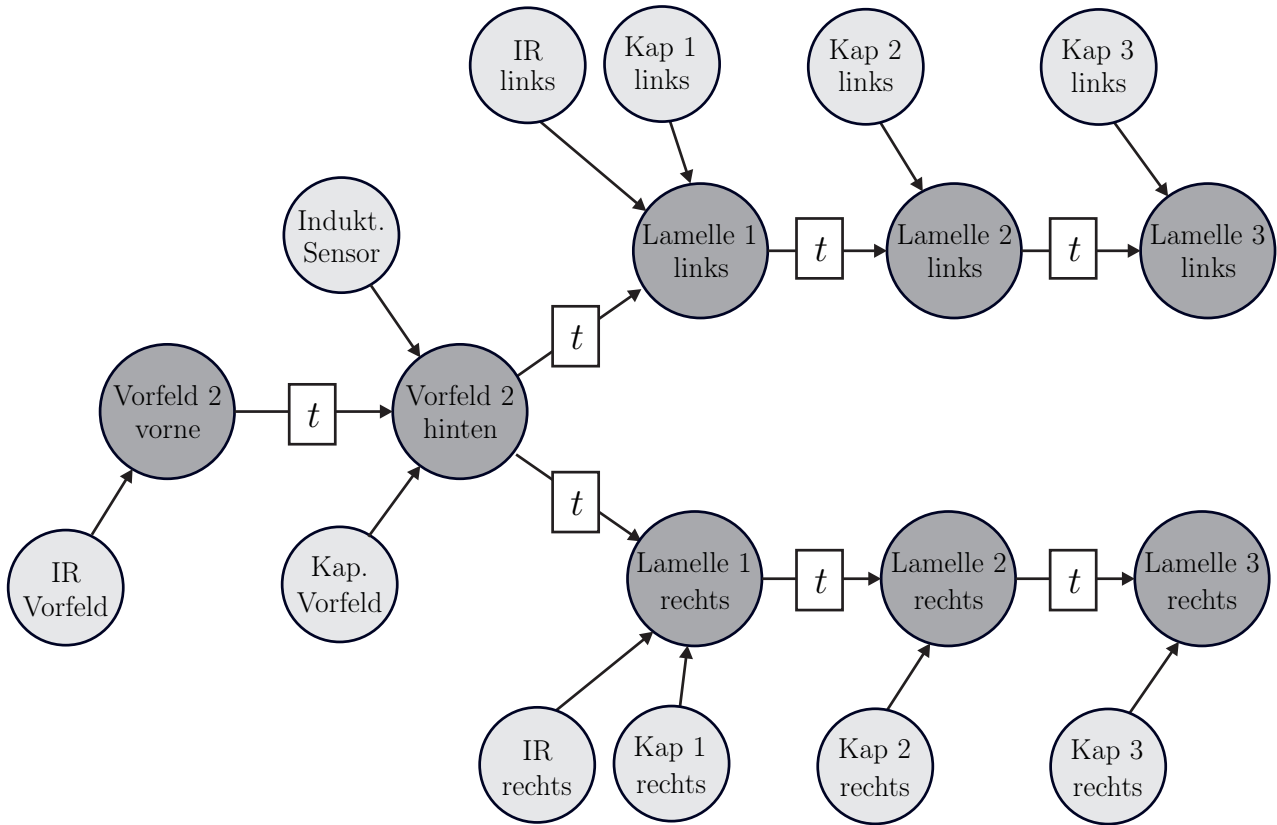


Abbildung 4.15: Bayes'sches Netz mit Zustandsspeicherung

höher als die der vorangegangenen Wahrscheinlichkeit, so wird die aktuelle Wahrscheinlichkeit übernommen. Ansonsten wird der alte Wert beibehalten. Zur Inferenz werden dann die Wahrscheinlichkeiten $P_{t_n,aktiv}(A)$ verwendet. Nach einem festen Zeitintervall $[0, k_{max}]$ werden die Wahrscheinlichkeiten dann wieder erneuert, auch wenn kein neues Maximum vorliegt. Das Zeitintervall muss so gewählt sein, dass langsame Handbewegungen ($v < 0.2 \text{ m/s}$) noch sicher erkannt werden können. Dies ist für $k_{max} = 20$ gewährleistet.

$$P_{t_n,aktiv}(A) = \begin{cases} P_{t_{n-k}}(A) & \text{für } P_{t_n}(A = hand) < P_{t_{n-k}}(A = hand) \text{ und } k < k_{max} \\ P_{t_n}(A) & \text{für } P_{t_n}(A = hand) \geq P_{t_{n-k}}(A = hand) \text{ oder } k \geq k_{max} \end{cases}$$

In Abbildung 4.15 ist das gesamte Netz zur Informationsfusion mit integrierter Zustandsspeicherung dargestellt.

4.4.4 Einbindung der Bildverarbeitung

Mit den Informationen, die aus der Bildverarbeitung der Kamerabilder stammen (s. Kapitel 3), wird das Bayes'sche Netz, wie in Abbildung 4.16 dargestellt, erweitert. Hierzu ist ein neuer Informations- und einer neuer Hypothesenknoten notwendig. Die Kamera (Informationsknoten) überwacht dabei das *Vorfeld 1* (Hypothesenknoten).

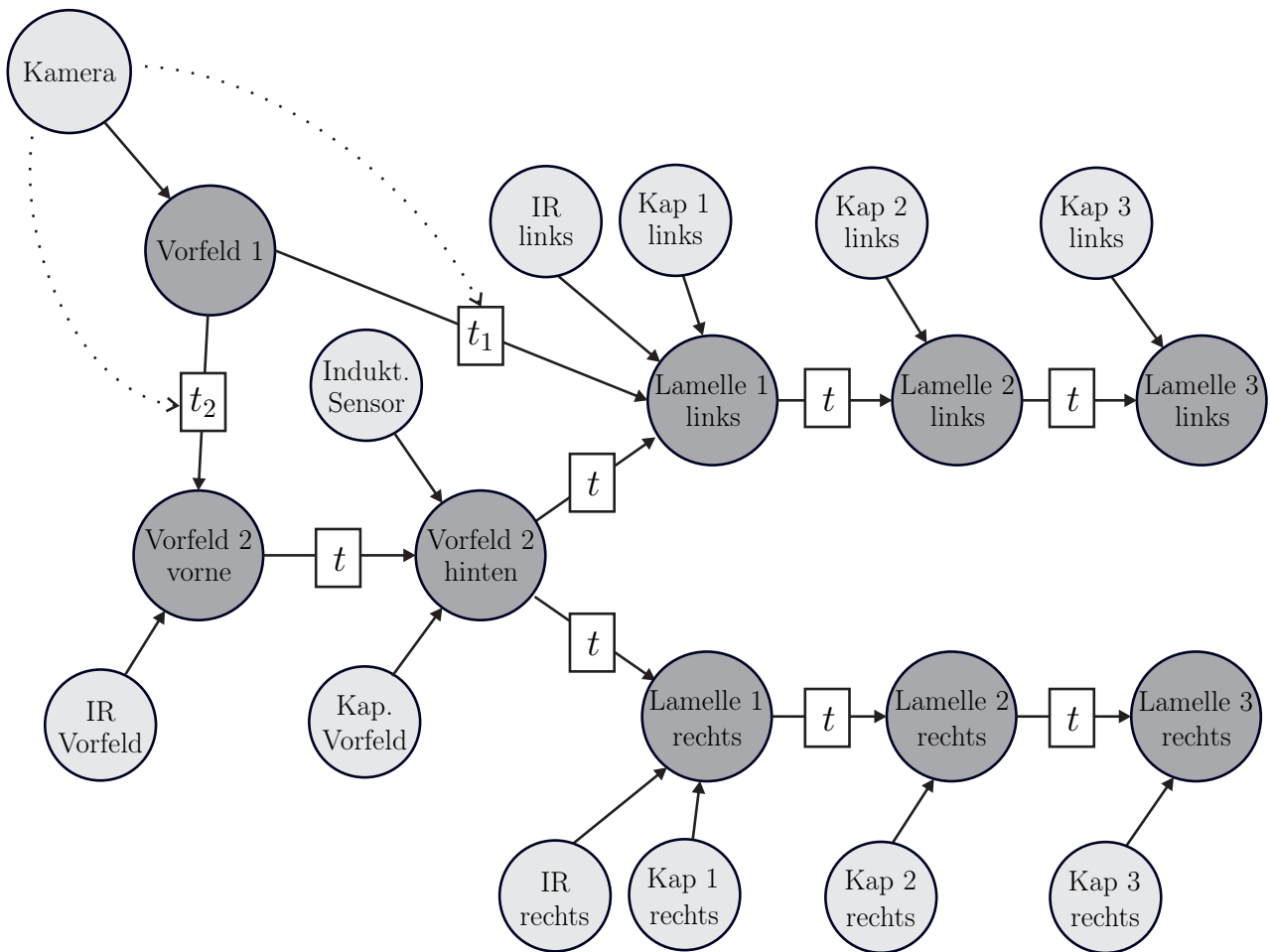


Abbildung 4.16: Bayes'sches Netz mit eingebundener Bildverarbeitung

Durch die Bildverarbeitung kann frühzeitig eine Hand im *Vorfeld 1* erkannt werden und die Bewegungsrichtung bestimmt werden. Diese Information und der berechnete Eintrittszeitpunkt t_2 wird an den Knoten *Vorfeld 2 vorne* weitergeleitet.

Die Hand kann auch direkt vom *Vorfeld 1* in die *Gefährdungszone* eindringen. Um dies zu berücksichtigen ist eine Verbindung von Knoten *Vorfeld 1* zu Knoten *Lamelle 1 links* notwendig. Des Weiteren erlaubt die Bildverarbeitung die Bestimmung des Eintrittszeitpunktes t_1 in den Bereich der *Lamelle 1 links*.

Die Variablen und Zustände der neuen Knoten sind in Tabelle 4.4 zu finden. Der Knoten *Kamera* besitzt nur die beiden Zustände *hand* und *keine hand*, da die Bildverarbeitung eine Hand zuverlässig entdeckt. Im Fehlerfall, also bei Anwesenheit einer Hand im *Vorfeld 1* und ein Nichterkennen der Hand durch die Bildverarbeitung, liefert zwar der Knoten *Vorfeld 1* den falschen Zustand *keine hand*, jedoch wird eine Hand trotzdem durch die Informationsfusion und die anderen involvierten Sensoren erkannt. Im Normalfall, also bei der richtigen Entdeckung einer Hand durch die Bildverarbeitung, reagiert die Informationsfusion

Knoten, Variablennamen	Zustand	Zustands- bezeichnung
Kamera (<i>KAM</i>)	keine Hand	<i>keine hand</i>
	Hand	<i>hand</i>
Vorfeld 1 (<i>V1</i>)	keine Hand	<i>keine hand</i>
	Hand	<i>hand</i>

Tabelle 4.4: Knoten, Variablennamen, Zustände und Zustandsbezeichnungen für die Netzerweiterung mit einer Kameraüberwachung

schneller, zuverlässiger und robuster als wenn keine Kamera vorhanden wäre. Die Initialisierungswahrscheinlichkeit für den Knoten *Kamera* ist damit

$$P(KAM) = P(\textit{keine hand}, \textit{hand}) = (1 ; 0) .$$

Durch die Anbindung des Knoten *Vorfeld 1* an die Knoten *Vorfeld 2 vorne* und *Lamelle 1 links* müssen deren Wahrscheinlichkeitstabellen erweitert werden. Der Knoten *Vorfeld 2 vorne* wird um den Zustand *hand* erweitert. Dieser ist aktiv, wenn eine Hand von der Bildverarbeitung detektiert und eine Temperaturerhöhung durch den Infrarot-Sensor registriert wird. Durch diese Erweiterung muss auch der Knoten *Vorfeld 2 hinten* ergänzt werden. Wird vom Knoten *Vorfeld 1 vorne* der Zustand *hand* und der Knoten *Kap Vorfeld* der Zustand *kleines ϵ_r* geliefert, ist der Zustand *hand* im Knoten *Vorfeld 2 hinten* aktiv.

Die Wahrscheinlichkeitstabelle des Knoten *Lamelle 1 links* wird um die Zustände vom neuen Elternknoten *Vorfeld 1* erweitert. Für den Zustand *keine hand* werden die vorhandenen Einträge übernommen. Der Zustand *hand* von Knoten *Vorfeld 1* bewirkt bei entsprechenden Informationen von den Knoten *IR links* und *Kap 1 links*, dass mit einer Wahrscheinlichkeit von eins der Zustand *hand* ausgegeben wird.

Die bedingten Wahrscheinlichkeitstabellen für das gesamte entworfene Bayes'sche Netz sind im Anhang F zu finden.

4.4.5 Bayes'sches Netz zur Gefahrenanalyse

Mit dem in Abbildung 4.16 vorgestellten Bayes'schen Netz kann mit den eingesetzten Sensoren eine Informationsfusion durchgeführt werden und die Gefahr für den Benutzer einer Formatkreissäge bestimmt werden. Dazu müssen die Wahrscheinlichkeitsverteilungen in den Hypothesenknoten ausgewertet werden. Dies erfolgt nach den Regeln der exakten Inferenz für Bayes'sche Netze mit der Berechnung der Wahrscheinlichkeiten, wie es in Abschnitt 4.2.2 festgelegt ist.

Die Auswertung der Wahrscheinlichkeiten lässt sich durch das Hinzufügen zweier weiterer Knoten visualisieren. In Abbildung 4.17 ist das entwickelte Bayes'sche Netz mit den beiden Knoten *Warnung* und *Stopp* visualisiert.

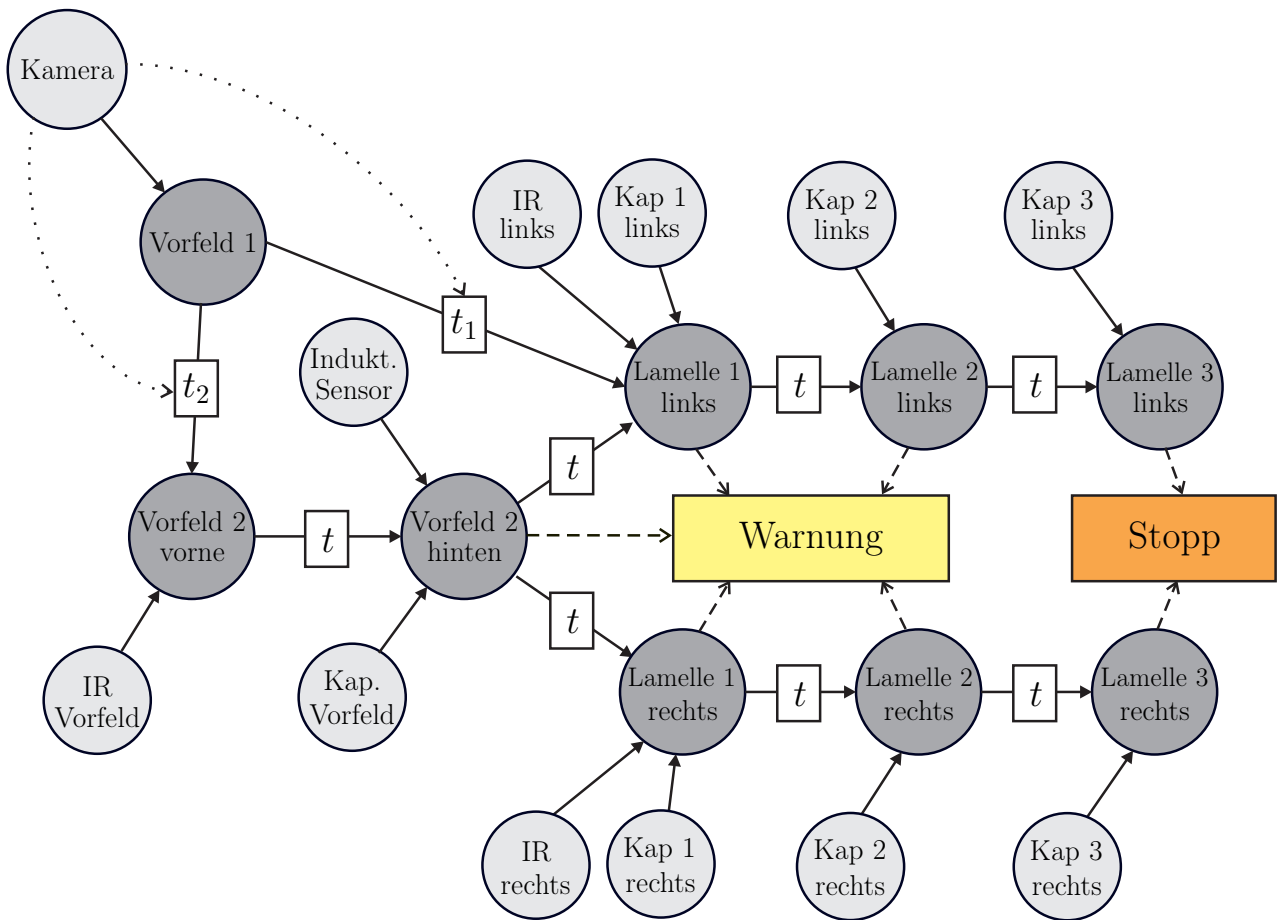


Abbildung 4.17: Bayes'sches Netz zur Gefahrenanalyse

Die Entscheidung, ob der Benutzer gewarnt oder das Sägeblatt gestoppt wird, ist abhängig von der Wahrscheinlichkeit für den Zustand *hand* in den entsprechenden Knoten. Mittels einer Grenzwertauswertung werden dann die Zustände *warnen* oder *stoppen* aktiv.

Für die Ausgabe einer frühzeitigen Warnung an den Benutzer, dass dessen Hand potentiell gefährdet ist, muss die Wahrscheinlichkeit für den Zustand *hand* in den Knoten *Vorfeld 2 hinten*, *Lamelle 1 links*, *Lamelle 2 links*, *Lamelle 1 rechts* und *Lamelle 2 rechts* den Wert 0,5 überschreiten.

$$\text{warnung} = \begin{cases} 1 & \text{für } P(V2H = \textit{hand}) > 0,5 \\ 1 & \text{für } P(LAM_L1 = \textit{hand}) > 0,5 \\ 1 & \text{für } P(LAM_L2 = \textit{hand}) > 0,5 \\ 1 & \text{für } P(LAM_R1 = \textit{hand}) > 0,5 \\ 1 & \text{für } P(LAM_R2 = \textit{hand}) > 0,5 \\ 0 & \textit{sonst} \end{cases} \quad (4.12)$$

Wird die Warnung vom Benutzer ignoriert und dessen Hand dringt in den Gefährdungsbereich bzw. in den Bereich unmittelbar vor das Sägeblatt ein, wird das Sägeblatt gestoppt. Für die Aktivierung des Zustandes *stopp* muss die Wahr-

scheinlichkeit für *hand* in den Knoten *Lamelle 3 links* und *Lamelle 3 rechts* den Wert 0,4 überschreiten.

$$\text{stopp} = \begin{cases} 1 & \text{für } P(LAM_L3 = \textit{hand}) > 0,4 \\ 1 & \text{für } P(LAM_R3 = \textit{hand}) > 0,4 \\ 0 & \textit{sonst} \end{cases} \quad (4.13)$$

Die beiden Werte 0,5 bzw. 0,4 für die Aktivierung von *warnung* bzw. *stopp* sind empirisch festgelegt worden und bieten den optimalen Kompromiss in Bezug auf Zuverlässigkeit, Robustheit und Reaktionszeit. Die Validierung und ein ausführliches Beispiel sind in Abschnitt 5.4 zu finden.

5 Funktionsstruktur des Gefahrenerkennungssystems

Aktive Sicherheitssysteme setzen sich aus zwei Aufgabengebieten zusammen. Ein Teil ist die Aktorik, mit der auf eine erkannte Gefahrensituation reagiert werden kann. Die Art und Weise der Reaktion ist vom eingesetzten Anwendungsfall abhängig. Das Erkennen von Gefahrensituationen ist der zweite Aufgabenteil. Hierbei wird durch eine Überwachung des Umfeldes der Gefahrenquelle das Gefahrenpotential für den menschlichen Benutzer kontinuierlich bestimmt. Zum einen muss eine Sensorik vorhanden sein, die die Überwachung durchführt. Zum anderen sind eine Auswertung der Sensorinformationen und eine Einschätzung der Gefahrenhöhe (Verletzungsrisiko) notwendig.

In den Kapiteln 2 und 3 sind perzeptive Systeme vorgestellt worden, mit denen eine menschliche Hand detektiert werden kann. Da jedes einzelne dieser Systeme unterschiedliche Detektionseigenschaften aufweist, ist es erforderlich die von den Sensorsystemen gelieferten Informationen zusammenzuführen. Dies ist in Kapitel 4 in Form einer auf dem Bayes'schen Netze basierenden Informationsfusion erklärt. Außerdem erfolgt in diesem Netzwerk die Einschätzung der aktuellen Gefahr für den Maschinenbenutzer sowie die Aktivierung der Aktorik.

In diesem Kapitel wird die Funktionsstruktur für eine Gefahrenanalyse am Beispiel einer Formatkreissäge vorgestellt. Die Konzepte können an beliebige andere Werkzeugmaschinen angepasst werden.

Zu Beginn dieses Kapitels werden geeignete Merkmale festgelegt, mit denen eine Gefahr für die menschliche Hand erkannt wird. Anschließend werden Bereiche an einer Formatkreissägemaschine festgelegt, für die eine technische und wirtschaftliche Überwachung sinnvoll ist. Die Realisierung der Überwachungssysteme, die Verarbeitungseinheit sowie der Aufbau an einer Formatkreissäge werden in Abschnitt 5.3 präsentiert. Das Kapitel schließt mit einer Validierung des Erkennungssystems.

5.1 Merkmale zur Gefahrenanalyse

Um eindeutige und reproduzierbare Entscheidungen für eine Gefahrenerkennung zu erhalten, werden Merkmale definiert, die eine Szene und die darin enthaltenen Dynamik abbilden. Unter einer Szene ist in diesem Zusammenhang eine Gefahrensituation für einen Menschen zu verstehen. Die Merkmale sind:

- **Objektidentifikation:** Die eindeutige Identifikation der Hand.
- **Objektposition:** Der aktuelle Ort einer Hand in Bezug auf die Gefahrenquelle (z. B. das Sägeblatt einer Formatkreissäge).
- **Objektbewegung:** Die Geschwindigkeit und Richtung einer Hand relativ zur Gefahrenquelle im überwachten Bereich.

Ein Überwachungssystem muss daher Sensorsysteme besitzen, welche diese drei Merkmale aus einer Szene extrahieren können. Dabei ist es möglich, Systeme einzusetzen, die alle drei Merkmale extrahieren können, wie z. B. ein Bildverarbeitungssystem. Die Objektbewegung kann aus einer Kombination von gleichartigen und auch verschiedenen Sensortypen bestimmt werden, wie z. B. kapazitive und Infrarot-Sensoren. Ebenso kann ein Objekt durch unterschiedliche Sensoren identifiziert werden. Die Kombination von Sensoren zur Merkmalsgewinnung erfolgt mittels einer Informationsfusion. Hierbei wird durch die einzelnen Sensorinformationen auf ein Merkmal geschlossen.

Die so gewonnen Merkmale einer Szene werden einer Analyseeinheit zugeführt. Diese bewertet die Merkmale, beispielsweise anhand von bekannten Grenzwerten. Dadurch kann auf eine potentielle Gefahr geschlossen und durch eine nachfolgende Entscheidungseinheit eine Reaktion des Sicherheitssystems ausgelöst werden. Für die Bewertung der Merkmale und die Entscheidungsfindung wird ein System zur Informationsfusion verwendet, s. Abschnitt 5.3.3.

In Abschnitt 5.3.2 werden Sensorsysteme beschrieben, mit denen die genannten drei Merkmale aus einer Szene an einer Formatkreissäge gewonnen werden. Dabei handelt es sich um Infrarot- und kapazitive Sensoren sowie einer Kamera mit einer Bildverarbeitung.

5.2 Gefahrenzonen und Überwachungsbereiche am Beispiel einer Formatkreissäge

Als Gefahrenzonen werden diejenigen Bereiche einer Maschine bezeichnet, in denen eine Person einer Gefährdungssituation ausgesetzt ist und ein Verletzungsrisiko besteht, s. Abschnitt 1.2. Die zu überwachenden Gebiete müssen deshalb so ausgelegt sein, dass eine Gefährdung des Benutzers durch das Sicherheitssystem verhindert werden kann. Andererseits müssen technische und ökonomische Aspekte berücksichtigt werden. Ein technisch aufwendiges bzw. ein überproportional zum Nutzen teures Sicherheitssystem wird vom Anwender bzw. Käufer nicht akzeptiert. Aus diesem Grund ist ein Kompromiss notwendig, der diese Aspekte berücksichtigt und gleichzeitig das tolerierbare Risiko [23] für den Anwender nicht übersteigt.

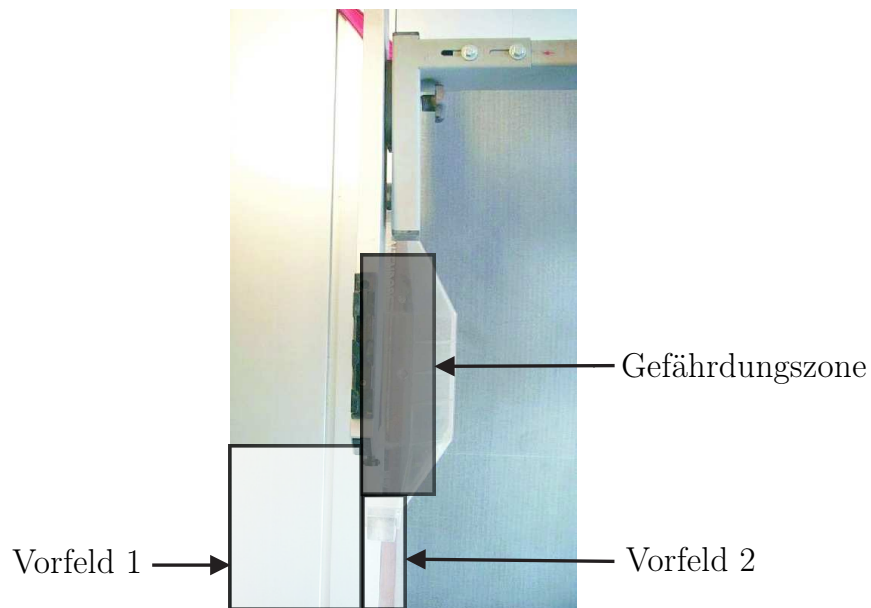


Abbildung 5.1: Zu überwachende Gebiete einer Formatkreissäge

Ein aktives Sicherheitssystem für Formatkreissägen muss diese Bedingungen einhalten. Im ersten Schritt werden die eigentlichen Gefahrenbereiche festgelegt. Im zweiten Schritt müssen geeignete Sensorsysteme zur Überwachung gefunden werden, s. Abschnitt 5.3.2, und die dabei erhaltenen Informationen verarbeitet und bewertet werden. Im letzten Schritt ist die Aktorik des Systems zu definieren, s. Abschnitt 5.3.4.

In den Abbildungen 1.3 und 1.4 werden die Bereiche einer Formatkreissäge, in denen ein potenzielles Verletzungsrisiko für die Hand eines Maschinenbenutzers besteht, vorgestellt. Basierend auf den dabei festgelegten Gefahrenbereichen, sind in Abbildung 5.1 die Gebiete skizziert, welche überwacht werden müssen damit obiger Kompromiss erfüllbar ist. In den markierten Gebieten finden auch die meisten Arbeitsabläufe statt [32]. Die verbleibenden Gebiete des Überwachungsbereiches werden nicht überwacht, da dies zusätzlichen technischen Aufwand und damit Kosten verursacht.

Die Größe der *Gefährdungszone* entspricht der der Schutzhaube. Der Überwachungsbereich ist in drei Zonen aufgeteilt, wobei die beiden Vorfelder aus Abbildung 5.1 überwacht werden. Das *Vorfeld 1* hat die Abmessungen $0,28 \times 0,38$ m. Das *Vorfeld 2* entspricht einem Rechteck mit den Maßen $0,25 \times 0,06$ m. Damit ist die Bewertungsgrundlage für das Merkmal *Objektposition* aus Abschnitt 5.1 definiert.

Das Gefahrenmerkmal *Objektbewegung* lässt sich durch die Objektgeschwindigkeit und die Bewegungsrichtung charakterisieren. Geschwindigkeiten bis $0,67$ m/s gelten als unkritisch, da dies der maximalen Vorschubgeschwindigkeit entspricht. Werte darüber werden als kritisch eingestuft. Reflexbewegungen bzw. schnelle Wischbewegungen besitzen i. A. Geschwindigkeiten bis 2 m/s [98].

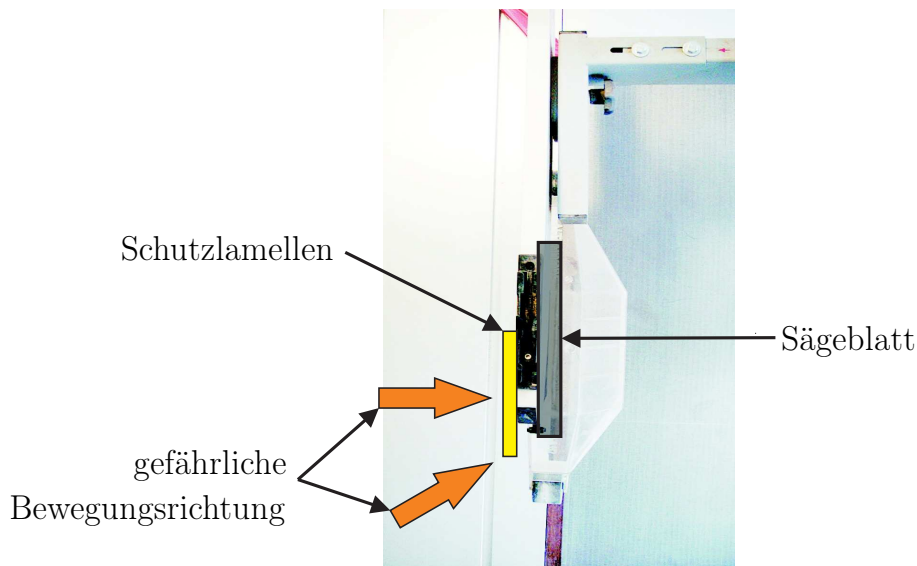


Abbildung 5.2: Gefährliche Bewegungsrichtungen an einer Formatkreissäge

5.3 Realisierung und Implementierung

Bereits in Abbildung 1.7 ist die Struktur des in dieser Arbeit entwickelten Sicherheitssystems dargestellt. In den vorangegangenen Kapiteln sind die eingesetzten Sensorsysteme und Datenverarbeitungsprinzipien vorgestellt worden. In diesem Abschnitt wird nun die Integration der Einzelsysteme ins Gesamtsystem präsentiert. Dazu wird das Konzept der hier verwendeten Schutzlamellen erläutert, die Sensorsysteme im Überblick sowie deren Anordnung gezeigt. Anschließend wird die Datenverarbeitung, die entwickelte Gefahrenanalyse und die eingesetzte Aktorik vorgestellt. Abschließend wird auf den modularen Aufbau und dessen Vorteile für das Gesamtsystem eingegangen.

5.3.1 Schutzlamellen

Die in Abbildung 5.2 skizzierten Handbewegungen in Richtung Sägeblatt sind als gefährlich einzustufen, da die Schutzhaube das Sägeblatt nicht abdeckt, so dass versehentlich in den Bereich eingebrachte Finger nicht geschützt sind. Eine rechtzeitige und zuverlässige Erkennung einer Gefahrensituation für die Hand ist mit einer berührungslos arbeitenden Sensorik in diesem Fall nicht gewährleistet. Es ist eine zuverlässige und präzise Handpositionsbestimmung erforderlich. Ebenso muss die Bewegungsrichtung genau erfasst werden. Die Schwierigkeit besteht nun darin, beide Informationen zuverlässig und frühzeitig zu bestimmen. Frühzeitig bedeutet in diesem Zusammenhang, dass nach der Datenverarbeitung noch eine ausreichende Reaktionszeit für die Aktorik zur Verfügung stehen muss. Des Weiteren muss die Gefahrenanalyse ausgehend vom Bewegungsmuster der Hand auf eine Gefahr schließen können, d. h. Arbeiten in unmittelbarer Nähe zum Sägeblatt, die z. B. bei Feinarbeiten notwendig sind, dürfen nicht

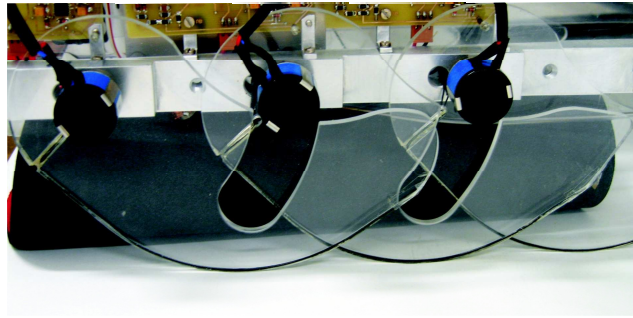


Abbildung 5.3: Schutzlamellen für eine Schutzhaube

als Gefahr interpretiert werden, da dies die Zuverlässigkeit des Systems reduzieren würde.

Diese Anforderungen ist durch ein Bildverarbeitungssystem erfüllbar, da die drei genannten Merkmale zur Gefahrenanalyse nur damit gewonnen werden können. Ein zuverlässiges Sicherheitssystem kann nicht auf einem Einzelsensorsystem basieren, das zusätzlich noch eine komplexe Datenverarbeitung erfordert, da Fehldetektionen nicht durch ein anderes System abgefangen werden können und somit die Zuverlässigkeit nicht gewährleistet ist.

Im ungünstigsten Fall bleiben durchschnittlich weniger als 3 cm zwischen Sägeblatt und Hand und diese Strecke ist noch teilweise durch die Schutzhaube verdeckt. Das Sicherheitssystem hat bei einer maximalen Handgeschwindigkeit von 2 m/s noch 15 ms als Reaktionszeit. Die zerstörungsfrei arbeitende Aktorik (Abschnitt 5.3.4) benötigt aber durchschnittlich 40 ms zum Stillsetzen des Sägeblatts ohne die Verarbeitungszeit der Daten. Unter diesen Umständen ist, wie oben erwähnt, eine zuverlässige Gefahrenerkennung und volle Wirkungsentfaltung der Aktorik nicht möglich.

Aus diesem Grund wurden vom Institut für Werkzeugmaschinen der Universität Stuttgart die in Abbildung 5.3 dargestellten Schutzlamellen für Formatkreissägen entwickelt. Jeweils drei dieser Lamellen werden an den Seiten der Schutzhaube angebracht, s. Abbildung 5.2. Dadurch werden die gefährlichen Handbewegungen zum Sägeblatt gestoppt bzw. verhindert. Die Schutzlamellen haben außerdem den Vorteil, dass diese bündig auf den Werkstücken aufliegt und deren Vorschubbewegung ohne Widerstand folgen. Da die Lamellen sich direkt an der Schutzhaube befinden, sind auch Feinarbeiten nahe am Sägeblatt möglich.

Zusätzlich besteht die Möglichkeit weitere Sensorik ins Überwachungssystem einzubauen. So werden die Schutzlamellen derart modifiziert, dass ein kapazitiver Sensor darin integriert wird, s. Abschnitt 2.3.3 und Abbildung 2.12. Dieser Aufbau ermöglicht eine genaue Überwachung des *Gefährdungsbereiches* indem das Einführen einer Hand registriert werden kann.

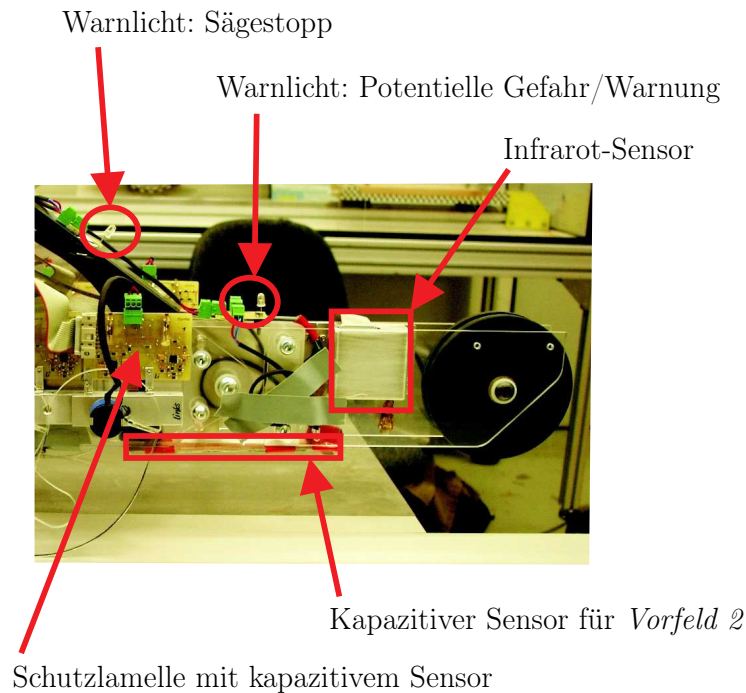
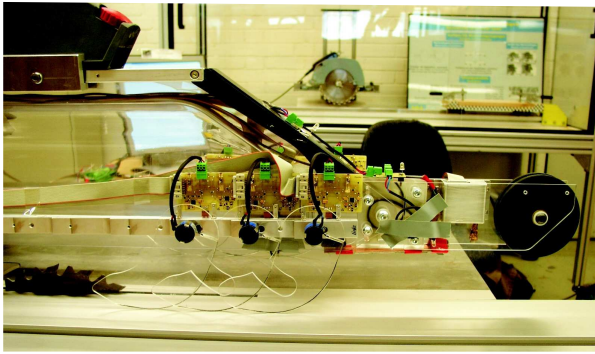


Abbildung 5.4: Modifizierte Schutzhaube mit Sensoren

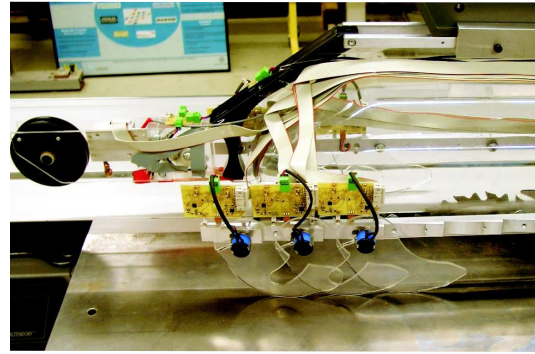
5.3.2 Sensorsysteme

Im Folgenden werden die Einsatzorte der einzelnen Sensorsysteme, die in dieser Arbeit vorgestellt werden, aufgezeigt. In Abbildung 5.4 ist ein Teil einer modifizierten Schutzhaube dargestellt, wie sie im Rahmen dieser Arbeit eingesetzt wird. Die Aktorik *Warnlicht* ist an dem oberen, vorderen Teil der Schutzhaube angebracht und dient zur optischen Signalisierung einer Warnung für den Maschinenbenutzer und des Sägestopps. Des Weiteren sind die für diesen Teil verwendeten Sensoren dargestellt. Dieser Aufbau ist als Versuchsstand am Institut für Industrielle Informationstechnik der Universität Karlsruhe (TH) aufgebaut. Im Vergleich zu einer Standardhaube ist die modifizierte Schutzhaube um insgesamt 20 cm verlängert. Dadurch wird zum einen erreicht, dass das zu überwachende *Vorfeld 2* (s. Abbildung 5.1) durch eine komplementäre Sensorik überwacht wird. Zum anderen ist dem Maschinenbenutzer allein durch die Bauform der Schutzhaube visualisiert, welches die für seine Hand kritischen Bereiche der Säge sind. Des Weiteren bietet sich durch diese Verlängerung eine weitere Möglichkeit an, zusätzliche Sensoren für die Überwachung zu integrieren. Durch die Verlängerung der Schutzhaube ist somit eine frühzeitige Erkennung einer Hand und Warnung des Benutzers gegeben.

In Abschnitt 2.3.2 ist ein kapazitiver Sensor zur Hauptfeldmessung eingeführt worden. Dieser überwacht zusammen mit einem Infrarot-Sensor (s. Abschnitt 2.2) das *Vorfeld 2*. Da die Position der Sensoren bekannt ist, können damit die Merkmale *Objektposition* und *Handidentifikation* zur Gefahrenanalyse gewonnen werden. Eine Bestimmung der Geschwindigkeit findet hier nicht statt, da dies durch



(a) links



(b) rechts

Abbildung 5.5: Ansichten der modifizierten Schutzhaube

den physikalischen Aufbau der Sensoren aufgrund von Streuung unpräzise ist. In Abbildung 5.4 ist der Einsatzort zu sehen.

Die Schutzlamellen dienen zur Abschirmung von lateralen Bewegungen und mit den integrierten kapazitiven Sensoren zur Überwachung der *Gefährdungszone*, s. Abschnitt 2.3.3. Zusätzlich ist zwischen der ersten und zweiten Lamelle ein Infrarot-Sensor angebracht, s. Abschnitt 2.2. Der Aufbau mit drei Lamellen, der kapazitiven und Infrarot-Sensoren ist auf der rechten und der linken Seite der Schutzhaube zu finden. Auf jeder Seite werden die Merkmale *Handidentifikation* und *Position* bestimmt. Die Position der Hand wird durch das sequentielle Berühren der einzelnen Lamellen erfasst. Die Bestimmung der Geschwindigkeit ist in der *Gefährdungszone* irrelevant, da das Eindringen einer Hand in diese Zone per Definition ein hohes Gefährdungspotential hat und unverzüglich nach der Erfüllung bestimmter Kriterien (s. Abschnitt 4.4) darauf reagiert werden muss. Die eingesetzte Schutzhaube ist in Abbildung 5.5 zu sehen.

Für die Überwachung des *Vorfeldes 1* vorgesehene Kamera und Bildverarbeitung (s. Kapitel 3) ist ein Testaufbau vorhanden, der im Versuchsstand integriert ist. Für einen Einsatz an einer Formatkreissäge ist ein zusätzlich angebrachter Schwenkarm erforderlich. Die Bildverarbeitung ist in der Lage sowohl die Objektposition, dessen Geschwindigkeit und Bewegungsrichtung zu bestimmen wie auch eine Handidentifikation durchzuführen. Eine vollständige Merkmalsbestimmung zur Gefahrenanalyse ist daher unter der Voraussetzung einer ausreichenden Beleuchtung (s. Abschnitt 3.2) und Datenverarbeitungskapazität möglich.

Weitere verwendete Sensoren sind ein Ultraschall-Sensor zur Bestimmung der Schutzhaubenhöhe sowie ein Temperatursensor zur Messung der aktuellen Umgebungstemperatur. Hiermit kann der Gültigkeitsbereich des Infrarot-Sensors überwacht werden.

5.3.3 Datenverarbeitung und Gefahrenanalyse

Alle Sensorinformationen, außer der kapazitiven Hauptfeldmessung (s. Abschnitt 2.3.2), stehen als analoge Signale im Bereich 0-10 V zur Verfügung und müssen im ersten Schritt für eine weitere Datenverarbeitung digitalisiert werden. Im nächsten Schritt erfolgen die Signalvorverarbeitung und Korrekturrechnungen für die einzelnen Sensorinformationen wie es in Abschnitt 4.3 beschrieben ist. Die Ausgangssignale dieser Datenverarbeitung stellen Wahrscheinlichkeiten für das Erkennen einer Hand im Erfassungsbereich der einzelnen Sensoren dar. Mithilfe des in Abschnitt 4.4 entwickelten Bayes'schen Netzes werden die Einzelinformationen zusammengefügt und die berechneten Endwahrscheinlichkeiten analysiert. In diesem Verarbeitungsschritt erfolgen also die Informationsfusion und die Gefahrenanalyse auf einmal. Anhand des Ergebnisses der Gefahrenanalyse erfolgt entweder die Warnung des Benutzers, dass dessen Hand gefährdet ist, oder die Gefahrenquelle wird neutralisiert, d. h. die Bremsaktorik wird ausgelöst und damit das Sägeblatt gestoppt.

Versuchsstand

Der Versuchsstand einer Formatkreissäge mit modifizierter Schutzhaube und Aktorik ist in Abbildung 5.6 zu sehen. Dieser ist am Institut für Werkzeugmaschinen der Universität Stuttgart aufgebaut, da dort die Aktorik gebaut und an der Säge getestet wird. Der Versuchsstand am Institut für Industrielle Informationstechnik besteht aus der modifizierten Schutzhaube mit einer dazugehörenden Halterung und einer Nachbildung des Arbeitstisches einer Formatkreissäge. Damit sind die wichtigsten Teile der Säge für die Validierung und Test des Sicherheitssystems gegeben.

Zur Digitalisierung der Sensordaten wird am Versuchsstand eine PC-basierte Wandlerkarte (National Instruments M-6254) verwendet. Anschließend wird die Vorverarbeitung, Informationsfusion und Gefahrenanalyse in einem PC mit den Programmen NI LabView und Visual C/C++ durchgeführt. Durch diesen Aufbau können die entwickelten Verfahren reproduzierbar validiert werden, s. Abschnitt 5.4.

Die Validierung der Bildverarbeitung und deren Ergebnisse sind in Abschnitt 3.9 zu finden. Der Testaufbau besteht aus Kamera (Basler A602fc) und einem PC und ist in den Versuchsstand integriert. Der vorgestellte Ablauf der Bildverarbeitung benötigt ca. 170 ms für die Verarbeitung eines Bildes. Dies ist für die Validierung des Verfahrens ausreichend, jedoch für die Integration ins Sicherheitssystem zu langsam. Aus diesem Grund arbeitet dieses System unabhängig von dem oben genannten. Um jedoch den Einfluss der Bildverarbeitung auf die Informationsfusion zu untersuchen, sind externe Eingabemöglichkeiten vorhanden. Damit können die Ergebnisse der Bildverarbeitung an die Informationsfusion

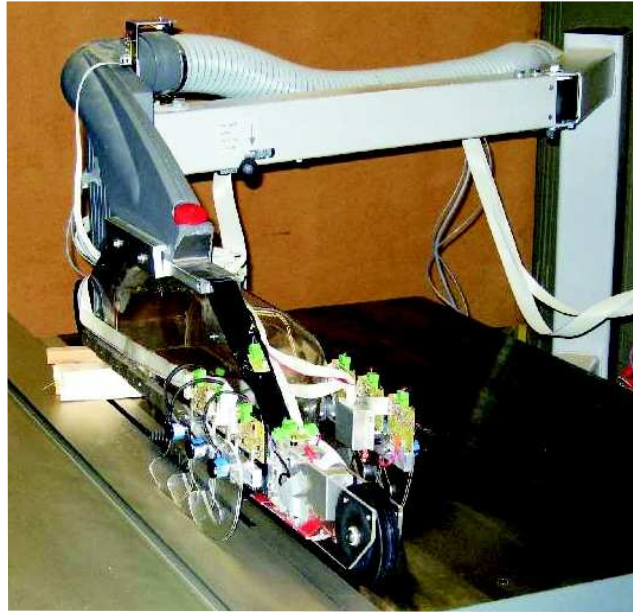


Abbildung 5.6: Versuchsstand

weitergegeben werden. Durch den technologischen Fortschritt der Bildverarbeitungssysteme ist die Integration in zukünftige Sicherheitssysteme aufgrund der gesteigerten Leistungsfähigkeit möglich [90]. Ebenso sinken die Preise für die Bildverarbeitungssysteme, was käuferseitig die Akzeptanz erhöht.

5.3.4 Aktorik

Damit ein Sicherheitssystem als *aktiv* bezeichnet werden kann, ist eine Aktorik notwendig, welche die Gefahrenquelle neutralisiert. In dem hier vorgestellten System wird eine Bremse zur Stillsetzung des Sägeblattes als Aktorik verwendet. Diese arbeitet zerstörungsfrei, d. h. nach einer erfolgten Auslösung der Bremse kann das Sägeblatt wieder verwendet werden. Entwickelt und getestet wurde das System am Institut für Werkzeugmaschinen, Universität Stuttgart [77]. Das System benötigt durchschnittlich 40 ms zum Abbremsen eines Sägeblattes. Das Sicherheitssystem ist so ausgelegt, dass der Maschinenbenutzer gewarnt wird, wenn dessen Hand an einer potentiell gefährlichen Stelle des Maschinentisches ist und er die Hand wieder aus dem kritischen Bereich heraus nehmen kann. Damit kann unnötiges Auslösen der Bremse vermieden werden. Die Warnung wird optisch durch LEDs angezeigt.

5.3.5 Modularer Aufbau des Systems

In Abbildung 1.7 ist der modulare Aufbau des entwickelten Systems zur Gefahrenerkennung zu sehen. Die Basisversion des System besteht aus der Überwachung der *Gefährdungszone*, der Informationsfusion und der Aktorik. Der *Über-*

wachungsbereich wird hierbei nicht berücksichtigt. Dieses minimale System verwendet drei Schutzlamellen mit integrierter kapazitiver Sensorik und Infrarot-Sensoren auf der rechten und linken Schutzhaubenseite. Diese Version ist kostengünstig, jedoch sind Fehlauflösungen z. B. aufgrund von Verbundwerkstoffen nicht auszuschließen. Die Zuverlässigkeit des Systems ist in diesem Fall gering. Um die Zuverlässigkeit des Systems zu steigern, ist es notwendig das *Vorfeld 1* und *Vorfeld 2* zu überwachen. Beides kann unabhängig voneinander in das System integriert werden, da die Informationsfusion mit einem Bayes'schen Netz ebenfalls modular aufgebaut ist. Einzelne Module können ohne Schwierigkeit hinzugefügt werden, s. Abschnitt 4.4. Die maximale Ausbaustufe des Systems ist erreicht, wenn beide Bereiche ins Sicherheitssystem integriert sind.

Das *Vorfeld 1* wird durch eine Kamera beobachtet. Das *Vorfeld 2* wird durch einen kapazitiven und einen Infrarot-Sensor überwacht. In diesem Bereich befindet sich ebenfalls der induktive Sensor zur Materialerkennung, s. Abschnitt 2.3.4. Die Schutzhaube ist ein zentraler Bestandteil der Überwachungsmöglichkeit an einer Formatkreissäge. Für Arbeitsgänge, bei denen diese demontiert werden muss, wie dem Einsetzschneiden, ist als einziges Sensorsystem eine Kamera einsetzbar, da diese eine separate Befestigung an der Maschine besitzt. In diesem Fall muss der zu beobachtende Bereich geändert werden, damit die Gefahrensituation besser eingeschätzt werden kann. Jedoch ist die Zuverlässigkeit dieses Einzelsensorsystems im Vergleich zum Multisensorsystem geringer. Ebenso ist ein leistungsstarke Bildverarbeitungseinheit notwendig.

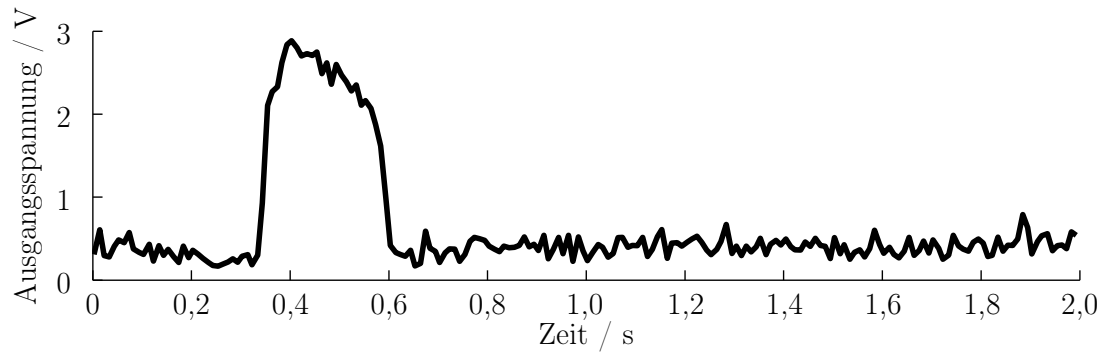
5.4 Validierung und Ergebnisse

Die Zuverlässigkeit des entwickelten Systems kann durch eine Validierung und Funktionstests nachgewiesen werden. Die Ergebnisse davon sind in diesem Abschnitt dargestellt. Zu Beginn wird am Beispiel einer einfachen Handdetektion der Ablauf der Informationsfusion und Gefahrenanalyse verdeutlicht. Anschließend wird die Reaktionsfähigkeit des Systems bei unterschiedlichen Handgeschwindigkeiten untersucht. Nachfolgend wird das Verhalten des Gesamtsystems bei unterschiedlichen Testfällen aufgezeigt.

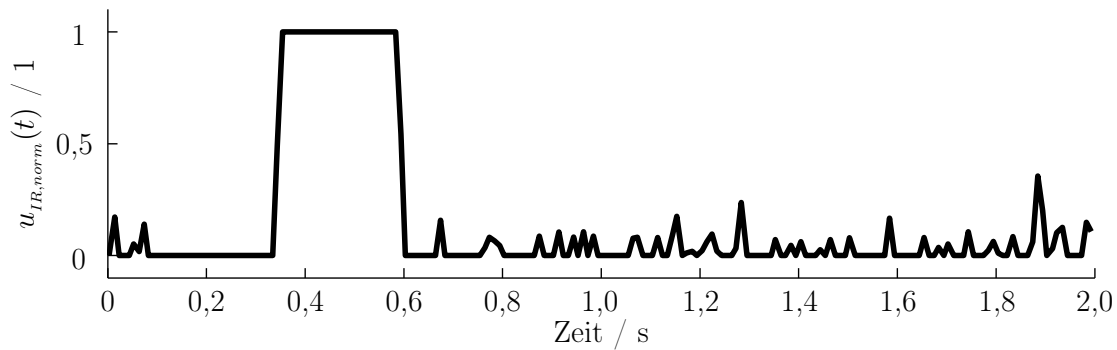
Die Validierung der Bildverarbeitung erfolgt in Abschnitt 3.10 und ist nicht Gegenstand der in diesem Abschnitt vorgestellten Untersuchung.

5.4.1 Beispiel: Detektion einer Hand

Bei diesem Beispiel wird der Ablauf der Gefahrenanalyse verdeutlicht. Dazu wird eine Hand langsam (mit 0,32 m/s, gemessen mit separaten Lichtschranken) durch das *Vorfeld 2* und unter den Schutzlamellen durch die *Gefährdungszone* bewegt. Die Höhe der Schutzhaube beträgt 6,21 cm.



(a) vor der Vorverarbeitung



(b) nach der Vorverarbeitung

Abbildung 5.7: Werte des Infrarot-Sensors im *Vorfeld 2*

In Abbildung 5.7 (a) ist der Werteverlauf des Infrarot-Sensors im *Vorfeld 2* dargestellt. Nach der Durchführung der Vorverarbeitung mithilfe der Gleichung (4.4) erhält man den in Abbildung 5.7 (b) dargestellten Verlauf.

Im nächsten Schritt wird der vorverarbeitete Sensorwert mithilfe der Wahrscheinlichkeitsfunktionen (Gleichungen (4.9) bis (4.11)) diskretisiert. Die daraus resultierenden Wahrscheinlichkeiten für die drei Zustände des Infrarot-Sensors *nichts*, *kalt* und *warm* sind in Abbildung 5.8 zu sehen.

Da bei diesem Test keine Kameraüberwachung eingesetzt wird, besitzt der Knoten *Kamera* seine Initialisierungswahrscheinlichkeit und damit auch der Hypothesenknoten *Vorfeld 1* die Wahrscheinlichkeitsverteilung

$$P(V1) = P(\text{keine hand}, \text{hand}) = (1 ; 0) .$$

Das Netz ist so entworfen, dass die Wahrscheinlichkeiten der Zustände des Knoten *IR Vorfeld*, wenn keine Informationen aus der Bildverarbeitung vorliegen, für die Zustände des Knoten *Vorfeld 2 vorne* übernommen werden. Aus diesem Grunde sind die in Abbildung 5.8 gezeigten Wahrscheinlichkeiten auch gleichzeitig die Wahrscheinlichkeiten der Zustände *nichts*, *kalt* und *warm* des Knoten *Vorfeld 2 vorne*. Der vierte Zustand dieses Knoten *hand* ist 0.

Die nächste Information, die mit in die Fusion einfließt, ist die des kapazitiven Sensors mit dem in Abbildung 5.9 dargestellten Verlauf.

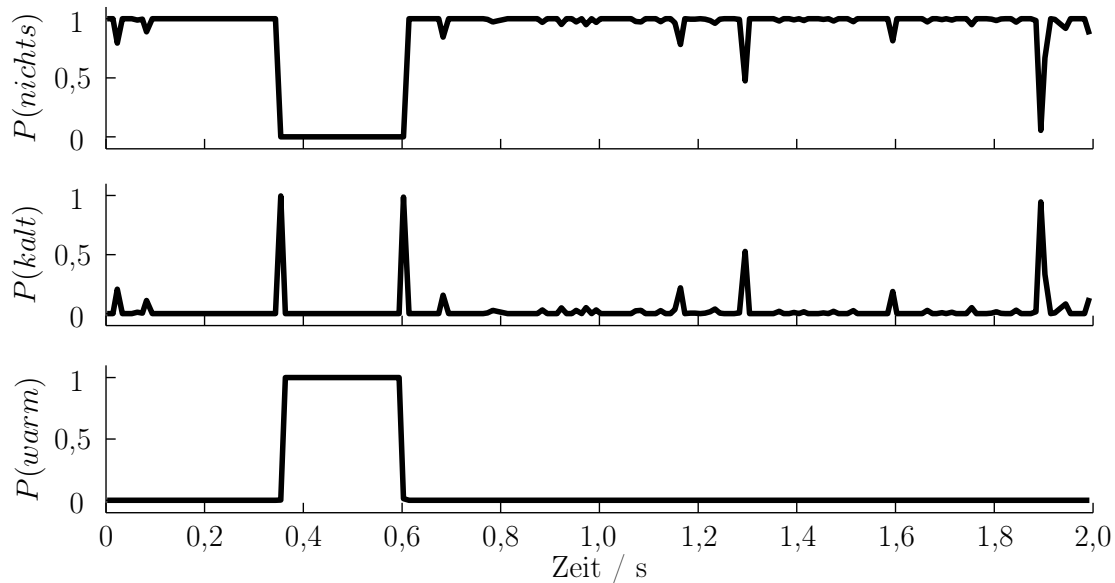


Abbildung 5.8: Wahrscheinlichkeiten der drei Zustände des Infrarot-Sensors

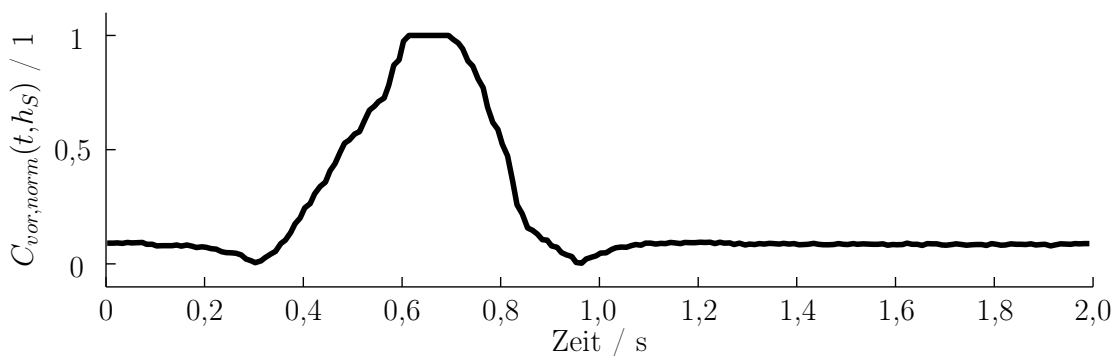


Abbildung 5.9: Wert des kapazitiven Sensors im Vorfeld 2 nach der Vorverarbeitung

Für die Materialerkennung wird die Initialisierungswahrscheinlichkeit für den Knoten *Indukt. Sensor* verwendet, da korrekterweise kein Metall erkannt wird.

Die Fusion der Informationen aus dem kapazitiven und dem Infrarot-Sensor findet im Knoten *Vorfeld 2 hinten* statt. Dieser Knoten besitzt die Zustände *nichts*, *holz*, *metall* und *hand*. Aufgrund der Lage des kapazitiven Sensors erreicht dieser sein Wertemaximum von 1 erst nach ca. 0,6 s. Da die Hand den Überwachungsbereich des Infrarot-Sensors zu diesem Zeitpunkt bereits wieder verlassen hat (s. Abbildung 5.7), ist die beschriebene Speicherung des Maximums zur Informationsfusion notwendig.

Um dies zu verdeutlichen ist in Abbildung 5.10 der Verlauf der einzelnen Zustandswahrscheinlichkeiten des Knoten *Vorfeld 2 hinten* ohne Zeitverzögerung gestrichelt, mit einer Zeitverzögerung für den Knoten *Vorfeld 2 vorne* durchgezogen und mit einer Zeitverzögerung für beide Knoten des *Vorfeld 2* gepunktet dargestellt.

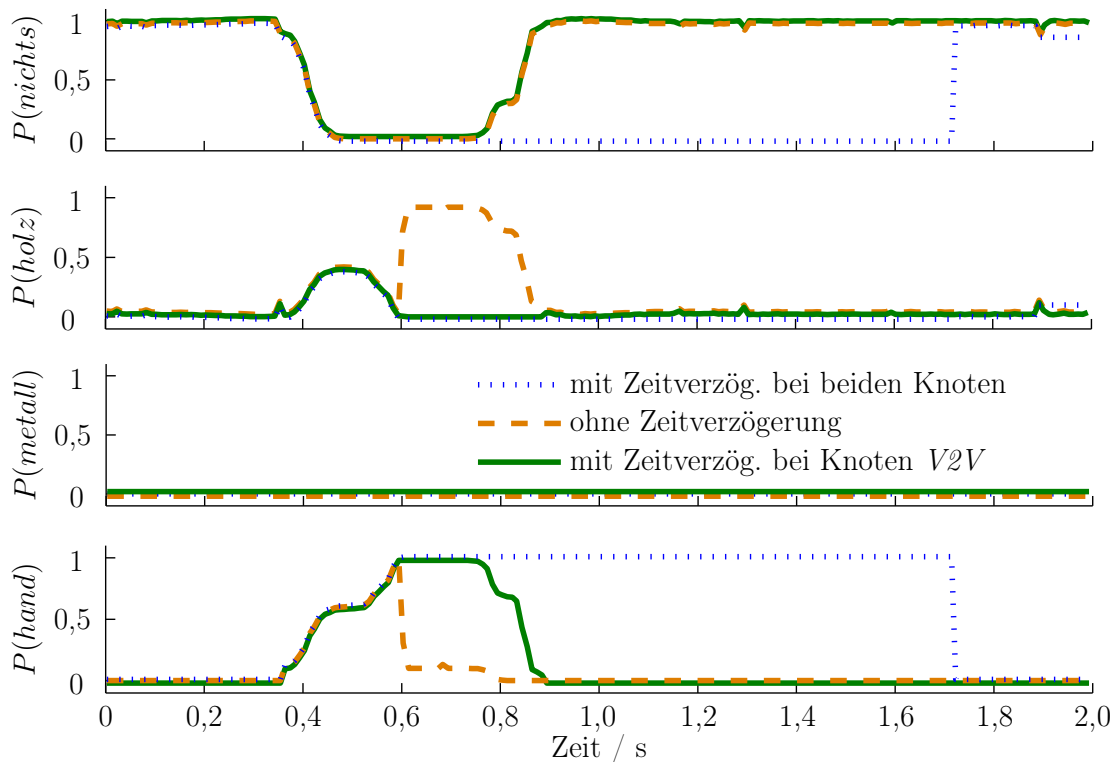


Abbildung 5.10: Wahrscheinlichkeiten der vier Zustände des Knoten *Vorfeld 2 hinten*

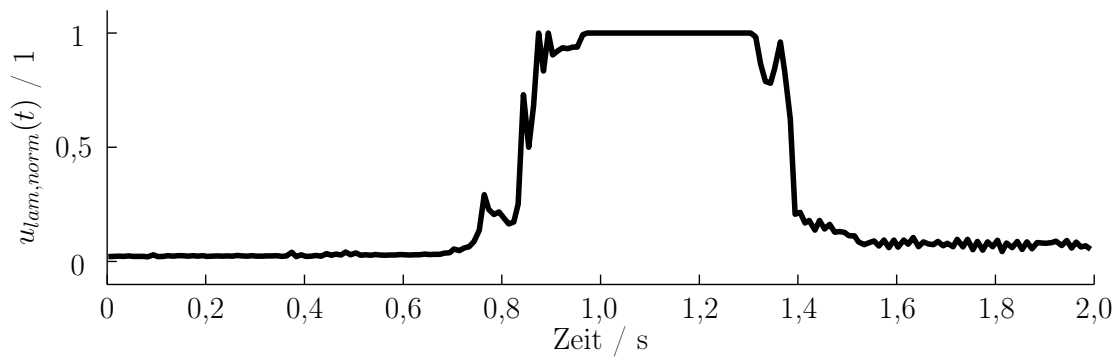
Wie beim Verlauf der Wahrscheinlichkeiten ohne Zeitverzögerung (gestrichelt) zu sehen ist, würde aufgrund des Verlustes der Information des Infrarot-Sensors in der Zeitspanne von 0,6 s bis 0,9 s auf den Zustand *holz* geschlossen, anstatt auf den richtigen Zustand *hand*. Der kapazitive Sensor liefert zwar für diese Zeitspanne seinen maximalen Ausschlag (s. Abbildung 5.9), ohne zusätzliche Information über eine Temperaturänderung wird jedoch zur Vermeidung eines Fehlalarms nicht auf eine Hand geschlossen.

Die durchgezogene Kurve stellt die beste Abschätzung für den Zustand *hand* des Knotens *Vorfeld 2 hinten* dar. Da aber die nächsten Sensoren wiederum die Hand erst detektieren, wenn sie den Bereich des kapazitiven Sensors des Vorfeldes verlassen hat, muss auch hier eine Zeitverzögerung eingefügt werden. Die durch Halten des Maximums für den Zustand *hand* erzielten Wahrscheinlichkeiten sind in gepunktet dargestellt und werden an die folgenden Hypothesenknoten weitergeleitet.

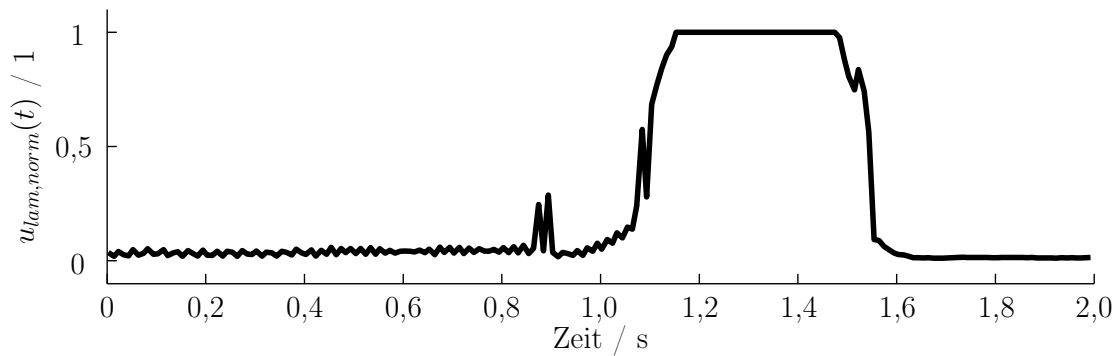
Da in diesem Beispiel die Hand auf der linken Seite der Säge in die *Gefährdungszone* eindringt, wird hier auf eine Betrachtung der Sensorwerte der Lamellen auf der rechten Seite verzichtet.

Nun werden die vorverarbeiteten Werte der kapazitiven Sensoren in den Lamellen auf der linken Seite aus Abbildungen 5.11 (a), (b) und (c) mit in die Fusion eingebracht.

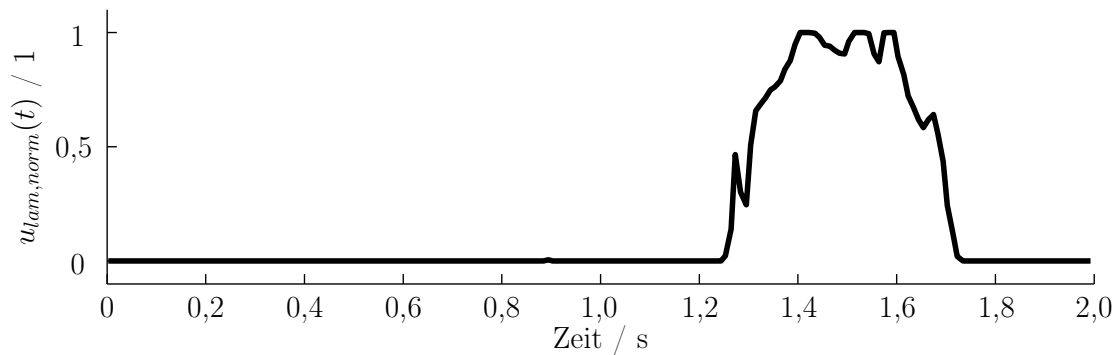
Für die Berechnung der Wahrscheinlichkeiten des Knoten *Lamelle 1 links* wird außer den Informationen aus dem vorhergehenden Hypothesenknoten *Vorfeld 2*



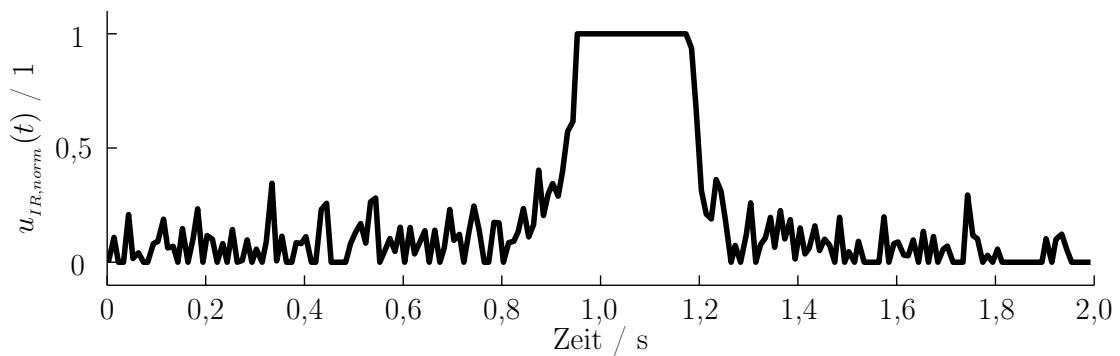
(a) Lamelle 1 links



(b) Lamelle 2 links



(c) Lamelle 3 links



(d) Infrarot-Sensor links

Abbildung 5.11: Vorverarbeitete Sensorwerte für die *Gefährdungszone*

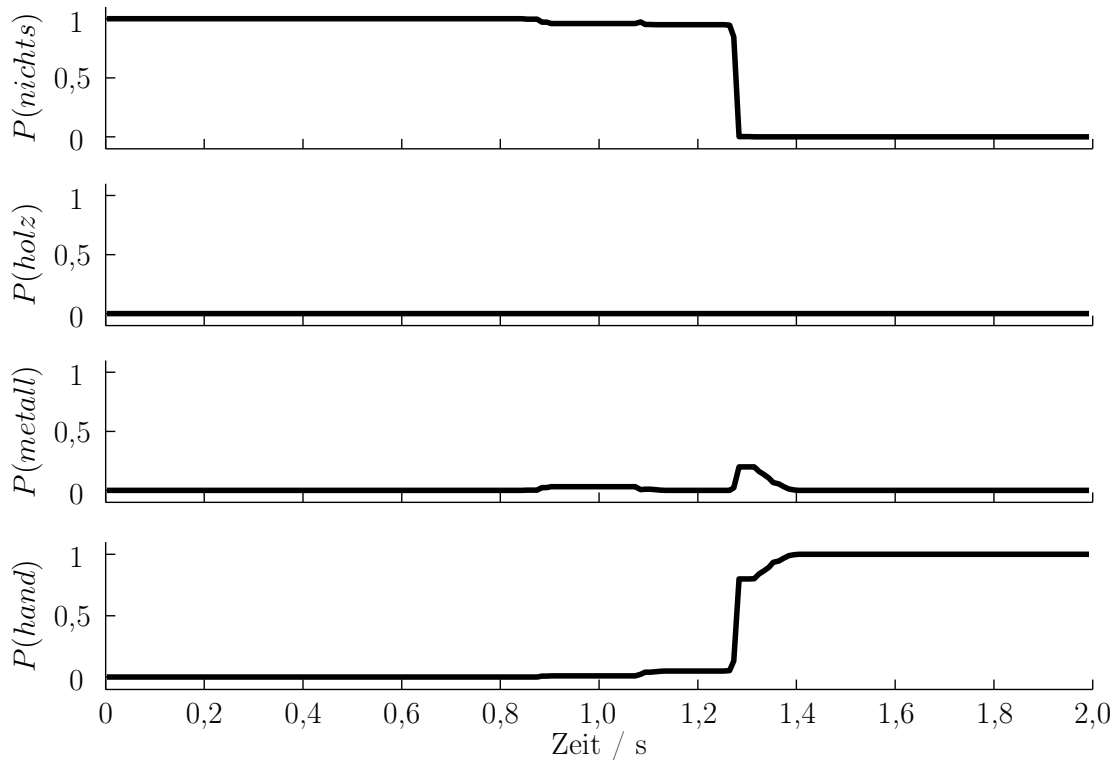


Abbildung 5.12: Wahrscheinlichkeiten der vier Zustände des Knoten *Lamelle 3 links*

hinten und den Wahrscheinlichkeiten des Informationsknoten *Kap 1 links* noch die Sensorwerte des Infrarot-Sensors, der parallel zur ersten Lamelle angebracht ist, benötigt. Der Verlauf dieses Sensorwerts nach der Vorverarbeitung ist in Abbildungen 5.11 (d) dargestellt.

Die Zustandswahrscheinlichkeiten des Hypothesenknoten *Lamelle 1 links* ergeben sich aus der Fusion der Wahrscheinlichkeiten des kapazitiven Sensors der ersten Lamelle, des linken Infrarot-Sensors und den Zustandswahrscheinlichkeiten des Hypothesenknotens *Vorfeld 2 hinten*. Die erhaltene Wahrscheinlichkeitsverteilung wird wiederum mit den Informationen aus dem kapazitiven Sensor der Lamelle 2 fusioniert.

Im letzten Fusionschritt werden dann die Informationen aus dem kapazitiven Sensor der Lamelle 3 ebenfalls mitverwendet. Dabei ergeben sich die Wahrscheinlichkeiten für die vier Zustände *nichts*, *holz*, *metall* und *hand* des letzten Hypothesenknoten *Lamelle 3 links*. Diese Verteilung ist in Abbildung 5.12 zu sehen. Nachdem für einen Zeitpunkt die Wahrscheinlichkeitsverteilungen aller Variablen in den verschiedenen Knoten berechnet wurden, wird noch die Überprüfung durchgeführt, ob eine Gefährdung der Hand vorliegt. Dies geschieht mithilfe der beiden Gleichungen (4.12) und (4.13). Mit deren Hilfe wird untersucht, ob eine Warnung an den Benutzer ausgegeben wird, oder gar die Säge abgestoppt werden muss.

Aufgrund der Überschreitung der Wahrscheinlichkeit von 0,5 für den Zustand *hand* im Knoten *Vorfeld 2 vorne* (Abbildung 5.10), wird bereits nach 0,434 s

eine Warnung an den Benutzer ausgegeben. Da sich die Hand jedoch weiter in Richtung Sägeblatt bewegt und die Wahrscheinlichkeit für den Zustand *hand* im letzten Hypothesenknoten zum Zeitpunkt 1,284 s den Wert von 0,4 überschreitet, wird zu diesem Zeitpunkt das Stoppen der Säge eingeleitet.

Für die Abbremsung des Sägeblatts werden durchschnittlich 40 ms benötigt. Die Hand bewegt sich in diesem Beispiel mit einer durchschnittlichen Geschwindigkeit von ca. 0,32 m/s. Die verbleibende Zeit bis das Sägeblatt abgebremst sein muss, variiert je nach dem Zeitpunkt zu dem die Wahrscheinlichkeit des Zustandes *hand* im letzten Hypothesenknoten den Wert von 0,4 überschreitet. In diesem Beispiel stehen zur Abbremsung 0,68 s zur Verfügung. Eine Verletzung der Hand wird also verhindert.

5.4.2 Handerkennung bei unterschiedlichen Bewegungsgeschwindigkeiten

Da an einer Formatkreissägen Bewegungsgeschwindigkeiten bis zu 2 m/s auftreten können, ist auch eine Untersuchung der maximalen Geschwindigkeit notwendig, bei der die Säge noch rechtzeitig abgebremst werden kann. Deshalb werden die Informationsfusion und die verwendeten Sensoren auf ihre Reaktionsgeschwindigkeit getestet. Dazu wurden an der Formatkreissäge zwei LED-Lichtschranken befestigt, mittels denen die Handgeschwindigkeit gemessen werden kann. Die Ergebnisse für unterschiedliche Geschwindigkeiten einer Hand sind in Tabelle 5.1 wiedergegeben.

Für die in der Tabelle aufgelisteten Versuchsergebnisse wurde eine Hand durch das *Vorfeld 2* in die Gefährdungszone auf das Sägeblatt zu bewegt. Die Informationen aus der Kameraüberwachung und der Materialerkennung wurden nicht mitverwendet. Die Abbremsung des Sägeblattes benötigt ca. 40 ms. Dies bedeutet, dass nach Aktivierung des Knoten *Stopp* noch mindestens 40 ms vergehen müssen, bis die Hand das Sägeblatt erreicht, um eine Verletzung zu verhindern. Die maximale Vorschubgeschwindigkeit bei Formatkreissäge liegt bei 0,67 m/s. Wie in der Tabelle 5.1 ersichtlich ist, reicht dies zum Abstoppen der Säge aus, um eine Verletzung zu vermeiden. Bei Geschwindigkeiten über 1 m/s hingegen, kann das Sägeblatt nicht mehr rechtzeitig abgebremst werden. Ab einer Geschwindigkeit von über 1,3 m/s ist beim Erreichen des Sägeblattes die Bremse noch nicht aktiviert. Jedoch sind diese hohen Geschwindigkeiten in diesem Bereich nur durch absichtliche Handbewegungen und nicht durch Reflex- oder Wischbewegungen zu erreichen.

Durch die Vergrößerung des Abstandes zwischen der letzten Lamelle und dem Sägeblatt, kann eine längere Zeit bis zum Stopp des Sägeblattes erreicht bzw. es kann auf höhere Geschwindigkeiten rechtzeitig reagiert werden.

Geschwin- digkeit in m/s	Hauben- höhe in cm	Zeit in ms	Zustandswahrscheinlichkeiten			
			<i>nichts</i>	<i>holz</i>	<i>metall</i>	<i>hand</i>
0,31	6,51	684,11	0	0	0	1
0,33	6,51	774,01	0	0	0,0090	0,9910
0,58	6,51	288,58	0	0	0,0040	0,9960
0,77	9,55	143,42	0,0004	0,0406	0,3935	0,5655
0,87	9,66	134,31	0,0038	0,0512	0,3514	0,5936
1,03	7,89	46,17	0	0	0,2066	0,7934
1,07	6,51	42,42	0	0	0,0016	0,9984
1,26	7,90	27,89	0	0	0,1072	0,8928
1,35	7,92	-67,50	0	0,0015	0,2115	0,7870
1,71	7,90	-85,71	0	0	0,0001	0,9999
1,88	7,88	-94,07	0	0,0003	0,2316	0,7681
1,92	7,89	-96,08	0	0	0	1

Tabelle 5.1: Ergebnisse der Informationsfusion bei unterschiedlichen Geschwindigkeiten.
Zeit: Zeitdauer von der Aktivierung des Knotens *Stopp* bis die Hand das Sägeblatt erreicht

In der Tabelle 5.1 ist außerdem zu sehen, dass die Grenze von 0,4 zur Aktivierung des Knoten *Stopp* bei allen Versuchen überschritten wurde und so eine Hand zuverlässig detektiert wird.

5.4.3 Verifikation der Gefahrenerkennung

Zur Überprüfung der Zuverlässigkeit der Informationsfusion mittels des entwickelten Bayes'schen Netzes wurden eine Vielzahl von Tests (ca. 3000) durchgeführt. Dabei wurde unter anderem festgestellt, dass die Handgeschwindigkeit Einfluss auf die Erkennung der Hand hat. Aus diesem Grund wird die Verifikation der Informationsfusion lediglich bei Handgeschwindigkeiten bis 0,67 m/s durchgeführt.

Für die Verifikation eines Systems hinsichtlich der Zuverlässigkeit sind vier Fälle zu unterscheiden [9]. Zum einen der Fall bei dem eine Hand richtig detektiert wird (*richtig positiv*) und der, in dem eine Hand nicht detektiert wird, obwohl eine Hand in Gefahr ist (*falsch negativ*). Außerdem muss untersucht werden, ob bei der Verarbeitung von unterschiedlichen Materialien richtig auf das Nicht-Vorhandensein einer Hand geschlossen wird (*richtig negativ*) oder eine Hand fälschlicherweise detektiert wird (*falsch positiv*). Eine Übersicht über die vier Fälle ist in Tabelle 5.2 dargestellt.

	Säge gestoppt	Säge nicht gestoppt
Hand in Gefahr	richtig positiv	falsch negativ
keine Hand in Gefahr	falsch positiv	richtig negativ

Tabelle 5.2: Vier Fälle zur Untersuchung der Zuverlässigkeit

Beschreibung des Testablaufs	richtig positiv Stopp aktiviert	falsch negativ Stopp nicht aktiviert
Hand durch <i>Vorfeld 2</i> in <i>Gefährdungszone</i>	>99 %	<1 %
Hand direkt seitlich in <i>Gefährdungszone</i> ohne Kameraüberwachung im <i>Vorfeld 1</i>	>96 %	<4 %
Hand direkt seitlich in <i>Gefährdungszone</i> mit simulierter Kameraüberwachung im <i>Vorfeld 1</i>	>97 %	<3 %
Ein Finger direkt seitlich in <i>Gefährdungszone</i> ohne Kameraüberwachung im <i>Vorfeld 1</i>	>92 %	<8 %
Ein Finger direkt seitlich in <i>Gefährdungszone</i> mit simulierter Kameraüberwachung im <i>Vorfeld 1</i>	>96 %	<4 %
Hand auf Holz durch <i>Vorfeld 2</i> in <i>Gefährdungszone</i>	>98 %	<2 %
Hand auf Metall durch <i>Vorfeld 2</i> in <i>Gefährdungszone</i>	>99 %	<1 %

Tabelle 5.3: Szenarien für die Fälle *richtig positiv* und *falsch negativ* und ihre Auftrittshäufigkeit

Zur Untersuchung der Fälle *richtig positiv* und *falsch negativ* wurde eine Hand auf unterschiedliche Arten in Richtung des Sägeblattes bewegt. Die verschiedenen Fälle und ihre Auftrittshäufigkeit sind in Tabelle 5.3 aufgeführt.

Die Informationsfusion funktioniert zuverlässig, wie anhand der hohen Auftrittshäufigkeiten für den Fall *richtig positiv* zu sehen ist. Im Fall, dass Arbeitshandschuhe verwendet werden, ist eine zuverlässige Handdetektion aufgrund der Wärmeisolation und einem zu großen Abstand für eine Kapazitätsmessung nicht mehr möglich. Das Tragen eines Handschuhs ist aus Gründen der Arbeitssicherheit jedoch verboten [32].

Die Verifikationsergebnisse der Informationsfusion für alle Fälle, in denen keine Hand in Gefahr ist, sind in Tabelle 5.4 wiedergegeben. Zum einen wird in der Tabelle die Auftrittshäufigkeit des Falls *falsch positiv* aufgelistet, in dem fälschlicherweise auf eine Hand geschlossen wird. Zum anderen ist die prozentuale Häufigkeit des Falls *richtig negativ* zu sehen, also des normalen Betriebs der Formatkreissäge ohne Gefahr für den Benutzer.

Beschreibung des Testablaufs	richtig negativ	falsch positiv
	normaler Betrieb	Notstopp aktiviert
Holz durch <i>Vorfeld 2</i> in <i>Gefährdungszone</i>	>99,5 %	<0,5 %
Metall durch <i>Vorfeld 2</i> in <i>Gefährdungszone</i>	>99,5 %	<0,5 %
Verbundwerkstoff durch <i>Vorfeld 2</i> in <i>Gefährdungszone</i>	>99,5 %	<0,5 %
warmes Holz durch <i>Vorfeld 2</i> in <i>Gefährdungszone</i>	>92 %	<8 %
warmes Metall durch <i>Vorfeld 2</i> in <i>Gefährdungszone</i>	≈ 40 %	≈ 60 %
Hand direkt neben den Lamellen an der <i>Gefährdungszone</i> vorbei (mit und ohne Kameraüberwachung)	>99,5 %	<0,5 %

Tabelle 5.4: Szenarien für die Fälle *falsch positiv* und *richtig negativ* und ihre Auftrittshäufigkeit

Wie in Tabelle 5.4 deutlich zu erkennen ist, funktioniert die Gefahrenerkennung bei normalem Betrieb und ohne eine vorhandene Gefährdung für den Anwender wie erwünscht im Hintergrund und beeinträchtigt den Betrieb nicht. Ausnahmen bilden dabei das Sägen von warmen Materialien, wobei warmes Holz zu 92 % richtig erkannt wird. Ein körperwarmes Metall ist jedoch für die kapazitiven und Infrarot-Sensoren von einer Hand nicht richtig zu unterscheiden, sodass auch mithilfe der Informationsfusion keine Unterscheidung getroffen werden kann. Ohne die im *Vorfeld 2* arbeitende induktive Sensorik wäre der *falsch positive* Wert noch höher.

Keine Schwierigkeiten bereitet dem System, wenn eine Hand nahe (weniger als 0,8 cm) an den Lamellen vorbei geführt wird. Dadurch lassen sich exakte Arbeiten in der Nähe des Sägeblattes durchführen, ohne eine Gefährdung der Hand zu riskieren und ohne dass das System die Bremsaktorik aktiviert.

5.4.4 Fazit

Durch die Kombination der kapazitiven und Infrarot-Sensoren in der Informationsfusion können 95 % (*richtig positiv*) der Unfälle verhindert werden. Das System arbeitet folglich zuverlässig. In über 99 % (*richtig negativ*) der Fälle funktioniert die Gefahrenerkennung ohne Störung, d. h. ohne Falschauslösung, und ist damit robust. Eine Optimierung, beispielsweise der Schutzlamellengeometrie erhöht die Zuverlässigkeit und Robustheit des Systems.

Durch die Vorverarbeitung der Sensordaten sowie der Struktur der Informationsfusion ist die Reaktionszeit des Systems ausreichend, s. Tabelle 5.1.

Das System deckt den Fall der Verarbeitung von warmen Metall nur ungenügend ab (40 %, *richtig negativ*). Dies kann durch den Einsatz der Bildverarbeitung jedoch signifikant verbessert werden.

Um die Bildverarbeitung an die Informationsfusion zu koppeln, ist ein spezialisierter Hochleistungsrechner erforderlich. Mit diesen so genannten *Embedded Vision Systems* (z. B. MaxxVision [64]) entsteht ein Echtzeit-fähiges System. Damit kann insgesamt die Zuverlässigkeit des Systems gesteigert werden. Allerdings sind die Kosten dafür erheblich (> 10000 €). Das Gesamtsystem, bestehend aus Gefahrenerkennung, Aktorik aber ohne Bildverarbeitung, ist für ca. 1800 € herstellbar (ca. 10-25 % des Preises einer Formatkreissägemaschine je nach deren Ausstattung).

6 Zusammenfassung

Sicherheitssysteme für Werkzeugmaschinen dienen der Erhöhung der Arbeitssicherheit und der Reduktion von Gefahren für den Maschinenbenutzer. Speziell bei Sägemaschinen sind diese Systeme notwendig, da die Arbeiten direkt an der offen liegenden Gefahrenquelle (Sägeblatt) durchgeführt werden und hierdurch ein erhöhtes Verletzungsrisiko besteht. So werden heutzutage eine Reihe von passiven Sicherheitseinrichtungen genutzt, Arbeitsschutzrichtlinien angewendet und die Maschinenbenutzer speziell geschult, damit das Unfallrisiko gesenkt wird.

Mit dem Fortschritt der Computer- und Sensortechnologie und den damit verbundenen sinkenden Kosten ist ein Einsatz elektronischer Überwachungstechnik in Kombination mit Aktoren für Sicherheitssysteme möglich. Diese Art von Systeme werden *aktive Sicherheitssysteme* genannt. In dieser Arbeit wurde ein neuartiges Konzept für die Gefahrenanalyse als Teil eines aktiven Sicherheitssystems entwickelt, das für den Einsatz an Formatkreissägen angepasst wurde und sich prinzipiell auf sämtliche Werkzeugmaschinen portieren lässt.

Die Gefahrenanalyse besteht aus zwei Aufgaben. Die erste ist das Erfassen einer Situation mit Sensoren. Die zweite Aufgabe ist die Verarbeitung und Bewertung der von den Sensoren erzeugten Information.

Sensorsysteme zur Handerkennung

Die Überwachung der Gefahrenbereiche durch die Sensoren erfolgt berührungslos, d. h. um eine Hand zu erkennen ist kein Kontakt zwischen Sensor und Hand notwendig. Dies ist erforderlich, damit ein Maschinenbenutzer in seinem gewohnten Arbeitsablauf nicht gestört ist und so die Akzeptanz für das Gesamtsystem steigt. Außerdem ist dadurch eine frühzeitige Erkennung einer Hand überhaupt erst möglich. Um diese Anforderungen zu erfüllen, sind optische und feldbasierte Sensoren geeignet.

Als besonders zweckmäßig haben sich thermopile Infrarot-Sensoren und kapazitive Sensoren herausgestellt. In Kapitel 2 werden diese im Detail vorgestellt. Die thermopilen Infrarot-Sensoren beruhen auf dem Seebeck-Effekt und zeichnen sich durch ihre im Vergleich zu pyroelektrischen Infrarot-Sensoren geringe Reaktionszeit aus. Die Infrarot-Sensoren eignen sich für die Erfassung der abgestrahlten Handwärme, als eine charakteristische Eigenschaft der Hand.

Kapazitive Sensoren reagieren auf die elektrische Feldänderung, die durch in den Detektionsbereich eingebrachte Materialien hervorgerufen wird. Menschliches Gewebe besteht hauptsächlich aus einem Elektrolyt, der die Impedanz des

Sensors ändert. Dabei sind Hauptfeld- und Streufeldmessung zu unterscheiden. Bei einer Hauptfeldmessung wird das zu untersuchende Material zwischen zwei Elektroden gebracht und die Impedanzänderung erfasst. Dieses Prinzip wird zur Vorfeldüberwachung eingesetzt.

Bei der Streufeldmessung sind die Elektroden parallel und in einer Ebene angeordnet. Hier kann nur das entstehende Streufeld um die Elektroden herum ausgewertet werden. Die dafür entwickelte Sensorik hat eine geringere Reichweite als die Hauptfeldmessung. Dies und die Art des Sensoraufbaus ermöglichen es, diesen Sensor in die Schutzlamellen zu integrieren und die Gefährdungszone einer Säge zu überwachen.

Eine Kamera ist als weiterer berührungslos arbeitender Sensor zu verstehen, da diese eine Array von Helligkeitssensoren darstellt und Bilder einer Szene zur Verfügung stellen kann. Die Auswertung der Bilder erfolgt durch die in Kapitel 3 vorgestellte Bildverarbeitung. Ziel ist die Erkennung einer Hand aus Bildfolgen und deren Lokalisation im Bild. Dazu ist als erster Schritt eine Bildvorverarbeitung mit einer Bewegungserkennung und Farbsegmentierung notwendig. Damit werden für die Handerkennung unnötige Bildinformationen, wie ein unbewegter Bildhintergrund, vorab entfernt.

Der Vorverarbeitung schließt sich eine Merkmalsextraktion an, die für das Bildverstehen wesentliche Informationen aus den Daten gewinnt. Hierzu wird in dieser Arbeit eine modifizierte Kreis-Hough-Transformation eingesetzt. Dabei wird das Merkmal Fingerspitze als eine Kombination von drei bis fünf Achtelkreisen betrachtet. Für eine Positionsbestimmung und eindeutige Handklassifikation werden unterschiedliche probabilistische Modelle eingesetzt. Dadurch kann dem System a priori Wissen über gültige Handgeometrien zur Verfügung gestellt werden.

Das im vorangegangenen Schritt erzeugte Wissen wird für die Initialisierung von Aktiven Konturen verwendet. Diese repräsentieren wiederum eine allgemeine Handform in einfacher Form und eignen sich für die nachfolgende Handverfolgung. Diese Aktive Konturen setzen sich aus B-Spline-Funktionen zusammen und erlauben daher eine Anpassung an die aktuell vorliegende Handform.

Die anschließende Handverfolgung dient dazu, die Trajektorie der gefundenen bewegten Hand zu bestimmen. Damit kann auf potentiell gefährliche Handbewegungen geschlossen werden. Die Verfolgung beruht auf einem Partikelfilter, der nach dem Condensation Algorithmus arbeitet.

Das Kamerasystem und die Bildverarbeitung erlauben die Lokalisation und Identifikation einer Hand sowie die Erfassung ihrer Bewegung im Bereich vor dem Sägeblatt einer Formatkreissäge. Voraussetzung für ein funktionierendes Bildverarbeitungssystem ist eine ausreichende Beleuchtung der Szene und eine leistungsfähige Verarbeitungseinheit.

Informationsfusion zur Gefahrenanalyse

Jedes der vorgestellten Sensorsysteme liefert nur eine Teilinformation über die gesamte aktuelle Situation im Gefahrenbereich einer Maschine. Daher ist es naheliegend, die einzelnen Informationen zu einer kompletten Situationsbeschreibung zusammenzufügen. Dieses Verfahren wird Informationsfusion genannt und in Kapitel 4 behandelt. Damit ist auch eine zuverlässige und robuste Gefahrenerkennung möglich.

Bevor die Informationen fusioniert werden ist jedoch eine Korrektur und Normierung der Eingangsdaten notwendig. So reagieren die kapazitiven Sensoren auf eine Änderung ihrer Geometrie, verursacht beispielsweise durch Auseinanderziehen von Elektroden oder Änderungen der Umgebungsbedingungen. Dies kann mithilfe von Kennfeldern korrigiert werden. Als weitere Eingangsgrößen für diese Kennfelder sind noch die Schutzlamellenposition und die Schutzhaubenhöhe zu erfassen. Nach der Korrektur und Normierung der Werte ist eine Diskretisierung in die entsprechenden Zustände sowie die Berechnung der Zustandswahrscheinlichkeiten erforderlich.

Nach dieser Vorverarbeitung können diese Zustandswahrscheinlichkeiten in einem Bayes'schen Netz fusioniert werden, bei dem die einzelnen Sensoren als Informationsknoten und die Fusionsergebnisse als Hypothesenknoten dargestellt sind. Die Struktur des Netzes orientiert sich in diesem Fall an den örtlichen Gegebenheiten einer Formatkreissäge. Um die Dynamik der Handbewegung zu berücksichtigen, wird die maximale Zustandswahrscheinlichkeit für eine erkannte Hand in jedem Hypothesenknoten über eine bestimmte Zeitdauer gespeichert und an den nachfolgenden weitergegeben. Das Netz ist modular aufgebaut, so dass das Hinzufügen oder Entfernen eines Informationsknotens möglich ist. Parallel zur Informationsfusion läuft im Bayes'schen Netz die Gefahrenanalyse. Diese basiert auf a priori Wissen über die Höhe der Zustandswahrscheinlichkeiten in den Endknoten. Diese Endknoten steuern die Aktivierung der Aktorik.

Sicherheitssystem

Das entwickelte aktive Sicherheitssystem für Formatkreissägen besteht aus der Gefahrenerkennung und einer Bremse als Aktorik. Diese wurde vom Institut für Werkzeugmaschinen, Universität Stuttgart, entwickelt. Das Gesamtsystem zur Gefahrenerkennung ist in Kapitel 5 dargestellt.

Für eine zuverlässige und reproduzierbare Gefahrenerkennung ist die Definition von Merkmalen erforderlich, welche eine Gefahrensituation beschreiben und mit den verfügbaren Sensoren gewonnen werden können. Diese Gefahrenmerkmale sind die Handidentifikation, Handposition und -bewegung (mit Geschwindigkeit und Bewegungsrichtung). Zulässige Grenzwerte dieser Merkmale ergeben sich aus den Größen der Überwachungsbereiche (Vorfeld 1 und 2) sowie des Gefährdungs-

bereiches direkt unter der Schutzhaube. Die maximale Vorschubgeschwindigkeit, bei der keine Gefahr angenommen wird, beträgt 0,67 m/s. Die stärksten Reflex- bzw. Wischbewegungen werden mit 2 m/s angegeben. Eine Bewegung in Richtung des Sägeblattes wird immer als gefährlich eingestuft.

Da seitlich auf das Sägeblatt zukommende Handbewegungen nur unzureichend erkannt werden können, sind neuartige Schutzlamellen an die Schutzhaube angebracht worden. Diese verhindern ein unbewusstes seitliches Eindringen in die Gefährdungszone. Zusätzlich ist die Kameraüberwachung für diesen Bereich vorgesehen.

Die oben erwähnten Sensorprinzipien sind an die Gegebenheiten einer Formatkreissäge angepasst und in die Schutzhaube integriert worden. Die Sensoren überwachen die für eine menschliche Hand potentiell gefährlichen Bereiche. Die dabei erzeugten Daten werden mit einem Bayes'schen Netz zusammengefasst, um die Zuverlässigkeit des Gesamtsystems zu erhöhen. Die parallel dazu im Netz ablaufende Gefahrenanalyse bewertet die von den Sensoren erfasste Szene. Entsprechend dem berechneten Verletzungsrisiko wird eine Benutzerwarnung ausgegeben oder die Neutralisierung der Gefahrenquelle eingeleitet, die in diesem Fall mit einer Sägeblattabbremmung erfolgt.

Fazit

Das Bildverarbeitungssystem arbeitet bisher zuverlässig mit einer Bildfrequenz von 6 fps. Dies ist für ein Sicherheitssystem noch nicht ausreichend, weshalb die Bildverarbeitung in einem separaten Testaufbau realisiert wurde. Eine Kamera kann über einen zusätzlichen Haltearm an einer Formatkreissäge angebracht werden und so einen Teil des Überwachungsbereiches beobachten. Mit den entwickelten Verfahren und einer dafür spezialisierten Verarbeitungseinheit kann die Überwachung zuverlässig mit Bildfrequenz über 50 fps erfolgen.

Die Informationsfusion mit dem Bayes'schen Netz ermöglicht die Integration der Sensordaten und die Gefahrenanalyse. Dieses System arbeitet in über 95 % der Fälle zuverlässig und verhindert dadurch eine Gefährdung des Benutzers. Störungen durch das Sicherheitssystem in Form von Falschauslösungen der Aktorik treten in weniger als 1 % der Fälle auf, so dass dieses System robust ist. Die Verarbeitungszeit der Daten und die damit verbundene Reaktionszeit des Systems sind ausreichend, um eine Gefahr rechtzeitig zu analysieren und darauf zu reagieren.

Das hier vorgestellte aktive Sicherheitssystem ermöglicht die Reduktion von potentiellen Gefahren für einen Maschinenbenutzer und trägt damit zu einer Steigerung der Arbeitssicherheit an gefährlichen Werkzeugmaschinen bei.

A Nomenklatur

Abkürzungen

CCD	Charge Coupled Device
CMOS	Complementary Metal Oxide Semiconductor
DAG	Directed Acyclic Graph
GHT	Generalized-Hough-Transformation
HSV	Hue-Saturation-Value
IEEE	Institute of Electrical and Electronic Engineers
IIC	Inter-Integrated-Circuit
KHT	Kreis-Hough-Transformation
LED	Light Emitting Diode (Leuchtdiode)
MMS	Mensch-Maschine-Schnittstelle
NI	National Instruments Corporation
PC	Personal Computer
RGB	Rot-Grün-Blau
ROI	Region of Interest, engl. betrachteter Ausschnitt eines Bildes
USB	Universal Serial Bus

Variablen, Symbole

Symbol	Einheit	Physikalische Größe
A		Ereignis
A	m^2	Fläche
a_i		Zustand i
B		Bildmatrix
B	T	Magnetische Flussdichte
B		Ereignis
$B_n(s)$		B-Spline Basisfunktion
B_κ		Parametermenge eines Objektmodells
b		Bild einer Sequenz
$bg(t)$		Hintergrundbild (Alpha-Segmentierer)
b_i		Zustand i
C	F	Kapazität
C_H		Menge von Farbwerten
$C_{ref}(h_S)$	F	Referenzkennlinie in Abhängigkeit von h_S
C_S		Menge von Farbwerten
C_{vor}	F	Gemessene Kapazität (<i>Vorfeld 1</i>)
$C_{vor,norm}$		Normiertes Ausgangssignal des kapazitiven Sensors (<i>Vorfeld 1</i>)

C_κ		Menge der Modellmerkmale
$c_{\kappa,l}$		Modellmerkmal
$C_t^{(n)}$		Kumulierte Wahrscheinlichkeit des n -Partikels zum Zeitschritt t
D_i		Dimension i einer Matrix oder Vektors
\mathbf{d}	m	Abstandsvektor
d	m	Abstand, Strecke
E	V/m	Elektrisches Feld
e		Abstände im Kennfeld in h_S -Richtung
\mathbf{F}		Binärbild
f	Hz	Frequenz
f_B	Bilder/s	Bildfrequenz
$fg(t)$		Vordergrundbild (Alpha-Segmentierer)
g_i		Abstände im Kennfeld in r_{SL} -Richtung
H		Hue-Wert (HSV-Farbraum)
h_S	m	Höhe der Schutzhaube
i		Zählvariable
j		Zählvariable
k		Zählvariable
k_t		Normierungskonstante
M		Anzahl an Normalen (Tracking)
m		Zählvariable
N		Anzahl Partikel, Messungen ,etc.
n_κ		Anzahl Modellmerkmale
\mathbf{O}		Menge der Beobachtungen
OR		Relatives Risiko
\mathbf{o}_k		Bildmerkmal
\mathbf{o}_κ		Merkmale eines lokalisierten Modells
p		Wahrscheinlichkeitsdichtefunktion
P		Wahrscheinlichkeit
Q		Kontrollvektor
\mathbf{R}		Rotationsmatrix
R_m	Ω	Messwiderstand
R_x		Filterkern (Schar-Operator)
R_y		Filterkern (Schar-Operator)
\mathbf{r}		Parametrische Spline-Kurve
r	m	Radius

r_{SL}	V	Position der Schutzlamelle
S		Saturation-Wert (HSV-Farbraum)
\bar{S}		Regularisierungsmatrix
S_{z_i}		Summe der Wahrscheinlichkeitsfunktionen
$\mathbf{s}_t^{(n)}$		n -tes Partikel zum Zeitschritt t
s_α		Schwellwert (Alpha-Segmentierer)
T	K	Temperatur
\mathbf{t}		Translationsvektor
t	s	Zeit
\mathbf{U}		Projektionsmatrix
U_{ind}	V	Induktionsspannung
U_{R_m}	V	Spannung am Messwiderstand R_m
U_{TH}	V	Thermospannung
$u_{IR,ref}$	V	Referenzwert des Infrarot-Sensors
$u_{IR,norm}$		Normiertes Ausgangssignal des Infrarot-Sensors
\hat{u}_{kl}^*	V	Interpolierter Korrekturwert (Lamelle)
u_{lam}	V	Ausgangsspannung des kapazitiven Sensors (Lamelle)
$u_{lam,norm}$		Normiertes Ausgangssignal eines kapazitiven Sensors (Lamelle)
V		Value-Wert (HSV-Farbraum)
\mathbf{v}	m/s	Geschwindigkeitsvektor
\mathbf{W}		Shape-Matrix
w		Anzahl der Leiterschleifen
\mathbf{X}		Shape-Vektor
x	m	Position
x		Koordinaten
\mathbf{x}		Zustände (Partikelfilter)
\mathbf{x}	m	Koordinatenvektor, Positionsvektor
y	m	Position
y		Koordinaten
\underline{Z}_C	Ω	Komplexe Impedanz
\mathbf{z}		Messung (Partikelfilter)
z_i		Zustände des Bayes'schen Netzes

Symbol	Einheit	Physikalische Größe
α		Vergessensfaktor (Alpha-Segmentierer)
α_s		Seebeck-Koeffizient
δ	m	Eindringtiefe
ϵ	F/m	Elektrische Feldkonstante
ϵ_r		Relative Permittivitätszahl
$\underline{\epsilon}_r$		Komplexe relative Permittivität
ϵ_V		Exzentrizität einer Ellipse
ϕ	Tm ²	Magnetischer Fluss
$\phi_{F,D}$		Orientierung einer Fingerspitze
ϕ_M		Drehwinkel des Modells
φ_{AK}		Objektrotation (Tracking)
κ	S	Leitfähigkeit
λ		Poissonparameter
μ		Permeabilität
$\bar{\mu}$		Mittelwert
μ_{z_i}		Zugehörigkeitsfunktion Zustand z_i
ϑ		Orientierung eines Kantengradienten
Σ		Kovarianzmatrix
σ^2		Varianz
σ		Standardabweichung
$\sigma_{S,x}^2$		Schwellwert (Varianzmethode)
$\sigma_{S,y}^2$		Schwellwert (Varianzmethode)
$\pi_t^{(n)}$		Gewichtung des n -te Partikels zum Zeitschritt t
ω	1/s	Rotationsgeschwindigkeit
Ω_κ		Objektklasse
Θ_H		Schwellwert
ν		Skalierungskonstante
ζ_κ		Zuordnungsvektor
ψ_{z_i}		Diskrete Zustandswahrscheinlichkeitsfunktion für z_i

B Parameterschätzung mit dem EM-Algorithmus

Für die Schätzung der Parameter der Handmodelle aus Abschnitt 3.6.1 auf Seite 78 werden normalverteilte Merkmale und $D_m = D_0$ (Dimension des Modellraums gleich der Dimension der Merkmalsvektoren) vorausgesetzt. Daher können die Schätzwerte in geschlossener Form angegeben werden. Die Herleitungen der folgenden angegebenen Formeln aus den Grundgleichungen des EM-Algorithmus sind in [35] zu finden.

Schätzung der diskreten Zuordnungswahrscheinlichkeiten

Die Iterationsformel zur Berechnung eines Schätzwertes für die diskreten Wahrscheinlichkeiten $\hat{p}_{\kappa,l}$ unter Verwendung des EM-Algorithmus lautet

$$\hat{p}_{\kappa,l}^{(i+1)} = \frac{1}{\sum_{\rho=1}^N \rho m} \sum_{\rho=1}^N \sum_{k=1}^{\rho m} \frac{\hat{p}_{\kappa,l}^{(i)} p(\rho \mathbf{o}_k | \hat{\mathbf{a}}_{\kappa,l}^{(i)}, \rho \mathbf{R}, \rho \mathbf{t})}{p(\rho \mathbf{o}_k | \hat{\mathbf{B}}_{\kappa}^{(i)}, \rho \mathbf{R}, \rho \mathbf{t})}$$

Durch Einsetzen der bereits angegebenen Dichtefunktionen für normalverteilte Merkmale ergibt sich:

$$\hat{p}_{\kappa,l}^{(i+1)} = \frac{1}{\sum_{\rho=1}^N \rho m} Q_{\kappa,l}^{(i)}$$

mit

$$Q_{\kappa,l}^{(i)} = \sum_{\rho=1}^N \sum_{k=1}^{\rho m} P_{\kappa,l,k,\rho}^{(i)} \tag{B.1}$$

und

$$P_{\kappa,l,k,\rho}^{(i)} = \frac{\hat{p}_{\kappa,l}^{(i)} \mathcal{N}(\rho \mathbf{o}_k | \rho \mathbf{R} \cdot \hat{\boldsymbol{\mu}}_{\kappa,l}^{(i)} + \rho \mathbf{t}, \rho \mathbf{R} \cdot \hat{\boldsymbol{\Sigma}}_{\kappa,l}^{(i)} \cdot \rho \mathbf{R}^T)}{\sum_{j=1}^{n_{\kappa}} \hat{p}_{\kappa,j}^{(i)} \mathcal{N}(\rho \mathbf{o}_k | \rho \mathbf{R} \cdot \hat{\boldsymbol{\mu}}_{\kappa,j}^{(i)} + \rho \mathbf{t}, \rho \mathbf{R} \cdot \hat{\boldsymbol{\Sigma}}_{\kappa,j}^{(i)} \cdot \rho \mathbf{R}^T)} \tag{B.2}$$

Schätzung der Mittelwerte der Normalverteilung

Die Mittelwerte werden ebenfalls iterativ geschätzt. Unter Verwendung des EM-Algorithmus kann folgende geschlossene Iterationsformel angegeben werden

$$\widehat{\boldsymbol{\mu}}_{\kappa,l}^{(i+1)} = \frac{\sum_{\rho=1}^N \sum_{k=1}^{\rho m} \frac{\hat{p}_{\kappa,l}^{(i)} p(\rho \mathbf{o}_k | \widehat{\mathbf{a}}_{\kappa,l}^{(i)}, \rho \mathbf{R}, \rho \mathbf{t})}{p(\rho \mathbf{o}_k | \widehat{\mathbf{B}}_{\kappa}^{(i)}, \rho \mathbf{R}, \rho \mathbf{t})} \cdot \rho \mathbf{R}^{-1} \cdot (\rho \mathbf{o}_k - \rho \mathbf{t})}{\sum_{\rho=1}^N \sum_{k=1}^{\rho m} \frac{\hat{p}_{\kappa,l}^{(i)} p(\rho \mathbf{o}_k | \widehat{\mathbf{a}}_{\kappa,l}^{(i)}, \rho \mathbf{R}, \rho \mathbf{t})}{p(\rho \mathbf{o}_k | \widehat{\mathbf{B}}_{\kappa}^{(i)}, \rho \mathbf{R}, \rho \mathbf{t})}}$$

Dies lässt sich wiederum unter Zuhilfenahme der Gleichungen (B.1) und (B.2) verkürzt darstellen

$$\widehat{\boldsymbol{\mu}}_{\kappa,l}^{(i+1)} = \frac{\sum_{\rho=1}^N \sum_{k=1}^{\rho m} P_{\kappa,l,k,\rho}^{(i)} \cdot \rho \mathbf{R}^{-1} \cdot (\rho \mathbf{o}_k - \rho \mathbf{t})}{Q_{\kappa,l}^{(i)}}$$

Schätzung der Kovarianzmatrix der Normalverteilung

Für den hier gültigen Fall $D_m = D_o$ kann folgende geschlossene Iterationsformel für die Kovarianzmatrix verwendet werden

$$\widehat{\boldsymbol{\Sigma}}_{\kappa,l}^{(i+1)} = \frac{\sum_{\rho=1}^N \sum_{k=1}^{\rho m} \frac{\hat{p}_{\kappa,l}^{(i)} p(\rho \mathbf{o}_k | \widehat{\mathbf{a}}_{\kappa,l}^{(i)}, \rho \mathbf{R}, \rho \mathbf{t})}{p(\rho \mathbf{o}_k | \widehat{\mathbf{B}}_{\kappa}^{(i)}, \rho \mathbf{R}, \rho \mathbf{t})} \cdot \rho \mathbf{R}^{-1} \cdot \rho \widehat{\mathbf{S}}_{\kappa,l,k}^{(i)} \cdot (\rho \mathbf{R}^T)^{-1}}{\sum_{\rho=1}^N \sum_{k=1}^{\rho m} \frac{\hat{p}_{\kappa,l}^{(i)} p(\rho \mathbf{o}_k | \widehat{\mathbf{a}}_{\kappa,l}^{(i)}, \rho \mathbf{R}, \rho \mathbf{t})}{p(\rho \mathbf{o}_k | \widehat{\mathbf{B}}_{\kappa}^{(i)}, \rho \mathbf{R}, \rho \mathbf{t})}}$$

mit

$$\rho \widehat{\mathbf{S}}_{\kappa,l,k}^{(i)} = (\rho \mathbf{o}_k - \rho \mathbf{R} \widehat{\boldsymbol{\mu}}_{\kappa,l}^{(i)} - \rho \mathbf{t}) \cdot (\rho \mathbf{o}_k - \rho \mathbf{R} \widehat{\boldsymbol{\mu}}_{\kappa,l}^{(i)} - \rho \mathbf{t})^T$$

Auch hier kann durch Einsetzen der Gleichungen (B.1) und (B.2) eine verkürzte Darstellung erreicht werden

$$\widehat{\boldsymbol{\Sigma}}_{\kappa,l}^{(i+1)} = \frac{\sum_{\rho=1}^N \sum_{k=1}^{\rho m} P_{\kappa,l,k,\rho}^{(i)} \cdot \rho \mathbf{R}^{-1} \cdot \rho \widehat{\mathbf{S}}_{\kappa,l,k}^{(i)} \cdot (\rho \mathbf{R}^T)^{-1}}{Q_{\kappa,l}^{(i)}} \quad (\text{B.3})$$

C Algorithmus zur Lösung des Fitting-Problems

Der komplette Algorithmus zur Lösung des Fitting-Problems nach [11] ist im Folgenden dargestellt. Dabei wird in den ersten beiden Schritten zunächst die Menge der Merkmalspunkte $\mathbf{r}_f(s_i)$ ermittelt, welche die gesuchte Objektkontur repräsentieren. Im dritten Schritt wird dann die *Informationsmatrix* S_i initialisiert, welche die Gewichtung jeder Schätzung $\hat{\mathbf{X}}_i$ bezüglich der ersten i Merkmalspunkte bestimmt. Auch die *Matrix der gewichteten Summe der Informationen* \mathbf{Z}_i wird initialisiert. In \mathbf{Z}_i wird der Einfluss der mittleren Form $\bar{\mathbf{X}}$ und der einzelnen gewichteten Messungen akkumuliert. Im Schritt 4 werden entlang der Kurve die jeweiligen Komponenten akkumuliert. Dabei wird jeweils nur die normale Komponente ν_i jeder Messung benutzt. σ_i stellt einen hier als konstant angenommenen Messfehler $\sigma_i = \sigma = \sqrt{N}$ dar. Schlussendlich werden die akkumulierten Messungen \mathbf{Z} zu einem neu geschätzten Shape-Vektor $\hat{\mathbf{X}}$ zusammengefügt.

Der Fitting-Algorithmus nach [11]:

Gegeben:

Schätzung der Anfangskurve $\bar{\mathbf{r}}(s)$ bzw. $\bar{\mathbf{X}}$ mit den Normalen $\bar{\mathbf{n}}(s)$ und die Regularisierungsmatrix $\bar{\mathbf{S}}$.

Gesucht:

Lösung des Minimierungsproblems

$$\min_{\mathbf{X}} \left\{ (\mathbf{X} - \bar{\mathbf{X}})^T \bar{\mathbf{S}} (\mathbf{X} - \bar{\mathbf{X}}) + \sum_{i=1}^N \frac{1}{\sigma_i^2} (\nu_i - \mathbf{h}(s_i)^T [\mathbf{X} - \bar{\mathbf{X}}])^2 \right\}$$

Algorithmus:

1. Wähle Stützstellen $s_i, i = 1, \dots, N$, so dass $s_1 = 0, s_{i+1} = s_i + h, s_N = L$
2. Finde für jedes i entlang der Kurvennormale $\bar{\mathbf{n}}(s_i)$ durch den Punkt $\bar{\mathbf{r}}(s_i)$ die Position des Merkmals $\mathbf{r}_f(s_i)$.
3. Initialisiere

$$\mathbf{Z}_0 = 0, \quad \mathbf{S}_0 = 0$$

4. Iteriere über $i = 1, \dots, N$:

$$\begin{aligned}\nu_i &= (\mathbf{r}_f(s_i) - \bar{\mathbf{r}}(s_i)) \cdot \bar{\mathbf{n}}(s_i) \\ \mathbf{h}(s_i)^T &= \bar{\mathbf{n}}(s_i)^T \mathbf{U}(s_i) \mathbf{W} \\ \mathbf{S}_i &= \mathbf{S}_{i-1} + \frac{1}{\sigma_i^2} \mathbf{h}(s_i) \mathbf{h}(s_i)^T \\ \mathbf{Z}_i &= \mathbf{Z}_{i-1} + \frac{1}{\sigma_i^2} \mathbf{h}(s_i) \nu_i\end{aligned}$$

5. Die akkumulierte Informationsmatrix und der akkumulierte Beobachtungsvektor sind damit

$$\mathbf{Z} = \mathbf{Z}_N, \quad \mathbf{S} = \mathbf{S}_N$$

6. Damit resultiert die optimal angepasste Kurve im Shape-Space:

$$\hat{\mathbf{X}} = \bar{\mathbf{X}} + (\bar{\mathbf{S}} + \mathbf{S})^{-1} \mathbf{Z}$$

D Condensation Algorithmus

Im Folgenden ist der Ablauf des Condensation Algorithmus nach [41] dargestellt. Eine detaillierte Beschreibung findet sich auch in Abschnitt 3.8.2 auf Seite 93.

Um aus der alten Partikelmenge $\{(\mathbf{s}_{t-1}^{(n)}, \pi_{t-1}^{(n)}, c_{t-1}^{(n)}), n = 1, \dots, N\}$ zum Zeitpunkt $t-1$ eine neue Partikelmenge $\{(\mathbf{s}_t^{(n)}, \pi_t^{(n)}, c_t^{(n)}), n = 1, \dots, N\}$ für den Zeitpunkt t zu konstruieren, müssen für jedes n -te neue Partikel die folgenden Schritte durchgeführt werden:

1. **Auswahl** eines Partikels $\mathbf{s}_t'^{(n)} = \mathbf{s}_{t-1}^{(j)}$ durch Ziehen mit Zurücklegen entsprechend der Wahrscheinlichkeit $\pi_{t-1}^{(j)}$:
 - a) Generiere eine gleichverteilte Zufallszahl $r \in [0,1]$
 - b) Finde durch eine Binäre Suche das kleinste j mit $c_{t-1}^{(j)} \geq r$
 - c) Setze $\mathbf{s}_t'^{(n)} = \mathbf{s}_{t-1}^{(j)}$

Dabei können hoch gewichtete Partikel ein- oder auch mehrmals ausgewählt werden, während niedrig gewichtete Partikel unter Umständen überhaupt nicht ausgewählt werden. Dieser Schritt wird auch als *Resampling* bezeichnet.

2. **Prädiktion** des neuen Objektzustandes für das ausgewählte Partikel $\mathbf{s}_t'^{(n)}$ nach $p(\mathbf{x}_t | \mathbf{x}_{t-1} = \mathbf{s}_t'^{(n)})$. Den neuen Objektzustand $\mathbf{s}_t^{(n)}$ erhält man damit durch Anwendung des Bewegungsmodells auf $\mathbf{s}_t'^{(n)}$.
3. **Messung** und dementsprechende Gewichtung des neuen Objektzustandes anhand des Beobachtungsmodells und der gemessenen Merkmale \mathbf{z}_t :

$$\pi_t^{(n)} = p(\mathbf{z}_t | \mathbf{x}_t = \mathbf{s}_t^{(n)})$$

Nachdem auf diese Weise iterativ alle N neuen Partikel erzeugt wurden, müssen die so gefundenen Gewichte noch normalisiert werden, so dass gilt:

$$\sum_{n=0}^N \pi_t^{(n)} = 1$$

Zudem muss jedem Partikel noch das kumulative Gewicht $c_t^{(n)}$ zugewiesen werden, wobei gelten muss:

$$\begin{aligned} c_t^{(0)} &= 0 \\ c_t^{(n)} &= c_t^{(n-1)} + \pi_t^{(n)} \quad (n = 1, \dots, N) \end{aligned}$$

E Beispiel eines Bayes'schen Netzes

Das folgende Beispiel ist [43] entnommen und erklärt anschaulich die Funktionsweise eines Bayes'schen Netzes.

Eines Morgens verlässt ein Herr Holmes (H) sein Haus und stellt fest, dass das Gras in seinem Vorgarten nass ist. Er überlegt, ob er am Vorabend vergessen hat, seinen Rasensprenkler (S) auszuschalten, oder ob es über Nacht geregnet (R) hat. Als er das ebenfalls nasse Gras seines Nachbarn Dr. Watson (W) bemerkt, schließt Herr Holmes, dass es mit sehr hoher Wahrscheinlichkeit über Nacht geregnet haben muss.

Eine graphische Veranschaulichung dieses Sachverhalts ist in Abbildung E.1 dargestellt.

Durch die Überlegung, warum sein Grass nass ist, führt Herr Holmes Schlussfolgerungen in entgegengesetzter Richtung der Pfeile durch. Da beide Pfeile auf Holmes zeigen, erhöht sich für beide Zustände Regen (R) und Sprenkler (S) die Wahrscheinlichkeit. Durch die erhöhte Wahrscheinlichkeit von Regen (R) wiederum erhöht sich auch die Wahrscheinlichkeit, dass der Rasen von Dr. Watson (W) nass ist. Aus diesem Grund überprüft Herr Holmes dessen Rasen und stellt fest, dass dieser ebenfalls nass ist. Dieses neue Wissen erhöht die Wahrscheinlichkeit für Regen (R) drastisch.

Der nächste Schritt ist für Maschinen sehr schwer zu berechnen, während er für Menschen absolut trivial erscheint. Er wird *explaining away* genannt. Der Grund für den nassen Rasen von Herr Holmes scheint durch Regen verursacht zu sein. Dadurch gibt es keinen Grund mehr zu glauben, dass er den Sprenkler am Vorabend nicht ausgeschaltet hat. Deshalb reduziert sich die Wahrscheinlichkeit des Sprenklers (S) auf seine Initialisierungswahrscheinlichkeit.

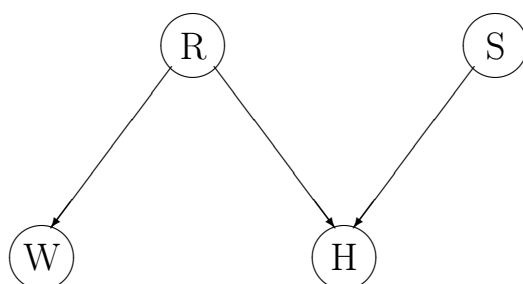


Abbildung E.1: Beispiel: Nasses Gras

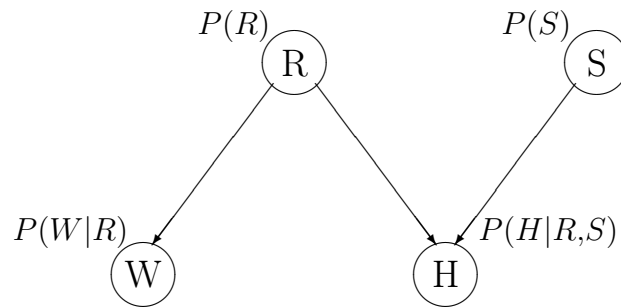


Abbildung E.2: Beispiel: Nasses Gras

In diesem einfachen Beispiel wird deutlich, wie durch das Wissen (Evidenz) über die Zustände von zwei beobachtbaren Knoten eines Bayes'schen Netzes, den Rasenstücken der beiden Nachbarn, auf die Wahrscheinlichkeit der anderen beiden Knoten geschlossen werden kann. Man erhält durch dieses logische Schließen Wahrscheinlichkeiten für die zwei nicht beobachtbaren Zustände Regen und Rasensprenkler.

Ebenso ist auch deutlich zu sehen, dass sich Abhängigkeiten in Bayes'schen Netzen ändern können. Während im Grundzustand die Wahrscheinlichkeiten für Regen (R) und Sprenkler (S) unabhängig voneinander sind, werden sie durch das Wissen, dass das Gras von Herrn Holmes nass ist, abhängig.

Mathematische Betrachtung des Beispiels

Die grafische Struktur des Beispiels ist in Abbildung E.2 mit den dazugehörigen Wahrscheinlichkeiten dargestellt. Alle Variablen dieses Bayes'schen Netzes besitzen genau zwei diskrete Zustände: „Ja“ und „Nein“. Um ein vollständig definiertes Bayes'sches Netz zu erhalten, werden allerdings noch die bedingten Wahrscheinlichkeitstabellen und die Initialisierungswahrscheinlichkeiten der Knoten ohne Eltern benötigt.

Die Initialisierungswahrscheinlichkeiten für R und S sind z. B. $P(R) = (0,2; 0,8)$ und $P(S) = (0,1; 0,9)$. Die Wahrscheinlichkeit, dass es über Nacht geregnet hat, liegt in diesem Beispiel also bei 20 %, während die, dass der Rasensprenkler versehentlich nicht abgestellt wurde, bei 10 % liegt. Die außerdem noch benötigten bedingten Wahrscheinlichkeitstabellen $P(W|R)$ und $P(H|R,S)$ sind in den Tabellen E.1 aufgelistet.

Mittels der in Abbildung E.2 dargestellten Struktur, der Initialisierungswahrscheinlichkeiten $P(R)$ und $P(S)$ und der Wahrscheinlichkeitstabellen in Tabelle E.1 ist das Bayes'sche Netz vollständig definiert.

Mithilfe der Gleichungen (4.2) und 4.3 auf Seite 125 können nun die Wahrscheinlichkeiten für einen nassen Rasen bei Dr. Watson, $P(W)$, und Herrn Holmes, $P(H)$, berechnet werden.

	R=Ja	R=Nein
W=Ja	1	0,2
W=Nein	0	0,8

(a) $P(W|R)$

	R=Ja	R=Nein
S=Ja	(1 ; 0)	(0,9 ; 0,1)
S=Nein	(1 ; 0)	(0 ; 1)

(b) $P(H|R,S)$

(α, β) steht für (H=Ja, H=Nein)

Tabelle E.1: Bedingte Wahrscheinlichkeitstabellen

	R=Ja	R=Nein
W=Ja	0,2	0,16
W=Nein	0	0,64

(a) $P(W,R)$

	R=Ja	R=Nein
S=Ja	(0,02 ; 0)	(0,072 ; 0,008)
S=Nein	(0,18 ; 0)	(0 ; 0,72)

(b) $P(H,R,S)$

(α, β) steht für (H=Ja, H=Nein)

Tabelle E.2: Verbundwahrscheinlichkeiten

Zuerst wird $P(W,R)$ durch Anwendung der Gleichung (4.2) berechnet. Ebenso wird die Verbundwahrscheinlichkeit $P(H,R,S)$ bestimmt, nur mit dem Unterschied, dass $P(H|R,S)$ von zwei Bedingungen abhängt

$$P(H,R,S) = P(H|R,S)P(R,S) . \quad (\text{E.1})$$

Da aber die Variablen R und S unabhängig voneinander sind, kann Gleichung (E.1) umgeformt werden zu:

$$P(H,R,S) = P(H|R,S)P(R)P(S) .$$

Die Ergebnisse sind in Tabelle E.2 aufgeführt.

Um nun die Wahrscheinlichkeiten für $P(W)$ und $P(H)$ zu erhalten, müssen noch R und S aus obigen Dichten marginalisiert werden. Das Ergebnis lautet:

$$P(W) = (0,36 ; 0,64) \quad \text{und} \quad P(H) = (0,272 ; 0,728) .$$

Durch die Entdeckung von Herrn Holmes, dass sein Rasen nass ist, wird Evidenz in das Netz eingefügt und die Wahrscheinlichkeiten müssen aktualisiert werden $P(H) = (1 ; 0)$. Dazu werden ausgehend von dem Knoten, für den Evidenz in das Netz eingefügt wurde, alle Wahrscheinlichkeiten des Netzes neu berechnet. Eine Möglichkeit ist, die Wahrscheinlichkeiten der Elternknoten mithilfe der Bayes'schen Regel (Gleichung (4.2)) zu berechnen. Anschließend würden dann, wie oben gezeigt, die Wahrscheinlichkeiten aller Kindknoten durch die Aktualisierung der Verbundwahrscheinlichkeiten vorgenommen und ausmarginalisiert.

	R=Ja	R=Nein
S=Ja	(0,02 ; 0)	(0,072 ; 0)
S=Nein	(0,18 ; 0)	(0 ; 0)

(a) $P(H,R,S) \cdot P^*(H)$

	R=Ja	R=Nein
S=Ja	(0,074 ; 0)	(0,264 ; 0)
S=Nein	(0,662 ; 0)	(0 ; 0)

(b) $P^*(H,R,S)$

Tabelle E.3: Berechnung und Ergebnis der aktualisierten Verbundwahrscheinlichkeit

	R=Ja	R=Nein
S=Ja	$0,02 \cdot \frac{0,736}{0,2}$	$0,16 \cdot \frac{0,264}{0,8}$
S=Nein	0	$0,64 \cdot \frac{0,264}{0,8}$

(a) $P^*(W,R) = P(W,R) \cdot \frac{P^*(R)}{P(R)}$

	R=Ja	R=Nein
W=Ja	0,736	0,0528
W=Nein	0	0,2112

(b) $P^*(W,R)$

Tabelle E.4: Berechnung und Ergebnis der aktualisierten Verbundwahrscheinlichkeit

Eine wesentlich effektivere Methode ist die Verbundwahrscheinlichkeit des Knoten, für den Evidenz vorliegt, mit seinen Elternknoten zu aktualisieren. Dies geschieht, indem man ihn mit dessen neuer Wahrscheinlichkeit, der Evidenz, multipliziert und durch die bisherige Wahrscheinlichkeit dividiert. Für das hier betrachtete Beispiel erhält man so:

$$P^*(H,R,S) = \frac{P(H,R,S)}{P(H)} \cdot P^*(H) .$$

Der Vorteil dieser Methode ist, dass bei harter Evidenz die neue Wahrscheinlichkeit entweder (0 ; 1) oder (1 ; 0) lautet und so die Multiplikation durch einfaches Überschreiben der Hälfte der Werte mit null durchgeführt werden kann.

So ergeben sich durch die Multiplikation die Werte in Tabelle E.3 (a) und durch die anschließende Division die neue Verbundwahrscheinlichkeit $P^*(H,R,S)$ in Tabelle E.3 (b).

Durch die Marginalisierung erhält man

$$P^*(R) = (0,736 ; 0,264) \quad \text{und} \quad P^*(S) = (0,338 ; 0,662) .$$

Die Aktualisierung der Wahrscheinlichkeit der Variablen W geschieht auf dieselbe Art und Weise.

$$P^*(W,R) = \frac{P(W,R)}{P(R)} \cdot P^*(R)$$

Die Berechnung und das Ergebnis sind in Tabelle E.4 dargestellt.

Ebenso kann auch die Zusatzinformation, dass der Rasen von Dr. Watson auch nass ist, mit eingebracht werden. Letztendlich erhält man bei diesem Beispiel die

Wahrscheinlichkeit für Regen von $P^{**}(R=Ja) = 0,93$ und für den Sprenkler von $P^{**}(S=Ja) = 0,161$. Die ausführliche Rechnung ist in [43] dargelegt.

Der Grund, weshalb die Wahrscheinlichkeit für den Sprenkler nicht auf seine Initialisierungswahrscheinlichkeit zurückgefallen ist, besteht darin, dass Dr. Watson ein vergesslicher Mensch ist und er eventuell zusätzlich seinen Sprenkler über Nacht angeschaltet hatte. Diese Vergesslichkeit spiegelt sich in der bedingten Wahrscheinlichkeit $P(W=Ja|R=Nein) = 0,2$ wieder.

Die Berechnung der Wahrscheinlichkeiten der einzelnen Zustände nach Einfügen von neuem Wissen (Evidenz) kann bei kleineren Netzen wie im vorangegangenen Beispiel Schritt für Schritt berechnet werden. Bei größeren Netzen erzeugt diese Art der Berechnung jedoch vor allem bei rückwärts gerichteter Inferenz einen großen Rechenaufwand, s. Abschnitt 4.2.3 auf Seite 132). In der Literatur existieren verschiedene Methoden zur Berechnung der Inferenz, s. [67].

F Bedingte Wahrscheinlichkeitstabellen für das Bayes'sche Netz

In den folgenden Tabellen werden die bedingten Wahrscheinlichkeitstabellen für das Bayes'sche Netz aus Abschnitt 4.4.4 aufgelistet.

Die erste Tabelle F.1 zeigt die bedingte Wahrscheinlichkeitstabelle des Knoten *Vorfeld 1*, der nur vom Knoten *Kamera* abhängt und dieselben Zustände wie dieser besitzt.

<i>KAM</i>	<i>V1</i>	
	<i>keine hand</i>	<i>hand</i>
<i>keine hand</i>	1	0
<i>hand</i>	0	1

Tabelle F.1: Bedingte Übergangswahrscheinlichkeitstabelle $P(V1|KAM)$

Die Abhängigkeit des Knoten *Vorfeld 2 vorne* von dessen Elternknoten *Vorfeld 1* und *IR Vorfeld* ist in der Tabelle F.2 aufgelistet.

<i>V1</i>	<i>IR_V</i>	<i>V2V</i>			
		<i>nichts</i>	<i>kalt</i>	<i>warm</i>	<i>hand</i>
<i>keine hand</i>	<i>nichts</i>	1	0	0	0
<i>hand</i>	<i>nichts</i>	1	0	0	0
<i>keine hand</i>	<i>kalt</i>	0	1	0	0
<i>hand</i>	<i>kalt</i>	0	0	0	1
<i>keine hand</i>	<i>warm</i>	0	0	1	0
<i>hand</i>	<i>warm</i>	0	0	0	1

Tabelle F.2: Bedingte Übergangswahrscheinlichkeitstabelle $P(V2V|V1,IR_V)$

In Tabelle F.3 sind die bedingten Übergangswahrscheinlichkeiten des Knoten *Vorfeld 2 hinten* verzeichnet.

<i>V2V</i>	<i>KAP_V</i>	<i>MAT</i>	<i>V2H</i>			
			<i>nichts</i>	<i>holz</i>	<i>metall</i>	<i>hand</i>
<i>nichts</i>	<i>nichts</i>	<i>kein metall</i>	1	0	0	0
<i>kalt</i>	<i>nichts</i>	<i>kein metall</i>	0,9	0,1	0	0
<i>warm</i>	<i>nichts</i>	<i>kein metall</i>	0,9	0	0	0,1
<i>hand</i>	<i>nichts</i>	<i>kein metall</i>	0,8	0	0	0,2
<i>nichts</i>	<i>kleines ϵ_r</i>	<i>kein metall</i>	0,3	0,7	0	0
<i>kalt</i>	<i>kleines ϵ_r</i>	<i>kein metall</i>	0	1	0	0
<i>warm</i>	<i>kleines ϵ_r</i>	<i>kein metall</i>	0	0,4	0	0,6
<i>hand</i>	<i>kleines ϵ_r</i>	<i>kein metall</i>	0	0	0	1
<i>nichts</i>	<i>großes ϵ_r</i>	<i>kein metall</i>	0	0,9	0	0,1
<i>kalt</i>	<i>großes ϵ_r</i>	<i>kein metall</i>	0	0,7	0	0,3
<i>warm</i>	<i>großes ϵ_r</i>	<i>kein metall</i>	0	0	0	1
<i>hand</i>	<i>großes ϵ_r</i>	<i>kein metall</i>	0	0	0	1
<i>nichts</i>	<i>nichts</i>	<i>metall</i>	1	0	0	0
<i>kalt</i>	<i>nichts</i>	<i>metall</i>	0,9	0,1	0	0
<i>warm</i>	<i>nichts</i>	<i>metall</i>	0,9	0	0	0,1
<i>hand</i>	<i>nichts</i>	<i>metall</i>	0,8	0	0	0,2
<i>nichts</i>	<i>kleines ϵ_r</i>	<i>metall</i>	0,3	0	0,7	0
<i>kalt</i>	<i>kleines ϵ_r</i>	<i>metall</i>	0	0	1	0
<i>warm</i>	<i>kleines ϵ_r</i>	<i>metall</i>	0	0	0,5	0,5
<i>hand</i>	<i>kleines ϵ_r</i>	<i>metall</i>	0	0	0,1	0,9
<i>nichts</i>	<i>großes ϵ_r</i>	<i>metall</i>	0	0	1	0
<i>kalt</i>	<i>großes ϵ_r</i>	<i>metall</i>	0	0	0,9	0,1
<i>warm</i>	<i>großes ϵ_r</i>	<i>metall</i>	0	0	0	1
<i>hand</i>	<i>großes ϵ_r</i>	<i>metall</i>	0	0	0	1

Tabelle F.3: Bedingte Übergangswahrscheinlichkeitstabelle $P(V2H|V2V,KAP_V,MAT)$

In Tabelle F.4 ist die bedingte Wahrscheinlichkeitstabelle des Knoten *Lamelle 1 links* in Abhängigkeit seiner Elternknoten *Vorfeld 2 hinten*, *IR rechts* und *Kap 1 rechts* sowie *Vorfeld 1* wiedergegeben.

<i>V2H</i>	<i>IR_L</i>	<i>KAP_L1</i>	<i>KAM</i>	<i>LAM_L1</i>			
				<i>nichts</i>	<i>holz</i>	<i>metall</i>	<i>hand</i>
<i>nichts</i>	<i>nichts</i>	<i>nichts</i>	<i>keine hand</i>	1	0	0	0
<i>holz</i>	<i>nichts</i>	<i>nichts</i>	<i>keine hand</i>	0,9	0,1	0	0
<i>metall</i>	<i>nichts</i>	<i>nichts</i>	<i>keine hand</i>	0,9	0	0,1	0
<i>hand</i>	<i>nichts</i>	<i>nichts</i>	<i>keine hand</i>	0,9	0	0	0,1
<i>nichts</i>	<i>kalt</i>	<i>nichts</i>	<i>keine hand</i>	0,9	0,1	0	0
<i>holz</i>	<i>kalt</i>	<i>nichts</i>	<i>keine hand</i>	0,9	0,1	0	0
<i>metall</i>	<i>kalt</i>	<i>nichts</i>	<i>keine hand</i>	0,9	0	0,1	0
<i>hand</i>	<i>kalt</i>	<i>nichts</i>	<i>keine hand</i>	0,8	0	0	0,2
<i>nichts</i>	<i>warm</i>	<i>nichts</i>	<i>keine hand</i>	0,9	0	0	0,1
<i>holz</i>	<i>warm</i>	<i>nichts</i>	<i>keine hand</i>	0,6	0,3	0	0,1
<i>metall</i>	<i>warm</i>	<i>nichts</i>	<i>keine hand</i>	0,8	0	0,1	0,1
<i>hand</i>	<i>warm</i>	<i>nichts</i>	<i>keine hand</i>	0,8	0	0	0,2
<i>nichts</i>	<i>nichts</i>	<i>kleines ϵ_r</i>	<i>keine hand</i>	0,3	0,7	0	0
<i>holz</i>	<i>nichts</i>	<i>kleines ϵ_r</i>	<i>keine hand</i>	0,1	0,9	0	0
<i>metall</i>	<i>nichts</i>	<i>kleines ϵ_r</i>	<i>keine hand</i>	0,1	0	0,9	0
<i>hand</i>	<i>nichts</i>	<i>kleines ϵ_r</i>	<i>keine hand</i>	0,1	0,4	0,4	0,1
<i>nichts</i>	<i>kalt</i>	<i>kleines ϵ_r</i>	<i>keine hand</i>	0,1	0,45	0,45	0
<i>holz</i>	<i>kalt</i>	<i>kleines ϵ_r</i>	<i>keine hand</i>	0	1	0	0
<i>metall</i>	<i>kalt</i>	<i>kleines ϵ_r</i>	<i>keine hand</i>	0	0	1	0
<i>hand</i>	<i>kalt</i>	<i>kleines ϵ_r</i>	<i>keine hand</i>	0	0	0,7	0,3
<i>nichts</i>	<i>warm</i>	<i>kleines ϵ_r</i>	<i>keine hand</i>	0,1	0,1	0,1	0,7
<i>holz</i>	<i>warm</i>	<i>kleines ϵ_r</i>	<i>keine hand</i>	0	0,1	0	0,9
<i>metall</i>	<i>warm</i>	<i>kleines ϵ_r</i>	<i>keine hand</i>	0	0	0	1
<i>hand</i>	<i>warm</i>	<i>kleines ϵ_r</i>	<i>keine hand</i>	0	0	0	1
<i>nichts</i>	<i>nichts</i>	<i>großes ϵ_r</i>	<i>keine hand</i>	0	0	0,95	0,05
<i>holz</i>	<i>nichts</i>	<i>großes ϵ_r</i>	<i>keine hand</i>	0	0	0,95	0,05
<i>metall</i>	<i>nichts</i>	<i>großes ϵ_r</i>	<i>keine hand</i>	0	0	1	0
<i>hand</i>	<i>nichts</i>	<i>großes ϵ_r</i>	<i>keine hand</i>	0	0	1	0
<i>nichts</i>	<i>kalt</i>	<i>großes ϵ_r</i>	<i>keine hand</i>	0	0	0,9	0,1
<i>holz</i>	<i>kalt</i>	<i>großes ϵ_r</i>	<i>keine hand</i>	0	0,2	0,6	0,2
<i>metall</i>	<i>kalt</i>	<i>großes ϵ_r</i>	<i>keine hand</i>	0	0	1	0
<i>hand</i>	<i>kalt</i>	<i>großes ϵ_r</i>	<i>keine hand</i>	0	0	0,7	0,3
<i>nichts</i>	<i>warm</i>	<i>großes ϵ_r</i>	<i>keine hand</i>	0	0	0	1
<i>holz</i>	<i>warm</i>	<i>großes ϵ_r</i>	<i>keine hand</i>	0	0	0	1
<i>metall</i>	<i>warm</i>	<i>großes ϵ_r</i>	<i>keine hand</i>	0	0	0	1
<i>hand</i>	<i>warm</i>	<i>großes ϵ_r</i>	<i>keine hand</i>	0	0	0	1
<i>nichts</i>	<i>nichts</i>	<i>nichts</i>	<i>hand</i>	1	0	0	0
<i>holz</i>	<i>nichts</i>	<i>nichts</i>	<i>hand</i>	0,9	0,1	0	0
<i>metall</i>	<i>nichts</i>	<i>nichts</i>	<i>hand</i>	0,9	0	0,1	0

<i>hand</i>	<i>nichts</i>	<i>nichts</i>	<i>hand</i>	0,9	0	0	0,1
<i>nichts</i>	<i>kalt</i>	<i>nichts</i>	<i>hand</i>	0	0	0	1
<i>holz</i>	<i>kalt</i>	<i>nichts</i>	<i>hand</i>	0	0	0	1
<i>metall</i>	<i>kalt</i>	<i>nichts</i>	<i>hand</i>	0	0	0	1
<i>hand</i>	<i>kalt</i>	<i>nichts</i>	<i>hand</i>	0	0	0	1
<i>nichts</i>	<i>warm</i>	<i>nichts</i>	<i>hand</i>	0	0	0	1
<i>holz</i>	<i>warm</i>	<i>nichts</i>	<i>hand</i>	0	0	0	1
<i>metall</i>	<i>warm</i>	<i>nichts</i>	<i>hand</i>	0	0	0	1
<i>hand</i>	<i>warm</i>	<i>nichts</i>	<i>hand</i>	0	0	0	1
<i>nichts</i>	<i>nichts</i>	<i>kleines ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>holz</i>	<i>nichts</i>	<i>kleines ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>metall</i>	<i>nichts</i>	<i>kleines ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>hand</i>	<i>nichts</i>	<i>kleines ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>nichts</i>	<i>kalt</i>	<i>kleines ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>holz</i>	<i>kalt</i>	<i>kleines ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>metall</i>	<i>kalt</i>	<i>kleines ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>hand</i>	<i>kalt</i>	<i>kleines ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>nichts</i>	<i>warm</i>	<i>kleines ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>holz</i>	<i>warm</i>	<i>kleines ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>metall</i>	<i>warm</i>	<i>kleines ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>hand</i>	<i>warm</i>	<i>kleines ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>nichts</i>	<i>nichts</i>	<i>großes ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>holz</i>	<i>nichts</i>	<i>großes ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>metall</i>	<i>nichts</i>	<i>großes ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>hand</i>	<i>nichts</i>	<i>großes ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>nichts</i>	<i>kalt</i>	<i>großes ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>holz</i>	<i>kalt</i>	<i>großes ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>metall</i>	<i>kalt</i>	<i>großes ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>hand</i>	<i>kalt</i>	<i>großes ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>nichts</i>	<i>warm</i>	<i>großes ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>holz</i>	<i>warm</i>	<i>großes ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>metall</i>	<i>warm</i>	<i>großes ϵ_r</i>	<i>hand</i>	0	0	0	1
<i>hand</i>	<i>warm</i>	<i>großes ϵ_r</i>	<i>hand</i>	0	0	0	1

Tabelle F.4: Wahrscheinlichkeitstabelle $P(LAM_L1|KAM,V2H,IR_L,KAP_L1)$

Die bedingten Übergangswahrscheinlichkeiten für den Knoten *Lamelle 2 links* sind in Tabelle F.5 aufgelistet. Die bedingten Übergangswahrscheinlichkeiten für die Knoten *Lamelle 3 links*, *Lamelle 2 rechts* und *Lamelle 3 rechts* sind identisch zu den bedingten Übergangswahrscheinlichkeit des Knoten *Lamelle 2 links*.

<i>LAM_L1</i>	<i>KAP_L2</i>	<i>LAM_L2</i>			
		<i>nichts</i>	<i>holz</i>	<i>metall</i>	<i>hand</i>
<i>nichts</i>	<i>nichts</i>	1	0	0	0
<i>holz</i>	<i>nichts</i>	0,95	0,05	0	0
<i>metall</i>	<i>nichts</i>	0,95	0	0,05	0
<i>hand</i>	<i>nichts</i>	0,95	0	0	0,05
<i>nichts</i>	<i>kleines ϵ_r</i>	0,1	0,45	0,45	0
<i>holz</i>	<i>kleines ϵ_r</i>	0	1	0	0
<i>metall</i>	<i>kleines ϵ_r</i>	0	0	1	0
<i>hand</i>	<i>kleines ϵ_r</i>	0	0	0,2	0,8
<i>nichts</i>	<i>großes ϵ_r</i>	0	0	0,9	0,1
<i>holz</i>	<i>großes ϵ_r</i>	0	0	0,9	0,1
<i>metall</i>	<i>großes ϵ_r</i>	0	0	0,9	0,1
<i>hand</i>	<i>großes ϵ_r</i>	0	0	0	1

Tabelle F.5: Übergangswahrscheinlichkeitstabelle $P(LAM_L2|LAM_L1,KAP_L2)$

In Tabelle F.6 auf der nächsten Seite findet sich die bedingte Wahrscheinlichkeitstabelle des Knoten *Lamelle 1 rechts* in Abhängigkeit seiner Elternknoten *Vorfeld 2 hinten*, *IR rechts* und *Kap 1 rechts*.

<i>V2H</i>	<i>IR_R</i>	<i>KAP_R1</i>	<i>LAM_R1</i>			
			<i>nichts</i>	<i>holz</i>	<i>metall</i>	<i>hand</i>
<i>nichts</i>	<i>nichts</i>	<i>nichts</i>	1	0	0	0
<i>holz</i>	<i>nichts</i>	<i>nichts</i>	0,9	0,1	0	0
<i>metall</i>	<i>nichts</i>	<i>nichts</i>	0,9	0	0,1	0
<i>hand</i>	<i>nichts</i>	<i>nichts</i>	0,9	0	0	0,1
<i>nichts</i>	<i>kalt</i>	<i>nichts</i>	0,9	0,1	0	0
<i>holz</i>	<i>kalt</i>	<i>nichts</i>	0,9	0,1	0	0
<i>metall</i>	<i>kalt</i>	<i>nichts</i>	0,9	0	0,1	0
<i>hand</i>	<i>kalt</i>	<i>nichts</i>	0,8	0	0	0,2
<i>nichts</i>	<i>warm</i>	<i>nichts</i>	0,9	0	0	0,1
<i>holz</i>	<i>warm</i>	<i>nichts</i>	0,6	0,3	0	0,1
<i>metall</i>	<i>warm</i>	<i>nichts</i>	0,8	0	0,1	0,1
<i>hand</i>	<i>warm</i>	<i>nichts</i>	0,8	0	0	0,2
<i>nichts</i>	<i>nichts</i>	<i>kleines ϵ_r</i>	0,3	0,7	0	0
<i>holz</i>	<i>nichts</i>	<i>kleines ϵ_r</i>	0,1	0,9	0	0
<i>metall</i>	<i>nichts</i>	<i>kleines ϵ_r</i>	0,1	0	0,9	0
<i>hand</i>	<i>nichts</i>	<i>kleines ϵ_r</i>	0,1	0,4	0,4	0,1
<i>nichts</i>	<i>kalt</i>	<i>kleines ϵ_r</i>	0,1	0,45	0,45	0
<i>holz</i>	<i>kalt</i>	<i>kleines ϵ_r</i>	0	1	0	0
<i>metall</i>	<i>kalt</i>	<i>kleines ϵ_r</i>	0	0	1	0
<i>hand</i>	<i>kalt</i>	<i>kleines ϵ_r</i>	0	0	0,7	0,3
<i>nichts</i>	<i>warm</i>	<i>kleines ϵ_r</i>	0,1	0,1	0,1	0,7
<i>holz</i>	<i>warm</i>	<i>kleines ϵ_r</i>	0	0,1	0	0,9
<i>metall</i>	<i>warm</i>	<i>kleines ϵ_r</i>	0	0	0	1
<i>hand</i>	<i>warm</i>	<i>kleines ϵ_r</i>	0	0	0	1
<i>nichts</i>	<i>nichts</i>	<i>großes ϵ_r</i>	0	0	0,95	0,05
<i>holz</i>	<i>nichts</i>	<i>großes ϵ_r</i>	0	0	0,95	0,05
<i>metall</i>	<i>nichts</i>	<i>großes ϵ_r</i>	0	0	1	0
<i>hand</i>	<i>nichts</i>	<i>großes ϵ_r</i>	0	0	1	0
<i>nichts</i>	<i>kalt</i>	<i>großes ϵ_r</i>	0	0	0,9	0,1
<i>holz</i>	<i>kalt</i>	<i>großes ϵ_r</i>	0	0,2	0,6	0,2
<i>metall</i>	<i>kalt</i>	<i>großes ϵ_r</i>	0	0	1	0
<i>hand</i>	<i>kalt</i>	<i>großes ϵ_r</i>	0	0	0,7	0,3
<i>nichts</i>	<i>warm</i>	<i>großes ϵ_r</i>	0	0	0	1
<i>holz</i>	<i>warm</i>	<i>großes ϵ_r</i>	0	0	0	1
<i>metall</i>	<i>warm</i>	<i>großes ϵ_r</i>	0	0	0	1
<i>hand</i>	<i>warm</i>	<i>großes ϵ_r</i>	0	0	0	1

Tabelle F.6: Bedingte Wahrscheinlichkeitstabelle $P(LAM_R1|V2H,IR_R,KAP_R1)$

Literaturverzeichnis

- [1] ANALOG DEVICES: *Datenblatt: Linear Output Magnetic Field Sensor (AD22151)*, 2003.
- [2] ANALOG DEVICES: *Datenblatt: 24-bit Capacitance to Digital Converter (AD7747)*, 2005.
- [3] BASLER AG, Ahrensburg: *User's manual Basler A600f*, März 2004.
- [4] BAUMANN, J.: *Einspritzmengenkorrektur in Common-Rail-Systemen mit Hilfe magnetoelastischer Drucksensoren*. Dissertation, Universität Karlsruhe (TH), 2006.
- [5] BAYES, T.: *An Essay Towards Solving a Problem in the Doctrine of Chance*. In: *Philosophical Transactions of the Royal Society*, Bd. 53, S. 370–418, 1763.
- [6] BEDWORTH, M., J. O'BRIEN und M. JEMITY: *The Omnibus model: a new model of data fusion?*. In: *IEEE Aerospace and Electronic Systems Magazine*, Bd. 15, S. 30 – 36, April 2000.
- [7] BERUFGENOSSENSCHAFTLICHE ZENTRALE FÜR SICHERHEIT UND GESUNDHEIT (BGZ), Sankt Augustin: *BGR 500 - Betreiben von Arbeitsmitteln*, Oktober 2006.
- [8] BERUFGENOSSENSCHAFTLICHES INSTITUT FÜR ARBEITSSCHUTZ UND FACHBEREICH ANGEWANDTE INFORMATIK DER FH BONN-RHEIN-SIEG: *Berührungslos wirkende Schutzeinrichtung zur Fingererkennung an Kreis-sägen (Projekt-Nr. BGIA0073)*. Sankt Augustin, 2005.
- [9] BEYERER, J. (Hrsg.): *Informationsfusion in der Mess- und Sensortechnik*. Universitätsverlag Karlsruhe, 2006.
- [10] BLACK, M. J. und A. D. JEPSON: *A Probabilistic Framework for Matching Temporal Trajectories: CONDENSATION-Based Recognition of Gestures and Expressions*. Lecture Notes In Computer Science; Proceedings of the 5th European Conference on Computer Vision-Volume I, 1406:909–924, 1998.
- [11] BLAKE, A. und M. ISARD: *Active contours*. Springer-Verlag, 1998.

- [12] BLATTER, C.: *Wavelets - Eine Einführung*. Vieweg Verlagsgesellschaft mbH, 1998.
- [13] BRAMMER, K. und G. SIFFLING: *Kalman-Bucy-Filter: Deterministische Beobachtung und stochastische Filterung*. Oldenbourg Verlag, 4. verb. Aufl., 1994.
- [14] BRONSTEJN, I. N., K. A. SEMENDJAJEW, G. MUSIOL und H. MÜHLIG: *Taschenbuch der Mathematik*. Verlag Harri Deutsch, 3. überarb. und erw. Aufl., 1997.
- [15] BURT, P. J. und E. H. ADELSON: *The Laplacian Pyramid as a Compact Image Code*. In: *IEEE Transactions on Communications*, Bd. 31, April 1983.
- [16] CAMPOS, T. E. DE: *3D Hand and Object Tracking for Intention Recognition*. Dissertation, University of Oxford, 2003.
- [17] DELLAERT, F.: *The Expectation Maximization Algorithm*. Techn. Ber. GITGVU-02-20, College of Computing, Georgia Institute of Technology, February 2002.
- [18] DEMANT, C., B. STREICHER-ABEL und P. WASZKEWITZ: *Industrielle Bildverarbeitung: Wie optische Qualitätskontrolle wirklich funktioniert*. Springer-Verlag, 2. Aufl., 2002.
- [19] DEMPSTER, A. P., N. M. LAIRD und D. B. RUBIN: *Maximum Likelihood from Incomplete Data via the EM-Algorithm*. *Journal of Royal Statistical Society. Series B*, 39(1):1–38, 1977.
- [20] DEXTER RESEARCH CENTER, INC.: *Datenblatt IR-Sensor ST60R*. Laser Components GmbH, Februar 2002.
- [21] DHAWALE, P., M. MASOODIAN und B. ROGERS: *Bare-hand 3D gesture input to interactive systems*. In: *CHINZ '06: Proceedings of the 6th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction*, S. 25–32, New York, NY, USA, 2006. ACM Press.
- [22] DIKICH, E.: *Verfahren zur automatischen Gesichtserkennung*. Dissertation, Universität Karlsruhe (TH), 2003.
- [23] DIN EN 61508-4 (VDE 0803 TEIL 4): 2002-08: *Funktionale Sicherheit elektrischer/elektronischer/programmierbar elektronischer sicherheitsbezogener Systeme - Teil 4: Begriffe und Abkürzungen*. Beuth Verlag Berlin, 2002.

- [24] FACHAUSSCHUSS „EINWIRKUNGEN UND ARBEITSBEDINGTE GESUNDHEITSGEFAHREN“ DER BG 7: *BG Regel: Natürliche und künstliche Beleuchtung von Arbeitsstätten BGR 131-1 (Teil 1)*. Hauptverband der gewerblichen Berufsgenossenschaften, Oktober 2006.
- [25] FACHAUSSCHUSS HOLZ: *Interner Untersuchungsbericht Nr.: 200120763*. HBG (Holz-Berufsgenossenschaft), Stuttgart, 2001.
- [26] GERTHSEN, C., D. MESCHÉDE und H. VOGEL: *Gerthsen Physik*. Springer Verlag, 23. überarb. Aufl., 2006.
- [27] GONZALEZ RAFAEL C., WOODS RICHARD E., E. S. L.: *Digital Image Processing*. Addison-Wesley, 1992.
- [28] GRAUEL, A.: *Neuronale Netze*. BI-Wiss.-Verl., 1992.
- [29] HALL, D. L.: *Mathematical techniques in multisensor data fusion*. Artech House, Inc., 2. Aufl., 2004.
- [30] HARDENBERG, C. VON und F. BRARD: *Bare-hand human-computer interaction*. In: *PUI '01: Proceedings of the 2001 workshop on Perceptive user interfaces*, S. 1–8, New York, NY, USA, 2001. ACM Press.
- [31] HBG (HOLZ-BERUFSGENOSSENSCHAFT), Stuttgart: *Auswertung der Risikoanalysen. Interner Untersuchungsbericht*, 2003.
- [32] HBG (HOLZ-BERUFSGENOSSENSCHAFT): *Holzbearbeitungsmaschinen, Handhabung und sicheres Arbeiten. Lehrgangsbegleitheft*, 1 2003.
- [33] HIENZ, H., K. GROBEL und G. OFFNER: *Real-time hand-arm motion analysis using a single video camera*. In: *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, S. 323 – 327, 1996.
- [34] HONEYWELL: *Datenblatt: 1-Axis Magnetic Sensor (HMC1001)*, 2000.
- [35] HORNEGGER, J. M.: *Statistische Modellierung, Klassifikation und Lokalisation von Objekten*. Dissertation, Universität Erlangen-Nürnberg, 1996.
- [36] HOUGH, P. V. C.: *Method and means for recognizing complex patterns*. U.S. Patent 3.069.654, Dezember 1962.
- [37] IFM ELECTRONIC GMBH, Essen: *Schulungsunterlagen: Kapazitive Näherungsschalter und Füllstandssensoren*, 2003.
- [38] ILIEV, B. und G. MEIJER: *An impedance-measurement system for electrical characterization of rockwool substrates*. Proceedings of the ISA/IEEE Sensors for Industry Conference, SIcon/04:46–49, 2004.

- [39] INNONET -FÖRDERUNG VON INOVATIVEN NETZEN, VDI-VDE-IT, BERLIN: *Entwicklung eines Sicherheitssystems für Formatkreissägen (Cut-Stop)*. Projektbeschreibung, 2005.
- [40] INTEL CORPORATION: *Open Source Computer Vision Library*, 2006. <http://www.intel.com/technology/computing/opencv>.
- [41] ISARD, M. A.: *Visual Motion Analysis by Probabilistic Propagation of Conditional Density*. Dissertation, Oxford University, 1998.
- [42] JASPARD, F., M. NADI und A. ROUANE: *Dielectric properties of blood: an investigation of haematocrit dependence*. *Physiological Measurement*, 24:137–147, 2003.
- [43] JENSEN, F. V.: *An introduction to Bayesian networks*. UCL Press, London, 2000.
- [44] JENSEN, F. V.: *Bayesian Networks and Decision Graphs*. Springer Verlag, New York, Inc., 2001.
- [45] JÄHNE, B.: *Digitale Bildverarbeitung*. Springer-Verlag, 6. überarb. und erw. Aufl., 2005.
- [46] JÄHNE, B., H. HAUSSECKER und P. GEISSLER: *Handbook of computer vision and applications*, Bd. 1: Sensors and Imaging. Academic Press, 1999.
- [47] JÄHNE, B., H. HAUSSECKER und P. GEISSLER: *Handbook of computer vision and applications*, Bd. 2: Signal Processing and Pattern Recognition. Academic Press, 1999.
- [48] JONDRAL, F.: *Grundlagen der Wahrscheinlichkeitsrechnung und stochastischer Prozesse für Ingenieure*. Teubner Verlag, 2000.
- [49] KAKUMANUA, P., M. S. und N. BOURBAKIS: *A survey of skin-color modeling and detection methods*. In: *Pattern Recognition*, Bd. 40, S. 1106–1122, 2007.
- [50] KASS, M., A. WITKIN und D. TERZOPOULOS: *Snakes: Active contour models*. *International Journal of Computer Vision*, 1(4):321–331, Januar 1988.
- [51] KIENCKE, U.: *Ereignisdiskrete Systeme*. Oldenbourg Verlag, 2. überarb. u. erw. Aufl., 2006.
- [52] KIENCKE, U. und R. EGER: *Messtechnik*. Springer Verlag, 6. durchges. u. korr. Aufl., 2005.

-
- [53] KIENDL, H.: *Fuzzy Control methodenorientiert*. Oldenbourg Verlag, 1997.
- [54] KLAUS, F.: *Einführung in Techniken und Methoden der Multisensor-Datenfusion*. Habilitationsschrift, Universität Siegen, 2003.
- [55] KLEIN, L. A.: *Sensor and data fusion: a tool for information assessment and decision making*. SPIE Press, 2004.
- [56] KOCH, K.-R.: *Einführung in die Bayes-Statistik*. Springer Verlag, 2000.
- [57] KRAMER, F.: *Passive Sicherheit von Kraftfahrzeugen*. Vieweg Verlag, 1998.
- [58] LEHNER, G.: *Elektromagnetische Feldtheorie für Ingenieure und Physiker*. Springer Verlag, 5. Aufl., 2006.
- [59] LETESSIER, J. und F. BERARD: *Visual tracking of bare fingers for interactive surfaces*. In: *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*, S. 119–122, New York, NY, USA, 2004. ACM Press.
- [60] LIEDTKE, R. J.: *Principles of Bioelectrical Impedance Analysis*. Techn. Ber., RJL Systems, Inc., 1997.
- [61] LIU, N. und B. LOVELL: *Hand Gesture Extraction by Active Shape Models*. In: *Proceedings of Digital Image Computing: Techniques and Applications (DICTA 2005)*, S. 59 – 64, December 2005.
- [62] LUCAS, B. und T. KANADE: *An iterative image registration technique with an application to stereo vision*. In: *DARPA Image Understanding Workshop*, S. 121–130, 1981.
- [63] LUO, R. und M. KAY: *Multisensor integration and fusion in intelligent systems*. IEEE Transactions on Systems, Man and Cybernetics, 19:901 – 931, Sept.-Oct. 1989.
- [64] MAXXVISION: *Firmen-Homepage: <http://www.maxxvision.com/>*, 2007.
- [65] MENZ, B.: *Modellbasierte Objekterkennung mit 3-dimensionalen Ultraschall-Arrays*. Dissertation, Universität Karlsruhe (TH), 1998.
- [66] MÜLLER, R.: *Detektion bewegter Objekte durch Frequenzanalyse der Bildfolge*. Diplomarbeit, Institut für Prozeßrechner-technik und Robotik, Universität Karlsruhe (TH), 1993.
- [67] NEAPOLITAN, R. E.: *Learning Bayesian networks*. Pearson Prentice Hall, Upper Saddle River, NJ, 2004.

- [68] NIEBERIE, J. und S. HAUER: *Aktives Sicherheitssystem macht Tischkreissägen "kindersicher"*. In: *Die Industrie der Steine und Erde*, Bd. 4, 1998.
- [69] NIXON, M. S. und A. S. AGUADO: *Feature Extraction and image processing*. Newnes, 2002.
- [70] NÖLKER, C.: *Ein System zur Visuellen Erkennung von Handposturen*. Dissertation, Technische Fakultät, Universität Bielefeld, 2000.
- [71] PEARL, J.: *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann Publishers, Inc., 2. Aufl., 1997.
- [72] PEARL, J.: *Causality : models, reasoning, and inference*. Cambridge Univ. Press, 2001.
- [73] PIERGA, W., T. SIWEK und A. SHISHEGAR: *Sicherheitsdetektions- und Schutzsystem für kraftangetriebene Werkzeuge*. EP000001622748A2, 2006.
- [74] SCHARR, H.: *Optimale Operatoren in der Digitalen Bildverarbeitung*. Dissertation, Universität Heidelberg, 2000.
- [75] SCHAUMBURG, H.: *Sensoren*. Teubner Verlag, 1992.
- [76] SCHMIDT, R. F. (Hrsg.): *Physiologie des Menschen*. Springer Verlag, 28. korrigierte und aktualisierte Aufl., 2000.
- [77] SCHUMACHER, K.: *Endbericht: Entwicklung eines Sicherheitssystems für Formatkreissägen - Cut-Stop (Teilvorhaben: Entwicklung einer Bremse für das rotierende Werkzeug)*. Techn. Ber., Universität Stuttgart, Institut für Werkzeugmaschinen, 2006.
- [78] SHAFER, G.: *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [79] SHAN, C., Y. WEI, T. TAN und F. OJARDIAS: *Real Time Hand Tracking by Combining Particle Filtering and Mean Shift*. In: *Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition*, S. 669 – 674, 2004.
- [80] SOILLE, P.: *Morphologische Bildverarbeitung*. Springer Verlag, Heidelberg, 1998.
- [81] SPICKERMANN, D.: *Werkstoffe der Elektrotechnik und Elektronik*. Schlembach Verlag, 2002.

-
- [82] STENGER, B. D. R.: *Model-Based Hand Tracking Using A Hierarchical Bayesian Filter*. Dissertation, University of Cambridge, 2004.
- [83] STEPHEN, G. und A. JOHNSTON: *Brake positioning system*. Patent: US000006857345 B2, 2001.
- [84] TIETZE, U. und C. SCHENK: *Halbleiter-Schaltungstechnik*. Springer-Verlag, 12. Aufl., 2002.
- [85] TRÖGER, J. und K. SCHUMACHER: *Untersuchung zur Anwendbarkeit des SawStop-Systems für Formatkreissägen*. Universität Stuttgart, IfW, Stuttgart, 2003.
- [86] TRÄNKLER, H.-R. (Hrsg.): *Sensortechnik*. Springer-Verlag, 1998.
- [87] UENOHARA, M. und T. KANADE: *Use of Fourier and Karhunen-Lóeve Decomposition for Fast Pattern Matching With a Large Set of Templates*. In: *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Bd. 19, August 1997.
- [88] ULRICH, M.: *Hierarchical Real-Time Recognition of Compound Objects in Images*. Dissertation, Technische Universität München, 2003.
- [89] VARSHNEY, P. K.: *Distributed Detection and Data Fusion*. Springer Verlag, 1997.
- [90] VDMA (VERBAND DEUTSCHER MASCHINEN- UND ANLAGENBAU E.V.): *Industrielle Bildverarbeitung 2007/2008: Anwendungen - Produkte - Bezugsquellen*, 12 2006.
- [91] WAHL, F. M.: *Digitale Bildsignalverarbeitung*. Springer Verlag, 1989.
- [92] WENDEL, J.: *Integrierte Navigationssysteme: Sensordatenfusion, GPS und Inertiale Navigation*. Oldenburg Verlag, 2007.
- [93] WOJEK, C., K. NICKEL und R. STIEFELHAGEN: *Activity Recognition and Room-Level Tracking in an Office Environment*. In: *Proceedings of the International Conference on Multisensor Fusion and Integration for Intelligent Systems*, S. 25–30. IEEE Press, 2006.
- [94] YÖRÜK, E., E. KONUKOGLU, B. SANKUR und J. DARBON: *Shape-Based Hand Recognition*. *IEEE Transactions on Image Processing*, 15:1803 – 1815, 7 2006.

Patente und Patentanmeldungen

- [95] BARRHO, J., K. SCHUMACHER und BETEILIGTE FIRMEN: WILHELM ALTENDORF GMBH & CO. KG , OTTO MARTIN MASCHINENBAU GMBH & CO. KG: *Nothaltevorrichtung für Kreissägeblatt*. Gebrauchsmuster angemeldet 2006.
- [96] MATSUNAGA, T., M. HIEMER und J. BARRHO:
Fahrzeugunfallanalysevorrichtung. DE102005012119B4, 2006.
Accident analysis device for vehicle. JP002006131092AA, 2006 (Offenlegung).
Vehicle accident analyzing apparatus. US000007162345B2, 2006.
- [97] MATSUNAGA, T., M. HIEMER und J. BARRHO:
System for determining road surface slope. JP002006133076AA, 2006 (Offenlegung).
Vehicle accident analyzing apparatus. US020060100767A1, 2006 (Offenlegung).

Eigene Veröffentlichungen

- [98] BARRHO, J.: *Abschlussbericht: Entwicklung eines Sicherheitssystems für spanende Werkzeugmaschinen - Cut-Stop (Teilvorhaben: Erkennung von Gefahrensituationen)*. Techn. Ber., Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2006.
- [99] BARRHO, J.: *Object Segmentation in Image Sequences*. In: KIENCKE, U. und K. DOSTERT (Hrsg.): *Reports on Industrial Information Technology*, Bd. 10, S. 97–102, Aachen, 2007. Shaker Verlag.
- [100] BARRHO, J., M. ADAM und U. KIENCKE: *Finger Localization and Classification in Images based on Generalized Hough Transform and Probabilistic Models*. In: *9th International Conference on Control, Automation, Robotics and Vision, Singapore*, 2006.
- [101] BARRHO, J., M. ADAM und U. KIENCKE: *Konzeption eines Praktikumsversuches zur Verarbeitung von Bildsequenzdaten*. In: *Virtuelle Instrumente in der Praxis (VIP 2006)*, S. 348 – 351, Heidelberg, München, 2006. Hüthig Verlag.
- [102] BARRHO, J. und J. BAUMANN: *Concepts of Establishing Homogeneous Illumination in Machine Tool Working Areas Considering Image Processing and Safety Principles*. In: KIENCKE, U. und K. DOSTERT (Hrsg.):

- Reports on Industrial Information Technology*, Bd. 8, S. 73–81, Aachen, 2005. Shaker Verlag.
- [103] BARRHO, J. und J. BAUMANN: *Feature Extraction Using Transform Based Methods*. In: KIENCKE, U. und K. DOSTERT (Hrsg.): *Reports on Industrial Information Technology*, Bd. 9, S. 93–101, Aachen, 2006. Shaker Verlag.
- [104] BARRHO, J., J. HAUGER und U. KIENCKE: *Information Fusion for Hazard Analysis in Multi-Sensor Systems based on Bayesian Networks*. In: *Proceedings of the World Congress on Engineering 2007*. IAENG, 2007.
- [105] BARRHO, J. und E. HEINZE: *Multi Sensor Data Fusion for Hazard Analysis using Voting Logic*. In: KIENCKE, U. und K. DOSTERT (Hrsg.): *Reports on Industrial Information Technology*, Bd. 10, S. 87–96, Aachen, 2007. Shaker Verlag.
- [106] BARRHO, J. und M. HIEMER: *Integrated Safety Systems for Circular Saws: Concepts for Detection Systems of Human Limbs*. In: KIENCKE, U. und K. DOSTERT (Hrsg.): *Reports on Industrial Information Technology*, Bd. 7, S. 73–81, Aachen, 2004. Shaker Verlag.
- [107] BARRHO, J. und M. HIEMER: *Wheel Speed Signal Analysis to Assess Operativeness of Anti-Lock-Braking Systems (ABS)*. In: KIENCKE, U. und K. DOSTERT (Hrsg.): *Reports on Industrial Information Technology*, Bd. 8, S. 27–36, Aachen, 2005. Shaker Verlag.
- [108] BARRHO, J., M. HIEMER, U. KIENCKE und T. MATSUNAGA: *Estimation of Elevation Difference based on Vehicle's Inertial Sensors*. In: *16th IFAC World Congress, Prague, Czech Republic*, 2005.
- [109] BARRHO, J., J. KALLENBERG und U. KIENCKE: *Visual Tracking of Human Hands for a Hazard Analysis based on Particle Filtering*. In: *11th IPMU International Conference, Paris, France*, 2006.
- [110] BARRHO, J. und U. KIENCKE: *Neues Sicherheitskonzept für Formatkreissägen (Teil 2)*. HOB - Die Holzbearbeitung, 3:70–73, 2007.
- [111] BARRHO, J., U. KIENCKE, K. SCHUMACHER und J. TRÖGER: *Non-Contact Working and Non-Interfering Safety System for Sliding Table Saws*. In: *6th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, Beijing, PRC*, 2006.
- [112] BARRHO, J. und V. SEEMANN: *An Ultrasonic Sensor System for Object Localization*. In: KIENCKE, U. und K. DOSTERT (Hrsg.): *Reports on*

- Industrial Information Technology*, Bd. 9, S. 103 – 107, Aachen, 2006. Shaker Verlag.
- [113] BARRHO, J., V. SEEMANN und U. KIENCKE: *Entwicklung einer Ultraschallsensorik zur Positionsbestimmung von Objekten an Formatkreissägen*. In: *Virtuelle Instrumente in der Praxis (VIP 2006)*, S. 81– 84, Heidelberg, München, 2006. Hüthig Verlag.
- [114] HIEMER, M. und J. BARRHO: *Determination of Crash Parameters for Low Speed Crash Reconstruction*. In: KIENCKE, U. und K. DOSTERT (Hrsg.): *Reports on Industrial Information Technology*, Bd. 7, S. 13–21, Aachen, 2004. Shaker Verlag.
- [115] HIEMER, M. und J. BARRHO: *Observer Design for Road Gradient Estimation*. In: KIENCKE, U. und K. DOSTERT (Hrsg.): *Reports on Industrial Information Technology*, Bd. 7, S. 23–30, Aachen, 2004. Shaker Verlag.

Betreute Diplom- und Studienarbeiten

- [116] ADAM, M.: *Konzeption und Realisierung eines Versuches zur Verarbeitung von Bildsequenzdaten einer Kamera*. Studienarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2005.
- [117] ADAM, M.: *Merkmalsbasierte Modellierung, Lokalisation und Klassifikation von komplexen Objekten*. Diplomarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2006.
- [118] BECKER, J.: *Erweiterung einer Bewegungsanalyse in Videosequenzen durch Object-Tracking mittels Aktiver Konturen*. Studienarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2006.
- [119] BECKER, J.: *Shadow Detection and Overlapping Shape Recovery in Image Sequences*. Diplomarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH) in Kooperation mit Denso IT Laboratory, Japan, 2007.
- [120] DEISSLER, T.: *Extraktion und Lokalisation komplexer geometrischer Strukturen*. Studienarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2005.
- [121] DEISSLER, T.: *Bilddatensegmentierung unter Berücksichtigung anthropogener Strukturelemente am Beispiel von Satellitenbildern*. Diplomarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH) in Kooperation mit FGAN-FOM, Ettlingen, 2006.

-
- [122] GEHRING, P.: *Materialidentifikation basierend auf dem Wirbelstromeffekt*. Studienarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH) und HS Karlsruhe, 2006.
- [123] HARTMANN, J.: *Untersuchung und Implementierung von Farbsegmentierungsverfahren zur Merkmalsgewinnung aus Einzelbildern*. Studienarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2004.
- [124] HAUGER, J.: *Konzeption und Implementierung eines Bayes'schen Netzes zur Informationsfusion*. Diplomarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2006.
- [125] HEINZE, E.: *Entwicklung einer redundanten, berührungslos arbeitenden Sensorik zur Gefahrenerkennung mittels Datenfusion*. Diplomarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2006.
- [126] HUA, R.: *Entwicklung eines Positionssystems für Schutzlamellen*. Diplomarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2006.
- [127] IMPELLIZZERI, A.: *Untersuchung von transformationsbasierten Methoden zur Extraktion von Bildmerkmalen*. Studienarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2005.
- [128] IMPELLIZZERI, A.: *Optimierung der Merkmalsextraktion für ein probabilistisches Lokalisations- und Klassifikationsverfahren*. Diplomarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2007.
- [129] KALLENBERG, J.: *Bewegungsanalyse von Objekten im Gefahrenbereich einer Werkzeugmaschine*. Studienarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2005.
- [130] KALLENBERG, J.: *Optimierung und Erweiterung einer Bewegungsanalyse basierend auf probabilistischen Methoden*. Diplomarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2007.
- [131] KLEIN, T.: *Bewegungsdetektion in realen Bildsequenzen zur Gefahrenerkennung*. Studienarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2006.
- [132] ROSENAU, M.: *Bilderfassung und -auswertung zur Überwachung von Gefahrenbereichen bei Werkzeugmaschinen*. Diplomarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2004.

- [133] SEEMANN, V.: *Positionsbestimmung mittels Ultraschall-Sensorik*. Studienarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2005.
- [134] WURM, C.: *Entwurf und Simulation einer auf Photometrie basierenden Positionsschätzung*. Diplomarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH) in Kooperation mit Mechaless GmbH, Karlsruhe, 2005.
- [135] ZHANG, L.: *Autonomes Lernen kontextabhängiger Handstellungen am Beispiel des Zeigens auf und des Anforderns von Objekten*. Diplomarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH) in Kooperation mit Honda Research Institute Europe, Offenbach, 2007.
- [136] ZHANG, X.: *Objektsegmentierung und -lokalisierung mittels Korrelation*. Studienarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2006.
- [137] ZHU, Y.: *Verbesserung der Bildqualität und Implementierung von Verfahren zur Vorder-/Hintergrundschtzung*. Studienarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2005.
- [138] ZHU, Y.: *Entwicklung und Aufbau eines μC -Steuergeräts für ein Sicherheitssystem*. Masterarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2006.
- [139] ZUMBÜLTE, M.: *Bewegungserkennung in Bildfolgen zur Merkmalsgewinnung*. Studienarbeit, Institut für Industrielle Informationstechnik, Universität Karlsruhe (TH), 2005.

Index

A

Abstraktionsebene 121
Aktive Kontur 84
Aktorik 167
Alpha-Segmentierer 54
Azyklischer gerichteter Graph ... 127

B

Bayes Statistik 122
Bayes'sches Netz 125
Bayes'sches Theorem 126
Bayes-Tracker 93
Berührungslose Handerkennung .. 15,
43
Bildanalyse 102
Bildmerkmal 75
Bildverarbeitungskette 44
Bildvorverarbeitung 62
Binominalfilter 60

C

Clustering 70
Condensation Algorithmus 97
 Beobachtungsmodell 100
 Bewegungsmodell 99

D

d-Separation 128
Deterministischer Drift 97
Differenzbildverfahren 51
Differenzielle Kantendetektion 61

F

Farbsegmentierung 58
Farbsensor 18

Flimmern 48
Formatkreissäge 1
Fusionsarchitektur 121

G

Gauß Filter 60
Gefährdungssituation 3
Gefährdungszone 3, 161
Gefahrenmerkmal 159
Gefahrenzone 160

H

Hall-Effekt 35
Hypothesenknoten 132

I

ICondensation Algorithmus 99
Induktiver Sensor 36
Informationsfusion 117
Informationsknoten 132
Infrarot-Sensor 19
 Korrektur 135

J

JDL-Modell 118

K

Kamera 46
Kapazitätsmessung 30
Kapazitiver Sensor 24
 Hauptfeldänderung 26
 Korrektur (Lamellen) 137
 Korrektur (Vorfeld) 136
 Streifelfeldänderung 28
Klassifikation 81
Kreis-Hough-Transformation 67

L

Lichtquelle 48
Lokalisation 80

M

Marginalisierung 126
Merkmalskurve 90
Modellmerkmal 73
Musterkurve 90

N

Normal Displacement 90
Normierung von Sensorsignalen .. 134

O

Omnibus-Modell 119

P

Parametrische Modelldichte 77
Partikelfilter 94

S

Scharr-Operator 62
Schiebestock 5
Schutzhaube 5
Schutzlamelle 29, 162
 Positionsbestimmung 35
Seebeck-Effekt 19
Sensorintegration 121
Shape-Space 88
Situationsmerkmal 159
Sobel-Operator 62
Spaltkeil 5
Stochastische Diffusion 97

T

Thermopiler Sensor
 siehe Infrarot-Sensor
Tolerierbares Risiko 4
Tracking 93

U

Überwachungsbereich 3, 161
Ultraschall-Sensor 16

V

Varianzmethode 81
Vorfeld 1 161
Vorfeld 2 161

W

Wirbelstromsensor 36

Z

Zustandswahrscheinlich-
 keitsfunktion 144

ISBN: 978-3-86644-156-9

www.uvka.de