

Innovative Sensorkonzepte und Signalverarbeitungsstrategien zur Bewegungserkennung und Präsenzkontrolle von Personen

Vom Fachbereich Elektrotechnik und Informatik der
Universität Siegen
zur Erlangung des akademischen Grades

Doktor-Ingenieur (Dr.-Ing.)

genehmigte Dissertation

vorgelegt von

Dipl.-Ing. Thomas P. U. Krieger
geboren am 26.09.1966 in Kirchen/Sieg

1. Gutachter: Prof. Dr.-Ing. habil. Otmar Loffeld
2. Gutachter: Prof. Dr.-Ing. Rudolf Schwarte

Vorsitzender der Prüfungskommission: Prof. Dr.-Ing. Hubert Roth

Tag der mündlichen Prüfung: 11. Juni 2002

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter im Zentrum für Sensorsysteme (ZESS) der Universität Siegen.

Mein besonderer Dank gilt meinem Doktorvater Herrn Prof. Dr.-Ing. Otmar Loffeld für die Anregung und wissenschaftliche Betreuung meiner Arbeit und für die Geduld, die er mir in der Endphase meiner Promotion entgegengebracht hat.

Herrn Prof. Dr.-Ing. Rudolf Schwarte, dem Leiter des Instituts für Nachrichtenverarbeitung (INV) der Universität Siegen, danke ich für die bereitwillige Übernahme des Korreferats, die kritische Durchsicht meiner Arbeit und die sich daraus ergebenden Anregungen. Ich werde die zahlreichen Stunden in seinem Büro, in denen er mir im Vertrauen auch seine neuesten Erfindungen erläuterte, und die gute Zusammenarbeit in der S-TEC nicht vergessen.

Bei Herrn Prof. Dr.-Ing. Hubert Roth bedanke ich mich für die Leitung und Durchführung des Promotionsverfahrens.

Ferner danke ich allen Kollegen, den wissenschaftlichen und nicht-wissenschaftlichen Mitarbeitern des ZESS und INV sowie den Mitarbeitern der Firmen Aicoss GmbH, Asentics GmbH, INSTA Elektro GmbH, S-TEC GmbH und PMD-Technologies GmbH für die gute Zusammenarbeit der letzten Jahre und ihre Unterstützung bei der Durchführung dieser Arbeit. An dieser Stelle sollen auch die Studenten, die durch Studien- und Diplomarbeiten sowie als Praktikanten und studentische Hilfskräfte zum Gelingen beigetragen haben, nicht vergessen werden.

Besonderen Dank schulde ich weiterhin Herrn Dr.-Ing. Klaus Hartmann, dem Geschäftsführer des Zentrums für Sensorsysteme. Er hat mir als Berufseinsteiger eine Chance gegeben. Durch seinen fachlichen Rat bei der Durchführung von Industrieprojekten, auch im Hinblick auf die Promotion, hat er einen wesentlichen Teil zu meiner beruflichen Entwicklung beigetragen. Es war nie langweilig im ZESS und ich hoffe, er hat es nicht bereut.

Mein ganz besonderer Dank gilt meinen Eltern Rudolf und Reinhilde Krieger. Sie haben mich auf den richtigen Weg gebracht. Nach dem Tod meines Vaters hat meine Mutter viele Mühen und Entbehren auf sich genommen, um mir das Studium zu ermöglichen. Das war der entscheidende Grundstein für meinen bisherigen Lebensweg.

Ich möchte diese Arbeit meiner Frau Brigitte und meinen beiden Kindern Mario André und Louisa Christin widmen, die davon in vielerlei Hinsicht am unmittelbarsten betroffen waren. Es war für uns alle eine entbehrungsreiche Zeit. Ihr Verständnis, ihre Rücksicht und der Rückhalt in meiner Familie hat mir sehr geholfen.

Siegen, im März 2002

Thomas Krieger

Abstract

Motion observers and presence controllers are essential parts of modern concepts for building automation and energy management. Especially the low-cost passive infrared sensors, based on the pyroelectric effect, are suitable for reducing the overall power consumption by providing information about the presence or absence of persons. The first part of this documentation presents a review of recent and new sensor systems for motion and occupancy detection.

Because of the lower power consumption, passive sensing devices are more appropriate for energy management and building control than systems, based on an active sensor concept. However, especially in applications with a reduced field of view (e. g. automatic door opener systems) these sensors are frequently used today. With increasing field of view, for example in well known wide range indoor and outdoor surveillance applications, passive sensor concepts become dominant.

Especially in indoor applications, conventional motion observers suffer from the low resolution of commercially available pyroelectric detectors. On the other hand, CCD cameras for the visible and near infrared region of light provide extremely high resolutions, but for comparatively high cost. Newer developments in CMOS imager technology enable the design of low cost camera sensors for the consumer market.

In conjunction with a motion observer with pyroelectric detectors, advantages of both sensor principles can be combined. The result will be a highly integrated, high resolution occupancy detection system with additional features. These include for example a scene dependent light management or, in addition with suitable means for data communication, the transmission of single informations or, if necessary, of complete video frames to an external terminal.

In the scope of this work, preliminary investigations on both, the CMOS camera and the motion observer with pyroelectric detectors, are presented. Both parts can be realized as separate sensor systems or, as planned for future products, in combination. However, the image processing algorithms and the evaluation strategy for the pyroelectric sensor part, are not yet implemented on the target hardware. This will be part of further research and development at the Center of Sensor Systems (ZESS).

Many years of experience in optical sensor design at the ZESS and closer investigations on signal processing strategies for motion observers together with our industrial partners, culminated in a new superior motion observer. Especially digital signal processing algorithms reduce the susceptibility to false alarms through a very special evaluation of signal-specific characteristics. The drop in prices and the high performance of modern microcontrollers speed up the transition from recent analogue to modern digital techniques.

For future improvements, a new evaluation strategy for motion observers was developed in the scope of this work. Based on the foundations of estimation theory, an innovative signal processing concept with statistical filter is presented in the following. This filter internally uses a discrete-time linear state space model of the real world process. This leads to a linear, optimum estimation algorithm, which is called “Kalman-filter” according to its explorer.

The Kalman-filter is used as suppressor for unwanted background signals. Deviations between the predicted behaviour and the behaviour observed at the real system are used for motion detection. This procedure yields into an increased sensitivity with simultaneously improved false alarm immunity as will be shown in this work.

Inhaltsverzeichnis

Abkürzungen und Begriffserläuterungen	XIII
Konstanten	XIV
Formelzeichen und Symbole	XV
1 Einleitung	1
1.1 Motivation	1
1.2 Zielsetzung	3
1.3 Gliederung	4
2 Sensorsysteme zur Anwesenheitserkennung von Personen	6
2.1 Einleitung	6
2.2 Messgrößen zur Anwesenheitserkennung	6
2.2.1 Personenerkennung in Videobildern	8
2.2.1.1 Trennung und Adaption komplexer Hintergrundszenen	9
2.2.1.2 Modellbasierte Ansätze zur Erkennung und Verfolgung von Personen	10
2.2.2 Messtechnische Erfassung körpereigener Vitalfunktionen	13
2.3 Präsenzkontrolle durch Bewegungserkennung	14
2.3.1 Ultraschallsensoren zur Präsenzkontrolle	15
2.3.2 Mikrowellensensoren zur Bewegungserkennung	19
2.3.3 Intensitätsmessende optische Sensoren im Bereich des sichtbaren Lichts und des nahen Infrarot	23
2.3.3.1 Lichttaster als Ersatz für taktile Lichtschalter und Dimmer	25
2.3.3.2 Zusätzliche Informationen bei Verwendung von von optischen Sensoren	27
2.3.4 Photomischdetektoren zur Messung der Laufzeit des Lichts	30
2.3.4.1 Funktionsweise eines Photogate-PMD	30
2.3.4.2 PMD-Grundstrukturen und Kenndaten	32
2.3.4.3 Entfernungsmessung mit CW-Verfahren	35
2.3.4.4 Systemkonzept für PMD-basierte Entfernungsmessgeräte	37
2.3.4.5 Entfernungsmessergebnisse mit einem 1D-PMD-Sensor	38
2.3.4.6 Parallele Entfernungsmessung mit PMD-Arrays	40
2.3.5 Sensoren zur Erfassung von Wärmestrahlung	42
2.3.5.1 Thermoelemente und Thermopiles zur Anwesenheitskontrolle	44
2.3.5.2 Bewegungserkennung mit pyroelektrischen Detektoren	47

2.4	<i>Literatur</i>	49
3	Bildgebende Sensoren zur Anwesenheitserkennung	52
3.1	<i>Einleitung</i>	52
3.2	<i>Strahlungsphysikalische Grundlagen</i>	54
3.2.1	Senderseitige Kenngrößen und Definition des Raumwinkels.....	54
3.2.2	Empfängerseitige Kenngrößen.....	57
3.2.3	Wellenausbreitung und Energieübertragung.....	58
3.2.4	Geometrische Optik.....	60
3.2.5	Berücksichtigung einer Linse im Übertragungsweg.....	63
3.2.6	Strahlungsrauschen.....	64
3.3	<i>Bildgebende Sensoren</i>	65
3.3.1	Lichtempfindliche Elemente in CCD- und CMOS-Technologie.....	65
3.3.2	Kenngrößen zur Beurteilung der Sensorperformance.....	69
3.3.2.1	<i>Geometrische Daten</i>	69
3.3.2.2	<i>Pattern Noise</i>	70
3.3.2.3	<i>Dunkelstrom und Dunkelstromrauschen</i>	70
3.3.2.4	<i>Rauschquellen</i>	71
3.3.2.5	<i>Signal-Rausch-Verhältnis und Dynamikbereich</i>	72
3.3.2.6	<i>Spektrale Empfindlichkeit</i>	73
3.3.2.7	<i>Zusätzliche Eigenschaften</i>	74
3.3.3	Kamera- und Weltkoordinatensystem.....	75
3.3.4	Gesichtsfeld und Bereichsauflösung.....	77
3.3.5	Systemkonzept des Bildwächters.....	81
3.4	<i>Literatur</i>	83
4	Auswerteverfahren für bildgebende Sensoren	84
4.1	<i>Einleitung</i>	84
4.2	<i>Differenzbildverfahren zur Bewegungserkennung</i>	84
4.3	<i>Bewegungserkennung mit Zeilen- und Spaltenmittelwerten</i>	90
4.4	<i>Varianzauswertung bei konstanter Grauwertinterpolation</i>	92
4.5	<i>Varianzauswertung bei linearer Grauwertinterpolation</i>	95
4.6	<i>Matrixorientiertes Verfahren zur Grauwertinterpolation</i>	100
4.7	<i>Anwendung des t-Tests zur Bewegungsdetektion</i>	103
4.8	<i>Anwendung des F-Tests zur Bewegungsdetektion</i>	106
4.9	<i>Literatur</i>	113

5	Komponenten von PIR-Bewegungsmelder	114
5.1	<i>Pyroelektrische Detektoren</i>	114
5.1.1	Pyroelektrische Festkörper und pyroelektrisches Verhalten	114
5.1.2	Ferroelektrizität, Polarisation und pyroelektrischer Effekt	115
5.1.3	Thermisches Verhalten eines pyroelektrischen Sensors	117
5.1.4	Elektrisches Verhalten eines pyroelektrischen Sensors	119
5.1.5	Übertragungsfunktion eines Einzelementsensors	121
5.1.6	Rauschquellen bei pyroelektrischen Detektoren	122
5.1.6.1	<i>Temperaturrauschen</i>	123
5.1.6.2	<i>Dielektrisches Rauschen</i>	123
5.1.6.3	<i>Widerstandsrauschen</i>	124
5.1.6.4	<i>Verstärkerrauschen</i>	124
5.1.6.5	<i>Mikrophonie</i>	124
5.1.7	Kenngrößen von pyroelektrischen Detektoren	125
5.1.7.1	<i>Rauschäquivalente Strahlungsleistung</i>	125
5.1.7.2	<i>Detektivität und spezifische Detektivität</i>	125
5.1.8	Übertragungsverhalten eines Doppелеlementsensors	125
5.1.9	Simulationen zum Übertragungsverhalten eines Dualsensors	127
5.2	<i>Hardwarekomponenten von Bewegungsmeldern</i>	134
5.3	<i>Fresnellinsen zur Bündelung langwelliger Temperaturstrahlung</i>	138
5.4	<i>Literatur</i>	141
6	Systementwicklung von Passiv-Infrarot-Bewegungsmeldern	142
6.1	<i>Simulationstool zum Entwurf von Fresnellinsenarrays</i>	142
6.1.1	Simulation des Erfassungsfeldes	143
6.1.2	Berücksichtigung der Systemparameter und der Erfassungsfeldspezifikationen	147
6.1.2.1	<i>Sensorpositionen und Orientierung</i>	147
6.1.2.2	<i>Geometrie und Positionierung der Gesamtlinsenfläche</i>	148
6.1.3	Initialisierung und Optimierung eines Fresnellinsenarrays	149
6.1.4	Beispiele zur Erfassungsfeldsimulation	151
6.1.4.1	<i>270°-Wächter mit 3 Doppелеlementdetektoren</i>	151
6.1.4.2	<i>90°-Wächter mit einem 4x2-PIR-Array</i>	153
6.1.5	Anwendungsmöglichkeiten des Simulationstools	154
6.2	<i>Aufzeichnung und Verarbeitung von Wächtersignalen</i>	156
6.2.1	Datenaufnahme, Kommunikationsprotokoll und Dateihandling	157
6.2.2	Softwareentwicklung, Auswertung und Ergebnisausgabe	159
6.2.3	Parameteroptimierung und Beurteilung des Schaltverhaltens	160

6.3	Literatur	160
7	Konventionelle Signalverarbeitungskonzepte für PIR-Bewegungsmelder	161
7.1	<i>Signale aus unterschiedlichen Bewegungsrichtungen</i>	161
7.2	<i>Referenzpegelnachführung und Schwellenauswertung</i>	163
7.3	<i>Ermittlung typischer Signaleigenschaften</i>	166
7.4	<i>Bewertung charakteristischer Signaleigenschaften</i>	172
7.5	<i>Korrelationstechniken zur Bewegungserkennung</i>	179
7.5.1	Autokorrelation zur Erkennung von deterministischen Signalanteilen	179
7.5.2	Kreuzkorrelationsfunktion zum Vergleich von Signalen	181
7.6	<i>Literatur</i>	182
8	Modellbasierter Ansatz zur Bewegungserkennung	184
8.1	<i>Einleitung</i>	184
8.2	<i>Systemkonzept mit Kalman-Filter</i>	184
8.3	<i>Abgleichprozess zwischen Modell und realem Verhalten</i>	185
8.4	<i>Zeitkontinuierliche und zeitdiskrete Zustandsraummodelle</i>	187
8.5	<i>Modellierung von Nutz- und Störsignalen</i>	191
8.6	<i>Sensorielles Beobachtungsmodell</i>	193
8.7	<i>Allgemeine Beschreibung des Kalman-Filters</i>	194
8.8	<i>Anwendungsmöglichkeiten zur Bewegungserkennung</i>	197
8.9	<i>Literatur</i>	198
9	Entwicklung eines zeitdiskreten Systemmodells zur Signalvorverarbeitung	199
9.1	<i>Modellierung des Bewegungsmelderverhaltens</i>	199
9.1.1	Übertragungsfunktion und Stoßantwort des vereinfachten Bewegungsmeldermodells	199
9.1.2	Herleitung des zeitkontinuierlichen Zustandsraummodells	202
9.1.3	Ableitung des äquivalenten zeitdiskreten Zustandsraummodells	204
9.1.4	Berücksichtigung von Rauschprozessen	204
9.1.5	Filtergleichungen des Kalman-Filters	205
9.1.6	Simulationsergebnisse mit dem vereinfachten Bewegungsmeldermodell	206
9.1.6	Modellerweiterung des einfachen Bewegungsmeldermodells	208
9.2	<i>Allgemeines Modell zur Signalverfolgung</i>	210
9.2.1	Herleitung des zeitkontinuierlichen Zustandsraummodells	210

9.2.2	Ableitung des Systemmodells.....	213
9.2.3	Simulationsergebnisse mit dem allgemeinen Signalmodell.....	214
9.3	Allgemeines Signalmodell mit Berücksichtigung von Drifteinflüssen.....	220
9.2.1	Herleitung des erweiterten Signalmodells.....	220
9.2.2	Simulationsergebnisse mit dem erweiterten Signalmodell.....	222
10	Anwendung des Kalman-Filters zur Bewegungserkennung.....	229
10.1	Einleitung.....	229
10.2	Parametrierung der Kalman-Filters anhand von typischen Sensorsignalen.....	230
10.2.1	Kovarianz des Ausgangsrauschprozesses.....	230
10.2.2	Kovarianzmatrix des Eingangsruschprozesses.....	234
10.2.2.1	Störmodell ohne Objekt im Erfassungsfeld.....	234
10.2.2.2	Allgemeines Signalmodell ohne Objekt im Erfassungsfeld.....	235
10.2.2.3	Allgemeines Signalmodell mit Objekt im Erfassungsfeld.....	236
10.3	Bewegungserkennung anhand von internen Größen des Kalman-Filters.....	237
10.3.1	Schwellenauswertung der Residuensequenz.....	237
10.3.2	Auswertung der 1. Komponente der Zustandsprädiktion.....	240
10.3.3	Auswertung der 2. Komponente der Zustandsprädiktion.....	242
10.4	Schaltverhalten der verschiedenen Auswerteverfahren.....	243
10.4.1	Schaltverhalten bei konventioneller Signalauswertung.....	243
10.4.2	Schaltverhalten bei Auswertung der Residuensequenz.....	247
10.4.3	Bewertung der 1. Komponente der Zustandsprädiktion.....	251
10.4.4	Bewertung der 2. Komponente der Zustandsprädiktion.....	256
10.4.5	Fehlschaltverhalten bei starken Störungen.....	260
11	Zusammenfassung und Ausblick.....	263

Abkürzungen und Begriffserläuterungen

AC	Alternating Current (Wechselanteil)
ADC	<u>A</u> nalog- <u>D</u> igital- <u>C</u> onverter
AKF	Autokorrelationsfunktion
APS	<u>A</u> ctive <u>P</u> ixel <u>S</u> ensor
CAN	<u>C</u> ontroller <u>A</u> rea <u>N</u> etwork
CCD	<u>C</u> harge <u>C</u> oupled <u>D</u> evice
CDS	<u>C</u> orrelated <u>D</u> ouble <u>S</u> ampling
CMOS	<u>C</u> omplementary <u>M</u> etal <u>O</u> xide <u>S</u> emiconductor
CPLD	<u>C</u> omplex <u>P</u> rogrammable <u>L</u> ogic <u>D</u> evice
CS	<u>C</u> orrelated <u>S</u> ampling
CW	<u>C</u> ontinuous <u>W</u> ave
DC	Direct Current (Gleichanteil)
ESD	elektrostatische Kontaktentladung
EVG	Elektronisches Vorschaltgerät
FET	<u>F</u> eldeffekttransistor
FPGA	<u>F</u> ield <u>P</u> rogrammable <u>G</u> ate <u>A</u> rray
FT	<u>F</u> rame <u>T</u> ransfer
Ge	Germanium
H0	Hypothese H0: keine Bewegung erkannt
H1	Hypothese H1: Bewegung erkannt
IR	Infrarot
IR-A	Elektromagnetische Strahlung im Spektralbereich des nahen Infrarot
IR-B	Elektromagnetische Strahlung im Spektralbereich des mittleren Infrarot
IR-C	Elektromagnetische Strahlung im Spektralbereich des fernen Infrarot
IT	<u>I</u> nterline <u>T</u> ransfer
KKF	Kreuzkorrelationsfunktion
KVG	Konventionelles Vorschaltgerät
LED	<u>L</u> ight <u>E</u> mitting <u>D</u> iode
LTI	<u>L</u> inear <u>T</u> ime <u>I</u> nvariant
MC	Mikrocontroller
MOD	minimale Objektdistanz eines Objektivs
MOS	<u>M</u> etall- <u>O</u> xide- <u>S</u> emiconductor
NIR	naher Infrarot-Bereich (IR-A)

OP	Operationsverstärker
PIR	Passiv-Infrarot
PMD	<u>Photo-Misch-Detektor</u> (<u>Photonic Mixer Device</u>)
PPS	<u>Passive Pixel Sensor</u>
PRNU	<u>Photo Response Non-Uniformity</u>
SBI	Hintergrundlichtunterdrückung (<u>Suppression of Background Illumination</u>)
Si	Silizium
SNR	Signal-Rausch-Verhältnis
SoC	Systems on Chip
UART	<u>Universal Asynchronous Receiver Transmitter</u>
VIS	Elektromagnetische Strahlung im Wellenlängenbereich des sichtbaren Lichts
2D	zweidimensional
3D	dreidimensional

Konstanten

c	Lichtgeschwindigkeit ($c = 2.99792438 \cdot 10^{10}$ cm/s)
h	Plancksches Wirkungsquantum ($h = 6.626176 \cdot 10^{-34}$ Ws ²)
k	Boltzmann-Konstante ($k = 1.380662 \cdot 10^{-23}$ Ws/K)
q	Elementarladung ($q = 1.6021 \cdot 10^{-19}$ C)

Symbole und Formelzeichen

Kapitel 1

Kapitel 2

A	Fläche eines PMD-Modulationsgates bzw. eines Detektorelements
C	Modulationskontrast beim PMD
C_{int}	Integrationskapazität
D	Dämpfungsfaktor
f_m	Modulationsfrequenz
$G(x)$	Generationsrate von Ladungsträgern
I_a	mittlerer Strom auf die PMD-interne Integrationskapazität a
I_b	mittlerer Strom auf die PMD-interne Integrationskapazität b
m_E	Modulationsgrad eines Empfängers
m_S	Modulationsgrad eines Senders
$P_{\text{opt}}(t)$	modulierte Lichtleistung
$P_e(t)$	am PMD-Empfänger auftreffende Strahlungsleistung
$P_s(t)$	ausgesendete Strahlungsleistung
R	Abstand zwischen PMD-Sensor und Target
R_m	gemessener Abstand zwischen PMD-Sensor und Target
R_s	Soll-Abstand zwischen PMD-Sensor und Target
S	PMD-Empfindlichkeit
t	Zeit
$T_a(f_m, u_{\text{am}}(t))$	PMD-Transfercharakteristik, Pixelhälfte A
$T_b(f_m, u_{\text{bm}}(t))$	PMD-Transfercharakteristik, Pixelhälfte B
T_{int}	Integrationsdauer
U_a	PMD-Ausgangsspannung, Ausgang a
$u_{\text{am}}(t)$	Spannungsverlauf am Modulationsgate am
U_b	PMD-Ausgangsspannung, Ausgang v
$u_{\text{bm}}(t)$	Spannungsverlauf am Modulationsgate bm
Z_{max}	Eindeutigkeitsbereich
φ_L	Phasenverschiebung infolge der Lichtlaufzeit τ_L
φ_0	konstante Phasenverschiebung
λ_m	Modulationswellenlänge

τ_L Lichtlaufzeit

Kapitel 3

A Fläche

A_{ph} photoempfindliche Fläche eines Pixels

A_E Fläche eines Strahlungsempfängers

A_L Fläche einer Linse bzw. eines Objektivs

A_R Fläche eines reflektierenden Targets

A_S Fläche eines Strahlungssenders

d Abstand zwischen Sender und Empfänger

D Durchmesser der Eintrittspupille einer Optik

E, E_{ph} Bestrahlungsstärke

H Belichtung

I Strahlstärke

I_{dk} Dunkelstrom

I_{ph} Photostrom

k Blendenzahl

L Strahldichte eines Senders

L_B Strahldichte einer Hintergrundstrahlung

L_R an einer diffusen Oberfläche reflektierte Strahldichte

$L_{\lambda S}(\lambda, T)$ spektrale Strahldichte eines schwarzen Körpers

M spezifische Ausstrahlung

$M_{\lambda S}(\lambda, T)$ spektrale spezifische Ausstrahlung eines schwarzen Körpers

$M_{p\lambda S}(\lambda, T)$ photonenbezogene spektrale spezifische Ausstrahlung

M_S gesamte spezifische Ausstrahlung eines schwarzen Körpers

$n_{p\lambda}(\lambda)$ zur spektralen Strahlungsleistung äquivalente Photonenzahl

n_{ph} generierte Ladungsträgermenge in Abhängigkeit der Photonenzahl n_p

T Temperatur

W Strahlungsenergie

$W_p(\lambda)$ Photonenenergie in Abhängigkeit der Wellenlänge λ

$\alpha(\lambda)$ spektraler Absorptionsgrad

$\varepsilon(\lambda)$ spektraler Emissionsgrad

φ_A Winkel zwischen der Strahlungseinfallsrichtung und der Normale zu A_R

φ_E Winkel zwischen der Strahlungsausbreitungsrichtung und der Normale zu A_E

φ_S Winkel zwischen der Strahlungsausbreitungsrichtung und der Normale zu A_S

λ	Wellenlänge
$\rho(\lambda)$	spektraler Reflexionsgrad
$\tau(\lambda)$	spektraler Transmissionsgrad
Φ	Strahlungsleistung, Strahlungsfluss
$\Phi_\lambda(\lambda)$	spektralen Strahlungsleistung in Abhängigkeit der Wellenlänge λ
$\Phi_{\lambda\max}$	maximaler spektraler Strahlungsfluss
$\Phi_{\lambda S}(\lambda, T)$	spektraler Strahlungsfluss eines schwarzen Körpers
Ω	Raumwinkel
Ω_E	Raumwinkel, der durch ein Empfangselement aufgespannt wird
Ω_S	Raumwinkel, der durch ein Sendeelement aufgespannt wird

Kapitel 4

e^2	quadratische Abweichung der Interpolationsfehler im Grauwertfenster G_w
$\underline{f}_0, \underline{f}_1, \underline{f}_2$	Zeilenvektor mit Interpolationsparametern für die Grauwertmengen $\underline{g}_0, \underline{g}_1, \underline{g}_2$
$f(x_w, y_w)$	Interpolationsprofil im Grauwertfenster G_w in Abhängigkeit der Ortskoordinaten
f_{00}, f_{01}, f_{02}	Mittelwerte der Grauwertmengen $\underline{g}_0, \underline{g}_1, \underline{g}_2$
f_{x0}, f_{x1}, f_{x2}	linearer Interpolationsparameter der Grauwertmengen $\underline{g}_0, \underline{g}_1, \underline{g}_2$ in Spaltenrichtung
f_{y0}, f_{y1}, f_{y2}	linearer Interpolationsparameter der Grauwertmengen $\underline{g}_0, \underline{g}_1, \underline{g}_2$ in Zeilenrichtung
f_{xx}, f_{xx1}, f_{xx2}	quadratischer Interpolationsparameter der Mengen $\underline{g}_0, \underline{g}_1, \underline{g}_2$ in Spaltenrichtung
f_{yy}, f_{yy1}, f_{yy2}	quadratischer Interpolationsparameter der Mengen $\underline{g}_0, \underline{g}_1, \underline{g}_2$ in Zeilenrichtung
f_{xy}, f_{xy1}, f_{xy2}	quadratischer Interpolationsparameter der Mengen $\underline{g}_0, \underline{g}_1, \underline{g}_2$
F	F-verteilte Zufallsvariable
F_N	Freiheitsgradanzahl des Nenners bei t-Verteilung bzw. F-Verteilung
F_Z	Freiheitsgradanzahl des Zählers bei t-Verteilung bzw. F-Verteilung
\underline{g}	Spaltenvektor mit den Grauwerten g_i mit $i \in [1, I]$
\underline{g}_0	Grauwertvektor, der aus der Vereinigung der Spaltenvektoren \underline{g}_1 und \underline{g}_2 entsteht
\underline{g}_1	Spaltenvektor mit den Grauwerten g_{i1} mit $i \in [1, I]$ aus dem Fenster G_w in Bild 1
\underline{g}_2	Spaltenvektor mit den Grauwerten g_{i2} mit $i \in [1, I]$ aus dem Fenster G_w in Bild 2
$G(k)$	Grauwertbild/-matrix zum Zeitpunkt k
\underline{G}_w	Fenster bzw. Teilmatrix aus dem Grauwertbild $G(k)$
k	diskrete Zeitachse
$L(\underline{g}_i, \mu, \sigma^2)$	Verbundverteilungsdichtefunktion der Grauwerte g_i mit den Parametern μ und σ^2
$L(g_{i1}, \mu_1, (\sigma_1)^2)$	Verbundverteilungsdichtefunktion der Grauwertmenge \underline{g}_1 aus Bild 1
$L(g_{i2}, \mu_2, (\sigma_2)^2)$	Verbundverteilungsdichtefunktion der Grauwertmenge \underline{g}_2 aus Bild 2
$L(g_{i1}, f_{01}, f_{x1}, f_{y1}, (\sigma_1)^2)$	Verbundverteilungsdichtefunktion der Grauwertmenge \underline{g}_1 aus Bild 1

$L(g_{i2}, f_{01}, f_{x2}, f_{y2}, (\sigma_2)^2)$	Verbundverteilungsdichtefunktion der Grauwertmenge g_2 aus Bild 2
S	empirische oder aus der t- bzw. F-Verteilung ermittelte Auswerteschwelle
t	t-verteilte Zufallsvariable
$\underline{x}, \underline{y}$	Spaltenvektoren mit Spalten- und Zeilenkoordinaten zu Grauwertmengen g_1 und g_2
x_w, y_w	Ortskoordinaten in Spalten- und Zeilenrichtung
α	Vertrauensniveau (Parameter der F-Verteilung)
μ	linearer Mittelwert der Teilmatrix G_w / des Spaltenvektors g
μ_0	Mittelwert der Grauwertmenge g_0 , die aus der Vereinigung von g_1 und g_2 entsteht
μ_1	linearer Mittelwert der Teilmatrix G_w / des Spaltenvektors g_1 in Bild 1
μ_2	linearer Mittelwert der Teilmatrix G_w / des Spaltenvektors g_2 in Bild 2
σ^2	Varianz der Teilmatrix G_w / des Spaltenvektors g relativ zum Interpolationsprofil
$(\sigma_0)^2$	Varianz der Grauwertmenge g_0 , die aus der Vereinigung von g_1 und g_2 entsteht
$(\sigma_1)^2$	Varianz der Teilmatrix G_w / des Spaltenvektors g_1 im Bild 1
$(\sigma_2)^2$	Varianz der Teilmatrix G_w / des Spaltenvektors g_2 im Bild 2

Kapitel 5

a, A	Oberfläche des Dielektrikums bzw. Detektorelements
C_{py}	Kapazität des pyroelektrischen Detektorelements
C_{el}	resultierende Kapazität eines pyroelektrischen Detektors
D	Detektivität
D^*	spezifische Detektivität
G_{el}	resultierender Leitwert eines pyroelektrischen Detektors
c_{th}	spezifische Wärmekapazität eines Körpers
C_{th}	Wärmekapazität eines Körpers
d	Dicke des Dielektrikums
E_E	Betrag der äußeren elektrischen Feldstärke infolge von Oberflächenladungen Q
E_I	Betrag der inneren elektrischen Feldstärke infolge der Polarisation P
g_{th}	Wärmeübergangskoeffizient
G_{ft}	Eingangleitwert eines Feldeffekttransistors
G_{py}	Leitwert des hochohmigen Detektordielektrikums
G_{th}	totaler Wärmeübergangskoeffizient
$h_s(t)$	zeitabhängige Stoßantwort eines pyroelektrischen Detektors
$H_{el}(s)$	elektrischer Teil der Übertragungsfunktion eines pyroelektrischen Detektors
$H_{py}(s)$	Übertragungsfunktion zur Berücksichtigung des pyroelektrischen Effekts
$H_s(s)$	Übertragungsfunktion eines pyroelektrischen Detektors

$ H_s(\omega) $	Betragsamplitudenspektrum der Übertragungsfunktion $H_s(s)$
$H_{th}(s)$	thermischer Teil der Übertragungsfunktion eines pyroelektrischen Detektors
$H_{uv}(s)$	vereinfachte Übertragungsfunktion des Feldeffekttransistors
$I(s)$	Laplace-Transformierte des Stromes $i(t)$
K	konstanter Verstärkungsfaktor
$L\{\dots\}$	Laplace-Transformierte
$L\{\dots\}^{-1}$	inverse Laplace-Transformierte
m	Betrag des Dipolmoments
NEP	<u>N</u> oise <u>E</u> quivalent <u>P</u> ower
$p(T)$	pyroelektrischer Koeffizient
P	Betrag des Vektors der Polarisation
P_s	Betrag der remanenten Polarisation
Q	Ladung
$R(f)$	Empfindlichkeit eines pyroelektrischen Detektors
R_{el}	resultierender Widerstand eines pyroelektrischen Detektors
$R_\lambda(f)$	spektrale Detektorempfindlichkeit
s	komplexe Variable
t	Zeit
T	Temperatur des Dielektrikums
T_U	Umgebungstemperatur
$u(t)$	zeitabhängige Eingangsspannung des Feldeffekttransistors
$U(s)$	Laplace-Transformierte der Spannung $u(t)$
$v(t)$	zeitabhängige Ausgangsspannung des Feldeffekttransistors
V	Volumen einer quaderförmigen Probe
$V(s)$	Laplace-Transformierte der Spannung $v(t)$
V_U	Spannungsverstärkung des Feldeffekttransistors
$\alpha(\lambda)$	spektraler Absorptionskoeffizient
$\varepsilon(\lambda)$	spektraler Emissionskoeffizient
λ	Wellenlänge
$\varphi_s(s)$	Phasenspektrum der Übertragungsfunktion $H_s(s)$
ρ	Dichte
σ_a	Oberflächenladungsdichte auf den Detektorelektroden
τ_{el}	elektrische Zeitkonstante eines pyroelektrischen Detektors
τ_{th}	thermische Zeitkonstante eines pyroelektrischen Detektors
ω	Kreisfrequenz

$\Phi(t)$	zeitlich variierende Leistung der einfallenden Temperaturstrahlung
$\Phi_a(t)$	vom Detektormaterial absorbierte Strahlungsleistung
$\Phi_d(t)$	Teil von $\Phi_a(t)$, der zur Temperaturerhöhung des Dielektrikums beiträgt
$\Phi_u(t)$	Leistungsverluste durch Wärmeabstrahlung an die Umgebung
$\Phi_s(t)$	Leistungsverluste durch Wärmeableitung an das Substrat
$\vartheta(t)$	Temperaturänderung des pyroelektrischen Dielektrikums
$\Theta(s)$	Laplace-Transformierte der Temperaturdifferenz $\vartheta(t)$
$\Psi(s)$	Laplace-Transformierte der Strahlungsleistung $\Phi(t)$

Kapitel 6

Kapitel 7

ALARM(k)	Verlauf des Alarmflags über der diskreten Zeitvariable k
AMP(k)	Variable zur Verfolgung der maximalen Signalamplitude auf einer Flanke
AMPMAX	Maximal zulässige Signalamplitude zur Alarmerzeugung
AMPMIN	Minimal erforderliche Amplitude zur Alarmerzeugung
ABS(k)	absolute Abweichung verschiedener Signale vom Bezugspegel REF(k)
BOT(k)	adaptiver unterer Schwellwert relativ zum Bezugspegel REF(k)
GRD(k)	Variable zur Verfolgung des maximalen Signalgradienten auf einer Flanke
GRDMAX	Maximal zulässiger Gradient zur Alarmerzeugung
GRDMIN	Minimal erforderlicher Gradient zur Alarmerzeugung
k	Diskreter Zeitpunkt
LEV(k)	Zwischengröße zur Bestimmung der Schwellwerte TOP(k) und BOT(k)
LEVADJ	Divisor zur Festlegung der Geschwindigkeit bei der Schwellwertadaption LEV(k)
MAX(k)	Variable zur Maximalwertverfolgung während einer steigenden Signalflanke
MIN(k)	Variable zur Minimalwertverfolgung während einer fallenden Signalflanke
NEG(k)	Variable zur Verfolgung der Dauer einer fallenden Signalflanke
PIR(k)	zeitdiskretes Ausgangssignal eines Bewegungsmelders
POL_AMP(k)	Polarität des Signals PIR(k) relativ zum Referenzpegel REF(k)
POL_CHG(k)	Polaritätswechsel des Signals PIR(k) relativ zum Referenzpegel REF(k)
POS(k)	Variable zur Verfolgung der Dauer einer steigenden Signalflanke
REF(k)	adaptiver Referenzpegel REF(k)
REFADJ	Divisor zur Festlegung der Adaptionsgeschwindigkeit des Bezugspegels REF(k)
SLPMAX	Maximal zulässige Flankendauer zur Alarmerzeugung
SLPMIN	Minimal erforderliche Flankendauer zur Alarmerzeugung

THRMAX	Maximal zulässige Abweichung LEV(k) zur Berechnung von TOP(k) und BOT(k)
THRMIN	Minimal zulässiger Abweichung LEV(k) zur Berechnung von TOP(k) und BOT(k)
TOP(k)	adaptiver oberer Schwellwert relativ zum Bezugspegel REF(k)
WID(k)	Variable zur Verfolgung der Dauer zwischen benachbarten gleichartigen Extrema
WIDMAX	Maximal zulässige Signaldauer zur Alarmerzeugung
WIDMIN	Minimal erforderliche Signaldauer zur Alarmerzeugung

Kapitel 8

A_c	System-/Zustandsübergangsmatrix eines zeitkontinuierlichen Zustandsraummodells
A	System-/Zustandsübergangsmatrix eines zeitdiskreten Zustandsraummodells
B_c	deterministische Steuermatrix eines zeitkontinuierlichen Zustandsraummodells
B	deterministische Steuermatrix eines zeitdiskreten Zustandsraummodells
C_c	Beobachtungsmatrix eines zeitkontinuierlichen Zustandsraummodells
C	Beobachtungsmatrix eines zeitdiskreten Zustandsraummodells
$d(k)$	Störanteil im Ausgangssignal $y(k)$ infolge des Eingangsräuschens $w_d(k)$
$E\{\dots\}$	Erwartungswertbildung
$f_d(\xi)$	Verteilungsdichtefunktion von $w_d(k)$ in Abhängigkeit der Schwelle ξ
$f_{\underline{x}(k+1) \underline{Y}(k)}(\xi \underline{Y}(k))$	Prädiktionsdichte (Kalman-Filter), bedingt auf Messwerthistorie $\underline{Y}(k)$
$f_{\underline{x}(k+1) \underline{Y}(k+1)}(\xi \underline{Y}(k+1))$	Filterdichte (Kalman-Filter), bedingt auf Messwerthistorie $\underline{Y}(k+1)$
$f_v(\xi)$	Verteilungsdichtefunktion von $v(k)$ in Abhängigkeit der Schwelle ξ
$F_d(\xi)$	Wahrscheinlichkeitsverteilung von $w_d(k)$ in Abhängigkeit der Schwelle ξ
$F_v(\xi)$	Wahrscheinlichkeitsverteilung von $v(k)$ in Abhängigkeit der Schwelle ξ
G_c	stochastische Steuermatrix eines zeitkontinuierlichen Zustandsraummodells
G	stochastische Steuermatrix eines zeitdiskreten Zustandsraummodells
k	diskrete Zeitachse
$K(k)$	Kalman-Gain-Matrix zur Gewichtung der Residuensequenz $r(k)$
$L\{\dots\}$	Laplace-Transformierte
$L\{\dots\}^{-1}$	inverse Laplace-Transformierte
$m_d(k)$	linearer Erwartungswert des Störprozesses $w_d(k)$
$m_v(k)$	linearer Erwartungswert des Rauschprozesses $v(k)$
$\underline{m}_x(k)$	Mittelwerte des zeitdiskreten Zustandsvektors $\underline{x}(k)$
$\underline{m}_x(t)$	Mittelwerte des zeitkontinuierlichen Zustandsvektors $\underline{x}(t)$
$R, R(k)$	Kovarianz des Rauschprozesses $v(k)$
$P_{dd}(k)$	Kovarianz des Störprozesses $w_d(k)$

$P_{vv}(k)$	Kovarianz des Rauschprozesses $v(k)$
$P_{xx}(k)$	Kovarianzmatrix des zeitdiskreten Zustandsvektors $\underline{x}(k)$
$P_{xx}(t)$	Kovarianz des Zustandsvektors $\underline{x}(t)$
$P^-(k)$	Prädiktionsfehlerkovarianz der Zustandsvorhersage
$P^+(k)$	Estimationsfehlerkovarianz der Zustandsschätzung
$Q, Q(k)$	zeitdiskrete Kovarianzmatrix des vektoriellen Rauschprozesses $\underline{w}(k)$
$Q(t)$	Varianz des Eingangsrauschprozesses $w(t)$
$r(k)$	Residuensequenz
$s(k)$	Nutzsignalanteil im Ausgangssignal $y(k)$ infolge des Eingangsrauschens $w_s(k)$
t	kontinuierliche Zeitachse
T	Abtastintervall
$u(k)$	skalares, deterministisches Eingangssignal eines zeitdiskreten Zustandsraummodells
$\underline{u}(k)$	vektorieller, deterministischer Input eines zeitdiskreten Zustandsraummodells
$u(t)$	deterministisches Eingangssignal eines zeitkontinuierlichen Zustandsraummodells
$v(k)$	skalarer Ausgangsrauschprozess eines zeitdiskreten Zustandsraummodells
$v(t)$	Ausgangsrauschprozess (Messrauschen)
$\underline{w}(k)$	vektorieller Eingangsrauschprozess eines zeitdiskreten Zustandsraummodells
$w(k)$	skalarer Eingangsrauschprozess eines zeitdiskreten Zustandsraummodells
$w_s(k)$	skalarer Nutzsignalanteil im Eingangsrauschprozess $w(k)$
$w_d(k)$	skalarer Störsignalanteil im Eingangsrauschprozess $w(k)$
$w(t)$	Eingangsrauschprozess (Driving Noise), stochastisches Eingangssignal
$\underline{x}(t)$	Zustandsvektor eines zeitkontinuierlichen Zustandsraummodells
$\dot{\underline{x}}(t)$	erste Ableitung des Zustandsvektors $\underline{x}(t)$
$\hat{\underline{x}}^-(k)$	vorhergesagter zeitdiskreter Zustandsvektor
$\hat{\underline{x}}^+(k)$	geschätzter zeitdiskreter Zustandsvektor
$y(k)$	äquidistant abgetastetes Ausgangssignal $y(kT)$ mit normierter Zeitachse ($T = 1$)
$y(kT)$	äquidistant abgetastetes Ausgangssignal $y(t)$ eines Bewegungsmelders
$\hat{y}^-(k)$	vorhergesagtes zeitdiskretes Ausgangssignal
$\underline{Y}(k)$	Messwerthistorie der Messwerte $y(k)$
$y(t)$	zeitkontinuierliche Ausgangsspannung eines Bewegungsmelders
$\Phi(t, t_0)$	globale Zustandsübergangsmatrix
ξ	Schwelle zur Ermittlung der Verteilungsfunktionen

Kapitel 9

A_c	Systemmatrix eines zeitkontinuierlichen Zustandsraummodells
A_i	Systemmatrix des zeitdiskreten Zustandsraummodells i mit $i = 1..4$
B_c	deterministische Steuermatrix eines zeitkontinuierlichen Zustandsraummodells
B_i	deterministische Steuermatrix des zeitdiskreten Zustandsraummodells i mit $i = 1..4$
C_c	Beobachtungsmatrix eines zeitkontinuierlichen Zustandsraummodells
C_i	Beobachtungsmatrix des zeitdiskreten Zustandsraummodells i mit $i = 1..4$
$d(k)$	Störsignalanteil im Ausgangssignal $y(k)$ infolge des Eingangsrauschens $w_d(k)$
$E\{\dots\}$	Erwartungswertbildung
f	Frequenz
$F\{\dots\}$	Fourier-Transformierte
$F^{-1}\{\dots\}$	inverse Fourier-Transformierte
$FDK(i,j)$	Kovarianz der Filterdichte, Komponente (i,j) mit $i = 1..4$ und $j = 1..4$
$FDM(i,1)$	Mittelwert der Filterdichte, Komponente i mit $i = 1..4$
G_c	stochastische Steuermatrix eines zeitkontinuierlichen Zustandsraummodells
G_i	stochastische Steuermatrix des zeitdiskreten Zustandsraummodells i mit $i = 1..4$
$h(t)$	Stoßantwort des vereinfachten Bewegungsmelders
$H(s)$	vereinfachte Übertragungsfunktion eines Bewegungsmelders
$ H(f) $	Betragsamplitudenspektrum
$ H(f) ^2$	Energiedichtespektrum
$H_s(s)$	Übertragungsfunktion eines pyroelektrischen Detektors
$\text{Im}\{\dots\}$	Imaginärteil einer komplexen Zahl
k	diskrete Zeitachse
K	Verstärkungsfaktor in der Übertragungsfunktion $H(s)$ des Bewegungsmelders
$K(k)$	Kalman-Gain-Matrix zur Gewichtung der Residuensequenz $r(k)$
$KGM(i,1)$	Kalman-Gain-Matrix, Komponente i mit $i = 1..4$
K_s	Verstärkungsfaktor in der Übertragungsfunktion $H_s(s)$ des Detektors
$L\{\dots\}$	Laplace-Transformierte
$L^{-1}\{\dots\}$	inversen Laplace-Transformierte
m_w	linearer Erwartungswert des Rauschprozesses $w(t)$
m_v	linearer Erwartungswert des Rauschprozesses $v(t)$
n_w	Leistungsdichte eines stochastischen Nutzsignals am Systemeingang
N_w	Leistungsdichte eines stochastischen Störprozesses am Systemeingang
$PDK(i,j)$	Kovarianz der Prädiktionsdichte, Komponente (i,j) mit $i = 1..4$ und $j = 1..4$

$PDM(i,1)$	Mittelwert der Prädiktionsdichte, Komponente i mit $i = 1 \dots 4$
P_{ww}	Varianz des Rauschprozesses $w(t)$
P_{vv}	Varianz des Rauschprozesses $v(t)$
P_{yy}	Leistung des Ausgangssignals $y(t)$
$P^-(k)$	bedingte Prädiktionsfehlerkovarianz der Zustandsvorhersage
$P^+(k)$	bedingte Estimationsfehlerkovarianz der Zustandsschätzung
Q	Kovarianzmatrix eines zeitdiskreten Rauschprozesses $\underline{w}(k)$ am Systemeingang
Q_i	Kovarianzmatrix des Rauschprozesses $\underline{w}(k)$, Kalman-Filter i mit $i = 1 \dots 4$
R	Kovarianzmatrix eines zeitdiskreten Rauschprozesses $v(k)$ am Systemausgang
R_i	Kovarianzmatrix des Rauschprozesses $v(k)$, Kalman-Filter i mit $i = 1 \dots 4$
$\text{Re}\{\dots\}$	Realteil einer komplexen Zahl
$r(k)$	Residuensequenz des Kalman-Filters
$RSS(1,1)$	Residuensequenz des Kalman-Filters
s	komplexe Variable
$s(k)$	Nutzsignalanteil im Ausgangssignal $y(k)$ infolge des Eingangsrauschens $w_s(k)$
T	Abtastintervall
T_c	Korrelationszeitkonstante
$ U(f) ^2$	Energiedichtespektrum des Signals $u(t)$
$U(s)$	Laplace-Transformierte einer determinierten Anregungsfunktion
$V(s)$	Laplace-Transformierte der Ausgangsspannung des Bewegungsmelders
V_{ac}	konstanter Verstärkungsfaktor
$v(t)$	Messrauschen am Systemausgang
$w(t)$	zeitkontinuierliche stochastische Anregung (Driving Noise)
$\underline{w}(k)$	vektorieller zeitdiskreter Rauschprozess am Systemeingang
$w_1(k)$	1. Komponente des Rauschvektors $\underline{w}(k)$
$w_2(k)$	2. Komponente des Rauschvektors $\underline{w}(k)$
$\underline{x}(t)$	Zustandsvektor eines zeitkontinuierlichen Zustandsraummodells
$\dot{\underline{x}}(t)$	erste Ableitung des Zustandsvektors $\underline{x}(t)$
$\ddot{\underline{x}}(t)$	zweite Ableitung des Zustandsvektors $\underline{x}(t)$
$x_1(t)$	erste Komponente des Zustandsvektors $\underline{x}(t)$
$x_2(t)$	zweite Komponente des Zustandsvektors $\underline{x}(t)$
$\hat{\underline{x}}^-(k)$	vorhergesagter zeitdiskreter Zustandsvektor
$\hat{\underline{x}}^+(k)$	geschätzter zeitdiskreter Zustandsvektor
$\underline{x}_s(k)$	Zustandsvektor $\underline{x}(k)$ des allgemeinen Signalmodells
$\underline{w}_s(k)$	Rauschvektor $\underline{w}(k)$ des allgemeinen Signalmodells

$\underline{x}_d(k)$	Zustandsvektor $\underline{x}(k)$ des Störmodells
$\underline{w}_d(k)$	Rauschvektor $\underline{w}(k)$ des Störmodells
$X(s)$	Laplace-Transformierte einer Hilfsgröße $x(t)$
$y(k)$	äquidistant abgetastetes Ausgangssignal $y(kT)$ mit normierter Zeitachse ($T = 1$)
$y(t)$	Ausgangssignal eines Bewegungsmelders
$ Y(f) ^2$	Energiedichtespektrum des Signals $y(t)$
α	inverse thermische Zeitkonstante
β	inverse elektrische Zeitkonstante
γ	Korrelationsparameter
$\delta(\tau)$	Dirac-Impuls
$\varphi(f)$	Phasenspektrum
$\varphi_{hh}(\tau)$	Autokorrelationsfunktion der Stoßantwort $h(t)$
$\varphi_{uu}(\tau)$	Autokorrelationsfunktion des Eingangssignals $u(t)$
$\varphi_{ww}(\tau)$	Autokorrelationsfunktion des weißen Rauschprozesses $w(t)$ am Systemeingang
$\varphi_{yy}(\tau)$	Autokorrelationsfunktion des Ausgangssignals $y(t)$
$(\sigma_{dd})^2$	Varianz des Störsignalanteils $d(k)$
$(\sigma_{ss})^2$	Varianz des Nutzsignalanteils $s(k)$
$(\sigma_{vv})^2$	Varianz des Signals $v(t)$
$(\sigma_{ww})^2$	Varianz des Signals $w(t)$
$(\sigma_{yy})^2$	Varianz des Signals $y(t)$
τ	zeitliche Verschiebung
τ_{th}	thermischen Zeitkonstante eines pyroelektrischen Detektors
τ_{el}	elektrische Zeitkonstante eines pyroelektrischen Detektors
ω	Kreisfrequenz
$\phi_{uu}(f)$	Leistungsdichtespektrum eines Leistungssignals $u(t)$
$\phi_{yy}(f)$	Leistungsdichtespektrum eines Leistungssignals $y(t)$
$\Phi(t-t_0)$	globale Zustandsübergangsmatrix eines Zustandsraummodells

Kapitel 10

A_4	Kalman-Filter 4: zeitinvariante Systemmatrix
AMPMAX	Maximal zulässige Signalamplitude zur Alarmerzeugung
AMPMIN	Minimal erforderliche Amplitude zur Alarmerzeugung
ABS(k)	absolute Abweichung verschiedener Signale vom Bezugspegel REF(k)
BOT(k)	adaptiver unterer Schwellwert relativ zum Bezugspegel REF(k)
C_4	Kalman-Filter 4: zeitinvariante Beobachtungsmatrix

d(k)	Störsignalanteil im Ausgangssignals $y(k)$ infolge der Störanregung $\underline{w}_d(k)$
EST	in den Beobachtungsraum abgebildeter, geschätzter Zustandvektor $\hat{\underline{x}}^+(k)$
FDM(i,1)	Komponente i des geschätzten Zustandsvektors mit $i = 1 \dots 4$
GRDMAX	Maximal zulässiger Gradient zur Alarmerzeugung
GRDMIN	Minimal erforderlicher Gradient zur Alarmerzeugung
k	diskrete Zeitachse
K(k)	Kalman-Filter 4: Kalman-Gain-Matrix
KGM(i,1)	Komponente i der Kalman-Gain-Matrix mit $i = 1 \dots 4$
K_s	konstanter Verstärkungsfaktor (hier $K_s = 1$)
LEV(k)	Zwischengröße zur Bestimmung der Schwellwerte TOP(k) und BOT(k)
LEVADJ	Divisor zur Festlegung der Geschwindigkeit bei der Schwellwertadaption LEV(k)
m_y	Mittelwert der Abtastwerte $y(k)$ im Fenster mit N Abtastwerten
n_w	auf den Bewegungsmeldereingang umgerechnete Nutzsinalvarianz $(\sigma_{ss})^2$
N_w	auf den Bewegungsmeldereingang umgerechnete Störsignalvarianz $(\sigma_{dd})^2$
P(k)	Kalman-Filter 4: Kovarianzmatrix des Prädiktionsfehlers
P ⁺ (k)	Kalman-Filter 4: Kovarianzmatrix des Schätzfehlers
PDM(i,1)	Komponente i des vorhergesagten Zustandsvektors mit $i = 1 \dots 4$
PIR(k)	allgemeines Ausgangssignal eines Bewegungsmelders (Kanal 1, 2 oder 3)
PIRi(k)	Digitalisiertes Ausgangssignal des Bewegungsmelders, Kanal i mit $i = 1 \dots 3$
PRE	in den Beobachtungsraum abgebildeter, vorhergesagter Zustandvektor $\hat{\underline{x}}^-(k)$
$Q_4(k)$	Kalman-Filter 4: Kovarianzmatrix des Eingangsruschprozesses $\underline{w}(k)$
r(k)	Kalman-Filter 4: Residuensequenz
R_4	Kalman-Filter 4: zeitinvariante Kovarianzmatrix des Ausgangsruschprozesses $\underline{v}(k)$
REF(k)	adaptiver Referenzpegel REF(k)
REFADJ	Divisor zur Festlegung der Adaptionsgeschwindigkeit des Bezugspegels REF(k)
RSS(k)	Residuensequenz r(k)
RSSi(k)	Residuensequenz des Bewegungsmeldersignals, Kanal i = 1..3
s(k)	Nutzsinalanteil im Ausgangssignals $y(k)$ infolge der Anregung $\underline{w}_s(k)$
SLPMAX	Maximal zulässige Flankendauer zur Alarmerzeugung
SLPMIN	Minimal erforderliche Flankendauer zur Alarmerzeugung
T	Abtastintervall
T_c	Korrelationszeitkonstante
THRMAX	Maximal zulässige Abweichung LEV(k) zur Berechnung von TOP(k) und BOT(k)
THRMIN	Minimal zulässiger Abweichung LEV(k) zur Berechnung von TOP(k) und BOT(k)

TOP(k)	adaptiver oberer Schwellwert relativ zum Bezugspegel REF(k)
v(k)	Kalman-Filter 4: skalarer Ausgangsrauschprozess
$\underline{w}(k)$	Kalman-Filter 4: vektorieller Eingangsrauschprozess
$\underline{w}_d(k)$	Kalman-Filter 4: vektorieller Eingangsrauschprozess (Störung)
WIDMAX	Maximal zulässige Signaldauer zur Alarmerzeugung
WIDMIN	Minimal erforderliche Signaldauer zur Alarmerzeugung
$\underline{w}_s(k)$	Kalman-Filter 4: vektorieller Eingangsrauschprozess (Nutzsignal)
$\hat{\underline{x}}^-(k)$	Kalman-Filter 4: Vorhersagewert des Zustandsvektors $\underline{x}(k)$
$\hat{\underline{x}}^+(k)$	Kalman-Filter 4: Schätzwert des Zustandsvektors $\underline{x}(k)$
y(k)	Kalman-Filter 4: gemessenes Ausgangssignal des Bewegungsmelders
α	inverse thermische Zeitkonstante τ_{th}
β	inverse elektrische Zeitkonstante τ_{el}
γ	Korrelationsparameter = T_c^{-1}
$(\sigma_{dd})^2$	Varianz des Störsignalanteils d(k) am Ausgang des Bewegungsmelders
$(\sigma_{ss})^2$	Varianz des Nutzsignalanteils s(k) am Ausgang des Bewegungsmelders
$(\sigma_{vv})^2$	Varianz des überlagerten Ausgangsrauschprozesses v(k)
$(\sigma_{yy})^2$	Varianz des gemessenen Ausgangssignals y(k) des Bewegungsmelders
τ_{th}	thermische Zeitkonstante des pyroelektrischen Detektors
τ_{el}	elektrische Zeitkonstante des pyroelektrischen Detektors

Kapitel 11

1 Einleitung

1.1 Motivation

Sensoren zur Bewegungs- und Anwesenheitserkennung von Personen sind ein wesentlicher Bestandteil moderner Gebäudeautomatisierungskonzepte. Sie ermöglichen die intelligente, bedarfsgerechte Steuerung zentraler Gebäudefunktionen und gewährleisten einen wirtschaftlichen Umgang mit den zur Verfügung stehenden Ressourcen. Die Liste möglicher Anwendungsfelder ist lang, sie umfasst unter anderem

- die Regelung von Heizungs-, Lüftungs- und Klimaanlage,
- den ordnungsgemäßen Betrieb von Transportanlagen (Fahrstühle, Rolltreppen),
- die Überwachung des Gebäudes unter sicherheitstechnischen Gesichtspunkten
- und die Steuerung von Türöffnungs- und Beleuchtungsanlagen.

Neben der automatischen Lichtabschaltung bei ausreichendem Tageslicht, die sicherlich ein gewisses Einsparungspotential bietet, kann der Energieverbrauch durch eine bedarfsabhängige Beleuchtungssteuerung weiter reduziert werden. Die technische Voraussetzung dafür sind Sensoren, welche die Anwesenheit von Personen erkennen und signalisieren. Da die "echte" Personenerkennung mit erheblichem technischen Aufwand verbunden ist, erfolgt die Präsenzkontrolle bei den heute üblichen Systemen auf dem Umweg über die Bewegungserkennung.

Wegen des geringeren Leistungsbedarfs bieten sich insbesondere passive Sensoren für Anwendungen in der Gebäudetechnik an. Je nach Applikation und Einsatzort ergeben sich allerdings sehr unterschiedliche Anforderungen hinsichtlich der geforderten Funktionalität, der Empfindlichkeit und der Störsicherheit, so dass auch aktive Sensorkonzepte wie beispielsweise die als Türöffner bekannten Mikrowellen-Bewegungsmelder oder die in Hochsicherheitsbereichen eingesetzten Ultraschall-Bewegungsmelder durchaus ihre Berechtigung haben.

Zur Überwachung größerer Flächen oder Räume werden insbesondere Passiv-Infrarot-Bewegungsmelder (kurz: PIR-Bewegungsmelder) eingesetzt. Diese früher vorwiegend im Outdoorbereich eingesetzten Melder registrieren Änderungen der Temperaturstrahlung in ihrem Erfassungsfeld. Ein Nachteil dieser Wächter ist die stark unterschiedliche Reichweite für Bewegungen in Richtung bzw. quer zum Sensor.

In den letzten Jahren wurden auch PIR-Bewegungsmelder zur Bewegungs- und Anwesenheitserkennung für den Indoorbereich entwickelt. Speziell zur Anwesenheitskontrolle ist eine deutlich höhere Erfassungsfelddichte erforderlich, da selbst feinmotorische Bewegungen erkannt werden müssen. Sensorarrays mit mehrerer Detektoren in Verbindung mit einem feiner unterteilten Fresnellinsenarray sorgen für die im Vergleich zu Outdoor-Anwendungen höherauflösende Bereichsabdeckung. Eine eindeutige örtliche Zuordnung von Temperaturände-

rungen ist infolge der Überlagerung mehrerer Erfassungsfeldbereiche auf den einzelnen Sensorelementen nicht möglich.

Deutlich höhere Auflösungen und eine eindeutige örtliche Zuordnung von Veränderungen lassen sich mit bildverarbeitenden Sensorsystemen erzielen. Im Gegensatz zu der aus sicherheitstechnischen Anwendungen bekannten Videoüberwachung erfordern einfache Automatisierungstechnische Aufgabenstellungen allerdings nicht unbedingt die Aufnahme und Übertragung von qualitativ hochwertigen Videobildern. Vielmehr ist eine Sensor-Aktor-Kombination, die im folgenden als "Bildwächter" bezeichnet wird, zur Aktivierung bzw. Deaktivierung eines oder mehrerer elektrischer Verbraucher ausreichend.

Als Sensorelement des "Bildwächters" bieten sich anstelle von CCD-Matrixsensoren mittlerweile CMOS-Bildsensoren an, deren Bildqualität durch die intensive Forschungs- und Entwicklungsarbeit der letzten Jahre zunehmend verbessert wurde. Wesentliche Komponenten zur Steuerung der Bildaufnahme, zum Auslesen und zur Speicherung von Bilddaten sowie zur Datenvorverarbeitung können auf dem Chip integriert werden, was sich sehr positiv auf den Leistungsverbrauch und auch auf die Herstellungskosten des Bildwächters auswirkt.

CMOS-Bildsensoren ermöglichen die Bewegungserkennung anhand von Intensitäts- oder Farbunterschieden in aufeinanderfolgenden Bildern. Weiterhin ist auch die Erfassung von geänderten Kontur- oder Texturverläufen denkbar. Prinzipiell ist allerdings zu beachten, dass bei der optischen Abbildung einer 3D-Szene auf den 2D-Bildsensor die Tiefeninformation der Szene verloren geht. Dieser Umstand wirkt sich beispielsweise dahingehend aus, dass die Unterscheidung zwischen Personen und deren Schattenbild nicht ohne weiteres möglich ist. Bewegt sich die Person außerhalb des Gesichtsfeldes, während der Schatten ins Erfassungsfeld fällt, sind die Änderungen nicht ohne weiteres als Störungen zu identifizieren. Als weiterer Nachteil eines einfachen, CMOS-basierten Bildwächters muss die bei völliger Dunkelheit erforderliche Ergänzungsbeleuchtung und der damit einhergehende Leistungsverbrauch betrachtet werden.

Die vorher beschriebenen Nachteile eines CMOS-Bildwächters und eines PIR-Bewegungsmelders bzw. PIR-Präsenzmelders lassen sich durch die Kombination beider Sensorprinzipien vermeiden. Dabei ist es zunächst unerheblich, ob beide Sensorprinzipien in einem Gerät integriert oder als zwei getrennte Geräte im Multisensorsystem betrieben werden. Der CMOS-Bildwächter kann in diesem Fall als passive Komponente zur hochauflösenden Anwesenheitskontrolle verwendet werden, nachdem der PIR-Melder die ohnehin vorhandene Raumbeleuchtung aktiviert hat. Großflächigere Änderungen im Intensitätsbild des Bildwächters können anhand des PIR-Melders – und umgekehrt – verifiziert werden, so dass insgesamt eine erhöhte Fehlalarmimmunität durch die multisensorielle Datenfusion zu erwarten ist.

1.2 Zielsetzung

Im Rahmen eines mehrjährigen Industrieprojekts wurden im Zentrum für Sensorsysteme (ZESS) der Universität Siegen intensive Untersuchungen zur Bewegungs- und Anwesenheitserkennung von Personen für unterschiedliche Anwendungen in der Gebäudetechnik durchgeführt. Im Anschluss an die Bewertung bekannter Sensorprinzipien bildete insbesondere die Verbesserung des Schaltverhaltens von PIR-Bewegungsmeldern einen Schwerpunkt der Entwicklungstätigkeit.

Zur Bewältigung dieser Aufgabenstellung war es erforderlich, alle Komponenten des Bewegungsmelders optimal aufeinander abzustimmen. Prinzipiell ist festzustellen, dass das optische System einen entscheidenden Einfluss auf das Schaltverhalten des Bewegungsmelders hat. Fehler bei der Auslegung des Sensor-Linse-Systems verschlechtern die Leistungsbilanz und sind nachträglich nicht kompensierbar. Zur Vereinfachung des Entwurfs von Fresnellinsenarrays sollte daher ein Simulationstool zur automatisierten Berechnung der Linsenanordnung in Abhängigkeit der Erfassungsfeldspezifikationen und der vom Wächterdesign vorgegebenen Linsenfläche erstellt werden.

Im Hinblick auf das Schaltverhalten und die Störsicherheit hat die konventionelle Signalauswertung mit Fensterkomparatoren entscheidende Nachteile. Moderne Bewegungsmelder werden daher durchweg mit Mikrocontrollern oder Mikroprozessoren ausgestattet, die eine adaptive Anpassung an die zugrundeliegenden Umgebungsbedingungen und eine sehr differenzierte Signalauswertung mit digitalen Signalverarbeitungsalgorithmen gestatten. Außerdem lassen sich ohne weiteren Hardwareaufwand zusätzliche Funktionen implementieren, welche die Funktionalität und den Bedienkomfort des Melders erhöhen. Der technische Fortschritt in der Controllertechnik und der damit einhergehende Preisverfall in den letzten Jahren hat diese Entwicklung beschleunigt.

Zur Entwicklung und Optimierung geeigneter Signalverarbeitungskonzepte war es erforderlich, objektive Kriterien zur Beurteilung des Schaltverhaltens bei Nutzsensoren und Störungen unter unterschiedlichsten Umgebungsbedingungen abzuleiten. Um die zeit- und damit kostenoptimale Entwicklung auch im Hinblick auf zukünftige Verbesserungen zu ermöglichen, sollte der Algorithmenentwurf und die anschließende Beurteilung und Parameteroptimierung rechnergestützt erfolgen. Neben der Implementierung und Untersuchung von Signalverarbeitungsverfahren sollte ein modular erweiterbares "Wächterentwicklungssystem" erstellt werden, das den Entwickler von zeitaufwendigen Arbeiten bei der Datenaufnahme, der Datenverwaltung sowie bei der Parameteroptimierung unterstützt.

Viele Entwicklungsergebnisse für PIR-Bewegungsmelder im Outdoor-Bereich lassen sich unmittelbar auf die Präsenzkontrolle im Indoorbereich übertragen. Im Hinblick auf das in Kap. 1.1 erläuterte Multisensorkonzept waren daher zusätzliche Untersuchungen zum CMOS-

Bildwächter erforderlich. Neben grundlegenden Überlegungen zum Systemkonzept sollte die Dimensionierung des optischen Systems unter Beachtung der technischen Randbedingungen sowie die Tauglichkeit bekannter Bildverarbeitungsverfahren für die vorliegende Aufgabenstellung untersucht werden.

Aufgrund der zur Bildverarbeitung anfallenden Datenmengen und des damit verbundenen Rechenaufwands ist insbesondere für den CMOS-Bildwächter eine Hardwareimplementierung des Auswerteverfahrens erforderlich. In diesem Zusammenhang bietet es sich an, die Algorithmik für den PIR-Bewegungsmelder ebenfalls in Hardware nachzubilden.

Das gute Preis-Leistungs-Verhältnis digitaler programmierbarer Bausteine sowie der hohe Integrationsgrad moderner CMOS-Sensoren bilden dabei die Grundlage zur Realisierung eines Low-Cost-Sensorsystems mit CMOS-Bildwächter und PIR-Bewegungsmelder. Die Implementierung beider Signalverarbeitungsverfahren auf der Zielhardware sowie die Entwicklung eines Verfahrens zur multisensoriellen Datenfusion ist allerdings nicht mehr Gegenstand dieser Arbeit.

1.3 Gliederung

Als Einstieg in die vorliegende Thematik sind in **Kap. 2** zunächst grundsätzliche Messgrößen zur Bewegungs- und Anwesenheitsdetektion von Personen beschrieben. Anschließend werden die nach dem Stand der Technik gebräuchlichsten Sensorkonzepte zur Personenerfassung in der Gebäudetechnik sowie Stärken und Schwächen der einzelnen Verfahren zusammengefasst. Dabei wurden auch Erfassungsprinzipien berücksichtigt, die sich prinzipiell ebenfalls für die vorliegende Aufgabenstellung eignen, die aber aus verschiedenen Gründen für einfache Aufgaben in der Gebäudetechnik nicht angewendet werden.

Aufgrund der besonderen Bedeutung für die vorliegende Arbeit wurde den CMOS-Bildsensoren ein eigenes Kapitel gewidmet. In **Kap. 3** werden zunächst einige allgemeine, strahlungsphysikalische und optische Grundlagen zusammengefasst, bevor die Eigenschaften von CMOS-Bildsensoren im Vergleich zu konventionellen CCD-Matrixsensoren erläutert werden. Die Beschreibung der Vorteile und Anwendungsmöglichkeiten der CMOS-Technologie im Hinblick auf die Realisierung eines Low-Cost-Bildwächters runden dieses Kapitel ab. In **Kap. 4** folgen dann die zur Erkennung von Grauwertänderungen in Videobildern untersuchten Bildverarbeitungsstrategien.

Der weitaus größte Teil der heute kommerziell erhältlichen Bewegungsmelder verwendet pyroelektrische Detektoren. Die hohe Empfindlichkeit und das einfache Systemkonzept haben wesentlich zu deren Verbreitung beigetragen. In **Kap. 5** wird die Funktionsweise pyroelektrischer Detektoren und die Vorgehensweise zur Ermittlung der Übertragungsfunktion erläutert.

Kap. 6 beschreibt das im Rahmen dieser Arbeit erstellte System zur Wächterentwicklung. Anhand von zwei Beispielen werden zunächst die Anwendungsmöglichkeiten des Simulationstools zum Fresnellinsendesign vorgestellt. Im zweiten Teil wird das Konzept und das Applikationspotential der zur Algorithmenentwicklung und Parameteroptimierung verwendeten Software erläutert, die sich im praktischen Einsatz durchaus bewährt hat.

Eine Auswahl konventioneller Auswertestrategien für PIR-Bewegungsmelder wird in **Kap. 7** präsentiert. Im Vergleich zur Signalauswertung mit Fensterkomparatoren ermöglichen diese Verfahren eine bessere Trennung von Stör- und Nutzsignalen. Außerdem wird die Vorgehensweise zur Parameteroptimierung am Beispiel eines ausgewählten Verfahrens verdeutlicht.

Basierend auf den Grundlagen der Estimationstheorie wird in **Kap. 8** ein innovatives Signalverarbeitungskonzept mit einem statistischen Filter vorgestellt. Dieses Filter verwendet intern ein zeitdiskretes lineares Zustandsraummodell des realen Prozesses und führt auf einen linearen, optimalen Estimationsalgorithmus, der nach seinem Entdecker "Kalman-Filter" genannt wird. Aufgrund der Optimalitätseigenschaften von Kalman-Filtern und des systematischen Ansatzes ergibt sich eine optimale Verarbeitungsgüte, die analytisch und quantitativ erfassbar ist.

Aufbauend auf der allgemeinen Beschreibung nach Kap. 8 wird in **Kap. 9** die Vorgehensweise zur Ableitung des zeitdiskreten Systemmodells für das Kalman-Filter erläutert. In **Kap. 10** wird darauf aufbauend die Anwendung des Kalman-Filters zur Vorverarbeitung von Bewegungsmeldersignalen beschrieben. In Verbindung mit typischen Wächtersignalen aus dem Indoor- und Outdoor-Bereich werden abschließend die Ergebnisse der Bewegungserkennung anhand von internen Größen des Kalman-Filters präsentiert.

Kap. 11 enthält eine Zusammenfassung der im Rahmen dieser Arbeit erzielten Ergebnisse.

2 Sensorkonzepte zur Anwesenheitserkennung

2.1 Einleitung

Während die Anwesenheitserkennung und sogar die Identifizierung von Personen für einen Menschen einfach ist und nahezu unbewusst abläuft, stellt diese Aufgabe an technische Systeme sehr hohe Anforderungen. Obwohl der Aufbau und die Funktion des menschlichen Auges mittlerweile recht gut bekannt sind, ist es bisher nicht gelungen, mit technischen Systemen nur annähernd die Leistungsfähigkeit des menschlichen visuellen Systems zu erreichen. Das liegt daran, dass der Mensch nicht nur "sieht", sondern er "nimmt wahr" [Lit. 2.1].

Neben der eigentlichen Bildaufnahme laufen in unserem visuellen System komplexe Prozesse zum Erkennen (Überführung von Bildpunkten in aussagefähigere verdichtete Strukturen) und zur Entscheidung (vom Erkennen abgeleitete Schlussfolgerungen über den Bildinhalt) ab [Lit. 2.1]. Diese Prozesse der Informationsverdichtung und -verarbeitung sind für die hohe Leistungsfähigkeit unseres biologischen Sehsystems verantwortlich. Bis heute sind sie aber nur ansatzweise verstanden.

2.2 Messgrößen zur Anwesenheitserkennung

Im folgenden werden daher zunächst die heute mit technischen Systemen möglichen Sensor-konzepte zur Anwesenheitskontrolle von Personen beschrieben. Abb. 2.1 enthält eine Zusammenfassung der prinzipiellen Möglichkeiten, erhebt aber keinen Anspruch auf Vollständigkeit. Unabhängig vom zugrundeliegenden Detektionsprinzip kann festgestellt werden, dass bei Anwendungen in der Gebäudetechnik meistens "natürliche Szenen" zur Bewertung vorliegen. Im Vergleich zu den bei industriellen Aufgabenstellungen oftmals anzutreffenden, künstlichen Arrangements besteht der wesentliche Unterschied darin, dass nur selten spezielle Randbedingungen zur Steigerung der Auswertesicherheit spezifiziert werden können.

Vom Sensorprinzip her ist die Personenerkennung mit bildverarbeitenden Systemen zur Erfassung von bewegten und unbewegten Personen geeignet. Unter praktischen Gesichtspunkten, die in Kap. 2.2.1 erläutert werden, erweist sich die uneingeschränkte Detektion allerdings als sehr schwierig. Insbesondere zur modellbasierten Erkennung werden meist körperspezifische Merkmale (aufrechte Gestalt, Gesichtskontur, Hautfarbe, Bewegung etc.) extrahiert und akkumuliert. Auch einzelne Körpermerkmale können bereits für eine im Vergleich zum Gesamtmodell einfachere Anwesenheitskontrolle herangezogen werden.

Eine weitere Möglichkeit zur Erkennung von Personen ist die messtechnische, berührungslose Erfassung von körpereigenen Vitalfunktionen. In Kap. 2.2.2 wird ein System vorgestellt, das über Entfernungen von 30 m die Atembewegungen und den Herzschlag von Lebewesen registrieren kann.

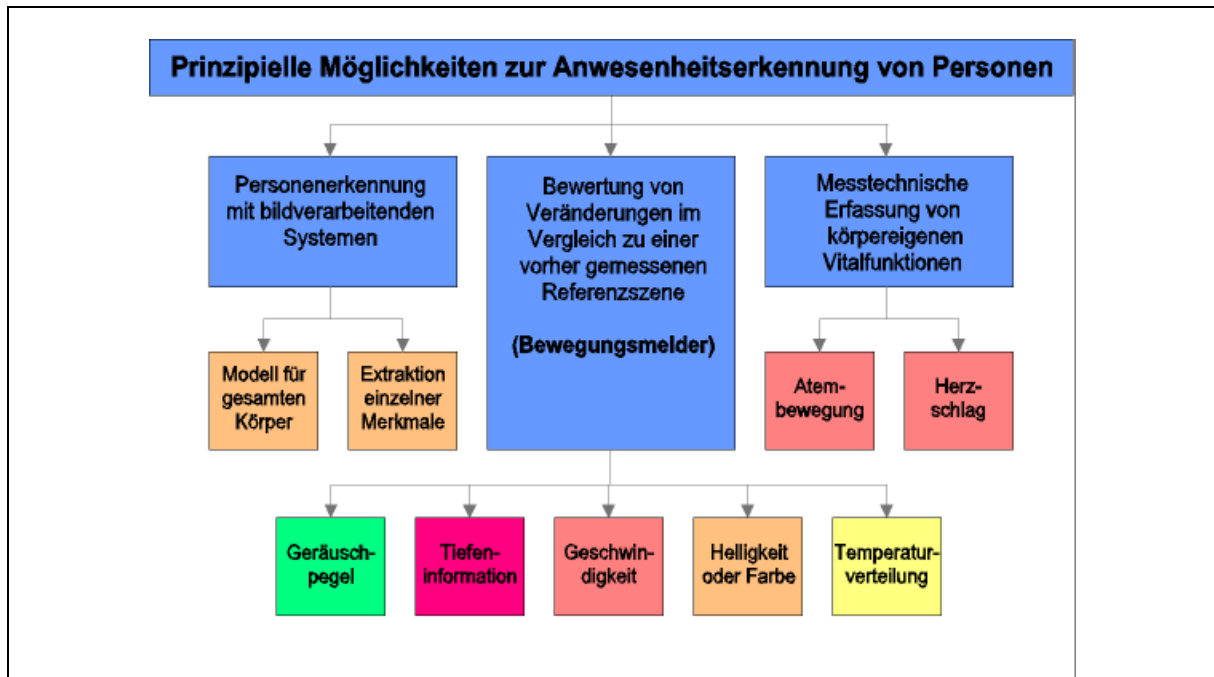


Abb. 2.1: Verfahren und Messgrößen zur Anwesenheitserkennung von Personen

Eine mit deutlich geringerem Systemaufwand verbundene, allerdings auch dementsprechend fehleranfällige Präsenzkontrolle ist auf dem Umweg über die Bewegungserkennung möglich. Hierbei werden Veränderungen im Vergleich zu einer vorher gemessenen und adaptiv nachgeführten Referenzszene beurteilt [Abb. 2.1]. Behelfsmäßig wird meist ein längerer Zeitraum überwacht, wobei das erneute Auftreten einer Bewegung als Indikator für die ununterbrochene Anwesenheit einer Person gewertet wird.

So ist es beispielsweise in einem ersten Ansatz denkbar, mit hochempfindlichen Mikrofonen die von Schritten oder sonstigen Bewegungen stammenden Geräusche aufzunehmen und zu bewerten. Dem Bereich "Sicherheitstechnik" sind weiterhin Sensoren zuzuordnen, welche die Luftdruckschwankungen innerhalb von Gebäuden, die z. B. beim Öffnen und Schließen von Fenstern und Türen auftreten, mit sehr empfindlichen Infraschallsensoren messen. Diese Sensoren registrieren Luftdruckschwankungen im Frequenzbereich von 0.2 ... 4.0 Hz und liegen somit außerhalb des vom Menschen wahrnehmbaren Schallbereichs. In [Pat. 2.1] wird ein Verfahren zum Überwachen von Innenräumen vorgeschlagen, das Luftdruckschwankungen innerhalb und außerhalb von Gebäuden auswertet, um die Fehlalarmrate zu reduzieren.

Verfahren zur Gewinnung von Tiefeninformationen (Positions-/Entfernungsdaten) oder von Bewegungsgeschwindigkeiten erfordern meist aktive Sensoren. Im Gegensatz zu passiven Systemen, die lediglich die vom detektierten Objekt selbst ausgehende Strahlung bewerten, beinhalten aktive Systeme stets die erforderlichen Sende- und Empfangskomponenten. Der Sender, der nicht zwingend in unmittelbarer Nähe des Empfängers angeordnet ist, erzeugt ein Fremdfeld. Bewegt sich ein Objekt in diesem Feld, ändern sich charakteristische Feldeigenschaften. Die in Kap. 2.3 beschriebenen Radar- und Ultraschallmelder arbeiten nach diesem

Prinzip, wobei insbesondere der dort beschriebene Doppler-Effekt zur Bewegungserkennung dient.

Verzichtet man auf aufwendige Bildverarbeitungslösungen zur "echten" Personenerfassung, lassen sich mit modernen, hochauflösenden CMOS-Bildsensoren kostengünstige Bewegungsmelder realisieren. Diese Melder werten Intensitäts- oder Farbunterschiede anhand von Differenzbildfolgen aus. Des Weiteren ist auch die Erfassung von geänderten Kontur- oder Texturverläufen denkbar. Insgesamt existiert eine Vielzahl von optischen Sensoren, die im Wellenlängenbereich des sichtbaren Lichts oder im nahen Infrarot empfindlich sind und die sich für unterschiedlichste Aufgabenstellungen und Erfassungsfeldgeometrien innerhalb der Gebäudetechnik eignen [Kap.2.3.3]. Es gibt sowohl aktive als auch passive Systeme, wobei die letztgenannten ausschließlich das Tageslicht bzw. eine ohnehin vorhandene Raumbeleuchtung als Lichtquelle verwenden. Bei völliger Dunkelheit "sehen" diese Sensoren allerdings nichts. Sie werden daher vorwiegend als zusätzliche Komponenten in Multisensoranordnungen eingesetzt.

Ein wesentlicher Nachteil aktiver Sensorsysteme ist der zusätzliche Leistungsbedarf zur Erzeugung des Fremdfeldes und der im Vergleich zu passiven Systemen erhöhte Bauteilauwand für den Sender. Der weitaus größte Teil der heute eingesetzten Bewegungsmelder basiert auf einem passiven Erfassungsprinzip, dem sogenannten pyroelektrischen Effekt [s. Kap. 5]. Diese Melder registrieren Änderungen der Temperaturstrahlung im Überwachungsbereich des Sensors. In Kap. 0 werden Anwendungsmöglichkeiten und Eigenschaften von thermischen Sensoren, darunter auch von pyroelektrischen Detektoren, vorgestellt.

2.2.1 Personenerkennung in Videobildern

Im Vergleich zu Sensoren, die lediglich Bewegungen im Erfassungsbereich registrieren, bieten bildverarbeitende Systeme grundsätzlich die Möglichkeit, bewegte und unbewegte Personen automatisch zu erkennen. Dazu muss der Algorithmus bekannte, personenspezifische Merkmale aus den Bilddaten extrahieren.

Bereits die Bildaufnahme ist dabei mit einem hohen Informationsverlust verbunden. Bei der optischen Abbildung der räumlichen Szene auf den Kamerachip entsteht ein Intensitätsbild ohne Tiefeninformation, so dass dem Algorithmus ein wesentliches Kriterium zur Separation zwischen Personen und dem Hintergrund fehlt. Ohne zusätzliche Maßnahmen ist es daher auch nicht mehr möglich, Verzerrungen durch die Projektion zu korrigieren oder die geometrischen Abmessungen der im Bild extrahierten Objekte zu ermitteln. Um die Tiefeninformation zurückzugewinnen, können zwei Videokameras an unterschiedlichen Raumpositionen verwendet werden (Stereo-Vision). Durch geeignete Kalibriertechniken, bei denen korrespondierende Punkte im Überwachungsbereich gesucht und anhand der bekannten Kamerapo-

sitionen in 3D-Koordinaten umgerechnet werden, kann die Rekonstruktion der 3D-Szene erfolgen. Diese Vorgehensweise bedingt einen hohen System- und Berechnungsaufwand.

Im folgenden werden daher zunächst Bildverarbeitungslösungen betrachtet, die nur das Intensitätsbild auswerten, und es wird eine Auswahl verschiedener, in der Literatur beschriebener Ansätze zur Personenerkennung in Videobildern vorgestellt. Der erste Schritt ist dabei meist die Separation von Personen (Vordergrund) und Hintergrund.

2.2.1.1 Trennung und Adaption komplexer Hintergrundszenen

Bei bekannten, unveränderlichem Hintergrund ist eine einfache Schwellwertoperation zur Trennung ausreichend. Sollen lediglich bewegte Objekte erfasst werden, liefert ein simples Differenzbildverfahren bereits eine erste Diskriminierung [Lit. 2.2]. Der Vordergrund wird dabei durch die Subtraktion von zeitlich aufeinanderfolgenden Bildern berechnet, wobei allerdings vorausgesetzt wird, dass die Person ständig in Bewegung bleibt.

Dieses Verfahren ähnelt der Vorgehensweise zur Bewegungserkennung durch die Bewertung des optischen Flusses [Lit. 2.3]. Die Berechnung und Bewertung des optischen Flusses ist aufwendig und erfordert für Echtzeitanwendungen spezielle Hardwarekonzepte. In [Lit. 2.4] werden CMOS-Bildsensoren zur Bewegungserkennung beschrieben, darunter auch einige zum Online-Processing des optischen Flusses, bei denen alle erforderlichen Berechnungsschritte in Hardware realisiert sind.

Die uneingeschränkte Erkennung und Verfolgung von Personen ist unter völlig allgemeinen Umgebungsbedingungen sehr anspruchsvoll. Variable Beleuchtungsverhältnisse durch natürliche und künstliche Lichtquellen, stark strukturierte Hintergründe mit zum Teil bewegten Objekten sowie bewegte und unbewegte Personen und deren Schattenbilder stellen hohe Anforderungen an ein Bildverarbeitungssystem. Bei einer ortsfesten Kamera können Personen zusätzlich von Objekten verdeckt werden, die eigentlich dem Hintergrund zuzuordnen sind. Umgekehrt verdecken Personen einen Teil des Hintergrunds, so dass in diesem Fall keine Information über den betreffenden Hintergrundabschnitt vorliegt.

Der Hintergrund kann daher nur in Ausnahmefällen als konstant angenommen werden. Im allgemeinen Fall ist eine ständige Adaption des Hintergrunds erforderlich, nicht zuletzt unter besonderer Beachtung der vorhandenen Beleuchtungsverhältnisse. In [Lit. 2.5] werden lediglich einfache Grauwertänderungen zur Anpassung des Hintergrunds verwendet, während in [Lit. 2.6] ein aufwendigerer Algorithmus zur Adaption vorgeschlagen wird.

Ausgehend von Grauwertänderungen und deren erster Ableitung wird dort ein Kalman-Filter zur Hintergrundberechnung verwendet. In Bildbereichen mit bewegten Objekten erfolgt die Adaption langsam, während dem Hintergrund zugeordnete Änderungen schnell nachgeführt werden. Nachteilig bei diesem Verfahren ist die um einen Zeitschritt verzögerte Umschaltung

zwischen Vordergrund und Hintergrund [Lit. 2.7], die unter Umständen bewirkt, das eine Person zweimal erkannt wird, nämlich an der gegenwärtigen Position und an der Stelle, an der sie sich vorher befand.

In [Lit. 2.7] wird ein Lösungsansatz mit einem verbesserten Adaptionsverhalten vorgestellt. Zur Hintergrundestimation dient ebenfalls ein Kalman-Filter. Im Vergleich zu [Lit. 2.6] wird allerdings nicht nur die Differenz zwischen dem aktuellen und dem geschätzten Grauwert zur Übernahme in den Vordergrund betrachtet. Zusätzlich wird die Grauwertabweichung unter der Bedingung, dass das aktuelle Pixel dem Hintergrund zuzuordnen ist, bewertet.

Mögliche Auswirkungen von Schattenbildern und Unsicherheiten infolge verdeckter Hintergrundbereiche werden anhand von dynamischen Schwellwerten berücksichtigt. Bei größerer Unsicherheit werden die zum Grauwertverlauf symmetrischen Schwellen zur Trennung von Vorder- und Hintergrund automatisch erhöht.

2.2.1.2 Modellbasierte Ansätze zur Erkennung und Verfolgung von Personen

Neben der zuvor beschriebenen Trennung von Vordergrund und Hintergrund unter bestmöglicher Berücksichtigung der Umgebungsbedingungen sind weitere Signalverarbeitungsschritte zur Personenerkennung erforderlich. Zur Verfeinerung der Erkennungsstrategie werden in den zuvor extrahierten Vordergrundregionen typische, personenspezifische Merkmale gesucht. Um die Klasse der detektierbaren Personen nicht zu stark einzuschränken, sollten dabei nur sehr allgemeine Annahmen bezüglich wesentlicher Körpermerkmale getroffen werden.

Eine wichtige Eigenschaft zur Personenerkennung ist die Bewegung. In Verbindung mit geeigneten Körpermodellen kann die Bewegungsrichtung und die Orientierung des Körpers im Videobild bestimmt werden, auch wenn die Person teilweise oder kurzfristig komplett von ortsfesten Gegenständen verdeckt wird. Aufgrund der Variabilität des menschlichen Körpers und möglicher Eigenverdeckungen sowie unter Beachtung unterschiedlicher Kleidungsgeohnheiten ist es bisher nicht gelungen, ein allgemeingültiges Modell zur Anwesenheitserkennung vorzugeben, das für alle Alltagssituationen eine zufriedenstellende Erkennungsqualität gewährleistet.

In [Lit. 2.8] und [Lit. 2.9] wird ein zweidimensionales, auf Spline-Interpolationen basierendes Modell zur Erkennung und Verfolgung von Fußgängern vorgestellt. Das "deformierbare" Modell beinhaltet Einschränkungen hinsichtlich der typischen, aufrechten Gestalt und typischer Bewegungsabläufe von Fußgängern [Abb. 2.2]. Das Modell und die zulässigen Abweichungen werden während einer Trainingsphase automatisch eingelernt. Diese Form der Modellierung gestattet eine flexible Anpassung des Modells in Abhängigkeit des Bewegungsablaufs und des Betrachtungswinkels. Die Kontur bestimmt den Suchbereich für die weitere Objektverfolgung. Teilweise Verdeckungen der Silhouette werden vom System toleriert.

Zum Abgleich zwischen dem Modell und den Bilddaten wird ein Kalman-Filter eingesetzt, das den Objektmittelpunkt und die Bewegungsgeschwindigkeit, die Skalierung und Orientierung des Objekts sowie typische Gestaltparameter rekursiv berechnet.

Im Gegensatz zu der vorher beschriebenen Erkennung von Fußgängern anhand der aufrechten Gestalt werden in [Lit. 2.10] zunächst Kopfformen in den extrahierten Vordergrundregionen gesucht. Dabei werden sowohl eindeutige Vordergrundkonturen als auch charakteristische Kantenverläufe in dieser Region berücksichtigt.

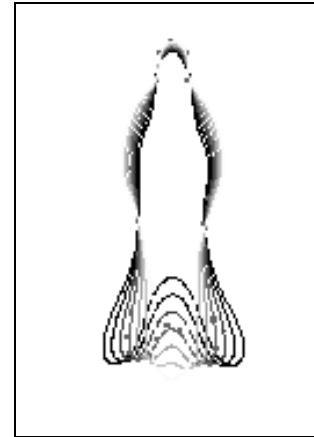


Abb. 2.2: "Fußgängermodell" nach [Lit. 2.8]

Ausgehend von gefundenen Köpfen werden in [Lit. 2.10] weitere körperspezifische Eigenschaften, beispielsweise die Schultern, der Oberkörper und die Arme, unter Verwendung eines im System hinterlegten Personenmodells geprüft. In diesem Modell sind unter anderem die Suchbereiche für weitere Körperteile sowie typische Längenverhältnisse, bezogen auf die Abmessungen des identifizierten Kopfes, hinterlegt. Der Kopf wird im Rahmen des 2D-Modells als Ellipse, die Schultern werden als Geraden mit unterschiedlichen Winkeln und der Oberkörper sowie die Arme werden jeweils als Rechteck modelliert.

Die bisher beschriebenen Verfahren verwenden 2D-Modelle zur videobasierten Personenerkennung. In [Lit. 2.11] wird ein 3D-Körpermodell vorgestellt, das nicht nur die Verfolgung von Personen, sondern auch die Analyse des gesamten Bewegungsablaufs ermöglicht. Das 3D-Modell beinhaltet eine Beschreibung der internen Struktur eines Objekts, die weitgehend dem menschlichen Skelett nachempfunden ist, eine Beschreibung der 3D-Gestalt des Objekts anhand geometrischer Primitive sowie eine Beschreibung typischer Objekteigenschaften, bezogen auf die einzelnen Körperteile [Lit. 2.11]. So wird der Kopf beispielsweise als Ellipsoid mit der zusätzlichen Eigenschaft "Hautfarbe" modelliert.

Das interne Objektmodell besteht aus einer Anzahl von Teilobjekten. Jedes Teilobjekt wird durch eine Gerade E_y zwischen zwei benachbarten Gelenkpunkten J_x und J_y beschrieben [Abb. 2.3]. Diese Struktur wurde speziell im Hinblick auf die Verfolgung von komplexen Bewegungsabläufen gewählt. Für die Gelenkpunkte werden zunächst die korrespondierenden 3D-Positionen bestimmt.

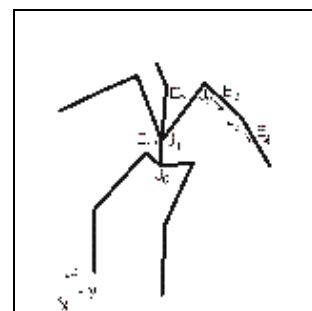


Abb. 2.3: Internes 3D-Objektmodell [Lit. 2.11]

Anschließend werden der Rumpf, die Arme und die Beine als konische Zylinder modelliert [Abb. 2.4], deren Symmetrieachsen mit den Verbindungsgeraden zwischen benachbarten Gelenkpunkten übereinstimmen. Das geometrische Modell beschreibt das Volumen des modellierten Körpers. Es kann zur Plausibilitätsprüfung für berechnete Gelenkstellungen dienen. Außerdem lassen sich Projektionen des Körpers auf beliebige Bildebenen berechnen.

Abb. 2.5 und Abb. 2.6 zeigen eine mit charakteristischen Markierungen versehene Testperson und das berechnete Objektmodell. Die Testperson wurde aus zwei verschiedenen Kameraperspektiven aufgenommen und daraus die 3D-Koordinaten der markanten Punkte bestimmt.

Eine allgemeine Anwendung zur Anwesenheitserkennung von Personen wird in [Lit. 2.11] zusätzlich beschrieben. Personen werden dabei anhand der typischen Kopfform und der zugehörigen Eigenschaft "Hautfarbe" identifiziert. Die Farbinformation wird dabei in einen speziellen, intensitätsunabhängigen Farbraum transformiert.

Mit einer im Vergleich zum Stereo-Vision-Konzept größeren Ungenauigkeit lässt sich auch die Objektposition im Raum abschätzen, indem der Kopf des 3D-Modells mit der im Videobild detektierten Kopfform zur Deckung gebracht wird [Abb. 2.7].

In den vorhergehenden Abschnitten wurden wesentliche Probleme bei der automatischen Erkennung und Verfolgung von Personen in Videobildern und verschiedene Lösungsansätze für spezifische Aufgabenstellungen zusammengefasst.

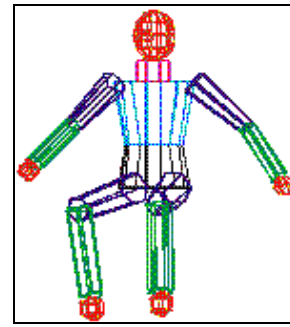


Abb. 2.4: Geometrisches 3D-Körpermodell [Lit. 2.11]



Abb. 2.5: Matching des internen Objektmodells [Lit. 2.11]

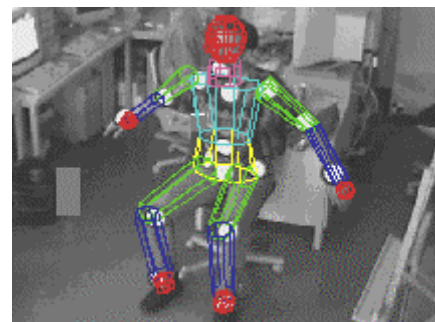


Abb. 2.6: Matching des geometrischen Modells [Lit. 2.11]

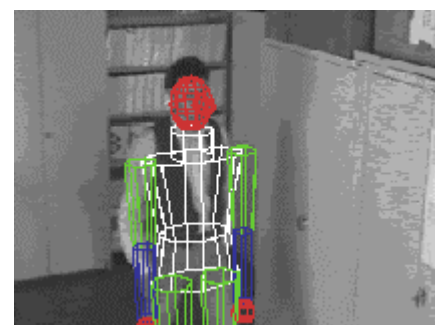


Abb. 2.7: Personendetektion durch Kopferkennung und Modellabgleich [Lit. 2.11]

Die uneingeschränkte Präsenzkontrolle unter völlig allgemeinen Umgebungsbedingungen ist bisher ein ungelöstes Problem, obwohl in den letzten Jahren zahlreiche Forschergruppen mit dieser Problematik beschäftigt waren und auch wichtige Fortschritte auf diesem Gebiet erzielt wurden. Für einfache Applikationen in der Gebäudeautomatisierung sind die vorgestellten Verfahren aufgrund des erforderlichen Systemaufwands und der damit verbundenen Kosten nicht geeignet.

2.2.2 Messtechnische Erfassung körpereigener Vitalfunktionen

Reine Bewegungsmelder reagieren äußerst empfindlich auf Bewegungen. Völlig ruhig sitzende Menschen können mit diesen Geräten allerdings nicht erkannt werden. Im folgenden wird ein aktives System zur Messung von körpereigenen Vitalfunktionen vorgestellt, das auch in diesem Fall eine sichere Personendetektion gewährleistet. Ausgenutzt wird der Umstand, dass ein von lebender Materie reflektiertes Mikrowellensignal eine von den biologischen Funktionen Herzschlag und Atembewegung beeinflusste Amplituden- und Phasenmodulation erfährt [Lit. 2.12].

Ausgewertet wird insbesondere die Phasenmodulation, da sie im Vergleich zur Amplitudenmodulation ein günstigeres Signal-Rausch-Verhältnis bietet. Das an ortsfesten Gegenständen der Umgebung zurückgestreute Signal besitzt im Gegensatz dazu lediglich eine konstante, von den örtlichen Gegebenheiten abhängige Phasenverschiebung zum Sendesignal.

Aufbau und Funktion des Sensorsystems sollen anhand von Abb. 2.8 erläutert werden. Das um 10 dB verstärkte Oszillatorsignal durchläuft das Dämpfungsglied α_1 , den Zirkulator und wird über die Antenne mit einem Abstrahlwinkel von 15° ausgesendet. Sender und Empfänger verwenden dieselbe Antenne, die Sendeleistung wird mit der Dämpfung α_1 eingestellt.

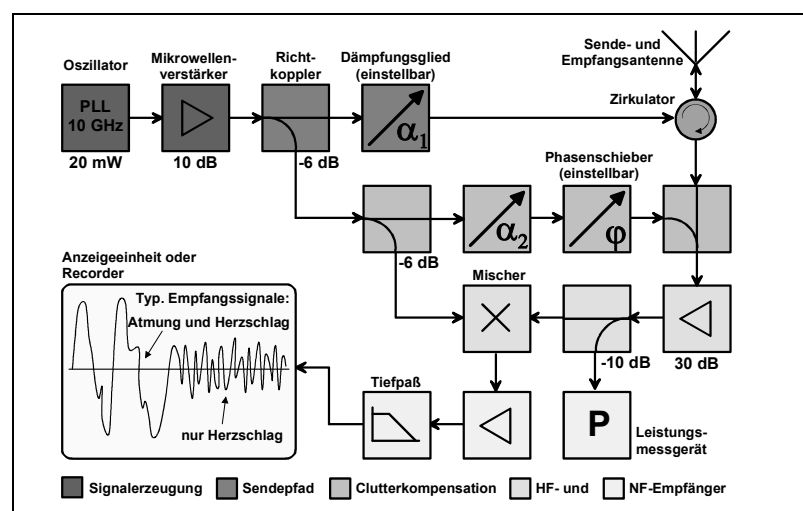


Abb. 2.8: Blockschaltbild eines elektronischen Lebensdetectionssystems [Lit. 2.12]

Aus dem Sendesignal werden zusätzlich die Referenzsignale zur Unterdrückung von Cluttersignalen sowie für den Mischer abgeleitet [Lit. 2.12]. Mit dem einstellbaren Dämpfungsglied α_2 und dem Phasenschieber ϕ werden Cluttersignale, die durch Reflektion des Sendestrahlens an ortsfesten Gegenständen in der Umgebung entstehen und daher eine konstante Phasenver-

schiebung zum Sendesignal aufweisen, weitgehend unterdrückt. Die Qualität der Kompensation kann mit dem Leistungsmessgerät P kontrolliert werden.

Das von Cluttersignalen bereinigte Empfangssignal wird anschließend einem Mischer zugeführt. Durch den Mischprozess mit dem ausgekoppelten Sendesignal und anschließende Tiefpassfilterung (Grenzfrequenz 4 Hz) entsteht ein niederfrequentes Signal, das auf die Atembewegungen des Brustkorbes und den Herzschlag zurückzuführen ist. Typische Signalformen von Empfangssignalen sind in Abb. 2.8 ebenfalls angedeutet [Lit. 2.12].

Eine Weiterentwicklung dieses Messverfahrens ist in den Patentschriften "Elektronisches Lebensdetektionssystem" [Pat. 2.2] und "Elektronisches Überwachungssystem" [Pat. 2.3] dargestellt. Neben der reinen Anwesenheitskontrolle dient das System zur Identifikation von Personen mit dem Ziel, die Aufenthaltsberechtigung festzustellen sowie zur Überprüfung des Gesundheitszustands. Vorausgesetzt wird dabei, dass jeder Mensch und jedes größere Tier infolge seiner Atmung, der Herzfunktion und weiterer Körperfunktionen ein spezifisches mechanisches Schwingungsspektrum aufweist, das ihn oder es ähnlich einem Fingerabdruck eindeutig charakterisiert. Eine Unterscheidung anhand des Spektrums ist daher möglich [Pat. 2.3].

Im Zuge der eigentlichen Auswertung wird ein mit Mikrowellen erstelltes Signalbild des Gebäudes einschließlich befugter Personen und Tiere (Sollzustand) mit dem aktuellen Signal (Istzustand) verglichen. Das Maß der zulässigen Abweichung ist einstellbar, wodurch sich die Empfindlichkeit des Systems beeinflussen lässt.

Ursprünglich wurden diese Systeme für militärische Zwecke entwickelt, beispielsweise zur Überwachung von Hochsicherheitsbereichen oder zur Prüfung des Gesundheitszustandes von Verwundeten oder Gefangenen. Des weiteren sind Anwendungen bei der Suche nach Verschütteten sowie im klinischen Bereich zur Überwachung von Patienten und von suizidgefährdeten Personen bekannt.

2.3 Präsenzkontrolle durch Bewegungserkennung

Bewegt sich eine Person im Überwachungsbereich eines Sensors, bieten sich, im Vergleich zu den Verhältnissen vor Eintritt der Person und auch während des unbewegten Verweilens im Erfassungsfeld, verschiedene physikalische Messgrößen zur Detektion an [Abb. 2.9]. In der Gebäudeautomatisierung haben sich bislang im wesentlichen drei Sensorprinzipien zur Bewegungserkennung durchgesetzt, die Schallwellen (Ultraschallbewegungsmelder) oder elektromagnetische Wellen (Mikrowellen- und Passiv-Infrarot-Bewegungsmelder) verwenden.

Die in Kap. 2.3.1 beschriebenen Ultraschallbewegungsmelder und die in Kap 2.3.2 vorgestellten Mikrowellenbewegungsmelder messen vorwiegend die Geschwindigkeit von bewegten Objekten. Darüber hinaus bieten insbesondere Ultraschallsensoren die Möglichkeit, mit rela-

tiv einfachen technischen Mitteln entfernungsmessende Systeme zu realisieren, welche die Anwesenheit von Objekten unmittelbar identifizieren können. Ultraschall- und Mikrowellenbewegungsmelder zählen zu den aktiven Systemen, da zusätzliche Komponenten zur Erzeugung des jeweiligen Feldes erforderlich sind.

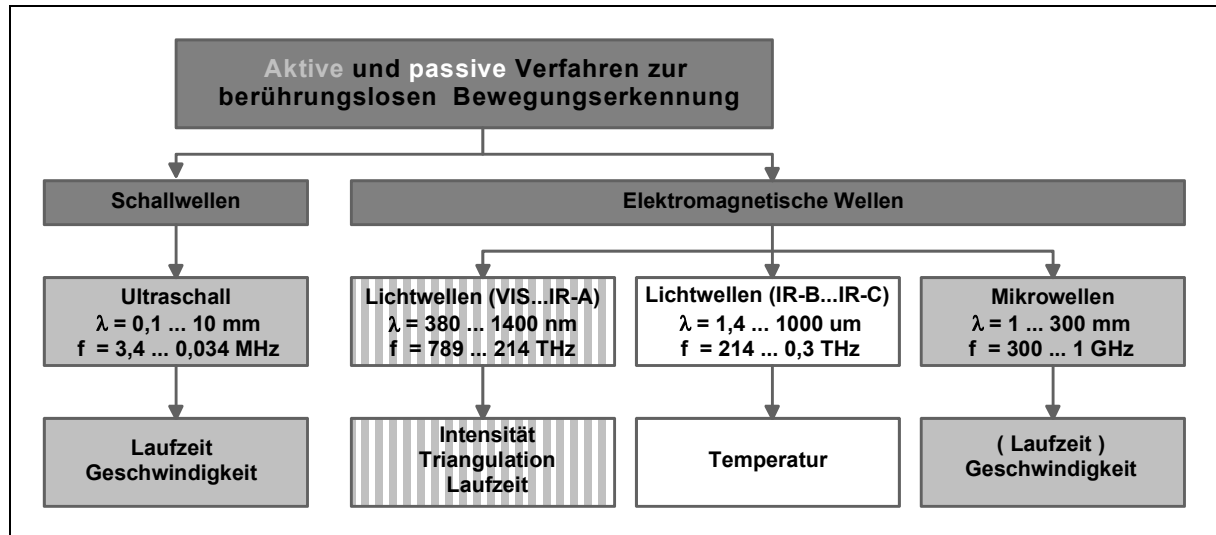


Abb. 2.9: Messprinzipien zur berührungslosen Bewegungserkennung

Passiv-Infrarot-Bewegungsmelder [Kap. 0] gehören dagegen zu den passiven Systemen. Sie erfassen die von beliebigen Objekten mit Temperaturen oberhalb des absoluten Nullpunkts ausgehende Temperaturstrahlung breitbandig. Speziell zur Personenerkennung ist dabei zu berücksichtigen, dass die aufgrund der Körpertemperatur emittierte Strahlung ein Maximum bei einer Wellenlänge von 10 μm, also im Bereich des fernen Infrarot (IR-C), aufweist.

Die vor allem in sicherheitstechnischen Anwendungen anzutreffenden optischen Sensoren (Lichtschranken, -gitter, etc.) sowie auch die Videobewegungsmelder arbeiten im Bereich des sichtbaren Lichts (VIS) oder im nahen Infrarot (IR-A) [Kap. 2.3.3]. Sie registrieren die emittierte, die transmittierte oder die an beliebigen Objekten reflektierte Strahlung natürlicher und künstlicher Lichtquellen. Je nach Anwendungsfall kann zwischen passiven und aufwendigeren aktiven Systemen mit zusätzlicher Beleuchtungseinheit unterschieden werden.

2.3.1 Ultraschallsensoren zur Präsenzkontrolle

Im Gegensatz zu elektromagnetischen Wellen zählen Ultraschallwellen zu den mechanischen Schwingungen und sind demzufolge an das Vorhandensein von Materie gebunden. Für beide Wellenformen gilt allerdings, dass das Produkt aus der Wellenlänge λ und der Frequenz f gleich der Phasengeschwindigkeit v der Welle ist:

$$(Gl. 2.1) \quad \lambda \cdot f = v$$

Ein wesentlicher Unterschied ist die Ausbreitungsgeschwindigkeit. Während sich elektromagnetische Wellen in Luft mit der konstanten Lichtgeschwindigkeit c des Vakuums fort-

pflanzen ($c = 3 \cdot 10^8$ m/s), breiten sich Ultraschallwellen in diesem Medium mit der deutlich geringeren Schallgeschwindigkeit ($v \approx 340$ m/s) aus. Aus diesem Grund bieten sich Ultraschallsensoren zunächst für Entfernungsmessungen nach dem Pulslaufzeitprinzip an. Abb. 2.10 zeigt ein einfaches System zur Bestimmung der Echolaufzeit.

Der Oszillator erzeugt ein Burstsinal, das in den Überwachungsbereich abgestrahlt wird. Das reflektierte Signal (Echo) wird verstärkt, digitalisiert und zur Zeitmessung verwendet. Der Taktgeber sorgt für die Umschaltung zwischen dem Sende- und dem Empfangsbetrieb, außerdem startet er die Zeitmessung.

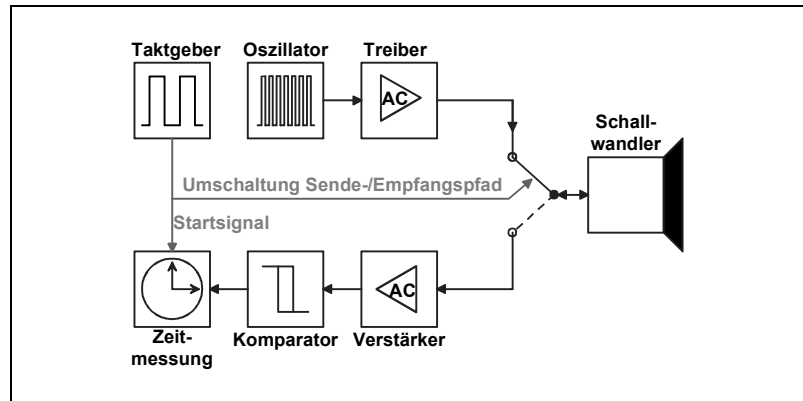


Abb. 2.10: Prinzipschaltbild eines Ultraschallsensors zur Messung der Echolaufzeit

Aus der Echolaufzeit T_E und der Schallgeschwindigkeit v lässt sich der Abstand a zwischen dem Sensor und dem reflektierenden Objekt ermitteln. Dabei ist zu berücksichtigen, dass der Burst diesen Abstand zweimal durchläuft:

$$(Gl. 2.2) \quad a = \frac{v \cdot T_E}{2}$$

Zur Vermeidung der Überlagerung des gesendeten und des empfangenen Signals ist ein gewisser Mindestabstand erforderlich. Das Entfernungsauflösungsvermögen e zwischen mehreren Targets ist durch die Burstdauer T_S gegeben:

$$(Gl. 2.3) \quad e = v \cdot T_S$$

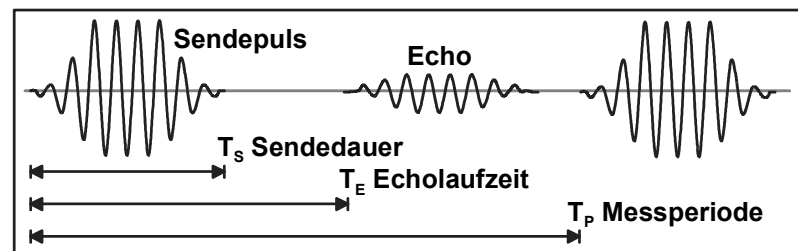


Abb. 2.11: Bestimmung der Echolaufzeit beim Pulslaufzeitverfahren

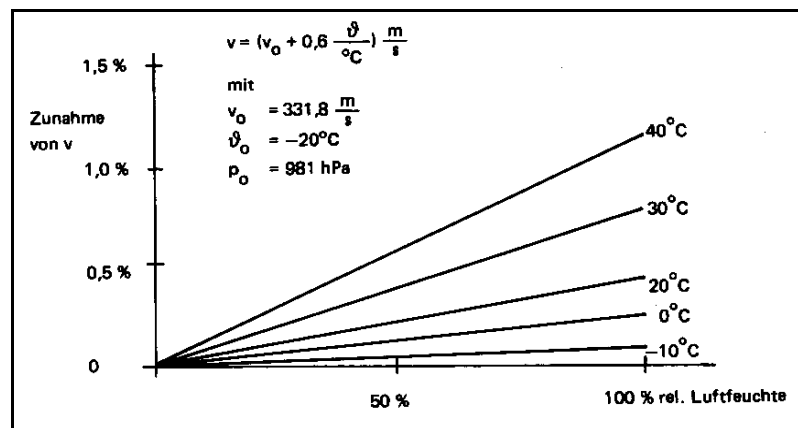


Abb. 2.12: Schallgeschwindigkeit als Funktion der relativen Luftfeuchtigkeit mit der Temperatur als Parameter [Lit. 2.13]

Die Genauigkeit dieses Systems wird von der Schallgeschwindigkeit beeinflusst.

In Abb. 2.12 ist die Zunahme der Schallgeschwindigkeit in Abhängigkeit der Luftfeuchtigkeit und der Temperatur dargestellt. Bei höheren Genauigkeitsanforderungen müssen diese Einflüsse kompensiert werden. Die von der Frequenz abhängige Dämpfung von Schallwellen zeigt Abb. 2.13. Bei höheren Frequenzen steigen die Verluste durch innere Reibung und Wärmeleitung.

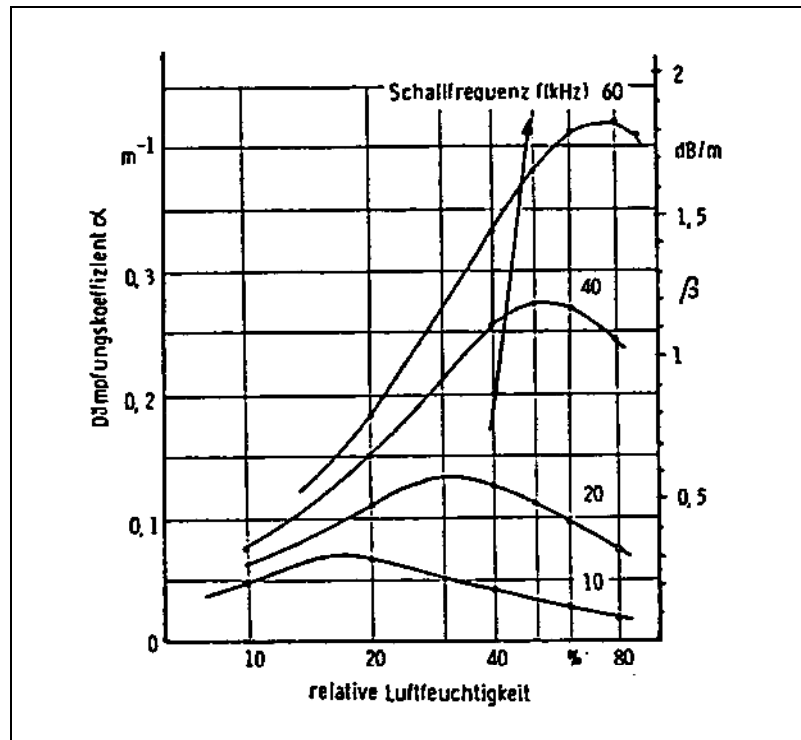


Abb. 2.13: Dämpfung in Abhängigkeit der relativen Luftfeuchtigkeit und der Frequenz [Lit. 2.13]

Bewegungsmelder auf Ultraschallbasis verwenden daher bevorzugt niedrige Schallfrequenzen bis ca. 40 kHz. In [Lit. 2.14] wird ein Ultraschallsensor vorgestellt, der nach dem zuvor beschriebenen Pulslaufzeitverfahren arbeitet. Aufgrund der durchgeführten Abstandsmessung ist eine präzise Trennung zwischen dem Erfassungsbereich und dem Hintergrund möglich. Insgesamt wird ein Raumwinkel von 60° mit einer zwischen 0,8 ... 2 m einstellbaren Reichweite überwacht.

Das System enthält einen Mikrocontroller, der die von den örtlichen Gegebenheiten und der Belegung des Überwachungsbereichs abhängigen Echosignale als digitales Rasterbild abspeichert und anschließend einen Mustervergleich zur Szenenauswertung durchführt. Zur Auswertung werden immer mehrere, nacheinander aufgezeichnete Signalverläufe verwendet, um die Zuverlässigkeit der Anwesenheitserkennung zu steigern. Im Messbereich bis 1,5 m lassen sich Auflösungen von 1 mm erzielen (Impulslaufzeit ca. 6 μs), so dass bereits die Atembewegungen einer Person zur Detektion ausreichen. Darüber hinaus sind deutlichere Körperbewegungen zur Anwesenheitserkennung erforderlich.

Neben der Messung von Pulslaufzeiten wird in Verbindung mit Ultraschallsensoren vorwiegend der Dopplereffekt zur Bewegungserkennung angewendet. Anstelle von einzelnen Impulsen oder Pulsfolgen wird ein unmoduliertes Dauerstrichsignal in den Überwachungsbereich ausgestrahlt (CW-Verfahren). Die Reflektion an ortsfesten Gegenständen bewirkt dabei lediglich eine konstante, entfernungsabhängige Phasenverschiebung, während durch die Reflektion

an bewegten Objekten zusätzliche, geschwindigkeitsproportionale Frequenzänderungen infolge des Dopplereffekts auftreten.

Zur Erläuterung des Dopplereffekts soll folgendes Szenario betrachtet werden. Bewegt sich ein Empfänger mit konstanter Geschwindigkeit im Wellenfeld eines stationären Senders, registriert er eine im Vergleich zur tatsächlichen Frequenz f höhere Schwingungsfrequenz, wenn er sich der Quelle nähert. Entfernt er sich vom Sender, wird eine geringere Frequenz wahrgenommen. Im allgemeinen Fall eines bewegten Senders und eines bewegten Empfängers gilt für die beobachtete Frequenz f' (Gl. 2.4), für den stationären Sender lässt sich daraus (Gl. 2.5) ableiten.

$$(Gl. 2.4) \quad f' = \frac{v - v_E}{v - v_S} \cdot f$$

$$(Gl. 2.5) \quad f' = \frac{v - v_E}{v} \cdot f$$

Die Variablen v_S und v_E sind die Geschwindigkeitskomponenten von Sender und Empfänger in Ausbreitungsrichtung der Welle, f und v repräsentieren die Frequenz und die Phasengeschwindigkeit der Welle. Für den stationären Sender nach (Gl. 2.5) ergibt sich eine Dopplereffrequenz f_D gemäß (Gl. 2.6). Unter Berücksichtigung der physikalischen Grundgleichungen lässt sich weiterhin für eine ortsfeste Sender-Empfänger-Anordnung (Transceiver) und ein mit der konstanten Geschwindigkeit v_R bewegtes, reflektierendes Objekt (Gl. 2.7) ableiten.

$$(Gl. 2.6) \quad f_D = f' - f = \pm \left| \frac{v_E}{v} \right| \cdot f$$

$$(Gl. 2.7) \quad f_D = f' - f = \pm 2 \left| \frac{v_R}{v} \right| \cdot f$$

Bei allen Gleichungen ist zu beachten, dass die zur Wellenausbreitungsrichtung senkrechten Geschwindigkeitskomponenten keinen Beitrag zur Dopplerverschiebung liefern. In (Gl. 2.6) und (Gl. 2.7) gilt das positive Vorzeichen jeweils für eine Annäherung des Empfängers bzw. des Reflektionsobjekts an den Sender. Entfernt sich der Empfänger bzw. das Objekt, ist die Dopplerverschiebung negativ, so dass sich eine geringere resultierende Frequenz am Beobachtungsort ergibt. Der Vergleich von (Gl. 2.6) und (Gl. 2.7) zeigt, dass die Dopplereffrequenz im Fall einer ortsfesten, integrierten Sende-Empfangseinheit von der doppelten Objektgeschwindigkeit bestimmt wird. Die vorstehenden Ausführungen zeigen, dass das Dauerstrichverfahren nur zur Bewegungserkennung geeignet ist, absolute Abstände bzw. Entfernungsänderungen können mit diesem Verfahren nicht ermittelt werden.

Das Dopplersignal wird entweder durch den direkten Vergleich der Sende- und der Empfangsfrequenz oder durch die Mischung beider Signale und anschließende Demodulation ausgewertet. Insbesondere der Mischprozess wird in Kap. 2.3.2 näher erläutert. Mit modernen Mikrocontrollern lassen sich problemangepasste Auswertestrategien implementieren, so dass hochwertige und zuverlässige Bewegungsmelder realisiert werden können [Lit. 2.15].

Zusammenfassend kann festgestellt werden, dass sich mit Ultraschallsensoren sehr empfindliche Bewegungsmelder, basierend auf dem Pulslaufzeitprinzip oder nach dem Dopplerverfahren realisieren lassen. Fehlalarme sind vor allem infolge von Schallwellen im Ultraschallbe-

reich (z. B. Ultraschall-Fernbedienungen, Hundepfeifen), durch nicht von Personen verursachte Bewegungen im Erfassungsfeld (wehende Gardinen oder pendelnde Blätter von Pflanzen) und durch starke Luftturbulenzen (Heizkörper) zu erwarten. Um die Fehlalarmimmunität zu erhöhen, können Ultraschallsensoren mit anderen Sensorprinzipien kombiniert werden.

Ultraschallsensoren werden vorwiegend in Indooranwendungen eingesetzt. Ein hermetisch dichtes Sensorgehäuse ist nur mit erhöhtem Aufwand möglich, da Ultraschall als mechanische Schwingung eine starke Dämpfung beim Durchgang durch feste Medien erfährt. Zur Überwachung größerer Bereiche sind in der Regel mehrere Transceiver oder zumindest mehrere Sender in einem einzigen Gerät integriert. Sollen mehrere, voneinander unabhängige Ultraschallbewegungsmelder betrieben werden, sind besondere Maßnahmen zur Verhinderung von gegenseitigen Beeinflussungen erforderlich.

2.3.2 Mikrowellensensoren zur Bewegungserkennung

Während die Laufzeitmessung von Impulsen oder Impulspaketen bei Ultraschallsensoren mit einfachen technischen Mitteln realisierbar ist, erfordern derartige Messungen mit elektromagnetischen Wellen einen sehr hohen Systemaufwand. Wegen der unterschiedlichen Ausbreitungsgeschwindigkeiten benötigt eine Schallwelle beispielsweise für eine Strecke von 1 mm eine Laufzeit von ca. 3 μ s, während eine elektromagnetische Welle diesen Abstand in nur 3,3 ps durchläuft. Mikrowellenbewegungsmelder, die umgangssprachlich auch als Radarmelder bezeichnet werden, nutzen daher den in Kap. 2.3.1 beschriebenen Dopplereffekt zur Anwesenheitserkennung von Personen.

Radarmodule für die vorliegende Applikation arbeiten vorwiegend im unteren GHz-Bereich. Von der Fa. SIEMENS werden zwei Module bei 2,45 GHz (KMY -24) [Lit. 2.18] und bei 9,35 GHz (SMX-1) [Lit. 2.17] angeboten. Ferner sind Sensoren für das K-Band unterhalb von 26,5 GHz bekannt. Wegen der geringen Reichweite bis ca. 10 m sind atmosphärische Dämpfungen bei diesen Frequenzen von untergeordneter Bedeutung [Abb. 2.14].

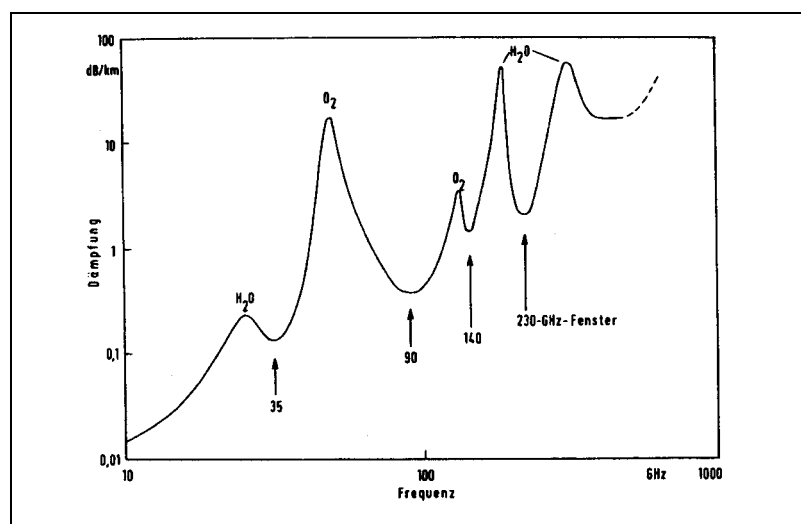


Abb. 2.14: Atmosphärische Dämpfung [dB/km] von Mikrowellen als Funktion der Frequenz [GHz] [Lit. 2.13]

In Abb. 2.15 ist das Blockschaltbild eines einfachen Doppler-Radars dargestellt. Das Mikrowellensignal wird zunächst verstärkt und über eine Antenne ausgestrahlt. Aus dem Sendesignal nach (Gl. 2.8) wird ebenfalls das Referenzsignal (Gl. 2.9) für den Mischer ausgekoppelt. Das Empfangssignal nach (Gl. 2.10) enthält einen zusätzlichen bewegungsabhängigen Anteil (Dopplerverschiebung) sowie eine konstante Phasenverschiebung durch Reflektionen an ortsfesten Gegenständen.

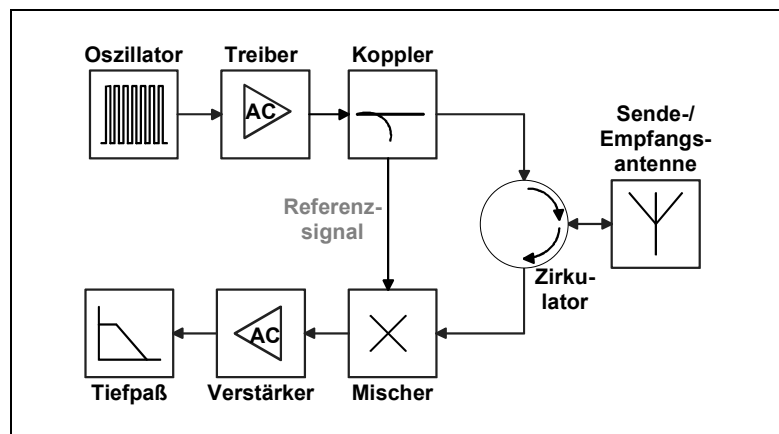


Abb. 2.15: Prinzipschaltbild eines Doppler-Radars zur Geschwindigkeitsmessung bei bewegten Objekten

(Gl. 2.8) Sendesignal $s(t) = A_S \cos(\omega t)$

(Gl. 2.9) Referenzsignal $r(t) = A_R \cos(\omega t)$

(Gl. 2.10) Empfangssignal $e(t) = A_E \cos(\omega t \pm \frac{2v}{c} \omega t + \varphi)$

Der konstante Anteil variiert je nach Entfernung zwischen $0 \dots 2\pi$, wobei zu berücksichtigen ist, dass der Eindeutigkeitsbereich infolge der hohen Sendefrequenz sehr klein ist. Der Mischprozess nach (Gl. 2.11) liefert Summanden mit der Differenzfrequenz sowie mit der Summenfrequenz von $r(t)$ und $e(t)$. Der rechte Term wird durch einen Tiefpass unterdrückt.

$$(Gl. 2.11) \quad m(t) = r(t)e(t) = \frac{A_R A_E}{2} \left[\cos\left(\mp \frac{2v}{c} \omega t + \varphi\right) + \cos\left(2\omega t \pm \frac{2v}{c} \omega t + \varphi\right) \right]$$

Die konstante Phasenverschiebung im Differenzfrequenzsignal hat zunächst keine Bedeutung für die reine Bewegungserkennung, da es nur um das Auftreten dieses Signals bzw. um die Geschwindigkeitsbestimmung anhand der Schwingungsdauer geht.

Abb. 2.16 und Abb. 2.17 zeigen typische Signalformen bei radialer und tangentialer Gehrichtung in einem Abstand von 3 m. Die Kurven wurden mit einem Radarmelder aufgezeichnet, der mit dem Sensor SMX-1 der Fa. SIEMENS bestückt war. Anhand von Abb. 2.16 lässt sich eine mittlere Periodendauer von ca. 16 ms ablesen, was einer Dopplerfrequenz von 62.3 Hz bzw. einer Gehgeschwindigkeit von nahezu 1 m/s entspricht. Bei tangentialer Bewegung nach Abb. 2.17 ist die Dopplerfrequenz kleiner und das Signal unsymmetrischer, da es lediglich durch geringfügige Abstandsänderungen beim Quergehen verursacht wird. Die Empfindlichkeitscharakteristik des Sensors SMX-1 zeigt einen Öffnungswinkel oberhalb von $\pm 60^\circ$, die Sendeleistung beträgt 1 mW.

Diese Leistung ist nach [Lit. 2.16] als gesundheitlich unbedenklich einzustufen. Mit zusätzlicher Verstärkung des Dopplersignals lassen sich frontal Reichweiten von 6 ...7 m realisieren, in Gehrichtungen von $\pm 60^\circ$ beträgt die Reichweite noch 3 ...4 m. Mit zunehmender Entfernung zum Sensor müssen die Bewegungen für eine sichere Detektion immer deutlicher werden.

Einen größeren Öffnungswinkel oberhalb von $\pm 90^\circ$ bietet der SIEMENS-Sensor KMY-24. Im Vergleich zum Doppler-Radar nach Abb. 2.15 enthält der Sensor zwei Mischer und eine zusätzliche Verzögerungsleitung.

Mit dieser Anordnung kann auch die Bewegungsrichtung erkannt werden. Abb. 2.18 verdeutlicht die Funktionsweise [Lit. 2.18]. Unter Berücksichtigung des Sendesignals aus (Gl. 2.8) ergibt sich am Ausgang von Tiefpass 1 das Signal $t_1(t)$ gemäß (Gl. 2.12), wobei beachtet wurde, dass sich beim Mischprozess die Auswirkungen der Verzögerungsleitung aufheben. Der Tiefpass 2 liefert $t_2(t)$ nach (Gl. 2.13).

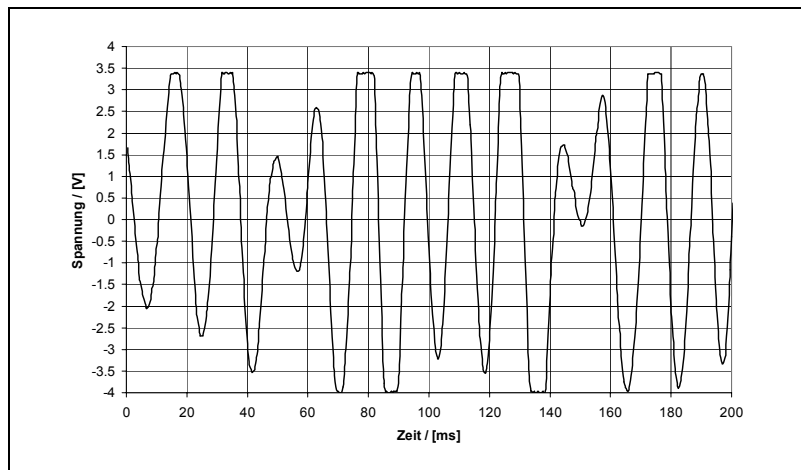


Abb. 2.16: Dopplersignal bei radialer Bewegungsrichtung in 3 m Entfernung (Gehgeschwindigkeit 1 m/s)

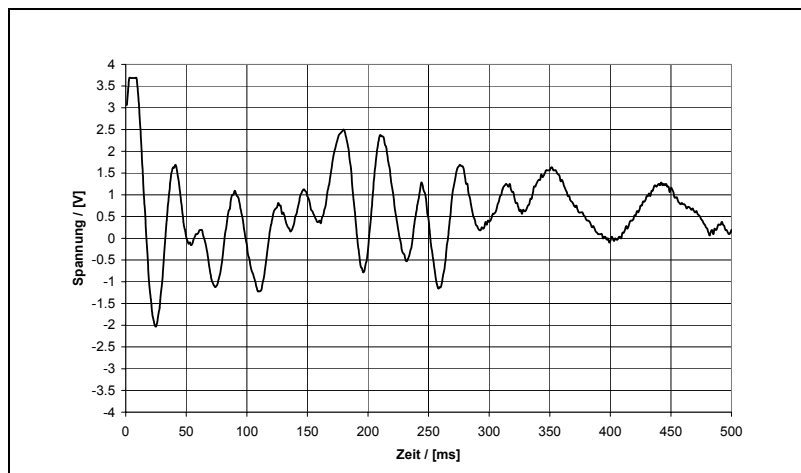


Abb. 2.17: Dopplersignal bei tangentialer Bewegung in 3 m Entfernung (Gehgeschwindigkeit 1 m/s)

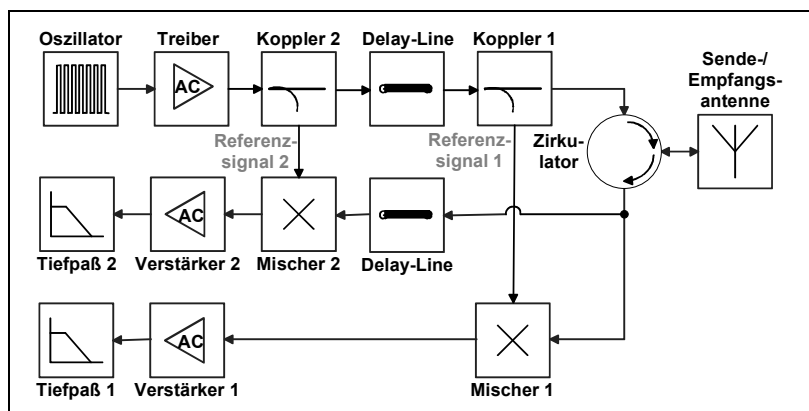


Abb. 2.18: Prinzipschaltbild eines Doppler-Radars mit Richtungserkennung

$$(Gl. 2.12) \quad t_1(t) = r_1(t)e(t) = s(t - \delta)e(t) = \frac{A_{R1} A_E}{2} \cos\left(\mp \frac{2v}{c} \omega t + \delta - \delta + \varphi\right)$$

$$(Gl. 2.13) \quad t_2(t) = r_2(t)e(t - \delta) = s(t)e(t - \delta) = \frac{A_{R2} A_E}{2} \cos\left(\mp \frac{2v}{c} \omega t + 2\delta + \varphi\right)$$

Vergleich man (Gl. 2.12) und (Gl. 2.13), erkennt man, das bei einer Phasenverzögerung $\delta = \pi/4$ das Signal $t_2(t)$ dem Signal $t_1(t)$ je nach Vorzeichen der Dopplerfrequenz entweder um 90° voreilt oder nacheilt. Nähert sich die Person dem Sensor, ist das Vorzeichen im Argument positiv und das Signal 2 eilt dem Signal 1 voraus, entfernt sich die Person, ist Signal 2 nacheilend.

Im Gegensatz zu anderen Sensoren reagieren Dopplersensoren ausschließlich auf Bewegungen und sind somit weitgehend unabhängig von Umgebungseinflüssen. Die Mikrowellen durchdringen Glas, Holz und dünne Wände, so dass auch eine verborgene Montage möglich ist. Eine Reichweiteneinstellung ist infolge dieser Eigenschaft allerdings kritisch. Es kann durchaus vorkommen, dass auch vorbeifahrende Fahrzeuge oder Bewegungen in Nachbarräumen zu ungewollten Alarmen führen.

Da die Mikrowellensensoren außerdem im Dauerstrichbetrieb arbeiten, sind zusätzliche Maßnahmen erforderlich, um mehrere Radarmodule innerhalb begrenzter Räumlichkeiten zu betreiben. Radarbewegungsmelder werden heute vorwiegend in automatischen Türöffnungsanlagen eingesetzt, wobei in der Bundesrepublik Deutschland zum Betrieb dieser Sensoren eine Postzulassung erforderlich ist.

Genau wie bei den Ultraschallmodulen überwacht meist ein einzelner Radarwächter einen größeren räumlichen Bereich. Der Forderung nach möglichst kleinen, unauffälligen Sensoren kommt der Umstand entgegen, das kleine Antennenaperturen eine breite Empfindlichkeitscharakteristik bedingen. Zur Berechnung des Öffnungswinkels 2α für eine runde Antenne mit dem Durchmesser d kann (Gl. 2.14) verwendet werden.

$$(Gl. 2.14) \quad \alpha = \arcsin\left(1,22 \frac{\lambda}{d}\right)$$

(Gl. 2.14) beschreibt die beugungsbegrenzte Winkelauslösung einer Antenne, wobei λ die Wellenlänge der Strahlung ist. Argumente ≥ 1 sind gleichbedeutend mit einem Öffnungswinkel von mindestens $\pm 90^\circ$ und daher mit einer sehr weiträumigen Abstrahlung verbunden. Die in der Gebäudetechnik eingesetzten Ultraschallsensoren und Radarmodule weisen typische Öffnungswinkel oberhalb von $\pm 30^\circ \dots \pm 90^\circ$ auf. Eine laterale Bereichsauflösung zur Überwachung einzelner Raumsegmente ist bei beiden Erfassungsprinzipien nur bedingt möglich und mit erheblichem Mehraufwand verbunden.

2.3.3 Intensitätsmessende optische Sensoren im Bereich des sichtbaren Lichts und des nahen Infrarot

Mit optischen Sensoren, die im Bereich des sichtbaren Lichts und im nahen Infrarot (vgl. Abb. 2.9) arbeiten, lassen sich im Vergleich zu Ultraschall- und Mikrowellensensoren lateral hochauflösende Bewegungsmelder realisieren. Abb. 2.19 vermittelt einen Überblick zu verschiedenen aktiven Messprinzipien mit optischen Sensoren und den zugehörigen Erfassungsfeldgeometrien. Prinzipiell lassen sich Sensoren unterscheiden, die lediglich die Intensität der empfangenen Strahlung bewerten, und solche, die alternativ oder auch zusätzlich Tiefeninformationen über die betrachtete Szene zur Verfügung stellen.

Zu den intensitätsmessenden Sensoren zählen einfache Lichtschranken für den Durchlicht-, den Reflektor- oder den Tasterbetrieb, die mit einzelnen oder auch zeilenförmig angeordneten lichtempfindlichen Elementen oder auch mit Kleinarrays ausgestattet sind, sowie die aus sicherheitstechnischen Anwendungen bekannten Lichtgitter und Lichtvorhänge zur Überwachung größerer Flächen. Auch die in konventionellen Videoanwendungen verwendeten Matrixsensoren gehören zu dieser Gruppe. Mit geeigneten Farbfiltern lässt sich dabei nicht nur die Intensität (Helligkeits- oder Grauwertinformation) der einfallenden Strahlung erfassen, sondern auch deren spektrale Zusammensetzung (Farbinformation).

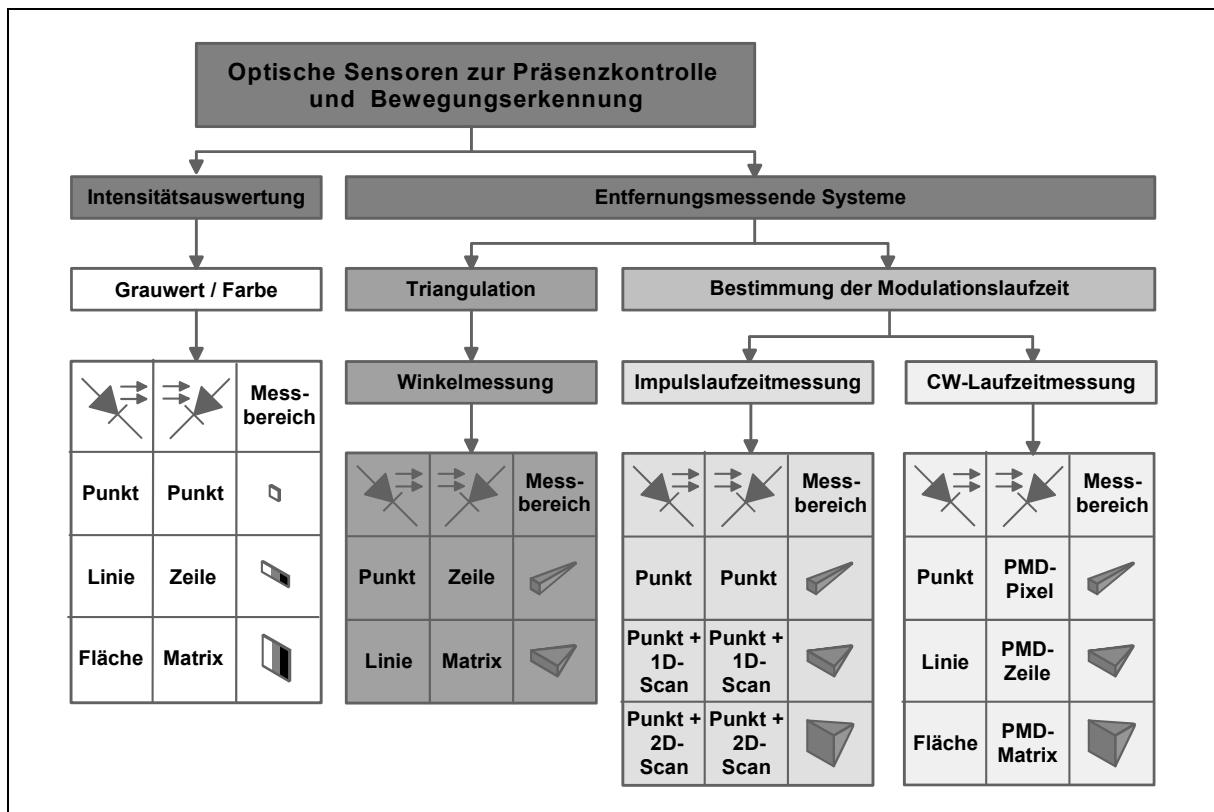


Abb. 2.19: Messprinzipien und Erfassungsfeldgeometrien von optischen Sensoren zur Präsenzkontrolle und Bewegungserkennung

Zur Anwesenheitserkennung von Objekten werten die zuvor beschriebenen Sensorsysteme lediglich Änderungen der Helligkeits- oder Farbinformation im Vergleich zu einer adaptiv nachgeführten Hintergrundszene aus. Langsame Veränderungen werden dabei dem Hintergrund zugeordnet, während schnelle und ausreichend starke Abweichungen als Objekteintritt oder Objektbewegung interpretiert werden. Objekte, die sich nur unwesentlich vom Hintergrund unterscheiden, können mit diesen Sensoren nicht erfasst werden.

Speziell mit Hinblick auf das sehr kostensensitive Marktsegment der Leuchtenindustrie wurde im Rahmen dieser Arbeit ein kostengünstiger, mikrocontrollergesteuerter Lichttaster zum berührungslosen Schalten und Dimmen von Kunstlichtquellen entwickelt. Der Aufbau und die Funktion dieses Tasters werden in Kap. 2.3.3.1 als Beispiel für ein intensitätsauswertendes System dargestellt. Anwendungen von höherauflösenden Sensorarrays in einfachen bildverarbeitenden Systemen sind in Kap. 3 und 4 detaillierter beschrieben und werden daher an dieser Stelle nicht weiter berücksichtigt.

Eine von einem Objekt verursachte globale oder lokale Änderung der Tiefeninformation im Überwachungsbereich, die als wesentliches Kriterium zur Anwesenheitserkennung zu betrachten ist, geht bei der optischen Abbildung der räumlichen Szene auf einen konventionellen Bildwandler verloren. Die Rekonstruktion oder die direkte Messung der Tiefeninformation bedingt einen höheren Systemaufwand.

Neben den intensitätsmessenden Systemen sind in Abb. 2.19 verschiedene Sensorkonzepte aufgeführt, die eine Abstandsbestimmung zwischen dem Sensor und einem detektierten Objekt ermöglichen. Während bei Triangulatoren die Abstandsberechnung anhand bekannter geometrischer Beziehungen erfolgt, werten andere Entfernungsmessgeräte die Laufzeit von einzelnen Lichtpulsen oder von Impulsfolgen aus. In Verbindung mit ein- oder zweidimensionalen Scannern bietet sich insbesondere das Pulslaufzeitverfahren zur Überwachung größerer Flächen oder Volumina an. Wegen des erforderlichen Hardwareaufwands sind diese Systeme für einfache Aufgabenstellungen in der Gebäudetechnik allerdings zu kostenintensiv, so dass hier geeignete Alternativen erforderlich sind.

Neben den Pulslaufzeitsystemen, die aufgrund der hohen Lichtgeschwindigkeit eine sehr hohe zeitliche Auflösung erfordern, haben sich in der Vergangenheit vor allem Entfernungsmessgeräte durchgesetzt, welche die Phasenlaufzeit von CW-modulierten Lichtwellen bestimmen. Aufgrund der im Vergleich zu Pulslaufzeitsystemen deutlich längeren Messdauer ist deren Einsatz in Verbindung mit Scannern zur Raumüberwachung allerdings weniger sinnvoll. Vielmehr wäre bei der Verwendung dieses Verfahrens die simultane Messung mit vielen parallelen Entfernungsmessern erforderlich, was aber bislang technisch nicht möglich war.

Ein sehr vielversprechender Ansatz zur Lösung dieses Problems wurde an der Universität Siegen gefunden. Die innovative, von Hrn. Prof. Schwarte erfundene und von der Fa. S-TEC

entwickelte PMD-Technologie bietet die Möglichkeit, entfernungsmessende Pixel in unterschiedlichen geometrischen Anordnungen auf einem CMOS-Chip zur integrieren. Die zum Verständnis wichtigsten Grundlagen dieser Technologie und das Anwendungspotential zur Gebäudeautomatisierung werden zu einem späteren Zeitpunkt dargestellt.

2.3.3.1 Lichttaster als Ersatz für taktile Lichtschalter und Dimmer

Als Beispiel für einen intensitätsauswertenden optischen Sensor wird ein im Rahmen von Projektarbeiten entwickelter, punktuell messender Lichttaster mit automatischer Hintergrundanpassung nach Abb. 2.20 beschrieben. Das zugrundeliegende Hardwarekonzept zeigt Abb. 2.21.

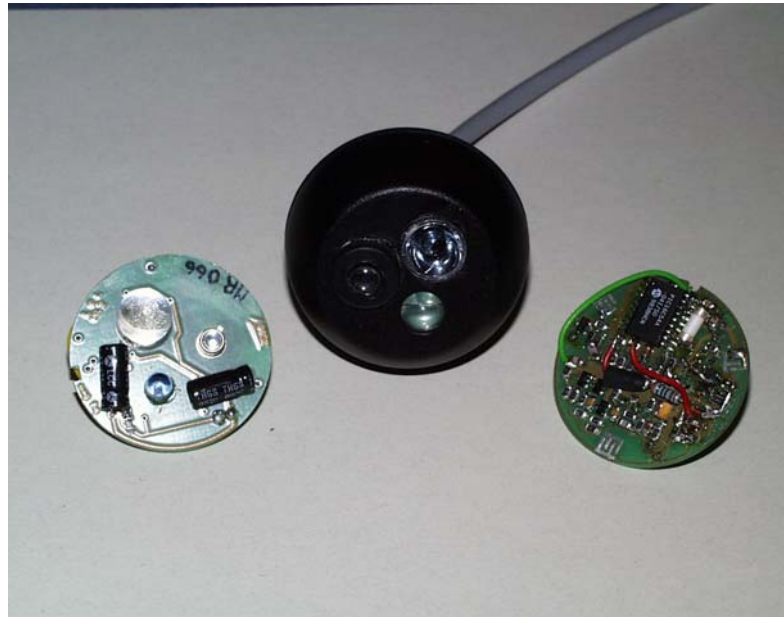


Abb. 2.20: Lichttaster als Ersatz für taktile Lichtschalter und Dimmer

Im unteren Teil von Abb. 2.21 erkennt man den LED-Sender, der infrarotes Licht mit einer Wellenlänge von 890 nm aussendet. Es wird gepulstes Licht mit einer Pulsdauer von 10 μ s und einer Wiederholrate von 2 ms abgestrahlt. Während des Impulses fließt ein maximaler Strom von ca. 2 A durch die LED, in den Impulspausen fließt kein Strom. Nach (Gl. 2.15) ergibt sich ein mittlerer Strombedarf des Senders von ungefähr 10 mA, das gesamte System verbraucht 15 mA.

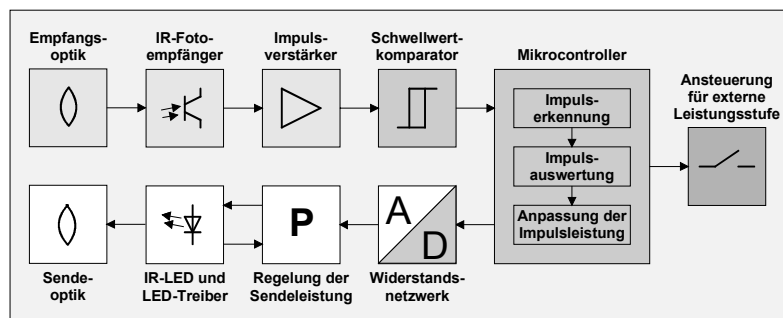


Abb. 2.21: Lichttaster zur Präsenzkontrolle und Bewegungserkennung

$$(Gl. 2.15) \quad \bar{I} = \left[\frac{T_{PULS}}{T_{PAUSE} - T_{PULS}} \right] \cdot I_{PULS} = \left[\frac{10 \mu s}{2 \text{ ms} - 10 \mu s} \right]_{500 \text{ HZ}} \cdot 2.0 \text{ A} \approx 10 \text{ mA}.$$

Die von der Sendediode emittierte Lichtleistung wird von einem Mikrocontroller in Abhängigkeit des aktuellen Empfangspegels während des Lichtimpulses gesteuert. Die Einstellung der Sendeleistung erfolgt durch Stromsteuerung der Sendediode, wobei ein einfaches Widerstandsnetzwerk zur Wandlung der digitalen Sollwerte in die analoge Steuerspannung dient.

Im oberen Teil von Abb. 2.21 ist der Empfangspfad dargestellt, der lediglich einen Fototransistor und einen mit diskreten Transistoren realisierten Impulsverstärker enthält. Der Komparator zur Digitalisierung des empfangenen Signals ist bereits Teil der Eingangsstufe des Mikrocontrollers.

Da lediglich ein digitalisierter Empfangspegel bewertet wird, stellt sich die Sendeleistung bei konstantem Hintergrund gerade so ein, dass sie genau um eine Leistungsstufe schwankt. Liegt kein Empfangssignal vor, wird die Sendeleistung erhöht. Im entgegengesetzten Fall wird sie reduziert [Abb. 2.22]. Die Regelung bewirkt, dass die Sendeleistung automatisch an geänderte Reflektivitätsbedingungen im Erfassungsfeld angepasst wird.

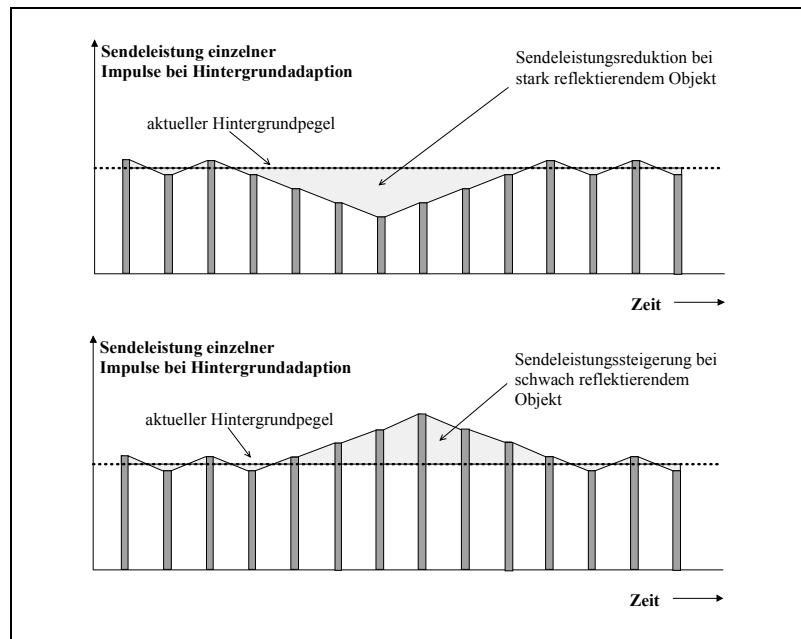


Abb. 2.22: Anpassung der gesendeten Lichtleistung in Abhängigkeit der Reflektivität des Hintergrunds

Der in digitaler Form vorliegende Wert der aktuellen Sendeleistung und der zeitliche Verlauf ermöglicht die Unterscheidung von kurzfristigen Bewegungen und einer längeren Anwesenheit im Erfassungsfeld. Außerdem kann das Vorzeichen der Reflektivitätsänderung bestimmt werden. Der Reflex-taster wurde zur Ansteuerung einer dimmbaren Lichtquelle entwickelt. Er unterscheidet anhand der Signaldauer beim Durchgang mit der Hand zwischen Schalt- und Dimmanforderungen seitens des Benutzers. Kurzzeitige Signaländerungen bewirken das Ein- bzw. Ausschalten der Lampe, während längere Abweichungen bei eingeschalteter Lampe den Dimmbetrieb aktivieren. Überschreitet die Dimmdauer eine vordefinierte Zeit, erfolgt die Zwangsabschaltung der Lampe.

Es hat sich gezeigt, dass für Reichweiten bis zu 1 m keine besonderen Anforderungen an das optische System bestehen. Die in den verwendeten Bauteilen ohnehin vorhandenen Elemente zur Lichtbündelung erweisen sich in Verbindung mit einem preiswerten Reflektor für die

Sendediode sowie einer einfachen Kollimatorlinse mit Abschattungsrohr für den Fototransistor als völlig ausreichend.

Mit einem etwas aufwendigeren optischen System lassen sich Reichweiten zwischen 2 ...3 m problemlos realisieren. Eine weitere Empfindlichkeitssteigerung ist möglich, wenn während eines Messzyklus mehrere Impulse unmittelbar vor dem Komparator aufintegriert werden.

2.3.3.2 Zusätzliche Informationen bei Verwendung von optischen Sensoren

Neben den vorstehend beschriebenen optischen Sensoren zur Bewegungs- und Anwesenheitserkennung sind in kommerziellen Bewegungsmeldern oftmals lichtempfindliche Sensoren zum tageslichtabhängigen Betrieb von Kunstlichtquellen integriert. Meist werden relativ langsame Fotowiderstände verwendet. Zusätzliche Informationen zur bedarfsabhängigen Beleuchtungssteuerung lassen sich, wie unten gezeigt wird, mit den schnelleren Fotodioden gewinnen.

Verwendet man optische Sensoren zur Bewegungserkennung, ist generell die Möglichkeit einer je nach Sensorkonfiguration sehr differenzierten zeitlichen oder örtlichen Helligkeitsmessung ohne zusätzlichen Systemaufwand gegeben.

Im folgenden wird ein Sensor zur Unterscheidung von Tageslicht und Kunstlicht vorgestellt, der den von künstlichen Lichtquellen emittierten Wechselanteil bewertet. Bei Glühlampen und bei Leuchtstofflampen mit konventionellen Vorschaltgeräten treten Intensitätsschwankungen mit der doppelten Netzfrequenz (in Deutschland 100 Hz) auf, während bei Leuchtstofflampen mit elektronischen Vorschaltgeräten typische Lichtwelligkeiten im 20 ...40 kHz-Bereich vorhanden sind. In Verbindung mit dimmbaren elektronischen Vorschaltgeräten werden heute Frequenzen von 70 ...80 kHz erreicht.

Der in Abb. 2.23 dargestellte Prototyp eines Lux-Sensors ermittelt sowohl den aktuellen Helligkeitswert als auch die von Kunstlichtquellen verursachten Schwankungen der Beleuchtungsstärke im Frequenzbereich von 100 Hz ...100kHz. Der Aufbau und die grundsätzliche Funktionalität des Sensors sollen anhand von Abb. 2.24 erläutert werden.

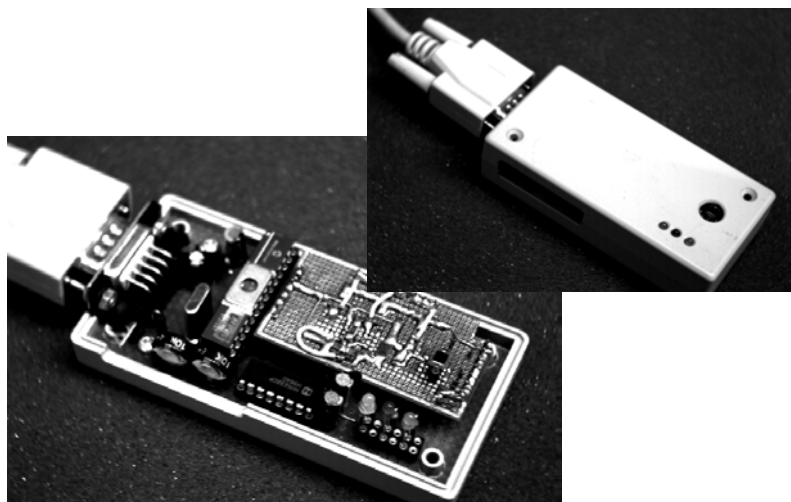


Abb. 2.23: Prototyp eines Sensors zur Unterscheidung von natürlichen und künstlichen Lichtquellen

Über einen externen Taster wird dem Sensor die Kunstlichtanforderung seitens des Benutzers mitgeteilt. In diesem Fall vergleicht der Sensor den aktuellen Helligkeitspegel mit einem zuvor eingelernten oder voreingestellten Schwellwert und schaltet bei Bedarf die Raumbeleuchtung ein.

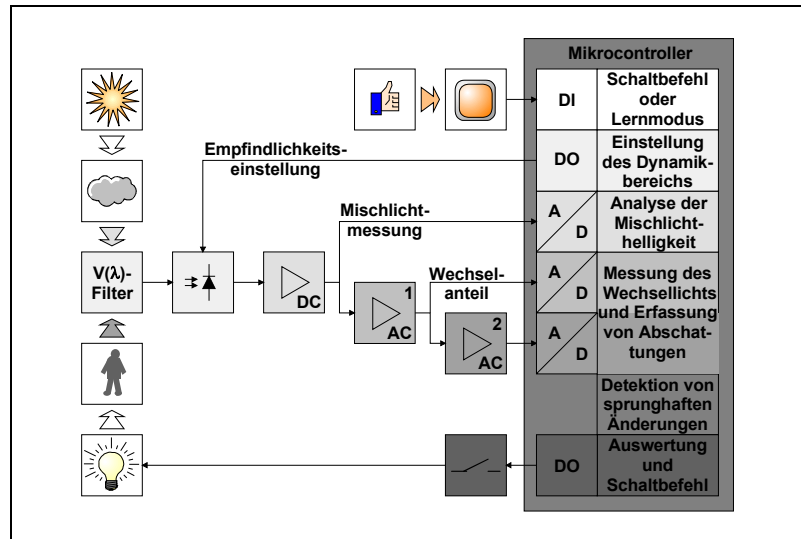


Abb. 2.24: Prinzipschaltbild des Helligkeitssensors

Um einen großen Dynamikbereich für die Helligkeitsmessung zur Verfügung zu stellen, erfolgt die Einstellung der Empfindlichkeit durch den Mikrocontroller.

Bei eingeschalteter Beleuchtung verfolgt der Sensor den zeitlichen Verlauf des Summenlichts und er schaltet die Lichtquellen aus, sobald die Beleuchtungsstärke am Sensorort eine extern vorgebbare Helligkeitsschwelle für einen längeren Zeitraum überschreitet.

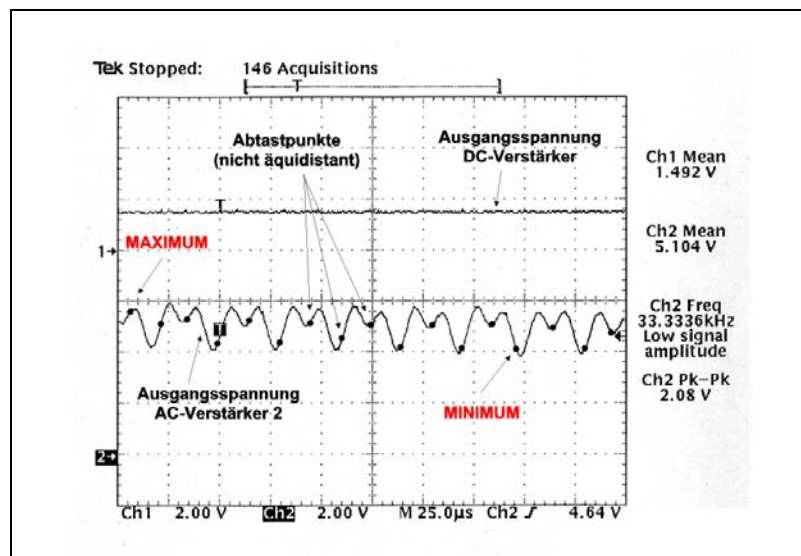


Abb. 2.25: Ermittlung des Lichtflimmerns von Kunstlichtquellen in einem definierten Zeitraum

Der Bewertungszeitraum wird erneut gestartet, wenn der eingestellte Schwellwert unter Berücksichtigung einer geringen Hysterese unterschritten wird. Je nach Montageort muss der Sensor dabei zwischen natürlichen Schwankungen des Tageslichts und lokalen Abschattungen oder Reflektivitätsänderungen in seinem Erfassungsbereich unterscheiden. Der Mikrocontroller wählt außerdem den zur Erkennung von Wechsellichtanteilen erforderlichen Verstärkerkanal aus. Zur Anhebung von kleinen Signalamplituden stehen zwei zusätzliche AC-Verstärker zur Verfügung. Die Vorgehensweise zur Erkennung von Wechselanteilen wird anhand von Abb. 2.25 erläutert. Der untere Signalverlauf zeigt typische Beleuchtungsstärkeschwankungen beim Betrieb zweier 36W-Leuchtstofflampen an einem elektronischen Vorschaltgerät (EVG). Die Abtastpunkte wurden nachträglich eingezeichnet und sollen nur zur

Verdeutlichung des Prinzips dienen. Der Sensor verfolgt über einen Zeitraum von 20 ms den Minimal- und den Maximalwert an dem selektierten Verstärkerausgang. Innerhalb dieser Zeit wird das analoge Signal nicht-äquidistant abgetastet. Die nicht-äquidistante Abtastung ermöglicht in gewissen Grenzen auch die Feststellung von Wechselanteilen, deren Frequenz höher ist als die verwendete Abtastfrequenz. Tendenzielle Aussagen bezüglich des Auftretens und der zeitlichen Änderungen von höherfrequenten Signalen sind daher auch in diesem Fall möglich.

Vorteilhaft anwenden lässt sich das zuvor beschriebene Verfahren zur Abschattungserkennung. Um eine bedarfsgerechte Steuerung der Raumbeleuchtung vorzunehmen, muss der Sensor prinzipiell zwischen Helligkeitsänderungen infolge von schwankenden Tageslichtverhältnissen und Abschattungen oder Reflektionsgradänderungen innerhalb seines Erfassungsbereiches unterscheiden. Während die relativ langsamen Tageslichtänderungen nur am Ausgang des DC-Verstärkers auftreten, beeinflussen Abschattungen und Reflektionsgradänderungen im Überwachungsbereich auch den Wechselanteil der Beleuchtungsstärke am Empfangsort. Auf diese Weise lassen sich Abschattungen identifizieren, die ansonsten die Einschaltdauer der Lampe verlängern würden.

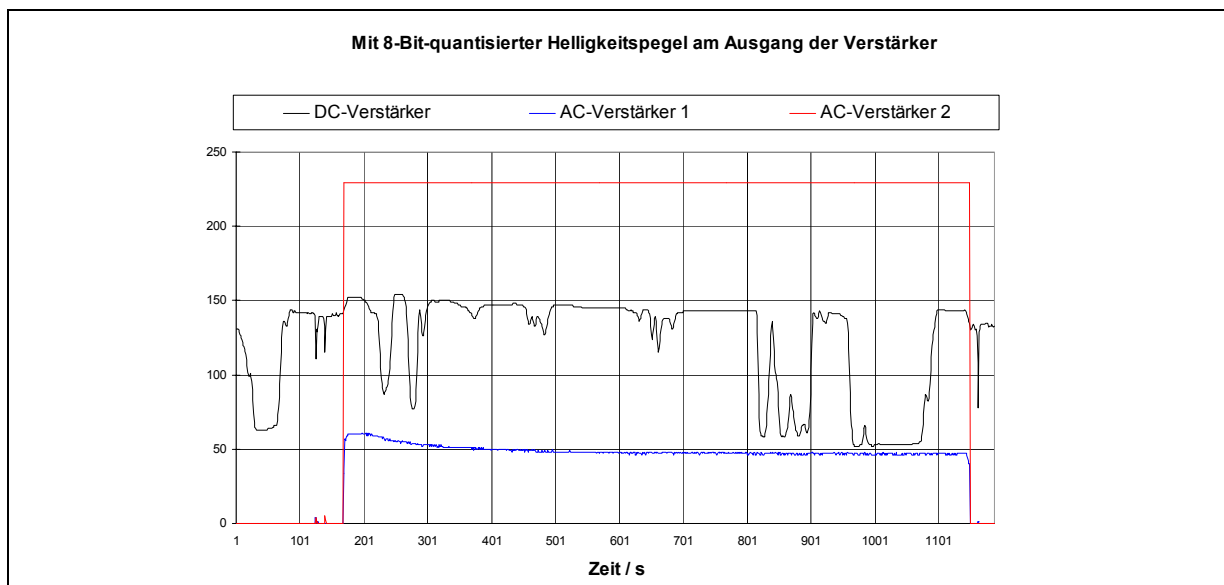


Abb. 2.26: Ein- und Ausschalten der Raumbeleuchtung (Leuchtstofflampen mit KVG) bei gleichzeitig stark variierendem Tageslichtanteil

Abb. 2.26 zeigt die mit einem 8-Bit-AD-Wandler digitalisierten Ausgangsspannungen des DC-Verstärkers und der beiden AC-Kanäle bei stark schwankendem Tageslichtpegel. Nach ca. 180 s wurden zwei 36 W-Leuchtstofflampen mit konventionellem Vorschaltgerät (KVG) eingeschaltet, die automatische Abschaltung erfolgte nach etwa 1160 s infolge eines zu hohen Tageslichtanteils. Aufgrund der hohen Wechselspannungsverstärkung kann das von den Kunstlichtquellen hervorgerufene Lichtflimmern problemlos von den Tageslichtschwankungen diskriminiert werden.

Abb. 2.27 zeigt das Blockschaltbild eines intelligenten, berührungslos arbeitenden Lichtschalters, der aus der Kombination des Reflex-tasters nach Abschnitt 2.3.3.1 und des gerade beschriebenen Lux-Sensors zur Unterscheidung von Tageslicht- und Kunstlicht hervorgeht.

Mit Hinblick auf eine möglichst kostengünstige Realisierung dieses Sensorkonzepts bietet sich die Integration der optischen Empfänger, der analogen Verstärkerstufen, der Leistungsregelung und des AD-Wandlers in CMOS-Technologie an, um die Anzahl der in den diskreten Aufbauten verwendeten Bauteile zu minimieren. Preiswerte Mikrocontroller mit einer für die vorliegende Aufgabenstellung ausreichenden Rechenleistung stehen bereits heute zur Verfügung.

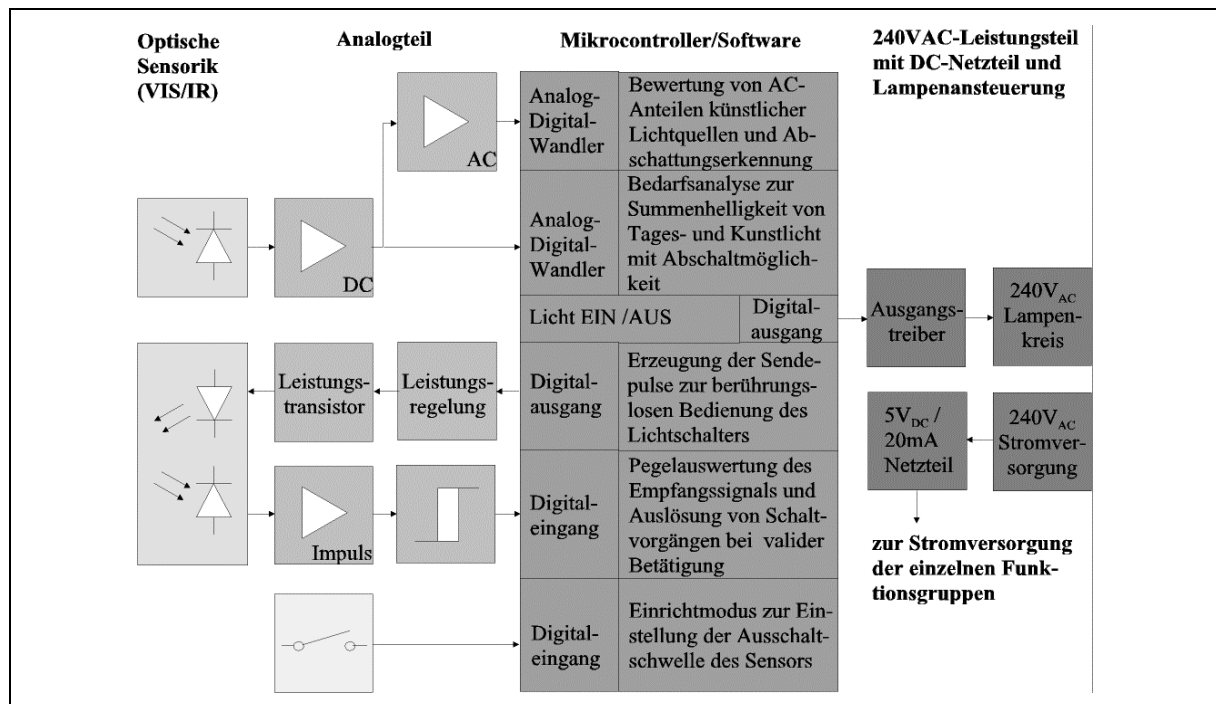


Abb. 2.27: Kombination des Reflex-tasters nach Abschnitt 2.3.3.1 und des Helligkeitssensors nach Abschnitt 2.3.3.2

2.3.4 Photomischdetektoren zur Messung der Laufzeit des Lichts

2.3.4.1 Funktionsweise eines Photogate-PMD

Photomischdetektoren (PMD's) sind neuartige optoelektronische Empfangselemente, die zur Messung der Phasenlaufzeit von modulierten Lichtwellen eingesetzt werden können. Das eigentliche PMD-Element wird zusammen mit der pixelnahen Ansteuer- und Ausleseelektronik auf einem CMOS-Chip integriert. Die Möglichkeit, eine Vielzahl von parallelen Entfernungsmessern in Zeilen-, Doppelzeilen- oder Matrixform auf einem winzigen CMOS-Chip anzuordnen, eröffnet völlig neue Perspektiven für die gesamte Automatisierungstechnik.

Bevor einzelne Anwendungen in der Gebäudetechnik aufgezeigt werden, soll die prinzipielle Funktionsweise des PMD anhand von Abb. 2.28 und Abb. 2.29 vorgestellt werden. Das optische Eingangssignal $P_{opt}(t)$ wird im PMD mit dem elektrischen Gegentakt-Bezugssignal $\pm u_m(t)$, das mit einer Phasenverschiebung von 180° an den Modulationsgates am und bm nach Abb. 2.29a angelegt wird, gemischt. Die optisch transparenten Modulationsgates werden mit einer Vorspannung von ca. -0.5 V bei einer Modulationssignalamplitude unterhalb von 1 V betrieben.

Im Anschluss an den Mischprozess werden die Mischprodukte auf den Sperrschichtkapazitäten der in Abb. 2.29b dargestellten Ausledioden integriert und an den gepufferten PMD-Ausgängen als niederfrequente Spannungen U_a und U_b zur Verfügung gestellt. Vor Beginn des Messzyklus werden die Integrationskapazitäten über ein gemeinsames Resetsignal auf einen positiven Spannungspegel aufgeladen.

Abb. 2.29c/d verdeutlichen den Transport der photogenerated Ladungsträger in Abhängigkeit der Polarität der Gegentaktspannung an den Modulationsgates sowie den Akkumulationsprozess auf den Sperrschichtkapazitäten.

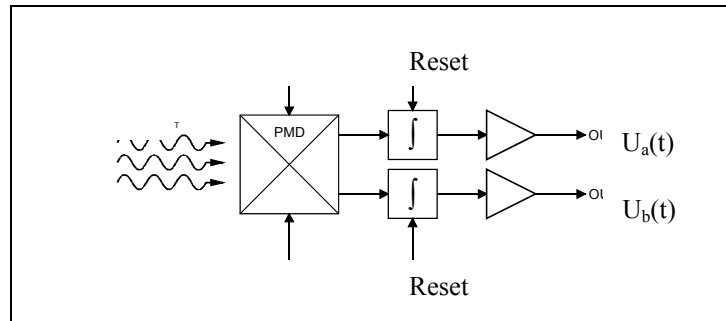


Abb. 2.28: Blockschaltbild zur Funktion des PMD

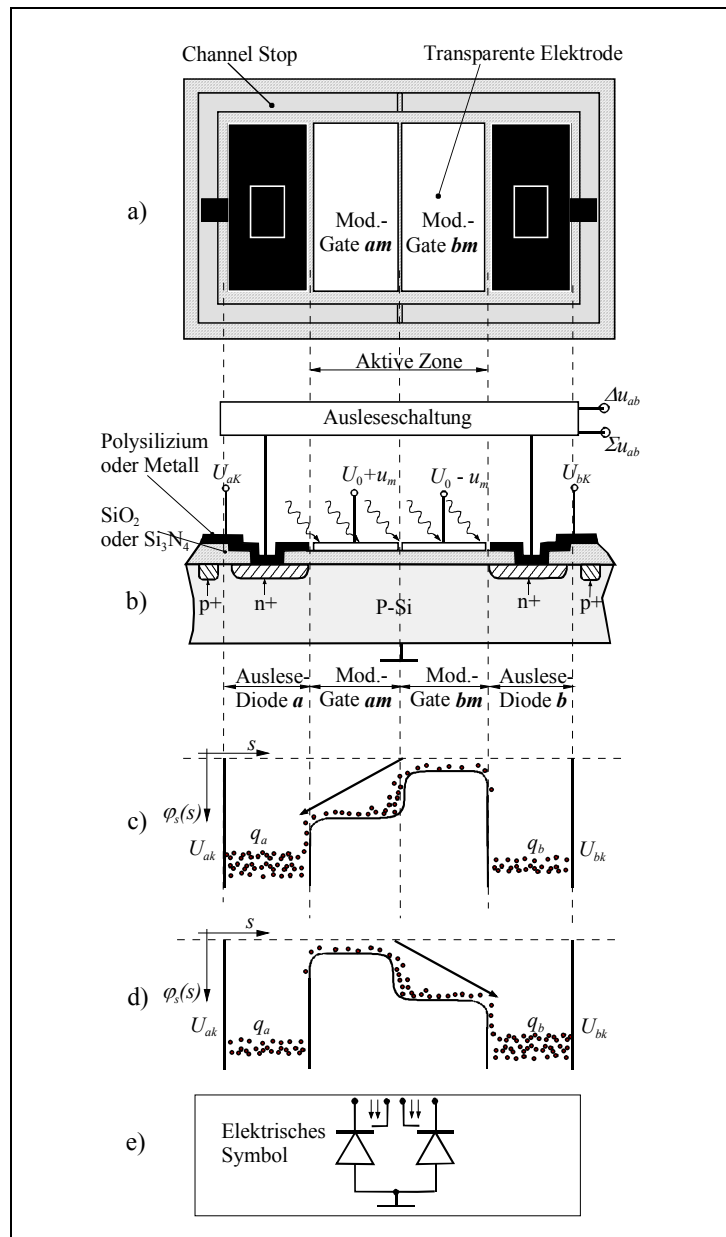


Abb. 2.29: Schematischer Aufbau, Potentialverläufe und Symbol eines Photomischdetektors

Die von der moduliert einfallenden Lichtleistung $P_{\text{opt}}(t)$ erzeugten Elektronen werden im Takt der Modulationsspannungen auf die Integrationskapazitäten verteilt. Durch die Mischung mit anschließender Integration erfüllt das PMD die Funktion eines Korrelators. Das Mischprodukt erhält man aus der Differenz $U_a - U_b$ der Ausgangsspannungen, während die Summe $U_a + U_b$ den von konventionellen optischen Sensoren erfassten Intensitätswert der Strahlung liefert.

Bereits an dieser Stelle wird klar, dass ein PMD unabhängig vom farblichen Kontrast zwischen dem Hintergrund und dem Vordergrund arbeitet. Die aus den Mischprodukten gebildeten Entfernungsdaten ermöglichen eine zuverlässige Trennung auch bei verschwindendem Kontrast. Bewegte Schatten oder bewegte Lichtquellen wie vorbeifahrenden Autos, die bei einer intensitätsauswertenden Bildverarbeitung als Grauwertänderungen erkannt werden, beeinflussen die Bewegungserkennung mit PMD-Sensoren nicht.

2.3.4.2 PMD-Grundstrukturen und Kenndaten

Abb. 2.30 und Abb. 2.31 zeigen zwei wichtige PMD-Grundstrukturen. Während das 2-Gate-PMD insbesondere für höherauflösende bildgebende Sensoren vorgesehen ist, ermöglicht die Fingerstruktur die Realisierung großer lichtempfindlicher Flächen unter Beibehaltung der hohen Modulationsbandbreite. Letztere dient daher vorwiegend zur Realisierung von PMD-Einzelpixeln, Zeilenanordnungen und Kleinarrays.

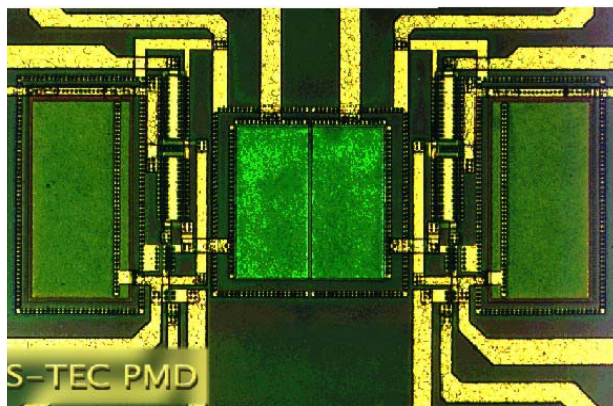


Abb. 2.30: 2-Gate-PMD-Grundstruktur

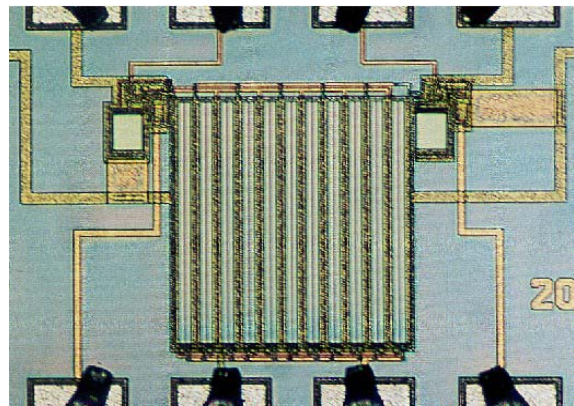


Abb. 2.31: PMD-Fingerstruktur

Mit optimierten PMD-Strukturen wurden im Labor bisher 3dB-Grenzfrequenzen bis 80 MHz nachgewiesen. Abb. 2.32 zeigt die Abhängigkeit des Modulationskontrasts des älteren PMD-Pixels STP17.5-2F-SG von der Modulationsfrequenz bei einem mittleren Photostrom von 3 nA. Die Grenzfrequenz dieses Pixels liegt bei ca. 30 MHz. Die Definition des in Abb. 2.32 verwendeten Modulationskontrasts C ergibt sich aus der PMD-Korrelationskennlinie, die in Abb. 2.34 für zwei unterschiedliche Modulationsfrequenzen dargestellt ist. Allgemein gilt

$$(Gl. 2.16) \quad C = \frac{U_{\max} - U_{\min}}{U_{\max} + U_{\min}} = \frac{1}{2} \cdot m_s \cdot m_E$$

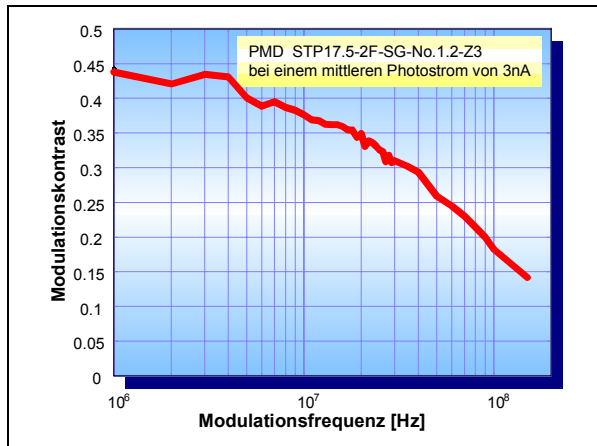


Abb. 2.32: Modulationskontrast in Abhängigkeit der Modulationsfrequenz

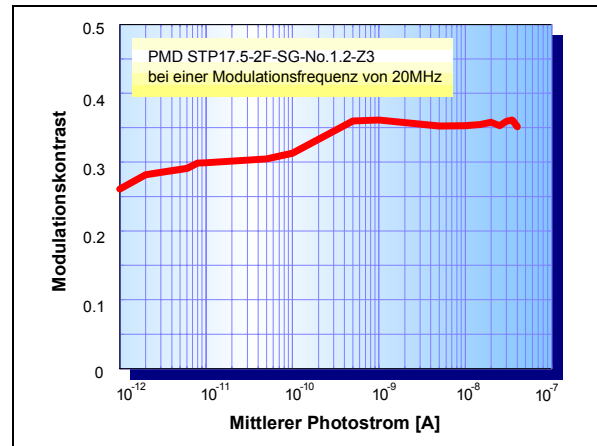


Abb. 2.33: Modulationskontrast als Funktion des mittleren Photostromes

Die Variablen m_s und m_e in (Gl. 2.16) sind die Modulationsgrade des Senders und des Empfängers. Mit dem Offset M_0 bzw. E_0 und den Modulationsamplituden M_m und E_m erhält man die Beziehungen

$$(Gl. 2.17) m_s = \frac{M_m}{M_0} \quad (Gl. 2.18) m_e = \frac{E_m}{E_0}$$

Selbst wenn Sender und Empfänger zu 100% durchmoduliert sind, ergibt sich nur der Kontrast $C = 50\%$. Nach Abb. 2.33 ist der Kontrast auch vom Photostrom abhängig.

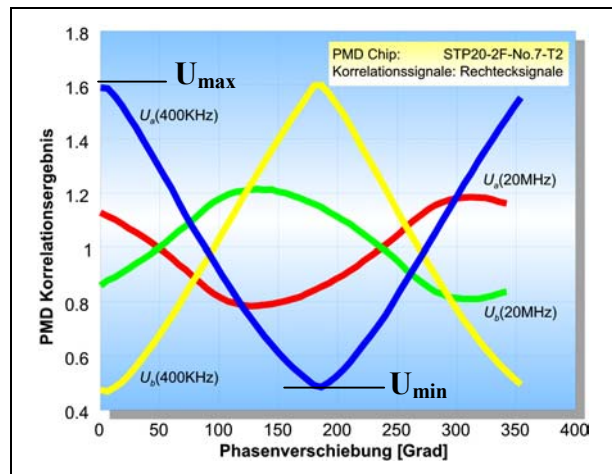


Abb. 2.34: Typische Korrelationskennlinie

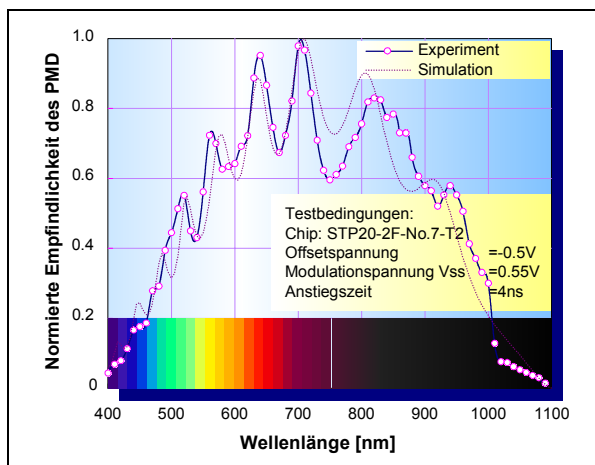


Abb. 2.35: Verlauf der normierten spektralen Empfindlichkeit ($S_{max} \approx 0.25A/W$)

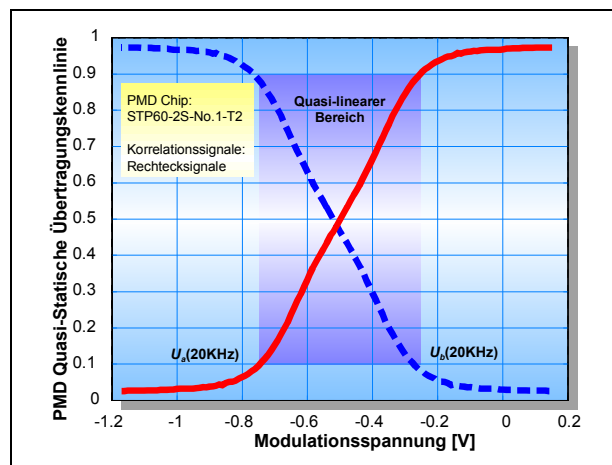


Abb. 2.36: Transferkennlinien $T_a(u_{am}(t))$ und $T_b(u_{bm}(t))$ beider PMD-Ausgänge

Abb. 2.35 zeigt die spektrale Empfindlichkeit eines PMD-Sensors. Der wellenförmige Verlauf entsteht infolge von Interferenzen in den verschiedenen Halbleiterschichten. Aus dieser Kurve lässt sich die optimale Strahlungswellenlänge des Senders entnehmen.

Das Eingangssignal des PMD ist ein optisches Signal, welches ein elektrisches Ausgangssignal hervorruft. Die in Abb. 2.37 abgebildeten Transfercharakteristiken $T_a(f_m, u_{am}(t))$ und $T_b(f_m, u_{bm}(t))$ beider PMD-Ausgänge, die eine Funktion der Modulationsfrequenz f_m sowie der Steuerspannungen $u_{am}(t)$ bzw. $u_{bm}(t)$ sind, wurden folgendermaßen definiert:

$$(Gl. 2.19) \quad T_a(f_m, u_{am}(t)) = \frac{1}{2A \cdot e \cdot \int G(x) dx} \cdot \left\langle \frac{\partial q_a(u_{am}(t))}{\partial t} \right\rangle$$

$$(Gl. 2.20) \quad T_b(f_m, u_{bm}(t)) = \frac{1}{2A \cdot e \cdot \int G(x) dx} \cdot \left\langle \frac{\partial q_b(u_{bm}(t))}{\partial t} \right\rangle$$

In (Gl. 2.19) und (Gl. 2.20) berücksichtigt A die Fläche eines Modulationsgates, die Konstante e ist die Elementarladung und $G(x)$ beschreibt die Generationsrate der Ladungsträger im aktiven Bereich. Der Ausdruck im Nenner beider Gleichungen entspricht daher der Gesamtzahl der in der Zeiteinheit ∂t unter beiden Modulationsgates erzeugten Ladungsträger. Der Term $\langle \partial q_a / \partial t \rangle$ bzw. $\langle \partial q_b / \partial t \rangle$ in (Gl. 2.19) und (Gl. 2.20) berücksichtigt den Erwartungswert der mittleren Konzentrationsänderung der Ladungsmenge unter dem jeweiligen Modulationsgate.

Infolge der Normierung der mittleren Konzentrationsänderung der Ladungen auf die insgesamt generierte Ladungsmenge sind die Transfercharakteristiken einheitenlose Größen. Die ebenfalls einheitenlose Übertragungskennlinie $T(t)$ des PMD-Empfängers erhält man aus Differenz $T_a(t) - T_b(t)$ beider Transferkennlinien.

Im folgenden wird davon ausgegangen, dass am Empfänger eine zu den Modulationsspannungen um die Laufzeit τ_L phasenverschobene Strahlungsleistung $P(t - \tau_L)$ auftritt. Für die mittleren Ströme I_a bzw. I_b auf die Integrationskapazitäten ergeben sich mit der spektralen Sensorempfindlichkeit S dann die Beziehungen

$$(Gl. 2.21) \quad I_a = S \cdot \int_0^{T_{int}} P(t - \tau_L) \cdot T_a(t) dt \quad \text{und} \quad (Gl. 2.22) \quad I_b = S \cdot \int_0^{T_{int}} P(t - \tau_L) \cdot T_b(t) dt .$$

Für die mittleren PMD-Ausgangsspannungen U_a und U_b auf den Integrationskapazitäten C_{int} folgt daraus

$$(Gl. 2.23) \quad U_a = S \cdot \frac{T_{int}}{C_{int}} \cdot \int_0^{T_{int}} P(t - \tau_L) \cdot T_a(t) dt \quad \text{sowie}$$

$$(Gl. 2.24) \quad U_b = S \cdot \frac{T_{int}}{C_{int}} \cdot \int_0^{T_{int}} P(t - \tau_L) \cdot T_b(t) dt .$$

2.3.4.3 Entfernungsmessung mit CW-Verfahren

Die im nächsten Abschnitt dargestellten Entfernungsmessungen wurden mit der Modulationsfrequenz $f_m = 20$ MHz durchgeführt. In diesem Fall ergibt sich der Eindeutigkeitsbereich

$$(Gl. 2.25) \quad Z_{\max} = \frac{\lambda_m}{2} = \frac{c}{2 \cdot f_m} = \frac{3 \cdot 10^8 \text{ m/s}}{2 \cdot 20 \cdot 10^6 \text{ Hz}} = 7.50 \text{ m}$$

für Entfernungsmessung. Darin bedeutet λ_m die Modulationswellenlänge der inkohärenten optischen Strahlung, die Konstante c ist die Lichtgeschwindigkeit.

Das Grundprinzip der Entfernungsmessung soll im weiteren Verlauf anhand des in Abb. 2.37 dargestellten CW-Verfahrens erläutert werden. Die von dem Sender ausgestrahlten Lichtwellen werden an dem Target im Abstand R reflektiert und Erreichen nach nochmaligem Durchlaufen der Strecke R den Empfänger. Für die Laufzeit τ_L der Welle gilt in diesem Fall die Beziehung

$$(Gl. 2.26) \quad \tau_L = \frac{2 \cdot R}{c} = \frac{2 \cdot R}{\lambda_m \cdot f_m}.$$

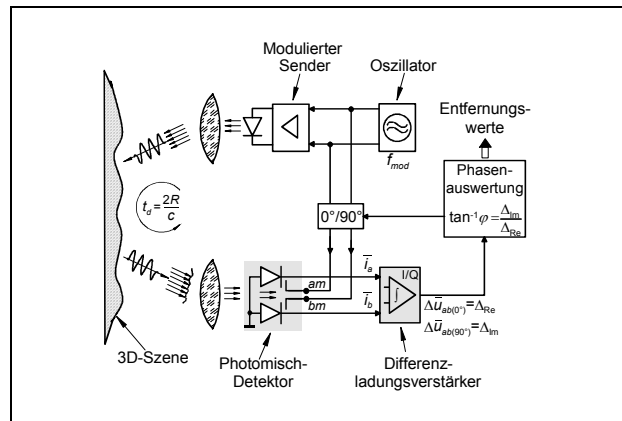


Abb. 2.37: Entfernungsmessung mit dem CW-Verfahren

Die Lichtlaufzeit τ_L ergibt die Phasenverschiebung $\varphi_L = \omega_m \cdot \tau_L = 2\pi \cdot f_m \cdot \tau_L$ der am PMD-Empfänger auftreffenden Strahlungsleistung $P_e(t) = D \cdot P_s(t - \tau_L)$, bezogen auf das Sendesignal $P_s(t)$. Der Faktor D berücksichtigt dabei die Entfernungsabhängigkeit der Strahlungsleistung $P_s(t)$ sowie Dämpfungen im Übertragungsweg. Der Reflektionsfaktor des Targets geht ebenfalls an dieser Stelle ein. Bei harmonischer Modulation des Senders folgt für die Empfangsleistung

$$(Gl. 2.27) \quad P_e(t) = D \cdot P_s(t - \tau_L) = D [P_0 + P_m(t - \tau_L)] = D [P_0 + P_m \cdot \cos(\omega_m t - \varphi_L)].$$

Für eine ebenfalls harmonische Anregung gleicher Frequenz an den Modulationsgates gilt

$$(Gl. 2.28) \quad u_{am}(t) = U_0 + u_m(t) = U_0 + U_m \cdot \cos(\omega_m t - \varphi_0) \quad \text{und}$$

$$(Gl. 2.29) \quad u_{bm}(t) = U_0 - u_m(t) = U_0 - U_m \cdot \cos(\omega_m t - \varphi_0).$$

Arbeitet man im linearen Bereich der PMD-Übertragungskennlinie, ergeben sich unmittelbar die Transfercharakteristiken der PMD-Ausgänge:

$$(Gl. 2.30) \quad T_a(t) = T_0 + T_m(t) = T_0 + T_m \cdot \cos(\omega_m t - \varphi_0) \quad \text{und}$$

$$(Gl. 2.31) \quad T_b(t) = T_0 - T_m(t) = T_0 - T_m \cdot \cos(\omega_m t - \varphi_0).$$

Für die PMD-Ausgangsspannungen U_a und U_b gelten dann unter Beachtung von (Gl. 2.23) und (Gl. 2.24) K die Beziehungen

$$(Gl. 2.32) \quad U_a = S \cdot \frac{D \cdot T_{\text{int}}}{C_{\text{int}}} \cdot \int_0^{T_{\text{int}}} [P_0 + P_m \cdot \cos(\omega_m t - \varphi_L)] \cdot [T_0 + T_m \cdot \cos(\omega_m t - \varphi_0)] dt \quad \text{und}$$

$$(Gl. 2.33) \quad U_b = S \cdot \frac{D \cdot T_{\text{int}}}{C_{\text{int}}} \cdot \int_0^{T_{\text{int}}} [P_0 + P_m \cdot \cos(\omega_m t - \varphi_L)] \cdot [T_0 + T_m \cdot \cos(\omega_m t - \varphi_0)] dt ,$$

woraus sich nach der Integration über die Zeit T_{int} die entfernungsabhängigen Ausgangsspannungen nach (Gl. 2.34) und (Gl. 2.35) ableiten lassen. Dabei wurde berücksichtigt, dass die Multiplikation der mittleren Ströme I_0 bzw. I_m mit der Integrationszeit T_{int} die Ladungsmengen $Q_0 = I_0 T_{\text{int}}$ und $Q_m = I_m T_{\text{int}}$ liefert.

$$(Gl. 2.34) \quad U_a = \frac{S \cdot D}{C_{\text{int}}} \cdot \left[P_0 Q_0 + \frac{P_m Q_m}{2} \cdot \cos(\varphi_0 - \varphi_L) \right] \quad \text{und}$$

$$(Gl. 2.35) \quad U_b = \frac{S \cdot D}{C_{\text{int}}} \cdot \left[P_0 Q_0 - \frac{P_m Q_m}{2} \cdot \cos(\varphi_0 - \varphi_L) \right] .$$

Mit (Gl. 2.34) und (Gl. 2.35) sind zwei Gleichungen zur Bestimmung der drei unbekanntenen Terme $P_0 \cdot Q_0$, $P_m \cdot Q_m$ und $\varphi_L - \varphi_0$ vorhanden. Zwei weitere Gleichungen ergeben sich, wenn eine weitere Messung, beispielsweise mit einer zusätzlichen 90° -Verschiebung zwischen dem Sende- und Empfangspfad, durchgeführt wird. Unter Beachtung der gegenphasigen Ansteuerung der Modulationsgates gilt mit der zusätzlichen Phasenverzögerung φ_k allgemein die Beziehung

$$(Gl. 2.36) \quad U_k = \frac{K}{C_{\text{int}}} \left[P_0 Q_0 + \frac{P_m Q_m}{2} \cos(\varphi_0 - \varphi_L + \varphi_k) \right] \quad \text{mit} \quad \varphi_k = \frac{2\pi}{N} k \quad \text{und} \quad k = 1, 2, \dots, N .$$

Bei insgesamt 4 Phasen mit einer Differenz von jeweils 90° , wobei zwei um 180° versetzte Phasen jeweils gleichzeitig gemessen werden, erhält man nach einigen Umformungen und mit der Annahme $\varphi_0 = 0$ eine einfache Gleichung zur Entfernungsbestimmung:

$$(Gl. 2.37) \quad R = \frac{c \cdot \tau_L}{2} = \frac{c \cdot \varphi_L}{4\pi f_m} = \frac{c}{4\pi f_m} \arctan \left(- \frac{U_1 - U_3}{U_2 - U_4} \right)$$

In (Gl. 2.37) fällt unmittelbar auf, dass der konstante Offset nach (Gl. 2.36) bei der Differenzbildung entfällt. Dieser Umstand sollte aber nicht darüber hinwegtäuschen, dass sich ein zusätzlicher Fremdlichtanteil nachteilig auf den Dynamikbereich des PMD-Empfängers auswirkt. Aus diesem Grund enthalten aktuelle PMD-Empfänger eine SBI-Schaltung (Suppression of Background Illumination), die eine automatische Gleichlichtunterdrückung bewirkt. Neuere Auslesetechniken und neue PMD-Konzepte lassen hier zusätzliche Verbesserungen erwarten.

Die zusätzliche Phase φ_0 entsteht aufgrund von zusätzlichen Phasenverschiebungen im Übertragungsweg. Sie ist von den Betriebsbedingungen abhängig und bewirkt daher nur in Ausnahmefällen einen konstanten Offsetfehler in den Entfernungsdaten. Zur Kompensation sind daher geeignete Referenzierungskonzepte erforderlich.

2.3.4.4 Systemkonzept für PMD-basierte Entfernungsmessgeräte

Abb. 2.38 zeigt ein erstes PMD-basiertes 1D-Entfernungsmessgerät, dessen elektronische Komponenten in Abb. 2.39 abgebildet sind. Neben dem Optikkblock, der die Sende- und Empfangsoptiken einschließlich der PMD-Frontend-Platine und dem Lasersender enthält, besteht das System aus einer CPLD-Platine, einer Mikrocontrollerplatine (MC) mit C164-Mikrocontroller sowie einer kombinierten Netzteil- und Schnittstellenplatine (DC). Neben einem RS232-Interface stehen eine CAN-Schnittstelle, ein digitaler 24V-Triggereingang und drei 24V-Schaltausgänge zur Verfügung.



Abb. 2.38: PMD-basiertes 1D-Entfernungsmessgerät

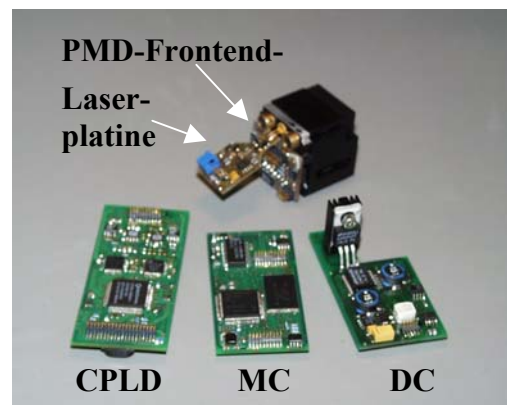


Abb. 2.39: Elektronische Komponenten des 1D-Entfernungsmessgeräts

Das Blockschaltbild in Abb. 2.40 verdeutlicht die Funktionalität der einzelnen Baugruppen sowie das Zusammenspiel der optischen und elektronischen Komponenten. Zur Distanzmessung wird moduliertes Licht verwendet, wobei die von der Monitordiode des Lasers empfangene Intensität zur Überwachung und zur Stabilisierung der insgesamt emittierten Strahlungsleistung dient. Die ausgesendete Laserstrahlung wird zunächst von einer Kollimatorlinse fokussiert. In Verbindung mit PMD-Zeilensensoren und PMD-Matrixsensoren erfolgt anschließend die Aufweitung des Lichtflecks mit einer diffraktiven Optik.

Mit Hinblick auf die optische Leistungsbilanz wird nur das unmittelbar auf den PMD-Empfänger abgebildete Objektfeld ausgeleuchtet. Die geometrischen Abmessungen des PMD-Sensors und die Brennweite der verwendeten Empfangslinse bestimmen die erforderliche Strahlaufweitung, während die Apertur der Empfangslinse die auf dem PMD-Chip tatsächlich vorhandene Bestrahlungsstärke beeinflusst. Im Empfangspfad wurde zusätzlich ein

Rotfilter und ein Sperrfilter für den nahen Infrarot-Bereich (NIR) vorgesehen. Die Transmissionseigenschaften des so realisierten Bandfilters sind auf die Laserwellenlänge abgestimmt.

Zur Ansteuerung des Lasersenders und des SoC-PMD-Empfängers wird ein programmierbarer Logikbaustein verwendet. Während zur Erzeugung der sende- und empfangsseitigen Modulationssignale ein CPLD (Complex Logical Device) ausreicht, erfordert die für zukünftige Anwendungen vorgesehene, hardwaregestützte Berechnung der Entfernungsinformationen die Implementierung in einem FPGA (Field Programmable Gate Array).

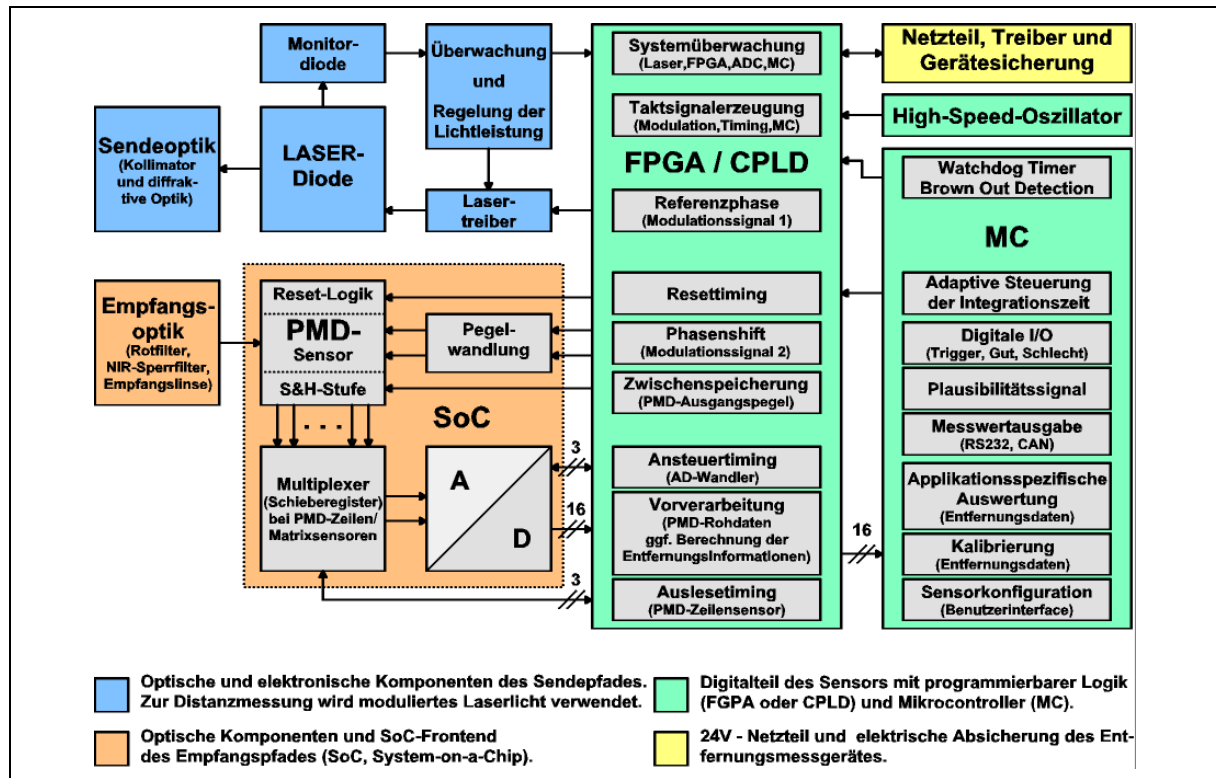


Abb. 2.40: Systemkonzept eines PMD-basierten Entfernungsmessgerätes

In der aktuellen Version des Entfernungsmessgeräts übernimmt in der Mikrocontroller die Steuerung des gesamten Messablaufs sowie alle anfallenden Berechnungen. Außerdem stellt er einen UART und ein CAN-Interface zur Verfügung. Ausgangsseitig müssen daher lediglich die Treiber zur Pegelanpassung an die Spezifikationen der RS232-Schnittstelle und des CAN-Busses vorgesehen werden.

2.3.4.5 Entfernungsmessergebnisse mit einem 1D-PMD-Sensor

Mit dem 1D-Messgerät aus Abb. 2.38 wurden die in Abb. 2.41 und Abb. 2.42 dargestellten Messergebnisse erzielt. Dabei dienen digitale Modulationssignale zur Ansteuerung des Lasersenders und des PMD-Empfängers. Zur Auswertung der Rohdaten wurde allerdings kein 4-Phasen-Algorithmus, sondern ein komplexerer Algorithmus verwendet, der bessere Eigenschaften bei der Unterdrückung von Nichtlinearitäten und bei der Filterung von Oberwellen bietet.

Die Messungen erfolgten mit einer Laserleistung von 1 mW bei einer Wellenlänge von 658 nm. Der Laser und das PMD-Element wurden digital mit einer Frequenz von 20 MHz moduliert. Als Target dienten die in Tab. 2.1 aufgeführten Pappen. Der Reflektionsfaktor der grauen, braunen und schwarzen Pappe wurde auf den der weißen Pappe, der die Remission 100% zugeordnet wurde, normiert.

Targets zur Entfernungsmessung	Reflektivität (@ 658 nm)
Weißer Pappe	100 %
Graue Pappe	33 %
Braune Pappe	27 %
Schwarze Pappe	8 %

Tab. 2.1: Targets mit unterschiedlichen Reflektionsfaktoren

Zur Kalibrierung des Demonstrators wurde zunächst die Korrekturkennlinie $R_S = f(R_M)$ bestimmt, welche den Sollwert R_S in Abhängigkeit des tatsächlichen Messwerts R_M angibt. Zur Festlegung der Korrekturkennlinie wurden für jede gemessene Entfernung (0.2m, 0.4m...3.0m) alle Einzelmessungen mit den Reflektionsfaktoren 100%, 33%, 27% und 8% gemittelt. Dadurch erhält man die Erwartungswerte der Entfernungsmessung bei einem mittleren Remissionsfaktor. Statistische Messfehler werden infolge der Mittelung über jeweils 400 Messwerte weitgehend eliminiert, während systematische Abweichungen erhalten bleiben.

Ausgehend von diesen Mittelwerten wurde die Regressionsgerade (Gl. 2.38), welche gleichzeitig die unkorrigierte Sensorkennlinie $R_M = f(R_S)$ darstellt, nach der Methode der kleinsten Fehlerquadrate ermittelt.

$$(Gl. 2.38) \quad \frac{R_M}{[m]} = \left(0.9917 \cdot \frac{R_S}{[m]} - 0.0451 \right) \quad (Gl. 2.39) \quad \frac{R_S}{[m]} = \left(1.0084 \cdot \frac{R_M}{[m]} + 0.0455 \right)$$

Die Achsenspiegelung von (Gl. 2.38) an der idealen Sensorkennlinie $R_M = R_S$ liefert die gewünschte Korrekturkennlinie $R_S = f(R_M)$ nach (Gl. 3.39), welche zur Ermittlung der korrigierten Entfernungsmesswerte in Abb. 2.42 mit den Standardabweichungen aus Abb. 2.41 verwendet wurde. Betrachtet man (Gl. 3.38) bzw. (Gl. 3.39), fällt zunächst auf, dass nahezu keine Korrektur der Kennliniensteigung erforderlich ist. Anders verhält es sich mit dem Kennlinienoffset, der infolge einer zusätzlichen Phasenverschiebung im Messgerät entsteht. Dieser Offset ist wie vorher schon erläutert wurde, im allgemeinen nicht konstant, sondern er hängt von den Betriebsbedingungen des Geräts ab. Für hochgenaue Entfernungsmessungen sind daher geeignete Referenzierungskonzepte erforderlich, um die geforderte Genauigkeit selbst bei unterschiedlichsten Umgebungsbedingungen einzuhalten.

Mit Hinblick auf Anwendungen in der Gebäudetechnik kann festgestellt werden, dass für viele Aufgabenstellungen keine hochgenauen Absolutmessungen erforderlich sind. Vielmehr sind Entfernungsänderungen, die auch mit einem nicht offsetkorrigierten Sensor gemessen werden können, für eine Bewegungsdetektion völlig ausreichend.

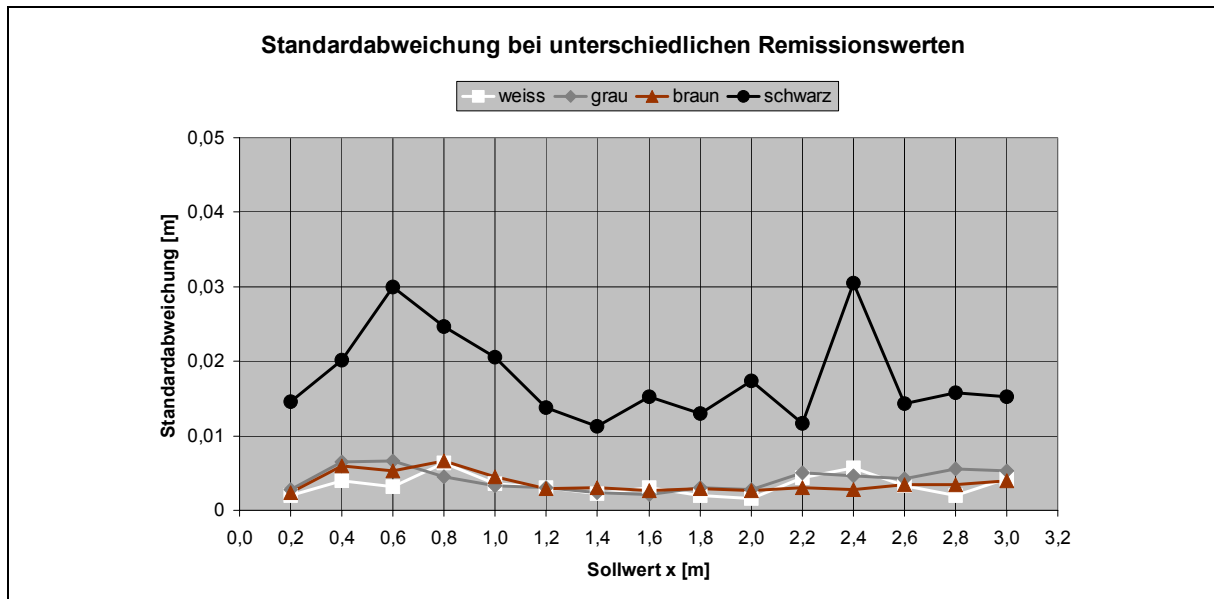


Abb. 2.41: Standardabweichung der korrigierten Entfernungsmesswerte als Funktion der Remission des Targets

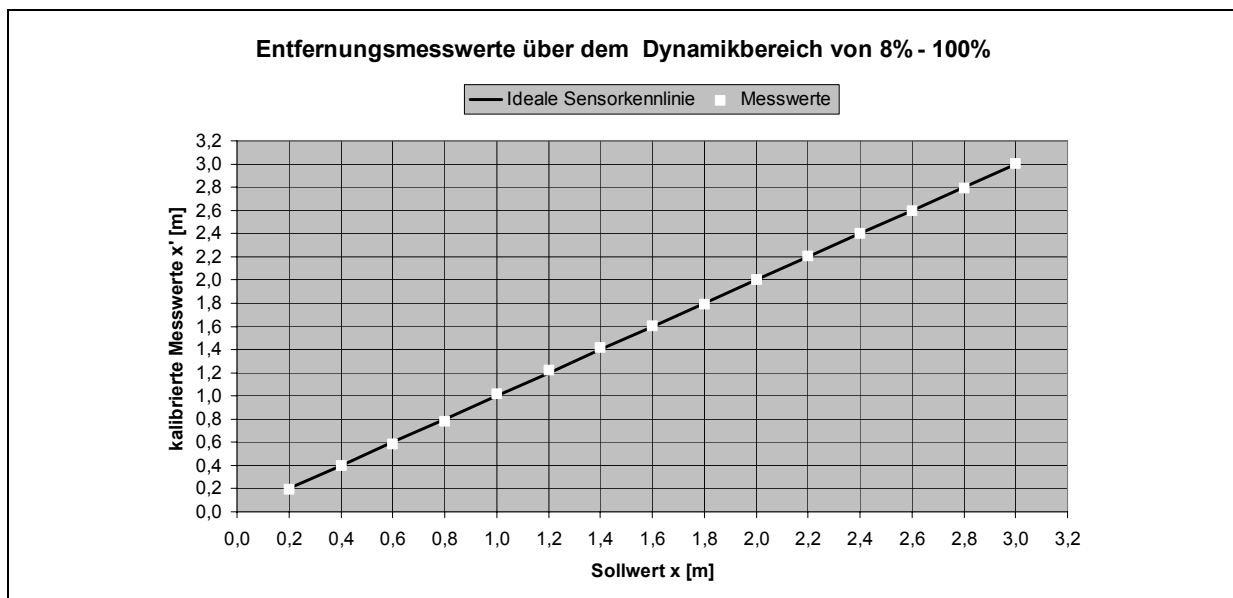


Abb. 2.42: Entfernungsmesswerte des kalibrierten 1D-PMD-Demonstrators im Dynamikbereich von 8% - 100%

2.3.4.6 Parallele Entfernungsmessung mit PMD-Arrays

Entfernungsmessende Lichtschranken mit PMD-Einzelpixeln oder Kleinarrays mit wenigen Pixeln, mit denen die Durchgangsrichtung erkannt werden kann, stellen keine besonderen Ansprüche an die Signalauswertung. Für höherauflösende Zeilensensoren oder Matrixsensoren steigt der Berechnungsaufwand mit der Pixelzahl entsprechend an. Abb. 2.43 und Abb. 2.44 zeigen Beispiele für PMD-Empfänger mit mehreren parallelen, entfernungs-messenden Pixeln.

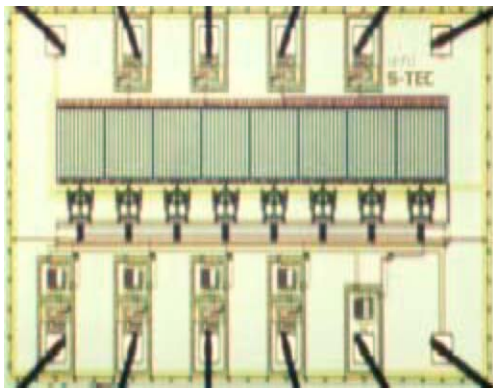


Abb. 2.43: Zeilensensor mit PMD-Fingerpixel-Strukturen

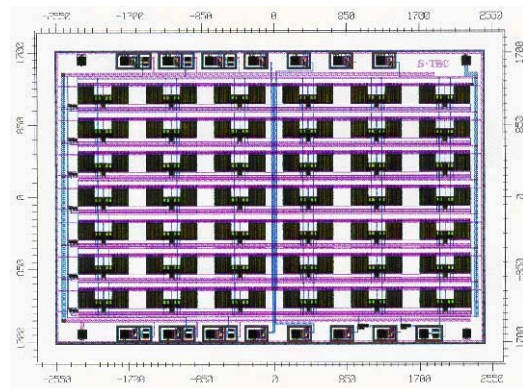


Abb. 2.44: PMD-Matrix mit 12x7 Pixeln

Zur Verifikation der parallelen Entfernungsmessung wurde im Labor der in Abb. 2.45 abgebildete T-förmige Prüfkörper mit dem Matrixsensor aus Abb. 2.44 vermessen. Während die laterale Auflösung des Sensors für komplexe 3D-Szenen nicht ausreicht, zeigen die Messergebnisse, dass einfache Messobjekte durchaus erkannt werden können.

Die Verschleifung der Kanten resultiert daraus, dass sowohl Abschnitte des Prüfkörpers als auch Teile des Hintergrunds auf dasselbe Pixel abgebildet wurden. Zur Beleuchtung der 40 cm x 50 cm großen Szene wurde ein LED-Array verwendet. Prinzipiell ist festzuhalten, dass die zur Vermessung einer Szene erforderliche Lichtleistung sich an der Größe des zu überwachenden Volumens und der minimalen Targetremission orientiert. Zur Überwachung typischer Raumgrößen in der Gebäudetechnik ist das PMD daher zwar durchaus geeignet, wegen der zur Zeit vorhandenen Sensorempfindlichkeit wäre allerdings eine hohe LED-Anzahl und ein vergleichsweise hoher Stromverbrauch erforderlich.

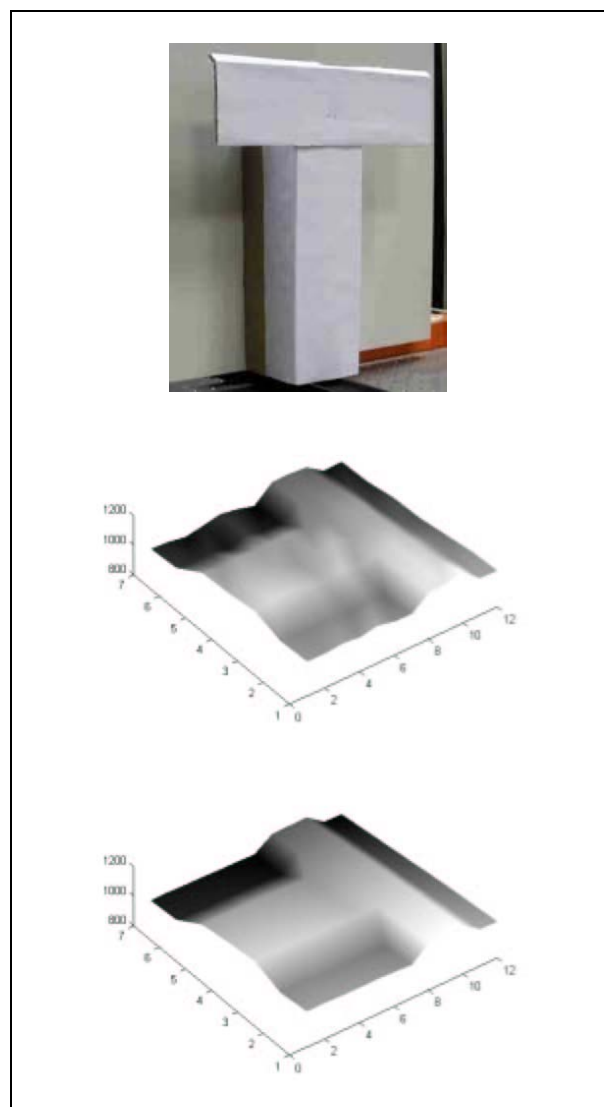


Abb. 2.45: T-förmiger Prüfkörper und ungefilterte sowie gefilterte Messwerte

Anwendungen mit eingeschränktem Gesichtsfeld, beispielsweise zur Zugangs- oder Durchgangskontrolle, zur Überwachung von Fahrstühlen etc. sind dagegen mit dem aktuellen Entwicklungsstand der PMD-Technologie durchaus realisierbar. Ein wesentlicher Faktor für das preissensitive Marktsegment der Gebäudeautomatisierung sind allerdings die Systemkosten.

Im Hinblick auf ein kostenoptimales Systemkonzept ist es wichtig, möglichst viele Komponenten auf dem PMD-Chip zu integrieren. Zukünftige hochintegrierte PMD-Chips werden daher neben dem Sensorelement und der unmittelbaren Sensorperipherie weitere Funktionsgruppen wie z.B. einen hochauflösenden AD-Wandler, eine Schaltung zur Erzeugung phasenverschobener Modulationssignale für den Sende- und Empfangspfad incl. Referenzierungseinheit und ggf. sogar die Treiber- und Regelstufen für externe Laser- oder LED-Beleuchtungen enthalten.

Aufgrund des attraktiven Preis-Leistungs-Verhältnisses moderner logischer Bausteine ist es außerdem sinnvoll, wesentliche Teile der Datenvorverarbeitung und auch der Applikationssoftware vom Mikrocontroller ins FPGA zu verlagern. In Verbindung mit dem für größere Zeilen- oder Matrixsensoren erforderlichen Rechenaufwand hat die Hardwareimplementierung nicht nur Vorteile bei den Systemkosten, sondern auch bezüglich der Auswertegeschwindigkeit und der damit verbundenen Reaktionsfähigkeit des Systems. Außerdem sollte berücksichtigt werden, dass mittlerweile IP-Cores erhältlich sind, die es ermöglichen, leistungsfähige Prozessorarchitekturen zusammen mit dem applikationsspezifischen Teil in einem FPGA zu implementieren. Dadurch ergibt sich die Möglichkeit, für viele Anwendungsfälle maßgeschneiderte Lösungen zur Verfügung zu stellen.

2.3.5 Sensoren zur Erfassung von Wärmestrahlung

Im Vergleich zu den vorher beschriebenen optischen Sensoren werden in diesem Abschnitt Detektoren zur Erfassung von Wärmestrahlung beschrieben. Tab. 2.2 zeigt eine Übersicht zu verschiedenen Typen und einige charakteristische Daten. Zu den thermischen Detektoren gehören auch die pneumatischen Zellen (Golay-Zelle und Kapazitätsmikrofon) und die Bolometer.

In Golay-Zellen dehnt sich das eingeschlossene Xenongas infolge der Erwärmung durch einfallende IR-Strahlung aus. Die daraus resultierende Druckänderung bewegt einen Spiegel im Strahlengang zwischen der integrierten Lichtquelle und einem Fotoempfänger. Das Ausgangssignal der Zelle wird von der am Empfänger auftreffenden Strahlungsleistung bestimmt. Bei den Kapazitätsmikrophonen wird die von der einfallenden Strahlung abhängige Gasausdehnung unmittelbar in eine Kapazitätsänderung umgewandelt [Lit. 2.19].

Bolometer ändern dagegen ihren Widerstand in Abhängigkeit der einfallenden Strahlung. Sie besitzen einen recht großen Temperaturkoeffizienten. Im Gegensatz zum photoelektrischen

Effekt wird die einfallende Strahlung in Bolometern zunächst in Wärme umgewandelt, die ihrerseits eine Widerstandsänderung im Material bewirkt [Lit. 2.19].

Von den in Tab. 2.2 aufgeführten thermischen Detektoren werden im folgenden allerdings nur die Thermopiles [Abschnitt 2.3.5.1] und die pyroelektrischen Detektoren [Abschnitt 2.3.5.2] näher beschrieben. Mit diesen Sensoren lassen sich einfache und vor allem kostengünstige Präsenz- und Bewegungsmelder realisieren.

Detektorarten		Detektoren	Spektrale Empfindlichkeit (um)	Betriebstemperatur (K)	Detektivität D* [cm·Hz ^{1/2} /W]	
Thermische Detektoren	Pneumatische Zelle	Golay-Zelle	Keine Abhängigkeit von der Wellenlänge	300	D*(λ,10,1) = 1 x 10 ⁹	
	Bolometer			300	D*(λ,10,1) = 1 x 10 ⁸	
	Thermoelement,-pile	TGS, LiTaO ₃		300	D*(λ,10,1) = 6 x 10 ⁸	
	Pyroelektr. Detektor			300	D*(λ,10,1) = 2 x 10 ⁸	
Quantendetektoren	Intrinsische Typen	Foto-leitende Typen	PbS	1.0 3.6	300	D*(500, 600,1)= 1 x 10 ⁹
			PbSe	1.5 5.8	300	D*(500, 600,1)= 1 x 10 ⁸
			HgCdTe	2.0 ... 16.0	77	D*(500,1000,1)= 2 x 10 ¹⁰
		Foto-voltaische Typen	Ge	0.8 1.8	300	D*(λ _p , 1000,1)= 1 x 10 ¹¹
			InGaAS	0.7 1.7	300	D*(λ _p) = 5 x 10 ¹²
			InAs	1.0 3.1	77	D*(500,1200,1)= 1 x 10 ¹⁰
	InSb		1.0 5.5	77	D*(500,1200,1)= 2 x 10 ¹⁰	
	Extrinsische Typen	HgCdTe	Ge:Au	1.0 ... 10.0	77	D*(500,900,1) = 1 x 10 ¹¹
			Ge:Hg	2.0 ... 14.0	4.2	D*(500,900,1) = 8 x 10 ⁹
			Ge:Cu	2.0 ... 30.0	4.2	D*(500,900,1) = 5 x 10 ⁹
			Ge:Zn	2.0 ... 40.0	4.2	D*(500,900,1) = 5 x 10 ⁹
			Si:Ga	1.0 ... 17.0	4.2	D*(500,900,1) = 5 x 10 ⁹
			Si:As	1.0 ... 23.0	4.2	D*(500,900,1) = 5 x 10 ⁹

Tab. 2.2: Detektoren zur Erfassung von Temperaturstrahlung [Lit. 2.19]

Generell lässt sich feststellen, dass die Empfindlichkeit aller angegebenen, thermischen Detektoren über einen großen Bereich unabhängig von der Wellenlänge ist (s. Abb. 2.46). Sie können bei Raumtemperatur betrieben werden. Im Vergleich zu den unten beschriebenen Quantendetektoren ist die Reaktionszeit auf externe Anregungen allerdings langsamer und die spezifische Detektivität D* ist um den Faktor 10 ...1000 geringer.

Die in Tab. 2.2 und in Abb. 2.46 angegebene spezifische Detektivität D* beschreibt das Signal-Rausch-Verhältnis bei einer Strahlungsleistung von 1 W. D* wird meist auf eine aktive Fläche von 1 cm² bei einer Rauschbandbreite von 1 Hz normiert und ist daher unabhängig von der Größe und der Form des Sensorelements. Zur Interpretation der Detektivitätsangabe werden wie in Tab. 2.2 gewöhnlich die Temperatur der Strahlungsquelle in [K] oder die Wellenlänge in [um] sowie die Chopperfrequenz und die Rauschbandbreite B in [Hz] aufgeführt. Je höher der D*-Wert, desto empfindlicher ist der Detektor. Allgemein gilt

$$(Gl. 2.40) \quad D^* = \frac{S/N \cdot \sqrt{B} \cdot \sqrt{A}}{P}$$

Die Detektivität von Quantendetektoren ist deutlich höher als die der thermischen Detektoren. Allerdings zeigt sich in Abb. 2.46 eine ausgeprägte Abhängigkeit von der Wellenlänge der Strahlung und von der Betriebstemperatur. Je nach Sensormaterial und gewünschter Wellenlänge ist eine Kühlung des Detektors unumgänglich.

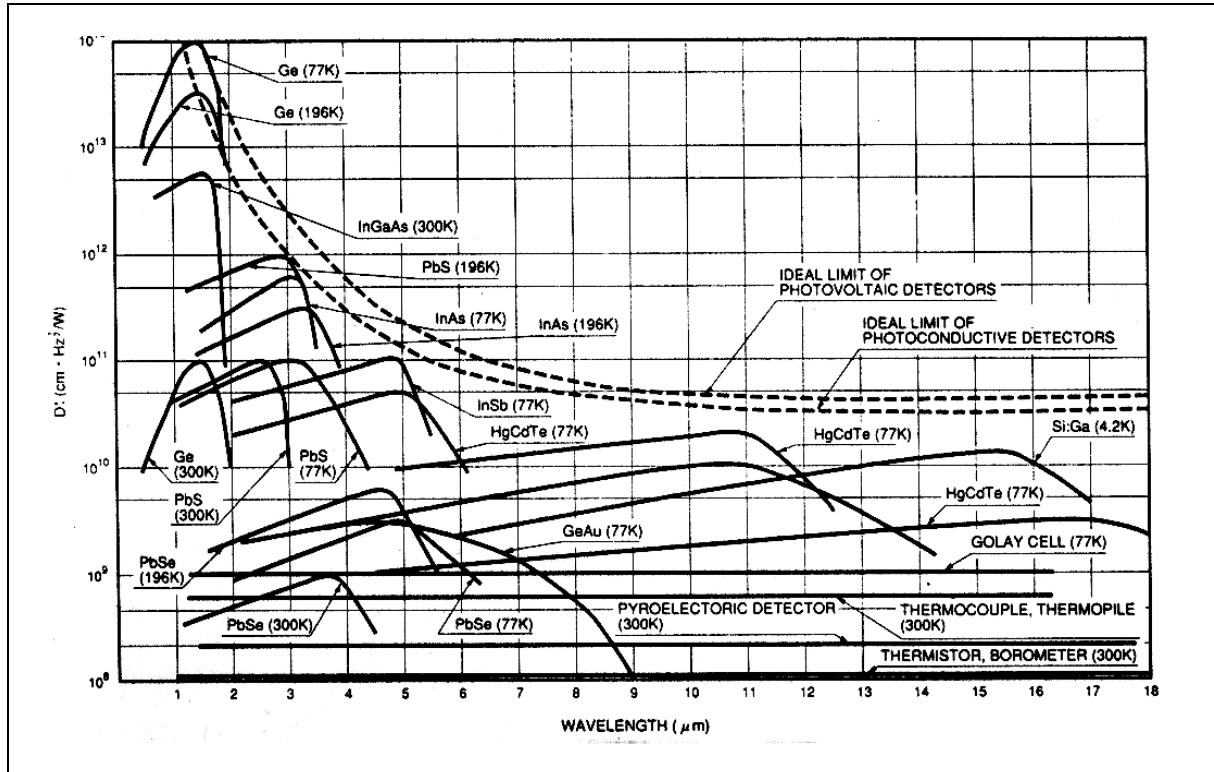


Abb. 2.46: Detektivität von Strahlungsdetektoren als Funktion der Wellenlänge [Lit. 2.19]

Bei intrinsischen Quantendetektoren werden die detektierbaren Wellenlängen vom Bandabstand des (intrinsischen) Materials bestimmt. Die Empfindlichkeit fällt oberhalb der Grenzwellenlänge steil ab. Generell unterscheidet man zwischen fotoleitenden und fotovoltaischen Detektoren. Die erstgenannten ändern ihre Leitfähigkeit in Abhängigkeit der einfallenden Strahlung, während bei den anderen ein elektrischer Strom erzeugt wird. Ternäre HgCdTe- und PbSnTe-Materialien zeichnen sich dadurch aus, dass die Grenzwellenlänge über einen großen Bereich durch das Mischungsverhältnis eingestellt werden kann [Lit. 2.19]. Die Wellenlängenbeschränkungen extrinsischer Sensoren sind von der Dotierungskonzentration der Fremdatome im Ge- oder Si-Substrat abhängig. Im Vergleich zu intrinsischen Detektoren ist die Betriebstemperatur deutlich niedriger [Lit. 2.19]. Aufgrund des zur Kühlung erforderlichen Systemaufwands im zur Personenerkennung erforderlichen Wellenlängenbereich um 10 um werden Quantendetektoren im folgenden nicht weiter berücksichtigt.

2.3.5.1 Thermoelemente und Thermopiles zur Anwesenheitskontrolle

Ein Thermoelement ist eine elektrische Verbindung zweier leitender Materialien, die unterschiedliche Werte auf der thermoelektrischen Spannungsreihe aufweisen. Tritt zwischen den

Verbindungsstellen der leitenden Materialien und den nicht bestrahlten Referenzstellen eine Temperaturdifferenz auf, wird eine für die Materialpaarung typische und zur Temperaturabweichung proportionale Thermospannung generiert (Seebeck-Effekt).

Ein Thermopile enthält typischerweise 10 ... 100 in Serie geschaltete Thermopaare, wobei die strahlungsempfindlichen "heißen" Kontakte auf einer als Trägermaterial dienenden Schwarzschrift hoher Absorptionsfähigkeit aufgebracht sind. Die "kalten" Referenzkontakte sind vor der Bestrahlung geschützt und befinden sich ohne zusätzliche Kühlung auf Umgebungstemperatur. Entsprechend der Anzahl der zusammengeschalteten Thermopaare vervielfacht sich die generierte Thermospannung, allerdings lässt sich die Empfindlichkeit von Thermopiles nicht beliebig durch eine Erhöhung der Elementanzahl steigern.

Prinzipbedingt ist das Thermopile ein Detektor, der im Gegensatz zu pyroelektrischen Sensoren auch auf Gleichlicht reagiert. Die Ausgangsspannung S ist nach (Gl. 2.41) über einen großen Wellenlängenbereich proportional zur absorbierten Strahlungsleistung P [Lit. 2.20]. Die Konstante S_R ist die Empfindlichkeit (Responsivity) des Sensors.

$$(Gl. 2.41) \quad S = S_R \cdot P$$

$$(Gl. 2.42) \quad N = \sqrt{4kTB R}.$$

Das Rauschen wird im wesentlichen durch das thermische Rauschen des Detektorwiderstandes R bestimmt und steigt mit wachsendem Sensorwiderstand an. Mit der Boltzmannkonstante k und der Rauschbandbreite B des Systems lässt sich mit (Gl. 2.42) der Effektivwert der Rauschspannung N bei der absoluten Temperatur T bestimmen. Für die Temperatur $T = 300 \text{ K}$, einen Innenwiderstand $R = 45 \text{ k}\Omega$ und bei einer Bandbreite $B = 10 \text{ Hz}$ liegt der Effektivwert in der Größenordnung von 100 nV .

Signalspannungen unterhalb des Rauschens können nicht detektiert werden. Diese Grenze beschreibt die "Noise Equivalent Power" (NEP) nach (Gl. 2.43), wobei der Kehrwert nach (Gl. 2.44) als Detektivität D bezeichnet wird. Der Zusammenhang zur spezifischen Detektivität D^* nach (Gl. 2.40) ist ebenfalls angegeben. A ist die Sensorfläche.

$$(Gl. 2.43) \quad NEP = \frac{N}{S_R \sqrt{B}}$$

$$(Gl. 2.44) \quad D = \frac{1}{NEP} = \frac{D^*}{\sqrt{A}}$$

Das Thermopile benötigt eine gewisse Zeit, um im Anschluss an eine Strahlungsänderung das thermische Gleichgewicht zu erreichen. Die Zeitkonstante ist anhängig von der Sensorgeometrie und von den thermischen Eigenschaften des verwendeten Materials. Typische Zeitkonstanten liegen im $10 \dots 100 \text{ ms}$ -Bereich. Reaktionsgeschwindigkeit und Empfindlichkeit sind konkurrierende Eigenschaften.

Der von einem Thermopile gelieferte Gleichanteil kann vorteilhaft zur Anwesenheitskontrolle von Personen verwendet werden. Im Vergleich zu Bewegungsmeldern mit pyroelektrischen Sensoren [Kap. 2.3.5.2] soll im folgenden der erforderliche Systemaufwand zur Auswertung

des Gleichanteils und ein mögliches Konzept für einen Präsenzmelder vorgestellt werden. Um ein ausreichendes Detektionsverhalten zu erzielen, ist eine mehr als 10.000-fache Gleichspannungsverstärkung erforderlich.

Der hohe Gleichspannungsverstärkungsfaktor erfordert Operationsverstärker (OP) mit möglichst geringen Offsetspannungen und extrem niedriger Temperaturdrift. Abb. 2.47 zeigt zwei DC-Verstärker, von denen insbesondere der erste OP wegen der 1000-fachen Verstärkung sehr hochwertig sein muss. Der nachgeschaltete Differenzverstärker bewirkt nur noch eine 20-fache Verstärkung, allerdings muss das Bezugspotential so eingestellt werden, das der Spannungspegel am Ausgang nicht in die Sättigung geht. Die Einstellung des Bezugspotentials erfolgt mikrocontrollergesteuert unter Beachtung des aktuellen Signals und der Innentemperatur des Thermopiles, welche anhand des internen Thermistors ermittelt werden kann.

Der in Abb. 2.47 dargestellte Wechselspannungsverstärker (Verstärkungsfaktor 50) steigert die Empfindlichkeit der Bewegungsdetektion. Weiterhin wurde ein Fotowiderstand zur Lichtmessung vorgesehen, um eine tageslichtabhängige Betriebsart zu ermöglichen. Abb. 2.48 zeigt die Ausgangsspannung des Thermopiles TPS634, wenn sich eine Person im Gesichtsfeld aufhält. Die Messungen wurden bei einer Zimmertemperatur von 22 °C durchgeführt. Der Sensor war in einer Höhe von 1 m horizontal ausgerichtet, die Gehrung wurde frontal in Richtung größter Empfindlichkeit gewählt. Die Messungen wurden ohne zusätzliche Linse im Strahlengang zwischen dem Sensor und der Person durchgeführt.

Bei stabilen Umgebungsbedingungen ist das Thermopile, auch ohne zusätzliche Optik, zur Anwesenheitskontrolle von Personen im Nahbereich bis etwa 3 m geeignet. Der maximale Öffnungswinkel des Erfassungsfeldes beträgt je nach Thermopile bis zu $\pm 45^\circ$. Überwacht das Thermopile einen größeren räumlichen Bereich, wird die zu detektierende Person meist nur einen Teil des Erfassungsbereichs ausfüllen.

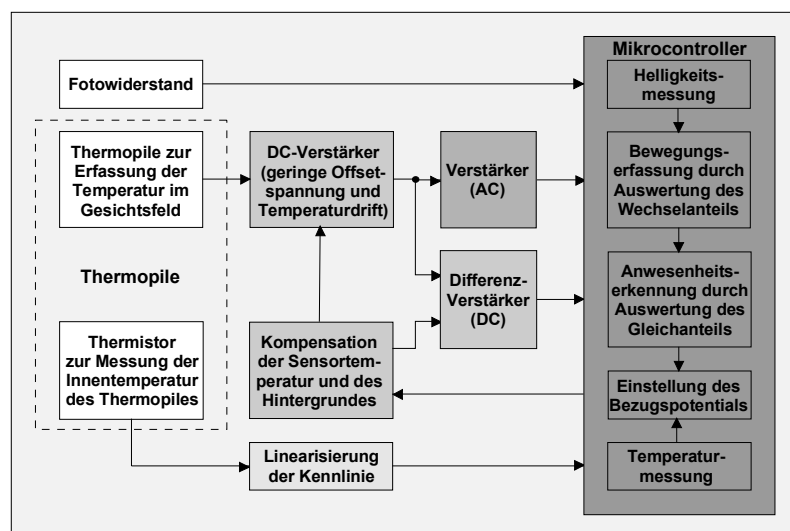


Abb. 2.47: Systemkonzept für einen Präsenzwächter mit Thermopile und Mikrocontroller

In diesem Fall ist eine absolute Temperaturmessung und damit die Identifizierung eines Menschen anhand seiner Körpertemperatur nicht möglich. Das Ausgangssignal entspricht einer mittleren Temperaturverteilung im gesamten Erfassungsfeld. Durch die Mittelung über das gesamte Gesichtsfeld lassen sich außerdem die langsamen Signalgradienten eines Thermopiles im Fernfeld erklären. Aus diesem Grund empfiehlt sich die getrennte Auswertung von Gleich- und Wechselanteil, um die Reaktionszeiten zu verkürzen.

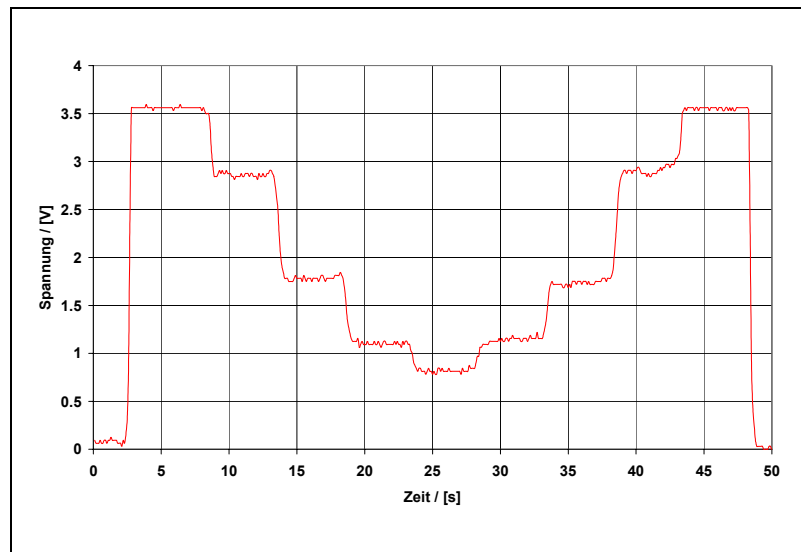


Abb. 2.48: Ausgangsspannungen des Thermopiles bei ca. 5 s Verweildauer in Abständen von 0.5 m, 1.0 m, 1.5 m, 2.0m und 2.5 m

Infolge der hohen Gleichspannungsverstärkung ist die gesamte Schaltung sehr anfällig gegenüber Temperaturdriften in den Verstärkerstufen, aber auch bezüglich langsamen Änderungen der Referenzstellentemperatur im Thermopile. Die Optimierung der Schaltungsanordnung hinsichtlich einer bestmöglichen Kompensation von Effekten, die auf Änderungen der Innentemperatur zurückzuführen sind, wurde bisher nur ansatzweise durchgeführt. Bei Versuchen im Klimaschrank konnte gezeigt werden, dass die Kompensation zumindest für einen bestimmten Temperaturbereich möglich ist.

Eine weitere wesentliche Störkomponente sind langsame Temperaturänderungen im Erfassungsbereich des Sensors. Selbst geringfügige Änderungen der Umgebungstemperatur äußern sich in einer langsamen Drift des Ausgangssignals. Für eine sichere Objekterkennung und -verfolgung mit dem Ziel der Präsenzkontrolle ist es daher wichtig, Änderungen der Hintergrundtemperatur bei der Auswertung im Mikrocontroller zu berücksichtigen.

2.3.5.2 Bewegungserkennung mit pyroelektrischen Detektoren

Im Gegensatz zu den Thermopiles reagieren pyroelektrische Detektoren ausschließlich auf Änderungen der Wärmestrahlung im Erfassungsbereich. Die absorbierte Strahlung ruft eine Erwärmung des pyroelektrischen Materials hervor und bewirkt eine Veränderung der Kristallpolarisation (Pyroelektrischer Effekt) im Sensorelement. Die dabei auftretenden Feldände-

rungen werden durch Ausgleichsströme zwischen den Sensorelektroden kompensiert, die letztendlich zum Nachweis von Strahlungsänderungen dienen. Aufgrund der Bedeutung für diese Arbeit erfolgt in Kap. 5 eine detaillierte Darstellung der Funktionsweise und der Vorgehensweise zur Modellierung des Übertragungsverhaltens pyroelektrischer Detektoren.

Der weitaus größte Teil kommerziell erhältlicher Bewegungsmelder verwendet pyroelektrische Detektoren. Die hohe Empfindlichkeit und das einfache Systemkonzept, das in dargestellt ist, haben wesentlich zur weiten Verbreitung beigetragen. Ein weiterer Vorteil ist das passive Erfassungsprinzip, das einen geringen Stromverbrauch der Geräte und damit niedrige Betriebskosten garantiert.

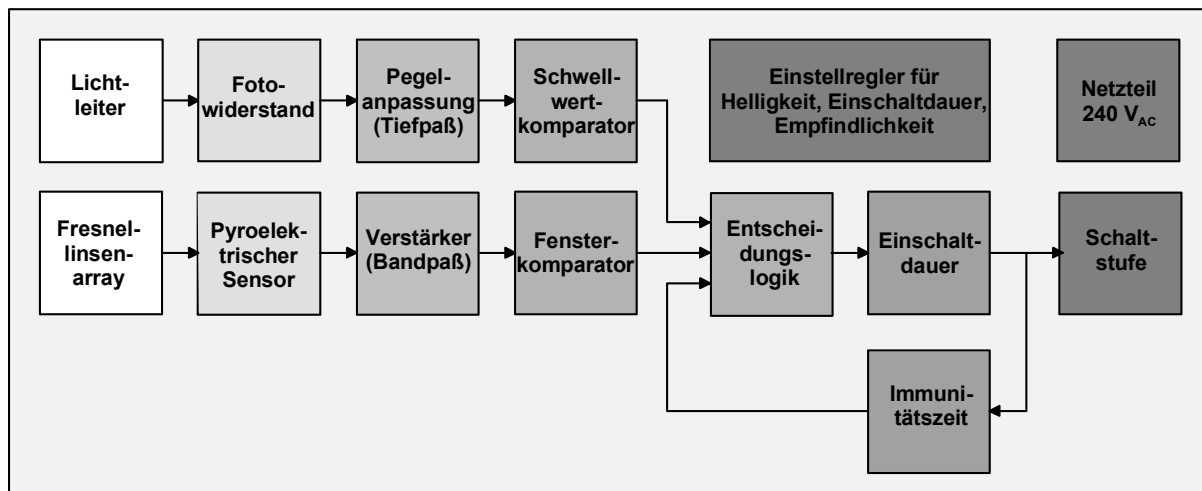


Abb. 2.49: Systemkomponenten eines konventionellen PIR-Bewegungsmelders

Im Vergleich zu der in Kap. 2.3.5.1 beschriebenen Problematik bei der Auswertung von Gleichspannungssignalen ist bei Bewegungsmeldern mit pyroelektrischen Sensoren lediglich eine hohe Wechselspannungsverstärkung erforderlich. Verwendet werden oftmals ein oder zwei Operationsverstärker, die als Bandpässe mit einer unteren Grenzfrequenz um 0.1 Hz und mit einer oberen Grenzfrequenz von 5 ... 10 Hz ausgelegt sind. Die Gesamtverstärkung liegt im Bereich von 500 ... 5000.

Die einfachste Form der Signalauswertung enthält nach Abb. 2.49 einen Fensterkomparator, der die Analogsignale pegelbewertet digitalisiert. Bei ausreichender Signalamplitude erfolgt die Aktivierung der Schaltstufe, wobei die Einschalt-dauer von einem nachtriggerbaren Timer gesteuert wird. Nach dem Abschalten ist eine kurze Immunitätszeit zur Vermeidung von Fehlauflösungen durch die angeschlossene Last vorgesehen. Während dieser Zeit ist die Entscheidungslogik, die im einfachsten Fall aus einer UND-Verknüpfung der in Abb. 2.49 gezeigten, digitalen Signalen besteht, gesperrt. Um eine tageslichtabhängige Betriebsart zu ermöglichen, wird von der Entscheidungslogik ebenfalls der aktuelle Helligkeitspegel im Vergleich zu einem extern einstellbaren Schwellwert berücksichtigt.

Mit Hinblick auf das Schaltverhalten und die Störsicherheit hat die Signalauswertung mit Fensterkomparatoren Nachteile gegenüber modernen Bewegungsmeldern, die durchweg mit Mikrocontrollern oder Mikroprozessoren zur digitalen Signalverarbeitung ausgestattet sind. Der technische Fortschritt dieser Systeme und deren Preisverfall in den letzten Jahren hat diese Entwicklung beschleunigt.

Mikrocontrollergesteuerte Bewegungsmelder verwenden im Vergleich zur oben beschriebenen Pegelauswertung deutlich aufwendigere Signalverarbeitungsstrategien, um das Ansprechverhalten bei gleichzeitig reduzierter Fehlalarmrate zu verbessern. In Kap. 7 werden verschiedene konventionelle Verfahren beschrieben, während in Kap. 8-10 ein signaltheoretisch optimaler, modellbasierter Ansatz zur Signalauswertung bei PIR-Bewegungsmeldern präsentiert wird. Ein weiterer Vorteil von Bewegungsmeldern mit Mikrocontrollern ist die Möglichkeit, ohne weiteren Hardwareaufwand zusätzliche Funktionen zu implementieren, welche die Funktionalität und den Bedienkomfort des Melders erhöhen. Beispiele für den gesteigerten Funktionsumfang sind verschiedene Betriebsmodi wie die Anwesenheitssimulation, der Dauerlichtbetrieb oder das automatische Ein- und Ausschalten der angeschlossenen Lichtquellen zu den vom Anwender vorgegebenen Uhrzeiten.

Unabhängig vom Aufwand der Signalverarbeitung hat das optische System einen entscheidenden Einfluss auf das Schaltverhalten von Bewegungsmeldern. Jede einzelne Linse des Fresnellinsenarrays muss unter Beachtung der gewünschten Erfassungsfeldspezifikationen und der vom Wächterdesign vorgegebenen Beschränkung der Gesamtlinsenfläche berechnet werden. Um ein möglichst kleines, unauffälliges Wächtergehäuse zu gewährleisten, ist eine optimale Ausnutzung der Gesamtlinsenfläche erforderlich. Sowohl das Wächterdesign als auch die Erfassungsfeldspezifikationen orientieren sich dabei meist an marktpolitischen Überlegungen. Kap. 6 beschreibt den Entwurf von Fresnellinsenarrays für Bewegungsmelder.

2.4 Literaturverzeichnis

Literatur

- [Lit. 2.1] B. Jähne, R. Massen, B. Nickolay, H. Scharfenberg: Technische Bildverarbeitung - Maschinelles Sehen, Springer Verlag, 1995
- [Lit. 2.2] Maylor K. Leung, Yee-Hong Yang: Human Body Motion Segmentation In A Complex Scene, Pattern Recognition, Volume 20, pp. 55-64, 1987
- [Lit. 2.3] H. Nagel: On the estimation on optical flow: Relations between different approaches and some new Results, Artificial Intelligence 33, pp. 299-324, 1987
- [Lit. 2.4] A. Moini: Vision Chips or Seeing Silicon, 2nd Ed., Centre for GaAs VLSI Technologies, University of Adelaide, Australia, 1996

- [Lit. 2.5] S. Brofferio, L. Carnimeo, D. Comunale, G. Mastronardi: A Background Updating Algorithm for Moving Object Scenes
in: V. Cappellini: Time-varying Image Processing and Moving Object Recognition, 2nd Ed., pp. 289-296, Elsevier Publishers B.V., Amsterdam, 1990
- [Lit. 2.6] K. Karmann, A. von Brandt: Moving Object Recognition using an Adaptive Background Memory
in: V. Cappellini: Time-varying Image Processing and Moving Object Recognition, 2nd Ed., pp. 289-296, Elsevier Publishers B.V., Amsterdam, 1990
- [Lit. 2.7] C. Ridder, O. Munkelt, H. Kirchner: Adaptive Background Estimation and Foreground Detektion using Kalman-Filtering, Proceedings of International Conference on Recent Advances in Mechatronics, ICRAM 1995, Istanbul, Turkey
- [Lit. 2.8] A. Baumberg: Leeds People Tracker. A short readable overview of the techniques behind the program, including a discussion of how mutual occlusions with cars are handled, <http://www.scs.leeds.ac.uk/imv/publications.html>
- [Lit. 2.9] P. Remagnino, A. Baumberg, T. Grove, D. Hogg, T. Tan, A. Worrall, K. Baker: An Integrated Traffic and Pedestrian Model-Based Vision System, Department of Computer Science, University of Reading, Berkshire RG6 6AY, School of Computer Studies, University of Leeds, LS2 9JT, UK
- [Lit. 2.10] K. Huggle, W. Eckstein: Extraktion von Personen in Videobildern, Technische Universität München
- [Lit. 2.11] O. Munkelt, C. Ridder, D. Hansel, W. Hafner: A Model-Driven Three-Dimensional Image-Interpretation System Applied to Person Detection in Video Images
in B. Jähne: Handbook of Computer Vision and Applications, Vol. 3. pp. 467-484, Academic Press, 1999
- [Lit. 2.12] K.-M. Chen, D. Misra, H. Wang, H.-R. Chuang, E. Postow: An X-Band Microwave Life-Detection System, IEEE Transactions on Biomedical Engineering, BME-33, No. 7, July 1986
- [Lit. 2.13] W. Eißler et al.: Praktischer Einsatz von berührungslos arbeitenden Sensoren, Sensorik Band 6, 2. Auflage, expert verlag, 1996
- [Lit. 2.14] Schulz & Walter Elektronik: Personendetektor PeD-05M - Hochempfindlicher Bewegungsdetektor mit exakt einstellbarer Reichweite, Produktbeschreibung der Firma Schulz & Walter Elektronik GmbH, Brandteichstraße 19, 17489 Greifswald
- [Lit. 2.15] H. Schicktanz, G. Hädrich: Elektrotechnische und elektronische Sicherungstechnik, FEINGERÄTETECHNIK, Berlin 40 (1991) 1

- [Lit. 2.16] G. Lohninger: Beleuchtungssteuerung mit Mikrowellensensoren, Funkschau, Heft 12/1992
- [Lit. 2.17] G. Lohninger: Microwave sensor SMX-1: Opening doors by invisible hands, Siemens AG, Components XXIX (1994), No. 4
- [Lit. 2.18] G. Lohninger: Doppler module KMY-24: Radar sensor detects motion and direction, Siemens AG, Components XXXI (1996), No. 4
- [Lit. 2.19] Hamamatsu: Infrared Detectors, Cat. No. KIRD0001E04, Sep. 95 T
- [Lit. 2.20] J. Schilz: Thermoelectric infrared sensors (thermopiles) for remote temperature measurements; pyrometry, EG & G Heiman Optoelectronics, Firmenschrift vom 13-Aug-1999
- [Lit. 2.21] Laser Components: IR-Detektoren und Zubehör, Produktkatalog 1995
- [Lit. 2.22] W. Heywang: Sensorik, Springer-Verlag 1984, Reihe Halbleiterelektronik, Band 17, ISBN 3-540-12767-4
- [Lit. 2.23] E. Dereniak, D. Crowe: Optical Radiation Detectors, John Wiley & Sons, Wiley series in pure and applied optics, III, ISBN 0-471-89797-3

Patente und Offenlegungsschriften

- [Pat. 2.1] Schaaf, N.: Verfahren zum Überwachen vom Innenräumen durch Erfassen von Druckschwankungen, Offenlegungsschrift DE3731593A1, 1987
- [Pat. 2.2] D. Hablov, O. Fisun, L. Lupichev, V. Osipov, V. Sestiperow, R. Schimko: Elektronisches Lebensdetektionssystem, Patentschrift DE4241664C2, 1992
- [Pat. 2.3] D. Hablov, O. Fisun, L. Lupichev, V. Osipov, V. Sestiperow, R. Schimko: Elektronisches Überwachungssystem, Patentschrift DE4241648C1, 1992

3 Bildgebende Sensoren zur Anwesenheitserkennung

3.1 Einleitung

Im Gegensatz zu den Passiv-Infrarot-Bewegungsmeldern (PIR-Bewegungsmeldern), die auf Temperaturänderungen in ihrem Erfassungsfeld reagieren, zählen die in Kap. 2 aufgeführten Ultraschallsensoren und Radarmodule zur Gruppe der aktiven Systeme. Sie benötigen zusätzliche Energie zur Erzeugung eines Fremdfeldes, dessen charakteristische Änderungen zur Bewegungs- und Anwesenheitserkennung von Objekten dienen.

Bildverarbeitende Sensorsysteme mit zusätzlicher externer oder integrierter Beleuchtungseinheit zählen ebenfalls zu den aktiven Systemen. In Verbindung mit der Ausnutzung des Tageslichts oder einer meist vorhandenen Raumbelichtung ist aber auch der Einsatz als passive Sensorkomponente in einem Multisensorsystem denkbar. Aufgrund des geringeren Energiebedarfs bieten sich insbesondere passive Sensorkonzepte für Anwendungen im Bereich der Gebäudetechnik an.

Speziell bei pyroelektrischen Detektoren erweist sich die sensorspezifische Kompensation der Hintergrundstrahlung als Nachteil. Bewegungen im Erfassungsfeld verursachen meist Änderungen der Temperaturstrahlung und können ohne großen Aufwand über Entfernungen bis 20 m und mehr detektiert werden. Es ist dagegen mit demselben Sensor nicht möglich, einen ruhig sitzenden Menschen im unmittelbaren Nahbereich des Wächters zu erkennen. Dennoch verwenden heute übliche Präsenzmelder pyroelektrische Detektoren zur Anwesenheitserkennung von Personen. Die Anwesenheitskontrolle erfolgt dabei auf dem Umweg über eine hochauflösende Bewegungserkennung innerhalb eines vordefinierten Zeitfensters.

Um auch geringfügige Bewegungen zu detektieren, erfordert ein PIR-Präsenzmelder im Vergleich zu PIR-Bewegungsmeldern für den Outdoorbereich eine deutlich feinere Unterteilung des Erfassungsfeldes. Da die Einzellinsen unter Beachtung der Leistungsbilanz nicht beliebig verkleinert und die Gesamtlinsenfläche mit Rücksicht auf das Gerätedesign nicht beliebig vergrößert werden kann, bieten sich insbesondere Sensorarrays mit mehreren pyroelektrischen Detektoren zur Erzielung einer möglichst hohen Erfassungsfelddichte an. Ein vielversprechender Ansatz zur Realisierung von pyroelektrischen Detektorarrays einschließlich der erforderlichen Ansteuer- und Ausleseelektronik wird beispielsweise in [Lit. 3.1] präsentiert.

Noch höhere Auflösungen lassen sich mit bildverarbeitenden Sensorsystemen (BV-Sensoren) erzielen. Beispielsweise werden zur Gebäudesicherung vorwiegend CCD-Kameras mit externer oder eigener Intelligenz eingesetzt, welche die Erkennung von bewegten Objekten und gegebenenfalls auch die Objektverfolgung automatisch durchführen. Im Alarmfall kann das Wachpersonal auf die Vorgänge aufmerksam gemacht und gleichzeitig die Aufzeichnung der Szene veranlasst werden. Derartige Überwachungssysteme sind sehr kostenintensiv, zumal

die Kameras mit geeignetem Zubehör ausgestattet werden müssen, um diesen Anforderungen gerecht zu werden.

Im Gegensatz zur vorher beschriebenen Videoüberwachung erfordern einfache automatisierungstechnische Aufgabenstellungen nicht unbedingt die Aufnahme und Übertragung von qualitativ hochwertigen Videobildern an eine ggf. weit entfernte Alarmzentrale. Vielmehr ist eine Sensor-Aktor-Kombination, die im folgenden als "Bildwächter" bezeichnet wird, zur Aktivierung bzw. Deaktivierung eines oder mehrerer elektrischer Verbraucher in Abhängigkeit der Raumbelastung ausreichend.

Als Sensorelement eines "Bildwächters" stehen neben konventionellen CCD-Matrixsensoren mittlerweile auch CMOS-Bildsensoren zur Verfügung, deren Bildqualität durch die intensive Forschungs- und Entwicklungsarbeit in den letzten Jahren zunehmend verbessert wurde.

Zur Realisierung eines Bildwächters bieten sich insbesondere CMOS-Bildsensoren an, da wesentliche Komponenten zur Steuerung der Bildaufnahme, zum Auslesen und zur Speicherung von Bilddaten sowie zur Datenvorverarbeitung auf dem Chip integriert werden können. Diese Integrationsmöglichkeiten bietet die CCD-Technologie nicht. Es sind zusätzliche externe Komponenten zum Betrieb eines CCD-Sensors erforderlich. Abschnitt 3.3 gibt einen kurzen Überblick zur Technologie von CCD- und CMOS-Bildsensoren und vermittelt einen Eindruck vom erforderlichen Systemaufwand und vom Anwendungspotential für die vorliegende Aufgabenstellung.

Die im Anschluss an die Bildaufnahme ablaufenden Verarbeitungsalgorithmen sind weitgehend unabhängig vom verwendeten Bildaufnehmer. Im weiteren Verlauf dieses Kapitels wird ein einfaches Signalverarbeitungskonzept zur Anwesenheitserkennung vorgestellt, das nur einen geringen Systemaufwand erfordert und das auch für eine Hardwareimplementierung geeignet ist. Die Anwesenheitskontrolle erfolgt, wie auch bei den heute üblichen PIR-Präsenzmeldern, auf dem Umweg über die Bewegungserkennung.

Als Einstieg in die zugrundeliegende Problematik werden in Kap. 3.2 zunächst wichtige strahlungsphysikalische Grundlagen zusammengefasst. Die Gesetze der Strahlungsausbreitung sowie die aufgeführten sender- und empfängerseitigen Kenngrößen gelten sowohl für Sensorsysteme im Bereich des sichtbaren Lichts und des nahen Infrarot (IR-A) als auch für die ab Kap. 5 beschriebenen Bewegungsmelder mit pyroelektrischen Detektoren. Wie bereits erwähnt wurde, ist zur Anwesenheitserkennung von Personen mit pyroelektrischen Detektoren speziell der Wellenlängenbereich um 10 μm im fernen Infrarot (IR-C) von besonderem Interesse.

3.2 Strahlungsphysikalische Grundlagen

3.2.1 Senderseitige Kenngrößen und Definition des Raumwinkels

Jeder Körper mit einer Temperatur oberhalb des absoluten Nullpunkts sendet infolge seiner eigenen Temperatur elektromagnetische Strahlung aus. Die abgegebene Strahlungsleistung Φ , die auch als Strahlungsfluss bezeichnet wird, beschreibt dabei die gesamte, pro Zeiteinheit dt in den umgebenden Raum abgestrahlte Energie dW . Für die Strahlungsleistung Φ gilt somit (Gl. 3.1). Ist die pro Zeiteinheit $\Delta t = dt$ abgestrahlte Energie $dW = W$ konstant, gilt die in (Gl. 3.1) angegebene Vereinfachung.

$$(Gl. 3.1) \quad \Phi = \frac{dW}{dt} \cong \frac{W}{\Delta t}.$$

Zur Erstellung von Leistungsbilanzen ist oftmals der Anteil des Strahlungsflusses zu ermitteln, der in einen bestimmten Raumbereich emittiert wird.

Als Maßzahl für diesen räumlichen Bereich dient der Raumwinkel Ω nach Abb. 3.1, der als Verhältnis der Oberfläche A des eingeschlossenen Kugelabschnitts zum Quadrat des Kugelradius R definiert ist.

$$(Gl. 3.2) \quad \Omega = \frac{A}{R^2} \text{ mit } [\Omega] = 1 \text{ sr (Steradian)}$$

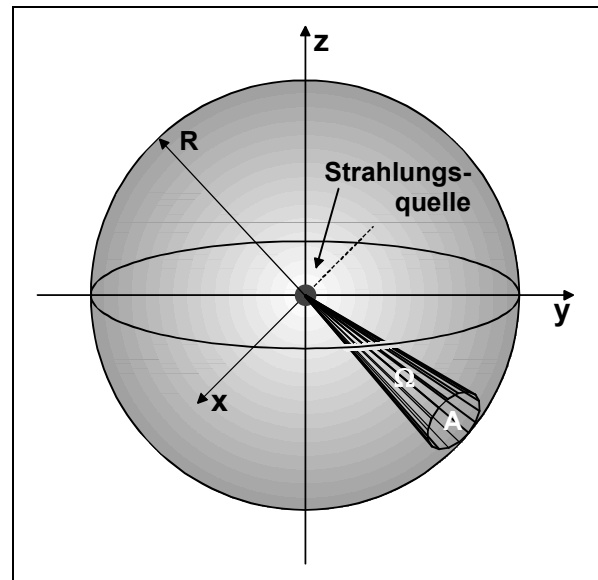


Abb. 3.1: Definition des Raumwinkels

Bei geringen Öffnungswinkeln kann die Oberfläche des Kugelabschnitts näherungsweise durch eine ebene Fläche ersetzt werden. Wird der räumliche Bereich durch einen Kegel mit dem Öffnungswinkel 2α begrenzt, lässt sich der Raumwinkel Ω nach (Gl. 3.3) berechnen.

$$(Gl. 3.3) \quad \Omega = 2\pi \cdot (1 - \cos(\alpha))$$

Mitunter ist die Darstellung eines infinitesimalen Raumwinkelelements in Kugelkoordinaten erforderlich. Ausgehend von Abb. 3.2 kann die Fläche dA eines kleinen Flächenelements nach (Gl. 3.4) bestimmt werden, woraus sich das Raumwinkelelement $d\Omega$ nach (Gl. 3.5) ableiten lässt.

$$(Gl. 3.4) \quad dA = R \cdot d\theta \cdot R \cdot \sin(\theta) \cdot d\varphi \quad (Gl. 3.5) \quad d\Omega = \sin(\theta) \cdot d\theta \cdot d\varphi$$

Für Leistungsberechnungen in Verbindung mit punktförmigen Strahlungsquellen wurde die Strahlstärke I nach (Gl. 3.6) als Quotient aus dem Strahlungsfluss $d\Phi$ und dem durchstrahlten Raumwinkelelement $d\Omega$ definiert. Bei konstanter Strahlungsleistung gilt wiederum die in (Gl.

3.6) angegebene Vereinfachung. Die Strahlstärke ist im allgemeinen abhängig von der Strahlungsrichtung, die in Kugelkoordinaten durch die Winkel θ und φ festgelegt ist.

$$(Gl. 3.6) \quad I = \frac{d\Phi}{d\Omega} \cong \frac{\Phi}{\Delta\Omega}$$

Flächige Strahler werden im Rahmen dieses Modells als Ansammlung von nahezu punktförmigen, ebenen Strahlungsquellen betrachtet [Abb. 3.3]. Die Strahldichte L nach (Gl. 3.7) charakterisiert dabei die Strahlstärke I , die von einem infinitesimalen Flächenelement dA ausgeht oder durch dieses hindurchtritt. Der Winkel θ in (Gl. 3.7) berücksichtigt die Projektion des Flächenelements dA in Ausbreitungsrichtung der Strahlung.

$$(Gl. 3.7) \quad L = \frac{I}{\cos\theta \cdot dA} = \frac{d\Phi}{\cos\theta \cdot dA \cdot d\Omega}$$

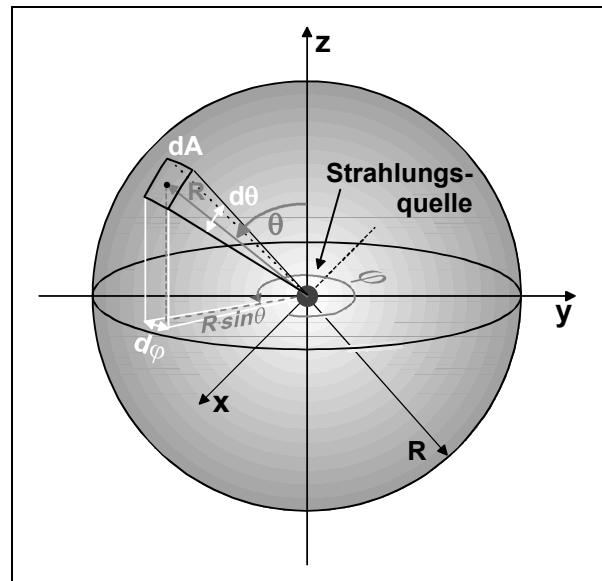


Abb. 3.2: Infinitesimales Raumwinkelelement in Kugelkoordinaten

Im Gegensatz zu einer punktförmigen Strahlungsquelle wird die von einem ebenen Flächenelement dA ausgehende Strahlung lediglich in einen Bereich mit dem Raumwinkel $\Omega = 2\pi$ (Halbraum) abgestrahlt [Abb. 3.3].

Die spezifische Ausstrahlung M beschreibt in diesem Fall den pro Flächenelement dA in den Halbraum emittierten Strahlungsfluss $d\Phi$. M ist eine Funktion des Ortes auf der Kugeloberfläche. Die Integration von M über alle Flächenelemente dA liefert die insgesamt von der Oberfläche emittierte Strahlungsleistung Φ .

$$(Gl. 3.8) \quad M = \frac{d\Phi}{dA}$$

$$(Gl. 3.9) \quad I = I_0 \cdot \cos(\theta)$$

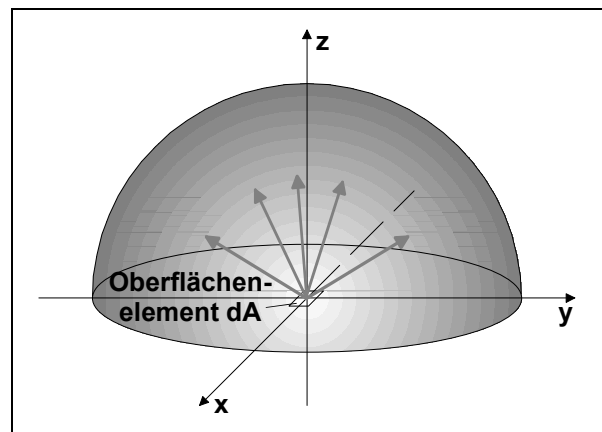


Abb. 3.3: Spezifische Ausstrahlung eines Flächenelements

Für theoretische Betrachtungen wird oftmals ein idealer, maximal diffus strahlender Körper verwendet, der als Lambertscher Strahler bezeichnet wird. Die Strahlstärke I eines beliebigen Flächenelements dA des Lambertschen Strahlers variiert nach (Gl. 3.9) mit dem Kosinus des Winkels θ zwischen der Normale des Flächenelements und der Ausstrahlungsrichtung und hängt somit ausschließlich von der Fläche des in Strahlungsrichtung projizierten Flächenelements ab.

Der Vergleich von (Gl. 3.9) und (Gl. 3.7) zeigt, dass ein Lambertischer Strahler eine konstante Strahldichte L aufweist. Die Strahldichte ist somit unabhängig von der Strahlungs- oder Betrachtungsrichtung. Berechnungen zur Leistungsübertragung lassen sich dadurch stark vereinfachen. Für die spezifische Ausstrahlung M eines Flächenelements dA des Lambertischen Strahlers folgt (Gl. 3.10), wobei die Integration über die Strahlstärke auf den Halbraum begrenzt wird.

$$(Gl. 3.10) \quad M = \frac{d\Phi}{dA} = \int_{\Omega=0}^{2\pi} \frac{I}{dA} d\Omega = \int_{\theta=0}^{\pi/2} \int_{\varphi=0}^{2\pi} \frac{L \cdot dA \cdot \cos\theta}{dA} \cdot \sin\theta d\theta d\varphi = \pi \cdot L.$$

Die spektrale spezifische Ausstrahlung $M_{\lambda S}(\lambda, T)$ eines schwarzen Strahlers wird durch das Plancksche Strahlungsgesetz (Gl. 3.11) beschrieben. Der schwarze Strahler ist definitionsgemäß auch ein Lambertischer Strahler. Er emittiert in jede Richtung für jede Wellenlänge λ die maximal mögliche Energie im Vergleich zu allen anderen Temperaturstrahlern mit gleicher Temperatur und identischen geometrischen Formen und Abmessungen [Lit. 3.2].

$$(Gl. 3.11) \quad M_{\lambda S}(\lambda, T) = \frac{d\Phi_{\lambda S}(\lambda, T)}{dA_S} = \frac{2\pi \cdot c^2 \cdot h}{\lambda^5} \cdot \left[\exp\left(\frac{h \cdot c}{\lambda \cdot kT}\right) - 1 \right]^{-1}$$

Lichtgeschwindigkeit $c = 2.99792438 \cdot 10^{10}$ cm/s, Plancksches Wirkungsquantum $h = 6.626176 \cdot 10^{-34}$ Js, Boltzmann-Konstante $k = 1.380662 \cdot 10^{-23}$ Js/K

In Abb. 3.4 ist die spektrale Strahldichte $L_{\lambda S}(\lambda, T)$ des schwarzen Körpers nach (Gl. 3.11) dargestellt, wobei der Zusammenhang zwischen der Strahldichte und der spezifischen Ausstrahlung in den Halbraum nach (Gl. 3.10) berücksichtigt wurde.

Die Strahldichte ist in dieser Abbildung als Funktion der Wellenlänge und mit der Temperatur als Parameter aufgetragen. Je höher die Temperatur ist, umso stärker verschiebt sich das Maximum der Ausstrahlung bzw. der Strahldichte zu kürzeren Wellenlängen. Dieser Zusammenhang wird durch das Wiensche Verschiebungsgesetz ausgedrückt:

$$(Gl. 3.12) \quad \lambda_{\max} \cdot T = 2898 \text{ um} \cdot \text{K}$$

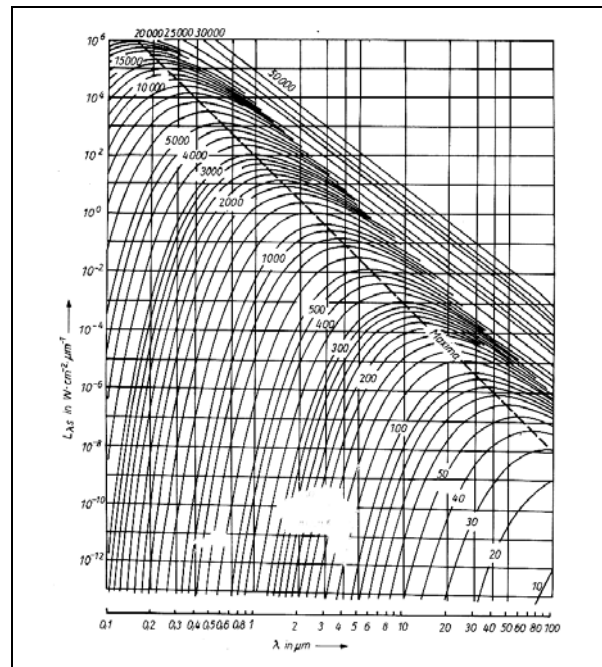


Abb. 3.4: Spektrale spezifische Strahldichte des schwarzen Körpers mit $L_{\lambda S} = M_{\lambda S} / \pi$ [Lit. 3.3]

Die gesamte spezifische Ausstrahlung M_S des schwarzen Körpers erhält man nach dem Stefan-Boltzmann-Gesetz (Gl. 3.13) durch Integration von (Gl. 3.11) über alle Wellenlängen.

(Gl. 3.13) $M_S = \sigma \cdot T^4$ mit $\sigma = 5,67032 \cdot 10^{-12} \frac{W}{cm^2 K^4}$

Das Emissionsverhalten realer (nicht schwarzer) Körper wird durch den wellenlängenabhängigen Emissionsgrad $\epsilon(\lambda)$ beschrieben, der als Verhältnis der spezifischen Ausstrahlung des Prüflings zur maximalen Ausstrahlung eines schwarzen Körpers definiert ist.

(Gl. 3.14) $\epsilon(\lambda) = \frac{M_\lambda(T)}{M_{\lambda S}(T)}$

Substanz	Bandemissionsgrad
Sand	0,93
Grünes Laub	0,90
Holz	0,98
Menschliche Haut	0,98
Beton	0,96

Tab. 3.1: Bandemissionsgrad verschiedener Materialien im Bereich von $\lambda = 8 \dots 14$ [Lit. 3.3]

Der Emissionsgrad eines schwarzer Strahlers ist unabhängig von der Wellenlänge. Nach (Gl. 3.14) gilt $\epsilon(\lambda) = \epsilon_S = 1$. Körper mit einem kleineren Emissionsgrad als dem des schwarzen Strahlers werden als grau bezeichnet, wenn die Unabhängigkeit des Emissionsgrades von der Wellenlänge gewährleistet ist.

3.2.2 Empfängerseitige Kenngrößen

In Kap. 3.2.1 wurden verschiedene Definitionen zur Beschreibung senderspezifischer Eigenschaften zusammengefasst. In diesem Abschnitt soll die Empfangsseite betrachtet werden.

Eine wesentliche Größe zur Erstellung von Leistungsbilanzen ist die Bestrahlungsstärke E . Sie beschreibt den pro Flächenelement dA einer bestrahlten Fläche auftreffenden Strahlungsfluss $d\Phi$.

(Gl. 3.15) $E = \frac{d\Phi}{dA} \cong \frac{\Phi}{A}$

Bei konstanter Strahlungsleistung Φ durch die Gesamtfläche A können die differentiellen Elemente in (Gl. 3.15) wiederum entsprechend ersetzt werden.

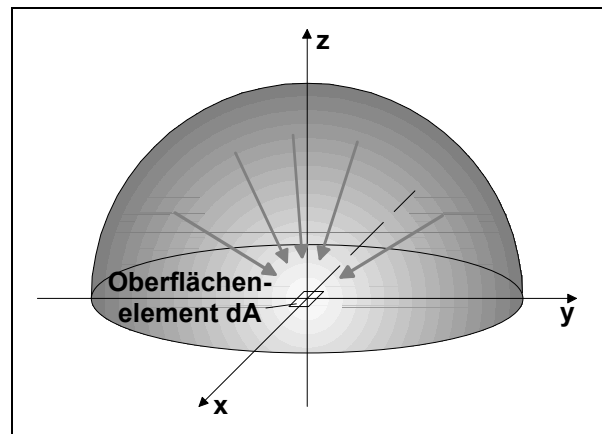


Abb. 3.5: Bestrahlungsstärke eines Flächenelements

Integriert man die Bestrahlungsstärke über die Gesamtfläche A , erhält man die insgesamt verfügbare Strahlungsleistung. Das zeitliche Integral über die Bestrahlungsstärke wird als Belichtung H bezeichnet. Die Belichtung ist ein Maß für die pro Flächenelement in der Empfangsebene vorhandene Strahlungsenergie.

Wird die Oberfläche dA , die um den Winkel φ_E gegen die Strahlungsausbreitungsrichtung geneigt ist, mit der konstanten Strahlstärke I einer punktförmigen Lichtquelle beleuchtet, liefert das fotometrische Entfernungsgesetz nach (Gl. 3.16) die resultierende Bestrahlungsstärke. Der Term $dA \cdot \cos(\varphi_E)/d^2$ entspricht dabei dem Raumwinkel, den das Flächenelement dA im Abstand d von dem Punktstrahler aufspannt.

$$(Gl. 3.16) \quad E = \frac{d\Phi}{dA} = I \cdot \frac{dA \cdot \cos(\varphi_E)}{dA \cdot d^2} = I \cdot \frac{\cos(\varphi_E)}{d^2}$$

Das Verhalten des Körpers für den auftreffenden Strahlungsfluss wird durch den Absorptionsgrad $\alpha(\lambda)$, den Reflexionsgrad $\rho(\lambda)$ und den Transmissionsgrad $\tau(\lambda)$ gekennzeichnet. Die Koeffizienten beschreiben das Verhältnis der absorbierten, der reflektierten und der transmittierten Strahlungsleistung zur einfallenden Leistung. Aus der Leistungsbilanz folgt:

$$(Gl. 3.17) \quad \alpha(\lambda) + \rho(\lambda) + \tau(\lambda) = 1$$

Alle drei Größen hängen von der Temperatur, der chemischen Zusammensetzung und der Oberflächenbeschaffenheit des Körpers sowie auch von den Eigenschaften des Strahlungsfeldes ab. Nach dem Kirchhoffschen Gesetz gilt außerdem der Zusammenhang, dass der Absorptionsgrad $\alpha(\lambda)$ und der Emissionsgrad $\varepsilon(\lambda)$ eines Körpers identisch sind.

Für den Fall einer vollkommenen diffusen Strahlungsreflektion an dem Flächenelement dA im Abstand d zur Strahlungsquelle, wobei zur Vereinfachung im folgenden der Reflektionsfaktor $\rho(\lambda) = \rho = \text{const.}$ gewählt wird, lässt sich die reflektierte Strahldichte L_R nach (Gl. 3.18) berechnen [Lit. 3.2]. Wie bereits vorher erwähnt wurde, ist der Reflektionsfaktor $\rho = d\phi_R/d\phi$ als Verhältnis der reflektierten Strahlungsleistung $d\phi_R$ zu der auf der Oberfläche auftreffenden Strahlungsleistung $d\phi$ definiert. Der Winkel φ_A berücksichtigt dabei die Einfallrichtung der Strahlung, relativ zur Normale des Flächenelements dA . Mit der Definitionsgleichung der Bestrahlungsstärke $E = d\Phi/dA$ des bestrahlten Flächenelements kann der in (Gl. 3.18) angegebene Zusammenhang zwischen der Bestrahlungsstärke des Flächenelements und der reflektierten Strahldichte abgeleitet werden.

$$(Gl. 3.18) \quad L_R = \frac{1}{\pi [\text{sr}]} \cdot \frac{d\Phi_R}{dA \cdot \cos(\varphi_A)} = \frac{1}{\pi [\text{sr}]} \cdot \frac{\rho \cdot d\Phi}{dA \cdot \cos(\varphi_A)} = \frac{\rho \cdot E}{\pi [\text{sr}] \cdot \cos(\varphi_A)}$$

3.2.3 Wellenausbreitung und Energieübertragung

Bei der Ausbreitung elektromagnetischer Wellen in der Luft tritt eine wellenlängenabhängige Absorption der Strahlung auf. Abb. 3.6 zeigt den Transmissionsgrad der Luft, der entlang einer Strecke von 1.6 km in der Erdatmosphäre gemessen wurde. Ungünstige Wetterbedingungen wie Nebel, Regen oder Schnee bewirken eine zusätzliche Dämpfung in dem für die Personenerfassung mit pyroelektrischen Detektoren verwendeten Wellenlängenbereich um 10 μm . Unter diesen Umgebungsbedingungen verschlechtert sich das Detektionsverhalten infolge der reduzierten Strahlungsleistung deutlich.

Zur Berechnung der übertragenen Leistung wird im folgenden die homogene Ausstrahlung bzw. die konstante Strahldichte eines Lambertischen Strahlers vorausgesetzt. In diesem Fall können die infinitesimalen Flächenelemente dA_S und dA_E des Senders und des Empfängers durch die jeweilige Gesamtfläche A_S bzw. A_E ersetzt werden. Außerdem soll der Abstand zwischen Sender und Empfänger größer als die fotometrische Grenzentfernung sein, d. h. der Abstand ist mindestens zehnmals größer als der Durchmesser beider Flächen. Bei der Ausbreitung der Strahlung muss der Divergenzwinkel nach Abb. 3.7 berücksichtigt werden. Infolge des Öffnungswinkels ändert sich die durch- oder bestrahlte Fläche mit wachsender Entfernung vom Sender.

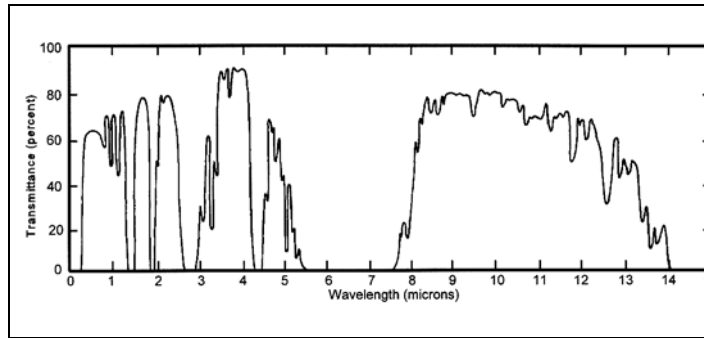


Abb. 3.6: Transmissionsgrad der Luft in Abhängigkeit der Wellenlänge der Strahlung]

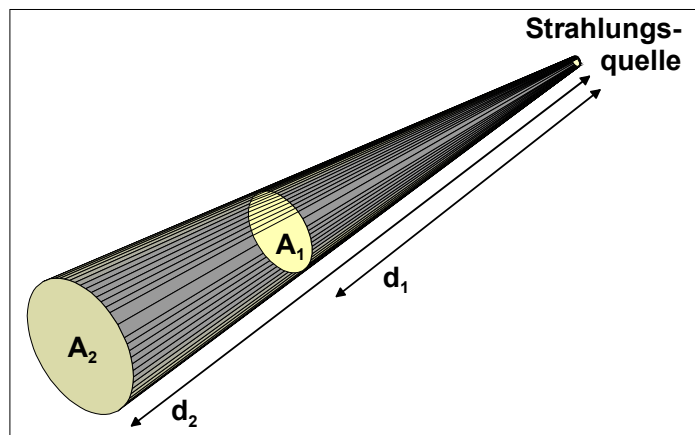


Abb. 3.7: Bestrahlungsstärke in verschiedenen Abständen zur Quelle

$$(Gl. 3.19) \quad \frac{E_1}{E_2} = \frac{d_2^2}{d_1^2}$$

Bei konstanter Strahldichte kann das Verhältnis der Bestrahlungsstärken E_1 und E_2 der parallelen Flächen A_1 und A_2 im Abstand d_1 und d_2 von einer punktförmigen Strahlungsquelle mit dem Entfernungsgesetz nach (Gl. 3.19) ermittelt werden. Im doppelten Abstand zur Quelle ist nur noch ein Viertel der Leistung pro bestrahltem Flächenelement vorhanden.

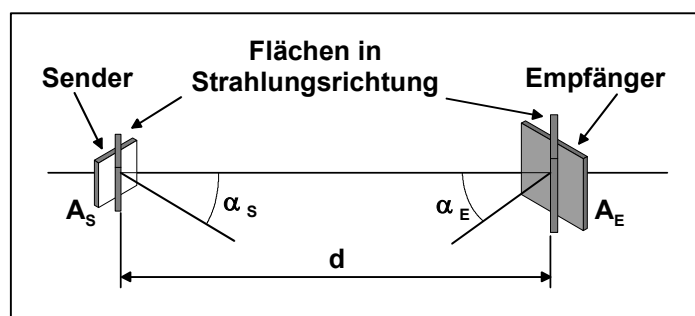


Abb. 3.8: Allgemeine Anordnung des Senders und des Empfängers

$$(Gl. 3.20) \quad \Phi = L \frac{A_S \cos(\alpha_S) \cdot A_E \cos(\alpha_E)}{d^2}$$

Im allgemeinen Fall, wenn die Sender- und die Empfängerfläche A_S und A_E um einen Winkel α_S bzw. α_E gegen die Ausbreitungsrichtung der Strahlung geneigt sind, müssen nach dem photometrischen Grundgesetz die Projektionen beider Flächen senkrecht zur Ausbreitungs-

richtung der Strahlung berücksichtigt werden. Die insgesamt am Empfänger auftreffende Strahlungsleistung lässt sich in diesem Fall mit (Gl. 3.20) bestimmen.

Der Term $I_S = L \cdot A_S \cdot \cos(\alpha_S)$ in (Gl. 3.20) entspricht der Strahlstärke des Senders in Richtung eines Empfängers, der den Raumwinkel $\Omega_E = A_E \cdot \cos(\alpha_E) / d^2$ aufspannt. Umgekehrt kann der Term $I_E = L \cdot A_E \cdot \cos(\alpha_E)$ als diejenige Strahlstärke aufgefasst werden, die der Empfänger unter dem Raumwinkel von $\Omega_S = A_S \cdot \cos(\alpha_S) / d^2$ von dem Sender "sieht". Ist das Gesichtsfeld Ω_B des Empfängers größer als der von der Senderfläche vorgegebene Raumwinkel Ω_S , tritt am Empfangsort zusätzlich ein Leistungsanteil $\Phi_B = L_B \cdot A_E \cdot \cos(\alpha_E) \cdot (\Omega_B - \Omega_S)$ infolge der Hintergrundstrahlung mit der Strahlstärke L_B auf.

Zur Ableitung des Kosinusetzes der Radiometrie soll Abb. 3.9 dienen. Bei der dargestellten Anordnung des Senders ergeben sich je nach Empfängerausrichtung und Abstand unterschiedliche Abhängigkeiten der Strahlungsleistung vom Einfallswinkel α .

Fall I	Keine Cos-Abhängigkeit
Fall II	$\cos^1(\alpha)$ -Abhängigkeit
Fall III	$\cos^2(\alpha)$ -Abhängigkeit
Fall IV	$\cos^3(\alpha)$ -Abhängigkeit
Fall V	$\cos^4(\alpha)$ -Abhängigkeit

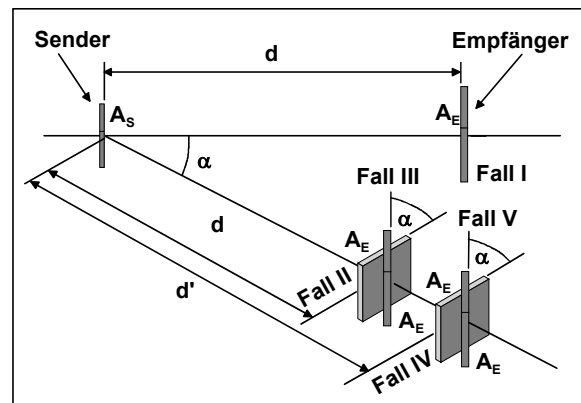


Abb. 3.9: Ableitung des Kosinusetzes der Radiometrie [Lit. 3.4]

Betrachtet man nur die durch die Empfängeranordnungen I und V aufgespannte Ebene, ergibt sich mit $d'^2 = d^2 / \cos(\alpha)$ in dieser Ebene die in (Gl. 3.21) angegebene Variation der Strahlungsleistung mit dem Faktor $\cos^4(\alpha)$.

$$(Gl. 3.21) \quad \Phi = L \cdot \frac{A_S \cdot A_E}{d'^2} \cdot \cos^2(\alpha) = L \cdot \frac{A_S \cdot A_E}{d^2} \cdot \cos^4(\alpha)$$

3.2.4 Geometrische Optik

Ausgehend von der Descartes'schen Linsenformel für dünne Linsen werden im folgenden zunächst einige grundlegende Definitionen und Beziehungen im Zusammenhang mit der optischen Abbildung einer Szene im Gegenstandsraum auf einen Bildsensor erläutert. In der Praxis werden diese Gleichungen auch zur Abschätzung der Abbildungseigenschaften von realen Objektiven verwendet. Dabei ist zu beachten, dass diese Gleichungen strenggenommen nur im Rahmen der in der geometrischen Optik getroffenen Vereinfachungen gelten. Als Einstieg in diese Thematik ist in Abb. 3.10 die Bündelung eines aus großen Entfernungen einfallenden, parallelen Strahlenbündels mit einem Linsensystem dargestellt. Fällt die Strahlung parallel zur optischen Achse ein, wird sie im bildseitigen Brennpunkt der Linse fokussiert. Den Abstand zwischen der hinteren Hauptebene H' des optischen Systems und diesem Brennpunkt

bezeichnet man als bildseitige Brennweite f' . Bei dünnen Linsen fallen die in Abb. 3.10 dargestellten Hauptebenen H und H' des optischen Systems zusammen.

Unter Anwendung der in Abb. 3.11 gezeigten, geometrischen Beziehungen im Objekt- und im Bildraum nach

$$(Gl. 3.22) \quad m' = \frac{y'}{y} = \frac{a'}{a} \quad \text{und}$$

$$(Gl. 3.23) \quad \frac{y'}{y} = \frac{a' - f'}{f'}$$

erhält man die für dünne Linsen gültige Linsenformel von Descartes (Gl. 3.24). Sie beschreibt den grundlegenden Zusammenhang zwischen der bildseitigen Brennweite f' der Linse sowie der Bildweite a' und der Gegenstandsweite a (s. Abb. 3.11).

Das in (Gl. 3.22) angegebene Verhältnis zwischen der Bildgröße y' und der Gegenstandsgröße y ist der Abbildungsmaßstab m' des optischen Systems.

Setzt man (Gl. 3.22) in (Gl. 3.24) ein, kann die erforderliche Objektivbrennweite f' bei bekannter Gegenstandsweite a und bei bekanntem Abbildungsmaßstab m' nach (Gl. 3.25) berechnet werden. Analog lassen sich die Gleichungen (Gl. 3.26), (Gl. 3.27), (Gl. 3.28) und (Gl. 3.29) zur Bestimmung der Objektgröße y , der Bildgröße y' , der Objektweite a oder der Bildweite a' ableiten.

$$(Gl. 3.25) \quad f' = \left(1 + \frac{1}{m'}\right)^{-1} \cdot a = \left(1 + \frac{y}{y'}\right)^{-1} \cdot a \quad (Gl. 3.28) \quad a = \left(1 + \frac{1}{m'}\right) \cdot f' = \left(1 + \frac{y}{y'}\right) \cdot f'$$

$$(Gl. 3.26) \quad y = \left(\frac{a}{f'} - 1\right) \cdot y' \quad (Gl. 3.29) \quad a' = (1 + m') \cdot f' = \left(1 + \frac{y'}{y}\right) \cdot f'$$

$$(Gl. 3.27) \quad y' = \left(\frac{a}{f'} - 1\right)^{-1} \cdot y \quad (Gl. 3.30) \quad k = \frac{f'}{D}$$

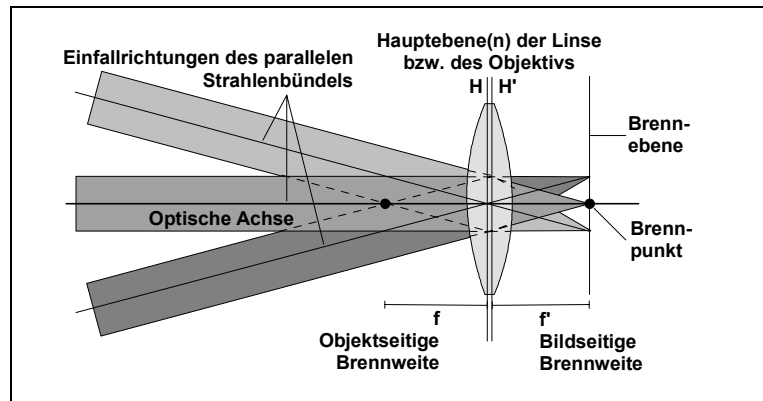


Abb. 3.10: Abbildung paralleler Strahlbündel durch eine dünne Linse

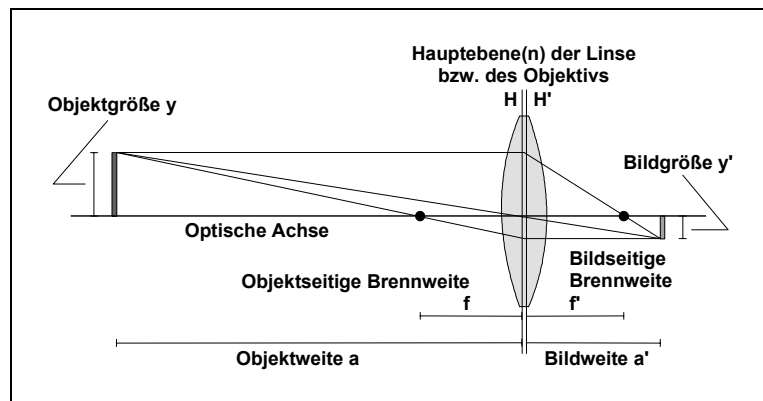


Abb. 3.11: Optische Grundgrößen nach DIN 1335 [Lit. 3.11]

$$(Gl. 3.24) \quad \frac{1}{f'} = \frac{1}{a'} + \frac{1}{a} \quad (\text{Linsenformel von Descartes})$$

Die Blendenzahl k nach (Gl. 3.30) ist ein Maß für die Beleuchtungsstärke eines Objektivs. Sie ergibt sich als Quotient der Brennweite f' und des Durchmessers D der Eintrittspupille des Objektivs. Die Blendenzahl wird vom Objektivhersteller angegeben und ist bei vielen Objektiven manuell oder automatisch einstellbar. Aus der Fotografie sind die Blendenzahlen $k = 1, 1.4, 2, 2.8, 4, 5.6, 8, 11, 16, 22, 32$ bekannt. Jede Stufe unterscheidet sich von der vorhergehenden um den Faktor $\sqrt{2}$, wodurch sich jeweils eine Halbierung der Fläche der Eintrittspupille und damit auch der durchgelassenen Lichtmenge ergibt.

Die Blendenzahl beeinflusst unmittelbar die Schärfentiefe der Abbildung. Darunter versteht man den Entfernungsbereich, den das Objektiv in der gewählten Einstellung gerade noch scharf auf die Bild- bzw. Sensorebene abbilden kann. Abb. 3.12 zeigt die Abbildung eines unendlich weit entfernten Objektpunktes auf die Sensorebene, die in diesem Fall in der Brennebene des Objektivs liegt..

Bewegt man den Objektpunkt näher zum Objektiv, wird die Strahlung in einem Punkt hinter der Brennebene fokussiert.

In der Bildebene entsteht ein Unschärfekreis, der von der tatsächlichen Objektweite und von der Fläche der Eintrittspupille des Objektivs abhängt.

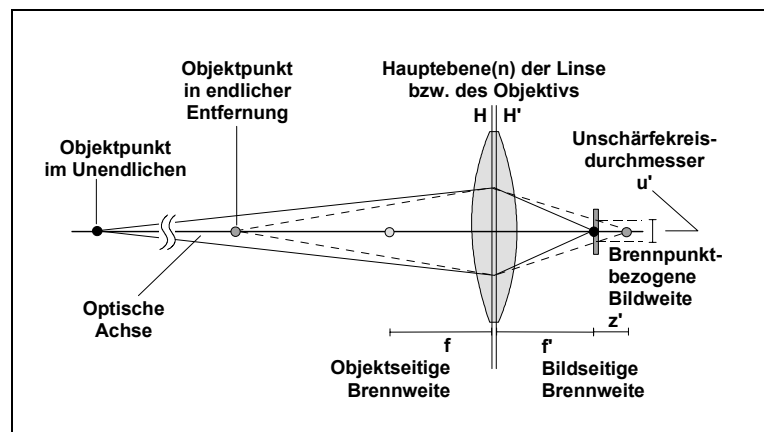


Abb. 3.12: Unschärfekreis und Schärfentiefe

Mit der Blendenzahl k und dem maximal zulässigen Durchmesser u' des Unschärfekreises, der in Verbindung mit bildgebenden Sensoren beispielsweise anhand einer Seitenlänge oder der Diagonale eines Pixels vorgegeben werden kann, gilt für die Schärfentiefe a_f im Fernpunkt bzw. für die Schärfentiefe a_n im Nahpunkt:

$$(Gl. 3.31) \quad a_f = \frac{a \cdot f^2}{f^2 - u' \cdot k \cdot (a - f)}$$

$$(Gl. 3.32) \quad a_n = \frac{a \cdot f^2}{f^2 + u' \cdot k \cdot (a - f)}$$

Zur Fokussierung auf einen anderen Entfernungsbereich kann die in Abb. 3.12 dargestellte, brennpunktbezogene Bildweite z' mit Hilfe der Entfernungseinstellung am Objektiv, durch Zwischenringe oder durch die Verwendung von Nahlinsen optimiert werden. Durch den begrenzten Hub der Entfernungseinstellung bei Standardobjektiven ergibt sich die vom Hersteller angegebene, minimal einstellbare Objektdistanz (MOD). Bildgebende Sensoren werden normalerweise in der Brennebene eines auf unendlich fokussierten Objektivs angeordnet. Die Bildweite und die bildseitige Brennweite stimmen in diesem Fall überein (s. Abb. 3.13).

Der horizontale, vertikale oder diagonale Bildwinkel w_h , w_v oder w_d kann folgendermaßen berechnet werden:

(Gl. 3.33)

$$w_{h,v,d} = 2 \cdot \tan^{-1} \left(\frac{y'_{h,v,d}}{f} \right).$$

Die Länge $y'_{h,v,d}$ entspricht jeweils der Hälfte der horizontalen, der vertikalen oder der diagonalen Chipabmessungen.

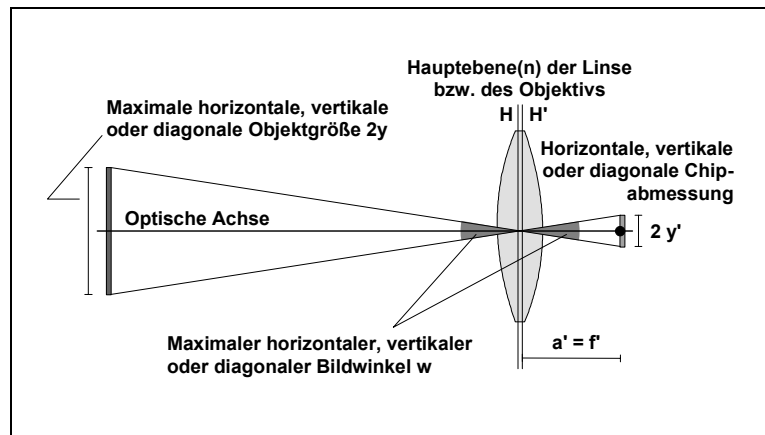


Abb. 3.13: Definition des Bildwinkels für ein auf unendlich fokussiertes Objektiv

3.2.5 Berücksichtigung einer Linse im Übertragungsweg

Wird nach Abb. 3.14 eine Linse im Strahlengang zwischen dem Sender und dem Empfänger verwendet, kann anhand von (Gl. 3.21) zunächst die auf der Linse auftreffende Strahlungsleistung bestimmt werden. Dazu wird die Linsenfläche A_L anstelle der Empfängerfläche A_E in (Gl. 3.21) eingesetzt und man erhält

(Gl. 3.34)

$$\Phi = L \cdot \frac{A_S \cdot A_L}{d^2} \cdot \cos^4(\alpha)$$

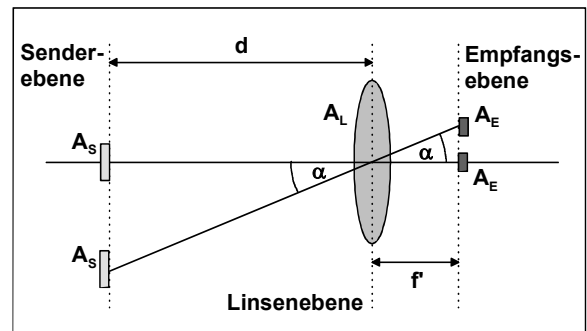


Abb. 3.14: Bestrahlungsstärke in der Detektorebene bei Verwendung einer Linse

Abbildungsfehler und Transmissionsverluste durch die Linse sollen im weiteren Verlauf vernachlässigt werden. Unterstellt man daher, dass die insgesamt an der Linse auftreffende Strahlungsleistung verlustlos auf das Empfangselement abgebildet wird, lässt sich mit (Gl. 3.34) die Bestrahlungsstärke am Ort des Detektors nach (Gl. 3.35) ermitteln. Stellt die Fläche A_S nur den Anteil einer ausgedehnten Strahlungsquelle dar, der von der Linse vollflächig auf die Detektorfläche A_E abgebildet wird, stimmen die Raumwinkel $\Omega_E = \Omega_S = \Omega$ auf beiden Seiten der Linse überein. In diesem Fall ist zur Berechnung der Bestrahlungsstärke in (Gl. 3.35) lediglich die Linsenapertur A_L und die Brennweite f sowie die Strahldichte L des Senders erforderlich [Lit. 3.4].

(Gl. 3.35)

$$E = \frac{\Phi}{A_E} = L \cdot \frac{A_S \cdot A_L}{A_E \cdot d^2} \cdot \cos^4(\alpha) = L \cdot \frac{A_L}{f^2} \cdot \cos^4(\alpha) \quad \text{mit } \Omega = \frac{A_E}{f'^2} = \frac{A_S}{d^2}$$

(Gl. 3.35) wird allgemein als das "cos⁴-Gesetz" der Radiometrie bezeichnet. Berücksichtigt man die Abbildungsfehler einer realen Linse, ergibt sich experimentell ein Faktor zwischen

$\cos^{2.5}(\alpha) \dots \cos^4(\alpha)$ in Abhängigkeit der Größe des Gesichtsfeldes und der Qualität der Linse. [Lit. 3.4]. Mit dem Durchmesser D_L der Linse bzw. der Eintrittspupille am Objektiv erhält man die Gesamtfläche $A_L = \pi D_L^2/4$ der Eintrittsapertur. Berücksichtigt man zusätzlich, dass der Quotient D_L/f nach (Gl. 3.30) dem Kehrwert der Blendenzahl k entspricht, folgt für die Bestrahlungsstärke des Empfängers in Abhängigkeit der Strahldichte des Senders die Beziehung

$$(Gl. 3.36) \quad E = \frac{\Phi}{A_E} = L \cdot \frac{\pi \cdot D_L^2}{4} \cdot \frac{A_E}{A_E \cdot f^2} [\text{sr}] \cdot (\cos \varphi)^4 = L \cdot \frac{\pi}{4 \cdot k^2} [\text{sr}] \cdot (\cos \varphi)^4.$$

Bei der Berechnung der insgesamt vom Hintergrund auf einem Sensorelement hervorgerufenen Beleuchtungsstärke muss das gesamte Erfassungsfeld des Detektors berücksichtigt werden. Mit dem halben Öffnungswinkel θ_B , wobei es unerheblich ist, ob das Erfassungsfeld durch eine mechanische Blende oder durch eine Linse begrenzt wird, ergibt sich nach [Lit. 3.4] bei konstanter Strahldichte L_B die Beziehung

$$(Gl. 3.37) \quad E = \frac{d\Phi}{dA} = \int_{\theta=0}^{\theta_B} \int_{\varphi=0}^{2\pi} \frac{L_B \cdot dA \cdot \cos \theta}{dA} \cdot \sin \theta \, d\theta \, d\varphi = \pi \cdot L_B \cdot (\sin \theta_B)^2.$$

3.2.6 Strahlungsrauschen

Das Strahlungsrauschen beschreibt die statistische Unsicherheit der Strahlungsmenge, die von einer Strahlungsquelle in einem definierten Zeitintervall ausgesendet wird. Neben den in den vorherigen Abschnitten verwendeten energiebezogenen Größen ist zur Beschreibung des Strahlungsrauschens eine Umrechnung in äquivalente, photonenbezogene Größen zweckmäßig. Ausgehend von der Energie $W_P(\lambda) = h \cdot f = h \cdot c/\lambda$ eines einzelnen Photons lässt sich aus der spektralen Strahlungsenergie $W_\lambda(\lambda)$ die äquivalente Photonenzahl $n_{P\lambda}(\lambda)$ bestimmen.

(Gl. 3.38)

$$n_{P\lambda}(\lambda) = \frac{W_\lambda(\lambda)}{W_P(\lambda)} = \frac{\lambda}{c \cdot h} \cdot W_\lambda(\lambda) = k_P(\lambda) \cdot W_\lambda(\lambda) \quad \text{mit} \quad k_P(\lambda) = \frac{\lambda}{c \cdot h}$$

Die Photonenzahl $n_{P\lambda}(\lambda)$ entspricht dabei der Strahlungsenergie $W_\lambda(\lambda)$, welche die Grundlage für die Berechnung der energiebezogenen Größen in den vorherigen Abschnitten darstellt. Zur Umrechnung in photonenbezogene Größen müssen die spektralen energiebezogenen Größen mit dem Faktor $k_P(\lambda)$ nach (Gl. 3.38) multipliziert werden. Beispielsweise ergibt sich der zur spektralen Strahlungsleistung $\Phi_\lambda(\lambda)$ korrespondierende spektrale Photonfluss (Photonenzahl pro Zeitintervall t) aus

$$(Gl. 3.39) \quad \Phi_P(\lambda) = \frac{n_P(\lambda)}{t} = \frac{\lambda}{c \cdot h} \cdot \Phi(\lambda).$$

Die Integration von $n_{P\lambda}(\lambda)$ über alle Wellenlängen liefert die insgesamt vorhandene Photonenzahl n_P pro Zeitintervall t .

Nach [Lit. 3.4] folgt in photonenbezogenen Größen für die spektrale spezifische Ausstrahlung $M_{P\lambda S}(\lambda, T)$ eines schwarzen Strahlers das Plancksche Strahlungsgesetz nach (Gl. 3.40), für die spezifische Ausstrahlung $M_{PS}(T)$ gilt das Stefan-Boltzmann-Gesetz gemäß (Gl. 3.41).

$$(Gl. 3.40) \quad M_{P\lambda S}(\lambda, T) = \frac{d\Phi_{P\lambda S}(\lambda, T)}{dA_S} = \frac{2\pi \cdot c}{\lambda^4} \cdot \left[\exp\left(\frac{h \cdot c}{\lambda \cdot kT}\right) - 1 \right]^{-1}$$

$$(Gl. 3.41) \quad M_{PS}(T) = \sigma_P \cdot T^3 \quad \text{mit } \sigma_P = 1,52 \cdot 10^{11} \frac{\text{Photonen}}{\text{s} \cdot \text{cm}^2 \cdot \text{K}^4}$$

Das Wiensche Verschiebungsgesetz lautet unter Berücksichtigung der photonenbezogenen Ausstrahlung:

$$(Gl. 3.42) \quad \lambda_{\max} \cdot T = 3669 \text{ um} \cdot \text{K}$$

Aus (Gl. 3.13) erhält man für einen schwarzen Strahler der Temperatur 300 K eine mittlere Strahlungsleistung von ungefähr 46 mW pro cm^2 . Nach (Gl. 3.41) entspricht das einem mittleren Photonfluss $n_P = \bar{n}_P$ von ca. $4.1 \cdot 10^{18}$ Photonen pro Sekunde und pro cm^2 der strahlenden Oberfläche.

Der tatsächliche Photonfluss kann erheblich von diesem Mittelwert abweichen. Bei einer großen Photonenzahl und einem ausreichend langen Beobachtungszeitraum kann eine Gaußförmige Verteilungsdichte des am Empfangsort auftretenden Photonflusses angenommen werden [Lit. 3.4]. Die Varianz σ_P^2 der Gaußverteilung ist unter der Annahme einer Poissonverteilung bei der Photonemission gleich der mittleren Photonenzahl \bar{n}_P . Diese Näherung gilt für Temperaturen $T < 500\text{K}$ und für Wellenlängen von 300nm ...30um.

Für die Standardabweichung σ_P , die gleichzeitig dem Effektivwert des Rauschsignals entspricht, folgt in diesem Fall $\sigma_P = \sqrt{\bar{n}_P}$. Daraus resultiert ein Signal-Rausch-Verhältnis $S/N = \bar{n}_P / \sigma_P = \sqrt{\bar{n}_P}$, das auch bei der Verwendung eines idealen (nicht rauschenden) Detektors nicht verbessert werden kann.

3.3 Bildgebende Sensoren

3.3.1 Lichtempfindliche Elemente in CCD- und CMOS-Technologie

Zur Umwandlung eines optischen Signals in ein elektrisch messbares Signal ist ein optoelektronischer Wandler erforderlich. Als Grundstrukturen zur Realisierung von bildgebenden Sensoren werden häufig pn-Übergänge oder MOS-Kapazitäten eingesetzt, deren prinzipieller Aufbau in Abb. 3.15 dargestellt ist.

Ausgangspunkt für die nachfolgenden Betrachtungen ist die beleuchtungsabhängige Erzeugung von beweglichen Ladungsträgern in der Raumladungszone (RLZ) von Halbleiterbauelementen. Die von der Bestrahlung hervorgerufene Ladungsträgermenge (Anzahl der Elekt-

ron-Loch-Paare) ist über einen weiten Intensitätsbereich proportional zur einfallenden Lichtleistung. Anhand ihrer spektralen Zusammensetzung kann die optische Gesamtleistung in eine äquivalente Photonenzahl n_p pro Zeiteinheit t umgerechnet werden. Die Energie eines Photons ist dabei umgekehrt proportional zur Wellenlänge der Strahlung [s. Abschnitt 3.2.6]. Ohne Beleuchtung sind keine frei beweglichen Ladungsträger in der Raumladungszone vorhanden.

Zur Erzeugung eines Elektron-Loch-Paares muss die Energie eines Photons mindestens so groß sein wie die Energiedifferenz zwischen dem Leitungs- und dem Valenzband des Halbleitermaterials (bei Silizium 1.12 eV). Falls eine ausreichende Energie vorliegt, können auch mehrere Ladungsträgerpaare von einem einzelnen Photon erzeugt werden.

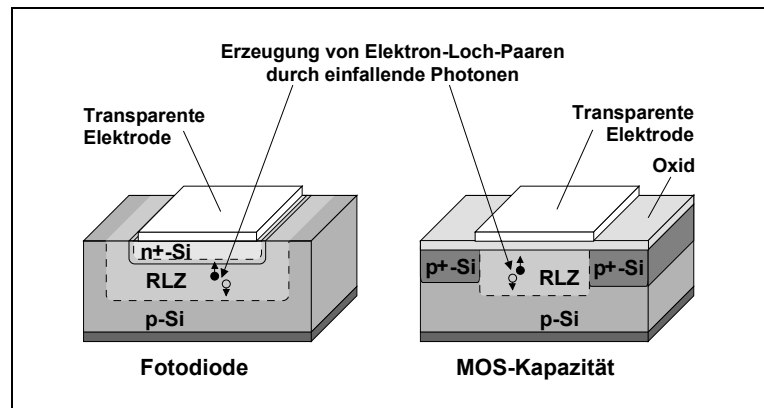


Abb. 3.15: Fotodiode (pn-Übergang) und MOS-Kapazität als Grundstrukturen zur Umwandlung von optischen in elektrische Signale

Aufgrund von Reflektionen an den verschiedenen Grenzflächen, durch Absorptionen an Defektstellen oder durch zu große Eindringtiefen sind allerdings nicht alle einfallenden Photonen am Umwandlungsprozess beteiligt. Als Maßzahl für die Effektivität dieses Prozesses wurde daher der Quantenwirkungsgrad η als Verhältnis der erzeugten Ladungsträgermenge n_{ph} zur Zahl der einfallenden Photonen n_p definiert (Gl. 3.43).

$$(Gl. 3.43) \quad \eta = \frac{n_{ph}}{n_p}$$

(Ladungsträgeranzahl n_{ph} ,
Photonenanzahl n_p)

$$(Gl. 3.44) \quad R = \frac{I_{ph}}{\Phi} = \frac{n_{ph} \cdot q/t}{n_p/t \cdot h \cdot c/\lambda} = \eta \frac{\lambda \cdot q}{h \cdot c}$$

(Lichtgeschwindigkeit $c = 2.99792438 \cdot 10^8$ m/s,
Plancksches Wirkungsquantum $h = 6.626176 \cdot 10^{-34}$ Ws²,
Elementarladung $q = 1.6021 \cdot 10^{-19}$ C)

Durch ein externes elektrisches Feld werden die in der Raumladungszone erzeugten Elektron-Loch-Paare voneinander getrennt. Schließt man den äußeren Stromkreis der pn-Diode, ist der Sperrstrom $I_{ph} = n_{ph} \cdot q/t$ durch den in Sperrrichtung betriebenen pn-Übergang direkt proportional zur einfallenden Strahlungsleistung ϕ . Für den Proportionalitätsfaktor R , der auch als Photostrom-Empfindlichkeit bezeichnet wird, gilt (Gl. 3.44).

Während die in Abb. 3.15 dargestellte MOS-Kapazität in CCD-Sensoren zur Akkumulation von Ladungsträgern verwendet wird, ermöglichen pn-Dioden die Realisierung von integrierenden und nicht integrierenden CMOS-Pixeln. Als Beispiel für einen nichtintegrierenden

Sensor soll an dieser Stelle ein CMOS-Imager mit logarithmischer Empfindlichkeitscharakteristik erwähnt werden, der aufgrund seiner nichtlinearen Kennlinie einen Dynamikbereich von mehr als 120 dB für die einfallende Strahlung bereitstellt. Die Ausgangsspannung derartiger Pixel ändert sich typischerweise um 30...40 mV pro Dekade der optischen Eingangsleistung. Bei normalen Beleuchtungsverhältnissen ergeben sich daher sehr kontrastarme Videobilder, die außerdem mit einer im Vergleich zu anderen Bildsensoren höheren Fixed-Pattern-Noise beaufschlagt sind. Eine ausführlichere Zusammenfassung ist in [Lit. 3.5] zu finden.

Zur Erläuterung der prinzipiellen Funktion eines integrierenden CMOS-Pixels soll zunächst das passive Pixel (PPS) nach Abb. 3.16 betrachtet werden. Die gestrichelt eingezeichnete Sperrschichtkapazität des pn-Übergangs ist dabei zunächst auf das Potential V der Videoleitung aufgeladen.

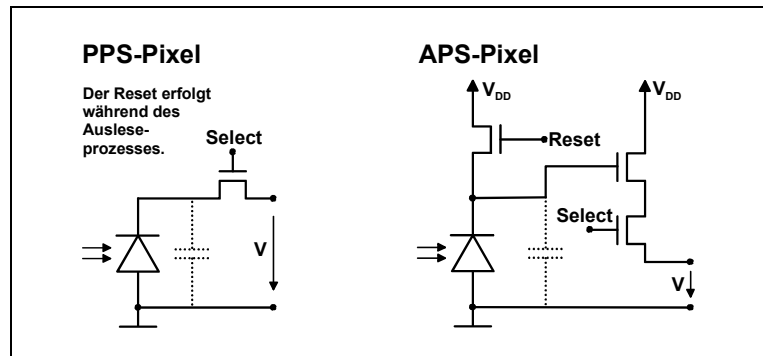


Abb. 3.16: Konzept eines CMOS-PPS- und eines CMOS-APS-Pixels (PPS = Passiv-Pixel-Sensor, APS = Aktiv-Pixel-Sensor)

Während der Integrationsphase ist die Fotodiode durch den gesperrten MOSFET vom restlichen, nicht dargestellten Teil der Ausleseschaltung isoliert. Die von der einfallenden Strahlung erzeugten, beweglichen Ladungsträger werden infolge der gespeicherten Sperrspannung getrennt, wobei die Elektronen zu der positiv vorgespannten Kathode wandern. Dort tragen sie zu einer Reduktion der während der Resetphase gespeicherten Ladungsmenge und damit zu einer Verringerung der über der Diode anliegenden Sperrspannung bei. Die Weite der Sperrschicht und damit auch die Sperrschichtkapazität ist abhängig von der Spannung über dem pn-Übergang. Mit zunehmender Entladung ergibt sich somit ein nichtlineares Übertragungsverhalten.

Am Ende der Integrationsphase wird der MOSFET leitend und verbindet die Photodiode mit einem Verstärker, der den Reset der Sperrschichtkapazität durchführt und gleichzeitig die zum erneuten Aufladen erforderliche Ladungsmenge misst. Im Anschluss an die Resetphase ist die Ausgangsspannung des Verstärkers proportional zu der gemessenen Ladungsmenge und steht zur weiteren Verarbeitung zur Verfügung.

Ein wesentlicher Nachteil konventioneller PPS-Matrixsensoren ist die hohe Kapazität der Videoleitung zwischen dem Pixel und der Eingangsstufe des Ausleseverstärkers. Insbesondere im Vergleich zu CCD-Bildsensoren weisen sie daher deutlich schlechtere Rauscheigenschaften auf [Lit. 3.5] und außerdem eine geringere Empfindlichkeit auf [Lit. 3.7]. Vorteil-

haft ist dagegen der wahlfreie Zugriff auf einzelne Pixel des Arrays, der sich aber auch mit APS-Sensoren realisieren lässt.

Im rechten Teil von Abb. 3.16 ist die typische Struktur eines CMOS-APS-Pixel dargestellt. Im Vergleich zum PPS-Pixel wird die Ausgangsspannung des APS-Pixels über einen oder mehrere MOSFET-Source-Folger gepuffert, bevor die Durchschaltung auf die Videoleitung erfolgt. Die im Vergleich zur Kapazität der Videoleitung deutlich kleinere Eingangskapazität des Source-Folgers wirkt sich sowohl auf das Rauschverhalten als auch auf die Empfindlichkeit des Pixels positiv aus. Die von einem Elektron auf der resultierenden Integrationskapazität hervorgerufene Spannungsänderung wird dabei als "conversion gain" bezeichnet. Infolge der Pufferung muss ein zusätzlicher Reset-Transistor eingesetzt werden. Aufgrund der höheren Transistoranzahl im Vergleich zur PPS-Technologie ergeben sich kleinere lichtempfindliche Flächen, wenn man eine gleichbleibende Gesamtpixelfläche unterstellt. Das Verhältnis der lichtempfindlichen Fläche A_{ph} zur Gesamtfläche A des Pixels wird als Füllfaktor FF bezeichnet. Der Füllfaktor nach (Gl. 3.45) wird gewöhnlich in Prozent angegeben:

$$(Gl. 3.45) \quad FF = \frac{A_{ph}}{A} \cdot 100\%$$

Abb. 3.17 zeigt den grundsätzlichen Aufbau von Imagern mit PPS- bzw. APS-Pixeln. Einzelne Pixel werden über den integrierten Zeilen- und Spaltendecoder ausgewählt. Obwohl dieses Konzept grundsätzlich einen wahlfreien Pixelzugriff ermöglicht, stellen manche CMOS-Bildsensoren nur einen kompletten Frame zur Verfügung. Die gesamte Timing- und Ansteuerlogik sowie ggf. die Umwandlung in ein normkonformes Videosignal ist in diesem Fall zusätzlich auf dem Chip integriert.

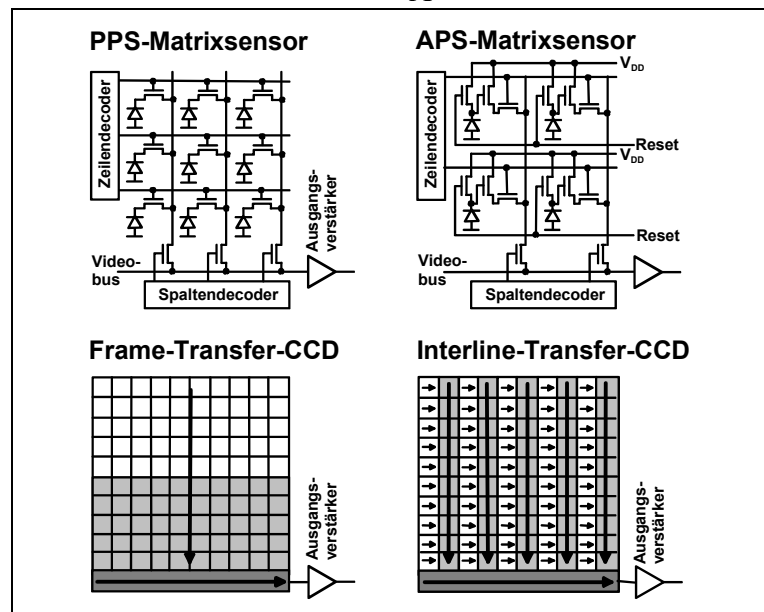


Abb. 3.17: Auslesetechniken von PPS- und APS-Matrixsensoren sowie von Frame-Transfer- und Interline-Transfer-CCD's

Zum Vergleich wurde in Abb. 3.17 die prinzipielle Architektur eines Frame-Transfer-CCD's (FT-CCD) und eines Interline-Transfer-CCD's (IT-CCD) dargestellt. Die auf den weiß dargestellten MOS-Kapazitäten integrierten Ladungen werden beim FT-CCD zeilenweise in die grau dargestellten, optisch abgedeckten Bereiche geschoben, bevor das Austakten beginnt. Diese Struktur hat einen Füllfaktor von 100%, allerdings kann eine Nachbeleuchtung während der Umspeicherung erfolgen. Beim IT-CCD werden die Nachbeleuchtungseffekte reduziert,

da das komplette Bild zunächst in die gegen Beleuchtung geschützten vertikalen Schieberegister übertragen wird. IT-CCDs dominieren heute den Markt bei Standardanwendungen. Der Füllfaktor beim IT-CCD beträgt 25-30% [Lit. 3.7]. Ein CCD liefert ein komplettes Bild (progressive scan) oder aber zwei Felder (interlaced mode), von denen eines die ungeraden und das zweite die geraden Zeilen eines Bildes enthält.

Zum Ladungstransport innerhalb der CCD-Schieberegister, der in [Lit. 3.8] ausführlich beschrieben ist, sind je nach CCD-Typ 2...4 zusätzliche Spannungen erforderlich, die mit speziellen Pegelwandlern erzeugt werden müssen. Im Vergleich zur 3.3 V- oder 5.0 V-Spannungsversorgung von CMOS-Sensoren ergibt sich daher ein höherer Bauteil Aufwand für das Netzteil. Neben einem Timinggenerator zur Erzeugung der CCD-Taktsignale ist außerdem ein externer Video-AD-Wandler erforderlich, der technologiebedingt nicht auf dem CCD-Chip integriert werden kann. Insbesondere die externen Komponenten erhöhen den Strombedarf des gesamten CCD-Sensorfrontends erheblich.

Die Integrationsmöglichkeiten bei CMOS-Bildsensoren wurden bereits vorher erwähnt. Die gesamte Ansteuerelektronik (Zeilen-, Spaltendecoder, automatische Verstärkungsregelung, etc.) sowie ein 8-, 10- oder 12-Bit-AD-Wandler sind meist bereits auf dem Chip vorhanden. Als externe Kamerakomponenten sind oftmals lediglich eine 3.3 V- oder 5 V-Spannungsversorgung und eine Taktquelle erforderlich. Im Vergleich zu einem konventionellen CCD-Frontend folgt eine mindestens um den Faktor 2...3 reduzierte Leistungsaufnahme.

Ausgehend von den zuvor beschriebenen Standard-CMOS-Sensoren findet man in der Literatur hochintegrierte Smart-Pixel-Sensoren, die zusätzliche Schaltkreise zur Signalvorverarbeitung enthalten. In [Lit. 3.5] werden beispielsweise intelligente Pixel mit integrierter Offsetkorrektur zur Erhöhung der Dynamikbereichs oder mit interner Verstärkung zur Steigerung der Sensorempfindlichkeit vorgestellt.

3.3.2 Kenngrößen zur Beurteilung der Sensorperformance

3.3.2.1 Geometrische Daten

Die Gesamtfläche eines CMOS-Sensors ist theoretisch nur von der Wavergöße begrenzt. Aufgrund der zusätzlich integrierten Komponenten ist die lichtempfindliche Sensorfläche, die auch für Erfassungsfeldberechnungen in Verbindung mit einer abbildenden Optik benötigt wird, meist deutlich kleiner als die gesamte Chipfläche.

Innerhalb der Sensorfläche sind die einzelnen Pixel angeordnet, wobei der lichtempfindliche Teil aus dem Füllfaktor nach (Gl. 3.45) und der Gesamtpixelfläche berechnet werden kann. Je nach CMOS-Chip kann der Füllfaktor durch die Verwendung von Mikrolinsen vergrößert sein. Der Mittenabstand zwischen benachbarten, bei CMOS-Sensoren oftmals quadratischen Pixeln wird im folgenden als Pitch bezeichnet.

3.3.2.2 Pattern Noise

Die im Vergleich zu CCD-Kameras höhere Fixed Pattern Noise (FPN) von CMOS-Sensoren ist auf prozessbedingte Variationen der elektrischen Eigenschaften aller elektronischen Komponenten im Pixel und in den Auslestufen zurückzuführen. Für Videoanwendungen muss die FPN auf jeden Fall unterdrückt werden, was beispielsweise durch die Subtraktion eines vorher gespeicherten Referenzbildes realisiert werden kann.

Andere Möglichkeiten zur Unterdrückung der FPN wie "correlated sampling" (CS) bzw. das aus der CCD-Technik bekannte "correlated double sampling" (CDS) sind auf integrierende Bildwandler beschränkt. Der Unterschied zwischen CS und CDS besteht in dem Zeitpunkt, an dem der Resetpegel des Pixels gemessen wird. Beim CS wird der auf die Integrationsphase folgende Resetpegel des Pixels ermittelt, während beim CDS-Verfahren die Bestimmung des Resetpegels unmittelbar vor Beginn des Integrationsvorgangs erfolgt. Beide Verfahren bilden anschließend die Differenz zwischen der Ausgangsspannung des Pixels am Ende des Integrationsprozesses und dem jeweiligen Resetpegel, wodurch statische Abweichungen im gesamten Übertragungsweg wirkungsvoll unterdrückt werden.

Mit Hinblick auf die vorliegende Anwendung soll angemerkt werden, dass eine separate FPN-Unterdrückung bei Auswerteverfahren, die ohnehin Differenzen zwischen verschiedenen Videobildern auswerten, nicht unbedingt notwendig ist.

Ein weiterer Beitrag zur Pattern Noise entsteht durch die "photo response non-uniformity" (PRNU), die neben technologischen Parametern auch von der Wellenlänge der einfallenden Strahlung bestimmt wird.

3.3.2.3 Dunkelstrom und Dunkelstromrauschen

Außer den von den Photonen generierten Ladungsträgern werden an Verunreinigungen des Halbleiters, an den Materialgrenzen des Substrats, etc. zusätzliche Ladungsträger erzeugt, die nicht von dem Nutzsignal unterschieden werden können. Diese Ladungsträger sind unter dem Begriff "Dunkelstrom" zusammengefasst. Der Dunkelstrom wird zwar von ortsfesten Verunreinigungen erzeugt, allerdings ist deren Verteilung über den Chip nicht regelmäßig. Hinzu kommt, dass der Dunkelstrom stark temperaturabhängig ist. Er verdoppelt sich alle 8-9 °C.

Tab. 3.2 vermittelt einen Eindruck von typischen Dunkelströmen bei Raumtemperatur. Der sehr geringe Dunkelstrom hochwertiger CCD-Sensoren ist darauf zurückzuführen, dass diese Technologie speziell mit Hinblick auf niedrige Dunkelströme optimiert wurde.

Technologie	Dunkelstrom
CCD	10-50 pA/cm ²
DRAM (0.5-um-Prozess)	25 pA/cm ²
Optimized CMOS	50-200 pA/cm ²
Standard CMOS	1000 pA/cm ²

Tab. 3.2: Typische Dunkelströme bei Raumtemperatur [Lit. 3.7]

In Verbindung mit integrierenden Bildaufnehmern ergibt sich eine vom Dunkelstrom bestimmte, maximale Integrationsdauer. Eine Maßnahme zur Vermeidung hoher Dunkelströme ist die Kühlung des Chips, was aber für den vorliegenden Anwendungsfall nicht praktikabel ist. Dennoch sollte beachtet werden, dass dem Bildsensor keine zusätzliche Wärme durch die übrigen Komponenten des Bildwächters zugeführt wird.

3.3.2.4 Rauschquellen

Sowohl der von den Photonen erzeugte Photostrom I_{ph} und als auch der Dunkelstrom I_{dk} liefern einen Beitrag zum Gesamttrauschen des Sensors. Der Effektivwert des von der Beleuchtungsstärke E_{ph} hervorgerufenen Rauschanteils $n_{rms,ph}$ kann mit (Gl. 3.46) berechnet werden, während (Gl. 3.47) den dunkelstromabhängigen Rauschanteil $n_{rms,dk}$ des Schrotrauschens beschreibt [Lit. 3.7]. Die Gleichungen liefern jeweils die Anzahl der beteiligten Rauschelektronen.

$$(Gl. 3.46) \quad n_{rms,ph} = \sqrt{\eta \cdot k_p \cdot E_{ph} \cdot A_{ph} \cdot t} \quad (Gl. 3.47) \quad n_{rms,dk} = \sqrt{I_{dk} \cdot t/q}$$

(k_p = Faktor zur Umwandlung der Strahlungsleistung in einen äquivalenten Photonenfluss, t = Integrationszeit, A_{ph} = Pixelfläche, q = Elementarladung)

Zusätzlich zum Schrotrauschen des Photo- und des Dunkelstromes sind weitere Rauschquellen vorhanden. Infolge des Johnson-Rauschens im Kanalbereich des Reset-Transistors treten statistische Schwankungen des Resetpegels auf, die als Resetrauschen bzw. allgemein als kTC-Rauschen bezeichnet werden. (Gl. 3.48) dient zur Abschätzung der Rauschelektronenanzahl $n_{rms,res}$.

$$(Gl. 3.48) \quad n_{rms,res} = \frac{\sqrt{k \cdot T \cdot C}}{q} \quad (Gl. 3.49) \quad n_{rms,1/f} \propto \frac{C}{q} \cdot \sqrt{\frac{I_{ph}^\alpha \cdot B}{f^\beta}}; \alpha = \frac{1}{2} \dots 2, \beta \approx 1$$

(k = Boltzmann-Konstante, T = absolute Temperatur, C = effektive Integrationskapazität, B = Rauschbandbreite, f = Frequenz, q = Elementarladung)

Das nach (Gl. 3.49) zu $1/\sqrt{f}$ proportionale Flickerrauschen (1/f-Rauschen) entsteht durch statistische Schwankungen der Ladungsträgerbeweglichkeit und der Ladungsträgerkonzentrationen im Kanalbereich des Source-Folgers. Die Fluktuationen treten insbesondere an den Kontaktstellen unterschiedlicher Materialien auf, die Ursachen dieses Rauschterms sind aber noch nicht abschließend geklärt.

Schwankungen des Ausgangssignals infolge des Reset- und Flickerrauschens können durch das vorne beschriebene CDS-Verfahren reduziert werden.

Bei höheren Frequenzen ist das Flickerrauschen aufgrund der $1/\sqrt{f}$ -Abhängigkeit im Vergleich zum thermischen Rauschen aller im Übertragungsweg vorhandenen Komponenten vernachlässigbar. Reduziert man diesen Rauschanteil auf das thermische Rauschen im Kanalbe-

reich des Source-Folgers, lässt sich mit (Gl. 3.50) die thermisch bedingte Rauschelektronenzahl ermitteln. Mit dem auf den Eingang bezogenen Kanalwiderstand R' folgt die thermisch generierte Rauschelektronenzahl

$$(Gl. 3.50) \quad n_{\text{rms,th}} = \frac{C}{q} \cdot \sqrt{4 \cdot k \cdot T \cdot B \cdot R'}$$

(k = Boltzmann-Konstante, T = Integrationszeit, C = effektive Integrationskapazität, B = Rauschbandbreite, q = Elementarladung)

Die Reduktion des thermischen Rauschens erfordert eine Optimierung der technologischen Parameter. Alternativ bietet sich eine Verringerung der Sensortemperatur an, die sich weiterhin vorteilhaft auf die Höhe des Dunkelstromes und des zugehörigen Rauschanteils auswirkt.

3.3.2.5 Signal-Rausch-Verhältnis und Dynamikbereich

Die insgesamt vorhandene effektive Rauschelektronenzahl $n_{\text{rms,N}}$ ergibt sich aus der geometrischen Addition aller beteiligten Rauschterme gemäß

$$(Gl. 3.51) \quad n_{\text{rms,N}} = \sqrt{n_{\text{rms,shot}}^2 + n_{\text{rms,read}}^2}$$

Der Term $n_{\text{rms,shot}}$ beinhaltet das Schrotrauschen des Nutzsignals und des Dunkelstromes, während die Readout-Noise $n_{\text{rms,read}}$ alle übrigen Rauschbeiträge zusammenfasst. Auch das Quantisierungsrauschen eines im Übertragungsweg vorhandenen AD-Wandlers kann an dieser Stelle berücksichtigt werden. An dieser Stelle sollte allerdings beachtet werden, dass das Schrotrauschen des Photostromes zur Definition des Dynamikbereichs des Sensors gewöhnlich nicht berücksichtigt wird.

Abb. 3.18 zeigt die grundlegenden Definitionen der Empfindlichkeit und des Dynamikbereichs eines integrierenden Bildwandlers.

Der Dynamikbereich des Sensors wird durch das Signal-Rausch-Verhältnis SNR in (Gl. 3.52) charakterisiert.

$$(Gl. 3.52) \quad \text{SNR} = 20 \cdot \log_{10} \frac{S}{N}$$

Der optische Dynamikbereich ist dagegen durch die minimal erforderliche und durch die maximale Beleuchtungsstärke zur Signaldetektion festgelegt, wobei sich letztere durch die maximal mögliche Ladungsträgeranzahl auf der Integrationskapazität ergibt.

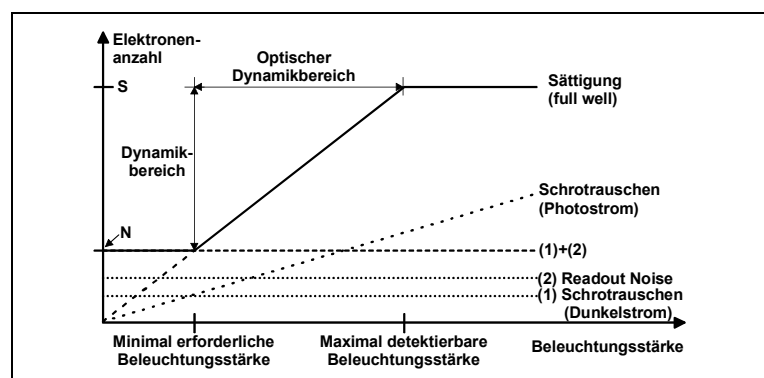


Abb. 3.18: Dynamikbereich von integrierenden Bildsensoren bei gegebener Integrationszeit

3.3.2.6 Spektrale Empfindlichkeit

Abb. 3.19 zeigt den Verlauf der auf den Maximalwert $S_{\lambda, \max}$ normierten, spektralen Empfindlichkeit $S_{\lambda}(\lambda)$ des monochromen Bildsensors OV6120. Zum Vergleich ist zusätzlich das Empfindlichkeitsdiagramm des identischen, allerdings mit RGB-Farb- und IR-Sperrfilter ausgestatteten Farbchips OV6620 angegeben. Die prinzipiellen Verläufe sind typisch für konventionelle CMOS-Bildsensoren.

In Verbindung mit einer IR-Beleuchtung muss ein monochromer Bildsensor ohne IR-Sperrfilter für den Bildwächter verwendet werden. Ein weiterer Vorteil im Vergleich zum Farbchip ist die insgesamt höhere Empfindlichkeit im Bereich von 400..1100 nm.

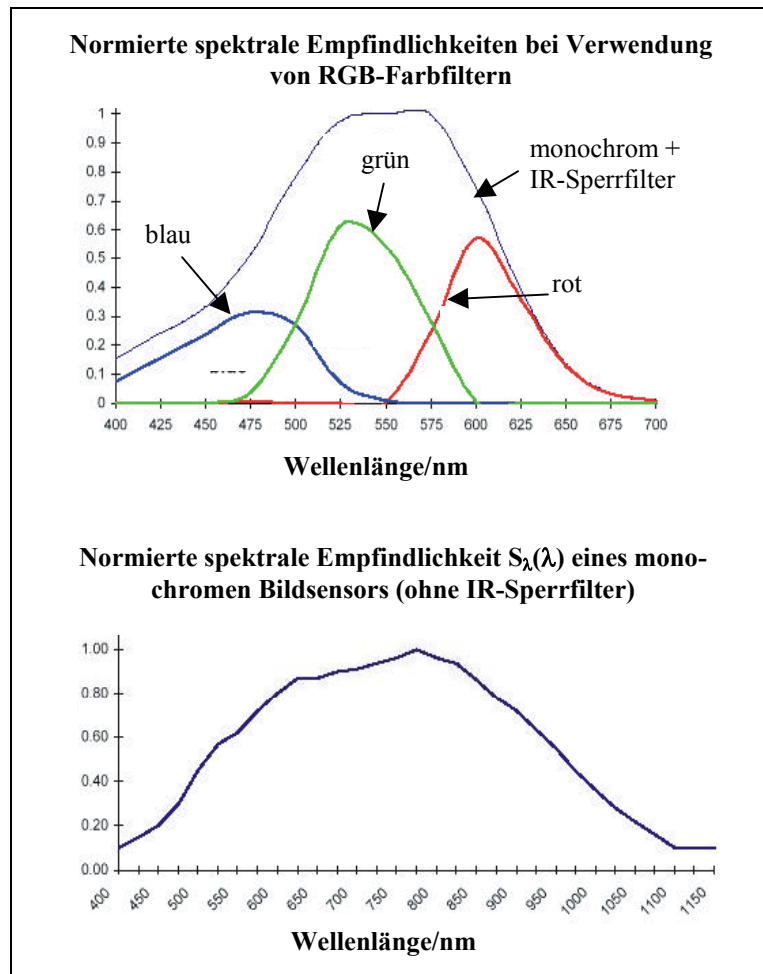


Abb. 3.19: Verlauf der auf den Maximalwert normierten, spektralen Empfindlichkeit der CMOS-Sensoren OV6620 und OV6120

Die Empfindlichkeit beschreibt ganz allgemein den Zusammenhang zwischen der Ausgangsgröße des Sensors, z. B. einer analogen oder digitalisierten Ausgangsspannung, und der optischen Anregung. Als Eingangsgröße wird oft die Bestrahlungsstärke E des Sensors verwendet. Da die Ausgangsspannung bei integrierenden Sensoren allerdings nicht nur von der Bestrahlungsstärke, sondern auch von der Integrationsdauer T_{int} abhängt, wird die Empfindlichkeit gewöhnlich auf den Integrationszeitraum normiert.

Mit der spektralen Empfindlichkeit $S(\lambda) = S_{\lambda, \max} \cdot S_{\lambda}(\lambda)$, der spektralen Zusammensetzung des Strahlungsflusses $\phi(\lambda) = \phi_{\lambda, \max} \cdot \phi_{\lambda}(\lambda)$ sowie der Pixelfläche A_{ph} kann die Beleuchtungsstärke $E(\lambda) = E_{\lambda, \max} \cdot E_{\lambda}(\lambda) = \phi_{\lambda, \max} \cdot \phi_{\lambda}(\lambda) / A_{\text{ph}}$ und damit die Ausgangsspannung U in Abhängigkeit der Beleuchtungsstärke des Pixels bestimmt werden.

$$(Gl. 3.53) \quad U = T_{\text{int}} \cdot E_{\lambda, \max} \cdot S_{\lambda, \max} \int_{\lambda_{\min}}^{\lambda_{\max}} E_{\lambda}(\lambda) \cdot S_{\lambda}(\lambda) d\lambda$$

Im Vergleich zur breitbandigen spektralen Zusammensetzung des Sonnenlichts oder der einer konventionellen Raumbeleuchtung ist die Strahlung einer LED sehr schmalbandig. Um die maximale Sensorempfindlichkeit für eine LED-Beleuchtung bereitzustellen, sollte die Peak-Wellenlänge der LED mit dem Maximum der spektralen Empfindlichkeit $S_{\lambda_{\max}}$ übereinstimmen. Der Umstand, dass dieses Maximum bei Verwendung eines CMOS-Sensors ohne IR-Sperrfilter im nahen IR-Bereich liegt, kommt der Forderung nach einer für den Menschen unsichtbaren Ergänzungsbeleuchtung entgegen.

Die in den Datenblättern der Hersteller oftmals angegebene Empfindlichkeit S in der Einheit $V/(Lux \cdot s)$ darf nicht mit der maximalen spektralen Empfindlichkeit $S_{\lambda_{\max}}$ verwechselt werden. Die Empfindlichkeit S wird mit einer fotometrisch bewerteten Strahlungsleistung oder Beleuchtungsstärke ermittelt und beschreibt die Sensorausgangsspannung in Abhängigkeit der Integrationszeit und der Strahlungsleistung einer Lichtquelle mit der Farbtemperatur T .

Beispielsweise entspricht die Normlichtart A (Normlichtarten der Internationalen Beleuchtungs-Kommission CIE) dem Licht eines schwarzen Strahlers mit der Temperatur 2856 K. Die Normlichtart A ist definiert im Wellenlängenbereich von ca. 300 nm...830 nm und kann in sehr guter Näherung mit einer Glühlampe der Farbtemperatur 2856 K erzeugt werden. Diese Normlichtart wird angewendet, wenn Beleuchtungsbedingungen entsprechend dem Glühlampenlicht hergestellt werden sollen [Lit. 3.14].

3.3.2.7 Zusätzliche Eigenschaften

Wie bereits vorher erwähnt wurde, bietet die CMOS-Technologie die Möglichkeit, weitere Komponenten auf dem Chip zu integrieren. Bereits angesprochen wurden intelligente Pixel, die Funktionen zur Signal- bzw. Bildvorverarbeitung enthalten und der wahlfreie Pixelzugriff, der allerdings nicht von allen CMOS-Imagern unterstützt wird. Der wahlfreie Zugriff ist hauptsächlich in Verbindung mit nicht integrierenden Pixeln anzutreffen, während bei integrierenden Imagern oftmals der von CCD-Matrixsensoren bekannte Frame-Transfer dominiert.

Im Gegensatz zu CCD-Empfängern, bei denen der Integrationsprozess für alle Pixel eines Frames oder Fields gleichzeitig abläuft, verwenden CMOS-Sensoren bis auf wenige Ausnahmen einen "Rolling-Shutter"-Mechanismus zur Steuerung der Belichtungszeit. Dabei werden lediglich alle Pixel innerhalb einer Zeile des CMOS-Sensors zeitlich synchron belichtet. Die aus der konventionellen Bildverarbeitung bekannte Aufnahme von schnell bewegten Objekten mit einer stroboskopischen Beleuchtung ist mit diesen Sensoren nicht möglich.

Während bei Standard-CMOS-Sensoren die Integrationsphase durch den Auslesevorgang beendet und die beteiligten Pixel in den Ausgangszustand zurückgesetzt werden, gibt es spezielle Bildaufnehmer, bei denen die Ausgangsspannung während der Ladungsträgerakkumulation mehrfach abgetastet wird. Vorteilhaft anwenden lässt sich dieses Prinzip zur Einstellung der Integrationszeit in Abhängigkeit der Beleuchtungsverhältnisse im Erfassungsfeld.

Die maximale Bildwiederholfrequenz von CMOS-Imagern für Standardanwendungen orientiert sich an der Bildwiederholfrequenz konventioneller Videostandards (RS-170 (USA, Japan): 30 Hz, CCIR (Europa): 25 Hz). Je nach Pixelanzahl und auch als zusätzliche Eigenschaft bieten viele CMOS-Sensoren geringere Bildwiederholraten an. Bei der Verwendung einer niedrigeren Bildwiederholrate in Verbindung mit einer zur Netzfrequenz asynchronen Integrationsphase müssen die Auswirkungen des Lichtflimmerns konventioneller Raumbelichtungen bei der Anwendung eines differenzbildenden Auswerteverfahrens beachtet werden.

Die Auslesegeschwindigkeit der Pixel, die durch die elektrische Bandbreite des gesamten Übertragungspfades begrenzt ist, und die Pixelanzahl innerhalb des Frames bestimmen die maximale Bildwiederholfrequenz. Insbesondere bei großen Pixelanzahlen werden mehrere, parallele Auslesepfade verwendet, um eine zufriedenstellende Auslesegeschwindigkeit zu erreichen. Der Flaschenhals im gesamten Übertragungsweg ist der Onchip-AD-Wandler. Die hohe Pixelrate (Größenordnung 10 MHz) erfordert sehr schnelle Wandler mit Auflösungen von 8...12 Bit für Standardanwendungen.

Zur Erhöhung der Bildrate bei konventionellen CMOS-Sensoren ist es oftmals möglich, ein oder mitunter auch mehrere, kleine Auslesefenster zu definieren (Windowing). Ein anderer Ansatz zur Steigerung der Framerate oder zur Datenreduktion sind Subsampling-Techniken, die allerdings mit einer reduzierten Ortsauflösung im Erfassungsfeld verbunden sind.

Der im Vergleich zur CCD-Technik geringere Aufwand an externen Komponenten zur Realisierung eines bildverarbeitenden Systems und der damit einhergehende, deutlich geringere Stromverbrauch des Sensorfrontends wurde bereits vorher erwähnt. Sie sind eine grundlegende Voraussetzung zur Realisierung eines CMOS-Bildwächters.

3.3.3 Kamera- und Weltkoordinatensystem

Bildverarbeitende Systeme bilden eine 3D-Szene auf einen 2D-Bildsensor ab. Bei dieser Projektion geht die Tiefeninformation der räumlichen Szene verloren. Im folgenden wird eine Bildmatrix mit der Spaltenzahl M und der Zeilenzahl N nach Abb. 3.20 betrachtet. Der Index m dient zur Spalten- und der Index n zur Zeilenadressierung des Pixels $P(m,n)$.

Unter Beachtung der Mittenabstände $p_m = h/M$ in horizontaler und $p_n = v/N$ in vertikaler Richtung ergeben sich folgende Beziehungen zwischen den Indizes m und n und den in Abb. 3.20 dargestellten Kamerakoordinaten x_b und y_b :

$$(Gl. 3.54) \quad x_b = \left(\frac{1-M}{2} + m \right) \cdot \frac{h}{M}$$

$$(Gl. 3.55) \quad y_b = \left(\frac{N-1}{2} - n \right) \cdot \frac{v}{N}$$

Der Bildsensor ist nach Abb. 3.21 in der Brennebene des Objektivs angeordnet, während der Ursprung des Weltkoordinatensystems in die Hauptebene der abbildenden Optik gelegt wurde.

Für den Mittelpunkt eines beliebigen Pixels $P(x_b, y_b, z_b)$ in Weltkoordinaten folgt dann die Beziehung

(Gl. 3.56)

$$P(x_b, y_b, z_b) = P(x_b, y_b, -f')$$

Der Winkel φ_{mn} , den die Gerade durch den Mittelpunkt des Pixels $P(x_b, y_b, z_b)$ und den Ursprung des Weltkoordinatensystems mit der optischen Achse bildet, kann anhand von (Gl. 3.54) und (Gl. 3.55) unter Beachtung von (Gl. 3.56) bestimmt werden:

$$(Gl. 3.57) \quad \varphi_{mn} = \tan^{-1} \left(\frac{\sqrt{x_b^2 + y_b^2}}{f'} \right) = \tan^{-1} \left(\sqrt{\left(\frac{1-M}{2} + m \right)^2 \cdot \frac{h^2}{M^2} + \left(\frac{N-1}{2} - n \right)^2 \cdot \frac{v^2}{N^2}} / f' \right)$$

Bildet man das Pixel mit dem Mittelpunkt $P(x_b, y_b, z_b)$ auf eine zur Bildebene parallele Ebene im Objektraum, die sich an der Stelle $z = a$ befindet, ab, folgt aus dem Strahlensatz der Zentralprojektion

$$(Gl. 3.58) \quad P(x_b, y_b, z_b) = P(x_b, y_b, -f') \rightarrow P(x_g, y_g, z_g) = P \left(\frac{-a}{f'} x_b, \frac{-a}{f'} y_b, a \right).$$

Bei der Projektion des Punktes $P(x_b, y_b, -f')$ in den Punkt $P(x_g, y_g, a)$ tritt somit ein konstanter Skalierungsfaktor $-a/f'$ zur Umrechnung der transversalen Kamerakoordinaten x_b und y_b in die transversalen Weltkoordinaten x_g und y_g auf. Das negative Vorzeichen berücksichtigt die am Projektionszentrum vorgenommene Punktspiegelung. Sinngemäß erfolgt die Berechnung der in Objektraum projizierten Pixelfläche mit dem Faktor $(a/f')^2$ (s. Kap. 3.3.4).

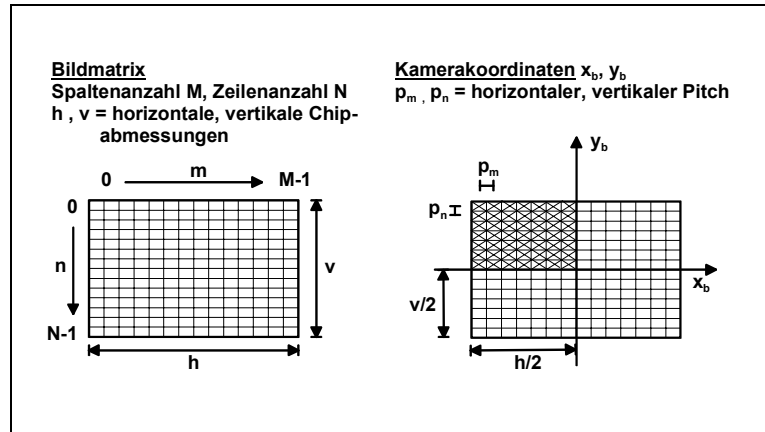


Abb. 3.20: Indizierung der Bildmatrix und Kamerakoordinatensystem unter Berücksichtigung des Pitch zwischen benachbarten Pixeln

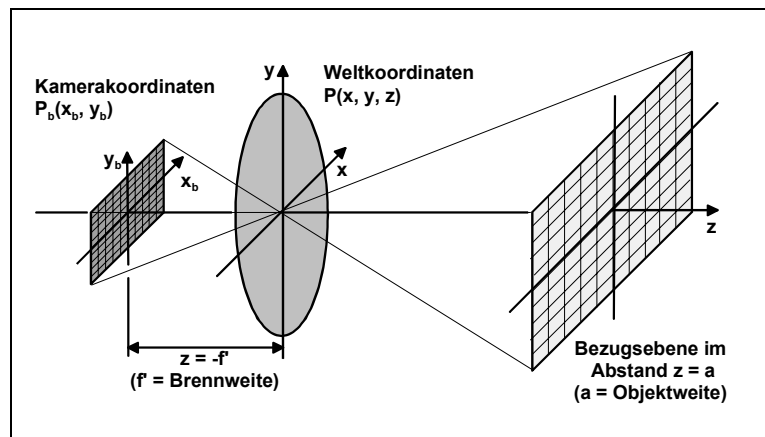


Abb. 3.21: Kamera- und Weltkoordinatensystem. Projektion der Pixel in eine zur Bildebene parallele Ebene im Objektraum.

3.3.4 Gesichtsfeld und Bereichsauflösung

Die heute für Kameraanwendungen verfügbaren Standardobjektive haben sich an den jeweils verfügbaren CCD-Bildformaten (s. Tab. 3.3) orientiert. Die größte Auswahl an Objektiven existiert für 1/2"- und 1/3"-Bildformate, wobei die Verwendung in Verbindung mit kleineren Imagerformaten durchaus möglich ist.

Der umgekehrte Fall sollte dagegen vermieden werden, was insbesondere beim Einsatz von CMOS-Bildsensoren zu berücksichtigen ist. Viele Objektivhersteller stellen Umrechnungstabellen zur Verfügung, denen der Bildwinkel in Abhängigkeit des verwendeten CCD-Formats entnommen werden kann

Ebenfalls in Verbindung mit der CCD-Technik wurde der C-Mount- und der CS-Mount-Standard für Objektive eingeführt (s. Abb. 3.22). Diese Standards spezifizieren unter anderem das Auflagemaß zwischen dem Objektivflansch und der Bildebene sowie das Kamera- bzw. Objektivgewinde. Durch den Einsatz eines 5 mm-Zwischenringes kann ein CS-Mount-Objektiv in ein C-Mount-Objektiv umgewandelt werden.

CCD-Format	Sensorfläche Breite x Höhe	Sensor- diagonale
1/6"	2.1 mm x 1.6 mm	2.64 mm
1/5"	2.6 mm x 1.9 mm	3.22 mm
1/4"	3.2 mm x 2.4 mm	4.00 mm
1/3"	4.8 mm x 3.6 mm	6.00 mm
1/2"	6.4 mm x 4.8 mm	8.00 mm
2/3"	8.8 mm x 6.6 mm	11.00 mm
1"	12.8 mm x 9.6 mm	16.00 mm

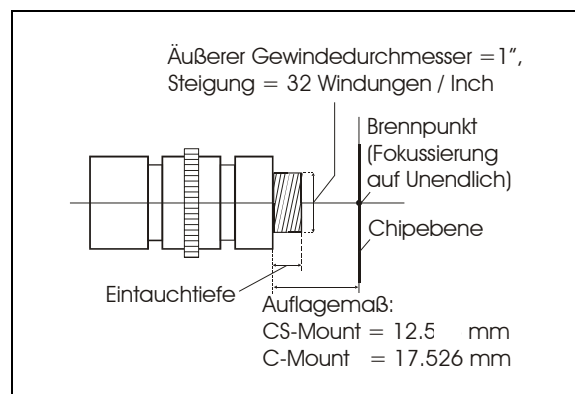


Abb. 3.22: Auflagemaß von C-Mount-/CS-Mount-Objektiven [Lit. 3.13]

Tab. 3.3: Bildformate von CCD-Matrixsensoren [Lit. 3.12] und daraus berechnete Bilddiagonlen

C-Mount- bzw. CS-Mount-Objektive sind im Vergleich zu den in Tab. 3.4 angegebenen S-Mount-Objektiven recht groß. Um ein möglichst kleines, unauffälliges und kostenoptimales Design eines Bildwächters zu ermöglichen, wird im folgenden die Verwendung der dort aufgeführten Kleinobjektive bevorzugt.

Im weiteren Verlauf werden die in Verbindung mit gängigen CCD-Formaten erzielbaren Bildwinkel in Abhängigkeit der Objektivbrennweite sowie die resultierenden Objektfeldauflösungen bei Verwendung von QCIF-, CIF-, QVGA- und VGA-Chips berechnet. Ausgangspunkt für die Abschätzung des resultierenden Bildwinkels ist (Gl. 3.33).

S-Mount-Objektive	B 1620	B 1220	B 1020	B 8020	B 6022	B 4318	B 4022	B 3620	B 3020	B 2525
CCD-Bildformat	1/3"	1/3"	1/3"	1/3"	1/3"	1/3"	1/3"	1/3"	1/3"	1/3"
Blendenzahl k (fest)	2.0	2.0	2.0	2.0	2.2	1.8	2.2	2.0	2.0	2.5
Brennweite f/[mm]	16	12	10.0	8.0	6.0	4.3	4.0	3.6	3.0	2.5
Abmessungen/[mm x mm]	∅14.0 x 15.5	∅14.0 x 14.0	∅14.0 x 14.8	∅14.0 x 13.8	∅14.0 x 16.0	∅15.0 x 14.3	∅14.0 x 16.0	∅14.0 x 16.5	∅14.0 x 15.6	∅15.0 x 17.4
Montagegewinde		M12	M12	M12	M12	M12	M12	M12	M12	M12

Tab. 3.4: S-Mount-Objektive

Die für CIF- und VGA-Auflösungen spezifizierte Pixelanzahl sowie die daraus abgeleitete Pixelzahl für das zugehörige 1/4-Bildformat sind in Tab. 3.5 zusammengefasst. Bei manchen CMOS-Sensoren ist die Umschaltung zwischen verschiedenen Auflösungen möglich. Die Reduktion erfolgt dabei durch Subsampling- oder Windowing-Techniken..

Format	Pixelanzahl	Speicherbedarf (8-Bit-ADC)
QCIF	176 x 144	24.75 kByte
QVGA	320 x 240	75.00 kByte
CIF	352 x 288	99.00 kByte
VGA	640 x 480	300.00 kByte

Tab. 3.5: Pixelzahl bei Standard-Bildsensoren und Speicherbedarf für einen kompletten Frame (8-Bit-ADC)

Viele Hersteller von CMOS-Imagern orientieren sich an gängigen CCD-Bildformaten. In Tab. 3.6 sind die resultierenden vertikalen Bildwinkel für verschiedene CCD-Größen in Abhängigkeit der Objektivbrennweite zusammengefasst. Tab. 3.7 enthält die horizontalen Bildwinkel.

CCD-Format Brennweite	1/6"	1/5"	1/4"	1/3"	1/2"	2/3"	1"
2.0 mm	43.6	50.82	61.93	83.97	100.39	117.56	134.76
2.5 mm	35.49	41.61	51.28	71.51	87.66	105.71	124.98
3.0 mm	29.86	35.14	43.6	61.93	77.32	95.45	115.99
3.5 mm	25.75	30.37	37.85	54.43	68.88	86.63	107.8
4.0 mm	22.62	26.72	33.4	48.46	61.93	79.05	100.39
4.5 mm	20.16	23.84	29.86	43.6	56.14	72.51	93.7
5.0 mm	18.18	21.52	26.99	39.6	51.28	66.85	87.66
5.5 mm	16.55	19.6	24.62	36.24	47.15	61.93	82.22
6.0 mm	15.19	17.99	22.62	33.4	43.6	57.62	77.32
6.5 mm	14.03	16.63	20.92	30.96	40.53	53.83	72.89
7.0 mm	13.04	15.46	19.46	28.84	37.85	50.48	68.88
7.5 mm	12.18	14.44	18.18	26.99	35.49	47.5	65.24
8.0 mm	11.42	13.54	17.06	25.36	33.4	44.83	61.93
8.5 mm	10.75	12.75	16.07	23.91	31.53	42.44	58.91
9.0 mm	10.16	12.05	15.19	22.62	29.86	40.27	56.14
9.5 mm	9.63	11.42	14.4	21.46	28.36	38.31	53.61
10.0 mm	9.15	10.85	13.69	20.41	26.99	36.53	51.28
10.5 mm	8.71	10.34	13.04	19.46	25.75	34.89	49.13
11.0 mm	8.32	9.87	12.45	18.59	24.62	33.4	47.15
11.5 mm	7.96	9.44	11.91	17.79	23.58	32.02	45.31
12.0 mm	7.63	9.05	11.42	17.06	22.62	30.75	43.6

Tab. 3.6: Vertikaler Bildwinkel in Abhängigkeit der Brennweite und der vertikalen Bildgröße bei Standard-CCD-Formaten

CCD-Format Brennweite	1/6"	1/5"	1/4"	1/3"	1/2"	2/3"	1"
2.0 mm	55.4	66.05	77.32	100.39	115.99	131.11	145.29
2.5 mm	45.56	54.95	65.24	87.66	104	120.79	137.33
3.0 mm	38.58	46.86	56.14	77.32	93.7	111.43	129.77
3.5 mm	33.4	40.75	49.13	68.88	84.87	103	122.65
4.0 mm	29.42	36.01	43.6	61.93	77.32	95.45	115.99
4.5 mm	26.27	32.23	39.15	56.14	70.83	88.71	109.78
5.0 mm	23.72	29.15	35.49	51.28	65.24	82.7	104
5.5 mm	21.62	26.6	32.44	47.15	60.38	77.32	98.65
6.0 mm	19.85	24.45	29.86	43.6	56.14	72.51	93.7
6.5 mm	18.35	22.62	27.66	40.53	52.42	68.19	89.11
7.0 mm	17.06	21.04	25.75	37.85	49.13	64.3	84.87
7.5 mm	15.94	19.67	24.09	35.49	46.21	60.8	80.95
8.0 mm	14.95	18.46	22.62	33.4	43.6	57.62	77.32
8.5 mm	14.08	17.39	21.32	31.53	41.26	54.74	73.96
9.0 mm	13.31	16.44	20.16	29.86	39.15	52.11	70.83
9.5 mm	12.61	15.58	19.12	28.36	37.23	49.7	67.94
10.0 mm	11.99	14.81	18.18	26.99	35.49	47.5	65.24
10.5 mm	11.42	14.12	17.33	25.75	33.9	45.47	62.73
11.0 mm	10.91	13.48	16.55	24.62	32.44	43.6	60.38
11.5 mm	10.43	12.9	15.84	23.58	31.1	41.87	58.19
12.0 mm	10	12.37	15.19	22.62	29.86	40.27	56.14

Tab. 3.7: Horizontaler Bildwinkel in Abhängigkeit der Brennweite und der horizontalen Bildgröße bei Standard-CCD-Formaten

Die berechneten Bildwinkel stimmen mit den Herstellerangaben für ein bestimmtes Objektiv nur näherungsweise überein, vermitteln aber einen ersten Eindruck von dem zu erwartenden Bildfeld. Ursache für diese Abweichungen sind die Abbildungseigenschaften realer Objektive, die nur näherungsweise durch die im Rahmen der geometrischen Optik getroffenen Vereinfachungen modelliert werden können.

Unter Beachtung typischer Erfassungsfelder von PIR-Präsenzwächtern wird im folgenden für den Bildwächter ein Öffnungswinkel von ca. 90° parallel zur längeren Seite des Bildsensors zugrundegelegt. Unter Beachtung der horizontalen und vertikalen Bildwinkel nach Tab. 3.6 und Tab. 3.7 erscheint es vorteilhaft, mindestens ein 1/3"-Format für den Bildaufnehmer in Verbindung mit einem möglichst kurzbrennweitigen Objektiv zu fordern. Verwendet man das in Tab. 3.4 angegebene S-Mount-Objektiv B2525 (Brennweite 2.5 mm), ergibt sich rechnerisch ein horizontaler Bildwinkel von 87.66° sowie ein vertikaler Bildwinkel von 71.51°.

1/3"-CMOS-Bildsensoren sind unter anderem mit den in Tab. 3.5 aufgeführten Auflösungen erhältlich. Unter Beachtung der zuvor für eine Brennweite von 2.5 mm spezifizierten Bildwinkel wurden die in Tab. 3.8 und Tab. 3.9 angegebenen Abstände der in die Objektebene projizierten Pixelmittelpunkte berechnet (vgl. Kap. 3.3.3). Als Parameter wurde der Abstand zwischen dem Bildwächter und der zur Bildebene parallelen Projektionsebene gewählt. Der Quotient der Seitenlänge nach Tab. 3.3 und der zugehörigen Pixelanzahl aus Tab. 3.5 diente dabei zur Ermittlung der Mittenabstände. Die Kombination eines 1/3"-Bildsensors mit einer

QVGA- oder VGA-Auflösung liefert einen identischen Pixelpitch $p = p_m = p_n$, für das QCIF- und das CIF-Format ergeben sich leicht unterschiedliche Mittenabstände p_m und p_n in horizontaler und vertikaler Richtung.

Abstand / [m]	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Auflösung										
QCIF	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
QVGA	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3
CIF	0.25	0.5	0.75	1	1.25	1.5	1.75	2	2.25	2.5
VGA	0.15	0.3	0.45	0.6	0.75	0.9	1.05	1.2	1.35	1.5

Tab. 3.8: Vertikale Auflösung / [cm] bei Verwendung eines 1/3"-Bildsensors mit QCIF-, QVGA-, CIF-, VGA-Auflösung in unterschiedlichen Entfernungen zum Bildwächter (Brennebene = Sensorebene || Objektebene, Brennweite 2.5 mm)

Abstand / [m]	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Auflösung										
QCIF	0.55	1.09	1.64	2.18	2.73	3.27	3.82	4.36	4.91	5.45
QVGA	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3
CIF	0.27	0.55	0.82	1.09	1.36	1.64	1.91	2.18	2.45	2.73
VGA	0.15	0.3	0.45	0.6	0.75	0.9	1.05	1.2	1.35	1.5

Tab. 3.9: Horizontale Auflösung / [cm] bei Verwendung eines 1/3"-Bildsensors mit QCIF-, QVGA-, CIF-, VGA-Auflösung in unterschiedlichen Entfernungen zum Bildwächter (Brennebene = Sensorebene || Objektebene, Brennweite 2.5 mm)

Ausgehend von Tab. 3.8 und Tab. 3.9 wurden auch die Flächen der in die Projektionsebene abgebildeten Pixel berechnet. Die Ergebnisse sind in Tab. 3.10 zusammengefasst.

Abstand / [m]	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Auflösung										
QCIF	0.28	1.09	2.46	4.36	6.83	9.81	13.37	17.44	22.1	27.25
QVGA	0.09	0.36	0.81	1.44	2.25	3.24	4.41	5.76	7.29	9
CIF	0.07	0.28	0.62	1.09	1.7	2.46	3.34	4.36	5.51	6.83
VGA	0.02	0.09	0.2	0.36	0.56	0.81	1.1	1.44	1.82	2.25

Tab. 3.10: Projizierte Pixelfläche / [cm²] bei Verwendung eines 1/3"-Bildsensors mit QCIF-, QVGA-, CIF-, VGA-Auflösung in unterschiedlichen Entfernungen zum Bildwächter (Sensorebene = Brennebene, Brennweite 2.5 mm)

CMOS-APS-Pixel besitzen im allgemeinen einen Füllfaktor von 20...30%. Es sollte daher beachtet werden, dass die lichtempfindliche Fläche A_{ph} meist kleiner ist als die Gesamtfläche $p_m p_n$ eines Pixels.

Abb. 3.23 zeigt typische Gesichtsfelder eines Bildwächters bei unterschiedlichen Kamerapositionen. Mit Hinblick auf eine zusätzliche Beleuchtungseinheit erscheint insbesondere die Deckenmontage vorteilhaft, da sich in diesem Fall zur optischen Achse symmetrische Gesichtsfelder und Beleuchtungsfelder ergeben. Eine konstante, maximale Bereichsauflösung ist nur in den zur Bildebene parallelen Projektionsebenen vorhanden. In den Ebenen, die nicht senkrecht zur optischen Achse sind, ist die Auflösung infolge der geometrischen Projektion abhängig vom Neigungswinkel gegen die optische Achse bzw. gegen die Bildebene.

Ein typischer Anwendungsfall eines Bildwächters ist die Erkennung von Personen, die sich in aufrechter Haltung im Erfassungsfeld bewegen. In diesem Fall sind relativ großflächige Bewegungen zu erwarten. Das andere Extrem ist die Erkennung von feinmotorischen Bewegungen, z. B. an Computerarbeitsplätzen. Insbesondere für die letztgenannte Anwendung erscheint die Deckenmontage als die geeignetere Montageform.

In Verbindung mit einer zusätzlichen LED-Beleuchtung ist es im Hinblick auf die Größe typischer Erfassungsfelder zur Raumüberwachung empfehlenswert, nicht den gesamten Raum, sondern nur den Außenbereich des Erfassungsfeldes ringförmig auszuleuchten.

Die reduzierte Größe des Beleuchtungsfeldes bewirkt einen geringeren Leistungsverbrauch des Bildwächters und wirkt sich daher vorteilhaft auf seine Betriebskosten. Weitere Einsparungen sind mit dem im nächsten Abschnitt beschriebenen passiven Sensorkonzept möglich.

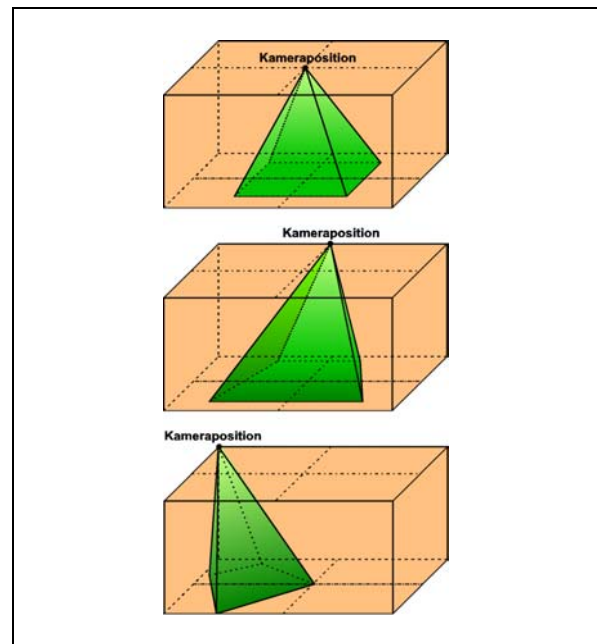


Abb. 3.23: Erfassungsfeld bei unterschiedlichen Kamerapositionen

3.3.5 Systemkonzept des Bildwächters

In diesem Abschnitt wird das Systemkonzept des Bildwächters als passive Komponente in einem Sensorsystem zur Anwesenheitserkennung vorgestellt. Um auch bei völliger Dunkelheit die Präsenzkontrolle durchzuführen, ist ein weiterer Sensor, z. B. ein Präsenz- oder Bewegungsmelder mit pyroelektrischen Detektoren, erforderlich, der die in vielen Anwendungen ohnehin vorhandene Raumbeleuchtung einschaltet.

Anschließend übernimmt der Bildwächter die hochauflösende Bewegungserkennung, wobei eine detektierte Änderung für das Nachtriggern der extern einstellbaren Ausschaltverzögerung sorgt. Der Bildwächter deaktiviert die Raumbeleuchtung erst wieder, wenn über den vorgegebenen Zeitraum keine weiteren Bewegungen erkannt wurden. Da das gerade beschriebene Multisensorsystem nur passive Sensorkomponenten verwendet, ergeben sich im Vergleich zu einem Bildwächter mit aktiver Beleuchtungseinheit deutliche Vorteile hinsichtlich der Leistungsaufnahme.

Zusätzlich ist es denkbar, dass der Bildwächter bei ausreichender Umgebungshelligkeit das Einschalten der Raumbeleuchtung ebenfalls beeinflusst. Der CMOS-Bildwächter ersetzt dabei den in konventionellen Bewegungsmeldern vorhandenen Helligkeitsmesser, der oftmals mit einem beleuchtungsabhängigen Widerstand (LDR) realisiert ist. Durch die hohe Auflösung

des CMOS-Sensors wird eine sehr differenzierte Beurteilung der Helligkeitsverhältnisse im Überwachungsbereich möglich.

Die logische UND-Verknüpfung der Alarmausgänge des Bildwächters und des Bewegungsmelders, welche im vorliegenden Fall zunächst als separate Geräte ausgeführt sind, erhöht außerdem die Fehlalarmimmunität. So lassen sich beispielsweise Helligkeitsänderungen durch externe Lichtquellen (vorbeifahrende Autos etc.) von Helligkeitsunterschieden infolge von Bewegungen innerhalb des vom Bewegungsmelder überwachten Bereichs trennen, indem man fordert, dass der pyroelektrische Bewegungsmelder zeitgleich einen Alarm auslösen muss. Erst wenn beide Bedingungen erfüllt sind, wird die Raumbelichtung eingeschaltet.

Werden beide Sensorprinzipien in einem einzigen Gerät zusammengefasst, sind weitere Verbesserungen der Fehlalarmimmunität zu erwarten. Die kombinierte Auswertung der Signalverläufe in unterschiedlichen Spektralkanälen (VIS oder IR-A und IR-C) ermöglicht die exakte zeitliche Zuordnung von Helligkeitsunterschieden und Temperaturänderungen im gemeinsam überwachten Bereich.

Durch die hohe Lateralauflösung des Bildwächters ist außerdem eine eindeutige örtliche Zuordnung der vom Bewegungsmelder registrierten Temperaturänderungen gegeben, weil die Überlagerung beider Erfassungsfelder bekannt ist. Außerdem können großflächigere Änderungen im Intensitätsbild des Bildwächters anhand des pyroelektrischen Erfassungsprinzips verifiziert werden.

Im Rahmen der vorliegenden Arbeit wurden verschiedene Auswerteverfahren zur Detektion von Grauwertänderungen in Videobildern untersucht. Die Ergebnisse dieser Untersuchungen, welche die Grundlage zur weiteren Entwicklung des Bildwächters bilden, sind in Kap. 4 zusammengefasst. Ausgangspunkt zur Beurteilung der einzelnen Verfahren sind typische Indoorszenen, die mit dem 1/3“-CMOS-Bildsensor VV5500 der Fa. ST Microelectronics aufgenommen wurden. Die technischen Daten des VGA-Chips VV5500 und einiger vergleichbarer CMOS-Sensoren sind in Tab. 3.11 zusammengefasst. Zur Übertragung und Speicherung der Bilddaten diente das ebenfalls von ST Microelectronics erhältliche Entwicklungskit zum VV5500. Die PC-seitig gespeicherten Bildsequenzen wurden anschließend offline mit dem Programmpaket MATLAB ausgewertet.

Aufgrund der zur Onlineanalyse der Bilddaten erforderlichen Rechenleistung erscheint eine hardwaregestützte Signalverarbeitung sinnvoll. Der hohe Integrationsgrad moderner FPGAs (Field Programmable Gate Array) und das attraktive Preis-Leistungs-Verhältnis dieser Bausteine ermöglicht die Realisierung eines Low-Cost-Bildwächters für Anwendungen im Bereich der Gebäudetechnik. Die Umsetzung des Auswerteverfahrens sowie die Implementierung auf der Zielhardware ist allerdings nicht mehr Gegenstand dieser Arbeit.

CMOS-Sensor	VV5500	LM9617	KAC0311
Hersteller	ST	NSC	Kodak/Motorola
Pixelanzahl (hor.)	640 352 320	648	640
Pixelanzahl (ver.)	480 288 240	488	480
Pixelgröße (hor. x ver.)	7.5 um x 7.5 um	7.5 um x 7.5 um	7.8 um x 7.8 um
Bildfläche (hor. x ver.)	4.89 mm x 3.66 mm	4.86 mm x 3.66 mm	5.00 mm x 3.70 mm
AD-Wandler	10 Bit	12 Bit	10 Bit
Dynamikbereich	81 dB	57 dB	49 dB
Kennlinie	Linear	Linear	Linear
Shutter	Rolling	Rolling	Rolling oder Full Frame
Gain	programmierbar	0...15 dB	-3 dB...17.4 dB
Bildrate	0.3...60 Hz	1...30 Hz	0...60 Hz
Auslesemodus	Prog. Scan oder Interlaced	Prog. Scan oder Interlaced	Prog. Scan oder Interlaced
Betriebsspannung	3.3...6 V	3.3 V	3.3 V
Leistungsverbrauch	139 mW	90 mW	215 mW
Power Down Mode	244 uW	5 mW	25 mW
Interface (Steuerung)	I ² C	I ² C	I ² C
Temperaturbereich	0...40 °C	0...50 °C	0...40°C

Tab. 3.11: Technische Daten von CMOS-Bildsensoren. Für den CMOS-Bildwächter wurde der VV5500 von ST Microelectronics ausgewählt.

3.4 Literatur

- [Lit. 3.1] R. Bruchhaus et. al.: "Thin Film Pyroelectric Detector Arrays with Surface Micromachined Pixels for Presence Detection", Proceedings IRS²2000
- [Lit. 3.2] Texas Instruments Deutschland GmbH: Das Opto-Kochbuch - Theorie und Praxis der Optoelektronik, 1. Auflage, Technik Marketing, München, 1975
- [Lit. 3.3] K. Herrmann, L. Walther: Wissensspeicher Infrarottechnik, 1. Auflage, Fachbuchverlag Leipzig, 1990
- [Lit. 3.4] E. Dereniak, D. Crowe: Optical Radiation Detectors, John Wiley & Sons, 1984
- [Lit. 3.5] P. Seitz: "Solid-State Image Sensing" in: B. Jähne et. al.: "Handbook of Computer Vision and Applications", Volume 1, Sensors and Imaging, Academic Press, 1999
- [Lit. 3.6] G. Meynants: "CMOS Image Sensors", Vortragsskripten, 5. FRAMOS CCD-Forum, 03.-04.05.2001
- [Lit. 3.7] N. Blanc, G. K. Lang: "CCD-Image Sensors versus CMOS-Image Sensors", Vortragsskripten zum 5. FRAMOS CCD-Forum, 03.-04.05.2001
- [Lit. 3.8] A. Theuwissen: "Solid State Imaging with Charge-Coupled Devices", Kluwer Academic Publishers, Dordrecht, Boston, London, 1995
- [Lit. 3.9] W. Ogiers: "Survey of CMOS Imagers", IMEC report P60280-MS-RP-002, 1997
- [Lit. 3.10] S. Kempainen: "CMOS image sensors: ECLIPSING CCDs in visual information", EDN Access 10.09.1997
- [Lit. 3.11] Firmenschrift, Jos. Schneider Optische Werke GmbH, Bad Kreuznach
- [Lit. 3.12] Produktkatalog BE BA 120/559, LINOS Photonics GmbH, Göttingen
- [Lit. 3.13] Produktkatalog Digitale Bildverarbeitung, DBS DIGITALE BILDVERARBEITUNG UND SYSTEMENTWICKLUNG GmbH, Bremen, 1999
- [Lit. 3.14] Schweizerische Lichttechnische Gesellschaft et. al.: "Handbuch für Beleuchtung", Verlag W. Girardet, Essen, 1975

4 Auswerteverfahren für bildgebende Sensoren

4.1 Einleitung

Beim Einsatz konventioneller, bildgebender Sensoren geht die Tiefeninformation der abgebildeten 3D-Szene verloren. Sie kann nur mit zusätzlichem technischen Aufwand rekonstruiert werden. Infolge der optischen Abbildung ergibt sich eine zweidimensionale Intensitätsverteilung in der Bildebene. Durch die Rasterung der Bildebene und die anschließende Quantisierung wird die Intensitätsverteilung in ein digitales Bild überführt.

Erfolgt die Bildaufnahme zu diskreten Zeitpunkten, entsteht eine zeitliche Bildfolge, die neben der Bewertung von statischen Eigenschaften auch die Beurteilung von dynamischen Vorgängen innerhalb des Beobachtungszeitraums zulässt. Grundlegende Betrachtungen zur Erfassung, Auswertung und Interpretation von bewegungsbedingten Veränderungen in Bildfolgen sind in [Lit. 4.1] und [Lit. 4.3] zusammengefasst.

4.2 Differenzbildverfahren zur Bewegungserkennung

Die einfachste und wahrscheinlich am weitesten verbreitete Methode zur Detektion von Grauwertänderungen basiert auf der Differenzbildung zeitlich aufeinanderfolgender Videobilder. Das so gewonnene Differenzbild vergleicht man pixelweise mit einem unterem und oberen Schwellwert, wobei Differenzen innerhalb des vorgegebenen Toleranzbandes als nicht signifikant zurückgewiesen werden (Hypothese H_0). Differenzwerte unterhalb bzw. oberhalb der zunächst fixen Schwellen werden als Indikator für eine Bewegung im Erfassungsfeld des Wächters interpretiert (Hypothese H_1).

Abb. 4.1 und Abb. 4.2 zeigen die Anwendung des gerade beschriebenen Verfahrens zur Erkennung von grob- und feinmotorischen Bewegungen. Selbst sehr feine Bewegungen, wie beispielsweise das Bewegen einer Maus oder die Bedienung der Tastatur an Bildschirmarbeitsplätzen, sind infolge der hohen Auflösung des CMOS-Bildsensors identifizierbar. Anhand von Abb. 4.1 ist außerdem zu erkennen, dass Grauwertveränderungen an der aktuellen und der vorherigen Position der Person sowie auch im Schattenbild auftreten. Voraussetzung für eine sichere Detektion ist allerdings ein ausreichend hoher Kontrast zwischen dem Hintergrund und der Person im Vordergrund der beobachteten Szene.

Entsprechend starke Grauwertänderungen können allerdings auch durch eine Veränderung der Beleuchtungsverhältnisse im Überwachungsbereich des Bildwächters hervorgerufen werden. Änderungen der natürlichen Lichtverhältnisse (Sonnenlicht) wirken sich wegen der im Vergleich zu typischen Bildraten langsamen Änderungsgeschwindigkeit eher global aus. Bei CMOS-Sensoren, die einen Rolling-Shutter-Mechanismus zur Steuerung ihrer Belichtungszeit verwenden, können zeilenweise durchaus abweichende Helligkeitsdifferenzen entstehen, wäh-

rend sich globale Beleuchtungsänderungen bei einem Sensor mit Full-Frame-Shutter auf alle Pixel gleich auswirken.

Im Gegensatz zum Sonnenlicht kann das Ein- oder Ausschalten von Kunstlichtquellen auch einen lokal begrenzten Einfluss auf die in den Pixeln akkumulierten Ladungsträger haben. Außerdem ist in diesem Zusammenhang zu beachten, dass insbesondere bei Leuchtstofflampen mit konventionellen Vorschaltgeräten Lichtwelligkeiten auftreten, die sich je nach Bildrate als periodische Helligkeitsänderungen im Grauwertbild bemerkbar machen. Bei Kunstlichtquellen mit elektronischen Vorschaltgeräten, die im Frequenzbereich von 40...100 kHz arbeiten, sind diese Störungen dagegen von untergeordneter Bedeutung.

Speziell in Verbindung mit konventionellen Kunstlichtquellen ist es daher empfehlenswert, eine zur Netzfrequenz synchrone Integrationsdauer zu wählen. Die Unterdrückung von natürlichen Helligkeitsschwankungen erfordert zusätzlich eine möglichst hohe Bildwiederholrate. Im Hinblick auf die vorliegende Aufgabenstellung muss allerdings beachtet werden, dass durch Bewegungen hervorgerufene Grauwertänderungen nicht beliebig schnell erfolgen. Je höher die Bildwiederholrate ist, umso geringer sind die in aufeinanderfolgenden Bildern detektierbaren Veränderungen.

Bei höheren Bildraten ist es unter Umständen erforderlich, Veränderungen über mehrere Bildern zu akkumulieren. Eine entsprechende Vorgehensweise wird beispielsweise in [Lit. 4.4] beschrieben. Mit Hinblick auf den dafür erforderlichen System- und Signalverarbeitungsaufwand wurde für die weiteren Untersuchungen zum Bildwächter eine Bildwiederholrate von 10 Hz eingestellt, wodurch ein ausreichender Spielraum zur Bewegungsdetektion in unmittelbar aufeinanderfolgenden Videobildern erhalten bleibt. Die mit dem ST5500-Evaluation-Kit maximal erzielbare Bildfrequenz liegt bei 15 Hz.

Während durch Objektbewegungen verursachte Grauwertänderungen normalerweise zu einer örtlichen Clusterung führen, treten Störungen infolge von Rauschen eher isoliert und über das gesamte Bild verteilt auf. Zur Verdeutlichung zeigt Abb. 4.3 typische Rauschstörungen für die Szene nach Abb. 4.1, wobei die Auswerteschwellen bewusst niedrig gewählt wurden.

Isolierte Rauschstörungen, die auf das Rauschen des Bildsensors und der übrigen Systemkomponenten zurückzuführen sind, lassen sich durch einfache Nachbarschaftsoperationen unterdrücken. In [Lit. 4.3] wird zunächst diese Vorgehensweise zur Diskriminierung von isolierten und örtlich geballten Grauwertänderungen vorgeschlagen. Änderungsbereiche zu geringer Größe werden dabei auf das Kamerarauschen zurückgeführt und somit von der Bewegungsdetektion ausgeschlossen. Zusätzlich zu den Schwellwerten zur Differenzbilddauswertung ist daher die Angabe einer weiteren Schwelle erforderlich, welche die minimale Bereichsgröße zur Bewegungserkennung definiert.

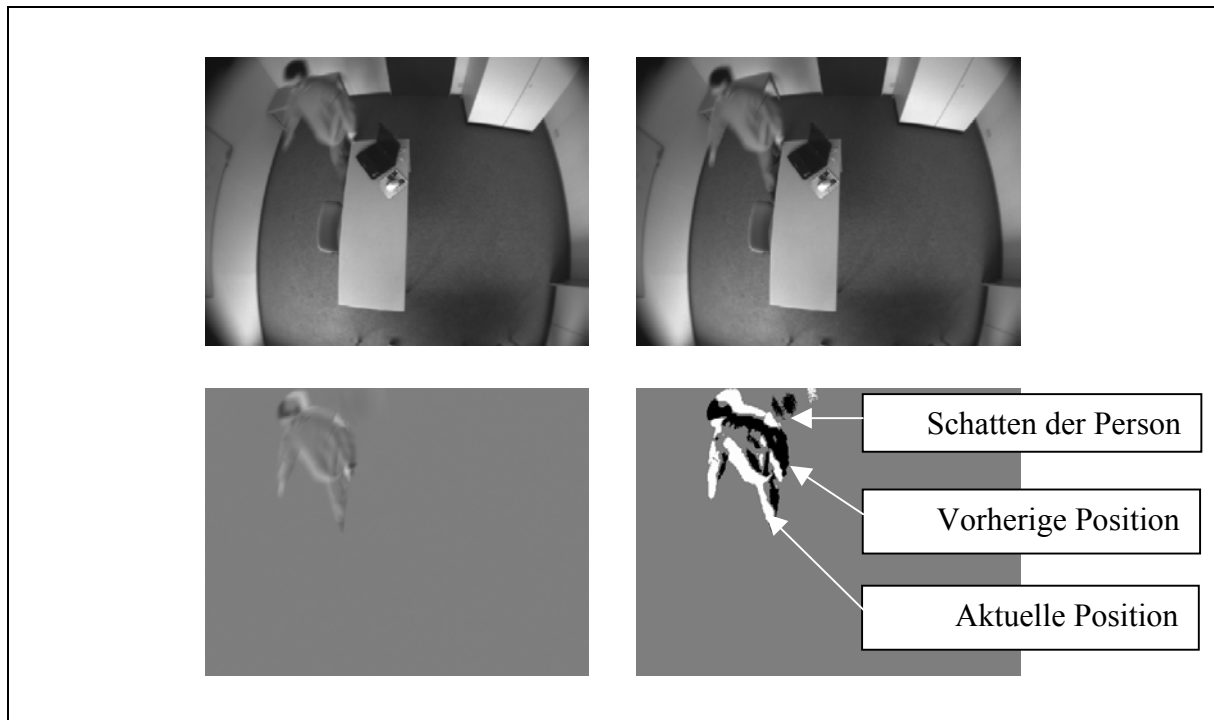


Abb. 4.1: Anwendung des Differenzbildverfahrens mit anschließendem Schwellwertvergleich zur Detektion von grobmotorischen Bewegungen. Originalbilder (oben) sowie Differenzbild (unten links) und mit Schwellwerten von ± 10 bewertetes Differenzbild (unten rechts). Geschwärtzte Bereiche zeigen eine Verdunkelung an, während weiß markierte Bereiche eine Aufhellung identifizieren.

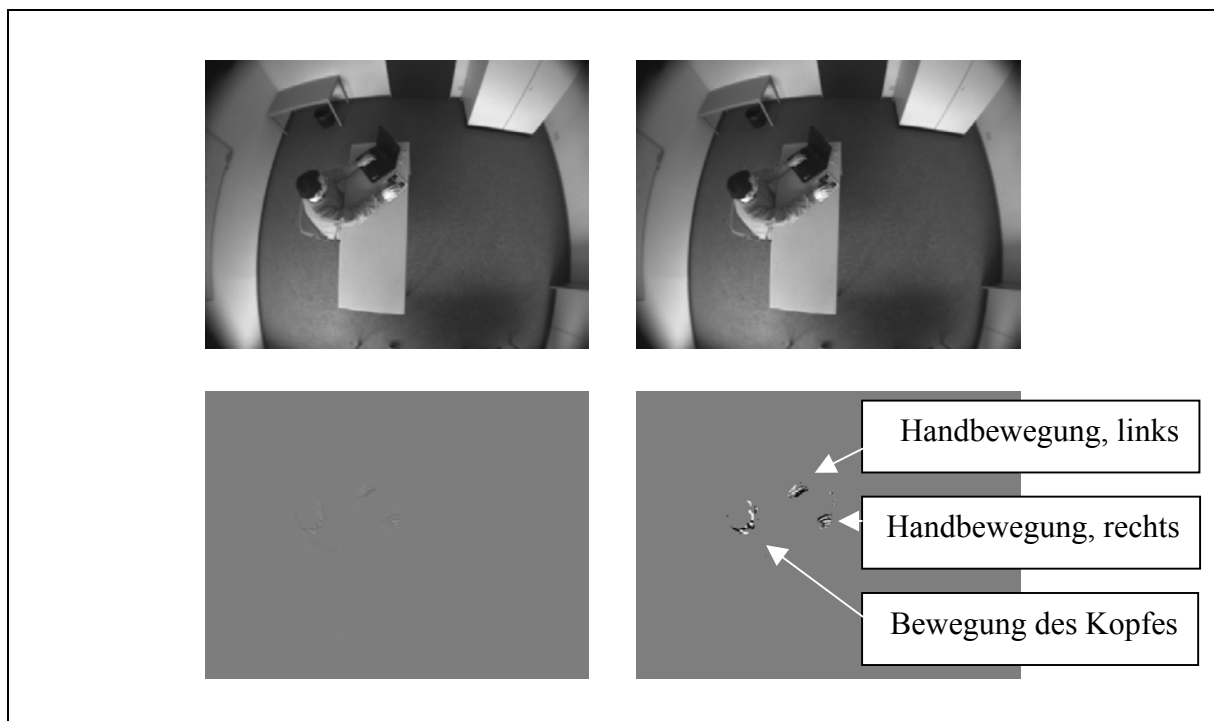


Abb. 4.2: Anwendung des Differenzbildverfahrens mit anschließendem Schwellwertvergleich zur Detektion von feinmotorischen Bewegungen. Originalbilder (oben) sowie Differenzbild (unten links) und mit Schwellwerten von ± 10 bewertetes Differenzbild (unten rechts). Schwarze Bereiche zeigen eine Verdunkelung an, während weiß markierte Bereiche eine Aufhellung identifizieren.

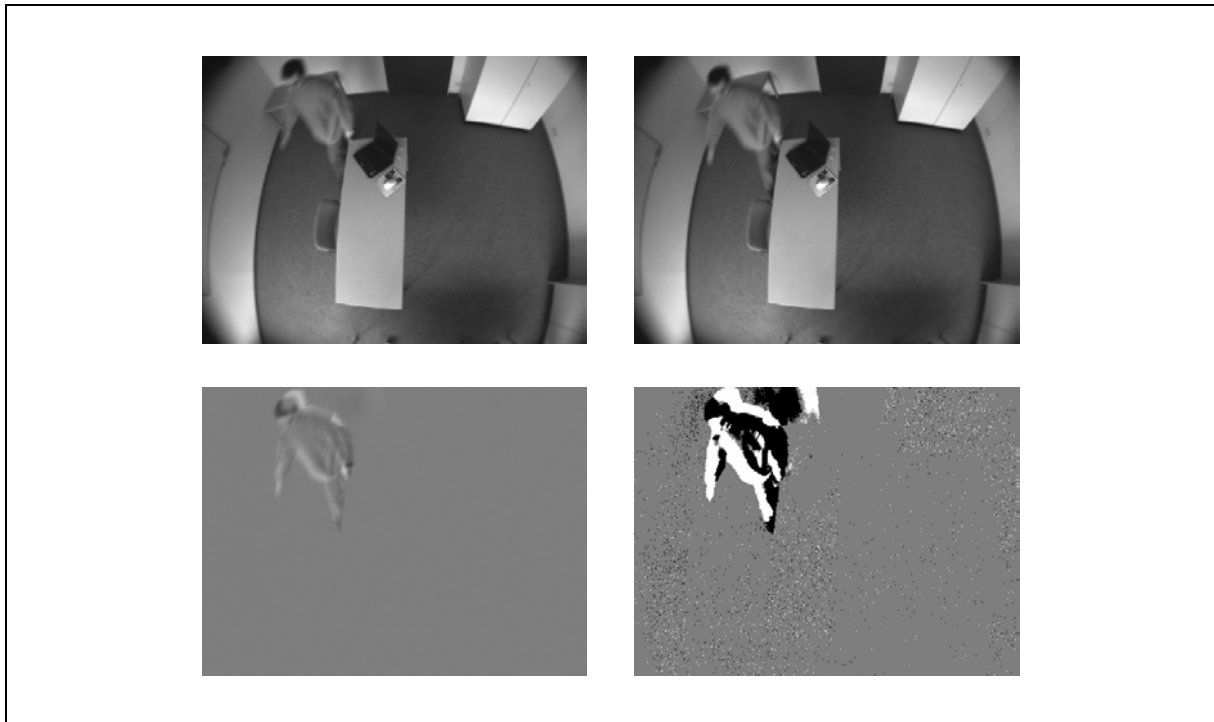


Abb. 4.3: Anwendung des Differenzbildverfahrens mit anschließendem Schwellwertvergleich. Originalbilder (oben) sowie Differenzbild (unten links) und mit Schwellwerten von ± 3 bewertetes Differenzbild (unten rechts). Deutlich zu erkennen sind Störungen durch Rauschen infolge der zu klein gewählten Auswerteschwellen.

Prinzipiell ist festzustellen, dass die Festlegung fixer Schwellwerte für allgemeine Anwendungen kritisch ist. Verkleinert man das Toleranzband zur Änderungsdetektion, steigt die Fehlalarmanzahl in Folge von Rauschstörungen kontinuierlich an. Werden dagegen die Schwellwerte und somit das Toleranzband vergrößert, reagiert der Sensor zunehmend unempfindlicher auf bewegungsbedingte Veränderungen. Außerdem wird die hohe Ortsauflösung des Bildsensors durch die minimal erforderliche Bereichsgröße herabgesetzt, wodurch sich das Detektionsverhalten für die feinmotorische Bewegungen verschlechtert. Insbesondere die Erkennung feinmotorischer Bewegungsabläufe ist zur Anwesenheitskontrolle nach dem vorne beschriebenen Prinzip von erhöhter Bedeutung.

Anstelle der Vorgabe von festen Schwellwerten ist es ebenfalls denkbar, für die Schwellen zur Auswertung des Intensitätsbildes eine adaptive Anpassung innerhalb definierter Grenzen vorzusehen. Während der Hypothese H_0 bietet sich beispielsweise eine fortlaufende Aktualisierung der Schwellwerte zur Rauschunterdrückung an. Der Einfluss der Lichtwelligkeit von zugeschaltetem Kunstlicht kann während der Einrichtphase des Bildwächters zusätzlich ermittelt und bei der Schwellwertfestlegung sinngemäß berücksichtigt werden.

Um die Robustheit des vorher beschriebenen Differenzbildverfahrens mit Schwellwert- und Bereichsgrößenvorgabe zu verbessern, wird in [Lit. 4.3] eine Vertauschung der bisherigen Bearbeitungsreihenfolge "Differenzbildauswertung – Clusterung" vorgeschlagen. Während die in [Lit. 4.3] bzw. [Lit. 4.4] beschriebenen Signalverarbeitungskonzepte in den folgenden

Abschnitten näher erläutert werden, soll an dieser Stelle zunächst die Mittelwertbildung innerhalb eines vordefinierten Teilbildes mit anschließendem Schwellwertvergleich betrachtet werden. Abb. 4.4 und Abb. 4.5 verdeutlichen diese Vorgehensweise für Fenstergrößen $M \times N$ von 4×4 und 8×8 Pixel für die Szene nach Abb. 4.2, wobei M die Zeilenanzahl und N die Spaltenanzahl innerhalb des Teilbildes repräsentiert.

Unter der Annahme, dass die Grauwerte innerhalb des betrachteten Fensters durch mittelwertfreies, Gaußverteiltes Rauschen bekannter Varianz gestört sind, folgt aus der Mittelwertbildung eine um den Faktor $(M \cdot N)^{-1/2}$ kleinere Varianz der Bereichsmittelwerte. Das Toleranzband zur Differenzbilddauswertung kann wegen der geringeren statistischen Abweichung verkleinert werden, wenn die zuvor getroffene Voraussetzung erfüllt ist.

Experimentell konnte bestätigt werden, dass die Schwellwerte bei einer Bildwiederholrate von 10 Hz entsprechend verringert werden können. Für Bildfrequenzen unterhalb von 5 Hz änderten sich die natürlichen Beleuchtungsverhältnisse bereits so stark, dass sich einzelne Mittelwerte in korrespondierenden Bildbereichen deutlich unterschieden. Eine entsprechende Schwellwertreduktion war in diesem Fall nicht möglich.

Vorteilhaft anwenden lässt sich die bereichsweise Mittelwertbildung zur Detektion von Veränderungen, die durch grobmotorische Bewegungen hervorgerufen werden. Betritt eine Person das Erfassungsfeld des Bildwächters, sind zunächst großflächigere Bewegungen zu erwarten, die außerdem das Ansprechen des integrierten pyroelektrischen Bewegungsmelders bewirken. Anschließend kann die Auflösung ggf. stufenweise bis zur maximalen Auflösung des Bildsensors erhöht werden. Auf diese Weise lassen sich im Anschluss an die Aktivierung des Bildwächters selbst feinmotorische Bewegungen zu erkennen, die auf das Verweilen der Person im Überwachungsbereich schließen lassen und damit eine erneute Initialisierung der Ausschaltverzögerung des Präsenzwächters bewirken.

Betrachtet man Abb. 4.4 und Abb. 4.5, erkennt man, dass bei einer Fenstergröße von 4×4 Pixel feinmotorische Bewegungen der Hände noch erkannt werden, während bei einer Fenstergröße von 8×8 Pixel nur noch die "großflächigere" Bewegung des Kopfes auftritt. Je größer der Teilbildbereich zu Auswertung gewählt wird, umso unempfindlicher reagiert der Bildwächter auf Veränderungen, selbst wenn die vorher beschriebene Herabsetzung der Auswerteschwellen durchgeführt wird.

Allgemein lässt sich feststellen, dass infolge der zur Mittelung verwendeten Bereichsgrößen entsprechend starke Grauwertänderungen innerhalb des Fensters erforderlich sind, damit eine bewegungsbedingte Veränderung erkannt werden kann. Prinzipiell müssen dabei zwei Extremfälle unterschieden werden. Valide Detektionsereignisse treten einerseits bei ausreichend deutlichen Grauwertänderungen vereinzelter Pixel auf und andererseits bei moderaten Intensitätsänderungen innerhalb des gesamten, zur Bewegungserkennung verwendeten Teilbildes.

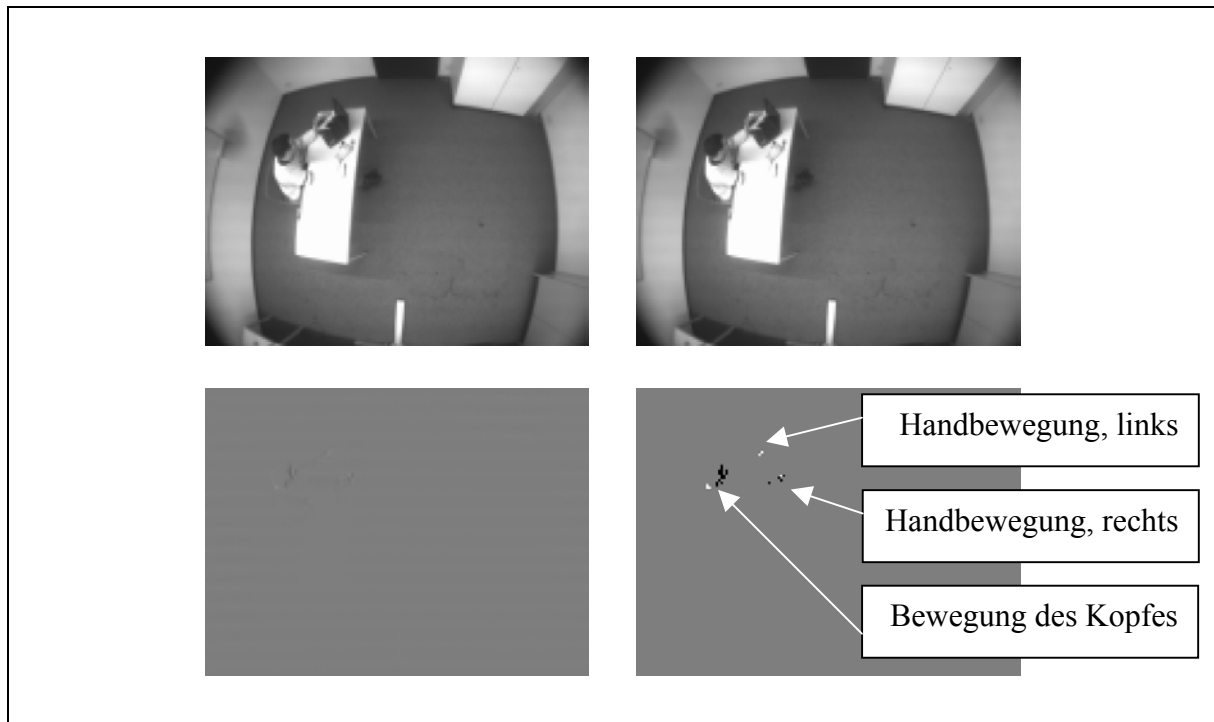


Abb. 4.4: Anwendung des Differenzbildverfahrens für eine Fenstergröße von 4 x 4 Pixel mit anschließendem Schwellwertvergleich zur Detektion feinmotorischer Bewegungen. Originalbilder (oben) sowie Differenzbild (unten links) und mit Schwellwerten von ± 2.5 bewertetes Differenzbild (unten rechts). Schwarze Bereiche zeigen eine Verdunkelung an, weiß markierte Bereiche identifizieren eine Aufhellung.

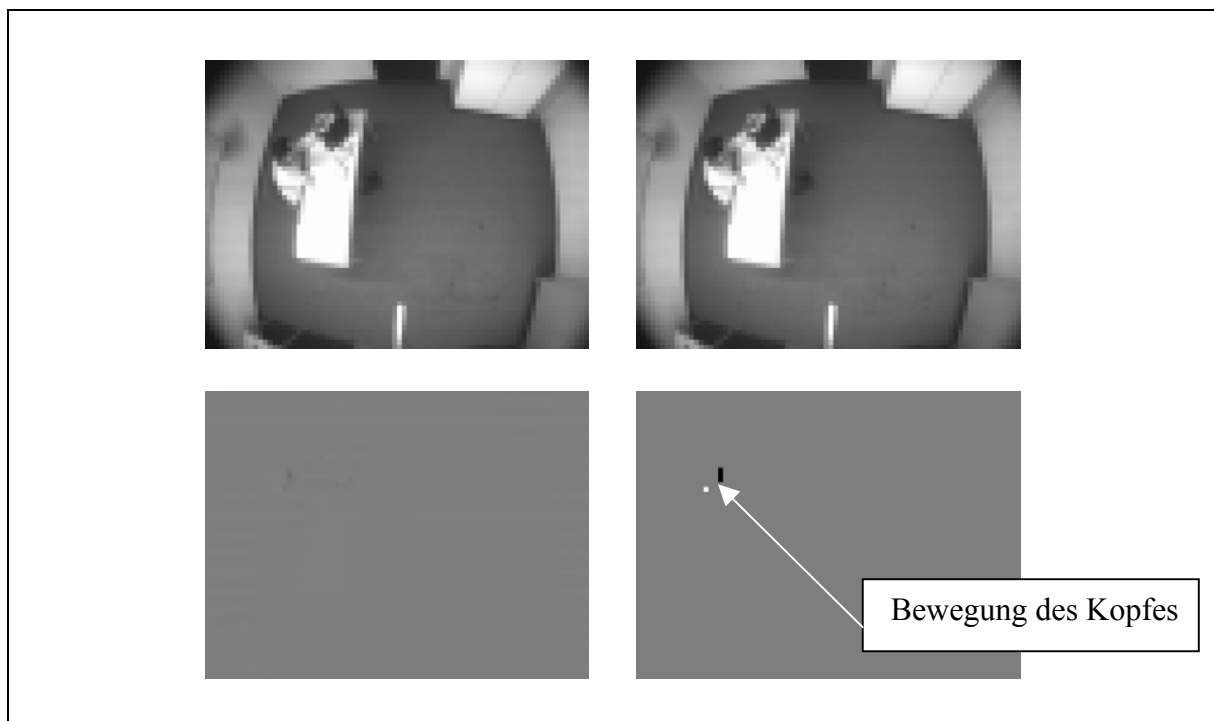


Abb. 4.5: Anwendung des Differenzbildverfahrens für eine Fenstergröße von 8 x 8 Pixel mit anschließendem Schwellwertvergleich zur Detektion feinmotorischer Bewegungen. Originalbilder (oben) sowie Differenzbild (unten links) und mit Schwellwerten von ± 1.25 bewertetes Differenzbild (unten rechts). Schwarze Bereiche zeigen eine Verdunkelung an, weiß markierte Bereiche identifizieren eine Aufhellung.

Es ist somit ein ausreichender Kontrast zwischen der Person im Vordergrund und der statischen Szene im Hintergrund erforderlich. Der tatsächlich in der Szene vorhandene Kontrast wird durch die Mittelwertbildung in den betrachteten Teilbildern reduziert. Im Gegensatz zu globalen Helligkeitsänderungen, die meist großflächige Aufhellungen oder Verdunkelungen bewirken, sind lokale, bewegungsbedingte Helligkeitsänderungen oftmals dadurch charakterisiert, dass hellere und dunklere Bereiche gleichzeitig im Differenzbild auftreten. Als zusätzliches Kriterium zur Bewegungsdetektion kann daher die prozentuale Verteilung von Aufhellungen und Verdunkelungen, bezogen auf die Anzahl der insgesamt veränderten Grauwerte, verwendet werden.

Aufgrund der sequentiellen Übertragung einzelner Pixel ist die Mittelung über quadratische Teilbildbereiche mit einem erhöhten Speicheraufwand während der Berechnung verbunden. Zur Speicherung des Ergebnisses für nachfolgende Differenzbildauswertungen sind bei einem VGA-Chip 300 kB erforderlich (s. Tab. 3.5). Bei der Fenstergröße $M \times N$ für die Mittelwertbildung reduziert sich der Speicherbedarf genau um diesen Faktor, so dass beispielsweise bei einer Fenstergröße von 2×2 Pixel insgesamt 75 kB und bei einer Fenstergröße von 4×4 Pixel insgesamt 18.75 kB zur Zwischenspeicherung der Bereichsmittelwerte erforderlich sind.

4.3 Bewegungserkennung mit Zeilen- und Spaltenmittelwerten

Einige CMOS-Sensoren bieten die Möglichkeit, anstelle des Intensitätsbildes nur die Zeilen- und Spaltenmittelwerte auszugeben. Da diese Eigenschaft in Hardware implementiert ist, erfolgt bereits im Chip eine hohe Datenreduktion. Entsprechend gering sind die Systemanforderungen an die Datenübertragung und die weitere Bildverarbeitung. Unabhängig davon, ob die Mittelwertbildung auf dem CMOS-Chip oder von einer nachfolgenden Signalverarbeitungsstufe durchgeführt wird, treten allerdings auch hier die im Zusammenhang mit der Mittelung über quadratische Bildbereiche beschriebenen Effekte auf. Da der ST5500 keine Zeilen- und Spaltenmittelwerte zur Verfügung stellt, wurde dieses Verfahren in Software nachgebildet. Abb. 4.6 zeigt die Auswertung der Zeilen- und Spaltenmittelwerte der Szene nach Abb. 4.1.

Zusätzlich zu den Bewertungsergebnissen des Differenzbildverfahrens nach Abb. 4.1 wurden in Abb. 4.6 jeweils die Zeilen- und Spaltenmittelwerte der beteiligten Originalbilder (oben), die um den Faktor 16 verstärkten Grauwertdifferenzen beider Vektoren (unten links) sowie die in Zeilen- und Spaltenrichtung mit unterschiedlichen Schwellwerten bewerteten Grauwertdifferenzen (unten rechts) dargestellt. Betrachtet man die jeweils anhand einer oberen und unteren Schwelle diskretisierten Grauwertdifferenzen, fällt auf, dass sich mitunter Aufhellungen und Verdunkelungen innerhalb einer Zeile bzw. Spalte kompensieren. Im Vergleich zur Mittelwertbildung über quadratische Teilbilder, wobei letztere deutlich weniger Pixel als eine komplette Zeile bzw. Spalte enthielten, ergibt sich eine geringere Empfindlichkeit für feinmo-

torische Bewegungen. Für die in Abb. 4.2 dargestellte Szene liefert das Differenzbildverfahren, angewendet auf die Zeilen- und Spaltenmittelwerte, keine Detektion.

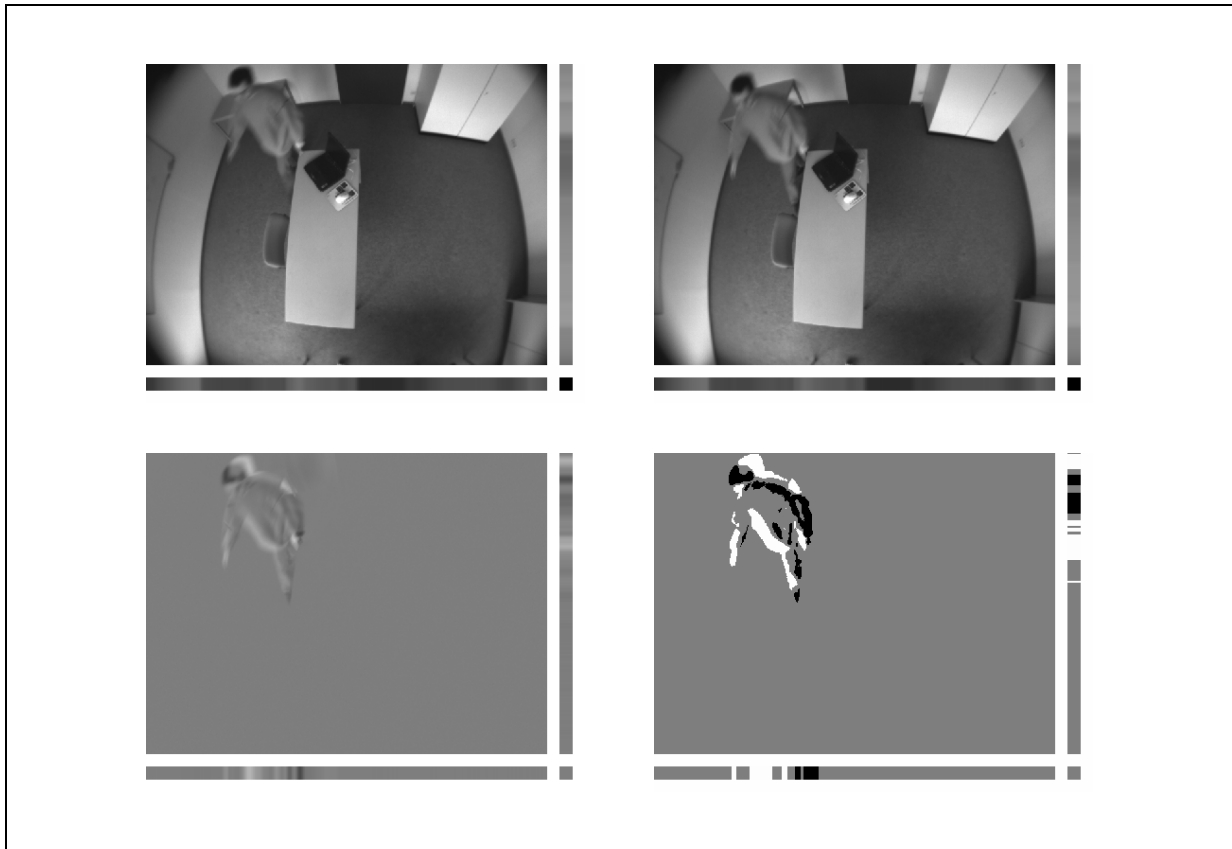


Abb. 4.6: Anwendung des Differenzbildverfahrens auf Zeilen- und Spaltenmittelwerte mit anschließendem Schwellwertvergleich zur Detektion grobmotorischer Bewegungen. Originalbilder (oben) sowie Differenzbild (unten links) und mit Schwellwerten bewertetes Differenzbild (unten rechts). Schwarze Bereiche zeigen eine Verdunkelung an, weiß markierte Bereiche identifizieren eine Aufhellung.

Zur Vermeidung der zuvor beschriebenen Kompensationseffekte bei der Mittelwertbildung, die auch bei quadratischen oder rechteckigen Teilbildern auftritt, bietet es sich an, die absolute Abweichung des tatsächlichen Grauwertverlaufs vom Bereichsmittelwert oder auch die quadratische Abweichung (Varianz) korrespondierender Bereiche in zeitlich aufeinander folgenden Bildern auszuwerten. Dieser Ansatz führt auf die in den nächsten Abschnitten vorgestellten Verfahren, die in [Lit. 4.3] und [Lit. 4.4] ausführlicher beschrieben sind. Prinzipiell interpolieren diese Verfahren den 2D-Grauwertverlauf in vordefinierten Bildausschnitten mit unterschiedlicher Güte (konstante, lineare und quadratische Interpolation).

Anschließend werden die Varianzen der Teilbilder ermittelt und als Grundlage zur Erkennung von bewegungsbedingten Grauwertänderungen verwendet. Mit Hinblick auf Anwendungen im Bereich der Gebäudetechnik soll die Empfindlichkeit dieser Verfahren im Vergleich zu dem vorne beschriebenen Differenzbildverfahren ermittelt werden.

4.4 Varianzauswertung bei konstanter Grauwertinterpolation

Ausgangspunkt für die im weiteren Verlauf vorgestellten Verfahren zur Änderungsdetektion ist die bereichsweise Modellierung des tatsächlichen Grauwertbildes \underline{G} durch analytische Funktionen $f(x_w, y_w)$, wobei x_w und y_w die diskreten Ortskoordinaten innerhalb des betrachteten Teilbildes \underline{G}_w darstellen.

Die nachfolgenden Berechnungen beziehen sich jeweils auf eine Teilmatrix \underline{G}_w der Größe $I = P \times Q$ und müssen daher für alle Blöcke in der Grauwertmatrix \underline{G} wiederholt werden. Einzelne Pixel $g(p,q)$ innerhalb des Fensters G_w werden mit $p = 0 \dots P-1$ in Zeilenrichtung und $q = 0 \dots Q-1$ in Spaltenrichtung adressiert.

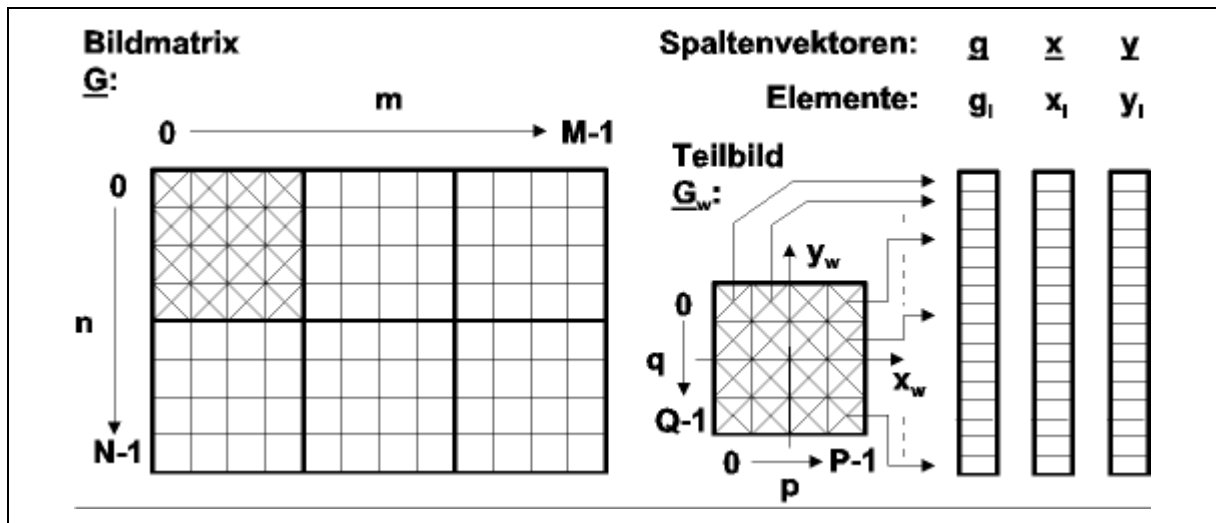


Abb. 4.7: Zerlegung der Bildmatrix \underline{G} in Teilmatrizen \underline{G}_w und Umwandlung der Grauwerte im Fenster \underline{G}_w in den Spaltenvektor \underline{g} . Die Elemente g_i des Spaltenvektors \underline{g} enthalten die einzelnen Grauwerte, beginnend mit der oberen, linken Ecke. Die Spaltenvektoren \underline{x} und \underline{y} mit den Elementen x_i und y_i beinhalten die zugehörigen Ortskoordinaten mit Bezug zur Fenstermitte.

Zur Vereinfachung der weiteren Betrachtungen werden die Grauwerte der Teilmatrix \underline{G}_w und die zugehörigen Koordinaten x_w und y_w nach Abb. 4.7 umsortiert. Daraus ergeben sich die Spaltenvektoren \underline{g} , \underline{x} , und \underline{y} . Korrespondierende Elemente g_i , x_i und y_i der Spaltenvektoren \underline{g} , \underline{x} , und \underline{y} werden im folgenden mit dem ganzzahligen Laufindex i mit $i \in [1, I]$ indiziert. Für die Summe aller Grauwerte $g(p,q)$ innerhalb des Fensters \underline{G}_w gilt somit

$$(Gl. 4.1) \quad \sum_{q=0}^{Q-1} \sum_{p=0}^{P-1} g(p,q) = \sum_{i=1}^I g_i$$

Die Interpolation des Grauwertverlaufs $\underline{G}_w(x_w, y_w)$ innerhalb des betrachteten Teilbildes soll zunächst mit der Konstante μ erfolgen. Betrachtet man die quadratische Abweichung e^2 des tatsächlichen und des interpolierten Grauwertverlaufs im Fenster \underline{G}_w , folgt

$$(Gl. 4.2) \quad e^2 = \sum_{i=1}^I e_i^2 = \sum_{i=1}^I (g_i - \mu)^2 .$$

woraus sich nach der Least-Square-Methode der Parameter μ bestimmen lässt. Differenziert man (Gl. 4.2) nach μ und setzt das Ergebnis gleich Null, erhält man aus

$$(Gl. 4.3) \quad \frac{\partial e^2}{\partial \mu} = 0 = -2 \cdot \sum_{i=1}^I (g_i - \mu) = -2 \cdot \left[\sum_{i=1}^I g_i - \mu \sum_{i=1}^I 1 \right] = -2 \cdot \left[\sum_{i=1}^I g_i - \mu \cdot I \right]$$

die Beziehung (Gl. 4.4). Erwartungsgemäß ist μ der mittlere Grauwert im Fenster \underline{G}_w . Die mittlere quadratische Abweichung e^2/I nach (Gl. 4.5) strebt für große I gegen die Varianz σ^2 , so dass im weiteren Verlauf $\sigma^2 = e^2/I$ verwendet wird.

$$(Gl. 4.4) \quad \mu = \frac{1}{I} \cdot \sum_{i=1}^I g_i = \bar{g}$$

$$(Gl. 4.5) \quad \frac{e^2}{I} = \frac{1}{I} \cdot \sum_{i=1}^I (g_i - \mu)^2 \xrightarrow{I \rightarrow \infty} \sigma^2 = \frac{1}{I-1} \cdot \sum_{i=1}^I (g_i - \mu)^2$$

Zur Änderungsdetektion in aufeinanderfolgenden Bildern werden die Mittelwerte μ_1 und μ_2 und die Varianzen $(\sigma_1)^2$ und $(\sigma_2)^2$ in korrespondierenden Teilbildern \underline{G}_{w1} und \underline{G}_{w2} betrachtet. Ausgehend von (Gl. 4.4) und (Gl. 4.5) werden diese Parameter nach (Gl. 4.6) – (Gl. 4.9) bestimmt. Die Indizes “1“ und “2“ bezeichnen die Zugehörigkeit zum jeweiligen Teilbild.

$$(Gl. 4.6) \quad \mu_1 = \frac{1}{I} \cdot \sum_{i=1}^I g_{i1} \quad (Gl. 4.7) \quad \sigma_1^2 = \frac{1}{I} \cdot \sum_{i=1}^I (g_{i1} - \mu_1)^2$$

$$(Gl. 4.8) \quad \mu_2 = \frac{1}{I} \cdot \sum_{i=1}^I g_{i2} \quad (Gl. 4.9) \quad \sigma_2^2 = \frac{1}{I} \cdot \sum_{i=1}^I (g_{i2} - \mu_2)^2$$

Anhand von (Gl. 4.6) - (Gl. 3.67) soll unterschieden werden, ob die Grauwertmengen \underline{G}_{w1} und \underline{G}_{w2} zu verschiedenen Grauwertmengen oder zu einer gemeinsamen Grauwertmenge \underline{G}_{w0} gehören, wobei die letztgenannte durch den Mittelwert μ_0 und die Varianz $(\sigma_0)^2$ charakterisiert wird. Für die Parameter μ_0 und $(\sigma_0)^2$ erhält man nach einigen Umformungen:

$$(Gl. 4.10) \quad \mu_0 = \frac{1}{2I} \cdot \left(\sum_{i=1}^I g_{i1} + \sum_{i=1}^I g_{i2} \right) = \frac{1}{2} \cdot (\mu_1 + \mu_2) \quad \text{und}$$

$$(Gl. 4.11) \quad \sigma_0^2 = \frac{1}{2I} \cdot \left(\sum_{i=1}^I (g_{i1} - \mu_0)^2 + \sum_{i=1}^I (g_{i2} - \mu_0)^2 \right) = \frac{\sigma_1^2 + \sigma_2^2}{2} + \left(\frac{\mu_1 - \mu_2}{2} \right)^2$$

Unter der Annahme, dass die Grauwerte g_i im ausgewählten Fenster \underline{G}_w homogen und nur von mittelwertfreiem Gaußverteilterm Rauschen bekannter Varianz gestört sind, sowie dass die beteiligten Rauschprozesse statistisch unabhängig sind, ergibt sich ganz allgemein die Verbundverteilungsfunktion $L(g_i, \mu, \sigma^2)$ nach (Gl. 4.12) als Verteilungsdichtefunktion [Lit. 4.5].

$$(Gl. 4.12) \quad L(g_i, \mu, \sigma) = \left[\frac{1}{2\pi \cdot \sigma^2} \right]^{I/2} \cdot \exp \left\{ -\frac{1}{2\sigma^2} \cdot \sum_{i=1}^I (g_i - \mu)^2 \right\}$$

$L(g_i, \mu, \sigma^2)$ wird auch als Likelihood-Funktion bezeichnet. Die Parameter μ und σ^2 können durch Differentiation von $L(g_i, \mu, \sigma^2)$ oder $\ln(L(g_i, \mu, \sigma^2))$ nach μ bzw. σ^2 und Null setzen der

resultierenden Gleichung berechnet werden. Diese Vorgehensweise wird als Maximum-Likelihood-Methode bezeichnet.

Das Maximum-Likelihood-Verfahren liefert dieselben Bestimmungsgleichungen für den Mittelwert μ und die Varianz $\sigma^2 = e^2/I$ wie (Gl. 4.4) und (Gl. 4.5), so dass an dieser Stelle auf eine erneute Ableitung verzichtet wird. Die so ermittelten Schätzwerte μ und σ^2 maximieren die Likelihood-Funktion und sind optimale Schätzwerte im Sinne der Least-Square-Estimation und des Maximum-Likelihood-Verfahrens.

Unter der Annahme, dass beide Grauwertmengen \underline{G}_{w1} und \underline{G}_{w2} inkompatibel sind (Hypothese H1), was darauf schließen lässt, dass sich die betrachtete Szene verändert hat, erhält man die Likelihood-Funktion $L|_{H1}$ nach (Gl. 4.13) als Produkt der einzelnen Verteilungsdichten $L(g_{i1}, \mu_1, (\sigma_1)^2)$ und $L(g_{i2}, \mu_2, (\sigma_2)^2)$:

$$(Gl. 4.13) \quad L|_{H1} = \left[\frac{1}{2\pi \cdot \sigma_1^2} \right]^{I/2} \cdot \left[\frac{1}{2\pi \cdot \sigma_2^2} \right]^{I/2} \cdot \exp \left\{ -\frac{1}{2\sigma_1^2} \sum_{i=1}^I (g_{i1} - \mu_1)^2 - \frac{1}{2\sigma_2^2} \sum_{i=1}^I (g_{i2} - \mu_2)^2 \right\}$$

Setzt man (Gl. 4.7) und (Gl. 4.9) im Exponent ein, folgt daraus

$$(Gl. 4.14) \quad L|_{H1} = \left[\frac{1}{4\pi^2 \cdot \sigma_1^2 \cdot \sigma_2^2} \right]^{I/2} \cdot \exp \left\{ -\frac{2 \cdot I}{2} \right\}.$$

Stammen die Grauwertmengen \underline{G}_{w1} und \underline{G}_{w2} aus einer identischen Grauwertmenge (Hypothese H0), ergibt sich die Likelihood-Funktion $L|_{H0}$ nach

$$(Gl. 4.15) \quad L|_{H0} = \left[\frac{1}{2\pi \cdot \sigma_0^2} \right]^{2 \cdot I/2} \cdot \exp \left\{ -\frac{1}{2\sigma_0^2} \left[\sum_{i=1}^I (g_{i1} - \mu_0)^2 + \sum_{i=1}^I (g_{i2} - \mu_0)^2 \right] \right\}$$

als resultierende Verbundverteilungsdichte. Mit $(\sigma_0)^2$ nach (Gl. 4.11) folgt

$$(Gl. 4.16) \quad L|_{H0} = \left[\frac{1}{4\pi^2 \cdot (\sigma_0^2)^2} \right]^{I/2} \cdot \exp \left\{ -\frac{2 \cdot I}{2} \right\}.$$

Zur Entscheidung, welche Hypothese erfüllt ist, wird in [Lit. 4.4] das Verhältnis

$$(Gl. 4.17) \quad \frac{L|_{H1}}{L|_{H0}} = \left[\frac{(\sigma_0^2)^2}{\sigma_1^2 \cdot \sigma_2^2} \right]^{I/2} \quad \text{bzw.} \quad \left[\frac{L|_{H1}}{L|_{H0}} \right]^{2/I} = \frac{1}{\sigma_1^2 \cdot \sigma_2^2} \cdot \left[\frac{(\sigma_1^2 + \sigma_2^2)}{2} + \left(\frac{\mu_1 - \mu_2}{2} \right)^2 \right]^I$$

gebildet und mit einem oberen Schwellwert S verglichen. Überschreitet das Verhältnis $(L|_{H1}/L|_{H0})^{2/I}$ den Schwellwert, wird Hypothese H1 (Bewegung erkannt) zugrundegelegt, sonst Hypothese H0 [Lit. 4.4]. Um undefinierte Ergebnisse bei der Berechnung von (Gl. 4.17) auszuschließen, werden die Varianzen $(\sigma_1)^2$ und $(\sigma_2)^2$ in [Lit. 4.4] auf den Minimalwert 1.0 begrenzt. Diese Vorgehensweise wird auch für die im weiteren Verlauf durchgeführten Untersuchungen beibehalten.

Abb. 4.8 und Abb. 4.9 zeigen die Bewertungsergebnisse der grob- und feinmotorischen Bewegungen nach Abb. 4.1 und Abb. 4.2. Insbesondere der Vergleich von Abb. 4.9, Abb. 4.2 und Abb. 4.4 zeigt, dass die konventionelle Schwellwertauswertung des einfachen bzw. des blockweise gemittelten Differenzbildes empfindlicher ist als die Bewertung der Varianz in Verbindung mit der stückweise konstanten Grauwertinterpolation.

Ursache für dieses zunächst unerwartete Verhalten ist die nicht adäquate Modellierung des Grauwertverlaufs im Rahmen des vorliegenden Modells. Grauwertkanten und stark strukturierte Hintergründe bewirken bei der Varianzberechnung einen erhöhten systematischen Fehler [Lit. 4.4]. Sie führen auf Varianzen, die deutlich größer sind als die Varianz des eigentlichen Rauschprozesses [Lit. 4.4], was in (Gl. 4.11) eine zu starke Gewichtung der Varianzen $(\sigma_1)^2$ und $(\sigma_2)^2$ gegenüber den Mittelwerten μ_1 und μ_2 ergibt. Um eine bessere Interpolation zu ermöglichen, wird der Grauwertverlauf im nächsten Abschnitt mit einem linearen Funktionsverlauf nachgebildet.

4.5 Varianzauswertung bei linearer Grauwertinterpolation

Anstelle des konstanten Funktionsverlaufs $f(x_w, y_w) = \mu$ erfolgt in diesem Abschnitt die Interpolation des Grauwertverlaufs im Fenster \underline{G}_w durch die lineare Funktion

$$(Gl. 4.18) \quad f(x_w, y_w) = f_0 + f_x \cdot x_w + f_y \cdot y_w.$$

Berücksichtigt man die in Abschnitt 4.4 beschriebene Spaltenvektordarstellung, folgt für den resultierenden Interpolationsfehler in Analogie zu (Gl. 4.2) die Beziehung

$$(Gl. 4.19) \quad e^2 = \sum_{i=1}^I e_i^2 = \sum_{i=1}^I (g_i - f_0 - f_x \cdot x_i - f_y \cdot y_i)^2.$$

Differenziert man (Gl. 4.19) nach den Parameter f_0 , f_x und f_y und setzt die Ableitungen gleich Null, erhält man wiederum die einzelnen Bestimmungsgleichungen nach dem Least-Square-Verfahren:

$$(Gl. 4.20) \quad \frac{\partial e^2}{\partial f_0} = 0 = -2 \cdot \sum_{i=1}^I (g_i - f_0 - f_x x_i - f_y y_i) \cdot (-1) \quad \Rightarrow \quad f_0 = \frac{1}{I} \cdot \sum_{i=1}^I g_i$$

$$(Gl. 4.21) \quad \frac{\partial e^2}{\partial f_x} = 0 = -2 \cdot \sum_{i=1}^I (g_i - f_0 - f_x x_i - f_y y_i) \cdot (-x_i) \quad \Rightarrow \quad f_x = \left[\sum_{i=1}^I g_i x_i \right] / \left[\sum_{i=1}^I x_i^2 \right]$$

$$(Gl. 4.22) \quad \frac{\partial e^2}{\partial f_y} = 0 = -2 \cdot \sum_{i=1}^I (g_i - f_0 - f_x x_i - f_y y_i) \cdot (-y_i) \quad \Rightarrow \quad f_y = \left[\sum_{i=1}^I g_i y_i \right] / \left[\sum_{i=1}^I y_i^2 \right]$$

Für die weiteren Betrachtungen werden verschiedene Summen über die Ortskoordinaten x_i und y_i benötigt, die im folgenden aufgelistet sind. Unter Beachtung der geometrischen Beziehungen und der Symmetrieeigenschaften nach Abb. 4.7 gelten innerhalb des Grauwertfensters die Beziehungen (Gl. 4.23) - (Gl. 4.27).

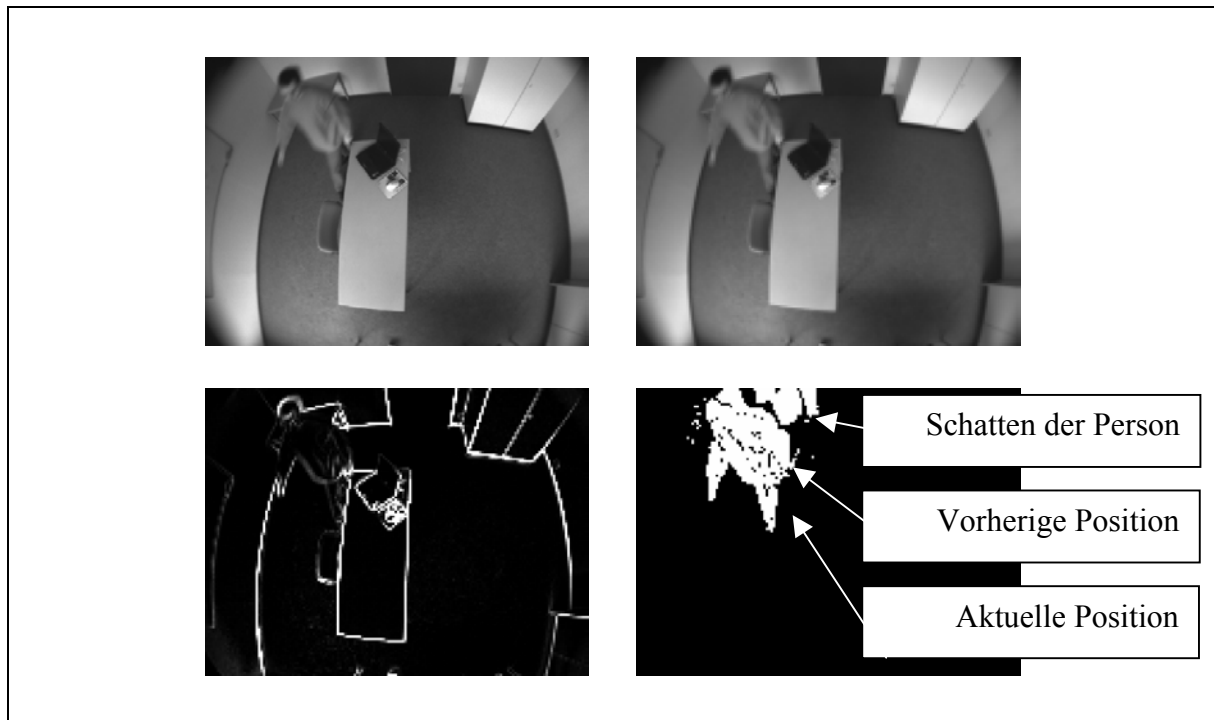


Abb. 4.8: Interpolation des ursprünglichen Grauwertbildes mit bereichsweise konstanten Funktionen und Bewertung der Varianz zur Erkennung von grobmotorischen Bewegungen für die Szene nach Abb. 4.1. Originalbild (oben links), interpoliertes Bild (oben rechts), Varianz (unten links) und mit dem Schwellwert $S_{\max} = 2.5$ bewertete Grauwertänderungen (unten rechts). Fenstergröße 4 x 4 Pixel.

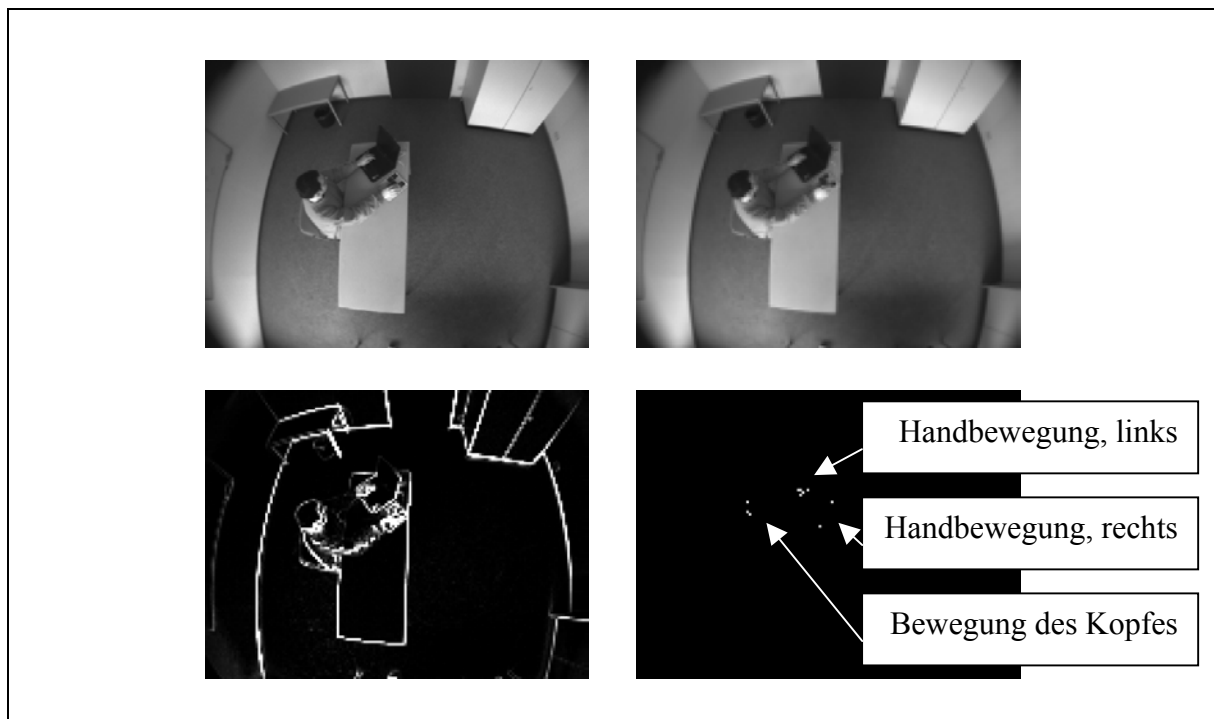


Abb. 4.9: Interpolation des ursprünglichen Grauwertbildes mit bereichsweise konstanten Funktionen und Bewertung der Varianz zur Erkennung von feinmotorischen Bewegungen für die Szene nach Abb. 4.2. Originalbild (oben links), interpoliertes Bild (oben rechts), Varianz (unten links) und mit dem Schwellwert $S = 2.5$ bewertete Grauwertänderungen (unten rechts). Fenstergröße 4 x 4 Pixel.

$$(Gl. 4.23) \quad \sum_{i=1}^I x_i = \sum_{q=0}^{Q-1} \left[\sum_{p=0}^{P-1} \left(\frac{1-P}{2} + p \right) \right] = 0 \quad (Gl. 4.24) \quad \sum_{i=1}^I y_i = \sum_{p=0}^{P-1} \left[\sum_{q=0}^{Q-1} \left(\frac{Q-1}{2} - q \right) \right] = 0$$

$$(Gl. 4.25) \quad \sum_{i=1}^I x_i y_i = \sum_{q=0}^{Q-1} \sum_{p=0}^{P-1} \left[\left(\frac{1-P}{2} + p \right) \cdot \left(\frac{Q-1}{2} - q \right) \right] = 0 \quad \text{sowie}$$

$$(Gl. 4.26) \quad \sum_{i=1}^I x_i^2 = \sum_{q=0}^{Q-1} \left[\sum_{p=0}^{P-1} \left(\frac{1-P}{2} + p \right)^2 \right] = \frac{P \cdot Q}{12} \cdot (P-1) \cdot (P+1) = \frac{I}{12} \cdot (I-1) \quad \text{für } P=Q=\sqrt{I}$$

$$(Gl. 4.27) \quad \sum_{i=1}^I y_i^2 = \sum_{p=0}^{P-1} \left[\sum_{q=0}^{Q-1} \left(\frac{Q-1}{2} - q \right)^2 \right] = \frac{P \cdot Q}{12} \cdot (Q-1) \cdot (Q+1) = \frac{I}{12} \cdot (I-1) \quad \text{für } P=Q=\sqrt{I}$$

In der Hypothese H1 (inkompatible Grauwertmengen) gelten unter Anwendung von (Gl. 4.26) und (Gl. 4.27) somit die Gleichungen (Gl. 4.28) – (Gl. 4.35) zur Ermittlung der Parameter der Interpolationsfunktion $f(x_w, y_w)$. In Analogie zu Abschnitt 4.4 ergibt sich außerdem die Likelihood-Funktion $L|_{H1}$ nach (Gl. 4.36) als Produkt der Einzelverteilungsdichten $L(g_{i1}, f_{01}, f_{x1}, f_{y1}, (\sigma_1)^2)$ und $L(g_{i2}, f_{02}, f_{x2}, f_{y2}, (\sigma_2)^2)$, wobei die Varianzen $(\sigma_1)^2$ und $(\sigma_2)^2$ durch (Gl. 4.34) und (Gl. 4.35) festgelegt sind.

$$(Gl. 4.28) \quad f_{01} = \frac{1}{I} \cdot \sum_{i=1}^I g_{i1}$$

$$(Gl. 4.29) \quad f_{02} = \frac{1}{I} \cdot \sum_{i=1}^I g_{i2}$$

$$(Gl. 4.30) \quad f_{x1} = \frac{12}{I \cdot (I-1)} \cdot \sum_{i=1}^I g_{i1} x_i$$

$$(Gl. 4.31) \quad f_{x2} = \frac{12}{I \cdot (I-1)} \cdot \sum_{i=1}^I g_{i2} x_i$$

$$(Gl. 4.32) \quad f_{y1} = \frac{12}{I \cdot (I-1)} \cdot \sum_{i=1}^I g_{i1} y_i$$

$$(Gl. 4.33) \quad f_{y2} = \frac{12}{I \cdot (I-1)} \cdot \sum_{i=1}^I g_{i2} y_i$$

$$(Gl. 4.34) \quad \sigma_1^2 = \frac{1}{I} \sum_{i=1}^I (g_{i1} - f_{01} - f_{x1} x_i - f_{y1} y_i)^2$$

$$(Gl. 4.35) \quad \sigma_2^2 = \frac{1}{I} \sum_{i=1}^I (g_{i2} - f_{02} - f_{x2} x_i - f_{y2} y_i)^2$$

(Gl. 4.36)

$$L|_{H1} = \left(\frac{1}{4\pi^2 \sigma_1^2 \sigma_2^2} \right)^{I/2} \exp \left\{ -\frac{1}{2} \left[\frac{1}{\sigma_1^2} \sum_{i=1}^I (g_{i1} - f_{01} - f_{x1} x_i - f_{y1} y_i)^2 + \frac{1}{\sigma_2^2} \sum_{i=1}^I (g_{i2} - f_{02} - f_{x2} x_i - f_{y2} y_i)^2 \right] \right\}$$

Während der Hypothese H0, d. h. beide Grauwertmengen \underline{G}_{w1} und \underline{G}_{w2} stammen aus derselben Grundgesamtheit \underline{G}_{w0} , gilt für die Parameter f_{00} , f_{x0} und f_{y0} :

$$(Gl. 4.37) \quad f_{00} = \frac{1}{2I} \cdot \left(\sum_{i=1}^I g_{i1} + \sum_{i=1}^I g_{i2} \right) = \frac{f_{01} + f_{02}}{2} \quad \text{sowie}$$

$$(Gl. 4.38) \quad f_{x0} = \left(\sum_{i=1}^I g_{i1} x_i + \sum_{i=1}^I g_{i2} x_i \right) / \left(2 \sum_{i=1}^I x_i^2 \right) = \frac{6}{I \cdot (I-1)} \left(\sum_{i=1}^I g_{i1} x_i + \sum_{i=1}^I g_{i2} x_i \right) = \frac{f_{x1} + f_{x2}}{2}$$

$$(Gl. 4.39) \quad f_{y0} = \left(\sum_{i=1}^I g_{i1} y_i + \sum_{i=1}^I g_{i2} y_i \right) / \left(2 \sum_{i=1}^I y_i^2 \right) = \frac{6}{I \cdot (I-1)} \left(\sum_{i=1}^I g_{i1} y_i + \sum_{i=1}^I g_{i2} y_i \right) = \frac{f_{y1} + f_{y2}}{2}$$

Für die resultierende Varianz $(\sigma_0)^2$ ergibt sich die Beziehung

$$(Gl. 4.40) \quad \sigma_0^2 = \frac{1}{2I} \left(\sum_{i=1}^I (g_{i1} - f_{00} - f_{x0}x_i - f_{y0}y_i)^2 + \sum_{i=1}^I (g_{i2} - f_{00} - f_{x0}x_i - f_{y0}y_i)^2 \right),$$

woraus man durch Einsetzen von f_{00} , f_{x0} und f_{y0} nach (Gl. 4.37) - (Gl. 4.39) und einigen Umformungen den Ausdruck

$$(Gl. 4.41) \quad \sigma_0^2 = \frac{\sigma_1^2 + \sigma_2^2}{2} + \left(\frac{f_{01} - f_{02}}{2} \right)^2 + \frac{I-1}{12} \left(\frac{f_{x1} - f_{x2}}{2} \right)^2 + \frac{I-1}{12} \left(\frac{f_{y1} - f_{y2}}{2} \right)^2$$

ableiten kann. Als Likelihood-Funktion $L|_{H0}$ erhält man in diesem Fall

$$(Gl. 4.42) \quad L|_{H0} = \left(\frac{1}{2\pi \cdot \sigma_0^2} \right)^{2I/2} \exp \left\{ -\frac{1}{2\sigma_0^2} \left[\sum_{i=1}^I (g_{i1} - f_{00} - f_{x0}x_i - f_{y0}y_i)^2 + \sum_{i=1}^I (g_{i2} - f_{00} - f_{x0}x_i - f_{y0}y_i)^2 \right] \right\}.$$

Entsprechend der Vorgehensweise in Abschnitt 4.4 wird das Verhältnis $L|_{H1}/L|_{H0}$ gebildet und die Varianzen nach (Gl. 4.34), (Gl. 4.35) und (Gl. 4.40) in den Exponentialterm eingesetzt. Verwendet man zusätzlich die Varianz $(\sigma_0)^2$ nach (Gl. 4.41) im Nenner von $L|_{H0}$ und potenziert das Likelihood-Verhältnis zur Vereinfachung der Berechnungen mit $2/I$, erhält man

$$(Gl. 4.43) \quad \left[\frac{L|_{H1}}{L|_{H0}} \right]^{2/I} = \frac{1}{\sigma_1^2 \cdot \sigma_2^2} \cdot \left[\frac{\sigma_1^2 + \sigma_2^2}{2} + \left(\frac{f_{01} - f_{02}}{2} \right)^2 + \frac{I-1}{12} \left(\frac{f_{x1} - f_{x2}}{2} \right)^2 + \frac{I-1}{12} \left(\frac{f_{y1} - f_{y2}}{2} \right)^2 \right]^{-2}.$$

(Gl. 4.43) dient wiederum zur Unterscheidung zwischen den Hypothesen H1 und H0, indem bei Überschreiten der Schwelle S die Hypothese H1 und sonst Hypothese H0 gewählt wird.

Abb. 4.10 und Abb. 4.11 zeigen die Bewertung der Szenen nach Abb. 4.1 und Abb. 4.2 in Verbindung mit der bereichsweise linearen Grauwertinterpolation. Sowohl anhand der stärker geschlossenen Änderungsbereiche bei grobmotorischen Bewegungen als auch anhand der Anzahl detektierter Grauwertänderungen bei feinmotorischen Bewegungsabläufen erkennt man unmittelbar, dass das zuletzt beschriebene Verfahren deutlich empfindlicher ist als die Auswertestrategien der Abschnitte 4.2 und 4.4. Aufgrund der höheren Interpolationsgüte treten kleinere systematische Abweichungen bei der Varianzberechnung auf, wodurch die vorher beschriebene Überbewertung der Varianzen $(\sigma_1)^2$ und $(\sigma_2)^2$ bei der Bewertung des Likelihood-Verhältnisses reduziert wird.

Allerdings beschreibt auch eine lineare Interpolationsfunktion den tatsächlichen Grauwertverlauf nur näherungsweise. Weiterführende Untersuchungen, die in [Lit. 4.3] und [Lit. 4.4] erwähnt sind, haben gezeigt, dass quadratische Interpolationsfunktionen für Grauwertbereiche mit starker Textur eine nochmalige Verbesserung liefern. In Verbindung mit diesem Ansatz ergibt sich eine so hohe Interpolationsgüte, dass die verbleibenden Varianzen hauptsächlich auf die Rauschprozesse bei der Bildaufnahme zurückgeführt werden können [Lit. 4.4].

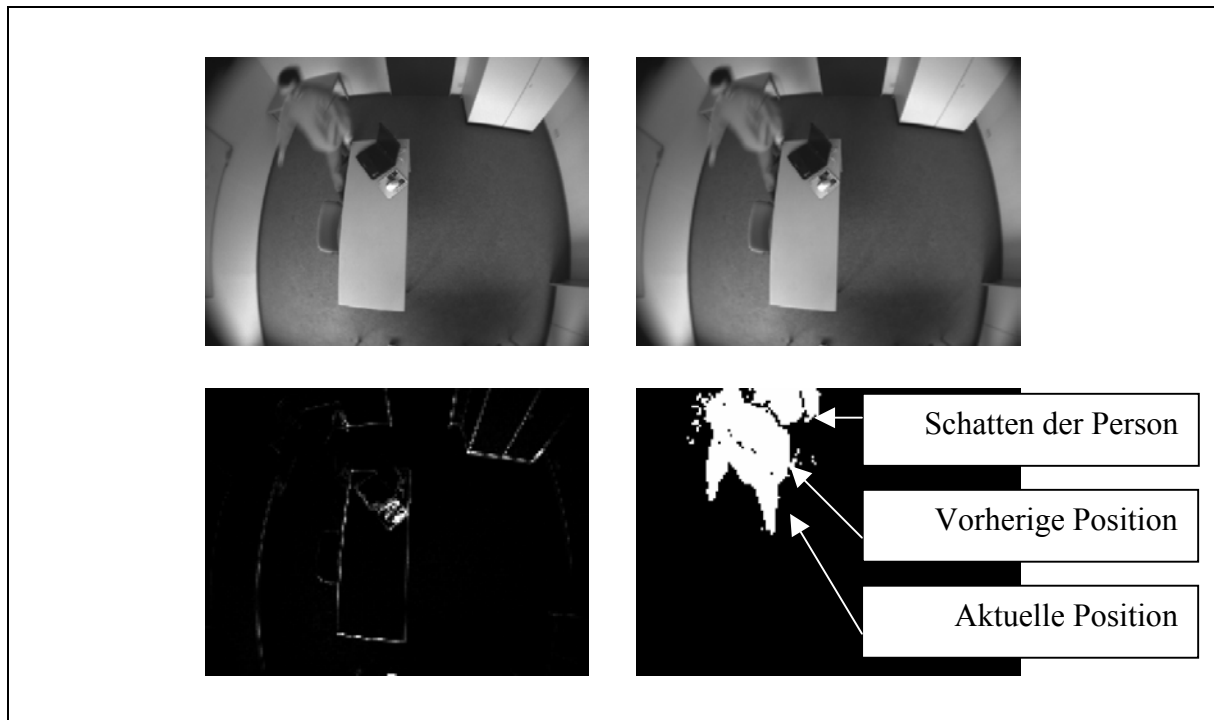


Abb. 4.10: Interpolation des ursprünglichen Grauwertbildes mit bereichsweise linearen Funktionen und Bewertung der Varianz zur Erkennung von grobmotorischen Bewegungen für die Szene nach Abb. 4.1. Originalbild (oben links), interpoliertes Bild (oben rechts), Varianz (unten links) und mit dem Schwellwert $S = 2.5$ bewertete Grauwertänderungen (unten rechts). Fenstergröße 4×4 Pixel.

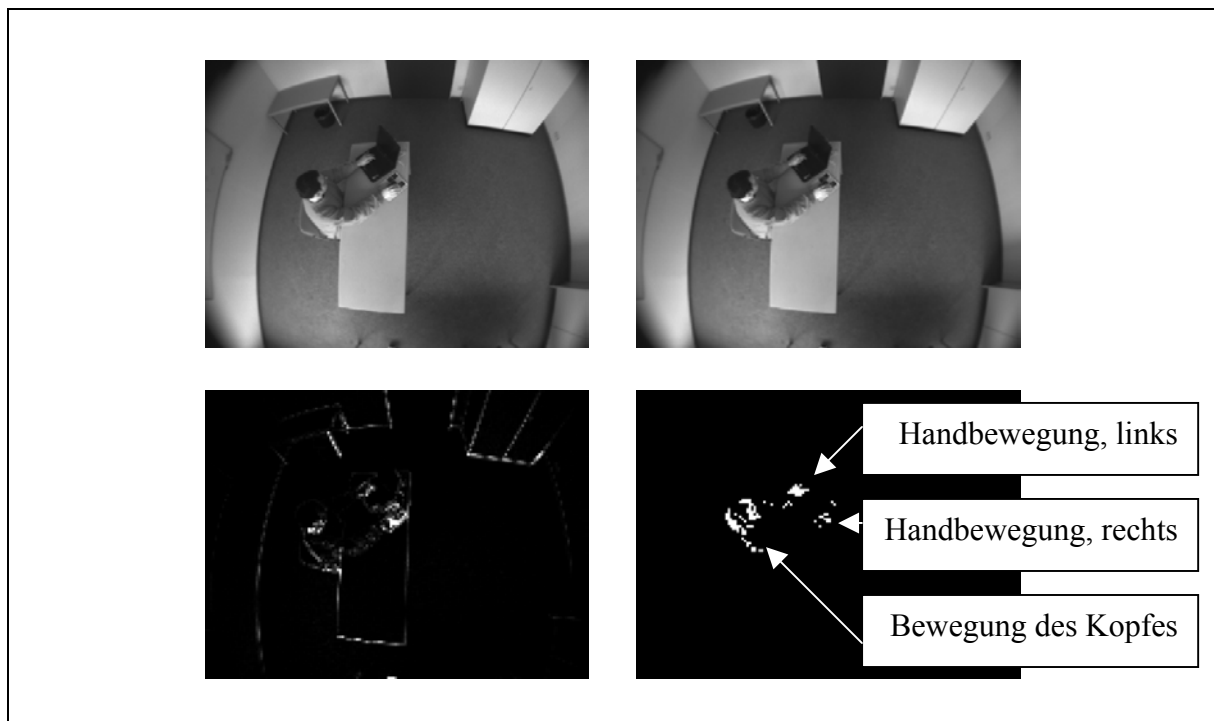


Abb. 4.11: Interpolation des ursprünglichen Grauwertbildes mit bereichsweise linearen Funktionen und Bewertung der Varianz zur Erkennung von feinmotorischen Bewegungen für die Szene nach Abb. 4.2. Originalbild (oben links), interpoliertes Bild (oben rechts), Varianz (unten links) und mit dem Schwellwert $S = 2.5$ bewertete Grauwertänderungen (unten rechts). Fenstergröße 4×4 Pixel.

Die höhere Empfindlichkeit für bewegungsbedingte Veränderungen, aber auch für Rauschstörungen erfordert weitere Maßnahmen zur Vermeidung von Fehlalarmen. Da bei der vorliegenden Applikation auch geringfügige Bewegungen erkannt werden müssen, können isolierte Störungen zwar durch Nachbarschaftsoperationen unterdrückt werden, allerdings darf die zur Rauschunterdrückung verwendete minimale Bereichsgröße nicht zu stark ausgeweitet werden.

Ein wesentlicher Nachteil aller bisher beschriebenen Verfahren ist die heuristische Festlegung der Auswerteschwellen. Im weiteren Verlauf werden daher zwei Möglichkeiten vorgestellt, welche den aus der Statistik bekannten t- bzw. F-Test zur Identifikation von systematischen Grauwertänderungen in der Szene nutzen. Der t-Test wird in Verbindung mit der konstanten Grauwertinterpolation verwendet (s. Abschnitt 4.7), während der F-Test nach Abschnitt 4.8 für lineare und quadratische Interpolationsfunktionen zum Einsatz kommt.

Anfang der 80-er Jahre wurden beide Methoden von Hans-Hellmut Nagel entwickelt. Sie sind z. B. in [Lit. 4.3] und [Lit. 4.4] beschrieben. Die zum Verständnis erforderlichen Grundlagen werden in Abschnitt 4.6 erläutert.

4.6 Matrixorientiertes Verfahren zur Grauwertinterpolation

Entsprechend den Definitionen nach Abb. 4.7 wird die Teilmatrix \underline{G}_w wiederum in den Spaltenvektor \underline{g} mit den Grauwerten g_i und $i \in [1, I]$ abgebildet. Die zugehörigen Ortskoordinaten sind in den Spaltenvektoren \underline{x} und \underline{y} mit den Elementen x_i und y_i enthalten. Für die weiteren Ableitungen wird außerdem der Spaltenvektor \underline{f} , dessen Komponenten aus den Interpolationsparametern bestehen, sowie die Matrix \underline{X} , deren Elemente aus den Ortskoordinaten x_i und y_i gebildet werden, benötigt.

Der Spaltenvektor \underline{f} wird für eine konstante Grauwertinterpolation durch den Vektor \underline{f}_k nach (Gl. 4.44) ersetzt, während für die lineare Interpolation der Vektor \underline{f}_l nach (Gl. 4.44) und für die quadratische Interpolation der Vektor \underline{f}_q nach (Gl. 4.44) gilt, wobei "T" die Transponierung des jeweiligen Vektors andeutet. Je nach Interpolationsverfahren wird die Matrix \underline{X} außerdem durch die Matrizen \underline{X}_k , \underline{X}_l bzw. \underline{X}_q nach (Gl. 4.45) ersetzt.

$$(Gl. 4.44) \quad \underline{f}_k = \mu \quad \underline{f}_l = [f_0 \quad f_x \quad f_y]^T \quad \underline{f}_q = [f_0 \quad f_x \quad f_y \quad f_{xx} \quad f_{yy} \quad f_{xy}]^T$$

$$(Gl. 4.45) \quad \underline{X}_k = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \underline{X}_l = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_1 & y_1 \end{bmatrix} \quad \underline{X}_q = \begin{bmatrix} 1 & x_1 & y_1 & x_1^2 & y_1^2 & x_1 y_1 \\ 1 & x_2 & y_2 & x_2^2 & y_2^2 & x_2 y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_1 & y_1 & x_1^2 & y_1^2 & x_1 y_1 \end{bmatrix}$$

Mit den Grauwerten $\underline{g} = [g_1 \ g_2 \ \dots \ g_I]^T$ und dem Interpolationsfehler $\underline{e} = [e_1 \ e_2 \ \dots \ e_I]^T$ folgt dann

$$(Gl. 4.46) \quad \underline{e} = \underline{g} - \underline{X} \cdot \underline{f} \quad \text{bzw.} \quad \sum_{i=1}^I e_i^2 = \underline{e}^T \cdot \underline{e} = (\underline{g} - \underline{X} \cdot \underline{f})^T \cdot (\underline{g} - \underline{X} \cdot \underline{f}),$$

woraus sich die Likelihood-Funktion $L(\underline{g}, \underline{f}, \sigma^2)$ der Grauwertmenge \underline{G} nach (Gl. 4.47) ableiten lässt.

$$(Gl. 4.47) \quad L(\underline{g}, \underline{f}, \sigma^2) = \left[\frac{1}{2\pi \sigma^2} \right]^{1/2} \cdot \exp \left\{ -\frac{1}{2\sigma^2} (\underline{g} - \underline{X} \cdot \underline{f})^T (\underline{g} - \underline{X} \cdot \underline{f}) \right\}$$

Logarithmiert man (Gl. 4.47), erhält man aus

$$\ln(L(\underline{g}, \underline{f}, \sigma^2)) = -\frac{1}{2} \ln(2\pi) - \frac{1}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} (\underline{g} - \underline{X} \cdot \underline{f})^T (\underline{g} - \underline{X} \cdot \underline{f})$$

durch Differentiation nach den Parametern \underline{f} und σ^2 und Null setzen der Ableitungen die Bestimmungsgleichungen (Gl. 4.48) und (Gl. 4.49).

$$(Gl. 4.48) \quad \frac{\partial \ln(L)}{\partial \underline{f}} = 0 = \frac{1}{\sigma^2} \cdot \underline{X}^T \cdot (\underline{g} - \underline{X} \cdot \underline{f}) \quad \Rightarrow \quad \underline{f} = (\underline{X}^T \cdot \underline{X})^{-1} \cdot \underline{X}^T \cdot \underline{g}$$

$$(Gl. 4.49) \quad \frac{\partial \ln(L)}{\partial \sigma^2} = 0 = -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4} (\underline{g} - \underline{X} \cdot \underline{f})^T (\underline{g} - \underline{X} \cdot \underline{f}) \quad \Rightarrow \quad \sigma^2 = \frac{1}{I} (\underline{g} - \underline{X} \cdot \underline{f})^T (\underline{g} - \underline{X} \cdot \underline{f})$$

Ersetzt man den Spaltenvektor \underline{f} und die Matrix \underline{X} durch \underline{f}_k und \underline{X}_k nach (Gl. 4.44), erhält man für die bereichsweise konstante Grauwertinterpolation den Mittelwert nach (Gl. 4.4) und die Varianz $\sigma^2 = e^2/I$ nach (Gl. 4.5). Für die lineare Grauwertinterpolation ergeben sich die Parameter f_0 , f_x und f_y nach (Gl. 4.20) - (Gl. 4.22), wenn \underline{f} und \underline{X} durch \underline{f}_1 und \underline{X}_1 nach (Gl. 4.44) ersetzt werden. Für die quadratische Grauwertinterpolation wurde bisher keine Summenschiebweise berechnet. Vielmehr wurden die weiter unten gezeigten Bewertungsergebnisse mit dem matrixorientierten Programmpaket MATLAB erzeugt, welches die unmittelbare Eingabe der Vektoren und Matrizen aus (Gl. 4.48) und (Gl. 4.49) gestattet.

Im Gegensatz zu den Abschnitten 4.4 und 4.5, wo zwei aufeinanderfolgenden Teilbildern \underline{G}_{w1} und \underline{G}_{w2} jeweils unterschiedliche Varianzen $(\sigma_1)^2$ und $(\sigma_2)^2$ zugeordnet wurden, wird im weiteren Verlauf davon ausgegangen, dass sich die Varianz $\sigma^2 = (\sigma_1)^2 = (\sigma_2)^2$ der zugrundeliegenden Rauschprozesse über der Zeit nicht ändert. In der Hypothese H1 folgt somit für die aus (Gl. 4.36) bekannte Verbundverteilungsdichte in Vektor-Matrix-Form die Beziehung

$$(Gl. 4.50) \quad L|_{H1} = \left(\frac{1}{2\pi \sigma^2} \right)^I \exp \left\{ -\frac{1}{2\sigma^2} \left[(\underline{g}_1 - \underline{X} \cdot \underline{f})^T (\underline{g}_1 - \underline{X} \cdot \underline{f}) + (\underline{g}_2 - \underline{X} \cdot \underline{f})^T (\underline{g}_2 - \underline{X} \cdot \underline{f}) \right] \right\},$$

aus der sich in bekannter Weise die Interpolationsparameter in den Vektoren \underline{f}_1 und \underline{f}_2 sowie die resultierende Varianz σ^2 ableiten lassen. In Analogie zu (Gl. 4.48) und (Gl. 4.49) erhält man die Interpolationsparameter nach (Gl. 4.51) und (Gl. 4.52) sowie die Varianz nach (Gl. 4.53) für den Fall inkompatibler Grauwertmengen \underline{G}_{w1} und \underline{G}_{w2} bzw. \underline{g}_1 und \underline{g}_2 .

$$(Gl. 4.51) \quad \underline{f}_1 = (\underline{X}^T \cdot \underline{X})^{-1} \cdot \underline{X}^T \cdot \underline{g}_1 \quad (Gl. 4.52) \quad \underline{f}_2 = (\underline{X}^T \cdot \underline{X})^{-1} \cdot \underline{X}^T \cdot \underline{g}_2$$

$$(Gl. 4.53) \quad \sigma^2 = \frac{1}{2I} \left[(\underline{g}_1 - \underline{X} \cdot \underline{f}_1)^T (\underline{g}_1 - \underline{X} \cdot \underline{f}_1) + (\underline{g}_2 - \underline{X} \cdot \underline{f}_2)^T (\underline{g}_2 - \underline{X} \cdot \underline{f}_2) \right]$$

Zur Bestimmung der Likelihood-Funktion $L|_{H_0}$ während der Hypothese H_0 definiert man zunächst einen neuen Grauwertvektor \underline{g}_0 , der die Vektoren \underline{g}_1 und \underline{g}_2 beinhaltet. Aufgrund der vergrößerten Grauwertmenge müssen die Matrix \underline{X} und der Fehlervektor \underline{e}_0 ebenfalls angepasst werden. Mit

$$(Gl. 4.54) \quad \underline{g}_0 = \begin{bmatrix} \underline{g}_1 \\ \underline{g}_2 \end{bmatrix} \quad (Gl. 4.55) \quad \underline{X}_0 = \begin{bmatrix} \underline{X} \\ \underline{X} \end{bmatrix} \quad (Gl. 4.56) \quad \underline{e}_0 = [\underline{e}_1 \quad \underline{e}_2 \quad \dots \quad \underline{e}_{2l}]^T$$

und dem Parametervektor \underline{f}_0 erhält man die Likelihood-Funktion

$$(Gl. 4.57) \quad L|_{H_0} = \left(\frac{1}{2\pi \sigma_0^2} \right)^{2l/2} \exp \left\{ -\frac{1}{2\sigma_0^2} (\underline{g}_0 - \underline{X}_0 \cdot \underline{f}_0)^T (\underline{g}_0 - \underline{X}_0 \cdot \underline{f}_0) \right\}.$$

Für den Parametervektor \underline{f}_0 ergibt sich mit der Maximum-Likelihood-Methode und unter Beachtung von (Gl. 4.48), (Gl. 4.49) sowie (Gl. 4.57) die Berechnungsvorschrift

$$(Gl. 4.58) \quad \underline{f}_0 = (\underline{X}_0^T \cdot \underline{X}_0)^{-1} \cdot \underline{X}_0^T \cdot \underline{g}_0 = \frac{\underline{f}_1 + \underline{f}_2}{2},$$

während sich für die Varianz $(\sigma_0)^2$ durch Einsetzen von (Gl. 4.58) die Beziehung

$$(Gl. 4.59) \quad \sigma_0^2 = \frac{1}{2l} \left[(\underline{g}_1 - \underline{X} \underline{f}_1)^T (\underline{g}_1 - \underline{X} \underline{f}_1) + (\underline{g}_2 - \underline{X} \underline{f}_2)^T (\underline{g}_2 - \underline{X} \underline{f}_2) + \frac{1}{2} (\underline{f}_1 - \underline{f}_2)^T \underline{X}^T \underline{X} (\underline{f}_1 - \underline{f}_2) \right]$$

ableiten lässt. Die ersten beiden Summanden in (Gl. 4.59) stimmen mit der Varianz σ^2 in der Hypothese H_1 nach (Gl. 4.53) überein. Somit lässt sich die Varianz $(\sigma_0)^2$ in der Hypothese H_0 aus der Varianz σ^2 und den Parametervektoren \underline{f}_1 und \underline{f}_2 der Grauwertmengen \underline{G}_{w1} und \underline{G}_{w2} in der Hypothese H_1 berechnen. Unter Anwendung von (Gl. 4.53) und (Gl. 4.59) ergibt sich aus (Gl. 4.50) und (Gl. 4.57) das Verhältnis

$$(Gl. 4.60) \quad \left[\frac{L|_{H_1}}{L|_{H_0}} \right]^{1/2} = 1 + \frac{1}{2} \frac{(\underline{f}_1 - \underline{f}_2)^T \underline{X}^T \underline{X} (\underline{f}_1 - \underline{f}_2)}{(\underline{g}_1 - \underline{X} \underline{f}_1)^T (\underline{g}_1 - \underline{X} \underline{f}_1) + (\underline{g}_2 - \underline{X} \underline{f}_2)^T (\underline{g}_2 - \underline{X} \underline{f}_2)}$$

der Likelihood-Funktionen $L|_{H_1}$ und $L|_{H_0}$. Je nach Interpolationsprofil erhält man daraus für die bereichsweise konstante Grauwertinterpolation das Verhältnis (Gl. 4.62), für die lineare Interpolation folgt (Gl. 4.63) und für die Interpolation mit quadratischen Funktionsverläufen der Form

$$(Gl. 4.61) \quad f(x_w, y_w) = f_0 + f_x \cdot x_w + f_y \cdot y_w + f_{xx} \cdot x_w^2 + f_{yy} \cdot y_w^2 + f_{xy} \cdot x_w y_w$$

erhält man nach [Lit. 4.4] die Beziehung (Gl. 4.64).

$$(Gl. 4.62) \quad \left[\frac{L|_{H_1}}{L|_{H_0}} \right]_k^{1/2} = 1 + \frac{1}{2} \cdot \frac{I \cdot (\mu_1 - \mu_2)^2}{\sum_{i=1}^l (g_{i1} - \mu_1)^2 + \sum_{i=1}^l (g_{i2} - \mu_2)^2}$$

$$(Gl. 4.63) \quad \left[\frac{L|_{HI}}{L|_{HO}} \right]_1^{1/2} = 1 + \frac{1}{2} \cdot \frac{I \cdot (f_{01} - f_{02})^2 + (f_{x1} - f_{x2})^2 \sum_{i=1}^I x_i^2 + (f_{y1} - f_{y2})^2 \sum_{i=1}^I y_i^2}{\sum_{i=1}^I (g_{i1} - f_{01} - f_{x1}x_i - f_{y1}y_i)^2 + \sum_{i=1}^I (g_{i2} - f_{02} - f_{x2}x_i - f_{y2}y_i)^2}$$

$$(Gl. 4.64) \quad \left[\frac{L|_{HI}}{L|_{HO}} \right]_q^{1/2} = 1 + \frac{1}{2} \cdot \frac{Z}{N} \quad \text{mit}$$

$$\begin{aligned} Z = & I \cdot (f_{01} - f_{02})^2 + (f_{x1} - f_{x2})^2 \sum_{i=1}^I x_i^2 + (f_{y1} - f_{y2})^2 \sum_{i=1}^I y_i^2 + \dots \\ & + (f_{xx1} - f_{xx2})^2 \sum_{i=1}^I x_i^4 + (f_{yy1} - f_{yy2})^2 \sum_{i=1}^I y_i^4 + (f_{xy1} - f_{xy2})^2 \sum_{i=1}^I x_i^2 y_i^2 \\ N = & \sum_{i=1}^I (g_{i1} - f_{01} - f_{x1}x_i - f_{y1}y_i - f_{xx1}x_i^2 - f_{yy1}y_i^2 - f_{xy1}x_i y_i)^2 + \dots \\ & + \sum_{i=1}^I (g_{i2} - f_{02} - f_{x2}x_i - f_{y2}y_i - f_{xx2}x_i^2 - f_{yy2}y_i^2 - f_{xy2}x_i y_i)^2 \end{aligned}$$

Je nach Interpolationsmethode dienen (Gl. 4.62), (Gl. 4.63) und (Gl. 4.64) zur Identifikation von Grauwertänderungen. Ausgehend von diesen Beziehungen werden in [Lit. 4.4] der t-Test für die bereichsweise konstante und der F-Test für das lineare und quadratische Interpolationsverfahren entwickelt. Beide Vorgehensweise sind in starker Anlehnung an [Lit. 4.4] in den nächsten Abschnitten zusammen mit den Bewertungsergebnissen grob- und feinmotorischer Bewegungsabläufe vorgestellt.

4.7 Anwendung des t-Tests

In diesem Abschnitt wird die Anwendung des t-Tests (s. z. B. [Lit. 4.5]) zur Erkennung von Grauwertänderungen in Verbindung mit (Gl. 4.62) erläutert. Der t-Test dient ganz allgemein zur Prüfung der Hypothese, ob zwei Proben aus einer gemeinsamen oder zwei unterschiedlichen Grundmengen stammen. Die Prüfung erfolgt insbesondere anhand der Mittelwerte der Proben, wobei vorausgesetzt wird, dass die Varianzen beider Grundmengen übereinstimmen.

Wie bereits vorher erwähnt wurde, werden die mit mittelwertfreiem, Gaußverteilterm Rauschen gestörten Grauwerte g_{i1} und g_{i2} mit $i \in [1, I]$ als normalverteilte, statistisch unabhängige Zufallsvariablen betrachtet. Außerdem wurde jeder Grauwertmenge \underline{G}_{w1} und \underline{G}_{w2} die über der Zeit unveränderliche Varianz σ^2 zugeordnet. Die Mittelwerte μ_1 und μ_2 der Grauwertmengen sind folglich normalverteilte Zufallsvariablen, die jeweils mit der Varianz σ^2/I um den wahren Wert schwanken. Die Differenz $\mu_\Delta = \mu_1 - \mu_2$ der Mittelwerte in (Gl. 4.62) ist somit ebenfalls eine erwartungswertfreie, normalverteilte Zufallsvariable, deren Varianz $(\sigma_\Delta)^2$ sich infolge der statistischen Unabhängigkeit aus der Summe der einzelnen Varianzen ergibt. Es gilt $(\sigma_\Delta)^2 = 2 \cdot \sigma^2/I$. Durch die Normierung von μ_Δ auf die Standardabweichung $\sigma_\Delta = \sigma \cdot (2/I)^{1/2}$ folgt dann die erwartungswertfreie, Gaußverteilte Zufallsvariable η_Δ mit der Varianz 1 [Lit. 4.4]:

$$(Gl. 4.65) \quad \eta_\Delta = \frac{\mu_\Delta}{\sigma_\Delta} = \frac{\mu_1 - \mu_2}{\sigma \cdot \sqrt{2/I}}$$

Normiert man beide Fehlersummen im Nenner von (Gl. 4.62) auf die Varianz σ^2 , ergeben sich zwei unabhängige χ^2 -Verteilungen mit jeweils I-1 Freiheitsgraden [Lit. 4.4].

$$\rho_{\Delta} = \sum_{i=1}^I (g_{i1} - \mu_1)^2 / \sigma^2 + \sum_{i=1}^I (g_{i2} - \mu_2)^2 / \sigma^2$$

Die resultierende Summe ist somit ebenfalls χ^2 -verteilt mit insgesamt $F_N = 2 \cdot (I-1)$ Freiheitsgraden. Daher weist die Zufallsvariable t in (Gl. 4.66) eine Studentische t-Verteilung auf, die nach [Lit. 4.5] dadurch charakterisiert ist, dass im Zähler eine erwartungswertfreie, normalverteilte, normierte Zufallsvariable auftritt, während im Nenner eine auf $(F_N)^{1/2}$ normierte χ^2 -Verteilung mit F_N Freiheitsgraden vorhanden ist.

$$(Gl. 4.66) \quad t = \frac{\eta_{\Delta}}{\sqrt{\rho_{\Delta} / (2 \cdot (I-1))}} = \frac{\frac{\mu_1 - \mu_2}{\sigma \sqrt{2/I}}}{\sqrt{\frac{1}{2 \cdot (I-1) \cdot \sigma^2} \left(\sum_{i=1}^I (g_{i1} - \mu_1)^2 + \sum_{i=1}^I (g_{i2} - \mu_2)^2 \right)}}$$

Mit der Zufallsvariable t kann (Gl. 4.62) folgendermaßen formuliert werden:

$$(Gl. 4.67) \quad \left[\frac{L|_{H1}}{L|_{H0}} \right]_k^I = 1 + \frac{1}{2} \cdot \frac{t^2}{2 \cdot (I-1)}$$

Da das Likelihood-Verhältnis nach (Gl. 4.67) eine monoton steigende Funktion von t ist, wird anstelle der Summe auf der rechten Seite lediglich t^2 zur Hypothesenprüfung verwendet [Lit. 4.4]. Überschreitet t^2 die Schwelle S^2 , wird die Hypothese H1 ausgewählt, ansonsten die Hypothese H0. Um den Rechenaufwand bei der Hypothesenprüfung zu verringern, wird

$$(Gl. 4.68) \quad I \cdot (I-1) \cdot (\mu_1 - \mu_2)^2 \begin{matrix} > \\ \leq \end{matrix} S^2 \left(\sum_{i=1}^I (g_{i1} - \mu_1)^2 + \sum_{i=1}^I (g_{i2} - \mu_2)^2 \right) \begin{matrix} \Rightarrow \text{Hypothese H1} \\ \Rightarrow \text{Hypothese H0} \end{matrix}$$

ausgewertet. Bei der Festlegung der Schwelle S wird die in [Lit. 4.5] in Abhängigkeit der Freiheitsgradanzahl F_N tabellierte Studentische t-Verteilung zugrundegelegt. Tab. 4.1 enthält eine Auswahl für unterschiedliche Fenstergrößen. Bei gegebenem Konfidenzniveau α , welches die Wahrscheinlichkeit angibt, das Hypothese H1 fälschlicherweise gewählt wurde, lässt sich darin der Schwellwert S ablesen.

T-TEST							
α	0.100	0.050	0.020	0.010	0.005	0.002	0.001
Pixelanzahl							
4 (2x2)	1.943	2.447	3.143	3.707	4.317	5.208	5.959
9 (3x3)	1.746	2.120	2.583	2.921	3.252	3.686	4.015
16 (4x4)	1.697	2.042	2.457	2.750	3.030	3.386	3.646
∞	1.645	1.960	2.326	2.576	2.807	3.090	3.291

Tab. 4.1: 100α -prozentige Werte S der Studentischen t-Verteilung [Lit. 4.5] in Abhängigkeit der Freiheitsgrade F_N und des Vertrauensniveaus α für diverse Fenstergröße

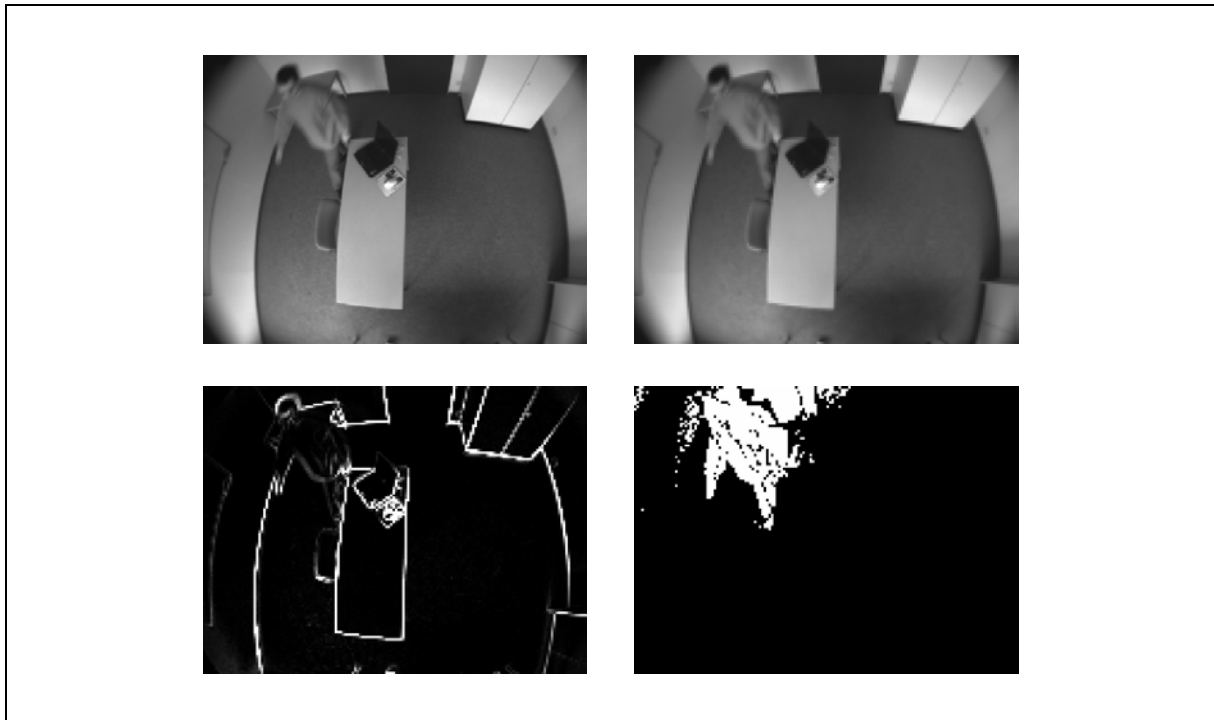


Abb. 4.12: Interpolation des ursprünglichen Grauwertbildes mit bereichsweise konstanten Funktionen und Anwendung des t-Tests zur Erkennung von grobmotorischen Bewegungen für die Szene nach Abb. 4.1. Originalbild (oben links), interpoliertes Bild (oben rechts), Varianz (unten links) und mit dem Schwellwert $S^2 = 3.030^2$ bewertete Grauwertänderungen (unten rechts). Fenstergröße 4 x 4 Pixel.

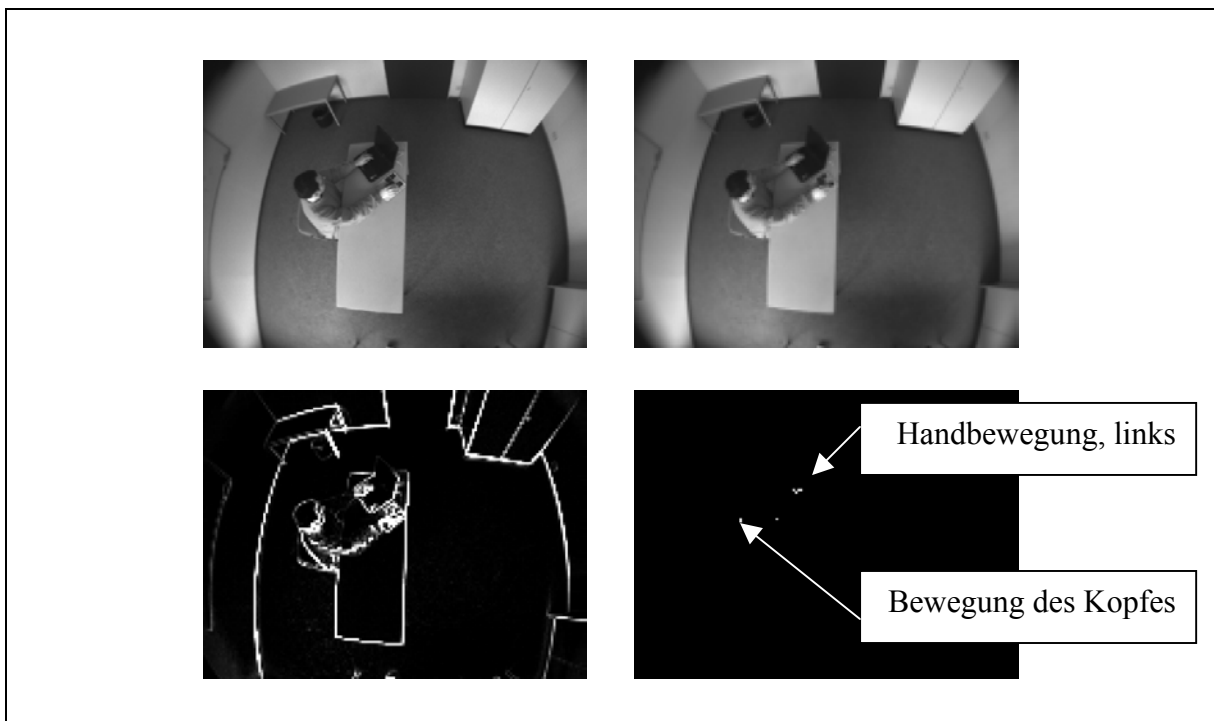


Abb. 4.13: Interpolation des ursprünglichen Grauwertbildes mit bereichsweise konstanten Funktionen und Anwendung des t-Tests zur Erkennung von feinmotorischen Bewegungen für die Szene nach Abb. 4.2. Originalbild (oben links), interpoliertes Bild (oben rechts), Varianz (unten links) und mit dem Schwellwert $S^2 = 3.030^2$ bewertete Grauwertänderungen (unten rechts). Fenstergröße 4 x 4 Pixel.

Abb. 4.12 und Abb. 4.13 zeigen die Ergebnisse des t-Tests bei der Auswertung der Szenen nach Abb. 4.1 und Abb. 4.2. Für eine Fenstergröße von 4x4 Pixel wurde in Tab. 4.1 der zum Vertrauensniveau $\alpha = 0.005$ gehörende Schwellwert $S = 3.030$ ermittelt.

Wie sich bereits anhand der Ergebnisse aus Abschnitt 4.4 vermuten ließ, liefert der t-Test in Verbindung mit der bereichsweise konstanten Grauwertinterpolation nur geringfügig bessere Ergebnisse als die Schwellenauswertung des Differenzbildes in 4.2. Bei der Varianzberechnung wirkt sich wiederum der systematische Fehler, insbesondere an Grauwertkanten, entsprechend stark aus. Im weiteren Verlauf wird daher der F-Test in Verbindung mit einer höheren Interpolationsgüte für die Bewegungserkennung bevorzugt.

4.8 Anwendung des F-Tests

In diesem Abschnitt soll zunächst der F-Test in Verbindung mit der linearen Grauwertinterpolation abgeleitet werden, bevor im zweiten Teil der Übergang auf quadratische Interpolationsprofile erfolgt. Als Vorlage dient wiederum [Lit. 4.4]. Der F-Test wird allgemein zur Prüfung der Hypothese, dass die Varianzen zweier Grundmengen identisch sind, angewendet. Dabei wird vorausgesetzt, dass beide Grundmengen normalverteilt sind [Lit. 4.5].

Zur Prüfung der letztgenannten Voraussetzung werden wieder die Grauwerte g_{i1} und g_{i2} mit $i \in [1, I]$ der Grauwertmengen G_{w1} und G_{w2} bzw. \mathbf{g}_1 und \mathbf{g}_2 betrachtet. Nach Abschnitt 4.7 sind die Grauwerte g_{i1} und g_{i2} Zufallsvariablen, die durch erwartungswertfreies, normalverteiltes Rauschen gestört sind. Die Varianz σ^2 beider Teilmengen \mathbf{g}_1 und \mathbf{g}_2 soll gleich sein und außerdem unveränderlich über der Zeit. Wie bereits erläutert wurde, sind dann die Mittelwerte f_{01} und f_{02} und auch die Steigungen f_{x1} , f_{x2} , f_{y1} und f_{y2} jeweils normalverteilte Zufallsvariablen mit den in (Gl. 4.69), (Gl. 4.70) und (Gl. 4.71) angegebenen Varianzen [Lit. 4.4].

$$(Gl. 4.69) \quad \sigma_{f_0}^2 = \frac{\sigma^2}{I} \quad (Gl. 4.70) \quad \sigma_{f_x}^2 = \frac{\sigma^2}{\sum_{i=1}^I x_i^2} \quad (Gl. 4.71) \quad \sigma_{f_y}^2 = \frac{\sigma^2}{\sum_{i=1}^I y_i^2}$$

Die Differenzen $f_{0\Delta} = f_{01} - f_{02}$, $f_{x\Delta} = f_{x1} - f_{x2}$ und $f_{y\Delta} = f_{y1} - f_{y2}$ sind somit ebenfalls normalverteilte Zufallsvariablen, deren Varianzen sich aufgrund der statistischen Unabhängigkeit aus der Summe der Einzelvarianzen zusammensetzen. Es gilt:

$$(Gl. 4.72) \quad \sigma_{\Delta 0}^2 = 2 \frac{\sigma^2}{I} \quad (Gl. 4.73) \quad \sigma_{\Delta x}^2 = \frac{2 \cdot \sigma^2}{\sum_{i=1}^I x_i^2} \quad (Gl. 4.74) \quad \sigma_{\Delta y}^2 = \frac{2 \cdot \sigma^2}{\sum_{i=1}^I y_i^2}$$

Außerdem sind die Zufallsvariablen $f_{0\Delta}$, $f_{x\Delta}$ und $f_{y\Delta}$ erwartungswertfrei, so dass die neue Zufallsvariable

$$(Gl. 4.75) \quad f_Z = \frac{f_{\Delta 0}^2}{\sigma_{\Delta 0}^2} + \frac{f_{\Delta x}^2}{\sigma_{\Delta x}^2} + \frac{f_{\Delta y}^2}{\sigma_{\Delta y}^2} = \frac{(f_{01} - f_{02})^2}{2 \sigma^2} I + \frac{(f_{x1} - f_{x2})^2}{2 \sigma^2} \sum_{i=1}^I x_i^2 + \frac{(f_{y1} - f_{y2})^2}{2 \sigma^2} \sum_{i=1}^I y_i^2$$

auch den Erwartungswert 0 aufweist. Darüber hinaus weist f_Z eine χ^2 -Verteilung mit insgesamt $F_Z = 3$ Freiheitsgraden auf.

Normiert man nach [Lit. 4.4] den Nenner von (Gl. 4.63) ebenfalls auf die Varianz σ^2 , ergibt sich die Zufallsvariable f_N , welche, wie die Summenterme selbst, χ^2 -verteilt ist. Die Zufallsvariable f_N weist insgesamt $F_N = I-3 + I-3 = 2(I-3)$ Freiheitsgrade auf.

$$(Gl. 4.76) \quad f_N = \frac{1}{\sigma^2} \sum_{i=1}^I (g_{i1} - f_{01} - f_{x1}x_i - f_{y1}y_i)^2 + \frac{1}{\sigma^2} \sum_{i=1}^I (g_{i2} - f_{02} - f_{x2}x_i - f_{y2}y_i)^2$$

Normiert man f_Z und f_N jeweils auf die Anzahl der zugehörigen Freiheitsgrade F_Z und F_N und bildet anschließend das Verhältnis F , ergibt sich eine F-verteilte Zufallsvariable:

$$(Gl. 4.77) \quad F = \frac{f_Z/F_Z}{f_N/F_N} = \frac{2(I-3) \left(\frac{(f_{01} - f_{02})^2}{2\sigma^2} I + \frac{(f_{x1} - f_{x2})^2}{2\sigma^2} \sum_{i=1}^I x_i^2 + \frac{(f_{y1} - f_{y2})^2}{2\sigma^2} \sum_{i=1}^I y_i^2 \right)}{\frac{3}{\sigma^2} \left(\sum_{i=1}^I (g_{i1} - f_{01} - f_{x1}x_i - f_{y1}y_i)^2 + \sum_{i=1}^I (g_{i2} - f_{02} - f_{x2}x_i - f_{y2}y_i)^2 \right)}$$

Für das Wahrscheinlichkeitsverhältnis nach (Gl. 4.63) folgt daraus (Gl. 4.78), die im weiteren Verlauf zur Hypothesenprüfung verwendet wird. Da es sich bei (Gl. 4.78) um eine monoton steigende Funktion der Variable F handelt, ist es ausreichend, F anstelle des gesamten Ausdrucks bezüglich der Überschreitung eines oberen Grenzwertes S zu testen. Aufgrund der einfacheren Berechnungsvorschrift wird (Gl. 4.79) zur Hypothesenprüfung verwendet. Zur Ableitung von (Gl. 4.79) wurden die Summen im Zähler von (Gl. 4.77) nach (Gl. 4.26) und (Gl. 4.27) eingesetzt.

$$(Gl. 4.78) \quad \left[\frac{L|_{H1}}{L|_{H0}} \right]_1^{1/I} = 1 + \frac{3}{2(I-3)} \cdot F$$

$$(Gl. 4.79) \quad \begin{aligned} I \frac{(I-3)}{3} \left((f_{01} - f_{02})^2 + \frac{(I-1)}{12} ((f_{x1} - f_{x2})^2 + (f_{y1} - f_{y2})^2) \right) &> \\ &\leq \\ S \cdot \left(\sum_{i=1}^I (g_{i1} - f_{01} - f_{x1}x_i - f_{y1}y_i)^2 + \sum_{i=1}^I (g_{i2} - f_{02} - f_{x2}x_i - f_{y2}y_i)^2 \right) &\Rightarrow \text{Hypothese H1} \\ &\Rightarrow \text{Hypothese H0} \end{aligned}$$

Der Schwellwert S wird anhand der tabellierten F-Verteilung in Abhängigkeit des Konfidenzniveaus α und in Abhängigkeit der Freiheitsgrade des Zählers und des Nenners ermittelt [Lit. 4.4]. Tab. 4.2 enthält eine Auswahl von Schwellwerten für die Konfidenzniveaus $\alpha = 0.05$ und $\alpha = 0.01$ nach [Lit. 4.5], wobei der für die lineare Grauwertinterpolation konstante Freiheitsgrad $F_Z = 3$ des Zählers berücksichtigt wurde. Der Freiheitsgrad $F_N = 2(I-3)$ des Nenners wird dagegen von der Pixelanzahl I im quadratischen Grauwertfenster der Größe $P \times Q$ bestimmt (s. Tab. 4.2).

In Verbindung mit quadratischen Interpolationsfunktionen ergibt sich, wie später gezeigt wird, der Zählerfreiheitsgrad $F_Z = 6$, während sich die Zahl der Freiheitsgrade des Nenners zu $F_N = 2(I-6)$ ergibt. Die entsprechenden Werte der F-Verteilung für die Vertrauensniveaus $\alpha = 0.05$ und $\alpha = 0.01$ wurden in Tab. 4.2 ebenfalls berücksichtigt.

F-TEST		Lineare Interpolation		Quadratische Interpolation	
	α	0.05 ($F_Z = 3$)	0.01 ($F_Z = 3$)	0.05 ($F_Z = 6$)	0.01 ($F_Z = 6$)
Pixelanzahl					
4	(2x2)	19.16	99.17	---	---
9	(3x3)	3.49	5.95	4.28	8.47
16	(4x4)	2.98	4.64	2.60	3.87
25	(5x5)	2.82	4.26	2.35	3.32
∞		2.60	3.78	2.10	2.80

Tab. 4.2: Obere 5-prozentige und obere einprozentige Werte der F-Verteilung [Lit. 4.5] in Abhängigkeit des Zählerfreiheitsgrades F_Z , des Nennerfreiheitsgrades F_N und des Vertrauensniveaus α . $F_Z = 3$ gilt für die lineare Grauwertinterpolation, während $F_Z = 6$ bei quadratischen Interpolationsprofilen zugrundegelegt wird

In Verbindung mit der linearen Grauwertinterpolation in einem Grauwertfenster mit 4x4 Pixel wurde der Schwellwert $S = 4.64$ für das Konfidenzniveau $\alpha = 0.01$ gewählt. Abb. 4.14 und Abb. 4.15 zeigen die Bewertung der grob- und feinmotorischen Bewegungen nach Abb. 4.1 und Abb. 4.2. Im Vergleich zur konstanten Grauwertinterpolation ergibt sich eine deutlich höhere Empfindlichkeit.

Neben bewegungsbedingten Grauwertänderungen treten allerdings auch vereinzelte Rauschstörungen auf, die sich durch die Wahl eines höheren Konfidenzniveaus unterdrücken lassen. Diese Vorgehensweise reduziert zwar auch die Empfindlichkeit der Bewegungsdetektion, allerdings sind die Auswirkungen bezüglich der Störungsunterdrückung deutlich stärker. Experimentell wurde ermittelt, dass bei einer Verdoppelung des Schwellwertes in den zugrundeliegenden Testdaten keine Rauschstörungen mehr vorhanden sind.

Eine Alternative zur Schwellwertanhebung ist die Elimination der isoliert auftretenden Störungen durch Nachbarschaftsoperationen. Fordert man, dass für ein valides Detektionsereignis Grauwertänderungen in mindestens 3 unmittelbar benachbarten Regionen vorhanden sein müssen, werden die Störungen auch ohne Schwellwertanhebung vollständig beseitigt.

Zum Vergleich mit der linearen Grauwertinterpolation soll im folgenden zusätzlich die Bewegungserkennung mit quadratischen Interpolationsprofilen beschrieben werden. In [Lit. 4.4] wurden die betreffenden Gleichungen in Vektor-Matrix-Form abgeleitet. An dieser Stelle werden die Summendarstellungen bevorzugt, da sie einen besseren Überblick hinsichtlich des erforderlichen Berechnungsaufwands bieten.

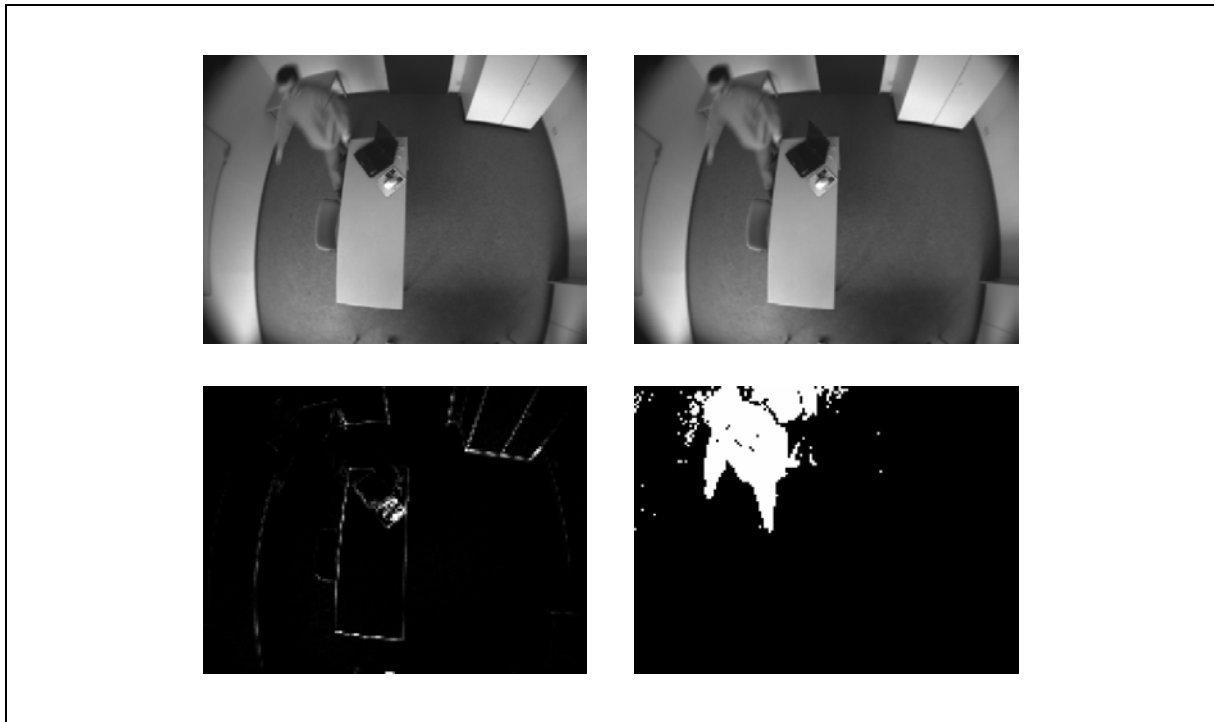


Abb. 4.14: Interpolation des ursprünglichen Grauwertbildes mit linearen Funktionen und Anwendung des F-Tests zur Erkennung von grobmotorischen Bewegungen für die Szene nach Abb. 4.1. Originalbild (oben links), interpoliertes Bild (oben rechts), Varianz (unten links) und mit dem Schwellwert $S = 4.64$ bewertete Grauwertänderungen (unten rechts). Fenstergröße 4×4 Pixel.

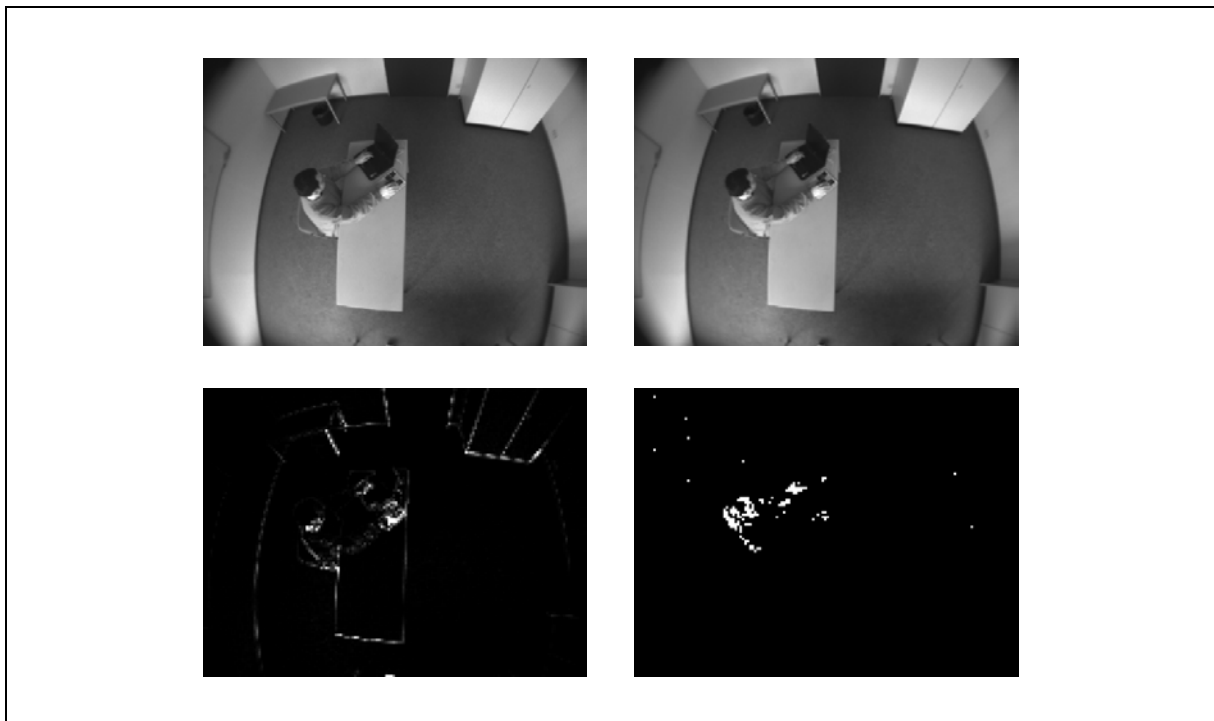


Abb. 4.15: Interpolation des ursprünglichen Grauwertbildes mit linearen Funktionen und Anwendung des F-Tests zur Erkennung von feinmotorischen Bewegungen für die Szene nach Abb. 4.2. Originalbild (oben links), interpoliertes Bild (oben rechts), Varianz (unten links) und mit dem Schwellwert $S = 4.64$ bewertete Grauwertänderungen (unten rechts). Fenstergröße 4×4 Pixel.

In (Gl. 4.61) wurde bereits die zur quadratischen Grauwertinterpolation verwendete Funktion $f(x_w, y_w)$ angegeben, die auf das Likelihood-Verhältnis nach (Gl. 4.64) führte. Ausgehend vom Quadrat des resultierenden Interpolationsfehlers lassen sich die Parameter $f_0, f_x, f_y, f_{xx}, f_{yy}$ und f_{xy} , beispielsweise mit dem Least-Square-Verfahren bestimmen:

$$(Gl. 4.80) \quad f_0 = \frac{7 \cdot I - 13}{2 \cdot I(I-4)} \sum_{i=1}^I g_i - \frac{15}{I(I-4)} \left(\sum_{i=1}^I g_i x_i^2 + \sum_{i=1}^I g_i y_i^2 \right)$$

$$(Gl. 4.81) \quad f_x = \frac{12}{I(I-1)} \sum_{i=1}^I g_i x_i \quad (Gl. 4.82) \quad f_y = \frac{12}{I(I-1)} \sum_{i=1}^I g_i y_i$$

$$(Gl. 4.83) \quad f_{xx} = \frac{180}{I(I-1)(I-4)} \sum_{i=1}^I g_i x_i^2 - \frac{15}{I(I-4)} \sum_{i=1}^I g_i$$

$$(Gl. 4.84) \quad f_{yy} = \frac{180}{I(I-1)(I-4)} \sum_{i=1}^I g_i y_i^2 - \frac{15}{I(I-4)} \sum_{i=1}^I g_i$$

$$(Gl. 4.85) \quad f_{xy} = \frac{144}{I(I-1)^2} \sum_{i=1}^I g_i x_i y_i$$

Zur Ableitung von (Gl. 4.80) und (Gl. 4.85) wurde die Summen nach (Gl. 4.26) und (Gl. 4.27) sowie die Beziehungen

$$(Gl. 4.86) \quad \sum_{i=1}^I x_i^4 = \frac{I(I-1)}{12} \cdot \frac{3I-7}{20} \quad (Gl. 4.87) \quad \sum_{i=1}^I y_i^4 = \frac{I(I-1)}{12} \cdot \frac{3I-7}{20}$$

$$(Gl. 4.88) \quad \sum_{i=1}^I x_i^2 y_i^2 = I \cdot \left(\frac{I-1}{12} \right)^2$$

verwendet. Die zu den zeitlich aufeinander folgenden Teilbilder gehörenden Parameter werden im weiteren Verlauf anhand der Indizes "1" und "2" unterschieden. Analog zur Vorgehensweise bei der Anwendung des F-Tests für lineare Interpolationsprofile erhält man unter Beachtung von (Gl. 4.64) die χ^2 -verteilte Zufallsvariable f_Z mit $F_Z = 6$ Freiheitsgraden.

$$(Gl. 4.89) \quad f_Z = \frac{(f_{01} - f_{02})^2}{2 \sigma^2} I + \frac{(f_{x1} - f_{x2})^2}{2 \sigma^2} \sum_{i=1}^I x_i^2 + \frac{(f_{y1} - f_{y2})^2}{2 \sigma^2} \sum_{i=1}^I y_i^2 + \dots$$

$$\frac{(f_{xx1} - f_{xx2})^2}{2 \sigma^2} \sum_{i=1}^I x_i^4 + \frac{(f_{yy1} - f_{yy2})^2}{2 \sigma^2} \sum_{i=1}^I y_i^4 + \frac{(f_{xy1} - f_{xy2})^2}{2 \sigma^2} \sum_{i=1}^I x_i^2 y_i^2$$

Im Nenner f_N ergibt sich mit

$$(Gl. 4.90) \quad f_N = \sum_{i=1}^I (g_{i1} - f_{01} - f_{x1} x_i - f_{y1} y_i - f_{xx1} x_i^2 - f_{yy1} y_i^2 - f_{xy1} x_i y_i)^2 + \dots$$

$$+ \sum_{i=1}^I (g_{i2} - f_{02} - f_{x2} x_i - f_{y2} y_i - f_{xx2} x_i^2 - f_{yy2} y_i^2 - f_{xy2} x_i y_i)^2$$

eine ebenfalls χ^2 -verteilte Zufallsvariable f_N mit $F_Z = 2 \cdot (I-6)$ Freiheitsgraden. Normiert man f_Z und f_N jeweils auf die Anzahl der zugehörigen Freiheitsgrade F_Z und F_N und bildet anschließend das Verhältnis F , ergibt sich wiederum eine F-verteilte Variable F , mit der sich das Maximum-Likelihood-Verhältnis nach (Gl. 4.64) folgendermaßen darstellen lässt:

$$(Gl. 4.91) \quad \left[\frac{L|_{H1}}{L|_{H0}} \right]_1^{1/I} = 1 + \frac{6}{2(I-6)} \cdot \frac{f_Z/6}{f_N/(2(I-6))} = 1 + \frac{3}{I-6} \cdot F$$

Wegen der Monotonie ist es erneut ausreichend, die Zufallsvariable F hinsichtlich der Überschreitung des Schwellwerts S zu prüfen. Ausgewertet wird dabei die Beziehung

(Gl. 4.92)

$$I \frac{(I-6)}{72} \left(12 \cdot (f_{01} - f_{02})^2 + (I-1) \cdot \left((f_{x1} - f_{x2})^2 + (f_{y1} - f_{y2})^2 + \frac{I-1}{12} \cdot (f_{xy1} - f_{xy2})^2 + \dots \right) \right) \begin{array}{l} > \\ \leq \end{array} \begin{array}{l} H1 \\ H0 \end{array}$$

$$S \cdot \left(\sum_{i=1}^I (g_{i1} - f_{01} - f_{x1}x_i - f_{y1}y_i - f_{xx1}x_i^2 - f_{yy1}y_i^2 - f_{xy1}x_iy_i)^2 + \dots \right)$$

$$\dots \sum_{i=1}^I (g_{i2} - f_{02} - f_{x2}x_i - f_{y2}y_i - f_{xx2}x_i^2 - f_{yy2}y_i^2 - f_{xy2}x_iy_i)^2$$

Abb. 4.16 und Abb. 4.17 zeigen die Bewertungsergebnisse der Szenen nach Abb. 4.1 und Abb. 4.2 für dieses Auswertefahren. Man erkennt unmittelbar, dass die Bewegungserkennung nochmals empfindlicher geworden ist. Allerdings haben auch die Störungen durch Rauschen deutlich zugenommen. Mit den in Tab. 4.2 angegebenen Konfidenzniveaus konnte keine vollständige Rauschunterdrückung erzielt werden. Es zeigte sich, dass erst ab dem Schwellwert $S = 3 \cdot 3.87 = 11.61$ eine vollständige Störungsunterdrückung in Verbindung mit den vorliegenden Testdaten möglich ist, was einem deutlich höheren Vertrauenswert entspricht. Die in Abb. 4.16 und Abb. 4.17 gezeigten Ergebnisse wurden mit dem Schwellwert $S = 2 \cdot 3.87 = 7.74$ ermittelt. Aufgrund des isolierten Auftretens der Rauschstörungen bietet es sich auch hier an, eine Mindestbereichsgröße vorzugeben, ab der eine Bewegung erkannt werden soll.

Obwohl das Verfahren noch empfindlicher auf die bewegungsbedingte Grauwertänderungen reagiert, wird für die Realisierung des Bildwächters die lineare Grauwertinterpolation mit anschließendem F-Test bevorzugt. Als Rechtfertigung für diese Wahl mag der Umstand dienen, dass die Verbesserung des Detektionsverhaltens beim Übergang von der konstanten zur linearen Interpolation deutlich stärker ausfiel als beim Übergang vom linearen zum quadratischen Profilverlauf. Entsprechende Ergebnisse sind auch in [Lit. 4.4] beschrieben.

Außerdem sollte beachtet werden, dass die quadratische Grauwertinterpolation im Vergleich zu den vorher beschriebenen Verfahren mit einem deutlich höheren Rechenaufwand verbunden ist, welcher den Leistungsverbrauch des Bildwächters in entscheidendem Maße beeinflusst. Bezüglich der Hardwareimplementierung des Verfahrens, die allerdings im Rahmen dieser Arbeit nicht mehr erfolgt, ist die lineare Grauwertinterpolation mit anschließendem F-Test ebenfalls deutlich weniger aufwendig.

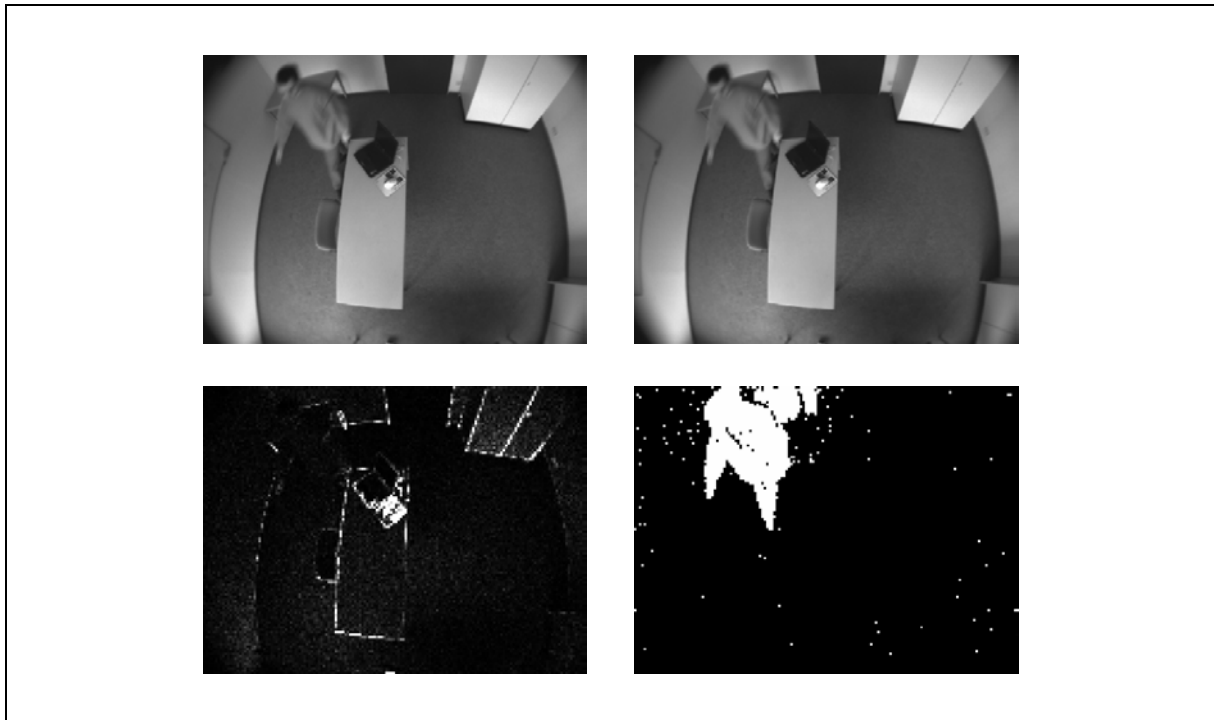


Abb. 4.16: Interpolation des ursprünglichen Grauwertbildes mit quadratischen Funktionen und Anwendung des F-Tests zur Erkennung von grobmotorischen Bewegungen für die Szene nach Abb. 4.1. Originalbild (oben links), interpoliertes Bild (oben rechts), Varianz (unten links) und mit dem Schwellwert $S = 7.74$ bewertete Grauwertänderungen (unten rechts). Fenstergröße 4 x 4 Pixel.

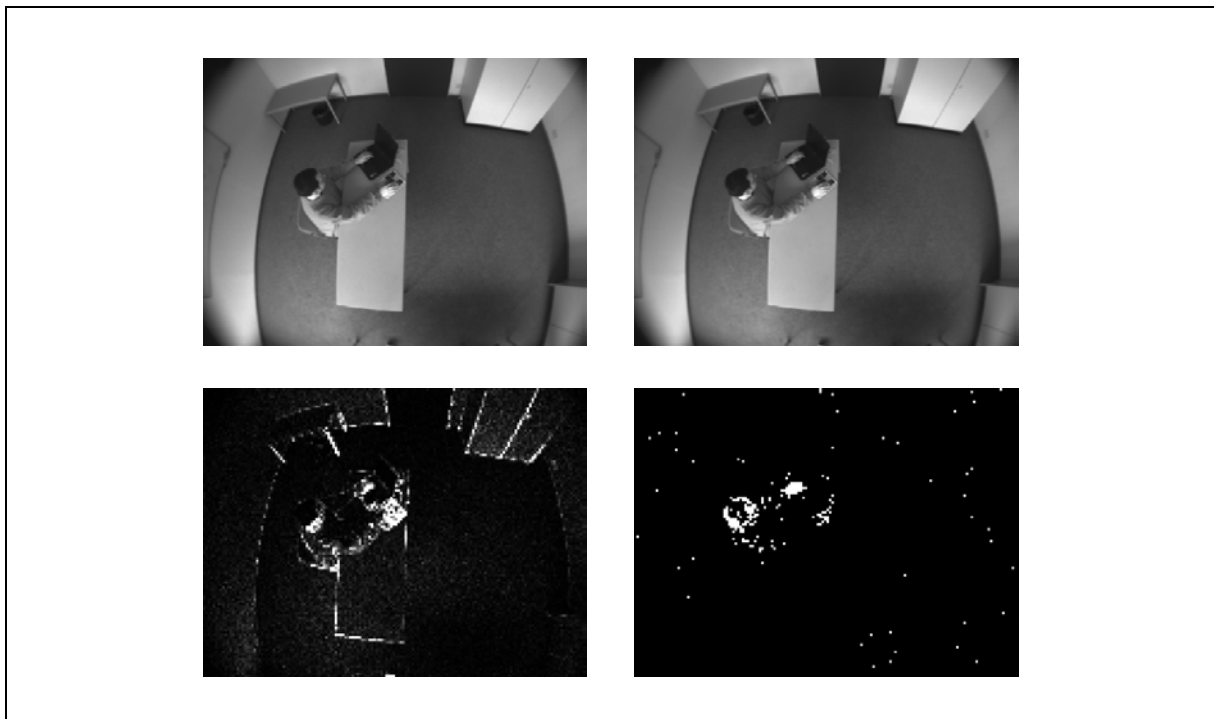


Abb. 4.17: Interpolation des ursprünglichen Grauwertbildes mit quadratischen Funktionen und Anwendung des F-Tests zur Erkennung von feinmotorischen Bewegungen für die Szene nach Abb. 4.2. Originalbild (oben links), interpoliertes Bild (oben rechts), Varianz (unten links) und mit dem Schwellwert $S = 23.87$ bewertete Grauwertänderungen (unten rechts). Fenstergröße 4 x 4 Pixel.

4.9 Literatur

- [Lit. 4.1] B. Jähne: "Digitale Bildverarbeitung", Springer Verlag Berlin Heidelberg, 1989
- [Lit. 4.2] P. Haberäcker: "Digitale Bildverarbeitung – Grundlagen und Anwendungen", Carl Hanser Verlag München Wien, 1991
- [Lit. 4.3] H.-H. Nagel: "Analyse und Interpretation von Bildfolgen", Informatik Spektrum 8/1985, S. 178-200, S. 312-327
- [Lit. 4.4] Y. Z. Hsu, H.-H. Nagel, G. Rekers: "New Likelihood Test Methods for Change Detection in Image Sequences", Computer Vision, Graphics and Image Processing 26/1984, pp. 73-106
- [Lit. 4.5] I. N. Bronstein, K. A. Semendjajew: "Taschenbuch der Mathematik", Teubner Verlagsgesellschaft Leipzig und Verlag Nauka, Moskau, 1979, Lizenzausgabe für den Verlag Harry Deutsch, Thum 1985

5 Komponenten von PIR-Bewegungsmeldern

5.1 Pyroelektrische Detektoren

5.1.1 Pyroelektrische Festkörper und pyroelektrisches Verhalten

Aufbauend auf den strahlungsphysikalischen Grundlagen nach Kap. 3 wird im folgenden das Funktionsprinzip und das Übertragungsverhalten von pyroelektrischen Sensoren beschrieben. Eine ausführlichere und weitergehende Beschreibung ist in [Lit. 5.3] und [Lit. 5.4] zu finden, an die sich die unten aufgeführte Zusammenfassung stark anlehnt.

Grundsätzlich lässt sich feststellen, dass ein pyroelektrischer Festkörper ein Dielektrikum ist, das ohne Einwirkung äußerer mechanischer Kräfte oder elektrischer Felder ein permanentes Dipolmoment aufweist, wie in Abb. 5.1 angedeutet wird. Der Betrag dieses Dipolmoments ist abhängig von der Temperatur des Dielektrikums. Mit steigender Temperatur verkleinert sich das Dipolmoment, bis die Polarisation oberhalb der Curie-Temperatur vollständig verschwindet.

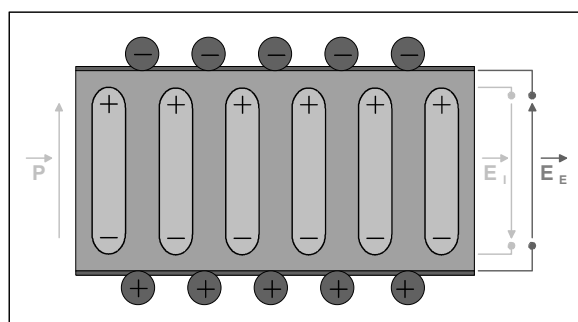


Abb. 5.1: Pyroelektrischer Festkörper im thermodynamischen Gleichgewicht

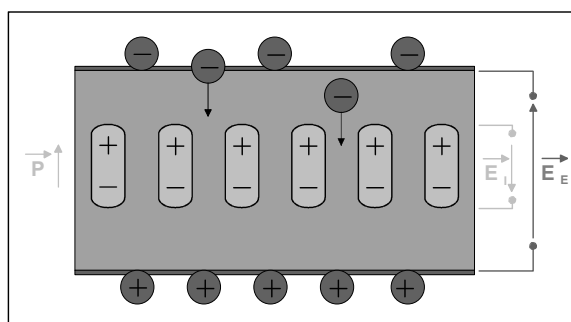


Abb. 5.2: Entstehung einer Differenzspannung infolge einer Temperaturerhöhung im Dielektrikum

Das infolge der Polarisation P vorhandene innere elektrische Feld E_I wird durch freie Oberflächenladungen aus der Umgebung kompensiert, so dass der Festkörper im thermodynamischen Gleichgewicht nach außen hin elektrisch neutral erscheint. Abb. 5.1 verdeutlicht den Compensationseffekt. Ändert sich das Dipolmoment infolge einer Temperaturänderung, erfolgt der Zufluss oder Abfluss von Oberflächenladungen nur langsam, da das Dielektrikum einen hohen Isolationswiderstand besitzt. Demzufolge tritt eine extern messbare Differenzspannung auf, die sich durch den geringen Isolationsstrom nur langsam abbaut und die ein Maß für die absolute Temperaturänderung ist (Abb. 5.2).

Die Temperaturänderung im Dielektrikum erfolgt mit einer gewissen Verzögerung, da sich die Materialtemperatur nicht sprunghaft ändern kann. Die Geschwindigkeit der Temperaturänderung wird von der Wärmekapazität des pyroelektrischen Materials, von der Wärmeableitung an das Substrat und von der Wärmeabstrahlung an die Umgebung bestimmt, wie später gezeigt wird. Die elektrischen Ausgleichsvorgänge unterliegen ebenfalls einer gewissen

Trägheit, die auch von der äußeren Beschaltung des Pyroelements abhängt. Zur Beschreibung des Übertragungsverhaltens eines pyroelektrischen Sensors muss demzufolge die thermische und die elektrische Zeitabhängigkeit beachtet werden.

Ferroelektrika wie Triglycinsulfat (TGS) und Lithiumtantalat (LiTaO_3) zeigen den gerade beschriebenen pyroelektrischen Effekt besonders ausgeprägt. Diese Stoffe besitzen ein ausgeprägtes Hystereseverhalten (s. Abb. 5.3). Die ferroelektrische Substanz befindet sich im Sättigungszustand P_S , wenn alle Domänen mit spontaner Polarisation gleich ausgerichtet sind, was durch das Anlegen eines äußeren elektrischen Feldes während des Produktionsprozesses erreicht werden kann.

Anschließend ist auch bei fehlendem äußeren Feld eine remanente Polarisation P_R vorhanden. Bei Temperaturen oberhalb der Curie-Temperatur verschwindet die spontane Polarisation und die Hysteresekurve entartet zu einer Geraden. Im undotierten Material ist der Prozess nach Überschreitung der Curie-Temperatur nicht reversibel.

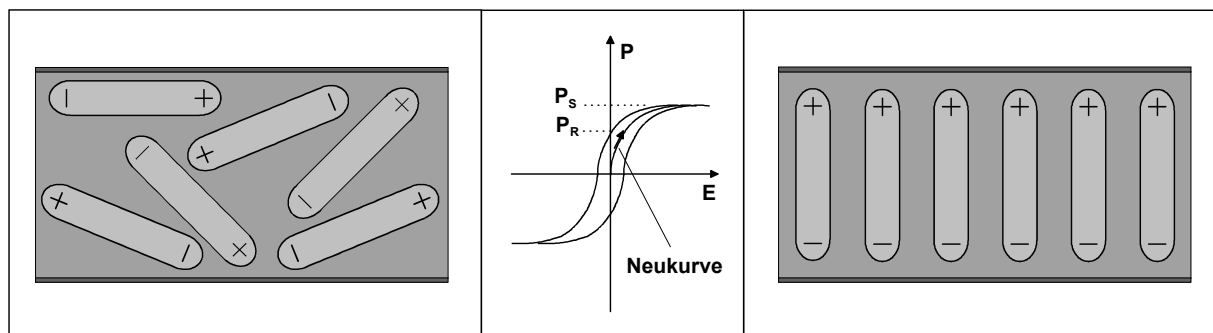


Abb. 5.3: Ungeordnete Ausrichtung der Domänen und Polarisation durch Anlegen eines äußeren elektrischen Feldes während des Produktionsprozesses (Neukurve).

5.1.2 Ferroelektrizität, Polarisation und pyroelektrischer Effekt

Zur Erklärung des ferroelektrischen Verhaltens geht man von der modellhaften Vorstellung aus, dass innerhalb des Dielektrikums „Elementarzellen“ vorhanden sind, in denen die Schwerpunkte von positiven und negativen Ladungen bereits ohne äußeres Feld gegeneinander verschoben sind. Ersetzt man diese „Punktladungswolken“ durch „Ersatzpunktladungen“ adäquater Größe in den Schwerpunkten der betreffenden Ladungsanhäufungen, können die Elementarzellen vereinfachend als kleine, permanente elektrische Dipole betrachtet werden.

Oberhalb der Curie-Temperatur verlieren die Elementarzellen durch Ionen- und Molekülverschiebungen ihren Dipolcharakter, unterhalb der kritischen Temperatur kommt es durch gegenseitige elektrische Wechselwirkung auch ohne äußeres Feld zu einer spontanen Parallelausrichtung der Dipole in einzelnen Teilbereichen. Die Teilbereiche mit gleicher Ausrichtung der Elementardipole werden als Domänen bezeichnet, bei der beschriebenen spontanen Ausrichtung der Dipole innerhalb einer Domäne spricht man von „spontaner Polarisation“.

Ohne externes elektrisches Feld erfolgt die Aufspaltung in Domänen aus energetischen Gründen in der Form, dass das resultierende Dipolmoment der gesamten Probe gleich Null wird. Ist beim Übergang von der ungeordneten zur geordneten Phase dagegen ein äußeres elektrisches Feld vorhanden, erfolgt eine Neuorientierung der Domänen. In diesem Fall erhält man eine Probe, deren resultierendes Dipolmoment \vec{m} von Null verschieden ist und die ein ausgeprägtes Hystereseverhalten der Polarisation in Abhängigkeit der anliegenden äußeren Feldstärke aufweist (s. Abb. 5.3). Das Dipolmoment \vec{m} ist ein Vektor, der in Richtung der Verbindungslinie von der negativen zur positiven Ladung zeigt und dessen Betrag dem Produkt aus Ladungsmenge q und Abstand d zwischen den Ladungen entspricht.

Normiert man das resultierende Dipolmoment $\vec{m} = m \cdot \vec{e}_m$ auf das Volumen $V = a^2 \cdot d$ einer quaderförmigen Probe, erhält man das mittlere Dipolmoment pro Volumeneinheit. Der Quotient ist ebenfalls eine gerichtete Größe und wird als Polarisationsvektor \vec{P} der elektrischen Polarisation bezeichnet. Er zeigt in Richtung des Dipolmoments der Polarisationsladungen und ist demzufolge genau entgegengesetzt zum elektrischen Feldstärkevektor zwischen den Ladungen gerichtet.

$$(Gl. 5.1) \quad \vec{P} = \frac{\vec{m}}{V} = \frac{q \cdot d}{a^2 \cdot d} \cdot \vec{e}_m = \frac{q}{a^2} \cdot \vec{e}_m = \sigma_a \cdot \vec{e}_m$$

Wie bereits vorher beschrieben wurde, influenziert die Polarisation des Dielektrikums Oberflächenladungen, deren elektrisches Feld dem der Polarisationsladungen entgegengesetzt ist. Das elektrische Feld der Oberflächenladungen kompensiert gerade das innere elektrische Feld der Polarisationsladungen.

Kontaktiert man nach Abb. 5.4 gegenüberliegende Flächen des Dielektrikums mit metallischen Elektroden, ändert sich die grundsätzliche Funktionalität nicht. Die Kompensationsladungen bilden sich lediglich an der Grenzfläche im Metalls aus. Aufgrund seines mechanischen Aufbaus kann ein pyroelektrischer Sensor als polarisierter Plattenkondensator betrachtet werden, wobei der Betrag der Polarisation von der Temperatur des Dielektrikums abhängt. Abb. 5.5 zeigt den prinzipiellen Verlauf der Polarisation $P = |\vec{P}|$ über der Temperatur.

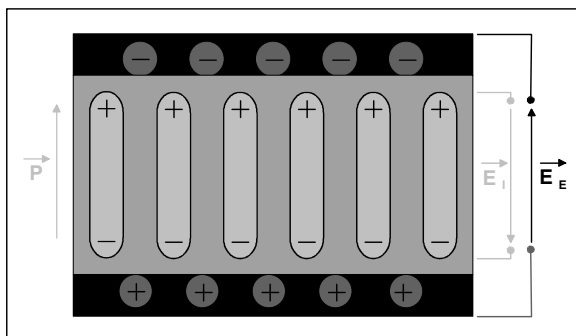


Abb. 5.4: Kontaktierung des pyroelektrischen Dielektrikums mit aufgedampften Metallelektroden

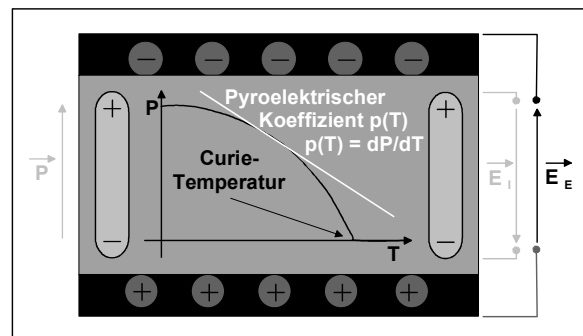


Abb. 5.5: Polarisation als Funktion der Temperatur und Definition des pyroelektrischen Koeffizienten

Der Betrag des Polarisationsvektors nach (Gl. 5.1) gibt die Flächendichte der Kompensationsladungen auf der Oberfläche des Dielektrikums an, welche senkrecht zur Richtung der Polarisation steht. Berechnet man daraus die auf einer ebenen Oberfläche A vorhandene Ladung Q , wobei die Richtungen der Flächennormale \vec{e}_A und des Polarisationsvektor \vec{P} nicht übereinstimmen, darf nur der senkrecht zu \vec{P} projizierte Flächenanteil zur Ermittlung der Gesamtladung verwendet werden:

$$(Gl. 5.2) \quad Q = \sigma_a \cdot \vec{e}_m \cdot A \cdot \vec{e}_A = \sigma_a \cdot A \cdot \cos(\langle \vec{e}_m, \vec{e}_A \rangle)$$

Im folgenden wird vorausgesetzt, dass beide Richtungsvektoren parallel sind. Die Kondensatorplatten stehen sich im Abstand d gegenüber.

Verändert sich der Betrag P der Polarisation \vec{P} um den Wert dP infolge einer Temperaturänderung dT des Dielektrikums, folgt daraus im stationären Zustand eine Änderung dQ der influenzierten Flächenladung um den Betrag

$$(Gl. 5.3) \quad \frac{dQ}{dT} = A \cdot \frac{d\sigma_a}{dT} = A \cdot \frac{dP}{dT} = A \cdot p(T) \quad \text{mit} \quad p(T) = \frac{dP}{dT}.$$

$p(T)$ wird als pyroelektrischer Koeffizient bezeichnet. Der Koeffizient beschreibt die Änderung der Polarisation in Abhängigkeit der Temperatur (s. Abb. 5.5). Ist der pyroelektrische Koeffizient eines ferroelektrischen Materials bekannt, kann die Ladung auf den Elektroden für kleine Änderungen dT der Temperatur T bestimmt werden.

Unter Beachtung der zeitlichen Abhängigkeit der Temperaturänderung dT/dt lässt sich der Vorgang der elektrischen Signalerzeugung in Abhängigkeit der thermischen Anregung ermitteln. Mit (Gl. 5.3) folgt

$$(Gl. 5.4) \quad i(t) = \frac{dQ(t)}{dt} = A \cdot p(T) \cdot \frac{dT(t)}{dt}.$$

Diese Gleichung beschreibt die thermisch-elektrische Kopplung beim pyroelektrischen Effekt. Aus (Gl. 5.4) folgt unmittelbar, dass mit einem pyroelektrischen Sensor nur Temperaturänderungen des Dielektrikums detektierbar sind. Polarisationsänderungen können durch Veränderungen der Temperatur in der unmittelbaren Sensorumgebung oder durch Absorption einer zeitveränderlichen Wärmestrahlung entstehen.

5.1.3 Thermisches Verhalten eines pyroelektrischen Sensors

Um die Ladungsänderungen in Abhängigkeit der zeitlichen Änderung der Temperatur $dT(t)/dt$ für konkrete Anwendungsfälle zu berechnen, muss zunächst die Temperaturänderung des Detektormaterials als Funktion der einfallenden Wärmestrahlung beschrieben werden. Unter Beachtung des Prinzips der Energieerhaltung lässt sich, wie im folgenden gezeigt wird, eine Differentialgleichung zur Beschreibung des thermischen Verhaltens ableiten. Ausgehend von der zeitlich variierenden Leistung $\Phi(t)$ der einfallenden Wärmestrahlung wird zunächst die

vom Detektormaterial absorbierte Strahlungsleistung $\Phi_a(t)$, also die Energie der Strahlung pro Zeiteinheit, bestimmt. Der Absorptionskoeffizient $\alpha(\lambda)$, der nach dem Kirchhoffschen Gesetz gleich dem Emissionskoeffizient $\varepsilon(\lambda)$ ist, beschreibt den grundlegenden Zusammenhang zwischen einfallender und absorbierter Strahlungsleistung. Der Einfachheit halber soll $\varepsilon(\lambda) = \varepsilon = \text{const.}$ gelten, so dass an dieser Stelle keine Abhängigkeit von der Wellenlänge λ der Strahlung auftritt.

$$(Gl. 5.5) \quad \Phi_a(t) = \varepsilon \cdot \Phi(t)$$

Ein Teil $\Phi_d(t)$ der absorbierten Strahlungsleistung bewirkt die Temperaturänderung des Dielektrikums, während die Verluste durch Wärmeabstrahlung an die Umgebung durch $\Phi_u(t)$ und durch Wärmeableitung an das Substrat durch $\Phi_s(t)$ beschrieben werden.

Der kontinuierliche Energieaustausch mit der Umgebung lässt sich unter Beachtung des Stefan-Boltzmannschen-Gesetzes bestimmen. Mit der Temperaturdifferenz $\vartheta(t) = T(t) - T_u(t)$ zwischen dem Dielektrikum und der Umgebung ergibt sich die an die Umgebung abgestrahlte bzw. von dort aufgenommene Strahlungsleistung $\Phi_u(t)$ (Gl. 5.6). A ist die Fläche des Detektorelements. Die in (Gl. 5.6) angegebene Näherung gilt für kleine Temperaturdifferenzen.

$$(Gl. 5.6) \quad \Phi_u(t) = A(\varepsilon \cdot \sigma \cdot T^4 - \alpha \cdot \sigma \cdot T_u^4) = A \cdot \varepsilon \cdot \sigma \cdot (T^4 - T_u^4) \approx 4 \cdot A \cdot \varepsilon \cdot \sigma \cdot \vartheta(t) \cdot T_u^3$$

Mit dem Wärmeübergangskoeffizient g_{th} lässt sich die durch den Temperaturunterschied $\vartheta(t)$ an das Substrat abgeführte Wärmeleistung $\Phi_s(t)$ bestimmen:

$$(Gl. 5.7) \quad \Phi_s(t) = g_{th} \cdot \vartheta(t)$$

Addiert man (Gl. 5.6) und (Gl. 5.7), lässt sich der gesamte Energieverlust pro Zeiteinheit nach (Gl. 5.8) beschreiben. G_{th} bezeichnet man als "totalen Wärmeübergangskoeffizient".

$$(Gl. 5.8) \quad \Phi_u(t) + \Phi_s(t) = (4A \cdot \varepsilon \cdot \sigma \cdot T_u^3 + g_{th}) \cdot \vartheta(t) = G_{th} \cdot \vartheta(t)$$

Die Temperaturerhöhung ϑ eines Körpers der Dichte ρ mit dem Volumen V kann anhand der spezifischen Wärmekapazität und der pro Zeiteinheit zugeführten Wärmeenergie berechnet werden. Die spezifische Wärmekapazität c_{th} gibt dabei die Wärmemenge an, die einem Körper der Masse $m = \rho \cdot V$ zugeführt werden muss, um die Temperaturerhöhung ϑ zu erreichen. C_{th} entspricht der Wärmekapazität des Elements.

$$(Gl. 5.9) \quad \Phi_d(t) = \rho \cdot V \cdot c_{th} \cdot \frac{d\vartheta(t)}{dt} = C_{th} \cdot \frac{d\vartheta(t)}{dt}$$

Betrachtet man die Energiebilanz pro Zeiteinheit, erhält man unter Beachtung der Einzelbeiträge nach $\Phi_a(t) = \varepsilon \cdot \Phi(t) = \Phi_d(t) + \Phi_u(t) + \Phi_s(t)$ eine lineare Differentialgleichung 1. Ordnung mit konstanten Koeffizienten, die das thermische Verhalten charakterisiert.

$$(Gl. 5.10) \quad \mu \cdot \Phi(t) = \vartheta(t) + \tau_{th} \cdot \frac{d\vartheta(t)}{dt} \quad \text{mit} \quad \mu = \frac{\varepsilon}{G_{th}}; \quad \tau_{th} = \frac{C_{th}}{G_{th}}$$

τ_{th} ist die thermische Zeitkonstante und $\vartheta(t)$ beschreibt den zeitlichen Verlauf der Differenztemperatur zwischen dem Detektormaterial und der Umgebung. Der Übergang zwischen der thermischen und der elektrischen Ebene erfolgt durch (Gl. 5.4). In (Gl. 5.4) stellt $T(t)$ die absolute Temperatur dar, durch die Bildung der Ableitung kann $T(t)$ durch $\vartheta(t)$ ersetzt werden, da $\vartheta(t)$ die Änderungen von $T(t)$ relativ zur aktuellen Umgebungstemperatur angibt.

$$(Gl. 5.11) \quad i(t) = A \cdot p(T) \cdot \frac{d\vartheta(t)}{dt} \quad (\text{Pyroelektrischer Effekt})$$

Bildet man die Laplace-Transformierten $L\{\dots\}$ von (Gl. 5.10), folgt unter Anwendung des Differentiationssatzes

$$(Gl. 5.12) \quad \mu \cdot \Psi(s) = \Theta(s) + \tau_{th} \cdot s \cdot \Theta(s) - \vartheta(0) \quad \text{mit} \quad \Psi(s) = L\{\Phi(t)\}, \quad \Theta(s) = L\{\vartheta(t)\}$$

Ohne Beschränkung der Allgemeinheit kann man $\vartheta(t=0) = \vartheta(0) = 0$ setzen. Die Umformung von (Gl. 5.12) liefert dann die Übertragungsfunktion $H_{th}(s)$ des thermischen Systems mit der Temperaturdifferenz $\Theta(s)$ als Ausgangsgröße in Abhängigkeit der einfallenden Strahlungsleistung $\Psi(s)$.

$$(Gl. 5.13) \quad H_{th}(s) = \frac{\Theta(s)}{\Psi(s)} = \mu \cdot \frac{1}{1 + s \cdot \tau_{th}}$$

$H_{th}(s)$ ist die Übertragungsfunktion eines Verzögerungsgliedes 1. Ordnung mit der thermischen Zeitkonstanten τ_{th} . Analog der Übertragung eines Spannungsimpulses über einen Tiefpass 1. Ordnung folgt der Temperaturverlauf bei einer sprungförmigen Anregung einer Exponentialfunktion der Form $(1 - e^{-t/\tau_{th}})$.

Der Faktor $A \cdot p(T)$ in (Gl. 5.11) ist für den betrachteten Arbeitspunkt nahezu konstant. Führt man die Laplace-Transformation von (Gl. 5.11) unter Beachtung der Differentiationsregel mit $\vartheta(t=0) = \vartheta(0) = 0$ durch, folgt daraus die Übertragungsfunktion $H_{py}(s)$ zur Umwandlung von Temperaturänderungen in einen elektrischen Strom:

$$(Gl. 5.14) \quad H_{py}(s) = \frac{I(s)}{\Theta(s)} = A \cdot p(T) \cdot s \quad \text{mit} \quad I(s) = L\{i(t)\}$$

Aus (Gl. 5.13) und (Gl. 5.14) lässt sich die Übertragungsfunktion $H_{pyth}(s)$ der bisher betrachteten Teilsysteme ableiten:

$$(Gl. 5.15) \quad H_{pyth}(s) = H_{py}(s) \cdot H_{th}(s) = \frac{I(s)}{\Psi(s)} = A \cdot p(T) \cdot \mu \cdot \frac{s}{1 + s\tau_{th}}$$

5.1.4 Elektrisches Verhalten eines pyroelektrischen Sensors

Im vorherigen Abschnitt wurde das thermische Verhalten eines pyroelektrischen Sensors und der pyroelektrische Effekt systemtheoretisch beschrieben. Der durch Temperaturänderungen hervorgerufene Strom $i(t)$ wird als Isolationsstrom durch das unvollkommene Dielektrikum des Sensors abgeführt. Wie bereits erläutert wurde, kann ein pyroelektrischer Sensor als Plattenkondensator mit temperaturabhängigem Dielektrikum betrachtet werden, der die Kapazität

C_{py} und den ohmschen Leitwert G_{py} besitzt. Zur Impedanzwandlung ist dem Sensor meist ein Feldeffekttransistor nachgeschaltet, dessen Eingangskapazität C_{ft} und dessen Eingangsleitwert G_{ft} nach Abb. 5.6 parallel zur Kapazität C_{py} und dem Leitwert G_{py} des Sensors liegen.

Die Kapazitäten addieren sich zur Gesamtkapazität $C_{el} = C_{py} + C_{ft}$, für den Gesamtleitwert folgt analog $G_{el} = G_{py} + G_{ft}$.

Eine Elektrode des Sensors ist durchlässig für Wärmestrahlung, die andere ist unmittelbar mit dem Sensorgehäuse verbunden, das gleichzeitig als Wärmesenke dient.

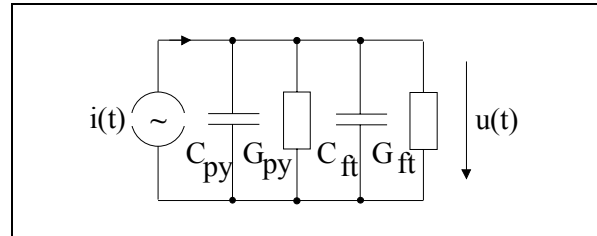


Abb. 5.6 Elektrisches Ersatzschaltbild eines pyroelektrischen Detektors mit Eingangsleitwert des Feldeffekttransistors

Die Ausgangsspannung $u(t)$ ist nur dann von Null verschieden, wenn ein Ausgleichsstrom infolge von Temperaturänderungen auftritt. Der Ausgleichsstrom fließt bei externer Beschaltung durchs Dielektrikum und über den Eingangswiderstand des Feldeffekttransistors. Unter Beachtung der Gesetze der Wechselstromlehre folgt

$$(Gl. 5.16) \quad u(t) = \frac{1}{G_{el} + j\omega C_{el}} i(t) = \frac{R_{el}}{1 + j\omega R_{el} C_{el}} i(t) = \frac{R_{el}}{1 + j\omega \tau_{el}} i(t)$$

mit $R_{el} = \frac{1}{G_{el}}$ und $\tau_{el} = R_{el} C_{el}$.

Aus (Gl. 5.16) folgt die Übertragungsfunktion $H_{el}(s)$ der elektrischen Seite, die ebenfalls ein Verzögerungsglied 1. Ordnung darstellt.

$$(Gl. 5.17) \quad H_{el}(s) = \frac{U(s)}{I(s)} = \frac{R_{el}}{1 + s \cdot \tau_{el}}$$

Vergleicht man die Laplace-Transformierte (Gl. 5.13) mit der zugehörigen Differentialgleichung (Gl. 5.10), läßt sich auch die zu (Gl. 5.17) gehörende Differentialgleichung unmittelbar angeben:

$$(Gl. 5.18) \quad R_{el} \cdot i(t) = u(t) + \tau_{el} \cdot \frac{du(t)}{dt} \quad \text{mit} \quad \tau_{el} = \frac{C_{el}}{G_{el}} = C_{el} \cdot R_{el}$$

(Gl. 5.18) beschreibt das zeitliche Verhalten der Spannung $u(t)$ am Ausgang eines RC-Tiefpasses mit der Zeitkonstante τ_{el} , der mit der Spannung $R_{el} \cdot i(t)$ angeregt wird. Die Wechselspannung $u(t)$ liegt am Eingang des Feldeffekttransistors an und steuert diesen um den eingestellten Gleichstromarbeitspunkt aus. Die Übertragung der Eingangsspannung $u(t)$ zum Ausgang $v(t)$ des Feldeffekttransistors ist von der Transistorbeschaltung abhängig. Im folgenden wird an dieser Stelle eine konstante Kleinsignalverstärkung V_u angesetzt, so dass an dieser Stelle die Übertragungsfunktion nach (Gl. 5.19) gilt.

$$(Gl. 5.19) \quad H_{uv}(s) = \frac{V(s)}{U(s)} = V_u$$

5.1.5 Übertragungsfunktion eines Einzelementdetektors

Multipliziert man die Teilübertragungsfunktionen nach (Gl. 5.19), (Gl. 5.17), (Gl. 5.14) und (Gl. 5.13), erhält man die Übertragungsfunktion (Gl. 5.20) eines pyroelektrischen Detektors:

$$H_s(s) = \frac{V(s)}{\Psi(s)} = H_{uv}(s) \cdot H_{el}(s) \cdot H_{pyth}(s) = H_{uv}(s) \cdot H_{el}(s) \cdot H_{py}(s) \cdot H_{th}(s)$$

$$(Gl. 5.20) \quad H_s(s) = \mu \cdot A \cdot p(T) \cdot R_{el} \cdot V_u \cdot \frac{s}{(1 + s\tau_{th}) \cdot (1 + s\tau_{el})}$$

Die Übertragungsfunktion beschreibt die Abhängigkeit der Ausgangsspannung des Sensors von der einfallenden Strahlungsleistung im Laplace-Bereich. Durch die Rücktransformation von (Gl. 5.20) in den Zeitbereich ergibt sich die Stoßantwort des Sensors:

$$(Gl. 5.21) \quad h_s(t) = \frac{\mu \cdot A \cdot p(T) \cdot R_{el} \cdot V_u}{\tau_{el} \cdot \tau_{th}} \cdot \left[\tau_{el} \cdot e^{-t/\tau_{th}} - \tau_{th} \cdot e^{-t/\tau_{el}} \right]$$

Das Amplitudenspektrum nach (Gl. 5.22) und das Phasenspektrum nach (Gl. 5.23) kann ebenfalls aus (Gl. 5.20) ermittelt werden, indem man die komplexe Variable s durch die komplexe Größe $j\omega = j2\pi f$ ersetzt:

$$(Gl. 5.22) \quad |H_s(\omega)| = K \cdot \frac{\omega}{\sqrt{1 + (\omega \tau_{th})^2} \cdot \sqrt{1 + (\omega \tau_{el})^2}} \quad \text{mit} \quad K = \mu \cdot A \cdot p(T) \cdot R_{el} \cdot V_u$$

$$(Gl. 5.23) \quad \varphi_s(\omega) = \tan^{-1} \left(\frac{1 - \omega^2 \tau_{th} \tau_{el}}{\omega \cdot (\tau_{th} + \tau_{el})} \right)$$

Durch die Nullstelle bei $f=0$ und die beiden Polstellen $f_1 = -1/(2\pi\tau_{th})$ und $f_2 = -1/(2\pi\tau_{el})$ ergibt sich eine Bandpasscharakteristik. Die größere Zeitkonstante bestimmt die Grenzfrequenz f_{HP} des Hochpasses. Beide Zeitkonstanten können anhand des Empfindlichkeitsdiagramms aus den Datenblättern des Detektors ermittelt werden.

Zur Simulation des Amplitudenspektrums (Abb. 5.7) und des Phasenspektrums (Abb. 5.8) wurden typische Grenzfrequenzen von $f_{HP}=0.05\text{Hz}$ und $f_{TP}=0.3\text{Hz}$ verwendet. Ordnet man die kleinere Grenzfrequenz dem thermischen Verhalten zu, folgen die Zeitkonstanten $\tau_{th}=3.18\text{s}$ und $\tau_{el}=0.53\text{s}$.

Der konstante Faktor in (Gl. 5.22) hat die Einheit $[V/W][s]$, wobei die Einheit $[s]$ gegen die Einheit $[Hz] = [1/s]$ des zeitabhängigen Teils der Übertragungsfunktion gekürzt werden kann. Der Faktor beschreibt das Verhältnis zwischen der Ausgangsspannung eines pyroelektrischen Sensors und der einfallenden Strahlungsleistung und ist somit ein Maß für die Detektorempfindlichkeit.

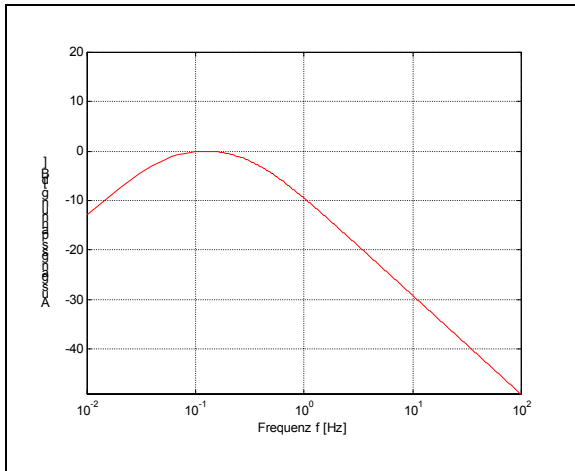


Abb. 5.7: Normiertes Amplitudenspektrum eines Einzelementdetektors

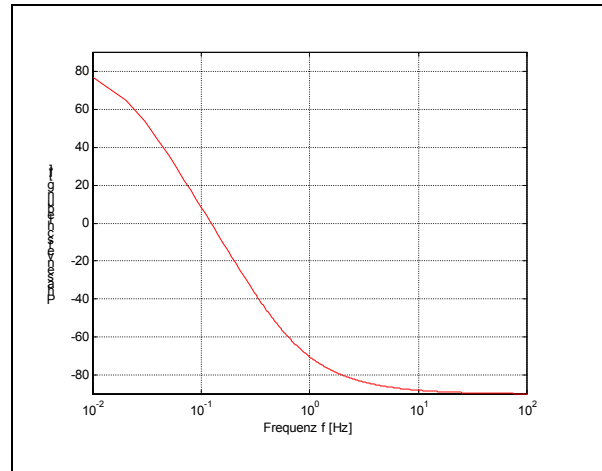


Abb. 5.8: Phasenverschiebung beim Einzelementdetektor

Die spektrale Empfindlichkeit $R_\lambda(f)$ oder die in den Datenblättern oftmals angegebene Empfindlichkeit $R(f) = R_T(f)$, die anhand der breitbandigen Abstrahlung eines schwarzen Körpers mit der Temperatur T ermittelt wird, ist im Vergleich zu (Gl. 5.22) als Verhältnis der Effektivwerte der Ausgangsspannung zur einfallenden Strahlungsleistung definiert.

$$(Gl. 5.24) \quad R(f) = \frac{U_{\text{rms}}(f)}{\Phi_{\text{rms}}}$$

Im des Transmissionsbereichs des Eintrittsfensters ist die Empfindlichkeit nahezu unabhängig von der Wellenlänge λ und damit auch von der Temperatur T der Strahlungsquelle [Lit. 5.3].

5.1.6 Rauschquellen bei pyroelektrischen Detektoren

Als Rauschquelle tritt bei pyroelektrischen Detektoren vorwiegend das Temperaturrauschen, das dielektrische Rauschen im pyroelektrischen Material, das Widerstandsrauschen am Parallelwiderstand und das Verstärkerrauschen des integrierten Feldeffekttransistors auf. Abb. 5.9 verdeutlicht die einzelnen Rauschbeiträge, die in diesem Abschnitt näher beschrieben werden sollen.

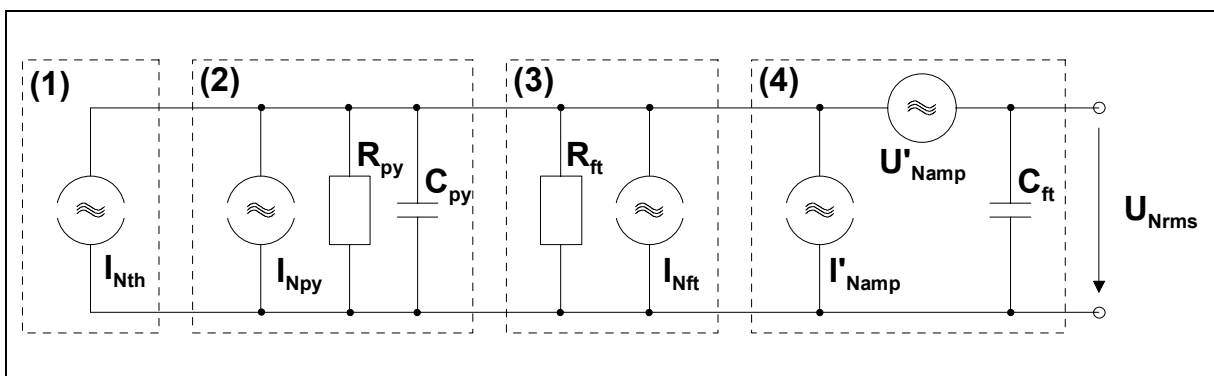


Abb. 5.9: Ersatzschaltbild zur Entstehung der Rauschspannung $U_{N\text{rms}}$ aus dem Temperaturrauschen $I_{N\text{th}}$ (1), dem dielektrischen Rauschen $I_{N\text{py}}$ (2), dem Widerstandsrauschen $I_{N\text{ft}}$ (3) und dem Verstärkerrauschen $I'_{N\text{amp}}$ und $U'_{N\text{amp}}$ (4) [Lit. 5.3]

Die in Abb. 5.9 dargestellten Rauschstromquellen erzeugen an dem Gesamtscheinleitwert

$$(Gl. 5.25) \quad Y = |\underline{Y}| = \sqrt{\left(\frac{1}{R_{py}} + \frac{1}{R_{ft}}\right)^2 + \omega^2 \cdot (C_{py} + C_{ft})^2}$$

voneinander unabhängige Rauschspannungsbeiträge, die sich quadratisch zum Effektivwert

$$(Gl. 5.26) \quad U_{Nrms} = \sqrt{U_{Nth}^2 + U_{Npy}^2 + U_{Nft}^2 + U_{Namp}^2}$$

der Ausgangsrauschspannung überlagern.

5.1.6.1 Temperaturrauschen

Das Temperaturrauschen ist auf Temperaturfluktuationen im Detektormaterial zurückzuführen, die infolge des Wärmeaustauschs mit der Umgebung (Wärmestrahlung) und mit dem Substrat (Wärmeableitung) entstehen. Aufgrund dieser Schwankungen tritt im Detektor eine Rauschleistung P_{Nth} nach (Gl. 5.27) auf. G_{th} ist der totale Wärmeübergangskoeffizient nach (Gl. 5.8), B ist die Bandbreite der Anordnung [Lit. 5.3].

$$(Gl. 5.27) \quad P_{Nth} = \sqrt{4kT^2 \cdot G_{th} \cdot B}$$

Die einfallende Strahlungsleistung muss eine Temperaturänderung oberhalb der vom Temperaturrauschen vorgegebenen Schwankungen hervorrufen, damit sie detektiert werden kann. Mit der Empfindlichkeit $R(f)$ des Detektors nach (Gl. 5.24) kann P_{Nth} in eine effektive Rauschspannung von

$$(Gl. 5.28) \quad U_{Nth} = \frac{I_{Nth}}{Y} = \frac{R(f)}{\varepsilon} \cdot \sqrt{4kT^2 \cdot G_{th} \cdot B}$$

umgerechnet werden. Der Faktor $1/\varepsilon$ berücksichtigt den Zusammenhang zwischen der einfallenden Strahlungsleistung und dem zur Temperaturänderung beitragenden Anteil.

5.1.6.2 Dielektrisches Rauschen

Ursache des dielektrischen Rauschens ist die thermische Bewegung von Ladungsträgern in einem elektrisch leitenden Material (thermisches Rauschen). Zur Berechnung des Rauschstromes wird der pyroelektrische Detektor als ein verlustbehafteter Kondensator mit der Kapazität C_{py} und dem Verlustwinkel δ betrachtet. Für kleine Verlustwinkel δ kann der reale Kondensator durch einen idealen Kondensator mit der Kapazität $C_p = C_{py} \cdot \cos(\delta) \approx C_{py}$ und dem Parallelwiderstand $R_p = R_{py} = (\omega C_{py} \cdot \tan(\delta))^{-1}$ nach Abb. 5.9 ersetzt werden. Nach [Lit. 5.4] ergibt sich mit dem Parallelwiderstand R_{py} ein effektiver Rauschstrom von

$$(Gl. 5.29) \quad I_{Npy} = Y \cdot U_{Npy} = \sqrt{4kT \cdot \frac{B}{R_{py}}} = \sqrt{4kT \cdot B \cdot \omega C_{py} \cdot \tan \delta} \quad .$$

5.1.6.3 Widerstandsrauschen

Das thermische Rauschen am Eingangswiderstand des Feldeffekttransistors liefert ebenfalls einen Beitrag zum Detektorrauschen. Analog zu (Gl. 5.29) folgt für den Effektivwert des Rauschstromes

$$(Gl. 5.30) \quad I_{Nft} = Y \cdot U_{Nft} = \sqrt{4kT \cdot B / R_{ft}}.$$

5.1.6.4 Verstärkerrauschen

Das Verstärkerrauschen wird nach Abb. 5.9 durch eine Stromquelle I'_{Namp} und eine Spannungsquelle U'_{Namp} am Verstärkereingang modelliert. Mit dem Scheinleitwert

$$(Gl. 5.31) \quad Y' = |Y'| = \sqrt{\left(\frac{1}{R_{py}} + \frac{1}{R_{ft}}\right)^2 + \omega^2 \cdot C_{py}^2}$$

lässt sich die Spannungsquelle in eine äquivalente Stromquelle $Y' \cdot U'_{Namp}$ umformen, so dass für den Gesamt rauschstrom des Verstärkers

$$(Gl. 5.32) \quad I_{Namp} = Y \cdot U_{Namp} = \sqrt{I'_{Namp}{}^2 + Y'^2 \cdot U'_{Namp}{}^2}$$

folgt. Bei bekannten Detektorparametern kann die Rauschspannung U_{Nrms} nach (Gl. 5.26) aus den einzelnen Rauschanteilen nach (Gl. 5.28) bis (Gl. 5.32) berechnet werden.

5.1.6.5 Mikrophonie

Neben den in den Abschnitten 5.1.6.1, 5.1.6.2, 5.1.6.3 und 5.1.6.4 beschriebenen Rauschquellen ändert sich die Polarisation von pyroelektrischen Detektoren nicht nur bei Temperaturwechseln, sondern auch bei Druckschwankungen (Piezoelektrischer Effekt). Bei mechanischen Vibrationen oder bei akustischen Signalen entstehen am Detektorausgang zusätzliche Rauschanteile, die in (Gl. 5.26) nicht berücksichtigt wurden.

Eine weitere Störquelle ist die optische Mikrophonie. Bewegt sich der Detektor aufgrund mechanischer Schwingungen in einem räumlich nicht konstanten Strahlungsfeld, erfolgt eine Modulation der einfallenden Strahlung.

Zur Kompensation von Mikrophonieeffekten sind in konventionellen pyroelektrischen Detektoren meist zwei Sensorelemente antiparallel oder in Serie geschaltet. Mit dieser Maßnahme lassen sich weiterhin unerwünschte Ausgangssignale durch global wirkende Hintergrundstörungen, die auf beiden Elementen auftreffen, weitgehend unterdrücken.

5.1.7 Kenngrößen von pyroelektrischen Detektoren

5.1.7.1 Rauschäquivalente Strahlungsleistung

Für praktische Anwendungen ist weniger die absolute Größe der Rauschspannung U_{Nrms} , sondern vielmehr das Signal-Rausch-Verhältnis $S/N = U_{rms}/U_{Nrms}$ interessant. Dieses Verhältnis wird bei pyroelektrischen Detektoren durch die Angabe der kleinsten, noch nachweisbaren Strahlungsleistung, dem NEP (Noise Equivalent Power), charakterisiert. Der NEP wird meist auf die Bandbreiteneinheit $\text{Hz}^{1/2}$ normiert und lässt sich mit der Empfindlichkeit $R(f)$ nach (Gl. 5.24) folgendermaßen berechnen [Lit. 5.3]:

$$(Gl. 5.33) \quad NEP(f) = \frac{U_{Nrms}}{R(f) \cdot \sqrt{B}}$$

Entspricht die einfallende Strahlungsleistung dem entnormierten NEP-Wert $\sqrt{B} \cdot NEP$, ergibt sich ein Signal-Rausch-Verhältnis $S/N = 1$.

5.1.7.2 Detektivität und spezifische Detektivität

Die Detektivität D entspricht dem Kehrwert des NEP. Sie dient als Gütemaß für das verwendete Detektormaterial. Je kleiner der NEP bzw. je größer die Detektivität ist, umso "besser" ist der Sensor. Einen objektiven Vergleich zwischen verschiedenen Sensoren ermöglicht die spezifische Detektivität D^* nach (Gl. 5.34). Die spezifische Detektivität ergibt sich aus dem auf die Flächeneinheit $A^{1/2}$ normierten NEP-Wert [Lit. 5.3].

$$(Gl. 5.34) \quad D^*(f) = \sqrt{A} \cdot D(f) = \frac{\sqrt{A}}{NEP(f)} = \frac{R(f) \cdot \sqrt{B} \cdot \sqrt{A}}{U_{Nrms}}$$

5.1.8 Übertragungsverhalten eines Doppелеlementensensors

Ausgehend von dem Modell eines Einzelelementdetektors wird in diesem Abschnitt das Systemverhalten eines pyroelektrischen Dualsensors mit parallelgeschalteten Sensoren betrachtet. Die thermische Kopplung zwischen beiden Einzelelementen über das Substrat sei vernachlässigbar, so dass im folgenden nur der Kompensationseffekt durch die gegenphasigen Ausgleichsströme berücksichtigt wird. Unter Beachtung der Gesetze der Wechselstromlehre kann die komplexe Ausgangsspannung $u(t)$ in Abhängigkeit der komplexen Ströme $i_1(t)$ und $i_2(t)$ berechnet werden (Gl. 5.35).

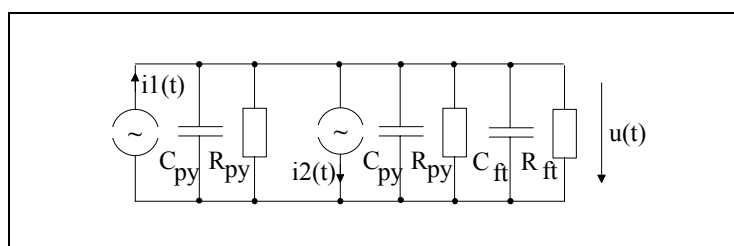


Abb. 5.10: Parallelschaltung von zwei pyroelektrischen Einzelelementen unter Beachtung des Eingangsleitwertes eines Feldeffekttransistors

$$(Gl. 5.35) \quad u(t) = u_{\text{par}}(t) = \frac{R_{\text{par}}}{1 + j\omega R_{\text{par}}(C_{\text{ft}} + 2C_{\text{py}})} [i_1(t) - i_2(t)]$$

$$\text{mit } R_{\text{par}} = \frac{R_{\text{py}}R_{\text{ft}}}{R_{\text{py}} + 2R_{\text{ft}}} = \frac{R_{\text{py}}}{2} \parallel R_{\text{ft}}$$

Der Vergleich von (Gl. 5.16) mit (Gl. 5.35) zeigt, dass sowohl der Einzelelementsensor als auch der Doppелеlementsensor durch ein Verzögerungsglied 1. Ordnung beschrieben werden können. Durch die gegenphasige Polung ist beim Doppелеlementsensor lediglich die Differenz der Einzelströme anzusetzen, während die grundsätzliche Form der Übertragungsfunktion nach (Gl. 5.20) unverändert erhalten bleibt. Der Verstärkungsfaktor und die Zeitkonstanten lassen sich aus dem in den Datenblättern angegebenen Empfindlichkeitsdiagramm $R(f)$ des jeweiligen Detektors ermitteln. In Analogie zu (Gl. 5.20) wird im folgenden der resultierende Gesamtwiderstand mit R_{el} und die wirksame Gesamtkapazität mit C_{el} bezeichnet.

Aus den vorangegangenen Kapiteln sind die Differentialgleichungen und die Übertragungsfunktionen der einzelnen Teilsysteme bekannt. Abb. 5.11 zeigt das gesamte Übertragungssystem zur Modellierung eines pyroelektrischen Doppелеlementensors.

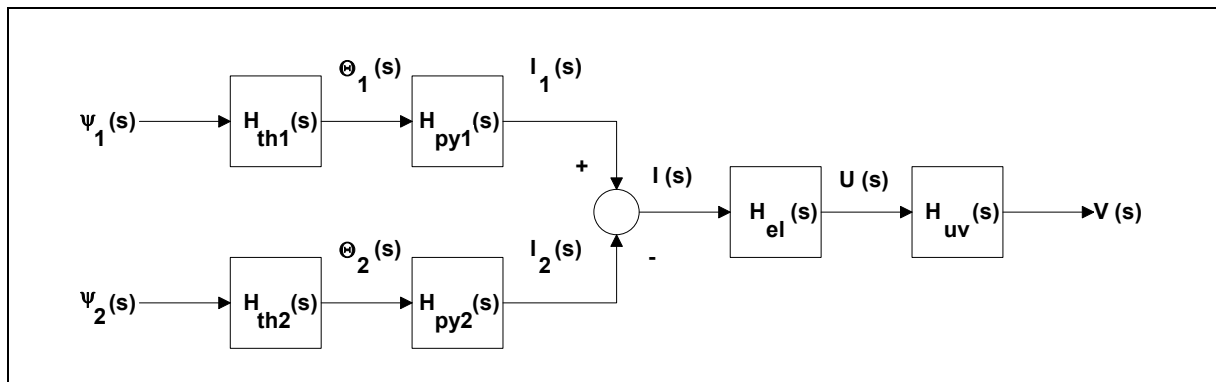


Abb. 5.11: Teilsysteme zur Ermittlung der Übertragungsfunktion $H_s(s)$ eines pyroelektrischen Doppелеlementensors

Zur Berechnung der Gesamtübertragungsfunktion $H_s(s)$ des Doppелеlements werden die Teilübertragungsfunktionen unter Beachtung des Summationspunktes für die Einzelströme miteinander multipliziert:

$$(Gl. 5.36) \quad H_s(s) = \frac{V(s)}{\Delta\Psi(s)} = K \cdot \frac{s}{(1 + s\tau_{\text{th}}) \cdot (1 + s\tau_{\text{el}})}$$

Zur Ermittlung von (Gl. 5.36) wurde (Gl. 5.20) mit der Konstante $K = \mu \cdot A \cdot p(T) \cdot R_{\text{el}} \cdot V_u$ nach (Gl. 5.22) verwendet. Ferner wurde vorausgesetzt, dass beide Sensorelemente identische Eigenschaften besitzen. In der Realität treten Abweichungen von dieser Annahme auf, die zu Fehlanpassungen führen. Sie bewirken eine unvollständige Kompensation der Ausgleichsströme bei globalen Störungen, die aber im folgenden nicht weiter berücksichtigt werden.

5.1.9 Simulationen zum Übertragungsverhalten eines Dualsensors

In diesem Abschnitt werden die Teilübertragungsfunktionen nach Abb. 5.11 verwendet, um das thermische und elektrische Verhalten eines pyroelektrischen Dualsensors zu simulieren. Die Zerlegung in einzelne Übertragungssysteme gestattet Einblicke in messtechnisch nicht zugängliche Teilsysteme des Detektors. Die Simulationen wurden mit dem Programm MATLAB unter Verwendung der SIGNAL PROCESSING TOOLBOX und des Programmpakets SIMULINK erstellt. Letzteres ermöglicht die grafische Eingabe der Funktionsblöcke und stellt Methoden zur Simulation zeitkontinuierlicher und zeitdiskreter Systeme bereit.

Für die nachfolgenden Simulationen wurden die Daten des pyroelektrischen Dualsensors Lhi 954 der Fa. Heimann verwendet (Abb. 5.14). In den Datenblättern ist der Verlauf der Sensorempfindlichkeit als Funktion der Frequenz angegeben, woraus die Grenzfrequenzen $f_{HP}=0.05\text{Hz}$ und $f_{TP}=0.3\text{Hz}$ ermittelt wurden. Ordnet man die kleinere Grenzfrequenz dem thermischen Verhalten zu, folgen die Zeitkonstanten $\tau_{th}=3.18\text{s}$ und $\tau_{el}=0.53\text{s}$. Mit der in den Datenblättern angegebenen Responsivity $R_{typ}(f=1\text{Hz}) = 3500 \text{ V/W}$ lässt sich die Konstante K aus (Gl. 5.36) bzw. (Gl. 5.22) bestimmen. Bei harmonischen Anregungen und unter Berücksichtigung der von den Signalformen abhängigen Effektivwerte (Abb. 5.12) gilt:

$$R(f) = \frac{U_{\text{rms}}(f)}{\Phi_{\text{rms}}(f)} = \frac{U(f)}{\sqrt{8}} \cdot \frac{\sqrt{8}}{\sqrt{3} \cdot \Phi} = \frac{|H(\omega)|_{f=1\text{Hz}}}{\sqrt{3}}$$

bzw.

(Gl. 5.37)

$$K = \frac{\sqrt{3}}{0.090} \cdot 3500 \text{ V/W} = 67112.433 \text{ V/W}$$

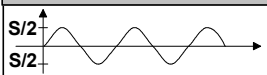
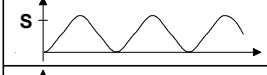
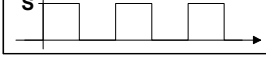
Signalverlauf über der Zeit		Effektivwert
	sin	$S_{\text{rms}} = S/\sqrt{8}$
	1- cos	$S_{\text{rms}} = S \cdot \sqrt{3/8}$
	rect	$S_{\text{rms}} = S/\sqrt{2}$

Abb. 5.12: Effektivwerte in Abhängigkeit der Signalform

Berechnet man das Amplitudenspektrum (Gl. 5.22) mit den Zeitkonstanten $\tau_{th}=3.18\text{s}$ und $\tau_{el}=0.53\text{s}$ und der Konstanten $K=67112.433 \text{ V/W}$, zeigt sich eine gute Übereinstimmung des simulierten Verlaufs nach Abb. 5.13 mit der vom Hersteller angegebenen Kurve in Abb. 5.14.

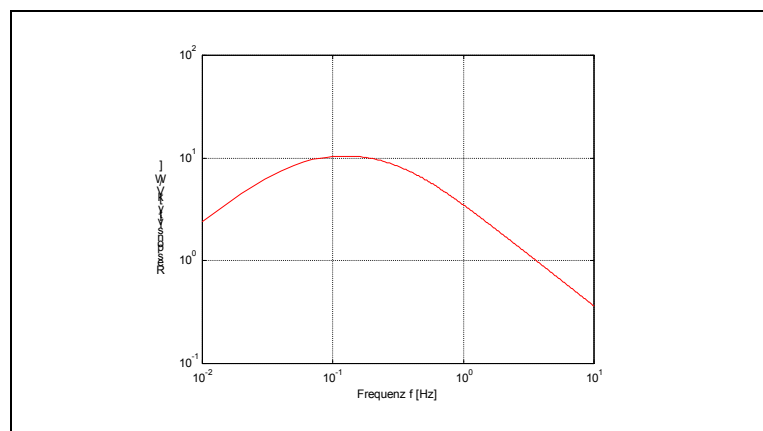


Abb. 5.13: Simulierter Empfindlichkeitsverlauf über der Frequenz, basierend auf den Herstellerangaben

Mit dem System aus Abb. 5.11 werden im folgenden unter Verwendung der Parameter des Lhi 954 typische Ausgangssignale bei unterschiedlichen Anregungen simuliert.

Abb. 5.15 zeigt das mit dem Programmpaket SIMULINK erstellte Blockschaltbild des Sensors sowie die Schnittstellen zur übergeordneten Software.

Am Eingang beider Sensorelemente ist jeweils eine Nutz- und eine Störsignalquelle vorgesehen. Zur Simulation des zeitlichen Verhaltens werden nur die Nutzsignale verwendet.

An Out1...Out14 stehen alle wichtigen Ausgangssignale zur Verfügung. Setzt man das Nutz- und das Störsignal eines Kanals auf Null, kann auch das zeitliche Verhalten eines einzelnen Sensorelements simuliert werden.

Abb. 5.16 zeigt das Übertragungsverhalten eines einzelnen Sensorelements bei einer Anregung mit kosinusförmig modulierter Strahlung. Zum Vergleich sind in Abb. 5.17 die einzelnen Signalverläufe bei rechteckförmiger Anregung darstellt. In Abb. 5.16 und Abb. 5.17 bezeichnen Out5 und Out6 die beiden Eingangssignale, wobei sich Out5 aus jeweils 5 Pulsen mit den dargestellten Pausen zusammensetzt.

Out11 ist die Differenz beider Eingangssignale, Out12 zeigt die Temperaturverläufe im Dielektrikum, Out13 die infolge des pyroelektrischen Effekts differenzierten Temperaturverläufe. Die Signale Out14 sind die Ausgangssignale des Sensors. Insbesondere in Abb. 5.17 erkennt man die frequenzbegrenzende Wirkung des Bandpasses bei schnellen Flanken.

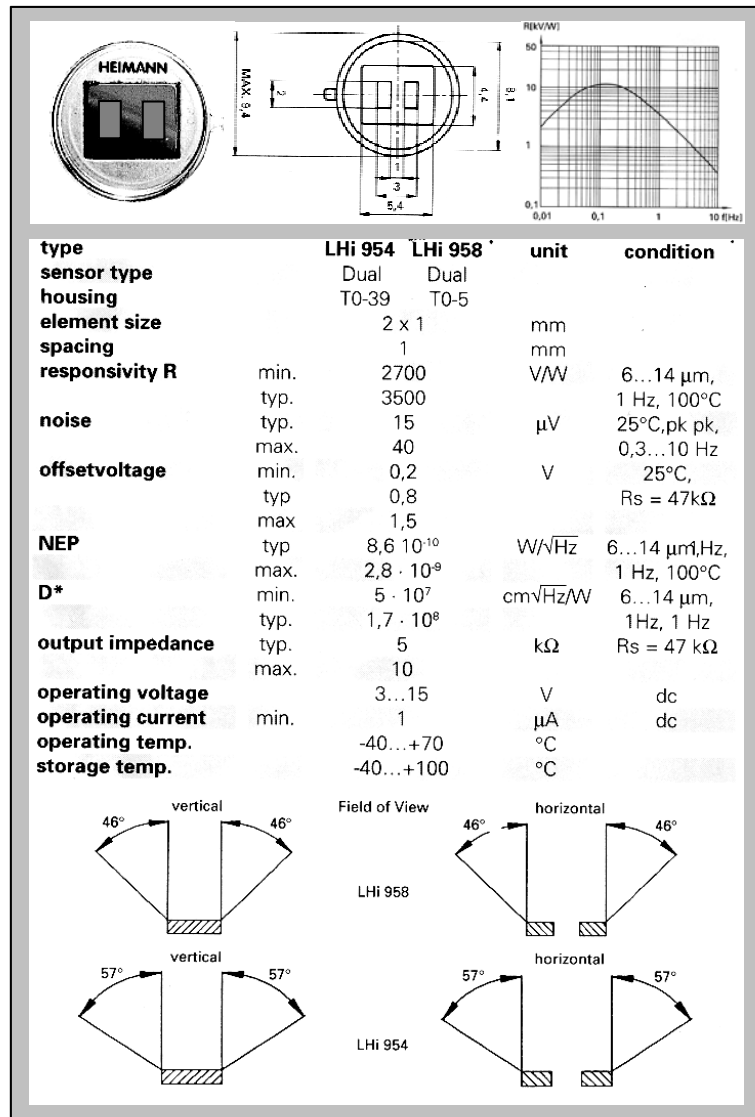


Abb. 5.14: Technische Daten des Doppelementensensors LHI 954 der Fa. HEIMANN

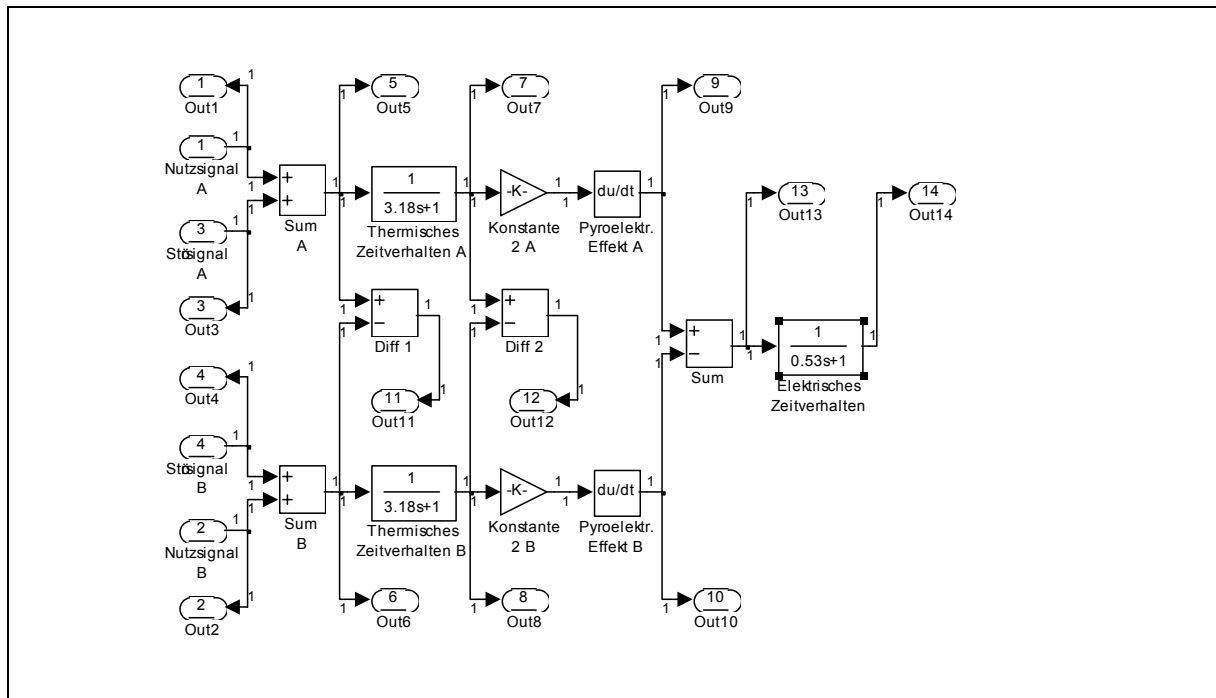


Abb. 5.15: SIMULINK-Blockschaltbild des pyroelektrischen Dualsensors Lhi954

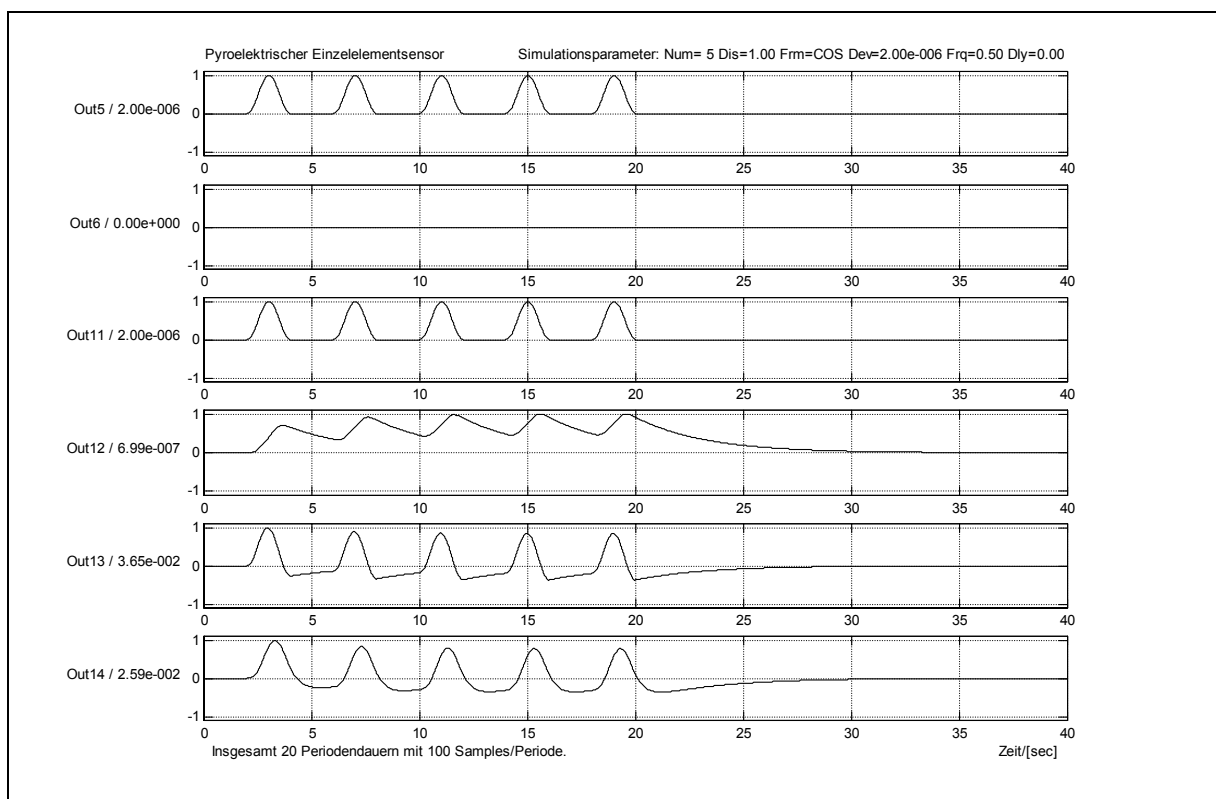


Abb. 5.16: Signale bei Anregung eines einzelnen Sensorelements mit kosinusförmig modulierter Strahlung

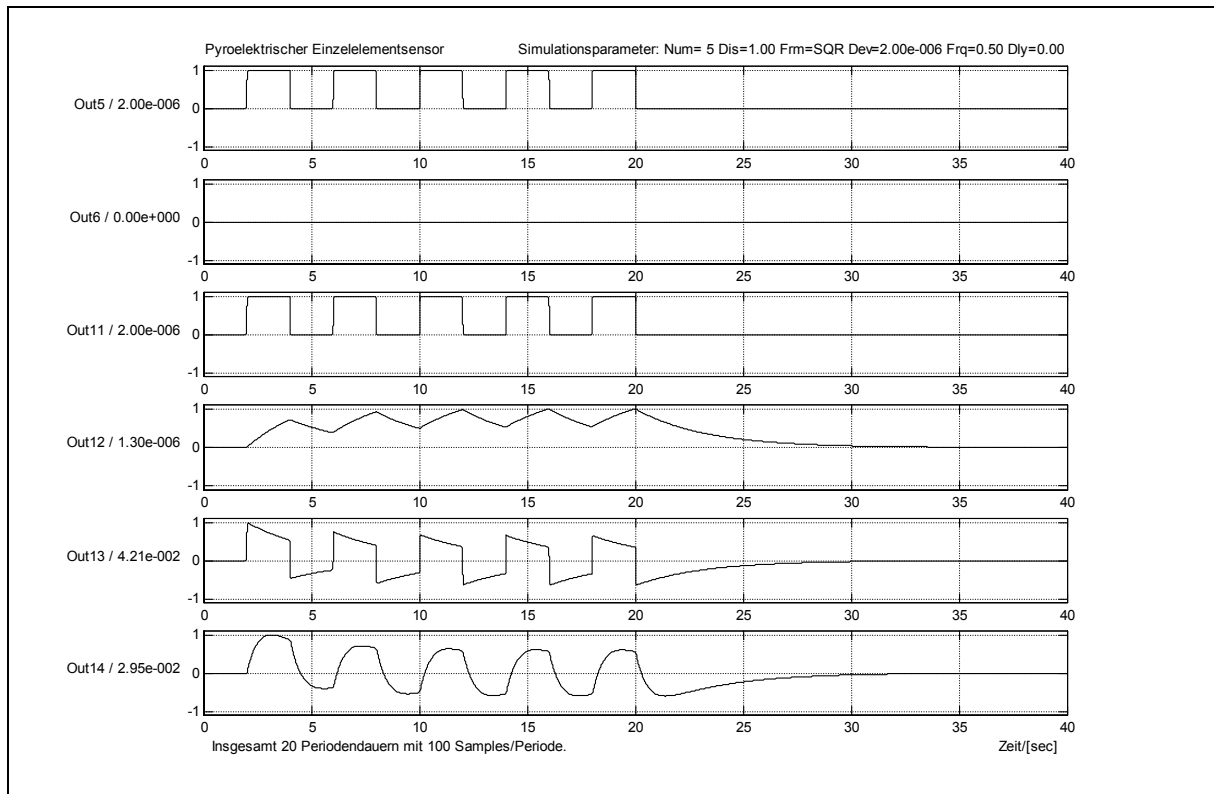


Abb. 5.17: Signale bei Anregung eines einzelnen Sensorelements mit rechteckförmig modulierter Strahlung

Abb. 5.18 und Abb. 5.19 zeigen das Verhalten eines Doppелеlementsensors, wenn beide Einzelsensoren nacheinander der harmonischen bzw. der rechteckförmigen Anregung ausgesetzt werden. Da nach Abb. 5.15 die Signale Out11, Out12 und Out13 die Signaldifferenzen zwischen beiden Einzelementen darstellen, ist in Abb. 5.18 und Abb. 5.19 nicht mehr die charakteristische Erwärmung der Dielektrika zu erkennen, sondern der Verlauf der Temperaturdifferenz zwischen beiden Sensorelementen.

Der zeitliche Versatz zwischen beiden Eingangssignalen des Doppелеlements und natürlich auch die Größe der Anregung hängt in der Realität von der Entfernung, der Gehrichtung, der Gehgeschwindigkeit etc. ab. Der starke zeitliche Versatz, der in Abb. 5.18 und Abb. 5.19 zugrundegelegt wurde, entspricht dabei eher den Verhältnissen beim Quergehen. Beim Draufzugehen ist diese Zeitverschiebung deutlich kleiner und je nach Gehrichtung können sich die Signale nahezu vollständig kompensieren.

Abb. 5.20 und Abb. 5.21 zeigen die Ausgangssignale eines Doppелеlementdetektors für beide Anregungsformen bei unterschiedlichen Signalfrequenzen (Simulationsparameter Frq). In den ersten beiden Diagrammen von Abb. 5.20 und Abb. 5.21 sind die Anregungen an beiden Systemeingängen dargestellt. Bei den Eingangssignalen wurde keine Zeitbasis angegeben, da diese sich aus den unterschiedlichen Zeitbasen der Ausgangssignale ergibt. Die Simulationen wurden jeweils für einen Zeitraum der 20-fachen Pulsdauer durchgeführt.

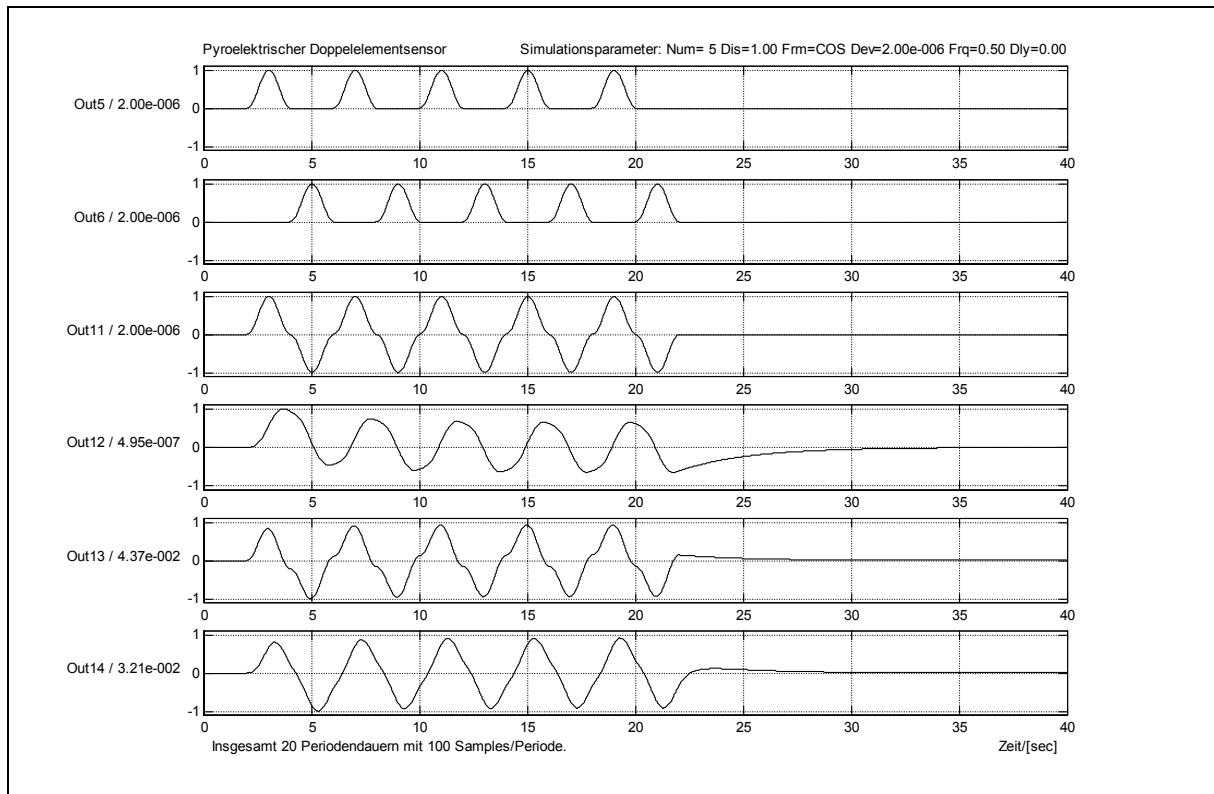


Abb. 5.18: Signale beim Doppелеlementdetektor bei Anregung mit kosinusförmigen, um eine Pulsdauer verzögerten Eingangsimpulsen auf beiden Kanälen

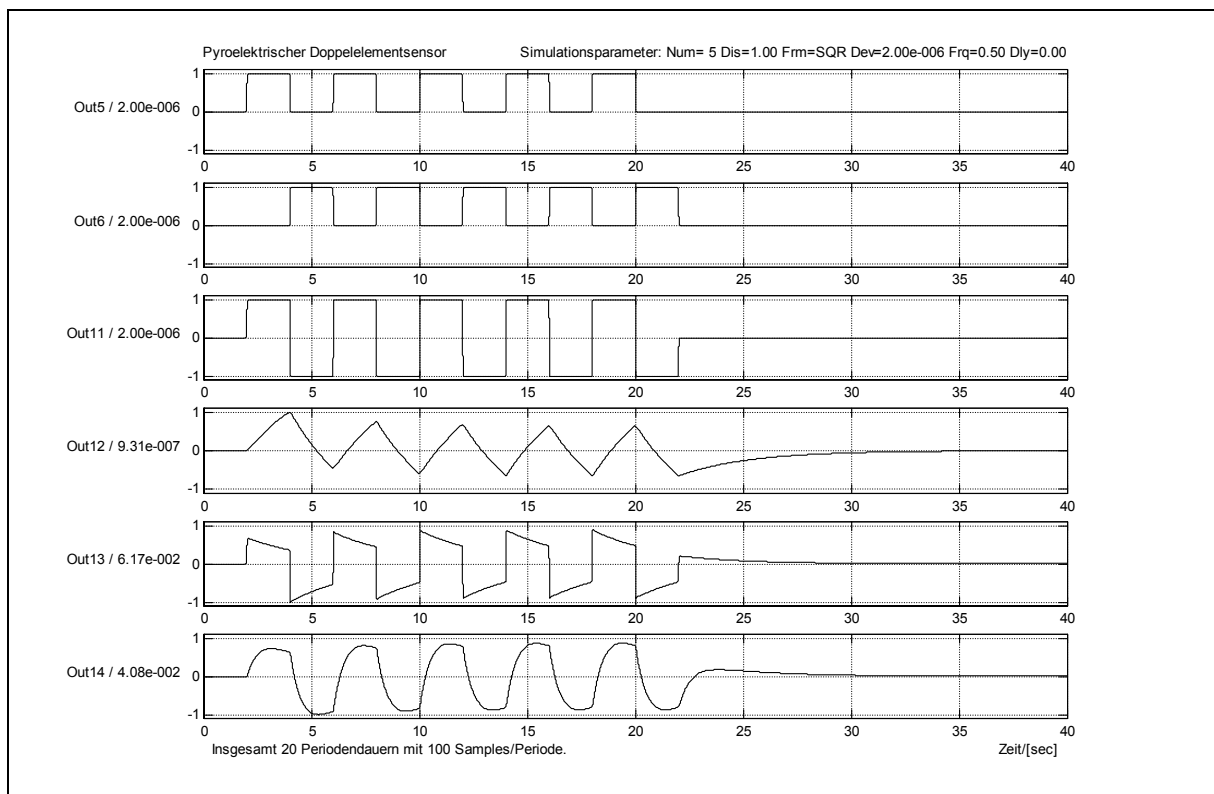


Abb. 5.19: Signale beim Doppелеlementdetektor bei Anregung mit rechteckförmigen, um genau eine Pulsdauer verzögerten Eingangsimpulsen auf beiden Kanälen

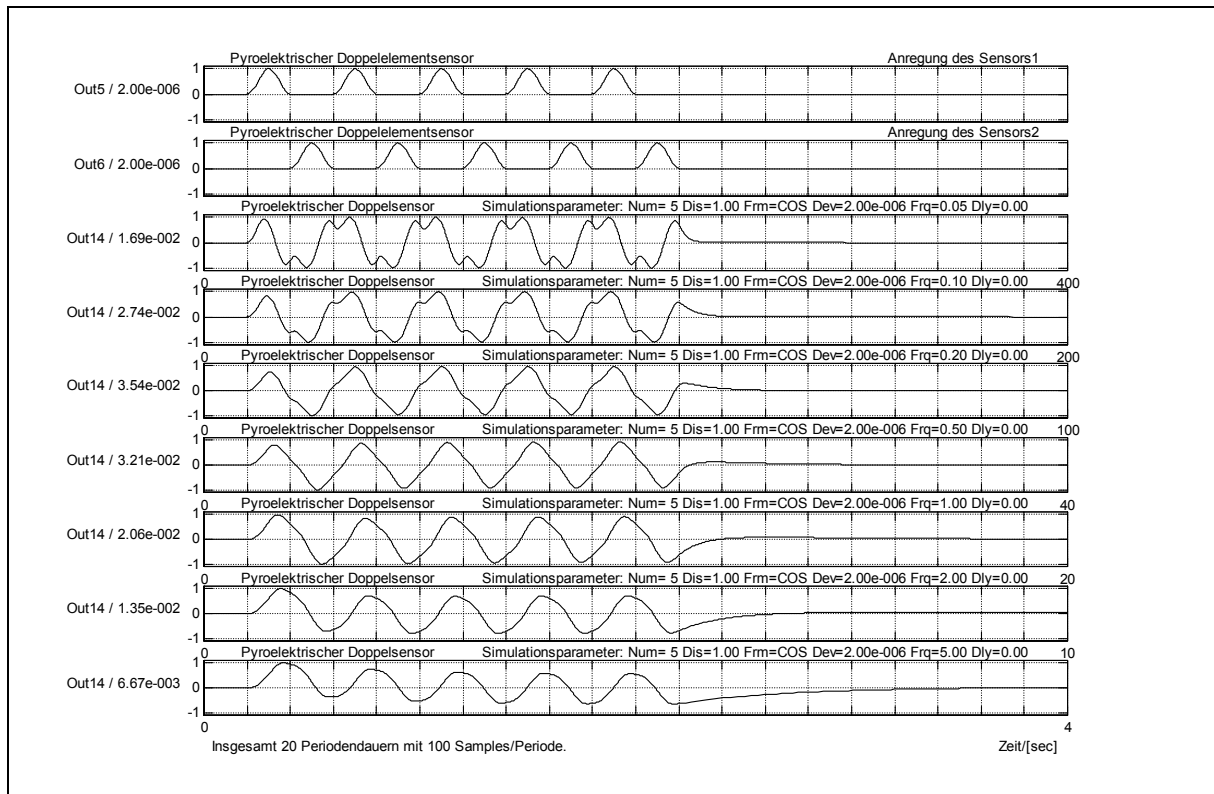


Abb. 5.20: Ausgangssignale des Doppелеlementdetektors bei Anregung mit kosinusförmigen Pulsen und unterschiedlichen Frequenzen (Frq = 0.05, 0.1, 0.2, 0.5, 1.0, 2.0 und 5.0 Hz)

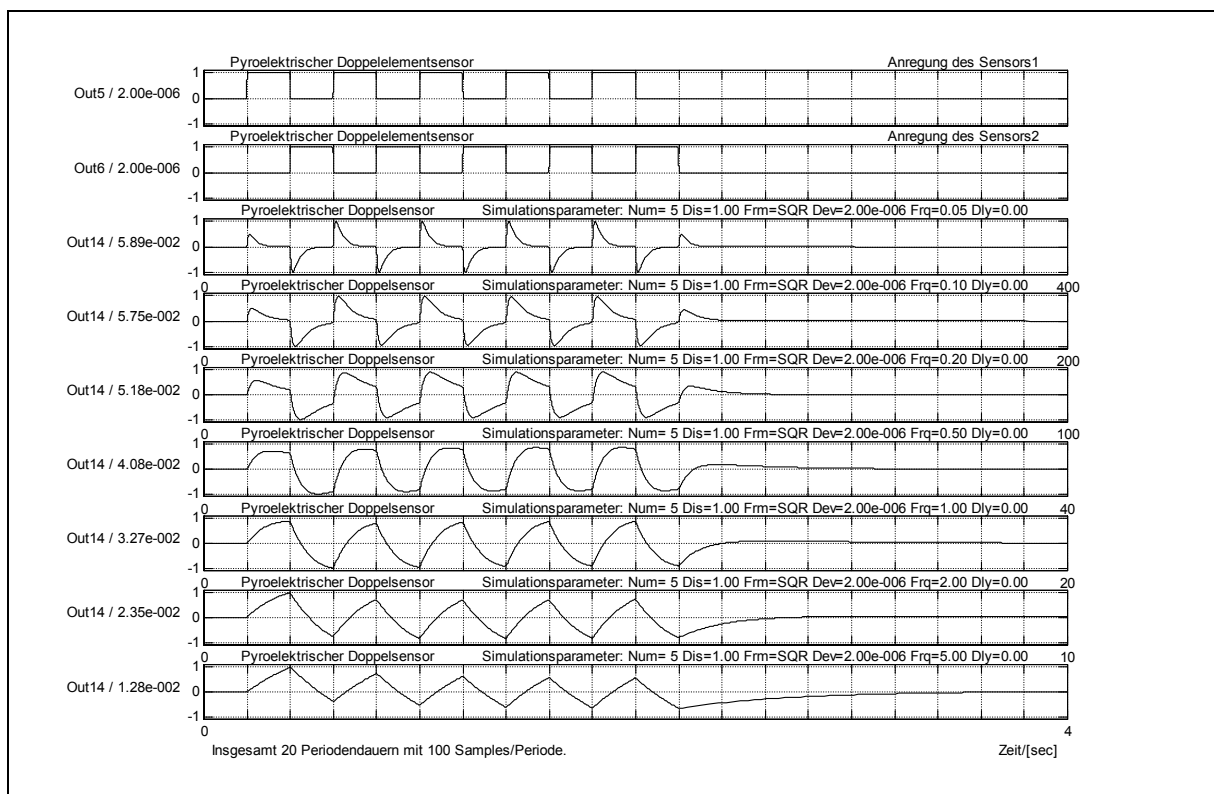


Abb. 5.21: Ausgangssignale des Doppелеlementdetektors bei Anregung mit rechteckförmigen Pulsen und unterschiedlichen Frequenzen (Frq = 0.05, 0.1, 0.2, 0.5, 1.0, 2.0 und 5.0 Hz)

Beide Abbildungen zeigen, dass Signale im unteren Frequenzbereich stärker von der thermischen Zeitkonstante beeinflusst werden, während bei höheren Frequenzen eine zunehmende Verschleifung der Signale infolge der elektrischen Zeitkonstante auftritt. Insbesondere bei rechteckförmiger Anregung sind deutliche Unterschiede im grundsätzlichen Signalverlauf zu erkennen. Langsame Signale werden stärker differenziert, während mit zunehmender Frequenz eine Verschleifung der Flanken einsetzt, da die Detektorerwärmung dem Wechsel nicht folgen kann. Dementsprechend werden auch die Signalamplituden kleiner.

Wie bereits vorher erwähnt, sind die Signalverläufe nicht nur von der Frequenz, sondern auch von der zeitlichen Verschiebung beider Eingangssignale abhängig. Die Auswirkungen der Verschiebung sollen anhand von Abb. 5.22 und Abb. 5.23 gezeigt werden. Die Simulationen wurden für eine feste Frequenz von 0.5 Hz mit der Zeitverschiebung (Simulationsparameter Dly) als Parameter durchgeführt. Die in Abb. 5.20 und Abb. 5.21 gezeigten Simulationen wurden mit einer Zeitverschiebung um eine Pulsdauer durchgeführt, wie der Vergleich der abgebildeten Eingangssignale zeigt. Für diesen Fall ist der Simulationsparameter Dly = 0.0.

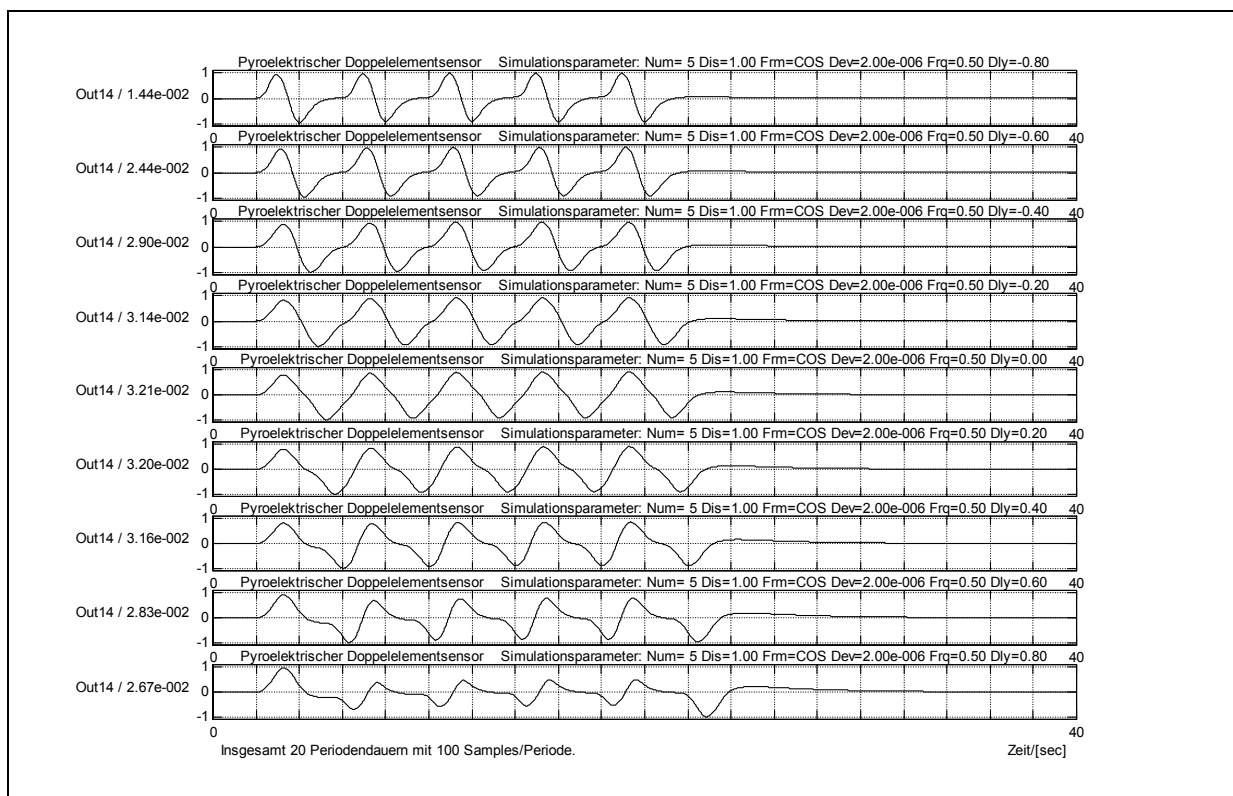


Abb. 5.22: Ausgangssignale des Doppelementdetektors bei Anregung mit kosinusförmigen Pulsen und unterschiedlichen Signalverzögerungen (Dly = 0.2-, 0.4-, 0.6-, 0.8-, 1.0-, 1.2-, 1.4-, 1.6- und 1.8-fache Pulsdauer)

Mit den in Abb. 5.22 und Abb. 5.23 angegebenen Werten wurden im Vergleich zu Dly=0.0 Zeitverschiebungen mit der 0.2-, 0.4-, 0.6-, 0.8-, 1.0-, 1.2-, 1.4-, 1.6- und 1.8-fachen Pulsdauer einstellt. Die Pausendauer wurde unverändert beibehalten, woraus sich die zum Teil unsymmetrischen Signale ergeben. Auf die Darstellung der Eingangssignale wurde der Über-

sichtlichkeit halber verzichtet. Die teilweise sprunghaften Übergänge in Abb. 5.23 entstehen durch die vollständige Kompensation bei der verschiebungsabhängigen Überlagerung der Rechteck-Impulse auf beiden Kanälen.

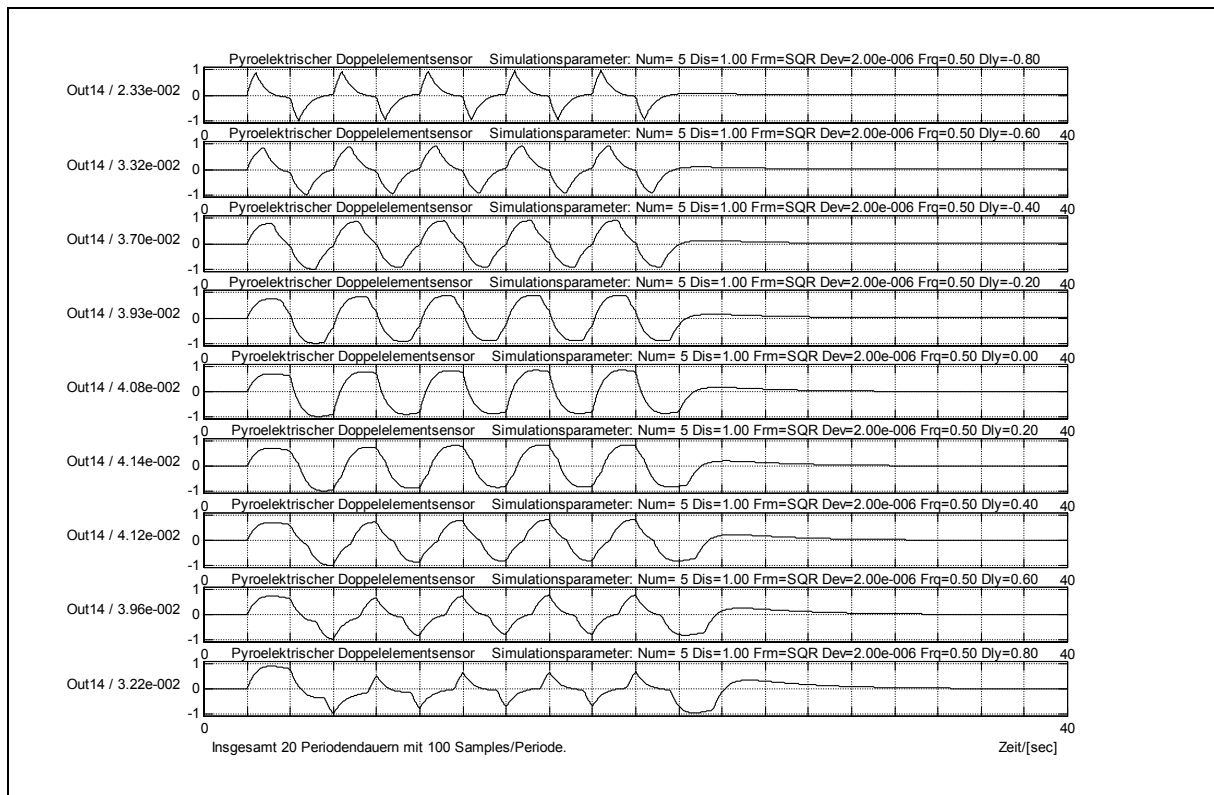


Abb. 5.23: Ausgangssignale des Doppелеlementdetektors bei Anregung mit rechteckförmigen Impulsen und unterschiedlichen Signalverzögerungen (Dly = 0.2-, 0.4-, 0.6-, 0.8-, 1.0-, 1.2-, 1.4-, 1.6- und 1.8-fache Pulsdauer)

5.2 Hardwarekomponenten von Bewegungsmeldern

Aufbauend auf dem zuvor beschriebenen Funktionsprinzip pyroelektrischer Detektoren werden in diesem Abschnitt die für einen PIR-Bewegungsmelder erforderlichen Hardwarekomponenten beschrieben. Im Vergleich zu anderen Sensorkonzepten erfordern diese Melder einen deutlich geringen Systemaufwand.

Ein weiterer Vorteil ist die Größe und die beim Entwurf in weiten Grenzen frei wählbare Form des Erfassungsfeldes. Pro 90°-Sektor wird meist ein Dualsensor eingesetzt. Typische Reichweiten liegen je nach Applikation im Bereich von 10 ...20 m. Die Reichweiten bei radialer Annäherung sind deutlich geringer und liegen oftmals unterhalb der halben, für Bewegungen parallel zum Wächter spezifizierten Reichweite.

Eine grundlegende Beschreibung der Signalquellen und des gesamten Systems soll anhand von Abb. 5.24 erfolgen. Beim Eintritt einer Person in den Überwachungsbereich des Melders ändern sich die Temperaturverhältnisse im Vergleich zum Hintergrund. Obwohl die insgesamt vorhandene Wärmestrahlung von einem Linsensystem gesammelt und auf den Detektor fo-

kussiert wird, registriert der Sensor aufgrund seines differenzierenden Verhaltens nur die von bewegten Objekten hervorgerufene Strahlungsänderung.

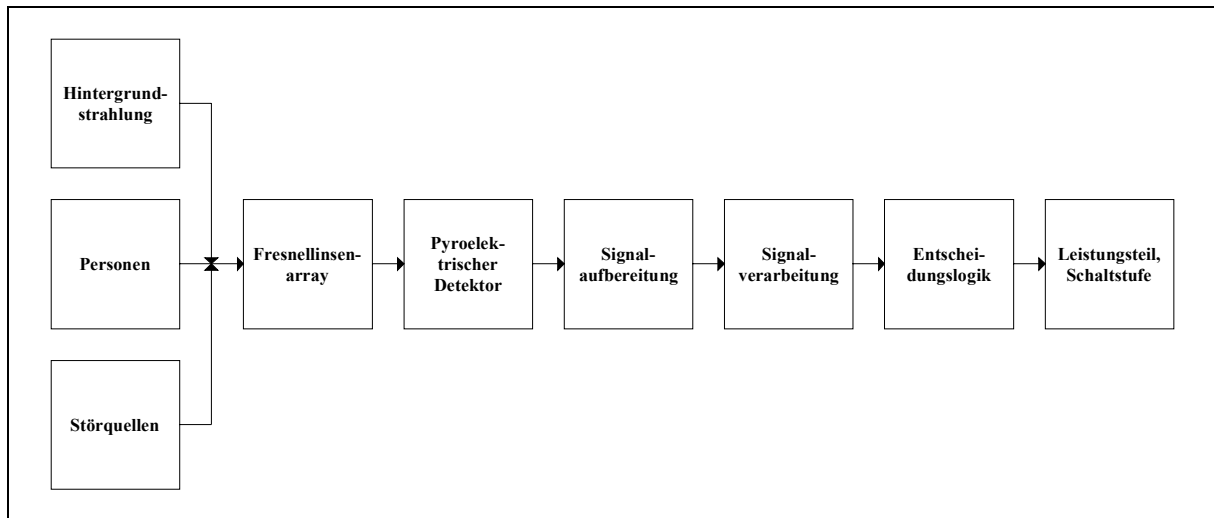


Abb. 5.24: Signalquellen und Systemkomponenten eines Passiv-Infrarot-Bewegungsmelders

Mit Hinblick auf die Größe des zu überwachenden Bereichs werden Fresnellinsenarrays zur optischen Abbildung verwendet, die aus einer größeren Anzahl von Einzellinsen bestehen. Die einzelnen Linsen zerlegen den gesamten Sektor in Teilabschnitte, wobei die aus den einzelnen Segmenten empfangene Strahlungsleistung auf den Sensorelementen überlagert wird. Infolge der Überlagerung ist eine ortsauflösende Bereichsüberwachung innerhalb eines Sektors nicht möglich.

Strahlungsänderungen und damit Fehlschaltungen können durch Tiere sowie durch vorbeifahrende Autos und Motorräder oder durch globale Temperaturänderungen infolge von Wind, Regen etc. verursacht werden. Großflächig auftretende Störungen werden zu einem großen Teil durch die gegenphasig betriebenen Sensorelemente unterdrückt. Infolge unvermeidlicher Fehlanpassungen beider Sensorelemente und wegen der meist unregelmäßigen Temperaturänderungen im Erfassungsfeld entstehen allerdings Störsignale, die durch eine intelligente Signalverarbeitung von den durch Personen hervorgerufenen Signalen diskriminiert werden müssen. In Kap. 7 werden verschiedene Signalverarbeitungsstrategien zur Unterscheidung zwischen Nutz- und Störsignalen vorgestellt.

Bevor das vom Detektor gelieferte Signal über eine geeignete Aktorik einen Schaltvorgang auslösen kann, muss es entsprechend aufbereitet werden. Eine rauscharme Analogschaltung dient zunächst zur Verstärkung und Bandbegrenzung des Sensorsignals. Dabei ist zu berücksichtigen, dass durch die analoge Filterung je nach Frequenz zusätzliche Signalverzerrungen auftreten, welche die abschließende Bewertung erschweren. Um diesen Effekt zu verdeutlichen, wurde das Übertragungsverhalten eines Bewegungsmelders mit Doppелеlementdetektor in Verbindung mit einer häufig in der Literatur anzutreffenden, zweistufigen Verstärkeranordnung simuliert [Abb. 5.25].

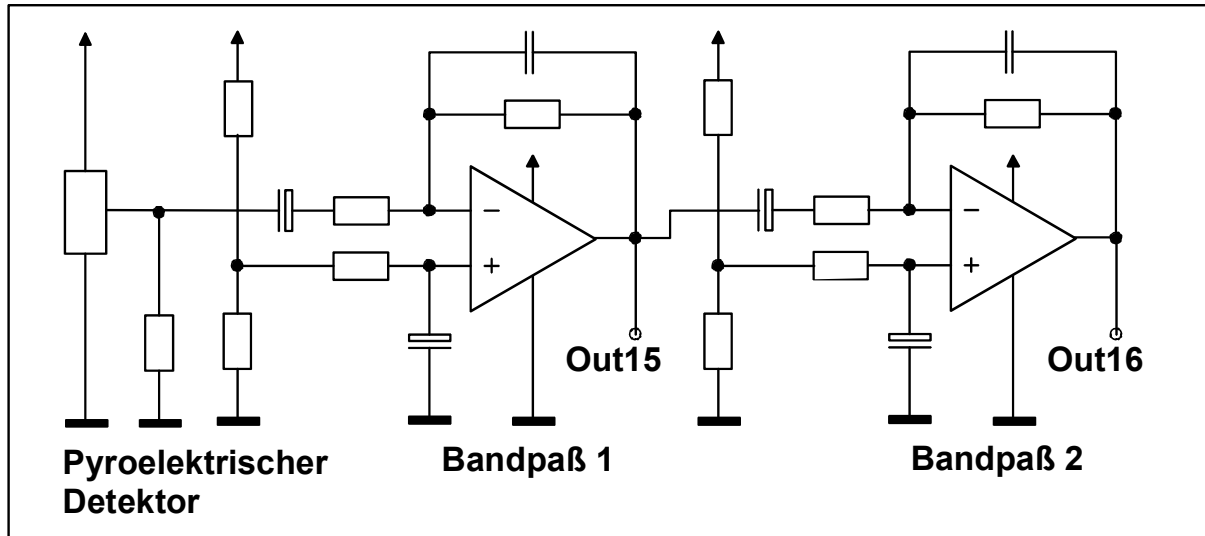


Abb. 5.25: Sensorsignalaufbereitung mit zwei invertierend beschalteten Operationsverstärkern mit Bandpasscharakteristik

Abb. 5.26 zeigt die Teilübertragungsfunktionen der Operationsverstärkerschaltung nach Abb. 5.25.

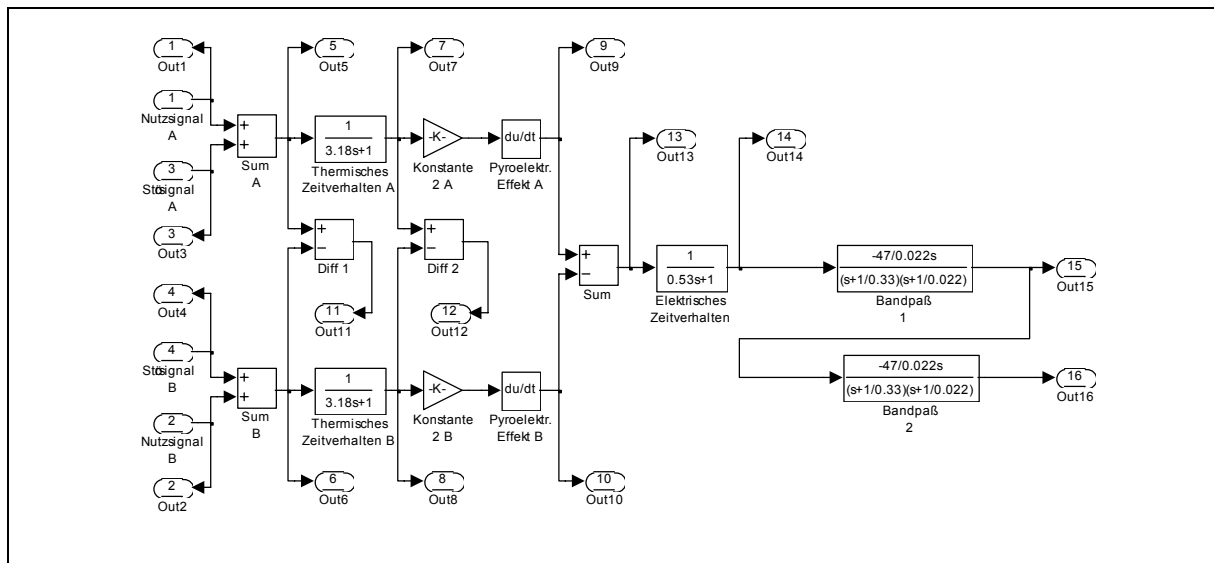


Abb. 5.26: Blockschaltbild eines Sensorsystems mit Doppелеlementsensor und zweistufigem Verstärker mit Bandpasscharakteristik

Abb. 5.27 und Abb. 5.28 zeigen die Signale am Ausgang der ersten (Out15) und der zweiten Verstärkerstufe (Out16) bei harmonischer Anregung mit unterschiedlichen Frequenzen. Aus den Simulationen ist ersichtlich, dass je nach Entfernung, Bewegungsrichtung und Bewegungsgeschwindigkeit in Verbindung mit der Bandpasscharakteristik des Sensors und der Verstärkerschaltung unterschiedliche Signalverläufe entstehen. Es zeichnet sich daher bereits an dieser Stelle ab, dass zur Wächterentwicklung und insbesondere zur Optimierung des Schaltverhaltens eine charakteristische, unter realen Bedingungen gemessene Auswahl an Testsignalen erforderlich ist.

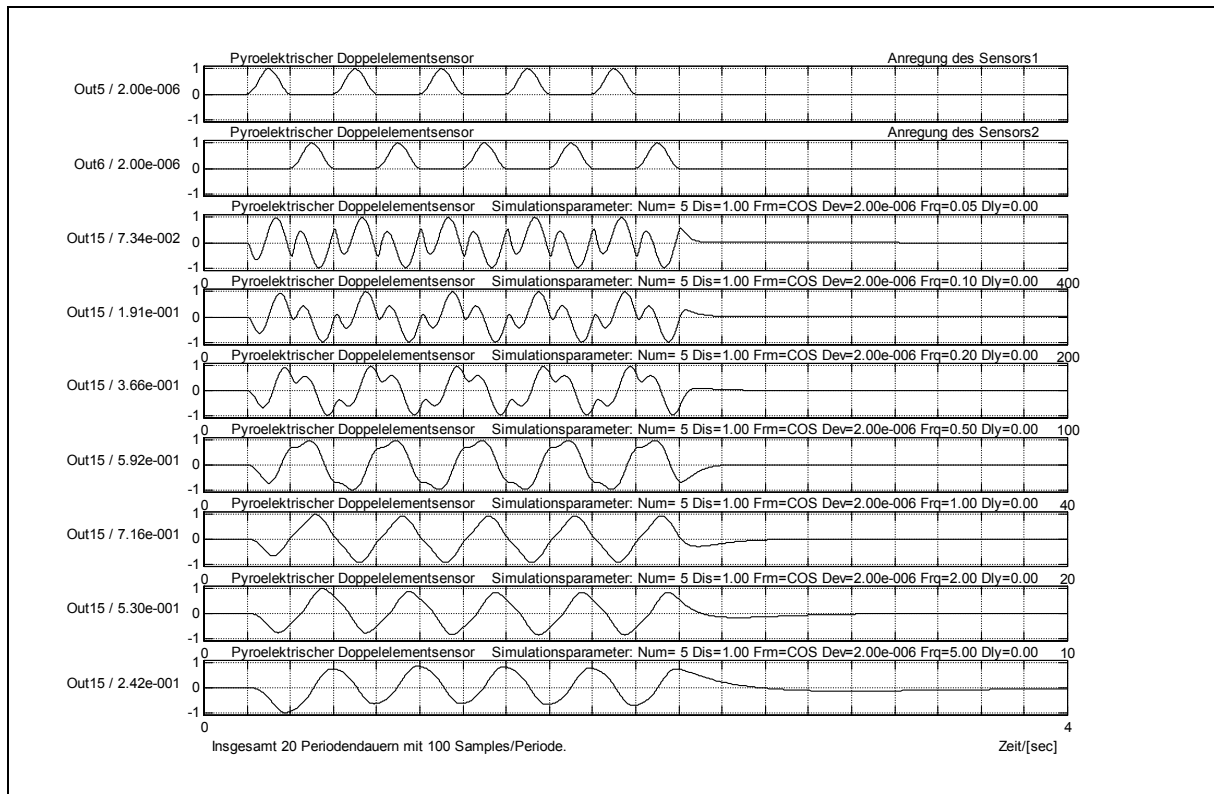


Abb. 5.27: Signale am ersten Verstärkerausgang Out15 bei Anregung mit kosinusförmigen Pulsen unterschiedlicher Frequenz (Frq = 0.05, 0.1, 0.2, 0.5, 1.0, 2.0 und 5.0Hz)

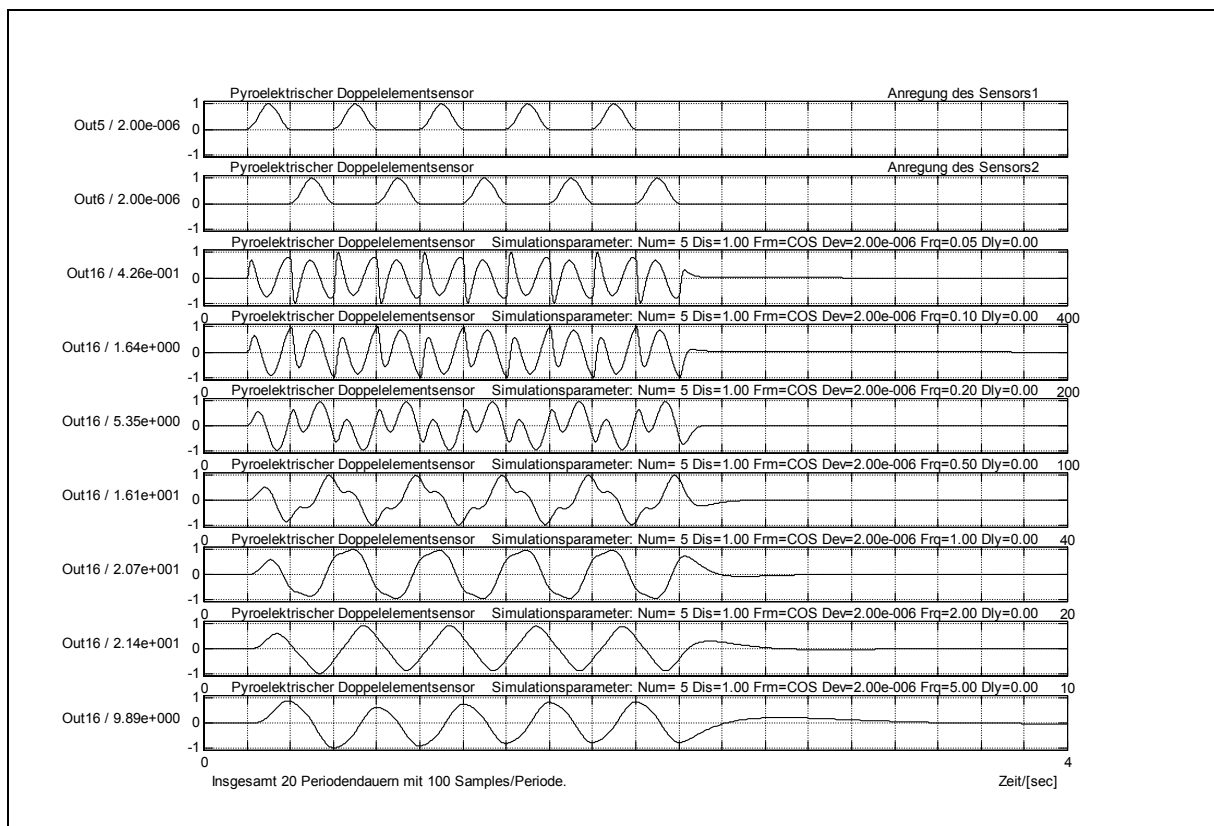


Abb. 5.28: Signale am zweiten Verstärkerausgang Out16 bei Anregung mit kosinusförmigen Pulsen unterschiedlicher Frequenz (Frq = 0.05, 0.1, 0.2, 0.5, 1.0, 2.0 und 5.0Hz)

Die in Abb. 5.24 abgebildete Entscheidungslogik kontrolliert die Einhaltung bestimmter Kriterien und führt gegebenenfalls eine Signalklassifikation durch. Während die Signalauswertung bei konventionellen Bewegungsmeldern meist mit einfachen Fensterkomparatoren erfolgt, werden in Wächtern der neuen Generation Mikrocontroller eingesetzt, die eine flexiblere Bewertung der gemessenen Signalverläufe zur Unterdrückung von Fehlalarmen ermöglichen. Als wesentliche Eigenschaften zur Reduktion der Fehlalarmrate müssen in diesem Zusammenhang eine adaptive Anpassung der Wächterempfindlichkeit an die Dynamik des Hintergrundsignals und die zusätzlichen Möglichkeiten durch multisensorielle Datenauswertung bei Meldern mit mehreren Detektoren betrachtet werden.

Unabhängig vom Aufwand der Signalverarbeitung hat das optische System einen entscheidenden Einfluss auf das Schaltverhalten von Bewegungsmeldern. Jede einzelne Linse des Fresnellinsenarrays muss unter Beachtung der gewünschten Erfassungsfeldspezifikationen und der vom Wächterdesign vorgegebenen Beschränkungen der insgesamt zur Verfügung stehenden Linsenfläche berechnet werden. Um ein möglichst kleines, unauffälliges Wächtergehäuse zu gewährleisten, ist eine optimale Ausnutzung der Gesamtlinsenfläche erforderlich. Sowohl das Wächterdesign als auch die Erfassungsfeldspezifikationen orientieren sich meist an marktpolitischen Überlegungen.

5.3 Fresnellinsen zur Bündelung langwelliger Temperaturstrahlung

Im nächsten Kapitel wird der Entwurf und die Berechnung von Fresnellinsenarrays unter Beachtung der Erfassungsfeldspezifikationen beschrieben. In diesem Abschnitt werden daher nur die grundlegenden Zusammenhänge zur Erläuterung der optischen Eigenschaften von Fresnellinsen und die aus der geometrischen Optik bekannten Abbildungsgleichungen für dünne Linsen zusammenfasst.

Für PIR-Bewegungsmelder werden zur optischen Signalverstärkung meist Linsen mit Brennweiten im Bereich von 20 ...40 mm eingesetzt. Zusätzlich erfordern die geringen Empfangsleistungen aus dem Fernfeld relativ große Linsenaperturen (20 ...50 mm), so dass konventionelle Linsen aufgrund der notwendigen Dicke sehr hohe Absorptionsverluste aufweisen würden. Höherwertige Linsen aus anderen Materialien, die im Wellenlängenbereich von 6 ...14 μm bessere Transmissionseigenschaften aufweisen, sind für die vorliegende Applikation zu kostenintensiv.

Stattdessen werden üblicherweise dünne Linsen z. B. aus Polyäthylen (Brechzahl $n \approx 1.5$) eingesetzt, bei denen der Umstand ausgenutzt wird, dass die Brechung der einfallenden Strahlung ausschließlich vom Verlauf der Grenzfläche zwischen dem Linsenmaterial und der Luft abhängt. Aufgrund der geringen Linsendicke (1 ...2 mm) können Materialien mit vergleichsweise schlechtem Transmissionsvermögen verwendet werden.

Das Dickenprofil einer Standardlinse nach Abb. 5.29 kann mit dem Malusschen Satz berechnet werden. Er besagt, dass die Laufzeiten einer Wellenfront von Punkt A bzw. von Punkt A' nach Punkt F identisch sind. Mit den verschiedenen Brechzahlen n_1 und n_2 im Lichtweg, der Lichtgeschwindigkeit c und der Ausbreitungsgeschwindigkeit $v=c/n$ im Linsenmaterial mit der Brechzahl n folgt:

$$(Gl. 5.38) \quad \frac{D(0)}{v} + \frac{f - D(0)}{c} = \frac{D(r)}{v} + \frac{\sqrt{[f - D(r)]^2 + r^2}}{c}$$

Für dünne Linsen mit $f \gg D(r)$ kann daraus die Linsendicke als Funktion des Linsenradius r berechnet werden:

$$(Gl. 5.39) \quad D(r) = D(0) + \frac{f - \sqrt{f^2 + r^2}}{n}$$

Unterteilt man die Linse in Streifen der Dicke $b \cdot \lambda/n$, wobei nur die Abschnitte mit dem ursprünglichen Randprofil weiterverwendet werden, erhält man die Fresnelstruktur nach Abb. 5.30.

Die Standardlinse und die Fresnellinse haben dieselben Brechungseigenschaften, während die Transmissionsverluste der Fresnellinse wegen der deutlich geringeren Dicke kleiner sind.

Die optimale Breite $D(0)$ einer Fresnellinse beträgt λ/n . Aus fertigungstechnischen Gründen und auch wegen der erforderlichen Stabilität werden beim Bewegungsmelderdesign aber üblicherweise Linsenstärken um 1 mm verwendet, wodurch sich die Transmissionsverluste im Vergleich zur optimalen Dicke deutlich erhöhen. Des weiteren nimmt auch die Beugungseffizienz mit wachsender Linsendicke ab.

Ausgehend von der Linsenstärke $b \cdot \lambda/n$ können die Radien r_m der einzelnen Zonen bzw. die Brennweite der Linse berechnet werden:

$$(Gl. 5.40) \quad r_m = \sqrt{(m\lambda)^2 + 2mbf\lambda}$$

$$(Gl. 5.41) \quad f = \frac{r_m^2 - (m\lambda)^2}{2m\lambda}$$

Aus (Gl. 5.40) folgt die maximale Zonenanzahl M einer Fresnellinse mit dem Radius R und der Blendenzahl $F = f/(2R)$:

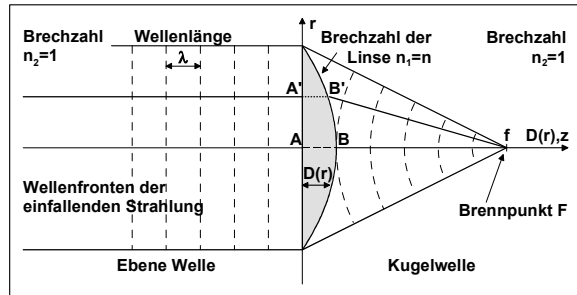


Abb. 5.29: Fokussierung einer ebenen Welle mit einer Standardlinse

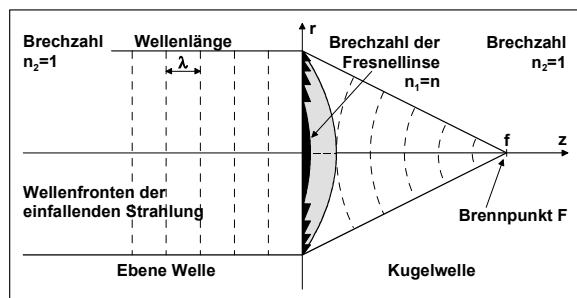


Abb. 5.30: Fokussierung einer ebenen Welle mit einer Fresnellinse

$$(Gl. 5.42) \quad M = \frac{1}{b\lambda} \cdot \left[\sqrt{f^2 + R^2} - f \right] = \frac{R}{b\lambda} \cdot \left[\sqrt{1^2 + 4F^2} - 2F \right]$$

Der maximale Radius der Fresnellinse ist durch den Winkel der Totalreflexion an der Grenzfläche vorgegeben. Eine weitere Vergrößerung des Linsendurchmessers bringt keine Vorteile [Lit. 5.3]. Für $M \gg 1$ und $b = 1$ kann mit (Gl. 5.40) der minimale Abstand $R_{\min} = r_M - r_{M-1}$ zweier Zonen ermittelt werden. Mit der numerischen Apertur NA erhält man die Beziehung

$$(Gl. 5.43) \quad R_{\min} = \frac{\lambda \cdot (f + M \cdot \lambda)}{R} = \frac{\lambda}{NA} \quad \text{mit} \quad NA = \frac{R}{\sqrt{R^2 + f^2}}$$

Die Intensitätsverteilung $I(p)$ in der Brennebene wird für den Fall einer ebenen einfallenden Welle durch die Airy-Funktion nach (Gl. 5.44) beschrieben. Abb. 5.31 zeigt den prinzipiellen Verlauf dieser Funktion mit dem beugungsbegrenzten Spottedurchmesser $2 \cdot w_{\text{Airy}} = 1.64 \cdot \lambda \cdot f$. Der Durchmesser wird definitionsgemäß an der Stelle gemessen, an der die Intensitätsverteilung auf den Faktor $1/e^2$ abgesunken ist [Lit. 5.5].

$$(Gl. 5.44) \quad I(p) = \left(\frac{2J_1(p)}{p} \right)^2 \quad \text{mit} \quad p = \frac{2\pi R}{\lambda f} \cdot r = \frac{\pi \cdot r}{\lambda \cdot F}$$

Eine Linse mit kleiner Blendenzahl F erzeugt einen geringeren Spottedurchmesser. Die minimale Blendenzahl wird durch den minimalen Abstand R_{\min} nach (Gl. 5.43) bestimmt:

$$(Gl. 5.45) \quad F_{\min} = \frac{1}{2} \sqrt{\left(\frac{\lambda}{R_{\min}} \right)^2 - 1}$$

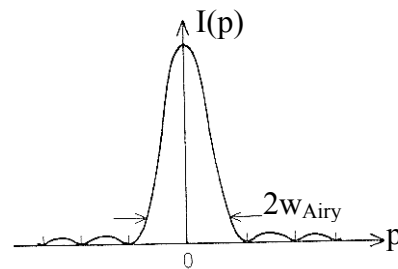


Abb. 5.31: Prinzipieller Verlauf der Intensitätsverteilung nach (Gl. 5.44)

Die Abbildungseigenschaften von Fresnellinsen sind durch die aus der geometrischen Optik bekannten Abbildungsgleichungen für dünne Linsen gegeben. Zwischen der Brennweite f der Linse, der Bildweite b und der Gegenstandsweite g gilt nach Kap. 3 der Zusammenhang

$$(Gl. 5.46) \quad \frac{1}{f} = \frac{1}{b} + \frac{1}{g},$$

für den Abbildungsmaßstab β folgt aus den geometrischen Beziehungen nach Abb. 5.32:

$$(Gl. 5.47) \quad \beta = \frac{b}{g} = \frac{B}{G} = \frac{b}{f} - 1 = \frac{1}{\frac{g}{f} - 1}$$

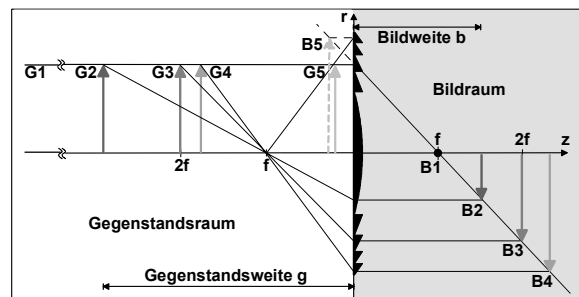


Abb. 5.32: Bildweiten in Abhängigkeit der Gegenstandsweite

Liegt die Gegenstandsweite g nach Abb. 5.32 unterhalb der Brennweite f , ergeben sich infolge der divergenten Strahlenbündel virtuelle "aufrechtstehende" Bilder (Beispiel G5-B5), während Gegenstandsweiten oberhalb der Brennweite jeweils reelle, gespiegelte Bilder (G2-B2, G3-B3, G4-B4) hervorrufen. Der Bereich vom Unendlichen bis zur doppelten Brennweite $2f$ wird dabei im Bildraum in das Intervall $[f, 2f]$ abgebildet.

Aufgrund der typischen Erfassungsfelder liegt bei konventionellen Bewegungsmeldern vorwiegend der Fall einer aus großen Entfernungen einfallenden Strahlung (G1) vor. Diese Strahlung wird im Brennpunkt B1 fokussiert. Der Detektor wird daher gewöhnlich in der Nähe der Brennebene positioniert.

Abweichend von den zuvor beschriebenen idealen Abbildungseigenschaften treten infolge des großen Variationsbereichs des Strahlungseinfallswinkels und des zur Bewegungserkennung verwendeten Wellenlängenbereichs sphärische und chromatische Abberationen auf, die aber im folgenden nicht weiter berücksichtigt werden sollen.

5.4 Literatur

- [Lit. 5.1] Texas Instruments Deutschland GmbH: Das Opto-Kochbuch - Theorie und Praxis der Optoelektronik, 1. Auflage, Technik Marketing, München, 1975
- [Lit. 5.2] K. Herrmann, L. Walther: Wissensspeicher Infrarottechnik, 1. Auflage, Fachbuchverlag Leipzig, 1990
- [Lit. 5.3] Valvo Unternehmensbereich Bauelemente der Philips GmbH, 1. Auflage, Dr. Alfred Hüthig Verlag, 1987
- [Lit. 5.4] E. Dereniak, D. Crowe: Optical Radiation Detectors, John Wiley & Sons, 1984
- [Lit. 5.5] E. Wolf: Progress In Optics, Volume XXIV, Elsevier Science Publishers B.V., 1987

6 Systementwicklung von PIR-Bewegungsmeldern

6.1 Simulationstool zum Entwurf von Fresnellinsenarrays

In Kap. 5 wurden die strahlungsphysikalischen Grundlagen, das Sensorprinzip und das Übertragungsverhalten von pyroelektrischen Detektoren beschrieben. Ausgehend von den zuvor beschriebenen Abbildungseigenschaften von Fresnellinsen wird in diesem Kapitel ein Softwaretool zur Dimensionierung des Erfassungsfeldes von Bewegungsmeldern vorgestellt, das zum Entwurf und zur Optimierung von Fresnellinsenarrays verwendet werden kann und das sich im praktischen Einsatz durchaus bewährt hat.

Abb. 6.1 zeigt als Beispiel das optische System eines Bewegungsmelders, das aus einer Vielzahl einzelner Fresnellinsen besteht. Jede Einzellinse bildet, wie später noch dargestellt wird, einen Teilbereich des gesamten Erfassungsfeldes auf die verschiedenen Sensorelemente ab.

Normalerweise sind mehrere Einzellinsen auf den Detektor fokussiert, der die günstigste Energieübertragung bietet.

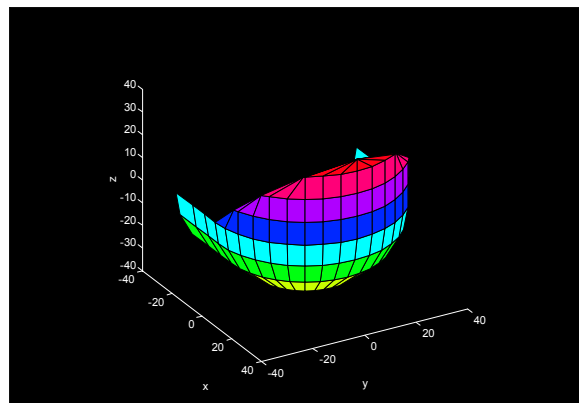


Abb. 6.1: Optisches System eines PIR-Bewegungsmelders

Auf den Sensorelementen entstehen daher Mehrfachüberlagerungen, so dass eine eindeutige räumliche Zuordnung von Strahlungsänderungen nicht mehr möglich ist. Relativ zu den übrigen Detektoren des Melders sind meist starke Defokussierungen der Linsen vorhanden. Die aus dieser Abbildung resultierenden Signalanteile liefern keinen nennenswerten Beitrag zum Schaltverhalten des Bewegungsmelders.

Die im folgenden vorgestellte Erfassungsfeldsimulation [Kap. 6.1.1] basiert auf den Grundlagen der geometrischen Optik und wurde mit dem Programmpaket "MATLAB" erstellt. Aufbauend auf den Entwicklungsergebnissen und den praktischen Erfahrungen beim Wächterdesign ermöglicht diese Software eine zeit- und kostenminimale Dimensionierung des optischen Systems von Bewegungsmeldern. Für zukünftige Entwicklungen wurde neben der Verwendung von Einzel- und Doppелеlementdetektoren auch der Einsatz von Sensorarrays mit größeren Elementanzahlen berücksichtigt.

Das Simulationstool berechnet aus den 3D-Sensorpositionen und den 3D-Positionen der einzelnen Fresnellinsen die Projektion der aktiven Elemente in den Überwachungsbereich. Dabei entstehen die für Bewegungsmelder typischen Erfassungsfelder, in denen sich überwachte und nicht überwachte Bereiche in den verschiedenen Entfernungszonen abwechseln.

Speziell für Linsengeometrien, die auf eine 2D-Abwicklung zurückzuführen sind, ist zusätzlich eine automatische Initialisierung des gesamten Fresnellinsenarrays anhand der geometrischen Parameter der Sensor-Linse-Anordnung und der Gesichtsfeldspezifikationen möglich. Letztere beschreiben die gewünschten Schaltreichweiten in Abhängigkeit der Erfassungsrichtung.

Unter Berücksichtigung einer auf Erfahrungswerten basierenden Überreichweite zur Erzielung des geforderten Schaltverhaltens werden daraus die Linsenzentren und die für eine Fokussierung erforderlichen Brennweiten der Einzellinsen berechnet. Bei vorgegebener Linsenhöhe innerhalb einer Entfernungzone wird auch die Breite der einzelnen Linsen unter Beachtung des Einfallswinkels der Strahlung ermittelt.

Anhand des Übertragungsverhaltens der Detektoren, der Sensorpositionen und anhand des optischen Systems ist es denkbar, typische Bewegungsmeldersignale zu simulieren. Zur Simulation sind Annahmen bezüglich der von Personen emittierten Strahlungsleistung, bezüglich der Definition von Bewegungsrichtungen, -geschwindigkeiten und der Körperform sowie zahlreiche Vereinfachungen hinsichtlich der Strahlungsübertragung und der Qualität der optischen Abbildung erforderlich.

Störsignale durch variable Umgebungsverhältnisse können aufgrund ihrer vielfältigen Ursachen und der vom jeweiligen Erfassungsfeld abhängigen örtlichen Gegebenheiten nur unzureichend berücksichtigt werden. Aus diesem Grund wurde die Simulation von Wächtersignalen im Rahmen der vorliegenden Arbeit nicht weiterverfolgt.

Aufgrund der Komplexität des gesamten Systems wurde stattdessen für die weitere Bewegungsmelderentwicklung eine empirisch orientierte Vorgehensweise gewählt, die auf realen, in verschiedenen Testfeldern gemessenen Bewegungsmeldersignalen basiert. Ein weiterer Schwerpunkt dieses Kapitels ist daher die Vorstellung eines Testsystems zur Aufnahme und zur Verarbeitung von Bewegungsmeldersignalen, das die gezielte Optimierung von Signalverarbeitungsstrategien und den Vergleich unterschiedlicher Verfahren sowie die Visualisierung der Wächtersignale in Verbindung mit den internen Zuständen der Signalverarbeitung ermöglicht.

6.1.1 Simulation des Erfassungsfeldes

In diesem Abschnitt werden einige grundlegende Überlegungen zur Simulation des Erfassungsfeldes von Bewegungsmeldern vorgestellt. An dieser Stelle soll die Kenntnis der 3D-Koordinaten der einzelnen Sensoren und Linsen vorausgesetzt werden [s. Kap 6.1.2]. Die Sensoren und die Linsen werden jeweils als ebene Flächen modelliert. Abbildungsfehler infolge der tatsächlich vorhandenen Biegung einzelner Linsen, chromatische Abberationen, Koma etc. werden im folgenden nicht weiter berücksichtigt.

Für dünne Linsen wird die Brennebene im Rahmen der geometrischen Optik als plane, zur Hauptachse der Linse parallele Ebene im Abstand der Brennweite f betrachtet [Abb. 6.2].

Diese Näherung gilt insbesondere für räumlich begrenzte Strahlenbündel, die parallel oder unter einem kleinen Winkel gegen die optische Achse der Linse einfallen.

Ein achsparallel einfallendes Strahlenbündel (ebene Wellenfront) wird in den Brennpunkt F der Linse abgebildet.

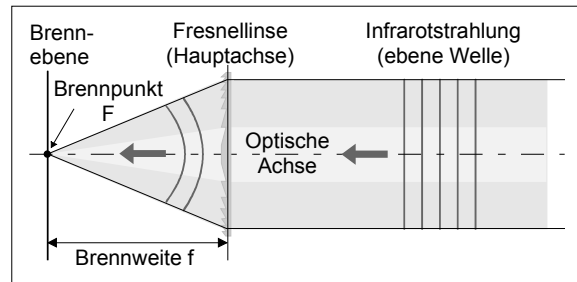


Abb. 6.2: Fokussierung einer parallel zur optischen Achse einfallenden ebenen Welle im Brennpunkt F

Abb. 6.3 zeigt den Querschnitt eines Erfassungsfeldes mit drei Entfernungszonen sowie eine mögliche Anordnung des Sensor-Linse-Systems. Aufgrund der im Vergleich zu den Sensor- und den Linsenabmessungen großen Reichweite von Bewegungsmeldern kann die aus den verschiedenen Entfernungszonen einfallende Strahlung grundsätzlich als ebene Welle betrachtet werden. Richtet man die Sensoren und die Linsen der äußeren Zone [Zone 3] jeweils auf den weitesten Entfernungsbereich aus, um in diesen Zonen ein Maximum an Empfindlichkeit bereitzustellen, sind die im Rahmen der geometrischen Optik getroffenen Vereinfachungen dort am besten erfüllt.

Die drei Linsen, die das Erfassungsfeld nach Abb. 6.3 erzeugen und deren Ausrichtung und Zentrenlage durch die optischen Achsen OA_1 , OA_2 und OA_3 angedeutet ist, sind parallel zur Sensorebene im Abstand der Brennweite f angeordnet. In allen Linsen sind die Zentren enthalten. Oftmals werden beim Linsendesign nur Teilbereiche einer Linse verwendet, die nicht notwendigerweise das Zentrum enthalten.

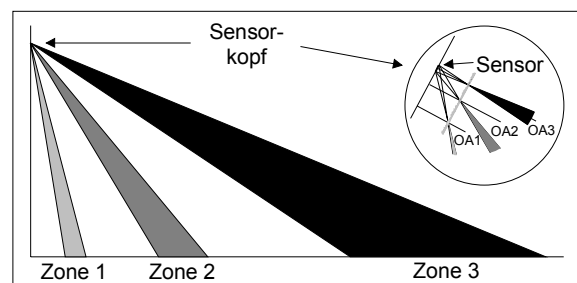


Abb. 6.3: Vertikaler Schnitt durch das Erfassungsfeld eines Bewegungsmelders mit drei Entfernungszonen

Bei der Anordnung nach Abb. 6.3 sind insbesondere für die beiden unteren Linsenebenen, die die Entfernungszonen 1 und 2 erzeugen, größere Abbildungsfehler aufgrund des im Vergleich zu Zone 1 größeren Einfallwinkels der Strahlung zu erwarten. Berechnet man auch für diese Zonen das Erfassungsfeld mit Hilfe der geometrischen Optik, müssen die resultierenden Ergebnisse als grobe Vereinfachung der tatsächlichen Verhältnisse betrachtet werden.

Der Umstand, das im Rahmen der Simulation lediglich eine möglichst homogene Verteilung der einzelnen Erfassungsfeldbereiche in den verschiedenen Entfernungszonen angestrebt

wird, mag als Rechtfertigung für diese Vorgehensweise dienen. Unter der Annahme einer ausreichende Linsengröße für den Nahbereich von 0.0 ... 6.0 m, der typischerweise von den inneren Zonen überwacht wird, ist weiterhin festzustellen, dass der Wächter in diesen Zonen aufgrund der hohen Wärmegradienten ohnehin bereits beim erstmaligen Eintritt einer Person schaltet. Systemoptimierungen zur Erzielung einer stabilen Quergehreichweite bei gleichzeitig möglichst hoher Reichweite für radial bewegte Objekte betreffen daher im wesentlichen die Linsen für den weitesten Entfernungsbereich.

Abb. 6.4 zeigt ein Sensorelement in der Brennebene der Linse. Der maximale Einfallswinkel gegen die optische Achse, unter dem eine aus großen Entfernungen einfallende Strahlung auf den Detektor abgebildet wird, kann aus den geometrischen Abmessungen des Sensors und der Linse sowie anhand der Lage der Brennebene ermittelt werden.

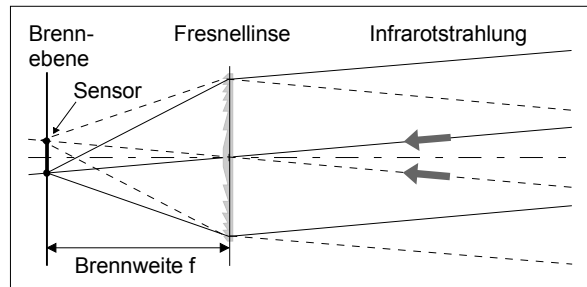


Abb. 6.4: Bestimmung des maximalen Einfallswinkels der Strahlung für ein Sensorelement in der Brennebene

Mit den Zentralstrahlen läßt sich die Projektion des Sensors in eine vordefinierte Bezugsebene berechnen. Die Zentralstrahlen verlaufen, wie der Name bereits andeutet, durch das Linsenzentrum und den jeweiligen Schnittpunkt der Randstrahlen mit der Brennebene. Die Randstrahlen sind in der 2D-Darstellung nach Abb. 6.4 durch die Randpunkte der Linse und die jeweils gegenüberliegenden Eckpunkte des Sensorelements vorgegeben.

Die Berechnung der Zentralstrahlen kann sinngemäß auf eine 3D-Anordnung übertragen werden, wie später noch dargestellt wird. Im Rahmen des Wächterdesigns ist die Anordnung des Sensors in der Brennebene nur für einzelne Linsen möglich, für die meisten Linsen treten Abweichungen von dieser idealen Lage auf. Abb. 6.5 und Abb. 6.6 verdeutlichen die Vorgehensweise zur Ermittlung der Zentralstrahlen, wenn die Sensorebene nicht mit der Brennebene übereinstimmt.

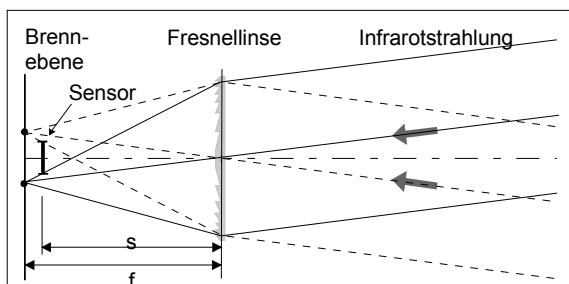


Abb. 6.5: Maximaler Einfallswinkel eines defokussierten Sensors ($s \leq f$)

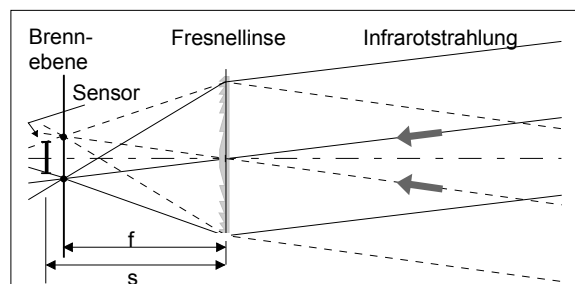


Abb. 6.6: Maximaler Einfallswinkel eines defokussierten Sensors ($s \geq f$)

Der wesentliche Unterschied zwischen beiden Abbildungen ist die Vertauschung der Randstrahlen durch die Sensoreckpunkte, welche zur Ermittlung der Schnittpunkte in der Brennebene verwendet werden.

Mit Hinblick auf die am Sensorort vorhandene Bestrahlungsstärke sollte die Defokussierung der Linsen in beiden Richtungen möglichst klein sein. Zur Dimensionierung von Fresnellinsenarrays müssen die Abstände der Sensorebene von der betrachteten Linsenebene bekannt sein.

Neben der Berechnung aller erforderlichen Schnittpunkte im 3D-Raum stellt das Simulationstool daher auch Informationen über die longitudinale und die laterale Deplazierung der Sensormittelpunkte, relativ zum zugehörigen Linsenzentrum, zur Verfügung. Der longitudinale, in Richtung der optischen Achse gemessene Abstand entspricht dabei der Brennweite für eine fokussierte Linse.

Abb. 6.7 zeigt die Abbildung der aktiven Elemente eines Doppelementensors in eine horizontale Bezugsebene. Es ist zweckmäßig, als Bezugsebene die "Bodenebene" zu wählen. Der Sensor befindet sich in der Montagehöhe M oberhalb dieser Ebene.

Beispielhaft wird im folgenden die Projektion des linken Sensorelements mit den Eckpunkten A_i ($i = 1, 2, 3, 4$) in die Bezugsebene betrachtet. Je nach Abstand s_i der Sensoreckpunkte von der Linsenebene (Linseneckpunkte L_i) werden die in die Bezugsebene projizierten Eckpunkte P_i ausgehend von den in Tab. 6.1 angegebenen Verbindungsgeraden (Randstrahlen) bestimmt:

Projizierter Eckpunkt	$s_i \leq f$ Gerade durch	$s_i \geq f$ Gerade durch
P1	L3 und A1	L1 und A1
P2	L4 und A2	L2 und A2
P3	L1 und A3	L3 und A3
P4	L2 und A4	L4 und A4

Tab. 6.1: Projektion der Sensorelemente in Abhängigkeit der Brennweite und des Abstands zur Linsenebene

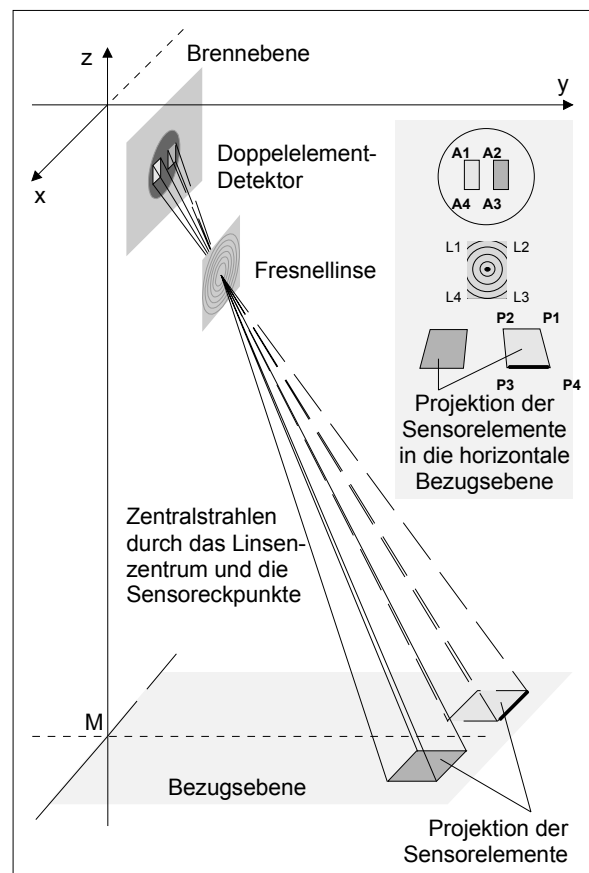


Abb. 6.7: Projektion der Sensorelemente in den Überwachungsbereich

Ausgehend von den Randstrahlen nach Tab. 6.1 werden zunächst die Schnittpunkte mit der zur Linsenebene parallelen Brennebene berechnet. Die Verbindungsgeraden durch diese Schnittpunkte und das Linsenzentrum schneiden die Bezugsebene in den Punkten P_i , die den Eckpunkten des in die Bezugsebene projizierten Sensorelements entsprechen.

Entsprechend der Linsenanzahl im Fresnellinsenarray wird die Zahl der aktiven Elemente eines mehrelementigen Sensors im Erfassungsfeld des Bewegungsmelders vervielfacht. In Abhängigkeit der Sensorgeometrie und der Lage der Linsenzentren entstehen daher überwachte und nicht überwachte Bereiche. Die Abstand zwischen benachbarten Erfassungsfeldbereichen, die durch verschiedene Linsen erzeugt werden, darf nicht zu groß sein, um eine lückenlose Überwachung zu gewährleisten.

6.1.2 Berücksichtigung der Systemparameter und der Erfassungsfeldspezifikationen

Im vorhergehenden Abschnitt wurde die Kenntnis der Sensor- und Linsenpositionen im 3D-Raum vorausgesetzt. In diesem Kapitel wird die Berechnung dieser Koordinaten, basierend auf der allgemeinen Festlegung der Systemparameter, beschrieben. Desweiteren wird eine auf den Erfassungsfeldspezifikationen beruhende, automatische Initialisierung eines Fresnellinsenarrays mit einer 2D-Abwicklung vorgestellt.

6.1.2.1 Sensorpositionen und Orientierung

Zunächst müssen die Abmessungen der einzelnen Sensorelemente und deren geometrische Anordnung im Detektor-Array, relativ zu dessen Mittelpunkt, spezifiziert werden [Abb. 6.8].

Um die Vorgabe beliebiger Elementgrößen und Zwischenräume zu ermöglichen, werden die Elemente einzeln anhand ihrer Eckpunkte identifiziert. Die Angaben werden in einer Datei gespeichert. Bei Bedarf erfolgt dann die Drehung der einzelnen Arrays um die verschiedenen Achsen im Sensorkoordinatensystem [Abb. 6.8].

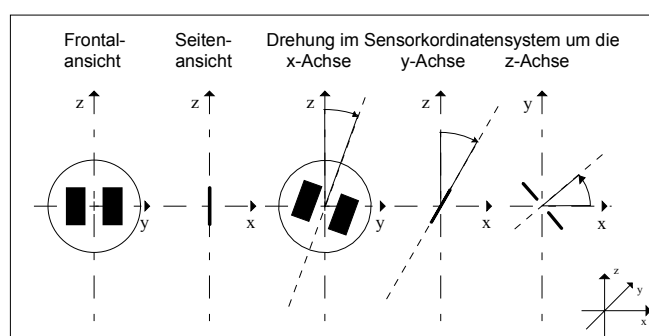


Abb. 6.8: Beschreibung und Drehung eines Arrays im Sensorkoordinatensystem

Zur Positionierung der Sensorarrays im Weltkoordinatensystem ist die Vorgabe der Ortsvektoren erforderlich [Abb. 6.9]. Die Komponenten der Ortsvektoren können unmittelbar oder durch die Drehung eines skalierten x-Einheitsvektors um die y- und die z-Achse angegeben werden.

Durch die Addition der Ortsvektoren in Weltkoordinaten und der Ortsvektoren der Eckpunkte im Sensorkoordinaten sind die Eckpunkte der Einzelemente im Weltkoordinatensystem festgelegt.

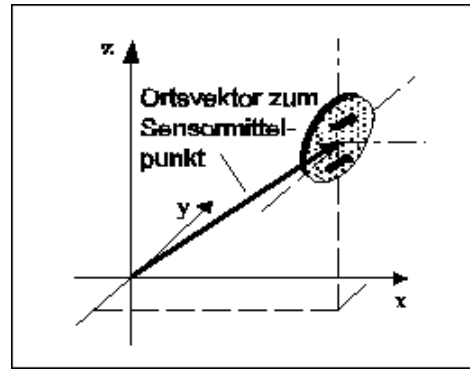


Abb. 6.9: Plazierung eines Arrays im Weltkoordinatensystem

6.1.2.2 Geometrie und Positionierung der Gesamtlinsenfläche

Die Form des optischen Systems von kommerziell erhältlichen Bewegungsmelder wird oftmals von markt- bzw. designpolitischen Überlegungen bestimmt. Meist ist ein relativ unauffälliges Wächterdesign gefordert, was sich im allgemeinen nachteilig auf die insgesamt zur Verfügung stehende Linsenfläche auswirkt. Eine möglichst optimale Ausnutzung der vorhandenen Gesamtfläche ist daher eine wesentliche Voraussetzung für ein zufriedenstellendes Schaltverhalten, zumal die Einzellinsenflächen mit Hinblick auf die Leistungsbilanz weder frei positioniert noch beliebig verkleinert werden können.

Neben aufwendigeren 3D-Linsen [z. B. Abb. 6.1] werden beim Bewegungsmelderdesign häufig Fresnellinsenarrays verwendet, die in eine zylindrische Form gebogen sind. Das optische System kann in diesem Fall auf eine 2D-Abwicklung zurückgeführt werden. Im Vergleich zu den 3D-Anordnungen verursacht die Herstellung dieser Linsen geringere Kosten und es existieren keine Restriktionen bezüglich der Entformbarkeit im Anschluss an den Prägeprozess.

Aufgrund der einfachen geometrischen Beziehungen wird im folgenden insbesondere die zylindrische Form mit einer 2D-Abwicklung zugrunde gelegt. Prinzipiell ist es allerdings auch möglich, komplexere Linsengeometrien durch die unmittelbare Angabe der 3D-Koordinaten der Eckpunkte und der Linsenzentren nachzubilden und das Erfassungsfeld zu simulieren.

Abb. 6.10 zeigt im oberen Teil die Abwicklung einer zylindrischen Linse. In der Abwicklung müssen zunächst die Eckpunkte und die Zentren der einzelnen Linsen definiert werden. Zur Vereinfachung dieses Schrittes ist es möglich, die Abwicklung automatisch anhand der Sensoranordnung, der Gesichtsfeldspezifikationen und einiger allgemeiner Vorgaben für die Linse (Positionierung und Orientierung im 3D-Raum (s.u.), Biegeradius der Gesamtfläche, Gesamtabwicklungsbreite, Höhe der Linsenebenen und Anzahl der Linsen pro Ebene) zu initialisieren. Bei der Initialisierung erfolgt die Berechnung der Linsenzentren und -breiten. Weiterhin werden auch die Brennweiten zur Fokussierung der Strahlung auf die Sensoren ermittelt. Die Berechnungsergebnisse werden in einer Datei gespeichert.

Im Anschluss an eine Initialisierungs- oder Optimierungsphase erfolgt die Erfassungsfeldsimulation, welche die Koordinaten der Linsen im 3D-Raum und deren Zuordnung zu den Sensorelementen erfordert.

Ausgehend von der 2D-Abwicklung werden daher zunächst die 3D-Koordinaten der Eckpunkte und der Linsenzentren unter Beachtung des Biegeradius R der Linse bestimmt. Die Linse wird dabei so angeordnet, dass der in Abb. 6.10 (oben) markierte Punkt in der Abwicklung dem Punkt $(R, 0, 0)$ im Weltkoordinatensystem entspricht. Die z -Achse fällt in diesem Fall mit der Längsachse des Zylinders zusammen.

Zur Positionierung und Ausrichtung der Linse ist zusätzlich eine Verschiebung in z -Richtung, eine Neigung gegen die z -Achse (Drehung um die y -Achse) sowie eine Drehung um die z -Achse vorgesehen.

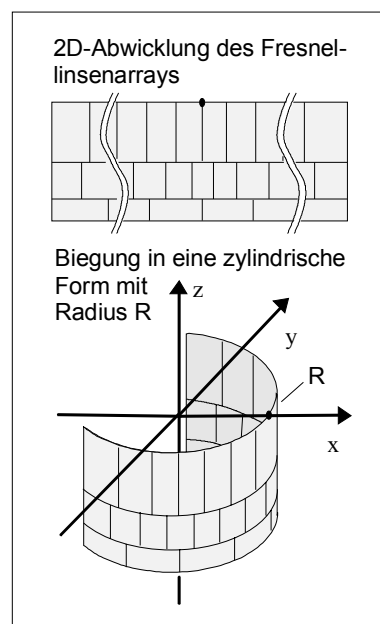


Abb. 6.10: Abwicklung, Biegung, Positionierung eines Fresnellinsenarrays

6.1.3 Initialisierung und Optimierung eines Fresnellinsenarrays

Zur Erzielung der Reichweite D für tangentielle Bewegungsrichtungen (Quergehen) ist eine gewisse Überdimensionierung des Erfassungsfeldes erforderlich.

In der Praxis hat sich die in [Abb. 6.11] angegebene Gleichung zur Berechnung der Überreichweite E für einen typischen Doppелеlementdetektor in Verbindung mit Linsenbrennweiten von 20... 30 mm bewährt.

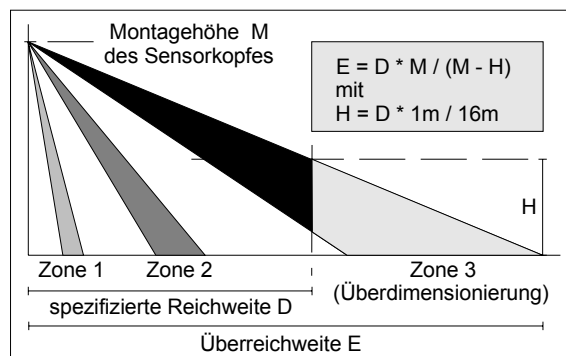


Abb. 6.11: Überdimensionierung eines Erfassungsfeldes zur Erzielung der spezifizierten Reichweite

Für eine spezifizierte Reichweite $D = 16$ m ergibt sich mit der oben angegebenen Dimensionierungsvorschrift eine Höhe der äußeren Bereichsgrenze von $H = 1$ m. Bei einer Montagehöhe $M = 2.5$ m folgt daraus die Überreichweite $E = 26.67$ m. Für die Reichweite $D = 8$ m ist dagegen nur noch eine Überreichweite von $E = 10$ m erforderlich. Das Simulationstool berücksichtigt die Überdimensionierung während der erstmaligen Berechnung eines Fresnellinsenarrays. Dieser Vorgang wird im folgenden als "Initialisierung" bezeichnet [Abb. 6.12].

Zur automatischen Berechnung müssen die Sensorkoordinaten im Weltkoordinatensystem, die Zuordnung der Schaltpunkte aus der Erfassungsfeldspezifikation sowie die allgemeinen Daten der Linse bekannt sein. Die Schaltpunkte ergeben sich aus der Montagehöhe und der Reichweitenspezifikation des Bewegungsmelders sowie aus der Linsenanzahl in den verschiedenen Linsenebenen. Um ein möglichst homogenes Schaltverhalten zu erzielen, sollte eine möglichst gleichmäßige Verteilung von überwachten und nicht überwachten Bereichen erfolgen.

Pro Sensorarray und pro Linse muß jeweils ein Schaltpunkt spezifiziert sein, die restlichen Schaltpunkte ergeben sich aus der geometrischen Anordnung der einzelnen Sensorelemente und der Linsen. Neben der Ermittlung der initialen Lage der Linsenzentren berechnet das Simulationstool auch die Breite der Linsen in Abhängigkeit der Einfallrichtung der Strahlung. Weiter außen gelegene Linsen sind demzufolge breiter als die Linsen, die sich unmittelbar vor den Detektoren befinden.

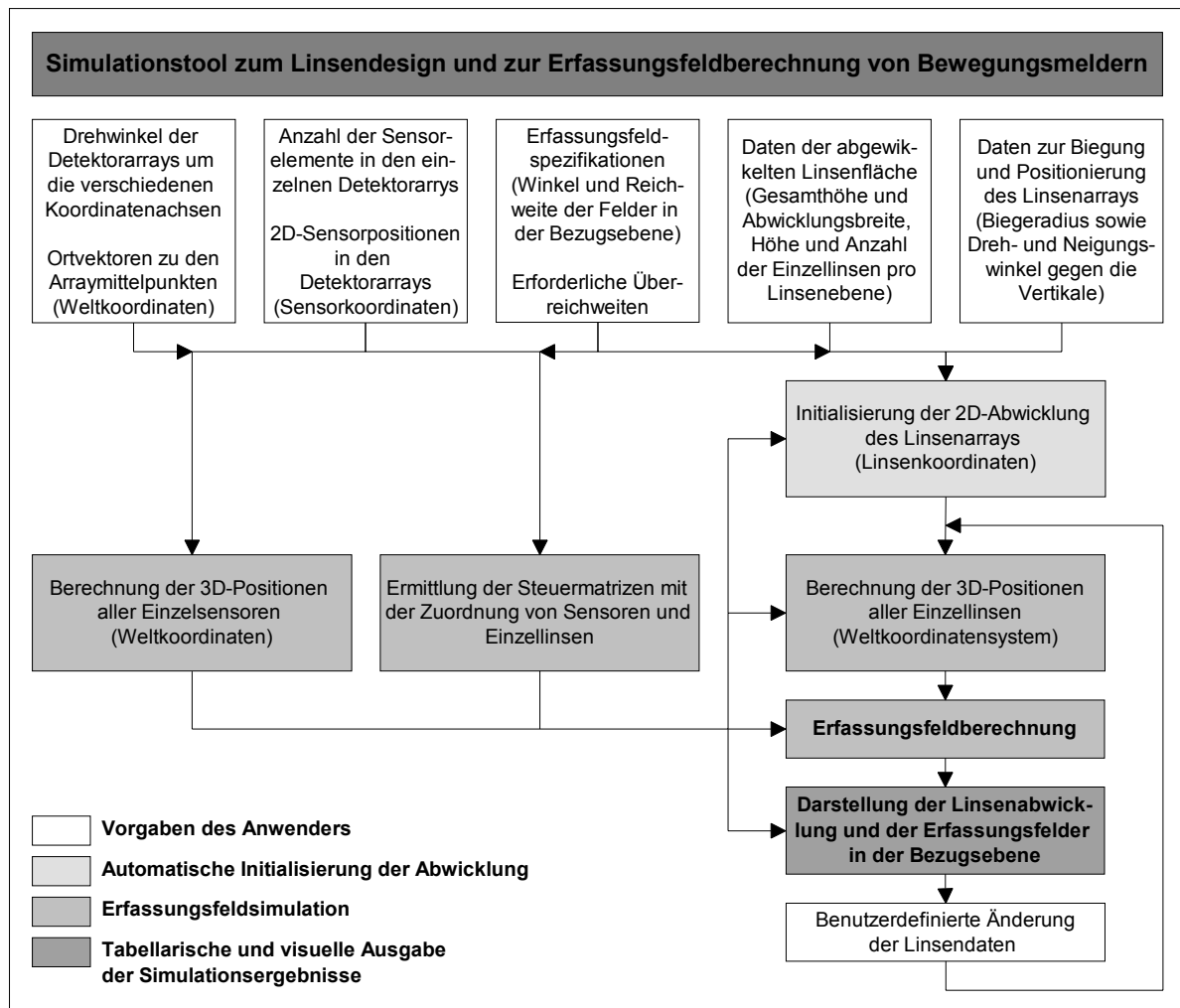


Abb. 6.12: Vorgehensweise zum Entwurf und zur Optimierung eines Fresnellinsenarrays

Im Anschluss an die Initialisierungsphase liegen die Abwicklungsdaten des Linsenarrays und die zur Fokussierung erforderlichen Brennweiten der Einzellinsen vor. Des weiteren ist zu

diesem Zeitpunkt auch die Zuordnung zwischen den einzelnen Sensorelementen und den Linsen bekannt. Ausgehend von diesen Daten erfolgt die Berechnung der 3D-Eckkoordinaten und der Zentren der Einzellinsen, bevor die Erfassungsfeldsimulation gemäß dem in Kap. 6.1.1 beschriebenen Verfahren durchgeführt wird.

Abschließend werden die Projektionen der aktiven Elemente in die Bezugsebene sowie die 2D-Abwicklung des Fresnellinsenarrays grafisch dargestellt. Ausgehend von diesen Darstellungen [Beispiel s. Kap. 6.1.4] lassen sich weitergehende Optimierungen ohne großen zeitlichen Aufwand durchführen. In der Datei mit den Abwicklungsdaten kann beispielsweise die Zentrenlage oder die Linsenbrennweite zur Einstellung einer gezielten Defokussierung modifiziert werden. Mit einem erneuten Simulationsdurchlauf lassen sich die Auswirkungen auf das Erfassungsfeld prüfen.

6.1.4 Beispiel zur Erfassungsfeldsimulation

Im folgenden werden zwei Beispiele zur Anwendung des Simulationstools vorgestellt. In Kap. 6.1.4.1 wird ein Fresnellinsenarray für einen 270°-Wächter mit drei Doppelementsensoren entworfen, während in Kap. 6.1.4.2 das optische System eines 90°-Melders mit einem 4x2-PIR-Array dimensioniert wird.

6.1.4.1 270°-Wächter mit drei Doppelementdetektoren

Ausgangspunkt zum Linsenentwurf ist eine Beschreibung der verwendeten Detektoren und die Spezifikation des gewünschten Erfassungsfeldes.

Im folgenden werden drei konventionelle Doppelementdetektoren mit einer aktiven Fläche von 1 mm x 2 mm pro Element und einem Spacing von 1 mm verwendet. Nach Abb. 6.13 ist jedem Detektor ein 90°-Sektor zugeordnet. Die Sensoren sind vertikal mit einem Drehwinkel von -90°, 0° und 90° um die z-Achse angeordnet. Im folgenden werden die Sektoren in dieser Reihenfolge mit Sektor 1, 2, 3 bezeichnet.

Die in Abb. 6.13 angegebenen Entfernungen entsprechen der gewünschten Schaltreichweite der äußeren Zone. Unterteilt man das gesamte Erfassungsfeld in 3 Entfernungszonen gemäß Abb. 6.11, ist bei der gewählten Vorgehensweise der Entwurf eines Fresnellinsenarrays mit 3 Linsenebenen erforderlich.

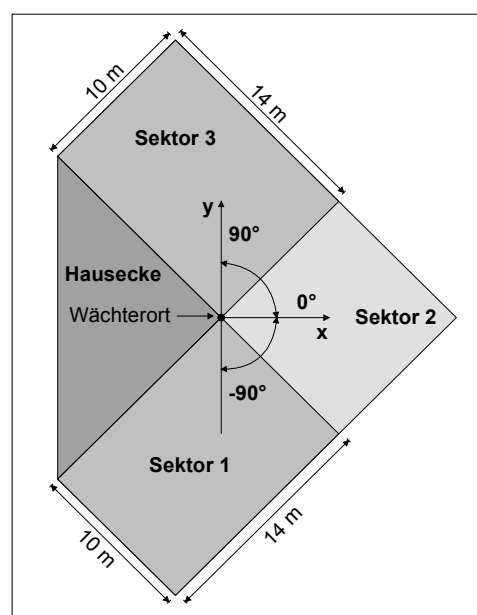


Abb. 6.13: Erfassungsfeldspezifikation eines 270°-Wächters zur Eckmontage

Die Reichweite der inneren Zone (Zone 1) und der mittleren Entfernungzone (Zone 2) wurde während des Designs aus den Vorgaben für die Maximalreichweite berechnet. Zone 1 wurde für die 0.2-fache und Zone 2 wurde für die 0.5-fache Maximalreichweite dimensioniert.

Die Längsachse der zylindrischen Gesamtlinsenfläche stimmt mit der z-Achse überein, der Biegeradius der Linsenfläche soll 35 mm betragen. Insgesamt steht eine Gesamtabwicklungsbreite von 164.94 mm zur Verfügung, für jeden 90°-Sektor folgt daraus eine Breite von 54.98 mm. Die Gesamthöhe der Abwicklung beträgt 30 mm, die folgendermaßen auf die drei Linsenebenen verteilt wird: Linsenhöhe: 5 mm (Zone 1), 10 mm (Zone 2), 15 mm (Zone 3). Um nicht zu große Zwischenräume im weitesten Entfernungsbereich zu erhalten, sollten mindestens 7 Einzelfelder pro Sektor in Zone 3 vorhanden sein. Nichtüberwachte Abschnitte der Zone 3 werden von den Feldern des mittleren Entfernungsbereichs (Zone 2) überwacht. Ein Unterkriechschutz oder eine Rückfeldüberwachung sind bei diesem Beispiel nicht vorgesehen. Für das Array wurden folgende Linsenanzahlen pro Ebene verwendet:

Linsenanzahl/Ebene	Entfernungzone 1	Entfernungzone 2	Entfernungzone 3
Sektor 1	5	6	7
Sektor 2	7	8	7
Sektor 3	5	6	7

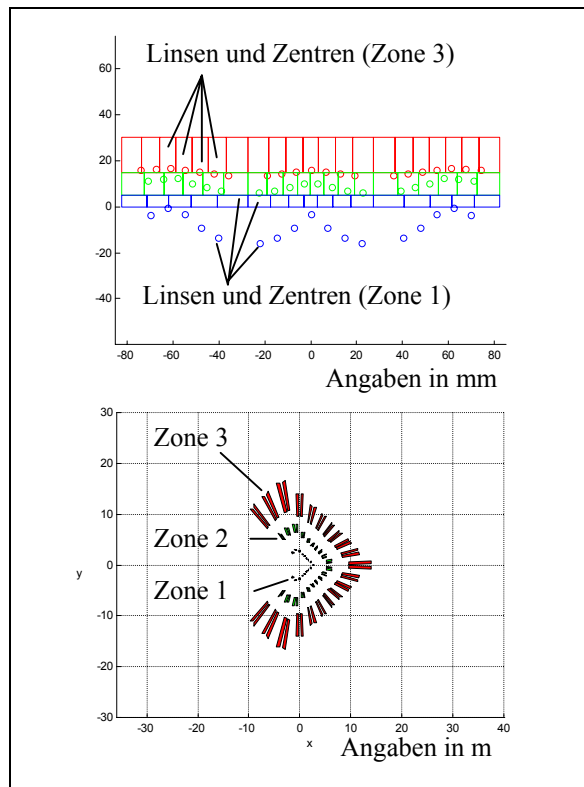


Abb. 6.14: Abwicklung des Fresnellinsenarrays und Erfassungsfeld ohne Überdimensionierung

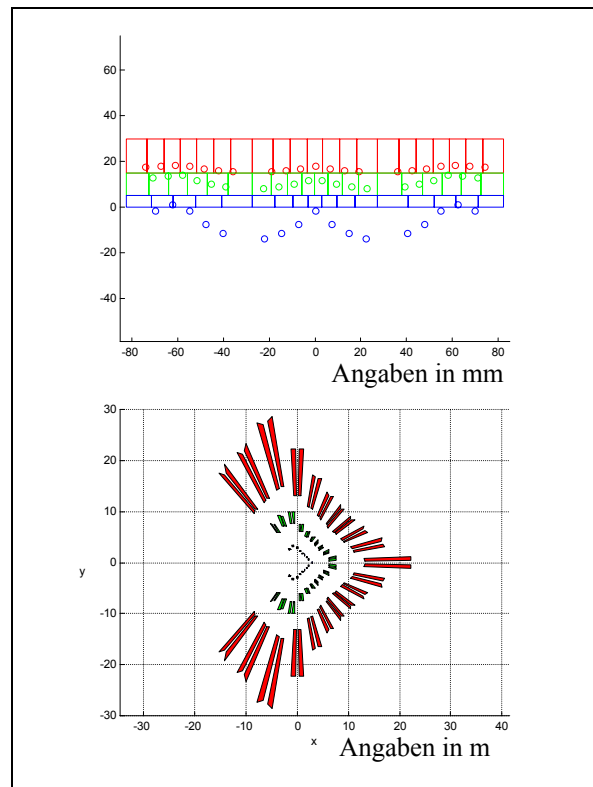


Abb. 6.15: Abwicklung des Fresnellinsenarrays und Erfassungsfeld mit Überdimensionierung

Abb. 6.14 und Abb. 6.15 zeigen die Abwicklung der mit dem Simulationstool berechneten Fresnellinsenarrays, wobei im ersten Fall auf die Überdimensionierung des Erfassungsfeldes bewusst verzichtet wurde. Die in Verbindung mit dieser Linse tatsächlich erzielbaren Schaltreichweiten liegen deutlich unterhalb der Spezifikationen. Mit dem optischen System nach Abb. 6.15, bei dem die Linsenzentren infolge der Überdimensionierung geringfügig "höher" liegen, können die spezifizierten Reichweiten eingehalten werden, allerdings vergrößert sich die Reichweite mit abnehmender Hintergrundtemperatur deutlich, so daß gegebenenfalls Maßnahmen zur elektronischen Reichweitenstabilisierung getroffen werden müssen.

An dieser Stelle soll noch einen weiteren Nachteil dieses Linsensystems erwähnt werden. Aufgrund der unsymmetrischen Feldverteilung in den Sektoren 1 und 3 ist das Störverhalten gegenüber großflächigen Änderungen der Hintergrundtemperatur schlechter als in Sektor 2.

6.1.4.2 90°-Wächter mit einem 4x2-PIR-Array

Als zweites Beispiel zur Anwendung des Simulationstools wird ein Fresnellinsenarray für ein 90°-Erfassungsfeld mit einer maximalen Reichweite von 8 m dimensioniert.

Als Sensor wird das Detektorarray nach Abb. 6.16 verwendet, wobei horizontal benachbarte Elemente jeweils parallel geschaltet sind.

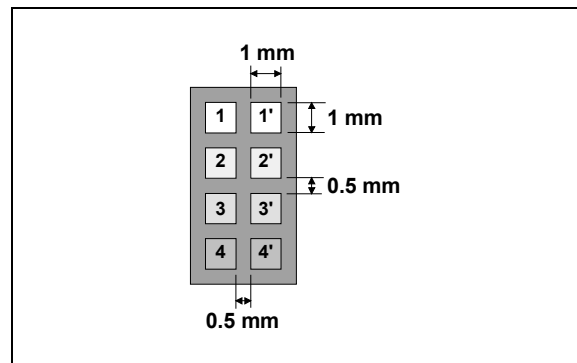


Abb. 6.16: Sensoranordnung im 4x2-PIR-Array

Die geometrische Anordnung des Arrays und der Linse soll anhand von Abb. 6.17 verdeutlicht werden. Der Biegeradius der Linse beträgt 50 mm, der Sensor ist in einem Abstand von 30 mm zur Linse platziert.

Es soll wiederum eine 3-Zonen-Linse erzeugt werden, wobei für den Nahbereich eine Linsenhöhe von 5 mm, für die mittlere Entfernungzone eine Höhe von 15 mm und für den weitesten Entfernungsbereich eine Linsenhöhe von 25 mm gewählt wird.

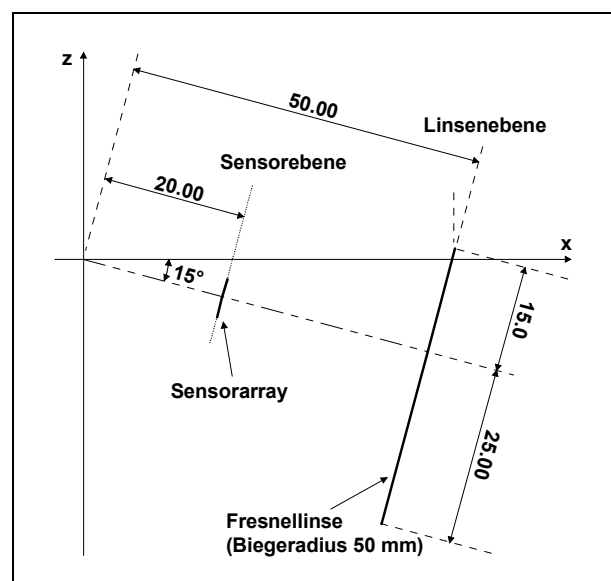


Abb. 6.17: Sensor-Linse-Konfiguration

Abb. 6.18 zeigt die Abwicklung des Fresnellinsenarrays und das Erfassungsfeld des Wächters unter Beachtung der Überdimensionierung nach Abb. 6.11. Verschiebt man den Sensor in positiver z-Richtung (vgl. Abb. 6.17), kann die Reichweite des Bewegungsmelders eingestellt werden. Abb. 6.19 zeigt die Reichweitenbegrenzung bei einer Verschiebung des Detektorarrays um 2.5 mm und 7.5 mm. Im Vergleich zu Abb. 6.18 erkennt man, dass sich die Projektionen der Sensorelement teilweise überlappen. Dieser Effekt ist auf die mit der Verschiebung anwachsende Defokussierung der Linsen zurückzuführen.

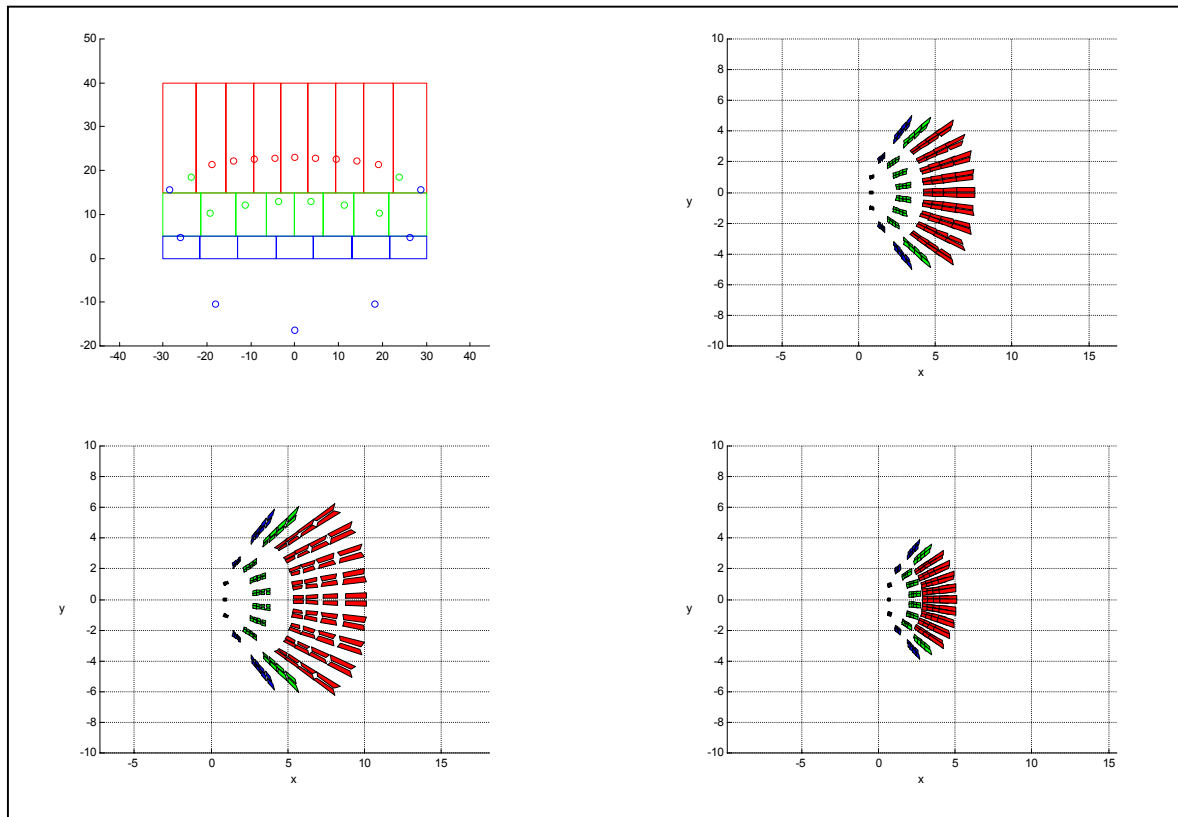


Abb. 6.18: Abwicklung des Fresnellinsenarrays und Erfassungsfeld für die Anordnung nach Abb. 6.17

Abb. 6.19: Reichweitereinstellung durch Verschiebung des PIR-Arrays um +2.5 mm und +7.5 mm in z-Richtung

6.1.5 Anwendungsmöglichkeiten des Simulationstools

In diesem Abschnitt sollen die wichtigsten Anwendungsmöglichkeiten des Simulationstools zusammengefasst werden. Im Vordergrund stand zunächst die Dimensionierung von Fresnellinsenarrays für Bewegungsmelder und die Simulation des zugehörigen Erfassungsfeldes. Infolge der definierten Vorgehensweise beim Linsenentwurf, der per Software durchführbaren Systemoptimierung und der Visualisierung wichtiger Teilergebnisse entstehen deutlich niedrigere Entwicklungskosten, kürzere Durchlaufzeiten bis zur Fertigstellung von Prototypen und ein geringerer Dokumentationsaufwand.

Insbesondere mit Hinblick auf die Fertigung von Prototypen liefert das Programm nicht nur die Abwicklungsdaten des Fresnellinsenarrays, sondern auch die auf das jeweilige Linsenzentrum bezogenen Koordinaten der Einzellinsen. Überträgt man diese Daten auf eine geeignete Vorlage (s. Abb. 6.20), die später auf den runden Linsenrohling aufgeklebt wird, lassen sich die Einzellinsen innerhalb kürzester Zeit zuschneiden.

Anschließend werden die Linsen unter Verwendung einer Vorlage der Gesamtabwicklung zusammengesetzt und verschweißt. Die zielscheibenartige Struktur in Abb. 6.20 erleichtert dabei das Aufkleben der Schneidvorlage. Selbst größere Linsenprototypen wie beispielweise in Abb. 6.14 lassen sich mit diesem Verfahren in weniger als einem Arbeitstag berechnen und herstellen.

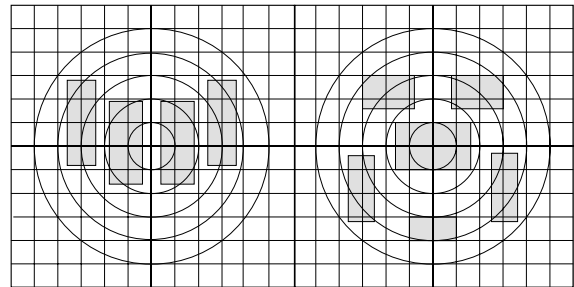


Abb. 6.20: Vorlage zum Aufkleben und Schneiden der Rohlinse

In Abb. 6.21 sind weitere Anwendungsmöglichkeiten des Programms dargestellt. Neben der Berechnung von Linsenarrays kann auch die geometrische Anordnung der Sensorelemente bei konstanten Linsenparametern variiert werden. Mit Hinblick auf anwenderspezifische Detektorarrays besteht daher die Möglichkeit, die Position und die Form der Einzelemente sowie deren Auswirkung auf das Erfassungsfeld vorab zu prüfen.

Bei Neuentwicklungen liefert die Simulation wichtige Erkenntnisse zur Optimierung des optischen Systems von PIR-Bewegungsmeldern, die beim Wächterdesign und bei der Gehäusekonstruktion berücksichtigt werden sollten. Im umgekehrten Fall können auch Restriktionen infolge eines bereits vorhandenen Wächtergehäuses in der Simulation beachtet werden.

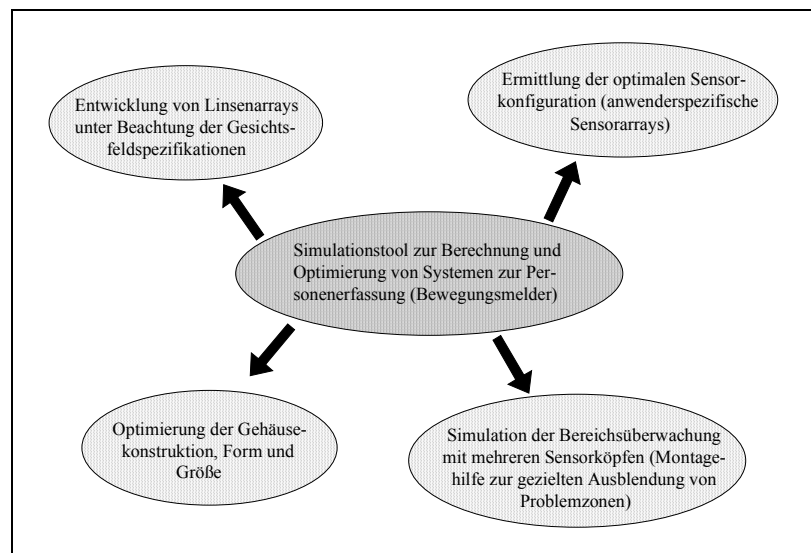


Abb. 6.21: Anwendungsmöglichkeiten des Simulationstools zur Erfassungsfeldberechnung

Ferner kann das Programm als Grundlage zur Simulation der Bereichsüberwachung mit mehreren Sensorköpfen oder als Montagehilfe zur Ausrichtung und zur gezielten Ausblendung von Problemzonen genutzt werden. Das Anwendungspotential des Simulationstools ist dabei

nicht auf den Bereich der PIR-Bewegungsmelder beschränkt. Durch geeignete Erweiterungen ist eine Übertragung auf weitere aktive und passive optische Sensorkonzepte möglich.

6.2 Aufzeichnung und Verarbeitung von Wächtersignalen

Bei der Auswertung der Signale von PIR-Bewegungsmeldern erschwert die Vielzahl der möglichen, im Voraus nicht bekannten Störsignale die Entwicklung eines optimalen Signalverarbeitungskonzepts. In der Literatur findet man zahlreiche Verfahren, von denen in Kapitel 7 eine kleine Auswahl präsentiert wird. Generell zeigt sich die Tendenz zu mikrocontrollerbasierten Systemen, die im Vergleich zur konventionellen Schwellwertauswertung deutlich komplexere Auswertestrategien und damit eine bessere Trennung von Stör- und Nutzsignalen ermöglichen.

Eine wesentliche Grundlage der Algorithmentwicklung ist dabei der objektive Vergleich von Signalverarbeitungskonzepten, die teilweise sehr unterschiedliche Signaleigenschaften auswerten. Zur Beurteilung der Qualität einzelner Auswerteverfahren und zur Ermittlung der jeweils optimalen Parametereinstellung bietet sich einerseits die erzielbare Schaltreichweite für Nutzsignale aus unterschiedlichen Gehrrichtungen (parallel, tangential, radial) sowie andererseits die Fehlalarmanzahl über einen längeren Untersuchungszeitraum an. Wichtig ist dabei, dass der Vergleich der Auswerteverfahren bzw. Parametereinstellungen anhand identischer Bewegungsmeldersignale erfolgt, was die Speicherung von digitalisierten Wächterdaten und eine nachträgliche Möglichkeit zur Offline-Auswertung erfordert.

Im folgenden wird das im Rahmen dieser Arbeit verwendete Testsystem zur Wächterentwicklung vorgestellt. In Abb. 6.22 sind die wichtigsten Eigenschaften zusammengefasst. Grundlage des Entwicklungssystems ist eine zuverlässige Datenaufnahme mit einem Prototyp des Bewegungsmelders und die gesicherte Datenübertragung zum PC. Durch die Verwendung der Zielhardware ist sichergestellt, dass das dynamische Systemverhalten des Wächters bereits während der Softwareentwicklung berücksichtigt wird.

Während der Datenaufzeichnung fallen sehr große Datenmengen an. Bei einem Abtastintervall von 10 ms treten beispielsweise pro Stunde 360000 Datensätze auf, pro Tag liefert jeder Wächter insgesamt 8640000 Datensätze. Um einzelne Signalverarbeitungskonzepte zu prüfen und Parametereinstellungen zu optimieren, müssen diese Daten dauerhaft gespeichert und später offline ausgewertet werden. Das Entwicklungssystem enthält daher Funktionen zur Speicherung und zur automatischen Verwaltung dieser Datenmengen. Zusätzlich sind im Anschluss an die Offline-Bewertung der Rohdaten verschiedene Ausgabe- und Analysemöglichkeiten zur Algorithmusprüfung und zur Beurteilung des Schaltverhaltens vorgesehen.

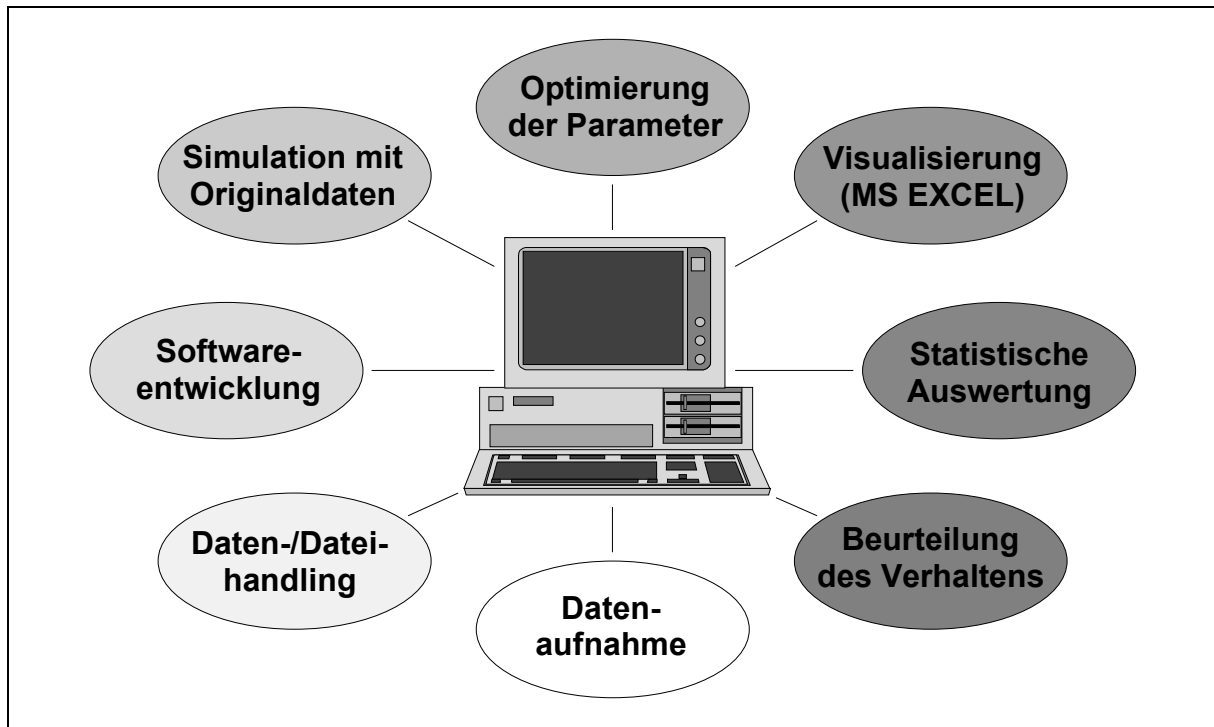


Abb. 6.22: Eigenschaften und Anwendungspotential des Systems zur Wächterentwicklung

6.2.1 Datenaufnahme, Kommunikationsprotokoll und Dateihandling

Ein Bewegungsmelder mit einer speziellen Software zur Datenaufnahme und -übertragung wird im folgenden als Datenlogger bezeichnet. Um von einem Datenlogger digitalisierte, aber falsch zum PC übertragene Werte von der Auswertung auszuschließen, wurde auf beiden Seiten ein einfaches Kommunikationsprotokoll zur Erkennung von Übertragungsfehlern implementiert. Aufgrund der Seltenheit von Übertragungsfehlern werden diese Werte auf der PC-Seite verworfen.

Der grundsätzliche Aufbau einer Message ist in Abb. 6.23 dargestellt. Beginnend mit dem Startbyte werden die einzelnen Zeichen einer Message über eine RS232-Schnittstelle übertragen. Pro Zeichen werden 1 Startbit, 8 Datenbits, kein Paritätsbit und 1 Stopbit mit einer Übertragungsrate von 57600 Baud gesendet. Bestimmte Bitkombinationen sind dabei für Steuerzeichen reserviert. Falls ein zu übertragendes Zeichen den Wert eines Steuerzeichens besitzt, wird dieses Zeichen in einen anderen Wert transponiert. Durch die Transponierung vergrößert sich die Länge der Message, PC-seitig wird die Transponierung wieder rückgängig gemacht. Die größere Message-Länge muss bei der Berechnung der Übertragungszeiten berücksichtigt werden.

Nach dem Startzeichen folgt ein fortlaufend aktualisierter Index für den Datensatz. Die Anzahl der ggf. durch Übertragungsfehler verlorenen Datensätze kann anhand eines Sprunges in diesem Index ermittelt werden. Die unmittelbar folgende Kommandokennung (s. Abb. 6.23) quittiert den korrekten Empfang des zuvor vom PC gesendeten Befehls, anschließend werden

die 5 Datenworte D0... D4 übertragen. Jedes Datenwort besteht dabei aus einem High-Byte und einem Low-Byte. Ohne Transponierung ist das Datenfeld mindestens 10 Byte lang. An das Ende der Message wird eine Checksumme über die gesamte Nachricht angehängt, die bei Bedarf ebenfalls transponiert wird.

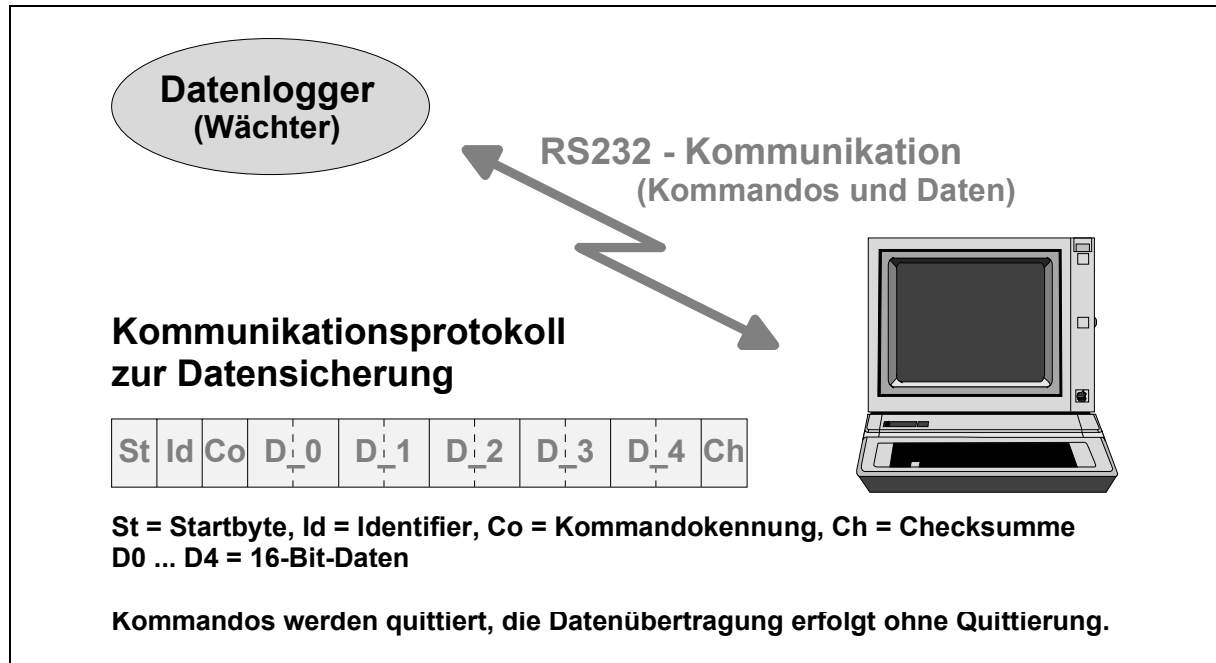


Abb. 6.23: Kommunikationsprotokoll zur Datenübertragung zwischen Sensor und PC

PC-seitig ist das Startzeichen, der Identifier, die Quittierung des zuvor übermittelten Kommandozeichens und die Länge der erwarteten Message ohne Transponierungen bekannt. Mit diesen Informationen können die vom Datenlogger eintreffenden Nachrichten dekodiert werden.

Im Anschluss an die Dekodierung erfolgt entweder die Speicherung der Sensordaten oder die Online-Auswertung mit einer benutzerdefinierten Signalverarbeitungsstrategie. In der gegenwärtigen Form der Implementierung enthalten die Datenworte D0... D4 die mit einer Auflösung von 10-Bit digitalisierten Signale von drei 3 PIR-Kanälen, ein Helligkeitssignal und verschiedene Potentiometereinstellungen, mit denen das Betriebsverhalten des Wächters beeinflusst werden kann.

Alternativ zur Online-Auswertung können auch die gespeicherten Rohdaten als Input für die Signalverarbeitung dienen. Bei dieser Form der Offline-Auswertung werden die Dateien in der Reihenfolge ihres Entstehens abgearbeitet. Für die Signalverarbeitung besteht kein Unterschied, ob die Daten vom Datenlogger geliefert oder aus einer Datei eingelesen werden.

Das gesamte Dateihandling erfolgt automatisch. Bei der Datenaufnahme wird PC-seitig eine mit MS EXCEL handhabbare Dateigröße mit maximal 16384 Datensätzen überwacht. Bei Bedarf wird eine neue Datei angelegt. Die Dateien werden dabei in der Reihenfolge ihres Ent-

stehens nummeriert. Maximal können 10000000 Dateien mit jeweils 16384 Datensätzen automatisch erzeugt bzw. eingelesen werden.

6.2.2 Softwareentwicklung, Auswertung und Ergebnisausgabe

Die gesamte Entwicklungsumgebung wurde mit MS Visual C erstellt. Erweiterungen und Änderungen im Zuge der Softwareentwicklung sind daher ohne großen Zeitaufwand möglich. Ein wesentlicher Vorteil der Hochsprachenimplementierung ist die Möglichkeit zur Untersuchung und zur Beurteilung des Wächterverhaltens, bevor die endgültige Implementierung auf dem Mikrocontroller des Bewegungsmelders erfolgt. Entwicklungs- und Testzeiträume lassen sich bei dieser Vorgehensweise deutlich reduzieren.

Aufbauend auf dem zuvor beschriebenen Konzept zum Datenaustausch und zur Dateiverwaltung, das für alle Applikationen gleich ist, können die implementierten Signalverarbeitungsstrategien und die Auswirkungen unterschiedlicher Parametereinstellungen im Offline-Betrieb bewertet werden. Die Vorgabe der Parameter erfolgt menugesteuert unter Beachtung der zulässigen Grenzwerte.

Zur Extraktion charakteristischer Signalverläufe und zur Ausgabe von algorithmusspezifischen Bewertungsergebnissen sind verschiedene Alternativen vorgesehen. Bei einem Alarmereignis werden im einfachsten Fall nur die Sensorsignale in einer separaten Datei gespeichert. Der extern vorgebbare Aufzeichnungszeitraum teilt sich dabei jeweils zur Hälfte auf die Zeit vor und nach der Alarmauslösung auf. Für den Fall der Online-Auswertung von Rohdaten ergibt sich mit dieser Option eine erhebliche Datenreduktion im Vergleich zur kontinuierlichen Datenaufnahme, im Offline-Betrieb werden dagegen die dem Alarm zugrundeliegenden Signalverläufe automatisch aus den Rohdaten extrahiert.

Neben der Signalspeicherung im Alarmfall ist es weiterhin möglich, das gesamte dynamische Verhalten der Algorithmen, z. B. eine adaptive Bezugspegel- oder Schwellwertnachführung, in der korrekten zeitlichen Zuordnung zu den Rohdaten zu protokollieren oder algorithmusspezifische Signaleigenschaften, wie beispielsweise Signalgradienten, Pulsdauern etc. zu extrahieren. Analog zur Dateiüberwachung beim Speichern von Rohdaten erfolgt auch für die Ausgabedateien die Überwachung und Begrenzung der Dateigröße.

Speziell für Gehtests im Online-Betrieb wurde eine weitere Option zur wirklichkeitsgetreuen Nachbildung des Wächterverhaltens implementiert. Unter Kontrolle des Entwicklungssystems kann die Initialisierungszeit bei der Inbetriebnahme, die Einschaltdauer nach einem Alarm sowie die Immunitätszeit nach dem Ausschalten der Lampe vorgegeben und damit der Schaltausgang des Datenloggers gezielt beeinflusst werden. Optional ist auch eine adaptive Verlängerung der Alarmdauer und der Immunitätszeit möglich.

6.2.3 Parameteroptimierung und Beurteilung des Schaltverhaltens

Die Offline-Auswertung der Wächterdaten ermöglicht die Simulation des Systemverhaltens unter realen Testbedingungen und liefert Aussagen über die Qualität der verwendeten Unterscheidungsmerkmale zwischen Stör- und Nutzsignalen. Mit dieser Form der Auswertung lassen sich daher nicht nur Parameteroptimierungen innerhalb kürzester Zeit durchführen, sondern es sind auch gezielte Anpassungen des zugrundeliegenden Signalverarbeitungskonzeptes in einer höheren Programmiersprache möglich. Die mit den ursprünglichen Signalverläufen zeitsynchrone Ausgabe von internen Zuständen der Signalverarbeitung reduziert dabei den gesamten Entwicklungs- und Prüfaufwand deutlich.

Typische Stör- und Nutzsignalklassen lassen sich mit statistischen Methoden aus den von der Signalverarbeitung gelieferten Entscheidungsmerkmalen ableiten. Neben Histogrammen zur allgemeinen Häufigkeitsverteilung der bewerteten Signaleigenschaften ergeben sich insbesondere aus deren Verbundverteilungen wichtige Ansätze zur Verbesserung des Schaltverhaltens, wie in den nächsten Kapiteln gezeigt wird.

Die Fehlalarmrate bei "reinen" Störsignalen und das Schaltverhalten bei "reinen" Nutzsignalen liefern die grundsätzlichen Aussagen zur Beurteilung des Bewegungsmelders. Während die Fehlalarmanzahl anhand typischer Störsignale aus einem realen Testfeld ermittelt wird, wurden die Nutzsignale in einem speziellen, von allen Störquellen bereinigten Prüffeld aufgezeichnet. Entwicklungsbegleitend entstand daher eine Sammlung von kritischen Störsignalen, die als Grundlage zur Untersuchung der später beschriebenen Signalverarbeitungskonzepte dient.

6.3 Literatur

- [Lit. 6.1] E. Hecht, A. Zajac: "Optics", Addison-Wesley Publishing Company, Inc., 1974
- [Lit. 6.2] F. L. Pedrotti et al.: "Introduction To Optics", 2nd Edition, Prentice Hall International, Inc., 1993
- [Lit. 6.3] Bergmann-Schaefer: "Lehrbuch der Experimentalphysik", Band III - Optik, 6. Auflage, Walter de Gruyter & Co., 1974
- [Lit. 6.4] M. Alonso, E. J. Finn: "Physik", Addison-Wesley Verlag (Deutschland) GmbH, 1977

7 Konventionelle Signalverarbeitungskonzepte

7.1 Signale aus unterschiedlichen Bewegungsrichtungen

Abb. 7.1 zeigt ein typisches Bewegungsmeldersignal am Ausgang der analogen Verstärkerstufe. Das Analogsignal wurde in äquidistanten Zeitschritten von $T = 10 \text{ ms}$ (Abtastfrequenz $F = 100 \text{ Hz}$) mit einem 10-Bit-AD-Wandler digitalisiert. Bei einem Eingangsspannungsbereich des AD-Wandlers von $0 \dots 5 \text{ V}$ ergibt sich das in Abb. 7.1 angegebene Quantisierungsintervall $Q = 5 \text{ V} / (2^{10} - 1) \approx 4.89 \text{ mV}$. Im Anschluss an den Abtastvorgang wurde das digitalisierte Signal mit dem Skalierungsfaktor $K = 32$ multipliziert. Die Skalierung entspricht einer Linksverschiebung des ursprünglichen 10-Bit-Abtastwertes um 5 Bit, wodurch sich bei der später beschriebenen Form der Mittelwertbildung eine höhere Rechengenauigkeit in Verbindung mit Fixed-Point-Divisionen erzielen lässt.

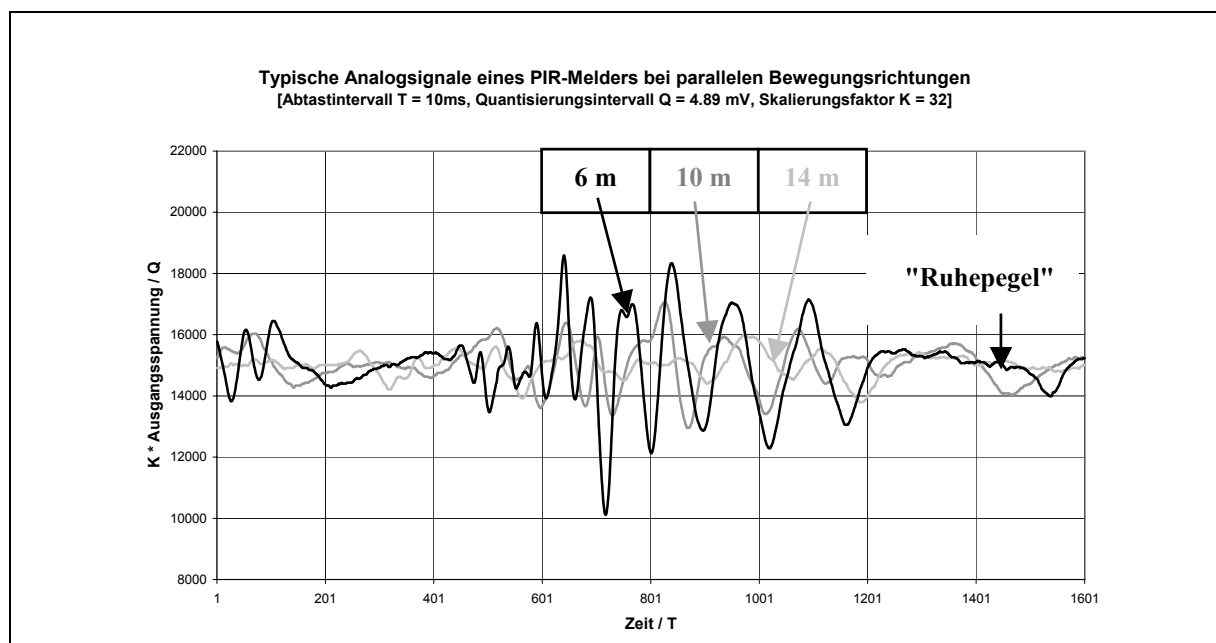


Abb. 7.1: Typisches Ausgangssignal eines PIR-Bewegungsmelders bei paralleler Bewegungsrichtung in Abständen von ca. 6 m, 10 m und 14 m (Gehgeschwindigkeit ca. 1 m/s).

Mit zunehmender Reichweite vom Bewegungsmelder werden die Amplituden der "Quergeräusche" nach Abb. 7.1 kleiner, was auf die Verringerung der einfallenden Strahlungsleistung gemäß dem Entfernungsgesetz nach Kap. 3 zurückzuführen ist. Die im Signalverlauf erkennbaren Schwingungen entstehen, wenn die antiparallel geschalteten Elemente eines Dualsensors sequentiell der von einer Person emittierten Temperaturstrahlung ausgesetzt werden.

Die ausgeprägte Strahlungsmodulation der "Quergeräusche" ist bei dem in Abb. 7.2 dargestellten "Draufzugehsignal" nur im unmittelbaren Nahbereich vorhanden. Die Person bewegt sich radial in Richtung des Melders. In diesem Fall sind die von dem Sensor registrierten Strahlungsänderungen deutlich kleiner, da einzelne Körperpartien zunächst fast permanent auf den Detektor abgebildet werden. Mit zunehmender Nähe zum Bewegungsmelder verändert

sich der betrachtete Körperrausschnitt, wobei beim Übergang an den inneren Zonengrenzen des Erfassungsfeldes und durch die mit der Schrittfrequenz einhergehenden Pendelbewegungen des Körpers eine zusätzliche Strahlungsmodulation hervorgerufen wird. Erschwerend kommt allerdings hinzu, dass die einfallende Wärmestrahlung bei der gleichzeitigen Abbildung auf beide Sensorelemente nahezu vollständig kompensiert werden kann.

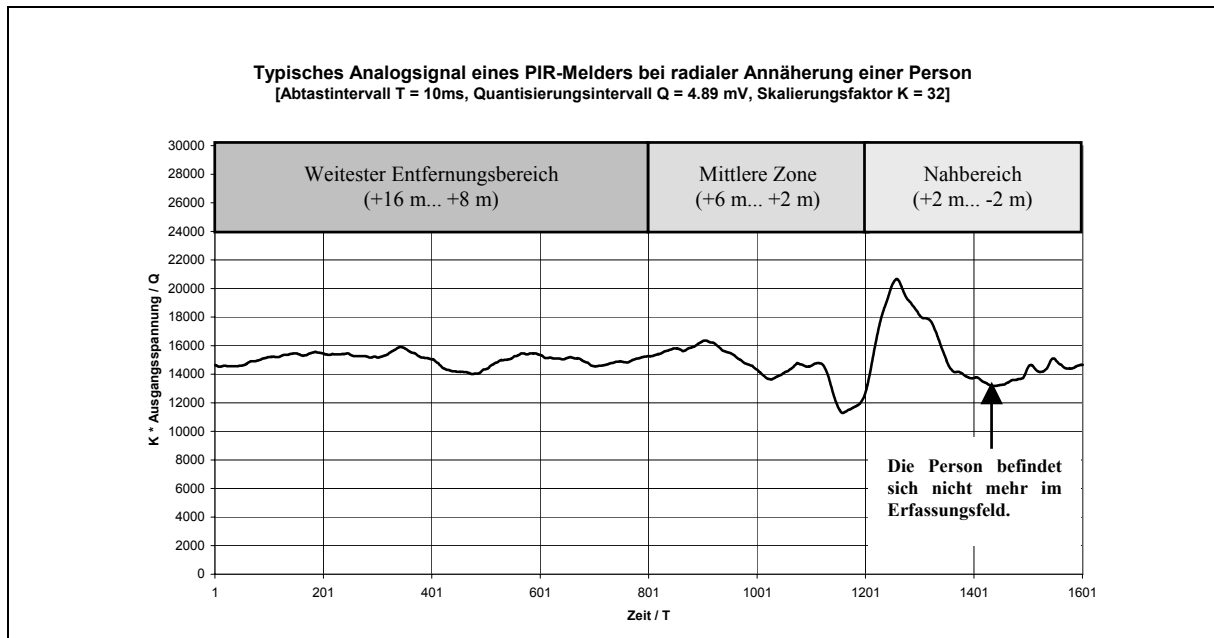


Abb. 7.2: Typisches Ausgangssignal eines PIR-Bewegungsmelders bei radialer Annäherung (Gehgeschwindigkeit ca. 1 m/s). Die im Bild angegebenen Abstände zeigen die im Vergleich zu Abb. 7.1 deutlich kleineren Signalamplituden beim "Draufzugehen".

Die Bewegungsrichtung hat einen entscheidenden Einfluss auf die Amplitude und auch auf die Form des Sensorsignals. Nur in unmittelbarer Sensornähe erhält man vergleichbare Signalamplituden für beide Bewegungsrichtungen. Je nach Bewegungsrichtung sind außerdem Mischformen beider Signalarten möglich.

Die Erzielung eines möglichst homogenen Schaltverhaltens im gesamten Erfassungsfeld bei gleichzeitiger Maximierung der Draufzugereichweite ist ein wesentliches Optimierungskriterium beim Bewegungsmelderdesign. Die je nach Bewegungsrichtung unterschiedlichen Signalamplituden erfordern unter Beachtung der zuvor genannten Ziele spezielle Auswertestrategien. Insbesondere die Steigerung der Draufzugereichweite bewirkt eine größere Empfindlichkeit gegenüber Störsignalen, so dass zusätzliche Unterscheidungskriterien zur Reduktion der Fehlalarmanfälligkeit bewertet werden müssen.

In den folgenden Abschnitten werden einige konventionelle Signalverarbeitungskonzepte für Bewegungsmelder vorgestellt. Eine wesentliche Voraussetzung für diese Untersuchungen ist die Zerlegung des Sensorsignals und die daraus abgeleitete Bewertung typischer Signaleigenschaften. Die Untersuchungen wurden ungeachtet der aktuellen patentrechtlichen Lage durch-

geführt. Dem Verfasser bekannte Offenlegungs- und Patentschriften sind an den betreffenden Stellen zitiert.

7.2 Referenzpegelnachführung und Schwellenauswertung

In [Pat. 7.1] wird ein Bewegungsmelder zur Raumüberwachung beschrieben, der neben der erforderlichen Optik mindestens zwei, im Gegentakt arbeitende Sensoren, einen Wechselspannungsverstärker sowie eine Auswerteschaltung und eine Alarmeinheit umfasst. Der Melder wird mit einer bipolaren Spannungsversorgung betrieben, so dass der Ruhepegel nahezu mit dem GND-Potential der angegebenen Schaltung übereinstimmt.

Die Auswerteschaltung besteht aus einem Fensterdiskriminator, der die positive bzw. negative Halbwelle des verstärkten Sensorsignals anhand eines oberen und unteren Schwellwerts digitalisiert. Die Schwellwerte sind konstant und nahezu symmetrisch zum Referenzpegel des Fensterdiskriminators, welcher sich adaptiv an das Ausgangssignal des Verstärkers anpasst.

Die Ausgangssignale des Fensterdiskriminators dienen zur Triggerung einer Zeitstufe, die ein Zeitfenster für die weitere Signalauswertung erzeugt. Tritt innerhalb des Zeitfensters am zweiten Ausgang des Fensterdiskriminators ebenfalls ein Signal auf, erfolgt die Alarmauslösung. Neben der durch die Schwellwerte festgelegten Mindestamplitude wird somit das Auftreten von zwei aufeinanderfolgenden Signalen mit entgegengesetzter Polarität und der zeitliche Abstand der zugehörigen Schwellwertüberschreitungen bewertet.

Eine weitere Verschärfung dieser Auswertekriterien wird in der Patentschrift DE 27 42 489 B2 vorgestellt [Pat. 7.1]. Dort dient eine Sequenz von mehreren, alternierenden positiven und negativen Impulsen als Referenzfolge für die Bewegungserkennung, was allerdings voraussetzt, dass zunächst eine gewisse Wegstrecke im Überwachungsbereich des Melders zurückgelegt wird. Das Verfahren bewirkt eine erhöhte Störsicherheit im Vergleich zu [Pat. 7.1], es hat allerdings Nachteile hinsichtlich eines schnellen Ansprechverhaltens im Alarmfall.

Ein wesentlicher Vorteil der zuvor erwähnten Referenzpegelnachführung ist die Unabhängigkeit der Auswertung von möglichen DC-Offsets im analogen Signalpfad. Diese Eigenschaft ist auch für einen mikrocontrollergesteuerten Bewegungsmelder durchaus erwünscht. Aus Kostengründen wird allerdings meist keine bipolare, sondern eine unipolare Spannungsversorgung verwendet. Das analoge Sensorsignal ist in diesem Fall mit einer konstanten Offsetspannung behaftet, die ungefähr der halben Betriebsspannung entspricht.

Im folgenden wird ein Verfahren zur Referenzpegeladaption und Schwellwertberechnung vorgestellt, das aufgrund seiner Einfachheit auch für Mikrocontroller mit geringer Rechenleistung geeignet ist. Ausgangspunkt für die Ermittlung des Referenzpegels $r_n(k)$ ist die Rechenvorschrift nach (Gl. 7.1), wobei der Zeitindex k dem auf die Abtastperiode T normierten Zeitschritt entspricht.

$$(Gl. 7.1) \quad r_n(k) = r_n(k-1) + \frac{s(k) - r_n(k-1)}{2^n} \quad \text{mit } n = 1, 2, 3$$

Die in (Gl. 7.1) verwendete Division des Differenzterms durch 2^n mit $n = 1, 2, 3, \dots$ kann durch eine Rechtsverschiebung des Differenzterms um n Bit realisiert werden, wodurch nur ein geringer Berechnungsaufwand für den Mikrocontroller anfällt. Unter Anwendung der Z-Transformation (s. [Lit. 7.1]) folgt aus (Gl. 7.1) zunächst die Beziehung

$$R_n(z) = R_n(z) \cdot z^{-1} + \frac{S(z) - R_n(z) \cdot z^{-1}}{2^n} \quad \text{mit } n = 1, 2, 3,$$

woraus sich die Z-Transformierte $R_n(z)$ des Bezugspegels $r_n(k)$ nach (Gl. 7.2) ableiten lässt.

$$(Gl. 7.2) \quad R_n(z) = \frac{1}{a_n - b_n \cdot z^{-1}} \cdot S(z) \quad \text{mit } a_n = 2^n, b_n = 2^n - 1 \quad \text{und } n = 1, 2, 3$$

Der Konvergenzbereich der Z-Transformierten beinhaltet den Einheitskreis in der komplexen Ebene, so dass die Übertragungsfunktion für $z = e^{j\omega}$ existiert. Mit der Substitution $z = e^{j\omega}$, wobei die Kreisfrequenz ω in Verbindung mit diskreten Signalen als Winkelschritt pro Abtastwert definiert ist, erhält man die komplexe Übertragungsfunktion $H_n(e^{j\omega})$ nach (Gl. 7.3).

$$(Gl. 7.3) \quad H_n(e^{j\omega}) = \frac{R_n(e^{j\omega})}{S(e^{j\omega})} = \frac{[a_n - b_n \cdot \cos(\omega)] - j \cdot [b_n \cdot \sin(\omega)]}{a_n^2 - 2 \cdot a_n \cdot b_n \cdot \cos(\omega) + b_n^2}$$

Da die Spektren diskreter Signale 2π -periodisch sind, ist eine Betrachtung im Intervall von $\omega = [-\pi, \pi]$ ausreichend. Dabei ist zu beachten, dass die höheren Frequenzen in der Nähe der ungeradzahlig Vielfachen von π auftreten, während niedrige Frequenzen bei geradzahlig Vielfachen von π einschließlich 0 vorhanden sind [Lit. 7.2]. Die Entnormierung der ω -Achse erfolgt mit der Abtastfrequenz F , die mit dem Winkelschritt 2π übereinstimmt. Entsprechend dem Nyquist-Kriterium muss die Abtastfrequenz mindestens doppelt so groß sein wie die höchste, vorkommende Signalfrequenz.

In Abb. 7.3 ist der Amplitudengang und der Phasengang der komplexen Übertragungsfunktion $H_n(e^{j\omega})$ nach (Gl. 7.3) dargestellt. Das Betragsamplitudenspektrum ist dabei durch (Gl. 7.4) und das Phasenspektrum durch (Gl. 7.5) gegeben.

$$(Gl. 7.4) \quad \left| H_n(e^{j\omega}) \right| = \left| \sqrt{\frac{1}{a_n^2 - 2 \cdot a_n \cdot b_n \cdot \cos(\omega) + b_n^2}} \right| \quad (Gl. 7.5) \quad \varphi(e^{j\omega}) = \arctan \left(\frac{-b_n \cdot \sin(\omega)}{a_n - b_n \cdot \cos(\omega)} \right).$$

Unter Beachtung von Abb. 7.3 bewirkt die durch (Gl. 7.1) gegebene Berechnungsvorschrift eine Tiefpassfilterung des digitalen Signals $s(k)$. Die Grenzfrequenz des Tiefpassfilters, aber auch die maximal erzielbare Dämpfung, ist eine Funktion der ganzzahligen Konstante n . Entsprechend dieser Funktionalität werden die "höherfrequenten" Signalanteile von $s(k)$ zunehmend gedämpft, während die niederfrequenten Anteile nahezu unverfälscht, aber je nach Frequenz mit einer deutlichen Phasenverschiebung am Ausgang $r_n(k)$ des Filters auftreten.

Ausgehend von dem Referenzsignal $r_n(k)$ kann durch die Addition und Subtraktion eines konstanten Wertes R_k jeweils ein oberer und ein unterer Schwellwert zum Vergleich mit dem ursprünglichen Sensorsignal $s(k)$ bestimmen. Überschreitet das Sensorsignal $s(k)$ die obere Schwelle $o(k) = r_n(k) + R_k$ oder unterschreitet $s(k)$ die untere Schwelle $u(k) = r_n(k) - R_k$, liegt ein Alarmfall vor. Anstelle des konstanten Offsets R_k zur Schwellwertberechnung ist auch eine Adaption von R_k an die zugrundeliegende Signaldynamik denkbar. Eine Möglichkeit zur Berechnung von adaptiven Schwellwerten wird im nächsten Abschnitt beschrieben.

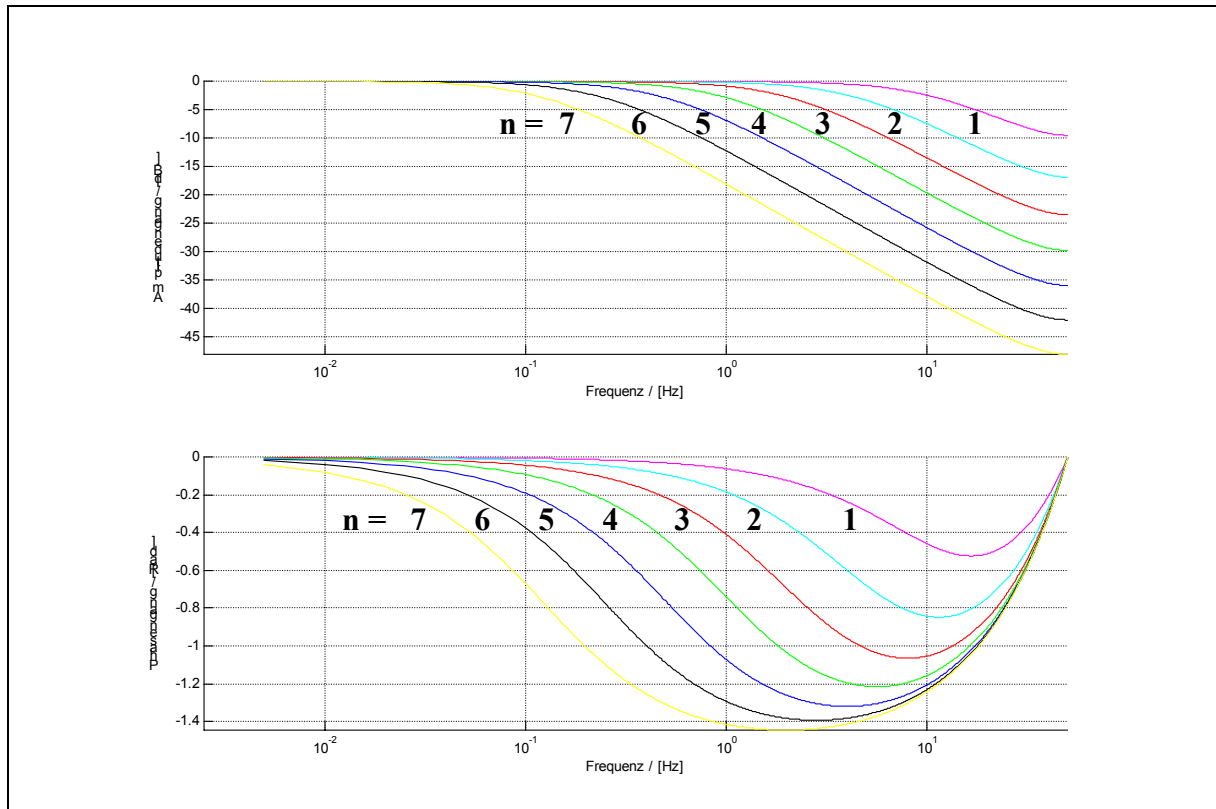


Abb. 7.3: Betragsamplituden- und Phasenspektrum der komplexen Übertragungsfunktion nach (Gl. 7.3)

Die mit der Konstante n beeinflussbare Dynamik des Referenzpegels $r_n(k)$ bestimmt entscheidend das Schaltverhalten des Wächters und kann je nach "Signalunruhe" auch online beeinflusst werden. Für große " n " ergibt sich ein relativ träges Adaptionsverhalten bei der Schwellwertberechnung, während bei kleinen " n " eine schnelle Schwellennachführung erfolgt. Beim Vergleich mit dem Sensorsignal $s(k)$ werden im letztgenannten Fall vorwiegend die höherfrequenten Anteile von $s(k)$ zur Signalbewertung verwendet.

Um die Störempfindlichkeit weiter zu erhöhen, kann zusätzlich zur Schwellwertprüfung eine minimale Über- bzw. Unterschreitungsdauer der Schwelle definiert werden, bevor tatsächlich ein Alarm ausgelöst wird. Bereits mit diesem einfachen Verfahren, das lediglich die "Signalunruhe" bewertet, lassen sich relativ gute Ergebnisse hinsichtlich der Störungsunterdrückung erzielen. Auf die Reichweite beim Draufzugehen wirkt sich diese Vorgehensweise allerdings

nachteilig aus, da typische niederfrequente Anteile insbesondere durch eine schnelle Bezugspegeladaption bzw. Schwellwertnachführung unterdrückt werden.

7.3 Ermittlung typischer Signaleigenschaften

Um ein zufriedenstellendes Schaltverhalten bei dem vorher beschriebenen Verfahren zu gewährleisten, ist ein Kompromiss zwischen einer schnellen Referenzpegel- bzw. Schwellwertadaption zur Störungsunterdrückung und einer langsamen Nachführung speziell mit Hinblick auf das Draufzugehverhalten erforderlich. Um die dabei verbleibende Fehlalarmanzahl weiter zu reduzieren, müssen zusätzliche, signalspezifische Eigenschaften zur Diskriminierung zwischen Stör- und Nutzsignalen gefunden werden.

Aufgrund der Vielzahl möglicher Störquellen existiert kein völlig eindeutiges Kriterium zur Unterscheidung, so dass eine vollständige Unterdrückung von Fehlalarmen auch mit diesen zusätzlichen Entscheidungskriterien nicht möglich sein wird. Durch die kombinierte Auswertung der im weiteren Verlauf verwendeten Signalcharakteristika kann die Detektionssicherheit allerdings soweit gesteigert werden, dass selbst bei ungünstigen Umgebungsbedingungen nur vereinzelte Fehlschaltungen auftreten.

Neben der bereits beschriebenen, adaptiven BezugspegelEinstellung in Verbindung mit von der Signalunruhe abhängigen, zum Referenzpegel symmetrischen Schwellwerten wird in diesem Abschnitt zusätzlich die Amplitudendifferenz und die Flankendauer zwischen den im Signalverlauf aufeinander folgenden Minimal- und Maximalwerten ausgewertet. Außerdem wird auch die zeitliche Differenz zwischen zwei gleichartigen Extremwerten (im folgenden als Pulsdauer bezeichnet) und der maximale Gradient auf der aktuellen Flanke bestimmt.

Ausgangspunkt für die Ermittlung der charakteristischen Signaleigenschaften ist die Bezugspegelnachführung nach (Gl. 7.1). Mit dem Referenzwert $r_n(k) = \text{REF}(k)$, der Konstante $n = 2$ und dem Sensorsignal $s(k) = \text{PIR}(k)$ erhält man den Bezugspegel

$$(Gl. 7.6) \quad \text{REF}(k) = \text{REF}(k-1) + \frac{\text{PIR}(k) - \text{REF}(k-1)}{2^2} .$$

Die Referenzpegelnachführung und die später beschriebene Form der Schwellwertberechnung ist in Abb. 7.4 beispielhaft an einem Mustersignal gezeigt. Das Signal $\text{REF}(k)$ dient dabei einerseits als Bezugspegel für die Bestimmung der symmetrischen Schwellen $\text{BOT}(k)$ und $\text{TOP}(k)$ (s. Abb. 7.4) und andererseits zur Erkennung der Polarität und von Polaritätswechseln des Originalsignals $\text{PIR}(k)$ relativ zum Referenzwert $\text{REF}(k)$ (s. Abb. 7.5). Spätestens bei einem Polaritätswechsel des Sensorsignals $\text{PIR}(k)$ wird außerdem ein Flankenwechsel detektiert. Die gesamte Auswertung ist mit der Erkennung eines Flankenwechsels synchronisiert, wobei das Verlassen des durch die Schwellwerte vorgegebenen Toleranzbandes bis zu diesem Zeitpunkt zwischengespeichert wird.

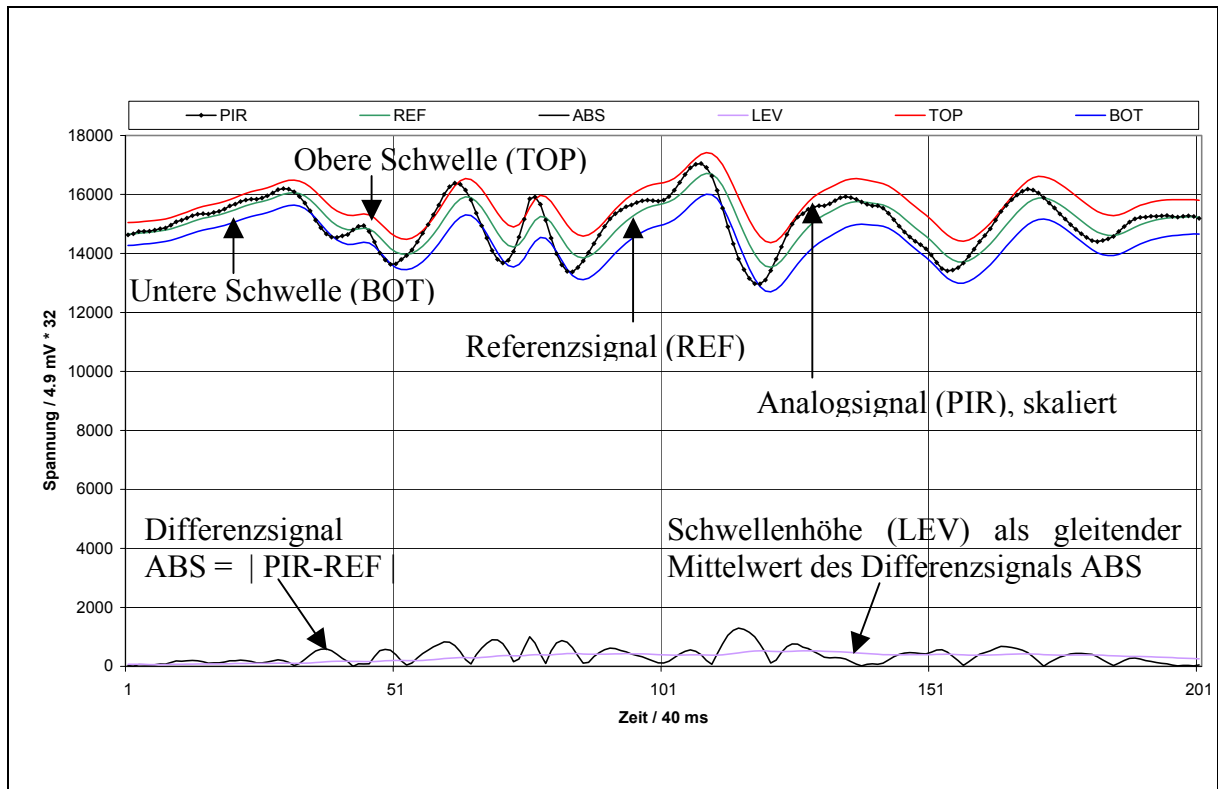


Abb. 7.4: Analogsignal PIR(k), Referenzpegel REF(k) und Berechnung der zum Bezugspegel symmetrischen Schwellen BOT(k) und TOP(k)

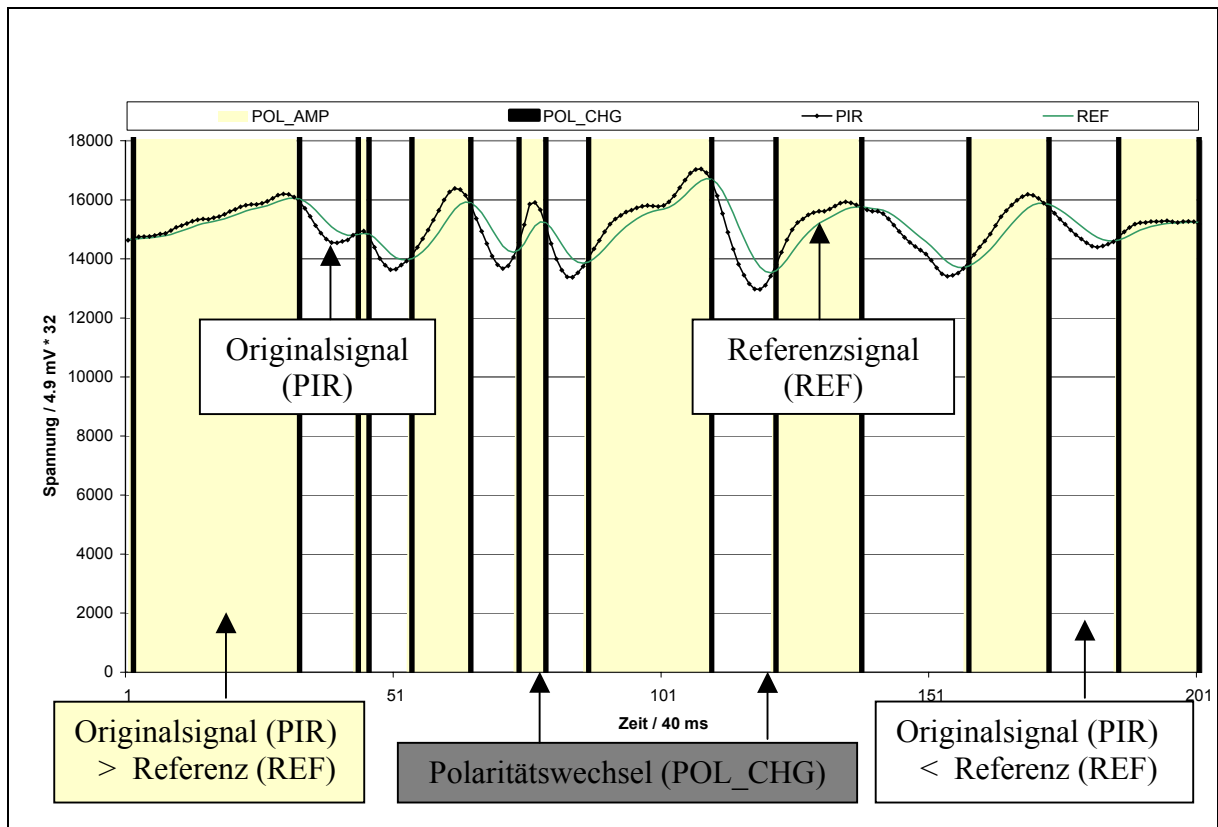


Abb. 7.5: Polarität POL_AMP(k) des Originalsignals PIR(k) relativ zum Referenzpegel REF(k) und Identifikation von Polaritätswechseln POL_CHG(k)

Zur Bestimmung der Schwellwerte BOT(k) und TOP(k) wird zunächst die absolute Differenz $ABS(k) = |PIR(k) - REF(k)|$ gebildet, woraus dann der ebenfalls zeitadaptive Offset

$$(Gl. 7.7) \quad LEV(k) = LEV(k-1) + \frac{ABS(k) - LEV(k-1)}{2^5}$$

für die Schwellen BOT(k) und TOP(k) berechnet wird. Unter Beachtung eines minimal erforderlichen Offsets THR_MIN und des maximal erlaubten Wertes THR_MAX werden die symmetrischen Amplitudenschwellen bestimmt:

$$(Gl. 7.8) \quad BOT(k) = REF(k) - \text{Minimum}[(THR_MIN + LEV(k)), (THR_MAX)]$$

$$(Gl. 7.9) \quad TOP(k) = REF(k) + \text{Minimum}[(THR_MIN + LEV(k)), (THR_MAX)]$$

Liegt der Signalwert PIR(k) außerhalb des durch die Schwellen BOT(k) und TOP(k) vorgegebenen Toleranzbandes, ist eine Alarmauslösung prinzipiell möglich. Allerdings muss der unmittelbar vorhergehende Signalverlauf die weiter unten beschriebenen Eigenschaften zusätzlich aufweisen, bevor der Alarmausgang tatsächlich aktiviert wird.

Abb. 7.6 zeigt die Maximal- bzw. Minimalwertverfolgung des Signalverlaufs, die wie alle Berechnungen online durchgeführt wird.

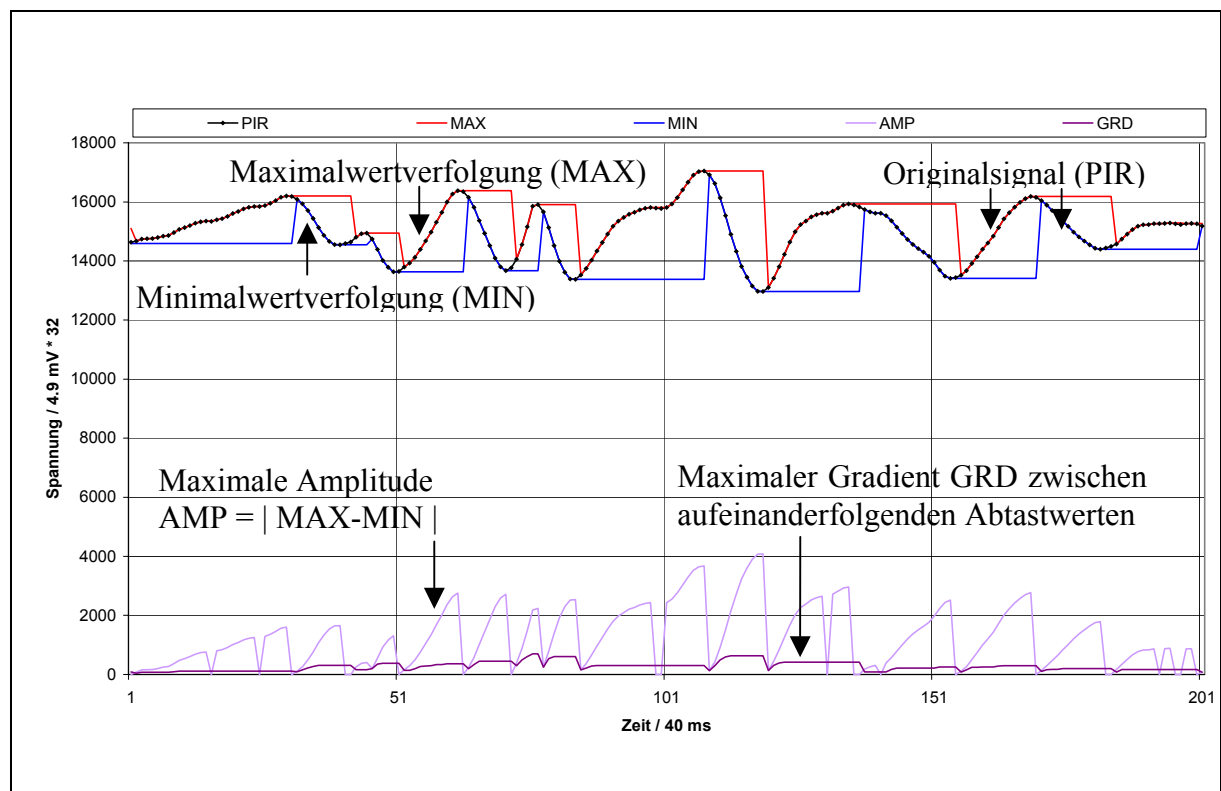


Abb. 7.6: Originalsignal PIR(k), Verfolgung der Maximalwerte MAX(k) bzw. der Minimalwerte MIN(k) im Signalverlauf sowie der maximalen Amplitude AMP(k) und des maximalen Gradienten GRD(k) der aktuellen Flanke

Auf einer steigenden Flanke wird das aktuelle Maximum MAX(k) fortlaufend aktualisiert, während das zuletzt aufgetretene Minimum MIN(k) unverändert erhalten bleibt. Bei einer

fallenden Flanke ist das Maximum $MAX(k)$ konstant, während in $MIN(k)$ die Minimalwertverfolgung durchgeführt wird. Für jeden Zeitschritt werden die absoluten Differenzen $|MAX(k) - MIN(k)|$ und $|PIR(k) - PIR(k-1)|$ gebildet und die maximalen Differenzen während einer Flanke in den Variablen $AMP(k)$ und $GRD(k)$ gespeichert.

Zur Unterscheidung zwischen steigenden und fallenden Flanken muss eine minimale Amplitudendifferenz zwischen dem letzten Extremwert (Minimum oder Maximum) und dem aktuellen Signalwert $PIR(k)$ vorliegen. Kurzfristig unveränderte Pegel oder kleinere "Peaks" während einer Flanke beeinflussen den aktuellen Zustand bei der Signalverfolgung daher nicht. Als weiteres Kriterium für einen Flankenwechsel, der letztendlich eine neue Bewertung des Signalverlaufs startet, ist ein Polaritätswechsel des Originalsignals $PIR(k)$ bezüglich des zeitlich nacheilenden Referenzpegels $REF(k)$ definiert. Insbesondere bei geringen Signalamplituden (elektronisches Rauschen oder Hintergrundrauschen) bewirkt diese Festlegung eine ständige Umschaltung der Flankenbewertung.

Sowohl die maximale Amplitude $AMP(k)$ als auch der maximale Gradient $GRD(k)$ während einer Flanke müssen innerhalb vorgegebener Toleranzgrenzen AMP_MIN und AMP_MAX bzw. GRD_MIN und GRD_MAX liegen, damit die Alarmauslösung erfolgt. Die Einhaltung der Toleranzgrenzen wird für jeden Abtastzeitpunkt neu bewertet. Das einmalige Überschreiten eines Maximalwertes verhindert die Alarmauslösung für die Dauer der gesamten Flanke.

Eine ähnliche Vorgehensweise zur Auswertung eines Bewegungsmeldersignals wird in [Pat. 7.2] beschrieben. Zur Verbesserung der Störfestigkeit gegenüber Bursteinkopplungen oder elektrostatischen Kontaktentladungen mit hoher Anstiegssteilheit wird die Steilheit des Sensorsignals mit Hinblick auf einen zulässigen oberen Grenzwert überwacht. Störungen durch langsame Temperaturänderungen im Erfassungsfeld werden dagegen mit einem unteren Grenzwert für die Steilheit ausgeblendet. Beide Grenzwerte definieren ein Toleranzband für die Steilheit typischer Nutzsingale, wobei außerdem das Auftreten eines minimalen Spannungshubs (Mindestamplitude) während einer Flanke als zweites Entscheidungskriterium angewendet wird [Pat. 7.2].

Ein wichtiger Vorteil der Differenzamplitudenauswertung und der Gradienten- bzw. Steilheitsbewertung ist die Unabhängigkeit vom absoluten Spannungspegel des Signals. Auch Signalüberlagerungen (langsame Störung + schnelleres Nutzsignal) können ohne großen Aufwand detektiert werden, sofern sie charakteristische Impulse hervorrufen.

Parallel zur Differenzamplituden- und Gradientenverfolgung wird im folgenden auch die Messung der Flankendauer durchgeführt. Bei einer steigenden Flanke erfolgt die Messung der Flankendauer mit dem Zähler $POS(k)$, während einer fallenden Flanke wird der Zähler $NEG(k)$ verwendet. Der jeweils nicht verwendete Zähler dient zur Zwischenspeicherung des Zählstandes für den Fall, dass das aktuelle Signal keine eindeutige Zuordnung des weiteren

Flankenverlaufs zulässt. Die entsprechenden Zählstände POS(k) oder NEG(k) werden nachträglich aktualisiert, sobald eine eindeutige Zuordnung möglich ist (s. Abb. 7.7).

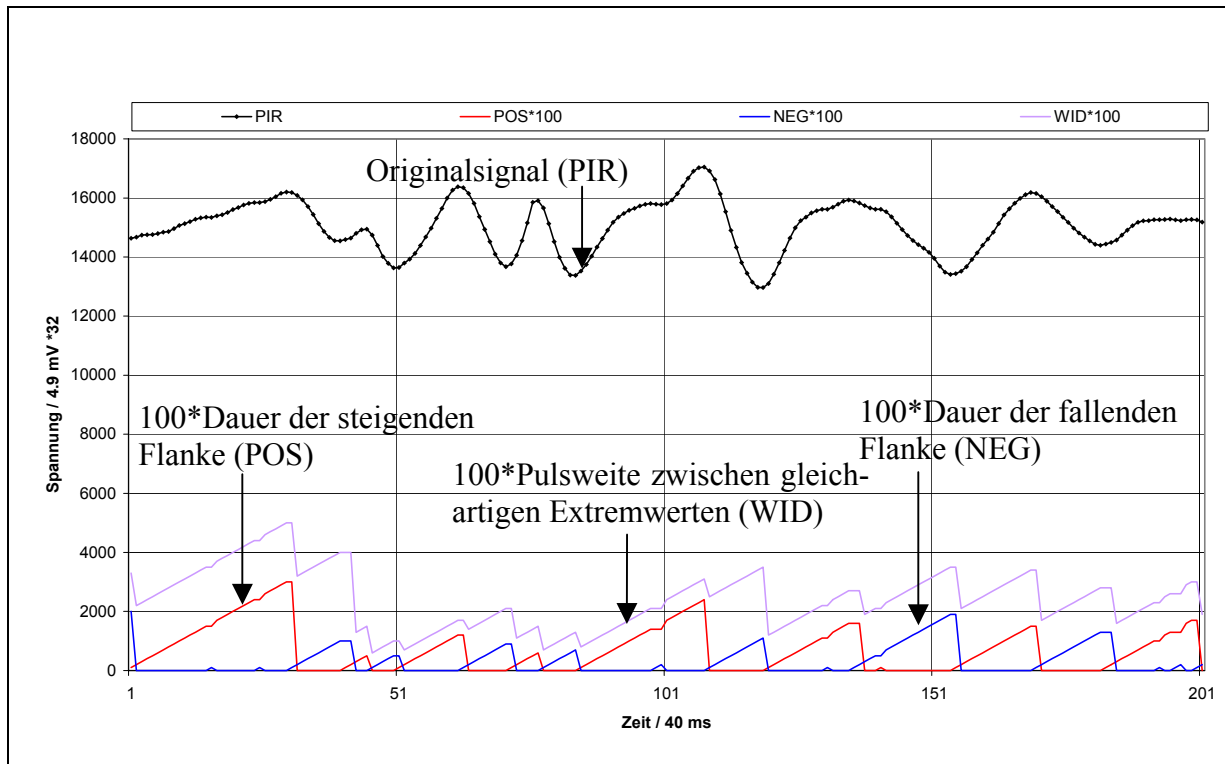


Abb. 7.7: Originalsignal PIR(k) und Verfolgung der Flankendauer POS(k) oder NEG(k) der steigenden oder fallenden Flanke sowie Messung der Pulsdauer WID(k) zwischen benachbarten, gleichartigen Extremwerten (Maximum - Maximum oder Minimum - Minimum). Die Skalierung der Zählerstände POS(k), NEG(k) und WID(k) mit dem Faktor 100 erfolgte lediglich zur Verbesserung der Darstellbarkeit.

Zusätzlich zur Messung der Flankendauer wird auch die Pulsdauer WID(k) zwischen zwei aufeinander folgenden, gleichartigen Extremwerten im Signalverlauf ermittelt. Während einer steigenden Flanke erfolgt die Pulsdauerermessung zwischen benachbarten Maxima, bei einer fallenden Flanke wird die Anzahl der Abtastzeitpunkte zwischen aufeinander folgenden Minima bestimmt. Zur Initialisierung des Zählers WID(k) dient dabei der Zählstand des zuletzt verwendeten Flanken Zählers POS(k) oder NEG(k). Analog zur Zwischenspeicherung des Zählstandes bei nicht eindeutigem Signalverlauf wird auch der Pulsdauerzähler WID(k) nachträglich aktualisiert, wenn der weitere Flankenverlauf identifiziert ist (s. Abb. 7.7).

Die Zählstände POS(k) bzw. NEG(k) und WID(k) müssen ebenfalls innerhalb vorgegebener Toleranzgrenzen SLP_MIN und SLP_MAX bzw. WID_MIN und WID_MAX liegen, um eine Alarmauslösung vorzubereiten. Auch hier wird die Einhaltung der Toleranzgrenzen für jeden Abtastzeitpunkt neu bewertet, wobei das einmalige Überschreiten eines Maximalwertes die Alarmauslösung für die Dauer der gesamten Flanke sperrt.

Abb. 7.8 zeigt den Zustand des globalen Alarmflags ALARM, der sich aus einer UND-Verknüpfung aller zuvor beschriebenen Alarmzustände ergibt. Ist nur eine Bedingung nicht erfüllt, bleibt der Alarmausgang des Bewegungsmelders deaktiviert.

Zur Ermittlung der Schaltzustände nach Abb. 7.8 wurden die Toleranzgrenzen für die Auswertung der zusätzlichen Signalparameter zunächst mit den minimal bzw. maximal möglichen Werten initialisiert. In diesem Fall wird lediglich die Referenzpegeladaption und die Schwellwertnachführung gemäß dem vorne beschriebenen Verfahren zur Diskriminierung zwischen Nutz- und Störsignalen verwendet. Im nächsten Abschnitt werden dann Möglichkeiten zur Verbesserung des Störverhaltens durch die zusätzliche Bewertung der vorher beschriebenen Signalcharakteristika aufgezeigt.

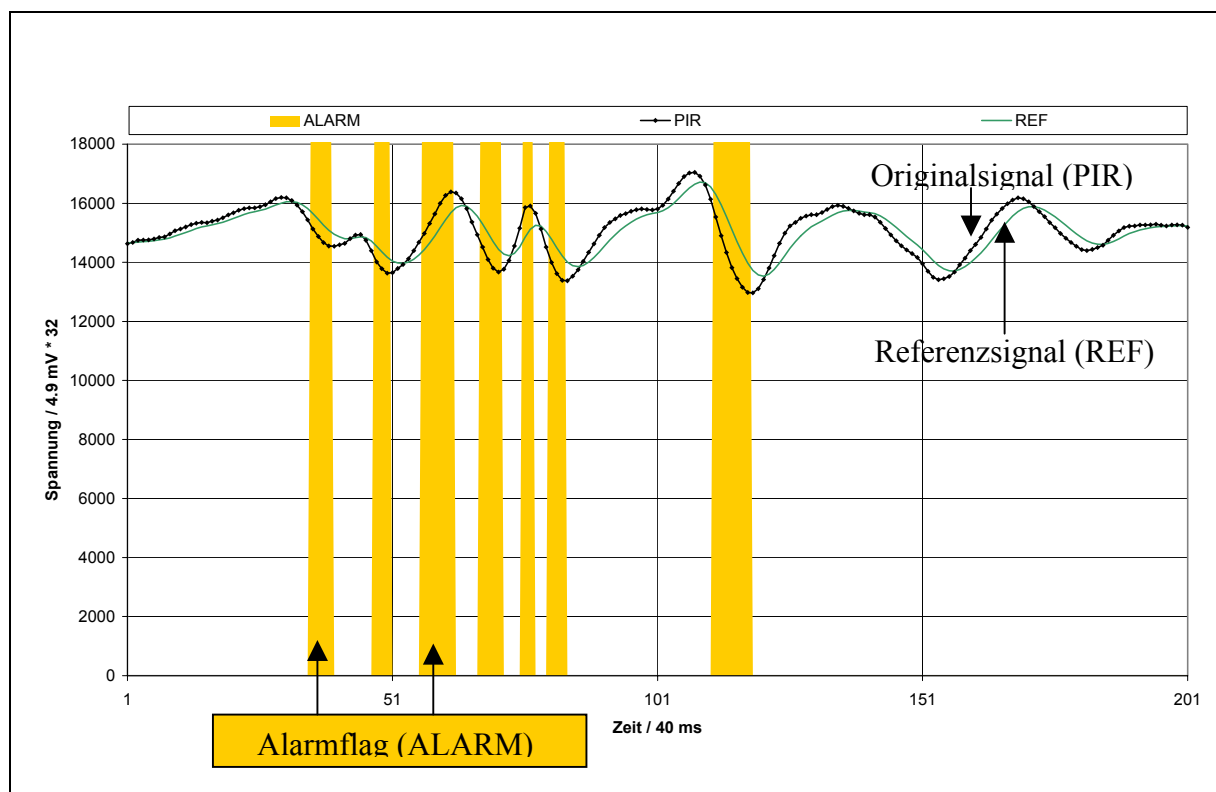


Abb. 7.8: Originalsignal PIR(k) und Zustand des globalen Alarmflags ALARM(k) zum Vergleich mit den in Abb. 7.4 dargestellten Über- bzw. Unterschreitungen der Schwellwerte TOP(k) und BOT(k)

Unter den vorgenannten Bedingungen wird das globale Alarmflag gesetzt, wenn das Signal PIR(k) das durch die Schwellen TOP(k) und BOT(k) vorgegebene Toleranzband verlässt (vgl. Abb. 7.4). Es wird wieder gelöscht, sobald ein Flankenwechsel detektiert wurde, der einen neuen Auswertezyklus startet. Das globale Alarmflag beeinflusst unmittelbar den Zustand des Schaltausgangs, der bei gesetztem Flag für eine vorgebbare Zeitspanne aktiviert wird. Treten weitere Alarmzustände während dieser Dauer auf, wird der Schaltausgang entsprechend nachgetriggert. Der Übersichtlichkeit halber wurde auf eine Darstellung des Zustands des Schaltausgangs in Abb. 7.8 verzichtet.

7.4 Bewertung charakteristischer Signaleigenschaften

In Verbindung mit einer Auswahl charakteristischer Nutz- und Störsignale wird im diesem Abschnitt die Optimierung des Schaltverhaltens des Bewegungsmelders rechnergestützt durchgeführt. Die Nutzsignale wurden im Indoorbereich aufgezeichnet, wobei die im folgenden verwendeten Richtungsdefinitionen anhand von Abb. 7.9 verdeutlicht werden sollen.

Abb. 7.9 zeigt ein radiales 180°-Erfassungsfeld mit einer maximalen Reichweite von 16 m. Die Richtungsangaben "parallel" und "vertikal" beziehen sich jeweils auf die innere Erfassungsfeldgrenze (Wand). In beiden Bewegungsrichtungen wurden Nutzsignale mit einer Abstufung von 0,5 m im gesamten Gesichtsfeld aufgenommen.

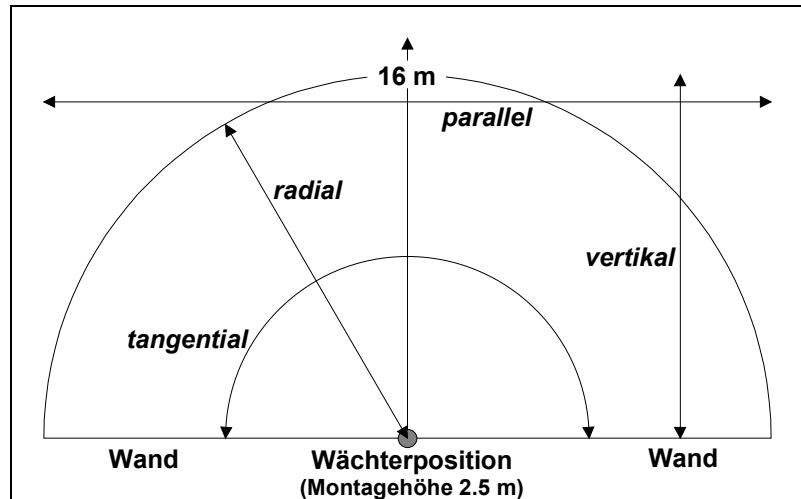


Abb. 7.9: Definition der Bewegungsrichtungen bei der Aufzeichnung der Nutzsignale

Die "radiale" und die "tangentiale" Bewegungsrichtung beziehen sich auf die Wächterposition. Der Bewegungsmelder war in einer Montagehöhe von 2.5 m montiert. Signale aus radialen Gehrichtungen (zum Wächter hin und wieder zurück) wurden mit einem Winkelschritt von 5° zwischen benachbarten Bewegungsrichtungen erzeugt, während der Radius bei tangentialer Bewegung wiederum in Schritten von 0,5 m variierte.

Abb. 7.10 zeigt die mit dem vorher beschriebenen Algorithmus ermittelten, charakteristischen Signaleigenschaften "Gradient" (GRD), "Flankendauer" (SLP) und "Pulsdauer" (WID) in Abhängigkeit der gemessenen "Differenzamplitude" (AMP). Nachträglich ergänzt wurde die ebenfalls dargestellte mittlere "Steigung" (STG), die am Impulsende als Quotient der Amplitude AMP und der Flankendauer SLP berechnet wird. Die Ermittlung der Steigung erfordert allerdings eine höhere Rechenleistung als die Bestimmung der übrigen Parameter.

Das im Indoorbereich vorhandene Rauschen wurde in einer weiteren Versuchsreihe ermittelt. Zum Vergleich mit den Nutzsignalen nach Abb. 7.10 wurden die resultierenden Signaleigenschaften in Abb. 7.11 angegeben. Erwartungsgemäß treten relativ kleine Rauschamplituden im Indoorbereich auf. Für die zugrundeliegenden Umgebungsbedingungen bildet das Indoorrauschen zunächst die untere Detektionsgrenze.

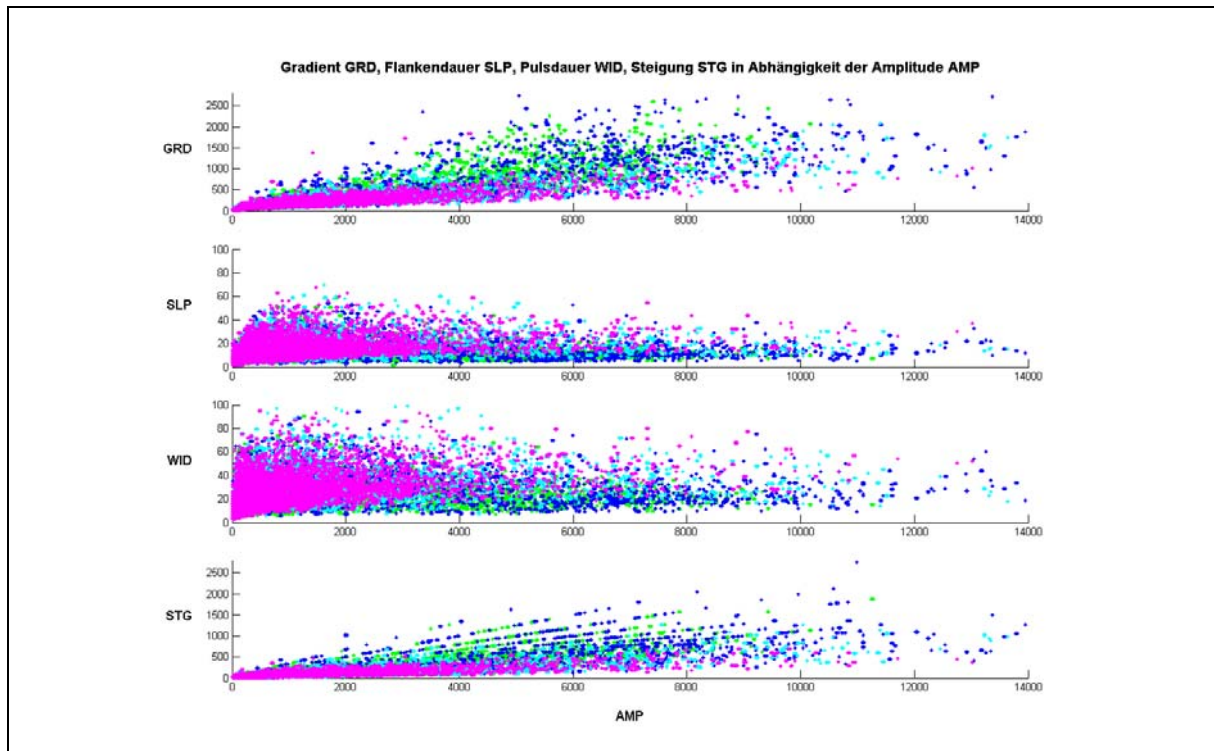


Abb. 7.10: Gradient, Flankendauer, Pulsdauer und mittlere Steigung der Nutzsignale aus den Indoormessungen in Abhängigkeit der Amplitude. Das Signalrauschen aus dem Indoorbereich ist in diesen Daten ebenfalls enthalten. (Farbcodierung: radial (magenta), vertikal (cyan), parallel (blau), tangential (grün))

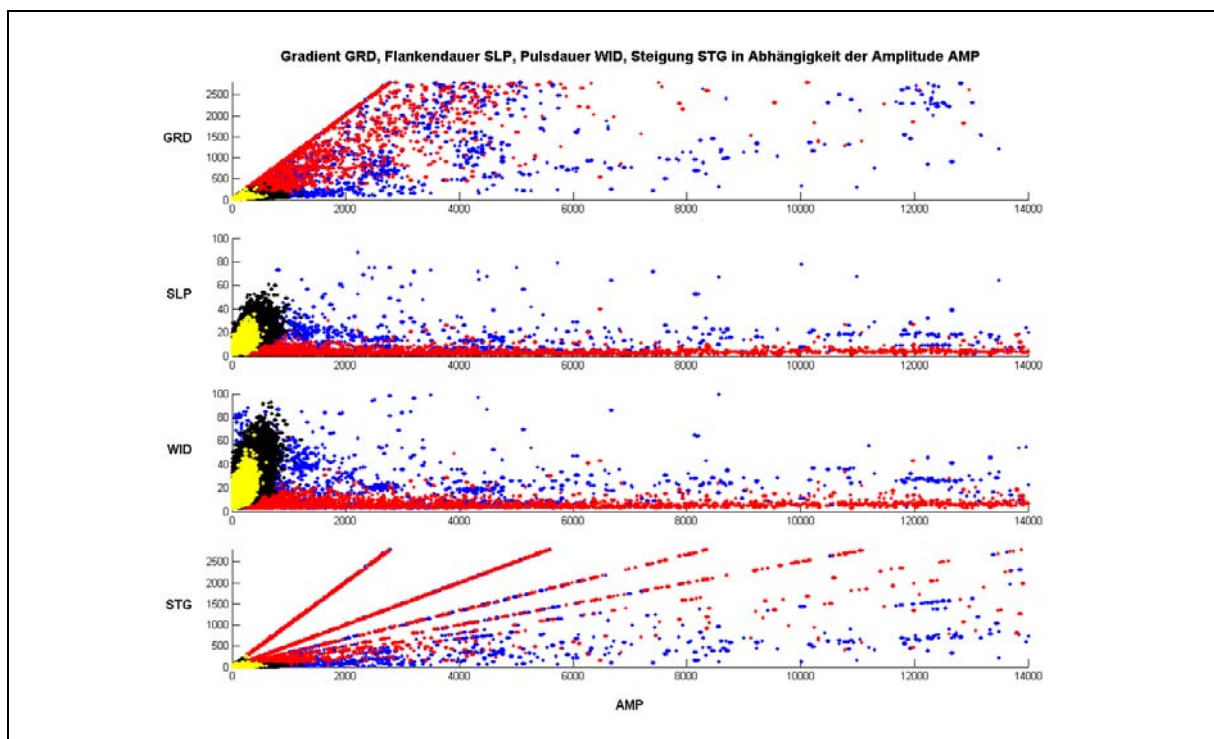


Abb. 7.11: Gradient, Flankendauer, Pulsdauer und mittlere Steigung der Störsignale von verschiedenen Störquellen in Abhängigkeit der Amplitude. (Farbcodierung: indoor (gelb), outdoor (schwarz), Bursteinkopplung und elektrostatische Kontaktentladung (rot), Handy im Nahbereich (blau))

Neben dem Rauschen aus dem Indoorbereich wurden in Abb. 7.11 weitere typische Störquellen berücksichtigt. Man erkennt, dass der Gradient und insbesondere die mittlere Steigung sehr gut zur Identifikation von steilflankigen Störungen durch elektromagnetische Einkopplungen (Burst) und elektrostatische Kontaktentladungen (ESD) geeignet sind. Werden derartige Störungen identifiziert, empfiehlt es sich, die Signalauswertung für einen vorgebbaren Zeitraum prinzipiell unempfindlicher auszulegen. Gleiches gilt für Signale, die auf den Handybetrieb in unmittelbarer Nähe des Bewegungsmelders zurückzuführen sind.

Fehlalarme durch Störungen im Outdoorbereich sind dagegen schwieriger von den tatsächlichen Auslösern zu unterscheiden. In Abb. 7.11 erkennt man, dass diese Störungen größere Amplituden als die Rauschsignale aus dem Indoorbereich aufweisen. Auch die Flankendauern und die Pulsdauern der Outdoor-Störungen weisen zum Teil deutlich größere Werte auf als die entsprechenden Indoor-Rauschsignale. An dieser Stelle soll darauf hingewiesen werden, dass die verwendeten Outdoor-Störungen aus einem recht anspruchsvollen Testfeld stammen und unter relativ ungünstigen Umgebungsbedingungen aufgenommen wurden. Wenn es daher gelingt, diese Störungen zu unterdrücken, sind bei "normalen" Umgebungsbedingungen kaum Fehlalarme zu erwarten, was durch entsprechende Vergleichsmessungen in anderen Testfeldern bestätigt wurde. Zur besseren Vergleichbarkeit wurden in Abb. 7.12 bis Abb. 7.14 alle Nutzsinalcharakteristiken aus Abb. 7.10 zusammen mit den Stördaten aus dem Indoor- und Outdoorbereich nach Abb. 7.11 in unterschiedlichen Kombinationen dargestellt.

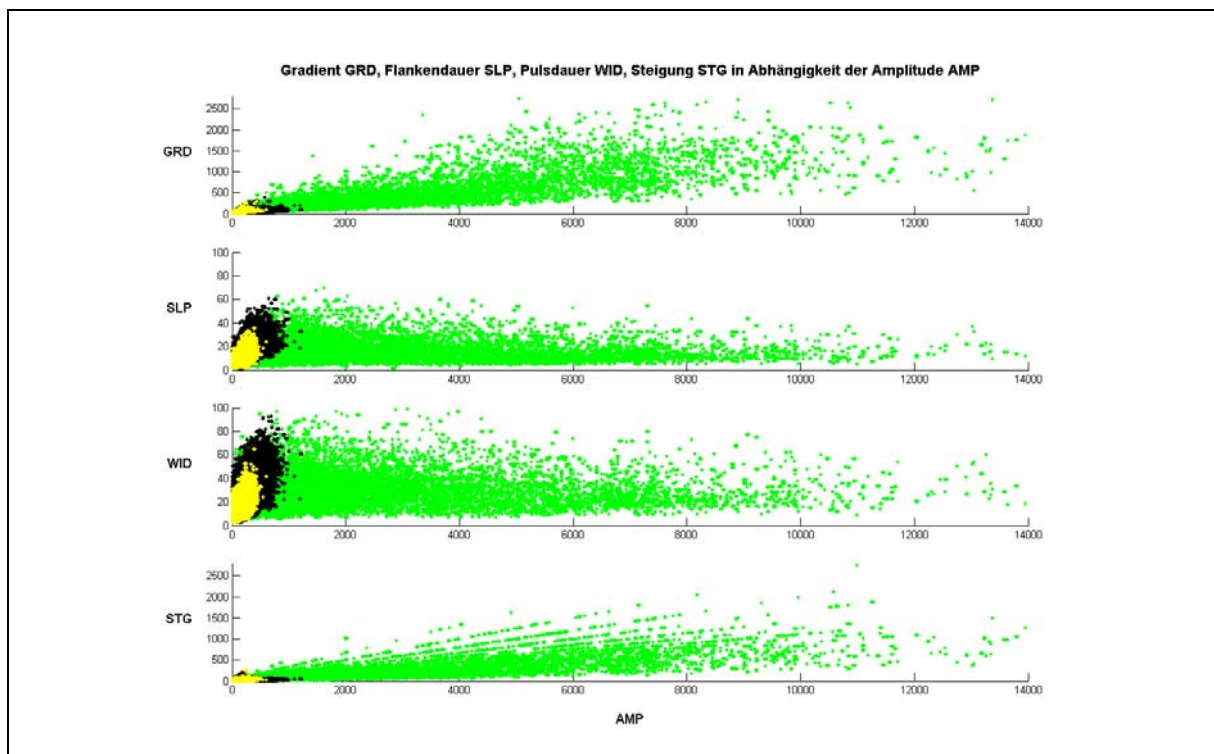


Abb. 7.12: Gradient, Flankendauer, Pulsdauer und mittlere Steigung der Signale aus den Indoor- und Outdoormessungen in Abhängigkeit der Amplitude. (Farbcodierung: Rauschen indoor (gelb), Rauschen outdoor (schwarz), Nutzsinal indoor (grün))

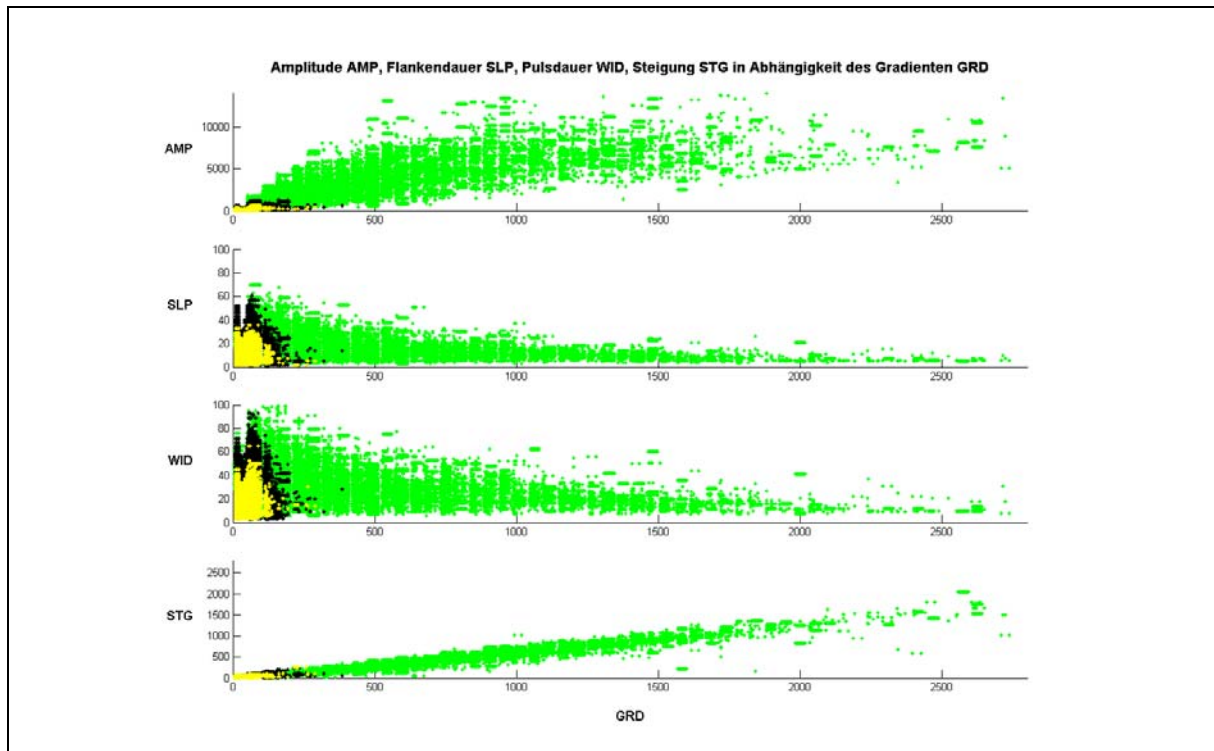


Abb. 7.13: Amplitude, Flankendauer, Pulsdauer und mittlere Steigung der Signale aus den Indoor- und Outdoormessungen in Abhängigkeit der Gradienten. (Farbcodierung: Rauschen indoor (gelb), Rauschen outdoor (schwarz), Nutzsignale indoor (grün))

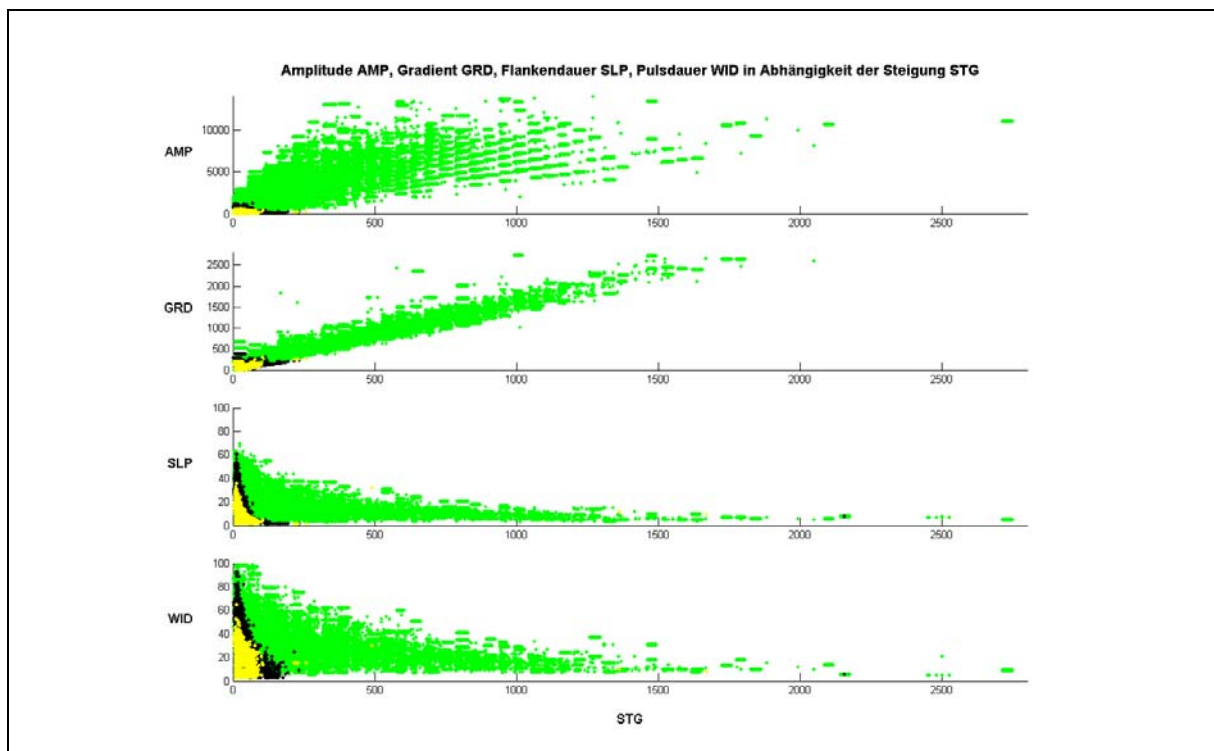


Abb. 7.14: Amplitude, Gradient, Flankendauer und Pulsdauer der Signale aus den Indoor- und Outdoormessungen in Abhängigkeit der mittleren Steigung. (Farbcodierung: Rauschen indoor (gelb), Rauschen outdoor (schwarz), Nutzsignale indoor (grün))

Anhand der Daten aus Abb. 7.12 bis Abb. 7.14 lässt sich eine erste Optimierungsstrategie ableiten, die im weiteren Verlauf vorgestellt wird. Darauf aufbauend werden die vorne beschriebenen Grenzwerte rechnergestützt durch eine Variation der Toleranzgrenzen ermittelt und die Auswirkungen auf das Schaltverhalten bei Störungen und Nutzsignalen präsentiert. An dieser Stelle soll darauf aufmerksam gemacht werden, dass das Schaltverhalten des optimierten Wächters von der Repräsentativität der zur Parametereinstellung verwendeten Signale abhängt.

Der im folgenden beschriebene Optimierungsprozess anhand der zugrundegelegten Messdaten sollte daher als Beispiel für die Anwendung und die Flexibilität des in Kap. 6 erläuterten Entwicklungstools verstanden werden. Für eine endgültige Optimierung sind weitere, unter extremen Umgebungsbedingungen und Temperaturkontrasten aufgenommene Testdaten erforderlich, die zum Zeitpunkt der Erstellung dieser Arbeit nur in begrenztem Umfang vorlagen. An der prinzipiellen Vorgehensweise ändert sich dadurch aber nichts.

Ein erster Schritt zur Parameteroptimierung ist die Einstellung geeigneter Adaptionsgeschwindigkeiten für die Referenzpegel- und die Schwellwertadaption sowie der minimalen und maximalen Grenzen bei der Schwellwertnachführung. Bei den entsprechenden Untersuchungen hat sich gezeigt, dass sich eine schnelle Referenzpegeladaption und eine trägere Schwellwertnachführung vorteilhaft auf das Schaltverhalten auswirken.

Die im folgenden verwendeten Adaptionsgeschwindigkeiten wurden bereits in (Gl. 7.6) und (Gl. 7.7) angegeben. Der minimale Schwellwert THR_MIN wurde anschließend so weit verringert, dass das Grundrauschen keinen Schaltvorgang auslöst. Es ergab sich die untere Grenze $THR_MIN = 188$. Als maximaler Schwellwert wurde $THR_MAX = 220$ gewählt, wodurch sich ein sehr gutes Schaltverhalten des Wächters in Verbindung mit Nutzsignalen aus dem Indoorbereich realisieren lässt. Abb. 7.15 zeigt die Auswirkungen dieser Einstellung auf die Stör- und Nutzsignale, wobei nur noch diejenigen Signalcharakteristika dargestellt sind, die weiterhin eine Alarmauslösung bewirkt hätten. Abb. 7.16 zeigt die zugehörigen Schaltpunkte des Wächters bei reiner Schwellenauswertung der Nutzsignale, es lassen sich allerdings nicht alle Fehlschaltungen unterdrücken (s. Tab. 7.1). Die verbleibenden Störsignale müssen durch die Auswertung der Gradienten bzw. Steigungen sowie durch entsprechende Toleranzgrenzen für die Flankendauer, die Pulsdauer und die Amplitude unterdrückt werden.

Man erkennt in Abb. 7.12 bis Abb. 7.14, dass die Vorgabe von Mindestwerten für die Flanken- und die Pulsdauer sich besonders auf die Signalverläufe auswirken wird, die ähnliche zeitliche Charakteristika wie das Grundrauschen aufweisen. Die maximalen Toleranzgrenzen wirken sich dagegen auf sehr langsame Signaländerungen aus, die typischerweise infolge von großflächigen Temperaturdriften des Hintergrunds entstehen. Die obere und die untere Grenze zur Flanken- und Pulsdauerbewertung sollte anhand typischer Nutzsignalcharakteristiken ermittelt werden, um die Klasse der detektierbaren Signale von vorne herein nicht zu sehr ein-

zuschränken. Die im Rahmen der Parametervariation ermittelten Grenzwerte für die Flanken- und die Pulsdauer sind in Tab. 7.1 angegeben. Die Pulsweite valider Schwingungen ist ungefähr doppelt so groß wie die korrespondierende Flankendauer. Insbesondere bei niedrigen Frequenzen werden durch diese Festlegung Anforderungen hinsichtlich der Symmetrie von steigender und fallender Flanke berücksichtigt.

Die zeitlichen Eigenschaften der Nutzsignale sind stark abhängig von der Bewegungsrichtung, der Gehgeschwindigkeit und der Entfernung beim Durchschreiten der einzelnen Erfassungsbereiche im Gesichtsfeld des Bewegungsmelders. Störsignale, deren Puls- und Flankendauer innerhalb des zuvor festgelegten Toleranzbandes liegen, müssen durch die amplitudenbezogenen Größen unterdrückt werden. Bei der Amplitude, dem Gradient bzw. der mittleren Steigung wirkt sich der Temperaturkontrast zwischen dem bewegten Objekt und dem Hintergrund stark aus.

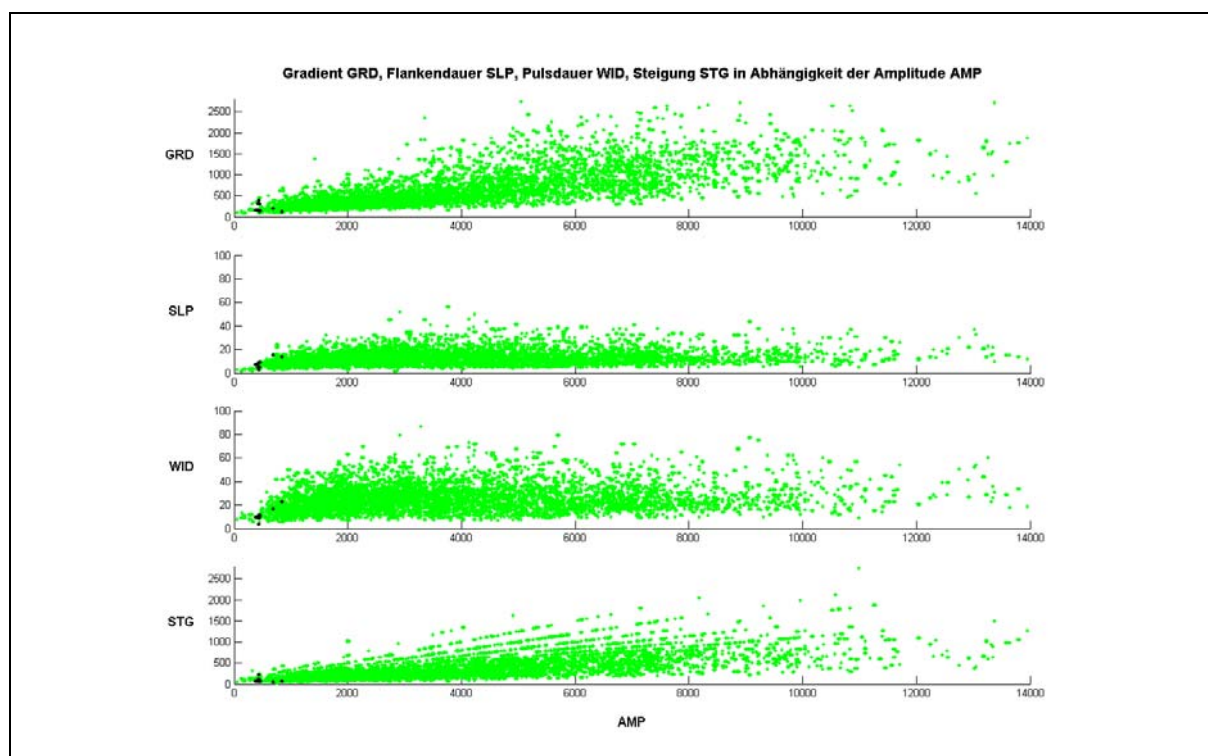


Abb. 7.15: Wirkung der Referenzpegel- und Schwellwertadaption in Verbindung mit dem minimalen und maximalen Grenzwert für die Schwellwerte TOP(k) und BOT(k) auf die Nutz- und Störsignaleigenschaften nach Abb. 7.12

Die für eine Alarmauslösung minimal erforderliche Amplitude AMP_MIN kann ebenfalls adaptiv an die zugrundeliegende Signalunruhe angepasst werden. Zur Steuerung bietet sich beispielsweise der dynamische Teil LEV(k) der Schwellwertnachführung aus (Gl. 7.7) an, der als Maß für die mittlere Signalunruhe betrachtet werden kann. Für die im folgenden gezeigten Ergebnisse der Signalauswertung wurde allerdings die konstante Mindestamplitude AMP_MIN = 600 zugrundegelegt. Außerdem wurde die obere Grenze AMP_MAX auf den Maximalwert 32767 gesetzt, so dass dieser Parameter die Auswertung nicht beeinflusst.

Anstelle der mittleren Steigung wurde außerdem der Gradient zur weiteren Fehlalarmunterdrückung verwendet. Im Rahmen der rechnergestützten Optimierungsphase ergab sich der minimale Grenzwert $GRD_MIN = 129$ zur Abgrenzung gegenüber Störsignalen aus dem Outdoorbereich sowie der Maximalwert $GRD_MAX = 2500$ zur Diskriminierung gegenüber steilflankigen, elektromagnetischen Störeinkopplungen.

Mit den gegebenen Toleranzgrenzen konnte die bei reiner Schwellenauswertung auftretende Fehlalarmzahl von 23 auf Null reduziert werden. Abb. 7.16 und Abb. 7.17 zeigen die Auswirkung der Optimierung auf die Reichweite in verschiedenen Bewegungsrichtungen. Die Einschränkungen reduzieren die Anzahl der Schaltpunkte deutlich. Sie wirken sich besonders auf die Draufzugehreichweite aus (vgl. Abb. 7.16 und Abb. 7.17 im Bereich von 7.5 ...10 m). Aufgrund der verbleibenden Schaltpunktanzahl und deren recht homogener Verteilung im Erfassungsfeld kann die vorgenommene Einstellung dennoch als geeigneter Kompromiss mit Hinblick auf das Schaltverhaltens betrachtet werden.

Algorithmusparameter	Nur Schwellwerte	Alle Parameter
THR_MIN	188	188
THR_MAX	220	220
AMP_MIN	1	600
GRD_MIN	1	129
GRD_MAX	32767	2500
SLP_MIN	1	13
SLP_MAX	32767	47
WID_MIN	1	23
WID_MAX	32767	80
Alarmanzahl	517	429
Fehlalarmzahl	23	0

Tab. 7.1: Alarm- und Fehlalarmanzahl der reinen Schwellenauswertung im Vergleich zur Schwellenauswertung in Verbindung mit zusätzlichen Toleranzgrenzen

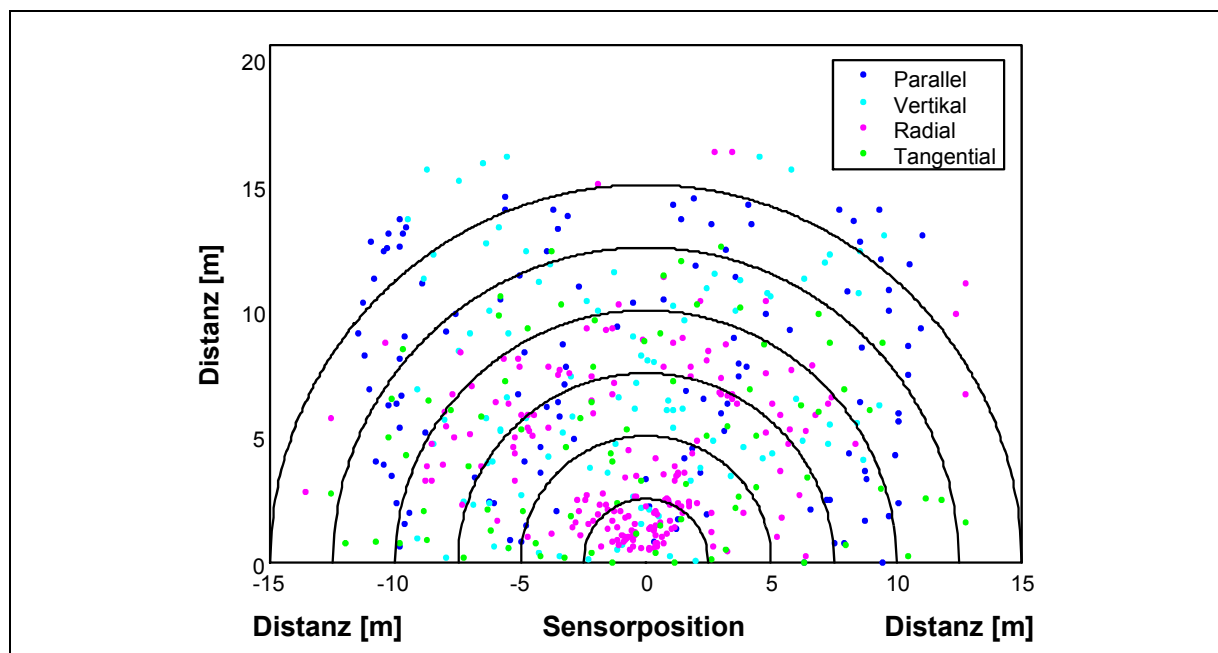


Abb. 7.16: Auswirkung der reinen Schwellenauswertung auf die Schaltreichweite des Bewegungsmelders

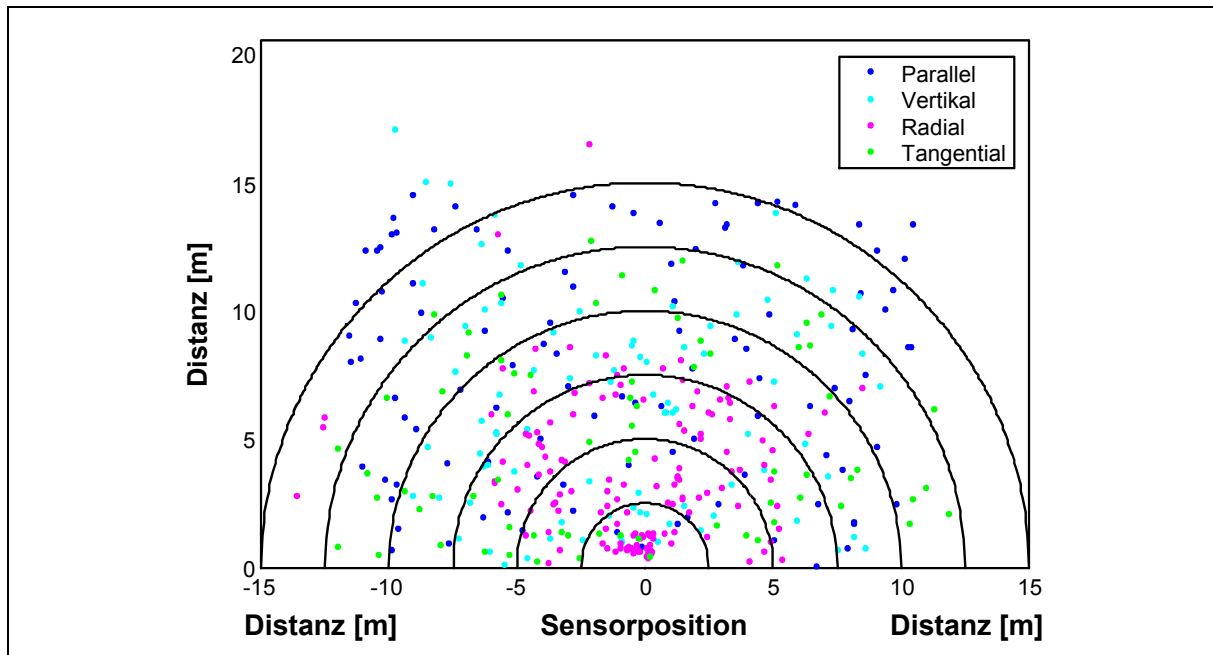


Abb. 7.17: Auswirkung der Schwellenauswertung in Verbindung mit der Bewertung zusätzlicher Signalcharakteristiken auf die Schaltreichweite

7.5 Korrelationstechniken zur Bewegungserkennung

Um dieses Kapitel über konventionelle Auswerteverfahren zur Bewegungserkennung abzuschließen, werden in den nächsten beiden Abschnitten zwei Verfahren beschrieben, welche die Autokorrelationsfunktion (AKF) und die Kreuzkorrelationsfunktion (KKF) zur Signalanalyse von Bewegungsmeldersignalen anwenden. Mit beiden Verfahren ergeben sich im Vergleich zu einfacheren Strategien, die auf einer reinen Schwellenbewertung beruhen, Verbesserungen hinsichtlich des Signal-Rausch-Verhältnisses, die sich zur Steigerung der Detektivität anwenden lassen.

Während bei der AKF Signalabschnitte desselben Signals miteinander verglichen werden und daher keine grundsätzlich neue Information über das zugrundeliegende Signal zur Verfügung steht, lassen sich mit der KKF zusätzliche Informationen über statistische Bindungen zwischen unterschiedlichen Signalen ableiten. Beide Verfahren verwenden implizit die Energie bzw. Leistung der korrelierten Signale, so dass es an dieser Stelle Berührungspunkte mit dem in Kap. 8-10 vorgestellten KALMAN-Filter gibt.

7.5.1 Autokorrelation zur Erkennung von deterministischen Signalanteilen

In [Pat. 7.3] wird ein Bewegungsmelder und ein Verfahren zur Bewegungsmeldung vorgestellt, dass die AKF des digitalisierten Sensorsignals $y(k)$ zur Unterdrückung von Störungen verwendet. Ausgangspunkt ist die Überlegung, dass sich das Ausgangssignal $y(k)$ des Bewegungsmelders als additive Überlagerung des Nutzsignals $s(k)$, des Störanteils $d(k)$ und des Rauschanteils $v(k)$ ergibt:

$$(Gl. 7.10) \quad y(k) = s(k) + d(k) + v(k)$$

Das Rauschsignal $v(k)$ wiederum repräsentiert die Überlagerung der thermischen und elektrischen Rauschkomponenten, wobei die effektive Rauschbandbreite durch eine nachgeschaltete Bandpass-Verstärker-Anordnung eingeschränkt wird. Bildet man die Gesamt-AKF $\varphi_{yy}(k)$ des Ausgangssignals $y(k)$, ergibt sich die Summe

(Gl. 7.11)

$$\varphi_{yy}(k) = \varphi_{ss}(k) + \varphi_{dd}(k) + \varphi_{vv}(k) + \varphi_{sd}(k) + \varphi_{ds}(k) + \varphi_{sv}(k) + \varphi_{vs}(k) + \varphi_{dv}(k) + \varphi_{vd}(k),$$

wobei die AKF anhand der Berechnungsvorschrift

$$(Gl. 7.12) \quad \varphi_{yy}(k) = \frac{\sum_{\mu=0}^{M-1} y(\mu) \cdot y(\mu+k)}{\sum_{\mu=0}^{M-1} y^2(\mu)}$$

gebildet wird. M entspricht der Anzahl der Abtastwerte im Korrelationsfenster, welches innerhalb eines gewissen Toleranzbandes an den zugrundeliegenden Signalverlauf angepasst wird. In [Pat. 7.3] wird vorausgesetzt, dass $s(k)$, $d(k)$ und $v(k)$ statistisch unabhängig sind. Berücksichtigt man zunächst nur die statistische Unabhängigkeit der Rauschkomponente $v(k)$, verschwinden die entsprechenden Kreuzkorrelationsterme in (Gl. 7.11), woraus folgt

$$(Gl. 7.13) \quad \varphi_{yy}(k) = \varphi_{ss}(k) + \varphi_{dd}(k) + \varphi_{vv}(k) + \varphi_{sd}(k) + \varphi_{ds}(k).$$

Im weiteren Verlauf wird durch betragsmäßige Differenzbildung der Gesamt-AKF $\varphi_{yy}(k)$ und der Rausch-AKF $\varphi_{vv}(k)$ eine um das Rauschen bereinigte AKF berechnet, aus der sich ein zeitlich adaptiver Mindestschwellwert zur Diskriminierung des Signal- und des Rauschanteils ableiten lässt [Pat. 7.3].

Aufgrund der Mittelwertbildung ist die AKF $\varphi_{vv}(k)$ des Rauschens sehr viel kleiner als die verbleibende Summe $\varphi_{gg}(k) = \varphi_{ss}(k) + \varphi_{dd}(k) + \varphi_{sd}(k) + \varphi_{ds}(k)$ bei Auftreten der Signale $s(k)$ bzw. $d(k)$. An dieser Stelle ergibt sich daher im Vergleich zu einfacheren Auswerteverfahren ein verbessertes Signal-Rausch-Verhältnis. Dieser Umstand erklärt sich aus der Tatsache, dass ein höheres a-priori-Wissen zur Signalauswertung verwendet wird, indem man die aktuelle Signalform mit einem Ausschnitt der Signalhistorie vergleicht.

Die Summen-AKF $\varphi_{gg}(k)$ enthält allerdings die AKF $\varphi_{ss}(k)$ der Nutzsingale, die AKF $\varphi_{dd}(k)$ der vom Hintergrund hervorgerufenen Störsingale sowie die entsprechenden Kreuzkorrelationsterme $\varphi_{sd}(k)$ und $\varphi_{ds}(k)$. Unabhängig davon, ob die Kreuzkorrelationsterme generell vernachlässigbar sind, bereitet in der Praxis gerade die Trennung von ähnlichen Signalverläufen der Nutzanteile $s(k)$ und der Hintergrundstörungen $d(k)$ sowie die Erkennung von möglichen Überlagerungen die meisten Probleme.

In [Pat. 7.3] werden daher zusätzliche Kriterien zur Unterdrückung von Fehlalarmen angewendet. Störungen $d(k)$, die keine periodischen Signalanteile besitzen, können beispielsweise

anhand ihrer monoton fallenden AKF detektiert werden [Pat. 7.3]. Problematisch sind aber Störungen $d(k)$, die ähnliche spektrale Leistungsdichten wie die zu detektierenden Nutzsignale $s(k)$ aufweisen und somit auch eine vergleichbare AKF besitzen. Hier ist die Unterscheidung zwischen Nutz- und Störsignalen ähnlich schwierig wie bei der vorher beschriebenen Auswertung signalspezifischer Eigenschaften.

Die AKF ist prinzipiell sehr gut geeignet, um deterministische Anteile in verrauschten Signalen zu detektieren. Bei periodischen Signalen, die insbesondere auf Bewegungen transversal zum Sensor zurückzuführen sind, entstehen ausgeprägte Maxima und Minima im AKF-Verlauf $\varphi_{yy}(k)$. Das Hauptmaximum tritt bei der Verschiebung $k = 0$ auf und entspricht bei gleichanteilfreien Signalen der Wechselleistung des betrachteten Signalabschnitts. Das Auftreten von insbesondere gegenpoligen Nebenmaxima in der AKF sowie deren Ausprägung und deren Bewertung im Vergleich zum Hauptmaximum sind die wesentlichen Entscheidungskriterien in [Pat. 7.3]. Außerdem sind durch den zeitlichen Abstand der Extrema Rückschlüsse auf die Winkelgeschwindigkeit des bewegten Objekts möglich.

Speziell mit Hinblick auf die AKF-Auswertung von Draufzugehsignalen wird außerdem die Verwendung eines "Doppeldualsensors" vorgeschlagen [Pat. 7.3]. Zusätzlich zu zwei nebeneinander liegenden Detektoren für transversale Bewegungen sind zwei vertikal angeordnete Sensorelementen für radiale Gehrchtungen vorgesehen, um eine verbesserte Strahlungsmodulation in dieser Bewegungsrichtung herbeizuführen.

7.5.2 Kreuzkorrelationsfunktion zum Vergleich von Signalen

Zum Vergleich von Signalverläufen unterschiedlicher Signalquellen wird in [Pat. 7.4] die KKF verwendet. Während die AKF nach Abschnitt 7.5.1 zur Detektion von periodisch wiederholten Anteilen innerhalb eines Signals dient, liefert die KKF ein Ähnlichkeitsmaß für zwei unterschiedliche Signale $s(t)$ und $r(t)$. In [Pat. 7.4] wird der normierte Korrelationskoeffizient $C(t)$

$$(Gl. 7.14) \quad C(t) = \frac{\int_{-T/2}^{T/2} s(\kappa) r(\kappa) d\kappa - \frac{1}{T} \cdot \int_{-T/2}^{T/2} s(\kappa) d\kappa \cdot \int_{-T/2}^{T/2} r(\kappa) d\kappa}{\sqrt{\left[\int_{-T/2}^{T/2} s^2(\kappa) d\kappa - \frac{1}{T} \cdot \left[\int_{-T/2}^{T/2} s(\kappa) d\kappa \right]^2 \right] \cdot \left[\int_{-T/2}^{T/2} r^2(\kappa) d\kappa - \frac{1}{T} \cdot \left[\int_{-T/2}^{T/2} r(\kappa) d\kappa \right]^2 \right]}}$$

als Ähnlichkeitsmaß verwendet. Man erhält den Korrelationskoeffizient $C(t)$ nach (Gl. 7.14), wenn man die Kovarianz (Wechselleistung)

(Gl. 7.15)

$$P = E\{[s(t) - E\{s(t)\}] \cdot [r(t) - E\{r(t)\}]\} = \frac{1}{T} \int_{-T/2}^{T/2} s(\kappa) r(\kappa) d\kappa - \frac{1}{T} \int_{-T/2}^{T/2} s(\kappa) d\kappa \cdot \frac{1}{T} \int_{-T/2}^{T/2} r(\kappa) d\kappa$$

auf die Faktoren $\sqrt{P_s}$ und $\sqrt{P_r}$ nach (Gl. 7.16) und (Gl. 7.17) normiert. P_s und P_r entsprechen dabei den Wechselleistungen der Signale $s(t)$ und $r(t)$.

$$(Gl. 7.16) \quad \sqrt{P_s} = \sqrt{E\{[s(t) - E\{s(t)\}]^2\}} = \frac{1}{T} \int_{-T/2}^{T/2} s^2(\kappa) d\kappa - \left[\frac{1}{T} \int_{-T/2}^{T/2} s(\kappa) d\kappa \right]^2$$

$$(Gl. 7.17) \quad \sqrt{P_r} = \sqrt{E\{[r(t) - E\{r(t)\}]^2\}} = \frac{1}{T} \int_{-T/2}^{T/2} r^2(\kappa) d\kappa - \left[\frac{1}{T} \int_{-T/2}^{T/2} r(\kappa) d\kappa \right]^2$$

In [Pat. 7.4] wird zwischen zwei Möglichkeiten zur Vorgabe eines Referenzsignals $r(t)$ unterschieden. Im ersten Fall wird das Referenzsignal $r(k)$ von einem zusätzlichen Sensor geliefert, der einen anderen Entfernungsbereich als der Sensor überwacht, von dem das Signal $s(t)$ stammt. Auf diese Weise kann beispielsweise eine Vorwarnzone realisiert werden, wobei der Bewegungsmelder erst dann schaltet, wenn in der zweiten Zone ein ähnliches Signal $s(k)$ auftritt. Neben einem ausreichend hohen Korrelationskoeffizient wird außerdem eine gewisse Mindestamplitude für eine Alarmauslösung gefordert.

Eine weitere Auswertestrategie beinhaltet die Ermittlung von typischen Referenzsignalen in unterschiedlichen Entfernungen und Bewegungsrichtungen zum Sensor sowie den Vergleich mit dem aktuellen Sensorsignal $s(k)$. Die Referenzsignale werden in einem Speichermedium abgelegt, wodurch sich aufgrund der Vielzahl der möglichen Kurvenverläufe ein großer Speicherbedarf und im Rahmen der Auswertung ein hoher Signalverarbeitungsaufwand ergibt. Verwendet man Rechtecksignale als typische Referenzfolgen, kann der Speicher- und Berechnungsaufwand deutlich reduziert werden. Die Repräsentativität der verwendeten Referenzfolgen hat dabei entscheidende Bedeutung für das Schaltverhalten des Wächters.

Die gerade beschriebene Vorgehensweise stellt bereits einen ersten, modellbasierten Ansatz zur Verarbeitung von Bewegungsmeldersignalen dar. Der Ähnlichkeitsvergleich wurde dabei auf den Vergleich mit typischen Nutzsensignalen reduziert, welche in diesem Fall das zu detektierende Systemverhalten modellieren. Im nächsten Kapitel wird ebenfalls ein modellbasierter Ansatz mit einem KALMAN-Filter zur Signalvorverarbeitung präsentiert, bei dem die vorher beschriebene Problematik der Auswahl von repräsentativen Referenzfolgen entfällt.

7.6 Literatur

Literatur

- [Lit. 7.1] H. D. Lüke: "Signalübertragung", 3. erw. Aufl., korrigierter Nachdruck - Berlin; Heidelberg; New York; London; Paris; Tokyo : Springer Verlag, 1988
- [Lit. 7.2] A. V. Oppenheim, R. W. Schaffer: "Zeitdiskrete Signalverarbeitung", 3. durchges. Auflage - München; Wien : Oldenbourg Verlag, 1999

Patente

- [Pat. 7.1] EP 0 070 364 B1: "Bewegungsmelder zur Raumüberwachung", Richard Hirschmann Radiotechnisches Werk, D-7300 Esslingen a. N. (DE), 1982
- [Pat. 7.2] EP 0 825 573 A1: "Verfahren und Vorrichtung zur Auswertung eines Signals eines Bewegungsmelders", Siemens AG, 80333 München (DE), 1996
- [Pat. 7.3] DE 196 07 607 A1: "Verfahren zur Bewegungsmeldung mit mindestens einem Infrarotsensor und Bewegungsmelder zur Durchführung des Verfahrens", ABB Patent GmbH, 68309 Mannheim, DE, 1996
- [Pat. 7.4] EP 0 107 042 A1: "Infrarot-Detektor zur Feststellung eines Eindringlings in einem Raum", CERBERUS AG, 8708 Männedorf, CH, 1983

8 Modellbasierter Ansatz zur Bewegungserkennung

8.1 Einleitung

Basierend auf den Grundlagen der Estimationstheorie wird im Folgenden ein innovatives Signalverarbeitungskonzept für Bewegungsmeldersignale mit einem statistischen Filter vorgestellt. Das Filter verwendet intern ein zeitdiskretes lineares Zustandsraummodell des realen Prozesses und führt auf einen linearen, optimalen Estimationsalgorithmus, der nach seinem Entdecker “Kalman-Filter“ genannt wird. Aufgrund der Optimalitätseigenschaften von Kalman-Filtern und des systematischen Ansatzes ergibt sich eine optimale Verarbeitungsgüte, die analytisch und quantitativ erfassbar ist.

Eine sehr empfehlenswerte Darstellung zur dynamischen Systemmodellierung im Zustandsraum, zu den wahrscheinlichkeits- und estimationstheoretischen Grundlagen sowie zur Funktionsweise und Anwendung von Kalman-Filtern ist in [Lit. 8.1] und [Lit. 8.2] zu finden.

Im Rahmen dieser Arbeit werden nicht die grundlegenden theoretischen Zusammenhänge des Kalman-Filters hergeleitet. Vielmehr soll, ausgehend von der stochastischen Modellierung des Eingangssignals eines Bewegungsmelders, ein estimationstheoretischer Ansatz zur Bewegungserkennung präsentiert werden, wobei sich das Anwendungspotential dieser Vorgehensweise durch die gewählte Form der Darstellung intuitiv erfassen lässt [Lit. 8.4].

8.2 Systemkonzept mit Kalman-Filter

Im Vergleich zu konventionellen Filtern, die das Frequenzspektrum von Signalen ändern, ist das Kalman-Filter ein statistisches Filter, das einen möglichst “optimalen“ Schätzwert eines unbekanntem Systemzustands aus verrauschten oder fehlerbehafteten Messdaten ermittelt. Ziel der “optimalen“ Estimation ist die Minimierung eines sinnvollen Fehlerkriteriums. In [Lit. 8.2] wird gezeigt, dass ein Kalman-Filter das optimale lineare Filter in Verbindung mit linearen Zustandsraummodellen darstellt.

Intern verwendet das in Abb. 8.1 dargestellte Kalman-Filter ein zeitdiskretes, lineares Zustandsraummodell des physikalischen (realen) Prozesses. Das Filter ist ein linearer, rekursiver Algorithmus, der Kenntnisse der Anfangswerte, über das dynamische Verhalten des Systems und der Sensoren sowie über die statistischen Eigenschaften der beteiligten Rauschprozesse benötigt.

Der Bewegungsmelder ist ein Teil des physikalischen Systems (s. Abb. 8.1). Wie bereits in Kap. 6 und 7 beschrieben wurde, verursachen Veränderungen der Temperaturverteilung im Erfassungsfeld entsprechende Variationen der Sensorausgangsspannung. Da weder der Zeitpunkt des Auftretens noch die Stärke der Anregung im Voraus bekannt sind, können Änderungen der Temperaturverteilung auch als stochastischer Prozess aufgefasst werden. Auf-

grund des differenzierenden Verhaltens von pyroelektrischen Detektoren ergibt sich bei stabilen thermischen Verhältnissen im Gesichtsfeld des Sensors eine konstante, temperaturunabhängige Ausgangsspannung.

Ausgehend von diesem Ansatz wird das Kalman-Filter insbesondere zur Vorverarbeitung der Sensordaten verwendet. Es bewirkt den Abgleich des modellierten und des tatsächlich beobachteten Systemverhaltens.

Im Anschluss an die Filterung sind weitere Stufen zur Verarbeitung charakteristischer Zustände des Kalman-Filters und zur Entscheidungsfindung erforderlich.

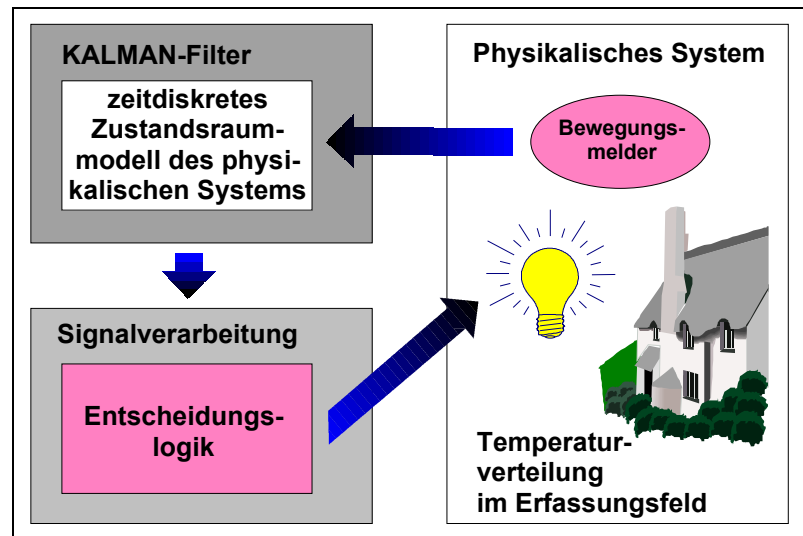


Abb. 8.1: Allgemeine Systembeschreibung bei Verwendung eines Kalman-Filters

8.3 Abgleichprozess zwischen Model und realem Verhalten

Ziel der Estimationstheorie ist die Gewinnung von Informationen aus gestörten, fehlerbehafteten Messwerten, in diesem Fall aus der "Beobachtung" der zeitkontinuierlichen Ausgangsspannung $y(t)$ eines Bewegungsmelders (s. Abb. 8.2). Der Ausdruck "Beobachtung" kennzeichnet dabei die Beeinflussung der Messergebnisse durch den Bewegungsmelder und somit die systematische Abhängigkeit der Messwerte von der verwendeten Sensorik. Die Korrektur des systematischen Fehlers erfordert Kenntnisse über das dynamischen Verhalten des Bewegungsmelders.

Unter dem Begriff "Rauschen" sind alle nichtreproduzierbaren Fehlereinflüsse zusammengefasst. Der Eingangsrauschprozess $w(t)$ und das ständig vorhandene Messrauschen $v(t)$ am Sensorausgang repräsentieren die stochastischen Komponenten der Messung.

Im Rahmen dieser Arbeit wird dem Eingangsrauschen, das auch als "Driving Noise" bezeichnet wird, eine besondere Bedeutung zugewiesen. Aufgrund des vorne beschriebenen, zufälligen Charakters der Temperaturänderungen im Erfassungsfeld werden diese Variationen selbst als stochastischer Rauschprozess $w(t)$ betrachtet. Abb. 8.2 verdeutlicht die Definitionen und vermittelt einen ersten Eindruck von der Wirkungsweise des Kalman-Filters.

Infolge der äquidistanten Abtastung des Signals $y(t)$ mit dem Abtastintervall T ergibt sich eine Folge von Messwerten $y(kT)$ mit $k = 0, 1, \dots \infty$. Zur Vereinfachung der Schreibweise wird im folgenden mit $T=1$ eine Normierung des Abtastintervalls durchgeführt. Die diskrete Zeitabhängigkeit wird fortan durch das Argument k angedeutet.

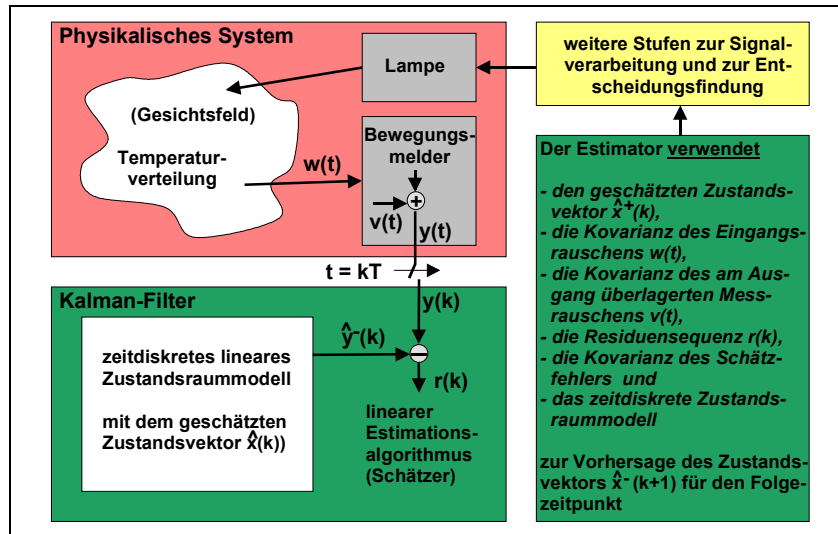


Abb. 8.2: "Beobachtung" der Ausgangssignale eines Bewegungsmelders und Abgleich mit dem Systemmodell

Für jeden Abtastzeitpunkt k berechnet das Kalman-Filter die Differenz $r(k)$ der Messgröße $y(k)$ und des "vorhergesagten" Beobachtungswertes $\hat{y}^-(k)$, wobei der zuletzt genannte Wert aus einer linearen Abbildung des vorhergesagten Zustandsvektors $\hat{x}^-(k)$ in den Beobachtungsraum hervorgeht. Insbesondere die Residuensequenz $r(k)$ des Kalman-Filters sowie die Kovarianzen des Prädiktionsfehlers und des Messrauschens dienen zur Zustandsestimation und demzufolge zum "Abgleich" zwischen modelliertem und realem Systemverhalten.

Zur Zustandsprädiktion wird der vorhergehende optimale Schätzwert unter Berücksichtigung des dynamischen Systemverhaltens und ggf. unter Beachtung determinierter Anregungen verwendet (Abb. 8.3). Beim Kalman-Filter muss streng zwischen der Prädiktion, also der Vorhersage von Zuständen vor Verfügbarkeit des nächsten Messwertes, und der Estimation von Systemzuständen unterschieden werden.

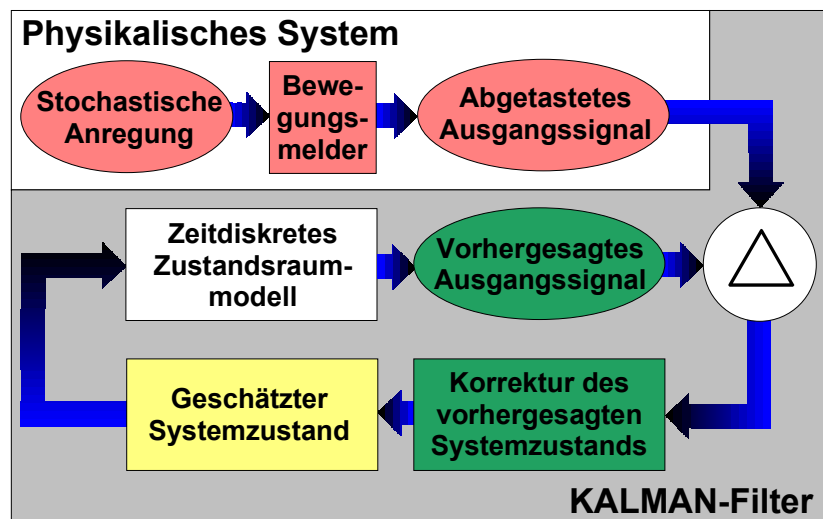


Abb. 8.3: Prädiktions-, Estimations- und Abgleichprozess innerhalb des Kalman-Filters. Das Zeichen " Δ " verdeutlicht die Bildung der Residuensequenz.

Die Schätzung (Estimation) des Zustandsvektors $\hat{x}^+(k)$ zum Zeitpunkt k erfolgt anhand des vorhergesagten Zustandsvektors $\hat{x}^-(k)$ zum Zeitpunkt k unter Berücksichtigung des aktuellen Messwertes $y(k)$. Der Prädiktionsprozeß wird im weiteren Verlauf durch ein hochgestelltes “-“-Zeichen angedeutet, während der optimale Schätzwert durch ein hochgestelltes “+“-Zeichen gekennzeichnet ist.

8.4 Zeitkontinuierliche und zeitdiskrete Zustandsraummodelle

In diesem Abschnitt wird die allgemeine Vorgehensweise zur Ableitung eines zeitdiskreten Zustandsraummodells für das Kalman-Filters beschrieben. In Kap. 9 wird diese Vorgehensweise zur Entwicklung eines vereinfachten Modells für den Vorgang der Bewegungserkennung angewendet. Ausgangspunkt für die folgenden Betrachtungen ist das allgemeine, zeitkontinuierliche Zustandsraummodell nach Abb. 8.4, das, wie später gezeigt wird, aus der allgemeinen Übertragungsfunktion eines linearen zeitinvarianten Systems abgeleitet werden kann. Abb. 8.4 soll zur Definition der später verwendeten Bezeichnungen dienen.

Der Zustandsvektor $\underline{x}(t)$ beschreibt die internen Zustände des Modells, die im vorliegenden Anwendungsfall mit der Beobachtungsmatrix C_c in ein skalares Signal abgebildet werden. Dem Ausgangssignal $y(t)$ ist zusätzlich das Messrauschen $v(t)$ überlagert.

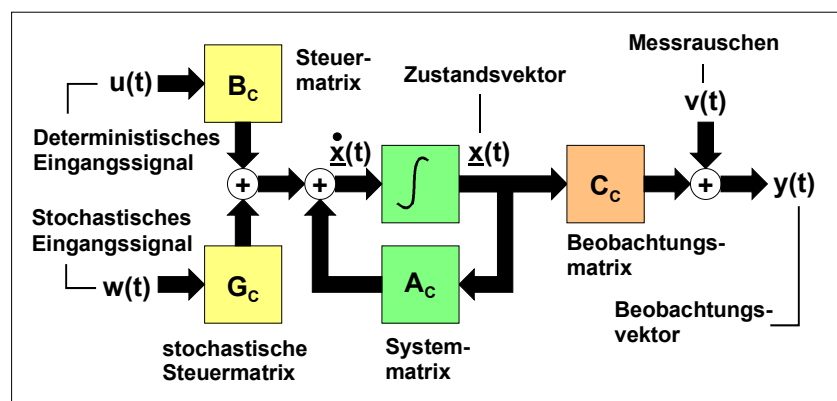


Abb. 8.4: Allgemeines zeitkontinuierliches Zustandsraummodell

Änderungen der internen Systemzustände, die durch die erste Ableitung $\dot{\underline{x}}(t)$ des Zustandsvektors $\underline{x}(t)$ repräsentiert werden, können ganz allgemein von deterministischen und stochastischen Eingangssignalen $u(t)$ und $w(t)$ hervorgerufen werden. Der Einfluss der hier als skalar angenommenen Steuersignale wird durch die beiden Steuermatrizen B_c und G_c berücksichtigt, ein direkter Durchgriff auf das Ausgangssignal $y(t)$ ist nicht vorgesehen.

Zustandsübergänge können auch ohne externe Anregung erfolgen. Ausgehend vom aktuellen Systemzustand $\underline{x}(t)$ erfolgen diese Änderungen unter Beachtung des dynamischen Systemverhaltens, das durch die Zustandsübergangsmatrix A_c modelliert wird. Unter Beachtung von Abb. 8.4 kann die allgemeine Zustandsdifferentialgleichung (Gl. 8.1), die auch als lokale Zustandsübergangsfunktion $f(\underline{x}(t), u(t), t)$ bezeichnet wird, und das Beobachtungsmodell (Gl. 8.2) ermittelt werden.

$$(Gl. 8.1) \quad \dot{\underline{x}}(t) = A_c \cdot \underline{x}(t) + B_c \cdot u(t) + G_c \cdot w(t) \quad (Gl. 8.2) \quad y(t) = C_c \cdot \underline{x}(t) + v(t)$$

Für die weiteren Betrachtungen wird zunächst die Existenz eines skalaren, deterministischen Steuersignals $u(t)$ und einer skalaren stochastischen Anregung $w(t)$ angenommen. Sie werden nach Abb. 8.4 über die Steuermatrizen B_c und G_c eingekoppelt. In [Lit. 8.2] wird gezeigt, dass aus der lokalen Zustandsübergangsfunktion (Gl. 8.1) die globale Zustandsübergangsfunktion $g(\underline{x}(t_0), u(t), w(t), t)$ nach (Gl. 8.3) ermittelt werden kann, welche die allgemeine Lösung der Zustandsdifferentialgleichung (Gl. 8.1) darstellt.

$$(Gl. 8.3) \quad \underline{x}(t) = \Phi(t, t_0) \cdot \underline{x}(t_0) + \int_{t_0}^t \Phi(t, \tau) \cdot B_c \cdot u(\tau) d\tau + \int_{t_0}^t \Phi(t, \tau) \cdot G_c \cdot w(\tau) d\tau$$

Die in (Gl. 8.3) auftretende Matrix $\Phi(t, t_0)$ heißt globale Zustandsübergangsmatrix. Für ein lineares, zeitinvariantes System ist diese Matrix nur abhängig vom betrachteten Zeitintervall $t-t_0$, so dass allgemein gilt $\Phi(t, t_0) = \Phi(t-t_0)$. In [Lit. 8.1] wird gezeigt, dass die globale Zustandsübergangsmatrix $\Phi(t, t_0)$ bei bekannter Zustandsübergangsmatrix A_c anhand von (Gl. 8.4) berechnet werden kann.

$$(Gl. 8.4) \quad \Phi(t, t_0) = \Phi(t-t_0) = L^{-1} \left\{ [s \cdot I - A_c]^{-1} \right\}$$

$L^{-1}\{\dots\}$ beschreibt dabei die inverse Laplace-Transformation der Matrix $[sI-A_c]^{-1}$, die durch Inversion der Matrix $[sI-A_c]$ entsteht. Die Matrix I ist eine Einheitsmatrix, deren Zeilen- und Spaltenzahl mit denen der Zustandsübergangsmatrix A_c korrespondiert. Die Matrix $\Phi(t, t_0)$ wird im weiteren Verlauf zur Ableitung des äquivalenten, zeitdiskreten Zustandsraummodells für das KALMAN-Filter benötigt.

Die globale Zustandsübergangsfunktion aus (Gl. 8.3) ermöglicht die Berechnung des Systemverhaltens im Zeitintervall $[t_0, t)$, wenn die Anfangsbedingungen $\underline{x}(t_0)$ und der Verlauf der deterministischen Steuerfunktion $u(t)$ in diesem Zeitraum bekannt sind. Der rechte Summand in (Gl. 8.3) muss als formale Schreibweise eines stochastischen Integrals betrachtet werden. Er berücksichtigt die Wirkung der stochastischen Anregung $w(t)$ im betrachteten Zeitintervall. Nach [Lit. 8.1] folgt für den Mittelwert $\underline{m}_x(t)$ des Zustandsvektors $\underline{x}(t)$ unter Beachtung der Linearität der Erwartungswertbildung $E\{\dots\}$ und der Integration aus

$$\begin{aligned} \underline{m}_x(t) &= E\{\underline{x}(t)\} = E\left\{ \Phi(t, t_0) \cdot \underline{x}(t_0) \right\} + E\left\{ \int_{t_0}^t \Phi(t, \tau) \cdot B_c \cdot u(\tau) d\tau \right\} + E\left\{ \int_{t_0}^t \Phi(t, \tau) \cdot G_c \cdot w(\tau) d\tau \right\} = \\ &= E\{\underline{x}(t)\} = \Phi(t, t_0) \cdot E\{\underline{x}(t_0)\} + \int_{t_0}^t \Phi(t, \tau) \cdot B_c \cdot u(\tau) d\tau + \int_{t_0}^t \Phi(t, \tau) \cdot G_c \cdot E\{w(\tau)\} d\tau \end{aligned}$$

für einen mittelwertfreien, weißen Rauschprozess mit $E\{w(t)\} = 0$ die Beziehung

$$(Gl. 8.5) \quad \underline{m}_x(t) = E\{\underline{x}(t)\} = \Phi(t, t_0) \cdot \underline{m}_x(t_0) + \int_{t_0}^t \Phi(t, \tau) \cdot B_c \cdot u(\tau) d\tau .$$

Für die Kovarianz $P_{\underline{x}\underline{x}}(t)$ des Zustandsvektors $\underline{x}(t)$ erhält man aus

$$P_{\underline{x}\underline{x}}(t) = E \left\{ \left[\underline{x}(t) - \underline{m}_{\underline{x}}(t) \right] \cdot \left[\underline{x}(t) - \underline{m}_{\underline{x}}(t) \right]^T \right\} =$$

$$E \left\{ \left[\Phi(t, t_0) \underline{x}(t_0) - \underline{m}_{\underline{x}}(t) + \int_{t_0}^t \Phi(t, \tau) G_c w(\tau) d\tau \right] \cdot \left[\Phi(t, t_0) \underline{x}(t_0) - \underline{m}_{\underline{x}}(t) + \int_{t_0}^t \Phi(t, \tau) G_c w(\tau) d\tau \right]^T \right\}$$

die Gleichung (Gl. 8.6), in welcher der Term $E\{w(t) \cdot w(t)^T\}$ durch die im allgemeinen Fall zeitabhängige Varianz $Q(t)$ des Eingangsruschprozesses $w(t)$ ersetzt werden kann. Zur Herleitung von (Gl. 8.6) wurde außerdem die statistische Unabhängigkeit des Rauschprozesses $w(t)$ vom Anfangszustand $\underline{x}_0(t)$ vorausgesetzt.

(Gl. 8.6)

$$P_{\underline{x}\underline{x}}(t) = \Phi(t, t_0) \cdot P_{\underline{x}\underline{x}}(t_0) \cdot \Phi(t, t_0)^T + \int_{t_0}^t \int_{t_0}^t \Phi(t, \tau) \cdot G_c \cdot E\{w(\tau) \cdot w(\nu)^T\} G_c^T \cdot \Phi(t, \nu)^T d\tau d\nu$$

Ein zeitdiskretes System ist äquivalent zu einem zeitkontinuierlichen Modell, wenn an den diskreten Abtastzeitpunkten $t = kT$ die Ausgangssignale und die Zustandsvektoren beider Systeme übereinstimmen. Zur umkehrbar eindeutigen Zuordnung zwischen beiden Modellen muss das Abtastintervall T das Nyquist-Kriterium erfüllen. Bei der Festlegung der Abtastrate muss das zeitliche Verhalten des Ausgangssignals und des Zustandsvektors berücksichtigt werden [Lit. 8.1]. Mit der Normierung $T = 1$ gelten für das allgemeine zeitdiskrete Zustandsraummodell nach Abb. 8.5 die Beziehungen

(Gl. 8.7) $\underline{x}(k+1) = A \cdot \underline{x}(k) + B \cdot u(k) + G \cdot w(k)$ (Gl. 8.8) $y(k) = C \cdot \underline{x}(k) + v(k)$

Die Matrizen A , B und C des zeitdiskreten Modells in Abb. 8.5 lassen sich mit der Übergangsmatrix $\Phi(t, t_0) = \Phi(t - t_0)$ des zeitkontinuierlichen Modells für $t_0 = kT$ und $t = (k+1)T$ berechnen.

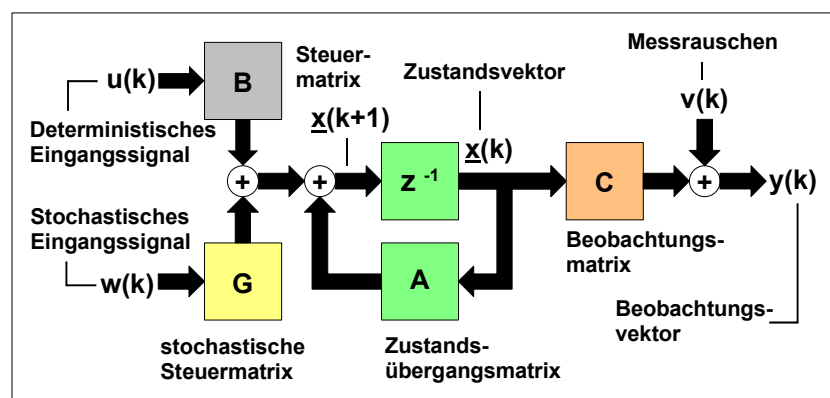


Abb. 8.5: Allgemeines zeitdiskretes Zustandsraummodell

Die Berechnungsvorschriften zur Umwandlung eines zeitkontinuierlichen in ein äquivalentes zeitdiskretes Zustandsraummodell sind in [Lit. 8.1] angegeben. Allgemein gelten für lineare zeitinvariante Systeme die Beziehungen:

(Gl. 8.9) $A = \Phi(t - t_0) = \Phi((k+1)T - kT) = \Phi(T) \approx I + A_c \cdot T$

$$(Gl. 8.10) \quad B = \int_{t_0}^t \Phi(t - \tau) \cdot B_c \, d\tau = \int_{kT}^{(k+1)T} \Phi((k+1)T - \tau) \cdot B_c \, d\tau \approx \left[I + \frac{1}{2} A_c \cdot T \right] \cdot B_c \cdot T$$

$$(Gl. 8.11) \quad C = C_c$$

Zur Umwandlung der Matrix B_c wurde eine im Abtastintervall konstante Steuergröße $u(t)$ angenommen. Ersetzt man die Terme $B \cdot u(k)$ und $G \cdot w(k)$ in (Gl. 8.7) durch die Vektoren $\underline{u}(k)$ und $\underline{w}(k)$, erhält man

$$(Gl. 8.12) \quad \underline{x}(k+1) = A \cdot \underline{x}(k) + \underline{u}(k) + \underline{w}(k).$$

Während die Matrix B nach (Gl. 8.10) immer berechnet werden kann, ist die Existenz der zeitdiskreten stochastischen Kontrollmatrix G nicht gesichert. Stattdessen wird oftmals der zeitdiskrete vektorielle Rauschprozess $\underline{w}(k)$ verwendet, der sich aus (Gl. 8.13) ergibt.

$$(Gl. 8.13) \quad \underline{w}(k) = \int_{t_0}^t \Phi(t - \tau) \cdot G_c w(\tau) \, d\tau = \int_{kT}^{(k+1)T} \Phi((k+1)T - \tau) \cdot G_c w(\tau) \, d\tau$$

(Gl. 8.12) besagt, dass der Zustandsvektor $\underline{x}(k+1)$ zum Zeitpunkt $(k+1)T$ aus den Werten des Zustandsvektors $\underline{x}(k)$, der vektoriellen Steuergröße $\underline{u}(k)$ und des zeitdiskreten Rauschprozesses $\underline{w}(k)$ zum unmittelbar vorhergehenden Zeitpunkt kT ermittelt werden kann. Aufgrund des zeitdiskreten und im folgenden als erwartungswertfrei angenommenen Rauschprozesses $\underline{w}(k)$ muss der Zustandsvektor $\underline{x}(k)$ als Zufallsvariable betrachtet werden. Für den zeitdiskreten Mittelwert $\underline{m}_x(k)$ gilt in diesem Fall:

$$(Gl. 8.14) \quad \underline{m}_x(k+1) = \Phi(T) \cdot E\{\underline{x}(0)\} + E\{\underline{u}(k)\} + \int_{kT}^{(k+1)T} \Phi((k+1)T - \tau) G_c E\{w(\tau)\} d\tau = \Phi(T) \cdot \underline{m}_x(k) + \underline{u}(k)$$

Mit der zeitdiskreten Kovarianz $Q(k)$ des Rauschprozesses $\underline{w}(k)$ erhält man die zeitdiskrete Kovarianz $\underline{P}_{xx}(k+1)$ des Zufallsvektors $\underline{x}(k+1)$ aus (Gl. 8.15), wobei die Herleitung analog zur Vorgehensweise bei der zeitkontinuierlichen Modellierung erfolgt.

$$(Gl. 8.15) \quad \underline{P}_{xx}(k+1) = \Phi(T) \cdot \underline{P}_{xx}(k) \cdot \Phi(T)^T + Q(k)$$

Wie später beschrieben wird, verwendet das Kalman-Filter intern die Matrizen A , B und C des zeitdiskreten Zustandsraummodells sowie die Kovarianzmatrix $Q(k)$ des Rauschprozesses $\underline{w}(k)$ am Systemeingang. Die Kovarianzmatrix $Q(k)$ des zeitdiskreten Systems erhält man aus der Kovarianz $Q(t)$ des zeitkontinuierlichen Modells anhand der Berechnungsvorschrift (Gl. 8.16) [Lit. 8.1].

$$(Gl. 8.16) \quad Q(k) = \int_{kT}^{(k+1)T} \Phi((k+1)T - \tau) \cdot G_c \cdot Q(\tau) \cdot G_c^T \cdot \Phi((k+1)T - \tau)^T \, d\tau$$

Im folgenden wird zunächst davon ausgegangen, dass ein äquivalentes zeitdiskretes Systemmodell vorliegt und dass die stochastische Steuermatrix G existiert. In einem nachfolgenden Kapitel wird dann gezeigt, dass anstelle des Summands $G \cdot w(k)$ der Rauschvektor $\underline{w}(k)$ zur Systemmodellierung verwendet werden muss. Im Rahmen der nachstehenden, allgemeinen Betrachtungen werden beide Beschreibungen zunächst als alternative Darstellungsformen betrachtet.

8.5 Modellierung von Nutz- und Störsignalen

Um eine möglichst gute Diskriminierung zwischen Nutzsignalen und Störsignalen zu erzielen, sollte die Signalverarbeitung modellgestützt erfolgen. Um die Klasse der detektierbaren Signalverläufe im Voraus nicht zu sehr einzuschränken, wird im weiteren Verlauf ein sehr allgemeiner Ansatz zur Modellierung von Hintergrundstörungen verwendet.

Da bei der vorliegenden Aufgabenstellung weder der Zeitpunkt des Auftretens noch die Größe der an sich skalaren, deterministischen Anregung $u(k)$ am Eingang des zeitdiskreten Zustandsraummodells nach (Gl. 8.7) bekannt ist, wird im folgenden lediglich eine stochastische Anregung $w(k)$ als Systeminput verwendet. Die stochastische Steuergröße werde dabei in derselben Weise wie das eigentlich deterministische Steuersignal eingekoppelt, so dass gilt:

$$(Gl. 8.17) \quad G_c = B_c$$

Der skalare Eingangsprozess $w(k)$ wird in einen Nutzsignalanteil $w_s(k)$ und einen Störanteil $w_d(k)$ zerlegt (s. Abb. 8.6). Der Nutzanteil $w_s(k)$ entsteht infolge von typischen Temperaturänderungen, die durch eine Person im Erfassungsfeld des Wächters hervorgerufen werden, während der Anteil $w_d(k)$ alle potentiellen Störsignalquellen berücksichtigt.

Ausgangsseitig verursacht das Signal $w_s(k)$ den Anteil $s(k)$ im digitalisierten Ausgangssignal $y(k)$. Der Störterm $w_d(k)$ bewirkt den Störanteil $d(k)$ im Ausgangssignal $y(k)$. Für die diskrete Zustandsdifferentialgleichung (Gl. 8.7) folgt in diesem Fall

$$(Gl. 8.18) \quad \underline{x}(k+1) = A \cdot \underline{x}(k) + B \cdot w(k) = A \cdot \underline{x}(k) + B \cdot [w_s(k) + w_d(k)]$$

Da insbesondere das Auftreten des Signals $w_s(k)$ detektiert werden soll, lassen sich prinzipiell zwei Fälle unterscheiden. In der Hypothese H_0 (kein bewegtes Objekt erkannt) gilt (Gl. 8.19), in der Hypothese H_1 (bewegtes Objekt erkannt) wird das zusätzlich vorhandene Eingangssignal durch (Gl. 8.20) berücksichtigt wird.

$$(Gl. 8.19) \quad w_s(k)|_{H_0} = 0 \qquad (Gl. 8.20) \quad w_s(k)|_{H_1} = w_s(k)$$

Der Rauschprozess $w_d(k)$ am Systemeingang, der alle Störeinflüsse zusammenfasst, wird im Rahmen des Modells als Gaußscher Rauschprozess (s. Abb. 8.7) mit der Verteilungsdichtefunktion $f_d(\xi)$ nach (Gl. 8.21) betrachtet.

Die Funktion $f_d(\xi)$ nach (Gl. 8.21) ist demzufolge durch die Angabe ihres Erwartungswertes $m_d(k)$ und der Kovarianz $P_{dd}(k)$ vollständig bestimmt.

Der Schwellwert ξ dient zur Ermittlung der Wahrscheinlichkeitsverteilung $F_d(\xi)$ des Rauschprozesses $w_d(k)$ (s. Abb. 8.8). Die Verteilungsfunktion $F_d(\xi)$ beschreibt dabei ganz allgemein die Wahrscheinlichkeit, das die Abtastwerte $w_d(k)$ kleiner oder gleich der Schwelle ξ sind.

Die Verteilungsdichtefunktion $f_d(\xi)$ erhält man als erste Ableitung der Verteilungsfunktion $F_d(\xi)$ nach der Pseudovariablen ξ . Die Gesamtfläche unter der Funktion $f_d(\xi)$ hat den Wert 1.

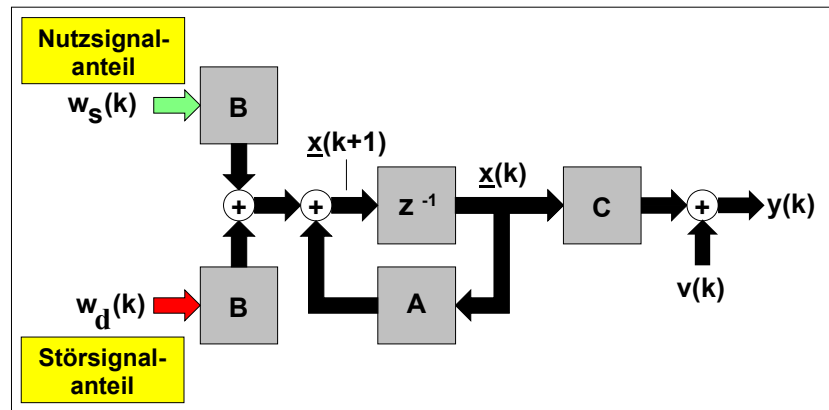


Abb. 8.6: Modellierung des Eingangssignals als stochastischer Prozess mit Nutzsiganalanteil $w_s(k)$ und Störsiganalanteil $w_d(k)$

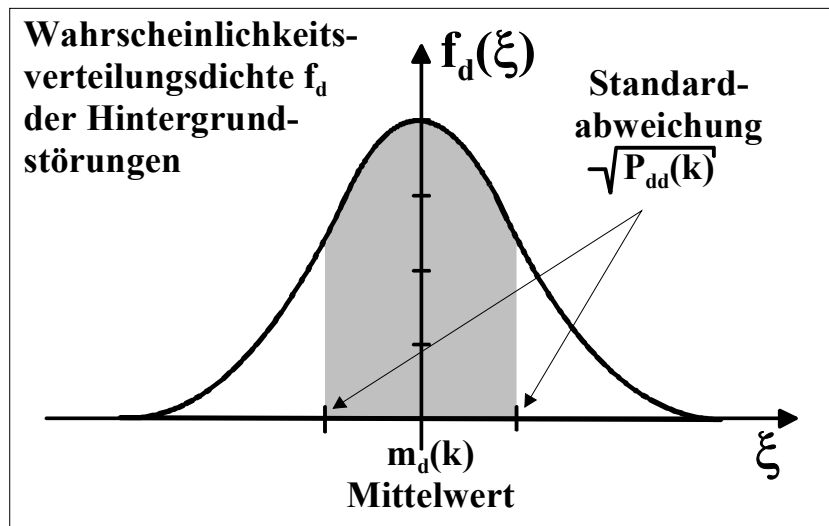


Abb. 8.7: Hintergrundstörungen als mittelwertfreier, gaußverteilter Prozess mit der Verteilungsdichte $f_d(\xi)$

$$(Gl. 8.21) \quad f_d(\xi) = \frac{1}{\sqrt{2\pi \cdot P_{dd}(k)}} \cdot \exp\left(-\frac{(\xi - m_d(k))^2}{2 \cdot P_{dd}(k)}\right)$$

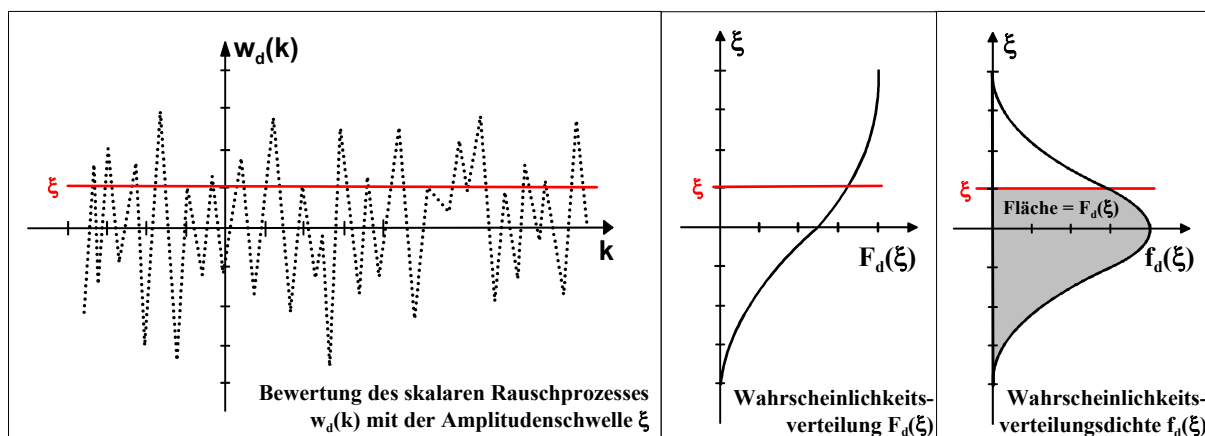


Abb. 8.8: Rauschprozess $w_d(k)$ und Ermittlung der Wahrscheinlichkeitsverteilung $F_d(\xi)$ sowie der zugehörigen Verteilungsdichtefunktion $f_d(\xi)$

Im folgenden wird ein mittelwertfreier Gaußprozess zur Modellierung des Hintergrundrauschens angenommen. Die Mittelwertfreiheit berücksichtigt dabei bereits das differenzierende Verhalten eines pyroelektrischen Detektors. Unter Berücksichtigung von (Gl. 8.21) ist die Verteilungsdichtefunktion in diesem Fall durch die Angabe des Mittelwertes $m_d(k)$, der nach (Gl. 8.22) als Erwartungswert erster Ordnung berechnet wird, und der zeitabhängigen Kovarianzmatrix $P_{dd}(k)$ nach (Gl. 8.23) vollständig beschrieben .

$$(Gl. 8.22) \quad m_d(k) = E\{w_d(k)\} = m_d = 0$$

$$(Gl. 8.23) \quad P_{dd}(k) = E \left\{ [w_d(k) - E\{w_d(k)\}]^2 \right\} = E \left\{ [w_d(k)]^2 \right\} = Q(k)$$

8.6 Sensorielles Beobachtungsmodell

Ausgehend von der zuvor beschriebenen Modellierung des Nutzsignals $w_s(k)$ und des Störsignals $w_d(k)$ am Systemeingang soll in diesem Abschnitt das beobachtbare (messbare) Ausgangssignal $y(k)$ nach (Gl. 8.8) analysiert werden. Wie bereits vorher beschrieben wurde, entsteht das Ausgangssignal $y(k)$ infolge von Temperaturänderungen im Erfassungsfeld des Sensors. Temperaturänderungen können dabei durch bewegte Objekte mit einer vom Hintergrund abweichenden Temperatur oder durch Änderungen der Hintergrundtemperatur, verursacht durch äußere Störeinflüsse wie Wind, Sonne, etc. hervorgerufen werden. Die Störsignale sind von langkorrelierter Natur und besitzen unter Umständen gleiche Eigenschaften wie das Nutzsignal, von dem sie diskriminiert werden müssen.

Im folgenden wird davon ausgegangen, dass das Eingangssignal $w_s(k)$ den Nutzsignalanteil $s(k)$ im Ausgangssignal $y(k)$ hervorruft, während die Störung $w_d(k)$ den Störterm $d(k)$ im Ausgangssignal $y(k)$ gemäß Abb. 8.9 bewirkt. Zusätzlich ist die Überlagerung von Nutz- und Störsignal verrauscht. Es kann angenommen werden, dass die Rauschstörungen $v(k)$ erwartungswertfrei und Gaußverteilt sind. Das skalare Ausgangssignal $y(k)$ ergibt sich als lineare Überlagerung des Signals $s(k)$, des Drift- bzw. Hintergrundsignals $d(k)$ und des Rauschanteils $v(k)$.

Für den Mittelwert $m_v(k)$ und die Kovarianz $P_v(k)$ folgt in diesem Fall:

$$(Gl. 8.24) \quad m_v(k) = E\{v(k)\} = 0$$

$$(Gl. 8.25) \quad P_{vv}(k) = E\{v(k)^2\} = R(k)$$

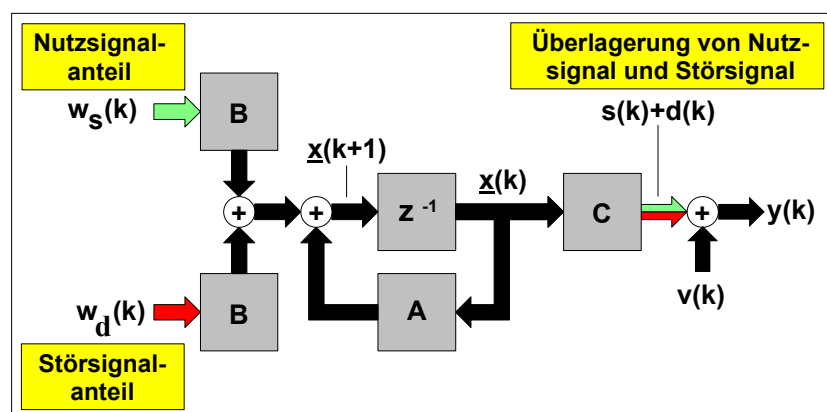


Abb. 8.9: Sensorielles Beobachtungsmodell unter Beachtung der Anregungen am Signaleingang

Unter Beachtung der Hypothesen H_0 und H_1 aus Abschnitt 8.5 lässt sich das sensorielle Beobachtungsmodell (Gl. 8.26) und (Gl. 8.27) formulieren. In der Hypothese H_0 (keine Bewegung erkannt) gilt

(Gl. 8.26)

$$y(k)|_{H_0} = d(k) + v(k),$$

in der Hypothese H_1 (Bewegung erkannt) folgt

(Gl. 8.27)

$$y(k)|_{H_1} = s(k) + d(k) + v(k)$$

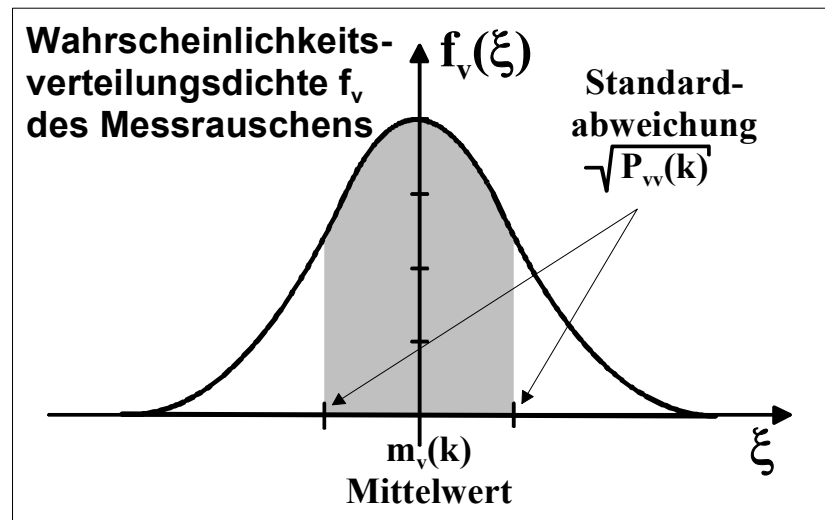


Abb. 8.10: Modellierung des Messrauschens am Systemausgang als mittelwertfreier, Gaußverteilter Prozess mit der Verteilungsdichte $f_v(\xi)$

An dieser Stelle soll darauf hingewiesen werden, dass Nutz- und Hintergrundsignale keineswegs als unkorreliert angenommen werden können. Das von beweglichen Objekten hervorgerufene Nutzsignal kann ähnliche spektrale Leistungsdichten aufweisen wie die Hintergrundstörungen.

8.7 Allgemeine Beschreibung des Kalman-Filters

In den vorhergehenden Abschnitten wurde die Berücksichtigung statistischer Messfehler im Ausgangssignal und die Modellierung der Hintergrundstörungen durch voneinander unabhängige, Gaußverteilte und erwartungswertfreie Rauschprozesse mit den zeitabhängigen Varianzen $R(k)$ und $Q(k)$ beschrieben. Bevor das weitere Verfahren zur Bewegungserkennung erläutert wird, sollen zunächst die Filtergleichungen des Kalman-Filters angegeben und dessen prinzipielle Wirkungsweise anhand von Abb. 8.11 und Abb. 8.12 erläutert werden.

Abb. 8.11 veranschaulicht die einzelnen Verarbeitungsschritte im Kalman-Filter. Man erkennt unmittelbar, dass die vom zeitdiskreten Systemmodell bekannten Matrizen A , B und C sinngemäß berücksichtigt sind. Der Block " z^{-1} " bezeichnet ein Verzögerungsglied erster Ordnung. Der eigentliche Schätzworgang läuft in mehreren Stufen ab, die im Folgenden nacheinander vorgestellt werden sollen. Die Zustandsprädiktion wird dabei durch ein hochgestelltes “-“-Zeichen angedeutet, um anzuzeigen, dass zu diesem Zeitpunkt noch kein neuer Messwert vorliegt. Der optimale Schätzwert wird demgegenüber durch ein hochgestelltes “+“-Zeichen gekennzeichnet, da er erst im Anschluss an eine neue Messung berechnet werden kann.

Zur Beschreibung ist die Definition der Messwert-historie $\underline{Y}(k)=[y_0 y_1 \dots y_k]^T$ mit den konkreten Realisationen $y_k = y(k)$ des Ausgangssignals erforderlich, die alle bis einschließlich zum Zeitpunkt k aufgetretenen Messwerte enthält.

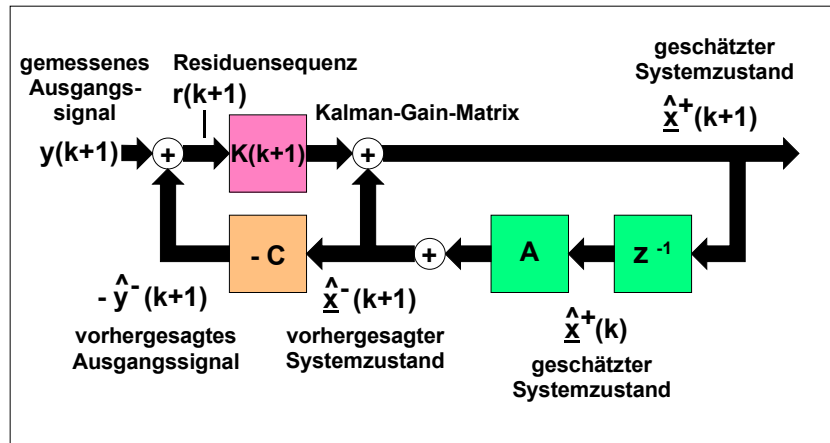


Abb. 8.11: Prinzipielle Struktur des Kalman-Filters

Der Algorithmus verwendet die auf die Messwertgeschichten zu den Zeitpunkten $k+1$ und k bedingten Verteilungsdichtefunktionen zur Zustandsprädiktion und -estimation (Abb. 8.12). Die Prädiktionsdichte $f_{\underline{x}(k+1)|\underline{Y}(k)}(\underline{\xi}|\underline{Y}(k))$ beinhaltet dabei die gesamte statistische Information, bedingt auf die Messwerthistorie $\underline{Y}(k)$, um den Zustandsvektor $\underline{\hat{x}}^-(k+1)$ für den unmittelbar folgenden Zeitpunkt vorherzusagen. Der anhand von (Gl. 8.28) ermittelte Prädiktionswert $\underline{\hat{x}}^-(k+1)$ entspricht dem linearen Erwartungswert $E\{\underline{x}(k+1)|\underline{Y}(k)\}$ des Zustandsvektors $\underline{x}(k+1)$, bedingt auf die Messwertgeschichte $\underline{Y}(k)$.

Nachdem ein neuer Messwert $y_{k+1} = y(k+1)$ vorliegt, wird die Filterdichte $f_{\underline{x}(k+1)|\underline{Y}(k+1)}(\underline{\xi}|\underline{Y}(k+1))$, bedingt auf die neue Historie $\underline{Y}(k+1)$ verwendet, um aus dem Prädiktionswert $\underline{\hat{x}}^-(k+1)$ den optimalen Schätzwert $\underline{\hat{x}}^+(k+1)$ zu bestimmen. Der optimale Schätzwert $\underline{\hat{x}}^+(k+1)$ nach (Gl. 8.28) ist der lineare Erwartungswert $E\{\underline{x}(k+1)|\underline{Y}(k+1)\}$ des Zustandsvektors $\underline{x}(k+1)$, bedingt auf die neue Messwerthistorie $\underline{Y}(k+1)$.

Die bedingten Verteilungsdichtefunktionen der Zufallsvariable $\underline{x}(k)$ sind Gaußförmig, wenn die beteiligten Verteilungsdichten der "Driving Noise" $w_d(k)$ und des Messrauschens $v(k)$ Gaußförmig sind [Lit. 8.2]. Zur Beschreibung sind dann lediglich die Mittelwerte nach (Gl. 8.28) und (Gl. 8.32) und die Kovarianzen aus (Gl. 8.29) und (Gl. 8.33) notwendig.

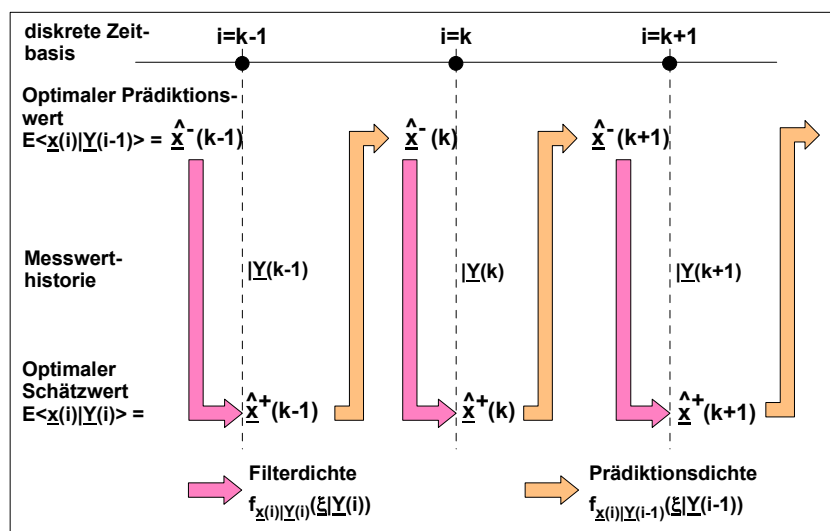


Abb. 8.12: Prädiktion und Estimation im KALMAN-Filter

Zur Zustandsprädiktion für den Zeitpunkt $k+1$ wird die Prädiktionsgleichung (Gl. 8.28) verwendet, die den optimalen Schätzwert $\hat{\underline{x}}^+(k)$ zum vorherigen Zeitpunkt k mit der lokalen Zustandsübergangsmatrix $A=\phi(T)$ gewichtet.

$$(Gl. 8.28) \quad \hat{\underline{x}}^-(k+1) = A \cdot \hat{\underline{x}}^+(k) \quad (\text{Prädiktionsgleichung})$$

Die bedingte Prädiktionsfehlerkovarianz ist identisch mit der Kovarianz der Zufallsvariablen $\underline{x}(k)$. Sie ist ein Maß für die Streuung der tatsächlichen Realisationen \underline{x}_k der Zufallsvariablen $\underline{x}(k)$ um den Prädiktionswert $\hat{\underline{x}}^-(k)$. An dieser Stelle wird auch die Varianz $Q(k)$ des Rauschprozesses am Systemeingang berücksichtigt.

$$(Gl. 8.29) \quad P^-(k+1) = A \cdot P^+(k) \cdot A^T + Q(k) \quad (\text{Prädiktionsfehlerkovarianz})$$

Mit dem Prädiktionswert und der Prädiktionsfehlerkovarianz kann die Berechnung der bedingten Verteilungsdichtefunktion $f_{\underline{x}(k+1)|Y(k)}(\underline{x}|Y(k))$ des Zustands $\underline{x}(k+1)$, bedingt auf die gesamte Messwerthistorie $Y(k)$ erfolgen [Lit. 8.2].

Die Residuensequenz $r(k+1)$ besteht aus der Differenz der neuen Messwertrealisation $y_{k+1} = y(k+1)$ und dem beobachtbaren Teil der Vorhersage. Sie enthält die in der neuen Messung vorhandene Korrekturinformation zum Prädiktionswert.

$$(Gl. 8.30) \quad r(k+1) = y(k+1) - C \cdot \hat{\underline{x}}^-(k+1) \quad (\text{Residuensequenz})$$

Die Kalman-Gain-Matrix bestimmt die Gewichtung der Korrekturinformation aus der Residuensequenz. Die Matrix berücksichtigt unter anderem das Verhältnis zwischen der Prädiktionsfehlerkovarianz und der Summe aus Prädiktionsfehlerkovarianz und Messfehlerkovarianz [Lit. 8.2].

$$(Gl. 8.31) \quad K(k+1) = P^-(k+1) \cdot C^T \cdot [C \cdot P^-(k+1) \cdot C^T + R(k+1)]^{-1} \quad (\text{Kalman-Gain})$$

Bedingt auf die spezielle Messwertgeschichte liefert das Kalman-Filter den optimalen Schätzwert, der aus einer linearen Kombination des Voraussageschätzwertes und der mit dem Kalman-Gain gewichteten Residuensequenz besteht [Lit. 8.2].

$$(Gl. 8.32) \quad \hat{\underline{x}}^+(k+1) = \hat{\underline{x}}^-(k+1) + K(k+1) \cdot r(k+1) \quad (\text{Estimationsgleichung})$$

Der Estimationswert und die Estimationsfehlerkovarianz nach (Gl. 8.33) gestatten die Berechnung der bedingten Verteilungsdichtefunktion $f_{\underline{x}(k+1)|Y(k+1)}(\underline{x}|Y(k+1))$ des Zustands $\underline{x}(k+1)$, bedingt auf die gesamte Messwerthistorie einschließlich des aktuellen Messwertes $y(k+1)$ [Lit. 8.2].

$$(Gl. 8.33) \quad P^+(k+1) = P^-(k+1) - K(k+1) \cdot C \cdot P^-(k+1) \quad (\text{Estimationsfehlerkovarianz})$$

8.8 Anwendungsmöglichkeiten zur Bewegungserkennung

In Kap. 8.5 wurde die Modellierung der Hintergrundstörungen als Rauschprozess mit der Kovarianzmatrix $Q(k)$ beschrieben. Für ein erstes einfaches Kalman-Filter ist es denkbar, $Q(k)$ anhand einer Reihe von Testmessungen offline zu bestimmen und für alle weiteren Auswertungen als Konstante anzusetzen. Über die Varianz des Rauschprozesses kann zusätzlich die Trägheit eines auf diesem Modell beruhenden Kalman-Filters beeinflusst werden.

Für das Nutzsignal $w_s(k)$ wurde demgegenüber kein Modell verwendet, um die Klasse der detektierbaren Objekte nicht von Anfang an durch eine Modellannahme für das Objektsignal einzuschränken. Die Bewegung eines Objektes im Erfassungsfeld des Sensors bewirkt nun das Auftreten eines Nutzsignals $s(k)$ am Sensorausgang, welches eine Abweichung zwischen dem Driftmodell und dem realen System bewirkt.

Die Verarbeitungsstrategie ist daher denkbar einfach. Verwendet man ein Kalman-Filter zur kontinuierlichen Schätzung des Hintergrundsignals aus den Sensordaten, werden sich Abweichungen des Driftmodells von den Sensordaten aufgrund des plötzlich vorhandenen Objektsignals $s(k)$ sofort in der Residuensequenz des Kalman-Filters auswirken. Die Residuensequenz verliert wesentliche charakteristische Eigenschaften, so dass ein einfacher Residuentest, wie er beispielsweise auch zur Plausibilitätskontrolle oder Ausreißerelimination von Messdaten verwendet wird, bereits eine erste Auswertestrategie darstellt [Lit. 8.3].

Alternativ kann die Kovarianz $Q(k)$ des Eingangsrauschprozesses durch eine Erweiterung des Systemmodells hinsichtlich der Anregung mit korreliertem Rauschen im Kalman-Filter online mitgeschätzt werden. Das gesamte Hintergrundmodell lässt sich dann beispielsweise aus einer "Leerlaufmessung" des Sensors ohne Objekt im Gesichtsfeld bestimmen. Solche Messungen können immer während der Abwesenheit von Objekten, also während der Hypothese H_0 , durchgeführt werden. Ein auf diesem Modell beruhendes Kalman-Filter ist in der Lage, sich selbständig an wechselnde Hintergrunddynamiken anzupassen [Lit. 8.3].

In Kap. 10 werden verschiedene Signalverarbeitungskonzepte vorgestellt, welche interne Zustände des Kalman-Filters auswerten. Angewendet werden insbesondere der Schwellwertvergleich der Residuensequenz sowie die Auswertung der um den Störsignalanteil bereinigten ersten und zweiten Zustandsvektorkomponente auf Basis des in Kap. 7 vorgestellten konventionellen Auswerteverfahrens. Bei der Auswertung der Residuensequenz wird eine konstante Varianz des Eingangsrauschprozesses angenommen, während die beiden letztgenannten eine dynamische Varianz Anpassung an die thermischen Verhältnisse im Erfassungsfeld in gewissen Grenzen zulassen.

Die Vielzahl der Auswertemöglichkeiten ist damit aber keineswegs erschöpft. Beispielsweise wird in [Lit. 8.3] vorgeschlagen, die Sprunghöhe und den Zeitpunkt von unerwarteten Zu-

standsänderungen infolge des Objektsignals $s(k)$ zu schätzen. Unter der Annahme Gaußförmiger Verteilungsdichten kann ein analytischer Zusammenhang zwischen der Entscheidungsschwellenhöhe zur Sprungidentifizierung und der Wahrscheinlichkeit des Auftretens eines Sprunges, bedingt auf die Messwerthistorie sowie die Sprunghöhe und den Sprungzeitpunkt, abgeleitet werden. Zur Unterscheidung zwischen den Hypothesen H_0 (keine Bewegung) und H_1 (bewegtes Objekt) dient der in Abb. 8.13 dargestellte Hypothesentester.

Das Verfahren basiert auf der Maximum-Likelihood-Estimation und erfordert zusätzlich zu dem Kalman-Filter zur Modellierung des Hintergrunds einen Hypothesentester und den in Abb. 8.13 dargestellten Sprungestimator.

Im Rahmen der vorliegenden Arbeit wird dieser vielversprechende Ansatz allerdings nicht weiterverfolgt.

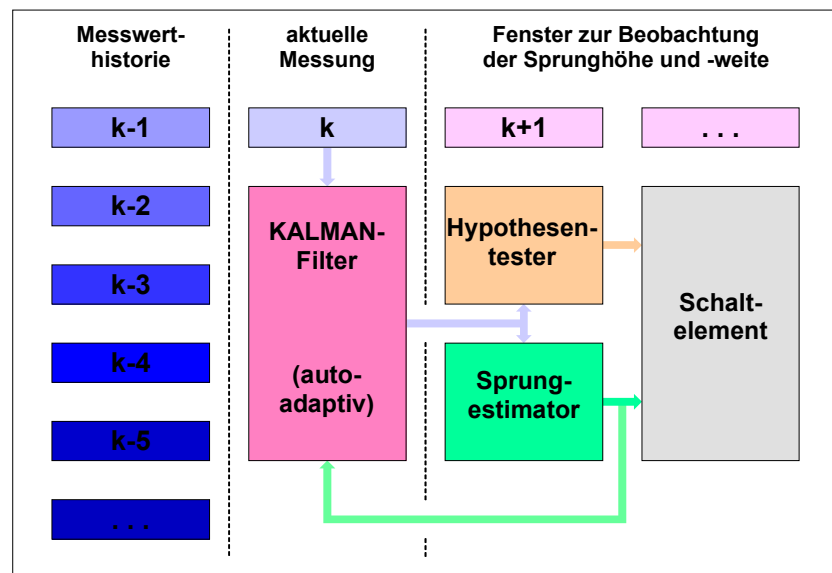


Abb. 8.13: Allgemeines System zur Bewegungserkennung mit Kalman-Filter, Hypothesentester, Sprungestimator und Schaltelement

8.9 Literatur

- [Lit. 8.1] Loffeld, Otmar: "Estimationstheorie I", R. Oldenbourg Verlag, München, Wien, 1990
- [Lit. 8.2] Loffeld, Otmar: "Estimationstheorie II", R. Oldenbourg Verlag, München, Wien, 1990
- [Lit. 8.3] Loffeld, Otmar: "Zur Signalverarbeitung von Bewegungsmeldern mit Infrarot-Sensoren", Interne Darstellung der Universität-Gesamthochschule Siegen, 1998
- [Lit. 8.4] T. Krieger, F. Holtz: "Signal processing of motion detectors with pyroelectric sensors", Proceedings IRS²2000, 6th Conference and Exhibition on Infrared Sensors & Systems, May 9-11, 2000, Erfurt

9 Entwicklung eines zeitdiskreten Systemmodells zur Signalvorverarbeitung

In diesem Kapitel wird die Vorgehensweise zur Ermittlung eines zeitdiskreten Zustandsraummodells beschrieben, das es gestattet, unerwartete Signaländerungen zu identifizieren und für die abschließende Bewegungserkennung bereitzustellen. Anhand der im weiteren Verlauf beschriebenen Verfahren bzw. Modelle soll insbesondere die Diskriminierung zwischen Nutzsignalen und typischen Hintergrundstörungen erfolgen. Dieser Vorgang muss als wesentlicher Teil der Signalvorverarbeitung betrachtet werden, während in Kap. 10 prinzipielle Möglichkeiten zur weiteren Signalanalyse und zur Entscheidungsfindung vorgestellt werden.

9.1 Modellierung des Bewegungsmelderverhaltens

Im Rahmen eines ersten Modells soll das in der Realität aufwendigere Hardwarekonzept von PIR-Bewegungsmeldern auf das eigentliche Sensorelement, den pyroelektrischen Detektor, und einen nachgeschalteten, idealen Wechselspannungsverstärker mit dem konstanten Verstärkungsfaktor V_{ac} reduziert werden. Komplexere Übertragungsfunktionen, die das dynamische Verhalten einer realen Verstärkerstufe berücksichtigen, sind ebenfalls denkbar. Allerdings erhöhen sie den Implementierungs- und Berechnungsaufwand für das Kalman-Filter erheblich, so dass an dieser Stelle auf ein vereinfachtes Bewegungsmeldermodell zurückgegriffen wird.

9.1.1 Übertragungsfunktion und Stoßantwort des vereinfachten Bewegungsmeldermodells

Ausgangspunkt für die nachfolgenden Betrachtungen ist die Übertragungsfunktion $H_s(s)$ eines pyroelektrischen Detektors nach (Gl. 4.62). Sie beschreibt allgemein das Verhältnis der Laplace-Transformierten des Ausgangssignal $V(s)$ des Bewegungsmelders bei einer zunächst determinierten Anregungsfunktion $\Delta\Psi(s) = U(s)$. Der vorne beschriebene Übergang auf die stochastische Anregung $w(t)$ soll erst im Anschluss an die Herleitung des zeitkontinuierlichen Zustandsraummodells erfolgen.

Unter Berücksichtigung des konstanten Verstärkungsfaktors V_{ac} folgt mit (Gl. 4.62) die idealisierte Übertragungsfunktion $H(s)$ des Bewegungsmelders:

$$(Gl. 9.1) \quad H(s) = \frac{Y(s)}{U(s)} = V_{ac} \cdot H_s(s) = \frac{V_{ac} \cdot V(s)}{\Delta\Psi(s)} = V_{ac} \cdot K_s \cdot \frac{s}{(1 + s\tau_{th}) \cdot (1 + s\tau_{el})}$$

Nach den Ausführungen in Kap. 5 entsprechen die Konstanten $\tau_{th} = 3.18$ s und $\tau_{el} = 0.53$ s der thermischen und der elektrische Zeitkonstante des pyroelektrischen Detektors, wobei für die Kehrwerte im folgenden die Abkürzungen $\alpha = 1/\tau_{th}$ und $\beta = 1/\tau_{el}$ verwendet werden. Für die Übertragungsfunktion $H(s)$ folgt dann mit der Abkürzung $K = V_{ac}K_s\alpha\beta$ die Beziehung

$$(Gl. 9.2) \quad H(s) = \frac{Y(s)}{U(s)} = V_{ac} K_s \alpha \beta \cdot \frac{s}{(s + \alpha) \cdot (s + \beta)} = K \cdot \frac{s}{(s + \alpha) \cdot (s + \beta)},$$

aus der die Stoßantwort $h(t)$ des Sensorsystems durch Rücktransformation in den Zeitbereich berechnet werden kann. Mit der inversen Laplace-Transformation $L^{-1}\{\dots\}$ folgt

$$(Gl. 9.3) \quad h(t) = K \cdot \frac{1}{\alpha - \beta} \cdot (\alpha \cdot e^{-\alpha t} - \beta \cdot e^{-\beta t})$$

Aus (Gl. 9.2) lässt sich weiterhin durch die Substitution $s = j\omega$ mit der Kreisfrequenz $\omega = 2\pi f$ die Übertragungsfunktion $H(f)$ nach (Gl. 9.4), das Phasenspektrum nach (Gl. 9.5) und das Betragsamplitudenspektrum $|H(f)|$ des vereinfachten Bewegungsmeldermodells nach (Gl. 9.6) ableiten.

$$H(f) = \frac{Y(f)}{U(f)} = K \cdot \frac{j\omega}{(\alpha + j\omega) \cdot (\beta + j\omega)} = K \cdot \frac{\omega^2(\alpha + \beta) - j\omega(\alpha\beta - \omega^2)}{\omega^4 + \omega^2(\alpha^2 + \beta^2) + \alpha^2\beta^2} \quad \text{folgt}$$

$$(Gl. 9.4) \quad H(f) = \text{Re}\{H(f)\} + j \cdot \text{Im}\{H(f)\} = \frac{K \cdot \omega^2(\alpha + \beta)}{(\omega^2 + \alpha^2)(\omega^2 + \beta^2)} + j \cdot \frac{K \cdot \omega(\omega^2 - \alpha\beta)}{(\omega^2 + \alpha^2)(\omega^2 + \beta^2)}$$

$$(Gl. 9.5) \quad \varphi(f) = -\arctan\left(\frac{\text{Im}\{H(f)\}}{\text{Re}\{H(f)\}}\right) = -\arctan\left(\frac{\omega^2 - \alpha\beta}{\omega(\alpha + \beta)}\right)$$

$$(Gl. 9.6) \quad |H(f)| = \sqrt{(\text{Re}\{H(f)\})^2 + (\text{Im}\{H(f)\})^2} = K \cdot \frac{\omega}{\sqrt{(\omega^2 + \alpha^2)(\omega^2 + \beta^2)}}$$

Aus dem Energiedichtespektrum $|H(f)|^2$ nach (Gl. 9.7), das dem Quadrat des Betragsamplitudenspektrums nach (Gl. 9.6) entspricht, lässt sich durch die inverse Fourier-Transformation $F^{-1}\{\dots\}$ außerdem die Autokorrelationsfunktion (AKF) $\varphi_{hh}(\tau)$ nach (Gl. 9.8) der Stoßantwort $h(t)$ aus (Gl. 9.3) bestimmen.

(Gl. 9.7)

$$|H(f)|^2 = [\text{Re}\{H(f)\}]^2 + [\text{Im}\{H(f)\}]^2 = K^2 \frac{\omega^4(\alpha + \beta)^2 + \omega^2(\omega^2 - \alpha\beta)^2}{(\omega^2 + \alpha^2)^2 \cdot (\omega^2 + \beta^2)^2} = \frac{K^2 \cdot \omega^2}{(\omega^2 + \alpha^2)(\omega^2 + \beta^2)}$$

(Gl. 9.8)

$$\varphi_{hh}(\tau) = 2 \frac{\alpha - \beta}{\alpha + \beta} \cdot \left(\frac{K}{2} \cdot \frac{1}{\alpha - \beta}\right)^2 (\alpha \cdot e^{-\alpha|\tau|} - \beta \cdot e^{-\beta|\tau|}) = \frac{K^2}{2 \cdot (\alpha^2 - \beta^2)} \cdot (\alpha \cdot e^{-\alpha|\tau|} - \beta \cdot e^{-\beta|\tau|})$$

Die AKF $\varphi_{hh}(\tau)$ kann nach (Gl. 9.9) zur Berechnung der AKF $\varphi_{yy}(\tau)$ am Ausgang eines linearen zeitinvarianten Systems (LTI-System) verwendet werden, wenn das System durch das Signal $u(t)$ mit der AKF $\varphi_{uu}(\tau)$ angeregt wird. “*” symbolisiert die Bildung des Faltungsprodukts.

$$(Gl. 9.9) \quad \varphi_{yy}(\tau) = \varphi_{uu}(\tau) * \varphi_{hh}(\tau) = \int_{-\infty}^{+\infty} \varphi_{uu}(v) \cdot \varphi_{hh}(\tau - v) dv \quad (\text{Wiener-Lee-Beziehung})$$

Da die Faltungsoperation im Zeitbereich einer Multiplikation im Frequenzbereich entspricht, erhält man bei Anwendung der Fourier-Transformation auf die Wiener-Lee-Beziehung einen

entsprechenden Zusammenhang im Frequenzbereich. Für Energiesignale $u(t)$ mit dem Energiedichtespektrum $|U(f)|^2$ entsteht bei der Übertragung über ein LTI-System mit der Stoßantwort $h(t)$ das Ausgangssignal $y(t)$, dessen Energiedichtespektrum $|Y(f)|^2$ nach (Gl. 9.10) berechnet werden kann.

$$(Gl. 9.10) \quad |Y(f)|^2 = |U(f)|^2 \cdot |H(f)|^2$$

Für Leistungssignale $u(t)$, die keine endliche Gesamtenergie aufweisen und die über eine mittlere Leistung in einem bestimmten Zeitintervall definiert werden, gilt ein ähnlicher Zusammenhang zwischen dem Leistungsdichtespektrum $\phi_{uu}(f)$ am Eingang und dem Leistungsdichtespektrum $\phi_{yy}(f)$ am Ausgang eines LTI-Systems mit der Stoßantwort $h(t)$ bzw. dem Energiedichtespektrum $|H(f)|^2$.

$$(Gl. 9.11) \quad \phi_{yy}(f) = \phi_{uu}(f) \cdot |H(f)|^2$$

Als Beispiel soll im folgenden das Leistungsdichtespektrum $\phi_{yy}(f)$ am Ausgang des Bewegungsmelders berechnet werden, wenn am Systemeingang ein weißer Rauschprozess $w(t)$ mit der Leistungsdichte N anliegt. Aus (Gl. 9.11) und (Gl. 9.7) folgt in diesem Fall

$$(Gl. 9.12) \quad \phi_{yy}(f) = \phi_{ww}(f) \cdot |H(f)|^2 = N \cdot |H(f)|^2 = N \cdot K^2 \frac{\omega^2}{(\omega^2 + \alpha^2)(\omega^2 + \beta^2)}$$

Die AKF $\varphi_{ww}(\tau)$ des weißen Rauschprozesses $w(t)$ am Systemeingang erhält man durch die inverse Fourier-Transformation des konstanten Rauschleistungsspektrums (s. (Gl. 9.13)). Setzt man (Gl. 9.13) in (Gl. 9.9) ein, folgt mit (Gl. 9.8) die AKF $\varphi_{yy}(\tau)$ des Signals $y(t)$ am Ausgang des Bewegungsmelders nach (Gl. 9.14).

$$(Gl. 9.13) \quad \varphi_{ww}(\tau) = N \cdot \delta(\tau)$$

$$(Gl. 9.14)$$

$$\varphi_{yy}(\tau) = N \cdot \delta(\tau) * \left(\frac{K^2}{2 \cdot (\alpha^2 - \beta^2)} \cdot (\alpha \cdot e^{-\alpha|\tau|} - \beta \cdot e^{-\beta|\tau|}) \right) = \frac{N \cdot K^2}{2 \cdot (\alpha^2 - \beta^2)} \cdot (\alpha \cdot e^{-\alpha|\tau|} - \beta \cdot e^{-\beta|\tau|})$$

Bei der Übertragung des Rauschprozesses $w(t)$ entsteht daher am Systemausgang ein farbiger Rauschprozess mit der Rauschleistungsdichte aus (Gl. 9.12) und der AKF nach (Gl. 9.14). Der Wert $\varphi_{yy}(0)$ entspricht der Leistung P_{yy} des Ausgangssignals $y(t)$. Allgemein gilt das Parsevalsche Theorem:

$$(Gl. 9.15) \quad P_{yy} = \varphi_{yy}(0) = \int_{-\infty}^{+\infty} y^2(t) dt = \int_{-\infty}^{+\infty} |Y(f)|^2 df \quad (\text{Parseval'sches Theorem})$$

Für $\tau = 0$ folgt daher aus (Gl. 9.14) die mittlere Rauschleistung am Systemausgang nach

$$(Gl. 9.16) \quad P_{yy} = \varphi_{yy}(0) = \frac{N \cdot K^2}{2 \cdot (\alpha^2 - \beta^2)} \cdot (\alpha - \beta) = \frac{N \cdot K^2}{2 \cdot (\alpha + \beta)},$$

wenn am Systemausgang ein weißer Rauschprozess mit der Leistungsdichte N anliegt. Bei bekannter Rauschleistung P_{yy} , die bei einem erwartungswertfreien Gaußverteilten Rauschprozess der Varianz $(\sigma_{yy})^2$ entspricht und daher eine reine Wechselleistung darstellt, gilt:

$$(Gl. 9.17) \quad P_{yy} = \sigma_{yy}^2 = \frac{N \cdot K^2}{2 \cdot (\alpha + \beta)}$$

Die Rauschleistung $P_{yy} = \sigma_{yy}^2$ wird später zur Parametrierung des Kalman-Filters benötigt.

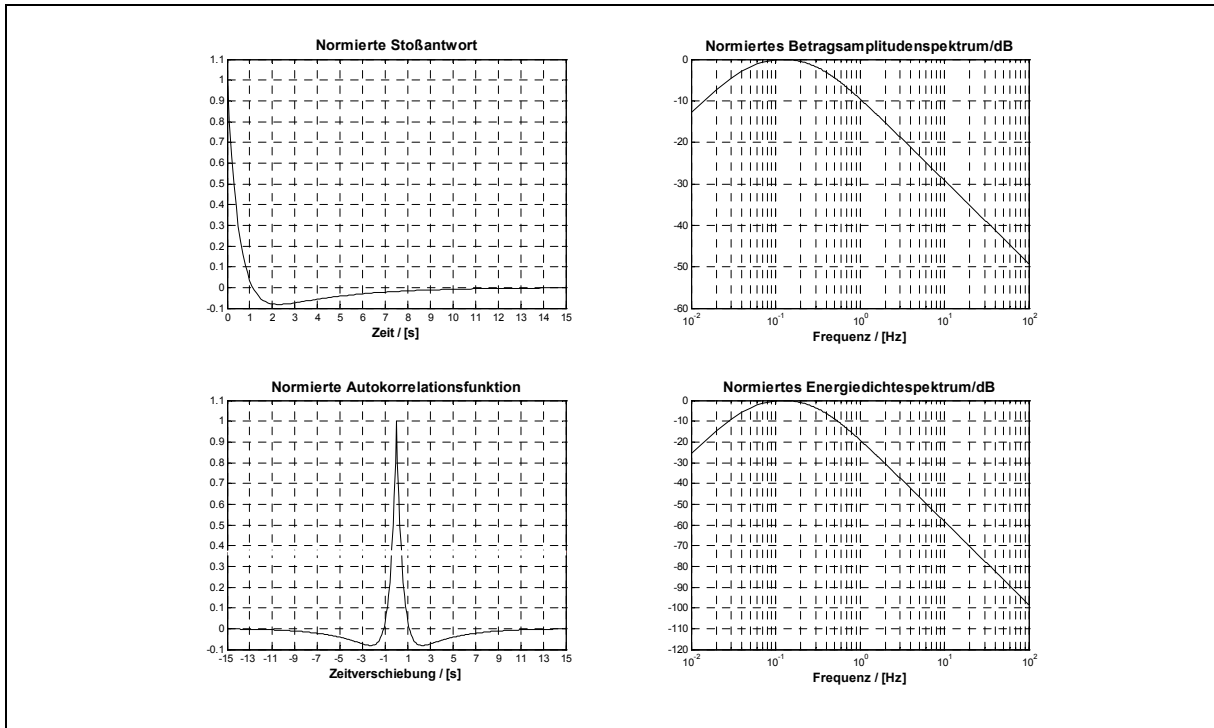


Abb. 9.1: Stoßantwort $h(t)$, Autokorrelationsfunktion $\phi_{hh}(\tau)$, Betragssamplitudenspektrum $|H(f)|$ und Energiedichtespektrum $|H(f)|^2$ des vereinfachten Bewegungsmeldermodells, normiert auf den jeweiligen Maximalwert

9.1.2 Herleitung des zeitkontinuierlichen Zustandsraummodells

Im Anschluss an die allgemeinen Betrachtungen aus Abschnitt 9.1.1 soll im folgenden zunächst ein zeitkontinuierliches Zustandsraummodell aus der Übertragungsfunktion nach (Gl. 9.2) abgeleitet werden. Die allgemeine Lösung dieses Differentialgleichungssystems wird dann im weiteren Verlauf zur Ableitung des äquivalenten zeitdiskreten Zustandsraummodells verwendet. Ausgehend von der Übertragungsfunktion (Gl. 9.2) wird zunächst die Hilfsgröße $X(s)$ nach (Gl. 9.18) definiert.

$$(Gl. 9.18) \quad Y(s) = K \cdot \frac{s}{(s + \alpha) \cdot (s + \beta)} \cdot U(s) = K \cdot s \cdot X(s) \quad \text{mit} \quad X(s) = \frac{s}{(s + \alpha) \cdot (s + \beta)} \cdot U(s)$$

Durch Ausmultiplizieren des Nenners erhält man (Gl. 9.19), die anschließend in den Zeitbereich zurücktransformiert wird. Setzt man alle Anfangsbedingungen gleich Null, erhält man mit der Zwischengröße $x(t)$ die Beziehung (Gl. 9.20).

$$(Gl. 9.19) \quad U(s) = s^2 X(s) + sX(s) \cdot (\alpha + \beta) + X(s) \cdot \alpha\beta$$

$$(Gl. 9.20) \quad u(t) = \ddot{x}(t) + \dot{x}(t) \cdot (\alpha + \beta) + x(t) \cdot \alpha\beta$$

Wählt man nun einen Zustandsvektor $\underline{x}(t) = [x_1(t) \quad x_2(t)]^T$ mit $x_1(t) = x(t)$ und $x_2(t) = \dot{x}(t)$, folgt mit

$$\begin{aligned} \dot{x}_1(t) &= \dot{x}(t) = 0 \cdot x_1(t) + 1 \cdot x_2(t) + 0 \cdot u(t) \\ \dot{x}_2(t) &= \ddot{x}(t) = -\alpha\beta \cdot x_1(t) - (\alpha + \beta) \cdot x_2(t) + 1 \cdot u(t) \end{aligned}$$

die zeitkontinuierliche Zustandsraumgleichung nach (Gl. 9.21) mit den Matrizen A_c und B_c nach (Gl. 9.22) und (Gl. 9.23).

$$(Gl. 9.21) \quad \dot{\underline{x}}(t) = A_c \cdot \underline{x}(t) + B_c \cdot u(t) \quad \text{mit} \quad \underline{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad \text{und} \quad \dot{\underline{x}}(t) = \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix}$$

$$(Gl. 9.22) \quad A_c = \begin{bmatrix} 0 & 1 \\ -\alpha\beta & -(\alpha + \beta) \end{bmatrix} \quad (Gl. 9.23) \quad B_c = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Die Matrix A_c wird als Systemmatrix bezeichnet, während B_c die deterministische Steuermatrix darstellt. Das Beobachtungsmodell lässt sich ebenfalls aus (Gl. 9.18) ableiten. Mit

$$y(t) = K \cdot \dot{x}(t) = 0 \cdot x_1(t) + K \cdot x_2(t)$$

erhält man die Ausgangsgleichung (Gl. 8.24) mit der Beobachtungsmatrix C_c nach (Gl. 8.25).

$$(Gl. 9.24) \quad y(t) = C_c \cdot \underline{x}(t) \quad (Gl. 9.25) \quad C_c = K \cdot [0 \quad 1]$$

Mit (Gl. 9.21) und (Gl. 9.24) sowie den Matrizen gemäß (Gl. 9.22), (Gl. 9.23) und (Gl. 9.25) ist das zeitkontinuierliche Zustandsraummodell vollständig beschrieben. Unter Anwendung von (Gl. 8.4) kann mit der Systemmatrix A_c die globale Zustandsübergangsmatrix $\Phi(t-t_0)$ des vereinfachten Bewegungsmeldermodells und damit die globale Zustandsübergangsfunktion nach (Gl. 8.3) ermittelt werden. Letztere ist die allgemeine Lösung des Differentialgleichungssystems (Gl. 9.21) und ermöglicht die Berechnung des Systemzustands $\underline{x}(t)$ für einen beliebigen Zeitpunkt t , wenn der Anfangszustand $\underline{x}(t_0)$ und die Wirkung der deterministischen Steuergröße $u(t)$ im Zeitintervall $[t_0, t)$ bekannt sind. Unter Anwendung von (Gl. 8.4) erhält man aus

$$\Phi(s) = [s \cdot I - A_c]^{-1} = \begin{bmatrix} s & -1 \\ \alpha\beta & s + (\alpha + \beta) \end{bmatrix}^{-1} = \frac{1}{(s + \alpha)(s + \beta)} \cdot \begin{bmatrix} s + (\alpha + \beta) & 1 \\ -\alpha\beta & s \end{bmatrix}$$

durch inverse Laplace-Transformation $L^{-1}\{\dots\}$ die gesuchte Matrix

(Gl. 9.26)

$$\Phi(t - t_0) = L^{-1}\{\Phi(s)\} = \frac{1}{\alpha - \beta} \cdot \begin{bmatrix} -\beta \cdot e^{-\alpha(t-t_0)} + \alpha \cdot e^{-\beta(t-t_0)} & -e^{-\alpha(t-t_0)} + e^{-\beta(t-t_0)} \\ \alpha\beta \cdot (e^{-\alpha(t-t_0)} - e^{-\beta(t-t_0)}) & \alpha \cdot e^{-\alpha(t-t_0)} - \beta \cdot e^{-\beta(t-t_0)} \end{bmatrix}$$

9.1.3 Ableitung des äquivalenten zeitdiskreten Systemmodells

Unter Beachtung der Ausführungen in Kap. 8 soll in diesem Abschnitt das zu (Gl. 9.21) - (Gl. 9.25) äquivalente zeitdiskrete Systemmodell ermittelt werden. Für die zeitdiskrete Systemmatrix A_1 der Zustandsdifferenzgleichung

$$(Gl. 9.27) \quad \underline{x}(k+1) = A_1 \cdot \underline{x}(k) + B_1 \cdot u(k) \quad \text{mit} \quad u(k) = \text{const. im Zeitintervall } [(k+1)T, kT)$$

folgt aus (Gl. 7.9) mit dem Abtastintervall $t-t_0 = T$ bzw. die Beziehung:

$$(Gl. 9.28) \quad A_1 = \Phi(T) = \frac{1}{\alpha - \beta} \cdot \begin{bmatrix} -\beta \cdot e^{-\alpha T} + \alpha \cdot e^{-\beta T} & -e^{-\alpha T} + e^{-\beta T} \\ \alpha\beta \cdot (e^{-\alpha T} - e^{-\beta T}) & \alpha \cdot e^{-\alpha T} - \beta \cdot e^{-\beta T} \end{bmatrix}$$

Aus (Gl. 7.10) erhält man die zeitdiskrete, deterministische Steuermatrix B_1 nach (Gl. 8.29), die an dieser Stelle nur der Vollständigkeit halber angegeben ist. Die zeitdiskrete Beobachtungsmatrix C_1 wird nach (Gl. 7.11) unverändert vom zeitkontinuierlichen Zustandsraummodell übernommen (Gl. 8.30).

$$(Gl. 9.29) \quad B_1 = \frac{1}{\alpha - \beta} \cdot \begin{bmatrix} -\frac{1}{\alpha}(1 - e^{-\alpha T}) + \frac{1}{\beta}(1 - e^{-\beta T}) \\ -e^{-\alpha T} + e^{-\beta T} \end{bmatrix} \quad (Gl. 9.30) \quad C_1 = C_c = K[0 \quad 1]$$

9.1.4 Berücksichtigung von Rauschprozessen

In Kap. 8 wurde bereits die Berücksichtigung eines weißen, erwartungswertfreien Rauschprozesses $w(t)$ am Systemeingang und eines ebenfalls weißen, erwartungswertfreien Rauschprozesses $v(t)$ am Systemausgang eingeführt. Im zeitkontinuierlichen Fall werden beide Rauschprozesse, die außerdem Gaußverteilt sein sollen und daher durch die Angabe ihres linearen Erwartungswerts m_w und m_v sowie der zugehörigen Varianzen P_{ww} und P_{vv} vollständig charakterisiert sind, folgendermaßen beschrieben:

$$(Gl. 9.31) \quad m_w = E\{w(t)\} = 0 \quad (Gl. 9.32) \quad P_{ww} = E\{[w(t) - E\{w(t)\}]^2\} = \sigma_{ww}^2 \cdot \delta(\tau)$$

$$(Gl. 9.33) \quad m_v = E\{v(t)\} = 0 \quad (Gl. 9.34) \quad P_{vv} = E\{[v(t) - E\{v(t)\}]^2\} = \sigma_{vv}^2 \cdot \delta(\tau)$$

Die Varianzen entsprechen der Wechselleistung der erwartungswertfreien Rauschprozesse und erzeugen ein über alle Frequenzen konstantes Leistungsdichtespektrum. Für die AKF φ_{ww} der "Driving Noise" und φ_{vv} des Ausgangsrauschprozesses gilt dann

$$(Gl. 9.35) \quad \varphi_{ww}(\tau) = N_w \cdot \delta(\tau) \quad (Gl. 9.36) \quad \varphi_{vv}(\tau) = N_v \cdot \delta(\tau)$$

Im weiteren Verlauf wird ein zu dem weißen Rauschprozess $w(t)$ äquivalenter zeitdiskreter Rauschprozess $\underline{w}(k)$ berechnet. Da der weiße Rauschprozess $w(t)$ anstelle des deterministischen Signals $u(t)$ in (Gl. 9.31) eingekoppelt wird, sind die zeitkontinuierliche stochastische Kontrollmatrix G_c und die deterministische Kontrollmatrix B_c identisch, was bereits in (Gl. 8.17) angedeutet wurde. Zur Berechnung der Kovarianzmatrix Q_1 des zeitdiskreten weißen Rauschprozesses $\underline{w}(k)$ am Systemeingang wird (Gl. 8.16) mit $G_c = B_c$ verwendet.

Außerdem wird angenommen, dass sich die Varianz $N_w(k) = N_w$ des zeitkontinuierlichen Rauschprozesses $w(t)$ über der Zeit nicht ändert, so dass N_w bei der Auswertung von (Gl. 8.16) vor das Integral gezogen werden kann:

$$Q_1 = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} = \int_{kT}^{(k+1)T} \Phi((k+1)T - \tau) \cdot B_c \cdot N_w \cdot B_c^T \cdot \Phi((k+1)T - \tau)^T d\tau$$

$$Q_1 = N_w \cdot \int_{kT}^{(k+1)T} \Phi((k+1)T - \tau) \cdot \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \cdot \Phi((k+1)T - \tau)^T d\tau$$

(Gl. 9.37)

$$Q_1 = \frac{N_w}{(\alpha - \beta)^2} \begin{bmatrix} \frac{1 - e^{-2\alpha T}}{2\alpha} + \frac{1 - e^{-2\beta T}}{2\beta} + \frac{2(1 - e^{-(\alpha+\beta)T})}{\alpha + \beta} & \frac{1}{2}(e^{-\alpha T} - e^{-\beta T})^2 \\ \frac{1}{2}(e^{-\alpha T} - e^{-\beta T})^2 & \frac{\alpha(1 - e^{-2\alpha T})}{2} + \frac{\beta(1 - e^{-2\beta T})}{2} + \frac{2\alpha\beta(1 - e^{-(\alpha+\beta)T})}{\alpha + \beta} \end{bmatrix}$$

Neben dem zeitdiskreten Gaußverteilten Eingangsruschprozess $\underline{w}(k)$, der durch den linearen Erwartungswert nach (Gl. 9.38) und die Kovarianzmatrix Q_1 nach (Gl. 9.37) bzw. (Gl. 9.39) vollständig beschrieben ist, wird auch für den zeitdiskreten Ausgangsruschprozess $v(k)$ eine Gaußverteilung angenommen.

Da das Ausgangssignal $y(k)$ eine skalare Größe ist, muss der Rauschprozess $v(k)$ ebenfalls eine skalare Größe sein. Das Ausgangsruschen soll außerdem nach (Gl. 9.40) erwartungswertfrei sein und die skalare Varianz R_1 nach (Gl. 9.41) aufweisen,

$$(Gl. 9.38) \quad E\{\underline{w}(k)\} = \underline{0} \qquad (Gl. 9.39) \quad E\{\underline{w}(k) \cdot \underline{w}(k)^T\} = Q_1 \cdot \delta(k)$$

$$(Gl. 9.40) \quad E\{v(k)\} = 0 \qquad (Gl. 9.41) \quad E\{v(k) \cdot v(k)\} = R_1 \cdot \delta(k)$$

9.1.5 Filtergleichungen des Kalman-Filters

In Kap. 8 wurden bereits die Filtergleichungen des Kalman-Filters angegeben und die prinzipielle Funktionalität beschrieben. Mit den zeitinvarianten Kovarianzen nach (Gl. 9.39) und (Gl. 9.41) sollen die Beziehungen nach (Gl. 8.28) - (Gl. 8.33) erneut formuliert und zusammenfassend dargestellt werden:

$$(Gl. 9.42) \quad \hat{\underline{x}}^-(k+1) = A_1 \cdot \hat{\underline{x}}^+(k) \qquad \text{(Prädiktionsgleichung)}$$

$$(Gl. 9.43) \quad P^-(k+1) = A_1 \cdot P^+(k) \cdot A_1^T + Q_1 \qquad \text{(Prädiktionsfehlerkovarianz)}$$

$$(Gl. 9.44) \quad r(k+1) = y(k+1) - C_1 \cdot \hat{\underline{x}}^-(k+1) \qquad \text{(Residuensequenz)}$$

$$(Gl. 9.45) \quad K(k+1) = P^-(k+1) \cdot C_1^T \cdot [C_1 \cdot P^-(k+1) \cdot C_1^T + R_1]^{-1} \qquad \text{(Kalman-Gain-Matrix)}$$

$$(Gl. 9.46) \quad \hat{\underline{x}}^+(k+1) = \hat{\underline{x}}^-(k+1) + K(k+1) \cdot r(k+1) \qquad \text{(Estimationsgleichung)}$$

$$(Gl. 9.47) \quad P^+(k+1) = P^-(k+1) - K(k+1) \cdot C_1 \cdot P^-(k+1) \qquad \text{(Estimationsfehlerkovarianz)}$$

Die (Gl. 9.42) - (Gl. 9.47) zeigen, dass das Kalman-Filter die Systemmatrix A_1 , die Beobachtungsmatrix C_1 sowie die Kovarianzen R_1 und Q_1 zur Zustandsprädiktion und -estimation verwendet. Um Rechenzeit einzusparen, soll an dieser Stelle noch eine Näherung für die Matrizen A_1 und Q_1 angegeben werden, die insbesondere bei kleinen Abtastintervallen T zutrifft. Mit dieser Annahme lassen sich die e-Funktionen in (Gl. 9.28) und (Gl. 9.37) in eine Reihe entwickeln. Bricht man die Reihenentwicklung nach dem linearen Term ab, folgt

$$(Gl. 9.48) \quad A_1 \approx A_1^* = \begin{bmatrix} 1 & T \\ -\alpha\beta T & 1 - (\alpha + \beta)T \end{bmatrix} \quad (Gl. 9.49) \quad Q_1 \approx Q_1^* = \begin{bmatrix} 0 & 0 \\ 0 & N_w T \end{bmatrix}$$

Die Varianz $R_1 = \sigma_{vv}^2$ des Ausgangsrauschprozesses kann messtechnisch bei konstanten Temperaturverhältnissen im Erfassungsfeld des Bewegungsmelders ermittelt werden. Die Auswertung kann offline erfolgen.

Zur Bestimmung der Konstanten N_w in (Gl. 9.37) misst man die Varianz $(\sigma_{dd})^2$ am Bewegungsmelderausgang in Verbindung mit typischen Drift- bzw. Störsignalen. Unter Beachtung der statistischen Unabhängigkeit des Eingangs- und des Ausgangsrauschprozesses und unter Anwendung von (Gl. 9.17) ergibt sich aus (Gl. 9.50) die Beziehung (Gl. 9.51) für die Rauschleistungsdichte N_w . Beim Übergang von (Gl. 9.50) auf (Gl. 9.51) wurde weiterhin angenommen, dass die am Systemausgang gemessene Varianz $(\sigma_{dd})^2$ des Eingangsruschprozesses sehr viel größer ist als die Varianz $(\sigma_{vv})^2$ des Ausgangsrauschprozesses.

$$(Gl. 9.50) \quad \sigma_{dd}^2 = \frac{N_w \cdot K^2}{2 \cdot (\alpha + \beta)} + \sigma_{vv}^2 \quad (Gl. 9.51) \quad N_w = \sigma_{dd}^2 \cdot \frac{2 \cdot (\alpha + \beta)}{K^2}$$

9.1.6 Simulationsergebnisse bei Verwendung des vereinfachten Bewegungsmeldermodells

In diesem Abschnitt soll die Auswertung von typischen Bewegungsmeldersignalen mit einem ersten Kalman-Filter beschrieben werden, welches intern das zeitdiskrete Modell nach Kap. 9.1.3 und 9.1.4 verwendet. Die Ausführungen in den vorangegangenen Kapiteln haben gezeigt, dass die Bewegungsmeldersignale, die im folgenden als Testsignale verwendet werden, neben einer hohen AC-Verstärkung, die durch den Faktor K in (Gl. 9.6) bzw. (Gl. 9.30) berücksichtigt wird, mit einem zusätzlichen DC-Offset oberhalb von 0V beaufschlagt sind.

Bereits an dieser Stelle ist zu vermuten, dass das Kalman-Filter, dessen internes Systemmodell letztendlich auf einem Bandpass beruht, das ursprüngliche Signal infolge des fehlenden DC-Pfades nicht nachführen kann. Abb. 9.2 verdeutlicht diesen Effekt, nachdem bereits alle Einschwingvorgänge abgeklungen sind.

Das auf dem Bandpassmodell beruhende Kalman-Filter 1 zeigt jeweils einen deutlichen Offset zwischen dem Originalsignal sowie dem vorhergesagten und dem geschätzten Ausgangssignal (s. Abb. 9.2 oben). Die letztgenannten Ausgangssignale entstehen durch Multiplikation

der Zustandsvektoren \hat{x}^- und \hat{x}^+ mit der Beobachtungsmatrix $C_1 = [0 \ 1]$, was einer Abbildung der Zustandsvektoren in den Beobachtungsraum entspricht.

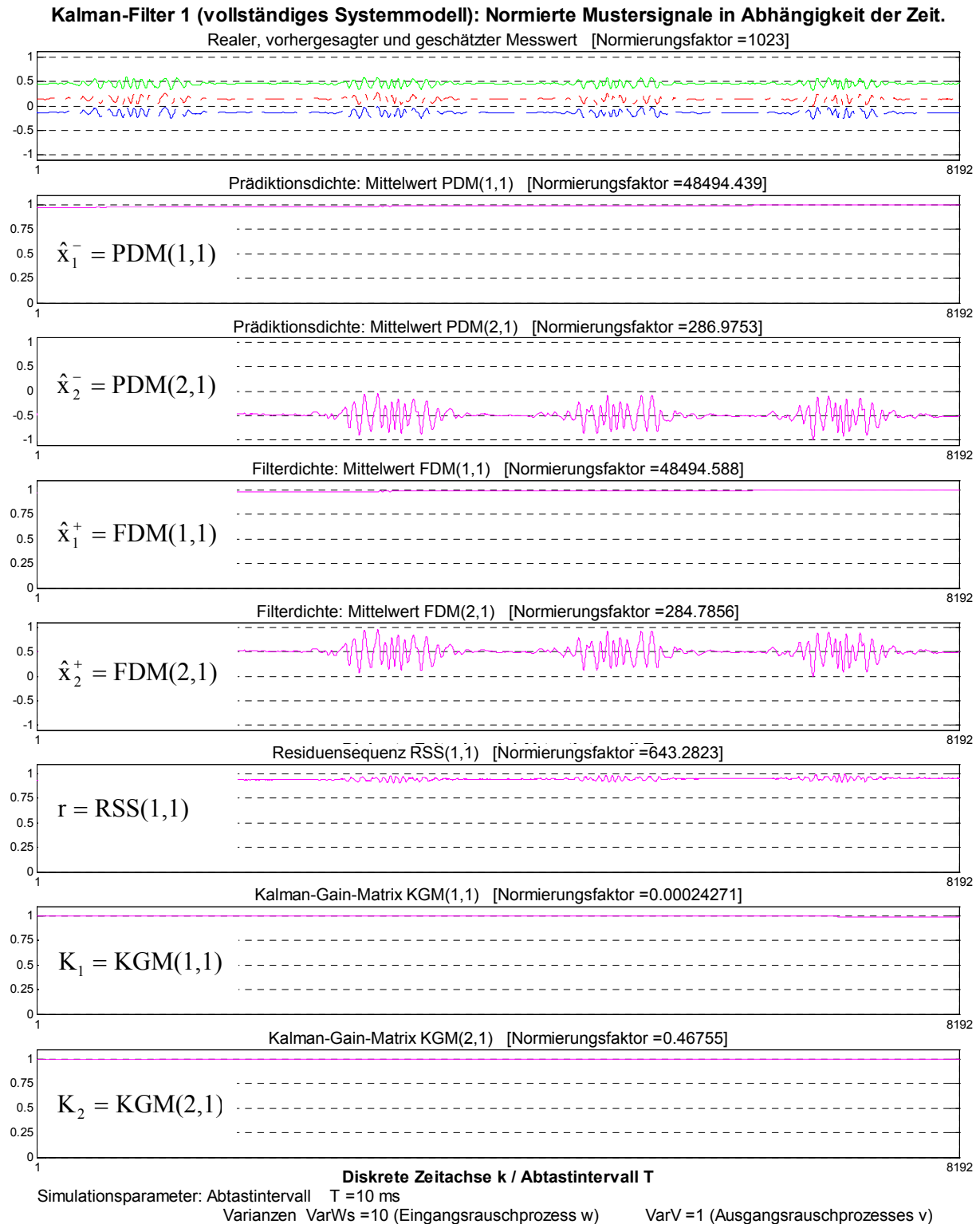


Abb. 9.2: Kalman-Filter 1. Normierte Ausgangssignale und Mittelwerte PDM und FDM der Prädiktions- und Filterdichte sowie Residuensequenz RSS und Kalman-Gain-Matrix KGM in Abhängigkeit der Zeit. Im oberen Diagramm sind das geschätzte Ausgangssignal (rot), das vorhergesagte Ausgangssignal (blau) und das gemessene Ausgangssignal (grün) und dargestellt.

Zusätzlich dargestellt sind die zeitlichen Verläufe der einzelnen Zustandsvektorkomponenten. Es fällt auf, dass der komplette Wechselanteil des ursprünglichen Signals jeweils vollständig in der 2. Komponente \hat{x}_2^- und \hat{x}_2^+ der Zustandsvektoren $\underline{\hat{x}}^-$ und $\underline{\hat{x}}^+$ enthalten ist.

Da die erste Komponente der Matrix C_1 Null ist, werden die Komponenten \hat{x}_1^- und \hat{x}_1^+ nicht bei der Bildung der Residuensequenz berücksichtigt (s. Abb. 9.2). Dort ist der fehlende DC-Offset deutlich erkennbar.

9.1.7 Modellerweiterung des einfachen Bewegungsmeldermodells

Um die gerade beschriebenen Probleme zu umgehen, bietet es sich zunächst an, die Matrix $C_2 = [1 \ 1]$ anstelle der ursprünglichen Matrix $C_1 = [0 \ 1]$ zu verwenden. Abb. 9.3 zeigt die zu Abb. 9.2 korrespondierenden Signalverläufe über der Zeit. Man erkennt, dass die in den Beobachtungsraum transformierten Zustandsgrößen in diesem Fall eine deutlich bessere Nachführung des Ausgangssignals ermöglichen. Allerdings ändert sich bei näherer Betrachtung das Frequenzverhalten des verwendeten Modells grundlegend. Mit

(Gl. 9.52)

$$H(s) = C_2 \cdot [s \cdot I - A_c]^{-1} \cdot B_c = [1 \ 1] \cdot \begin{bmatrix} s & -1 \\ \alpha\beta & s + (\alpha + \beta) \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{s + 1}{(s + \alpha) \cdot (s + \beta)},$$

wobei infolge der sonst unveränderten Systemeigenschaften die Matrizen A_c und B_c aus (Gl. 9.22) und (Gl. 9.23) eingesetzt wurden, ergibt sich mit der Substitution $s = j\omega$ und anschließender Betragsbildung das Amplitudenspektrum nach (Gl. 9.53) im Vergleich zur Bandpasscharakteristik nach (Gl. 9.6).

Für das Energiedichtespektrum erhält man (Gl. 9.54). Die aus dem zeitkontinuierlichen Systemmodell bekannte Konstante K wurde in (Gl. 9.53) und (Gl. 9.54) zusätzlich berücksichtigt.

$$(Gl. 9.53) \quad |H(f)| = \sqrt{(\operatorname{Re}\{H(f)\})^2 + (\operatorname{Im}\{H(f)\})^2} = K \cdot \sqrt{\frac{\omega^2 + 1}{(\omega^2 + \alpha^2)(\omega^2 + \beta^2)}}$$

$$(Gl. 9.54) \quad |H(f)|^2 = [\operatorname{Re}\{H(f)\}]^2 + [\operatorname{Im}\{H(f)\}]^2 = \frac{K^2 \cdot (\omega^2 + 1)}{(\omega^2 + \alpha^2)(\omega^2 + \beta^2)}$$

Nach Abb. 9.4 entsteht durch die Änderung der Beobachtungsmatrix C_1 in C_2 das Übertragungsverhalten eines Tiefpasses. Aufgrund der unterschiedlichen Stoßantworten (s. Abb. 9.4) beider Modelle sind ebenfalls veränderte Korrelationseigenschaften zu erwarten.

Das Tiefpassmodell (Kalman-Filter 2) wird im folgenden nicht weiter untersucht, da das im nächsten Abschnitt vorgestellte Kalman-Filter 3, das auf einem sehr allgemeinen Ansatz zur System bzw. Signalmodellierung beruht, eine bessere Anpassung an die zugrundeliegende Problemstellung zeigt.

Kalman-Filter 2 (vollständiges Systemmodell): Normierte Mustersignale in Abhängigkeit der Zeit.

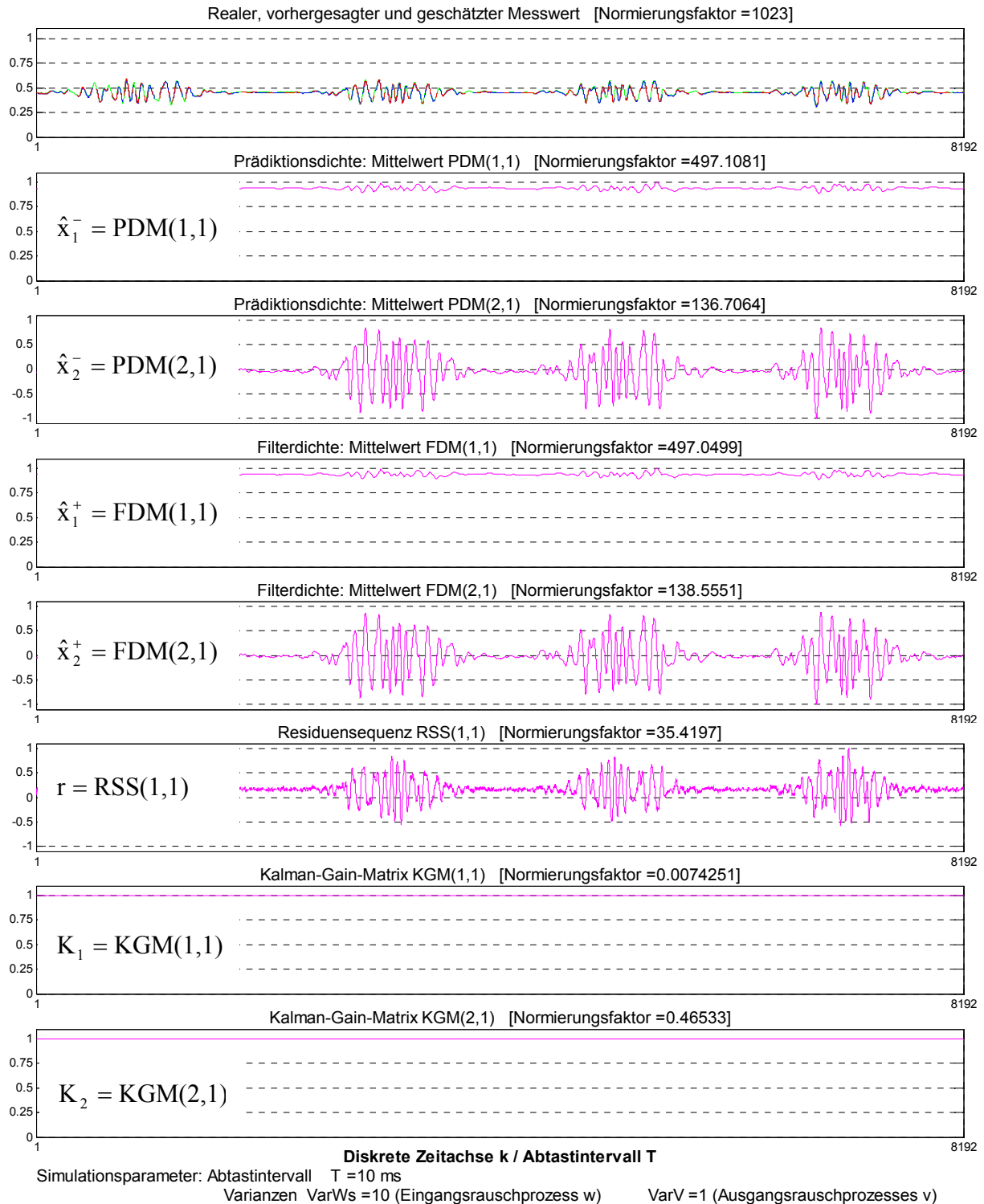


Abb. 9.3: Kalman-Filter 2. Normierte Ausgangssignale und Mittelwerte PDM und FDM der Prädiktions- und Filterdichte sowie Residuensequenz RSS und Kalman-Gain-Matrix KGM in Abhängigkeit der Zeit. Im oberen Diagramm sind das geschätzte Ausgangssignal (rot), das vorhergesagte Ausgangssignal (blau) und das gemessene Ausgangssignal (grün) und dargestellt.

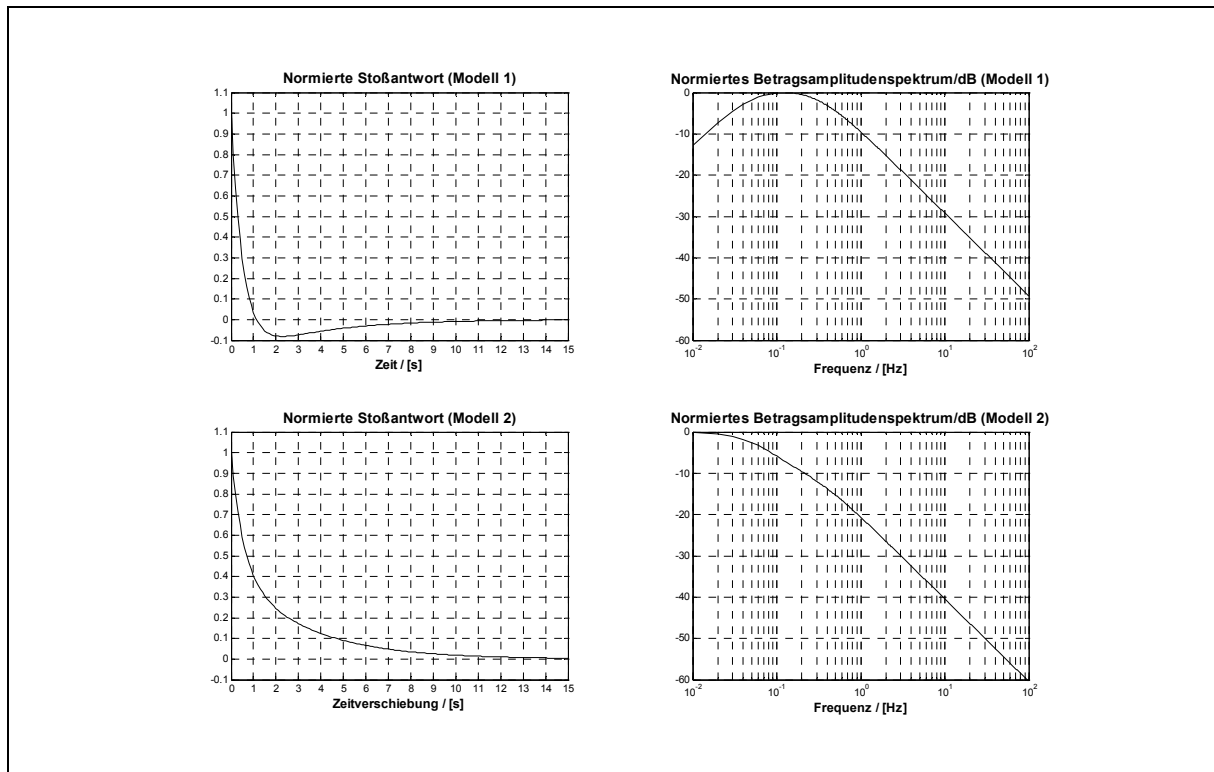


Abb. 9.4: Normierte Stoßantworten und Betragsamplitudenspektren des ursprünglichen Bandpassmodells (Modell 1) aus der PIR-Übertragungsfunktion ($C_1 = [0 \ 1]$) und des erweiterten Modells (Modell 2) mit $C_2 = [1 \ 1]$

9.2 Allgemeines Modell zur Signalverfolgung

9.2.1 Herleitung der zeitkontinuierlichen Zustandsraummodells

Ausgehend von den vorne beschriebenen Ergebnissen soll im weiteren Verlauf ein möglichst einfaches, allgemeines Signalmodell abgeleitet werden, welches eine DC-gekoppelte Verfolgung des gemessenen Ausgangssignals ermöglicht. In einem weiteren Schritt (s. Kap. 9.3) erfolgt die Berücksichtigung der Korrelationseigenschaften des Bewegungsmeldermodells nach Kap. 9.1.1 und 9.1.4. Die zeitkontinuierlichen Rauschprozesse $w(t)$ am Systemeingang und $v(t)$ am Systemausgang werden im Rahmen des zu entwickelnden Modells beibehalten.

In Abwesenheit von typischen Nutz- und Störsignalen, wobei letztere nach wie vor durch die vorne bereits beschriebenen Rauschprozesse modelliert werden, soll die in den Beobachtungsraum abgebildete Zustandsvektorkomponente $x_1(t)$ des Zustandsvektors $x(t) = [x_1(t) \ x_2(t)]^T$ dem Ausgangssignal $y(t)$ des Bewegungsmelders folgen.

Die erste Ableitung $\dot{x}_1(t) = dx_1(t)/dt$ beschreibt in diesem Fall die Änderung des Ausgangssignals $y(t) = x_1(t)$ im Idealfall ohne zusätzliche Rauschprozesse. Die erste Ableitung $\dot{x}_2(t) = dx_2(t)/dt$ der zweiten Zustandsvektorkomponente wird im folgenden als eine negativ autokorrelierte Größe mit dem Korrelationsparameter γ betrachtet. Der Eingangsrauschpro-

zess $w(t)$ soll sich insbesondere auf $\dot{x}_2(t)$ auswirken, so dass sich das zeitkontinuierliche Zustandsraummodell nach (Gl. 9.55), (Gl. 9.56) und (Gl. 9.57) aus dem Gleichungssystem

$$\begin{aligned}\dot{x}_1(t) &= \dot{x}(t) = 0 \cdot x_1(t) + 1 \cdot x_2(t) + 0 \cdot w(t) \\ \dot{x}_2(t) &= \ddot{x}(t) = 0 \cdot x_1(t) - \gamma \cdot x_2(t) + 1 \cdot w(t)\end{aligned}$$

ableiten lässt. Für das Beobachtungsmodell folgt unter den oben beschriebenen Voraussetzungen (Gl. 9.58) mit der Beobachtungsmatrix (Gl. 9.59). Die Konstante K in (Gl. 9.59) wird wie im vorherigen Kapitel gleich Eins gesetzt.

$$(Gl. 9.55) \quad \dot{\underline{x}}(t) = A_c \cdot \underline{x}(t) + G_c \cdot w(t) \quad \text{mit} \quad \underline{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad \text{und} \quad \dot{\underline{x}}(t) = \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix}$$

$$(Gl. 9.56) \quad A_c = \begin{bmatrix} 0 & 1 \\ 0 & -\gamma \end{bmatrix}$$

$$(Gl. 9.57) \quad G_c = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$(Gl. 9.58) \quad y(t) = C_c \cdot \underline{x}(t) + v(t)$$

$$(Gl. 9.59) \quad C_c = K \cdot [1 \quad 0]$$

Die allgemeine Lösung der Differentialgleichung für die Zustandsvektorkomponente $x_2(t)$ lautet unter Beachtung des Korrelationsparameters γ und des Startzeitpunkts t_0 :

$$(Gl. 9.60) \quad x_2(t) = e^{-\gamma(t-t_0)} \cdot x_2(t_0) + \int_{t_0}^t e^{-\gamma(t-\tau)} \cdot w(\tau) d\tau$$

(Gl. 9.60) besagt, dass die aktuelle Änderung $x_2(t)$ des Zustandsvektors ausgehend von einem vorherigen Wert, beispielsweise zum Startzeitpunkt t_0 , und aus der Wirkung der stochastischen Anregung im Zeitintervall $[t_0, t)$ bestimmt werden. Letztere ergibt sich aus dem mit der Zustandsübergangsmatrix gewichteten und über die Zeitdifferenz integrierten weißen Rauschen $w(t)$. Betrachtet man den linearen Erwartungswert $E\{x_2(t)\}$ für einen erwartungswertfreien Rauschprozess am Systemeingang, folgt mit $E\{w(t)\} = 0$ und unter Beachtung der statistischen Unabhängigkeit von $w(t)$ und $x_2(t_0)$

$$(Gl. 9.61) \quad E\{x_2(t)\} = e^{-\gamma(t-t_0)} \cdot E\{x_2(t_0)\}.$$

Im folgenden wird angenommen, dass der Erwartungswert $E\{x_2(t)\} = 0$ ist. Das bedeutet, dass im Ausgangssignal des Bewegungsmelders keine bevorzugte Richtung der Signaländerung auftritt. Steigende und fallende Flanken treten im statistischen Mittel gleich häufig auf.

Die Varianz der Zustandsvektorkomponente $x_2(t)$ erhält man bei statistischer Unabhängigkeit der Größen $x_2(t_0)$ und $w(t)$ aus der Berechnungsvorschrift

$$P_{x_2 x_2}(t) = E\{[x_2(t) - E\{x_2(t)\}]^2\} = E\{[x_2(t)]^2\}.$$

Setzt man $x_2(t)$ nach (Gl. 9.60) und die Varianz $E\{[w(t)]^2\} = n_w \cdot \delta(t)$ des Eingangsrauschprozesses $w(t)$ ein, folgt mit der Varianz $P_{x_2 x_2}(t_0)$ der Komponente $x_2(t)$ zum Zeitpunkt t_0

$$(Gl. 9.62) \quad P_{x_2x_2}(t) = e^{-2\gamma(t-t_0)} \cdot P_{x_2x_2}(t_0) + \frac{n_w}{2\gamma} [1 - e^{-2\gamma(t-t_0)}].$$

Betrachtet man den Grenzwert von (Gl. 9.62) für $\gamma \rightarrow 0$, erhält man eine linear über der Zeit anwachsende Kovarianz, was nicht den tatsächlichen Gegebenheiten entspricht. Strebt dagegen $t_0 \rightarrow -\infty$, kennzeichnet der Grenzwert nach (Gl. 9.63) bei konstantem Korrelationsparameter γ die stationären Signaländerungen.

$$(Gl. 9.63) \quad \lim_{t_0 \rightarrow -\infty} P_{x_2x_2}(t) = e^{-2\gamma(t-t_0)} \cdot P_{x_2x_2}(t_0) + \frac{n_w}{2\gamma} = \frac{n_w}{2\gamma} = \text{const.}$$

Während die Rauschleistungsdichte n_w in Kap. 9.3 als Parameter für den Feinabgleich des Kalman-Filters dient, soll n_w im folgenden zunächst als typische Varianz des Eingangsrauschprozesses $w(t)$ betrachtet werden. Anhand von Versuchsreihen kann beispielsweise die Varianz $(\sigma_{vv})^2$ des Ausgangsrauschprozesses bei fehlender Anregung sowie die Varianz $(\sigma_{ww})^2$ von $x_2(t)$ bei der zusätzlichen Anregung mit dem Rauschprozess $w(t)$ ermittelt werden. Als Testsignale zur Ermittlung von $(\sigma_{ww})^2$ werden wiederum typische Störsignale verwendet. Dieselbe Vorgehensweise wurde bereits in Kap. 9.1.5 bei der Herleitung von (Gl. 9.36) verwendet. Aus (Gl. 9.63) folgt dann

$$(Gl. 9.64) \quad n_w = (\sigma_{ww}^2 - \sigma_{vv}^2) \cdot 2\gamma \approx 2\gamma \cdot \sigma_{ww}^2.$$

Die oben aufgeführte Näherung gilt für den Fall, dass $\sigma_{ww}^2 \gg \sigma_{vv}^2$ ist, was bei der vorliegenden Aufgabenstellung durchaus unterstellt werden kann. Zur endgültigen Festlegung von n_w muss nun noch der Korrelationsparameter γ bestimmt werden. Dazu betrachtet man den Kovarianzkern $P_{x_2x_2}(t_1, t_2)$ der Komponente $x_2(t)$ für zwei unterschiedliche Zeitpunkte t_1 und t_2 .

Setzt man $x_2(t)$ nach (Gl. 9.60) in die Beziehung

$$P_{x_2x_2}(t_1, t_2) = E\{[x_2(t_1) - E\{x_2(t_1)\}] \cdot [x_2(t_2) - E\{x_2(t_2)\}]\}$$

ein, ergibt sich unter Beachtung der Unkorreliertheit von $x_2(t)$ und $w(t)$:

$$(Gl. 9.65) \quad P_{x_2x_2}(t_1, t_2) = \sigma_{ww}^2 \cdot e^{-\gamma(t_2-t_1)}$$

Normiert man den so berechneten Kovarianzkern auf die Varianz $(\sigma_{ww})^2$, erkennt man, dass die stationäre, normierte Kovarianz nur von der Zeitdifferenz $T_c = t_2 - t_1$ abhängt. Wird der Faktor γ so dimensioniert, dass die Kovarianzfunktion nach einer Sprunganregung im Zeitraum T_c auf $e^{-1} = 36,79\%$ ihres Maximalwertes abgefallen ist, folgt

$$(Gl. 9.66) \quad \frac{P_{x_2x_2}(t_1, t_2)}{\sigma_{ww}^2} = e^{-\gamma T_c} = e^{-1} \quad \Rightarrow \quad \gamma = \frac{1}{T_c}.$$

Die Konstante γ entspricht also dem Kehrwert der auch als Korrelationszeitkonstante bezeichneten Größe T_c . T_c und damit γ können ebenfalls offline anhand von Testdaten bestimmt werden.

9.2.2 Ableitung des zeitdiskreten Systemmodells

Analog zur Herleitung des zeitdiskreten Systemmodells in Kap. 9.1.3 aus dem zeitkontinuierlichen Modell nach Kap. 9.1.2 wird in diesem Abschnitt das äquivalente zeitdiskrete Systemmodell zu dem zeitkontinuierlichen System nach (Gl. 9.55) – (Gl. 9.59) berechnet. Die globale Zustandsübergangsmatrix $\Phi(t-t_0)$ des ebenfalls zeitinvarianten allgemeinen Modells nach (Gl. 9.68) erhält man aus der Laplace-Transformierten (Gl. 9.67) durch Rücktransformation $L^{-1}\{\dots\}$ in den Zeitbereich.

$$(Gl. 9.67) \quad \Phi(s) = [s \cdot I - A_c]^{-1} = \begin{bmatrix} s & -1 \\ 0 & s + \gamma \end{bmatrix}^{-1} = \frac{1}{s \cdot (s + \gamma)} \cdot \begin{bmatrix} s + \gamma & 1 \\ 0 & s \end{bmatrix}$$

$$(Gl. 9.68) \quad \Phi(t-t_0) = L^{-1}\left\{\Phi(s)\right\}_{t-t_0} = \frac{1}{\gamma} \cdot \begin{bmatrix} \gamma & 1 - e^{-\gamma(t-t_0)} \\ 0 & \gamma \cdot e^{-\gamma(t-t_0)} \end{bmatrix}$$

Mit der globalen Zustandsübergangsfunktion $\Phi(t-t_0)$ erhält man die Systemmatrix A_3 nach (Gl. 9.71) der zeitdiskreten Zustandsübergangsgleichung (Gl. 9.69). Der zeitdiskrete Rauschprozess wurde wieder als vektorielle Größe angesetzt. Die Beobachtungsmatrix C_3 in der Ausgangsgleichung (Gl. 9.72) ist identisch mit der zeitkontinuierlichen Beobachtungsmatrix C_c nach (Gl. 9.59). Die Konstante K in (Gl. 9.72) wird im folgenden gleich Eins gesetzt.

$$(Gl. 9.69) \quad \underline{x}(k+1) = A_3 \cdot \underline{x}(k) + \underline{w}(k) \quad (Gl. 9.70) \quad y(k) = C_3 \cdot \underline{x}(k) + v(k)$$

$$(Gl. 9.71) \quad A_3 = \Phi(T) = \frac{1}{\gamma} \cdot \begin{bmatrix} \gamma & 1 - e^{-\gamma T} \\ 0 & \gamma e^{-\gamma T} \end{bmatrix} \approx A_3^* = \begin{bmatrix} 1 & T \\ 0 & e^{-\gamma T} \end{bmatrix} \quad (Gl. 9.72) \quad C_3 = C_c = K \begin{bmatrix} 1 & 0 \end{bmatrix}$$

Trotz des skalaren Prozessrauschens $w(t)$ wird im zeitdiskreten Fall ein Rauschvektor $\underline{w}(k)$ zur Berücksichtigung der Rauscheigenschaften des zeitkontinuierlichen Prozesses benötigt, wie die nachstehende Ableitung zeigt. Aus (Gl. 8.16)

$$Q_3 = \int_{kT}^{(k+1)T} \Phi((k+1)T - \tau) \cdot G_c \cdot E\{w(\tau) \cdot w(\tau)\} \cdot G_c^T \Phi((k+1)T - \tau)^T d\tau$$

folgt mit dem linearen Erwartungswert $E\{w(t)\} = 0$ und mit der Kovarianz $E\{[w(t)]^2\} = n_w$ des Eingangsruschprozesses sowie der stochastischen Kontrollmatrix G_c nach (Gl. 9.57) die Beziehung

$$(Gl. 9.73) \quad Q_3 = \frac{n_w}{\gamma^3} \begin{bmatrix} \gamma T - 2(1 - e^{-\gamma T}) + \frac{1}{2}(1 - e^{-2\gamma T}) & \gamma \left[(1 - e^{-\gamma T}) - \frac{1}{2}(1 - e^{-2\gamma T}) \right] \\ \gamma \left[(1 - e^{-\gamma T}) - \frac{1}{2}(1 - e^{-2\gamma T}) \right] & \frac{1}{2} \gamma^2 (1 - e^{-2\gamma T}) \end{bmatrix} \approx Q_3^* = \frac{n_w}{2\gamma} \begin{bmatrix} 0 & 0 \\ 0 & 1 - e^{-2\gamma T} \end{bmatrix}$$

Die Näherung A_3^* in (Gl. 9.71) gilt ebenso wie die Näherung Q_3^* aus (Gl. 9.73) für kleine Abtastintervalle T . Die Elemente A_{22} und Q_{22} der Matrizen A_3 und Q_3 wurden im Gegensatz zu den übrigen Matrixelementen nicht vereinfacht, um die System- bzw. Korrelationseigenschaften nicht zu stark zu verändern. Die Konstante n_w kann nach (Gl. 9.64) aus der Varianz $(\sigma_{ww})^2$ des zeitkontinuierlichen Rauschprozesses ermittelt werden.

Bei der Näherungslösung Q_3^* liegen unkorrelierte Rauschkomponenten $w_1(k)$ und $w_2(k)$ vor, da die Nebendiagonalterme der Kovarianzmatrix gleich Null sind. Weiterhin verschwindet mit dem Hauptdiagonalterm Q_{11}^* der Kovarianzmatrix auch die Rauschkomponente $w_1(k)$. Als Näherungslösung tritt somit zeitdiskreter skalarer Rauschprozess zur Modellierung auf. In diesem Sonderfall existiert auch die zeitdiskrete stochastische Kontrollmatrix $G = G_c = [0 \ 1]$.

Betrachtet man die nochmals die zeitdiskrete Systemmatrix A_3^* , fällt auf, dass die Berechnung des Signalverlaufs $x_1(k+1)$ ausschließlich anhand des vorherigen Werts $x_1(k)$ und der letzten Änderung $x_2(k)$ durch Extrapolation im Zeitintervall T erfolgt.

9.2.3 Simulationsergebnisse bei Verwendung des allgemeinen Signalmodells

In diesem Abschnitt werden typische Bewegungsmeldersignale mit dem allgemeinen Signalmodell nach (Gl. 9.69) - (Gl. 9.72) verarbeitet (Kalman-Filter 3). Zur Erzeugung der in Abb. 9.5 und Abb. 9.6 dargestellten Signalverläufe wurde für den Eingangsrauschprozess die Varianz $(\sigma_{ww})^2 = \text{Var}W_s = 10.0$ verwendet, während für den Ausgangsrauschprozess eine relativ geringe Streubreite mit der Varianz $(\sigma_{vv})^2 = \text{Var}V = 1.0$ angenommen wurde. Die Darstellungen gelten jeweils für ein eingeschwungenes Kalman-Filter.

Da $(\sigma_{vv})^2 \ll (\sigma_{ww})^2$ ist, wird das gemessene Ausgangssignal $y(k)$ bei der Zustandsschätzung relativ stark gewichtet. Im umgekehrten Fall einer höheren statistischen Unsicherheit des Ausgangsrauschprozesses würde das Kalman-Filter seine vorangegangenen Schätzwerte stärker berücksichtigen als die Messwerte.

Abb. 9.5 zeigt im oberen Teil wiederum die in den Beobachtungsraum abgebildeten Zustandsvektorkomponenten im Zusammenhang mit dem gemessenen Ausgangssignal $y(k)$ sowie die linearen Erwartungswerte der Prädiktions- und der Filterdichte. Man erkennt, dass das Ausgangssignal $y(k)$ deutlich besser nachgeführt wird als bei dem Bandpassmodell (Kalman-Filter 1).

Die in Abb. 9.5 dargestellte Residuensequenz zeigt dieses Verhalten ebenfalls deutlich. Im Vergleich zu Kalman-Filter 2 verläuft die Residuensequenz nach Abb. 9.5 sogar nahezu offsetfrei, was als Indiz für eine bessere Anpassung des allgemeinen Signalmodells an die vorliegende Aufgabenstellung gewertet werden kann.

An dieser Stelle soll zusätzlich der Verlauf der zur Zustandsprädiktion und -estimation verwendeten Kovarianzmatrizen und des Kalman-Gains über der Zeit betrachtet werden. Abb. 9.6 zeigt, dass die Komponenten der Kovarianzmatrizen der von Prädiktions- und Filterdichte im Anschluss an die hier nicht dargestellte Einschwingphase des Filters jeweils gegen einen konstanten Wert streben. Gleiches gilt für die Komponenten der Kalman-Gain-Matrix nach Abb. 9.5.

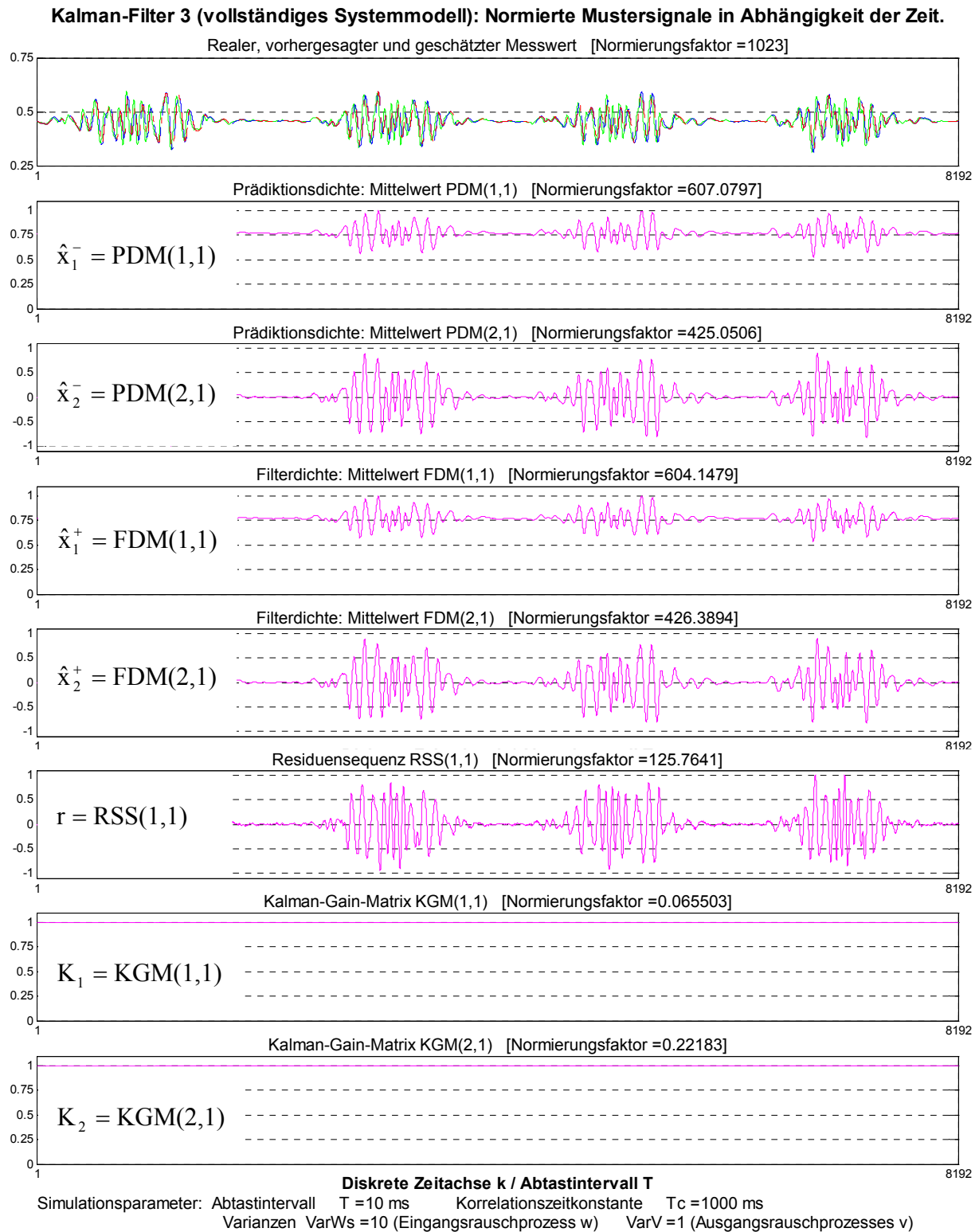


Abb. 9.5: Kalman-Filter 3. Normierte Ausgangssignale und Mittelwerte PDM und FDM der Prädiktions- und Filterdichte sowie Residuensequenz RSS und Kalman-Gain-Matrix KGM in Abhängigkeit der Zeit. Im oberen Diagramm sind das geschätzte, das vorhergesagte und das gemessene Ausgangssignal dargestellt.

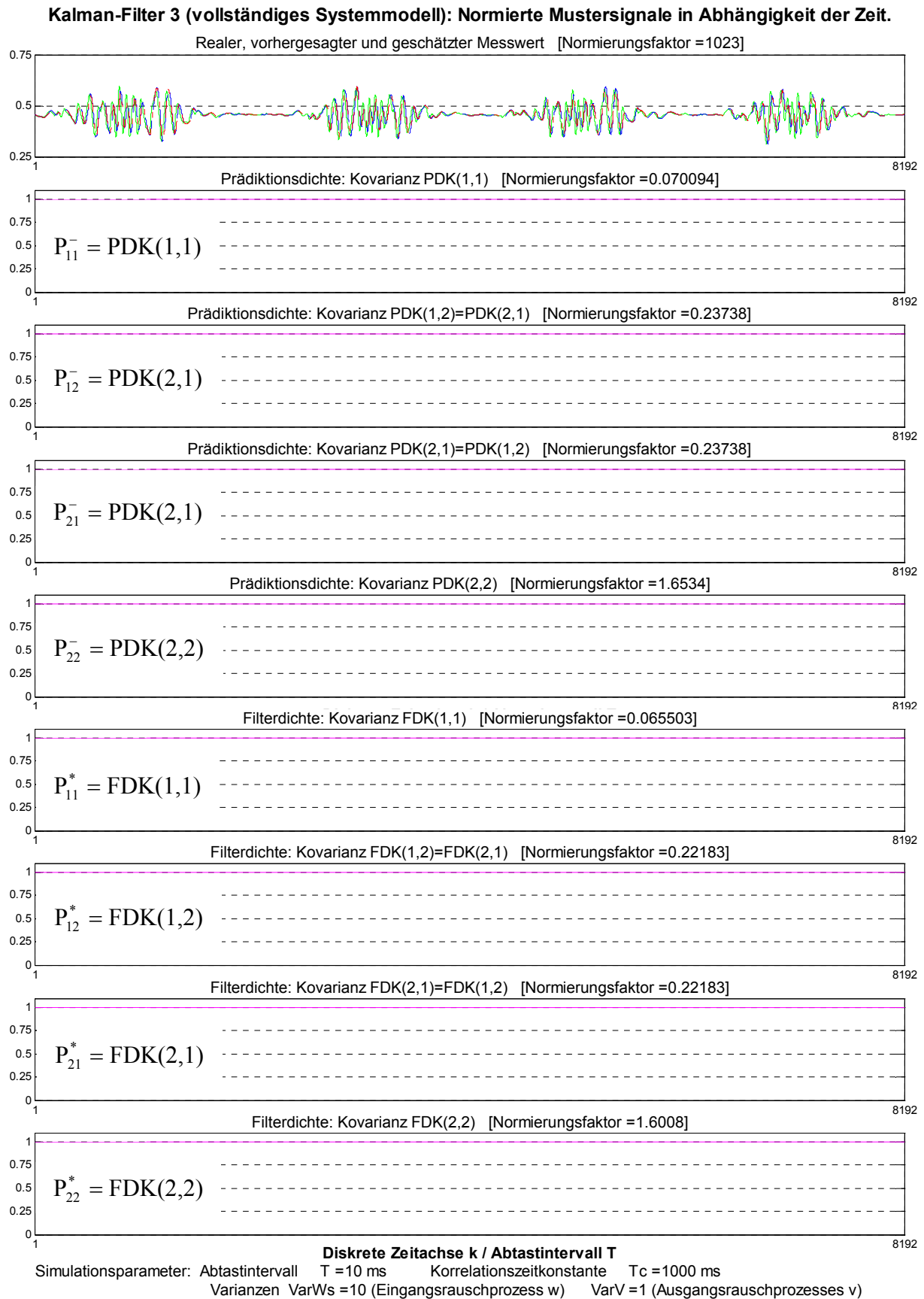


Abb. 9.6: Kalman-Filter 3. Normierte Ausgangssignale und Verlauf der Kovarianzmatrizen PDK und FDK der Prädiktions- und Filterdichte in Abhängigkeit der Zeit. Im oberen Diagramm sind das geschätzte, das vorhergesagte und das gemessene Ausgangssignal gleichzeitig dargestellt.

Von besonderem Interesse für die nachfolgenden Betrachtungen sind die Möglichkeiten zur Anpassung des Kalman-Filters an die reale Signalcharakteristik. Als Parameter zur Feineinstellung werden im weiteren Verlauf die Korrelationszeitkonstante T_c und die Kovarianz $(\sigma_{ww})^2$ des Eingangsruschprozesses $w(t)$ verwendet.

Die Auswirkungen einer Variation der Korrelationszeitkonstante T_c auf die Zustandsvektorkomponenten und damit auf das nachgeführte Ausgangssignal zeigt die obere Hälfte von Abb. 9.7. Zur Darstellung wurden Korrelationszeitkonstanten von $T_c = 0.1$ s, 0.5 s, 1.0 s, 5.0 s und 10.0 s gewählt. Abgebildet ist jeweils das gemessene Ausgangssignal $y(k)$ sowie die in den Beobachtungsraum abgebildeten Zustandsvektorkomponenten $C_3 \cdot \hat{x}^-(k)$ und $C_3 \cdot \hat{x}^+(k)$ des vorhergesagten und des geschätzten Zustandsvektors $\hat{x}^-(k)$ und $\hat{x}^+(k)$.

Die untere Hälfte von Abb. 9.7 zeigt den zeitlichen Verlauf der zugehörigen Residuensequenzen für die verschiedenen Korrelationszeitkonstanten. Die Residuensequenz entsteht, wie bereits vorne beschrieben wurde, aus der Differenz des Ausgangssignals $y(k)$ und des prädierten und abgebildeten Zustandsvektors $C_3 \cdot \hat{x}^-(k)$. Diese Korrekturinformation wird verwendet, um den optimalen Schätzwert $\hat{x}^+(k)$ zu bestimmen.

Anhand der in Abb. 9.7 angegebenen Normierungsfaktoren erkennt man, dass die Amplituden der Residuensequenzen mit einer Verringerung der Korrelationszeitkonstante T_c ebenfalls kleiner werden. Die Korrelationszeitkonstante T_c bzw. deren Kehrwert, der Korrelationsparameter γ , wirkt sich dabei insbesondere bei der Zustandsprädiktion mit der Systemmatrix A_3 aus. Die beste Übereinstimmung des vorhergesagten und des geschätzten Ausgangssignals mit dem gemessenen Signalverlauf erhält man nach Abb. 9.7 mit einer Korrelationszeitkonstante $T_c = 1.0$ s, wenn ein Eingangsruschprozess mit einer Varianz von $(\sigma_{ww})^2 = \text{VarWs} = 10$ angenommen wird.

Abb. 9.8 zeigt die Signalverläufe der gemessenen und der in den Beobachtungsraum abgebildeten Zustandsvektorkomponenten sowie die zugehörigen Residuensequenzen in Abhängigkeit der Varianz $(\sigma_{ww})^2 = \text{VarWs}$. Für die Darstellung wurden die Varianzen $(\sigma_{ww})^2 = \text{VarWs} = 1.0, 5.0, 10.0, 50.0, 100.0$ bei fester Korrelationszeitkonstante $T_c = 1.0$ s verwendet.

Insbesondere bei kleinen Varianzen des Eingangsruschprozesses ist das Kalman-Filter nicht in der Lage, den Ausgangssignalverlauf nachzuführen. Diese Eigenschaft wird später ausgenutzt, um die Trägheit des Kalman-Filters an die Signaleigenschaften typischer Störsignale anzupassen. Die Größe der Varianz $(\sigma_{ww})^2 = \text{VarWs}$ wirkt sich ebenfalls unmittelbar auf die Residuensequenz aus. Den in Abb. 9.8 aufgeführten Normierungsfaktoren ist zu entnehmen, dass die Amplituden in der Residuensequenz mit zunehmender Varianz des Eingangsruschprozesses abnehmen. Das Kalman-Filter kann in diesem Fall der ursprünglichen Signaldynamik besser folgen.

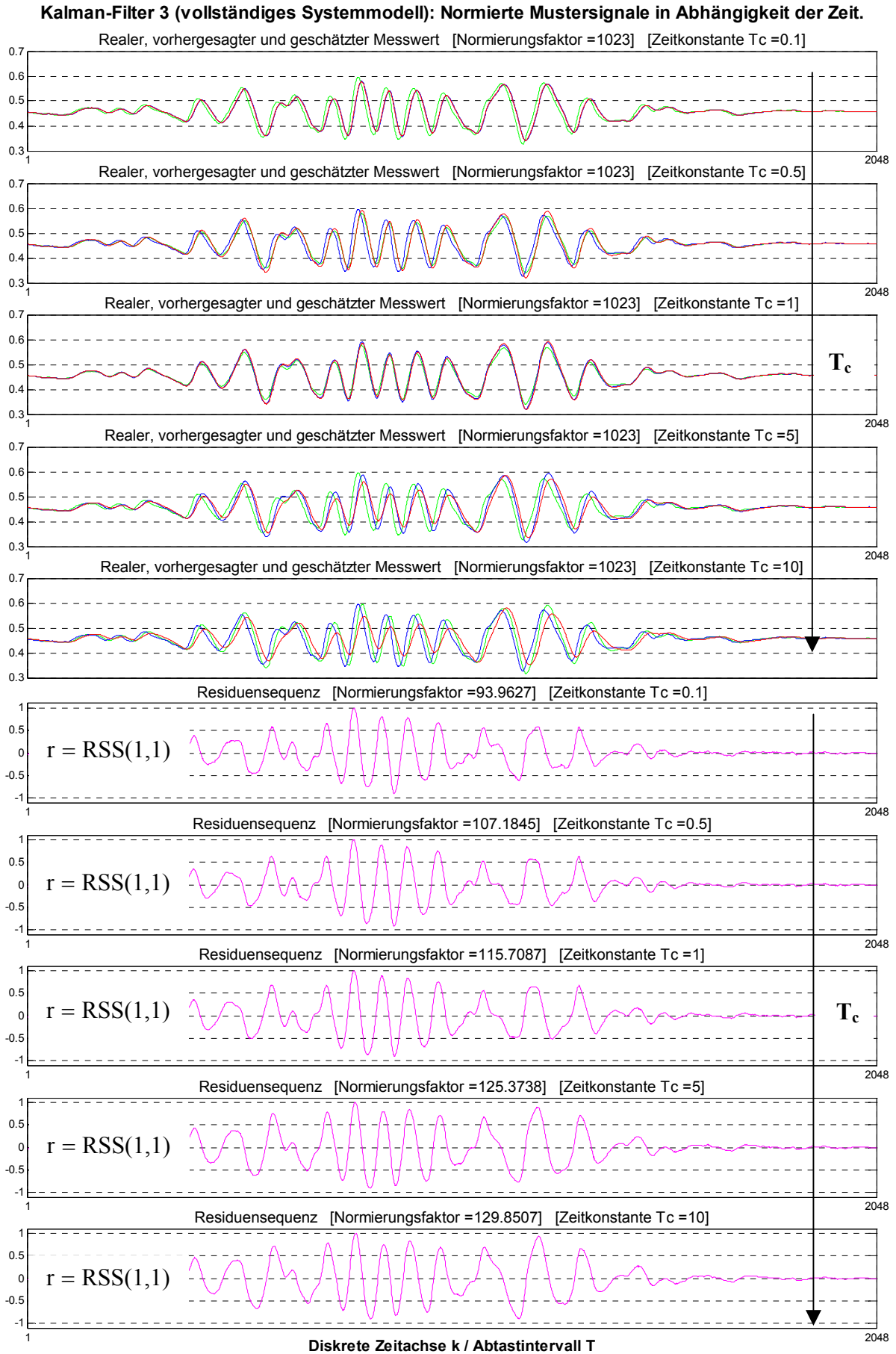


Abb. 9.7: Kalman-Filter 3. Normierte Ausgangssignale und Residuensequenz als Funktion der Korrelationszeitkonstante T_c . Eingangsräuschprozess: $\sigma_{ww}^2 = \text{Var}W_s = 10$.

Kalman-Filter 3 (vollständiges Systemmodell): Normierte Mustersignale in Abhängigkeit der Zeit.

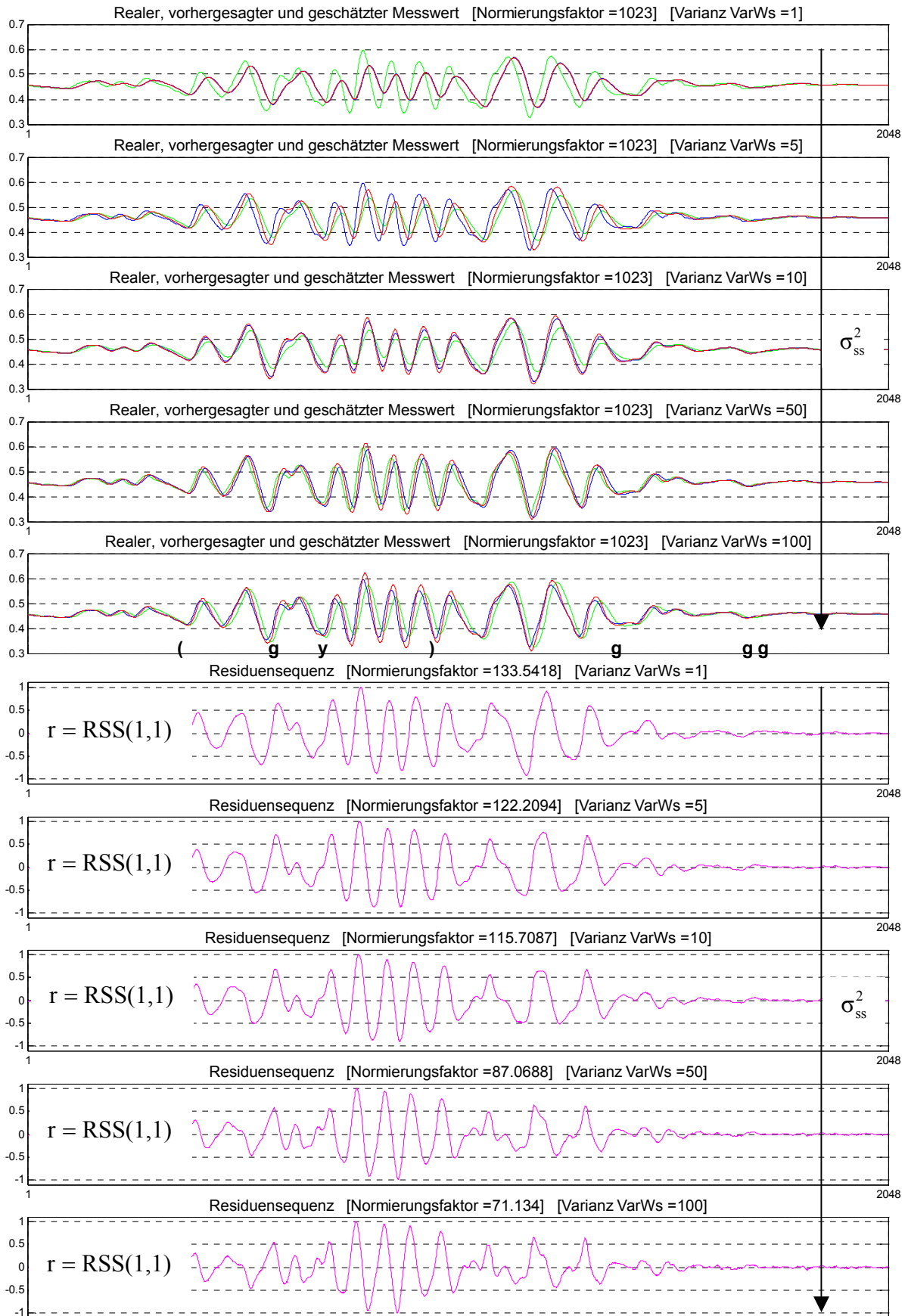


Abb. 9.8: Kalman-Filter 3. Normierte Ausgangssignale und Residuensequenzen in Abhängigkeit der Varianz $\sigma_{ww}^2 = \text{VarWs}$. Korrelationszeitkonstante: $T_c = 1.0$ s.

Wenngleich das allgemeine Signalmodell mit einer entsprechenden Anpassung der Korrelationszeitkonstante T_c und der Varianz $(\sigma_{ww})^2$ des Eingangsruschprozesses an die Dynamik des Hintergrundsignals bereits eine erste Möglichkeit zur Ausblendung des Hintergrunds darstellt, soll im folgenden mit Hinblick auf die Korrelationseigenschaften des realen Prozesses eine Erweiterung dieses allgemeinen Signalmodells vorgenommen werden.

9.3 Allgemeines Signalmodell mit spezieller Berücksichtigung von Drifteinflüssen

9.3.1 Herleitung des erweiterten Signalmodells

Im Rahmen der Erweiterung des allgemeinen Signalmodells wird die Bandpass-Charakteristik des vereinfachten Bewegungsmeldermodells nach Kap. 9.1.1 - 9.1.4 zur Modellierung der Drifteinflüsse des Hintergrundsignals verwendet, wodurch eine realitätsgetreuere Nachbildung des tatsächlichen Systemverhaltens möglich ist. Die Vorgehensweise zur Kombination beider Modelle wird im folgenden beschrieben.

Zur Ableitung des erweiterten Zustandsraummodells (Kalman-Filter 4) wird zunächst ein vergrößerter Zustandsvektor (Gl. 9.74) und ein vergrößerter Rauschvektor (Gl. 9.75) definiert.

$$(Gl. 9.74) \quad \underline{x}(k) = [\underline{x}_s(k) \mid \underline{x}_d(k)]^T \quad (Gl. 9.75) \quad \underline{w}(k) = [\underline{w}_s(k) \mid \underline{w}_d(k)]^T$$

$\underline{x}_s(k)$ und $\underline{w}_s(k)$ bezeichnen dabei den Zustandsvektor $\underline{x}(k)$ und den Rauschvektor $\underline{w}(k)$ des allgemeinen Signalmodells nach (Gl. 9.69), während unter $\underline{x}_d(k)$ und $\underline{w}_d(k)$ die Zustandsvektor- und Rauschkomponenten des vereinfachten Bewegungsmeldermodells aus (Gl. 9.27) zusammengefasst sind. Die Zustandsgleichung lautet in diesem Fall

$$(Gl. 9.76) \quad \underline{x}(k+1) = A_4 \cdot \underline{x}(k) + \underline{w}(k)$$

Mit den zeitdiskreten Systemmatrizen A_1 und A_3 gilt für die Systemmatrix A_4 des erweiterten, zeitdiskreten Modells die Beziehung (Gl. 9.77). Für die Kovarianzmatrix Q_4 ergibt sich bei derselben Vorgehensweise mit den Kovarianzmatrizen Q_1 und Q_3 der Zusammenhang nach (Gl. 9.78), wobei die einzelnen Komponenten der Kovarianzmatrix Q_4 in (Gl. 9.79) - (Gl. 9.84) angegeben sind.

$$(Gl. 9.77) \quad A_4 = \begin{bmatrix} A_3 & 0 \\ 0 & A_1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1-e^{-\gamma T}}{\gamma} & 0 & 0 \\ 0 & e^{-\gamma T} & 0 & 0 \\ 0 & 0 & \frac{-\beta \cdot e^{-\alpha T} + \alpha \cdot e^{-\beta T}}{\alpha - \beta} & \frac{-e^{-\alpha T} + e^{-\beta T}}{\alpha - \beta} \\ 0 & 0 & \frac{\alpha \beta \cdot (e^{-\alpha T} - e^{-\beta T})}{\alpha - \beta} & \frac{\alpha \cdot e^{-\alpha T} - \beta \cdot e^{-\beta T}}{\alpha - \beta} \end{bmatrix}$$

$$(Gl. 9.78) \quad Q_4 = \begin{bmatrix} Q_3 & 0 \\ 0 & Q_1 \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} & 0 & 0 \\ Q_{21} & Q_{22} & 0 & 0 \\ 0 & 0 & Q_{33} & Q_{34} \\ 0 & 0 & Q_{43} & Q_{44} \end{bmatrix}$$

$$(Gl. 9.79) \quad Q_{11} = \frac{n_w}{\gamma^3} \left[\gamma T - 2(1 - e^{-\gamma T}) + \frac{1}{2}(1 - e^{-2\gamma T}) \right]$$

$$(Gl. 9.80) \quad Q_{12} = Q_{21} = \frac{n_w}{\gamma^2} \left[(1 - e^{-\gamma T}) - \frac{1}{2}(1 - e^{-2\gamma T}) \right]$$

$$(Gl. 9.81) \quad Q_{22} = \frac{n_w}{2\gamma} [1 - e^{-2\gamma T}]$$

$$(Gl. 9.82) \quad Q_{33} = \frac{N_w}{(\alpha - \beta)^2} \left[\frac{1 - e^{-2\alpha T}}{2\alpha} + \frac{1 - e^{-2\beta T}}{2\beta} + \frac{2(1 - e^{-(\alpha+\beta)T})}{\alpha + \beta} \right]$$

$$(Gl. 9.83) \quad Q_{34} = Q_{43} = \frac{N_w}{2(\alpha - \beta)^2} [e^{-\alpha T} - e^{-\beta T}]^2$$

$$(Gl. 9.84) \quad Q_{44} = \frac{N_w}{(\alpha - \beta)^2} \left[\frac{\alpha(1 - e^{-2\alpha T})}{2} + \frac{\beta(1 - e^{-2\beta T})}{2} + \frac{2\alpha\beta(1 - e^{-(\alpha+\beta)T})}{\alpha + \beta} \right]$$

Aus (Gl. 9.79) - (Gl. 9.84) erkennt man, dass nunmehr zwei Eingangsruschprozesse mit den Varianzen n_w zur Einstellung des allgemeinen Signalmodells und N_w zur Anpassung des Driftmodells an die Eigenschaften von Störsignalen vorhanden sind. Beide Größen werden im weiteren Verlauf als Parameter zum Feinabgleich des Kalman-Filters verwendet, wobei vorausgesetzt wird, dass die zugrundeliegenden Eingangsruschprozesse erwartungswertfrei, weiß und Gaußverteilt sind.

Infolge des vergrößerten Zustandsvektors $\underline{x}(k)$ nach (Gl. 9.74) muss die Beobachtungsmatrix C_4 des zugehörigen Beobachtungsmodells ebenfalls erweitert werden. Geht man entsprechend den Ausführungen in Kap. 8 davon aus, dass sich im Ausgangssignal $y(k)$ der Anteil $d(k)$ des Hintergrundsignals und der Nutzsignalanteil $s(k)$, der ausschließlich in Verbindung mit einem bewegten Objekt im Erfassungsfeld des Melders auftritt, additiv überlagern, kann mit der vergrößerten Beobachtungsmatrix C_4 nach (Gl. 9.86) das Beobachtungsmodell (Gl. 9.85) des erweiterten Systems definiert werden.

$$(Gl. 9.85) \quad y(k) = C_4 \cdot \underline{x}_a(k) + \underline{v}(k) \quad (Gl. 9.86) \quad C_4 = [C_3 \mid C_1] = K \cdot [1 \ 0 \ 0 \ 1]$$

Der Driftanteil $d(k)$, der aus dem vereinfachten Bewegungsmeldermodell nach Kap. 9.1.1 - 9.1.4 abgeleitet wird, beeinflusst infolge des 4. Matrixelements der Beobachtungsmatrix C_4 die Berechnung des geschätzten Ausgangssignals und damit auch die Residuensequenz. Die 1. Komponente der Beobachtungsmatrix berücksichtigt den konstanten Offset des Bewegungsmeldersignals sowie alle Signalanteile, die nicht von dem Driftmodell erfasst werden, bei der Bildung der Residuensequenz.

Der zeitdiskrete Rauschprozess $v(k)$ aus (Gl. 9.85) wird im folgenden wiederum als erwartungswertfreier, weißer und Gaußverteilter Ausgangsrauschprozess mit der Varianz R betrachtet, so dass gilt:

$$(Gl. 9.87) \quad E\{v(k)\} = 0 \quad (Gl. 9.88) \quad E\{[v(k) - E\{v(k)\}]^2\} = E\{[v(k)]^2\} = R \cdot \delta(k)$$

9.3.2 Simulationsergebnisse bei Verwendung des allgemeinen Signalmodells

Abb. 9.9 und Abb. 9.10 zeigen beispielhaft den Signalverlauf eines gemessenen Ausgangssignals und der in den Beobachtungsraum abgebildeten Zustandsgrößen über der Zeit. Außerdem wurden die Zustandsgrößen selbst sowie die Kovarianzmatrizen der Prädiktions- und der Filterdichte, das Kalman-Gain und die zugehörige Residuensequenz abgebildet. Für das Nutzsinal wurde die Varianz $(\sigma_{ss})^2 = \text{Var}W_s = 10$ mit einer Korrelationszeitkonstante $T_c = 1.0$ s zugrundegelegt, während für den Störprozess am Eingang die Varianz $(\sigma_{dd})^2 = \text{Var}W_d = 1.0$ und für den Ausgangsrauschprozess die Varianz $(\sigma_{vv})^2 = \text{Var}V = 1.0$ angenommen wurde.

Betrachtet man den zeitlichen Verlauf des vorhergesagten oder des geschätzten Zustandsvektors nach Abb. 9.9, fällt unmittelbar auf, das jeweils nur die erste Komponente des vorhergesagten bzw. des geschätzten Zustandsvektors mit einem DC-Offset behaftet ist. Dieses Verhalten erscheint zunächst unerwartet, da bei dem vereinfachten Bewegungsmeldermodell nach Abb. 9.2 ebenfalls ein DC-Offset in der ersten Komponente vorhanden war.

Die dritte Zustandsvektorkomponente des erweiterten Modells, die mit der ersten Komponente des vereinfachten Bewegungsmeldermodells übereinstimmt, ist dagegen gleichanteilsfrei. Außerdem ist der konstante Offset, der beim vereinfachten Bewegungsmeldermodell zwischen der zweiten Komponente des prädierten und des geschätzten Zustandsvektorkomponente auftrat (s. Abb. 9.2), verschwunden.

Die einzelnen Komponenten der Kovarianzmatrizen der Prädiktions- und der Filterdichte streben nach Abb. 9.10 wieder gegen einen konstanten Endwert. Anhand der in den Abbildungen angegebenen Normierungsfaktoren kann man die Verteilung der Signalleistung auf die einzelnen Zustandsvektorkomponenten einschließlich der entsprechenden Kreuzleistungsterme erkennen.

Ähnlich wie die Elemente der Kovarianzmatrizen sind auch die einzelnen Elemente der Kalman-Gain-Matrix im Anschluss an die nicht abgebildete Einschwingphase des Filters konstant (s. Abb. 9.9, unterer Teil). Die ebenfalls in Abb. 9.9 dargestellte Residuensequenz entspricht wiederum dem unerwarteten Signalanteil, wenn die oben angegebenen Parameter ($(\sigma_{ss})^2$, $(\sigma_{dd})^2$, $(\sigma_{vv})^2$, T_c) zur Modellierung des Nutzsinal- und des Störsignalanteils verwendet werden. Im Anschluss an diese eher allgemeine Beschreibung der Signalverläufe des Kalman-

Filters 4 soll im weiteren Verlauf ein entscheidender Vorteil des erweiterten Zustandsraummodells im Vergleich zu den vorhergehenden Ansätzen erläutert werden.

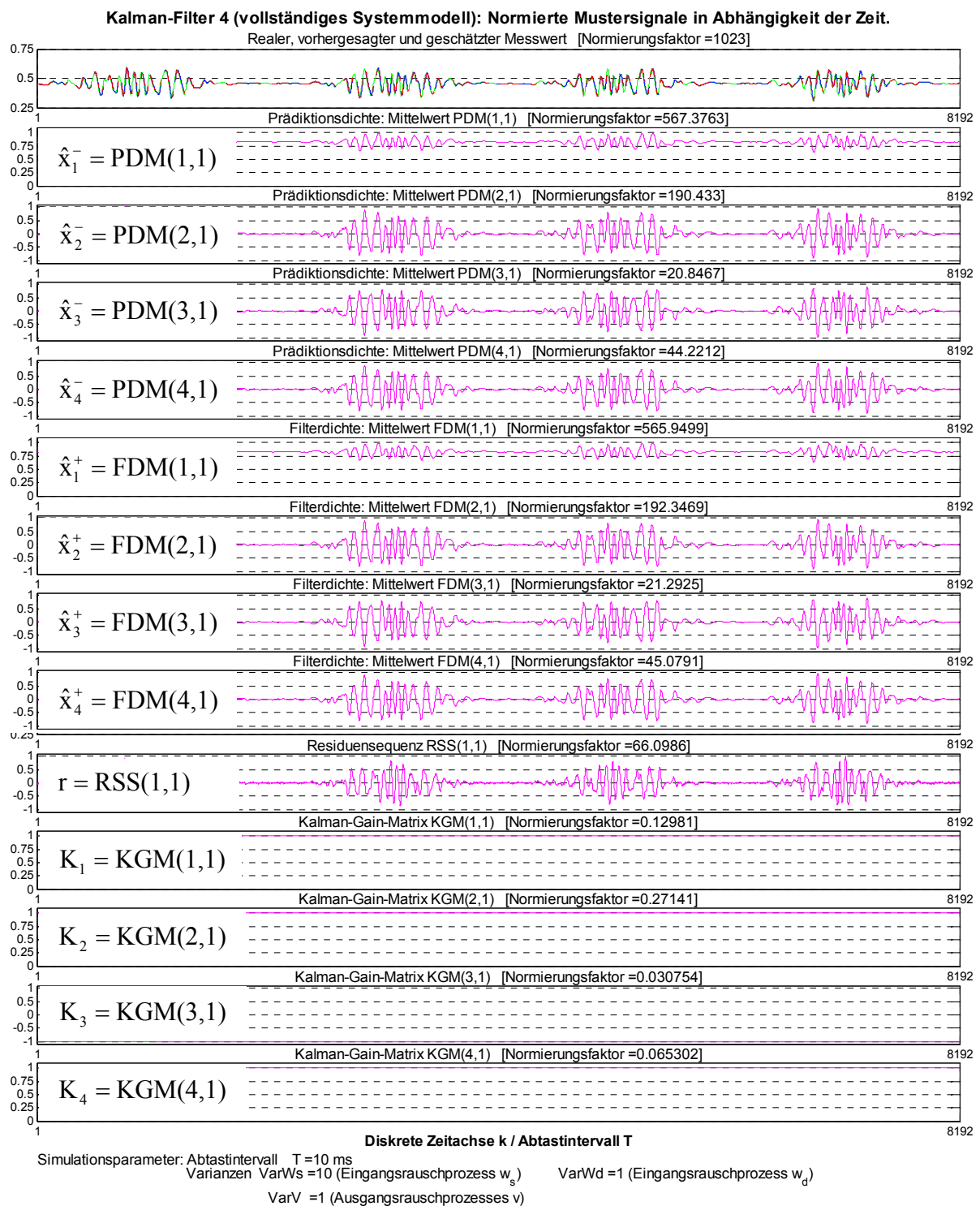


Abb. 9.9: Kalman-Filter 4. Normierte Ausgangssignale und lineare Erwartungswerte PDM und FDM der Prädiktions- und Filterdichte sowie Residuensequenz RSS und Kalman-Gain-Matrix KGM in Abhängigkeit der Zeit. Im oberen Diagramm sind das geschätzte, das vorhergesagte und das gemessene Ausgangssignal gleichzeitig dargestellt.



Abb. 9.10: Kalman-Filter 4. Normierte Ausgangssignale und Kovarianzen der Prädiktionsdichte PDK und der Filterdichte FDK in Abhängigkeit der Zeit.

Variiert man die Varianz $(\sigma_{dd})^2$ bei sonst festen Parametereinstellungen, wobei $(\sigma_{ss})^2$, T_c und $(\sigma_{vv})^2$ offline ermittelt werden können, bietet sich eine einfache Möglichkeit zur Steuerung der Empfindlichkeit der Bewegungserkennung. Abb. 9.11 - Abb. 9.14 verdeutlichen diesen Effekt. Als Einstieg in diese Thematik zeigt Abb. 9.11 zunächst wiederum die Nachführung des gemessenen Ausgangssignals durch die in den Beobachtungsraum abgebildeten Zustandsvektorkomponenten.

Erhöht man die Varianz $(\sigma_{dd})^2$ des Eingangsrauschprozesses, verringert sich der Nutzsignalanteil in der ersten und zweiten Komponente des Zustandsvektors (s. Abb. 9.12). Nach Abb. 9.13 steigt der Rauschanteil in der dritten und vierten Zustandsvektorkomponente, welche die Hintergrundstörungen und Driftprozesse modellieren, entsprechend an. Betrachtet man zusätzlich die Residuensequenz nach Abb. 9.14, fällt auf, dass diese mit steigender Störungsvarianz $(\sigma_{dd})^2$ ebenfalls kleiner wird. Die Residuensequenz kann also offensichtlich von den Störanteilen befreit werden, was für die nachfolgende Auswertung von entscheidender Bedeutung ist.

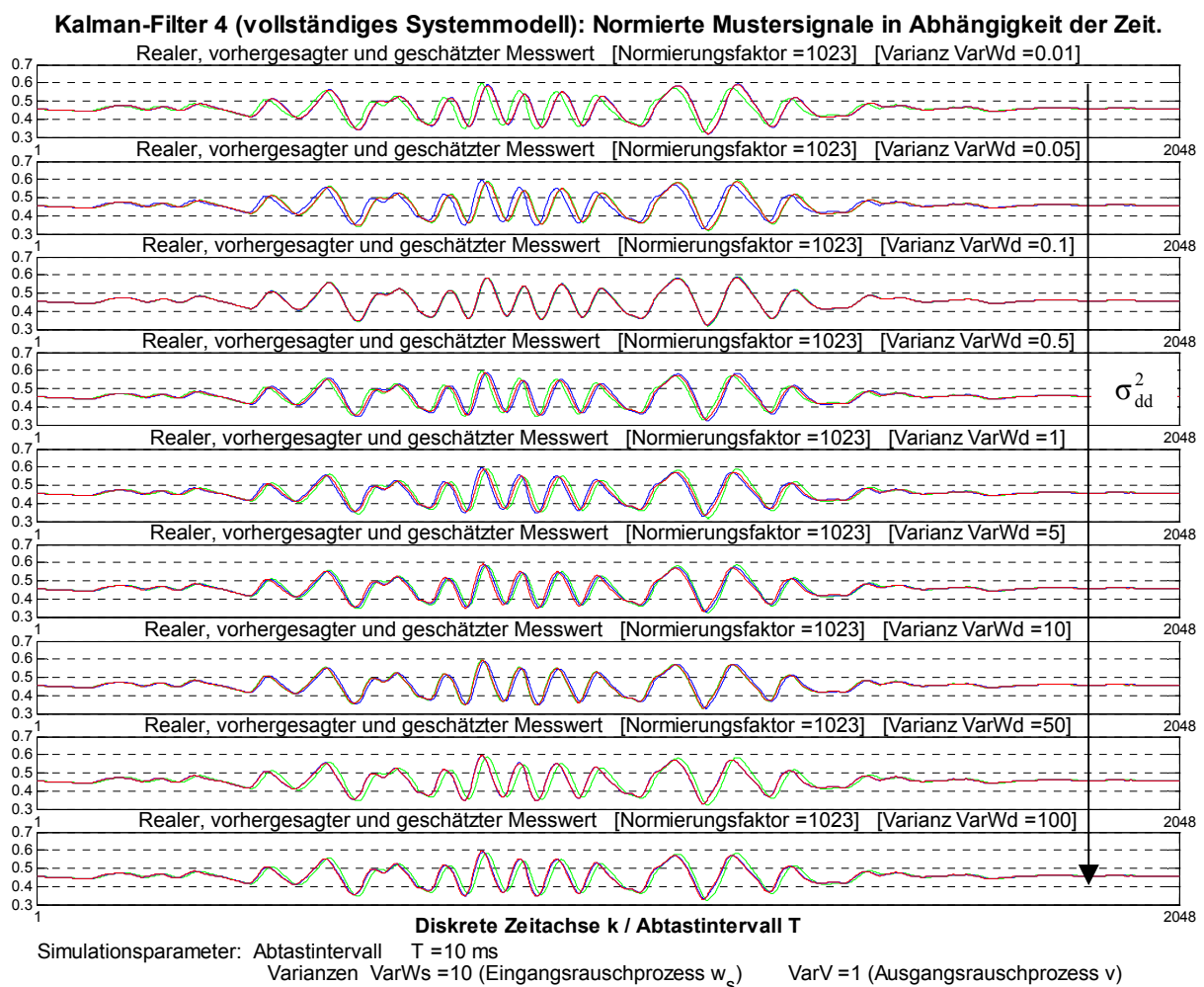


Abb. 9.11: Kalman-Filter 4. Normierte Ausgangssignale und in den Beobachtungsraum abgebildete Prädiktions- und Schätzwerte bei unterschiedlicher Varianz $\sigma_{dd}^2 = \text{VarWd}$ des Eingangsstörprozesses. Nutzsignal-Varianz. $\sigma_{ss}^2 = \text{VarWs} = 10$, Korrelationszeitkonstante $T_c = 1.0$ s.

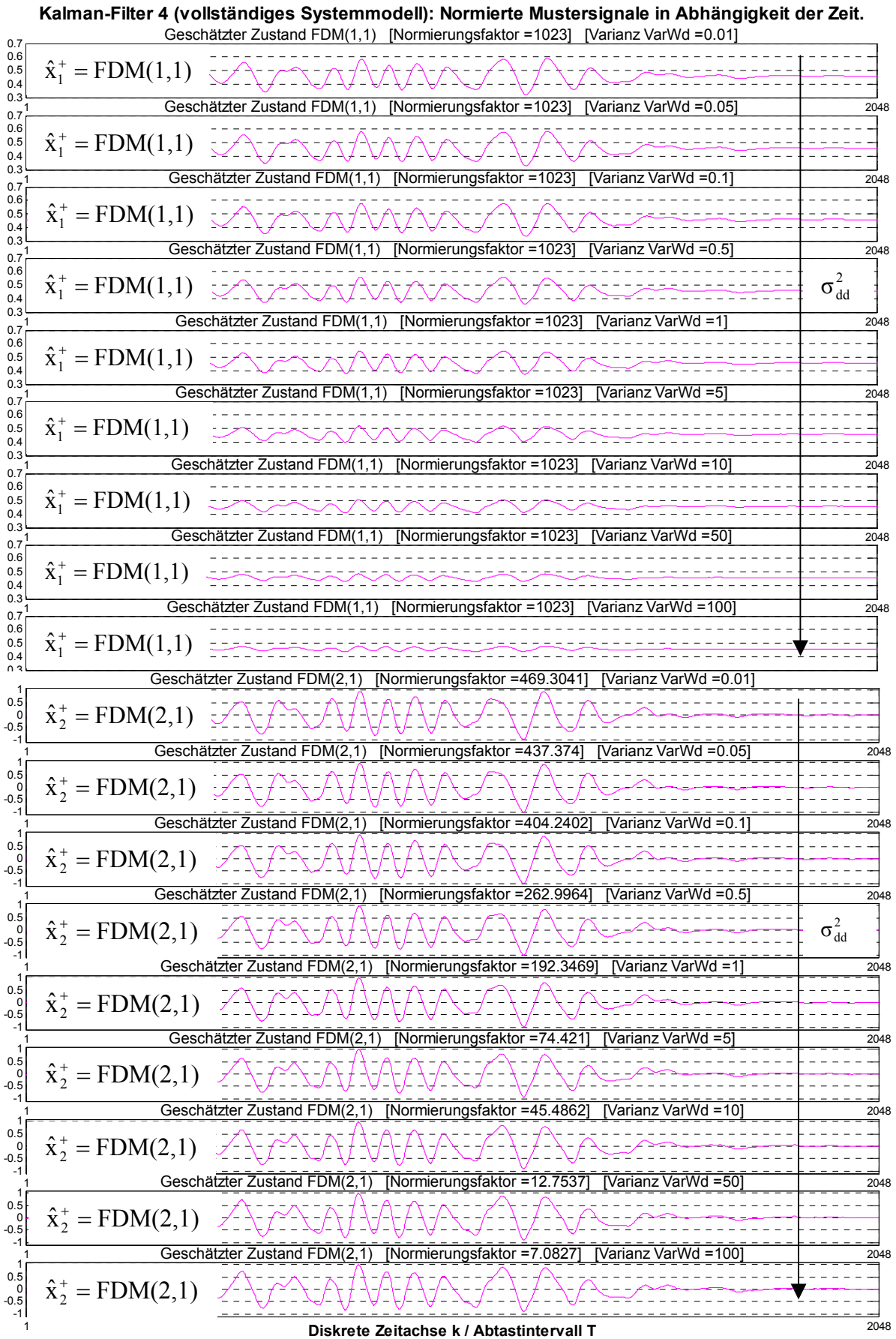


Abb. 9.12: Kalman-Filter 4. Zeitlicher Verlauf der 1. und 2. Komponente des geschätzten Zustandsvektors für die Ausgangssignale nach Abb. 9.11.



Abb. 9.13: Kalman-Filter 4. Zeitlicher Verlauf der 3. und 4. Komponente des geschätzten Zustandsvektors für die Ausgangssignale nach Abb. 9.11.

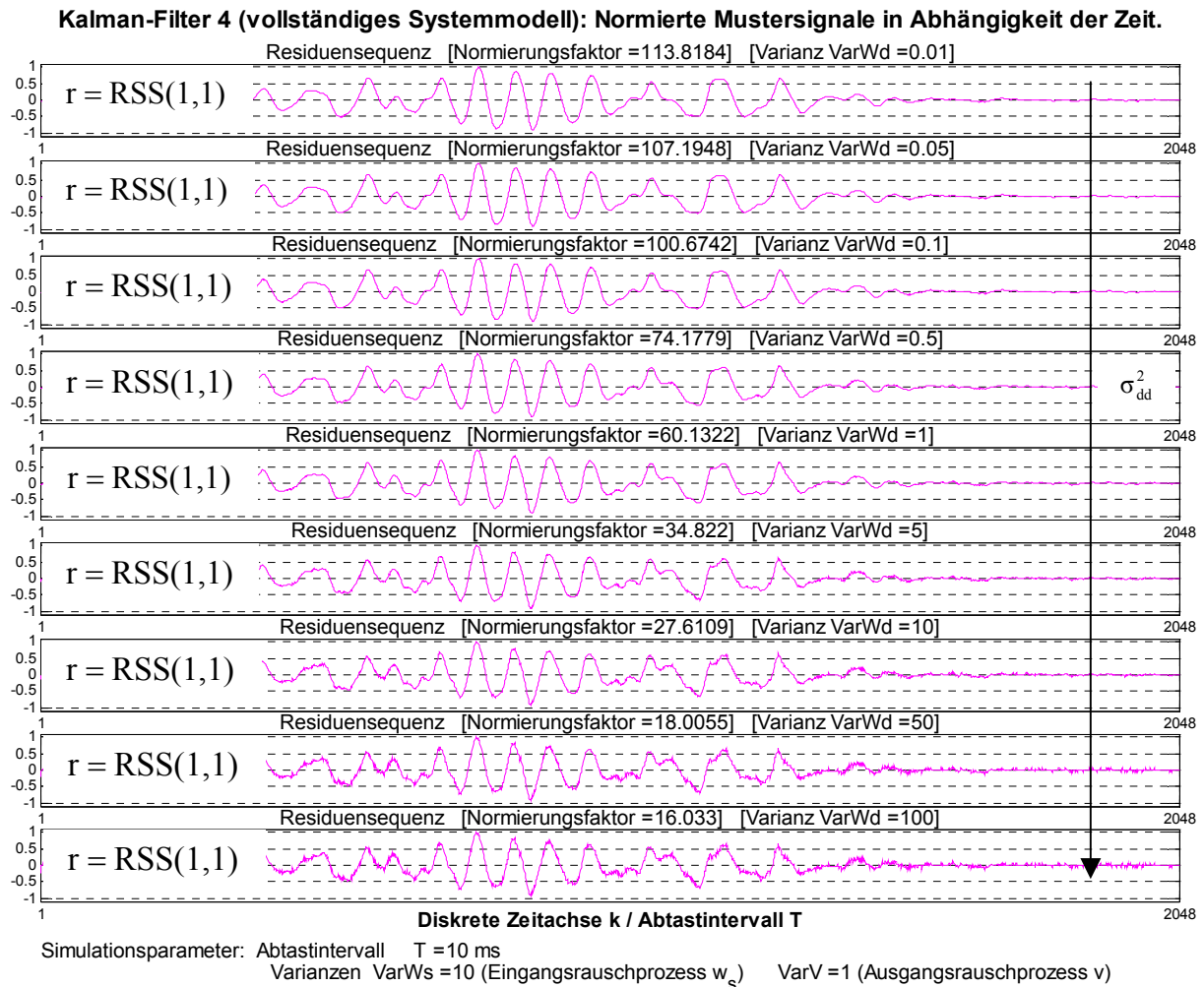


Abb. 9.14: Kalman-Filter 4. Residuensequenzen in Abhängigkeit der Varianz $\sigma_{dd}^2 = \text{Var}W_d$ für die Signalverläufe nach Abb. 9.11. Nutzsinal-Varianz $\sigma_{ss}^2 = \text{Var}W_s = 10$, Korrelationszeitkonstante $T_c = 1.0$ s.

Für eine optimale Störungsunterdrückung ist allerdings eine adaptive Anpassung der Störungsvarianz $(\sigma_{dd})^2$ an die tatsächlichen Verhältnisse im Erfassungsfeld nötig. Anwendungsmöglichkeiten adaptiver Kalman-Filter für diese Aufgabenstellung sind Gegenstand aktueller Untersuchungen, die bislang noch nicht abgeschlossen sind.

Prinzipiell bleibt aber festzuhalten, das sich auch ohne adaptives Verhalten bereits ein erstes Kalman-Filter zur Diskriminierung zwischen Nutz- und Störsignalen dimensionieren lässt. Ausgangspunkt ist die Kenntnis einer sinnvollen Obergrenze $(\sigma_{dd})^2$ typischer Störvarianzen, die offline aus aufgezeichneten Testdaten ermittelt werden kann. Im nächsten Kapitel wird dieser Ansatz verwendet, um mögliche Auswertestrategien für die so vorverarbeiteten Daten zu entwickeln.

10 Anwendung des Kalman-Filters zur Bewegungserkennung

10.1 Einleitung

In diesem Kapitel wird ein Verfahren zur Bewegungserkennung vorgestellt, welches die mit dem Kalman-Filter aus Kap. 9 vorverarbeiteten Sensorsignale als Ausgangspunkt für die Detektion verwendet. Als Einstieg in diese Thematik wird zunächst die Parametrierung des Kalman-Filters anhand von typischen Sensorsignalen aus dem Indoor- und dem Outdoorbereich sowie die experimentelle Optimierung dieser Einstellungen beschrieben.

Die Diskriminierung zwischen typischen Stör- und Nutzsignalen erfolgt beim Kalman-Filter 4 modellgestützt. Zur Anpassung des Filters an die zugrundeliegenden statistischen Eigenschaften des Hintergrundprozesses (Störungen) dient die Kovarianzmatrix Q_4 des Eingangsruschprozesses. Speziell das zur Modellierung des Störanteils verwendete Hintergrundmodell basiert auf dem vereinfachten Bewegungsmeldermodell, das in den vorangegangenen Kapiteln entwickelt wurde.

Ein erstes Verfahren zur Bewegungserkennung verwendet die Residuensequenz des Kalman-Filters. Wie bereits vorne erläutert wurde, wirken sich unerwartete Änderungen, welche beispielsweise durch die Bewegung von Personen im Erfassungsfeld des Melders hervorgerufen werden, in der Residuensequenz aus. Diese beschreibt die Abweichung des “vorhergesagten“ vom tatsächlich gemessenen Systemverhalten und wird daher auch als Innovationssequenz bezeichnet. Unerwartete Änderungen in der Residuensequenz sind unter Beachtung des zugrundeliegenden Störmodells daher ein sehr starkes Indiz für ein valides Detektionsereignis, wenngleich für eine völlig allgemeingültige Bewegungserkennung eine adaptive Anpassung des Störmodells an die zugrundeliegenden statistischen Eigenschaften des Hintergrundsignals erforderlich ist.

Eine adaptive Anpassung der Kovarianzmatrix Q_4 erfordert einen höheren Rechenaufwand als die zunächst verwendete Offline-Bestimmung dieser Matrix. Im Extremfall muss die Kovarianzmatrix für jeden Abtastzeitpunkt und für jeden Sensorkanal neu berechnet werden. Um Rechenzeit einzusparen, ist es allerdings auch denkbar, verschiedene Kovarianzmatrizen für unterschiedliche Hintergrunddynamiken offline zu bestimmen. Die Auswahl der im Filter zu verwendenden Kovarianzmatrix Q_4 des Eingangsruschprozesses kann dann in Abhängigkeit der Varianz der online ermittelten Störsignale erfolgen.

Neben der Bewertung der Residuensequenz werden im folgenden zwei weitere Verfahren zur Bewegungsdetektion vorgestellt. Unter Verwendung der bereits in Kap. 7 beschriebenen Auswertung charakteristischer Signaleigenschaften (Amplitude, Gradient, Flankendauer, Signaldauer) dienen die vom Kalman-Filter berechneten Prädiktionswerte zur Identifikation von

Signaländerungen. Zur Bewertung bieten sich insbesondere die 1. Komponente des vorhergesagten Zustandsvektors, welche das um das Hintergrundsignal bereinigte Nutzsignal darstellt, sowie die 2. Komponente des Zustandsvektors an. Letztere beschreibt näherungsweise die Änderung der 1. Zustandsvektorkomponente im betrachteten Zeitintervall. Beide Verfahren unterscheiden sich von den konventionellen Auswerteverfahren dadurch, dass die zu analysierenden Nutzsignale aus dem ursprünglichen Sensorsignal unter Verwendung des Hintergrundmodells berechnet werden. Störanteile im Sensorsignal werden daher wirkungsvoll unterdrückt, wenn die Dynamik des Störprozesses mit dem modellierten Systemverhalten übereinstimmt.

10.2 Parametrierung des Kalman-Filters anhand von typischen Sensorsignalen

Zur Anpassung des Kalman-Filters 4 (s. Kap. 9) an die verschiedenen Rauschprozesse werden im weiteren Verlauf typische Sensorsignale aus unterschiedlichen Testfeldern verwendet. Die Vorgehensweise bei der Einstellung der Filterparameter wird im folgenden beispielhaft erläutert. Im Gegensatz zu den Ausführungen in Kap. 7, wo eine Skalierung der ursprünglichen Sensorsignale mit dem Faktor 32 zur Erhöhung der Rechengenauigkeit für Fixed-Point-Operationen durchgeführt wurde, erfolgen alle Berechnungen in Verbindung mit der Kalman-Filterung mit Floating-Point-Arithmetik. Die in Kap. 7 verwendete Skalierung ist in diesem Fall nicht erforderlich. Beibehalten wird allerdings die blockweise Mittelwertbildung über vier aufeinanderfolgende Abtastwerte, so dass bei einer Abtastrate von 10 ms nach jeweils 40 ms ein neuer Datensatz zur Signalverarbeitung ansteht.

10.2.1 Kovarianz des Ausgangsrauschprozesses

Die zur Ermittlung der Kovarianz R_4 des Ausgangsrauschprozesses $v(k)$ verwendeten Messdaten wurden indoor bei nahezu konstanten Umgebungsbedingungen und ohne Personen im Erfassungsfeld des Bewegungsmelders aufgezeichnet. Betrachtet man die Standardabweichung dieser Messwerte in einem Zeitfenster mit jeweils 100 Datensätzen, ergeben sich Varianzen in der Größenordnung von 1.0 ... 20.0. Abb. 10.1 zeigt die Varianzen eines typischen Signalausschnitts der drei Sensorkanäle PIR1, PIR2 und PIR3.

Tab. 10.1 enthält eine Zusammenfassung der minimalen und maximalen Signalwerte der Messreihe. Zusätzlich angegeben sind die minimalen und maximalen Mittelwerte und Varianzen in einem Zeitfenster mit jeweils 100 Datensätzen. Obwohl diese Messdaten aus einem nahezu idealen, temperaturstabilen Erfassungsfeld stammen, muss beachtet werden, dass auch im Indoorbereich geringfügige Driften, z. B. der Hintergrund- oder der Detektortemperatur, vorhanden sind. Die Varianz dieser Messwerte sollte daher nicht unmittelbar als Kovarianz R_4 des Ausgangsrauschprozesses $v(k)$ verwendet werden. Vielmehr berücksichtigt $v(k)$ entsprechend dem zugrundeliegenden Systemmodell nur die statistischen Schwankungen der Wäch-

tersignale, die insbesondere auf das Detektor- und Verstärkerrauschen im Analogpfad sowie auf das Quantisierungsrauschen des verwendeten AD-Wandlers zurückzuführen sind.

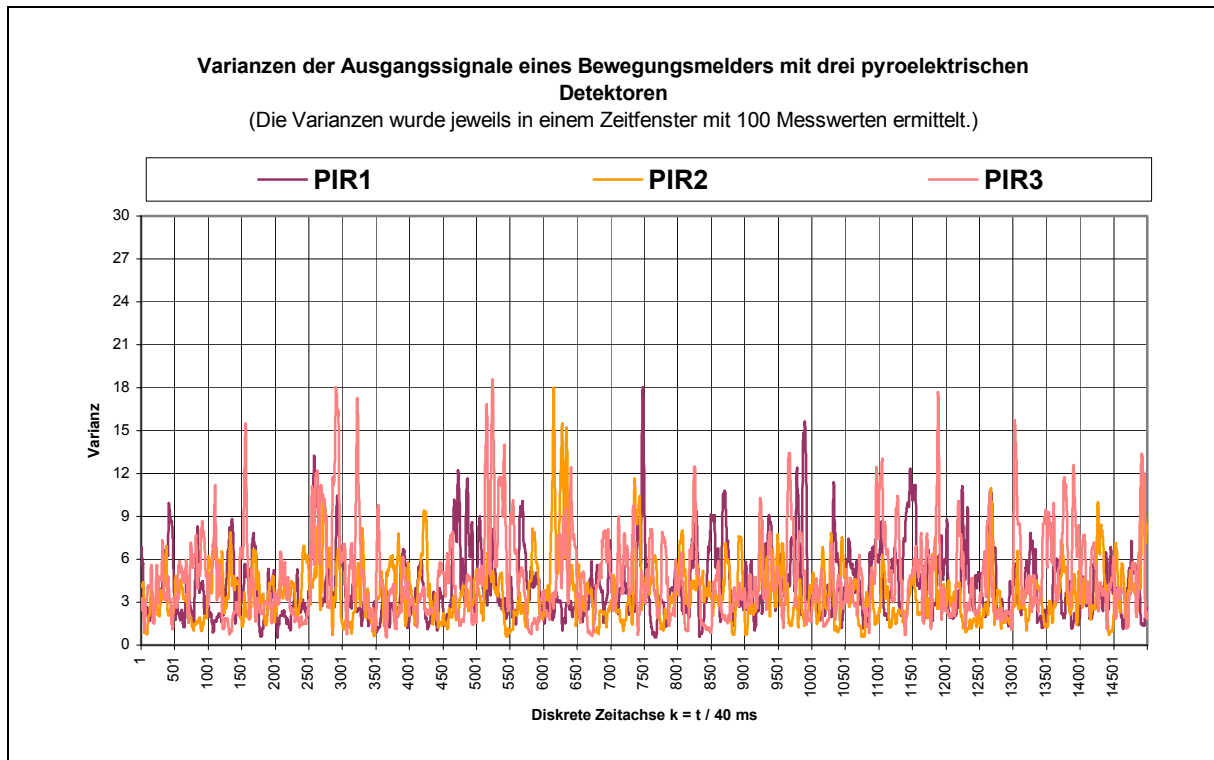


Abb. 10.1: Varianzen typischer Bewegungsmeldersignale aus dem Indoorbereich. Dargestellt sind die zeitlichen Verläufe der drei Sensorkanäle.

INDOOR-RAUSCHEN	PIR1	PIR2	PIR3
Min. Signalwert (Messreihe)	469.00	459.00	460.25
Max. Signalwert (Messreihe)	485.00	475.75	480.75
Min. Mittelwert (Zeitfenster))	475.88	466.58	469.88
Max. Mittelwert (Zeitfenster)	478.20	468.73	472.30
Min. Varianz (Zeitfenster)	1.84	1.22	1.10
Max. Varianz (Zeitfenster)	10.29	12.26	19.20

Tab. 10.1: Minimale und maximale Signalwerte sowie Mittelwerte und Varianzen der drei Sensorkanäle PIR1, PIR2, PIR3 in einem Zeitfenster von jeweils 100 Datensätzen.

Im vorliegenden Fall gestaltet sich die Diskriminierung zwischen dem statistischen Rauschanteil und den Signalschwankungen infolge geringfügiger Temperaturänderungen des Hintergrunds oder des Detektordielektrikums schwierig, so dass die Kovarianz R_4 im folgenden empirisch ermittelt wird. Zur Erläuterung der weiteren Vorgehensweise soll zunächst die Wirkung der Kovarianz R_4 des Ausgangsrauschprozesses $v(k)$ im Kalman-Filter betrachtet werden. Sie dient insbesondere zur Berechnung der Kalman-Gain-Matrix $K(k+1)$, so dass in Analogie zu (Gl. 9.45) folgt

$$(Gl. 10.1) \quad K(k+1) = P^-(k+1) \cdot C_4^T \cdot [C_4 \cdot P^-(k+1) \cdot C_4^T + R_4]^{-1} = P^-(k+1) \cdot C_4^T \cdot M.$$

Die Bildung der inversen Matrix $M = [\dots]^{-1}$ in (Gl. 10.1) reduziert sich im vorliegenden Fall auf die Berechnung einer skalaren Größe, die letztendlich als Skalierungsfaktor für die Kalman-Gain-Matrix dient. An dieser Stelle soll darauf hingewiesen werden, dass die einzelnen Elemente der Kovarianzmatrix $P^-(k+1)$ des Prädiktionsfehlers jeweils gegen einen konstanten Endwert streben, wenn keine adaptive Anpassung der Kovarianzmatrix Q_4 an die zugrundeliegende Hintergrunddynamik erfolgt.

In Verbindung mit der zeitinvarianten Beobachtungsmatrix C_4 und der ebenfalls als zeitinvariant angenommenen Kovarianz R_4 des Ausgangsrauschprozesses $v(k)$ streben dann auch die Elemente der Kalman-Gain-Matrix gegen konstante Endwerte. Die Kalman-Gain-Matrix $K(k+1)$ wird für jeden Abtastzeitpunkt neu berechnet. Sie bestimmt die Gewichtung der Residuensequenz $r(k+1)$ bei der Berechnung des optimalen Zustandsschätzwertes unter Beachtung des für den Zeitpunkt $k+1$ vorhergesagten Zustands. Nach (Gl. 8.47) gilt

$$(Gl. 10.2) \quad \hat{\underline{x}}^+(k+1) = \hat{\underline{x}}^-(k+1) + K(k+1) \cdot r(k+1)$$

Ausgehend von (Gl. 10.2) erkennt man unmittelbar, dass eine im Vergleich zum ersten Summand in (Gl. 10.1) große Kovarianz R_4 einen geringen Einfluss der Residuensequenz $r(k+1)$ auf die Zustandsschätzung bewirkt. Eine vergleichsweise große Kovarianz R_4 bedeutet, dass eine hohe Messunsicherheit infolge eines stark verrauschten Ausgangssignals vorliegt.

In diesem Fall gewichtet das Kalman-Filter seine Prädiktionswerte bei der Berechnung des optimalen Schätzwertes stärker als die Messwerte $y(k+1)$, die bei der Bildung der Residuensequenz $r(k+1)$ berücksichtigt werden. Je nach Größe der Kovarianz R_4 des Ausgangsrauschens $v(k)$ ergibt sich eine dauerhafte Abweichung zwischen dem gemessenen Signal $y(k)$ und dem in den Beobachtungsraum abgebildeten, optimalen Zustandsschätzwert, wie Abb. 10.2 zeigt.

In Abb. 10.2 ist die Nachführung des Ausgangssignals eines Bewegungsmelders bei unterschiedlichen Kovarianzen $R_4 = \text{Var}V = [0.0001, 0.001, 0.01, 0.1, 1.0]$ des Rauschprozesses $v(k)$ am Systemausgang dargestellt. Insbesondere bei den Varianzen $R_4 = \text{Var}V = [0.1, 1.0]$ ergeben sich deutliche Abweichungen zwischen dem gemessenen und dem geschätzten Ausgangssignal. Eine recht gute Übereinstimmung zwischen dem gemessenen und dem modellierten Systemverhalten zeigt sich für Varianzen $R_4 = \text{Var}V \leq 0.01$.

Für die weiteren Untersuchungen wird daher die vergleichsweise kleine Kovarianz $R_4 = 0.01$ verwendet, die eine starke Gewichtung der Residuensequenz $r(k+1)$ bei der Berechnung des optimalen Schätzwertes bewirkt und eine ausreichend gute Nachführung des Ausgangssignals gestattet. Abb. 10.3 zeigt die Auswirkung dieser Einstellung auf die Residuensequenz.

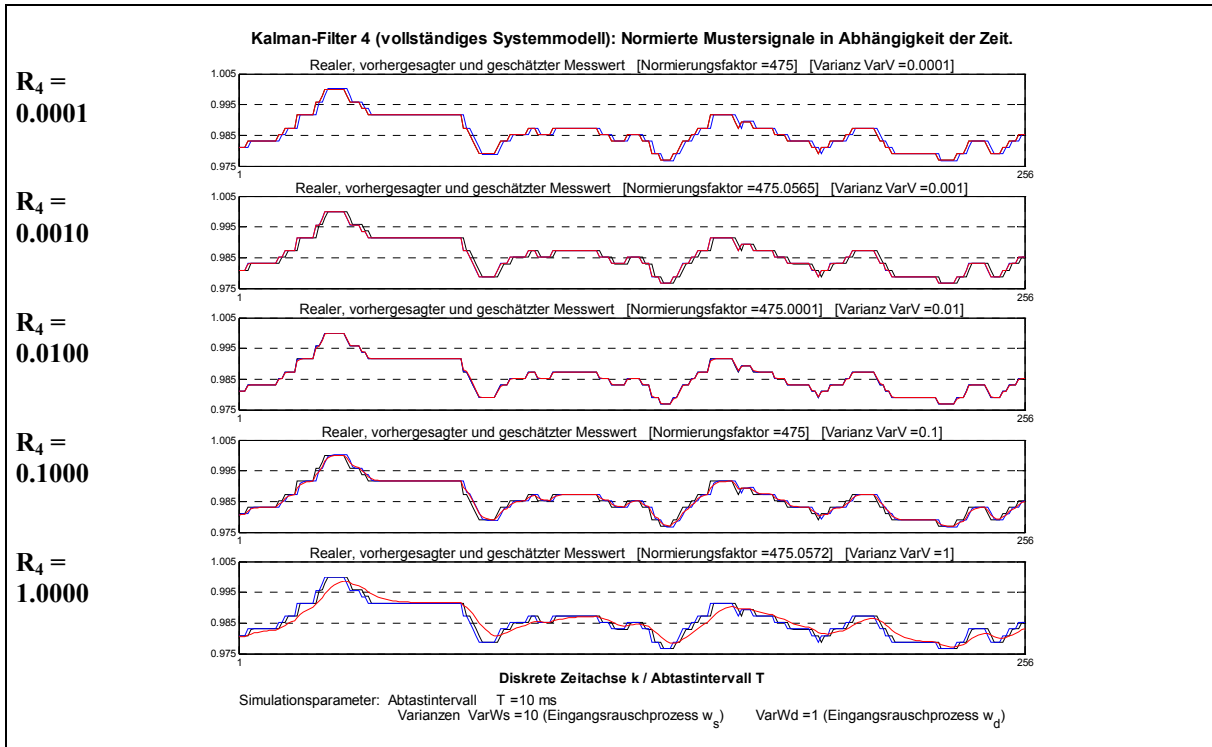


Abb. 10.2: Ausgangssignal eines Bewegungsmelders (schwarz) und in den Beobachtungsraum abgebildete Mittelwerte der Prädiktionsdichte (blau) und der Filterdichte (rot) bei unterschiedlichen Varianzen $R_4 = \text{VarV}$ des Ausgangsrauschens.

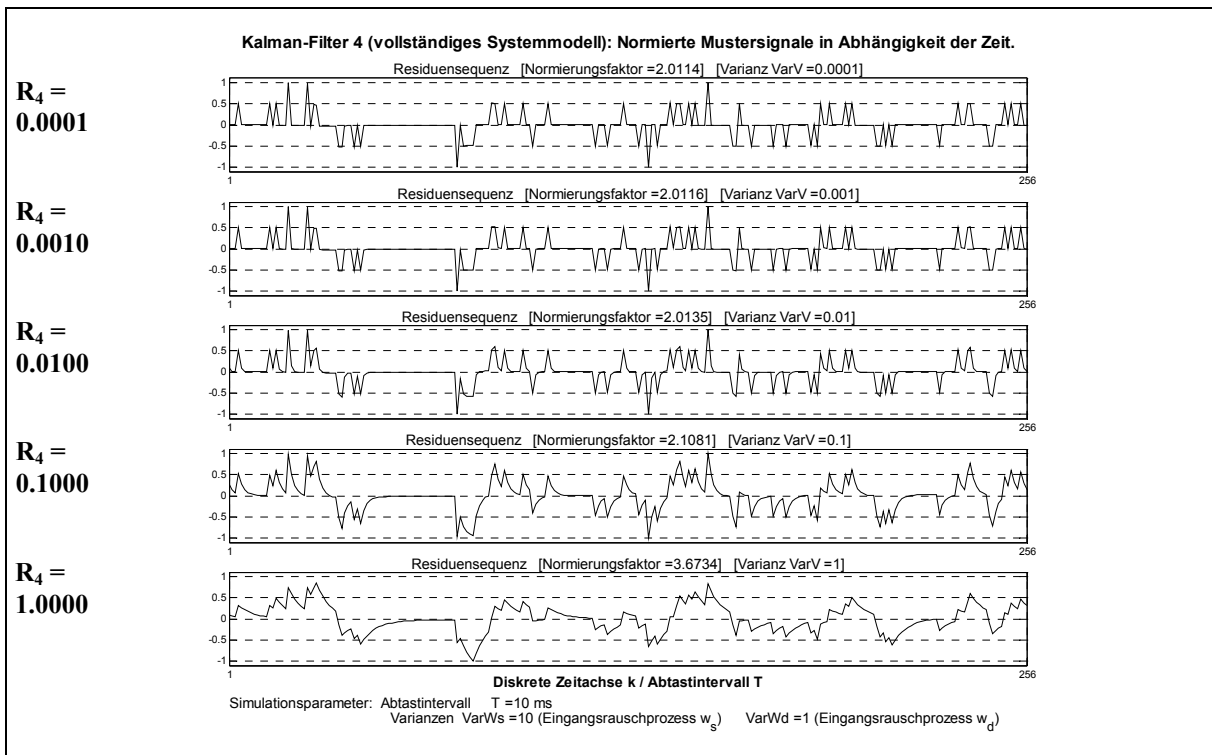


Abb. 10.3: Residuensequenzen bei der Nachführung des Ausgangssignals nach Abb. 10.2 in Abhängigkeit der Kovarianz $R_4 = \text{VarV}$.

Anhand der in Abb. 10.3 angegebenen Maximalwerte bei unterschiedlichen Kovarianzen R_4 und am zeitlichen Verlauf der einzelnen Residuensequenzen erkennt man, dass die in den

beiden unteren Diagrammen vorhandenen Überlagerungen für $R_4 = \text{Var}V \leq 0.01$ nahezu verschwunden sind. Da die Residuensequenz aber gerade die "unerwarteten" Signaländerungen charakterisiert, ist deren Minimierung für das vorliegende, nahezu ideale Testfeld eine wichtige Voraussetzung für ein zufriedenstellendes Detektionsverhalten.

10.2.2 Kovarianzmatrix des Eingangsrauschprozesses

Wie bereits in Kap. 9 erläutert wurde, wird die Kovarianzmatrix Q_4 des Eingangsrauschprozesses $w(k)$ bei der Berechnung der Prädiktionsfehlerkovarianz $P^-(k+1)$ berücksichtigt. In Analogie zu (Gl. 9.43) folgt für das Kalman-Filter 4 die Beziehung

$$(Gl. 10.3) \quad P^-(k+1) = A_4 \cdot P^+(k) \cdot A_4^T + Q_4,$$

wobei die zeitdiskrete Systemmatrix A_4 nach (Gl. 9.77) gewählt wird. Die Kovarianzmatrix Q_4 ist durch (Gl. 9.78) – (Gl. 9.84) definiert. In (Gl. 9.79) – (Gl.9.81) wird dabei die Varianz n_w zur Einstellung des allgemeinen Signalmodells verwendet, während in (Gl. 9.82) – (Gl. 9.84) die Varianz N_w zur Anpassung des Driftmodells an die Eigenschaften der Störsignale dient.

10.2.2.1 Störmodell ohne Objekt im Erfassungsfeld

Wenn kein Objekt im Erfassungsfeld des Bewegungsmelders erkannt wurde (Hypothese H_0), wird die gesamte Varianz $(\sigma_{yy})^2$ des beobachteten Ausgangssignals $y(k) = d(k) + v(k)$ dem Störmodell zugerechnet. Letzteres basiert auf dem allgemeinen Bewegungsmeldermodell nach Kap. 9 und wird in Kalman-Filter 4 sinngemäß verwendet. Setzt man voraus, dass der Summand $d(k)$ ganz allgemein den Anteil der Störungen im beobachteten Ausgangssignal $y(k)$ repräsentiert, während $v(k)$ den am Systemausgang überlagerten, skalaren Ausgangsrauschprozess beschreibt, so gilt unter der Annahme der Unkorreliertheit beider Störprozesse für die Varianzen die Beziehung

$$(Gl. 10.4) \quad \sigma_{yy}^2 = \sigma_{dd}^2 + \sigma_{vv}^2 \Rightarrow \sigma_{dd}^2 = \sigma_{yy}^2 - \sigma_{vv}^2 \quad \text{mit} \quad \sigma_{vv}^2 = R_4$$

Unter Beachtung von (Gl. 9.51) folgt in diesem Fall für die Varianz N_w des Eingangsrauschprozesses die Beziehung

$$(Gl. 10.5) \quad N_w|_{H_0} = \frac{2 \cdot (\alpha + \beta)}{K_s^2} \cdot \left(\sigma_{yy}^2|_{H_0} - \sigma_{vv}^2 \right) \quad \text{mit} \quad \alpha = \frac{1}{\tau_{th}}, \beta = \frac{1}{\tau_{el}}, K_s = 1.$$

In (Gl. 10.5) wird N_w aus der Varianz $(\sigma_{yy})^2$ des Ausgangssignals $y(k)$ unter Beachtung der Varianz $R_4 = (\sigma_{vv})^2$ des Ausgangsrauschprozesses $v(k)$ und der Bandpasscharakteristik des pyroelektrischen Detektors nach (Gl. 9.7) berechnet. In Analogie zu den Betrachtungen in Abschnitt 10.2.1 wird die Varianz $(\sigma_{yy})^2$ des Ausgangssignals $y(k)$ in einem Zeitfenster mit $N = 100$ aufeinanderfolgenden Datensätzen ermittelt. Bei einer Abtastfrequenz von 100 Hz ergibt sich demzufolge ein Zeitfenster von 4 s für die Varianzberechnung, da jeder Datensatz

wiederum als Mittelwert von vier aufeinanderfolgenden Abtastwerten gebildet wird. Zur Bestimmung des Mittelwertes m_y innerhalb des aktuellen Zeitfensters wird (Gl. 10.6) verwendet. (Gl. 10.7) dient zur Berechnung der Varianz $(\sigma_{yy})^2$ des beobachteten Ausgangssignals $y(k)$, woraus nach (Gl. 10.4) die Varianz $(\sigma_{dd})^2$ des Störanteils $d(k)$ folgt.

$$(Gl. 10.6) \quad m_y(k) = \frac{1}{N} \cdot \sum_{i=k-N+1}^k y(i) \quad \text{mit} \quad N = 100$$

$$(Gl. 10.7) \quad \sigma_{yy}^2(k) = \frac{1}{N-1} \cdot \sum_{i=k-N+1}^k (y(i) - m_y(k))^2 \quad \text{mit} \quad N = 100$$

Für eine erste Signalauswertung mit konstanter Varianz N_w des Störmodells (s. Abschnitt 10.3.1.1) wird im folgenden die Varianz $(\sigma_{yy})^2 = 20.0$ verwendet, die mit (Gl. 10.5) auf den Systemeingang umgerechnet werden kann. $(\sigma_{yy})^2$ wurde in der gleichen Größenordnung gewählt wie die maximale Varianz des Indoor-Rauschens nach Tab. 10.1.

In einem weiteren Schritt erfolgt dann eine adaptive Anpassung der Varianz N_w an die vorhandene Signaldynamik (s. Abschnitte 10.3.1.2 und 10.3.1.3). Zur Ermittlung der Varianz $(\sigma_{yy})^2$ der Messwerte wird wiederum ein Zeitfenster mit 100 Datensätzen gewählt. Bei dieser Vorgehensweise ist es sinnvoll, den zulässigen Variationsbereich von $(\sigma_{dd})^2$, der wiederum anhand von (Gl. 10.5) auf den Systemeingang bezogen wird, auf einen Maximalwert zu begrenzen. Eine sinnvolle obere Grenze kann anhand von Indoor- und Outdoormessungen experimentell bestimmt werden.

10.2.2.2 Allgemeines Signalmodell ohne Objekt im Erfassungsfeld

Ohne bewegtes Objekt im Erfassungsfeld sollen alle Abweichungen des beobachteten Ausgangssignals $y(k)$ vom Ruhepegel als Störungen interpretiert werden. Entsprechend den Ausführungen in Kap. 9 ist es daher in der Hypothese H_0 erforderlich, die Varianz $(\sigma_{ss})^2$ des allgemeinen Signalmodells im Vergleich zur Varianz $(\sigma_{dd})^2$ des Störmodells auf einen sehr kleinen Wert zu setzen.

Im weiteren Verlauf wird während der Hypothese H_0 die Varianz $(\sigma_{ss})^2$ nach (Gl 10.8) verwendet, die unter Beachtung von (Gl. 9.64) in die Varianz n_w am Systemeingang umgerechnet werden kann (s. (Gl. 10.9)). Die Varianz $(\sigma_{ss})^2$ ist kleiner als die minimale Varianz $(\sigma_{yy})^2$ des Indoor-Rauschens (s. Tab. 10.1), so dass sich Störungen entsprechend den Ausführungen in Kap. 9 insbesondere auf die offsetfreie, 4. Komponente des vorhergesagten Zustandsvektors auswirken.

$$(Gl. 10.8) \quad \sigma_{ss}^2 \Big|_{H_0} = 0.1$$

$$(Gl. 10.9) \quad n_w \Big|_{H_0} = 2 \gamma \cdot \sigma_{ss}^2 \quad \text{mit} \quad \gamma = \frac{1}{s}$$

10.2.2.3 Allgemeines Signalmodell mit Objekt im Erfassungsfeld

In (Gl. 8.26) und (Gl. 8.27) wurde die Zusammensetzung des Ausgangssignals $y(k)$ bei Abwesenheit (Hypothese H_0) und bei Anwesenheit (Hypothese H_1) eines Objekts im Erfassungsfeld des Bewegungsmelders beschrieben. Zusätzlich zu den Störanteilen $d(k)$ und $v(k)$ tritt in der Hypothese H_1 ein Objektsignal $s(k)$ im Ausgangssignal $y(k)$ auf, das den Rauschsignalen additiv überlagert ist. In diesem Fall vergrößert sich die Varianz $(\sigma_{yy})^2$ im Vergleich zur Hypothese H_0 . An dieser Stelle sind zwei unterschiedliche Vorgehensweisen denkbar. Im einfachsten Fall wird die Erhöhung der Varianz $(\sigma_{yy})^2$ im Vergleich zur Hypothese H_0 nicht beachtet (s. Abschnitt 10.3.1.1). Die höhere Signaldynamik im Vergleich zum Hintergrund zeigt entsprechende Auswirkungen auf die Residuensequenz. Eine Erhöhung der Störvarianz $(\sigma_{dd})^2$, die in der Hypothese H_0 durchgeführt wird, ist bei einem Detektionsereignis nicht unbedingt erforderlich. Verlässt die Person das Erfassungsfeld, reduziert sich die resultierende Varianz $(\sigma_{yy})^2$ im Idealfall wieder auf den Wert vor dem Objekteintritt.

Als zweite Möglichkeit kann die Erhöhung der Varianz $(\sigma_{yy})^2$ im Vergleich zur Hypothese H_0 durch eine Vergrößerung der Varianz $(\sigma_{ss})^2$ des allgemeinen Signalmodells berücksichtigt werden. Bei zunächst unkorrelierten Signalen $s(k)$, $d(k)$ und $v(k)$, die sich zum beobachteten Ausgangssignal $y(k) = s(k) + d(k) + v(k)$ überlagern, ergibt sich die Varianz $(\sigma_{yy})^2$ aus der Summe der einzelnen Varianzen

$$(Gl. 10.10) \quad \sigma_{yy}^2 = \sigma_{ss}^2 + \sigma_{dd}^2 + \sigma_{vv}^2 \Rightarrow \sigma_{ss}^2 = \sigma_{yy}^2 - \sigma_{dd}^2 - \sigma_{vv}^2 \quad \text{mit} \quad \sigma_{vv}^2 = R_4$$

In (Gl. 10.10) ist die Varianz $(\sigma_{vv})^2 = R_4$ bekannt (s. Abschnitt 10.2.1). Aus Abschnitt 10.2.2.1 folgt außerdem die Varianz $(\sigma_{dd})^2$ während der Hypothese H_0 , die während der Hypothese H_1 als unveränderlich angenommen wird. Für die Varianz N_w des Störmodells während der Hypothese H_1 gilt daher nach (Gl. 10.5) unverändert

$$(Gl. 10.11) \quad N_w \Big|_{H_1} = \frac{2 \cdot (\alpha + \beta)}{K_s^2} \cdot \left(\sigma_{yy}^2 \Big|_{H_0} - \sigma_{vv}^2 \right) = \frac{2 \cdot (\alpha + \beta)}{K_s^2} \cdot \sigma_{dd}^2 \Big|_{H_0}$$

Bei unkorreliertem Objektsignal $s(k)$ und Störsignal $d(k)$ folgt für die Varianz $(\sigma_{ss})^2$ während der Hypothese H_1 die Beziehung

$$(Gl. 10.12) \quad \sigma_{ss}^2 \Big|_{H_1} = \sigma_{yy}^2 \Big|_{H_1} - \sigma_{dd}^2 \Big|_{H_0} - \sigma_{vv}^2 \quad \text{mit} \quad \sigma_{vv}^2 = R_4,$$

welche zur Berechnung der Varianz n_w am Systemeingang infolge eines bewegten Objekts im Erfassungsfeld des Bewegungsmelders verwendet werden kann. Aus (Gl. 10.12) erhält man unter Beachtung von (Gl. 9.64) folgenden Zusammenhang für die resultierende Varianz:

$$(Gl. 10.13) \quad n_w \Big|_{H_1} = 2 \gamma \cdot \sigma_{ss}^2 \Big|_{H_1} = 2 \gamma \cdot \left(\sigma_{yy}^2 \Big|_{H_1} - \sigma_{dd}^2 \Big|_{H_0} - \sigma_{vv}^2 \right) \quad \text{mit} \quad \sigma_{vv}^2 = R_4 \quad \text{und} \quad \gamma = \frac{1}{s}$$

Bei korrelierten Signalanteilen $s(k)$ und $d(k)$ ist eine eindeutige Trennung der Varianzen der beteiligten Prozesse nicht möglich. Ein denkbarer Ansatz, der bisher lediglich experimentell verifiziert wurde, ist die Beibehaltung der Varianz N_w des Störmodells aus der Hypothese H_0 , während in der Hypothese H_1 die gesamte Varianz $(\sigma_{yy})^2 - (\sigma_{vv})^2$ zur Ermittlung der Varianz n_w am Eingang des allgemeinen Signalmodells verwendet wird (s. Abschnitte 10.3.1.2 und 10.3.1.3). In diesem Fall gilt

$$(Gl. 10.14) n_w|_{H_1} = 2\gamma \cdot \left(\sigma_{ss}^2|_{H_1} + \sigma_{dd}^2|_{H_1} \right) = 2\gamma \cdot \left(\sigma_{yy}^2|_{H_1} - \sigma_{vv}^2 \right) \text{ mit } \sigma_{vv}^2 = R_4 \text{ und } \gamma = \frac{1}{s}.$$

Durch die zusätzliche Verwendung der Varianz $(\sigma_{dd})^2$ im Rahmen des allgemeinen Signalmodells und des Störmodells, welches auf dem einfachen Bewegungsmeldermodell basiert, arbeitet das Kalman-Filter intern mit einer erhöhten Signaldynamik. Entsprechend den Ausführungen in Kap. 9 bewirkt das eine Verringerung der resultierenden Residuensequenz, so dass deren Auswertung während oder unmittelbar im Anschluss an die Hypothese H_1 insgesamt unempfindlicher wird. Bei der Auswertung der 1. und 2. Komponente des prädizierten Zustandsvektors sind die Amplituden der auszuwertenden Signale dagegen größer, so dass sich bei deren Bewertung in Verbindung mit korrelierten Signalanteilen eine etwas höhere Empfindlichkeit ergibt. Im nächsten Abschnitt werden die drei Auswerteverfahren näher erläutert.

10.3 Bewegungserkennung anhand von internen Größen des Kalman-Filters

10.3.1.1 Schwellenauswertung der Residuensequenz (Verfahren 1)

Ausgehend von der vorher beschriebenen Parametereinstellung des Kalman-Filters 4 soll in diesem Abschnitt die Bewegungserkennung anhand der Residuensequenz beschrieben werden. Nach den Ausführungen in Kap. 9 wird die Residuensequenz $r(k)$ aus der Differenz des beobachteten Ausgangssignals $y(k)$ und dem in den Beobachtungsraum abgebildeten Prädiktionwert (Zustandsvektor) gebildet. In Analogie zu (Gl. 9.44) erhält man die Residuensequenz des Kalman-Filters 4 aus

$$(Gl. 10.15) \quad r(k+1) = y(k+1) - C_4 \cdot \hat{\underline{x}}^-(k+1).$$

Unerwartete Signaländerungen wirken sich in der Residuensequenz aus, die in diesem Fall charakteristische Eigenschaften verliert. Bei einem optimal eingestellten Kalman-Filter ist die Residuensequenz eine nahezu mittelwertfreie, weiße Rauschsequenz, was im vorliegenden Anwendungsfall und ohne bewegtes Objekt (Hypothese H_0) im Erfassungsfeld des Bewegungsmelders durchaus erfüllt ist (s. auch Abb. 10.3).

Bewegt sich eine Person im Erfassungsfeld des Bewegungsmelders (Hypothese H_1), bewirkt das einen Ausschlag in der Residuensequenz $r(k+1)$. Eine einfache Schwellenauswertung liefert daher in diesem Fall ein erstes Kriterium zur Diskriminierung von Nutz- und Störsignalen, basierend auf dem verwendeten Systemmodell. Anhand des in Abb. 10.4 dargestellten

“Draufzugehnsignals“ soll die weitere Vorgehensweise erläutert werden. Die Abbildung verdeutlicht außerdem nochmals die Nachführung des Sensorsignals PIR durch die in den Beobachtungsraum abgebildeten Mittelwerte PRE und EST der Prädiktions- und Filterdichte.

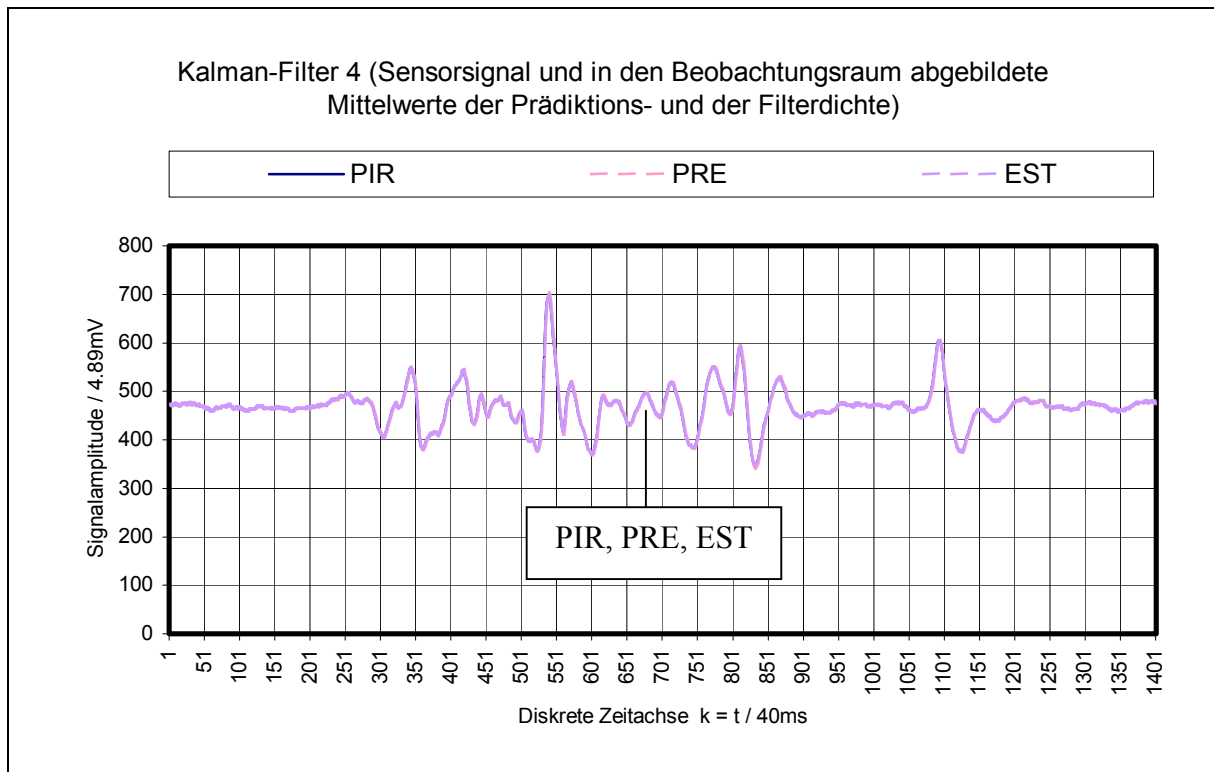


Abb. 10.4: Sensorsignal PIR (“Draufzugehnsignal“) und in den Beobachtungsraum abgebildete Mittelwerte PRE und EST der Prädiktions- und der Filterdichte.

Die Berechnung der Auswerteschwellen TOP(k) und BOT(k) erfolgt analog zur Vorgehensweise in Kap. 7. Zunächst wird der gleitende Mittelwert REF(k) nach (Gl. 10.16) gebildet. Anstelle des Sensorsignals PIR(k) wird diesmal die Residuensequenz $r(k) = \text{RSS}(k)$ verwendet, um zunächst den Bezugspegel REF(k) zu bestimmen. Durch den Nenner REFADJ = 2^{10} wird eine geringe Adaptionsgeschwindigkeit des Referenzpegels REF(k) eingestellt.

$$\text{(Gl. 10.16)} \quad \text{REF}(k) = \text{REF}(k-1) + \frac{\text{RSS}(k) - \text{REF}(k-1)}{\text{REFADJ}} \quad \text{mit} \quad \text{REFADJ} = 2^{10}$$

Anschließend wird die absolute Differenz $\text{ABS}(k) = |\text{RSS}(k) - \text{REF}(k)|$ gebildet, woraus sich der ebenfalls zeitadaptive Offset LEV(k) nach (Gl. 10.17) ergibt. Dieser dient zur Berechnung der Auswerteschwellen TOP(k) nach (Gl. 10.18) und BOT(k) nach (Gl. 10.19). Die Festlegung der zum Referenzwert REF(k) symmetrischen Amplitudenschwellen erfolgt wiederum unter Beachtung des minimal erforderlichen Offsets THRMIN und des maximalen Schwellenwerts THRMAX. Abb. 10.5 zeigt den Verlauf der Residuensequenz für den Signalabschnitt aus Abb. 10.4 in Verbindung mit der Schwellenauswertung, wobei zur Darstellung THRMIN = 4 und THRMAX = 10 gewählt wurde.

$$(Gl. 10.17) \quad LEV(k) = LEV(k-1) + \frac{ABS(k) - LEV(k-1)}{LEVADJ} \quad \text{mit} \quad LEVADJ = 2^5$$

$$(Gl. 10.18) \quad BOT(k) = REF(k) - \text{Minimum}[(THRMIN + LEV(k)), (THRMAX)]$$

$$(Gl. 10.19) \quad TOP(k) = REF(k) + \text{Minimum}[(THRMIN + LEV(k)), (THRMAX)]$$

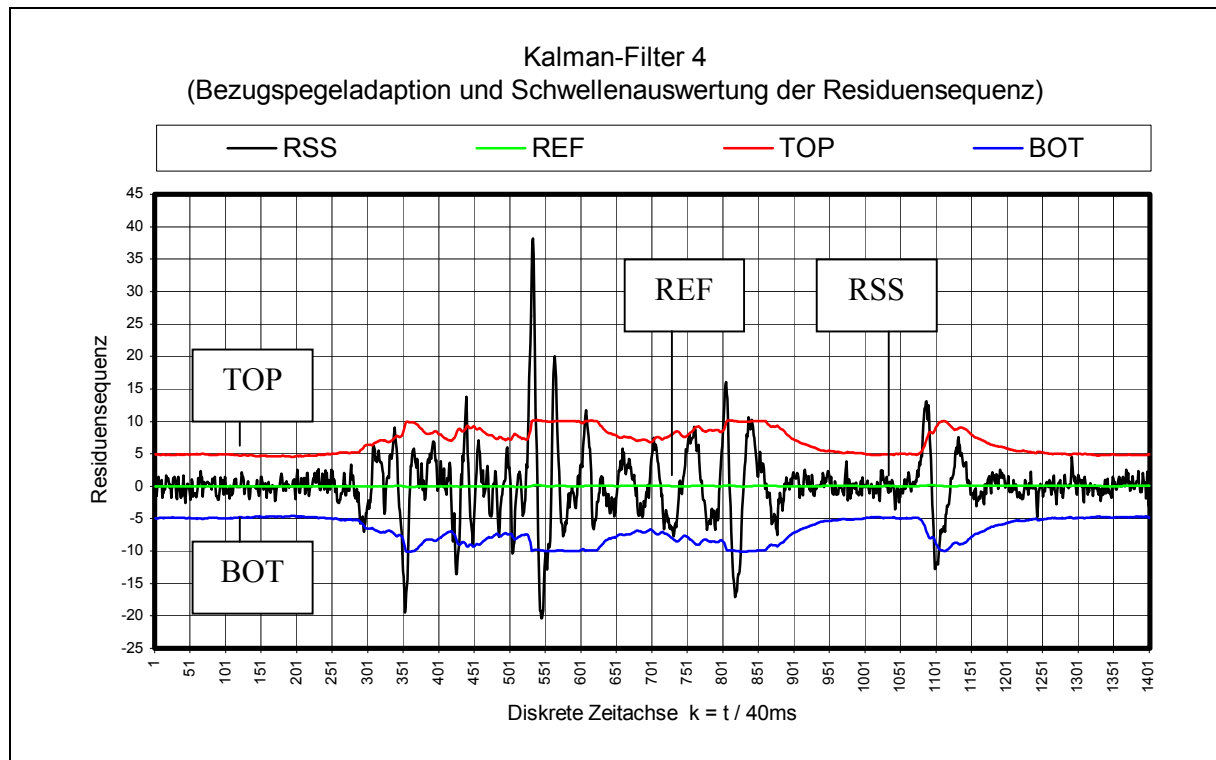


Abb. 10.5: Schwellenauswertung der Residuensequenz bei Auftreten eines Nutzsignals infolge eines bewegten Objekts in Verbindung mit den Varianzen $(\sigma_{ss})^2 = 0.1$, $(\sigma_{dd})^2 = 20.0$ und $R_4 = 0.01$.

Liegt die Residuensequenz $RSS(k)$ außerhalb des durch die Schwellen $BOT(k)$ und $TOP(k)$ vorgegebenen Toleranzbandes, erfolgt die Alarmauslösung. In Kap. 10.4.1 wird das Schaltverhalten des Bewegungsmelders bei Anwendung dieses Verfahrens im Vergleich zu der in Kap. 7 vorgestellten Signalverarbeitungsstrategie zusammengefasst. Dort wird gezeigt, dass die einfache Schwellenauswertung der Residuensequenz schlechtere Ergebnisse im Hinblick auf das Schaltverhalten des Wächters liefert als das Auswerteverfahren nach Kap. 7.

Eine Verringerung der als konstant angenommenen Varianz der Hintergrundstörungen bewirkt zwar grundsätzlich eine höhere Schaltpunktzahl, allerdings war ein zu der Auswertung nach Kap. 7 vergleichbares Schaltverhalten auch mit dieser Maßnahme nicht realisierbar.

Senkt man bei konstanter Varianz dagegen die minimalen und maximalen Schwellwerte ab, erhöht sich die Anzahl der Fehlalarme im Outdoorbereich deutlich. Die Varianz N_w des Störmodells und die minimalen und maximalen Schwellwertgrenzen können nicht unabhängig voneinander eingestellt werden.

An dieser Stelle soll angemerkt werden, dass das implementierte Kalman-Filter vorwiegend langsame Störungen, die im Outdoorbereich beispielsweise durch Temperaturvariationen des Hintergrunds hervorgerufen werden, unterdrückt. Schnelle Störsignale durch elektromagnetische Einstrahlungen erfordern weitere Maßnahmen, wie beispielsweise die zusätzliche Gradientenprüfung nach Kap. 7.

Eine weitere Alternative zur Unterdrückung steiflankiger Gradienten ist deren Berücksichtigung im Rahmen des Systemmodells, wodurch der Berechnungsaufwand allerdings ansteigt. Im weiteren Verlauf wird daher die zuerst beschriebene Methode angewendet.

10.3.1.2 Auswertung der 1. Komponente der Zustandsprädiktion (Verfahren 2)

Aufgrund der in Abschnitt 10.3.1.1 beschriebenen Ergebnisse bei der Auswertung der Residuensequenz wurden weitere Untersuchungen zur Verbesserung des Schaltverhaltens im Vergleich zur Auswertung nach Kap. 7 durchgeführt. Aufgrund der Ausführungen in Kap. 9 ist es naheliegend, einzelne Zustandsvektorkomponenten der Prädiktion zur Bewertung zu verwenden.

Abb. 10.6 zeigt beispielsweise die 1. und die 4. Komponente des vorhergesagten Zustandsvektors für den Signalabschnitt nach Abb. 10.4. Die 1. Komponente entspricht dabei dem prädizierten Nutzsignal PDM(1,1), das im Vergleich zum Originalsignal PIR um den in der 4. Zustandsvektorkomponente PDM(4,1) enthaltenen Störanteil bereinigt ist.

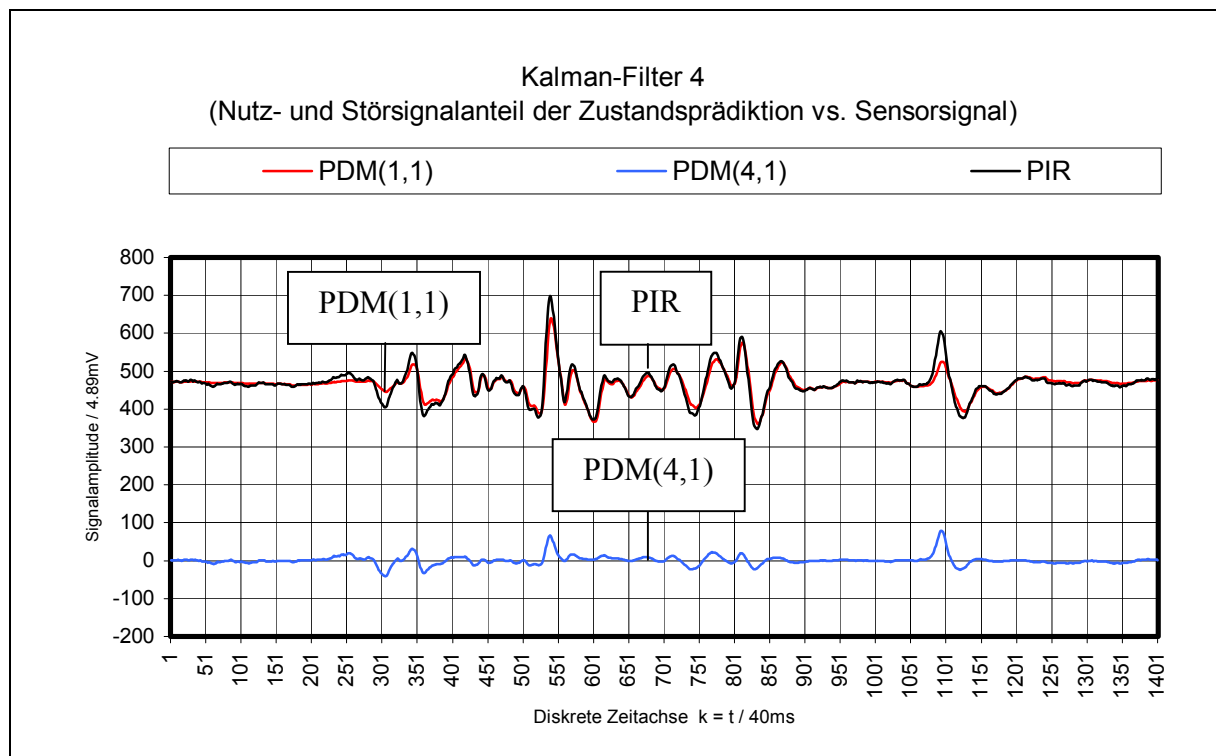


Abb. 10.6: Auswirkung des Störmodells auf die Zustandsprädiktion. 1. Komponente PDM(1,1) des vorhergesagten Zustandsvektors für die Varianzen $(\sigma_{ss})^2 = 0.1$, $(\sigma_{dd})^2 = 20.0$ und $R_4 = 0.01$. Die 4. Komponente PDM(4,1) des Zustandsvektors entspricht dem berechneten Störanteil.

In diesem Zusammenhang soll darauf aufmerksam gemacht werden, dass gerade diese beiden Komponenten bei der in (Gl. 10.15) vorgenommenen Abbildung in den Beobachtungsraum überlagert werden. Die resultierende Summe ergibt das vorhergesagte Gesamtsignal, welches anschließend zur Berechnung der Residuensequenz dient.

Die Wirkung der Störungsunterdrückung ist in Abb. 10.6 deutlich zu erkennen. Betrachtet man beispielsweise den Zeitabschnitt von $k = 201 \dots 251$, erkennt man einen Anstieg des Originalsignals $PIR(k)$. Bei der gewählten Parametereinstellung des Störmodells wird dieser Bereich nahezu vollständig als Störsignal interpretiert. In der 4. Zustandsvektorkomponente $PDM(4,1)$ zeigt sich daher ebenfalls ein deutlicher Signalanstieg, während dieser Bereich im Nutzsinal $PDM(1,1)$ stark bedämpft ist.

Nach Abb. 10.6 wirkt die Störunterdrückung somit insbesondere auf die minimale und maximale Signalamplitude und weniger auf die zeitbezogenen Größen wie die Flankendauer und die Signaldauer. Die in Kap. 7 verwendete Mindestamplitude AMP_MIN kann daher in Verbindung mit den Kalman-gefilterten Messdaten reduziert werden, während die zeitbezogenen Parameter der konventionellen Signalauswertung unverändert erhalten bleiben.

Zur abschließenden Bewegungserkennung wird der Nutzsinalanteil $PDM(1,1)$ mit dem Auswerteverfahren nach Kap. 7 ausgewertet. Aufgrund der Störungsunterdrückung durch das Kalman-Filter können die minimalen und maximalen Grenzen der Schwellenauswertung abgesenkt werden. Empirisch ermittelt wurden die Grenzwerte $THRMIN = 3.00$ und $THRMAX = 4.00$, wodurch sich in Verbindung mit einer hohen Adaptionsgeschwindigkeit des Bezugspegels ($REFADJ = 2^2$) ein relativ enges Toleranzband zur Signalauswertung ergibt. Die Konstante $LEVADJ = 2^5$ wurde nicht verändert. Die Adaptionsgeschwindigkeiten der Referenzpegel- und der Schwellwertadaptation zur Auswertung der 1. Komponente des vorhergesagten Zustandsvektors sind somit identisch zu dem in Kap. 7 beschriebenen Verfahren.

Anhand der vorstehenden Ausführungen ist bereits ersichtlich, dass auch in Verbindung mit der Signalvorverarbeitung im Kalman-Filter eine bestimmte Signalunruhe erforderlich ist, um einen Schaltvorgang auszulösen. Der leichte Signalanstieg des Signals $PIR(k)$ im vorgenannten Zeitabschnitt, der für das "Draufzugehen" in der äußeren Entfernungszone charakteristisch ist, wird vom Kalman-Filter unterdrückt. Es muss daher auch bei dieser Form der Signalauswertung ein geeigneter Kompromiss zwischen der Störanfälligkeit und der maximalen Empfindlichkeit für Nutzsignale gefunden werden.

Im Vorgriff auf das in Abschnitt 10.4.3 beschriebene Schaltverhalten bei Auswertung des Nutzsinalanteils $PDM(1,1)$ soll allerdings angemerkt werden, dass sich im Vergleich zur Auswertung nach Kap. 7 ein mindestens vergleichbares Schaltverhalten für die Nutzsignale bei gleichzeitig erhöhter Fehlalarmimmunität ergibt.

10.3.1.3 Auswertung der 2. Komponente der Zustandsprädiktion (Verfahren 3)

Eine weitere Verbesserung des Schaltverhaltens lässt sich erzielen, wenn anstelle des Nutzsignalanteils PDM(1,1) die zweite Zustandsvektorkomponente PDM(2,1) zur Bewegungserkennung verwendet wird. PDM(2,1) ist gleichanteilsfrei und nach (Gl. 9.69) und (Gl. 9.71) näherungsweise proportional zur Änderung der Nutzsignalanteils PDM(1,1) innerhalb des Zeitabschnitts T. Der Proportionalitätsfaktor T ergibt sich aus der in (Gl. 9.71) gegebenen Näherungslösung für kleine Abtastintervalle.

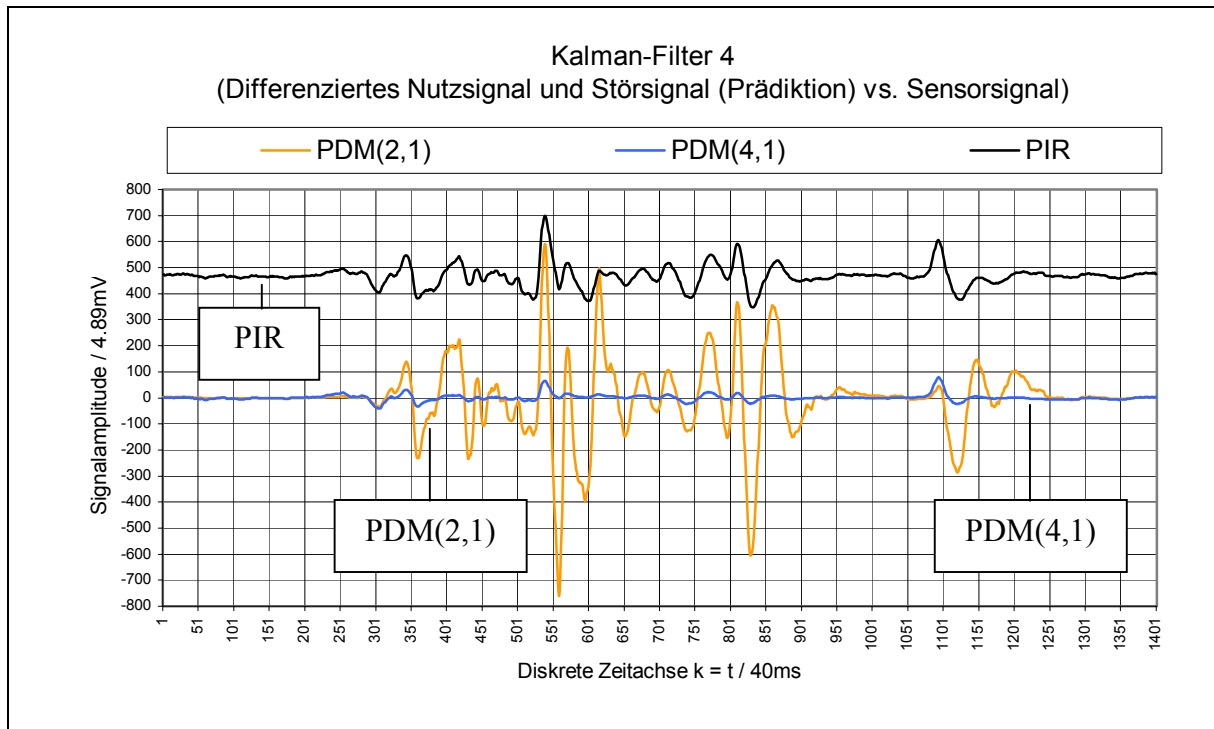


Abb. 10.7: Auswirkung des Störmodells auf die Zustandsprädiktion. 2. Komponente PDM(2,1) des vorhergesagten Zustandsvektors für die Varianzen $(\sigma_{ss})^2 = 0.1$, $(\sigma_{dd})^2 = 20.0$ und $R_4 = 0.01$. Die 4. Komponente PDM(4,1) des Zustandsvektors entspricht dem Störanteil in Abb. 10.6.

Abb. 10.7 zeigt den Verlauf der Zustandsvektorkomponente PDM(2,1) für den Signalabschnitt nach Abb. 10.4 bzw. Abb. 10.6. An dieser Stelle sollte die unterschiedliche Achsenkalibrierung der Abszissen von Abb. 10.6 und Abb. 10.7 beachtet werden.

Betrachtet man den Verlauf der 2. Zustandsvektorkomponente PDM(2,1) der Prädiktion, fällt auf, dass Abschnitte geringer Steigung des Sensorsignals PIR(k) im Vergleich zu den Abschnitten mit erhöhter Signalunruhe noch stärker unterdrückt werden (s. z. B. Signalabschnitt $k = 201 \dots 251$ vs. $k = 501 \dots 601$ in Abb. 10.7). Änderungen des Sensorsignals PIR(k), die nicht mit dem zugrundeliegenden Driftmodell übereinstimmen, verursachen deutlich stärkere Ausschläge in der Komponente PDM(2,1) als diejenigen, die ausgefiltert werden. In Abschnitt 10.4.4 sind die Ergebnisse bei der Bewertung der 2. Zustandsvektorkomponente der Prädiktion zusammengefasst.

10.4 Schaltverhalten der verschiedenen Auswerteverfahren

In diesem Abschnitt wird das Schaltverhalten eines Bewegungsmelders mit Kalman-Filter 4 im Vergleich zum Schaltverhalten des in Kapitel 7 vorgestellten Auswerteverfahrens betrachtet. Als Grundlage der Untersuchungen dienen die bereits vorher erwähnten Daten aus unterschiedlichen Testfeldern. Zur Einstellung der Grundempfindlichkeit werden zunächst typische Outdoor-Rauschsignale verwendet.

In Verbindung mit den indoor aufgezeichneten Nutzsignalen ergeben sich je nach Verfahren verschiedene Alarmanzahlen, die zusammen mit der Verteilung der Schaltpunkte im Erfassungsfeld zur Bewertung herangezogen werden können.

Qualitative Aussagen hinsichtlich der Störanfälligkeit lassen sich zusätzlich ableiten, wenn im Outdoor-Testfeld bei schwierigeren Witterungsverhältnissen aufgezeichnete Stördaten in diese Betrachtungen einbezogen werden.

10.4.1 Schaltverhalten bei konventioneller Signalauswertung

Für das Auswerteverfahren nach Kap. 7 wurden die bereits dort angegebenen Parametereinstellungen verwendet. Das gesamte Auswerteverfahren wurde allerdings auf Floating-Point-Arithmetik umgestellt.

Wie bereits vorher erwähnt wurde, wird die in Kap. 7 zur Erhöhung der Rechengenauigkeit für Fixed-Point-Arithmetik vorgenommene Skalierung der Rohdaten mit dem Faktor 32 hier nicht angewendet, so dass die amplitudenbezogenen Größen entsprechend umgerechnet werden müssen. In Tab. 10.2 sind die resultierenden Parametereinstellungen zusammengefasst.

PARAMETER-EINSTELLUNG	Kap. 7 (mit Skalierung)	Kap. 10 (ohne Skalierung)
MITTELUNGEN	4	4
THRMIN	188	5.88
THRMAX	220	6.88
AMPMIN	600	18.75
GRDMIN	129	4.03
GRDMAX	2500	78.13
SLPMIN	13	13.00
SLPMAX	47	47.00
WIDMIN	23	23.00
WIDMAX	80	80.00

Tab. 10.2: Vergleich der Parametereinstellung nach Kap. 7 und Kap. 10.

Tab. 10.3 enthält typische Signalcharakteristiken der für die Parametereinstellung des konventionellen Auswerteverfahrens verwendeten Outdoor-Störungen. Die Optimierung erfolgte derart, dass Fehlalarme vollständig unterdrückt werden.

In Tab. 10.4 sind Signalcharakteristiken typischer Rauschsignale aus dem Indoor-Bereich aufgeführt. Unterschiede im Vergleich zu Tab. 10.3 ergeben sich zunächst im minimalen und maximalen Signalwert sowie insbesondere bei der Varianz der zugrundeliegenden Signalverläufe innerhalb des verwendeten Zeitfensters. Da die Optimierung für die "unruhigeren" Sig-

nalverläufe im Outdoor-Bereich erfolgte, treten bei den Indoor-Signalen erwartungsgemäß keine Fehlalarme auf.

OUTDOOR-STÖRSIGNALE	PIR1	PIR2	PIR3
Min. Signalwert (Messreihe)	464.25	448.50	455.25
Max. Signalwert (Messreihe)	480.75	492.25	491.75
Min. Mittelwert (Zeitfenster)	470.87	468.84	471.29
Max. Mittelwert (Zeitfenster)	474.20	476.28	477.39
Min. Varianz (Zeitfenster)	0.62	0.65	0.92
Max. Varianz (Zeitfenster)	14.74	124.40	77.75
Fehlalarmzahl	0	0	0

Tab. 10.3: Outdoor-Störsignale in Verbindung mit dem Auswerteverfahren nach Kap. 7 (Parametereinstellung s. Tab. 10.2, Spalte 3). Minima und Maxima der gemittelten Sensorsignale PIR1, PIR2, PIR3, minimale und maximale Mittelwerte und Varianzen in einem Zeitfenster von 100 Datensätzen sowie Fehlalarmzahl.

INDOOR-RAUSCHEN	PIR1	PIR2	PIR3
Min. Signalwert (Messreihe)	469.00	459.00	460.25
Max. Signalwert (Messreihe)	485.00	475.75	480.75
Min. Mittelwert (Zeitfenster)	475.88	466.58	469.88
Max. Mittelwert (Zeitfenster)	478.28	468.73	472.30
Min. Varianz (Zeitfenster)	1.84	1.22	1.10
Max. Varianz (Zeitfenster)	10.29	12.26	19.20
Fehlalarmzahl	0	0	0

Tab. 10.4: Indoor-Rauschen in Verbindung mit dem Auswerteverfahren nach Kap. 7 (Parametereinstellung s. Tab. 10.2, Spalte 3). Minima und Maxima der gemittelten Sensorsignale PIR1, PIR2, PIR3, minimale und maximale Mittelwerte und Varianzen in einem Zeitfenster von 100 Datensätzen sowie Fehlalarmzahl.

Die Auswertung der Nutzsignale aus dem Indoorbereich, die in unterschiedlichen Bewegungsrichtungen aufgezeichnet wurden, liefert die in Tab. 10.5 angegebenen Alarmereignisse pro Sensorkanal. Anhand der angegebenen Signalcharakteristiken erkennt man die deutlich größere Signalunruhe im Vergleich zu den Störsignalen aus dem Indoor- und dem Outdoorbereich.

Die absolute Alarmanzahl in Tab. 10.3 - Tab. 10.5 wurde ermittelt unter Anwendung der in Kap. 7 beschriebenen Zerlegung der Sensorsignale PIR1, PIR2, PIR3 in einzelne Signalabschnitte, die jeweils auf die Existenz eines validen Detektionsereignisses geprüft werden. Die Signalzerlegung erfolgt dabei anhand der in Kap. 7 dargestellten Verfolgung von Minimal-

und Maximalwerten sowie anhand von Polaritätswechseln der Sensorsignale relativ zu den adaptiv nachgeführten Referenzpegeln.

INDOOR-NUTZSIGNALE	PIR1	PIR2	PIR3
Min. Signalwert (Messreihe)	222.25	93.50	123.25
Max. Signalwert (Messreihe)	785.25	720.50	745.50
Min. Mittelwert (Zeitfenster)	450.32	432.12	439.00
Max. Mittelwert (Zeitfenster)	503.63	506.80	509.90
Min. Varianz (Zeitfenster)	1.38	1.63	1.22
Max. Varianz (Zeitfenster)	15247.52	20713.47	13636.61
Alarmzahl	949	1075	738

Tab. 10.5: Indoor-Nutzsignale in Verbindung mit dem Auswerteverfahren nach Kap. 7 (Parametereinstellung s. Tab. 10.2, Spalte 3). Minima und Maxima der Sensorsignale PIR1, PIR2, PIR3, minimale und maximale Mittelwerte und Varianzen in einem Zeitfenster von 100 Datensätzen sowie Alarmanzahl.

Die absolute Summe der Alarme in Tab. 10.5 beträgt 2762. Infolge der ODER-Verknüpfung der Alarmereignisse auf den drei Sensorkanälen kann die resultierende Alarmzahl deutlich kleiner sein, da gleichzeitige Alarme auf verschiedenen Kanälen in diesem Fall nur einmal berücksichtigt werden. Im weiteren Verlauf werden daher die absoluten Summen bzw. die Zahl der Einzelalarme pro Sensorkanal zur Beurteilung des Verfahrens verwendet.

Wie bereits vorher beschrieben wurde, sind die zugrundeliegenden Nutzsignale in unterschiedlichen Bewegungsrichtungen aufgezeichnet worden. Entsprechend der Definition der verschiedenen Bewegungsrichtungen nach Abb. 6.9 erhält man die zugehörigen Alarmzahlen nach Tab. 10.6.

INDOOR-NUTZSIGNALE	PIR1	PIR2	PIR3
Parallele Bewegungsrichtung	251	290	204
Vertikale Bewegungsrichtung	193	219	145
Radiale Bewegungsrichtung	352	388	265
Tangentiale Bewegungsrichtung	153	178	124
Alle Richtungen (Summe)	949	1075	738

Tab. 10.6: Alarmzahl in Abhängigkeit des Sensorkanals und der Bewegungsrichtung.

Wählt man eine für den Einrichtbetrieb typische Alarmdauer von 2 s, wobei die Immunitätszeit auf 0.01 s eingestellt wird, reduziert sich die Alarmanzahl nochmals. Dies liegt daran, dass aufeinander folgende Alarme innerhalb des Zeitfensters von 2 s bei der Zählung nicht berücksichtigt werden. In Verbindung mit dem Auswerteverfahren nach Kap. 7 ergeben sich insgesamt 939 Alarmfälle. Die resultierende Verteilung der Schaltpunkte im Erfassungsfeld des Bewegungsmelders zeigt Abb. 10.8.

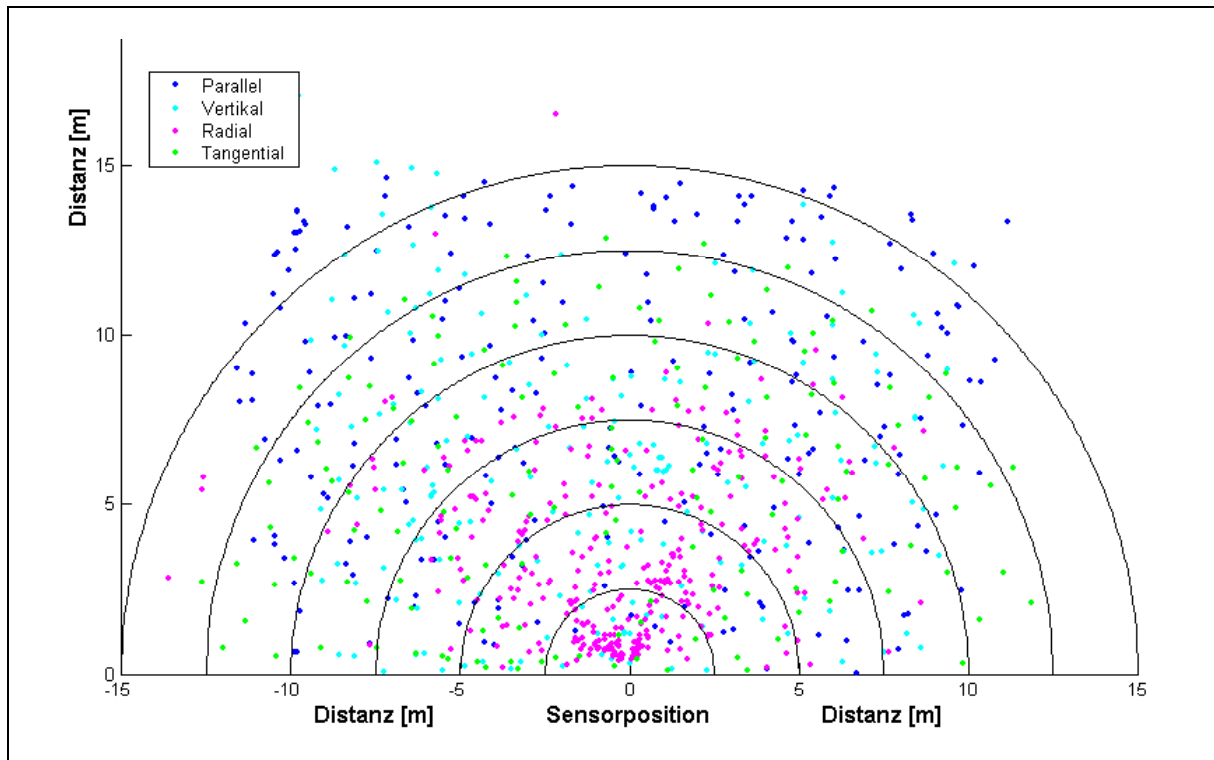


Abb. 10.8: Verteilung der Schaltpunkte im Erfassungsfeld des Bewegungsmelders in Verbindung mit dem konventionellen Auswerteverfahren nach Kap. 7. Für die Darstellung wurde eine Alarmdauer von 2 s und eine Immunitätszeit von 0.01 s gewählt.

Die zusätzlichen Auswerteparameter (AMPMIN, GRDMIN, GRDMAX, SLPMIN, SLPMAX, WIDMIN, WIDMAX) nach Tab. 10.2 wurden mit Hinblick auf eine Empfindlichkeitserhöhung für die Draufzugehsignale eingeführt, welche in größeren Entfernungen zum Sensor vergleichsweise kleine Amplituden aufweisen. Durch die zusätzlichen Parameter können die Auswerteschwellen abgesenkt werden, ohne dass sich die Anzahl der Fehlalarme deutlich erhöht. Allerdings werden durch die zusätzlichen Entscheidungskriterien auch viele Nutzsignale unterdrückt, wie die nachfolgenden Betrachtungen zeigen.

Verzichtet man auf die zusätzlichen Parameter, müssen die Schwellen THRMIN und THRMAX angehoben werden. Für die zur Optimierung verwendeten Störsignale wurden experimentell die Grenzwerte THRMIN = 9.00 und THRMAX = 10.00 ermittelt, ohne dass Fehlalarme auftreten. In Tab. 10.7 sind die resultierenden Alarmzahlen für die Nutzsignale in Abhängigkeit der Bewegungsrichtung zusammengefasst. Zum Vergleich wurden die Bewertungsergebnisse nach Tab. 10.6 bei Anwendung der Auswerteparameter nach Tab. 10.2 ebenfalls angegeben.

Betrachtet man nur die Alarmanzahl, ergeben sich deutliche Vorteile der reinen Schwellenauswertung. Ein Blick auf die in Abb. 10.9 dargestellte Verteilung der Schaltpunkte relativiert diese Aussage allerdings. Speziell im frontalen Bereich oberhalb von 7.5 m schneidet die einfache Schwellenauswertung für Signale aus vertikalen und radialen Bewegungsrichtungen

schlechter ab. Mit einer Alarmdauer von 2.0 s und einer Immunitätszeit von 0.01 s ergaben sich insgesamt 1052 Auslöser.

INDOOR-NUTZSIGNALE	Schwellenauswertung plus weitere Kriterien nach Tab. 10.2			Reine Schwellenauswertung (THRMIN = 9, THRMAX = 10)		
	PIR1	PIR2	PIR3	PIR1	PIR2	PIR3
Parallele Bewegungsrichtung	251	290	204	433	577	370
Vertikale Bewegungsrichtung	193	219	145	325	297	277
Radiale Bewegungsrichtung	352	388	265	553	604	414
Tangentiale Bewegungsrichtung	153	178	124	493	518	409
Alle Richtungen (Summe)	949	1075	738	1804	1996	1470

Tab. 10.7: Vergleich der Alarmanzahl für die Parametereinstellung nach Tab. 10.2 und der reinen Schwellenauswertung mit erhöhten Grenzwerten (THRMIN = 9, THRMAX = 10).

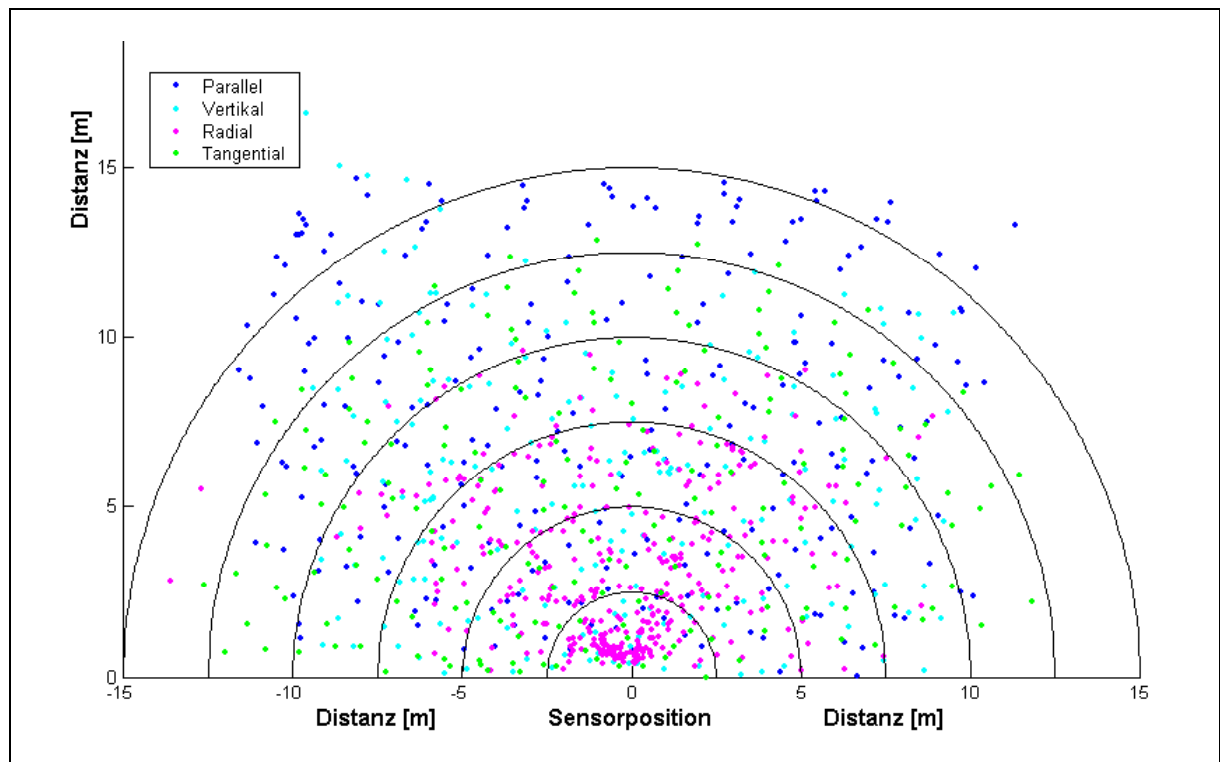


Abb. 10.9: Verteilung der Schaltpunkte im Erfassungsfeld des Bewegungsmelders in Verbindung mit einer reinen Schwellenauswertung der Sensorsignale. Für die Darstellung wurde eine Alarmdauer von 2 s und eine Immunitätszeit von 0.01 s gewählt.

10.4.2 Schaltverhalten bei Auswertung der Residuensequenz

Tab. 10.8 zeigt die für die im folgenden beschriebenen Untersuchungen verwendeten Einstellungen des Kalman-Filters 4. Im Vergleich zu dem Auswerteverfahren nach Abschnitt 10.4.1 werden nicht die Sensorsignale, sondern die Residuensequenz einer Schwellenauswertung

unterzogen. THRMIN und THRMAX legen das Toleranzband der adaptiven Schwellen für jeden Sensorkanal fest.

Die Varianzen $(\sigma_{ss})^2$ des allgemeinen Signalmodells, $(\sigma_{dd})^2$ des Störmodells und $(\sigma_{vv})^2$ des Ausgangsrauschens wurden für die Untersuchungen als konstant angenommen.

Tab. 10.9 zeigt typische Charakteristiken der Residuensequenzen RSS1, RSS2 und RSS3 sowie die Fehlalarmzahl pro Sensorkanal bei Verwendung der Outdoor-Störungen aus dem vorhergehenden Abschnitt.

PARAMETER-EINSTELLUNG	KALMAN-Filter
MITTELUNGSANZAHL	4
THRMIN	11.00
THRMAX	12.00
VARIANZEN $(\sigma_{ss})^2$	0.10
VARIANZEN $(\sigma_{dd})^2$	20.0
VARIANZEN $(\sigma_{vv})^2$	0.01

Tab. 10.8: Parametereinstellung des Kalman-Filters 4 mit Schwellenauswertung.

In Tab. 10.10 wurden wiederum die korrespondierenden Kenngrößen des Indoor-Rauschens aufgeführt. Im Vergleich zu Tab. 10.9 erkennt man die geringeren Abweichungen der Residuensequenz aus der Ruhelage sowie die kleinere maximale Varianz der Residuensequenz innerhalb des betrachteten Zeitfensters. Ausgehend von der für die Outdoor-Störungen optimierten Parametereinstellung werden im folgenden die Nutzsignale aus dem Indoor-Bereich mit dem nicht-adaptiven Kalman-Filter verarbeitet. In Tab. 10.11 ist die resultierende Alarmanzahl für die verschiedenen Sensorkanäle angegeben. Die Summe der Einzelalarme nach Tab. 10.11 beträgt 2938.

Zur Verbesserung des Schaltverhaltens wurde im weiteren Verlauf die als konstant angenommene Varianz $(\sigma_{dd})^2 = 20.0$ des Störmodells in zwei Stufen auf $(\sigma_{dd})^2 = 2.00$ bzw. $(\sigma_{dd})^2 = 0.20$ reduziert. Bei sonst unveränderten Parametern der Signalauswertung ergab sich für $(\sigma_{dd})^2 = 0.20$ ein Fehlalarm in Verbindung mit den Störsignalen aus dem Outdoorbereich, während bei $(\sigma_{dd})^2 = 2.00$ noch keine Fehlschaltungen aufgetreten sind. In Tab. 10.12 sind die Alarmzahlen in Abhängigkeit der Störvarianz am Systemeingang im Vergleich zu den Bewertungsergebnissen der konventionellen Auswertestrategie nach Abschnitt 10.4.1 zusammengefasst. Tab. 10.12 zeigt, dass sich durch eine Verringerung der Störvarianz $(\sigma_{dd})^2$ am Systemeingang die Empfindlichkeit des Bewegungsmelders steigern lässt. Für $(\sigma_{dd})^2 = 0.20$ werden deutlich mehr Alarmereignisse erkannt als bei dem konventionellen Auswerteverfahren. Zur weiteren Beurteilung wurde in Tab. 10.13 zusätzlich die Verteilung der Alarme für $(\sigma_{dd})^2 = 0.20$ auf die verschiedenen Bewegungsrichtungen angegeben.

Man erkennt in Tab. 10.13, dass die erhöhte Alarmzahl auf eine gesteigerte Empfindlichkeit für parallele und tangentielle Gehrichtungen zurückzuführen ist. Die dunkel grau unterlegten Felder enthalten dagegen Alarmzahlen für vertikale bzw. radiale Bewegungsrichtungen, welche im Vergleich zur Signalauswertung nach Kap. 10.4.1 deutlich geringer ausfallen.

OUTDOOR-RAUSCHEN	RSS1	RSS2	RSS3
Min. Residuensequenz	-7.19	-7.67	-7.93
Max. Residuensequenz	6.69	11.85	10.01
Min. Mittelwert (Zeitfenster)	-1.23	-1.19	-1.20
Max. Mittelwert (Zeitfenster)	0.17	0.31	0.26
Min. Varianz (Zeitfenster)	0.31	0.30	0.30
Max. Varianz (Zeitfenster)	2.79	4.18	3.13
Fehlalarmzahl	0	0	0

Tab. 10.9: Outdoor-Störsignale bei Schwellenauswertung der Residuensequenz (Parameter s. Tab. 10.8). Minima und Maxima der Residuensequenzen RSS1, RSS2, RSS3, minimale und maximale Mittelwerte und Varianzen der Residuensequenzen in einem Zeitfenster mit 100 Datensätzen sowie Fehlalarmzahl.

INDOOR-RAUSCHEN	RSS1	RSS2	RSS3
Min. Residuensequenz	-5.95	-5.40	-4.25
Max. Residuensequenz	7.92	4.12	4.99
Min. Mittelwert (Zeitfenster)	-1.21	-1.18	-1.25
Max. Mittelwert (Zeitfenster)	0.20	0.17	0.18
Min. Varianz (Zeitfenster)	0.49	0.48	0.62
Max. Varianz (Zeitfenster)	2.04	1.72	2.22
Fehlalarmzahl	0	0	0

Tab. 10.10: Indoor-Rauschen bei Schwellenauswertung der Residuensequenz (Parameter s. Tab. 10.8). Minima und Maxima der Residuensequenzen RSS1, RSS2, RSS3, minimale und maximale Mittelwerte und Varianzen der Residuensequenz in einem Zeitfenster mit 100 Datensätzen sowie Fehlalarmzahl.

INDOOR-NUTZSIGNALE	RSS1	RSS2	RSS3
Min. Residuensequenz	-88.74	-102.84	-114.08
Max. Residuensequenz	107.37	113.33	67.34
Min. Mittelwert (Zeitfenster)	-3.42	-3.47	-2.92
Max. Mittelwert (Zeitfenster)	3.31	3.56	3.99
Min. Varianz (Zeitfenster)	0.64	0.49	0.62
Max. Varianz (Zeitfenster)	1475.17	1847.04	1202.37
Alarmanzahl	1001	1091	846

Tab. 10.11: Indoor-Nutzsignale bei Schwellenauswertung der Residuensequenz (Parameter s. Tab. 10.8). Minima und Maxima der Residuensequenzen RSS1, RSS2, RSS3, minimale und maximale Mittelwerte und Varianzen der Residuensequenz in einem Zeitfenster mit 100 Datensätzen sowie Alarmanzahl.

INDOOR-NUTZSIGNALE	RSS1	RSS2	RSS3	Summe
Alarmzahl nach Abschnitt 10.4.1	949	1075	738	2762
Alarmzahl für $(\sigma_{ss})^2 = 0.10, (\sigma_{dd})^2 = 20.0$	1001	1091	846	2938
Alarmzahl für $(\sigma_{ss})^2 = 0.10, (\sigma_{dd})^2 = 2.00$	1025	1124	861	3010
Alarmzahl für $(\sigma_{ss})^2 = 0.10, (\sigma_{dd})^2 = 0.20$	1203	1303	986	3492

Tab. 10.12: Alarmereignisse in Verbindung mit Nutzsignalen aus dem Indoorbereich bei unterschiedlichen Varianzen $(\sigma_{dd})^2$ des Störprozesses am Systemeingang.

INDOOR-NUTZSIGNALE	Konventionelle Auswertung nach Abschnitt 9.3.1			Kalman-Filter 4 und Schwellenabfrage der Residuensequenz		
	PIR1	PIR2	PIR3	RSS1	RSS2	RSS3
Parallele Bewegungsrichtung	251	290	204	330	454	257
Vertikale Bewegungsrichtung	193	219	145	260	182	211
Radiale Bewegungsrichtung	352	388	265	186	239	196
Tangentiale Bewegungsrichtung	153	178	124	427	428	322
Alle Richtungen (Summe)	949	1075	738	1203	1303	986

Tab. 10.13: Alarmzahl bei Schwellenauswertung der Residuensequenz in Abhängigkeit der Bewegungsrichtung ($(\sigma_{dd})^2 = 0.02$).

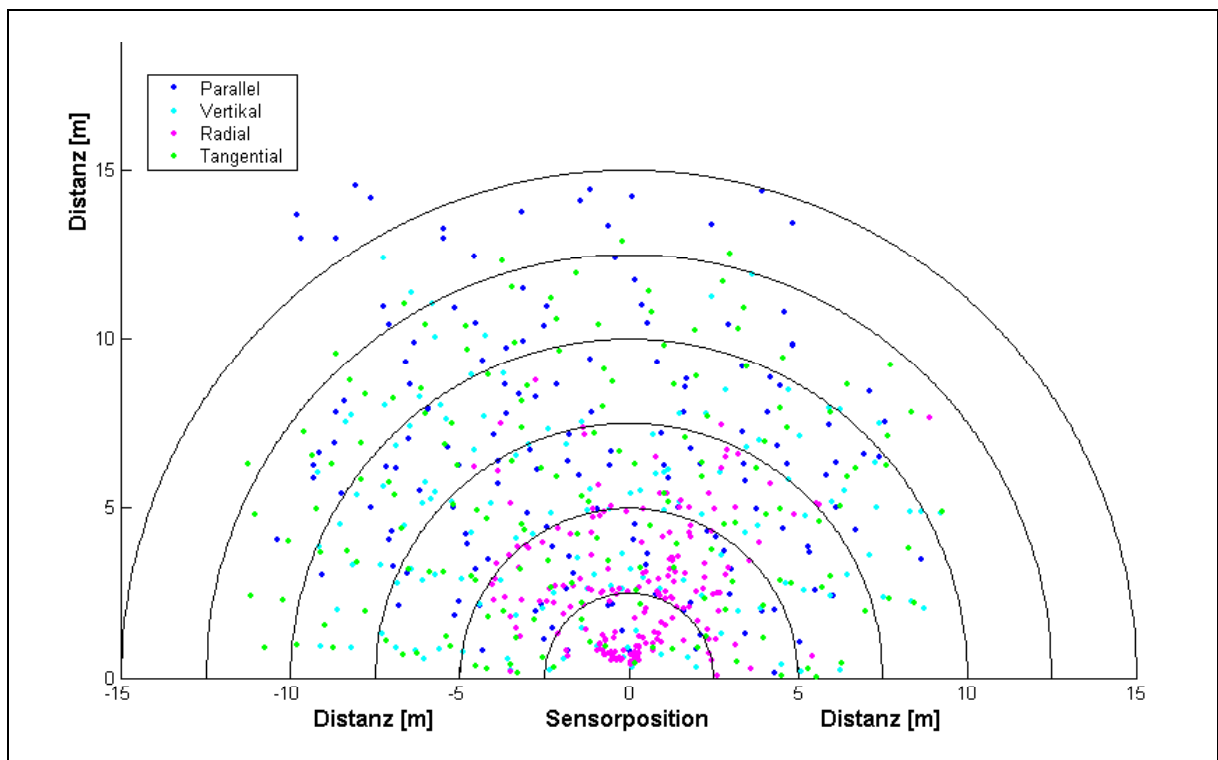


Abb. 10.10: Verteilung der Schaltpunkte im Erfassungsfeld des Bewegungsmelders bei Schwellenauswertung der Residuensequenz. Es wurde eine Alarmdauer von 2 s in Verbindung mit einer Immunitätszeit von 0.01 s gewählt.

In Analogie zur Vorgehensweise in Abschnitt 10.4.1 wurde auch im vorliegenden Anwendungsfall der Testbetrieb des Wächters mit einer Alarmdauer von 2 s und einer anschließenden Immunitätszeit von 0.01 s simuliert. Für $(\sigma_{dd})^2 = 20.0$ ergaben sich 585 Alarmfälle, für $(\sigma_{dd})^2 = 2.00$ traten 595 Alarmereignisse auf.

Mit $(\sigma_{dd})^2 = 0.20$ wurde die höchste Empfindlichkeit mit insgesamt 664 Alarmen und der in Abb. 10.10 dargestellten Verteilung der Schaltpunkte im Erfassungsfeld erzielt. Die vorher beschriebene Reduktion der Schaltpunkte für radiale und vertikale Bewegungsrichtungen wird anhand von Abb. 10.10 bestätigt.

10.4.3 Bewertung der 1. Komponente des Zustandsvektors der Prädiktion

Die in Abschnitt 10.4.2 verwendeten Varianzen $(\sigma_{ss})^2$ des allgemeinen Signalmodells und $(\sigma_{dd})^2$ des Störmodells wurden offline anhand von Testdaten ermittelt und dann als Konstanten im Rahmen des vereinfachten Systemmodells betrachtet. Im weiteren Verlauf werden beide Varianzen in einem Zeitfenster von 100 Datensätzen dynamisch berechnet.

Die Kovarianzmatrix Q_4 wird vor der Verarbeitung eines neuen Datensatzes jeweils aktualisiert. Durch die Nachführung der Kovarianzmatrix ist eine bessere Adaption an die in dem jeweiligen Detektorsignal vorhandene Dynamik möglich, so dass unerwartete Signaländerungen bereits frühzeitig erkannt werden können.

Nach den Ausführungen in den Kap. 10.2.2.1, 10.2.2.2 und 10.2.2.3 werden die Varianzen n_w des Signalmodells und N_w des Störmodells unter Beachtung sinnvoller Grenzwerte aus der Varianz $(\sigma_{yy})^2$ des Ausgangssignals $y(k)$ des Bewegungsmelders berechnet. Solange die Varianz $(\sigma_{dd})^2|_{H0} = (\sigma_{yy})^2|_{H0} - (\sigma_{vv})^2$ kleiner ist als $(\sigma_{dd,max})^2|_{H0}$, wird $(\sigma_{dd})^2|_{H0}$ zur Ermittlung der Störvarianz N_w am Systemeingang verwendet. Der Summand $(\sigma_{vv})^2$ ist sehr viel kleiner als $(\sigma_{yy})^2|_{H0}$, so dass er bei der Berechnung des Störmodells vernachlässigt wird.

Die Varianz n_w des allgemeinen Signalmodells wird im vorliegenden Fall nach (Gl. 10.9) bestimmt, wobei die minimale Nutzsinalvarianz $(\sigma_{ss,min})^2|_{H0}$ nach (Gl. 10.8) angenommen wird. Da $(\sigma_{ss,min})^2|_{H0}$ nach (Gl. 10.8) kleiner als die minimale Varianz aller verwendeten Testsignale ist, bewirkt diese Festlegung, dass das Ausgangssignal $y(k)$ hauptsächlich in der 4. Komponente des prädizierten Zustandsvektors nachgeführt wird. In der zur Bewegungserkennung verwendeten 1. Zustandsvektorkomponente sind diese Störanteile demzufolge stark bedämpft.

Übersteigt die Varianz $(\sigma_{dd})^2|_{H1} = (\sigma_{yy})^2|_{H1} - (\sigma_{vv})^2$ den Maximalwert $(\sigma_{dd,max})^2|_{H0}$, wird die Varianz n_w des allgemeinen Signalmodells vergrößert, während die Störvarianz N_w nach (Gl. 10.11) mit $(\sigma_{dd})^2|_{H1} = (\sigma_{dd,max})^2|_{H0}$ auf einen konstanten Maximalwert eingestellt wird. Zur Berechnung von n_w dient der Ansatz mit korrelierten Stör- und Nutzsinalen nach (Gl. 10.14), wobei für die Störvarianz $(\sigma_{dd})^2|_{H1}$ wiederum der Maximalwert $(\sigma_{dd,max})^2|_{H0}$ eingesetzt wird. Je größer die Varianz $(\sigma_{yy})^2|_{H1}$ des Ausgangssignals $y(k)$ ist, umso größer werden die

Varianzen $(\sigma_{ss})^2|_{H1}$ bzw. n_w , so dass das zugrundeliegende Ausgangssignal verstärkt in der 1. Komponente des vorhergesagten Zustandsvektors nachgeführt wird. Wie bereits in Kap. 10.3.1.2 beschrieben wurde, erfolgt anschließend die Schwellenauswertung.

In Tab. 10.14 sind die Parameter zur Auswertung der 1. Zustandsvektorkomponente zusammengefasst. Tab. 10.15 - Tab. 10.17 enthalten die Minima und Maxima der ausgewerteten Signalverläufe sowie die minimalen und maximalen Mittelwerte und Varianzen in einem Zeitfenster mit 100 Datensätzen.

Außerdem ist die Fehlalarmzahl für Indoor- und Outdoor-Störungen sowie die Alarmanzahl bei Verwendung der Nutzsignale angegeben. Die Summe der Einzelalarme nach Tab. 10.17 beträgt 3263.

Zum Vergleich mit den Ergebnissen aus Kap. 10.4.1 und 10.4.2 wurden die Alarmzahlen der verschiedenen Verfahren in Verbindung mit den Nutzsignalen aus dem Indoorbereich in Tab. 10.19 zusammengefasst.

PARAMETER-EINSTELLUNG	KALMAN-Filter
MITTELUNGSANZAHL	4
THRMIN	3.00
THRMAX	4.00
AMPMIN	16.00
GRDMIN	2.00
GRDMAX	78.13
SLPMIN	13.00
SLPMAX	47.00
WIDMIN	23.00
WIDMAX	80.00
VARIANZEN $(\sigma_{ss})^2$, Minimum	0.10
VARIANZEN $(\sigma_{dd})^2$, Maximum	2.00
VARIANZEN $(\sigma_{vv})^2$	0.01

Tab. 10.14: Parametereinstellung zur Auswertung der 1. Komponente des Zustandsvektors der Prädiktion.

Die dunkel grau unterlegten Felder in Tab. 10.18 enthalten die Alarmzahlen für vorwiegend radiale Bewegungsrichtungen. Die Anzahl der Alarme ist sowohl für radiale als auch für die übrigen Bewegungsrichtungen höher als bei der Auswertung nach Kap. 10.4.1.

OUTDOOR-RAUSCHEN	PDM1(1,1)	PDM2(1,1)	PDM3(1,1)
Min. 1. Zustandsvektorkomponente	469.33	459.66	464.60
Max. 1. Zustandsvektorkomponente	475.70	485.69	485.22
Min. Mittelwert (Zeitfenster)	471.60	469.96	472.56
Max. Mittelwert (Zeitfenster)	473.71	475.54	476.60
Min. Varianz (Zeitfenster)	0.09	0.10	0.11
Max. Varianz (Zeitfenster)	1.84	60.87	32.03
Fehlalarmzahl	0	0	0

Tab. 10.15: Outdoor-Störungen bei Schwellenauswertung der 1. Zustandsvektorkomponente der Prädiktion (Parameter s. Tab. 10.14). Minima und Maxima der Signale PDM1(1,1), PDM2(1,1), PDM3(1,1), minimale und maximale Mittelwerte und Varianzen in einem Zeitfenster mit 100 Datensätzen sowie Fehlalarmanzahl.

INDOOR-RAUSCHEN	PDM1(1,1)	PDM2(1,1)	PDM3(1,1)
Min. 1. Zustandsvektorkomponente	473.77	464.56	466.10
Max. 1. Zustandsvektorkomponente	480.06	470.96	475.16
Min. Mittelwert (Zeitfenster)	476.25	466.94	470.38
Max. Mittelwert (Zeitfenster)	477.68	468.13	471.71
Min. Varianz (Zeitfenster)	0.25	0.18	0.20
Max. Varianz (Zeitfenster)	1.50	1.84	3.62
Fehlalarmzahl	0	0	0

Tab. 10.16: Indoor-Rauschen bei Schwellenauswertung der 1. Zustandsvektorkomponente der Prädiktion (Parameter s. Tab. 10.14). Minima und Maxima der Signale PDM1(1,1), PDM2(1,1), PDM3(1,1), minimale und maximale Mittelwerte und Varianzen in einem Zeitfenster mit 100 Datensätzen sowie Fehlalarmanzahl.

INDOOR-NUTZSIGNALE	PDM1(1,1)	PDM2(1,1)	PDM3(1,1)
Min. 1. Zustandsvektorkomponente	201.40	76.75	83.50
Max. 1. Zustandsvektorkomponente	807.63	748.67	769.06
Min. Mittelwert (Zeitfenster),	449.24	431.10	431.50
Max. Mittelwert (Zeitfenster)	506.41	501.88	510.09
Min. Varianz (Zeitfenster)	0.20	0.25	0.19
Max. Varianz (Zeitfenster)	16787.10	20664.07	15704.48
Alarmanzahl	1099	1283	881

Tab. 10.17: Indoor-Nutzsignale bei Schwellenauswertung der 1. Zustandsvektorkomponente der Prädiktion (Parameter s. Tab. 10.14). Minima und Maxima der Signale PDM1(1,1), PDM2(1,1), PDM3(1,1), minimale und maximale Mittelwerte und Varianzen in einem Zeitfenster mit 100 Datensätzen sowie Alarmzahl.

	Konventionelle Auswertung nach Abschnitt 9.3.1			Bewertung der 1. Zustandsvektor- komp. (Parameter nach Tab. 10.14)		
	PIR1	PIR2	PIR3	PDM1(1,1)	PDM2(1,1)	PDM3(1,1)
INDOOR-NUTZSIGNALE						
Parallele Bewegungsrichtung	251	290	204	282	340	234
Vertikale Bewegungsrichtung	193	219	145	230	273	175
Radiale Bewegungsrichtung	352	388	265	417	474	329
Tangentiale Bewegungsrichtung	153	178	124	170	196	143
Alle Richtungen	949	1075	738	1099	1283	881

Tab. 10.18: Alarmzahl in Verbindung mit Nutzsignalen bei Auswertung der 1. Zustandsvektorkomponente der Prädiktion in Abhängigkeit der Bewegungsrichtung.

INDOOR-NUTZSIGNALE	Kanal 1	Kanal 2	Kanal 3	Summe	Alarmzahl (Testbetrieb)
Alarmzahl bei Verwendung der Parameter nach Tab. 10.2 (Kap. 10.4.1)	949	1075	738	2762	939
Alarmzahl reine Schwellenauswertung (Kap. 10.4.1)	1804	1996	1470	5270	1052
Alarmzahl bei Auswertung der Residuensequenz, $(\sigma_{dd})^2=0.20$ (Kap. 10.4.2)	1203	1303	986	3492	685
Alarmzahl für 1. Zustandsvektorkomponente und Parameter nach Tab. 10.14	1099	1283	881	3263	1087

Tab. 10.19: Alarmzahl bei Auswertung der 1. Zustandsvektorkomponente der Prädiktion in Abhängigkeit der Bewegungsrichtung im Vergleich zu den bisherigen Verfahren.

Abb. 10.11 zeigt die Verteilung der Schaltpunkte im Erfassungsfeld des Bewegungsmelders. Um den Vergleich mit Abb. 10.8 und Abb. 10.10 zu ermöglichen, wurde erneut eine Alarmdauer von 2.0 s und eine Immunitätszeit von 0.01 s eingestellt. Es ergaben sich insgesamt 1087 Alarme. Das wesentliche Ergebnis dieser Darstellung ist eine Reichweitenvergrößerung für die Draufzugehsignale. Während in Abb. 10.8 nur wenige Schaltpunkte außerhalb der 8 m-Grenze liegen, erkennt man in Abb. 10.11 mehr Schaltpunkte im Bereich von 8 ...11 m.

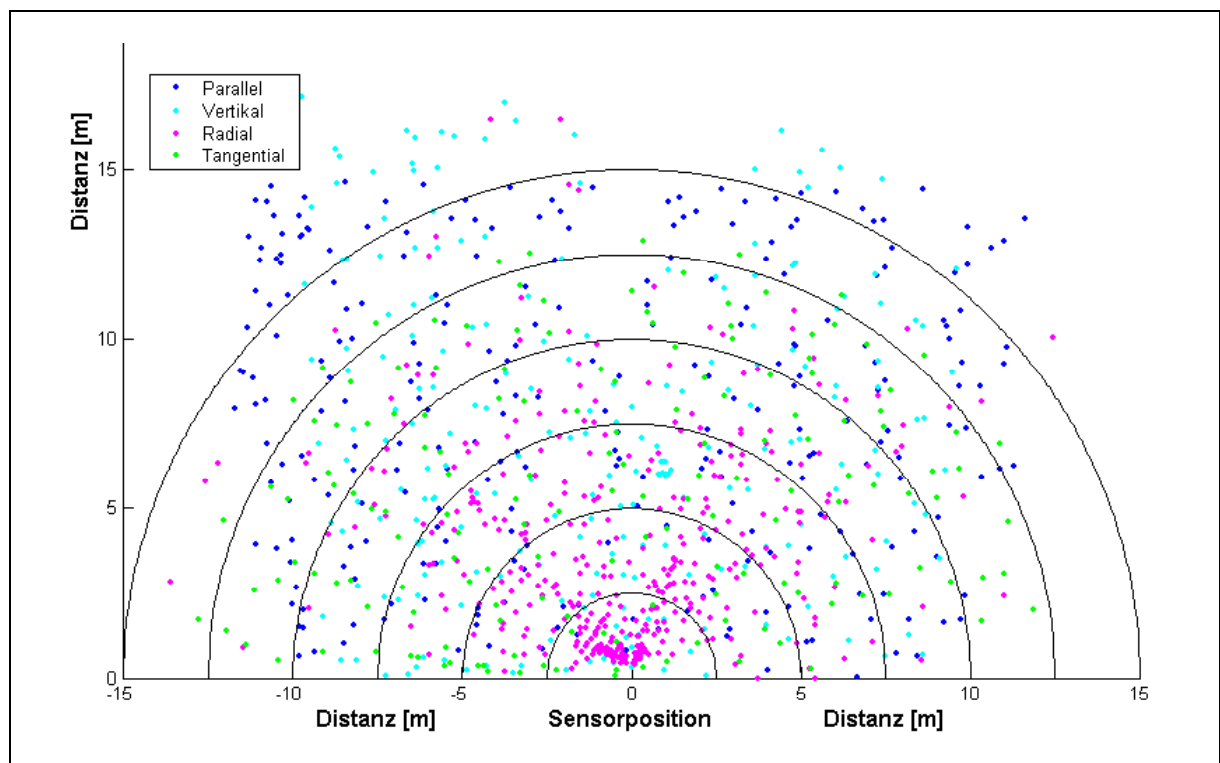


Abb. 10.11: Verteilung der Schaltpunkte im Erfassungsfeld des Bewegungsmelders bei Auswertung der 1. Komponente des Zustandsvektors der Prädiktion (Auswerteparameter nach Tab. 10.14). Es wurde eine Alarmdauer von 2 s in Verbindung mit einer Immunitätszeit von 0.01 s gewählt.

Im Vergleich zur reinen Schwellenauswertung nach Abschnitt 10.4.1 ist auch für die von den Störsignalen bereinigte 1. Komponente des Zustandsvektors die Auswirkung der zusätzlichen Entscheidungskriterien von Interesse. Tab. 10.20 zeigt die Alarmanzahlen bei Auswertung der 1. Komponente in Verbindung mit einer einfachen Schwellenauswertung. Auch im vorliegenden Fall müssen die Grenzwerte THRMIN und THRMAX erhöht werden, um die Fehlalarmanzahl für die zur Optimierung verwendeten Outdoor-Störungen auf Null abzusenken. Experimentell wurden die Grenzen THRMIN = 5 und THRMAX = 6 bestimmt.

INDOOR-NUTZSIGNALE	Bewertung der 1. Zustandsvektorkomp. (Parameter nach Tab. 10.14)			Bewertung der 1. Zustandsvektorkomp. (reine Schwellenauswertung)		
	PDM1(1,1)	PDM2(1,1)	PDM3(1,1)	PDM1(1,1)	PDM2(1,1)	PDM3(1,1)
Parallele Bewegung	282	340	234	488	635	386
Vertikale Bewegung	230	273	175	378	379	309
Radiale Bewegung	417	474	329	698	785	513
Tangentiale Bewegung	170	196	143	525	572	444
Alle Richtungen (Summe)	1099	1283	881	2089	2371	1652

Tab. 10.20: Vergleich der Auswerteverfahren mit der Parametereinstellung nach Tab. 10.14 und einer Schwellenauswertung ohne zusätzliche, einschränkende Kriterien.

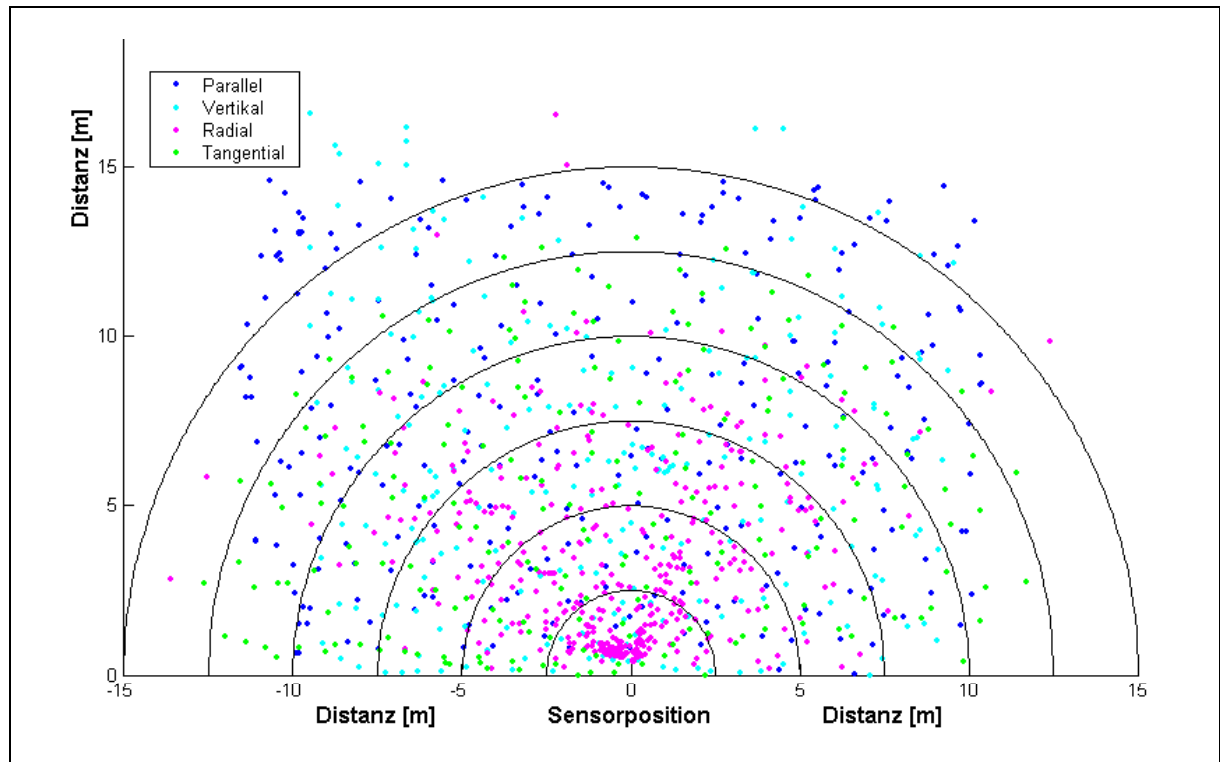


Abb. 10.12: Verteilung der Schaltpunkte im Erfassungsfeld des Bewegungsmelders bei Auswertung der 1. Komponente des vorhergesagten Zustandsvektors mit reiner Schwellenauswertung (THRMIN = 5, THRMAX = 6). Es wurde eine Alarmdauer von 2 s in Verbindung mit einer Immunitätszeit von 0.01 s gewählt.

Die Summe der Einzelalarme für die reine Schwellenauswertung beträgt 6112. Es treten somit 842 Schaltereignisse mehr auf als bei der einfachen Schwellenauswertung der Sensorsignale. Abb. 10.12 zeigt die zugehörige Verteilung der Schaltpunkte im Erfassungsfeld des Bewegungsmelders. In Verbindung mit der Alarmdauer von 2.0 s und der Immunitätszeit von 0.01 s ergaben sich insgesamt 1176 Schaltpunkte. Die einfache Schwellenauswertung nach 10.4.1 lieferte 1052 Auslöser.

10.4.4 Bewertung der 2. Komponente des Zustandsvektors der Prädiktion

In diesem Abschnitt werden die Ergebnisse bei Auswertung der 2. Komponente des vorhergesagten Zustandsvektors beschrieben. Das grundsätzliche Verfahren wurde bereits in Kap. 10.3.1.3 erläutert. An dieser Stelle soll darauf aufmerksam gemacht werden, dass die 2. Komponente des Zustandsvektors für kleine Abtastintervalle T näherungsweise der Ableitung der ersten Komponente entspricht. Daher ist eine erhöhte "Signalunruhe" zu erwarten, was sich positiv auf das Schaltverhalten des Wächters auswirken wird. Die adaptive Anpassung der Kovarianzmatrix Q_4 des Kalman-Filters erfolgt anhand der in Abschnitt 10.4.3 beschriebenen Vorgehensweise. Tab. 10.21 enthält die zur Auswertung verwendete Parametereinstellung.

PARAMETER-EINSTELLUNG	KALMAN-Filter
MITTELUNGSANZAHL	4
THRMIN	14.00
THRMAX	15.00
AMPMIN	60.00
GRDMIN	7.0
GRDMAX	2200.00
SLPMIN	3.00
SLPMAX	32.00
WIDMIN	6.00
WIDMAX	52.00
VARIANZEN (σ_{ss}) ² , Minimum	0.10
VARIANZEN (σ_{dd}) ² , Maximum	2.00
VARIANZEN (σ_{vv}) ²	0.01

Tab. 10.21: Parameter zur Auswertung der 2. Komponente des Zustandsvektors.

In Tab. 10.22, Tab. 10.23 und Tab. 10.24 sind die Minimal- und Maximalwerte der 2. Komponente des Zustandsvektors sowie die Minima und Maxima der Mittelwerte und der Varianzen in einem Zeitfenster mit 100 Datensätzen für die Outdoor-Störungen, für das Indoor-Rauschen und für die Nutzsignale angegeben. Die Parameteroptimierung erfolgte wiederum unter der Maßgabe, dass bei den Testsignalen aus dem Outdoorbereich keine Fehlauslöser entstehen.

Die Aufteilung der Alarme nach Tab. 10.24 auf die einzelnen Bewegungsrichtungen ist in Tab. 10.25 angegeben. In Verbindung mit der Auswertung der 2. Zustandsvektorkomponente ergibt sich wiederum eine höhere Empfindlichkeit des Wächters mit Kalman-Filter als bei dem Auswerteverfahren nach Kap. 10.4.1. Die Summe der Einzelalarme beträgt 8492. Abb. 10.13 zeigt die Verteilung der Schaltpunkte im Erfassungsfeld für die Parametereinstellung nach Tab. 10.21 im Testbetrieb. Es ergaben sich insgesamt 1279 Auslöser.

OUTDOOR-RAUSCHEN	PDM1(2,1)	PDM2(2,1)	PDM3(2,1)
Min. 2. Zustandsvektorkomponente	-3.54	-32.61	-21.16
Max. 2. Zustandsvektorkomponente	3.89	32.79	26.42
Min. Mittelwert (Zeitfenster)	-0.56	-4.75	-3.77
Max. Mittelwert (Zeitfenster)	0.53	5.01	3.43
Min. Varianz (Zeitfenster)	0.03	0.04	0.04
Max. Varianz (Zeitfenster)	2.56	393.83	183.52
Fehlalarmzahl	0	0	0

Tab. 10.22: Outdoor-Störungen bei Schwellenauswertung der 2. Zustandsvektorkomponente der Prädiktion (Parameter s. Tab. 10.21). Minima und Maxima der Signale PDM1(2,1), PDM2(2,1), PDM3(2,1), minimale und maximale Mittelwerte und Varianzen in einem Zeitfenster mit 100 Datensätzen sowie Fehlalarmzahl.

INDOOR-RAUSCHEN	PDM1(2,1)	PDM2(2,1)	PDM3(2,1)
Min. 2. Zustandsvektorkomponente	-3.23	-2.99	-4.55
Max. 2. Zustandsvektorkomponente	2.99	3.73	6.63
Min. Mittelwert (Zeitfenster)	-0.47	-0.40	-0.69
Max. Mittelwert (Zeitfenster)	0.49	0.40	0.79
Min. Varianz (Zeitfenster)	0.12	0.08	0.12
Max. Varianz (Zeitfenster)	1.77	2.39	5.68
Fehlalarmzahl	0	0	0

Tab. 10.23: Indoor-Rauschen bei Schwellenauswertung der 2. Zustandsvektorkomponente der Prädiktion (Parameter s. Tab. 10.21). Minima und Maxima der Signale PDM1(2,1), PDM2(2,1), PDM3(2,1), minimale und maximale Mittelwerte und Varianzen in einem Zeitfenster mit 100 Datensätzen sowie Fehlalarmzahl.

INDOOR-NUTZSIGNALE	PDM1(2,1)	PDM2(2,1)	PDM3(2,1)
Min. 2. Zustandsvektorkomponente	-4649.74	-3767.81	-5924.34
Max. 2. Zustandsvektorkomponente	4565.94	5667.76	4222.61
Min. Mittelwert (Zeitfenster),	-266.63	-285.36	-242.04
Max. Mittelwert (Zeitfenster)	237.17	330.33	272.55
Min. Varianz (Zeitfenster)	0.12	0.12	0.10
Max. Varianz (Zeitfenster)	4840689.91	4980887.47	3671827.41
Alarmanzahl	2820	3217	2455

Tab. 10.24: Indoor-Nutzsignale bei Schwellenauswertung der 2. Zustandsvektorkomponente der Prädiktion (Parameter s. Tab. 10.21). Minima und Maxima der Signale PDM1(2,1), PDM2(2,1), PDM3(2,1), minimale und maximale Mittelwerte und Varianzen in einem Zeitfenster mit 100 Datensätzen sowie Alarmanzahl.

	Konventionelle Auswertung nach Abschnitt 9.3.1			Bewertung der 2. Zustandsvektor- komp. (Parameter nach Tab. 10.21)		
	PIR1	PIR2	PIR3	PDM1(2,1)	PDM2(2,1)	PDM3(2,1)
INDOOR-NUTZSIGNALE						
Parallele Bewegungsrichtung	251	290	204	637	785	537
Vertikale Bewegungsrichtung	193	219	145	551	586	485
Radiale Bewegungsrichtung	352	388	265	1020	1187	881
Tangentiale Bewegungsrichtung	153	178	124	612	659	552
Alle Richtungen	949	1075	738	2820	3217	2455

Tab. 10.25: Alarmzahl in Verbindung mit Nutzsignalen bei Auswertung der 2. Zustandsvektorkomponente der Prädiktion in Abhängigkeit der Bewegungsrichtung.

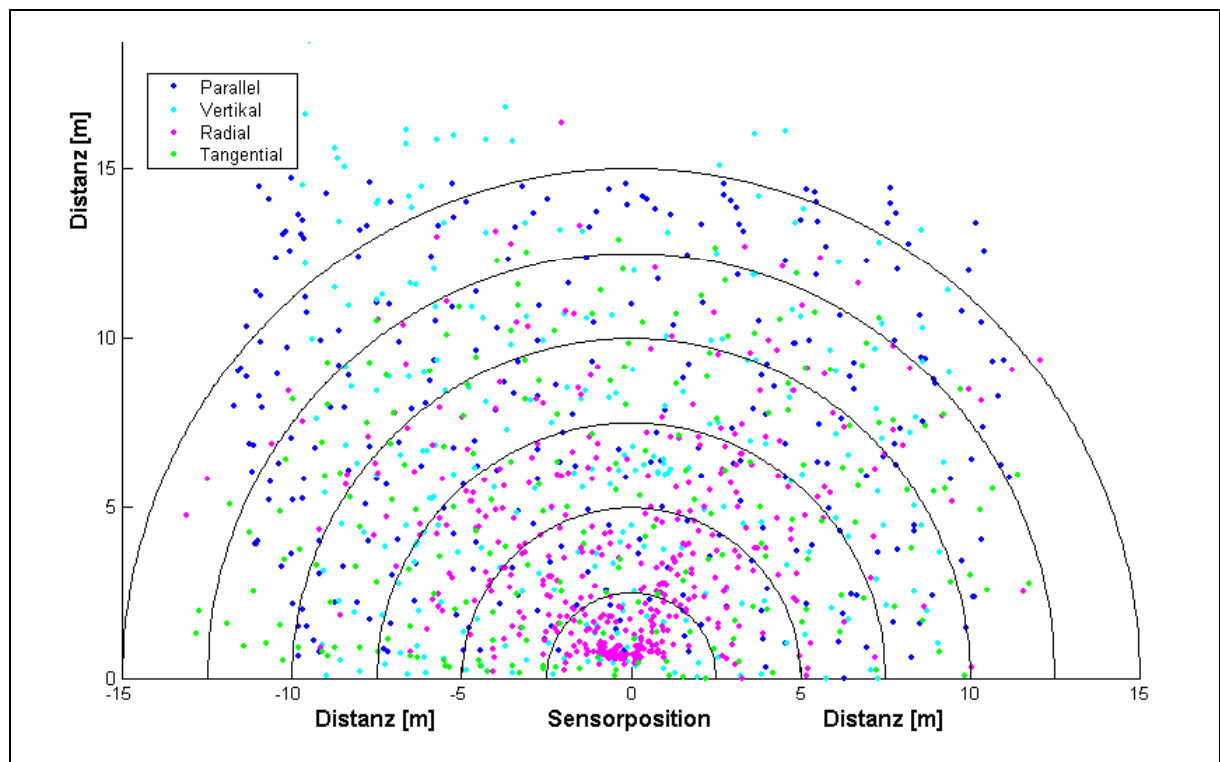


Abb. 10.13: Verteilung der Schaltpunkte im Erfassungsfeld des Bewegungsmelders bei Auswertung der 2. Komponente des Zustandsvektors der Prädiktion (Auswerteparameter nach Tab. 10.21). Es wurde eine Alarmdauer von 2 s in Verbindung mit einer Immunitätszeit von 0.01 s gewählt.

Zum Vergleich mit den vorher beschriebenen Untersuchungen wurde zusätzlich eine reine Schwellenauswertung der 2. Komponente des Zustandsvektors durchgeführt. Wählt man die Grenzwerte $THR_{MIN} = 20.0$ und $THR_{MAX} = 21.0$, ergeben sich keine Fehlschaltungen in Verbindung mit den bisher verwendeten Signalen aus dem Outdoor-Testfeld. In Tab. 10.26 sind die Alarme pro Sensorkanal im Vergleich zu den bisherigen Auswertergebnissen aufgeführt. Die Summe der Einzelalarme bei einer Schwellenauswertung ohne zusätzliche Parameter beträgt 9878 und ist damit höher als bei allen bisher beschriebenen Verfahren. Tab. 10.27

zeigt zusätzlich die Verteilung der Einzelalarme auf die unterschiedlichen Bewegungsrichtungen. In allen Gehrichtungen tritt eine Erhöhung der Alarmzahl im Vergleich zur Auswertung der 2. Zustandsvektorkomponente mit zusätzlichen Parametern auf.

Die Verteilung der Schaltpunkte im Erfassungsfeld des Bewegungsmelders ist für die einfache Schwellenauswertung in Abb. 10.14 wiedergegeben. Trotz der sehr hohen Einzelalarmanzahl ergeben sich im Testbetrieb lediglich 1262 Schaltsignale, wenn eine Alarmdauer von 2.0 s und eine Immunitätszeit von 0.01 s eingestellt wird.

INDOOR-NUTZSIGNALE	Kanal 1	Kanal 2	Kanal 3	Summe	Alarmzahl (Testbetrieb)
Alarmzahl bei Verwendung der Parameter nach Tab. 10.2 (Kap. 10.4.1)	949	1075	738	2762	939
Alarmzahl für reine Schwellenabfrage (Kap. 10.4.1)	1804	1996	1470	5270	1052
Alarmzahl bei Auswertung der Residuensequenz (Kap. 10.4.2)	1203	1303	986	3492	685
Alarmzahl für 1. Zustandsvektorkomponente und Parameter nach Tab. 10.14	1099	1283	881	3263	1087
Alarmzahl für 1. Zustandsvektorkomponente und reine Schwellenabfrage	2089	2371	1652	6112	1176
Alarmzahl für 2. Zustandsvektorkomponente und Parameter nach Tab. 10.21	2820	3217	2455	8492	1279
Alarmzahl für 2. Zustandsvektorkomponente und reine Schwellenabfrage	3286	3704	2888	9878	1262

Tab. 10.26: Alarmzahl bei Auswertung der 2. Zustandsvektorkomponente der Prädiktion in Abhängigkeit der Bewegungsrichtung im Vergleich zu den bisherigen Verfahren.

INDOOR-NUTZSIGNALE	Bewertung der 2. Zustandsvektorkomp. (Parameter nach Tab. 10.21)			Bewertung der 2. Zustandsvektorkomp. (reine Schwellenauswertung)		
	PDM1(1,1)	PDM2(1,1)	PDM3(1,1)	PDM1(1,1)	PDM2(1,1)	PDM3(1,1)
Parallele Bewegung	637	785	537	771	908	649
Vertikale Bewegung	551	586	485	640	716	552
Radiale Bewegung	1020	1187	881	1192	1375	1083
Tangentiale Bewegung	612	659	552	683	705	604
Alle Richtungen (Summe)	2820	3217	2455	3286	3704	2888

Tab. 10.27: Vergleich der Auswerteverfahren mit der Parametereinstellung nach Tab. 10.21 und einer Schwellenauswertung ohne zusätzliche einschränkende Kriterien.

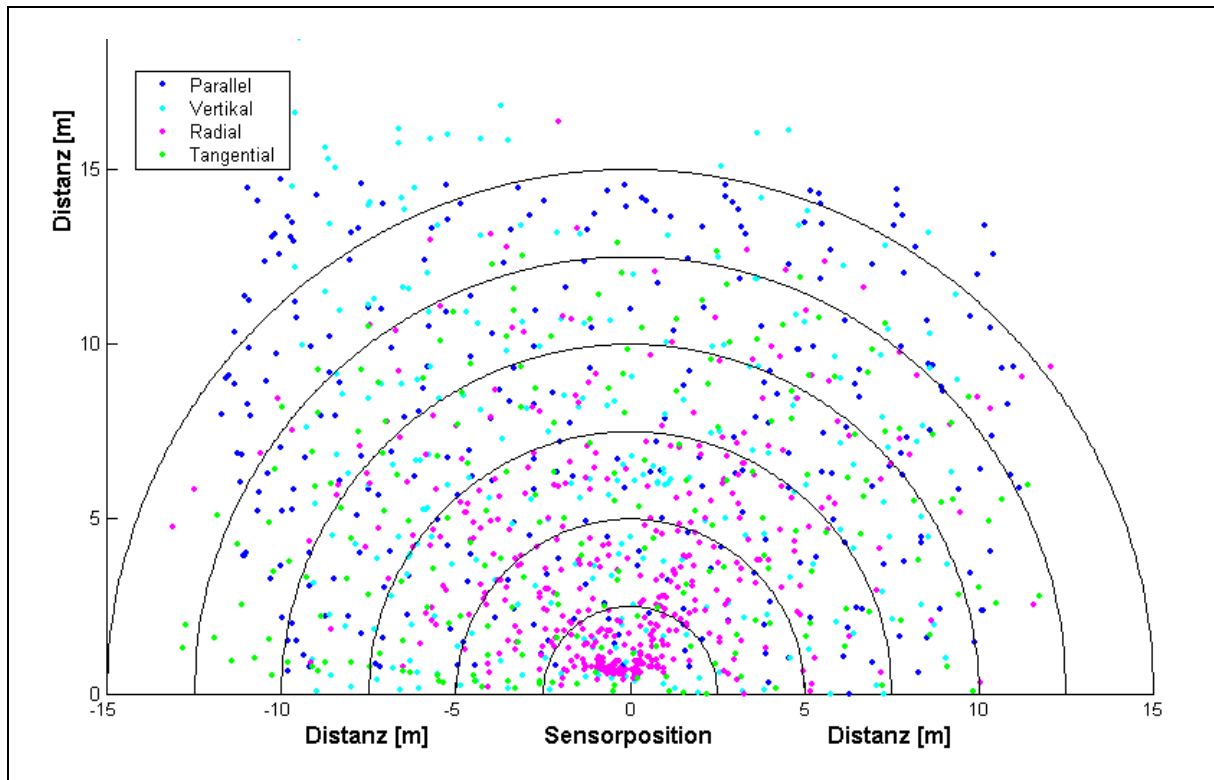


Abb. 10.14: Verteilung der Schaltpunkte im Erfassungsfeld des Bewegungsmelders bei Auswertung der 2. Komponente des Zustandsvektors der Prädiktion mit einer reinen Schwellenauswertung ($THR_{MIN} = 20$, $THR_{MAX} = 21$). Es wurde eine Alarmdauer von 2 s in Verbindung mit einer Immunitätszeit von 0.01 s gewählt.

Vergleicht man Abb. 10.13 und Abb. 10.14, so ist keine weitere Steigerung der Reichweite für radiale Gehrichtungen feststellbar, obwohl die Einzelalarmzahl auf den verschiedenen Sensorkanälen und in den unterschiedlichen Bewegungsrichtungen deutlich höher ist. Im Vergleich zu den in Abb. 10.8 und Abb. 10.9 dargestellten "Reichweiten" des Auswerteverfahrens nach Kap. 7 sowie zu der in Abb. 10.11 und Abb. 10.12 gezeigten Verteilung der Schaltpunkte bei Auswertung der 1. Zustandsvektorkomponente ist allerdings eine deutliche Reichweitensteigerung zu erkennen. Außerdem lässt die höhere Einzelalarmzahl auf eine schnellere Reaktion des Bewegungsmelders bei Bewegungen innerhalb des Erfassungsfeldes schließen.

10.4.5 Fehlschaltverhalten bei starken Störsignalen

Bisher wurden "typische" Störsignale aus dem Outdoor-Bereich zur Parametereinstellung der verschiedenen Auswerteverfahren verwendet. Mit den diesbezüglich optimierten Algorithmen wurden anschließend Nutzsignale, die indoor aufgezeichnet wurden, bewertet und die resultierende Schaltpunktverteilung im Erfassungsfeld des Wächters beurteilt. Zusammenfassend kann an dieser Stelle festgestellt werden, dass die Auswertung der Residuensequenz das unempfindlichste Schaltverhalten liefert. Deutliche Reichweiterhöhungen für die schwachen Signale aus radialen Bewegungsrichtungen ergeben sich dagegen im Vergleich zum Auswer-

teverfahren nach Kap. 7, wenn die 1. oder insbesondere die 2. Komponente des Zustandsvektors der Prädiktion zur Bewegungserkennung verwendet werden.

In diesem Abschnitt sollen die vorne beschriebenen Signalverarbeitungsstrategien mit den Parametereinstellungen nach Kap. 10.4.1 - 10.4.4 in Verbindung mit starken Störungen aus dem Outdoor-Testfeld bewertet werden. Diese Störsignale wurden unter extremen Umgebungsbedingungen aufgezeichnet.

Insbesondere im Erfassungsbereich des dritten Sensors war ein großer Busch vorhanden, dessen Äste vom Wind bewegt wurden. Der Abstand zwischen dem Busch und dem Bewegungsmelder betrug ca. 4 m. Daraus erklärt sich auch die größere Signalunruhe auf dem Sensorkanal 3 bei den vorne zur Optimierung verwendeten Störsignalen. Der erste Kanal zeigt dagegen ein sehr unauffälliges Verhalten. Die von Sensor 2 registrierten Turbulenzen liegen zwischen denen des ersten und des dritten Kanals.

In Tab. 10.28 ist die Fehlalarmzahl der drei im Bewegungsmelder vorhandenen Sensorkanäle in Verbindung mit den vorher beschriebenen Signalverarbeitungsstrategien zusammengefasst. Wie in den vorherigen Abschnitten ist sowohl die Summe der Einzelalarme als auch die im Testbetrieb (Alarmdauer 2.0 s, Immunitätszeit 0.01 s) aufgetretene Fehlalarmanzahl dokumentiert.

STARKE STÖRUNGEN IM OUTDOOR-TESTFELD	Kanal 1	Kanal 2	Kanal 3	Summe	Alarmanzahl (Testbetrieb)
Fehlalarmzahl bei Verwendung der Parameter nach Tab. 10.2 (Kap. 10.4.1)	0	0	19	19	19
Fehlalarmzahl für reine Schwellenabfrage (Kap. 10.4.1)	0	2	9	11	11
Fehlalarmzahl bei Auswertung der Residuensequenz (Kap. 10.4.2)	0	2	3	5	5
Fehlalarmzahl für 1. Zustandsvektorkomp. und Parameter nach Tab. 10.14	0	0	25	25	18
Fehlalarmzahl für 1. Zustandsvektorkomp. und reine Schwellenabfrage	0	0	11	11	11
Fehlalarmzahl für 2. Zustandsvektorkomp. und Parameter nach Tab. 10.21	0	0	12	12	8
Fehlalarmzahl für 2. Zustandsvektorkomp. und reine Schwellenabfrage	0	0	7	7	5

Tab. 10.28: Fehlalarmzahl bei Auswertung der 2. Zustandsvektorkomponente der Prädiktion in Verbindung mit starken Störsignalen aus dem Outdoor-Testfeld im Vergleich zu den bisherigen Signalverarbeitungsverfahren.

Vergleicht man die Anzahl der Fehlalarme in Tab. 10.28, kann als wesentliches Ergebnis festgehalten werden, dass die Auswertung der 1. und der 2. Zustandsvektorkomponente trotz eines empfindlicheren Schaltverhaltens für die Nutzsignale eine vergleichbare oder bessere Unterdrückung von Fehlschaltungen ermöglicht als das Auswerteverfahren nach Kap. 7. Entsprechende Resultate wurden auch in Verbindung mit weiteren Störsignalen aus unterschiedlichen Testfeldern bei variierenden Umgebungsbedingungen beobachtet.

Die Weiterführung dieser Untersuchungen, auch im Hinblick auf die weitere Optimierung und die Stabilität des Kalman-Filters für beliebige Nutz- und Störsignale, ist zum gegenwärtigen Zeitpunkt noch nicht abgeschlossen. Die bisherigen, auf Basis der zugrundeliegenden Testdaten ermittelten Resultate sind allerdings recht vielversprechend im Hinblick auf eine Angleichung der Schaltreichweiten für parallele und radiale Bewegungsrichtungen.

11 Zusammenfassung und Ausblick

Im Rahmen dieser Arbeit wurde zunächst ein Überblick zu bekannten und neuen Sensorkonzepten und Signalverarbeitungsstrategien zur Bewegungs- und Anwesenheitserkennung von Personen und das Applikationspotential für die Gebäudetechnik präsentiert.

Neben Mikrowellen- und Ultraschallbewegungsmeldern wurde schwerpunktmäßig die innovative PMD-Technologie vorgestellt, welche im Vergleich zu den vorgenannten Verfahren unter anderem eine höhere Ortsauflösung und damit eine differenziertere Bereichsüberwachung ermöglicht. Im Vergleich zu rein intensitätsmessenden Bildverarbeitungsverfahren bleibt die ursprüngliche 3D-Information der Szene erhalten, so dass der farbliche Kontrast zwischen dem Hintergrund und dem Objekt im Vordergrund keine Rolle spielt. Salopp formuliert ist es mit rein intensitätsmessenden Bildverarbeitungssensoren nicht möglich, den "schwarzen Adler über schwarzem Grund" zu erkennen. PMD-Sensoren werten dagegen die Phaseninformation der Lichtwellen aus und sind somit in der Lage, auch gleichfarbige Objekte anhand der zusätzlichen 3D-Information zu trennen. Bewegte Schattenbilder oder Lichtquellen, die in Intensitätsbildern entsprechende Grauwertänderungen hervorrufen, beeinflussen die Auswertung aufeinanderfolgender Tiefenbilder ebenfalls nicht.

Für größere Gesichtsfelder, die beispielsweise zur Präsenzkontrolle in Räumen oder für Überwachungsaufgaben im Outdoor-Bereich erforderlich sind, eignen sich aktive Sensorsysteme wegen des Leistungsbedarfs zum Aufbau des Fremdfeldes nur bedingt. Hier dominieren passive Sensorkonzepte. Zur Anwesenheitskontrolle im Indoor-Bereich wurde ein neuartiges Multisensorkonzept, der Bildwächter, vorgestellt, der die Vorzüge moderner CMOS-Bildsensoren mit denen eines Passiv-Infrarot-Bewegungsmelders zur Verbesserung des Schaltverhaltens und zur Erhöhung der Fehlalarmwahrscheinlichkeit kombiniert.

Aufbauend auf einigen grundlegenden Betrachtungen und einem Vergleich mit der im Sicherheitsbereich dominierenden CCD-Technologie wurde das Systemkonzept des Bildwächters abgeleitet und bekannte Bildverarbeitungsverfahren auf ihre Anwendbarkeit für die vorliegende Aufgabenstellung untersucht. Ein vielversprechender Ansatz zur Anwesenheitskontrolle ist die Auswertung der Varianzen korrespondierender Grauwertbereiche mit dem aus der Statistik bekannten F-Test.

Im Vergleich zur heuristischen Festlegung der Auswerteschwellen erfolgt die Schwellenfestlegung aus der für statistische Untersuchungen allgemein anerkannten und tabellierten F-Verteilung. Parameter bei der Dimensionierung ist die gerade noch akzeptable Wahrscheinlichkeit, dass eine Bewegung irrtümlicherweise erkannt wurde. Dasselbe Verfahren lässt sich ebenfalls zur Bewegungserkennung in aufeinanderfolgenden Tiefenbildern anwenden.

Das Auswerteverfahren und die dazu erforderliche lineare Grauwertinterpolation führen auf vergleichsweise einfache Berechnungsvorschriften, die für eine FPGA-Implementierung geeignet sind. Aufgrund der zur Onlineverarbeitung anfallenden Datenmengen und im Hinblick auf die Systemkosten ist eine hardwaregestützte Auswertung erforderlich. Die Umsetzung auf die Zielhardware und die Fusion der Daten aus unterschiedlichen Spektralbereichen, die nicht nur den CMOS-spezifischen Part sondern auch den PIR-Bewegungsmelder umfassen, ist Gegenstand weiterer Entwicklungsarbeiten im Zentrum für Sensorsysteme.

Ein weiterer Schwerpunkt dieser Arbeit beschäftigt sich mit den allseits bekannten Passiv-Infrarot-Bewegungsmeldern. Aufbauend auf einer Erläuterung des Funktionsprinzips und der einzelnen Komponenten stand zunächst die Wächterentwicklung im Vordergrund. Mit einem Simulationstool zum Design von Fresnellinsenarrays, die auf eine 2D-Abwicklung zurückzuführen sind, konnte der Entwicklungszeitraum von der Berechnung bis zur manuellen Fertigstellung eines Linsen-Prototypen deutlich verkürzt werden.

Die simulierten Erfassungsfelder zeigen eine überraschend gute Übereinstimmung mit dem tatsächlichen Gesichtsfeld, obwohl lediglich die geometrische Optik zur Ableitung der grundlegenden Beziehungen diente. Das Simulationstool wurde zur Dimensionierung des Fresnellinsenarrays eines kommerziell erhältlichen Bewegungsmelders verwendet. Allein durch die geschickte Anordnung und Auswahl der Linsen für den äußeren Entfernungsbereich konnte die Reichweite des Melders für radiale Bewegungsrichtungen im Vergleich zum Vorgängermodell um ca. 1 m gesteigert werden. Absolut gesehen ist das zunächst nicht viel, prozentual entsprach das immerhin einer Verbesserung von etwa 25%, was die Bedeutung der korrekten Auslegung des optischen Systems von PIR-Bewegungsmeldern unterstreicht.

Einen wesentlichen Einfluss auf das Schaltverhalten und die Störanfälligkeit von PIR-Bewegungsmeldern hat auch das Signalverarbeitungskonzept. Während bis vor wenigen Jahren die analoge Signalauswertung mit Fensterkomparatoren ggf. unter Beachtung weiterer Zeitbedingungen dominierte, sind moderne PIR-Bewegungsmelder mit Mikrocontrollern ausgestattet. Digitale Signalverarbeitungsverfahren ermöglichen eine deutlich differenziertere Auswertung und die Implementierung zusätzlicher Eigenschaften, die den Nutzen für den Anwender steigern. An dieser Stelle sei außerdem erwähnt, dass die Erfahrungen aus der Bewegungsmelderentwicklung im Outdoorbereich ohne weiteres auf Indoor-Präsenzmelder übertragbar sind.

Aufgrund der zentralen Bedeutung für das Schaltverhalten von PIR-Bewegungsmeldern wurden verschiedene digitale Signalverarbeitungsverfahren, von denen die vorliegende Arbeit nur eine Auswahl präsentiert, auf ihre Anwendbarkeit zur Verbesserung des Schaltverhaltens überprüft. Während bei früheren Untersuchungen zahlreiche Gehtests zur Beurteilung der Auswertestrategie erforderlich waren, wurde im Rahmen dieser Arbeit ein neuer Weg verfolgt.

Mit einem zum Datenlogger umgebauten Bewegungsmelder wurden typische Nutz- und Störsignale im Indoor- und Outdoorbereich aufgezeichnet. Zusammen mit den Nutzsignalen wurde außerdem die jeweils aktuelle Position der Person protokolliert. Nach einer entsprechenden Vorauswahl bilden diese Signalverläufe die Grundlage für eine Datenbank, die zur Überprüfung des Wächterverhaltens bei unterschiedlichen Bewegungsrichtungen und charakteristischen Störsignalen herangezogen werden kann.

Da alle Signalverarbeitungsverfahren in Verbindung mit diesen Testdaten bewertet wurden, konnten objektive Bewertungskriterien zum Vergleich der implementierten Auswertekonzepte abgeleitet werden. Neben der Fehlalarmanzahl ist das insbesondere die Schaltreichweite des Geräts in Abhängigkeit der Bewegungsrichtung.

Zur Parametrierung der untersuchten Algorithmen leistete das gerade beschriebene Entwicklungssystem ebenfalls eine wertvolle Hilfe. Durch die Bereitstellung der von den Algorithmen extrahierten Signalcharakteristiken kann eine für die jeweilige Signalverarbeitung optimale Parametereinstellung ermittelt werden, die im Testfeld unter realen Umgebungsbedingungen lediglich verifiziert werden muss. Natürlich ist die rein rechnergestützte Parameteroptimierung nur so repräsentativ wie die zugrundeliegenden Prüfsignale, so dass eine ständige Erweiterung und Pflege der Datenbank erforderlich ist.

Das gesamte Entwicklungssystem lieferte einen wesentlichen Beitrag zur Realisierung eines neuen, qualitativ hochwertigen Bewegungsmeldermodells, das mittlerweile in Serienstückzahlen produziert wird. Durch die intensive Zusammenarbeit im Rahmen des zugrundeliegenden Industrieprojektes ist es gelungen, ein vergleichsweise stabiles Schaltverhalten auch unter ungünstigen Umgebungsbedingungen zu erzielen, wie zahlreiche Testmessungen in unterschiedlichen Outdoor-Prüffeldern beweisen.

Aufgrund einer neuartigen, patentierten Sensor-Linse-Anordnung, welche die verstärkte Mehrfachnutzung einzelner Fresnellinsen im frontalen Bereich des Wächters ermöglicht, tritt im für das "Draufzugehen" wichtigen Erfassungsbereich eine hohe Felddichte auf. Daraus resultiert ein sehr geschlossenes Gesichtsfeld, das in Verbindung mit der implementierten Signalverarbeitungsstrategie das vergleichsweise sehr gute Schaltverhalten bewirkt.

Nachdem ein Datenlogger mit der Zielhardware zur Verfügung stand, konnten durch die vorwiegend rechnergestützte Entwicklung und Optimierung des Auswerteverfahrens weitere Verbesserungen der Wächterhardware und -software parallel durchgeführt werden. Der Entwicklungszeitraum für den neuen Wächter wurde dadurch deutlich verkürzt.

In den letzten Jahren wurden erhebliche Fortschritte auf dem Gebiet der Mikroprozessortechnik und bei programmierbaren logischen Bausteinen erzielt. Der daraus resultierende Preisverfall vorhergehender Generationen wird sich sicherlich auch in den nächsten Jahren fortset-

zen. Während die Onlineverarbeitung beim Bildwächter erhebliche Ansprüche an die Geschwindigkeit der zur Bildauswertung verwendeten Komponenten stellt, sind die Anforderungen hinsichtlich der Auswertegeschwindigkeit bei PIR-Bewegungsmeldern vergleichsweise gering. Da andererseits immer leistungsfähigere Mikrocontroller in dem zur Wächterentwicklung interessanten Preissegment angesiedelt werden, können zukünftige Produkte auch komplexere Auswertestrategien enthalten.

Mit Hinblick auf die allgemeine technische Entwicklung wurde daher für zukünftige Wächtergenerationen ein vielversprechender modellbasierter Ansatz zur Signalverarbeitung untersucht, der auf den Grundlagen der Estimationstheorie basiert. Während Kalman-Filter selbst sowie deren Anwendung für regelungstechnische Aufgabenstellungen oder für die Fernerkundung der Erde seit langem bekannt sind, ist die Übertragung auf die vorliegende Aufgabenstellung nach gegenwärtigem Kenntnisstand des Verfassers eine neue Applikation.

Im Anschluss an die grundsätzliche Beschreibung der Idee und der prinzipiellen Funktion des Kalman-Filters wurde zunächst ein zeitdiskretes lineares Zustandsraummodell des realen Prozesses hergeleitet. Dieses beinhaltet ein allgemein anwendbares Modell zur Signalverfolgung sowie ein Störmodell, das implizit die Übertragungsfunktion eines pyroelektrischen Dualsensors zur Modellierung von Störsignalen verwendet.

Zur Einstellung des Störmodells dienten die untere und obere Grenzfrequenz des pyroelektrischen Detektors sowie typische Störsignalvarianzen. Für eine höhere Modellierungsgüte ist es alternativ möglich, die nachfolgende analoge Filter- und Verstärkerstufe anhand der charakteristischen Pole der resultierenden Bandpass-Übertragungsfunktion einzubeziehen, ohne dass der Zustandsvektor des Kalman-Filters vergrößert werden muss.

Durch die gewählte Form der Modellierung erfolgte die Trennung von Nutz- und Störsignalen, basierend auf dem zugrundeliegenden Störmodell. Durch die Vorgabe einer oberen Grenze für die Störvarianz konnte sogar ein gewisses, adaptives Verhalten realisiert werden. Weitere Verbesserungen bezüglich der adaptiven Störungsunterdrückung sind zu erwarten, wenn die Varianz des Eingangsräuschprozesses im Kalman-Filter mitgeschätzt wird. Dieser Weg wurde im Rahmen der vorliegenden Arbeit nicht weiter verfolgt.

Im weiteren Verlauf wurden interne Zustände des Kalman-Filters, namentlich die Residuensequenz sowie die erste und zweite Komponente des Zustandsvektors der Prädiktion zur Bewegungserkennung verwendet. Während bei der Residuensequenz nur eine einfache Schwellwertabfrage implementiert wurde, erfolgte die Auswertung der störsignalbereinigten Zustandsvektorkomponenten anhand der aufwendigeren Bewertungskriterien des zuvor beschriebenen konventionellen Signalverarbeitungsverfahrens.

Es zeigte sich erwartungsgemäß, dass die dabei verwendeten Auswerteschwellen deutlich reduziert werden konnten, was auf die Wirkung der Störungsunterdrückung zurückzuführen ist. Allerdings werden durch die Störungsunterdrückung auch Nutzsignale bedämpft, so dass die Empfindlichkeitssteigerung infolge der Schwellwertabsenkung nicht unmittelbar als Empfindlichkeitssteigerung der Bewegungserkennung interpretiert werden kann.

In Verbindung mit typischen Wächtersignalen aus dem Indoor- und dem Outdoor-Bereich wurden die drei Verfahren im Vergleich zum Schaltverhalten einer vorher beschriebenen, konventionellen Auswertestrategie bewertet. Die Auswertung der Residuensequenz, in der sich insbesondere die unerwarteten Signaländerungen auswirken, schnitt dabei am schlechtesten ab. Bei Bewertung der ersten Komponente des vorhergesagten Zustandsvektors zur Bewegungserkennung ergab sich bereits ein besseres Schaltverhalten für Nutzsignale in Verbindung mit einer vergleichbaren Fehlalarmanzahl bei starken Störungen.

Ein deutlich besseres Schaltverhalten bei gleichzeitig erhöhter Fehlalarmimmunität erzielte die Auswertung der zweiten Zustandsvektorkomponente der Prädiktion, die der zeitlichen Ableitung der ersten Komponente entspricht. Für die beiden letztgenannten Verfahren ergaben sich in allen Bewegungsrichtungen mehr Auslöser, auch für radiale Bewegungsrichtungen. Anhand der ebenfalls vorliegenden Positionsdaten konnte außerdem festgestellt werden, dass die Schaltreichweiten für radiale Gehrichtungen bei den zugrundeliegenden Bewegungsmeldersignalen größer waren als die entsprechenden Reichweiten bei konventioneller Signalauswertung.

Bei den in dieser Arbeit vorgestellten Auswertefahren für PIR-Bewegungsmelder erfolgte die Festlegung der Schwellwerte und Varianzen jeweils unter Beachtung des empirischen Datenmaterials. In Verbindung mit einem zusätzlichen Sprungestimator und Hypothesentester ist es möglich, einen analytischen Zusammenhang zwischen der Sprunghöhe und dem Zeitpunkt des Auftretens der Signaländerung abzuleiten. Die Beurteilung der Auswirkungen auf die Schaltreichweite wird an dieser Stelle weiterführenden Untersuchungen überlassen.

Gleiches gilt für die vorher beschriebene Schätzung der Störvarianz am Systemeingang und für die kombinierte Auswertung mehrerer Sensorkanäle. Da viele Störungen großflächig auftreten und daher von mehreren Detektoren gleichzeitig registriert werden, bietet die Auswertung der Kreuzkorrelationsterme in der Kovarianzmatrix eines entsprechend erweiterten Systemmodells sicherlich weiteres Optimierungspotential zur Empfindlichkeitssteigerung und Störungsunterdrückung.