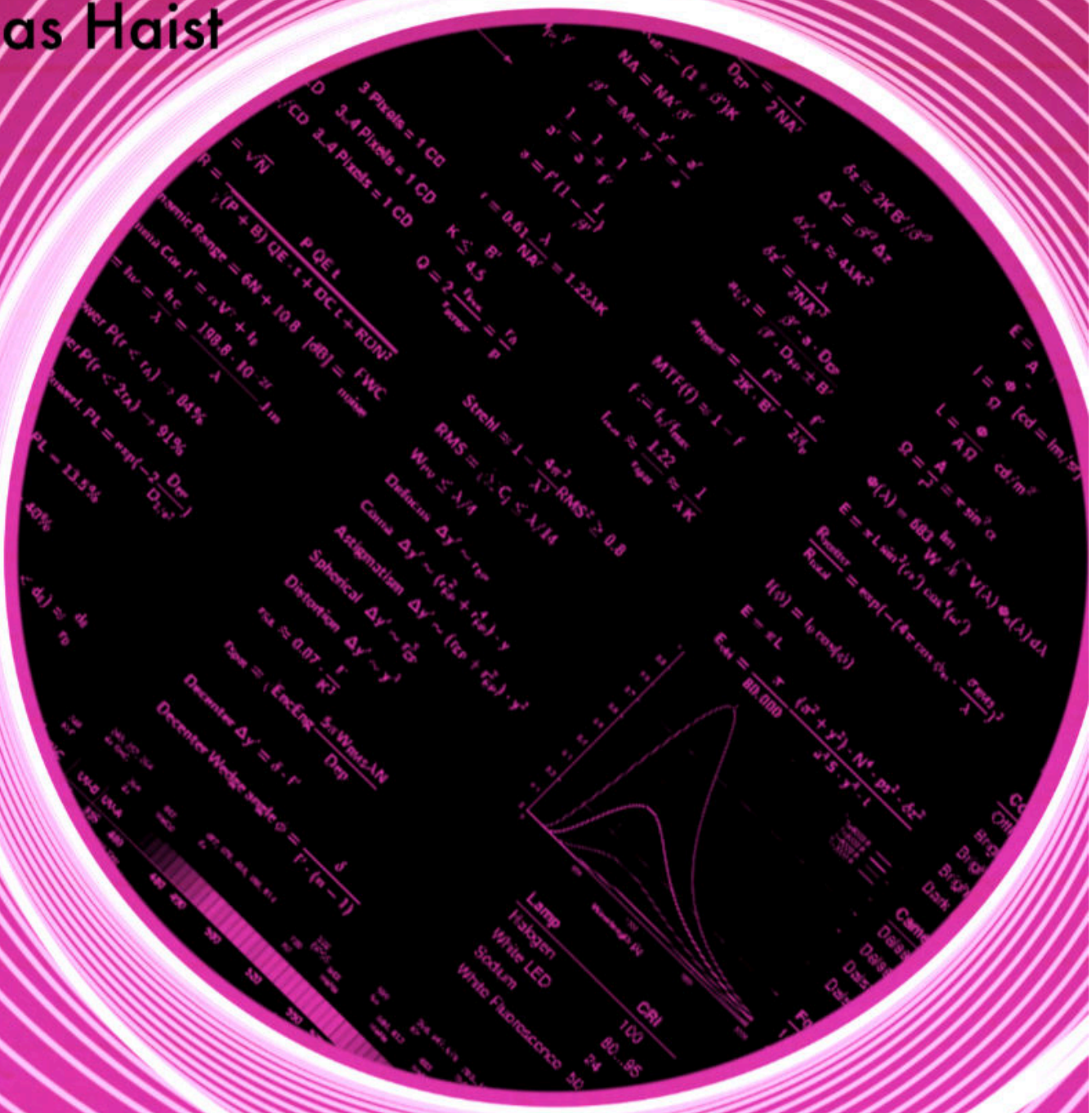
Tobias Haist



Optik in der Industriellen Bildverarbeitung

Inhaltsverzeichnis

1 Grundaufgaben der industriellen Bildverarbeitung **4**

1.1 Design des Gesamtsystems 9

2 Auslegung der Abbildung **13**

2.1 Abbildung mit einer Linsengruppe 14

2.2 Abbildung mit zwei Linsengruppen 22

2.3 Blende und Pupille 26

2.4 Schärfentiefe 32

2.5 Perspektive 35

2.6 Zusammenfassung wesentlicher Parameter 40

3 Mathematische Beschreibung der Bildentstehung und das Abtasttheorem **42**

3.1 Lineare, invariante Systeme 44

3.2 Eigenfunktionen und die Fouriertransformation 51

3.3 Die Fouriertransformation von Bildern 56

3.4 Berechnung 61

3.5 Einige Fourierpaare 63

3.6 Das Abtasttheorem 65

4 Transferfunktion **74**

4.1 Nochmals Faltung und Multiplikation 75

4.2 Zusammenschaltung von MTFs 81

4.3 Pixelgeometrie 85

4.4 Einige weitere Beispiele 88

4.5 Advanced: Grenzen 90

5 Laterale Auflösung **93**

5.1 Zur Definition von Auflösung 95

5.2 Zweipunkttrennung 97

5.3 Advanced: Überauflösung 102

5.4 Advanced: Dekonvolution 104

5.5 Bildfehler 105

5.6 Vignettierung 111

5.7 Pixelierung: Q 112

5.8 Advanced: Leere Vergrößerung 116

5.9 Modulationstransferfunktion 117

5.10 Advanced: Vorgehen zur Festlegung von Bildsensorauflösung und Auflösung der Optik 120

5.11 Advanced: Sensorgröße und Pixelgröße 128

5.12 Advanced: Die maximale Anzahl abbildbarer Pixel 133

5.13 Advanced: Weitere Auflösungskennzahlen 134

5.14 Advanced: Kohärente Abbildungen 140

5.15 Advanced: Auflösungsreduktion bei Farbsensoren 153

5.16 Advanced: Smartphones 154

5.17 Positionsauflösung 154
5.18 Herstellerangaben 159

6 Objektive 161

6.1 Weitere Beurteilungskriterien 164
6.2 Streulicht, Schleier und Geisterbilder 169
6.3 Foto- und Bildverarbeitungsobjektive 170
6.4 Telezentrische Objektive 172
6.5 Objektive für Zeilenkameras 173
6.6 Mikroskopische Abbildungen 174
6.7 Scheimpflug Bedingung 178
6.8 Mechanischer Anschluss 179
6.9 Filter und Schutzgläser 180
6.10 Infrarotoptiken 185
6.11 Advanced: Reinigung und Pflege 186

7 Fotometrie und Radiometrie 188

7.1 Lichttechnische Größen 188
7.2 Bestrahlungsstärke auf dem Bildsensor 198
7.3 Streuung und beleuchtete Flächen 199
7.4 Advanced: Notwendige Beleuchtungsstärke 200
7.5 Advanced: Wechselwirkung des Lichts mit dem Objekt 202
7.6 Spektralbereiche und Farbe 205

8 Beleuchtung 208

8.1 Beleuchtungsgeometrien und Beleuchtungsoptik 209

8.2 Advanced: Linienbeleuchtung 220
8.3 Fremdlicht 220
8.4 Lichtquellen 223
8.5 Advanced: Steuerung 237
8.6 Advanced: Spezielle Techniken 239

9 Bildsensoren 243

9.1 CMOS Sensoren 248
9.2 Advanced: CCD-Sensoren 250
9.3 Bildsensorgeometrie 253
9.4 Steuerung der Belichtung und Shutter 254
9.5 Rauschquellen 256
9.6 Advanced: Dynamik 267
9.7 Empfindlichkeit 272
9.8 Bereiche und Binning 281
9.9 Advanced: Triggerung 282
9.10 Advanced: Sensortransferfunktion 283
9.11 Farbsensoren 286
9.12 Advanced: Infrarotsensoren 291
9.13 Advanced: Zeilensensoren 294
9.14 Advanced: Kameraschnittstellen 296
9.15 Advanced: Intelligente Kameras 300
9.16 Advanced: Event-based Sensors 302
9.17 Zusammenfassung wesentlicher Parameter der Bildsensoren . . . 302

10 Advanced: Topografiesensoren **306**

10.1 Auflösungen und Messunsicherheiten 307

10.2 Was wird gemessen 308

10.3 Robustheit 309

10.4 Triangulation 310

10.5 Weißlichtinterferometrie 323

10.6 Shape from (De)Fokus 330

10.7 Konfokale Mikroskopie 331

10.8 Photometrisches Stereo (shape from shading) 335

10.9 Laufzeitsensorik 336

10.10 Digitale Holografie, Shearografie und ESPI 338

10.11 Computertomographie 341

10.12 Andere Entfernungssensoren 342

11 Spezifikation **345**

11.1 Musterkatalog 348

11.2 Messen, Prüfen und Rückführbarkeit 349

11.3 Spezifikation von Defekten 360

11.4 Prüfen und Klassifizieren 361

11.5 Statistik, Klassifikation und Prüfen 362

12 Abschließende Bemerkungen **369**

13 Anhang: Wahrscheinlichkeiten **371**

13.1 Oder und Und 371

13.2 Bedingte Wahrscheinlichkeit 373

13.3 Bayes Theorem 374

13.4 Das Baysche Theorem angewandt 375

.

Einführung

In diesem Buch soll ein eher stiefmütterlich behandelter Bereich der industriellen Bildverarbeitung, nämlich die Auswahl der Hardware von Bildverarbeitungssystemen besprochen werden. Konkret soll dargestellt werden, wie in systematischer Weise Abbildungs- und Beleuchtungsoptiken, Beleuchtungsquellen und Bildaufnehmer zu wählen und zu spezifizieren sind. Insbesondere für die Komponenten Objektiv, Beleuchtungsquelle und Bildsensor werden detailliert die technischen Angaben in Datenblättern besprochen, um eine Auswahl in der Praxis zu ermöglichen. Die hierfür notwendigen Grundlagen werden (ohne sonderliche theoretische Tiefe und möglichst anschaulich) dargestellt.

Obwohl es sich um ein extrem praxisnahes Thema handelt sind wir dennoch teilweise gezwungen, in sehr deutlicher Form von dem abzuweichen, was industrielle Praxis ist.

Die industrielle Bildverarbeitung ist ein interdisziplinäres und sehr erfahrungsorientiertes Arbeitsgebiet (ich vermeide hier den Begriff „Wissenschaft"). Während sich in der Bildverarbeitungsalgorithmik durchaus eine wissenschaftliche Vorgehensweise durchgesetzt hat, erfolgt die Auswahl der Bildverarbeitungshardware in vielen Anwendungsfällen basierend auf meist trivialen Grundüberlegungen und vor allem viel Erfahrung.

Dagegen ist zunächst nichts einzuwenden. Wenn genügend Erfahrung vorhanden ist, um sicher sagen zu können, welche Implementierung eine Aufgabe löst, dann sollte die Erfahrung auch genutzt werden.

Industrielle Bildverarbeitung ist aber auch sehr individuell. Der erfahrene Entwickler einer Bildverarbeitungsanwendung wird daher eine möglichst ähnliche Anwendung, die er bereits implementiert hat, als Grundlage seiner Überlegungen verwenden und basierend hierauf eine Lösung für die anstehende Aufgabenstellung wählen. Das kann, muss aber nicht, gut gehen und natürlich steht und fällt diese Vorgehensweise mit dem Maß der Ähnlichkeit zu bisherigen Aufgaben, der Schwierigkeit der Aufgabe und der Erfahrung des Entwicklers. Eine annähernd optimale Auswahl der Hardwarekomponenten eines BV-Systems wird man auf

diese Weise selten treffen können.

Es ist daher für den Neuling unerlässlich, die Grundlagen der Auswahl wirklich zu verstehen und sich nicht auf den Zufall bzw. das Glück zu verlassen. Dieses Wissen soll im folgenden in möglichst kurzer Form und parallel zur Vorlesung vermittelt werden. Dabei werde ich versuchen, die Theorie auf genau das gerade noch minimal Notwendige zu beschränken. Ableitungen und Beweise werden wir also meist auslassen und stattdessen im Sinne des puren Ingenieurvorgehens vor allem erklären, *wie* Systeme auszulegen sind und uns der Frage des *warum* eher anschaulich als mathematisch nähern.

Teilweise gehe ich für meisten Leser (und in Hinblick auf die mündliche Prüfung) deutlich zu weit ins Detail. Neben vielen nebensächlichen kleinen Details im Text habe ich auch versucht, entsprechende Abschnitte zu kennzeichnen. Unterkapitel mit dem Vorsatz „Advanced" sind also nicht prüfungsrelevant und damit aus meiner Sicht auch normalerweise für den typischen Bildverarbeiter nicht unbedingt notwendig. Scheuen Sie sich also nicht, diese Abschnitte zu übergehen. Ganz generell gilt: Überspringen, Sie die Kapitel, die für Sie irrelevant sind.

Die vorliegende Version des eBooks ist nicht perfekt (aber was ist schon perfekt?). Für jeden Hinweis auf Fehler oder Ungenauigkeiten, aber auch auf didaktisch schlechte Passagen bin ich sehr dankbar!

V1.3 Stuttgart, Oktober 2021

Kapitel 1

Grundaufgaben der industriellen Bildverarbeitung

Abbildung 1.0.1: *Viele Aufgaben in der industriellen Fertigung, die heute noch vom Mensch durchgeführt werden, werden zunehmend durch automatisierte Systeme erledigt. Bild mit freundlicher Genehmigung aus https://pixabay.com/de/workshop-operator-das-flie%C3%9Fband-802997/ unter der CC0 Creative Commons Lizens.*

Die industrielle Bildverarbeitung wird heute in nahezu allen industriellen Anwendungen eingesetzt. Von der Nahrungsmittelproduktion bis zur Überwachung von

Personen oder Fertigungsprozessen, von der Kontrolle von zugelieferten Materialien bis zur Endkontrolle von fertigen Maschinen: der Mensch wird zunehmend durch kamerabasierte oder anderweitig optisch erfassende Systeme abgelöst.

Die Hauptvorteile der maschinellen Kontrolle sind dabei natürlich eine Einsparung von Kosten und eine Erhöhung der Qualität. Entsprechende Systeme können rund um die Uhr ohne Ermüdung arbeiten und wesentlich schneller als Menschen hochgenau Objekte vermessen und das Ergebnis beurteilen. Subjektive werden durch objektive Beurteilungen ersetzt.

Während der Mensch bei der Analyse bzw. Deutung komplexer Szenen dem Computer nach wie vor weit überlegen ist, ist er gleichzeitig bei der fortlaufenden Beurteilung klar definierter metrischer Größen modernen Messsystemen natürlich hoffnungslos unterlegen. Mittels interferometrischer Methoden lassen sich heute leicht Höhenbestimmungen im Nanometerbereich durchführen und mit vollflächigen Verfahren lassen sich in Sekundenbruchteilen große Flächen auf kleinste Defekte untersuchen.

Die in der Praxis auftretenden Anwendungen sind so vielfältig wie die moderne Produktion selbst. Typische Grundaufgaben sind

- Objekte identifizieren, klassifizieren oder sortieren
- Objekte prüfen, z.B. Form, Oberfläche, Defekte im Volumen, Vollständigkeitsprüfung
- Objekte (bzw. ihre Eigenschaften) messen
- Objekte positionieren bzw. die Lage erkennen
- Steuerung von Anlagen
- Absicherung von Prozessen oder Anlagen
- Referenzmessungen, meist in einem gesonderten Messraum (*offline*).

Dabei ist zwischen den Aufgaben oft nicht eindeutig im Sinne einer (sinnvollen) Definition zu trennen. „Prüfen" kann z.B. meist als „Klassifikation" in „gut" und „schlecht" angesehen werden. Und In vielen Fällen ist für eine Prüfung zunächst eine Vermessung notwendig. Prüfen kann aber auch eine Vollständigkeitsprüfung

meinen und in diesem Fall ist wieder die Identifikation bzw. Klassifikation das Wesentliche.

Wichtig ist dabei in jedem Fall und zu aller erst, dass eine klare und eindeutige Vorstellung von der durch die Bildverarbeitung zu absolvierenden Funktion besteht. Da diese klare Beschreibung der Funktion oftmals alles andere als einfach und in vielen konkreten Anwendungen nicht gegeben ist werden wir hierauf später noch ausführlich zu sprechen kommen (vgl. Kapitel 11). Immer wenn Messungen Teil der Aufgabenstellung sind, dann ist der Fehlerbetrachtung und der sogenannten Rückführbarkeit viel Aufmerksamkeit zu widmen. Auch hierauf werden wir einzugehen haben.

Warum sollte man entsprechende Aufgabenstellungen aber überhaupt mittels Bildverarbeitung bzw. optischer Mess- und Inspektionstechnik lösen? Oder anders gefragt: Warum dominiert die optische Messtechnik heute die industrielle Fertigungsmesstechnik?

Als Grundvorteile kann man angeben:

- Berührungslose/zerstörungsfreie Messung,
- hohe Messgeschwindigkeit und
- parallele Erfassung vieler Messpositionen,
- Flexibilität durch Software,
- Messung aus Entfernung und
- Verschleißfreiheit

Viele dieser Vorteile, insbesondere in Kombination mit der hohen Messgeschwindigkeit bei der Erfassung vieler Datenpunkte lassen sich mit anderen Techniken (z.B. mechanische Antastung) nur schwer oder gar nicht erreichen. Natürlich sind diese immensen Vorteile auch der Grund warum eigentlich alle Lebewesen nennenswerter Größe sich so stark hinsichtlich ihrer Sensorik auf die optische Erfassung, also das Sehen, stützen.

Man sollte sich aber auch die Grenzen optischer Methoden bewusst machen. Zunächst werden die zu erfassenden Größen *indirekt*, eben über den Umweg des

Lichts, erfasst. Dies hat insbesondere bei der Erfassung sehr feiner Strukturen erhebliche Konsequenzen und es muss letztlich auf die eigentlich interessierende Größe (z.B. einen Abstand) rückgerechnet werden. Diese Rückrechnung kann schwierig sein und führt in jedem Fall zu Begrenzungen der optischen Messung (dazu später mehr).

Aber auch die oben genannten Vorteile müssen jeweils genau hinterfragt werden. Der Verschleißfreiheit steht die Verschmutzungsproblematik der Optik gegenüber. Hinsichtlich der berührungslosen Messung mit einem gewissen Abstand ist zu beachten, dass dadurch der Einfluss von Störeinflüssen (z.B. Fremdlicht) an Bedeutung gewinnt.

Nichtsdestotrotz überwiegen die technischen Vorteile. Der damit einhergehende Verlust an Arbeitsplätzen in der Inspektion und die zunächst anfallenden Investitionskosten werden zugunsten dieser Vorteile heute zunehmend in Kauf genommen, so dass die industrielle Bildverarbeitung ein stetig wachsender Markt ist.

Generell kann ein BV-System meist in die Bereiche

1. Sensorik bzw. Datenerfassung
2. Verarbeitung
3. Verwertung der Resultate (z.B. Steuerung, Qualitätssicherung, Visualisierung)

eingeteilt werden.

Der Schwerpunkt der Darstellung in diesem Buch wird der erste Punkt, also der Aufbau von Bildverarbeitungssystemen aus Sicht der Datenerfassung sein. Konkret: die Wahl der Optik, der Bilderfassungskomponenten sowie der Beleuchtung. Abb. 1.0.2 verdeutlicht diesen hardwareseitigen Aufbau eines typischen Bildverarbeitungssystem. Fragen der Verarbeitung bzw. Ergebnisbewertung werden wir nur insofern ansprechen, als dass sie für die Auslegung der Sensorik von Bedeutung sind.

Hinsichtlich der *optischen Grundlagen* werden wir uns mit den folgenden Grundfragen der optischen Abbildung beschäftigen:

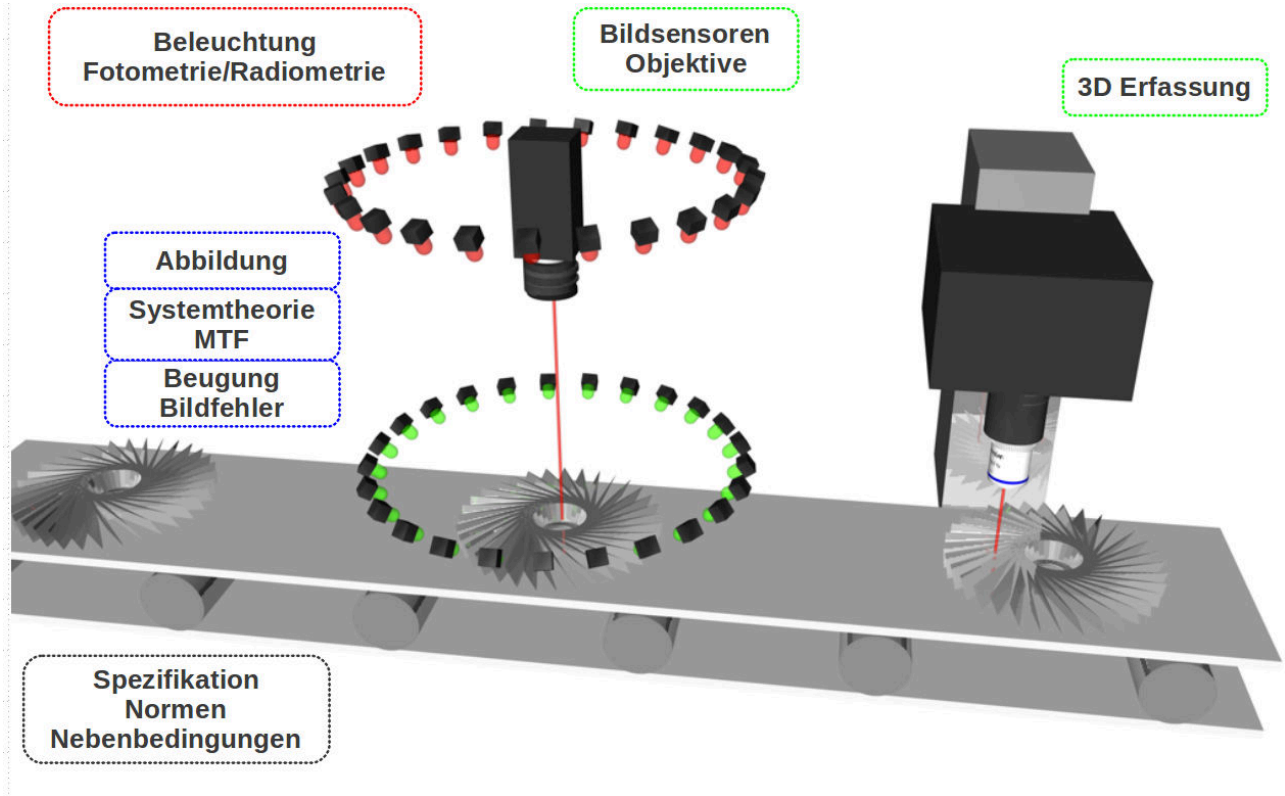


Abbildung 1.0.2: Wesentliche optische Teilaspekte der industriellen Bildverarbeitung

- Wo liegt der einzelne Bildpunkt (paraxiale Optik) ?
- Welche Form hat der einzelne Bildpunkt (Auflösung, Beugung und Bildfehler) ?
- Wie hell ist der einzelne Bildpunkt (Photometrie und Radiometrie) ?
- Wie erfolgt die Umwandlung des optischen Signals in ein digitales Signal (Bildsensoren und Interfaces) ?
- Wie kann die gesuchte Information optisch aufbereitet werden (Beleuchtungsgeometrien, Filter) ?

Neben der grundlegenden Theorie sind aber vor allem praktische Aspekte, z.B. Wahl der Leuchtmittel und konkrete Wahl von Objektiven, für uns wichtig.

1.1 Design des Gesamtsystems

Bildverarbeitungssysteme können vergleichsweise einfach aber auch sehr komplex werden. Je schwieriger die zu lösende Aufgabe, desto herausfordernder aber auch wichtiger ist das Gesamtdesign des Systems.

Bevor wir also in den einzelnen Kapiteln auf Einzelkomponenten von Bildverarbeitungssystemen eingehen, sollen hier einige allgemein zu beachtende Dinge besprochen werden. Im wesentlichen gelten sie für das Design praktisch jedes einigermaßen nicht-trivialen Systems. Umso wichtiger ist es, sie sich immer wieder klar zu machen und sie zu beachten:

Spezifikation: Von absolut fundamentaler Bedeutung ist es, zunächst exakt zu verstehen, was eigentlich genau gemessen oder detektiert werden soll. Die Aufgabenstellung muss komplett klar sein und — ganz besonders im Bereich der Bildverarbeitung — muss sie kritisch hinterfragt werden. Oftmals kommen Spezifikationen recht willkürlich zustande und machen bei genauer Betrachtung wenig Sinn. Darüber hinaus besteht in aller Regel zunächst keinerlei Konsens, wie geforderte Parameter genau definiert sind und wie sie verifiziert bzw. abgenommen werden sollen. Ein typisches Beispiel im BV-Umfeld ist ein unscharfer Begriff wie „Auflösung". Solange aber nicht absolut quantifiziert ist, was das System leisten soll, ist es kaum möglich, das System sinnvoll auszulegen.

Messmethodik: Der nächste entscheidende Punkt ist die Wahl der passenden Messmethodik. Der Bildverarbeiter denkt konservativ zunächst eben an eine konventionelle zweidimensionale Abbildung mit nachfolgender Auswertung per Software. Aber das ist natürlich nicht immer die beste Variante. Mess- bzw. Erfassungsmethodiken gibt es sehr viele. Abstandsmessungen müssen z.B. nicht zwangsweise optisch durchgeführt werden sondern können teilweise eleganter und preisgünstiger mittels anderer Sensoren erfolgen. Produktidentifikation muss z.B. nicht mittels Barcodes erfolgen, sondern es können evtl. RFID-Tags eingesetzt werden.

Aber auch hinsichtlich der optischen Erfassung gibt es viele Varianten, insbesondere sobald nicht nur zweidimensionale Bilder gewonnen werden sollen.

Die Wahl der passenden Messmethodik ist manchmal trivial und klar, manchmal aber auch sehr zweifelhaft. In diesem Fall lohnt es sich, Expertenmeinungen bzw. Erfahrungen (unter Umständen durch Literaturrecherche) einzuholen. Dies erscheint zunächst aufwändig, der Aufwand lohnt sich aber (vgl. nächster Punkt).

Verarbeitung: An dieser Stelle sollte man sich auch bereits wenn möglich über die Verarbeitung der Rohdaten Gedanken machen, denn die Auswertung hat oft einen entscheidenden Einfluss auf das Endresultat. Einfache, sehr schnelle Auswertungen liefern unter Umständen nur vergleichsweise schlechte Ergebnisse oder benötigen unter Umständen ein höheres Signal-Rausch-Verhältnis.

Fehler: Generell gilt bei jedem Projekt, dass Fehler bzw. Fehlentscheidungen a) kaum zu vermeiden sind und b.) umso teurer werden, je später sie entdeckt werden bzw. ihre negative Auswirkung zeigen. Ein zusätzlicher Tag Aufwand während der Planungsphase kann so unter Umständen Monate an unnützer Arbeit vermeiden. Wird z.B. beschlossen, dass die Messung mittels Streifenprojektion erfolgt und später zeigt sich in der Praxis, dass damit die geforderte Messunsicherheit nicht erzielt werden kann und daher ein konfokales Mikroskop eingesetzt werden muss, dann war die komplette Entwicklung des Streifenprojektionssystems letztlich nicht zielführend.

Dies gilt auf allen Ebenen - von dem Grunddesign bis zur Software. Fehler sollten so früh wie irgend möglich gefunden werden und es ist explizit einzuplanen, Fehler früh zu finden.

Unter Umständen sollten kleine Versuche zur experimentellen Verifikation (von Teilaspekten) eingeplant werden.

Mehrere Personen: Eine wichtige Methode, unnötige Fehler zu vermeiden ist es, mehrere Personen an der Aufgabe arbeiten zu lassen. Die Wahrscheinlichkeit, dass eine einzelne Person (egal wie erfahren und genial) bei einem komplexen System Fehler übersieht ist hoch. Es ist daher unumgänglich, mehrere Leute mit einzubeziehen. Je komplexer das System ist, desto mehr Teammitglieder sind notwendig. Verschiedene Methodiken sind möglich, wie diese Teammitglieder sich gegenseitig überprüfen können. Auch wenn der Aufwand dadurch zunächst steigt ergibt sich — außer bei vergleichsweise klaren und trivialen Aufgaben — letztlich ein deutlicher Gewinn.

Ein wichtiger positiver Nebeneffekt ist, dass die Verantwortung für das System verteilt wird. Je größer die mögliche negative Auswirkung des Versagen eines Systems ist (z.B. Personenschäden), desto wichtiger wird dieser Aspekt. Lastet die Verantwortung komplett auf einer Person, dann ist es für diese schwierig, unter einer solchen Belastung effektiv zu arbeiten.

Bedenken Sie: Fehler werden immer gemacht. Ein komplett fehlerfreies Arbeiten ist meist nicht realistisch und je größer der Zeitdruck ist, desto mehr Fehler sind zu erwarten. Planen Sie also Fehler (und ihre Erkennung) ein.

Protokollierung: Gewöhnen Sie sich an, anstatt eines Taschenrechners ein Skript für ihre Berechnungen zu verwenden. Dies kann z.B. in Matlab erfolgen, besser ist aber eine schnelle und freie Alternative wie octave. Damit erhalten Sie zwei Vorteile: 1. Sie können nachträglich einfach komplette Rechnungen, z.B. mit anderen Parametern oder wenn Sie einen Denkfehler entdeckt haben, wiederholen. 2. Die Arbeit ist protokolliert (für Sie selbst, aber auch für Teamkollegen, Auftraggeber usw.)

Nach diesen sehr generellen Designgrundsätzen kommen wir nun zu einigen eher spezifischeren Ratschlägen

Optische Systemauslegung: Zunächst sollte das optische System in jedem Fall lediglich paraxial ausgelegt werden. Aberrationen, Beugungsbegrenzung und Lichteffizienz werden zunächst separat abgeschätzt und dann erst später genauer betrachtet. Auch andere Systemteile sollten zunächst erst linearisiert betrachtet werden.

Gegenprüfung: Wann immer möglich sollten Plausibilitätsprüfungen durchgeführt werden. Nicht nur eigene, sondern auch professionelle Software ist voll von Fehlern, so dass z.B. auch Zemaxsimulationen hinterfragt werden sollten. Am besten ist es immer, mittels unterschiedlicher Tools einen Sachverhalt gegen zu prüfen (z.B. analytische Rechnung vergleichen mit Raytracing-Simulation).

Lichtbudget: Die Grundfragen bei jeder Abbildung sind: 1) Wo ist der Bildpunkt, 2) wie ist die Verteilung des Bildpunkts (Bildfehler) und 3) wie hell ist der Bildpunkt?

Optiker (und Bildverarbeiter) neigen dazu, den letzten Punkt zunächst zu vernachlässigen. Allerdings ist er von zentraler Bedeutung, denn viele Anwendungen der industriellen Bildverarbeitung sind hinsichtlich Beleuchtung bzw. Empfindlichkeit anspruchsvoll. Eine zu geringe Beleuchtungsstärke auf dem Bildsensor führt zu einem geringen Signal-Rausch-Verhältnis und damit zu einer schlechten Gesamtperformance bzw. einem Versagen des Systems. Es sollte also frühzeitig überprüft werden, ob bzw. wie genügend Licht zur Verfügung gestellt werden kann.

Fehlerbudget: Letztlich muss für alle wesentlichen Spezifikationen des Systems abgeschätzt werden, welche Fehlerquellen relevant sind und in welcher Größenordnung sie liegen. Für messende Systeme ist die Messunsicherheit in Form eines Fehlerbudgets aufzustellen.

In der ersten Entwurfsphase des Systems reichen also zunächst einfache Abschätzungen und Erfahrungswerte. Hier ist aber auch die größte Gesamterfahrung notwendig. Erst wenn man damit zu einem zufriedenstellenden System gelangt, lohnt es, in die weitere Designphase einzusteigen.

In diesem Fall ist es vor allem bei komplexen Systemen und Systemen am Rande der technischen Machbarkeit wesentlich, zumindest Teilaspekte oder aber das Gesamtsystem zu simulieren.

Kapitel 2

Auslegung der Abbildung

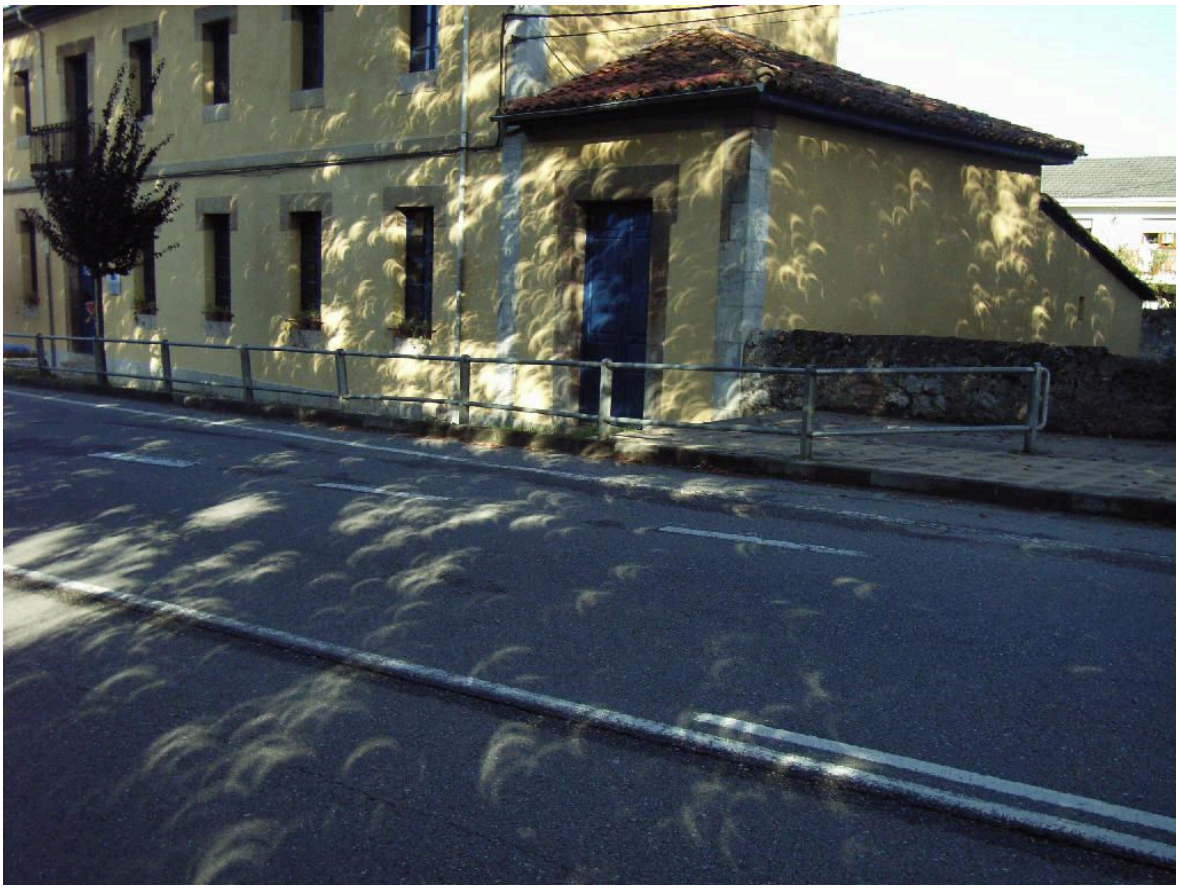


Abbildung 2.0.1: Abbildung der Sonne (während partieller Sonnenfinsternis) durch die vielen Löcher im Blätterdach eines Baums. Jedes Loch wirkt als Lochkamera. Foto mit freundlicher Genehmigung von Jan Liesener.

In diesem Kapitel werden Sie unter anderem lernen

- wie man eine einstufige Abbildungsgeometrie auslegt,

- welche Arten von Perspektiven es gibt und was hinsichtlich der Perspektive zu beachten ist,
- wie man die sehr gebräuchliche Kepleranordnung als zweistufiges System einsetzt,
- wie man die Öffnung der Optik spezifiziert und wie sie sich auf den Schärfentiefenbereich auswirkt und
- was die Pupillen eines optischen Systems sind und wozu wir diese benötigen.

2.1 Abbildung mit einer Linsengruppe

Im folgenden werden wir so kurz wie möglich an die wesentlichen Grundlagen der geometrisch-optischen Abbildung erinnern. Für Ableitungen oder Details verweisen wir auf die Vielzahl geeigneter Einführungen (z.B. [31]) und Referenzen (z.B. [19]). Sie werden feststellen, dass man mit lediglich zwei einfachen Gleichungen die Abbildungsgeometrie einer typischen Abbildung in der Bildverarbeitung festlegen kann.

Eine dieser Gleichungen ergibt sich ganz direkt aus dem Strahlensatz. Sie können dazu von einer Kamera mit sehr wenig geöffneter Blende bzw. einer *Lochkamera* ausgehen (Abb. 2.1.2).

Nach dem Strahlensatz gilt für den sogenannten *Abbildungsmaßstab*[1] (magnification) β', also das Verhältnis von Bildgröße y' zur Objektgröße y,

$$\beta' := \frac{y'}{y} = \frac{a'}{a} \tag{2.1}$$

[1]Umgangssprachlich und leider in großem Maße im Bereich der Bildverarbeitung wird oft der Begriff „Vergrößerung" verwendet. Dies ist problematisch, denn in der Optik bezeichnet „Vergrößerung" normalerweise eine Winkelvergrößerung. Selbst in den DIN Normen ergeben sich widersprüchliche Angaben zum Begriff der Vergrößerung, also die Vergrößerung, die z.B. ein Fernrohr bewirkt (angular Magnification). In der Mikroskopie wird „Vergrößerung" als „laterale Vergrößerung" für den (lateralen) Abbildungsmaßstab verwendet. Bei Okularen oder Fernrohren, die auf das Auge bezogen sind wird teilweise für die Vergrößerung eine Definition über den Tanges des Winkels und teilweise über den Winkel direkt verwendet (vgl. DIN 8039, DIN 1335).

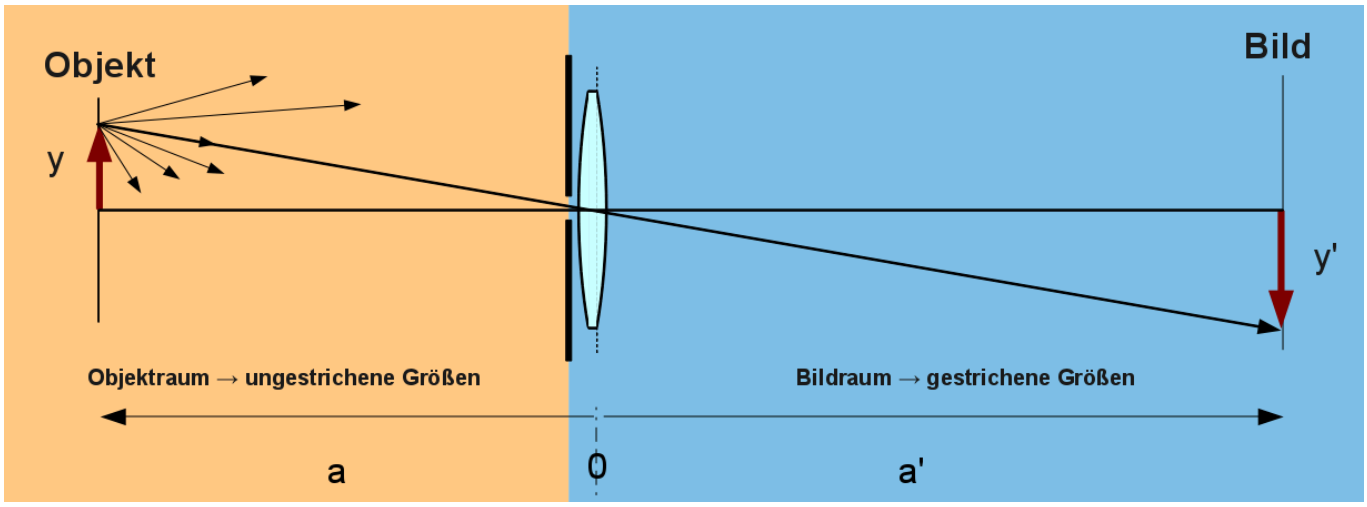


Abbildung 2.1.2: Typische optische Abbildungsgeometrie, hier am Beispiel einer Linsen- oder Lochkamera. Der (laterale) Abbildungsmaßstab y'/y ergibt sich direkt nach dem Strahlensatz zu $\beta' = y'/y = a'/a$. Der eingezeichnete Strahl ist der sogenannte „Hauptstrahl" des Objekt- bzw. Bildpunktes.

Hier und im folgenden verwenden wir gestrichene Größen für den Bildraum (in Abb. 2.1.2 *rechts* vom optischen System) und ungestrichene Größen für entsprechend Größen im Objektraum (*links* vom optischen System). Der Abstand des optischen Systems — hier des Loches — zum Objekt wird *Objektweite* genannt und mit a bezeichnet. In unserem einfachen Abbildungsmodell ist die Objektweite gleichbedeutend mit dem sogenannten *Arbeitsabstand* der Optik[2]. Entsprechend ist der Abstand des optischen Systems zur *Bildebene* die *Bildweite* und wird durch a' bezeichnet.

Beachten Sie bereits hier unbedingt die Definition des Koordinatensystems bzw. der Vorzeichen. Das Zentrum des Koordinatensystems liegt in der Linsenebene. Das heißt in Abb. 2.1.2 ist a eine negative Größe, während a' positiv ist. Dementsprechend ist der Abbildungsmaßstab β' negativ und in der Tat steht das Bild „auf dem Kopf", ist also invertiert. Wir werden diese Vorzeichenkonvention im gesamten Buch verwenden und alle Formeln basieren auf dieser Konvention.[3]

Aus Gründen der Lichteffizienz (durch das Loch der Lochkamera kann nur sehr

[2]Der Arbeitsabstand ist üblicherweise definiert als der Abstand der vorderen Begrenzung eines Systems (hier der Optik) zum Objekt.

[3]Vereinzelt findet sich in der Bildverarbeitungsliteratur auch eine Darstellung von Formeln basierend auf anderen Vorzeichendefinitionen. Beachten Sie dies!

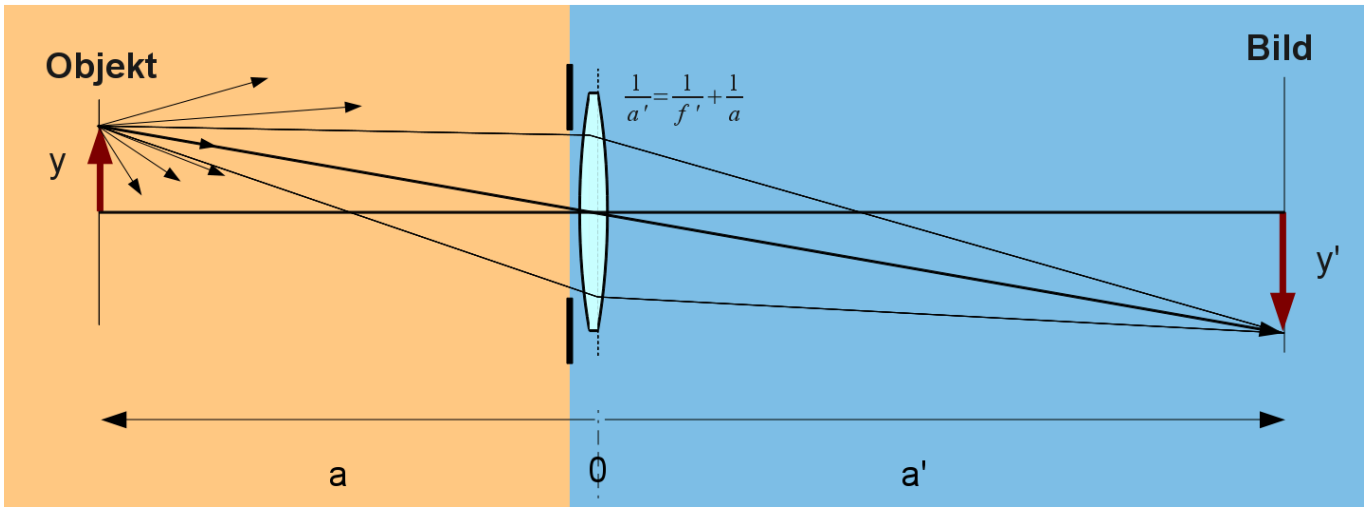


Abbildung 2.1.3: Die Linse (bzw. das Linsensystem) biegt die Strahlen vom Objekt kommenden Strahlen so um, dass sie im selben Bildpunkt konvergieren. Dazu müssen Objekt-, Bild- und Brennweite gemäß der Abbildungsgleichung miteinander verbunden sein.

wenig Licht dringen) verwenden nahezu alle Abbildungssysteme, sowohl biologische (z.B. Auge) als auch technische Systeme (z.B. Mikroskop), eine Linse bzw. ein Linsensystem anstatt des Lochs.[4] Die Linse sorgt letztlich dafür, dass auch Lichtstrahlen, die nicht durch das Loch der Lochkamera laufen, so umgelenkt werden, dass sie an derselben Position in der Bildebene auftreffen. Das funktioniert aber bei gegebenem Abbildungsmaßstab nur für *eine* Kombination aus Objektweite a und Bildweite a'. Diese Kombination an konkreten Abständen hängt mit einem, um nicht zu sagen *dem* elementaren Linsenparameter, der sogenannten *Brennweite* f', über eine weitere Gleichung, die sogenannten *Abbildungsgleichung* zusammen:

$$\frac{1}{a'} = \frac{1}{f'} + \frac{1}{a} \tag{2.2}$$

Beachten Sie bei der Anwendung unbedingt wieder die korrekten Vorzeichen, da ansonsten völlig falsche Werte resultieren. Die Brennweite ist hier die „bildseitige" Brennweite und wird daher — als Größe im Bildraum — mit einem Strich gekennzeichnet. Wenn man (etwas lax) von „der Brennweite" eines optischen Systems spricht ist meist diese bildseitige Brennweite gemeint.

[4] Auch die erzielbare Auflösung ist mit einem Lochkamerasystem deutlich begrenzt, vgl. Kapitel 5.

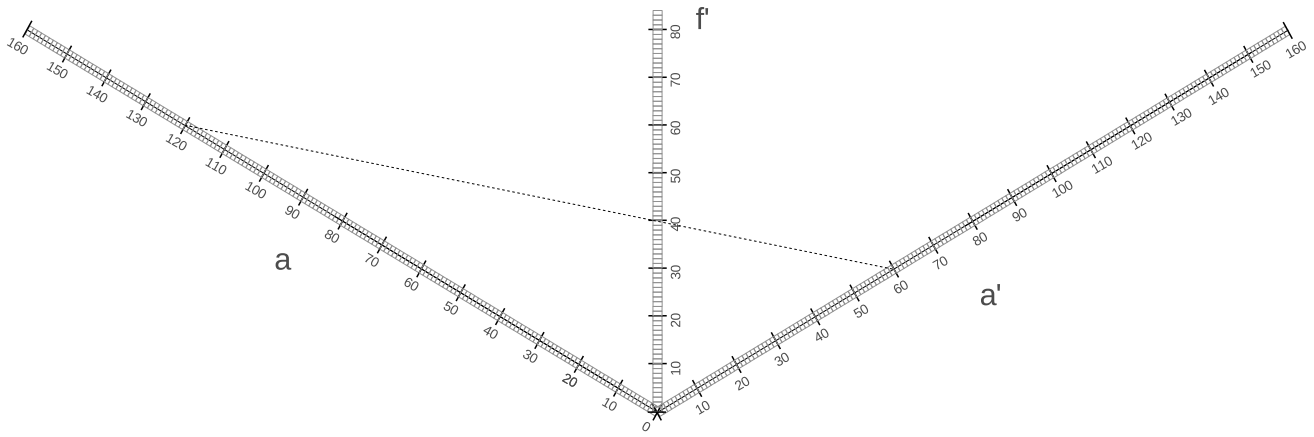
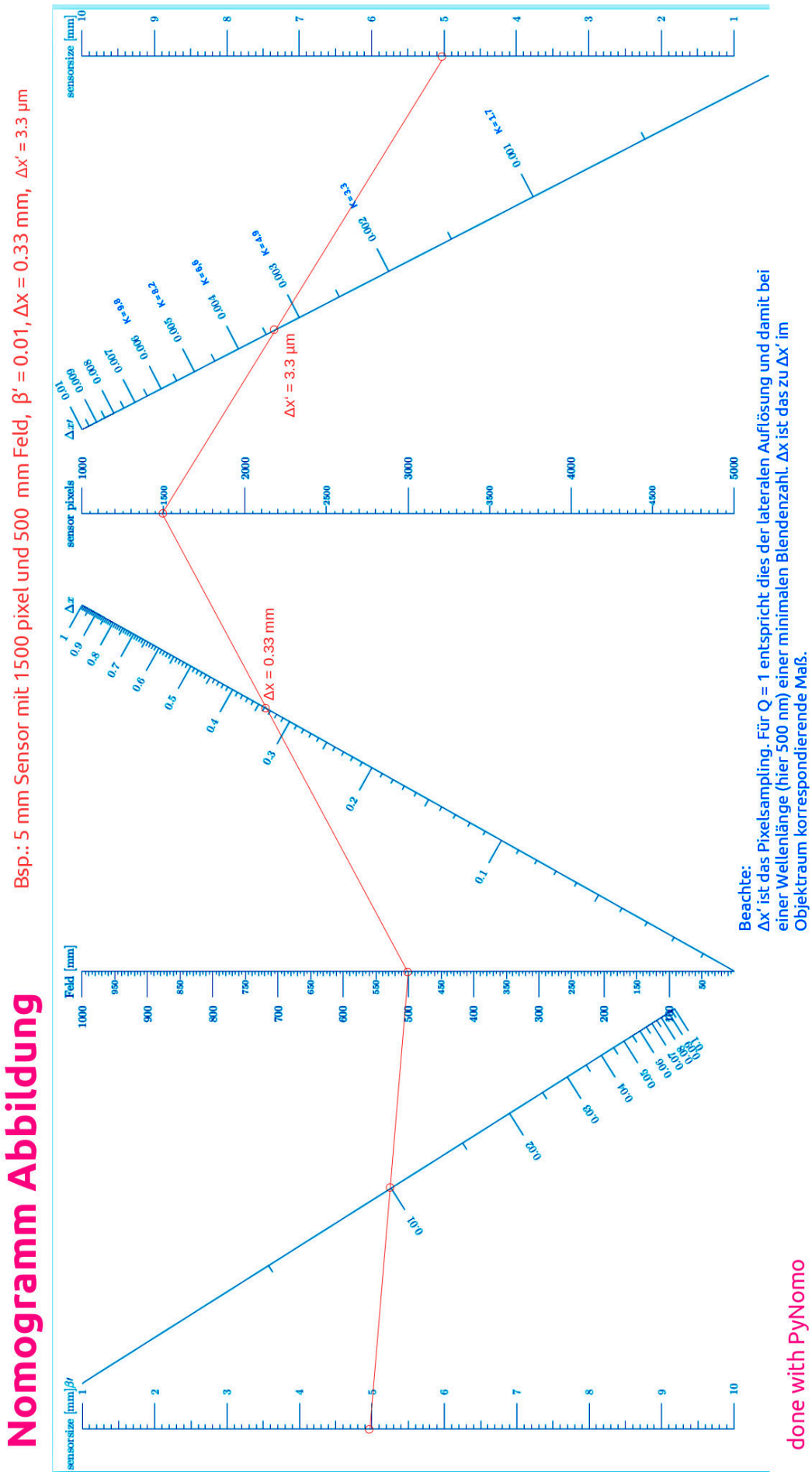


Abbildung 2.1.4: *Gaußsches Nomogramm zur grafischen Lösung der Abbildungsgleichung* 2.2. Im eingezeichneten Beispiel wird die Brennweite einer Abbildungslinse für die Objektweite -120 mm und die Bildweite 60 mm gesucht. Es ergibt sich $f' = 40$ mm. Die Einheiten sind beliebig und sie können die Achsen beliebig skalieren, d.h. für eine Objektweite von -600 mm und eine Bildweite von 300 mm ergibt sich eine Brennweite von 200 mm.

Falls Sie eine grafische Bestimmung bevorzugen können Sie Gleichung 2.2 anhand des Gaußschen Nomogramms lösen. Verbinden Sie hierzu einfach zwei der drei Größen a, a', f' durch eine Gerade und lesen Sie die dritte Größe ab (siehe Abb. 2.1.4).[5]

Ein etwas komplexeres Nomogramm zur Bestimmung von Auflösungen (im Sinne des Samplings) ist in Abb. 2.1.5 dargestellt.

[5]Im übrigen kann man ganz allgemein beim Auslegen von Systemen die Zusammenhänge verschiedener Parameter durch geeignete Kombination von Graphen als „Nomographen" gewinnbringend einsetzen (siehe z.B. [35]).


Nomogramm Abbildung
Bsp.: 5 mm Sensor mit 1500 pixel und 500 mm Feld, β' = 0.01, Δx = 0.33 mm, Δx' = 3.3 µm
sensorsize [mm]
β'
Feld [mm]
Δx
Δx = 0.33 mm
sensor pixels
Δx'
Δx' = 3.3 µm
sensorsize [mm]
Beachte:
Δx' ist das Pixelsampling. Für Q = 1 entspricht dies der lateralen Auflösung und damit bei einer Wellenlänge (hier 500 nm) einer minimalen Blendenzahl. Δx ist das zu Δx' im Objektraum korrespondierende Maß.
done with PyNomo

Typischerweise gehen Sie zur Grundauslegung der Abbildung folgendermaßen vor:

1. Sie berechnen den gewünschten Abbildungsmaßstab β' anhand der Aufgabengeometrie (meist eben das Verhältnis aus Bildaufnehmergröße zur maximalen Größe des Objekts).
2. Falls die Objektweite (bzw. der Arbeitsabstand) a gegeben ist können Sie mit Gl. 2.1 die Bildweite a' berechnen.
3. Aus a' und a berechnen Sie dann dann die benötigte Brennweite des optischen Systems.

Beispielhaft wollen wir die Abbildung eines 100 mm großen Maschinenteils auf einen Bildaufnehmer mit einer Kantenlänge von 10 mm betrachten. Offensichtlich ist der Abbildungsmaßstab β' = -10 mm / 100 mm = -0.1. Bei einem gewünschten Arbeitsabstand von -500 mm ergibt sich z.B. die Bildweite $a' = \beta' \cdot a$ = 50 mm. Mit Gl. 2.2 können wir die Brennweite des benötigten optischen Systems direkt zu f'=45.5 mm berechnen.

Es fällt auf, dass die Brennweite f' sehr nah bei der Bildweite a' liegt. Dies ist kein Zufall. Wenn die Objektweite a groß gegen die Bildweite a' ist, dann kann man für Überschlagsrechnungen einfach die Brennweite mit der Bildweite gleichsetzen. Je weiter das Objekt entfernt ist, desto besser ist diese Näherung (Grenzfall: Objekt im Unendlichen). Das ist übrigen auch die Definition der Brennweite: Parallel (also aus unendlicher Entfernung kommendes) einfallendes Licht wird im Abstand der Brennweite zur Linse fokussiert. Die entsprechende Ebene wird auch (bildseitige) *Brennebene* genannt, vgl. Abb. 2.1.6.

Für viele BV-Anwendungen ist das Objekt deutlich weiter entfernt als die Brennweite des Objektivs. Das bedeutet, dass das von einem Objektpunkt in die Optik einfallende Licht näherungsweise parallel ist und man daher das Bild in der Brennebene des Objektivs erhält. Es gilt also oft näherungsweise $a' \approx f'$.

Je nachdem, was bei einer konkreten Aufgabenstellung gegeben ist, kann man so durch einfaches Umformen bzw. Einsetzen der Gleichung 2.1 immer eine erste Grundauslegung der Optik vornehmen. Letztlich haben wir so die Geometrie der Abbildung und die Brennweite des Abbildungssystems festgelegt.

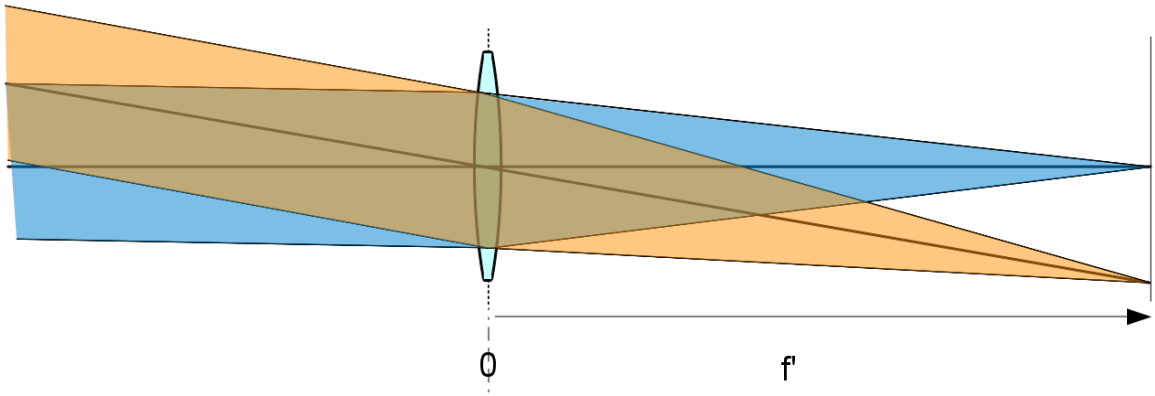


Abbildung 2.1.6: Parallel einfallendes Licht wird in der Brennebene gebündelt. Der Abstand der Brennebene von der Optik (genauer, der bildseitigen Hauptebene der Optik) wird Brennweite genannt. Für Objekte, die weit entfernt sind ist die Bildweite näherungsweise die Brennweite, also $a' \approx f'$.

Bei gegebener Brennweite einer Optik und gewünschtem Abbildungsmaßstab kann man mittels

$$a = f'(\frac{1}{\beta'} - 1) \tag{2.3}$$

auch direkt die Objektweite bestimmen. Die Gleichung ergibt sich direkt durch Einsetzen und Umformen unserer beiden Grundgleichungen.

Beispiel: Berechnen Sie die Objektweite für eine Anwendung mit Abbildungsmaßstab -0.1 und einer Objektivbrennweite von 100 mm.

Lösung: a = 100 mm (-10 - 1) = -1100 mm

Beispiel: Sie sollen ein Inspektionssysteme für Zahnräder mit einem Durchmesser von 100 mm auf einen 1/3 Zoll Bildsensor mit 1000 Pixeln abbilden (vertikale Ausdehnung: 3.6 mm). Der Arbeitsabstand soll 0.5 m betragen. a) Berechnen Sie den Abbildungsmassstab. b) Berechnen Sie die Brennweite der Abbildungsoptik.

Lösung: Der Abbildungsmaßstab ergibt sich direkt zu $\beta' = y'/y = -36 \cdot 10^{-3}\,a$. Daraus können wir wegen $\beta' = a'/a$ mit $a =$ -0.5 m die Bildweite berechnen. Es ergibt sich $a' = 18$ mm. Und damit über die Abbildungsgleichung die Brennweite $f' =$ 17.37 mm.

Wir sind bislang davon ausgegangen, dass das Objekt in einer klar definierten Objektebene liegt. Wenn das Objekt eine gewisse Tiefenausdehnung hat oder in verschiedenen Abständen liegen kann, ergibt sich nur für eine Ebene eine ideal scharfe Abbildung (vgl. Kapitel 5 und Abschnitt 2.3). Bei einer Verschiebung um Δz im Objektraum ist eine Verschiebung um

$$\Delta z' = \beta'^2 \, \Delta z \tag{2.4}$$

im Bildraum vorzunehmen, um wieder ein scharfes Bild zu erhalten (Nachfokussierung) [44].

Die Abbildungsmaßstäbe lateral und in der Tiefe sind demnach nicht gleich. Optische Abbildungen sind also nicht „volumentreu" und sie müssen den Zusammenhang eben bei einer etwaigen Nachfokussierung beachten. Beachten Sie, dass dieser *Tiefenabbildungsmaßstab* in keiner Form gleichbedeutend mit der Schärfentiefe (vgl. Abschnitt 2.4) ist, sondern lediglich die Lage der perfekten Bildpunkte im dreidimensionalen Raum beschreibt.

Beispiel: Sie bilden ein Objekt mit einem Durchmesser von 1 m auf einen 1/3 Zoll Sensor (vertikale Sensorausdehnung: 3.6 mm) ab. Die Schärfentiefe ist begrenzt aufgrund einer geringen verwendeten Blendenzahl. Gegenüber der ursprünglichen Objektposition wird nun das Objekt um 100 μm verschoben. Um wieviel müssen Sie den Bildsensor nachfokussieren?

Lösung: $|\beta'| = 3.6/1000 = 0.0036$ und damit $\Delta z'$ = 1.3 nm (also effektiv gar nicht).

Manchmal (wenn auch selten) denkt man eher in Winkeln statt in Objektfeldabmessungen. In diesem Fall lässt sich der entsprechende Objektwinkel eben schlicht mittels dem Tangens berechnen, also

$$\boxed{\tan \phi_{max} = y'_{max}/(2f') \qquad (2.5)}$$

2.2 Abbildung mit zwei Linsengruppen

Die bisher dünn angenommene Linse ist in der Praxis meist eine Kombination aus mehreren Linsen bzw. ein *Objektiv*, denn mit einer einzelnen Linse können wir ausgedehnte Objekte nur mit stark eingeschränkter Qualität abbilden. Erfreulicherweise können wir aber auch komplexe Objektive in erster Näherung sehr gut als eine Einzellinse betrachten.

Was ist aber bei einem entsprechenden Linsensystem die Bezugsebene von der z.B. die Objektweite oder auch die Brennweite anzugeben bzw. zu berechnen sind? Letztlich sind zwei Bezugsebenen, die sogenannten *Hauptebenen*, notwendig. Man denkt sich das komplexe optische Systeme mit den vielen Flächen an denen Lichtbrechung stattfindet, ersetzt durch zwei einzelne Ebenen, an denen die gesamte Brechung stattfinden soll. Das ist natürlich eine extreme Näherung, sie ist aber für die Auslegung 1. Ordnung eines Systems ausreichend.

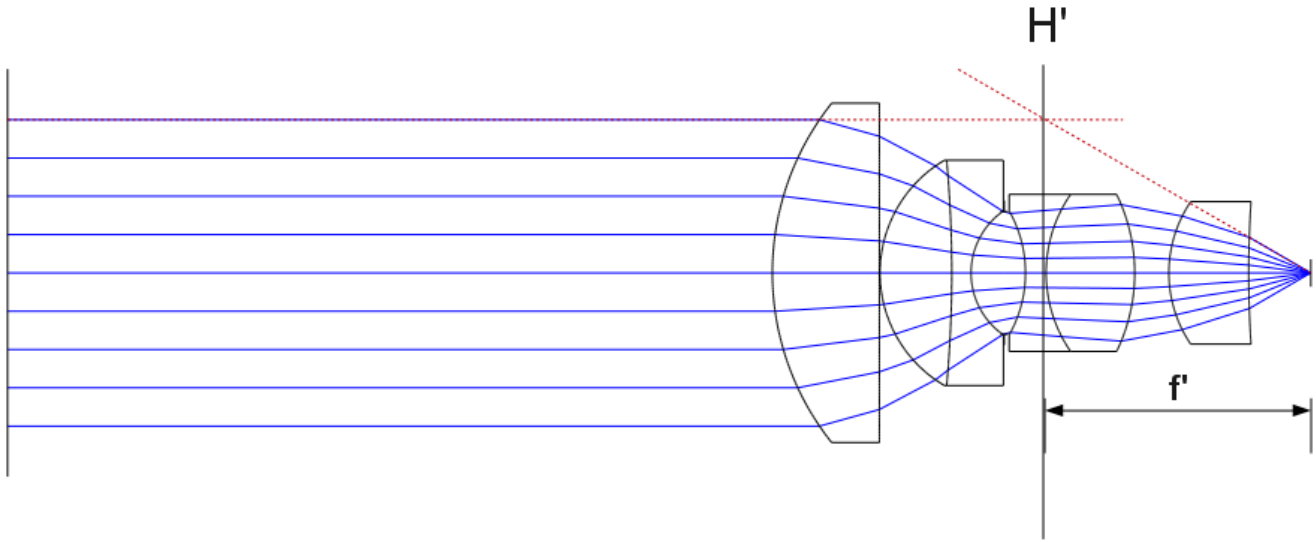


Abbildung 2.2.7: Hauptebenen sind Orte „äquivalenter Brechung", d.h. die vielen Brechungen des Lichts innerhalb eines komplexen optischen Systems werden durch eine einzelne virtuelle Brechung bzw. Strahlumlenkung an einer Ebene ersetzt. Hier dargestellt ist die Konstruktion zur Ermittlung der bildseitigen Hauptebene H'.

Abbildung 2.2.7 zeigt die bildseitige Hauptebene für ein Abbildungsobjektiv. Parallel einfallende Strahlen werden in Richtung des zugehörigen Bildpunkts gelenkt

und treffen sich dort. Wir greifen einen der Strahlen heraus und verlängern ihn sowohl objektseitig als auch bildseitig (beide Strahlverlängerungen sind rot dargestellt). Dort wo sich beide schneiden können wir vom Ort der „äquivalenten Brechung" reden, d.h. der einfallende Strahl wird nun nicht mehr an den vielen realen Linsenoberflächen gebrochen sondern nur einmal, an dem Kreuzungspunkt von einfallendem und dem Richtung Bildebene laufenden Strahl. Dieser Punkt ist ein Punkt der bildseitigen Hauptebene. Genauso können wir auch die objektseitige Hauptebene grafisch konstruieren indem wir von parallelem Licht auf der Bildseite ausgehen.

Auf die Berechnung der Position dieser Ebenen wollen wir hier nicht näher eingehen (siehe z.B. [31, 44]), für kommerzielle Objektive ist die Lage der Ebenen in den Datenblättern (teilweise) angegeben.

Für die Anwendung in der industriellen Bildverarbeitung ist es zunächst wesentlich zu wissen, dass die objektseitigen Abstände, also eben die Objektweite, von der objektseitigen Hauptebene aus gemessen werden und die bildseitigen Größen, also z.B. die bildseitige Brennweite f' und die Bildweite a', von der bildseitigen Hauptebene aus angegeben werden. Die dünne Linse spaltet gedanklich also in eine Box auf, die durch die objektseitige und die bildseitige Hauptebene begrenzt wird (Abb. 2.2.7). Der (virtuelle) Strahlverlauf zwischen den beiden Hauptebenen ist parallel zur *optischen Achse* (siehe Abb. 2.2.8).

Wir wollen im folgenden zwei zusammengesetzte optische Systemgeometrien, die in der industriellen Bildverarbeitung oft eingesetzt werden, besprechen. Das Teleobjektiv und das Teleskop.

Das *Teleobjektiv* dient dazu, einen vergleichsweise großen Abbildungsmaßstab auch bei kurzer Baulänge des optischen Systems (inklusive Bildaufnehmer) zu erreichen. Wenn bei einer konventionellen Linse ein großer Arbeitsabstand in Zusammenhang mit einem relativ großen Abbildungsmaßstab erzielt werden soll, dann sind aufgrund von Gleichung 2.1 auch die Bildweite a' und wegen Gleichung 2.2 die Brennweite groß. Eine bauliche Verkürzung ergibt sich aber, wenn die bildseitige Hauptebene künstlich in Richtung des Objekts verschoben wird, denn sowohl bildseitige Brennweite als auch Bildweite sind von dieser Ebene aus zu messen. Erreicht wird dies durch den Einsatz einer negativen Linse („Zerstreuungslinse", Brennweite $f' < 0$) vor der Bildebene.

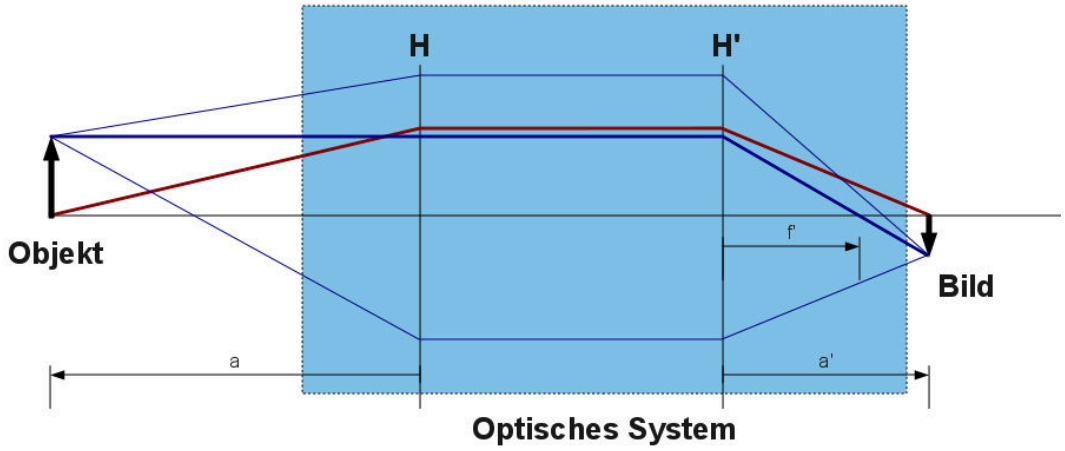


Abbildung 2.2.8: *Wenn die beiden Hauptebenen eines Systems bekannt sind, dann kann der paraxiale Strahlverlauf zu und von der Optik konstruiert werden. Objekt-, Bild- und Brennweite sind von den Hauptebenen aus zu messen.*

Oft werden auch verschiedene Abwandlungen von sogenannten *afokalen Systemen* bzw. Teleskopen zur Abbildung eingesetzt. Die typische Anordnung ist das *Kepler-Teleskop*. Umgangssprachlich wird mit „Teleskop" ein System verbunden, das weit entfernte Objekte vergrößert darstellt. In der Tat ist so eine Vergrößerung für das menschliche Sehen mit dem Kepler-Teleskop möglich. Für die industrielle Bildverarbeitung ist aber die Anwendung zur Abbildung kleiner und naher Objekte wichtiger. Abb. 2.2.10 zeigt die grundsätzliche Geometrie. Eine erste Linse bzw. Linsengruppe liegt im Abstand der Brennweite vom Objekt. Dadurch wird ein Bild des Objekts ins Unendliche projiziert. Die zweite Linsengruppe bildet dieses Zwischenbild aus dem Unendlichen dann in die Brennebene ab. Als Abbildungsmaßstab erhält man wieder mit dem Strahlensatz

$$\beta' = -\frac{f_2'}{f_1'} \tag{2.6}$$

Der Einsatz eines solchen Systems hat diverse Vorteile: Es kann auf vergleichsweise preisgünstige und einfache Art und Weise ein *doppelt-telezentrisches System* realisiert werden (vgl. Abschnitt 2.5), indem man den Abstand d zwischen beiden Linsengruppen gerade gleich der Summe der Brennweiten wählt, also $d = f_1 + f_2$, und an der Position der Brennweitenkreuzung die Systemblende des optischen Systems einbringt. Nur wenn der Abstand entsprechend gewählt wird, spricht man

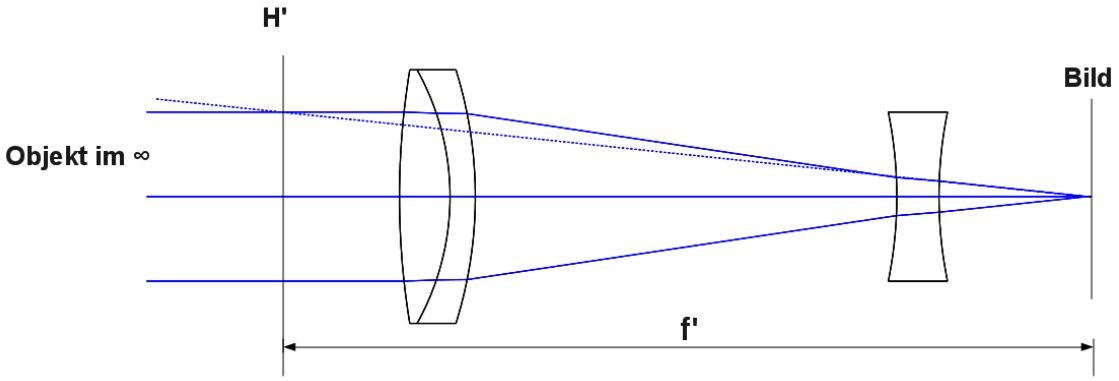


Abbildung 2.2.9: Teleobjektiv: Die Brennweite wird länger als die Baulänge des Objektivs. Erreicht wird dies durch eine Zerstreuungslinse (negative Brennweite f') vor der Bildebene, denn diese verschiebt die bildseitige Hauptebene vor die erste Linsengruppe.

im engeren Sinn von einem Teleskop, denn nur dann kann man das System auch zur Vergrößerung weit entfernter Objekte einsetzen. Außerdem ergibt die Anordnung mit zwei Objektiven letztlich eine um den Faktor 2 vergrößerte numerische Apertur (vgl. Abschnitt 2.3). Ein dritter Vorteil ist, dass sehr viele kommerziell verfügbare Standard-Optiken (z.B. typische Fotoobjektive, Standard-Achromate) für den Fall des Objekts im Unendlichen (also parallelen Lichteinfall) optimiert sind.

Beispiel: Sie wollen ein telezentrisches System mit $\beta' = -0.1$ aufbauen. Für die zweite Linse soll eine verhältnismäßig kurze Brennweite eingesetzt werden, um die Baulänge gering zu halten. Sie wählen f_2' = 10 und damit ergibt sich direkt f_1' = 100 mm.

Moderne Mikroskope nutzen letztlich dieselbe Grundanordnung bei der die Brennweite der ersten Linse, des sogenannten Mikroskopobjektivs, sehr kurz gewählt ist (z.B. 2 mm) und die Brennweite der zweiten Linse, der sogenannten Tubuslinse eher lang ist. Dementsprechend ergibt sich ein großer Abbildungsmaßstab (vgl. Abb. 2.2.10 und Abschnitt 6.6).

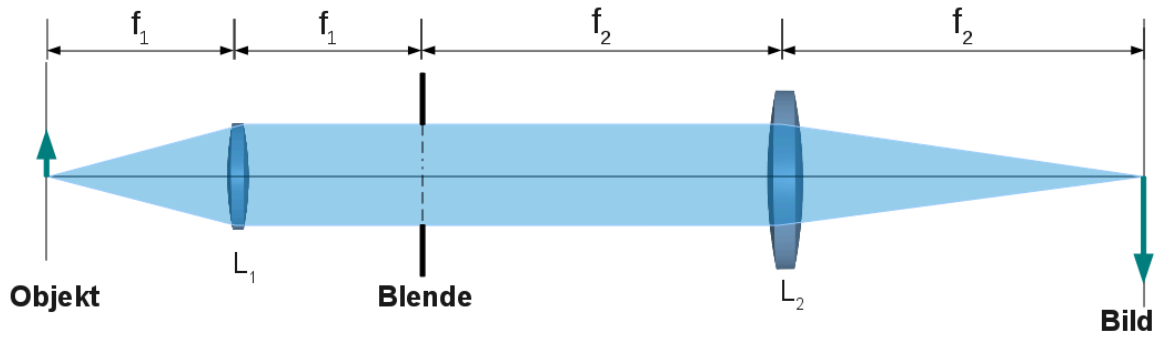


Abbildung 2.2.10: Telezentrische Abbildung durch ein Kepler Teleskop. Die Anordnung eignet sich sehr gut (wenn genügend Bauraum zur Verfügung steht), um mit zwei Standardobjektiven (beide auf unendlich korrigiert) eine gute bis sehr gute Abbildungsqualität zu erreichen. Diese Form der Abbildung ist außerdem die Grundgeometrie moderner Mikroskope. Dabei wird dann die erste Linse „Mikroskopobjektiv" und die zweite Linse „Tubuslinse" genannt.

2.3 Blende und Pupille

Der Raumwinkel des Lichtkegels, den wir von einem Objektpunkt erfassen ist einer der wichtigsten Parameter eines optischen Systems. Drei für uns ganz wesentliche Größen werden durch ihn beeinflusst:

- Aberrationen und Auflösung
- Lichtstärke[6], also die Menge des genutzten Lichts
- Schärfentiefe

Variieren können wir diesen Lichtkegel durch die Wahl des Durchmessers der *Blende*[7] (genauer *Aperturblende*) des optischen Systems. Sie ist dabei definitionsgemäß diejenige Komponente, die den von einem auf der Achse liegenden Punkt ausgehenden Lichtkegel am stärksten limitiert (siehe Abb. 2.3.11).

Man könnte zunächst meinen, dass die konkrete Lage der Blende uninteressant ist (Abb. 2.3.13). Solange lediglich ein Punkt auf der optischen Achse abgebildet wer-

[6]nicht zu verwechseln mit der Strahlstärke oder der Lichtstärke der Photometrie)

[7]Der Begriff „Blende" ist in diesem Zusammenhang üblich, aber etwas ungenau, denn Blende ist an sich der Überbegriff für verschiedene Blenden. Hier gemeint ist die Aperturblende. Es gibt auch Feldblenden, die das Objekt- bzw. Bildfeld begrenzen und Streulichtblenden, die Streulichtanteile vermindern. Wir werden — wie in der Bildverarbeitung üblich — im folgenden dennoch den Begriff „Blende" für die Aperturblende verwenden.

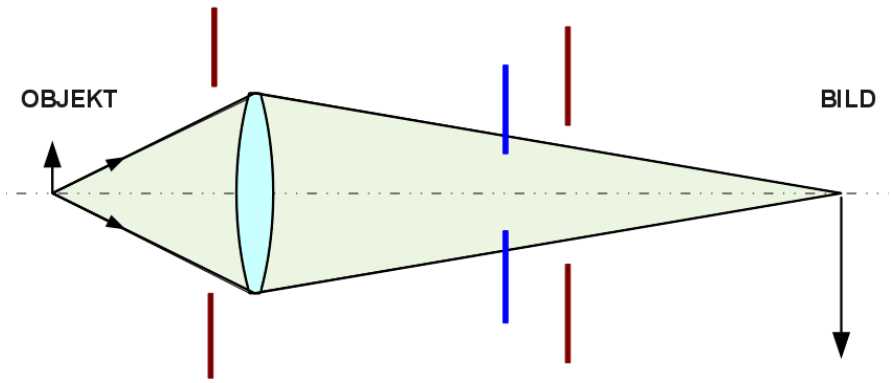


Abbildung 2.3.11: Eine Komponente limitiert den von einem auf der Achse liegenden Punkt ausgehenden Lichtkegel, der durch die Optik verarbeitet wird, am stärksten (hier blau eingezeichnet). Diese Komponente ist die Blende bzw. Aperturblende des Systems. Spezifiziert wird sie meist durch die Blendenzahl oder die numerische Apertur. Die Lichtkegellimitierung ist essentiell für die beugungsbedingte Auflösungsbegrenzung, die Schärfentiefe und die Lichteffizienz der Abbildung.

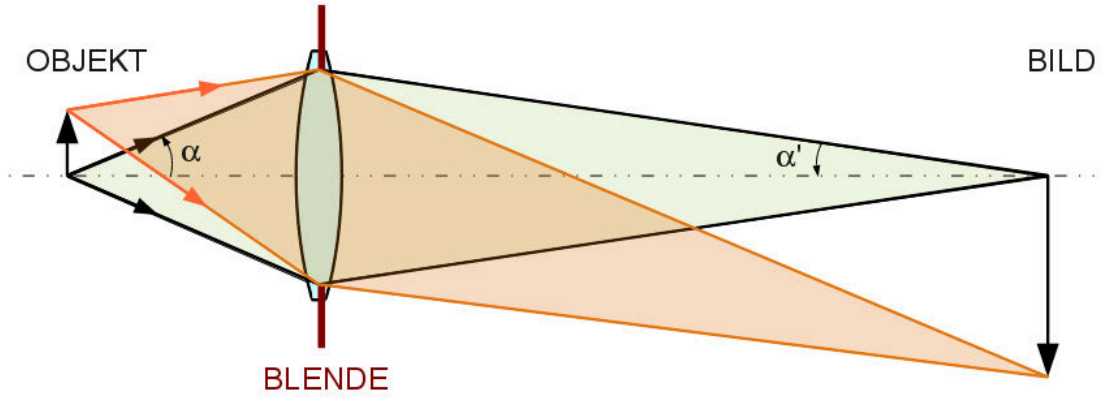
den soll ist dies auch zutreffend. Sobald aber ein ausgedehntes Feld mit mehreren Punkten relevant wird, ist die Lage sehr wichtig und damit ein entscheidendes Designkriterium. Abb. 2.3.12 zeigt für zwei verschiedene Blendenposition bei einer einfachen Abbildung mit einer Einzellinse, dass die Blendenposition sich maßgeblich auf den Lichtkegel unterschiedlicher Feldpunkte auswirkt. Dementsprechend variieren Lichtstärke, Auflösung, Aberrationen und Schärfentiefe mit dem Feld. Der Abfall der Lichtstärke mit dem Feld wird dabei als *Vignettierung* bezeichnet.

Den Raumwinkel des Lichtkegels geben wird im allgemeinen durch die sogenannte *numerische Apertur* an. Definiert ist diese in Vakuum (bzw. Luft) als der Sinus des halben Öffnungswinkels des Lichtkegels, also

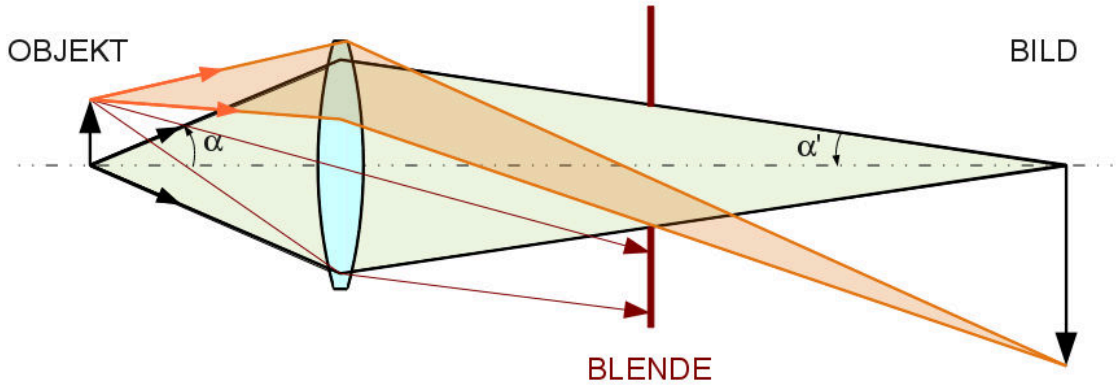
$$\boxed{\mathrm{NA} = \sin\alpha. \qquad (2.7)}$$

Diese Winkelangabe kann natürlich genauso im Bildraum erfolgen, also $NA' = \sin\alpha'$. Zwischen der numerischen Apertur NA im Objektraum und im Bildraum NA' gilt der einfache Zusammenhang

$$\mathrm{NA}' = \frac{\mathrm{NA}}{\beta'} \qquad (2.8)$$



(a) Blende bei der Linse



(b) Blende hinter Linse

Abbildung 2.3.12: Die Wahl der Blendenposition hat einen Einfluss auf den effektiven Lichtkegel der von jedem Feldpunkt genutzt wird. Je nach Blendenposition ergibt sich bei gleicher Blendenzahl (spezifiziert durch die in grün dargestellte Öffnung) ein unterschiedlicher Strahlkegel.

Die numerische Apertur zur Beschreibung der Öffnung erfasst in allen Situationen den wesentlichen Aspekt der Blende für die üblichen rotationssymmetrischen Abbildungssysteme. Als Maßzahl zur Beschreibung der Blende wird in der Fotografie und der Bildverarbeitung aber in aller Regel die *Blendenzahl* K (engl.: F-number oder F#) verwendet. K beschreibt für Anwendungen, bei denen das Objekt weit entfernt ist und somit die Bildweite der Brennweite entspricht, die bildseitige numerische Apertur. Definiert ist K für solche Fälle mittels

$$K = \frac{f'}{D}. \tag{2.9}$$

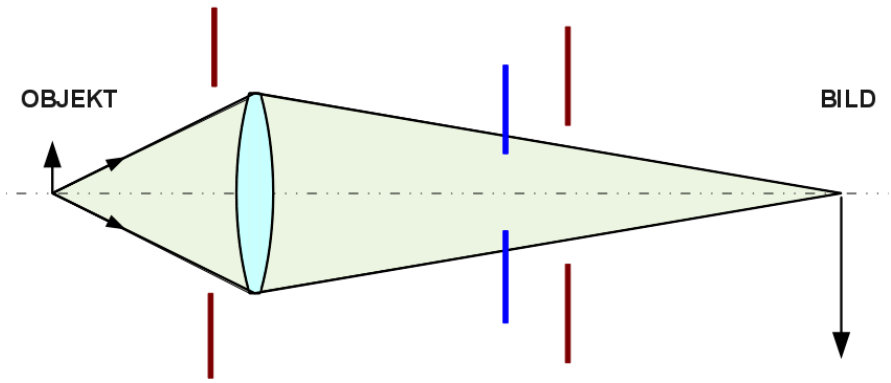


Abbildung 2.3.13: Betrachtet man nur einen auf der optischen Achse liegenden Objektpunkt, so ist die Lage der Blende irrelevant. Alle hier eingezeichneten Blenden haben exakt denselben Effekt.

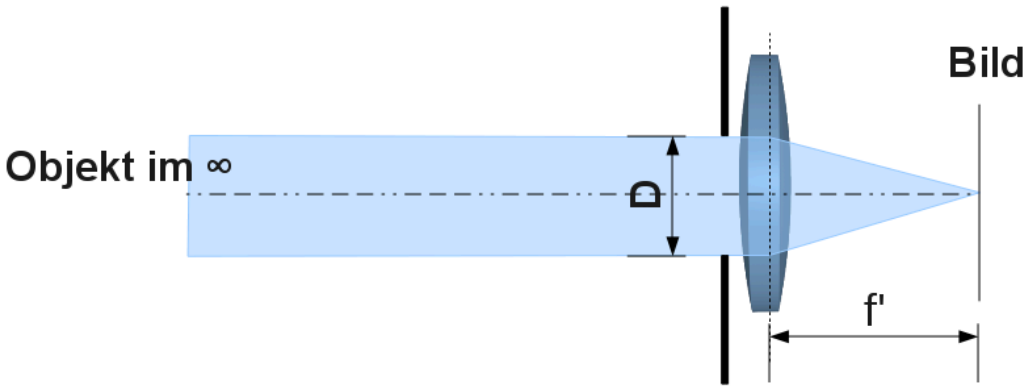


Abbildung 2.3.14: Die Blendenzahl K ist definiert als $K = f'/D$. Diese Definition ist nur sinnvoll für Objekte, die sich (näherungsweise) im Unendlichen befinden. Sie ist ein Maß für den bildseitigen Öffnungswinkel.

f' ist wieder die (bildseitige) Brennweite und D bezeichnet den „Durchmesser der Optik". Natürlich ist hier mit „Durchmesser" der effektiv wirksame, also der Licht sammelnde Durchmesser gemeint. Gedanklich denkt man sich das optische System also als eine Black-Box und der Eingang dieser Box ist die sogenannte *Eintrittspupille* der Optik (Abb. 2.3.15). Das klingt zunächst einfach, ist es aber letztlich nicht, denn die Lage und Größe dieser Eintrittspupille entspricht im allgemeinen nicht irgendeiner realen Komponente des optischen Systems sondern ist eine rein virtuelle Größe. Dazu aber gleich mehr.

Bei Objektiven werden die Blendenzahlen üblicherweise in einer mit $\sqrt{2}$ gestuften Reihe angegeben, also K=0.7, 1.0, 1.4, 2.0, 2.8, 4 usw., denn der Wechsel von

einer Blendenstufe zur nächsten (Faktor $\sqrt{2}$) führt zu einer Verdopplung der Lichtstärke.

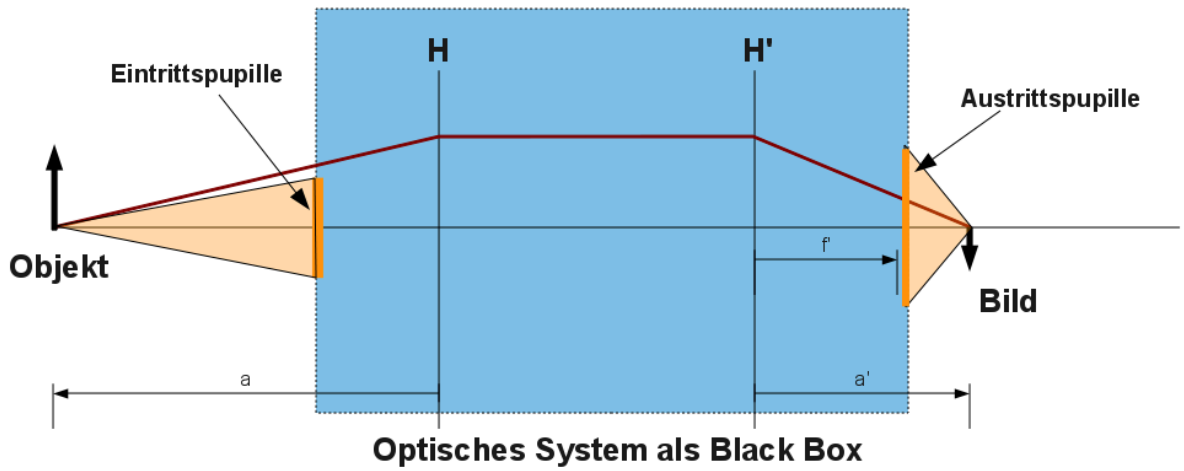


Abbildung 2.3.15: Hinsichtlich der Apertur und damit der vom System akzeptierten Strahlbündel bzw. der aus dem System austretenden Strahlbündel wird das System komplett durch die Eintritts– und die Austrittspupille beschrieben. Dies ist sowohl hinsichtlich des Strahlungsflusses als auch hinsichtlich der Wellenfrontdeformationen (Bildfehler) wesentlich. Man erspart sich so, für alle Betrachtungen das Innenleben des Systems zu analysieren. Beachten Sie: Durch die Hauptebenen wird der Strahlkegel nicht definiert. Beachten Sie außerdem: Die Pupillen können beliebig, also z.B. auch im Unendlichen, liegen. Auch die Hauptebenen können als die Ein- bzw. Ausgänge eines Black-Box Modells betrachtet werden. Es handelt sich aber um getrennte Modelle und entsprechend sind auch die Eingang- und Ausgangsports verschieden.

Für den geometrischen Optiker und den Optikdesigner ist ein vertieftes Verständnis der Pupillen und Blenden von großer Bedeutung. Neben dem Einfluss auf Lichtstärke, beugungsbedingte Auflösung und Schärfentiefe werden durch die Wahl der Blende bzw. der Pupillen auch die Perspektive (vgl. Telezentrie), Vignettierungen und Aberrationen wesentlich beeinflusst.

Für den Praktiker reicht es an dieser Stelle aus, sich klar zu machen, dass
a) letztlich der Raumwinkel der aufgefangenen oder zum Bildsensor fokussierten Strahlung entscheidend ist und dass wir
b) in der Mikroskopie diesen Winkel durch die numerische Apertur angeben während
c) für die meisten anderen Abbildungssysteme ($\beta' < 0.2$) meist die Blendenzahl zur Beschreibung verwendet wird.

Der Zusammenhang zwischen bildseitiger numerischer Apertur und Blendenzahl

wenn $|a| >> |a'|$ ist[8]

$$\sin\alpha' = \frac{D}{2\,f'} = \frac{1}{2K} \tag{2.10}$$

Wird das Objektiv nicht für eine Abbildung aus dem Unendlichen eingesetzt, wird teilweise die „effektive Blendenzahl" angegeben. Es gilt

$$K_{eff} := (1+\beta')K \approx K \tag{2.11}$$

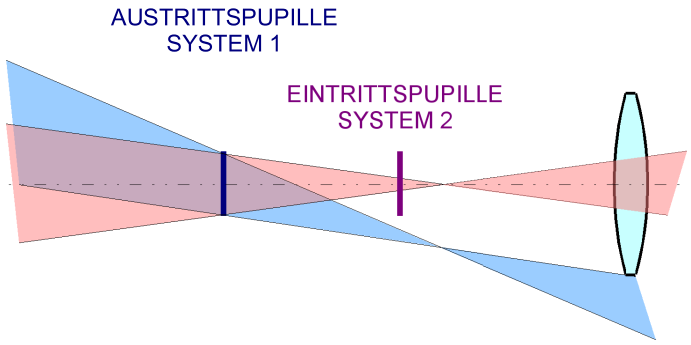
Wenn ein Hintereinanderschalten zweier optischer Teilsysteme ohne Lichtverlust erfolgen soll, ist darauf zu achten, dass die Eintrittspupille des zweiten Systems idealerweise in der Ebene der Austrittspupille des ersten Systems liegt und die Eintrittspupill des nachgeschalteten Systems mindestens so groß wie die Austrittspupille des ersten Systems ist (Abb. 2.3.16). So lässt sich ein Verlust an Licht und Information (Auflösung) vermeiden.

Es kann also wichtig sein, die Position und Größen der Pupillen von optischen Systemen zu berechnen. Letztlich ist dies vergleichsweise einfach möglich, wenn der Aufbau der Optik bekannt ist. Die *Eintrittspupille* ist das Bild der Blende des Systems durch die optischen Komponenten, die *vor* der Blende liegen (vgl. Abb. 2.3.17). Wenn Sie in eine Optik mit dem bloßen Auge schauen, dann *sehen* Sie genau diese Eintrittspupille.[9] In der Ebene der Eintrittspupille ergibt sich also ein scharfes Bild der Systemblende und die Größe des Bildes ist eben genau die Größe der Eintrittspupille.

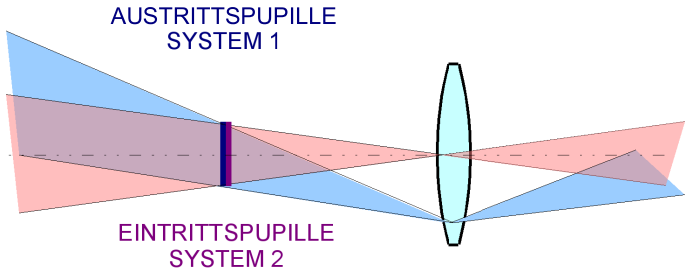
Ganz entsprechend ist die *Austrittspupille* das Bild der Systemblende durch die Optik, die *nach* Blende liegt und Sie können diese Austrittspupille sehen, wenn Sie von der anderen Seite her kommend — also vom Ausgang — in die Optik schauen.

[8] Bei der Betrachtung von Abb. 2.3.14 könnten Sie zunächst auf die Idee kommen, dass damit zwischen dem Winkel α' und der Blendenzahl K ein Zusammenhang gemäß dem Tangens besteht ($\tan\alpha = D/2/f'$) und daraus dann $\sin\alpha$ berechnen. Dies ist ein Trugschluss denn für reale Systeme, die hinsichtlich Bildfehler korrigiert sind, die Pupillen keine Ebenen sondern Kugeln mit den Objekt– bzw. Bildpunkten als Kugelmittelpunkt sind

[9] Die Blende selbst können Sie ja gar nicht *direkt* sehen, denn Linsen befinden sich zwischen Ihnen und der Blende.



(a) Die Eintrittspupille von System 2 fällt nicht in die Austrittspupille von System 1. Es ergibt sich Vignettierung.



(b) Wenn beide Pupillen aufeinander fallen und dieselbe Größe aufweisen wird Vignettierung vermieden.

Abbildung 2.3.16: Bei der Kombination von Systemen sollten die Pupillen der beiden Systeme aufeinander fallen.

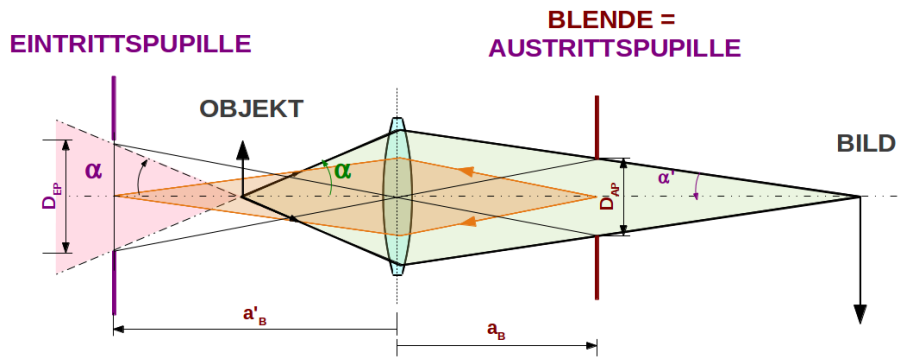
2.4 Schärfentiefe

Natürlich führt eine eng zugezogene Blende (entspricht großer Blendenzahl K) zu einer geringen Lichteffizienz. Vom Objekt ausgehendes Licht wird nur wenig genutzt. Wenn dies der einzige Einfluss der Blende wäre dann würde man natürlich immer eine stark geöffnete Blende (kleine Blendenzahl) verwenden.

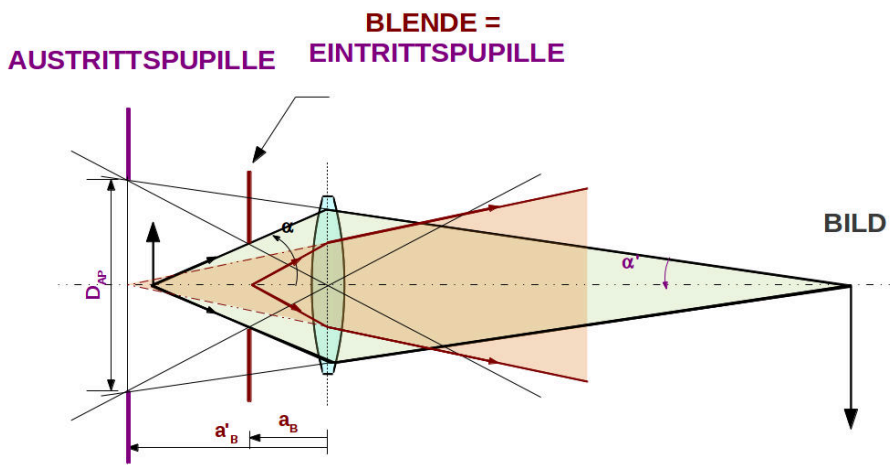
Eine große Öffnung bedeutet neben dem Vorteil der Lichtnutzung aber auch eine Verringerung der *Schärfentiefe* der Abbildung sowie meist auch eine Verringerung der praktisch erzielten Auflösung aufgrund von Bildfehlern. Auf die starke Zunahme der Bildfehler, die aus einer Vergrößerung der Öffnung resultieren, werden wir in Abschnitt 5.5 eingehen.

Die Schärfentiefe im Objektraum Δz ist also umso größer, je kleiner der Abbildungsmaßstab ist.

Dass die Schärfentiefe durch eine Vergrößerung der Blende verringert wird, ist



(a) Eintrittspupille



(b) Austrittspupille

Abbildung 2.3.17: Berechnung der Pupillen. Die Eintrittspupille ist das Bild der Aperturblende durch die vor der Aperturblende liegende Optik. Entsprechend ist die Austrittspupille das Bild der Aperturblende durch die nach der Aperturblende liegende Optik.

unmittelbar aus Abb. 2.4.18 einsichtig. Es zeigt sich (siehe z.B. [44]), dass bei einer Scharfstellung auf die Objektebene (Objektweite a) Objekte im Bereich zwischen a_v und a_h scharf abgebildet werden. Der (geometrische) *Schärfentiefenbereich* $\delta z = a_v - a_h$ kann durch eine einfache geometrische Überlegung (siehe z.B. [48] und Abb. 2.4.19) näherungsweise zu

$$\delta z \approx 2K\, B'/\beta'^2 \tag{2.12}$$

berechnet werden.

Dabei bezeichnet B' den erlaubten Durchmesser des (defokussierten) Bildes eines

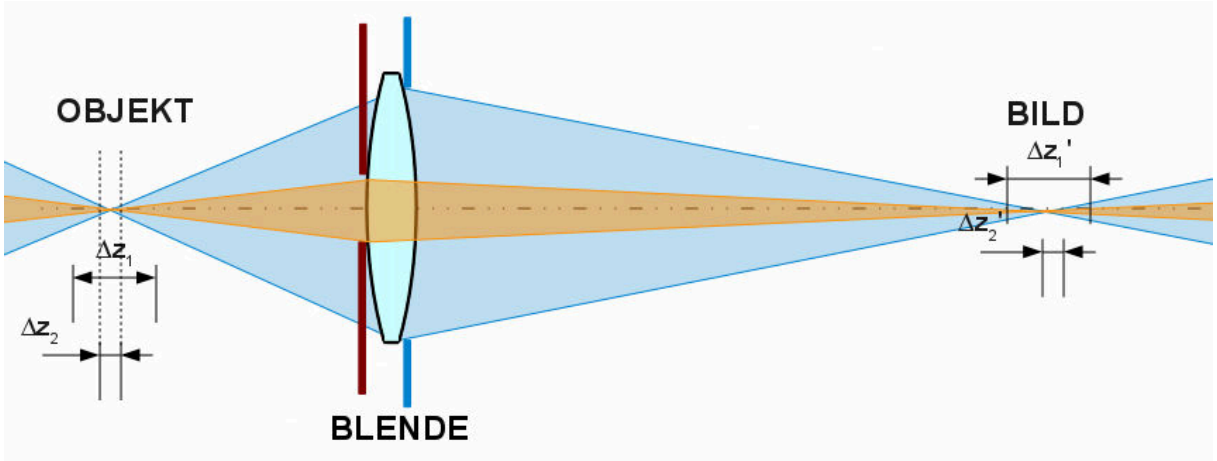


Abbildung 2.4.18: Je kleiner die numerische Apertur bzw. je größer die Blendenzahl (roter Strahlkegel), desto größer ist der Schärfentiefenbereich der noch zu einer akzeptablen Unschärfe (Unschärfescheibchen B') führt (Δz_2). Die Schärfentiefe kann sowohl im Objektraum als auch im Bildraum angegeben werden. Die Beziehung zwischen beiden Größen ist durch $\Delta z' = \beta'^2 \Delta z$ gegeben.

idealen Objektpunkts.

Nahdistanz und *Ferndistanz* sind die Entfernungen, bei denen wir noch ein scharfes Bild erhalten:

$$a_{v/h} = \frac{\beta' \cdot a \cdot D_{EP}}{\beta' \cdot D_{EP} \pm B'} \tag{2.13}$$

Für hohe numerische Aperturen ist dieser rein geometrische Wert allerdings wenig geeignet. Wenn man von einer maximalen Wellenfrontdeformation (Peak-to-Valley) aufgrund der Defokussierung (vgl. Kapitel 5.5) von $\lambda/4$ ausgeht, ergibt sich ein (wellenoptischer) Schärfentiefenbereich von

$$\boxed{\delta z' = \frac{\lambda}{2\mathrm{NA}'^2}} \tag{2.14}$$

Beispiel: Berechnen Sie die Schärfentiefe für ein Abbildungssystem mit $\beta' = -0.1$ bei einer Blendenzahl von 5. Der erlaubte Durchmesser des defokussierten Punktbildes darf 2 Pixel (10 μm) betragen.
Lösung: $\delta z = 2 \cdot 5 \cdot 0.01/0.01 = 10$ mm

Wir können mit Gl. 2.13 auch die *Nahdistanz* und *Ferndistanz*, bei der noch ein ausreichend scharfes Bild erzielt wird, berechnen.

Mit $\beta' = -0.1 = a'/a \approx f'/a = f'/1\,m$ ergibt sich f' = 100 mm. Mit K=5 folgt $D_{EP} = 20$ mm und damit a_h= 1.005 m und a_v = 0.995 m und somit wie erwartet wieder ein Schärfentiefenbereich von 10 mm.

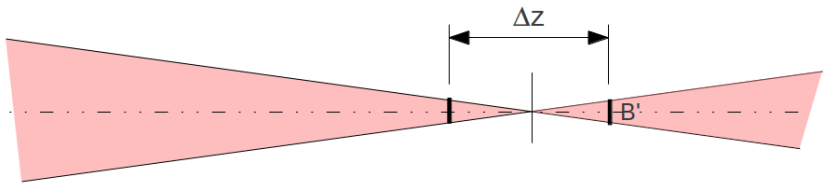


Abbildung 2.4.19: Der Schärfentiefenbereich ist linear abhängig vom erlaubten Unschärfekreisdurchmesser B' und der Blendenzahl.

Auch das menschliche Auge nutzt die Möglichkeit der vergrößerten Schärfentiefe durch Abblenden: Wenn wir Objekte sehr nah (und damit groß) betrachten, dann wird unsere Pupille, die die Blende des Auges ist, verkleinert. Dadurch wird der Schärfentiefenbereich der Abbildung vergrößert und wir können das zu betrachtende Objekt noch näher an das Auge bringen und es somit noch größer (es gilt ja $\beta' = a'/a$) betrachten.

Für die praktische Bestimmung der Schärfentiefe gibt es normierte keilförmige Objekte[10] mit feiner Strukturierung, die es erlauben die Schärfentiefe sehr exakt zu messen.

2.5 Perspektive

Für die industrielle Anwendung hat die bisher dargestellte Art der Abbildung oftmals einen gewissen Mangel. Nach Gleichung 2.1 ändert sich bei einer axialen Verschiebung ($a \rightarrow \tilde{a}$) nämlich leider — abgesehen von der Bildschärfenproblematik — auch der Abbildungsmaßstab β'. Wenn also aus dem Bild eine geometrische Größe bestimmt werden soll, dann führt eine axiale Verschiebung des Objekts ty-

[10] z.B. bei Edmund Scientific erhältlich, aber auch leicht selbst bastelbar

pischerweise zu einem Messfehler (siehe Abb. 2.5.20).

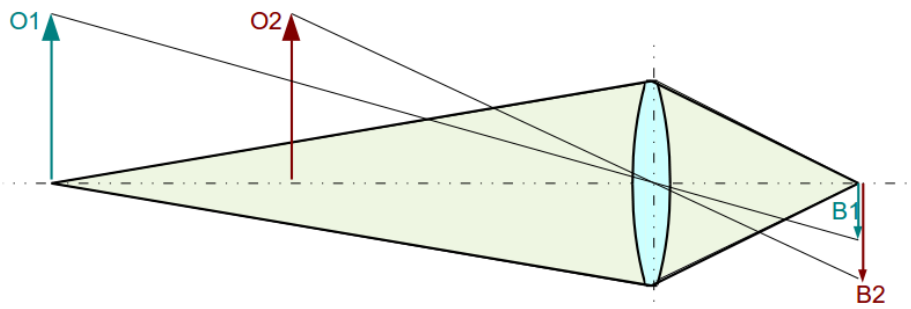


Abbildung 2.5.20: In aller Regel bilden optische Systeme gemäß der Zentralperspektive eine dreidimensionale Szene ab. Insbesondere werden weit entfernte Objekte mit einem geringeren Abbildungsmaßstab abgebildet und haben so ein entsprechend kleines Bild, also $\beta' = a'/a$

Die Tatsache, dass bei (konventionellen) Abbildungen weit entfernte ($|a|$ groß) Objekte stark verkleinert werden, ist für das menschliche Sehen ein Vorteil, denn wir sind so in der Lage, gleichzeitig eine große, ausgedehnte Szene zu überschauen und nahe Objekte im Detail beurteilen zu können. Für die (messende) Bildverarbeitung ist die Abstandsabhängigkeit des Abbildungsmaßstabs aber in aller Regel unerwünscht, denn einerseits macht sie Messungen in verschiedenen Ebenen schwieriger und andererseits führt sie oft zu komplexeren Bildern, die schwieriger zu verarbeiten sind (s.u.).

Zur Lösung der Problematik wird eine (oder mehrere) der drei folgenden Maßnahmen ergriffen:

- Sicherstellung (Mechanik), dass keine Variation der Objektweite auftritt (z.B. durch einen Anschlag). Dies ist natürlich nur möglich, wenn das Objekt selbst nicht eine große Tiefenausdehnung aufweist.
- Algorithmische Rückrechnung auf die wahre Objektgröße: Hierzu muss die real vorliegende Entfernung des Objekts in irgendeiner Form gemessen oder geschätzt werden. Durch geeignete vorangegangene Kalibrierung und dann nachfolgenden Ausgleich der Entfernungsabhängigkeit kann wieder auf die reale Objektgröße rückgerechnet werden.
- Einsatz einer Optik mit sogenannter (objektseitiger) *telezentrischer Perspektive*.

Wir wollen hier auf den letzten dieser Punkte, die telezentrische Perspektive, näher eingehen. Diese Art der Abbildung ist im täglichen Leben zwar unüblich, aber in der industriellen Bildverarbeitung aufgrund der bereits beschriebenen Problematik der Tiefenvariation des Abbildungsmaßstabs von enormer praktischer Bedeutung. Sofern eine messende Anwendung oder anderweitig komplexe Aufgabenstellung vorliegt und ein Einsatz von telezentrischen Objektiven möglich ist, wird man in aller Regel eben solche Optiken verwenden. Mikroskopobjektive sind nahezu immer objektseitig telezentrisch.

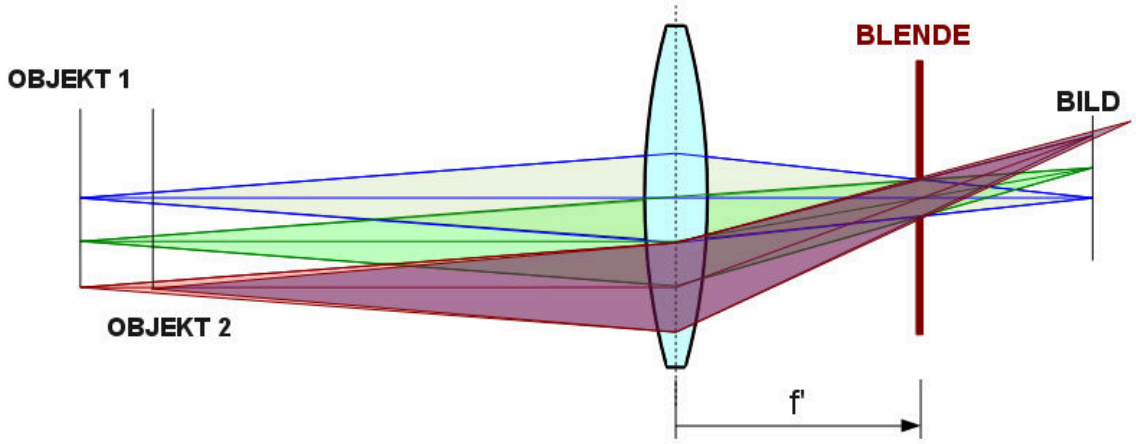


Abbildung 2.5.21: Objektseitig telezentrische Perspektive. Der Schwerpunkt des Strahlenbündels ist der Hauptstrahl. Dieser läuft im Objektraum offensichtlich parallel zur optischen Achse. Eine Verschiebung des Objekts entlang der optischen Achse (violett dargestellt für tiefsten Feldpunkt) ändert am Verlauf des Hauptstrahls nicht. Die Abbildung des entsprechenden Objektpunkts wird zwar unscharf, der Schwerpunkt des unscharf abgebildeten Punkts bleibt aber unverändert. Erreicht wird diese telezentrische Perspektive durch die Position der Blende in der Brennebene des ersten Linsengruppe (bzw. durch Positionierung der Eintrittspupille im Unendlichen).

Abb. 2.5.21 zeigt eine typische *objektseitig telezentrische Abbildung*. Entscheidend, ist die Position der Blende (vgl. Abschnitt 2.3). Bei der telezentrischen Abbildung wird die Blende genau in die Brennebene des Systems gebracht. Der mittlere Strahl durch die Blende läuft daher im Objektraum *parallel* zur optischen Achse. Dieser Strahl ist von besonderem Interesse, denn er ist quasi der „Schwerpunkt" (begrifflich genauer: „Schwerstrahl") des Strahlenbündels, das von der Optik aufgefangen wird. Alle anderen Strahlen des Strahlenbündels sind leicht anders geneigt, aber die mittlere Strahlneigung des Strahlenbündels ist eben Null bzw. parallel

zur optischen Achse orientiert.

Was passiert nun, wenn das Objekt leicht axial verschoben wird? Natürlich wird die Abbildung unscharf, aber diese leichte Unschärfe muss kein Problem in der Anwendung darstellen, denn trotzdem können z.B. Schwerpunkte in der Bildebene sehr genau bestimmt werden (Faustregel: auf ca. 1/10 Pixel genau). Da der Schwerstrahl des Strahlenbündels parallel zur Verschiebung läuft, ändert sich die Position des Auftrefforts des Schwerstrahl in der Bildsensorebene nicht. Nach wie vor wird der Objektpunkt an exakt derselben Position auf dem Bildsensor registriert! Trotz Variation der Entfernung bleibt das Bild also gleich groß, ein für messtechnische Anwendungen oftmals immenser Vorteil.

Wie gut der Strahlengang telzentrisch ist wird durch den *Telezentriefehler* angegeben. Dieser sagt aus, wie stark der Schwerstrahl für einen bestimmten Feldpunkt von der perfekten Parallelität zur optischen Achse der Abbildung abweicht. Wenn sich also z.B. für einen Punkt ein Telezentriefehler von 1 mrad ergibt dann bedeutet das, dass ein axiale Verschiebung des Objekts um 10 mm eben einer lateralen Objektpunktverschiebung von $0.001 \cdot$ 10 mm = 10 μm entspricht.

Da der Schwerstrahl im Objektraum parallel verläuft spricht man von objektseitiger Telezentrie. Entsprechend gibt es auch eine *bildseitige Telezentrie*, bei der der entsprechende Strahl eben im Bildraum parallel läuft. Dies hat sensorseitig, insbesondere bei Sensoren mit Mikrolinsen, Vorteile. Hat man gleichzeitig sowohl objekt- wie auch bildseitige Telezentrie spricht man von beidseitiger Telezentrie.

Aus der Konstruktion ergibt sich aber leider auch, dass das abzubildende Objekt kleiner als der Durchmesser der Frontlinse (genauer: Eintrittspupille, vgl. Abschnitt 2.3) sein muss. Dies ist eine erhebliche Einschränkung, denn größere Objekte lassen sich daher leider nur mit großen (und dementsprechend teuren) Objektiven telezentrisch abbilden. Einfach aufgebaute telezentrische Objektive sind im Vergleich zu ihrem Arbeitsabstand (und im Gegensatz zu nicht-telezentrischen Objektiven) vergleichsweise lang.

Aber auch für nicht-messende Aufgaben kann die Telezentrie deutlich vereinfachend wirken. Abb. 2.5.22 zeigt die konventionelle sowie eine telezentrischen Perspektive eines Bauteils. Bei Verwendung der telezentrischen Perspektive wird die Szene deutlich einfacher. Es sind weniger Linien zu beachten, die Analyse ist

viel weniger anspruchsvoll. Insbesondere ist zu beachten, dass gleiche Objektteile (hier z.B. Schlitze), je nachdem an welcher Position sie liegen, bei konventioneller Perspektive unterschiedlich im Bild erscheinen.

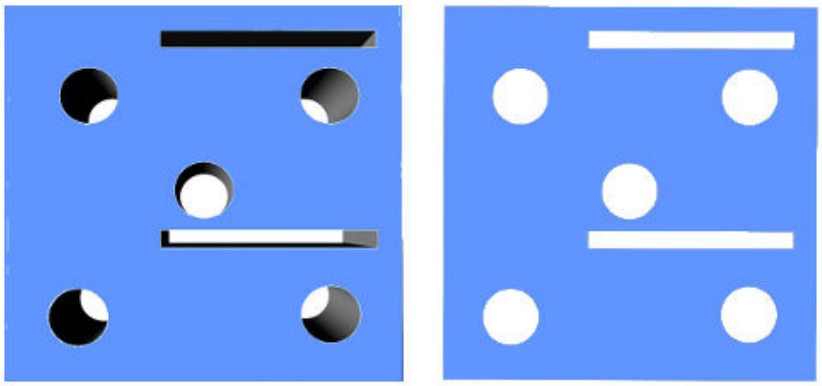


Abbildung 2.5.22: Gegenüberstellung der konventionellen entozentrischen (links) und der telezentrischen (rechts) Abbildung eines Bauteils mit Bohrungen und Schlitzen. Die telezentrische Abbildung liefert direkt die Aufsicht im Sinne einer Parallelprojektion und ist deutlich einfacher zu verarbeiten (weniger Kanten, weniger relevante Punkte).

Und selbst wenn das Objekt keine Tiefenausdehnung aufweist werden in der Praxis gerne telezentrische Objektive eingesetzt. Zunächst sollte man in diesem Fall keinen Vorteil erwarten. In der Praxis werden telezentrische Objektive aber eben speziell für messende Anwendungen optimiert und haben daher vergleichsweise geringe Verzeichnungen (vgl. Abschnitt 5.5). Die Kalibrierung ist dementsprechend oftmals stark vereinfacht.

Neben der Zentralperspektive (*entozentrische Perspektive*) und der telezentrischen Abbildung wird manchmal auch eine *hyperzentrische Abbildung* (bzw. *perizentrische Abbildung*) benötigt. Dies ist dann der Fall, wenn man gleichzeitig die Stirnfläche und die parallel oder aufeinander zu laufenden Seitenfläche eines Körpers abbilden will (siehe Abb. 2.5.23). Weit entfernte Teile werden also vergrößert dargestellt. Entscheidend ist hierbei, dass die Eintrittspupille aus Sicht der Optik vor oder bei dem Objekt, das abgebildet werden soll, liegt. Allerdings ist die Optik aus rein geometrischen Gründen (vgl. Abb. 2.5.23) deutlich größer als die Objektausdehnung zu wählen und die Abbildung der Seitenwände erfolgt verzerrt.

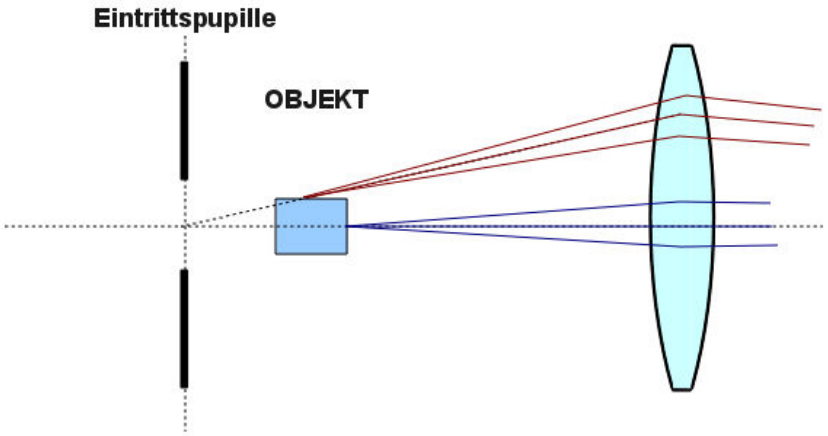


Abbildung 2.5.23: Hyperzentrische Perspektive. Die Eintrittspupille liegt „vor" dem Objekt. Daher können gleichzeitig die Seitenwände (verzerrt) und die Deckfläche abgebildet werden.

2.6 Zusammenfassung wesentlicher Parameter

Für die Spezifikation einer Abbildung müssen Sie sich zunächst über einige Grundparameter der Anwendung und den sich daraus ergebenden Folgen klar werden. In jedem Fall müssen Sie festlegen:

Objektfeld bzw. Objektvolumen: Welches Feld muss von der Abbildung erfasst werden und welcher Schärfentiefenbereich ist notwendig?

Objektentfernung (Arbeitsabstand): Gibt es Einschränkungen hinsichtlich des Abstands von Objekt zur Optik, insbesondere den minimal erlaubten Abstand (*Minimal object distance*) ?

Bild- bzw. Sensorgröße: Welche Sensorgröße wird angestrebt?

Brennweite: Die Brennweite ist der für uns zunächst wichtigste Parameter der Abbildungsoptik. Für einen sich aus Objektfeld und Bildgröße ergebenden Abbildungsmaßstab und der Objektentfernung kann die Brennweite berechnet werden.

Blendenzahl: Die Blenden ist der zweitwichtigste Parameter der Abbildungsoptik. Große Blendenzahlen führen zu hohen Schärfentiefen aber auch zu ge-

ringer Lichteffizienz. Außerdem hat die Wahl der Blende einen wesentlichen Einfluss auf die Auflösung der Abbildung (vgl. Kapitel 5).

Perspektive: Falls möglich sollte bereits bei der Grundauslegung der Abbildungsgeometrie klar sein, ob eine konventionelle, eine telezentrische oder eine hyperzentrische Abbildung eingesetzt werden soll.

Zusätzliche Anforderungen: Falls zusätzliche, besondere Anforderungen an die Abbildung bestehen (z.B. Zoom, spezielle Wellenlängenbereiche, Farbe, Preis), sollten diese bereits bei der Grundauslegung der Optik protokolliert und ggf. berücksichtigt werden.

Oft sind verschiedene Varianten möglich. Z.B. ändern sich für die gleiche Aufgabenstellungen natürlich Abstände und Brennweiten, wenn eine andere Sensorgröße verwendet wird. Unter Umständen sollten Sie also verschiedene Varianten an Grundauslegungen berücksichtigen. Und natürlich kann es bei der Auslegung eines Gesamtsystems auch vorkommen, dass die Grundauslegung der Abbildung iterativ mehrmals an das Gesamtsystem angepasst werden muss. Unter Umständen zeigt sich später, dass mehrere Kameras statt einer eingesetzt werden sollen und dementsprechend ändern sich unter Umständen z.B. die Objektfelder. Am besten (auch im Sinne einer Protokollierung) ist also wie immer die Auslegung in einer einfachen skriptbasierten Rechenumgebung (z.B. Octave, Matlab).

Erst wenn die oben dargestellten Parameter (zumindest näherungsweise) klar sind, kann eine konkrete Auswahl kommerzieller Komponenten erfolgen. Die konkrete Wahl wird in Kapitel 6 dargestellt. Zunächst müssen wir allerdings die zur näheren Spezifikation notwendigen Grundlagen in den folgenden drei Kapiteln betrachten.

Kapitel 3

Mathematische Beschreibung der Bildentstehung und das Abtasttheorem

Abbildung 3.0.1: Farbmoire durch nicht ausreichende Abtastung (Aliasing)

In diesem Kapitel werden Sie unter anderem lernen

- wie die Antwort linearer, verschiebungsinvarianter Systeme beschreiben wird,

- wie lineare Filter im Ortsraum und im Ortsfrequenzraum beschrieben werden,
- warum und wie die Basiszerlegung eines Bilds in harmonische Funktionen (Fourier-Transformation) und in abgetastete Punkte (Pixel) vorgenommen wird,
- wie zwischen dem Ortsraum und dem Ortsfrequenzraum gewechselt wird und
- was bei der Abtastung eines Bilds beachtet werden muss.

Dazu werden wir uns mit etwas mathematischer Theorie beschäftigen müssen. Es gibt sehr wenige Stellen in der industriellen Bildverarbeitung wo dies notwendig wird. Aber ohne die lineare Systemtheorie ist es kaum möglich, ein wirkliches Verständnis der Beschreibung der Optik, aber auch der Algorithmik von Bildverarbeitung zu erlangen. Es lohnt sich an dieser Stelle daher, in eine gewisse theoretische Basis zu investieren. Für uns essentiell sind dabei die Fouriertransformation, das Abtasttheorem sowie die Beschreibung der Bildqualität mittels der Modulationstransferfunktion (MTF).

Lineare Systeme bzw. lineare Filter sind für uns in der Bildverarbeitung aus zweierlei Gründen entscheidend. Zum einen werden die für uns wesentlichen Komponenten der Bildentstehung durch lineare Systeme bzw. Filter beschrieben, aber auch viele der für die Bildverarbeitung wesentlichen Filter sind linear bzw. es wird der Ausgang von verschiedenen linearen Filter dann kombiniert.

Weil die lineare Systemtheorie so vielseitig ist und in so vielen verschiedenen Bereichen von Bedeutung ist, werden Studierende der Ingenieurs- oder Naturwissenschaften vermutlich auf die eine oder andere Weise bereits mit dem Gebiet in Berührung gekommen sein, meist im Sinne einer Beschreibung von Zeitsignalen. In der Optik begegnet uns die Fouriertransformation ebenfalls bei Zeitsignalen (z.B. Spektralzerlegung durch Prisma oder Gitter). Für die Theorie der Abbildung und die Algorithmik der Bildverarbeitung sind aber zweidimensionale, räumliche Fouriertransformationen von Belang, und hierauf wollen wir uns im folgenden vor allem konzentrieren.

Wir starten mit einer kleinen systemtheoretischen Betrachtung, um näher zu verstehen, warum Fouriertransformationen so wichtig für die Beschreibung von Systemen sind bevor wir dann sehen, wie immens wichtig sie vor allem in der Optik und in der Bildverarbeitung sind.

3.1 Lineare, invariante Systeme

In Abb. 3.1.2 ist ein System abstrakt dargestellt. Das System liefert ein Ausgangssignal o wenn es mit einem Eingangssignal i beaufschlagt wird. „System" kann dabei sowohl ein aus Hardware bestehendes System (z.B. ein Objektiv) oder auch ein Algorithmus sein. In der Regel werden die Signale von bestimmten Variablen (z.B. der Zeit, der Frequenz oder dem Ort) abhängen. Die Signale selbst können ganz verschiedene Größen sein, z.B. Spannungen oder Intensitäten.

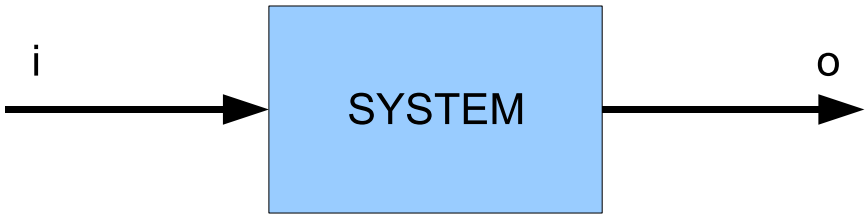


Abbildung 3.1.2: Black–Box Modell

Eine besondere Bedeutung haben *lineare Systeme*. Ein lineares System liefert einfach die Summe einzelner Ausgangssignale wenn es mit der Summe von Eingangssignalen beaufschlagt wird:

$$i_1 \rightarrow o_1 \tag{3.1}$$

$$i_2 \rightarrow o_2 \tag{3.2}$$

$$\Rightarrow$$

$$i_1 + i_2 \rightarrow o_1 + o_2 \tag{3.3}$$

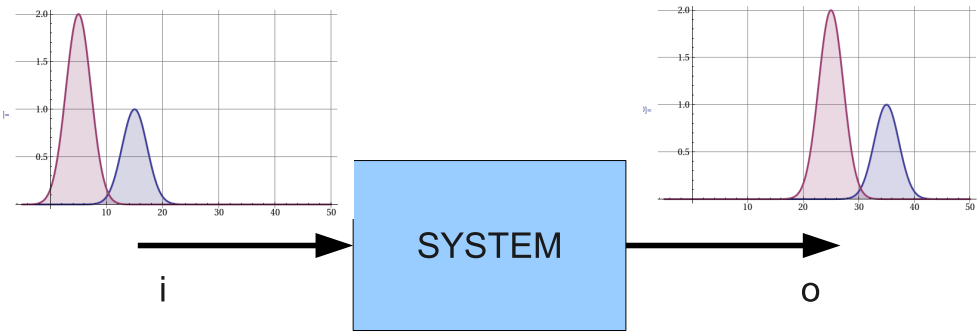
D.h. die Signal i_j werden unabhängig voneinander durch das System geleitet.

Daraus ergibt sich (noch etwas allgemeiner ausgedrückt)

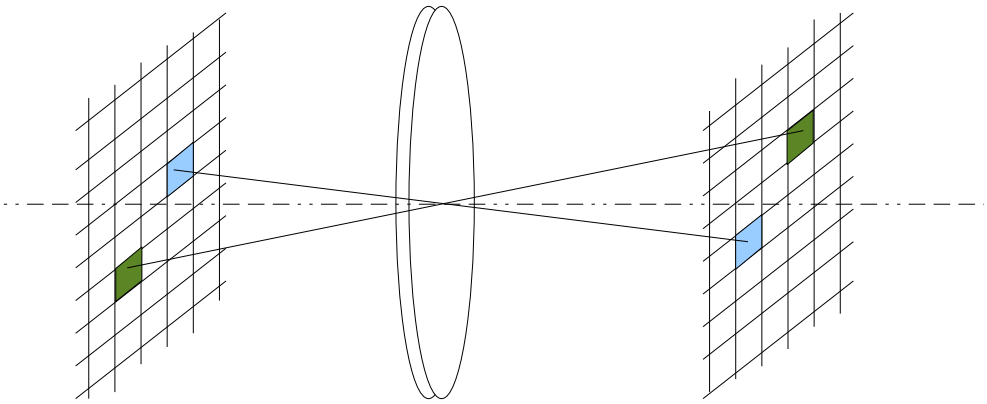
$$\boxed{a_1\, i_1 + a_2\, i_2 \rightarrow a_1\, o_1 + a_2\, o_2} \tag{3.4}$$

mit den Konstanten a_1 und a_2.

Sehr viele praktische Systeme sind linear oder werden in erster Näherung durch linearisierte Systeme ausreichend gut beschrieben. Nehmen Sie als Beispiel ein Telefon: Wenn Sie gleichzeitig sprechen und Klavier spielen, dann wird am anderen Ende der Leitung die Summe beider Signale „Klavier + Sprache" den Lautsprecher verlassen. Erst wenn das Gesamtsignal zu groß wird, ist die Linearität nicht mehr gegeben („Übersteuerung").



(a) Pulsübertragung durch elektrisches System



(b) optische Abbildung

Abbildung 3.1.3: Lineare, verschiebungsinvariante Systeme

In der Optik tritt die *Linearität* in vielfältiger Weise auf. Z.B. werden bei einer Abbildung die einzelnen Objektpunkte unabhängig voneinander durch das Abbildungssystem abgebildet. Man präsentiert also quasi am Eingang eine Summe elementarer Signale (z.B. einzelne Objektpunkte) und erhält am Ausgang die Summe aller entsprechender Bildpunkte. Grundsätzlich ergibt sich die Linearität der Optik aus den elementaren Grundgleichungen der Elektrodynamik[1], den Maxwell-Gleichungen. Da diese linear sind, sind auch die aus ihnen abgeleiteten Gleichun-

[1]Die Optik ist klassisch eine Untermenge der Elektrodynamik.

gen der Optik linear.[2]

In vielen Fällen hat man es darüber hinaus mit sogenannten *invarianten Systemen* zu tun. Das Telefon kann wieder gut als anschauliches Beispiel für ein zeitinvariantes System dienen: Zwei exakt gleich klingende Worte, im Abstand einer Sekunde gesprochen, werden am Ausgang ebenfalls den Abstand 1 Sekunde aufweisen und sich hinsichtlich des Klangs (bzw. der Signalform) nicht unterscheiden. Allgemeiner ausgedrückt: Das System liefert ein Ergebnis, das invariant hinsichtlich einer (zeitlichen) Verschiebung ist. Beachten Sie: Die Invarianz bezieht sich **NICHT** darauf, dass eine Sekunde Unterschied am Eingang auch eine Sekunde am Ausgang bewirkt. Vielmehr ist entscheidend, dass das Signal zu den beiden unterschiedlichen Zeiten gleich bleibt.

Bei der optischen Abbildung haben wir entsprechend eine (näherungsweise) Invarianz gegenüber *räumlichen* Verschiebungen.[3] Der Abstand zwischen den Bildpunkten kann durchaus vom Abstand zwischen den Objektpunkten differieren[4]. Die Form der Bildpunkte (ausgedrückt z.B. durch die Punktbildverwaschungsfunktion, vgl. Kapitel 4) ist aber näherungsweise konstant und gleiche Abstände im Objektraum werden in gleiche Abstände im Bildraum transformiert (s. Abb. 3.1.7).

Wir wollen zunächst etwas näher auf die Folgen eingehen, die sich *ganz allgemein* für die Beschreibung eines linearen–invarianten Systems ergeben.

Hierzu zerlegen wir eine beliebige Eingangsfunktion, die auf unser System trifft in unendlich viele unendlich kurze ideale Peaks, d.h. sogenannte δ–Funktionen[5]. Die δ–Funktion ist dabei einerseits unendlich kurz ausgedehnt, andererseits aber auch unendlich hoch so dass die Fläche unter der δ–Funktion letztlich 1 ergibt, also $\int_{-\infty}^{\infty} \delta(x)dx = 1$. Wie dem auch sei: Anschaulich zerlegen wir unser Objekt in unendlich viele Einzelpunkte, die wir getrennt abbilden. Weil wir ein lineares System annehmen, können wir dann am Systemausgang die Summe der Systemantworten auf die einzelnen Punkte bilden.

[2]Natürlich bedeutet das nicht, dass keine nichtlinearen Ausdrücke in Gleichungen auftreten. Linearität meint hier die Linearität in der oben dargestellten Art für Systeme.

[3]Das Gebiet in dem diese näherungsweise Konstanz gilt wird üblicherweise „isoplanatische Region" genannt.

[4]Nur bei einer 1:1 Abbildung sind die Abstände im Bildraum exakt gleich den Abständen im Objektraum.

[5]Mathematisch ist diese Sprechweise extrem lax. Es müsste eigentlich von δ–Distributionen gesprochen werden. Wir werden hier aber — unserer allgemeinen Fahrlässigkeit folgend — auf eine genauere (mathematisch korrekte) Beschreibung mit Grenzwerten zu Gunsten der Anschaulichkeit und Einfachheit verzichten.

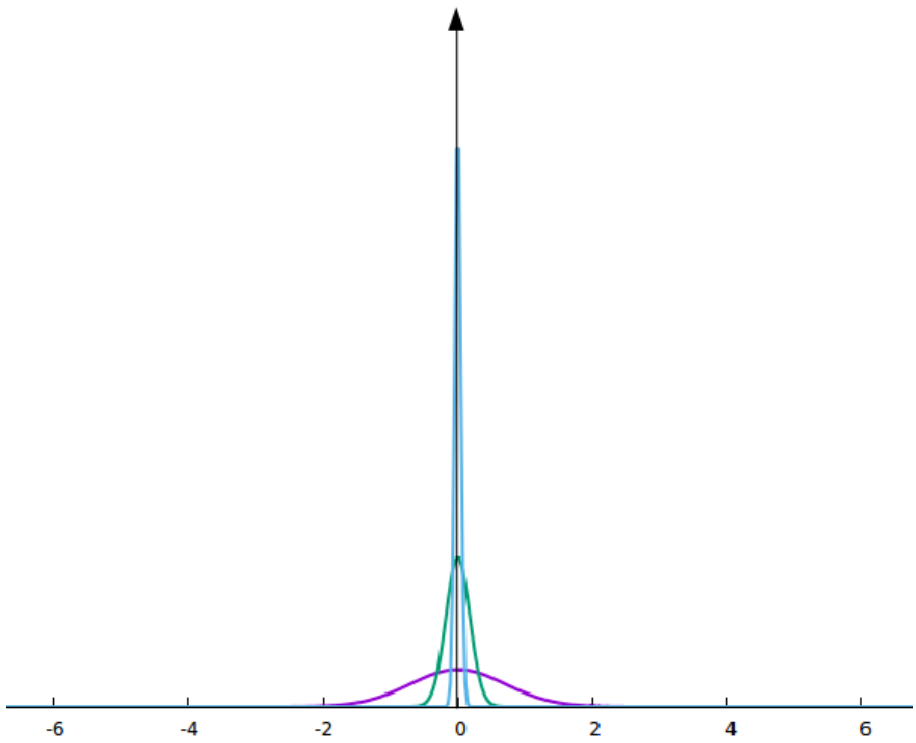


Abbildung 3.1.4: Veranschaulichung der Delta-Funktion als Grenzwert von Gaußfunktionen mit abnehmender Breite (aber konstanter Fläche): Der Wert der Deltafunktion ist an der Stelle x=0 unendlich und an allen anderen Stellen ($x \neq 0$) Null.

Natürlich sind die δ–Funktionen bei unserer Basiszerlegung gegeneinander verschoben. Für Objekte (bzw. Bilder) ist dies besonders anschaulich. In diesem Fall kann man sich die einzelnen δ–Funktionen ja einfach als ideale Punkte in der Objektebene bzw. (unendlich feine) Pixel vorstellen[6]. Wir zerlegen also gemäß Gl. 3.13 bzw. Abb. 3.1.5

$$i(x) = a_0\,\delta(0) + a_1\,\delta(x-1) + a_2\,\delta(x-2) + ... = \sum_j a_j\,\delta(x - j\,\delta x_s). \tag{3.5}$$

Die a_j sind bei einem solchen Objekt einfach die Helligkeiten der entsprechenden Objektpunkte s an den Stellen $x_s \cdot j$.

Wenn wir die *Antwort* auf einen elementaren δ-Impuls an der Stelle 0 mit $h(x)$ be-

[6]Für Zeitsignale ist $\delta(t)$ entsprechend ein Zeitpunkt.

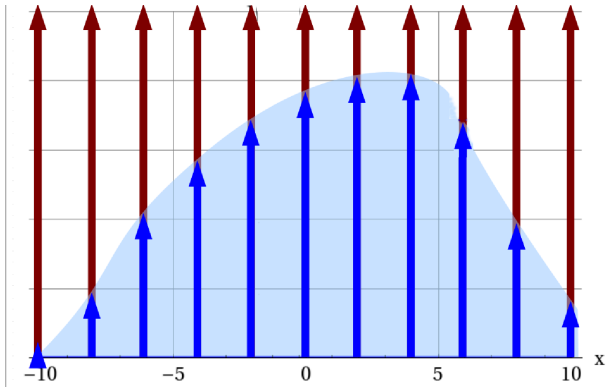


Abbildung 3.1.5: Zerlegung mit Deltas. Die Multiplikation einer kontinuierlichen Funktion (hellblau) mit einem Deltakamm (rot) führt zu einer Diskretisierung nach Gleichung 3.5

.

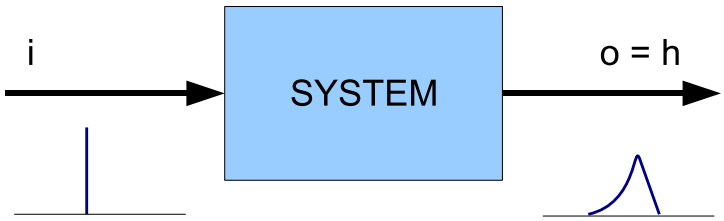
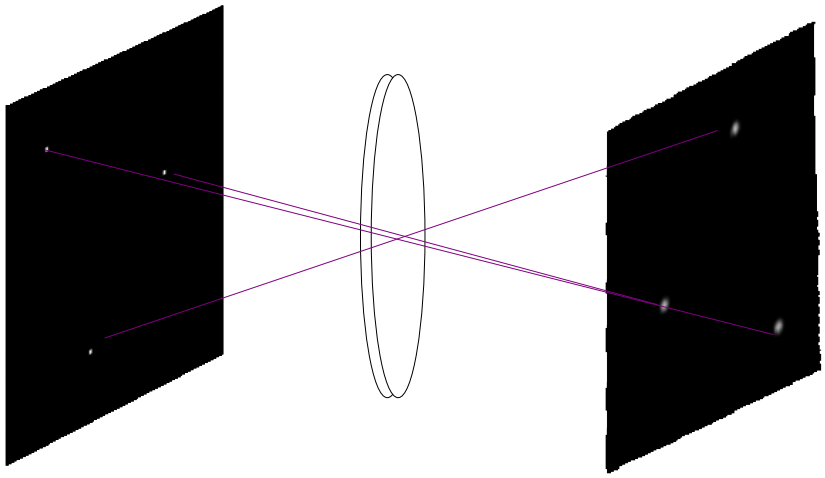


Abbildung 3.1.6: Black–Box Modell

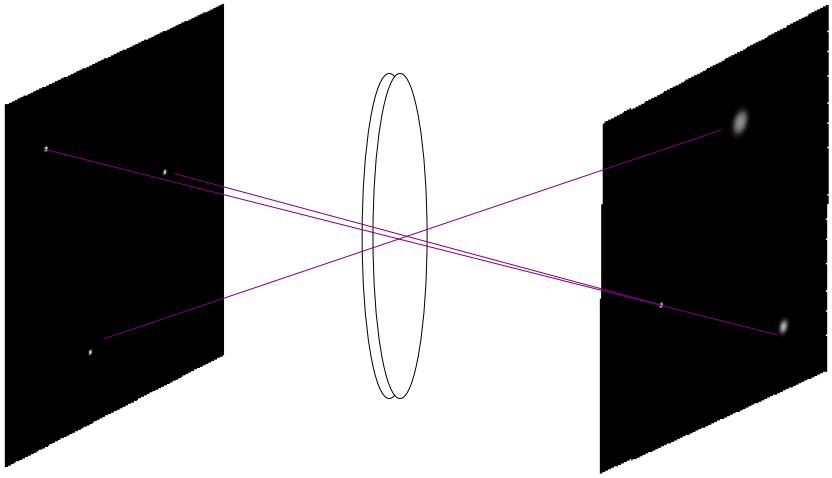
zeichnen, also $\delta(0) \to h(x)$, und fordern, dass eine Verschiebung am Eingang auch nur zu einer Verschiebung am Ausgang führt (invariantes System) dann ergibt sich für das lineare System am Ausgang (also z.B. im Bild)

$$\boxed{o(x) = a_0\, h(x) + a_1\, h(x-1) + a_2\, h(x-2) + \ldots = \sum_j a_j\, h(x-j) \qquad (3.6)}$$

Hierbei ist das Sampling am Eingang und Ausgang jeweils „1". Natürlich könnten wir hier auch beliebig andere Abtastungen verwenden.



(a) Invarianz gegenüber räumlicher Verschiebung, Isoplanasie



(b) Varianz gegenüber räumlicher Verschiebung

Abbildung 3.1.7: *Invarianz der Antwort bei räumlicher Verschiebung (Verschiebungsinvarianz)*

Für den kontinuierlichen Fall mit unendlich vielen Objekt– und Bildpunkten wird die Summe natürlich durch ein Integral ersetzt und die Gleichungen 3.4 und 3.6 werden zu

$$i(x) = \int_{-\infty}^{\infty} i(x')\,\delta(x-x')dx' \tag{3.7}$$

$$o(x) = \int_{-\infty}^{\infty} i(x')\,h(x-x')dx', \tag{3.8}$$

dabei haben wir die bisher diskreten Koeffizienten a durch die „kontinuierlichen Koeffizienten" (sprich: eine Funktion) i ersetzt.

Beide Gleichungen sind sogenannte *Faltungsintegrale*. Sie sind so gebräuchlich, dass es sich lohnt, eine eigene Schreibweise für die Faltung einzuführen:

$$\boxed{f(x) * g(x) := \int_{-\infty}^{\infty} f(x')g(x-x')dx'. \tag{3.9}}$$

Gleichung 3.8 wird damit zu dem wichtigen Ergebnis

$$\boxed{o(x) = i(x) * h(x), \tag{3.10}}$$

d.h. das Ausgangssignal eines invarianten linearen Systems bzw. Filters ergibt sich durch Faltung des Eingangssignals mit der Systemantwort $h(x)$ auf einen δ–Impuls. Als Merkregel: „Jedes (lineare,zeitinvariante) Filter ist ein Falter."

D.h. all unsere Systemkomponenten und Algorithmen in der Bildverarbeitung werden — solange sie näherungsweise als linear verschiebungsinvariant betrachtet werden können — durch Faltungen beschrieben. Natürlich sind nicht alle Algorithmen der Bildverarbeitung linear. Aber die linearen Algorithmen bilden, wenn man so will, praktisch den Kern der Theorie und viele *nichtlinearen Algorithmen* (z.B. der Kantendetektion) basieren letztlich wieder auf der nichtlinearen Kombination von linearen Filtern.

Wenn wir also das Verhalten $h(x)$ am Ausgang des (linearen invarianten) Systems für einen unendlich kurzen Eingangsimpuls kennen, sind wir in der Lage, die Antwort des Systems auf ein *beliebiges* Eingangssignal i zu berechnen. Wir wissen also alles über das System, wenn wir seine Antwort auf einen δ–Impuls kennen.

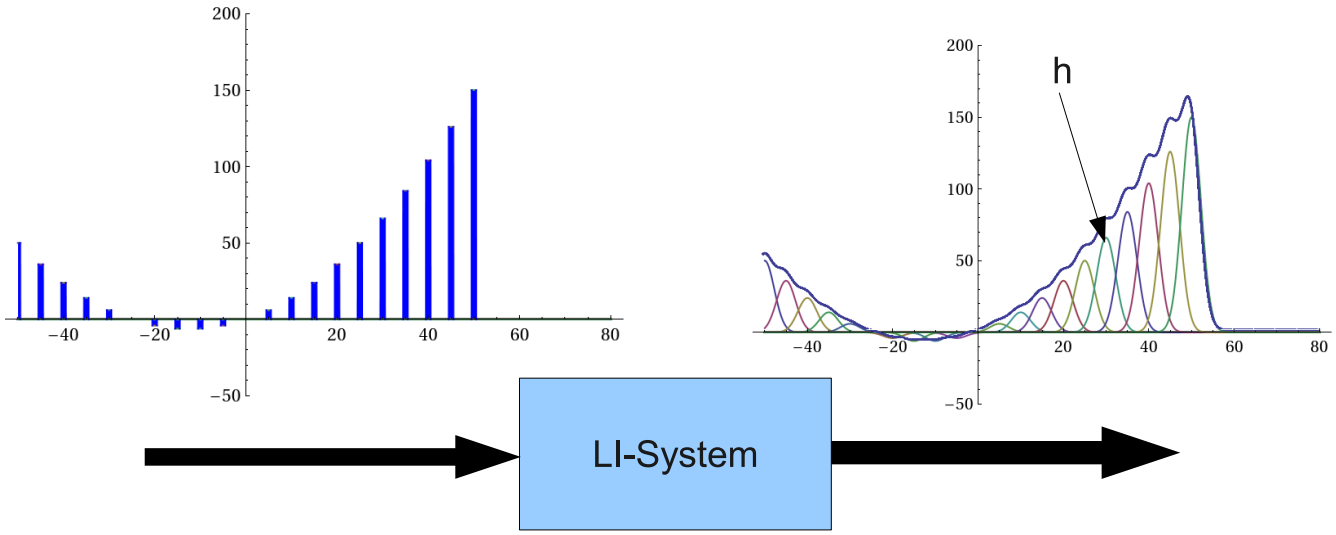


Abbildung 3.1.8: Das Verhalten des LIS auf einen unendlich kurzen Eingangspuls ist ausreichend zur vollständigen Beschreibung der Systemantwort auf ein beliebiges Eingangssignal. Man denkt sich das Eingangssignal als aus einer Summe gewichteter Delta-Funktionen bestehend. Am Ausgang resultiert entsprechend eine Summe (Linearität !) von gewichteten Punktbildfunktionen. Beachten Sie: Zur Visualisierung wurde hier als Eingangssignal ein diskretes Signal gewählt. Im Normallfall sind Ein- und Ausgangssignal aber kontinuierlich.

3.2 Eigenfunktionen und die Fouriertransformation

Betrachten Sie die eindimensionale Funktion $f(x) = \sin(x^2)$. Es mag Ihnen ganz natürlich erscheinen, dass dieses $f(x)$ *die* Funktion selbst ist. Um die Funktion beispielsweise an der Stelle $x = 2,5$ zu berechnen setzen wir einfach den Wert $x = 2,5$ in die Funktion ein und erhalten sofort $f(x = 2,5) = \sin(2,5^2) \approx -0,03318$. Soweit so gut.

Sie können zu demselben Ergebnis aber auch auf wesentlich kompliziertere Weise gelangen. Zum Beispiel könnten Sie $f(x)$ mittels

$$\sin(x^2) \approx x^2 - \frac{1}{6}x^6 + \frac{1}{120}x^{10} - .... \tag{3.11}$$

bestimmen.

Warum sollten wir diesen Umweg gehen? Eine mögliche Antwort ist sicher, dass wir damit eine Berechnung komplizierter Funktionen (hier: Sinus) vermeiden.[7] Das ist aber bei weitem noch nicht alles.

Man kann hier streiten, was die Funktion „ist". Letztlich wird der Sinus meist geometrisch definiert, aber das muss nicht so sein. Man kann ihn z.B. eben auch über eine Reihenentwicklung definieren.

Mathematisch betrachtet haben wir die zu berechnende Funktion mit Gl. 3.11 durch eine Summe aus gewichteten Elementarfunktionen (bzw. Basisfunktionen) $g_j(x)$ beschrieben, d.h.

$$f(x) = \sum_{j=0}^{j=N} a_j g_j(x) \tag{3.12}$$

$$= \sum_{j=0}^{j=N} a_j x^j \tag{3.13}$$

N muss dabei theoretisch Unendlich sein, um eine exakte Lösung zu erzielen. In der Praxis gibt man sich natürlich mit Näherungen zufrieden und bricht die Summation an geeigneter Stelle ab. Als Basisfunktionen haben wir hier $g_j(x) = x^j$ gewählt, denn diese Funktionen lassen sich sehr einfach berechnen (Arithmetik durch Addition und Multiplikation).

Wir können aber auch ganz andere Basisfunktionen verwenden. So haben wir im letzten Abschnitt z.B. verschobene Delta-Funktionen gewählt. Auch dort ist das Orginalsignal *die gewichtete Summe* vieler Basisfunktionen (verschobene Delta-Impulse).

Grundsätzlich sind viele solcher Systeme von Basisfunktionen möglich. Nur wenige sind allerdings wirklich praxisrelevant und ergeben einen direkten Nutzen.

[7] Und in der Tat rechnet Ihr Taschenrechner letztlich intern über eine entsprechende Reihendarstellung die Funktion aus.

Das System mit der größten Bedeutung ist das System *ebener Wellen* (bzw. harmonischer Schwingungen), das uns zur Fouriertransformation führt[8]

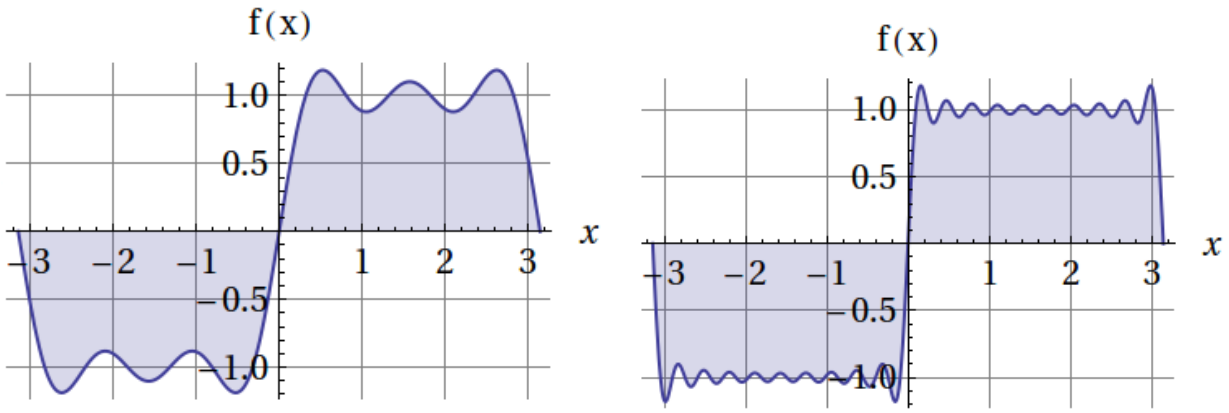
$$g_u(x) = e^{2\pi iux} = \cos(2\pi ux) + i\sin(2\pi ux) \tag{3.14}$$

Eine Funktion wird in diesem System eben durch die Summe

$$f(x) = \sum_{u=-\infty}^{\infty} a_u\, e^{2\pi iux} \tag{3.15}$$

dargestellt.

Ein Beispiel zeigt Abb. 3.2.9 bei der die Zerlegung einer periodischen Rechteckschwingung in Sinusschwingungen (der Imaginärteil von Gleichung 3.15) dargestellt ist. Bereits die Überlagerung (blaue Kurve) von 3 Sinusschwingungen nähert das Signal einigermaßen brauchbar an.[9]



(a) 3 Komponenten (b) 7 Komponenten

Abbildung 3.2.9: Durch Addition von vielen Sinusschwingungen passender Amplitude und Phase kann man nahezu jede Funktion nähern.

Die besondere Bedeutung der harmonischen Basisfunktionen liegt in ihrer Eigenschaft, *Eigenfunktionen* von linearen Systemen zu sein. D.h. beim Durchgang durch das lineare System behält das Sinussignal seine Form. Es kann zwar phasengeschoben werden und die Amplitude des Sinus darf sich ändern, aber es bleibt ein Sinus und auch die Frequenz des Sinus ändert sich nicht (siehe Abb. 3.2.10).

[8] Die verwendeten Konstanten in Gl. 3.14 wurden hier mehr oder weniger willkürlich gewählt.

[9] Es zeigt sich, dass eine reine Überlagerung von reellen Sinusschwingungen nur einen Teil aller in der Praxis auftretenden Funktionen beschreiben kann. Um wirklich alle (mathematisch gutartigen) Funktionen darstellen zu können benötigt man zusätzlich phasengeschobene Sinusfunktionen (bzw. den Kosinus) und eine imaginäre Gewichtung von einigen Funktionen. Oder aber man geht direkt zu der Zerlegung mit komplexen Exponentialfunktionen (Gl.3.15) über.

Das ist mit Sicherheit eine sehr außergewöhnliche Eigenschaft, denn im allgemeinen ändert sich eine Signalform beim Durchgang durch ein lineares System ganz erheblich. Z.B. wird ja aus einem ideal abgebildeten Punkt bei der optischen Abbildung nicht wieder ein idealer Punkt. Und weil die Antwort des Systems auf diese speziellen harmonischen Basisfunktionen so einfach ist, macht es Sinn, genau diese Basisfunktionen als Grundkomponenten zu verwenden, in die wir unsere Eingangssignale zerlegen.

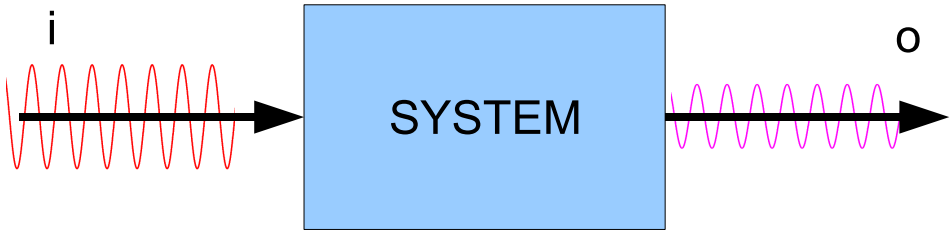


Abbildung 3.2.10: Sinus– und Kosinusschwingungen sind Eigenfunktionen von linear invarianten Systemen und werden daher beim Durchgang durch das System nicht verzerrt sondern lediglich abgeschwächt/verstärkt und eventuell phasengeschoben.

Mathematisch ist für die Eigenfunktion $i(x)$ zu fordern:

$$o(x) = h(x) * i(x) \stackrel{!}{=} k \cdot i(x) \tag{3.16}$$

$$= \int_{-\infty}^{\infty} i(x')h(x-x')dx' \tag{3.17}$$

Die Grundidee zur Berechnung der Antwort des Systems auf ein beliebiges Eingangssignal ist also folgende: Wir zerlegen das Eingangssignal in die Basisfunktionen $g_u = \exp(2\pi ux)$, berechnen dann die Antwort des linearen Systems auf die einzelnen Basisfunktionen (lediglich Multiplikationen mit einem konstanten Faktor $a(u) = A_u \cdot \exp(i\phi_u)$, und summieren schließlich diese Antworten.

Dadurch sparen wir uns die oft mühsame Berechnung der Faltung des Eingangssignals mit der Punktantwort.

Diese Vorgehensweise kann zum einen rein mathematisch vorteilhaft sein, aber teilweise können auch die Systemantworten auf die verschiedenen Basisfunktionen direkt messtechnisch erfasst werden. Ein wichtiges Beispiel ist die *Modulationstransferfunktion*, die wir noch ausführlicher zu besprechen haben. Letztlich werden dem System, z.B. einem Objektiv, dabei die Eigenfunktionen (also sinusförmige Gitter) angeboten. Das Bild ist (Eigenfunktion) wieder ein si-

nusförmiges Gitter, aber der Kontrast des Gitters nimmt mit zunehmender Gitterfrequenz ab.

Ganz konkret bedeutet diese Eigenschaft der Eigenfunktion, dass ein sinusförmiges Gitter auch z.B. bei Defokussierung oder anderen Aberrationen, die wir als lineares System bzw. Filter betrachten können, eine sinusförmige Modulation (mit verringertem Kontrast) ergibt.[10]

Wir können nun Gl. 3.15 für ein kontinuierliches Signal $f(x)$ verallgemeinern. Die Summe wird dabei zu einem Integral:

$$f(x) = \int_{-\infty}^{\infty} a(u)\, e^{2\pi i u x} du \tag{3.18}$$

Mathematisch betrachtet haben die verwendeten Variablen u und x noch überhaupt keine Bedeutung. x steht zwar normalerweise bei uns in der Bildverarbeitung für den Ort, aber letztlich können wir hier beliebige physikalische Größen verwenden, z.B. die Zeit, wie es beispielsweise bei der zeitlichen Veränderung einer elektrischen Spannung notwendig wäre.

In Gl. 3.18 haben wir zunächst ein uns unbekanntes $a(u)$ angesetzt. Die Frage ist natürlich, wie wir dieses $a(u)$ konkret wählen müssen, um ein ganz spezielles $f(x)$ nach Gl. 3.18 zu erzielen. Das geschieht eben durch die *Fouriertransformation*

$$\boxed{a(u) = \int_{-\infty}^{\infty} f(x)\, e^{-2\pi i u x} dx \qquad (3.19)}$$

Gleichung 3.19 ist die Fouriertransformation für räumliche Signale[11]. Gleichung 3.18 ist entsprechend die *inverse Fouriertransformation*. Wie man durch Einsetzen verifiziert gilt $\mathrm{FT}^{-1}[\mathrm{FT}[f(x)]] = f(x)$.

Teilweise werden unterschiedliche Normierungskonstanten bei der (inversen) Fouriertransformation verwendet. Wir verwenden hier die in der Optik übliche Nomenklatur nach Goodman [17]. Letztlich ergeben sich bei anderer Wahl der Normierung aber keine relevanten Änderungen der Mathematik.

[10] Eine Tatsache, die messtechnisch zuerst von H. Frieser 1936 genutzt wurde [37].

[11] Ganz entsprechend kommt man zu zeitlichen Signalen, indem man das x durch t sowie das u durch f ersetzt.

Wir haben damit formal die Fouriertransformation und ihre Umkehrung im Sinne einer (motivierten) Definition eingeführt. Die Basiszerlegung unseres Signals ist also eine inverse Fouriertransformation und die Koeffizienten unserer Zerlegung ergeben sich gerade durch die Fouriertransformation.

3.3 Die Fouriertransformation von Bildern

In der Bildverarbeitung interessieren wir uns meist für zweidimensionale räumliche Fourierzerlegungen. Das Fourierpaar der Zerlegung in Basisfunktionen (inverse Fouriertransformation) und Bestimmung der Zerlegungskoeffizienten F (Fouriertransformation) wird also geschrieben als

$$\mathrm{IFT : f(x, y)} = \int_{-\infty}^{\infty} F(u,v)\, e^{2\pi i(ux+vy)} du\, dv \tag{3.20}$$

$$\mathrm{FT : F(u, v)} = \int_{-\infty}^{\infty} f(x,y)\, e^{-2\pi i(ux+vy)} dx\, dy \tag{3.21}$$

offensichtlich sind beide Formeln nahezu symmetrisch zueinander (Vorzeichenwechsel des Exponenten).

Wie in den Gleichungen 3.20 und 3.21 verwendet man üblicherweise Klein- und Großbuchstaben, um ein zugehöriges Fourierpaar zweier Funktionen zu kennzeichnen. D.h. wir bezeichnen z.B. mit $A(u)$ die Fouriertransformierte von $a(x)$.

Die Basisfunktion e^{iux} der Zerlegung ist eine periodische Funktion. Bei Zeitsignalen kann sie als zeitliche „Schwingung" verstanden werden. Bei Bildern handelt es sich um eine periodische Variation der Lichtamplitude in der Raumkoordinate. Das klingt recht abstrakt, ist aber eigentlich sehr simpel, nämlich ein periodisches Gitter (Abb. 3.3.11). Sowohl der Realteil, als auch der Imaginärteil und der Betrag von e^{iux} ist ein einfaches Kosinus- bzw. Sinusgitter. In Analogie zur gewöhnlichen (Zeit-)Frequenz nennt man u und v die Raumfrequenzen (in x- und in y-Richtung).[12]

Im letzten Abschnitt wurde bereits angedeutet, dass die Fourierkomponenten nicht nur rechnerisch, sondern auch direkt messtechnisch ermittelt werden

[12] Oft verwendet man in der Optik auch den Wellenzahlvektor $\vec{k}$, der einfach das 2π-fache des Raumfrequenzvektors ist: $\vec{k} = 2\pi(u, v, w)$

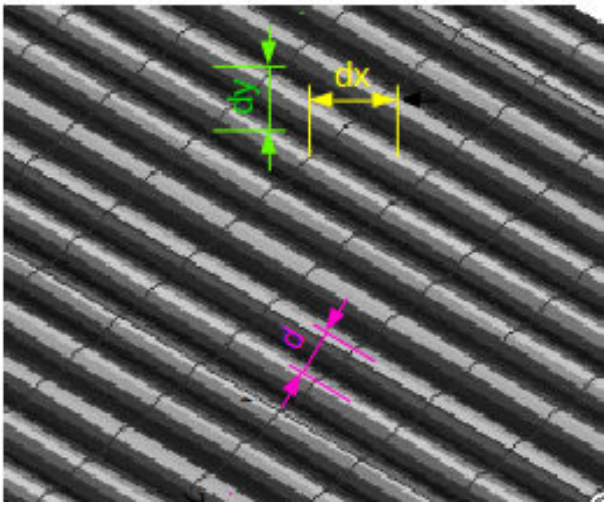


Abbildung 3.3.11: Eine konkrete Raumfrequenz (u,v) beschreibt ein Gitter der Periode d. Die „Frequenz" des Gitters ist das Inverse der Gitterperiode. Ein Gitter in x-Richtung mit der Periode d entspricht also der Raumfrequenz u=1/d.

können. Ohne Beweis (siehe z.B. [17]) wollen wir einige wichtige Ergebnisse der Fourieroptik hier nennen:

Fouriertransformation durch Linsen: Die (räumliche) Fouriertransformation wird durch eine einfache Linse (bzw. ein Objektiv) bewerkstelligt. Eine Lichtverteilung in der vorderen Brennebene wird durch die Linse in die hintere Brennebene fouriertransformiert (vgl. Abb. 3.3.13). Das bekannte Airy-Scheibchen, das uns in Kapitel 5 ausführlich beschäftigen wird ergibt sich so z.B. einfach aus der Fouriertransformierten einer Kreisblende, nämlich der Lichtverteilung in der Ebene vor der Linse.

„Lichtverteilung" bezieht sich dabei aber nicht auf die Intensität sondern auf Amplitude und Phase des elektrischen Felds des Lichts. Um die für uns in der Bildverarbeitung üblicherweise relevante Intensität (darauf sind Bildsensoren sensitiv) zu erhalten, muss das elektrische Feld noch quadriert werden.

Andere Ebenen: Der Zusammenhang für die Intensität der Fouriertransformierten gilt — abgesehen von einer einfachen linearen Skalierung — auch für andere Ebenen. Eine beliebige Ebene, die nicht konjugiert zum Objekt ist, also keine Bild- oder Zwischenbildebene, liefert durch Fouriertransformation und anschließende Intensitätsbildung die Intensität in der Bildebene.

Beugung am Gitter: Wenn wir als Eingangsverteilung ein sinusförmiges Gitter

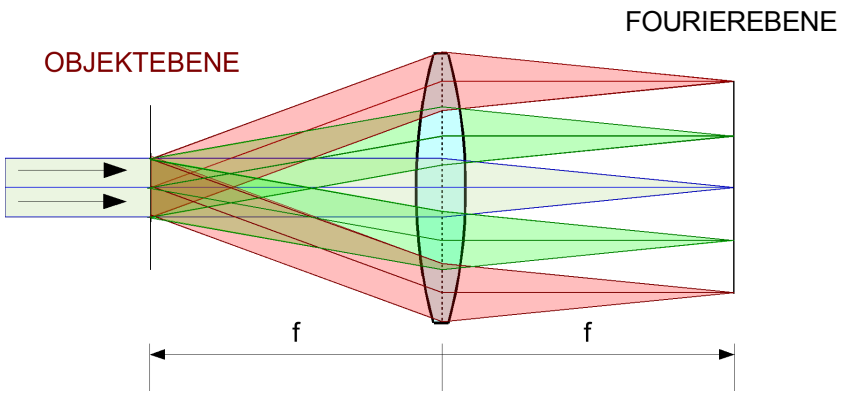


Abbildung 3.3.12: Eine Linse (bzw. Objektiv) führt eine räumliche Fourierzerlegung des Lichtfelds in der vorderen Brennebene durch. Das Ergebnis ergibt sich in der bildseitigen Brennebene. Hier dargestellt ist ein Objekt, das einer Überlagerung von zwei sinusförmigen Gittern unterschiedlicher Frequenz entspricht. Entsprechend ergeben sich für jeden Sinus zwei Peaks (Fouriertransformation eines Sinus ergibt zwei Delta-Peaks). Das rote Strahlenbündel entspricht dem Gitter mit hoher Raumfrequenz (kleine Gitterperiode).

annehmen, dann ergeben sich zwei Peaks in der Fourierebene. Da die Fouriertransformation eines Sinus zu zwei Delta-Impulsen führt erwarten wir in der Fourierebene (vorderer Teil des Systems in Abb. 3.3.14) also zwei Helligkeitspeaks. Das macht Sinn denn wir wissen aus der elementaren Beugungstheorie (Beugung am Gitter), dass bei der Beleuchtung eines (Sinus-)gitters mit der Periode p das Licht gemäß $\sin(\phi) = \pm\lambda/p$ abgelenkt wird. Die Linse führt dann dazu, dass dieses abgelenkte aber parallele Licht in der hinteren Brennebene der Linse in zwei getrennte Punkte fokussiert wird (unsere Delta-Peaks). Je feiner das Gitter ist, also je höherfrequent bzw. je geringer die Periode p des Gitters, desto weiter außen liegen die sich ergebenden Peaks.

Abbildung: Eine Abbildung kann dann als doppelte Fouriertransformation aufgefasst werden. Besonders auffällig ist dies bei dem beidseitig telezentrischen Abbildungssystem, das in Abb. 3.3.13 dargestellt ist.

Auflösung und Tiefpassfilterung Stellen wir uns vor, dass in der Fourierebene in Abb. 5.2.4 die Blende des optischen Systems liegt. Wenn nun ein abzubildendes Gitter hochfrequent genug ist, dann werden die Peaks von der Blende geblockt. Zur Bildebene gelangt also kein Licht und dementsprechend wird das Gitter nicht abgebildet.[13] Hochfrequente Gitter bzw. feine Strukturen führen

[13] Bei einem realen Amplitudengitter ist die Transmission in der Objektebene nicht $\sin(k \cdot x)$ sondern $1 + \sin(k \cdot x)$ denn negative Amplituden werden nicht realisiert. Dementsprechend ergeben sich in der Fouriertransformation drei Peaks und einer der Peaks, der ungebeugte Lichtanteil, kann die Blende passieren, so dass das Bild nicht komplett dunkel erscheinen wird.

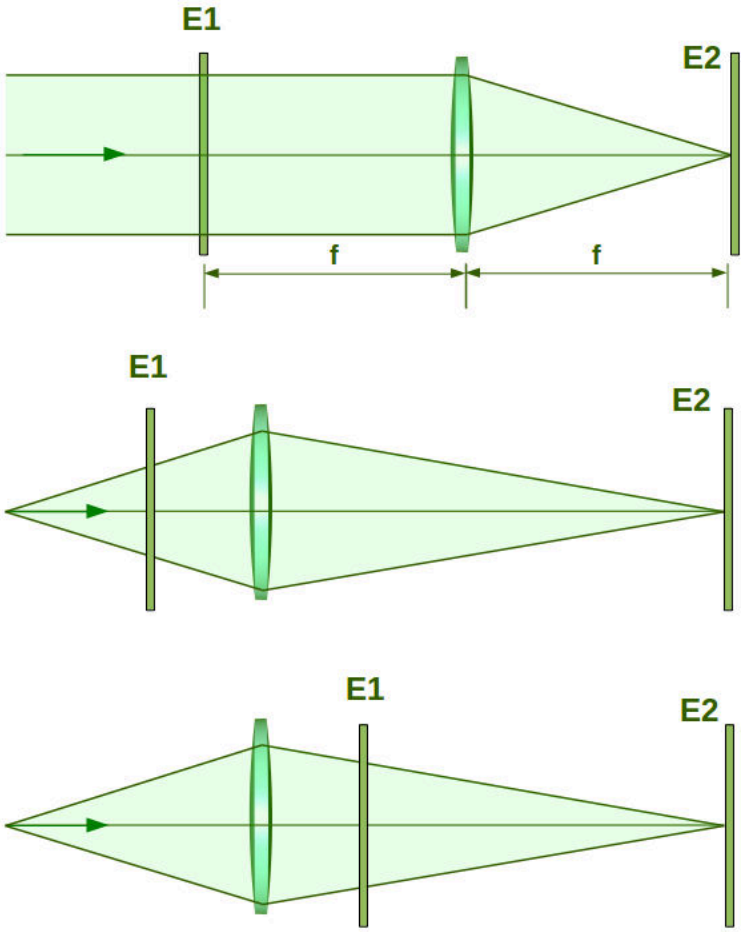


Abbildung 3.3.13: Die Linse als einfaches System zur räumlichen Fouriertransformation: Die Intensität in der Ebene E2 entspricht in allen drei Fällen der Intensität der Fouriertransformierten der Lichtverteilung (komplexe Amplitude) in der Ebene E1. Der Abstand von E1 zur Linse ist dabei irrelevant. Wenn der Abstand gerade der Brennweite entspricht (oben dargestelltes System), dann gilt zusätzlich, dass die komplexe Amplitude in E2 der Fouriertransformation der komplexen Amplitude in E1 entspricht.

so zu weit außen (hohe Frequenzen) liegenden Peaks und können so unter Umständen die nachfolgende Optik nicht passieren bzw. werden ausgeblendet.

Die Begrenzung des Lichtdurchgangs durch die „Apertur" des optischen Systems führt so anschaulich zu einer Begrenzung der maximal transmittierten Frequenz und damit zu einer Begrenzung der (lateralen) Auflösung des Systems. Dies ist der wesentliche Inhalt der Beugungsbegrenzung der optischen Abbildung (vgl. Kapitel 5).

Das optische System ist dementsprechend ein *Tiefpassfilter*, denn nur Gitteranteile mit niedriger (Raum-)Frequenz können das System passieren. Mathe-

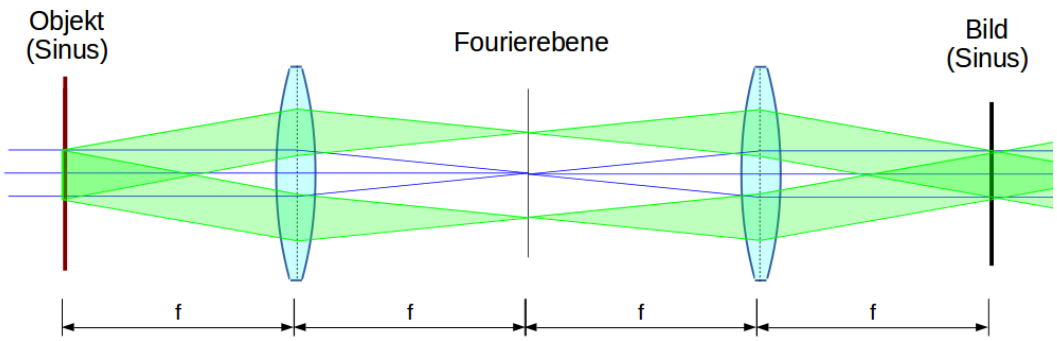
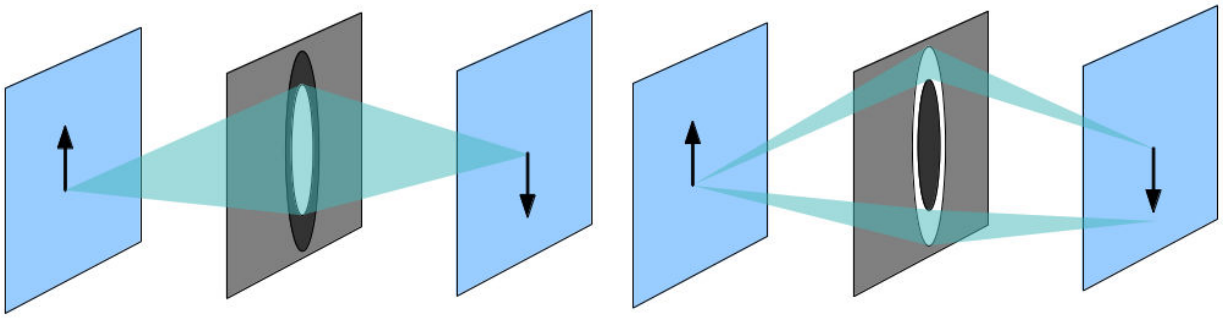


Abbildung 3.3.14: Ein 4-f System führt nacheinander zwei Fouriertransformation durch und erzielt so letztlich eine (invertierte) Abbildung. Wenn als Objekt ein Sinusgitter verwendet wird, dann ergibt die Fouriertransformation zwei Delta-Peaks, die in der Fourierebene als isolierte Peaks beobachtet werden.

matisch bedeutet Tiefpassfilterung also, dass die Fouriertransformation eines Signals beschränkt wird (Abb. 3.3.15). Tiefpassfilter lassen tiefe Frequenzen passieren und blockieren hohe Frequenzen. Feine Details werden also geblockt und es ergibt sich eine Verschmierung bzw. Mittelung im Ergebnis (Anwendung z.B. Rauschunterdrückung).

Ganz entsprechend ergibt sich ein *Hochpassfilter*, wenn die tiefen Frequenzen geblockt werden. Hochpassfilter lassen also hohe Frequenzen passieren und verstärken somit effektiv alle Bildanteile, die hohe Frequenzanteile beinhalten (insbesondere Kanten).



(a) Tiefpass (b) Hochpass

Abbildung 3.3.15: Hoch- und Tiefpassfilter können optisch realisiert werden (hier dargestellt für kohärentes Licht). Hohe Raumfrequenzen, wie sie in feinen Strukturen „enthalten" sind liegen in der Blendenebene weit außen.

3.4 Berechnung

Nach diesem kurzen Ausflug in die optische Informationsverarbeitung kommen wir nun zurück zu unserer ursprünglichen Aufgabe, den Ausgang des LIS zu berechnen. Wir denken uns also das Eingangssignal $i(x)$ dargestellt durch seine Fouriertransformierte $I(u)$.

$$i(x) := \int_{-\infty}^{\infty} I(u) e^{2\pi i u x} du \tag{3.22}$$

Jede Komponente $e^{2\pi i u x}$ wird um einen komplexen Faktor $H(u)$ durch das System modifiziert und daher können wir schreiben (Linearität):

$$o(x) = \int_{-\infty}^{\infty} H(u) I(u) e^{2\pi i u x} du \tag{3.23}$$

$o(x)$ kann aber natürlich genauso wie $i(x)$ zerlegt werden:

$$o(x) := \int_{-\infty}^{\infty} O(u) e^{2\pi i u x} du \tag{3.24}$$

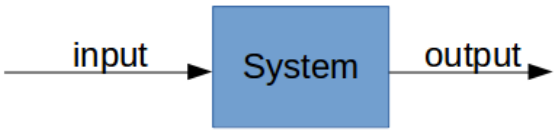
Durch Gleichsetzen von Gl. 3.23 und 3.24 ergibt sich dann direkt

$$\boxed{O(u) = H(u) I(u)} \tag{3.25}$$

Dies ist ein bedeutsames Resultat, der sogenannte *Faltungssatz*. Es sagt uns, dass wir den Ausgang des LIS nicht nur durch die recht komplizierte Faltung (immerhin eine Integration) sondern auch durch eine einfache Multiplikation beschreiben können. Dies ist aber eben nur dann möglich, wenn wir uns im (Orts–)Frequenzraum befinden. Der Faltungssatz gilt ganz generell für beliebige Funktionen, unabhängig von der systemtheoretischen Betrachtung:

$$\boxed{\mathrm{FT}[g * h] = \mathrm{FT}[g] \cdot \mathrm{FT}[h]} \tag{3.26}$$

Abb. 3.4.16 zeigt nochmals zusammengefasst die beiden grundsätzlichen Möglichkeiten zur Berechnung der Antwort eines LIS. Abgesehen von der reinen Möglichkeit der Berechnung der Faltung ist dieses Ergebnis der zentrale Ausgangspunkt für die zur Beschreibung der Qualität von Optiken üblicherweise verwendeten Modulationstransferfunktion (Kapitel 4).



V1: $o(x) = i(x) * h(x) = \int_{-\infty}^{\infty} i(x - x')h(x')dx'$

V2: Im Fourierraum: O(u) = I(u) H(u)

Abbildung 3.4.16: Im wesentlichen haben wir zwei Möglichkeiten, das Ausgangssignal eines linearen invarianten Systems zu berechnen. Der direkte Weg bleibt im Orts– bzw. Zeitraum und berechnet die Faltung mit der Punktbildfunktion. Der indirekte (aber teilweise wesentlich einfachere und anschaulichere) Weg berechnet die Faltung durch eine Multiplikation im Orts– bzw. Zeitfrequenzraum (Faltungssatz).

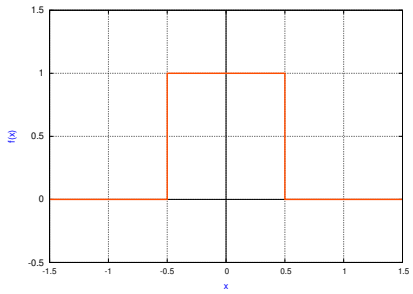
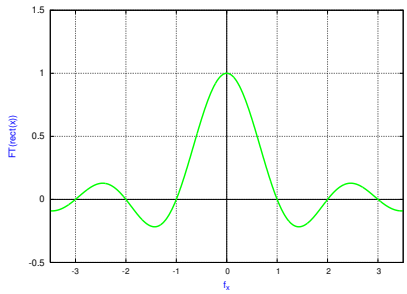
3.5 Einige Fourierpaare

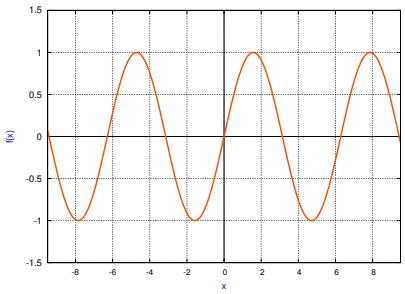
Die Beschreibung optischer Systeme mittels der Fouriertheorie wird insbesondere dadurch mathematisch praktikabel, dass vergleichsweise leistungsfähige Theoreme die Berechnung von Fouriertransformationen erheblich vereinfachen.

In der Praxis ist eine symbolische Integration der Fouriertransformation selten notwendig. Für numerische Probleme stehen leistungsfähige Algorithmen (vor allem die Fast Fourier Transform, FFT) zur Verfügung und einfache Funktionen lassen sich meist mittels der elementaren Fourier Theoreme berechnen. Wir wollen hierauf an dieser Stelle nicht näher eingehen (Details siehe [17] oder [3]).

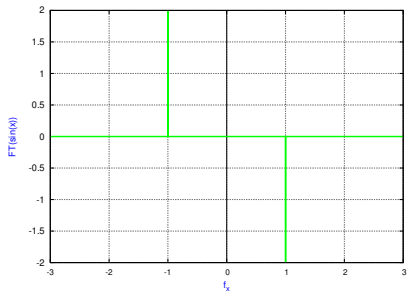
Allerdings wollen wir zumindest die Fouriertransformation für einige wichtige Grundsignale angeben. Aufgrund der Symmetrie zwischen Fouriertransformation und inverser Fouriertransformation sind die Fourierpaare jeweils in beide Richtungen lesbar. Das bedeutet, dass z.B. eine Rechteckfunktion im Ortsbereich, die zu einer sinc–Funktion im Ortsfrequenzbereich führt, genauso im Frequenzbereich betrachtet werden kann. So wird dann z.B. eine Rechteckfunktion im Ortsfrequenzbereich zu einer sinc–Funktion im Ortsbereich.

Wir listen die Fourierpaare für die eindimensionale räumliche Fouriertransformation.

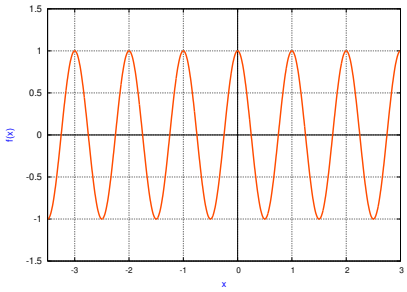
Funktion im Ortsraum	Fouriertransformation
 $\mathrm{rect}(x)$	 $\mathrm{sinc}(u)$



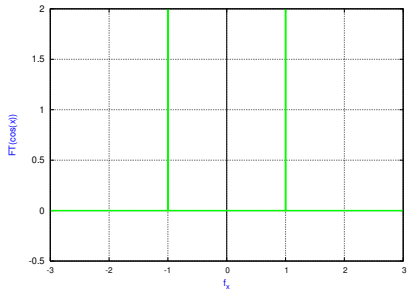
$\sin(x)$



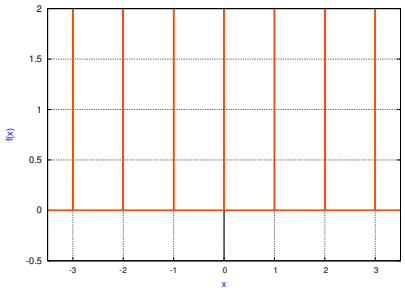
$i\left(\delta(u+1)-\delta(u-1)\right)$



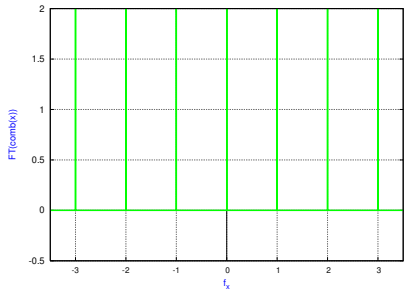
$\cos(x)$



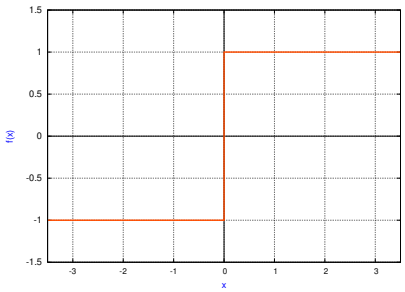
$\delta(u \pm 1)$



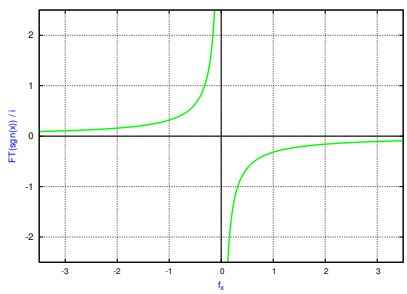
$\mathrm{comb}_{\Delta x}(x)$



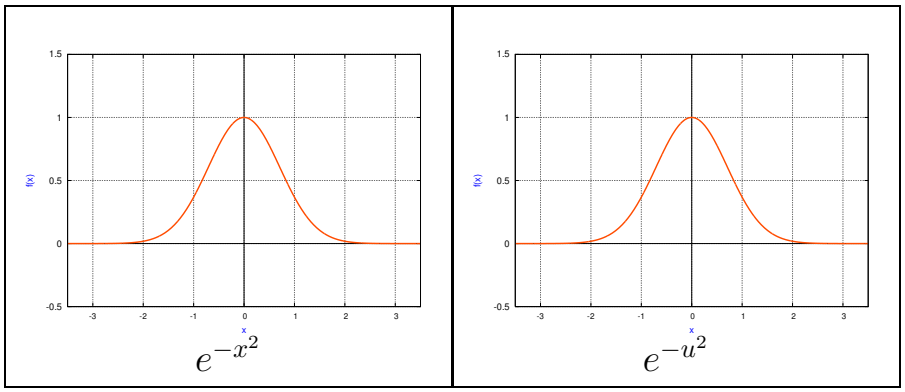
$\Delta x\,\mathrm{comb}_{1/\Delta x}(u)$



$\mathrm{sgn}(x)$



$\frac{-i}{\pi u}$



3.6 Das Abtasttheorem

Bisher haben wir kontinuierliche Funktionen betrachtet. Das ist — zumindest auf den ersten Blick — auch sinnvoll, denn bei einer Abbildung bilden wir ein kontinuierliches Objekt ab. Eine diskrete Beschreibung ist dennoch von Bedeutung. Zum einen sind wir bei einer Bearbeitung im Computer meist auf eine numerische Behandlung angewiesen und müssen daher eine *Diskretisierung* durchführen, zum anderen ist die eigentliche Messung eine Diskretisierung. Bei der Abbildung tastet beispielsweise der Bildaufnehmer das abgebildete Objekt an N Stützstellen (entsprechend den Pixeln des Bildes) ab.[14]

Die Abtastung (*Sampling*) durch Bildaufnehmer wird uns an verschiedenen Stellen noch beschäftigen. So führt die Mittelung über die einzelne Pixelfläche zu einer Auflösungreduktion (Abschnitt [**?**]) und die Wahl von Abtastungsperiode in Bezug zur optischen Auflösung des Abbildungssystems ist ebenfalls von großer Bedeutung (Abschnitt 5.7). Hier soll es aber zunächst um die ideale punktweise Abtastung von idealen Bildsignalen gehen. Unschärfen der Abbildungsoptik und praktische Bildsensoreffekte werden wir also zunächst außen vor lassen.

Wir wollen also jetzt von der kontinuierlichen Beschreibung zu einer diskreten Beschreibung übergehen und gehen im folgenden davon aus, dass die Stützstellen,

[14] Dasselbe Problem haben wir bei elektrischen Signalen: Wir können (und wollen es im übrigen auch nicht) einen Spannungsverlauf nicht unendlich genau vermessen. Wir haben uns auf eine begrenzte, ausreichende Anzahl von Abtastpunkten zu beschränken .

auf denen die Messpunkte liegen, in äquidistantem Abstand Δx zueinander liegen. Das muss nicht immer der Fall sein. Manchmal wäre sogar eine andere Wahl der Stützstellen, z.B. in bestimmten Regionen eine erhöhte Anzahl, sinnvoller.[15] In der Regel weisen aber die in der Optik verfügbaren Aus– und Eingabemedien, z.B. LCDs und CCDs, ein gleichmäßiges Raster von Stützstellen auf.

Zunächst werden wir das Problem der Abtastung — ganz anschaulich — im Ortsbereich verstehen, bevor wir dann — eher mathematisch — auch die Deutung im Ortsfrequenzraum, also der Fourierebene, betrachten.

Wir versuchen das sogenannte *Samplingtheorem* zunächst anhand eines einfachen Beispiels anschaulich zu machen. Es sagt aus, dass wir mit der doppelten Frequenz der maximal im Signal vorhandenen Frequenz abtasten sollen.

In Abb. 3.6.17 ist eine Sinusschwingung im Ortsbereich mit der Periode 50 μm skizziert. Diese wird mit einer Abtastperiode von 40 μm punktweise abgetastet. Offensichtlich ist also das Abtasttheorem verletzt, denn dieses würde verlangen, dass wir das Signal mit einem Abtastabstand von maximal 25 μm abtasten.

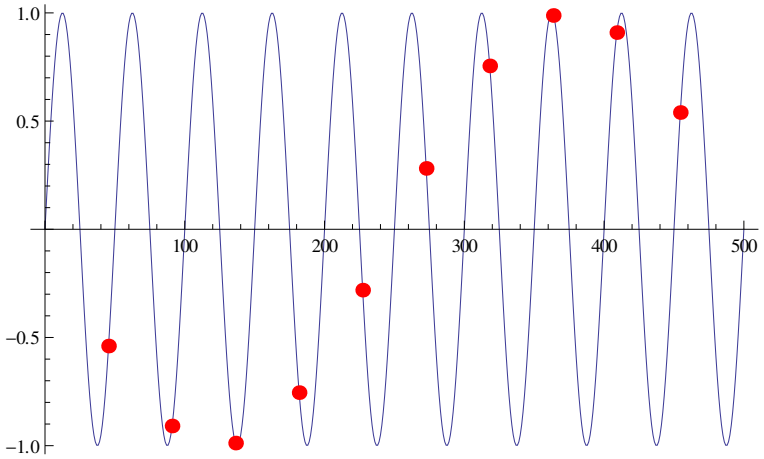


Abbildung 3.6.17: Aliasing bzw. Moire im Ortsraum. Die Abtastung mit einer zu geringen Samplingfrequenz führt zu dem Auftreten einer falschen, tieffrequenzen Frequenz.

Bei der Betrachtung des abgetasteten Signals ergibt sich der Eindruck, dass die Signalperiode deutlich größer (ca. 500 μm) ist. Es ergibt sich eine drastische Fehleinschätzung hinsichtlich der Signalfrequenz. Es ist also nicht nur so, dass Frequenzen oberhalb der durch das Abtasttheorem gegebenen Grenzfrequenz durch die Abtastung nicht erfasst werden. Vielmehr wird durch das Vorhandensein ent-

[15] So wie es auch auf unserer Netzhaut praktiziert wird.

sprechender Frequenzen oberhalb der Grenzfrequenz das Ergebnis auch im eigentlichen Messbereich verfälscht. Es hat den Anschein, dass im Signal Frequenzen enthalten sind, die in Wirklichkeit eben nicht vorhanden sind, sondern nur durch die mangelhafte Abtastung hervorgerufen werden. Man sprich von Aliasing , einem Sonderfall des *Moire*.[16]

Anschaulich ist klar, was bei der Abtastung schief läuft. Es ist aber erhellend, sich den Vorgang der Abtastung im Orstfrequenzraum anzuschauen.

Dazu wollen wir die Fouriertransformation des abgetasteten Signal berechnen. Die einfachste und sinnvollste Möglichkeit, die Integration der Fouriertransformation über die äquidistanten Messpunkte durchzuführen, ist, sie einfach zu summieren. Wir wissen zunächst nicht, was sich zwischen den Messpunkten abspielt und nehmen durch diese Summation einfach an, dass sich der Signalverlauf zwischen zwei Messstellen nur wenig ändert.

Wir setzen also

$$a(\mu) = \int_{-\infty}^{\infty} A(x) e^{-2\pi i \mu x} dx \approx \sum_{n=-\infty}^{\infty} A(n\Delta x) e^{-2\pi i \mu n \Delta x} \tag{3.27}$$

Dies ist die diskrete Fouriertransformation (DFT). Die schnelle Fouriertransformation (engl. Fast Fourier Transform, *FFT*) ist einfach ein Algorithmus, der es erlaubt, die DFT sehr effizient zu berechnen. Alle Ergebnisse, die für die DFT gewonnen werden sind dementsprechend genauso für die FFT gültig.

Es stellt sich sofort die Frage, wieviele Stützstellen benötigt werden, um den durch die Approximation gegebenen Fehler nicht zu groß werden zu lassen.

Mathematisch können wir die ideale, punktförmige Abtastung einer Funktion $f(x)$ durch

$$g(x) = \mathrm{comb}_{\Delta x}(x) \cdot f(x) = \sum_{k=-\infty}^{\infty} \delta(x - k\Delta x) f(x) \tag{3.28}$$

beschreiben[17].

[16] Moires ergeben sich bei additiver oder multiplikativer Überlagerung von Gittern (meist im Orts- oder Zeitebereich). Im Fall der Bilderfassung ist ein Gitter durch das Objekt gegeben, das zweite Gitter ist die regelmäßige Pixelstruktur.

[17] Die Abtastung mit endlich ausgedehnten Pixeln wird durch eine zusätzliche Faltung des Delta–Kammes mit einer der Pixelbreite entsprechenden rect–Funktion beschrieben. Dazu später mehr.

Die Kammfunktion $\mathrm{comb}_{\Delta x}(x)$ist dabei eine Folge von äquidistanten Delta-Peaks mit dem Abstand Δx (siehe Abb. 3.1.5).

Im Frequenzraum erhalten wir damit mit dem Faltungssatz (Gl. 3.26).

$$G(u) = F(u) * \mathrm{comb}_{1/\Delta x}(u), \tag{3.29}$$

denn die Fouriertransformation der Kammfunktion ist wieder eine Kammfunktion.

Die Faltung einer Funktion mit einem einzelnen, verschobenen Delta-Peak führt zu einer Verschiebung der Funktion:

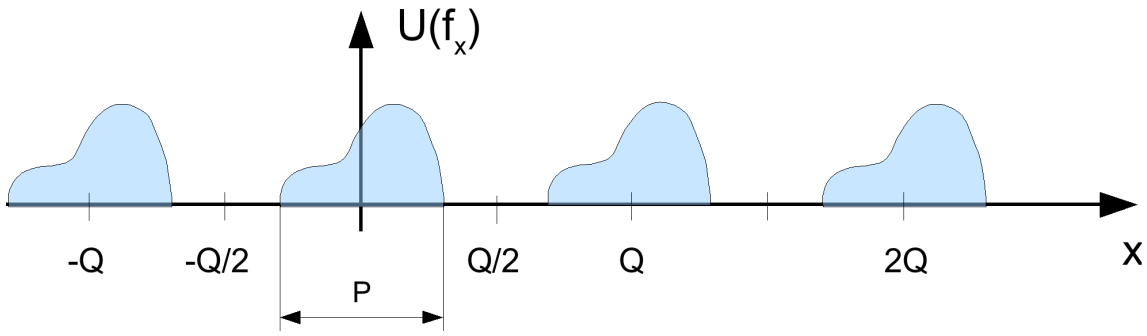
$$F(u) * \delta(u - \Delta u) = F(u - \Delta u), \tag{3.30}$$

so dass wir letztlich in Gleichung 3.29 eine Überlagerung vieler gegeneinander verschobener Kopien von $F(u)$ erhalten. Oder in anderen Worten: Durch die Abtastung ergibt sich automatisch eine periodische Wiederholung des Spektrums der abgetasteten kontinuierlichen Funktion (vgl. Abb. 3.6.18). Die Wiederholungsperiode ist direkt durch $1/\Delta x$ gegeben.

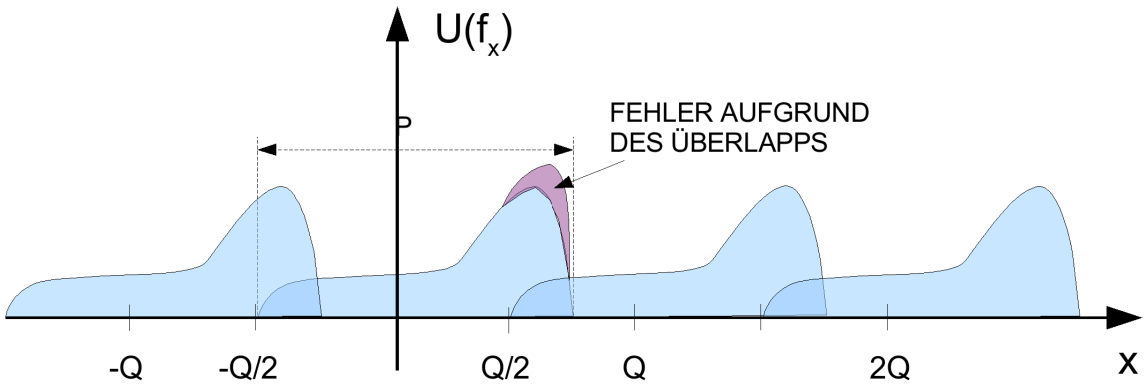
Es kommt ganz offensichtlich zu Fehlern, wenn das Spektrum von f eine größere Ausdehnung als dieses $1/\Delta x$ hat. In diesem Fall überlagern (addieren) sich dann nämlich die einzelnen spektralen Wiederholungen. Ganz verschiedene Einzelüberlagerung können zur selben Summe führen und ein Rückschluss auf das ursprüngliche Signale ist so nicht mehr möglich.

Dies ist in Abbildung 3.6.18 für zwei verschiedene Signale demonstriert. Der Fehler kann anschaulich (periodische Wiederholung des Spektrums) durch ein „Umklappen" des Spektrums an der sogenannten *Nyquistfrequenz* u_N (s.u.) begriffen werden. Frequenzen mit $u + u_N$ erscheinen bei $u - u_N$. Man spricht von *Aliasing* bzw. *Moires*.

Den wichtigsten Sonderfall haben wir bei einer sogenannten „Tiefpass–Abtastung": Das Signal wird als begrenzt hinsichtlich hoher Frequenzen angenommen (ein Fall, der sich sowohl in der Optik als auch in der Elektronik bei den meisten Systemen praktisch automatisch ergibt). Eine Begrenzung tiefer Frequenzen findet nicht statt. Wenn ein reelles Signal vorliegt, dann ist weiterhin zu beachten, dass das Spektrum symmetrisch bezüglich $f = 0$ ist. Insgesamt wird die Breite des Spektrums also bei einer maximalen Frequenz von f_g genau $2f_g$ sein.



(a) Abtastfrequenz ausreichend



(b) Abtastfrequenz zu niedrig (Aliasing)

Abbildung 3.6.18: Aliasing durch Nichtbeachtung des Abtasttheorems bei Abtastung mit der Samplingfrequenz Q. Äquidistante Abtastung (Sampling) führt zu einer periodischen Wiederholung des Grundspektrums im Frequenzraum. Sobald sich die Kopien überlappen, also nicht mehr trennbar sind, ergeben sich Aliasing Fehler.

Damit es zu keinen Überlappungen durch die Abtastung kommt, muss demnach $\frac{1}{\Delta x} \geq 2f_g$ bzw.

$$\boxed{\Delta x \leq \frac{1}{2f_g} \qquad (3.31)}$$

gelten. Mit anderen Worten: Damit es zu keinen Fehlern bei der Abtastung kommt, darf das Signal nur Frequenzen unterhalb der sogenannten *Nyquistfrequenz* $f_N = 1/(2\Delta x)$ enthalten.

Abtasttheorem: Eine korrekte Schätzung der in einem Signal enthaltenen Frequenz ist nur möglich, wenn die Abtastfrequenz mindestens doppelt so hoch ist wie die maximale im Signal enthaltene Frequenz. Eine korrekte Schätzung der Amplitude der Schwingung ist damit allerdings zunächst noch nicht möglich.

Hierfür ist eine etwas höhere Abtastfrequenz notwendig. Die maximal im Signal erlaubte Frequenz wird *Nyquist-Frequenz* genannt.[18]

In der visuellen Wirkung besonders problematisch sind Moires bei Farbsensoren mit Farbmasken (s. Abb. 3.6.19). In diesem Fall ergeben sich — auch bei unfarbigen Objekten — deutliche Farbstörungen.

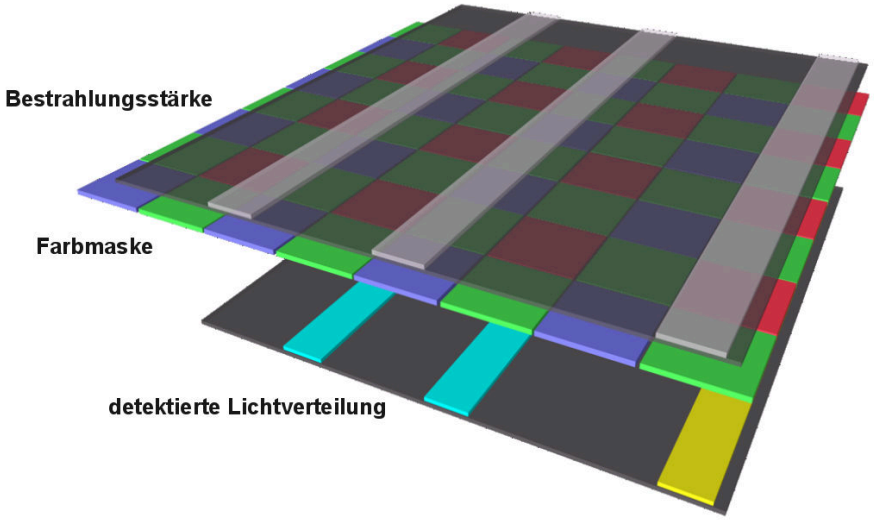


Abbildung 3.6.19: Die überwiegende Anzahl von Farbsensoren nutzt ein absorbierendes Farbfilter vor den lichtempfindlichen Pixeln. Aus den sich so ergebenden Daten werden die jeweils fehlenden beiden Farben an jedem Pixel durch Interpolation gewonnen. Sehr fein strukturierte Helligkeitsverteilungen können zu Farbaliasing führen (hier dargestellt: eine Schwarz-Weiß Eingangslichtverteilung führt zu einer Cyan/Gelb Detektion). Es ergeben sich stark störende Farbeffekte. Zur Vermeidung dieser Effekte werden Unschärfen und optische Tiefpässe eingesetzt.

Bei unterabgetasteten Bildern ergeben sich also vergleichsweise tieffrequente Störungen, die insbesondere dann ins Auge stechen, wenn die Objekte periodische Strukturen aufweisen (Moires).

Um entsprechende Fehler zu vermeiden, ist es daher unerlässlich, das Eingangssignal hinsichtlich zu hoher Frequenzen zu begrenzen. Aus diesem Grund werden bei akustischen Signalen vor dem Abtasten (Sampeln) entsprechende Tiefpassfilter eingesetzt.

In der Optik muss ebenfalls eine entsprechende Tiefpassfilterung vorgenommen werden. Ein typisches Beispiel ist der Tiefpassfilter bei *Farbsensoren* (vgl. Abb.

[18] Der Zusammenhang wurde bereits 1928 von H. Nyquist gefunden.

3.6.19). Meist ist bei hohen Abtastfrequenzen, wie sie bei modernen Digitalkameras z.B. üblich sind, die Abtastfrequenz so hoch, dass bereits leichte Unschärfen, wie sie durch Aberrationen, Beugung oder Defokussierung automatisch auftreten, einen ausreichenden Tiefpasseffekt ergeben.

Dennoch kann immer noch leichtes Moire auftreten. Abb. 3.6.20 zeigt ein mit einem Samsung A5 aufgenommenes regelmäßiges Raster. Die Moires sind deutlich sichtbar. Abb. 3.6.21 und 3.6.22 zeigen einen Bildausschnitt vergrößert, in Abb. 3.6.22 mit erhöhter Farbsättigung wodurch die Farbfehler deutlicher sichtbar werden.

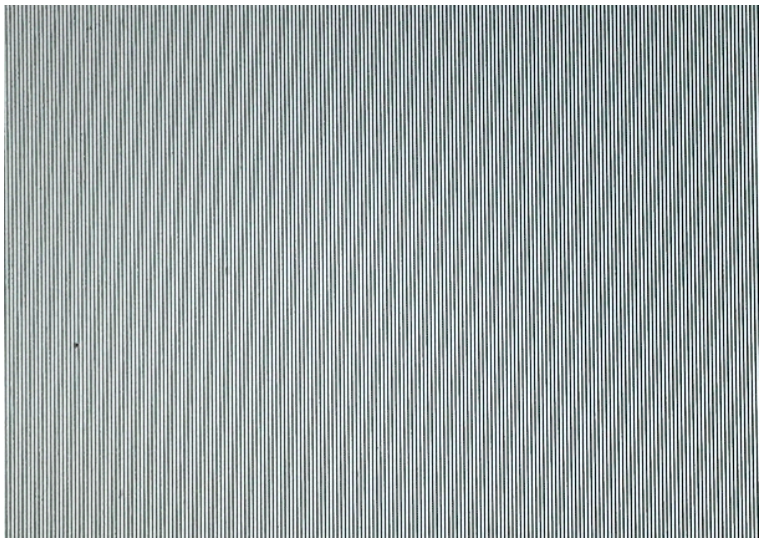


Abbildung 3.6.20: Moire bei Fotografie eines regelmäßigen Binärgitters mit einem Samsung A5.

Gehen wir davon aus, dass durch eine entsprechende wie auch immer geartete Tiefpassfilterung im Eingangssignal nur Frequenzen unterhalb von u_g auftreten:

$$G(u) = F(u) \cdot \mathrm{rect}(u/u_g) \tag{3.32}$$

Im Ortsraum erhalten wir dann

$$g(x) = f(x) * \mathrm{sinc}(f_g x) = \sum_{-\infty}^{\infty} x[nP]\mathrm{sinc}(x/P - n) \tag{3.33}$$

mit der Abtastperiode P.

Dies ist die *Rekonstruktionsformel* für abgetastete *bandbegrenzte Funktionen*. Damit wird es möglich, die Orginalfunktion g(x) nicht nur an den eigentlich gemessenen Abtastpunkten zu beschreiben, sondern auch die Zwischenstellen zu berechnen.

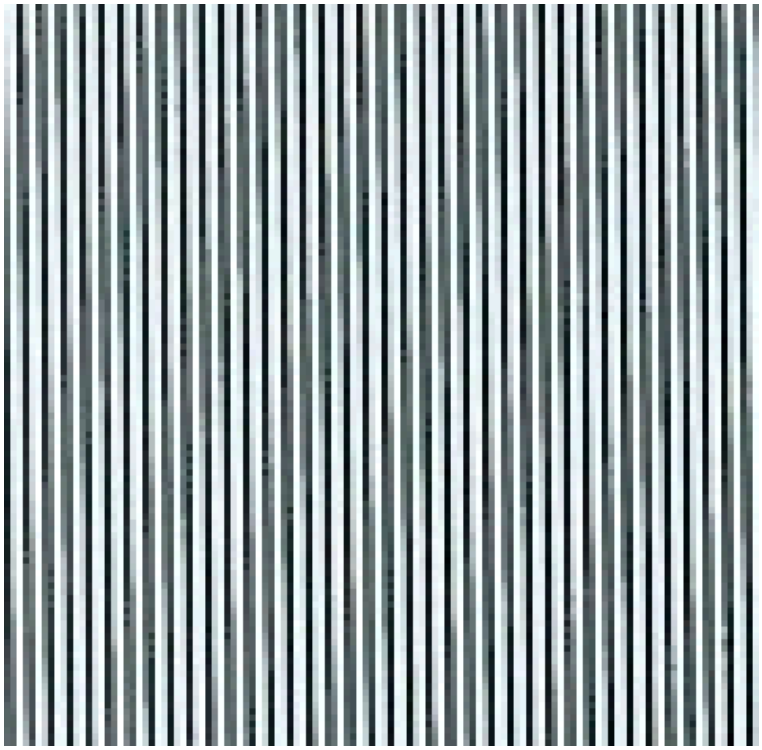


Abbildung 3.6.21: Moire bei Fotografie eines regelmäßigen Binärgitters mit einem Samsung A5. Vergrößerte Darstellung.

In vielen Fällen ist eine entsprechende Rekonstruktion von Zwischenstellen gar nicht nötig oder kann einfach linear zwischen Abtaststellen erfolgen. Man „ahnt" praktisch die Funktion korrekt allein durch Betrachtung der Abtaststellen. Für eine exakte Bestimmung der Zwischenstellen müsste Gl. 3.33 allerdings beachtet werden. In der Praxis wird entsprechender Aufwand praktisch nie getrieben, zumal die Abtastung in der Optik nie punktweise sondern mit ausgedehnten Pixeln erfolgt.

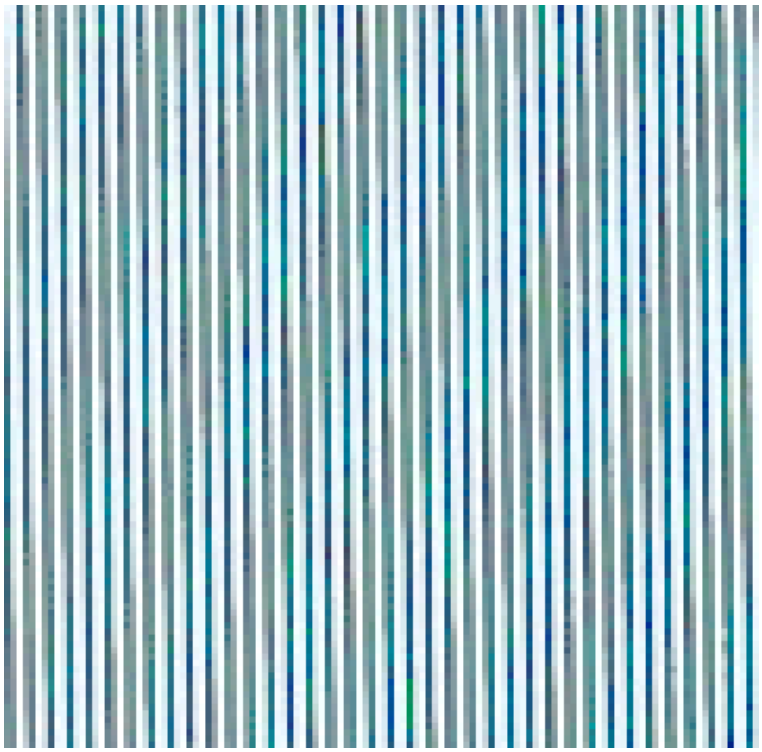
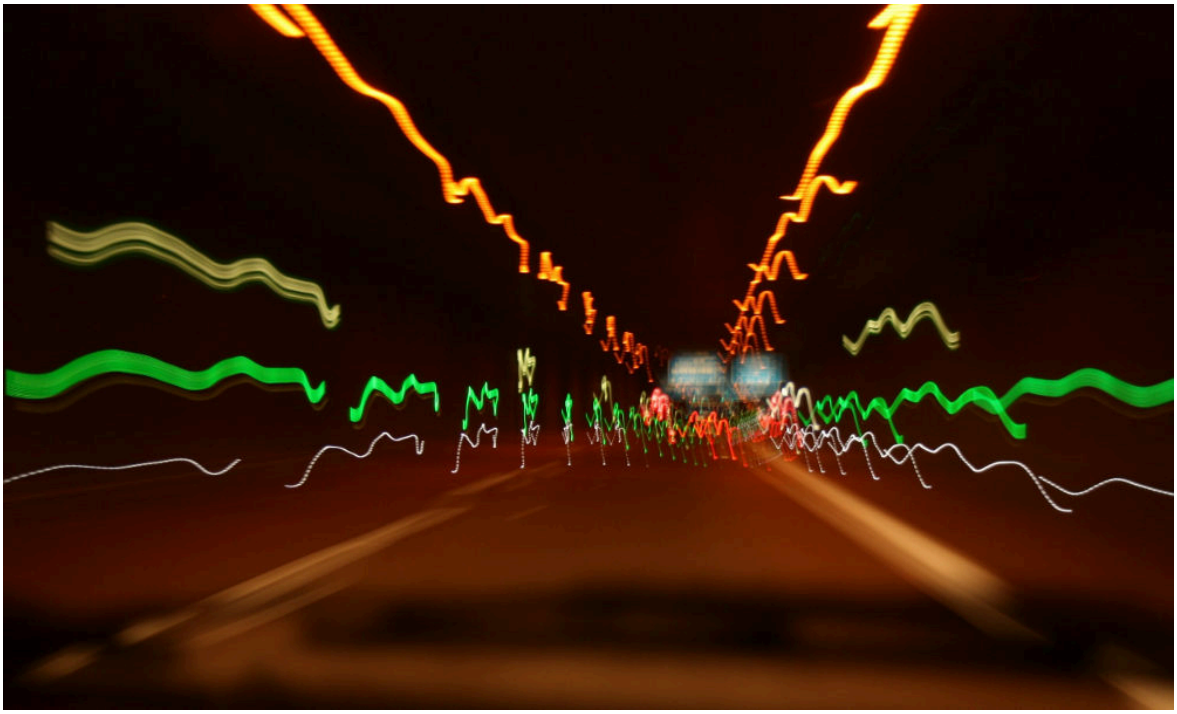


Abbildung 3.6.22: Moire bei Fotografie eines regelmäßigen Binärgitters mit einem Samsung A5. Vergrößerte Darstellung mit Erhöhung der Farbsättigung.

Kapitel 4

Transferfunktion



Basierend auf der mathematischen Beschreibung des letzten Kapitels sind wir nun in der Lage, die in der industriellen Bildverarbeitung übliche Beschreibung der Bildqualität, nämlich die sogenannte *Modulationstransferfunktion* (MTF), zu verstehen und für eine konkrete Applikation anzuwenden.

In diesem Kapitel werden Sie unter anderem lernen

- wie die Abbildungsleistung einer Optik üblicherweise beschrieben wird (optische Transferfunktion),
- was die Modulationstransferfunktion anschaulich bedeutet,

- wie die Modulationstransferfunktion der Kombination verschiedener Komponenten berechnet wird und
- wie die Modulationstransferfunktion eines Bildaufnehmers und typischer Bildaufnahmefehler (Defokus, Verwackeln) berechnet wird.

4.1 Nochmals Faltung und Multiplikation

Ohne uns größere Gedanken über die zugrunde liegende Physik zu machen, wollen wir die Optik (z.B. ein Fotoobjektiv) als ein linear invariantes System annehmen. Wir wissen bereits aus Abschnitt 3.1, dass wir in diesem Fall den Ausgang des Systems durch die Faltung des Eingangs i mit der Punktbildfunktion h beschreiben können, d.h. $o(x) = i(x) * h(x)$

Außerdem wissen wir, dass die Berechnung auch im Ortsfrequenzraum über den Faltungssatz ausgeführt werden kann. D.h. wir transformieren $i(x)$ und $h(x)$ und erhalten dann (vgl. Gl. 3.25)

$$O(u) = I(u)H(u) \tag{4.1}$$

$H(u)$ ist die also die Fouriertransformierte der Punktbildfunktion $h(x)$ und wird als optische Transferfunktion (OTF) bezeichnet. Der Betrag von H ist die Modulationstransferfunktion, kurz MTF. Allgemein, also z.B. in der Elektrotechnik, wird H *Übertragungsfunktion* genannt.

Um die Bedeutung der MTF anschaulich zu verstehen, betrachten wir die Abbildung eines Kosinusgitters der Raumfrequenz u_0. D.h. wir bilden das Gitter $1 + \cos(2\pi u_0 x)$ ab. (Der zusätzliche Konstantanteil „1" ist notwendig weil der Kosinus von -1 bis 1 läuft und wir keine negativen Amplituden zulassen wollen.)

Wir wissen bereits, dass Sinus– bzw. Kosinusschwingungen Eigenfunktionen des linearen invarianten Systems sind. Wir erwarten also, dass wir am Ausgang des Systems ebenfalls eine Kosinusschwingung mit der grundsätzlichen Form $\cos(2\pi u_0 x)$ erhalten. Allerdings wird der Kontrast dieser Schwingung in der Regel verringert sein, das Signal wird beim Durchgang durch das passive optische

System gedämpft[1].

In der Ortsfrequenzebene besteht unser Kosinus aus zwei Delta–Impulsen und einem konstanten Anteil

$$\mathrm{FT}[\cos(2\pi u_0 x)] = 1 + \delta(2\pi(u - u_0)) + \delta(2\pi(u + u_0)) \tag{4.2}$$

Durch das System werden Kosinussignale mit unterschiedlichen Frequenzen (Gl. 4.2) unterschiedlich stark abgeschwächt. Der Abschwächungsfaktor ist gerade die MTF, also der Betrag von $H(u)$.

Für typische Abbildungssysteme ist die MTF eine stetig abfallende Funktion (siehe z.B. Abb. 4.1.1). Daraus ergibt sich, dass höhere Ortsfrequenzen im Vergleich zum Konstantanteil ($u = 0$) stärker abgeschwächt werden. In erster (und etwas zu optimistischer) Näherung ist die MTF eines *beugungsbegrenzten Systems* eine linear abfallende Funktion[2]:

$$MTF(u) \approx 1 - \frac{u}{u_{max}} \tag{4.3}$$

Die *Grenzfrequenz* ist gerade

$$u_{max} = \frac{1}{\lambda K} \tag{4.4}$$

Oft wird die MTF so normiert, dass sie für den Konstantanteil 1 wird.[3]

Mit der Normierung der MTF auf 1 für $u = 0$ wird für unsere angenommene Kosinusschwingung der Ausgang des Systems in der Ortsfrequenzebene

$$O(u) = H \cdot I = 1 + |H(u_0)|\delta(u - u_0) + |H(-u_0)|\delta(u + u_0) \tag{4.5}$$

sein. Im Ortsraum ergibt sich damit für $|H(u)| = |H(-u)|$

$$o(x) = 1 + |H(u_0)| \cos(2\pi x u_0) \tag{4.6}$$

[1] Aktive optische Systeme, z.B. Laserresonatoren oder phasenkonjugierende Spiegel, insbesondere aber nichtlineare Filter der digitalen Bildverarbeitung, können das Signal auch verstärken.

[2] Die exakte MTF findet sich z.B. in [12, 36]

[3] Die Normierung auf 1 bei tiefen Frequenzen macht nicht immer Sinn denn Streulicht kann dazu führen, dass auch bei sehr tiefen Frequenzen der Kontrast der Abbildung merklich sinkt. Wir vernachlässigen in diesem Abschnitt aber zunächst Störlichtanteile (z.B. Streuung).

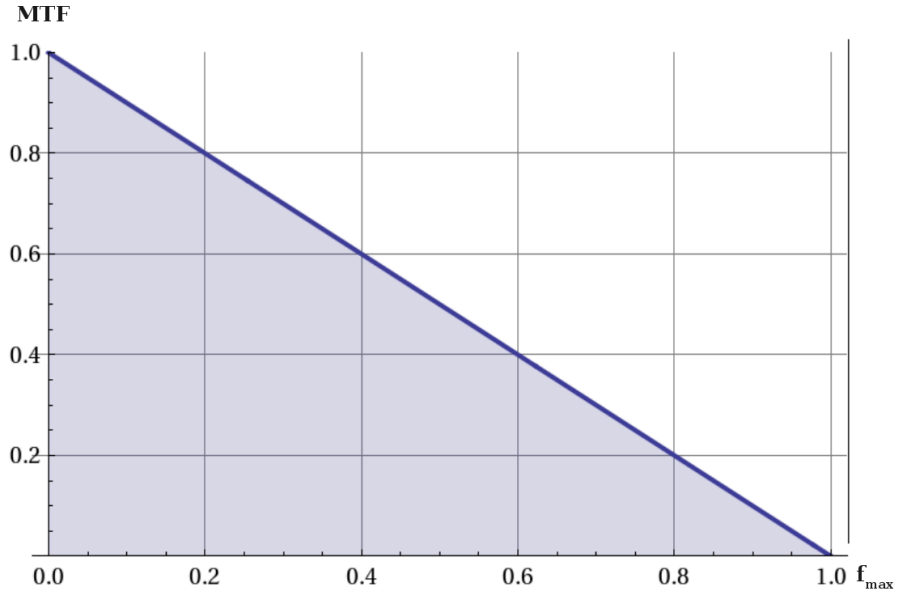


Abbildung 4.1.1: MTF eines idealen (keine Aberrationen) beugungsbegrenzten Abbildungssystems in eindimensionaler (guter) Näherung.

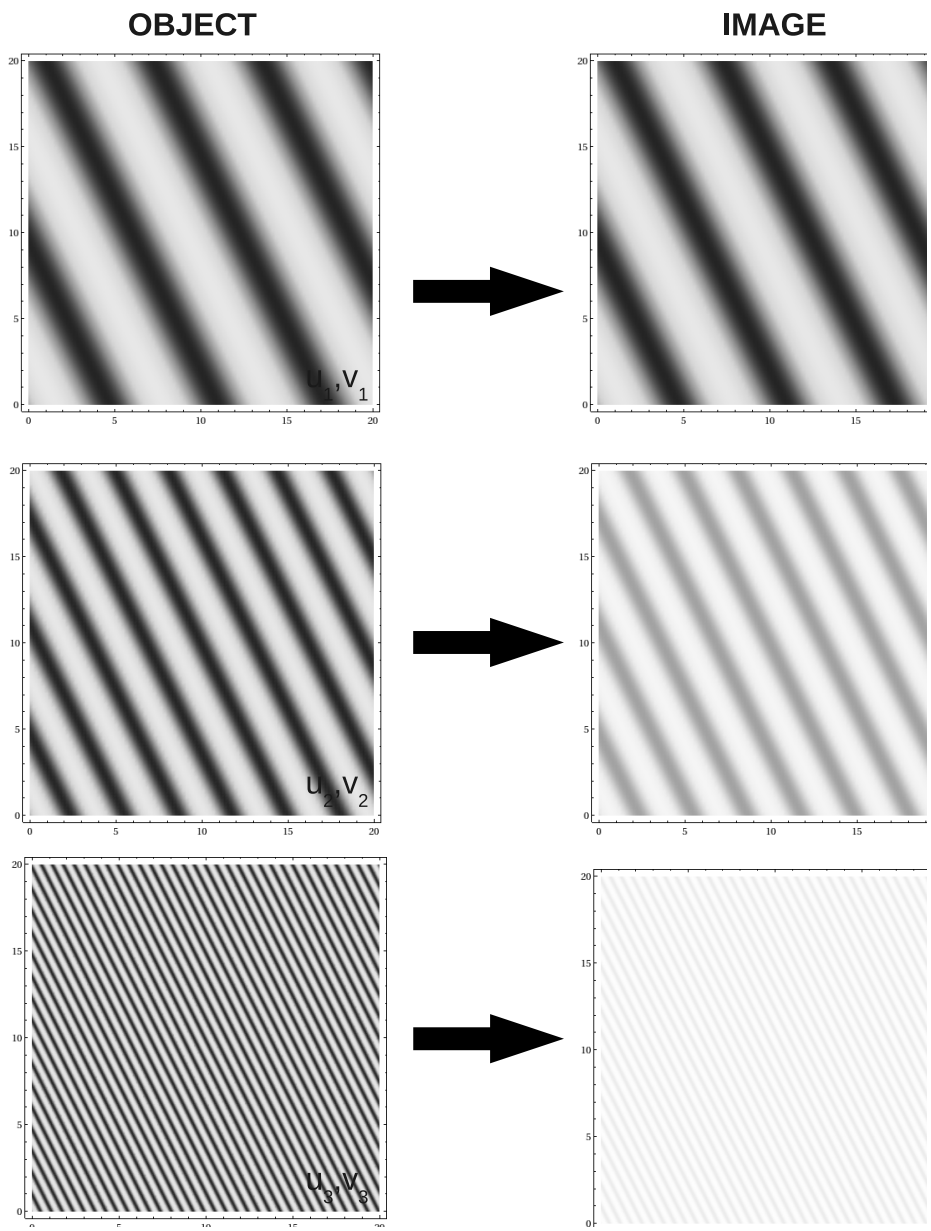


Abbildung 4.1.2: Die MTF gibt den Kontrast des Bildes eines periodischen Sinusgitters an. Üblicherweise nimmt dieser Kontrast mit zunehmender Gitterfrequenz ab.

Der *Kontrast*[4] (engl.: Modulation) dieser Schwingung ist für das angenommene $|H(u_0)| < 1$ verringert, denn es gilt (für $H := |H(u_0)|$)

$$K := \frac{I_{max} - I_{min}}{I_{max} + I_{min}} = \frac{1 + H - (1 - H)}{1 + H + 1 - H} = \frac{2H}{2} = H \tag{4.7}$$

Wir haben so also das wichtige Resultat erhalten, dass der Kontrast einer durch das System übertragenen Kosinusschwingung direkt durch die MTF des Systems bei der Frequenz der Kosinusschwingung gegeben ist. Natürlich gilt dasselbe auch für einen phasenverschobenen Kosinus, also den Sinus.

Die MTF gibt also an, welcher Kontrast bei der Abbildung eines Gitters mit einer bestimmten Raumfrequenz bzw. Gitterperiode im Bild erzielt wird.

Abb. 4.1.2 verdeutlicht nochmals die Zusammenhänge.

Für sehr geringe Gitterfrequenzen ergibt sich so ein Kontrast von nahezu 1, während hohe Frequenzen schlecht übertragen werden und der Kontrast dementsprechend abfällt. Ein Beispiel zeigt Abb. 4.1.3. Die Gitterfrequenz wird in Linienpaaren pro mm (lp/mm) angegeben. So bedeuten 100 lp/mm, dass pro Millimeter 100 Perioden in der Bildebene auftreten bzw. dass die Periodenlänge 10 μm beträgt.

Verschiedene Auftragungen für die MTF sind üblich. Gegenüber der bisherigen Darstellung (Kontrast über Raumfrequenz) wird bei Objektiven oft eine Darstellung der MTF über die (Bild-)Feldkoordinate dargestellt (s. Abb. 4.1.4), denn natürlich hängt die Abbildungsleistung eines Systems meist von der Position im Bildfeld ab.[5] In diesem Fall geben die Kurven dann jeweils die MTF für vorher definierte Raumfrequenzen (z.B. 50 lp/mm) an. Zu beachten ist auch, dass jeweils zwei Kurven für verschiedene Gitterorientierungen („sagital" und „tangential") dargestellt sind. D.h. je nach Gitterorientierung kann sich durchaus ein anderer Kontrast im Bild ergeben. Asymmetrische Punktbildfunktionen, wie sie z.B. bei Koma auftreten machen dies anschaulich. In diesem Fall ist es klar, dass es einen starken Unterschied macht, ob eine (periodische) Struktur entlang der langen Komaausdehnung (Schweif) ist oder senkrecht hierzu.

[4] Der Begriff „Kontrast" ist mit etwas Vorsicht zu verwenden, denn des gibt durchaus unterschiedliche Definitionen. Im angelsächsischen werden hier verschiedene Begriffe für die verschiedenen Definitionen verwendet.

[5] Dabei geht streng genommen die Ortsinvarianz verloren, dennoch kann man in lokalen Regionen (sogenannten isoplanatischen Bereichen) wieder die linear invariante Theorie und damit die MTF anwenden.

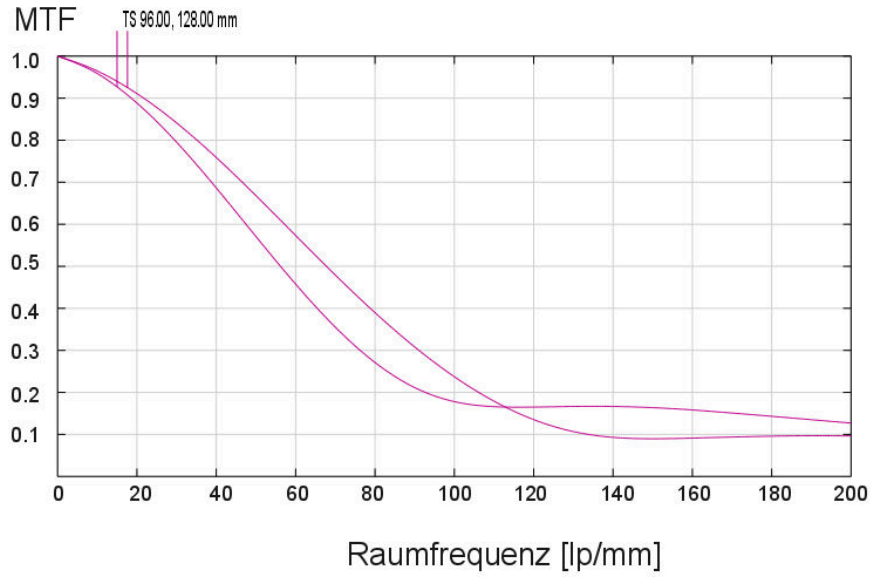


Abbildung 4.1.3: Typische MTF einer Abbildungsoptik, dabei werden zueinander senkrecht orientierte Gitter mit (leicht) unterschiedlichem Kontrast abgebildet.

Die Phase der OTF wird auch als PTF bezeichnet. In der Praxis ist sie meist von untergeordneter Bedeutung und wird nur selten benutzt bzw. angegeben. Solange der durch die MTF gegebene Kontrast größer 0.2 bleibt, ist die PTF auf jeden Fall geringer als $\pi/4$ und hat dabei bei abtastenden Bildaufnehmern praktisch keinen Effekt [6, 5].[6]

Anschaulich gibt die PTF an, wie stark ein Sinus mit einer bestimmten Frequenz am Ausgang lateral verschoben wird (s. Abb. 4.1.5). Für Systeme, bei denen diese PTF deutlich für verschiedene Frequenzen variiert muss man auch mit erheblichen Abbildungsartefakten rechnen. Eine linear über die Pupille variierende PTF führt zu einer gesamten Verschiebung in der Bildebene. Dies ist zunächst kein Problem, wird aber problematisch, wenn ortsabhängige Aberrationen eben verschiedene lineare PTF-Anteile haben und damit verschiedene Verschiebungen in der Bildebene, also Verzeichnung, resultiert. Für hochgenaue Positionsbestimmung kann die PTF damit relevant werden.

[6] Wichtig wird sie bei Systemen mit unsysmmetrischen Punktbildfunktionen bei denen eine Dekonvolution zur eigentlichen Bildgewinnung durchgeführt wird.

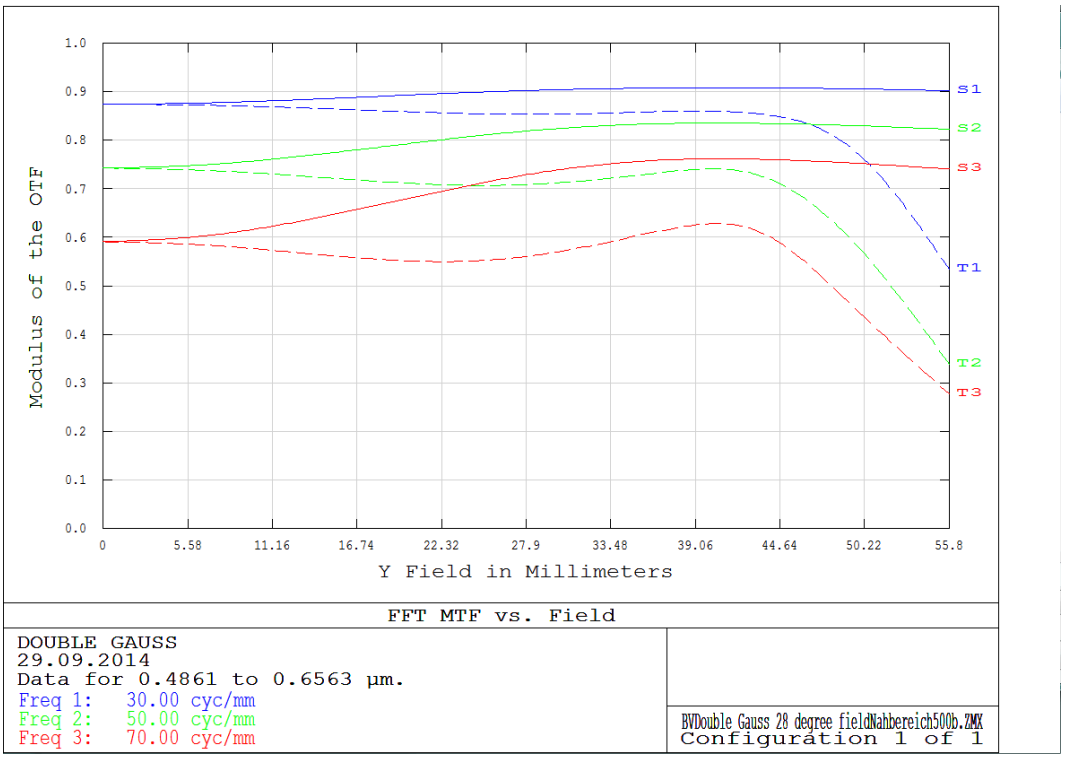
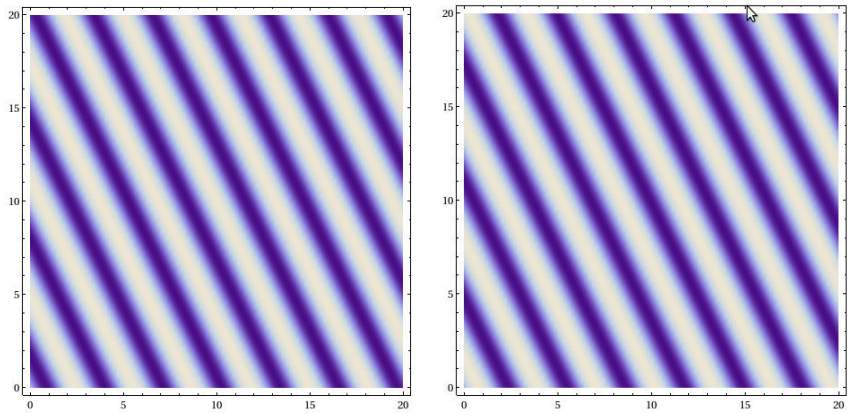


Abbildung 4.1.4: MTF in Auftragung über das Feld. Für drei verschiedene Raumfrequenzen wird die MTF in Abhängigkeit der Feldkoordinate dargestellt.

4.2 Zusammenschaltung von MTFs

MTF Kurven sind das Standardmittel zur Beschreibung der Abbildungsleistung von Optiken, aber auch kompletten Bildgebungssystemen (inklusive der Kamera und der elektronischen Verarbeitung). Bedenken Sie aber immer: Eine Beschreibung oder Analyse von Systemen mittels der MTF ist nur dann sinnvoll, wenn es sich um (zumindest in guter Näherung) lineare, verschiebungsinvariante Systeme handelt! . Nur in diesem Fall wird das Superpositionsintegral bei der Beschreibung des Ausgangssignals (Gl. 3.10) zu einem Faltungsintegral und nur dann macht (eben über den Faltungssatz) eine Übertragungsfunktion Sinn. Solange das System linear bleibt kann man in einzelnen sogenannten isoplanatischen Bereichen dennoch eine näherungsweise Verschiebungsinvarianz annehmen und so die MTF verwenden.

Die meisten nicht–trivialen Systeme bestehen aus *Teilsystemen*. Eine Analyse des Gesamtsystems kann in diesem Fall (oft) durch getrennte Analyse der Teilsysteme erfolgen. Die Frage ist dabei nur, wie Sie die Teilergebnisse zu kombinieren haben.



(a) PTF ist 0

(b) PTF $\neq 0$ bei relevanter Raumfrequenz

Abbildung 4.1.5: Die Phasentransferfunktion führt anschaulich zu einer Verschiebung des (hier dargestellten) abgebildenten Gitters. Beachten Sie, die PTF ist für unterschiedliche Gitterperioden unterschiedlich.

Wenn wir die gesamte Abbildungskette in einzelne Teile zerlegen, z.B. in die Optik, den optischen Tiefpassfilter, die Abtastung und Mittelung durch die CCD sowie die weitere (lineare, verschiebungsinvariante) Verarbeitung durch die Elektronik, dann erhalten wir:

$$
\begin{aligned}
o_{Optik} &= h_{Optik} * o && (4.8)\\
o_{TP} &= h_{TP} * o_{optik} && (4.9)\\
o_{CCD} &= h_{CCD} * o_{TP} && (4.10)\\
o_{El} &= h_{Elektronik} * o_{CCD} && (4.11)\\
&= h_{Elektronik} * h_{CCD} * h_{TP} * h_{Optik} * o && (4.12)\\
& && (4.13)
\end{aligned}
$$

Daraus folgt mit dem Faltungssatz aber direkt, dass die Gesamttransferfunktion durch die Multiplikation der Teiltransferfunktionen gegeben ist

$$H = H_{Elektronik}\, H_{CCD}\, H_{TP}\, H_{Optik} \tag{4.14}$$

Für den Betrag von H, also die MTF gilt damit[7]

$$|H| = |H_{Elektronik}|\,|H_{CCD}|\,|H_{TP}|\,|H_{Optik}| \tag{4.15}$$

Ganz allgemein erhalten wir also bei der Hintereinanderschaltung von N Systemen die OTF durch das Produkt der Einzeltransferfunktionen:

$$\boxed{H = \Pi_1^N H_i \qquad (4.16)}$$

Das Systemverhalten wird zwar maßgeblich durch die MTF mit der stärksten Abschwächung festgelegt, es sind aber eben wirklich alle Teil-MTFs von Bedeutung, so dass auch die Optimierung eines Systemteils, der schon bereits eine gute MTF aufweist eine weitere, wenn auch geringe, Gesamtverbesserung zur Folge hat.

4.2.1 Advanced: Zusammengesetzte Systeme und Kohärenz

Allerdings funktioniert die Vorgehensweise in der Optik leider nicht immer so reibungslos. Sie könnten auf die Idee kommen, dass wir bei einem zusammengesetzten optischen System, z.B. einem Mikroskop bestehend aus Tubuslinse und Mikroskopobjektiv, ganz analog die Übertragungsfunktionen multiplizieren können. Da der Betrag der Teilübertragungsfunktionen in jedem Fall kleiner eins ist würde das bedeuten, dass die MTF der Kette schlechter als die der Einzelglieder ist. Dies ist aber natürlich nicht der Fall, denn einzelne Teile der Optik gleichen Aberrationen anderer Teile aus so dass ein zusammengesetztes optisches System eine sehr viel bessere Abbildungsqualität als eine Einzelkomponente erzielen kann (deshalb bestehen Objektive ja auch aus mehreren Linsen).

Hier funktioniert die multiplikative Verbindung einzelner MTFs also nicht. Funktionieren würde sie, wenn zwischen den einzelnen Teilkomponenten jeweils die Intensität gebildet wird (und damit die Phase verloren geht) oder aber die Phase anderweitig gestört wird (z.B. Mattscheibe für Zwischenbild).

[7] Der Betrag der Multiplikation zweier komplexer Zahlen $A\,e^{i\phi}$ und $A\,e^{i\psi}$ ist $|c| = |A\,e^{i\phi} \cdot B\,e^{i\psi}| = |A||B||e^{i(\phi+\psi)}| = |A|\,|B|$.

Für das zusammengesetzte System (man denke sich ein Zwischenbild) gilt bei zunächst kohärentem Licht für das Eingangsfeld i:

$$o = h_1 * h_2 * i \tag{4.17}$$

$$O = H_1 \cdot H_2 \cdot I \tag{4.18}$$

h_1 und h_2 können sich gegenseitig durch die Faltung zu einem δ-Impuls wegheben. Nämlich genau dann, wenn eben H_2 die Phasenkonjugierte zu H_1 ist, also $H_2 = H_1^*$. Das zweite System gleicht also die Fehler der ersten Systems aus.

Wenn das Gesamtsystem inkohärent betrieben wird dann ändert sich daran nichts, denn erst am Ende der Kette wird die Intensität gebildet: $h_i = |h_1 * h_2|^2$. Bisher sind wir im Laufe dieses Buches immer von inkohärenten Punktbildverwaschungsfunktionen ausgegangen, die an sich grundlegendere bzw. einfachere Beschreibung ist aber die kohärente Punktbildfunktion. Das Betragsquadrat der kohärenten Punktbildfunktion ergibt die (in der Regel für uns relevante) inkohärente Übertragungsfunktion.

Üblicherweise werden die Teilsysteme in der abbildenden Optik auch genau durch diese inkohärente Antwort, also z.B. $|h_1|^2$ und $|h_2|^2$ beschrieben.

Wir betrachten zunächst aber erst die kohärente Übertragsfunktion (vgl. Gl. 4.18):

$$H_k(u) = H_1(u) \cdot H_2(u) \tag{4.19}$$

Und in unserem Fall (System 2 gleicht Aberration von System 1 aus) ergibt sich $H_k(u) = 1$.

Kommen wir zum inkohärenten Fall:

$$H_i(u) = \mathrm{FT}[|h(x)|^2] \tag{4.20}$$

$$= \mathrm{FT}[h(x) \cdot h^*(x)] \tag{4.21}$$

$$= H_k(u) * H_k^*(u) \tag{4.22}$$

also die Autokorrelation der kohärenten Übertragungsfunktion.

Da die *kohärente Übertragungsfunktion* H_k ja aber hier 1 ist, wird auch die inkohärente Übertragungsfunktion 1, also aberrationsfrei sein.

Was passiert, wenn wir eine Mattscheibe in die Zwischenbildebene einbringen (bei einigen optischen Systemen wird das gemacht)? In diesem Fall muss man beide

Teilsysteme getrennt betrachten, d.h.

$$|o_1|^2 = |h_1|^2 * |i|^2 \tag{4.23}$$

$$|o_2|^2 = |h_2|^2 * |o_1|^2 = |h_2|^2 * |h_1|^2 * |i|^2 \tag{4.24}$$

und damit die Übertragungsfunktion

$$H'(u) = (H_1 * H_1^*) \cdot (H_2 * H_2^*), \tag{4.25}$$

die eben nicht eins ergibt sondern in der Tat das Produkt zweier inkohärenter Übertragungsfunktionen darstellt.

Wir können die Teil-Transferfunktion eines Systems also nur dann zu einer Gesamt-Transferfunktion multiplizieren, wenn wir a) entweder ein kohärentes System betrachten oder b) die Teilsysteme ohne eine klare Phasenbeziehung gekoppelt sind (im Normalfall über eine Intensitätsbildung) oder c) keine Phasenänderung durch das nachgeschaltete System auftritt.

4.3 Pixelgeometrie

Der *Füllfaktor* eines Pixels hat maßgeblichen Einfluss auf das sich ergebende digitale Bild. Bisher hatten wir nur den Einfluss der Optik mittels der MTF beschrieben und die eigentlich Bildabtastung mit Delta-Impulsen, also idealen punktförmigen Pixeln angenommen. Natürlich ist das nicht realistisch, denn die lichtempfindliche Fläche wird aus Gründen der Lichteffizienz möglichst groß gewählt.[8]

Es ergibt sich also eine Mittelung des optischen Signals über die aktive Fläche des Pixels und diese hat natürlich auch einen Verschmierungseffekt zur Folge, der zu einer Kontrastreduzierung führt.

Im Sinne eines optimalen Kontrasts (nicht aber hinsichtlich der Lichteffizienz) wäre also eine rein punktförmige Abtastung zu bevorzugen. In diesem Fall ist die Punktbildfunktion des Sensors eben ein $\delta-$Impuls und führt daher zu einer MTF mit $MTF(u) = 1$. Da das Rauschen bei Digitalkameras meist die Hauptbegrenzung hinsichtlich der Bildqualität darstellt geht man aber eher den umgekehrten

[8] Ein Einsatz von Mikrolinsen und punktförmigen Pixeln ändert hieran nichts, in diesem Fall sind die aktiven Flächen letztlich die Mikrolinsenflächen.

Weg und verwendet zusätzlich Mikrolinsen vor den Sensoren, um eine möglichst maximale effektive lichtempfindliche Fläche pro Pixel zu erzielen.[9]

Das Verhältnis von lichtempfindlicher Fläche zur Gesamtfläche wird (linearer) Füllfaktor FF genannt.[10] In einer Richtung ergibt sich für einen Sensor mit Pixelabstand Δx damit

$$h(x) = \mathrm{rect}(x/(FF\,\Delta x)). \tag{4.26}$$

Die MTF ergibt sich wieder durch Fouriertransformation:

$$MTF(u) = |\mathrm{sinc}(FF\,\Delta x\,u)|. \tag{4.27}$$

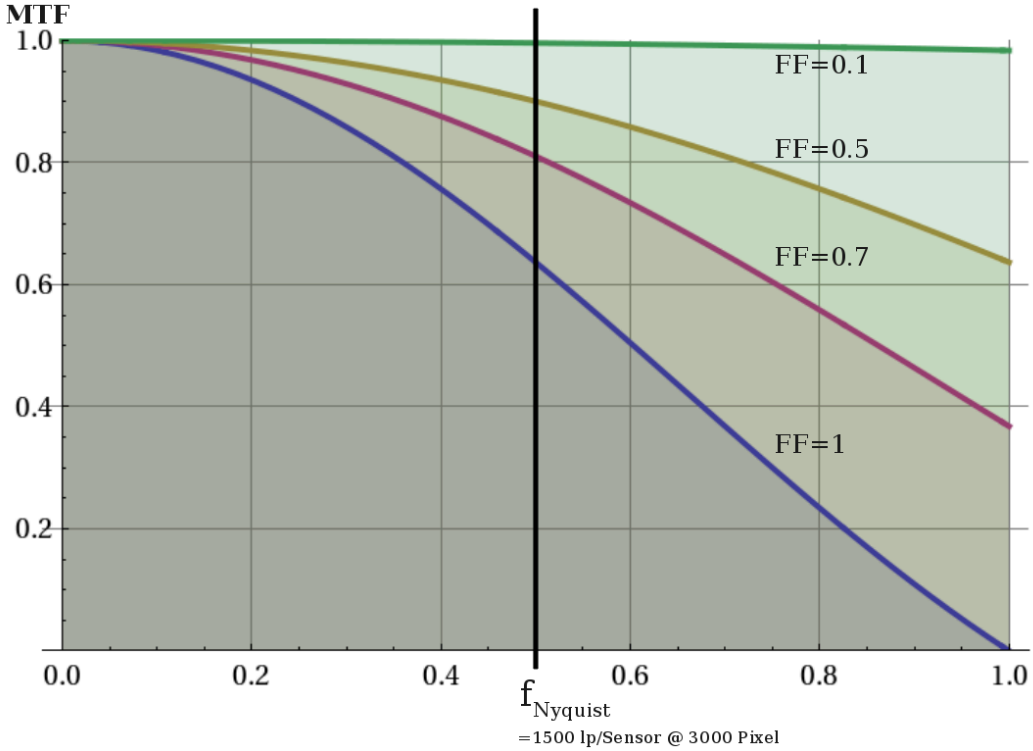


Abbildung 4.3.6: Füllfaktor bedingter Abfall der MTF für einen Detektor in eindimensionaler Näherung, vgl. Gleichung 4.27.

Für den Grenzfall eines Füllfaktors mit 100%, also einer vollkommenen Ausnutzung der Fläche ergibt sich damit ein Kontrastabfall für hohe Raumfrequenzen. Objektfrequenzen unterhalb der Nyquistfrequenz werden zwar korrekt erfasst,

[9] Ab einer Pixelgröße von ca. 8 Mikrometern nimmt in der Praxis die Effizienz der Linsen wohl ab und der Gewinn beim Füllfaktor wird (bei konstanter Elektronikgröße) gering.

[10] Oft wird der Füllfaktor in Bezug zur Pixelfläche definiert. Dieser dann quadratische Füllfaktor ergibt sich für rechteckige Pixelbereiche somit aus dem Quadrat des linearen Füllfaktors.

allerdings mit einem deutlich reduzierten Kontrast. Effektiv reduziert sich also gegenüber einem Sensor mit punktförmiger Abtastung die „Auflösung" merklich.

Die Grenzfrequenz liegt beim ersten Nulldurchgang des sinc. Für einen Füllfaktor von 1 und eine Pixelgröße von z.B. 5 μm ist die erste Nullstelle gegeben durch

$$\mathrm{sinc}(5\mu\mathrm{m} \cdot u) = 0 \tag{4.28}$$

also

$$\pi\, 5\mu\mathrm{m} \cdot u = \pi \tag{4.29}$$

$$u = \frac{1}{5\mu\mathrm{m}} = 0.2 \cdot 10^6 \frac{1}{m} = 200\,\mathrm{lp/mm} \tag{4.30}$$

Natürlich tritt aufgrund des periodischen Abtastvorgangs für entsprechend hohe Raumfrequenzen (kleine Strukturen) Aliasing auf. Wie bereits besprochen kann das Aliasing vermieden werden, wenn Signale nur bis zur Nyquistfrequenz $1/(2 \cdot 5)$ μm = 100 lp/mm vorkommen. Der Kontrast für diese Raumfrequenzen beträgt dann $\mathrm{sinc}(5\mu\mathrm{m} \cdot 100\,\mathrm{lp/mm}) = 0.64$.

Der Einfluss des Pixels in Zusammenhang mit der multiplikativen Kombination der Teil-MTFs (Gl. 4.15) erklärt auch, warum es durchaus Sinn machen kann, sehr viele Pixel zur Abtastung eines optischen Signals zu verwenden. Man verwendet also mehr Pixel, als man zunächst gemäß der begrenzten optischen Auflösung der Abbildung überhaupt benötigen würde. Ein System, mit gerade so vielen Pixel wie nach Nyquist notwendig sind, erzielt bei sehr feinen Strukturen aufgrund der endlichen Pixelfläche einen geringeren Kontrast als ein System mit *Überabtastung*. Besonders häufig und stark wird diese Überabtastung in der Mikroskopie eingesetzt, bei der oft viele Pixel eine einzelne optische Auflösungszelle aufnehmen.

Die Empfindlichkeit des Pixels muss auch nicht konstant über die aktive Pixelfläche sein sondern kann räumlich variieren, so dass dann eine MTF des Pixels resultiert, die eben nicht durch eine sinc-Funktion beschrieben werden kann. Insbesondere kann das auch dazu führen, dass das Verhalten des Sensors in horizontaler Richtung und vertikaler Richtung deutlich unterschiedlich ist (wenn die räumliche Pixelempfindlichkeit in beiden Richtungen unterschiedlich aufgeteilt ist). Und letztlich hängt die Pixelempfindlichkeit auch von der vorgeschalteten Mikrolinse ab bzw. von der Art der Beleuchtung (Winkelverteilung), die auf diese Mikrolinse trifft.

Unterschiedliche MTFs in horizontaler und vertikaler Richtung können natürlich ebenfalls ihren Ursprung in der Art des Auslesens haben, denn die zu transferierende Ladungen von benachbarten Pixeln nehmen (zumindest bei CCDs, aber auch bei Systemen mit analoger Videoübertragung) einen unterschiedlichen Weg, bis sie letztlich digital verglichen werden (vgl. Kapitel 9). Am extremsten ist der Unterschied bei analogen Videokameras. Hier ergibt sich eine gute Trennung der unterschiedlichen Zeilen während innerhalb der Zeile oft eine deutliche Verschmierung der Information auftritt. Aber auch bei modernen digitalen Sensoren (insbesondere bei sehr kleinen Pixeln) kann das *Übersprechen* zu benachbarten Pixeln (*Crosstalk*) in verschiedene Richtungen unterschiedlich stark sein [18]). Es kann für verschiedene Aufgabenstellungen also durchaus einen Unterschied machen, wie der Bildsensor orientiert ist. Besonders kritisch kann das elektronische Übersprechen bei Farbsensoren mit den typischerweise eingesetzten Bayer-Farbmasken werden.

4.4 Einige weitere Beispiele

Wir wollen im folgenden einige für die industrielle Bildverarbeitung wichtige Modulationstransferfunktionen beispielhaft besprechen. Wir starten hierzu mit der *Defokussierung*:

Dazu nehmen wir ein vereinfachtes geometrisch-optisches Strahlenmodell für den Defokus an (Kreisscheibe), d.h. ein defokussierter Objektpunkt wird zu einer Kreisscheibe mit dem Durchmesser d.

Wir können damit sagen, dass die (inkohärente) Punktbildfunktion h durch

$$h = \mathrm{circ}(r, d) \tag{4.31}$$

gegeben ist.

Für rotationssymmetrische Funktionen geht die Fourier-Transformation in die Fourier-Bessel Transformation über [17]. Und damit

$$MTF(u) \sim |\int_0^\infty \mathrm{rect}(\rho/(D/2))J_0(2\pi u\rho)\rho d\rho| \tag{4.32}$$

$$= |2\frac{J_1(\pi u D)}{\pi d u}| \tag{4.33}$$

mit der Besselfunktion 1. Ordnung, 1. Art J_1).

Näherungsweise kann man den Defokus natürlich auch eindimensional betrachten. In diesem Fall wird aus dem circ ein rect und es ergibt sich als MTF ein sinc. Wenn man den Blurkreisdurchmesser mit B Pixeln annimmt dann wird entsprechend die MTF

$$MTF(u) = |\mathrm{FT}[\mathrm{rect}(x/(\Delta x \cdot B))]| \sim \mathrm{sinc}(B \cdot u\Delta x) \tag{4.34}$$

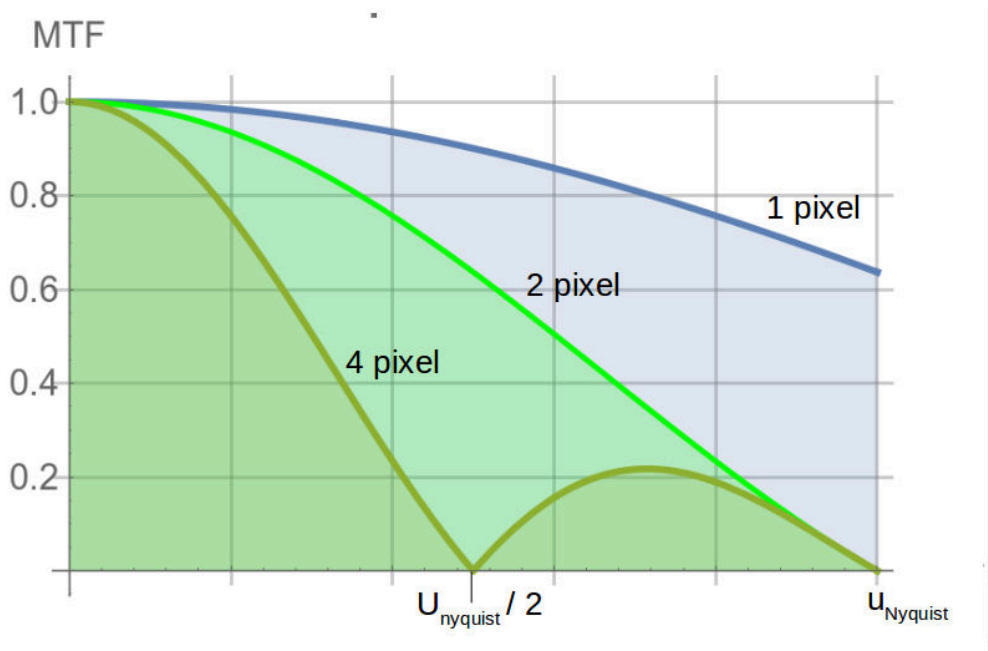


Abbildung 4.4.7: *Defokusbedingter Abfall der MTF für einen punktabtastenden Detektor in eindimensionaler Näherung, vgl. Gleichung* 4.34. Dargestellt sind Kurven für verschiedene Blurkreisdurchmesser (angegeben in Pixel).

Abb. 4.4.7 zeigt für einige verschiedene Defokussierungen diesen Zusammenhang[11]. Beachten Sie: der Defokusanteil für einen Durchmesser von einem Pixel entspricht dem MTF Abfall aufgrund der endlichen Pixelausdehnung (Gl. 4.27)

[11]Eine genauere Rechnung ist in [12] sowie in [21] dargestellt. Es ergibt sich eine MTF die proportional zu $1/u\ \sin(\alpha u - \beta u^2)$ ist.

und muss nicht mehrfach einberechnet werden. Wenn also die MTF für die Defokussierung verwendet wird kann die Pixel-MTF vernachlässigt werden.

Es zeigt sich in der MTF ein Bereich mit Kontrasten von Null (bzw. nahe Null) während höhere Frequenzen dann wieder einen nennenswerten Kontrast aufweisen. Man spricht von *spurious resolution*.

Ganz ähnlich gehen wir zur Berechnung der MTF für das *Verwackeln*, also die *Bewegungsunschärfe* bei der Bildaufnahme vor, die oft bei der Kontrolle von Teilen auf Fließbändern auftritt. Es ergibt sich in der Bildebene eine eindimensionale Verschmierung, d.h. die Punktbildfunktion ist wieder eine rect–Funktion und entsprechend wird die MTF eine sinc–Funktion:

$$MTF(u) = |\mathrm{FT}[\mathrm{rect}(x/\Delta x)]| = |\mathrm{sinc}(u\Delta x)| \tag{4.35}$$

Eine komplett zufällige Bewegung (*jitter*) wird in [12] beschrieben und ein schönes (und extremes) Beispiel sehen Sie am Kapitelanfang.

Bildfehler werden ebenfalls mittels der MTF erfasst. Beim Vergleich verschiedener Objektive ist dies der zentrale Sinn der MTF-Kurven. *Aberrationen* eines Bildpunktes führen ja direkt zu einer Verschmierung des idealen (delta-förmigen) Punktes mit der Punktbildfunktion $h(x)$. Damit ist klar, dass die MTF sich entsprechend ändert, denn sie kann wieder durch Fouriertransformation von $h(x)$ gewonnen werden. Die MTF ist damit ein sehr sensitives Beurteilungskriterium zur Beurteilung von Aberrationen.

4.5 Advanced: Grenzen

Solange man sich auf kleine Regionen innerhalb des Felds bezieht liefert die faltungsbasierte Beschreibung (bzw. gleichbedeutend das Transferfunktionsmodell), die wir hier vorgestellt haben und die Grundlage der optischen Abbildungstheorie darstellt, in der Regel exakte Ergebnisse.

Probleme treten allerdings dann auf, wenn die relevanten Objektstrukturen sehr klein werden (Mikroskopie). Grundlage unserer Theorie war die Zerlegung des

Objekts in unendlich viele ideale Punkte, die wir (gewichtet mit der Objektfunktion) summiert haben. Die Fouriertransformation des idealen Punkts ist 1. Gleichzeitig sollte das die Feldverteilung des Lichts in der Pupillenebene des Abbildungssystems sein.

In der Praxis haben wir aber eben nicht einen idealen Punkt sondern sehr eng benachbart unendlich viele ideale Punkte. In diesem Fall interferieren die Felder aufgrund der einzelnen Punkte. Interferenz bedeutet hier, dass anstatt der Intensität die elektrischen Felder summiert werden und erst nach der Summation die Intensität durch Quadrieren gebildet wird. Die eng benachbarten Punkte sind dabei (teil-)kohärent. Dies führt dann zu einer komplizierteren Lichtverteilung in der Pupillenebene.

Ganz generell kann man sagen, dass die Abbildungstheorie es uns erlaubt, das Lichtfeld in der Bildebene basierend auf dem Lichtfeld in der Objektebene zu berechnen. Aber das Objektfeld in der Objektebene ist *nicht* das Objekt selbst sondern nur die Lichtverteilung *nach* dem Objekt. Der Zusammenhang dieser Lichtverteilung mit dem Objekt ist hochgradig kompliziert wenn die relevanten Objektstrukturen in der Größenordnung der Wellenlänge des Lichts oder darunter liegen.

Eine komplette Berechnung des Bildes beinhaltet also zunächst eine Simulation der Wechselwirkung des Beleuchtungslichts mit der Objektstruktur und die Berechnung des sogenannten Nahfelds nach der Struktur bevor die Abbildungsmethodik angewendet werden kann.

Für die meisten Bildverarbeitungsprobleme müssen wir diese Details nicht beachten. Bei stark vergrößernden mikroskopischen Anwendungen werden sie aber relevant.

Aber auch bei hochgenauen Positionsbestimmungen können Probleme auftreten. Wie genau die Pupille der Optik ausgeleuchtet wird hängt in komplexer Weise sowohl von der lokalen Objektstruktur als auch der Beleuchtung ab. D.h. es ist für uns meist nicht vorhersehbar, welcher Objektpunkt wie die Pupille des Abbildungssystems ausleuchtet.

Solange die Abbildung ohne Bildfehler erfolgt ist das nur dahingehend ein Problem, dass die lokale beugungsbedingte Auflösung variiert. Sobald aber Aberra-

tionen auftreten — und die in der Bildverarbeitung eingesetzten Objektive sind bei kleinen Blendenzahlen alle aberrationsbehaftet — können lokale Verschiebungen des Bildpunktes auftreten. Abb. 4.5.8 verdeutlicht das Problem. Wenn z.B. Licht lediglich auf die dargestellte Pupillenzone fällt, dann führt die dort vorhandene Aberration zu einer minimalen Strahlverkippung[12]. Diese Strahlablenkung führt dann zu einer Verschiebung in der Bildebene und damit zu einer Fehlmessung der Position. Die Positionsbestimmung wird damit abhängig von der Mikrostruktur des Objekts. Daher lässt sich die Variation auch nicht durch eine Kalibrierung umgehen.

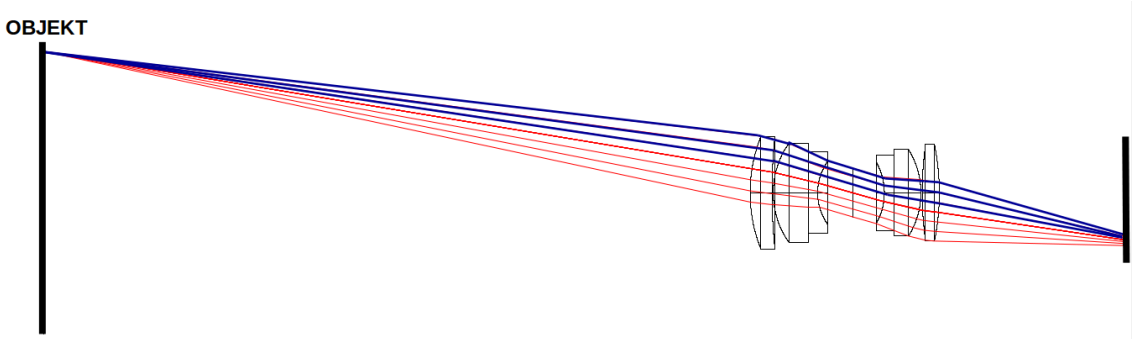


Abbildung 4.5.8: Eine ungleichmäßige Pupillenausleuchtung führt in der Kombination mit Aberrationen des Abbildungssystems zu einer Verschiebung des Bildpunkts. Angenommen, ein Objekt streut Licht vermehrt in die blau dargestellten Zonen, dann ergibt sich in der Bildebene ein im Mittel höher liegender Bildpunkt (Schwerpunkt). Streut das Objekt bevorzugt in die rot dargestellten Bereich ist entsprechend der Bildpunkt eher tiefer liegend. Das Messergebnis bzw. das Bild hängt so von der Streucharakteristik der Objektoberfläche ab.

Je kleiner die Eintrittspupille ist, desto geringer ist der Effekt, aber natürlich führt die geringe Eintrittspupille auch zu einer generell reduzierten Auflösung der Abbildung. Probleme aufgrund ungleichförmiger Pupillenausleuchtung potenzieren sich, wenn Abbildungssysteme geteilte Pupillen verwenden, d.h. unterschiedliche Pupillenbereiche unterschiedliche optische Funktionen erfüllen.

[12]Die lokale Aberration in der Pupille können sie als lokalen Phasenkeil betrachten, der analog einem Prisma das Licht ablenkt.

Kapitel 5

Laterale Auflösung

In diesem Kapitel werden Sie unter anderem lernen

- welche verschiedene Arten der „Auflösung" in der Bildverarbeitung relevant sind,
- wie man laterale Auflösungen spezifiziert und berechnet,
- welche lateralen Auflösungen für verschiedene Aufgabenstellungen gewählt werden,

- aus welchen Anteilen sich die laterale Auflösungsbeschränkung zusammensetzt,
- welche typischen Bildfehler beim Einsatz von Objektiven zu erwarten sind,
- welche Pixelgröße und welche Pixelzahl für eine Anwendung zu wählen sind und
- wie sich kohärente Abbildungen von inkohärenten Abbildungen unterscheiden.

Die *Bildqualität* der letztlich mit einem BV-System gewonnenen Bilder hängt von einer Vielzahl von Parametern ab. Wesentlich sind die Abbildungsoptik, die Beleuchtung und natürlich die eigentliche Bildaufnahme durch den Bildsensor. Alle Teilsysteme haben einen Einfluss und sind zu betrachten. Noch wesentlicher ist aber zunächst, exakt zu definieren, was man für eine gegebene Anwendung unter „Bildqualität" konkret versteht. In der Regel wird die Qualität des Bildes für die nachfolgende computergestützte Auswertung in komplexer Weise vorrangig von

- Auflösungen,
- Kontrasten,
- Dynamiken und
- Störsignalen (vor allem Rauschen)

abhängen.

Ein für viele Aufgabenstellungen sehr wichtiger, wenn nicht *der* zentrale Parameter ist dabei die *Auflösung*. Und vermutlich werden Sie sich wundern, dass wir für einen Parameter ein ganzen Kapitel mit sehr vielen Seiten verwenden.

Und das, obwohl der Schwerpunkt dieses Kapitels nur der (für den Bildverarbeiter) Hauptaspekt der Auflösung, nämlich die laterale Auflösung bei inkohärenten Abbildungen, ist. Dabei werden wir zunächst den wichtigsten Fall, nämlich die

Auflösung im Sinn der Fähigkeit zur Trennung feiner Details ausführlich besprechen. Für nahezu alle industriellen Bildverarbeitungssysteme muss diese laterale Auflösung beachtet werden.

Und Vielleicht ist das Kapitel wirklich zu ausführlich. Das Allerwichtigste kann man in etwa folgendermaßen zusammen:

- Es gibt ein theoretisches Limit für die laterale optische Auflösung. Wir verwenden dabei (für die bildseitige laterale Auflösung) die Formel

$$r'_A = 0.61\frac{\lambda}{\mathrm{NA}'} \approx 1.22\frac{\lambda\, f'}{D} = 1.22\lambda\, K \tag{5.1}$$

- Die Pixelierung wählt man meist passend zur optische Auflösung, typischerweise 1..3 Pixel pro r'_A.
- Neben der Auflösung zur Trennung bzw. Detektion feiner Strukturen sind für viele industrielle Aufgabenstellungen aber auch geometrische Messungen notwendig. Hierbei ist meist die Positionsauflösung bzw. meist genauer: Die Messunsicherheit einer Positionsbestimmung, entscheidend.

Warum, das so ist, was es genau bedeutet und wann konkret welche Auflösungen von Bedeutung sind und wie sie zu wählen sind, darum gehts in diesem — wie gesagt länglichen und teilweise auch anstrengenden — Kapitel.

Wir werden die beiden Hauptaufgabenstellungen — Detektion und Messung — getrennt betrachten. In der Praxis ergeben sich oft kombinierte Anforderungen und entsprechend muss man bei der Auslegung beide Bereiche berücksichtigen.

5.1 Zur Definition von Auflösung

Auflösung ist ein extrem allgemeiner und damit problematischer Begriff. Der Begriff „Auflösung" wird in der Optik von verschiedenen Leuten und anwendungsabhängig völlig unterschiedlich definiert bzw. benutzt.

Eine gute Definition liefert Hobbs: „Resolution is the ability to look at an object and see what is there, in an unambigous way that does not depend on our choice

of model.” [24]

Damit wird schon deutlich, dass es sich um eine Art Rekonstruktionsproblem handelt, bei dem man ein Modell des Objekts hat und basierend auf dem Bild Parameter dieses Modells schätzt, z.B. den Abstand zwischen zwei idealen Punkten oder aber die Anzahl vorhandener Punkte. Klar ist damit auch, dass die Auflösung untrennbar verbunden mit dem Signal-Rausch-Verhältnis des Bildes bzw. der Messung ist.

Meist macht man sich in der Bildverarbeitung allerdings nicht so detaillierte Gedanken, sondern unterscheidet einige praxisrelevante Sonderfälle. Unterscheiden müssen wir vor allem zwischen Auflösungen im Sinne von

- der Unterscheidung bzw. Trennung eng benachbarter Objekte (*Zweipunkt Trennung*),
- der Bestimmung der Position eines Punkts oder Linie (*Positionsauflösung*), entweder als Auflösung oder als Messunsicherheit,
- der Auflösung einer axialen Messung (Höhen- bzw. „3D-”Messung) und
- der reinen Auflösung eines Bildsensors als Anzahl der verfügbaren Pixel.

Diese unterschiedlichen Auflösungen hängen teilweise miteinander zusammen, sind aber in keinem Fall durcheinander ersetzbar. Als Beispiel soll die Interferometrie betrachtet werden: Ein gutes Interferometer kann Höhenabweichungen auf optischen Flächen mit einer Auflösung von z.B. 1 nm messen. Die laterale Auflösung im Sinne einer Zweipunkt Trennung kann aber unter Umständen lediglich 1 mm betragen. Darüber hinaus macht es oft einen erheblichen Unterschied, wie ein näher spezifizierter Auflösungsparameter genau gemessen wurde (z.B. Höhenmessung an Spiegel oder an rauer Oberfläche).

Am allerwichtigsten ist zunächst aber: Verwechseln Sie nicht Auflösung mit *Genauigkeit* bzw. *Messunsicherheit*! Wir werden im Kapitel 12 noch ausführlich auf die Unterschiede eingehen. Nehmen Sie aber zunächst hin, dass eine Auflösung von 1 nm *nicht* bedeutet, dass sie etwas auf 1 nm *genau* gemessen haben.

5.2 Zweipunkttrennung

Wenn ein Optiker im Rahmen der Bildverarbeitung ohne nähere Angabe von „Auflösung" redet, dann ist zunächst in der Regel die laterale Auflösung des zweidimensionalen Bildes im Sinne einer Zweipunkt-Auflösung gemeint. Die Frage ist letztlich: Wenn mein Objekt sehr eng lateral benachbarte feine Strukturen aufweist, kann ich diese noch „auflösen" bzw. (mit ausreichendem Kontrast) getrennt detektieren?

Auflösung in diesem Sinn ist somit die Fähigkeit eines Systems, feine Details zu unterscheiden.

Die in diesem Sinne verstandene Auflösung (wie man sie auch immer im Detail definiert, s.u.), hängt entscheidend sowohl von der Optik als auch dem Bildsensor und der Algorithmik ab. Dementsprechend müssen alle Faktoren bei der Beurteilung der Auflösung in Betracht gezogen werden.

Für die Auslegung der Hardware sind zunächst drei bzw. vier Faktoren bei der Frage nach der lateralen Auflösung im Sinne der Zweipunkttrennung zu beachten:

1. Können die Strukturen überhaupt theoretisch nach den Gesetzen der Physik noch trennbar sein?
2. Ist die eingesetzte Optik ausreichend gut korrigiert, um *in der Praxis* die Strukturen noch trennen zu können?
3. Ist der eingesetzte Bildsensor in der Lage, diese Trennung noch zu ermöglichen? Wird also die optische Abbildung ausreichend gut abgetastet?
4. Ist das Signal-Rausch-Verhältnis hoch genug, um die Trennung durchzuführen?

Im Sinne der Trennung ist die Auflösung des optischen Signals damit untrennbar mit der Abtastung des Signals (Pixelierung und Bildsensoreffekte), dem Rauschen und der Algorithmik verbunden. Das macht die Thematik unübersichtlich und schwierig.

Vergleichen Sie hierzu Abb. 5.2.1 bei der zwei eng benachbarte ideale Punkte unterschiedlicher Intensität mit unterschiedlichem Rauschen bei einer gewissen optischen Auflösung (diese führt zu einer Verbreiterung des Bildes) aufgenommen werden. Eine etwaige zusätzliche Pixelierung durch den Sensor wird hierbei zunächst vernachlässigt[1]. Klar ist, dass es sehr schwierig ist, bei stark verrauschten Bildern, eng benachbarte Punkte zu trennen.

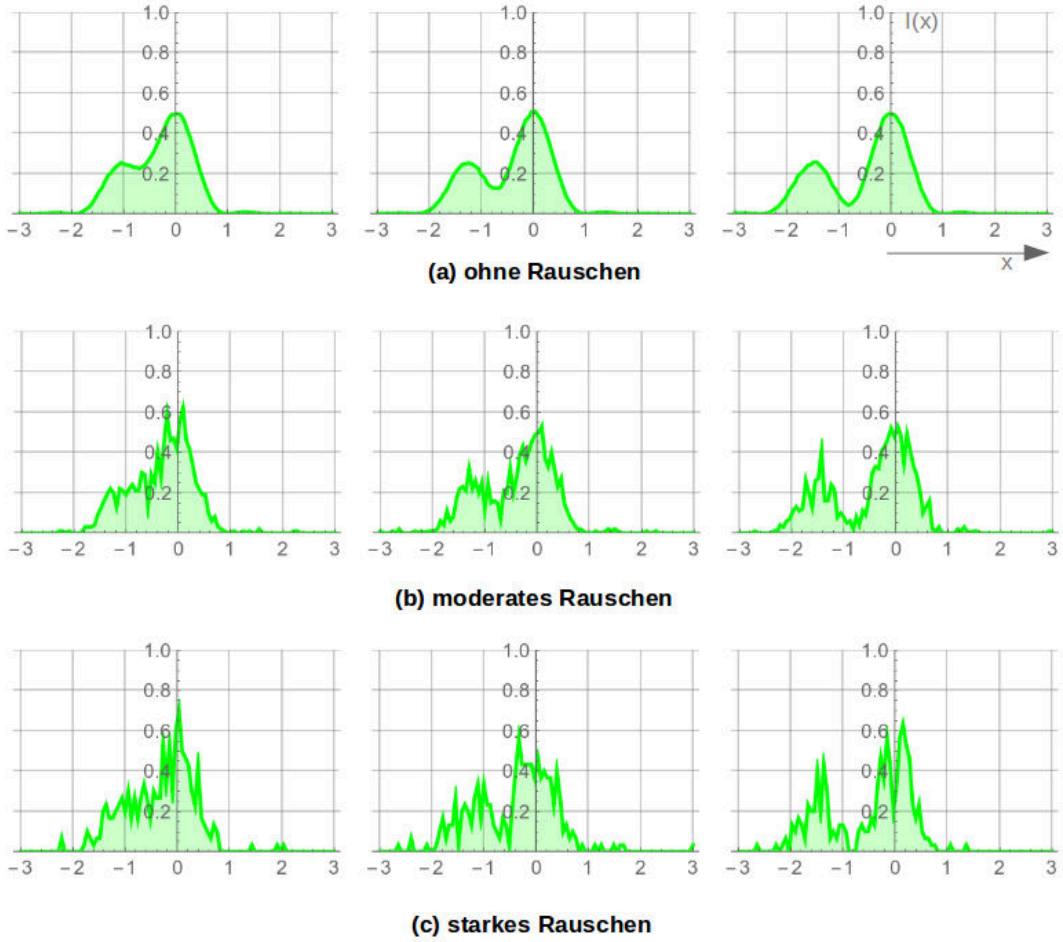


Abbildung 5.2.1: Die Auflösung im Sinne einer Zweipunkttrennung kann nicht unabhängig vom Rauschen betrachtet werden. Optische Auflösung, Pixelierung, Rauschen, verschiedenen Effekten des Bildsensors und Auswertealgorithmik entscheiden zusammen, *ob eine Punkttrennung gelingt. Die Zeilen entsprechend unterschiedlichen Signal-Rausch-Verhältnissen, die Spalten unterschiedlichen Abständen der beiden Punkte. Einer der beiden Objektpunkte ist dabei doppelt so hell wie der zweite Punkt.*

[1]bzw. Sie können sich einen Sensor mit sehr vielen Pixeln vorstellen

Abbildung 5.2.2: Eine Reduktion der Auflösung ergibt sich zum einen durch den Bildsensor (blaue Quadrate), zum anderen durch das optische System. Das (ebenfalls relevante) Rauschen kann sowohl durch das Licht selbst wie auch den Bildsensor eingebracht sein.

Laien denken üblicherweise beim Begriff „Auflösung" lediglich an die „Sensorauflösung" (ein weiterer ungenauer Begriff) bzw. die Anzahl der Pixel des Bildsensors. Dieser Einfluss ist für einen idealen monochromatischen Bildsensor vergleichsweise trivial wenn wir die MTF des Pixels vernachlässigen. Aus dem Datenblatt entnimmt man die Anzahl der Pixel und den Abstand einzelner Pixel, den sogenannten *Pixelpitch* p und kann so die Auflösung r_o auf dem Objekt angeben: $p/\beta' = r_o$. Man rastert also einfach das Objektfeld mit der Anzahl der Pixel (z.B. 1600 x 1200 bei einem 2 MPixel Sensor).

Wesentlich interessanter ist die theoretische Grenze der Auflösung. Wir können sie dennoch relativ kurz abhandeln. Ein zentrales Ergebnis der Beugungstheorie (siehe z.B. [17]) ist, dass sich bei der Abbildung eines idealen Objektpunkts ein ausgedehnter Bildpunkt, das sogenannte *Airy-Scheibchen*, mit dem Durchmesser

$$D'_A = 1.22\frac{\lambda}{\mathrm{NA}'} \approx 2.44\frac{\lambda\, f'}{D} = 2.44\,\lambda\, K \tag{5.2}$$

ergibt (siehe z.B. [22], Abb. 5.2.3). K bezeichnet wieder die Blendenzahl und D den Durchmesser der Eintrittspupille der Optik. Den Durchmesser des äquivalenten Bereichs in der Objektebene D_A finden wir dann einfach wieder über den Abbildungsmaßstab, also $D'_A = \beta' D_A$.

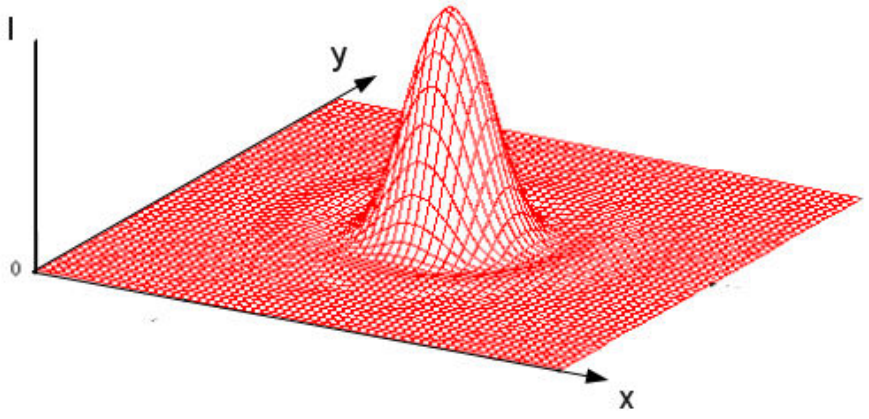


Abbildung 5.2.3: Ein aberrationsfreies System generiert als Bild eines idealen Objektpunkts eine ausgedehnte Intensitätsverteilung. Der Durchmesser (von Nullstelle zu Nullstelle) des „Airy-Patterns" ist gerade 2.44 $\lambda\, K$.

Zwei eng benachbarte ideale Punkte führen dementsprechend zu einer

Überlagerung der beiden ausgedehnten Bildpunkte. *Rayleigh* hat definiert, dass wir die Punkte als noch trennbar betrachten sollten, wenn das Maximum des einen Punktbildes gerade in das Minimum des zweiten Punktbildes fällt. Eine genaue Analyse zeigt, dass dann noch ein Kontrast von 15% das Vorhandensein von zwei statt einem Punkt anzeigt bzw. dass das Minimum der beiden Punkte 73.5% der Intensität der Maxima aufweist. Als *Auflösung nach Rayleigh* haben wir also $r'_A = D'_A/2$:

$$r'_A = 0.61\frac{\lambda}{\mathrm{NA}'} \approx 1.22\frac{\lambda\, f'}{D} = 1.22\lambda\, K \tag{5.3}$$

Innerhalb dieses Radius findet sich 84% der Leistung des Spots. Der Kontrast der Eindellung (diese liegt bei ca. 75% vom Maximum) ist etwa 15%.

Beispiel: Sie benutzen eine korrigierte Optik die ein beugungsbegrenztes Abbildungsverhalten bei der verwendeten Blendenzahl von 9 erzielt. Welchen Durchmesser erwarten Sie für die Abbildung eines idealen Punkts?

Lösung: $2\,r' = 2.44\,\lambda\,K = 2.44 \cdot 500\,\mathrm{nm} \cdot 9 = 10.9\ \mu\mathrm{m}$

Natürlich ist das Kriterium zunächst willkürlich. Letztlich kann die Auflösung im Sinne einer Unterscheidung von eng benachbarten Objekten nicht unabhängig vom Rauschen beurteilt werden. Nehmen Sie an, dem Bildsignal zweier gemäß Rayleigh benachbarter Punkte sei starkes Rauschen überlagert und es sind nur wenige Pixel über die relevante Ausdehnung vorhanden, um z.B. das Rauschen durch Mittelung zu reduzieren. Dann können — aufgrund des Rauschens — die benachbarten Punkte eben offensichtlich nicht unterschieden werden (vgl. Abb. 5.2.1).

Eine wirklich objektive Beurteilung, ob Punktetrennung möglich oder nicht ist, ist nur möglich, wenn sowohl die optische Auflösung also auch die Sensorauflösung, das Rauschen und die Algorithmik bekannt sind.

Dieses sogenannte „Beugungslimit" ist für uns in der industriellen Bildverarbei-

tung eine unumgehbare theoretische Grenze. Da die numerische Apertur bei Systemen in Luft auf 1 begrenzt ist können wir mit sichtbarem Licht ($\lambda_{min} > 400$ nm) in keinem Fall Strukturen kleiner 250 nm auflösen.[2] Für typische Objektive im makroskopischen Bereich ist eine Blendenzahl von K $<$ 1.4 schon sehr schwierig zu erzielen und K kleiner 1 ist kaum zu erreichen. Generell ist immer daran zu denken, dass hohe numerische Aperturen auch kleine Schärfentiefen bedingen.

Bei der mikroskopischen Bildgebung kann ein zusätzlicher Auflösungsgewinn durch eine hochaperturige Beleuchtung $\mathrm{NA_{ILL}}$, erzielt werden (siehe z.B. [39] und Abschnitt 5.14) Die sich ergebende Auflösung kann näherungsweise durch

$$r'_A = 0.61 \frac{\lambda}{\mathrm{NA} + \mathrm{NA_{ILL}}} \tag{5.4}$$

berechnet werden.[3]

Gegenüber der Zweipunktauflösung von Rayleigh hat *Abbe* die Abbildung eines periodischen Gitters untersucht. Die Herangehensweise ist also eng verwandt mit der Grundidee der MTF. Entscheidend ist bei der Analyse, dass die Beugung von Licht an einem periodischen Gitter der Periode d zu einer Lichtablenkung um den Winkel $\alpha = \arcsin(\lambda/d)$ führt [22]. Abb. 5.2.4 zeigt, dass eine Blende (hier dargestellt am konzeptionell besonders einfach darzustellenden Beispiels des doppeltelezentrischen Systems bzw. Mikroskops) dazu führen kann, dass das am Gitter gebeugte Licht nicht die Bildebene erreichen kann und somit kein Bild des Gitters entsteht. Als Auflösung ergibt sich gemäß Abb. 5.2.4 für die kleinste halbe Gitterperiode

$$r'_A = 0.5 \frac{\lambda}{\mathrm{NA}'} \qquad \text{bzw.} \qquad r_A = 0.5 \frac{\lambda}{\mathrm{NA}} \tag{5.5}$$

5.3 Advanced: Überauflösung

Wir wollen noch kurz den Begriff *Überauflösung* ansprechen. Dieser wird von unterschiedlichen Personen sehr unterschiedlich verwendet.

[2] Bei periodischen Strukturen kann die minimal auflösbare Periode gemäß *Abbe* zu 200 nm angenommen werden.

[3] Anschaulich gesprochen: Schräg auf die Struktur einfallendes Licht wird in die Apertur des Mikroskops gebeugt. Information vom Objekt kann so erfasst werden.

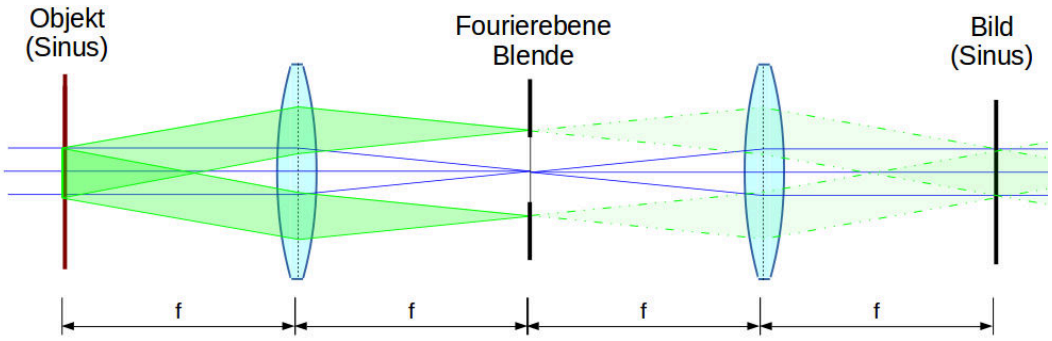


Abbildung 5.2.4: Zur Auflösung nach Abbe: Die Abbildung eines periodischen Objekts (Gitter, hier Sinus) führt zu Peaks in der Blendenebene. Wenn die Peaks außerhalb des Durchlassbereichs der Blende liegen kommt kein die Struktur beschreibendes Licht bei der Bildebene an. Die Struktur des Gitters ist also zu fein und daher wird das Gitter nicht abgebildet.

Im Sinne unseres Kapitels über laterale Auflösung bedeutet „Überauflösung", dass feine Strukturen, die (deutlich) enger als nach Rayleigh benachbart sind, noch getrennt beobachtbar sind.

Zunächst scheint das unmöglich zu sein und das Modell von Abbe zeigt deutlich, dass z.B. das Licht, das an einem sehr feinen Gitter gebeugt wird, nicht durch das optische Systeme transferiert wird und daher dieses Gitter definitiv nicht abgebildet werden kann.

Ein Ausweg ist das direkte Rastern ohne optische Abbildung im sogenannten *Nahfeld* des Objekts, eine Technik, die im Rahmen der industriellen Bildverarbeitung nicht sinnvoll einsetzbar ist. Sobald Nichtlinearitäten bei der Objekt-Licht Wechselwirkung genutzt werden können (z.B. in der sogenannten *STED Mikroskopie*) kann ebenfalls die Auflösung deutlich verbessert werden.

Ein anderer Ausweg ist rein mathematischer Natur. Wenn das abzubildende Objekt begrenzt ist dann ist die Fouriertransformation des Objekts analytisch fortsetzbar. Das bedeutet, man bestimmt die Fouriertransformierte des Bildes und setzt das sich so ergebende Spektrum analytisch zu höheren Ortsfrequenzen fort. Dieses ausgedehnte Spektrum wird dann invers fouriertransformiert, um ein höher aufgelöstes Bild zu erhalten.

Am Beispiel des Gitters: Wenn das Gitter räumlich begrenzt ist dann besteht das Spektrum des Gitters nicht nur aus zwei diskreten Punkten (kohärenter Falle) sondern auch ein wenig Licht, das zu tieferen Raumfrequenzen gelenkt wurde und so

das optische System passieren kann. Basierend auf diesem letztlich an der Objektberandung des Gitters gestreuten bzw. gebeugten Lichts muss man dann auf das Gitterspektrum rückschließen.

Dies ist in der Praxis umso schwieriger (bzw. unmöglicher), je stärker das Rauschen der detektierten Information ist. Ebenfalls ist es notwendig, die optische Transferfunktion des Abbildungssystems hochgenau zu kennen.

In der Praxis sind entsprechende echte Überauflösungsansätze, die die *beugungsbegrenzte* Auflösung im Sinne einer Punktetrennung verbessern daher leider nicht praktikabel und weit von einer industriellen Anwendung entfernt.

5.4 Advanced: Dekonvolution

Weniger schwierig ist es, die Auflösung bei aberrierten Bildern (z.B. defokussierte Aufnahmen) deutlich zu verbessern. Entsprechende Ansätze werden *Dekonvolutionsverfahren* genannt. Die einfachste Variante hierzu ist der *inverse Filter*. Wenn die Abbildung mit der Übertragungsfunktion $H(u)$, also dem Punktbild $h(x)$ erfolgt, dann kann das Orginalobjekt einfach erhalten werden indem im Fourierraum mit der inversen Übertragungsfunktion multipliziert wird:

$$i(x) = h(x) * o(x) \tag{5.6}$$

$$\rightarrow I(u) = H(u)\,O(u) \tag{5.7}$$

$$I'(u) = I(u) \cdot \frac{1}{H(u)} = O(u) \tag{5.8}$$

$$o'(x) = \mathrm{FT}[I'(u)] = o(x) \tag{5.9}$$

Probleme treten an den Nullstellen der Übertragungsfunktion $H(u)$ auf, denn an diesen Stellen wird durch Null dividiert und etwaiges Rauschen wird stark verstärkt.[4]

Dieses Grundprinzip der Dekonvolution wird durch eine Vielzahl verbesserter Methoden angegangen, in der industriellen BV werden entsprechende Ansätze bislang allerdings kaum genutzt. Stattdessen wird versucht, die Aberrationen an sich auf ein sinnvolles Maß zu begrenzen.

[4] Details zur Anwendung und Theorie: siehe [41]

In der industriellen Bildverarbeitung wird in aller Regel eine inkohärente Beleuchtung eingesetzt. Bei kohärenter Bildgebung wird die Definition der Auflösung noch ein wenig komplizierter, denn das Bild hängt entscheidend nicht nur vom Abstand zweier feiner Strukturen sondern auch von derer Phasendifferenz ab. Ein Vorschlag für eine Vorgehensweise zur Charakterisierung findet sich bei Horstmeyer et al. [28].

Schließlich soll darauf hingewiesen werden, dass bei zusätzlicher Information auch höhere Auflösungen möglich sind. Wenn z.B. bekannt ist, dass garantiert nur zwei gleich helle Punkte vorhanden sind und nur der Abstand zwischen den beiden Punkten gesucht wird, dann kann man für diese Fragestellung ein entsprechendes Modell des Bildes aufstellen und die freien Parameter des Modells (also z.B. Helligkeit und Abstand) basierend auf den real aufgenommenen Bildern schätzen. In diesem Fall kann eine Auflösung deutlich unter der nach Rayleigh definierten erzielt werden. Natürlich muss bei entsprechenden „modellbasierten Rekonstruktionen" das Modell ausreichend gut die Realität beschreiben. O'Connor et al zeigen ein Beispiel für die Defekterkennung in periodischen Strukturen [38]. Auch hier gilt, dass entsprechende Ansätze in der industriellen Bildverarbeitung selten zum Einsatz kommen.[5]

5.5 Bildfehler

Während es vergleichsweise einfach ist, einen einzelnen Punkt oder ein sehr kleines Objektfeld beugungsbegrenzt abzubilden[6], ist die in der Bildverarbeitung übliche Abbildung ausgedehnter Objektfelder bei geringen Blendenzahlen deutlich schwieriger zu realisieren. Sowohl aufgrund des Designs als auch aufgrund von Fertigungstoleranzen ergeben sich für reale Objektive damit *Bildfehler* (*Aberrationen*). Typischerweise variieren diese Bildfehler — und damit dann auch die Auflösung — über das Feld.

Idealerweise sollten alle Strahlen, die von einem einzelnen Objektpunkt ausgehen sich im zugeordneten (konjugierten) Bildpunkt schneiden. Wäre das Brechungsgesetz, das die Änderung der Strahlrichtung an einer Grenzfläche beschreibt,

[5] Eine Ausnahme bildet die Halbleiterindustrie mit den dort vorherrschenden extrem kleinen Strukturen.
[6] das schafft jeder preiswerte CD-Player

$n \sin\phi = n' \sin\phi'$, linear, also $n\phi = n'\phi$, dann wäre diese ideale Abbildung auch ohne Probleme für ausgedehnte Felder möglich. Die Nichtlinearität des Brechungsgesetzes (Sinus) führt aber dazu, dass sich nur Strahlen unter kleinen Winkeln ϕ bzw. ϕ' in einem Bildpunkt schneiden. Je größer die Winkel, desto stärker sind die Abweichungen vom idealen Verhalten. Damit ist bereits klar, dass im allgemeinen Bildfehler mit zunehmender Apertur (bzw. abnehmender Blendenzahl) stärker werden.

Es gibt nun verschiedene Möglichkeiten der mathematischen Beschreibung dieser Zusammenhänge, die wir hier aber nicht näher betrachten wollen. Im Sinne einer elementaren Vorstellung genügt es für uns, typische Grundeinteilungen und Zusammenhänge zu verstehen (bzw. zur Kenntnis zu nehmen). Eine etwas ausführlichere Darstellung findet sich in [51].

Für ein konkretes Objektiv hängt die Aberration jedes Strahls im wesentlichen davon ab, von welcher Feldposition y er startet und durch welche Position (r, ϕ) der Strahl in der Eintrittspupille er läuft. Man kann die Abweichungen der Strahldurchstoßpunkte in der Bildebene von ihrer eigentlichen Zielposition durch eine Polynomentwicklung beschreiben und die wesentlichen Terme dieser Polynomentwicklung sind Defokus, sphärische Aberration, Koma, Astigmatismus, Bildfeldwölbung und Verzeichnung. Zusätzlich zu diesen „primären (monochromatischen) Aberrationen" ergeben sich höhere Ordnungen, Aberrationen für verschiedene Wellenlängen (chromatische Aberrationen) sowie Kombinationen der monochromatischen mit den chromatischen Aberrationen. In der Feld- und Pupillenabhängigkeit der Stärke der Aberrationen unterscheiden sich die unterschiedlichen Aberrationen.

Die *sphärische Aberration* ist (neben dem Defokus, der linear mit der Apertur r ansteigt) die einzige Aberration, die auch für einen Objektpunkt auf der optischen Achse, also $y = 0$ auftritt. Dafür ist die Abhängigkeit mit der Radialkoordinate in der Pupille r sehr groß ($\Delta y' \sim r^3$), d.h. durch Abblenden (r wird verkleinert) kann die sphärische Aberration extrem reduziert werden bzw. wenn große numerische Aperturen verwirklicht werden sollen ist der sphärischen Aberration besondere Aufmerksamkeit zu widmen.

Die *Verzeichnung* hängt dagegen überhaupt nicht von der Pupillenposition ab und verschwindet daher auch nicht durch Abblenden. Sie steigt dafür mit der Feldko-

ordinate ($\Delta y' \sim y^3$). Üblicherweise kann man (bei rotationssymmetrischen Systemen) zwischen kissenförmiger und tonnenförmiger Verzeichnung unterscheiden. Zwei Definitionen für die Verzeichnung sind üblich. Die gewöhnliche Verzeichnung ist das Verhältnis von realer zur paraxialen Bildkoordinate (minus 1). „TV distortion" ist der Mittelwert zwischen der y-Koordinate des rechten und linken Bildrandes im Verhältnis zur y-Koordinate im Zentralbereich. (minus 1).

Komplett symmetrische Abbildungssysteme sind frei von Verzeichnung[7]. Beachten Sie, dass diese Verzeichnung aufgrund der Optik nicht mit der perspektivischen Verzeichnung oder einer rein geometrischen Verzeichnung (Objekt verkippt) verwechselt werden darf.

Die *Koma* hängt stark von der Position der Blende des optischen Systems ab und variiert quadratisch mit der Apertur und linear mit der Feldposition ($\Delta y' \sim y\, r^2$). Sie hat ihren Namen aufgrund des kometenhaften Punktbilds.

Astigmatismus und *Bildfeldwölbung* sind miteinander verwandt. Generell steigen sie mit $\Delta y' \sim r\, y^2$ an. Astigmatismus führt dazu, dass Strukturen je nach Orientierung in verschiedenen Ebenen scharf abgebildet werden. Die Bildfeldwölbung bzw. Feldkrümmung ändert ebenfalls die Fokussierung, allerdings in Abhängigkeit der Feldposition. D.h. weit außen liegende Feldpunkte werden in einer anderen Ebene fokussiert als Punkte nahe der optischen Achse, näherungsweise liegen die scharfen Bildpositionen auf einer Kugelschale. Ohne eine spezielle Korrektur würden die Bildpunkte auf einer Kugelschale liegen. Besonders relevant ist sie bei Weitwinkelobjektiven.

Bei den chromatischen Fehlern ist zunächst zwischen der *chromatischen Längsaberration* und der *chromatischen Vergrößerungsdifferenz* zu unterscheiden. Die chromatische Längsaberration bedeutet schlicht, dass der Bildpunkt in Abhängigkeit von der Wellenlänge in unterschiedlichen Ebenen liegt. Dies ist nicht weiter verwunderlich, denn die Brennweite einer Linse hängt natürlich vom Brechungsindex der Linse ab und der Brechungsindex variiert wieder mit der Wellenlänge.

Die chromatische Vergrößerungsdifferenz hat ihren Ursprung ebenfalls in der Bre-

[7]Dabei ist aber zu beachten, dass auch Objekt- und Bildweite symmetrisch sein müssen, d.h. die perfekte Verzeichnungsfreiheit ist nur für 1:1 Abbildungssysteme automatisch gegeben. Dennoch haben symmetrisch aufgebaute Systeme auch für 1:x Systeme verbesserte Verzeichnungseigenschaften.

chungsindexvariation. Es ergibt sich (s. Abb. 5.5.5), dass der Abbildungsmaßstab für den Feldpunkt von der Farbe abhängt. Dies führt bei *Farbbildaufnehmern* zu *Farbsäumen* vor allem am Bildrand und bei monochromatischen Sensoren zu entsprechenden Unschärfen. Die einfachste Möglichkeit zur Eliminierung der Farbfehler in der industriellen BV ist natürlich, eine mehr oder weniger monochromatische Beleuchtung zu verwenden.

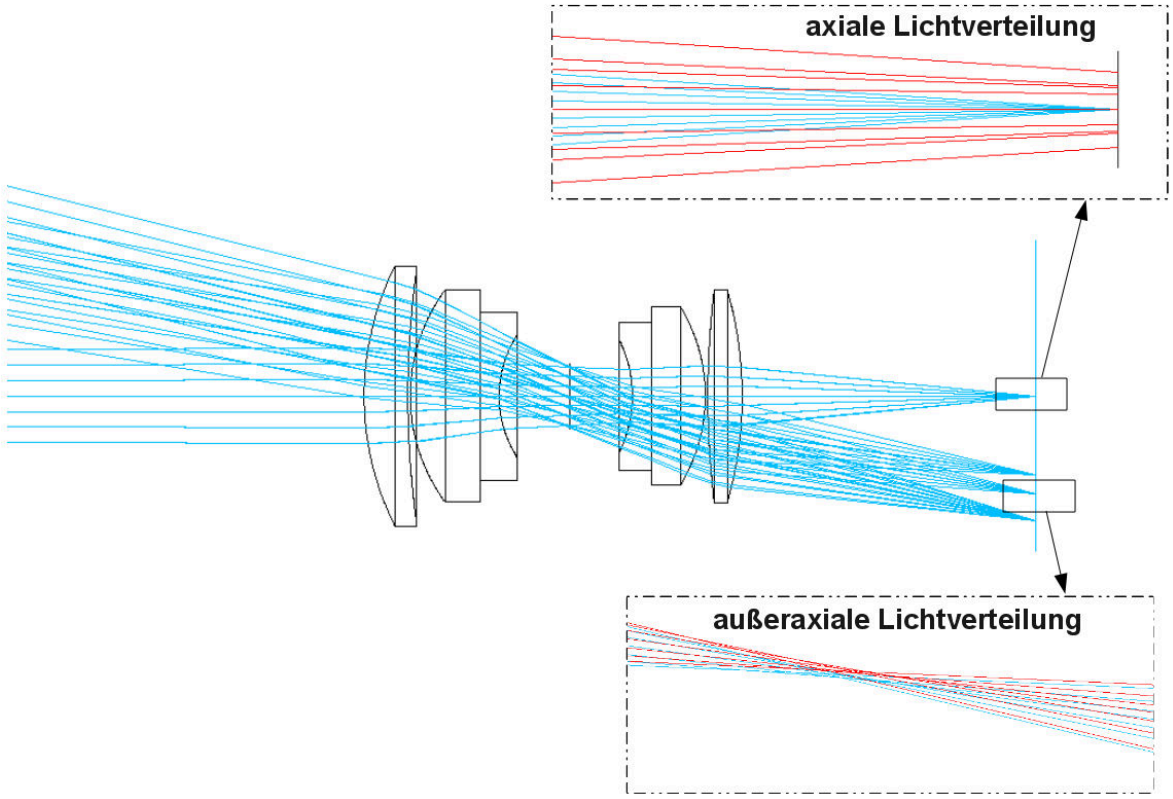


Abbildung 5.5.5: Chromatische Aberrationen dargestellt für zwei verschiedene Wellenlängen. Im Allgemeinen ergibt sich sowohl eine seitliche, als auch eine axiale Verschiebung des Bildpunkts in Abhängigkeit der Farbe. Die seitliche Verschiebung führt zu einer Variation des Abbildungsmaßstabs mit der Farbe. Die axiale Verschiebung ist der Farbfehler, der sich selbst für die Position x=y=0, also die optische Achse ergibt.

Generell kann gesagt werden, dass die Schwierigkeit sowohl im Design als auch bei der Herstellung von Optiken mit zunehmender Lichtstärke und Anzahl abzubildender Punkte steigt. Die Korrektur eines optischen Systems ist also umso problematischer, je mehr Punkte in hoher Qualität und guter Lichteffizienz gleichzeitig abgebildet werden müssen. Daher ist eben die beugungsbegrenzte Abbildung eines Punktes, wie sie zum Beispiel bei einem DVD Abspielgerät notwendig ist, sehr einfach realisierbar, während die ebenfalls beugungsbegrenzte Abbildung von Milliarden von Punkten, wie wir sie in einem Lithografieobjektiv benötigen,

entsprechend schwierig und teuer zu erzielen ist.[8]

Die einfachste wirklich brauchbare Kenngröße, die angibt, wie komplex ein Abbildungssystem bei einer gegebenen Lichtstärke gewählt werden muss, ist genau diese Anzahl der abzubildenen „Pixel" bzw. damit gleichbedeutend: die zu übertragende Informationsmenge. Wir wollen dabei Pixel als kleine Kreise in der Bild– bzw. Objektebene betrachten. Vornehmer (und auch exakter) ausgedrückt ist die „Anzahl der Pixel" das *Ortsbandbreitenprodukt*. Ein großes Ortsbandbreitenprodukt bedeutet entweder, dass die Bildpunkte sehr klein werden müssen oder aber dass das Bildfeld eine große Ausdehnung aufweist. Beides bedingt einen gewissen Aufwand. Informationstheoretisch kann man sagen: Soll die Information von vielen Objektpunkten ins Bild übertragen werden, dann ist hierzu eine aufwändigere Optik notwendig als bei wenigen Punkten.[9]

Das Ortsbandbreitenprodukt ist also in gewissem Sinn ein einfaches Maß für die notwendige „Komplexität" des optischen Systems, das dieses Ortsbandbreitenprodukt übertragen soll. Für die meisten Bildverarbeitungsanwendungen ist es allerdings nicht geeignet, denn wir erwarten dort üblicherweise, dass in allen Bildbereichen eine ausreichende Auflösung erzielt wird. D.h. maßgeblich ist meist (wenn auch nicht immer) die erzielte Auflösung im am schlechtesten abgebildeten Feldbereich, meist am Bildrand.

Die Auflösung nach Rayleigh kann sinnvoll nur verwendet werden, wenn die Bildfehler, die aufgrund des optischen Designs und der Toleranzen bei der Fertigung die Abbildungsqualität verringern, klein gegenüber der Beugungsbegrenzung sind. Die Aberrationen führen dann bei einer Abbildung eines idealen Punktes und unter Vernachlässigung der Beugung zu einem Bildpunkt dessen Radius (deutlich) kleiner als die nach Gleichung 5.2 berechnete Punktbildausdehnung ist.[10]. Für relativ weit ausgedehnte Objektfelder und hohe Lichtstärken ist das mit preisgünstiger Optik in der Regel nicht zu erreichen.

Andererseits werden die meisten hochwertigen Objektive beugungsbegrenzt,

[8] Natürlich können auch andere Parameter, z.B. ein ausgedehnter Wellenlängenbereich oder hohe Anforderungen an die Verzeichnungsfreiheit, zu erheblichen Schwierigkeiten und damit Kosten führen.

[9] Teilweise wird auch die Area-weighted average resolution verwendet, um das Gesamtauflösungsverhalten eines Objektivs zu kennzeichnen.

[10] Für den sogenannten Wellenfrontfehler (gemessen z.B. mit einem Interferometer) der Optik ergibt sich dann ein Peak-to-Valley Wellenfrontfehler kleiner ein Viertel der Wellenlänge ist [20]

wenn man genügend abblendet ($K > 8$). Für große Blendenzahlen K, also eine geringe Öffnung des Systems bzw. achsnahe Strahlen durch das System erzielen viele Systeme somit eine beugungsbegrenzte Abbildungsqualität (Gl. 5.2). Für die meisten Objektive gilt weiterhin, dass die beste Abbildungsqualität nicht bei der geringsten Blende erreicht wird. Ein besseres Bild erhält man meist, wenn die Blende um 2 bis 3 Stufen zugezogen wird. Die beugungsbegrenzte Auflösung (und natürlich die Lichtstärke) wird dabei zwar reduziert, aber die geometrischen Aberrationen werden stärker reduziert.

Wenn die Aberrationen ausschlaggebend werden, dann können wir als Beurteilungskriterium für die Bildqualität den Radius des Bildpunkts bei der Abbildung eines idealen Punkts verwenden. Üblicherweise wird der Radius in diesem Fall im Sinne eines RMS (root-mean-squared) Radius verwendet (siehe z.B. Abb. 5.5.6).

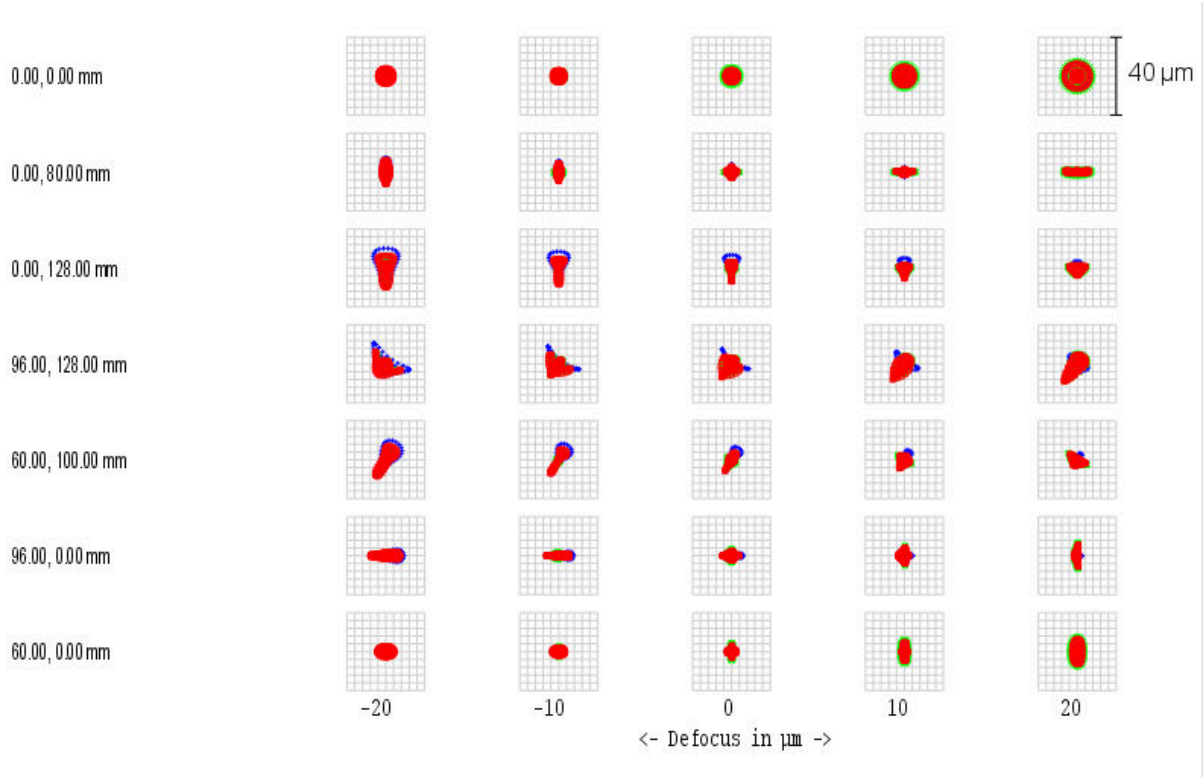


Abbildung 5.5.6: Beispielhafte Variation des Spotbilds bei der Abbildung idealer Punkte in Abhängigkeit der Feldposition (Zeilen) und der Fokussierung (Spalten).

5.6 Vignettierung

Abb. 2.3.12 („relative illumination") zeigt ein Phänomen, das als *künstliche Vignettierung* bezeichnet wird. Strahlen, die von einem außeraxialen Objektpunkt kommen werden teilweise von Blenden absorbiert und erreichen so nicht die Bildebene. Als Resultat ergibt sich für diese Punkte eine verminderte *Bestrahlungsstärke* in der Bildebene.[11]

Dieser Effekt ist teilweise gewollt und wird vom Designer des Objektivs berücksichtigt. Generell ist es — wie bereits erwähnt — schwierig, bei vertretbaren Kosten Objektive mit hoher Lichtstärke und großem Feld mit hervorragender Abbildungsqualität zu realisieren. Je weiter die Blende des Systems geöffnet wird, desto mehr Strahlen führen zur Bildentstehung. Es ist dementsprechend schwierig, bei offener Blende alle Strahlen, die von einem Objektpunkt ausgehen auch wieder in einem Bildpunkt zu bündeln. Die künstliche Vignettierung hilft hier, denn es werden Strahlen ausgeblendet. Diese können somit nicht zu einer Unschärfe des Bildpunkts führen. Das Bild bleibt trotz geöffneter Blende akzeptabel.

Beachtet werden muss aber, dass das verringerte Strahlenbündel dann für den entsprechenden Feldpunkt auch eine verringerte numerische Apertur und damit eine verringerte Auflösung nach Rayleigh zur Folge hat.

Wird die Blende deutlich geschlossen dann ergibt sich keine künstliche Vignettierung weil das Strahlbündel sowieso von der Blende des Systems deutlich beschnitten ist. Die künstliche Vignettierung hilft so also auch bei weit geöffneter Blende noch eine gute bis akzeptable Bildschärfe im gesamten Feld zu erzielen. Ohne sie wäre oftmals nur das Zentrum des Bilds scharf.

Auch wenn ein Objektiv keine künstliche Vignettierung aufweist ergibt sich automatisch eine Vignettierung, die sogenannte *natürliche Vignettierung*. Die *Bestrahlungsstärke* in der Bildebene variiert mit dem Bildwinkel ω' hierbei (bei entozentrischen Abbildungen) gemäß

$$E(\omega') = E(0)\,\cos^4\omega' \tag{5.10}$$

[11] Zusätzlich wird die effektive numerische Apertur für diese Strahlen reduziert und damit die beugungsbedingte Unschärfe vergrößert. (Dieser Effekt ist allerdings in der Regel ohne praktische Bedeutung da die aberrationsbedingte Bildunschärfe für außeraxiale Bildpunkte bei den meisten Objektiven und Anwendungen der ausschlaggebende Anteil für die Bildverschlechterung ist.

über das Feld. Der Bildwinkel ist für ein dünnes System gleich dem Feldwinkel im Objektraum. Für ein endlich ausgedehntes System kann durch die Blendenlage der Feldwinkel und damit die Vignettierung beeinflusst werden.

Die Ursache für dieses Verhalten liegt in der Verkippung des Hauptstrahls gegenüber der Objekt- und Zielebene um den Winkel ϕ sowie der Abnahme der Intensität aufgrund der Energieerhaltung mit dem Abstand zum Quadrat. Eine detaillierte Behandlung findet sich z.B. in [31]. Der Effekt ist natürlich vergleichsweise stark bei Weitwinkelobjektiven aufgrund der sich dort ergebenden starken Winkel. Die künstliche Vignettierung ist vor allem bei stark geöffneten Objektiven (kleine Blendenzahl) von Bedeutung.

Die Korrektur der sich aufgrund der Vignettierung ergebenden inhomogenen *Bestrahlungsstärke* erfolgt üblicherweise per Software (Normierung durch ein „leeres" Grundbild). Dies ist ohne Probleme möglich, wenn genügend Dynamik und ein genügend hohes *Signal-Rausch-Verhältnis* vorliegt. Es ist auch möglich, durch geeignete optische Filter vor der Bildaufnahme die die Variation der Bestrahlungsstärke auf dem Sensor auszugleichen, hierbei wird dann aber natürlich ein Teil der sonst nutzbaren Lichtmenge eliminiert.

5.7 Pixelierung: Q

Neben der Festlegung der optischen Auflösung ist natürlich auch die Wahl der Pixelgröße des Bildsensors entscheidend. Das Verhältnis von Auflösungbeschränkung durch die Optik und durch den Bildaufnehmer wird durch das sogenannte Q spezifiziert:

$$Q = 2\,\frac{r_{Optik}}{r_{Bildsensor}} \tag{5.11}$$

Die Auflösung des Bildsensors wird dabei im Sinne einer Ortsfrequenz definiert, so dass $r_{Bildsensor}$ als das Doppelte des Pixelabstands p angenommen werden kann.

In [12] wird dabei für r_{optik} der Faktor 1.00 statt 1.22 wie in unserer Definition mittels Rayleigh verwendet.[12] Die Auflösung der Optik kann dabei im einfachsten Fall über die Beugungsbedingung (z.B. nach Auflösung nach Rayleigh), den

[12] Die Unterschiede in der Definition sind für die Praxis wenig relevant, es zeigt sich aber, dass wie meist in der Bild-

RMS Radius des Punktbilds oder eben mittels der MTF (Grenzfrequenz bei einem bestimmten Kontrast, z.B. 30%) definiert werden. In diesem Fall ist Q dann das Verhältnis von optischer Auflösung (z.B. Rayleigh) und Pixelabstand und es ergibt sich Q=1 wenn ein Pixel der Auflösung nach Rayleigh entspricht. Für ein beugungsbegrenztes System ist Q somit das Verhältnis des Radius des Punktbildes nach Airy und der Pixelgröße p:

$$Q = \frac{r_A}{p} \tag{5.12}$$

Klar ist, dass es meist vergleichsweise wenig Sinn macht, wenn die Auflösungen extrem differieren. Das Punktbild einer schlechten Optik mit 100 Pixeln abzutasten ist meist keine gute Idee, genauso wenig ist es sinnvoll, mittels einer aufwändigen Optik ein Punktbild zu generieren, das deutlich kleiner als ein einzelner Pixel ist. Ein ausgewogenes Design ergibt sich, wenn beide Auflösungen gleich sind, wenn also Q = 2 ist. In diesem Fall sampelt man den Radius des Airys mit 2 Pixeln (siehe Abb. 5.7.7. Hochfrequente Muster können mit vergleichsweise gutem Kontrast aufgelöst werden. Aliasing wird aber vermieden.

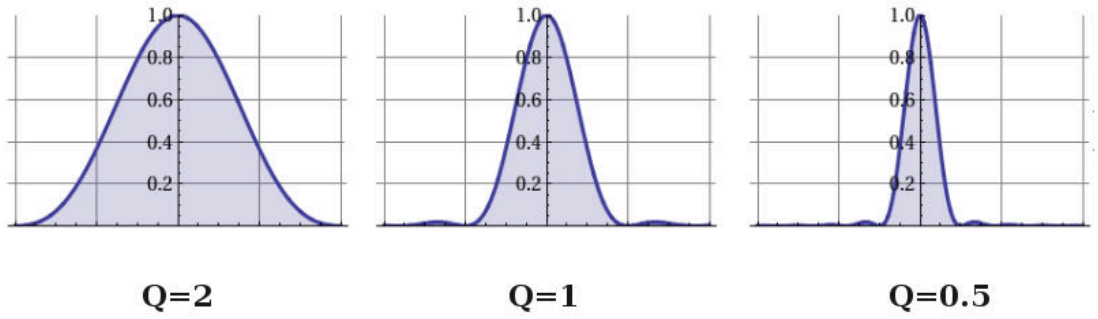


Abbildung 5.7.7: Zur Definition von Q. Q=2 bedeutet, dass der Durchmesser des Airy-Punktbilds (Nullstellen) mit 4 Pixeln abgetastet wird (bzw. der Radius mit 2 Pixeln).

Ein Q kleiner 2 bedeutet, dass die Optik ein sehr gutes Bildsignal liefert, dieses aber gemäß einer nach dem *Samplingtheorem* nicht ausreichenden Anzahl von Pixeln abgetastet wird. Es können sich in diesem Fall Moires ergeben (Aliasing).

verarbeitung die exakte Definition der Kenngrößen überprüft werden sollte wenn verschiedene Werte miteinander verglichen werden.

Trotzdem wird dies bei vielen Designs (Q = 1 ... 1.5) in Kauf genommen. Der Gewinn an Kontrast bei feinen Details überwiegt bei einer gegebenen Pixelanzahl oft die Möglichkeit, dass sich Moires ausbilden. Die konkrete Wahl des Q hängt von den konkreten Anforderungen der Aufgabenstellung ab. Beim Vergleich verschiedener Werte ist aber wieder auf die exakte Definition von Q zu achten.

In den letzten Jahren hat sich ein Trend zu vielen kleinen Pixeln (Pixelgröße: 1... 1.5 μm) bei Bildaufnehmern durchgesetzt, so dass zwangsläufig mit großen Q gearbeitet wird. Man muss allerdings klar zwischen verschiedenen Anwendungen und den sich daraus ergebenden Anforderungen differenzieren.

Es ergeben sich bei der Abbildung auch erhebliche Unterschiede der Abbildungsleistung, je nachdem, wie das abzubildende Objekt konkret aussieht. Bei periodischen Objekten ist die Definition über die MTF am sinnvollsten, denn diese gibt direkt den Kontrast bei der Abbildung der periodischen Struktur an. Leichte Unterschiede ergeben sich dabei für periodische Binärstrukturen gegenüber den in der üblichen MTF Definition verwendeten sinusförmigen Streifen.[13] Aperiodische isolierte Strukturen werden letztlich mit einer verbesserten Qualität abgebildet, denn im Gegensatz zu den periodischen Sinusstreifen ist die MTF nicht nur an einer Stelle, der Grenzfrequenz, sondern über einen weiten Bereich ausschlaggebend. Da für tiefere Raumfrequenzen die MTF meist höher ist, ergibt sich dann ein Gewinn an „Auflösung" bzw. Kontrast. Abb. 5.7.8 zeigt deutlich, dass Strukturen sehr gut sichtbar werden, obwohl der Strukturabstand im Bereich der Auflösung nach Rayleigh liegt. In diesem Fall ist es wesentlich, dass das sich ergebende Bild vergleichsweise hoch aufgelöst abgetastet wird (großes Q).

Abb. 5.7.8 zeigt beispielhaft eine entsprechende Simulation für ein Bildverarbeitungsobjektiv (Doppel-Gauß) mit 500 mm Arbeitsabstand und einer Feldgröße von 500 mm. Das Objektiv wurde optimiert für polychromatische Anwendungen (RGB), aber korrekt für die hier simulierte Beleuchtung mit quasimonochromatischem Licht ($\lambda = 600$ nm) fokussiert. Die Blendenzahl beträgt 3.

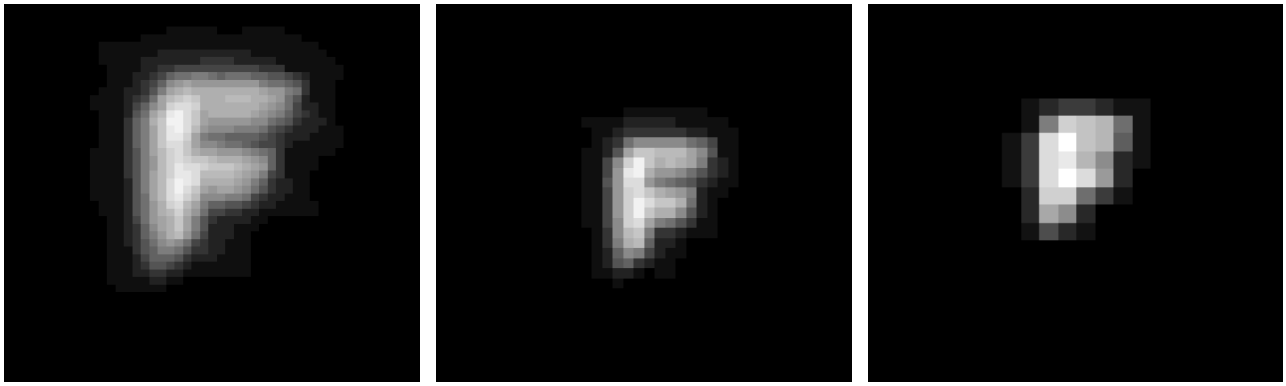
Nach Airy ergibt sich eine Auflösung von ca. r_A = 2.2 μm . Die Aberrationen führen aber on-axis zu einem RMS Bildpunktradius von 4 μm . Daher sollte für weitere Abschätzungen dieser Wert verwendet werden. Das Objekt soll einen guten Kontrast aufweisen daher würde man gemäß Abb. 5.10.14 von einer minimalen kriti-

[13] Der Kontrast für binäre Strukturen ist gegenüber Sinusstreifen leicht angehoben.

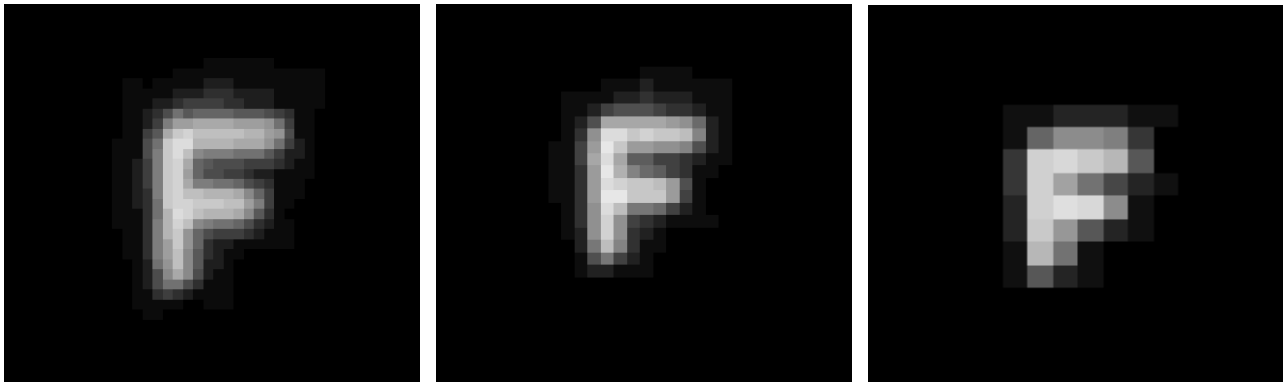
schen Dimension von $1...1.5 \cdot 4 = 4..6\ \mu\mathrm{m}$ (bildseitig) ausgehen. Dabei definieren wir für die Simulationen die kritische Dimension als den Dunkelabstand zwischen den beiden Querbalken des „F".

Entsprechende Simulationsergebnisse für verschiedene Pixelgrößen sind in der oberen Reihe dargestellt. Für die Wahl der Anzahl an Pixeln pro kritischer Dimension sind offensichtlich etwa 2 bis 2.5 Pixel ausreichend.

In der unteren Reihe sind Simulationsergebnisse für eine gröbere Struktur dargestellt (CD= 6 $\mu\mathrm{m}$ statt 4 $\mu\mathrm{m}$).



(a) CD=4 $\mu\mathrm{m}$, Pixelsize=0.5$\mu\mathrm{m}$ → Q=8, r_{RMS} =CD, 8 Pixel pro CD (b) CD=4 $\mu\mathrm{m}$, Pixelsize=1.5$\mu\mathrm{m}$ → Q=2.6, r_{RMS} =CD (c) CD=4 $\mu\mathrm{m}$, Pixelsize=2.0$\mu\mathrm{m}$ → Q=2, r_{RMS} =CD, 2 Pixel pro CD



(d) CD=6 $\mu\mathrm{m}$, Pixelsize=1.0$\mu\mathrm{m}$ → Q=6, r_{RMS} =CD/1.5, 6 Pixel pro CD (e) CD=6 $\mu\mathrm{m}$, Pixelsize=1.5$\mu\mathrm{m}$ → Q=4, r_{RMS} =CD/1.5, 4 Pixel pro CD (f) CD=6 $\mu\mathrm{m}$, Pixelsize=3$\mu\mathrm{m}$ → Q=2, r_{RMS} =CD/1.5, 2 Pixel pro CD

Abbildung 5.7.8: *Simulation eines Doppel-Gauß Objektivs mit Arbeitsabstand 500 mm und einem Objektfeld von ca. 250 mm bei monochromatischer Beleuchtung mit einem RMS Bildpunktradius von r_{RMS}=4 $\mu\mathrm{m}$. Die kritische Dimension wird als Abstand zwischen den beiden waagrechten F-Strichen (Dunkelbereich !) angenommen.*

Eine ideale Wahl lässt sich nur erzielen, wenn das sich ergebende Bild (inklusive Beugungseffekte und Aberrationen) für verschiedene Pixelgrößen simuliert (oder gemessen) wird. Abb. 5.7.8 zeigt sehr deutlich, dass es manchmal durchaus Sinn

macht, die Pixelauflösung deutlich besser als die optische Auflösung zu wählen, also ein größeres Q zu verwenden. In diesem Fall ergibt sich eine klar verbesserte Darstellung feiner isolierter Details. Dies ist z.B. in der Mikroskopie sehr gebräuchlich, macht aber auch für viele Bildverarbeitungsaufgaben, bei denen sehr feine Details erfasst werden sollen Sinn. In diesem Fall sollte ein Q von 2 bis 3 (eventuell sogar 4) gewählt werden. Eine Abtastung mit einer Frequenz oberhalb der Nyquistfrequenz macht also durchaus Sinn wenn feinste Strukturen abgebildet werden sollen. Im bereits dargestellten MTF-Bild macht das Sinn, denn somit verringert dann die Pixel-MTF die Gesamt-MTF nicht weiter.

Objektseitig entspricht dies im übrigen bei dem System einer kritischen Dimension von 173 $\mu\mathrm{m}$. Über das gesamte Feld (250 mm) würde sich so eine notwendige Pixelzahl von ca. 33 MPixel ergeben[14].

Es ist allerdings zu beachten, dass die Ergebnisse monochromatisch berechnet wurden (LED) und vor allem keine Fertigungsfehler in der Simulation enthalten sind. Die in Abb. 5.7.8 angenommene Auflösung von 4 $\mu\mathrm{m}$ bildseitig ist also in der Praxis selten so erreichbar.

Die Ergebnisse zeigen aber dennoch klar, dass eine deutliche Überabtastung mit vielen Pixeln Sinn machen kann und dass man mit den Parametern optische Auflösung und Abtastung (Pixelauflösung) spielen kann, um gegebene Strukturen noch sinnvoll zu erfassen.

5.8 Advanced: Leere Vergrößerung

Gemäß dem Samplingtheorem lohnt sich ein großes Q nur sehr bedingt (über die leicht verbesserte Pixel-MTF). Bilder mit hohem Q sehen zunächst aber deutlich besser oder auch hochaufgelöster aus.

Dies ist aber bis zu einem gewissen Grad ein Trugschluss. Wenn ein mit $Q \geq 3$ gesampeltes, einigermaßen rauscharmes Bild auf höhere Auflösungen upgesampelt wird, dann ergeben sich oft Bilder, die kaum von Bildern mit größerem Q unterschieden werden können.

[14] $(250\,\mathrm{mm}\ /\ 173\ \mu\mathrm{m} \cdot 4)^2$

Für die Interpolation sollen dabei kubische Splines oder besser noch eine Lanczos Interpolation verwendet werden.

5.9 Modulationstransferfunktion

Obwohl der RMS-Radius eines Bildpunkts für die Beurteilung sehr wertvoll ist (s.u.) wird er leider für die meisten Optiken nicht angegeben. Stattdessen werden fast alle detaillierten Angaben zur Auflösung optischer Systeme anhand der *Modulationstransferfunktion* (MTF), die wir im vorangegangenen Kapitel näher untersucht haben, vorgenommen. Dabei zeigt sich, dass die MTF sehr deutlich auf Aberrationen reagiert, so dass eine Betrachtung der MTF sehr gut zeigt, ob ein System näherungsweise beugungsbegrenzt ist oder nicht.

Natürlich hängt der Kontrast bzw. die MTF von der Einstellung bzw. Verwendung der Optik ab, denn die Aberrationen variieren mit diesen Parametern. Für verschiedene Arbeitsabstände oder Blendeneinstellungen wird man also unterschiedliche MTF-Kurven erhalten.

Da die Aberration auch mit der Position im Feld variiert bekommt man entsprechend auch unterschiedliche Kurven für die unterschiedlichen Feldposition. Aus diesem Grund ist eine weitere Darstellung der MTF sinnvoll (Abb. 5.9.9). Hierbei wird über das Feld aufgetragen, welcher Kontrast bei einer ganz bestimmten Raumfrequenz (z.B. 50 lp/mm) erzielt wird.

Als MTF-basierte Auflösung wird oft die Grenzfrequenz verwendet, bei der der MTF Kontrast (über das relevante Feld) noch 10% oder 30% beträgt, also die kleinste Gitterperiode, die noch mit 10%, 30% oder 50% Kontrast abgebildet wird (MTF10, MTF30, *MTF50*).

Die MTF ist sicher das Mittel der Wahl zur Beurteilung, wenn periodische Strukturen abgebildet werden sollen. Dies ist aber in der industriellen Bildverarbeitung eher selten der Fall.

Es sei aber hier davor gewarnt, das Punktbild einfach mit der MTF basierten Auflösung gleichzusetzen. Bei der Abbildung aperiodischer Objekte werden gleichzeitig viele verschiedene Raumfrequenzen durch die Optik übertragen

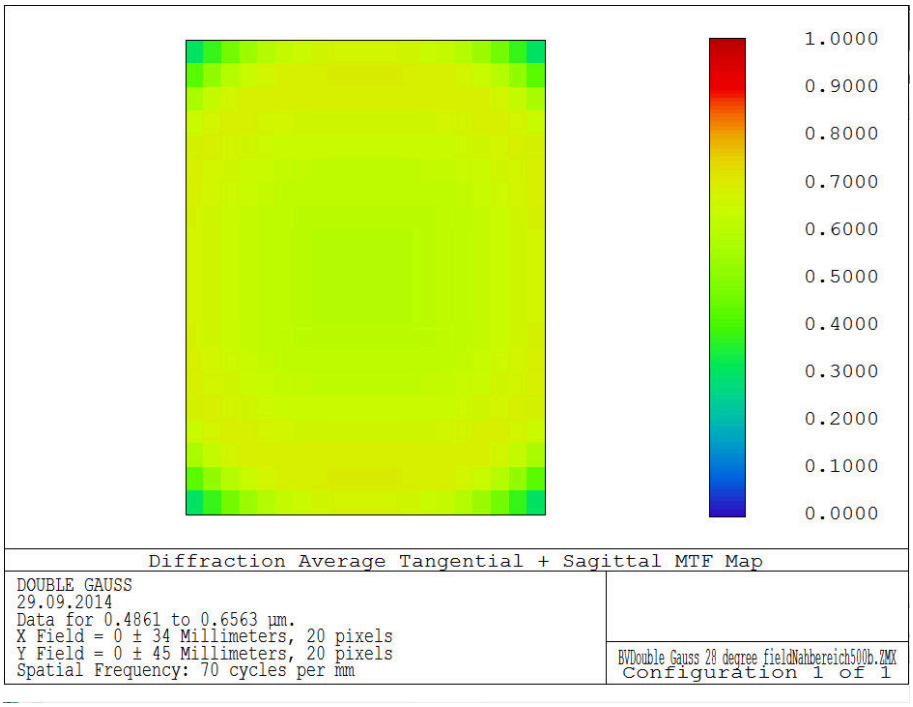


Abbildung 5.9.9: Variation der MTF über das Feld. Am Rand ist die MTF geringer als in zentralen Bereichen.

und dementsprechend kann ein einzelner MTF-Wert eben nicht die Qualität des sich ergebenden Bilds sinnvoll beschreiben. So kommt es, dass teilweise erhebliche Diskrepanzen bei MTF-basierter Bildbeurteilung und anderen Auflösungsbeurteilungen zustande kommen (siehe z.B. [43]). Bei einer (oft notwendigen) heuristischen Umrechnung von MTF-basierten in Spotradius-basierten Maßzahlen macht man also immer einen (anwendungsabhängigen) mehr oder weniger starken Fehler.

In einfachster Näherung wird für die MTF ein lineares Verhalten gemäß

$$MTF(f) = 1 - f \tag{5.13}$$

mit der normalisierten Frequenz $f = f_u/f_{max}$ als Modell verwendet. Bei beugungsbegrenzter Auflösung gemäß Rayleigh kann die maximale bzw. Grenzfrequenz gemäß

$$f_{max} \approx \frac{1.22}{r_{spot}} = \frac{1}{\lambda\, K} \tag{5.14}$$

berechnet werden.

Beispiel: Wir benötigen für eine Anwendung einen RMS-Durchmesser der Bildpunkte (bei der Abbildung idealer Objektpunkt) von 11 μm . Zur Verfügung steht ein Objektiv das laut Spezifikation einen Kontrast von 30% bei 60 lp/mm liefert. Kann dieses Objektiv eingesetzt werden? **Lösung:** $f_{max} = 1.22/11$ microns = 110.000 lp/m = 110 lp/mm. Mit Gl. 5.13 ergibt sich hier $MTF(f) = 0.3 \approx 1 - f$ und für den Kontrast von 30% also f=0.7 bzw. f_{max} = 60/0.7 = 85 lp/mm.

Andererseits fordern wir aber gemäß der Aufgabe $f_{max} = 1.22/11 = 110$ lp/mm. Daher wird das Objektiv die Anforderung nicht erfüllen können.

Natürlich ist insgesamt ein möglichst hohes Kontrastniveau zu wünschen. D.h. eine Linse die über alle Ortsfrequenzen einen Kontrast von lediglich 30% liefert ist sicher nicht erstrebenswert. Stattdessen soll bei möglichst vielen Frequenzen ein möglichst hoher Kontrast erzielt werden. Quantifiziert werden kann diese Aussage durch die Fläche unterhalb der MTF Kurve wenn diese im Sinne einer Kontrastangabe über die Raumfrequenz dargestellt wird. Ein möglichst großer Flächeninhalt ist in diesem Fall wünschenswert. Verschiedene Kenngrößen können basierend auf dieser Idee abgeleitet werden, siehe z.B. [33, 37], der „subjective quality factor"[15], die „Heynacher Kennzahlen" und das „Squareroot Integral" (SQRI). Beim Vergleich unterschiedlich großer Bildsensoren bzw. Pixelgröße sollte eine entsprechende Skalierung der MTFs vorgenommen werden [33]; auch eine Angabe der Frequenzen in „lp pro Bildhöhe" anstatt lp/mm ist sinnvoll.

Aber auch die Fläche ist nicht alles entscheidend. Teilweise wird auch ein möglichst konstanter Verlauf der MTF bevorzugt und ganz allgemein muss gesagt werden, dass der subjektive Schärfeeindruck eines Bildes nicht zufriedenstellend in eine einzige Zahl gepresst werden kann.[37]

[15] dabei muss die Ortsfrequenzachse logarithmisch skaliert sein

Der geforderte Kontrast bei der Abbildung von periodischen Strukturen variiert üblicherweise mit der Raumfrequenz der Struktur. So erfordert das menschliche Sehsystem z.B. bei hohen Raumfrequenzen einen deutlich höheren Kontrast als bei tiefen Raumfrequenzen [20]. Ganz entsprechend muss auch in der maschinellen Bildverarbeitung bedacht werden, dass feine Strukturen schwieriger zu erkennen sind (weniger Pixel) als grobe Strukturen und dementsprechend der Kontrast vergleichsweise hoch sein muss. Die konkrete Variation des geforderten Kontrasts mit der Raumfrequenz ist natürlich algorithmen- und rauschabhängig.

Darüber hinaus ist auch die geforderte *Bestrahlungsstärke* von der Raumfrequenz abhängig denn auch hier gilt, dass bei wenigen Pixeln Rauschen schlechter toleriert werden kann als bei vielen Pixeln. Daher können eine starke Vignettierung oder dunkle Bildbereiche dazu führen, dass die Abbildungsqualität letztlich nicht mehr ausreichend ist.

5.10 Advanced: Vorgehen zur Festlegung von Bildsensorauflösung und Auflösung der Optik

Nach diesen etwas längeren Vorüberlegungen nähern wir uns dem zentralen Punkt dieses Kapitels, nämlich der konkreten Festlegung von optischer Auflösung und Bildsensorauflösung. Entscheidend ist aber zunächst, dass letztlich für die konkrete Anwendung zuerst definiert werden muss, welche Anforderungen an die Auflösung gestellt werden. Ohne eine genaue Definition dieser aufgabenspezifischen Anforderungen ist es unmöglich, gezielt ein passendes Objektiv auszuwählen (oder zu entwickeln).

Diese Definition der Anforderung ist in der Regel alles andere als einfach. Aber erst wenn sie erfolgt ist, kann in einem zweiten Schritt die optische Auflösung und die Bildsensorauflösung festgelegt werden. Folgende Fragen sollten zunächst beantwortet werden:

1. Geht es um die Auflösung im Sinne einer Trennung benachbarter Strukturen oder soll vielmehr eine geometrische Länge gemessen werden?
2. Welche Geometrie haben die kleinsten noch aufzulösenden Strukturen?

3. Handelt es sich um (quasi-)periodische (→ MTF-basierte Auflösungsbeschreibung) oder eher isolierte Strukturen (→ Rayleigh-basierte Beschreibung)?

4. Wie ist der Kontrast (bei einer perfekten mathematischen Abbildung) dieser Strukturen? Hierzu sind oft Vorexperimente notwendig da dieser Kontrast stark von der Aufgabenstellung (Objekt) aber auch der Beleuchtung abhängt. Durch ein hochauflösendes Referenzmesssystem (z.B. hochauflösendes Mikroskop) lassen sich oftmals sinnvolle Anhaltspunkte für den in der Anwendung vorliegenden Kontrast finden.

5. Welchen Kontrast kann die bildverarbeitende Anwendung noch sinnvoll akzeptieren?

6. Muss mit Moires (vgl. Kapitel 3) gerechnet werden?

7. Muss die Auflösung über das gesamte Feld erzielt werden?

8. Kann zu quasimonochromatischem Licht übergegangen werden (weniger Aberrationen)?

9. Ist die Blendenzahl aus anderen Überlegungen (Lichtstärke und Schärfentiefe) bereits festgelegt oder kann sie noch frei gewählt werden?

10. Ist der Arbeitsabstand festgelegt oder kann er noch frei gewählt werden?

11. Gibt es andere Anforderungen an die Optik (insbesondere Baulänge und Gewicht)?

Im folgenden gehen wir zunächst von Detektionsaufgaben aus. Die Messung geometrischer Strukturen wird im Abschnitt 5.17 dargestellt.

Idealerweise ist zwischen periodischen und isolierten Strukturen zu unterscheiden. An sich wäre für die meist auftretenden isolierten Strukturen eine Definition anhand der real vorliegenden Ausdehnung des Punktbilds (z.B. des realen RMS-Wertes des Spotradius) der geeignete Ausgangspunkt. Leider wird dieser Wert in aller Regel von den Herstellern der Optik nicht angegeben und ist meist nur bei eigenen Optikdesigns verfügbar. Wenn allerdings die Blende klein genug gewählt werden kann (Lichtstärke und Beugungsbegrenzung), dann ist das System oft

annähernd beugungsbegrenzt und die vorliegende Spotausdehnung kann einfach berechnet werden kann (Gl. 5.2).

Abb. 5.10.10 zeigt die sich ergebende Intensitätsverteilung für zwei benachbarte Punkte in verschiedenen Abständen für ein beugungsbegrenztes optisches System.

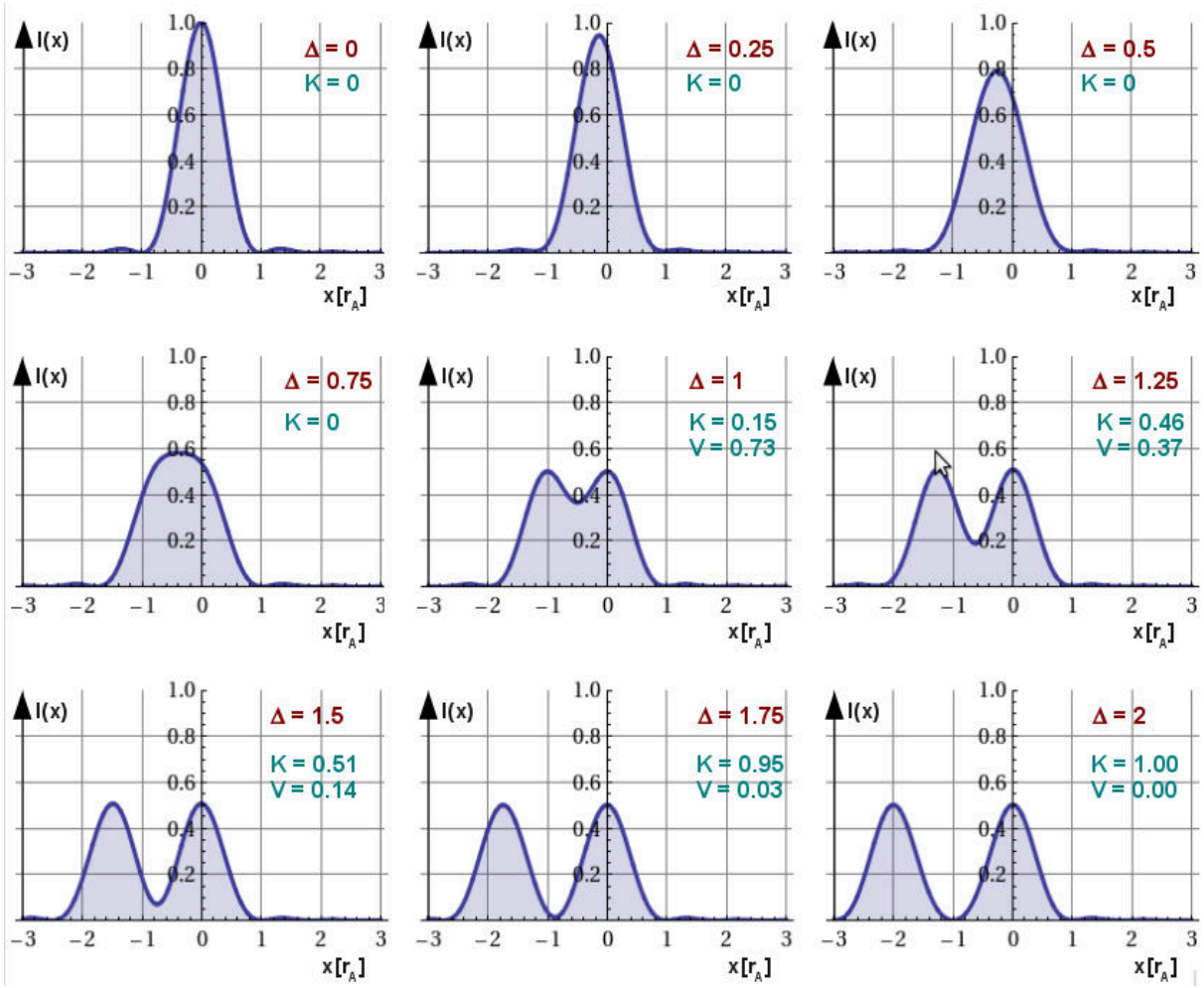


Abbildung 5.10.10: Intensitätsverteilung bei der inkohärenten beugungsbegrenzten Abbildung von zwei benachbarten Punkten. Der Abstand der beiden Punkte ist rot gekennzeichnet und in Vielfachen des Airyradius angegeben. Angegeben sind jeweils Kontrast K und das Verhältnis des Zwischenminimums zum Maximum.

Für Objektstrukturen mit einem sehr guten Kontrast (z.B. Loch in Chrommaske) ist in der Regel der sich für $\Delta = 1 \cdot r_A$ ergebende Kontrast im Bild ausreichend. D.h. die Optik kann so ausgelegt werden, dass die kritische Dimension CD in etwa der Auflösung nach Rayleigh (Gl. 5.2) entspricht. Wenn dagegen der Objektkontrast geringer ist, dann ist es sinnvoll, zu $CD = \Delta = 1.5 \cdot r_A$ zu wechseln. Abbildung

5.10.11 zeigt beispielhaft das Verhalten für zwei benachbarte Punkte mit unterschiedlicher Reflektivität ($R_1 = 2R_2$).

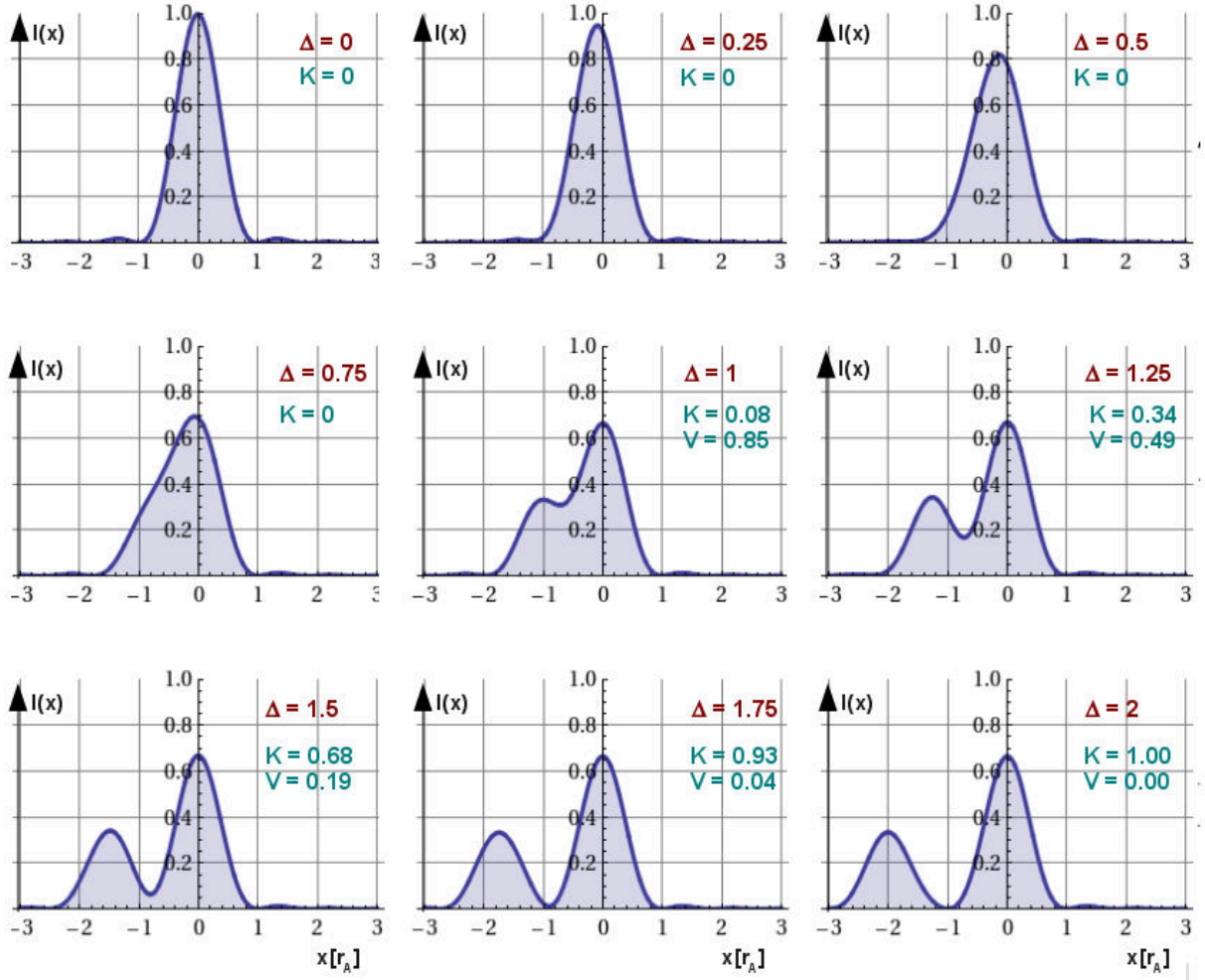


Abbildung 5.10.11: Intensitätsverteilung bei der inkohärenten beugungsbegrenzten Abbildung von zwei unterschiedlich hellen, benachbarten Punkten. Der Abstand der beiden Punkte ist rot gekennzeichnet und in Vielfachen des Airyradius angegeben. Angegeben sind jeweils der Kontrast K und das Verhältnis des Zwischenminimums zum Maximum.

Für deutlich aberrationsbehaftete Systeme ist stattdessen der RMS-Radius des Punktbilds zu verwenden. Natürlich hängt die Punktbildform von der konkret vorliegenden Aberration ab, am ehesten macht es aber Sinn, diese durch eine Gaußkurve anzunähern. In diesem Fall ergeben sich bei der Überlagerung die in Abb. 5.10.12 dargestellten Intensitäten.

Offensichtlich kann der RMS Spotradius nicht direkt mit dem Airyradius der Auflösung nach Rayleigh gleichgesetzt werden. Für Strukturen mit sehr gutem

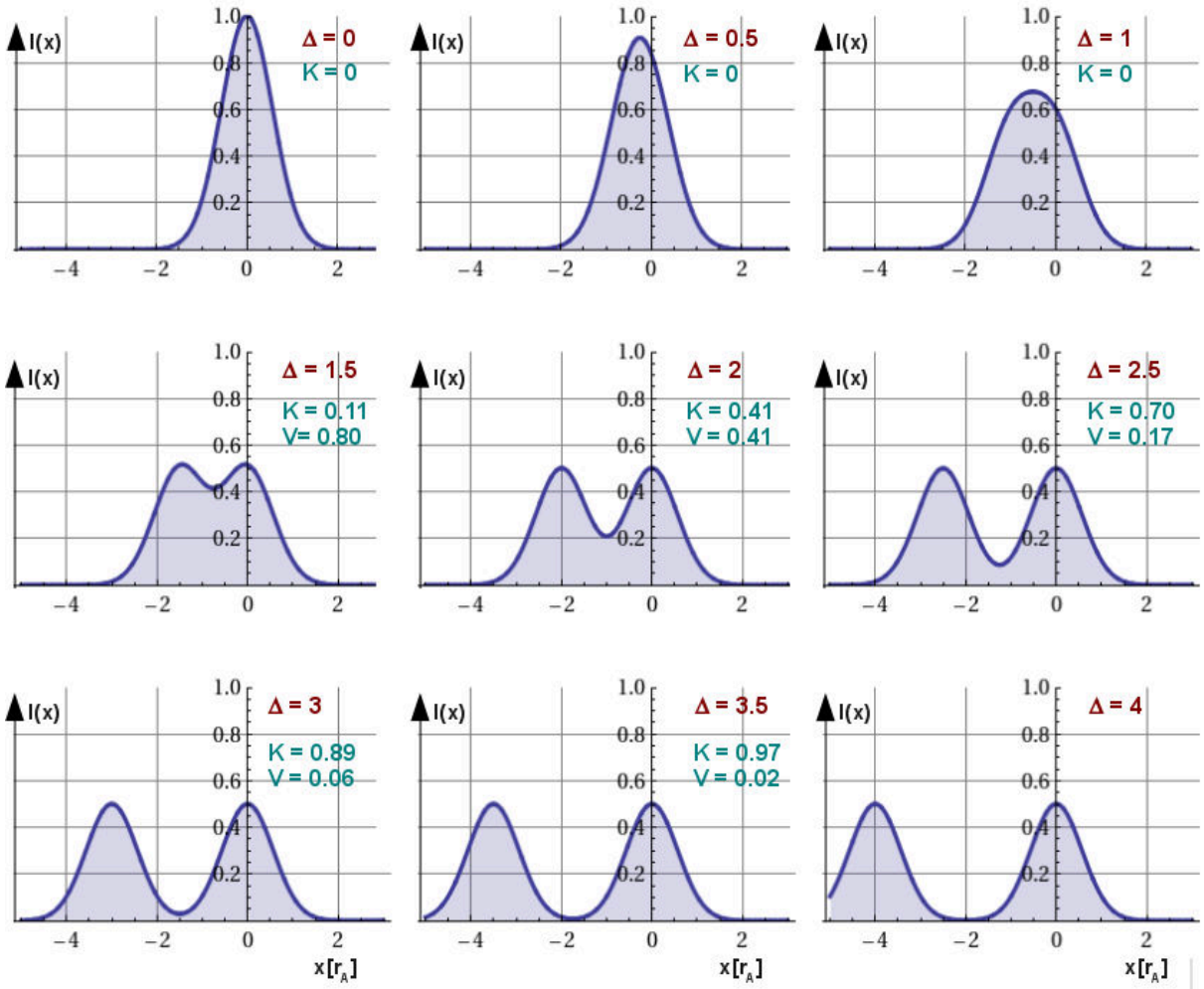


Abbildung 5.10.12: Intensitätsverteilung in der Bildebene bei der Abbildung von zwei eng benachbarten Objektpunkten mit einem aberrationsbehafteten Abbildungssystem. Der Abstand der beiden Punkte ist rot gekennzeichnet und in Vielfachen des Airyradius angegeben.

und schlechtem Kontrast kann als Faustregel

$$CD_{gut} \approx 1.7 \cdot r_{RMS} = 1\, r_A \tag{5.15}$$

$$CD_{schlecht} \approx 2.5 \ldots 3\, r_{RMS} = 1.5\, r_A \tag{5.16}$$

angenommen werden.

Schneider-Kreuznach gibt für Bildverarbeitungsobjektive verwendet eine Zielstellung von einem Kontrast von 30% bei 2/3 der Nyquistfrequenz.[16] Für einen Sensor mit 2.5 μm großen Pixel Pixeln ist $f_{max} = 200$ lp/mm. Der 33% Kontrast wäre dann in der Tat etwa bei 2/3 von diesen 200 lp/mm. D.h. die maximale Blendenzahl K ist dann einfach nach 200 lp/mm = 200.000 lp/m $= 1/(\lambda K)$ bestimmbar und ergibt

[16] S. Mahler, Invsion TechTalk November 2020

$K = 10$ (natürlich ohne Aberrationen).

Bei der Abbildung von Linien ergibt sich gegenüber der Abbildung von Punkten kaum eine nennenswerte Änderung. Abb. 5.10.13 zeigt die sogenannte inkohärente Linespreadfunction [47], beispielhaft für eine numerische Apertur von 0.1 bei einer Wellenlänge von λ =500 nm. Der Unterschied zur zum Vergleich dargestellten Airy-Verteilung ist gering, so dass in aller Regel für benachbarte idealisiert dünne *Linien* ebenfalls die Auflösungsfaustregeln der Gleichungen 5.15 und 5.16 angenommen werden können.

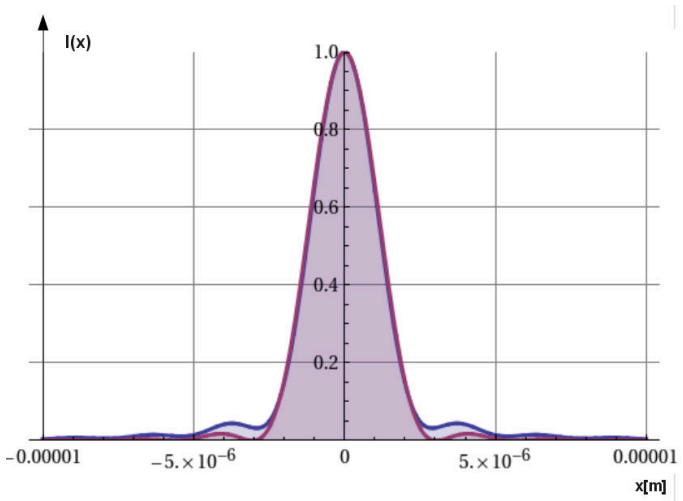


Abbildung 5.10.13: Intensitätsverteilung bei der beugungsbregenzten Abbildung einer idealen Linie (im Vergleich zum idealen Punkt).

In Abb. 5.10.14 sind Vorschläge für eine Wahl von Abtastfrequenzen und erforderlichen Auflösungen (MTF, Spotradius) angegeben. Es zeigt sich aber, dass die Ergebnisse stark von der konkreten Anwendung, den Strukturen und der Algorithmik abhängen. Am besten ist die Simulation entsprechender Bilder zur gezielten Auslegung von Auflösungen und Pixelgrößen.

Wesentlich bei entsprechenden Rechnungen und Angaben ist eine genaue Definition der kritischen Dimension. Hätte man als kritische Dimension z.B. den Abstand zwischen den Mitten der beiden F Querbalken angenommen, hätten sich die Werte schnell um den Faktor 1.5 geändert.

In Abb. 5.10.15 ist ein möglicher Algorithmus für Detektionsaufgaben zur Auswahl von Pixelgröße und optischer Auflösung dargestellt.

		Optische Auflösung			Pixelgröße
		Airy bzw. Blendenzahl r_A=1.22 λ K	r_{RMS}	MTF	
Isolierte Einzelpunkte & Einzellinien					Q = 2..3
CD'	hoher Objektkontrast	r_A= 1...2 CD	r_{RMS}= 1..2 CD	$1/GF_{10}$ = 2 CD	Oder mehr Pixel bei exakter Positionsbestimmung.
	geringer Objektkontrast	r_A= CD	r_{RMS}= CD	$1/GF_{30}$ = 2 CD	stark Rauschabhängig !
Benachbarte Einzelpunkte & Einzellinien					
F CD'	hoher Objektkontrast	r_A= CD	r_{RMS}= CD / 1.5	$1/GF_{10}$ = CD	Q= 2..4
	geringer Objektkontrast	r_A= CD / 1.5	r_{RMS}= CD / 2.5	$1/GF_{30}$ = CD	
(Quasi-)periodische Strukturen					Q = 2..4
CD'	hoher Objektkontrast	r_A= CD	r_{RMS}= CD / 1.5	$1/GF_{10}$ = CD	
	geringer Objektkontrast	r_A= CD / 1.5	r_{RMS}= CD / 2.5	$1/GF_{30}$ = CD	
Kanten		r_A oder r_{RMS} = ca. 1... 10 Pixel (Q=1 .. 10)			

Bildsensor MPixel = (Objektfeld / CD)² • 9...16

Abbildung 5.10.14: Wahl von optischer Auflösung und Pixelgröße basierend auf der kleinsten kritischen Dimension. Die grün dargestellten Parameter sind jeweils bei der Auswahl zu bevorzugen. Rote Kennzahlen sollten nur verwendet werden, wenn die anderen Kennzahlen für die jeweilige Anwendung nicht bekannt sind. Die Angabe für Kanten bezieht sich auf eine modellgestützte Positionsbestimmung der Kante.

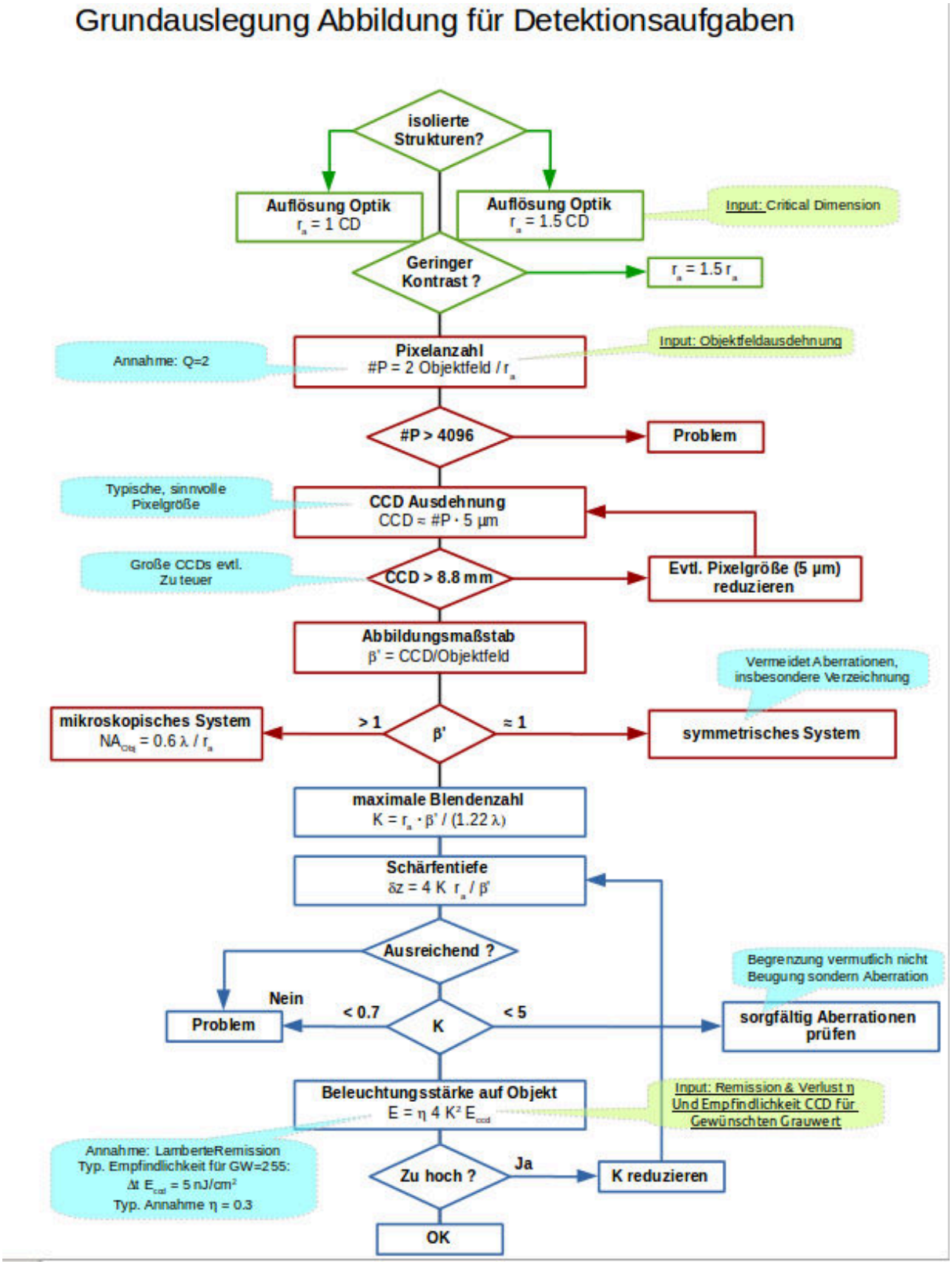


Abbildung 5.10.15: Algorithmus zur Auslegung der Abbildung für detektierende Bildverarbeitungslösungen.

5.11 Advanced: Sensorgröße und Pixelgröße

Eng verbunden mit der Festlegung der optischen Auflösung und Q ist natürlich zwangsläufig die Pixelgröße des Bildsensors und es ist klar, dass die Sensorgröße einen erheblichen Einfluss auch auf die Auslegung der Optik hat: Wenn dasselbe Objektfeld auf einen Sensor mit kleineren Pixeln abgebildet werden soll ist zwangsläufig bei gleichem Objektabstand die Brennweite der Abbildungsoptik reduziert. Bevor wir also die Optik auslegen bzw. auswählen können, muss die *Sensorgröße* definiert sein.

Generell besteht bei Bildaufnehmern ein Trend zu immer kleineren Sensoren und aufgrund der Massenfertigung entsprechender Sensoren ist die Qualität solcher Sensoren mit kleinen Pixeln inzwischen erstaunlich hoch. Weil die Sensorgröße einen so immensen Einfluss auf das gesamte Abbildungssystem hat, sollte ihre Wahl aber gut durchdacht werden.

Prinzipiell kann gemäß der geometrischen Optik ein Objektivdesign mit geänderter Brennweite einfach dadurch gewonnen werden, dass alle geometrischen Größen (Linsenradien, Durchmesser, Abstände) mit dem Skalierungsfaktor der Brennweiten multipliziert wird. Gemäß der geometrischen Optik würde sich das skalierte Objektiv mit dem neuen Bildsensor also exakt gleich verhalten wie die ursprüngliche Kombination von Objektiv und Bildsensor und wir könnten bei oberflächlicher Betrachtung Optik und Bildsensorgröße getrennt betrachten. Aus verschiedenen Gründen ist diese Betrachtung zu einfach und wir wollen etwas ausführlicher auf die sich ergebende Komplexität eingehen, insbesondere da oftmals falsche Schlussfolgerungen gezogen werden.

Zunächst muss festgelegt werden, was man eigentlich vergleicht bzw. was konstant gehalten wird. Wir werden im folgenden davon ausgehen, dass die Anzahl der Pixel konstant bleibt, d.h. wir gehen von zwei Sensoren mit z.B. 10 MPixeln aus und fragen uns, ob es günstiger ist, einen kleinen Sensor oder einen großen Sensor zu verwenden. Dabei soll das Objektfeld natürlich unverändert bleiben und komplett durch den Sensor erfasst werden.

Aber damit ist immer noch nicht klar spezifiziert, wie verglichen werden soll. Wesentlich ist, ob wir wir die Schärfentiefe oder aber die Blendenzahl konstant halten

wollen.

Nehmen wir zunächst den Fall der *konstanten Blendenzahl*. Der große Sensor hat wegen der damit einhergehenden größeren Brennweite damit letztlich eine größere Eintrittspupille. D.h. es wird mehr Licht aufgesammelt, wir können also von einem verbesserten Signal-Rausch-Verhältnis ausgehen (Vorteil 1). Außerdem ist bei beiden Sensoren die theoretisch maximale Auflösung (1.22 λK) gleich, d.h. das Verhältnis von Auflösung zu Pixelgöße ist beim Sensor mit großen Pixeln natürlich besser, bzw. die maximal mögliche Auflösung ist verbessert (Vorteil 2). Und schließlich sind die fertigungsbedingten Toleranzen bei der großen Optik günstiger (Vorteil 3). Im Vergleich mit sehr kleinen Pixeln ist mit die Sättigungskapazität deutlich erhöht (maximales Signal-Rausch-Verhältnis, siehe Abschnitt 9, Vorteil 4) und das Übersprechen zwischen benachbarten Pixel reduziert (Vorteil 5). Dem stehen natürlich die erhöhten Sensorkosten (Nachteil 1), eine verringerte Schärfentiefe (Nachteil 2) und ein größeres optisches System (Nachteil 3) entgegen.

Für den Fall einer *konstanten Schärfentiefe* ist zu beachten, dass aufgrund des größeren Abbildungsmaßstabs bei großen Sensoren dieselbe Schärfentiefe nur durch Abblenden erreicht wird. Damit verringert sich die Fläche der Eintrittspupille. Letztlich ergibt sich dieselbe Eintrittspupillenfläche wie für den kleinen Sensor. Somit erzielen beide Sensoren dieselbe Lichtnutzung und somit (hinsichtlich des Photonenrauschens) dasselbe Signal-Rausch-Verhältnis. Die Vorteile 3 bis 5 sowie Nachteil 1 bleiben. Zusätzlich ergeben sich aber noch durch die vergrößerten Blendenzahl verringerte Aberrationen (oder verminderte Kosten für die Optik).

Summa summarum sind somit große Sensoren fast immer zu bevorzugen, wenn die Kosten und der notwendige Bauraum keine Probleme darstellen.

Wer eine einfache Grundaussage sucht: Wenn möglichst viel Information übertragen, also viele Pixel abgebildet werden sollen, dann sind große Systeme zu bevorzugen.

Abgesehen von diesen sehr grundsätzlichen Überlegungen, sind auch noch rein praktische Fertigungsaspekte zu beachten. Bei einer Skalierung des Abbildungssystems skalieren natürlich auch die Toleranzen hinsichtlich der Fertigung und Justage und die Herstellung sehr kleiner Systeme verteuert sich dadurch unter

Umständen. Wenn also auch die Abbildungsfehler entsprechend skaliert werden sollen, also so, dass sie z.B. nach wie vor einem Pixel entsprechen, dann werden die Anforderungen an die Fertigung stark erhöht. Andererseits führen sehr große Systeme ganz entsprechend auch zu hohen Herstellungskosten, nicht zuletzt wegen des Materialeinsatzes und der notwendigen Mechanik.

Kleine Pixel in Verbindung mit den verkleinerten Optiken können hinsichtlich der Abbildungsqualität daher sehr problematisch sein, auch wenn durch die Skalierung natürlich auch der Durchmesser des Bildfelds skaliert wird.

Ein weiteres Problem kleiner Pixel ist das Rauschverhalten. Die *Bestrahlungsstärke* (Leistung pro Flächeneinheit) ist zwar konstant aber letztlich ist das Signal-Rausch-Verhältnis prinzipiell oft durch die Anzahl der genutzten Photonen pro Pixel begrenzt (vgl. Abschnitt 9.5).

Bei der Skalierung der Brennweite der Optik führt eine konstante Blendenzahl natürlich gemäß $K = f'/D$ auch zu einer Skalierung der Eintrittspupille. Die Eintrittspupille wird also verkleinert und somit wird die aufgenommene Lichtmenge reduziert. Oftmals kann der Verlust an Lichtstärke nicht toleriert werden. Dadurch werden sehr geringe Blendenzahlen notwendig. Bei Consumerprodukten ergeben sich teilweise Blendenzahlen in der Größenordnung 1,2 bis 1,5. Dies führt dann aber — bei einer brauchbaren Abbildungsqualität — zu entsprechend aufwändigen Optiken. In Kombination mit den erhöhten Anforderungen an die Toleranzen bei der Fertigung kann so der Vorteil der preiswerteren Sensorfläche insgesamt schnell ins Gegenteil verkehrt werden, insbesondere wenn die notwendige Stückzahl für eine entsprechende Produktion nicht gegeben ist. Für Smartphone Kameras werden so teilweise 8 asphärische Flächen verwendet, ein für Produkte mit geringer Stückzahl kaum wirtschaftlich realisierbarer Wert und die asphärischen Flächen führen zu erhöhten Anforderungen hinsichtlich der Justagetoleranzen.

Und schließlich ergibt sich noch das Problem, dass bei sehr kleinen Pixeln das optische und elektrische Übersprechen benachbarter Pixel deutlich zunimmt, so dass sich bei Abbildung feiner Strukturen der Kontrast verschlechtert [18]. Allerdings wird nicht alles durch kleine Pixel schlechter. Bei gleicher Blendenzahl vergrößert sich die Schärfentiefe denn der Abbildungsmaßstab wird verringert.

Letztlich ist es — trotz teilweise anderweitiger Aussagen — schwierig, dieselbe Performance in einer Anwendung mit kleinen Pixeln bei gleicher Pixelanzahl zu realisieren. Der Kostenvorteil beim Sensor und beim Volumen der Optik führt meist zu einer Verringerung der Gesamtabbildungsleistung oder aber zu aufwändigen Optiken. Für die meisten Anwendungen kann in der industriellen Bildverarbeitung daher von Pixelgrößen zwischen p =3 und 7 Mikrometern ausgegangen werden. Die im Consumerbereich üblichen extrem kleinen Pixel im Bereich zwischen 1 und 2 Mikrometern führen bei hohen Qualitätsanforderungen nicht zu einer Kostenreduktion und sind vor allem der dort vorherrschenden starken Bauraumreduktion und weniger strikten Bildqualitätsanforderungen geschuldet.

Das bedeutet aber nicht, dass kleine Pixel generell gemieden werden sollten. Bisher sind wir von einer konstanten Anzahl von Pixeln ausgegangen. Geht man aber von einer konstanten Sensorfläche aus, dann können kleine (und damit viele) Pixel durchaus Sinn machen. Dabei steigt zwar die zu verarbeitende Informationsmenge (Hauptnachteil), die Gesamtauflösung kann aber eben aufgrund der Überabtastung steigen und es ergeben sich keine Probleme mit Moires, insbesondere bei Farbanwendungen.

Letztlich bleibt die Wahl der Sensorgröße und Pixelgröße schwierig und in der Regel ist es sinnvoll, mehrere Varianten der Kombination aus Bildaufnehmer und Abbildungsoptik durchzuspielen.

Wenn die Pixelgröße festliegt kann basierend auf der kritischen Dimension CD der kleinsten aufzulösenden Details direkt bei gegebenem Q der Abbildungsmaßstab berechnet werden (vgl. Abb. 5.11.16):

$$\beta' \, CD = Q \, p \tag{5.17}$$

Für Q=4 ist der sich ergebende Abbildungsmaßstab in Abb. 5.11.16 dargestellt.

Beispiel: Sie sollen die Optik eines Spionagesatelits auslegen. Dabei soll Q=2 bei 10 μm Pixeln erreicht werden. Der Abbildungsmaßstab beträgt $\beta' = -1 \cdot 10^{-4}$. Welche Auflösung (im Sinne der Zweipunkt-Auflösung) wird am Boden (Erde) erreicht ?
Lösung: $CD = 2 \cdot 10\mu m \cdot 10^4 = 0.2$ m

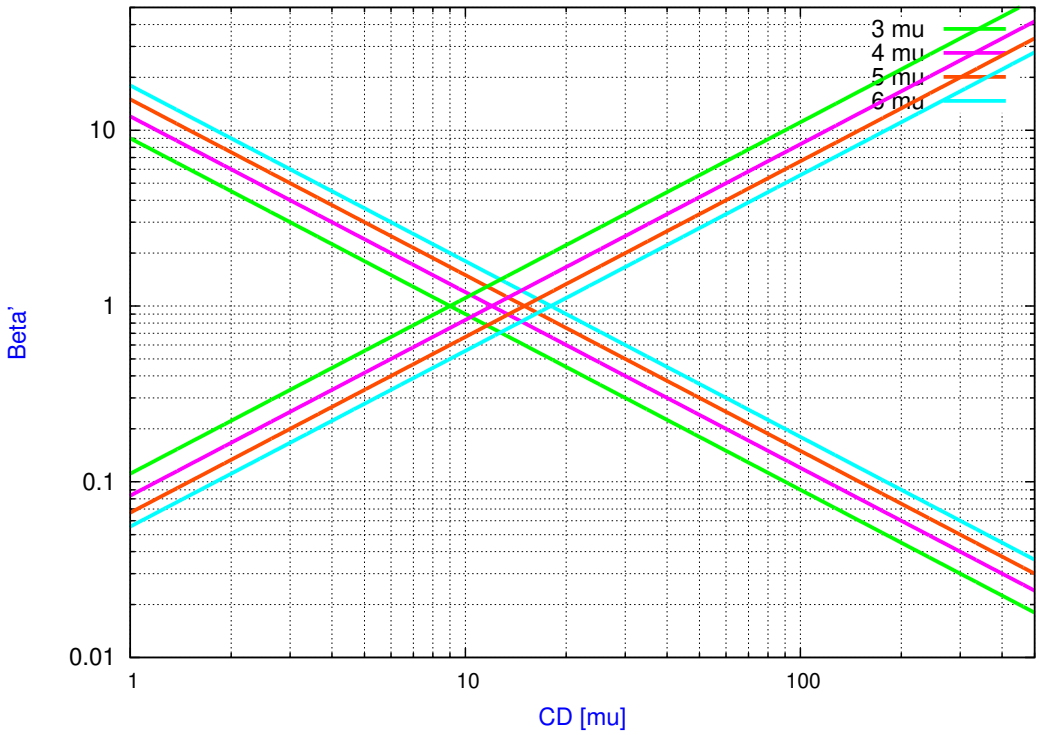


Abbildung 5.11.16: Wahl des Abbildungsmaßstabs bei Q=4 für verschiedene kritische Dimensionen der aufzulösenden Defekte bei verschiedenen Pixelgrößen (3 ... 6 μm). Q = 4 bedeutet, dass 4 Pixel zum Abtasten der Auflösungsbegrenzung (z.B. Rayleigh Radius) verwendet werden. Bei isolierten Strukturen mit gutem Kontrast werden soll pro Struktur zwischen 2 und 4 Pixel (linear) verwendet.

Beispiel: Sie wollen benachbarte Linien auf einem Objekt getrennt detektieren. Der minimale Linienabstand beträgt 50 μm . Sie setzen eine Kamera mit 5 μm Pixeln ein. Was müssen Sie hinsichtlich der Auflösung beachten?

Lösung: Gemäß Abb. 5.10.14 fordern wir $r_A = 0.8 \cdot 50 = 40\ \mu\mathrm{m}$ auf dem Objekt. Außerdem fordern wir, dass den 50 μm 3.5 Pixel entsprechen, also objektseitig einem Pixel 14 μm entsprechen. Damit erhalten wir einen Abbildungsmaßstab von $\beta' = -14/50$ = -0.28. Die Auflösung nach Airy im Bildraum wird daher 40 · 0.28 = 11.2 μm . Damit berechnen wir die Blendenzahl zu $K = 11.2/1.22{\cdot}500\mathrm{nm} = 18$. D.h. hinsichtlich der Auflösung müssen wir keine große Blende (kleine Blendenzahl) verwenden (evtl. aber aufgrund der Lichtstärke). Aufgrund der großen Blendenzahl lässt sich mit konventionellen Optiken bei begrenztem Feld auch ohne Probleme die beugungsbegrenzte Abbildungsqualität erzielen. Die Anforderungen an die Abbildung lassen sich also hinsichtlich der geforderten Auflösung einfach erfüllen.

5.12 Advanced: Die maximale Anzahl abbildbarer Pixel

Wieviel Pixel könnten rein theoretisch abgebildet werden wenn wir Aberrationen vernachlässigen und von einer beugungsbegrenzten Auflösung ausgehen?

Wir folgen dazu den Ausführungen von Lohmann bzw. Cossair [7] und betrachten die maximal aufgelöste Anzahl von Bildpunkten, die letztlich dem sogenannten „Ortsbandbreitenprodukt" (engl. space-bandwidth product, SBWP) entspricht:

$$SBWP \sim \frac{\Delta x \Delta y}{(\lambda K)^2} \tag{5.18}$$

Dabei haben wir eine beugungsbegrenzte Auflösung λK (vgl. Gl. 5.2) sowie eine Bildausdehnung $\Delta x\, \Delta y$ angenommen.

Das System wird nun um einen Faktor M, z.B. $M = 10$, skaliert, d.h. alle geometrischen Größen (z.B. Brennweite, Durchmesser Linsen, Bildausdehnung) werden mit M multipliziert. Dadurch bleibt zunächst die Blendenzahl erhalten.

Wenn wir zunächst (eher unrealistisch) annehmen, dass es sich in beiden Fällen um beugungsbegrenzte Systeme handelt, dann ist klar, dass das SBWP sich um den Faktor M^2 vergrößert:

$$SBWP_M \sim \frac{M\, \Delta x\, M\, \Delta y}{(\lambda K)^2} = M^2\, SBWP \tag{5.19}$$

In der Praxis werden durch die Skalierung des Systems natürlich auch etwaige Aberrationen um den Skalierungsfaktor vergrößert. Wenn das skalierte System also im Grenzfall nicht durch die Beugung sondern durch die Aberrationen begrenzt ist, dann ist schlimmstenfalls nichts gewonnen denn die Bildpunktfläche steigt somit mit M^2 und damit bleibt das Ortsbandbreitenprodukt konstant.

Der entscheidende Punkt ist nun, dass die Blendenzahl vergrößert wird, um die Aberrationen konstant (idealerweise sogar unterhalb der Beugungsbegrenzung) zu halten. Wenn wir die Blendenzahl um den Faktor M steigern würden, dann

wäre wieder das Ortsbandbreitenprodukt konstant, denn auch in diesem Fall würde wegen $r \approx \lambda K$ ja die Bildpunktfläche wieder mit M^2 skalieren.

In der Praxis ist aber eine wesentliche geringere Reduzierung der Apertur (Vergrößerung der Blendenzahl) notwendig, um die Abbildungsleistung konstant zu halten. In [7] ist dies explizit für sphärische Aberrationen berechnet. Oftmals werden in der Praxis bei kommerziellen Objektiven die Blendenzahlen proportional zur dritten Wurzel der Brennweite (in mm) gewählt, also z.B. Blendenzahl 4.6 für Brennweite 100 mm, um die Spotgröße konstant zu halten.

Bei einer Skalierung mit $M = 10$ ändert sich daher die Blendenzahl lediglich um den Faktor $10^{0.33} = 2.1$. Damit ergibt sich

$$SBWP_{10} =\sim \frac{10\Delta x\, 10\, \Delta y}{(2.1\lambda K)^2} \approx 25\, SBWP \tag{5.20}$$

Es werden also 25 mal so viele Pixel durch das optische System übertragen. Bei gleicher Pixelgröße ist allerdings zu beachten, dass die vergrößerte Blendenzahl zu einer reduzierten Bestrahlungsstärke jedes Pixels und damit zu einem verringerten Signal-Rausch-Verhältnis führt. Nach Gl. 7.11 ist die Bestrahlungsstärke proportional zu $1/K^2$ und daher ergibt sich hier eine Reduzierung der Bestrahlungsstärke um den Faktor $2.1^2 \approx 4$. Will man die Anzahl der Photonen pro Pixel bei gleicher *Belichtungszeit* konstant halten, muss man offensichtlich die Pixelfläche um den Faktor 4 vergrößern und damit reduziert sich natürlich das Ortsbandbreitenprodukt um den Faktor 4. Letztlich bleibt aber dann dennoch eine deutliche Vergrößerung des Ortsbandbreitenprodukts bestehen:

$$SBWP'_M = \frac{25}{4} SBWP \approx 6\, SBWP \tag{5.21}$$

5.13 Advanced: Weitere Auflösungskennzahlen

In der Bildverarbeitung werden in aller Regel Auflösungen gemäß Rayleigh (s.o.) oder aber mittels der Modulationstransferfunktion (siehe Kapitel 4) definiert. Dennoch werden teilweise auch andere Angaben verwendet. Die wichtigsten sind:

Auflösung nach Abbe: Die Auflösung nach Abbe ist formal bis auf einen Vorfak-

tor mit der Auflösung nach Rayleigh (Gl. 5.2) identisch (1 statt 1.22). Letztlich wird sie aus der Abbildung eines periodischen Gitters abgeleitet (siehe Abschnitt 5.3 und z.B. [47]) und macht hier auch als Angabe Sinn. Sie wird besonders häufig in der Mikroskopie eingesetzt.

Sparrow Kriterium: Eine weitere Abwandlung des Rayleigh Kriteriums stellt das Sparrow Kriterium dar. Dabei ergibt sich gerade ein Minimum für die Einhüllende bei der Abbildung von zwei eng benachbarten Punkten auszubilden. Da dieses Minimum aber zunächst infinitesimal klein ist, ist das Kriterium in der Praxis meist wenig hilfreich. Gegenüber der Auflösung nach Rayleigh ergibt sich ein Vorfaktor von 0.93 anstatt 1.22.

Strehl Kriterium: Die Strehlzahl bzw. Strehlverhältnis ist das Verhältnis aus Intensität im Spotmaximum bei der Abbildung eines einzelnen Punktes im Vergleich zur Intensität an der selben Position für ein System ohne Aberration. Somit ist die Strehlzahl vor allem geeignet, um die Korrektur bzw. Qualität einer optischen Abbildung zu beurteilen (z.B. Strehlzahl $> 80\%$). Für die Auslegung einer konkreten BV-Anwendung ist der Parameter sehr ungünstig bzw. muss erst näherungsweise in einen Spotradius umgerechnet werden.

Three-Bar Resolution: Sehr gebräuchlich ist die Beurteilung der Bildqualität an sogenannten *Three-bar resolution targets*, die aus unterschiedlich skalierten Gruppen von jeweils drei Balken bestehen. Am gebräuchlichsten sind die *USAF 1951* targets.

RMS-Spotradius: Der mittlere Radius der Punktbildverwaschungsfunktion kann sowohl simulativ als auch experimentell bestimmt werden und ist ein sehr anschaulicher Parameter.

Heynacher Kennzahlen: Für den subjektiven Bildeindruck ist vor allem der Kontrast bei mittleren Frequenzen entscheidend. Die Heynacher Kennzahl trägt dem Rechnung[34]:

$$H = 14 \log(\frac{1}{40} \int_0^{40/mm} MTF(f) df \tag{5.22}$$

Subjective Quality Factor: Ebenfalls eine Kennzahl, die versucht, den subjektiven Bildeindruck zu erfassen. Die Zusammenhänge sind einigermaßen kompliziert (siehe [34]).

Eine Umrechnung verschiedener Qualitätskennzahlen ineinander ist meistens nicht bzw. eben nur im Sinne einer Näherung möglich.

Sobald die kompletten Komponenten der Abbildungskette vorhanden sind wird man für die Charakterisierung der Auflösung in der Praxis die letztlich zu betrachtenden Merkmale mit dem kompletten Abbildungssystem abbilden und so entscheiden, ob die Abbildungsqualität insgesamt ausreichend ist.

Für einen quantitativen Vergleich der Abbildungsleistung aber auch für eine Beurteilung bei nicht-vorhandenen Objektdetails (z.B. Kratzer der richtigen Größe) kann es aber Sinn machen, standardisierte Testobjekte zu verwenden. Meist werden binäre Teststrukturen eingesetzt, denn zum einen ist die Herstellung vergleichsweise einfach und zum anderen ist der Kontrast des Objekts damit klar definiert (nahezu 1). Gebräuchlich und kommerziell einfach verfügbar sind:

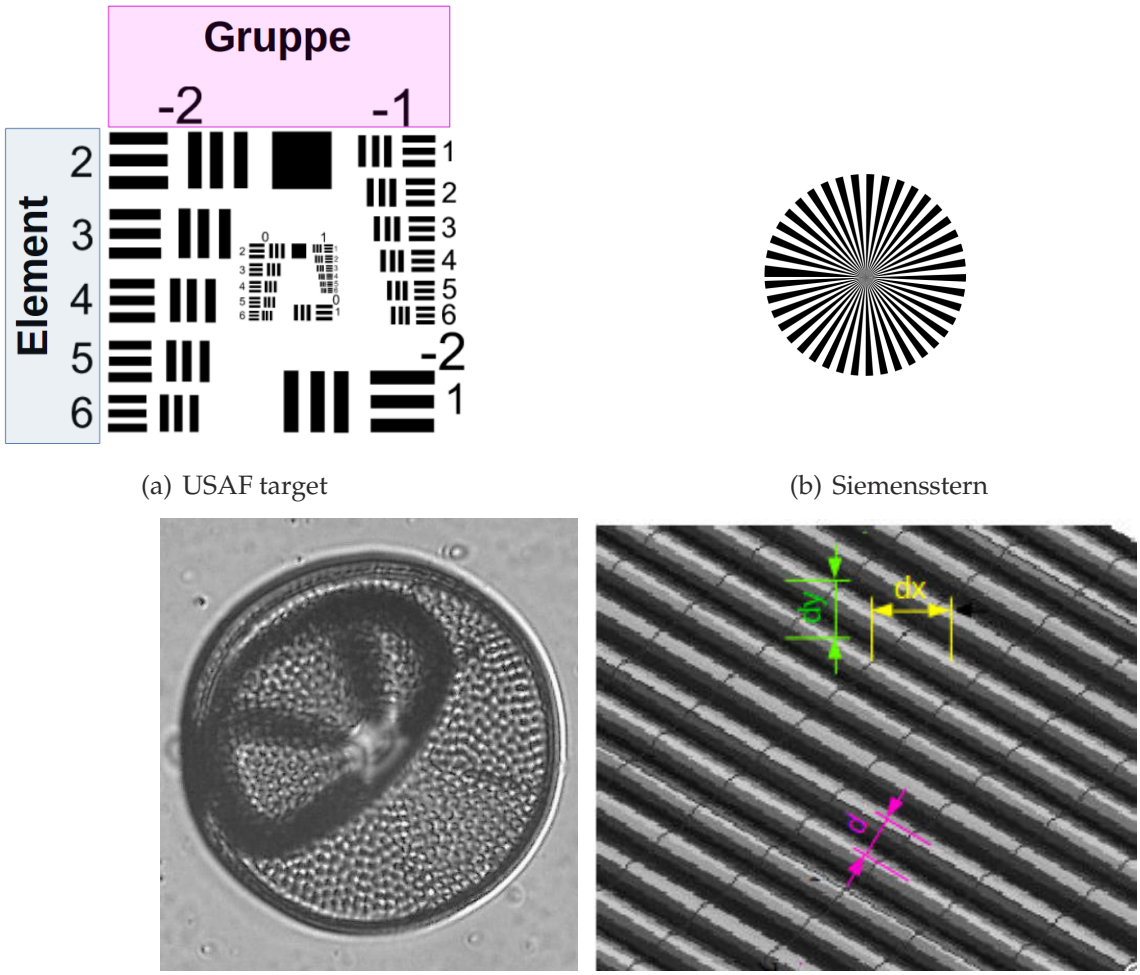
- USAF 1351 Three-Bar Teststrukturen[17]
- Gitter mit verschiedenen Frequenzen ($\rightarrow$ MTF)
- Siemenssterne
- IEEE Resolution Target
- NBS 1963A Resolution Target
- ISO 12233 test chart
- standardisierte Bilder von Szenen (z.B. IEEE)
- Kieselalgen (Diatomeen)

Für ortsinvariante Systeme sind Siemenssterne sehr gut geeignet. Das kreisförmige Testmuster besteht aus 16 oder 32 Schwarz-Weiß-Speichen. Im Zentrum des Kreises wird die lokale Streifendichte sehr hoch, so dass man zur Auflösungsbeurteilung den Radius verwenden kann, bei dem die Speichen gerade noch aufgelöst werden. Aufgrund der radialen Symmetrie lässt sich sehr gut beurteilen, wie die Auflösung mit der Orientierung variiert.

[17] Das USAF 1951 Testmuster liefert rein aufgrund der Dreierstrukturen eine etwas andere Auflösung als ein regelmäßiges ausgedehntes Gitter.

Ebenfalls kann man anhand eines einzelnen Bilds im Falle einer ortsinvarianten Abbildungsqualität (dies ist allerdings eine oftmals nicht realistische Nebenbedingung) die MTF für alle möglichen Streifenfrequenzen und Orientierungen messen.

Oft werden gängige bzw. altbekannte Bilder zur Untersuchung insbesondere von Algorithmen und Auflösungen untersucht. Direkt verfügbar sind viele unter http://www.hlevkin.com/06testimages.htm (vgl. Abb. 5.13.18). Der Vorteil der Verwendung von Standardtestbildern ist, dass damit ein Vergleich der Ergebnisse aus unterschiedlichen Quellen einfacher möglich ist. Eigentlich wäre also eine Standardisierung dieser Muster sinnvoll.



(a) USAF target

(b) Siemensstern

(c) Diatom

(d) Gitter

Abbildung 5.13.17: verschiedene typische Teststrukturen.

Wenn das feldabhängige Verhalten geprüft werden soll muss entweder das Testobjekt verschoben werden oder aber das Testobjekt wird direkt repliziert über das gesamte Feld verwendet.

Generell gilt natürlich, dass man Teststrukturen zur Prüfung verwenden sollte, die einerseits eine quantitative Beurteilung ermöglichen, andererseits aber möglichst ähnlich zur geplanten Anwendung sind.

(a) Lena

(b) Airplane

(c) Baboon

(d) crowd

(e) couple

(f) cameraman

Abbildung 5.13.18: Verschiedene gebräuchlich Testbilder

5.14 Advanced: Kohärente Abbildungen

Kohärente Abbildungen werden in der industriellen Bildverarbeitung vergleichsweise selten eingesetzt[18]. Insbesondere ergeben sich durch die *Kohärenz* Probleme an Kanten (Überschwingen) und Rauschen aufgrund von *Granulation* (*Speckles*) (siehe Abb. 5.14.21). Von der rauen Oberfläche, die mit kohärentem Licht beleuchtet wird, starten Kugelwellen mit (aufgrund der *Rauheit*) zufälliger Phase zur Abbildungsoptik (*subjektive Speckles*) oder einem Bildaufnehmer (*objektive Speckles*) und interferieren dort (Abb. 5.14.20).

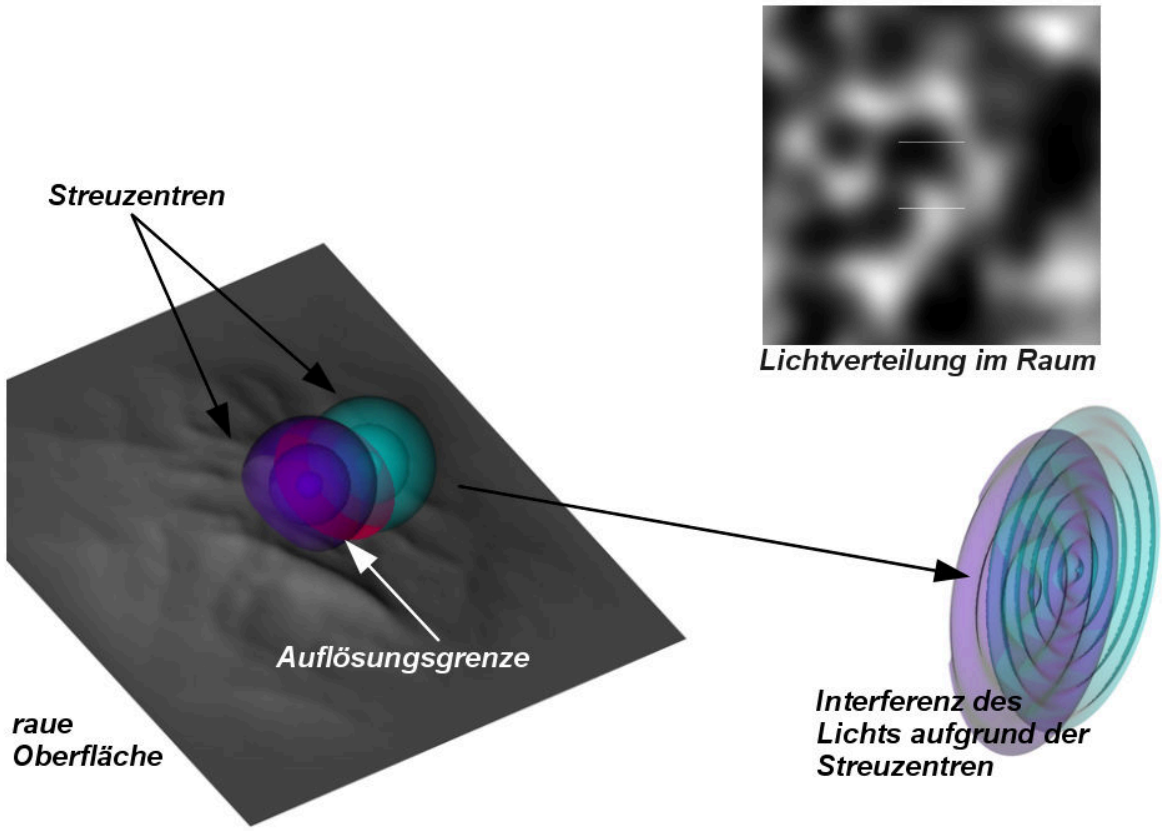


Abbildung 5.14.19: Prinzip der Speckleentstehung: Die von eng benachbarten Punkte auf der rauen Oberfläche ausgehenden Wellen überlagern sich in der Bildebene aufgrund der begrenzten Auflösung (Bildfehler und Beugungsbegrenzung). Die zufälligen Phasen dieser Wellen (raue Oberfläche) führen bei der Interferenz zu zufälligen Helligkeitsmuster. Voraussetzung ist allerdings ausreichende zeitliche und räumliche Kohärenz.

Die mathematische Beschreibung der Abbildung erfolgt durch eine *kohärente*

[18]Eine Ausnahme stellen strukturierte Beleuchtungen dar, die oft mittels diffraktiven optischen Elementen und kohärenter Beleuchtung realisiert werden.

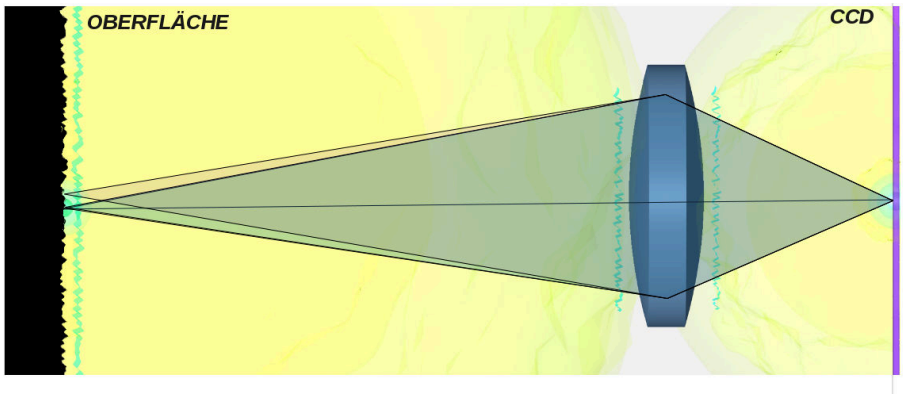


Abbildung 5.14.20: Entstehung von Speckles durch Interferenz bei rauen Oberflächen (→ zufällige Phase) in Verbindung mit kohärenter Beleuchtung. Die von eng benachbarten Punkten der Oberfläche ausgehenden Kugelwellen interferieren in einem Punkt auf dem Sensor. Aufgrund der zufälligen Phasen dieser Kugelwellen ergeben sich nicht vorhersehbare Interferenzen.

Übertragungsfunktion. Die inkohärente und kohärente Übertragungsfunktion sind über die Autokorrelation miteinander verbunden.[19]

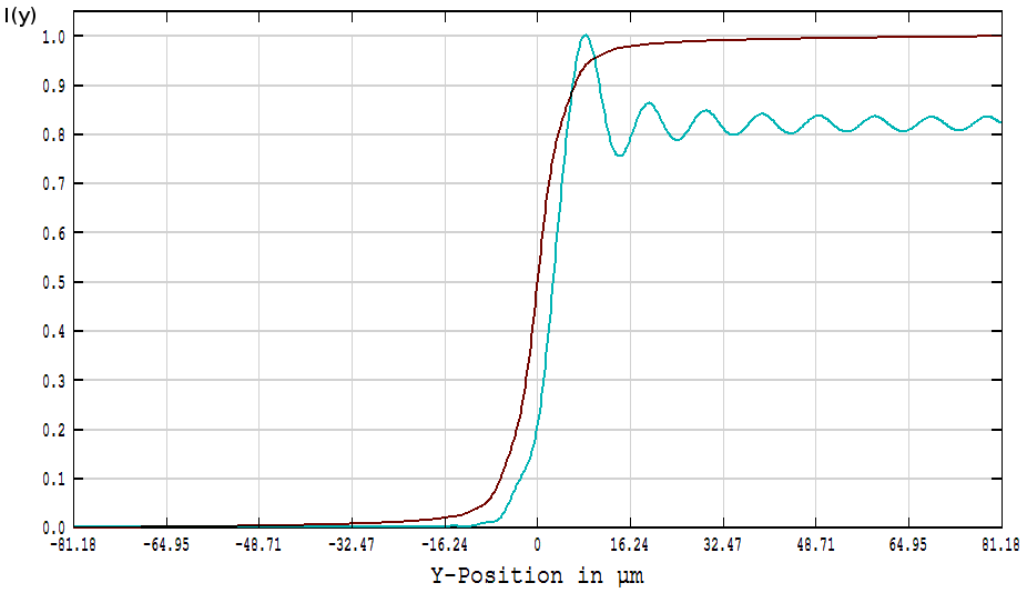


Abbildung 5.14.21: Überschwinger an Kanten bei kohärenter Abbildung eines Objekts (im Vergleich zur inkohärenten Abbildung).

Abb. 5.14.19 und 5.14.20 verdeutlichen das Zustandekommen der Granulation (Speckles). Eine raue Oberfläche wird mit kohärentem Licht beleuchtet. Hierdurch

[19] Im kohärenten Fall gilt die Linearität für das elektrische Feld während im inkohärenten Fall die Linearität für die Intensität, also das elektrische Feld im Quadrat gilt. Daher ist das inkohärente Punktbild das Quadrat des kohärenten Punktbilds und aufgrund des Faltungssatzes ist dann die inkohärente Übertragungsfunktion die Autokorrelation der *kohärenten Übertragungsfunktion*.

gehen von der Oberfläche Kugelwellen mit unterschiedlicher Phase aus, denn die Position der Kugelwellen ist aufgrund der Rauigkeit zufällig. Bei der kohärenten Überlagerung (s.u.) werden nun die elektrischen Felder aufgrund aller Kugelwellen in der Sensorebene addiert und das sich so ergebende Gesamtfeld wird quadriert, um die Intensität zu erhalten (Abb. 5.14.22). Dabei ergeben sich Stellen „destruktiver Interferenz" bei denen die Intensität Null wird.

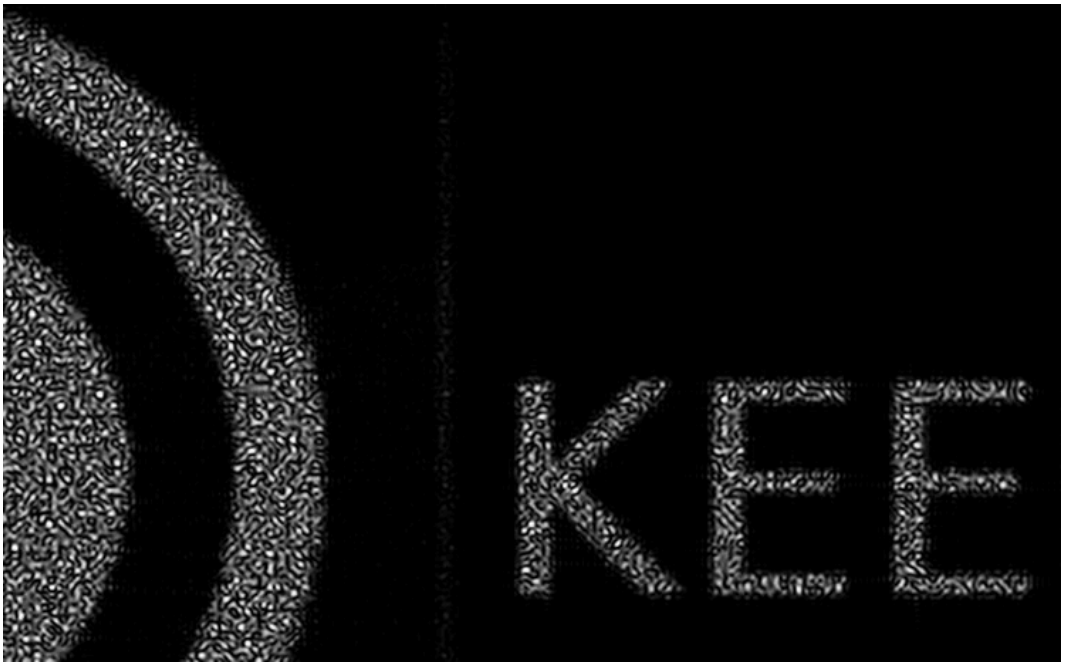


Abbildung 5.14.22: Ausschnitt eines (simulierten) Bildes bei vollständig kohärentem Licht und einer Abbildung mit geringer Blende. Geringe Intensitäten kommen dabei deutlich weniger häufig als hohe Intensitäten vor. Die mittlere Größe eines Speckles entspricht gerade der Auflösung nach Rayleigh.

Ob eine kohärente, eine inkohärente oder aber eine teilkohärente Abbildung vorliegt, muss basierend auf Parametern zur Beschreibung der *zeitlichen* und *räumlichen* Kohärenz entschieden werden. Nur wenn *sowohl* die zeitliche wie auch die räumliche Kohärenz gegeben ist muss mit Problemen aufgrund von Speckles oder anderen kohärenten Störungen gerechnet werden.

Zentral ist zunächst die zeitliche Kohärenz. Sie wird spezifiziert durch die sogenannte *Kohärenzlänge* L die sich letztlich aus der spektralen Bandbreite $\Delta\lambda$ des genutzten Lichts nach

$$L = \frac{\lambda^2}{\Delta\lambda} \tag{5.23}$$

berechnet.[20] Anschaulich gibt die Kohärenzlänge direkt die Länge der Wellenpakete an, die die Lichtquelle aussendet. Eine starke Einschränkung der Wellenlänge, z.B. zur starken Fremdlichtreduktion, mit einem sehr engen Bandpassfilter führt zu einer großen Wellenpaketlänge[21] und damit zu hoher zeitlicher Kohärenz und unter Umständen Speckles.

Damit Speckles auftreten können muss die Rauheit der Oberfläche deutlich kleiner als die Kohärenzlänge sein (z.B. um den Faktor 5..10 kleiner). Der Kontrast der Speckles hängt in komplexer Weise vom Oberflächenprofil sowie der Geometrie der Abbildung (Winkel) ab. Persson gibt z.B. für raue Oberflächen mit einer arithmetischen Rauheit R_a von 50 nm einen Kontrast von ca. K=0.1 an [40]. Bei 120 nm sind es bereits K=0.42. Gleichzeitig darf die Rauheit aber auch nicht zu klein sein, um Speckles zu ergeben. Für sehr kleine Rauheiten ($< \lambda/20$) verschwinden daher die Speckles.

Beispiel: : Sie benutzen eine Laserdiode bei $\lambda = 600$ nm mit einer spektralen Halbwertsbreite des Spektrum von $\delta\lambda = 0.1$ nm. Das beleuchtete Objekt hat eine Rauheit von 50 nm RMS. Schätzen Sie ab, ob mit Speckles zu rechnen ist oder nicht.

Lösung: Die Kohärenzlänge beträgt $L = 600^2/1$ = 360.000 nm = 0.36 mm und liegt daher deutlich oberhalb der Oberflächenrauheit. Die Oberflächenrauheit ist also „klein genug". Diese Voraussetzung für die Ausbildung von Speckles ist daher gegeben. Wir müssen aber zusätzlich untersuchen, ob die Rauheit stark genug ist um überhaupt ausreichend zufällige Phasen zu generieren. Dies ist der Fall denn 50 nm $> \lambda/20 = 30$ nm.

Der zweite entscheidende Parameter, die *räumliche Kohärenz*, ist letztlich für in der Bildverarbeitung anfallende Aufgabenstellungen durch die (effektive) räumliche Ausdehnung der Beleuchtungsquelle gegeben. Für *Laser* ist die räumliche Kohärenz immer gegeben. Man kann sich das den Laser verlassende Licht in geometrischer Näherung am einfachsten durch das in Abb. 5.14.23 dargestellte

[20] Je nachdem, wie man die Interferenz beurteilt wird teilweise (siehe z.B. [41]) auch die Hälfe des Werts angegeben (in diesem Fall entspricht dies einer maximalen Phasenschiebung von 90° der Wellen zueinander).

[21] Im Grenzfall einer reinen monochromatischen Schwingung, also ein Sinus, ergibt sich eine unendliche Ausdehnung

Modell vorstellen, d.h. das Licht kommt scheinbar aus einem Punkt. Für andere Quellen, z.B. LEDs, ergibt sich räumliche Kohärenz wenn die Winkelausdehnung, mit der das Objekt beleuchtet wird klein gegen die laterale Ausdehnung des kohärenten Bereichs auf dem Objekt ist. Für eine kreisförmige Quelle mit dem Durchmesser D ergibt sich ein Durchmesser des in sich kohärenten Bereichs im Abstand z von etwa $\lambda \cdot z/D$.

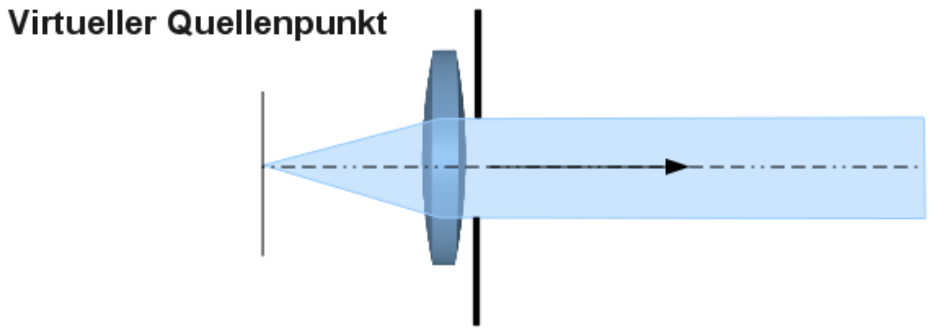


Abbildung 5.14.23: Das Lichtfeld eines Laser ist räumlich und zeitlich kohärent. Man kann es sich aus einem Punkt (räumliche kohärent) kommend und einfarbig (zeitlich kohärent) denken.

In aller Regel sind Speckles bei der industriellen Bildverarbeitung unerwünscht (Ausnahme: Messverfahren, die auf Speckles basieren. Daher werden selten Laser zur Beleuchtung eingesetzt. Andererseits haben Laser den Vorteil, dass die gesamte Energie eben bei nahezu einer Wellenlänge abgestrahlt wird und daher eine Fremdlichtunterdrückung sehr effektiv durchgeführt werden kann.

Wenn Speckles also bei einer Laserbeleuchtung vermieden werden sollen muss eine wie auch immer geartete Mittelung verwendet werden. Im einfachsten Fall erfolgt eine räumlich Mittelung auf den einzelnen Pixeln des Bildsensors. Entscheidend hierbei ist dann natürlich der mittlere Durchmesser eines Speckles. Die mittlere Größe der Speckles ist gerade durch die Auflösung des Abbildungssystems nach Airy gegeben. Eine Abbildung mit der Blendenzahl K führt also zu Speckles mit einem mittleren Durchmesser von 2.44 $\lambda\, K$:

$$D_{speckle} \approx 1.22 \frac{\lambda}{\mathrm{NA}'} = 2.44\, \lambda\, K \tag{5.24}$$

Beispiel: Für eine Anwendung mit kohärenter Beleuchtung sollen im Durchschnitt ca. 25 Speckles auf einen einzelnen Pixel fallen, um eine ausreichende Reduktion des specklebedingten Rauschens zu erreichen. Die Pixelgröße beträgt 8 $\mu\mathrm{m}$. Welche Blendenzahl können wir verwenden?

Lösung: 25 Speckles bedeuten, dass 5 x 5 Speckles auf einen Pixel fallen. Dementsprechend sollte die mittlere Specklegröße 8 $\mu\mathrm{m}$ / 5 = 1,3 $\mu\mathrm{m}$ betragen.

Für den Radius des mittleren Speckles bedeutet das, dass dieser kleiner 0,65 $\mu\mathrm{m}$ sein sollte. Und damit ergibt sich

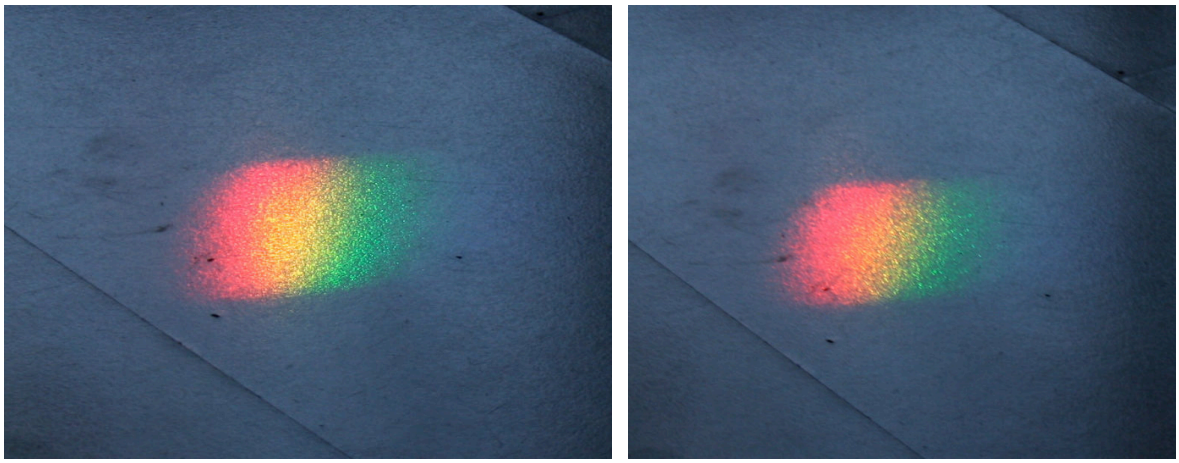
$$0.65 = 0.61\lambda/\mathrm{NA}, \tag{5.25}$$

also $\mathrm{NA} \approx 1$. Daher würde die Blendenzahl 0.5 betragen, ein leider nicht sonderlich realistischer Wert (theoretisches Maximum). Die geplante Mittelung kann so also nicht durchgeführt werden. Vielmehr muss zusätzlich per Software oder Binning (siehe Abschnitt 9.8.13) räumlich über mehrere Pixel gemittelt werden.

Neben der räumlichen Mittelung kann auch eine zeitliche Mittelung die Speckles effektiv reduzieren. Hierzu werden z.B. rotierende Diffusoren im Beleuchtungsstrahlengang eingesetzt. Damit wird erreicht, dass sich die Phasen auf dem Objekt zufällig im Zeitverlauf ändern und dementsprechend das sich ergebende Specklemuster zeitlich zufällig variiert und somit über die *Belichtungszeit* des Sensors mittelt. Neben mechanisch rotierenden Elementen können auch elektrisch bzw. elektronisch variierbare Elemente (z.B. Flüssigkristallmodulatoren) eingesetzt werden.

Bei geeigneten Bedingungen ergeben sich durchaus auch ohne die Verwendung kohärenter Lichtquellen (leichte) Speckles. Abb. 5.14.24 zeigt einen von spektral aufgespaltenem Sonnenlicht beleuchteten Laminatboden. Die Rauheit ist ausreichend gering, um bei der kurzen Kohärenzlänge des spektral aufgespalteten Lichts (einige zehn Mikrometer) dennoch Speckles auszubilden. Die räumliche Kohärenz (Ausdehnung der Sonne) ist ebenfalls hoch genug. Je größer die Lichtquellenausdehnung (Variation des Beleuchtungswinkels auf die abzubildenden Oberfläche), desto geringer sind Speckleeffekte.

Sobald feine Strukturen abgebildet werden, muss auch bei Beleuchtung mit einer LED und der damit einhergehenden begrenzten Kohärenzlänge mit deutlichen



(a) Blende 8 (b) Blende 16

Abbildung 5.14.24: Specklemuster auf einem Laminatboden bei verschiedenen Blendenzahlen. Die mittlere Specklegröße variiert direkt proportional zur Blendenzahl. Beleuchtung mit Sonnenlicht

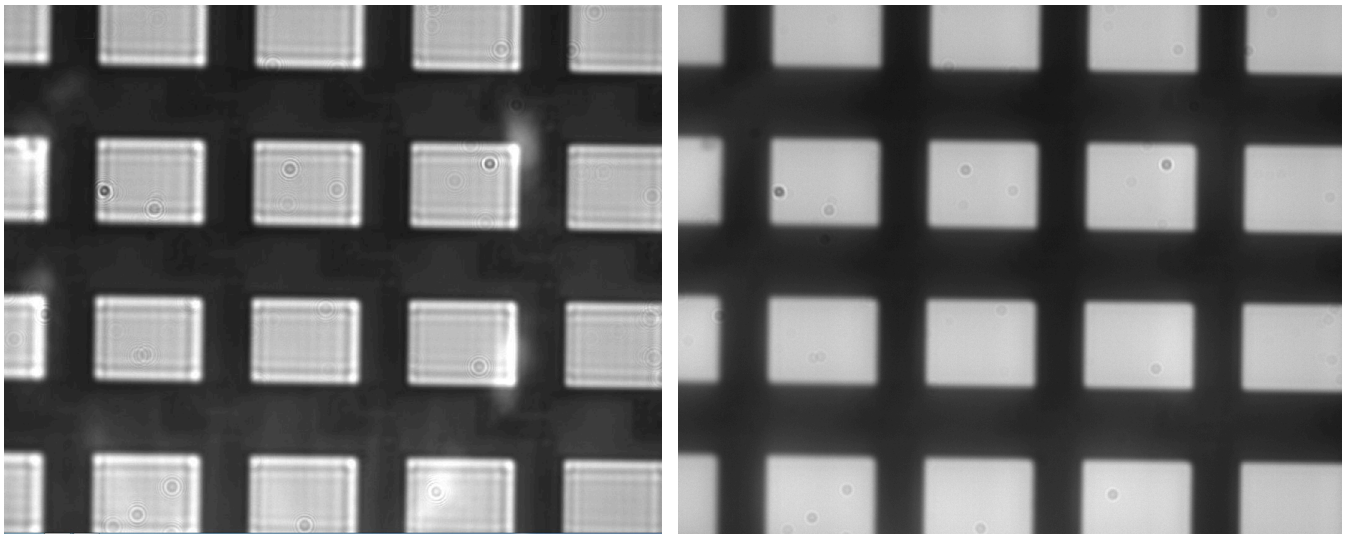
kohärenten Effekten gerechnet werden.

Abb. 5.14.27 zeigt die Abbildung eines Gitters mit der Periode 40 μm unter Beleuchtung einer roten LED. Der Abstand zwischen dem konventionellen LED Beleuchtungsmodul und Objekt beträgt ca. 200 mm und die effektive Ausdehnung der Lichtquelle ca. 3 mm. Es ergibt sich das in Abb. 5.14.27 (a) dargestellte Bild. Deutliche kohärente Überschwingereffekte an den Kanten werden sichtbar. Kleine Dreckpartikel auf der Oberfläche führen zu entsprechenden ringförmigen Beugungsringen mit vergleichsweise hohem Kontrast. Es handelt sich um eine teilkohärente Abbildung.

Abb. 5.14.27 (b) zeigt dieselbe Abbildung bei der zusätzlich in den Beleuchtungsstrahlengang in kurzer Entfernung zum Objekt ein Diffusor (Milchglasscheibe) eingebracht wurde. Die Kohärenz und damit die kohärenten Artefakte (Kantenüberschwinger und Ringstrukturen) ist deutlich reduziert. Das Bild ist verbessert.

Abb. 5.14.26 verdeutlicht den prinzipiellen Unterschied der beiden Abbildungen.

Wir gehen beispielhaft von 2 Atomen unserer Lichtquelle aus und betrachten jeweils einen von diesen Atomen ausgehenden Lichtstrahl, der durch einen bestimmten Punkt in der Objektebene läuft. Jedes der beiden Atome macht hier für sich betrachtet ein kohärentes Bild. Entscheidend ist hierfür zunächst die ausreichende zeitliche Kohärenz (s. Abb. 5.14.26 (b). Wir haben eine sphärische Welle,



(a) ohne Diffusor (b) mit Diffusor

Abbildung 5.14.25: Abbildung eines Gitters mit 40 μm Periode bei Beleuchtung mit einer roten LED.

die von dem Atom ausgeht und exemplarisch zeichnen wir zwei Lichtstrahlen, die dieser Welle entsprechen (Strahl senkrecht zur Wellenfront) ein. Die beiden Strahlen (natürlich gibt es in Wirklichkeit unendlich viele, aber wir Konzentrieren uns zur Erklärung auf zwei) treffen das Objekt an benachbarten Positionen. Dort wird das Licht gestreut (oder falls sich an der entsprechenden Stelle keine Struktur befindet eben nicht gestreut).

Z.B. mag der blaue Strahl gerade auf eine Kante fallen und so gestreut werden während der schwarze Strahl durch eine Transparenz des Gitters fällt und seine Richtung nicht ändert. Aufgrund der begrenzten Auflösung der Abbildung (Beugungsbegrenzung und/oder Aberrationen) werden die beiden benachbarten Punkte verschmiert bzw. unscharf abgebildet, d.h. ein Punkt in der Bildebene empfängt Licht von beiden Objektpunkten. Zur Interferenz der entsprechenden beiden Lichtanteile kommt es, wenn die unterschiedlichen Weglängen, kleiner als die Kohärenzlänge der Lichtquelle sind. Davon (kleine Strukturen, quasimonochromatische LED oder Laser Beleuchtung) wollen wir im folgenden ausgehen.

Was passiert nun, wenn ein weiteres, räumlich getrenntes Atom der Lichtquelle vorhanden ist? Auch dieses wird zu einer kohärenten Abbildung der Struktur führen. Das Licht der unterschiedlichen Atome interferiert hier nicht, denn die Atome stimmen sich nicht untereinander ab, d.h. sie emittieren unabhängig voneinander und im zeitlichen Mittel wird so zwischen den Atomen keine Kohärenz

zu beobachten sein.

Solange aber alle Abbildungen aufgrund der einzelnen Atome zum selben Intensitätsmuster führen, wird man dennoch die Interferenzen bzw. die kohärenten Abbildungscharakteristiken sehen. Bei N Atomen wird die Abbildung aufgrund des einzelnen Atoms einfach N mal heller sein.

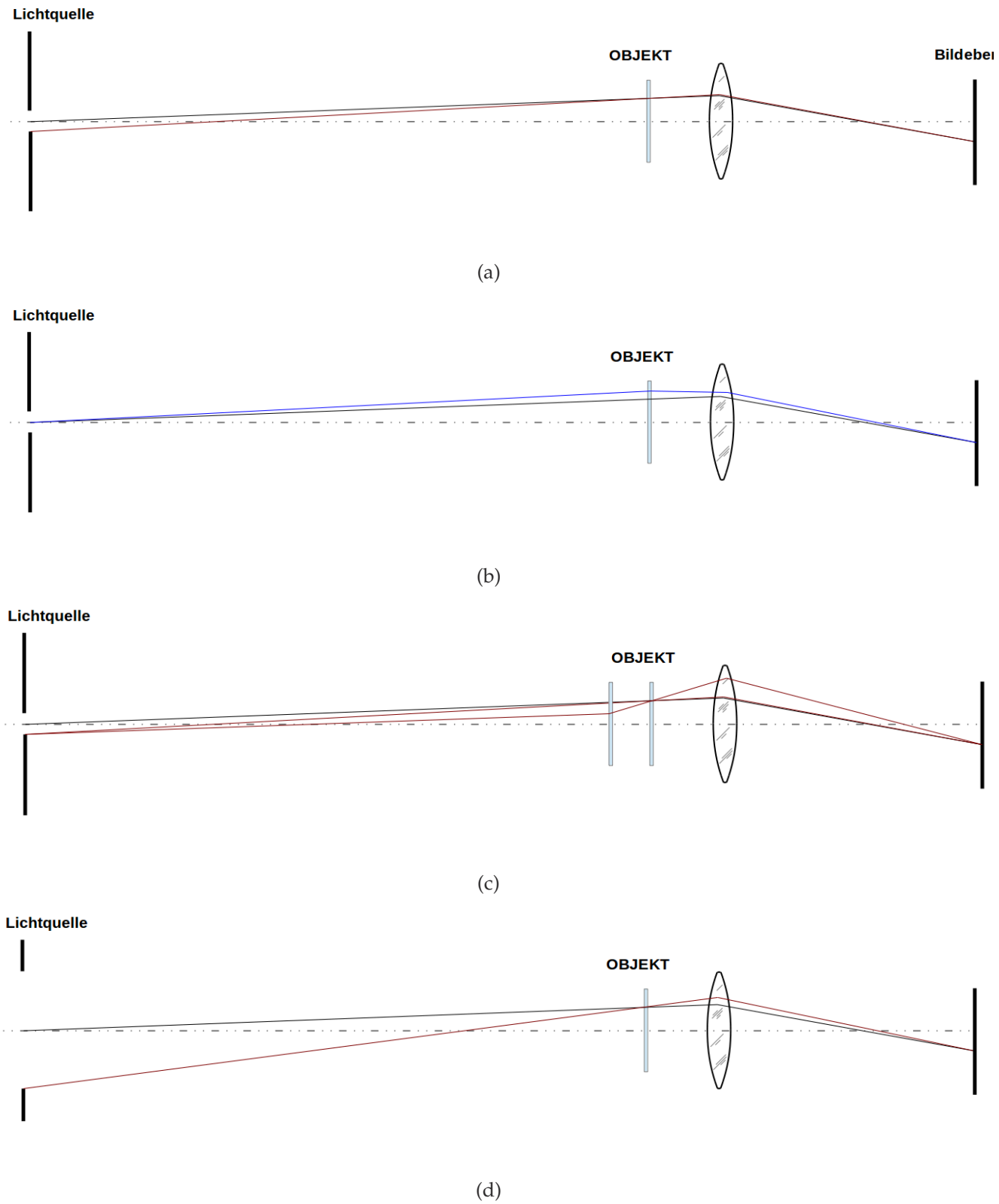


Abbildung 5.14.26: Zur Erläuterung der Kohärenzreduktion mit einem Diffusor

Bei Verwendung eines zusätzlichen Diffusors bzw. Streuers vor dem Objekt (Abb. 5.14.26 (c)) wird derselbe Objektpunkt vom selben Atom der Lichtquelle auf unterschiedlichen geometrischen Wegen beleuchtet (die beiden roten Strahlen). Aufgrund der Abbildung kommen diese unterschiedlichen Strahlen in der Bildebene wieder zusammen und können dort interferieren. Allerdings ergibt sich eine Interferenz wieder nur, wenn die Wegunterschiede der Strahlen kleiner als die Kohärenzlänge sind. Je näher der Diffusor beim Objekt liegt, desto eher wird dies nicht der Fall sein, denn die Wege, die das Licht läuft sind für die einzelnen Strahlen sehr unterschiedlich. Der Diffusor hat so die Kohärenz reduziert. Letztlich handelt es sich hier um eine Reduktion der räumlichen Kohärenz. Ganz entsprechend hätte man auch eine größere Grundausdehnung der Lichtquelle wählen können.

Das Beispiel zeigt, dass kohärente Artefakte durchaus bei scheinbar inkohärenten Abbildungen, also auch ohne Laser, auftreten können und dass es sich lohnen kann, die Beleuchtung zu hinterfragen. Zu beachten ist aber weiterhin, dass bei der Beleuchtung eines rauen Objekts der statische Diffusor oft durch einen rotierenden Diffusor ersetzt werden muss. Die kohärenten Artefakte (Speckles) ändern sich dann je nach Diffusorstellung und im zeitlichen Mittel heben sie sich gegenseitig weg und man erhält wieder eine inkohärente Abbildung.

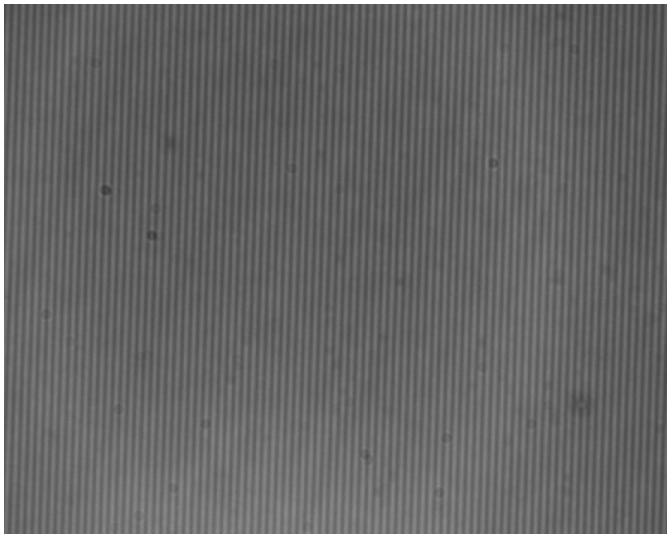
An dem Beispiel lässt sich auch die Auflösungsreduktion in Abhängigkeit von der Beleuchtung nochmals erläutern. Abb. 5.14.27 (a) zeigt die Abbildung eines Gitters der Periode 1.9 $\mu\mathrm{m}$ bei Verwendung des Diffusors. Ohne Diffusor (wird bei der verwendeten objektseitigen numerischen Apertur von von 0.266) keine Abbildung des Gitters erreicht. Dies ist auch nicht verwunderlich, denn die minimal auflösbare Gitterperiode beträgt gemäß Abbe für ein beugungsbegrenztes System $\lambda/\mathrm{NA} = 2.4\,\mu\mathrm{m}$.

Der Diffusor sorgt dafür, dass die effektive numerische Apertur der Beleuchtung deutlich vergrößert wird (s. Abb. 5.14.28). Schräg auf das Objekt einfallende Strahlen werden durch die Beugung am Gitter so beeinflusst, dass sie durch die Optik transmittiert werden. Das Gitter wird abgebildet. Das Beispiel zeigt sehr schön, dass die Auflösung einer Abbildung nicht nur von der Abbildung selbst, sondern bei feinen Strukturen auch wesentlich durch die Beleuchtung beeinflusst wird. Dementsprechend lässt sich die numerische Apertur der Beleuchtung in den meisten Mikroskopen variieren und als Daumenregel kann man die

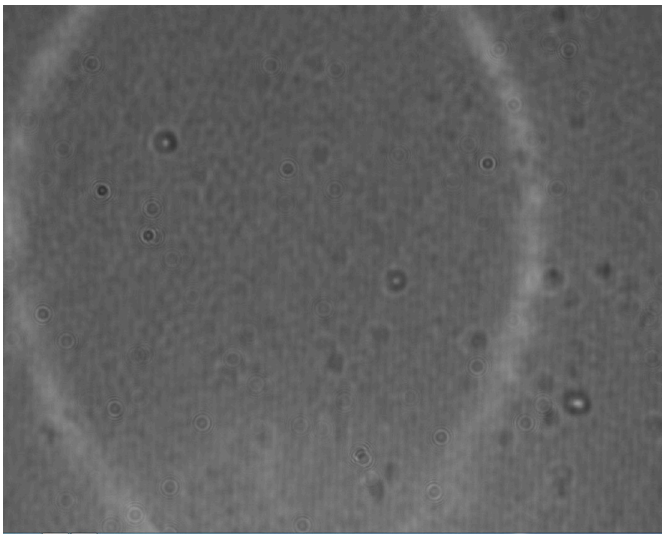
Orginalformel von Abbe, die eben für kollimiertes Beleuchtungslicht gilt, gemäß $d = \lambda/(\mathrm{NA}_{\mathrm{Bel.}} + \mathrm{NA}_{\mathrm{Abb.}})$ abwandeln.

Abb. 5.14.27 (c) zeigt, dass sich der Kontrast bei einer Halbierung der Beleuchtungs-NA (realisiert durch Verschieben des Diffusors) nochmals deutlich reduziert. Und die Kontrastreduktion (ebenfalls gegeben durch eine Auflösungsreduktion) in Abb. 5.14.27 (d) erfolgt weil die numerische Apertur auf 0.2 reduziert wurde.[22].

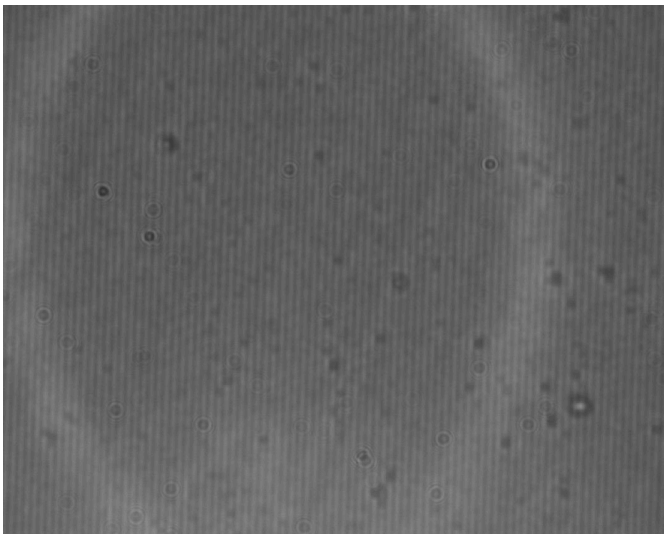
[22]Eine entsprechende Verringerung ergibt sich für das verwendete System im übrigen auch, wenn die numerische Apertur auf ca. 0.4 erhöht wird, denn in diesem Fall werden die Aberrationen zu stark)



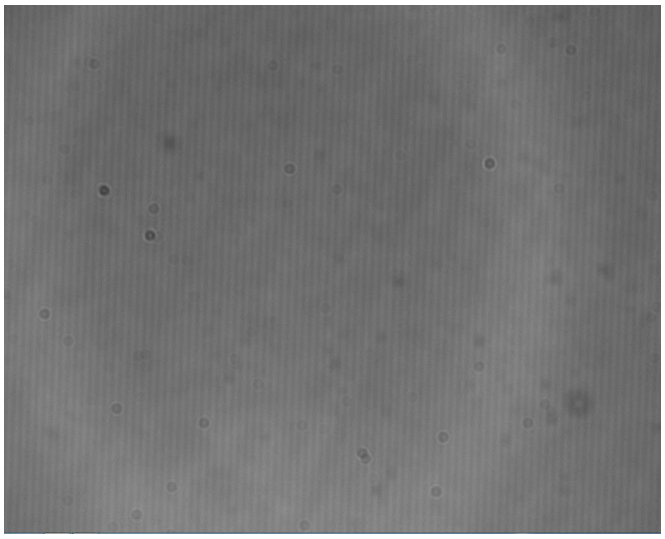
(a) Diffusor bei Objekt



(b) ohne Diffusor



(c) Diffusor zwischen Objekt und Quelle



(d) Diffusor bei Objekt aber verringerte numerische Apertur (0.2)

Abbildung 5.14.27: *Zur Erläuterung der Auflösungverbesserung durch einen Diffusor*

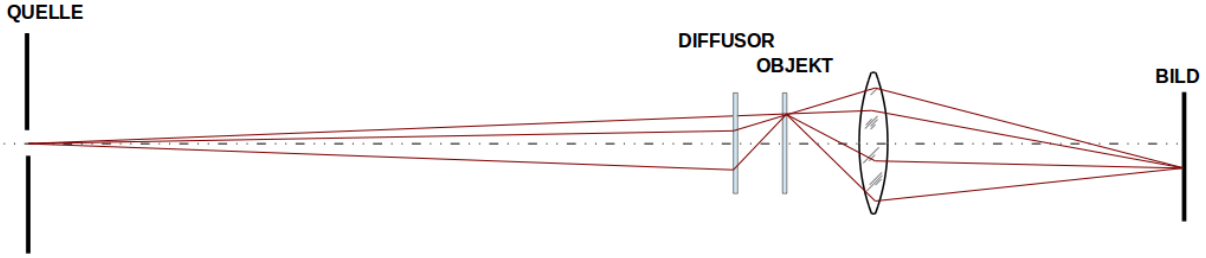


Abbildung 5.14.28: *Ein Diffusor führt zu einer erhöhten numerischen Apertur der Beleuchtung und kann so die Auflösung des Systems erhöhen.*

5.15 Advanced: Auflösungsreduktion bei Farbsensoren

Bei *Farbsensoren* ergeben sich teilweise sehr störende Farbartefakte, die letztlich ihre Ursache in der *Dispersion* des Objektivs haben. Unterschiedliche Wellenlängen werden durch die optischen Elemente leicht unterschiedlich abgelenkt. Dies führt unter anderem zu einer wellenlängenabhängigen Lage der Bildebene (longitudinale chromatische Aberration). Wenn auf den grünen Spektralbereich perfekt fokussiert wird dann sind sowohl der Rot- als auch der Blaukanal leicht defokussiert. Ein feines Detail oder eine Kante erhält dementsprechend einen violetten Halo. Der Effekt ist natürlich umso stärker, je offener die Blende ist, denn dadurch ergibt sich eine sehr geringe Schärfentiefe. Bei monochromatischen Sensoren ergeben sich natürlich im digitalen Bild keine Farbsäume, der Effekt führt aber bei einer Beleuchtung mit weißem Licht eben zu einer reduzierten Auflösung.

Die monochromatischen MTFs der einzelnen Spektralanteile haben durchaus einen deutlich unterschiedlichen Verlauf. Wenn nun die MTF im roten Spektralkanal bei mittleren und hohen Frequenzen im Vergleich zu Blau und Grün ausgeprägt hoch ist, dann ergibt sich zwangsläufig im Kantenbereich (das Ortsfrequenzspektrum der Kante enthält viele hohe Frequenzen) eine Verstärkung der Rottöne. Bei monochromatischen Sensoren führt dies „nur" zu einer Verringerung der Auflösung. Bei Farbanwendungen können entsprechende Farbsäume aber darüber hinaus erhebliche Probleme bereiten.

Zur Vermeidung entsprechender Störungen muss sichergestellt werden, dass die Information räumlich gemittelt bzw. verschmiert wird. Bei sehr kleinen Pixeln, wie sie in vielen Consumerprodukten zu finden sind reicht hierfür in aller Regel bereits die durch die Optik (Aberrationen und Beugungsbegrenzung) gegebene Unschärfe. Ansonsten kann aber auch ein optisches Tiefpassfilter direkt vor dem eigentlichen Sensor verwendet werden.

Die Auflösung leidet natürlich unter entsprechender Tiefpassfilterung, aber auch die Farbinterpolationsalgorithmik führt zu einer Verschlechterung der Auflösung. Farbsensoren sollten daher bei auflösungskritischen Aufgabenstellungen nur eingesetzt werden, wenn dies wirklich notwendig ist.

5.16 Advanced: Smartphones

Es ist erstaunlich, welche Bildqualitäten mit heutigen High-End Smartphones erreicht wird. Trotz kleiner Pixel und preiswerten Optiken werden oft sehr rauscharme Bilder mit hoher Auflösung erzielt.

Für die optische Abbildung werden sehr komplexe Optiken bestehend aus Asphären verwendet. Die Herstellung ist durch die hohen Stückzahlen nahezu perfekt optimiert. Sowohl die design- wie auch die herstellungsbedingten Aberrationen sind sehr gering.

Dennoch können auch die Hersteller nicht zaubern. Der Haupttrick liegt in der massiven Nachbearbeitung der Bilder. Bei nicht bewegten Szenen Szenen können und werden sogar mehrere Bilder zu einem verbesserten Bild kombiniert.

5.17 Positionsauflösung

Oftmals interessiert in der industriellen Bildverarbeitung nicht die Auflösung im Sinne einer Abbildung feinster Details sondern primäres Ziel ist die Messung geometrischer Größen, z.B. Durchmesser einer Bohrung oder Abstand von Kanten zueinander.

Eine sich ebenfalls oft ergebende Anwendung ist die Abbildung von *Kanten*. Abb. 5.17.29 zeigt die sich ergebende Intensitätsverteilung (wieder für eine numerische Apertur von 0.1 bei $\lambda = 500$ nm und mit der Vergleichsverteilung nach Airy). Offensichtlich ist die Kantenausdehnung[23] gering vergrößert gegenüber dem Airy Radius.

Dies wird bewerkstelligt, indem die Position verschiedener Merkmale mit möglichst geringer Messunsicherheit bestimmt wird. Wenn hier also von *Positionsauflösung* die Rede ist, dann ist das eigentlich meist der falsche Begriff, denn eigentlich ist man dann an der „Messunsicherheit der Positionsbestimmung" interessiert. Die Messunsicherheit wird dabei zum einen durch die Kalibrierung und die Rückführbarkeit auf eine Referenzmessung oder ein kalibriertes Objekt be-

[23] bei einer Definition der Kantenbreite über 10% und 90% der Intensität

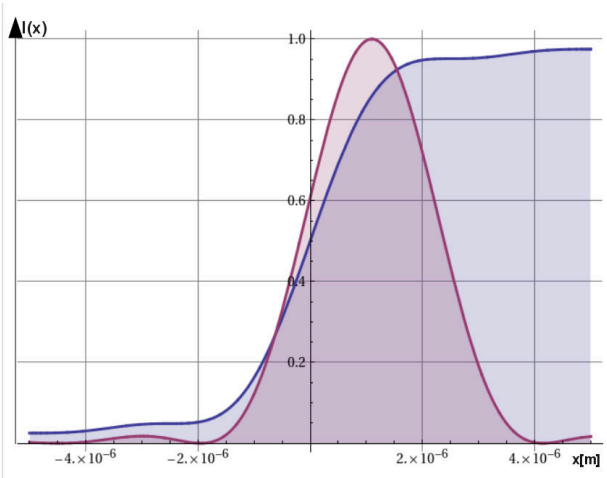


Abbildung 5.17.29: Intensitätsverteilung an einer Kante. Die Positionsauflösung, die sich bei Antastung einer Kante ergibt kann wesentlich höher als die optische Auflösung im Sinne einer Zweipunkttrennung sein. Die Kantenausdehnung im Bild t ist von derselben Größenordnung wie die Auflösung nach Rayleigh.

stimmt, zum anderen hängt sie aber auch von rein statistischen Fehlern ab. Diese Zusammenhänge können insgesamt sehr komplex werden. So kann z.B. in einer konkreten Anwendung die Messunsicherheit immer im Januar erhöht sein weil die Temperatur der Anlage dann leicht anders ist (thermische Ausdehnung), oder aber der Fremdlichtanteil führt mittags im Januar zu einer erhöhten Messunsicherheit.

Die Fehlerquellen für entsprechende, systematische Fehler sind vielfältig und wir werden im Kapitel 11 noch auf diese Problematik näher eingehen. Gehen wir aber davon aus, dass entsprechende systematische Fehler vermieden bzw. durch Kalibrierung eliminiert worden sind. Dennoch wird die Bestimmung der Position eines Merkmals (z.B. einer Kante) natürlich nicht beliebig genau möglich, sondern durch das Bildaufnahmesystems und die Algorithmik begrenzt sein.

Es ist wichtig, sich klar zu machen, dass die Messunsicherheit aufgrund der optischen Abbildung und der Bildaufnahme aber nicht gleich der Auflösung im Sinne der Trennung feiner Details ist. Wir können also nicht einfach die Auflösung nach Rayleigh oder die Pixelgröße verwenden, um die Positionsauflösung zu schätzen.

Vielmehr ergibt sich auf dem Bildsensor im Bereich der abgebildeten Kante ein Helligkeitsverlauf und basierend auf diesem Helligkeitsverlauf wird die Kantenposition geschätzt. Hierzu stehen verschiedene mehr oder weniger aufwändige Verfahren zur Verfügung, in der Praxis erreicht man unter günstigen Bedingun-

gen aber durchaus Positionsauflösungen von ca. 1/10 ... 1/50 Pixel im Sinne einer Standardabweichung von der wahren Position. Wieder ist hier der Begriff „Auflösung" allerdings mit Vorsicht zu betrachten bzw. er wird unterschiedlich verwendet. Die Auflösung im Sinne des kleinsten (reproduzierbar) detektierbaren Unterschieds kann durchaus noch deutlich unterhalb dieser Auflösung im Sinne einer Standardabweichung, also der statistischen Messunsicherheit, liegen. Es sollte also jeweils klar sein sein, wie der verwendete Wert genau definiert ist (vgl. Kapitel 12).

Sie müssen weiter beachten, dass oftmals verschiedene Positionsbestimmungen gemittelt ein Endergebnis ergeben. Wenn beispielsweise der Abstand zweier paralleler Kanten ermittelt werden soll dann kann entlang der Kante gemittelt werden und bei N *Mittelungen* kann der statistische Fehler um $\sqrt{N}$ reduziert werden. Somit lassen sich dann letztlich geometrische Parameter mit einer sehr viel höheren Genauigkeit ermitteln als man zunächst gemäß der lateralen Auflösung erwarten würde.

Als (sehr grobe) Abschätzung für die Messunsicherheit der Position, die sie letztlich bei einer gegebenen Pixelgröße p und einer gegebenen Anzahl von Mittlungen N erzielen, können Sie also bei optimalen Bedingungen (hohe Abtastrate des Bildes, geringe Aberrationen, sehr gute Kamerakalibrierung) von

$$\Delta X = 0.1...0.01\,\frac{p}{\sqrt{N}} \tag{5.26}$$

ausgehen. Dabei wurde angenommen, dass die Position lokal mit 1/10 bis 1/100 der Pixelauflösung bestimmt werden kann. Allerdings ist der in der Praxis erzielbare Wert stark von der konkreten Algorithmik und der Anwendung abhängig, so dass es sich lohnt, bei entsprechenden Aufgabenstellungen ausführlicher die Problematik zu beleuchten.

Für eine Charakterisierung können typische Kalibriertargets eingesetzt werden. Diese sind für kleine Objektfelder kommerziell in vielfältiger Form, z.B. als schachbrettartige Strukturen, Punktarrays oder konzentrische Kreise kommerziell verfügbar.

Wie bei der Auslegung der Abbildung für Detektionsaufgaben (vgl. Abschnitt 5.10) stellt sich auf für messende Aufgaben der Bildverarbeitung die Frage, wie konkret optische Auflösung und Sensorauflösung zu wählen sind.

Während bei den Detektionsaufgaben die optische Auflösung die zunächst maßgebliche Größe von der aus gerechnet wird, darstellt, ist diese optische Auflösung für Geometrie-messende Aufgabenstellungen weniger relevant. Dort werden letztlich Positionen von Kanten oder isolierten Strukturen hochgenau bestimmt. Der Kantenverlauf im Bild kann und soll dabei durchaus über viele (Faustregel: ca. 7 bis 15 Pixel) verlaufen.[24]

Abb. 5.17.30 zeigt die durchaus komplexe Vorgehensweise bei der Auslegung. Andere Vorgehensweisen sind möglich und je nach Aufgabenstellung kann auch ein modifizierter Ansatz sinnvoller sein. Die dargestellte Grundvorgehensweise kann in drei Hauptbereich unterteilt werden: 1) Auflösung basierend auf Pixeln, 2) Auflösung der Optik, 3) Blendenzahl. Der Ablauf muss mit verschiedenen potentiellen Sensoren durchgespielt werden.

[24] Zu wenige Pixel sind sogar ungünstig, denn letztlich wird in den Grauwertverlauf im Bild eine modellhafte Kante angenähert.

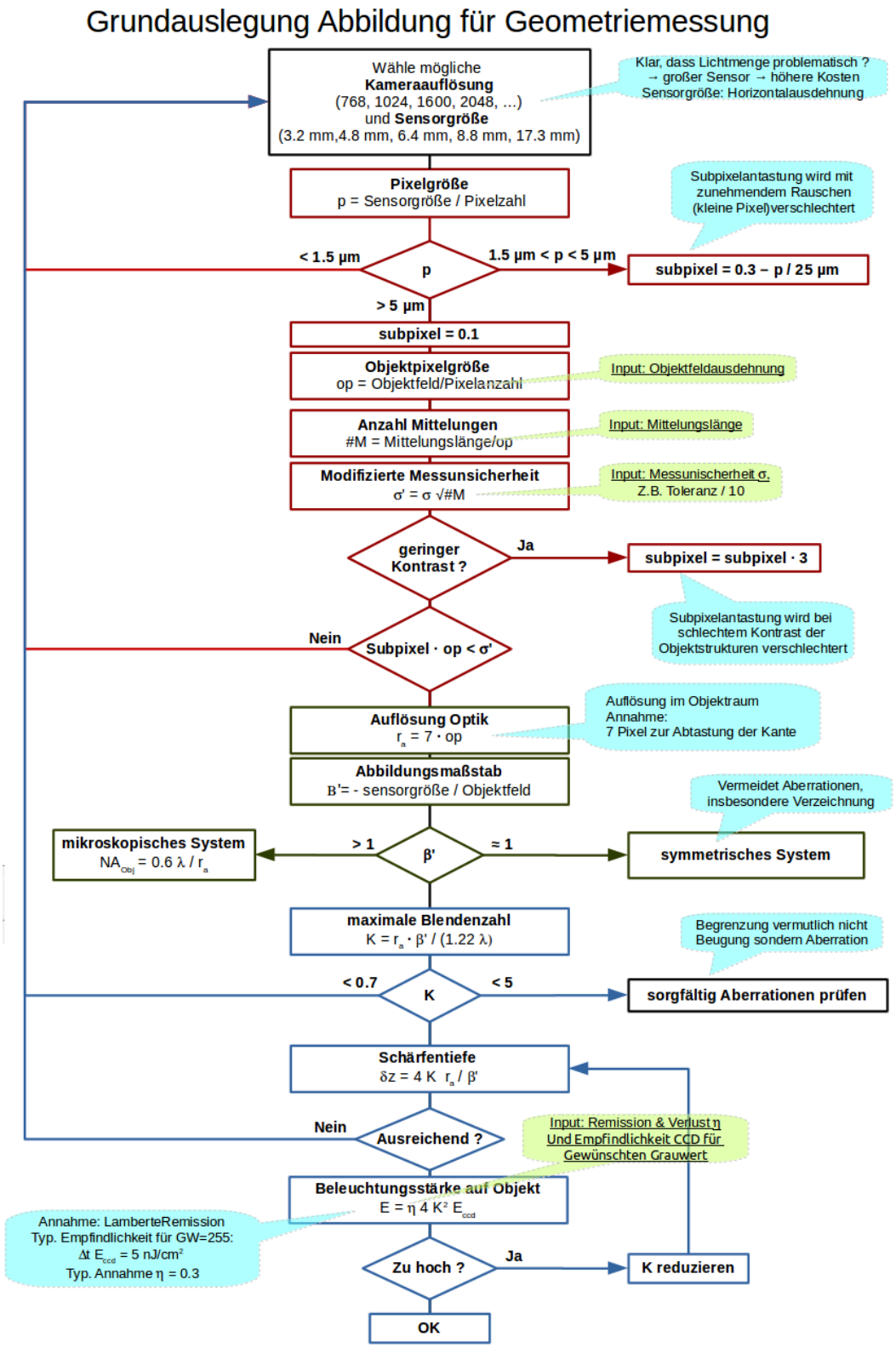


Abbildung 5.17.30: Algorithmus zur Auslegung der Abbildung für 2D-messende Bildverarbeitungslösungen.

Wir verdeutlichen den Algorithmus anhand eines einfachen Beispiels: Das Testobjekt habe eine Ausdehnung von 50 mm und die Breiten von 2 mm langen Strukturen sollen ermittelt werden. Die Toleranz soll 50 μm betragen.

Wir gehen von einem Standardsensor (1/2 Zoll, 1600 x 1200 Pixel) aus und berechnen damit die Pixelgröße zu 5 μm und den Subpixelfaktor zu 0.14. Wir gehen also davon aus, dass an einzelnen Positionen die Kanten auf $0.14 \cdot 5\ \mu\mathrm{m} \approx 0.7\ \mu\mathrm{m}$ genau bestimmt werden können. Die Pixelgröße auf dem Objekt beträgt p_o =31 μm und die Strukturlänge entspricht damit 64 Pixeln. Als Messunsicherheit fordern wird optimistisch bei der Toleranz von 50 μm einen Wert von 7 μm . Aufgrund der Mittelung reicht eine Messunsicherheit pro Position von $7 \cdot \sqrt{64} = 56\ \mu\mathrm{m}$.

Wegen $0.14 \cdot 31\mu\mathrm{m} < 56\mu\mathrm{m}$ scheint die Aufgabe gut lösbar.

Wir fahren mit der optischen Auflösung fort: Es ergibt sich für die Auflösung nach Rayleigh im Objektraum $r_A \approx 7 \cdot p_o \approx 200\ \mu\mathrm{m}$ und der Abbildungsmaßstab $\beta' = -6.4/50 = -0.128$. Damit wird die maximale Blendenzahl $K_{max} = 35$. Es ist bereits hier klar, dass eine kleinere Blendenzahl sinnvoller ist (Lichteffizienz). Wir gehen also zu K=10 über, den hierbei kann noch von einer guten, beugungsbegrenzten Abbildungsqualität ausgegangen werden. Damit wird die Schärfentiefe 62 mm.

5.18 Herstellerangaben

Die Herstellerangaben in den „technischen Daten" zu Abbildungssystemen helfen leider oftmals kaum, die Bildqualität ausreichend gut zu beurteilen. Oftmals erfolgt eine pauschalierte Angabe der Art „5-Megapixel Objektiv". Eine entsprechende Angabe ohne detaillierte Nennung der Einsatzbedingungen und der Art der Definition ist lediglich als Werbeaussage zu verstehen. Die gleiche Optik kann — je nach Definition der MPixel gut und gern zwischen 2 und 30 MPixel variieren. Eine einfache Möglichkeit ist es, z.B. statt Werte basierend auf einem Kontrast von 30% Werte für 10% anzugeben. Selbst wenn MTF-Kurven zur Verfügung gestellt werden ist zu hinterfragen, ob es sich um rein designbasierte Kurven oder aber um echte Messungen handelt. Bei designbasierten Kurven werden teilweise Beugung und Beschichtung nicht berücksichtigt („geometrische MTF") so dass insbesondere bei höheren Blendenzahlen das Datenmaterial mehr oder weniger nutzlos wird

und die Fertigungstoleranzen sind natürlich ebenfalls nicht enthalten.

In aller Regel geben die Objektivhersteller leider aber überhaupt keine sinnvollen Kennwerte an und — wenn überhaupt — kann man nur per Rückfrage an die MTF Kurven gelangen. Diese Kurven variieren massiv mit den Einsatzbedingungen für das Objektiv. Es ist also darauf zu achten, dass Blendenzahl, Arbeitsabstand, Schärfentiefe, Wellenlängen von MTF und geplantem Einsatz möglichst exakt übereinstimmen. Ansonsten kann leicht eine erhebliche (Faktor 2...10) Variation der Spezifikation auftreten. So ist es denkbar, dass der Bildsensor sich unnötig verteuert und die Bildwiederholfrequenz drastisch und unnötig eingeschränkt wird (Transfer von unnötig vielen Pixeln). Teilweise, z.B. bei der Schärfentiefe, sind sehr unterschiedliche Definitionen gebräuchlich, so dass hier genau überdacht werden muss, ob Werte wirklich vergleichbar sind.

Bei allen Überlegungen zum Kontrast von Strukturen ist zu beachten, dass der sich letztlich im digitalen Bild ergebende Kontrast eine multiplikative Kombination des Kontrast aufgrund des idealen Bildes der Optik-MTF und der Sensor-MTF ist. Elektronische Nachbearbeitung wirkt sich natürlich ebenfalls auf den Kontrast dramatisch aus, ein Vergleich ist dann aber nur bei gleichem Signal-Rausch-Verhältnis sinnvoll. Bei aggressiver Kontrastverstärkung verstärkt sich eben in der Regel auch (bildabhängig) das Rauschen.

Bei MTF Kurven ist zu beachten, dass der Kontrast bei hohen Ortsfrequenzen stark von der Orientierung des abzubildenden Gitters abhängt. Es kann so durchaus sein, dass der Kontrast verschiedener Gitterorientierungen um den Faktor zwei variiert bzw. dass bei ungünstiger Gitterorientierung der Kontrast deutlich unter die an sich veranschlagte 30% Marke fällt. Weiterhin muss bedacht werden, dass die Ergebnisse für eine perfekte Fokussierung angegeben werden. Die MTF ändert sich auch mit leichtem Defokus sehr drastisch im Bereich hoher Ortsfrequenzen. Da in den meisten Aufgabenstellungen ausgedehnte (tiefe) Objekte abgebildet werden und darüber hinaus oftmals auch Lagetoleranzen durch die Zuführung der Objekte gegeben sind ist unter Umständen die durch den Defokus induzierte Reduktion der Auflösung entscheidend.

Kapitel 6

Objektive

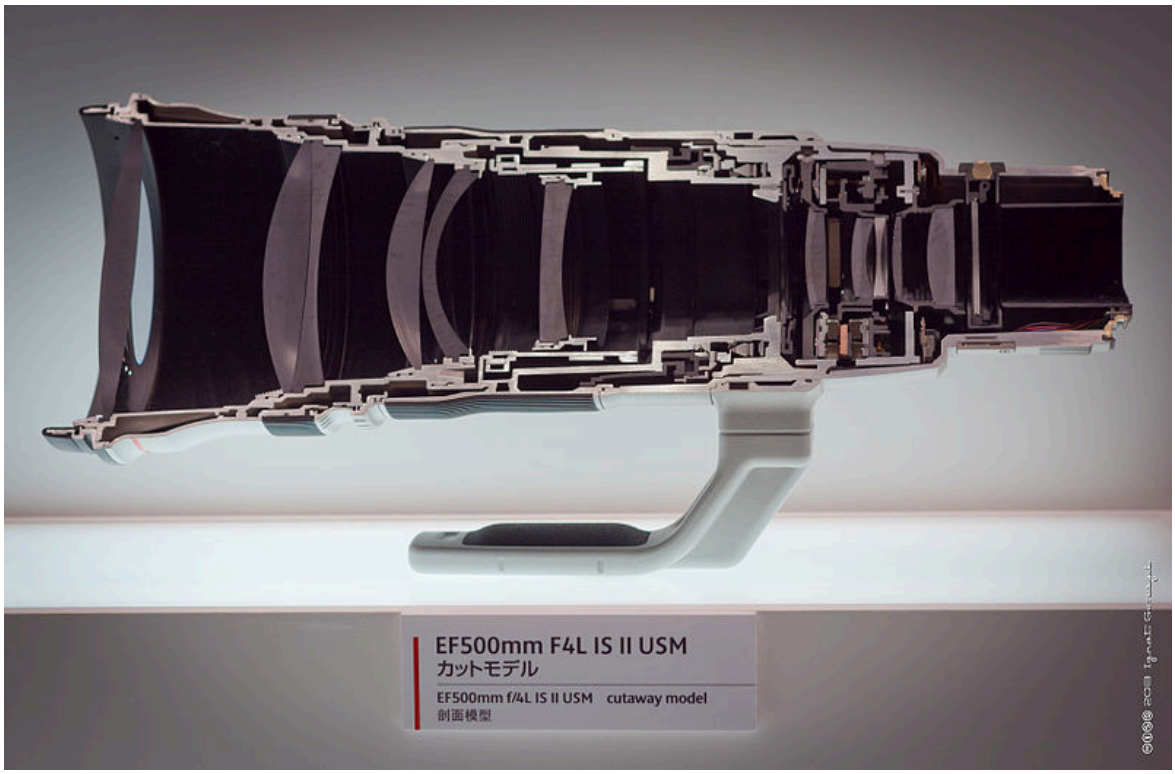


Abbildung 6.0.1: Bild mit freundlicher Genehmigung von Ignat Gorazd unter der Creative Commons Attribution-Share Alike 2.0 Generic license.

In diesem Kapitel werden Sie unter anderem lernen

- welche unterschiedlichen Objektivtypen am Markt verfügbar sind,
- wie die Abbildungsleistung der Objektive spezifiziert wird,

- was bei der Auswahl zu beachten ist,
- welche Filter zum Einsatz kommen können und
- was hinsichtlich Pflege und Reinigung zu beachten ist.

Objektive gibt es in einer Vielzahl von Ausführungen, mit ganz unterschiedlichen Bauformen und zu unterschiedlichen Preisen. Die Auswahl des für eine Anwendung passenden Objektivs ist alles andere als einfach. Wesentlich ist, die Abbildungsaufgabe zunächst exakt zu definieren bzw. die verschiedenen möglichen Varianten zu betrachten.

Basierend auf möglichen Anwendungsgeometrien und Sensorgrößen berechnen Sie — wie bereits besprochen — die gewünschte Brennweite und Blendenzahl. Damit kommen Sie dann zunächst zu den wesentlichen Grundparametern, die es bei einer Auswahl zu betrachten gilt:

- Brennweite
- Blendenzahl
- Arbeitsabstand
- Sensorgröße
- Wellenlängenbereich / Farbe
- Perspektive (Telezentrisch oder Nicht-telezentrisch)
- mechanischer Anschluss

Basierend auf den Überlegegungen in Kapitel 2 wird das Objektiv zunächst durch Brennweite und Blendenzahl definiert. Aber das reale Objektiv muss hinsichtlich seiner Abbildungsleistung optimiert werden und diese Optimierung orientiert sich an einem speziellen Arbeitsabstand und für ein vorgegebenes Bildfeld. Sie können daher z.B. nicht erwarten mit einem 50/1.8 Objektiv, das für einen 1/3 Zoll Sensor berechnet wurde, gute Ergebnisse für einen 2/3 Zoll Sensor zu erhalten.

Grundsätzlich stehen uns verschiedene Varianten zur Realisierung einer Abbildungsoptik zur Verfügung:

Standard-BV Objektiv: Diese Variante ist sicher in den meisten Anwendungen das Mittel der Wahl. Die Objektive sind etwas teurer als konventionelle Fotoobjektive, aber hinsichtlich ihrer Eigenschaften auf BV Anwendungen ausgelegt. Insbesondere sind typische BV-Anschlußgewinde vorgegeben und in der Regel lassen sich Einstellungen wie Fokus, Blende oder Zoom feststellen, so dass eine ungewollte Verstellung während des Betriebs (z.B. durch Vibrationen) nicht erfolgt.

Während im Fotobereich meist Zoomobjektive verwendet werden, ist in der industriellen Bildverarbeitung das Fix-Fokus Objektiv (also ein Objektiv mit einer festen Brennweite) Standard. Hierdurch kann es erst gar nicht passieren, dass sich die Brennweite und damit der Abbildungsmaßstab verändert. Teilweise werden *Zoomobjektive* auch als *Varifocal Objektive* bezeichnet.

Fotoobjektive: Zeichnen sich aufgrund der hohen Stückzahl durch ein sehr gutes Preis-Leistungsverhältnis aus, d.h. eine gute bis sehr gute Abbildungsleistung wird preisgünstig erreicht. In der Regel sind die Objektive allerdings für eher große (oft unendlich) Arbeitsabstände optimiert (Ausnahme: Makroobjektive). Anschluss, manuelle Einstellung und Feststellung können problematisch sein.

Kombination zweier Fotoobjektive: Die Kombination zweier auf unendlich korrigierter (Foto-)Objektive im Sinnes eines Keplerteleskops erzielt eine große numerische Apertur und die Möglichkeit der telezentrischen Perspektive bei guter Abbildungsqualität. Nachteilhaft ist die recht große Baulänge der Kombination (vgl. Abschnitt 2.2) und unter Umständen der geringe Arbeitsabstand. Weiterhin ist diese Art der Abbildung nur für vergleichsweise kleine Objektfelder (Bildfeld der Fotoobjektivs) möglich.

Einzellinsen: Achromate (=Doublets) sind vergleichsweise günstig erhältlich und lassen sich bei begrenztem Feld und begrenzter Apertur auch sinnvoll für einfache Abbildungsaufgaben einsetzen. Eine Kombination zweier Achromate ist ganz analog zur Kombination der Fotoobjektive möglich und wird auch bereits gefasst als „Achromatic Pair" kommerziell angeboten. Gegenüber ein-

zelnen Achromaten werden bei geringen Kosten auch sogenannte Triplets (3 zusammengekittete Elemente) und Elemente mit einer asphärischen Fläche eingesetzt.

Relay-Linsen: Für 1:1 Abbildungen wird üblicherweise ein voll symmetrischer Aufbau der Optik angestrebt. Dadurch werden insbesondere Verzeichnungen automatisch eliminiert. Entsprechende Relay-Linsen werden kommerziell angeboten.

Mikroskopische Abbildungen: Mikroskopische Abbildungen, die zu einer Vergrößerung[1] führen werden durch die Kombination eines speziellen Mikroskopobjektivs und einer Tubuslinse (meist genügt ein Achromat) realisiert.

Spezialanfertigungen: Spezialanfertigungen sind in der Regel dann notwendig, wenn eine unübliche Abbildungsgeometrie vorliegt oder durch ein Medium abgebildet werden muss. Auch ungewöhnliche Wellenlängenbereich (z.B. UV) machen oft ein Spezialdesign notwendig.

In aller Regel werden die meisten Anwendungen also mit Standard-BV Objektiven oder Standard-Mikroskopobjektiven realisiert und wir werden uns vorrangig die Spezifikation dieser Objektive anschauen. Bevor wir detaillierter auf die einzelnen Objektivtypen eingehen, werden wir zunächst die wesentlichen Parameter der Spezifikation besprechen.

Für die Wahl des Objektivs ist es wesentlich, dass für alle diese Parameter klar ist, welchen Wertebereich das Objektiv dabei jeweils abdecken muss.

6.1 Weitere Beurteilungskriterien

Neben den im letzten Abschnitt bereits dargestellten Grundparametern sind eine Reihe weiterer Parameter essentiell bei der Objektivwahl.

Auflösung: Natürlich ist Auflösung im Sinne der gerade noch auflösbaren feinsten Strukturen („limiting resolution") ein für sehr viele Aufgaben immens wichtiger

[1] oder auch Abbildungsmaßstab nahe ± 1

Parameter. Die Spezifikation sollte hierbei an die konkrete Aufgabenstellung angepasst sein. Beachten Sie weiterhin, dass — wie im vorangegangenen Kapitel dargestellt — die sich letztlich ergebende Gesamtauflösung nur teilweise von der Optik abhängt. Bedenken Sie auch die Schärfentiefe und die Kontrastverminderung durch den konkret geplanten Bildaufnehmer. Eng verbunden mit der Auflösung (je nach Spezifikation der Auflösung) ist natürlich der Kontrast der Abbildung. Nicht nur Aberrationen und die Beugungsbegrenzung sondern auch Streulicht kann zu einer Kontrastverminderung führen, so dass auch bei der Spezifikation der Auflösung über Bildpunktradien dennoch die MTF beachtet werden sollte.

Letztlich interessiert auch der erzielbare Kontrast bei weniger feinen Strukturen. Bei der visuellen Betrachtung des Bildes ist er oftmals der maßgebliche Parameter zur subjektiven Beurteilung. Für einen hohen Kontrast ist wichtig, dass die MTF insgesamt einen möglichst hohen Wert hat. Das System sollte vor allem bei niedrigen und mittleren Frequenzen gute Werte erzielen.

In den meisten Fällen wird die Auflösung bei BV-Objektiven mittels der MTF spezifiziert. Beachten Sie aber unbedingt, dass die MTF Kurven oder Grenzfrequenzen bei den von Ihnen spezifizierten Bedingungen gelten bzw. bei verschiedenen Objektiven auch dahingehend vergleichbar sind. MTF-Kurven für unterschiedliche Blendenzahlen, Feldwinkel oder Wellenlängen sind schlicht nicht miteinander vergleichbar. Sie müssen ebenfalls darauf achten, ob die dargestellten MTF-Kurven wirklich Messkurven sind. Einige Hersteller sparen sich die MTF Messung (oder sind mit den Ergebnissen nicht zufrieden) und geben stattdessen die theoretischen, aus dem Design stammenden MTF-Kurven an. Es ist klar, dass dies kaum eine Aussagekraft hinsichtlich des wirklichen Verhaltens des realen Objektivs hat. Und selbst wenn es sich um an sich vergleichbare Messkurven handelt bestehen oftmals erhebliche Unterschiede wenn unterschiedliche Messmethodiken angewendet werden. Im Zweifelsfall sollten Objektive geliehen werden und dann die Auflösung mittels eines einfachen Binärgitters (besser: Binärgitter mit zunehmender Periode) geprüft werden.[2]

Bei vielen Produkten ist nur ein einzelner Wert (z.B. „Auflösung") ohne Angabe der Einsatzbedingungen angegeben. Diese Angabe ist in aller Regel nicht nutzbar. Teilweise finden sich auch lediglich Angaben der Art „geeignet für 5 MPixel".

[2] Beachten Sie: der Kontrast für binäre Strukturen ist gegenüber Sinusstreifen leicht angehoben.

Auch dies eine Angabe, die schlicht nichts aussagt. Seriöse Hersteller hochqualitativer Optiken (z.B. Schneider-Kreuznach) stellen die notwendigen Angaben zur Verfügung oder können zumindest auf Nachfrage die relevanten Daten liefern.

Lichtstärke (Transmission und Vignettierung): Die Lichtstärke ist zunächst direkt proportional zu $1/K^2$. Aber auch die Transmission des Objektivs und seine Vignettierung können entscheidend sein.[3] Bei Mikroskopobjektiven kann hier durchaus die Hälfte des Lichts durch Reflexionen an den Teillinsen verloren gehen. Die Vignettierung kann zwar durch Beleuchtung oder Software ausgeglichen werden, aber dies geht immer mit einem erhöhten Aufwand und einer Verringerung des Signal-Rausch Verhältnisses einher. Es sollte also überdacht werden, welches Maß hier noch akzeptiert wird. Andererseits sollte bedacht werden, dass auch ein perfektes entozentrisches Objektiv in fast allen Fällen der sogenannten natürlichen Vignettierung unterliegt und dementsprechend einen $\cos^4$ Abfall der Bestrahlungsstärke mit dem Bildfeldwinkel nach sich zieht (vgl. Abschnitt 2.3.12).[4]

Verzeichnung: *Verzeichnungen* können im Prinzip durch Kalibrierung und nachträglichen Ausgleich per Software reduziert werden. Je stärker die notwendige Korrektur aber wird, desto mehr Aufwand ist hierfür zu fordern. Für messende Aufgaben reduziert man die Probleme daher gerne, indem man verzeichnungskorrigierte Objektive verwendet.

Für gewöhnliche visuelle Betrachtungen sind Verzeichnungen von einigen Prozent unproblematisch (Ausnahme: Szenen mit langen parallelen Strukturen). In der Bildverarbeitung sind Verzeichnungen meist deutlich problematischer, insbesondere wenn geometrische Größen, z.B. Abstände oder Durchmesser, gemessen werden sollen. Hierbei sind oftmals Abweichungen von Subpixeln also Zehntel oder Hundertstel Promille des Bildfeldes von Bedeutung und kleinste Verzeichnungen wirken sich störend aus.

In aller Regel wird man für solche Anforderungen eine Korrektur der Verzeichnung bzw. Kalibrierung des Felds vornehmen müssen. Wenn die Verzeichnung aber vergleichsweise klein ist (Promille-Bereich) dann ist es ausreichend, eine ein-

[3] Selten wird daher eine Angabe, die sowohl Transmission als auch Blendenzahl beinhaltet, die sogenannte t-number bzw. t/stops angegeben.

[4] Für telezentrische Objektive ist die Vignettierung deutlich reduziert.

fache lineare Korrektur vorzunehmen. Dementsprechend sollte man nicht denken, dass wenn sowieso eine Korrektur oder Kalibrierung notwendig ist, dann jegliche Verzeichnungsparameter egal sind.

Eine absolut verzeichnungsfreie Abbildung ergibt sich (automatisch) bei einem komplett symmetrischen Abbildungssystem (Abb. 6.1.2). Allerdings ist zu beachten, dass eine perfekte Symmetrie eben auch nur gelingt, wenn es sich um eine 1:1 Abbildung mit gleichem Objektabstand und Bildabstand handelt. Eine weitere Möglichkeit ist die Lochkamera, die aber außer für extreme Spezialanwendungen aufgrund der prinzipbedingten Unschärfe und extremen Lichtschwäche kaum einsetzbar ist. Sehr gute Verzeichnungseigenschaften haben üblicherweise auch telezentrische Bildverarbeitungsobjektive da typischerweise Benutzer, die Telezentrie benötigen eine gute Verzeichnungskorrektur zu schätzen wissen (meist handelt es sich um messende Aufgabenstellungen).

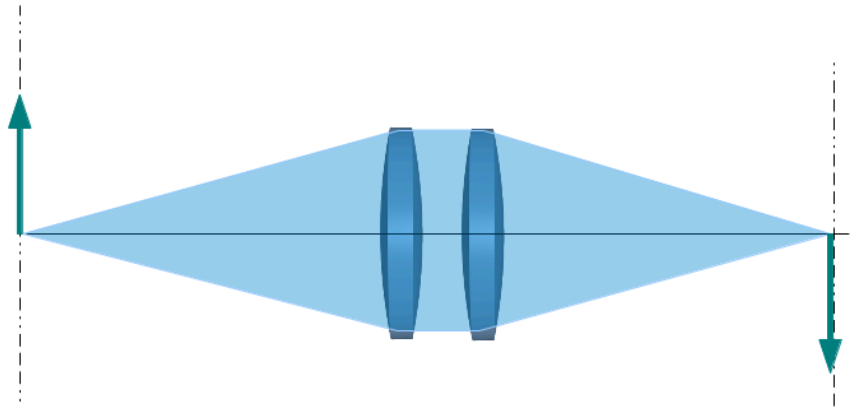


Abbildung 6.1.2: Symmetrische Systeme vermeiden jegliche Verzeichnungen

Streulichtempfindlichkeit und Beschichtung: Streulicht kann einerseits bei der Abbildung selbst durch ungewollte Reflexionen und Streuungen im Innern des Objektivs auftreten. Anderseits kann ein Objektiv empfindlich oder unempfindlich gegen das schräge Eindringen von zusätzlichem Fremdlicht/Streulicht sein. Eine Schwärzung der Linsenränder und der zusätzliche Einbau von Blenden führt zu einer generellen Verbesserung des Kontrasts bzw. Verminderung des Streulichts. Teilweise kann die Empfindlichkeit durch eine aufgesetzte Kappe (lens hood) reduziert werden.

Wesentlich für die Qualität des Bildes ist vor allem auch die sogenannte *Vergütung*,

also die Entspiegelung der optischen Gläser. An jedem brechenden Übergang von einem Glas zu Luft oder zu einem anderen Glas ergeben sich zunächst ungewollte Reflexionen, die natürlich möglichst gut unterdrückt werden sollten.

Wenn man von einem Objektiv aus 5 Linsen mit nicht entspiegelten Gläsern ausgeht und an jeder der $2 \cdot 5 = 10$ Grenzschichten eine typische 4% Reflexion annimmt, dann bedeutet das, dass lediglich $0.96^{10} = 66\%$ des Lichts korrekt abgelenkt werden. Schlimmer als der Lichtverlust ist allerdings, dass ein Teil des restlichen Lichts (hier 34%) nicht gemäß der gewünschten Abbildung den Bildsensor erreicht und so a) den Kontrast reduziert, b) Geisterbilder erzeugt, c) als Reflex sichtbar wird.

Eine gute Entspiegelung ist also für eine hochwertige Bildqualität entscheidend. Die Antireflexbeschichtungen müssen zur Wellenlänge des genutzten Lichts passen. Hierdurch lassen sich bei breitbandigen Beschichtungen die üblichen 4% Reflexionen pro Fläche bei senkrechtem Lichteinfall auf ca. 0.5 % reduzieren. Eine weitere Reduktion ist möglich, wenn der Wellenlängenbereich eingeschränkt wird. Die Grundidee der Beschichtung ist die Interferenz des an der Vorderseite der Beschichtung und der Hinterseite der Beschichtung reflektierten Lichts, so dass sich destruktive Interferenz ergibt (Abb. 6.1.3).

Die Beurteilung kann teilweise anhand der MTF erfolgen, am besten ist aber die direkte Betrachtung von Testbildern. Nachträglich kann das Störlichtverhalten durch Blenden vor dem Objektiv und durch (die richtige) monochromatische Beleuchtung oder einen Farbfilter verbessert werden. Die Beschichtungen sind in jedem Fall farbabhängig, daher sind die Beschichtungen passend zur geplanten Lichtquelle (und etwaigem Umgebungslicht) zu wählen.

Auch die Reinigung von Optiken kann in Bezug auf Beschichtungen zu Problemen führen. Unter Umständen macht es daher Sinn, vor dem Objektiv ein (ebenfalls entspiegeltes) leicht auszutauschendes Deckglas zu verwenden[5]

Ebenheit des Felds bzw. Bildfeldwölbung/Objektfeldwölbung: Im Prinzip sind wir bei unseren Überlegungen immer von ebenen Objekt- und Bildfeldern ausgegangen. Der Schärfentiefenbereich gibt somit ein Messvolumen im Objektraum

[5]Je dünner das Glas und je geringer die Blendenzahl der Optik, desto geringer die dadurch induzierten Aberrationen.

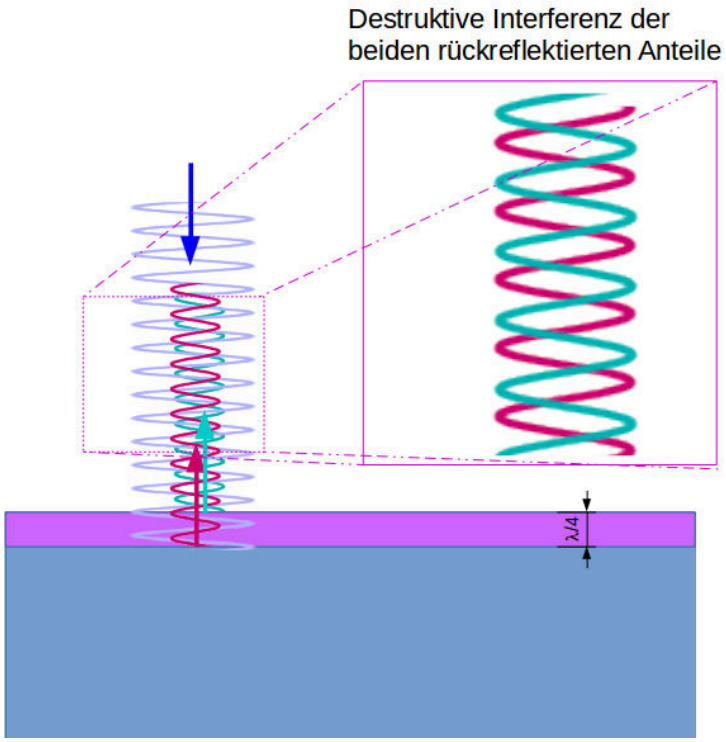


Abbildung 6.1.3: Prinzip einer Antireflexbeschichtung.

um diese Ebene an. Oftmals ist das Objektfeld aber gewölbt. Dies kann zu Problemen führen und muss ggf. beachtet werden.

Mikrolinsen: Der Bildfeldwinkel, unter dem Strahlen auf den eigentlichen Sensor treffen hat insbesondere bei Sensoren mit Mikrolinsen einen deutlichen Einfluss auf das Übersprechen benachbarter Pixel. Insbesondere bei Sensoren mit einer Farbmaske kann dies zu unschönen Bildstörungen führen. Ein Objektiv, das für einen Bildsensor ohne Mikrolinsen optimiert wurde und dort hervorragende Ergebnisse erzielt, kann bei einem Einsatz mit Mikrolinsen deutlich schlechtere Ergebnisse liefern.

6.2 Streulicht, Schleier und Geisterbilder

Neben den bereits besprochenen Bildfehlern können sich weitere deutliche Beeinträchtigungen der Bildqualität durch Streueffekte und ungewünschte Reflexionen ergeben.

Insbesondere sehr heller Lichteinfall auf das Objektiv kann zu deutlichen Störungen führen. Seitlicher Einfall von unerwünschtem Licht, das an sich nicht zur Abbildung genutzt wird kann teilweise durch Hauben bzw. *Gegenlichtblenden* („Sonnenschutzblenden"), die auf das Objektiv gesteckt werden, reduziert werden.

Je nachdem, ob ein lokaler Lichteinfall zu einer lokal begrenzten Helligkeit in der Bildebene führt oder nicht unterscheidet man zwischen Geisterbildern (*ghosts*) sowie *Blendenflecken* (*lens flare*) und eher ausgedehnten Störungen durch gestreutes Licht (*veiling glare*). Zur Beurteilung des veiling glare sollte ein kleines, ideal schwarzes Objekt vor einem sehr ausgedehnten hellen Untergrund (ideal: Ulbrichtkugel) aufgenommen werden. Der Kontrast des schwarzen Objekts zur Umgebung ist ein geeignetes Maß zur Beurteilung[6]

Insbesondere bei der Analyse von Farben stören Streuanteile deutlich, da sie die Farbsättigung verringern. Aber auch generell wird der Kontrast durch die Streuanteile verringert, ganz besonders in dunklen Szenenbereichen, d.h. die MTF wird dann auch bei tiefen Raumfrequenzen kleiner eins.

Insgesamt lassen sich entsprechende Fehler bzw. Kontrastreduktionen vor allem durch passende Antireflexbeschichtungen, Beschichtungen im Inneren des Objektivs, Vermeidung zusätzlicher Störlichtquellen sowie Blenden reduzieren. Entsprechend ist eine Hauptursache von kontrastarmen, flauen Bildern der Einsatz eines Objektivs im falschen Wellenlängenbereich (meist Objektiv, das für den sichtbaren Spektralbereich gedacht ist, im NIR). Die Beschichtungen passen nicht zum Spektrum und daher ist der Reflexionsgrad pro Fläche entsprechend hoch. Dieses Störlicht führt dann zu einem mangelhaften Kontrast.

6.3 Foto- und Bildverarbeitungsobjektive

Der grundsätzliche Aufbau der meisten Bildverarbeitungsobjektive ist derselbe wie bei Fotoobjektiven. Lediglich die spezielle Optimierung (insbesondere auf einen nahen Arbeitsabstand und speziellen Wellenlängenbereich) sowie die mechanischen Parameter (Feststellung, Anschlußgewinde) unterscheiden die Typen.

[6] Eine genaue Definition ist in der ISO 9358 zu finden.

Die grundsätzlichen Bezeichnungen sind allerdings gleich.

Objektive, die einen Feldwinkel größer 60 ... 70° erfassen werden meist als *Weitwinkelobjektive* bezeichnet. Wenn noch deutlich größere Feldwinkel erfasst werden spricht man von *Hyperfeldwinkeln* oder im Extremfall von nahezu 180° oder mehr von *Fisheye-Objektiven*. Mit entsprechenden Objektiven lassen sich sehr große Szenen auf den Bildsensor abbilden und entsprechend werden sie z.B. stark in der Überwachungstechnik eingesetzt. Wenn die Brennweite eines Objektivs in etwa der Diagonale des Bildsensors entspricht werden Objektive als *Normalobjektiv* bezeichnet. Dadurch ergibt sich ein Feldwinkel von ca. 45°.

Wenn die Brennweite länger als die Baulänge ist, dann spricht man von einem *Teleobjektiv*. Der sogenannte *Telefaktor* bezeichnet gerade das Verhältnis von Brennweite zu Baulänge. Ein Telefaktor größer 1 wird erreicht, indem eine Kombination aus einer sammelnden und einer zerstreuenden Baugruppe (negative Brennweite) verwendet wird. Die zerstreuende Baugruppe liegt dabei relativ nah bei der Bildebene (Abb. 2.2.9).

Der umgekehrte Fall wird *Retrofokusobjektiv* genannt und vergrößert den Abstand zur Bildebene. Dadurch kann zusätzlicher Bauraum (z.B. für Filter) geschaffen werden. Je kleiner die Sensoren und damit auch die Brennweiten von Bildgebungssystemen werden, desto eher ist eine Retrofokuskonstruktion notwendig, da ansonsten kaum noch genug Platz zwischen letztem Objektivelement und Bildsensor bleibt.

Oftmals werden noch *Zwischenringe* zur Erweiterung der Einsatzmöglichkeiten von Objektiven empfohlen. Der Zwischenring wird zwischen Kamera und Objektiv angebracht und ändert somit die Bildweite. Damit ändert sich automatisch auch der Abbildungsmaßstab. Während dieses Vorgehen bei geringeren Anforderungen an die Auflösung sinnvoll ist, ist der Einsatz bei den heute üblichen hohen Auflösungen kaum zu empfehlen. Die Objektive sind schlicht für einen anderen Anwendungsfall, also ohne Zwischenring, optimiert und dementsprechend kann die spezifizierte Abbildungsqualität mit dem Zwischenring nicht erreicht werden.

Letztlich berechnet man also für die minimale Objektdistanz (MOD) des Objektivs den Bildabstand a'_{MOD} (dabei geht man näherungsweise davon aus, dass MOD

und Objektweite a näherungsweise gleich sind[7]) und für den gewünschten Abstand a ebenfalls den Bildabstand a' (Formel: $a' = af'/(a + f')$, Vorzeichen beachten). Die Differenz $a - a'_{MOD}$ ist dann die Zwischenringdicke.

Generell sind beim Einsatz von Fotoobjektiven für industrielle Kameras die Bildfelder zu beachten. Klassische Fotoobjektive für analoge Filmaufnahmen sind für ein *Filmformat* von 36 x 24 mm optimiert. Wenn sie mit einem vergleichsweise kleinen Sensor verwendet werden, dann wird entsprechend nur ein kleiner Teil des Bildfelds verwendet. Dies ist im Prinzip zunächst kein Problem, beachtet werden muss aber, dass von der Abbildungsleistung des Objektivs nur ein kleiner Teil genutzt wird. Eine gute bis sehr gute typische klassische Filmobjektivauflösung liegt in der Größenordnung von 30 bis 50 Linienpaaren pro Millimeter. Für moderne Bildverarbeitungsbildsensoren mit typischen Pixelgrößen im Bereich von 5 μm ist das meist nicht ausreichend.

Darüber hinaus sind die Objektive oft weit ab von einer bildseitigen Telezentrie. Dementsprechend sind die Hauptstrahlen für außeraxiale Bildpunkte geneigt. Dies führt in Zusammenhang mit den vor den meisten Bildsensoren liegenden Mikrolinsen unter Umständen zu Vignettierung und Übersprechen der Pixel.

Bei Bildverarbeitungsobjektiven gibt es nicht immer eine Möglichkeit zur Fokussierung am Objektiv sondern es wird teilweise davon ausgegangen, dass mittels des Objektabstands fokussiert werden kann. Dies muss bei der Auswahl bedacht werden.

6.4 Telezentrische Objektive

Die Vorteile der *Telezentrie* wurden bereits in Kapitel 2 dargestellt. Insofern werden telezentrische Objektive gern und oft vor allem in der messenden Bildverarbeitung eingesetzt. Dabei ist aber zu beachten, dass der Bereich, in dem die Telezentrie ausreichend gut funktioniert oft vergleichsweise klein ist. Daher sollte man beim Einsatz eines entsprechenden Objektivs die entsprechende Spezifikation genau prüfen und diese entsprechend einhalten. Bei der Angabe des Bereichs ist darauf zu achten, welche Größenänderung (z.B 10 μm) diesem Bereich entspricht.

[7]Das ist nicht immer der Fall (vgl. Teleobjektiv). Im Zweifelsfall muss man die Hauptebene des Objektivs kennen.

Da der Telezentriefehler sehr klein, aber dennoch von Bedeutung ist, ist oftmals eine Messung notwendig. Eine Vorgehensweise hierzu ist in [45] zu finden.

Zu unterscheiden ist zwischen objektseitig und beidseitig telezentrischen Objektiven. Bei letzteren ist zusätzlich der Strahlengang sensorseitig parallel (siehe Abb. 6.4.4). Dadurch ist der Einfallswinkel des Strahlkegels auf den Sensor (bzw. dessen Mikrolinsen) unabhängig von der Feldposition, so dass keine sensorseitigen Vignettierungen auftreten und auch das etwaige optische Übersprechverhalten zwischen benachbarten Pixeln minimiert und feldunabhängig wird. In der Regel sind Objektive, die für den Betrieb mit aktuellen digitalen Bildsensoren angeboten werden zumindest näherungsweise *bildseitig telezentrisch*.

Teilweise werden auch *teiltelezentrische Objektive* angeboten. In diesem Fall ist der *Telezentriefehler* zwar vergleichsweise hoch, der Telezentriebereich ist aber dennoch deutlich erweitert gegenüber einem konventionellen entozentrischen Objektiv.

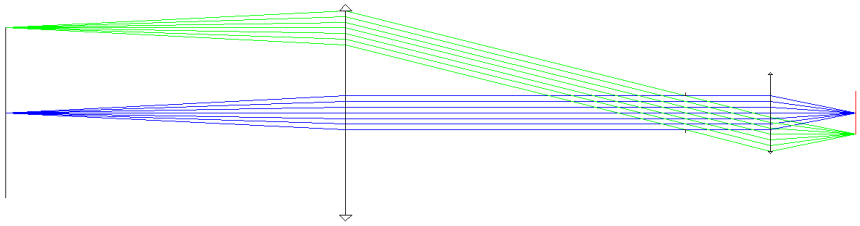


Abbildung 6.4.4: Beidseitige Telezentrie: Der Hauptstrahl jedes Objektpunkts läuft sowohl im Objektraum wie auch im Bildraum parallel zur optischen Achse.

6.5 Objektive für Zeilenkameras

Für die Abbildung auf eine *Zeilenkamera* kann natürlich ein gewöhnliches Objektiv eingesetzt werden. Genutzt wird die Abbildungsleistung dann eben nur für eine Zeile. Andererseits gibt es aber speziell für die Abbildung einer Zeile optimierte Objektive, die oft für diesen Fall dem konventionellen Objektiv in der Abbildungsqualität überlegen sind.

6.6 Mikroskopische Abbildungen

Für vergrößernde Abbildungen werden heute üblicherweise *Keplerteleskope* mit speziellen, für die mikroskopische Abbildung optimierten Objektiven (*Mikroskopobjektive*) verwendet. Der Grundaufbau einer mikroskopischen Abbildung ist in Abb. 6.6.5 dargestellt.[8] Das Objektiv ist dabei für eine Abbildung ins Unendliche optimiert[9] und eine nachfolgende, vergleichsweise einfache Linse[10] bildet dann das Zwischenbild aus dem Unendlichen auf den Bildsensor ab. Diese sogenannte *Tubuslinse* sollte idealerweise zum verwendeten Mikroskopobjektiv passen, um eine bestmögliche Abbildungsqualität zu erreichen. In guter Näherung sind die meisten Mikroskopobjektive telezentrisch aufgebaut.

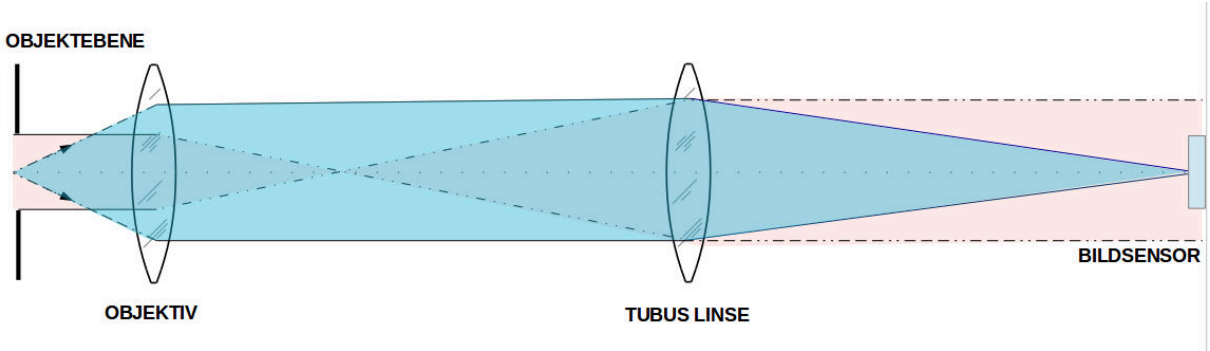


Abbildung 6.6.5: Aufbau eines abbildenden Mikroskops. Sowohl Objektiv als auch Tubuslinse sind auf „unendlich", also paralleles Licht, optimiert. Der Beleuchtungsstrahlengang ist rötlich gezeichnet, der Abbildungsstrahlengang ist blau dargestellt.

Ein genaues Verständnis der Kennzeichnung bzw. Spezifikation von entsprechenden Objektiven ist unerlässlich wenn hohe Fehlinvestitionen vermieden werden sollen, denn Mikroskopobjektive können durchaus schnell mehrere tausend Euro kosten.

Das Verhältnis der Brennweiten von Tubuslinse und Objektiv ergibt den Abbildungsmaßstab (bzw. in der Mikroskopie „Vergrößerung"). Die Angabe der *Vergrößerung* auf den Objektiven kann dementsprechend in eine Brennweite umgerechnet werden. Die Brennweite eines Zeiss Objektivs mit 100-facher Vergrößerung wird aufgrund der bei Zeiss üblichen Tubuslinsenbrennweite von 165

[8] Es gibt auch Objektive, die für einen Einsatz ohne Tubuslinse optimiert sind. In diesem Fall kann der Einsatz von dicken Filtern oder Strahlteiler problematisch werden da kein kollimierter Strahlengang vorliegt.

[9] Alte Mikroskope wurden auf einen festgelegte Position, z.B. 160 mm, des Bilds optimiert.

[10] Die numerische Apertur für diese Linse ist vergleichsweise klein, so dass kein komplexes System notwendig ist.

Abbildung 6.6.6: Typisches Bildverarbeitungsmikroskop

mm letztlich 1,65 mm.[11]

Die zweitwichtigste Angabe zur Spezifikation eines Mikroskopobjektivs ist die (objektseitig) *numerische Apertur*, die letztlich der Blendenzahl bei anderen Objektiven entspricht. Generell gilt, dass Objektive mit starken „Vergrößerungen" ein kleines Objektfeld abdecken und Objektive mit hoher numerischer Apertur eine geringe Feldausdehnung und eine sehr kleine Schärfentiefe ($\Delta z \approx 0.5\,\lambda/\mathrm{NA}^2$) aufweisen.

Neben diesen beiden Grundparametern (Vergrößerung und numerische Apertur) sind eine Reihe weiterer Angaben von Belang:

- *Immersion*: Wird das Objektiv in Luft verwendet oder handelt es sich um ein sogenanntes Immersionsobjektiv. Der Vorteil der Immersion ist die dadurch vergrößerte numerische Apertur ($\mathrm{NA} = \mathrm{n}\sin\alpha$). Nachteil, insbesondere bei der *Ölimmersion*, ist die Verschmutzung und natürlich ganz generell die problematische Arbeitsweise in einem Immersionsmedium.

  Mit „W" werden *Wasserimmersionsobjektive* gekennzeichnet während ein Arbeiten in Öl mit „oil" gekennzeichnet wird. Wenn keine spezielle Angabe er-

[11] Die Vergrößerung wird auf den Objektiven auch durch einen farblichen Ring gekennzeichnet, siehe DIN 8578.

folgt, dann ist das Immersionsmedium Luft.

Ein Einsatz des Objektivs mit dem falschen Immersionsmedium führt zu massiven (sphärischen) Aberrationen und sollte tunlichst vermieden werden.

- *Deckglas*: In der Mikroskopie ist es in vielen, vor allem biomedizinischen Anwendungen, üblich und sinnvoll, das Präparat durch ein Deckglas abzudecken. Die Deckglasdicke hat dabei erneut einen massiven Einfluss auf die Abbildung und dementsprechend muss die Deckglasdicke bei der Optimierung des Objektivs berücksichtigt werden. Wie auch bei der Immersion führt die falsche Deckglasdicke zu einer deutlich reduzierten Abbildungsqualität.

- *Tubuslänge*: Für Objektive mit einem Zwischenbild (welches üblicherweise dann visuell durch ein Okular betrachtet wird) wird die Tubuslänge in mm gefolgt von der Deckglasdicke angegeben. Also z.B. „160/0.17". Für Objektive ohne endliches Zwischenbild (das ist heute wie besprochen der Standardfall) wird ∞ angegeben.

- *Long Working Distance* (L, LD oder LDW) Objektive erzielen einen deutlich vergrößerten Arbeitsabstand. Dies kann ein immenser Vorteil sein, denn hochaperturige Mikroskopobjektive haben oft einen Arbeitsabstand von lediglich wenigen hundert Mikrometern.

- *Polarisation*: Für quantitative Polarisationsmikroskopie sind ebenfalls spezielle Objektive notwendig („POL", „P" oder „PO").

Daneben sind die *Ebenheit* des Objektfelds („Plan" oder „PL") und die chromatische Korrektur[12] wesentliche Merkmale, die den Preis eines Mikroskopobjektivs bestimmen. Für spezielle Mikroskopieverfahren, insbesondere die Abbildung von transparenten Objekten werden verschiedene *Phasenkontrastobjektive* angeboten („DIC" oder „PH"). Für die doch im Rahmen der industriellen Bildverarbeitung doch recht spezielle Anwendung entsprechender Verfahren wird hier nicht eingegangen.

Teilweise wird die Gesamtkorrektur der Aberration durch die Kombination von Mikroskopobjektiv und Tubuslinse erzielt. Ist dies der Fall dann sollte natürlich

[12]Eine gebräuchliche Bezeichnung ist „Apochromat". Dadurch wird zum Ausdruck gebracht, dass drei Wellenlängen (im sichtbaren Bereich) gleichzeitig korrigiert sind, allerdings natürlich nicht, wie gut die Abbildungsleistung bei diesen Wellenlängen wirklich ist.

die passende Kombination beider Komponenten verwendet werden. für Bildverarbeitungsanwendungen sind einfache Tuben erhältlich, bei denen bereits ein Mikroskopobjektiv und eine dazu passende Tubuslinse integriert sind und an die direkt ein passender Bildsensor angeschraubt werden kann. Extern ist dann lediglich eine passende Beleuchtung sowie geeignete mechanische Verstellmöglichkeit vorzusehen.

Die sehr übliche *Dunkelfeldbeleuchtung* (vgl. Abschnitt 8.1.1) kann mit konventionellen Objektiven realisiert werden, oft werden aber auch spezielle Dunkelfeldobjektive eingesetzt.

Die Preise von Mikroskopobjektiven variieren sehr stark, denn sehr gute Abbildungsleistungen über große Felder lassen sich nur mit sehr hohem Aufwand realisieren. *Apochromate* ergeben dabei bessere Abbildungsergebnisse als Achromate. „Plan" Objektive erzielen ein ebenes Objektfeld. „Epi"oder „M" bedeutet Auflicht (im Gegensatz zur in der Biologie traditionellen *Durchlichtmikroskopie* ('Kurzzeichen: „D").

Bei Bildverarbeitungsmikroskopen herrscht — genau wie bei 40 Euro Mikroskopen von Aldi oder Lidl — leider die Unsitte, gesamte Mikroskope oder Mikroskopobjektive rein über eine „Vergrößerung" zu spezifizieren. So kauft man dann z.B. ein 10.000-fach Mikroskop. Letztlich ist das natürlich Blödsinn, denn sobald Sie einen doppelt so großen Monitor verwenden, würde dann automatisch Ihr 10.000-fach Mikroskop zu einem 20.000-fach Mikroskop mutieren.

Die Vergrößerung zwischen Objekt und Ausgabepixel ist also kaum ein sinnvoller Parameter. Wichtiger ist eben die (objektseitige) Auflösung oder eben die numerische Apertur, denn diese erlaubt uns bei qualitativ hochwertigen Objektiven, die Auflösung zu berechnen bzw. abzuschätzen ($r = 0.61\lambda/\mathrm{NA}$). Ein 10.000-fach Mikroskop kann aber durchaus die mikroskopische Struktur eines Objekts deutlich schlechter auflösen als ein 1.000-fach Mikroskop.

Es ist schade, dass selbst sehr teure Mikroskope jenseits der 50.000 Euro Marke oft keine, für den Fachmann nachvollziehbaren technischen Angaben machen. Vermutlich wird dem Kunde nicht zugetraut, mit sinnvollen Angaben umzugehen.

6.7 Scheimpflug Bedingung

Wenn die Bildebene zur Objektebene bei einer Abbildung verkippt ist, dann ergibt sich das Problem, dass hierdurch verschiedene Bereiche auf dem Objekt in unterschiedlicher Entfernung liegen und dementsprechend Teile eines ausgedehnten Objekts unter Umständen unscharf abgebildet werden. Ob dies der Fall ist kann mit der Abschätzung der Schärfentiefe (vgl. Abschnitt 2.4) getestet werden.

Falls die Schärfentiefe ein Problem ist, kann durch Verkippen des Objektivs ein entsprechend verkipptes scharfes Objektfeld generiert werden. Hierzu wird die Senkrechte zur Symmetrieachse des Objektivs so gelegt, dass ihre Verlängerung durch den Schnitt der Objektebene und der Bildebene verläuft (vgl. Abb. 6.7.7).

Zu beachten ist allerdings, dass sich eine deutliche Verzerrung des Bildes ergibt, die per Software ausgeglichen werden muss.

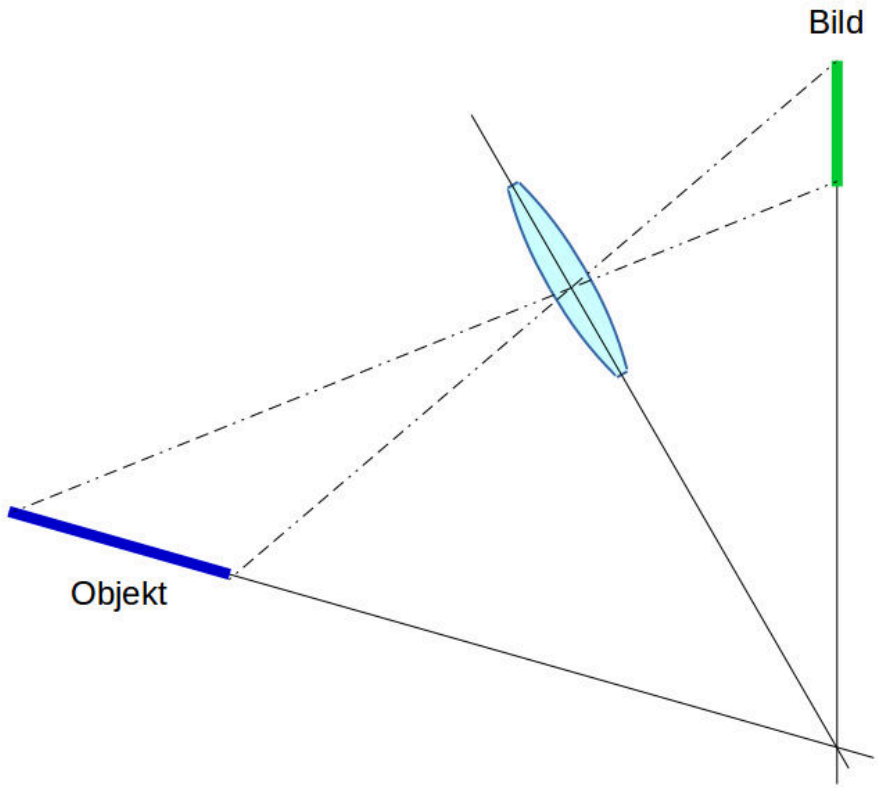


Abbildung 6.7.7: Wenn die Objektebene zur Bildebene verkippt ist, dann lässt sich dennoch eine scharfe Abbildung erzielen indem die Objektivebene so verkippt wird, dass sich alle Ebenen in einer Gerade schneiden. Zu beachten ist, dass dennoch das Bild perspektivisch verzerrt ist.

6.8 Mechanischer Anschluss

Natürlich muss das Objektiv in irgendeiner Form im optischen System, meist direkt an einer Kamera montiert werden. Mit Abstand am gebräuchlichsten sind hierfür *C-Mount* und *CS-Mount* Anschlüsse. Beide nutzen dasselbe 1 Zoll Gewinde mit einer Steigung von 1/32 Zoll. Unterschiedlich ist aber das Auflagemaß, also der Abstand vom Objektiv-Auflageflansch der Kamera bis zum Bildsensor. Bei C-Mount Objektiven beträgt dieser Abstand 17,526 mm und bei CS-Mount Objektiven 12,5 mm (vgl. Abb. 6.8.8). Durch den Einsatz eines geeigneten Zwischenrings kann ein C-Mount Objektiv also zum CS-Mount Objektiv umfunktioniert werden bzw. bei einer CS-Mount Kamera kann man auch C-Mount Objektiv einsetzen (wenn man einen 5 mm Zwischenring verwendet).

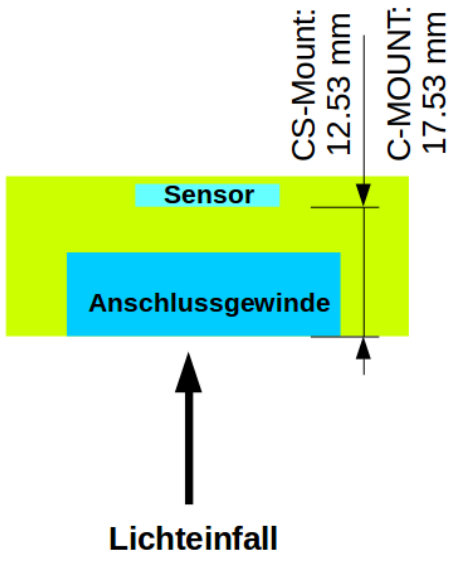


Abbildung 6.8.8: Rein mechanisch passen C-Mount Objektive auch auf CS-Mount Kameras (selber mechanischer Anschluss). Allerdings ist das Auflagemaß unterschiedlich.

Teilweise, vor allem für große Sensoren wird auch der sogenannte *F-mount* verwendet und zunehmend kommt auch aus der Consumer-Optik der *Micro-Four-Third Standard* (M43) in industrielle Kameras. Hierbei ist natürlich die Hoffnung, die durch Massenproduktion vergleichsweise preisgünstigen und dennoch hochwertigen Consumerobjektive nutzen zu können.

Wie der Name sagt, sind entsprechende Objektive für einen vergleichsweise großen Bildkreis, also ca. 1.3 Zoll Sensoren (immer beachten: Das entspricht nicht

echten 1.3 Zoll, vgl. Abschnitt 9.3) optimiert. Natürlich können sie auch mit kleineren Sensoren genutzt werden, ein Teil der Performanz bleibt dabei aber eben ungenutzt. Die Objektive sollen dabei gemäß dem Standard möglichst gut bildseitig telezentrisch sein.

Und die bildseitige Telezentrie ist — bei gegebenem Mount — rein geometrisch um so schwieriger zu erzielen, je größer das Bild sein soll.

Der noch wenig verbreitete *S-Mount* bzw. *M12* wird eher für kleinformatige Boardkameras (bis 1/2 oder 1/3 Zoll) und vor allem im Bereich der eher preisgünstigen Optiken verwendet.

6.9 Filter und Schutzgläser

Schutzgläser und *Filter* werden aus verschiedenen Gründen eingesetzt. Generell unterscheiden wir zwischen:

- Schutzgläsern, z.B. zur Abschirmung von Dreck oder hohen Drücken,
- *Spektralfiltern*, meist zur Unterdrückung von *Fremdlicht* oder zur spektralen Analyse,
- *Polarisationsfiltern* zur Unterdrückung oder Hervorhebung von *Reflexionen* und *Metallen*
- Graufilter zur reinen Lichtabschwächung

Zu beachten ist, dass ein zusätzliches Glas vor dem eigentlichen Objektiv in aller Regel im Optikdesign nicht berücksichtigt ist. Minderwertige Filter mit leichten Unebenheiten führen natürlich in jedem Fall zu Bildfehlern. Aber auch eine perfekte Planparallelplatte führt beim nicht-parallelen Durchgang zu Fehlern (Defokus, sphärische Aberration, Astigmatismus). Je paralleler der auf das Objektiv einfallende Strahlverlauf ist, desto unkritischer ist diese zusätzliche Planparallelplatte wenn sie in hoher Qualität ausgeführt ist. Insbesondere bei geringen Blendenzahlen in Verbindung mit geringen Objektdistanzen sollte die Wirkung beachtet werden. Auch bei parallelem Strahlverlauf (also großer Blendenzahl) können sich

aber (leichte) Verzeichnungen ergeben. Die Effekte sind umso stärker, je dicker der Filter ist.

6.9.1 Farbfilter

Farbfilter werden sehr häufig eingesetzt. Drei wesentliche positive Effekte lassen sich erzielen:

- Breitbandiges Licht führt zu chromatischen Aberrationen. Durch Einengung des Spektrum kann also tendentiell die Bildqualität verbessert werden.
- Bei monochromatischer Beleuchtung kann andersfarbiges (meist weißes) Licht deutlich reduziert werden. Dadurch steigt der Kontrast im Bild (*Fremdlichtunterdrückung*).
- Licht der falschen Wellenlänge wird von den Beschichtungen der Optik meist nicht perfekt transmittiert, so dass Streulicht und Ghosts entstehen können. Der Filter eliminiert dieses Licht.

Zusätzlich erfordern einige Anwendungen, vor allem Fluoreszenz, nahezu zwingend den Einsatz eines passenden Filters.

Grundsätzlich unterscheidet man bei spektralen Filtern zwischen

Tiefpass: bzw. *Langpass*, (in Reflexion: Hot mirrors), d.h. tiefe Frequenzen, also große Wellenlängen, werden durchgelassen.

Hochpass: (in Reflexion: *Cold mirrors*), d.h. hohe Frequenzen, also niedrige Wellenlängen, werden durchgelassen, z.B. als IR-Sperrfilter vor Bildsensoren.

Bandpass: Ein gewisser Spektralbereich wird transmittiert, z.B. zur Sichtbarmachung spezieller spektraler Details mit hohem Kontrast.

Bandsperre: (*Notch-Filter*): Ein bestimmter Bandbereich wird geblockt, z.B. zur Blockierung einer zur Anregung verwendeten Laserwellenlänge.

Komplexere Charakteristiken: Mehrere Spektralbänder werden transmittiert (z.B. Blau und Rot, nicht aber grün).

Zu unterscheiden ist auch, ob die Filter in Transmission oder Reflexion genutzt werden.

Bei den „Cut-on" bzw. „Cut-off" Wellenlängen werden üblicherweise 50% des maximalen Lichts transmittiert. Die *Halbwertsbreite* ist ganz entsprechend die spektrale Bandbreite, innerhalb derer die Transmission im Vergleich zur maximalen Transmission mindestens 50% beträgt.

Oftmals kann man durch den Einsatz geeigneter Filter eine Farbkamera und eine Farbbildverarbeitung vermeiden und sehr gute Kontraste erzielen. Unter kontrollierten Beleuchtungsbedingungen kann man allerdings auch statt einer abbildungsseitigen Farbfilterung eine Beleuchtung mit entsprechender spektraler Charakteristik verwenden. Diese kann entweder ebenfalls über Farbfilter, oder aber mittels passender LED-Beleuchtung realisiert werden.

Neben dem Verlauf der spektralen Charakteristik ist auch auf die generelle Transmission im Passbereich zu achten.

Die Funktionsweise von Farbfiltern basiert entweder auf Absorption oder aber Interferenz. Absorbierende Filter haben vergleichsweise breite Transmissions- bzw. Reflektionscharakteristiken und setzen die zu filternde Strahlung in Wärme um. Die dadurch bei hohen Lichtleistungen entstehende Wärme kann zu verschiedenen praktischen Problemen führen und muss dann durch Luftkonvektion oder Wasserkühlung abgeführt werden. Abbildungsseitig können die sich dabei ergebenden *Schlieren* in der Luft bzw. dem Wasser zu erheblichen Aberrationen und damit einer erheblichen Bildqualitätsminderung führen. Entsprechende Filterelemente sind aber preiswert, arbeiten nahezu winkelunabhängig und bieten eben breite Spektralcharakteristika.

Schmalbandige und sehr steile Filterkennlinien werden dagegen üblicherweise durch Interferenzfilter[13] realisiert. Die generelle Funktionsweise ist ähnlich der Farbentstehung an Seifenblasen, d.h. dünne Schichten unterschiedlicher Brechzahl, die auf ein Trägermaterial (Glas oder Kunststoff) aufgebracht sind führen zu Reflexionen an den Schichtgrenzen und die Interferenz der entsprechenden Lichtanteile ist wellenlängenabhängig.

[13] Von „Dichroiten Filtern" spricht man bei Interferenzfiltern, wenn ein ausgedehntes Spektralband transmittiert und das (farb-)komplementäre Feld reflektiert wird. Im Prinzip ist das allerdings für alle Interfernzfilter gegeben.

Hierbei ergibt sich eine mehr oder weniger starke Abhängigkeit der spektralen Charakteristik mit dem Winkel (siehe Abb. 6.9.9). Durch leichte Variation des Winkels, unter dem der Filter verwendet wird, kann die (zentrale) Wellenlänge eingestellt werden. Es gilt:

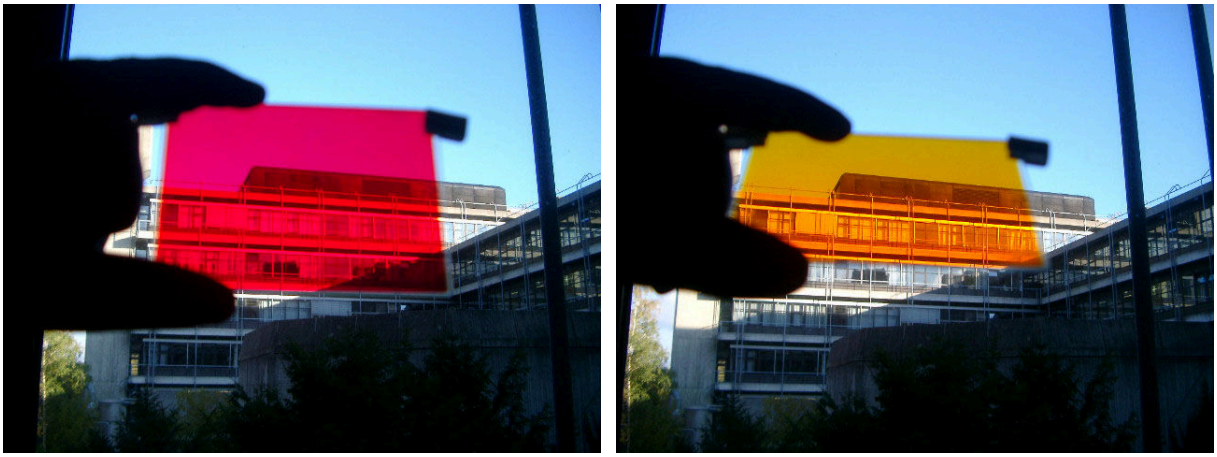
$$\lambda_{phi} = \lambda\sqrt{1 - \frac{\sin^2\phi}{n^2}} \tag{6.1}$$

Durch Drehung wird der Filter also zu kürzeren Wellenlängen verstimmt (*Blue Shift*). Dabei ist allerdings zu beachten, dass deutliche Winkel zu diversen Problemen (Reflexionsverluste, Aberrationen, Strahlversatz) führen. Auch ohne Drehung ergibt sich für die typische Position eines Filters (vor dem eigentlichen Objektiv) der Effekt, dass für nicht-telezentrische Objektive die Strahlenbündel von unterschiedlichen Feldpunkten eben unterschiedlich geneigt den Filter treffen und damit eine unterschiedliche spektrale Transmission erfahren. Natürlich ist der Effekt umso stärker, je größer der Feldwinkel (also Vorsicht bei Weitwinkelobjektiven) ist.

Das kann zu erheblichen Problemen führen. Typisch bei monochromatischer Beleuchtung ist es, dass dann am Feldrand die Bestrahlungsstärke deutlich abnimmt oder sogar ganz auf Null fällt. Dementsprechend gibt es Filter, die durch Kombination aus Absorption und Intererenz einigermaßen eng begrenzte Bandpassfilter erreichen und den Effekt deutlich minimieren.

Aus diesem Grund wird der Filter auch teilweise zwischen Objektiv und Bildsensor angebracht. Hierzu gibt es Filter mit C-Mount Gewinde, die — vorausgesetzt es ist ausreichend Platz — eingebracht werden können. Die Grundidee ist, bei (ganzer oder zumindest teilweiser) bildseitiger Telezentrie die Winkeländerungen mit dem Feld zu vermeiden und damit eine konstante (über das Feld) spektrale Transmission zu erreichen. Aber auch hierbei können sich Probleme ergeben. Zunächst ändert sich die spektrale Transmission (Verbreiterung). Insbesondere bei kleinen Brendenzahlen muss aber beachtet werden, dass sich die Auflösung verringern kann. Zum einen ändert sich effektiv die numerische Apertur (bzw. Blendenzahl), denn die stark geneigten Strahlen werden abgeschwächt. Es ergibt sich also eine sogeannte Apodisation. Zum anderen ist (je nach Filterdicke und Blendenzahl) mit Aberrationen zu rechnen, denn es wird stark fokussiert durch eine Planparallelplatte transmittiert.[14]

[14] Objektseitig ist die numerische Apertur deutlich kleiner und damit dieser Aberrationsbeitrag viel geringer.



(a) Orientierung 1 (b) Orientierung 2

Abbildung 6.9.9: Interferenzfarben am Beispiel eines Interferenzfilters: Durch die Reflexion an Vorder– und Rückseite an dünnen Schichten ergibt sich eine – je nach Wellenlänge und Einfallswinkel – konstruktive oder destruktive Überlagerung und damit eine Verstärkung oder Abschwächung. Bei Verkippung ergibt sich ein „Blue-Shift".

Typische Halbwertsbreiten von Interferenzfiltern liegen im Bereich von 5...100 nm, so dass auch breitbandiges Fremdlicht sehr effektiv von der eigentlichen Beleuchtungswellenlänge (Laser oder LED) getrennt werden kann. Ein weiterer Vorteil ist die hohe Transmission, die im Durchlassbereich möglich ist ($> 90\%$). Achten Sie darauf, dass teilweise die Filtercharakteristik nur sehr ungenau garantiert ist und dann z.B. die Zentralwellenlänge um bis zu 20 nm schwanken kann.

Die Dämpfung in Sperrbereichen von Filtern wird üblicherweise als *Optische Dichte* (*Optical Density*, „Blocking") logarithmisch angegeben, d.h. ein OD 3 Filter führt zu einer Abschwächung um den Faktor $10^3 = 1000$.

6.9.2 Advanced: Schutzgläser

Bei verschiedenen Anwendungen können *Schutzgläser* essentiell für die Sicherung der Funktionsweise der Abbildungsoptik sein. Verschmutzungen (z.B. Ölnebel, Staub) oder Umgebungsbedingungen (z.B. hohe Temperatur, Drücke) sind oftmals von der Optik oder dem Rest des BV-Systems abzuschirmen. Meist ist eine *Antireflexbeschichtung* notwendig, um *Doppelbilder* oder sonstige Störungen zu vermeiden. Es gibt aber zusätzlich Beschichtungen, die gegen Beschlagen (anti-fog, hydrophobic) oder Verschmierung (oleophobic) helfen.

6.9.3 Advanced: Polarisationsfilter

Typischerweise werden in der Bildverarbeitung auf Absorption basierende *Polarisationsfilter* eingesetzt. Hauptnachteil ist, dass auch ein Teil des an sich zu transmittierenden Lichts absorbiert wird. Die Lichteffizienz also negativ beeinflusst wird und das Element kann überhitzen. Kristallpolarisatoren vermeiden diese Nachteile, sind aber in der Bildverarbeitung aufgrund des sehr hohen Preises nicht üblich.

Wie bei jedem Filter ist zunächst die Qualität des Substrats entscheidend. Inhomogenitäten der Dicke oder des Materials führen zu Aberrationen. Weiterhin ist teilweise wichtig, wie das Auslöschungsverhältnis von gewolltem zu ungewolltem Licht ist.

Details zur Polarisation finden sich in [34].

6.10 Infrarotoptiken

Natürlich ist es problematisch, Objektive deutlich außerhalb ihrer Spezifikation zu betreiben. Dementsprechend ist es meist wenig sinnvoll, die Wellenlängen deutlich anders als für das Objektiv spezifiziert zu wählen. In einem Teil des nahen *Infrarot* (780 - 1100 nm) können oft noch konventionelle Objektive eingesetzt werden. Es ergibt sich zwar eine verringerte Abbildungsqualität und auch die Transmission ist (aufgrund der unpassenden Beschichtungen der Linsen) oft verringert, dennoch ist eine sinnvolle Abbildung möglich.

Problematischer sind natürlich deutlich längere Wellenlängen. Oftmals kommt erschwerend hinzu, dass der Wellenlängenbereich sehr groß ist und daher die Korrektur chromatischer Aberrationen schwierig wird. Da im mittleren und längeren Infrarot meist aber auch die Auflösung des Bildsensors deutlich reduziert ist, kommt man dennoch zu ausgewogenen Systemen.

6.11 Advanced: Reinigung und Pflege

Verschmutzungen können die Abbildungsleistung von Optiken erheblich beeinträchtigen. Daher kann es notwendig werden, die Optiken zu reinigen. Verschmutzungen, die in bzw. nahe der Bildebene oder einer Zwischenbildebene liegen, zeigen sich als lokalisierte Defekte im Bild. Oft stellt sich die Frage, ob der Bildsensor selbst verschmutzt ist. Dies kann einfach durch Rotation des Sensors gegenüber dem Rest der Optik untersucht werden. Bleibt die Position der Verschmutzungen im Bild erhalten dann kann davon ausgegangen werden, dass das Problem auf dem Bildsensor zu suchen ist.

Verschmutzungen, die in anderen Ebenen auftreten führen zu einer Kontrastminderung des Bildes aufgrund von Streulicht.

Generell gilt, dass die Reinigung von Optiken nur erfolgen sollte, wenn es unbedingt notwendig ist. Daher ist oberste Priorität, die Verschmutzung zunächst zu vermeiden. Im Zweifelsfall sollten sie passende Handschuhe bei der Montage/Justage tragen, denn insbesondere *Fingerabdrücke* können vergleichsweise schlecht entfernbare Verschmutzungen erzeugen. Ebenfalls zur Prävention können Schutzfilter vor dem eigentlichen Objektiv verwendet werden. Bei Nichtbenutzung sollten die Optiken abgedeckt werden (im einfachsten Fall durch Objektivdeckel).

Verschmutzungen können teilweise auch durch geeignete Luftführung im Betrieb vermieden werden, z.B. durch Schutzgase, die einen leichten Überdruck eines gehausten Inspektionssystem erzeugen. Typischerweise sollten Objektive eher nach unten gerichtet eingebaut werden weil sich Staub dann deutlich reduziert auf der Frontlinse absetzt.

Sollte dennoch eine Reinigung notwendig werden, sind einige generelle Dinge zu beachten, um Beschichtungen nicht zu zerstören und nicht oberflächliche Verschmutzungen zu unwiderbringbaren Fehlern zu machen. Details finden sich in [30] und für den speziellen Fall von Mikroskopobjektiven in [50]. Im Zweifelsfall ist der Hersteller der Optik zu befragen.

- Es sollte immer mit der wenigsten aggressiven Behandlung begonnen werden. Ein typischer Start ist das Abblasen mit einem speziellen Blasebalg. Hier-

durch soll schlecht anhaftender Staub oder andere partikuläre Verschmutzungen entfernt werden.

Teilweise wird auch Druckluft verwendet. Wenn der Druck zu hoch ist, ist dieser Ansatz allerdings gefährlich, denn er kann dazu führen, dass harte Partikel auf der Oberfläche in die Oberfläche eingetrieben werden und so zu Kratzern führen.

- Fingerabdrücke sollten möglichst schnell entfernt werden.
- Zunächst wird *Isopropanol* oder eine spezielle Linsenreinigungsflüssigkeit verwendet, um die Oberfläche weiter zu säubern. Hierzu wird mit einem Papierlinsenputztuch ein Tropfen der Reinigungsflüssigkeit so langsam abgezogen, dass an der Begrenzung des Linsenputztuchs die Reinigungsflüssigkeit gerade verdunstet. Natürlich ist das Papierreinigungstuch entsprechend der Linsengeometrie zu zuschneiden.
- Hartnäckigerer Schmutz kann mit *Aceton* behandelt werden. Hierbei besteht aber bereits eine deutlich vergrößerte Gefahr der Beschädigung von Beschichtungen. Insbesondere kann dadurch auch Material der Linsenbefestigung angegriffen bzw. aufgelöst werden, das sich dann im ungünstigsten Fall über die Linse verteilt.
- Lokale Verschmutzungen auf sehr kleinen Optiken (Mikroskopobjektive) reinigt man am besten mittels speziellen Baumwollstäbchen oder synthetischen Putzstäbchen (keinesfalls konventionelle Q-Tipps verwenden).

Aber immer gilt: Reinigung vermeiden, wenn nicht unbedingt notwendig !

Kapitel 7

Fotometrie und Radiometrie

In diesem Kapitel werden Sie unter anderem lernen

- wie basierend auf einer physikalischen Strahlungsbeurteilung photometrische Größen abgeleitet bzw. berechnet werden,
- welche photometrischen und radiometrischen Grundgrößen von Bedeutung sind,
- wie man die Bestrahlungsstärke auf dem Bildsensor berechnet und
- wie man ein Objekt beleuchten muss, um eine gewisse Bestrahlungsstärke auf einem Sensor zu erzielen.

7.1 Lichttechnische Größen

Die Lichttechnik ist ein eher unübersichtliches und wenig geliebtes Gebiet der Bildverarbeitung. An sich sind die wichtigsten Zusammenhänge gut überschaubar. Im wesentlichen sind vier Grundgrößen für uns wesentlich. In der Praxis ergeben sich aber eine Vielzahl von Bezeichnungen und Einheiten, denn diese Grundgrößen werden sowohl radiometrisch, also mittels SI-Einheiten, als auch photometrisch, also in Bezug auf das menschliche Sehsystem, definiert. Somit ergeben sich bereits acht Größen und üblich sind sowohl englische als auch deutsche Bezeichner und historisch bedingt sehr viele verschiedene Einheiten[1].

[1] Details siehe DIN 5031

Die Wellenlängen von für den Menschen sichtbarem Licht liegen im Bereich von ca. 380 bis 780 nm (siehe Abb. 7.1.1). Ultraviolettes Licht entspricht kürzeren Wellenlängen (ca. 100 bis 380 nm) und das infrarote Licht liegt oberhalb des sichtbaren Bereichs.

Man könnte der Ansicht sein, dass diese *Empfindlichkeit* des menschlichen Sehsystems für uns ohne Belang ist, denn es geht in der industriellen Bildverarbeitung um maschinelles Sehen. Aus zweierlei Gründen ist dies nicht der Fall. Zum einen erfolgt die Spezifikation vieler Beleuchtungsquellen und auch Kameraempfindlichkeiten oft in Einheiten, die auf das menschliche Sehsystem bezogen sind. Insofern ist eine Kenntnis der entsprechenden Einheiten von Bedeutung. Zum anderen haben aber natürlich auch technische Detektoren eine spektrale Empfindlichkeit. Dementsprechend gelten die grundlegenden Zusammenhänge, die sich bei menschlichen Sehen ergeben genauso für technische Systeme. Der einzige Unterschied ist die veränderte spektrale Empfindlichkeit, die wir durch $V(\lambda)$ bezeichnen.

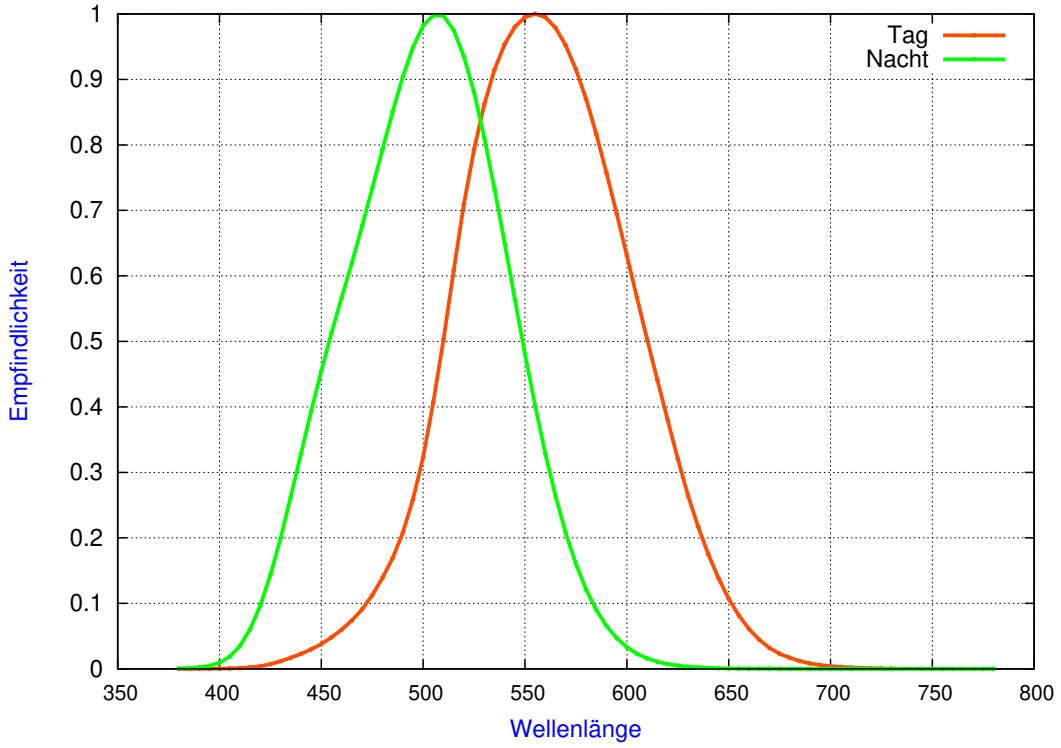


Abbildung 7.1.1: Empfindlichkeit $V(\lambda)$ des menschlichen Auges. Tagsüber liegt die maximale Empfindlichkeit bei ca. 550 nm (gelbgrün), nachts verschiebt sie sich in den blaugrünen Bereich (Purkinje–Effekt). Die Kurven sind jeweils normiert. Nachts ist die Empfindlichkeit natürlich erhöht.

Um die Reaktion eines lichtempfindlichem Sensors, z.B. eines Auges, auf ein be-

liebiges Eingangsspektrum $\Phi_e(\lambda)$ zu berechnen starten wir zunächst mit einer Wellenlänge. Offensichtlich ist das Signal des Sensors auf einfallendes der Wellenlänge λ umso höher, je höher die Empfindlichkeit ist, d.h.

$$\Phi(\lambda) \sim V(\lambda)\ \Phi_e(\lambda) \tag{7.1}$$

Die Proportionalität[2] beschreiben wir wie üblich durch eine Konstante, konkret K_{M}, also

$$\boxed{\Phi(\lambda) = K_{\mathrm{M}} \cdot V(\lambda)\ \Phi_e(\lambda) \qquad (7.2)}$$

Wenn wir als Einheit für das Eingangsspektrum Watt verwenden, dann ergibt sich für das Ergebnis Φ ebenfalls die Einheit Watt, zumindest wenn die Empfindlichkeit V und die Proportionalitätskonstante als einheitslose Gewichtsfaktoren betrachtet werden.

Da nun aber die Größe Φ auf das menschliche Sehen bezogen ist, wäre eine solche physikalische Einheit ungünstig. Stattdessen wird eine weitere Einheit eingeführt, das Lumen. Um die Konversion von dieser Einheit *Lumen* (lm) auf Watt durchzuführen benötigt man den Konversionsfaktor (genauer: *photometrisches Strahlungsäquivalent*) K_{M}. Die Einheit dieses Faktors muss dann lm/W sein. Aus historischen Gründen ist der numerische Wert nicht 1 lm/W sondern 683 lm/W. Eine *Strahlungsleistung* von 1 W im gelb-grünen Spektralbereich ($V(555\mathrm{nm}) \approx 1$) führt somit zu einem Lichtfluss von 683 lm.[3]

Wenn nun eine zweite Wellenlänge gleichzeitig auf den Sensor fällt, dann wird auch für diese Wellenlänge Gl. 7.2 gelten. Da nun zwei Wellenlängen vorhanden sind haben wir

$$\Phi = K_{\mathrm{M}} \cdot V(\lambda_1)\ \Phi_e(\lambda_1) + K_{\mathrm{M}} \cdot V(\lambda_2)\ \Phi_e(\lambda_2), \tag{7.3}$$

also die Summe.

[2] die letztlich eigentlich die Definition von $V(\lambda)$ selbst darstellt

[3] Da die Empfindlichkeit des menschlichen Sehsystems bei schlechten Lichtverhältnisse (nachts) doch deutlich verändert ist (erhöhte Empfindlichkeit und spektrale Empfindlichkeit blau-verschoben) kann auf die Augempfindlichkeit dann mit einem anderen Konversionsfaktor (1699 lm/W statt 683 lm/W) umgerechnet werden. Entsprechend werden alle lichttechnischen Größen im Prinzip sowohl für das Tag- als auch das Nachtsehen definiert (vgl. DIN 5031).

Wenn wir gleichzeitig eine kontinuierliches Spektrum aus vielen Wellenlängen haben, dann wird die Summe eben zu einem Integral, dem *Überlappintegral* zwischen der normierten spektralen Empfindlichkeit $V(\lambda)$ und dem Eingangsspektrum $\Phi_e(\lambda)$:

$$\Phi = K_{\mathrm{M}} \int_0^\infty V(\lambda)\, \Phi_e(\lambda)\, d\lambda \tag{7.4}$$

Auf diese Weise kann zwischen beliebigen *radiometrischen Größen*, die in physikalischen Einheiten angegeben sind, auf *fotometrische Größen* umgerechnet werden. Man unterscheidet also zwischen *Fotometrie* mit auf das menschliche Sehsystem bezogenen Größen und *Radiometrie*, in der rein energetische Größen Verwendung finden. Ganz entsprechend kann auch mit einer veränderten Empfindlichkeit auch das Verhalten von anderen Bildsensoren berechnet bzw. spezifiziert werden.

Verschiedene Größen sind dabei von Belang und finden sich in den technischen Daten von Beleuchtungsquellen. Grundsätzlich werden die selben Symbole bzw. Variablennamen für radiometrische und fotometrische Größen verwendet. Zur Unterscheidung der beiden Systeme bekommt die radiometrische Größe jeweils ein „e" als Subscript („e" für „energetic").

Unsere Basisgröße ist der *Lichtfluss* bzw. *Lichtstrom* Φ. Er entspricht – aber eben auf das Auge bezogen – der gesamten *Strahlungsleistung* Φ_e. Seine Einheit ist das Lumen.

Die *Beleuchtungsstärke* E ist der Lichtfluss pro Flächeneinheit (vgl. Abb. 7.1.2):

$$E = \frac{\Phi}{A} \tag{7.5}$$

In radiometrischen Größen sprechen wir von der *Bestrahlungsstärke* E_e, die eben ganz entsprechend die Strahlungsleistung pro Flächeneinheit ist.[4]

[4] Wenn wir die Abstrahlung pro Flächeneinheit verwenden, dann muss man bedenken, dass die effektiv in eine Richtung ϕ wirksame bzw. sichtbare Fläche natürlich rein geometrisch noch mit $\cos\phi$ gewichtet werden muss (Projektion).

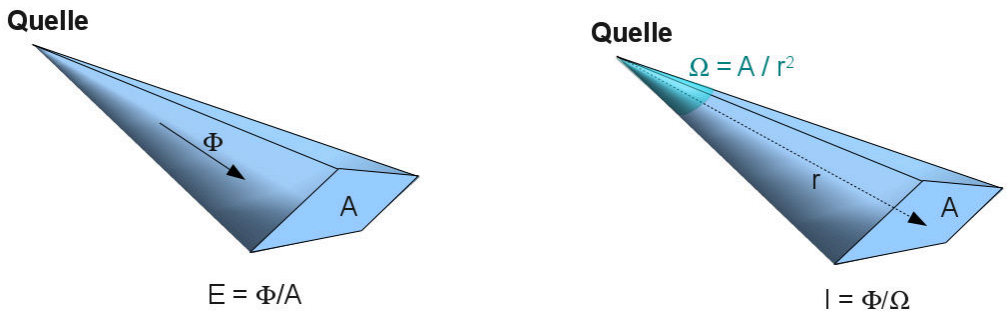


Abbildung 7.1.2: Links: Die Beleuchtungsstärke (Bestrahlungsstärke) E ist der Lichtfluss (Strahlungsfluss) pro Flächeneinheit. Rechts: Die Lichtstärke (Strahlstärke) I ist definiert als der Lichtfluss (Strahlungsfluss) pro Raumwinkel.

Die Definition der fotometrischen bzw. radiometrischen Größen gilt streng genommen immer nur für infinitesimale Größen. Genauer müssten wir also an sich schreiben: $dE = d\Phi/dA$. Bei Variation über den Raum muss man integrieren, also z.B. $E = \int dE$. Allerdings wird man in der industriellen Bildverarbeitung in der Regel die Größen als näherungsweise konstant annehmen, so dass man einfach die Differentiale weg lässt, also z.B. $E = \Phi/A$. Man sollte sich der dadurch entstehenden Näherung aber bewusst sein.

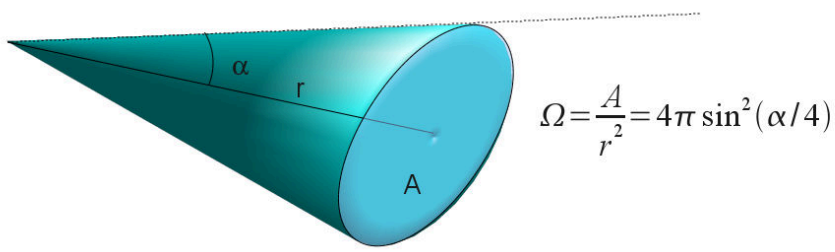


Abbildung 7.1.3: Raumwinkel geben eine Art räumlichen Winkel an. Spezifiziert wird dieser Winkel im Raum, indem man sich eine Kugel mit Radius 1 um den Quellpunkt denkt und den Flächenanteil, der von der Strahlungsverteilung diese Kugel durchstößt bestimmt. Dieser Flächenanteil ist gerade der Raumwinkel. Für den Vollraum ergibt sich somit als Raumwinkel die Kugeloberfläche, also 4π.

Die *Lichtstärke* I (I_e) ist der Lichtfluss (Strahlungsfluss) in einen Raumwinkelbereich, also

$$I = \frac{\Phi}{\Omega} \tag{7.6}$$

(vgl. Abb. 7.1.2 rechts). Der *Raumwinkel* ist die Fläche einer Einheitskugel, die vom Strahlungsbündel durchstoßen wird:

$$\Omega = \frac{A}{r^2} \tag{7.7}$$

Für einen Lichtkegel (Abb. 7.1.3) gilt

$$\Omega = 4\pi \sin^2(\alpha/4) \approx \pi \sin^2 \alpha \tag{7.8}$$

Die *Leuchtdichte* (Strahldichte) schließlich ist der Lichtfluss (Strahlungsfluss) pro Raumwinkel und Fläche. Sie gibt letztlich die subjektiv wahrgenommene Helligkeit einer Licht emittierenden Fläche an.[5] Bei zunehmender Entfernung des Auges von dieser Fläche nimmt der das Auge erreichende Lichtfluss zwar aufgrund der *Energieerhaltung* (*quadratisches Abstandsgesetzes*) ab, die Fläche des Bildes auf der Netzhaut nimmt aber ebenfalls quadratisch ab, so dass letztlich die Beleuchtungsstärke auf der Netzhaut konstant bleibt. Wenn uns also eine Fläche „hell" erscheint, dann bedeutet das, dass die Fläche in unseren Raumwinkelbereich viel Licht pro Flächeneinheit abstrahlt. Verkompliziert wird die Thematik auch dadurch, dass es einen „projizierten Raumwinkel" gibt, bei dem die Projektion mittels des Kosinus bereits einberechnet ist. Für kleine Winkel der Oberflächennormale zur Beobachtung entspricht der Raumwinkel dem projizierten Raumwinkel und der Kosinus ist 1. Vorsicht ist aber eben bei größeren Beobachtungswinkeln dahingehend gegeben, dass der Kosinus genau einmal (und nicht etwa doppelt) einberechnet wird.

Bei nicht senkrechter Betrachtung der Fläche muss für die Normierung auf die Fläche wieder die effektiv wirksame bzw. sichtbare Fläche verwendet werden (vgl. Fußnote zur Beleuchtungsstärke). Es ergibt sich also

[5] Aufgrund der Nichtlinearität der Helligkeitsempfindung des Auges die Leuchtdichte noch mit 0.33 (dunkles Umfeld) oder 0.5 (helles Umfeld) zu potenzieren um ein besseres Maß für die subjektive Helligkeit zu erhalten (siehe [34])

$$L = \frac{I}{A \cos\phi} \tag{7.9}$$

Die Leuchtdichte (Strahldichte) bleibt beim Durchgang durch ein optisches System ohne Verluste konstant — unabhängig in welcher Ebene des Systems man die Dichte bestimmt.

Nun sind diese acht Größen zwar im Prinzip vergleichsweise einfach miteinander verbunden, aber die Vielfalt der Namen und Einheiten führt in aller Regel dann doch zur Verwirrung bzw. Unübersichtlichkeit. Abb. 7.1.4 zeigt nochmals den Zusammenhang und in der folgenden Tabelle sind alle wesentlichen Bezeichnungen aufgeführt. Zumindest kann man sich merken, dass der hintere Teil des Namens für radiometrische und fotometrische Größen jeweils gleich sind. Der vordere Teil hat für fotometrische Größen immer mit „Licht" oder „leuchten" zu tun während er für radiometrische Größen eine Wortverwandschaft mit „Strahlung" aufweist.

Abbildung 7.1.4: Photometrische (radiometrische) Größen ergeben sich jeweils durch Angabe des Lichtfluss (bzw. der Leistung) pro Raumwinkel und/oder pro Fläche. Bild nach W.J. Cassarly, „Nonimaging Optics".

Größe	**Formel**	**Einheit**	**engl. Bez.**
Lichtfluss, Lichtstrom Φ		lm	luminous flux
Strahlungsfluss, Leistung Φ_e		W	(radiant) flux, power
Beleuchtungsstärke E	$E = \Phi/A$	lx = lm/m^2	illuminance
Bestrahlungsstärke[6] E_e		W/m^2	irradiance
Lichtstärke I	$I = \Phi/\Omega$	cd	luminous intensity
Strahlstärke I_e		W/sr	radiant intensity
Leuchtdichte L	$L = I/(A\cos\phi)$	cd/m^2	luminance
Strahldichte L_e		W/sr/m^2	radiance

Wenn man die Bestrahlungsstärke oder Beleuchtungsstärke auf eine emittierende Fläche bezieht ist nach DIN 5031 der Begriff *spezifische (Licht-)Ausstrahlung* zu wählen.

Besondere Vorsicht ist beim Begriff *Intensität* gegeben. Dieser bedeutet im Deutschen in aller Regel eine Bestrahlungsstärke, also Leistung pro Flächeneinheit. Im angelsächsischen wird der Begriff innerhalb der Optik unterschiedlich gebraucht. Teilweise bezeichnet er auch dort eine Bestrahlungsstärke, in der Fotometrie und Radiometrie ist er aber eine Strahlungsstärke, also die Leistung auf den Raumwinkel bezogen.

Nach [1] geben wir einige typische Werte für Beleuchtungsstärken und Leuchtdichten an:

	E_e	E	L
Direktes Sonnenlicht (am Mittag)	1000 W/m^2	100.000 lx	1.6 $\cdot 10^9$ cd/m^2
Bewölkter Himmel	100 W/m^2	10.000 lx	2000 cd/m^2
Vollmond	0.002 W/m^2	0.2 lx	2.500 cd/m^2
Mondlose Nacht 10^{-5} W/m^2	0.001 lx	0.001 cd/m^2	
Büro	5 W/m^2	300..500 lx	100 cd/m^2
Xenon Kurzbogenlampe			bis 5 $\cdot 10^9$ cd/m^2
100 W Glühbirne, matt			200 cd/m^2
LED-TV			500 cd/m^2
weißes Papier @ 500 lx			100 cd/m^2
Gehweg, beleuchtet @ nacht			2 cd/m^2
Straße im Abblendlicht			1000 cd/m^2

Teilweise, insbesondere bei Beleuchtungssystemen, ist der Begriff der *Etendue* bzw.

Lichtleitwert von Bedeutung. Die Etendue ist das Produkt aus Raumwinkel der Strahlung und (wirksamer) Fläche[7], durch die die Strahlung propagiert. Dieser Lichtleitwert kann niemals (durch ein passives optisches System) ohne Lichtverlust abnehmen. werden. Abnehmen kann der Lichtfluss zwar durchaus, aber das ist dann eine Folge von Verlusten. Man kann also durchaus die Bestrahlungsstärke ändern, z.B. indem man die Fläche verringert. Gleichzeitig vergrößert sich dadurch aber der Raumwinkel, um den Lichtleitwert konstant zu halten.[8]. Anschaulich kann man sich das am einfachsten anhand eine Keplerteleskops klar machen.

Für Lichtquellen ist es meist ideal aber eben auch schwierig, möglichst viel Licht aus einem kleinen Bereich in einen möglichst kleinen Raumwinkelbereich abzustrahlen. Das Extrembeispiel dafür ist der Laser. Man spricht in diesem Zusammenhang auch von *Brillianz*. Das entsprechende Gegenteil ist ein diffus abstrahlender wolkenbehangener Himmer.

Um mit den Größen vertraut zu werden, sollten Sie einige (an sich sehr einfache) Übungsaufgaben durchspielen:

Beispiel: Die Beleuchtungsstärke auf dem Sensor beträgt E=0.5 lx. Sie benutzen rotes Licht bei einer Wellenlänge von 600 nm. Berechnen Sie die Bestrahlungsstärke. ($V(600$ nm) = 0.631).

Lösung: Für monochromatisches Licht ergibt sich

$$E(\lambda) = \frac{\mathrm{lm}}{\mathrm{W}} \mathrm{V}(\lambda) \Phi_{\mathrm{e}} \tag{7.10}$$

und damit $E_e = 0.5/683/0.631$ = 1.1 mW/m^2.

Beispiel: Berechnen Sie die Beleuchtungsstärke auf Ihrem Bildsensor (2/3 Zoll) wenn ein Lichtfluss von 0.001 lumen auf den Sensor fällt.

Lösung: $E = 0.001/A$ mit $A = 8.8 \cdot 6.6$ mm^2 ergibt $E = 17.2$ lx.

[7] also wieder die Fläche multipliziert mit $\cos\phi$

[8] In der geometrischen Optik ist diese Aussage praktisch gleichbedeutend mit der sogenannten Lagrangeinvarianten

---

Beispiel: Ihre Lichtquelle hat eine Lichtstärke von 100 candela und beleuchtet ein Objekt im Abstand von 1 m. Die Lichtstärke ist mehr oder weniger konstant über den Raumwinkelbereich. Berechnen Sie die Beleuchtungsstärke auf dem Objekt.

Lösung: Der Raumwinkel ist durch $\Omega = A/1^2 = A$ sr/m^2 gegeben. Mit $I = \Phi/\Omega$ erhalten wir $\Phi = I \cdot A$ und damit $E = \Phi/A = I \cdot lx/cd$ = 100 lx.

Beispiel: Die Strahldichte eines beleuchteten Objekts beträgt 50 W/(m^2 sr). Sie nehmen eine homogen verteilte Lichtaussendung an und die Fläche des Objekts beträgt 10 x 10 mm. Berechnen Sie die Leistung, die das Objekt in Richtung der Eintrittspupille des Kameraobjektivs remittiert wenn die Kamera in einem Abstand von 1 m steht und einen Durchmesser der Eintrittspupille von 10 cm aufweist.

Lösung: Mit $\Omega = A/r^2 = \pi r^2/1^2 = 0.0079$ erhalten wir $\Phi = L \cdot 0.01^2 \cdot \Omega = 39\ \mu$W

7.2 Bestrahlungsstärke auf dem Bildsensor

Für die Aufnahme von Bildern ist natürlich letztlich die Bestrahlungsstärke auf unserem Bildsensor die maßgebliche Größe, denn bei einer gegebenen Empfindlichkeit des Sensors kann damit dann berechnet werden, welche Grauwerte sich digital bei der Belichtung des Sensors ergeben.

Wir geben hier ohne Beweis die wesentliche Formel der Beleuchtungs- bzw. Bestrahlungsstärke auf dem Sensor an, wenn ein Objekt der *Helligkeit* (genauer: Leuchtdichte bzw. Strahldichte) L von einem Abbildungssystem mit der numerischen Apertur NA erfasst wird.

$$E = \pi\, L \cdot \mathrm{NA}'^2 \cos^4(\omega') = \frac{\pi}{4K^2} \cos^4(\omega') \tag{7.11}$$

ω' ist dabei der bildseitige Feldwinkel, also der Winkel des Hauptstrahls vom Zentrum der Austrittspupille zum jeweiligen Bildpunkt. NA' ist die bildseitige numerische Apertur, also der Sinus des halben Öffnungswinkels von der Austrittspupille zur Bildebene.

Der $\cos^4$ Anteil wird „natürliche" Vignettierung genannt und ist (normalerweise) bei allen (entozentrischen) Objektiven vorhanden[9], vgl. Abschnitt 2.3.12.

Beispiel: Ein Objekt emittiert Licht (weil es beleuchtet wird) mit einer Strahldichte von 10 W/m^2/sr. Das Kameraobjektiv arbeitet bei Blende 5. Berechnen Sie die on-axis Bestrahlungsstärke auf dem CCD Sensor.

Lösung: K=5 entspricht einer numerischen Apertur von $\sin\alpha' = 1/(2K) = 0.1$. On-axis bedeutet, dass der Feldwinkel ω' Null ist. Daher gibt $E = \pi \cdot 10 \cdot 0.1^2$ = 0.314 W/m^2.

Beispiel: Sie nutzen ein Objektiv mit einer Brennweite von 20 mm in Verbindung mit einem 1/3 Zoll Sensor (6 mm Diagonale). Wie stark ist die natürliche Vignet-

[9] Wenn keine entsprechenden Gegenmaßnahmen ergriffen werden.

tierung aufgrund des Objektivs? (Nehmen Sie an, dass der Objektabstand deutlich größer als die Brennweite ist.)

Lösung: Der maximale bildseitige Feldwinkel beträgt $\omega' = 3/20$ und daher $\omega' = 0.15$ rad. Dadurch ergibt sich die natürliche Vignettierung zu $\cos^4 \omega' \approx 0.95$ was bedeutet, dass in den Ecken des Bildsensors die Bestrahlungsstärke um 5% reduziert sein wird.

7.3 Streuung und beleuchtete Flächen

Um die Bestrahlungsstärke auf dem Sensor zu berechnen ist es nach Gleichung 7.11 notwendig, die Strahldichte L_e bzw. Leuchtdichte L des beleuchteten Objekts zu kennen.

Wird das Objekt mit der Beleuchtungsstärke E beleuchtet, kann bei *Lambertscher Streuung* die Leuchtdichte nach

$$E = \pi L \tag{7.12}$$

berechnet werden.[10] (Im übrigen ist ein preiswertes (< 50 Euro) Luxmeter oft sinnvoll, um Beleuchtungen zu überprüfen aber auch die Stärke des Fremdlichts abzuschätzen.)

Lambertsche Streuung kann praktisch als mehr oder weniger „perfekte" Streuung einer Oberfläche angesehen werden. Einfallende Lichtteilchen werden unabhängig von ihrer Einfallsrichtung auf die Oberfläche in eine beliebige Ausfallsrichtung gestreut. Für die Lichtstärke gilt:

$$I = I_0 \cos \phi \tag{7.13}$$

In der Praxis zeigen z.B. mattes Papier oder sandgestrahlte Oberflächen nahezu Lambertsche Streuung.

Die „Helligkeit" bzw. Leuchtdichte ist konstant, unabhängig davon aus welcher Richtung die Fläche betrachtet wird[11].

[10] Für eine Ableitung siehe z.B. [44].

[11] Unter schräger Betrachtung ist ein Flächenelement um den Faktor cos ϕ verkleinert, so dass sich damit das cos ϕ

Beispiel: Sie bilden eine diffus streuende Oberfläche mit einer Fläche von 0.5 m^2 ab. Die Fläche wird mit 1 W/m^2 beleuchtet. 50% des Lichts wird absorbiert. Berechnen Sie die Strahlstärke in Richtung 20° zur Oberflächennormalen.

Lösung: $I_e = L_e \cdot A = E_e \cdot 0.5/\pi \cdot 0.5 = 0.079$ W/sr on axis und multipliziert mit cos(20) ergibt sich I(20°) = 0.075 W/sr.

In der Praxis ergibt sich oft eine Mischform aus idealer Reflexion und idealer Streuung, so dass man das Gesamtverhalten im einfachsten Fall eben durch einen gerichteten Anteil (Reflexion) und einen ideal ungerichteten Anteil (Streuung) beschreibt. Dies ist allerdings auch nur eine Näherung, da sich in der Realität eine Streukeule ergibt die mathematisch am besten durch die *BRDF* (Bidirectional reflectance distribution function, vgl. Abschnitt 7.5) beschrieben wird.

7.4 Advanced: Notwendige Beleuchtungsstärke

Für eine gegebene Empfindlichkeit und Auflösung eines Bildsensors kann man für den sich oft ergebenden Fall eines bewegten Objekts (Fließband) den Zusammenhang zwischen notwendiger Beleuchtung und Geometrie der Aufgabenstellung in eine Formel packen.

Man startet zunächst mit dem Abbildungsmaßstab, der sich aus der Feldhöhe, der Pixelgröße ps und der Anzahl Pixel N in eine Richtung berechnet:

$$\beta' = -\frac{ps \cdot N}{y_{max}}, \tag{7.14}$$

Um einen Vollausschlag einer 8-Bit Kamera, also 255, der *Sensitivität* S (Angabe in Digitalwerten („Digital Numbers") pro Pixel pro nJ) zu erzielen benötigt man bei der Belichtungszeit τ eine sensorseitige Bestrahlungsstärke von

$$E_{CCD} = \frac{255}{S \cdot \tau} \tag{7.15}$$

Gesetz erklärt (konstantes L bei Flächenverkleinerung führt dazu, dass auch I gemäß $L := I/A$ verkleinert sein muss.

D.h., wenn man den Sensor 1 Sekunde lang mit E_{CCD} belichtet, dann ergibt sich als ausgegebener Digitalwert $E_{CCD} \cdot S \cdot 1\,s$.

Basierend auf einer geforderten Schärfentiefe δz kann man die Blendenzahl berechnen:

$$K = \beta'^2 \frac{\delta z}{2ps} \tag{7.16}$$

Außerdem ist die Brennweite des Abbildungsobjektivs durch den Abbildungsmaßstab und den geforderten Arbeitsabstand a gegeben:

$$f = \beta' \cdot a \tag{7.17}$$

Dadurch können wir den maximalen Feldwinkel berechnen:

$$\omega = \arctan \frac{y_{max}}{2a} \tag{7.18}$$

Wir lösen

$$E_{CCD} = \pi L \, \mathrm{NA}'^2 \cos^4 \omega' \tag{7.19}$$

nach L auf und nutzen für die Lambertsche Oberfläche

$$E_o = \pi L \tag{7.20}$$

und erhalten

$$\boxed{E_{obj} = \frac{\pi}{80.000} \frac{(a^2 + y_{max}^2) \cdot N^4 \cdot ps^2 \cdot \delta z^2}{a^4 \, S \cdot y_{max}^4 \cdot \tau}} \tag{7.21}$$

Beispiel: Sie nutzen eine hochwertige Kamera (Dalsa Pantera) zur Abbildung eines Objekts (100 x 100 mm), das sich auf einem Fließband bewegt. Aufgrund der Bewegung sollte die Belichtungszeit kurz gewählt werden (1 ms). Sie nutzen 1000x 1000 Pixel mit einer Pixelgröße von 12 x 12 μm . Der Arbeitsabstand beträgt 1 m. Aufgrund des Fließbands (Wackeln) muss die Schärfentiefe mindestens 5 mm betragen. Berechnen Sie die notwendige Bestrahlungsstärke auf dem (streuenden) Objekt.

Lösung:

$$E_e = \frac{\pi}{80000} \frac{(1^2 + 0.1^2)1000^4 \cdot (12 \cdot 10^{-6})^2 0.005^2}{1^4 \cdot 20 \cdot 0.1^4 \cdot 0.001} = 0.071\,\mathrm{W/m^2} \tag{7.22}$$

7.5 Advanced: Wechselwirkung des Lichts mit dem Objekt

Licht, das auf Atome oder Moleküle fällt, regt die Elektronen der Atome auf höhere Energieniveaus an. Nach einer gewissen — in der Regel extrem kurzen — Zeit wird die Anregungsenergie wieder abgegeben. Dies kann in der Form von Wärme erfolgen. In diesem Fall wird also Lichtenergie in Wärmeenergie umgesetzt, man spricht von *Absorption*.

Wenn dagegen die Anregungsenergie wieder direkt in Lichtenergie umgewandelt wird, dann ergibt sich Streuung, Reflexion oder Transmission des Lichts.[12]

Fällt Licht auf eine Oberfläche, dann wird in der Regel also ein Teil des Lichts absorbiert, ein Teil wird transmittiert bzw. dringt in das Material ein und ein Teil wird von der Oberfläche wieder ausgesendet.

Die modellhafte Beschreibung dieser Anteile durch die Modelle „Reflexion" oder „Streuung" sind Idealisierung der sehr komplexen Wechselwirkung der Lichtteilchen mit den Atomen der Oberfläche. Dennoch sind diese Modelle sehr nützlich, um phänomenologisch das Verhalten des Lichts an den Oberflächen zu beschreiben.

Abb. 7.5.5 verdeutlicht die Zusammenhänge. An einer optisch glatten Oberfläche, d.h. einer Oberfläche deren Rauheit deutlich kleiner als die Wellenlänge des Lichts ist, wird das Licht komplett reflektiert und gebrochen. Streuung findet nicht statt, d.h. die Richtung des Lichts nach der Oberfläche ist eindeutig durch das Reflexionsgesetz und das Brechungsgesetz gegeben. Sobald die Oberfläche eine gewisse Rauheit aufweist, ergibt sich zusätzlich eine gestreuter Lichtanteil, d.h. zufällige, nicht vorhersehbare Lichtrichtungen treten zusätzlich auf.

Natürlich muss aus Gründen der Energieerhaltung die Summe aus gestreutem, reflektiertem und transmittierter Energie gleich der Energie des einfallenden Lichts

[12] Bei einer teilweisen Umsetzung in Wärme spricht man Fluoreszenz oder Phosphoreszenz. Kurzwelliges Licht wird in langwelliges Licht umgesetzt.

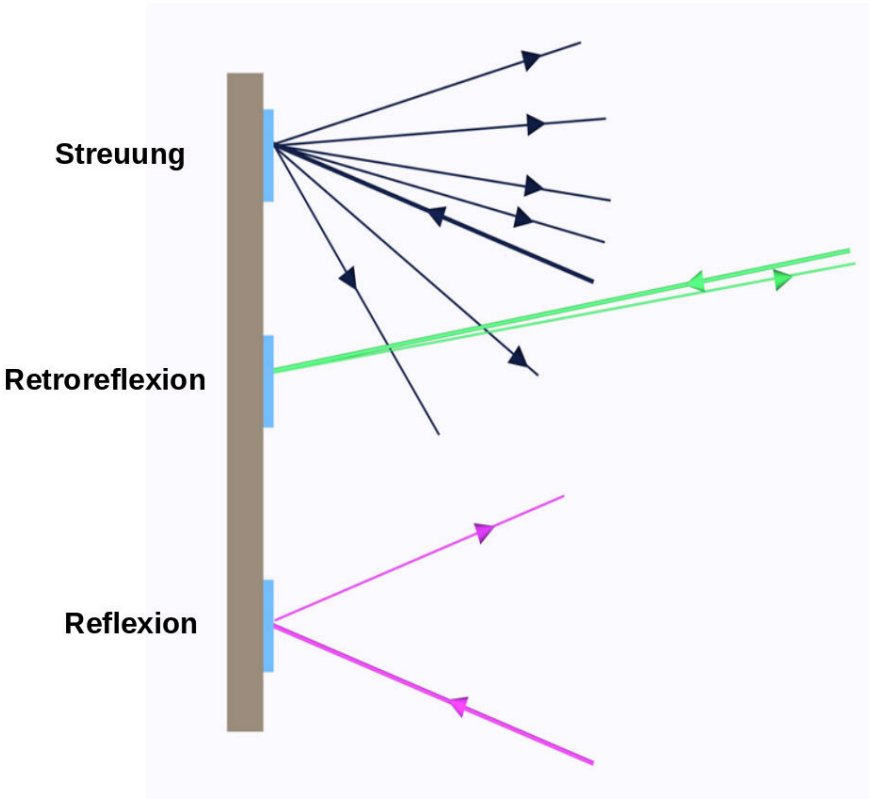


Abbildung 7.5.5: Möglichkeiten der Interaktion von Licht an technischen Oberflächen: Streuung, gerichtete Reflexion und Retroreflexion

sein.

Die winkelabhängige Lichtverteilung nach der Interaktion mit der Oberfläche wird über die *bidirektionale Reflektanz* (bidirectional reflectance distribution function, *BRDF*)[13] beschrieben. Für optisch glatte Oberflächen bei dielektrischen Medien (z.B. Glas) kann die Stärke der beiden Lichtanteile (Reflexion und Brechung) direkt durch die „Fresnelgleichungen" berechnet werden. Für raue Oberflächen sind verschiedene Modelle zur Berechnung der BRDF gebräuchlich.

Für Oberflächen mit einem nennenswerten Anteil an gerichteter Reflexion kann (im Sinne einer Faustregel) nach [36] der Anteil des insgesamt gestreuten (und nicht reflektierten) Lichts (*Total integrated scatter* abgeschätzt werden:

$$TIS = 1 - \exp(\frac{2\pi\,\Delta n\,\sigma_{RMS}\cos\theta}{\lambda})^2 \tag{7.23}$$

Dabei bezeichnet σ_{rms} die RMS Rauigkeit der Oberfläche. Δn ist 2 für den Übergang von Metall (n=-1) nach Luft (n=1) und θ is der Einfallswinkel (von der

[13] Alternativ: BSDF, bidirectional scatter distribution function

Normale aus gemessen).

Sehr gut poliertes Glas erreicht ein σ_{RMS} von besser 1 nm und für optische Spiegel sollte idealerweise so ein TIS von ca. 1 Promille erreicht werden. Eine genauere Beschreibung wird letztlich durch die BRDF gegeben.

Die BRDF ist vom Material und der Oberflächenstruktur aber auch vom Einfallswinkel abhängig (s. Abb. 7.5). Letztlich gibt die BRDF somit an, welche Strahldichte im Verhältnis zur einfallenden Bestrahlungsstärke sich winkelabhängig ergibt. Verschiedene Modelle sind verfügbar, um die BRDF näherungsweise zu berechnen. In der Praxis wird aber oft nur eine grobe Unterteilung in einen gerichteten und einen komplett diffusen Anteil vorgenommen.

Die Absorption der Oberfläche kann natürlich nicht im allgemeinen beschrieben werden sondern ist extrem vom Material abhängig. Dabei ist nicht nur das Grundmaterial selbst zu beachten sondern auch Schichten, die sich auf dem Material bilden. Typische Beispiele sind Oxidschichten, die sich durch Reaktion mit der Umgebungsluft ergeben. Einige typische Werte für verschiedene Materialien finden sich in [27] Kap. 3.

7.6 Spektralbereiche und Farbe

Weite Bereiche des elektromagnetischen Spektrum können für Bildgebungen genutzt werden. Die überwiegende Anzahl von Anwendungen findet sich im visuellen Spektralbereich, also zwischen ca. 400 und 700 nm[14]. Leistungsfähige Detektoren sind in diesem Spektralbereich kostengünstig mit hohen Pixelzahlen verfügbar und wir können eben direkt, das was wir sehen, mit dem in Verbindung bringen, was eine Kamera „sieht".

Aber auch andere Spektralbereiche lassen sich nutzen:

Röntgenstrahlung: einige nm. Wird nicht im Sinne von Abbildungen, aber zur Durchleuchtung eingesetzt.

Ultraviolettes Licht: 100 - 380 nm (meist genutzt: 250 - 380 nm), wird vor allem zur *Fluoreszenzanregung* genutzt. Organische Moleküle (z.B. *Öle*, *Fingerabdrücke*) fluoreszieren unter ultravioletter Bestrahlung und können so sichtbar gemacht werden.

sichtbares Licht: 380 - 780 nm, (meist genutzt: 420 - 780 nm)

Nahinfrarot-A: 780 - 1400 nm, Beachte: *Silizium*-basierte Sensoren arbeiten aber nur bis ca. 1100 nm.

Short Wave Infrared SWIR: 1100 - 1800 nm (nicht in der der DIN enthalten)

Nahinfrarot-B: 1400 - 3000 nm, insbesondere zur Visualisierung von Temperaturen.

Mittleres Infrarot: 3000 - 50.000 nm, insbesondere zu Visualisierung von Temperaturen

Für die Mehrzahl der Anwendungen in der industriellen Bildverarbeitung wird jeweils nur ein Spektralbereiche betrachtet, d.h. es werden nicht unterschiedliche Spektralbereiche unterschiedlich detektiert bzw. prozessiert. Oder in einfachen Worten: Meist wird „die Farbe" nicht beachtet denn bei vielen Objekten (vor allem Metallteilen) ist in der Farbe praktisch keine relevante Information kodiert.

[14] Nach der DIN Norm 5031 genauer: 380 bis 780 nm

Mit dem zunehmenden Trend, Consumersensoren und Consumeroptiken industriell einzusetzen und der Möglichkeit, auch die gestiegene Informationsmenge von Farbsensoren in kurzer Zeit auszuwerten, wird allerdings auch der Anteil an Farblösungen in Zukunft vermutlich weiter deutlich zunehmen.

Für Anwendungen zur Beurteilung bzw. Überprüfung von Farben ist ein möglichst kontinuierliches Spektrum sinnvoll. Typischerweise werden also thermische Quellen (meist Halogen) eingesetzt. Teilweise werden auch spezielle Kombinationen verschiedener LEDs verwendet. Spezifiziert wird die Qualität zur Farbfähigkeit durch den sogenannten *Farbwiedergabeindex* bzw. *Color Rendering Index* (CRI). Eine Lichtquelle, die dasselbe Spektrum wie eine angegebene Referenzlichtquelle (z.B. thermischer Strahler mit einer bestimmten Temperatur) erzielt hat demnach einen Index von 100 — ein Wert, der näherungsweise von Glühbirnen erreicht wird. Für die Beleuchtung von Innenräumen wird z.B. ein Farbwiedergabeindex besser 80 verlangt, was nicht von allen Leuchtstoffröhren erreicht wird.

In der DIN 6169 sind 14 verschiedene Spektren definiert, die als Basis für die Berechnung der Abweichung dienen. Dementsprechend kann auch mit überlagerten Peaks im Spektrum ein sehr guter CRI ($> 95\%$) erzielt werden (z.B. bei Normlichtquellen, die durchaus auch auf Leuchtstoffröhren basieren können). Man sollte sich hier aber bewusst sein, dass nicht alle technischen Anwendungen (z.B. Kalibrierung von Hyperspektralsystemen oder Spektrometern) damit sinnvoll angegangen werden können. Der Farbwiedergabeindex ist eben letztlich für den menschlichen Beobachter definiert.

Die Farbbeurteilung selbst ist ein komplexes Thema. Auch für typische farbfilterbasierte (meist Bayer) Bildsensoren ist eine genaue Farbmessung schwierig, denn die Bildsensoren nutzen jeweils spezielle unterschiedliche Spektren für die Filter und auch die Bildsensoren selbst haben eine spektrale Empfindlichkeit bzw. Quanteneffizient. Die Umrechnung auf normierte Farbräume (CIE-Normvalenzsystem und daraus abgeleitete Räume) ist daher alles andere als trivial bzw. eben nur näherungsweise möglich.

Neben der Fähigkeit, den Sensor gut zur Farbbeurteilung einsetzen zu können ist eine möglichst hohe Lichteffzienz für viele Anwendungen auch wichtig, so dass hier Kompromisse eingegangen werden müssen. Einen guten Einstieg findet man in [34].

Nahinfrarotlicht kann unter Umständen Unterschiede sichtbar machen, die mit dem menschlichen Auge oder RGB-Kameras kaum sichtbar werden. Zu unterscheiden ist der Wellenlängenbereich bis ca. 1100 nm in dem konventionelle Silizium-basierte Sensoren eingesetzt werden können und der Bereich oberhalb von 1100 nm.[15]. Dort sind Bildsensoren (meist InGaAs-basiert oder InSb-basiert) mit höherer Auflösung leider vergleichsweise teuer.

Während Kühlen des Detektors eigentlich immer die Sensorperformance mittels Rauschverminderung verbessert, kann sie im Wellenlängenbereich oberhalb von 1000 nm bei Silizium-basierten Detektoren zu Problemen führen. Die durch die Kühlung verringerte Ladungsträgerbeweglichkeit führt zu einer effektiven Reduzierung der Quanteneffizienz. Somit sinkt die Empfindlichkeit.

Auf den wichtigen infraroten Spektralbereich gehen wir in einem gesonderten Abschnitt (s. Abschnitt 9.12) näher ein.

[15]Nicht alle Bildsensoren, die auf Silizium basieren haben einen entsprechenden spektralen Empfindlichkeitsbereich, denn oftmals werden IR- oder UV-Sperrfilter eingesetzt. Es sind also jeweils die Datenblätter der jeweiligen Sensoren zu konsultieren. Außerdem ist die Quanteneffizienz bei 1100 effektiv Null, d.h. in der Praxis macht es wenig Sinn höher als 1050 nm zu gehen)

Kapitel 8

Beleuchtung

In diesem Kapitel werden Sie unter anderem lernen

- welche verschiedenen Lichtquellen und Bauformen von *Lichtquellen* in der Bildverarbeitung eingesetzt werden,
- wann Sie welche Beleuchtungsgeometrie verwenden,
- welche Parameter bei der Beurteilung einer Beleuchtung relevant sind,
- welche Vor- und Nachteile unterschiedliche Lichtquellen aufweisen,
- was bei der Wahl der Wellenlänge zu beachten ist,
- wie man Fremdlichtanteile unterdrücken kann,
- wie man Fluoreszenz und Polarisation einsetzen kann, um die Bildgebung zu verbessern.

Industrielle Bildverarbeitung ist mehr als Abbildungsoptik, Kamera und Software. Von zentraler Bedeutung, aber oftmals unterschätzt, ist die Beleuchtung. Eine gut gewählte *Beleuchtung* ist für viele Aufgaben essentiell, um eine Bildverarbeitungsanwendung erfolgreich zu implementieren. Eine sinnvolle Wahl der Beleuchtung entscheidet letztlich sehr stark über die Kosten eines BV-Systems und über den Schwierigkeitsgrad der softwaretechnischen Umsetzung.

Generell gilt, dass die Beleuchtung so gewählt werden sollte, dass die interessierenden Merkmal „gut" (in der Regel mit hohem Kontrast) im Bild erscheinen während gleichzeitig die nicht-interessierenden Merkmal möglichst wenig sichtbar sein sollten. Die Geometrie der Beleuchtung sowie die Wahl der Wellenlänge spielen hierfür eine zentrale Rolle.

Darüber hinaus ist durch geeignete Wahl der Beleuchtung dafür zu sorgen, dass genügend Licht auf den Bildsensor fällt und der Anteil an *Störlicht* (z.B. Umgebungslicht) vernachlässigbar ist. Letztlich muss es dabei das Ziel sein, die optische Leistung der Lichtquelle möglichst effizient zu nutzen. Eine erhöhte Leistung der Lichtquelle führt zu höheren Kosten oder einer beschleunigten Alterung und ganz generell zu Problemen.

Neben der sinnvollen Wahl der eigentlichen Lichtquelle sind natürlich die Beleuchtungsoptik aber auch verschiedene Techniken (z.B. eine Pulsung zur Fremdlichtunterdrückung) wesentlich zur Auslegung des Beleuchtungssystems.

8.1 Beleuchtungsgeometrien und Beleuchtungsoptik

Es ist nicht im geringsten ausreichend, die Beleuchtungsquelle für die Beleuchtung zu wählen. Absolut entscheidend sind sowohl die Geometrie der Gesamtanordnung des BV-Systems (inklusive Objekt) als auch die Optik, die für die Beleuchtung verwendet wird.

Es sei nochmals daran erinnert, dass es die zentrale Aufgabe des Beleuchtungssystems ist, für relevante Merkmale einen möglichst hohen abbildungsseitigen Kontrast zu generieren während irrelevante Merkmale möglichst kontrastarm abgebildet werden sollten.

Abbildung 8.1.1 verdeutlicht dies anhand einer sehr oft eingesetzten Beleuchtungsmethodik, der sogenannten *Dunkelfeldbeleuchtung*. Das Beleuchtungslicht wird bei einer ebenen, fehlerfreien Oberfläche weg von der Kamera gestreut bzw. reflektiert. Lediglich an lokalen Defekten (z.B. *Kratzern*) ergibt sich Streuung in Richtung der Kamera. Die Defekte werden daher als helle Strukturen auf einem ansonsten dunklen Bild sehr gut sichtbar. Ein Beispiel für die Hervorhebung von

Kanten zeigt Abb. 8.1.2.

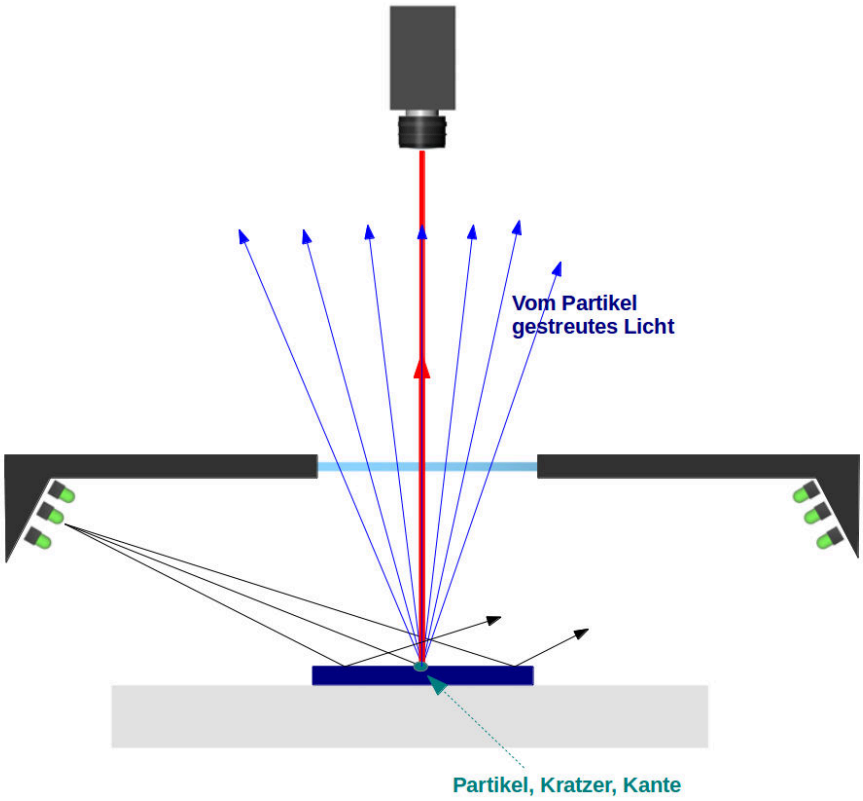


Abbildung 8.1.1: Dunkelfeldbeleuchtung mittels Ringlicht. An der ebenen Oberfläche reflektiertes Licht tritt nicht in die Eintrittspupille der Abbildung ein. Daher ergibt sich an den entsprechenden Stellen ein dunkles Bild. Defekte und Kanten führen zur Streuung und ein Teil des Beleuchtungslichts findet so den Weg in das Kameraobjektiv bis zu Bildebene.

(a) diffuse Beleuchtung

(b) näherungsweise Dunkelfeldbeleuchtung durch tiefes Ringlicht

Abbildung 8.1.2: Beleuchtung einer Dose mit einem Ringlicht

Selbst diese an sich klare und einfache Geometrie ist in der Praxis nicht immer un-

problematisch. Welcher Winkel bzw. welche Winkelverteilung ist ideal, wie wirken sich Fehler in der Positionierung des Objekts aus und was passiert bei komplexeren Objektgeometrien? Vor allem hängt die Streuung an der lokalen Struktur massiv von den Details der Struktur ab. In der Praxis muss man daher meist experimentell vorgehen und unter verschiedenen Richtungen beleuchten und die sich ergebenden Bilder analysieren. Ideal hierzu ist z.B. eine Schwanenhalsleuchte. Basierend auf den Ergebnissen zeigt sich dann, welche Gesamtbeleuchtungsgeometrie günstig ist.

Nach diesem Beispiel der Dunkelfeldbeleuchtung wollen wir etwas allgemeiner werden. Grundsätzlich ist zwischen einer Beleuchtung in *Auflicht* und (seltener) in *Durchlicht* zu unterscheiden.

8.1.1 Durchlicht

Durchlichtbeleuchtungen sind sehr gut geeignet, die (Außen-)Konturen im Sinne einer Geometrieprüfung zu untersuchen. Zu beachten ist dabei allerdings, dass an den Kanten bei einfachen Beleuchtungssystemen Reflexionen auftreten können, die unter Umständen das Ergebnis der Messung deutlich verfälschen.

Hinsichtlich der Vermeidung solcher Effekte sollte eine *telezentrische Beleuchtungsoptik* eingesetzt werden. Hierdurch wird sichergestellt, dass das Objekt nur von parallelem Licht beleuchtet wird (Abb. 8.1.3). In diesem Fall muss dann allerdings auch eine telezentrische Abbildungsoptik verwendet werden. Abbildungs- und Beleuchtungsoptik müssen exakt aufeinander justiert werden, um Fehler zu vermeiden.

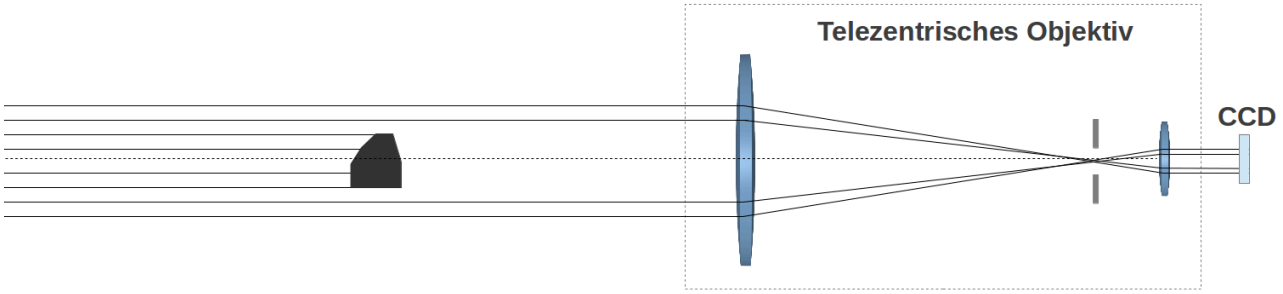


Abbildung 8.1.3: Telezentrische Durchlichtbeleuchtung mit telezentrischer Abbildungsoptik

Eine weitere wichtige Anwendung von Durchlichtbeleuchtungen sind (halb-)transparente Proben, wie sie sehr häufig in der biomedizinischen Mikroskopie auftreten. In diesem Fall ist die Durchlichtbeleuchtung der klassische Standard. Man unterscheidet in diesem Fall zwischen kritischer Beleuchtung, bei der die Quelle in die Objektebene abgebildet wird, und *Köhlerscher Beleuchtung*, bei der die Quelle in einer Fourierebene des Objekts liegt.

Wir wollen auf diese Beleuchtungen nicht weiter eingehen und verweisen z.B. auf [39] da die überwiegende Anzahl an Anwendungen der industriellen Bildverarbeitung ohne Zweifel in *Auflicht* erfolgt.

8.1.2 Auflicht

Grundsätzlich ist bei Auflichtbeleuchtungen zwischen *diffusem* und *gerichtetem Licht* zu unterscheiden. Diffus bedeutet, dass ein Objektpunkt aus unterschiedlichen Richtungen beleuchtet wird. Je mehr Richtungen, desto diffuser die Beleuchtung (s. Abb. 8.1.4).

Eine ideal *diffuse Beleuchtung* ist ein komplett wolkenbedeckter Himmel, natürlich eine für die industrielle Bildverarbeitung nicht realistische Beleuchtungsquelle. Da das Licht idealerweise gleichzeitig aus allen Richtungen kommen soll benötigt man eine ausgedehnte Quelle. Sehr diffuses Licht erzielt man mit einer *Dombeleuchtung* (Abb. 8.1.4), allerdings ist, vor allem bei ausgedehnten Objekten, der Platzbedarf (und damit dann auch die Kosten) sehr hoch. In diesem Fall werden zunehmend sogenannte *Flatdomes* verwendet (s. Abb. 8.1.5) oder aber man verwendet mehrere Flächen-, Linien- oder Punktstrahler (unter Umständen in Verbindung mit einer diffus streuenden Fläche in Transmission oder Reflexion).

Man kann nicht generell sagen, welcher Grad der Diffusion für eine bestimmte Erfassung von Merkmalen am besten geeignet ist. Glanzlichter und damit Überstrahlungen werden aber in jedem Fall vermieden. Dadurch wird oft die Auswertung vereinfacht, insbesondere, wenn der Reflexionsanteil einer Oberfläche hoch ist und leichte Verformungen der Oberfläche erlaubt sind.

Abb. 8.1.7 zeigt ein typisches Beispiel der Verbesserung, die durch eine diffuse Dombeleuchtung erzielt werden kann. Zur Kontrolle der Geometrie der Dose ist

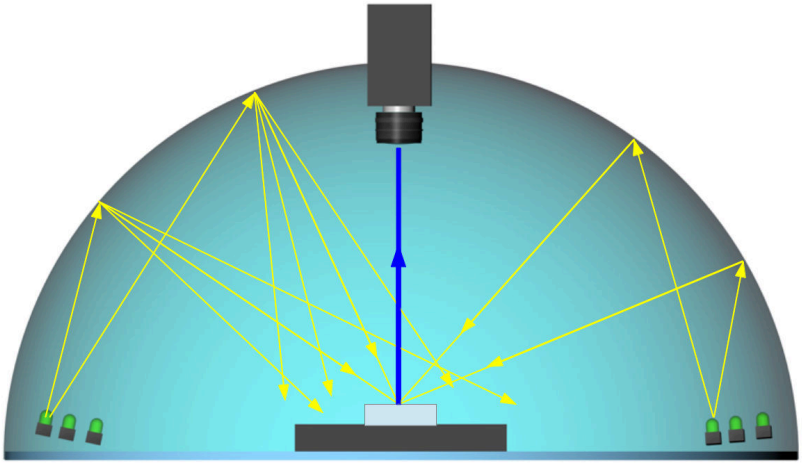


Abbildung 8.1.4: Diffuse Beleuchtungsgeometrie in Auflicht mit einer Dombeleuchtung. Bei der diffusen Beleuchtung trifft Licht aus allen möglichen Richtungen auf jede Stelle des Objekts. Bei der (total) gerichteten Beleuchtung ergibt sich an jeder Position nur eine Beleuchtungsrichtung. Oft wird ein Dom mit streuender Oberfläche verwendet, um eine möglichst ideale diffuse Beleuchtung zu realisieren.

das diffuse Bild deutlich besser geeignet, denn die Reflexe im Bild mit teilweise gerichteter Beleuchtung erschweren die Auswertung, insbesondere da die Position und Stärke der Reflexionsanteile mit leichten Formänderungen variiert. Wenn dagegen gerade die Formänderung detektiert werden soll ist eine gerichtete Beleuchtung natürlich zu bevorzugen. Es hängt also immer davon ab, was detektiert bzw. gemessen werden soll.

Ringbeleuchtungen sind weniger stark gerichtet als einzelne Spotstrahler, denn zumindest fallen unterschiedliche radiale Richtungen auf das Objekt. Dennoch sind sie ohne zusätzliche Diffusoren teilgerichtet. Selbst mit zusätzlicher Diffusorschicht (wie in Abb. 8.1.6 verwendet) ist der gerichtete Anteil stark, so dass deutliche Reflexionsstörungen auftreten können.

Natürlich kann die nicht-diffuse Beleuchtung auch Vorteile ergeben. Nämlich dann, wenn gerade (geänderte) Reflexionen und Streuungen sichtbar gemacht werden sollen. Abb. 8.1.8 zeigt ein Beispiel eines Rohres mit einer deutlicheren und einer sehr minimalen Eindellung. Unter diffuser Beleuchtung ist die Eindellung nicht sichtbar, unter gerichteter Beleuchtung aber verhältnismäßig gut.

Die Abbildung zeigt sehr gut, dass selbst bei passender Beleuchtung (hier: gerichtet) eine Auswertung durchaus anspruchsvoll werden kann wenn die Ober-

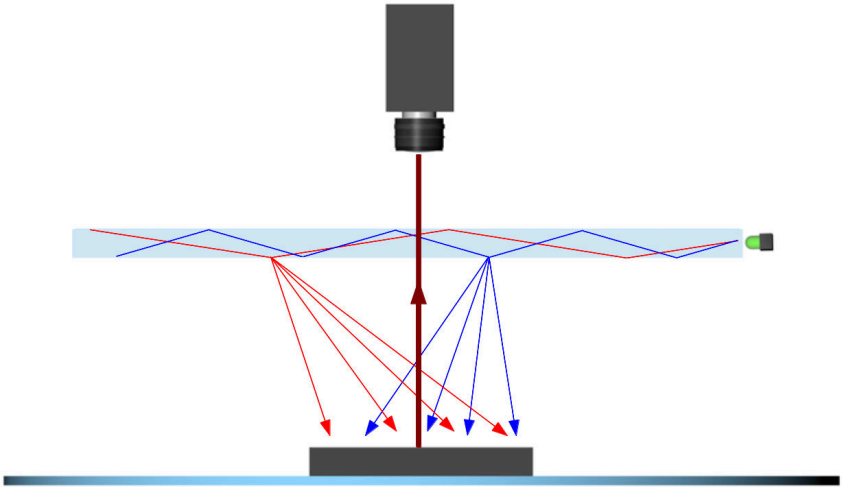
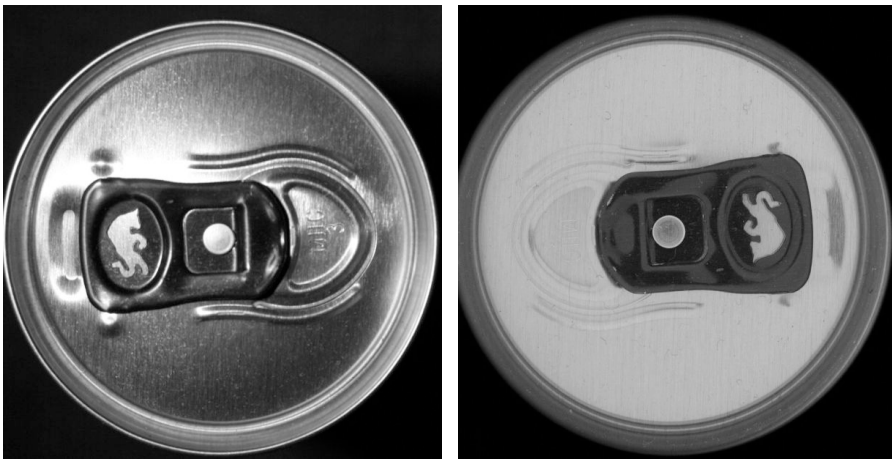


Abbildung 8.1.5: Flatdom Beleuchtung. Seitlich eingekoppeltes Licht wird fortlaufend an der Lichtleiterplatte reflektiert. Kleine Unregelmäßigkeiten der Platte sorgen dafür, dass ein Teil des Lichts nicht reflektiert sondern gestreut wird und so das Objekt unter verschiedenen Winkeln erreicht. Durch die Wahl der Höhe des Flatdomes kann das Winkelspektrum des Beleuchtungslichts variiert werden. Die Kamera inspiziert die Oberfläche durch ein Loch in der Lichtleiterplatte.

fläche eines korrekten Bauteils gewollte Variationen aufweist (hier kleiner Eindellungen). Andererseits zeigt das Beispiel abermals, dass bei falsch gewählter Beleuchtung (hier: diffus) überhaupt keine Chance besteht, eine stabile Auswertung zu realisieren. Eine oberflächige Betrachtung von Abb. 8.1.8 (b) würde zu dem Schluss führen, dass ein dreidimensionales Messverfahren zur Detektion entsprechender Dellen eingesetzt werden muss.

Für die Hervorhebung von Details in bestimmten Orientierungen ist natürlich die *Beleuchtungsrichtung* wesentlich. Abb. 8.1.9 zeigt als Beispiel eine Aufnahmeplatte mit angerissenen Positionen. Diese werden in *gerichteter Beleuchtung* dann besonders gut sichtbar, wenn die Beleuchtungsrichtung senkrecht zur Orientierung der Struktur gewählt wird. Von *koaxialer Beleuchtung* spricht man, wenn die gerichtete Beleuchtungsrichtung parallel zur optischen Achse der Abbildung erfolgt. Die teilgerichtete Beleuchtung mit einem (tiefen) Ringlicht führt in diesem Fall ebenfalls zu einer guten Sichtbarkeit der Linien.

Idealerweise würde eine diffuse Beleuchtung eher verhindern, dass diese Linien überhaupt sichtbar werden (je nach Aufgabenstellung kann das schlecht oder gut sein). Es zeigt sich deutlich, dass dies hier nicht der Fall ist. Die Wechselwirkung von Licht mit *Kratzern* ist sehr komplex und für die sich ergebenden Helligkeits-



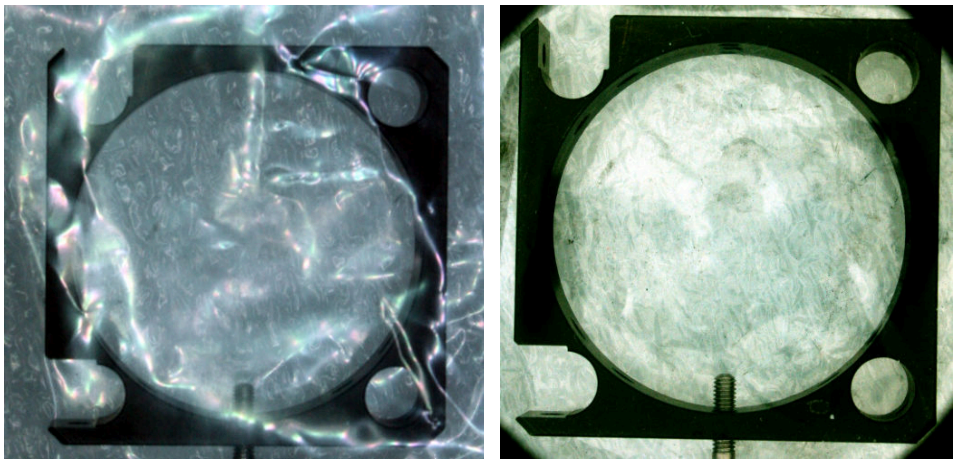
(a) gerichtete Beleuchtung (b) diffuse Dombeleuchtung

Abbildung 8.1.6: Diffuse Beleuchtung vermeidet bzw. reduziert Glanzlichter und kann daher insbesondere bei Oberflächen mit ausgeprägten Reflexionsanteilen das Bildergebnis deutlich verbessern.

unterschiede im Bild sind neben dem Kratzer selbst auch die Rauheit der Umgebung und die Kameraöffnung entscheidend. Es ist also oftmals schwierig, sicher vorherzusagen, welchen Effekt eine bestimmte Beleuchtung im Bild letztlich ergibt. Selbst eine Simulation der entsprechenden Lichtwechselwirkung mit der *Mikrostruktur* ist nur eingeschränkt möglich. Entscheidend ist also beim Stand der Technik, dass eine Mischung aus heuristischem und experimentellem Vorgehen notwendig ist, um die Beleuchtungsgeometrie passend zu wählen.

Generell gilt, dass eine gerichtete Beleuchtung auch zu vom Objekt kommenden gerichtetem Licht führt wenn das Objekt nicht sehr diffus streuend ist. Die Lichtrichtung hängt dabei natürlich aufgrund des Reflexionsgesetzes von der lokalen Oberflächenneigung des Objekts ab. Dabei ist nicht nur die makroskopische Topografie sondern auch die *Mikrotopografie* entscheidend. Die Mikrotopografie führt zu einer „Streuung" an der Oberfläche. Aus einer definierten Richtung einfallendes Licht wird so in verschiedene Richtungen abgelenkt bzw. gestreut. Welcher Anteil des Lichts im statistischen Mittel in welche Richtung streut wird durch die bereits besprochene BRDF (vgl. Abschnitt 7.5) spezifiziert.

Bei Objekten mit einem stark gerichteten Reflexionsanteil und gerichteter Beleuchtung kann es je nach Geometrie von Objekt, Beleuchtung und Abbildung zu einzelnen Reflexen im Bild kommen. An diesen Stellen wird der Bildsensor meist komplett überstrahlt und in der Regel wird man die Reflexe vermeiden wollen. In diesem Fall ist eine (mehr oder weniger ausgeprägte) diffuse Beleuchtung vorteil-



(a) Ring-Beleuchtung (teilgerichtet) (b) Diffuse Dombeleuchtung

Abbildung 8.1.7: Bauteil in verpacktem Zustand (Klarsichtfolie). Die gerichtete Beleuchtung durch ein Ringlicht führt in (a) zu einem deutlichen Störlichtanteil in Form von Reflexionen, die eine automatisierte Auswertung unnötig erschweren. Die diffuse Dombeleuchtung vermeidet den Nachteil.

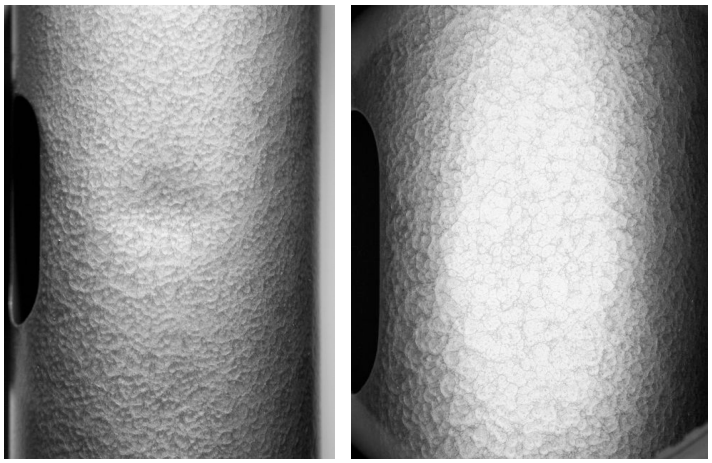
haft.

Andererseits kann die gerichtete Beleuchtung bei solchen Objekten sehr gut eingesetzt werden, um Abweichungen von ansonsten ebenen Objekten zu finden. Das klassische Beispiel ist die bereits in der Einleitung dieses Abschnitts eingeführte *Dunkelfeldbeleuchtung*. Da die Orientierung der Defekte meist zufällig ist, erfolgt die schräge Beleuchtung in der Regel von allen Richtungen und wird dann mit einem *Ringlicht* realisiert (s.Abb. 8.1.10). Der Einfallswinkel wird über die axiale Position des Ringlichts variiert. Je größer das zu inspizierende Objekt ist, desto aufwändiger ist ein solcher Ringlichtansatz.

Durch Variation des Ringlichtabstandes von dem abzubildenden Objekt kann das Bild deutlich variiert werden. Der Dunkelfeldcharakter geht mehr und mehr verloren (Abb. 8.1.10 und 8.1.11).

Oft einfacher ist der Einsatz von einzelnen, linearen Beleuchtungsmodulen, sogenannten *Linienlichtern*. Das Verhalten ist dann natürlich nicht mehr isotrop, d.h. je nach Orientierung werden Strukturen unterschiedlich auf der Kamera dargestellt. Teilweise kann das aber auch ein Vorteil sein (z.B. Unterdrückung von Bearbeitungsriefenstrukturen).

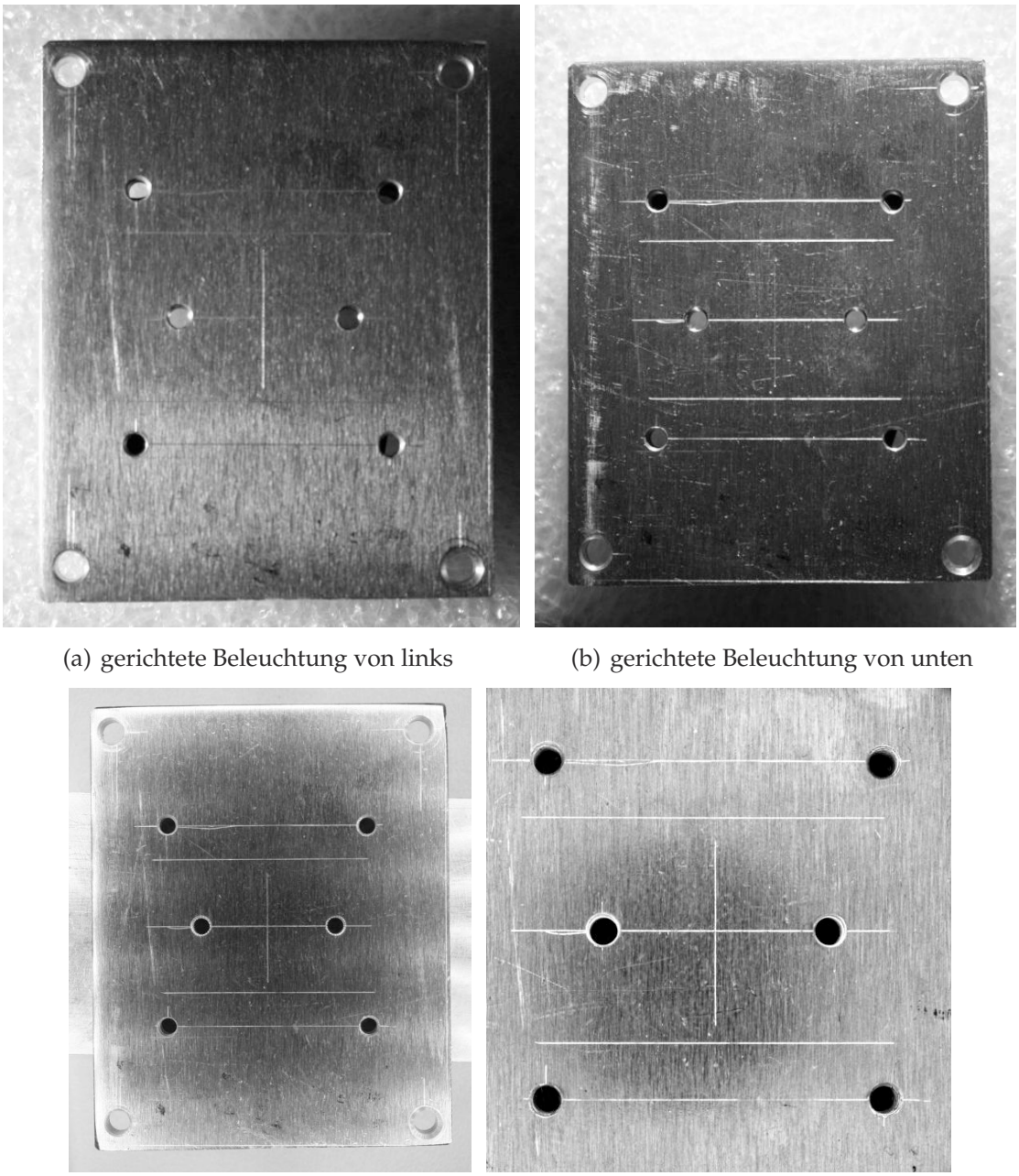
Viele verschiedene Auflichtquellen mit einer teilweise gerichteten Abstrahlung sind verfügbar und bilden einen Kompromiss aus komplett gerichteter und kom-



(a) gerichtete Beleuchtung (b) Diffuse Dombeleuchtung

Abbildung 8.1.8: *Leichte* Eindellungen einer Oberfläche mit Teilreflexionsanteil können unter gerichteter Beleuchtung sichtbar gemacht werden und verschwinden bei diffuser Beleuchtung.

plett diffuser Beleuchtung. Hohe Strahlungsleistungen lassen sich auch unter beengten Platzverhältnissen über *Lichtleiter* zum Objekt bringen.



(a) gerichtete Beleuchtung von links

(b) gerichtete Beleuchtung von unten

(c) Teilgerichtete Beleuchtung mit Ringlicht

(d) Dombeleuchtung

Abbildung 8.1.9: Anrisse auf einer Aluminiumoberfläche bei unterschiedlicher Beleuchtung.

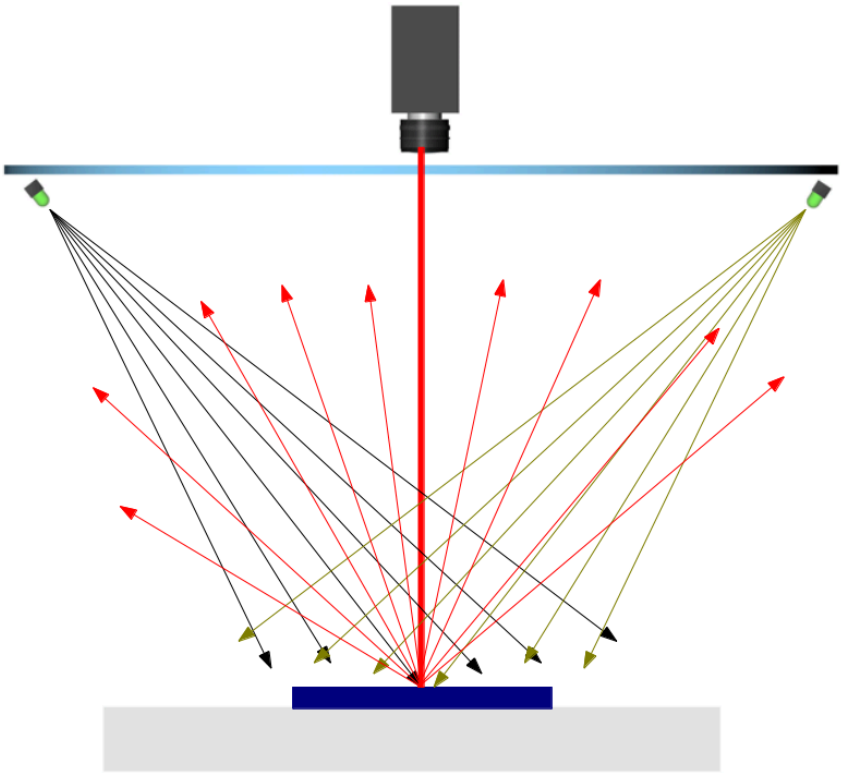


Abbildung 8.1.10: Beleuchtung mittels Ringlicht bei mittlerer Höhe. Der Dunkelfeldcharakter geht beinahe verloren. Geringe Streuungen führen aber zu Aufhellungen im Bild.

(a) Ringlicht in falscher Höhe: starke Störreflexe aufgrund der einzelnen Ringlicht-LEDs

(b) sinnvolle Höhe: die Reflexionen sind reduziert und liegen in wenig störenden Bereichen

Abbildung 8.1.11: Beleuchtung einer Dose mit einem Ringlicht

8.2 Advanced: Linienbeleuchtung

Insbesondere wenn Objekte mit konstanter Geschwindigkeit unter einem Inspektionssystem bewegt werden und hohe Lichtleistungen notwendig sind, ist es sinnvoll, eine eindimensionale Bildaufnahme mit dementsprechend eindimensionaler Beleuchtung zu überdenken. Diese sogenannten Linienbeleuchtungen werden also meist in Kombination mit *Zeilenkameras* eingesetzt.

Die Bewegung des Objekts erfordert zwangsweise eine vergleichsweise kurze Belichtungszeit τ, denn zu lange Belichtungen führen natürlich zu *Bewegungsunschärfen*. Mit einem konventionellen Flächensensor müssen also während der Belichtungszeit Millionen von Pixeln genug Licht bekommen. Dementsprechend muss sehr viel Energie in einen kurzen Beleuchtungspuls gepackt werden.

Bei einem eindimensionalen Sensor ist die effektive Bildaufnahmezeit deutlich verlängert, denn bei z.B. 1000 x 1000 Pixeln im Gesamtbild ist die Gesamtaufnahmedauer dann 1000 τ. Dementsprechend steht letztlich mehr Licht pro Bild zur Verfügung (s. Abb. 8.2.12).

Der Gewinn fällt allerdings in der Praxis etwas geringer aus, denn die Beleuchtung kann nicht exakt nur eine Auflösungszelle der Abbildungsoptik ausleuchten[1]. Außerdem muss bedacht werden, dass eine ungleichförmige Bewegung des Objekts natürlich zu Bildverzerrungen führt. Die Geschwindigkeit des Objekts muss also sehr konstant sein.

Eine Parallelisierung des Ansatzes ist möglich indem mehrere Zeilen parallel beleuchtet und aufgenommen werden.

8.3 Fremdlicht

Ungewünschtes Licht, das nicht vom eigentlichen Beleuchtungssystem generiert wird, ist ein ernstzunehmendes Problem in vielen BV-System. Dieses sogenannte *Fremdlicht* steht nicht unter der direkten Kontrolle vom BV-System. Das Hauptpro-

[1] Abgesehen davon, dass bei leichter Höhenvariation des Objekts aufgrund des Triangulationswinkels zwischen Beleuchtung und Abbildung ein Versatz auftritt der ebenfalls ausgeleuchtet werden muss.

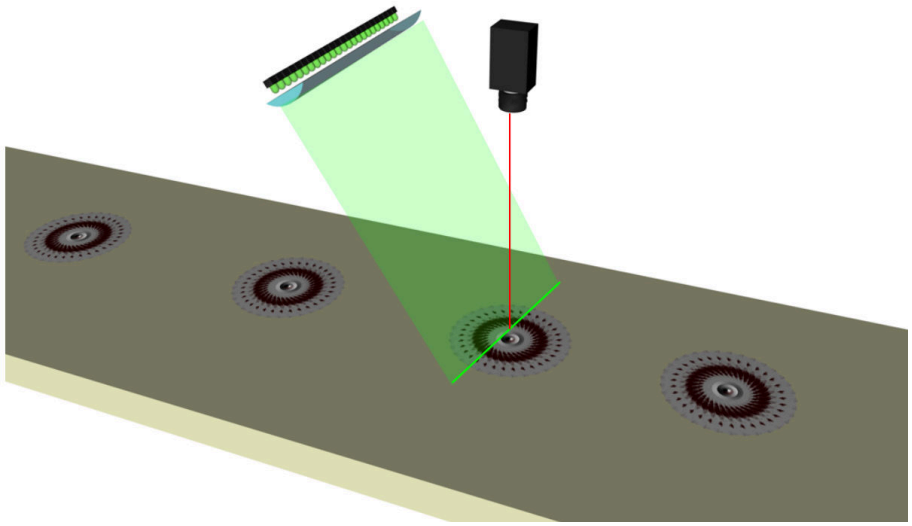


Abbildung 8.2.12: Bei Bewegungen mit konstanter Geschwindigkeit kann die effektive Belichtungszeit pro Gesamtbild deutlich gesteigert werden indem die Zeilen zeitsequentiell aufgenommen werden.

blem ist dabei meist, dass das Fremdlicht in nicht vorhersehbarer Weise zeitlich variiert und damit die Funktion des BV-Systems in erheblichem oder gar fatalem Maße stört.

Es ist daher dafür zu sorgen[2], dass der Fremdlichtanteil deutlich unterhalb des Signallichtanteils liegt. Verschiedene Methoden sind hier gebräuchlich:

Kapselung: Die direkteste Möglichkeit der Vermeidung von Fremdlicht ist, den optisch relevanten Anteil des BV-Systems mit einem lichtdichten Gehäuse zu kapseln. Diese an sich sehr preisgünstige Lösung ist aus verschiedensten Gründen (z.B. Zuführung der Teile, Temperaturerhöhung) leider oft nicht praktikabel.

Pulse: Typischerweise ist das Fremdlicht (näherungsweise) konstant[3]. Daher kann man eine deutliche Reduzierung des prozentualen Fremdlichtanteils erreichen, wenn man es schafft, die für die Beleuchtung verwendete Energie in einem kurzen Puls abzugeben und den Bildsensor nur während der *Pulsung* zu belichten. LEDs und viele Laser sind hierfür geeignete Beleuchtungsquellen, die solch eine Pulsung mit kurzzeitig sehr hoher Leistung ermöglichen.

Natürlich ist eine exakte Synchronisierung von Bildsensor und Beleuchtung

[2] algorithmenabhänig

[3] Eine wichtige und beachtenswerte Ausnahme hiervon ist das durch Leuchtstoffröhren (Fabrikbeleuchtung) induzierte Fremdlicht, das mit 50 Hz periodisch schwankt.

notwendig. Eine gewisse Toleranz wird in der Praxis möglich, wenn der Beleuchtungspuls etwas (z.B. Faktor 2) länger als der Belichtungspuls gewählt wird.

Wellenlänge: Auch die spektrale Abstrahlcharakteristik einer Quelle kann zur effektiven Fremdlichtunterdrückung genutzt werden. Wenn als Beleuchtungsquelle des BV-Systems ein Spektralbereich verwendet wird, der anteilsmäßig im Fremdlicht nur gering vorhanden ist, kann vor dem Bildsensor ein passender Spektralfilter eingesetzt werden, der das Fremdlicht stark reduziert.

Ein typisches Beispiel ist das breitbandige *Sonnenlicht*. Wenn zur Beleuchtung eine vergleichsweise monochromatische LED (Halbwertsbreite z.B. 10 nm) eingesetzt wird, dann kann ein entsprechender Spektralfilter eine Sonnenlichtabschwächung um den Faktor 30 erreichen[4].

Berechnet werden kann die Abschwächung, wenn man ein gewissen Spektrum des Fremdlichts annimmt. Für den Strahlungsfluss nach dem Filter ergibt sich

$$\Phi_e[W] = (\lambda[\mu m] - 0.4)\frac{FWHM[nm]}{10} \;\; @\,3000\,\mathrm{K},\; 1000\;\mathrm{lm} \tag{8.1}$$

$$\Phi_e[W] = FWHM[nm] \cdot 0.02 \;\; @\,6000\,\mathrm{K},\; 1000\;\mathrm{lm} \tag{8.2}$$

Dabei bezeichnet $FWHM$ die spektrale Halbwertsbreite des eingesetzten Filters.

Beispiel: Sie messen Sonnenlicht mit einer Beleuchtungsstärke von 10.000 lux auf den Objekten, die zu inspizieren sind. Natürlich ist diese Beleuchtungsstärke nicht konstant sondern variiert und führt so zu Fehlern bei der Bildauswertung. Daher verwenden Sie eine LED mit einer spektralen Halbwertsbreite von 10 nm zur Beleuchtung und kameraseitig einen entsprechenden Bandpassfilter. Wie stark sollte die durch die LED realisierte Bestrahlungsstärke auf dem Objekt gewählt werden,

[4] Unter der Annahme einer konstanten spektralen Empfindlichkeit des Bildsensors über einen Bereich von 300 nm.

wenn die Beleuchtungsstärke aufgrund der Sonne maximal 10% der LED-Stärke erreichen darf ?

Lösung: Für die Bestrahlungsstärke aufgrund der Sonne (T=6000 K) ergibt sich mit Gl. 8.2 $\Phi_e = 10 \cdot 0.02 = 0.2\,\mathrm{W/m^2}$. Das bedeutet, dass das LED-Beleuchtungssystem mindestens 2 $\mathrm{W/m^2}$ erreichen sollte.

Beispiel: Um die selbe Fremdlichtunterdrückung zu erzielen wollen Sie anstatt des Farbfilters eine Pulsbeleuchtung verwenden. Wie kurz müssen die Pulse (und die entsprechende Belichtungszeit) sein? Gehen Sie von einer Empfindlichkeit der Kamera von 300 nm (Halbwertsbreite) aus.

Lösung: Die 10.000 lx entsprechen $10 \cdot 0.02 \cdot 300 = 60$ $\mathrm{W/m^2}$ bzw. 60 $\mathrm{J/m^2}$ in einer Sekunde. Diese Energie muss entsprechend um den Faktor 300 reduziert werden (60 / 0.2). Dies erreicht man durch die Pulsung. d.h. die Pulslänge bzw. die Belichtungszeit darf 3,3 ms nicht übersteigen.

8.4 Lichtquellen

Bevor wir auf die in der industriellen Bildverarbeitung wesentlichen Lichtquellentypen und ihre Eigenschaften eingehen, wollen wir zunächst generelle Beurteilungskriterien angeben.

Leider wird kaum ein Bereich der Bildverarbeitung so stiefmütterlich behandelt wie die Beleuchtung. Daher fehlen typischerweise — auch bei sehr teuren Beleuchtungssystemen — konkrete technische Daten und es ergeben sich Spezifikation der Art „sehr hell" oder ähnlich. Oftmals finden sich keine Angaben, die über die bloße Nennung der elektrischen Anschlussleistung und der Geometrie des Systems hinausgehen. Letztlich ist der Grund für den eklatanten Informationsmangel, dass die Beurteilung lichttechnischer Größen für die überwiegende Anzahl an Bildverarbeitungsexperten nicht möglich ist und daher entsprechende Daten praktisch nie hinterfragt bzw. verlangt werden.

Sie sollten daraus nicht den Schluss ziehen, dass diese Angaben nicht essentiell wären und es nicht der Mühe lohnt, sich mit den korrekten Angaben zu beschäftigen, denn es wird Ihnen nicht gelingen, eine geeignete Beleuchtung zu spezifizieren, wenn die entsprechenden Kenntnisse nicht vorhanden sind. Vielmehr müssen Sie in der Regel beim Hersteller eines Beleuchtungssystems im Zweifelsfall schlicht die relevanten Daten erfragen.

Generell sind nicht alle Angaben für alle Lichtquellen bzw. Beleuchtungsmodule relevant und die gewünschten Daten hängen davon ab, ob eine zusätzliche Beleuchtungsoptik eingesetzt wird. Im folgenden folgt eine kleine Übersicht der wesentlichen technischen Daten von Lichtquellen:

Abgestrahlte optische Leistung: Meist ist leider nur die elektrische Leistungsaufnahme angegeben. Für die Auslegung des Systems ist aber natürlich zunächst die optisch abgegebene Leistung von Bedeutung. Unter Umständen kann die Anforderung an die CW (kontinuierliche, continuous wave) Leistung reduziert werden, wenn ein gepulster Betrieb möglich ist. Beachten Sie, dass die Lebensdauer und Stabilität positiv beeinflusst werden, wenn Sie die Lichtquelle bei reduzierter Leistung betreiben können.

Möglichkeit zur Pulsung: Sowohl zur Fremdlichtunterdrückung als auch zur Vermeidung von *Bewegungsunschärfen* durch kurze Belichtungszeiten kann es sinnvoll sein, die Lichtquelle synchronisiert mit der Bildaufnahme gepulst betreiben zu können.

Ansteuerelektronik: Insbesondere für die Pulsung ist eine entsprechende Ansteuerelektronik notwendig.

Spektrale Abstrahlcharakteristik: In welchem Spektralbereich wird das Licht emittiert.

Zeitliche Konstanz der Abstrahlung: Viele Lichtquellen haben eine deutliche zeitliche Variation der emittierten Lichtleistung. Besonders stark, aber für viele Anwendungen nicht entscheidend, ist dabei die Variation innerhalb der „Aufwärmphase". Unter Umständen sind die charakteristischen Variationszeiten deutlich kürzer als die Belichtungszeit, so dass die Variation teilweise gemittelt wird.

Alterung und Lebensdauer: Die Lebensdauer ist ein in der industriellen Praxis besonders wichtiger Parameter, denn der Ausfall eines Systems in der Produktion ist oft mit erheblichen Folgekosten verbunden. Dementsprechend ist auf eine geringe Ausfallwahrscheinlichkeit zu achten. Üblicherweise lässt sich die Lebensdauer der Lichtquelle deutlich verlängern indem die Quelle bei einer abgesenkten Leistung betrieben wird. Die Lichtquelle ist in diesem Fall also überdimensioniert auszulegen.

Abstrahlgeometrie: Hinsichtlich der Abstrahlgeometrie ist zu beachten, dass die Objektbeleuchtung ausreichend homogen ist, es kann aber auch entscheidend sein, ob zuviel überschüssiges Licht generiert wird. Durch eine nachgeschaltete Optik lässt sich die Abstrahlcharakteristik natürlich beeinflussen.

Effizienz: Eine mangelhafte Effizienz kann bei Anwendungen mit hoher Lichtleistung zu sehr viel Abwärme und hohen Anforderungen an die Stromversorgung führen. Insbesondere die Abwärme kann zu einer Vielzahl von Sekundärproblemen (z.B. frühzeitige Alterung von Komponenten, thermische Ausdehnung von System und Objekt) führen

Homogenität: Die *Homogenität* der Beleuchtung kann, muss aber nicht wichtig sein. Oftmals ergibt sich sowieso aufgrund der Abbildungsoptik durch Vignettierung eine inhomogene sensorseitige Bestrahlungsstärke, die per Software ausgeglichen wird. Deutliche Inhomogenitäten können aufgrund des eventuell verringerten Signal-Rausch-Verhältnisses in bestimmten Bildbereichen aber zu subtilen Problemen bei der Auswertung führen.

Abwärme und Infrarotlichtanteil: siehe Effizienz.

Color Rendering Index: siehe Abschnitt 7.6

Preis: Bilden Sie eine Gesamtbilanz bei der etwaige Wartungs-/Austauschkosten verschiedener Quellen einbezogen sind. Meist sind Austausch/Wartung sehr viel teurer als die eigentliche Quelle.

Bauform: Die Bauform muss natürlich die gegebenen aufgabenspezifischen Randbedingungen erfüllen

notwendige Spannung, Strom, EMV, Verpolungsschutz: Subtile Probleme können aufgrund elektromagnetischer Störeinstrahlung bei industriellen An-

lagen oder des Inspektionssystems selbst auftreten. Daher ist insbesondere bei hohen (gepulsten) Lichtleistungen und den damit verbundenen hohen elektrischen Leistungen zu prüfen, ob Probleme auftreten können.

Vibrationsempfindlichkeit: *Vibrationen* können bei verschiedenen Leuchtmitteln und Leuchtkörpern zu unterschiedlichen Störungen führen. Diese können vom reinen mechanischen Defekt, über eine Störung der elektrischen Zuleitung (Kontaktprobleme) bis zur Vibration der Lage der Lichtemission (vibrierende Glühwendel) reichen.

Staubschutz: *Verschmutzungen* können sich auf transparenten Flächen der Lichtquelle oder des Beleuchtungssystems niederschlagen. Bei hoher Lichtleistung kann dies durch Absorption des Lichts an den Verschmutzungen zu einem Einbrand mit Schädigung der entsprechenden Flächen oder Brandgefahr führen. Geeignete Vorrichtungen zur Vermeidung der Verschmutzung sind in diesem Fall notwendig.

Wasserempfindlichkeit: Eventuell muss die Bauform der Leuchte so gestaltet sein, dass Spritzwasser toleriert wird. Hierzu ist eine geeignete Kapselung oder Lichtleitung notwendig.

Explosionsgefahr: Lampen, die unter hohem oder tiefem Druck arbeiten führen unter Umständen zu einem erheblichen Schädigungspotential aufgrund einer möglichen Explosion/Implosion.

Hinsichtlich der Bauformen von Lichtquellen werden vor allem folgende Grundtypen unterschieden:

Ringlicht: Sehr gebräuchliche Bauform bei der die Lichtquelle ringartig zentriert zum Objekt und zur Kamera angeordnet wird (Abb. 8.1.10). Das Licht strahlt also radial unter denselben Winkeln auf das Objekt. Die vertikale Position kann in weiten Bereichen selbst gewählt werden[5]. Rein geometrisch ergibt sich durch die Höhe der Quelle und die Gesamtgeometrie ein deutlich unterschiedliches Beleuchtungsverhalten. Es ist weiterhin zu unterscheiden, ob eine zusätzliche Diffusorschicht vor den LEDs verwendet wird.

[5]Bei sehr naher Entfernung zum Objekt, z.B. als Dunkelfeldbeleuchtung, ist allerdings eine andere Abstrahlrichtung für die einzelnen LEDs zu wählen. Die geplante Position ist also zu beachten.

Dombeleuchtung: Die Dombeleuchtung liefert sehr homogene und diffuse Beleuchtungen und schirmt gleichzeitig das Objekt deutlich von Fremdlicht ab. Das Grundprinzip ist die Streuung an der Kugeloberfläche (vgl. Abb. 8.1.4). Problematisch sind aber insbesondere bei etwas größeren Objekten die Kosten sowie der Platzbedarf.

Flat-Dom: Eine Alternative mit geringerem Platzbedarf und reduzierten Kosten, bei der die Abschirmwirkung weitgehend entfällt und das Winkelspektrum der Beleuchtungsrichtungen reduziert ist (Abb. 8.1.5).

Flächenstrahler: Werden ebenfalls zur homogenen und diffusen Ausleuchtung, vor allem großer Flächen eingesetzt. Je nach Homogenisierungsgrad können sie recht preiswert sein.

Linienlichter: Linienhafte Bauform (z.B. „Neonröhre" oder LED-Zeile) zur gerichteten Beleuchtung bei zweidimensionaler Abbildung oder zur Beleuchtung einer Zeile.

Punktstrahler: In verschiedenen Bauformen (z.B. Fasern, Lichtleiter, Halogenspot) zur gerichteten Beleuchtung oder in Verbindung mit einer Diffusorfläche als Flächenstrahler.

8.4.1 Thermische Quellen: Halogenlampen

Die meisten in der Natur vorkommenden Quellen (z.B. Sonne, Blitz, Feuer) sind thermische Quellen. Die Energiezufuhr, die nachfolgend dann zu einer Lichtemission führt, wird per Wärme zugeführt. Trotz der völlig unterschiedlichen thermischen Quellen wird ihr Verhalten doch im wesentlichen durch dieselben Grundgleichungen bestimmt. Diese Gleichungen sollen hier ohne Ableitung[6] kurz erläutert werden.

Streng genommen gelten sie für einen sogenannten *schwarzen Strahler*. Einen entsprechenden Strahler kann man sich als ein kleines Loch in einem beheizten Körper denken. Die Unterschiede zu der Abstrahlung anderer beheizter Körper sind aber für unsere Zwecke vernachlässigbar.

[6] z.B. Haken/Wolf: Atom– und Quantenphysik

Das *Plancksche Strahlungsgesetz* beschreibt die Strahldichte L, also den Strahlungsfluss (Einheit W) einer Einheitsfläche, die in einen bestimmten Raumwinkel emittiert, in Abhängigkeit von der Wellenlänge und der Temperatur des Strahlers:

$$L_e(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/(\lambda kT)} - 1} \cdot \frac{1}{\Omega_0} \tag{8.3}$$

$\Omega_0 = 1/sr$ ist dabei lediglich eine Einheitenkonstante, also eine Konstante, die dafür sorgt, dass in der Gleichung die Einheiten korrekt sind. $h = 6,626 \cdot 10^{-34}$ Js ist die Plancksche Konstante, $c = 3 \cdot 10^8$ m/s die Lichtgeschwindigkeit und $k = 1,38 \cdot 10^{-23}$ J/K die *Boltzmann Konstante*. Mit dieser Gleichung ist es somit möglich, die spektrale Abstrahlung zu berechnen, wenn die Temperatur der Quelle bekannt ist.

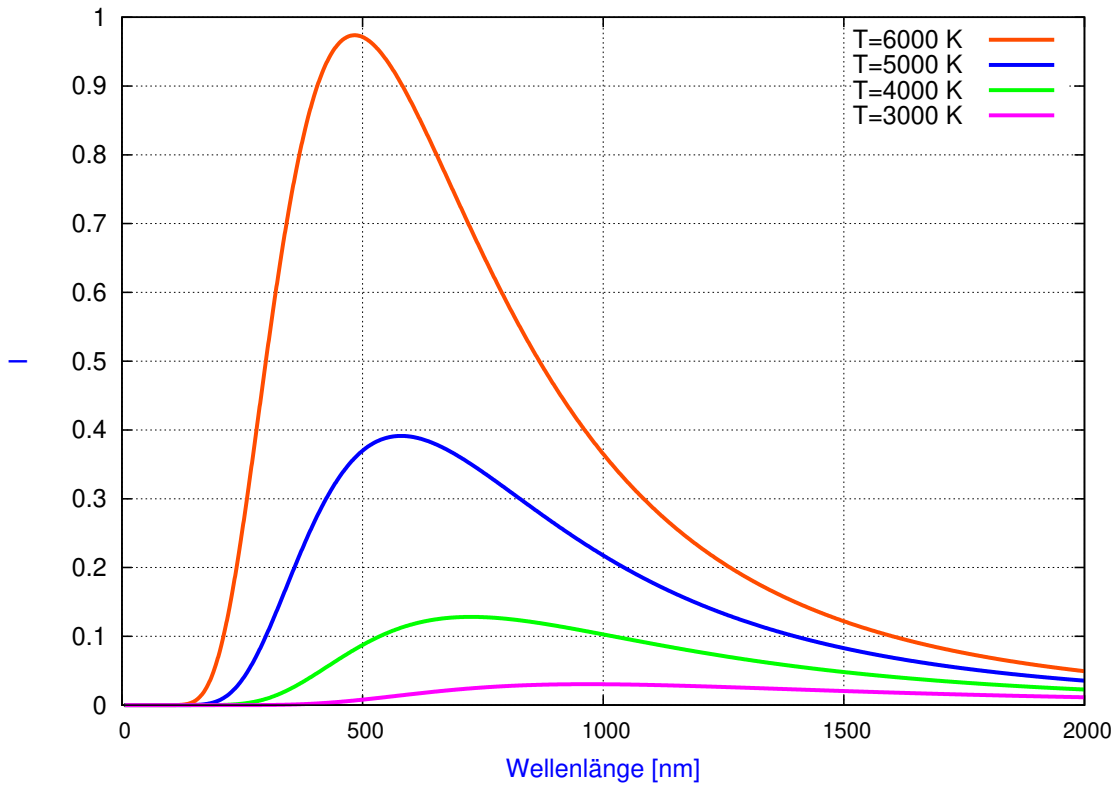


Abbildung 8.4.13: Spektrale Abstrahlcharakteristik einer thermischen Lichtquelle (Plancksches Strahlungsgesetz). Dabei wurden die Einzelkurven nicht zusätzlich normiert (heiße Quellen strahlen mehr Energie ab als kalte Quellen).

Oft interessiert man sich nicht für die Abstrahlung in einen bestimmen Raumwinkelbereich bei einer Frequenz, sondern für die gesamte abgestrahlte Leistung. Man kommt durch Integration dann zum *Stefan-Boltzmann-Gesetz*:

$$P(T) = \frac{\sigma}{\pi} T^4 \tag{8.4}$$

Dabei ist $\sigma = 5,67 \cdot 10^{-8} W/m^2/K^4$ eine Konstante. Man erkennt, dass umso mehr Leistung abgestrahlt wird, je höher die Temperatur ist (vgl. Abb. 8.4.13). Das theoretische Maximum an Effizienz, das mit einem thermischen Strahler realisiert werden kann liegt bei 95 lm/W. In der Praxis kommt man bei Wolfram zu einem Wert von maximal 54 lm/W. Im Vergleich hierzu lässt sich nicht-thermisches, „weißes" Licht theoretisch mit Wirkungsgraden von bis zu 200 lm/W generieren.

Verschiedene Stoffe derselben Temperatur haben zwar dieselbe spektrale Abstrahlcharakteristik (Plancksches Gesetz), strahlen aber sehr verschiedene Leistungen ab.[7] Dabei hängt der Emissionsgrad nicht nur vom Material, sondern auch von der Oberflächenbearbeitung ab[8].

Bei der von *Edison* und unabhängig davon von *Swan* entwickelten *Glühbirne* wird ein dünner Draht durch Stromfluss erwärmt. Für den Draht kommen dabei nur Materialien in Frage, die bei hoher Temperatur (ca. 2000 °C bis 3000 °C) nicht sofort verdampfen. *Wolfram* ist hierfür ein sehr geeignetes Material. Um die Lebensdauer in sinnvolle Größenordnungen (ca. 1.000 Stunden) zu bringen, ist die Glühbirne mit inertem Gas (z.B. Stickstoff) gefüllt.[9]

Die im Vergleich zu Sonnenlicht (ca. 6000 K) geringe Temperatur führt zu einem recht „warmen" Licht. Die *Farbtemperatur* gibt einfach die Temperatur eines Planckschen Strahlers an, der zu einem ähnlichen Spektrum führt. Licht, das von einem Strahler mit einem Spektrum gemäß Planck mit geringerer Temperatur als die Sonne herrührt, wird als „warm" bezeichnet. Farbtemperaturen werden aber dementsprechend nicht nur bei thermischen Quellen angegeben. Insbesondere die Normlichtquellen, z.B. *D65* für 6500K, sollen zwar die passende Farbtemperatur haben, dennoch können Peaks im Spektrum vorhanden sein (z.B. bei D65-Leuchtstoffröhren).

Gewöhnliche Glühbirnen werden praktisch kaum in der industriellen Bildverarbeitung eingesetzt, können aber oft als an sich unerwünschtes Fremdlicht auftreten. Stattdessen wird oft eine Abwandlung, die *Halogenlampe* eingesetzt. Sie funktioniert im Prinzip gleich wie eine herkömmliche Glühbirne. Allerdings werden

[7] Viele Strahler sind nicht perfekt „schwarz", sondern emittieren weniger Leistung. Solche Strahler werden auch als „graue Strahler" bezeichnet, wenn die Form des Emissionsspektrums weitgehend dem Planckschen Strahlungsgesetz folgt.

[8] Eine Tabelle einiger Materialien findet man unter http://www.thermographie-online.com/emissionsgrad.htm

[9] Eine Evakuierung der Glühbirne funktioniert natürlich ebenfalls.

dem Füllgas Halogene beigemischt. Dadurch wird eine Schwärzung des Glaskolbens verhindert. Sowohl die Lebensdauer (ca. 1,500 bis 2000 Stunden) als auch die Effizienz (ca. 15 bis 25 lm/W) ist gegenüber der Glühbirne (ca. 15 lm/W) besser. Die etwas höhere Temperatur der Glühwendel (ca. 3000 °C) ergibt ein etwas natürlicheres, d.h. näher beim Sonnenlicht liegendes Spektrum. Ein Betrieb mit niedrigen Spannungen ist möglich.

Eine Veränderung der Farbtemperatur (in der Regel zu kälteren Farbtemperaturen) wird oft durch spezielle Reflektorbeschichtungen oder nachgeschaltete Filter (Angaben in mired) realisiert. Rote Lichtanteile werden dabei teilweise vom Reflektor der Lampe transmittiert und daher nicht ins zu beleuchtende Gebiet reflektiert.

Vor- und Nachteile von Halogenbeleuchtungen:

+ sehr preisgünstig

+ kontinuierliches Spektrum ($\rightarrow$ guter Farbwiedergabeindex)

+ geringe Ansteuerspannung

o Farbtemperatur ca. 3000 K

- Lebensdauer gering ($<$ 2000 Stunden)

- starker Abfall des Lichtflusses mit Alter

- Anfällig gegen Vibrationen

- geringe Effizienz, viel Abwärme

- keine Pulsung möglich

8.4.2 Bogenlampen

Bei *Bogenlampen* werden zwei Elektroden durch ein Gas (z.B. Luft, Xenon, meist unter hohem Druck) voneinander getrennt. Bei hoher Spannung ergibt sich ein Lichtbogen zwischen den Elektroden. Die ionisierte Luft führt einerseits zur Lichtaussendung, andererseits zu einer hohen Leitfähigkeit zwischen den Elektroden.

Neben einem kontinuierlichen Betrieb können die Lampen auch geblitzt eingesetzt werden und erreichen dann extrem hohe Leuchtdichten von bis zu 350.000 candela/cm^2. Die Farbtemperatur ist hoch (5.000 bis 10.000 K) und es ergibt sich ein einigermaßen kontinuierliches Spektrum. Die Effizienz beträgt bis zu 60 lm/W und die Quellen sind auch für sehr hohe Leistungen (mehrere kW elektrische Leistung) erhältlich. Sehr oft eingesetzt werden *Kurzbogenlampen*, z.B. *XBO* von Osram.

Beim Einsatz in der Bildverarbeitung kann — vor allem bei hohen Leistungen — der hohe Anteil an Infrarotlicht zu Problemen (z.B. Erwärmung der Szene) führen. Infrarotfilter müssen eine hohe Leistung abführen und entsprechend gekühlt werden. Ein Ausfall entsprechender Kühlkomponenten kann unter Umständen zu drastischen Problemen innerhalb der gesamten Produktion führen.

Vor- und Nachteile von Bogenlampen:

+ sehr hohe Strahl bzw. Leuchtdichten erzielbar

- Lebensdauer 10.000 Stunden (meist aber deutlich kürzer)

- Explosionsgefahr durch hohen Druck ($\rightarrow$ potentielle Sicherheitsprobleme)

- teuer

- starke Wärmeproduktion

- UV Anteile im Licht ($\rightarrow$ potentielle Sicherheitsprobleme)

- hohe Spannungen notwendig ($\rightarrow$ potentielle Sicherheitsprobleme)

- Endhelligkeit wird erst nach Aufwärmphase erreicht.

8.4.3 Leuchtstofflampen

Die umgangssprachlich als *Neonröhren* bezeichneten Lichtquellen sind *Leuchtstofflampen* und damit *Entladungslampen*. Ein von Strom durchflossenes Gas (z.B. *Quecksilberdampf*) in einer teilevakuierten Röhre gibt dabei (in der Regel ultraviolette) Strahlung[10] ab, die durch eine fluoreszierende *Leuchtschicht* in sichtba-

[10] Bei Quecksilber vor allem λ=254 nm.

res Licht konvertiert wird[11], daher auch der Name *Fluoreszenzlampe*. Verschiedene Beschichtungen, die zu verschiedenen Spektren führen, sind gebräuchlich und verfügbar. Nur wenn die Lampe rötlich leuchtet, wird dabei Neon als Leuchtgas verwendet. Bei der Verwendung von Argon ergibt sich ein blaugrünes Leuchten, Krypton führt zu Lavendelfärbungen und Xenon leuchtet bläulich.[12]

Damit es zum Stromfluss in der Röhre kommt, werden zunächst Elektronen durch *Glühemission*[13] erzeugt. Die emittierten Elektronen werden zur positiven Gegenelektrode beschleunigt und führen dabei durch Stöße zu Ionen ($\rightarrow$ erhöhte Leitfähigkeit) und zur Atomanregung mit nachfolgender Lichtemission. Die Lebensdauer ist gegenüber der Glühbirne um den Faktor 10 (ca. 10.000 bis 50.000 Stunden) besser und die Effizienz ist mit ca. 65 bis 95 lm/W sehr hoch. Durch die Beschichtung lässt sich die Farbe der Lampe variieren. Die Farbtemperatur liegt dabei zwischen ca. 3.000 und 6.000 K.

Stromsparlampen sind letztlich gewöhnliche Leuchtstofflampen. Lediglich die Röhren selbst sind gebogen, um eine geringere Länge der Lampe zu erzielen. Das elektronische Vorschaltgerät für die Zündung ist in der Regel im Sockel der Lampe integriert.

Vor- und Nachteile von Leuchtstofflampen:

+ Lebensdauer größer 10.000 Stunden

+ sehr preisgünstig

+ sehr hohe Effizienz

+ homogene Beleuchtung großer Flächen

o verschiedene Farbtemperaturen von 3000 bis 6000 K

- sehr unflexible, große Bauform

[11] Wenn diese Leuchtschicht beschädigt ist, dann tritt ultraviolettes Licht aus (auch wenn vom Glas selbst einiges absorbiert wird) und es besteht dementsprechend eine Gesundheitsgefahr. Wenn die mit Quecksilber (giftig !) gefüllte Röhre zerplatzt, dann ist das natürlich ebenfalls ungünstig.

[12] Durch zusätzliche Fluoreszenzbeschichtungen wird allerdings das Spektrum der Entladungslampe nochmals (teilweise deutlich) geändert.

[13] Der hierbei ablaufende Vorgang ist im Detail auch recht komplex. Bei alten Röhren wird ein sogenannter „Starter" benötigt, neuere Röhren werden meist mit einem elektronischen Vorschaltgerät betrieben (bessere Effizienz, kein Flimmern, kein Brummen, längere Lebensdauer).

- stark diskontinuierliches Spektrum
- keine Pulsung möglich
- Zeitliches Flackern

8.4.4 Lichtemittierende Dioden (LED)

Das Grundprinzip von *Leuchtdioden* (Lichtemittierende Dioden, LEDs) basiert auf einem stromdurchflossenen pn–Übergang. Es handelt sich bei der LED also um eine Diode, die in Durchlassrichtung betrieben wird. Beim Stromfluss durch die Grenzschicht zwischen n– und p–dotiertem Halbleitermaterial findet eine Rekombination von Elektronen und Löchern statt. Dabei wird Energie frei, die in Form von Licht abgestrahlt wird. Man spricht von *Elektrolumineszenz*.

Abbildung 8.4.14: LEDs werden heute in einer Vielzahl von Anwendungen eingesetzt.

LEDs haben im Vergleich zu anderen Quellen eine Reihe herausragender Merkmale. Bei geringen Kosten wird eine lange Lebensdauer und eine hohe Robustheit bei guten bis sehr guten Effizienzen erzielt. Zusätzlich benötigen sie nur wenig Bauraum.

Die Energiekonversion (elektrisch $\rightarrow$ optisch) findet mit einer Effizienz von nahezu 100% statt. Die Gesamteffizienz der LED ist allerdings deutlich geringer und liegt bei ca. 30%, denn letztlich ist es schwierig, alle der in der dünnen Grenzschicht (einige Angström bis ca. 1 μm) generierten Photonen verlustfrei nach außen zu bringen.

Trotzdem ist die Effizienz im Vergleich mit anderen Lichtquellen sehr hoch. Inzwischen lassen sich Effizienzen von mehr als 100 lm/W erzielen[14].Ebenfalls möglich sind hohe Leistungen von z.B. 200 lm bei 10 W (OSTAR Lighting). Bei vielen LEDs ist ein gepulster Betrieb möglich. Dadurch kann die Leistung im Puls um bis zu den Faktor 10 gesteigert werden.

Die Frequenz der abgestrahlten Photonen wird durch die Bandstruktur der Halbleiter bestimmt und führt zu relativ schmalbandigem (ca. 10 bis 50 nm Halbwertsbreite) Licht. Um zu weißen LEDs zu kommen sind verschiedene Methoden im Einsatz. Am gebräuchlichsten ist die Nutzung einer fluoreszierenden *Konversionsschicht*, die das kurzwellige Licht der LED in ein breitbandiges *Fluoreszenzlicht* umwandelt. Sehr gebräuchlich (und kostengünstig) ist die Kombination einer blauen LED und einer gelben Konversionsschicht. Es handelt sich also um eine *dichromatische Farbmischung*, Blau + Gelb = Weiß. Ausgewogenere Spektren erhält man durch die Mischung von drei oder vier Spektren. Natürlich lassen sich auch mehrere LEDs mit unterschiedlicher Farbe direkt mischen. Auch dieser Ansatz wird oft eingesetzt. Generell ergibt sich dadurch ein etwas „künstlicheres" Spektrum (Peaks), aber die Energieeffizienz ist deutlich höher (beim Einsatz von fluoreszierenden Konversionsschichten wird ja ein Teil der Energie der blauen Photonen in Wärme umgewandelt und steht so für die Beleuchtung nicht mehr zur Verfügung). Einzelne LEDs sind vergleichsweise schlecht zur Beleuchtung von Objekten für eine sinnvolle Farbbeurteilung geeignet. Erst eine Kombination von mindestens vier LEDs liefert brauchbare Farbwiedergabewerte. Sehr viele ver-

[14]Für kleine Flächen sind die Effizienzen teilweise deutlich höher (bis 300 lm/W) und die Hoffnung besteht, diese Effizienzen auch auf größere Flächen übertragen zu können.

schiedene Emissionscharakteristika sind verfügbar und es wird versucht, ein einigermaßen kontinuierliches und konstantes Spektrum (sogenannte neutral-white LEDs) zu erzielen.

Angesteuert werden die LEDs über den fließenden Strom, denn die optische Leistung variiert in guter Näherung linear mit dem Strom, werden der Zusammenhang mit der Spannung extrem nichtlinear ist und eine Regelung bzw. Stabilisierung so schwierig macht.

Vor- und Nachteile von LEDs:

+ Lebensdauer bis 100.000 Stunden

+ preisgünstig

+ sehr hohe Effizienz

+ einfache Ansteuerung

+ pulsbar

+ Punktstrahler mit hohen Leuchtdichten, geringe Baugröße

0 spektrale Abstrahlcharakteristik sehr eingeschränkt $\rightarrow$ gut für Fremdlichtunterdrückung

- sehr hohe Strahlungsflüsse nicht möglich

- oft problematisches Spektrum $\rightarrow$ Farbwiedergabeindex schlecht.

Aktuell kann man sagen, dass die Vorteile der LED so immens sind, dass in der industriellen BV meistens LEDs eingesetzt werden wenn ein Einsatz überhaupt möglich ist. Hauptkontrakriterium für den Einsatz ist die begrenzte Lichtleistung. Aufgrund der sehr hohen Konzentration der Leistung muss für hohe Lichtleistungen die Augensicherheit gewährleistet werden (siehe IEC 62471).

8.4.5 Laser

Kohärentes Licht, wie es von Lasern ausgesendet wird kann Vor- aber auch Nachteile bei der Beleuchtung haben. Holografische und interferometrischer Messverfahren arbeiten prinzipbedingt mit kohärentem Licht, aber auch die gewöhnliche

Beleuchtung kann kohärent erfolgen. Insbesondere wenn auf extrem kleinen Flächen eine hohe Bestrahlungsstärke erzielt werden soll werden oft Laser eingesetzt. Ein weiterer Vorteil ist die vergleichsweise hohe Schärfentiefe bei der Generierung von einzelnen Lichtpunkten. Dies ist z.B. bei der *Lasertriangulation* von Vorteil.

Aber auch andere inhomogene Lichtverteilungen lassen sich oft vorteilhaft und mit hoher Effizienz mit Lasern realisieren. Die Elemente werden auch als *Diffraktiv-optische Elemente, DOE* bezeichnet. Zwischen dem eingebrachten Hologramm, einem strukturierten Plastik- oder Glaselement, und der zu beleuchtenden Ebene besteht (näherungsweise) eine zweidimensionale Fouriertransformationsbeziehung. Ein passend strukturiertes Hologramm führt so zu der gewünschten Bestrahlungsstärke auf dem Objekt. Typische Muster (Schachbretter, Kreuze, Kreise, Linien) sind kommerziell erhältlich und es lassen sich auch individualisierte Muster realisieren. In Zusammenhang mit LCDs können auch programmierbare Lichtmuster generiert werden.

Problematisch an einer Beleuchtung rauer Oberflächen, wie sie typischerweise in der industriellen Bildverarbeitung vorkommen, sind die sich ergebenden Speckles (vgl. Abschnitt 5.14.19) aber auch der zu beachtende Laserschutz, der schon bei vergleichsweise kleinen Leistungen (oberhalb vom 1 mW) beachtet werden muss.

Das vom Laser letztlich emittierte Licht kann aus einer einzelnen sogenannten „Mode" oder einer Kombination verschiedener Moden bestehen. Klassischerweise denkt man zunächst an die einzelne Mode. Die niedrigste Mode, die sogenannte TEM_{00} Mode entspricht gerade einer Gaußschen Verteilung und die Fokussierbarkeit dieser Mode ist optimal. Andere Moden erzielen dagegen nur vergrößerte Spots und das Verhältnis der realisierten Spotfläche im Vergleich zur Spotfläche für einen Gaußmode wird mit M^2 bezeichnet. Für Anwendungen, bei denen die beste Fokussierbarkeit wesentlich ist, sollte also das M^2 nahe 1 bleiben.

Wesentlich für die Spezifikation sind

- Betriebsmodus (Continous Wave oder gepulst)
- Leistung (pro Puls oder Continious Wave (CW))
- Wellenlänge

- Emittierte Mode (evtl. Angabe M^2)
- Kohärenzlänge
- Polarisationsverhalten
- zeitliche Stabilität (Rauschen)

Vor- und Nachteile:

+ extrem hohen Leuchtdichten, geringe Baugröße

+ extrem monochromatisch $\rightarrow$ sehr gute Femdlichtunterdrückung

+ (teilweise) pulsbar

+ (teilweise) hohe Effizienz

- eventuell Laserschutz notwendig

- Speckles und andere kohärente Artefakte

8.5 Advanced: Steuerung

Natürlich kann man eine Lampe einfach nur einschalten. Insbesondere mit den heute üblicherweise eingesetzten LED Beleuchtungsquellen werden oft aber spezielle Beleuchtungscontroller verwendet.

Zunächst ist zum Betrieb einer LED eine Steuerung über den Strom zu bevorzugen, denn die Kennlinie der LED ist stark nichtlinear. Während die emittierte optische Leistung in guter Näherung linear mit dem Strom steigt, führen kleine Spannungsänderungen schon zu massiven Änderung der Leistung. Letztlich muss also eine möglichst konstante Stromquelle zur Ansteuerung verwendet werden.

Die Konstanz ist wichtig, denn Änderungen in der Helligkeit führen oft zu Änderungen der Messung. Auch wenn die Software die Helligkeit normiert, ist dies nur bis zu einem gewissen Grad möglich und subtile Effekte können — insbesondere bei komplizierten Aufgabenstellungen — zu Messfehlern führen.

Bei aller Konstanz ist allerdings auch die Alterung der Lichtquelle zu beachten und so kann es sinnvoll sein, wenn sich der Ansteuerstrom eben automatisiert regeln lässt. Und hier sind wir dann bereits bei der einfachsten Anwendung für den Controller der Lichtquelle.

Teilweise wünscht man auch einen gepulsten Betrieb der Beleuchtung. Gründe hierfür können vielfältig sein. Oft soll die Bestrahlungsstärke während der Belichtungszeit erhöht werden und dies ist mit einem gepulsten LED Betrieb möglich. Oder aber das Hintergrundslicht soll explizit detektiert (für eine nachfolgende Subtraktion) werden. Die LED emittiert also nur dann Licht, wenn die Kamera auch wirklich belichtet. Entsprechend muss der Strom geschaltet werden und diese Schaltung muss dann wieder mit dem Bildsensor synchronisiert werden. Sobald mehrere Bildsensoren und/oder mehrere Beleuchtungsquellen (unter Umständen zeitversetzt) geschaltet werden sollen, desto komplexer wird die Steuerung. In der Regel wird ein spezieller LED Controller eingesetzt, aber es gibt auch Kameras, die direkt LEDs ansteuern können. Allerdings sind natürlich Leistung und Stromstärke im Pulsmodus begrenzt.

Ein Betrieb mit Pulsen kann im übrigen auch sinnvoll sein, wenn eine variable Beleuchtungsstärke erzielt werden soll. Viele Quellen können nur (sinnvoll) in einem vergleichsweise engen Ansteuerbereich (z.B. 1 bis 1.5 A) genutzt werden. Über ein geeignetes (und mit der Kamera synchronisiertes) Pulsmuster kann dann im Mittel (während einer Belichtungszeit) eine deutlich geringere Beleuchtungsstärke realisiert werden. Wenn man die Pulsbreiten zur Anstuerung nutzt spricht man von *Pulse-Width-Modulation* (PWM).

8.6 Advanced: Spezielle Techniken

8.6.1 Fluoreszenz

Die Anregung von Elektronen in Atomen kann auch durch Licht erfolgen. Eingestrahlte Photonen führen zu angeregten Energieniveaus. Nach einer gewissen Verweildauer der Elektronen in diesen Niveaus fallen die Elektronen zurück auf ihre Grundniveaus. Dabei werden Photonen abgestrahlt. Unter Umständen erfolgt der Übergang in den Grundzustand (strahlungslos) über Zwischenniveaus, so dass die Energie (und damit die Frequenz) der abgestrahlten Photonen geringer ist als die Energie der anregenden Photonen. Bei einer Einstrahlung mit blauem Licht wird dementsprechend z.B. rotes, grünes oder gelbes Licht abgestrahlt.

Die Unterscheidung zwischen *Fluoreszenz* und *Phosphoreszenz* erfolgt anhand der Zeitdauer zwischen Anregung und Remission. Bei kurzen Zeiten kleiner 10^{-4} s spricht man von Fluoreszenz. Bei der Phosphoreszenz kann die Dauer der Abstrahlung sehr lange (bis zu mehrere Stunden) sein und findet Anwendung in einer Vielzahl von Spielzeug und Notbeleuchtungen. Phosphoreszenz kann im übrigen nicht nur durch Bestrahlung mit Licht, sondern auch mit Beschuss durch Elektronen (*Kathodolumineszenz*, *Brownsche Röhre* z.B. beim Fernsehgerät) auftreten.

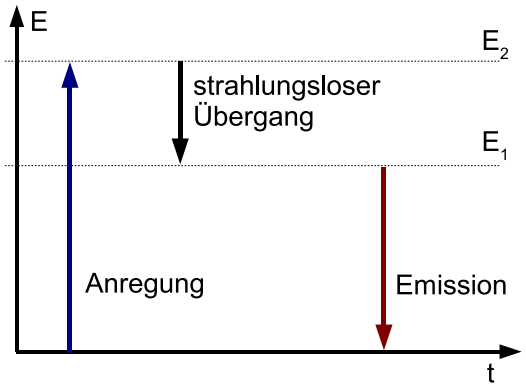


Abbildung 8.6.15: Fluoreszenz: Die Anregung von Elektronen mit energiereichen Photonen führt nach einer gewissen Zeit zu einer Remission von Photonen mit verringerter Energie und daher längerer Wellenlänge.

Leuchtfarben, wie man sie z.B. von *Rettungswesten* kennt, erhalten ihre Leuchtkraft durch Fluoreszenz. Kurzwelliges, blaues und ultraviolettes Licht wird z.B. in orangefarbenes Licht (gelb und rot) konvertiert. Damit wird *mehr* Licht im gelb–

roten Spektralbereich von der Rettungsweste emittiert als überhaupt im Anregungslicht vorhanden war. Die Weste leuchtet damit deutlich heller orange als eine mit gewöhnlicher orangener Farbe lackierte Fläche.

Fluoreszierende Stoffe finden sich auch in den sogenannten *Aufhellern* von *Waschmitteln*. Sie sorgen für strahlendes Weiß, indem ultraviolettes Licht in weißes Licht konvertiert wird. Viele Stoffe in der Natur (vor allem *Bakterien* und *Mineralien*) fluoreszieren ebenfalls und zum Einsatz kommen fluoreszierende Pigmente auch bei *Sicherheitsanwendungen* (z.B. auf *Geldscheinen*).

Für die industrielle Bildverarbeitung wesentlich ist, dass *organische Stoffe* in der Regel fluoreszieren. Es wird so möglich, organische Rückstände (z.B. Öle, Fingerabdrücke) auf technischen Oberflächen per Fluoreszenz sichtbar zu machen. Die Beleuchtung erfolgt mit intensivem kurzwelligem Licht (z.B. blau oder grün) und bei der Detektion wird die Anregungswellenlänge durch einen Spektralfilter geblockt. Das den Bildsensor erreichende Licht ist somit nur das fluoreszierende Licht, das von Positionen mit organischen Rückständen herrührt.

8.6.2 Polarisation

Wir wollen in diesem Buch nicht weiter auf die elektromagnetische Beschreibung des Lichts eingehen, denn in aller Regel können wir sich daraus ergebende Effekte vernachlässigen. Die wichtigste Ausnahme hiervon ist für uns sicherlich die Beugungsbegrenzung der Auflösung (vgl. Abschnitt 5).

Es gibt aber eine weitere Eigenschaft des Lichts, die sich im naiven Teilchenmodell der Strahlenoptik nicht erschließt: die Polarisation. Im Wellenmodell wird Licht als die kombinierte Schwingung eines elektrischen und eines magnetischen Felds betrachtet. Die beiden Felder stehen dabei senkrecht aufeinander und auch beide senkrecht zur Ausbreitungsrichtung[15]. In welcher Ebene das elektrische Feld schwingt gibt die *Polarisation* des Lichts an.

Die Schwingunsebenen der Feldkomponenten einer elektromagnetischen Welle werden von uns nicht (bzw. nahezu nicht) direkt gesehen. Den Photodetektoren in unseren Augen ist es egal, unter welchem Winkel das elektrische Feld der Welle

[15] Zumindest gilt dies in homogenen Medien bzw. schwach inhomogenen Medien, siehe z.B. [49]

auftrifft und daher sind wir eben nicht in der Lage, Polarisation direkt zu sehen. Andere Lebewesen, z.B. Bienen, unterliegen dieser Beschränkung nicht und nutzen so die Polarisation — analog zur Farbe — als zusätzlichen Informationskanal.

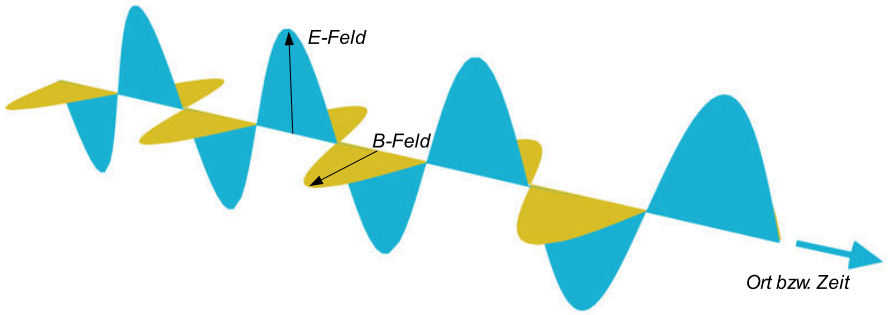


Abbildung 8.6.16: Licht als elektromagnetische Welle. Der Vektor des elektrischen Felds steht senkrecht auf dem Vektor des magnetischen Felds. Beide stehen senkrecht zur Lichtausbreitungsrichtung. Das Diagramm kann sowohl als eine Momentaufnahme (Zeit t=konstant) als auch eine Aufnahme des zeitlichen Verlaufs (Ort x=konstant) betrachtet werden. Die horizontal verlaufende Achse ist dann jeweils die Ortsachse (x) oder die Zeitachse (t).

Solange das elektrische Feld immer in derselben Richtung schwingt, spricht man von *linearer Polarisation*.

Die meisten natürlichen und künstlichen Lichtquellen (z.B. Sonne, Glühbirne) senden *unpolarisiertes Licht* aus.[16] „Unpolarisiert" bedeutet dabei, dass die Photonen des Lichtfelds eine „zufällige" Polarisation aufweisen.

Entscheidend in der industriellen Anwendung ist nun, dass der Polarisationsgrad bei einer Reflexion mehr oder weniger erhalten bleibt, bei der Streuung aber (materialbhängig) teilweise verloren geht. Dadurch lassen sich *Störreflexe* durch geeigneten Einsatz von Polarisationsfilter reduzieren oder Materialen, insbesondere *Metalle* (Polarisation bleibt erhalten) von Nichtmetallen, voneinander unterscheiden. Besonders die Differenz von zwei mit unterschiedlicher (bildseitiger) Polarisation aufgenommenen Bildern erlaubt teilweise die Separierung von Objekten.

Es gibt verschiedene Möglichkeiten, aus unpolarisiertem Licht teilweise oder komplett polarisiertes Licht herzustellen bzw. das Licht anhand seiner Polarisation zu filtern. Bei jeder Reflexion an einer dielektrischen Grenzschicht (also z.B.

[16] Laser als bedeutende technische Lichtquellen liefern je nach Bauart polarisiertes oder unpolarisiertes Licht.

an einer Glasscheibe) ergibt sich eine teilweise Polarisierung. Der Reflexionsfaktor hängt stark vom Einfallswinkel, den Brechzahlen der Medien und eben der Polarisation ab, so dass im reflektierten Licht in der Tat eine der beiden Polarisationsrichtungen stärker vertreten ist (vgl. Fresnelgleichungen, z.b. in [22]). Diese Grenzschichten lassen sich als dünne Schichten auslegen und führen so z.B. zu polarisierenden Strahlteilerwürfeln.

Eine andere, qualitativ sehr hochwertige aber auch teure Variante sind Prismen aus doppelbrechenden Materialien[17] bei denen die Polarisationsrichtungen unterschiedlich abgelenkt werden.

Gebräuchlich sind in der Bildverarbeitung aber vor allem absorbierende Polarisationsfilter bei denen jeweils eine Polarisationsrichtung komplett absorbiert, also in Wärme umgesetzt, wird. Meist werden diese aus Kunststoffen hergestellt (typische preiswerte Fotopolarisatoren) aber auch Elemente mit besserer Qualität aus Glas sind verfügbar. Natürlich kann die Umsetzung in Wärme bei der beleuchtungsseitigen Verwendung von leistungsstarken Lichtquellen zu Problemen bis zur Zerstörung der Elemente führen.

Ganz generell ist bei der Auswahl der Polarisatoren auf den zulässigen Winkelbereich der Polarisatoren, das Auslöschungsverhältnis von zu filternder und zu transmittierender Polarisation, die Gesamttransmission und die Wellenlängenabhängigkeit zu achten. Weiterhin muss die optische Qualität des Elements ausreichend sein, um keine nennenswerten Bildfehler in das System einzubringen. Daher wird der Einsatz in der Regel vor dem Frontelement des Objektivs (paralleler Strahlengang) erfolgen.

[17] unterschiedliche Polarisationsrichtungen „sehen" verschiedene Brechzahlen und dementsprechend ergibt sich eine Richtungsaufspaltung für unterschiedliche Polarisationsrichtungen

Kapitel 9

Bildsensoren

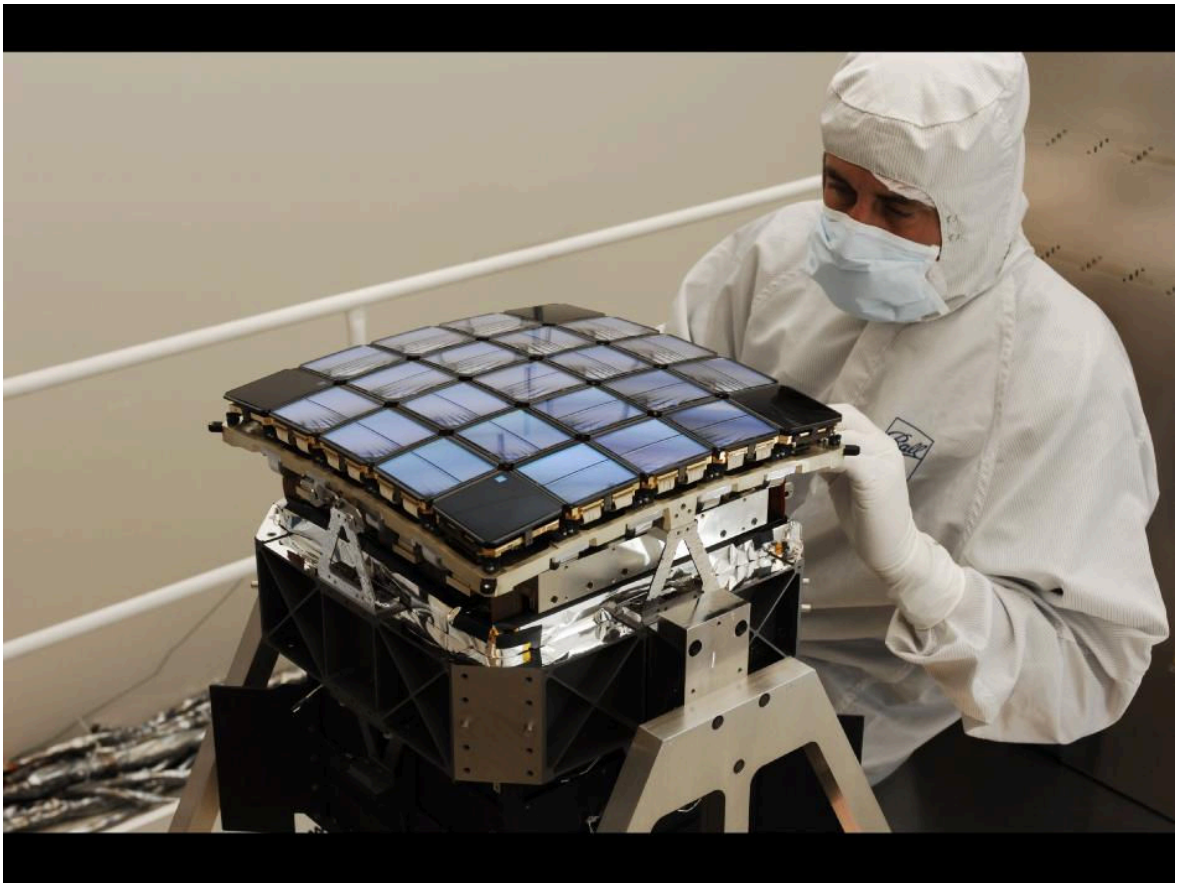


Abbildung 9.0.1: Focal Plane Array des Kepler Teleskop. Bild mit freundlicher Genehmigung der NASA.

In der industriellen Bildverarbeitung werden heute im wesentlichen CMOS Sensoren eingesetzt. Wir werden vorrangig auf diesen Sensortyp eingehen, aber im „Advanced-Bereich" auch etwas zu CCDs und anderen Sensoren sagen.

Der Sensor allein ist allerdings noch nicht die Kamera. Rein begrifflich wird unter

„Kamera" entweder der gesamte Bildaufnehmer mit oder eben ohne Optik (Objektiv) verstanden. Das zentrale Element des Bildaufnehmers selbst ist natürlich der Bildsensor, der in hoher Stückzahl von wenigen, meist asiatischen, Firmen hergestellt wird. Allerdings muss der Sensor auch elektronisch angesteuert werden.

Abb. 9.0.2 zeigt schematisch, wie entsprechende moderne Bildaufnehmer in der Regel aufgebaut sind. Der Bildsensor wird meist von einem FPGA angesteuert und ausgelesen. Unter Umständen wird ein zusätzlicher Mikrocontroller verwendet, um die Gesamtsteuerung zu übernehmen und natürlich braucht es ein passendes Interface zum Computer oder der in den Computer eingesteckten Bilderfassungskarte (Framegrabber). An verschiedenen Stellen können die eigentlichen Bilddaten also verarbeitet werden. Für Low-Cost und Low-Power Lösungen bei großer Stückzahl ist es auch möglich, mehrere der Funktionalitäten in einen applikationsspezifischen Schaltkreis (ASIC) zu integrieren und so sehr hohe Auswertegeschwindigkeiten zu erzielen (z.B. bei Sensoren für Lichtschnittanwendungen (Triangulation)).

Basis ist zunächst natürlich die Lichtdetektion. Wir wollen uns an dieser Stelle nicht in den Tiefen der Halbleiterphysik verlieren und nur die wesentlichen Punkte, die letztlich zu einem Signal im Computer führen, betrachten.

Der zentrale Prozess ist zunächst der innere *Photoeffekt*: Ein einfallendes Photon kann seine Energie zur Generierung eines Elektron-Lochpaares verwenden. Das so entstandene Elektron wird durch geeignete von außen eingebrachte elektrische Felder lokalisiert gespeichert (Pixel). Die einfallenden Photonen führen somit zu gespeicherten Ladungen, die dann in nachfolgenden Prozessen abtransportiert werden. Der schematische Aufbau eines Bildafunehmerpixels ist in Abb. 9.0.3 dargestellt.

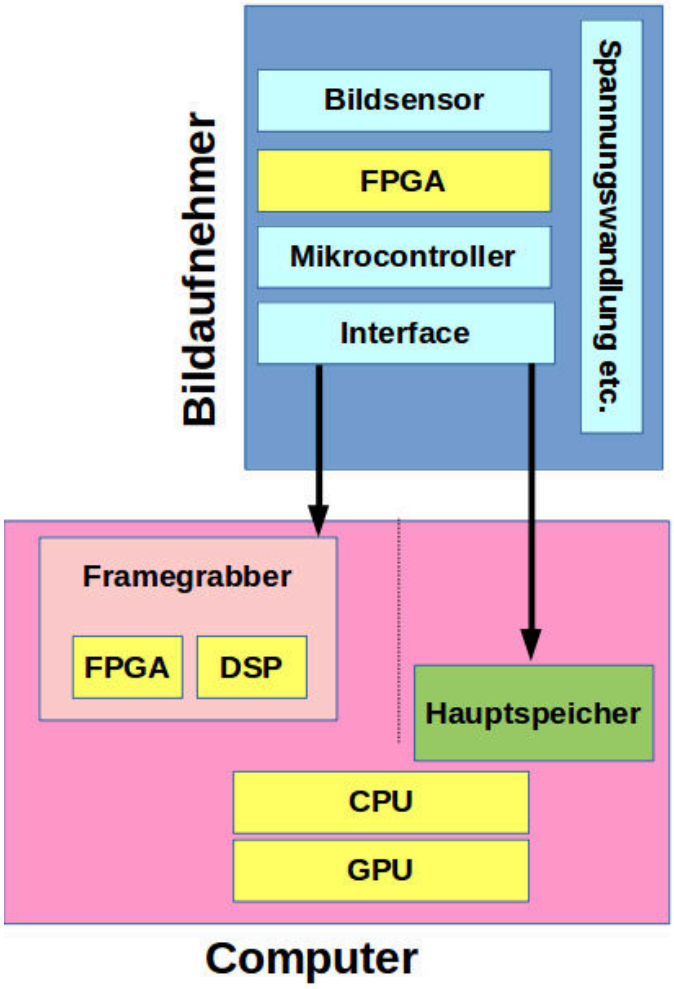


Abbildung 9.0.2: Typischer Aufbau eines Bildaufnehmers und Anschluss an den Computer. Die Anbindung des Bildaufnehmers erfolgt direkt über eine im PC integrierte Schnittstelle (USB, Ethernet) oder aber über einen Framegrabber (z.B. CameraLink). In den gelb dargestellten Blöcken können Berechnungen stattfinden. Verschiedene Speicherbereiche (hier nicht eingezeichnet) werden zur Speicherung der Bilddaten verwendet (in Kamera, in Framegrabber, auf Grafikkarte).

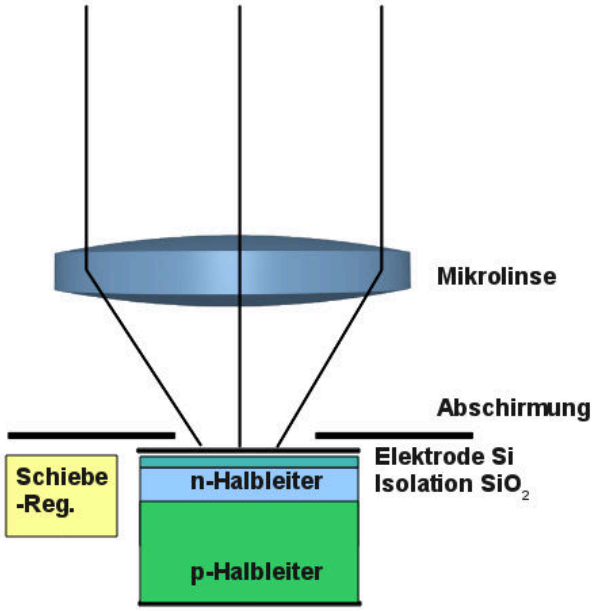


Abbildung 9.0.3: Schematischer Aufbau eines CCD-Pixels. Als Elektroden und Substrat wird dotiertes Silizium verwendet. Eine isolierende Schicht (Siliziumoxid) liegt zwischen den Ansteuerelektroden und dem eigentlichen Pixel. Damit lässt sich das elektrische Feld im Pixel räumlich steuern.

Der Pixel kann als Kondensator aufgefasst werden. Je nach einfallender Lichtmenge ist der Kondensator unterschiedlich stark geladen und beim Auslesen fließt die Ladung ab. Die zentrale Größe ist dabei zunächst die Kapazität des Kondensators, die sogenannte *Sättigungskapazität* (engl. *Quantum Well Capacity*, QWC), die schlicht angibt, wie viele Ladungen pro Pixel gespeichert werden können. Je kleiner die Pixel sind, desto geringer ist diese Sättigungskapazität. Typische Werte liegen zwischen 2.000 und 100.000 Elektronen.[1] Je größer die Sättigungskapazität, desto größer die mögliche Dynamik. Auch das Signal-Rausch-Verhältnis hängt wesentlich von ihr ab (vgl. Abschnitt 9.5).

Nicht jedes einfallende Photon führt allerdings zu einer detektierbaren Ladung. Die Wahrscheinlichkeit für den Prozess wird *Quanteneffizienz* genannt. Sie hängt — genau wie die des menschlichen Auges — stark von der Wellenlänge des Lichts, der Geometrie des Pixels und natürlich den Materialien, die verwendet werden, ab.[2] In Abb. **??** ist eine typische Abhängigkeit gezeigt. Beachten Sie dabei, dass es durchaus auch bei kleinen Variationen der Wellenlänge deutliche Variationen geben kann. Die in Datenblättern angegeben Kurven sind teilweise schlicht nicht mit hoher spektraler Auflösung gewonnen.

Den Standard bilden auf Silizium basierende Sensoren. Für diese ist prinzipbedingt nur eine Empfindlichkeit von ca. $\lambda = 300$ nm bis 1100 nm gegeben (mit nennenswerter Empfindlichkeit eher bis 950 nm). Längerwellige Photonen haben nicht genug Energie, um Elektronen im Silizium in das sogenannte „Leitungsband" zu heben und damit verfügbar zu machen. Sehr kurzwelliges Licht wird bereits in oberen Schichten absorbiert.

Je nachdem, wie die Ladung abtransportiert und weiterverarbeitet wird unterscheidet man bei kommerziellen Sensoren heute vor allem zwischen CCD (Charged Coupled Device) und CMOS (Complementary Metal Oxide Semiconductor) Sensoren. Bei CCD Sensoren wird die im Pixel gespeicherte Ladung zu einem (oder wenigen) zentralen Ausgangselementen „geschoben" und dann dort in eine Spannung gewandelt. Bei CMOS Sensoren wird dagegen direkt im bzw. am Pixel die Ladung in eine Spannung gewandelt. Auf die sich dadurch ergebenden we-

[1] Allerdings wird die QWC von der verschiedenen Herstellern teilweise unterschiedlich definiert, z.B. in Bezug auf den linearen Bereich des Pixelverhaltens oder aber den gesamten Bereich oder auch die Lichtmenge, bevor Blooming auftritt. Weiterhin werden teilweise typische oder auch maximal mögliche Werte angegeben.

[2] Vor allem im NIR hängt sie auch stark von der Temperatur ab.

sentlichen Unterschiede in der Anwendung werden wir im folgenden Abschnitt zu sprechen kommen.

9.1 CMOS Sensoren

Sensoren basierend auf *CMOS* Transistoren haben gegenüber CCDs eine Reihe von Vorteilen. Entscheidend war und ist aber zunächst vor allem die Möglichkeit der Nutzung konventioneller Herstellungstechniken für integrierte Schaltkreise (IC). Die Herstellungsprozesse können zwar nicht komplett und unverändert übernommen werden, eine erhebliche Einsparung bei der Herstellung ist aber dennoch möglich. Fortschritte bei der lithografischen Herstellung von ICs können vergleichsweise einfach genutzt werden, so dass inzwischen sehr kleine Sensoren möglich sind. Vorteilhaft ist auch, dass die CMOS-basierten Sensoren mit geringen Spannungen auskommen während CCDs meist mehrere, vergleichsweise hohe Spannungen benötigen. Auch die Energieeffizienz ist verbessert und es lässt sich auf einfache Weise zusätzliche Funktionalität, vor allem hinsichtlich der Verarbeitung der Signale, integrieren.

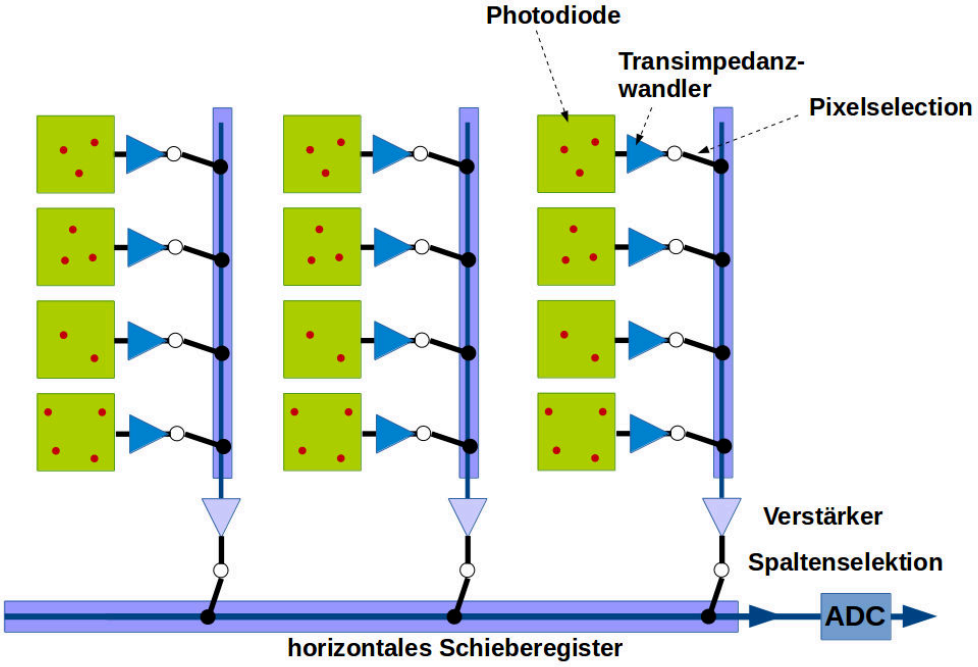


Abbildung 9.1.4: Im Gegensatz zur CCD wird bei CMOS Sensoren bereits im Pixel eine Konversion der Ladung in eine Spannung vorgenommen.

Ein wesentlicher Schritt zur Verbesserung der Qualität von CMOS Sensoren war dabei die Integration von einem oder mehreren Transistoren pro Pixel (*Active Pixel Sensor*, APS). Hierdurch konnte das Rauschverhalten, die traditionelle Schwachstelle von CMOS Sensoren, deutlich verbessert werden.

Auch die üblicherweise durchgeführte Analog-Digital Wandlung kann direkt auf dem CMOS Chip durchgeführt werden. Durch die Integration solcher Funktionalitäten auf dem Chip kann das Gesamtsystem „Kamera" deutlich verkleinert werden und auch winzige Kameras von nur wenigen Kubik-Millimetern Bauraum lassen sich realisieren (z.B. für endoskopische Anwendungen). Die Verkleinerung typischer Sensoren hat ihren Grund aber vor allem in einer weiteren Reduktion der Kosten. Letztlich sind Halbleiterelemente umso teurer, je größer sie sind, denn man kann pro gefertigtem Wafer eben deutlich weniger Elemente nutzen.

Die integrierbare Funktionalität kann auch für andere Dinge genutzt werden. Z.B. kann pro Zeile oder Spalte ein separater AD-Konverter eingesetzt werden, um einen sehr schnellen Bildsensor zu generieren. Auch die Kennlinie der Pixel kann variiert werden. So können Sensoren mit hoher oder sogar programmierbare Dynamik (vgl. Abschnitt 9.6) realisiert werden. Vielfach von Bedeutung ist auch die Möglichkeit, einzelne Gebiete (*Regions of Interest*) mit erhöhter Geschwindigkeit auszulesen (vgl. Abschnitt 9.8).

Generell kann gesagt werden, dass CMOS-basierte Sensoren historisch gesehen eine schlechtere Bildqualität als CMOS Sensoren liefern, dafür aber preiswerter sind und Vorteile hinsichtlich des Energieverbrauchs und der Funktionalität (z.B. Regions-of-interest, Spezialfunktionen) haben. Der Nachteil bei der Bildqualität tritt allerdings zunehmend in den Hintergrund und ist heute im wesentlich für hohe Empfindlichkeiten bzw. Anwendungen mit geringer Beleuchtungsstärke und mittleren bis hohen Anforderungen an das Signal-Rausch-Verhältnis von Bedeutung.

Insbesondere aufwändigere CMOS Sensoren, wie z.B. die sogenannten „scientific CMOS" Sensoren (*sCMOS*) erzielen aufgrund einer verbesserten Ausleseelektronik mit einem AD-Wandler pro Spalte ein sehr gutes Rauschverhalten und können hinsichtlich des Ausleserauschens CCD Sensoren übertreffen.

Früher galt, dass qualitativ hochwertige Sensoren immer CCD Sensoren sein

müssen. Diese Zeiten sind vorbei. Vielmehr muss man die technischen Daten der Sensoren betrachten. Für die industrielle Bildverarbeitung sollte man insbesondere auf das Rauschverhalten und die Geschwindigkeit beim Vergleich achten.

9.2 Advanced: CCD-Sensoren

Grundsätzlich werden bei *CCD Sensoren* die generierten Ladungsträger (Also entweder Elektronen oder Löcher) entlang einer Spalte des Bilds verschoben bis sie in ein horizontales Schieberegister gelangen. Dann wird dieses Schieberegister durch eine horizontale Verschiebung ausgelesen (Abb. 9.2.5). Die ausgelesenen Ladungen werden schließlich über einen *Transimpedanzwandler* in Spannungen konvertiert.[3] Bei analogen Sensoren können diese Spannungen direkt als Videosignal verwendet werden. Digitale Sensoren digitalisieren das Analogsignal. Teilweise erfolgt noch ein digitaler Control Loop mittels Digital-Analog Wandlung zur Grundstabilisierung des Signals.

Der eigentliche Ladungstransport ist nicht trivial und verschiedene Techniken werden genutzt, um mit extrem hoher Effizienz die Ladungen weiter zu schieben. Letztlich müssen die Ladungen über mehrere tausend Zellen geschoben werden. Wenn dieser Transport pro Zelle nicht nahezu perfekt erfolgt, also jedes mal einige wenige Elektronen nicht transportiert werden, dann ergibt sich eine deutliche Bildverschlechterung im Sinne einer Kontrastreduktion. Über die Ansteuerelektroden oberhalb des Pixels, z.B. vier pro Pixel, lässt sich durch eine fortlaufende periodische Änderung der elektrischen Felder dieser Transport bewerkstelligen.

Aus Sicht des Anwenders ist wesentlich, ob es sich um einen *Interlaced* oder *Non-interlaced* (= *Progressive Scan*, *Vollbildverfahren*) Sensor handelt. Interlaced Sensoren liefern zunächst nacheinander alle geraden Zeilen des Bilds bevor sie danach alle ungeraden Zeilen ausgeben. Die klassische Fernsehtechnik sowie die üblichen analogen Videonormen arbeiten nach dem Interlace-Verfahren. Es bietet bei einer visuellen Darstellung den Vorteil, dass auch bei geringen Bildwiedergabefrequen-

[3] Davor wird der sich ergebende Gleichspannungsanteil durch einen Hochpassfilter (Kondensator) entfernt, es erfolgt eine Vorverstärkung und dann eine Restauration des Gleichstromanteils, um eine etwaige Hintergrundshelligkeit wieder herzustellen. Rauschen wird unter Umständen durch sogenanntes „correlated double sampling" reduziert. Einige CCDs nutzen die unbelichteten „Black pixels" an den Zeilenrändern, um die Grundspannung einzustellen.

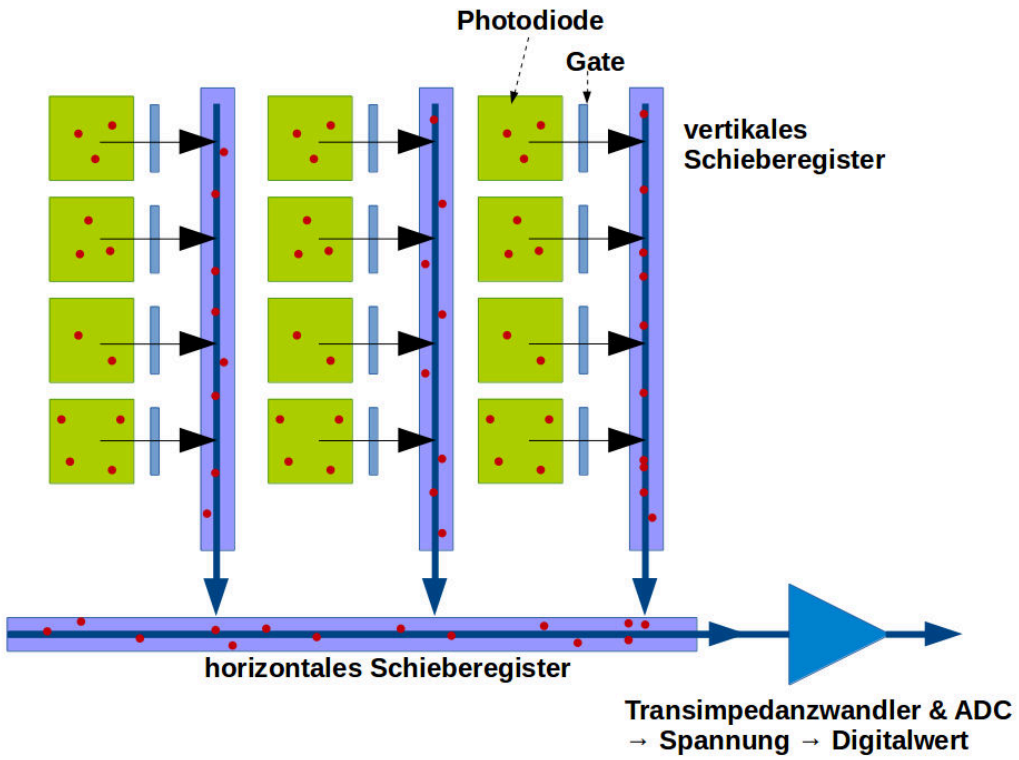
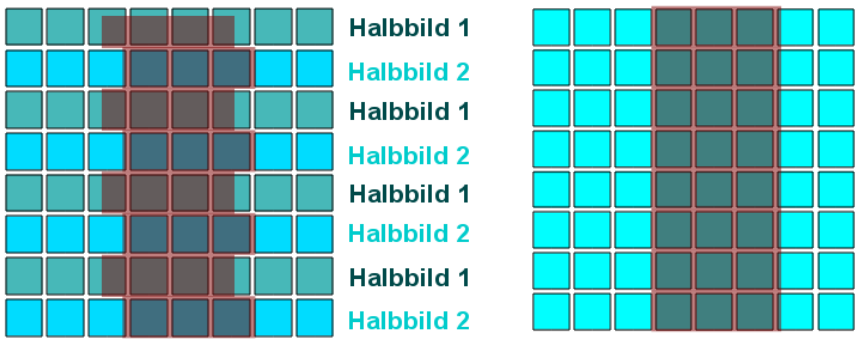


Abbildung 9.2.5: Fließband- oder Eimerketten Prinzip des Auslesens der Ladungen in einem CCD Sensor. Die Ladungen der Pixel werden jeweils horizontal und dann über ein Schieberegister vertikal geschoben bevor sie ausgelesen werden. Somit reicht eine Ausleseelektronik, um die Ladungen in Spannungen zu wandeln und sie dann zu digitalisieren.

zen (z.B. 50 Hz) ein Update des Bildes mit dieser Frequenz zumindest für die Hälfte des Bildes erfolgt.



(a) Interlaced (b) Non-Interlaced

Abbildung 9.2.6: Interlace und Non-interlaced (progressive scan) Sensoren: Beim Interlace-Verfahren wird das Bild — auch bei der Belichtung — in zwei Halbbilder aufgeteilt. Dementsprechend ergeben sich bei bewegten Objekten (hier: ein horizontal bewegter Balken der Breite 3 Pixel) Artefakte (Zeilenversatz).

Für ein bewegtes Objekt ergibt sich bei der maschinellen Bildverarbeitung das Problem, dass das Gesamtbild sich aus zwei zeitlich versetzten Einzelbildern und damit aus zwei Bildern mit dem Objekt bei verschiedenen Positionen ergibt (siehe Abb. 9.2.6). Für bewegte Objekte ist der Interlace-Modus in der industriellen Bildverarbeitung daher kaum zu gebrauchen und man wird stattdessen immer einen Progressive-Scan Sensor verwenden.

Die eigentliche Bewegung der Ladung zum horizontalen Schieberegister kann ebenfalls auf verschiedene Weise erfolgen. Konzeptionell am einfachsten sind *Full-Frame* Sensoren. Ein belichtetes Bild wird Zeile für Zeile in das Schieberegister geladen und dann von dort zum Ausgang geschoben. Während des gesamten Schiebevorgangs muss natürlich dafür gesorgt werden, dass die lichtempfindlichen Zellen (Pixel) nicht weiter belichtet werden. Daher wird entweder ein *Shutter* oder aber eine gepulste Beleuchtung verwendet. Entsprechende Sensoren werden heute selten in der industriellen Bildverarbeitung verwendet.

Bei *Frame-Transfer* Sensoren wird zusätzlich zum eigentlichen Sensorarray noch ein gleich großes Speicherarray verwendet. Nach der Belichtung wird innerhalb sehr kurzer Zeit (ca. 0,5 ms) die aufgenommene Ladung in das Speicherarray transferiert. Nur während dieser kurzen Zeit kann es zum Verschmieren aufgrund von Nachbelichtung kommen. Wenn daher die eigentliche Belichtungszeit deutlich länger als diese Transferzeit ist, ist kein Shutter bzw. kein Pulsen der Beleuchtung notwendig. Frame-Transfer Sensoren sind prinzipbedingt immer non-interlaced Sensoren und liefern eine ideale Lichtnutzung. Natürlich wird die doppelte Chip-Fläche benötigt (teuer) und ein mechanischer Shutter (oder gepulste Beleuchtung) ist — zumindest bei kurzen Belichtungszeiten — notwendig, denn während des Schiebens der Ladungen in den Speicherbereich bleiben die Pixel selbstverständlich lichtempfindlich.

Am gebräuchlichsten im Massenmarkt sind aber *Interline-Transfer* Sensoren. Hierbei wird zwischen jeder Spalte eine zusätzliche Speicherspalte eingezogen. Dadurch muss die Ladung jedes Pixels nur um eine sehr geringe Strecke (eben die benachbarte Speicherzelle) verschoben werden. Dies kann extrem schnell (1 μs) erfolgen und ein Shutter ist nicht notwendig. Damit wird die Verschmierung aufgrund einer möglichen Nachbelichtung während dieser Zeit ausgeschlossen. Auf einen mechanischen Shutter kann damit verzichtet werden. Die eigentliche lichtemp-

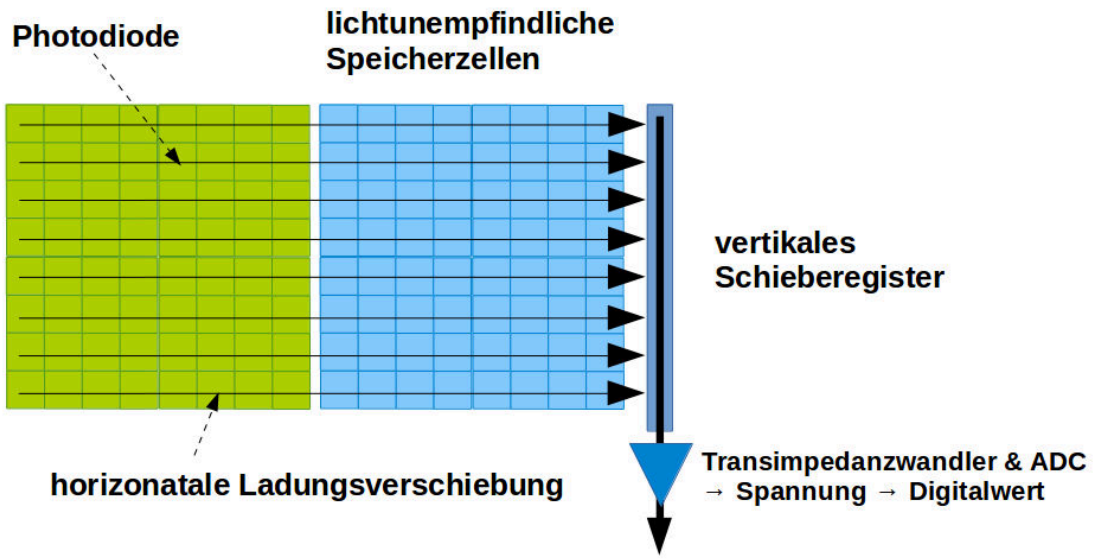


Abbildung 9.2.7: Frame transfer Sensoren verwenden zusätzlich zum eigentlichen lichtempfindlichen Bereich einen Speicherbereich. Das gesamte Bild wird also zunächst in den Speicherbereich verschoben bevor dieser ausgelesen wird.

findliche Fläche wird natürlich deutlich reduziert, aber durch *Mikrolinsen* kann die effektiv wirksame Fläche wieder auf nahezu 100% gebracht werden.

Allgemein kann gesagt werden, dass heute in der industriellen Bildverarbeitung als CCDs vor allem Interline-Transfer Sensoren oder Frame-Transfer Sensoren (teurer aber lichtempfindlicher) eingesetzt werden. Meist als als non-interlaced Sensoren.

9.3 Bildsensorgeometrie

Ein wesentliches Merkmal von Flächensensoren ist natürlich zunächst ihre Grundgeometrie, d.h. wie viele Pixel sind vorhanden und wie ist der Abstand der Pixel zueinander (*Pixelpitch*) bzw. wie groß ist die lichtempfindliche Fläche. Diese wird in einem etwas ungewöhnlichen bzw. verwirrenden Maß angegeben, nämlich in Bezug auf eine Umrechnung auf heute praktisch nicht mehr verwendete Vidicon Röhren. Die Umrechnung ist letztlich ohne Belang. Es ergeben sich aber für die gebräuchlichsten Größen folgende geometrischen Größen[4]:

[4] Allerdings können die Größen leicht variieren, d.h. ein 1/3 inch Sensor kann bei unterschiedlichen Herstellern leicht unterschiedliche Geometrien haben.

Format	Diagonale [mm]	Geometrie [mm x mm]
1/8 inch	2	1.6 x 1.2
1/6 inch	3	2.4 x 1.8
1/4 inch	4	3.2 x 2.4
1/3 inch	6	4.8 x 3.6
1/2 inch	8	6.4 x 4.8
2/3 inch	11	8.8 x 6.6
4/3 inch	21	17.3 x 13

In der industriellen Bildverarbeitung werden aus Qualitätsgründen (Rauschen) in der Regel Sensoren ab 1/3 inch verwendet. Kleinere Sensoren sind in preisgünstigen Consumerprodukten (z.B. Webcams) zu finden. Durch die stetig steigenden Qualitäten der Sensoren nimmt aber auch die Verbreitung von Sensoren mit kleinen Diagonalen stetig zu. Für die Pixel gilt entsprechend, dass sich ein Trend zu immer kleineren Pixelgeometrien ergibt. Inzwischen sind Sensoren mit Pixeln im Bereich 1 μm bis 1.5 μm verfügbar

Vergleichsweise preiswerte große Bildsensoren können mit dem 4/3 inch Standard (sogenannter *Micro-Four-Third Standard*), realisiert werden da hierfür sowohl Optiken als auch Bildsensoren aus dem Consumer-Markt übernommen werden können. Dabei ist auch eine motorisierte Änderung wesentlicher Objektivparameter (Fokus, Zoom, Blende) sehr günstig möglich. Die in der Consumer-Fotografie übliche Steuerbarkeit der Objektive ist (bislang) in der industriellen Bildverarbeitung wenig gebräuchlich, letztlich aber dann sehr vorteilhaft, wenn in einer Applikation eine Anpassung an unterschiedliche Helligkeitsverhältnisse ($\rightarrow$ Apertursteuerung) oder unterschiedliche Objektweiten ($\rightarrow$ Fokussierung) notnwendig wird.

9.4 Steuerung der Belichtung und Shutter

Grundsätzlich muss die Belichtungszeit der Sensoren regelbar sein, um den Sensor bei verschiedenen Beleuchtungsverhältnissen einsetzen zu können. In aller Regel werden *elektronische Shutter* verwendet, die wesentlich genauer als mechanische *Shutter* arbeiten.

Für CCD und hochwertige CMOS Sensoren ist die gleichzeitige Belichtung aller Pixel (sogenannter *global Shutter* oder *Snapshot Modus*) üblich und in der Regel die beste Wahl.

Wenn die Belichtung aller Pixel bei einem CMOS Sensor gleichzeitig erfolgt und die Pixel dann ausgelesen werden sollen ergibt sich das Problem, dass das Adressieren, aber auch die Verarbeitung (AD-Konversion, Weiterleitung an Schnittstelle) nur sequentiell erfolgen kann. Es ergibt sich also hinsichtlich der Verarbeitung ein Flaschenhals.

Konzeptionell die einfachste, aber lichtineffiziente Lösung, ist es, eben entsprechend lange mit der Aufnahme des nächsten Bildes zu warten. Das bedeutet aber, dass während der sich ergebenden Totzeit der Belichtung einfallende Photonen nicht genutzt werden. Es ist auch möglich, für jeden Pixel eine zusätzliche Speicherzelle beim Pixel zu verwenden und so einen globalen Shutter zu realisieren, dies führt aber zu deutlich erhöhten Kosten und einem eventuell erhöhten Rauschen.

Üblicherweise wird daher ein *Rolling Shutter* Modus (oder *Focal plane shutter*) eingesetzt. Hierbei werden die Zeilen zu unterschiedlichen Zeiten aktiv geschaltet bzw. belichtet. Dadurch gibt es einen Zeitversatz zwischen den Zeilen. Problematisch ist diese Vorgehensweise bei bewegten Objekten und kann zu massiven Verfälschungen der Geometrie führen (Abb. 9.6.11) die dann digital so gut wie möglich ausgeglichen werden müssen.

Andererseits ergibt sich beim Rolling-Shutter Modus ein verringertes Ausleserauschen, eine verbesserte Quanteneffizienz und evtl. auch eine höhere Bildwiederholfrequenz (frames per second, fps). Viele Sensoren bieten beide Bleichtungsmodi an und es ist stark anwendungsabhängig, welcher dann die bessere Wahl ist.[5]

Beachten Sie, dass ein erheblicher Unterschied zwischen Belichtungszeit, Bildwiederholfrequenz und Latenz besteht. Zunächst muss die Belichtungszeit natürlich kürzer sein als das Inverse der Bildwiederholfrequenz. Sie kann aber natürlich erheblich kürzer sein.

Unabhängig davon dauert es, bis ein belichtendes Photon letztlich von der Soft-

[5] Und schließlich sei erwähnt, dass einige Spezialanwendungen für hohe Geschwindigkeiten bzw. Scans, die zeitlich variable Belichtung des Sensors ganz gezielt ausnutzen.

Abbildung 9.4.8: Aufnahme eines Flugzeugpropellers mittels Rolling Shutter. Bild mit freundlicher Genehmigung von Jacoba Thissen.

ware registriert wird. Die komplette Signalkette innerhalb der Kamera über die Bilderfassungskarte über den Kameratreiber bis zum Betriebssystem kostet Zeit. Diese Zeit ist die *Latenz* und für echtzeitkritische Anwendungen ist diese Latenz natürlich von enormer Wichtigkeit. Bei konventionellen (nicht echtzeitfähigen) Betriebssystemen ist sie leider auch nur näherungsweise konstant. Einige Kameras protokollieren innerhalb der Bilder den genauen Belichtungszeitpunkt, so dass zumindest nachträglich eine Korrektur möglich ist. Bei getriggerten Kameras kann der Takt der Aufnahme aber auch von außen vorgegeben werden.

9.5 Rauschquellen

Das *Signal-Rausch-Verhältnis* (engl. signal-to-noise ratio, SNR) ist einer der zentralen Faktoren zur Beurteilung von Bildsensoren. Wie im Kapitel 5 dargelegt hängt die in der Praxis erzielbare Auflösung bei der Abbildung feiner Details stark vom SNR ab. Starkes Rauschen reduziert die effektive Auflösung und erschwert ganz

generell die Auswertung von Bildern.

Konkrete Werte für die einzelnen Rauschangaben oder die Angabe des Signal-Rausch-Verhältnisses hängen *massiv* von der exakten Definition der Größen und vor allem der Vorgehensweise zu ihrer experimentellen Bestimmung ab. Beim Vergleich entsprechender Daten unterschiedlicher Hersteller muss man also sehr genau überprüfen, ob bzw. wie die Werte vergleichbar sind.

Die Spezifikationen zum Rauschen sind generell nicht immer einheitlich, meist erfolgt aber eine Angabe in RMS-Rauschelektronen. In der Regel gehen wir davon aus, dass die einzelnen Rauschanteile unkorreliert zueinander sind. In diesem Fall kann man dann direkt das Gesamtrauschen[6] als quadratisch gewichtete Summe der einzelnen Rauschquellen annehmen:

$$R = \sqrt{R_1^2 + R_2^2 + R_3^2 + ...} \tag{9.1}$$

Bei den Spezifikation ganz allgemein, von Bildsensoren aber insbesondere, ist darauf zu achten, ob die Angaben sich auf die typische Performance („nominal") oder aber eine garantierte Performance beziehen. Insbesondere hinsichtlich des Rauschens können einzelne Sensoren im Sinne einer Exemplarstreuung deutlich voneinander abweichen. Teilweise werden die Sensoren in verschiedenen Qualtitäsgraden angeboten.

Der maßgebliche Standard zur Definition der Rauschbeurteilung von Kameras ist die *EMVA1288* [2]. Idealerweise werden die einzelnen Rauschparameter (siehe unten) entsprechend dieser Richtlinie gemessen und angegeben, so dass dann eine Vergleichbarkeit verschiedener Kameras gewährleistet ist.

Im wesentlichen müssen zwei Hauptbestandteile des Rauschens in Betracht gezogen werden: Photonenrauschen und elektronisches Rauschen. Insbesondere das elektronische Rauschen wird dann üblich- und sinnvollerweise weiter unterteilt.

[6] bzw. den Erwartungswert des Gesamtrauschens

9.5.1 Photonenrauschen

Wir beginnen mit dem *Photonenrauschen*. Dieses Rauschen ist prinzipieller Natur und — zumindest in der industriellen Bildverarbeitung[7] — bestimmt es eine unüberwindbare Grenze für das SNR. Konventionelle Lichtquellen (eingeschlossen Laser) emittieren die Lichtteilchen nicht in gleichmäßigen Abständen sondern rein statistisch. Wenn während eines Zeitintervalls Δt im Mittel N Photonen an einem Detektor ankommen dann bedeutet das nicht, dass in einem weiteren Intervall Δt wieder N Photonen ankommen werden. Die Lichtaussendung erfolgt statistisch und N bezieht sich hier auf einen Mittelwert.

Die Standardabweichung der Photonenzahl ergibt sich dagegen zu $\Delta N = \sqrt{N}$. Das bedeutet, dass mit ca. 68% Wahrscheinlichkeit die gemessene Photonenzahl im Bereich $N - \Delta N \ldots N + \Delta N$ liegt.

Das SNR ergibt sich somit zu

$$SNR_P = \frac{N}{\Delta N} = \sqrt{N} \tag{9.2}$$

wenn wir die Unsicherheit in der Photonenzahl als „Rauschen" auffassen: **Je größer die Photonenzahl ist, desto höher wird das SNR.**

Zunächst ist die Photonenzahl pro Pixel durch die Sättigungskapazität des Pixels begrenzt. Wenn ein Sensor also Pixel mit einer Sättigungskapazität von z.B. 8.000 Elektronen aufweist, dann kann das erreichbare SNR unter idealen Bedingungen (ideale Belichtung und keinerlei elektronisches Rauschen) maximal $\sqrt{8000} \approx 90$ betragen. Generell gilt also, dass mit kleinen Pixeln, die dementsprechend eine geringe Sättigungskapazität aufweisen, kein hohes SNR erzielbar ist!

Obwohl es sich hier nun um Elektronen und nicht Photonen handelt ändert sich bei der statistischen Betrachtung nichts, denn die Photonen werden mit einer gewissen Wahrscheinlichkeit (*Quanteneffizienz*) in Elektronen gewandelt.

Künstlich kann man diese prinzipbedingte Grenze für das SNR allerdings durch Mehrfachbelichtung mit nachfolgender Mittelung erhöhen. Wenn also die aufzunehmende Szene zeitlich nicht variiert kann man z.B. in kurzer Folge 9 Bilder auf-

[7] Im Bereich der Quantenoptik kann durch sogenanntes „squzeed light" die Grenze (prinzipiell) durchbrochen werden

nehmen und diese Mitteln. Dadurch ergibt sich eine 9-fach erhöhte Photonenzahl und daher ein um den Faktor $\sqrt{9} = 3$ verbessertes SNR.

Die Photonenzahl selbst (so man sie z.B. für photometrische Rechnungen benötigt) bekommt man über den Umweg der Energie aus Leistungsangaben mittels folgender beiden Relationen für einzelne *Photonen*:

$$E = h\,\nu = 6.626 \cdot 10^{-34}\,\mathrm{J/s}\ \cdot \nu \tag{9.3}$$

$$c = 3.0 \cdot 10^{8}\,\mathrm{m/s} = \lambda\,\nu \tag{9.4}$$

Die Leistung ist natürlich die Energie pro Zeit, d.h. 1 mW entspricht bei einer Belichtungszeit von 1 Sekunde genau 1 mJ.

Mit den beiden Formeln können wir z.B.berechnen, dass λ = 500 nm einer Frequenz ν von 6·10^{14} Hz entspricht und sicht damit eine Energie pro Photon von 4.0 ·10^{-19} J (bzw. 2.5 eV) ergibt.

In guter Näherung kann man die Anzahl der Photonen pro Watt bei einer Wellenlänge mit

$$N = \lambda \cdot 5 \cdot 10^{18} \tag{9.5}$$

berechnen. D.h. dieselbe Leistung führt im längerwelligen Bereich zu mehr Photonen (und damit z.B. einem geringeren Photonenrauschen).

9.5.2 Dunkelstrom

Das *Dunkelstromrauschen* ist der Rauschanteil, der selbst dann auftritt, wenn keinerlei Licht auf den einzelnen Pixel trifft. Es ist praktisch — wie der Name sagt — das Rauschen des *Dunkelstroms*. Um den Dunkelstrom messen und dann vom eigentlichen Nutzsignal abziehen zu können werden meist einige zusätzliche, vom Licht abgeschirmte, Pixel auf dem Sensor verwendet.[8] Problematisch wird der Dunkelstrom vor allem bei längeren Belichtungszeiten (Sekunden). Er kann so viele Elektronen generieren, dass die Sättigungskapazität des Pixels überschritten wird und damit keine Detektion von Licht mehr möglich wird.

[8] Für bessere Resultate kann der Dunkelstrom bei Betriebsbedingungen für jeden Pixel gemessen und gespeichert werden.

Der Dunkelstrom und damit auch das Dunkelrauschen, das proportional zur Wurzel des Dunkelstrom ist, steigt (exponentiell) mit der Temperatur an, es handelt sich daher um ein *thermisches Rauschen*. Daher werden Anwendungen, die eine hohe Anforderung an die Empfindlichkeit oder das SNR haben oft mit gekühlten Sensoren[9] durchgeführt. Für Raumtemperaturen verdoppelt sich der Dunkelstrom mit der sogenannten *doubling temperature*. Sie liegt in der Größenordnung von 5 bis 10 K. Das Dunkelstromrauschen wird damit bei längeren Belichtungszeiten (z.B. mehrere Sekunden) und hohen Temperaturen oft der Hauptrauschanteil. Gekühlte Bildsensoren haben also vor allem bei geringem Licht mit langen Belichtungszeiten Vorteile. In diesem Fall sind auch CCD Sensoren teilweise immer noch den CMOS Sensoren überlegen.

Generell steigt der Dunkelstrom mit der Größe des Pixels. Allerdings kann sich für sehr kleine Pixel ein umgekehrter Effekt ergeben, weil die in diesen Pixeln sich ergebenden hohen elektrischen Feldstärken wieder zu höheren Rauschanteilen führen können.

Auch hohe Bildraten begünstigen das Auftreten des Dunkelstroms, denn hohe Geschwindigkeiten führen zu einer deutlichen Erwärmung des Sensors und damit zu einem hohen Dunkelstrom. So kann sich der Bildsensor durchaus auf über 40°C erhitzen. Daraus folgt, dass gegenüber den technischen Daten, die unter Umständen bei 20°C angegeben sind, mit einer starken Vergrößerung des Dunkelstroms zu rechnen ist.

Meist wird der Dunkelstrom in Elektronen pro Zeiteinheit angegeben, damit ein Vergleich mit der Sättigungskapazität des Pixels einfach möglich ist.

9.5.3 Advanced: Ausleserauschen

Rauschen kann natürlich entlang des gesamten Signalverarbeitungsweges entstehen. Für CMOS-Sensoren ist ein ausführliches Rauschmodell in [18] dargestellt. Wir wählen hier eine vereinfachte Betrachtung.

Von besonderer Bedeutung ist für typische (einige ms bis einige zehn ms) Belichtungszeiten die eigentliche Auslesestufe und die sich dort ergebenden Rausch-

[9] vergleichsweise unproblematisch für einen praktischen Betrieb sind Peltier-gekühlte Elemente, die für zweistufige Systeme z.B. -25° C erreichen und dann Dunkelströme von wenigen Elektronen pro Pixel und Sekunde liefern

quellen werden meist als *Ausleserauschen* (engl. *Read-out noise*, RON) bezeichnet.

Hauptanteile sind:

- *thermisches Rauschen* des Resetschalters des Feldeffekttransistors und aller Widerstände (Rauschspannung proportional $\sqrt{4kTRB}$ mit der effektiv genutzten Bandbreite B, die deutlich oberhalb der Pixelfrequenz (einige hundert MHz) des Bildaufnehmers liegt). Auch die letztlich zur Digitalisierung eingesetzten AD-Wandler haben einen thermischen Rauschanteil [18].
- *1/f-Rauschen*: Der Transistor bzw. Verstärker selbst hat zusätzliches 1/f-Rauschen[10]
- Rauschen des *Ausgangsverstärkers* (dieses ist ebenfalls teilweise thermisch)
- *Quantisierungsrauschen* (und zusätzliches Rauschen) des *AD-Wandlers*

Teilweise werden die thermischen Anteile allerdings auch nicht dazu gezählt, so dass das Ausleserauschen dann unabhängig von der Temperatur ist.

Das Ausleserauschen steigt mit zunehmender Pixelfrequenz. Für die technische Angabe des Ausleserauschens ist es also wesentlich, bei welcher Auslesefrequenz es angegeben ist. Teilweise werden unrealistisch tiefe Frequenzen verwendet, denn bei diesen Frequenzen ist das Ausleserauschen deutlich reduziert.

9.5.4 Advanced: Fixed-pattern Rauschanteile

Zusätzlich zu diesen Rauschanteilen, die letztlich für alle Pixel gleich sind ergibt sich noch Rauschen, das — bei gleicher Belichtung — für unterschiedliche Pixel variiert. Dieses *Fixed-Pattern Rauschen* (engl. fixed-pattern noise, FPN) ist somit ein Rauschmuster, das über dem Bild liegt. Es hat sowohl einen vom optischen Signal unabhängigen Anteil, die sogenannte *Dark signal non-uniformity*, DSNU, als auch einen Anteil, der mit zunehmender Anzahl an Elektronen zunimmt (*Photo Response Nonuniformity*, PRNU), also eine Variation der Empfindlichkeit der einzelnen Pixel. Insbesondere bei CMOS Sensoren, bei denen jeder Pixel eine eigene

[10] Ursache sind Kristalldefekte bzw. Verunreinigungen im Halbleitermaterial

Verstärkung hat ist dieser Effekt ausgeprägt.[11] Entsprechend können sich auch leicht streifenartige Unregelmäßigkeiten ergeben, wenn die jeweiligen Verstärker einer Zeile unterschiedlich sind.

Aber auch Unregelmäßigkeiten in den optischen Schichtdicken oder in der Dotierung der Halbleitermaterialien können zu dem Rauschen führen. Aufgrund der Komplexität der Ursachen ist das PRNU in der Regel nicht rein zufällig sondern weist komplexe Muster (z.B. Periodizitäten) auf. Insbesondere bei hohen Signalstärken ist das PRNU ein entscheidender wenn nicht sogar dominierender Rauschanteil und das PRNU nimmt mit der Photonzahl linear zu (im Gegensatz zum Photonenrauschen, das ja „nur" mit der Wurzel der Photonenzahl steigt). Zusammen mit dem Photonenrauschen ist es dann die entscheidende Größe. Dennoch wird das Signal-Rausch-Verhältnis bei hohen Signalstärken besser.

Teilweise werden beide dieser Rauschanteile, also DSNU und PRNU, als *scene noise* bezeichnet. Ein Ausgleich kann entweder schon auf dem Bildsensor (bei CMOS Sensoren) oder aber nachträglich in der Kamera oder im PC erfolgen (*Nonuniformity correction*)[12]. Bei sehr einfachen Systemen, die keine entsprechende Korrektur nutzen, ist der Effekt sehr deutlich störend sichtbar. Eine zeitliche Mittelung (Addieren mehrerer Bilder), die ansonsten bei Rauschen (oft) eine Verbesserung erzielt, kann nicht angewendet werden. Teilweise ergeben sich auch periodische Störungen, die besonders stark im Bild auffällig sind.

Für die Angabe der FPN sind sehr unterschiedliche Definitionen in der Praxis anzutreffen. Hierauf ist — wie immer — beim Vergleich verschiedener Sensorspezifikationen zu achten.

Einzelne defekte Pixel kann man (muss man aber nicht) als eine Sonderform des Fixed-pattern noise betrachten. Insbesondere bei Sensoren im SWIR Spektralbereich muss mit vielen defekten Pixeln gerechnet werden. Bei herkömmlichen Bildsensoren im sichtbaren Spektralbereich liegt die Anzahl der defekten Pixel bei wenigen Pixeln pro Sensor. Komplett defekte Pixel werden auch als *tote Pixel* be-

[11] An sich führt die Variation der Kennlinie nicht nur zu einer Änderung der Empfindlichkeit gegenüber Licht sondern auch der Dunkelstromanteil (und sein Rauschen) wird unterschiedlich stark verstärkt. Dennoch erfolgt die Definition eben in der Regel (und auch der EMVA1288) über den rein lichtabhängigen Anteil.

[12] Problematisch ist bei der Korrektur, dass aufgrund der unterschiedlichen Eindringtiefe verschiedener Wellenlängen der Effekt von der spektralen Zusammensetzung abhängt. Daher muss idealerweise die Kalibrierung mit derselben spektralen Zusammensetzung wie bei der Zielapplikation erfolgen. Dementsprechend ist eine feste Korrektur auf der Kamera für gute Korrekturen problematisch.

zeichnet. *Heiße Pixel* erzielen das volle Signal auch bei Dunkelheit und *warme Pixel* liefern einfach deutlich falsche Werte bei Belichtung.

Sogenanntes „Zeilenrauschen" (engl.: *row noise* bedeutet, dass alle Pixel einer Zeile noch einen bestimmten Rauschanteil zugeschlagen bekommen (für jede Zeile ein anderer)[13].

Und auch die Spalten können sich teilweise synchron verhalten und so einen *Column-Level noise* Anteil generieren. Wenn z.B. für Spalten jeweils nochmals eigene Verstärker verwendet werden, dann können diese so wieder eine unterschiedliche Kennlinie aufweisen und so ergibt sich wieder ein Fixed-pattern noise Anteil. Entsprechende linienhafte Störungen sind natürlich besonders auffällig bei visueller Betrachtung. Dementsprechend werden die Effekte idealerweise bereits in der Kamera ausgeglichen (Nachbearbeitung).

9.5.5 Advanced: Quantisierungsrauschen

Durch die Quantisierung bei der Digitalisierung auf diskrete Werte ergibt sich ebenfalls ein Rauschen, das sogenannte *Quantisierungsrauschen*. Die Standardabweichung dieses Rauschens beträgt (siehe z.B. [12])

$$N_Q = \frac{1}{\sqrt{12}} \tag{9.6}$$

Allerdings ist zu beachten, dass beim eigentlichen Analog-Digital Wandlungsprozess oft noch deutliches zusätzliches Rauschen entsteht, das weit über das rein mathematische Quantisierungsrauschen hinaus geht. Qualitativ hochwertige AD-Wandler sind hier also das entscheidende Mittel zur Rauschreduktion.

9.5.6 Zusammenfassendes Rauschmodell

Bildsensoren sind komplexe Bauteile bestehend aus vielen Millionen von elektronischen Komponenten. Dementsprechend komplex ist das Rauschverhalten des

[13] Ursachen sind nach Gow das (vor der Belichtung notwendige) Zurücksetzen der Pixel über eine (pro Zeile) gemeinsame Leitung [18]

einzelnen Bildsensors. Für Details sei z.B. auf den hervorragenden Artikel von Gow et al. verweisen [18].

In den allermeisten Fällen ist es ausreichend, diese Komplexität durch ein vergleichsweise einfaches Modell, z.B. EMVA1288, zu erfassen.

Abb. 9.5.9 zeigt ein entsprechendes (leicht abgewandeltes[14]) Modell einer typischen Signalkette eines Bildaufnehmers hinsichtlich des Rauschens. Leichte Variationen werden von verschiedenen Autoren vorgenommen (z.B. PRNU Einbeziehung bereits vor Dunkelstromanteil).

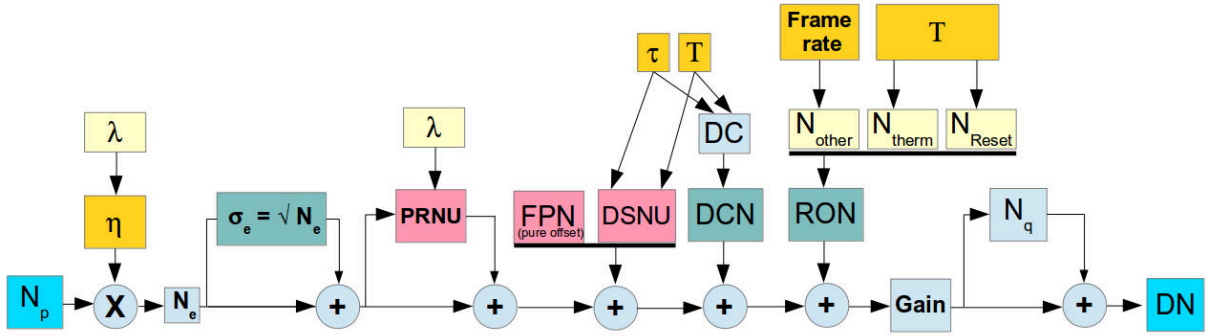


Abbildung 9.5.9: Darstellung der Signalkette für einen Pixel zur Rauschbestimmung. In Orange dargestellt sind die jeweiligen Einflußfaktoren (λ = Wellenlänge, T = Temperatur, τ = Belichtungszeit. In Rot dargestellt sind Fixed-Pattern Rauschanteile. N_P und N_e bezeichnen die Anzahl von Photonen und Elektronen.

Aber auch diese „einfachen" Modell sind oftmals schon vergleichsweise komplex und nicht jeder Hersteller gibt entsprechende Parameter für seine Sensoren an.

Generell gilt: Wenn genügend Licht vorhanden ist, dann ist das Photonenrauschen die maßgebliche Größe. Wenn dieses Photonenrauschen „deutlich" größer als die elektronischen Rauschquellen ist spricht man vom Schrotrausch-begrenzten Fall (shot-noise limited detection). Natürlich sollte man genau diesen Fall (möglichst viel Licht) anstreben, denn dadurch wird das SNR maximiert.

Bei kurzen und mittleren Belichtungszeiten ist ein Hauptaugenmerk auf das Ausleserauschen zu richten. Bei langen Belichtungszeiten ist vor allem das thermische Rauschen entscheidend.

Zusammenfassend kann man die verschiedenen Rauschanteile durch die folgende

[14] Kombination von EMVA1288 [2] und [29]

Formel näherungsweise beschreiben:

$$\mathrm{SNR} = \frac{\mathrm{S}}{\mathrm{N}} = \frac{\mathrm{P\ QE\ t}}{\sqrt{(\mathrm{P}+\mathrm{B})\ \mathrm{QE}\cdot\mathrm{t} + \mathrm{DC\ t} + \mathrm{RON}^2}} \tag{9.7}$$

Üblicherweise rechnet man in Elektronen und pro Pixel. QE bezeichnet die Quanteneffizienz der Lichtwandlung, also die Wahrscheinlichkeit der Ladungsgenerierung pro einfallendem Photon. Sie liegt typischerweise im sichtbaren Spektralbereich zwischen 0.4 und 0.9. DC ist der Dunkelstrom (dark current), der in Elektronen pro Sekunde angegeben ist. P ist die einfallende Signalleistung auf den Pixel während B die Leistung des (meist störenden) Hintergrundslichts angibt. RON schließlich bezeichnet das Ausleserauschen (read-out noise) des Pixels und t ist die Belichtungszeit.

Fixed-pattern Noise Anteile können prinzipiell kalibriert werden und sind in der Formel nicht enthalten, können aber einfach (quadratisch) hinzugefügt werden. Bei eigenen Rechnungen muss darauf geachtet werden, ob in Photonen, Elektronen oder Digitalwerten gerechnet wird. Alle drei Varianten sind möglich, aber natürlich müssen dann eben alle Anteile auch jeweils passend vorliegen.[15] Das Quantisierungsrauschen ist meist ohne Belang denn oftmals werden AD-Wandler mit hoher Bitzahl eingesetzt und im Vergleich zu den anderen Rauschquellen ist dieser Anteil dann vernachlässigbar.

Für brauchbare Bilder ist ein SNR von 10 erforderlich und „gute" Bilder benötigen ein SNR von 40 [26]. Ein SNR von 1 wird erzielt, wenn man mit der *noise equivalent exposure* (NEE) bzw. *noise equivalent power* belichtet.

Für die reinen Photodetektoren wird oft das Verhältnis von NEE zur Wurzel der Pixelfläche und Bandbreite angegeben, denn üblicherweise steigt das quadrierte Rauschen linear mit der Detektorfläche und der Bandbreite.

Abb. 9.5.10 zeigt den Vergleich von drei Bildsensoren hinsichtlich des Signal-Rausch-Verhältnisses in Abhängigkeit von der Bestrahlungsstärke, simuliert bei einer Wellenlänge von 800 nm und einer Belichtungszeit von 1/60 Sekunde in doppeltlogarithmischer Darstellung. Die sehr teure Scientific-CMOS Kamera (PCO edge 3.1) liefert generell ausgezeichnete Ergebnisse. Diesen Vorteil kann die Kamera aber vorrangig bei sehr wenig Licht ausspielen. Bei mittleren Bestrah-

[15] Man kann natürlich nicht einfach das Rauschen in Digitalwerten mit dem Rauschen in Photonen addieren.

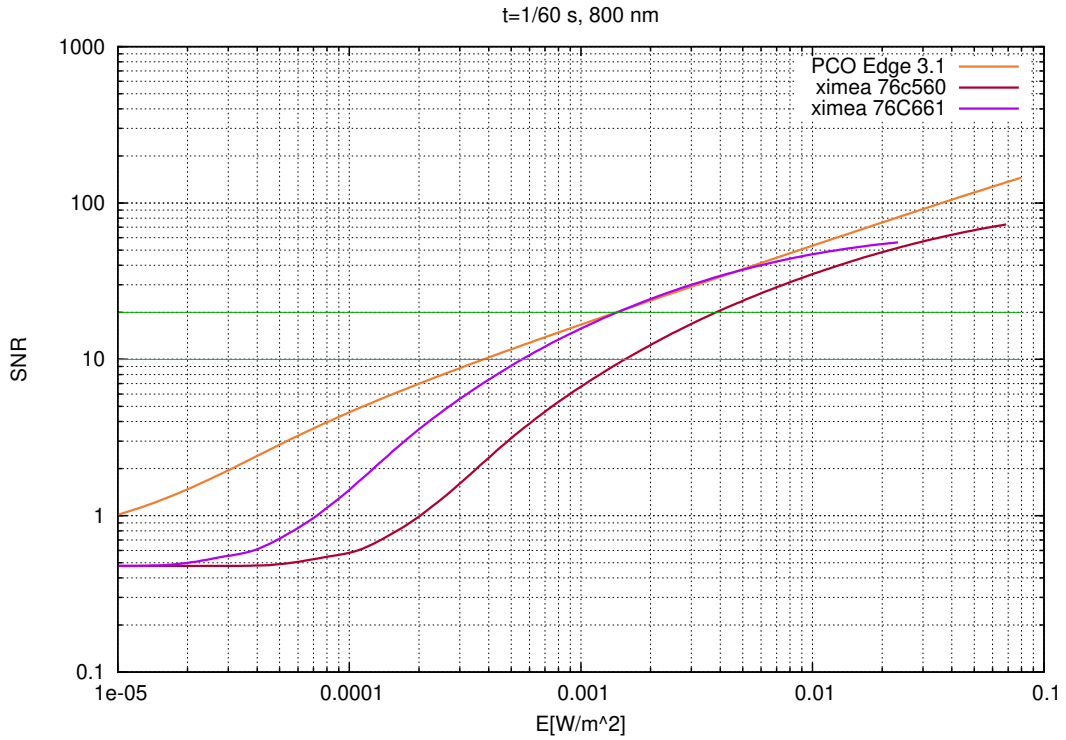


Abbildung 9.5.10: Das erzielbare Signal-Rausch-Verhältnis hängt stark von den jeweiligen Einsatzbedingungen ab (simulierte Kurven basierend auf den technischen Daten). Beachten Sie die doppelt-logarithmische Auftragung. Die Unterschiede sind beträchtlich.

lungsstärken ist die NIR-enhanced Ximea Kamera praktisch genauso gut (kostet aber nur 1/10 der PCO). Bei hohen Bestrahlungsstärken wird das PRNU wieder ausschlaggebend und die PCO kann wieder einen Vorteil verbuchen.

Deutlich wird, dass das erzielbare Signal-Rausch-Verhältnis extrem vom jeweiligen Einsatz abhängt. Es ist also ratsam, genau zu analysieren, was genau benötigt wird und welche Bedingungen vorliegen.

9.6 Advanced: Dynamik

Die *Dynamik* eines Sensors ist das Verhältnis von maximaler zur minimalen Bestrahlungsstärke, die (bei gleicher Belichtungszeit) erfasst werden kann. Sollen Szenen erfasst werden, bei denen gleichzeitig sehr geringe und sehr hohe Strahldichten vorliegen, ist entsprechend eine hohe Dynamik notwendig.

Die pro Belichtung maximal mögliche Anzahl an pro Pixel gespeicherten Elektronen, also die Sättigungskapazität, liegt üblicherweise im Bereich von ca. 1.000 bis 2.000 Elektronen pro $\mu\mathrm{m}^2$.[16] Zeilensensoren mit sehr großen Pixel erreichen so Sättigungskapazitäten von bis zu einigen Millionen Elektronen pro Pixel. Die Belichtung, die gerade diese Sättigungskapazität mit Ladungsträgern füllt wird als *saturation equivalent exposure* (SEE) angegeben. Für sehr kleine Pixel, wie sie z.B. in einfachen Handys Verwendung finden, ergeben sich Sättigungskapazitäten von einigen 1000 Elektronen. Bei einem Rauschanteil von einigen 10 Elektronen werden so Dynamik und SNR (vgl. Abschnitt 9.5) massiv beschränkt. Bei Standard-Bildverarbeitungssensoren liegen die Sättigungskapazitäten meist im Bereich von ca. 10.000 Elektronen. Sensoren, die ein sehr gutes Signal-Rausch-Verhältnis erzielen sollen, nutzen Kapazitäten von einigen 10.000 Elektronen.

Wenn mehr Elektronen (z.B. aufgrund einer sehr starken Belichtung) an einem Pixel generiert werden, dann können diese Elektronen in benachbarte Pixel abfließen. Bei CCDs ist der Effekt in Richtung der Ladungsschiebung, also in Spaltenrichtung extrem stark während geeignete Abschirmungen zwischen den verschiedenen Spalten möglich sind. Dementsprechend ergibt sich im Bild eine vertikale helle Linie an der Stelle, wo die starke Überbelichtung stattfindet. Man spricht von *Blooming*.

Systeme mit großer Dynamik werden oft unter dem Begriff *High Dynamik Range - HDR* beworben.

Für die quantitative Angabe der Dynamik ist es nun entscheidend, was man als minimales, kleinstes Signal annimmt. Hier bestehen verschiedene Möglichkeiten und man sollte sich klar machen, was für die konkrete Anwendung relevant ist und beim Vergleich verschiedener Sensoren auf die konkrete Definition achten.

[16] Für CMOS Sensoren ist der Wert aktuell besser als für CCD Sensoren.

Als kleinstes Signal kann man annehmen:

Rauschen: Wenn als Bezugsgröße die Stärke des Rauschens, die sogenannte *noise equivalent power*, verwendet wird dann entspricht die Dynamik dem maximal erzielbaren SNR.

1 Bit: Oftmals wird vom digitalen Signal ausgegangen und als Bezugsgröße ein Bit (bzw. „1 Digital Number" (DN)), also die kleinste digitale Quantisierungsstufe, verwendet.

Ausleserauschen: Hier wird das Ausleserauschen als Bezugsgröße gewählt. Das Dunkelstromrauschen ist stark abhängig von Temperatur und Belichtungszeit und stört daher beim Vergleich. Das Ausleserauschen ist dann sinnvoll, wenn kurze Belichtungszeiten verwendet werden können.[17]

Quantisierungsrauschen: Hier wird das Quantisierungsrauschen $1/\sqrt{12}$ als Bezugsgröße gewählt.

Die Definitionen unterscheiden sich ganz erheblich und mit der letzten Definition lassen sich natürlich die beeindruckendsten Werte angeben. Man kann hier auch zwischen *intra-scene* und *pixel-basierter* Dynamik unterscheiden. Wenn man die erste Definition (mit dem gesamten Rauschen) verwendet, dann ist das unter Umständen zu negativ. Denn: in einem dunklen Bildbereich ist natürlich auch das Photonenrauschen geringer und dort ist eben vor allem das Ausleserauschen und thermisches Rauschen relevant. Wesentlich ist jedenfalls, nicht Äpfel mit Birnen zu vergleichen, d.h. die Definition der Dynamik (oder auch des SNR) muss klar sein.

Die Dynamik wird meist logarithmisch in Dezibel (dB) mittels

$$D = 20 \log \frac{I_{max}}{I_{min}} \tag{9.8}$$

angegeben.[18]

[17] Allerdings ist auch das Ausleserauschen nicht wirklich konstant sondern abhängig von der Bildwiederholfrequenz (*nicht* der Belichtungszeit).

[18] Der Faktor 20 statt 10 kommt dadurch zustande, dass letztlich der Grauwert des Kamerabildes proportional zu einer Spannung ist, Dynamiken aber traditionell auf Leistungen bezogen sind.

Für den meist angegebenen Fall der Definition über das Quantisierungsrauschen (größte Dynamikwerte) erhält man für einen N-Bit Sensor mit $I_{max} = 2^N$ und $I_{min} = 1/\sqrt{12}$

$$D = 20 \log_{10} 2^N \sqrt{12} = 20 \cdot N \cdot \log_{10} 2 + 20 \cdot 0.5 \cdot \log_{10} 12 \tag{9.9}$$

$$= = 6\,N + 10.8\,\mathrm{dB} \tag{9.10}$$

Für einen 8 Bit Sensor ergeben sich so 59 dB und für ein 12 Bit System 83 dB. In der Praxis ist diese Definition der Dynamik leicht irreführend, denn letztlich nutzbar ist selbst bei sonstiger kompletter Rauschfreiheit bei einem 8 Bit System eben nur eine Dynamik von 8 Bit während die 59 dB einer effektiven Dynamik von fast 10 Bit entsprechen würden (ld 887 $\approx$ 10). Wenn man als Bezugspunkt 1 Bit werden entfällt die Addition von 10.8. Üblicherweise (Dominanz des Photonenrauschen) erhält man damit eine hohe Dynamik nur für Pixel mit einer ausreichend hohen Sättigungskapazität.

Beachten Sie, für optische Leistungen wird meist die Definition

$$\mathrm{x}\,[\mathrm{dB}] = 10\, log_{10} \frac{P_2}{P_1} \tag{9.11}$$

verwendet.

3 dB entsprechen damit gerade einem Faktor 2, 10 dB sind ein Faktor 10 und 1 dB ist ein Faktor von 1.25. Für gute Bilder fordert man hinsichtlich der optischen Leistung in etwa 20 dB, d.h. eine Hell-zu-Dunkel Dynamik von etwa $P_2/P_1 = 100$. Bei einer Angabe für die Kamera werden daraus dann aber eben 40 dB. Achten Sie also in jedem Fall auf die Definition bei Angaben in dB.

9.6.1 Wieviele Bits?

Wieviele Bits sollte nun ein Bildsensor haben? Für ein überall gut ausgeleuchtetes Bild ist das Signal-Rausch-Verhältnis ja gerade SNR = $\sqrt{(}QWC)$. Damit würde man für übliche Sensoren gut mit 8 Bit auskommen. Vielfache von 8 Bit (= 1 Byte) sind in der Computertechnik besonders angenehm, denn die üblichen Prozessierungs- und Speicherarchitekturen sind eben in solchen Vielfachen, also in Bytes, organisiert. Für 10 Bits braucht man also in der Regel dann schon zwei

Bytes, d.h. im Vergleich zu 8 Bit ergibt sich oft eine verdoppelte Speicher- oder Prozessierungsanforderung.

Aber: In Szenenbereichen mit wenig Licht ist zwar das SNR, aber eben auch das Rauschen reduziert. Wenn ich in einem dunklen Bereich lediglich 100 Elektronen in einem Pixel sammle, dann erwarte ich dort ein Photonenrauschen von 10 Elektronen. Bezogen auf eine Sättigungskapazität von z.B. 10.000 Elektronen sind das 0.1 %. In diesem Fall macht es also Sinn (unter Vernachlässigung von anderen Rauschquellen), 10 Bits zu verwenden. Bei der Verwendung von 8 Bit wird in dunklen Szenenbereichen das SNR nicht optimal genutzt.

Allerdings ist es stark anwendungsabhängig, ob dieser Wechsel von 8 nach z.B. 10 Bit mit der Verbesserung in dunklen Bereichen, den erhöhten Aufwand rechtfertigt.

Eine Variante kann sein, die Daten zunächst (vor dem Transport und der Speicherung) durch eine nichtlineare Funktion zu transformieren. Z.B. bietet sich bei Dominanz von Photonenrauschen eine Wurzelfunktion an.[14]

9.6.2 Nichtlineare Kennlinien

Meist wünscht man bei der industriellen Bildverarbeitung eine lineare Kennlinie und dementsprechend sollte man — je nach Algorithmik, die auf die Bilder angewendet wird — auf eine entsprechende Angabe achten. Insbesondere nahe der Sättigung und im Bereich kleinster Signal (hier dominiert unter Umständen der Dunkelstrom) können die Pixel ein nichtlineares Verhalten zeigen.

Es gibt aber auch Anwendungen, bei denen eine deutliche Nichtlinearität gewünscht ist, z.B. bei der Betrachtung einer Szene mit sehr hellen Anteilen (z.B. Glanzlichter) und gleichzeitig Schattenbereichen, die ausgewertet werden sollen. Verschiedenste Varianten sind für die Kennlinie möglich.

Bei *nichtlinearen Kamerakennlinien* ist wichtig, dass eine hohe Dynamik allein oft nicht ausreicht, um eine Bildverarbeitungsaufgabe zu lösen. Entscheidend ist auch, mit wie vielen quantisierten Stufen der Helligkeitsverlauf beschrieben wird.

Bei einer äquidistanten Wahl der Quantisierungsstufen ergibt sich das „Problem",

dass geringe Helligkeitsunterschiede in dunklen Bereichen einer Szene kaum oder gar nicht zu getrennten digitalen Pixelwerten führen. Üblicherweise werden entsprechend nichtlineare Kennlinien meist für die qualitative Bildverarbeitung eingesetzt während man innerhalb der üblicherweise gut kontrollierbaren Bedingungen innerhalb einer industriellen Anwendung eine ausreichende lineare Dynamik einsetzt. Allerdings ist zu beachten, dass es auch hier gewisse unerwünschte Abweichungen von der Linearität (z.B. durch Nichtlinearitäten des AD-Konverters bedingt) gibt, die im Zweifelsfall experimentell bestimmt werden müssen, um sie nachfolgend ausgleichen zu können.

Der Dynamikbereich kann bei einem stationären Objekt auch durch eine Mehrfachaufnahme mit variabler Belichtungszeit erhöht werden. Hierbei werden die aufgenommenen Bilder digital zu einem Einzelbild mit erhöhter Dynamik kombiniert. Diese kann sowohl PC-seitig als auch schon kameraseitig erfolgen. Teilweise werden die Ergebnisse für zwei unterschiedliche Belichtungszeiten in einen einzelnen 16 Bit Grauwert kodiert.

Nichtlineare Kennlinie sind nicht nur zur Vergrößerung der Dynamik üblich sondern werden bei analogen Videokameras, insbesondere im Überwachungsbereich, oft eingesetzt. Monitore haben üblicherweise eine nichtlineare Kennlinie hinsichtlich der Helligkeit, um dem menschlichen Betrachter, dessen Sehsystem ebenfalls eine nichtlineare Kennlinie aufweist, einen möglichst ausgewogenen Seheindruck mit hohem Informationsgehalt zu liefern. Wenn dementsprechend mittels analoger Adressierung ein Monitor direkt angesteuert wird, dann werden oft Kameras mit entsprechend nichtlinearer Kennlinie verwendet, um diesen Effekt auszugleichen. Meist wird die Kennline gemäß einer *Gamma-Kurve*

$$I' = \alpha V^{\gamma} + I_0 \tag{9.12}$$

gewählt. Teilweise kann auch zwischen einer linearen (Gamma = 1) und einer nichtlinearen (Gamma $\neq$ 1) Kennlinie umgeschaltet werden.

Moderne Kameras bieten manchmal die Möglichkeit, direkt eine *Lookup-Tabelle* zu programmieren, um einen Spannungswert in einen beliebigen Digitalwert zu wandeln.

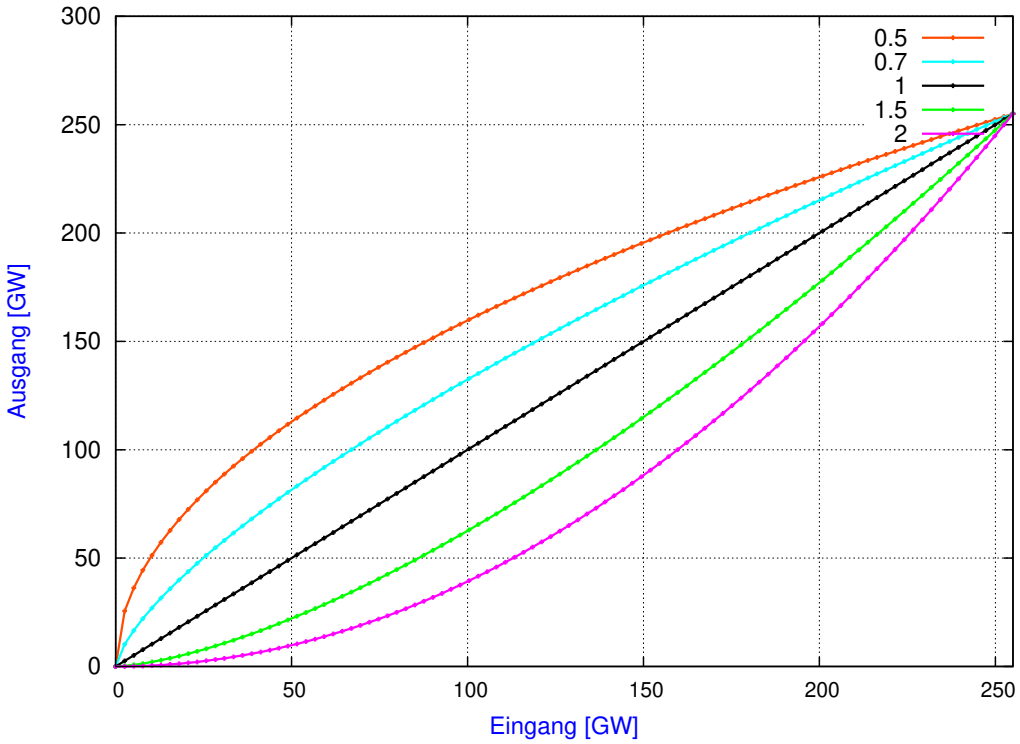


Abbildung 9.6.11: Verschiedene Gamma-Kennlinien

9.7 Empfindlichkeit

Die *Sensitivität* (engl.: sensitivity) gibt die für eine gerade noch detektierbare Änderung notwendige Bestrahlungs- oder Beleuchtungsstärke des Sensors an. Was dabei gerade noch detektierbar ist, ist wieder nicht klar definiert. Üblich ist als Definition ein SNR von 1 oder aber die Erzielung eines bestimmten prozentualen Ausgabewert (z.B. 10 % des maximalen Werts). Die Angaben können sich auf photometrische Größen (also z.B. lux s) oder radiometrische Größen (J/cm^2) Größen beziehen.

Noch vielfältiger sind die Angaben bei der *Responsitivity*. Sie gibt an, zu welche Änderung eine Variation der einfallen Bestrahlungs- oder Beleuchtungsstärke führt. In aller Regel verwendet man Sensoren mit linearer Kennlinie so dass die Angabe der Steigung der Kennlinie ausreichend ist. Eine typische Angabe ist 5 DN/(nJ/cm^2). DN bezeichnet dabei *Digital Number*, ist also einfach die Änderung des digitalisierten Werts. 5 DN/(nJ/cm^2) bedeutet also, das sich die Ausgabe des Sensors um 5 (Digitalwerte) ändert, wenn sich die Bestrahlungsstärke um 1 nJ/cm^2 ändert.

Teilweise erfolgt eine Angabe der Sensitivität in der Form 80IRE @ xx lux sekunden. In diesem Fall wird z.B. 80% der maximalen Signalstärke (z.B. 80% von 256 bei einem 8 Bit System oder 80% von 1 V bei einem analogen Videosystem) bei einer Beleuchtungsstärke von xx lux und einer Belichtungszeit von 1 Sekunde erhalten. Alternativ kann die Angabe auch eine Signalstärke (in lux sekunden) sein, die zu einem gewissen Signal-Rausch-Verhältnis, z.B. 30, führt [33].

Zu beachten ist bei entsprechenden Angaben, dass ein Vergleich nur bei gleicher elektronischer *Verstärkung* (*Gain*) bzw. eigentlich gleichem Signal-Rausch-Verhältnis sinnvoll ist. Natürlich lässt sich die Signalstärke durch den Verstärker des Bildsensors steigern. Gleichzeitig wird aber das Rauschen gesteigert, das Signalrauschverhältnis wird also durch die Verstärkung nur verbessert, wenn das Quantisierungsrauschen einen wesentlichen Anteil bildet.

Eine automatische Anpassung der Verstärkung (*Auto Gain Control*, AGC) kann aber dennoch bei variablen Beleuchtungsverhältnissen günstig sein, um die weitestgehend konstante Bildhelligkeiten für die nachfolgende Bildverarbeitung sicherzustellen und die Sättigungskapazität des Sensors sinnvoll zu nutzen (Photonenrauschen).

Ganz generell ist bei der Angabe der Empfindlichkeit darauf zu achten, ob diese pro Photon, also als Quanteneffizienz oder aber in Bezug auf die Beleuchtungsstärke (Photometrie) oder die Bestrahlungsstärke (Radiometrie) angegeben ist. Bei der Angabe in Bezug auf das Photon ist zu beachten, dass aufgrund der Photonenenergie $E = h\nu = hc/\lambda$ kurzwelliges Licht bei gleicher Leistung weniger Photonen enthält als langwelliges Licht.

9.7.1 Advanced: Details zur Empfindlichkeit

Die spektrale Empfindlichkeit eines Bildsensors hängt von verschiedenen Faktoren ab, zunächst entscheidend ist aber das Grundmaterial. Außer für Spezialanwendungen werden in der Regel Polysilizium Detektoren verwendet und diese sind von ca. 300 bis ca. 1100 nm empfindlich (siehe z.B. Abb. 9.7.12). Während der Abfall nur höheren Wellenlängen bei traditionellen CCD Sensoren vergleichsweise hoch ist, erzielen aktuelle CMOS Sensoren in diesem Bereich leicht verbesserte Werte.

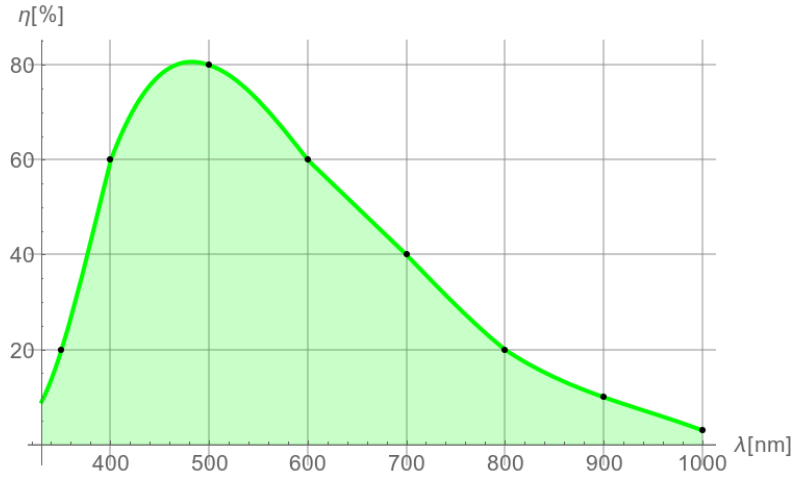


Abbildung 9.7.12: Typische spektrale Empfindlichkeit eines siliziumbasierten Bildsensors. Das Maximum kann irgendwo im Bereich 500 bis 800 nm liegen; besonders häufig liegt es aber im grünen Spektralbereich (ca. 500 nm). Beachten Sie: Bei einer Auftragung des Stroms pro einfallendem Watt optischer Leistung (anstatt wie hier der Quanteneffizienz) ergibt sich ein komplett anderer Verlauf bei dem das Maximum im nahen Infrarot liegt. Dies liegt daran, dass pro W Leistung wesentlich mehr Infrarot-Photonen als z.B. blaue Photonen emittiert werden ($E = h\,\nu$).

Für *UV-empfindliche Sensoren* kann eine *Fluoreszenzschicht* (z.B. *Lumogen*) vor den eigentlichen Bildsensor gebracht werden. Die Quanteneffizienz beträgt für entsprechende UV-Photonen (ab ca. $\lambda = 120$ nm) dann in etwa 10 bis 15 %. Allerdings muss mit einer Reduktion der räumlichen Auflösung gerechnet werden, denn eine gewisse Dicke der Schicht ist notwendig, um eine ausreichende Effizienz für die Fluoreszenz zu erhalten.

Traditionell werden Sensoren mit Beleuchtung „von vorne" (*front-side illumination*) eingesetzt. Dabei muss das einfallende Licht zunächst durch die Adressierungselektronik des Sensors propagieren, bevor es in der eigentlichen photosensitiven Schicht in Ladungsträger gewandelt wird. Damit diese Adressierungsschichten nicht stören wird durch ein passendes Design der *Mikrolinsen* vor dem Pixel dafür gesorgt, dass das Licht möglichst wenig mit den Adressierungsstrukturen wechselwirkt. Je kleiner die Pixelgeometrien werden, desto problematischer ist dieser Ansatz, insbesondere wenn die Pixelbreite in den Bereich von wenigen Wellenlängen vor stößt und ein nicht geringer Teil des einfallenden Lichts an den Adressierungsstrukturen gebeugt wird und so wieder zu einem Übersprechen zu benachbarten Pixeln führen kann. Ein Ansatz ist die Verwendung von lichtleitenden Strukturen auf dem Pixel, aber diese führen zu einer verminderten Quantenef-

fizienz.

Darüber hinaus wird kurzwelliges Licht bereits sehr nah an der Isolationsschicht in Elektronen gewandelt. Rekombinationsprozesse führen dann zu einer effektiven Reduzierung der generierten Ladungsträger.

Da die Transmission von Silizium (insbesondere in diesem Fall der Silizium Elektroden) für sichtbares Licht nicht sonderlich gut ist[19] gibt es spezielle dünne Sensoren, bei denen dann eine Beleuchtung „von hinten" erfolgt. Diese *back-side illumination* (bzw. *back-thinned Detektoren*) führt zu höheren Effizienzen von bis zu 90% denn in diesem Fall muss das zu detektierende Licht nicht die Metallverdrahtung und die Elektronik passieren sondern trifft direkt die Photodiode. Allerdings ist die Herstellung entsprechender Sensoren aufwändiger.[20] Vor allem kurzwelliges Licht ist problematisch hinsichtlich einem Übersprechen zu benachbarten Pixeln und einer verringerten Quanteneffizienz.

Bei kleinen Pixelgeometrien ist auch zu beachten, dass die Abschirmung benachbarter Pixel eine Herausforderung darstellt.Für hohe Quanteneffizienzen sollte die Dicke der photosensitiven Schichte mehrere Mikrometer betragen, andererseits sind die Abstände Pixel sehr klein, eventuell nur 1 Mikrometer. Das Übersprechen kann hier durchaus im Bereich von 10% liegen.

Sensoren in höheren Wellenlängenbereichen benötigen andere Halbleitermaterialien. Besonders wichtig ist dabei *Indium-Galium-Arsenid* (InGaAs). Hiermit lassen sich Sensoren bis ca. 1700 nm herstellen. In diesem unteren „short-wave infrared" (SWIR) Spektralbereich lassen sich viele interessante Materialidentifikationen durchführen (vgl. Abschnitt 9.12). Leider sind die Sensoren bei relativ schlechter Auflösung sehr teuer (meist im Preisbereich deutlich oberhalb 10.000 Euro), so dass der Einsatz in der industriellen Bildverarbeitung nicht sehr verbreitet ist. Längere Wellenlängen werden selten eingesetzt, eine Übersicht entsprechender Sensoren findet sich bei Rogalski [42].

[19] sonst wäre eine effiziente Lichtdetektion an sich ja nicht möglich

[20] Z.B. muss eine kürzere Brennweite des Mikrolinsenarrays realisiert werden, was zu einer dickeren Schicht des Mikrolinsenmaterials führt. Außerdem muss ein zusätzliches Metallgitter zur lateralen Abschirmung der Pixel voneinander eingebracht werden.

9.7.2 Advanced: Ein Beispiel

Auch wenn die einzelnen Zusammenhänge, die sich für das Rauschen und die Empfindlichkeit ergeben, an sich nicht schwierig sind, sind die Überlegungen, die zu einer konkreten Auswahl führen in der Praxis dann doch meist komplex. Die sich ergebende Problematik soll an einer relativ simplen Anwendung aufgezeigt werden.

Hierzu nehmen wir an, dass es unsere Aufgabe ist, in einem Abstand von 1 m über einen Schärfentiefenbereich von 150 mm ein Objekt der Ausdehnung 200 mm zu erfassen. Ganz wesentlich ist dabei, welche objektseitige Auflösung wir dabei fordern. Wir gehen hier vereinfacht von einem Durchmesser eines verschmierten Punktbilds von 0.2 mm aus.

Für einen Sensor mit 5 mm Kantenlänge ergibt sich sofort der Abbildungsmaßstab $\beta' =$ -5 mm / 200 mm = -0.025. Damit wird im Bildraum die Unschärfe $B' = B \cdot \beta' = 5\,\mu\mathrm{m}$. [21]

Mit Gl.2.12 können wir nun die minimal mögliche Blendenzahl berechnen:

$$\delta z = 150 \text{ mm } = 2K\frac{B'}{\beta'^2} \tag{9.13}$$

Also

$$K = \frac{\delta z\ \beta'^2}{4B'} \approx 4.7 \tag{9.14}$$

Wir können damit einen konkreten Bildsensor und eine konkrete Optik suchen. Betrachten wir den Bildsensor: Im einfachsten Fall (wir machen uns hier nun keine detaillierten Gedanken zur lateralen Auflösung) können wir hier von ca. 1000 x 1000 Pixeln ausgehen (1 Pixel pro Verschmierungskreis).[22]

Die *Belichtungszeit* muss (anwendungsbedingt kurz sein, z.B. 1/100 Sekunde). Eine mögliche preisgünstige Kamera wäre die Ximea xiQ mit 1280 x 1024 Pixeln (5.3 μm Pixelgröße).

Nun ist es entscheidend festzulegen, welches Signal-Rausch-Verhältnis eigentlich

[21] Gegencheck: Im Objektraum werden 1000 Punkte getrennt (Feld = 200 mm und Auflösung 0.2 mm. Im Bildraum werden ebenfalls 1000 Punkte getrennt (Feld = 5 mm, Auflösung 0.005 mm).

[22] Wie im Kapitel zur Auflösung besprochen ist das nicht sonderlich sinnvoll, aber es soll hier vorrangig um die Empfindlichkeit gehen.

für die konkrete Anwendung notwendig ist. Dies hängt natürlich von der Anwendung selbst und der Auflösung ab. Hier kommen in komplexer Weise alle Anforderungen der Anwendung aber auch Ergebnisse der Systemauslegung (Pixelzahl) in Spiel und eine seriöse Auslegung würde an sich verlangen, dass hier — basierend auf Simulationen mit der geplanten eingesetzten Bildverarbeitung — untersucht wird, welches SNR zu fordern ist. Das würde hier jetzt zu weit führen und wir nehmen als brauchbaren Wert 20 an. Für den gewählten Sensor ergibt sich, dass wir eine Bestrahlungsstärke (bei λ =600 nm) von mindestens $E_{e,sensor}$ =0.002 W/m^2 benötigen.

Als nächstes stellt sich die Frage, ob wir genügend Beleuchtungslicht E_e in unserer Anwendung zur Verfügung haben. Hierzu können wir die Gleichungen 7.11 und 7.12 nutzen:

$$E_{e,sensor} = 0.002\,\mathrm{W/m^2} = \frac{R\,E_e}{4K^2}, \tag{9.15}$$

dabei haben wir R die typische Reflektanz des Objekts bezeichnet. Mit R=0.5 und K=4.7 folgt hieraus $E_e = 1.4$ W$/m^2$.

(Für gelb grünes Licht wäre das immerhin eine Beleuchtungsstärke von 0.035 W$/m^2 \cdot 683$ lm/W = 960 lx).

Nehmen wir an, dass wir im Sinne einer Low-Light Anwendung nur 10 x weniger Licht zur Verfügung haben.

Naheliegend wäre zunächst nach einer Kamera mit einem verbesserten SNR zu suchen. Durch geeignete CCDs kann hier aufgrund der höheren Gesamteffizienz (meist rein als Quanteneffizienz angegeben) bereits ein Faktor von 2 erzielt werden. eine weitere klassische Lösung wäre es, einen Sensor mit größeren Pixeln oder aber doppelt so vielen Pixeln in Kombination mit Binning zu nutzen.

Eine letztlich größere Fläche A des Bildaufnehmers führt wegen $\Phi_e = E_e \cdot A$ zu einer Aufnahme von mehr Licht und damit einer Verbesserung des Rauschens.

Wir könnten eine sehr gute scientific CMOS Kamera, z.B. PCO Edge 3.1 einsetzen. Es zeigt sich, dass wir hier das geforderte SNR bereits bei $E_{e,sensor}$ =0.001 erreichen.[23] Wenn wir nun noch ein 2x2-fach Binning verwenden (die Kamera hat ca. 2000 x 2000 Pixel), dann ergibt sich ein weiterer Lichtgewinn um den Faktor 4,

[23]Die eigentliche Stärke der Kamera, auch bei langen *Belichtungszeiten* durch die eingesetzte Kühlung ein sehr rauscharmes Signal zu liefern ist hier nicht wirklich relevant.

so dass wir gegenüber der Ximea xiQ nun bereits einen Gewinn um den Faktor 8 erzielt haben.

Wesentlich war dabei die vergrößerte Sensorfläche. Allerdings führt diese leider auch zu einer Änderung der minimal möglichen Blendenzahl, denn mit der Sensorgröße ändert sich der Abbildungsmaßstab und damit dann bei gleichem Schärfentiefenbereich die minimal mögliche Blendenzahl. Bei einer Vergrößerung des Sensors um den Faktor 2 vergrößert sich dann also auch die Blendenzahl um den Faktor 2 und es ergibt sich kein Lichtgewinn und damit auch kein Gewinn beim SNR.

Der Vorteil zum großformatigeren Sensor hat hier also keinen Erfolg gezeigt weil die Anwendung letztlich durch die Schärfentiefe begrenzt war. Für andere Anwendungen wo die Blende durch die Auflösung (beugungsbedingt oder aberrationsbedingt) begrenzt ist, wäre der großformatige Sensor allerdings ein sehr guter Lösungsweg.

Hier muss stattdessen überdacht werden, ob bereits optisch mehr Signal gewonnen werden kann. Z.B. könnte evtl. ein Autofokussystem eingesetzt werden, um die Anforderung an die Schärfentiefe zu reduzieren.

Ein weiterer Punkt ist wichtig: Wenn in der Anwendung auch die halbe Auflösung noch möglich wäre, dann könnte man zunächst — auch mit dem xiQ Sensor — ein 2x2 Binning anwenden und hätte schon die 4-fache Lichtmenge bzw. ein um den Faktor 2 verbessertes SNR. Darüber hinaus würde aber auch die Blendenzahlbegrenzung relaxiert (ebenfalls um den Faktor 2). Dadurch würde ein weiterer Faktor 4 in der Lichtmenge gewonnen werden. Alles in allem hätte die Halbierung der Auflösungsanforderung also eine Verbesserung des SNR um den Faktor 4 zur Folge. Und das wiederum könnte durchaus für die Applikation (je nach eigentlicher Fragestellung) eine erhebliche Verbesserung darstellen. Wenn es z.B. eher um die Detektion von sehr kontrastarmen Features geht, dann kann der Gewinn an SNR einen Gewinn an lateraler Auflösung unter Umständen deutlich überwiegen.

Im folgenden Abschnitt wird eine andere Lösungsvariante, nämlich spezielle Sensoren, die auch bei kürzeren Belichtungszeiten gute SNR erzielen, besprochen.

9.7.3 Advanced: Bildsensoren mit hoher Empfindlichkeit

Konventionelle CCDs haben meist eine ausreichende Empfindlichkeit für die meisten industriellen Anwendungen. Die Beleuchtung wird ausreichend stark gewählt und meist ist es preisgünstiger, die Beleuchtungsstärke anstatt die Kameraempfindlichkeit um den Faktor zwei zu erhöhen. Für besonders kritische Anwendungen (meist die Detektion von sehr kleinen Defekte durch Dunkelfeldbeleuchtung bei kurzer Belichtungszeit) sind aber ebenfalls verschiedene Sensortechnologien verfügbar, mit denen deutlich geringere Beleuchtungsstärken aufgenommen werden können.

Generell können geringe Lichtsignale natürlich durch eine verlängerte Belichtungszeit ausgeglichen werden. In der industriellen Bildverarbeitung ist das logischerweise nicht immer möglich (begrenzte Zeit, Bewegungsunschärfe). Aber wenn es möglich wird, dann ist zu beachten, dass der Dunkelstrom linear (und damit auch das *Dunkelstromrauschen* mit der Wurzel) der Belichtungszeit anwächst. Oftmals ist daher eine *Kühlung* notwendig. Und natürlich ist jegliches Rauschen bei kleinen Signalen besonders nachteilhaft. Es ist damit dann meist eine gekühlte Kamera notwendig. Diese Kühlung kann rein *thermoelektisch* (*Peltier*) oder aber — für wissenschaftliche Fragestellungen und aufwändig — mittels flüssigem Stickstoff erfolgen. Bei thermoelektrischer Kühlung unterscheidet man hinsichtlich der Stufen.[24] Mit einer (typischen) Zweistufenkühlung können immerhin -25°C und damit bereits Belichtungszeiten im Minutenbereich erreicht werden. Bei einer Vierstufenkühlung kommt man in den (für die industrielle Bildverarbeitung uninteressanten) Bereich nahe einer Stunde. Eine starke Kühlung reduziert allerdings auch die Quanteneffizienz .

Für sehr hohe Empfindlichkeiten sind sogenannte *electron bombarded* CCDs (EBCCD) verfügbar. Dabei werden die generierten Elektronen durch eine hohe Spannung beschleunigt. Allerdings ergibt sich gegenüber back-thinned CCDs eine deutlich reduzierte Quanteneffizienz sowie eine Verschlechterung der Sensortransferfunktion, also eine Reduktion des Kontrastes für hohe Raumfrequenzen. Auch die Dynamik ist oftmals reduziert.

Die *Intensified CCDs* (ICCD, I^2CCD) nutzen den Lawineneffekt. Hierzu werden

[24] dabei kühlt jede Stufe die Abwärmeseite der vorangegangenen Stufe

die primären Elektronen auf eine *Mikrokanalplatte* abgebildet. Diese besteht aus sehr feinen beschichteten Kanälen. Bei jedem Auftreffen eines Elektrons auf eine Wand werden mehrere Sekundärelektronen heraus geschlagen, so dass sich insgesamt nach den wenigen Zehntel Millimetern Länge einer Mikrokanalplatte eine mehrere tausendfache Gesamtvervielfachung ergibt. Die Sekundärelektronen treffen dann auf eine Phophorschicht bei der letztlich wieder Photonen generiert werden. Diese werden dann (meist mittels Glasfasern) auf eine konventionelle CCD gebracht. Das ursprüngliche Bild wird also über den Umweg von Elektronen verstärkt. Damit ergeben sich Empfindlichkeiten bis in den Bereich von 0,001 lux (mondlose Nacht). Die Verstärkung kann über die an der Mikrokanalplatte angelegte Spannung auf einfache Weise gesteuert werden. Entsprechend kann so über die Spannung auch einfach ein sehr schneller elektronischer Shutter realisiert werden. Allerdings ist mit Übersprechen zwischen den einzelnen Kanälen zu rechnen und die Kosten für entsprechende Sensoren sind hoch.

Bei *Electron-Multiplying CCDs* (EMCCD) wird die Elektronenvervielfachung am Ende des Schieberegisters der CCD, also vor der eigentlichen Ausleseeinheit, eingebracht. Dementsprechend wird auch das *Dunkelstromrauschen* mit verstärkt. Eine Kühlung ist also nahezu immer notwendig. Die Quanteneffizienz der Sensoren kann sehr hoch (größer 90%) sein, das SNR ist aber dennoch nicht zwangsläufig besser oder schlechter als bei ICCDs denn der Auslesevorgang ist eben deutlich anders. Die Vervielfältigung führt auch zu einer Verringerung des Signal-Rausch-Verhältnis.

Für entsprechend verstärkende Sensoren ist die Schrotrauschgrenze wieder durch die Anzahl von Elektronen bedingt und diese lässt sich entsprechend

$$N_e = G \cdot F \sqrt{\eta \Phi_p \tau} \tag{9.16}$$

bei der Belichtungszeit τ berechnen. G ist der Verstärkungsfaktor (gain) und F bezeichnet einen Rauschfaktor, der angibt, wie stark das Rauschen durch die Verstärkung ansteigt (typisch $1.3 < F < 2$, siehe [10]).

Aufgrund der höheren Quanteneffizienz können teilweise back-illuminated CCDs bei geringen Lichtverhältnissen ein besseres Signal-Rausch-Verhältnis als EMCCDs oder ICCDs erreichen. Eine genaue Analyse der technischen Daten der Sensoren ist bei entsprechenden Kameras unerlässlich (insbesondere in Anbetracht des hohen Preises).

Inzwischen sind auch erste, echte *Single-Pixel Kameras* verfügbar, die es erlauben, räumlich aufgelöst, einzelne Photonen zu zählen. Hierbei verlassen wir aber definitiv unsere bisherigen Anwendungsbereiche, die durch das Photonenrauschen dominiert werden und landen in der Quantenoptik.

9.8 Bereiche und Binning

Oftmals ist es von großem Vorteil (geringere Datenmenge, höhere Geschwindigkeit), nur bestimmte Bereiche des belichteten Bildes auslesen zu können, z.B. Bereiche, in denen relevante Objekte liegen. Man spricht von *Regions of Interest* (ROI) oder *Area of Interest*. Für CMOS Sensoren ist es aufgrund der beliebigen Adressierung der einzelnen Pixel in der Regel unproblematisch, entsprechende Funktionalitäten auf dem Bildsensor zu integrieren. Teilweise sind auch komplexere Bereiche (mehrere Rechtecke) verfügbar.

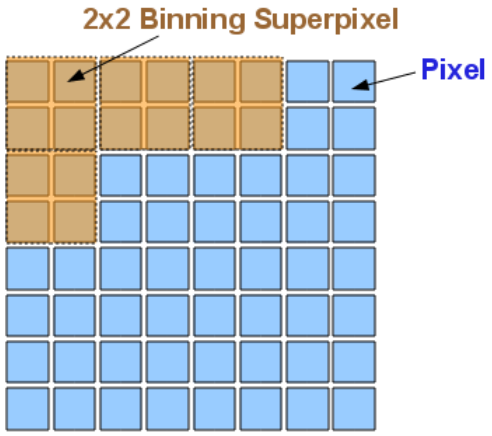


Abbildung 9.8.13: Beim 2x2 Binning werden jeweils vier Pixel zu einem Superpixel (4-fache Sättigungskapazität und 4-fache lichtempfindliche Fläche) zusammengefasst.

Bei CCD Sensoren ist eine zufällige Adressierung schwierig. Vergleichsweise einfach ist die Beschränkung auf bestimmte Zeilen, denn es werden dann lediglich die entsprechenden Zeilen durch das horizontale Schieberegister zum AD-Konverter geleitet und so weiterverarbeitet. Meist wird die Funktion als *partial scan* bezeichnet. Dementsprechend ist bei vielen Sensoren eine entsprechende Funktionalität implementiert. Komplexere Geometrien, z.B. rechteckige Bereiche,

sind aber nicht möglich bzw. führen kaum zu einer Beschleunigung.

Eine andere Variante beim Auslesen ist das *Binning*. Hierbei werden die Signale mehrerer, z.B. vier, benachbarter Pixel addiert. Dadurch wird zwar die Auflösung des Sensors effektiv verringert, aber Empfindlichkeit und Signal-Rausch-Verhältnis sowie die Auslesegeschwindigkeit steigen. Wenn 4 Pixel addiert werden sollte man für unkorreliertes Rauschen eine Verbesserung um den Faktor $\sqrt{4} = 2$ erwarten. Dies kann auch im PC per Postprocessing realisiert werden. Erfolgt die Integration bereits auf dem Sensor kann der Gewinn sogar noch etwas höher ausfallen da in diesem Fall das Ausleserauschen nur ein- und nicht viermal zustande kommt.[25] Zu beachten ist aber, mit welcher Dynamik das Auslesesystem ausgelegt wird. Wenn 4 Pixel zusammengefasst werden sollen, dann ergibt sich der volle Vorteil z.B. nur, wenn die Kapazität des Auslesesystems mindestens das Vierfache der Sättigungskapazität des Pixels entspricht oder eben mit eher geringen Signalen gearbeitet werden muss.

Von *Hardware-Binning* spricht man, wenn die Binning Funktionalität bereits in der Kamera durchgeführt wird. Nur dort lässt sich durch das Binning das Ausleserauschen bei CCD Sensoren optimal verringern.

9.9 Advanced: Triggerung

Die Bildaufnahme kann sowohl „freilaufend" als auch durch externe Ereignisse ausgelöst („getriggert") werden. Wenn z.B. Teile auf einem Fließband erfasst werden sollen ist es sinnvoll, die Kamera mit der Position des Teils (z.B. durch eine Lichtschranke) zu synchronisieren.

Im sogenannten *Edge Pre-select* Modus liefert die Flanke des *Triggerimpulses* das Signal für den Start der Bildaufnahme und die Dauer der Belichtung ist voreingestellt. Dies ist der Normalfall.

Der *Puls-Width Control* Modus regelt die Belichtungsdauer über die Breite des Triggerpulses.

[25] Für CMOS Sensoren ist der Vorteil nicht ganz so ausgeprägt denn hierbei werden die Ladungen nicht direkt vor der Elektronik additiert sondern es erfolgt erst die Konversion in Spannungen.

Bei der Verwendung mehrerer Kameras, die gleichzeitig ausgelöst wurden kann es notwendig werden, die Auslesezeitpunkte der Kameras unterschiedlich zu wählen bzw. nochmals extra zu triggern. Dies wird möglich, wenn eine entsprechende Zwischenspeicherung der Daten kameraseitig erfolgt. Man spricht von *Frame Delay Readout*.

Weiterhin ist zu bedenken, dass die maximale Bildfrequenz bei getriggertem Betrieb im Vergleich zum sogenannten „Freilaufbetrieb" oft verringert ist.

9.10 Advanced: Sensortransferfunktion

Wie haben im Kapitel 4 bereits darauf hingewiesen, dass nicht nur die Optik sondern letztlich alle verschiebungsinvarianten linearen Systemteile durch eine Transferfunktion beschrieben werden. Das Gesamtverhalten ist dann über die Multiplikation der einzelnen Transferfunktionen gegeben.

Für den Bildsensor ergibt sich die Transferfunktion, die sogenannte *Sensortransferfunktion*, ebenfalls aus der Berücksichtigung verschiedener Effekte. Entscheidend sind

1. Mikrolinsen (bei den meisten Sensoren)
2. *polarisationsoptischer* Tiefpass (bei vielen Farbsensoren)
3. Pixelgeometrie
4. ortsabhängige Variation der Pixelempfindlichkeit
5. Übersprechen zu benachbarten Pixeln

Wieder gilt, dass alle diese passiven Effekte zu einer Modulationstransferfunktion kleiner 1 führen also den Kontrast letztlich verringern.

Der Effekt der Mikrolinsen wird meist vernachlässigt. Im Sinne einer Transferfunktion ist die Mikrolinse meist vernachlässigbar denn ihre Geometrie ist passend zur Geometrie des Pixels gewählt und wird durch diesen Beitrag ausreichend gut beschrieben. Wesentlicher ist, dass das Mikrolinsenverhalten stark

winkelabhängig ist. Mit anderen Worten: Es ergibt sich für ausreichend starke Winkel ein starker verschiebungsvarianter Anteil, der durch eine einzelne Transferfunktion eben nicht beschrieben werden kann. Dementsprechend muss man für verschiedene Feldpositionen (diese entsprechend verschiedenen Winkel des Hauptstrahl (außer bei bildseitig telezentrischen Designs)) verschiedene Abschwächungen annehmen. Dies bezieht sich aber mehr auf die generelle Helligkeit in diesem Bereich als auf den Kontrast, der bei der Aufnahme von feinen und mittleren Strukturen erzielt werden kann. Ein Verhalten, das also eher durch den Begriff „Vignettierung" (vgl. Abschnitt 2.3.12) als durch eine MTF zu beschreiben ist. Problematisch sind die Effekte auch ganz besonders bei Farbsensoren.

Bei Sensoren, die für die Verwendung mit einem ganz speziellen (und nicht bildseitig telezentrischen) Objektiv optimiert sind (z.B. Raspberry Pi-Camera) werden die Mikrolinsen so lokal verschoben, dass die Vignettierungseffekte minimiert werden. Werden solche Sensoren mit einem anderen Objektiv betrieben ist unter Umständen mit deutlichen Störeffekten zu rechnen.

Und schließlich muss noch beachtet werden, dass die Mikrolinsen die maximal mögliche numerische Apertur beschränken können. Letztlich nimmt die Empfindlichkeit des Pixels mit zunehmendem Einfallswinkel (auf die Mikrolinse) ab. Hierfür gibt der Hersteller Transferkurven an. Damit ist klar, dass große Winkel (entsprechend einer hohen numerischen Apertur bzw. kleinen Blendenzahl) dann stark abgeschwächt werden.

Hinsichtlich der *Pixelgeometrie* ist entscheidend, dass Licht, das auf einen rechteckigen Pixel der Breite b fällt letztlich über einen rechteckigen Bereich integriert wird, d.h. im Bildraum ist die Punktbildfunktion in erster Näherung eine rect-Funktion[26].

Das bedeutet nach Kapitel 4, dass die Transferfunktion ein sinc wird:

$$H_P \sim \mathrm{sinc}(u \cdot b) \tag{9.17}$$

Die lokale Lichtempfindlichkeit über einen Pixel ist allerdings in der Praxis nicht konstant sondern aufgrund der doch komplexen Geometrie ortsabhängig [27].

[26] An sich ist der Vorgang dieser Integration nichtlinear und eine Angabe einer Systemfunktion bzw. einer „Punktbildfunktion" ist problematisch. Nichtsdestotrotz genügt das einfache lineare Modell für eine gute Abschätzung des Sensorverhaltens.

Wenn man zusätzlich berücksichtigt, dass die Elektronen im Pixel diffundieren ergibt sich eine weitere Verschmierung[27], die zu einer leichten, aber stark wellenlängenabhängigen Verringerung der Gesamttransferfunktion des Sensors führt.

In vertikaler Richtung, also in Richtung des Ladungstransfers ergibt sich ebenfalls eine Verschmierung aufgrund des nicht perfekten Ladungstransfers.

Die MTF des Sensors hängt auch von der Wellenlänge ab denn längere Wellenlängen dringen tiefer ein und die sich ergebenden Ladungsträger diffundieren daher mehr. Daher wird die MTF bei großen Wellenlängen verschlechtert.

Auch eine reduzierte Ladungsverschiebung bei der CCD kann zu einer Verschmierung und damit einer Verringerung der MTF führen.

Um Aliasing bei der Abstastung hochfrequenter Details zu vermeiden (vgl. Abschnitt 3.6 und Abschnitt 5.7) werden bei Farbsensoren mit Bayermaske (oder ähnlicher Farbmaske) oft zusätzliche Tiefpassfilter direkt vor dem Sensor verwendet. Das Ziel ist, die lokale Bildinformation so zu verteilen, dass auf die jeweils benachbarten roten, grünen und blauen Pixel dieselbe Beleuchtungsstärke fällt.

Eine Standardvariante zur Erfüllung dieses Ziels ist der Einsatz eines polarisationsoptischen Tiefpasses: Ein auf einen doppelbrechenden Kristall einfallender unpolarisierter Strahl wird zunächst in zwei lateral versetzte Kopien aufgespalten[28]. Ein zweiter doppelbrechender Kristall führt dann zu einer erneuten Aufspaltung jeder der beiden Strahlen, so dass insgesamt vier räumlich versetzte Strahlen auftreten.

In jede Richtung ist das Punktbild also die Summe von zwei (versetzten) Delta-Peaks Dementsprechend ist die Transferfunktion ein Kosinus:

$$H_{PTP}(u, v) = \cos(2u/b) \tag{9.18}$$

[27] z.B. berechenbar mit dem Modell von Sieb, [12]

[28] Wie es zu dieser als „Doppelbrechung" genannten Erscheinung kommt ist nicht ganz einfach und soll uns hier nicht weiter beschäftigen (siehe z.B. [31])

9.11 Farbsensoren

Die meisten Bildverarbeitungsanwendungen sind nach wie vor monochrom. Farbe kann genutzt werden, um Erkennungsaufgaben zu erleichtern, z.B. die Erkennung einer Seriennummer oder aber auch die Unterscheidung von verschiedenen Materialen. Teilweise ist die Farbe aber auch schlicht Teil der zu erfüllenden Aufgabe (z.B. Farbkontrolle). Wenn keine Farbfunktionalität notwendig oder hilfreich ist, macht es in der Regel auch keinen Sinn, einen Farbsensor einzusetzen denn Farbe führt in jedem Fall bei ansonsten gleichen Parametern zu erhöhten Anforderungen (z.B. Reduzierung chromatischer Aberrationen bei den Objektiven, erhöhten Datenaufkommen $\rightarrow$ Datentransfer, Speicherung und Verarbeitung, geringere Lichtnutzung).

Generell ist die Vorgehensweise bei der Farbbildaufnahme eine einfache Erweiterung der monochromatischen Aufnahme auf mehrere Kanäle. Die Abbildungsoptik wird so ausgelegt, dass alle relevanten Spektralbereiche in ausreichender Bildqualität übertragen werden und der Bildsensor nutzt eine wie auch immer geartete spektrale Diskriminierung. Grundsätzlich stehen dabei folgende Varianten zur Auswahl:

Farbmasken: Vor den eigentlichen lichtempfindlichen Pixeln werden Farbmasken (meist mit drei unterschiedlichen Spektralkanälen) verwendet, vgl. Abb. 3.6.19.

Zeitliches Multiplexing: Es wird entweder ein zeitlich wechselnder Filter vor dem Objektiv der Kamera verwendet oder aber die Szene wird zu unterschiedlichen Zeiten mit unterschiedlichen Spektren beleuchtet (meist getaktete LED Quellen). Zeitsequentiell werden somit Bilder mit unterschiedlicher spektraler Information aufgenommen. Natürlich macht das nur Sinn, wenn sich die Szene während den unterschiedlichen Aufnahmen nicht ändert.

Mehrere Kameras/Bildsensoren: Es werden verschiedene Bildsensoren so in einer Kamera kombiniert (sogenannte „3-Chip Kameras"), dass das spektral aufgespaltene Licht auf die Sensoren (in der Regel: drei) fällt. Dadurch ergibt sich für jeden Spektralkanal direkt ein eigenständiges Bild.

Dispersive Aufspaltung: Bei abtastender Bildaufnahme kann durch ein dispersives Element im Abbildungsstrahlengang eine eindimensionale Bildinformation auf ein zweidimensionales Signal transformiert werden. Die zweite Dimension dient somit zur Erfassung des Spektrum. Es werden also nicht nur wenige Spektralkanäle sondern sehr viele (z.B. 50 ... 100) Spektralbänder gleichzeitig erfasst. Natürlich wird dies letztlich durch Einschränkung der räumlichen oder zeitlichen Auflösung erreicht.

Für jeden Kanal gilt aber, dass die Antwort des photosensitiven Elements (z.B. ein Pixel) wieder durch die spektrale Verteilung des Objektpunkts gewichtet mit der spektralen Empfindlichkeit des Spektralelements berechnet wird. D.h. es wird ganz analog zu Gleichung 7.4 wieder das Signal berechnet. Lediglich die spektrale Empfindlichkeit ist dann eben für jeden Kanal anders zu wählen.

Mit Abstand die gebräuchlichste Variante ist dabei die Farbfiltermaske direkt vor den Pixeln des Bildaufnehmers. Abb. 9.11.14 zeigt das meist eingesetzte Bayer-Pattern bei dem — entsprechend der Empfindlichkeit des menschlichen Sehens — doppelt so viele grüne wie rote oder blaue Pixel verwendet werden. Die beste laterale Auflösung und Lichtnutzung eines solchen Sensors erhält man also im grünen Spektralbereich.

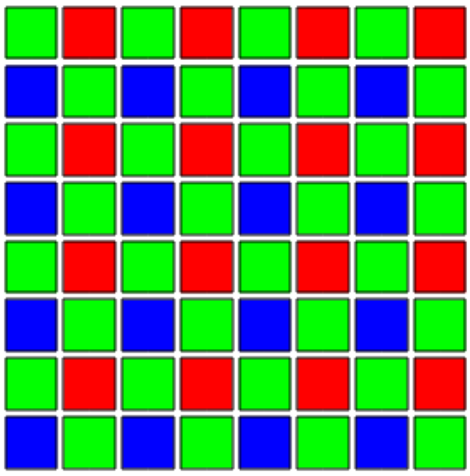


Abbildung 9.11.14: Bayer Maske zur Farbdetektion

Klar ist, dass in jedem Fall die laterale Auflösung reduziert wird, denn die vorhandenen Pixel werden für unterschiedliche Farben verwendet. Pro Spektralkanal stehen also weniger Pixel als im monochromen Fall zur Verfügung. Wenn eine

Farbkamera mit Farbmaske (bzw. der Treiber bzw. der Framegrabbers) dennoch für jeden Pixel RGB Werte ausgibt, d.h. jedem Pixel drei Werte zuordnet (obwohl natürlich nur ein Wert gemessen werden kann) dann wurde zwangsweise die jeweils fehlende Information durch Interpolation bestimmt. Verschiedene solche *Farbinterpolationsalgorihtmen* werden eingesetzt bzw. müssen ggf. noch selbst implementiert werden.

Eine klassische Lösung ist der Einsatz von drei Bildaufnehmern für drei Spektralkanäle. Dazu wird das optische Bildsignal zunächst spektral aufgespalten bevor es von den drei getrennten Sensoren erfasst wird. Problematisch ist — insbesondere bei hohen Auflösungen — die exakte Justage der Sensoren zueinander. Daher wird diese Lösung (aufgrund der immer kleiner werdenden Pixel) zunehmend seltener eingesetzt. Vorteilhaft ist aber, dass kein Lichtverlust (im Vergleich zur Absorption in den Filtern der Bayermaske) auftritt.

Durchsetzen konnte sich auch nicht die Foveon Lösung, bei der ein Pixel letztlich aus drei getrennten Pixeln in der Tiefe besteht, d.h. an derselben lateralen Position können wirklich unterschiedliche Farben gemessen werden. Langwellige Photonen (rot) dringen tiefer in den Pixel ein. Blaue Photonen kommen am wenigsten tief. So ist also über die Tiefe eine Farbtrennung möglich. Der an sich elegante Ansatz leidet aber unter deutlicher Beeinträchtigung der Qualität.

Generell ist die typischerweise eingesetzte Wahl von drei Spektralkanälen bei Rot, Grün und Blau, natürlich technisch nur dann sinnvoll, wenn die Anwendung in irgendeiner Form auf das menschliche Sehen bezogen ist (z.B. Kontrolle von Druckbild). Für reine Identifikationsaufgaben können auch andere Spektralkanäle sinnvoll sein.

Aber auch für klassische Farbanwendungen sind verschiedene Varianten des Filters möglich. Z.B. kann man die grünen Pixel durch gelbe Pixel ersetzen und damit dann noch etwas (in diesem gelb-grünen Kanal) an Lichtstärke gewinnen. Das eigentliche Grünsignal wird dann berechnet, indem vom gelben Pixelwert der benachbarte rote Pixelwert subtrahiert wird. Bei gleicher Belichtungszeit ist dieser aber (wegen der verringerten Lichttransmission) verrauschter und so schlägt auch dieses Rauschen im Farbbereich dann wieder auf das Grünsignal durch. Letztich ist der Gewinn an effektiver Lichtstärke also eher gering. Bei Automotive Sensoren werden teilweise auch Bereiche ohne Filter zur verbesserten Lichtausnutzung

verwendet. So besteht z.B. ein RGCC-Sensor aus 25% roten, 25% grünen und 50% Pixeln ohne Filter.

Anstatt einem Gewinn an Lichtstärke kann man auch versuchen eine verbesserte Farbtreue zu erzielen. Hierzu können dann z.B. vier statt drei Farbfilter einesetzt werden.

Am gebräuchlichsten und qualitativ besten ist meist die Verwendung von gepulsten Lichtquellen unterschiedlicher Wellenlänge (gepulste LEDs) in Verbindung mit einer Mehrfachaufnahme. Dadurch wird sichergestellt, dass die Bildaufnahme der unterschiedlichen Spektralkanäle völlig pixelsynchron erfolgt. Natürlich müssen Mehrfachaufnahme und Pulsung bei bewegten Objekten schnell genug sein, um Störungen zu verhindern.

Bei der Kontrolle von Farbabweichungen bzw. Farbänderungen geht es meist um eine vom Menschen unter Umständen detektierbare Farbänderung.[29]

Letztlich gilt es also zu detektieren, ob eine Änderung der spektralen Charakteristik eines Objekts zu einer für den Mensch wahrnehmbaren Änderung führt. Dabei ist zu beachten, dass die wahrgenommene Farbe natürlich nicht nur vom Objekt selbst sondern auch von der Beleuchtung abhängt (*Metamerie*). So kann weißes Licht in unterschiedlicher spektraler Zusammensetzung entsprechend auch unterschiedliche Farbwahrnehmungen bewirken.

9.11.1 Advanced: Hyperspektral- und Multispektral-Sensorik

Hyperspektralsensoren können als ein Sonderfall von Farbsensoren aufgefasst werden. Bei diesen wird für jeden Objektpunkt ein Spektrum von vielen Spektralkanälen aufgenommen. Zum Einsatz kommen können im Prinzip die bereits bekannten Methodiken (insbesondere zeitliches Pulsen der Beleuchtung und Farbmosaikmasken),

Aufgrund den inzwischen einfach verfügbaren Sensoren mit hoher Auflösung werden zunehmend immer stärker auf Mosaikfiltern basierende Sensoren eingesetzt. Diese werden meist *Hyperspektral Snapshot Imager* genannt, da das gesamte

[29] Streng genommen ist auch der Begriff „Farbe" ein Begriff, der sich rein auf die Wahrnehmung des Menschen bezieht. Im Slang der Bildverarbeitung wird dies aber selten so eng gesehen.

Spektrum (→ Hyperspektral) eben ortsaufgelöst (→ Imager) während einer Belichtung (→ Snapshot) erfasst wird.

Dabei geht durch die Farbfilter natürlich viel Licht verloren und es muss weiterhin ein Trade-off zwischen räumlicher und spektraler Auflösung eingegangen werden.

Sogenannte *Push-Broom* Sensoren erfassen zu einem Zeitpunkt lediglich die Spektren einer Zeile, diese dadurch dann aber spektral und örtlich hochaufgelöst und ohne Lichtverlust. Hierzu wird einfach ein dispersiven Element (Gitter oder Prisma) im Strahlengang verwendet.

Zunächst stellt sich die Frage, welche Daten überhaupt benötigt werden. Ein klassisches *hyperspektrales Imaging* führt bei hoher Kanalzahl und hoher lateraler Auflösung zu einer großen Informationsmenge pro Bild und damit unter Umständen zu Problemen bei Datentransfer, Datenspeicherung oder Echtzeitverarbeitung.

In den allermeisten Fällen wird in der industriellen Anwendung daher eine Hyperspektralkamera nur zur einmaligen Identifikation der spektralen Information, also zur Beantwortung der Frage, welche spektralen Bereiche besonders wichtig sind, verwendet. Nachfolgend kann dann mit geringer Kanalzahl die eigentliche Inspektionsaufgabe durchgeführt werden.

Die (konzeptionell) einfachste Variante zur Erweiterung der Kanalzahl ist sicher, die üblicherweise eingesetzten Farbfilterarrays der Farbkamera (Bayermaske, vgl. Abschnitt 9.11.14) auf mehr Kanäle zu erweitern, also z.B. 5x5 Filter zu verwenden. Dadurch reduziert sich die Auflösung des Bildsensors natürlich in beide Richtungen um den Faktor 5; aber Pixel sind heute vergleichsweise preiswert und dafür ergeben sich eben z.B. 25 Spektralkanäle. Für vergleichsweise wenig Kanäle, z.B. 10, wird teilweise auch der Begriff *Multispektral-Imaging* verwendet.

Kameras mit spektralen Filterarrays sind (noch) vergleichsweise selten (und teuer) und da die Filter keine idealen Bandpasstransmissionen haben ist eine Nachbearbeitung im Computer (spektrale Korrektur) notwendig. Hinsichtlich der Lichtintensität wird natürlich ebenfalls eine Menge an Licht eliminiert, so dass das Signal-Rausch Verhältnis im Vergleich zu anderen Lösungen geringer ist. Ein großer Vorteil ist dagegen, dass die spektrale Information wirklich gleichzeitig für alle Spek-

tralbänder erfasst wird.

Dies ist bei dem nächsten konzeptionell einfachen Ansatz nicht der Fall. Nämlich der zeitlich modulierten Beleuchtung (z.B. mittels unterschiedlicher monochromatischer LEDs) oder aber einem zeitlich variablen Filter kameraseitig. Der Filter kann dabei entweder rein mechanisch (Farbrad) arbeiten[30] oder aber elektronisch schaltbar (z.B. mittels Flüssigkristall-basierten Filtern) sein. Vorteil hierbei ist, dass die hohe Auflösung des Bildsensors erhalten bleibt. Die spektrale Information wird durch Einschränkung bei der Zeitauflösung gewonnen.

Eine optimale Lichtnutzung erhält man durch Ansätze, bei denen Filter vermieden werden. In diesem Fall muss das Licht spektral (durch Prismen oder Gitter) aufgespalten werden. Mit einem zweidimensionalen Bildaufnehmer lässt sich so das Spektrum einer Zeile des Objekts bestimmen. Die Methodik ist also zunächst gut für Teile auf Fließbändern oder auch rotierte Teile anwendbar. Andererseits ist für viele Anwendungen dann die Geschwindigkeit (2D Bildaufnehmer statt Zeilenkamera) nicht ausreichend. Hauptvorteil ist aber, dass die spektrale Auflösung sehr hoch sein kann und das Licht optimal genutzt wird.

9.12 Advanced: Infrarotsensoren

Die Nutzung von längerwelligem Licht außerhalb des durch das Silizium limitierten Spektralbereichs (bis ca. 1050 nm) ist für verschiedene Anwendungen von Interesse. Verschiedene Materialien bzw. Stoffe verhalten sich in nicht sichtbaren Spektralbereichen deutlich verändert und dementsprechend kann eine Identifikation unter Umständen deutlich einfacher werden. Interessant ist auch, dass verschiedene Materialien im short-wave Infrared (SWIR) Bereich transparent werden. So lassen sich dann z.B. Füllstände in Gefäßen von außen beurteilen wenn das Gefäßmaterial transparent für die Strahlung ist (Wasser absorbiert z.B. besonders gut bei 1450 nm). Ein weiteres Beispiel ist die Inspektion von Strukturen durch Siliziumabdeckungen, wie sie in der Mikrosystemtechnik Verwendung finden.

Je länger die Wellenlänge ist, desto mehr profitiert man bei entsprechenden Sensoren von einer Kühlung. Thermische Elektronen sind solche, die — aufgrund der

[30] In diesem Fall ist natürlich eine Synchronisation mit dem Bildsensor notwendig.

Temperatur — das Valenzband verlassen und sich dann frei im Material bewegen können. Es ergibt sich ein Dunkelstrom (und damit einhergehend auch ein Dunkelstromrauschen). Eine Kühlung verringert die Anzahl dieser Elektronen stark. Die freien Elektronen entstehen dann wirklich dann und nur dann, wenn Licht einfällt und die Anzahl der Elektronen ist dann eben proportional zur einfallenden Lichtmenge (Photonenanzahl).

Längere Wellenlängen im mittleren Infrarotbereich sind aus den üblichen „Wärmebildaufnahmen" (*Thermografie*) bekannt. Hier kann sowohl mit einer thermischen Lichtquelle in Transmission oder Reflexion gearbeitet werden oder aber die Eigenemission des Körpers (im Sinne eines Plankschen Strahlers) wird genutzt. Bei aktiver Erwärmung kann so auch der Wärmefluss aufgezeichnet werden und z.B. damit auf Defekte im Inneren von Körpern rückgeschlossen werden. Selbst geringe Temperaturunterschiede im Millikelvin Bereich lassen sich detektieren.

Letztlich gibt es verschiedene Messverfahren, die eine Detektion von infraroter Strahlung notwendig machen. Einen guten Überblick gibt Bauer [4]. Leider sind die entsprechenden Bildsensoren vergleichsweise teuer und zwangsläufig gegenüber thermischem Rauschen empfindlich. Ganz generell ist oftmals das Rauschverhalten entsprechender Sensoren problematisch. Teilweise ist auch mit einer erheblichen Anzahl „toter Pixel" zu rechnen. Durch eine aktive Kühlung lässt sich das thermische Rauschen der Detektoren sehr stark beeinflussen, so dass für rauscharme Anwendungen meist gekühlte Detektoren verwendet werden.

Von *Quantendetektoren* spricht man wenn der eigentliche Detektionsprozess letztlich wie beim CMOS oder CCD Bildsensor funktioniert, d.h. Lichtquanten führen über den inneren Photoeffekt zu Ladungsträgern. Für Silizium ist man hier aber eben auf Wellenlängen unterhalb 1100 nm begrenzt. Mit *Indium-Gallium-Arsenid* (InGaAs) Sensoren lassen sich Bereiche bis 1700 nm abdecken. Längere Wellenlängen werden mit Platinsilizid Sensoren oder Indiumantimonid Sensoren abgedeckt.

Thermische Detektoren reagieren dagegen nicht auf einzelne Photonen sondern hier wird die insgesamt aufgenommene Energie (in Form von Wärme) detektiert. Bei einem *Bolometer* führt der Wärmeeintrag so z.B. zu einer Widerstandsänderung. Bevor ein neues Signal detektiert werden kann muss aber die Wärme abtranspor-

tiert werden. Die Sensoren sind also langsamer (Größenordnung einige 10 ms) und auch meist weniger empfindlich, dafür aber preiswerter. Weiterhin führt der Wärmefluss zu einer effektiven Verringerung der Auflösung.

Insbesondere im SWIR können aber teilweise erhebliche Unterschiede der Absorption verschiedener Stoffe eine Erkennung erleichtern. Wasser wird z.B. sehr stark in diesem Spektralbereich absorbiert. Der ferne Infrarotbereich wird sensorseitig meist durch sogenannte *Mikrobolometer* abgedeckt und wird vor allem zur *Temperaturbestimmung* genutzt. Die Anwendungen sind vielfältig und reichen von der Detektion von Fieber bis zu laser-gesteuerten Bomben. Überall, wo sich Objekte durch eine geänderte Temperatur zum Hintergrund auszeichnen sind sie einsetzbar.

Eine wichtige Klasse von Anwendungen ist die zerstörungsfreie Prüfung mit Defekten unterhalb der Oberfläche. Solche sind normalerweise mittels optischer Methoden ja zunächst nicht sichtbar. Man kann aber nun „eine Frage an das Objekt stellen", z.B. eine pulsartige Erwärmung mittels Infrarotlicht durchführen, und dann messen, wie die Oberfläche reagiert. Diese Reaktion kann z.B. in einer sehr feinen (Mikrometer) Verschiebung der Oberfläche bestehen. Entsprechend werden dann eben hochgenaue 3D-Messverfahren eingesetzt. Aber auch die Oberflächentemperatur liefert wichtige Ergebnisse. Oftmals ergibt sich oberhalb eines Defekts eben eine andere Temperaturabstrahlung, die mit einer Infrarotkamera dann detektiert werden kann. Auch Leckagen (z.B. bei Getränkedosen oder Gaspipelines) sind sehr gut über die geänderte Temperatur an der Austrittsstelle detektierbar.

Beim Einsatz solcher Sensoren ist in jedem Fall auch auf eine an den Spektralbereich angepassten Optik (Gläser, Optikdesign und Beschichtung) zu achten. Ansonsten sind Aberrationen, Streulicht und Transmission sehr schlecht.

Eine gute Einführung in die verschiedenen *Infrarotsensoren* und ihre Leistungsdaten bzw. Einsatzmöglichkeiten findet man bei [4]. Für Sensoren im langwelligen Infrarotbereich kann [**?**] nützlich sein.

Für quantitative Temperaturmessungen geht man davon aus, dass die Objekte gemäß dem Planckschen Strahlungsgesetz (Abb. 8.4.13) abstrahlen. Allerdings ist dabei die Stärke der Emission materialabhängig. Von opake (nicht transparente)

Materialien ist die Summe aus Emissivität ϵ und Absorption gerade 1. Die Emissivität ist das Verhältnis aus abgestrahlter Leistung im Vergleich zur abgestrahlten Leistung eines ideal schwarzen Körpers:

$$\epsilon(\lambda) = \frac{W_{Objekt}}{W_{SchwarzerKoerper}} \tag{9.19}$$

Aus der Planckschen Strahlungsformel lässt sich (durch Integration) berechnen, wieviel Leistung nun ein Körper abstrahlt:

$$P = \epsilon \cdot \sigma \cdot T^4 \tag{9.20}$$

Dabei bezeichnet σ die Stefan Boltzmann Konstante ($\sigma = 5,67 \cdot 10^{-8}$ W/m^2/K^4).

Man kann somit über eine Leistungsmessung auf die Temperatur rückschließen wenn man die Emissivität kennt bzw. kalibriert.[31]

Wichtig kann auch die Frage sein, bei welcher Wellenlänge ein Objekt am meisten Strahlung emittiert. Dies kann mittels

$$\lambda T = 2989 \mu m\, K \tag{9.21}$$

abgeschätzt werden.

Wenn es auf die maximale Anzahl an Photonen (im Gegensatz zur Leistung in Watt) ankommt, dann wird die Konstante zu 3670 μm K (siehe [35]).

9.13 Advanced: Zeilensensoren

Wir sind bisher implizit immer von Flächensensoren (*area sensors*), also typischen Kameras, wie man sie üblicherweise kennt, ausgegangen. Manchmal kann es aber auch sinnvoll sein, einen eindimensionalen Detektor, also eine Zeile, einzusetzen.

Zeilensensoren (engl. *line scan sensors*) sind vor allem dann vorteilhaft, wenn in der konkreten Anwendung eine Linearbewegung ohnehin vorhanden ist (Teil auf

[31] Unter Umständen muss man auch das Rückstreuverhalten des Umgebungslichts mit einbeziehen, denn das von Objekten der Umgebung emittierte Licht wird am Objekt ja auch gestreut und ist dem Eigenleuchten überlagert.

Fließband). In diesem Fall können Beleuchtungs- und Abbildungsoptik hinsichtlich der (einfacheren) Abbildungsoptik optimiert werden. Wesentlich ist für eine korrekte Geometrieerfassung natürlich, dass die Linearbewegung ausreichend gleichförmig ist. Eine Verbesserung der Empfindlichkeit kann durch Zeilenkameras mit Time-Delay and Integration (TDI) Funktionalität erzielt werden. Hierbei liegen mehrere Zeilen lichtempfindlicher Sensoren (in Vorschubrichtung) hintereinander und eine exakte zeitliche Synchronisation von Vorschub und Taktfrequenz des Sensors erlaubt die Integration der vom Objekt ausgehenden Lichts auf verschiedenen Zeilen.

Zeilenkameras werden üblicherweise dann eingesetzt, wenn große, sich bewegende Objekte mit hoher Auflösung erfasst werden sollen. Das klassische Beispiel sind Bahnwaren (z.B. Papier oder Stoffbahnen[32]). In diesem Fall können durch die Bewegung viele Pixel zur Erfassung des Objekts Verwendung finden. Neben linearen Bewegungen sind auch Rotationsbewegungen typische Anwendungen, d.h. ein Zeilensensor erfasst eine Linie auf dem Objekt und das Objekt wird unter dem Sensor rotiert.

Damit die Bilderfassung ohne Verzerrungen auch bei nicht perfekt gleichförmiger Bewegung funktioniert müssen Bilderfassung und Beleuchtung basierend auf der Position (z.B. über Winkelencoder des Bandvorschubs) getriggert werden.

Darüber hinaus ist zu beachten, dass entsprechende Systeme oft vergleichsweise empfindlich hinsichtlich Vibrationen sind. Wenn die Kamera leicht vibriert ergibt sich natürlich eine massive Geometrieänderung. Aber auch die linienförmige Beleuchtung ist problematisch, denn auch diese muss stabil zur Kamera sein. Die Justage selbst ist nicht ganz so einfach wie bei einem zweidimensionalen Bildsensor, denn die Schärfe kann nur entlang einer Zeile (oder eben in Bewegung) beurteilt werden.

Entsprechende Zeilensensoren werden teilweise mit sehr vielen Pixeln (z.b. 8000 Pixel) realisiert. Dementsprechend ergibt sich ein großer Sensor und damit ein großes Bildfeld. Konventionelle C- oder CS-Mount Anschlüsse sind dann oft mechanisch nicht mehr einsetzbar und es muss zu F-Mount oder M72-Anschlüssen übergegangen werden.

[32] Außerhalb der industriellen Messtechnik sind Satellitenaufnahmen ein schönes Beispiel. Der Satellit erfasst fortlaufend während seiner Bewegung die Erdoberfläche.

Zeilenkamera-basierte Bilderfassung hat also Vorteile (insbesondere bei Bahnware), führt aber generell zu einer vergleichsweise hohen Komplexität bzw. möglichen praktischen Problemen. Daher sollten für konventionelle BV-Aufgaben ohne extreme Anforderungen soweit es möglich ist eher Flächensensoren zum Einsatz kommen.

Die Zeilenfrequenzen sind natürlich wesentlich höher wie die Bildwechselraten von Flächensensoren. Mehrere 10 KHz sind üblich, aber auch mehrere 100 kHz sind heute machbar. Weitere Details zu Zeilenkameras und ihrer Anwendung findet sich bei Demant et al. [8].

9.14 Advanced: Kameraschnittstellen

Natürlich besteht eine nahezu unbegrenzte Möglichkeit, eine Kamera mit einem externen System (meist ein PC) kommunizieren zu lassen. Teilweise verwenden Hersteller eigene Formate, in der Industrie haben sich dennoch mehrere Formate als Standards durchgesetzt.

9.14.1 Analoges Video

Der frühere Standard, die Daten analog als kontinuierliche Spannungen zu übertragen, wird heute kaum noch verwendet und sollte vermieden werden. Typische analoge Interfaces sind der *CCIR* Standard und der *EIA 170* bzw. *RS 170* Standard. Abgesehen von den geringen Auflösungen, die erzielt werden können ergeben sich natürlich zusätzliche Rausch- bzw. Fehlerquellen und PC-seitig ist eine Digitalisierung der Analogsignal notwendig. An sich wird nahezu der gesamte Prozess der Bildaufnahme hin zur Verarbeitung für den Anwender komplizierter. Der Hauptvorteil der *analogen Kameratechnik* ist, dass die Bildsignale direkt an einfache Videomonitore ausgegeben werden können. Aber auch dieser Vorteil verliert sich immer mehr, denn preiswerte Monitore haben heute oft keinen Videoeingang.

Beim CCIR Verfahren handelt es sich um ein sogenanntes *Interlaced*-Verfahren, d.h. das Bild wird in Form von zwei Halbbildern, zeilenweise verschachtelt

übertragen. Die Auflösung ist auf 767 x 575 Pixel bei einer Übertragungsfrequenz von ca. 15 MHz begrenzt. Zur eigentlichen Bilderfassung ist ein geeigneter Framegrabber notwendig. In der Regel wird man heute keine Systeme basierend auf analoger Videotechnik für die industrielle BV auslegen.

Insbesondere bei analogen Videokameras ergibt sich das Problem, das ein pixelsynchrones Abtasten oft problematisch ist. Natürlich muss auch bei digitalen Bildsensoren eine pixelsynchrone Abtastung erfolgen. Bei analogen Videosignalen ergibt sich — genau wie beim analogen Fernsehen — aber nur ein zeilenweises Synchronisationssignal. Wann der Framegrabber also innerhalb einer Zeile die einzelnen Pixel abtasten soll ist nicht exakt festgelegt und muss vom Framegrabber selbst ermittelt werden. Fehler sind unvermeidlich und können bis zu mehrere (horizontale) Pixel betragen. Selbst wenn der Fehler nicht direkt sichtbar ist oder an sich vergleichsweise gering ist (digitale Systeme) kann er für Anwendungen, die eine Genauigkeit im Subpixelbereich erfordern, problematisch werden. Teilweise liefern analoge Videokameras daher noch ein Pixelclock-Signal, das dann die horizontale Taktung der Pixelinformation liefert und von geeigneten Framegrabbern genutzt werden kann.

9.14.2 Digitale Schnittstellen

Das viele Jahre vorherrschende *FireWire* bzw. *IEEE1394b* Interface (64 MByte/s) wird heute nur noch selten eingesetzt. Aktuell sind vor allem fünf Standards verbreitet (siehe Tab. 9.14.15). Allerdings kann sich dieser Zustand in wenigen Jahren wieder komplett ändern, so dass wir an dieser Stelle nicht detailliert auf die einzelnen Varianten eingehen wollen (teilweise findet sich eine detaillierte Beschreibung in [8]). Letztlich ist neben den technischen Daten hierbei oft entscheidend, wie sehr sich ein Standard durchsetzt. Nur dann ist nämlich ausgereifte Hard- und vor allem Software bei geringen Kosten verfügbar.

	Camera Link	**Camera Link HS**	**CoaXPress**	**GigE**[33]	**USB3**[34]
Bandbreite[35]	bis zu 850 MB/s	bis zu 6.8 GB/s	6.3 GB/s	128 MB/s	350 MB/s
Framegrabber nötig	Ja	Ja	Ja	Nein	Nein
Stromvers. Kam.	Optional	Nein	Ja	Optional	Ja

Abbildung 9.14.15: Einige gebräuchliche Kameraschnittstellen, Stand 2021

Hinsichtlich der möglichen Kabellängen ist eine generelle Aussage schwierig, denn teilweise ist es möglich, mit Erweiterungen die Grundlängen deutlich zu erhöhen. Insbesondere bei Längen über 10 Metern sollte man sich aber Gedanken machen. Je höher die Kabellänge und je höher die Datenrate, desto eher sollten optische Verbindungen genutzt werden. Bei hohen Datenraten ist auch zu prüfen, ob teilweise Datenpackete verloren gehen; vor allem bei einem Nicht-Realttimefähigen Betriebssystem (z.B. Windows). Oft können zur Steigerung der Datentransferrate auch mehrere Kanäle eines Interfacestandards (z.B. Dual-10GigE) genutzt werden.

Der Hauptvorteil von *Gigabit Ethernet*(*GigE*) und *USB-3* ist sicherlich, dass kein zusätzlicher Framegrabber zur Datenerfassung notwendig ist.[36] Dadurch werden die Kosten der Gesamthardware zur Bilderfassung deutlich reduziert. Allerdings ist die maximale Geschwindigkeit entsprechend ebenfalls reduziert. Bedenken Sie aber auch, dass die mit hoher Geschwindigkeit PC-seitig ankommenden Daten verarbeitet werden müssen. Ein weiterer Vorteil ist die starke Verbreitung beider Schnittstellen im PC-Bereich, die eine langfristige Verfügbarkeit garantiert.

Ein insbesondere im professionellen Umfeld sehr gebräuchlicher Standard ist das *CameraLink* Interface. Die Bandbreite ist mit 850 MBytes/s hoch, das Timing ist klar definiert. Da dieser Standard vergleichsweise alt ist, sind professionelle Framegrabber und Kameras in großer Zahl am Markt vorhanden. Eine schnellere Version mit (inzwischen) bis zu 6800 MByte/s, *CameraLink HS*, ist verfügbar und konkurriert mit dem *CoaXPress* Standard (6300 MByte/s).

Ein Nachteil des sehr gebräuchlichen CameraLink Standards ist bzw. war es, dass Framegrabber-spezifisch die Ansteuerung der Kamera programmiert werden muss. Dies kann trotz der vom Hersteller üblicherweise zur Verfügung gestellten SDKs und Treiber einen nicht geringen Aufwand bedeuten und es kann zu Problemen bei bestimmten Kombinationen von Kameras und Framegrabber führen. Soll die Ansteuerung mit einer Bildverarbeitungsbibliothek, z.B. unter Labview, erfolgen, ist unter Umständen ebenfalls mit Problemen oder zusätzlichen Kosten für die Software zur Bilderfassung mit der speziellen Kamera zu rechnen.

[36] Teilweise ist es aber aus Performance Gründen notwendig, eine zusätzliche Interfacekarte im PC-einzusetzen, denn die auf den Hauptplatinen vieler PCs vorhandenen Interfaces können teilweise nicht die volle Übertragungsrate des Standards erfüllen. Dies ist allerdings sehr preisgünstig möglich.

Das *GenICam* Protokoll will solche Nachteile vermeiden. Die Idee ist dabei ein generisches Programmierinterface, so dass GigE, *CoaXPress* bzw. *CXP*, Camera Link (HS) und USB 3 Kameras softwareseitig über dasselbe Interface angesprochen werden können. So sollen die Softwareentwicklungskosten deutlich reduziert werden. Generell kann aber auch der Einsatz einer geeigneten Bildverarbeitungsbibliothek, z.B. Common Vision Blox, hinsichtlich einer Vereinheitlichung der Bilderfassungshardware Sinn machen.

Der Nachteil des notwendigen Framegrabbers bei Camera Link und CoaXPress kann aber auch ein Vorteil sein. Nämlich dann, wenn sowieso eine Vorverarbeitung auf dem Framegrabber vorgenommen werden soll (um hohe Auswertegeschwindigkeiten zu erzielen). Produkte mit integrierten digitalen Signalprozessoren (*DSP*) oder programmierbaren Logikelementen (*FPGAs*) sind am Markt verfügbar, um entsprechend schnelle Verarbeitungen zu ermöglichen. Außerdem kann der Framegrabber (mittels DMA) direkt in den Hauptspeicher des Computers die Daten schreiben. (Letztlich geht das natürlich auch mit GigE Systemen).

Im Bereich von preiswerten Consumersensoren wird vor allem der *MIPI CSI-2* Standard verwendet. Im industriellen Umfeld ist er dagegen nur bedingt verbreitet. Pro Verbindung kann damit maximal (und nur über kurze Entfernungen) 2.5 GBit/s übertragen werden.

Heutige Bildsensoren erzielen intern inzwischen sehr hohe Geschwindigkeiten (z.B. Sony Pregius 4: 38 GBit/s), die nicht mehr über konventionellen Schnittstellen übretragbar sind. Kompression kameraseitig und Dekompression beim Computer ist eine Möglichkeit, die effektive Datenrate zu erhöhen. Zu bedenken ist aber auch, dass extrem hohe Datenraten zu mehr Wärmeproduktion führen und damit einem erhöhten Stromverbrauch und einem vergrößerten Bildsensor (Wärmeabführung) zur Folge haben.

Welches Interface für eine Anwendung am besten geeignet ist, sollte jeweils durch die aktuell vorliegenden Anforderungen und die Marktsituation bestimmt werden. Ganz entscheidend ist dabei oftmals die Marktdominanz und die bereits vorhandene Erfahrung. Die wesentlichen Beurteilungskriterien sind:

- Verfügbarkeit der passenden Hardware (Kamera und PC-Anschluss) und Software (Treiber) und langfristige Investitionssicherheit

- Kosten (PC-seitig, Kameraseitig, Software)
- Komplexität/"Ease-of-use"
- Übertragungsgeschwindigkeit
- Möglichkeit des Anschlusses mehrerer Kameras (und Synchronisation)
- Maximale Kabellänge (evtl. Repeater)
- Störsicherheit der Übertragung (evtl. Glasfaser)
- Leistungsversorgung über Anschlusskabel
- Triggermöglichkeiten
- Zeitliche Stabilität
- *Hot-Pluggable* (kann die Schnittstelle ohne Probleme im Betrieb abgekoppelt werden)

Wesentlich ist aber auch, dass bei sehr hohen Datenraten meist auch die Verarbeitung ein Problem darstellt. In diesem Fall ist es sinnvoll zu prüfen, ob nicht FPGA-basiert eine (Vor-)Auswertung sehr sensornah vorgenommen werden kann.

9.15 Advanced: Intelligente Kameras

Sogenannte *intelligente Kameras* bzw. *Smart-Cameras* bestehen letztlich aus einem Bildsensor in Kombination mit einer digitalen Verarbeitungseinheit. Meist wird eine gewöhnliche CPU oder aber ein *digitaler Signalprozessor* (DSP) in das Gehäuse der Kamera integriert. Teilweise sind noch *FPGAs* für eine Bildvorverarbeitung vorhanden. Ganz generell spricht man von *embedded Vision* wenn letztlich die gesamte (oder zumindest der wesentliche Teil) der Bildverarbeitungslösung in einem speziell dafür abgestellten (meist vergleichsweise kleinen) Rechner — teilweise mit Kamerahardware und oder Beleuchtung — integriert ist. Entsprechende Rechner basieren entweder auf sogenannten „embedded PCs" (meist mit Windows als Betriebssystem) oder Linux PCs (z.B. Raspberry Pi oder NVidia TX). Die

Grenzen von „fest verdrahteten" zu eher frei programmierbaren Systemen sind fließend.

Dadurch ergibt sich ein einfach zu montierendes kleines System, das aus Sicht eines Systemintegrators perfekt einbindbar ist. Darüber hinaus können sich Vorteile dadurch ergeben, dass ein Echtzeitbetrieb gegeben ist. Zeitliche Probleme, wie sie sich bei konventionellen Desktop-Betriebssystemen gegeben sind, werden somit vermieden. Die Prozessierung muss sich also wirklich nur um die Kameradaten kümmern.

Diesen Vorteilen stehen aber auch gravierende Nachteile gegenüber: 1. Die Flexibilität sinkt (oder kann zumindest oft sinken), d.h. man kann nicht einfach z.B. den Bildaufnehmer selbst ändern bzw. austauschen, oft ist auch die Algorithmik stark auf das eingeschränkt, was der Hersteller vorgesehen hat. Dies ist insbesondere vor dem Hintergrund der enormen Vielseitigkeit der durch Bildverarbeitung zu erledigen Lösungen ein Problem. 2. Die Verarbeitungsgeschwindigkeit ist gegenüber einer konventionellen Lösung meist reduziert (genau wie der Platzbedarf) somit lassen sich nicht alle Bildverarbeitungsfunktionalitäten verwirklichen. 3. Es wird eine hohe Abhängigkeit zum Hersteller der Kamera eingegangen, denn man ist auf die Tools zur Softwaregenerierung des Herstellers und die Langzeitverfügbarkeit angewiesen, d.h. ein guter Support vom Hersteller ist unbedingt erforderlich (aber nicht zwangsweise gegeben[37]). Selbst wenn sich die Anwendung über die Jahre nicht ändert und die Hardware in Takt bleibt kann es notwendig werden, das System anzupassen. Ein wichtiger Punkt sind die zunehmenden Anforderungen aufgrund von Sicherheitsproblemen, die z.B. bei vernetzten Kameras zu beachten sind.

Vor dem Einsatz einer intelligenten Kamera ist dementsprechend genau zu prüfen, ob hier langfristig mit erheblichen Einschränkungen bei der Flexibilität und Abhängigkeit zu rechnen ist.

Eine Sonderform stellen die sogenannten *Vision-Sensoren* dar. Diese bezeichnen intelligente Kameras, die für eine spezielle Klasse von Aufgabenstellung vorkonfiguriert bzw. optimiert sind und nur in sehr geringem Maße programmierbar bzw. parametrierbar sind.

[37] insbesondere bei Einzellösungen, bei denen nicht viele Kameras verbaut werden

9.16 Advanced: Event-based Sensors

Sogenannte Event-based Sensoren geben nicht etwa fortlaufend ganze Bilder aus, sondern melden letztlich Pixel, die ihre Helligkeit (deutlich genug) geändert haben. Dies hat vor allem dann einen Vorteil, wenn nur wenige Pixel sich im Laufe der Zeit ändern. Entsprechend muss dann wesentlich weniger Information übertragen/gespeichert werden, so dass Detektionen auch im höheren kHz-Bereich und bei sehr geringer Latenz einfach möglich werden.

Man hat also nicht mehr eine feste Bildwiederholrate. Stattdessen lösen Pixel, bei denen sich die Helligkeit (deutlich genug) ändert ein „Event" aus. Einen guten (wenn auch aus meiner Sicht unrealistisch positiven) Überblick gibt [15].

Anstatt dann jeweils den gesamten Grauwert eines sich änderenden Pixels zu senden kann man auch nur die Änderung senden. Allerdings ergibt sich dann natürlich eine Akkumulation von Fehlern mit der Zeit. Den letztlich summierten Grauwerten ist also nicht wirklich zu „trauen" und der Sensor muss regelmäßig wieder zurück gesetzt werden. Es sind auch Modelle verfügbar, bei denen das Prinzip mit einer herkömmlichen Bildausgabe verbunden sind, so dass jedes zweite Pixel ein „Event-Pixel" ist.

Problematisch ist der Ansatz, wenn auch geringere Intensitätsänderungen von Belang sind, z.B. bei der Verschiebung eines ausgedehnten Spots und nachfolgender Schwerpunktsbestimmung. Generell ist die radiometrische Genauigkeit der rekonstruierten Bilder prinzipbedingt problematisch, vor allem auch vor dem Hintergrund der Rauschproblematik.

Neben der Geschwindigkeit ist aber auch die Dynamik ein großer Vorteil der Sensoren, da sich der Sensor praktisch automatisch pixelweise an die Bestrahlungsstärke lokal anpasst.

9.17 Zusammenfassung wesentlicher Parameter der Bildsensoren

Üblicherweise wird man heute für die meisten Anwendungen einen CMOS Sensor mit nicht zu kleinen Pixeln wählen. Die konkrete Pixelgröße hängt von der

gewünschten Gesamtpixelzahl und den vertretbaren Kosten ab. Full-frame Sensoren werden eher bei erhöhten Anforderungen an die Empfindlichkeit gewählt. In jedem Fall sind progressive-scan (non-interlaced) Sensoren mit globalem Shutter zu wählen wenn eine Bewegung nicht ausgeschlossen werden kann.

Details zu den hier besprochenen Sensorparametern sowie Messvorschriften finden sich in der EMVA 1288 [2]. [38]

Für die meisten Anwendungen sind neben Pixelanzahl, Sensorgröße und Preis vor allem Rauschverhalten, Sensitivität und Geschwindigkeit von Bedeutung. Sensorgröße und Schnittstelle bestimmen oft wesentlich über den Preis. Aber auch die Einfachheit des Einsatzes ist oft wichtig. Hier sind Fragen hinsichtlich der Software wichtig.

Wir listen hier nochmal die wesentlichen Parameter, die ggf. zu prüfen bzw. zu überdenken sind:

- CCD oder CMOS
- Anzahl Pixel
- Pixelgröße
- Sättigungskapazität
- Sensorformat

- Spektrale Empfindlichkeit, Quanteneffizienz
- Elektronische Rauschanteile, Signal-Rausch-Verhältnis (ggf. Doubling Temperatur, fixed-pattern noise)
- Empfindlichkeit und Verstärkung
- Dynamik und Quantisierung
- ggf. Farbfilter, Transmission

---

[38] Beispielspezifikationen basierend hierauf sind in https:www.zenodo.orgcollectionuser-emva1288 zu finden.

- Bildwiederholfrequenz
- Shutter
- Triggerung
- minimale und maximale Belichtungsdauer

- Interface
- Software (SDK, Support, Demoprogramme, Einbindung in vorhandene Systeme)
- Region-of-Interest bzw. partial scan
- Binning
- Programmierbarkeit
- Spezialfunktionen: z.B. Look-up Tables
- Kennlinie und Linearität

- Objektivgewinde
- Deckglas: Beschichtung, Transmission
- Mikrolinsen optimiert auf welche Objektive (Vignettierung ?)

- Leistungsbedarf (und Erwärmung)
- zulässige Arbeitsbedingungen
- Ansteuerung externer Komponenten (vor allem Blende, Fokus, Zoom) direkt durch die Kamera
- erlaubter Temperaturbereich
- Wasserresistenz (IP-Schutzklasse)
- Shockresistenz (in g)

- Dimension, Gewicht, Robustheit (Vibration und Schock) und mechanische Befestigung
- Lebensdauer bzw. Mean-Time-Between-Failure
- Langzeitverfügbarkeit
- Kundenservice
- Preis

Es sind also sehr viele Parameter zu überprüfen und so erklärt sich auch die Vielfalt der verfügbaren Sensoren. Beachten Sie, dass ein Teil dieser Parameter (Temperaturbereich, Shockresistenz etc.) meist notgedrungen dann auch für das Objektiv gelten müssen.

Kapitel 10

Advanced: Topografiesensoren

Die dreidimensionale Erfassung von Objekten nimmt in der industriellen Bildverarbeitung kontinuierlich an Bedeutung zu. Sobald direkt dreidimensionale Größen erfasst werden *müssen* ist der Einsatz einer dreidimensionalen Datenerfassung natürlich unumgänglich.

Oft kann die *Dreidimensionalität* aber auch für konventionelle zweidimensionale Aufgabenstellungen starke Vorteile bieten. Ein typisches Beispiel ist die Erkennung von an sich dreidimensionalen Defekten (z.B. Kratzer) oder Merkmalen bei überlagerten zweidimensionalen Störgrößen (z.B. Waschmittelflecken, Beschriftungen). Durch die dreidimensionale Erfassung kann so ermittelt werden, an welchen Stellen wirklich eine relevante Abweichung (eben z.B. eine Delle oder ein Kratzer) vorliegt. Starke Oberflächentexturierung können ausgeblendet werden. Letztlich bietet die dreidimensionale Erfassung einen zusätzlichen Datenkanal der die Auswertung unterstützen bzw. vereinfachen kann.

Leider wird dieser Vorteil in der Regel mit einer deutlich komplexeren Hardware zur Datenerfassung und oft auch verlängerten Datenerfassungszeit erkauft. Auch die möglichen Fehler bzw. Probleme bei der Erfassung nehmen deutlich zu und letztlich ist auch die Auswertung teilweise deutlich komplexer.

Im Vergleich zur konventionellen zweidimensionalen Bilderfassung ist eine hochwertige 3D Erfasssung also meist vergleichsweise teuer, oft deutlich zeitintensiver und häufig ergeben sich auch starke Messfehler oder gar nicht auswertbare Bereiche auf dem Objekt. Die reine Messwerterfassung ist weniger robust und die Beurteilung des lokalen Ergebnisses deutlich schwieriger als bei gewöhnlichen

Bildern.

Insgesamt sind *Topografiesensoren* somit hinsichtlich ihrer Eigenschaften und Implementierung meist deutlich komplexer als konventionelle Bildsensoren. Die höhere Komplexität zeigt sich dabei nicht zuletzt auch in der schwierigeren Beurteilung der Leistung entsprechender Sensoren.

Sehr viele unterschiedliche Messmethoden sind zur dreidimensionalen Erfassung gebräuchlich. Wir können diese Methoden allerdings auf wenige im industriellen Umfeld gebräuchliche Grundprinzipien rückführen. Mehr als 99% aller industriellen 3D Anwendungen werden aktuell durch entsprechende Sensoren realisiert. Dabei beschränken wir uns auf sogenannte *technische Oberflächen*, d.h. Oberflächen, die eine gewisse *Rauheit* aufweisen und somit nicht spiegelnd sind.

Gemäß unserer Darstellungsweise bei den anderen Hardwarekomponenten (Objektive und Bildsensoren) werden wir auch hier nicht darauf eingehen *wie im Detail* entsprechende 3D-Sensoren zu realisieren sind, sondern die Anwendung der Sensoren und ihre prinzipielle Funktionsweise darstellen. Einen 3D Sensor wird man üblicherweise nicht selbst bauen, sondern — genau wie einen Bildsensor auch — zukaufen.

10.1 Auflösungen und Messunsicherheiten

Absolut wesentlich ist
a) zwischen *Auflösungen* und *Messunsicherheiten* zu unterscheiden und
b) sich klar zu machen, dass zwischen lateraler und axialer Messunsicherheit bzw. Auflösung zu unterscheiden ist.

Je nach Messprinzip sind laterale und axiale Größen gekoppelt oder auch nicht. Ein extremes Beispiel ist die Interferometrie bei der oft axiale Auflösungen bis in den sub-Nanometer Bereich erzielt werden während die laterale Auflösung bei vielen Geräten eher im Bereich von Millimetern liegt. Bei vielen Messsystem ist andererseits eine sehr gute axiale Auflösung oft durch laterale Mittelung, also auf Kosten der lateralen Auflösung, erreichen.

Hersteller geben oft anstelle einer ausführlichen Angabe den „schönsten" Wert

(meist die axiale Auflösung) unter günstigen Bedingungen an. Der Kunde muss sich aber im klaren sein, dass die axiale Messunsicherheit deutlich schlechter als die Auflösung ist und die laterale Auflösung (im Sinne z.B. einer Zweipunkttrennung, vgl. Kapitel 5) ebenfalls deutlich verringert ist. Es macht auch einen gewaltigen Unterschied, ob die Höhe eines Spiegels oder die Position einer rauen Kante im Raum ermittelt werden soll. Oftmals sind die Werte kaum von verschiedenen Aufgabenstellungen ineinander übertragbar. Es ist also wichtig, genau zu wissen, wie ein Hersteller seine Angaben ermittelt hat und was sie wirklich bedeuten.

Für alle Verfahren ist die Kalibrierung hinsichtlich der erzielbaren Messunsicherheiten ganz wesentlich, falls rückführbare Messungen durchgeführt werden sollen (teilweise ist das nicht nötig, denn für Detektionsaufgaben ist oft eine exakte Bestimmung metrischer Größen nicht notwendig, z.B. Kratzerdetektion).

Kalibrierung und Rückführbarkeit (vgl. Abschnitt 11.2) werden gegenüber reinen zweidimensionalen Aufgaben deutlich komplexer. Teilweise wird zwischen „echten" 3D Messungen, die eine dreidimensionale Punktwolke liefern und sogenannten „$2\frac{1}{d}$ D" Messungen unterschieden. $2\frac{1}{d}$ D bedeutet dabei, dass auf einem zweidimensionalen Raster ein Höhenwert bestimmt wird. Hinterschneidungen oder ähnliches können so nicht ermittelt werden. [1]

In jedem Fall ist aber nötig, die Kalibrierung von messenden Verfahren generell und von dreidimensional messenden Verfahren im Besonderen ganz explizit als zentralen Punkt wahrzunehmen. Letztlich entscheidet oft die Kalibrierung und nicht das Messprinzip über die schlussendlich erzielbare Genauigkeit.

10.2 Was wird gemessen

Gedanklich trennt man bei Werkstücken meist die *Mikrogeometrie* von der *Makrogeometrie*. Die Mikrogeometrie besteht aus sehr feinen (Größenordnung Mikrometer) regelmäßigen oder zufälligen Strukturen. Die zufälligen Anteile sind im wesentlichen die Rauheit und da diese für viele Anwendungen sehr wichtig ist (Beispiel: Laufflächen, Reibung) können optische 3D Messmethoden eingesetzt wer-

[1] Nahezu alle kommerziell im industriellen Umfeld eingesetzten Sensoren fallen in diese Kategorie.

den, um diese zu messen.[2]

Aber auch wenn die eigentliche *Mikrotopographie* uninteressant ist, kann sie dennoch die Messung der *Makrotopographie* stark beeinflussen. Wir werden auf diesen Aspekt bei den Triangulationsverfahren etwas weitergehend eingehen, prinzipiell ist eine Beeinflussung aber bei jedem der Messverfahren zu bedenken.

Eine weitere Beeinflussung ist durch dünne *Schichten* auf der Oberfläche gegeben. Diese führen für das Licht zu Phasensprüngen und damit zu Messfehlern bei allen gängigen Verfahren, wenn hochgenaue Messungen (sub-Mikrometer Bereich) notwendig sind.

Entscheidend ist also bei der dreidimensionalen Erfassung — noch viel mehr als es ohnehin schon bei der zweidimensionalen Bildgebung der Fall ist —, genau zu wissen bzw. zu spezifizieren, was man eigentlich genau benötigt und wie die konkreten Bedingungen sind. Ein Vergleich der Sensoren basierend auf wenigen Sensorparametern führt leider in vielen Fällen nicht zu einer optimalen Lösung. Eine sehr praxisnahe Einführung in die gängigen 3D Verfahren für technische Objekte findet man bei [4].

10.3 Robustheit

In der fertigungsnahen Messtechnik ist oft mit *Vibrationen* zu rechnen. Diese stellen für hochgenaue Topografiemessungen daher zunächst ein Problem dar. Sobald ein scannender Ansatz Verwendung findet muss beachtet werden, dass unterschiedliche Objektpunkte zu unterschiedlichen Zeiten gemessen werden und Vibrationen daher Topografien vortäuschen können. Eine Variante ist es, parallel zur optischen Topografiemessung dann die Vibration zu erfassen, um so nach (oder auch bei) der Messung die Messdaten zu korrigieren.

Vibrationen können aber bei vielen der Messverfahren noch deutlich problematischere Folgen nach sich ziehen. Viele Messverfahren benötigen für die Ermittlung eines *einzigen* Höhenwerts bereits mehrere Messungen zu unterschiedlichen Zeiten. Wenn während dieser Zeiten vibrationsbedingte Höhenschwankungen auf-

[2] Mittlere Rauheiten sind oft sehr gering. Wenn Sie sich noch (schlecht) aber sichtbar in der Oberfläche spiegeln können, dann liegt die mittlere Rauheit unter ca. 150 nm.

treten sind die Folgen oft kaum zu korrigieren.

Daher bevorzugt man in diesem Fall sogenannte *Einzelschuss* (*single-shot*) Verfahren bei denen zumindest dieses Problem umgangen wird.

Vibrationen sind ein wichtiger Aspekt hinsichtlich der *Robustheit*. In der Praxis ist für viele Anwendungen nicht der genauste Sensor am besten sondern ein "robuster" Sensor, der eine noch ausreichende Genauigkeit unter allen relevanten Bedingungen erzielt. Insbesonder variiert die erzielbare Genauigkeit oft stark mit der Oberfläche des Objekts aber eben auch mit diversen Nebenbedingungen (z.B. den bereits angesprochenen Vibrationen).

10.4 Triangulation

Triangulationsbasierte Messverfahren werden in vielfältiger Form in der industriellen Bildverarbeitung eingesetzt. Sie sind vergleichsweise einfach und daher einigermaßen kostengünstig realisierbar. Grundsätzlich unterscheidet man zwischen passiven und aktiven Verfahren. *Aktiv* bedeutet dabei schlicht, dass das Objekt in irgendeiner Form unabhängig vom Umgebungslicht und einer nicht-trivialen Weise beleuchtet wird. Bei Triangulationsverfahren werden so z.B. ein Punkt, eine Linie oder auch ein ausgedehntes Muster (z.B. Streifen) verwendet.

Die aktive Beleuchtung ist tendenziell schwierig bei sehr großen Objekten bzw. Szenen (Szenendimensionen größer einige Meter), liefert aber in der Regel eine deutlich robustere Erfassung.

Abb. 10.4.1 verdeutlicht die triangulationsbasierte Entfernungsbestimmung am Beispiel der vom Menschen eingesetzten "Stereo-Vision": Letztlich ist für das Gehirn klar, dass das Objekt sowohl auf dem Sehstrahl (Hauptstrahl der optischen Abbildung) für das rechte und für das linke Auge liegt. Also muss das Objekt im Raum beim Schnittpunkt der beiden Geraden liegen.[3]

[3] Natürlich macht das Gehirn nur indirekt eine Schnittpunktsberechnung. Die Zusammenhänge sind vielmehr erlernt.

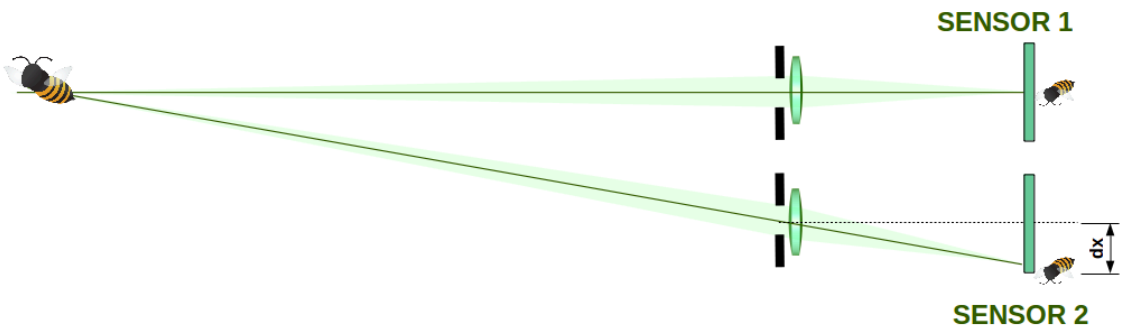


Abbildung 10.4.1: Binokulares Sehen bzw. passives Stereo-Vision: Aus der Differenz der Bildposition im rechten und im linken Auge/Kamera kann auf die Entfernung rückgerechnet werden. Letztlich muss der Schnittpunkt beider Sehstrahlen (linkes und rechtes Auge) ausgehend von zueinander korrespondierenden Punkten gefunden werden.

10.4.1 Aktive Verfahren

Eine einfache technische Umsetzung für die Messung einer einzelnen Entfernung ist in Abbildung 10.4.2 am Beispiel der (aktiven) *Punkttriangulation* dargestellt. Der an der Oberfläche gestreute (!) Lichtpunkt wird auf die (Zeilen-)Kamera abgebildet und aus der Position des Bildpunkts kann offensichtlich auf die Entfernung des Punkts rückgeschlossen werden.

Für eine (objekt-seitig) telezentrische Erfassung des Bildes wird der mathematische Zusammenhang von lateraler Position des Bildpunkts in der Kameraebene und Tiefe besonders einfach. In diesem Fall können wir für den Triangulationswinkel α direkt angeben, dass ein Höhenversatz um Δz zu einem Δx in der Objektebene der Kamera mit (vgl. auch Abb. 10.4.3)

$$\Delta x = \Delta z \, \sin(\alpha) \tag{10.1}$$

führt. Entsprechend ist dann das zugeordnete $\Delta x'$ in der Bildebene einfach wieder $\Delta x' = \beta' \, \Delta x$.

Für andere Abbildungsgeometrien (insbesondere nicht-telezentrische Abbildungen) wird die Geometrie leicht anders, das Grundprinzip bleibt aber natürlich und auch der generelle Zusammenhang, dass die Empfindlichkeit mit zunehmendem Triangulationswinkel steigt bleibt erhalten.

Oft wird der Sensor mit einem Laser realisiert, man spricht auch von *Lasertriangulation*. Die Ausdehnung des Lasers im Objektraum ist natürlich nicht Null. Idea-

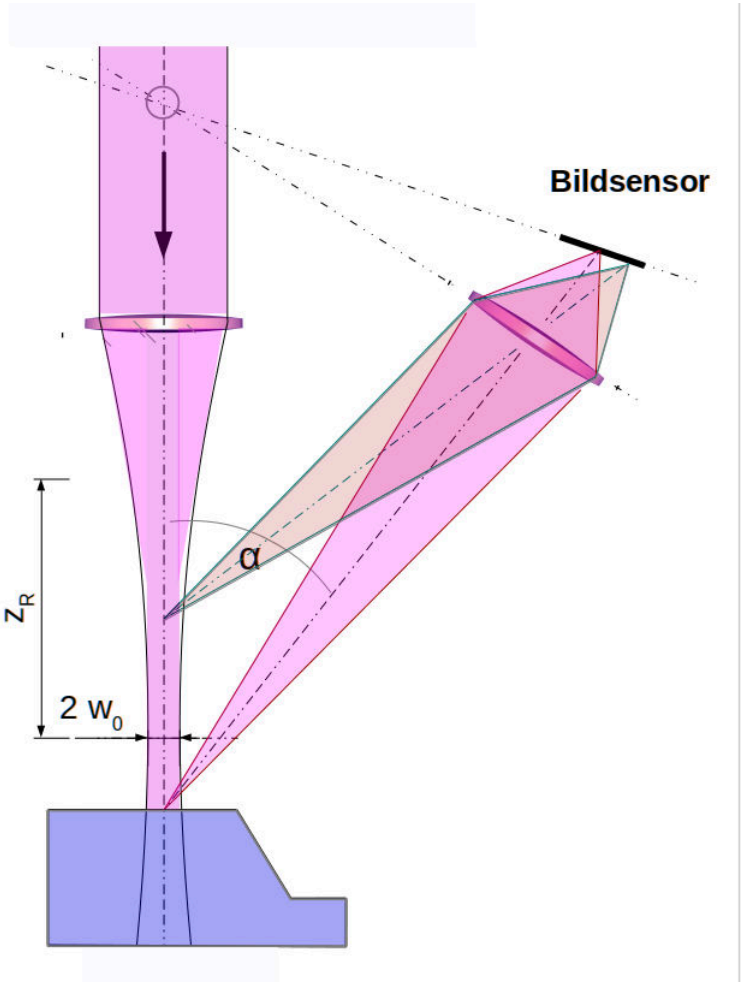


Abbildung 10.4.2: Prinzip der Punkttriangulation (dargestellt mit Scheimpflugkorrektur).

lerweise fokussiert man den Laser auf die Mitte des Objektbereichs. Entscheidend ist dann die sich dort ergebende Taille, die über einen gewissen Bereich (Rayleigh-Länge) dann relativ konstant bleibt. In jedem Fall wird durch diese Ausdehnung die laterale Auflösung der 3D-Messung begrenzt.

In der in Abb. 10.4.2 dargestellten Geometrie ist die die Objektebene der Kamera gegenüber der Grundausrichtung des Objekts gekippt. Dies kann zu Problemen hinsichtlich der Schärfentiefe der Abbildung führen, d.h. unterschiedliche Objekthöhen führen dann zu unterschiedlich scharfen Abbildungen. Dies kann (insbesondere bei nicht punkt- sondern flächig-messenden Systemen zu Problemen führen. In diesem Fall wird dann eine Abbildung nach Scheimpflug verwendet (vgl. Abschnitt 6.7).

Für eine hohe Messrate bei der scannenden Messung mit einem Punktsensor muss

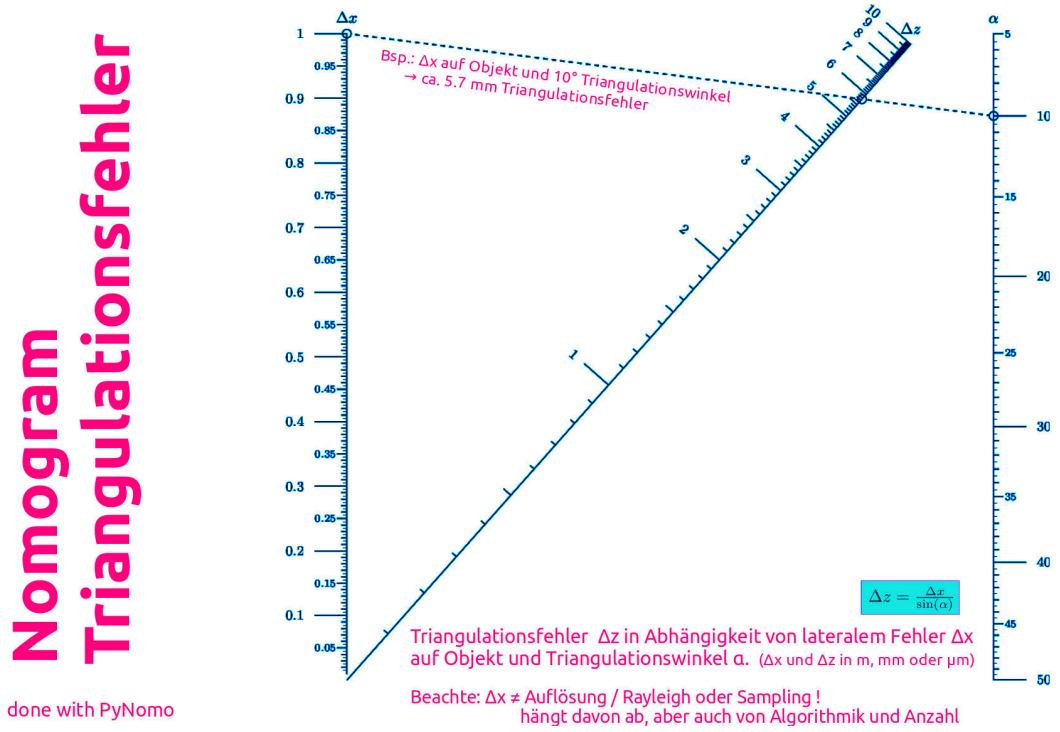


Abbildung 10.4.3: Messunsicherheit bei der Triangulationsmessung (ohne Speckles)

der eigentliche Detektor zur Messung der Position des Lichtspots entsprechend schnell sein. Hier können Positionsempfindliche Detektoren/Dioden/Devices (engl. „position sensing diode/device/detector", *PSD*) verwendet werden. Diese liefern je nach Auftreffort auf der Sensorfläche direkt ein (üblicherweise analoges) Signal, das die Position repräsentiert. Alternativ werden schnelle Zeilenkameras eingesetzt, so dass Messfrequenzen bis in den Größenordnungsbereich von mehreren 100 kHz möglich sind.

Soll mehr als ein Punkt vermessen werden kann ein Punktsensor zu einem Scanner (zusätzlicher Scanspiegel) erweitert werden. Alternativ kann man teilweise schon vorhandene Bewegungen des Werkstücks (Fließband, Rotation, ...) nutzen, muss dann aber die notwendige Gleichformigkeit der Bewegung beachten bzw. sicherstellen. Bei Verwendung von Zeilenkameras mit einer Zeilenfrequenz von 100 kHz können so im Prinzip 100.000 dreidimensionale Abtastungen pro Sekunde vorgenommen werden (vorausgesetzt der Scanner ist entsprechend schnell).

Eine einfache Erweiterung von der Erfassung einzelner Punkte hin zur eindimensionalen Linie führt zum sogenannten *Lichtschnitt*, einem sehr gebräuchlichen Verfahren (Abb. 10.4.4). In diesen Fall werden gleichzeitig die Höhenwerte entlang

einer Linie gemessen. Anstatt des Zeilensensors wird also ein flächiger Bildsensor eingesetzt.

Für hohe Geschwindigkeiten (z.B. 30 KHz) werden CMOS Sensoren eingesetzt, die aus dem zweidimensionalen Bild in hoher Geschwindigkeit (Prozessierung auf Sensor) die Höhendaten entlang der projizierten Linie berechnen.

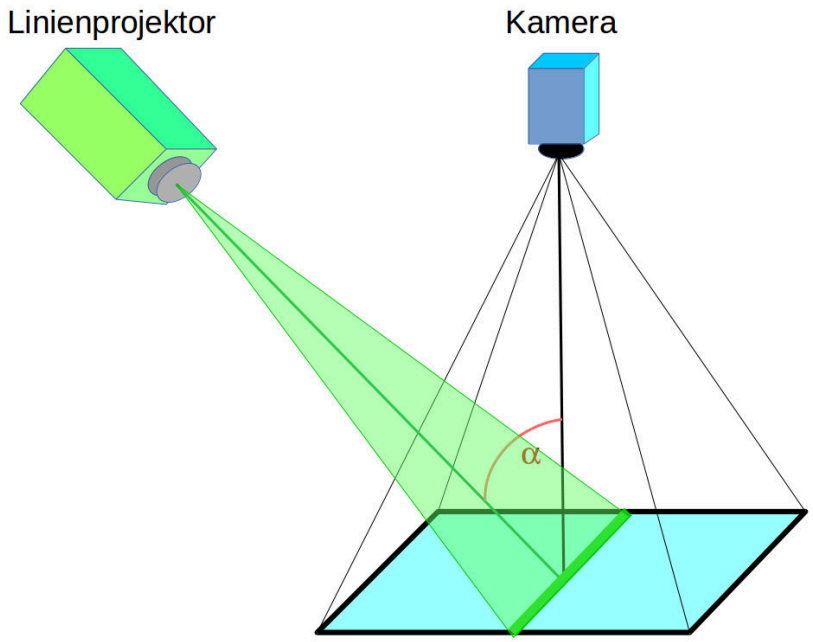


Abbildung 10.4.4: Prinzip der Linientriangulation („Lichtschnitt")

Noch mehr Daten können gleichzeitig erfasst werden, indem gleichzeitig viele Linien auf das Werkstück projiziert werden. Entsprechende Verfahren werden meist *strukturierte Beleuchtung* oder *Streifenprojektion* genannt. Meist werden für eine optimale Auswertung Aufnahmen mit mehreren verschiedenen projizierten Mustern durchgeführt, z.B. einer Kombination aus binären Mustern und sinusförmigen Streifen. Generell führen mehr Bildaufnahmen zu einer verbesserten Messunsicherheit und einer Erhöhung des sogenannten Eindeutigkeitsbereichs der Messung.

Oft werden anstelle einer Kamera und eines Projektors auch zwei Kameras eingesetzt, so dass eigentlich ein Stereoverfahren Verwendung findet. Die Projektion dient dann „lediglich" dazu, dass Strukturen auf der Oberfläche entstehen. Zum Einsatz zur Patterngenerierung kommen meist Laser in Verbindung mit diffraktiv-optischen ELementen oder Mikrolinsenarrays, es sind aber auch andere Muster,

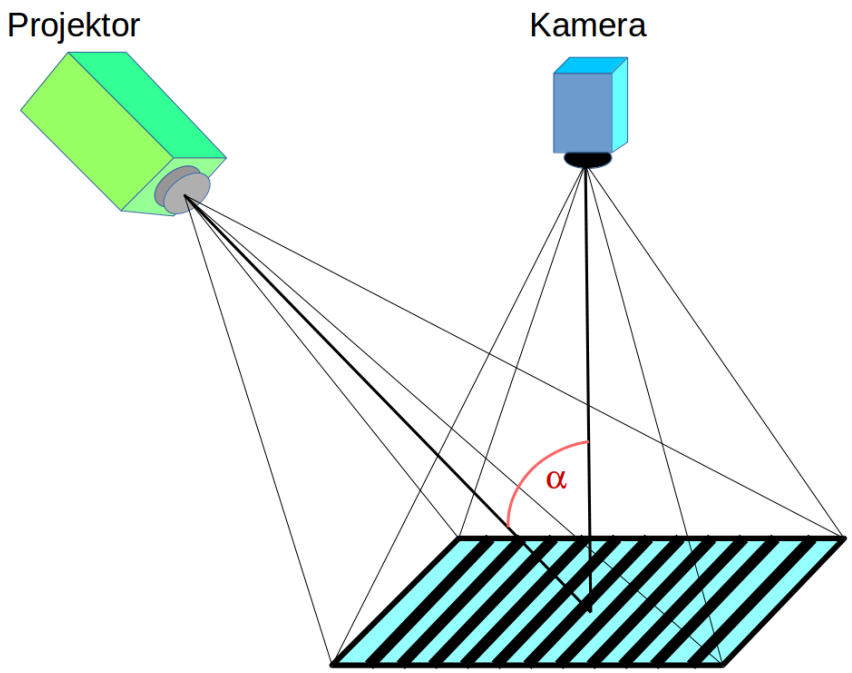


Abbildung 10.4.5: Weitere Parallelisierung: „Streifenprojektion" bzw. „strukturierte Beleuchtung".

z.B. Specklemuster, verwendbar. Wieder kann durch eine zeitliche Änderung des Musters und die Nutzung mehrerer Bilder bzw. Bildpaare eine Verbesserung der lateralen Abtastungm der Entfernungsbestimmung und Robustheit erreicht werden.

Die Verfahren lassen sich in weiten Größenbereichen einsetzen und können durchaus geringe bis sehr geringe axiale Messunsicherheiten, z.B. unter 1 Mikrometer bei der mikroskopischen Streifenprojektion, erreichen — natürlich nur bei einem entsprechend kleinen Objektfeld im mm-Bereich. Die Beurteilung entsprechender Systeme ist nicht immer trivial. Zu achten ist insbesondere auf die Kombination der lateralen und der Höhenauflösung. Für viele Anwendung ist eine Höhenauflösung von wenigen Mikrometern nicht sinnvoll, wenn gleichzeitig die Messfleckgröße einige Hundert Mikrometer beträgt (siehe Beispiel oben).

Für die aktiven Triangulationssysteme können wir grob folgende Vor- und Nachteile annehmen:

+ gut eingeführte und verstandene Technik

+ vergleichsweise einfache Implementierung $\rightarrow$ preiswert

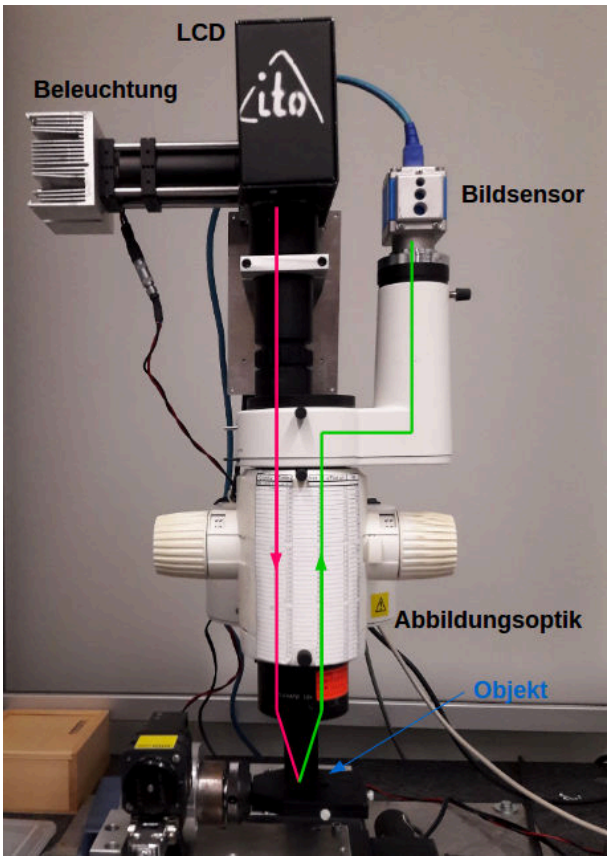


Abbildung 10.4.6: Streifenprojektionssystem für Mikroobjekte

+ Möglichkeit der Messung in verschiedenen Skalen von Mikro bis Makro.

- Problem von Abschattungen

- Hohe Dynamik an Metalloberflächen notwendig (Reflexionen)

- Für spiegelnde Oberflächen nicht geeignet.

10.4.2 Messunsicherheit

Die Messunsicherheit ist also umso besser, je kleiner das gerade noch sicher zu bestimmende $\delta x'$ in der Kameraebene ist und je größer der Triangulationswinkel α ist (Gl. 10.1). Natürlich können wir in der Praxis α nicht beliebig groß machen, denn es treten bei dreidimensionalen Objekten schnell Abschattungen auf.

Aber auch das $\delta x'$ unterliegt Beschränkungen. Zunächst haben wir einen pixelierten Bildaufnehmer. $\delta x'$ ist aber nicht durch die Pixelgröße gegeben, denn

wir können bei einem einzelnen ausgedehnten Punkt auf dem Bildsensor mittels Schwerpunktsbildung die Position wesentlich genauer als auf einen Pixel bestimmen. In der Praxis sind hier Genauigkeiten bis zu 1/50 Pixel denkbar[4] und natürlich hängt hier die Genauigkeitsbestimmung auch vom Rauschen des Sensors ab.

Wesentlich ist aber, dass es noch eine fundamentalere Begrenzung gibt, die vor allem bei den üblicherweise eingesetzten Lasern (aber auch anderen eher schmalbandingen Lichtquellen) das Limit für die Genauigkeit festlegt.

Das Problem liegt letztlich in der *Kohärenz* begründet. Sobald wir eine optisch raue Oberfläche, also eine Oberfläche, die uns nicht spiegelnd erscheint (man spricht hier auch von „technischen Oberflächen") mit kohärentem Licht abbilden, ergeben sich in der Bildebene Speckles. Diese Speckles sind zunächst in ihrer Form zufällig. Im Mittel haben sie gerade die Größe, einer Auflösungszelle, aber aufgrund der Zufälligkeit der Struktur der Speckles schwankt natürlich auch der Schwerpunkt des Speckles erheblich. Damit wird die Bestimmung von $\delta x'$ letztlich prinzipiell ungenau solange die exakte Mikrotopografie nicht bekannt ist.[5]

Anschaulich können wir nicht wie sonst bei Positionsmessungen die Mittelung von N Photonen zum Einsatz bringen sondern wir sind auf die Größenordnung der Spotradius begrenzt. Der objektseitige Spotradius der Abbildung ist $0.61\,\lambda/\mathrm{NA}_{\mathrm{Abb}}$ und es ergibt sich so im kohärenten Fall

$$\delta z = \frac{1}{2\pi}\frac{\lambda}{\mathrm{NA}_{\mathrm{Abb}}}\frac{1}{\sin\alpha} \tag{10.2}$$

Um die Messunsicherheit zu reduzieren müssen somit die Speckles reduziert bzw. gemittelt werden. Hierzu kann sowohl bei der räumlichen wie auch der zeitlichen Kohärenz angesetzt werden oder aber man mittelt statistisch unabhängige Specklemuster umso letztlich zum selben Punktbild wie bei einer inkohärenten Abbildung zu kommen.

[4] Eine leicht ausgedehntes (z.B. mittels Defokussierung) Punktbild ist dabei eher von Vorteil.

[5] Und selbst wenn sie bekannt wäre (was natürlich die Aufgabe der 3D-Vermessung unnötig machen würde), gelänge es uns zwar theoretisch, nicht aber in der Praxis, für sinnvolle Größen eine Rückrechnung vorzunehmen.

Gl. 10.2 wird mit dem *Specklekontrast* C dann zu (vgl. Abb. 10.4.7)

$$\delta z = \frac{C}{2\pi} \frac{\lambda}{\mathrm{NA}_{\mathrm{Abb}}} \frac{1}{\sin \alpha} \tag{10.3}$$

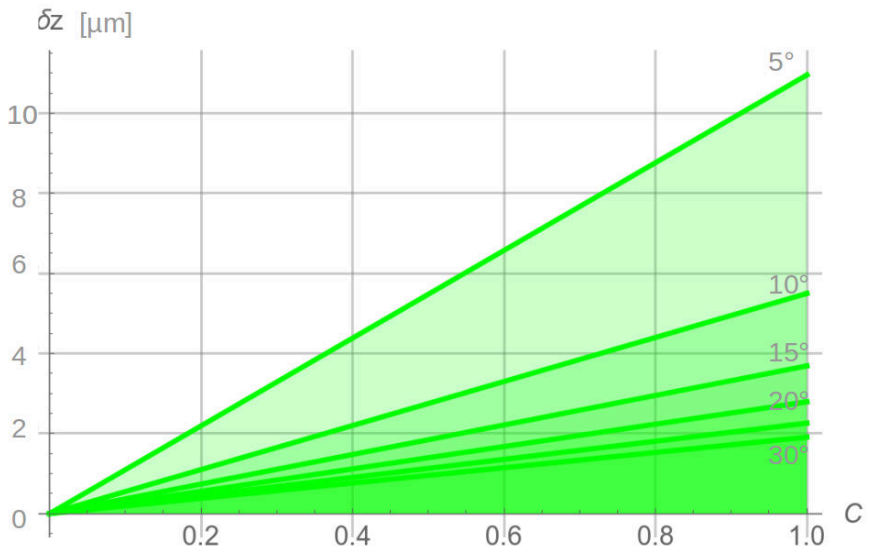


Abbildung 10.4.7: *Axiale Messunsicherheit aufgrund von Speckles in Abhängigkeit des zeitlichen* Kohärenzgrades C_T für eine objektseite numerische Abbildungapertur von 0.1 und $\lambda = 600$ nm bei verschiedenen Triangulationswinkeln. Für halb so große numerische Aperturen verdoppelt sich die Unsicherheit (vgl. Gl. 10.3).

Der Kohärenzgrad C hängt von der Ausdehnung des genutzten Spektralbereichs (zeitliche Kohärenz) sowie der Geometrie der Beleuchtung (räumliche Kohärenz) ab. Günstig sind also sehr polychromatische räumlich ausgedehnte Lichtquellen. Eine große weiße LED ist somit zunächst einer kleinen monochromatische LED vorzuziehen.

Der Kohärenzgrad aufgrund der zeitlichen Kohärenz C_T ist bei Beleuchtung mit einer *Kohärenzlänge* L einer rauen Oberfläche mit der Rauheit (Standardabweichung) σ_z gegeben durch [9] (vgl. Abb. 10.4.8):

$$C_T^2 = \frac{1}{\sqrt{1 + (4\sigma_z/L)^2}} \tag{10.4}$$

Anschaulich sollte die Kohärenzlänge deutlich geringer als die Rauheit der Oberfläche sein, um die Speckles zu vermeiden. Dies ist in der Praxis allerdings für relativ glatte (wenn auch matte) Oberflächen in der Praxis kaum erzielbar.[6]

[6]Beachten Sie auch, dass das effektiv nutzbare Spektrum meist auf wenige hundert Nanometer begrenzt ist; un-

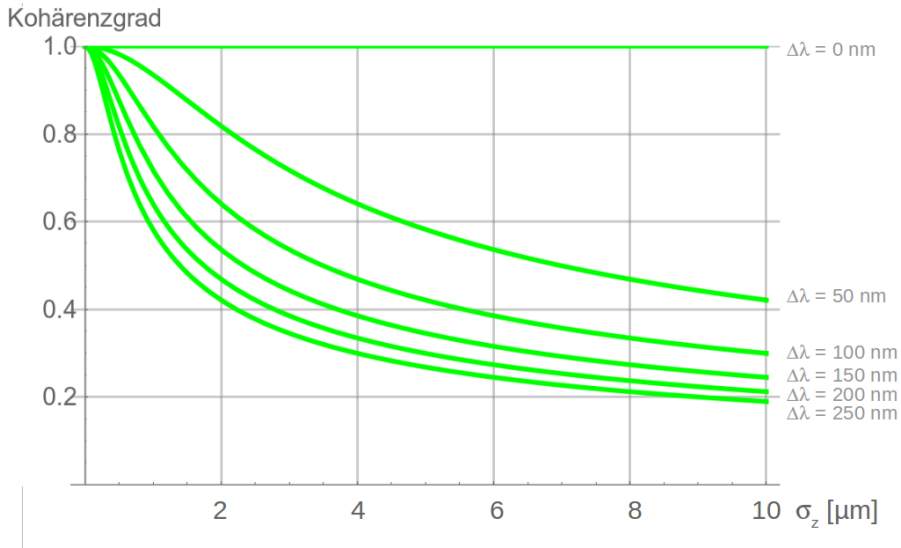


Abbildung 10.4.8: Der Kohärenzgrad der zeitlichen Kohärenz bei Beleuchtung einer rauen Oberfläche mit Rauheit σ_z unter verschiedenen Kohärenzlängen bei einer Wellenlänge von $\lambda = 600$ nm.

Aber auch die räumliche Kohärenz kann genutzt werden. Die einzelnen Punkte einer ausgedehnten Lichtquelle sind im allgemein nicht korreliert (bzw. kohärent zueinander). Das bedeutet, dass wenn wir in eine Auflösungszelle der Abbildungsoptik mehrere Lichtquellenpunkte getrennt abbilden, dann ergibt sich ein Mittelungseffekt. Dementsprechend muss die numerische Apertur der Beleuchtung dann größer sein (kleine Beleuchtungsspots) als die numerische Apertur der Beobachtung. Flächenmäßig bekommen wir $(\mathrm{NA_{bel}}/\mathrm{NA_{Abb.}})^2$ unabhängige Punkte in die Auflösungszelle des Abbildungssystems und damit ergibt sich ein Mittelungsvorteil von $\sqrt{(\mathrm{NA_{bel}}/\mathrm{NA_{Abb.}})^2} = (\mathrm{NA_{bel}}/\mathrm{NA_{Abb.}})$ wenn eine ausgedehnte Lichtquelle verwendet wird (z.B. LED).

In diesem Fall wird dann also die axiale Unsicherheit

$$\delta z = \frac{C}{2\pi} \frac{\lambda}{\mathrm{NA_{Bel}}} \frac{1}{\sin\alpha} \tag{10.5}$$

Andererseits ist es natürlich mit einer großen Lichtquelle schwieriger einen kleinen Punkt zu beleuchten. Letztlich bedeutet eine geringe räumliche Kohärenz, dass wir den Punkt der Oberfläche des zu vermessenden Objekts mit einem relativ großen Strahlenkegel, also einer großen beleuchtungsseitigen numerischen Apertur, beleuchten. Dies bedingt aber eine verringerte Schärfentiefe für den Be-

abhängig von der Lichtquelle! Denn letztlich ist das *genutzte* Spektrum ausschlaggebend und dieses wird vom Sensor (meist empfindlich im sichtbaren Spektralbereich) und der Lichtquelle bestimmt.

leuchtungsspot. Damit wird entweder der axiale Messbereich reduziert oder aber ein Autofokussystem ist notwendig.

Man kann auch mittels zeitlicher Mittelung, z.B. durch rotierende Diffusoren, zeitlich variable Lichtquellen (z.B. modulierte Laserdioden) oder aber auch einfach eine Bewegung des Messobjekts die Speckles reduzieren. So führt eine Erfassung von Objekten, die z.B. mittels eines Fließbands zugeführt werden bereits zu einer Specklereduktion. Andererseits kann die Bewegung natürlich je nach Belichtungszeit (und damit einhergehend Specklereduktion) zu einer Verschmierung in lateraler Richtung, also einer Reduzierung der lateralen Auflösung führen.

Auch mehrere Aufnahmen aus unterschiedlicher Orientierung (z.B. Rotationsanordnung) liefern natürlich unterschiedliche Speckles (und generell unterschiedliche Reflexionen/Fehler) und lassen sich so mitteln.

Die Zusammenhänge sind in der Praxis kompliziert und die genau erzielte axiale Messunsicherheit bei gegebener lateraler Auflösung korrekt vorherzusagen ist oft sehr schwer. Gemessenen Werten sollte hier klar der Vorzug vor (vielleicht zu stark simplifizierenden) Abschätzungen gegeben werden.

Üblich bei triangulationsbasierten Messsystemen (aber teilweise auch anderen 3D Messverfahren) ist letztlich die Angabe einer sogenannten *flächenbezogenen Auflösung*. Ein Wert von 1:20.000 bedeutet so z.B. dass bei einer Messfeldgröße von 100 mm die Höhenauflösung 100 mm / 20.000 = 5 μm beträgt. Die laterale Auflösung wird (bei statischen Objekten) durch die Größe des Lichtpunkts auf der Oberfläche und damit die gewöhnliche Rayleighauflösung und Aberrationen bestimmt.

Wenn man nun z.B. fordert, dass der axiale Messbereich 10 mm beträgt, und bei einer Blende von K gearbeitet wird, dann ergibt sich vereinfacht eine Messfleckgröße $\Delta \approx 2.44\lambda/K + 5$ mm / K. In diesem Fall für $\lambda = 500$ nm ergibt sich das Optimum für $K = 63$ mit ca. 150 μm .[7]

[7] K=63 bedeutet bei einem Arbeitsabstand von 1 m einen Durchmesser der Optik von 16 mm.

10.4.3 Passive Verfahren

Passive Verfahren, die ohne eine strukturierte Beleuchtung des zu vermessenden Objekts auskommen, werden als *Stereoverfahren* bzw. *Stereovision* bezeichnet. Letztlich ist das Prinzip dasselbe, die Projektion wird schlicht durch eine weitere Abbildung ersetzt. Ein Punkt in der Szene wird von beiden Kameras unter einem unterschiedlichen Winkel gesehen. Es ergibt sich so ein Versatz der Bildpunktpositionen für die beiden Kameras und aus diesem Versatz kann auf die axiale Position des Objektpunkts rückgerechnet werden.

Diese Rückrechnung ist an sich trivial. Schwierig ist dagegen das finden zueinander "korrespondierender" Punkte in den beiden Kamerabildern. Für homogene Flächen ohne Textur können natürlich prinzipiell keine korrespondierenden Punkte gefunden, denn es werden ganz generell dort keine Punkte gefunden. Abhilfe schafft das Anbringen künstlicher Marken auf dem Objekt oder aber eben wieder die aktive Beleuchtung der Oberfläche mit einem strukturierten Muster (Streifen, Speckles, Punktemuster etc.).

Bei mehr als zwei Kameras spricht man von *Multistereo*, wenn sehr viele Ansichten verwendet werden von *plenoptischen Kameras* oder *integral imaging*.

Prinzipbedingt lässt sich auf diese Weise nur dreidimensionale Information aus Szenenbereichen mit genügend Strukturen gewinnen. Die Genauigkeit ist gegenüber aktiven Verfahren reduziert.

Der einzige Vorteil ist, dass eben auf eine aktive Beleuchtung verzichtet werden kann und daher die Systemkosten etwas reduziert sind. Dieser Vorteil wiegt umso schwerer, je größer der zu erfassende (und damit auszuleuchtende) Bereich ist. In der industriellen Bildverarbeitung werden entsprechende Verfahren auf Grund der Nachteile und der meist gering ausgedehnten Szenen selten eingesetzt.

Grundsätzlich bleibt der Triangulationsansatz derselbe. Entscheidend ist nun, dass jeweils korrespondierende Punkte in den beiden Bildern gefunden werden, z.B. eine Ecke auf dem Objekt. Für homogene Flächen ist es dann notwendig, künstliche Marken auf dem Objekt anzubringen. Die Genauigkeit der Positionsbestimmungen in den Bildern ergibt dann (in Kombination mit der Genauigkeit der Kalibrierung) die Messunsicherheit. Wenn mehr als zwei Kamerabilder zur

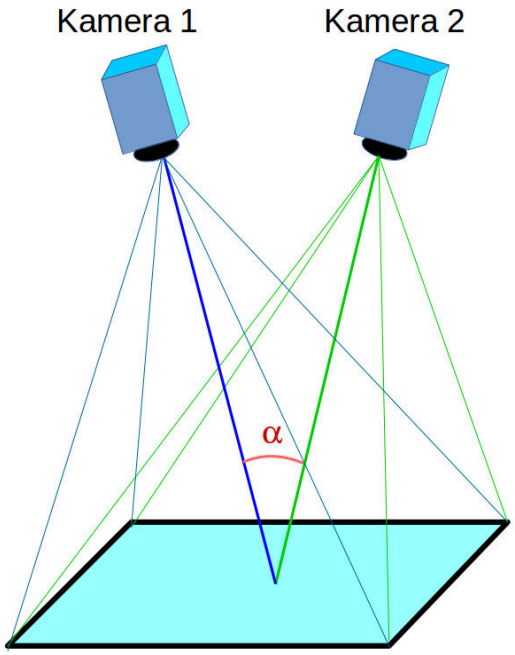


Abbildung 10.4.9: Stereoerfassung

Verfügung stehen steigt meist die Robustheit der Auswertung.

Folgende Vor- und Nachteile ergeben sich:

+ Gut eingeführte Technik

+ sehr preiswert

+ keine aktive Beleuchtung notwendig $\rightarrow$ einfaches System

- Probleme an untexturierten Oberflächen

- Problem von Abschattungen[8]

- Höhere Messunsicherheit als bei aktiven Verfahren

[8]Durch sogenannte „Dual-Sensor" Konfigurationen lassen sich diese deutlich vermindern. Letztlich werden zwei Detektoren aus unterschiedlichen Richtungen eingesetzt.

10.5 Weißlichtinterferometrie

Ein typischer Aufbau eines *Interferometers* ist in Abb. 10.5.10 dargestellt. Das Licht wird am halbdurchlässigen Spiegel in zwei Pfade aufgeteilt und später am selben Spiegel auch wieder vereint bevor es zum Detektor wandert.

Zur Berechnung der am Detektor einfallenden Intensität werden die elektrischen Felder beider Wellen addiert und das Ergebnis quadriert. Für sinus- bzw. cosinusförmige Schwingungen erhalten wir so

$$I \sim |E_1 + E_2|^2 = |A_1 \cos(\phi_1(t)) + A_2 \cos(\phi_2(t))|^2 \tag{10.6}$$

$$= \ldots = E_1^2 + E_2^2 + 2E_1E_2\cos(\phi_2 - \phi_1) \tag{10.7}$$

Wenn also die beiden interferierenden Teilwellen in Phase sind, dann erhalten wir eine hohe Intensität, wenn beide Teilwellen in Gegenphase ($\phi_2 - \phi_1 = \pi$) sind fällt die Intensität auf Null.

Die Phasendifferenz $\phi_2 - \phi_1$ kommt durch Längenunterschiede der beiden Lichtpfade zustande[9]

Die Phasen selbst sind mit der Wellenlänge skaliert. Eine Wellenlänge entspricht ja gerade einer (örtlichen) Periode der Schwingung des elektrischen Felds, also 2π. Entsprechend führt eine Verschiebung einer der beiden Spiegel in Abb. 10.5.10 um Δz zu einer Phasenänderung von[10]

$$\Delta\phi = 2 \cdot 2\pi \frac{\Delta z}{\lambda} \tag{10.8}$$

und zeigt ganz generell die extreme Empfindlichkeit der Interferometerie für Längenmessungen.[11]

Solange nur winzige Verschiebungen gemessen werden sollen funktioniert die bisher dargestellte interferometrische Messung sehr gut. Sobald die Verschiebung aber mehr als eine halbe Wellenlänge betragen kann haben wir das Problem, das unklar ist, ob die Verschiebung Δz oder aber $\Delta z + \lambda/2$ oder $\Delta z + 2 \cdot \lambda/2$ bzw.

[9] Alternativ kann man die Phasendifferenz für andere Aufgabenstellungen auch durch Brechzahlunterschiede induzieren.

[10] Beachte: Ein Verschieben des Spiegels führt (Hin- und Rückweg) zu einer doppelten Änderung der Lichtwegs, daher zusätzlicher Faktor 2

[11] Das extremste Beispiel findet sich in der Nutzung dieses Prinzips zur Detektion von Gravitationswellen.

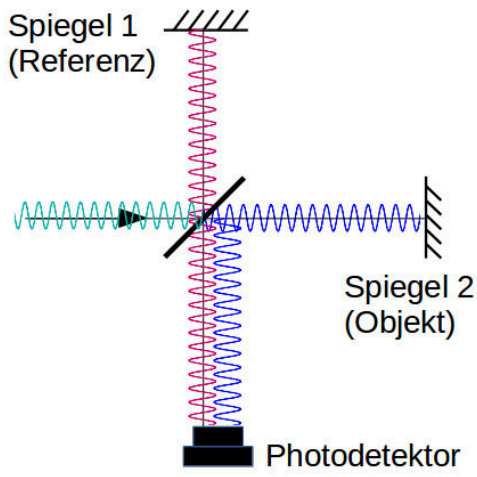


Abbildung 10.5.10: Grundprinzip der Interferometrie, hier am Beispiel eines Michelson Interferometers

allgemein

$$\Delta z + N \cdot \lambda/2 \text{ , mit} N \in \mathbb{N}, \tag{10.9}$$

betragen hat. Alle diese Verschiebungen würden zur selben Intensität führen. Das Messergebnis ist mehrdeutig.

Bei typischen Längenmessinterferometern, wie sie auch in der industriellen Praxis zum Einsatz kommen[12], gibt es zwei einfache Varianten, das Problem der Mehrdeutigkeit zu umgehen.

Variante 1 nutzt, dass die Bewegung nicht sprunghaft sondern kontinuierlich erfolgt und erfasst zeitlich hochaufgelöst (entsprechend schneller Detektor und Elektronik) das Messsignal so, dass zwischen zwei Messwerten nur eine minimaler Versatz entsteht.

Die zweite Variante benutzt zwei Laser unterschiedlicher Wellenlänge. Die Phasendifferenzen, die sich für die beiden Wellenlängen ergeben sind natürlich (Gl. 10.8) unterschiedlich, so dass aus der Kombination beider Intensitäten (bzw. Phasen) auf das N in Gl. 10.9 rückgeschlossen werden kann. Man spricht von *Heterodyninterferometerie*.

Unsere eigentliche Aufgabe ist es aber in diesem Kapitel, eine unbekannte Topografie zu messen. Und die eben dargestellte zweite Variante kann in der Tat

[12] häufig in hochgenau arbeitenden Führungsschienen

auch genutzt werden, um solche eine Topografiemessung durchzuführen. Man kann sich leicht denken, dass das Ergebnis stabiler wird, wenn mehr als zwei Wellenlängen verwendet werden und wenn wir sehr viele Wellenlängen aus einem gewissen Wellenlängenintervall $\Delta\lambda$ verwenden, dann landen wir bei der „Weißlichtinterferometrie."

Es wird in diesem Fall also — im Gegensatz zur konventionellen Interferometrie — kein monochromatisches Licht (Laser) verwendet, sondern eine vergleichsweise breitbandige Lichtquelle (z.B. eine LED).

Es gibt nun verschiedene Varianten, wie entsprechende Systeme technisch aufgebaut werden können und verschiedene Namen haben sich für diese weißlichtinterferometrischen Verfahren eingebürgert. Egal ob *Optische Kohärenztomografie* (*OCT*), *Kohärenzradar*, *Spektralinterferometrie* oder *Weißlichtinterferometrie*; das Grundprinzip ist immer dasselbe.

Technisch ist zunächst vor allem zwischen „spektralen" Systemen (*Spektralinterferometrie*) und Systemen „im Ortsraum" zu unterscheiden. Hinsichtlich der Ausführung zur Heterodyninterferometrie ist vermutlich das sogenannte *swept-source OCT* Verfahren am einfachsten zu durchschauen. In diesem Fall ändern wir die Wellenlänge der Lichtquelle in kleinen Schritten und nehmen jeweils das passende Interferenzsignal auf, das dann nachfolgend rückgerechnet wird.

Alternativ können wir aber die durchstimmbare Lichtquelle auch durch eine breitbandige Lichtquelle mit nachfolgender spektralen Aufspaltung, also einem Spektrometer ersetzen („spektral OCT"). Die einzelnen Photodioden der Spektrometerzeile sind somit eben Interferometerdetektoren für verschiedene Wellenlängen, die durch das Interferometer laufen.

Systeme im Ortsraum erklären sich zunächst einfacher über den Zugang der Kohärenz. Das Interferometer in Abb. 10.5.10 funktioniert wie besprochen nur, solange die Pfaddifferenz der beiden Pfade kleiner als die Kohärenzlänge der verwendeten Lichtquelle ist. Die *Kohärenzlänge* gibt letztlich direkt die Länge eines typischen emittierten Wellenpakets an. Andere Wellenpakete (ausgesendet von anderen Atomen) können zu diesem Wellenpaket eine beliebige Phasendifferenz haben, so dass die Interferenz zufällig ist und sich im Mittel weg hebt, d.h. der Interferenzterm in Gl. 10.7 wird im Mittel Null.

Die Kohärenzlänge berechnet sich aus der Bandbreite der Lichtquelle zu

$$L = \frac{\lambda^2}{\Delta\lambda}, \tag{10.10}$$

d.h. schmalbandige Lichtquellen (z.B. Laser) haben eine lange Kohärenzlänge während breitbandiges Licht (z.B. Halogenlampen) nur eine sehr kurze Kohärenzlänge (thermische Strahlen typischerweise wenige Mikrometer) haben.

Wir dürfen also bei unserer Interferenzerklärung (Addition der beiden Teilwellen) de-facto immer nur ein ausgesandtes Wellenpaket betrachten und natürlich kann dieses aufgespaltene und wieder vereinte Wellenpaket nur dann eine Interferenz generieren, wenn beide Teile mehr oder weniger gleichzeitig ankommen (vgl. Abb. 10.5.11 und 10.5.12).

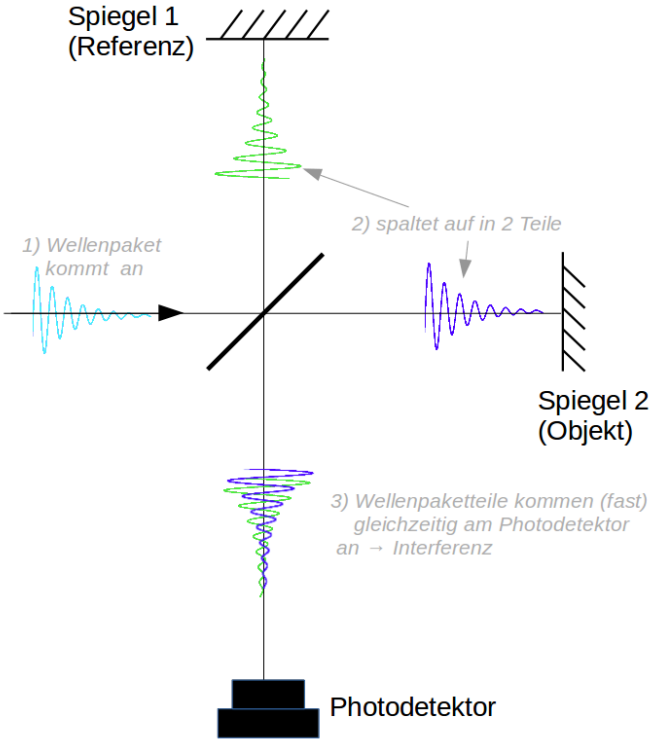


Abbildung 10.5.11: Weißlichinterferometrie bei gleichen Armlängen

Letztlich ergibt sich damit, dass bei starken Pfaddifferenzen die Interferenz verschwindet und wir haben eine Intensität gemäß Abb. 10.5.13.

Die einfachste (und stabilste) Form der Auswertung nutzt die Einhüllende dieses Signal. Hierbei wird das Maximum der Kohärenz gesucht, indem das Maximum

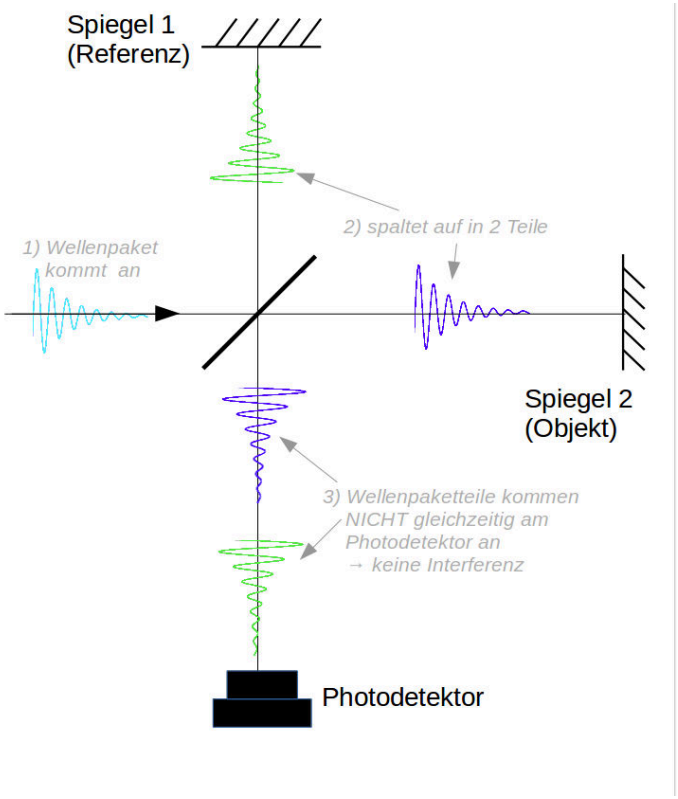


Abbildung 10.5.12: Weißlichinterferometrie bei ungleichen Armlängen. Eine Interferenz findet nicht statt.

des entsprechenden Signals (siehe gebildet wird. Die Einhüllende selbst ist genau über die Kohärenzlänge ausgedehnt.

Die Phasenauswertung versucht basierend auf dieser Position des Einhüllendenmaximums eine weitere feinere Positionierung durch Auswertung der Phase bzw. der Analyse der Feinstruktur des Interferenzmusters (Cosinus-Schwingung).

Die Messunsicherheit für die Topografiebestimmung liegt damit unter geeigneten Bedingungen im sub-lambda Bereich. Allerdings muss — unabhängig vom Messverfahren — in solchen Größenordnungen ganz generell mit allerlei möglicher Probleme gerechnet werden und auch die Kalibrierung bzw. Rückführung ist entsprechend schwer. Probleme können an Kanten, speziellen Geometrien, aber auch bei *benetzten Oberflächen* (z.B. *Ölfilme*) auftreten.

Zunächst ist aber ein großer Vorteil, dass auch *Schichten* messbar sind. D.h. es ergibt sich für jede Schicht zunächst eine Interferenz und bei bekannter Brechzahl des Schichtmaterials kann auf die Schichtdicke rückgeschlossen werden.

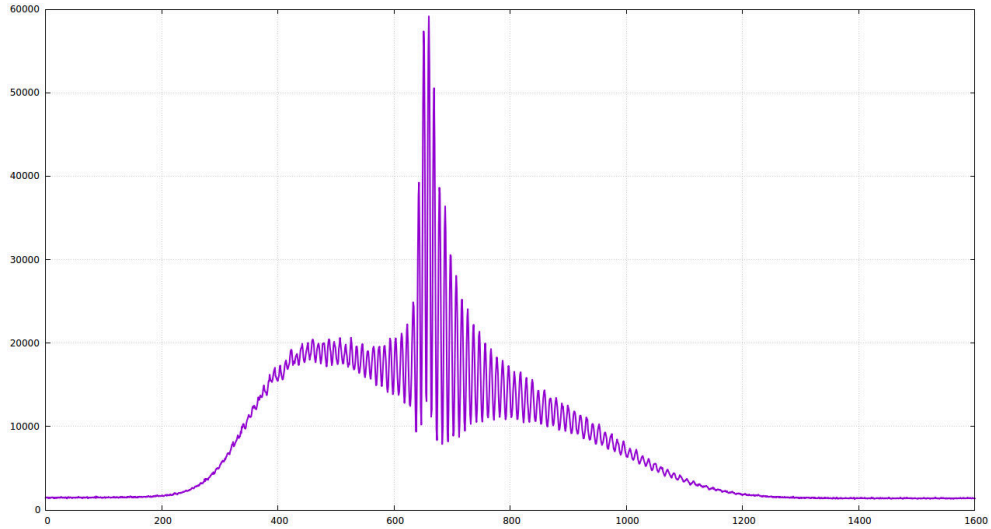


Abbildung 10.5.13: Weißlichtinterferometriesignal (mit zusätzlicher konfokalen Diskriminierung)

Unabhängig von diesen Basisprinzipien können unterschiedliche Interferometerkonfigurationen zum Einsatz kommen. Letztlich ist für das Grundprinzip nur die Zweistrahlinterferenz entscheidend. Wie die zwei Pfade zustande kommen ist zunächst egal. Hinsichtlich der technischen Umsetzbarkeit und den daraus sich ergebenden Unterschiede bestehen natürlich dann deutliche Unterschiede. Sehr hohe laterale Auflösungen lassen sich meist in der sogenannten *Linnik* Anordnung (letztlich eine Erweiterung des *Michelson* Interferometers mit speziell abgeglichenen Mikroskopobjektiven) erzielen. Industriell ist die *Mirau* Anordnung am beliebtesten (s. Abb. 10.5.14). Beide Pfade sind nahezu identisch so dass externe Einflüsse (Schwankungen von Temperatur oder Luftdruck, *Vibrationen*) einen eher geringen Einfluss haben. Die eigentliche Interferenz wird im zentralen Element, dem Mireauobjektiv, realisiert (Abb. 10.5.14).

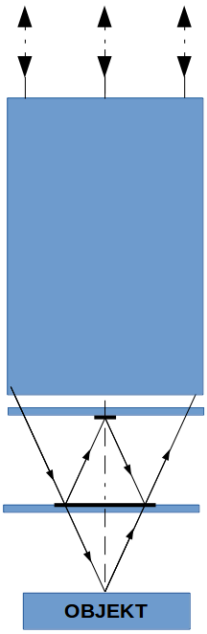


Abbildung 10.5.14: Nahezu common-path Geometrie (und damit hohe Stabilität) lässt sich mittels Mireau Objektiven erzielen.

+ Sehr geringe axiale Messunsicherheiten erzielbar
+ Messung durch Schichten oder Schichtbestimmungen
- Im Vergleich zur Triangulation deutlich mehr Rohdaten zu erfassen → langsamer
- Vibrationsanfällig[13]
- vergleichsweise teuer
- komplexe Effekt bei nichttrivialen Oberflächen

[13]Single-Shot Systeme (meist spektral realisiert) vermeiden diesen Nachteil, können aber dafür keine Flächen während einem Messvorgang detektieren.

10.6 Shape from (De)Fokus

Die Schärfentiefe der Abbildung ist begrenzt und somit können wir Entfernungen messen, indem wir einfach solange die Fokuslage ändern, bis das Bild scharf wird.

Gemäß Abschnitt 2.4 wird der Schärfentiefenbereich umso geringer, je kleiner die Blendenzahl bzw. je größer die numerische Apertur des Objektiv ist ($dz \sim \lambda/\mathrm{NA}^2$). Für mikroskopische Abbildungen mit hochaperturigen Objektiven erzielt man so Schärfentiefen unterhalb von 1 $\mu\mathrm{m}$ und damit sehr genaue Abstandsbestimmungen.

Wenn man nun eine gesamte Topografie bestimmen will, dann geht man folgendermaßen vor: Man verfährt in gleichbleibenden Abständen das Objekt axial durch den Schärfentiefenbereich und nimmt entsprechende zweidimensionale Bilder des Objekts auf. Bei der nachfolgenden Auswertung sucht man für jeden Oberflächenpunkt des Objekts das Bild, das die höchste Schärfe aufweist. Die entsprechende Bildnummer entspricht natürlich genau einer Position des Objekts bei der Aufnahme und damit einer bestimmten Höhe.

Zur Beurteilung der Bildschärfe sind viele verschiedene Algorithmen einsetzbar. „Scharf" bedeutet dabei, dass auch feine und feinste Strukturen mit hohem Kontrast vorhanden sind und dementsprechend kann man z.B. die Intensität des hochpassgefilterten Bilds verwenden. Zur Ermittlung des der Position (x, y) zugeordneten Höhenwerts $z(x, y)$ wird somit die lokale Nachbarschaft des Punkts entsprechend beurteilt.

Nur für objektseitig telezentrische Abbildungssysteme bleibt bei den einzelnen Bildern der Zusammenhang zwischen Objektpunkt (x, y) und Bildpunkt (x', y') für die verschiedenen Aufnahmen konstant. Will man also aufwendige Umrechnungen vermeiden, dann empfiehlt es sich dementsprechend, ein telezentrisches Abbildungssystem einzusetzen.

Ganz generell sind die Fokusverfahren vor allem für kleine Skalen interessant. Eine Skalierung auf große Geometrien ist schwierig bzw. erfordert vergleichsweise große Optiken. Die *Depth-from-Defocus* bzw. *shape from Defocus* Verfahren sind letztlich auf unterster Abstraktionsebene auch Triangulationsverfahren (genau wie das konfokale Mikroskop, vgl. Abschnitt 10.7). In der Praxis sind die Sy-

steme aber natürlich dennoch zu unterscheiden und haben ganz unterschiedliche Eigenschaften.

Die Ähnlichkeit zeigt sich auch bei der Unterscheidung von „aktiven" und „passiven" Systemen. Natürlich funktioniert die dargestellte Methodik ja nur, wenn das Objekt in der auflösungsbedingten Skala überhaupt genug feine Strukturen aufweist, um eine Schärfenbeurteilung zuzulassen. Je nach Objekt kann durch geeignete Wahl der Skale und der Abbildungsleistung eine Messung vorgenommen werden, d.h. die Abbildung muss im Zweifelsfall an das Objekt angepasst sein.

Aber auf einer komplett homogenen Fläche wird man keinen Unterschied zwischen fokussiertem und nicht-fokussiertem Bild erkennen und dementsprechend versagt dann die Methodik. Sobald wir aber (aktiv) ein feines Muster auf die Oberfläche projizieren, können wir wieder entsprechend arbeiten. Die Grenzen zur konfokalen Mikroskopie sind dabei fließend und Zwischenvarianten der beiden Grundtechniken sind möglich und werden z.B. in der *structured illumination microscopy* angewendet.

Generell haben wir folgende Vor- und Nachteile:

+ prinzipiell vergleichsweise einfacher Aufbau, aber hochgenaue Verschiebung oder Defokussierung (z.B. elektrisch schaltbare Linsen) ist notwendig.

+ Hohe Robustheit bei technischen Oberflächen in der Mikroskale

- Im Vergleich zur Triangulation deutlich mehr Rohdaten zu erfassen → langsamer

o Vorrangig im mikroskopischen Bereich einsetzbar.

10.7 Konfokale Mikroskopie

Die *konfokale Mikroskopie* kann (wenn auch nicht historisch) als Fortsetzung des Depth from Defocus betrachtet werden. Die Basis ist also wieder die defokussierte optische Abbildung.

Für einen Punktsensor wird ein Spot mittels mittlerer bis hoher numerischer Aper-

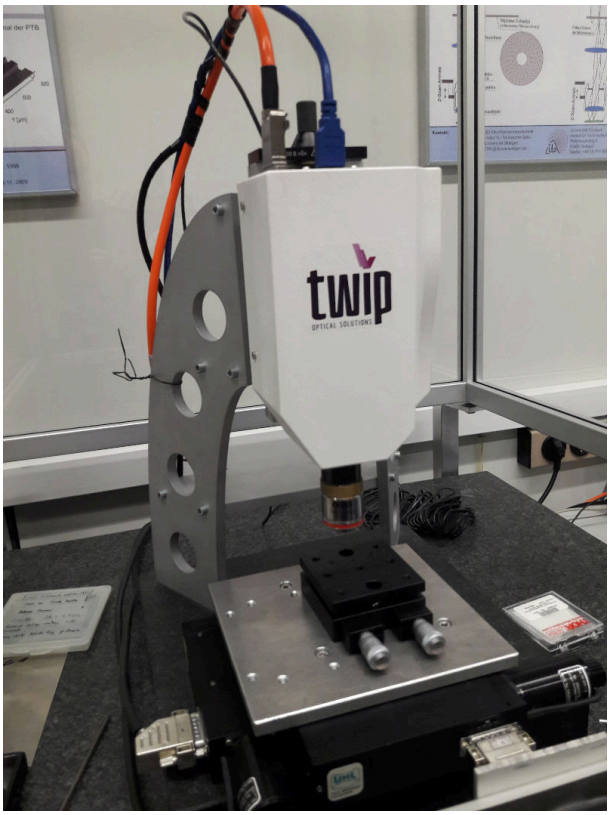


Abbildung 10.7.15: Kommerzielles Konfokalmikroskop zur Topografiemessung

tur in den Objektraum projiziert. Wenn die Objektoberfläche exakt an der korrekten axialen Position ist dann ergibt sich dort ein kleiner (idealerweise beugungsbegrenzter) Beleuchtungspunkt. Dieser Punkt wird durch dieselbe Optik auf einen Punktdetektor (oder einen flächigen Detektor mit vorgeschalteter Lochblende) abgebildet. Es ergibt sich am Detektor ein starkes Signal (s. Abb. 10.7.16).

Eine leichte axiale Fehlposition der Oberfläche führt zunächst zu einer deutlichen Verschmierung der Lichtintensität auf dem Objekt, d.h. die lokale Beleuchtungsstärke auf dem Objekt wird reduziert. Gleichzeitig wird das Licht, das nun vom eigentlich interessierenden Objektpunkt kommt auch noch unscharf abgebildet. Es kommt also abbildungsseitig wenig Licht durch die Lochblende. Insgesamt kommt also sowohl aufgrund der Beleuchtung wie auch aufgrund der Abbildung weniger Licht beim Detektor an.

Abb. 10.7.17 zeigt ein typisches konfokales Signal für einen Punktsensor. Der Schwerpunkt der Kurve[14] ist das Maß für die lokale axiale Position.

[14] auch andere, fitting-basierte Auswertungen sind natürlich möglich

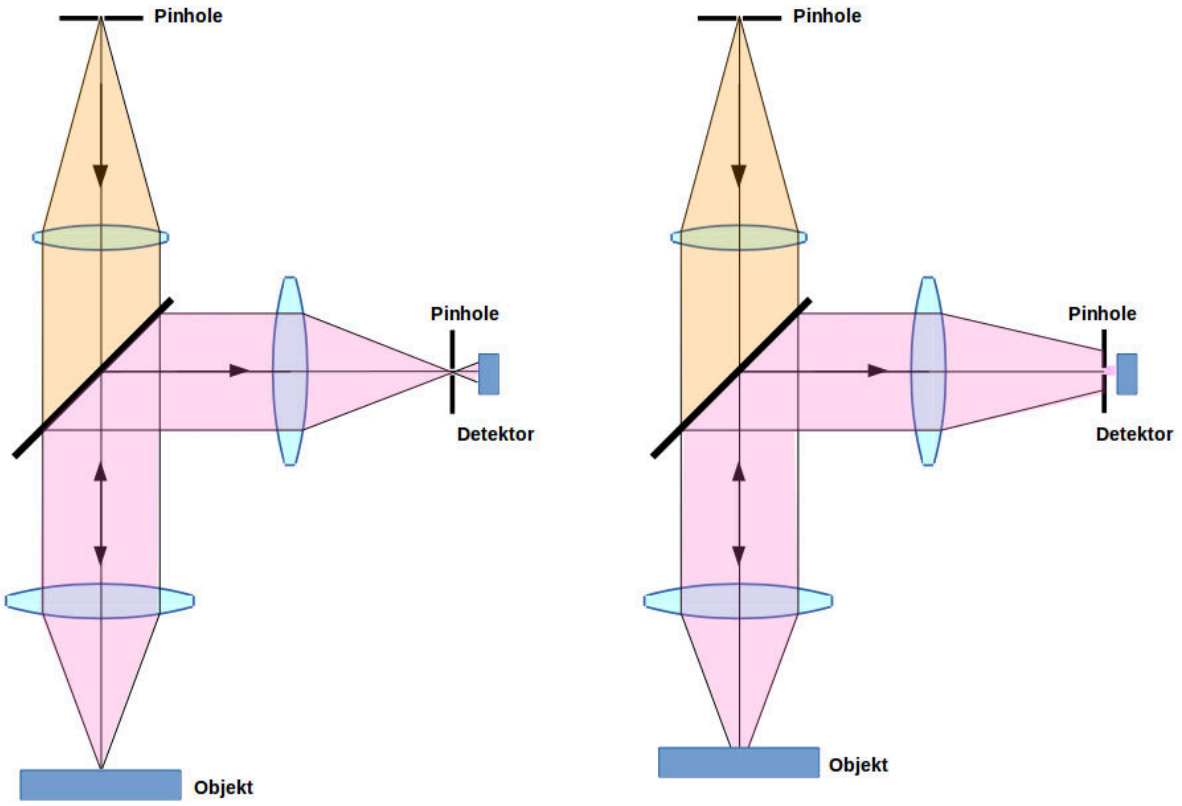


Abbildung 10.7.16: Prinzip der konfokalen Detektion: Wenn das Objekt exakt im Fokus ist (linke Seite), dann erreicht viel Licht durch das Pinhole den Detektor (Beleuchtungspinhole ist konjugiert zur Objektoberfläche und zum Beobachtungspinhole). Im defokussierten Fall ergibt sich eine geringe Intensität am Detektor. Das Objekt wird in axialer Richtung zum Mikroskop verfahren. Das detektierte Maximum gibt die Objekthöhe an.

Der mechanische axiale Scan lässt sich durch eine chromatische Detektion vermeiden. Entsprechende *chromatisch konfokale Sensoren* sind vergleichsweise gebräuchlich. Hierbei wird ein Objektiv mit starker chromatischer Längsaberration verwendet. Dementsprechend haben unterschiedliche Wellenlängen unterschiedliche Fokuslagen. Bei der Detektion wird dann also die axiale Position durch die Wellenlänge kodiert. Entweder man nutzt dementsprechend ein Spektrometer (oder aber eine Hyperspektralkamera) oder das Beleuchtungslicht wird hinsichtlich seiner Wellenlänge variiert (sogenanntes *swept source*).

Zur Messung ausgedehnter Oberflächen wird meist Scanning verwendet. Es sind aber auch Systeme mit rotierenden Lochblendenscheiben (Nipkow-Scheiben) oder Mikrolinsen (bessere Lichtausnutzung) vorhanden (siehe z.B. Abb. 10.7.15). In jedem Fall sind aber gegenüber triangulierenden Systemen wieder mehr Messdaten erforderlich.

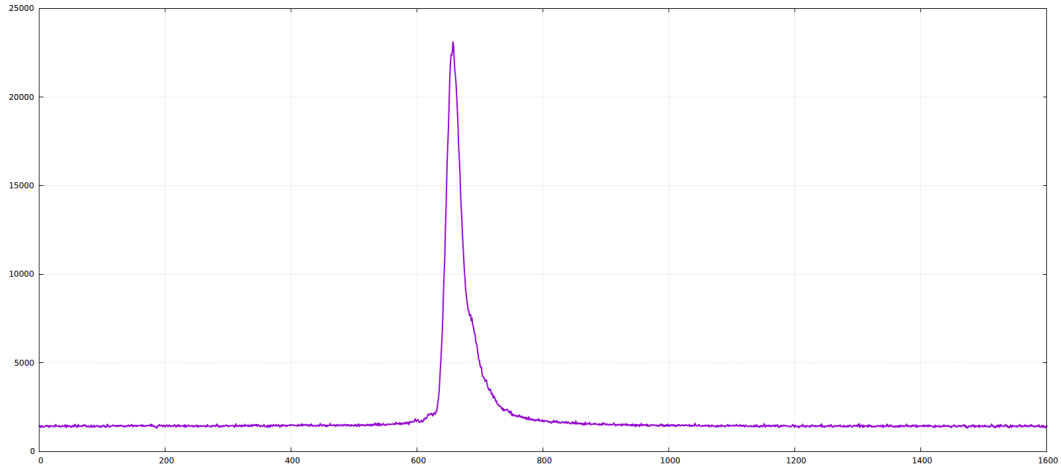


Abbildung 10.7.17: *Konfokales Signal: Helligkeit über axiale Position*

Konfokale Mikroskope werden vorrangig für andere Anwendungen, vor allem in der biomedizinischen Forschung, eingesetzt. In diesem Fall werden keine Topografien ermittelt sondern es kann bei hoher Auflösung die Streuung aus benachbarten Objektebenen stark reduziert werden.

Generell ergeben sich für die Topografiemessung folgende Vor- und Nachteile:

+ Hohe Robustheit bei technischen Oberflächen in der Mikroskale[15]

- Im Vergleich zur Triangulation deutlich mehr Rohdaten zu erfassen → langsamer

- Vergleichsweise teuer

o Vorrangig im mikroskopischen Bereich einsetzbar.

[15] Allerdings können bei bestimmten Topografien deutliche Messfehler auftreten.

10.8 Photometrisches Stereo (shape from shading)

Wenn eine Oberfläche aus einer bestimmten Richtung beleuchtet wird, dann streut bzw. reflektiert sie Licht unterschiedlicher Stärke in unterschiedliche Richtungen.

Natürlich ist das Streuverhalten abhängig von Material und Mikrotopografie der Oberfläche. Spezifiziert wird es durch die sogenannte Bi-Reflectance-Distribution-Function (*BRDF*). Diese gibt die Lichtstärke in einen bestimmten Winkelbereich (α, β) bei vorgegebenem Einfallswinkel (ϕ, θ) an (vgl. Abschnitt 7.5). BRDFs von verschiedenen Oberflächen lassen sich näherungsweise durch Modelle beschreiben, meist werden aber konkrete Messdaten zur Angabe der BRDF verwendet (vgl. Abschnitt 7.5).

Wenn die BRDF bekannt ist dann kann durch Beleuchtung der Oberfläche aus unterschiedlichen Richtungen mit Bildaufnahme aus einer (anderen) Richtung auf die lokale Oberflächenneigung des Objekts rückgerechnet werden. Aus der Neigung ergibt sich dann (letztlich durch Integration) die Oberflächentopografie.

Bei dieser Art der Messung kann eine Menge schief gehen und so haben hinsichtlich der der Messunsicherheit entsprechende Shape-from-Shading Systeme keinen guten Ruf. Die Empfindlichkeit gegenüber lokalen Änderungen der Topografie kann allerdings (wie auch beim Dunkelfeldverfahren) durchaus hoch sein.

Die Methode ist preiswert und schnell und dementsprechend eben sehr gut geeignet, wenn es nicht auf hochgenaue Topografiemessungen ankommt, sondern die Topografie als Merkmal genutzt werden soll.

+ Preiswert

+ Schnell

o Vorrangig nicht zur hochgenauen Messung sondern eher zur Nutzung der Topografie als Merkmal

- Messdaten fehleranfällig

10.9 Laufzeitsensorik

Die *Laufzeitmessung* nutzt das aus der Akustik gut bekannte Echolot-Prinzip. Anstatt akustischer werden elektromagnetische Wellen verwendet. Für sehr langwellige Strahlung nennt sich die Technik *Radar*, für sichtbares oder nah-infrarotes Licht *Lidar* („light detection and ranging"). Gebräuchlich ist vor allem auch die englische Bezeichnung *Time-of-flight sensing* und viele flächige arbeitende Sensoren werden auch (aufgrund des Prinzips) als *PMD-Sensoren* (Photonic Mixing Device) bezeichnet.

Ein Lichtpuls wird ausgesendet und vom Objekt zum Bildsensor zurück reflektiert. Die verstrichene Dauer bis zur Wiederankunft des Pulses beim Bildsensor wird detektiert. Mittels „Weg = Geschwindigkeit mal Zeit" wird auf den Weg rückgerechnet.

Grundsätzlich ergeben sich drei Möglichkeiten, die Laufzeitmessung technisch umzusetzen:

- Interferometrie (in allen Spielarten)
- Pulsbasierte Messung
- Intensitätsmodulation

Eine interferometrische Lösung haben wir bereist im Abschnitt 10.5 kennengelernt. Konventionelle Interferometer werden zur Messung von optisch glatten (spiegelnden) Oberflächen eingesetzt und sollen hier nicht weiter besprochen werden. Holografische Techniken basierend letztlich auf der Interferometrie, werden aber gesondert (Abschnitt 10.10) besprochen. Im folgenden konzentrieren wir uns daher auf die beiden anderen Varianten, bei denen ohne Interferenz die Entfernung gemessen wird.

Naheliegend ist die Verwendung von kurzen Lichtpulsen. In Verbindung mit einer CMOS-basierten Detektion lassen sich auch vergleichsweise hohe räumliche Auflösungen ohne Mehrdeutigkeitsprobleme erreichen. Nachteil ist aber die begrenzte Messunsicherheit der Entfernungsbestimmung.

Es ist aber ebenfalls möglich (und preisgünstiger realisierbar), die Intensität des

Lichts mit einem zusätzlichen Signal (üblicherweise ein Sinus) in der Amplitude zu modulieren. Bei der Detektion wird dann die Phase dieses aufmodulierten Signal bestimmt und mit der Referenzphase verglichen. Die sich so ergebende Phase ist dann direkt proportional zur Entfernung.

In der industriellen Messtechnik wird die Technik bisher kaum eingesetzt, eher in der Robotik[16]. Dies liegt an der begrenzten Genauigkeit im typischerweise mm-Bereich, die die Systeme erreichen. Ein Hauptproblem ist bei ausgedehnten Szenen genug reflektiertes Licht von den Objekten zur Detektion aufzufangen. Insbesondere bei dunklen oder das Licht weg reflektierenden Flächen ergeben sich Probleme. Entsprechend viel Licht sollte in den Puls gelegt werden, was die Beleuchtungsquelle teuer macht.

Flächig arbeitende Systeme mit vielen Pixeln leiden letztlich bei großen Abständen noch stärker unter der mangelnden Lichteffizienz.

Für scannende, auf Lasern basierende Systeme können ohne Probleme auch sehr große Entfernungen vermessen werden. Hier sind die Sensoren anderen Techniken in verschiedener Hinsicht überlegen. Die direkte Konkurrenztechnologie, die Triangulation, würde für große Entfernungen von z.B. 50 m bereits eine vergleichsweise große Stereobasis benötigen was zu einem großen Sensor und praktischen Problemen bei der Kalibrierung führt. Andererseits gilt aber natürlich, dass für eine hohe Punktdichte bei guter Auflösung und großen Entfernungen für das Lidarprinzip sehr viel Licht in den Raum abgegeben werden muss. Entsprechend ergeben sich zwangsläufig Beschränkungen aufgrund der Lasersicherheit und deutliche Nachteile gegenüber der rein passiven Stereovision (Triangulation).

Bei Szenen mit großer (Helligkeits-)Dyanmik können Probleme durch Streulicht und ungewollte Reflexe innerhalb der Abbildungsoptik (Ghosts, Flare) auftreten.

Als Vor- und Nachteile ergeben sich dementsprechend für die Laufzeitsensoren:

+ preiswert

+ keine Abschattungsproblematik

[16] Außerhalb der industriellen Produktion werden entsprechende Sensoren auch für Fahrerassistenzsysteme bzw. autonomes Fahren, die militärische Aufklärung und viele Anwendungen zur Untersuchung der Atmosphäre eingesetzt.

+ kleiner Sensor, gut für die Integration
+ Entfernungsmessung auch auf große Entfernung möglich.
- geringe Genauigkeit axial (einige mm)
- begrenzte lateral Auflösung
- starke gepulste Lichtquelle notwendig.

10.10 Digitale Holografie, Shearografie und ESPI

Eng verwandt mit der Interferometrie ist die Holografie (letztlich kann die konventionelle Interferometrie als ein Spezialfall der Bildebenenholografie (s.u.) aufgefasst werden). Ganz offensichtlich ist in einem Hologramm dreidimensionale Information gespeichert (in der Phase des Lichtfelds) und diese kann zur Messung genutzt werden.

Man lässt bei der Holografie ganz allgemein zwei Wellen interferieren. Die eine Welle wird üblicherweise Objektwelle o, die zweite "Referenzwelle" r genannt. Interferenz bedeutet, dass beide Wellen addiert werden. Sowohl klassische Filme wie auch elektronische oder biologische Photodetektoren reagieren auf die Intensität des Lichts, die sich proportional zum Quadrat des elektrischen Felds ergibt. Für das Hologramm haben wir somit

$$H(x,y) = |r(x,y) + o(x,y)|^2 = |r|^2 + |o|^2 + ro^* + or^* \tag{10.11}$$

Dabei sind wir nun zu einer komplexen Schreibweise für r und o übergegangen (dadurch wird die Mathematik einfacher). Man beschreibt also ein periodisches elektrisches Feld $E = A\cos(\phi)$ durch den Realteil von $A\exp(i\phi)$, also $E = A\cos(\phi) = \mathrm{Re}(A\exp(i\phi))$.

Wenn man hier an einem Punkt (x,y) im Raum für r und o wieder Kosinusschwingungen ansetzen, dann ergibt sich in diesem (bzw. jedem) Punkt Gl. 10.11. Die Phasendifferenz von r und o ist in der Intensität gespeichert. Und zwar für jeden Punkt. Holografie ist also eine Sonderform der Interferometrie, bei der raue Oberflächen (und damit einhergehend eine komplizierte Objektwelle) erfasst werden.

Zur Rekonstruktion des so aufgezeichneten Hologramms wird wieder mit der Referenzwelle beleuchtet. Das Licht "geht durch" das Hologramm und dementsprechend wird das Lichtfeld mit der Transmission des Hologramms multipliziert:

$$\begin{aligned} rH &= r(|r|^2 + |o|^2 + ro^* + or^*) && (10.12) \\ &= r|r|^2 + r|o|^2 + rro^* + |r|^2 o, && (10.13) \end{aligned}$$

Eines dieser Lichtfelder (letzter Term) ist dabei aber direkt gleich dem ursprünglich aufgezeichneten Objektfeld o, denn $|r|^2$ ist für eine typische Referenzwelle (z.B. ebene oder sphärische Welle) konstant.[17] Somit haben wir eine Kopie des Objektlichtfelds hergestellt. Im Vergleich zu einer Fotografie ist im Hologramm die komplette Information des Lichtfelds gespeichert.

Im Strahlenmodell bedeutet das, dass sowohl die lokale Lichtintensität wie auch die Strahlrichtung im Hologramm gespeichert ist (bei einem Foto ist dagegen nur die Intensität gespeichert). Ein Betrachter, der im rekonstruierten Lichtfeld steht, bemerkt keinen Unterschied zu dem Lichtfeld, das durch das Orginalobjekt zustande kommen würde.

Nun kann man natürlich das Hologramm auch — anstatt auf klassischen Film — direkt auf einem Bildaufnehmer aufzeichnen und die Rekonstruktion des Hologramms im Computer rein digital simulieren. Man spricht dann von *digitaler Holografie*. Insbesondere kann man dann bei der Simulation die Rekonstruktion in verschiedenen Ebenen berechnen und so mit nur einem Hologramm ein gesamtes Volumen aufzeichnen bzw. rekonstruieren.

Besonders interessant ist dabei die Möglichkeit, nicht nur die Amplitude (bzw. Intensität) sondern auch die Phase des Lichtfelds rekonstruieren zu können. Damit eignen sich die digitalholografischen Verfahren auch für die Messung von Phasenstrukturen oder auch die Detektion von sehr feinen axialen Verschiebungen (besonders prominent natürlich bei der Interferometrie). In ihrer Grundform ist die digitale Holografie ein Einzelbildverfahren und somit sehr schnell.[19]

Trotz der Eleganz und der Vorteile ist die Methode allerdings nicht immer sinnvoll einsetzbar. Für gute Auflösungen bei großen Feldern sind sehr viele Hologramm-

[17] Durch Einfügen geeigneter Winkel[18] kann man dafür sorgen, dass die anderen Terme sich nicht mit dem gewünschten Lichtfeld o überlagern.

[19] auch wenn das Ergebnis durch Mehrbilderweiterungen (z.B. mit Phasenschiebung) verbessert werden kann.

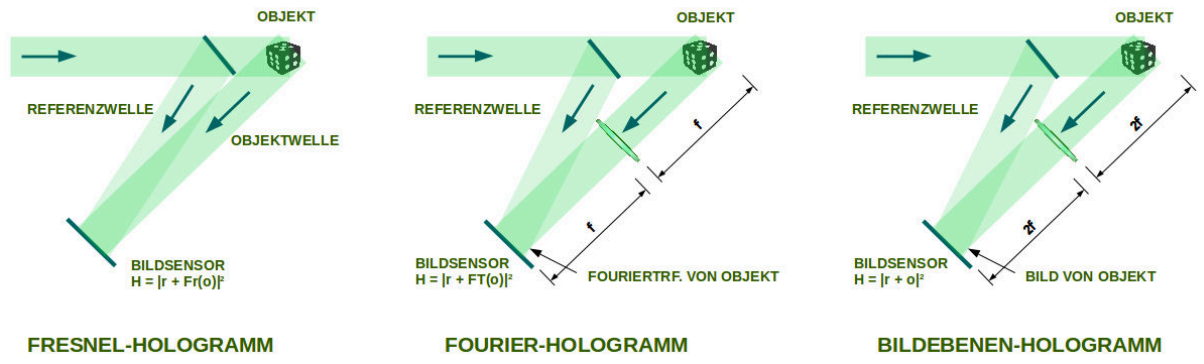


Abbildung 10.10.18: *Geometrien für die (digitale) Holografie: Aufgezeichnet wird das Lichtfeld am Ort des Hologramms. Dieses Lichtfeld kann das vom Objekt propagierte Lichtfeld (Fresnelhologramm), die Fouriertransformierte des Objektes (Fourierhologramm) oder ein Bild des Objekts (Bildebenenhologramm) sein.*

pixel nötig. Vor allem sind die Rekonstruktionen bei rauen Objektoberflächen aber verspeckelt, also stark verrauscht. Eine Phasenauswertung macht bei der rauen Oberfläche zunächst keinen Sinn, denn die Rauheit führt zu einer zufälligen Phase, die sich im Hologramm eben als Speckles zeigt. Deformationen als Differenz von zwei Messungen (mit natürlich derselben Mikrotopografie) lassen sich so aber gut erfassen.[20] Man spricht auch von *Speckle Interferometrie* oder *ESPI* (electronic speckle pattern interferometry) bzw. *DSPI* (digital speckle pattern interferometry) . Erfolgen die Hologrammaufnahmen mit geringem zeitlichen Abstand lassen sich somit natürlich auch sehr gut Schwingungen detektieren. Oft wird das Hologramm als Bildebenenhologramm aufgezeichnet, d.h. die Oberfläche wird auf den Bildaufnehmer abgebildet, vgl. Abb. 10.10.18.

Man kann das Problem der Rauheit und der Phase auch anders betrachten. Letztlich ist der geringe Eindeutigkeitsbereich der interferometrischen Messung die Ursache der Probleme. Wieder kann der Eindeutigkeitsbereich erhöht werden indem zwei (oder mehr) Hologramme mit unterschiedlichen Wellenlängen aufgenommen und ausgewertet werden. Damit steigt der Eindeutigkeitsbereich wieder mit der sogenannten synthetischen Wellenlänge (vgl. Abschnitt 10.5). Alternativ kann man auch Hologramme mit Beleuchtung aus leicht unterschiedlichen Winkeln verwenden oder die Wellenlänge künstlich erhöhen, indem das Licht unter

[20]Die Empfindlichkeit gegenüber Deformationen hängt vom konkreten Versuchsaufbau ab. So lassen sich sowohl sogenannte in-plane wie auch *out-of-plane* Deformationen messen.

einem großen Einfallswinkel auf die Oberfläche fällt.

Aufgezeichnet werden kann sowohl das Objektfeld nach freier Propagation durch den Raum (*Fresnelhologramm*), das abgebildete Objektlichtfeld (*Bildebenenhologramm*) oder die Fouriertransformation des Objekts (*Fourierhologramm*), siehe Abb. 10.10.18. Je nach Wahl von Beleuchtungs- und Beobachtungsgeometrie kann man Deformationen in unterschiedlichen Empfindlichkeitsrichtungen erfassen (siehe z.B. [34]).

Von *holografischer Interferometrie* spricht man, wenn man ein Hologramm des Objekts aufnimmt, rekonstruiert und die rekonstruierte Lichtwelle mit der ursprünglichen Objektlichtwelle interferieren lässt. Änderungen des Objekts werden dann direkt in Interferenzstreifen sichtbar.[21] Alternativ kann man auch zwei Hologramme (in kurzer Folge) des Objekts im selben Hologramm aufnehmen (*Doppelpuls-Holografie*).

Bei der *Shearografie* oder *Shearinginterferometrie* wählt man als Referenz das verschobene Objektlichtfeld. Das Hologramm ist somit $H = |o(x)+o(x+\Delta x)|^2$. Damit wird man auf den Gradient des Objekts sensitiv. In der Mikroskopie wird das Verfahren bei der Implementierung mit Abbildung (also Bildebenenholografie) dann *Differentieller Interferenzkontrast* (*DIC*) genannt.[22]

Die *konoskopische Holografie* verfolgt eine ähnliche Idee. Auch hier wird das Lichtfeld mit einer Kopie seiner selbst zur Interferenz gebracht. Die Kopie ist dabei gegenüber dem Original (polarisationsoptisch) defokussiert.[46]

10.11 Computertomographie

Die *Computertomographie* (CT) nimmt eine Sonderrolle in der „optischen" 3D-Messtechnik ein. Aufgrund der extrem kurzen Wellenlänge (Röntgenlicht) können viele Objekte durchleuchtet werden und es können so letztlich echte Volumendaten, z.B. Kavitäten im Inneren von Objekten, erfasst werden. Auch wenn die Objekte keine inneren, unzugänglichen Strukturen aufweisen, kann dieses Verfahren

[21] Eine lokale Änderung der Objektentfernung um $\lambda/2$ führt zu einem Streifen.

[22] Die Realisierung erfolgt dann meist nicht über ein klassisches Interferometer sondern polarisationsoptisch (siehe z.B. [39]).

die Objekterfassung verbessern, denn komplexe Objektgeometrien lassen sich mit herkömmlichen optischen Methoden oft nur schwer erfassen.

Grundprinzip der Methode ist die Aufnahme von zweidimensionalen Durchstrahlungsbildern unter vielen unterschiedlichen Richtungen. Hierzu wird meist das zu vermessende Objekt gedreht. Aus diesen Projektionen kann auf das Innere des Objekts, z.B. mittels des berühmten Fourier-Slice Theorems, rückgerechnet werden.

Nachteile des Verfahrens sind die vergleichsweise hohen Kosten und der hohe Zeitbedarf. Für einige Anwendungen ist die Methodik aber die einzige Möglichkeit zur Qualitätsprüfung.

Bildsensoren für Röntgenlicht arbeiten mit einer Szntillatorschicht in der die hochenergetischen Röntgenphotonen in sichtbares Licht gewandelt wird. Das sichtbare Licht wird dann meist durch eine faseroptische Bildleitung (fiber-optic plate) zum Bildsensor transportiert. Dabei kann auch ein Abbildungsmaßstab (z.B. 1:2) verwendet werden, um die Szintillatorschicht an den Bildsensor anzupassen. Aufgrund dem Mangel ans Abbildungsoptiken werden Objekte bei Röntgenanendungen klassischerweise durchleuchtet, das Objekt muss also kleiner als die Szintillatorschicht (z.B. Cäsiumiodid) sein (und außerdem wird natürlich eine Röntgenquelle benötigt).

+ Echte volumetrische Daten (nicht nur Oberfläche)

+ Erfassung von Defekten im Volumen

- hohe Kosten

- aufwändige Messung

10.12 Andere Entfernungssensoren

In diesem Abschnitt sollen einige weitere optische Methoden zur Entfernungsbestimmung kurz im Sinne des allgemeinen Überblicks angesprochen werden. In der Regel finden die Sensoren beim Stand der Technik entweder nur selten oder aber nur für sehr spezielle Aufgaben Einsatz in der industriellen Messtechnik.

Radiometrische Entfernungssensoren (bzw. *Intensitätsbasierte Entfernungssensoren*) können sehr genau berührungslos die Antastung an ein Objekt, also den Abstand zu einem Objekt detektieren. Der Arbeitsabstand muss hierzu allerdings sehr gering (eben typische Anwendung Antastung) sein, meist im Millimeterbereich. Für kleine Abstände lassen sich sehr hohe Auflösungen im Bereich von Mikrometern erzielen.

Eine Oberfläche wird beleuchtet (z.B. mittels einer Faser). Das an der Oberfläche gestreute Licht wird dann von einer Optik (meist ebenfalls eine Faser) aufgefangen und die aufgefangene Lichtmenge nimmt quadratisch mit dem Abstand ab.

Aufgrund der Einfachheit des Messprinzips können sehr preiswerte und schnelle Sensoren gebaut werden. Natürlich ist das Prinzip aber anfällig gegenüber Verschmutzungen oder Änderungen der Oberflächenrauheit aber auch Verkippung. Eine genaue Kalibrierung ist in jedem Fall für jeden Objekttyp notwendig.

Self-mixing Interferometrie ist eine interessant Alternative zur klassischen Interferometrie bei der genutzt wird, dass in den Laser rückgestreutes Licht das emittierte Licht beeinflusst. Letztlich hat man nicht nur im Laser selbst einen Resonator sondern auch das Messobjekt selbst bildet mit dem Laser einen Resonator und dementsprechend ändert sich das Verhalten des Lasers wenn sich dieser externe Resonator ändert.

Im einfachsten Fall kann hierzu einfach die (ohnehin oft im Laser vorhandene) Referenzphotodiode zur Analyse verwendet werden. Diese misst die durch das Objekt entstehende Leistungsschwankung.

Digitale Holografie kann letztlich als Interferometrie an einer rauen Oberfläche betrachtet werden. Dabei kann die Oberfläche auf den Sensor abgebildet werden (Bildebenenhologramm) oder aber das aufgezeichnete Lichtfeld ist über die Fresneltransformation (Fresnelhologramm) oder Fouriertransformation (Fourierhologramm) mit dem Objekt verbunden.

In jedem Fall kann im Rechner die Rekonstruktion des Hologramms simuliert werden. Und zwar für verschiedene Entfernungen, so dass damit letztlich eine dreidimensionale Rekonstruktion möglich wird. Problematisch sind die sich in der Rekonstruktion ergebenden Speckles und die vergleichsweise hohen Anforderungen an das Ortsbandbreitenprodukt (Anzahl Pixel des Hologramms bzw.

Bildsensors), um bei sinnvollem Signal-Rausch Verhältnis ausgedehnte Objekte mit guter Genauigkeit messen zu können.

Besonders geeignet ist die Technik zur Untersuchung von Deformationen. Gebräuchlich ist hier für die Bildebenengeometrie die Bezeichnung *Electronic Speckle Pattern Interferometry*. Zur dreidimensionalen Messung werden meist Hologramme mit verschiedenen Wellenlängen (analog zur Heterodyninterferometrie) oder unter verschiedenen Winkeln vorgenommen.

Kapitel 11

Spezifikation

Die Beauftragung sowie Realisierung einer Bildverarbeitungslösung ist in aller Regel ein sehr anspruchsvolles und fehlerträchtiges Unterfangen. Wesentlich ist, dass sich der Kunde zunächst klar macht, was er genau benötigt. Diese Aufgabe ist alles andere als trivial und bedarf in aller Regel der Unterstützung von externen Fachleuten (im einfachsten Fall dem potentiellen Anbieter der Bildverarbeitung).

Fehler, die in der Spezifikationsphase gemacht werden führen zu überproportional teuren Änderungen im späteren Projektverlauf oder können gar einen sinnvollen Einsatz unmöglich machen. Es ist also bei der Spezifikation der Aufgabe äußerste Sorgfalt notwendig. Eine generelle Einführung aus Kundensicht mit wertvollen Hinweisen zur Vorgehensweise findet man in [4].

In diesem Kapitel werden Sie unter anderem lernen

- wie die Anforderung an ein Bildverarbeitungssystem zu spezifizieren ist,
- was die Grundbegriffe der Messtechnik bedeuten,
- wie die Leistung von klassifizierenden Systemen beurteilt wird,
- wie Defekte zu spezifizieren sind,
- was bei der Abnahme zu beachten ist.

Bildverarbeitungssysteme sind in aller Regel keine Produkte „von der Stange" sondern werden für eine ganz spezielle Anwendung realisiert. Daher muss bei der

Kommunikation und Vertragsgestaltung zwischen Auftraggeber und Auftragnehmer (also dem Anbieter der Bildverarbeitungslösung) klar definiert werden, welche Aufgabe wie gelöst werden soll und wie letztlich entschieden wird, ob das realisierte System letztlich die Anforderungen erfüllt. Grundsätzlich werden diese Festlegungen in zwei zentralen Dokumenten vorgenommen:

Lastenheft: Hier definiert der Auftraggeber möglichst exakt, welche Aufgaben das System (aktuell und zukünftig) zu erfüllen hat. Alle wesentlichen Nebenbedingungen werden genannt. Es sollte auch der bisherige Stand innerhalb der Fertigung beschrieben sein. Es werden also der Ist-Zustand sowie der Soll-Zustand definiert. Unter Umständen muss das Dokument zusammen mit einem unabhängigen Fachmann oder aber unter Mithilfe des Anbieters erstellt werden.

Das Lastenheft kann an verschiedene mögliche Auftragnehmer gegeben werden. Angebote basierend auf dem Lastenheft lassen sich so vergleichen.

Pflichtenheft: Basierend auf dem Lastenheft generiert der Anbieter ein Pflichtenheft indem die Aufgabe aus Sicht des Anbieters dargestellt wird und definiert wird, wie die Aufgabe prinzipiell gelöst werden soll. Hierbei werden in der Regel zentrale Komponenten des Systems bereits spezifiziert (z.B. Kameras, Objektive usw.).

Sinn beider Dokumente ist es, zum einen vertragliche Klarheit zu schaffen, zum anderen — und das ist an sich zentral — soll sichergestellt werden, dass der Anbieter die Aufgabenstellung und Wünsche des Auftraggebers voll verstanden hat und dass dem Auftraggeber mögliche Beschränkungen klar sind.

In der VDI Richtlinie 2632 sind Details zur Ausarbeitung von Lasten- und Pflichtenheften für die industrielle Bildverarbeitung angegeben. Generell ist aber zu sagen, dass eine Vielzahl an — oft anwendungs- bzw. branchenabhängigen — Richtlinien und Normen definieren, welche Parameter wie anzugeben sind. Ohne Anspruch auf Vollständigkeit wollen wir einige zentrale Punkte, die spezifiziert werden sollten, angeben:

- Grundaufgabenstellung (1D, 2D oder 3D, Klassifikation, dimensionelles Messen, Lageerkennung, Identifikation, Vollständigkeitsprüfung, Zählung, Code-

Lesen, Soll-Ist Vergleich, usw.)

- Zusätzliche Angaben zur Aufgabenstellung (z.B. 100% Kontrolle oder Stichprobenkontrolle, inline, offline oder at-line, Laborbetrieb, Experimentalsystem, mobil, stationär)
- Bei messenden Systemen: Angaben der zu messenden Größen und zur Messunsicherheit oder ähnlicher Angaben (s.u.), evtl. getrennt für verschiedene Objektbereiche
- Gesamtszene (Objektfeld, Hintergrund)
- Ziel und bisherige Vorgehensweise
- Angaben zu möglichen Störmerkmalen
- Objekt
  - Objektgeometrie
  - mögliche Verdeckung
  - Objektposition und Orientierung (Lage- und Orientierungstoleranzen)
  - Zuführung (z.B. Fließband, Roboter, Angabe von Toleranzen), Bewegungszustand (insbesondere Variabilität der Geschwindigkeit), Vibrationen
  - Oberflächenbeschaffenheit und Material (matt, glänzend, BRDF, strukturiert, Rauheit, wellig, homogen, bedruckt usw., Farbe, Transluzenz)
  - Verschmutzung (partikuläre, filmische Verschmutzung, Fasern, Staub, Öl, Reinigungsrückstände usw.)
  - verschiedene Objektklassen, Variabilität der Objekte und ihrer Oberflächen[1].
- Defektspezifikation
- Defektstatistik bei aktueller Produktion
- Platzangebot und Arbeitsabstand, sonstige geometrische Nebenbedingungen
- Angaben zum Durchsatz und andere zeitliche Anforderungen, z.B. *Zykluszeit*.

[1] z.B. durch Alterung oder Produktvielfalt

- Angaben zur Fehlerfindungsstatistik (Detektionswahrscheinlichkeit, Falschklassifikationswahrscheinlichkeit, Falschalarmrate usw.)
- Umgebungsbedingungen, z.B.
  - Fabrikhalle, Messraum usw.
  - Fremdlicht
  - Elektromagnetische Verträglichkeit (EMV)
  - Vibrationen
  - Luftfeuchtigkeit
  - Reinraum
  - Temperatur(schwankung)
  - Staub, Ruß, Öl, Gase usw.
- Angaben zu Schnittstellen (Qualitätsmanagement, Prozesssteuerung, Datenbanken, Statistik, Mensch-Maschine Schnittstelle, Fernwartung).
- Angaben zur Inbetriebnahme
- Angaben zur Medienversorgung (elektrische Leistung, Gase, Kühlwasser usw.)
- Angaben zur Abnahme
- Ausfallwahrscheinlichkeit des Systems (Mean-time between failure, Umrüstbarkeit, Mean-time to Repair) bzw. Verfügbarkeit
- Schulung und Dokumentation

11.1 Musterkatalog

Wie zuverlässig eine automatisierte Beurteilung per Bildverarbeitung funktioniert hängt in erheblichem Ausmaß natürlich von den konkreten Objekten und Defekten, die erkannt werden sollen, ab. Oftmals ist es daher notwendig, genügend repräsentative Testobjekte mit genügend repräsentativen Fehlern oder Merkmalen zur Verfügung zu haben. Hierzu dienen sogenannte *Musterkataloge*.

Idealerweise sollten für alle zu detektierende Defekte auf allen relevanten Bauteilen genügend Musterteile mit entsprechenden Defekten zur Verfügung stehen. Dies führt aber schnell zu einer kombinatorischen Explosion an notwendigen Musterteilen und insbesondere bei den Defekten ist es oft mehr als schwierig, Grenzdefekte für die Defektspezifikationen zur Verfügung zu haben. Beispielsweise kann spezifiziert sein, dass Kratzer mit einer Mindestbreite von 10 μm und einer Mindestlänge von 20 μm zu detektieren sind. Um das System zu testen, müssen entsprechende Bauteile mit Defekten in dieser Größenordnung vorliegen. Wenn die Defekte in der Produktion selten sind oder die Produktion eventuell noch nicht läuft, dann müssen entsprechende Defekte künstlich hergestellt werden und auf einem Referenzsystem vermessen werden.

Weiterhin ist zu beachten, dass sich wesentliche Parameter der Objekte im Laufe der Zeit ändern können. Dies kann sich sowohl auf die produzierten Teile (z.B. Modifikation der Objekte für neue Serie, geänderte Produktionsbedingungen) als auch auf die Teile des Musterkatalogs, die im Laufe der Zeit altern (z.B. oxidieren oder sich deformieren), beziehen.

Der Musterkatalog ist daher ein wichtiges (aber auch problematisches) Instrument, das zur Realisierung (und oft zur Abnahme) einer Bildverarbeitungslösung verwendet wird.

11.2 Messen, Prüfen und Rückführbarkeit

Sehr viele Systeme in der industriellen BV werden zur Prüfung von Objekten eingesetzt. Beim Prüfen wird untersucht, ob ein Objekt die an das Objekt gestellten Anforderungen erfüllt oder nicht. Oft können diese Anforderungen durch *Messung* überprüft werden, teilweise sind aber auch nicht-messende Anforderungen relevant (z.B. Vollständigkeitsprüfung). Begrifflich ist die Messung also eine Möglichkeit, eine Prüfung durchzuführen und ob das bildverarbeitend messende Systeme seine Messung für eine Anwendung „ausreichend" ordentlich vollzieht kann als „Messfähigkeit" bezeichnet werden.

Die Prüfung selbst kann man im übrigen auch als Klassifikationsaufgaben (z.B. in die beiden Klassen „Teil gut" bzw. „in Ordnung (IO)" oder „Teil schlecht" bzw.

„Nicht in Ordnung (NIO)") ansehen. Dazu später mehr. Wenn die Prüfung auf einer Messung beruht, dann muss aber natürlich zunächst klar sein, wie aussagekräftig das Messergebnis überhaupt ist. Jede Messung hat eine Unsicherheit und ohne eine Angabe dieser Unsicherheit ist das Messergebnis in aller Regel wertlos.

Wesentlich sind hierbei zufällige und systematische Fehler, aber auch die Fehler aufgrund der Rückführbarkeit (z.B. Unsicherheiten der Kalibrierkörper). Die Gesamtheit dieser Fehler ergibt letztlich die sogenannte „Messunsicherheit". Diese ist für uns der zentrale Begriff zur Beurteilung des möglichen Fehlers einer Messung.

Bevor wir uns im Detail mit der Messunsicherheit beschäftigen sollen die für uns wesentlichen Begriffe der Messtechnik hier kurz dargestellt werden. Details und Definitionen finden sich in DIN 1319.[2]

Prüfung: Vergleich mit einer Forderung. z.B. kann gefordert werden, dass der Messwert innerhalb eines bestimmten Intervalls liegt. Allerdings ist Prüfung nicht auf Messergebnisse beschränkt sondern kann z.B. das Vorhandensein von Merkmalen „prüfen".

Auflösung: Kleinste Differenz der zu messenden Größe, die durch die Messung eindeutig (mit einer gewissen Wahrscheinlichkeit[3]) unterschieden werden kann.[4] Beachten Sie: Der Begriff der Auflösung wird in verschiedenen Kontexten, insbesondere der Optik, sehr unterschiedlich verwendet.

Messabweichung: Differenz von Messwert zu wahrem Wert

Messfehler: frühere/traditionelle Bezeichnung für Messabweichung

Messunsicherheit: Gibt zusammen mit dem Messwert ein Intervall an, indem der wahre Wert mit „hoher Wahrscheinlichkeit" liegt.[5] Umgekehrt kann angenommen werden, dass bei einer Messung mit hoher Wahrscheinlichkeit ein Messwert im Intervall der Messunsicherheit um den wahren Wert erhalten

[2] Einen guten Überblick findet man in [23]

[3] z.B. 95 %

[4] Z.B. kann hierfür bei der Annahme einer Normalverteilung eine wiederholte Messung zur Bestimmung der Standardabweichung durchgeführt werden.

[5] Ein quantitativer Vergleich verschiedener Messunsicherheiten macht daher nur Sinn, wenn dieselben statistischen Annahmen (z.B. 95% Vertrauensintervall) gemacht werden. Oft erfolgen die Angaben in Bezug zur Standardabweichung, also z.B. „3σ".

wird. Teilweise wird die Messunsicherheit auch als relative Messunsicherheit, z.B. in Prozent, angegeben.

Maximum Permissible Error (MPE): Vor allem im angloamerikanischen Raum gebräuchlich. Gibt eine garantierte Grenze für den Messfehler an.

Genauigkeit: Ist ein nur qualitativ zu verwendender Begriff für die Messunsicherheit. Bei quantitativen Angaben sollte man also von Messunsicherheiten sprechen. Genauigkeit wird teilweise synonym mit Messunsicherheit verwendet.

Systematischer Fehler: Abweichung des Erwartungswerts vieler Messungen vom wahren Wert. Statistische Fehler sind also nicht enthalten.

Richtigkeit: Der Messwert kann zwar durchaus statistisch stark streuen, im Mittel ergibt sich aber der wahre Werte der zu messenden Größe.

Wiederholbarkeit (Repeatability): Am selben Bauteil wird vom selben System unter selben Bedingungen wiederholt gemessen[6].

Reproduzierbarkeit (Reproducibility): Die *Reproduzierbarkeit* bezieht sich auf dieselbe Situation wie die der Wiederholbarkeit. Allerdings kann ein anderer Prüfer, die Messung/Prüfung durchführen oder aber das Messobjekt wird neu zugeführt. Im Gegensatz zur Wiederholbarkeit wird also der *gesamte* Messvorgang neu gestartet.

Wiederholstandardabweichung: Standardabweichung der Messwerte unter Wiederholbedingungen.

Präzision (precision): Der Messwert kann zwar durchaus falsch sein, aber bei wiederholter Messung ergibt sich immer derselbe (unter Umständen falsche) Messwert. Statistische Fehler sind also bei einer präzisen Messung klein.

Abb. 11.2.1 verdeutlicht den wichtigen Unterschied zwischen Präzision und Genauigkeit.

Bei all diesen Definition, die zunächst sehr objektiv wirken, muss immer beachtet werden, dass immer noch eine genaue Definition der Bedingungen zur Ermittlung

[6]Entscheidend ist aber oft, welche Zeitdauer zwischen den Messungen liegt bzw. ob sich Nebenbedingungen geändert haben könnten. So ist z.B. für eine ausreichen „Warmlaufzeit" zu sorgen

der Kenngrößen vorhanden sein muss. Teilweise ist diese in Anwendungsnormen und Richtlinien geregelt, teilweise definieren die Hersteller von Messgeräten sie selbst. So macht es natürlich erhebliche Unterschiede, ob z.B. die Messung einer Topografie auf eher kooperativen oder unkooperativen Oberfläche erfolgt, ob die Topografie eben oder kompliziert deformiert ist etc.

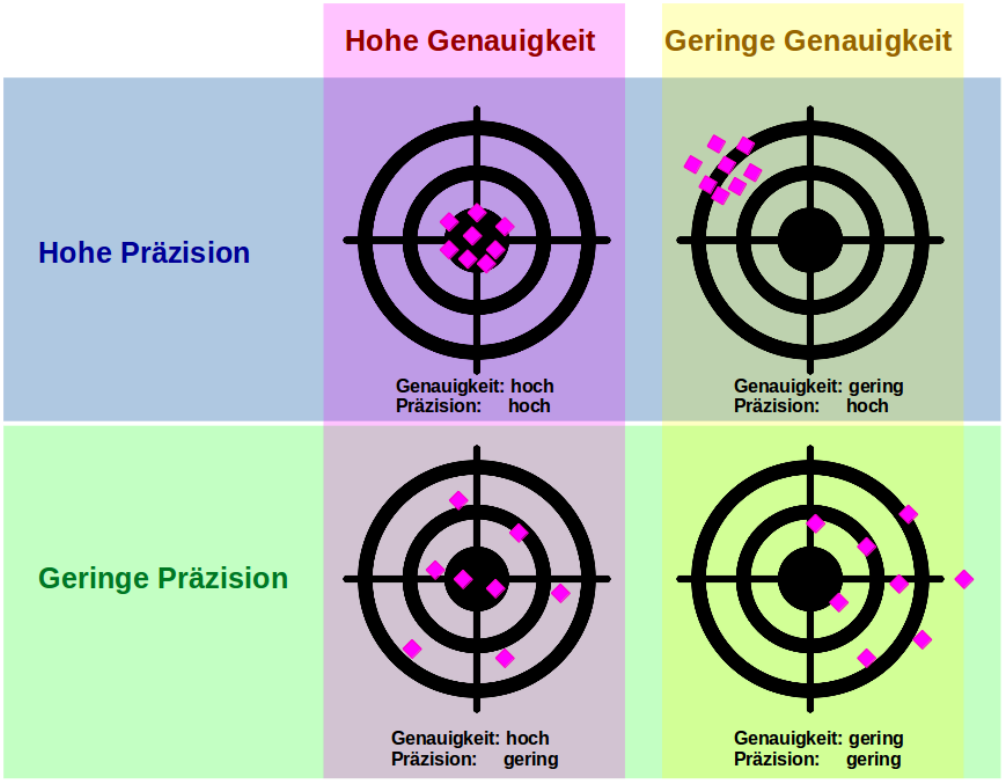


Abbildung 11.2.1: Zwischen Genauigkeit (bzw. Messunsicherheit als klar definiertem Begriff) und Präzision muss unterschieden werden. Eine Messung kann z.B. ungenau aber präzise sein (typischer Fall bei starkem systematischem Fehler).

Natürlich kann man die Wiederholmessungen unter ganz unterschiedlichen Bedingungen (z.B. gleicher Bediener aber unterschiedlicher Ort, unterschiedliches Messgerät etc.) durchführen und entsprechend gibt es verschiedene Begrifflichkeiten (z.B. *Vergleichspräzision*), die hier eine genauere Kennzeichnung erlauben.

11.2.1 Messunsicherheit

Wie bereits dargestellt setzt sich die *Messunsicherheit* aus systematischen und statistischen Fehlern sowie Fehlern der Rückführbarkeit zusammen.

Gehen wir zunächst von dem Fall aus, dass wir den wahren Wert mit extrem hoher Genauigkeit kennen. Wir können dann den systematischen vom statistischen Fehler trennen indem wir sehr oft die Messung wiederholen (idealerweise unendlich oft). In diesem Fall wäre die Differenz von wahrem Wert und dem gemittelten Messwert dann ein systematischer Fehler und kalibrierbar.

Wenn wir nicht unendlich oft messen, dann können wir bei derselben Berechnung aber natürlich nicht sicher sein, ob wir wirklich komplett systematische von statistischen Fehlern getrennt haben. Schließlich gibt es ja vermutlich eine (kleine) Differenz von Mittelwert bei unendlich vielen Messungen zu Mittelwert bei z.B. „nur" 100 Messungen.

In der Praxis werden wir den wahren Wert einer Größe also nicht unbedingt exakt kennen und daher bezieht man sich zur Kennzeichnung der Abweichungen oft eben auf den „Schätzwert der wahren Größe." Dabei geht man dann aber davon aus, dass etwaige systematische Fehler bereits eliminiert sind. Verbleibende systematische Fehleranteile werden dann (zumindest in der GUM, s.u.) wie statistische Fehler behandelt.

Das vollständige Messergebnis wird angegeben als Messwert $\pm$ Messunsicherheit (bei einer bestimmten Vertrauensangabe, s.u.). Der zentrale Begriff bei der Spezifikation einer Messabweichung ist also heute die Messunsicherheit. Ihre Bestimmung bzw. Spezifikation ist exakt im *Guide of Uncertainty of Measurements* (GUM) geregelt [11].[7] Allerdings ist sie als rein formale Methode durchaus nicht unproblematisch und beinhaltet insbesondere die Annahme der Linearisierung und setzt für verbleibende systematische Fehler eben auch das Gaußsche Fehlerfortpflanzungsgesetz an. Dies führt dann zu einer zu positiven Bewertung der Unsicherheit (vgl. [25] S. 576).[8]

Beachten Sie unbedingt, dass dabei die systematischen Fehler bereits wegkalibriert sind und wir gedanklich nur noch von statistischen Fehlern ausgehen; auch wenn ein Teil dieser Fehler noch nicht-erfassbare systematische Ursprünge haben mögen. Diese Restanteile an systematischen Fehlern sind dann schlicht für uns zu

[7] Es existieren durchaus leicht unterschiedliche Definitionen der Messunsicherheit, die sich allerdings nicht wirklich widersprechen, sondern eher verschiedene Aspekte betonen. Die traditionellen Definitionen gehen dabei von den an sich nicht bekannten wahren Werten von Messgrößen aus.

[8] Wenn es wirklich darauf ankommt, dann lohnt es sich, ein komplettes Modell der Einflussgrößen und Ungenauigkeiten aufzustellen und per Monte-Carlo Simulation zu simulieren.

komplex zu erfassen und somit auszugleichen.

Die Bestimmung der Messunsicherheit für eine Größe, die in komplexer Weise von anderen (mit Fehler behafteten) Größen abhängt wird schnell ein relativ aufwändiges und fehleranfälliges Unterfangen. In [25] sind einige gute Beispiele für entsprechende Rechnungen angegeben. Wie so oft steckt auch hier der Teufel im Detail.

Wenn die Messunsicherheit bestimmt ist, dann gibt man sie am klarsten als $\pm$ an. Erfolgt eine Einzelwertangabe, dann bezieht sich diese in der Regel auf das halbe Gesamtintervall der Unsicherheit.[9]

Die Messunsicherheit selbst bezieht sich immer auf ein Vertrauensintervall. Wir betrachten hier rein statistische Fehler, gehen also davon aus, dass etwaige systematische Fehler durch Kalibrierung eliminiert oder aber zumindest bekannt sind.

Wenn wir die zu betrachtende Messung unendlich oft wiederholen würden, dann würde sich ein mittlerer Messwert x_m (Erwartungswert) ergeben und diesen Wert nehmen wir als wahren Wert der zu messenden Größe an. Einzelmessungen x_i streuen um diesen mittleren Wert und üblicherweise wird die (empirische) Standardabweichung zur Kennzeichnung des Streubereichs verwendet:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} |x_i - x_m|^2}{N-1}} \tag{11.1}$$

Bei einer zugrundeliegenden Gaußverteilung wird ein einzelner Messwert x_i mit einer Wahrscheinlichkeit von 68% im Intervall $x_i - \sigma \ldots x_i + \sigma$ liegen. Mit 95,5% Wahrscheinlichkeit liegt er im Intervall $x_i - 2\sigma \ldots x_i + 2\sigma$.[10]

D.h. die Angabe eines Messwerts in der Form $a \pm b$ @ 2 σ bedeutet, dass der wahre Wert mit einer Wahrscheinlichkeit von 95,5% im Bereich $a - b \ldots a + b$ liegt.[11] Wenn keine explizite Angabe des Vetrauensbereichs erfolgt, dann ist meist von 1 σ, also von einer Wahrscheinlichkeit von 68% auszugehen.

[9] Eine Messunsicherheit von 1 mm bedeutet also, dass sich mit 68% Wahrscheinlichkeit das Messgerät einen Wert liefert, der im Bereich ± 1 mm (also Gesamtbereich: 2 mm) vom wahren Wert abweicht.

[10] Für kleine N ist die Unsicherheit der Schätzung der Standardabweichung selbst allerdings hoch und man muss dann ggf. nochmals mit einem Erweiterungsfaktor, dem t-Faktor (aus der Studentverteilung) multiplizieren (siehe [25]).

[11] $3\,\sigma \rightarrow 99.7\%$

Wenn vom Wert 1 σ abgewichen wird (sehr gebräuchlich ist in der Praxis 2 σ $\rightarrow$ 95,5%) spricht man auch von „erweiterter Messunsicherheit".[12].

Oft wird diese Erweiterung auch im Formelbezeichner angegeben. Für die Messunsicherheit schreibt man so in Tabellenwerken U, U_{95} oder U_{99}, um Messunsicherheiten mit den Erweiterungsfaktoren 1,2 und 3 anzugeben (der Index gibt also das Vertrauensintervall für die Messung an).

Ebenfalls zu bedenken ist, dass sich der letztlich für den Kunden relevante Fehler nicht nur aus dem Messgerät selbst sondern auch aus dem Fehler der *Rückführung* σ_R und anderen möglichen Fehlerquellen (z.B. Umgebungseinflüssen) ergibt. Für diese gesamte Messunsicherheit σ_e wird dann (im Sinne von unabhängigen Fehlern) quadratisch gemittelt:

$$\sigma_e = \sqrt{\sigma^2 + \sigma_R^2 + \sigma_T^2 + ...} \tag{11.2}$$

In der Bildverarbeitung tritt vergleichsweise oft der Fall auf, dass sich eine zu messende Größe als Mittelwert von Einzelmessungen ergibt. Ein typisches Beispiel ist der Durchmesser einer Bohrung. Letztlich können hier N Einzeldistanzen zu einem mittleren Durchmesser kombiniert werden. Es stellt sich also die Frage, um wieviel sich die Streuung des Mittelwerts durch die Mittelung verringert. Wenn die Fehler der Einzelmessungen unabhängig voneinander und zufällig sind (keine systematischen Fehler), dann wird sich der Fehler bei N Einzelmessungen gerade wieder um $\sqrt{N}$ verringern. Dies ist soweit auch anschaulich, denn nach dem Gaußschen Fehlerfortpflanzungsgesetz haben wir für die den Fehler einer Addition von N Termen (Mittelwertbildung $x_g = x_1/N + x_2/N + ... + x_N/N$:

$$\Delta x_g = \sqrt{|\frac{\Delta x_1}{N}|^2 + |\frac{\Delta x_2}{N}|^2 + ... + |\frac{\Delta x_N}{N}|^2} = \sqrt{N \cdot \frac{\Delta x_1^2}{N^2}} = \frac{\Delta x_1}{\sqrt{N}} \tag{11.3}$$

Wenn die Anzahl der Mittelungen nicht sehr groß (etwa > 50) ist, dann ist das Ergebnis noch mit einem sogenannten „Vertrauensfaktor" t_{N-1} zu multiplizieren. Für das Vertrauensintervall von 68% gilt z.B. $t_4 = 1.14$, $t_9 = 1.06$, $t_{99} = 1.01$.[13] Für größere Vertrauensintervall gelten aber andere Werte, insbesondere für 95%:

[12] In diesem Beispiel ist der „Erweiterungsfaktor" 2.

[13] Die Korrektur hat letztlich ihren Ursprung in dem Unterschied zwischen Gaußscher Normalverteilung und der eigentlich zu zugrundeliegenden „Student-Verteilung".

$t_4 = 2.8$, $t_9 = 2.3$, $t_{99} = 2.01$

Es gilt also bei Mittelwertbildungen

$$\boxed{\sigma_N = t_{N-1}\frac{\sigma}{\sqrt{N}} \approx \frac{\sigma}{\sqrt{N}} \qquad (11.4)}$$

Wir bleiben nochmals beim Fehlerfortpflanzungsgesetz. Wenn sich eine für uns relevante Größe y aus Messwerten x_i durch eine Formel ergibt, also $y := f(x_1, x_2, ...x_N)$ dann stellt sich natürlich ebenfalls die Frage, wie sich die erwartete Unsicherheit von y aus der Unsicherheit der x_i ergibt. Man spricht von „Messunsicherheit einer indirekten Größe" bzw. „kombinierter Messunsicherheit".

Bei exakter Kenntnisse der Fehler bzw. Abweichungen δx_i der x_i vom wahren Wert $x_{i,m}$ würde man natürlich schreiben

$$\delta y = \frac{\partial f}{\partial x_1}\delta x_1 + \frac{\partial f}{\partial x_2}\delta x_2 + ... + \frac{\partial f}{\partial x_N}\delta x_N \qquad (11.5)$$

Das ist auch der Ansatz, wenn wir von systematischen Fehlern ausgehen [25].

Wenn die Fehler aber unabhängig voneinander sind, dann erwartet man, dass sie sich teilweise gegeneinander wegheben und man verwendet das Gaußsche Fehlerfortpflanzungsgesetz

$$\boxed{\delta y = \sqrt{|\frac{\partial f}{\partial x_1}\delta x_1|^2 + |\frac{\partial f}{\partial x_2}\delta x_2|^2 + ... + |\frac{\partial f}{\partial x_N}\delta x_N|^2} \qquad (11.6)}$$

Nach der GUM Methode A werden die δx_i statistisch bestimmt (Standardabweichung). Bei der GUM Methode B erhält man sie auf anderem Wege, z.B. durch eine Herstellerangabe eines Messgeräts oder aber durch ein ausreichend genaues Modell des Messvorgangs bzw. der variablen Größe x_i. Z.B. könnte eines der x_i für eine Temperatur stehen und es ist in der Anwendung (warum auch immer) klar, dass diese Temperatur um maximal ± 1 K variieren kann. In diesem Fall wäre dann das passende δx_i eben als 1 K anzunehmen.

Eine Kalibrierung[14] ist an sich für alle messenden Anwendungen notwendig, denn in irgendeiner Form muss eine Umrechnung von Pixeln, z.B. nach metrischen Größen vorgenommen werden.Bei kamerabasierten Systemen muss so eigentlich immer eine Korrektur von Verzeichnungen erfolgen. In diesem Fall ist die Umrechnung auch noch positionsabhängig.

Bei der Kalibrierung wird die Ausgabe des BV-Systems auf eine Maßgröße, z.B. Meter, „rückgeführt". Die *Rückführung* erfolgt in der Praxis mit einem *Referenzmesssystem*, dessen Messunsicherheit bekannt ist, oder mittels *Kalibrierobjekten*, deren Toleranzen bekannt sind. Nur wenn eine entsprechende Rückführung durchgeführt wird, ist eine Angabe von Gesamtmessunsicherheiten sinnvoll. Als Faustregel fordert man in der Praxis oft für die Unsicherheit der Referenz maximal ein Zehntel der letztlich in der Anwendung notwendigen Messunsicherheit. Zu bevorzugen ist allerdings eine exakte Betrachtung gemäß GUM.

In aller Regel besteht die Rückführung darin, dass geeignete kalibrierte Testobjekte (z.B. Kalibriertargets) vermessen werden. Für zweidimensionale geometrische Kalibrierungen werden also Maßstäbe benötigt, die kalibriert sind. Entsprechende Masken (in der Regel Chrom auf Glas) werden in vielfältiger Form (Punkte, Linien, Kreis, Strichmarken) und Größe kommerziell angeboten und sind unerlässlich. Auch für die Helligkeitskalibrierung (Linearität der Bildaufnahme, radiometrische Kalibrierung) gibt es Kalibrierobjekte, bei denen unterschiedliche Felder exakt definierte optische Abschwächungen aufweisen.

Bei Systemen, die nicht direkt in der Produktion bzw. inline eingesetzt werden, kann die Kalibrierung vergleichsweise oft und regelmäßig durchgeführt werden. In der Produktion selbst ist die Kalibrierung oft mit erhelblichen Kosten verbunden, so dass es wesentlich ist, nicht zu oft kalibrieren zu müssen.

Welche Messunsicherheit ist nun aber überhaupt zu fordern. Dies ist natürlich stark anwendungsabhängig aber einen einfachen Überblick zeigt Abb. 11.2.2. Letztlich werden Faustregeln angewendet, mit denen entschieden wird, ob eine Messung zur Sicherstellung einer gewissen Bauteiltoleranz „fähig" ist. Eine typische Faustregel würde z.B. aussagen, dass $0.15\,T/6\sigma > 1.3$ ist, Mit 6σ ist dann die Messunsicherheit für ein Vertrauensintervall von 6σ anzunehmen und T ist die er-

[14] Teilweise wird auch von „Eichung" geredet. Dieser Begriff sollte aber nur bei amtlichen Kalibrierungen verwendet werden.

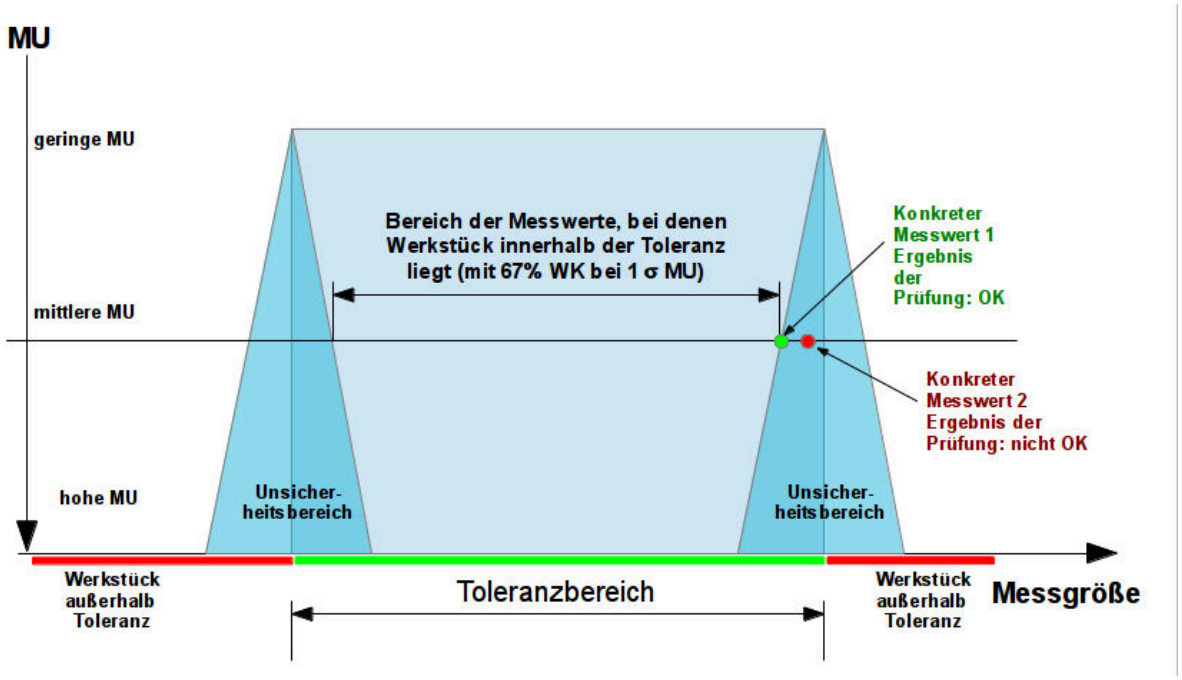


Abbildung 11.2.2: Zusammenhang zwischen Messunsicherheit und erlaubten Toleranzen bei der Bestimmung einer Größe. Zu den gemessenen Werten muss jeweils die Messunsicherheit addiert (bzw. subtrahiert) werden, um zu entscheiden, ob die wahre Größe mit hoher Wahrscheinlichkeit noch innerhalb der Toleranzen liegt. Darstellung modifiziert nach [13].

laubte Toleranz. Wenn also z.B. die Messunsicherheit für die Messung einer Länge 10 μm (bei 1 σ) beträgt, müssen wir von 60 μm ausgehen. Dementsprechend wäre das Messsystem (unter dieser willkürlichen Definition !) nur sinnvoll im Sinne einer Messprozessfähigkeit einsetzbar wenn $T > 1.3 \cdot 60/0.15 = 520\ \mu\mathrm{m}$.

Die Anforderung mit 6σ ist auch deutlich härter als eine andere, sehr bekannte heuristische Regel, die *Goldene Regel der Messtechnik* nach der die zu prüfende Toleranz mindestens 10 mal der Messunsicherheit entspricht.

Teilweise wird die grobe Messunsicherheit über die Anzahl signifikanter Stellen einer Angabe gemacht. D.h. die angegebenen Stellen sind in etwa als vertrauenswürdig zu sehen. Dementsprechend würde man bei einer Angabe von 1,00 m davon ausgehen, dass der Wert in etwa auf einen cm genau angegeben ist. Eine sinnvollere und exaktere Angabe wäre x = 1,000 m $\pm$ 0.011 m.

Zu beachten ist, dass die Anzahl der angegebenen Stellen des Werts und der Unsicherheit übereinstimmen sollten. Eine noch exaktere Angabe in der Art x = 1,00000 m $\pm$ 0.01123 m macht — außer bei der Messung von Naturkonstanten — in der

Praxis in aller Regel keinen Sinn, denn es ist völlig ausreichend die Unsicherheit auf ca. 10% der Unsicherheit genau zu kennen.

Neben der eigentlichen Angabe der Messunsicherheit muss immer hinterfragt werden, unter welchen Bedingungen — insbesondere mit welchen Objekten — diese Messunsicherheiten denn ermittelt wurden, wie und wie oft kalibriert werden muss bzw. wie die Rückführung dauerhaft sichergestellt wird und welche Nebenbedingungen eingehalten werden müssen.

Wir sind bisher davon ausgegangen, dass die sich ergebenden Fehler einer Normalverteilung genügen. Das muss natürlich nicht immer der Fall sein. Insbesondere bei sehr seltenen Fehlern können auch gänzlich andere Verteilungsfunktionen ins Spiel kommen. Entsprechend schwieriger wird dann die mathematische Behandlung.

Vor allem die Abschätzung von Abhängigkeiten bei Fehlermodellen wird sehr schnell kompliziert wenn mehrere Fehlereinflüsse in einem bestimmten mathematischen Modell zu einer Beeinflussung der Zielgröße führen.

Am besten (weil straight forward) betrachtet man solche Abhängigkeiten anhand einer Monte-Carlo Simulation. D.h. man generiert viele Systeme mit Parametervariation und simuliert die Ausgangsgröße. Ein typisches Beispiel aus der Optik ist die Beurteilung von fertigungsbedingten Bildfehlern eines Objektivs aufgrund von Toleranzen der Linsen und ihrer Einbauparameter. Man kann in diesem Fall nicht einfach die Auswirkungen der Einzelfehler (z.B. Radientoleranz der ersten und der zweiten Fläche) quadratisch gemittelt addieren sondern es besteht eine (oft starke) — in diesem Fall nichtlineare — Abhängigkeit.

Ein schönes Beispiel dafür, dass selbst bei einfachsten Messungen man nicht einfach die Unsicherheiten addieren kann gibt Hennes an [23]. Nehmen Sie an, Sie messen mit zwei Meterstäben ein z.B. 2 m langes Regal. Hierzu legen Sie die Meterstäbe hintereinander. Dabei machen Sie einen Fehler wenn die beiden Meterstäbe nicht exakt parallel zueinander orientiert sind (vgl. Abb. 11.2.3).

Wenn Sie nun davon ausgehen, dass Sie eine Normalverteilung um die Längensumme der beiden Meterstäbe als Fehler annehmen, dann wäre ihre Länge also 2 m $\pm$ x mm.

Nun ist es aber so, dass bei Fehlausrichtung die Länge nur kleiner 2 m werden kann (nicht größer!). Die Angabe würde also keinen Sinn machen. Stattdessen müssten Sie von einer einseitigen Fehlerverteilung (Angabe eher im Sinn 2 m - x mm) ausgehen.

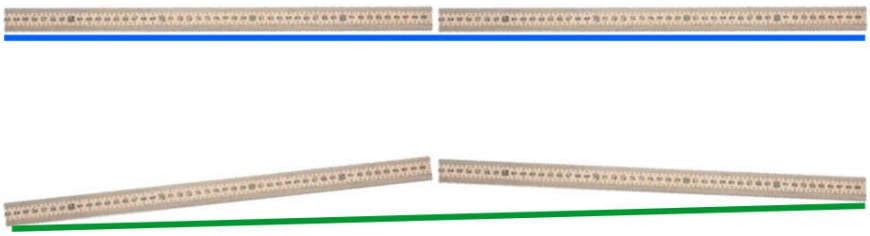


Abbildung 11.2.3: Bei manchen Aufgabenstellungen sind die Fehler einseitig. Bei der Kombination von zwei Meterstäben kann die Gesamtlänge bei Anschluss Ende 1 zu Ende 2 (grün dargestellt) nicht länger als die Summe der beiden Meterstäbe werden (blau dargestellt).

11.3 Spezifikation von Defekten

Letztlich kann alles, was in irgendeiner Form zu einem Objekt gehört fehlerhaft sein. Es besteht also eine immense Vielfalt an Eigenschaften, die von Bildverarbeitungssystemen geprüft werden müssen. Dennoch sind in vielen Fällen die selben Merkmale entscheidend und daher existieren für wesentliche praxisrelevante Prüfmerkmale verschiedene Normen.

Von besonderer Bedeutung ist die sogenannte „Geometrische Produktspezifikation" (GPS) nach EN ISO 25178 und EN ISO 8785, die sehr ausführlich auf verschiedene Aspekte der Definition geometrischer Spezifikationen von Objekten und ihrer Fehler eingeht. Wichtige Grundgrößen sind

- Maße
- Abstände
- Radien
- Winkel und Richtungen
- Formen

- Profile
- Oberflächenunvollkommenheiten
- Oberflächenbeschaffenheiten
- Mikrostrukturen (u.a. Rauheit)

Oberflächenunvollkommenheiten sind für die optische Prüfung natürlich von besonderem Interesse, denn direkt kann ein optischer Zugriff nur auf Oberflächenmerkmale erfolgen.

Entscheidend sind Länge, Breite, und Höhe/Tiefe der jeweiligen Oberflächenunvollkommenheit sowie die Anzahl entsprechender Defekte. Teilweise kann es auch relevant sein, wie die Defekte auf die zu prüfende Oberfläche verteilt sind (z.B. Porennest). Direkt abgeleitet sind die Fläche und das Volumen des Defekts.

Wie wollen nicht auf die einzelnen spezifizierten Defekte (z.B. Riefe, Kratzer, Lunker, Riss, Beule usw.) eingehen. In der Regel kennt der Auftraggeber geeignete Defektmodelle bzw. Beschreibungen der für ihn relevanten Defekte.

11.4 Prüfen und Klassifizieren

Für klassifizierende Systeme sind andere Parameter relevant. Allerdings ist zu beachten, dass die zur Klassifikation verwendeten Größen (z.B. Kratzerlänge) oft aus Messungen abgeleitete Größen sind und so auch die Messunsicherheiten bei vielen klassifizieren Systemen von zentraler Bedeutung sind.

Wenn ein Bildverarbeitungssystem zur Prüfung von Eigenschaften von Objekten eingesetzt werden soll, dann sind in der Regel klare Vorgaben hinsichtlich der Sicherstellung dieser Prüfung notwendig. Das BV-System ist in diesem Fall ein *Prüfmittel* und muss Anforderungen an die *Prüfmittelfähigkeit* erfüllen. Letztlich bedeutet das, dass die messenden Systembestandteile sinnvoll und regelmäßig kalibriert sind und klar ist (vgl. Rückführbarkeit), wie die wesentlichen Parameter (z.B. Messunsicherheit) sind.

Was jeweils konkret wie nachgewiesen werden muss variiert von Anwendung zu Anwendung und Hersteller zu Hersteller. Letztlich muss sich das BV-System also an den Forderungen des Auftraggebers orientieren. Aufgrund der Vielfalt an Anwendungen der industriellen Bildverarbeitung ist es wenig sinnvoll, allgemeine Richtlinien zur konkreten Definition des Nachweis der Prüfmittelfähigkeit anzugeben.

11.5 Statistik, Klassifikation und Prüfen

In sehr vielen Bildverarbeitungssystemen wird eine Klassifikation von den zu untersuchenden Objekten vorgenommen. Diese Klassifikation kann in ganz unterschiedliche Klassen erfolgen. Typische Beispiele sind Gutteil/Schlechtteil, verschiedene Modelltypen oder Zu-Reinigen/Nicht-Reinigen oder ein Sortieren (z.B. nach Farbe).

Oftmals basiert die Klassifikation indirekt auf Messungen, z.B. wenn eine Geometrien kontrolliert werden oder wenn Kratzer z.B. eine bestimmte maximale Länge aufweisen dürfen. Für die Messungen selbst gelten die bereits dargestellten Unsicherheitsangaben.

Bei typischen Objekterkennungsaufgaben, also z.B. die Prüfung, ob ein gewisses Sicherheitsmerkmal korrekt auf dem Objekt vorhanden ist, macht eine Messunsicherheit aber keinen Sinn. Stattdessen ist man an der Klassifikationsleistung des Systems interessiert.

Diese Klassifikationsleistung wird durch verschiedene Kenngrößen beschrieben. Ganz generell können wir die Leistung des Systems in einer Tabelle wie in Abb. 11.5 dargestellt beschreiben. Hierbei wird für jede Klasse i angegeben, in wie vielen Fällen eine korrekte *Klassifikation* (Klasse i) erfolgt und in wie vielen Fällen die Klassifikation in eine Falsche Klasse j erfolgt.

	Teil iO	Teil NiO
Detektion iO	Korrekt	Schlupf
Detektion NiO	Fehlausschuss	Korrekt

In Produktionsanwendungen ist sicher die Klassifikation in gute und schlechte

Teile am häufigsten von Interesse. Man hat also eine Zuordnung in zwei Klassen und es ergibt sich eine 2x2 Matrix zur Beurteilung der Klassifikationsleistung. Die Diagonalelemente entsprechen den korrekten Klassifikationen. Die beiden anderen Fälle (Außerdiagonalelemente) werden als

Fehlausschuss: Gute Teile werden irrtümlich als fehlerhaft klassifiziert. klassifiziert (*Pseudofehler*, Fehler 1. Art, α-Fehler, *Falschalarm*)

Schlupf: bzw. *Durchschlupf* Falsche Teile werden irrtümlich als Gutteil klassifiziert (Fehler 2. Art, β-Fehler)

bezeichnet. Mit den Begrifflichkeiten ist bei entsprechenden Anwendungen generell sehr sorgfältig umzugehen, insbesondere, da die Größen oft unterschiedlich normiert werden bzw. auf unterschiedliche Grundgrößen bezogen werden.

Wie die Klassifikationen „in Ordnung" (iO) bzw. „nicht in Ordnung" (NIO) zustande kommen ist sehr anwendungsabhängig und kann wieder von Klassifizierungen und Messungen abhängen. Eine genaue Analyse kann durchaus sehr kompliziert sein, orientiert sich aber letztlich an der hier dargestellten Logik.

Eine Vielzahl von Begriffen werden in der Praxis verwendet, um verschiedene Wahrscheinlichkeiten bzw. Raten anzugeben. Dabei ist wichtig, sich jeweils klar zu machen, auf welchen Fall sich der Begriff bezieht. Die *Spezifität* z.B. kann entweder in obiger Tabelle den richtig-negativ Wert oder aber den richtig-positiv Wert angeben. Der Zusammenhang entscheidet also oftmals hinsichtlich der Definition.

Wesentlich ist außerdem, dass die reine Erkennungsleistung des BV-Systems meist nicht die letztlich relevante Größe für den Kunden des BV-Systems darstellt. Dieser ist nämlich in der Regel am real auftretenden Fehlausschuss bzw. Schlupf in der Produktion interessiert.

Nur auf den ersten Blick wird dies durch Tab. 11.5 spezifiziert. Dort sind nämlich Klassifikationswahrscheinlichkeiten angegeben. Konkret: Wenn von 100 defekten Produkten z.B. 10 Produkte irrtümlicherweise als „in Ordnung" klassifiziert wurden dann ergibt sich dort ein Eintrag von 10%.

Entscheidend ist nun aber *auch*, wie hoch der Anteil der fehlerhaft produzierten Teile überhaupt ist. Wenn dieser Anteil sehr gering ist (Extremfall: Null), dann

ist die Nichterkennungsrate von nur 10% unter Umständen sehr gut akzeptabel während sie bei einem hohen Anteil an fehlerhaften Teilen intolerabel sein kann.

Beispiel: Hersteller A produziert 50% seiner Sicherheitsgurte falsch. Eine Nichterkennungsrate von 10% bedeutet, dass letztlich von 1000 Sicherheitsgurten 50 fehlerhaft ausgeliefert werden (Schlupf 5%). Hersteller B verwendet das selbe Inspektionssystem, seine Produktion ist aber deutlich sorgfältiger und es werden dort lediglich mit 5% Wahrscheinlichkeit defekte Gurte produziert. Damit wird Hersteller B trotz Verwendung desselben Inspektionssystems bei seiner Auslieferung nur 5 fehlerhafte Gurte (Schlupf 0.5 %) ausliefern.

Letztlich müssen also die Erkennungsraten r_{ij} noch mit der Eintrittswahrscheinlichkeit p_i für das entsprechende Ereignis i multipliziert werden, um eine für den Kunden des BV-Systems relevante Größe zu erhalten:

$$g_{ij} = p_i \, r_{ij} \tag{11.7}$$

Diese Größe g_{ij} wird „gewichtete Erkennungsrate" für die Klasse i genannt.

Schließlich ist noch zu berücksichtigen, dass die Auswirkungen von verschiedenen Fehlklassifikationen unterschiedlich sein können. Zunächst sind Teilklassifikationen unter Umständen gänzlich ohne direkte Auswirkung auf die finale Beurteilung der Produktqualität, aber auch wenn die Auswirkung vorhanden ist, kann sie unterschiedlich stark ausfallen. Z.B. ist bei einem Sicherheitsgurt Schlupf sicher als deutlich weniger wünschenswert zu werten als Fehlausschuss. Der Fehlausschuss ist zwar ärgerlich aber weit weniger relevant. Die Auswirkungen können also sehr unterschiedlich sein und müssen spezifiziert werden. Entsprechend kann dann das Klassifikationssystem anhand dieser Vorgaben ausgelegt werden.

Entscheidenden Einfluss hat diese Kombination aus Produktionsstatistik (wieviele Teile werden fehlerhaft produziert) und (ungewichteten) Erkennungsraten (wie gut kann das BV-System etwas klassifizieren) also ganz wesentlich auf die Auslegung, die prinzipielle Realisierbarkeit und den Nachweis der Leistungsfähigkeit einer BV-Lösung.

Bevor wir uns um diesen Aspekt bemühen, wollen wir die Zusammenhänge noch etwas allgemeiner angeben und verwenden den Satz von *Bayes* (vgl. Anhang) für zwei Ereignisse A und B

$$P_{A|B} = \frac{P_{B|A}P_A}{P_B} = \frac{P_{B|A}P_A}{P_{B|\neg A}(1-P_A) + P_{B|A}P_A} \tag{11.8}$$

P_A - Wahrscheinlichkeit, dass A wahr ist
P_B - Wahrscheinlichkeit, dass B wahr ist
$P_{A|B}$ - Wahrscheinlichkeit, dass A wahr ist, wenn B wahr ist
$P_{B|A}$ - Wahrscheinlichkeit, dass B wahr ist, wenn A wahr ist
$P_{B|\neg A}$ Wahrscheinlichkeit, dass B wahr ist wenn A nicht wahr ist

Beispiel: Unser BV-System findet fehlerhafte Teile mit einer Wahrscheinlichkeit von 99%. Die Produktion ist vergleichsweise gut und liefert (OHNE INSPEKTION) nur 2% Ausschuss (fehlerhafte Teile). Wenn nun unser System ein Fehlteil anzeigt, wie hoch ist dann die Wahrscheinlichkeit, dass dieses Teil wirklich ein fehlerhaftes Teil ist ?

Wir bezeichnen mit A das Ereignis, dass das Teil fehlerhaft ist und mit B das Ereignis, dass das Teil fehlerhaft klassifiziert wurde. Demnach interessieren wir uns für $P_{A|B}$.

Wir haben P_A = 0.02 und $P_{B|A}$ = 0.99.

Wir benötigen zur Berechnung aber noch $P_{B|\neg A}$, also die Wahrscheinlichkeit, dass das System einen Defekt anzeigt, wenn gar kein Defekt vorliegt. Hierfür nehmen wir eine Wahrscheinlichkeit von 1% (Pseudodefekt Fehlklassifikationsrate) an.

Und damit

$$P_{A|B} = \frac{0.99 \cdot 0.02}{0.99 \cdot 0.02 + 0.01 \cdot (1 - P_A)} = 0.67 \tag{11.9}$$

Das bedeutet, dass die Wahrscheinlichkeit, dass das Teil wirklich fehlerhaft ist statistisch nur in 67% aller Fälle korrekt ist bzw. dass 33% der Teile, die als fehlerhaft markiert wurden in Wirklichkeit nicht fehlerhaft sind (CA. 33% Fehlausschuss bzw. Pseudodefekte), ein Wert mit dem der Produzent der Teile unter Umständen ganz und gar nicht zufrieden ist.

Das Ergebnis mag auf den ersten Blick merkwürdig erscheinen aber letztlich sind wir im Abschätzen solcher Zahlen bei niedrigen Wahrscheinlichkeiten sehr schlecht und eine echte Rechnung ist praktisch immer notwendig.

Wir kommen nun zur Frage, wie man die Leistungsfähigkeit eines entsprechenden Inspektionssystems statistisch nachweist.

Nehmen wir z.B. an, dass letztlich der Schlupf für eine Anwendung kleiner 1 % gehalten werden muss. Das bedeutet, dass im Mittel von 100 produzierten Teilen lediglich ein defektes Teil als „gut" das System passieren darf. Wenn die Produktion so gut ist, dass auch nur 1% der Teile defekt produziert werden, dann ist eine Kontrolle nicht notwendig[15]. Wenn die Produktion aber $p_p = 5\%$ Ausschuss generiert, dann müssen die „meisten" der entsprechend defekt detektierten Teile gefunden werden.

Wie hoch muss die Erkennungsrate p_f für die Fehlerdetektion des BV-Systems sein? Die Wahrscheinlichkeit für das fälschliche Durchlassen des defekten Teils ist $p_p \cdot (1 - p_f)$ und dies soll kleiner 1% bleiben. Damit ergibt sich $p_f > 1 - 0.01/0.05 = 80\%$.

Dies ist eine klare Forderung an die Implementierung des BV-Systems und bei der Abnahme muss diese Erkennungsrate von 80% nachgewiesen werden. Dies ist bei einer Abnahme anhand der laufenden Produktion unter Umständen nicht ganz trivial denn hierfür werden viele produzierte Teile notwendig sein.

Entscheidend für die Spezifikation, aber auch die Planung der Abnahme des BV-Systems, ist die Konfidenz bzw. „Vertrauenswürdigkeit", mit der das System eine Entscheidung bzw. eine Klassifikation durchführt. Die Konfidenz sagt also etwas darüber aus, wie zuverlässig das System arbeitet.

Die Prüfung von 100 Teilen zum Nachweis der Leistungsfähigkeit des Systems ist natürlich in unserem Beispiel nicht ausreichend, denn bei 100 Teilen wird man im Mittel nur 1 defektes Schlupfteil erwarten und eine Statistik mit nur einem Ereignis ist sicher nicht sonderlich aussagekräftig.

Wenn wir mit N die zu betrachtenden Teile bezeichnen von denen k Teile durchgelassene fehlerhafte Teile sind (Schlupf), dann ist die relative Häufigkeit, die wir

[15] vorausgesetzt, es ist sichergestellt, dass sich die Produktionsergebnisse nicht ändern

bei der Abnahme nachweisen $h = p_p = 0.01 = k/N$

Je mehr Teile wir untersuchen (N groß), desto sicherer wird unsere Entscheidung hinsichtlich der Gesamtleistung unseres Systems sein. Für ein *Vertrauensniveau* von $1 - \alpha$ kann nach Agresti und Coull ein Vertrauensbereich

$$\Delta p = z_{1-\alpha/2} \sqrt{\frac{q(1-q)}{N+4}} \tag{11.10}$$

$$\approx z_{1-\alpha/2} \frac{p}{\sqrt{N}} \tag{11.11}$$

angegeben werden, wobei die Näherung für große N korrekt ist. $z_{1-\alpha/2}$ bezeichnet dabei die entsprechende Quantile der *Normalverteilung*. q entspricht letztlich einer leicht geänderten Häufigkeit

$$q = \frac{k+2}{N+4} \approx p \tag{11.12}$$

Im Fall eines ca. 70% großen Vertrauensinterval sagt die Formel schlicht aus, dass die erwartete Streuung der Häufigkeit eben gerade der Standardabweichung der Normalverteilung entspricht. Dies ist ganz analog zu der Streuung eines Messwerts zu sehen. Für den Fall vieler, zufälliger Fehler erwartet man mit einer 70%-igen Wahrscheinlichkeit den Messwert in einem 1 σ Intervall der Messdaten.

Für eine 95% Vertrauensinterval ist $\alpha = 0.05$ und damit 1-$\alpha/2$ = 0.975. Die entsprechende Quantile der standardisierten Normalverteilung kann man nicht analytisch berechnen und liegt tabeliert vor[16].

Die Werte für einige wichtige Vetrauensintervalle seien hier beispielhaft angegeben: z(50%) = 0.68, z(70%) = 1.04, z(80%) = 1.28, z(90%) = 1.605, z(95%) = 1.96.

Wenn wir 1000 Teile bei der Abnahme testen, dann ergibt Gl. 11.10 bei eine 90% Vertrauensinterval somit wegen $q = (10+2)/1004 = 0.12$ %

$$\Delta p = 1.605 \sqrt{\frac{q(1-q)}{1004}} = 0.5\% \tag{11.13}$$

Bei einer Messung von 1000 Teilen, bei denen wir 10 Schlupfteile gefunden haben, können wir also statistisch mit 90% Wahrscheinlichkeit sagen, dass der echte Schlupf im Bereich 1 ± 0.56 %, grob also zwischen 0.5 und 1.5 Prozent liegen wird. (Bei 10.000 Teilen würde sich entsprechend das Intervall auf 1 ± 0.16 % verbessern.)

[16] siehe z.B. http:de.wikibooks.orgwikiMathematik:_Statistik:_Normalverteilungstabelle

Wenn der Kunde fordert, dass der Schlupf kleiner 0.1 % mit einer Wahrscheinlichkeit von 90% liegen muss dann können wir entweder mehr Teile zur Verifikation bei der Abnahme verwenden, oder aber wir verbessern die Erkennungsrate unseres Systems, so dass entsprechend der Schlupf verbessert wird.

Die Darstellung zeigt, dass die Zusammenhänge schnell sehr komplex werden und dass für eine brauchbare statistische Absicherung der Erkennungsleistung oft sehr viele Teile geprüft werden müssen. Teilweise ergibt sich eine Anzahl notwendiger Prüfungen, die nicht praktikabel durchgeführt werden kann. Vor der Vereinbarung hinsichtlich der Abnahme eines Systems ist daher eine exakte Berechnung unerlässlich.

Falls die Anzahl notwendiger Prüfungen zu hoch ist kann ein Musterkatalog mit Fehlern Abhilfe schaffen. In diesem Fall werden nicht N zufällige Teile überprüft, sondern man einigt sich auf eine Stichprobe mit einer geänderten Statistik. In dieser Stichprobe sind dann fehlerhafte Teile, die ansonsten in der Produktion nur selten vorkommen in ausreichender Zahl vorhanden.

Kapitel 12

Abschließende Bemerkungen

Der Erfolg eines industriellen Bildverarbeitungssystems entscheidet sich oftmals nicht durch die verwendete Software sondern bereits bei der eigentlichen Datenerfassung. Je besser das Bildmaterial, desto einfacher, robuster und schneller die Bildverarbeitung.

Und dabei ist die gesamte Aufnahmekette von der Beleuchtung bis zur Datenerfassung mit dem Computer oder einer Bilderfassungskarte relevant. Zu wenig Licht oder z.B. auch eine mangelhafte Unterdrückung von Fremdlicht führt zu Rauschen und teilweise unvorhersehbaren Bildern. Rauschen resultiert in einer Reduktion der effektiv nutzbaren Auflösung und bei messenden Systemen zu einer Erhöhung des Messunsicherheit. Aber das ist nur einer von vielen Zusammenhänge. Für jede konkrete Aufgabenstellung ist neu zu überlegen, wie das passende Inspektions- oder Messsystem idealerweise ausgelegt werden sollte.

Viele Parameter unseres Erfassungssystems sind dabei in komplexer Weise gekoppelt, so dass die Änderung eines Parameters (z.B. Blendenzahl) immense Auswirkungen auf viele weitere Parameter hat (z.B. bei der Blendenzahl: Auflösung, Rauschen, Belichtungszeit, Schärfentiefe). Diese Abhängigkeiten führen dann natürlich zu weiteren Abhängigkeiten (hier z.B. unter anderem: Wahl des Bildsensors und der Beleuchtung).

In jedem Fall gilt es, Beleuchtung, Optik und Bildsensorik passend und aufeinander abgestimmt auszuwählen. Und das ist nicht einfach. In den meisten Fällen wird man daher mit einem vielleicht nicht optimalen aber doch sehr gut geeigneten System zufrieden sein. Aber auch um ein solches System auszulegen bedarf

es entweder sehr viel Erfahrung oder ein nicht geringes Maß an Wissen über die zugrunde liegenden Zusammenhänge.

Ich hoffe, dass ich Ihnen mit diesem Buch helfen konnte, genau diese Zusammenhänge zu verstehen. Vor allem sollte Ihnen klar geworden sein, wie komplex die Zusammenhänge sind und dass sehr viele Parameter ganz wesentlich sind. Hüten Sie sich vor dem naiven Bildverarbeiteransatz „Ich nehme irgendeine Kamera und den Rest krieg ich schon per Software hin". Stattdessen sollte für Sie immer Priorität haben, zunächst dafür zu sorgen, möglichst gut verarbeitbare Bilder zu gewinnen. Optimieren sie also immer den Kontrast der sie interessierenden Information, reduzieren Sie das Rauschen und sorgen sie für stabile Verhältnisse.

Kapitel 13

Anhang: Wahrscheinlichkeiten

Was hat dieses Kapitel überhaupt in diesem Buch zu suchen? Nun, wenig (deshalb ist es in den Anhang gewandert :)). Aber: Sobald es eben um die Auslegung von inspizierenden, klassifizierenden oder detektierenden Systemen geht, werden die Wahrscheinlichkeiten wichtig (siehe Kapitel 11) und daher hier ein paar Basics, die das vielleicht eingerostete Verständnis erhöhen.

13.1 Oder und Und

Wenn wir zwei Versuchsausgänge A und B betrachten, dann ergibt sich direkt

$$P(A \text{ oder } B) = P(A) + P(B) \tag{13.1}$$

Das typische Beispiel ist ein Gewinnspiel mit einem Würfel. Ich gewinne, wenn das Ereignis A, z.B. "Wurf der Zahl 5", *oder* das Ereignis B, z.B. "Wurf der Zahl 6" eintritt. Die Gewinnwahrscheinlichkeit ist dann logischerweise 1/6 + 1/6 = 1/3.

Wir ändern nun minimal etwas an den Versuchsbedingungen: Wir nehmen zwei Würfel und würfeln hintereinander. A sei das Ereignis "Würfel 1 ergibt 5" und B sei das Ereignis "Würfel 2 ergibt auch 5".

Wie ist nun die Wahrscheinlichkeit, dass A *oder* B eintritt. Jetzt hängt es ein wenig davon ab, was wir mit "oder" meinen. Wir meinen damit im folgenden *nicht* das sogenannte *exklusives oder*, bei dem wirklich nur das Eine oder eben das Andere eintreten darf.

Bei uns meint "oder", dass das Eine oder das Andere oder auch Beide gleichzeitig eintreten dürfen.

Jetzt müssen wir aber aufpassen: Es kann nämlich leicht passieren, dass man eine Ereignisse "doppelt zählt". Dazu schauen wir uns alle möglichen Versuchsausgänge in Abb. 13.1.1 an.

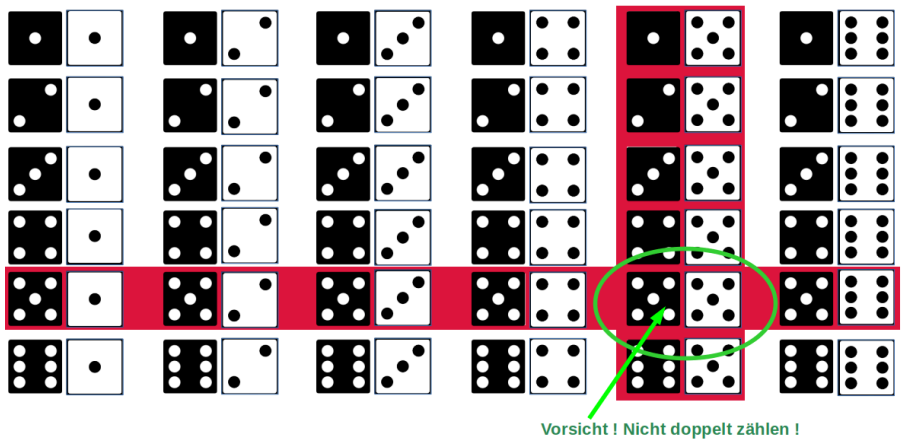


Abbildung 13.1.1: Bei der Verknüpfung von Ereignissen mit "oder" muss man aufpassen, dass man Ereignisse nicht doppelt zählt (hier grün markiert). Dargestellt sind die zwei Würfelwürfe und gesucht ist die Wahrhscheinlichkeit dafür, mindestens eine 5 zu würfeln, also P(W1=5 oder W2=5.

Die günstigen bzw. relevanten Ereignisse sind rot hinterlegt. Ganz offensichtlich haben wir 6 + 6 - 1 = 11 Versuchsausgänge, die unsere Bedingung WürfelA=5 (Ereignis A) *oder* Würfel2=5 (Ereignis B) erfüllen. Mathematisch

$$P(A \text{ oder } B) = P(A) + P(B) - P(A \text{ und } B) \tag{13.2}$$

$$= \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36} \tag{13.3}$$

Wie kommt die Diskrepanz zwischen unserer ersten Gleichung (Gl. 13.1) und dieser Gleichung zustande?

Nun, bei unserem ersten Versuch mit einem Würfel kann es nicht vorkommen, dass gleichzeitig A und B eintreten, denn der eine Würfel muss sich ja für eine Zahl entscheiden. Also ist P(A und B) = 0. Vergessen Sie also am besten die erste Gleichung und halten sich generell an die Zweite.[1]

[1] Sie können die erste Gleichung immer dann nehmen, wenn Sie eben sicher sind, dass sich zwei Ereignisse A und B gegenseitig ausschließen, also nicht gleichzeitig eintreten können.

13.2 Bedingte Wahrscheinlichkeit

Immer wenn wir Vorinformation haben, dann können wir die natürlich nutzen. Wir können damit unserer Aussagen über die Wahrscheinlichkeit von Ereignissen oder Ereigniskombinationen verbessern.

Durch das Vorwissen wird aus der Gesamtheit aller möglichen Ereigniskombinationen letztlich ein Teil eliminiert. Und damit ändern sich natürlich die Wahrscheinlichkeiten.

Konkret: Nehmen wir an, dass wieder zwei Würfel geworfen werden und jemand verrät mir, dass mit dem zweiten Würfel eine 5 oder eine 6 geworfen wurde. Dadurch ändern sich natürlich die Möglichkeiten (siehe Abb. 13.2.2).

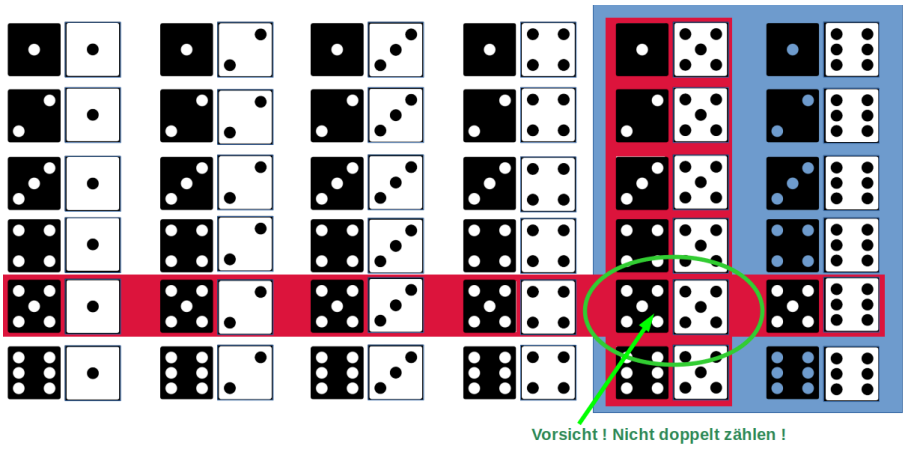


Abbildung 13.2.2: Als Vorinformation wissen wir, dass mit dem zweiten Würfel eine 5 oder eine 6 gewürfelt wurde. Es sind also nur die blau hinterlegten Felder relevant.

Jetzt ist nur noch der blau hinterlegte Bereich an Ereignissen überhaupt relevant. Wenn es uns wieder darum geht, mindestens eine 5 zu würfeln, dann haben wir 7 Ereignisse, bei denen das der Fall ist und wir haben 12 Ereignisse insgesamt, also ist die Wahrscheinlichkeit 7/12.

Man braucht nun einen Formalismus, mit dem man nicht jedesmal graphisch eine Lösung ermitteln muss.

Wir schreiben $P(A|B)$ für die Wahrscheinlichkeit, dass Ereignis A eintritt, wenn Ereignis B eingetreten ist (auf Englisch: "A given B"). in unserem Beispiel gerade wäre B also "Würfel 2 zeigt eine 5 oder eine 6". Und Ereignis A wäre dann: "Mindestens eine 5 wird gewürfelt".

$P(A|B)$ bedeutet ja, dass gleichzeitig A und B eingetreten sind. Aber eben gewichtet mit der Wahrscheinlichkeit, dass B eintritt.

$$P(A|B) = \frac{P(A \text{ und } B)}{P(B)} \tag{13.4}$$

In unserem Beispiel

$$P(A|B) = \frac{P(A \text{ und } B)}{P(B)} = \frac{7/36}{1/3} = \frac{7}{12} \tag{13.5}$$

13.3 Bayes Theorem

Wir können bei Gl. 13.4 einfach auf beiden Seiten mit P(B) multiplizieren und die logische Tatsache nutzen, dass "A und B" dasselbe wie "B und A" ist. Also:

$$P(B) \cdot P(A|B) = P(A \text{ und } B) = P(B \text{ und } A) \tag{13.6}$$

Für $P(B \text{ und } A)$ muss aber eben nach Gl. 13.4 auch gelten (wir vertauschen nur die Buchstaben)

$$P(A) \cdot P(B|A) = P(B \text{ und } A) \tag{13.7}$$

Also letztlich das *Baysche Theorem* [32]:

$$P(A) \cdot P(B|A) = P(B) \cdot P(A|B) \tag{13.8}$$

Dieses Theorem ist trotz seiner Einfachheit beeindruckend. Es erlaubt uns bedingte Wahrscheinlichkeiten aus anderen bedingten Wahrscheinlichkeiten zu berechnen.

Nehmen wir an wir haben Halsschmerzen (A) und würden gerne wissen, wie hoch die Wahrscheinlichkeit ist, dass die Ursache eine Erkältung ist (B). Wir suchen also $P(A|B)$. Wenn wir wissen, wie hoch die Wahrscheinlichkeit ist, dass jemand Halsschmerzen hat, wenn er erkältet ist, also $P(B|A)$, und wir die Grundwahrscheinlichkeiten für Halsschmerzen $P(A)$ und Erkältung $P(B)$ ist, dann

können wir mittels des Bayschen Theorem genau die gesuchte Wahrscheinlichkeit berechnen.

Zurück zu Gl. 13.8. Wenn A und B *voneinander unabhängig* sind, dann gilt

$$P(A|B) = P(A) \tag{13.9}$$
$$P(B|A) = P(B) \tag{13.10}$$

Das ist in unserem Beispiel nicht der Fall. Aber wenn es gilt, dann ergibt sich die einfache Regel

$$P(A \text{ und } B) = P(A) \cdot P(B) \tag{13.11}$$

Das ist z.B. der Fall, wenn wir zweimal mit einem Würfel würfeln und uns nach der Wahrscheinlichkeit fragen, dass beide Würfel eine 6 zeigen (P = 1/6 $\cdot$1/6 = 1/36).

13.4 Das Baysche Theorem angewandt

Wir starten mit Gl. 13.8 und nutzen, dass wir für $P(B)$ schreiben können:

$$P(B) = P(B|A) \cdot P(A) + P(B|\neg A)P(\neg A) \tag{13.12}$$

Diese Gleichung sagt in Worten, dass die Wahrscheinlichkeit für B eben die Summe der Wahrscheinlichkeit von "B und A" und "B und nicht-A" ist. A und nicht-A zusammen sind alle Ereignisse und daher trifft die Aussage zu.

Damit also (Gl. 13.8):

$$P(A|B) = \frac{P(A)\,P(B|A)}{P(B)} \tag{13.13}$$
$$= \frac{P(A)\,P(B|A)}{P(B|A) \cdot P(A) + P(B|\neg A)P(\neg A)} \tag{13.14}$$

Machen wir ein konkretes Beispiel, das sinngemäß in [16] zu finden ist, und die Wichtigkeit des Bayschen Theorems verdeutlicht:

Sie machen im Rahmen einer Routineuntersuchung beim Arzt einen Test auf Bauchspeicheldrüsenkrebs. Der Fehler des Tests betrage 5% (Vorischt, eine etwas lakse Ausdrucksweise). Der Krebs zeigt also mit 95% Wahrscheinlichkeit bei einer kranken Person den Krebs an. Bei einer gesunden Person zeigt er mit 5% Wahrscheinlichkeit dennoch "Krebs" an (falsch-positiv). Bei Ihnen ist der Test leider positiv. Sollten Sie sich Sorgen machen?

Nein. Letztlich interessiert Sie ja nun die Wahrscheinlichkeit, dass Sie Krebs haben. Dabei müssen Sie aber — außer den Daten der Aufgabenstellung — auch noch berücksichtigen, wie denn die Grundwahrscheinlichkeit ist, dass ein Mensch Bauchspeicheldrüsenkrebs hat.

Diese liegt bei lediglich 0.00012.

Vereinfacht (nicht ganz exakt) müssen wir nun also beide Wahrscheinlichkeiten, die für uns von Interesse sind, abschätzen.

Variante 1 bedeutet, dass ich Krebs habe und der Test das anzeigt. Hierfür ist die Wahrscheinlichkeit P(Krebs) = 0.00012 · 0.95 = 0.000114.

Variante 2 sagt, dass der Test zwar positiv ist, aber ich — zum Glück — keinen Krebs habe. Hierfür ist die Wahrscheinlichkeit P(kein Krebs) = (1-0.00012) · 0.05 = 0.05.

Offensichtlich ist Variante 2 um den Faktor 0.05 / 0.000114 = 438 wahrscheinlicher!

Das erscheint vielleicht merkwürdig, aber unser "gesunder Menschenverstand" versagt oft dann, wenn es um sehr kleine Wahrhscheinlichkeiten geht.

Falls Sie Zweifel haben: Durchdenken Sie die Fragestellung mit einer männlichen Testperson und einem Schwangerschaftstest. Auch wenn der Schwangerschaftstest positiv sein sollte gehen Sie vermutlich nicht davon aus, dass der Mann wirklich schwanger ist.

Wie würden wir die Krebs-Fragestellung mit dem Bayschen Theorem lösen?

Wir nennen B das Ereignis "Test positiv". A nennen wir das Ereignis "Patient hat Krebs". Gesucht ist also $P(A|B)$.

Wir nutzen Gleichung 13.14 mit folgenden Werten:
$P(A)$ = 0.00012
$P(B|A)$ = 0.95
$P(B|\neg A)$ = 0.05
$P(\neg A) = 1 - P(A)$ = 0.99988

$$
\begin{aligned}
P(A|B) &= \frac{P(A)\,P(B|A)}{P(B|A) \cdot P(A) + P(B|\neg A)P(\neg A)} && (13.15)\\
&= \frac{0.00012 \cdot 0.95}{0.95 \cdot 0.00012 + 0.05 \cdot 0.99988} && (13.16)\\
&= 0.00228 \quad = 1/440 && (13.17)
\end{aligned}
$$

Literaturverzeichnis

[1] Angelo V Arecchi, R John Koshel, and Tahar Messadi. Field guide to illumination. SPIE, 2007.

[2] European Machine Vision Association et al. Standard for characterization of image sensors and cameras. *EMVA Standard*, 1288, 2010.

[3] Wolfgang Bachmann. *Signalanalyse: Grundlagen und mathematische Verfahren*. Springer-Verlag, 2013.

[4] Norbert Bauer. *Handbuch zur industriellen Bildverarbeitung: Qualitätssicherung in der Praxis*. Fraunhofer-Allianz Vision, 2007.

[5] Vikrant R Bhakta, Manjunath Somayaji, and Marc P Christensen. Effects of sampling on the phase transfer function of incoherent imaging systems. *Optics express*, 19(24):24609–24626, 2011.

[6] Vikrant R Bhakta, Manjunath Somayaji, and Marc P Christensen. Applications of the phase transfer function of digital incoherent imaging systems. *Applied optics*, 51(4):A17–A26, 2012.

[7] Oliver S Cossairt, Daniel Miau, and Shree K Nayar. Scaling law for computational imaging using spherical optics. *JOSA A*, 28(12):2540–2553, 2011.

[8] Christian Demant, Bernd Streicher-Abel, and Axel Springhoff. *Industrielle Bildverarbeitung: wie optische Qualitätskontrolle wirklich funktioniert*. Springer-Verlag, 2011.

[9] Rainer G Dorsch, Gerd Häusler, and Jürgen M Herrmann. Laser triangulation: fundamental uncertainty in distance measurement. *Applied optics*, 33(7):1306–1314, 1994.

[10] David Dussault and Paul Hoess. Iccds edge out electron-multiplying ccds in low light. *Laser focus world*, 40(9):69–75, 2004.

[11] DIN ENV. 13005. leitfaden zur angabe der unsicherheit beim messen (gum), 1999.

[12] Robert D Fiete. *Modeling the imaging chain of digital cameras*. SPIE press Bellingham, 2010.

[13] David Fleischle. *Konzepte und Implementierung fÃ$\frac{1}{4}$r die prozessnahe Integration optischer Sensoren*. Dissertation Universität Stuttgart, 2013.

[14] Wolfgang Förstner. Image preprocessing for feature extraction in digital intensity, color and range images. In *Geomatic method for the analysis of data in the earth sciences*, pages 165–189. Springer, 2000.

[15] Guillermo Gallego, Tobi Delbruck, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew Davison, Jörg Conradt, Kostas Daniilidis, et al. Event-based vision: A survey. *arXiv preprint arXiv:1904.08405*, 2019.

[16] Larry Gonick and Woollcott Smith. *Statistik in Cartoons*. Vahlen, 2009.

[17] Joseph W Goodman. *Introduction to Fourier optics*. Roberts and Company Publishers, 2005.

[18] Ryan D Gow, David Renshaw, Keith Findlater, Lindsay Grant, Stuart J McLeod, John Hart, and Robert L Nicol. A comprehensive tool for modeling cmos image-sensor-noise performance. *IEEE Transactions on Electron Devices*, 54(6):1321–1329, 2007.

[19] H Gross. Handbook of optical systems–fundamentals of technical optics, vol. 1, 2005.

[20] H Gross, H Zugge, M Peschka, and F Blechinger. Handbook of optical systems: Aberration theory and correction of optical systems, vol. 3, 2007.

[21] Gerd Häusler. A method to increase the depth of focus by two step image processing. *Optics Communications*, 6(1):38–42, 1972.

[22] Eugene Hecht. *Optik*. McGraw-Hill, 1987.

[23] Maria Hennes. Konkurrierende genauigkeitsmaße–potential und schwächen aus der sicht des anwenders. *Allgemeine Vermessungs-Nachrichten*, 7:136–146, 2007.

[24] Philip CD Hobbs. *Building electro-optical systems: making it all work*, volume 71. John Wiley & Sons, 2011.

[25] Jörg Hoffmann. *Taschenbuch der Messtechnik*. Taschenbuchverlag Leipzig, 2002.

[26] Gerald C Holst and Terrence S Lomheim. *CMOS/CCD sensors and camera systems*, volume 408. JCD publishing USA, 2007.

[27] Alexander Hornberg. *Handbook of machine vision*. John Wiley & Sons, 2007.

[28] Roarke Horstmeyer, Rainer Heintzmann, Gabriel Popescu, Laura Waller, and Changhuei Yang. Standardizing the resolution claims for coherent microscopy. *Nature Photonics*, 10(2):68, 2016.

[29] Kenji Irie, Alan E McKinnon, Keith Unsworth, and Ian M Woodhead. A technique for evaluation of ccd video-camera noise. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(2):280–284, 2008.

[30] Emily Kubacki. The dirt on cleaning optics. *Photonics Spectra*, 40(3):56, 2006.

[31] Dietrich Kühlke. *Optik: Grundlagen und Anwendungen; mit Tabellen, Beispielen und Aufgaben mit Lösungen*. Harri Deutsch Verlag, 2004.

[32] Will Kurt. *Bayesian statistics the fun way: understanding statistics and probability with Star Wars, Lego, and Rubber Ducks*. No Starch Press, 2019.

[33] Chi-Shao S Lin, Bimal P Mathur, David L Standley, and Mau-Chung F Chang. Figure of merit characterizes electronic imager performance. *Laser focus world*, 38(6):107–111, 2002.

[34] Martin Löffler-Mang, Helmut Naumann, and Gottfried Schröder. *Handbuch Bauelemente der Optik: Grundlagen, Werkstoffe, Geräte, Messtechnik*. Carl Hanser Verlag GmbH Co KG, 2020.

[35] Brian K McComas. The art and science of missile defense sensor design. In *Sensors and Systems for Space Applications VII*, volume 9085, page 90850F. International Society for Optics and Photonics, 2014.

[36] John Lester Miller, Edward Friedman, John N Sanders-Reed, Katie Schwertz, and Brain McComas. *Photonics rules of thumb*. Number PUBDB-2021-03249. SPIE Press Bellingham, Washington, 2020.

[37] H.N. Nasse. How to read mtf curves.

[38] Daniel O-Connor, Andrew J Henning, Ben Sherlock, Richard K Leach, Jeremy Coupland, and Claudiu L Giusca. Model-based defect detection on structured surfaces having optically unresolved features. *Applied optics*, 54(30):8872–8877, 2015.

[39] W Oettel. Handbuch der mikroskopie, 1977.

[40] Ulf Persson. Real time measurement of surface roughness on ground surfaces using speckle-contrast technique. *Optics and lasers in engineering*, 17(2):61–67, 1992.

[41] Pramod K Rastogi. *Optical measurement techniques and applications*. Artech House, 1997.

[42] A Rogalski. Optical detectors for focal plane arrays. *Optoelectronics review*, 12(2):221–246, 2004.

[43] John R Rogers. Three-bar resolution versus mtf: how different can they be anyway? In *Optical Engineering+ Applications*, pages 707109–707109. International Society for Optics and Photonics, 2008.

[44] Gottfried Schröder and Hanskarl Treiber. *Technische optik*, volume 8. Vogel Buchverlag, 2007.

[45] Norbert Schuster and Simone Weber. Telezentriefehler und telezentriebereich. *Optik & Photonik*, 3(3):57–59, 2008.

[46] Michael Schuth and Wassili Buerakov. *Handbuch Optische Messtechnik: Praktische Anwendungen für Entwicklung, Versuch, Fertigung und Qualitätssicherung*. Carl Hanser Verlag GmbH Co KG, 2017.

[47] Wolfgang Singer, Michael Totzeck, and Herbert Gross. *Handbook of optical systems, physical image formation*, volume 2. John Wiley & Sons, 2006.

[48] Warren J Smith. *Modern optical engineering*. Tata McGraw-Hill Education, 1966.

[49] M. Wagner. *Elemente der theoretischen Physik* 2. Vieweg, 1991.

[50] M. Zölffel. Das saubere mikroskop. *pdf via google*.

[51] M. Zölffel. Melles griot technical guides. *pdf via cvimellesgriot.com*.

Index

Index

Öle, 205
Ölfilme, 327
Ölimmersion, 175
Überabtastung, 87
Überauflösung, 102
Überlappintegral, 191
Übersprechen, 88
Übertragungsfunktion, 75
1/f-Rauschen, 261

Abbe, 102
Abbildungsgleichung, 16
Abbildungsmaßstab, 14, 15
Aberrationen, 90, 105
Absorption, 202
Abtasttheorem, 70
Abwärme, 225
Aceton, 187
Active Pixel Sensor, 249
AD-Wandlers, 261
afokalen Systemen, 24
Airy-Scheibchen, 100
Aktiv, 310
Aliasing, 68
Alterung, 225
analogen Kameratechnik, 296
Antireflexbeschichtung, 184
Antwort, 47
Aperturblende, 26
Apochromate, 177
Arbeitsabstand, 15
Area of Interest, 281
area sensors, 294
Astigmatismus, 107
Aufheller, 240
Auflösung, 58, 94, 350
Auflösung nach Rayleigh, 101
Auflösungen, 307
Auflicht, 211, 212
Ausgangsverstärkers, 261
Ausleserauschen, 261
Austrittspupille, 31
Auto Gain Control, 273

back-side illumination, 275
back-thinned Detektoren, 275
Bakterien, 240
bandbegrenzte Funktionen, 71
Bandpass, 181
Bandsperre, 181
Bayes, 364
Baysche Theorem, 374
Beleuchtung, 208
Beleuchtungsrichtung, 214
Beleuchtungsstärke, 191
Belichtungszeit, 134, 145, 276
Belichtungszeiten, 277
benetzten Oberflächen, 327
Beschichtung, 167

Bestrahlungsstärke, 111, 112, 120, 130, 191
beugungsbegrenzten Systems, 76
Bewegungsunschärfe, 90
Bewegungsunschärfen, 220, 224
bidirektionale Reflektanz, 203
Bildebene, 15
Bildebenenhologramm, 341
Bildfehler, 105
Bildfeldwölbung, 107, 168
Bildqualität, 94
bildseitig telezentrisch, 173
bildseitige Telezentrie, 38
Bildweite, 15
Binning, 282
Blende, 26
Blendenflecken, 170
Blendenzahl, 28
Blooming, 267
Blue Shift, 183
Bogenlampen, 230
Bolometer, 292
Boltzmann Konstante, 228
BRDF, 200, 203, 335
Brennebene, 19
Brennweite, 16
Brillianz, 196
Brownsche Röhre, 239

C-Mount, 179
CameraLink, 298
CameraLink HS, 298
CCD Sensoren, 250
CCIR, 296
chromatisch konfokale Sensoren, 333
chromatischen Längsaberration, 107
chromatischen Vergrößerungsdifferenz, 107
CMOS, 248
CoaXPress, 298, 299
Cold mirrors, 181
Color Rendering Index, 206
Column-Level noise, 263
Computertomographie, 341
Crosstalk, 88
CS-Mount, 179
CXP, 299

D65, 229
Dark signal non-uniformity, 261
Deckglas, 176
Defekte, 360
Defokussierung, 88
Dekonvolutionsverfahren, 104
Depth-from-Defocus, 330
DIC, 341
dichromatische Farbmischung, 234
Differentieller Interferenzkontrast, 341
Diffraktiv-optische Elemente, DOE, 236
diffuse Beleuchtung, 212
diffusem, 212
Digital Number, 272
Digitale Holografie, 343
digitaler Holografie, 339
digitaler Signalprozessor, 300
Diskretisierung, 65
Dispersion, 153
Dombeleuchtung, 212

Doppelbilder, 184
Doppelpuls-Holografie, 341
doppelt-telezentrisches System, 24
doubling temperature, 260
Dreidimensionalität, 306
DSP, 299
DSPI, 340
Dunkelfeldbeleuchtung, 177, 209, 216
Dunkelstromrauschen, 259, 279, 280
Dunkelstroms, 259
Durchlicht, 211
Durchlichtbeleuchtungen, 211
Durchlichtmikroskopie, 177
Durchschlupf, 363
Dynamik, 267

ebener Wellen, 53
Ebenheit, 176
Edge Pre-select, 282
Edison, 229
EIA 170, 296
Eigenfunktionen, 53
Eindellungen, 217
Eintrittspupille, 29, 31
Einzelschuss, 310
electron bombarded, 279
Electron-Multiplying CCDs, 280
Electronic Speckle Pattern Interferometry, 344
Elektrolumineszenz, 233
elektronische Shutter, 254
embedded Vision, 300
Empfindlichkeit, 189
EMVA1288, 257
Energieerhaltung, 193
Entladungslampen, 231
entozentrische Perspektive, 39
ESPI, 340
Etendue, 195

F-mount, 179
Füllfaktor, 85
Falschalarm, 363
Faltungsintegrale, 50
Faltungssatz, 61
Farbbildaufnehmern, 108
Farbinterpolationsalgorihtmen, 288
Farbsäumen, 108
Farbsensoren, 70, 153
Farbtemperatur, 229
Farbwiedergabeindex, 206
Fehlausschuss, 363
Ferndistanz, 34
FFT, 67
Filmformat, 172
Filter, 180
Fingerabdrücke, 186, 205
FireWire, 297
Fisheye-Objektiven, 171
Fixed-Pattern Rauschen, 261
flächenbezogenen Auflösung, 320
Flatdomes, 212
Fluoreszenz, 239
Fluoreszenzanregung, 205
Fluoreszenzlampe, 232
Fluoreszenzlicht, 234
Fluoreszenzschicht, 274
Focal plane shutter, 255
Fotometrie, 191
fotometrische Größen, 191

Fotoobjektive, 163
Fourierhologramm, 341
Fouriertransformation, 55
FPGAs, 299, 300
Frame Delay Readout, 283
Frame-Transfer, 252
Fremdlicht, 180, 220
Fremdlichtunterdrückung, 181
Fresnelhologramm, 341
front-side illumination, 274
Full-Frame, 252

Gain, 273
Gamma-Kurve, 271
Gegenlichtblenden, 170
Geldscheine, 240
Genauigkeit, 96, 351
GenICam, 299
gerichtetem Licht, 212
gerichteter Beleuchtung, 214
ghosts, 170
Gigabit Ethernet, 298
GigE, 298
Glühbirne, 229
Glühemission, 232
global Shutter, 255
Goldene Regel der Messtechnik, 358
Granulation, 140
Grenzfrequenz, 76
Guide of Uncertainty of Measurements, 353

Halbwertsbreite, 182
Halogenlampe, 229
Hardware-Binning, 282
Hauptebenen, 22
Heiße Pixel, 263
Helligkeit, 198
Heterodyninterferometerie, 324
High Dynamik Range - HDR, 267
Hochpass, 181
Hochpassfilter, 60
holografischer Interferometrie, 341
Homogenität, 225
Hot-Pluggable, 300
Hyperfeldwinkeln, 171
Hyperspektral Snapshot Imager, 289
hyperspektrales Imaging, 290
Hyperspektralsensoren, 289
hyperzentrische Abbildung, 39

IEEE1394b, 297
Immersion, 175
Indium-Galium-Arsenid, 275
Indium-Gallium-Arsenid, 292
Infrarot, 185
Infrarotsensoren, 293
integral imaging, 321
intelligente Kameras, 300
Intensified CCDs, 279
Intensität, 195
Intensitätsbasierte Entfernungssensoren, 343
Interferometers, 323
Interlaced, 250, 296
Interline-Transfer, 252
intra-scene, 268
invarianten Systemen, 46
inverse Filter, 104
inverse Fouriertransformation, 55
Isopropanol, 187

jitter, 90

Köhlerscher Beleuchtung, 212
Kühlung, 279
künstliche Vignettierung, 111
Kalibrierobjekten, 357
Kanten, 154, 210
Kapselung, 221
Kathodolumineszenz, 239
Kepler-Teleskop, 24
Keplerteleskope, 174
Klassifikation, 362
koaxialer Beleuchtung, 214
kohärente Übertragungsfunktion, 84
kohärenten Übertragungsfunktion, 141
Kohärenz, 140, 317
Kohärenzgrades, 318
Kohärenzlänge, 142, 318, 325
Kohärenzradar, 325
Koma, 107
konfokale Mikroskopie, 331
konoskopische Holografie, 341
Kontrast, 79
Konversionsschicht, 234
Kratzern, 209, 214
Kurzbogenlampen, 231

Lambertscher Streuung, 199
Langpass, 181
Laser, 143, 235
Lasertriangulation, 236, 311
Lastenheft, 346
Latenz, 256
Laufzeitmessung, 336
Lebensdauer, 225
lens flare, 170
Leuchtdichte, 193
Leuchtdioden, 233
Leuchtschicht, 231
Leuchtstofflampen, 231
Lichtfluss, 191
Lichtleiter, 217
Lichtleitwert, 196
Lichtquellen, 208
Lichtschnitt, 313
Lichtstärke, 192
Lichtstrom, 191
Lidar, 336
line scan sensors, 294
lineare Polarisation, 241
Lineare Systeme, 43
lineare Systeme, 44
Linearität, 45
Linien, 125
Linienlichtern, 216
Linnik, 328
Lochkamera, 14
Long Working Distance, 176
Lookup-Tabelle, 271
Lumen, 190
Lumogen, 274

M12, 180
Makrogeometrie, 308
Makrotopographie, 309
Maximum permissible error, 351
Messabweichung, 350
Messfehler, 350
Messung, 349
Messunsicherheit, 96, 350, 352

Messunsicherheiten, 307
Metalle, 241
Metallen, 180
Metamerie, 289
Michelson, 328
Micro-Four-Third Standard, 179, 254
Mikrobolometer, 293
Mikrogeometrie, 308
Mikrokanalplatte, 280
Mikrolinsen, 169, 253, 274
Mikroskopobjektive, 174
Mikrostruktur, 215
Mikrotopografie, 215
Mikrotopographie, 309
Mineralien, 240
Minimal object distance, 40
MIPI CSI-2, 299
Mirau, 328
Mittelungen, 156
Modulationstransferfunktion, 54, 74, 117
Moire, 67
Moires, 68
MTF50, 117
Multispektral-Imaging, 290
Multistereo, 321
Musterkataloge, 348

Nahdistanz, 34
Nahfeld, 103
natürliche Vignettierung, 111
Neonröhren, 231
nichtlinearen Algorithmen, 50
nichtlinearen Kamerakennlinien, 270
noise equivalent exposure, 265
noise equivalent power, 265, 268
Non-interlaced, 250
Nonuniformity correction, 262
Normalobjektiv, 171
Normalverteilung, 367
Notch-Filter, 181
numerische Apertur, 27, 175
Nyquist–Frequenz, 69
Nyquist-Frequenz, 70
Nyquistfrequenz, 68

Objektiv, 22
objektive Speckles, 140
objektseitig telezentrische Abbildung, 37
Objektweite, 15
OCT, 325
offline, 5
Optical Density, 184
Optische Dichte, 184
Optische Kohärenztomografie, 325
optischen Achse, 23
organische Stoffe, 240
Ortsbandbreitenprodukt, 109
out-of-plane, 340

partial scan, 281
Passive Verfahren, 321
Peltier, 279
perizentrische Abbildung, 39
Perspektive, 35
Pflichtenheft, 346
Phasenkontrastobjektive, 176
Phosphoreszenz, 239
Photo Response Nonuniformity, 261
Photoeffekt, 244

photometrisches Stereo, 335
photometrisches Strahlungsäquivalent, 190
Photonen, 259
Photonenrauschen, 258
pixel-basierter, 268
Pixelpitch, 100, 253
Plancksches Strahlungsgesetz, 228
plenoptischen Kameras, 321
PMD-Sensoren, 336
Polarisation, 176, 240
Polarisationsfilter, 185
Polarisationsfiltern, 180
polarisationsoptischer, 283
Positionsauflösung, 96, 154
Präzision, 351
Prüfmittel, 361
Prüfmittelfähigkeit, 361
Prüfung, 350
Progressive Scan, 250
PSD, 313
Pseudofehler, 363
Puls-Width Control, 282
Pulse-Width-Modulation, 238
Pulsung, 221
Punkttriangulation, 311
Push-Broom, 290

Q, 112
quadratisches Abstandsgesetzes, 193
Quantendetektoren, 292
Quanteneffizienz, 247, 258
Quantisierungsrauschen, 261, 263
Quantum Well Capacity, 247
Quecksilberdampf, 231

räumliche Kohärenz, 143
Rückführung, 355, 357
Radar, 336
Radiometrie, 191
Radiometrische Entfernungssensoren, 343
radiometrischen Größen, 191
Rauheit, 140, 307
Raumwinkel, 193
Rayleigh, 101
Read-out noise, 261
Referenzmesssystem, 357
Reflexionen, 180
Regions of Interest, 249, 281
Reinigung, 186
Rekonstruktionsformel, 71
Relay Linsen, 164
Reproduzierbarkeit, 351
Responsitivity, 272
Retrofokusobjektiv, 171
Rettungswesten, 239
Richtigkeit, 351
Ringbeleuchtungen, 213
Ringlicht, 216
RMS Spotradius, 135
Robustheit, 310
Rolling Shutter, 255
row noise, 263
RS 170, 296

S-Mount, 180
Sättigungskapazität, 247
Sampling, 65
Sampling–Theorem, 69
Samplingtheorem, 66, 113

saturation equivalent exposure, 267
scene noise, 262
Schärfentiefe, 32
Schärfentiefenbereich, 33
Scheimpflug, 178
Schichten, 309, 327
Schlieren, 182
Schlupf, 363
Schutzgläser, 180, 184
schwarzer Strahler, 227
sCMOS, 249
Self-mixing Interferometrie, 343
Sensitivität, 200, 272
Sensorgröße, 128
Sensortransferfunktion, 283
shape from Defocus, 330
shape from shading, 335
Shearinginterferometrie, 341
Shearografie, 341
Shutter, 252, 254
Sicherheitsanwendungen, 240
Signal-Rausch-Verhältnis, 112, 256
Silizium, 205
Single-Pixel Kameras, 281
single-shot, 310
Smart-Cameras, 300
Snapshot Modus, 255
Sonnenlicht, 222
Sparrow Kriterium, 135
Speckle Interferometrie, 340
Specklekontrast, 318
Speckles, 140
Spektralfiltern, 180
Spektralinterferometrie, 325
spezifische (Licht-)Ausstrahlung, 195
Spezifität, 363
sphärische Aberration, 106
spurious resolution, 90
Störlicht, 209
Störreflexe, 241
Staub, 226
STED Mikroskopie, 103
Stefan-Boltzmann-Gesetz, 228
Stereoverfahren, 321
Stereovision, 321
Strahlungsleistung, 190, 191
Strehl, 135
Streifenprojektion, 314
Streulicht, 169
Stromsparlampen, 232
structured illumination microscopy, 331
strukturierte Beleuchtung, 314
subjektive Speckles, 140
Swan, 229
swept source, 333
swept-source OCT, 325
SWIR, 205
Systematischer Fehler, 351

technische Oberflächen, 307
Teilsystemen, 81
teiltelezentrische Objektive, 173
Telefaktor, 171
Teleobjektiv, 23, 171
Telezentrie, 172
Telezentriefehler, 38, 173
telezentrische Beleuchtungsoptik, 211
telezentrischer Perspektive, 36

Temperaturbestimmung, 293
Thermische Detektoren, 292
thermisches Rauschen, 260, 261
thermoelektisch, 279
Thermografie, 292
Three-bar resolution targets, 135
Tiefenabbildungsmaßstab, 21
Tiefpass, 58, 181
Tiefpassfilter, 59
Time-of-flight sensing, 336
Topografiesensoren, 307
Total integrated scatter, 203
tote Pixel, 262
Transimpedanzwandler, 250
Triangulationsbasierte Messverfahren, 310
Triggerimpulses, 282
Tubuslänge, 176
Tubuslinse, 174

unpolarisiertes Licht, 241
USAF 1951, 135
USB-3, 298
UV-empfindliche Sensoren, 274

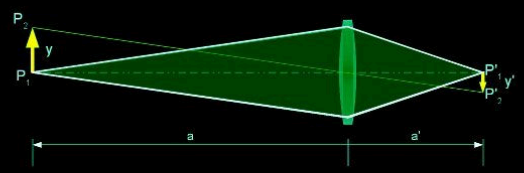
Varifocal Objektive, 163
veiling glare, 170
Vergütung, 167
Vergleichspräzision, 352
Vergrößerung, 174
Verschmutzungen, 186, 226
Verstärkung, 273
Vertrauensniveau, 367
Verwackeln, 90
Verzeichnung, 106
Verzeichnungen, 166
Vibrationen, 226, 309, 328
Vignettierung, 27, 111
Vision-Sensoren, 301
Vollbildverfahren, 250

warme Pixel, 263
Waschmittel, 240
Wasserimmersionsobjektive, 175
Weißlichtinterferometrie, 325
Weitwinkelobjektive, 171
Wiederholbarkeit, 351
Wolfram, 229

XBO, 231

Zeilenkamera, 173
Zeilenkameras, 220
Zeilensensoren, 294
Zoomobjektive, 163
Zweipunkt Trennung, 96
Zwischenringe, 171
Zykluszeit, 347

$$\frac{1}{a'} = \frac{1}{a} + \frac{1}{f'} \qquad a = f'\left(\frac{1}{\beta'} - 1\right) \qquad F\# = K = \frac{f'}{D_{EP}} = \frac{1}{2\,NA'}$$



$K_{eff} := (1+\beta')K$

$NA = NA'\,\beta'$

$\beta' = M := \frac{y'}{y} = \frac{a'}{a}$

$\delta z \approx 2K\,B'/\beta'^2$

$\Delta z' = \beta'^2\,\Delta z$

$\delta z'_{\lambda/4} \approx 4\lambda K^2$

$\delta z' = \frac{\lambda}{2NA'^2}$

	Radiometric	Photometric
Φ	Flux [W]	Luminous flux [lm]
$E = \Phi/A$	Irradiance [W/m²]	Illuminance [lx = lm/W]
$I = \Phi/\Omega$	Radiant Intensity [W/sr]	Luminous intensity [cd = lm/sr]
$L = I/A$	Radiance [W/(m²· sr]	Luminance [cd/m²]

$a_{1/2} = \frac{\beta' \cdot a \cdot D_{EP}}{\beta' \cdot D_{EP} \pm B'}$

$a_{Hyperf} = \frac{f'^2}{K \cdot B'}$

$\Omega = \frac{A}{r^2} \approx \pi \sin^2\alpha$

$\Phi(\lambda) = 683\,\frac{lm}{W}\int_0^\infty V(\lambda)\,\Phi_e(\lambda)\,d\lambda$

$E = \pi L \sin^2(\alpha')\cos^4(\omega')$

$\frac{R_{scatter}}{R_{total}} = 1 - \exp(-(4\pi\cos\phi_{in}\cdot\frac{\sigma_{RMS}}{\lambda})^2)$

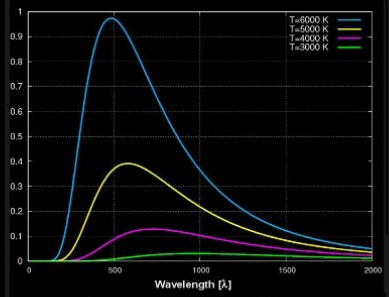
$I(\phi) = I_0\cos(\phi)$

$E = \pi L$

$E_{obj} = \frac{\pi}{80.000}\,\frac{(a^2+y^2)\cdot N^4\cdot ps^2\cdot\delta z^2}{a^4\,S\cdot y^4\cdot t}$

$\Phi_e[W] = FWHM[nm]\cdot 0.02$ @ 6000 K, 1000 lm

$\Phi_e[W] = (\lambda[\mu m] - 0.4)\frac{FWHM[nm]}{10}$ @ 3000 K



Lamp	CRI
Halogen	100
White LED	80...95
Sodium	24
White Fluorescence	50 ... 80

Material	Reflect
Aluminimum polished	0.65
Aluminimum unpolished	0.55
Chrome polished	0.65
White Paint	0.8
Granite	0.15
Paper	0.7
White Silk	0.3
Skin	0.45
Snow	0.7
Dark Wood	0.07
Bright Wood	0.4
Black Paint	0.025

λ	$V(\lambda)$
370	0.00001
380	0.00004
390	0.00012
400	0.00040
410	0.0012
440	0.0230
450	0.0380
470	0.0910
490	0.4080
500	0.3230
510	0.5030
520	0.7100
530	0.8620
555	1.0000
570	0.9154
590	0.7570
600	0.6310
620	0.3810
650	0.1070
670	0.0320
700	0.0041
720	0.0010
740	0.0002
770	0.00003
785	0.00001

Conditions	Illuminance
Office	500
Bright sunlight	100.000
Bright day, shadow	10.000
Bright Moon	0.45
Dark Moon	0.001

Camera	Resonsitivity
Dalsa Piranha 3	14 DN/J/cm²
Dalsa Pantera	20 DN/J/cm²
Dalsa Piranha ES	246 DN/J/cm²
Dalsa Spyder 3	129 DN/J/cm²

Format	Diagonal [mm]	Geometry [mm x mm]
1/8 inch	2	1.6 x 1.2
1/6 inch	3	2.4 x 1.8
1/4 inch	4	3.2 x 2.4
1/3 inch	6	4.8 x 3.6
1/2 inch	8	6.4 x 4.8
2/3 inch	11	8.8 x 6.6
4/3 inch	21	17.3 x 13

Interface	Datarate	Max. Cable L
CameraLink	240 MB/s per Chipset	10 m
GigE	100 MB/s	100 m
CoaXPress	780 MB/s	100 m
IEEE 1394a	50 MB/s	4.5 m
IEEE 1394b	83 MB/s	100 m
USB 2.0	38 MB/s	10 m
USB 3.0	350 MB/s	3 m

plated points or lines	$r_A \approx 1..2 \cdot CD$	3 Pixels = 1 CD
ighbouring points or lines	$r_A \approx 0.6..1 \cdot CD$	3..4 Pixels = 1 CD
riodic structures	$MTF_{10\%} \approx 1/CD$	3..4 Pixels = 1 CD

$r = 0.61\frac{\lambda}{NA'} = 1.22\lambda K$

$\beta' = -\frac{f_2}{f_1}$

$Q = \frac{r_A}{p}$

$MTF(f) \approx 1 - f$

$f := f_x/f_{max}$

$f_{max} \approx \frac{1.22}{r_{spot}} \approx \frac{1}{\lambda K}$

$V \approx 6..10 \cdot RMS$

ncertainty $\approx 10 \cdot$ Resolution

$\sigma \to 68\%$

$\sigma \to 95\%$

$\sigma \to 99.7\%$

$_{|B}P_B = P_{B|A}P_A$

$_{|B} = \frac{P_{B|A}P_A}{P_{B|\neg A}(1-P_A) + P_{B|A}P_A}$

ometimes : $P_{B|\neg A} = 1 - P_{B|A}$

$= \frac{\lambda^2}{\Delta\lambda}$

$_{ver} \approx 0.61\frac{\lambda}{NA'}$

$(I) \sim \exp(-\frac{I}{<I>})$

$_{phi} = \lambda\sqrt{1 - \frac{\sin^2\phi}{n^2}}$

$\Delta N = \sqrt{N}$

$SNR = \frac{P\,QE\,t}{\sqrt{(P+B)\,QE\cdot t + DC\,t + RON^2}}$

Dynamic Range $= \frac{FWC}{noise}$

Dynamic Range dB $= 6\,N$ [dB]

$E_{Ph} = h\nu = \frac{h\,c}{\lambda} = \frac{198.8\cdot 10^{-27}}{\lambda}\,J\,m$

AiryPower $P(r < r_A) \to 84\%$

AiryPower $P(r < 2r_A) \to 91\%$

GaussianPowerl. $PL = \exp(-2\frac{D_{EP}}{D_{1/e^2}})$

$\frac{D_{EP}}{D_{1/e^2}} = 1 \to PL = 13.5\%$

$\frac{D_{EP}}{D_{1/e^2}} = 0.5 \to PL = 37\%$

Strehl $\approx 1 - 4\pi^2 W_{RMS}^2 \geq 0.8$

RMS $\approx \sqrt{\sum C_j^2}$

$W_{PV} \leq \lambda/4$

Defocus $\Delta y' \sim r_{EP}$

Coma $\Delta y' \sim r_{EP}^2\,y$

Astigmatism $\Delta y' \sim r_{EP}\,y^2$

Spherical $\Delta y' \sim r_{EP}^3$

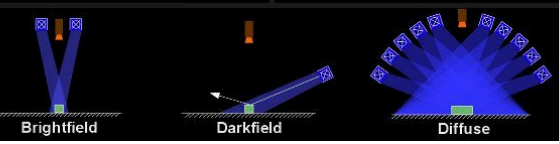
Distortion $\Delta y' \sim y^3$

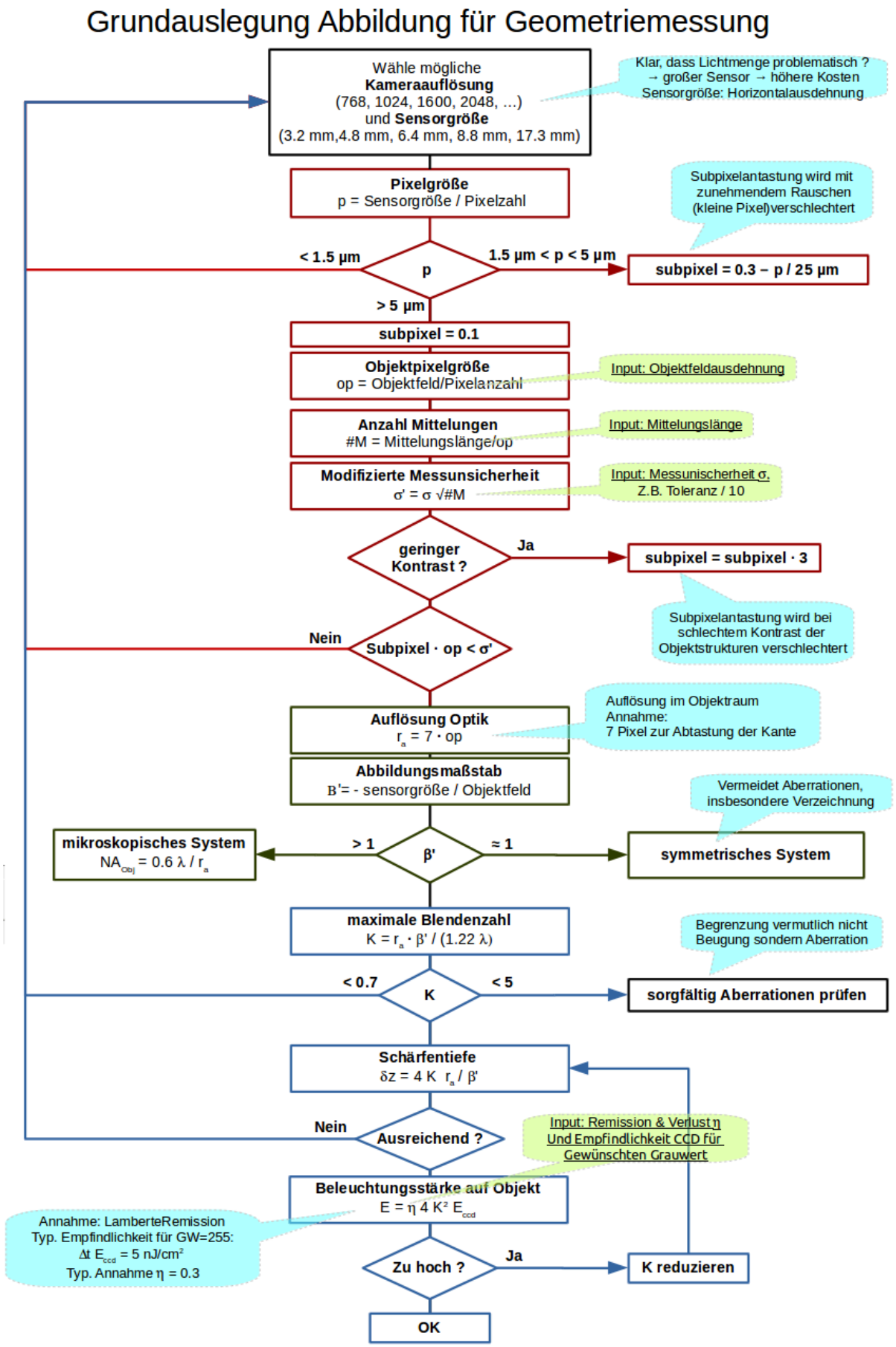
$r_{SA} \approx 0.07\cdot\frac{f'}{K^3}$

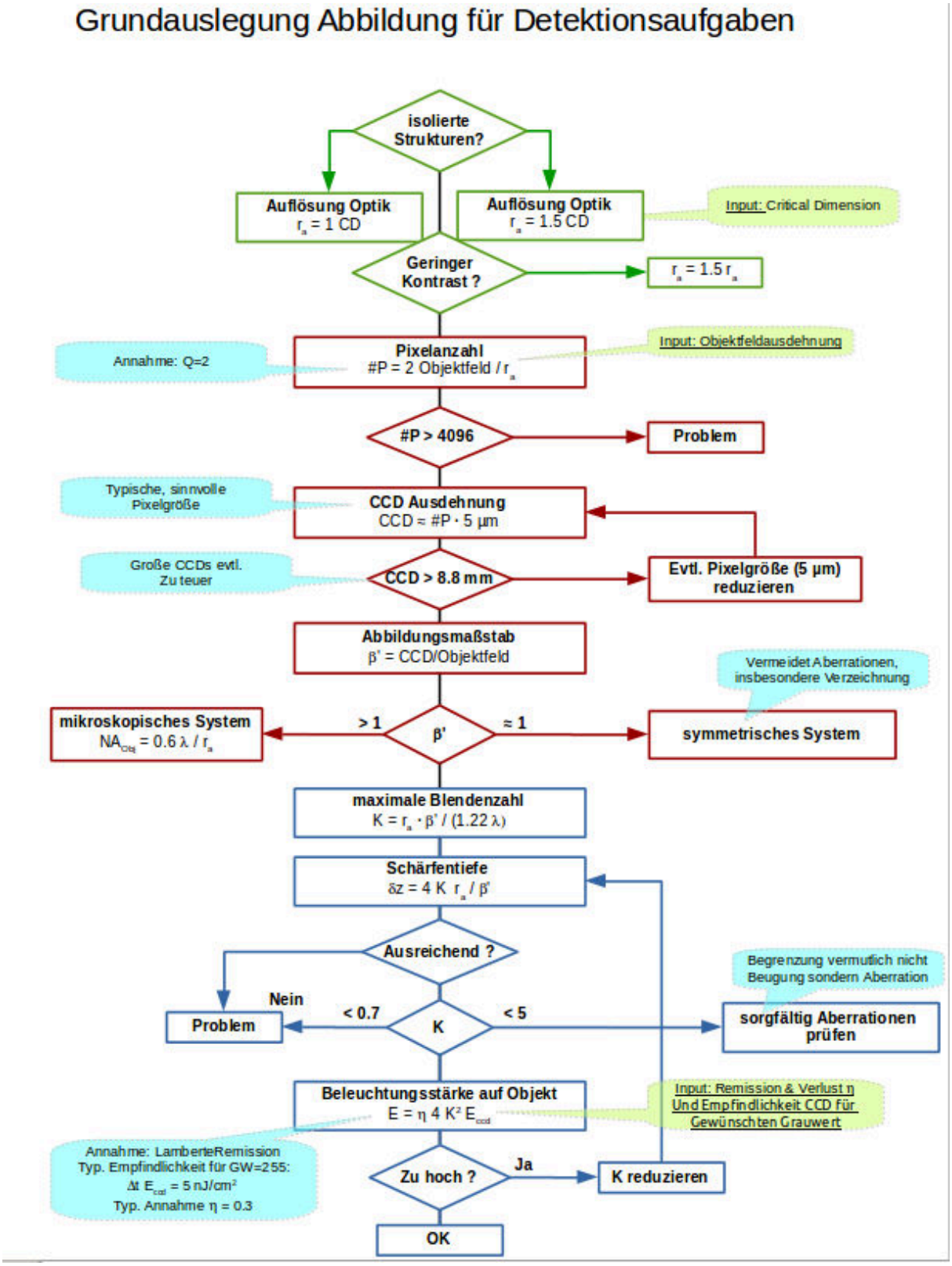
$r_{Spot} \approx \sqrt{EncEng\,\frac{5}{2}\pi W_{RMS}\lambda N\cdot K}$

Decenter $\Delta y' = \delta\cdot f'$

Decenter Wedge angle $\phi = \frac{\delta}{f'\cdot(n-1)}$




Grundauslegung Abbildung für Geometriemessung
Wähle mögliche
Kameraauflösung
(768, 1024, 1600, 2048, ...)
und Sensorgröße
(3.2 mm,4.8 mm, 6.4 mm, 8.8 mm, 17.3 mm)
Klar, dass Lichtmenge problematisch ?
→ großer Sensor → höhere Kosten
Sensorgröße: Horizontalausdehnung
Pixelgröße
p = Sensorgröße / Pixelzahl
Subpixelantastung wird mit
zunehmendem Rauschen
(kleine Pixel)verschlechtert
< 1.5 µm
p
1.5 µm < p < 5 µm
subpixel = 0.3 – p / 25 µm
> 5 µm
subpixel = 0.1
Objektpixelgröße
op = Objektfeld/Pixelanzahl
Input: Objektfeldausdehnung
Anzahl Mittelungen
#M = Mittelungslänge/op
Input: Mittelungslänge
Modifizierte Messunsicherheit
σ' = σ √#M
Input: Messunischerheit σ,
Z.B. Toleranz / 10
geringer
Kontrast ?
Ja
subpixel = subpixel · 3
Subpixelantastung wird bei
schlechtem Kontrast der
Objektstrukturen verschlechtert
Nein
Subpixel · op < σ'
Auflösung Optik
r_a = 7 · op
Auflösung im Objektraum
Annahme:
7 Pixel zur Abtastung der Kante
Abbildungsmaßstab
B'= - sensorgröße / Objektfeld
Vermeidet Aberrationen,
insbesondere Verzeichnung
mikroskopisches System
NA_Obj = 0.6 λ / r_a
> 1
β'
≈ 1
symmetrisches System
maximale Blendenzahl
K = r_a · β' / (1.22 λ)
Begrenzung vermutlich nicht
Beugung sondern Aberration
< 0.7
K
< 5
sorgfältig Aberrationen prüfen
Schärfentiefe
δz = 4 K r_a / β'
Nein
Ausreichend ?
Input: Remission & Verlust η
Und Empfindlichkeit CCD für
Gewünschten Grauwert
Beleuchtungsstärke auf Objekt
E = η 4 K² E_ccd
Annahme: LamberteRemission
Typ. Empfindlichkeit für GW=255:
Δt E_ccd = 5 nJ/cm²
Typ. Annahme η = 0.3
Zu hoch ?
Ja
K reduzieren
OK


Grundauslegung Abbildung für Detektionsaufgaben
isolierte Strukturen?
Auflösung Optik
r_a = 1 CD
Auflösung Optik
r_a = 1.5 CD
Input: Critical Dimension
Geringer Kontrast ?
r_a = 1.5 r_a
Annahme: Q=2
Pixelanzahl
#P = 2 Objektfeld / r_a
Input: Objektfeldausdehnung
#P > 4096
Problem
Typische, sinnvolle Pixelgröße
CCD Ausdehnung
CCD ≈ #P · 5 µm
Große CCDs evtl. Zu teuer
CCD > 8.8 mm
Evtl. Pixelgröße (5 µm) reduzieren
Abbildungsmaßstab
β' = CCD/Objektfeld
Vermeidet Aberrationen, insbesondere Verzeichnung
mikroskopisches System
NA_{Obj} = 0.6 λ / r_a
> 1
β'
≈ 1
symmetrisches System
maximale Blendenzahl
K = r_a · β' / (1.22 λ)
Schärfentiefe
δz = 4 K r_a / β'
Ausreichend ?
Nein
Problem
< 0.7
K
< 5
Begrenzung vermutlich nicht Beugung sondern Aberration
sorgfältig Aberrationen prüfen
Beleuchtungsstärke auf Objekt
E = η 4 K² E_{ccd}
Input: Remission & Verlust η Und Empfindlichkeit CCD für Gewünschten Grauwert
Annahme: LamberteRemission
Typ. Empfindlichkeit für GW=255:
Δt E_{ccd} = 5 nJ/cm²
Typ. Annahme η = 0.3
Zu hoch ?
Ja
K reduzieren
OK