

Diese Arbeit wurde vorgelegt am
Institut für Textiltechnik der RWTH Aachen University
Univ.-Prof. Prof. h.c. (MGU)
Dr.-Ing. Dipl.-Wirt. Ing. Thomas Gries

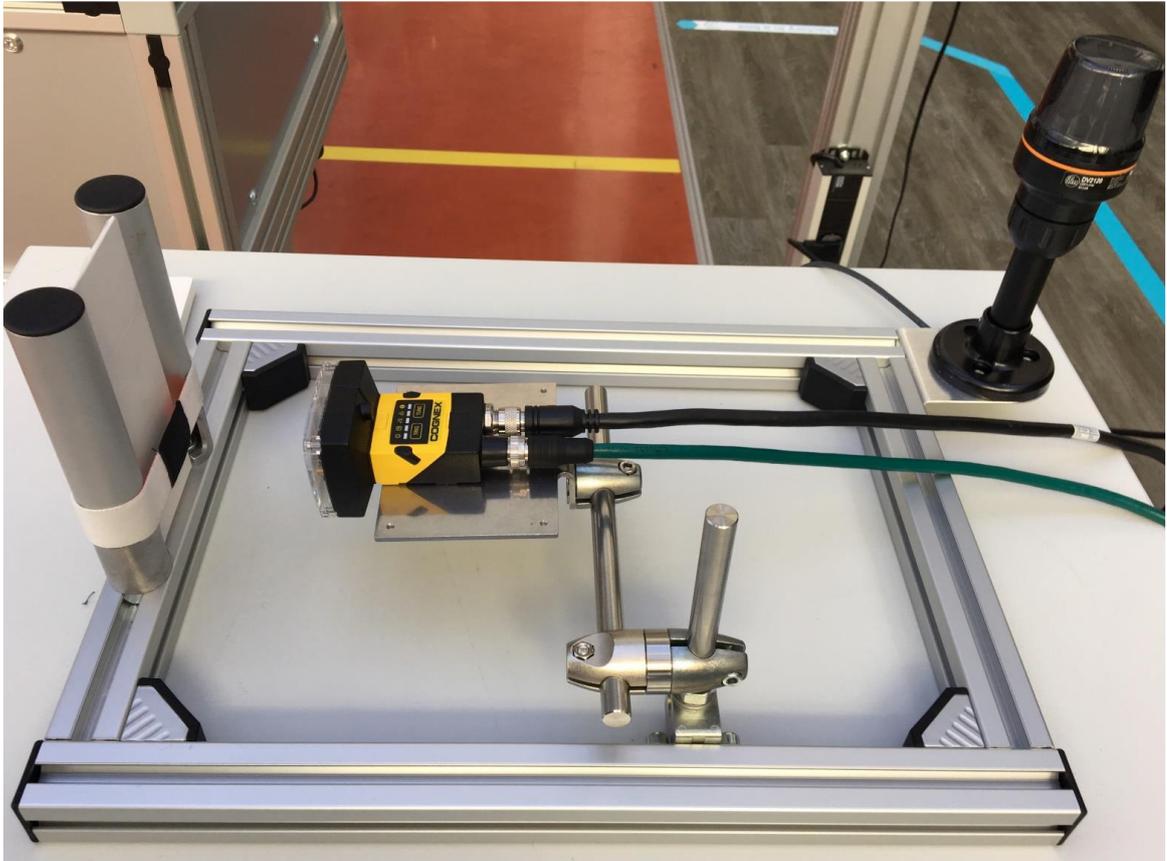
Industrielle Bildverarbeitung zur Qualitätskontrolle in Produktions-
linien: Entwicklung einer Entscheidungslogik zur Anwendungsfall-
spezifischen Auswahl von Hard- und Software

Vorgelegt als: Masterarbeit
Von: Michael Müller
Matr.-Nr. 345306

1. Prüfer: Univ.-Prof. Prof. h.c. (MGU) Dr.-Ing. Dipl.-Wirt. Ing. Thomas Gries
2. Prüfer: Dr.-Ing. Dieter Veit

Ansprechpartner: Kai Müller, M. Sc. RWTH

Aachen, Mai 2020



2020

Masterarbeit

**Industrielle Bildverarbeitung zur
Qualitätskontrolle in Produktionslinien:
Entwicklung einer Entscheidungslogik
zur Anwendungsfall-spezifischen
Auswahl von Hard- und Software**

Michael Müller

Kurzfassung

Ziel der Arbeit:

Das Ziel der Arbeit ist die Entwicklung einer Entscheidungslogik zur Anwendungsfall-spezifischen Auswahl von Hard- und Software für Bildverarbeitungssysteme zur Qualitätsüberprüfung in industrieller Produktion. Auf Hardwareseite werden dabei die Komponenten Kamera, Objektiv und Beleuchtungssystem betrachtet. Auf Softwareseite wird in Abhängigkeit eines Anwendungsfalls entschieden, ob konventionelle Algorithmen oder Methoden des Deep Learning besser geeignet sind.

Lösungsweg:

Im Rahmen einer Literaturrecherche werden zunächst relevante Beschreibungsgrößen zur standardisierten Charakterisierung von Technologien und Anwendungsfällen aufgearbeitet. Dabei werden weiterhin Interdependenzen zwischen einzelnen Technologiekomponenten und Eigenschaften des Anwendungsfalls identifiziert. Mittels einer Marktrecherche wird eine Datenbank mit konkreten Produktinformationen aufgebaut. Auf diesen Arbeitsschritten aufbauend, wird ein Regelwerk zur Auswahl der Hard- und Softwaretechnologien abgeleitet und an einem Anwendungsfall im Digital Capability Center Aachen überprüft. Die Entscheidungslogik zur Auswahl von Hardwarekomponenten wird abschließend in eine nutzerfreundliche Computerapplikation überführt.

Zentrale Ergebnisse:

Zur Auswahl geeigneter Hardwarekomponenten wird eine fünfstufige Entscheidungslogik entwickelt und als Softwareapplikation implementiert, welche dem Nutzer in Abhängigkeit des spezifizierten Anwendungsfalls geeignete Komponenten vorschlägt und diese nach Listenpreis priorisiert. In einer simulativen Bewertung werden damit Komplexitätsreduktionen zwischen 73 und 98 % und Kosteneinsparungen zwischen 46 und 93 % erzielt. Eine Entscheidung zwischen Deep Learning und konventionellen Algorithmen kann anhand der gegebenen Entwicklungsumstände sowie der Komplexität von Bildmerkmalen getroffen werden.

Schlagwörter:

Industrielle Bildverarbeitung, Technologieauswahl, Deep Learning, Maschinelles Sehen, Qualitätskontrolle

Inhaltsverzeichnis

Inhaltsverzeichnis	IV
Abbildungsverzeichnis	VI
Tabellenverzeichnis	IX
Abkürzungs- und Formelverzeichnis.....	X
1 Einleitung.....	1
1.1 Problemstellung.....	1
1.2 Zielsetzung.....	1
1.3 Aufbau der Arbeit	2
2 Einführung in die industrielle Bildverarbeitung	3
2.1 Existierende Ansätze zur Technologieauswahl.....	3
2.2 Ansatz zur Technologieauswahl im Rahmen dieser Arbeit	5
3 Vorstellung und Parametrisierung von Bildverarbeitungstechnologien	7
3.1 Hardware.....	7
3.1.1 Kamerasensor	7
3.1.2 Objektiv	19
3.1.3 Beleuchtungssystem	26
3.1.4 Zwischenergebnis Hardwaretechnologie	36
3.2 Bildverarbeitungsalgorithmen	38
3.2.1 Konventionelle Bildverarbeitungsalgorithmen	38
3.2.2 Maschinelles Lernen und Deep Learning für die Bildverarbeitung	50
4 Marktrecherche von Bildverarbeitungstechnologien	57
4.1 Hardware.....	57
4.1.1 Verfügbarkeit von Produktinformationen.....	57
4.1.2 Produktübersicht Kameras	58
4.1.3 Produktübersicht Objektive.....	60
4.1.4 Produktübersicht Beleuchtungssysteme.....	64
4.2 Software.....	64
4.2.1 Open Source Software	64
4.2.2 Kommerzielle Software.....	66
5 Formale Charakterisierung von Anwendungsfällen industrieller Bildverarbeitung.....	68
5.1 Allgemeine Charakterisierung von Anwendungsfällen gemäß VDI / VDE 2632 Blatt 2	68
5.2 Konkrete Charakterisierung von Anwendungsfällen im Rahmen der Entscheidungslogik	69

6	Methodik zur Technologieauswahl	72
6.1	Erläuterung der hardwarebezogenen Entscheidungslogik	72
6.1.1	Kameravorauswahl.....	73
6.1.2	Detaillierte Kameraauswahl.....	75
6.1.3	Objektivvorauswahl	78
6.1.4	Detaillierte Objektivauswahl.....	78
6.1.5	Auswahl des Beleuchtungssystems.....	85
6.2	Erläuterung der softwarebezogenen Entscheidungslogik	91
7	Anwendung der Entscheidungslogik.....	96
7.1	Hardwareauswahl.....	97
7.2	Überprüfung der Eignung der Hardware	98
7.3	Softwareauswahl.....	99
7.4	Überprüfung der Eignung verschiedener Softwarelösungen	100
7.4.1	Konventionelle Bildverarbeitung mit Cognex In-Sight Explorer	101
7.4.2	Kombination aus konventioneller Bildverarbeitung und Deep Learning mit Cognex In-Sight Vidi	101
7.4.3	Deep Learning mit Teachable Machine	103
8	Umsetzung der Auswahllogik als Softwareapplikation	104
8.1	Funktionsweise der softwaregestützten Auswahllogik	104
8.2	Bewertung des Nutzens der Auswahllogik.....	106
9	Zusammenfassung.....	109
10	Ausblick	111
	Literaturverzeichnis	112
11	Anhang.....	122
11.1	Beschreibungsparameter von Hardwarekomponenten	122
11.2	Übersicht konventioneller Bildverarbeitungsalgorithmen.....	123
11.3	Übersicht erfolgreicher Netzarchitekturen der ImageNet Challenge	125
11.4	Anwendung der Auswahllogik, Schritte 1 und 2	126
11.5	Anwendung der Auswahllogik, Schritt 5.....	127
11.6	Berechnungsdetails der angepassten Sichtfeldgröße	127
11.7	Benutzeroberfläche Entscheidungslogik.....	128
	Erklärung	131

Abbildungsverzeichnis

Abb. 2.1: Komponenten eines optischen Inspektionssystem i. A. a. [MPZ03]	3
Abb. 2.2: Konzeption der Entscheidungslogik zur Technologieauswahl	5
Abb. 3.1: Bildaufnahmeprozess i. A. a. [Jäh12]	8
Abb. 3.2: Aufbau CCD-Sensor i. A. a. [Mat06].....	9
Abb. 3.3: Aufbau CMOS-Sensor i. A. a. [DSS11]	10
Abb. 3.4: Quantisierung eines kontinuierlichen Spannungssignals i. A. a. [Jäh12].....	11
Abb. 3.5: Beispielformate und Bilddiagonalen von Bildsensoren i. A. a. [Ste18]	13
Abb. 3.6: Beispiele unterschiedlicher Graustufen i. A. a. [Ste18].....	14
Abb. 3.7: Konventioneller Aufnahmemodus (links) und Overlapped Aufnahmemodus (rechts) i. A. a. [DSS11].....	15
Abb. 3.8: Aufnahme Kerzenlicht bei geringem Dynamikumfang und unterschiedlichen Belichtungszeiten (links) sowie hohem Dynamikumfang (rechts) [Ste18].....	17
Abb. 3.9: Strahlengang einer dünnen Linse i. A. a. [ST02]	20
Abb. 3.10: Relative Öffnung eines Objektivs i. A. a. [DSS11]	22
Abb. 3.11: Beispiel normierte Linienabbildung (hier skaliert abgebildet) zur MTF-Bestimmung [Ste18]	24
Abb. 3.12: Beispiel Helligkeitsverteilung Techspec Compact 6 mm Objektiv bei unterschiedlichen Blendenzahlen i. A. a. [Edm14].....	24
Abb. 3.13: Unterschiedliche Arten der Verzeichnung i. A. a. [Fer19]	25
Abb. 3.14: Zusammenhang zwischen Sensor- und Objektivformat i. A. a. [Mat06].....	25
Abb. 3.15: Zusammenhang zwischen photometrischen Größen i. A. a. [Jah06].....	27
Abb. 3.16: Mögliche Arten der Interaktion von Licht mit Objekten i. A. a. [Jah06].....	28
Abb. 3.17: Vergleich unterschiedlicher Lichtquellen zur industriellen Bildverarbeitung von 1 = schlecht bis 5 = sehr gut i. A. a. [KUH14; Jah06].....	29
Abb. 3.18: Spektrale Emission von weißem LED-Licht i. A. a. [Lic18].....	30
Abb. 3.19: Schalten eines LED-Blitzlichts i. A. a. [Ste18].....	30
Abb. 3.20: Unterschiedliche Beleuchtungswinkel i. A. a. [Per10]	32
Abb. 3.21: Farbdarstellung in monochromatischen Bildern in Abhängigkeit der Lichtfarbe i. A. a. [Ste18].....	33
Abb. 3.22: Relative Helligkeitsverteilung eines Ringlichtes in Abhängigkeit vom Abstand zum Werkstück (LWD) [CCS18]	35
Abb. 3.23: Übersicht und Interdependenzen von erarbeiteten Beschreibungsgrößen (Eigendarstellung)	37
Abb. 3.24: Beispiel einer Kontrastnormierung [DSS11]	41
Abb. 3.25: Referenzbild (links), zu untersuchendes Bild (Mitte), Bilddifferenz (Rechts) [DSS11]	42
Abb. 3.26: Ausgangsbild (links), Bild nach Glättungsfilter (mittig), Bild nach Medianfilter (rechts) [DSS11]	43

Abb. 3.27: Ausgangsbild (links), Bild nach Kantenfiter (rechts) [DSS11]....	43
Abb. 3.28: Ausgangsbild (links), Bild nach Koordinatentransformation (rechts) [SUW18].....	44
Abb. 3.29: Histogramm-basierte Schwellwertanalyse zur Bildsegmentierung [DSS11]	46
Abb. 3.30: Parameter eines Kantenmodells i. A. a. [Ste06; DSS11]	48
Abb. 3.31: Grundsätzlicher Aufbau eines Perzeptrons i. A. a. [Ste06]	52
Abb. 3.32: Struktur eines künstlichen Neurons mit einer Sigmoid Aktivierungsfunktion i. A. a. [Bur19].....	52
Abb. 3.33: Struktur eines Convolutional Neural Network i. A. a. [LGT+97] .	54
Abb. 3.34: Unterschied zwischen konventioneller Bildverarbeitung und Deep Learning i. A. a. [WMZ+18].....	55
Abb. 4.1: Statistische Verteilung von Auflösung, Pixelgröße, Sensordiagonale und Bildrate von Kameras (Eigendarstellung)	59
Abb. 4.2: Übersicht monochromatischer Kameras (Eigendarstellung)	60
Abb. 4.3: Pinbaugruppe betrachtet durch ein Festbrennweitenobjektiv (links) und ein telezentrisches Objektiv (rechts) [Fer19]	61
Abb. 4.4: Übersicht Festbrennweitenobjektive (Eigendarstellung)	62
Abb. 4.5: Festbrennweiteobjektiv und telezentrisches Objektiv i. A. a. [Ste18]	63
Abb. 4.6: Übersicht telezentrische Objektive (Eigendarstellung).....	64
Abb. 5.1: Parameter zur Charakterisierung des Anwendungsfalls	71
Abb. 6.1: Logik zur Vorauswahl geeigneter Kameratechnologien	74
Abb. 6.2: Logik zur detaillierten Kameraauswahl.....	77
Abb. 6.3: Logik zur Objektivvorauswahl.....	78
Abb. 6.4: Zusammenhang von Sensor- und Sichtfeldgröße.....	80
Abb. 6.5: Detaillierte Logik zur Auswahl von Festbrennweitenobjektiven..	82
Abb. 6.6: Detaillierte Logik zur Auswahl von telezentrischen Objektiven ..	84
Abb. 6.7: Logik zur Auswahl der Beleuchtungsanordnung.....	87
Abb. 6.8: Übersicht möglicher Beleuchtungsanordnung Teil 1 i. A. a. [Ste18; Per10].....	88
Abb. 6.9: Übersicht möglicher Beleuchtungsanordnung Teil 2 i. A. a. [Ste18; Per10].....	89
Abb. 6.10: Übersicht möglicher Beleuchtungsanordnung Teil 3 i. A. a. [Ste18; Per10].....	90
Abb. 6.11: Logik zur Auswahl eines geeigneten Softwareansatzes	93
Abb. 6.12: Entwicklungsaufwand für konventionelle Bildverarbeitung und Deep Learning in Abhängigkeit der Aufgabenkomplexität (Eigendarstellung)	95
Abb. 7.1: Beispiel des am DCC produzierten Armbandes (links Vorderseite, rechts Rückseite)	96
Abb. 7.2: Beispiele auftretender Produktionsfehler beim Vernähen der Armbandtaschen.....	96
Abb. 7.3: Spezifikationsübersicht Cognex In-Sight 2000	97
Abb. 7.4: Prototyp Prüfstand mit Cognex In-Sight 2000.....	98
Abb. 7.5: Aufnahme Armbandtasche mit Cognex In-Sight 2000	99

Abb. 7.6:	Aufnahme Armbandtasche mit Cognex In-Sight 7802	99
Abb. 7.7:	Konventionelle Bildverarbeitung mit Cognex In-Sight Explorer	101
Abb. 7.8:	Anlernen fehlerhafter Nahtbereiche mit Cognex In-Sight Vidi..	102
Abb. 8.1:	Screenshot Auswahltool: Spezifizierung des Anwendungsfalls	104
Abb. 8.2:	Screenshot Auswahltool: Ergebnisvisualisierung und Spezifizierung weiterer Auswahlparameter.....	105
Abb. 8.3:	Screenshot Auswahltool: Tabellarische Übersicht geeigneter Kameras	106
Abb. 9.1:	Entwickeltes und angewandtes Vorgehen zur Technologieauswahl	109
Abb. 11.1:	Beispielhafter Kurvenverlauf eines SNR Graphen i. A. a. [EMVA 1388].....	122
Abb. 11.2:	Übersicht konventioneller Bildverarbeitungsalgorithmen Teil 1 (Eigendarstellung i. A. a. [DSS11; SUW18; PKB13; Ste06]).....	123
Abb. 11.3:	Übersicht konventioneller Bildverarbeitungsalgorithmen Teil 2 (Eigendarstellung i. A. a. [DSS11; SUW18; PKB13]).....	124
Abb. 11.4:	Anwendung der Logik zur Auswahl einer geeigneten Beleuchtung	127
Abb. 11.5:	Screenshot Berechnung Sichtfeld	127
Abb. 11.6:	Benutzeroberfläche Entscheidungslogik, Schritt 1.....	128
Abb. 11.7:	Benutzeroberfläche Entscheidungslogik, Schritt 3.....	128
Abb. 11.8:	Benutzeroberfläche Entscheidungslogik, Schritt 4 Teil 1 (Beispiel Festbrennweitenobjektiv)	128
Abb. 11.9:	Benutzeroberfläche Entscheidungslogik, Schritt 4 Teil 2 (Beispiel Festbrennweitenobjektiv)	129
Abb. 11.10:	Benutzeroberfläche Entscheidungslogik, Schritt 5.....	130

Tabellenverzeichnis

Tab. 2.1:	Literatur zu Auswahllogiken von Bildverarbeitungstechnologien..	4
Tab. 3.1:	Übersicht gebräuchlicher Kameraschnittstellen i. A. a. [AEJ18].	18
Tab. 3.2:	IP-Schutzarten [DIN 60529]	19
Tab. 4.1:	Verfügbarkeit von Produktinformationen (Eigendarstellung)	58
Tab. 4.2:	Übersicht kommerzieller BV-Softwareprodukte (Eigendarstellung)	67
Tab. 6.1:	Mögliche Variationsgrade bei Objektmerkmalen (Eigendarstellung)	94
Tab. 7.1:	Auswahlschritt 3 zur Vorauswahl eines Objekttyps	97
Tab. 8.1:	Kosteneinsparung und Komplexitätsreduktion der Auswahllogik bei Kameraauswahl.....	107
Tab. 8.2:	Kosteneinsparung und Komplexitätsreduktion der Auswahllogik bei Objektivauswahl	108
Tab. 11.1:	Übersicht erfolgreicher Architekturen der Imagenet Challenge [CPC17]	125
Tab. 11.2:	Anwendung der Hardware-Auswahllogik, Schritte 1 und 2	126

Abkürzungs- und Formelverzeichnis

b	Bildweite
B	Bildgröße
BV	Bildverarbeitung
c	Lichtgeschwindigkeit
cd	Candela
C	Durchmesser Zerstreuungskreis (auch: Pixelgröße)
CCD	Charge-coupled device
CDS	Correlated Double Sampling
CMOS	Complimentary metal-oxide-semiconductor
CNN	Convolutional Neural Network
DL	Deep Learning
DN	Digital Number
DSNU	Dark signal nonuniformity
E	Energie
E_v	Beleuchtungsstärke
f	Brennweite
f	Frequenz
FOV	Field of View / Sichtfeld
g	Gegenstandsweite
g	Grauwert
g_0	Schwarzwert
G	Objektgröße
h	Plancksches Wirkungsquantum
i. A. a.	in Anlehnung an
k	Blendenzahl
I_v	Lichtstärke
lp	Linienpaare
lx	Lux
mm	Millimeter
ML	Maschinelles Lernen / Machine Learning
MOD	Minimale Objektdistanz
MP	Megapixel
MTF	Modulationstransferfunktion
nm	Nanometer
px	Pixel
PRNU	Photo response nonuniformity
sr	Steradian
SNR	Signal-Noise-Ratio, Signal / Rausch-Verhältnis
V	Volt
W	Watt
z. B.	zum Beispiel
β	Vergrößerung (Magnification)
Δg	Schärfentiefe
λ	Wellenlänge
μm	Mikrometer
ρ	Reflexionsanteil
τ	Transmissionsanteil

1 Einleitung

In der Einleitung wird eine kurze Übersicht über Problemstellung, Zielsetzung und Aufbau der vorliegenden Arbeit gegeben.

1.1 Problemstellung

Industrielle Bildverarbeitungsmethoden ermöglichen produzierenden Unternehmen die automatisierte Qualitätsüberwachung von Produktionslinien, wodurch Produktivitätssteigerungen und die Senkung von Ausschusskosten erreicht werden können [Tel06]. Typische Anwendungsfälle von optischen Inspektionssystemen beinhalten die Überprüfung von Kennzeichnungen, Produktmaßen, Oberflächenqualitäten oder Vollständigkeitskontrollen [DSS11; Vdm16]. Mit einem durchschnittlichen jährlichen Wachstum von 6,5 % wird ein globaler Markt von 14,7 Mrd. USD für industrielle Bildverarbeitungsanwendungen im Jahr 2025 vorausgesagt [Gvr18].

Die Zuverlässigkeit des Ergebnisses einer Bildverarbeitungsapplikation hängt dabei maßgeblich von der ausgewählten Hardware zur Bildaufnahme und den Algorithmen zur Bildanalyse (Software) ab. Diese Technologiekombination aus Hard- und Software muss spezifisch auf den Anwendungsfall angepasst werden [Sac12]. Aufgrund des breiten Spektrums von Anwendungsfällen existiert eine hohe Anzahl an unterschiedlichen Technologielösungen. Durch jüngere technische Entwicklungen wie Deep Learning gewinnt der Angebotsmarkt der industriellen Bildverarbeitung dabei zusätzlich an Komplexität. Diese Komplexität erschwert die effiziente Auswahl und Implementierung von Bildverarbeitungssystemen für Anwender, welche mit der Thematik wenig vertraut sind.

Während Hersteller von Bildverarbeitungssystemen Kunden bei der Auswahl passender Hard- und Softwarelösungen im Rahmen ihrer eigenen Vertriebsstrategie mit entsprechender Produktexpertise unterstützen, existiert in der wissenschaftlichen Literatur noch kein allgemeiner, unabhängiger Ansatz zur Auswahl von optischen Kontrollsystemen für produzierende Unternehmen.

1.2 Zielsetzung

Ziel dieser Arbeit ist es, eine Entscheidungslogik zur Auswahl von Bildverarbeitungstechnologien zur Qualitätsüberwachung von Werkstücken in der industriellen Fertigung zu entwickeln. Die Entscheidungslogik versetzt den Anwender in die Lage, unterschiedliche Anwendungsfälle visueller Qualitätskontrolle systematisch zu kategorisieren und Empfehlungen zur Auswahl von passender Hardware und

Software zu geben. Dazu werden konkret die folgenden beiden Forschungsfragen geklärt:

- 1) **Hardware:** Wie lassen sich eine geeignete Kamera, zugehöriges Objektiv und Beleuchtung eines Bildverarbeitungssystems in Abhängigkeit des Anwendungsfalls bestimmen?
- 2) **Software:** Wie lässt sich in Abhängigkeit des Anwendungsfalls entscheiden, ob für die Bildverarbeitung konventionelle Algorithmen oder Deep Learning verwendet werden soll?

Im Rahmen dieser Masterarbeit beschränken sich die analysierten Bilderfassungstechnologien auf digitale 2D-Bilderfassung im menschlich sichtbaren Spektralbereich zur Qualitätsüberprüfung industrieller Waren. Technologien wie die 3D Bilderfassung, Zeilenkameras, hyperspektrale Bildverarbeitung oder industrielle Bildverarbeitung zur aktiven Prozesssteuerung sind nicht im Bearbeitungsumfang enthalten.

1.3 Aufbau der Arbeit

Zu Beginn der Arbeit wird in Kapitel 2 eine kurze Übersicht zu existierenden Ansätzen zur Technologieauswahl gegeben und anschließend eine eigene Methodik zur Entwicklung der Auswahllogik entworfen. In Kapitel 3 wird der Stand der Technik visueller Inspektionssysteme und zugehöriger Algorithmen aufgearbeitet. Dabei werden Beschreibungsparameter zum formalen Vergleich der Technologien definiert sowie Interdependenzen zwischen diesen Parametern aufgezeigt. Darauf aufbauend wird in Kapitel 4 unter Verwendung eines Webcrawlers eine Marktrecherche zu existierenden Technologien durchgeführt. Die Technologien werden dabei, soweit Informationen verfügbar sind, anhand der zuvor definierten Beschreibungsparameter verglichen und in eine Produktdatenbank überführt. In Kapitel 5 werden in Anlehnung an Richtlinie VDI / VDE 2632 Blatt 2 Parameter zur formalen Charakterisierung von Anwendungsfällen definiert. Die in den vorherigen drei Kapiteln erworbenen Kenntnisse werden in Kapitel 6 zu einer Entscheidungslogik zur Anwendungsfall-spezifischen Auswahl von Hard- und Software gemäß Forschungsfrage 1 und 2 zusammengeführt. Diese Entscheidungslogik wird in Kapitel 7 anhand eines konkreten Anwendungsfalls am DCC Aachen veranschaulicht und überprüft. Die Entscheidungslogik zur Auswahl geeigneter Hardware (Forschungsfrage 1) wird in Kapitel 8 zudem in eine benutzerfreundliche Computerapplikation überführt. Mittels dieser Applikation wird abschließend der Nutzen, der durch die entworfene Entscheidungslogik generiert wird, quantitativ bewertet.

2 Einführung in die industrielle Bildverarbeitung

Der Prozess der industriellen Bildverarbeitung kann in die beiden grundlegenden Schritte der Bilderfassung und der Bildverarbeitung unterteilt werden [FMM11]. Je nach Anwendungsfall des Systems können die Ergebnisse aus der Bildverarbeitung anschließend zur aktiven Prozesssteuerung genutzt werden. Dieser Prozessschritt wird im Rahmen der vorliegenden Arbeit jedoch nicht weiter betrachtet. Aus technologischer Sicht umfasst der klassische Bildaufnahmeprozess eine Kamera mit Optik (Objektiv), sowie eine Lichtquelle zur Beleuchtung des zu untersuchenden Objekts. Die durch die Kamera erfassten Daten werden an eine Rechneinheit zur weiteren Bildverarbeitung übertragen (siehe Abb. 2.1). [MPZ03]

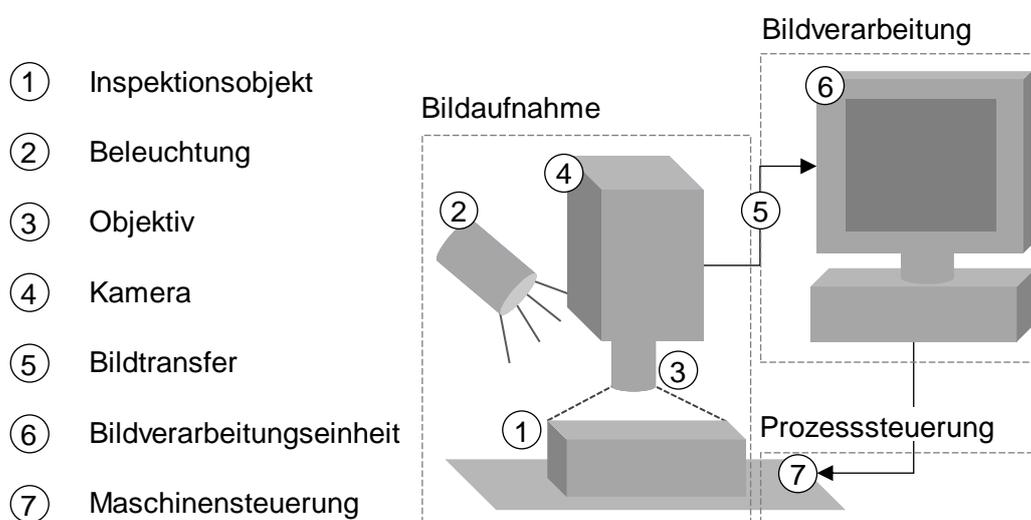


Abb. 2.1: Komponenten eines optischen Inspektionssystem i. A. a. [MPZ03]

Bei einem Inspektionsvorgang mit Hilfe eines Bildverarbeitungssystems wird nicht das Objekt selbst, sondern das visuelle Abbild des Objekts geprüft. Oftmals ermöglicht erst die geschickte Ausnutzung der Eigenschaften von geeigneter Lichtquelle, Prüfobjekt und Kamera die Lösung schwieriger Applikationen [Hen08]. Ziel des Prozessschrittes Bildaufnahme ist es daher, ein visuelles Objektabbild zu erzeugen, das Rückschlüsse auf die tatsächliche Objektbeschaffenheit ermöglicht.

Die in Kapitel 1.2 definierte erste Forschungsfrage zur Auswahl geeigneter Hardware bezieht sich gemäß Abb. 2.1 auf den Prozess der Bildaufnahme. Forschungsfrage 2 hingegen bezieht sich auf den Prozessschritt der Bildverarbeitung.

2.1 Existierende Ansätze zur Technologieauswahl

Zur Einführung in das Thema werden im Rahmen einer Literaturrecherche Quellen untersucht, welche Logiken zur Auswahl geeigneter Technologien für industrielle

Bildverarbeitungsapplikationen thematisieren. Eine Übersicht von Quellen und deren Herangehensweise ist in Tab. 2.1 aufgeführt.

Tab. 2.1: Literatur zu Auswahllogiken von Bildverarbeitungstechnologien

Autor	Titel	Fokus der Arbeit	Eignung
[DSS11]	Handbuch der industriellen Bildverarbeitung	Holistische Übersicht relevanter Bildverarbeitungshardware und deren Interdependenzen	Keine konkrete Auswahllogik
[KST90]	An Expert System for Industrial Machine Vision	Entworfenes Expertensystem unterstützt Auswahl und Design geeigneter Musterklassifizierungsalgorithmen	Sehr hoher Spezialisierungsgrad, veralteter Stand der Technik
[Mat06]	Camera Systems in Machine Vision	Abstrakte Übersicht relevanter Aspekte zur Auswahl eines Bildverarbeitungssystems und technischer Interdependenzen	Keine konkrete Auswahllogik
[MCC+19]	Deep Learning vs. Traditional Computer Vision	Grundlegende Unterscheidung zwischen Deep Learning und konventionellen Methoden zur Bildanalyse	Keine konkrete Auswahllogik
[Nov91]	The Lighting and Optics Expert System for Machine Vision	Entworfenes Expertensystem gibt Anwendungsfall-spezifische Empfehlungen bzgl. Beleuchtungstechnik, Lichtquelle und zu verwendender Linse	Veralteter Stand der Technik
[Per10]	Auswahl der geeigneten Beleuchtung für Anwendungen in der industriellen Bildverarbeitung	Übersicht verschiedener Beleuchtungslösungen und deren Funktionsweisen	Keine konkrete Auswahllogik

Das Prinzip einer Entscheidungslogik, welche dem Anwender konkrete relevante Technologielösungen in Abhängigkeit seines Anwendungsfalls vorschlägt, wird in

dieser Form nur von [Nov91] und [KST90] aufgegriffen. Die dabei entworfene Entscheidungslogik ist thematisch jedoch stark eingegrenzt. In Anbetracht der Veröffentlichungszeitpunkte sind die beiden Publikationen zudem technologisch nicht auf dem aktuellen Stand, da Technologien wie CMOS-Sensoren, LED-Beleuchtungen und Deep Learning noch nicht am Markt etabliert waren [Ste18]. Andere Quellen wie z. B. [DSS11] geben zeitgemäße Überblicke relevanter Technologien und deren Anwendungsbereiche. Sie erläutern dem Leser die technischen Zusammenhänge eines Bildverarbeitungssystems. Es handelt sich dabei jedoch nicht um eine konkrete Entscheidungslogik, mit der dem Nutzer in Abhängigkeit seiner spezifischen Anwendung die Technologieauswahl erleichtert wird.

2.2 Ansatz zur Technologieauswahl im Rahmen dieser Arbeit

Da keine der recherchierten Literaturquellen zum Beantworten der definierten Forschungsfragen ausreicht, wird in dieser Arbeit ein eigenes, regelbasiertes Vorgehen (Entscheidungslogik) zur anwendungsfallspezifischen Auswahl von Hard- und Softwaretechnologien für visuelle Qualitätsinspektionen entwickelt. Das grobe Vorgehen im Rahmen dieser Arbeit ist Abb. 2.2 zu entnehmen.

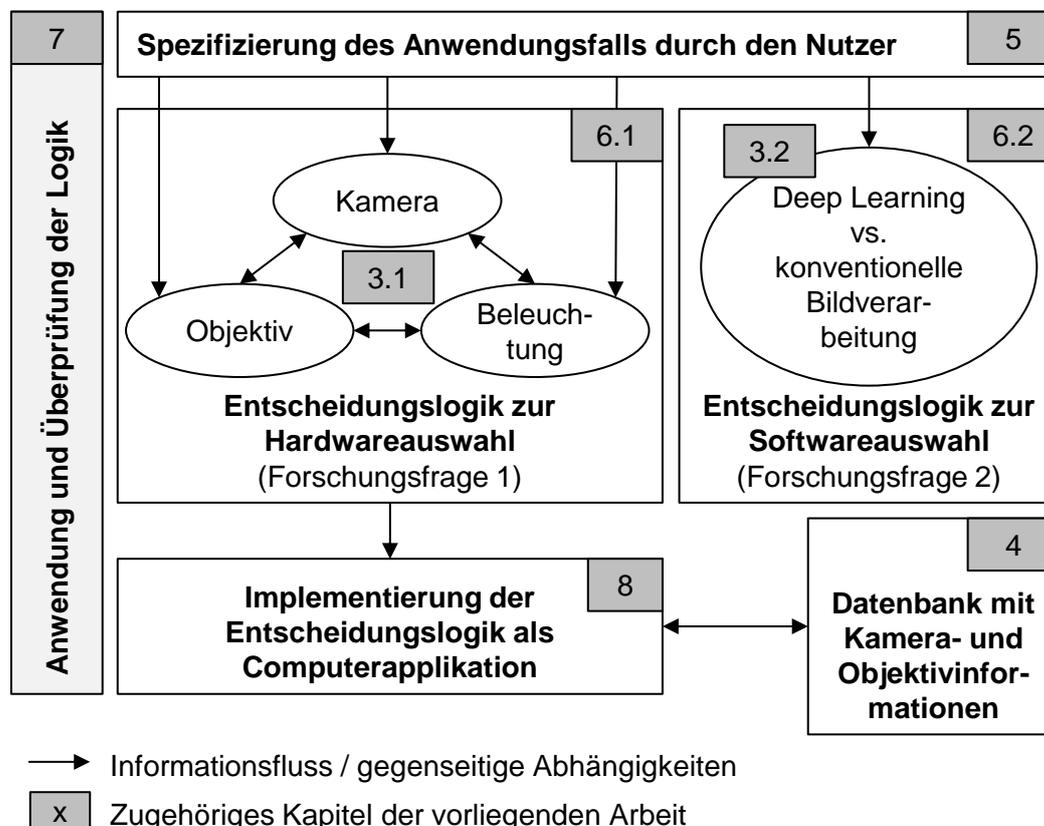


Abb. 2.2: Konzeption der Entscheidungslogik zur Technologieauswahl

Aufbau und Vorgehensweise

Um die Entscheidungslogik anwenden zu können, bedarf es zunächst konkreter Informationen über den gegebenen Anwendungsfall. Dazu werden in Kapitel 5 Beschreibungsgrößen zur formalen Charakterisierung definiert. Anhand dieser Charakteristika werden dem Nutzer die geeignete Hardware (Kamera, Objektiv, Beleuchtung) und Software (Deep Learning vs. konventionelle Bildverarbeitung) empfohlen. Um dies zu ermöglichen, müssen in der Logik Relationen zwischen Anwendungsfallcharakteristika und Technologieeigenschaften sowie technische Abhängigkeiten zwischen den verschiedenen Hardwarekomponenten hinterlegt sein (siehe schwarze Pfeile in Abb. 2.2). Das dazu notwendige technische Grundverständnis wird in Kapitel 3 erarbeitet und bei der Entwicklung der Logik in Kapitel 6 wieder aufgegriffen.

Da für die Entscheidungslogik zur Hardwareauswahl verschiedene Berechnungsschritte und die Einbindung von konkreten Produktinformationen notwendig sind, ist die konzeptionierte Logik in schriftlicher Form nicht anwenderfreundlich. Daher wird die Hardware-bezogene Entscheidungslogik in Kapitel 8 in eine nutzerfreundliche Computerapplikation überführt. Dabei wird eine Datenbank mit Kamera- und Objektivprodukten verwendet, welche im Rahmen einer Marktrecherche in Kapitel 4 aufgebaut wird.

Zielgrößen

In der Literatur zu Projekt- und Produktionsmanagement wird häufig das Spannungsfeld zwischen Kosten, Qualität und Zeit angeführt. Diese drei Zielgrößen sind von einem Entscheider untereinander zu priorisieren, da selten alle drei Zielgrößen gleichzeitig optimiert werden können. [MSL+03] Wird z. B. eine qualitativ hochwertige Lösung mit geringem Zeitaufwand zur Implementierung angestrebt, so führt diese zu entsprechend hohen Kosten.

Im Rahmen der Entscheidungslogik der vorliegenden Arbeit werden Technologien zunächst nach qualitativen Aspekten ausgewählt. Eine Technologie (Hardware und Software) muss alle Mindestanforderungen, die bei der Spezifizierung des Anwendungsfalls definiert werden, vollständig erfüllen.

Werden bei der Hardwareauswahl die Mindestanforderungen von mehreren Produkten erfüllt, so sollen diese nach Listenpreisen priorisiert werden (siehe Kapitel 8.1). Unter der Grundvoraussetzung der qualitativen Eignung werden somit die Anschaffungskosten optimiert werden.

Bei der Auswahl zwischen Deep Learning und konventionellen Algorithmen zur Bildverarbeitung sind die monetären Kosten hingegen schwer zu bewerten. Sind für eine Anwendung sowohl konventionelle Algorithmen als auch Deep Learning grundsätzlich geeignet, so werden die Lösungen anhand ihres zu erwartenden Entwicklungsaufwandes (Optimierung der Zielgröße Zeit) verglichen.

3 Vorstellung und Parametrisierung von Bildverarbeitungstechnologien

In diesem Kapitel werden die wesentlichen technischen Grundlagen der industriellen BV vorgestellt. Im Bereich der Bilderfassung werden relevante Hardwaretechnologien und ihre charakteristischen Parameter vorgestellt. Im Bereich der Bildanalyse werden verschiedene algorithmische Methoden zur Auswertung der Bilddaten aufgearbeitet und nach Anwendungsbereich kategorisiert. Hierbei wird insbesondere auf die Unterschiede zwischen traditionellen BV-Algorithmen und Deep Learning im Kontext von Bilderkennung eingegangen.

3.1 Hardware

Das Kapitel Hardware beschreibt die physikalischen Grundlagen sowie die in der Fachliteratur geläufigen Beschreibungsparameter für Kamerasensoren, Objektive und Beleuchtungssysteme. Weitere Hardwarekomponenten in diesem Kontext sind eine Rechneinheit zur Bildverarbeitung, Frame Grabber zur Datenzwischenspeicherung und -übertragung, Kabel und Stecker sowie Montageelemente. Gemäß Themeneingrenzung werden diese jedoch hier nicht weiter betrachtet.

3.1.1 Kamerasensor

In diesem Abschnitt wird zunächst die grundlegende Funktionsweise von Kamerasensoren erläutert. Anschließend werden Parameter zur technischen Beschreibung vorgestellt. Der Fokus liegt dabei auf Matrixbildsensoren zur Aufnahme von monochromatischen Bildern. Für eine detailliertere Behandlung der physikalischen Grundlagen sei auf die Literaturquellen [Jäh12; EMVA 1288] verwiesen.

3.1.1.1 Grundlegende Funktionsweise von Bildsensoren

Ein Bildsensor besteht aus mehreren lichtempfindlichen Elementen (Pixel), die linienförmig (Zeilensensor) oder flächig (Matrixsensor) angeordnet sind. Der Bildsensor konvertiert ein Bild, d.h. eine räumlich kontinuierlich variierende Bestrahlung, die durch ein optisches System auf den Sensor fällt, in einen digitalen Datenstrom [Jäh12]. Diese Konvertierung lässt sich in fünf Prozessschritte gemäß Abb. 3.1 untergliedern, die im Folgenden näher erläutert werden:

Prozessschritte der digitalen Bildaufnahme

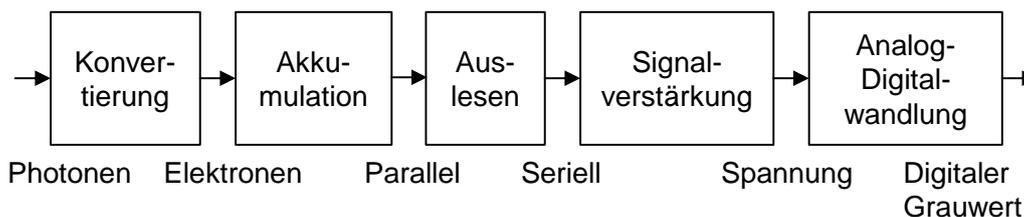


Abb. 3.1: Bildaufnahmeprozess i. A. a. [Jäh12]

Konvertierung: Bei der Konvertierung wird Strahlung mittels des photoelektrischen Effekts in elektrische Ladungen (Elektronen) umgewandelt. Elektromagnetische Strahlung ist in Form von Photonen quantisiert. Mit der Lichtgeschwindigkeit c und dem planckschen Wirkungsquantum h ergibt sich für ein Photon mit der Frequenz f bzw. der Wellenlänge λ folgende Energie:

$$E_{\text{Photon}} = h \times f = \frac{h \times c}{\lambda} \quad (3.1)$$

Elektromagnetische Strahlung im Wellenlängenbereich von 380 bis 780 nm ist für das menschliche Auge sichtbar und wird im allgemeinen Sprachgebrauch als Licht bezeichnet. Trifft ein Photon auf einen Halbleiter, so kann dieser das Photon absorbieren. Dies geschieht jedoch nur, wenn das Photon ausreichend Energie besitzt, um ein Elektron des Halbleiters auf ein höheres Energieniveau zu bringen. Die notwendige Energie zur Elektronenanregung bezeichnet man als Bandlückenenergie. Sie ist abhängig von der Zusammensetzung des Halbleiters. Für Silizium liegt die Bandlückenenergie bei etwa 1,1 eV, woraus eine maximale Wellenlänge von etwa 1100 nm zur Elektronenanregung folgt. Photonen höherer Wellenlänge bzw. geringerer Energie können von Silizium nicht aufgenommen werden, das Material wirkt transparent. Reicht die Energie des Photons hingegen aus, wird das Elektron in einen höheren Energiezustand im sogenannten Leitungsband versetzt. Dort ist es frei beweglich. [Jäh12; DSS11; Jer09]

Akkumulation: Gemäß dem photoelektrischen Effekt werden auf der Sensorfläche eines jeden Pixels (Fotodiode) über eine bestimmte Zeit (Belichtungs- oder auch Integrationszeit) Photonen in frei bewegliche Elektronen konvertiert. Die freigesetzten Elektronen werden über die Belichtungszeit mittels einer angelegten Spannung in einem sogenannten Potenzialtopf gesammelt und gespeichert. Die maximale Anzahl an Elektronen, die im Potenzialtopf gespeichert werden kann, bezeichnet man als Full-Well-Kapazität. Digitalkameras mit Halbleitersensoren besitzen keinen mechanischen Belichtungsmechanismus, sodass ein elektrischer Mechanismus zum Starten und Beenden der Belichtungszeit verwendet werden muss. Dies geschieht über das Löschen / Auslesen der Ladungen in den Fotodioden aller Pixel. Werden die elektrischen Ladungen (zum Auslesen) aus der

Fotodiode weitergeleitet, beendet dies den vorherigen Belichtungszeitraum, gleichzeitig beginnt eine neue Belichtungszeit. [Mat06]

Auslesen: Beim Auslesen der Ladungen aus der Fotodiode eines jeden Pixels unterscheidet man bei Halbleiterbildsensoren zwischen zwei wesentlichen Funktionsweisen, Complimentary Metal-Oxide-Semiconductor (CMOS) und Charge-Coupled Device (CCD) Sensoren. Bei CCD-Sensoren liegen dazu zwischen jeder Spalte der matrixförmig angeordneten Pixel sogenannte vertikale Transportregister, siehe Abb. 3.2. Die Ladungen eines jeden Pixels werden am Ende der Belichtungszeit gleichzeitig in das Transportregister verschoben. Das Transportregister ist vor zusätzlichem Lichteinfall weitgehend geschützt, sodass keine weiteren Ladungsträger erzeugt werden. Nach dieser zeitlich parallelen Verschiebung der Ladungsträger in die vertikalen Transportregister werden sie anschließend seriell (zeitlich nacheinander) weiterverarbeitet. Die Ladungen eines Pixels werden dabei Zeile für Zeile zu einem horizontalen Transportregister verschoben. Am Ausgang des horizontalen Transportregisters werden über einen Ausgangsverstärker die Ladungen eines jeden Pixels in einen proportionalen Spannungswert umgewandelt. Da die in den Transportregistern gespeicherten Ladungen seriell ausgelesen werden müssen, schränkt dies die Auslesegeschwindigkeit und somit auch die maximale Bildrate ein (ein Belichtungsvorgang kann erst gestoppt werden, wenn der Auslesevorgang der vorherigen Belichtung abgeschlossen wurde). Durch die Verwendung mehrerer Ausgangsverstärker können mehrere Pixel parallel ausgelesen und die Geschwindigkeit somit erhöht werden. Dies bezeichnet man als Multi-Tap-Sensoren. [Mat06; DSS11; Jäh12]

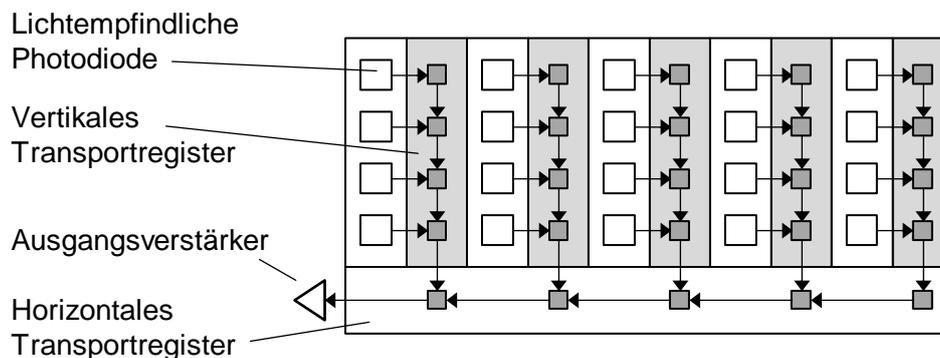


Abb. 3.2: Aufbau CCD-Sensor i. A. a. [Mat06]

CCD-Sensoren haben einen oder wenige Ladungsverstärker zur Umwandlung der Ladungen in Spannungswerte. Im Gegensatz dazu hat bei CMOS-Sensoren jedes Pixel einen eigenen Ladungsverstärker. Helligkeitsinformationen liegen somit bereits als Spannung an jedem Pixel vor, die akkumulierten Ladungen müssen also nicht mehr über den Bildsensor zum Ausgangsverstärker transportiert werden. Stattdessen können die Spannungswerte des Pixels über eine Transistorschaltung (siehe Abb. 3.3) in beliebiger Reihenfolge oder in Teilbereichen ausgelesen

werden. Die einzelnen Ladungsverstärker an den Pixeln unterscheiden sich untereinander leicht hinsichtlich Verstärkungsfaktor und Offset, sodass aus gleicher Ladungszahl unterschiedliche Spannungen resultieren. [Mat06; Jäh12; HKT07] Dies macht sich in einem ortsfesten Bildrauschen (Fixed Pattern Noise) bemerkbar. CMOS-Sensoren sind in der Fertigung günstiger, schneller im Auslesen und geringer im Energieverbrauch als CCD-Sensoren. Aufgrund des geringeren Bildrauschens wurden CCD-Sensoren traditionell bevorzugt für Industrieanwendungen eingesetzt (63% Marktanteil in 2008) [CPS10]. Mit zunehmender Verbesserung der Bildqualität ist inzwischen jedoch die Mehrheit aktueller Industriekameras mit CMOS-Sensoren ausgestattet [Ste18]. So wird z. B. durch die Methode des Correlated Double Sampling (CDS) der pixelspezifische Offset herausgerechnet, indem vom Spannungswert der aktuellen Helligkeit der Spannungswert bei nicht-Belichtung subtrahiert wird.

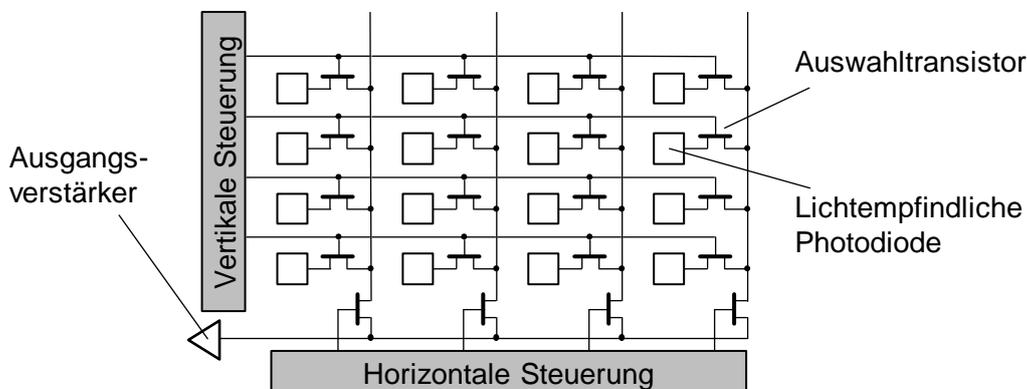


Abb. 3.3: Aufbau CMOS-Sensor i. A. a. [DSS11]

Signalverstärkung und Analog-Digitalwandlung: Nach dem Auslesen liegen die Helligkeitswerte der einzelnen Pixel als Spannungen vor. Lässt der Anwendungsfall der Bildaufnahmen keine Optimierung der Beleuchtung und Erhöhung der Belichtungszeit zu, können die gemessenen Spannungswerte künstlich verstärkt werden. So erhält man höhere Helligkeitswerte, jedoch wird neben dem Bildsignal auch das Rauschen (siehe Abschnitt Rauschen) verstärkt [DSS11]. Bei der abschließenden Analog-Digitalwandlung werden die kontinuierlichen Spannungswerte in diskrete Helligkeitswerte gewandelt, man spricht hierbei auch von einer Quantisierung mittels nicht-linearer Stufenfunktion (siehe Abb. 3.4) [Jäh12]. Die Bilddaten werden standardmäßig in 256 Graustufen quantisiert, dies entspricht einer 8 Bit Architektur. Der geringste Grauwert von 0 (schwarz) entspricht dabei der Grundspannung, die ohne Lichteinfall herrscht. Der Maximalwert 255 ergibt sich für die maximal speicherbare Anzahl an Ladungsträgern in einem Pixel. [Jäh12; Ste18]

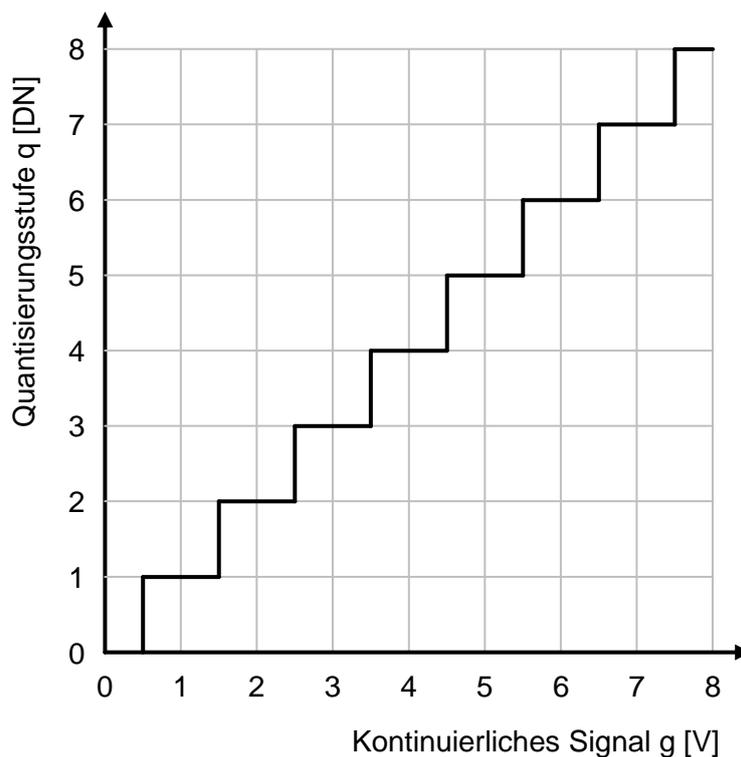


Abb. 3.4: Quantisierung eines kontinuierlichen Spannungssignals i. A. a.
[Jäh12]

Bei den hier vorgestellten Prozessschritten wird nicht zwischen den unterschiedlichen Wellenlängen des einfallenden Lichtes unterschieden, es entsteht somit ein monochromatisches Bild. Zur Aufnahme von Farbbildern müssen Helligkeitswerte jedoch nach Farbbereich getrennt erfasst werden. Im Gegensatz zum Konsumgütermarkt werden Farbkameras in der Produktion allerdings nur bei Bedarf für konkrete Farbinformationen verwendet. Es sei an dieser Stelle daher eine kurze Übersicht geeigneter Technologien gegeben:

- **3-Chip-Farbkameras** verfügen über drei verschiedene Chips. Mittels eines Farbteilers wird das einfallende Licht in seine rote, grüne und blaue Komponente aufgeteilt und auf den jeweiligen Chip geleitet. Somit stehen für jeden Pixel Informationen zu allen drei Farbkomponenten zu Verfügung. Solche Kameras liefern hochwertige Farbbilder, sie sind jedoch auch aufwendig und teuer. [DSS11]
- **1-Chip-Farbkameras mit Farbfiltermaske** verfügen nur über einen Bildsensor. Die vor dem Sensor angebrachte Filtermaske sorgt dafür, dass ein Pixel immer nur einen Farbkanal erfassen kann. Jedes Pixel empfängt somit entweder rotes, grünes oder blaues Licht. Um für jedes Pixel alle drei Farbinformationen zu erhalten, werden diese über Helligkeitswerte der Nachbarpixel mit einer anderen spektralen Empfindlichkeit interpoliert. 1-Chip-Farbkameras

sind kostengünstig und stellen die am häufigsten verwendete Farbbildtechnologie dar. [Mat06]

- **Dreilagfarbsensoren** haben drei übereinander angeordnete Pixellagen. Jede Pixellage absorbiert nur einen bestimmten Wellenlängenbereich, andersfarbiges Licht wandert durch den Halbleiter auf die nächste Pixelebene. Es handelt sich hierbei jedoch lediglich um eine Nischentechnologie. [DSS11]

3.1.1.2 Technologieparametrisierung von Bildsensoren

Die folgenden technischen Parameter zur Charakterisierung und Leistungsbeschreibung von Bildsensoren werden in diesem Abschnitt vorgestellt:

- Auflösung
- Pixelgröße und Füllfaktor
- Optisches Format
- Bittiefe
- Belichtungszeit
- Bildrate
- Quantenausbeute und spektrale Empfindlichkeit
- Dunkelstrom
- Signal / Rausch-Verhältnis
- Sättigungskapazität
- Empfindlichkeitsschwelle
- Dynamikumfang
- Linearität
- DSNU und PRNU
- Übertragungsschnittstelle
- Schutzgrad

Eine Auswahl dieser Beschreibungsparameter wird in Kapitel 6.1 für die Entscheidungslogik zur Technologieauswahl verwendet.

Auflösung: Für Matrixsensoren wird die Pixelanzahl entweder in horizontaler und vertikaler Richtung sowie als Gesamtanzahl in Megapixeln angegeben. Das typische Seitenverhältnis beträgt 4:3.

Die benötigte Sensorauflösung eines Kamerasystems ergibt sich aus Größe des kleinsten Objektdetails, der Größe des Sichtfeldes (FOV) sowie der Anzahl an Pixeln, die das kleinste Objektdetail in der Bildebene repräsentiert. In Anlehnung an das Abtasttheorem von Nyquist und Shannon muss ein Merkmal durch mindestens 2 Pixel repräsentiert werden können, woraus die Multiplikation mit dem Faktor 2 in Gleichung 3.2 resultiert. Diese Berechnung muss für die Horizontale und Vertikale durchgeführt werden. [Ste18; Jäh12; SHB99]

$$\text{Sensorauflösung} = \frac{\text{FOV} \times 2}{\text{minimale Objektgröße}} \quad (3.2)$$

Pixelgröße und Füllfaktor: Je größer ein Pixel ist, desto größer ist auch dessen Lichtempfindlichkeit (siehe Quanteneffizienz) und die maximale Kapazität zur Aufnahme elektrischer Ladungen (siehe Sättigungskapazität). Dadurch wird der Dynamikumfang eines Bildes erhöht, jedoch sind solche Sensoren auch höher im Preis. [Hol10] Der Füllfaktor gibt den prozentualen Anteil der lichtempfindlichen Fläche eines Pixels an. Bei Sensoren mit komplexen Schaltstrukturen zum Auslesen bzw. Umwandeln der Ladungsträger kann der Füllfaktor auf etwa 30% herabsinken. Durch die Verwendung von Mikrolinsen vor den einzelnen Pixeln kann der Füllfaktor erhöht werden. [Ste18]

Optisches Format: Das optische Format bezieht sich auf die Diagonallänge eines Matrixbildsensors. Dieses Maß ist relevant, um ein zu der Sensorgröße passendes Objektiv bestimmen zu können. Üblicherweise wird das optische Format in Zoll angegeben. Diese Angabe entspricht jedoch nicht der tatsächlichen Zolllänge der Bilddiagonale, sondern hat historische Gründe. [Ste18] Eine Übersicht an üblichen Sensorgrößen kann Abb. 3.5 entnommen werden.

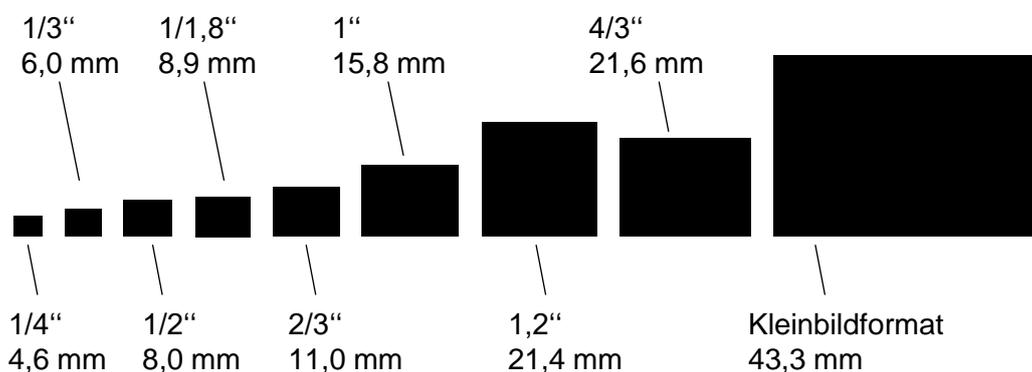


Abb. 3.5: Beispielformate und Bilddiagonalen von Bildsensoren i. A. a. [Ste18]

Bittiefe: Die Bittiefe gibt an, mit wie vielen Bits die Helligkeitswerte der Pixel quantisiert werden. Sie beträgt standardmäßig 8 Bit, was 256 Graustufen entspricht. Das Auflösungsvermögen des menschlichen Auges nimmt Unterschiede dunkler Bereiche deutlich feiner wahr als helle Bereiche. Für Spezialanwendungen mit sehr feinen Unterscheidungen im Dunkelbereich werden Quantisierungen mit mehr als 8 Bit benötigt. Die meisten A / D-Wandler von Sensoren können mit bis zu 14 Bit konvertieren. [Mat06; Jäh12; Ste18]

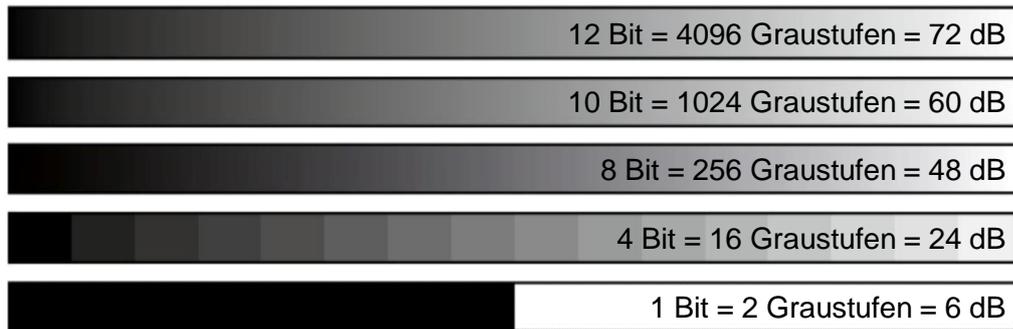


Abb. 3.6: Beispiele unterschiedlicher Graustufen i. A. a. [Ste18]

Belichtungszeit: Die Belichtung des Bildsensors bzw. die Anzahl absorbiertes Photonen ist vereinfacht ausgedrückt das Produkt aus Zeit und Strahlungsintensität [Jäh12]. Die Belichtungszeit ist somit neben der Strahlungsintensität der wesentliche Faktor zur Steuerung der Helligkeitswerte. Sie ist im eigentlichen Sinne kein Parameter eines Bildsensors, da sie flexibel einstellbar ist. Jedoch ist sie für das Verständnis weiterer Bildparameter unabdingbar und sei daher hier erwähnt. Eine Verdopplung der Belichtungszeit gleicht einer Halbierung der Strahlungsintensität bei konstanten Helligkeitswerten. Je nach Anwendungsfall ist die maximale Belichtungszeit jedoch eingeschränkt. So muss etwa bei bewegten Teilen zur Minimierung der Bewegungsunschärfe eine geringe Belichtungsdauer gewählt werden. [Jah06]

Bildrate: Die Bildrate gibt an, wie viele Bilder eine Kamera pro Sekunde aufnehmen kann. Aus Anwendersicht ergibt sich die benötigte Bildrate aus der Taktung der zu untersuchenden Objekte. Durchlaufen etwa 10 Teile pro Sekunde die Inspektionsstelle, so wird eine Bildrate von 10 Bildern pro Sekunde benötigt. [Ste18] Aus technischer Sicht existieren verschiedene Faktoren, welche die maximal mögliche Bildrate limitieren:

- **Bildaufnahmemodus:** Traditionell erfolgen Belichtung und Auslesevorgang nacheinander, die Bildrate ist dann durch den Kehrwert der Summe der beiden Vorgangszeiten begrenzt. CCD-Sensoren können auch im sogenannten Overlapped-Modus betrieben werden (siehe Abb. 3.7). Hierbei kann die Belichtung eines Bildes parallel zum Auslesen der vorherigen Bildwerte erfolgen. Die Bildrate ist dann theoretisch durch den Kehrwert des Maximums von Belichtungszeit und Auslesezeit limitiert. [DSS11]
- **Schnittstellentechnologie:** Unterschiedliche Schnittstellen und die damit verbundenen Technologien ermöglichen verschiedene Übertragungsgeschwindigkeiten für Bilddaten. Das Datenvolumen eines Bildes ergibt sich aus der ausgelesenen Pixelanzahl und der Bittiefe. [AEJ18]
- **Sensortechnologie:** CCD-Sensoren bedürfen generell einer höheren Auslesezeit, da hier die elektrischen Ladungen Pixel für Pixel über den Chip zum

Ladungsverstärker transportiert werden müssen. Bei CMOS-Sensoren werden die Ladungen bereits im Pixel in Spannungen konvertiert. [Mat06]

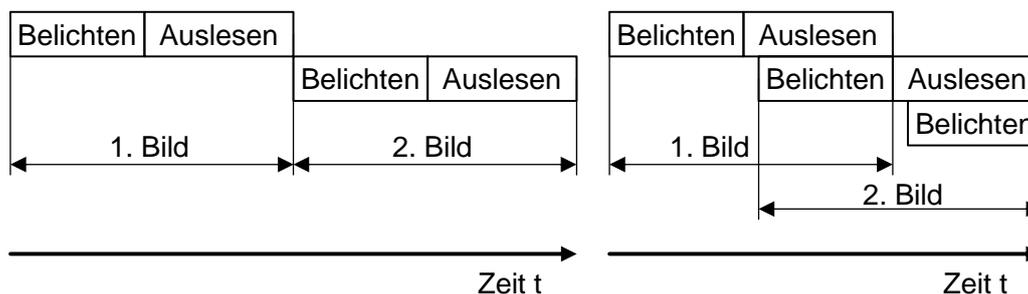


Abb. 3.7: Konventioneller Aufnahmemodus (links) und Overlapped Aufnahme-Modus (rechts) i. A. a. [DSS11]

Quantenausbeute und spektrale Empfindlichkeit: Die experimentell bestimmbare Quantenausbeute beschreibt die Anzahl kumulierter Elektronen N_e im Verhältnis zur Photonenzahl N_p , die auf die Pixeloberfläche aufgetroffen ist. Man bezeichnet dieses Verhältnis auch als Quanteneffizienz. Die Quanteneffizienz variiert mit der Photonenwellenlänge λ , sodass man auch von spektraler Empfindlichkeit spricht.

$$H(\lambda) = \frac{N_e}{N_p} \quad (3.3)$$

Die Quantenausbeute bezieht sich auf die komplette Pixeloberfläche, nicht nur auf den lichtempfindlichen Teil. Für niederenergetische Strahlung kann sie maximal 1 sein, da hier ein absorbiertes Photon höchstens in ein Elektron umgewandelt wird. In der Praxis ist sie aufgrund verschiedener Effekte (z. B. Füllfaktor, Reflektion von Photonen an Sensoroberfläche) kleiner als 1. [Jäh12; DSS11; EMVA 1288]

Dunkelstrom: Freie Ladungsträger werden in einem Sensorpixel nicht nur durch Photonen, sondern auch durch thermische Effekte erzeugt. Ohne eine Signalkorrektur (Offset) g_0 würden sich an einem Pixel folglich auch Helligkeitswerte ohne jeglichen Photoneneinfall messen lassen. Diesen auf unerwünschten thermischen Effekten basierenden Strom bezeichnet man als Dunkelstrom. Er unterscheidet sich von Pixel zu Pixel und variiert mit der Umgebungstemperatur und der Belichtungszeit. In der Praxis ist der Dunkelstrom meistens erst bei Belichtungszeiten von über einer halben Sekunde relevant. [EMVA 1288; Ste18]

Signal / Rausch-Verhältnis (SNR): Der Bildaufnahmeprozess unterliegt verschiedenen Rauscheffekten, die keinen Bezug zum tatsächlichen Bildinhalt, dem Bildsignal g haben. Rauschen führt somit zu einer Verfälschung des aufgenommenen Signals g_m . Mit dem Mittelwert des Rauschens g_0 ergibt sich das Bildsignal g zu:

$$g = g_m - g_0 \quad (3.4)$$

Mittels statistischer Verfahren lassen sich gemessene Helligkeitswerte in ihren Rausch- und Signalanteil aufgliedern und zueinander in Verhältnis setzen. Dies bezeichnet man als SNR. Das SNR entspricht dem Quotienten aus dem rauschbereinigten Bildsignal g (siehe Gleichung 3.4) und der Standardabweichung σ_g des Rauschens.

$$\text{SNR} = \frac{g}{\sigma_g} \quad (3.5)$$

Folgende Arten des Rauschens finden in der Messung des SNR Beachtung:

- **Photonenrauschen:** Das Photonenrauschen resultiert aus der natürlichen Schwankung der Lichtintensität. Es ist somit unabhängig vom Sensor bedingt.
- **Dunkelrauschen:** Dieses entspricht dem thermischen Rauschen, das auch mittels Dunkelstrom beschrieben wird.
- **Quantisierungsrauschen:** Dieses Rauschen resultiert aus der Umwandlung von kontinuierlichen Spannungswerten in diskrete Stufen im Rahmen der Quantisierung. Es lässt sich durch eine höhere Bittiefe reduzieren.

Das SNR wird in Abhängigkeit von der Belichtung (Photonenanzahl) gemessen. Mit zunehmender Belichtung wächst der Signalanteil überproportional zum Rauschanteil, sodass sich das SNR verbessert. Der maximale SNR-Wert ergibt sich für die Sättigungskapazität eines Pixels. [YG99; EMVA 1288; Mat06] In Datenblättern werden die SNR-Werte oft in einem doppeltlogarithmischen Diagramm über die Belichtung aufgetragen (siehe Anhang 11.1).

Sättigungskapazität: Die maximale Sättigungskapazität entspricht der maximalen Anzahl an absorbierbaren Elektronen pro Flächeneinheit (Full-Well-Kapazität) abzüglich der Anzahl an rauschbedingten Ladungen. Der Bildsensor nimmt also noch mehr Ladungsträger als die Sättigungskapazität auf. Jedoch sind diese zusätzlichen Ladungen nicht auf das eigentliche Signal, sondern auf dessen Rauscheffekte zurückzuführen. Eine hohe Sättigungskapazität ermöglicht einen höheren Dynamikumfang und ein höheres SNR. [EMVA 1288; Jäh12]

Empfindlichkeitsschwelle: Die Empfindlichkeitsschwelle beschreibt die minimale Anzahl an Photonen, die benötigt wird, um ein Signal vom Rauschen zu unterscheiden. Sie entspricht einem quantisierten Helligkeitswert sowie einem SNR von 1. Der Empfindlichkeitsschwellwert beinhaltet die Information der Quanteneffizienz sowie des Dunkelrauschens. Er ist daher von höherer Aussagekraft als einer dieser Werte für sich isoliert betrachtet. [EMVA 1288]

Dynamikumfang: Der Dynamikumfang ist der Quotient aus Sättigungskapazität und Empfindlichkeitsschwelle. Ein hoher Dynamikumfang bedeutet anschaulich vereinfacht, dass zwischen einer hohen Anzahl an Helligkeitswerten differenziert werden kann (siehe Abb. 3.8). Allerdings ist der genaue Zahlenwert des Dynamikumfangs nicht aussagekräftig, da das absolute Rauschen des

Empfindlichkeitsschwellwertes deutlich geringer als das absolute Rauschen der Sättigungskapazität ist. Das Signal / Rausch-Verhältnis (SNR) bei Sättigungskapazität liefert hingegen die tatsächliche maximale Anzahl an Quantisierungsstufen. [Ste18; EMVA 1228]



Abb. 3.8: Aufnahme Kerzenlicht bei geringem Dynamikumfang und unterschiedlichen Belichtungszeiten (links) sowie hohem Dynamikumfang (rechts) [Ste18]

Linearität: Die Menge an Elektronen in einem Pixel steigt linear mit der Anzahl an einfallenden Photonen bis zum Erreichen der Full-Well-Kapazität. Im Rahmen der Spannungsumwandlung innerhalb der Ladungsverstärker kann es jedoch zu geringen Nichtlinearitäten kommen. Die Linearitätsabweichung wird über verschiedene Belichtungen im Intervall von 5% bis 95% Sättigung in Prozent (positive und negative Abweichung) gemessen. Als Kennwerte werden die minimale und maximale Abweichung innerhalb des Intervalls angegeben. In der Praxis ist die Linearitätsabweichung hauptsächlich bei sehr präzisen Farbmessungen von Relevanz. [EMVA 1288; DSS11]

DSNU und PRNU: Die bisherigen Parameter zur Sensorempfindlichkeit beziehen sich auf die Werte einzelner Pixel bzw. werden diese Größen über die Werte aller vermessenen Pixel hinweg gemittelt. Um räumlich bedingte Unterschiede zwischen den einzelnen Pixeln beschreiben zu können, hat die EMVA die beiden folgenden Größen definiert:

- **Dark signal nonuniformity (DSNU):** Die DSNU ist die ortsbedingte Standardabweichung des Dunkelstroms über alle Pixel.
- **Photo response nonuniformity (PRNU):** Die PRNU ist die ortsbedingte Standardabweichung der Helligkeitswerte aller Pixel geteilt durch die mittlere Pixelhelligkeit.

Da diese örtlichen Abweichungen per Definition als zeitlich konstant gelten, lassen sie sich in der Praxis durch eine Offsetkorrektur einfach beheben. [EMVA 1228]

Übertragungsschnittstelle: In der BV-Industrie hat sich eine Reihe von Übertragungsschnittstellen etabliert, die den Austausch von Informationen zwischen lokalem Speicher der Kamera und externen Computern ermöglichen. Die Übertragungsschnittstellen unterscheiden sich z. B. bezüglich Übertragungstechnik, maximal möglicher Übertragungsgeschwindigkeit oder maximal möglicher Kabellänge. Eine Übersicht relevanter Übertragungsstandards und deren Eigenschaften ist Tab. 3.1 zu entnehmen. [AEJ18]

Tab. 3.1: Übersicht gebräuchlicher Kameraschnittstellen i. A. a. [AEJ18]

	Maximale Übertragungsrate [Mbytes / s]	Maximale Kabellänge [m]	Echtzeitfähigkeit	Schnittstelle am PC
CoaXPress	> 5000	< 50	Gut	Frame Grabber
GigE Vision	< 1000	< 120	Gut	Standard
USB 3 Vision	< 500	10 -120	Mäßig	Standard
IEEE1394 + IIDC	< 200	< 10	Mäßig	Teilweise
Camera Link	< 1000	< 10	Sehr gut	Frame Grabber
Camera Link HS	< 5000	10 - 120	Gut	Frame Grabber

Schutzgrad: In industriellen Anwendungen sind Kameras häufig ungünstigen Umweltbedingungen ausgesetzt. Gelangen Staub oder Feuchtigkeit in die Komponenten des Kamerasystems, kann sich dies negativ auf die Bildqualität und Langlebigkeit der Kamera auswirken. Zum Schutz werden Kamera und Objektiv von speziellen Gehäusen umschlossen, die nach IP-Standard geschützt sind. Die zweistellige Zahlencodierung lässt sich dabei gemäß Tab. 3.2 übersetzen [DSS11; DIN 60529].

Tab. 3.2: IP-Schutzarten [DIN 60529]

Ziffer	Schutz vor Fremdkörper [1. Ziffer]	Schutz vor Wasser [2. Ziffer]
1	Fremdkörper ≥ 50 mm	Tropfwasser
2	Fremdkörper $\geq 12,5$ mm	Fallendes Tropfwasser
3	Fremdkörper $\geq 2,5$ mm	Fallendes Sprühwasser
4	Fremdkörper $\geq 1,0$ mm	Allseitiges Spritzwasser
5	Staub in schädigender Menge	Allseitiges Strahlwasser
6	staubdicht	Starkes Strahlwasser unter erhöhtem Druck
7	-	Zeitweiliges Untertauchen
8	-	Dauerndes Untertauchen
9	-	Hochdruck- / Dampfstrahlreiniger

3.1.2 Objektiv

Damit ein scharfes und kontrastreiches Bild auf dem Bildsensor entstehen kann, bedarf es einer Optik in Form eines Objektivs. Im Folgenden werden mit dem Abbildungsgesetz zunächst die physikalischen Grundlagen von Optiken erklärt, anschließend werden technische Parameter zur Charakterisierung von Objektiven erläutert.

3.1.2.1 Abbildungsgesetz: Grundlegende Funktionsweise von Optiken

Zur Erklärung des Abbildungsgesetzes wird an dieser Stelle das Modell der dünnen Linse verwendet. Bei einer dünnen Linse spielt die Linsendicke keine Rolle für die Abbildung, es handelt sich vielmehr um einen Idealfall. Die Beschreibungsparameter von realen Objektiven werden mittels komplizierterer Abbildungsmodelle bestimmt. Jedoch eignet sich das Modell der dünnen Linse gut, um ein grundlegendes Verständnis von Objektivkennwerten zu vermitteln. [DSS11]

Die Konstruktion des Bildes einer dünnen Linse folgt den folgenden drei Gesetzmäßigkeiten [ST02]:

- 1) Ein parallel zur optischen Achse einfallender Strahl verläuft hinter der Linse durch den Brennpunkt F.
- 2) Ein Strahl, der durch den vor der Linse liegenden Brennpunkt F' einfällt, verläuft hinter der Linse parallel zur optischen Achse.

- 3) Ein Strahl, der durch den Mittelpunkt der Linse verläuft, verändert seine Richtung nicht.

Aus diesen Regeln folgt, dass die Strahlen, die von einem Gegenstandspunkt A mit Abstand g (Gegenstandsweite) zur Mittelebene der Linse ausgehen, sich hinter der Linse wieder in einem Punkt A' treffen (siehe Abb. 3.9). In der Entfernung b (Bildweite) hinter der Linsenmittelebene wird der ursprüngliche Punkt wieder als Punkt, und somit in maximaler Schärfe abgebildet. Den Abstand zwischen Linsenmittelebene und Brennpunkt F / F' bezeichnet man als Brennweite f . Die Brennweite f ist eine charakteristische Größe der Linse und kann gemäß der Linsenmachergleichung aus den Krümmungsradien ermittelt werden. [DSS11; ST02; Hec05]

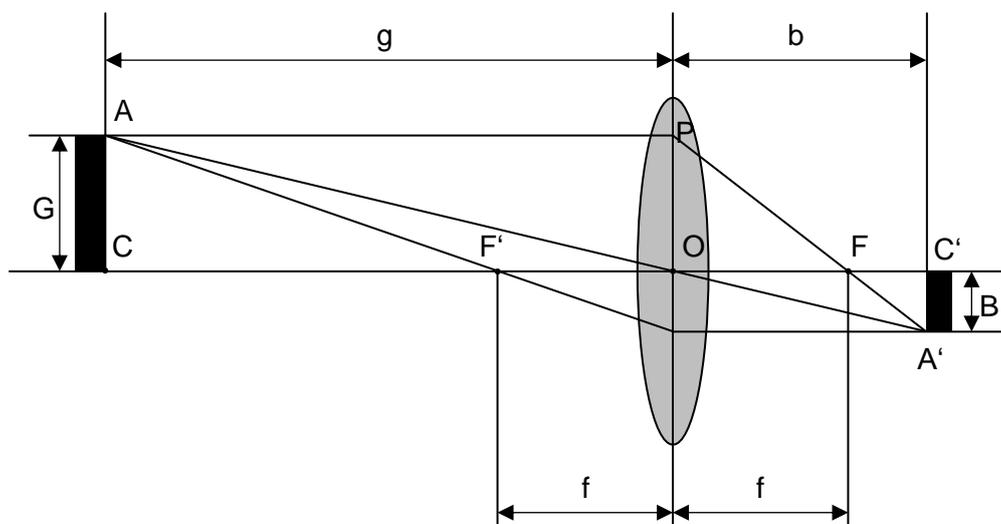


Abb. 3.9: Strahlengang einer dünnen Linse i. A. a. [ST02]

Über die Ähnlichkeiten der Dreiecke ACO und $A'C'O$ sowie von Dreieck POF zu $A'C'F$ lässt sich die Abbildungsgleichung für dünne Linsen herleiten:

$$\frac{1}{f} = \frac{1}{g} + \frac{1}{b} \leftrightarrow b = \frac{g \times f}{g - f} \quad (3.6)$$

Ein interessanter Grenzfall ergibt sich für Abbildungen von Objekten aus sehr großer Entfernung. Hier strebt der Kehrwert der Gegenstandsweite gegen null, sodass die Bildweite b der Brennweite f entspricht.

Da das Verhältnis von Bildgröße B zu Gegenstandsgröße G dem Verhältnis von Bildweite b zu Gegenstandsweite g gleich ist, ergibt sich beim Einsetzen dieser Relation in Gleichung 3.6 folgende Beziehung zwischen Bildgröße B und Gegenstandsgröße G in Abhängigkeit von f und g :

$$B = \frac{f}{g - f} \times G \quad (3.7)$$

Mittels Gleichung (3.7) lassen sich zwei Phänomene von Linsenabbildungen erklären. Einerseits führt eine Erhöhung der Brennweite f zu einer Vergrößerung der Bildgröße B . Dies ist der Grund für die Länge von Teleobjektiven. Andererseits

wird das Bild größer, wenn man den Abstand zum Objekt verringert (Gegenstandsweite g). Wird g verringert, so muss man nach Gleichung 3.6 die Bildweite b erhöhen, um weiterhin eine scharfe Abbildung zu erhalten. In der Praxis lässt sich eine Erhöhung von b durch das Herausdrehen des Objektivtubusses oder durch das Verwenden von Zwischenringen zwischen Kamera und Objektiv bewirken.

3.1.2.2 Beschreibungsgrößen für Objektive

Die folgenden Größen werden in der Fachliteratur zur quantitativen Beschreibung von Objektiven aufgeführt:

- Brennweite
- Vergrößerung
- Blendenzahl
- Schärfentiefe
- Minimale Objektdistanz
- Auflösungsvermögen
- Homogenität der Ausleuchtung
- Verzeichnung
- Objektivgröße / Lichtkreisdurchmesser
- Kameraanschluss

Die Brennweite und Vergrößerung werden bereits im Rahmen des Abbildungsgesetzes erklärt, die restlichen Größen werden im Folgenden genauer vorgestellt.

Blendenzahl: Die Blendenzahl k ist der Quotient aus Blendendurchmesser d und Brennweite f (siehe Abb. 3.10).

$$k = \frac{d}{f} \quad (3.8)$$

Gemäß der international standardisierten Blendenreihe kann die Blende in folgenden Blendenzahl-Schrittweiten eingestellt werden:

1,0 1,4 2,0 2,8 4,0 5,6 8,0 11,0 16,0 22,0 32,0

Aufeinanderfolgende Blendenzahlen stehen im Verhältnis von $\sqrt{2}$ zueinander. Da die einfallende Lichtmenge proportional zum Quadrat des Blendendurchmessers ist, lässt sich mit einem Schritt zur nächsthöheren Blendenzahl die einfallende Lichtmenge halbieren. Die größte Blendenöffnung eines Objektivs bezeichnet man auch als Lichtstärke. [DSS11]

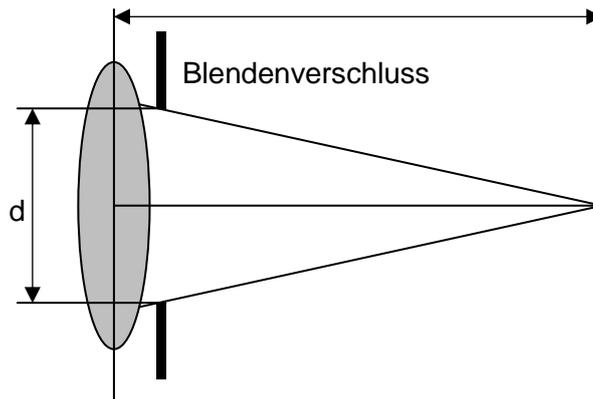


Abb. 3.10: Relative Öffnung eines Objektivs i. A. a. [DSS11]

Schärfentiefe: Gemäß der Abbildungsgleichung lässt sich der benötigte Abstand von Objekt zu Kamera bei gegebener Brennweite und Bildweite exakt berechnen. Dies geschieht unter der Annahme, dass der Gegenstandspunkt A zur scharfen Abbildung auf der Bildseite auch punktförmig dargestellt werden muss. In der Praxis sind die Sensorelemente zur Bildaufnahme (Pixel) jedoch nicht infinitesimal klein, sodass eine gewisse Unschärfe in der Punktabbildung akzeptabel ist. Die Abbildung muss somit nicht punktförmig sein, sondern kann flächig als sogenannter Zerstreuungskreis mit Durchmesser C abgebildet werden. Solange der Zerstreuungskreisdurchmesser kleiner als ein Sensorelement ist, kann ein Punkt mit einem einzigen Pixel abgebildet werden. Nach dem Abbildungsgesetz ist ein Gegenstand folglich nicht nur bei der exakten Gegenstandsweite g_0 scharf abbildbar, sondern innerhalb eines Entfernungsbereiches. Diesen Entfernungsbereich bezeichnet man als Schärfentiefe, die sich wie folgt berechnen lässt:

$$\Delta g \approx \frac{2 \times C \times k \times g_0 \times (g_0 - f)}{f^2} \text{ für } g_0 \ll \infty \quad (3.9)$$

Neben dem Zerstreuungskreisdurchmesser C , Blendenzahl k und Brennweite f ist der Schärfentiefebereich auch vom Nennabstand g_0 abhängig. Je größer der Nennabstand ist, desto größer ist auch der Schärfentiefebereich. Bei großen Nennabständen ($g_{\text{Unendlich}}$) werden alle Objekte scharf in der Brennebene abgebildet. Dies ist der Fall, wenn folgendes gilt:

$$g_{\text{Unendlich}} > \frac{f^2}{C \times k} \quad (3.10)$$

Eine zunehmende Brennweite führt hingegen zu einer abnehmenden Schärfentiefe, was sich für Teleobjektive als problematisch erweisen kann. Diesem Effekt lässt sich mittels der Blendensteuerung entgegenwirken. Durch das Verkleinern der Blende werden weniger Lichtstrahlen eingefangen, sodass die Aufweitung der Strahlenbündel und somit der Zerstreuungskreis kleiner wird. [DSS11]

Minimale Objektdistanz: Nach dem Abbildungsgesetz hängt die Bildweite von der Brennweite und der Gegenstandsweite ab. Rückt ein Objekt dichter an die

Optik heran, so muss die Bildweite erhöht werden. Der Abstand zwischen Linse und Bildsensor ist maximal, wenn der Objektivtubus bis zum Anschlag herausgedreht wird. Aus der maximal erreichbaren Bildweite sowie den Regeln der Schärfentiefe lässt sich die mindestens benötigte Distanz zwischen Objekt und Linse ableiten, um ein Objekt scharf abbilden zu können. Diese minimale Objektdistanz (MOD) wird für Objektive als Kennwert angegeben. Durch das Verwenden von Zwischenringen zwischen Kamera und Objektiv lässt sich die minimale Objektdistanz zusätzlich reduzieren. Jedoch können dann Gegenstände aus großer Entfernung nicht mehr scharf abgebildet werden. [DSS11; Neu05]

Auflösungsvermögen: Eine ideale Linse ist in der Lage, einen Objektpunkt auch als einzelnen Punkt in der Bildebene abzubilden. In der Realität begrenzen allerdings folgende Abbildungsfehler das Auflösungsvermögen einer Linse [Ste18; Thö04]:

- **Defektabweichungen:** Unebenheiten an der Oberfläche einer Linse führen zu einem veränderten Brechungsverhalten.
- **Chromatische Abweichungen:** Unterschiedliche Lichtwellenlängen werden vom Linsenmaterial unterschiedlich gebrochen.
- **Sphärische Abweichungen:** Bei sphärischen Linsen ändert sich der Brennpunkt in Abhängigkeit vom Lichteinfallort.

Die quantitative Beschreibung des tatsächlichen Auflösungsvermögens erfolgt durch die sogenannte Modulationstransferfunktion (MTF). Die MTF gibt an, welche Ortsfrequenzen zu einem bestimmten Kontrastwert mit dem Objektiv aufgelöst werden können. Die Ortsfrequenz wird dabei in Linienpaaren pro Millimeter (lp / mm), bezogen auf die Objektebene angegeben. Mittels normierter Linienabbildung (siehe Abb. 3.11) kann eine solche MTF-Kennkurve aufgenommen werden. Die Auflösungsbeschreibung eines Objektivs unterscheidet sich in diesem Sinne von der Auflösungsbeschreibung eines Bildsensors, die in Pixeln erfolgt. Der Kehrwert der Ortsfrequenz entspricht der Größe des kleinsten auflösbaren Strukturdetails. So entspricht z. B. eine Ortsfrequenz von 5 lp / mm einer auflösbaren Detailgröße von 0,1 mm (da 5 Linienpaare 10 Linien entsprechen). Über den Vergrößerungsmaßstab lässt sich dieser Wert auf die Bildebene umrechnen und dementsprechend die minimale Pixelgröße berechnen, die das Objektiv noch auflösen kann. In der Praxis wird das Auflösungsvermögen eines Objektivs daher auch durch die minimale Pixelgröße beschrieben, die mit einem Kontrastvermögen von 30% aufgelöst werden kann. Je kleiner dieser Wert ist, desto hochauflösender ist das Objektiv. [Ste18; Haf03; Thö04; KUB14]

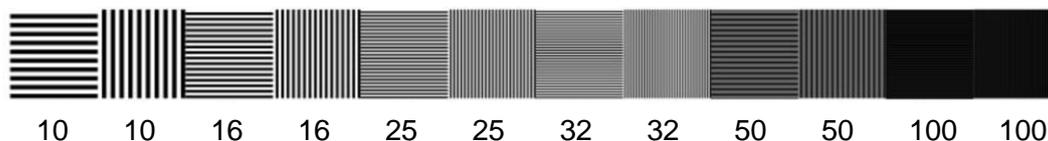


Abb. 3.11: Beispiel normierte Linienabbildung (hier skaliert abgebildet) zur MTF-Bestimmung [Ste18]

Homogenität der Ausleuchtung: Aufgrund unterschiedlich langer Lichtwege nimmt die Lichtintensität bei einem Objektiv vom Bildmittelpunkt in Richtung Rand hin ab. Objektivhersteller geben die Intensitätsverteilung in ihren Datenblättern in Diagrammen gemäß Abb. 3.12 an. Dem Abfall der Lichtintensität zum Bildrand hin kann mit einer Erhöhung der Blendenzahl entgegengewirkt werden. [Edm14]

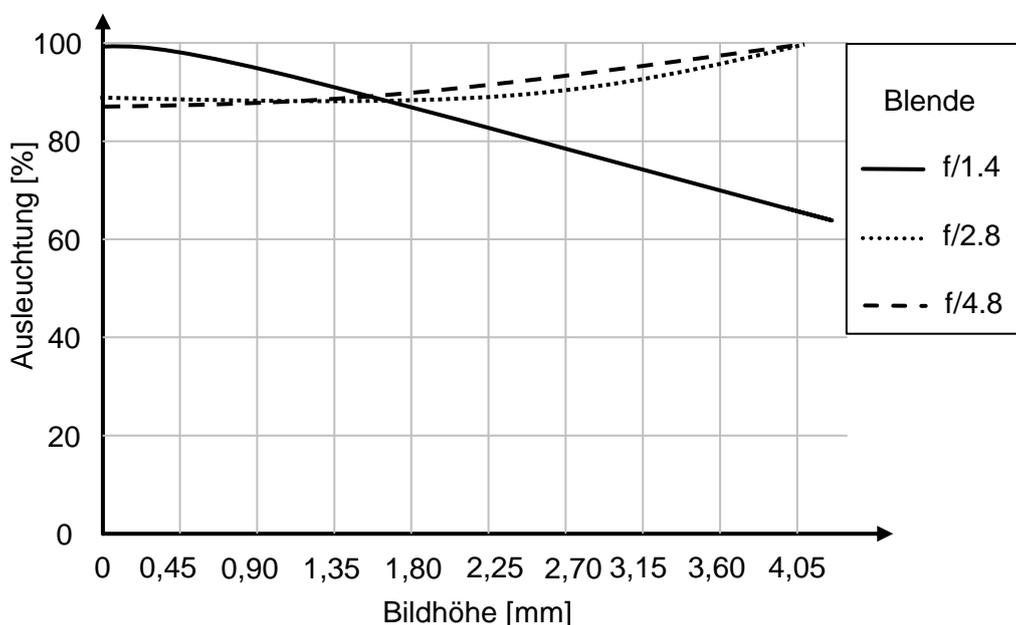


Abb. 3.12: Beispiel Helligkeitsverteilung Techspec Compact 6 mm Objektiv bei unterschiedlichen Blendenzahlen i. A a. [Edm14]

Verzeichnung: Im Idealfall erzeugt eine Optik eine exakte Wiedergabe des Objekts im Bildfeld ohne Verzerrungen. In der Realität weisen Linsen jedoch eine Verzeichnung auf, die exakte Messungen erschwert. Das Bild kann dabei entweder nichtlinear gedehnt oder gestaucht werden, siehe Abb. 3.13. Die Verzeichnung wird als Prozentsatz angegeben und gemäß folgender Formel berechnet:

$$\text{Verzeichnung} = \frac{\text{Ist-Distanz} - \text{Soll-Distanz}}{\text{Ist-Distanz}} \times 100 \quad (3.11)$$

Für die Berechnung der Verzeichnung bezieht man sich auf die Distanz vom Bildmittelpunkt zu den Bildecken, da hier die maximale Verzeichnung auftritt. Die Verzeichnung kann in der BV mittels mathematischer Verfahren reduziert werden,

jedoch ist dies aufgrund fehlender Informationen zur Bildtiefe nur eingeschränkt möglich. Aufwendige Linsenschleifungen und -anordnungen ermöglichen nahezu verzeichnungsfreie Aufnahmen. [Ste18; NH09]

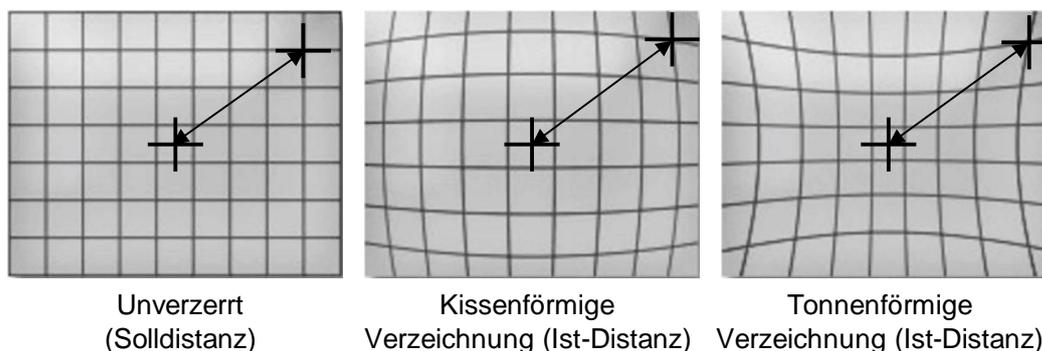


Abb. 3.13: Unterschiedliche Arten der Verzeichnung i. A. a. [Fer19]

Objektivgröße / Lichtkreisdurchmesser: Bei der Charakterisierung von Bildsensoren in Kapitel 3.1.1.2 wird auf die unterschiedlichen Sensorformatgrößen eingegangen. Diese müssen mit der Objektivgröße abgestimmt werden. Ist der Bildsensor größer als der Lichtkreis des Objektivs, so wird ein Teil der Pixel nicht ausgeleuchtet. Dieser als mechanische Vignettierung bezeichnete Effekt wird im linken Bild in Abb. 3.14 dargestellt [Ste18]. Die Größenangabe von Objektiven erfolgt wie bei den Bildsensoren in Zoll. Zu einem Objektiv mit einem Lichtkreisdurchmesser von einem Zoll passt dementsprechend ein Bildsensor, dessen Diagonale kleiner gleich ein Zoll ist.

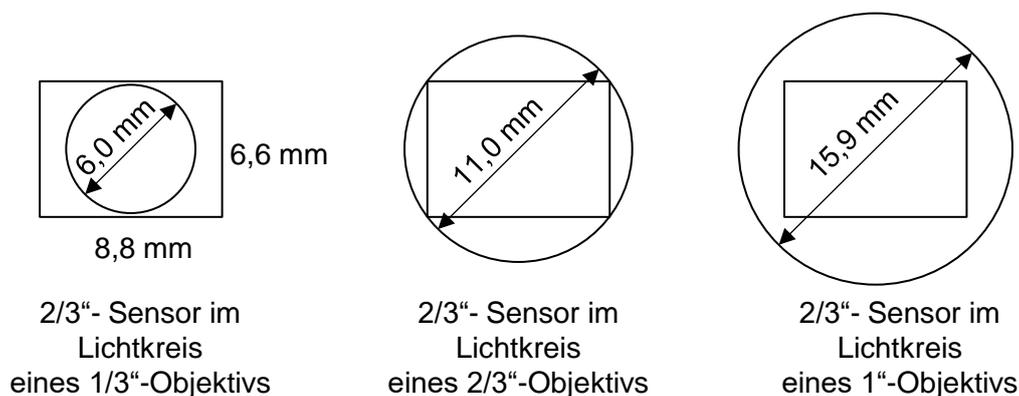


Abb. 3.14: Zusammenhang zwischen Sensor- und Objektivformat i. A. a. [Mat06]

Kameraanschluss: Zur Montage von Objektiven an Kameras sind standardisierte Befestigungsschnittstellen notwendig. Dies bezieht sich einerseits auf den Gewindeanschluss, andererseits auf das sogenannte Auflagemaß. Das Auflagemaß ist der Abstand von der Objektivauflage zur Sensorebene. Es ist ein konstruktives Maß der Kamera. Hierdurch wird der minimal mögliche Abstand zwischen Linse und Bildsensor und somit auch die minimale Bildweite b definiert. Ist das Objektiv

auf ein kleineres Auflagemaß als das tatsächliche Auflagemaß der Kamera ausgelegt, so lässt sich auch beim vollständigen Eindrehen des Objektivbusses die Bildweite b nicht auf die Brennweite f verkürzen. Objekte aus dem Unendlichen können also nicht mehr scharf abgebildet werden. [DSS11] In der industriellen BV ist der C-Mount mit einem Auflagemaß von 17,525 mm der am häufigsten verwendete Kameraanschluss. [Ste18]

3.1.3 Beleuchtungssystem

Unter einem Beleuchtungssystem wird im Kontext dieser Arbeit zusammenfassend das System aus Lichtquelle (Beleuchtungsmittel), zusätzlichen Elementen zur Lichtmanipulation sowie die relative Positionierung von Beleuchtungselementen zu Prüfobjekt und Bilderfassungseinheit verstanden. Nach einer grundlegenden Funktionserklärung werden in diesem Kapitel Beschreibungsgrößen zur Parametrisierung von Beleuchtungssystemen erläutert.

3.1.3.1 Grundlegende Funktionsweise von Beleuchtungsmitteln

Zum besseren Verständnis von Beleuchtungsmitteln werden in den beiden folgenden Abschnitten grundlegende physikalische Eigenschaften von Licht sowie die Funktionsweisen von LEDs erläutert.

Physikalische Eigenschaften des Lichtes

Licht ist der Informationsträger aller visuellen Informationen und somit auch die physikalische Grundlage von industriellen BV-Systemen. Licht tritt in der Form elektromagnetischer Strahlung auf, die über die Wellenlänge charakterisiert werden kann. Licht im menschlich sichtbaren Bereich hat eine Wellenlänge von 380 bis 780 nm. Den Wellenlängenbereich von 15 bis 380 nm bezeichnet man als ultraviolettes Licht, der Infrarotlichtbereich liegt bei Wellenlängen zwischen 780 nm und 1 mm. Bei der Erklärung der Wirkweise von Licht wird auf folgende Axiome zurückgegriffen: [Jah06]

- Licht breitet sich in Form elektromagnetischer Wellen aus. Es hat somit eine elektrische und eine magnetische Komponente.
- Licht breitet sich in homogenen und isotropen Medien in Form gerader Linien aus.
- Es besteht keine Interaktion zwischen Licht aus unterschiedlichen Lichtquellen.
- Am Übergang zwischen zwei verschiedenen Medien treten Brechungs- und Reflexionserscheinungen der Lichtstrahlen auf.
- Die Ausbreitungsgeschwindigkeit des Lichtes ist abhängig vom Ausbreitungsmedium.

- Die Ausbreitungsgeschwindigkeit c berechnet sich aus der Wellenlänge λ und der Frequenz f zu:

$$c = \lambda \times f \quad (3.12)$$

Die maximale Lichtgeschwindigkeit von 299.792 km / s wird im Vakuum erreicht.

Bei der Beschreibung von Lichtintensität unterscheidet die Literatur zwischen den Konzepten Radiometrie und Photometrie. Radiometrie ist die energetische Bewertung der Ausbreitung elektromagnetischer Wellen, die von einer Quelle ausgehend durch ein Medium zu einem Strahlungsdetektor gelangen. Energieangaben werden dabei in Joule gemessen. Das photometrische Intensitätskonzept entspricht dem der Radiometrie, jedoch wird hier beim Strahlungsdetektor von der spektralen Empfindlichkeit des menschlichen Auges ausgegangen. Im Kontext der BV ist daher zunächst das radiometrische Konzept von Relevanz, da die spektrale Empfindlichkeit von Kamerasensoren nicht der des menschlichen Auges gleicht. [Jah06; KBV14] In industriellen Datenblättern werden jedoch häufig auch photometrische Größen zur Beschreibung von weißem Licht verwendet. [Ste18] Photometrische und radiometrische Größen können über das von der Wellenlänge abhängige photometrische Strahlungsäquivalent ineinander umgerechnet werden.

Die grundlegende photometrische Größe ist die Lichtstärke I , sie wird in Candela (cd) angegeben. Ein Candela ist definiert als die Lichtstärke einer monochromatischen Lichtquelle mit einer Wellenlänge von 555 nm, die 1 / 683 Watt pro Steradian (W / sr) abstrahlt. Hieraus lassen sich, wie in Abb. 3.15 aufgeführt, die weiteren photometrischen sowie ihre radiometrischen Äquivalente herleiten. [Jah06]

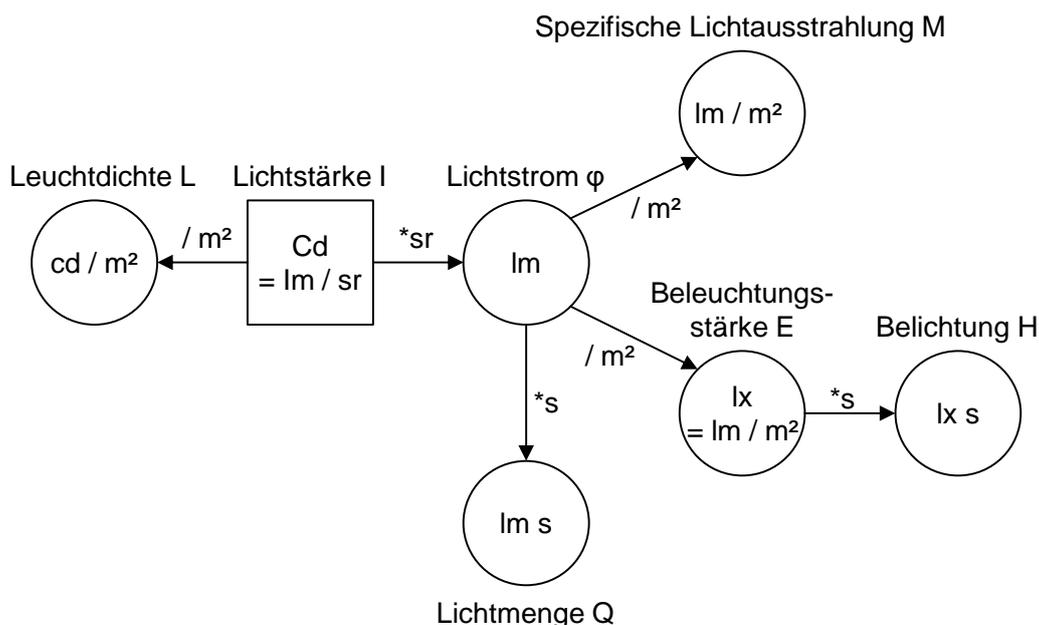


Abb. 3.15: Zusammenhang zwischen photometrischen Größen i. A. a. [Jah06]

Man unterscheidet drei grundlegende Arten der Interaktion zwischen Licht und Objekt:

- Reflexion: Das Objekt reflektiert das Licht von der Oberfläche zurück. Je nach Oberflächeneigenschaft wird ein Lichtstrahl direkt oder diffus reflektiert.
- Transmission: Das Licht wandert durch das Objekt hindurch. Es können dabei am Übergang Brechungseffekte auftreten.
- Absorption: Das einfallende Licht wird vom Objekt absorbiert.

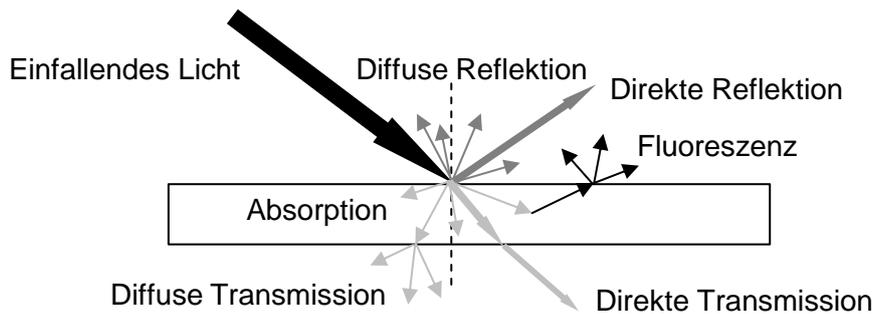


Abb. 3.16: Mögliche Arten der Interaktion von Licht mit Objekten i. A. a. [Jah06]

Diese lassen sich gemäß Abb. 3.16 in weitere Untergruppen untergliedern. Ein gewisser Anteil des Lichtes wird in der Praxis immer absorbiert. Zudem können auch alle drei Arten der Interaktion gleichzeitig auftreten. Verständlicher Weise kann ein Bildsensor nur Licht aufnehmen, das vom Objekt transmittiert oder reflektiert wird. Die Wellenlänge des Lichtes, der Einfallswinkel, die Materialzusammensetzung sowie die Oberflächenbeschaffenheit des Objekts sind wesentliche Faktoren, die über die anteilmäßige Aufteilung des Lichtes entscheiden. Gemäß des Energieerhaltungssatzes muss die Summe der reflektierten, transmittierten und absorbierten Anteile 100 % der einfallenden Strahlungsenergie entsprechen. Für reale Testobjekte sind die genauen jeweiligen Anteile unbekannt. [Jah06]

LEDs als Technologiegrundlage von Beleuchtungssystemen

Während in der älteren Literatur noch auf eine Bandbreite von unterschiedlichen Technologien als mögliche Lichtquelle eingegangen wird, hat sich in der Industrie inzwischen die LED als primäre Lichtquelle durchgesetzt [Jah06; KUH14]. Dies ist auf einen rapiden Preisverfall sowie eine technische Überlegenheit der LED gegenüber traditionellen Leuchtmitteln zurückzuführen, siehe Abb. 3.17 [KUH14; Jah06]. Einen Sonderfall stellen Laser dar, die in der industriellen BV weiterhin Einsatz finden. Sie werden zur strukturierten Beleuchtung von Oberflächen verwendet, gemäß Themendefinition stehen sie jedoch nicht im Fokus dieser Arbeit.

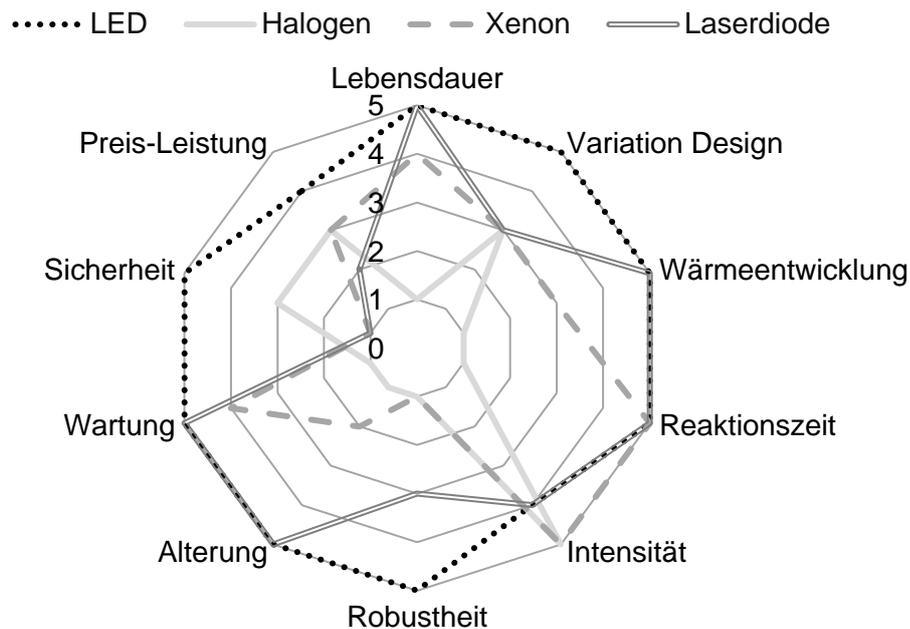


Abb. 3.17: Vergleich unterschiedlicher Lichtquellen zur industriellen Bildverarbeitung von 1 = schlecht bis 5 = sehr gut i. A. a. [KUH14; Jah06]

LEDs basieren auf dem Prinzip von sogenannten P/N-Halbleiterdioden (positiv / negativ). Bei der Anlegung einer Spannung in Durchgangsrichtung wandern Elektronen von der negativ dotierten Seite auf die positiv dotierte Seite, wobei das Elektron einen energetisch niedrigeren Zustand im Valenzband annimmt. Zur Energieabgabe beim Übergang strahlt das Elektron Photonen ab, die sich in einem schmalen Wellenlängenband befinden. [KBV14; Lic18] Sie werden daher als monochromatisch wahrgenommen.

Die Wellenlänge der Photonen ist im Wesentlichen vom gewählten Halbleitermaterial abhängig. [Lic18] Vom menschlichen Auge als weiß wahrgenommenes Licht setzt sich aus einem breiten Spektrum von Licht verschiedener Wellenlängen zusammen. Auch weißes Licht kann durch LEDs mittels einer Beschichtung aus phosphorsezierendem Material erzeugt werden. Das ursprünglich monochromatische Licht einer blauen LED wird dadurch in ein breites Wellenlängenspektrum gemäß Abb. 3.18 aufgefächert. [Ste18]

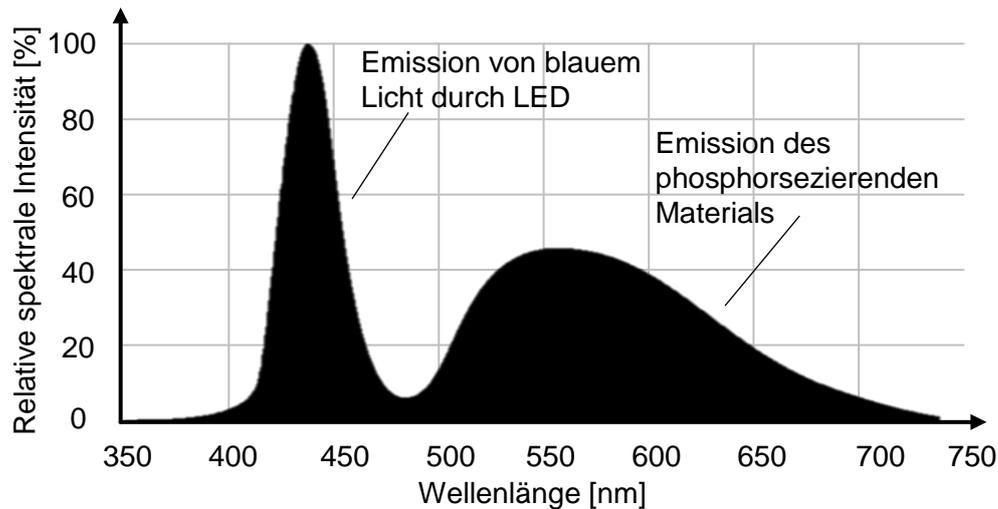


Abb. 3.18: Spektrale Emission von weißem LED-Licht i. A. a. [Lic18]

LEDs können kontinuierlich Licht aussenden. Ihre Lebenszeit lässt sich jedoch erhöhen, wenn man sie nur für den Moment der eigentlichen Belichtung einschaltet und anschließend direkt wieder abschaltet. [YH05] Die volle Leuchtstärke kann von einer LED ohne nennenswerte Verzögerung erreicht werden, siehe Abb. 3.19.

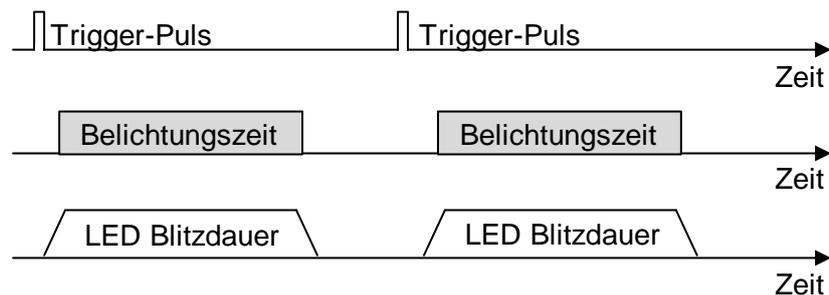


Abb. 3.19: Schalten eines LED-Blitzlichts i. A. a. [Ste18]

Werden LEDs nur für kurze Zeitintervalle eingeschaltet, so können sie auch mit einer Leistung deutlich über dem Nominalwert des Dauerbetriebs verwendet werden. Diese erhöhte Lichtintensität ist jedoch nur möglich, wenn die LEDs zwischen den einzelnen Pulsen ausreichend Zeit zum Abkühlen haben. Da eine Überhitzung zur Beeinträchtigung oder gar vollständigen Schädigung einer LED führen kann, werden in der Industrie je nach Leistungsbedarf unterschiedliche Konzepte zur Kühlung angewendet. [LCB09]

3.1.3.2 Beschreibungsgrößen von Beleuchtungssystemen

Die hier aufgelisteten Größen dienen der Charakterisierung von Beleuchtungssystemen und werden anschließend genauer erklärt:

- Beleuchtungswinkel
- Wellenlänge
- Lichtausbreitung
- Beleuchtungsstärke
- Homogenität
- Beleuchtungsfläche
- Beleuchtungssteuerung
- Lichtmanipulation

Die wissenschaftliche Literatur führt weitaus mehr Größen zur Beschreibung von Lichtquellen auf. Diese sind jedoch häufig eher theoretischer Natur und daher nicht direkt für die Auswahl eines Beleuchtungssystems im industriellen BV-Kontext anwendbar.

Beleuchtungswinkel: Bei der Beleuchtungsanordnung unterscheidet man zwischen Auflicht- und der Hintergrundbeleuchtung als grundlegende Techniken. Wie Abb. 3.20 zu entnehmen, befindet sich bei einer Auflichtbeleuchtung die Lichtquelle auf der gleichen Seite wie die Kamera. Bei der Hintergrundbeleuchtung ist diese hinter dem Prüfobjekt orientiert. Diese beiden Beleuchtungsanordnungen lassen sich jeweils in Hellfeld und Dunkelfeld Anwendungen unterteilen. Bei einem Hellfeld reflektiert eine ebene Oberfläche die einfallenden Strahlen in Richtung der Optik. Durch diesen Lichteinfall wird die Oberfläche im Bild entsprechend hell dargestellt. Bei einer Dunkelfeldbeleuchtung reflektiert eine glatte Oberfläche einfallendes Licht nicht in Richtung der Optik, sie erscheint im Bild folglich als dunkel. Nur an Unebenheiten oder Kanten des Objekts wird das Licht gestreut und ein Teil davon in Richtung Optik reflektiert (siehe Abb. 3.20). Dadurch werden Unebenheiten im Bild gezielt hell dargestellt. [Jah06; JIA LI-001; Per10]

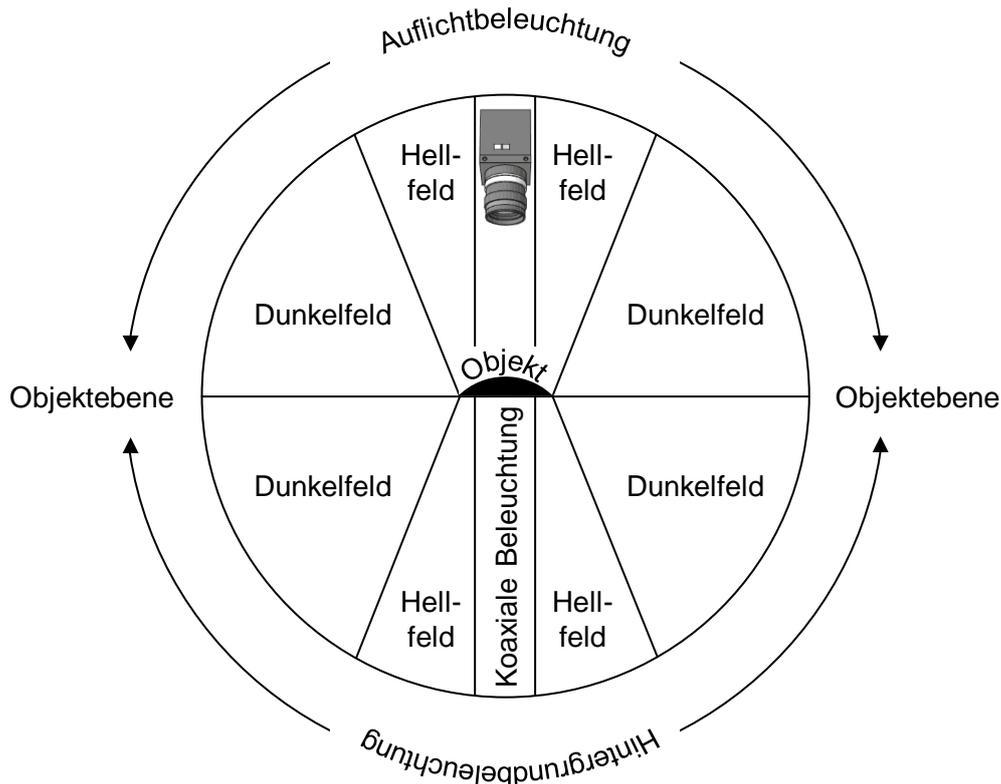


Abb. 3.20: Unterschiedliche Beleuchtungswinkel i. A. a. [Per10]

Wellenlänge: Die Wellenlänge / Farbe des Lichts spielt nicht nur für Farbanwendungen, sondern auch für die Aufnahme monochromatischer Bilder eine wichtige Rolle. Abb. 3.21 zeigt, wie sich die monochromatische Helligkeitsdarstellung der Farben rot, grün und blau in Abhängigkeit der Lichtfarbe verändert. Anwendungen, bei denen einzelne Objektfarben nicht signifikant hervorgehoben werden sollen, werden daher mit weißem Licht ausgeleuchtet. Eine gezielte Verringerung von Farbeffekten lässt sich zudem mit Infrarotlicht erzielen [Ste18; VDI 2632]. Die Verwendung von Licht im kurzwelligen Infrarotbereich und Kamerasensoren mit einer entsprechenden spektralen Empfindlichkeit ermöglicht außerdem die Aufnahme von Merkmalen, die im sichtbaren Spektralbereich nicht erkennbar wären. Das auch als Chemical Colour Imaging bezeichnete Verfahren basiert auf den unterschiedlichen Absorptionsverhalten chemischer Elemente im Infrarotlichtbereich. Wasser etwa absorbiert Infrarotlicht im Bereich vom 1450 nm. Mit einer Beleuchtung von 1450 nm können Füllstände von Flüssigkeiten in undurchsichtigen Behältern überprüft werden. [Ste18]

Für das menschliche Auge nicht sichtbares ultraviolettes Licht (UV) kann zur Vermessung sehr kleiner Merkmale verwendet werden. So ist gemäß des Rayleigh-Kriteriums die minimale auflösbare Objektgröße durch die Lichtwellenlänge limitiert. Mit sichtbarem Licht können Merkmale bis zu einer Größe von etwa 500 nm vermessen werden. Eine Reduktion der Wellenlänge in den UV Bereich

gilt als einfache Möglichkeit zur Erzielung eines höheren Auflösungsvermögens. [Ste18]

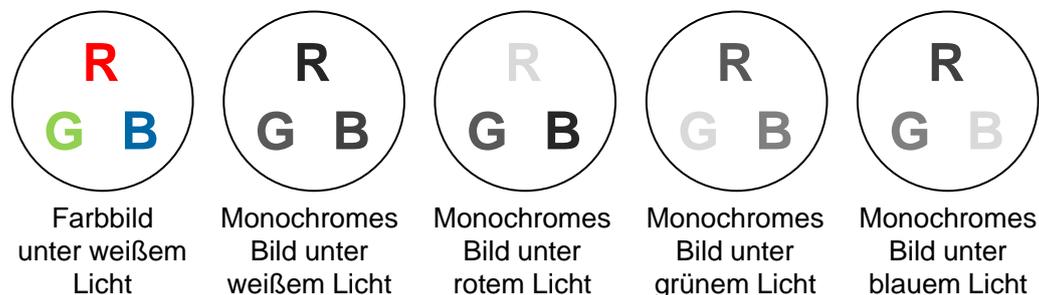


Abb. 3.21: Farbdarstellung in monochromatischen Bildern in Abhängigkeit der Lichtfarbe i. A. a. [Ste18]

Lichtausbreitung: Die Ausbreitungsrichtung von Licht kann auf unterschiedliche Weisen manipuliert werden. In der Literatur unterscheidet man hierzu zwischen direkter, diffuser, fokussierter, kollimierter und strukturierter Beleuchtung.

- **Direktes Licht** fällt von der Lichtquelle direkt auf das Prüfobjekt, ohne in seiner Ausbreitung manipuliert zu werden. Ist die Lichtquelle wie bei einer LED annähernd punktförmig, so erzeugt sie auf dem Objekt ein Lichtfeld, das vom Zentrum abnehmend stark an Helligkeit verliert. Aufgrund dieser geringen Homogenität findet direktes Licht in der industriellen BV kaum Anwendung. [Jah06; Ste18]
- **Diffuses Licht** wird durch die Verwendung von transparenten Diffusoren erzeugt. Der Diffusor dämpft und streut direktes Licht. Das Licht wird vom Diffusor gemäß eines Lambert-Strahlers in alle Richtungen gleichmäßig ausgestrahlt, sodass ein homogenes Helligkeitsfeld entsteht. [Jah06]
- **Fokussiertes Licht** erlaubt die gezielte punkt- oder linienförmige Ausleuchtung von Objekten. Mit einer Fokussierlinse werden Lichtstrahlen gebündelt, sodass der fokussierte Objektbereich homogen und mit hoher Intensität ausgeleuchtet werden kann. Dies ist z. B. für Objekte relevant, bei denen aufgrund der hohen Durchlaufgeschwindigkeit sehr kurze Belichtungszeiten notwendig sind. Zudem kann mit einer Fokussierlinse nur der für die BV relevante Teilbereich eines Objekts ausgeleuchtet werden. [Jah06]
- **Kollimiertes Licht** bezeichnet parallel verlaufende Lichtstrahlen. Die von einer Lichtquelle radial ausgehenden Lichtstrahlen werden dabei mittels einer Richtlinse parallelisiert. Kollimiertes Licht erlaubt die Erzeugung von scharfen Kontrasten auch im Fall von fließenden Bildübergängen. Für präzise Vermessungsaufgaben wird kollimiertes Licht in Kombination mit einer telezentrischen Optik verwendet. [Ste18; Jah06]

- **Strukturiertes Licht** bezeichnet die Projektion von Mustern auf das Prüfobjekt. Über die Verzerrung des Musters und die relative Position von Kamera zu Prüfobjekt lassen sich Rückschlüsse auf die Objektgeometrie berechnen. Dadurch lassen sich auch dreidimensionale Merkmale oder Kanten mit sehr geringen Kontrasten mittels zweidimensionaler Bildaufnahme analysieren. [Ste18; DSS11]

Beleuchtungsstärke: Die geläufigste Größe zur Beschreibung der Intensität des Lichtes im Kontext der Bildaufnahme ist die Beleuchtungsstärke E_v . Sie entspricht dem Lichtstrom je Flächeneinheit und wird in Lux (lx) angegeben. Sie lässt sich gemäß der in Abb. 3.15 vorgestellten Relationen in andere photometrische und radiometrische Größen überführen. Die von einer punktförmigen Strahlungsquelle ausgehende Beleuchtungsstärke E_v nimmt proportional mit dem Quadrat der Distanz d ab.

$$E_v \sim \frac{1}{d^2} \quad (3.13)$$

Die an der Objektoberfläche erreichte Beleuchtungsstärke entspricht jedoch noch nicht der am Bildsensor gemessenen Beleuchtungsstärke. Je nach Objektbeschaffenheit verändert sich das Verhältnis von reflektiertem, transmittiertem und absorbiertem Licht. Nur der vom Objekt reflektierte Lichtanteil erreicht die Kameraoptik. Die Kameraoptik wiederum transmittiert nur einen Teil des einfallenden Lichtes. Die am Bildsensor erreichte Beleuchtungsstärke E_{Sensor} lässt sich in Abhängigkeit des Objektreflexionsanteils ρ , des Transmissionsanteils τ der Optik mit Blendenzahl k und der in der Objektebene auftretenden Lichtstärke E_v wie folgt ausdrücken:

$$E_{\text{Sensor}} = E_v \times \rho \times \frac{\tau}{4 \times k^2} \quad (3.14)$$

Es handelt sich hierbei um eine vereinfachte Approximation. Die tatsächlich am Sensor auftretende Beleuchtungsstärke wird u.a. von der Wellenlänge, dem Einfallswinkel des Lichtes und perspektivischen Eigenschaften der Linse beeinflusst. [Jah06; JIIA LI-002; VDI 2632]

Homogenität: Die gemäß Gleichung 3.13 von einer Lichtquelle auf einer Oberfläche mit Abstand d erzeugte Beleuchtungsstärke ist in der Realität nicht, wie zunächst angenommen, homogen. Sie variiert stattdessen lokal in Abhängigkeit vom Strahlengang. Die örtliche Homogenität der Ausleuchtung wird von Herstellern mittels grafischer Darstellungen wie in Abb. 3.22 angegeben. Für die meisten Anwendungen ist eine möglichst homogen ausgeleuchtete Objektoberfläche erwünscht. [Jah06; VDI 2632; CCS18]

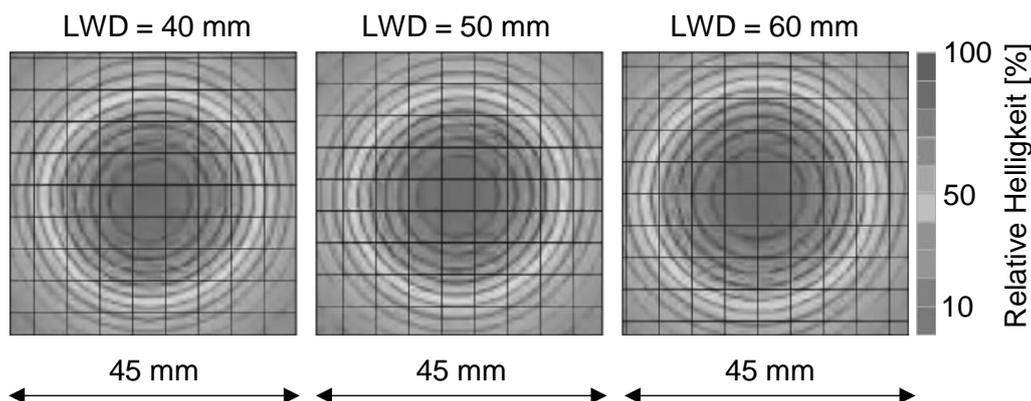


Abb. 3.22: Relative Helligkeitsverteilung eines Ringlichtes in Abhängigkeit vom Abstand zum Werkstück (LWD) [CCS18]

Beleuchtungsfläche: Die Größe der beleuchteten Fläche hängt neben der Art der Lichtausbreitung auch von der Größe des lichtgebenden Mediums ab. Hersteller von Beleuchtungssystemen bieten daher die Ausführung eines Produktes (z. B. Ringleuchten oder Lichtfelder) in mehreren Größen an. [JIA LI-002; VDI 2632]

Beleuchtungssteuerung: Wie bereits in Abb. 3.19 erläutert, lässt sich die Einschaltdauer einer LED variieren. Wird die Beleuchtung nur für sehr kurze Zeit aktiviert, so bezeichnet man dies auch als Blitzbetrieb. Der Blitzbetrieb kann über die maximale Pulsdauer und die Wiederholungsrate charakterisiert werden. Die Beleuchtungsintensität einer LED kann zudem variabel gesteuert werden. Dies eignet sich einerseits zur einfachen Anpassung der benötigten Helligkeit an eine Prüfaufgabe. Andererseits kann die Intensitätskontrolle genutzt werden, um Alterungseffekte und den damit verbundenen Intensitätsabfall auszugleichen. [Ste18; LUW13]

Lichtmanipulation: Die Qualität des Lichtes lässt sich durch verschiedene Manipulationstechniken verändern. Dazu verwendet man Filter, die vor der Lichtaustrittsfläche und / oder im Objektiveneingang angebracht werden. Man unterscheidet zwischen den folgenden Filterkategorien:

- **Polarisationsfilter** unterdrücken Lichtwellen, die in eine bestimmte Richtung schwingen. Dies ist nützlich zur Unterdrückung von spiegelnden Reflexionen an ebenen Oberflächen, die häufig parallel zur Reflexionsoberfläche polarisiert sind. [Ste18; Jah06]
- **Farbfilter** unterdrücken ausgewählte Wellenlängenbereiche. Man unterscheidet dabei zwischen Kurzpassfiltern, Langpassfiltern und Bandpassfiltern. Kurzpassfilter blockieren Licht mit Wellenlängen größer einem spezifizierten Schwellwert. Entsprechend blockieren Langpassfilter Wellenlängen kleiner dem definierten Schwellwert. Bandpassfilter ermöglichen das gezielte Durchlassen von Licht in einem Wellenlängenintervall. Solche Filter sind nicht nur auf

den menschlich sichtbaren Farbbereich beschränkt, sondern kommen auch im UV- und IR-Bereich zum Einsatz. [Jah06; Ste18]

3.1.4 Zwischenergebnis Hardwaretechnologie

Auf Basis der Literaturlaufarbeitung der vorherigen Kapitel werden in Abb. 3.23 die vorgestellten Beschreibungsparameter für Hardwarekomponenten und deren Interdependenzen zusammenfassend dargestellt. Eine mit einem Kreuz markierte Zeilen-Spalten Paarung bedeutet dabei, dass diese beiden Größen direkt miteinander zusammenhängen. So hängt z. B. das optische Format eines Bildsensors (ID 17) von der Auflösung (ID 15) und Pixelgröße (ID 16) ab. Zusätzlich sind in der Übersicht Beschreibungsgrößen zur Spezifizierung des Anwendungsfalls aufgeführt. Diese werden systematisch in Kapitel 5 hergeleitet.

3.2 Bildverarbeitungsalgorithmen

Nach der Bilderfassung bilden Algorithmen zur Bildanalyse den zweiten Technologiebereich, der in dieser Arbeit genauer untersucht wird. Hierbei soll gemäß der wissenschaftlichen Fragestellung (siehe Kapitel 1.2) zwischen konventionellen Algorithmen und Methoden des maschinellen Lernens (ML), insbesondere des Deep Learning (DL), unterschieden werden.

Unter einem Algorithmus versteht man in der Informatik eine eindeutige Vorschrift zur Lösung eines Problems. Algorithmen bestehen dabei aus endlich vielen wohldefinierten Einzelschritten. [LRS10]

Maschinelles Lernen ist ein Teilgebiet der Informatik und dient als Oberbegriff für künstliche Generierung von Wissen aus Erfahrung. Das künstliche System lernt dabei zunächst aus Beispielen und kann diese nach Beendigung der Lernphase verallgemeinern. [DLV18] Auch das maschinelle Lernen basiert auf Algorithmen. Diese definieren jedoch nicht explizite Regeln zur Lösung eines spezifischen Problems. Stattdessen beschreiben die Algorithmen hierbei Vorschriften, anhand derer das System eigenständig Regeln zur Problemlösung erlernt.

In Abgrenzung dazu sollen unter konventionellen Algorithmen daher jene Algorithmen verstanden werden, bei denen spezifische Vorschriften zur Problemlösung explizit vom menschlichen Nutzer definiert werden. [WMZ+18]

In Kapitel 3.2.1 wird zunächst die auf konventionellen Algorithmen basierende etablierte Vorgehensweise für Bildinspektionen vorgestellt und eine Übersicht entsprechender Algorithmen gegeben. Im Anschluss wird in Kapitel 3.2.2 auf die Besonderheiten von Methoden des Deep Learning im Kontext der industriellen BV eingegangen und erläutert, inwiefern diese Methoden konventionelle Algorithmen ergänzen oder ersetzen können.

3.2.1 Konventionelle Bildverarbeitungsalgorithmen

In Praxis und Fachliteratur existiert eine große Anzahl an verschiedenen Algorithmen zur Bildverarbeitung, auf die die Definition der konventionellen Algorithmen zutrifft. So verfügt z. B. die BV-Software Cognex Insight Explorer über ca. 90 verschiedene Algorithmen zur Bildanalyse. Aufgrund dieser Komplexität ist es im Rahmen der Arbeit nicht möglich, alle in der Literatur aufgeführten Algorithmen und ihre jeweiligen Parameter im Detail vorzustellen. Stattdessen werden hier generische Klassen von Algorithmen kurz vorgestellt. Für eine detailliertere Erläuterung einzelner Algorithmen sei auf die zugrundeliegende Fachliteratur verwiesen [DSS11; Ste06; SUW18; GW08; SHR08].

Zur besseren Kategorisierung von Algorithmen wird zunächst der typische Bildverarbeitungsprozess dargestellt. Die Gesamtaufgabe der Bildüberprüfung lässt sich

in mehrere Teile zerlegen, die unabhängig voneinander durchgeführt werden können. Dazu unterscheidet [DSS11] zwischen den folgenden Schritten:

- 1) **Bildaufnahme:** Dieser Schritt entspricht dem in Kapitel 3.1.1.1 vorgestelltem Vorgehen zur Generierung von digitalen Bildern.
- 2) **Bildvorverarbeitung:** Hierbei werden die Eigenschaften von aufgenommenen Bildern nachträglich verändert, um nachfolgende Untersuchungsschritte zu vereinfachen. So können z. B. Kontraste nachträglich erhöht oder ausgewählte Strukturen hervorgehoben werden.
- 3) **Setzen von (manuell selektierten) Arbeitsbereichen:** In vielen Fällen ist es sinnvoll, den Bereich des Bildes für weitere Untersuchungen einzuschränken. So kann etwa der Bereich limitiert werden, in dem ein Schriftzug zu lesen oder die Anwesenheit eines Gegenstandes kontrolliert werden soll. Dadurch lässt sich der weitere Rechenaufwand reduzieren. Zudem können für unterschiedliche Überprüfungsaufgaben in einem Bild mehrere Arbeitsbereiche definiert werden. Diese Bereiche bezeichnet man auch als Region of Interest (ROI).
- 4) **Erzeugung von Objekten:** Um Aussagen über ein Objekt im Bild machen zu können, muss dieses Objekt zunächst definiert werden. Möchte man etwa den Durchmesser einer Bohrung messen, so muss zunächst ein Kreisobjekt aus dem Bild isoliert werden. Diesen Schritt bezeichnet man auch als Segmentierung.
- 5) **Berechnung von Objektmerkmalen:** Nach der Isolierung eines einzelnen Objekts können dessen charakteristische Objektmerkmale berechnet werden. Dies können z. B. die eingeschlossene Fläche, Abstände oder räumliche Objektorientierung sein.
- 6) **Klassifizierung:** Anhand der berechneten Objektmerkmale wird abschließend eine Aussage über das Objekt getroffen. Diese Aussage ist das finale Qualitätsurteil der Bildinspektion.

Nicht für jede Prüfaufgabe wird auch jeder dieser Prozessschritt benötigt. So kann es z. B. bei einer Oberflächenprüfung sein, dass lediglich die Helligkeitsverteilung im definierten Arbeitsbereich von Relevanz ist. Es müssen also keine speziellen Objekte definiert werden. [DSS11]

Dieses prozessorientierte Schema gibt die grobe Klassifizierung von konventionellen Algorithmen vor, welches im Folgenden weiter ausdetailliert wird. Eine zusammenfassende Übersicht verschiedener konventioneller Bildverarbeitungsalgorithmen ist Anhang 11.2 zu entnehmen

Prozessschritt 3 zur Definition von Arbeitsbereichen wird hierbei mit in die Oberklasse der Objektsegmentierung eingeordnet, da sich diese beiden Verarbeitungsschritte bzgl. ihrer zugrundeliegenden Algorithmen stark ähneln. Die verschiedenen Algorithmenklassen werden im Folgenden weiter erklärt.

3.2.1.1 Bildvorverarbeitung

Bei den Algorithmen der Bildvorverarbeitung unterscheidet man zwischen drei unterschiedlichen Arten der Bildmanipulation:

- **Photometrische Bildvorverarbeitung** umfasst die gezielte Manipulation von Helligkeitswerten des Ausgangsbildes.
- **Geometrische Bildvorverarbeitung** (auch Koordinatentransformation) beschreibt die Veränderung der Lage von Bildpunkten zueinander.
- **Integraltransformationen** überführen die örtliche pixelgebundene Darstellungsform von Helligkeitswerten in den Frequenzraum, z. B. mittels Fourier- oder Wavelet-Transformation. Bild Darstellungen im Frequenzraum eignen sich insbesondere zur translations- / rotationsunabhängigen Analyse von Oberflächentexturen. Aufgrund der fortgeschrittenen Mathematik wird dieser Algorithmenbereich hier nicht weiter ausgeführt. Detaillierte Erklärungen finden sich z. B. bei [GW08; Rus07; NFH07].

Photometrische Bildvorverarbeitung

Bei photometrischen Bildvorverarbeitungsalgorithmen wird weiterhin zwischen Punktoperationen und lokalen Operationen unterschieden. Bei Punktoperationen wird zur Helligkeitsberechnung des Ergebnispixels ausschließlich die Helligkeit des korrespondierenden Punktes aus dem Eingangsbild verwendet. Bei lokalen Operationen hingegen werden zusätzlich die benachbarten Pixel des korrespondierenden Punktes zur Berechnung des Pixelhelligkeitswertes hinzugezogen. [DSS11]

Punktoperationen

Die photometrischen Punktoperationen unterteilt [DSS11] in Operatoren der Grauwerttransformation und der Bildarithmetik. Bei Grauwerttransformationen wird der Helligkeitswert eines Pixels mittels derselben ortsunabhängigen Funktion in einen neuen Helligkeitswert umgerechnet:

$$o_{ij} = f(i_{ij}) \quad (3.15)$$

Darin ist o_{ij} der Grauwert an den Pixelkoordinaten $x = i$ und $y = j$ im Ergebnisbild, i_{ij} ist der Grauwert an der gleichen Stelle im Eingangsbild. Die Abbildungsfunktion $f()$ kann dabei auf verschiedene Arten realisiert werden. Einige Beispiele für solche Abbildungsfunktionen sind:

- **Lineare Grauwertskalierungen**, bei denen Bildhelligkeit und Bildkontrast über die Koeffizienten c_1 und c_2 manipuliert werden:

$$o_{ij} = c_2 \times (c_1 + i_{ij}) \quad (3.16)$$

Ist $c_1 > 0$, so wird das Bild insgesamt heller. Für $c_1 < 0$ wird das Bild entsprechend dunkler. Mit $|c_2| > 1$ wird das Ausgangsbild kontrastreicher, für $|c_2| < 1$ kontrastärmer.

- **Kontrastnormierungen**, die einzelne Helligkeitswerte über den kompletten Grauwertbereich (bei 8-Bit Codierungen von 0 bis 255) normieren und so den Bildkontrast erhöhen (siehe Abb. 3.24). Wenn i_{\min} der minimale Grauwert und i_{\max} der maximale Grauwert des Eingangsbildes sind, so ergibt sich für die Pixel des Ausgangsbildes bei einer 8-Bit Signal-Codierung folgender Helligkeitswert:

$$o_{ij} = \frac{255}{i_{\max} - i_{\min}} \times (i_{ij} - i_{\min}) \quad (3.17)$$

- **Look-Up-Tabellen**, in denen für jeden eingehenden Helligkeitswert ein Ausgangshelligkeitswert hinterlegt wird. Dies ermöglicht die Erzeugung beliebiger, auch nichtlinearer Umwandlungen.

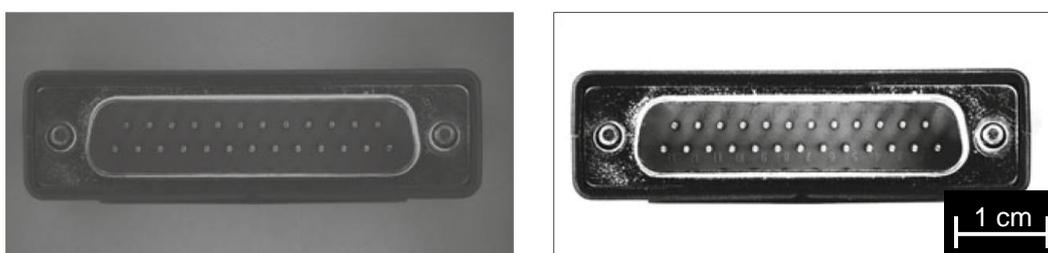


Abb. 3.24: Beispiel einer Kontrastnormierung [DSS11]

Bei der Bildarithmetik werden zwei oder mehrere Bilder punktweise miteinander verrechnet. Die Abbildungsfunktion ist also weiterhin punktgebunden, im Gegensatz zur Grauwerttransformation jedoch nicht mehr ortsunabhängig, da die Pixel der verrechneten Bilder lokal unterschiedliche Helligkeitswerte besitzen können. Formal ergibt sich der Helligkeitswert des Pixels o_{ij} des Ausgangsbildes aus den n Pixelhelligkeitswerten der Eingangsbilder i_{ij} :

$$o_{ij} = f(i_{0,ij}, i_{1,ij}, \dots, i_{n,ij}) \quad (3.18)$$

Mittels geeigneter Anwendung von Addition, Subtraktion oder Mittelung können so ein zu untersuchendes Bild und ein Referenzbild miteinander verrechnet werden, um etwa ungewollte Schwankungen der Lichtverhältnisse oder verschiedene Druckfehler zu identifizieren. Dies ist beispielhaft in Abb. 3.25 dargestellt, in der das Bild auf der linken Seite als Referenzbild dient. Das mittlere Bild ist das zu untersuchende Bild, welches einen Fehldruck beinhaltet. Es fehlt der Buchstabe r , zudem ist ein heller Fleck auf dem Objekt. Durch Subtraktion des mittleren Bildes vom Referenzbild ergibt sich das rechte Bild, in dem lediglich die beiden fehlerhaften Objekte zu sehen sind. In diesem Beispiel muss nach der Subtraktion noch der Absolutbetrag der berechneten Grauwerte gebildet werden. Ansonsten würden zu helle Pixel im mittleren Bild (siehe heller Fleck) in der Bilddifferenz zu einem nicht darstellbaren negativen Grauwert führen. [DSS11]



Abb. 3.25: Referenzbild (links), zu untersuchendes Bild (Mitte), Bilddifferenz (Rechts) [DSS11]

Lokale Operatoren

Im Gegensatz zu den Punktoperationen beziehen lokale Bildoperatoren eine Gruppe von mehreren Pixeln im Eingangsbild zur Berechnung des Helligkeitswertes eines Pixels im Ausgangsbild ein. Soll z. B. der Helligkeitswert an der Pixelkoordinate (6,7) berechnet werden und alle acht benachbarten Pixel in die Berechnung mit einbezogen werden, so ergibt sich der neue Helligkeitswert aus der Verrechnung der neun Pixelwerte an den Koordinaten (5,6), (5,7), (5,8), (6,6), (6,7), (6,8), (7,6), (7,7), (7,8) des Eingangsbildes. Diese Auswahl benachbarter Pixel bezeichnet man auch als lokale Umgebung. Die Größe der lokalen Umgebung muss vom Benutzer festgelegt werden, sehr typisch sind z. B. Filter der Größe 3 x 3 und 5 x 5 Pixel. Bezüglich der Verrechnungsfunktion der Pixelregion lässt sich zwischen verschiedenen Filterarten unterscheiden [SUW18].

Lineare Filter resultieren aus der gewichteten Summierung von Helligkeitswerten in einer definierten Pixelnachbarschaft. Die beiden häufigsten Formen der linearen Filter sind Glättungsfiler und Kantenfilter. Glättungsfiler bilden einen Mittelwert aus den einzelnen Helligkeitswerten der Pixelregion, wodurch Bildrauschen oder feine Strukturdetails unterdrückt werden können. Dies lässt sich in Abb. 3.26 am mittleren Bild veranschaulichen. Die Anwendung linearer Glättungsfiler führt hier dazu, dass einzelne Störpixel sich weniger vom Hintergrund absetzen. Gleichzeitig umfassen diese Störungen nun aber auch mehrere Pixel, da der Störpixel in die Mittelwertbildung der benachbarten Pixel mit einfließt. Dieser Effekt führt auch zur Kantenunschärfe des schwarzen Rechtecks. Im Gegensatz zu Glättungsfilern werden Kantenfilter verwendet, um Variationen im Bild bewusst hervorzuheben. Dies wird durch die Differenzbildung von Grauwerten benachbarter Pixel ermöglicht, was in Abb. 3.27 veranschaulicht wird. An den Übergängen des hellen Hintergrunds zum dunklen Oktaeder ist die Differenz der Helligkeitswerte benachbarter Pixel (weiß zu schwarz) sehr hoch, was zu einer Kantendarstellung mit entsprechend hohen Helligkeitswerten führt. Innerhalb des weißen Hintergrundes und des dunklen Farbverlaufs des Oktaeders sind die Helligkeitswerte benachbarter Pixel gleich oder ähnlich hoch, sodass die Differenz entsprechend gering ist und dunkel dargestellt wird. (Anmerkung: Der hier dargestellte Filter deckt nur Kanten in

vertikaler Richtung auf, sodass die beiden horizontalen Kanten nicht detektiert werden). [SUW18; DSS11]

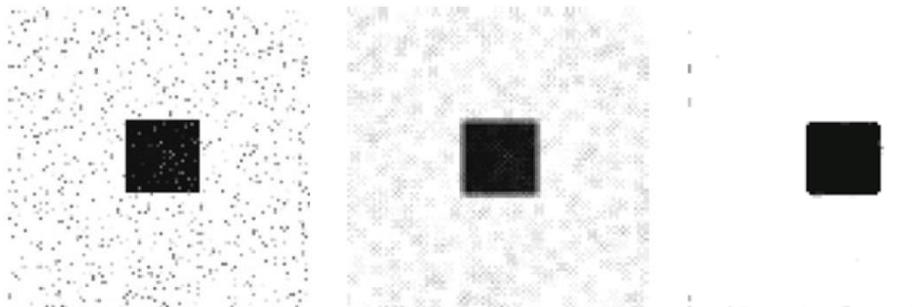


Abb. 3.26: Ausgangsbild (links), Bild nach Glättungsfilter (mittig), Bild nach Medianfilter (rechts) [DSS11]

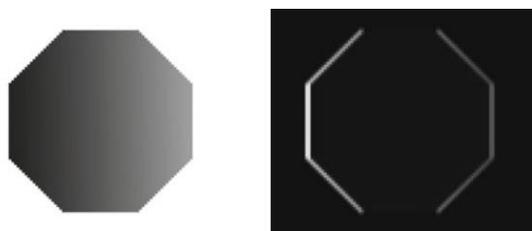


Abb. 3.27: Ausgangsbild (links), Bild nach Kantenfilter (rechts) [DSS11]

Medianfilter basieren auf der Medianwertbildung von benachbarten Helligkeitswerten. Während bei Filtern mit Mittelwertbildung einzelne Störpixel zu einer deutlichen Veränderung des Grauwertes der benachbarten Pixelumgebung führen, können mit Medianfiltern einzelne Störpixel vollständig unterdrückt werden (siehe Abb. 3.26 rechtes Bild). Zudem bleibt hier auch die Kantenschärfe überwiegend erhalten. [SUW18; DSS11]

Morphologische Filter beruhen auf den mengentheoretischen Basisoperationen der Erosion und Dilatation. Im Falle der Erosion ergibt sich der neue Helligkeitswert eines Pixels aus dem Minimum aller Helligkeitswerte innerhalb der betrachteten Pixelnachbarschaft. Dies führt zu einem gezielten Verdünnen von hellen Strukturen, da bereits ein einzelner dunkler Punkt innerhalb des Filteroperators zur Verdunklung des Zentralpunktes führt. Eine Dilatation ergibt sich im Gegensatz zur Erosion aus dem Maximalwert innerhalb der Pixelumgebung. Entsprechend werden dunkle Pixelwerte unterdrückt und helle Strukturen vergrößert. [NFH07]

Geometrische Bildvorverarbeitung

Algorithmen zur geometrischen Bildverarbeitung verändern die Koordinaten der einzelnen Bildpunkte, die Grauwerte der Bildpunkte bleiben jedoch unverändert. Es handelt sich somit um das Gegenstück zu photometrischen Algorithmen. Formal lassen sich solche Transformationen für einen Eingangsbildpunkt mit den Koordinaten (x_i, y_i) und dem Helligkeitswert g_i sowie den Ausgangsbildpunkt an der Stelle (x_o, y_o) mit dem Helligkeitswert g_o wie folgt beschreiben:

$$\begin{aligned}x_o &= f_x(x_i, y_i) \\y_o &= f_y(x_i, y_i) \\g_o &= g_i(x_i, y_i)\end{aligned}\tag{3.19}$$

Koordinatentransformationen kommen insbesondere zur Korrektur perspektivischer Verzerrungen und objektivbedingter Verzeichnungen zum Einsatz. Abb. 3.28 zeigt das Beispiel eines Matrix-Codes, der unter einem ungünstigen Blickwinkel aufgenommen und mittels (projektiver) Koordinatentransformation entzerrt wird. [SUW18]

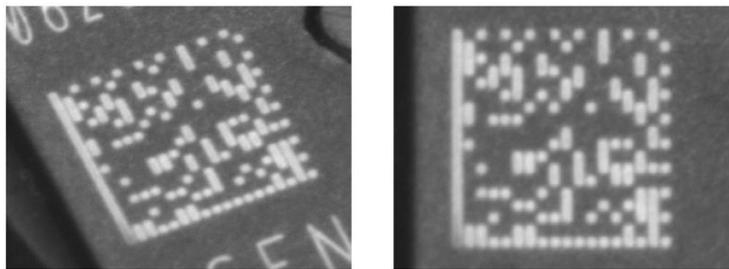


Abb. 3.28: Ausgangsbild (links), Bild nach Koordinatentransformation (rechts) [SUW18]

Innerhalb der Koordinatentransformation unterscheidet [SUW18] zwischen den folgenden beiden Transformationsarten:

- **Affine Transformationen** eignen sich zur Veränderung von Darstellungen, die lediglich zweidimensionale Freiheitsgrade haben. Es werden somit keine Tiefeninformationen für die Transformation benötigt. Eine solche Situation existiert, wenn Objekte ohne perspektivische Verzerrungen aufgenommen werden. Das zweidimensionale Bild kann dann mittels Operationen der zweidimensionalen Translation, Rotation, Skalierung und Krümmung manipuliert werden.
- **Projektive Transformationen** eignen sich zur Veränderung von zweidimensionalen Darstellungen, die jedoch dreidimensionale Freiheitsgrade haben. Ein Beispiel hierfür ist Abb. 3.28, da hier das Objekt gegenüber der Sensorebene dreidimensional im Raum geneigt ist.

Beide Transformationsarten können mit entsprechenden Matrixoperationen beschrieben und damit die neuen Bildpunktpositionen des Ausgangsbildes berechnet werden. Dabei kommt es häufig vor, dass die berechneten Bildpunktkoordinaten nicht exakt auf den diskreten Pixelkoordinaten liegen. Die Koordinatenpositionen des Ausgangsbildes müssen in diesem Fall interpoliert werden.

3.2.1.2 Bildsegmentierung

Dieses Unterkapitel beschreibt algorithmische Verfahren zur Isolierung von Objekten aus einer Bildszene. Relevante Objekte können z. B. Konturen einer Bohrung, Buchstaben oder spezifische Grauwertmuster sein. In der Praxis und Literatur existiert dazu eine Vielzahl verschiedener Methoden. Eine Auswahl der folgenden Methoden wird hier vorgestellt:

- Grauwertsegmentierung
- Konturverfolgung
- Template Matching
- Kantendetektion

Neben diesen algorithmischen Methoden können Objekte natürlich auch manuell vom Benutzer definiert werden. Dies ist sinnvoll, wenn sich zu verarbeitende Objekte im Bild immer an der gleichen Stelle befinden. Liegt etwa die Seriennummer eines Werkstückes immer im selben Bildbereich, so könnte dieser Bereich manuell als Arbeitsobjekt definiert werden. Innerhalb dieses manuell definierten Bildbereiches könnten nun weitere Segmentierungsalgorithmen angewendet werden, um die Ziffern der Seriennummer als einzelne Objekte zu segmentieren. Dem Beispiel entsprechend kann die stufenweise Segmentierung von Objekten oft einfacher implementierbar sein, als ein Objekt (in dem Fall einzelne Ziffern) direkt aus dem Gesamtbild zu isolieren. [SUW18]

Grauwertsegmentierung

Bei der Grauwertsegmentierung wird ein Objekt anhand seiner Grauwerte isoliert. Im einfachsten Fall unterscheidet sich das gesuchte Objekt gegenüber dem Hintergrund dabei durch vergleichsweise niedrigere oder höhere Grauwerte. In diesem Fall der einfachen Zweiklassen-Segmentierung spricht man von einer Binärsegmentierung. Zur Differenzierung der Objekte muss dann lediglich ein Schwellwert gefunden werden, der die Grauwertgrenze zwischen Objekt und Hintergrund definiert. Bei komplexeren Grauwertsegmentierungen (wenn z. B. ein Teil des Hintergrunds heller als das Objekt, ein anderer Teil dunkler ist) müssen entsprechend mehrere Schwellwerte definiert werden. Zur Identifizierung geeigneter Schwellwerte werden in der Praxis häufig die folgenden beiden Methoden verwendet: [DSS11]

- **Statische Schwellwerte** werden manuell vom Nutzer definiert. Liegt z. B. bei einem hellen Objekt der Grauwert eines Pixels unter dem definierten Schwellwert, so wird dieser Pixel als Hintergrund klassifiziert.
- **Histogramm-basierte Schwellwerte** können aus einer Histogrammdarstellung der Bildgrauwerte gewonnen werden, wie es in Abb. 3.29 beispielhaft dargestellt wird. [DSS11] Die hellen Pixel des Hintergrundes formen in der Histogrammdarstellung das rechte Maximum, während die dunkleren Pixel des Objekts das linke Maximum bilden. Der Schwellwert entspricht dem Grauwert des lokalen Minimums zwischen den beiden Maxima.

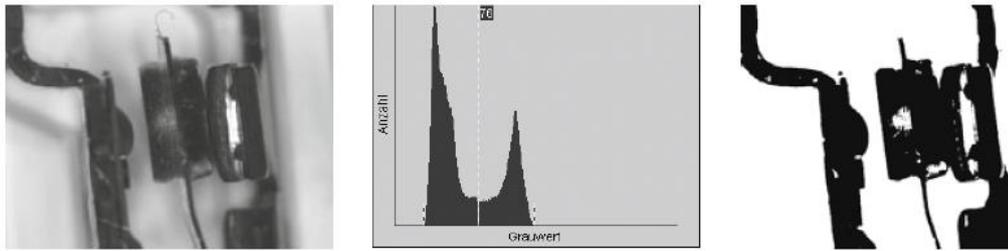


Abb. 3.29: Histogramm-basierte Schwellwertanalyse zur Bildsegmentierung [DSS11]

Konturverfolgung

Eine Objektkontur ist eine geschlossene Linie, die vollständig am Rande des Objekts verläuft. Sie besteht ausschließlich aus Randpunkten des Objekts [SUW18]. Im Gegensatz zur Grauwertsegmentierung ist das resultierende Objekt also kein Flächenelement, sondern ausschließlich dessen Kontur. Dies ermöglicht andere Methoden zur Berechnung von Objekteigenschaften (siehe Kapitel 3.2.1.3). Zum Auffinden dieser Objektkonturen existieren verschiedene Verfahren, beispielhaft sei hier der Algorithmus nach [Par94] aufgeführt. Auch hier muss das Objekt zunächst mittels Grauwertsegmentierung vom Hintergrund isoliert werden. Die Objektkontur lässt sich dann wie folgt identifizieren: [Par94]

- 1) Finde einen dem Objekt zugehörigen Bildpunkt, der einen benachbarten Bildpunkt hat, der aber nicht dem Objekt zugehörig ist.
- 2) Finde einen benachbarten Bildpunkt, der a) ebenfalls als objektzugehörig klassifiziert ist, der b) noch nicht besucht wurde und der c) ein benachbartes Pixel besitzt, das nicht dem Objekt zugehörig ist.
- 3) Wiederhole Schritt 2) solange, bis kein weiterer Bildpunkt gemäß der oben genannten Definition gefunden werden kann.

Diese Vorgehensweise entspricht lediglich der Grundform eines Algorithmus zur Konturverfolgung. Für komplexerer Objekte bedarf es einer komplizierteren Version.

Template Matching

Beim Template Matching werden Objekte im Bild anhand einer existierenden Bildvorlage identifiziert. Im Gegensatz zu den vorherigen Methoden der Objektsegmentierung muss hier also schon eine Objektbeschreibung vorliegen, nach der konkret gesucht wird. Das Bild des gesuchten Objekts (Template) wird dazu über das zu untersuchende Bild geschoben. In jeder Position wird dann ein Ähnlichkeitsmaß zwischen dem Template und dem aktuellen Bildausschnitt berechnet. Überschreitet das berechnete Ähnlichkeitsmaß einen vorgegebenen Schwellwert, so liegt das Objekt im aktuellen Bildabschnitt vor. [SUW18; DSS11] Zur konkreten Berechnung der Ähnlichkeitsmaße existiert eine Vielzahl verschiedener Methoden, die hier nicht alle im Detail vorgestellt werden können. Grundsätzlich kann dabei jedoch zwischen grauwertbasierten und formbasierten Ähnlichkeitsmaßen unterschieden werden. [PKB13; SUW18]

- **Grauwertbasierte Ähnlichkeitsmaße** berechnen die Korrelation zwischen den Grauwerten des Templates und denen des betrachteten Bildausschnittes. Dabei werden alle Bildpunkte der Template-Fläche in die Ähnlichkeitsberechnung mit einbezogen.
- **Formbasierte Ähnlichkeitsmaße** berechnen die Ähnlichkeit ausschließlich anhand der Kanteninformation. Dazu müssen mit den zuvor vorgestellten Filter- oder Segmentierungsverfahren das Template und der zu untersuchende Arbeitsbereich in eine Kantendarstellung umgewandelt werden. Das Template Matching anhand von Kanteninformationen ist grundsätzlich weniger rechenaufwendig als grauwertbasierte Verfahren. Das kantenbasierte Verfahren ist jedoch nur für Objekte anwendbar, die eindeutig anhand ihrer Kontur identifiziert werden können.

Kantendetektion

Algorithmen der Kantendetektion dienen zum gezielten Auffinden von Kanten in einem zuvor definierten Arbeitsbereich. Algorithmen der Kantendetektion ähneln in ihrem Resultat einem Kantenfilter oder einer Konturverfolgung, jedoch bedürfen sie aufgrund ihres konkreten Kantenmodells eines erheblich geringeren Rechenaufwandes. [DSS11] Die Kantendetektion erfolgt dabei auf einer im Bild definierten Suchlinie. Für jeden Bildpunkt auf der Suchlinie wird der Grauwert ausgelesen, sodass sich ein Grauwertprofil entlang der Suchlinie abbilden lässt. Ergibt sich entlang dieses Profils eine Grauwertänderung gemäß des definierten Kantenmodells, so liegt an diesem Ort eine Kante vor. Ein einfaches Kantenmodell lässt sich anhand folgender Parameter definieren, die in Abb. 3.30 visualisiert werden: [DSS11]

- **Richtung:** Es wird zwischen ansteigenden und abfallenden Kanten differenziert. Eine abfallende Kante entspricht einem Abfall des Grauwertes, also

einem Übergang von hell nach dunkel. Im umgekehrten Fall handelt es sich um eine ansteigende Kante.

- **Kantenhöhe:** Um das Kriterium für eine gültige Kante zu erfüllen, muss entlang der Suchlinie eine bestimmte Mindstdifferenz zwischen den Grauwerten vorliegen. Diese Mindstdifferenz bezeichnet man als Kantenhöhe.
- **Kantenbreite:** Die Kantenbreite definiert die Strecke entlang der Suchlinie, auf der die festgelegte Kantenhöhe überbrückt werden muss.

Im digitalen Bild können Kanten nur auf der Grenze von benachbarten Bildpunkten verlaufen. Dies entspricht in der Regel jedoch nicht dem tatsächlichen Ort der Kante in der Bildszene. Interpoliert man allerdings den diskreten Grauwertverlauf, so lassen sich Kantenpositionen mit einer Genauigkeit besser als ein Bildpunkt bestimmen. Solche Interpolationsmethoden bezeichnet man als subpixelgenaue Kantendetektion. [Ste06]

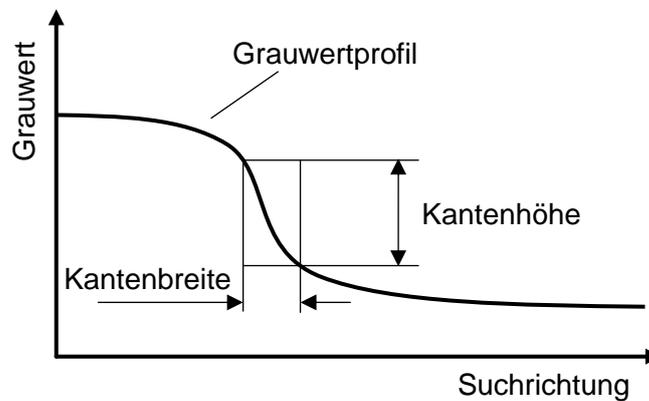


Abb. 3.30: Parameter eines Kantenmodells i. A. a. [Ste06; DSS11]

3.2.1.3 Berechnung von Objektmerkmalen

Nachdem mit den bisher vorgestellten Methoden ein Objekt zunächst isoliert wird, können in einem weiteren Schritt dessen charakteristische Merkmale berechnet werden. Diese Merkmale eines Objekts werden mittels quantitativer Werte beschrieben. Man bezeichnet diese auch als Features. Den Prozess der Berechnung von Objektmerkmalen bezeichnet man entsprechend als Feature Extraction. Die Menge aller Merkmale, die zur Beschreibung verwendet werden, werden in einem Merkmalsvektor (feature vector) zusammengefasst. In Anlehnung an die unterschiedlichen Arten isolierter Objekte lassen sich die folgenden beiden Klassen an Merkmalen für Bildobjekte definieren: [DSS11; SUW18]

- **Flächenbasierte Objektmerkmale** beschreiben die Flächeneigenschaften eines Objekts. Die primitivste Eigenschaft ist dabei die Fläche selbst, die sich aus der Aufsummierung aller dem Objekt zugehörigen Pixel ergibt. Mittels der Berechnung von Momenten höheren Grades lassen sich z. B. der Schwerpunkt einer Fläche oder dessen Hauptträgheitsachsen berechnen, womit die

translatorische und rotatorische Objektorientierung im 2D-Raum bestimmt werden kann. Eine Übersicht weiterer Größen kann der Gesamtübersicht in Anhang 11.2 entnommen werden. Grauwertstatistiken zweiter Ordnung beinhalten zusätzlich Informationen über räumliche Distanzen, sodass sie sich zur charakteristischen Beschreibung einfacher Texturen eignen. Weiterhin lassen sich wie bei den flächenbasierten Objektmerkmalen Momente definieren, die zusätzlich mit den Grauwerten der Pixel gewichtet werden.

- **Konturbasierte Objektmerkmale** charakterisieren konturförmige Merkmale. Ähnlichen den flächenbasierten Objektmerkmalen lassen sich hier Momente unterschiedlichen Grades definieren, wobei jedoch nur die Bildpunkte der Kontur in die Berechnung mit einfließen. Im Vergleich zur flächenbasierten Momentenberechnung werden ähnlich gute Ergebnisse mit geringerem Rechenaufwand erzielt. Eine zusätzliche typische Konturengröße ist der Konturenumfang. Hierzu wird die euklidische Distanz eines jeden Konturpixels berechnet und aufsummiert. Für Pixelelemente mit einem horizontalen oder vertikalen Nachbar ist die Distanz 1, Pixel mit ausschließlich in der Diagonalebene liegenden Nachbarn werden mit der Distanz $\sqrt{2}$ in die Summe eingerechnet. [SUW18]

Es sei angemerkt, dass die Berechnung von Objektmerkmalen nicht in jedem BV-Prozess angewendet werden muss. So reicht es z. B. im Rahmen einer visuellen Anwesenheitskontrolle bereits aus zu wissen, ob ein gesuchtes Objekt im Bild isoliert werden konnte oder nicht. Es bedarf hier also keiner Berechnung weiterer Eigenschaften.

3.2.1.4 Klassifikation

Bei der Klassifikation wird das Bildobjekt anhand der extrahierten Merkmale (siehe Kapitel 3.2.1.3) einer von mehreren vorgegebenen Klassen zugeordnet. Es wird also die Menge aller quantitativen Objektmerkmale genutzt, um eine finale logische Aussage über das Objekt zu treffen. [Ste06] Beispiele für Klassifikationsaufgaben sind:

- Erkennung eines Schriftzeichens in einem segmentierten Bild
- Bezeichnung eines Oberflächenfehlers anhand der Farbe und Maße der Fehlstelle
- Bezeichnung einer Schraube in einer Sortieranlage mit unterschiedlichen Schrauben anhand Länge und Breite des Schraubenkopfes

Wie sich an diesen Beispielen erkennen lässt, kann eine Klassifikation unterschiedlich komplex sein. Im einfachsten Fall lässt sich eine deterministische Aussage über ein Objekt anhand eines einzelnen Merkmals treffen. Wird etwa ein Werkstück anhand der Größe eines Bohrloches für gut oder schlecht befunden werden, so ist der Bohrlochdurchmesser das einzige relevante Merkmal, das isoliert und berechnet werden muss. Im Rahmen der Klassifizierung muss dann

lediglich geprüft werden, ob der Durchmesser innerhalb eines vorgegebenen Sollmessbereiches liegt oder nicht.

Neben solchen deterministischen Klassifikationsregeln sind in der BV insbesondere statistische Klassifikatoren für komplexere Merkmalsvektoren von Relevanz. [Ste06] Diese geben an, mit welcher Wahrscheinlichkeit ein Objekt (repräsentiert durch seinen Vektor an Merkmalen) einer Klasse zugehört. Es lässt sich zwischen zwei Arten von statistischen Klassifikatoren unterscheiden: [DSS11]

- **Multireferenzklassifikatoren** benötigen eine Anzahl von Prototypenmustern für die einzelnen Klassen. Um einen beobachteten Merkmalsvektor zu klassifizieren, wird dessen Ähnlichkeit mit den verschiedenen Prototypen berechnet und der Klasse des Prototypens mit der höchsten Ähnlichkeit zugeordnet. Multireferenzklassifikatoren sind dem Vorgehen des Template Matching im Rahmen der Objektsegmentierung sehr ähnlich.
- **Funktionalklassifikatoren** verwenden ebenso wie Multireferenzklassifikatoren klassifizierte Trainingsmuster. Funktionalklassifikatoren nutzen die Trainingsmuster jedoch nicht als Prototypen zur Ähnlichkeitsberechnung, sondern versuchen stattdessen, die Entscheidungsfunktion direkt zu approximieren. Ein geläufiges Approximationsmodell ist das Multilayer-Perzeptron, welches in Kapitel 3.2.2.1 noch ausführlicher erläutert wird.

Statistische Klassifikatoren fallen gemäß der zuvor eingeführten Definition bereits in den Bereich des maschinellen Lernens. Die Bedeutung des maschinellen Lernens für die industrielle Bildverarbeitung wird in Kapitel 3.2.2 genauer thematisiert.

3.2.2 Maschinelles Lernen und Deep Learning für die Bildverarbeitung

In Kapitel 3.2.1 werden Ablauf und Algorithmen des konventionellen Bildverarbeitungsprozesses vorgestellt. Dabei werden bereits im Bereich der Objektsegmentierung sowie der abschließenden Klassifizierung Methoden des maschinellen Lernens flüchtig aufgegriffen. Die Mehrheit der Verarbeitungsschritte beruht jedoch auf konventionellen Algorithmen.

In der jüngeren Fachliteratur gewinnen fortgeschrittene Methoden des maschinellen Lernens, die allgemein als Deep Learning bezeichnet werden, zunehmend an Bedeutung [BK19; WSS16]. Diese Methoden erlauben eine erhebliche Reduktion des manuellen Programmieraufwandes, sodass sie auch in industriellen Softwareapplikationen zunehmend Beachtung finden [WMZ+18]. Für ein besseres Verständnis dieser Methoden werden zunächst die Konzepte des Multi-Layer Perzeptrons und des Convolutional Neural Network (CNN) vorgestellt, bevor eine vergleichende Einordnung von konventionellen Algorithmen, maschinellem Lernen und Deep Learning für BV-Anwendungen gegeben wird.

3.2.2.1 Maschinelles Lernen mittels Multilayer-Perzeptron

Das Multilayer-Perzeptron ist eine populäre Methode des maschinellen Lernens zur Klassifizierung von Merkmalsvektoren (siehe Kapitel 3.2.1.3) und basiert auf dem Konzept eines künstlichen neuronalen Netzes. [DSS11; RRK89] Aufgrund des Umfangs kann auch an dieser Stelle keine vollständige mathematische Darstellungsweise erfolgen, stattdessen werden grundlegende Konzepte primär qualitativ beschrieben.

Aufbau

In Abb. 3.31 wird die grundlegende Struktur eines neuronalen Netzes in der Form eines Multilayer-Perzeptrons dargestellt. Dieses besteht aus mehreren Verarbeitungseinheiten (auch Knoten oder Neuronen genannt), welche in Schichten angeordnet sind. Die Knoten aufeinanderfolgender Schichten sind miteinander über gewichtete Kanten verknüpft. Man unterscheidet hierbei zwischen drei Arten von Schichten:

- **Eingangsschicht (Input Layer):** Diese entspricht im Fall der Bildverarbeitung dem Vektor an zuvor extrahierten Bildmerkmalen, der klassifiziert werden soll.
- **Verborgene Schicht (Hidden Layer):** Knoten in der verborgenen Schicht verarbeiten eingehende Werte gemäß der im Folgenden dargestellten Funktionsweise (siehe unten). Der berechnete Output eines Knotens wird von Knoten der darauffolgenden Schicht wiederum weiterverarbeitet.
- **Ausgangsschicht (Output Layer):** Die Anzahl der Knoten in der Ausgangsschicht entspricht der Anzahl definierter Klassen. Der in einem Output-Knoten berechnete Wert wird an keine weiteren Knoten weitergeleitet, sondern kann stattdessen in einen Wahrscheinlichkeitswert umgerechnet werden. Dieser gibt an, mit welcher Wahrscheinlichkeit der Merkmalsvektor der jeweiligen Klasse entspricht.

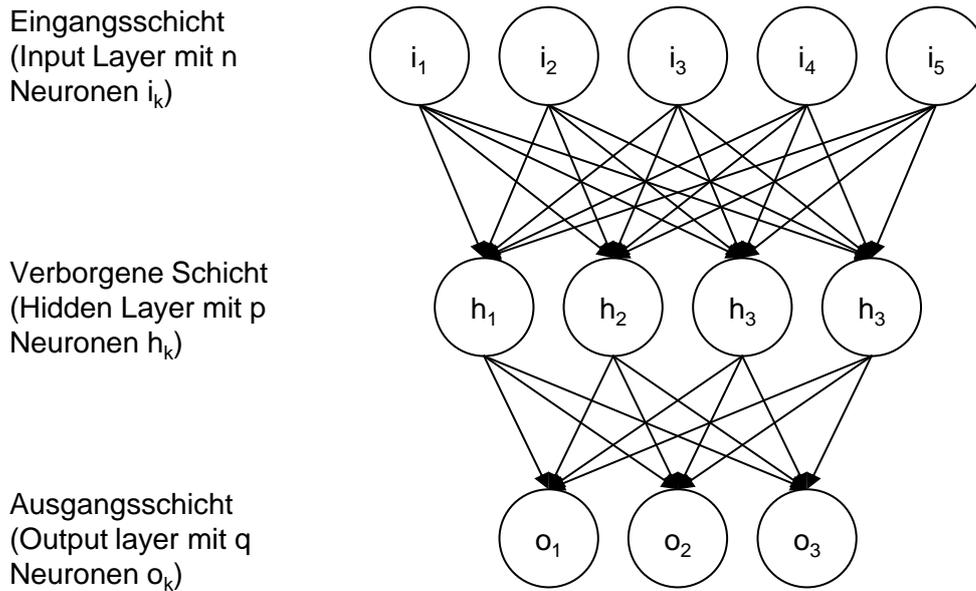


Abb. 3.31: Grundsätzlicher Aufbau eines Perzeptrons i. A. a. [Ste06]

Ein einzelner Knoten verarbeitet die eingehenden Signale nach dem folgenden Prinzip (siehe Abb. 3.32).

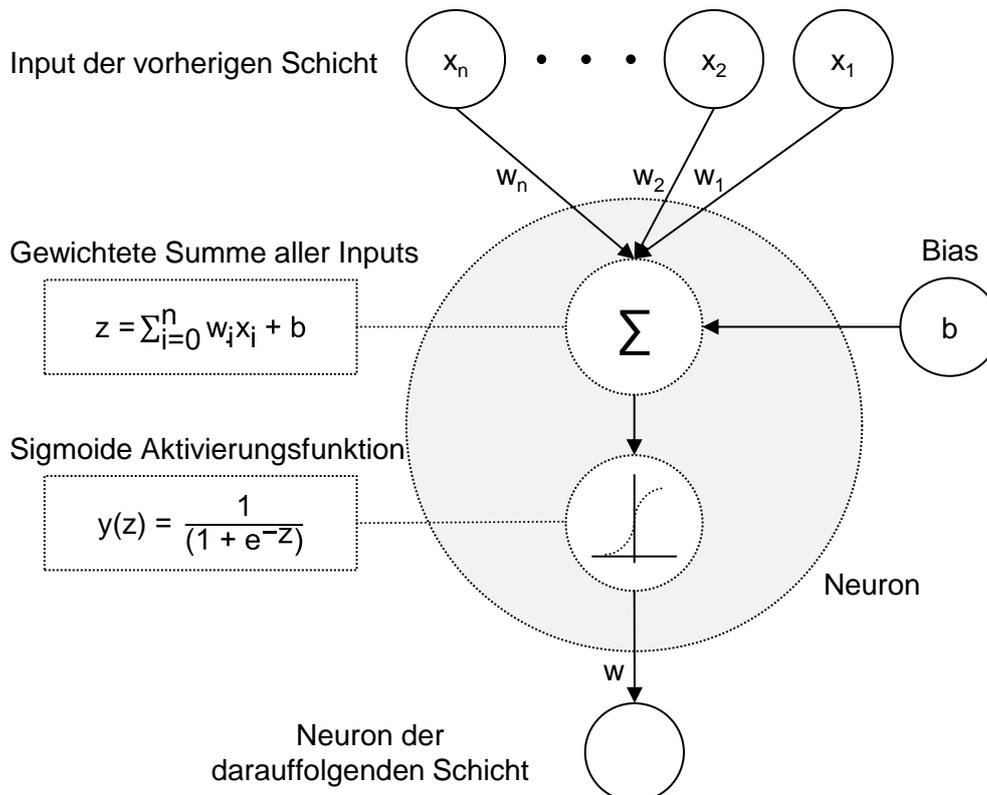


Abb. 3.32: Struktur eines künstlichen Neurons mit einer Sigmoid Aktivierungsfunktion i. A. a. [Bur19]

Die Eingangssignale x_i aus den Knoten der vorherigen Schicht werden jeweils mit dem zu trainierenden Gewicht w_{ij} multipliziert und anschließend aufsummiert. Da es für den Trainingsprozess des Netzes nichtlinearer Zusammenhänge bedarf, wird die gewichtete Summe anschließend mittels einer nichtlinearen Aktivierungsfunktion in das Ausgangssignal y überführt (für das Beispiel in Abb. 3.32 wird hierfür eine sigmoide Funktion verwendet). [Ste06; DSS11; Bur19]

Trainieren des Netzes

Um die vorgestellte Netzstruktur für Klassifizierungsaufgaben verwenden zu können, muss das Netz mit Beispieldaten trainiert werden. Ziel des Trainierens eines neuronalen Netzes ist es, geeignete Werte für alle Kantengewichte zu finden, so dass möglichst viele Merkmalsvektoren richtig klassifiziert werden. Hierzu müssen Trainingsdaten in der Paarung (x, y) vorliegen, wobei x der Merkmalsvektor und y seine Klassifizierung ist. Zum Anlernen des Netzes müssen also bereits klassifizierte Beispieldaten vorliegen. Man bezeichnet diese Art des maschinellen Lernens deswegen auch als Supervised Learning. Das eigentliche Trainieren erfolgt in den folgenden zwei Schritten:

- **Vorwärtsvermittlung (Feed-forward):** Bei der Vorwärtsvermittlung wird der Merkmalsvektor eines Trainingsdatensatzes mittels der zuvor vorgestellten Netzstruktur klassifiziert. Dabei wird zunächst von einer zufälligen Gewichtsverteilung ausgegangen, sodass die resultierende Klassifizierung zu Beginn nicht sehr akkurat ist.
- **Fehler-Rückvermittlung (Backpropagation):** Die Differenz aus mittels Vorwärtsvermittlung berechneter Klassifizierung und tatsächlicher Klassifizierung bezeichnet man als Klassifizierungsfehler. Der Klassifizierungsfehler muss folglich über die Summe aller Trainingsdatensätze minimiert werden. Dies geschieht mit dem sogenannten Backpropagation-Verfahren, bei dem die einzelnen Kantengewichte in Richtung abnehmenden Fehlers verändert werden.

Nachdem der komplette Datensatz diese beiden Schritte einmalig durchlaufen hat, spricht man von dem Abschluss einer Epoche. Dieser Vorgang wird mehrfach wiederholt, sodass die Kantengewichte über mehrere Epochen iterativ angepasst werden. Für eine detaillierte mathematische Ausführung sowie weitere Besonderheiten, die während des Trainingsprozesses zu beachten sind, sei auf [Ert13; WFH+17] verwiesen.

3.2.2.2 Deep Learning mittels CNN

Deep Learning bezeichnet eine Klasse künstlicher neuronaler Netze, die mehrere verborgene Schichten zwischen Eingabeschicht und Ausgabeschicht haben und dadurch eine umfangreiche innere Struktur aufweisen [GBC16]. Das Convolutional

Neural Network (CNN) ist eine solche Deep Learning Struktur, die in der maschinellen BV sehr populär ist [SLF18; SSP+18].

Bei einem CNN handelt es sich um ein neuronales Netz (siehe vorheriges Kapitel), dem zusätzlich sogenannte Convolutional Layers vorgeschaltet sind. Die Struktur eines CNN ist in Abb. 3.33 visualisiert [LGT+97].

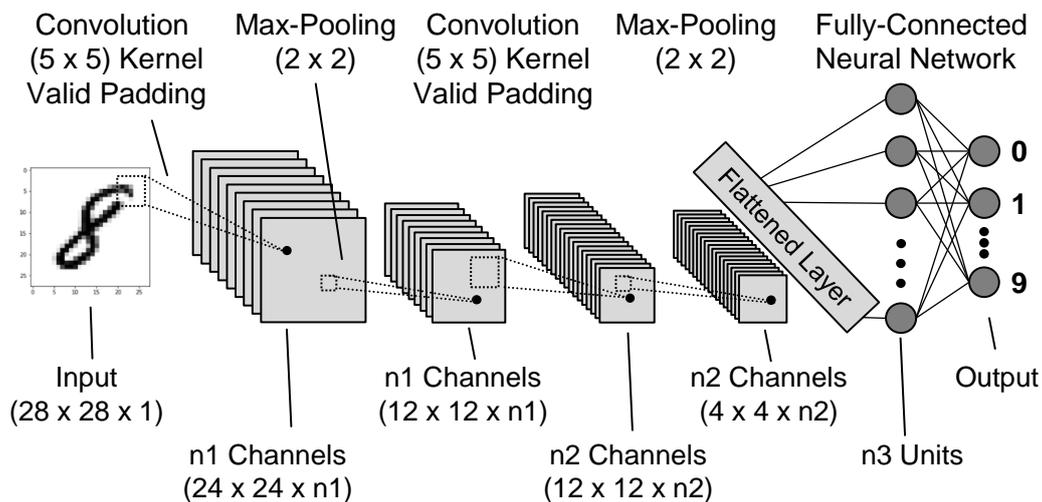


Abb. 3.33: Struktur eines Convolutional Neural Network i. A. a. [LGT+97]

Convolutional Layers basieren auf dem Prinzip der diskreten Faltung. Dabei wird ein Kernel bzw. Filter definiert, der schrittweise Ausschnitte des zu untersuchenden Bildes auswertet. Je nach Parametrisierung des Kernels werden unterschiedliche Eigenschaften des Bildes, z. B. Kanten, extrahiert und in einer sogenannten Feature Map gespeichert. Diese Feature Map kann wiederum von einer weiteren Convolutional Layer gefiltert werden, sodass Muster auf höheren Metaebenen identifiziert werden können. Zwischen mehreren Convolutional Layers werden in der Praxis häufig sogenannte Pooling-Layers verwendet, welche Feature Maps komprimieren. [CPC17; LGT+97]

Die Vorgehensweise der Convolution ist den in Kapitel 3.2.1.1 vorgestellten Filtermechanismen grundsätzlich sehr ähnlich. Jedoch werden beim CNN die Filterfunktionen nicht explizit ausgewählt. Stattdessen wird eine Vielzahl verschiedener Filter auf das Ausgangsbild angewendet, was zu einer deutlich höheren Informationskomplexität als bei manuell definierten Filtern führt. Nicht alle daraus resultierenden Feature Maps beinhalten für die Klassifizierung auch relevante Informationen. Zur Bewältigung dieser Informationsfülle dient nun das darauffolgende Neuronale Netz. Dieses weist im Laufe des Trainingsprozesses relevanten Features höhere Kantengewichte zu, wohingegen die Kantengewichte von unwichtigen Features reduziert werden [CPC17; LGT+97].

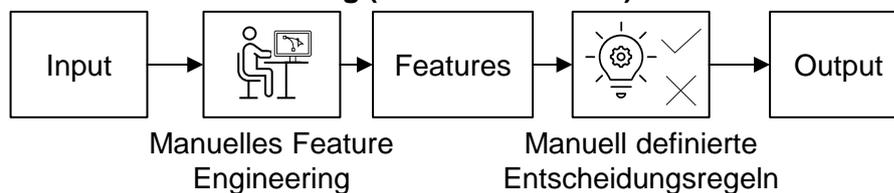
Das vorgestellte CNN bildet die Grundlage für viele Deep Learning Klassifizierungsmethoden im BV-Bereich. Darauf basierend wurden im Laufe der Zeit komplexere Architekturen entwickelt, mit denen bessere Vorhersagegenauigkeiten erreicht werden können. Anhang 11.3 gibt eine Übersicht erfolgreicher Architekturen der ImageNet Challenge (Wettbewerb zur Bildklassifizierung mit 1000 verschiedenen Klassen) [CPC17].

Neben der Klassifizierung von Bildern, also der Einordnung eines Bildes in eine zuvor definierte Klasse, eignen sich CNN-basierte Deep Learning Methoden auch zur Objektdetektion. Hierbei wird eine abgewandelte CNN-Struktur verwendet, um einzelne Objekte innerhalb eines Bildes zu finden und deren Position zu markieren. [ZZX+19] Die Objektdetektion ähnelt in ihrer Funktion daher der manuellen Arbeitsbereichsdefinition und Objektsegmentierung des konventionellen Bildverarbeitungsablaufs. Formal betrachtet ist die CNN-basierte Objektdetektion allerdings auch eine Klassifizierung. Dabei wird nicht das Bild als Ganzes klassifiziert, sondern verschiedene algorithmisch erzeugte Bildsubbereiche.

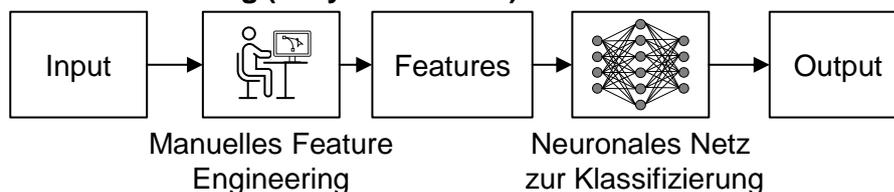
3.2.2.3 Abgrenzung unterschiedlicher Bildverarbeitungsvorgehen

Zur besseren Abgrenzung von Deep Learning und konventionellen Bildverarbeitungsalgorithmen dient der zusammenfassende Überblick der unterschiedlichen Vorgehensweisen (siehe Abb. 3.34).

Konventionelle Bildverarbeitung (White Box Modell)



Hybride Bildverarbeitung (Grey Box Modell)



Deep Learning (Black Box Modell)

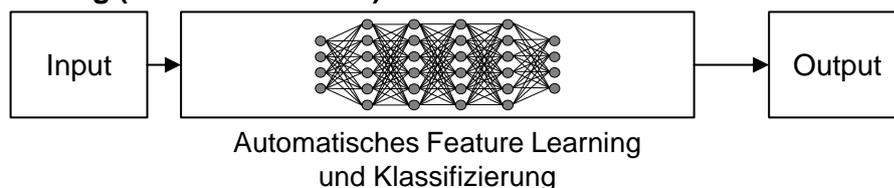


Abb. 3.34: Unterschied zwischen konventioneller Bildverarbeitung und Deep Learning i. A. a. [WMZ+18]

Ziel aller kamerabasierten Qualitätsinspektionen ist eine Aussage über das / die im Bild befindliche(n) Objekt(e). Um diese Aussage treffen zu können, müssen aussagerelevante Merkmale aus dem Bild extrahiert werden. Bei konventionellen Bildverarbeitungsalgorithmen erfolgt dieser Prozess schrittweise und auf Basis menschlicher Expertise. [WMZ+18] Beim reinen Deep Learning bedarf es keiner manuell programmierten Merkmalsextraktion mehr. Stattdessen werden relevante Merkmale über die Anpassung der Gewichte automatisch aus dem Eingangsbild extrahiert. Zwischen reinem Deep Learning und konventioneller BV existieren weitere hybride Ansätze, die sich bzgl. ihres Automatisierungsgrades unterscheiden lassen. Bei einem hybriden Ansatz geringer Automation werden lediglich einzelne Schritte des BV-Prozesses durch einfache Methoden des maschinellen Lernens ersetzt, etwa im Rahmen einer finalen Musterklassifizierung mittels Multilayer-Perzeptron. Zur Merkmalsextraktion bedarf es dabei weiterhin konventionell programmierter Algorithmen. Ein hybrider Ansatz mit hohem Automatisierungsgrad hingegen wäre z. B. die Verwendung eines vollständigen CNNs, dem lediglich ein manueller Bildvorverarbeitungsalgorithmus zur Grauwerttransformation vorgeschaltet wird. Die Vor- und Nachteile dieser unterschiedlichen Vorgehensweisen werden in Kapitel 6.2 im Rahmen der Entscheidungslogik detaillierter aufgegriffen.

4 Marktrecherche von Bildverarbeitungstechnologien

Die im vorherigen Kapitel theoretisch vermittelten technischen Grundlagen und Charakterisierungsparameter werden nun mittels einer Marktrecherche zu existierenden Produkten um eine praxisorientierte Perspektive erweitert. Die daraus gewonnenen Kenntnisse über tatsächlich existierende Produkte (Stand Februar 2020) fließen in die Entscheidungslogik zur Technologieauswahl mit ein und erhöhen deren Praxistauglichkeit. Wie in Kapitel 3 wird auch hier zwischen Hardware- und Softwaretechnologien unterschieden.

4.1 Hardware

Die Marktrecherche zu Hardwareprodukten thematisiert zwei verschiedene Fragestellungen. In Kapitel 4.1.1 wird zunächst dargestellt, welche der in Kapitel 3.1 vorgestellten technischen Beschreibungsgrößen in der Praxis vom Hardwareanbieter zur Verfügung gestellt werden. In den darauffolgenden drei Unterkapiteln wird eine Übersicht von Kameras, Objektiven und Beleuchtungsmitteln gegeben.

4.1.1 Verfügbarkeit von Produktinformationen

Die in Kapitel 3 vorgestellten Größen zur technischen Charakterisierung von Kameras, Objektiven und Beleuchtungssystemen werden anhand der wissenschaftlichen Fachliteratur zusammengestellt. Sie sind somit zunächst theoretischer Natur. In der Praxis werden jedoch nicht alle dieser Angaben von Herstellern zur Verfügung gestellt. Die im Rahmen der Marktrecherche wahrgenommene Verfügbarkeit von Produktinformationen ist in Tab. 4.1 dargestellt. Dabei wird zwischen drei Stufen der Informationstransparenz differenziert:

- 1) **Hohe Informationstransparenz** ist gegeben, wenn die Beschreibungsgröße in vergleichenden Produktübersichten oder etwa in der Eingabemaske einer Produktdatenbank aufgeführt wird. Der Nutzer kann folglich sehr einfach verschiedene Produkte bzgl. dieser Beschreibungsgröße unterscheiden.
- 2) **Mittlere Informationstransparenz** ist gegeben, wenn die Beschreibungsgröße mit geringem zusätzlichem Aufwand eingesehen werden kann. Diese Informationsstufe entspricht Beschreibungsgrößen, die vom Hersteller in einem Produktdatenblatt zur Verfügung gestellt werden.
- 3) **Geringe Informationstransparenz** liegt vor, wenn eine Beschreibungsgröße nicht im technischen Produktdatenblatt aufgeführt wird.

Für die Marktrecherche und folgende Logik zur Technologieauswahl werden primär Beschreibungsparameter verwendet, die eine hohe oder mittlere Informations-transparenz haben.

Tab. 4.1: Verfügbarkeit von Produktinformationen (Eigendarstellung)

	Beschreibungsgröße	Transparenz		Beschreibungsgröße	Transparenz
Kameraparameter	Auflösung	Hoch	Objektivparameter	Brennweite	Hoch
	Pixelgröße und Füllfaktor	Hoch		Vergrößerung	Hoch
	Optisches Format	Hoch		Blendenzahl	Hoch
	Bittiefe	Mittel		Schärfentiefe	Gering
	Bildrate	Hoch		Minimale Objektdistanz	Hoch
	Quantenausbeute / Spektrale Empfindlichkeit	Mittel		Auflösungsvermögen	Hoch
	Dunkelstrom	Gering		Homogenität der Ausleuchtung	Gering
	Signal / Rausch-Verhältnis	Mittel		Verzeichnung	Mittel
	Sättigungskapazität	Mittel		Objektivgröße / Lichtkreisdurchmesser	Hoch
	Empfindlichkeitsschwelle	Mittel		Kameraanschluss	Hoch
	Dynamikumfang	Mittel	Beleuchtungsparameter	Beleuchtungswinkel	Hoch
	Linearität	Mittel		Wellenlänge	Hoch
	DSNU und PRNU	Gering		Lichtausbreitung	Mittel
	Übertragungsschnittstelle	Hoch		Lichtintensität	Mittel
	Schutzgrad	Mittel		Homogenität	Mittel
	Sensortyp	Hoch		Beleuchtungsfläche	Gering
Shuttermodus	Hoch	Beleuchtungssteuerung		Mittel	
		Lichtmanipulation		Mittel	

4.1.2 Produktübersicht Kameras

Für die Durchführung der Marktanalyse im Bereich Hardware werden primär Daten der Stemmer Imaging AG, Puchheim, verwendet. Stemmer Imaging ist ein internationaler Anbieter von Bildverarbeitungstechnologien und bietet Kunden individuell zugeschnittene Lösungen für industrielle Bildverarbeitung. [Ste20] Auf ihrem Internetauftritt bietet Stemmer Imaging Informationen über eine Vielzahl an Produkten im Bereich Kameras, Optiken und Beleuchtungen von ausgewählten Herstellern. Für die Marktanalyse werden diese Informationen zunächst mittels eines Web-Crawlers (Webscraper.io als Google Chrome Erweiterung) extrahiert, bereinigt und anschließend ausgewertet [Web20].

4.1.2.1 Angebotsspektrum Industriekameras

Für die Analyse wird ein um unvollständige Daten reduzierter Datensatz aus 708 Industriekameras mit Matrixsensoren von 8 verschiedenen Herstellern verwendet. Diese lassen sich zunächst nach ihrer spektralen Empfindlichkeit unterscheiden. 338 Modelle nehmen Bilddaten im menschlich sichtbaren Spektralbereich auf und geben diese monochromatisch wieder, weitere 312 Modelle geben diese als Farbbilder wieder. 58 Modelle arbeiten in unterschiedlichen Infrarotbereichen. Eine Übersicht verschiedener Kameraparameter ist Abb. 4.1 zu entnehmen. Da vereinzelte Werte im oberen Leistungsspektrum die Achsendarstellung in Abb. 4.1 stark verzerren würden, sind diese dort nicht abgebildet.

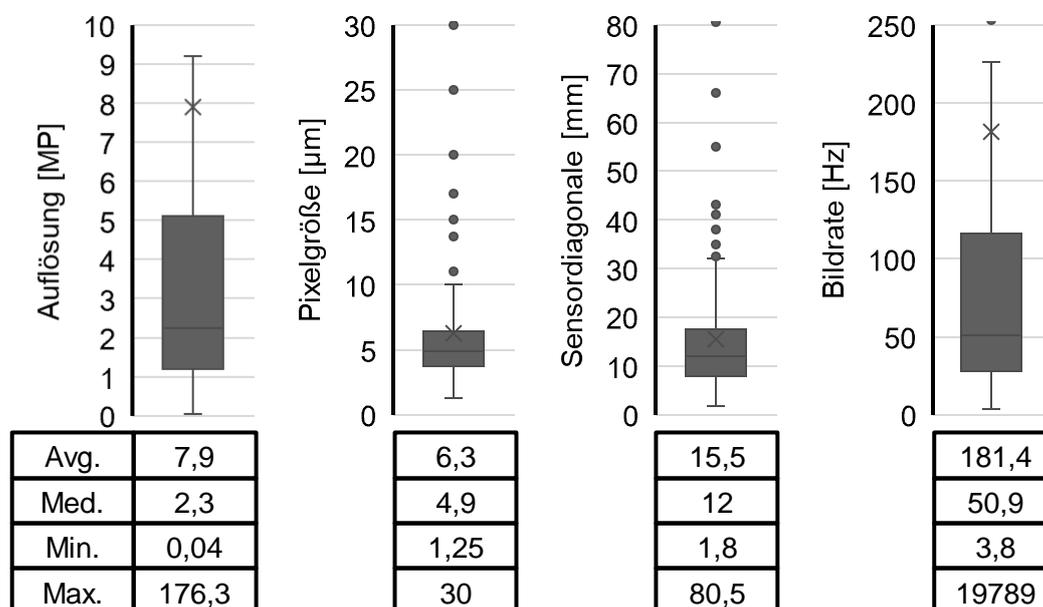


Abb. 4.1: Statistische Verteilung von Auflösung, Pixelgröße, Sensordiagonale und Bildrate von Kameras (Eigendarstellung)

Bereits bei der Betrachtung von Auflösung, Pixelgröße und Bildrate lässt sich schnell feststellen, dass ein sehr breit gefächertes Angebot an Produkten existiert. Während im Konsumgüterbereich der Nutzer oft eine hohe Flexibilität bzgl. der Einsatzmöglichkeiten einer Kamera fordert, muss bei industriellen Qualitätskontrollen eine Kamera speziell für einen ausgewählten Anwendungsfall geeignet sein [Ste18]. Es bedarf somit keiner unnötigen Überdimensionierung, sondern einer Lösung, die für den gegebenen Anwendungsfall gerade so gut wie nötig ist. In diesem Sinne hat eine durchschnittliche Industriekamera lediglich eine Auflösung von 2,3 MP und eine Bildrate von 51 Hz. Der durchschnittliche Kamerapreis beträgt 4242 €, der Preismedian liegt bei 1800 €. Eine Einordnung nach Auflösung, Bildfrequenz, Pixelgröße (siehe relative Blasengröße) und Preis ist beispielhaft für alle monochromatischen Kameras in Abb. 4.2 dargestellt.

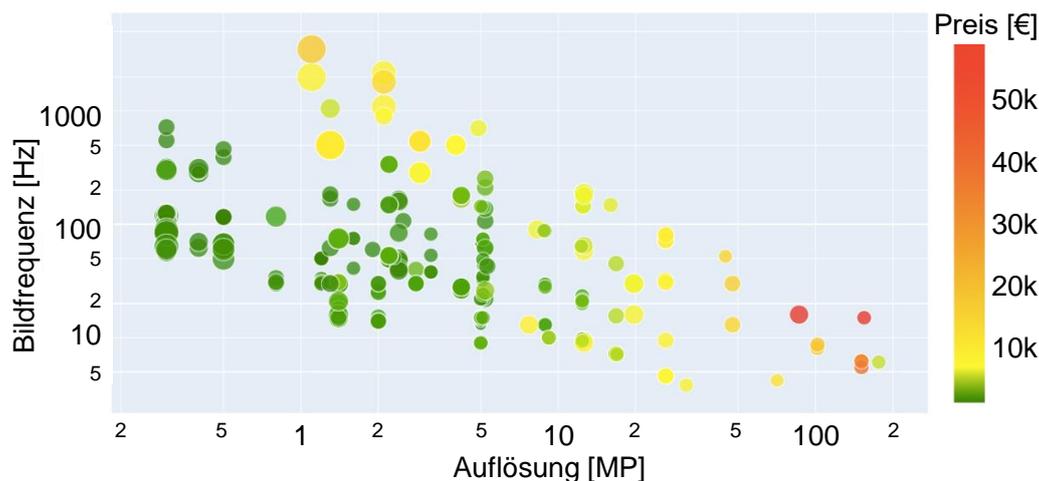


Abb. 4.2: Übersicht monochromatischer Kameras (Eigendarstellung)

4.1.2.2 Sonstige auswahlrelevante Produkteigenschaften: Intelligente Kameras

An dieser Stelle wird kurz auf die Bedeutung von Kameras mit eingebetteten Verarbeitungssystemen, sogenannten intelligenten Kameras (Smart Cameras), eingegangen. Während herkömmliche Kameras die über den Sensor aufgenommenen Bilddaten zur Auswertung an eine externe Rechereinheit weiterleiten, verfügen intelligente Kameras über einen eingebauten Computer zur BV. Nach der lokalen Bildauswertung kann das modifizierte Bild oder lediglich eine extrahierte Bildinformation (z. B. Qualitätsurteil) an andere Systeme weitergeleitet werden. [HCG+06] Für den Anwender reduzieren sich dadurch die technologische Systemkomplexität sowie der physische Platzbedarf. Bei der Auswahl einer geeigneten intelligenten Kamera ist insbesondere die Kompatibilität mit geeigneten Softwareprodukten zu beachten, da einige Hersteller wie z. B. die Cognex Corporation, Natick, ihre Hardware lediglich in Kombination mit eigenen Softwaremodulen anbieten. Neben der Kamera und zugehöriger Bildverarbeitungseinheit bietet Cognex zudem integrierte Optik- und Beleuchtungssysteme an, sodass hier ein vollständig aufeinander abgestimmtes Hard- und Softwarepaket aus einer Hand angeboten wird.

4.1.3 Produktübersicht Objektive

Bei Objektiven für die industrielle BV lässt sich zwischen den folgenden vier Kategorien unterscheiden [Ste18]:

- **Festbrennweitenobjektive** werden bei Anwendungen mit einem festen Arbeitsabstand bevorzugt eingesetzt. Sie sind der gebräuchlichste Objektivtyp für industrielle Anwendungen. Sie besitzen eine feste Brennweite und einen

entozentrischen Strahlengang. Makroobjektive stellen eine Unterklasse der Festbrennweitenobjektive dar, die auf die Fokussierung bei Nahaufnahmen optimiert sind.

- **Telezentrische Objektive** werden für Applikationen eingesetzt, bei denen perspektivische Verzerrungen oder Verzeichnungen unerwünscht sind (siehe Abb. 4.3). Bei telezentrischen Objektiven werden perspektivische Verzerrungen durch die Verwendung von ausschließlich achsparallelen Strahlen und einer komplexen Optik vermieden. Gleich große Objekte mit unterschiedlichem Abstand zum Objektiv erscheinen durch eine telezentrische Optik betrachtet dadurch immer noch gleich groß. Telezentrische Objektive kommen daher insbesondere bei präzisen Messaufgaben von dreidimensionalen Objekten zum Einsatz.
- **Zoomobjektive** verfügen über mehrere Linsen, deren relative Position zueinander verstellt werden kann. Dadurch lassen sich unterschiedliche Brennweiten realisieren. Zoomobjektive eignen sich für den Einsatz in Applikationen mit dynamischen Abbildungsverhältnissen. Für Applikationen mit hoher Wiederholgenauigkeit oder präzise Messaufgaben werden jedoch Festbrennweitenobjektive bevorzugt.
- **Spezialobjektive** sind Optiken, die in ausgewählten Sonderfällen zur Anwendung kommen. Aufgrund des geringen Marktanteils existieren hierzu vergleichsmäßig wenige Produkte, sodass diese hier nicht detaillierter analysiert werden. Zu den Spezialoptiken gehören z. B. Fish-Eye Objektive, welche u. a. zur Aufnahme von Bildern aus dem Inneren zylindrischer Körper geeignet sind.

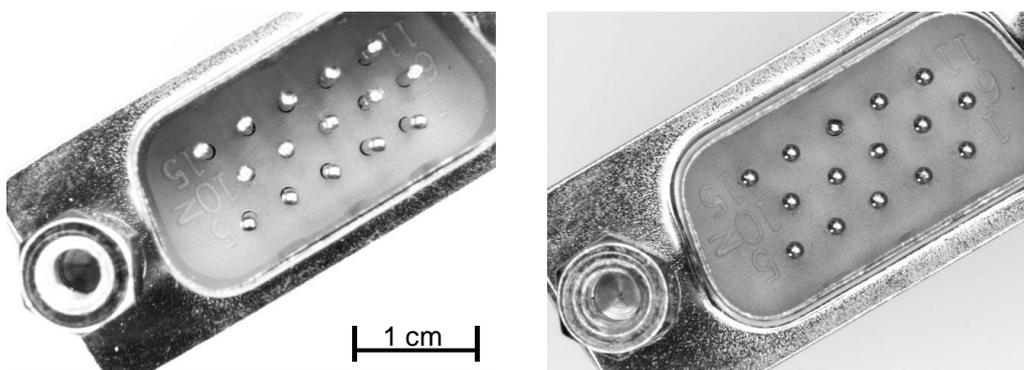


Abb. 4.3: Pinbaugruppe betrachtet durch ein Festbrennweitenobjektiv (links) und ein telezentrisches Objektiv (rechts) [Fer19]

In der industriellen BV werden primär Festbrennweitenobjektive und telezentrische Objektive verwendet [Ste18]. Diese beiden Produktkategorien werden daher im Folgenden näher vorgestellt.

4.1.3.1 Festbrennweitenobjektive

Im Rahmen der Marktanalyse werden Daten von 316 Festbrennweitenobjektive von 10 verschiedenen Herstellern erfasst. 250 dieser Objektive sind für die Aufnahme von Farbbildern geeignet, während 21 Objektive ausschließlich für monochromatische Bildaufnahmen verwendet werden können. Monochromatische Aufnahmen lassen sich jedoch auch mit Objektiven für Farbkameras machen. Weitere 45 Festbrennweitenobjektive sind für Spezialanwendungen mit Infrarot- oder UV-Licht geeignet. Die Verteilung der Merkmalsausprägungen Brennweite, minimale Pixelgröße, Bilddiagonale (relative Blasengröße) und Listenpreisen können Abb. 4.4 entnommen werden. Dabei ist erkennbar, dass Objektive mit einem höheren Auflösungsvermögen (geringe minimale Pixelgröße und große Bilddiagonale) tendenziell teurer sind.

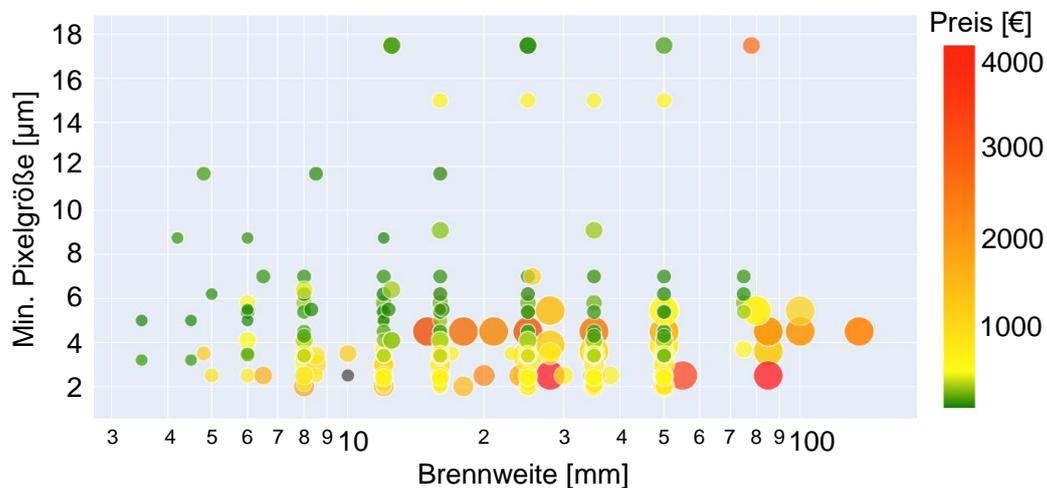


Abb. 4.4: Übersicht Festbrennweitenobjektive (Eigendarstellung)

4.1.3.2 Telezentrische Objektive

Obwohl telezentrische Objektive lediglich für genaue Vermessungsaufgaben benötigt werden und somit einen kleineren Marktanteil als Festbrennweitenobjektive haben [Ste18], wird im Rahmen der Marktanalyse eine höhere Anzahl an telezentrischen Produkten identifiziert (471 telezentrische Objektive vs. 316 Festbrennweitenobjektive). Dies ist der speziellen Optik von telezentrischen Objektiven geschuldet, die gegenüber konventionellen Objektiven noch stärker auf den spezifischen Anwendungsfall zugeschnitten sein muss. Im Vergleich zu Festbrennweitenobjektiven werden zusätzlich noch folgende Spezifikationsparameter verwendet. [Ste18]

- **Objektivdurchmesser:** Aufgrund des parallelen Strahlenganges muss der Objektivdurchmesser zur Objektseite hin mindestens so groß wie das betrachtete Objekt sein. Da je nach Anwendungsfall der benötigte Sichtfeldbereich

unterschiedlich groß ist, existiert entsprechend eine Vielzahl an verschiedenen Durchmesser Ausführungen. [Ste18]

- **Messabstand:** Aufgrund des parallelen Strahlenganges lässt sich unabhängig von der Objektgröße ein optimaler Arbeitsabstand definieren, auf den ein Objektiv ausgelegt ist. [Sil18]
- **Vergrößerungsverhältnis:** Das Vergrößerungsverhältnis wird bei telezentrischen Systemen anstelle der Brennweite zur Beschreibung des Abbildungsverhältnisses verwendet.
- **Bi-Telezentrie:** Einfache telezentrische Systeme weisen lediglich auf der Objektseite einen parallelen Strahlengang auf. Bi-telezentrische Systeme haben stattdessen auch auf der Bildseite einen parallelen Strahlengang. Dadurch wird eine noch höhere Vergrößerungskonstanz und Schärfentiefe gegenüber einfachen telezentrischen Systemen erreicht.

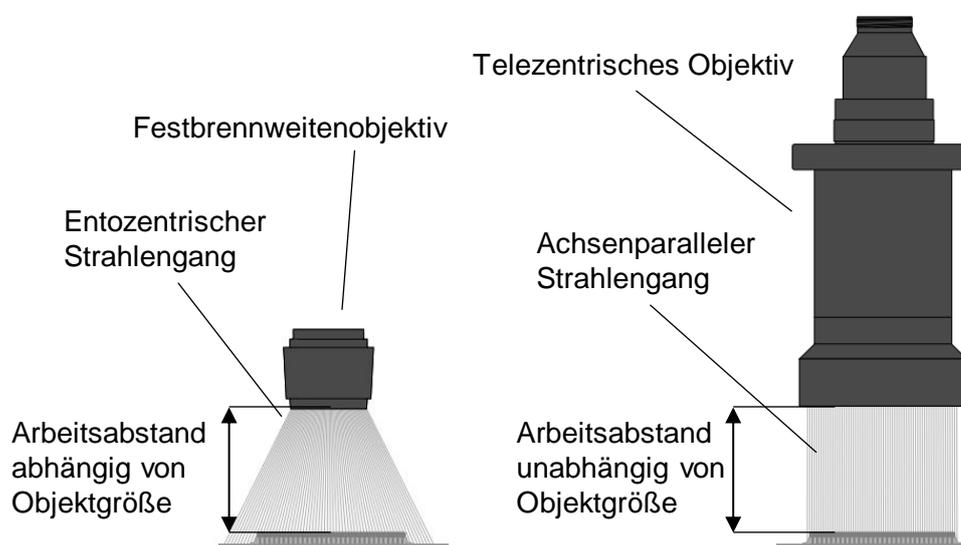


Abb. 4.5: Festbrennweiteobjektiv und telezentrisches Objektiv i. A. a. [Ste18]

Eine Übersicht der Ausprägungen telezentrischer Objektive bezüglich Messabstand, Objektivdurchmesser, Auflösungsvermögen (relative Blasengröße) und Listenpreis ist Abb. 4.6 zu entnehmen. Insgesamt werden dabei 471 telezentrische Objektive betrachtet, von denen 352 bi-telezentrisch sind. Mit einem durchschnittlichen Preis von 2690 € sind telezentrische Objektive deutlich teurer als Festbrennweitenobjektive (Durchschnittspreis von 680 €).

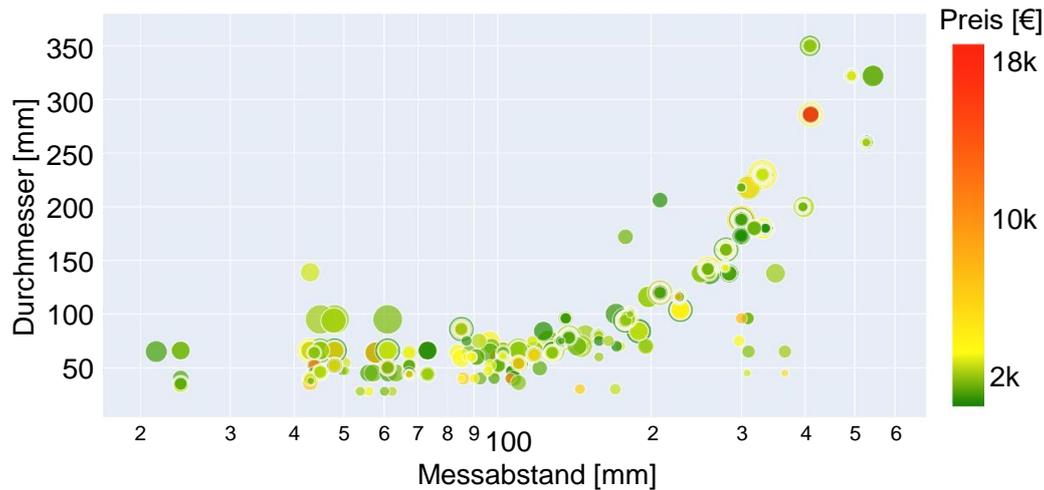


Abb. 4.6: Übersicht telezentrische Objektive (Eigendarstellung)

4.1.4 Produktübersicht Beleuchtungssysteme

Wie in Kapitel 4.1 bereits erwähnt, weisen Herstellerangaben von Beleuchtungssystemen im Vergleich zu Kameras und Objektiven wenig standardisierte Beschreibungsgrößen auf. Aus diesem Grund wird hier kein quantitativer Vergleich von Produkten durchgeführt. Eine generische Übersicht industrieller Beleuchtungsprodukte ist Abb. 6.8 in Kapitel 6.1.5 zu entnehmen.

4.2 Software

Im Rahmen der Software-fokussierten Marktanalyse werden unterschiedliche Möglichkeiten zur Programmierung von Inspektionsanwendungen herausgearbeitet. Neben der Differenzierung zwischen konventionellen und Deep Learning basierten Algorithmen (siehe zweite wissenschaftliche Fragestellung) wird zudem zwischen frei verfügbarer Software (Open Source) und kommerzieller Software unterschieden.

4.2.1 Open Source Software

Open Source Software wird im Kontext dieser Arbeit als frei verfügbare BV-Software verstanden, für deren Verwendung keine initialen oder dauerhaften Nutzungsgebühren anfallen.

In der Praxis handelt es sich dabei um frei verfügbare Code-Bibliotheken, die dem Nutzer Zugriff auf vordefinierte Bildverarbeitungsfunktionen bieten. Diese Bibliotheken sind nicht speziell auf die Entwicklung von Bildverarbeitungsapplikationen für industrielle Qualitätskontrollen ausgelegt, sondern eignen sich zur

selbstständigen Applikationsentwicklung in unterschiedlichen Kontexten. Zur Implementierung von BV-Anwendungen mittels Open Source-Bibliotheken sind folglich entsprechende Programmierkenntnisse notwendig.

Unterschiedliche Code-Bibliotheken sind in den gängigen objektorientierten Programmiersprachen verfügbar. Dazu zählen insbesondere C / C++, Java und Python. C / C++ sind kompilierte Programmiersprachen. Das bedeutet, dass in C / C++ geschriebene Programme vor der Anwendung einmalig aufgebaut bzw. kompiliert werden müssen. Bei der eigentlichen Anwendung muss das kompilierte Programm dann nur geladen und ausgeführt werden. Java und Python sind hingegen interpretierte Programmiersprachen. Diese werden nicht einmalig im Voraus kompiliert, sondern müssen bei jeder Anwendung abschnittsweise interpretiert und ausgeführt werden. Aus diesem Grund haben kompilierte Programmiersprachen eine höhere Laufzeiteffizienz als interpretierte Sprachen. [Daw14; Cas14]

Weiterhin lässt sich zwischen Bibliotheken mit konventionellen BV-Algorithmen sowie speziellen Bibliotheken zur Erstellung von ML- / DL-Applikationen unterscheiden. In den beiden folgenden Unterkapiteln wird eine kurze Übersicht entsprechender Bibliotheken gegeben.

4.2.1.1 Open Source Bibliotheken für konventionelle Bildverarbeitung

Im Rahmen einer Internetrecherche lässt sich eine Vielzahl verschiedener Bibliotheken für konventionelle Bildverarbeitung identifizieren, die nicht nur im industriellen Kontext Anwendung finden. Die Verbreitung von Bibliotheken wird dabei anhand der Anzahl an Nennungen in Onlineforen approximiert. Einige der populärsten Bibliotheken seien an dieser Stelle kurz vorgestellt:

- **OpenCV** kann mit C / C++, Python und Java verwendet werden. Die umfangreiche Bibliothek ist aufgrund ihrer Laufzeiteffizienz für Echtzeitapplikationen sehr gut geeignet und wird für akademische und kommerzielle Anwendungen häufig verwendet. [PBK+12]
- **NumPy** ist eine Python Bibliothek und ebenso wie OpenCV im akademischen und industriellen Umfeld weit verbreitet. NumPy verfügt über eine einfachere Programmiersyntax als OpenCV, was jedoch auch in einer geringeren Laufzeiteffizienz resultiert. [VCV11]
- **ImageJ** ist eine in Java geschriebene Bibliothek, die mit mehr als 400 verfügbaren Erweiterungen zu einem umfangreichen Bildverarbeitungsprogramm erweitert werden kann. ImageJ verfügt über eine geringere Laufzeiteffizienz als OpenCV und wird insbesondere für medizinische und wissenschaftliche Bildanalysen verwendet. [SRE12; AMR04]

4.2.1.2 Open Source Bibliotheken für Machine Learning Applikationen

Ähnlich der in Kapitel 4.2.1.1 vorgestellten Methodik lassen sich Open Source Bibliotheken für die Entwicklung von Machine Learning Applikationen recherchieren. Drei Beispiele für populäre Bibliotheken sind hier aufgeführt: [Klo19; War17]

- **TensorFlow** ist eine 2015 von Google LLC, Menlo Park, veröffentlichte Bibliothek. Sie ist in verschiedenen Programmiersprachen verfügbar, basiert jedoch auf einer C++ Infrastruktur. Diese gibt Entwicklern ein hohes Maß an Freiheitsgraden.
- **Keras** ist eine in und für Python geschriebene Bibliothek und baut auf TensorFlow auf. Keras hat aufgrund seiner einfachen Handhabung eine sehr große Nutzercommunity, bietet im Vergleich zu TensorFlow aber auch weniger Freiheitsgrade in der Programmierung.
- **Caffe** ist eine in C++ geschriebene und für Python, Matlab und C++ verfügbare Bibliothek. Durch die Einbindung in Apache Spark kann Caffe über mehrere Rechner verteilt verwendet werden, was zur erheblichen Steigerung der Verarbeitungsgeschwindigkeit beim Modelltraining führt.

4.2.2 Kommerzielle Software

Im Gegensatz zur Open Source Software ist eine Vielzahl der kommerziellen Produkte speziell auf Applikationen im industriellen BV-Kontext zugeschnitten. Die Produkte werden in dieser Arbeit dabei hinsichtlich der folgenden Kriterien kategorisiert:

- **Benutzerschnittstelle:** Hierbei wird zwischen Anwendungen mit einer grafischen Benutzeroberfläche (GUI) zur Erstellung von Applikationen und Code-Bibliotheken unterschieden. Für die Erstellung von Applikationen mittels GUI bedarf es keiner besonderen Programmierkenntnisse.
- **Zugrundeliegende Algorithmen:** Hierbei wird gemäß wissenschaftlicher Fragestellung zwischen Softwareprodukten unterschieden, die primär entweder auf Deep Learning oder auf konventionellen Algorithmen basieren. Während auf konventionellen Algorithmen basierende Produkte schon lange am Markt sind, wird inzwischen auch eine Vielzahl von DL-Produkten angeboten.
- **Spezialisierungsgrad:** Dieser gibt an, ob ein Produkt speziell für die industrielle Qualitätskontrolle oder für einen breiteren Anwendungskontext gedacht ist. Die Mehrheit der hier vorgestellten Produkte ist auf die industrielle Qualitätskontrolle ausgelegt. In der Unterkategorie der Deep Learning Produkte werden jedoch auch Produkte mit einem breiteren Anwendungskontext betrachtet. Dies ist der Tatsache geschuldet, dass Technologieunternehmen wie Google oder Microsoft im Bereich des maschinellen Lernens als führend gelten [NVK+15]. Produkte solcher Unternehmen ermöglichen das Trainieren sehr

leistungsfähiger Deep Learning Klassifikationen. Da sie jedoch nicht speziell für die industrielle BV gedacht sind, bedürfen solche Produkte eines zusätzlichen Programmieraufwands zur Prozesseinbettung.

Eine Übersicht der Rechercheresultate ist in Tab. 4.2 gegeben.

Tab. 4.2: Übersicht kommerzieller BV-Softwareprodukte (Eigendarstellung)

Produktname	Hersteller	Schnittstelle	Zugrundeliegende Algorithmen	Spezialisierungsgrad
Amazon Rekognition	AWS	GUI	Deep Learning	Allgemein
Auto ML Vision	Google	GUI	Deep Learning	Allgemein
Custom Vision	Microsoft Azure	GUI	Deep Learning	Allgemein
VisionPro Vidi	Cognex	GUI	Deep Learning	Industrielle BV
Open eVision	Euresys	GUI	Deep Learning	Industrielle BV
Matrox Image Library 10	Matrox	GUI	Deep Learning	Industrielle BV
MVTec Halcon	MVTec	GUI	Deep Learning	Industrielle BV
NeuroCheck	NeuroCheck	GUI	Deep Learning	Industrielle BV
VisionSpace	Omron	GUI	Deep Learning	Industrielle BV
Pekat Vision	Pekat Vision	GUI	Deep Learning	Industrielle BV
Coake 7	SAC Sirius	GUI	Deep Learning	Industrielle BV
Sentin Explorer	Sentin	GUI	Deep Learning	Industrielle BV
CVB Polimago	Stemmer Imaging	GUI	Deep Learning	Industrielle BV
Frame Grabber Software	Basler Web	GUI	Konventionelle BV	Industrielle BV
Insight Explorer	Cognex	GUI	Konventionelle BV	Industrielle BV
VisionPro (incl. Designer)	Cognex	GUI	Konventionelle BV	Industrielle BV
VISIONexpert 5	Compar	GUI	Konventionelle BV	Industrielle BV
Design Assistent	Matrox	GUI	Konventionelle BV	Industrielle BV
MVTec Merlic	MVTec	GUI	Konventionelle BV	Industrielle BV
CV-Inspect	Octum	GUI	Konventionelle BV	Industrielle BV
OCTUMISEr	Octum	GUI	Konventionelle BV	Industrielle BV
Open Vino	Intel	Code library	Deep Learning	Allgemein
Matlab Vision Toolbox	Mathworks	Code library	Deep Learning	Allgemein
Matlab Vision Toolbox	Mathworks	Code library	Konventionelle BV	Allgemein
Vision Library	Cognex	Code library	Konventionelle BV	Industrielle BV
Imaging Library	Matrox	Code library	Konventionelle BV	Industrielle BV
Common Vision Blox	Stemmer Imaging	Code library	Konventionelle BV	Industrielle BV

5 Formale Charakterisierung von Anwendungsfällen industrieller Bildverarbeitung

Um im Rahmen einer Entscheidungslogik Empfehlungen für geeignete Technologien geben zu können, bedarf es neben entsprechender Technologiekenntnisse (siehe Kapitel 3) eines detaillierten Verständnisses des jeweiligen Anwendungsfalls. Hierzu muss der Anwendungsfall eindeutig beschrieben und charakterisiert werden. Zur Charakterisierung von Anwendungsfällen industrieller Bildverarbeitung wird in Kapitel 5.1 zunächst eine allgemeine Herangehensweise in Anlehnung an Richtlinie VDI / VDE 2632 Blatt 2 beschrieben. Darauf aufbauend wird in Kapitel 5.2 das konkrete Modell zur Charakterisierung von Anwendungsfällen im Rahmen dieser Arbeit definiert.

5.1 Allgemeine Charakterisierung von Anwendungsfällen gemäß VDI / VDE 2632 Blatt 2

Die Richtlinie VDI / VDE 2632 Blatt 2 ist ein Leitfaden zur Erstellung eines Lastenheftes und eines Pflichtenheftes im Kontext industrieller Bildverarbeitung [VDI 2632-2]. In einem Lastenheft werden alle Anforderungen an eine technische Lösung einschließlich der Randbedingungen aus Anwendersicht definiert [VDI 3694]. Lastenhefte werden daher normalerweise zur Unterstützung der Kommunikation zwischen einem Anbieter von BV-Systemen (Dienstleister) und dem Anwender (Auftragsgeber) verwendet. Die Beschreibungskomponenten eines Lastenheftes eignen sich allerdings auch für diese Arbeit gut, um einen Anwendungsfall im Rahmen der zu entwickelnden Entscheidungslogik zu spezifizieren.

Die VDI / VDE 2632 Blatt 2 nennt insgesamt 45 zu spezifizierende Einflussgrößen, die in den folgenden sechs Klassen zusammengefasst werden [VDI 2632-2]:

- **Aufgabenstellung:** Zunächst bedarf es einer Einordnung der zur erfüllenden Aufgabenstellung. Diese beinhaltet eine allgemeine Beschreibung des objektbezogenen Prüfungsziels (z. B. Vermessung, Lageerkennung oder Oberflächenprüfung) sowie die Definition konkreter Fehlerklassen. Zudem ist die bisherige Vorgehensweise (z. B. manuelle Sichtkontrolle) zu erläutern und das Prüfungsziel in den produktionstechnischen und höheren betriebswirtschaftlichen Kontext einzuordnen. Es ist weiterhin ein Verfahren zur Validierung der Funktionstüchtigkeit des kamerabasierten Prüfverfahrens zu definieren.
- **Prüfobjekt:** Hierbei gilt es, alle für die Aufgabenstellung relevanten Eigenschaften des zu untersuchenden Objekts zu spezifizieren. Dies umfasst das Erstellen eines repräsentativen Musterkatalogs und das Definieren von

Grenzfällen. Bei der Spezifizierung sind makroskopische und mikroskopische Eigenschaften wie Abmessungen, Rauheit, Farbe, Textur und Reflexionsverhalten zu beschreiben. Außerdem ist eine mögliche Streubreite von Objekten durch zusätzliche externe Effekte wie z. B. Korrosion, Objektverschmutzung oder thermische Instabilität zu benennen.

- **Szene:** Die Charakterisierung der Bildszene dient der Identifikation umgebungsbedingter Einflüsse auf den BV-Prozess. Dabei sind Handhabung, Bewegung und Positionierung des Prüfobjekts zu beschreiben. Weitere relevante Einflüsse können Vibrationen, unterschiedliche Fremdlichteinwirkungen, Umgebungstemperatur, Feuchtigkeit oder Fremdstoffe in der Luft sein.
- **Prozessintegration:** Die Prozessintegration beschreibt, wie das BV-System in existierende Prozesse eingefügt wird. Dies betrifft einerseits physische Aspekte wie räumliche Randbedingungen sowie mechanische und elektrische Schnittstellen. Andererseits muss bestimmt werden, wie die Daten aus dem BV-System weiter genutzt werden sollen. Generierte Daten können so etwa auf Monitoren visualisiert oder in einer Datenbank langfristig gespeichert werden, wofür es unterschiedlicher Übertragungsschnittstellen und Datenprotokolle bedarf.
- **Mensch-Maschine-Schnittstelle:** Hierbei ist das Bedienkonzept des BV-Systems für den Nutzer zu spezifizieren. Dies beinhaltet z. B. die Definition grafischer Benutzeroberflächen sowie von Sprache, Zugriffsrechte und die Prüfprogrammverwaltung.
- **Sonstiges:** Unter dieser Kategorie werden weitere Anforderungen bezüglich der Inbetriebnahme des BV-Systems festgelegt. Dazu können z. B. Anforderungen bzgl. physischer Installation, durchzuführende Schulungen, abzugebende Dokumentation sowie zukünftige Wartungsdienstleistungen festgelegt werden.

Für eine vollständige Erläuterung der einzelnen Einflussgrößen sei auf die entsprechende Richtlinie verwiesen.

5.2 Konkrete Charakterisierung von Anwendungsfällen im Rahmen der Entscheidungslogik

Nicht alle der in VDI / VDE 2632 Blatt 2 definierten Einflussgrößen können auch im Rahmen der Entscheidungslogik betrachtet werden, da sie den Umfang dieser Arbeit deutlich übersteigen würden. Andererseits ist die VDI Richtlinie in einigen Punkten sehr allgemein gehalten, sodass es hier einer stärkeren Konkretisierung bedarf. Für diese Arbeit wird daher eine eigene Charakterisierung von

Anwendungsfällen verwendet, die sich jedoch an der VDI Richtlinie orientiert. Die wichtigsten Einflussgrößen zur Charakterisierung des Anwendungsfalls sind dabei:

- **Art der Inspektionsaufgabe:** Für die Software-Auswahllogik ist zwischen den Aufgabentypen Vermessung, Barcodeauslesen, Farbtreueüberprüfung, Objektdetektion, Objekt- bzw. Musterklassifizierung und Oberflächeninspektion zu unterscheiden. Der Bereich der Objekt- und Musterklassifizierung beinhaltet weitere industrietypische Anwendungen wie optische Buchstabenerkennung oder Montagekontrollen.
- **Variationsgrad der Objekte:** Hierbei ist zu bestimmen, wie stark sich die zu untersuchenden Objekte untereinander unterscheiden. Eine spezifische Übersicht von Variationsformen wird in Kapitel 6.2 vorgestellt.
- **Sichtfeldgröße:** Die Sichtfeldgröße beschreibt Länge und Breite des relevanten Objekts bzw. Objektbereiches aus Kameraperspektive (siehe Abb. 5.1). Ist ein Objekt in seiner Position variabel, so muss die Sichtfeldgröße ausreichend überdimensioniert werden, um das Objekt in jeder Position vollständig zu beinhalten.
- **Minimal benötigte Detailgröße:** Hierbei ist zu spezifizieren, wie groß das kleinste noch scharf aufzulösende Detail ist. Handelt es sich etwa um eine Messanwendung mit einer benötigten Genauigkeit von 0,1 mm, so ist diese als minimal benötigte Detailgröße anzugeben (siehe Abb. 5.1).
- **Objektform:** Im Rahmen der Auswahllogik ist zu spezifizieren, ob ein Objekt eine flache (z. B. Münze) oder eine komplexere, dreidimensionale Form hat (z. B. Kugel).
- **Objekttiefe:** Hierbei ist anzugeben, welcher Tiefenbereich des Objekts aus Kameraperspektive gleichmäßig scharf abzubilden ist (siehe Abb. 5.1). Dabei ist zu beachten, dass lediglich für die Inspektionsaufgabe relevante Bereiche scharf dargestellt sein müssen.
- **Reflexionseigenschaften des Objekts:** Im Rahmen der Auswahllogik ist vereinfacht lediglich zwischen Licht reflektierenden und Licht transmittierenden Objekten zu unterscheiden.
- **Kontrastumfang:** Dieser bezeichnet das Verhältnis von hellstem zu dunkelstem Farbwert, der in einem Bild wiedergegeben werden müssen. Bei Objekten mit einheitlicher Oberfläche ist dieser gering und für die Anwendung unkritisch. Ein hoher Kontrastumfang ist jedoch bei präzisen Farbbestimmungen oder Anwendungen wichtig, bei denen im Sichtbereich sowohl stark reflektierende als auch stark absorbierende Oberflächen vorkommen.
- **Arbeitsabstand:** Der Arbeitsabstand bezeichnet den Abstand zwischen Schärfeebene des Objekts und Objektivfront (siehe Abb. 5.1). Dieser richtet sich nach räumlichen Einschränkungen in der gegebenen

Produktionsumgebung. Für den Anwendungsfall sind minimal und maximal erlaubter Arbeitsabstand zu spezifizieren.

- **Anzahl zu überprüfender Objekte pro Zeiteinheit:** Die Bildaufnahme rate entspricht der Taktrate, mit welcher Objekte visuell zu kontrollieren sind. Sie hängt von der Prozessintegration des BV-Systems in die Produktionslinie ab.
- **Objektgeschwindigkeit:** Je nach Prozessintegration kann die Geschwindigkeit des Objekts relativ zur Kamera im Moment der Bildaufnahme gleich oder größer null sein. Objektgeschwindigkeiten größer null treten z. B. auf, wenn Objekte kontinuierlich auf einem Fließband an der Kamera vorbeibewegt werden.
- **Vibrationen:** Hierbei ist anzugeben, ob die Umgebung der Bildaufnahme starken Vibrationen ausgesetzt ist. Dies ist z. B. bei Presswerken der Fall.
- **Fremdkörperpartikel:** Eine Übersicht zur Spezifizierung der Größe von Fremdkörpern in einer Produktionsumgebung wird in Tab. 3.2 in Kapitel 3 gegeben.
- **Feuchtigkeit:** Eine Übersicht zur Spezifizierung der Feuchtigkeit, die z. B. durch Reinigungsstandards bedingt sein kann, wird in Tab. 3.2 in Kapitel 3 gegeben.

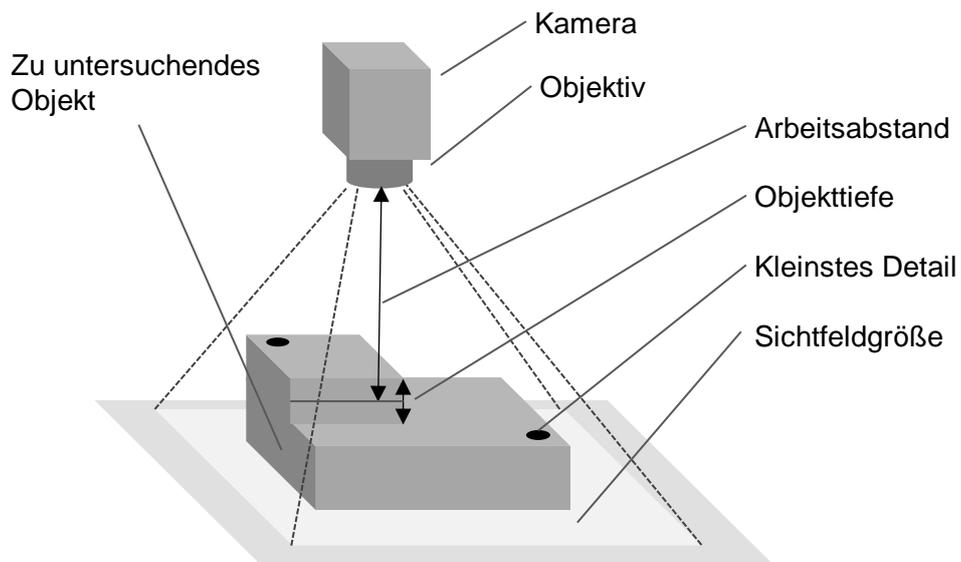


Abb. 5.1: Parameter zur Charakterisierung des Anwendungsfalls

Diese Größen werden in der Entscheidungslogik in Kapitel 6 zur Definition von technischen Anforderungen wieder aufgegriffen.

6 Methodik zur Technologieauswahl

In den vorherigen Kapiteln werden relevante technische Grundlagen, am Markt existierende Produkte sowie Beschreibungsgrößen zur Charakterisierung von Anwendungsfällen vorgestellt. Die dabei erarbeiteten Kenntnisse werden nun zu einer Entscheidungslogik zur Technologieauswahl zusammengeführt. In Unterkapitel 6.1 wird zunächst eine detaillierte Logik zur Auswahl der Hardwarekomponenten Kamera, Objektiv und Beleuchtungssystem vorgestellt. In Unterkapitel 6.2 wird anschließend eine softwarebezogene Auswahllogik eingeführt. Diese strukturiert die Entscheidung, ob für ein Inspektionsproblem Methoden der konventionellen Bildverarbeitung oder Deep Learning verwendet werden sollen.

6.1 Erläuterung der hardwarebezogenen Entscheidungslogik

Die entwickelte Logik zur Auswahl von Kamera, Objektiv und Beleuchtung setzt sich aus fünf größeren Auswahlritten zusammen. Es ist zu empfehlen, die Auswahlritte in der vorgestellten Reihenfolge abzuarbeiten, da diese auf Informationen der vorherigen Schritte aufbauen.

- 1) Die **Kameravorauswahl** dient zur Überprüfung, ob eine 2D-Bildaufnahme die am besten geeignete Technologie zur visuellen Qualitätskontrolle darstellt.
- 2) Falls die 2D-Bildaufnahme geeignet ist, dient der Schritt der **detaillierten Kameraauswahl** zur Bestimmung konkret geeigneter Kameramodelle. Ist eine 2D-Bildaufnahme nicht geeignet, so ist die Entscheidungslogik beendet und die Schritte 2) bis 5) sind nicht auszuführen.
- 3) Bei der **Objektivvorauswahl** wird eine Empfehlung zur zu verwendenden Objektivart ermittelt.
- 4) Soll ein telezentrisches oder ein Festbrennweitenobjektiv verwendet werden, so können geeignete Modelle im Rahmen der **detaillierten Objektivauswahl** bestimmt werden. Wird in Auswahlritt 3) hingegen ein Zoom- oder Spezialobjektiv empfohlen, so ist ein Experte zur weiteren Anforderungsspezifizierung zu konsultieren.
- 5) Die **Auswahl des Beleuchtungssystems** unterstützt die Selektion geeigneter Beleuchtungsmethoden.

In jedem der fünf Auswahlritte werden Charakteristika des Anwendungsfalls in erforderliche Technologieanforderungen überführt. Die Auswahlritte 2 und 4 zur detaillierten Bestimmung geeigneter Kameras und Objektive gehen dabei über die reine Bestimmung von Technologieanforderungen hinaus. Hier werden anhand der Logik und der im Rahmen der Marktrecherche zusammengestellten Datenbank

konkrete Produkte empfohlen und im Umkehrschluss deren Implikationen für die weitere Hardwareauswahl beachtet.

So lässt sich z. B. anhand der benötigten Sichtfeldgröße und der Größe des kleinsten aufzulösenden Details eine theoretisch benötigte Auflösung für die Kamera berechnen. Die tatsächliche Auflösung der Kamera ist jedoch größer gleich der theoretisch benötigten Auflösung, sodass im Umkehrschluss das tatsächliche Sichtfeld auch größer gleich dem theoretisch benötigten Sichtfeld ist. Dies hat wiederum Einfluss auf die anschließende Objektivauswahl. Für eine korrekte Hardwareauswahl ist es also wesentlich, dass nicht nur Informationen über den Anwendungsfall, sondern auch konkrete Produktinformationen mit einfließen.

Die einzelnen Auswahlsschritte werden in den folgenden Unterkapiteln vorgestellt. Hierbei wird zunächst ausschließlich bewertet, welcher technischen Anforderungen es zur Realisierung des gewünschten Anwendungsfalls bedarf. Eine praktische Implementierung der Auswahllogik mit Einbindung einer konkreten Produktdatenbank wird anschließend in Kapitel 8 vorgestellt. Dabei werden neben den technischen Größen auch die Listenpreise der Komponenten betrachtet. Dem Anwender wird dann automatisch die günstigste Komponente vorgeschlagen, welche alle technischen Anforderungen erfüllt.

6.1.1 Kameravorauswahl

Bei der Kameravorauswahl wird überprüft, welche grundlegende Kameratechnologie für den gegebenen Bildverarbeitungsanwendungsfall am besten geeignet ist. Der Anwendungsfall wird anhand von sechs Fragen klassifiziert und damit einer von neun grundlegenden Kameratechnologien zugeordnet (siehe Abb. 6.1). Nur wenn die Anwendung als ein Fall der 2D-Bildaufnahme klassifiziert wird, sind die weiteren Schritte der Entscheidungslogik anwendbar.

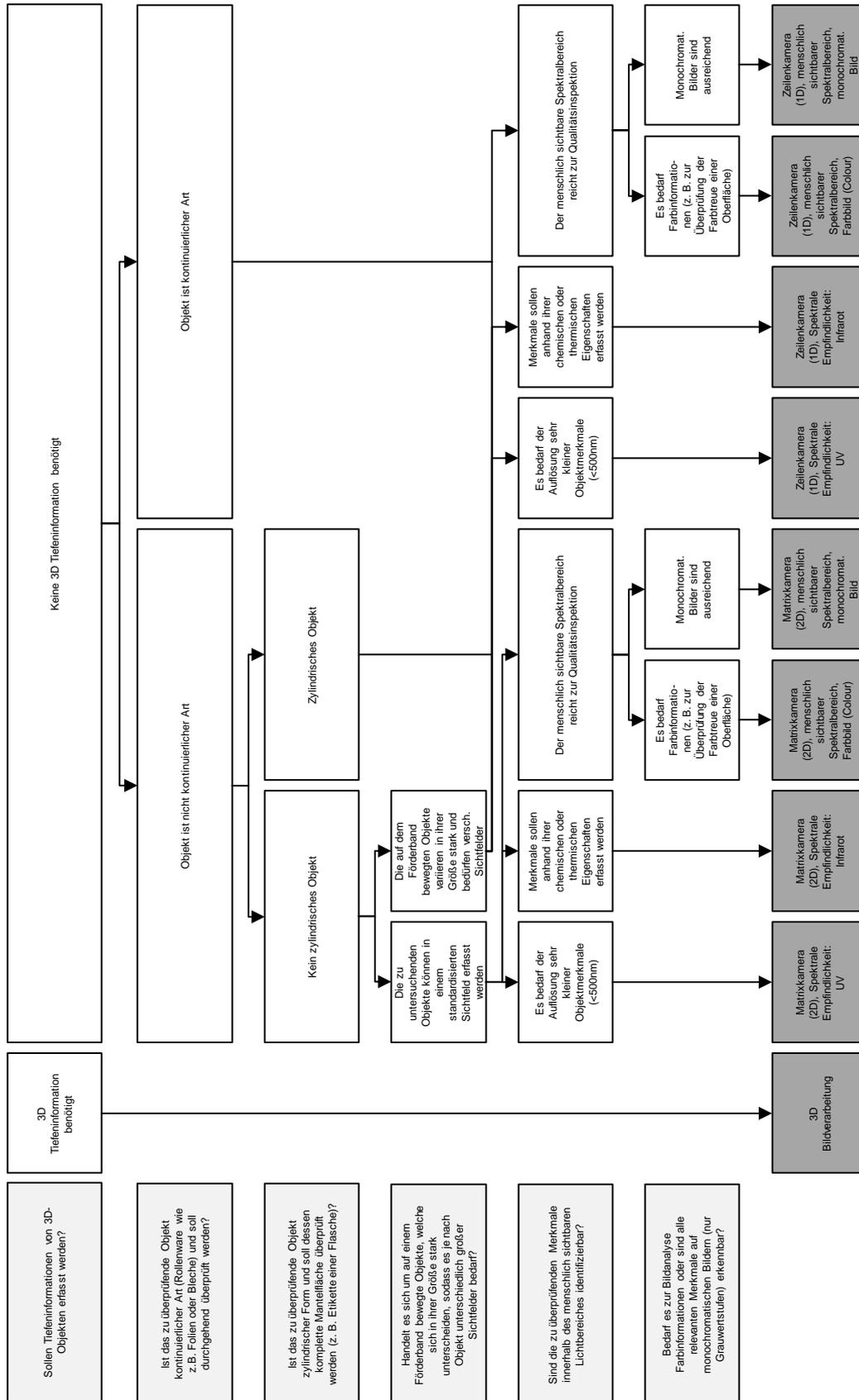


Abb. 6.1: Logik zur Vorauswahl geeigneter Kameratechnologien

6.1.2 Detaillierte Kameraauswahl

In Abb. 6.2 wird eine Übersicht der einzelnen Auswahlsschritte zur Bestimmung einer geeigneten Kamera gegeben. Die Übersicht ist dazu in drei Bahnen eingeteilt. Die oberste Bahn umfasst Parameter zur Beschreibung des Anwendungsfalles (siehe Kapitel 5.2). Zwischenschritte, mittels derer die externen Parameter in Anforderungen auf Technologieebene überführt bzw. umgerechnet werden, sind auf der zweiten Bahn dargestellt. In der dritten Bahn werden die konkreten Anforderungen an die Auswahl der Kamera dargestellt, welche sich aus den Informationsverarbeitungsschritten in den vorherigen beiden Bahnen ergeben. Diese Anforderungen sind von 1 bis 8 durchnummeriert und werden im Folgenden kurz erläutert.

- 1) Zunächst muss anhand der gewünschten Sichtfeldgröße und der Größe des kleinsten aufzulösenden Details die mindestens benötigte Kameraauflösung in horizontaler und vertikaler Richtung berechnet werden. Dies erfolgt anhand der in Gleichung 3.2 vorgestellten Relation. Die tatsächliche Kameraauflösung muss in horizontaler und vertikaler Richtung größer gleich den berechneten Mindestwerten sein.
- 2) Die mindestens benötigte Bildrate der Kamera ergibt sich direkt aus der Anzahl an aufzunehmenden Bildern pro Sekunde.
- 3) Der Nutzer hat zu spezifizieren, mit welcher Geschwindigkeit sich das Objekt während der Bildaufnahme bewegt. Bewegt sich das Objekt während der Aufnahme, so sind Kameras mit einem Rolling Shutter Verschluss für die Anwendung auszuschließen. Aus der Objektgeschwindigkeit lässt sich zudem die maximale Belichtungszeit herleiten, welche für die Auswahl des Beleuchtungssystems von Relevanz ist (siehe Kapitel 3.1.1.2).
- 4) Wie in Kapitel 3.1.1.2 vorgestellt, sind Signal / Rausch-Verhältnis (SNR) und Empfindlichkeit eines Bildsensors die relevanten Größen zur Beschreibung des Kontrastumfangs und der mindestens benötigten Belichtung eines Bildsensors. Diesen beiden Größen sind für die Auswahllogik jedoch in zweierlei Hinsicht problematisch. Einerseits werden Angaben zu diesen Werten oft nicht in Produktübersichten aufgeführt, sodass zunächst umständlich der Bildsensor und das zugehörige Datenblatt ermittelt werden muss. Andererseits sind die beiden Größen für eine Vielzahl von industriellen Anwendungen auch nicht weiter relevant, da die Bildszene gut ausgeleuchtet werden kann und es nur selten eines besonders hohen Kontrastumfangs bedarf. Zur Auswahl stehende Kameramodelle sollten also nur hinsichtlich dieser beiden Größen untereinander verglichen werden, wenn die Informationen einfach zur Hand sind oder Empfindlichkeit bzw. Kontrastumfang bei der Anwendung von besonderer Bedeutung sind. Ist dies nicht der Fall, so können Empfindlichkeit und SNR mittels der Pixelgröße approximiert werden (siehe Anforderung 5).
- 5) Die Pixelgröße eines Bildsensors kann als Approximation für dessen relative Empfindlichkeit und SNR verwendet werden. Dabei haben kleine Pixel ($< 5 \mu\text{m}$)

gegenüber großen Pixeln ($> 8 \mu\text{m}$) eine bessere Empfindlichkeit bei schlechten Lichtverhältnissen, aber ein geringeres SNR. Große Pixel eignen sich entsprechend besser für Anwendungen mit Anforderungen an einen hohen Kontrast. Pixel mittlerer Größe bieten ein ausgewogenes Verhältnis von SNR und Empfindlichkeit. [Hol10] Liegen bei einer BV-Anwendung keine besonderen Anforderungen bzgl. Kontrastumfang oder Empfindlichkeit vor, so sind Pixel mittlerer Größe zu empfehlen. Sensoren mit mittleren oder großen Pixeln sind aufgrund der entsprechend größeren Sensorfläche teurer als Sensoren mit kleinen Pixeln. Jedoch stellen größere Pixel auch eine geringere Anforderung an die notwendige Auflösungsqualität des Objektivs, sodass dieses wiederum günstiger als bei kleinen Pixeln zu erwerben ist [Ste18].

- 6) Eine weitere, zu spezifizierende Anforderung an die Kamera ist die zu verwendende Schnittstelle zur Datenübertragung. Da die zu übertragende Datenmenge bei der Mehrheit der Bildverarbeitungsanwendungen keine Einschränkung bzgl. der Schnittstellenauswahl darstellt, wird die Integrationsfähigkeit in bestehende Systeme als primäres Auswahlkriterium herangezogen.
- 7) In diesem Auswahlschritt ist der gewünschte Schutzgrad des Kamerasystems zu bestimmen. Hierzu muss zunächst die Produktionsumgebung bzgl. Feuchtigkeit und auftretenden Fremdkörpern klassifiziert werden. Mittels Tab. 3.2 kann dann der notwendige IP-Schutzgrad des Systems abgeleitet werden. Auch die anderen Hardwarekomponenten (insbesondere elektrische Komponenten der Beleuchtung) müssen diesen Schutzgrad erfüllen.
- 8) Zusätzliche Kameraanforderungen, die nicht über die bisherigen Beschreibungsgrößen standardisiert spezifiziert werden, können ggf. über eine Einschränkung der Kameraauswahl auf einzelne Hersteller formuliert werden. Dazu muss der Anwender der Auswahllogik jedoch das grobe Produktportfolio bzw. die spezielle Expertise des Unternehmens bereits kennen. So ist z. B. der Hersteller Cognex bekannt für Kameralösungen mit integriertem Prozessor und dazu abgestimmter Software für lokale Bildverarbeitung.

6.1.3 Objektivvorauswahl

Bei der Objektivvorauswahl ist der Anwendungsfall gemäß den Fragen in Abb. 6.3 zu klassifizieren. Anhand dieser Fragen wird festgelegt, welcher der in Kapitel 4.1.3 vorgestellten Objektivtypen zu verwenden ist. Wenn für die Anwendung ein telezentrisches oder ein Festbrennweitenobjektiv geeignet ist, kann dieses im nächsten Auswahlschritt detailliert bestimmt werden. Auf die detaillierte Auswahl von Zoom- und Spezialobjektiven wird nicht weiter eingegangen, da es sich hierbei um in der Industrie selten vorkommende Sonderfälle handelt.

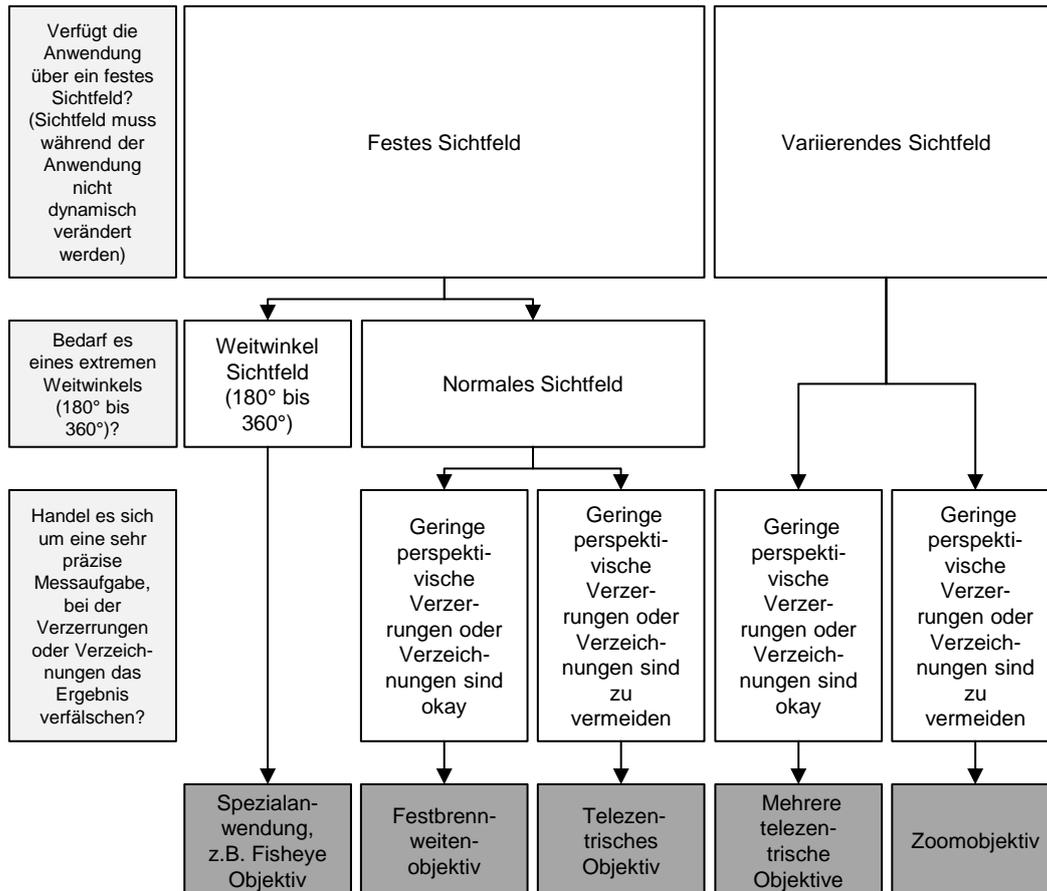


Abb. 6.3: Logik zur Objektivvorauswahl

6.1.4 Detaillierte Objektivauswahl

In der detaillierten Objektivauswahl werden, je nach Resultat der vorherigen Auswahlstufe, konkret geeignete telezentrische bzw. Festbrennweitenobjektive für den jeweiligen Anwendungsfall bestimmt. Die Vorgehensweisen unterscheiden sich dabei je nach Objektivart leicht. Daher wird zunächst die Vorgehensweise für Festbrennweitenobjektive vorgestellt. Anschließend wird auf Besonderheiten bei der Auswahl von telezentrischen Objektiven eingegangen.

6.1.4.1 Detaillierte Auswahl von Festbrennweitenobjektiven

Eine Übersicht der Auswahlsschritte ist in Abb. 6.5 dargestellt. Auch hier werden mehrere Bahnen zur Kategorisierung der einzelnen Zwischenschritte verwendet. Im Vergleich zur Kameraauswahl wird für die Darstellung der detaillierten Objektivauswahl noch eine zusätzliche Bahn hinzugefügt. Diese enthält technische Parameter der zuvor ausgewählten Kamera, welche einen Einfluss auf die Auswahl des Objektivs haben. Die Auswahlsschritte setzen sich wie folgt zusammen:

- 1) Je nach Beschichtungsart und weiteren Linseneigenschaften variiert die spektrale Eignung von Objektiven [Ste18]. Hersteller spezifizieren daher für ihre Objektive, ob diese auf herkömmliche monochromatische bzw. Farbanwendungen oder auf Spezialanwendungen im UV- oder Infrarotbereich ausgelegt sind. Bei der Auswahl ist daher zu beachten, dass die Empfindlichkeit des Objektivs mit der spektralen Empfindlichkeit der zuvor ausgewählten Kamera übereinstimmt.
- 2) Um das Objektiv sicher an die Kamera anschließen zu können, muss dieses über die gleiche mechanische Anschlusschnittstelle wie die Kamera verfügen (z. B. C-Mount, F-Mount).
- 3) Das Auflösungsvermögen des Objektivs muss mindestens so hoch wie das Auflösungsvermögen des ausgewählten Kamerasensors sein, um die Bildqualität nicht zu reduzieren. Dazu muss die minimale Pixelgröße des Objektivs kleiner gleich der Pixelgröße des Bildsensors sein [Len06]. Hierdurch wird gewährleistet, dass Details pixelgenau abgebildet werden können. Damit dies für alle Pixel des Bildsensors gilt, muss zusätzlich Bedingung 4) erfüllt sein.
- 4) Der Lichtkreisdurchmesser des Objektivs muss größer gleich der Bilddiagonale des Bildsensors sein, um Vignettierungseffekte zu vermeiden (siehe Kapitel 3.1.2.2).
- 5) Der mit dem Objektiv realisierbare Arbeitsabstand muss innerhalb des vom Anwender gewünschten Arbeitsabstandsintervalls liegen. Da der realisierbare Arbeitsabstand bei Festbrennweitenobjektiven von mehreren externen Faktoren abhängt, bedarf es zu dessen Bestimmung einiger Zwischenschritte. Bei Umstellung der Gleichung (3.7) ergibt sich für den Arbeitsabstand g in Abhängigkeit der Brennweite f , der Sichtfeldgröße G und der Bildgröße B folgender Wert:

$$g = \frac{B+G}{B} \times f \quad (6.1)$$

Die Sichtfeldgröße G bezieht sich auf die Diagonale des tatsächlichen Sichtfeldes, welches bei vielen Anwendungen größer als die Diagonale des benötigten Sichtfeldes ist (siehe Anhang 11.6). Es muss daher zunächst anhand der horizontalen und vertikalen Auflösung der Kamera, sowie der benötigten Sichtfeldgröße, die tatsächliche Sichtfeldgröße bestimmt werden (Berechnungsdetails siehe Code in Anhang 11.6). Die Bildgröße B entspricht der Diagonalen

des ausgewählten Kamerasensors. Die Brennweite wird vom Objektivhersteller angegeben. Die Brennweite wird dabei im Produktdatenblatt unter der Annahme angegeben, dass die Bildsensordiagonale dem Lichtkreisdurchmesser des Objektivs entspricht. Ist der Lichtkreisdurchmesser jedoch größer als der Bildsensor, so wird nur ein Teil des Lichtkreisdurchmessers für die Bildaufnahme genutzt. Der tatsächlich auf dem Bildsensor abgebildete Sichtfeldbereich ist also kleiner als der im Lichtkreis abgebildete Sichtbereich (siehe Abb. 6.4)

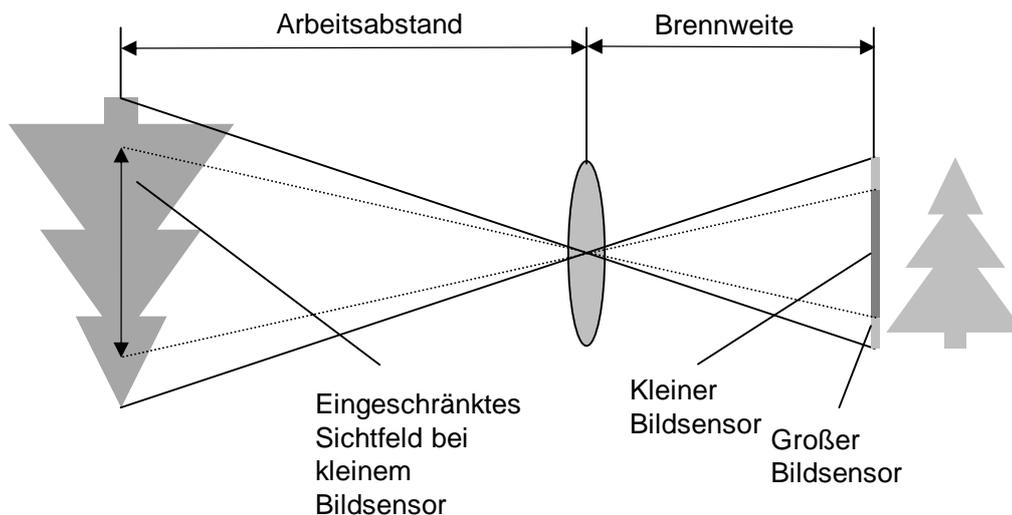


Abb. 6.4: Zusammenhang von Sensor- und Sichtfeldgröße

Zur korrekten Arbeitsabstandberechnung muss daher noch das Verhältnis von Lichtkreisdurchmesser zu Bilddiagonale einbezogen werden, sodass sich mit Gleichung 6.1 folgender Ausdruck ergibt:

$$G = \frac{B+G}{B} \times f \times \frac{\text{Lichtkreisdurchmesser}}{\text{Sensordiagonale}} \quad (6.2)$$

Das Verhältnis von Lichtkreisdurchmesser zu Sensordiagonale bezeichnet man auch als Formatfaktor. [Len06]

- 6) In Schritt 5) wird der mit einem ausgewählten Objektiv realisierbare Arbeitsabstand zunächst berechnet und anschließend mit dem gewünschten Arbeitsabstandsintervall verglichen. Zusätzlich muss nun noch überprüft werden, ob der berechnete Arbeitsabstand größer als die objektivspezifische minimale Objektdistanz (MOD, siehe Kapitel 3.1.2.2) ist. Ist dies nicht der Fall, so kann der zuvor theoretisch berechnete Arbeitsabstand aufgrund des begrenzten Objektivtubusses nicht realisiert werden.
- 7) Die Blendenzahl eines Objektivs muss ausreichend hoch sein, um die für die Anwendung geforderte Schärfentiefe realisieren zu können. Die benötigte Blendenzahl k lässt sich in Abhängigkeit des berechneten Arbeitsabstandes g_0 , der Pixelgröße C , der tatsächlichen Brennweite f und der geforderten Schärfentiefe Δg gemäß der Relation aus Gleichung 3.7 berechnen. Zusätzlich

muss der Anwender nun noch spezifizieren, ob die Anwendung starken Erschütterungen bzw. Vibrationen unterliegt. Ist dies nicht der Fall, so kann ein Objektiv mit verstellbarer Blende verwendet werden. Hier muss überprüft werden, ob die maximal einstellbare Blendenzahl größer als die benötigte Blendenzahl ist. Die Blendenzahl kann dann bei der Erprobung des BV-Systems noch verstellt werden. Liegen hingegen starke Vibrationen vor, so ist ein Objektiv ohne veränderbare Blende zu verwenden. Die Blende ist dann nicht spontan verstellbar, sodass bereits beim Kauf auf eine ausreichend hohe Blendenzahl geachtet werden muss. Es dürfen jedoch auch keine unnötig hohen Blendenzahlen verwendet werden, da sich mit zunehmender Blendenzahl die einfallende Lichtmenge reduziert (siehe Kap. 3.1.2.2). Zum Ausgleich würden dann leistungsstärkere und entsprechend teurere Beleuchtungssysteme benötigt werden [DSS11]. Zudem können bei sehr hohen Blendenzahlen ungewünschte Beugungserscheinungen auftreten, die fehlerhafte Bilder erzeugen [Sil18].

- 8) Wie bereits bei der Kameraauswahl erläutert, können sonstige, eher qualitativ beschreibbare Produkthanforderungen über eine Eingrenzung der betrachteten Hersteller in die Logik eingebracht werden.

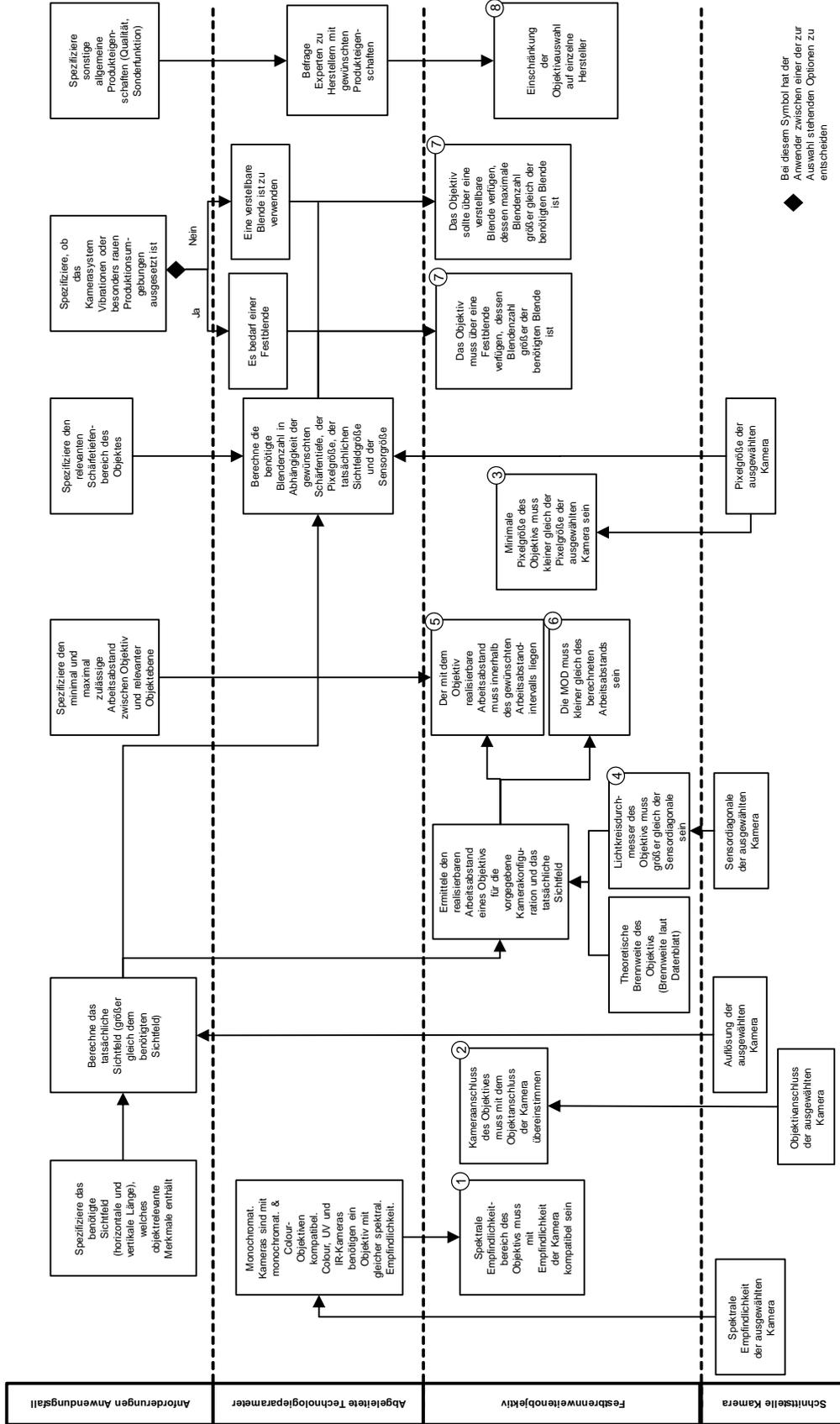


Abb. 6.5: Detaillierte Logik zur Auswahl von Festbrennweitenobjektiven

6.1.4.2 Detaillierte Auswahl von telezentrischen Objektiven

Die Logik zur Auswahl von telezentrischen Objektiven ähnelt der Auswahllogik von Festbrennweitenobjektiven. Im Vergleich zu Festbrennweitenobjektiven sind jedoch folgende Unterschiede zu beachten:

- **Größe des objektseitigen Objektivdurchmessers (siehe Anforderung 5):** Aufgrund des achsparallelen Strahlengangs muss bei telezentrischen Objektiven der objektseitig nutzbare Durchmesser größer gleich der Diagonalen des Sichtfeldes sein (siehe Abb. 3.14). Der objektseitig nutzbare Objektivdurchmesser berechnet sich dabei anhand der Diagonalen des Bildsensors und der charakteristischen Vergrößerung (Magnification) des Objektivs wie folgt [Sil18]:

$$\text{Nutzbarer Objektivdurchmesser} = \frac{\text{Sensordiagonale Kamera}}{\text{Magnification Objektiv}} \quad (6.3)$$

- **Arbeitsabstand (siehe Anforderung 6):** Im Gegensatz zu Festbrennweitenobjektiven können Hersteller bei telezentrischen Objektiven aufgrund des vorgegebenen achsenparallelen Strahlengangs den optimalen Arbeitsabstand objektunabhängig bestimmen. Dieser optimale Arbeitsabstand wird als Kennwert in den Produktangaben veröffentlicht. Bei der Objektivauswahl muss also lediglich beachtet werden, ob der optimale Arbeitsabstand des Objektivs innerhalb des geforderten Abstandintervalls liegt.
- **Approximation der Blendenzahl (siehe Anforderung 8):** Die allgemeinen Hinweise zur Auswahl der Blende bei Festbrennweitenobjektiven gelten bei telezentrischen Objektiven ebenso. Lediglich die Berechnung zur Bestimmung der erforderlichen Blendenzahl k ist hierbei anders. Diese kann gemäß folgender Formel in Abhängigkeit der Vergrößerung β , der Pixelgröße C und der gewünschten Schärfentiefe Δg grob approximiert werden [Sil18]:

$$k = \frac{\Delta g \times \beta^2}{8 \times C} \quad (6.4)$$

- **Spezifizierung der Genauigkeit (siehe Anforderungen 9 & 10):** Telezentrische Objektive werden insbesondere für Vermessungsaufgaben verwendet, bei denen es einer hohen Abbildungsgenauigkeit ohne perspektivische Verzerrungen bedarf. Um verschiedene telezentrische Objektive bezüglich ihrer Genauigkeit vergleichen zu können, eignen sich die Herstellerangaben zur prozentualen Verzeichnung und zum MTF-Wert. Die für eine Anwendung maximal erlaubte Verzeichnung lässt sich mit Gleichung 3.11 spezifizieren. Der für die Anwendung mindestens benötigte MTF-Wert lässt sich aus der kleinsten benötigten Detailgröße am Prüfobjekt ableiten. Hierzu kann in der Praxis ein visuelles Testchart wie in Abb. 3.11 verwendet werden.

Eine Übersicht der kompletten Auswahllogik für telezentrische Objektive ist Abb. 6.6 zu entnehmen.

6.1.5 Auswahl des Beleuchtungssystems

Aufgrund der grundlegend verschiedenen Bauformen von Beleuchtungssystemen werden in Produktdatenblättern auch unterschiedliche Parameter zur Produktbeschreibung verwendet. Verschiedene Beleuchtungsprodukte können daher kaum mittels quantitativer Parameter verglichen werden, wie es bei Kameras und Objektiven der Fall ist. Zudem kann mittels der Parameter des Beleuchtungssystems die Interaktion von ausgestrahltem Licht und zu untersuchendem Objekt nicht ausreichend präzise beschrieben werden. In der Praxis werden potenzielle Beleuchtungslösungen stattdessen oft experimentell für den spezifischen Anwendungsfall verglichen [Ste18].

Trotz mangelnder quantitativer Vergleichbarkeit lassen sich zumindest einige qualitative Überlegungen zur Vorauswahl von geeigneten Beleuchtungssystemen treffen. Diese umfassen die folgenden Größen des Beleuchtungssystems:

- Beleuchtungswinkel und Art der Lichtausbreitung
- Wellenlänge des ausgestrahlten Lichtes
- Beleuchtungssteuerung

6.1.5.1 Beleuchtungswinkel und Art der Lichtausbreitung

Die Kombination aus Beleuchtungswinkel (relative Position der Beleuchtung zu Objekt und Kamera) und der Art der Lichtausbreitung (siehe Kapitel 3.1.3.2) kann zusammenfassend als Beleuchtungsanordnung verstanden werden. In der Fachliteratur werden unterschiedliche Beleuchtungsanordnungen anhand ihrer Strahlengänge klassifiziert und mit konkreten Beispielen veranschaulicht [Jah06; Ste18]. Dabei wird von einer Beleuchtungsanordnung ausgehend deren Wirkweise auf zu inspizierende Objekte unterschiedlicher Form erläutert.

Zur Erstellung der hier vorgestellten Entscheidungshilfe wird die Darstellungsweise der Literatur umgekehrt. Es wird also nicht von Beleuchtungsanordnung und Objekteigenschaften ausgehend auf die Beleuchtungswirkung zurückgeschlossen. Stattdessen wird hier anhand von gewünschter Beleuchtungswirkung und Eigenschaften des zu inspizierenden Objekts ausgehend eine Empfehlung für geeignete Beleuchtungsanordnungen gegeben. Der Anwender muss dazu Objekt und gewünschte Beleuchtungswirkung bezüglich folgender vier Kriterien definieren:

- **Objektform:** Ist das zu untersuchende Objekt bzw. der relevante Objektbereich vollkommen flach (z. B. bedrucktes Blatt Papier) bzw. weist es lediglich geringe Unebenheiten oder Erhöhungen (z. B. gestanzte Münze) auf? Oder ist das zu untersuchende Objekt bzw. der relevante Objektbereich eher tief, wellig oder gewölbt (z. B. Kugel, zerknittertes Papier, Flasche)?

- **Reflektionseigenschaften des Objekts:** Reflektiert (z. B. Münze) oder transmittiert (z. B. gefülltes Wasserglas) das Objekt primär Licht?
- **Helligkeit paralleler Flächen:** Sollen Flächen des Objekts, welche sich parallel zum Bildsensor der Kamera befinden, im Vergleich zu nicht-parallelen Flächen hell oder dunkel dargestellt werden?
- **Helligkeit nicht-paralleler Flächen:** Sollen Flächen des Objekts, welche nicht-parallel zum Bildsensor der Kamera sind, hell oder dunkel dargestellt werden? Nicht-parallele Flächen können z. B. Kantenflächen, Konturen, Wölbungen oder kleine Unebenheiten sein.

Die Entscheidung, welche Flächen im Bild hell oder dunkel erscheinen sollen, ist von der beabsichtigten Inspektionsaufgabe abhängig. Sollen etwa Konturen oder Kratzer vermessen werden, so sollten parallele und nicht-parallele Flächen unterschiedlich hell sein (also hell-dunkel oder dunkel-hell). Soll stattdessen das Muster einer gewölbten Verpackung untersucht werden, so sollten sowohl parallele als auch nicht-parallele Flächen gleichmäßig hell dargestellt werden. Eine Übersicht zulässiger Kombinationsmöglichkeiten von Kriterien und daraus folgende Vorschläge zur Beleuchtungsanordnung ist Abb. 6.7 zu entnehmen. Schematische Darstellungen der Beleuchtungsanordnungen und zugehörige Anwendungsbeispiele sind in den Abb. 6.8 bis Abb. 6.10 dargestellt.

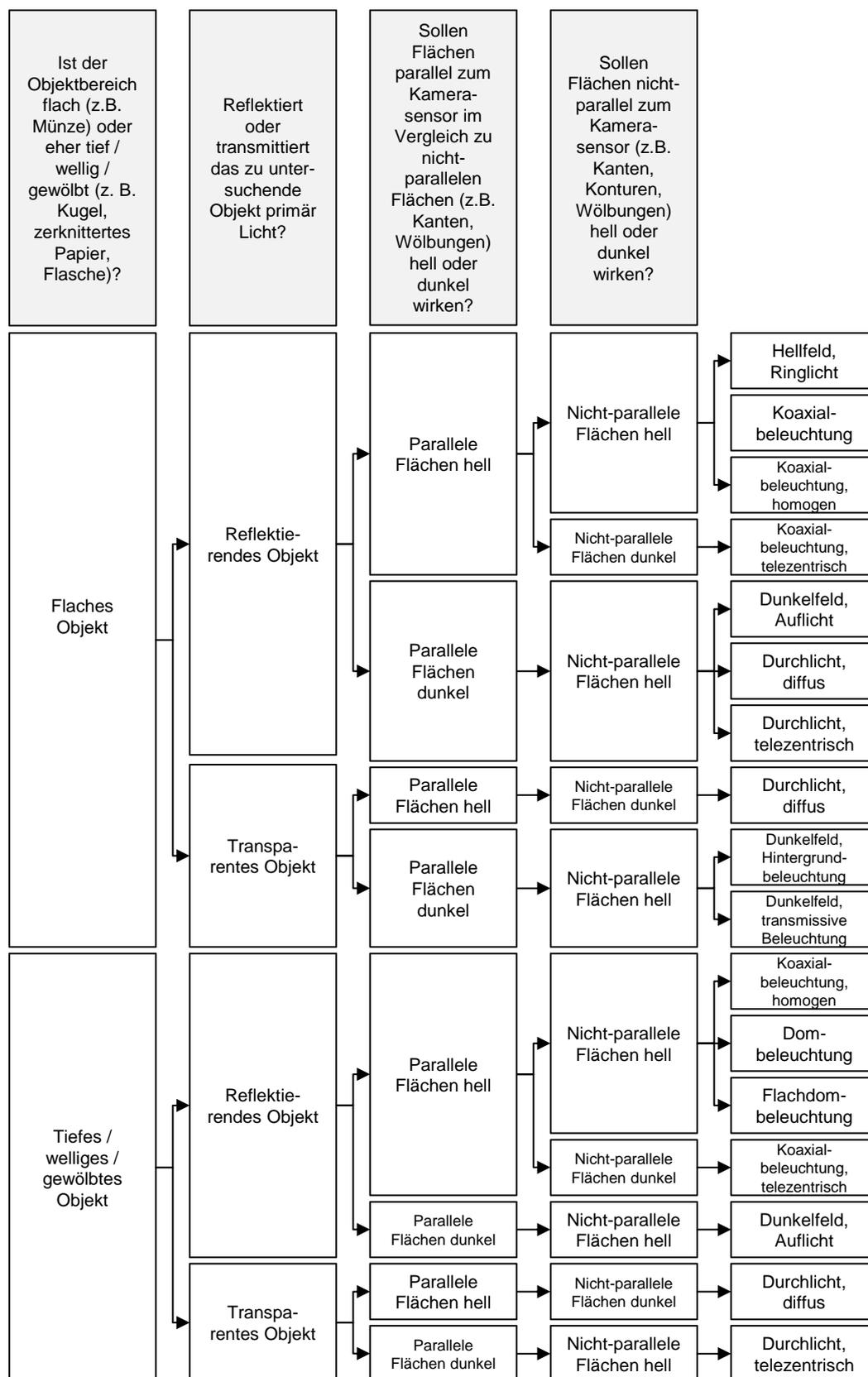


Abb. 6.7: Logik zur Auswahl der Beleuchtungsanordnung

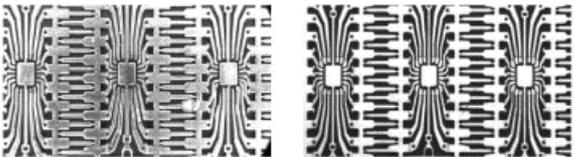
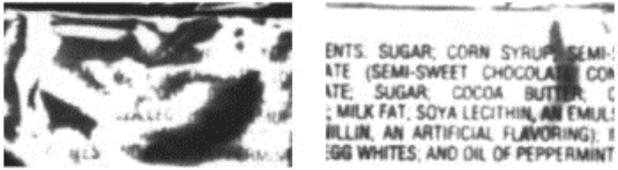
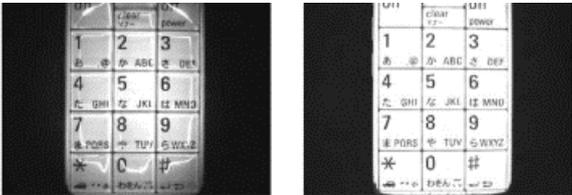
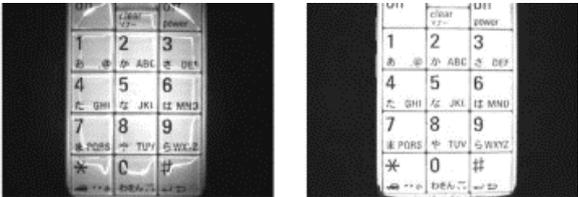
	Anordnung	Anwendungsbeispiel
Hellfeld (Ringlicht)		 Geprägte Münze unter normaler Hellfeldbeleuchtung (Ringlicht).
Koaxialbeleuchtung		 Reflektierendes Objekt wird mit Ringlicht (links) ungleichmäßig ausgeleuchtet. Rechts erfolgt eine Ausleuchtung mit Koaxialbeleuchtung.
Koaxialbeleuchtung homogen		 Aufdruck einer gewellten Folienverpackung unter Ringlicht (links) und homogener Koaxialbeleuchtung (rechts).
Dombeleuchtung		 Tastenfeld unter Ringlicht (links) und Dombeleuchtung (rechts).
Flachdombeleuchtung		 Tastenfeld unter Ringlicht (links) und Flachdombeleuchtung (rechts).

Abb. 6.8: Übersicht möglicher Beleuchtungsanordnung Teil 1 i. A. a. [Ste18; Per10]

	Anordnung	Anwendungsbeispiel
Koaxialbeleuchtung telezentrisch	<p>Fokussieroptik LED</p>	<p>Batterie mit einer flachen Delle unter einer normalen (links) und einer telezentrischen Koaxialbeleuchtung (rechts).</p>
Dunkelfeld Auflicht	<p>LED</p>	<p>Geprägte Münze unter Dunkelfeldbeleuchtung.</p>
Dunkelfeld Hintergrundbeleuchtung	<p>Merkmale LED</p>	<p>Kratzer und geprägter Schriftzug einer transparenten Schüssel mit normaler Hintergrundbeleuchtung (links) und mit Dunkelfeldbeleuchtung (rechts).</p>
Dunkelfeld transmissive Beleuchtung	<p>Merkmale</p>	<p>Geätztes Glas mit einer Koaxialbeleuchtung (links) und einer transmissiven Dunkelfeldbeleuchtung (rechts).</p>
Durchlicht diffus	<p>Diffusor</p>	<p>Elektronikbauteil mit Hellfeldbeleuchtung (links) und diffuser Durchlichtbeleuchtung (rechts).</p>

Abb. 6.9: Übersicht möglicher Beleuchtungsanordnung Teil 2 i. A. a. [Ste18; Per10]

	Anordnung	Anwendungsbeispiel
Durchlicht telezentrisch	<p>Kollimiertes Licht Fokussieroptik</p>	<p>Transparenter Nadelkopf mit diffuser Durchlichtbeleuchtung (links) und telezentrischer Durchlichtbeleuchtung (rechts).</p>

Abb. 6.10: Übersicht möglicher Beleuchtungsanordnung Teil 3 i. A. a. [Ste18; Per10]

Bei einigen der Kombinationen sind mehrere Beleuchtungsanordnungen denkbar. Hier ist zunächst die einfachste bzw. kostengünstigste Lösung auszuprobieren. So sind z. B. bei einfachen Vermessungsaufgaben von Objektkonturen diffuse Durchlichtbeleuchtungen ausreichend, lediglich bei sehr präzisen Messaufgaben mit telezentrischen Objektiven muss eine teurere, telezentrische Durchlichtbeleuchtung verwendet werden.

Die Komponenten der einzelnen Beleuchtungsanordnungen sind in unterschiedlichen Größen zu erwerben. Hierbei ist insbesondere bei Ringlichtern, Koaxial- und Dombeleuchtungen sowie Durchlichtern darauf zu achten, dass die von ihnen ausgeleuchtete Fläche größer als das zu untersuchende Objekt ist.

6.1.5.2 Wellenlänge des ausgestrahlten Lichtes

Bei der Auswahl der Wellenlänge bzw. Farbe des Beleuchtungssystems sind folgende Grundsätze zu beachten, die sich an den Vorgaben der Literatur orientieren [Ste18; Fer20]:

- **Weißes LED-Licht** gilt als die Standardlichtquelle für die Aufnahme von farbigen und monochromatischen Bildern, da es einen dem natürlichen Sonnenlicht ähnlich breiten Spektralbereich umfasst. Bildaufnahmen mit weißem Licht entsprechen dem menschlichen Farbempfinden.
- Durch **farbiges Licht** können bei monochromatischen Bildaufnahmen einzelne Objektfarben gezielt hervorgehoben oder unterdrückt werden, wodurch der anschließende Bildverarbeitungsprozess erleichtert werden kann. Abb. 3.21 gibt eine Übersicht der Wirkung von verschiedenen Lichtfarben. In der Praxis kommt farbiges Licht nur bei geeigneten Sonderfällen zum Einsatz.
- Werden spezielle **UV- oder Infrarotlicht**-empfindliche Kameras verwendet (siehe Kameravorauswahl), so ist auch eine Lichtquelle mit der gleichen spektralen Abstrahlung zu verwenden.

6.1.5.3 Beleuchtungssteuerung

Es ist grundsätzlich zu empfehlen, eine LED-Beleuchtung auszuwählen, welche im gepulsten Blitzbetrieb betrieben werden kann. Dies hat die folgenden Vorteile:

- Durch die eingeschränkten Beleuchtungszeiten werden im Vergleich zum Dauerbetrieb langfristig Energiekosten eingespart.
- Es wird eine Überhitzung der LED vermieden, wodurch deren Langlebigkeit erhöht und Wartungskosten reduziert werden.
- Im gepulsten Blitzbetrieb kann die LED kurzzeitig einen Lichtstrom abgeben, der über der Nennleistung im Dauerbetrieb liegt. Mit einer höheren Belichtung kann einerseits die Belichtungszeit reduziert werden, um Bewegungsunschärfe bei bewegten Objekten zu vermeiden. Andererseits kann die Blende erhöht werden, um so die Schärfentiefe der Bildaufnahme zu erhöhen.

Bei der Auswahl eines LED-Blitzcontrollers ist darauf zu achten, dass die einstellbare Pulsdauer und Pulsfrequenz größer gleich der benötigten Belichtungszeit und Bildrate sind. Wird bei sehr kurzen Belichtungs- bzw. Pulsdauern trotz überhöhtem Lichtstrom keine ausreichende Belichtung erreicht, so ist eine LED mit einer höheren Nennleistung zu verwenden. [Ste18]

6.2 Erläuterung der softwarebezogenen Entscheidungslogik

Gemäß der zweiten wissenschaftlichen Fragestellung wird in diesem Unterkapitel eine Logik vorgestellt, welche dem Anwender bei der grundsätzlichen Entscheidung hilft, ob für eine bildbasierte Inspektion konventionelle Algorithmen oder Deep Learning verwendet werden soll. Die Logik basiert einerseits auf dem grundsätzlichen Verständnis der unterschiedlichen Verarbeitungsweisen von Bilddaten, welche in Kapitel 3.2 detailliert beschrieben werden. Andererseits fließen in die Logik praktische Erfahrungen ein, welche bei verschiedenen Bildverarbeitungsprojekten im Kontext des DCC Aachen gemacht wurden. Eine konkrete Anwendung dieser Logik wird in Kapitel 7.3 vorgestellt.

Entscheidungslogik

Die Entscheidungslogik ist vollständig in Abb. 6.11 dargestellt. Dabei wird zunächst anhand der folgenden Kriterien überprüft, ob eine Deep Learning Klassifizierung ggfs. grundsätzlich ausgeschlossen werden kann:

- **Transparenz:** Deep Learning Klassifikationen basieren auf automatisch erlernten statistischen Zusammenhängen. Für den Anwender ist die Grundlage für eine getroffene Klassifizierung daher nicht unmittelbar nachvollziehbar. Soll

ein Werkstück auf präzise definierte Qualitätskriterien überprüft werden, wie sie z. B. in Lieferantenverträgen spezifiziert sein können, so sind manuell programmierte Klassifizierungsalgorithmen besser geeignet.

- **Trainingsdaten:** Zum Trainieren von Deep Learning Klassifikationsmodellen bedarf es für jede Klasse ausreichender Trainingsdaten. In industriellen Produktionen mit geringen Losgrößen oder sehr geringen Fehlerquoten kann es problematisch sein, mit vertretbarem Aufwand ausreichend Bildmaterial zu Fehlerklassen zu sammeln.
- **Rechenleistung:** Das Trainieren von Deep Learning Modellen erfordert mehr Rechenleistung als das Programmieren von konventionellen Algorithmen. Ist die lokal verfügbare Rechenleistung eingeschränkt und soll z. B. zum Schutz der Daten kein externer Cloud-Anbieter verwendet werden, so kann ein ausreichendes Trainieren des Modells zu zeitaufwendig werden.

Wenn es sich bei der Inspektionsaufgabe, bzw. einem zugehörigen Subproblem zur Ausführung der Inspektionsaufgabe, um ein Problem der Objektdetektion, Objektklassifikation oder Anomaliedetektion handelt, so können sowohl Deep Learning als auch konventionelle Algorithmen verwendet werden [Cog18]. Hier gilt es abzuschätzen, welche der beiden Herangehensweisen mit einem geringeren Entwicklungsaufwand verbunden ist.

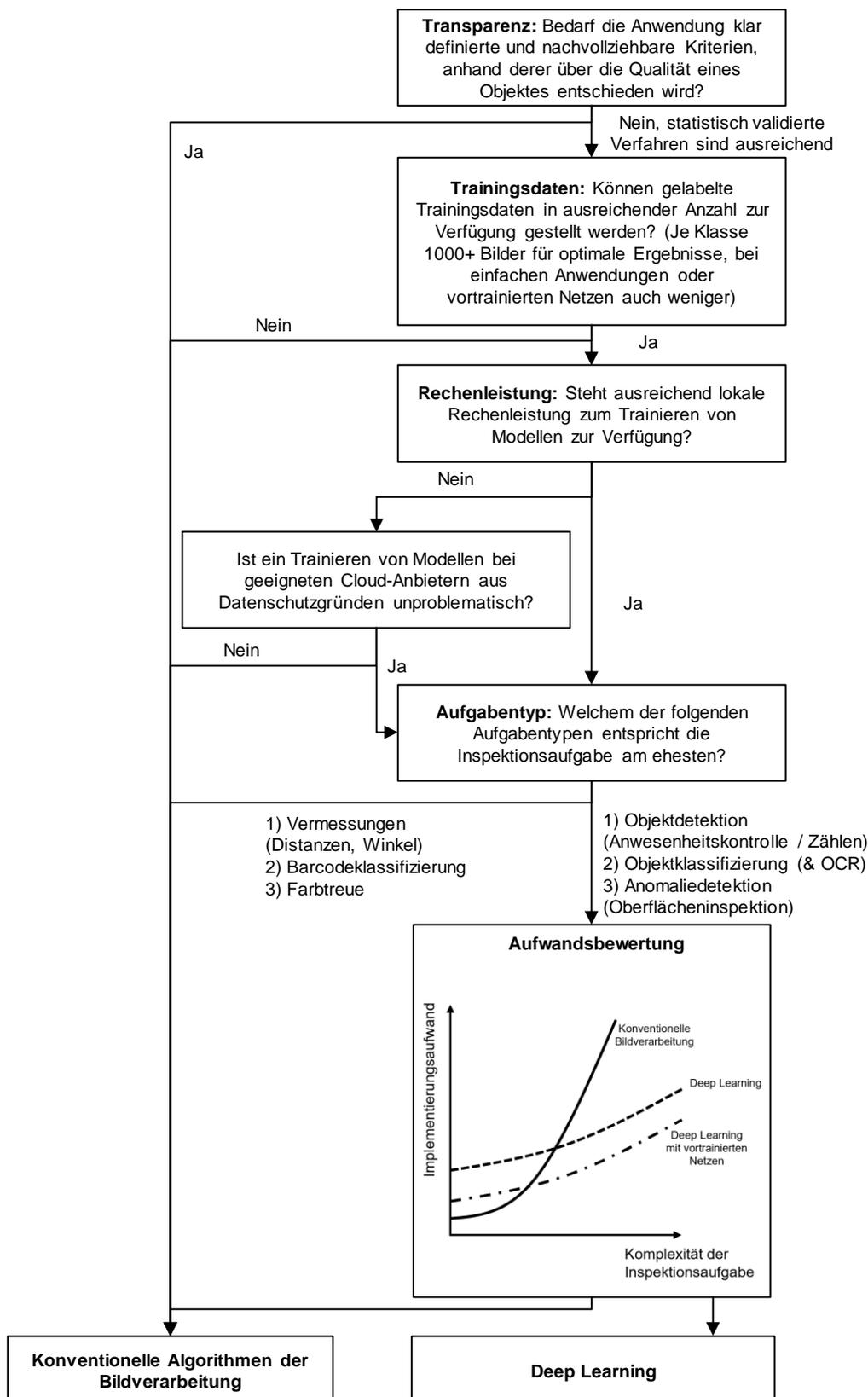


Abb. 6.11: Logik zur Auswahl eines geeigneten Softwareansatzes

Abschätzung des Entwicklungsaufwandes

Für die Abschätzung des Entwicklungsaufwandes wird angenommen, dass geeignete Netzarchitekturen bzw. Bibliotheken mit konventionellen Algorithmen bereits zur Verfügung stehen. Es gilt also nur, diese Architekturen zu trainieren bzw. Algorithmen auszuwählen und zu parametrisieren. Im Kontext des DCC Aachen werden für Deep Learning externe Cloud Services von Microsoft und Google verwendet. Für konventionelle Algorithmen wird die Software Cognex In-Sight Explorer genutzt, welche eine Bibliothek an konventionellen Algorithmen in einer grafischen Benutzeroberfläche bietet.

Der Entwicklungsaufwand für konventionelle Bildverarbeitung und Deep Learning unterscheidet sich in Abhängigkeit von der Komplexität der Inspektionsaufgabe. Es gilt daher zunächst, die Komplexität einer BV-Aufgabe grob abzuschätzen. Die Komplexität einer Aufgabe steigt mit folgenden Faktoren:

- **Anzahl klassifizierungsrelevanter Objektmerkmale:** Im einfachsten Fall reicht ein einzelnes Merkmal aus, um ein Urteil über die Güte des Objekts zu treffen. Oftmals entscheidet jedoch eine Vielzahl von Merkmalen über die Objektqualität.
- **Anzahl der Variationsgrade je Merkmal:** Um ein Merkmal für die Klassifikationsaufgabe verwenden zu können, muss dieses zunächst aus dem Bild extrahiert werden (siehe Kapitel 3.2.1) Dies erweist sich als umso aufwendiger, in je mehr Variationsgraden das relevante Merkmal vorliegen kann. Treiber für Variationen sind in Tab. 6.1 aufgeführt.

Tab. 6.1: Mögliche Variationsgrade bei Objektmerkmalen (Eigendarstellung)

Geometrische Variationen	Photometrische Variationen
Translation / Rotation im Sichtbereich	Variantenbedingte Farbvariationen
Formvarianten des Objekts (z. B. bei verschiedenen Produktvarianten)	Unbeabsichtigte Farbvariationen (z. B. Rost, Verschmutzung)
Formsteifigkeit	Variabilität im Beleuchtungssystem
Überlappungen / Verdeckungen	Externe Lichteinflüsse

Bei Auswertungen mit einer geringen Komplexität sind konventionelle Methoden einfacher zu implementieren (siehe Abb. 6.12). Aufgrund der geringen Anzahl an Varianten lassen sich relevante Merkmale mit wenigen manuell ausgewählten Algorithmen extrahieren und klassifizieren. Mit wenigen Beispielbildern werden dabei sehr hohe Genauigkeiten erreicht. Solche einfachen Anwendungen können auch mit Deep Learning implementiert werden. Um dabei jedoch eine ausreichend hohe statistische Signifikanz zu erreichen, bedarf es hier deutlich mehr

klassifizierter Trainingsdaten. Deren Erhebung und manuelle Klassifizierung bedeuten einen höheren Arbeitsaufwand.

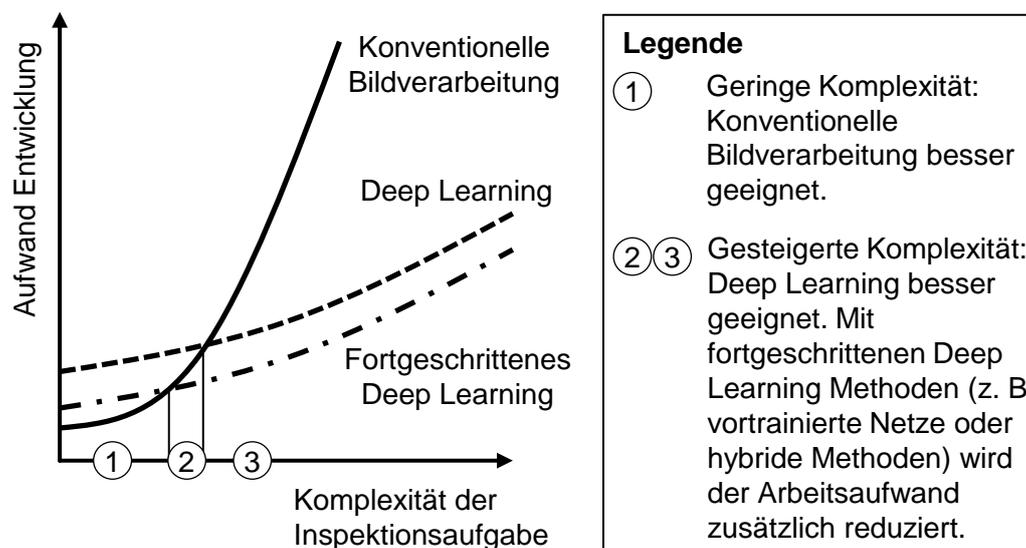


Abb. 6.12: Entwicklungsaufwand für konventionelle Bildverarbeitung und Deep Learning in Abhängigkeit der Aufgabenkomplexität (Eigendarstellung)

Mit zunehmender Komplexität der Aufgabe steigt der Aufwand für die Implementierung bei konventionellen Algorithmen jedoch überproportional an, da jede neue Variante eines relevanten Merkmals zusätzlicher oder komplizierterer Algorithmen bedarf [Cog19]. Ab eines gewissen theoretischen Komplexitätsgrades ist der Implementierungsaufwand nun höher als bei Deep Learning. Bei Deep Learning werden alle relevanten Objektmerkmale automatisch extrahiert, es erfordert also keinen menschlichen Aufwand zur Definition von komplexeren Algorithmen. Mit zusätzlicher Variantenkomplexität sind lediglich mehr Trainingsdaten zu erheben. Die Aufwandskurve steigt folglich auch hier mit zunehmender Komplexität noch an, jedoch deutlich flacher als bei konventioneller Bildverarbeitung. Aufgrund dieser nur flach ansteigenden Aufwandskurve können mittels Deep Learning auch Aufgaben gelöst werden, deren Komplexität einen nicht vertretbaren Aufwand für konventionelle Inspektionsaufgaben darstellen würde.

Auch bei der Verwendung eines Deep Learning Ansatzes kann es sinnvoll sein, zusätzlich noch konventionelle Algorithmen einzubinden. So kann z. B. mittels konventioneller Algorithmen der für die Deep Learning Klassifikation relevante Bildbereich eingegrenzt werden und so das Training vereinfacht werden.

7 Anwendung der Entscheidungslogik

In diesem Kapitel werden anhand eines Anwendungsbeispiels des DCC Aachen die unterschiedlichen Komponenten der zuvor entworfenen Auswahllogik überprüft und veranschaulicht.

In der Demonstrationsfabrik des DCC Aachen werden in einem mehrstufigen Prozess Armbänder produziert. Im letzten Produktionsschritt werden dabei Stofftaschen mit einem RFID-Chip in Handarbeit an das Armband genäht (siehe Abb. 7.1).

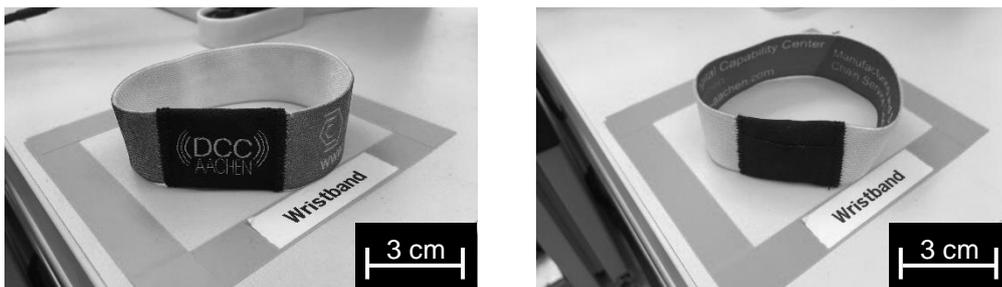


Abb. 7.1: Beispiel des am DCC produzierten Armbandes (links Vorderseite, rechts Rückseite)

Dabei kann es durch eine unpräzise Positionierung der Stofftasche und Ungenauigkeiten bei der Nahtführung zu optischen Qualitätsfehlern an der Naht kommen (siehe Abb. 7.2). Zur visuellen Überprüfung der angenähten Taschen wird daher eine geeignete kamerabasierte Qualitätskontrolle der vernähten Armbandrückseiten implementiert.

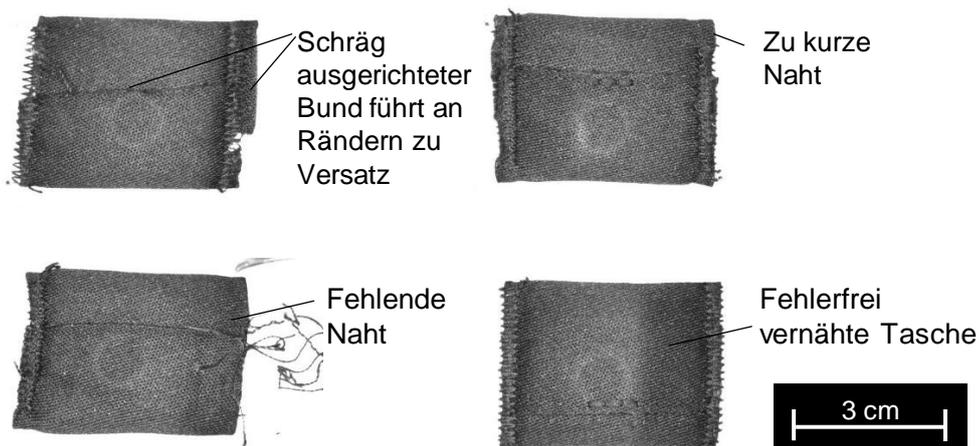


Abb. 7.2: Beispiele auftretender Produktionsfehler beim Vernähen der Armbandtaschen

7.1 Hardwareauswahl

Zur Auswahl einer geeigneten Hardwarelösung werden die Auswahlsschritte der Entscheidungslogik durchlaufen. Dazu wird in den ersten beiden Auswahlsschritten zur Vor- und Detailauswahl einer Kamera der Anwendungsfall gemäß Anhang 11.4 detailliert charakterisiert.

Anhand der daraus resultierenden technischen Anforderungen sowie der expliziten Herstellerpräferenz wird das Cognex Kameramodell In-Sight 2000-230 als günstigste geeignete Lösung befunden. Das Modell hat die in Abb. 7.3 aufgeführten Spezifikationen.

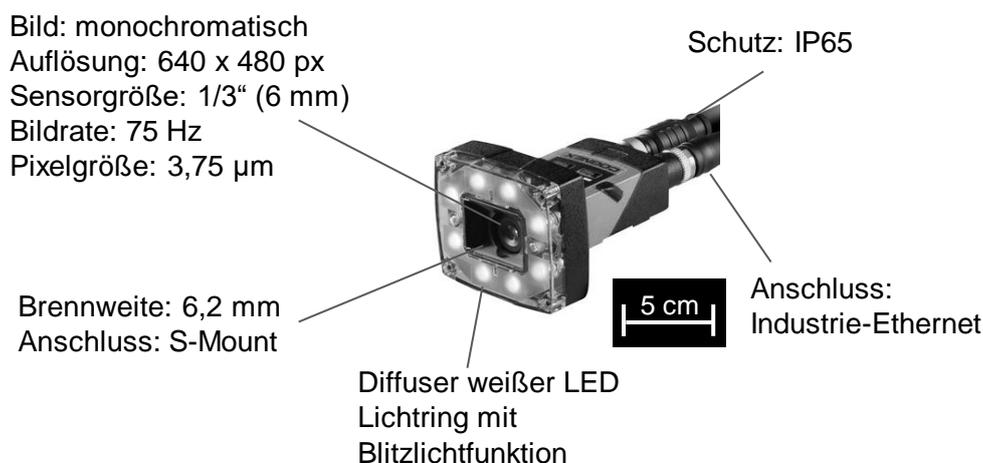


Abb. 7.3: Spezifikationsübersicht Cognex In-Sight 2000

Im dritten Auswahlsschritt wird gemäß Tab. 7.1 ein Festbrennweitenobjektiv für geeignet befunden.

Tab. 7.1: Auswahlsschritt 3 zur Vorauswahl eines Objektivtyps

Frage Auswahllogik	Antwort	Anmerkung
Verfügt die Anwendung über ein festes Sichtfeld?	Ja	Armband wird in Fixierung gespannt.
Bedarf es eines extremen Weitwinkels?	Nein	
Handelt es sich um eine sehr präzise Messaufgabe?	Nein	Es soll primär die Positionierung von Tasche und Nähten kontrolliert werden. Dabei können Vermessungen hilfreich sein, jedoch sind leichte Verzerrungen tolerierbar.
Resultat Logik:	Festbrennweitenobjektiv	

Die anschließende detaillierte Objektivauswahl lässt sich in einer stark reduzierten Form durchführen, da Cognex bereits eine auf die Kamera abgestimmte Auswahl an Objektiven bietet. Da der Prüfstandaufbau insgesamt möglichst kompakt sein soll, liegt der bevorzugte Arbeitsabstand zwischen 5 und 20 cm. Die kleinste

verfügbare Brennweite mit Autofokus beträgt 6,2 mm, sodass sich hiermit nach Gleichung 6.2 ein theoretischer Arbeitsabstand von 76 mm realisieren lässt.

Zur Auswahl des Beleuchtungssystems wird die in Kapitel 6.1.5 vorgestellte Logik angewendet (siehe Anhang 11.5). Als Resultat kommen verschiedene Formen der Hellfeldbeleuchtung (Ringlicht, Koaxialbeleuchtung und homogene Koaxialbeleuchtung) in Frage. Da Cognex für die Kamera ein integriertes Ringlicht anbietet, wird diese Option ausgewählt.

7.2 Überprüfung der Eignung der Hardware

Die zuvor aufgestellten theoretischen Überlegungen zur Eignung der Hardware werden im Rahmen der praktischen Umsetzung überprüft. Dabei wird ein tatsächlicher Arbeitsabstand von 74 mm bei leicht verändertem Sichtfeld gemessen (siehe Abb. 7.4).

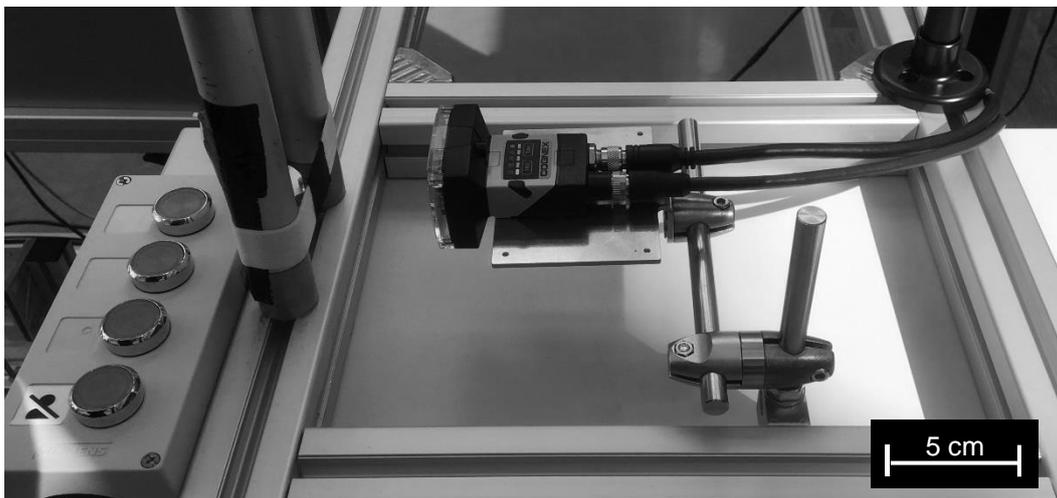


Abb. 7.4: Prototyp Prüfstand mit Cognex In-Sight 2000

Zur Beleuchtung der Prüfscene wird der integrierte Beleuchtungsring der Kamera mit 13 % der Maximalleistung bei einer Belichtungszeit von 5 ms verwendet. Da das Armband während der Bildaufnahme nicht bewegt wird, existieren keine Einschränkungen bezüglich der maximalen Belichtungszeit. Die hier genutzten Werte werden daher rein experimentell bestimmt.

Mit diesen Einstellungen werden die in Abb. 7.5 abgebildeten Aufnahmen der Armbandtaschen erzeugt. Der weiße Objekthintergrund und die weiße Innenseite des Armbandes werden dabei absichtlich überbelichtet, sodass sich die dunkle Armbandtasche und ihre Nähte deutlich abheben. Einzelne Nähte sind trotz der geringen Bildauflösung noch eindeutig zu erkennen (siehe rechtes Bild Abb. 7.5).

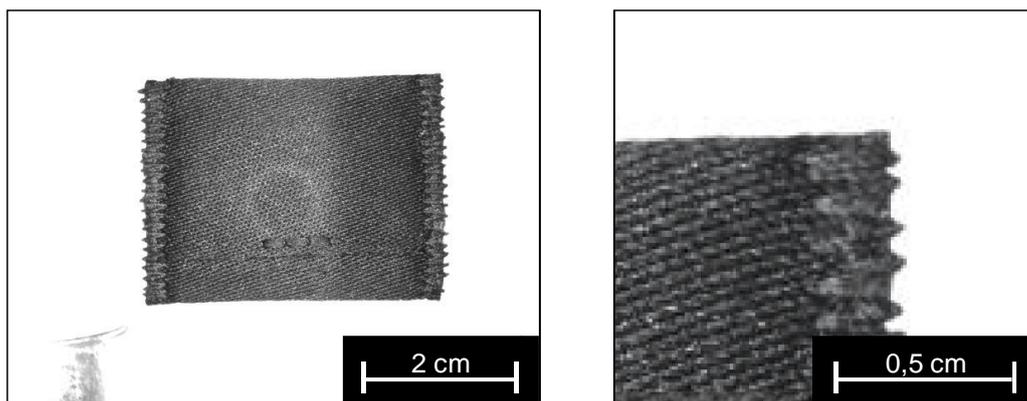


Abb. 7.5: Aufnahme Armbandtasche mit Cognex In-Sight 2000

Eine Aufnahme mit einem zusätzlich getesteten Kameramodell (Cognex In-Sight 7802) mit einer Auflösung von 2 MP ist in Abb. 7.6 zu sehen.

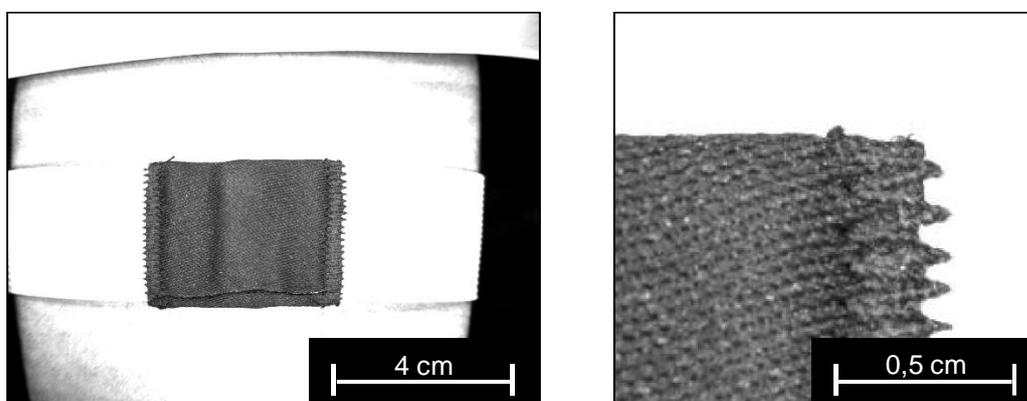


Abb. 7.6: Aufnahme Armbandtasche mit Cognex In-Sight 7802

Mit der höher auflösenden Kamera sind Nähte nur etwas schärfer aufgelöst als zuvor, gleichzeitig lässt sich jedoch ein größeres Sichtfeld realisieren. Dies ermöglicht zusätzliche Flexibilität in der Positionierung des Armbandes. Im gegebenen Anwendungsfall werden die Armbänder jedoch beim manuellen Einspannen in ihrer vertikalen Position an einem Anschlag ausgerichtet. Somit ist eine ausreichend präzise Positionierung möglich, um die BV-Applikation mit dem kleineren Sichtbereich der Cognex In-Sight 2000 zu realisieren.

7.3 Softwareauswahl

Des Weiteren wird der Anwendungsfall auf geeignete Herangehensweisen zur Bildanalyse bewertet. Bezüglich Transparenz, Trainingsdaten und Rechenleistung existieren im gegebenen Anwendungsfall keine relevanten Einschränkungen. Da es sich bei dem Grundproblem um ein Klassifizierungsproblem handelt (gut oder schlecht vernähte Armbandtasche), kommen sowohl Deep Learning als auch

konventionelle Bildverarbeitung in Betracht. Es gilt daher zu bewerten, welche der beiden Methoden einfacher zu implementieren ist. Zur Komplexitätsbewertung lassen sich gemäß Tab. 6.1 folgende Überlegungen aufstellen:

- **Translation / Rotation im Sichtbereich:** Ein Anschlag fixiert die Position von Armbändern in vertikaler Richtung. Aufgrund des verbleibenden horizontalen Freiheitsgrades kann die Taschenposition beim manuellen Einsetzen jedoch weiterhin variieren. Aufgrund des hellen Hintergrundes lässt sich die genaue Position jedoch mit konventionellen Algorithmen schnell bestimmen
- **Formvarianten des Objekts:** Bei dem vorliegenden Anwendungsfall werden lediglich schwarze Taschen mit schwarzem Garn betrachtet. Es liegt somit nur eine Produktvariante vor. Aufgrund der menschlichen Prozessführung unterscheidet sich jedoch selbst bei qualitativ guten Nähten deren Form im Detail. Da genau dieses Merkmal wesentlich für das Qualitätsurteil ist, erzeugt dies ein hohes Maß an Komplexität für die Prüfanwendung.
- **Formsteifigkeit:** Die für Textilprodukte typische Formnachgiebigkeit führt bei jeder Tasche zu einer leicht unterschiedliche Objektform. Dies verstärkt den Effekt der verschiedenen Naht-Formvarianten zusätzlich.
- **Überlappungen / Verdeckungen:** Das Prüfobjekt ist im Sichtfeld ungestört zu erkennen.
- **Photometrische Farbvariationen:** Entscheidungsrelevante photometrische Variationen liegen im Anwendungsfall keine vor. Zwar variieren die Lichtbedingungen in der Produktionsumgebung, jedoch sind diese durch den Einsatz des Blitzlichtes vernachlässigbar gering.

Die Formvariationen der Nähte bedeuten ein hohes Maß an Komplexität, welche mit konventionellen Algorithmen schwierig zu beherrschen ist. Entsprechend ist eine Verwendung von Deep Learning für die Klassifikationsaufgabe zu empfehlen.

7.4 Überprüfung der Eignung verschiedener Softwarelösungen

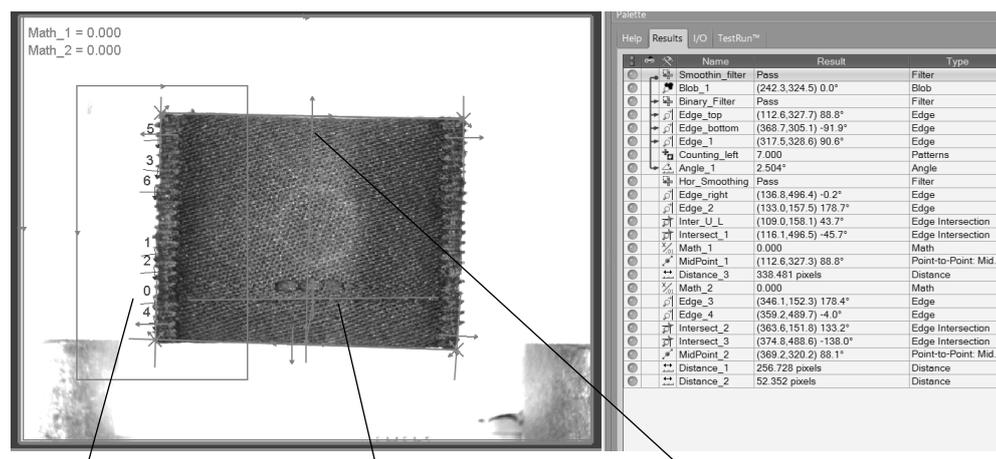
Zur Überprüfung dieser Empfehlung werden drei verschiedene Herangehensweisen für den Anwendungsfall getestet:

- Konventionelle Bildverarbeitung mit Cognex In-Sight Explorer
- Deep Learning mit Teachable Machine von Google LLC, Menlo Park
- Kombination aus konventioneller Bildverarbeitung und Deep Learning mit Cognex In-Sight Vidi

Die Vorgehensweisen und Resultate der unterschiedlichen Ansätze werden im Folgenden kurz vorgestellt.

7.4.1 Konventionelle Bildverarbeitung mit Cognex In-Sight Explorer

Cognex In-Sight Explorer umfasst eine Bibliothek mit verschiedenen konventionellen Algorithmen zur Bildanalyse. Mit diesen wird das Bild zunächst gefiltert und die relative Position der Tasche im Bild bestimmt. Darauf aufbauend können die Positionen von Außenkanten und Stoffbund ermittelt werden, sodass anschließend Abstände und Winkel zwischen diesen Elementen berechnet werden können. Zudem wird mit einem einfachen Mustererkennungswerkzeug überprüft, ob Nähte an einem Rand grundsätzlich vorhanden sind (siehe Abb. 7.7).



Anzahl vorhandener
Nähte

Relativer Winkel
von Stoffbund zu
Außenkante

Taschenhöhe und
Breite

Abb. 7.7: Konventionelle Bildverarbeitung mit Cognex In-Sight Explorer

Nach mehreren Iterationen zur Parameteranpassung kann mit den Werkzeugen Höhe und Breite der Tasche sowie die relative Position des Bundes korrekt extrahiert werden. Anhand dieser Merkmale lässt sich entscheiden, ob eine Tasche beim Nähen korrekt positioniert wurde. Es handelt sich dabei jedoch lediglich um eines von mehreren möglichen Fehlermerkmalen. Das verwendete Mustererkennungswerkzeug erweist sich auch nach mehreren Iterationen der Kalibrierung als fehleranfällig, sodass die Anzahl vorhandener Nähte nicht zuverlässig bestimmt werden kann. Eine Anwendung, welche alle auftretenden Fehlermerkmale verlässlich erkennt, kann nicht unter vertretbarem Zeitaufwand entwickelt werden.

7.4.2 Kombination aus konventioneller Bildverarbeitung und Deep Learning mit Cognex In-Sight Vidi

Im Rahmen einer Machbarkeitsstudie wird der Anwendungsfall vom technischen Kundensupport der Firma Cognex mit der Bildverarbeitungssoftware Cognex In-Sight Vidi untersucht. Die lizenzierte Software beinhaltet sowohl Werkzeuge zur konventionellen Bildverarbeitung als auch verschiedene Deep Learning

Algorithmen. Beide Arten von Algorithmen können dabei über eine grafische Benutzeroberfläche miteinander kombiniert werden. Für den gegebenen Anwendungsfall werden damit folgende Verarbeitungsschritte durchgeführt:

- 1) Mit einem konventionellen Mustererkennungsalgorithmus wird die Tasche im Bild lokalisiert
- 2) Der linke und rechte Randbereich der Armbandtasche werden als separate Arbeitsbereiche definiert.
- 3) Für die beiden definierten Arbeitsbereiche wird ein Deep Learning Algorithmus auf die Nahtüberprüfung antrainiert. Dazu sind vom Anwender als gut und schlecht klassifizierte Trainingsbilder bereitzustellen. Bei jenen Bildern, die als schlecht definiert sind, markiert der Anwender zusätzlich manuell den fehlerhaften Nahtbereich (siehe Abb. 7.8). Durch diese Spezifizierung des relevanten Merkmalsbereichs lässt sich der Trainingsprozess beschleunigen, da es nun weniger Trainingsaufwand für die Ermittlung relevanter Bildbereiche erfordert.

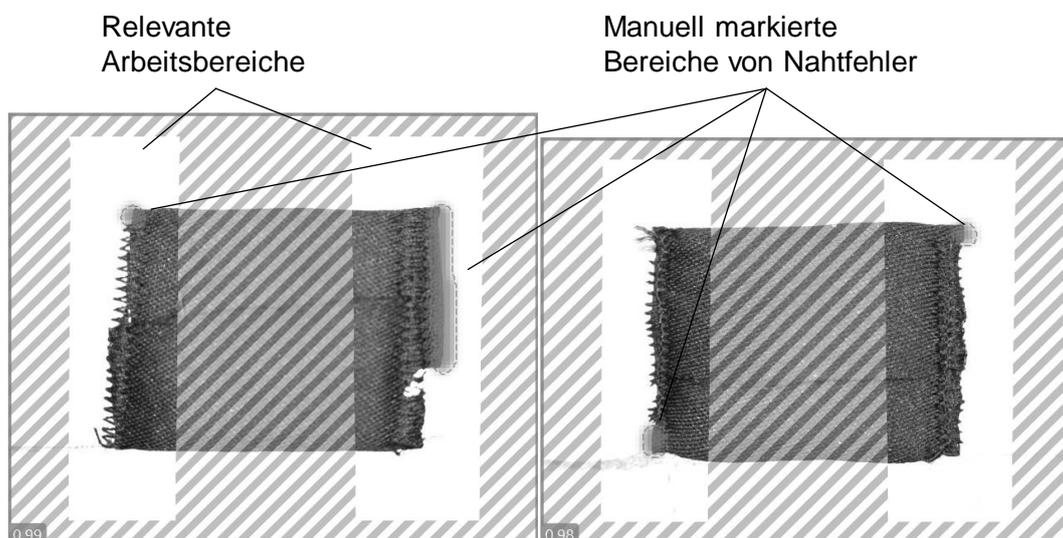


Abb. 7.8: Anlernen fehlerhafter Nahtbereiche mit Cognex In-Sight Vidi

Für den beschriebenen Trainingsprozess werden im Rahmen der Machbarkeitsstudie lediglich 10 als fehlerfrei definierte Bilder sowie 4 Bilder mit Nahtfehlern verwendet. Im Anschluss wird das trainierte Modell mit 184 weiteren Bildern überprüft, von denen 177 richtig als fehlerfrei bzw. fehlerhaft erkannt werden. Dies entspricht einer Klassifizierungsgenauigkeit von 96,2 %. Die Kombination aus konventioneller Bildverarbeitung und Deep Learning erscheint in diesem ersten Machbarkeits-test als eine attraktive Möglichkeit, um eine robuste Klassifikationsanwendung mit geringem Trainingsaufwand zu erhalten.

7.4.3 Deep Learning mit Teachable Machine

In einem dritten Ansatz wird eine Deep Learning Applikation mit Teachable Machine (Version 2.1.2 – Image Project) Cloud Service von Google antrainiert. Hierbei werden lediglich als fehlerfrei bzw. fehlerhaft definierte Bilder vorgegeben, weitere manuelle Vorverarbeitungsmethoden werden nicht benötigt. Um eine direkte Vergleichbarkeit zum vorherigen Ansatz mit Cognex In-Sight Vidi zu ermöglichen, wird das gleiche Trainingsdatensatz aus lediglich 14 Bildern verwendet (Epochen: 50, Learning rate: 0,001). Zum Validieren wird entsprechend der gleiche Validierungsdatensatz wie zuvor verwendet. Erwartungsgemäß wird hier eine geringere Klassifizierungsgenauigkeit von lediglich 76,1 % erreicht.

8 Umsetzung der Auswahllogik als Softwareapplikation

Zur Steigerung der Benutzerfreundlichkeit wird die entworfene Auswahllogik in eine Softwareapplikation überführt. Deren grundlegende Funktionsweise wird in Kapitel 8.1 vorgestellt. Die erstellte Applikation wird anschließend in Kapitel 8.2 verwendet, um den Nutzen der Auswahllogik statistisch zu quantifizieren.

8.1 Funktionsweise der softwaregestützten Auswahllogik

Die in Kapitel 6.1 vorgestellten Entscheidungsregeln zur Auswahl von Hardwarekomponenten werden im Kontext dieser Masterarbeit mittels Python und Jupyter Notebook in eine Webanwendung mit grafischer Benutzeroberfläche überführt. Die Regeln der Entscheidungslogik werden dabei mit der Produktdatenbank für Kameras und Objektive verknüpft, welche im Rahmen der Marktrecherche in Kapitel 4.1 aufgebaut wird.

Gemäß der Entscheidungslogik spezifiziert der Nutzer zunächst seinen Anwendungsfall. Dies ist beispielhaft für die detaillierte Kameraauswahl (siehe Kapitel 6.1.2) in Abb. 8.1 dargestellt.

2) Auswahl Kameramodell

A) Spezifikation Anwendungsfall

1) Wie groß ist das benötigte Sichtfeld in x-Richtung? (Objekt inkl. evtl. Randbereich)

x-Wert [cm]

2) Wie groß ist das benötigte Sichtfeld in y-Richtung?

y-Wert [cm]

3) Wie groß ist das kleinste aufzulösende Detail des Objektes?

Details [mm]

4) Wie viele Bilder müssen maximal pro Sekunde gemacht werden?

Wert [1/s]

5) Bewegt sich das Objekt kontinuierlich fort (Fließmontage)? Falls ja, mit welcher Geschwindigkeit?

Wert [cm/s]

Kontinuierliche Bewegung

6) Bitte wählen Sie die zutreffendste Beschreibung der Lichtverhältnisse der Prüfzene:

Kontrast: Wenig Licht Normal Hoher Kontrast (HD...

Abb. 8.1: Screenshot Auswahltool: Spezifizierung des Anwendungsfalls

In Abhängigkeit der spezifizierten Parameter werden alle geeigneten Kameramodelle in einer grafischen Übersicht visualisiert (siehe Abb. 8.2). Der Nutzer kann hier weitere Präferenzen bzgl. Preis, Hersteller und Gehäusetyp spezifizieren.

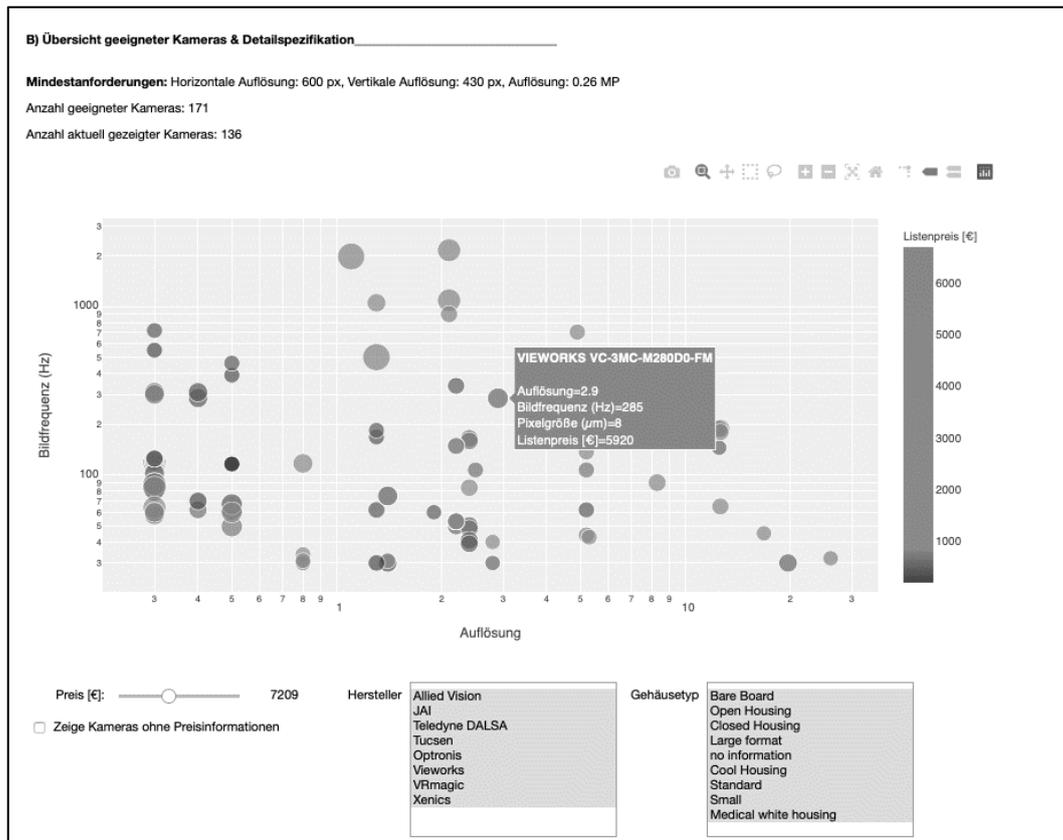


Abb. 8.2: Screenshot Auswahltool: Ergebnisvisualisierung und Spezifizierung weiterer Auswahlparameter

Zudem wird eine tabellarische Übersicht geeigneter Kameras erzeugt (siehe Abb. 8.3). In dieser sind die Kameras nach Listenpreis sortiert, sodass dem Anwender die günstigsten geeigneten Modelle zuerst angezeigt werden. Die Tabelle enthält weitere Produktinformationen sowie einen Link zum Produkt. Darauf aufbauend kann der Anwender sein präferiertes Kameramodell auswählen. Die Eigenschaften des ausgewählten Modells werden bei der anschließenden Objektivauswahl durch die hinterlegte Logik dann automatisch beachtet.

#Resultate 5

Hersteller	Name	Listenpreis [€]	Auflösung	Pixelgröße (µm)	Bildfrequenz (Hz)	Gehäusotyp	Objektivanschluss	Horizontale Aufl. (Px)	Vertikale Aufl. (Px)	Sensortyp	Sensorgroße (mm)	Sensormode
31	Allied Vision AV ALVIUM 1800U-050M-BB	189.0	0.5	4.8	116.0	Bare Board	None	800	600	CMOS	4.8	Python 48
32	Allied Vision AV ALVIUM 1800U-050M-CH-C	219.0	0.5	4.8	116.0	Closed Housing	C-Mount	800	600	CMOS	4.8	Python 48
33	Allied Vision AV ALVIUM 1800U-050M-CH-CS	219.0	0.5	4.8	116.0	Closed Housing	CS-Mount	800	600	CMOS	4.8	Python 48
34	Allied Vision AV ALVIUM 1800U-050M-CH-S	219.0	0.5	4.8	116.0	Open Housing	S-Mount	800	600	CMOS	4.8	Python 48
35	Allied Vision AV ALVIUM 1800U-050M-OH-C	219.0	0.5	4.8	116.0	Open Housing	C-Mount	800	600	CMOS	4.8	Python 48

C) Kameraauswahl

Wählen Sie aus den zur Verfügung stehenden Kameras ein Modell aus, welches für die weiteren Berechnungen verwendet werden soll.

Auswahl:

Abb. 8.3: Screenshot Auswahltool: Tabellarische Übersicht geeigneter Kameras

Weitere Screenshots zur vollständigen Darstellung aller Auswahlsschritte sind Anhang 11.7 zu entnehmen.

8.2 Bewertung des Nutzens der Auswahllogik

Mit dem entwickelten Auswahltool kann abgeschätzt werden, welchen Nutzen die entwickelte Auswahllogik für Anwender hat, die mit den technischen Zusammenhängen von industriellen Kamerasystemen wenig vertraut sind. Dazu werden drei fiktive Anwendungsfälle erstellt, welche geringe (Szenario 1), mittlere (Szenario 2) und hohe technische Anforderungen (Szenario 3) implizieren. Für jeden dieser drei Anwendungsfälle wird bestimmt, wie stark die Modellauswahl an geeigneten Kameras eingeschränkt wird (Komplexitätsreduktion) und welche Kosteneinsparung dabei erzielt wird (siehe Tab. 8.1). Die Komplexitätsreduktion und Kosteneinsparung werden dabei wie folgt berechnet:

$$\text{Komplexitätsreduktion} = 1 - \frac{\text{Anzahl geeigneter Kameras}}{\text{Anzahl Kameras Datenbank}} \quad (8.1)$$

$$\text{Kostenreduktion} = 1 - \frac{\text{Preis günstigste geeignete Kamera}}{\text{Durchschnittspreis aller geeigneten Kameras}} \quad (8.2)$$

Tab. 8.1: Kosteneinsparung und Komplexitätsreduktion der Auswahllogik bei Kameraauswahl

		Szenario 1	Szenario 2	Szenario 3
Spezifikationen	Farbbereich	monochromatisch		
	Sichtfeld x	8 cm	10 cm	20 cm
	Sichtfeld y	4 cm	10 cm	20 cm
	Größe des kleinsten Details	0,25 mm	0,1 mm	0,1 mm
	Bildrate	5	30	30
	Kontinuierlich bewegtes Objekt?	Nein	Ja	Ja
	Lichtverhältnisse Prüfszene	Normal		
	Sonstige Einschränkungen	Gehäuse darf nicht "bare board" oder "open housing" sein		
	Anzahl Kameras Datenbank	702		
	Anzahl geeigneter Kameras	186	34	12
Resultat Logik	Preis der günstigsten geeigneten Kamera	219 €	660 €	4.230 €
	Durchschnittspreis der 5 günstigsten geeigneten Kameras	267 €	1.163 €	6.482 €
	Durchschnittspreis aller geeigneten Kameras	3.298 €	5.402 €	8.465 €
	Prozentuale Komplexitätsreduktion	73,5%	95,2%	98,3%
	Prozentuale Kostenreduktion	93,4%	87,8%	50,0%

In Szenario 1 sind die Spezifikationen des Anwendungsfall einfach gehalten, da Sichtfeld und benötigte Bildrate relativ klein und die Größe des kleinsten aufzulösenden Details hoch sind. Entsprechend sind die resultierenden technischen Anforderungen gering. Aufgrund des monochromatischen Farbbereichs kann zwar eine Vielzahl von Kameras ausgeschlossen werden, die Komplexitätsreduktion ist mit 73,5 % aber noch vergleichsweise gering. Viele der verwendbaren Kameras sind für den Anwendungsfall jedoch überdimensioniert, sodass sich erhebliche Preisunterschiede ergeben. Anhand der hinterlegten Listenpreise kann dem Anwender die günstigste unter der Vielzahl an geeigneten Optionen empfohlen werden. Der Mehrwert für den Anwender wird hierbei also primär durch die Preistransparenz des Auswahltools generiert. Mit zunehmender Komplexität des Anwendungsfalls steigt die durch das Auswahltool erzielte Komplexitätsreduktion, die prozentuale Kosteneinsparung nimmt hingegen ab. In solchen Fällen wird der Mehrwert für den Anwender primär durch die hinterlegte technische Auswahllogik und weniger durch die Preistransparenz generiert.

Die günstigste Kamera aus Szenario 2 wird weiterhin verwendet, um die Komplexitäts- und Kostenreduktion für die anschließende Objektivauswahl zu ermitteln. Hierbei wird insbesondere die Komplexität der Auswahl deutlich reduziert (siehe Tab. 8.2).

Tab. 8.2: Kosteneinsparung und Komplexitätsreduktion der Auswahllogik bei Objektivauswahl

	Festbrennweiten- objektiv	Telezentrisches Objektiv
Arbeitsabstand Objektiv	20 - 80 cm	
Relevanter Tiefenbereich	5 mm	
Sonstige Einschränkungen	Keine	
Anzahl vorhandener Objektive	316	464
Anzahl geeigneter Objektiv	20	8
Preis der günstigsten Empfehlung	333 €	1.032 €
Durchschnittspreis der 5 günstigsten Empfehlungen	482 €	1.460 €
Durchschnittspreis aller geeigneten Empfehlungen	637 €	1.897 €
Prozentuale Komplexitäts- reduktion	94%	98%
Prozentuale Kostenreduk- tion	48%	46%

9 Zusammenfassung

Qualitätskontrollen mittels industrieller Bildverarbeitung ermöglichen produzierenden Unternehmen die Senkung von qualitätsbezogenen Kosten. Technische Lösungen werden dabei selten von der Stange gekauft, sondern müssen für optimale Ergebnisse spezifisch auf den jeweiligen Anwendungsfall zugeschnitten werden. [Sac12] Marktseitig existiert daher eine hohe Anzahl an verschiedenen Produkten, die modular miteinander kombiniert werden können. Diese Produktkomplexität ist für Anwender, welche mit der Thematik wenig vertraut sind, schwer zu erfassen und erschwert die effiziente Auswahl und Implementierung einer Bildverarbeitungsapplikation.

In der vorliegenden Arbeit wird daher eine Logik zur Anwendungsfall-spezifischen Auswahl von Bildverarbeitungstechnologien entwickelt. Konkret werden dabei die folgenden beiden Technologiebereiche thematisiert:

- 1) **Hardware:** Wie lassen sich eine geeignete Kamera, zugehöriges Objektiv und Beleuchtung eines Bildverarbeitungssystems in Abhängigkeit des Anwendungsfalls bestimmen? (Forschungsfrage 1)
- 2) **Software:** Wie lässt sich in Abhängigkeit des Anwendungsfalls entscheiden, ob für die Bildverarbeitung konventionelle Algorithmen oder Deep Learning verwendet werden soll? (Forschungsfrage 2)

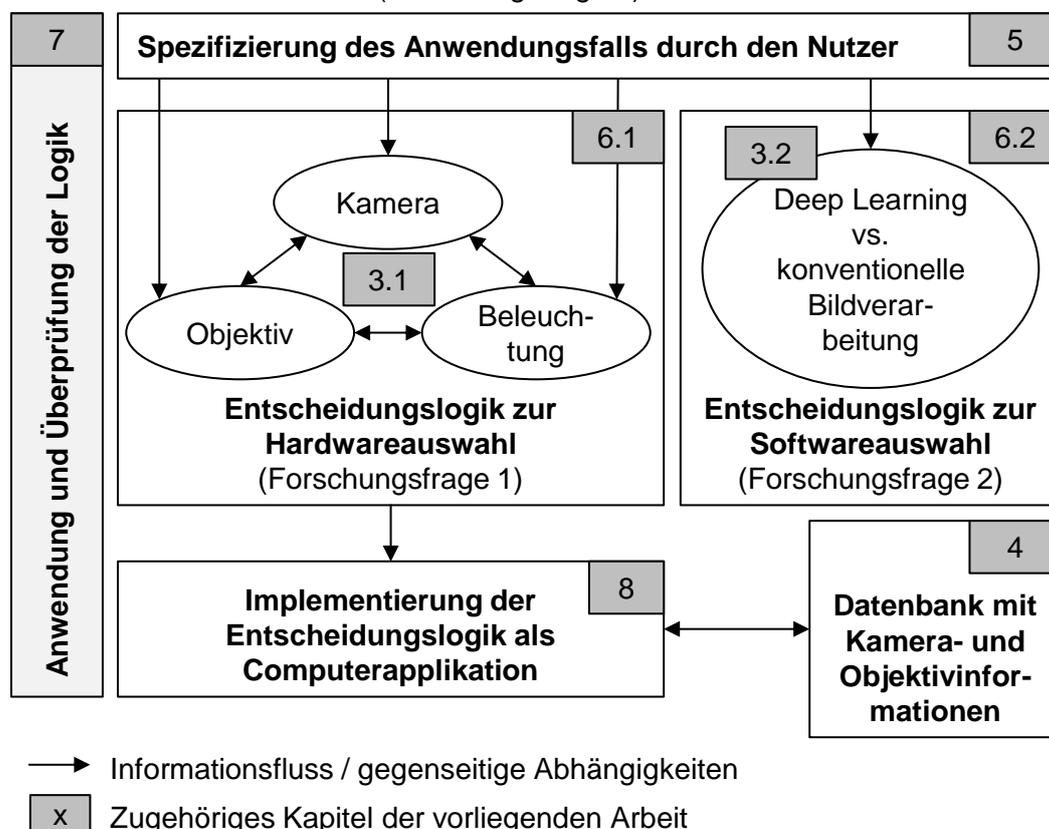


Abb. 9.1: Entwickeltes und angewandtes Vorgehen zur Technologieauswahl

Zur Beantwortung dieser Fragestellungen wird in Kapitel 3 zunächst eine Einführung in die zugrundeliegenden Technologien gegeben und Beschreibungsparameter der Literatur zur Technologiespezifikation ausgearbeitet. Als Zwischenergebnis wird dabei eine detaillierte Übersicht der Interdependenzen der Hardwaretechnologien entwickelt. Auf der Softwareseite wird eine klare Abgrenzung von Deep Learning und konventionellen Algorithmen vorgenommen sowie eine Übersicht konventioneller Algorithmen erstellt.

In Kapitel 4 werden im Rahmen einer Marktrecherche die für die Entscheidungslogik relevanten Beschreibungsparameter eingegrenzt und eine Produktdatenbank mit 1482 Hardwarekomponenten (Kameras und Objektive) aufgebaut.

Auf diesen Literatur- und Marktrecherchen sowie einer Vorlage zur standardisierten Charakterisierung von Anwendungsfällen (Kapitel 5) baut die anschließend entwickelte Entscheidungslogik auf (Kapitel 6). Zur Beantwortung von Forschungsfrage 1 wird ein fünfteiliges Vorgehen zur Bestimmung von Anwendungsfall-spezifischer Hardware (Kamera, Objektiv, Beleuchtungssystem) entworfen. Die Entscheidung zwischen Deep Learning und konventionellen Algorithmen (Forschungsfrage 2) kann anhand der gegebenen Entwicklungsumstände sowie der Komplexität von Bildmerkmalen getroffen werden. Die entwickelte Logik wird in Kapitel 7 anhand eines Anwendungsfalls zur Qualitätsüberprüfung von vernähten Armbändern erfolgreich erprobt und veranschaulicht.

Die Hardware-bezogene Entscheidungslogik wird in Kapitel 8 zusammen mit der Produktdatenbank in eine benutzerfreundliche Computerapplikation überführt und für eine simulative Bewertung des generierten Nutzens verwendet. Dabei werden mit der entworfenen Entscheidungslogik in Abhängigkeit des Anwendungsfalls Komplexitätsreduktionen zwischen 73 und 98 % und Kosteneinsparungen zwischen 46 und 93 % erzielt.

10 Ausblick

Aufbauend auf den gesammelten Erkenntnissen bei der Erstellung und Anwendung der Entscheidungslogik bieten sich folgende Möglichkeiten zur Weiterentwicklung des Vorgehens an.

Erweiterung der Produktdatenbank

Nicht alle der zu Beginn definierten Parameter zur Beschreibung von Hardwaretechnologien konnten vollständig in die Auswahllogik eingebunden werden, da diese Produktinformationen im Rahmen der Marktrecherche nicht einfach zugänglich waren. Mit einer aufwendigeren Marktrecherche könnten zusätzliche Produktinformationen eingeholt und in die Produktdatenbank der Auswahllogik integriert werden.

Ausdetaillierung der Entscheidungslogik zur Auswahl geeigneter Software

Im Rahmen dieser Arbeit wird eine Logik zur grundsätzlichen Entscheidung über die Verwendung von Deep Learning vs. konventionellen Algorithmen zur Bildverarbeitung erarbeitet. Die Logik reicht jedoch noch nicht aus, um dem Anwender in Abhängigkeit seines Inspektionsproblems konkrete Algorithmen bzw. ein konkretes Softwareprodukt zu empfehlen. Mit der Übersicht zu konventionellen Algorithmen und der Software-Marktrecherche werden in der vorliegenden Arbeit dazu bereits hilfreiche Grundlagen ausgearbeitet.

Einbettung der Entscheidungswahllogik in eine Investitionsrechnung

Die bisherige Auswahllogik findet unter der Prämisse statt, dass ein kamerabasiertes Inspektionssystem für die spezifizierte Anwendung grundsätzlich geeignet ist. In der Praxis ist jedoch zu prüfen, ob die Aufgabe mit anderen Verfahren und Technologien (z. B. einer manuellen Sichtkontrolle) nicht kostengünstiger zu implementieren ist. Dazu werden Methoden der Investitionsrechnung herangezogen. Da in der Datenbank der entworfenen Computerapplikation Listenpreise bereits hinterlegt sind, könnte die Anwendung um eine Funktion zur Berechnung von Amortisationszeiten erweitert werden.

Literaturverzeichnis

- [AEJ18] N. N.: AIA; EMVA; JIA (Hrsg.):
Guide to understanding machine vision standards
Michigan; Barcelona; Tokyo: AIA; EMVA; JIA, 2018
[Firmenschrift]
- [AMR04] Abràmoff, M.; Magalhães, P.; Ram, S.:
Image processing with ImageJ
Biophotonics International 7 (2004), H. 11, S. 36 - 42
- [BK19] Bhanare, H.; Khachane, A.:
Quality Inspection of Tire using Deep Learning based Computer
Vision
International Research Journal of Engineering and Technology 6
(2019), H. 11, S. 3555 - 3558
- [Bur19] Burkov, A.:
The hundred-page machine learning book
1. Aufl. - Quebec City: Andriy Burkov, 2019
- [Cas14] Cass, S.:
The top 10 programming languages spectrum's 2014 ranking
IEEE Spectrum 51 (2014), H. 7, S. 68
- [CCS18] N. N.: CCS Inc. (Hrsg.):
Ring Lights LDR2 series
Kyoto: CCS Inc., 2018
[Produktdatenblatt]
- [Cog18] N. N.: Cognex Corporation (Hrsg.):
Deep Learning für die automatische Fertigung
Natick: Cognex Corporation, 2018
[Firmenschrift]
- [Cog19] N. N.: Cognex Corporation (Hrsg.):
Deep Learning im Vergleich zur industriellen Bildverarbeitung
Natick: Cognex Corporation, 2019
[Firmenschrift]
- [CPC17] Canziani, A.; Paszke, A.; Culurciello, E.:
An analysis of deep neural network models for practical applica-
tions (2017)

- [CPS10] Chaki, A.; Prashant, M.; Sen, P.:
A Comprehensive Market Analysis on Camera and Illumination
Sensors for Image
IEEE Computer Society (Hrsg.):
2010 International Conference on Computational Intelligence and
Communication Networks, Bhopal, 26.-27. November 2010
New York: IEEE, 2011
- [Daw14] Dawson-Howe, K.:
A practical introduction to Computer Vision with OpenCV.
1. Aufl. – New York: Wiley, 2014
- [DIN
60529] DIN EN 60529
Schutzarten durch Gehäuse
- [DLV18] Döbel, I.; Leis, M.; Vogelsang, M.:
Maschinelles Lernen – Kompetenzen, Anwendungen und For-
schungsbedarf
Berlin: Fraunhofer Gesellschaft, 2018
[Forschungsbericht]
- [DSS11] Demant, C.; Streicher-Abel, B.; Springhoff, A.:
Industrielle Bildverarbeitung - Wie optische Qualitätskontrolle
wirklich funktioniert.
3. Aufl. – Heidelberg; Berlin: Springer., 2011
- [Edm14] N. N.: Edmund Optics (Hrsg.):
Techspec compact fixed focal length lense
Mainz: Edmund Optics, 2014
[Produktdatenblatt]
- [EMVA
1288] EMVA Standard 1288 Release 3.1
Standard for Characterization of Image Sensors and Cameras
- [Ert13] Ertel, W.:
Grundkurs künstliche Intelligenz – Eine praxisorientierte Einfüh-
rung.
3. Aufl. - Wiesbaden: Springer, 2013
- [Fer19] Fermum, L.:
Telezentrische Objektive. – München: Vision Doctor, 2019, URL:
<https://www.vision-doctor.com/telezentrische-objektive.html>,
Zugriff am 19.12.2019

- [Fer20] Fermum, L.:
Beleuchtung: Merkmale & Fehler sichtbar machen. – München:
Vision Doctor, 2020, URL: <https://www.vision-doctor.com/farbiges-licht-farbige-koerper.html>, Zugriff am 12.04.2020
- [FMM11] Fernandes, A.; Moreira, L.; Mata, J.:
Machine Vision Applications and Development Aspects
IEEE Control Systems Society (Hrsg.):
9th IEEE Conference on Control and Automation, Santiago,
19.-21. Dezember 2011
New York: IEEE, 2011
- [GBC16] Goodfellow, I.; Bengio, Y.; Courville, A.:
Deep learning.
1. Aufl. - Cambridge, Massachusetts: The MIT Press, 2016
- [Gvr18] N. N.: Grand View Research (Hrsg.):
Machine Vision Market Analysis Report by Offering, by Application, by Product, by End Use, and Segment Forecasts, 2018 - 2025
San Francisco: Grand View Research, 2018
[Marktforschungsanalyse]
- [GW08] Gonzalez, R.; Woods, R.:
Digital Image Processing.
3. Aufl. – Upper Saddle River: Pearson, 2008
- [Haf02] Haferkorn, H.:
Optik – Physikalisch-technische Grundlagen und Anwendungen.
4. Aufl. – Weinheim: Wiley-VCH, 2002
- [HCG+06] Hudson, E.; McCormick, J.; Genise, R.; Dahl, J.:
Smart Camera. US Patent 6 988 008, Veröffentlichungstag
17.01.2006
- [Hec05] Hecht, E.:
Optik.
4. Aufl. – München; Wien: Oldenbourg, 2005
- [Hen08] Henzler, T.:
Richtig beleuchtet, richtig erkannt – Beleuchtungstechniken für
die industrielle Bildverarbeitung
Optik & Photonik (2008), H. 3, S. 34 - 37

-
- [HKT07] Hain, R.; Kähler, C.; Tropea, C.:
Comparison of CCD, CMOS and intensified cameras
Experiments in Fluids (2007), H. 3, S. 403 - 411
- [Hol10] Holst, G.:
Wie groß darf's sein? – Pixelgröße beeinflusst Bildqualität und
Empfindlichkeit von Bildsensoren
Optik & Photonik (2010), H. 3, S. 37 - 39
- [Jah06] Jahr, I.:
Lighting in Machine Vision
In Hornberg, A.:
Handbook of Machine Vision. 1. Aufl.- Weinheim: Wiley-VCH,
2006, S. 73-204
- [Jäh12] Jähne, B.:
Digitale Bildverarbeitung und Bildgewinnung.
7. Aufl. – Heidelberg; Berlin: Springer-Verl., 2012
- [Jer09] Jeremias, R.:
CMOS Bildsensoren mit Kurzzeitverschluss zur Tiefenerfassung
nach dem Lichtlaufzeit-Meßprinzip, Duisburg, Universität Duis-
burg-Essen, Dissertationsschrift, 2009
- [JIIA LI-
001] JIIA LI-001-2018
Lighting for machine vision / Image Processing System – Funda-
mentals of Design and Specification of Brightness of Optical Irridi-
ation
- [JIIA LI-
002] JIIA LI-002-2017
Light Performance Specifications
- [KBV14] Khan, T.; Bodrogi, P.; Vinh, Q.:
LED Lighting: Technology and Perception.
1. Aufl. – Weinheim: Wiley-VCH, 2014
- [Klo19] Klostermann, A.:
Machine Learning Landscape: An ML showdown in search for the
best tool. – Medium, 2019, URL: [https://medium.com/@Aiko-
Path/machine-Learning-landscape-an-ml-showdown-in-search-of-
the-best-tool-9db2345861fb](https://medium.com/@Aiko-Path/machine-Learning-landscape-an-ml-showdown-in-search-of-the-best-tool-9db2345861fb), Zugriff am 06.3.2020

- [KST90] Kitamura, A.; Sato, H.; Tamura, H.:
An expert system for industrial machine vision
IEEE (Hrsg.):
Proceedings of the 10th International Conference on Pattern
Recognition, Atlantic City, 16.-21. Juni 1990
New York: IEEE, 1990
- [KUB14] Kuther, M.; Hülck, S.:
Megapixelobjektive. – Würzburg: Elektronik Praxis, 2014, URL:
<https://www.elektronikpraxis.vogel.de/welche-aufloesung-die-bv-anwendung-wirklich-braucht-a-443804/index2.html>, Zugriff am
14.1.2020
- [KUH14] Kucera, B.; Brückner, C.:
LEDs. – Würzburg: Elektronik Praxis, 2014, URL: <https://www.elektronikpraxis.vogel.de/so-erhalten-sie-eine-hoehere-lichtausbeute-bei-geringeren-kosten-a-463241/>, Zugriff am
17.1.2020
- [LCB09] Lai, Y.; Cordero, N.; Barthel, F.:
Liquid cooling of bright LEDs for automotive applications
Applied Thermal Engineering Vol. 29 (2009), H. 5, S. 1239 - 1244
- [Len06] Lenhardt, K.:
Optical Systems in Machine Vision
In Hornberg, A.:
Handbook of Machine Vision. 1. Aufl.- Weinheim: Wiley-VCH,
2006, S. 205-332
- [LGT+97] Lawrence, S.; Giles, C.; Tsoi, A.; Back, A.:
Face Recognition: A convolutional Neural-Network Approach
IEEE Transactions on Neural Networks 8 (1997), H. 1, S. 98 -
113
- [Lic18] N. N.: Licht.de (Hrsg.):
LED: Grundlagen – Applikation - Wirkung
Frankfurt: licht.wissen, 2018
[Firmenschrift]
- [LRS10] Leier, C.; Rivest, R.; Stein, C.:
Algorithmen – Eine Einführung.
1. Aufl. – München: Oldenbourg, 2010

-
- [LUW13] Lin, W.; Uang, C.; Wang, P.:
LED strobe lighting for Machine Vision Inspection
IEEE Computer Society (Hrsg.):
2nd International Symposium on Next-Generation Electronics,
Kaohsiung, 25.-26. Februar 2013
New York: IEEE, 2013
- [Mat06] Mattfeldt, H.:
Camera Systems in Machine Vision
In Hornberg, A.:
Handbook of Machine Vision.
1. Aufl. – Weinheim: Wiley-VCH, 2006, S. 361-426
- [MCC+19] Mahony, N.; Campell, S.; Carvalho, A.; Harapanahalli, S.; Hernandez, G.:
Deep Learning vs. Traditional Computer Vision
Arai, K.; Kapoor, S. (Hrsg.):
Proceedings of the 2019 Computer Vision Conference (CVC),
Las Vegas, 25.-26. April 2019
New York: Springer, 2019
- [MPZ03] Malamas, E.; Petrakis, E.; Zervakis, M.:
A survey on industrial vision systems, applications and tools
Image and Vision Computing (2003), H. 2, S. 171 - 188
- [MSL+03] Müller, S.; Stricker, H.; Lindemann, U.; Zäh, M:
Kostenoptimale Technologieauswahl in der Montagesystem-
planung
ZWF 98 (2003), H. 9, S. 431 - 435
- [Neu05] Neumann, B.:
Bildverarbeitung für Einsteiger.
1. Aufl. – Berlin; Heidelberg: Springer, 2005
- [NFH07] Nischwitz, A.; Fischer, M.; Haberäcker, P.:
Computergrafik und Bildverarbeitung.
2. Aufl. – Wiesbaden: Vieweg, 2007
- [NH09] Nasse, H.; Hönlinger, B.:
Verzeichnungen. – Oberkochen: Carl Zeiss AG, 2009, URL:
<https://lenspire.zeiss.com/photo/app/uploads/2018/12/Distortion-2009-DE.pdf>, Zugriff am 13.1.2020

- [Nov91] Novini, R.:
Lighting and optics expert system for machine vision
SPIE (Hrsg.):
Proc. SPIE 1386: Machine vision systems integration in industry,
Boston, 1990
Bellingham: SPIE, 1991
- [NVK+15] Najafabadi, M; Villanustre, F.; Khoshgoftaar, T; Wald, R.:
Deep learning applications and challenges in big data analytics
Journal of Big Data 2 (2015), H. 1, S. 1-21
- [Par94] Parker, J.:
Practical Computer Vision using C.
1. Aufl. – New York: Wiley, 1994
- [PBK+12] Pulli, K.; Baksheev, A.; Korniyakov, K.; Eruhimov, V.:
Real-time computer vision with OpenCV
Communications of the ACM 55 (2012), H. 6, S. 61 - 69
- [Per10] Perrot, S.:
Auswahl geeigneter Beleuchtung für die Anwendungen in der in-
dustriellen Bildverarbeitung
Optik & Photonik (2010), H. 3, S. 51 - 55
- [PKB13] Pervin, N.; Kumar, D.; Bhardwaj, I.:
An overview of template matching methodologies and its applica-
tions
International Journal of Research in Computer and Communica-
tion Technology (2013), H. 10, S. 988 - 995
- [RRK89] Ruck, D.; Rogers, S.; Kabrisky, M.:
Feature Selction using a Multilayer Perceptron
Journal of Neural Network Computing 2 (1989), H. 2, S. 40 - 48
- [Rus07] Russ, J.:
The Image Processing Handbook.
5. Aufl. – Boca Raton: CRC Press, 2007
- [Sac12] Sackewitz, M.:
Leitfaden zur industriellen Bildverarbeitung.
3. Aufl. – Stuttgart: Fraunhofer, 2012
- [SHB99] Sonka, M.; Hlavac, V.; Boyle, R.:
Image Processing, Analysis, and Machine Vision.
2. Aufl. – Pacific Grove: PWS Publishing, 1999

-
- [SHR08] Sonka, M.; Hlavac, V.; Roger, B.:
Image Processing, Analysis, and Machine Vision.
3. Aufl. – Stamford: Cengage Learning, 2008
- [Sil18] N. N.: Sill Optics GmbH & Co. KG (Hrsg.):
Machine Vision
Wendelstein: Sill Optics, 2018
[Firmenschrift]
- [SLF18] Staar, B.; Lütjen, M.; Freitag, M.:
Anomaly detection with convolutional neural networks for industrial surface inspection
ScienceDirect (Hrsg.):
12th CIRP Conference on Intelligent Computation in Manufacturing Engineering, Gulf of Naples, 18.-20. Juli 2018
New York: CIRP, 2018
- [SRE12] Schneider, C.; Rasband, W.; Eliceiri, K.:
NIH Image to ImageJ: 25 years of image analysis
Nature Methods 9 (2012), H. 7, S. 671
- [SSP+18] Sustika, R.; Subetki, A.; Pardede, H.; Suryawati, E.; Mahendra, O.; Yuwana, S.:
Evaluation of Deep Convolutional Neural Network Architectures for Strawberry Quality Inspection
International Journal of Engineering & Technology (2018), H. 7, S. 75 - 80
- [ST02] Schröder, G.; Treiber, H.:
Technische Optik.
9. Aufl. – Würzburg: Vogel, 2002
- [Ste06] Steger, C.:
Machine Vision Algorithms
In Hornberg, A.:
Handbook of Machine Vision. 1. Aufl.- Weinheim: Wiley-VCH, 2006, S. 511-692
- [Ste18] N. N.: Stemmer Imaging AG:
Handbuch der Bildverarbeitung
Puchheim: Stemmer Imaging, 2018
[Firmenschrift]

- [Ste20] N. N.: Stemmer Imaging:
Über Stemmer Imaging. – Graz: Stemmer Imaging, 2020, URL:
<https://www.stemmer-imaging.com/de-at/ueber-stemmer-imaging/>, Zugriff am 01.3.2020
- [SUW18] Steger, C.; Ulrich, M; Wiedemann, C.:
Machine Vision Algorithms and Applications.
2. Aufl. – Würzburg: Vogel, 2018
- [Tel06] Telljohan, A.:
Introduction to Building a Machine Vision System
In Hornberg, A.:
Handbook of Machine Vision. 1. Aufl.- Weinheim: Wiley-VCH,
2006, S. 35-72
- [Thö04] Thöniß, T.:
Abbildungsfehler und Abbildungsleistung optischer Systeme. –
Göttingen: PhotonikNet, 2004, URL: https://www.researchgate.net/profile/Thomas_Thoeniss/publication/273697559_Abbildungsfehler_und_Abbildungsleistung_optischer_Systeme/links/5509590a0cf27e990e0e59e4.pdf, Zugriff am
14.1.2020
- [VCV11] Van Der Walt, S., Colbert, S.; Varoquaux, G.:
The NumPy array: a structure for efficient numerical computation
Computing in Science & Engineering 13 (2011), H. 2, S. 22
- [VDI 2632-1] VDI / VDE 2632 Blatt 1
Industrielle Bildverarbeitung – Grundlagen und Begriffe
- [VDI 2632-2] VDI / VDE 2632 Blatt 2
Industrielle Bildverarbeitung – Leitfaden für die Erstellung eines
Lastenhefts und eines Pflichtenhefts
- [VDI 3694] VDI / VDE 3694
Lastenheft / Pflichtenheft für den Einsatz von Automatisierungssystemen
- [Vdm16] N. N.: VDMA Robotics + Automation Association:
Machine Vision 2017/18 – key technology for automation solutions
Frankfurt am Main: VDMA Verlag, 2016
[Firmenschrift]

-
- [War17] Wartala, R.:
Praxiseinstieg Deep Learning.
1. Aufl. – Heidelberg: O'Reilly, 2017
- [Was06] Waszkewitz, P.:
Machine Vision in Manufacturing
In Hornberg, A.:
Handbook of Machine Vision. 1. Aufl.- Weinheim: Wiley-VCH,
2006, S. 693-781
- [Web20] N. N.: Web Scraper:
Making web data extraction easy and accessible for everyone. –
Riga: Web Scraper, 2020, <https://webscraper.io/>, Zugriff am
7.5.2020
- [WFH+17] Witten, I.; Frank, E.; Hall, M.; Pal, C.:
Data Mining – Practical Machine Learning Tools and Techniques.
4. Aufl. - Cambridge: Kaufmann, 2017
- [WMZ+18] Wang, J.; Ma, Y.; Zhang, L.; Gao, R.; Wu, D.:
Deep learning for smart manufacturing: Methods and applications
Journal of Manufacturing Research (2018), H. 46, S. 144 – 156
- [WSS16] Weimer, D.; Scholz-Reiter, B.; Shpiltani, M.:
Design of deep convolutional neural network architectures for au-
tomated feature extraction in industrial inspection
CIRP Annals 65 (2016), H. 1, S. 417 - 420
- [YG99] Yang, D.; Gamal, A.:
Comparative analysis of SNR for image sensors with enhanced
dynamic range
SPIE (Hrsg.):
Proc. SPIE 3649: Sensors, cameras, and systems for scientific /
industrial application, San Jose, 23.-29. Januar 1999
Bellingham: SPIE, 1999
- [YH05] Yam, F.; Hassan, Z.:
Innovative advances in LED technology
Microelectronics Journal 36 (2005), H. 36, S. 129 - 137
- [ZZX+19] Zhao, Z.; Zheng, P.; Xu, S.; Xindong, W.:
Object Detection with Deep Learning: A Review
IEEE Transactions on Neural Networks and Learning Systems 30
(2019), H. 11, S. 3212 - 3232

11 Anhang

11.1 Beschreibungsparameter von Hardwarekomponenten

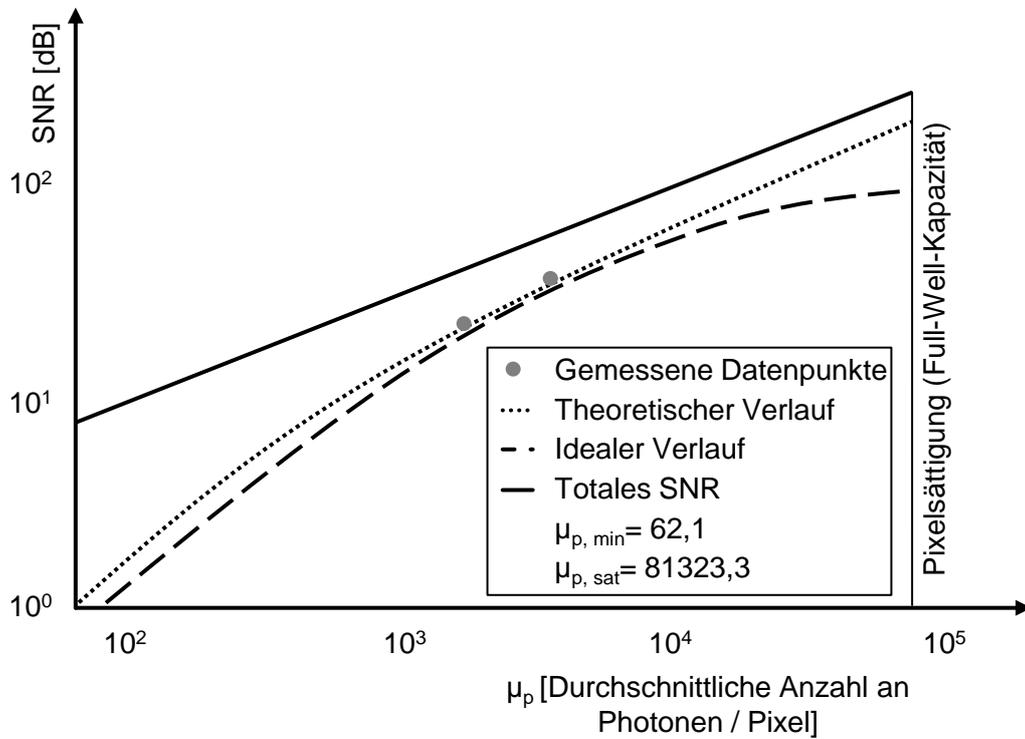


Abb. 11.1: Beispielhafter Kurvenverlauf eines SNR Graphen i. A. a.
[EMVA 1388]

11.2 Übersicht konventioneller Bildverarbeitungsalgorithmen

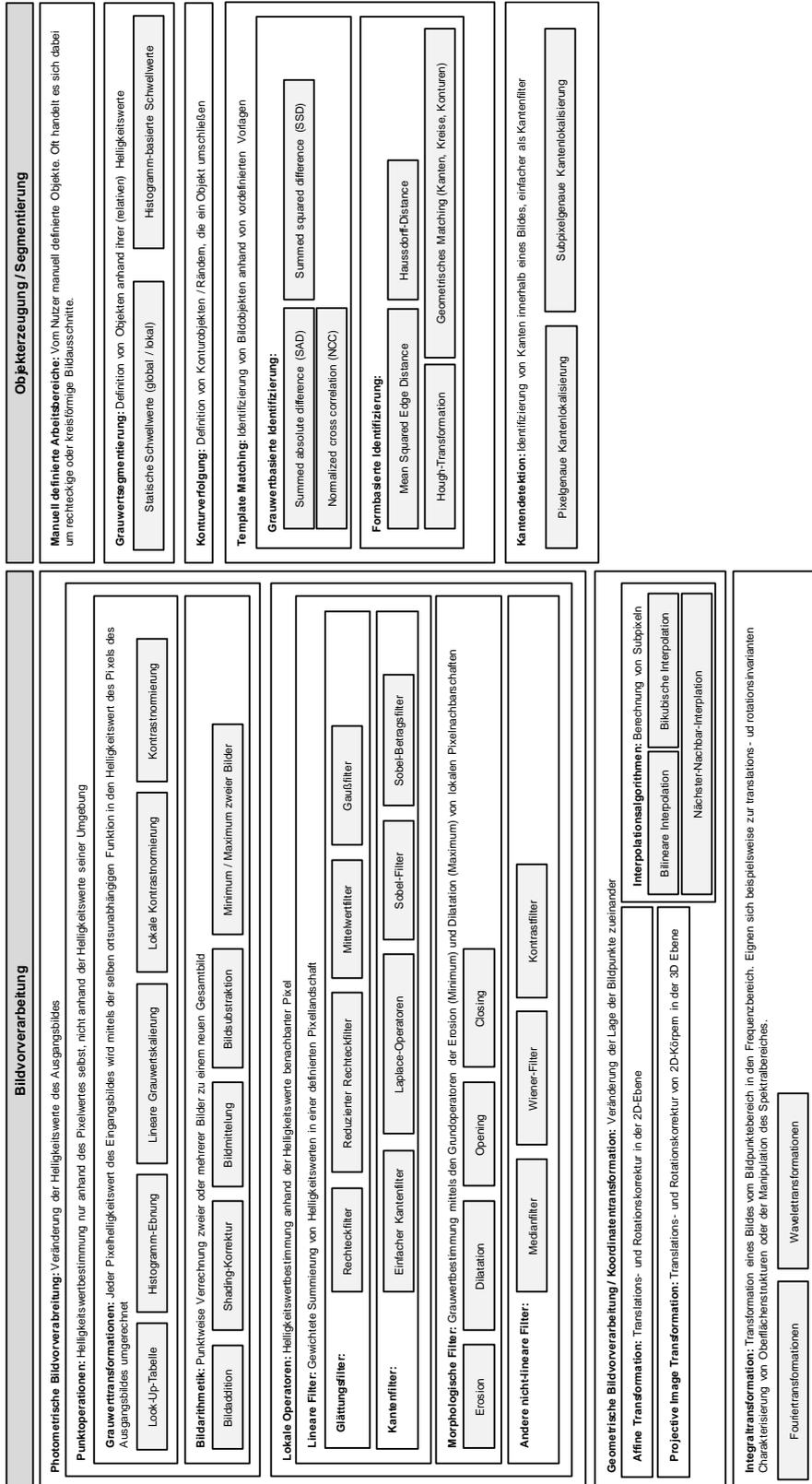


Abb. 11.2: Übersicht konventioneller Bildverarbeitungsalgorithmen Teil 1 (Eigendarstellung i. A. a. [DSS11; SUW18; PKB13; Ste06])

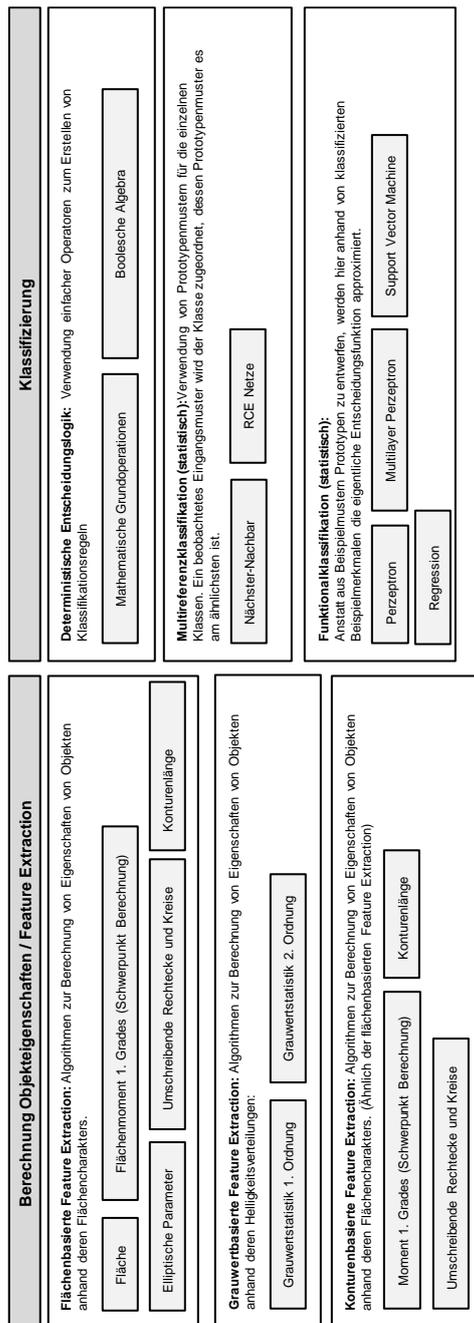


Abb. 11.3: Übersicht konventioneller Bildverarbeitungsalgorithmen Teil 2 (Eigendarstellung i. A. a. [DSS11; SUW18; PKB13])

11.3 Übersicht erfolgreicher Netzarchitekturen der ImageNet Challenge

Tab. 11.1: Übersicht erfolgreicher Architekturen der Imagenet Challenge [CPC17]

Architektur	Lernmethode	Kostenfunktion	Softmax Layer	End-to-End Training	Plattform	Sprache
R-CNN	SGD, BP	Hinge loss (classification), Bounding box regression	Ja	Nein	Caffe	Matlab
SPP-net	SGD	Hinge loss (classification), Bounding box regression	Ja	Nein	Caffe	Matlab
Fast RCNN	SGD	Class Log loss + Bounding box regression	Ja	Nein	Caffe	Python
Faster R-CNN	SGD	Class Log loss + Bounding box regression	Ja	Ja	Caffe	Python / Matlab
R-FCN	SGD	Class Log loss + Bounding box regression	Nein	Ja	Caffe	Matlab
Mask R-CNN	SGD	Class Log loss + Bounding box regression + Semantic sigmoid loss	Ja	Ja	TensorFlow / Keras	Python
FPN	Synchronized SGD	Class Log loss + Bounding box regression	Ja	Ja	TensorFlow	Python
YOLO	SGD	Class sum-squared error loss + bounding box regression + object confidence + background confidence	Ja	Ja	Darknet	C
SSD	SGD	Class softmax loss + bounding box regression	Nein	Ja	Caffe	C++
YOLOv2	SGD	Class sum-squared error loss + bounding box regression + object confidence + background confidence	Ja	Ja	Darknet	C

11.4 Anwendung der Auswahllogik, Schritte 1 und 2

Tab. 11.2: Anwendung der Hardware-Auswahllogik, Schritte 1 und 2

	Frage / Kriterium gemäß Auswahllogik	Antwort	Anmerkung
Kameravorauswahl	Sollen Tiefeninformationen von 3D-Objekten erfasst werden?	Nein	Alle relevanten Informationen im 2D-Sichtfeld erkennbar.
	Ist das zu überprüfenden Objekt kontinuierlicher Art (Rollenware wie z. B. Folien oder Bleche) und wird mit hoher Geschwindigkeit im Sichtfeld der Kamera bewegt?	Nein	Armband wird vom Werker zur visuellen Kontrolle in vorgesehene Einrichtung eingespannt.
	Ist das zu überprüfenden Objekt zylindrischer Form?	Nein	
	Handelt es sich um auf einem Förderband bewegte Objekte?	Nein	
	Sind die zu überprüfenden Merkmale innerhalb des menschlich sichtbaren Lichtbereiches identifizierbar?	Ja	
	Bedarf es zur Bildanalyse Farbinformationen oder sind monochromatische Bilder (schwarz-weiß) ausreichend?	Monochromatisch	Qualitätsurteil bedarf keiner speziellen Farbinformationen, Nähte sind auch in Graustufen auf dunklem Hintergrund erkennbar.
	Resultat Logik:	Matrixkamera (2D), menschlich sichtbarer Spektralbereich, monochromatisches Bild	
Detaillierte Kameraauswahl	Benötigte Sichtfeldgröße, horizontal	6 cm	Armbandtasche ist etwa 3,5 cm breit. Das Sichtfeld muss jedoch größer sein, da die Armbandposition aufgrund des manuellen Einsetzens in die Spannvorrichtung leicht variieren kann.
	Benötigte Sichtfeldgröße, horizontal	4,3 cm	Armbandtasche ist etwa 2,7 cm hoch. Das Sichtfeld muss jedoch größer sein, da die Armbandposition aufgrund des manuellen Einsetzens in die Spannvorrichtung leicht variieren kann.
	Benötigte minimale Detailgröße	0,2 mm	Der Durchmesser des zum Vernähen verwendeten Fadens beträgt ca. 0,2 mm.
	Aufzunehmende Bilder pro Sekunde	< 1 Bild / s	Bildaufnahme wird manuell vom Werker ausgelöst. Einsetzen, auslösen und austauschen des Armbandes dauert mindestens 5 Sekunden.
	Bewegt sich das Objekt während der Bildaufnahme?	Nein	s.o.
	Bedarf es eines besonders hohen Kontrastes oder liegen schlechte Lichtverhältnisse vor?	Nein	Es soll eine externe Beleuchtung für konstante Belichtungsverhältnisse verwendet werden.
	Präferenzen bzgl. Datenschnittstelle	Industrial Ethernet	Kamerasystem soll mit intern verwendetem Datenanalyse-Rechner (ia: factory cube) kompatibel sein, welcher über Ethernet und I/O-Link Schnittstellen verfügt.
	Gewünschter Schutzgrad	Schutz gegen Staub und Spritzwasser	Produktionsbedingungen unterliegt keinen besonderen Verunreinigungen. Aufgrund von Demonstrationszwecken soll jedoch ein industrietypischer Schutzgrad eingehalten werden.
	Eingrenzung geeigneter Hersteller aufgrund sonstiger Präferenzen	Cognex	Für eine möglichst anwenderfreundliche Integration werden Hersteller bevorzugt, die ein Gesamtpaket aus notwendiger Hardware und kompatibler Software bieten. Der Hersteller Cognex ist für seine Komplettlösungen aus Smart Camera mit verbauter Optik und integrierter Beleuchtung bekannt.
	Resultat Logik:	min. horizontale Auflösung: 600 px; min. vertikale Auflösung: 430 px; min. Bildrate: 1 Hz; Global oder Rolling Shutter; mittlere Pixelgröße (5-8 µm); Industrial Ethernet; IP54; Cognex	

11.5 Anwendung der Auswahllogik, Schritt 5

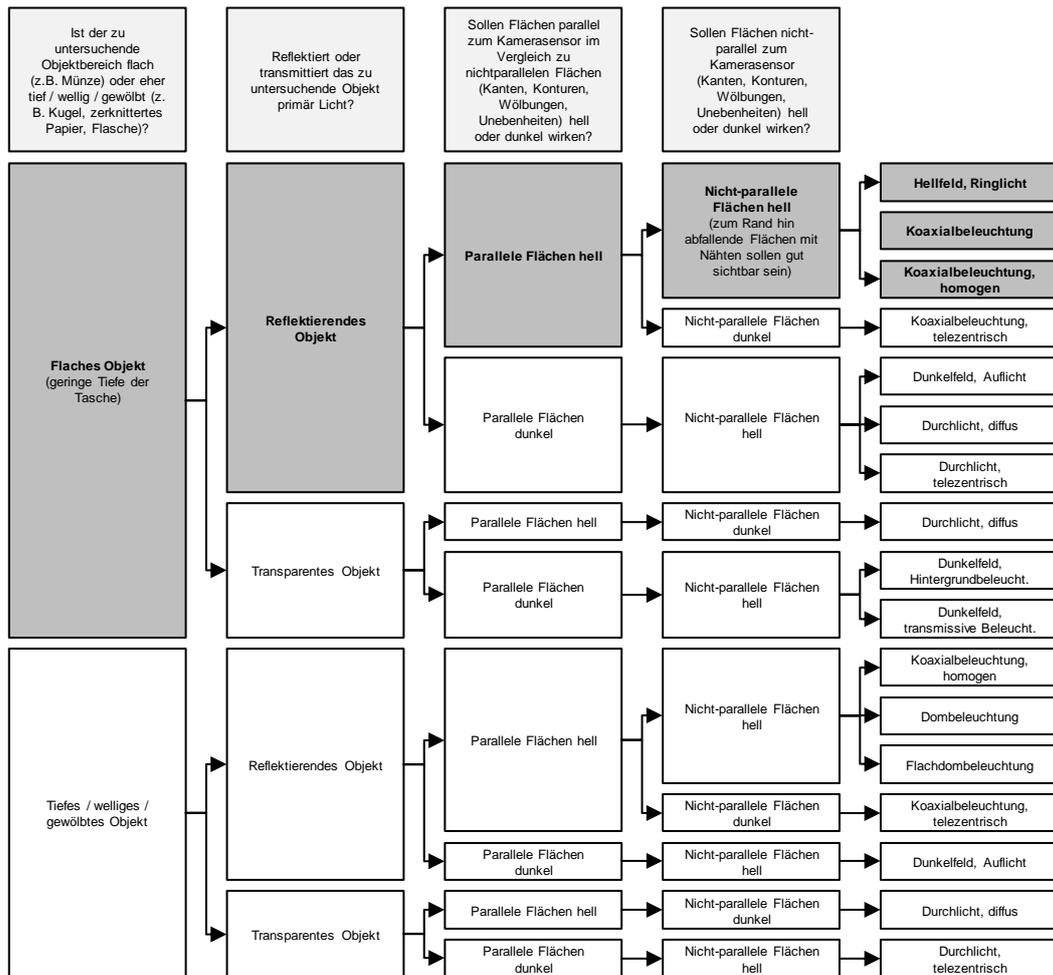


Abb. 11.4: Anwendung der Logik zur Auswahl einer geeigneten Beleuchtung

11.6 Berechnungsdetails der angepassten Sichtfeldgröße

```
# calculate adjusted field of view according to selected sensor
FOV_horizontal = max(distance_x.value, distance_y.value) # FOV in cm
FOV_vertical = FOV_horizontal * cam_selected['Vertikale Aufl (Px)'] / cam_selected['Horizontale Aufl (Px)']
FOV_diameter = (FOV_horizontal**2 + FOV_vertical**2)**0.5
```

Abb. 11.5: Screenshot Berechnung Sichtfeld

11.7 Benutzeroberfläche Entscheidungslogik

Entscheidungslogik industrielle Bildverarbeitung

1) Vorauswahl Kamera

Bitte beantworten Sie die folgenden Fragen zur allgemeinen Charakterisierung des Anwendungsfalls:

1) Sind die zu überprüfenden Merkmale innerhalb des menschlich sichtbaren Lichtbereiches identifizierbar?

Antwort:

2) Bedarf es zur Bildanalyse Farbinformationen oder sind monochromatische Bilder (schwarz-weiß) ausreichend?

Antwort:

3) Sollen Tiefeninformationen von 3D-Objekten erfasst werden?

Antwort:

4) Ist das zu überprüfende Objekt kontinuierlicher Art (Rollenware wie z. B. Folien oder Bleche) und wird mit hoher Geschwindigkeit im Sichtfeld der Kamera bewegt?

Antwort:

5) Ist das zu überprüfende Objekt zylindrischer Art, und soll dessen komplette Mantelfläche überprüft werden (z.B. Etikett auf einer Flasche)?

Antwort:

2D Verarbeitung, monochromatisch

Abb. 11.6: Benutzeroberfläche Entscheidungslogik, Schritt 1

3) Vorauswahl Objektiv

1) Verfügt die Anwendung über ein festes Sichtfeld (Objekt befindet sich immer in einem bestimmten Abstand und definierten Arbeitsbereich)? - D.h., das Sichtfeld muss während der Anwendung nicht dynamisch angepasst werden?

Antwort:

2) Bedarf es einem extremen Weitwinkel / 360° Blickfeld für die Anwendung (z. B. bei der Inspektion der Innenseite eines zylindrischen Körpers)?

Antwort:

3) Handelt es sich um eine sehr präzise Messaufgabe, bei der perspektivische Verzerrungen oder Verzeichnungen das Ergebnis verfälschen würden?

Antwort:

Festbrennweitenobjektiv

Abb. 11.7: Benutzeroberfläche Entscheidungslogik, Schritt 3

4) Detailauswahl Objektiv

Festbrennweite

Telezentrisch

A) Ausgewählte Kamera

Hersteller	Name	Listenpreis [€]	Auflösung	Pixelgröße (µm)	Bildfrequenz (Hz)	Gehäusetyp	Objektivanschluss	Horizontale Aufl (Px)	Vertikale Aufl (Px)	Sensortyp	Sensorgröße (mm)	Sensorm
680	VRMAGIC C-9+/BW-PRIO WINDOW V2	299,0	1,3	5,2	30,0	Small	C-Mount	1288	1032	CMOS	8,6	MT9V

B) Weitere Spezifikationen des Anwendungsfall

1) Was ist der minimal und maximal zulässige Abstand zwischen Objekt und Objektiv?

Abstand [cm]:

2) Was ist der relevante Tiefenbereich des Objektes, der scharf abgebildet werden muss?

Tiefe [cm]:

Abb. 11.8: Benutzeroberfläche Entscheidungslogik, Schritt 4 Teil 1 (Beispiel Festbrennweitenobjektiv)

C) Übersicht geeigneter Objektive & Detailspezifikation

Mindestanforderungen: Minimale Brennweite: 2.65 mm, Maximale Brennweite: 158.90 mm

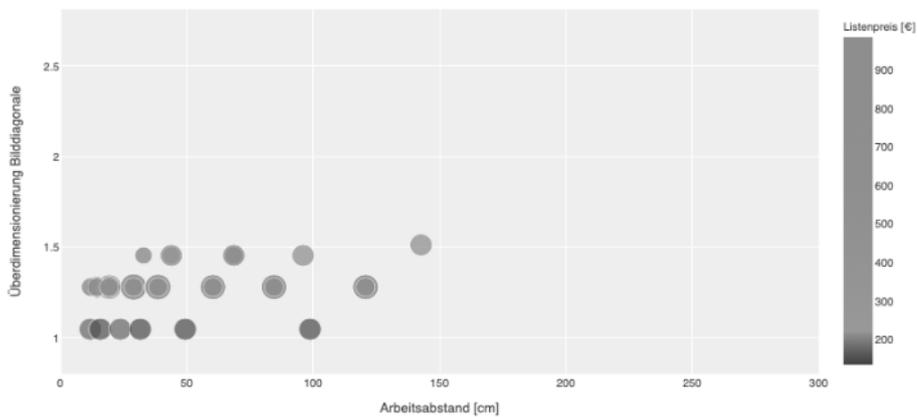
Empfohlene einzustellende Blende: 5.7

Warnung Blende: Blende unkritisch

Diagonale Bildsensor 8.6

Anzahl geeigneter Objektive: 130

Anzahl aktuell gezeigter Objektive: 76



Spektrale Empfindlichkeit:

- Monochrom
- Colour
- Infrarot
- UV

Blendentyp:

- verstellbar
- fixiert

Übermaß: 1.80

Preis max: 1000

Hersteller:

- Common Vision Optics
- Computar
- Fujifilm
- FUJINON
- Kowa
- Qioptiq
- Ricoh
- Schneider Kreuznach
- Tamron
- Zeiss

	Name	Hersteller	Listenpreis [€]	Brennweite [mm]	MOD [mm]	Min Pixelgröße [µm]	Bilddiagonale [mm]	Max Sensorgöße	Spektrale Empfindlichkeit	Blendenzahl	Blendentyp	Objektivanschluss
13	CVO GMTHR33520MCN	Common Vision Optics	135.0	35.0	250.0	4.3	11.0	2/3"	Colour	2.0	flexible	C-Mou
14	CVO GMTHR35028MCN	Common Vision Optics	135.0	50.0	500.0	4.3	11.0	2/3"	Colour	2.8	flexible	C-Mou
261	TAMRON M118FM25	Tamron	145.0	25.0	100.0	3.5	9.0	1/1.8"	Monochrome	1.6	flexible	C-Mou
262	TAMRON M118FM50	Tamron	145.0	50.0	200.0	3.5	9.0	1/1.8"	Monochrome	2.8	flexible	C-Mou
260	TAMRON M118FM16	Tamron	145.0	16.0	100.0	3.5	9.0	1/1.8"	Monochrome	1.4	flexible	C-Mou

#Resultate 5

Abb. 11.9: Benutzeroberfläche Entscheidungslogik, Schritt 4 Teil 2 (Beispiel Festbrennweitenobjektiv)

5) Auswahl Beleuchtungsart

Entscheidungshilfe

Übersicht

1) Ist das zu untersuchende Objekt flach oder eher wellig / konkav / konvex?

Antwort:

2) Reflektiert oder transmittiert das zu untersuchende Objekt Licht?

Antwort:

3) Sollen Flächen parallel zur Kamera im Vergleich zu Kanten / (nichtparallelen Flächen) hell oder dunkel wirken? (siehe Beispiel)

Antwort:

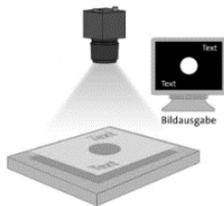
4) Sollen zur Kamera nichtparallele Flächen / Kanten / Konturen / Unebenheiten hell oder dunkel wirken?

Antwort:

Diffuses Durchlicht

Durchlicht, Diffus

Beleuchtungsanordnung:



Beispielapplikation:



Normale Hellfeldbeleuchtung



Diffuses Durchlicht zur Erkennung des Füllstandes

Abb. 11.10: Benutzeroberfläche Entscheidungslogik, Schritt 5

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig angefertigt habe. Es wurden nur die in der Arbeit ausdrücklich benannten Quellen und Hilfsmittel benutzt. Wörtlich oder sinngemäß übernommenes Gedankengut habe ich als solches kenntlich gemacht.

Ort, Datum

Unterschrift