

13. JAHRGANG  
AUGUST 2012

4

++++ SPECIAL +++ SPECIAL +++ SPECIAL +++++

# INSPECT

Methodik der industriellen  
Bildverarbeitung

Front-End der Bildverarbeitung

Back-End der Bildverarbeitung

Messtechnik und 3D

**GRUNDLAGEN**

DER BILDVERARBEITUNG  
OPTISCHEN MESSTECHNIK

PARTNER OF:

 **VISION**  **AUTOMATICA**



**GIT VERLAG**

[www.inspect-online.com](http://www.inspect-online.com)

# xiQ

SuperSpeed

USB 3.0

ximea



## xiQ USB 3.0 - Industriekameras

VGA mit 600 Bildern/Sek bis 4 MPixel mit 90 Bildern/Sek  
Sensoren mit hoher NIR-Empfindlichkeit  
Geringste Größe und Leistungsaufnahme  
Unterstützung für Windows und Linux  
Mit nur einer Platine perfekt für die OEM-Integration

Kompatibel mit über 30 gängigen Bildverarbeitungs-Bibliotheken ...

[ximea.com/usb3](http://ximea.com/usb3)

# Sehr geehrte Leserinnen, sehr geehrte Leser,

plant man, ein Gebäude zu errichten, ist es eine Selbstverständlichkeit, zunächst an ein sicheres Fundament zu denken. Anders als das eigentliche Gebäude entzieht sich dieses Fundament jedoch später den Blicken seiner Nutzer. Mit anderen Worten, die tragende Rolle des Fundaments wird nicht mehr wahrgenommen und auch wenig gewürdigt.

Bleibt man bei diesem Bild, widmet das Team der INSPECT im Falle der industriellen Bildverarbeitung dem Fundament seine ganz besondere Aufmerksamkeit. Mit der vorliegenden Sonderausgabe „Die Grundlagen der Bildverarbeitung und optischen Messtechnik“ lenken wir den Blick gezielt auf das Wissensfundament der Bildverarbeitung. In diesem Heft haben wir für unsere Leser die zwischen 2001 und 2010 in der INSPECT veröffentlichte Fachartikelserie von Professor Dr. Christoph Heckenkamp in einer kompakten und vom Autor überarbeiteten Fassung zusammengeführt.

Alle diejenigen, welche im Rahmen industrieller Projekte oder Forschungsvorhaben die Bildverarbeitung bereits einsetzen, oder planen, diese leistungsfähige Technologie in einem zukünftigen Projekt zu verwenden, finden hier einen verständlichen und sicheren Einstieg in die Thematik, um sich das erforderliche Basiswissen anzueignen.

Denn selbst ein gut geplantes Bauwerk von feinsten Eleganz und hoher Funktionalität wird nach kürzester Zeit in sich zusammenstürzen, wenn es kein tragfähiges Fundament hat. Schnell führt der Mangel an Grundwissen zu einer erheblichen Verschwendung wertvoller Arbeitszeit, indem Komponenten schlecht kombiniert werden oder sogar nach der Methode von Versuch und Irrtum vorgegangen wird. Was nützen intelligente Kameras der neuesten Generation, extreme Rechnerleistung und eine superschnelle Schnittstellentechnologie, wenn der Anwender nicht weiß, wie die Ergebnisse seiner Bemühungen zustande kommen und welche Parameter seine Ergebnisse beeinflussen.

Abgesehen von der Befriedigung des reinen Wissensdurstes kann der Leser somit auch einen praktischen Nutzen aus der Lektüre dieser INSPECT Sonderausgabe ziehen: Er spart Geld, gewinnt Zeit und kann beeindruckende Erfolge mit dem Einsatz der modernen Bildverarbeitung für sich verbuchen.

Die Redaktion der INSPECT dankt Herrn Professor Heckenkamp ganz besonders für seine Bereitschaft, dieses Vorhaben sowohl konzeptionell wie auch als Autor zu unterstützen, und es bis zur endgültigen Fertigstellung fachlich zu begleiten.

Ich wünsche Ihnen beim Lesen der „Grundlagen der Bildverarbeitung und optischen Messtechnik“ viel Vergnügen, einen großen Wissenszuwachs und den Nutzen bei der täglichen Arbeit.

Bernhard Schroth  
Chefredakteur der INSPECT

[www.inspect-online.com](http://www.inspect-online.com)



MICRO-EPSILON



## INDUSTRIELLE FARBMESSUNG

**NEU** colorCONTROL ACS 7000  
zur Qualitätskontrolle (390 - 780 nm)

- Berührungslos messen, 50 mm Abstand
- Online Farbmessung: 25 Hz - 2000 Hz
- Farbabstand  $\Delta E$  0,01 messen
- Extrem hohe Auflösung, 5 nm
- Reflektivitätsspektrum erfassen
- Bis zu 15 Farben einlernen und vergleichen
- Messgeometrie: 30°/0°
- Farbräume: XYZ; CIE L\*a\*b\*; L\*u\*v\*; wählbar
- Ethernet/EtherCAT, RS 422, Digital-Ausgänge
- Lichtarten: A, C, D65, D50, D75, E, F4, F7, F11

[www.micro-epsilon.de](http://www.micro-epsilon.de)

MICRO-EPSILON Eltrotec GmbH  
73066 Uchingen · Tel. 0 7161/98872-300  
[eltrotec@micro-epsilon.de](mailto:eltrotec@micro-epsilon.de)

# INHALT

## Methodik der industriellen Bildverarbeitung

- 6 Einführung: Die Bildverarbeitungskette
- 12 Blob-Analyse: Das „Arbeitspferd“ der Bildverarbeitung
- 16 Klassifizierung: Eine Klasse für sich
- 20 Front-End der Bildverarbeitung

## Front-End der Bildverarbeitung

- 20 Beleuchtung: In Lichtgeschwindigkeit gerechnet
- 26 Objektive: Abbildung mit Maß
- 32 Der Hintergrund: Tabula rasa
- 34 Kontrast und Schärfe: Einfach scharf!

## Back-End der Bildverarbeitung

- 40 Farbbildverarbeitung: Nachts sind alle Katzen grau
- 43 Filteroperationen: Glätten oder Falten?
- 46 Konturcodierung: Kettencode und Konturanalyse
- 49 Kantenerkennung: Klare Kante
- 51 Konturbasiertes Pattern-Matching: Konturverfolgung
- 53 Die Hough-Transformation: Auf der Geraden ins Ziel

## Messtechnik und 3D

- 58 Stereo-Vision: Nicht mono, sondern stereo
- 62 Das PMD-Prinzip: Ein magisches Auge
- 65 Kameramodelle: Von Kameraparametern zu Weltkoordinaten
- 68 Kamerakalibrierung: Weltkoordinaten im Griff
- 70 Triangulation: Nicht für die Schule, sondern für das Leben ....
- 72 Radiometrie: Strahlende Präsenz
- 75 Kameraempfindlichkeit: Dynamisches Rauschen

**31, 37–39, 56–57, 76–77** Produkte

**78** Index

**78** Impressum



REVOLUTIONARY  
PERFORMANCE  
FOR THE PRICE

LOW POWER  
MINIMUM HEAT  
HIGH IMAGE QUALITY

2K, 4K, 6K, 8K, 12K  
OPTIONS

GIGE AND CAMERA  
LINK INTERFACES

STANDS OUT.  
FITS IN.



**ONLY THE RACER PACKS SO MUCH PERFORMANCE INTO SO LITTLE SPACE.**

Basler's newest series of line scan cameras shrinks both the space required and the cost involved in securing superior speed, reliability and image quality. Measuring a slim 56mm, it's a perfect fit for most any system—including multi-camera arrangements. It also fits your budget, with its pared down price and by delivering exceptional performance under low voltage LED lighting.

Sometimes it's the smallest things that have the biggest impact.  
Learn more about the racer at [baslerweb.com](http://baslerweb.com).



read more



**BASLER**  
the power of sight

# Einführung

## Die Bildverarbeitungskette

Die Grundlagen eines technischen Fachgebiets sind diejenigen Bereiche, die für das Verständnis der wichtigsten Zusammenhänge unabdingbar sind und die sich im Lauf der Zeit nicht wesentlich ändern. Unter diesem Gesichtspunkt sind in der Zeitschrift INSPECT in den letzten Jahren die Artikel zu den „Grundlagen der Bildverarbeitung“ erschienen. Die Einführung stellt einige übergreifende Aspekte dar und bietet als „roten Faden“ für die Lektüre die sog. „Bildverarbeitungskette“ an.



► **Der Autor**  
**Prof. Dr. Christoph Heckenkamp**  
 Hochschule Darmstadt  
 Studiengang Optotechnik und  
 Bildverarbeitung  
 heckenkamp@h-da.de  
 www.fbmn.h-da.de

## Überblick

In diesem Heft ist eine Auswahl der Artikel zu den „Grundlagen der Bildverarbeitung“ zusammengestellt, die seit 2007 in der Zeitschrift INSPECT erschienen sind. Einige Beiträge sind zusammengefasst oder überarbeitet, bei manchen ist die Literaturliste aktualisiert, viele sind unverändert übernommen und geben so auch einen Einblick in die Entwicklung der letzten Jahre. Die Artikel sind einzeln erschienen und für sich allein verständlich. Dabei ist es auch in diesem Heft geblieben. Die einzelnen Kapitel bauen nicht aufeinander auf, und wer möchte, kann das Heft irgendwo aufschlagen und sofort mit der Lektüre beginnen. Auch die Literaturangaben sind weiterhin unmittelbar im Anschluss an jeden Artikel aufgeführt, damit sie ggf. sofort zur Hand sind. Auf ein Schlagwortverzeichnis haben wir verzichtet, denn beim Gesamtumfang von knapp 80 Seiten erscheint uns der Zugang über das Inhaltsverzeichnis akzeptabel.

Für die ersten beiden Artikel zu den Themen „Blob-Analyse“ und „Klassifizierung“ haben wir auf Beiträge aus dem Jahr 2002 zurückgegriffen. Sie bieten einen allgemeinen Einblick in die Methodik der industriellen Bildverarbeitung und sind als schneller Einstieg geeignet. Diese Lektüre erleichtert das Verständnis der übrigen Beiträge. Bei der Auswahl der Themen haben wir uns auf Wesentliches beschränkt und wollen damit einen einfachen Zugang zu den Kernbereichen des Gebiets eröffnen. Wer tiefere Einsichten erlangen möchte, wird sich damit zwar den langen Weg über Weiterbildungskurse, Lehrbücher oder ein Studium nicht ersparen können. Für einen kurzen Rundflug über das Gebiet können die Artikel jedoch vielleicht eine Orientierungshilfe bieten.

## Abgrenzung

Die Artikel in diesem Heft behandeln Themen, die in erster Linie für den Einsatz der Bildverarbeitung in der industriellen Fertigung von Bedeutung sind. Die entsprechende Disziplin wird

oft als „Industrielle Bildverarbeitung“ (IBV) bezeichnet, im englischen Sprachgebrauch wird der Oberbegriff „machine vision“ benutzt. Die dominierende Anwendung in der Produktion ist die Qualitätskontrolle im Fertigungstakt mit dem Ziel einer 100%-Prüfung an verschiedenen Stellen der Wertschöpfungskette. Beispielsweise wird geprüft, ob das Gewinde einer Getränkeflasche in Ordnung ist, bevor der Schraubverschluss aufgesetzt wird. Die 100%-Kontrolle ist somit ein Instrument, das sowohl eine effiziente Fertigung ermöglicht als auch gewährleistet, dass beim Endkunden nur Gutteile ankommen. Das typische Portfolio dieser klassischen IBV besteht aus der Anwesenheitskontrolle, Lage- und Positionserkennung, Vollständigkeitsprüfung, Vermessung, Barcode- und Flächencoderkennung, optischer Zeichenerkennung (OCR) und Oberflächenprüfung. Die hohen Taktraten und damit die Rationalisierung und Produktivitätssteigerung in der industriellen Fertigung sind nur möglich, weil für diese Prüfaufgaben erprobte Bildverarbeitungslösungen zur Verfügung stehen. Mit manueller Kontrolle wäre eine 100%-Prüfung in vielen Industrien gar nicht realisierbar. Automatische Systeme ermöglichen zudem die lückenlose Protokollierung des Prüfprozesses, eine wichtige Eigenschaft, wenn die Rückverfolgbarkeit gefordert ist.

Zwei weitere Bereiche haben in letzter Zeit das Anwendungsspektrum der IBV erweitert. Ein wichtiges Gebiet für Vision-Anwendungen ist inzwischen die Handhabungstechnik, speziell die Robotik. Viele Robotersysteme, die z.B. bei der Verpackung, aber auch z.B. beim Schweißen oder bei der Kontrolle von Kleberaunen, Schweißnähten usw. eingesetzt werden, arbeiten mit der Unterstützung von Bildverarbeitungssystemen. Häufig handelt es sich dabei um Mess- und Prüfaufgaben im dreidimensionalen Raum. Daher ist hier das Gebiet der 3D-Bildverarbeitung oder Stereo-Vision von Bedeutung. Der zweite Bereich, der mehr und mehr von der Bildverarbeitung durchdrungen wird, ist die Mess-, Steuer- und Regeltechnik (MSR). Bei einer klassischen Prüfaufgabe in der laufenden Produktion interessiert

zwar nur die Aussage „in Ordnung“ oder „nicht in Ordnung“. Diese Aussage wird aber auf der Basis von Messwerten getroffen, die in der Bildverarbeitung als „Merkmale“ bezeichnet werden. Typische Merkmale sind Fläche und Umfang eines Teils oder der Durchmesser einer Bohrung. Es liegt nahe, solche Parameter in einer Produktion über die Zeit zu verfolgen und einzugreifen, wenn ein Trend erkennbar ist, dass eine Messgröße aus dem Toleranzbereich herausläuft. Wer realisiert hat, dass Bildverarbeitung eine Messtechnik ist, wird die verfügbaren Messwerte bald zur systematischen Optimierung des Prozesses und zur Regelung verwenden wollen. Daher ist seit einigen Jahren zu beobachten, dass etablierte Firmen aus dem Bereich der MSR-Technik sich den Methoden der Bildverarbeitung nähern, indem sie ihre Lichtschranken oder Abstandssensoren mit Zeilen- oder Flächensensoren ausstatten, einen Controller integrieren und damit zu „intelligenten“ Systemen aufrüsten.

Auch außerhalb der Industrie gewinnen Bildverarbeitungsanwendungen an Bedeutung. Ein typisches Beispiel ist die Kontrolltechnik im deutschen LKW-Mautsystem, einem Vorzeigeprojekt der deutschen Bildverarbeitungsindustrie. Hier werden bei Nacht und Nebel, bei strahlendem Sonnenschein und bei strömendem Regen, im fließenden Verkehr und im Stop-and-go-Betrieb die Kennzeichen der vorbeifahrenden Fahrzeuge (und noch einige andere Parameter) erfasst und gelesen, eine Aufgabe, vor der man gar nicht genug Respekt haben kann. Weitere Anwendungen außerhalb der Industrie sind die Barcodeerkennung im Supermarkt, kamerabasierte Fahrerassistenzsysteme, Apps zum Lesen von Flächencodes mit Smartphones oder die Bestimmung der Laufwege von Fußballspielern aus Videosequenzen. Diese Anwendungen erfordern oft weit komplexere Methoden und können deutlich schwieriger sein als die typischen Verfahren der IBV. Sie fallen ebenfalls noch unter den Begriff der „machine vision“, und die Themen der Beiträge in diesem Heft erleichtern auch den Einstieg in dieses Gebiet. Eine qualifizierte Gesamtschau der „machine vision“ würde jedoch den Rahmen dieses Sonderhefts sprengen.

Nicht abgedeckt sind hier außerdem alle Fragestellungen im Zusammenhang mit der Programmierung, der Softwareentwicklung und dem Software-Engineering. Sie gehören zu den technischen Grundlagen vieler Ingenieurdisziplinen und sind zudem ständig im Wandel. Auch die Problematik der Schnittstellen bzw. Bussysteme für die Datenübertragung von der Kamera zur Recheneinheit wurde ausgespart. Auf diesem Gebiet gibt es ebenfalls eine rapide technische Entwicklung, die zudem von Marketing-Aktivitäten überlagert wird. Die Möglichkeiten der Implementierung eines BV-Systems in Hardware – als PC-System, als intelligente Kamera, als Kompakt-system auf Signalprozessor-Basis, mit FPGA-Unterstützung usw. – werden ebenso wenig behandelt wie die Struktur von Bilddatenformaten. Die genannten Themenbereiche sind zwar sämtlich für die Realisierung von Bildverarbeitungslösungen von Bedeutung. Sie gehören aber nicht zum Kern der Grundlagenkenntnisse, über die potentielle Anwender der IBV im Dialog mit Anbietern verfügen sollten. Ebenso wenig werden betriebswirtschaftliche und organisatorische Fragen behandelt, wie z.B. die Amortisationszeiten von BV-Prüfanlagen oder die Abwicklung von BV-Projekten von der Anfrage bis zum site-acceptance-test.

Ein wesentlicher Teilbereich ist jedoch bewusst ausgespart: die Grundlagen der Kamera- und Detektortechnik. Dieses Gebiet hat sich in den letzten zehn Jahren sehr dynamisch entwickelt. Im Detail ist damit eine erhebliche Vielfalt von verschiedenen Bildaufnahmeverfahren und Kamertypen entstanden. Ein kurzer Grundlagenartikel ist selbst für eine Übersicht nicht mehr angemessen. In einigen Beiträgen wer-

den spezielle Aspekte behandelt, z.B. das PMD-Prinzip oder die Bayer-Filtermatrix. Wer sich den erforderlichen Überblick verschaffen will, kann auf die diversen „white papers“ der einschlägigen Hersteller zurückgreifen oder sich in die mittlerweile gut zugängliche Literatur zu diesem Thema vertiefen [1] [2].

## Struktur des Heftes

Die Anordnung der Artikel unterstützt eine durchgehende Lektüre von der ersten bis zur letzten Seite. Die beiden ersten Beiträge geben einen Überblick über die Methodik der Bildverarbeitung. Sie zeigen, dass die grundlegenden Verfahren für manche Prüfaufgaben sehr einfach sein können. Zugleich wird aber auch deutlich, dass eine Bildverarbeitungslösung ein komplexes System aus Zuführung, Beleuchtung, Optik,ameratechnik, Integration in den Fertigungsfluss, Rechner- und Softwareentwicklung ist. Diese Komponenten müssen zu einem funktionsfähigen System zusammengeführt werden, das an die jeweilige Problemstellung angepasst sein muss. Der Beitrag zur „Blob-Analyse“ erläutert, wie Messgrößen aus einem Bild entnommen werden können, und der Artikel zur „Klassifizierung“ befasst sich mit der schwierigen Problematik, dass schließlich in einer Prüfstation eine Entscheidung gefällt werden muss, ob der Prüfling in Ordnung ist oder nicht.

Jeder Messtechniker weiß, dass aus einem Signal nur die Information gewonnen werden kann, die eingangs mit erfasst wird. Die vier Artikel des zweiten Kapitels befassen sich daher mit Problemen am „front-end“ eines Bildverarbeitungssystems. Bevor das Bild als Datei vorliegt, sind bereits wesentliche Bildverarbeitungsprozesse abgelaufen, die massiven Einfluss auf die Bilddaten haben. Von überragender Bedeutung ist die Beleuchtung, die Prüfmerkmale herausarbeiten und Störeinflüsse unterdrücken soll. Aber auch die Abbildungsoptik hat Auswirkungen auf die Information, die letztlich in der Bilddatei enthalten ist.

Die dann folgenden sechs Artikel behandeln einige Methoden, die unmittelbar mit der Auswertung der Bilddaten zu tun haben, wenn das Bild bereits aufgenommen ist und im Rechner als Datei vorliegt. Gelegentlich wird „Bildverarbeitung“ auf solche Operationen auf der Bilddatei reduziert und erscheint als Teilgebiet der Mathematik oder der Informatik. Diese Vorstellung ist derartig irreführend und für industrielle Anwendungen geradezu sinnentstellend, dass man ihr gar nicht deutlich genug widersprechen kann. Bildverarbeitung in der industriellen Anwendung ist eine Systemtechnik, in diesem Fall sogar eine interdisziplinäre Aufgabe, die auf Kenntnissen aus verschiedenen Fachrichtungen aufbaut. Es handelt sich demnach um eine typische Ingenieurdisziplin. Es genügt definitiv nicht, irgendeine Kamera und eines der vielen BV-Softwarepakete zu kaufen, um eine Bildverarbeitungslösung zu realisieren. Manche Anwender sind dieser Illusion verfallen und mussten unter vielfältigen Schwierigkeiten feststellen, dass die industrielle Bildverarbeitung eine anspruchsvolle, erklärungsbedürftige Technologie ist, die breite Fachkenntnisse und viel Erfahrung erfordert. Wenn überhaupt eine Zuordnung zu einem klassischen Ingenieurfach sinnvoll ist, kann man die Bildverarbeitung in der industriellen Anwendung am besten als Teilgebiet der Automatisierungstechnik oder als spezielle Ausprägung der Mechatronik auffassen. Spätestens im nächsten Jahrzehnt wird sich in diesem Fachgebiet die Mehrzahl der Ingenieure mit Bildverarbeitungssystemen als selbstverständlichen Komponenten der MSR-Technik konfrontiert sehen und auf Grundkenntnisse der Bildverarbeitung nicht verzichten können. Die nicht-industriellen Anwen-



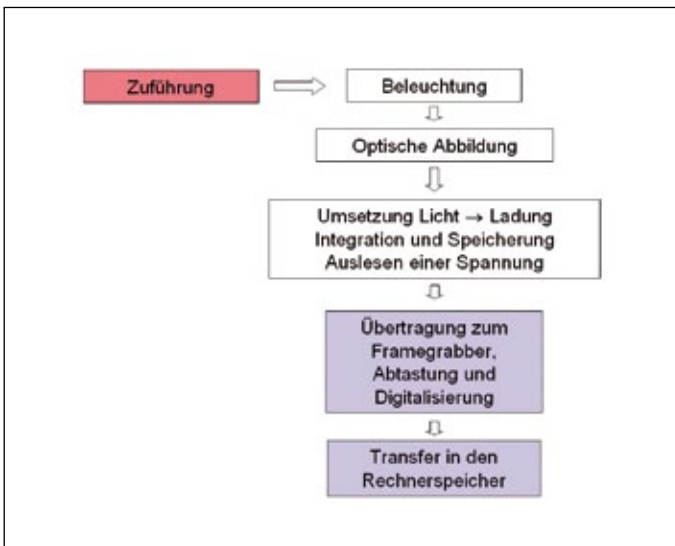


Abb. 1: Das „front-end“ der Bildverarbeitungskette

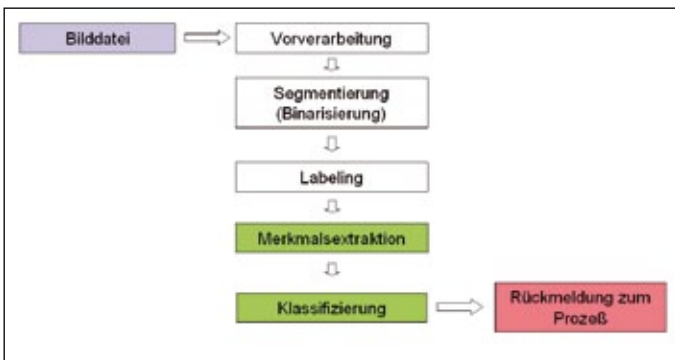


Abb. 2: Das „back-end“ der Bildverarbeitungskette

dungen der Bildverarbeitung gehen darüber weit hinaus und werden von der Automatisierungstechnik gar nicht erfasst, so dass sich die „machine vision“ über kurz oder lang zu einer eigenständigen Ingenieurdisziplin entwickeln wird. In Darmstadt geben wir unseren Studenten im Studienprogramm „Optotechnik und Bildverarbeitung“ [3] das Handwerkszeug dafür schon seit einigen Jahren mit auf den Weg.

Im vierten Kapitel werden einige Themen aufgegriffen, bei denen Bildverarbeitung als Methode der Messtechnik deutlich in Erscheinung tritt. Offensichtlich ist dieser Aspekt, wenn es um die Gewinnung von 3D-Daten im realen Raum aus den zweidimensionalen Bilddaten einer Kamera geht. Aber auch die Bestimmung von relativen und absoluten „Helligkeiten“ ist als messtechnische Fragestellung in die Bildverarbeitung ein-

gedungen und hat zur intensiven Auseinandersetzung mit den Eigenschaften der Bildsignale angeregt, schließlich sogar zu einer objektiven Definition des Begriffs Empfindlichkeit geführt. Es ist ein Indiz für die weiter fortschreitende Professionalisierung der IBV, dass man auch diese Fragestellungen mittlerweile zu den Grundlagen der Bildverarbeitung zählen muss und sich ihnen nicht mehr entziehen kann.

### Der „rote Faden“

Wer das Heft nicht durchgängig von vorn bis hinten lesen möchte, sondern eine andere Reihenfolge vorzieht, kann sich bei der Lektüre gut an der Vorstellung einer „Bildverarbeitungskette“ orientieren. Bei vielen Bildverarbeitungssystemen kann man sich die Abläufe in der Prüfstation wie die lineare Abarbeitung von Teilschritten vorstellen, die wie die Glieder einer Kette ineinander greifen. Abbildung 1 zeigt schematisch das „front-end“ dieser Bildverarbeitungskette, Abbildung 2 das „back-end“. Ein Prüfling wird zunächst der Prüfstation zugeführt, dann beleuchtet, die Szene wird optisch auf einen Detektor abgebildet, und die Helligkeitsverteilung wird dort über die Pixelstruktur räumlich diskret abgetastet und in ein Ladungsmuster umgesetzt. Die Ladungen werden dann in Spannungen umgesetzt und bei einer Analogkamera als Video-Signal an einen Frame-Grabber übertragen. Dort wird die Pixelinformation durch Abtastung des Analogsignals zurückgewonnen und die Amplitude als Maß für die Helligkeit digitalisiert. Die Daten werden anschließend geordnet in einem Bilddatenfile abgelegt. Bei einer Digitalkamera werden die Signalspannungen aus dem Detektor direkt digitalisiert und als digitaler Datenstrom an die Recheneinheit übertragen. Das „back-end“ beginnt mit dem digitalen Bilddatenfile. Meist werden die Bilddaten vorverarbeitet, z. B. durch Filterung oder Entzerrung, anschließend wird der Vordergrund vom Hintergrund getrennt (Segmentierung), danach erfolgt oft eine Zusammenhangsanalyse, so dass auf die verschiedenen voneinander getrennten Objekte im Bild zugegriffen werden kann (labeling). Im nächsten Schritt werden charakteristische Messgrößen für die Objekte ermittelt (Merkmalsextraktion), die als Basis für eine Entscheidung (Klassifizierung) dienen, z. B. „in Ordnung“, „nicht in Ordnung“. Am Ende wird das Ergebnis an den Prozess zurückgemeldet, z. B. über eine SPS.

Mit dem Hilfsmittel der Bildverarbeitungskette wird besonders deutlich, dass die wesentliche Aufgabe der industriellen Bildverarbeitung in einer ungeheuren Reduktion der Information besteht. Die komplexe Szene aus der realen Welt, die sowohl in der Ortsauflösung als auch in der Helligkeit de facto analog ist und vielfache weitere Information trägt, wird im Extremfall auf ein einziges Bit reduziert, nämlich auf die Information „in Ordnung“ oder „nicht in Ordnung“. Diese Informationsreduktion beginnt weit vor der Ablage der Bilddaten im Rechnerpeicher.



Abb. 3: Eine reale Szene wird optisch auf den Bildaufnehmer abgebildet (links). Durch die räumliche Abtastung und die Digitalisierung der Helligkeitsverteilung wird die Information drastisch reduziert (Mitte) und in einem Bilddatenfile geordnet abgelegt (rechts).



# MIT USB 3 uEye® CP KAMERAS AUF DER ÜBERHOLSPUR

Schnell, leistungsstark, zukunftssicher



BEREITS SCHON JETZT  
LIEFERBAR!

**ids**

■ **Wir sind USB.** Kein anderer Hersteller hat so viel Erfahrung in der Entwicklung und im Einsatz von USB-Industriekameras. In der neuen Generation unserer uEye Kamerafamilie mit USB 3.0 stecken die Erkenntnisse von einem Jahrzehnt USB-Technologie. Denn Erfahrung macht den Unterschied.

[www.ids-imaging.de/usb3](http://www.ids-imaging.de/usb3)



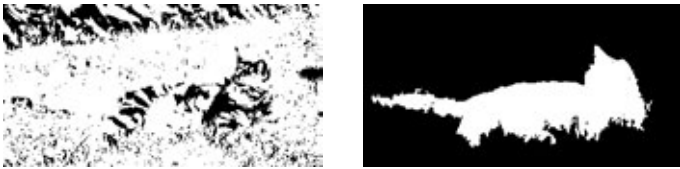


Abb. 4: Eine binarisierte Aufnahme einer Katze in einer realen Szene (links). Die Segmentierung des Binärbildes in Objekt und Hintergrund (rechts) ist mit den Mitteln der Bildverarbeitung nicht zu leisten, der Mensch erkennt jedoch mit einem Blick, dass es sich um eine Katze handelt.



Abb. 5: Ein Panther-Butt in der Bucht von El Quadim, El Quseir, Ägypten

Aus einer realen Szene, die z.B. etwa so aussehen könnte wie der linke Teil der Abbildung 3, wird bereits im „front-end“ allein durch die räumlich diskrete Abtastung und die Umsetzung in ein digitales Graustufenbild ein großer Teil der tatsächlich vorhandenen Information entfernt. Das mittlere Teilbild der Abbildung 3 verdeutlicht dies anhand eines Graustufenbildes mit bewusst schlechter Pixelauflösung. Letztlich wird die Information dann auf die digitalen Helligkeitswerte an diskreten Stellen der Bildebene reduziert, wie sie in der Tabelle im rechten Teilbild der Abbildung 3 dargestellt sind. Nur in dieser dramatisch reduzierten Form, nämlich räumlich in Pixeln (oder Kacheln) diskretisiert und in der Intensität digitalisiert, können Bildinformationen überhaupt in einem Rechner verarbeitet werden. Das eigentliche Problem der Bildverarbeitung besteht nun darin, dass zwar Information verworfen werden darf, sogar verworfen werden muss, denn der größte Teil der Bildinformation ist uninteressant. Die „richtige“ Information soll jedoch übrig bleiben! Welche Information relevant ist, lässt sich aber nicht pauschal festlegen, sondern hängt von der jeweiligen Anwendung ab. Die Methodik der IBV besteht nun darin, jede Systemkomponente abgestimmt auf die jeweilige Prüfaufgabe so auszulegen, dass irrelevante Information unterdrückt wird und relevante Information möglichst sicher und einfach extrahiert werden kann. Das beginnt bei der Zuführung, geht weiter über die Beleuchtung, die Abbildungsoptik, über alle folgenden Elemente der Bildverarbeitungskette bis zur Auswahl und Implementierung der angemessenen Algorithmen für die Verarbeitung der Bilddatei.

## Bildverarbeitung als Messtechnik

Bildverarbeitungssysteme funktionieren grundsätzlich anders als das menschliche visuelle System. Kein Mensch kann die relative Strahlungsleistung nennen, die auf ein bestimmtes Stäbchen seiner Retina fällt. In einem Bildverarbeitungssystem hingegen sind die digitalen Grauwerte sämtlicher Pixel präsent. Die industrielle Bildverarbeitung und auch die nicht-industriellen Anwendungen der „machine vision“ setzen bei der Bildanalyse in der Regel an Grauwertdifferenzen im Bild an, die letztlich auf die Beleuchtung zurückgehen. Bei der IBV im Speziellen und der „machine vision“ im Allgemeinen handelt es sich um eine Messtechnik. Ein IBV-System liefert präzise Messergebnisse, die ein biologisches Seh-System gar nicht liefern kann und auch nicht liefern muss. Die Stärke natürlicher visueller Systeme liegt in ihrer hervorragenden Klassifizierungsleistung. Ein Beispiel ist in Abbildung 4 dargestellt, eine binarisierte Aufnahme einer natürlichen Szene. Ein menschlicher Betrachter wird problemlos erkennen, dass es sich hier um eine Katze handelt – spätestens, wenn man es ihm gesagt hat. Auf der Basis des Binärbilds ist mit den Methoden der Bildverarbeitung nicht einmal die Trennung in Vordergrund und Hintergrund (wie im rechten Teilbild, das aus dem ursprünglichen Farbbild gewonnen wurde) zu leisten. Der Mensch jedoch kann aus dem Binärbild sogar unmittelbar erkennen, dass es sich um eine Katze handelt. Diese Fähigkeit beruht auf Vorwissen – die meisten Menschen haben schon viele Katzen gesehen. Dass Vorwissen dabei eine wichtige Rolle spielt, ist für manche Leser vermutlich in Abbildung 5 erkennbar. Diese Unterwasser-Aufnahme zeigt einen Panther-Butt aus dem roten Meer, den man häufig nur dann im Bild findet, wenn man ihn schon einmal gesehen hat. Angesichts dieser Fähigkeiten des Menschen sollte man sich der „machine vision“ besser nicht mit der Erwartung nähern, dass mit einem technischen Vision-System ähnliche Leistungen vollbracht werden können. Potentielle Anwender sind oft enttäuscht, wenn bestimmte Prüfaufgaben mit den Methoden der Bildverarbeitung nicht gelöst werden können, obwohl „man das doch sofort sieht“. Andererseits sind potentielle Anwender oft überrascht, welche Fülle von quantitativen Daten mit geringem Aufwand aus Bildern von Objekten gewonnen werden können, wie genau und zuverlässig Positionen und Drehlagen gemessen werden oder auch nur wie schnell ein IBV-System mehrere Dutzend Objekte in einem Bild lokalisieren und zählen kann. Bildverarbeitung ist eine Messtechnik und liefert präzise Messergebnisse auf der Basis von Grauwertdifferenzen, die im Wesentlichen durch die Beleuchtung entstehen, biologische visuelle Systeme hingegen bieten einen Überlebensvorteil auf der Basis von Vorwissen.

## Weitere Quellen

Diese Zusammenstellung von Grundlagenartikeln erhebt nicht den Anspruch auf Vollständigkeit, sondern möchte das Wesentliche knapp und möglichst klar darstellen. Daher waren in mancher Hinsicht Kompromisse nicht zu vermeiden. Anders als in einem Lehrbuch werden nicht alle verwendeten Begriffe sorgfältig eingeführt, sondern gelegentlich als bekannt vorausgesetzt. Manche Zusammenhänge werden nicht so ausführlich erläutert wie es unter didaktischen Gesichtspunkten sinnvoll wäre. Auf mathematische Begriffsbildungen wird so weit wie möglich verzichtet, und einige Beispiele, die zum Verständnis nützlich sein könnten, werden im Text nur kurz angerissen und weder weiter ausgeführt noch mit Bildern oder Graphiken ergänzt. Es kann hilfreich sein, solche



Artikel mit einem Schreibstift in der Hand zu lesen und sich die Zusammenhänge anhand selbst erdachter einfacher Beispiele oder Skizzen zu verdeutlichen. Manche Unklarheit wird sich auch mit der Lektüre weiterer Artikel auflösen. Wer tiefer in die Materie einsteigen will, kann vielleicht von den einschlägigen Weiterbildungskursen profitieren [4] oder mag sich mit einem Lehrbuch [5-7] beschäftigen. Auch die Vortragsforen auf diversen Fachmessen [8] bieten Gelegenheit, sich weiter in die Grundlagen der Bildverarbeitung einzuarbeiten.

### Danksagung

Für die hervorragende Unterstützung durch das gesamte INSPECT-Team und die angenehme Zusammenarbeit während nunmehr über zehn Jahren möchte ich mich herzlich bedanken. Viele der Gedanken und Beispiele in den Artikeln sind im Zusammenhang mit meinen Vorlesungen und Laborveranstaltungen zur industriellen Bildverarbeitung im Studiengang „Optotechnik und Bildverarbeitung“ an der Fachhochschule Darmstadt (nun umbenannt in Hochschule Darmstadt) entstanden. Ohne die zahlreichen Anregungen aus der Arbeit mit den Studenten, aus vielen interessanten Gesprächen mit meinen Kollegen im Studiengang, namentlich vor allem mit Prof. Dr. Scharfenberg und Prof. Dr. Sandau, sowie mit unseren Laboringenieuren, Dipl.-Ing. Udo Häberle und Dipl.-Ing. Frank Friehl, und nicht zuletzt auch ohne die viel-

fältigen Kontakte zu unseren Firmenpartnern aus der Industrie in Form von Abschlussarbeiten und Projekten wäre die intensive und kontinuierliche Verfolgung der Entwicklungen in diesem Fachgebiet gar nicht möglich. Allen, die daran Anteil haben, bin ich sehr dankbar, dass ich in einem solchen Umfeld arbeiten kann.

### Literatur

- [1] G. C. Holst, T. S. Lomheim, CMOS/CCD-Sensors and Camera Systems, SPIE-Press, 2nd ed. 2011
- [2] H. Mattfeldt, Camera Systems in Machine Vision, in A. Hornberg (ed.), Handbook of Machine Vision, Wiley-VCH 2006
- [3] [www.fbmh-h-da.de](http://www.fbmh-h-da.de)
- [4] Weiterbildungskurse zur Bildverarbeitung bieten z.B. Optence, das Optik-Kompetenznetz Hessen-Rheinland-Pfalz, [www.optence.de](http://www.optence.de), oder das Heidelberger Bildverarbeitungsforum, [www.bv-forum.de](http://www.bv-forum.de)
- [5] Christian Demant, Bernd Streicher-Abel, Axel Springhoff, Industrielle Bildverarbeitung, 3. Auflage, Springer 2011
- [6] Wilhelm Burger, Mark James Burge, Digitale Bildverarbeitung, 2. Auflage, Springer 2006
- [7] Carsten Steger, Markus Ulrich, Christian Wiedemann, Machine Vision Algorithms and Applications, Wiley-VCH 2008
- [8] Überschaubare und empfehlenswerte Fachmessen mit hohem Anteil an Bildverarbeitungsfirmen sind die VISION, die CONTROL, die AUTOMATICA und die SPS

## Embedded Solutions für alle OEM-Applikationen Klein, kompakt, leistungsstark.



### Ob maßgeschneidert oder Standardmodell:

Die embedded Smart Kameras von VC sind frei programmierbar und bieten dadurch höchstmögliche Flexibilität in Echtzeit.



Erfahren Sie mehr über unsere vielfältige Produktpalette im Internet.



# Blob-Analyse

## Das „Arbeitspferd“ der Bildverarbeitung

Bei vielen Fragestellungen der industriellen Bildverarbeitung befinden sich im Sichtbereich der Kamera mehrere verschiedene Objekte, deren Position, Größe oder Form bestimmt werden müssen. In solchen Fällen wird häufig die „Blob-Analyse“ eingesetzt, ein weitverbreitetes Verfahren der Bildverarbeitung, das von vielen Bildverarbeitungsbibliotheken unterstützt wird.

Als Beispiel ist in Abbildung 1 ein Tablettenblistert zu sehen, bei dem zwei Nöpfe nicht ordnungsgemäß mit Tabletten gefüllt sind. Im einfachsten Fall interessiert hier lediglich die Anzahl der vorhandenen Tabletten im Bild. Diese Prüfaufgabe ist schon mit sehr einfachen Methoden, z.B. mit „Pixelzählen“, zuverlässig lösbar, weil die Sollgröße der Prüfobjekte mit geringen Toleranzen bekannt ist [1]. Das Verfahren versagt jedoch, wenn Prüfobjekte mit unterschiedlichen Flächen und an undefinierten Positionen im Bild vorliegen. Bereits das einfache Zählen von isolierten Objekten im Bild, etwa bei einer Aufnahme mehrerer regellos vorliegender Schrauben und Muttern, erfordert weitergehende Methoden der Bildverarbeitung.

Eine Analyse der „Blobs“ liefert in solchen Fällen außer der Anzahl der voneinander getrennten Objekte im Bild zusätzlich eine Reihe von quantitativen Objektmerkmalen, die Rückschlüsse auf Prüfgrößen für die Sichtprüfung ermöglichen oder für Anwendungen in der Robotik interessant sind. Dazu gehören z.B. die Position, die Fläche oder der Umfang von Objekten.

### Segmentierung und Pixelzählen

Ein „Blob“ ist eine zusammenhängende Menge von Punkten im Vordergrund eines Binärbildes. Wenn sich mehrere voneinander getrennte Objekte in einem Bild befinden, kann man diese Objekte als einzelne „Blobs“ auffassen und analysieren. Voraussetzung für die Blob-Analyse ist zunächst die Trennung von „Vordergrund“ und „Hintergrund“. Diese Unterscheidung, die sog. „Segmentierung“, wird meist als Trennung der eigentlich interessierenden Objekte vom Untergrund verstanden. Beispielsweise sollen Tabletten vom Tablettenblistert oder Schrauben von der Montagefläche unterschieden werden. In einem Bild bedeutet diese Forderung, dass von jedem Pixel bekannt sein muss, ob es zum Vordergrund, also zu den Objekten, oder zum Hintergrund gehört.

Zu diesem Zweck erzeugt man in der Bildverarbeitung aus dem Quellbild ein neues Bild mit denselben Abmessungen und codiert die Zugehörigkeit eines Pixels zum Vorder- oder Hintergrund in seinem Grauwert. Beispielsweise werden Vordergrundpixel auf Weiß und Hintergrundpixel auf Schwarz gesetzt. Das Ergebnis ist ein Binärbild, in dem nur noch die beiden Grauwerte für Schwarz und Weiß vorkommen. Die Segmentierung eines Bildes in Vorder- und Hintergrund ist manchmal alles andere als einfach. In vielen Fällen führt jedoch bereits eine einfache globale Binarisierung zum Ziel. Dabei wird der Grauwert jedes Pixels mit einer vorher festgelegten Schwelle verglichen. Liegt der Grauwert des aktuellen Pixels über der Schwelle, wird es auf Weiß gesetzt, anderenfalls auf Schwarz.

Dieses Verfahren funktioniert bei dem oben genannten Beispiel, dem Tablettenblistert, ausgezeichnet.

Schon nach diesem Bildverarbeitungsschritt können viele einfache Prüfaufgaben bearbeitet werden. Die Zahl der weißen Pixel in Abbildung 1 ist ein Maß für die Fläche aller Tabletten im Bild. Wenn ein Algorithmus die weißen Pixel zählt und das Ergebnis mit der Anzahl der Pixel für ein Bild vergleicht, in dem der Blistert korrekt mit acht Tabletten befüllt ist, stellt sich heraus, ob wirklich acht Tabletten vorhanden sind. Beispielsweise haben die Tabletten im Binärbild in Abbildung 1 jeweils eine Fläche von etwa 660 Pixeln. Die exakte Anzahl hängt von der Position der Tablette im Bild, von den Beleuchtungsbedingungen sowie von der perspektivischen Verzerrung, der Verzeichnung des Objektivs und vom Rauschen ab. In Abbildung 1 sind die Aufnahmebedingungen so optimiert, dass die Abweichung der Flächen der Tabletten vom Mittelwert lediglich  $\pm 8$  Pixel beträgt, also weniger als  $\pm 1,5\%$ . Daher kann hier allein durch Zählen der Vordergrundpixel bereits plausibel überprüft werden, ob alle Tabletten vorhanden sind. Wenn zusätzlich die Sollpositionen der Tabletten im Bild hinreichend genau bekannt sind, kann das Verfahren separat auf kleine „regions of interest“ angewendet werden, die jeweils eine einzelne Tablette enthalten. Damit kann die Vollständigkeit einzelner Tabletten geprüft werden. Formabweichungen, die die Gesamtfläche nicht verändern, oder der Bruch einer Tablette können auf diese Weise jedoch nicht erkannt werden.

### Labeling

Nach der Segmentierung ist zwar für jedes Pixel aus dem zugeordneten Grauwert „Weiß“ oder „Schwarz“ sofort ersichtlich, ob es zu einem Objekt oder zum Untergrund gehört. Wenn aber mehrere weiße Objekte auf schwarzem Untergrund vorliegen, ist nicht unmittelbar klar, zu welchem Objekt ein bestimmtes

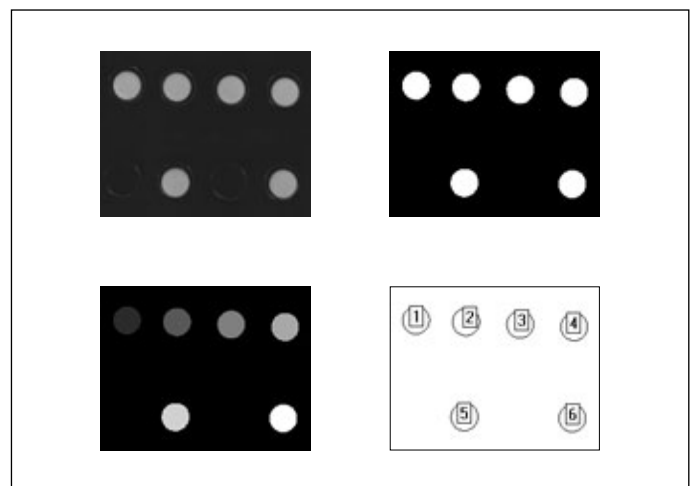


Abb. 1: Oben links das Grauwertbild eines Tablettenblisters, daneben das zugehörige Binärbild für die Grauwertschwelle 100. Unten zwei Darstellungen des Labelbildes.



weißes Pixel gehört. Wenn ein Bildverarbeitungsalgorithmus ein weißes Pixel aus dem Bilddatenfile ausliest, kann es zunächst keiner bestimmten einzelnen Tablette, sondern nur der Klasse aller Tabletten im Bild zugeordnet werden.

Damit die Blobs im Bild, also die Bereiche zusammenhängender Objektpixel, einzeln analysiert werden können, ist ein weiterer Bildverarbeitungsschritt erforderlich, das sog. Labeling. Das Labeling ist eine systematische Zusammenhangsanalyse aller Objektpixel im Bild. Objektpixel, die miteinander zusammenhängen, werden demselben Objekt zugeordnet und bilden einen „Blob“. Zwei Objektpixel gelten als zusammenhängend, wenn sie mindestens eine Pixelkante oder mindestens eine Pixelecke gemeinsam haben. [2]

Zur Unterteilung der Menge der Vordergrundpixel in Blobs wird das gesamte Bild Pixel für Pixel abgetastet und jeweils in einer kleinen Umgebung geprüft, ob ein Pixel Nachbarn besitzt, die mit einem bereits bekannten Blob zusammenhängen oder ob ein neuer Blob angelegt werden muss. Der Labeling-Algorithmus vergibt für jeden neuen Blob, der bei der Abtastung des Bildes gefunden wird, in der Reihenfolge des Auffindens eine Nummer, das „Label“. Wenn das ganze Bild durchlaufen ist, gibt die größte als Label vergebene Zahl unmittelbar die Anzahl der Blobs im Bild an. Auf diese Weise werden die Objekte im Bild beim Labeling gezählt. Das Verfahren ist insgesamt algorithmisch aufwendig und muss eine Reihe von Sonderfällen berücksichtigen, damit es zuverlässig funktioniert. Eine Erläuterung des Labeling in voller Allgemeinheit ist daher an dieser Stelle nicht möglich. Labeling-Verfahren sind im Detail in der Literatur beschrieben. [2] [3] [4] [5]

Die Vergabe von Labels dient aber nicht nur zum Zählen der Blobs im Bild, sondern ermöglicht auch die Zuordnung der Vordergrundpixel zu den einzelnen Objekten. Dazu wird beim Labeling ein neues Bild angelegt, das Label-Bild. Der Labeling-Algorithmus „färbt“ die Pixel im Label-Bild mit dem Grauwert, der der Nummer ihres Blobs entspricht. Das Ergebnis des Labeling ist daher ein Grauwertbild, bei dem jeder Blob einen einheitlichen Grauwert besitzt, der seiner Blobnummer entspricht. Ein Algorithmus, der anschließend ein Pixel aus dem Label-Bild ausliest, kann damit anhand des Grauwertes des Pixels feststellen, zu welchem Blob das Pixel gehört.

Das Labelbild lässt sich auch visuell interpretieren. Ein Beispiel zeigt Abbildung 1. Allerdings sind die Blobs in dieser Darstellung nicht unmittelbar mit den Grauwerten eingefärbt, die ihrer Labelnummer entsprechen, denn im Labelbild würden dann nur die Grauwerte 0 (für den Untergrund) und 1 bis 6 für die Objekte vorkommen. Das Bild wäre damit auf einem üblichen Monitor für einen menschlichen Betrachter praktisch vollkommen schwarz. Die Label-Grauwerte werden deshalb gleichmäßig über den gesamten verfügbaren Grauwertbereich von 1 bis 255 verteilt (gespreizt). Alternativ kann man das Labelbild visualisieren, indem wie in Abbildung 1 ganz rechts die Kontur jedes Blobs dargestellt und die Labelnummer am Blob vermerkt wird. Solche Darstellungsvarianten sind sehr nützlich und werden von den meisten Bildverarbeitungsprogrammen [6] automatisch unterstützt.

### Auswertung des Labelbildes

Im Labelbild kann ein Algorithmus gezielt auf die Pixel eines bestimmten Objekts zugreifen, weil die Blobnummer im Grauwert der Vordergrundpixel codiert ist. Beispielsweise kann die Fläche des Objekts mit der Blobnummer 1 ermittelt werden, indem einfach alle Pixel mit dem Grauwert 1, die im Labelbild vorkommen, gezählt werden. Im Unterschied zum einfachen „Pixelzählen“ ohne Zusammenhangsanalyse ist nach dem La-

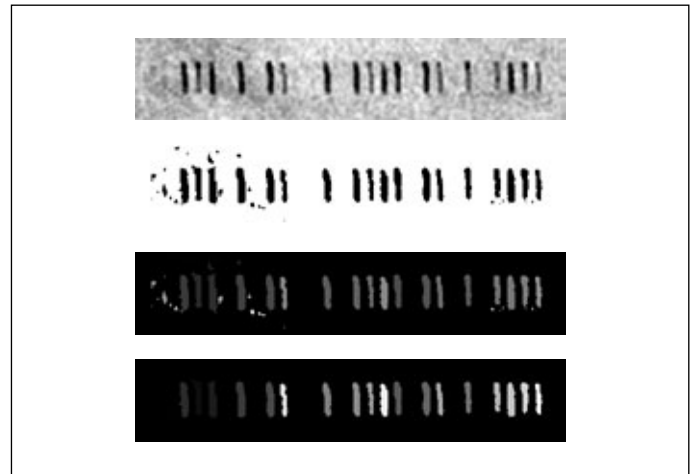


Abb. 2: Oben das Grauwertbild einer Strichcodierung, darunter das Binärbild für die Grauwertschwelle 155. Anschließend das gesamte Labelbild, ganz unten das Labelbild für eine Mindestgröße von 50 Pixeln für die Blobs.

Blob-Index	Fläche	Schwerpunkt		Schwerpunkt	Abstand
		x	y	x geordnet	
1	217	65	37	50	
2	263	80	41	65	15
3	264	50	40	80	15
4	278	109	40	109	29
5	257	140	40	140	31
6	220	272	38	154	14
7	241	302	39	199	45
8	173	345	37	229	30
9	219	199	40	244	15
10	220	229	40	258	14
11	151	244	38	272	14
12	164	316	40	302	30
13	164	376	37	316	14
14	244	390	41	345	29
15	180	404	37	376	31
16	162	154	40	390	14
17	223	258	41	404	14
18	148	418	39	418	14

Tab. 1: Ergebnisse der Blob-Analyse für das letzte Labelbild in Abbildung 2. In der letzten Spalte der Abstand aufeinanderfolgender x-Koordinaten der Blob-Schwerpunkte.

beling sichergestellt, dass die gezählten Pixel tatsächlich sämtlich zu einem zusammenhängenden Blob gehören, so dass die Flächenbestimmung nicht durch verstreute Pixel gestört wird, die sich beim einfachen Pixelzählen zufällig noch im Hintergrund im jeweiligen Zählfenster befinden könnten.

Verstreute Pixel oder Pixelgruppen im Hintergrund werden beim Labeling als eigenständige Blobs ermittelt und können meist wegen ihrer deutlich geringeren Größe von den eigentlich interessierenden Objekten unterschieden werden. Die meisten Bildverarbeitungsbibliotheken bieten eine Schwelle für die minimale Größe der Blobs an, die im Ergebnisbild erscheinen sollen. Ein Beispiel dafür zeigt die Abbildung 2, in der eine Aufnahme einer Strichcodierung mit relativ schlechtem Kontrast

und starken Hintergrundstrukturen dargestellt ist. Die Segmentierung mit einer globalen Binarisierungsschwelle, hier 155, gelingt nur unvollständig. Das Labeling ohne Größentoleranz ermittelt insgesamt 45 Blobs. Das Labeling mit einer Mindestgröße von 50 Pixeln für einen Blob (Abb. 2 ganz unten) führt auf 18 Blobs und markiert genau die Strichcodebalken im Originalbild. Der kleinste Balken hat eine Fläche von etwa 150 Pixeln, so dass die Diskriminierung der Blobs mit weniger als 50 Pixeln in diesem Fall problemlos ist und Störungen aus dem Hintergrund zuverlässig eliminiert.

### Der Schwerpunkt

Die Information ist bei dem Strichcode in Abbildung 2 im Abstand zwischen den einzelnen Balken enthalten. Zur Decodierung muss also die Position der Balken im Bild ermittelt werden. In der Blob-Analyse bietet sich dafür der Schwerpunkt an. Der Schwerpunkt eines Objekts ist in einem Binärbild der Mittelwert der x- und y-Koordinaten aller Pixel, die zum Objekt gehören. Tabelle 1 zeigt die Ergebnisse für die Bestimmung der Schwerpunkte der Blobs aus dem letzten Labelbild in Abbildung 2. Weil in diesem Beispiel der Strichcode bereits horizontal ausgerichtet ist, streuen die y-Koordinaten der Schwerpunkte nur geringfügig um etwa 5%. Die Flächen der einzelnen Striche, die nominell gleich groß sind, streuen dagegen mit etwa 30% wesentlich stärker. Das zeigt, dass der Schwerpunkt ein relativ robustes Merkmal ist.

Bei genauerer Betrachtung der Tabelle 1 fällt auf, dass die Balken nicht in der Reihenfolge gelabelt sind, in der sie im Bild von links nach rechts vorkommen. Das liegt daran, dass der Labeling-Algorithmus zeilenweise von links oben nach rechts arbeitet und dabei zuerst auf den zweiten, anschließend auf den dritten und erst dann auf den ersten Balken stößt. Die Liste der x-Koordinaten für die Schwerpunkte muss deshalb in aufsteigender Reihenfolge geordnet werden. In der letzten Spalte der Tabelle sind die Abstände der x-Koordinaten der Schwerpunkte aufeinanderfolgender Balken aufgelistet. Offensichtlich sind die Abstände ganzzahlige Vielfache von 15 Pixelkantenlängen. Die drei Klassen mit 15, 30 und 45 Pixelkantenlängen sind deutlich erkennbar und können bei einer Streuung von  $\pm 1$  Pixelkantenlänge sicher voneinander getrennt werden. Die Decodierung des Strichcodes ist deshalb bei diesem Beispiel trotz des relativ schlechten Ausgangsbildes mit Hilfe der Blob-Analyse problemlos möglich.

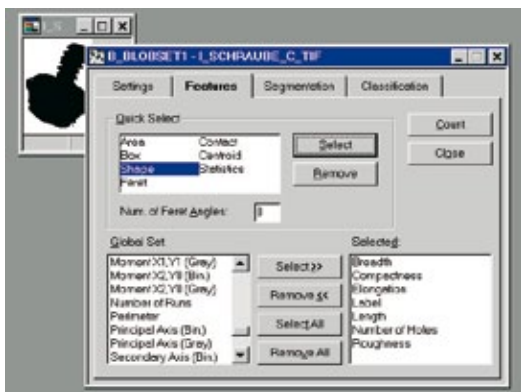


Abb. 3: Auswahlmenü für die Blob-Analyse in einem Bildverarbeitungsprogramm [7].

### Weitere Möglichkeiten der Blob-Analyse

Fläche und Schwerpunkt sind nur zwei Beispiele für die vielen Objektmerkmale, die die Blob-Analyse zur Verfügung stellen kann. Abbildung 3 zeigt einen Blick auf das Menü für die Blob-Analyse in einem typischen Bildverarbeitungsprogramm [7] mit einer kleinen Auswahl der mehreren Dutzend dort wählbaren Merkmale. Oft verwendete Kenngrößen sind z.B. der Umfang, das umschreibende Rechteck, der mittlere Durchmesser oder die Trägheitsmomente eines Blobs. Auch Formparameter, z.B. die sog. Rundheit (das ist im Wesentlichen das Verhältnis zwischen dem Quadrat des Umfangs und der Fläche) oder Parameter wie die Hauptachsenrichtung, die Rückschlüsse auf die Orientierung des Objekts zulassen, können für jeden Blob aus dem Labelbild berechnet werden.

Wenn das „Pixelzählen“ allein bei einer Prüfaufgabe nicht mehr zum Ziel führt und aufwendigere Verfahren erforderlich werden, ist oft die Blob-Analyse der nächstliegende Schritt. Das Labelbild gestattet den gezielten Zugriff auf einzelne Objekte im Bild. Damit werden viele Prüfparameter zugänglich, die häufig auch bei schwierigen Bildverarbeitungsaufgaben robuste Lösungen ermöglichen.

### Literaturangaben

- [1] Siehe GIT Inspect 2/2001, S. 10 ff.
- [2] Der Zusammenhang wird hier im Vordergrund über die sog. „Achter-Nachbarschaft“ definiert. Zum Begriff des Zusammenhangs und zum Labeling siehe z.B. Gerhard X. Ritter, Joseph N. Wilson, Computer Vision Algorithms in Image Algebra, CRC-Press, 2nd ed. 2001, p. 173
- [3] Rafael C. Gonzales, Richard E. Woods, Digital Image Processing, Addison-Wesley, 1993, p. 42 ff., p. 532 ff.
- [4] Nick Efferd, Digital Image Processing, Addison-Wesley, 1st ed. 2000, p. 260 ff.
- [5] John C. Russ, The Image Processing Handbook, CRC-Press, 3rd ed. 1999, p. 440 ff.
- [6] Die Bearbeitung der Bilder in Abb.1 und 2 wurde mit dem Programm AdOculus von The Imaging Source, www.theimagingsource.com, durchgeführt
- [7] Inspector von Matrox Electronic Systems



**GIG**  
VISION

# DIE KLEINSTE



Die Flea3 GigE CCD-Kameras sind in  
12 Modellen von VGA bis 5 Megapixel erhältlich.

**PREISANFRAGEN** über [www.ptgrey.com/GIGE](http://www.ptgrey.com/GIGE)

**POINT GREY**

*15 years*  
of Innovation in Imaging  
— 1997 - 2012 —

# Klassifizierung

## Eine Klasse für sich

Am Ende der Bildverarbeitungskette steht bei vielen Bildverarbeitungssystemen eine Entscheidung: das Kugellager ist vollständig mit Kugeln bestückt und somit in Ordnung oder nicht, ein Fahrzeug darf in das Parkhaus einfahren, die Tomate ist noch nicht rot genug zur Ernte. Die Basis für solche Entscheidungen ist eine Klassifizierung. Dabei werden numerische Werte für verschiedene Größen, die aus dem Bild extrahiert worden sind, bestimmten Kategorien zugeordnet, den sogenannten Merkmalsklassen. Bei der industriellen Produktionskontrolle gibt es letztlich meist nur zwei Klassen: „in Ordnung“ und „nicht in Ordnung“. In vielen Fällen müssen aber mehrere Klassen unterschieden werden, z. B. bei der Einteilung eines Luftbildes in Gewässer, Wälder, bewirtschaftete und bebaute Bereiche oder bei der Klarschifterkennung, die eine Zuordnung der Objekte zu den Buchstabenklassen erfordert.

Die Problematik der Klassifizierung, die im weitesten Sinne zum Gebiet der Mustererkennung gehört, wird seit Jahrzehnten intensiv untersucht. [1,2,3] Ob eine bestimmte Klassifizierungsmethode erfolgreich ist oder nicht, hängt sehr stark von der spezifischen Aufgabenstellung ab. Die Klassifizierung ist daher eines der schwierigsten Probleme der Bildverarbeitung. Zur Auslegung einer guten, problemangepassten Klassifizierungslösung ist so viel Erfahrung erforderlich, dass manche Akteure diese Disziplin sogar eher einer Kunst als einer Wissenschaft zuordnen möchten.

### Ein Beispiel: Strichcodeerkennung

Ein einfaches Beispiel, an dem viele Aspekte der Klassifizierung deutlich werden, zeigt Abbildung 1. Dabei handelt es sich um die Aufnahme einer Strichcodierung vom Typ „2-aus-5-interleaved“, bei der die Information sowohl in den Balken als auch in den Lücken codiert ist. Bei diesem Strichcode kommen nur zwei Breiten für die Elemente vor. Wenn die Codierung gelesen werden soll, muss für jedes Element lediglich festgestellt werden, ob es zur Klasse „breit“ oder zur Klasse „schmal“ gehört. Die Abfolge von breiten und schmalen Balken bzw. Lücken kann dann entsprechend der Syntax des Codes in Ziffern umgesetzt werden.

Bei einem Blick auf die Häufigkeitsverteilung der Elementbreiten wird die Problematik sofort deutlich. Abbildung 1b zeigt die Verteilungen für die drei Codierungen aus Abbildung 1a, die mit unterschiedlicher Auflösung gedruckt wurden. Auf der x-Achse ist die Elementbreite aufgetragen, die hier mit einem speziellen Verfahren mit einer Genauigkeit von einer halben Pi-

xelkantenlänge bestimmt werden konnte, auf der y-Achse findet sich die absolute Häufigkeit, mit der die jeweiligen Breiten im Strichcode vorkommen. Die Unschärfe der Bildaufnahme, Inhomogenitäten in der Beleuchtung, der relativ schlechte Kontrast und die Diskretisierung in Pixel führt zu einer Verschmierung der Verteilung mit einer Spannweite von etwa 2 bis 3 Pixelkantenlängen. Bei der oberen und der mittleren Codierung zerfällt die Verteilung noch deutlich in zwei getrennte Bereiche, bei der unteren Codierung ist jedoch nicht einmal die Entscheidung möglich, ob die beiden Verteilungen sich überlappen oder lediglich aneinandergrenzen.

### Cluster im Merkmalsraum

Solange eine Lücke zwischen den Verteilungen erkennbar ist, liegt es nahe, die gemessenen Breiten mit einem Schwellwert zu klassifizieren, der innerhalb der Lücke liegt. Breiten oberhalb des Schwellwerts gehören dann zur Klasse „breit“, Breiten unterhalb zur Klasse „schmal“. Die zugehörige Funktion, die den Merkmalswerten (hier den gemessenen Breiten) bestimmte Klassen eindeutig zuweist, wird als „Klassifikator“ bezeichnet. Die Gesamtheit der Werte, die die verwendeten Merkmale annehmen können, ist der „Merkmalsraum“. Die Anhäufungen der Merkmalswerte in voneinander getrennten Verteilungen heißen „Punktwolken“ oder „Cluster“. Das eigentliche Problem der Klassifizierung besteht nun darin, geeignete Merkmale für die Entscheidung auszuwählen und ein Verfahren zu finden, mit dem die Merkmalswerte zu Clustern zusammengefasst werden können.

In dem hier diskutierten Beispiel sind die Verhältnisse überschaubar. Die Betrachtung der Merkmalsverteilung suggeriert bereits die Aufteilung in zwei getrennte Cluster. Wo aber soll die Trennung zwischen den Clustern verlaufen, welcher Schwellwert für die Klassifizierung der Breiten ist optimal? Bei dieser Anwendung war außerdem eine dynamische Klassifizierung gefordert, d. h. der Schwellwert soll bei jeder Aufnahme erneut direkt aus dem Bild errechnet werden. Das Verfahren wird damit robuster und flexibler, weil die Decodierung weitgehend unabhängig von der Breite eines Strukturelements funktioniert, also skalierungsinvariant ist. Es liegt nahe, als Schwellwert die Mitte der Lücke zwischen den beiden Verteilungen zu verwenden. Allerdings ist es nicht einfach, einen Algorithmus zu schreiben, der die Lücke unter allen auftretenden Umständen sicher lokalisiert. Insbesondere bei kurzen Codierungen sind die Maxima häufig nicht stark ausgeprägt, und es können mehrere Lücken in der Verteilung auftreten. Eine andere Möglichkeit besteht darin, die Mitten der beiden getrennten Verteilungen zu bestimmen und die Schwelle in die Mitte zwischen diese Zentren zu legen. Dazu müssen die Cluster jedoch mit einem anderen Verfahren bereits so gut getrennt worden sein, dass sie gesondert behandelt werden können.

Solange nur zwei Klassen vorkommen, ist auch die Mitte der Gesamtverteilung eine sinnvolle Schwelle. Geeignete Größen sind z. B. der Mittelwert oder der Schwerpunkt der Gesamtverteilung. Strenggenommen ist beim Strichcode „2-aus-5-interleaved“ allerdings zu beachten, dass die schmalen Elemente häufiger vorkommen. Mit Annahmen über die Verteilung der



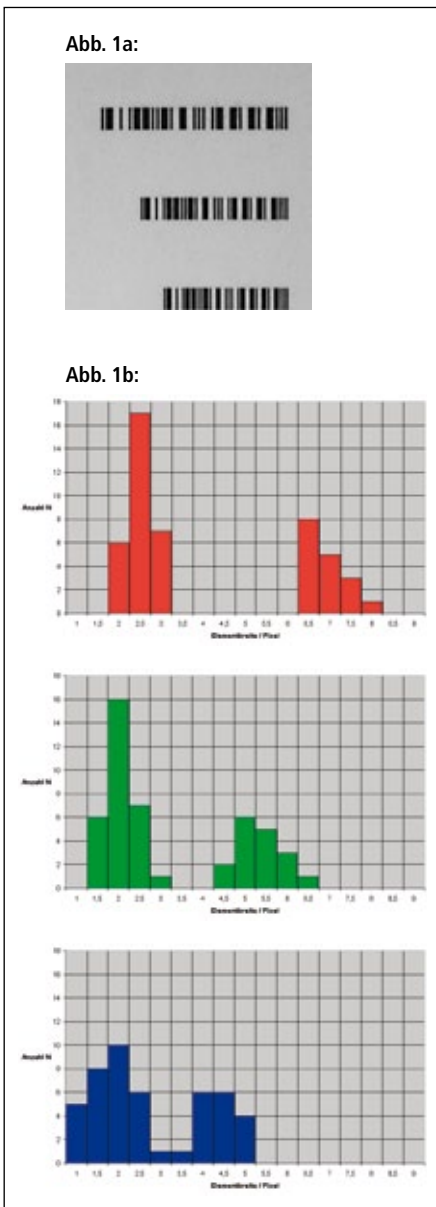


Abb. 1a: Eine Strichcodierung „2-aus-5-interleaved“, gedruckt in drei verschiedenen Strichstärken  
 Abb. 1b: Häufigkeitsverteilungen der Elementbreiten für die drei Strichcodes aus Abb. 1a

Merkmalswerte können aus statistischen Überlegungen heraus Korrekturen für diesen Einfluss berechnet werden. Pragmatiker ziehen jedoch in der Regel die simple Lösung der theoretisch besseren vor: der Mittelwert aus dem kleinsten und dem größten vorkommenden Wert in der Verteilung ist ebenfalls ein geeigneter Schwellwert. Dieser verblüffend einfache Klassifikator lässt sich sehr schnell und zuverlässig berechnen und hat sich im praktischen Einsatz erstaunlich gut bewährt.

### Klassifizierung mit Vorkenntnis

Bei den gerade angestellten Überlegungen ist Vorkenntnis über die Struktur des Merkmalsraums in die Konstruktion des Klassifikators eingeflossen. Weil bekannt ist, dass nur zwei Strukturbreiten vorkommen, sind nur zwei Cluster sinnvoll. Die Zahl der Klassen liegt also fest. In der Mustererkennung spricht man dann von einer überwachten Klassifizierung. Es gibt auch Methoden, mit denen die Clusterstruktur im Merkmalsraum unab-

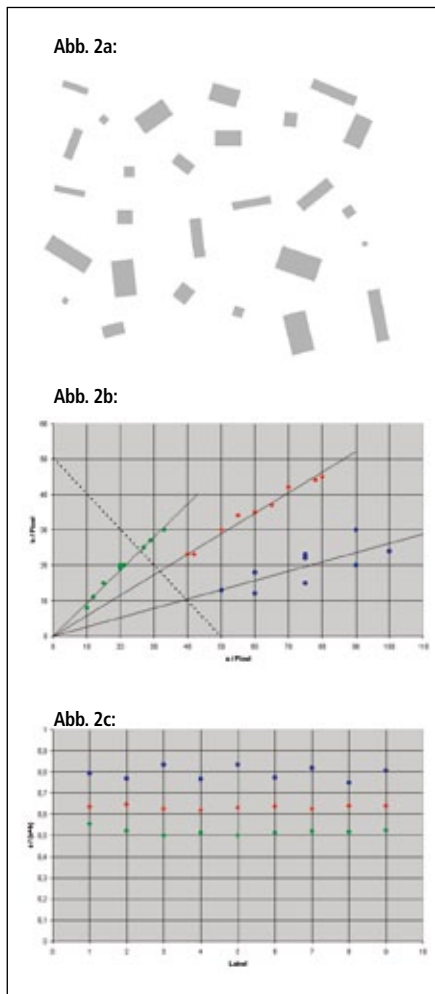


Abb. 2a: Drei Klassen von Rechtecken  
 Abb. 2b: Merkmalsraum für die größere Kantenlänge a und die kleinere Kantenlänge b für die Rechtecke aus Abb. 2a  
 Abb. 2c: Reduktion des zweidimensionalen Merkmalsraums aus Abb. 2b auf die Größe  $a/(a+b)$ . Die Elemente der drei Cluster sind jeweils mit einem Label versehen, also durchnummeriert worden.

hängig von Vorkenntnis analysiert werden kann [1]. Das ist sinnvoll, wenn Fehlerbilder gefunden werden sollen, die bisher noch nicht bekannt waren, z.B. bei der Prüfung von Oberflächen. Solche Verfahren werden als unüberwachte Klassifizierung bezeichnet.

Bei vielen industriellen Prüfaufgaben können die Klassen objektiv charakterisiert und in Lernläufen mit Mustern realisiert werden, und in den meisten Fällen ist die Zahl der Klassen bei der Prüfaufgabe bekannt, so dass sehr weitgehende Vorkenntnis über die Struktur des Merkmalsraums vorhanden ist. Allerdings findet ein Klassifikator, der auf Vorkenntnis beruht, grundsätzlich nur die Klassen, die ihm bekannt gemacht wurden. Oft ist es sehr schwierig, repräsentative Muster zu finden, die wirklich die gesamte Bandbreite der Merkmalsausprägungen abbilden. Auch bei der oben beschriebenen Strichcodeerkennung treten in Wirklichkeit gar nicht zwei, sondern vier Klassen auf. Die Lücken sind nämlich in den aufgenommenen Bildern systematisch schmaler als die Striche. Dies wird sowohl durch das Druckverfahren als auch durch die Aufnahmetechnik hervorgerufen. Bei getrennter Vermessung von Lücken und Strichen wird die Güte der Klassifizierung durch getrennte Behandlung der beiden Elementtypen tatsächlich weiter gesteigert.

### Überlappende Cluster

Wenn die Cluster im Merkmalsraum sich überlappen, sind Fehlklassifizierungen sehr wahrscheinlich. Eine solche Situation liegt im Prinzip für den unteren Strichcode aus Abbildung 1a vor. Tatsächlich grenzen die beiden Verteilungen hier ohne Lücke aneinander. Die Zuordnung der Elementbreiten 3,0 und 3,5 zu einer der beiden Breitenklassen ist jedoch nicht von vornherein klar. Bei einer Fehlklassifizierung wird ein Element einer Klasse zugeordnet, obwohl es nicht dazu gehört, oder es wird von einer Klasse ausgeschlossen, obwohl es hineingehört. Diese beiden Formen von Fehlklassifizierungen werden als „Fehler erster und zweiter Art“ bezeichnet. Jede der beiden Fehlerarten kann unangenehme Folgen haben. Ein gutes Teil, das als schlecht klassifiziert wird und im Ausschuss landet, mindert die Produktivität der Anlage, und ein schlechtes Teil, das als gut klassifiziert und weiterverwendet wird, kann in einer nachfolgenden Stufe der Wertschöpfungskette oder im Endprodukt zu erheblichen Schäden führen. Beide Fälle sind selbstverständlich unerwünscht, können jedoch unter Umständen nicht vollständig verhindert werden.

Auch im Überlappbereich der beiden Merkmalsverteilungen kann man eine Klassifizierung vornehmen [1,2]. Ein einfaches Verfahren dafür ist das Prinzip des nächsten Nachbarn („minimum distance“): Für einen Merkmalswert wird diejenige Klasse gewählt, die den geringsten Abstand zum Merkmalswert hat. Die Positionen der Klassen im Merkmalsraum können z.B. die in einem Lernlauf bestimmten Mittelpunkte der Clus-

ter sein. Eine Weiterentwicklung dieses Konzepts ist die Zuordnung nach dem Prinzip der „maximum likelihood“. Dabei wird diejenige Klasse gewählt, der das Merkmal mit höherer Wahrscheinlichkeit angehört. Dazu müssen aber die Wahrscheinlichkeitsverteilungen der Merkmalswerte bekannt sein. In einer weiteren Komplexitätsstufe kann bei dieser Vorgehensweise berücksichtigt werden, dass die Schäden bei einem Fehler erster oder zweiter Art verschieden hoch sein können. Auf diese Weise kann das Risiko bei einer Fehlklassifizierung quantifiziert werden, und es wird diejenige Klasse ausgewählt, bei der das Risiko minimal ist. Diese Methode wird als Bayes-Klassifikator bezeichnet [2].

Die Problematik dieser Verfahren liegt darin, dass der Aufwand für eine belastbare empirische Ermittlung der Wahrscheinlichkeitsverteilungen sehr hoch ist. Darüber hinaus müssen Fehlklassifikationen im praktischen Einsatz bewusst in Kauf genommen werden. Bei manchen Anwendungen ist dies jedoch kein Hinderungsgrund. Soll beispielsweise ein Bewegungsmelder an einem Privathaus nur bei Anwesenheit von Menschen die Beleuchtung einschalten und Katzen ignorieren, so ist eine gelegentliche Fehlklassifizierung, bei der Katzen als Menschen eingeschätzt werden, gewiss akzeptabel, denn der damit verbundene Schaden ist vermutlich vernachlässigbar. Bei vielen Anwendungen sollen Fehlklassifizierungen allerdings unter allen Umständen vermieden werden. Beispiele dafür sind die Produktionskontrollen in der Pharmaindustrie oder Zugangssysteme für Sicherheitsbereiche. Ein möglicher Ausweg ist hier die Bildung einer neuen Klasse, der alle Objekte zugeordnet werden, bei denen das Fehlklassifizierungsrisiko zu groß ist. Die fraglichen Objekte können dann z.B. in einer anderen Station mit größerer Prüftiefe oder aber manuell erneut untersucht werden. Bei der Realisierung dieses Konzepts werden zuerst in einem Lernlauf die Merkmalswerte für definierte Musterteile aufgenommen und Cluster Grenzen für die verschiedenen Klassen definiert. Die Grenzen können dabei je nach Erfahrung oder Sicherheitsbedürfnis gegenüber dem Lernlauf erweitert oder auch zusammengezogen werden. In der laufenden Kontrolle werden neue Merkmalswerte nur dann klassifiziert, wenn sie innerhalb eines der vorher definierten Cluster liegen, andernfalls gelten sie als nicht klassifizierbar.

### Ein Ausweg: Weitere Merkmale?

Bei überlappenden Merkmalsverteilungen kann es sinnvoll sein, weitere Merkmale zur Entscheidung heranzuziehen. Der Merkmalsraum wird dann mehrdimensional. Ein Beispiel liefert die Abbildung 2a, in der mehrere Rechtecke dargestellt sind. Für jedes Rechteck wurden aus diesem Bild die beiden Kantenlängen bestimmt.  $a$  ist die größere,  $b$  die kleinere Kantenlänge. In Abbildung 2b sind die Kantenlängen gegeneinander aufgetragen. Im jetzt zweidimensionalen Merkmalsraum bilden sich drei deutlich getrennte langgestreckte Cluster.

Die Kantenlänge  $a$  für sich allein ist nicht geeignet, um diese Klassen zu trennen. Damit lassen sich in dem vorliegenden Bild lediglich die Quadrate von den beiden anderen Klassen unterscheiden, nicht jedoch die verbleibenden Klassen untereinander. Auch die Hoffnung, dass durch das zusätzliche Merkmal  $b$

die Klassifizierung möglich sein könnte, wird enttäuscht, denn für  $b$  überlappen sich sogar die Verteilungen aller drei Klassen. Dennoch zeigt Abbildung 2b, dass die Cluster im zweidimensionalen Merkmalsraum getrennt sind. Die Information über die Unterscheidung muss also in den Merkmalen  $a$  und  $b$  enthalten sein. Selbstverständlich ist das auch tatsächlich der Fall, denn bei einem etwas längeren Blick auf Abbildung 2a identifiziert man die drei Klassen als Quadrate, Rechtecke mit einem Seitenverhältnis von etwa 2:1 und Rechtecke mit einem Seitenverhältnis von etwa 4:1. Es liegt nahe, zur Klassifizierung ein Merkmal zu verwenden, das die Streckung der Objekte beschreibt. Geeignete Größen sind z.B. das Verhältnis von  $b$  zu  $a$  oder das Verhältnis zwischen der Fläche des Objekts und dem Quadrat seines Umfangs. Die Clustergeometrie im Merkmalsraum legt jedoch unabhängig von der Visualisierung des Problems in Form von Rechtecken eine weitere Größe nahe. Die Cluster gruppieren sich nämlich um Strahlen, also um Richtungen im Merkmalsraum. Eine Zentralprojektion mit dem Ursprungspunkt als Zentrum wirft die Cluster deshalb auf der gestrichelt eingezeichneten Geraden in drei kompakte, deutlich getrennte Punkthaufen. Algebraisch entspricht diese Zentralprojektion der Bildung der Größe  $a/(a+b)$ . In Abbildung 2c ist dieses Merkmal auf der  $y$ -Achse des Diagramms aufgetragen. Zur Identifikation sind alle Objekte, die zur selben Klasse gehören, mit einem Label versehen, also durchnummeriert worden. Es ist deutlich erkennbar, dass dieses Merkmal für die drei Klassen zu drei getrennten Clustern führt. Auf diese Weise wird die Klassifizierung also auf ein einziges Merkmal zurückgeführt. Ähnliche Methoden können auch in drei- oder mehrdimensionalen Merkmalsräumen verwendet werden. Ein bekanntes Beispiel ist die Farbbildverarbeitung im dreidimensionalen RGB-Merkmalsraum. Dort wird die Klassifizierung oft durch Transformationen wesentlich vereinfacht, die aus der Farbmeterik stammen, z.B. von RGB nach HSI. Weitere Möglichkeiten zur Konstruktion von Klassifikatoren bieten neuronale Netze und Verfahren der sogenannten Fuzzy-Logik [3]. Am oben beschriebenen Beispiel ist deutlich geworden, dass zusätzliche Merkmale zwar mehr Information enthalten können. Die Klassifizierung wird aber nicht unbedingt und unmittelbar vereinfacht oder verbessert, wenn weitere Merkmale hinzugezogen werden. An Merkmalen herrscht in der Bildverarbeitung üblicherweise kein Mangel – die Kunst besteht vielmehr darin, die richtigen Merkmale mit hinreichender Genauigkeit und Stabilität zu erfassen und den Merkmalsraum dann geschickt so zu transformieren, dass eine möglichst einfache, sichere und robuste Klassifizierung für das vorliegende Problem möglich wird.

### Literatur

- [1] P. Haberäcker, Digitale Bildverarbeitung, 4. Auflage, Hanser, 1991, S. 248 ff.
- [2] Rafael C. Gonzales, Richard E. Woods, Digital Image Processing, Addison-Wesley, 1993, p. 574 ff.
- [3] Eine gute Einführung in die Methodik der neuronalen Netze und der Fuzzy-Logik gibt P. Haberäcker, Praxis der Digitalen Bildverarbeitung und Mustererkennung, Hanser, 1995, Kapitel 8 und 9



## Heiße Liebe

Weshalb werden Sie sich in die Prosilica GT verlieben? Wegen ihrer Robustheit? Vielleicht eher wegen ihrer hochempfindlichen und schnellen Sensoren? Oder bevorzugen Sie Fernbeziehungen mit langen Kabellängen und Power over Ethernet? Vielleicht lieben Sie sie für alles, was sie anzubieten hat. Machen Sie sich keine Sorgen, zu viel von ihr zu verlangen. Die Prosilica GT kommt auch in einer heißen Umgebung klar – bis zu 60°C. Es darf aber auch mal kalt werden – bis zu -20°C. Und dank ihrer Temperaturkontrolle sagt sie Ihnen immer, wie sie sich fühlt. Ist die Prosilica GT Ihr Traumpartner? Finden Sie es heraus unter [www.AlliedVisionTec.com/HeisseLiebe](http://www.AlliedVisionTec.com/HeisseLiebe)



SEEING IS BELIEVING



# Beleuchtung

## In Lichtgeschwindigkeit gerechnet

Beleuchtung ist eine der wichtigsten Komponenten eines Bildverarbeitungssystems. Eine gut gewählte Beleuchtungsmethode hebt Prüfmerkmale hervor und unterdrückt störende Strukturen. Die Weiterverarbeitung des aufgenommenen Bildes wird damit wesentlich einfacher, die Applikation wird robuster. Die Beleuchtung muss jedoch stets an die Anwendung angepasst sein, und es müssen eine Reihe von technischen Randbedingungen berücksichtigt werden. Eine geschickt ausgelegte Beleuchtung ist eine sehr effiziente, unschlagbar schnelle Methode der Bildvorverarbeitung.

### Was ist daran schwierig?

Dass wir im Tageslicht der Sonne die Helligkeit und die Farbigkeit unserer natürlichen Umwelt über unsere Augen wahrnehmen können, erscheint uns als Selbstverständlichkeit. In einer ansonsten dunklen Nacht können wir uns im Licht des Vollmondes orientieren, obwohl die Beleuchtungsstärke um mehrere Größenordnungen geringer ist als gegen Mittag an einem klaren Sommertag. Selbst massive Inhomogenitäten der Beleuchtung nehmen wir bewusst gar nicht wahr, und auch Schatten stören uns kaum bei der Interpretation der Szenen aus unserer natürlichen Umgebung. Sogar die erhebliche Variation der spektralen Verteilung des Sonnenlichts im Lauf eines Tages fällt uns kaum auf, und wir sehen die weiß verputzte Fassade eines Hauses als weiße Fläche, gleichgültig ob sie in das Rot eines Sonnenuntergangs getaucht ist oder vom weißen Mittagslicht beleuchtet wird. Dass diese Erkennungsleistungen des menschlichen visuellen Systems schier atemberaubend sind, kann man Außenstehenden kaum verständlich machen. Nur wenige Berufsgruppen, etwa Photographen oder Innenarchitekten, haben eine professionelle Sensibilität für diese Problematik. Selbst erfahrene Bildverarbeitungsingenieure sind erstaunt, wenn ihnen diese Zusammenhänge erstmals bewusst werden. Wir unterschätzen die Bedeutung der Beleuchtung für eine Bildverarbeitungslösung, weil uns die Erkennungsleistung unseres menschlichen visuellen Systems so trivial erscheint, dass wir von einem technischen System instinktiv vergleichbare Fähigkeiten erwarten.

### Ohne Beleuchtung keine Bildverarbeitung

Auf den ersten Blick wenig spektakulär, tatsächlich aber eine der wesentlichen Entwicklungen im Bereich der Bildverarbeitung in den letzten zehn Jahren ist das wachsende Bewusstsein über die grundlegende Bedeutung der Beleuchtung. Ein Bildverarbeitungssystem kann nur das verarbeiten, was in einer Szene beleuchtet wird oder was selbst leuchtet. Diese Erkenntnis erscheint selbstverständlich und ist keineswegs neu. Neu ist, dass mittlerweile jeder, der auf diesem Gebiet arbeitet, mit dieser Einsicht konfrontiert wird. Schon 1991 behandelt Hartmut Ernst in seinem exzellenten Lehrbuch [1] das Thema Beleuchtungsmethoden in einem eigenen Kapitel, wenn auch nur auf

zwei von insgesamt ca. 300 Seiten, und schreibt: „Eine wesentliche Komponente für die Bildverarbeitung ist eine dem Problem angepasste Beleuchtung. Eine gute Beleuchtung der zu erfassenden Szene erlaubt oft die Verwendung einfacherer und damit schnellerer Algorithmen zur Bildauswertung. Generell wird man die Beleuchtung so einrichten, dass sich ein hoher Kontrast ergibt. In der Regel fordert man auch eine Minimierung von Reflex- und Schattenbildung“. Das ist heute noch genauso richtig und wichtig wie damals. Die Beleuchtung dient dazu, Merkmale herauszuarbeiten und störende Strukturen zu unterdrücken. Sie ist in diesem Sinne notwendiger und höchst effizienter Bestandteil der Bildverarbeitungskette. Heute wird man das natürliche Umgebungslicht nur dann als Beleuchtung in Kauf nehmen, wenn es dazu keine akzeptable Alternative gibt.

### Die weitere Entwicklung

Über Jahrzehnte wurde die Bildverarbeitung in Lehrbüchern in der Breite primär als „digital image processing“ verstanden, also als algorithmische Behandlung der Bilddatei. Filteroperationen, Grauwerttransformationen und morphologische Verfahren wurden in epischer Breite dargelegt. Noch im Jahr 2005 wurde in einem ansonsten ausgezeichneten einführenden Lehrbuch [2] das Thema Beleuchtung in einem einzigen Satz abgehandelt: „Der eigentliche Prozess der Entstehung von Bildern ist oft kompliziert und meistens für die Bildverarbeitung auch unwesentlich“. Schon im selben Jahr kann man jedoch in der 6. Auflage eines bekannten Lehrbuchs [3] bereits 35 (von 600) Seiten lesen, auf denen wesentliche Aspekte der Wechselwirkung zwischen Beleuchtung und Objekt behandelt werden. Ein Lehrbuch aus dem Jahr 2008 [4] stellt die Bedeutung der Beleuchtung als Systemkomponente für die Bildverarbeitung ausdrücklich heraus und widmet dem Thema auf 12 von ca. 300 Seiten ein eigenes Kapitel mit praktischen Hinweisen und Beispielen. Und im Jahr 2010 spendiert ein bedeutendes Vertriebshaus für Bildverarbeitungs-komponenten in seinem ca. 300 Seiten umfassenden Katalog [5] für das Thema „Beleuchtungs-Technologie“ bereits 20 Seiten mit Basisinformationen im Stil eines Applikationsratgebers und präsentiert weitere 40 Seiten Produktinformationen zu diesem Thema, also etwa 20% des Gesamtumfangs. Dort findet sich der bemerkenswerte Hinweis: „Die Auswahl der geeigneten Beleuchtung ist ausschlaggebend für die Qualität der Bildaufnahme und kann die anschließende Bildauswertung stark vereinfachen. Meist wird jedoch die Bedeutung der richtigen Lichtquelle für ein Bildverarbeitungs-System unterschätzt. Die passende Beleuchtung zu finden, ist eine anspruchsvolle Aufgabe.“

### Kontrast

Bildverarbeitung lebt vom Kontrast: Meist zerfällt das Bild in Objekte, die sich deutlich von einem Hintergrund abheben. Erst wenn Objektpixel und Hintergrundpixel getrennt sind, können weitere Informationen extrahiert werden, z.B. die Anzahl der Objekte, ihre Fläche, ihr Durchmesser und ähnliche Größen. Die Segmentierung, also die Unterscheidung von Vordergrund- und Hintergrundpixeln, wird umso einfacher, je größer der Grauwertunterschied zwischen Objekten und Hintergrund ist. Im Idealfall



ist dieser Kontrast im gesamten Bildfeld annähernd gleich groß und der Hintergrund hat einen annähernd konstanten Grauwert. Aber auch bei inhomogenem Hintergrund ist eine Segmentierung möglich, solange ein lokaler Kontrast zu den Objekten im Vordergrund vorhanden ist. Es ist offenkundig, dass die Beleuchtung einen großen Einfluss auf den Kontrast im Bild haben kann.

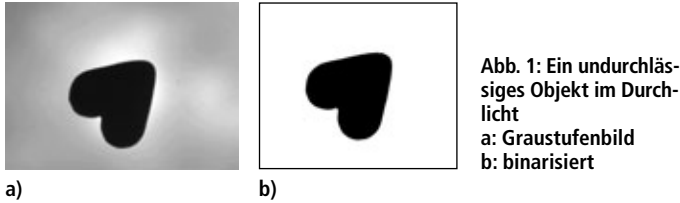


Abb. 1: Ein undurchlässiges Objekt im Durchlicht  
a: Graustufenbild  
b: binarisiert

## Durchlicht und Auflicht

Hervorragenden Kontrast liefert eine Bildaufnahme bei Beleuchtung im Durchlicht. Dabei befindet sich das Objekt zwischen Lichtquelle und Kamera. Im Gegensatz dazu spricht man von Beleuchtung im Auflicht, wenn die Lichtquelle und die Kamera auf derselben Seite des Objekts angeordnet sind. Für undurchlässige Gegenstände erscheinen die Objektpixel im Durchlicht sehr dunkel auf sehr hellem Hintergrund. Abbildung 1 zeigt ein Beispiel. Diese Beleuchtungsmethode ist gut geeignet, wenn der Umfang, der Durchmesser, die Fläche, die Form und ähnliche Eigenschaften von Objekten geprüft werden sollen, die Oberflächenbeschaffenheit aber keine Rolle spielt. Meist kann das Bild dann mit einer einfachen, globalen Binarisierungsschwelle in Vordergrund und Hintergrund segmentiert werden, ohne dass die Farbe, die Struktur oder die Reflektivität der Oberfläche sich auf das Ergebnis auswirken. Auf diese Weise kann mit derselben Prüfeinrichtung beispielsweise der Innendurchmesser einer Unterlegscheibe geprüft werden, gleichgültig, ob die Scheibe aus Stahl oder Messing gefertigt ist, ob sie hochglänzend oder matt erscheint oder ob Drehriefen vorhanden sind oder nicht. Bei einer Beleuchtung im Auflicht würden sich diese Variationen in der Oberflächenbeschaffenheit des Objekts als Strukturen im aufgenommenen Bild wieder finden und müssten mit Bildverarbeitungsalgorithmen eliminiert werden. Die Variationsbreite von Oberflächeneigenschaften ist jedoch auch bei einer industriellen Fertigung so groß, dass bei Auflichtaufnahmen immer wieder überraschende Effekte auftreten, die beim Systemtest nicht vollständig vorausgesehen werden können. Eine Durchlichtanordnung hingegen eliminiert alle Oberflächeneffekte bereits bei der Bildaufnahme und ist damit effizienter, sicherer und schneller als jede Bildvorverarbeitung einer Auflichtaufnahme.

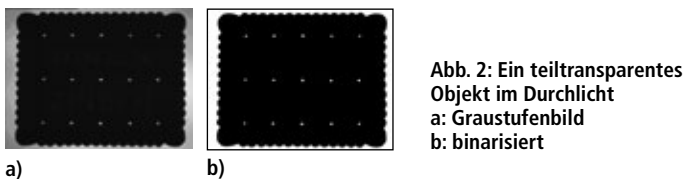


Abb. 2: Ein teiltransparentes Objekt im Durchlicht  
a: Graustufenbild  
b: binarisiert

Auch bei einer Durchlichtanordnung können bei der Bildaufnahme Strukturen innerhalb des Objekts entstehen, wenn der Gegenstand nicht vollständig lichtundurchlässig ist. Abbildung 2 zeigt ein Beispiel. Objekt und Hintergrund können bei der Aufnahme aus Abbildung 2 jedoch immer noch sicher getrennt werden, denn die Grauwerte des durchscheinenden Schriftzugs liegen weit unter der Binarisierungsschwelle, die zur Segmen-

tierung verwendet wurde. Auch bei diesem Beispiel wirken sich der Bräunungsgrad des Kekses, die Beschaffenheit des verwendeten Teigs oder ein zusätzlicher, vollständig oder auf Teilen der Oberfläche aufgebracht Schokoladenüberzug nur geringfügig auf das Ergebnis der Bildaufnahme aus. Wenn die verbleibenden Löcher stören, z. B. weil es bei der Prüfaufgabe nur auf die Zahl der Zähne ankommt, können die transparenten Stellen mit einer einfachen Bildverarbeitungsoperation geschlossen werden. Andererseits kann man die Durchlichtaufnahme gut verwenden, wenn gerade die Zahl und Lage der Löcher im Keks geprüft werden soll. Im Auflicht ist es wesentlich schwieriger, die Löcher so prominent hervorzuheben wie im Durchlicht.

In Abbildung 1 sind schon mit bloßem Auge Inhomogenitäten in der Hintergrundbeleuchtung erkennbar. Da der Kontrast zwischen Objekt und Hintergrund hier sehr hoch ist, wirken sich diese Schwankungen bei der Weiterverarbeitung der Bilddaten nicht störend aus. Die Lichtquelle war in diesem Fall ein kommerzielles ebenes LED-Leuchtfeld, das aus einem LED-Array und einer vorgeschalteten Streuplatte besteht. Die unterschiedlichen Helligkeiten an verschiedenen Stellen entstehen dadurch, dass einige LEDs heller leuchten als andere. Andere Leuchtfelder für Durchlichtbeleuchtungen verwenden LED-Zeilen, deren Lichtemission seitlich in eine Kunststoffscheibe eingekoppelt wird, so dass sich die Strahlung durch Totalreflexion im Innern der Scheibe ausbreitet. Wenn die Oberfläche angehaut wird, tritt ein Teil der Strahlung aus. Grundsätzlich kann durch eine gezielte Oberflächenbehandlung die Intensitätsverteilung auf dem Leuchtfeld definiert beeinflusst werden. Als Leuchtmittel sind auch Leuchtstoffröhren geeignet, wie sie z. B. bei den Leuchtischen für technische Zeichner, für die Reproduktionstechnik oder für die Betrachtung von Röntgenaufnahmen eingesetzt werden.

## Richtige Belichtung

Bei der Durchlichtbeleuchtung sollte man nicht nur auf den Kontrast, sondern auch auf die Aussteuerung achten. Wie bei jeder Beleuchtungsmethode, ist es auch beim Durchlicht sinnvoll, das Bildsignal in den hellen Bereichen nicht in die Sättigung zu treiben („Überbelichtung“), denn dadurch geht Information verloren. Beispielsweise kann es bei Überbelichtung an den Rändern von Objekten zu Überstrahlungen kommen, die sich störend auf die Vermessung des Durchmessers, der Fläche oder des Umfangs auswirken. Andererseits darf das Bildsignal auch nicht zu gering ausgesteuert sein. Wenn sich Hintergrund und Objekt nur um wenige Graustufen unterscheiden, ist eine robuste Segmentierung mit einer globalen Segmentierungsschwelle nicht möglich. Es ist also sinnvoll, mit der Beleuchtung den Dynamikumfang der Kamera gut auszuschöpfen, aber einen sicheren Abstand zur Übersteuerung und zur Unterbelichtung einzuhalten. Eine typische, für die meisten Anwendungen hinreichend stabile Aussteuerung für eine 8-bit-Kamera, die 256 Graustufen liefert, sind Grauwerte in der Umgebung von etwa 200 Graustufen für den Hintergrund und von etwa 50 Graustufen für den Vordergrund.

## Richtcharakteristik der Beleuchtung

Sowohl Auflicht- als auch Durchlichtbeleuchtung können gerichtet oder ungerichtet (diffus) ausgeführt werden. Es gibt eine große Zahl von verschiedenen Anordnungen, die sich hinsichtlich der Divergenz und der Orientierung des Strahlenbündels zur Probe und zur Kamera unterscheiden. Beispielsweise kann eine diffuse Beleuchtung von einem LED-Leuchtfeld mit einer Streuplatte kommen, das unter 45° zur Oberfläche eines Prüflings angeordnet ist. Ein Halogen-Spot kann als Punktlicht-

quelle verwendet werden und definierte Reflexe von Kugeln erzeugen, beispielsweise bei der Prüfung von Kugellagern auf vollständige Bestückung. Eine diffus reflektierende Halb-Hohlkugel, in die Licht eingekoppelt wird, kann eine diffuse, von allen Seiten gleichmäßig auf die Probe einfallende Beleuchtung erzeugen, die wie natürliches, diffuses Tageslicht an einem wolkenverhangenen Herbsttag wirkt und daher auch als „cloudy-sky“-Beleuchtung bezeichnet wird. Andererseits sind Laserlichtquellen und Kollimatoroptiken verfügbar, die gerichtete Lichtbündel mit sehr geringen Divergenzen erzeugen. Mittlerweile gibt es eine Reihe von Firmen, die sich auf Beleuchtungseinrichtungen für die Bildverarbeitung spezialisiert haben und ein breites Produktspektrum anbieten.

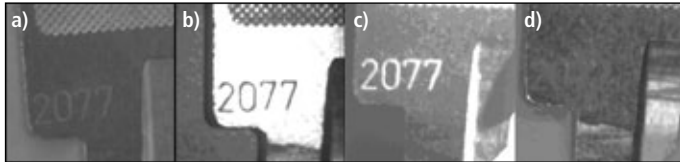


Abb. 3: Aufnahme einer Prägung auf einem Schlüssel  
a: unter undefiniertem Tageslicht; b: mit gerichteter, kollimierter Beleuchtung senkrecht von oben; c: mit gerichteter, kollimierter Beleuchtung schräg von der Seite; d: mit diffuser Beleuchtung

### Hellfeld und Dunkelfeld

Eine hilfreiche Klassifizierung ist die Unterscheidung in Hellfeld- und Dunkelfeldbeleuchtung. Abbildung 3 zeigt ein Beispiel für die unterschiedlichen Verhältnisse bei der Ausleuchtung einer Prägung in einer reflektierenden Schlüsselselfläche. Abbildung 3a ist eine Aufnahme unter undefiniertem Tageslicht in einem Laborraum. Die Prägung ist erkennbar, zeigt aber nur geringen Kontrast. Abbildung 3b ist eine Aufnahme mit einer gerichteten, kollimierten Beleuchtung. Das Strahlenbündel fällt senkrecht auf die Oberfläche ein, und die Kamera ist ebenfalls senkrecht auf die Prägung gerichtet. Die Prägung streut das senkrecht einfallende Licht diffus in alle Richtungen, während die glänzende Oberfläche des Schlüssels das Licht direkt in die Kamera reflektiert. Die Prägung erscheint daher dunkel auf hellem Untergrund. Hier handelt es sich um eine Hellfeldanordnung. Für Abbildung 3c wurde eine Dunkelfeldanordnung realisiert. Ein kollimiertes Lichtbündel fällt schräg auf die Oberfläche des Schlüssels, die Kamera ist wieder senkrecht auf die Prägung gerichtet. Das einfallende Licht wird von der glänzenden Oberfläche spiegelnd reflektiert und kann nicht in das Objektiv der Kamera gelangen. Nur im Bereich der Prägung wird das Licht diffus in alle Richtungen gestreut, so dass ein Teil der Strahlung von der Oberfläche senkrecht nach oben remittiert wird. Die Prägung erscheint daher hell auf dunklem Untergrund. Abbildung 3d zeigt die Verhältnisse bei einer diffusen Beleuchtung, bei der das Licht aus allen Richtungen gleichmäßig auf den Schlüssel fällt. Sowohl von der spiegelnd reflektierenden Fläche als auch aus dem Bereich der Prägung kommen nun Lichtstrahlen, die sich in alle Raumrichtungen ausbreiten. Die Kamera ist wieder senkrecht auf die Oberfläche gerichtet ist. Im entstehenden Bild sind die Unterschiede zwischen der Prägung und der spiegelnd reflektierenden Fläche kaum zu erkennen. Dieses Beispiel zeigt: eine diffuse Beleuchtung unterdrückt Oberflächenstrukturen, eine gerichtete Beleuchtung hebt Oberflächenstrukturen hervor. Die Ausprägung der gerichteten Beleuchtung als Hellfeld- oder Dunkelfeldanordnung ermöglicht wahlweise einen hellen oder einen dunklen Untergrund. Kratzer, Riefen, Erhebungen und andere topographische Oberflächenfehler werden besonders hervorgehoben, wenn das Licht gerichtet und unter sehr flachem Winkel auf die Oberfläche fällt (Streiflicht).



Abb. 4: Aufnahme einer Münze mit ringförmiger Beleuchtung zur Hervorhebung der Kanten

### Ringlicht

Bei schrägem Einfall entstehen durch gerichtete Lichtbündel helle Bildbereiche an Kanten oder Prägungen. Wenn das Licht jedoch nur von einer Seite auf das Objekt fällt, bilden sich an den abfallenden Kanten dunkle Bildbereiche und manchmal ausgeprägte Schatten. Bei Prüflingen, die nicht in einer definierten Lage an der Prüfstation ankommen, kann dies bei der weiteren Behandlung des Bildes zu Schwierigkeiten führen. Eine ringförmige, gerichtete Beleuchtung führt in einer solchen Situation zu besseren Ergebnissen. Abbildung 4 zeigt ein Beispiel für eine solche Aufnahme. Zur Beleuchtung wurde hier eine Halogen-Kaltlichtquelle benutzt, deren Strahlung in einen Lichtleiter-Querschnittswandler eingekoppelt wurde. Auf der Auskoppelseite ist der verwendete Querschnittswandler als schmaler, ringförmiger Ausschnitt einer Kegelfläche so ausgeführt, dass das Licht unter einem Winkel von ca. 45° von allen Seiten gerichtet zur Mittelachse des Rings hin austritt. Die Kamera ist auf der Achse des Rings montiert. Je nach Abstand zwischen Ringfläche und Probe wird ein größerer oder kleinerer kreisförmiger Bereich gerichtet ausgeleuchtet. In Abbildung 4 ist gut zu erkennen, dass auf diese Weise eine rotations-symmetrische Beleuchtung entsteht, die die Kanten unabhängig von ihrem Verlauf auf der Münze nahezu gleichmäßig heraushebt, während ebene Flächen dunkel bleiben. Besonders gut ist dies beim Schriftzug und beim Prägejahr sowie bei den Sternen erkennbar, die teilweise hochgeprägt und teilweise tiefgeprägt sind. Das Bild hat große Ähnlichkeit mit einem Kantenbild, das aus einem konventionell aufgenommen Bild durch Nachbearbeitung mit Kantenfiltern entstanden sein könnte. Das „Kantenbild“ in Abbildung 4 ist jedoch allein mit den Mitteln der Beleuchtungstechnik entstanden, ohne dass eine einzige Zeile Programmcode erforderlich ist – und das in einer „Rechenzeit“, die auch die besten Prozessoren nicht erreichen können.

### Blitzbeleuchtung

Für industrielle Anwendungen ist die Zeitstruktur der Beleuchtung oft von großer Bedeutung. Häufig werden Teile in der industriellen Fertigung auf Förderbändern durch die Prüfstation transportiert und müssen in der Bewegung aufgenommen werden. Geschwindigkeiten in der Größenordnung von mehreren Metern pro Sekunde sind keine Seltenheit. Schon bei 1 m/s Transportgeschwindigkeit wird ein Objekt während 1 ms um 1 mm versetzt. Eine Belichtungszeit von 20 ms, die bei einer 50-Hz-„progressive-scan“-Kamera ohne weitere Maßnahmen gegeben ist, führt also zu einer Verschmierung des Bildes um 20 mm in Transportrichtung. Für die meisten Bildverarbeitungsmethoden ist diese Verschmierung nicht akzeptabel. Meist strebt man eine Bildverschmierung um weniger als eine Pixelkantenlänge an. Eine gute Möglichkeit zur Kontrolle der Bildverschmierung ist eine definiert gepulste Beleuchtung. Mit LEDs sind definierte Pulsbreiten bis herunter zu einigen Mikrosekunden und Pulsfrequenzen im Bereich einiger Kilohertz möglich. Im Pulsbetrieb können LEDs außerdem mit deutlich

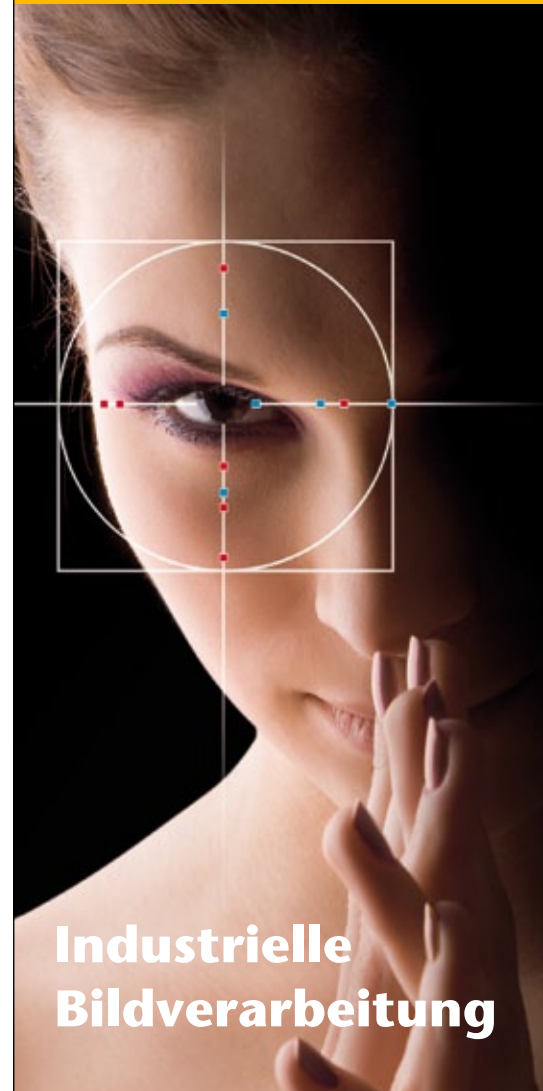
höherem Strom betrieben werden als im DC-Modus. Die Lichtleistung in einem LED-Puls ist daher oft so hoch, dass das Restlicht aus der Umgebung deutlich überstrahlt werden kann. Die Bewegung wird dann durch den LED-Blitz „eingefroren“, denn das Objekt remittiert nur während der Dauer des Blitzes genügend Licht für ein Bildsignal. Dennoch ist es wie bei jeder anderen Beleuchtung auch bei LED-Blitzbeleuchtungen sinnvoll, eine Abdeckhaube, einen Tunnel oder ähnliche Abschirmungen zu verwenden, damit kein störendes Umgebungslicht auf die Probe fallen kann und die Beleuchtung allein durch die definierten Lichtquellen der Prüfeinrichtung erzeugt wird.

## LEDs

LEDs sind mittlerweile für Blitzbeleuchtungen das Leuchtmittel der Wahl. Nur in Ausnahmefällen wird man auf klassische Blitzlampen, z.B. Xe-Gasentladungslampen, zurückgreifen, wenn extreme Anforderungen an die Lichtleistung während des Pulses gestellt werden. Während Gasentladungslampen mit Hochspannung betrieben werden, wegen ihrer Ausführung in Glas oder Keramik vergleichsweise fragil sind und auch im Hinblick auf die elektromagnetische Verträglichkeit spezielle Maßnahmen in der Schaltungstechnik erfordern, handelt es sich bei LEDs um Bauelemente, die jeder Elektroniker ohne Vorbehalte elektrisch und mechanisch in eine Schaltung integrieren wird. Zudem liegt die Lebensdauer von LEDs bei korrekter Auslegung der Versorgung bei 10.000 Stunden und mehr, die Helligkeitsschwankungen von Puls zu Puls sind deutlich geringer als bei Gasentladungslampen, und die Pulsform kann gezielt eingestellt werden. LEDs gibt es außerdem inzwischen für praktisch jeden Spektralbereich, der für die industrielle Bildverarbeitung von Bedeutung ist, vom nahen UV über Blau, Grün und Rot bis ins Infrarot und auch als breitbandige weiße LED. Der Aufbau von LED-Beleuchtungen aus Einzeldioden ermöglicht zudem die Konstruktion von Flächen, Ringen, und Zeilen aus Elementen mit unterschiedlichen Abstrahlcharakteristiken, so dass es kaum geometrische Anforderungen an eine Beleuchtung gibt, die mit LEDs nicht zu erfüllen wären.

## Leuchtstoffröhren

Beliebte Leuchtmittel für großflächige, homogene Beleuchtungen oder für die Ausleuchtung von Linien sind nach wie vor Leuchtstoffröhren. Die klassische Ausführung als dünne, zylindrische Röhre führt bereits zu einer sehr homogenen Helligkeitsverteilung in Längsrichtung. Mit einer zweiten, parallel im richtigen Abstand angeordneten Leuchtstoffröhre kann eine sehr homogene Ausleuchtung einer Fläche erreicht werden. Bei richtiger Beschaltung haben Leuchtstoffröhren eine Lebensdauer von mehreren Tausend Stunden. Für viele industrielle Anwendungen ist dieser Zeitraum akzeptabel, wenn konstruktiv sichergestellt ist, dass der Austausch ohne großen Aufwand erfolgen kann. Leuchtstoffröhren gibt es auch als Ringleuchten, außerdem in verschiedenen Farben und als „Schwarzlichtlampe“ für das nahe UV. Asynchroner Blitzbetrieb ist mit Leuchtstoffröhren jedoch nicht ohne weiteres möglich. Sie werden daher meist als Dauerstrich-Beleuchtung eingesetzt, so dass bei bewegten Objekten der „electronic shutter“ der Kamera für das Einfrieren der Bewegung verwendet werden muss. Leuchtstoffröhren können zwar grundsätzlich mit Gleichstrom betrieben werden und emittieren dann Licht ohne Zeitstruktur. Eine kommerzielle Leuchtstoffröhrenlampe für den Consumer-Bereich arbeitet jedoch mit 50-Hz-Wechselspannung. Die Lichtemission folgt dem Zeitverhalten des Stroms, so dass die Lichtintensität mit einer Frequenz von 100 Hz moduliert ist und



## Industrielle Bildverarbeitung

## Komponenten, Systeme & Beratung

[www.polytec.de/bv](http://www.polytec.de/bv)

- LED-Beleuchtung
- Faseroptische Beleuchtung
- Stroboskope
- Kameras
- Vision-Sensoren
- Hochgeschwindigkeitskameras
- Objektive
- Vision Packages & Software



bei den Nulldurchgängen des Wechselstroms nahezu auf Null absinkt. Für Standard-Kameras, die mit einer Bild- oder Halbbildfrequenz von 50 Hz betrieben werden, entstehen dadurch stark ausgeprägte Schwebungen in der Helligkeit der aufgenommenen Bilder. Bei asynchron getriggerten Kameras mit Shutter-Zeiten von einigen Millisekunden oder weniger ist es purer Zufall, ob die Leuchtstoffröhre während der Belichtungszeit gerade im Minimum oder im Maximum der Lichtintensität leuchtet, so dass erhebliche Helligkeitsschwankungen von Bild zu Bild entstehen. Leuchtstoffröhren müssen folglich mit Frequenzen betrieben werden, die so hoch sind, dass die beschriebenen Effekte in der Anwendung nicht auftreten können. Auch bei einer Lichtquelle mit Hochfrequenz-Vorschaltgerät ist es ratsam, sich mit einer (wenig aufwendigen) Messung mit Photodiode und Oszilloskop Gewissheit über die Zeitstruktur der Beleuchtung zu verschaffen.

### Spektralverteilung

Auch die Spektralverteilung der Beleuchtung muss auf die Prüfaufgabe abgestimmt und an die Spektralverteilung der Empfindlichkeit der Kamera angepasst werden. Die Kombination dieser beiden Größen kann den Kontrast ebenso stark beeinflussen wie die Richtung und die Geometrie der Beleuchtung. Beispielsweise ergibt ein rot gedruckter Strichcode auf weißem Untergrund unter Tageslicht für den menschlichen Betrachter meist einen akzeptablen Kontrast. Bei Beleuchtung mit rotem LED-Licht wird der Kontrast hingegen deutlich geringer, denn sowohl die rote Druckfarbe als auch der weiße Untergrund remittieren einen großen Teil des roten Lichts. Ein hoher Kontrast entsteht, wenn die Beleuchtung im Wesentlichen blaues Licht enthält, weil der weiße Untergrund auch im Blauen den größten Teil der einfallenden Strahlung remittiert, während die rote Druckfarbe das blaue Licht weitgehend absorbiert. Eine Beleuchtung im Infraroten wird in diesem Fall den Kontrast nahezu zum Verschwinden bringen, weil die Remission der meisten roten Druckfarben sich im Infraroten nur geringfügig vom Rückstreuverhalten eines weißen Papiers oder Kartons unterscheidet. Entsprechend kann ein blauer Aufdruck unter der breitbandigen Beleuchtung einer Glühlampe, aufgenommen mit einer Standard-Graustufen-Kamera, zu hohem Bildkontrast führen, während ein roter Aufdruck nahezu im Untergrund verschwindet. Abbildung 5 zeigt ein Beispiel. Hinzu kommt hier, dass eine Glühlampe im Infraroten mit deutlich höherer Strahlungsleistung emittiert als im Blauen. Da eine Standard-Kamera auch im Bereich von 700 nm bis 900 nm noch eine nennenswerte Empfindlichkeit hat, wird der Bildkontrast durch die spektralen Verhältnisse im Roten und im Infraroten dominiert. Durch ein IR-Sperrfilter vor der Kamera kann dieser Effekt bei Bedarf stark reduziert werden. Wahlweise kann also durch gezielte Abstimmung der Beleuchtung auf die spektralen Eigenschaften der Prüflinge und der Kamera der Kontrast im Bild angehoben oder stark abgeschwächt werden. Wenn ein Druck-

bild auf weißem Untergrund geprüft werden soll, ist eine Beleuchtung angemessen, deren Spektralverteilung komplementär zur Remission der Druckfarbe verläuft. Wenn der Aufdruck unterdrückt werden soll, z.B. bei der Überprüfung von bedruckten Textilien auf Webfehler, sollte in einem Spektralbereich beleuchtet werden, in dem die Remission oder die Transparenz der Druckfarbe hoch ist.

### Weitere Gesichtspunkte

In diesem kurzen Artikel können nur einige der Aspekte behandelt werden, die für die Auslegung einer Beleuchtung bei einer Bildverarbeitungslösung wichtig sind. Im Einzelfall können andere Gesichtspunkte dominieren. Bei der Präzisionsvermessung ist die Homogenität der Beleuchtung oft wesentlich für die erreichbare Genauigkeit. Inhomogene Beleuchtungsverteilungen können jedoch in Grenzen algorithmisch korrigiert werden („shading-Korrektur“). Auch die Kurz- und Langzeitstabilität einer Beleuchtungslösung ist für viele Methoden in der industriellen Bildverarbeitung wichtig. Häufig wird eine Binarisierungsschwelle in Abhängigkeit von der Bildhelligkeit dynamisch festgelegt, und gelegentlich wird die Helligkeit der Beleuchtung geregelt. Die vielfältigen Fragen im Zusammenhang mit der „Intensität“ wurden hier nur gestreift, können jedoch in manchen Anwendungen entscheidend sein. In diesem Zusammenhang sind auch die Leistungsaufnahme und die Wärmeentwicklung von Bedeutung. Polarisation und Kohärenz sind weitere Parameter, die bei speziellen Prüfaufgaben beachtet werden müssen. Beispielsweise können Reflexe an Folien, Glas und metallischen Oberflächen unter Umständen mit Polarisationsfiltern unterdrückt werden, und bei Verwendung von Laserlichtquellen können lästige Tupfen (Speckle) entstehen. Andererseits können Laser sehr gut zur Ausleuchtung von Linien oder von komplexeren Mustern eingesetzt werden, z.B. beim Lichtschnittverfahren. Auch die gezielte Erzeugung von Schatten bei der 3D-Vermessung („shape from shading“) und die strukturierte Beleuchtung (Projektion von Streifenmustern mit Hilfe von Flüssigkristall-Arrays) können hier nur genannt, aber nicht erläutert werden. Nicht zuletzt sind auch die Verfügbarkeit und die Kosten der verschiedenen Beleuchtungseinrichtungen wichtige Kriterien bei der Systemauslegung.

### Professionalisierung

Noch vor zehn Jahren gab es nur wenige Firmen, die spezielle Beleuchtungseinrichtungen für die Bildverarbeitung gebaut haben. Mittlerweile wird man auf jeder einschlägigen Messe mehrere Firmen finden, die sich auf diesen Sektor spezialisiert haben. Interessant ist, dass darunter auch renommierte Hersteller von hochwertigen Abbildungsoptiken sind, die ihr Produktspektrum um LED-Beleuchtungen ergänzen und dabei ihre Optik-Kompetenz einbringen. Beleuchtung wird mittlerweile als lukrativer Teil des Bildverarbeitungsmarktes wahrgenommen und als anspruchsvolle Ingenieurdisziplin begriffen. Tatsächlich hat die entscheidende Entwicklung in den letzten zehn Jahren also in den Köpfen stattgefunden, indem „...die Bedeutung der Beleuchtung ins Bewusstsein innovativer Hersteller von Systemen zur industriellen Bildverarbeitung ... gedrungen ist ...“ [6]. Es ist kein Zufall, dass parallel die Berechnung von nicht-abbildenden Optiken Eingang in die professionellen Werkzeuge des „optical design“ gefunden hat. Beleuchtungskonzepte für Fußballstadien oder moderne Autoscheinwerfer werden heute mit Optik-CAD-Programmen gerechnet und optimiert. Diese Methoden lassen sich zwanglos auf Beleuchtungen für die Bildverarbeitung übertragen. LEDs mit ihrer kompakten Bauform und

Abb. 5: Ein blau gedruckter Barcode (oben) und ein rot gedruckter Barcode (unten), beleuchtet mit einer Glühlampe und aufgenommen mit einer Graustufenkamera



der Kompatibilität zu Standard-Elektroniken eröffnen Möglichkeiten für Beleuchtungskonzepte, die erst mit Optik-Design-Programmen ausgeschöpft werden können. Die wichtigste technische Entwicklung der letzten zehn Jahre in diesem Bereich ist jedoch die allgemeine Verfügbarkeit von hinreichend hellen weißen LEDs. Deren Erfinder, Shuji Nakamura, wurde im Jahr 2006 mit dem Millennium Technology Prize ausgezeichnet [7].

### Die Zukunft

Während Licht streng genommen auf den Wahrnehmungsbereich des Menschen begrenzt ist, sind die Silizium- und InGaAs-Detektoren unserer Standardkameras bis in den nahen IR-Bereich hinein empfindlich. In den nächsten Jahren werden voraussichtlich die spektrale Zusammensetzung der Beleuchtung und die Ausweitung auf den Strahlungsbereich außerhalb des Sichtbaren weiter in den Mittelpunkt des Interesses rücken. Farbbildverarbeiter schauen bereits auf den Farbwiedergabeindex von LED-Beleuchtungen, und der IR-Bereich, der mit Silizium-Detektoren abtastbar ist, wird schon in vielen Anwendungen genutzt. Schon jetzt kann man spektral programmierbare Lichtquellen kaufen, und es wird langsam deutlich, dass eine spektrale Diskriminierung nicht zwingend auf der Empfängerseite erfolgen muss, sondern auch durch die Beleuchtung bewerkstelligt werden kann. Spektakuläre Technologiesprünge sind jedoch nicht zu erwarten, denn Innovation ist ein langsamer Prozess und findet in erster Linie in den Köpfen statt.

### Literaturangaben

- [1] H. Ernst, „Einführung in die digitale Bildverarbeitung, Grundlagen und industrieller Einsatz mit zahlreichen Beispielen“, Franzis-Verlag, München 1991
- [2] W. Burger, M. J. Burge, „Digitale Bildverarbeitung“, Springer 2005
- [3] B. Jähne, „Digitale Bildverarbeitung“, Springer 2005, Kapitel 6, „Quantitative Visualisierung“, und Kapitel 7.5, „Radiometrie der Abbildung“
- [4] C. Steger, M. Ulrich, C. Wiedemann, „Machine Vision Algorithms and Applications“, Wiley-VCH, Weinheim 2008
- [5] Stemmer Imaging, „Handbuch der Bildverarbeitung“, Puchheim 2010
- [6] Schott Info 100/2002, S. 21 ff.
- [7] Der Millennium Technology Prize ist eine renommierte Auszeichnung für technische Erfindungen, s. [www.millenniumprize.fi](http://www.millenniumprize.fi)

# Perfekte Farben - jederzeit



## 3-CCD-Kameras von JAI

Fordert Ihre Anwendung höchste Farbgenauigkeit? Dann entscheiden Sie sich am besten für die präziseste Farbkamera der Branche. Die 3-CCD Kameras der AT-Serie von JAI setzen fortschrittliche Prismentechnologie ein, für eine Farbgenauigkeit, die weit über die der Bayer-Kameras liegt - ohne dass Details durch Farbinterpolationen verloren gehen.

Für scharfe Konturen oder das Lesen kleiner Barcodes ist keine Verdoppelung oder Verdreifachung der Bayer-Auflösung erforderlich. Mit Kameras der AT-Serie erkennen Sie selbst feinste Farbnuancen und winzige Elemente ohne Leistungseinbußen, verursacht durch höhere Auflösungen. Rufen Sie uns an oder besuchen Sie unsere Website, wenn Sie mehr über die farbenfreudigen Details erfahren möchten.



### AT-140

- 1392 x 1040 (1,4 MP)
- 4,65 µm quadratische Pixel
- 25 fps Camera Link oder 20 fps GigE Vision
- Bis zu 36-Bit-RGB-Ausgang

### AT-200

- 1628 x 1236 (2,0 MP)
- 4,40 µm quadratische Pixel
- 20 fps Camera Link oder 15 fps GigE Vision
- Bis zu 36-Bit-RGB-Ausgang



Weitere Informationen

Europa, Naher Osten & Afrika: +45 4457 8888  
 Nord-, Mittel- und Südamerika: +1 800 445-5444  
 Asien-Pazifik: +81 45-440-0154  
[www.jai.com](http://www.jai.com)



See the possibilities

# Objektive

## Abbildung mit Maß

Beleuchtung und optische Abbildung bilden das „front-end“ eines Bildverarbeitungssystems. Ein Objektiv setzt die beleuchtete 3D-Szene im Gegenstandsraum in ein 2D-Bild auf dem Bildaufnehmer um. Bildverarbeiter stehen oft vor der Aufgabe, den Arbeitsabstand bei vorgegebener Objektgröße zu bestimmen, die Objektivbrennweite zu spezifizieren oder einen Zwischenring auszusuchen. Dabei werden sie mit Begriffen wie „Format“, „C-Mount“, „Schärfentiefe“ oder „Telezentrie“ konfrontiert. Eine kurze Darstellung der wesentlichen Zusammenhänge bei der optischen Abbildung und die Erläuterung einiger optischer Grundbegriffe im Hinblick auf typische Fragestellungen bei der Bildverarbeitung soll die Orientierung erleichtern.

Ein Teil der Strahlung, die von einem beleuchteten Objekt ausgeht, tritt in die Objektivöffnung ein und erzeugt ein Bild in der Bildebene. Die Symmetrieachse des Objektivs, die „Blickrichtung“, wird als optische Achse bezeichnet. Der Detektor einer Kamera muss senkrecht auf der optischen Achse stehen und darauf zentriert sein.

### Format

Die Abmessungen des Detektorchips legen fest, welcher Ausschnitt aus der Bildebene als digitales Bild erfasst wird. Das Objektiv muss passend zur verwendeten Chipgröße, dem sog. Format, gewählt werden.

Ein Format wird durch eine Zollangabe spezifiziert. Üblich sind die Angaben 1“, 2/3“, 1/2“, 1/3“ und 1/4“. Die Bezeichnung bezog sich ursprünglich auf den Außendurchmesser einer Kameraröhre und hat wenig mit der tatsächlich nutzbaren Bilddiagonale zu tun. Die traditionellen Bezeichnungen wurden auf die Halbleiter-Detektorchips übertragen. Die Abmessungen der lichtempfindlichen Fläche eines sog. „1-Zoll“-Chips betragen heute etwa 12,7 mm in der Breite und etwa 9,5 mm in der Höhe. Die Diagonale ist mit etwa 15,9 mm deutlich kürzer als die 25,4 mm, die die Formatbezeichnung von 1 Zoll nahelegen. Grob gerechnet ist die lange Seite eines Detektorchips etwa halb so lang wie das Maß in der Bezeichnung des Formats. Die genauen Abmessungen variieren je nach Hersteller und sollten bei Bedarf erfragt werden.

Für eine Kamera mit einem „1-Zoll-Chip“ kann (und sollte) demnach ein „1-Zoll-Objektiv“ geordert werden. Es kann problemlos auch für eine Kamera mit kleinerem Format verwendet werden. Manchmal werden höherformatige Objektive eingesetzt, weil im Innern des vorgesehenen Bildkreises bessere Abbildungseigenschaften erwartet werden als bei einem Objektiv mit angepasstem Format. Wird umgekehrt eine Kamera mit größerem Format verwendet, sind zum Bildrand hin erhebliche Verzerrungen, starker Helligkeitsabfall oder sogar Vignettierung möglich.

### Auflagemaß: „C-mount“ und „CS-mount“

Ein typisches Kameraobjektiv für die Bildverarbeitung besteht zwar aus mehreren Linsen, verhält sich jedoch in mancher Hinsicht in guter Näherung wie eine einfache, dünne Sammellinse. Parallel zur optischen Achse einfallende Strahlen werden auf der Bildseite im Brennpunkt gesammelt, so dass für ein weit entferntes Objekt ein Bild in der Brennebene entsteht. Wird das Objekt auf die Kamera zu bewegt, wandert das Bild von der Linse weg. Damit auf dem feststehenden Detektor ein scharfes Bild entstehen kann, muss die Linse auf das Objekt zu bewegt, also aus dem Objektiv herausgezogen werden. Dazu dient der Scharfstellring am Objektiv. Am mechanischen Anschlag ist der maximale Objektivauszug erreicht, so dass es einen Mindestabstand zwischen Objekt und Objektiv gibt, den minimalen Arbeitsabstand (im Englischen „minimum object distance“, MOD). Unterhalb dieser Entfernung können Objekte nicht mehr scharf abgebildet werden.

Auch in Richtung auf den Detektorchip ist der Verfahrweg der Linse begrenzt. Sehr weit entfernte Objekte müssen scharf abgebildet werden können. Der minimale Abstand zwischen Linse und Detektor darf daher nicht größer sein als die Brennweite. Damit der Objektivhersteller diese Bedingung einhalten kann, ist der Abstand zwischen der Anlagefläche am Kameragehäuse und dem Detektor, das sog. Auflagemaß, genormt. Ein Objektiv muss passend zum Auflagemaß der Kamera gewählt werden. Abbildung 1 verdeutlicht die Verhältnisse.

In der Bildverarbeitung dominieren Objektive mit „C-mount“ (Auflagemaß etwa 17,5 mm) und mit „CS-mount“ (Auflagemaß ca. 12,5 mm). Eine C-Kamera, die mit einem CS-Objektiv versehen ist, wird für weit entfernte Objekte kein scharfes Bild liefern, weil der Brennpunkt vor dem Detektorchip liegt und die Linse nicht mehr näher an den Detektor gebracht werden kann. Umgekehrt kann eine CS-Kamera durchaus mit einem C-Objektiv betrieben werden. Für weit entfernte Objekte muss die Linsengruppe jedoch bereits aus dem Objektivtubus herausgezogen werden, so dass der Verfahrweg eingeschränkt und damit der minimale Arbeitsabstand vergrößert wird. Die bessere Lösung ist ein ca. 5 mm dicker Adapterring zwischen Kamerakörper und Objektiv, der aus der CS-Kamera eine C-Kamera macht.

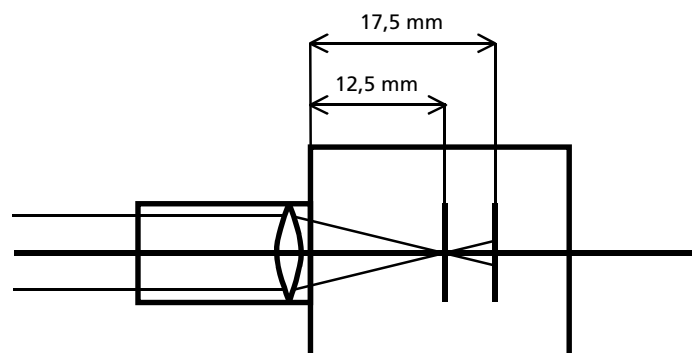


Abb. 1: Ein CS-mount-Objektiv, ausgelegt für ein Auflagemaß von 12,5 mm, kann an einer C-mount-Kamera mit Auflagemaß 17,5 mm weit entfernte Objekte nicht scharf auf den Detektor abbilden.



## Zwischenringe

Wenn der minimale Arbeitsabstand für die Anwendung zu groß ist, hilft ein Ring zwischen Kameragehäuse und Objektiv. Der Abstand zwischen Linse und Detektor wird damit vergrößert, so dass auch näher gelegene Objekte scharf abgebildet werden können. Zwischenringe werden deshalb auch als Makro-Ringe bezeichnet. Durch die Montage des Zwischenrings vergrößert sich jedoch das Auflagemaß, so dass es auch einen maximalen Arbeitsabstand gibt. Jenseits dieses Abstandes ist keine scharfe Abbildung möglich. Die optisch bessere Lösung ist ein Objektiv mit geringerer Brennweite.

## Welche Brennweite?

Welche Objektivbrennweite ist erforderlich, wenn bei gegebenem Chipformat und Arbeitsabstand ein bestimmter Ausschnitt eines Objekts scharf abgebildet werden soll? Die erforderliche Brennweite kann in guter Näherung mit einer einzigen Formel bestimmt werden.

Abbildung 2 zeigt die typische Abbildungssituation der Bildverarbeitung. Das Objekt mit der Größe  $y$  wird verkleinert in das invertierte Bild mit der Größe  $y'$  abgebildet. Entsprechend DIN 1355 wird der Abstand zwischen Objekt und Linsenmitte, die Gegenstandsweite  $a$ , negativ gerechnet. Der Abstand  $a'$  zwischen Linsenmitte und Bild, die Bildweite, ist positiv, ebenso die bildseitige Brennweite  $f'$ . Die nach oben gerichtete Gegenstandsgröße  $y$  ist positiv, die Bildgröße  $y'$  negativ. Zur Vereinfachung sind nur zwei Strahlen gezeichnet, die von der Spitze des Objekts ausgehen. In Wirklichkeit tragen alle Strahlen zur Abbildung bei, die von einem Objektpunkt aus in das Objektiv einfallen und die Objektivblende passieren.

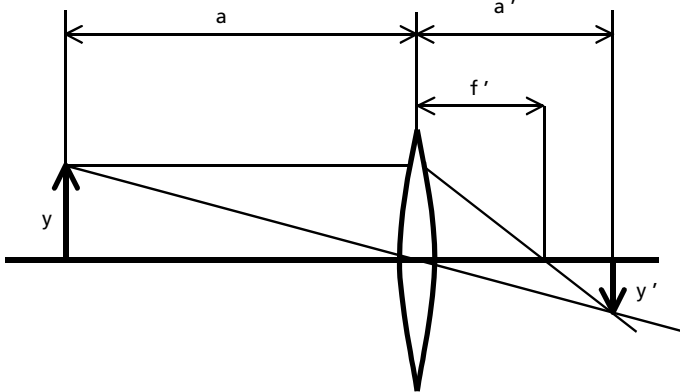


Abb. 2: Abbildung mit einer dünnen Sammellinse. Das Objekt befindet sich links, das Bild rechts von der Linse.  $y$ : Gegenstandsgröße (positiv);  $y'$ : Bildgröße (negativ);  $a$ : Gegenstandsweite (negativ);  $a'$ : Bildweite (positiv);  $f'$ : Brennweite (positiv)

Die geometrischen Verhältnisse werden quantitativ durch die Abbildungsgleichung beschrieben:

$$1/a' - 1/a = 1/f'$$

Außerdem gilt die Beziehung:

$$\beta' = y'/y = a'/a$$

$\beta'$  ist der Abbildungsmaßstab. Wegen der Inversion bei der Abbildung ist  $\beta'$  negativ.

Eine praktische Umformung dieser Beziehungen ist:

$$f' = -a/[1-(y'/y)] = a\beta'/(1-\beta')$$

Die Gegenstandsweite  $a$ , also der Abstand zum Objekt, ist meist bekannt. Die Objektgröße  $y$  ist z.B. eine Kantenlänge des gewünschten rechteckigen Ausschnitts im Objektraum. Die Größe  $y'$  im Bildraum ist dann die entsprechende Kantenlänge des Detektorchips, die sich aus dem Kameraformat ergibt. Aus  $y$  und  $y'$  wird dann  $\beta'$  berechnet. Bei der Bestimmung von  $f'$  ist dann zu beachten, dass  $a$  und  $\beta'$  negativ sind. Die Bildweite  $a'$  ergibt sich beim „Scharfstellen“ und ist belanglos, solange sie durch den Objektivauszug mechanisch realisiert werden kann.

Die Formel darf nur angewendet werden, wenn die Gegenstandsweite  $a$  deutlich größer als die Länge des Objektivs ist. Handelsübliche Objektive sind sehr kompakt, so dass in der industriellen Bildverarbeitung die Bedingung meist - aber nicht immer - erfüllt ist. Nur dann kann die Gegenstandsweite mit genügender Genauigkeit bis zur Mitte des Objektivs gemessen werden. Der exakte Bezugspunkt, der Hauptpunkt des Objektivs, wird nur selten spezifiziert.

Ein Beispiel: Ein quadratischer Bereich mit 5 cm Kantenlänge soll von einer 1“-Kamera im Abstand von 30 cm erfasst werden. Die Länge der kurzen Kante des lichtempfindlichen Bereichs des Detektorchips wird mit 9,6 mm angegeben. Folglich ist  $\beta' = y'/y = -9,6 \text{ mm} / 50 \text{ mm} = -0,192$ . Mit  $a = -30 \text{ cm} = -300 \text{ mm}$  ergibt sich:

$$f' = a\beta'/(1-\beta') = (-300 \text{ mm})(-0,192)/1,192 = 48,3 \text{ mm}$$

Wenn das Objektiv 30 mm lang ist, ist die Objektweite  $a$  auf 15 mm entsprechend ca. 5% genau bekannt. Die o.g. Näherung ist daher gerechtfertigt. Das Resultat für  $f'$  ist nur auf etwa 5% genau, so dass Objektive mit Brennweiten zwischen etwa 46 mm und etwa 51 mm in Frage kommen.

Die Formel zeigt, dass der Arbeitsabstand verringert werden kann, ohne dass sich der Abbildungsmaßstab ändern muss. Dazu ist nur eine kleinere Brennweite erforderlich. So kann z.B. die Baulänge einer Prüfstation und damit der mechanische Aufwand bei gleichbleibendem Gesichtsfeld deutlich reduziert werden. Allerdings muss geprüft werden, ob die erforderliche Güte der optischen Abbildung noch gewährleistet ist und ob die kleinere Bildweite nicht zu mechanischen Problemen führt. Bei festem Abstand zwischen Kamera und Objekt liegt auch der Abbildungsmaßstab fest. Ein größerer Abbildungsmaßstab wird dann nur durch Verkleinerung der Brennweite erzielt. Wenn lediglich ein größerer Ausschnitt aus dem Objektraum abgebildet werden soll, kann unter sonst gleichen Bedingungen eine Kamera mit einem größeren Chip eingesetzt werden. Der Detektor erfasst dann einen größeren Bereich des Bildfeldes, ohne dass Abbildungsmaßstab, Brennweite und Arbeitsabstand geändert werden müssen. Das Objektiv muss jedoch für dieses Format geeignet sein.

## Kontrollrechnung

Zoom-Objektive, bei denen die Brennweite in Grenzen kontinuierlich verstellbar ist, werden in der Bildverarbeitung wegen der höheren Kosten, der größeren Baulänge und der geringeren Lichtstärke meist nur im Labor verwendet. Grundsätzlich kann ein Hersteller innerhalb des technisch zugänglichen Bereichs ein Objektiv beliebiger Brennweite herstellen. Es gibt Fälle, in denen dies sinnvoll ist, und die Kosten sind für Kleinserien oft geringer als befürchtet. Für Einzelanwendungen wird jedoch die erforderliche Brennweite berechnet und dann ein Objektiv aus einer Stan-

Standardserie mit möglichst ähnlicher Brennweite verwendet. Die erforderlichen Korrekturen können dann meist durch „Scharfstellen“, mit Zwischenringen oder durch Verschiebung der Kamera vorgenommen werden. Es ist jedoch sinnvoll, die Auswirkungen auf die anderen Bestimmungsgrößen zu berechnen und zu prüfen, ob die Werte akzeptabel sind. Die Beziehung für  $f'$  kann dazu nach den anderen beiden Größen aufgelöst werden:

$$a = f' (1 - \beta') / \beta'$$

$$\beta' = f' / (f' + a) = y' / y$$

Die Größen  $\beta'$ ,  $a$  und  $y'$  tragen negatives Vorzeichen,  $f'$  und  $y$  sind positiv.

Im Zweifelsfall ist eine kürzere Brennweite günstiger, weil der gewünschte Ausschnitt damit eher vom Detektor erfasst werden kann. Die Ungenauigkeit für die Bestimmung des Arbeitsabstands sollte berücksichtigt werden. Geringfügige Änderungen der Ausgangsgrößen können sich im Ergebnis stark auswirken. Manche Hersteller bieten für diese Berechnungen

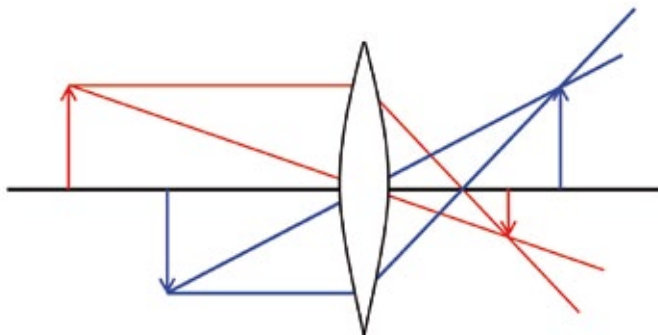


Abb. 3: Abbildung zweier gleich großer Objekte mit einem Standardobjektiv: Die Bildgröße hängt von der Objektentfernung ab

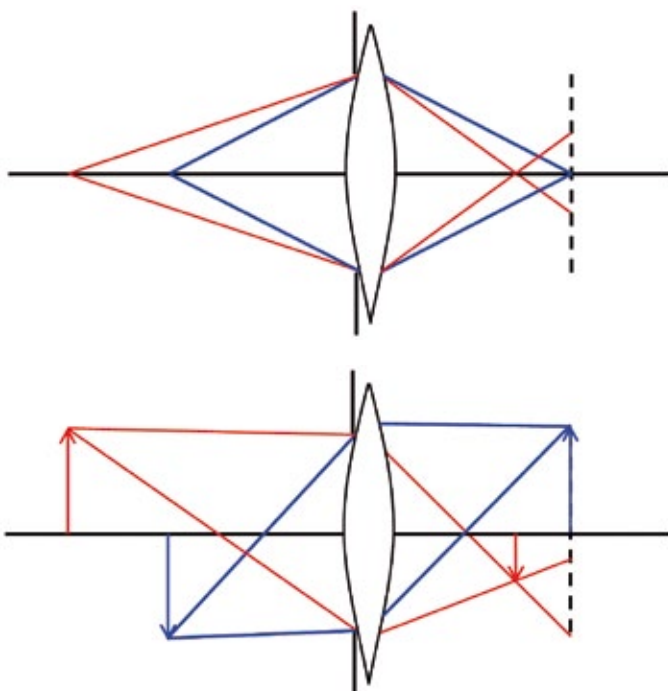


Abb. 4: Wenn der Detektor in der Bildebene des blauen Pfeils liegt, wird der rote Pfeil unscharf abgebildet; Abb. 4a zeigt den Strahlengang für den Fußpunkt, 4b für die Spitze des roten Pfeils

PC-Programme an, in denen bereits alle erforderlichen Objektivdaten abgelegt sind.

### Blendenzahl und Schärfentiefe

Objektive sind mit einer verstellbaren Irisblende versehen. Bei größerer Blendenöffnung fällt eine größere Lichtleistung auf den Detektor, so dass das Bild heller wird. Meist wird als Kenngröße die Blendenzahl  $k$  verwendet, das Verhältnis zwischen der Brennweite  $f'$  und dem Durchmesser DEP der Eintrittspupille:

$$k = f' / \text{DEP}$$

Je größer  $k$ , desto kleiner ist der Blendendurchmesser und desto weniger Lichtleistung fällt auf den Detektor. Wer ein helleres Bild möchte, muss folglich eine kleinere Blendenzahl einstellen. Viele Objektive haben Rasterungen für feste Blendenzahlen nach der Normreihe, z.B. für  $k = 0,71; 1; 1,4; 2$ , die sich jeweils um den Faktor  $\sqrt{2}$  unterscheiden. Die Lichtleistung ist proportional zur Fläche der Blendenöffnung und ändert sich von Stufe zu Stufe um den Faktor 2.

Der Sichtbereich des Objektivs wird durch eine kleinere Blendenöffnung nicht eingeschränkt. Die Abbildung wird dadurch lediglich auf achsennahe Strahlen beschränkt. Damit wird auch die Schärfentiefe größer. Bei fester Einstellung des Objektivs gibt es nur einen Punkt auf der optischen Achse, der optimal scharf abgebildet wird. Wird dieser Punkt längs der optischen Achse verschoben, ohne dass die Lage der Objektivlinse nachkorrigiert wird, entsteht in der Bildebene ein Scheibchen, der Unschärfekreis. Wenn der Unschärfekreis nicht zu groß ist, wird man immer noch von einer scharfen Abbildung sprechen. Oberhalb und unterhalb des optimal scharf abgebildeten Punktes gibt es also eine Zone, in der Objekte im Bild ebenfalls als scharf abgebildet erscheinen. Die Ausdehnung dieses Bereichs, gemessen längs der optischen Achse, ist die Schärfentiefe. Je größer die Blendenöffnung, je kleiner also die Blendenzahl, desto kleiner ist die Schärfentiefe. Der Absolutwert für die Schärfentiefe hängt vom akzeptablen Durchmesser des Unschärfekreises ab. In der Bildverarbeitung wird dafür bislang noch kein einheitliches Maß verwendet. Die Schärfentiefe hängt außerdem von der Brennweite und vom Arbeitsabstand ab.

### Telezentrische Abbildung

Standardobjektive für die Bildverarbeitung bilden in Zentralprojektion ab. Ein Objekt erscheint im Bild umso kleiner, je weiter es von der Kamera entfernt ist. Ein dreidimensionales Objekt, das sich auch längs der optischen Achse erstreckt, kann damit zwar bei ausreichender Schärfentiefe in seiner gesamten Tiefe scharf abgebildet werden. Der Abbildungsmaßstab für eine Kante am oberen Rand des Objekts ist jedoch ein anderer als am unteren Rand.

Für Vermessungszwecke sind telezentrische Objektive günstiger. Sie lassen nur Strahlen zur Abbildung zu, die parallel zur optischen Achse verlaufen. Damit wird eine Parallelprojektion erreicht. Das Bild eines Objekts, das sich längs der optischen Achse bewegt, bleibt gleich groß.

Um die Situation genauer zu verstehen, betrachten wir zunächst zwei gleich große Objekte, die sich in verschiedenen Abständen vom Objektiv befinden. In Abbildung 3 ist diese Situation dargestellt. Zur Verdeutlichung sind ein roter und ein blauer Pfeil senkrecht zur optischen Achse gezeichnet, jeweils mit ihrem Fußpunkt auf der optischen Achse, aber entgegengesetzt gerichtet. Für jeden der beiden Pfeile kann man die Lage des scharfen Bildes konstruieren, wenn man weiß, dass Licht-

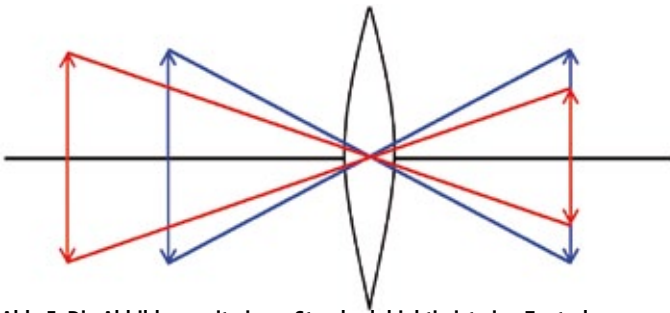


Abb. 5: Die Abbildung mit einem Standardobjektiv ist eine Zentralprojektion

strahlen, die parallel zur optischen Achse in das Objektiv einfallen, durch den bildseitigen Brennpunkt verlaufen und Lichtstrahlen, die durch das Zentrum des Objektivs verlaufen, nicht gebrochen werden. Zur Konstruktion verwendet man die Spitze des Pfeils; dort, wo sich die beiden Konstruktionsstrahlen im Bildraum schneiden, liegt das scharfe Bild der Pfeilspitze. Entsprechend kann man für den Mittelpunkt und weitere Punkte des Pfeils konstruieren und stellt fest, dass das Bild in einer Ebene senkrecht zur optischen Achse entsteht. Insbesondere wird auch der Fußpunkt des Pfeils von der optischen Achse aus wieder auf die optische Achse abgebildet. Die Konstruktion zeigt sofort, dass die beiden gleich großen Pfeile zu verschiedenen großen scharfen Bildern abgebildet werden: das weiter entfernte Objekt erscheint im Bild kleiner als das näher liegende Objekt. Der Abbildungsmaßstab ist für die beiden Positionen verschieden, er hängt von der Objektweite ab. Die beiden scharfen Bilder sind jedoch nicht nur unterschiedlich groß, sondern entstehen zudem in verschiedenen Bildweiten, also in unterschiedlichen Bildebenen. Wenn der Objektabstand variiert, wie im oben beschriebenen Anwendungsszenario, müsste man folglich das Objektiv nachfokussieren, damit ein scharfes Bild entsteht. In der Praxis ist das nur in Ausnahmefällen handhabbar. Man wird vielmehr mit einer festen Bildebene arbeiten und in Kauf nehmen, dass das Bild unscharf wird.

Abbildung 4 zeigt, wie die Unschärfe entsteht. Die geometrischen Verhältnisse sind dieselben wie in Abbildung 3, es ist jedoch nur der Fußpunkt der beiden Pfeile gezeichnet. Aus der Konstruktion in Abbildung 3 ist bereits bekannt, wo die jeweiligen Bildebenen liegen. In Abbildung 4a sind nur die Randstrahlen eines Strahlenbündels gezeichnet, das zur Bildentstehung beitragen kann, wenn die Fassung der Linse als Blende wirkt. Wir nehmen an, dass sich der Detektorchip der Kamera in der Bildebene des näher liegenden Objekts befindet. Da sich die Strahlen für das weiter entfernte Objekt bereits vor der Bildebene schneiden, werden sie als divergentes Strahlenbündel auf den Detektor fallen und dort ein Scheibchen als Bild erzeugen. Je weiter das Objekt entfernt ist, desto größer wird das Scheibchen, desto unschärfer das Bild. Abbildung 4b verdeutlicht dieselbe Situation, jedoch nun für einen Objektpunkt außerhalb der optischen Achse. Die Spitze des weiter entfernten Pfeils wird ebenfalls unscharf abgebildet. Qualitativ ist erkennbar, dass das entstehende Bild kleiner ist als das Bild des scharf abgebildeten Objekts. Folglich wird auch bei feststehender Bildebene ein Objekt mit zunehmender Entfernung von der Kamera in ein kleiner werdendes Bild abgebildet – eine Erkenntnis, die man innerhalb kürzester Zeit empirisch gewinnen kann, wenn man im Labor eine Kamera von einem Objekt wegbewegt. Alternativ genügt es, mit offenen Augen bewusst in die weite Welt zu schauen, denn unsere Augen verhalten sich dann in dieser Hinsicht ebenfalls wie Standardobjektive. Quantitativ können die Verhältnisse nun jedoch nicht mehr mit der elementaren Abbildungsgleichung beschrieben werden, sondern erfordern kom-

plexere Methoden der technischen Optik. Insbesondere müssen die wirksamen Blenden im System und Beugungseffekte berücksichtigt werden [1].

## Zentralprojektion und Parallelprojektion

Die Abbildung mit einem Standardobjektiv ist eine Zentralprojektion. Zur Veranschaulichung sind in Abbildung 5 beide Objekte aus Abbildung 3 in beiden Abständen gezeichnet und nur die Strahlen eingetragen, die durch das Zentrum der Linse verlaufen. In dieser einfachen Darstellung wird sofort klar, dass die Bildgröße bei der Zentralprojektion vom Objektabstand abhängt. Wenn wir die Zentralprojektion durch eine Parallelprojektion ersetzen könnten, wäre diese Abhängigkeit eliminiert. Wir benötigen folglich eine optische Anordnung, bei der nur diejenigen Strahlen zur Bildentstehung beitragen, die parallel zur optischen Achse verlaufen. Bei einer einfachen Sammellinse verlaufen alle Strahlen, die parallel zur optischen Achse in die Linse eintreten, im Bildraum durch den Brennpunkt – das ist die Definition des Brennpunkts. Folglich wird eine Blende, die im bildseitigen Brennpunkt positioniert ist, dafür sorgen, dass nur die gewünschten Strahlen das System bis zum Detektor durchlaufen können, während alle anderen Strahlen abgeblockt werden. Abbildung 6a zeigt diese Situation: das Konzept eines objektseitig telezentrischen Objektivs. Man kann zeigen, dass mit diesem Verfahren in einer feststehenden Bildebene die Bildgröße unabhängig von der Objektweite ist, der Abbildungsmaßstab also unabhängig von der Objektentfernung gleich bleibt [1]. Aus Intensitätsgründen kann der Blendendurchmesser nicht beliebig weit reduziert werden, so dass die Telezentrie

# designing views

60 years of superior optical experience

in developing solutions for your success

KAIZEN – 100% japanese quality

150 different standard lenses, e.g.

3.5 – 850 mm focal length

SWIR – day & night

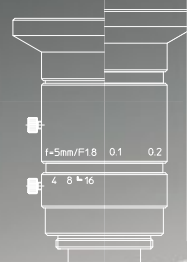
up to 10MP lenses

1/3" c-mount lenses

HD-Multi-Megapixel

custom-made lenses

special coating solutions



Kowa Optimed Deutschland GmbH  
Immermannstraße 43B  
40210 Düsseldorf  
Germany  
fn +49-(0)211-179354-0  
lens@kowaeurope.de  
www.kowa.eu/lenses





nur näherungsweise realisiert werden kann. Telezentrische Objektive sind meist für einen festen Arbeitsabstand ausgelegt und haben einen definierten Telezentriebereich um diesen Arbeitsabstand herum. Innerhalb des Telezentriebereichs ist die verbleibende Änderung des Abbildungsmaßstabs spezifiziert. Für ein gutes telezentrisches Objektiv kann beispielsweise im Telezentriebereich eine Bildgrößenänderung von maximal 1 µm im gesamten Bildkreis angegeben sein. Wenn man diese Grenze bei einer Vermessungsaufgabe nutzen will, ist es sinnvoll, auch die übrigen optischen Parameter eines solchen Objektivs zu optimieren. Ein Restfehler, der wesentlich zur Ungenauigkeit einer Vermessung beitragen kann, ist die Verzeichnung. Daher sind telezentrische Objektive häufig auch in dieser Hinsicht hochwertig und mit Verzeichnungen von weit unter einem Prozent spezifiziert.

### Beidseitige Telezentrie

Die Konstanz des Abbildungsmaßstabs im Telezentriebereich eines objektseitig telezentrischen Objektivs gilt nur für eine feste Bildebene senkrecht zur optischen Achse. Außerdem können Kanten im Bild asymmetrisch werden, weil die Strahlen schräg auf die Detektorebene fallen. Da sich die Bildgröße im Telezentriebereich in der Regel nur um Bruchteile einer Pixelkantenlänge ändert, werden meist Subpixelverfahren zur Vermessung eingesetzt, die empfindlich auf Veränderungen im Kantenprofil reagieren. In einem zweistufigen System können diese Nachteile verhindert werden. Dazu wird die Blende in den Brennpunkt eines zweiten Linsensystems gestellt und so ins Unendliche abgebildet, d.h. die Strahlen treten parallel zur optischen Achse aus der zweiten Linsengruppe aus und fallen symmetrisch auf den Detektor, unabhängig von der Höhe im Bildkreis (s. Abb. 6b). Man kann zeigen, dass bei diesen sog. beidseitig telezentrischen Objektiven die Bildgröße auch bei variabler Bildebene konstant bleibt [1]. Der Abbildungsmaßstab ist dann das Verhältnis der beiden Brennweiten. Solche Objektive werden auch mit der Möglichkeit zur Nachfokussierung angebo-

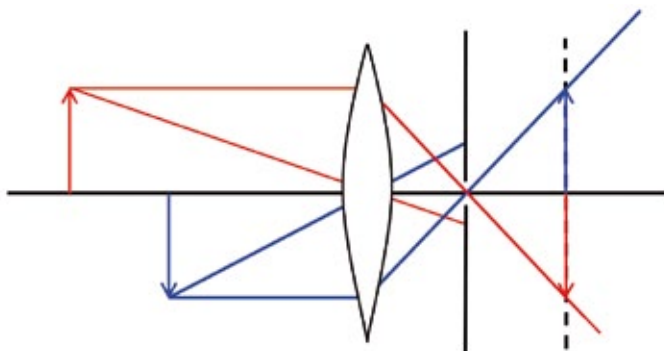


Abb. 6a: Ein objektseitig telezentrisches System entsteht durch eine Blende im Brennpunkt; für die feste Detektorebene ist der Abbildungsmaßstab unabhängig von der Objektentfernung

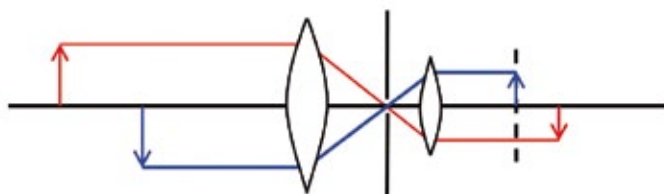


Abb. 6b: Beim beidseitig telezentrischen System ist zusätzlich die Bildgröße unabhängig von der Lage der Detektorebene

ten, so dass der Arbeitsabstand an die Applikation angepasst werden kann. Insgesamt hat dieses Konzept große Ähnlichkeit mit dem Kepler-Fernrohr, allerdings mit dem Unterschied, dass im telezentrischen Objektiv nur Strahlen parallel zur optischen Achse zugelassen werden. Eine für die Anwendung wesentliche Konsequenz dieser Einschränkung auf eine Parallelprojektion ist die Tatsache, dass der Durchmesser der Eintrittsoptik eines telezentrischen Objektivs mindestens so groß sein muss wie das Objekt, das damit abgebildet werden soll. Für die Prüfung von Bahnware über Breiten in der Größenordnung von Metern oder andere ausgedehnte Objekte sind telezentrische Objektive daher schwerlich als ein einziges Bauelement realisierbar. Ein Ausweg ist eine gestaffelte Phalanx von mehreren telezentrischen Objektiven, deren Bildbereiche sich überlappen. Aufgrund ihrer optischen Konstruktion haben telezentrische Objektive zudem im Vergleich zu Standardobjektiven eine relativ große Baulänge [2], können eine Masse von mehreren Kilogramm haben, sind lichtschwach und teuer. Für hochgenaue Vermessungsaufgaben gibt es jedoch oft keine Alternative zu diesen Spitzenprodukten der technischen Optik.

### Weitere Objekteigenschaften

Eine Reihe weiterer Gesichtspunkte, die bei der Auswahl eines Objektivs von Bedeutung sind, können hier nur kurz angesprochen werden [3]. Generell sollten in der Bildverarbeitung nur Objektive eingesetzt werden, die speziell für Kameras mit Halbleiterdetektoren ausgelegt sind. Für Detektorchips mit hoher Pixelauflösung (Megapixel-Kameras) müssen unter Umständen andere Objektive verwendet werden als für Standard-Videokameras. Wer besonders feine Strukturen abbilden muss, sollte sich beim Hersteller nach der Übertragungsfunktion (MTF) des Objektivs für hohe Ortsfrequenzen erkundigen und auch nachfragen, ob in der Kombination mit dem gewählten Detektorchip Artefakte durch Unterabtastung (aliasing) entstehen können. Ein qualifizierter Kundenberater wird die Verhältnisse gewiss gern erläutern. Die Abbildungsfehler von Objektiven sind bei Vermessungsaufgaben oft wichtig und können ebenso quantifiziert werden wie die Lichtleistung, die ein Objektiv von einer beleuchteten Szene auffangen kann. Die meisten Objektive sind für den sichtbaren Spektralbereich ausgelegt. Wenn Kameras bzw. Detektoren ohne IR-Sperrfilter eingesetzt werden, kann die Güte der optischen Abbildung erheblich leiden. Auch der Streulichtanteil eines Objektivs kann für einige Anwendungen wichtig sein. Schließlich sollte auch bei Monochromkameras bedacht werden, dass Objektive, die schlecht farbkorrigiert sind, an einer Kante Farbsäume erzeugen, die sich im Monochrombild als Verschmierung äußern. Für Anwendungen in der Industrie ist auch die mechanische Robustheit eines Objektivs wichtig. Verlässliche Feststellelemente für die Fokussierung und die Blendeneinstellung sind nützlich, und die Rüttel- und Schockfestigkeit kann in rauer Industrieumgebung entscheidend sein.

Objektive für die Bildverarbeitung sind mechanisch und optisch hochwertige Hightech-Komponenten. Das Objektiv für eine Bildverarbeitungsanwendung sollte folglich mit ähnlicher Sorgfalt ausgewählt werden wie die Kamera und die Beleuchtung.

### Literatur

- [1] Lenhardt, K., Optical measurement techniques with telecentric lenses, Schneider-Kreuznach, [www.schneiderkreuznach.com/knowhow.htm](http://www.schneiderkreuznach.com/knowhow.htm)
- [2] Schuhmann, R., Thöniß, T., Technisches Messen 65 (1998) 4, S. 131 ff.
- [3] Einen guten Überblick gibt Lenhardt, K., Optical Systems in Machine Vision, in A. Hornberg (ed.), Handbook of Machine Vision, Wiley-VCH 2006

**Zoom-Objektiv jetzt auch motorisiert**



Das Navitar-Objektiv Zoom 7000 ist jetzt in neuer Version auch mit motorisierter Blende und Fokus bei Polytec verfügbar. Die Motorisierung kann mittels der beiliegenden Windows- oder Labview-kompatiblen Software über den seriellen RS-232- oder USB-Anschluss am Controller angesteuert werden, der als Platinen- oder Gehäuseversion verfügbar ist.

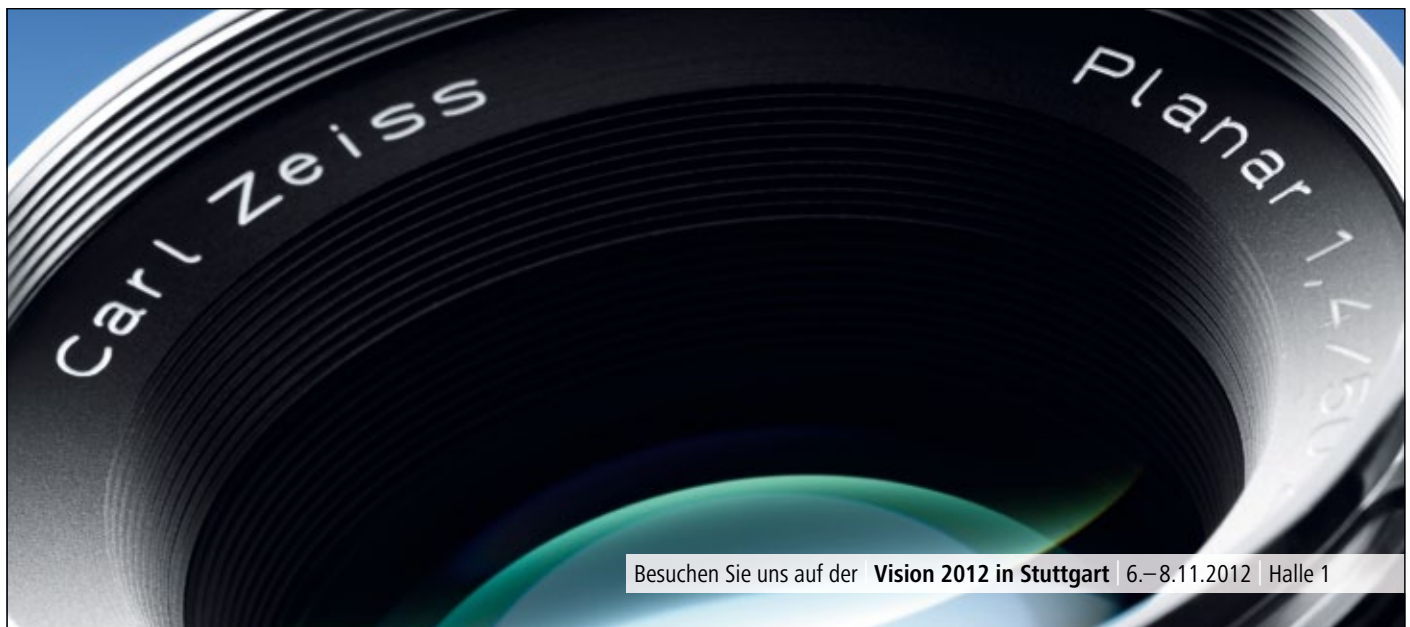
Wahlweise können Zoom und Iris, nur Zoom oder Zoom, Fokus und Iris motorisiert werden. Dabei sind die Zoom-, Fokus- und Blendenpositionen reproduzierbar. Die Motorkonstruktion ist für raue Einsatzbedingungen ausgelegt. Navitar hat für die Positionsreferenzierung magnetische Halleffektsensoren integriert. Halleffektsensoren sind elektronische Komponenten ohne bewegliche Teile. Die Funktionsweise beruht auf dem Halleffekt, der 1879 entdeckt und nach seinem Entdecker Edwin Hall benannt wurde. Der Halleffekt-Abtastmechanismus löst mit Hilfe eines Magnetfelds einen Impuls in einer Halbleiterschaltung sehr präzise aus. Im Gegensatz zu fotoelektrischen Sensoren werden Halleffektsensoren durch schmutzige oder staubige Umgebungen, Umgebungslicht oder Netzspannungsschwankungen nicht beeinträchtigt. Typische Anwendungen sind darüber hinaus die automatisierte Qualitätskontrolle und Montage, biomedizinische Anwendungen sowie die Bauteilkontrolle von Leiterplatten- und elektronischen Komponenten.

[www.polytec.delzoom7000](http://www.polytec.delzoom7000)

**Neue RGB-Farbzeilenkameras mit 3 x 7600 Pixel**

Die neuen Farbzeilenkameras von Schäfter+Kirchhoff SK22800GJRC-XC, SK22800VJRC-XC und SK22800CJRC-XC sind mit 3 x 7.600 Pixel (RGB) für hochauflösende Oberflächeninspektionen z. B. in der Holz- und Druckindustrie prädestiniert. Die Kameras verfügen über eine Gigabit Ethernet Schnittstelle. Die Pixel werden mit 120 MHz ausgelesen. Die maximale Zeilenrate beträgt 4,93 kHz. Die SK22800CJRC-XC wird über eine CameraLink-Schnittstelle mit einem Grabber verbunden und über diesen gesteuert. Mit Pixelfrequenzen von bis zu 150 MHz (3x 50 MHz) werden Zeilenfrequenzen von 6,17 kHz erreicht. In allen drei Zeilenkamerateypen arbeitet der gleiche Farbzeilensensor. Aufgrund der Pixelgröße von 9,3 x 9,3 µm sind die Kameras sehr lichtempfindlich. Der Triple-Line Sensor hat zwischen den RGB-Sensorzeilen ein geringes Line-Spacing der einfachen Pixelhöhe. Dadurch ist der Farbkonvergenzfehler bei frei bewegten Objekten gering, z. B. beim Einsatz in der Schüttgutortierung. In Verbindung mit hochauflösenden Objektiven sind die Farbzeilenkameras auch für den Einsatz in der Mikroskopie geeignet.

[www.SuKHamburg.de](http://www.SuKHamburg.de)



Besuchen Sie uns auf der **Vision 2012 in Stuttgart** 6.–8.11.2012 Halle 1

**Perfektion für hohe Ansprüche**



**ZF-Objektive von Carl Zeiss – Fokussiert auf Qualität**

Hochmoderne Technologie erfordert exakte visuelle Kontrollen, damit die heutigen Qualitätsstandards erfüllt werden können. Carl Zeiss liefert Ihnen die optische Technologie, die zu Ihrer Innovation passt. Unsere ZF-Objektive begeistern durch ihre optische Leistung und mechanische Robustheit. Das bedeutet für Sie einen zuverlässigen Partner bei der Erreichung Ihrer Ziele.

[www.zeiss.com/lenses4industry](http://www.zeiss.com/lenses4industry)



We make it visible.

# Der Hintergrund

## Tabula rasa

Bei industriellen Anwendungen ist ein Bild mit einem homogenen Hintergrund oft besonders günstig für die nachfolgenden Verarbeitungsschritte. Es ermöglicht eine sichere, robuste Trennung von Vordergrund und Hintergrund. Dieser Beitrag behandelt einige einfache Verfahren der Bildverarbeitung, die in dieser Hinsicht für die Charakterisierung einer Beleuchtungssituation nützlich sein können.

In der industriellen Bildverarbeitung soll die Beleuchtung die relevanten Merkmale von Objekten hervorheben und die irrelevanten Eigenschaften unterdrücken. Im einfachsten Fall erzeugt die Beleuchtung ein nahezu binäres Bild, in dem sich z.B. dunkle Objekte mit deutlichem Kontrast von einer hellen Unterlage abheben. Abbildung 1a zeigt ein typisches Beispiel. Die meisten Bildverarbeiter sind sehr zufrieden mit der Beleuchtung, wenn mit einer globalen Binarisierungsschwelle das Graustufenbild in ein brauchbares Binärbild überführt werden kann und der Schwellwert innerhalb weiter Grenzen verändert werden darf, ohne dass die Segmentierung nennenswert leidet.

### Grauwertverteilung

Oft wird eine solche Situation anhand des Grauwert-Histogramms beurteilt. In Abbildung 1b ist das Histogramm der Szene aus Abbildung 1a dargestellt. Das Histogramm zerfällt in diesem Fall in zwei deutlich getrennte Bereiche. Die Häufung um den Grauwert 25 herum gehört zu den dunklen Objektpixeln, die Häufung um den Grauwert 215 zu den hellen Hintergrundpixeln. Eine Binarisierung mit der globalen Schwelle 104 führt zu dem Binärbild in Abbildung 1d mit klarer Trennung von Objekten und Hintergrund.

Damit scheint es sich um eine ideale Beleuchtungssituation zu handeln. Bei näherer Betrachtung ist die Angelegenheit jedoch komplizierter. Das Bild in Abbildung 1a zeigt Münzen, die im Durchlicht aufgenommen worden sind. Münzen sind vollständig undurchlässig. Folglich sollten nur zwei scharfe Linien im Histogramm erscheinen, die jeweils zum Grauwert für die dunklen Münzen und für den hellen Hintergrund gehören. Wie kann also dort eine Verteilung mit einer Breite von jeweils etwa 20 Graustufen entstehen? Noch deutlicher wird die Problematik bei einem Blick auf Abbildung 1c. Dort ist auf der Ordinate nach oben nicht die Anzahl, sondern die Wurzel aus der Anzahl der Pixel mit dem jeweiligen Grauwert aufgetragen. Im Bereich zwischen den beiden Häufungen kommen offenbar sämtliche möglichen Grauwerte tatsächlich im Bild vor. Ein Grund für die Zwischenwerte im Histogramm ist die Bildgebung an den Kanten der Münzen. Dort entstehen Reflexionen und Schatten, zusätzlich trägt die optische Abbildung zu einer Verschleifung der Kanten bei. Dadurch entsteht ein kontinuierliches Kantenprofil, das mit dem Detektorraster diskret abgetastet wird. Dabei können grundsätzlich alle möglichen Grauwerte zwischen dem Wert für den hellen Hintergrund und dem dunklen Objekt erzeugt werden. Hinzu kommt das Rauschen des Bildsignals, das bei Standardkameras nicht

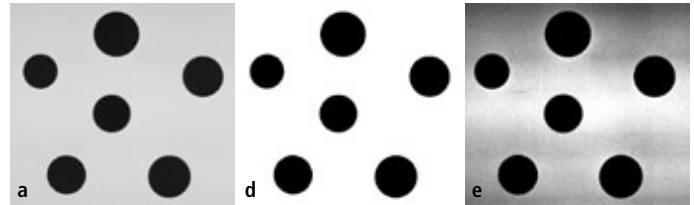
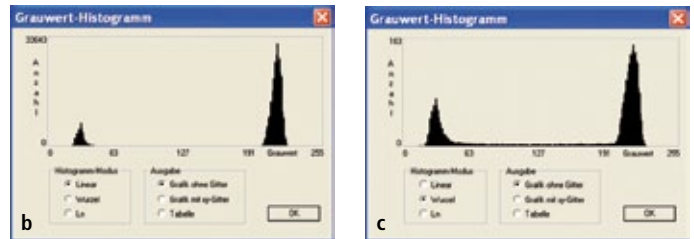


Abb. 1: Eine Durchlichtaufnahme, das zugehörige Grauwert-Histogramm mit linearer Skala und mit Wurzelskalierung, das daraus entstehende Binärbild und das Ergebnis einer Grauwertspreizung



zu unterschätzen ist. Auch die unterschiedliche Empfindlichkeit der Detektorpixel und die unterschiedlichen Dunkelsignale („fixed pattern noise“) tragen zur Verteilung der Graustufen bei.

Das Histogramm des Prüfbildes erlaubt demnach keine eindeutigen Rückschlüsse auf die Güte der Beleuchtung. Mehr Information liefert eine Aufnahme von einer beleuchteten Fläche, ohne dass sich Objekte im Bild befinden. Als Beispiel ist in Abbildung 2a ein Ausschnitt aus dem Hintergrund von Abbildung 1a dargestellt. Das zugehörige Histogramm besteht aus einer glockenförmigen Häufung im Bereich zwischen den Grauwerten 206 und 225. Grauwerte außerhalb dieses Intervalls kommen im Bildausschnitt gar nicht vor. Die Spannweite beträgt lediglich 20 Grauwerte, der Median liegt bei 216, und die Standardabweichung ist 2,7. Diese globalen statistischen Größen oder davon abgeleitete Größen, die z.B. auf den Zentralwert der Verteilung normiert sind, können für manche Anwendungen durchaus zur Charakterisierung der Güte einer Beleuchtung geeignet sein.

### Lokale Strukturen

Globale Kenngrößen geben aber nicht unmittelbar Aufschluss über Strukturen im Bild. Selbst bei so günstig erscheinenden statistischen Parametern wie für Abbildung 2a kann man Überraschungen erleben. Abbildung 2b zeigt denselben Bildausschnitt, jedoch nach einer Kontrastanhebung. In diesem Beispiel wurden die Grauwerte zwischen 210 und 225 auf den vollen Grauwertumfang gespreizt, also alle Grauwerte unterhalb 211 auf Null gesetzt, alle Grauwerte oberhalb von 224 auf 255, und alle Grauwerte dazwischen linear auf den Bereich von 0 bis 255 verteilt. Die Bildinformation hat sich dadurch nicht geändert, sie ist lediglich für einen menschlichen Betrachter besser zugänglich. Es ist deutlich erkennbar, dass die Helligkeitsverteilung keineswegs homogen und auch nicht nur statistisch verrauscht ist, sondern dass es einen Helligkeitsverlauf und Flecken im Bild gibt. Noch deutlicher wird die Situation, wenn diese Operation auf das gesamte Bild angewandt wird. Abbildung 1e zeigt das Ergebnis für eine Spreizung des Grauwertintervalls von 190 bis 220. Hier treten Strukturen und Effekte in



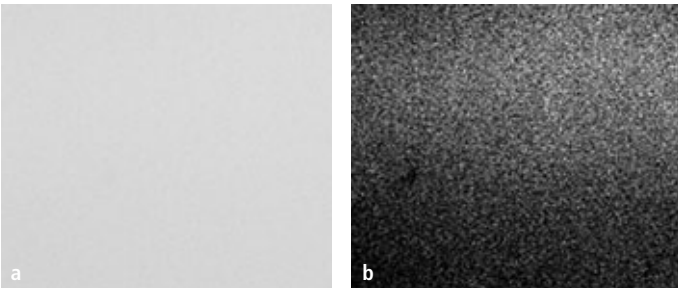


Abb. 2: Ein Ausschnitt aus dem Hintergrund von Abbildung 1 sowie das Ergebnis einer Grauwertspreizung für diesen Ausschnitt

Erscheinung, die allein mit globalen statistischen Parametern nicht aufgedeckt werden können. Ein weiteres, spektakuläres Beispiel zeigt Abbildung 3a, eine Aufnahme eines kommerziell erhältlichen Durchlichtfeldes. In diesem Bild kommen Grauwerte im Intervall von 59 bis 93 vor, die Standardabweichung beträgt 5,1. Die Spreizung der Grauwerte zwischen 70 und 90 in Abbildung 3b zeigt jedoch sehr deutlich, dass es massive Strukturen in der Beleuchtung gibt.

### Vielfältige Einflussfaktoren

Bei dieser Vorgehensweise handelt es sich nicht um die Charakterisierung der Beleuchtung selbst, sondern um die Kombination der Einflüsse der Beleuchtung, der Abbildungsoptik und der Eigenschaften des Detektors sowie der nachfolgenden Signalverarbeitung. Beispielsweise kann eine Verzeichnung durch das Objektiv dazu führen, dass die tatsächliche Verteilung der Bestrahlungsstärke auf der Probe nicht gleichmäßig in ein Bild-signal umgesetzt wird. Außerdem haben viele Objektive, die für die Bildverarbeitung eingesetzt werden, einen Randabfall, d.h. zum Bildfeldrand hin wird von einer ideal homogen leuchtenden Fläche weniger Licht auf den Detektor übertragen als im Zentrum. Letztlich geht es bei einer Applikation aber nicht in erster Linie um die Eigenschaften der Beleuchtung selbst, sondern um die Ausprägung der Bildinformation, die durch das Zusammenspiel von Beleuchtung, optischer Abbildung und Detektor samt Folgeelektronik erzeugt wird. Oft wirkt sich eine kontinuierliche Variation der Helligkeitsverteilung im Bild auch gar nicht störend aus. Wenn beispielsweise die Verarbeitung der Bilddaten auf Kantenerkennung beruht, kommt es auf den globalen Verlauf der Helligkeit gar nicht wesentlich an. Letztlich ist bei den meisten Bildverarbeitungsaufgaben ein guter lokaler Kontrast entscheidend, denn die Bildverarbeitung setzt weitgehend an lokalen Grauwertunterschieden an.

Inhomogenitäten der Beleuchtung dürfen also nicht zu vorgeäuschten lokalen Kontrasten führen. Solche Artefakte können jedoch in sehr vielen verschiedenen Ausprägungen auftreten, vergleichbar der Problematik bei der Prüfung von Oberflächen auf Kratzer, Einschlüsse und andere Fehler, die nicht vollständig vorher beschrieben werden können.

### Systematische Analyseverfahren

Als systematischer Ansatz ist ein Gradientenbild geeignet, das die lokalen Steigungen zwischen benachbarten Grauwerten erfasst. Auch Verfahren aus der Texturanalyse können für diesen Zweck geeignet sein. Als Beispiel zeigt Abbildung 4a die Grauwertübergangsmatrix, die sog. Co-occurrence-Matrix, für die Szene aus Abbildung 1a. Auf den beiden Achsen des Bildes sind die möglichen Grauwerte im Bild aufgetragen, in diesem Beispiel der Bereich von 0 bis 255. Für jedes Pixel im Bild wird nun der Grauwert ausgelesen und untersucht, welchen Grau-

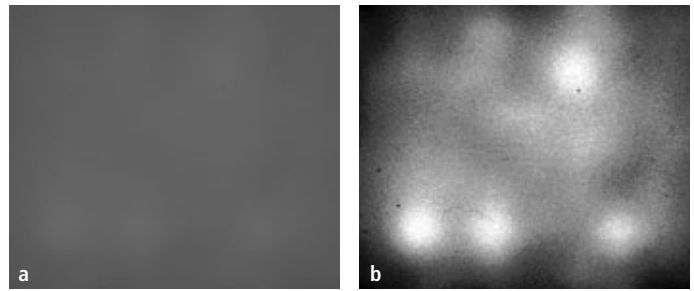


Abb. 3: Die Aufnahme eines kommerziellen Durchlichtfeldes und das Ergebnis einer Grauwertspreizung

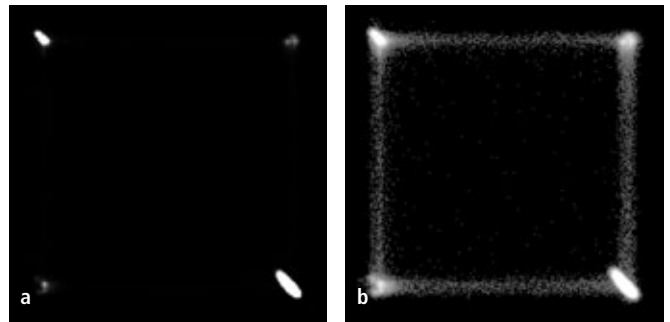


Abb. 4: Die Grauwertübergangsmatrix für Abbildung 1 und eine Nachbearbeitung zur Hervorhebung weiterer Strukturen in der Matrix

wert das benachbarte Pixel hat. Hat das aktuelle Pixel z.B. den Grauwert 100 und das Nachbarpixel den Grauwert 120, wird im Ergebnisbild der Zähler für die Matrixposition (100, 120) um eins erhöht. Für eine perfekt homogene Verteilung, bei der alle Grauwerte im Bild identisch sind, wird also nur ein einziger Punkt in der Matrix belegt. Die Co-occurrence-Matrix in Abbildung 4a ist etwas komplizierter. Es gibt zwei sehr helle Bereiche auf der Diagonalen, die den homogenen dunklen Flächen der Objekte und dem hellen Hintergrund entsprechen. Auf der Gegendiagonalen sind zwei schwächer ausgeprägte Bereiche zu sehen, die den Übergängen an den Kanten zuzuordnen sind. Bei der Erzeugung dieser Matrix wurde nicht der Grauwert des direkt benachbarten Pixels abgeprüft, sondern eines Pixels, das sich 16 Pixel rechts und 16 Pixel unterhalb des aktuellen Pixels befindet. Auf diese Weise tritt der Einfluss der Kanten stärker hervor, die Reichweite der lokalen Betrachtung wird erhöht. In Abbildung 4b wurde diese Matrix noch einmal kontrastverstärkt, d.h. die Helligkeit der Bereiche entspricht nun nicht mehr linear der Zahl der Eintragungen in der Matrix. Dadurch wird erkennbar, welche weiteren Grauwertsprünge im Bild vorhanden sind. Diese Darstellung ist ein nützliches Hilfsmittel bei Überlegungen zur Festlegung eines Toleranzbereichs für die Graustufenschwelle bei einer Binarisierung, für die Auslegung einer adaptiven Schwelle oder für die Kantenerkennung. Auch die lokalen Variationen eines Hintergrundbildes kann man damit genauer charakterisieren. Die Co-occurrence-Matrix ist jedoch immer noch ein komplexes Gebilde. Es ist nahe liegend, aus dieser Matrix Kenngrößen zu berechnen, die zur Beurteilung der Grauwertverteilung im Hintergrund für die jeweilige Applikation geeignet sind. Beispielsweise könnte abgeprüft werden, dass die Matrix nur innerhalb einer vorgegebenen Zone in einem Teilbereich der Diagonale belegt ist. Andere Kenngrößen werden in der Texturanalyse verwendet und sind in der Literatur beschrieben [1].

### Literatur

- [1] Gonzales, R. C., Woods, R. E., Digital Image Processing, Addison-Wesley 1993, p. 508 ff.

# Kontrast und Schärfe

Einfach scharf!

Ob ein optisches System – und damit auch das Objektiv eines Bildverarbeitungssystems – bestimmte Details einer Szene „scharf“ abbilden wird, kann man anhand der sog. Modulations-Übertragungsfunktion (MTF, modulation transfer function) beurteilen. Diese Erkenntnis aus der Systemtheorie nützt jedoch einem Techniker, Betriebsingenieur oder Betriebsleiter wenig, der an der Linie steht und sich fragt: Wie kann ich „das Bild scharf stellen“? In diesem Artikel werden einige Begriffe erläutert, die mit dieser Fragestellung zusammenhängen. Wer mag, kann anschließend seine Bildverarbeitungsapplikation um eine Routine erweitern, die ein Gütemaß für die Bildschärfe ausgibt oder eine Autofokus-Funktion implementieren.

Jeder digitale Photoapparat, der sich auf dem Consumer-Markt behaupten will, nimmt dem Kunden das „Scharfstellen“ per „Autofokus“ ab. Bei Bildverarbeitungssystemen für industrielle Anwendungen ist diese Funktionalität jedoch keineswegs allgemein üblich. Ein wesentlicher Grund für diesen, auf den ersten Blick verblüffenden, Zustand ist die Tatsache, dass bei industriellen Anwendungen der Abstand zwischen Kamera und Objektebene meist mechanisch gut definiert ist. Daher genügt es, beim Einrichten der Anlage einmalig für die ausreichende Abbildungsqualität zu sorgen. Anwender sind gut beraten, die dabei festgelegte Einstellung des Objektivs gegen Veränderungen zu schützen, die z.B. durch Vibrationen oder durch Hantieren in der Anlage entstehen können. Objektive für industrielle Anwendungen haben zu diesem Zweck häufig spezielle Feststellschrauben oder sind in Käfigen montiert. Schwankungen der Gegenstandsweite sind zwar auch bei Industrieanwendungen

oft unvermeidbar, beispielsweise bei Zuführung der Teile mit einem Förderband, die Auswirkung einer Abweichung von der optimalen Gegenstandsweite auf die Schärfe im Bild hängt jedoch davon ab, wie weit die Blende des Objektivs geöffnet ist. Wenn die Beleuchtung ausreichend hell ist, kann die Blende weitgehend zugezogen werden, und die Schärfentiefe ist groß, d.h. die Abbildung ist auch dann noch scharf, wenn das Objekt längs der optischen Achse um eine bestimmte Strecke verschoben wird. Muss die Blende jedoch weit geöffnet werden, kann der Schärfentiefebereich sehr schmal sein, unter Umständen nur wenige Millimeter. Wer eine Anwendung plant, die in dieser Hinsicht robust sein muss, sollte sich daher um eine intensive Beleuchtung kümmern.

## Schärfentiefe und Blende

Die Abhängigkeit der Schärfentiefe von der Blende ist nützlich, wenn man die Abbildung manuell „scharf“ einstellen möchte. Dazu wird zunächst die Blende weit geöffnet und die Beleuchtung so weit abgedunkelt, dass ein noch nicht übersteuertes Bild entsteht. In dieser Situation ist der Schärfentiefebereich sehr schmal, und geringe Abweichungen vom optimalen Objektvorauszug haben bereits deutliche Bildunschärfen zur Folge, die man im Livebild gut erkennen kann. Auch ungeübte Bediener finden auf diese Weise schnell und intuitiv die korrekte Einstellung des Objektvorauszugs. Im Idealfall wird die Blende für den Messbetrieb dann wieder zugezogen und der Prüfbereich mit höherer Intensität ausgeleuchtet. Der Schärfentiefebereich wird dann groß sein, und Schwankungen im Objektabstand werden sich wegen der optimalen Einstellung des Objektivs nur sehr geringfügig auf die Bildschärfe auswirken.

## Was ist scharf?

Manchmal ist man jedoch gezwungen, die Einstellungen bei weitgehend geschlossener Blende vorzunehmen, wenn z.B. die

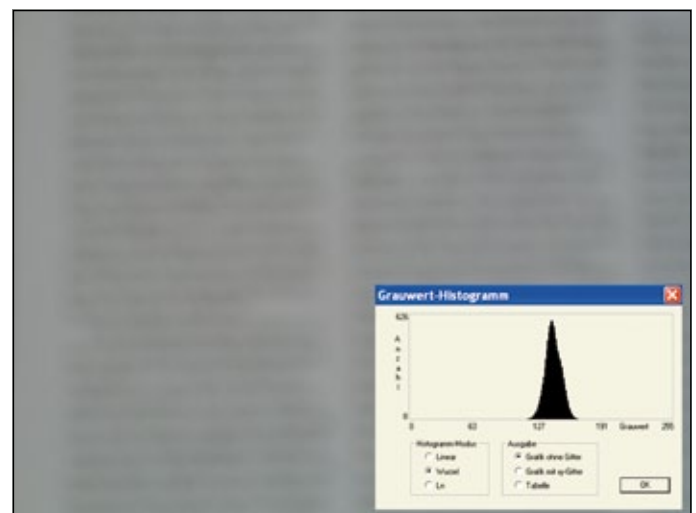
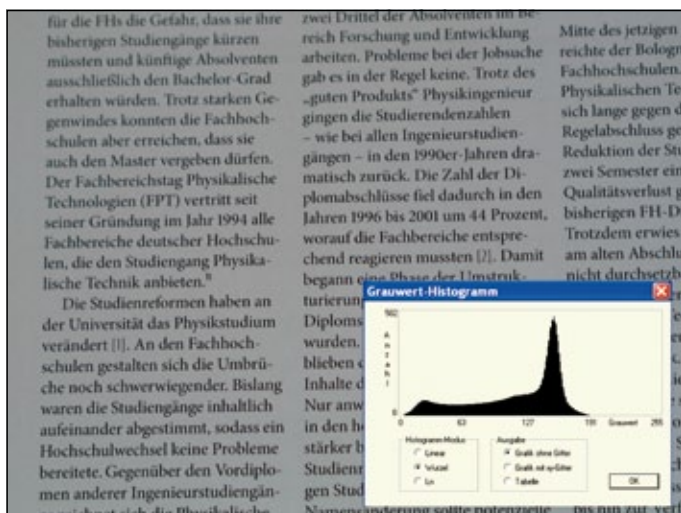


Abb. 1: Zwei Bilder derselben Szene, aufgenommen mit unterschiedlichen Objektvorauszügen unter ansonsten gleichen Bedingungen. In den zugehörigen Grauwert-Histogrammen ist zur Verdeutlichung der Strukturen die Wurzel aus der absoluten Häufigkeit aufgetragen (erstellt mit AdOculus von The Imaging Source)

Installation einer Hilfsbeleuchtung zu kompliziert, aus Platzgründen nicht möglich oder zu langwierig ist. Bei großem Schärfentiefebereich ist es manuell sehr schwierig, den optimalen Objektivauszug durch Beobachtung des Bildes auf einem Monitor zu finden. In solchen Fällen ist es hilfreich, einen numerisch bewerteten Parameter aus dem Bild zu ermitteln, der als Maß für die „Bildschärfe“ dienen kann. Jeder hat a priori eine Vorstellung davon, was ein unscharfes und was ein scharfes Bild ist. Deshalb sind wir bisher in diesem Artikel auch ohne eine weitere Erläuterung oder Definition des Begriffs „Bildschärfe“ ausgekommen. Wenn Sie jedoch überlegen, was Sie eigentlich genau tun, wenn Sie ein Bild „scharf stellen“, und wie Sie merken, dass es noch unscharf ist oder bereits hinreichend scharf, werden Sie feststellen, dass Sie es nicht wirklich wissen. Da wir mit unserem körpereigenen visuellen System permanent und unwillkürlich scharfe Bilder auf unserer Netzhaut einregeln, erscheint uns dieser Vorgang als vollkommen selbstverständlich, und die Frage, nach welchem Gütekriterium wir uns dabei richten, löst manchmal interessante, gelegentlich sogar stark emotional gefärbte Reaktionen aus. Betrachten wir also ein scharfes und ein sehr unscharfes Bild einer Szene, z.B. die beiden Bilder in Abbildung 1. Beim unscharfen Bild sind die in der Szene vorhandenen Strukturen kaum noch erkennbar, die Buchstaben und die Zwischenräume verschwimmen zu beinahe ununterscheidbaren Flächen. Einerseits sind die Kanten der Strukturen in der Szene sehr stark verschliffen, also weniger steil als beim scharfen Bild. Andererseits ist der Unterschied zwischen den Grauwerten der dunklen Zeichen und der hellen Zwischenräume beim unscharfen Bild wesentlich geringer als beim scharfen Bild. Bei weiterer Defokussierung der optischen Abbildung werden die Kanten noch flacher, und die Zeichen verschmelzen mit den Zwischenräumen zu Zonen mit einheitlichem Grauwert. Aus diesen Beobachtungen ergeben sich zwei Ansätze für die Berechnung eines Güteparameters aus dem Bild, der die „Bildschärfe“ quantitativ wiedergibt.

### Güteparameter für die Bildschärfe

Viele Bildverarbeiter werden sofort daran denken, die Steilheit der Kanten im Bild zu bestimmen. Ein Maß dafür ist die Steigung eines Grauwertprofils durch die Kante senkrecht zur Kantenrichtung, z.B. als Differenzenquotient  $(g(x+dx)-g(x-dx))/2dx$ .

Im einfachsten Fall bildet man für eine Kante, die parallel zur y-Achse des Bildes verläuft, lediglich  $g(x+1)-g(x)$ ; je größer dieser Wert, desto steiler ist die Kante. Entsprechende zweidimensionale Verallgemeinerungen sind als Kantenfilter in vielfältiger Form aus der Literatur bekannt. Ihre Anwendung auf das Quellbild ergibt als Ergebnis ein neues, genauso großes Bild, in dem die Kantensteilheit für jedes Pixel als Grauwert codiert ist. Homogene Flächen erscheinen darin dunkel, unabhängig von ihrem ursprünglichen Grauwert, weil die Steigung des Grauwertverlaufs darin sehr gering ist. Lediglich im Bereich von Kanten ergeben sich hellere Zonen. Dieses Kantenbild kann man nun weiter verarbeiten. Beispielsweise kann man es mit einer empirisch ermittelten Schwelle binarisieren und einfach alle Pixel mit Grauwerten oberhalb der Schwelle zählen, also alle Pixel an Kanten mit einer Steigung größer als ein vorgegebener Wert. Nimmt man nun bei konstanter Intensität der Beleuchtung und konstanter Blendeneinstellung eine Fokussierungsreihe mit verschiedenen Objektivauszügen auf, wird die Zahl der Pixel oberhalb der Binarisierungsschwelle beim optimal scharf eingestellten Bild maximal. Wenn das Verfahren jedoch unabhängig vom Bildinhalt und unabhängig von der Intensität der Beleuchtung gute Ergebnisse liefern soll, muss die optimale Schwelle für die Binarisierung des Kantenbildes dynamisch aus dem Bild berechnet werden. Die entsprechenden Ansätze sind vielen Bildverarbeitern bekannt: Sie werden bei der Segmentierung von Graustufenbildern in Vordergrund und Hintergrund mittels Binarisierung häufig eingesetzt. Selbstverständlich sind auch komplexere Methoden als das einfache Pixelzählen zur Aufbereitung des Kantenbildes geeignet. Schon das hier skizzierte einfache Verfahren funktioniert recht gut, es ist jedoch wegen der Filteroperation mit erheblichem Rechenaufwand verbunden.

### Hinweise aus dem Histogramm

Eine einfachere Methode ergibt sich aus dem zweiten oben diskutierten Gesichtspunkt. Wenn bei der Defokussierung letztlich die Grauwertdifferenzen zwischen Vordergrund und Hintergrund zumindest lokal immer geringer werden, muss sich dies in einer gleichmäßig ausgeleuchteten Szene unmittelbar im Grauwert-Histogramm äußern. In Abbildung 1 ist zu jedem der beiden Bilder auch das zugehörige Histogramm dargestellt. In

## Applikationsspezifische Kameraserien

**Kappa optronics GmbH**  
Germany | USA | France | Australia  
www.kappa.de

realize visions .



dem scharf eingestellten Bild ist die Spannweite der Grauwerte deutlich größer als im unscharfen Bild: Der kleinste Grauwert ist dort 0, der größte 223, während im unscharfen Bild nur der Bereich von 108–190 belegt ist. Der Mittelwert der Grauwerte ist in beiden Bildern übrigens annähernd gleich groß, nämlich 136,2 im scharfen Bild und 138,2 im unscharfen Bild. Beide Bilder sind also im Mittel annähernd gleich hell, sie unterscheiden sich jedoch erheblich in der Verteilung der Grauwerte. Eine beliebte und bekannte Quantifizierung einer Verteilung ist die Standardabweichung. Beim scharfen Teilbild der Abbildung 1 beträgt die Standardabweichung der Grauwerte 30,7 und beim unscharfen Teilbild 5,5. Tatsächlich kann man auch mit Hilfe dieses globalen statistischen Parameters eine brauchbare ObjektivEinstellung erreichen. Allerdings ist die Verteilung beim scharf eingestellten Bild stark asymmetrisch, und ein Statistiker würde die Standardabweichung daher vielleicht nicht als optimale Kenngröße ansehen. Sie kann jedoch direkt aus dem Grauerthistogramm berechnet werden und ist somit schnell und einfach zugänglich. Noch einfacher wird die Charakterisierung der Verteilung über ihre Spannweite, also die Differenz  $g_{max}-g_{min}$  zwischen dem größten und dem kleinsten auftretenden Grauwert im Bild. Im scharf eingestellten Bild aus Abbildung 1 beträgt die Spannweite 223, im unscharf eingestellten Bild nur 82. Schon mit dieser sehr einfachen Kenngröße gelingt es, bei dieser Szene zu einem hinreichend scharf eingestellten Bild zu kommen. Allerdings ist dieser Ansatz nicht allgemeingültig und auch recht gefährlich. Schon ein einziger Glanzpunkt oder ein einziges schwarzes Pixel im Bild können die Spannweite drastisch erweitern, obwohl das Bild in großen Teilen unscharf ist. Ein unscharfes Bild kann außerdem auch ohne solche Artefakte eine große Spannweite haben. Stellen Sie sich z.B. ein Bild vor, dessen obere Hälfte weiß und dessen untere Hälfte schwarz ist. Bei Defokussierung der Abbildung wird zwar die Kante verschliffen, aber am oberen und am unteren Rand wird das Bild weiß bzw. schwarz bleiben. Die globale Spannweite bleibt dann unverändert, und über diesen Parameter wird es bei einem solchen Bild nicht gelingen, ein scharfes Bild zu erhalten.

Von der Spannweite ist es nur noch ein kleiner Schritt zu einem weiteren, häufig verwendeten Begriff, dem Kontrast. Eine sinnvolle Definition des Kontrasts ist  $c=(g_{max}-g_{min})/(g_{max}+g_{min})$ , also die Spannweite bezogen auf die Summe aus größtem und kleinstem Grauwert. Das scharf eingestellte Bild aus Abbildung 1 hat den Kontrast  $c=(223-0)/(223+0)=1$ , das unscharf eingestellte Bild den Kontrast  $c=(190-108)/(190+108)=0,275$ . Auch mit diesem Gütemaß kann man die optimale ObjektivEinstellung für die Szene aus Abbildung 1 finden.

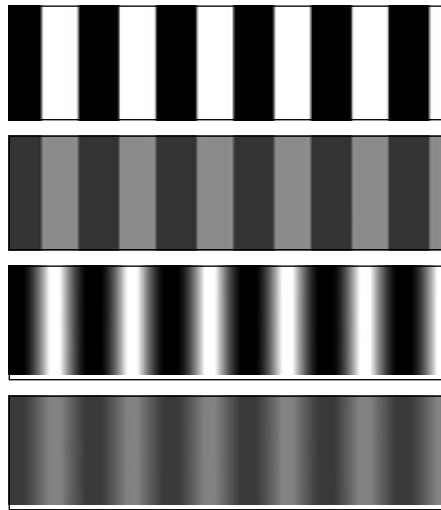


Abb. 2: Scharfe und unscharfe Strichraster, jeweils mit hohem und niedrigem Kontrast

### Autofokus mit Kontrast

Den Kontrast kann man auch als Modulationstiefe auffassen. Eine Struktur mit hohem Kontrast, z.B. ein Linienraster, ist sehr tief durchmoduliert, im Extremfall bis herunter zum Grauwert 0. Dann nimmt der Kontrast den Maximalwert 1 an, und zwar unabhängig von der absoluten Höhe der Modulation und unabhängig davon, ob die Kanten der Struktur steil oder weniger steil sind. Allgemein kann es unscharfe Bilder mit hohem Kontrast geben ebenso wie scharfe Bilder mit geringem Kontrast, scharfe Bilder mit hohem Kontrast und unscharfe Bilder mit geringem Kontrast. Abbildung 2 zeigt dies anhand eines einfachen Beispiels. Der globale Kontrast oder auch die globale Spannweite, sogar lokaler Kontrast und lokale Spannweite, berechnet in kleinen Bildausschnitten, sind folglich nicht bei allen Szenen geeignete Gütemaße für die Optimierung der Bildschärfe. Bei einer gleichmäßig ausgeleuchteten Szene mit guter Aussteuerung der Grauwertskala,

also einem mittleren Grauwert von etwa 130 und ohne Überlauf und Unterlauf, und bei genügend feinen Strukturen im Bild sind Kontrastmaße jedoch gut geeignet für die Ansteuerung einer ObjektivEinstellung mit scharfer Abbildung der Szene, wie das Beispiel in Abbildung 1 demonstriert. Auf dieser Basis ist grundsätzlich auch eine automatische „Scharfeinstellung“ (Autofokus) realisierbar, sofern der Objektivauszug motorisch verändert werden kann und der gewählte Parameter bei optimaler Einstellung ein Minimum oder ein Maximum annimmt. Eine Autofokusroutine, die ein Gütemaß aus dem aktuellen Bild berechnet, liefert allerdings im Allgemeinen nicht unmittelbar einen Zielwert für den Objektivauszug, sondern wird den optimalen Wert durch einen Regelvorgang finden müssen, der in einem Extremum des Güteparameters endet. Die Programmierung einer guten Regelung für diesen Zweck ist nicht einfach, und es müssen immer mehrere Bilder aufgenommen und ausgewertet werden, bis der optimale Zustand erreicht ist. Bei Anwendungen mit Prüfräten von mehreren Teilen pro Sekunde ist dies eine wesentliche Erschwernis. Hinzu kommt, dass das Objektiv zwischen je zwei Bildaufnahmen der Regelschleife mechanisch verfahren werden muss. Die Anforderungen an die Feinwerktechnik für ein Objektiv, das eine solche Belastung dauerhaft aushält, sind erheblich. Eine „Autofokusfunktion“ für jedes einzelne Bild wie in der digitalen Photographie wird man in der industriellen Bildverarbeitung schon aus diesen Gründen nur selten finden. Sie wird jedoch auch nicht oft benötigt, denn für viele Verfahren der industriellen Bildverarbeitung ist ein hoher Kontrast bzw. eine ausreichende Spannweite der Graustufen zwischen Vordergrund und Hintergrund und vor allem eine gleichmäßige Ausleuchtung wichtiger als ein scharfes Bild. Als Hilfsmittel bei der Einrichtung einer Prüfstation wäre ein „Schärfemonitor“ aber gewiss eine nützliche zusätzliche Funktionalität.

### Fünf-Megapixel-Optiken mit P-Iris Steuerung



Kowa erweitert sein Standardprogramm an hochauflösenden IR-korrigierten 2/3" Optiken. LM16JC5MM-IR ist für den Einsatz mit 2/3" 5MP Sensoren konzipiert. Die P-Iris Blendensteuerung erfolgt über einen Schrittmotor wodurch präzise Einstellungen des Blendenwerts realisiert werden. Das IR-korrigierte Design der Optik liefert ein scharfes und kontrastreiches Bild für den Wellenlängenbereich zwischen 400nm bis 900nm ohne Fokusverschiebung. Zusätzlich zur 10MP Serie ist ab sofort eine neue 2/3" C-mount Serie für 5MP Sensoren mit 3.5µm Pixelgröße erhältlich. Die neue JC5M2 Series um-

fasst 4 verschiedene Brennweiten zwischen 12 bis 35mm. In allen Modellen wird spezielles XD-Gals mit besonderen Dispersionseigenschaften sowie asphärische Linsen verarbeitet. Das spezielle optische Design und durch den Einsatz präziser Feinmechanik liefern die Optiken hochauflösende, kontrastreiche Bilder von der Naheinstellung bis ins Unendliche. Alle neuen Optiken für höchste Anforderung an Vibration-, Rüttel- und Temperaturbeständigkeit im industriellen Umfeld ausgelegt.

[www.kowa.eu](http://www.kowa.eu)

### Kompakte 4 MPixel USB 3.0 Kamera

Ximea liefert seine USB 3.0 Kameras (xiQ-Serie) nun auch mit 4 MPixel Auflösung auf Basis des CMOSIS CMV4000 und rundet damit das Lieferspektrum nach oben ab. Mit nur 26 mm Kantenlänge, geringster Leistungsaufnahme, dem Aufbau mit nur einer Platine, Treibern für Windows und Linux sowie der Unterstützung von über 30 der gängigsten Bildverarbeitungslibraries zeichnen sich die Kameras aus und empfehlen sich auch für den OEM-Einsatz. Das neue Programm xiCOP ermöglicht die Analyse der USB 3.0 Kameras sowie der Infrastruktur und unterstützt Softwareupdates und die Optimierung der USB 3.0-Kommunikation. Neben den USB3.0 Kameras umfasst das Lieferspektrum von Ximea u.a. auch PC-Kameras (Currera-Serie) und Mikro-USB 2.0 Kameras mit nur 15 x 15 x 9 mm Größe bei 5 MPixeln Auflösung.

[www.ximea.com](http://www.ximea.com)



### Multispektralkamera mit hoher Auflösung und Sensitivität



Mit 2-CCD-Multispektralkamera AD-130GE hat JAI bekannt, ein Kameramodell mit deutlich höherer Auflösung und Lichtempfindlichkeit als die Vorgängermodelle auf den Markt gebracht. Die Kamera verfügt über ein dichroitisches Prisma für die simultane Erfassung von sowohl Farbe als auch Nah-Infrarot (NIR)- Digitalvideo auf einem einzigen optischen Pfad, für effiziente multispektrale Prüfung und Analyse. Sie bietet eine

Auflösung von 1,3 Megapixeln (1.296 x 966 Pixel). Mit Hilfe zweier präzise ausgerichteter 1/3" Sony ICX447 CCDs bietet die Kamera eine der AD-080-Serie vergleichbare Erfassungsrate (31 fps), bei gleichzeitig erhöhter Sensitivität, insbesondere im NIR-Bereich des Spektrums. Für das sichtbare Licht liegt die Sensitivität ca. 20% über derjenigen der AD-080 Modelle; die NIR-Sensitivität ist mehr als doppelt so hoch, wie bei den Vorgängermodellen. Dies ermöglicht unter typischen Einsatzbedingungen eine höhere Bildqualität, bei gleichzeitig verminderten Beleuchtungsanforderungen für die Erzeugung angemessener Signalrauschabstände bei Inspektionsaufgaben

[www.jai.com](http://www.jai.com)

### GigE-Flächenkameras mit CMOS-Sensoren



Der Kameraspzialist Basler startet mit der Serienproduktion seiner ace GigE-Kameramodelle mit den leistungsstarken CMOS-Sensoren CMV2000 und CMV4000 von Cmosis. Kurz darauf werden die beiden Monochrom-Modelle auch als NIR (Nah-Infrarot)-Varianten verfügbar sein. Diese ace GigE-Kameras gibt es mit 2 MP Auflösung und 50 Bildern pro Sekunde (acA2000-50) und mit 4 MP und 25 Bildern pro Sekunde (acA2040-25), jeweils in Monochrom und Farbe. Sie zeichnen sich durch niedriges Rauschen und eine sehr hohe Empfindlichkeit aus. Sie eignen sich beispielsweise für den Einsatz in der Fabrikautomation, im Bereich Halbleiter- und Elektronikfertigung, für Verkehrsanwendungen, Solarbranche und Medizinsektor. Die

beiden Monochrom-Modelle der ace Cmosis GigE Kameras gibt es auch als NIR-optimierte Varianten. Im Vergleich zu den bisher verfügbaren monochromen Kameramodellen, weisen die neuen Modelle bei einer Wellenlänge von 900 nm eine doppelt so hohe Empfindlichkeit auf.

[www.baslerweb.com](http://www.baslerweb.com)

[www.inspect-online.com](http://www.inspect-online.com)

# Linux-basierter Intelligenter Analog Videokonverter



OEM-Version  
erhältlich

Verwandelt jede PAL/NTSC Kamera in  
... eine intelligente Kamera  
... eine IP/Ethernet-Kamera  
... eine digitale Kamera mit HDMI-Ausgang  
... eine Standbild- oder Videokameraeinheit

- Standard Debian Linux
- 300 MHz ARM9 Prozessor
- 600 MHz DSP mit 4800 MIPs
- 256 MB RAM
- 512 MB Flash-Speicher

- Fast Ethernet
- US-Host und RS232
- General purpose I/Os
- Analog-/HDMI-Ausgang

- H.264, MPEG4, JPG Codecs
- GStreamer-Unterstützung

Mehr Informationen unter:



VRmagic GmbH  
Augustaanlage 32  
68165 Mannheim  
Germany

Phone +49 621 400 416 - 20  
Fax +49 621 400 416 - 99

[www.vrmagic-imaging.com](http://www.vrmagic-imaging.com)  
[info.imaging@vrmagic.com](mailto:info.imaging@vrmagic.com)



### USB-Singlesensor-Kameras mit hochauflösenden Sensoren

VRmagic vergrößert sein Angebot an USB-Singlesensor-Kameras. Drei neue Modelle mit hochauflösenden Sensoren sind ab sofort Teil der Produktserie. Die CMOS-Sensoren CMV2000 und CMV4000 von Cmosis sind besonders lichtempfindlich und bieten auch bei großen Helligkeitsunterschieden ausgezeichnete Bildqualität bei hohen Frameraten. Durch die verwendete Global-Shutter-Technologie sind sie sowohl für die Aufnahme von stillstehenden als auch schnell bewegten Objekten geeignet. Der 2/3 Zoll-Sensor CMV2000 in der VRmFC-22 liefert 44 Bilder pro Sekunde bei einer Auflösung von 2.048 x 1.088 Pixel. Der in der VRmFC-42 verbauten 1 Zoll-Sensor CMV4000 erzielt bis zu 24 FPS bei einer Auflösung von bis 4,2 Megapixel. Der Sony ICX445 CCD-Sensor der VRmC-14 kann bis zu 22 Bilder pro Sekunde bei einer Auflösung von 1.296 x 966 aufnehmen. Der 1/3 Zoll-Sensor zeichnet sich durch eine erhöhte Lichtempfindlichkeit im infraroten Bereich aus. Daher ist er für den Einsatz bei Vermessungs- und Positionierungsaufgaben ideal.

[www.vrmagic.com](http://www.vrmagic.com)

### Kleine Kamera jetzt mit GigE-Schnittstelle



Die neue Board-Level Industriekamera uEye LE von IDS Imaging Development Systems GmbH misst 45x45 mm. Aufgrund der Gigabit Ethernet-Schnittstelle sind Bildraten von 50 Bildern pro Sekunde bei voller 1,3 Megapixel

Auflösung möglich. Dank einer maximalen Kabellänge von 100 Metern ist sie flexibel einsetzbar. Digitale Ein- und Ausgänge für Trigger, Blitz, Pulsweiten-Modulation, I2C-Bus, sowie 2 General Purpose I/Os runden die, besonders für OEM-Kunden spezifizierte, Kamera ab. Die Stromversorgung (12-24 V) erfolgt über ein externes Netzteil. Die GigE uEye LE Kamera ist mit oder ohne M12 Linsenhalter von Lensation erhältlich. Die GigE uEye LE ist mit leistungsstarken CMOS-Sensoren von e2v ausgestattet. Der 1,3 Megapixel Sensor ist als Mono-, Farb- sowie, ganz neu als NIR-Version, mit extrem hoher Lichtempfindlichkeit besonders im nahen Infrarotbereich erhältlich.

[www.ids-imaging.de](http://www.ids-imaging.de)

### Kamerareihe um Farb-Modelle erweitert



Teledyne Dalsa hat seine Kamerareihe Genie TS um drei neue Farb-Modelle erweitert. Die Genie TS setzt die neueste CMOS Imaging-Sensor-Technologie ein. Zu den Farbkameras der Baureihe Genie TS gehören 5M-, 8M- und 12M-Modelle mit Bildraten bis zu 29 fps. Alle Funktionen sind mit den Softwareprogrammen von Teledyne Dalsa oder mit GigE Vision-konformer Software anderer Hersteller zu nutzen.

Wie alle Genie-Kameras ist die Baureihe Genie TS GigE Vision-konform auf der Basis des AIA (Automated Imaging Association) GigE Vision-Standards. Die neue Genie-Kamerabaureihe nutzt Gigabit-Ethernet und sendet Daten über Standard-CAT-5e- und CAT-6-Kabel bis zu 100 m weit. Dadurch werden größere Abstände zwischen Kamera und Inspektionssystem möglich.

[www.teledynedalsa.com](http://www.teledynedalsa.com)

### GigE-Vision Digitalkamera-Serie erweitert



Point Grey hat ein neues Modell seiner Flea3-GigE-Vision Digitalkamera-Serie angekündigt. Das neueste Flea3 Modell zeichnet sich durch seine VGA Auflösung bei 120 FPS aus. Die FL3-GE-03S1C/M basiert auf den Sony ICX618 CCD Sensor mit EXview Global

Shutter zur Maximierung der Quanteneffizienz sowie naher IR-Empfindlichkeit. Die Kameras schaffen 120 FPF bei voller 648 x 488 Auflösung und sogar darüber hinaus bei kleinerem „Regions of Interest“ (ROI). Sie eignen sich für Anwendungen in den Bereichen Fabrikautomation, Elektronikfertigung sowie Lebensmittel- und Verpackunginspektion.

[www.ptgrey.com](http://www.ptgrey.com)

### Vielseitige ultrakompakte Smart Kameras



Die äußerst kompakten intelligenten Kameras der Serie VC nano von Vision Components sind als Platinenkameras oder wahlweise auch im IP68-Schutzgehäuse erhältlich. Speziell in Kombination mit den Software-Tools von Vision Components bieten sie Anwendern kosteneffiziente, leistungsfähige Bildverarbeitungs-Komplettpakete, die sich leicht in Maschinen und Anlagen integrieren lassen. Zur Auswahl stehen mehrere Programme für verschiedenste

Anwendungsbereiche: der VC Barcode Reader für die schnelle und zuverlässige Erkennung von 1D-Codes, der VC Smart Reader für die Erfassung von Data-Matrix-Codes und der VC Smart Finder zur konturbasierten Objekterkennung. Dank einer beidseitig bestückten Platine erreichen VC nano-Kameras Abmessungen von lediglich 40 x 65 mm bzw. 80 x 45 x 20 mm. Einige Modelle sind mit einem abgesetzten Kamerakopf ausgestattet, der nur 22 x 22 x 19,5 mm misst und mit einem bis zu mehrere Meter langen Anschlusskabel viel Flexibilität bei der Installation bietet. Je nach Modell sind sie mit 300-, 400- oder 700-MHz-DSPs mit einer maximalen Rechenleistung von 5.600 MIPS ausgestattet. Die Bildaufnahme erfolgt über einen CMOS-Sensor mit Global Shutter und einer maximalen Auflösung von 5 Megapixeln. Zur Programm- und Bildspeicherung dienen bis zu 128 MB DDRAM und bis zu 32 MB Flash. Eine frei programmierbare 100 Mbit-Ethernetschnittstelle ermöglicht eine Livebild-Ausgabe am PC. Zur Einbindung in vorhandene Automatisierungsumgebungen stehen zudem zwei digitale Eingänge und vier Ausgänge zur Verfügung. In Kürze können die Smart Kameras außerdem per Power over Ethernet mit Strom versorgt werden.

[www.vision-components.com](http://www.vision-components.com)

### Neues Produktportfolio für 3D Machine Vision

Zur Automatica wird Isra mit einem neuen Produktportfolio für 3D Machine Vision starten. Die neue Ready-to-Use-Produktlinie ermöglicht es Anwendern, diese deutlich einfacher in ihre Anlagen und Prozesse zu installieren und zu integrieren. „Plug & Automate“ nennt Isra das neue Konzept: In die Anlage integriert wird ohne Programmierung, Geräte werden einfach eingesteckt und sind dann funktionsfähig. Die Sensoren verfügen bereits über alle wichtigen Kommunikationsschnittstellen zur Anbindung an Roboter und übergeordnete Automatisierungssysteme, zum Beispiel GigE und Profinet. Die Software ist geschrieben im Metro Style, wie für Windows 8 definiert, und basiert damit auf Software-Technologie der neuen Generation. Das Portfolio enthält Lösungen für die 3D Roboterführung, für die 3D In-line Messtechnik sowie die 3D Formmessung. Das smarte Mensch-Maschine-Interface setzt auf Touch-Bedienung.

[www.isravision.com](http://www.isravision.com)



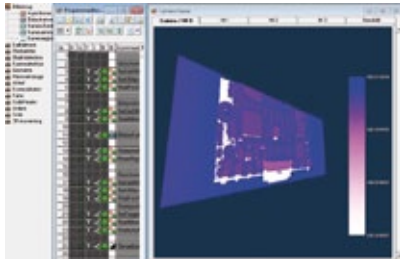
### HD-Kamera- Modul mit 30-fachen Zoom



Sony Europe's Image Sensing Solutions Division kündigt ein neues Kameramodul an und erweitert damit sein Angebot an OEM-Digital-Videokameras. Das neue Modul vereint Sonys Exmor CMOS-Sensor mit einem 30-fach optischen Zoom, was eine hohe Bildqualität für Anwendungen in den Bereichen Transportwesen, Sicherheit und bei schlechten Sichtverhältnissen ermöglicht. Der 1/2.8-Typ-Farbsensor der FCB-EH6500 bietet echte HD-Bilder (1080p) bei 30fps. Im Vergleich zum Basismodell, der FCB-EH6300, erweitert die neue Kamera den Betrachtungswinkel um über 7% (59,5° bei 1.080p; 40,6° bei 720p), erhöht die Brennweite um 50% (f = 4,3 bis 129 mm) und hält dennoch eine mindestens erforderliche Lichtstärke von nur 0,5 lx\* bei. Das Kameramodul ist eines der effizientesten am Markt und benötigt nur 3,4 W\* Leistung (4,7 W bei aktivem Motor). Der Betriebstemperaturbereich erstreckt sich von -5 bis +60 °C. Die Kamera bietet eine Reihe von Funktionen und Bildeffekte, z.B. einen 12-fach Digitalzoom (360-fach insgesamt), Privacy Zone Masking (Ausblendung bestimmter Bereiche), Bewegungserkennung, Farbverbesserung, E-Flip, Spiegelbild, Nega-Art und zahlreiche Weißabgleich-Modelle.

[www.sonybiz.net/vision](http://www.sonybiz.net/vision)

### Software mit 3D-Befehlssatz



Die EyeVision Software von EVT verfügt jetzt über einen neuen 3D-Befehlssatz. In Kombination mit z.B. dem EyeScan LT 3D Sensor können jetzt Anwendungen wie Bin-Picking, 3D-Matching und Roboterführung umgesetzt werden. Mit dem neuen EyeVision 2.6 Release stellt EVT

dem Anwender Bildverarbeitungsbeefehle zur Verfügung welche in Kombination mit einem der implementierten 3D Sensoren die dritte Dimension erschließen. Mit dem neuen Softwarerelease – EyeVision 2.6 – wurden 3D Befehle entwickelt, um klassische Aufgabenstellungen vieler Anwender zu lösen, wie z.B. das Greifen von ungeordneten Teilen von Transportbändern, Paletten oder aus Kisten. Für die einfache 3D-Matching Befehle die dem EyeVision Standard folgen, stehen Drag-and-Drop-Icons zur Verfügung und ermöglichen dem Anwender, auch solche Lösungen umzusetzen.

[www.evt-web.com](http://www.evt-web.com)

### 3D-Dokumentation zur Sanierung historischer Gebäude



Die Kirche San Michele Arcangelo aus dem Jahr 1735 liegt im Zentrum von Borgo di Terzo nahe Bergamo in Italien. An der Vorderseite der Fassade zeichneten sich deutliche Mauerwerksrisse ab, die auf die auf schwere Schäden des Gebäudes hinwiesen. Im Zuge der 3D-Dokumentation mit dem Faro Laser Scanner wurde die Ursache für die beobachteten Schäden deutlich: Als Folge einer früheren Restaurierung kam es zu einer Setzung der hölzernen Dachkonstruktion. Davon

sind sowohl die Anker eines Binders als auch der auf dem Gewölbe aufliegende Gewölberücken betroffen. Das führte zu einer starken Durchbiegung und zu schweren Schäden des darunter liegenden Gewölbemauerwerks. Mit der FEM-Analyse wurde der aktuelle Belastungsverlauf im Gewölbe bewertet und der Verlauf nach der geplanten Restaurierung überprüft. Sie ergab, dass sowohl die Auflager der Bogenpfeiler als auch die korrespondierenden Anker saniert werden müssen. Zudem ist der mittlere Bogen mit einem Stahlnetz und Kalkmörtel zu verstärken. Die Verformung des Gewölbes unter Last zeigt außerdem, dass ein Bindeglied zum Gewölberücken nachgegeben hat, was den Austausch dieses Ankers nahelegte. Nachdem diese Verstärkungsmaßnahmen in das Bestands-FEM-Modell eingefügt worden waren, zeigte sich eine allgemeine Verbesserung der gesamten statischen Struktur.

[www.faro.com/germany](http://www.faro.com/germany)

[www.inspect-online.com](http://www.inspect-online.com)

### Neue GigE-Kamerafamilie im Miniaturgehäuse

The Imaging Source präsentiert eine neue GigE-Kamerafamilie mit einem Footprint von 29 x 29 mm, der zum Footprint vieler Analogkameras kompatibel ist. Ein Wechsel von analogen Kameras zu digitalen Kameras kann somit problemlos erfolgen. Die Kameras verfügen über ein robustes Industriegehäuse mit C/CS-Mount Objektiv-Anschluss, Triggereingang und digitalen I/Os. Die Farb-, Monochrom- und Bayer-Modelle sind in VGA, SXGA und 2MP-Auflösungen verfügbar und somit prädestiniert für Machine Vision-Anwendungen in der Automatisierungstechnik, Verkehrsüberwachung, Qualitätskontrolle, Medizin, Logistik, Mikroskopie und Sicherheitstechnik. Die Kameras mit nur wenigen Codezeilen in neue oder bestehende Applikationen integriert. Treiber für z.B. LabView, Halcon, DirectX, Twain, WDM werden mitgeliefert. Sämtliche Kamerafunktionen werden softwaremäßig konfiguriert. Außerdem stehen leistungsfähige Automodi zur Verfügung, die eine optimale Bildqualität bei wechselnden Lichtverhältnissen garantieren. Im Lieferumfang enthalten sind: Kamera, Treiber für Windows XP, Vista und 7, das SDK IC Imaging Control 3.2 (.NET und C++ Bibliothek) und IC Capture.

[www.theimagingsource.com](http://www.theimagingsource.com)



### High-Speed-Kamera mit 4 Megapixel



Mikrotron stellt mit der Eo-Sens 4CXP eine kompakte, digitale Vier-Megapixel-Hochgeschwindigkeitskamera mit CoaXPress Echtzeit-Datenschnittstelle vor. Über die Schnittstelle mit hoher Datenübertragungsrate bis 2.4 Gbyte/s (bis 40 m) erfolgen mit nur einem Kabel sowohl Datentransfer, Kommunikation und Kontrolle zum PC als auch die Stromversorgung. Das sind entscheidende Vorteile für die einfache Systemintegration und Bedienung. Die Kamera erfasst 500 Bilder pro Sekunde mit einer Auflösung von 4 Megapixeln (2.336 x 1.728 Pixel). In Full HD-Auflösung von 1.920 x 1.080 Pixel liefert sie 900 Bilder pro Sekunde. Damit kann die Kamera selbst kleinste Details der Bauteile präzise erfassen. Durch die stufenlos variable Gestaltung des Bildfeldes in Lage und Auflösung (RoI) wird eine hohe Anwendungsflexibilität erzielt und Aufnahmefrequenzen von über 100.000 Bildern/Sekunde gewährleistet. Hinzu kommt der Vorteil hoher Lichtempfindlichkeit von 3200 ASA monochrom und 2400 ASA in RGB mit Bayer-Filter.

[www.mikrotron.de](http://www.mikrotron.de)



## Der neue pictor®!

### VISION&CONTROL

SYSTEMS • LIGHTING • OPTICS

- Intelligente Kameras
- Mehrkamerasysteme
- Software
- Optik & Beleuchtung



QUALITÄT  
MADE IN  
GERMANY

[www.vision-control.com](http://www.vision-control.com)

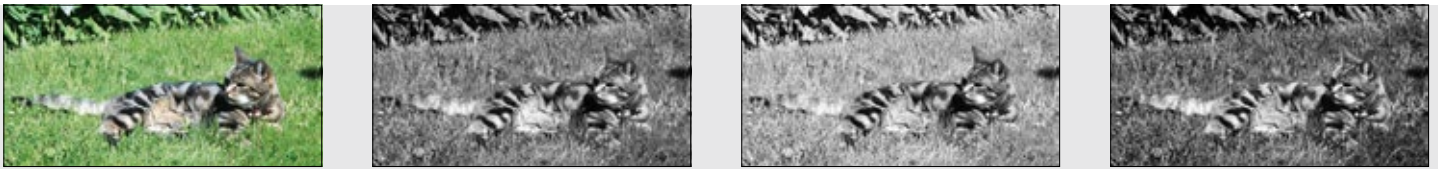


Abb. 1: Ein RGB-Bild mit den Farbausügen für R, G und B und ein Graustufenbild derselben Szene. Die Binarisierung des Graustufenbildes ergibt keine brauchbare Binarisierung des Differenzbildes führt zum gewünschten Resultat.

# Farbbildverarbeitung

Nachts sind alle Katzen grau

In der industriellen Bildverarbeitung werden Farbbilder nur verwendet, wenn dies unbedingt erforderlich ist. Gelegentlich können Szenen jedoch nur mit Hilfe der Farbinformation segmentiert werden. Weitaus schwieriger ist die Überprüfung der Farbeigenschaften von Objekten auf der Basis eines Farbbildes. Dieses Problem fällt eher in den Bereich der optischen Messtechnik (Radiometrie) als in das Gebiet der typischen Bildverarbeitung.

Sprachlich beschreiben wir den menschlichen Sinneseindruck „Farbe“ mit Hilfe von drei Attributen: Farbton, Helligkeit und Sättigung. Der Farbton ist die Farbart (rot, gelb, violett usw.). Die Helligkeit entspricht der Graustufe bei einer Aufnahme mit einer Monochromkamera. Unter der Sättigung kann man die Bunttheit einer Farbe verstehen. Reine Spektralfarben, etwa die rote Farbe eines He-Ne-Lasers, sind vollständig gesättigt, unbunte Farben wie Grau, Weiß und Schwarz sind vollständig ungesättigt. Es gibt sowohl helle als auch dunkle, tief gesättigte und weniger gesättigte Farben mit unterschiedlichen Farbtönen, z.B. ein dunkles, tiefgesättigtes Rot oder ein helles, wenig gesättigtes Blau (Pastellblau). Alle drei Farbattribute sind also unabhängig voneinander, und man kann eine Farbe als Punkt in einem dreidimensionalen Raum auffassen.

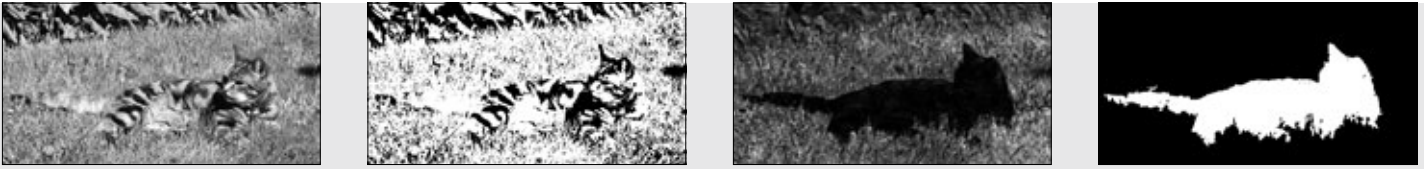
Farbkameras erfassen Farben über ihre Anteile an drei Primärfarben, z.B. Rot, Grün und Blau (RGB). Weiß ist dann durch gleiche große Farbanteile gekennzeichnet, die jeweils voll ausgesteuert sind ( $R=G=B=1$ ). Die RGB-Farbanteile kann man auch in drei andere Parameter umrechnen, denn derselbe Punkt in einem dreidimensionalen Parameterraum kann in verschiedenen Koordinatensystemen dargestellt werden. In der Bildverarbeitung ist das sog. HSI-Farbsystem gebräuchlich, das an die menschliche Farbwahrnehmung angelehnt ist. H steht dabei für den Parameter Farbton (engl. hue), S für die Sättigung (engl. saturation) und I für die Helligkeit (engl. intensity).

Farbkameras geben normalerweise für jedes Pixel das RGB-Tripel aus, liefern also drei Bilddateien einer Szene. Abbildung 1 zeigt ein Beispiel für ein RGB-Farbbild, außerdem das korrespondierende Graustufenbild und die drei Teilbilder (die sog. Farbauszüge) für die RGB-Farbanteile. Farbbildverarbeitung bedeutet also bei gleicher Pixelauflösung die dreifache Datenmenge im Vergleich zur Graustufenbildverarbeitung, mit allen Konsequenzen für die Bandbreite der Übertragungspfade und

die Rechenzeit. Bei Analogkameras kann die Information über Farbton und Sättigung, die sog. Chroma, in das Helligkeitssignal eingebettet sein (FBAS-Signal) und muss im Framegrabber durch Demodulation zurück gewonnen werden. Das kann zu Übersprechen zwischen den Farbkanälen führen, so dass sich z.B. Teile der Rotinformation plötzlich im Blaubild wieder finden, beispielsweise an Kanten. Eine FBAS-Kamera ist daher für die meisten Fragestellungen der Bildverarbeitung nicht zu empfehlen, gewiss nicht für die Bewertung von Farben bei der Qualitätskontrolle. Die nächst bessere Variante ist die getrennte Übertragung des Helligkeits- und des Chromasignals (Y/C-Signal), noch besser die getrennte Übertragung der drei RGB-Signale.

## Farbseparierung

Auch die getrennte Übertragung der RGB-Signale kann jedoch nicht über einige gravierende Probleme bei der Farbsignalgewinnung in der Kamera hinwegtäuschen. Die typische Farbkamera enthält einen Detektorchip mit einer darüber liegenden Farbfiltermatrix. Weit verbreitet ist das sog. Bayer-Filter-Muster [1]. Die Elementarzelle dieser Filtermatrix besteht aus einem Quadrat mit vier Pixeln, von denen zwei mit einem Grünfilter und je eines mit einem Rot- und Blaufilter belegt sind (s. Abb. 2). Eine solche sog. 1-Chip-Farbkamera, für die im Datenblatt z.B. eine Pixelauflösung von 1.000 x 1.000 Pixeln angegeben ist, liefert in der Regel drei komplette Bilder mit je 1.000 x 1.000 Pixeln in den Farbbereichen R, G und B – der Detektorchip enthält jedoch insgesamt lediglich 1.000 x 1.000 Pixel! In der Kamera werden für jedes Pixel die beiden jeweils fehlenden Farbanteile aus den Daten der umgebenden Pixel berechnet und zusammen mit dem tatsächlich aufgenommenen Farbanteil auf den Signalausgang des Pixels gelegt. Was dabei in einem extremen Fall passieren kann, z.B. an Schwarz-Weiß-Übergängen oder bei Reflexen, zeigen die Bildausschnitte in Abbildung 2. Hier wird deutlich, dass im Farbsignal durch die Interpolation Farben vorkommen können, die in der Realität gewiss nicht vorhanden sind. Für Bildverarbeitungsanwendungen gibt es seit einigen Jahren 1-Chip-Farbkameras, die den unmittelbaren Zugriff auf die Rohbilddaten zulassen, d.h. es wird nur der Farbwert ausgegeben, der vom jeweiligen Pixel auch tatsächlich erfasst wird; die übrigen Farbwerte werden ausgespart. Auf diese Weise wird für den Anwender die tatsächliche Ortsauflösung der Farbsignale erkennbar, und er kann z.B. verlässlich den mittleren Farbwert einer hinreichend großen Fläche bestimmen, ohne dass die Quellsignale durch Interpolationen verfälscht werden. Allerdings muss dazu die Farbmaskierung des Detektorchips bekannt sein und bei der Programmierung berücksichtigt werden. Alternativ kann eine 3-Chip-Kamera verwendet werden. Sol-



bare Segmentierung. Das Differenzbild zwischen dem grünen und dem blauen Farbauszug liefert im Objekt zusammenhängend sehr niedrige Werte. Die Binari-

che Kameras enthalten drei getrennte Detektorchips und einen Farbteiler, der die drei Farbanteile optisch aus dem einfallenden Strahlenbündel abtrennt und jeweils auf separate Detektoren abbildet. Die drei Detektorchips müssen sehr sorgfältig aufeinander justiert werden. Dies stellt sehr hohe Anforderungen an die Fertigungsgenauigkeit. Hinzu kommt, dass die Farbteiler mit Aufdampfschichten realisiert werden, deren spektrale Reflektivität systembedingt vom Einfallswinkel der Strahlung abhängt. Daher sind die Bilder aus 3-Chip-Farbkameras oft von großflächigen Farbverläufen durchzogen, die für eine Farbbeurteilung erst herausgerechnet werden müssen. Ein beinahe banales Problem bei 3-Chip-Farbkameras ist die unangenehme Tatsache, dass bei manchen Standard-Objektiven die Linsen so weit in die Kamera hineinragen können, dass sie bei einigen Kameramodellen auf den Farbteilerprismen aufsetzen.

Eine weitere Möglichkeit zur Gewinnung von Farbbildern mit hoher Ortsauflösung bietet das Prinzip der Foveon-Detektoren. Dabei wird ausgenutzt, dass die Eindringtiefe des Lichts in den Silizium-Detektorchip von der Wellenlänge abhängt. Blaues Licht wird in den oberen Schichten, grünes Licht in den mittleren Schichten und rotes Licht in den unteren Schichten absorbiert. Wenn es gelingt, die ausgelösten Ladungen unter jedem Detektorpixel in separaten Speichertaschen in diesen drei Tiefenbereichen zu speichern und gezielt auszulesen, ist ein Detektor mit hoher Pixelauflösung und drei Farbsignalen pro Pixel realisiert. Kameras mit solchen Detektoren sind verfügbar.

Farbauszüge entstehen auch, wenn eine Monochromkamera mit breiter spektraler Empfindlichkeit mit einem Filterrad vor dem Objektiv oder vor dem Detektor ausgerüstet wird. Die Bilder in den verschiedenen Spektralbändern werden dann sukzessive aufgenommen, und zwischen zwei Bildaufnahmen wird das Filterrad weitergedreht. Auf diese Weise kann die volle Pixelauflösung des Detektors für jeden Farbauszug genutzt werden. Für viele industrielle Anwendungen kommt diese Methode

nicht in Frage, etwa bei schnell bewegten Objekten. Für eine Reihe von Prüfaufgaben sind auf diese Weise jedoch sehr präzise orts aufgelöste radiometrische bzw. farbmetrische Messungen realisiert worden. Im Unterschied zu den 1-Chip-Kameras, bei denen innerhalb einer Charge mit erheblichen Streuungen der spektralen Empfindlichkeit gerechnet werden muss, hat der Anwender bei den Filterrad-Kameras die volle Kontrolle über die spektralen Eigenschaften der Farbauszüge [2].

Eine selten angewandte, aber gut funktionierende Methode zur Farbbildgewinnung ergibt sich auch aus drei Aufnahmen einer Szene mit einer gewöhnlichen Graustufenkamera unter sukzessiver Beleuchtung mit blauem, grünem und rotem Licht. Bei geschickter Auswahl der Lichtquellen entstehen so ebenfalls Farbauszüge für die RGB-Farbanteile, analog zu den Bildern, die bei einer Filterradkamera mit einem RGB-Filterrad entstehen. Wir haben in unseren Labors eine solche Beleuchtungseinheit mit LEDs aufgebaut und damit gute Ergebnisse erzielt. Die Farbkanäle können über die LED-Ansteuerung sehr viel schneller geschaltet werden als mit einem mechanischen Filterrad, so dass jeweils drei unmittelbar aufeinander folgende Kameraframes genutzt werden können.

## Beleuchtung

Der Farbeindruck, der von einer diffus remittierenden Fläche entsteht, ist von der Spektralverteilung der Beleuchtung abhängig. Farben können überhaupt nur dann sinnvoll in einem Farbbild erfasst werden, wenn die Beleuchtung alle Wellenlängen des sichtbaren Spektrums mit signifikanten Intensitäten enthält. Eine Gleichverteilung der Intensitäten über das gesamte Spektrum ist jedoch nicht zwingend erforderlich. Bei einem Weißabgleich an der Kamera wird in der aktuell verwendeten Beleuchtung eine weiße Fläche ins Sichtfeld der Kamera gebracht, und die RGB-Signale werden auf den gleichen

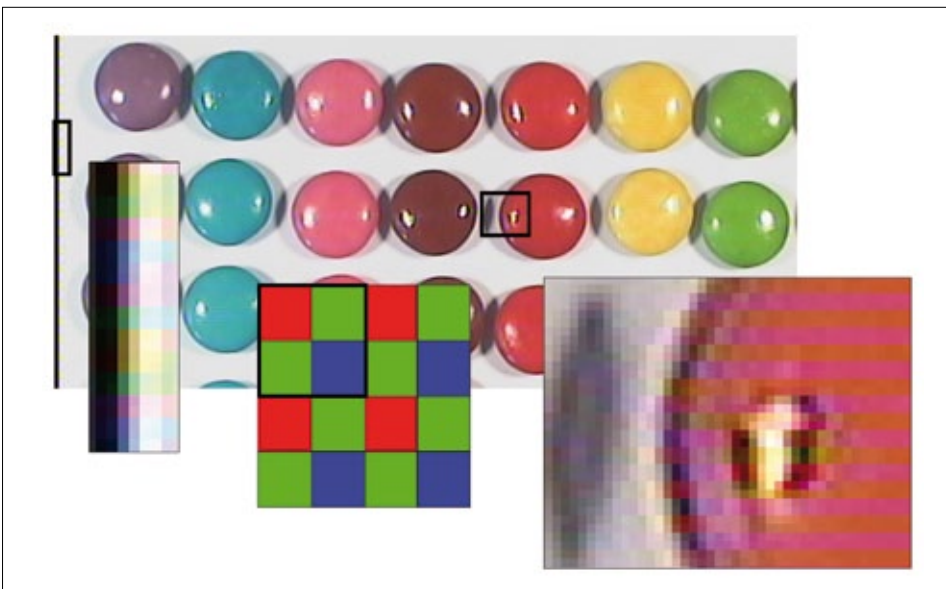


Abb. 2: Zwei Bildausschnitte mit Artefakten bei einem Bild aus einer 1-Chip-Farbkamera, in der Mitte ist das Bayer-Mosaikfilter schematisch dargestellt.



Pegel eingeregelt. Sobald sich die spektrale Zusammensetzung der Beleuchtung ändert, muss dieser Weißabgleich wiederholt werden. In der Farbbildverarbeitung müssen also nicht nur die Intensitätsschwankungen der Beleuchtung kompensiert werden, sondern auch die spektrale Zusammensetzung der Beleuchtung muss kontrolliert und möglichst konstant gehalten werden. Auch wenn der Weißabgleich unterschiedli-

che Intensitäten in den Spektralbändern der Beleuchtung und die spektrale Empfindlichkeitskurve des Detektors ausregelt, dürfen keinesfalls ganze Spektralbereiche fehlen. Eine weiße Fläche wird in einem technischen System nie weiß erscheinen, wenn die Beleuchtung keinen Rotanteil enthält. Ebenso kann eine grüne Farbe nicht zu einem auswertbaren Bildsignal führen, wenn das grüne Spektralband in der Beleuchtung

fehlt, sondern wird schwarz erscheinen. Für die Beleuchtung müssen folglich spektral breitbandige Lichtquellen eingesetzt werden, z.B. Glühlampen oder weiße LEDs mit möglichst geringer Struktur in der Spektralverteilung.

### Segmentierung mittels Farbe

Abbildung 1 zeigt ein Beispiel für eine Szene, bei der die Segmentierung im Grau-

stufenbild nicht gelingt, die Auswertung der Farbinformation jedoch zum Ziel führt. Eine Binarisierung im Graustufenbild ergibt ein völlig unbrauchbares Ergebnis. Allerdings dominiert im Hintergrund der Grünanteil, so dass der Grünauszug dort sehr hell erscheint, der Blauauszug hingegen dunkel. Im Unterschied dazu ist das Fell der Katze zwar gestreift, aber nahezu unbunt, so dass in einem Objektpixel die drei Farbanteile nahezu denselben Wert annehmen. Der Absolutwert variiert in den unterschiedlichen Regionen durchaus deutlich, für jedes Pixel im Objekt ist aber annähernd  $R=G=B$ . Das Differenzbild zwischen dem grünen und dem blauen Farbauszug ergibt folglich im Hintergrund vorwiegend helle Werte, während im Objekt die dunklen Grauwerte dominieren. Zur Verdeutlichung sind in Abbildung 1 im Differenzbild die Grauwerte gespreizt worden. Eine einfache Binarisierung im Differenzbild, gekoppelt mit dem Ausschluss relativ kleiner Objekte, ergibt eine gute Trennung von Vordergrund und Hintergrund.

### Farbbewertung

Die Bewertung von Farben auf der Basis von Farbbildern ist ein schwieriges Teilgebiet der Bildverarbeitung und führt weit über den Rahmen dieses Artikels hinaus. Schon die relative Beurteilung von Farbeigenschaften, z.B. die Prüfung von Farbcodierungen, ist mit vielen Schwierigkeiten verbunden, erst recht die objektive Messung von Farbwerten im Sinne der Farbmessung [2].

### Literatur

- [1] B.F. Bayer: Color Imaging Array. US Patent No. 3971065.
- [2] U. Krüger, F. Schmidt: Ortsaufgelöste Farbmessung. GIT INSPECT Spezial, p. 8 ff., 2004.



# Holen Sie sich die Maximalpunktzahl.

Mit Industriekameras und Vision Sensoren von Baumer.



Neugierig?

[www.baumer.com/vision](http://www.baumer.com/vision)



Besuchen Sie uns vom 06. bis 08.11.2012  
auf der VISION in Stuttgart  
Halle 1 Stand F32

# Filteroperationen

## Glätten oder Falten?

Glättungs- und Kantenfilter sind sog. lineare Filter. Diese Filteroperationen gehören zweifellos zum Standardrepertoire und somit zu den Grundlagen der Bildverarbeitung. In Lehrbüchern werden lineare Filter oft sehr ausführlich behandelt, denn sie können umfassend mit einem eleganten theoretischen Instrumentarium beschrieben werden. Wir verzichten hier auf diesen Ansatz und betrachten die wesentlichen Aspekte mit deutlich geringerem mathematischem Aufwand anhand der Gruppe der Glättungsfilter.

### Filter als Nachbarschaftsoperation

Filter sind Operatoren, die jedem Bildpunkt einen neuen Grauwert zuordnen. Aus dem Quellbild entsteht durch die Anwendung eines Filters („Filterung“) ein neues, gleich großes Bild. Filter berücksichtigen auch die Umgebung eines Pixels zur Berechnung des neuen Grauwerts und werden daher oft als „Nachbarschaftsoperationen“ bezeichnet. Welche Pixel aus der Umgebung verwendet werden, hängt vom jeweiligen Filter ab. Die relevante Umgebung wird durch eine Maske spezifiziert, die mit ihrem Zentrum über den aktuell betrachteten Bildpunkt gelegt wird.

Ein Filter wird durch zwei Festlegungen definiert:

- die Filtermaske,
- die Vorschrift, mit der aus den Grauwerten innerhalb der Maske der Grauwert des aktuell betrachteten Pixels im Zielbild berechnet wird.

Die Masken werden oft auch als Strukturelemente bezeichnet. Abbildung 1 zeigt als Beispiel eine typische, sog. „3x3-Maske“. Sie besteht aus einem Quadrat mit drei Pixeln Kantenlänge.

Die Maske wird so über das Bild gelegt, dass das Pixel, dessen neuer Grauwert berechnet werden soll, im Zentrum der Maske liegt. Der neue Grauwert wird bestimmt, indem die Grauwerte der Pixel im Quellbild betrachtet werden, die innerhalb der Maske liegen. Bei linearen Filtern entsteht der neue Grauwert aus einer Linearkombination der Grauwerte innerhalb der Filtermaske. Im einfachsten Fall ist das der Mittelwert dieser Grauwerte.

Bei der Anwendung der Filteroperation auf das ganze Bild werden nacheinander alle Pixel aus dem Quellbild von links nach rechts und von oben nach unten abgetastet. Man greift sich also ein Pixel aus dem Quellbild heraus, legt die Filtermaske so auf das Quellbild, dass das aktuell betrachtete Pixel im Zentrum der Maske liegt, berechnet den neuen Grauwert aus den Grauwerten der Pixel, die innerhalb der Maske liegen, und ordnet diesen Grauwert dem aktuell betrachteten Pixel zu – aber im Ergebnisbild!

Beim nächsten Schritt wird die Filtermaske über das nächste Pixel zentriert, das beim Abtasten des Quellbildes an der Reihe

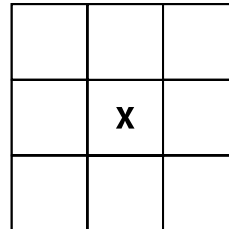


Abb. 1:  
Eine 3x3-Maske

ist. Das vorher behandelte Pixel wird dann meist noch innerhalb der Filtermaske liegen und würde die Berechnung des neuen Grauwerts beeinflussen. Bei den meisten Filtern darf man daher nicht mit den bereits modifizierten Grauwerten rechnen, sondern muss die Grauwerte aus dem ursprünglichen Bild benutzen. Es ist deshalb sinnvoll, mit zwei völlig unabhängigen, gleich großen Bildern zu arbeiten, dem Quellbild und dem Zielbild. Abbildung 2 zeigt die ersten fünf Abtastschritte für eine Filteroperation mit einer 3x3-Maske. Links ist das Quellbild mit der jeweiligen Maskenposition gezeigt, rechts das Zielbild. Die bereits bearbeiteten Pixel im Zielbild sind mit dem Grauwert ausgefüllt, der sich als Mittelwert der Grauwerte in der Filtermaske ergibt. Übrigens verwendet man mit Vorliebe symmetrische Masken mit einer ungeraden Zahl von Pixeln. Dann ist unmittelbar klar, welches das zentrale Pixel ist, dem der neue Grauwert zugeordnet wird. Masken können jedoch grundsätzlich beliebige Größen und beliebige Formen haben. Bei asymmetrischen Masken muss aber explizit verabredet werden, welchem Pixel der neue Grauwert zugeordnet werden soll.

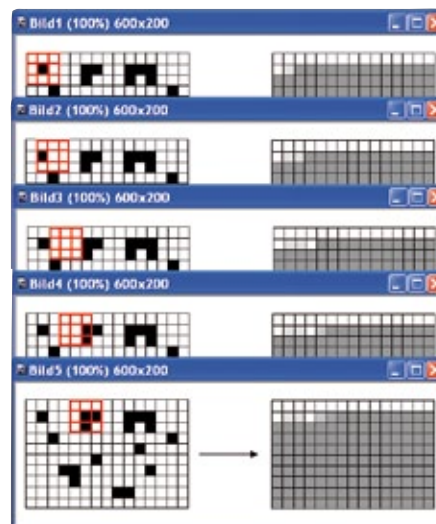


Abb. 2: Abtasten des Quellbildes (links) mit einer 3x3-Filtermaske und Abspeichern der Resultate im Ergebnisbild (rechts).

Für Pixel in der Nähe der Bildränder ragt die Maske möglicherweise über das Bild hinaus. Dort fehlen also Grauwerte, die für die Berechnung des neuen Grauwerts des Pixels im Zentrum der Maske erforderlich sind. Es gibt mehrere Möglichkeiten, mit diesem Problem sinnvoll umzugehen. Die sicherste Methode besteht darin, für die betreffenden Pixel im Ergebnisbild gar keinen Grauwert zuzuordnen. Das Ergebnisbild ist dann etwas kleiner als das Quellbild. Eine andere Vorgehensweise ergänzt das Bild an den Rändern durch einen genügend großen Rahmen, so dass die Maske immer mit Grauwerten gefüllt ist. Der Bildinhalt sollte dabei möglichst „sinnvoll“ nach außen fortgesetzt werden. Wenn z.B. der obere Bildrand um zwei Zeilen ergänzt werden muss, könnte man einfach die erste Bildzeile unverändert zweifach kopieren und oben hinzufügen. An den Ecken nimmt man den Mittelwert der benachbarten Zeilen- bzw. Spaltenpixel. Das Bild wird auf diese Weise über den Rand hinaus „periodisch ergänzt“. Andererseits könnte man den Bildinhalt auch an der

Wert zuzuordnen. Das Ergebnisbild ist dann etwas kleiner als das Quellbild. Eine andere Vorgehensweise ergänzt das Bild an den Rändern durch einen genügend großen Rahmen, so dass die Maske immer mit Grauwerten gefüllt ist. Der Bildinhalt sollte dabei möglichst „sinnvoll“ nach außen fortgesetzt werden. Wenn z.B. der obere Bildrand um zwei Zeilen ergänzt werden muss, könnte man einfach die erste Bildzeile unverändert zweifach kopieren und oben hinzufügen. An den Ecken nimmt man den Mittelwert der benachbarten Zeilen- bzw. Spaltenpixel. Das Bild wird auf diese Weise über den Rand hinaus „periodisch ergänzt“. Andererseits könnte man den Bildinhalt auch an der

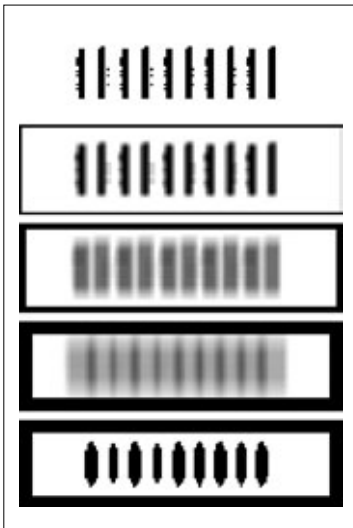


Abb. 3: Dasselbe Quellbild (oben), gefiltert jeweils mit einem Mittelwertfilter mit 3x3-, 7x7- und 13x13-Maske.

Bei der 13x13-Maske tritt Kontrastumkehr auf, deutlich erkennbar im binarisierten Bild (unten).

oberen Kante spiegeln, also die erste Zeile des Originalbildes direkt oberhalb des Bildrandes und darüber die zweite Zeile des Originalbildes ergänzen. Sowohl die Spiegelung als auch die periodische Ergänzung werden in Bildverarbeitungsprogrammen verwendet. Es hängt jedoch vom verwendeten Filtertyp ab, welche Methode besser geeignet ist.

### Glättungsfilter

Ein einfacher Ansatz zur Erkennung von Objekten in einem Bild ist die Vorstellung, dass Objekte Regionen mit konstantem Grauwert sind. Damit dieser Ansatz funktioniert, müssen sich die Objekte untereinander und vom Untergrund in Bezug auf den Grauwert genügend stark unterscheiden. Selbst unter optimalen Bedingungen sind die Grauwerte innerhalb einer Region jedoch nicht konstant. Der Mittelwert dieser Grauwerte ist deshalb zur Charakterisierung einer Region möglicherweise besser geeignet. Diese Überlegung ist ein Beispiel, bei dem ein Mittelwertfilter (oder allgemeiner ein Glättungsfilter) nützlich sein kann.

Glättungsfilter dienen hauptsächlich dazu, das statistische Rauschen in einem Bild zu unterdrücken und filigrane Störungen zu beseitigen. Abgesehen davon sollten die Auswirkungen auf das Bild möglichst gering sein. Glättungsfilter sollten daher folgende Anforderungen erfüllen:

- Die Lage des Objekts soll durch die Filterung nicht verändert werden.
- Der mittlere Grauwert soll erhalten bleiben.
- Die Glättung sollte unabhängig von der Richtung einer Struktur dieselbe Wirkung haben, sie sollte isotrop sein.
- Feinere Strukturen sollten stärker abgeschwächt werden als grobe Strukturen.

Die beiden letzten Forderungen erscheinen auf den ersten Blick selbstverständlich. Sie sind jedoch durchaus erwähnenswert, weil sie keineswegs einfach zu erfüllen sind. Die Isotropie kann wegen der Diskretisierung grundsätzlich nicht vollständig erfüllt werden, und es gibt durchaus Filter, deren Dämpfungsverhalten in Bezug auf die „Feinheit“ der Struktur (genauer: als Funktion der Ortsfrequenz) zunächst überraschend ist. Ein Beispiel eines solchen Filters wird weiter unten dargestellt.

### Mittelwertfilter

Die einfachste Idee zur Realisierung einer Glättung ist eine Mittelwertbildung über die Grauwerte in der Umgebung eines Pi-

xels. Ein Beispiel dafür ist das eindimensionale Rechteckfilter  ${}^3R_x$  mit einer 1x3-Filtermaske, d.h. die Maske umfasst drei Pixel in x-Richtung und ein Pixel in y-Richtung, besteht also aus einer Zeile und drei Spalten. Eine beliebige Darstellungsform dafür ist:

$${}^3R_x = \frac{1}{3} [111]$$

Die eckige Klammer symbolisiert die Filtermaske, die über das Quellbild geschoben wird. Das mittlere Pixel der Maske liegt jeweils über dem Bildpixel, für das gerade der neue Grauwert im Ergebnisbild berechnet werden soll. Die Zahlen innerhalb der eckigen Klammern und der Vorfaktor beschreiben die Zuordnungsvorschrift, nach der der neue Grauwert aus den Grauwerten der Umgebung berechnet wird. In diesem Beispiel sollen die Grauwerte innerhalb der Maske jeweils mit 1 multipliziert, dann aufsummiert und schließlich durch 3 geteilt werden. Die Notation mit den eckigen Klammern ist also eine Kurzschreibweise für die Zuordnungsvorschrift:

$$\bar{g}(x,y) = \frac{1}{3} [1 \cdot g(x-1,y) + 1 \cdot g(x,y) + 1 \cdot g(x+1,y)]$$

mit der dem Pixel (x,y) im Ergebnisbild der mittlere Grauwert aus einer 1x3-Umgebung des Pixels (x,y) im Quellbild zugeordnet wird. Der Faktor 1/3 führt dazu, dass bei der Filterung der Mittelwert erhalten bleibt: nach der Filterung ist der Mittelwert genauso groß wie vorher. Dieses Mittelwertfilter macht aus einer scharfen Kante parallel zur y-Achse eine verschliffene Kante. Die Lage der Kante bleibt aber erhalten. Eine Kante parallel zur x-Achse wird von dem Filter gar nicht beeinflusst. Dieses Filter ist also keineswegs isotrop, als Glättungsfilter ist es daher nicht gut geeignet – es sei denn, die selektive Glättungswirkung längs der x-Achse ist bewusst gewünscht. Entsprechend gibt es das Mittelwertfilter, das nur parallel zur y-Achse glättet. Die zweidimensionale Verallgemeinerung ist:

$${}^3R = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Dieses Mittelwertfilter mit einer 3x3-Filtermaske liefert eine wesentlich bessere Annäherung an die Isotropie als die beiden eindimensionalen Glättungsfilter. Es bewertet jedoch die Nachbarn in den Diagonalrichtungen genauso stark wie die Nachbarn in x- und y-Richtung, obwohl die Diagonalnachbarn weiter entfernt sind. Die Isotropie ist auch mit diesem Filter noch nicht optimal auf dem diskreten Bildrastrer realisiert.

Die Mittelwertfilter haben einen weiteren gravierenden Nachteil. Strukturen werden durch diese Filter zunächst immer stärker geschwächt, je feiner sie werden. Dann wird bei einer bestimmten räumlichen Periode aus einer periodischen Struktur, also einem Gittermuster, eine vollständig homogene Fläche mit einheitlichem Grauwert. Wird die Struktur dann noch feiner, wird sie nach der Filterung wieder erkennbar, unter Umständen sogar mit umgekehrtem Kontrast, wenn auch stark geschwächt gegenüber dem Originalbild. Ein Beispiel zu diesem Verhalten, der sog. Kontrastumkehr, zeigt Abbildung 3. Oben ist das Quellbild dargestellt, darunter das Ergebnis der Filterung mit einer 3x3-, einer 7x7- und schließlich einer 13x13-Filtermaske. Zur Verdeutlichung ist für die 13x13-Maske das Ergebnis der Filteroperation binarisiert worden. Bei dieser Maskengröße kehrt sich der Kontrast um, und aus den ursprünglich vorhandenen 10 werden neun Balken! Ein ideales Glättungsfilter sollte ein solches Verhalten nicht zeigen, sondern durchgän-



gig feinere Strukturen stärker dämpfen als weniger feine Strukturen. Es kann für die nachfolgenden Bildverarbeitungsstufen sehr unangenehm sein, wenn Muster mit einer bestimmten räumlichen Periode von einem Filter völlig unterdrückt werden, etwas feinere Muster dann aber wieder sichtbar sind.

### Binomialfilter, Gaußfilter

Die gewünschte Isotropie kann man erreichen, indem die Gewichtsfaktoren in der Filtermaske eine rotationssymmetrische Verteilung um das zentrale Pixel nachbilden. Das unerfreuliche Dämpfungsverhalten des Mittelwertfilters hängt damit zusammen, dass die Mittelwertbildung mit einer Gewichtsfunktion realisiert wird, die zu den Rändern hin abrupt abbricht. Trägt man die Gewichtsfaktoren graphisch über der x-y-Ebene auf, entsteht beim Mittelwertfilter ein Quader mit scharf abfallenden Seitenflächen. Ein ideales Filter müsste aber eine sanft zu den Rändern abfallende Gewichtsfunktion benutzen. Die genauere Betrachtung im Rahmen der Systemtheorie liefert das Resultat, dass eine rotationssymmetrische Gaußfunktion eine sehr gute Gewichtsfunktion für ein Glättungsfilter wäre [1], [2]. Auf der diskreten Ebene lässt sich eine Gaußfunktion niemals perfekt realisieren, sondern lediglich approximieren. Die diskrete Approximation der Gaußverteilung ist die Binomialverteilung. Diese Gruppe von Filtern wird daher als Binomialfilter oder auch als Gaußfilter bezeichnet. Eine einfache zweidimensionale Approximation eines Gaußfilters ist:

$${}^2B = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Schon dieses einfache Gaußfilter liefert für viele Aufgabenstellungen ausreichende Isotropie. Kanten werden durch das Gaußfilter verschliffen, ihre Lage bleibt jedoch erhalten. Das statistische Rauschen wird durch diese Filtergruppe erheblich reduziert, allerdings auf Kosten der Detailauflösung. Die Filterkoeffizienten des Gaußfilters entsprechen den Werten einer diskreten Binomialverteilung, die für sehr große Masken in die Gaußverteilung übergehen würde. Solche Filteroperationen sind sehr zeitaufwendig. Für die Maskengröße gibt es deshalb praktische, rechentechnisch bedingte Grenzen.

### Nachbemerkung

Die hier behandelten Filter gehören zur Gruppe der linearen Filter, die auch oft als „diskrete Faltungen“ bezeichnet werden. Streng genommen wurden hier jedoch nicht Faltungen, sondern Korrelationen betrachtet. Der Unterschied ist erst bei asymmetrischen Masken erkennbar und auch dann für viele Anwendungen in der Praxis irrelevant. In diesem Artikel wurden nur symmetrische Masken benutzt. Schon bei Kantenfiltern sind die Masken jedoch nicht mehr symmetrisch. In der weiterführenden Behandlung der Filter unter systemtheoretischen Aspekten wird der Unterschied zwischen Faltung und Korrelation herausgearbeitet [1], [2].

### Literatur

- [1] B. Jähne, Digitale Bildverarbeitung, Springer-Verlag
- [2] W. Burger, M. J. Burge, Digitale Bildverarbeitung, Springer-Verlag 2005, S. 79 ff.

## Y.HDR-Inspect Neue Dimension in der Röntgenprüfung

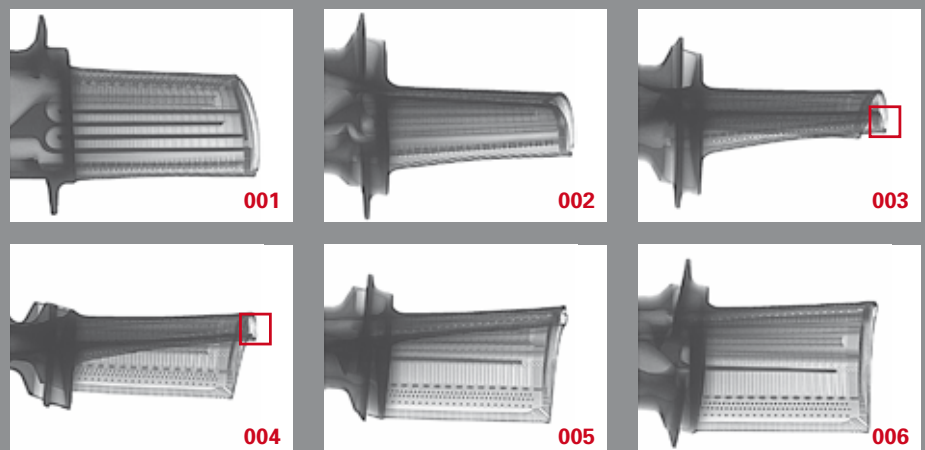
Y.HDR-Inspect bietet:

- Information über die räumliche Lage eines Fehlers im Prüfteil
- Information über die dreidimensionale Charakteristik des Fehlers
- Unverzögliche Sichtbarkeit des Fehlers in allen Materialstärken des Prüfteils durch Spezialfilter

Bitte nehmen Sie mit uns Kontakt auf:  
YXLON International GmbH  
Essener Bogen 15  
22419 Hamburg, Deutschland  
Telefon +49 40 527 29-101  
oder besuchen Sie uns auf  
unserer Website [www.yxlon.de](http://www.yxlon.de)



**YXLON**  
Technology with Passion



# Konturcodierung

## Kettencode und Konturanalyse

Bei industriellen Anwendungen der Bildverarbeitung sollen meist quantitative Eigenschaften von Objekten aus dem aufgenommenen Bild gewonnen werden, z. B. der Umfang, die Fläche oder die Lage des Schwerpunkts. Im Idealfall unterscheiden sich die Objekte vom Untergrund deutlich durch ihren Grauwert und treten dann durch ihre Kanten und Ecken, also mit ihrer Kontur, hervor. Dieser Artikel beschreibt die Codierung von Objekten anhand ihrer Kontur in einem Kettencode und gibt einige Beispiele für Objektmerkmale, die sehr einfach und effizient aus einer Konturcodierung berechnet werden können.

### Kontur

Die folgende Aussage ist unmittelbar einleuchtend: Wenn man die Länge der Umrandung eines Objekts aus einem Bild entnimmt, wird man ein gutes Maß für seinen Umfang erhalten. Ein einfaches Maß für die Länge der Umrandung ist die Zahl der Objektpixel, die auf dem Rand liegen. Besonders vielversprechend ist dieser simple Ansatz bei Binärbildern. Wenn sich die Grauwerte innerhalb eines Objekts hinreichend von den Grauwerten im Untergrund unterscheiden, also z. B. deutlich kleiner sind, kann das Objekt durch eine globale Grauwertschwelle vom Hintergrund getrennt werden. Alle Pixel mit Grauwerten unterhalb der Schwelle werden auf Null und alle Pixel mit Grauwerten ab der Schwelle werden auf 255 gesetzt. Das Ergebnis dieser Binarisierung ist ein Bild, in dem nur noch die Grauwerte 0 und 255 vorkommen: ein homogenes schwarzes Objekt liegt auf einem homogenen weißen Untergrund. Wenn es nun gelingt, alle Pixel zu identifizieren, die im Binärbild auf der Umrandung des Objekts liegen, ist die Aufgabe gelöst. Dazu sind zunächst einige Begriffsbildungen erforderlich, die auf den ersten Blick sehr formal erscheinen. Sie sind jedoch typisch für die Vorgehensweise in vielen Bereichen der Bildverarbeitung und erweisen sich auch im Zusammenhang mit der Analyse der Kontur als nützlich.

Zuerst muss definiert werden, was unter der „Kontur“ eines Objekts zu verstehen sein soll. Dabei spielen die Begriffe „Nachbarschaft“ und „Zusammenhang“ eine Rolle. Bei einem Recht-

eck ist es einfach, die Randpunkte und die vollständig im Innern des Objekts liegenden Punkte zu unterscheiden (s. Abb. 1). Komplizierter ist die Frage nach den Randpunkten bei dem zweiten in Abbildung 1 dargestellten Objekt. Gehören die beiden Bereiche, die lediglich „über Eck“ aneinander grenzen, überhaupt zu einem Objekt, oder handelt es sich um zwei getrennte Objekte, die lediglich durch die Aufnahme „zusammengewachsen“ sind? Diese Frage lässt sich selbstverständlich objektiv klären, wenn man einen Blick auf die realen Objekte werfen kann. Wenn aber nur noch die Bilddaten verfügbar sind, muss für diesen Fall eindeutig definiert sein, ob die beiden Bereiche als ein Objekt oder als zwei Objekte behandelt werden sollen. Üblicherweise (aber nicht notwendigerweise) werden in dieser Situation die beiden Bereiche als zu einem Objekt gehörig, als „zusammenhängend“ betrachtet. Der „Zusammenhang“ von Objektbereichen wird über den Begriff der „Nachbarschaft“ definiert. Auch der Begriff der „Nachbarschaft“ eines Bildpunktes muss definiert werden, denn es gibt kein intuitives Kriterium, um zu entscheiden, ob diagonal angrenzende Bildpunkte zur Nachbarschaft eines Pixels gehören sollen oder nicht. Deshalb werden zwei unterschiedliche Nachbarschaftsbegriffe auf einem quadratischen Raster verwendet (s. Abb. 2). Zur 4er-Nachbarschaft eines Bildpunktes gehören die vier unmittelbar, d. h. über eine Seite angrenzenden Pixel. Zur 8er-Nachbarschaft gehören zusätzlich die vier Pixel, die lediglich über eine Ecke an den Bildpunkt grenzen. Damit Vordergrundbereiche, die lediglich „über Eck“ miteinander zusammenhängen, zum selben Objekt gehören, muss innerhalb eines Objekts die 8er-Nachbarschaft als Zusammenhangskriterium verwendet werden. In den oben skizzierten Beispielen sind die Pixel, die zum Vordergrund gehören, schwarz, die Hintergrundpixel sind weiß. Wenn wir ein schwarzes Pixel herausgreifen und ein weiteres schwarzes Pixel in seiner 8er-Nachbarschaft finden, dann gehören beide Pixel zum selben Objekt, sie hängen zusammen. Für die Hintergrundbereiche wird dagegen die 4er-Nachbarschaft als Zusammenhangskriterium benutzt. Wenn wir also ein weißes Pixel herausgreifen und ein weiteres weißes Pixel in seiner Nachbarschaft finden, dann gehören beide Pixel nur dann zum selben Untergrundbereich, wenn sie innerhalb der 4er-Nachbarschaft zusammenhängen – aber nicht, wenn sie lediglich innerhalb der 8er-Nachbarschaft, also „über Eck“ zusammenhängen. Das ist deshalb sinnvoll, weil eine schwarze Gerade, die diagonal durch ein weißes Quadrat verläuft, das Quadrat in zwei voneinander getrennte, weiße dreieckige Untergrundbereiche zerlegen

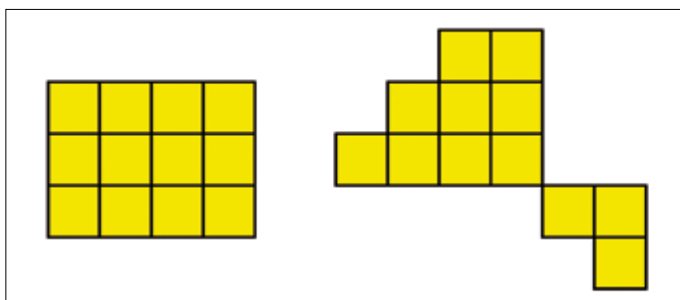


Abb. 1: Ein Rechteck und ein komplexeres Objekt in einem Binärbild. Zur Verdeutlichung sind die Objektpixel gelb gezeichnet.

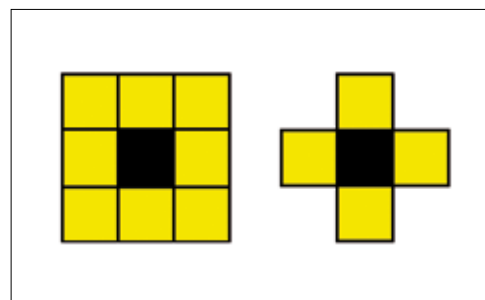


Abb. 2: Die Vierer-Nachbarschaft (rechts) und die Achter-Nachbarschaft (links) eines Pixels.

sollte. Wäre die 8er-Nachbarschaft als Zusammenhangskriterium im Untergrund zugelassen, würden die beiden Dreiecke zusammenhängen, der Untergrund würde also durch das Objekt hindurchfließen. Das würde zu Problemen führen. Beispielsweise könnte bei einer Unterlegscheibe die Bohrung im Innern nicht mehr vom Untergrund außerhalb der Scheibe unterschieden werden, wenn die Umrandung an einer Stelle nur noch „über Eck“ geschlossen wäre. Solche Aufnahmesituationen treten in der Praxis durchaus auf. Manchmal wird die Umrandung auch bewusst auf die Breite eines einzelnen Pixels verdünnt, z.B. bei einigen Verfahren der Zeichenerkennung. In der digitalen Bildverarbeitung wird daher die folgende, allgemein übliche und leistungsfähige Definition benutzt:

**Im Objekt gilt die 8er-Nachbarschaft (d.h. auch Punkte, die nur über eine Ecke mit anderen Objektpunkten verbunden sind, hängen mit dem Rest des Objektes zusammen), im Untergrund gilt dagegen die 4er-Nachbarschaft als Zusammenhangskriterium.**

Konturpunkte oder Randpunkte hängen zwar auch mit den übrigen Objektpixeln zusammen. Sie unterscheiden sich von innen liegenden Punkten aber dadurch, dass sich in ihrer Nähe auch Bildpunkte befinden, die zum Hintergrund gehören. In der digitalen Bildverarbeitung wird für die Kontur eines Objektes, das sich in einem Rasterbild mit rechteckigen Pixeln befindet, die folgende praktikable, allgemein übliche Definition benutzt:

**Die Kontur eines Objektes ist die Menge von Objektpixeln, die bezüglich der 8er-Nachbarschaft zusammenhängend sind (also mindestens ein Objektpixel als 8er-Nachbarn haben) und in der 4er-Nachbarschaft mindestens einen Nachbarn im Untergrund haben.**

Punkte, die unmittelbar am Bildrand liegen, müssen ebenfalls als Konturpixel zugelassen werden, auch wenn sie keinen Nachbarn im Untergrund haben. Für das komplexere Objekt in Abbildung 1 folgt aus den angegebenen Definitionen, dass es sich um ein einziges Objekt handelt.

### Freeman-Code

Grundsätzlich können nun alle Objektpixel untersucht und die Konturpixel aufgrund ihrer Nachbarschaftsrelation zum Untergrund identifiziert werden. Das Ergebnis ist eine Liste der Koordinatenpaare  $(x, y)$  der Konturpunkte. Die Zahl der Konturpixel ist dann bereits eine erste Schätzung für den Umfang des Objektes. Einfacher und effizienter ist es jedoch, zunächst einen Konturpunkt durch „Antasten“ zu bestimmen. Man beginnt „oben links“ im Bild und testet zeilen- und spaltenweise alle Pixel, bis man auf das erste schwarze Pixel trifft. Dieses Pixel gehört bestimmt zur Kontur, es ist ein sicheres Konturpixel. Von diesem Pixel aus kann man sich an der gesamten Kontur „entlangtasten“ und nacheinander alle anderen Konturpixel bestimmen. Wenn vom ersten, durch „Antasten“ ermittelten, Konturpixel aus die weiteren Konturpunkte nacheinander bestimmt sind, liegt die Kontur als Folge aneinandergrenzender Konturpixel vor. Man kann sich die Kontur dann als eine in sich geschlossene Kette vorstellen, die das Objekt vollständig umschließt. Da die Konturpunkte nun von einem Anfangspunkt aus als Liste unmittelbar aufeinander folgender Pixel ermittelt worden sind, kann die Kontur mit Hilfe eines Richtungscode als kompakte Folge von Zahlen abgelegt werden. Jeder Richtungscode  $f_p$  gibt die Richtung an, in die man gehen muss, damit man vom aktuellen Konturpunkt  $P$  zum nächsten Konturpunkt

kommt. Zur Richtungs-codierung auf einem rechteckigen Pixelraster sind nur acht Richtungen notwendig, die üblicherweise wie in Abbildung 3 links dargestellt bezeichnet werden. Wichtig ist dabei die Konvention, dass man die Kontur vom Startpunkt aus entgegen dem Uhrzeigersinn durchläuft. Der gesamte Konturcode besteht dann aus dem Koordinatenpaar  $(x, y)$  des Startpunktes und der Folge der Richtungscode  $f_p$ . Der Konturcode wird oft auch als Kettencode bezeichnet, mit den hier dargestellten Konventionen speziell als Freeman-Code [1]. Ein Beispiel für die Freeman-Codierung der Kontur ist in Abbildung 3 für das dort dargestellte Objekt angegeben. Der Startpunkt, dessen Koordinaten  $(x, y)$  explizit im jeweiligen Koordinatenraster bestimmt werden müssen, ist rot markiert.

Der Freeman-Code kann direkt durch eine Konturverfolgung berechnet werden, die vom ersten, durch „Antasten“ ermittelten, Konturpixel ausgeht. Dabei wird Schritt für Schritt die Richtung ermittelt, in der das nächste Konturpixel liegt, bis die gesamte Kontur abgearbeitet ist. Auf den ersten Blick erscheint diese Aufgabe recht einfach. Die Umsetzung in einen Algorithmus ist jedoch keineswegs trivial. Außerdem gibt es eine Reihe von Sonderfällen, z.B. Löcher in Objekten, dünne Linien oder einzelne Pixel, die verlässlich behandelt werden müssen. Eine detaillierte Beschreibung geeigneter Verfahren, z.B. der Konturverfolgung nach Pavlidis [2], geht jedoch über den Rahmen dieses Artikels hinaus. Die Konturverfolgung ist eine sehr effiziente und potentiell schnelle Methode, denn es werden nur die Konturpixel und ihre unmittelbare Umgebung untersucht. Die Pixel im Innern eines Binärobjektes müssen gar nicht abgerufen werden. Außerdem kann die Konturverfolgung zu einem Labeling-Verfahren erweitert werden, mit dem mehrere getrennte Objekte in einem Bild identifiziert und gekennzeichnet werden können. Solche Labeling-Methoden sind erheblich schneller als klassische Verfahren, die den Zusammenhang bei allen Bildpixeln sequentiell überprüfen.

### Auswertung des Konturcodes

In einem Binärbild enthält die Kontur die gesamte Information über das Objekt. Dies wird unmittelbar klar, wenn man sich vorstellt, dass aus der Kenntnis der Kontur grundsätzlich das gesamte Objekt rekonstruiert werden kann, z.B. durch Füllen der Kontur (eventuelle Löcher müssen dabei zusätzlich codiert sein). Folglich muss es möglich sein, aus der Kontur jede Größe zu berechnen, die man aus einer Liste der  $(x, y)$ -Koordinaten aller Objektpixel in einer Rastercodierung bestimmen könnte, z.B. die Fläche, den Schwerpunkt oder ein umschrei-

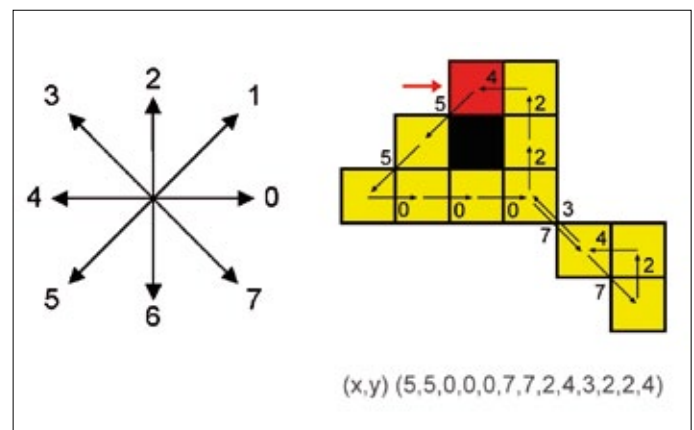


Abb. 3: Ein komplexes Objekt und die Freeman-Codierung seiner Kontur. Die Konturpixel sind gelb gezeichnet, das rote Pixel ist das durch Antasten von links und von oben gefundene erste Konturpixel. Links die Definition der Richtungscode für die acht möglichen Konturschritte.



bendes Rechteck. Das ist tatsächlich der Fall, und Freeman hat schon vor fast 50 Jahren die entsprechenden Algorithmen angegeben [3]. Diese Verfahren sind sämtlich potentiell sehr viel schneller als die direkte Berechnung aus der Rastercodierung aller Objektpixel, denn es muss lediglich der Konturcode ausgewertet werden.

Ein Beispiel ist die Berechnung der Konturlänge. Eine erste Schätzung ist einfach die Gesamtzahl der Einträge im Konturcode, entsprechend der Zahl der Konturpixel. Eine bessere Schätzung berücksichtigt, dass die „geraden“ Freeman-Richtungen den Schritten parallel zu den Koordinatenachsen entsprechen, die „ungeraden“ Freeman-Richtungen den Diagonalschritten. Bei quadratischen Pixeln ist ein Diagonalschritt um den Faktor  $\sqrt{2}$  länger als ein Schritt längs der beiden Achsen. Die Konturlänge ist dann die Summe aus der Zahl der „geraden“ Schritte und der mit  $\sqrt{2}$  gewichteten Zahl der „ungeraden“ Schritte.

$$U_c = n_g + n_u \sqrt{2}$$

Die Konturlänge ist häufig kein robustes Merkmal, weil sie stark vom Bildrauschen abhängt. Unter Umständen glättet man die Kontur vorher. Auch das ist direkt im Freeman-Code möglich.

Sehr einfach können aus dem Konturcode auch die maximalen und minimalen Werte für die x- und y-Koordinaten ermittelt werden, die bei einem Umlauf um die Kontur von den Konturpunkten angenommen werden. Die Differenz von Maximum und Minimum für die beiden Koordinaten x und y ist die Breite bzw. die Höhe des Objekts (genauer: die Länge der Projektion der Objektkontur auf die jeweilige Achse). Das Objekt liegt innerhalb eines Rechtecks mit diesen Abmessungen, dessen Seiten parallel zu den Koordinatenachsen verlaufen. Zur Berechnung betrachtet man nur Differenzen von Koordinaten und stellt zunächst eine Tabelle auf, die jedem Richtungscode die Änderung der x- und y-Koordinate zuordnet, die sich ergibt, wenn man mit diesem Richtungscode zum nächsten Konturpixel geht (s. Tab. 1). (dx, dy) kann man als Richtungsvektor interpretieren, der die Richtung zum nächsten Konturpunkt angibt. Über die gesamte Konturkette wird dann für jeden Konturpunkt, vom Startpunkt ausgehend, die Summe der jeweils bis zum aktuellen Konturpunkt aufgelaufenen Änderungen dx und dy bestimmt und am Ende jeweils das Maximum und Minimum dieser Summen ermittelt. Aus dem Konturcode wird also zunächst eine Folge von Richtungsvektoren (dx, dy), einer für jeden Konturpunkt, erzeugt. Dann werden für jeden Konturpunkt j die Summen  $b_j$  und  $h_j$  und daraus die Breite B und die Höhe H berechnet.

$$b_j = \sum_{i=0}^j dx_i$$

$$h_j = \sum_{i=0}^j dy_i$$

$$B = \max(b_j)$$

$$H = \max(h_j)$$

Auch die Fläche des Objekts erhält man aus dem Konturcode. Zur Berechnung erweitert man die Tabelle um eine weitere Spalte, die jeweils die Flächenänderung dA der Fläche zwischen Kontur und x-Achse für jeden Richtungscode enthält. Beginnend mit dem Startwert 0 für die Fläche wird die Kontur vollständig durchlaufen: auf dem „Hinweg“ werden überflüssige Flächenstücke zwischen Kontur und x-Achse mitgenommen, die auf dem „Rückweg“ wieder herausfallen. Übrig bleibt die Fläche zwischen den beiden „Wegen“. Die Summanden  $\frac{1}{2}$

entstehen dadurch, dass die Flächenelemente bei den Diagonalschritten Trapeze sind, bei den übrigen Schritten jedoch Rechtecke. Zunächst wird dazu nach der Tabelle für den nächsten Richtungscode die Änderung dy und daraus die neue y-Koordinate berechnet, das ist die y-Koordinate nach dem Konturschritt. Damit wird die Flächenänderung dA entsprechend der Tabelle bestimmt.

Tab. 1: dx, dy und dA für die acht möglichen Richtungscode

Richtungscode	dx	dy	dA
0	1	0	y
1	1	-1	y + 1/2
2	0	-1	0
3	-1	-1	-(y + 1/2)
4	-1	0	-y
5	-1	1	-(y - 1/2)
6	0	1	0
7	1	1	y - 1/2

Beginnend mit A = 0 am Anfangspunkt addiert man jetzt die jeweiligen dA beim Umlauf um die Kontur auf und erhält am Ende die eingeschlossene Fläche. (Die Spalte dx ist für diesen Zweck in der Tabelle gar nicht erforderlich). Die Fläche kann bereits simultan mit der Bestimmung des Konturcodes durch Konturverfolgung berechnet werden.

Zum Verständnis dieser Verfahren ist es sinnvoll, sich für einige einfache Pixelmuster die Konturlänge, das umschreibende Rechteck und die Fläche mit Bleistift und Papier auszurechnen. Bei der Interpretation der Resultate ist zu beachten, dass die Kontur im Sinne der Freeman-Codierung ein Polygonzug ist, der die Mittelpunkte der Konturpixel mit Geradenstücken entsprechend den acht möglichen Richtungen in einem rechteckigen Pixelraster verbindet. Die Fläche eines Objekts ist die Fläche innerhalb des geschlossenen Polygonzugs, und der Umfang eines Objekts ist die Länge des geschlossenen Polygonzugs. Ein Konturpixel trägt folglich nie vollständig zur Fläche des Objekts bei. Ein Quadrat, das aus vier Pixeln besteht, hat daher die Fläche 1 (und nicht 4), und eine Gerade, die aus einer Aneinanderreihung einzelner Pixel besteht, hat die Fläche 0.

Fazit

Die Beschreibung eines Binärobjekts über seine Kontur mit Hilfe eines Kettencodes ermöglicht sehr effiziente und schnelle Verfahren zur Berechnung von Objektmerkmalen wie Fläche, Umfang, Schwerpunkt oder umschreibendes Rechteck. Die Implementierung entsprechender Methoden erfordert die Auseinandersetzung mit zunächst abstrakt erscheinenden Begriffen der Bildverarbeitung. Der Aufwand lohnt sich jedoch, denn grundsätzlich ist die gesamte Information über ein Binärobjekt in seiner Kontur und damit im Kettencode enthalten.

Literatur

[1] H. Freeman, Boundary Encoding and Processing, in: B. S. Lipkin, A. Rosenfeld, Eds., Picture Processing and Psychopictonics, Academic Press, New York, 1970, p. 241 ff.  
 [2] T. Pavlidis, Algorithmen zur Grafik und Bildverarbeitung, Heise Verlag, 1990, S. 181 ff.  
 [3] s. Zitate 3-5 der Konferenzbeiträge von H. Freeman aus den Jahren 1961 und 1962 in [1]

# Kantenerkennung

## Klare Kante

Bei industriellen Anwendungen der Bildverarbeitung sollen meist quantitative Eigenschaften von Objekten aus dem aufgenommenen Bild gewonnen werden, z. B. der Durchmesser oder die Lage einer Bohrung. Im Idealfall unterscheiden sich die Objekte vom Untergrund deutlich durch ihren Grauwert und treten dann durch ihre Kanten erkennbar hervor. Dieser Artikel beschreibt einige einfache Methoden zur Erkennung von Kanten in Bildern.

### Konturen in Binärbildern

Bei sorgfältiger Auslegung der Aufnahmesituation unterscheiden sich die Grauwerte innerhalb eines Objekts hinreichend von den Grauwerten im Untergrund. Das Objekt kann dann durch eine globale Grauwertschwelle vom Hintergrund getrennt und in ein Binärbild überführt werden. Alle Pixel mit Grauwerten unterhalb der Schwelle werden auf Null und alle Pixel mit Grauwerten ab der Schwelle auf 255 gesetzt. Damit ist das Bild in Vordergrund und Hintergrund segmentiert, die Kanten des Objekts sind damit jedoch noch nicht gefunden. Diese Aufgabe kann aber ein wesentlicher Zwischenschritt der Bildverarbeitungskette sein. Beispielsweise basiert die Decodierung von Barcodes auf der Bestimmung der Kantenpositionen der Balken und Lücken. Auch die Methoden der Hough-Transformation, z. B. zur Erkennung von Geraden, werden auf Bilder angewendet, in denen die Kanten herausgearbeitet und möglichst auf die Breite eines Pixels reduziert sind. Generell wird an Kanten die grundsätzliche messtechnische Basis der Bildverarbeitung deutlich: hier treten lokale Kontrastunterschiede hervor, die letztlich durch die Beleuchtung entstehen.

Ein gut präpariertes Ausgangsbild und die Extraktion von Kanten ist in Abbildung 1 zu sehen. Das Quellbild kann gut mit einer konstanten Binarisierungsschwelle segmentiert werden. Im nächsten Schritt, dem Labeling, wird der Zusammenhang zwischen den Vordergrundpixeln untersucht. Das Ergebnis ist nicht nur eine eindeutige Zuordnung jedes Vordergrundpixels zu einem Konglomerat von zusammenhängenden Pixeln („Blob“, in diesem Fall sechs getrennte Objekte), sondern auch die Identifikation der sog. Konturpixel, also der Kanten der Objekte. Der Labeling-Algorithmus liefert diese Information als Nebenprodukt, denn Konturpixel sind zwar wie alle anderen Objektpixel auf den Grauwert 255 gesetzt, haben aber außerdem, anders als die inneren Objektpixel, mindestens einen Nachbarn im Untergrund. Eine Möglichkeit der Kantenerkennung ist also

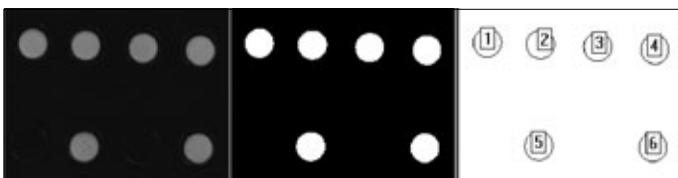


Abb. 1: Graustufenbild, Binärbild und Labelbild mit Konturausgabe

das Labeling mit Darstellung der Konturpixel der gelabelten Objekte. Auch die Konturcodierung im Kettencode, die auf der Konturverfolgung basiert, liefert unmittelbar die Kanten im Bild.

### „Antasten“

Die Labeling-Funktion und die Konturverfolgung werden zwar von vielen Bildverarbeitungsbibliotheken als Funktionalität angeboten, sind aber keineswegs trivial. Auch das Labeling muss zunächst einmal ein Objekt im Bild lokalisieren. Dazu wird das Binärbild zeilenweise von links oben nach rechts unten abgetastet und für jedes Pixel der Grauwert geprüft. Das erste Pixel, dessen Grauwert 255 ist, gehört zu einem Objekt und ist zwangsläufig ein Konturpixel, liegt also auf einer Kante. Diese Methode bezeichnet man als „Antasten“. Durch Antasten von links, rechts, oben und unten kam man diejenigen Konturpixel finden, die den Bildrändern am nächsten liegen. In vielen industriellen Anwendungen werden Kanten mit dieser simplen, sehr effizienten Methode detektiert. Das funktioniert ausgezeichnet, wenn aufgrund von Vorkennntnis vorausgesetzt werden kann, dass Kanten in ganz bestimmten Bildbereichen liegen müssen, ihre Orientierung und genaue Position aber aufgrund der Zuführung der Prüfteile nicht exakt definiert sind. Solche Kanten werden dann durch Antasten in vorher festgelegten Toleranzbereichen (ROIs) „gefangen“. Obwohl in manchen Fällen schon eine einzige Suchlinie ausreichen würde, um die Kante zu lokalisieren, sucht man sicherheitshalber in mehreren eng benachbarten Suchlinien und testet auf Konsistenz, damit nicht ein verstreutes Störpixel eine Kante vortäuschen kann. Zusätzlich wird man noch für einige Pixel jenseits des Kantenpixels prüfen, ob man sich noch innerhalb eines Objekts befindet. Alternativ kann man auch über mehrere Zeilen mitteln und auf diese Weise Störungen eliminieren. In jedem Fall ist das Antasten eine sehr effiziente Methode, wenn man Vorkennntnis über die Szene nutzen kann. Schließlich muss nicht das gesamte Bild untersucht werden, wie das beim Labeling erforderlich ist, sondern man kann die Kantensuche einstellen, sobald man die Kante gefunden hat. Eine Erweiterung des simplen Antastens längs Zeilen und Spalten sind „Antastpfeile“ in verschiedenen Richtungen, die auch auf ein gemeinsames Zentrum weisen oder von einem gemeinsamen Zentrum ausgehen können, etwa um die Rundheit einer Bohrung zu überprüfen.

### Antasten im Graustufenbild

Wer je ein „Antasten“ programmiert, wird schnell gemerkt haben, dass die Einschränkung auf Binärbilder meist gar nicht erforderlich ist. Es spricht nichts dagegen, gleich mit der Analyse der Grauwerte auf der Suchlinie zu starten und die Kante beim ersten Grauwert zu lokalisieren, der oberhalb der Schwelle für die Binarisierung liegt. Die nächste Verallgemeinerung besteht darin, auch die Absolutschwelle für den Grauwert aufzugeben; es genügt, den Grauwertunterschied zwischen dem nächsten Pixel auf der Abtastlinie und dem aktuellen Pixel gegen eine Schwelle zu testen! Die Kante wird dann immer zuverlässig gefunden, solange die Objekte sich überhaupt vom Hintergrund hinreichend abheben. Selbst inhomogene Beleuchtung oder den Randabfall von Objektiven verkraftet diese Methode sehr



Abb. 2: Graustufenbild, dilatiertes Graustufenbild (oben) und Ergebnis der Differenzbildung (links unten); zum Vergleich rechts unten Ergebnis einer Sobel-Filterung mit Detailausschnitt aus den Kantenbildern

gut, denn es kommt immer nur auf den lokalen Grauwertunterschied zwischen Objekt und Hintergrund an. Selbstverständlich darf es weder extrem dunkel noch extrem hell (übersteuert) im Bild werden, und die Kantenschwelle muss deutlich über dem Rauschen im Bild liegen. Abgesehen davon ist diese Kantenerkennung erstaunlich robust, denn sie ist eine differentielle Methode, die den lokalen Kontrast auswertet.

Die systematische Weiterentwicklung dieser Idee ist das klassische Kantensfilter. Jeder Mathematiker würde von vornherein eine Kante mit einem Anstieg oder Abfall im Grauwertprofil assoziieren und die Kante über die Steigung dieses Profils vom Rest des Bildes isolieren. Die Position der Kante ist dann durch das Maximum der Steigung gegeben. Die Differenzbildung zwischen den Grauwerten aufeinander folgender Pixel ist nichts anderes als eine primitive Implementierung der ersten Ableitung der Graustufenfunktion längs der betrachteten Suchrichtung, also z.B. längs der x-Achse, auf dem diskreten Raster der Bildebene. Um sicherzustellen, dass man Kanten unabhängig von ihrer Richtung im Bild findet, muss man isotrope Filter konstruieren, z.B. den Betrag des Gradientenvektors, den man als Grauwert in einem Ergebnisbild darstellen kann. Kanten erscheinen darin hell, Bereiche mit konstantem Grauwert dunkel. Kantensfilter sind in der Literatur ausführlich beschrieben und werden daher hier nicht weiter erläutert [1, 2].

### Ausdünnung

Als Beispiel für ein typisches Resultat einer Kantensfilter-Operation zeigt Abbildung 2 im unteren rechten Teilbild das Ergebnis einer Sobel-Filterung. Ein Sobel-Filter bildet die Ableitung in einer Richtung und glättet in der dazu senkrechten Richtung. Die Kanten sind als helle Regionen im Kantenbild sichtbar. Offenkundig isoliert das Sobel-Filter zwar die Kanten im Bild recht gut, eine Kante erstreckt sich aber über mehrere Pixel. Zur Lokalisierung der Kante muss das Kantenbild also nachbearbeitet werden. Das ist grundsätzlich möglich; beispielsweise kann man das Maximum in der Kantenregion suchen. Solche Verfah-

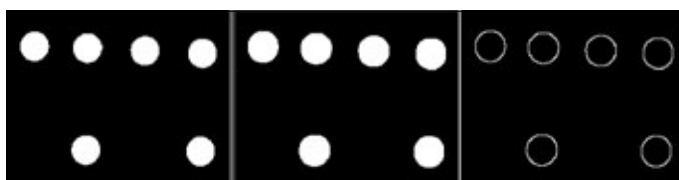


Abb. 3: Binärbild, dilatiertes Binärbild und Ergebnis der Differenzbildung

ren sind jedoch zeitaufwendig, und in der industriellen Anwendung müssen sie außerdem robust programmiert werden. Geschickter sind Verfahren, die sofort verdünnte Kanten liefern. In Binärbildern gibt es dafür verlässliche Methoden. Ein Beispiel zeigt Abbildung 3. Hier wurde auf das binäre Quellbild eine Binärdilatation angewandt, die dazu führt, dass die Objekte an der Peripherie um ein Pixel ausgedehnt werden. Anschließend wird die Differenz zwischen dem Originalbild und dem dilatierten Bild gebildet, so dass im Ergebnis der angebaute Rand stehen bleibt und die Kontur als Kantenbild mit einer Breite von einem Pixel wiedergibt. Auch im Graustufenbild gibt es die Dilatation, sie hat aber eine etwas andere Bedeutung als im Binärbild und führt nicht einfach zu einer Dehnung des Objekts um ein Pixel. Das Resultat der Subtraktion vom Originalbild ergibt daher nicht einfach ein ausgedünntes Konturbild, sondern ähnelt stark dem Resultat einer Kantensfilteroperation, siehe Abbildung 2, Detailausschnitt. Generell ist auch die Dilatation als Rangordnungsfilter zeitaufwendig. Wer keinesfalls filtern möchte, kann einfach das gesamte Bild um ein Pixel nach rechts und ein Pixel nach unten verschieben und das so verschobene Bild mit dem Quellbild logisch verknüpfen, z.B. subtrahieren. Das Ergebnis ist für viele Anwendungen als Kantenbild völlig ausreichend und kann je nach Hardwareplattform sehr schnell berechnet werden.

### Fazit

Ausgedünnte Kantenbilder können aus Binärbildern durch Labeling, Konturverfolgung oder Antasten in einfacher Weise berechnet werden. Alternativ kann man eine Binärdilatation verwenden und die Differenz zum Quellbild berechnen; diese Operation liefert direkt ein ausgedünntes Kantenbild. In Graustufenbildern kann man Kantenpositionen durch Antasten aus der Steigung des Grauwertverlaufs ermitteln. Diese Methode wird meist eingesetzt, wenn Vorkenntnis über die Lage und Richtung der Kanten vorliegt. Allgemeingültige Verfahren müssen auf Kantensfilter zurückgreifen. Die Kanten erscheinen dann als helle, über mehrere Pixel ausgedehnte Regionen, die ggf. anschließend verdünnt werden müssen.

### Literatur

- [1] B. Jähne, Digitale Bildverarbeitung, Springer-Verlag
- [2] W. Burger, M. J. Burge, Digitale Bildverarbeitung, Springer-Verlag 2005, S.111 ff.



# Konturbasiertes Pattern-Matching

## Konturverfolgung

Beim Pattern-Matching wird ein Bild nach einem vorgegebenen Muster abgesucht: dem „template“ oder dem „Modell“. Das Verfahren liefert grundsätzlich für jeden Bildpunkt ein Ähnlichkeitsmaß. Je höher dieser „score“, desto höher ist die Wahrscheinlichkeit, dass sich das gesuchte Muster tatsächlich an dieser Stelle befindet. Es gibt Vergleichsmethoden, die sich auf die Kontur von Objekten beziehen. Sie setzen nicht an Grauwertmustern an, sondern verarbeiten die Kanten, die in einem Bild vorhanden sind.

### Konturen

Die Kontur ist die Randlinie eines Binärobjekts. Die Begriffsbildung setzt folglich voraus, dass Objekte bereits vom Hintergrund segmentiert sind, z.B. durch eine Binarisierung. Hintergrundpixel können z.B. schwarz sein, Vordergrundpixel weiß. Abbildung 1 zeigt ein Beispiel für diese Situation. Eine Zusammenhangsanalyse, das „labeling“, liefert dann die einzelnen Binärobjekte und in der Regel gleich auch deren Umrandung, wie im rechten Teil von Abbildung 1 zu sehen ist. Konturpixel sind dabei diejenigen Objektpixel, die sowohl Nachbarn im Hintergrund als auch im Objekt haben (genauer: die mit dem Objekt über die Achter-Nachbarschaft und mit dem Hintergrund über die Vierer-Nachbarschaft zusammenhängen). Aufgrund dieses Kriteriums kann man diese speziellen Objektpunkte abtrennen und separat als Kontur abspeichern. Eine bekannte und effiziente Methode zur Codierung der Kontur ist ein Richtungscode. Dabei werden die einzelnen Konturpixel, ausgehend von einem Startpunkt, längs der Kontur verfolgt und der jeweils erforderliche Richtungsschritt zum nächsten Konturpunkt als ganze Zahl zwischen 0 und 7 abgelegt. Abbildung 2 zeigt ein Beispiel, das den sog. Freeman-Code verwendet. Den Startpunkt (x, y), im Bild rot gezeichnet, ermittelt man durch Antasten von links nach rechts und von oben nach unten. Er ist auf diese Weise eindeutig definiert. Die Kontur enthält die gesamte Information über das Binärobjekt. Das wird verständlich, wenn man sich klar macht, dass man das Objekt aufgrund der Kontur vollständig rekonstruieren kann. Es muss folglich möglich sein, ein Objekt allein aufgrund seiner Kontur zu erkennen. Betrachten wir z.B. ein ideales Quadrat, dessen Kanten parallel zu den Koordinatenachsen liegen. Der Startpunkt (x, y) der Codierung ist dann die linke obere Ecke. Die Kontur wird gegen den Uhrzeigersinn abgearbeitet, so dass die erste Kante die Richtung 6 hat, die zweite die Richtung 0, die dritte die Richtung 2 und die vierte die Richtung 4. Je nach Kantenlänge des Quadrats wird folglich der Freeman-Code mit einer Gruppe von z.B. 20 Richtungen mit Code 6 beginnen, gefolgt von 20 Richtungen 0 usw. Vier gleich lange Gruppen von Richtungen, bei denen am Ende ein Wechsel um 90° entgegen dem Uhrzeigersinn auftritt, codieren demnach ein Quadrat. Die Richtungsänderung um 90° entspricht einer Addition von +2 (modulo 8) auf die aktuelle Richtung, ist also algorithmisch sehr einfach feststellbar. Selbst ein

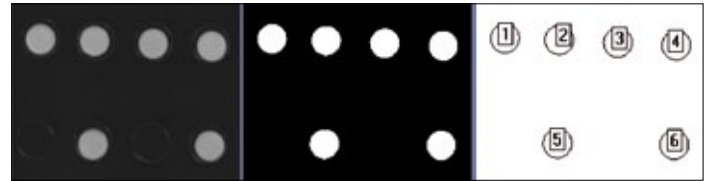


Abb. 1: Graustufenbild, Binärbild und Labelbild mit Konturen der Binärobjekte

ideales Rechteck wird jedoch nach einer realen Bildaufnahme nicht mehr nur vier einheitliche Richtungsgruppen aufweisen, erst recht nicht bei freier Drehlage. Allerdings kann man den Freeman-Code glätten und dann auf Geradenstücke mit entsprechenden Richtungswechseln am Ende zurück schließen. Mit diesem einfachen Beispiel ist bereits die grundlegende Idee des konturbasierten Pattern-Matching beschrieben. Aus der Kontur wird ein geometrisches Modell des Objekts gewonnen, das nicht als Grauwertmuster, sondern in abstrakten Begriffen beschrieben wird. In diesem Sinne liegt ein Quadrat vor, wenn die Kontur aus vier gleich langen Geradenstücken besteht, die an den Endpunkten aneinandergrenzen, und wenn dort ein Richtungswechsel um 90° gegen den Uhrzeigersinn auftritt. Es ist unmittelbar klar, dass diese Beschreibung unabhängig von der Drehlage und unabhängig von der Größe des Objekts immer ein Quadrat beschreibt. Das Modell ist also a priori lage- und skalierungsinvariant, eine sehr nützliche Eigenschaft.

### Korrelation auf der Kontur

Wie oben angedeutet, kann direkt der geglättete Konturcode mit dem Konturcode des gesuchten Objekts verglichen werden. Sofern das Verfahren skalierungsinvariant sein soll, also auch gleichartige Objekte unterschiedlicher Größe gefunden werden sollen, muss der Konturcode zunächst auf eine Referenzlänge normiert und entsprechend umgerechnet werden. Wenn auch die Drehlage frei ist, beginnen die Konturcodierungen in der Regel bei verschiedenen Startpunkten auf der Kontur. Die Vergleichsprozedur wird dadurch wesentlich komplexer. Gelegentlich behilft man sich mit der Berechnung von Merkmalen aus dem Konturcode, z.B. einem Histogramm der Richtungs-codes. Dabei ist jedoch zu beachten, dass zwangsläufig Information verloren geht und die Klassifizierung im Allgemeinen nicht mehr eindeutig ist. Eine Alternative ist die Codierung der Kontur als Radiusfunktion. Dazu wird der Schwerpunkt des Objekts ermittelt und, ausgehend von einer Referenzrichtung, in festen Winkelabständen der Abstand zwischen Schwerpunkt und Kontur bestimmt. Dieser „Radius“ oder auch Polarabstand wird über dem Winkel aufgetragen. Die Radiusfunktion ist charakteristisch für das Objekt und wird gelegentlich als Signatur bezeichnet. Ein Beispiel zeigt Abbildung 3. Die rote Kurve im rechten oberen Bereich der Abbildung ist die geglättete Radiusfunktion für das fünfeckige Objekt, das links abgebildet ist. Für ein zweites, in einem Suchbild gefundenes Objekt schiebt man die vorgegebene über die gefundene Radiusfunktion und berechnet jeweils die Korrelation. Liegt die Korrelation über einer vorher empirisch festgelegten Schwelle, gilt dies als Übereinstimmung. Damit ist das Objekt gefunden und außerdem seine Drehlage bestimmt, denn es ist bekannt, bei welcher Win-

kelposition die Übereinstimmung aufgetreten ist. Dieses Verfahren kann man als die Übertragung des klassischen „pattern matching“ mit Berechnung der Grauwert-Korrelation zwischen „template“ und Bildausschnitt auf die Auswertung der Kontur betrachten und somit als „konturbasiertes Pattern Matching“ bezeichnen. Die abstrakten Eigenschaften des Modells werden dabei jedoch gar nicht verwendet.

### Geometric Model Finder

Das Pattern Matching, angewandt auf die Konturcodierung, setzt ein Binärobjekt im Bild voraus, also eine erfolgreiche Segmentierung und eine durchgängige Randlinie. Die Beschreibung des gesuchten Objekts in Form eines geometrischen Modells hingegen eröffnet eine Möglichkeit, das Modell direkt mit Strukturen im Graustufenbild zu vergleichen. Dazu wird das Graustufenbild zunächst in ein Kantenbild umgerechnet, und im Idealfall werden die Kanten auf eine Breite von einem Pixel verdünnt. Die Konturen von Objekten treten dann als Kanten hervor. Allerdings gibt es auch Kanten, die nicht unmittelbar zu Objekträndern gehören, sondern aus der Textur von Oberflächen oder aus Schatten entstehen, und manche Konturen sind unterbrochen. Auf diesem Kantenbild wird im nächsten Schritt nach den geometrischen Grundformen gesucht, auf die sich das Modell des gesuchten Objekts bezieht. Handelt es sich z.B. um das oben diskutierte Quadrat, so wird das Kantenbild zunächst nach Geraden abgesucht. Für diesen Zweck gut geeignet ist die Hough-Transformation, die für Geradenstücke und auch für unterbrochene Geraden im Bild jeweils einen „score“ berechnet, der es erlaubt, die besten Kandidaten für tatsächlich vorliegende Geraden im Bild von zufälligen Übereinstimmungen mit dem Geradenmodell zu trennen [1]. Auch für Kreise gibt es eine Variante der Hough-Transformation, die wir im Rahmen dieser Artikelserie bereits behandelt haben [2]. Geraden und Kreise reichen völlig aus, um reale Objekte im Sinne einer Polygondarstellung, ergänzt um Kreissegmente, zu modellieren. Für das oben genannte Beispiel eines Quadrats bedeutet das geometrische Modell, dass vier gleich lange Geraden im Bild vorkommen müssen, die sich an den Endpunkten berühren und für die am Ende jeweils ein Richtungswechsel von 90° gegen den Uhrzeigersinn auftritt. Werden im Kantenbild nur drei gleich lange Geraden gefunden, und die vierte Gerade

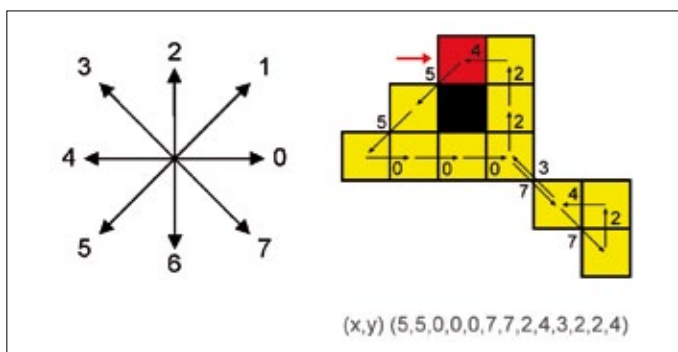


Abb. 2: Richtungs- und Konturcodierung nach Freeman (links) und Beispiel für die Konturcodierung (rechts). Die Kontur wird vom Startpunkt (x,y) aus (rot) entgegen dem Uhrzeigersinn abgearbeitet und die Richtung des jeweils nächsten Konturschritts notiert.

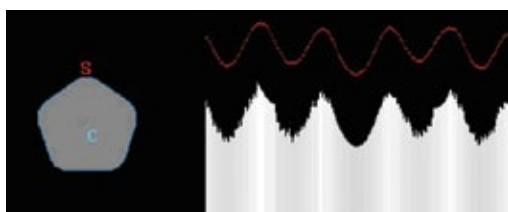


Abb. 3: Ein Objekt und seine zugehörige Radialfunktion. C ist der Schwerpunkt des Objekts, S das erste Konturpixel im Polarabstandsplot.



Abb. 4: Graustufenbild (oben), Kantenbild (Mitte) und Ergebnis der Hough-Transformation für Kreise (unten); eingeblendet in das Originalbild sind als Punkte die Mittelpunkte der gefundenen Kreise. Auch der teilweise verdeckte Schlüsselring wird gefunden.

ist nur halb so lang wie die drei anderen, die Anforderungen in Bezug auf die Richtungen sind jedoch erfüllt, und es gibt außerdem drei Berührungspunkte an den Enden, so ist möglicherweise lediglich ein Teil der Kontur des Quadrats von einem anderen Objekt überdeckt oder durch einen Schatten verdunkelt. Man wird diese Struktur dann vielleicht nicht mit Sicherheit als Quadrat klassifizieren, aber sehr wohl als einen Kandidaten für weitergehende Analyse in Betracht ziehen. Viele „Geometric Model Finder“ arbeiten mit mehr oder weniger komplexen Varianten der Hough-Transformation für geometrische Primitive und nutzen die Relationen zwischen diesen Strukturelementen im Rahmen eines vorher aufgestellten Modells oder verwenden eine Verallgemeinerung der Hough-Transformation [3]. Dadurch wird erreicht, dass auch Objekte gefunden werden, deren Kontur nicht durchgängig

im Kantenbild erscheint oder die teilweise von Fremdobjekten überdeckt sind. Je mehr Parameter zugelassen werden, desto aufwendiger wird die Implementierung, und die Ausführungszeit stößt schnell an die Grenzen, die zumindest bei der Bildverarbeitung in der laufenden Produktion üblich sind.

### Fazit

Wer von konturbasiertem Pattern Matching spricht, meint meist Suchverfahren in Graustufenbildern, die auf Kantenbildern aufsetzen und geometrische Modelle von Objekten verwenden. Solche „geometric model finder“ funktionieren auch bei teilweise überdeckten Objekten und können skalierungs- und lageinvariant sein. Wie die klassische Grauwertkorrelation liefern auch die konturbasierten Verfahren lediglich eine (gelegentlich recht geringe) Wahrscheinlichkeit für die Übereinstimmung zwischen den Strukturen im Suchbild und dem Modell. Es gibt eine große Variantenvielfalt für die Konzeption und Implementierung dieser Methoden [4]. Abbildung 4 verdeutlicht am Beispiel der Hough-Transformation für Kreise, dass auch einfache Varianten bereits stabil gegen Reflexe und inhomogene Beleuchtungsverteilung sein können, Fremdstrukturen gut diskriminieren und auch bei Überdeckung funktionieren. Komplexere Modelle geraten in Bezug auf die Rechenzeit schnell außer Kontrolle und müssen sorgfältig optimiert werden [4]. Zweifellos sind die „geometric model finder“ eine wertvolle Ergänzung des Instrumentariums der Bildverarbeitung und gehören mittlerweile zu den Grundlagen dieses Gebietes.

### Literatur

[1] W. Burger, M. J. Burge, Digitale Bildverarbeitung, Springer, 2005, S.152–167  
 [2] INSPECT 3/2008, S. 24 ff.  
 [3] D.H. Ballard, Generalizing the Hough-transform to detect arbitrary shapes, Pattern Recognition Vol. 13, No. 2, p 111–122, 1981  
 [4] Steger, C., Ulrich, M., Wiedemann, C., Machine Vision Algorithms and Applications, Wiley-VCH, Weinheim 2008, p. 222 ff.

# Die Hough-Transformation

## Auf der Geraden ins Ziel

Der Begriff „Hough-Transformation“ weckt weitgehend die Assoziation mit einer Methode zum Auffinden von Geraden in einem Bild. Wer sich zum ersten Mal eingehend mit dem Verfahren befasst, kann sich leicht im Gestrüpp der mathematischen Formalitäten verlieren, die für die rigorose Umsetzung unumgänglich sind. Die grundlegende Idee der Hough-Transformation ist jedoch sehr einfach. Sie wird unmittelbar deutlich, wenn man als Einstieg die Hough-Transformation für Kreise anstatt für Geraden wählt. Wer das Konzept auf diese Weise einmal erfasst hat, wird dem Gedankengang zur Verallgemeinerung auf andere analytisch beschreibbare Kurven, inklusive Geraden, problemlos folgen können. Dabei wird der wesentliche Vorzug der Hough-Transformation deutlich: sie ermöglicht die Erkennung von Formen und Kurven, auch wenn sie unvollständig oder teilweise überdeckt sind.

### Die grundlegende Idee

Wir betrachten zunächst ein einfaches Binärbild wie im linken Teil der Abbildung 1, eine vollständige Kreislinie aus schwarzen Pixeln auf weißem Hintergrund. Angenommen, der Radius des Kreises sei vorab bekannt, und es solle der Mittelpunkt des Kreises gefunden werden. Ein Kreis ist die Menge aller Punkte, die von einem bestimmten Punkt, dem Mittelpunkt, sämtlich denselben Abstand haben, den Radius des Kreises. Den Mittelpunkt kann man folgendermaßen finden: Das Bild wird zeilenweise von links oben nach rechts unten Pixel für Pixel abgerastert. Um jedes Pixel wird ein Testkreis mit dem Radius des gesuchten Kreises gezeichnet. Dann wird die Überdeckung zwischen dem Testkreis und dem tatsächlichen Kreis im Bild ermittelt. Dazu wird das Bild längs des Testkreises abgetastet und gezählt, wie viele schwarze Pixel dabei gefunden werden. Dieses Ergebnis wird in einem neuen Bild als Zahl am aktuellen Mittelpunkt des Testkreises abgelegt. Für viele Stellen im Bild wird es keinerlei Überdeckung geben, der resultierende Wert ist Null; an einigen Stellen gibt es ein oder zwei Schnittpunkte zwischen Testkreis und dem Binärkreis im Bild; sobald jedoch der Mittelpunkt des Binärkreises im Bild als Mittelpunkt des Testkreises verwendet wird, entsteht die größte Überdeckung, die mit diesem Verfahren möglich ist. Die Zahlen, die im Ergebnisbild gespeichert sind, können als Grauwerte interpretiert werden, und der Mittelpunkt des Binärkreises erscheint als heller Fleck auf dunklem Hintergrund.

Dieser Ansatz liefert gute Ergebnisse, ist jedoch nicht sonderlich effizient. In vielen Fällen werden damit große Bereiche im Untergrund abgetastet, obwohl die schwarzen Pixel des Binärkreises weit entfernt liegen. Offensichtlich wäre es viel geschickter, die Vordergrundpixel heranzuziehen anstatt das gesamte Bild abzurastern. Statt um jedes Bildpixel, ob nun Vor-

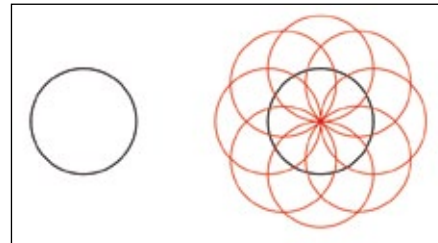


Abb. 1: Ein schwarzer Kreis und einige rote Kreise mit demselben Radius, deren Mittelpunkte auf der schwarzen Kreislinie liegen

dergrund oder nicht, einen Kreis zu zeichnen und die Überdeckung mit dem Binärkreis zu bestimmen, werden nun Kreise mit dem gesuchten Radius nur um die schwarzen Vordergrundpixel gezeichnet. Einige dieser Kreise sind im rechten Teil der Abbildung 1 gezeichnet, hervorgehoben durch die rote Farbe. Die roten Kreise haben sämtlich einen Punkt gemeinsam: den Mittelpunkt des gesuchten Kreises. Dort schneiden sich alle so gezeichneten Kreise, und die Zahl der Schnitte ist gleich der Zahl der Pixel auf dem schwarzen Binärkreis. Das ist dieselbe Zahl, die sich nach dem oben geschilderten Verfahren für die maximale Überdeckung ergibt. Wer kurz darüber nachdenkt, stellt fest, dass beide Methoden äquivalent sind. Das zweite Verfahren ist jedoch wesentlich effizienter, weil es Kreise nur um die Vordergrundpixel zeichnet und nicht um jedes Pixel im Bild. Dazu wird zunächst ein Ergebnisbild angelegt und mit dem Grauwert Null initialisiert. Das Quellbild wird wieder zeilenweise von links oben nach rechts unten Pixel für Pixel abgerastert. Nur wenn der Algorithmus dabei auf ein schwarzes Vordergrundpixel stößt, wird im Ergebnisbild an der entsprechenden Stelle ein Kreis mit dem gesuchten Radius gezeichnet und an jedem Pixel im Ergebnisbild, das auf diesem Kreis liegt, der Grauwert um 1 erhöht. Nachdem diese Prozedur für alle Vordergrundpixel im Quellbild abgearbeitet ist, hat sich im Ergebnisbild an jedem Pixel die Anzahl der Schnitte zwischen den eingezeichneten Kreisen durch Hochzählen angesammelt. Das Ergebnisbild wird daher auch als Akkumulatorarray bezeichnet, die Anzahl der Schnitte an jedem Pixel wird durch den Grauwert repräsentiert. Der hellste Fleck im Akkumulatorbild ist der Mittelpunkt des gesuchten Kreises. Falls der Testradius falsch gewählt wurde, wird es immer noch Schnitte geben, ihre Zahl ist aber deutlich geringer als für den korrekten Suchradius. Die Mittelpunkte von realen Kreisen im Bild werden sich im Akkumulator folglich deutlich als Spitzen über einem Rauschteppich zeigen. Das ist das Prinzip der Hough-Transformation.

Wir lassen nun die Annahme fallen, dass der Radius des Kreises bereits bekannt ist. Gesucht ist nun nicht nur der Mittelpunkt, sondern auch der Radius. Dazu gibt man ein Intervall für die Radien der Testkreise vor und führt die oben beschriebene Prozedur für alle Radien aus dem Intervall (oder auch für eine Auswahl daraus) durch. Das Resultat ist dann nicht nur ein Ergebnisbild, sondern ein Bildstapel mit je einem Bild für jeden Radius. Wie zuvor gibt es jedoch in dem Stapel ein Bild, dessen Helligkeitsmaximum größer ist als bei allen anderen Akkumulatorarrays. Die Position dieses Maximums ist der Mittelpunkt des gesuchten Kreises, und der zugehörige Testradius ist der gesuchte Radius. Selbst wenn der Bildstapel lediglich zu einem einzigen Ergebnisbild aufaddiert wird, entsteht häufig ein heller Fleck am Mittelpunkt des tatsächlichen Kreises.



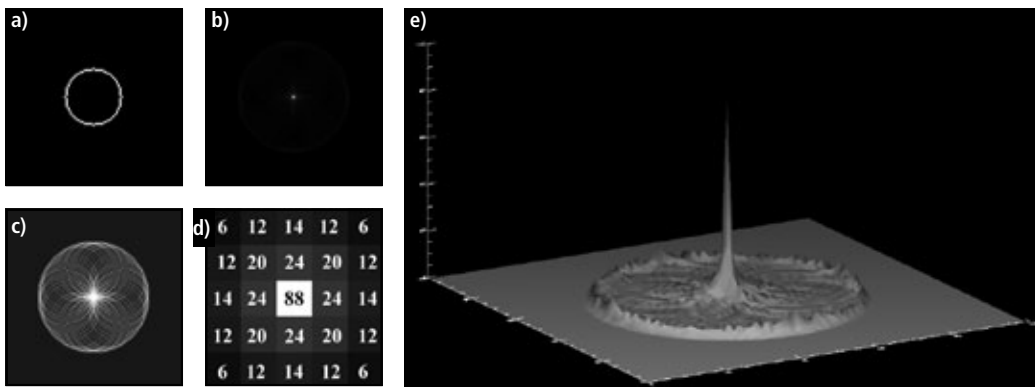


Abb. 2: Hough-Transformation für einen Kreis  
 a) Binär-Kreis in der diskreten Bildebene; b) Akkumulator-Bild, Ergebnis der Hough-Transformation; c) Akkumulator-Bild, kontrastverstärkt; d) Zentraler Bereich des Akkumulators, Zahlen sind die Grauwerte der Pixel; e) 3D-Plot der Grauwerte des Akkumulators



Abb. 3: a) Quellbild; b) Konturen sind herausgearbeitet und verdünnt; c) Kontrastverstärktes Akkumulator-Bild nach Anwendung der Hough-Transformation auf das verdünnte Konturbild

### Ein einfaches Beispiel

In Abbildung 2 sind die Ergebnisse dieses Verfahrens für einen synthetisch erzeugten Binärkreis gezeigt. Die Pixelstruktur in der diskreten Bildebene ist deutlich erkennbar. Das entstehende Akkumulator-Bild mit dem hellen Fleck an der Stelle des Kreismittelpunkts hat große Ähnlichkeit mit dem rechten Teil der Abbildung 1. Es gibt darin weitere, jedoch deutlich schwächere Strukturen, die in einem Teilbild von Abbildung 2 zu sehen sind, bei dem der Kontrast verstärkt ist. Zur Veranschaulichung sind die mittleren 5x5 Pixel des Akkumulators vergrößert gezeichnet. Die Zahlen sind die zugehörigen Grauwerte der Pixel. Das letzte Teilbild vermittelt in einem 3D-Plot der Grauwerte über den Bildkoordinaten einen Eindruck vom Signal-Rausch-Verhältnis der Methode. Bilder von realen Szenen wie in Abbildung 3 können ähnliche Strukturen aufweisen. Um die Hough-Transformation effizient umzusetzen, sollten zunächst die Konturen im Bild herausgearbeitet werden, am besten verdünnt auf ein Pixel, wie beim zweiten Teilbild der Abbildung 3. Die Hough-Transformation wird dann auf das so vorverarbeitete Bild angewandt. Im Mittelpunkt der Felge entsteht im Akkumulator ein ausgeprägter heller Fleck.

### Vorteile

Offenkundig kann dieses Verfahren sehr viel Rechenzeit in Anspruch nehmen, es hat jedoch einige wesentliche Vorteile. Es funktioniert nicht nur für einen vollständigen Kreis, sondern auch für Kreise, bei denen Teile fehlen, sowohl in Form von abgedeckten oder ausgeschnittenen Sektoren als auch bei unterbrochenen Kreislinien. Das Überdeckungsmaß wird dann allerdings einen kleineren Wert annehmen als bei einem vollständigen Kreis, so dass der Abstand zwischen dem hellsten

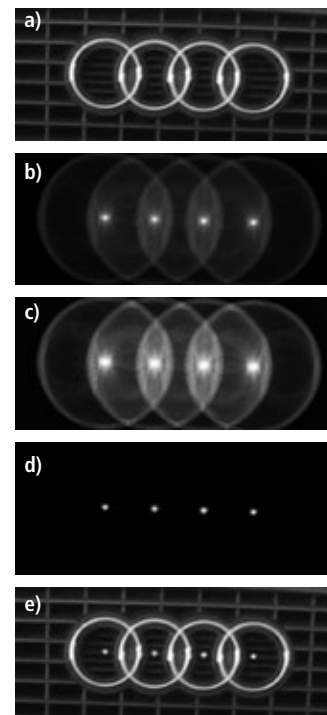


Abb. 4: Hough-Transformation für überlagerte Ringe  
 a) Quellbild; b) und c) Akkumulator-Bilder nach Hough-Transformation, Rohbild und kontrastverstärkt; d) Akkumulator nach der Binarisierung; e) Überlagerung von a) und d)

Fleck und dem Untergrund im Akkumulator absinkt und das Maximum irgendwann nicht mehr verlässlich gefunden werden kann, wenn die Zahl der Pixel auf der Kreislinie zu gering wird. Ein weiterer Vorteil ist, dass die gesuchten Kreise nicht als separate Objekte voneinander getrennt vorliegen müssen wie bei der Blob-Analyse. Zwei Kreise, die sich berühren oder teilweise überlagert sind, können über die Hough-Transformation im Allgemeinen als zwei Kreise klassifiziert werden.

Abbildung 4 zeigt ein Beispiel für Kreise, die sich überlagern. Schon im Quellbild erscheinen diese Kreise auch ohne Vorverarbeitung als gut herausgearbeitete Kreislinien. Auch hier ist das Akkumulatorarray zusätzlich in einer zweiten, kontrastgehobenen Version dargestellt. Als weiteres Beispiel wurden für Abbildung 5 mehrere Münzen aufgenommen. Die meisten dieser Münzen erscheinen als vollständige, voneinander getrennte kreisförmige Objekte im Bild. Zwei Münzen berühren sich jedoch, und bei einer Gruppe von drei Münzen überdecken sich die Kreise teilweise. Die Hough-Transformation wurde auf die Binärkonturen dieser Objekte angewandt und ermöglicht die Lokalisierung aller Kreise als helle Flecken im Akkumulator-Bild.

Es gibt viele weitere Beispiele, die die Leistungsfähigkeit dieser Methode belegen. In unkontrollierten Szenen können jedoch störende Strukturen auftreten, z.B. bei der Suche nach kreisförmigen Verkehrszeichen in Bildern von belebten Innenstadtstraßen, und selbst auf einem leistungsfähigen PC kann eine Hough-Transformation bei manchen Bildern mehrere Sekunden Rechenzeit benötigen.

### Die Hough-Transformation für Geraden

Während die Grundidee der Hough-Transformation für Kreise und die Konstruktion des Akkumulators intuitiv verständlich

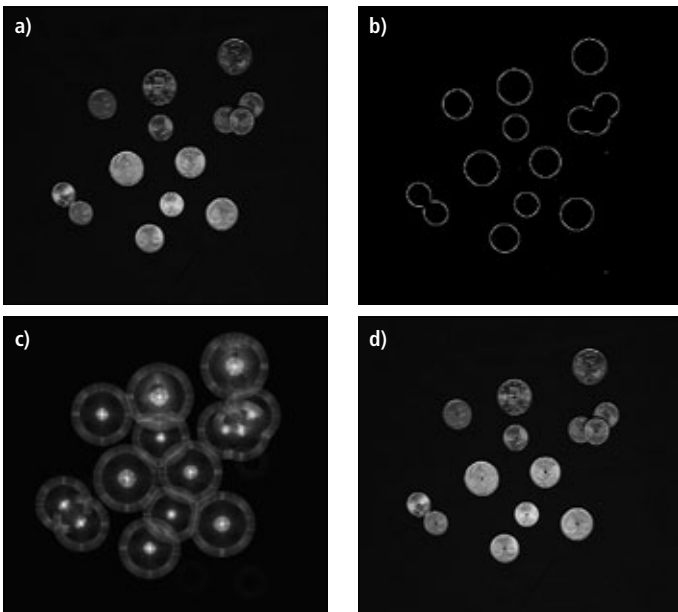


Abb. 5: Hough-Transformation für teilweise abgedeckte Münzen  
a) Quellbild; b) Konturen sind herausgearbeitet und verdünnt; c) Kontrastverstärktes Akkumulator-Bild nach Anwendung der Hough-Transformation auf das verdünnte Konturbild; d) Quellbild, überlagert mit den schwarzen Flecken im Zentrum der Münzen, die aus der Binarisierung des Akkumulators entstanden sind

sind und sogar unmittelbar implementiert werden können, erfordert die Hough-Transformation für Geraden einige mathematische Überlegungen. Analog zu der Suche nach Kreisen muss auch in diesem Fall die gesuchte Gerade nicht vollständig im Bild vorhanden sein. Die Hough-Transformation prüft vielmehr für die Menge der Vordergrundpixel in einem Bild, welche dieser Punkte kollinear sind, also auf einer gemeinsamen Geraden durch diese Punkte liegen. Wie bei der Suche nach Kreisen könnte man durch alle Bildpunkte, gleichgültig ob im Vordergrund oder im Hintergrund, sämtliche möglichen Geraden einzeichnen und die Vordergrundpixel zählen, die auf jeder dieser Geraden liegen. Auf diese Weise entsteht ein Gütemaß, mit dem die gewünschte Gerade von anderen Kandidaten getrennt werden kann. Offenkundig wäre das sehr aufwendig. Stattdessen orientiert sich ein effizienteres Verfahren wieder an den Vordergrundpixeln. Wir nehmen zunächst an, dass über die gesuchte Gerade bereits etwas bekannt ist, z.B. ihre Steigung. Dann könnte man durch jedes schwarze Vordergrundpixel eine Gerade mit dieser Steigung zeichnen, so dass eine Schar von parallelen Geraden entsteht. Wenn zwei oder mehr Pixel kol-

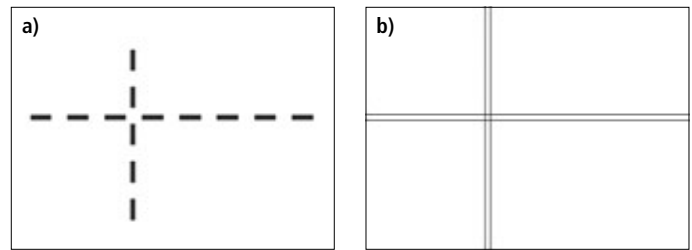


Abb. 6: Hough-Transformation für Geraden  
a) Quellbild; b) Geraden als Ergebnis der Hough-Transformation, angewendet auf die Konturen aus a

linear sind, fallen ihre zugehörigen Geraden aufeinander, und die Zahl der Überdeckungen ist identisch mit der Zahl der Pixel auf dieser speziellen Geraden. Allerdings ist nicht auf den ersten Blick klar, wie die Übereinstimmung zweier ganzer Linien in einem Bild überprüft werden kann. Aus der elementaren Geradenform in der x-y-Ebene,  $y=mx+b$ , ist jedoch unmittelbar ersichtlich, dass aus einer Schar paralleler Geraden mit der Steigung  $m$  diejenigen Geraden aufeinander fallen, die denselben Achsenabschnitt  $b$  haben. Um die Überdeckung zweier Geraden mit der Steigung  $m$  zu prüfen, müssen also lediglich die beiden zugehörigen Achsenabschnitte  $b_1$  und  $b_2$  berechnet und verglichen werden. Für einen bekannten Punkt  $(x, y)$  auf einer Geraden mit der Steigung  $m$  ist der Achsenabschnitt  $b=y-mx$ . Man kann also wie folgt vorgehen: Zunächst wird ein Akkumulator-Vektor angelegt, der so viele Elemente enthält wie Werte für den Achsenabschnitt  $b$  möglich sind. Der Vektor wird mit Null initialisiert. Das Bild wird nun zeilenweise von links oben nach rechts unten abgerastert. Wenn der Algorithmus auf ein schwarzes Pixel trifft, wird eine Gerade mit der Steigung  $m$  durch diesen Punkt gezeichnet und der Achsenabschnitt  $b$  berechnet. Das Element mit dem Index  $b$  im Akkumulatorvektor wird dann um 1 erhöht. Wenn diese Prozedur für alle Vordergrundpixel im Bild abgearbeitet ist, haben sich im Akkumulator die Überdeckungen zwischen den Geraden durch jedes Pixel durch Hochzählen angesammelt. Am Ende wird der hellste Fleck, also das Maximum im Akkumulator gesucht. Der zugehörige Index ist der Achsenabschnitt für die gemeinsame Gerade im Bild.

Wenn weder Achsenabschnitt noch Steigung bekannt sind, kann man den Ansatz verallgemeinern. Für jedes schwarze Pixel gibt es ein ganzes Bündel von Geraden mit verschiedenen Steigungen und verschiedenen Achsenabschnitten, die durch diese Pixel gezeichnet werden können. Steigung und Achsenabschnitt sind jedoch nicht willkürlich, sondern für ein Pixel  $(x, y)$  muss die Beziehung  $b=y-mx$  zwischen  $b$  und  $m$  für alle Geraden

Wir konzipieren, entwickeln und integrieren  
**3D- und 2D-Bildverarbeitungs-Systeme**  
zur Qualitätssicherung und Prozessoptimierung

- Oberflächenkontrolle
- Vermessung
- Sortieren und vergleichen
- Positionserkennung
- Anwesenheit, Vollständigkeit
- Klarschrift lesen
- Code lesen

**VISUELLE TECHNIK**  
INDUSTRIELLE BILDVERARBEITUNG

Visuelle Technik GmbH  
Industrielle Bildverarbeitung  
Ippichen 3 - 77709 Wolfach  
Tel.: +49 (0)7834 / 8670-0  
www.visuelle-technik.de

**Besuchen Sie uns auf der Vision 2012!**  
**Unser Team löst auch Ihre  
Bildverarbeitungsaufgabe!**

erfüllt sein, die durch dieses Pixel verlaufen. Wenn man nun alle Werte für  $m$  und  $b$ , die für ein bestimmtes Pixel  $(x_1, y_1)$  möglich sind, in einem Diagramm von  $b$  gegen  $m$  aufträgt, also in einer  $m$ - $b$ -Ebene, entsteht eine Gerade mit der Geradengleichung  $b = -x_1 m + y_1$ .  $-x_1$  ist die Steigung und  $y_1$  ist der Achsenabschnitt der Geraden im  $m$ - $b$ -Raum. Ein einzelnes Pixel  $(x_1, y_1)$  wird so in eine ganze Gerade im  $m$ - $b$ -Raum transformiert. Ein zweites einzelnes Pixel  $(x_2, y_2)$  wird durch eine andere Gerade im  $m$ - $b$ -Raum repräsentiert. Diese beiden Geraden werden sich im  $m$ - $b$ -Raum an einem bestimmten Punkt schneiden, an dem sie dann gemeinsame Werte für  $m$  und  $b$  haben. Dieser Schnittpunkt charakterisiert umgekehrt die Gerade im  $x$ - $y$ -Raum, die die beiden Pixel  $(x_1, y_1)$  und  $(x_2, y_2)$  verbindet. Offenbar ist die  $m$ - $b$ -Ebene ein Akkumulator-Bild, in dem man auf einfache Weise die Überdeckungen zwischen Geraden aufsummieren kann, die entstehen, wenn man die Büschel aller möglichen Geraden durch alle Vordergrundpixel im Bild zeichnet. Für jedes schwarze Vordergrundpixel  $(x, y)$  im Quellbild muss lediglich die zugehörige Gerade im  $m$ - $b$ -Raum gezeichnet werden. In jedem Pixel  $(m, b)$  auf dieser Geraden wird der Grauwert im Akkumulator-Array um 1 hochgezählt. Am Ende repräsentiert der hellste Fleck im Akkumulator-Bild die Steigung  $m$  und den Achsenabschnitt  $b$  der gesuchten Geraden. Mit diesen Daten kann man ins Quellbild zurückkehren und die vollständige Gerade einzeichnen.

In der Praxis wird die Hough-Transformation für Geraden nicht mit den Parametern Steigung und Achsenabschnitt implementiert, sondern über die Richtung und den senkrechten Abstand vom Ursprungspunkt entsprechend der sog. Normal- oder Hesse-Form der Geradengleichung. Dadurch werden die Komplikationen vermieden, die ansonsten bei Geraden parallel zur  $y$ -Achse auftreten. In diesem Parameter-Raum wird ein Punkt nicht mehr in eine Gerade, sondern in eine Sinuskurve transformiert. Kollinearität von Punkten aus dem Quellbild äußert sich jedoch immer noch als Schnittpunkt der zugehörigen Kurven im Parameterraum. Die Implementierung der Hough-Transformation für Geraden ist nicht trivial. Es ist ratsam, auf Lehrbücher zurückzugreifen [1] [2] und sich mit den verschiedenen weiteren Aspekten zu befassen, die weit über den Rahmen dieses einführenden Artikels hinausgehen. Abbildung 6 zeigt ein einfaches Beispiel für die Suche nach Geraden in einem Binärbild auf der Basis der Hough-Transformation. Die Zwischenschritte sind nicht dargestellt, lediglich die gefundenen Geraden in der  $x$ - $y$ -Ebene sind eingezeichnet.

## Weitergehende Verallgemeinerung

Die Grundidee der Hough-Transformation ist die Konstruktion eines Akkumulators im Parameter-Raum. Geraden werden durch zwei Parameter charakterisiert und führen auf einen zweidimensionalen Raum. Andere geometrische Figuren wie z.B. Ellipsen müssen analytisch durch mehr als zwei Parameter beschrieben werden. Der Akkumulator-Raum wird dann komplexer, und die Rechenleistung, die für eine sinnvolle Implementierung eines solchen Verfahrens erforderlich ist, mag vielleicht aktuell nicht realisierbar sein. An der Tragfähigkeit der Methode und dem eleganten Konzept der Hough-Transformation ändert das aber nichts. Hough hat seinen Ansatz im Jahr 1962 vorgeschlagen [1]. Damals konnte er sich wohl kaum vorstellen, welche Rechenleistung uns die aktuellen Prozessoren bieten. Etwa 20 Jahre später, im Jahr 1981, hat Ballard [3] eine Methode für die Hough-Transformation von beliebigen nicht-analytischen Formen publiziert. Heute haben wir enorme Rechenleistungen zu vernünftigen Kosten zur Verfügung, und die Hough-Transformation für Geraden und Kreise kann in manchen Anwendungen bei Standard-Video-Bildfrequenzen durchgeführt werden. Die Literatur ist vermutlich voll von solchen geschickten, wohl durchdachten Algorithmen aus den 60er, 70er und 80er Jahren des letzten Jahrhunderts. Sie schlummern in den Archiven und träumen von einer Implementierung in unseren schönen schnellen Bildverarbeitungssystemen.

## Literatur

- [1] R. C. Gonzales, R. E. Woods, Digital Image Processing, Addison-Wesley, 1993, p. 432-438
- [2] W. Burger, M. J. Burge, Digitale Bildverarbeitung, Springer, 2005, S.152-167
- [3] D. H. Ballard, Generalizing the Hough-transform to detect arbitrary shapes, Pattern Recognition Vol. 13, No. 2, p 111-122, 1981

## Danksagungen

Dank an unseren Diplom-Studenten Sven Schneider, der die Quellbilder für die Abbildungen zur Hough-Transformation für Kreise erstellt hat, und Dank an meinen Kollegen Stephan Nesper für den Hinweis auf die Relevanz der Literaturstelle [3].

## PRODUKTE

### Zeilenkamera auf dem iPad



Chromasens hat eine neue WebApp für das iPad entwickelt, die die neue Farbzeilenkamera in Bild und Animation vorstellt. Auf der Homepage von Chromasens befindet sich ein Video-Trailer, der die Möglichkeiten der AllPixa Smart WebApp erläutert. Dazu gehören eine 360°-Darstellung der Kamera und Bilderstrecken mit versteckten Inhalten, die der Anwender durch Drücken so genannter ‚Hotspots‘ aktiviert. Die Chromasens AllPixa Smart WebApp feierte auf der Fachmesse Laser World of Photonics China in Shanghai Premiere. Sie wurde auf iPads installiert und am Messestand öffentlich zugänglich gemacht. An den iPad-Ständern konnten sich die Besucher interaktiv mit der AllPixa beschäftigen. „Es hat sich gezeigt, dass das Interesse der Besucher und die Auseinandersetzung mit unserer neuesten Kamera durch die App positiv beeinflusst wurde“, so Markus Schnitzlein, Geschäftsführer von Chromasens.

[www.chromasens.de/allpixa](http://www.chromasens.de/allpixa)

### Bildsensor mit 20 Megapixel

Cmosis stellt einen 20-Megapixel- Bildsensor vor. Der neue Bildsensor CMV20000 ist für eine Auflösung von 20 Megapixeln (5.120 x 3.840) bei quadratischer Pixelstruktur mit 6,4 µm Kantenlänge ausgelegt. Dies resultiert in einer aktiven Sensorfläche von 32,8 x 24,6 mm² (entsprechend dem 35-mm Filmformat). Die Quanteneffizienz erreicht Spitzenwerte von mehr als 45%. Damit ergibt sich eine Empfindlichkeit von 8,29 V/Lux-s, in Kombination mit einem Dynamikbereich von 66 dB. Mittels Correlated Double Sampling (CDS) beim Betrieb mit Global Shutter unterdrückt die patentierte 8-Transistor Pixelarchitektur das Dunkelrauschen und die FPN-Uniformitätsabweichungen der Sensormatrix. Der neue Sensor bietet darüber hinaus eine hohe Shutter-Effizienz. Bei der vollen Auflösung von 20 Megapixeln und mit 12-bit Auflösung des A/D-Wandlers liefert der CMV20000 30 Full Frames pro Sekunde. Dies wird durch Verwendung von 16 LVDS-Ausgängen mit je 480 Mbit/s erzielt. Multiplexen auf nur acht Ausgangskanäle unterstützen niedrigere Frame-Raten. Partielles Auslesen, Windowing und Subsampling lassen sich für höhere Bildraten programmieren. Bei voller Auflösung und Frame-Rate beträgt der Leistungsverbrauch 1,1 W. Dieser Leistungsverbrauch lässt sich dynamisch steuern, wenn der Baustein bei niedrigeren Frame-Raten eingesetzt wird.

[www.cmosis.com](http://www.cmosis.com)



### Bildverarbeitung mit Sercos-Feldbus



Die neue Generation der Bildverarbeitungssysteme Pictor und Vicosys von Vision & Control ist ab sofort mit der Feldbusschnittstelle Sercos ausgerüstet. Besonders beim Einsatz in schnell und synchron laufenden Maschinen kommen die Vorteile der integrierten Feldbusschnittstelle Sercos zum Tragen. Die integrierte

Sercos-Funktionalität sorgt nicht nur eine einfache Einbindung der Bildverarbeitungssysteme in den Feldbus ohne zusätzliche Gateways. Sie bildet die Grundlage für eine zeitsynchrone und jitterfreie Triggerung des Bildverarbeitungssystems. Die Ergebnisse werden in Echtzeit mit kurzen Zykluszeiten (bis in  $\mu$ s-Bereich) über den Automatisierungsbuss an den Feldbus-Master übermittelt oder können auch in gleicher Weise zwischen Feldbus-Slaves ausgetauscht werden. Dies gewährleistet eine hohe Verarbeitungsgeschwindigkeit und Zeitsynchronisation beim Einsatz in schnell laufenden Maschinen und Prozessen. Wenn nötig können auch (Fehler-) Bilder im laufenden Betrieb über einen Nichtzeit-Kanal übertragen werden.

[www.vision-control.com](http://www.vision-control.com)

### Mikroskopielösungen mit Kameras



Omek Optics stellt Mikroskope her, die hochauflösende Bilder bei geringer Vergrößerung und einem besonders großflächigen Messfeld anzeigen können. So ist z. B. das Mic 401ET Mikroskop mit einer Tubuslinse ausgestattet, die bei geringer Brennweite Digitalkameras mit hochauflösenden Sensoren bis 1 Zoll unterstützt. Im Omek Mic 401ET Mikroskop ist mit einer AVT Manta G-504 CCD-Kamera verbaut. Gezeigt wurde die Kombination bereits auf der Optatec, wo ein CCD-Wafer bei zweifacher bzw. vierfacher Vergrößerung mit der 5-Megapixel Kamera geprüft wurde. Die

Manta G-504 ist mit einem 2/3" ICX655 CCD-Sensor von Sony ausgestattet mit einer Auflösung von 2.452 x 2.056 Pixeln. Dank der speziellen 80mm-Tubuslinse von Omek lässt sich die volle Auflösung des Sensors nutzen. Somit werden hochauflösende Bilder und eine hohe Detailschärfe möglich.

[www.alliedvisiontec.com](http://www.alliedvisiontec.com)

### Universeller Gehäusetypp vorgestellt



Das neue Lampenprogramm IL von Qioptiq besteht aus einem universellen Lampengehäusetypp, von dem je eine angepasste Version für Halogen-, Xenon- und Quecksilberdampflampen zur Verfügung steht. Unter Einbeziehung einer entsprechenden elektronischen Lampenversorgung der Serie PS bietet das Lampenprogramm IL ein hochwertiges, anspruchsvolles Lichtsystem für jede Applikation. Das universelle Lampengehäuse zeichnet sich durch seine Kompaktheit und die präzise und reproduzierbare Justierung der me-

chanischen Komponenten aus. Lampensockel, Spiegel und Kollektor sind hierbei von außen, über auf der Gehäuserückseite angeordnete Justierelemente, einstellbar. Als hervorragend stabil erweist sich die mechanische Lageposition im thermischen Betrieb. Durch Öffnungen an der Ober- und Unterseite des Gehäuses ist für die nötige Luftzirkulation, ohne jeglichen Lichtaustritt, gesorgt. Die geräuschlose und vibrationsfreie Konvektionskühlung bietet die Voraussetzung für einen stabilen Lichtaustritt der Lampen. Für die Aufnahme von Kondensoren dient eine Kollektorführung mit  $\varnothing$  60 mm. Darüber hinaus kann auch die Linos-Mikrobank direkt an das Lampengehäuse adaptiert werden.

[www.qioptiq.com](http://www.qioptiq.com)

[www.inspect-online.com](http://www.inspect-online.com)

### Sensoren jetzt mit Infrarotbeleuchtung



Baumer bietet für die Vision Sensoren der ID-, CS- und XF-Serie ab sofort auch Versionen mit integrierter (Nah-)Infrarotbeleuchtung an – mit einer Wellenlänge, die für den Menschen unsichtbar ist. Mitarbeiter der Produktion oder Maschinenbediener an den Fertigungslinien werden somit durch das typische „Blitzen“ der Beleuchtung bei einer visuellen Inspektion nicht mehr gestört. Ein integrierter Tageslichtsperrfilter sichert dabei eine stabile Applikation unabhängig vom Umgebungslicht. Verisens Vision-Sensoren sind in einem industrietauglichen Metallgehäuse mit Schutzart IP 67 integriert und erreichen hohe Taktraten mit bis zu 3.000 Auswertungen pro Minute. Die einheitliche Bediensoftware und einfache Integration der Verisens-Serien erlauben den flexiblen Einsatz bei sich ändernden Anforderungen und Applikationen.

[www.baumer.com](http://www.baumer.com)

### Objektiv mit telezentrischem Design

Das Computar Objektiv TEC-V7X - 5MP, 7x Zoom in telezentrischem Design vereint zwei Funktionen: Es kann als Macro Zoom Objektiv mit Vergrößerungen von 0,07x bis 0,5x und als telezentrisches Objektiv in einem Bereich von 0,25x bis 0,5x eingesetzt werden. Dieses Produkt bietet im gesamten Zoom-Bereich hervorragende Lichtdurchlässigkeit von mindestens 70 %, sogar in den Randbereichen des Bildes. Es deckt Arbeitsabstände von 182 bis 577mm ab und erlaubt eine präzise Einstellung von Tiefenschärfe und Kontrast dank seiner variablen manuellen Blende von F4.3 bis F32. Das TEC-V7X eignet sich für Kameras bis zu 5 MP Auflösung und Bildsensoren mit 1,1" optischem Format. Durch die einmalige Kombination spezifischer Leistungsmerkmale ist dieses Computar-Objektiv in der industriellen Bildverarbeitung äußerst vielseitig einsetzbar.

[www.framos.de](http://www.framos.de)

VISION  
VENTURES

---

Mergers &  
Acquisitions

in Machine Vision

M&A Intermediary • Strategy Advisory • Exit Planning

INTERNET  
E-MAIL

[www.vision-ventures.eu](http://www.vision-ventures.eu)  
[info@vision-ventures.eu](mailto:info@vision-ventures.eu)

# Stereo-Vision

Nicht mono, sondern stereo

Eine Kamera mit einem Standardobjektiv bildet eine dreidimensionale Szene in Zentralprojektion auf einen zweidimensionalen Detektorchip ab. Dabei geht die Tiefeninformation verloren.

Wird die Szene jedoch mit zwei Kameras aufgenommen, können die 3D-Koordinaten derjenigen Punkte im Raum ermittelt werden, die in beiden Bildern zu erkennen sind. Man spricht dann von „Stereo-Bildverarbeitung“, „Stereoskopie“ oder „Stereo-Vision“. Der Artikel behandelt die Grundzüge des Verfahrens.

## Zentralprojektion

Wenn ein Roboter Pralinen von einer Vorlagefläche aufpicken und im zugehörigen Nest einer Verpackung platzieren soll, muss ein BV-System, das diese Aufgabe unterstützt, die Koordinaten der fraglichen Praline und die Koordinaten der Zielmulde bestimmen. Für die Ansteuerung des Roboters sind dabei Absolutkoordinaten (sog. „Weltkoordinaten“) in Bezug auf einen festen Referenzpunkt erforderlich. Diese Aufgabe ist noch einigermaßen einfach zu lösen, wenn die gesuchten Punkte vollständig in einer Ebene liegen, die senkrecht auf der optischen Achse des Systems steht. Wenn jedoch eine Kamera mit einem Standard-Objektiv „schräg“ auf eine ebene Arbeitsfläche gerichtet wird, werden die Proportionen geometrischer Figuren perspektivisch verzerrt. Ein Beispiel dafür haben viele von Ihnen täglich vor sich, wenn sie über eine gerade verlaufende Landstraße fahren: Die in der Realität parallel verlaufenden Seitenlinien bilden in dem Bild, das wir sehen, zwei Geraden, die sich offenbar „irgendwo“ schneiden. Das hat im Übrigen nichts damit zu tun, dass wir üblicherweise mit zwei Augen sehen (wer es nicht glauben möchte, schließe bitte gelegentlich ein Auge: Die beiden Geraden erscheinen immer noch nicht parallel). Diese Verzerrung kommt dadurch zustande, dass sowohl das menschliche Auge als auch ein Standardobjektiv in Zentralprojektion arbeiten.

Eine Reihe von Problemstellungen lässt sich jedoch nicht auf die Vermessung in einer Ebene reduzieren. Beispielsweise muss ein Roboter, der an einer Autokarosserie Schweißpunkte setzt, alle drei Raumkoordinaten der Punkte erhalten. Ein Vision-System, das diesen Vorgang unterstützen soll, muss folglich die Koordinaten von Punkten im Raum vollständig liefern. Im Allgemeinen ist dies auf der Basis eines einzigen Bildes nicht möglich. Alle Punkte im Raum, die auf einer gemeinsamen Geraden durch das Projektionszentrum liegen, fallen im Bild auf denselben Punkt. Aus einem Bildpunkt können nur dann alle drei Raumkoordinaten des korrespondierenden Punktes im Objektraum berechnet werden, wenn mindestens eine der Koordinaten bekannt ist, bspw. die Entfernung zum Projektionszentrum. Wenn die Projektionen zweier Objekte auf die Bildebene gleich sind, kann weder die Entfernung noch die absolute Größe aus dem Bild berechnet werden. Die Bestimmung der Absolutkoordinaten erfordert die Bestimmung mindestens einer der unbekannt Koordinaten (z.B. der Entfernung oder „Tiefe“).

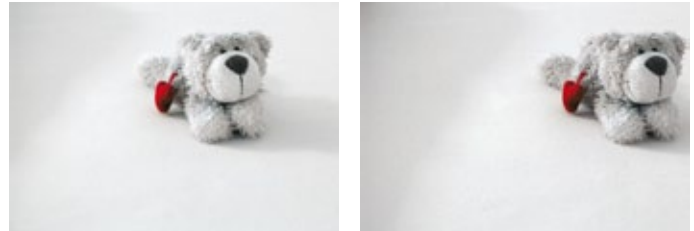


Abb. 1: Zwei Bilder derselben Szene, aufgenommen mit einem Stereo-System mit parallel ausgerichteten Kameras. Die Basislänge betrug 50 mm, die Brennweite des Objektivs 21 mm.

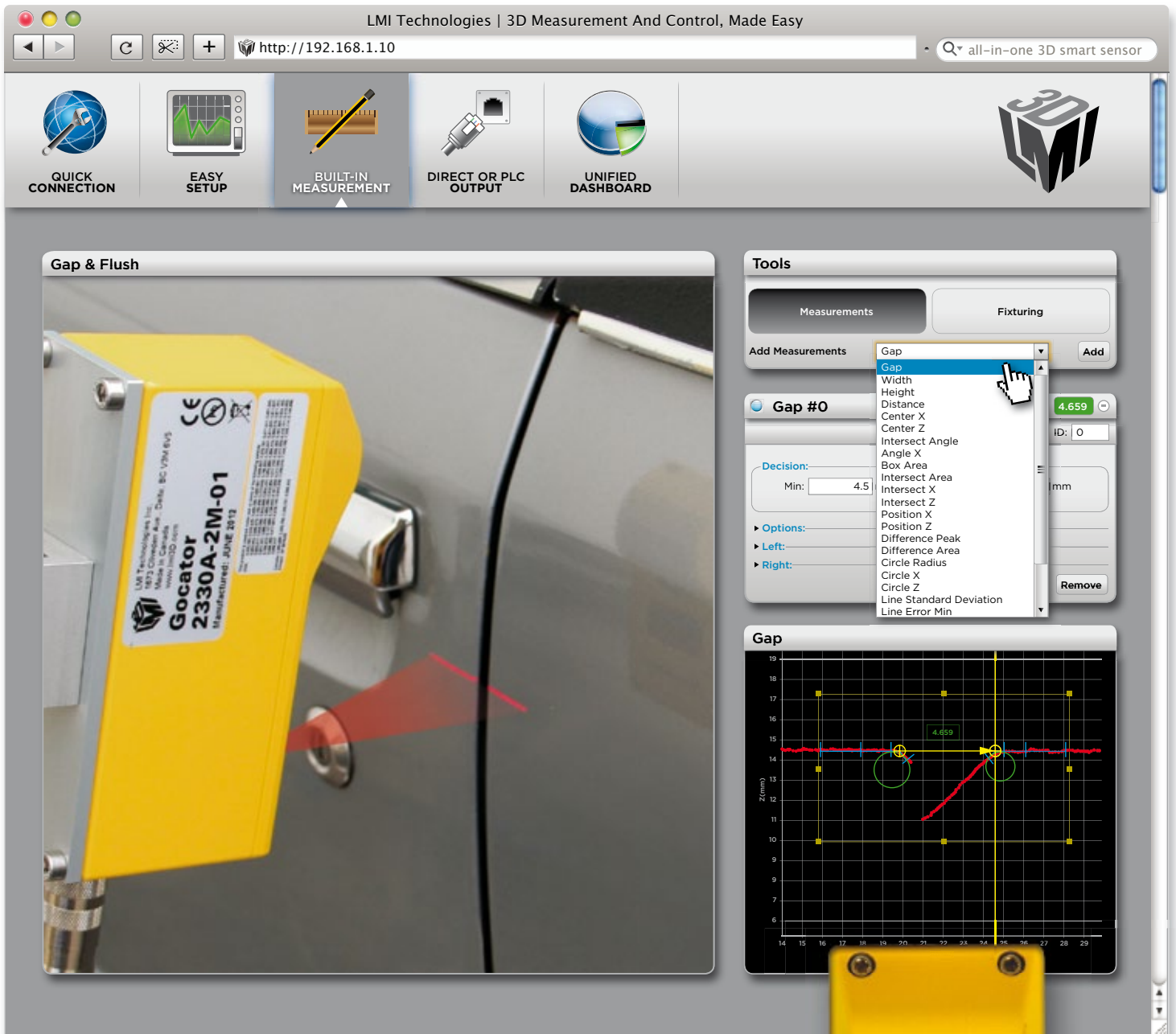
Ein etabliertes Verfahren, das dies leistet, ist die Stereoskopie. Dabei wird ein zweites Bild derselben Szene von einem anderen Punkt aus aufgenommen. Bei Kenntnis einiger Parameter des Aufnahmesystems können dann die Weltkoordinaten von prominenten Punkten, die in beiden Bildern zu sehen sind und gut gefunden werden können, errechnet werden. Alternativ können auch völlig andere Messverfahren ergänzend eingesetzt werden, bspw. Abstandsradar oder LIDAR. Stereoskopische Methoden sind jedoch etabliert und haben sich in der industriellen Praxis bewährt. Sobald die Weltkoordinaten bekannt sind, die zu Bildpunkten gehören, können im Prinzip auch 3D-Vermessungen von Objekten vorgenommen werden.

## Ein einfaches Stereo-System

Die einfachste stereoskopische Konfiguration besteht aus zwei identischen Kameras, deren optische Achsen parallel ausgerichtet sind. Der Abstand  $d_b$  der beiden optischen Achsen ist die sog. Basislänge. Eine weitere Einschränkung vereinfacht die Auswertung: Die beiden Kameras sollen so montiert sein, dass nicht nur die optischen Achsen parallel sind, sondern auch die Bildzeilen der Detektoren. Abbildung 1 zeigt Ausschnitte aus zwei Bildern derselben Szene, die auf diese Weise aufgenommen wurden. Die Basislänge betrug 50 mm, die Brennweite des Objektivs 21 mm. Wer die geänderte Perspektive zwischen den beiden Bildern sehen will, muss etwas genauer hinschauen. Wegen der diffusen Ränder des Stofftiers fallen die Unterschiede nicht sofort auf. Das Objekt wird jedoch in der linken Kamera erkennbar auf einen anderen Bereich des Detektors abgebildet als in der rechten Kamera. Diese Verschiebung enthält die Information über den Abstand des Objekts.

Liegt ein Objektpunkt  $P$  im Gegenstandsraum im Sichtbereich beider Kameras, z.B. die Nase des Stofftiers aus Abbildung 1, entsteht in beiden Kamerabildern ein Bildpunkt dieses Objektpunktes. Die Skizze in Abbildung 2 verdeutlicht die Verhältnisse. Eine Standardkamera, deren Objektiv in Zentralprojektion arbeitet, bildet die Punkte des Gegenstandsraums entsprechend den Beziehungen der geometrischen Optik auf die Empfängerebene ab. Wir verwenden hier drei Koordinatensysteme:

- Das **Weltkoordinatensystem**  $(X, Y, Z)$ , zur Verdeutlichung auch mit  $(X_w, Y_w, Z_w)$  bezeichnet, hat seinen Ursprung im op-



# Gocator®

## ALL-IN-ONE 3D SMART SENSOR

Wie die meisten Kamerasysteme ist Gocator schnell, zuverlässig, industrietauglich und SPS kompatibel. Im Gegensatz zu einfachen Kameras sieht Gocator die Sachen anders. Er scannt, misst, steuert in 3D und ermöglicht so eine neue Dimension der Industrieautomation und Qualitätssicherung. Das alles in einem kalibrierten Gerät.

Für weitere Informationen: [www.lmi3d.com/inspect](http://www.lmi3d.com/inspect)





tischen Hauptpunkt des Objektivs der linken Kamera. Dieser Hauptpunkt ist das Projektionszentrum. Die Z-Achse verläuft längs der optischen Achse dieser Kamera, und zwar in Richtung auf den Detektorchip. Die X- und Y-Achse haben dieselbe Orientierung wie die x- und y-Achse des Bildaufnehmers in der Kamera.

- Das **Sensorkoordinatensystem** ( $x_s, y_s$ ) hat seinen Ursprungspunkt im Durchstoßpunkt der optischen Achse mit der Ebene des Bildaufnehmers. Bei einer gut justierten Kamera liegt dieser Durchstoßpunkt im Mittelpunkt des Bildes, das als „Kamerabild“ auf dem Monitor des Rechners dargestellt wird. Die  $x_s$ - und  $y_s$ -Achse haben dieselbe Orientierung wie die x- und y-Achse des Bildaufnehmers in der Kamera.

- Das **Bildkoordinatensystem** ( $x, y$ ) oder **Pixelkoordinatensystem** ist das vertraute Koordinatensystem der Bildpixel, wie sie im Datenfile abgelegt sind. Es hat seinen Ursprungspunkt „links oben“, die x-Achse verläuft „nach rechts“, die y-Achse „nach unten“.

Sowohl die linke als auch die rechte Kamera haben je ein eigenes Bildkoordinatensystem und ein eigenes Sensorkoordinatensystem, in der Skizze mit den Indizes „L“ und „R“ gekennzeichnet. Dazu kommt das gemeinsame Weltkoordinatensystem. Jeder beliebige Punkt im Raum hat in diesem System eindeutige Weltkoordinaten. Damit die unten angegebenen Formeln stimmen, muss das Weltkoordinatensystem im optischen Hauptpunkt der in der Skizze als „links“ gekennzeichneten Kamera liegen, und die Koordinatenachsen müssen so verlaufen wie eingezeichnet. Grundsätzlich könnte man das Weltkoordinatensystem auch auf die rechte Kamera beziehen oder die Z-Achse aus der Kamera herauslaufen lassen. Dann ändern sich jedoch möglicherweise einige der unten angegebenen Formeln.

Die Koordinaten im Bildkoordinatensystem sind dimensionslos („... das sind einfach Zahlen ...“), die Welt- und Sensorkoordinaten tragen dagegen die Einheit Meter. Offenkundig kann man die Bildkoordinaten in Sensorkoordinaten umrechnen, wenn bekannt ist, wie viele Pixel der Detektor in x- und y-Richtung hat und welche Abmessungen ein Pixel hat. Im Folgenden wird angenommen, dass die optische Verzeichnung durch das Objektiv vernachlässigbar ist und die Übertragung der Daten in den Bildspeicher nicht zu Verzerrungen führt. Dann kann man ausrechnen, wie ein Punkt P im Gegenstandsraum auf einen Punkt mit den Sensorkoordinaten ( $x_s, y_s$ ) in der Bildebene **in der linken Kamera** abgebildet wird:

$$x_s = b (X_w / Z_w) \tag{1}$$

$$y_s = b (Y_w / Z_w)$$

Dabei ist b der Betrag der Bildweite, gerechnet vom Hauptpunkt des Objektivs aus. Die Bildweite kann aus den übrigen optischen Parametern des Systems berechnet werden. Für die Näherung dünner Linsen, die bei Objektiven für Industriekameras oft erfüllt ist, genügen dazu die Brennweite des Objektivs und die Gegenstandsweite. Insbesondere ist für weit entfernte Objekte b gleich der Brennweite f des Objektivs.

**Disparität**

Die obigen Gleichungen zeigen, dass aus einem Kamerabild allein nur das Verhältnis von je zwei Weltkoordinaten berechnet werden kann, nicht aber deren Absolutwerte. Eine davon unabhängige Information liefert die **Parallaxe** oder **Disparität**.

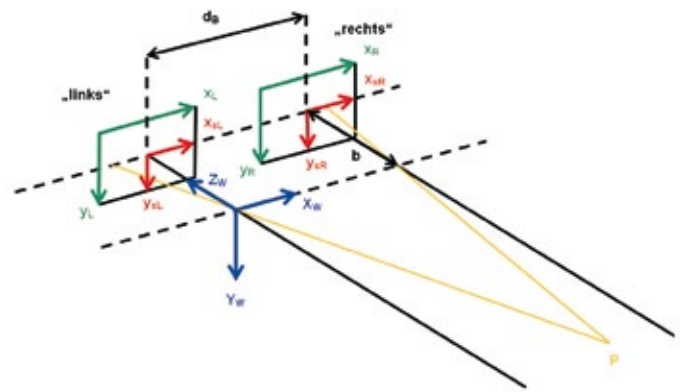


Abb. 2: Weltkoordinaten, Sensorkoordinaten und Bildkoordinaten bei der Abbildung eines Punktes P auf die beiden Detektoren eines Stereo-Systems mit parallel ausgerichteten Kameras.

Da der Punkt P von den beiden Kameras aus unterschiedlichen Richtungen anvisiert wird, haben die Punkte, auf die P in der linken und der rechten Kamera abgebildet wird, unterschiedliche Bildkoordinaten. Die Parallaxe ist dann die Differenz der x-Koordinaten dieser beiden korrespondierenden Bildpunkte im Sensorkoordinatensystem:

$$D = x_{sL} - x_{sR} \tag{2}$$

Aus den Abbildungsgleichungen für die beiden Kameras (oder auch direkt aus den geometrischen Verhältnissen) ergibt sich dann:

$$D = b d_B / Z_w \tag{3}$$

Folglich:

$$Z_w = (d_B b) / D \tag{4}$$

Damit ergibt sich  $Z_w$  aus Basisbreite, Parallaxe und Bildweite. Zusammen mit Gl. (1) ergibt sich dann:

$$X_w = (d_B x_{sL}) / D$$

$$Y_w = (d_B y_{sL}) / D \tag{5}$$

$$Z_w = (d_B b) / D$$

**Auswertung**

Für die Auswertung ist zu beachten, dass die Koordinaten  $x_s$  und  $y_s$  Absolutkoordinaten auf dem Detektorchip sind und nicht einfach die Pixelkoordinaten des Bildes im Rechner Speicher. Diese können aber aus den Pixelkoordinaten berechnet werden, wenn die absolute Kantenlänge eines Pixels bekannt ist. Dazu müssen die Pixelkoordinaten, deren Ursprungspunkt im linken oberen Eckpunkt des Bildes liegt, zunächst so umgerechnet werden, dass der Ursprungspunkt des neuen Pixelkoordinatensystems im Mittelpunkt des Bildes liegt. Von den ursprünglichen Koordinaten muss also lediglich die halbe Kantenlänge des Bildes in x- und y-Richtung subtrahiert werden, so dass nun auch negative Pixelkoordinaten im Bild vorkommen. Schon nach dieser einfachen Umrechnung können die Absolutwerte für  $X_w$  und  $Y_w$  berechnet werden, denn in den zugehörigen Formeln fällt die Absolutskalierung der Sensorkoordinaten durch die Division heraus. Lediglich die Basisbreite  $d_B$  vermittelt für diese beiden Weltkoordinaten die Anbindung an

die Absolutskalierung. Die Basisbreite muss also möglichst genau bekannt sein. Grundsätzlich kann man sie direkt messen, bei zwei identischen Kameras sogar anhand der Gehäusekanten, wenn sichergestellt ist, dass die Detektorchips genau genug auf das Gehäuse justiert sind. Die Absolutwerte für  $X_w$  und  $Y_w$  sind sogar unabhängig von der Bildweite  $b$ , so dass tatsächlich nur die Basisbreite für die Bestimmung dieser beiden Weltkoordinaten bekannt sein muss.

Erst die Berechnung von  $Z_w$  erfordert die Kenntnis der Absolutwerte für die Bildweite  $b$  und für die Disparität  $D$ . Bei weit entfernten Objekten kann für  $b$  die Brennweite  $f$  des Objektivs verwendet werden. Die Absolutanbindung für die Sensorkoordinaten erfolgt im Wesentlichen über die Pixelkantenlänge auf dem Detektorchip, die meist aus dem Datenblatt der Kamera entnommen werden kann. Bei den Bildern in Abbildung 1 handelt es sich um einen Chip mit einer Pixelkantenlänge von  $3 \mu\text{m}$ . Im linken Bild liegt die Nase des Stoffiers bei den Koordinaten (1385,376), im rechten Bild bei (1802,376). Die Disparität beträgt also 417 Pixel im Bildkoordinatensystem, entsprechend  $D = 1250 \mu\text{m} = 1,25 \text{ mm}$  im Sensorkoordinatensystem. Mit  $d_B = 50 \text{ mm}$  und der Näherung  $b = f = 21 \text{ mm}$  ergibt sich für den Abstand zum Objekt  $Z_w = 84 \text{ cm}$ . Sind die Basislänge und die Bildweite mit Ungenauigkeiten von je ca. 1 mm behaftet, führt dies zu einer Ungenauigkeit von ca. 5 cm für  $Z_w$ . Für manche Zwecke ist dies bereits ausreichend.

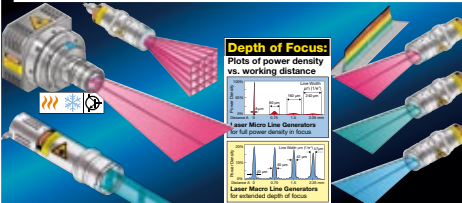
## Kamerakalibrierung

Bei höheren Anforderungen an die Genauigkeit werden die Kameraparameter, die in die Gleichungen eingehen, nicht direkt gemessen. Stattdessen werden Kalibriervorlagen mit Eichpunkten aufgenommen, deren Weltkoordinaten bekannt sind. Dazu werden die o. g. Gleichungen so umgestellt, dass die Kameraparameter aus den Bildkoordinaten und den bekannten Weltkoordinaten berechnet werden können. Anschließend werden die Weltkoordinaten unbekannter Punkte mit den so ermittelten Parametern berechnet. Diese Methode wird als Kamerakalibrierung bezeichnet. Zusätzlich werden dabei zunächst die optischen Verzeichnungen aufgrund der Abbildungsfehler des Objektivs korrigiert.

In diesem Artikel wurden die Grundlagen der Stereoskopie anhand einer einfachen Konfiguration erläutert. Da das Verfahren nur verwendet werden kann, wenn der interessierende Raumpunkt von beiden Kameras erfasst wird, muss der Überlapp der Sichtbereiche möglichst groß sein. Häufig werden daher die optischen Achsen der beiden Kameras gegeneinander geneigt („Binokularsystem“) oder mehr als zwei Kameras verwendet. Die Mathematik wird dann komplexer, das Grundprinzip der Stereoskopie bleibt jedoch dasselbe.

### Laser Line, Micro Focus, Laser Pattern Generators


Wavelength 405 - 2050 nm Made in Germany



**Application** **Laser Line Generators**

13LRM25S250 - 1.5 + 40TE - 640 - 500 - M33 - T12 - C - 6


**3D Profiling and Process Control**





**High power laser lines for laser light sectioning**

- Laser power up to 500 mW at 640 nm
- Constant width and uniform intensity distribution
- Digital interface
- Micro line versions for maximum power density in focus
- Macro line versions for extended depth of focus

Applications: 5 of 1000s







**VISION 2012**

Visit us at **Vision 2012**  
Hall 1.0, Booth 1.A.02  
November 6. - 8. 2012  
**STUTTGART**

### Line Scan Cameras

for Research and Machine Vision.  
Color, monochrome, or **TDI** sensors.  
[www.SuKHamburg.de/linescan](http://www.SuKHamburg.de/linescan) Made in Germany

from 512 to 12000 pixels



**Modular interface concept:**  
Analog: RS422  
Digital: LVDS  
Camera Link  
USB 2.0  
GigE  
Gigaset  
ETHERNET

Applications: 5 of 1000s

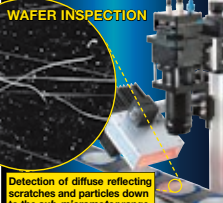






**TDI Line Scan Camera with Dark-field Illumination**

**WAFER INSPECTION**



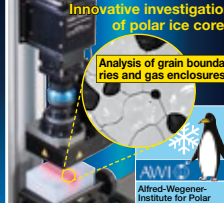
Detection of diffuse reflecting scratches and particles down to the sub-micrometer range.

**Applications**

**LASM - Large Area Scan Macroscopie**

Innovative investigation of polar ice cores


Analysis of grain boundaries and gas enclosures





AWI - Alfred-Wegener-Institute for Polar and Marine Research

**Application** **Line Scan Camera with integrated bright field illumination**

Turn-key System



**Filiform Corrosion Bavarian: Wurmfrass**

Schäfter+Kirchhoff develop and manufacture laser sources, line scan camera systems and fiber optic products for worldwide distribution and use.

# Das PMD-Prinzip

## Ein magisches Auge

Insbesondere in der Robotik und bei der Bildverarbeitung im Außenraum gibt es einen wachsenden Bedarf an Bildverarbeitungssystemen, die auch die Tiefeninformation für eine Szene liefern können. Mittlerweile sind von mehreren Herstellern sog. PMD-Kameras verfügbar, die einen wesentlichen Teil der Signalverarbeitung unmittelbar auf dem speziellen Detektorchip durchführen und ein orts aufgelöstes Abstandsbild erzeugen. Dazu muss die Helligkeit der Beleuchtung getaktet bzw. moduliert werden. Für viele Anwendungen sind diese Systeme bereits jetzt technisch geeignet und hinreichend kostengünstig, und eine weitergehende Optimierung ist in den nächsten Jahren zu erwarten. In diesem Artikel wird das Funktionsprinzip der PMD-Kameras erläutert.

Bei einer konventionellen Kamera wird eine Szene durch das Objektiv in Zentralprojektion auf einen ebenen Detektorchip abgebildet. Dabei geht die Tiefeninformation verloren. In der industriellen Anwendung der Bildverarbeitung dominieren daher Applikationen, bei denen Vorkenntnis über die Tiefe der Szene genutzt werden kann, z.B. weil alle Objekte auf einem Förderband liegen, dessen Abstand von der Kamera sich nicht wesentlich ändert. Klassische Stereo-Systeme nehmen mindestens zwei Bilder derselben Szene aus verschiedenen Positionen auf und suchen diejenigen Bildpunkte, die zum selben Punkt im realen Raum gehören. Nach einer geometrischen Kalibrierung kann man dann die 3D-Koordinaten dieses Punktes bestimmen. Damit das Verfahren überhaupt funktioniert, muss die Szene genügend Struktur aufweisen. Für eine weiße Wand wird es kaum gelingen, korrespondierende Bildpunkte zu berechnen – es sei denn, es wird eine Markierung aufgebracht. Eine elegante Lösung ist ein Laserstrahl, der einen hellen Lichtpunkt erzeugt. Mit einem Scanner kann der Strahl durch die gesamte Szene geführt werden, so dass grundsätzlich für jeden Bildpunkt aus dem gemeinsamen Sichtbereich der beiden Kameras eine Entfernung berechnet werden kann. Das Verfahren ist jedoch langsam und rechenaufwendig und daher für viele industrielle Anwendungen ungeeignet. Das Korrespondenzproblem kann man auch durch Projektion von definierten Mustern, z.B. Liniengittern, aus definierten Richtungen umgehen. Diese Methode wird als Lichtschnittverfahren bezeichnet, kommt mit einer Kamera aus und wird eingesetzt, wenn die Tiefeninformation auf wenigen Linien der Szene bereits ausreicht.

### Laufzeitmessung von Lichtimpulsen

Die Alternative sind aktive optische Sensoren auf der Basis der Messung der Laufzeit eines Lichtimpulses (LIDAR). Die Idee ist sehr einfach: Eine Lichtquelle sendet einen Lichtstrahl kurzer Dauer in Richtung eines Objektes, dessen Abstand bestimmt werden soll, und ein Photodetektor registriert das „Echo“, also

einen Teil der vom Objekt rückgestreuten Strahlung. Die Zeitspanne  $\Delta t$  zwischen dem Absetzen des Impulses und dem Eintreffen der remittierten Strahlung ergibt die Laufstrecke  $s = c\Delta t$ , folglich mit  $d = s/2$  den gesuchten Abstand zum Objekt. Bei industriellen Anwendungen sind Entfernungen im Bereich von wenigen Zentimetern bis zu mehreren Metern typisch, die meist mit Genauigkeiten von einigen Prozent gemessen werden sollen. Da das Licht in einer Nanosekunde eine Strecke von 30 cm zurücklegt, hat man es bei diesem Konzept zwangsläufig mit Pikosekunden-Elektronik zu tun. Das ist auch in unseren modernen Zeiten keine Kleinigkeit. Die elektronischen Anforderungen werden jedoch deutlich angenehmer, wenn das Sendersignal moduliert wird. Die Laufzeit wird dann nicht direkt bestimmt, sondern auf die Messung einer Phasenverschiebung zwischen der Modulation der Beleuchtung und der Modulation des Empfängersignals zurückgeführt. Das entsprechende elektrotechnische Konzept ist seit langem bekannt. Es wird in der Nachrichtentechnik als „Demodulation“ oder „Mischung“ bezeichnet und auch beim „lock-in“-Verstärker eingesetzt, der in der optischen Messtechnik beim Nachweis stark verrauschter Signale verwendet werden kann. Kern dieser Verfahren ist ein phasempfindlicher Detektor (auch „Demodulator“, „Mischer“, „Mixer“ oder „Korrelator“). Seit etwa zehn Jahren wird daran gearbeitet, einen solchen Mischer in jedem einzelnen Pixel eines Flächendetektors für optische Strahlung unterzubringen [1–3]. Das Resultat, das nun allmählich in anwendungsfähigen Produkten verfügbar wird, sind die sog. Photo-Misch-Detektor-Kameras (PMD), zur Zeit mit Auflösungen von etwa 160 x 120 Pixeln [4,5].

### Modulierte Beleuchtung

Anstelle eines Lichtpulses wird bei den PMD-Kameras eine kontinuierliche, aber mit fester Frequenz modulierte Beleuchtung eingesetzt. Wenn das Licht die Entfernung  $d$  zum Sender und zurück durchlaufen hat, hat die Modulation des Senders in der Regel einen anderen Schwingungszustand erreicht, und die Modulation des zu diesem Zeitpunkt ausgesandten Lichts

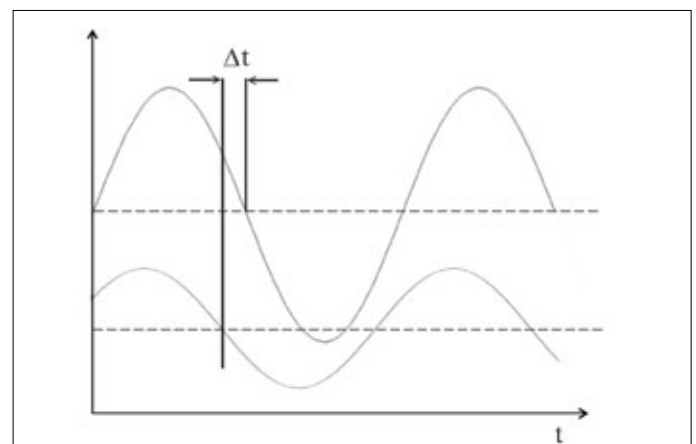


Abb. 1: Verlauf der modulierten Intensität der Beleuchtung (Sendersignal) und der gleichzeitig einlaufenden remittierten Strahlung (Empfängersignal) als Funktion der Zeit



$$\begin{aligned}
 \text{Sender:} & \quad \Phi(t) = \Phi_0(1 + \cos \omega t) \\
 \text{Empfänger:} & \quad I(t) = I_0(1 + \cos(\omega t - \Delta\varphi)) \\
 \text{Mischung:} & \quad \Phi(t)I(t) \\
 \\ 
 \Phi(t)I(t) &= \Phi_0 I_0 + \Phi_0 I_0 \cos \omega t + \Phi_0 I_0 \cos(\omega t - \Delta\varphi) + \Phi_0 I_0 \cos \omega t \cos(\omega t - \Delta\varphi) \\
 \\ 
 \cos \alpha \cos \beta &= \frac{1}{2}(\cos(\alpha - \beta) + \cos(\alpha + \beta)) \\
 \Rightarrow \\
 \Phi_0 I_0 \cos \omega t \cos(\omega t - \Delta\varphi) &= \frac{1}{2} \Phi_0 I_0 (\cos(\omega t - (\omega t - \Delta\varphi)) + \cos(\omega t + \omega t - \Delta\varphi)) = \\
 & \quad \frac{1}{2} \Phi_0 I_0 (\cos \Delta\varphi + \cos(2\omega t - \Delta\varphi)) \\
 \Rightarrow \\
 \Phi(t)I(t) &= \Phi_0 I_0 + \frac{1}{2} \Phi_0 I_0 \cos \Delta\varphi + \frac{1}{2} \Phi_0 I_0 \cos(2\omega t - \Delta\varphi) + \Phi_0 I_0 \cos \omega t + \Phi_0 I_0 \cos(\omega t - \Delta\varphi) \\
 \text{AC - Anteile mit } 2\omega t \text{ und } \omega t \\
 \text{DC - Anteil:} & \quad \Phi_0 I_0 + \frac{1}{2} \Phi_0 I_0 \cos \Delta\varphi = \Phi_0 I_0 (1 + \frac{1}{2} \cos \Delta\varphi)
 \end{aligned}$$

Abb. 2: Mischung des Sender- und Empfängersignals ergibt ein Signal mit einem DC-Anteil, der die Phasenverschiebung enthält

ist gegenüber der Modulation des zum selben Zeitpunkt empfangenen, rückgestreuten Lichts verschoben. Abbildung 1 zeigt schematisch den zeitlichen Verlauf des Sendesignals (oben) und des gleichzeitig aufgezeichneten Empfängersignals (unten). Die zeitliche Verschiebung  $\Delta t$  zwischen den beiden Modulationskurven ist die Laufzeit,  $\Delta t = 2d/c$ , entsprechend einer Phasenverschiebung  $\Delta\phi = 2\pi f \Delta t = 2\pi f \cdot 2d/c = (4\pi f/c)d$ . Die Entfernung  $d$  kann auf diese Weise nur für ein bestimmtes Intervall eindeutig aus der Phasenverschiebung ermittelt werden, nämlich für einen Abstandsbereich entsprechend der halben Wellenlänge der Modulation. Innerhalb dieses Bereichs ändert sich die Laufzeit vom Sender zur reflektierenden Fläche und zurück um weniger als eine Schwingungsdauer  $T$  der Modulation. Sobald die Laufzeitänderung gleich oder größer als  $T$  wird, kann ein Zustand mit Laufzeitänderung  $T + \Delta t$  nicht mehr vom Zustand mit Laufzeitänderung  $\Delta t$  unterschieden werden, denn die beiden zugehörigen Empfängerschwingungen sind dann in der Phase um  $2\pi$  gegeneinander verschoben und haben damit identische Zeitverläufe. Aktuell erhältliche PMD-Kameras arbeiten oft mit Modulationsfrequenzen von ca. 20 MHz. Ein kompletter Modulationszug hat dann eine Schwingungsdauer von  $T = 1/f = 50$  Nanosekunden, also eine Wellenlänge von  $\lambda = c/f = 15$  m. Der Bereich, innerhalb dessen eindeutige Abstandsmessungen möglich sind, ist unter diesen Bedingungen 7,5 m lang.

### Cleverer Detektorchip

„Mischen“ bedeutet im Jargon der Elektrotechnik die Multiplikation zweier Signale. Wenn zwei harmonische Signale derselben Frequenz miteinander multipliziert werden, entsteht ein Signal mit der doppelten Frequenz und einem zusätzlichen DC-Anteil proportional zur Phasendifferenz. Bei einer PMD-Kamera sind die Verhältnisse etwas komplizierter. Da es keine negative Lichtintensität gibt, ist der Sender nicht um die Nulllinie herum moduliert, sondern enthält einen Offset, der sich auch im Empfängersignal wieder findet. Die Umformungen in Abbildung 2 zeigen, dass in diesem Fall im DC-Anteil des Mischsignals neben dem Summanden, der proportional zur Phasendifferenz ist, auch das Produkt der Offsets der beiden Eingangssignale auftritt. Durch Integration bzw. Tiefpassfilterung wird der AC-Anteil eliminiert, und die gesuchte Phasenverschiebung  $\Delta\phi$  wird über den Gleichspannungsanteil des Signals messtechnisch zu-

gänglich. Dabei kann man im Prinzip beliebig lange integrieren und damit das Signal-Rausch-Verhältnis verbessern. Voraussetzung ist jedoch eine stabile Modulationsfrequenz. Es bleiben zwei Fragen: Wie kann man einen Multiplizierer in möglichst einfacher Form realisieren, und wie kann man den störenden Offset im DC-Anteil eliminieren? Die Lösung beider Probleme liegt in einer sehr cleveren Architektur eines Detektorchips, der in jedem einzelnen Pixel mit Hilfe des Prinzips der Ladungskopplung und der Ladungsverschiebung ein Mischsignal erzeugt und dieses zeitlich integriert [4]. Ein Detektorpixel besteht dabei aus zwei lichtempfindlichen Bereichen, die mit je einer lichtdurchlässigen Steuerelektrode versehen sind. Unterhalb dieser Elektrode entsteht wie bei einem CCD-Sensor eine Potentialmulde, in der die einfallenden Photonen in Elektronen umgesetzt werden. Die beiden lichtempfindlichen Bereiche liegen so nah zusammen, dass die beiden Potentialmulden bei entsprechender Steuerspannung miteinander verschmelzen und die entstehenden Elektronen darin frei verschoben werden können (Ladungskopplung). Die Steuerspannungen der beiden Elektroden werden nun im Gegentakt moduliert, so dass die entstehenden Ladungen „nach links“ oder „nach rechts“ verschoben werden, je nachdem, welche Steuerspannung „gewinnt“. „Links“ und „rechts“ werden die Ladungen jeweils in eine Speichertasche transferiert und dort während der Integrationszeit akkumuliert. Vereinfacht kann man sich eine rechteckförmige Modulation der Steuerelektroden vorstellen, entsprechend einem simplen Schaltimpuls. Wenn die Steuerspannung auf „high“ geschaltet ist, werden die entstehenden Ladungsträger nach links verschoben, beim Schaltzustand „low“ nach rechts. Ist die Intensität des einfallenden Lichts konstant, kommen auf diese Weise nach einem kompletten Schaltzyklus in der linken und in der rechten Speichertasche gleich viele Ladungsträger an. Wenn die Lichtintensität jedoch mit derselben Frequenz moduliert ist, mit der auch die Steuerelektroden getaktet werden, hängt die Zahl der links und rechts gesammelten Ladungsträger von der Phasenlage zwischen dem Lichtsignal und dem Steuersignal ab. Das Ergebnis ist das gewünschte Mischsignal, und die zeitliche Integration in den Speicherbereichen erledigt gleich auch noch die Unterdrückung der AC-Anteile. Die Ladungstrennung durch die Modulation der Steuerspannung sorgt außerdem dafür, dass gleich zwei unabhängige Signale entstehen, die etwas mit der Phasendifferenz zu tun haben. Die wesentlichen Teile der Signalverarbeitung sind bei diesem Konzept also bereits auf Chip-ebene im Prozess der Ladungstrennung, Ladungsverschiebung und Ladungsspeicherung implementiert, daher die Bezeichnung als photonischer Misch-Detektor (PMD). Diese Architektur ermöglicht nun problemlos die zeitliche Integration des Signals über mehrere Modulationszyklen. Bei dem oben genannten Beispiel mit einer Modulationsfrequenz von 20 MHz und einer Periodendauer von 50 Nanosekunden liefert eine Integrationszeit von einer Mikrosekunde bereits die Integration über 20 Schwingungsdauern. In vielen industriellen Anwendungen sind 10 oder 100 Mikrosekunden und länger als Belichtungszeit akzeptabel, so dass über viele Hundert Einzelmessungen gemittelt werden kann. Die Anforderungen an die elektronische Steuerung eines PMD-Detektors entsprechen demnach dem Stand, der aktuell bereits in der konventionellenameratechnik verwendet wird.

Bei dem oben genannten Beispiel mit einer Modulationsfrequenz von 20 MHz und einer Periodendauer von 50 Nanosekunden liefert eine Integrationszeit von einer Mikrosekunde bereits die Integration über 20 Schwingungsdauern. In vielen industriellen Anwendungen sind 10 oder 100 Mikrosekunden und länger als Belichtungszeit akzeptabel, so dass über viele Hundert Einzelmessungen gemittelt werden kann. Die Anforderungen an die elektronische Steuerung eines PMD-Detektors entsprechen demnach dem Stand, der aktuell bereits in der konventionellenameratechnik verwendet wird.

### Beeindruckender Stand der Technik

Eine cosinusförmige Modulation um einen Mittelwert herum kann in erster Näherung auch durch eine Rechteckfunk-

tion derselben Frequenz dargestellt werden (weil es sich dabei um die erste Komponente der Fourier-Zerlegung handelt). Die rechteckförmige Taktung der Spannung an der Steuerelektrode eines PMD-Detektors ergibt daher ein Mischsignal analog zu den Umformungen in Abbildung 2. Die Gegentaktmodulation führt dazu, dass im linken Kanal ein DC-Anteil proportional zu  $1+\Delta\phi/2$  entsteht, im rechten Kanal ein DC-Anteil proportional zu  $1-\Delta\phi/2$ . Die Differenz der beiden Signale liefert folglich ein Signal proportional zu  $\Delta\phi$ . Auch ohne Rückgriff auf diese Umformungen kann man sich das Resultat veranschaulichen, indem man direkt für Rechteckimpulse den zeitlichen Überlapp zwischen dem Zeitverlauf im ausgesandten Lichtimpuls und im remittierten Lichtimpuls betrachtet. Die Ladungstrennung im Detektorpixel sorgt dafür, dass während des Überlapps die entstehenden Ladungen nach links, in der verbleibenden Zeit nach rechts verschoben werden. Durch die Differenzbildung wird der Offset der Modulation eliminiert. Leider genügt zur Bestimmung der Phase der Differenzwert allein jedoch nicht, denn auch der Faktor  $\Phi_0 I_0$  geht in den Differenzwert ein. Die Amplitude  $I_0$  des Detektorstroms hängt z.B. vom Remissionsvermögen des Objekts ab, so dass auch bei gleicher Entfernung allein das Differenzsignal der beiden Kanäle eines Empfängerpixels noch nicht zur Bestimmung von  $\Delta\phi$  ausreicht. Nominell ergibt zwar die Summe der beiden Kanäle den gesuchten Proportionalitätsfaktor, und aus dem Quotienten würde der Faktor  $\Phi_0 I_0$  sogar ganz herausfallen. Das gilt aber nur, solange keine weiteren additiven Signalanteile vorkommen, die z.B. durch Umgebungslicht oder Dunkelstrom hervorgerufen werden können und kaum vermeidbar sind. Es müssen also weitere unabhängige Messgrößen aus dem Signal gewonnen werden. Eine Möglichkeit ist eine zweite, identische Detektorzelle in jedem Pixel, die wiederum aus zwei ladungsgekoppelten lichtempfindlichen Bereichen besteht, aber mit einem Modulationssignal angesteuert wird, das gegenüber der Modulation der Lichtquelle um  $90^\circ$ , also um  $\pi/2$  in der Phase verschoben ist. Dadurch entstehen zwei weitere, unabhängige DC-Anteile, die jeweils Information über die Phasendifferenz, die Modulationsamplitude und die Offsets enthalten. Aus den so gemessenen, insgesamt vier DC-Signalen des Mischsignals kann nun die Phasendifferenz  $\Delta\phi$  und damit der gesuchte Abstand  $d$  berechnet werden [1]. Ein beeindruckendes Beispiel für den mittlerweile erreichten Stand der Technik zeigt Abbildung 3 [6]. Im rechten Teil ist die Tiefeninformation aus einer PMD-Kamera dargestellt. Unterschiedliche Farben entsprechen unterschiedlichen Abständen.

Auch bei den PMD-Kameras spielen, wie bei anderen 3D-Messverfahren, im Detail eine Reihe von weiteren Effekten eine Rolle, z.B. die optische Verzeichnung, die Modulationsgüte bei der Beleuchtung oder unterschiedliche Dunkelströme von Pixel zu Pixel. Eine Kalibrierung des Detektorarrays für den genauen Zusammenhang zwischen gemessener Phasendifferenz und Entfernung kann daher notwendig und sinnvoll sein. Außer einer Tiefenkarte kann eine PMD-Kamera auch das Intensitätssignal einer klassischen Flächenkamera liefern, im Wesentlichen über die Summe der Signalkanäle eines Pixels (s. mittleres Teilbild in Abb. 3). Damit ist grundsätzlich die gesamte 3D-Information der Szene gleichzeitig aus einer Signalquelle zugänglich. Der wesentliche Rechenaufwand wird bereits auf Chipebene durch den photonischen Mischer hochgradig parallel erledigt, so dass die 3D-Daten mit sehr geringem zeitlichen Verzug nach der Bildaufnahme zur Verfügung stehen. Dadurch werden Applikationen möglich, die bisher wegen der langen Ausführungszeiten, der mechanischen Komplexität oder des apparativen Aufwands mit klassischen 3D-Systemen nicht durchführbar oder nicht wirtschaftlich waren. Wesentliche Impulse für die Weiterentwicklung der industriellen Bildverarbeitung sind in

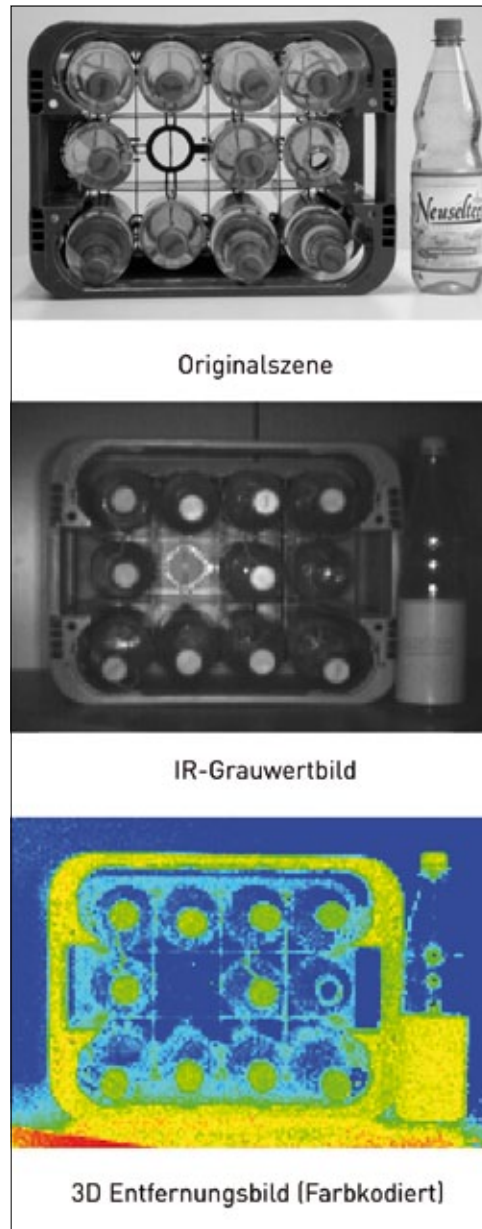


Abb. 3: Tiefenkarte (unten) und Intensitätsbild (Mitte) einer Szene, aufgenommen mit einer PMD-Kamera (mit freundlicher Genehmigung der PMDTechnologies GmbH, Siegen)

der Vergangenheit häufig durch Hardware-Innovationen ausgelöst worden. Auch die PMD-Technologie wird vermutlich in den nächsten Jahren in einigen Teilgebieten der Bildverarbeitung für Aufregung sorgen.

Literatur

- [1] T. Spirig, P. Seitz, O. Vietze, F. Heitger, IEEE Journal of Quantum Electronics 31, 1705 (1995)
- [2] R. Schwarte, Deutsche Patentanmeldung DE 197 04 496
- [3] R. Miyagawa, T. Kanade, IEEE Transactions on Electron Devices 44, 1648 (1997)
- [4] R. Schwarte, „PMD-Sensorik für Automotive – Möglichkeiten und Grenzen“, Optische Technologien in der Fahrzeugtechnik, Leonberg, 17./18.05.2006
- [5] siehe z. B. www.PMDTec.com; www.mesa-imaging.ch; www.canesta.com
- [6] freundlicherweise zur Verfügung gestellt von der Firma PMDTechnologies GmbH, Siegen

# Kameramodelle

## Von Kameraparametern zu Weltkoordinaten

Bei Anwendungen der Bildverarbeitung in der Robotik, bei Vermessungsaufgaben oder in der 3D-Bildverarbeitung müssen oft aus den Bilddaten die Koordinaten von Punkten im realen Raum, die sog. Weltkoordinaten, ermittelt werden. Dieser Rückschluss setzt ein geometrisches Kameramodell voraus. Im Unterschied zu radiometrischen Modellen wird dabei nicht der Signalpfad betrachtet, sondern lediglich beschrieben, wie Punkte aus dem Gegenstandsraum auf Pixelkoordinaten in der Bilddatei abgebildet werden.

### Zentralprojektion

Ein Bild einer Szene im dreidimensionalen Raum, das mit einem Standard-Objektiv aufgenommen wird, entsteht in Zentralprojektion, so dass die Proportionen geometrischer Figuren perspektivisch verzerrt erscheinen. Abbildung 1 zeigt ein Beispiel für diese Situation. Die Pflasterung verläuft in guter Näherung in einer Ebene. Die Kantenlänge der Steine ist in der Realität gewiss konstant, sie erscheint jedoch durch die perspektivische Verzerrung mit zunehmender Entfernung immer kleiner. Abbildung 2 verdeutlicht im oberen Teil den Zusammenhang im sog. Lochkameramodell der optischen Abbildung, das auch für Standardobjektive mit fester Brennweite in guter Näherung gilt. Die Objekte in der realen Welt befinden sich rechts von der Linse. Der Detektorchip links von der Linse ist in der Bildebene der Kamera fixiert. Zwei gleich große Objekte in verschiedenen Abständen vom Objektiv führen wegen der Zentralprojektion zu unterschiedlich großen Bildern. Nur für einen einzigen Objektstand entsteht ein optimales, scharfes Bild. In der Praxis kann man die Blendenzahl und die Brennweite jedoch oft so wählen, dass der Schärfentiefebereich hinreichend groß wird. Die quantitativen Zusammenhänge ergeben sich aus dem unteren Teil der Abbildung 2. Ein Punkt mit den Weltkoordinaten  $X_W$  und  $Z_W$  im realen Raum, bezogen auf das Projektionszentrum, wird auf dem Kamerasensor auf einen Punkt mit der Sensorkoordinate  $x_s$  im vertikalen Abstand  $b$  vom Projektionszentrum abgebildet. In Anlehnung an das Lochkameramodell wird  $b$  als Kammerkonstante bezeichnet. Offenbar ist

$$X_W/Z_W = x_s/b.$$

Dieser Zusammenhang gilt übrigens auch noch, wenn die Abbildung unscharf ist. Der Bildpunkt wird dann lediglich zu einem Scheibchen, seine Position entspricht in guter Näherung der Lage bei einer scharfen Abbildung. Im allgemeinen Fall kann ein Punkt in der Szene außerhalb der in Abbildung 2 gezeichneten Ebene liegen. Er hat dann zusätzlich eine von Null verschiedene Weltkoordinate  $Y_W$  senkrecht zur Zeichenebene, der zugehörige Bildpunkt hat eine von Null verschiedene Sensorkoordinate  $y_s$ . Die Zentralprojektion eines Punktes mit den Weltkoordinaten  $X_W, Y_W$  und  $Z_W$  auf den Bildpunkt mit den Sens-



Abb. 1: Perspektivische Verzerrung

orkoordinaten  $x_s$  und  $y_s$  wird daher durch die folgenden beiden Gleichungen beschrieben:

$$\begin{aligned} x_s &= -b (X_W/Z_W) \\ y_s &= -b (Y_W/Z_W) \end{aligned} \quad (1)$$

Alle Größen in den beiden Gleichungen tragen Maßeinheiten, z.B. Millimeter, auch die Sensorkoordinaten.  $x_s$  und  $y_s$  haben zunächst noch nichts mit den Pixelkoordinaten des Bildes in der Bilddatei zu tun, sondern sind Koordinaten in der realen Welt, allerdings in der Bildebene der Kamera. Die dimensionslosen Pixelkoordinaten der Bilddatei müssen über das bekannte Pixelraster des Sensors in die realen Koordinaten in der Sensorebene umgerechnet und dadurch an die realen Dimensionen angebunden werden.

### Externe und interne Parameter

Geometrische Kameramodelle verwenden zwei verschiedene Weltkoordinatensysteme. Eines der beiden Koordinatensysteme ist bereits in Abbildung 2 in Erscheinung getreten: das Weltkoordinatensystem  $X_W, Y_W, Z_W$  der Kamera. Die  $Z_W$ -Achse zeigt in Richtung der optischen Achse, und der Ursprungspunkt liegt im Projektionszentrum. Dieses Koordinatensystem ist fest mit der Kamera verknüpft und dreht bzw. verschiebt sich im Raum, wenn die Orientierung und die Position der Kamera (die sog. Pose) verändert werden. Zusätzlich arbeitet man meist noch mit einem zweiten Koordinatensystem, das z.B. an einer Arbeitsebene orientiert sein kann. Wir bezeichnen es hier mit  $X, Y, Z$ . Dieses Koordinatensystem kann durch zwei einfache Operationen in das Weltkoordinatensystem  $X_W, Y_W, Z_W$  der Kamera überführt werden. Die erste Operation ist eine Verschiebung des Koordinatenursprungs des Systems der Arbeitsebene in den Ur-



sprung des Kamera-Koordinatensystems. Der zugehörige Verschiebungsvektor wird durch die drei Koordinaten des Translationsvektors eindeutig beschrieben. Die zweite Operation ist eine Drehung des Koordinatensystems, so dass die Koordinatenachsen der beiden Systeme zusammenfallen. Zur Beschreibung dieser Drehung mit einer Drehmatrix sind drei weitere Parameter erforderlich. Diese sechs Größen werden als äußere Kameraparameter bezeichnet. Das Kameramodell enthält nun als Parameter die sechs Freiheitsgrade der äußeren Kamerarorientierung und als internen Parameter die Bildweite  $b$ .

Nun fehlt noch die Verknüpfung mit den dimensionslosen Pixelkoordinaten aus der Bilddatei. Traditionell liegt der Nullpunkt des Koordinatensystems im Bildfile in der linken oberen Ecke des Bildes, das man sieht, wenn man von der Kamera aus in den Objektraum schaut. Das Objektiv bildet diesen Punkt, nun vom Objektraum aus gesehen, auf die untere linke Ecke des Detektorchips ab. Abbildung 3 verdeutlicht die Zusammen-

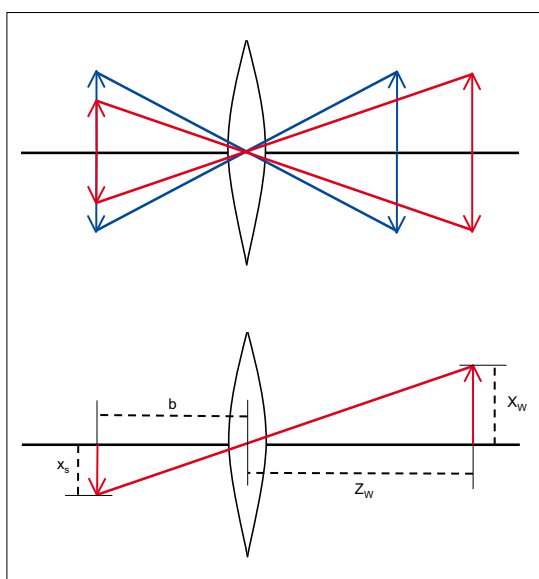


Abb. 2: Zentralprojektion im Lochkameramodell

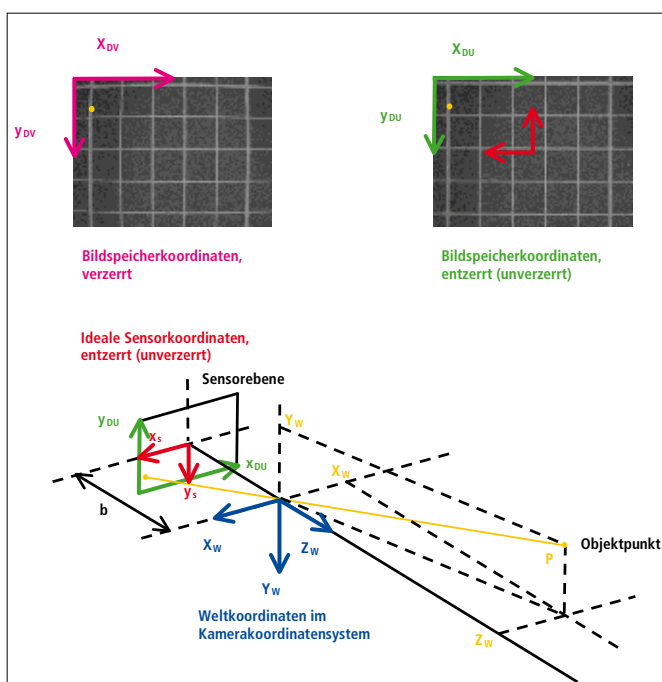


Abb. 3: Koordinatensysteme für die Transformation von Weltkoordinaten in Pixelkoordinaten in der Bilddatei sowie Einfluss der optischen Verzerrung

hänge. Der gelb gezeichnete Objektpunkt hat die Weltkoordinaten  $(X_W, Y_W, Z_W)$  im blau gezeichneten Weltkoordinatensystem der Kamera. Das Koordinatensystem der Arbeitsebene ist hier nicht eingezeichnet, damit die Darstellung übersichtlich bleibt. Der Objektpunkt wird auf einen Punkt in der Sensorebene abgebildet, der in der Bilddatei im Bereich der linken oberen Ecke erscheint. Zur Verdeutlichung ist im oberen Teil der Abbildung der Blick auf den Bildschirm dargestellt, auf dem die Bilddatei angezeigt wird. Im unteren Teilbild ist das grüne Koordinatensystem aus der Bilddatei in die Sensorebene eingeblendet. Zusätzlich zum Bilddaten-Koordinatensystem  $(x_D, y_D)$  wird nun das rot gezeichnete Sensor-Koordinatensystem  $(x_s, y_s)$  definiert. Der Ursprungspunkt dieses Systems liegt im Durchstoßpunkt der optischen Achse mit der Detektorebene, dem sog. Hauptpunkt. Die  $x$ - und  $y$ -Achse dieses Systems verlaufen parallel zur  $X_W$ - und  $Y_W$ -Achse des Weltkoordinatensystems. Die Koordinaten  $x_s$  und  $y_s$  tragen Einheiten, z.B. Millimeter oder Mikrometer. Für die Umrechnung der Pixelkoordinaten aus der Bilddatei in die Sensorkoordinaten muss zunächst der Ursprung des Koordinatensystems der Bilddatei in den Hauptpunkt des Sensors gelegt werden. Anschließend werden die so zentrierten Pixelkoordinaten mit der Pixelkantenlänge  $S_x$  und  $S_y$  in  $x$ - und  $y$ -Richtung multipliziert. Auf diese Weise entstehen skalierte Sensorkoordinaten. Die Skalierungsfaktoren entnimmt man meist aus dem Datenblatt des Kamerasensors. Bei Kameras mit digitaler Schnittstelle werden die Daten geometrietreu aus dem Sensor ausgelesen und in die Bilddatei übertragen, bei Analogkameras muss ggf. das Verhältnis zwischen der Pixeltaktfrequenz in der Kamera und der Abtastfrequenz des Frame Grabbers berücksichtigt werden. Die Pixelkoordinaten ergeben sich damit aus den Sensorkoordinaten über die folgende Beziehung:

$$x_D = -x_s/S_x + H_x$$

$$y_D = -y_s/S_y + H_y$$

(2)

$H_x$  und  $H_y$  sind dabei die Pixelkoordinaten des Hauptpunktes. In guter Näherung ist dies bei Kameras mit hinreichend gut justiertem Sensor einfach der Mittelpunkt des Detektorchips, so dass lediglich die Pixelzahl in  $x$ - und  $y$ -Richtung bekannt sein muss. Man kann die beiden Werte jedoch auch als freie Parameter im Modell weiter mitführen.

### Verzeichnungskorrektur

Im praktischen Einsatz ist häufig eine weitere Modifikation der Sensorkoordinaten erforderlich, die Verzeichnungskorrektur. Die in der industriellen Bildverarbeitung üblichen einfachen Objektiv weisen oft deutlich erkennbare Verzeichnungen auf. Ein Beispiel zeigt der obere linke Teil der Abbildung 3. Der Hauptanteil der Verzeichnung ist rotationssymmetrisch und hängt nur vom Abstand von der optischen Achse ab. Er kann mit einem einzigen Parameter beschrieben werden, z.B. in der Form:

$$x_{DU}^* = x_{DV}^*/(1+k r_V^2)$$

$$y_{DU}^* = y_{DV}^*/(1+k r_V^2)$$

(3)

mit  $r_V^2 = x_{DV}^{*2} + y_{DV}^{*2}$ . Der Parameter  $k$  beschreibt die Umrechnung von den verzerrten in die unverzerrten Pixelkoordinaten in der Bilddatei, jedoch unter Beachtung der Zentrierung der Koordinaten auf den Hauptpunkt, hier gekennzeichnet durch den Stern. Rechts oben in Abbildung 3 ist das Ergebnis der Entzerrung dargestellt. Die rot gezeichneten Sensorkoordinaten in

Abbildung 3 sind somit die idealen, unverzerrten Sensorkoordinaten, die bei einer Abbildung mit einer verzeichnungsfreien Linse in der Sensorebene entstehen würden. Das Kameramodell enthält nun fünf weitere interne Parameter.

### Schlussbemerkungen

Das hier vorgestellte Kameramodell mit insgesamt 12 Parametern geht auf die Arbeiten von Lenz [1] und Tsai [2] zurück, die bereits 1987 publiziert wurden. Die Bestimmung der Kenngrößen des Modells im Rahmen der Kamerakalibrierung aus Bildern von Vorlagen mit Referenzpunkten an genau bekannten Weltkoordinaten ist im Einzelnen mathematisch komplex. Der Zugang zu diesen Verfahren wird erleichtert, wenn die Abbildung mit den Methoden der linearen Algebra als Matrixoperation formuliert wird. Die entsprechende Vorgehensweise ist in der Literatur [3, 4] mittlerweile gut aufbereitet und beschrieben.

### Literatur

- [1] Lenz, R., Informatik-Fachberichte 149: Mustererkennung 1987, S. 212-216, Springer 1987
- [2] Tsai, R. Y., IEEE Journal of Robotics and Automation, Vol. RA-3 (4) 1987, 323-344
- [3] Steger, C., Ulrich, M., Wiedemann, C., Machine Vision Algorithms and Applications, Wiley-VCH, Weinheim 2008, p. 180 ff.
- [4] Azad, P., Gockel, T., Dillmann, R., Computer Vision, Elektor, Aachen 2007, S. 73 ff.

# ROMER

Jetzt bis zu  
23% genauer!



## Bewegungsfrei – der ROMER Absolute Arm mit integriertem 3D Laser Scanner.

Der ROMER Absolute Arm mit integriertem Laser Scanner ist ein portables Koordinatenmessgerät für eine Vielzahl von Anwendungen wie Inspektion von Bauteilen, Reverse Engineering, 3D-Modellierung oder Rapid Prototyping.

Produktivität steigern und Qualität sichern mit portablen Messarmen von Hexagon Metrology.

Demo? Ihren Ansprechpartner vor Ort finden Sie unter

[www.hexagonmetrology.de](http://www.hexagonmetrology.de)



# Kamerakalibrierung

## Weltkoordinaten im Griff

Bei Anwendungen der Bildverarbeitung in der Robotik oder bei der 3D-Vermessung müssen oft aus den Bilddaten die Koordinaten von Punkten im realen Raum ermittelt werden. Die Umrechnung von Bildkoordinaten in Weltkoordinaten setzt ein quantitatives Modell der Abbildung voraus. Bei der geometrischen Kamerakalibrierung werden die Parameter, die das Abbildungsmodell charakterisieren, aus Bildern von Kalibriervorlagen mit bekannten Referenzpunkten ermittelt. Der Artikel beschreibt die Grundzüge des Verfahrens.

### Zentralprojektion

Wenn ein Roboter eine Praline von einer Vorlagefläche aufnehmen und im zugehörigen Nest einer Verpackung ablegen soll, muss ein Bildverarbeitungssystem, das diese Aufgabe unterstützt, die Koordinaten der Praline und der Zielmulde bestimmen. Für die Ansteuerung des Roboters sind Absolutkoordinaten (sog. Weltkoordinaten) in Bezug auf einen festen Referenzpunkt erforderlich. Wir gehen zunächst davon aus, dass die gesuchten Punkte sämtlich in einer Ebene liegen, wie in der Szene in Abbildung 1. Ein Bild dieser Szene, das mit einem Standard-Objektiv aufgenommen wird, entsteht in Zentralprojektion, so dass die Proportionen geometrischer Figuren perspektivisch verzerrt erscheinen. Die Skalierung im Bild kann nicht allein mit einem konstanten Faktor in die Skalierung auf der realen Ebene übertragen werden. Besonders deutlich wird das bei schrägem Blick auf die Arbeitsebene wie in Abbildung 2. Die Kantenlänge der Steine ist in der Realität konstant, sie erscheint jedoch durch die perspektivische Verzerrung mit zunehmender Entfernung immer kleiner.

### Koordinatensysteme

Die Abbildung eines Punktes aus dem Objektraum in die Bildebene der Kamera kann mit einer Matrixtransformation beschrieben werden, dem sog. Kameramodell. Die Bilddaten im Rechnerpeicher sind lediglich mit dimensionslosen Pixelkoordinaten  $x_D$  und  $y_D$  verknüpft. Über die bekannte Pixelkantenlänge des Sensors können sie in die realen, mit Einheiten behafteten Sensorkoordinaten  $x_S$  und  $y_S$  des Bildes in der Sensorebene umgerechnet werden. Die Zentralprojektion eines Objektpunktes auf die Sensorebene wird am einfachsten im Kamerakoordinatensystem beschrieben. In diesem System hat ein Objektpunkt in der Szene die Weltkoordinaten  $X_W$ ,  $Y_W$  und  $Z_W$ , und der Ursprungspunkt liegt im Projektionszentrum des Objektivs. In einem pick-and-place-Szenario möchte man die Weltkoordinaten eines Zielpunktes jedoch direkt auf die Arbeitsfläche beziehen. Die X- und Y-Achsen dieses Systems liegen in der Arbeitsebene, die Z-Achse steht senkrecht auf der Ebene. Die Z-Koordinate aller Punkte in dieser Ebene ist dann Null und damit für alle Punkte in der Ebene bekannt. Abbildung 3 zeigt die verschiedenen Koordinatensysteme. Die gesamte Transformation der

Weltkoordinaten  $(X, Y, Z)$  eines Objektpunktes in die Pixelkoordinaten  $(x_D, y_D)$  des korrespondierenden Punktes in der Bilddatei kann man mathematisch mit einer einzigen Matrixoperation formulieren. Dieses mathematische Kameramodell enthält eine Reihe von Parametern, u. a. die sechs Freiheitsgrade der Kameraorientierung im Raum, die Bildweite der optischen Abbildung und einen Parameter zur Beschreibung der optischen Verzeichnung durch das Objektiv. Sobald die Modellparameter empirisch bestimmt sind, kann man die Matrixtransformation umkehren und von den Pixelkoordinaten auf die Weltkoordinaten zurück schließen. Aus den Pixelkoordinaten können jedoch nicht die Absolutwerte  $X_W$ ,  $Y_W$  und  $Z_W$  berechnet werden sondern nur die Verhältnisse  $(X_W/Z_W)$  und  $(Y_W/Z_W)$ . Anschaulich liegt das daran, dass zwei Objektpunkte, die von der Kamera aus in derselben Richtung, aber in unterschiedlichen Entfernungen liegen, durch die Zentralprojektion auf denselben Bildpunkt abgebildet werden. Sobald aber eine der Weltkoordinaten absolut bekannt ist, können die beiden verbleibenden Werte bestimmt werden.

### Kalibrierung

Zur Beschreibung der Transformation des Weltkoordinatensystems in der Arbeitsebene in das Pixelkoordinatensystem der Bilddatei sind 12 oder mehr freie Parameter erforderlich, je nach Detaillierungsgrad des Modells. Grundsätzlich könnten diese Kenngrößen durch separate Messungen ermittelt werden, z. B. die Bildweite oder der Verzeichnungsparameter. In der Praxis hat sich jedoch die Kalibrierung des Systems mit Hilfe eines Testmusters etabliert. Dazu wird mit der Kamera, die kalibriert werden soll, ein Objekt aufgenommen, das Markierungen mit bekannten Weltkoordinaten im Koordinatensystem des Objektraums enthält. Die Positionen der zugehörigen Bildpunkte werden mit den Methoden der Bildverarbeitung aus dem Bild entnommen, so dass sowohl die Weltkoordinaten der Markierungen als auch deren korrespondierende Pixelkoordinaten bekannt sind. Das Abbildungsmodell liefert die Verknüpfung zwischen den beiden Datensätzen. In den entsprechenden Gleichungen sind die Modellparameter nun die Unbekannten. Wenn eine hinreichende Zahl von Markierungen erfasst wird, entsteht ein Gleichungssystem, aus dem die 12 unbekannt Parameter berechnet werden können.

Wir verwenden in unserem Labor für das pick-and-place-Szenario aus Abbildung 1 eine ebene Kalibriervorlage, die in der Arbeitsebene platziert wird. Im Rahmen einer Diplomarbeit wurde ein geeignetes Testmuster entwickelt (Abb. 4). Es besteht aus kreisförmigen Objekten mit gutem Kontrast zum Untergrund, die in regelmäßigen, präzise eingehaltenen und in ab-



Abb. 1: Pick and Place in einer Ebene





Abb. 2: Perspektivische Verzerrung

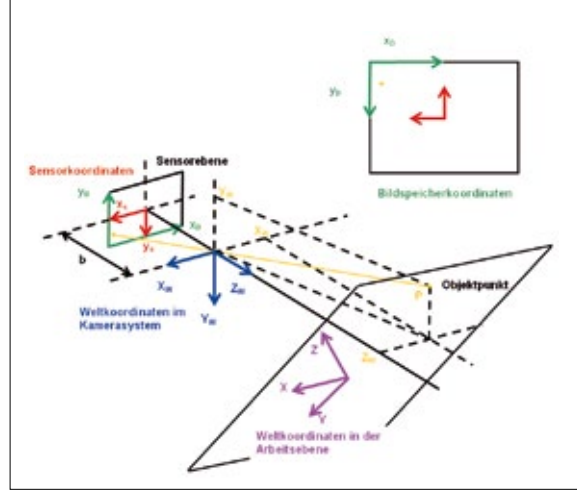


Abb. 3: Koordinatensysteme für die Transformation von Weltkoordinaten aus der Arbeitsebene in Pixelkoordinaten in der Bilddatei

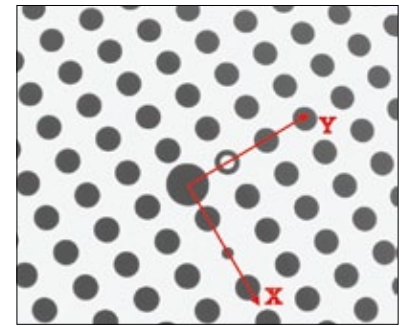


Abb. 4: Ebene Kalibriervorlage mit X- und Y-Weltkoordinatenachsen in der Arbeitsfläche

soluten Einheiten genau bekannten Abständen aufgebracht sind. Es ermöglicht außerdem die automatische Bestimmung des Koordinatenursprungs sowie der Achsenrichtungen des Weltkoordinatensystems in der Arbeitsebene. Die größere Kreisfläche im Zentrum definiert den Ursprung, das ringförmige Objekt und die kleinere Kreisfläche legen die Richtungen der X- und Y-Achse fest. Die Z-Achse steht senkrecht auf der Ebene. Die Position der Kreisflächen in der Bilddatei wird nach einer Binarisierung und anschließendem Labeling als Lage des Schwerpunktes der segmentierten Objekte berechnet. Diese Positionsbestimmung ist subpixelgenau und robust. Sie stellt keine hohen Anforderungen an die globale Homogenität der Beleuchtung. Ohne großen Aufwand erreichen wir im Labor eine Genauigkeit von 1/10 Pixelkantenlänge für die Bestimmung der Schwerpunktslagen.

### Ausgleichsrechnung

Die regelmäßige Struktur des Testmusters erleichtert die automatische Zuordnung zwischen den Punkten auf der Vorlage mit ihren bekannten Weltkoordinaten im System der Arbeitsebene und den nun bekannten, korrespondierenden Positionen in der Bilddatei. In Abbildung 4 könnte z.B. ein Array aus sechs mal sechs Kreisen verlässlich ausgewertet werden, also die Korrespondenzen von 36 bekannten Objektpunkten. Damit ist das Gleichungssystem für die 12 Modellparameter stark überbestimmt und muss angemessen ausgewertet werden. Die Werte mancher Parameter, z.B. die Pixelkantenlängen, finden sich mit hinreichender Genauigkeit im Datenblatt. Für die übrigen Kenngrößen wird mit den Methoden der Ausgleichsrechnung der optimale Parametersatz für die Anpassung der modellierten an die bekannten Daten berechnet. Geeignete Verfahren sind zwar mathematisch aufwendig, aber bekannt und numerisch handhabbar. Das Resultat einer solchen Auswertung ist jedenfalls ein Satz von Kameraparametern, der die Abbildung entsprechend dem Abbildungsmodell quantitativ beschreibt. Da in der Arbeitsfläche eine der Weltkoordinaten bekannt ist, nämlich die Z-Koordinate  $Z=0$  für alle Punkte in der Ebene, können nun die beiden übrigen Koordinaten X und Y für jeden Punkt in der Arbeitsebene aus den Pixelkoordinaten des korrespondierenden Bildpunktes mit Hilfe des Abbildungsmodells berechnet werden.

### Ungewissheit

Insbesondere kann man sich anschauen, wie gut die Kreise auf der Kalibriervorlage, aus denen die Kameraparameter berechnet wurden, mit dem Modell wieder gefunden werden. Ein geeignetes Gütemaß ist der Rückprojektionsfehler. Dazu werden die bekannten Weltkoordinaten mit Hilfe des Kameramodells in die Pixelkoordinaten der Bilddatei umgerechnet und mit den Pixelkoordinaten der Kreise im tatsächlich aufgenommenen Bild verglichen. Im Normalfall werden die Werte nicht genau über-

einstimmen, weil es sich bei der Kalibrierung um eine Ausgleichsrechnung handelt. Die Abweichung ist der Rückprojektionsfehler. Wahlweise kann man nun den mittleren oder den maximalen Rückprojektionsfehler als Gütemaß heranziehen. Im Labor erreichen wir bei Arbeitsabständen in der Größenordnung von einem halben Meter regelmäßig Rückprojektionsfehler in der Größenordnung von 1/10 bis 2/10 Pixelkantenlänge mit 8-bit-Standard-Analogkameras und handelsüblichen Objektiven. Die Ungewissheit des Verfahrens liegt also in derselben Größenordnung wie die Ungewissheit, mit der die Position der Schwerpunkte der Referenzobjekte in der Bilddatei bestimmt werden kann. Unter diesem Gesichtspunkt hat sich die Methode daher als tauglich erwiesen. Tatsächlich gelingt es dem Roboter aus Abbildung 1 auf diese Weise, die Schraubverschlüsse präzise aus dem linken Träger aufzunehmen und in den freien Mulden des rechten Trägers abzusetzen.

Die oben beschriebene Methode mit einer flächigen Kalibriervorlage in der Arbeitsebene beschränkt die Kalibrierung keineswegs auf diese Ebene. Die Kamera ist auf diese Weise vollständig kalibriert, solange ihre Orientierung nicht geändert wird. In der Arbeitsebene ist jedoch die vollständige Bestimmung der drei Weltkoordinaten (X, Y, Z) möglich, denn definitionsgemäß ist die Z-Koordinate für alle Punkte in dieser Ebene gleich Null und damit bekannt. Wenn für irgendeinen anderen Punkt in der Szene eine der drei Weltkoordinaten auf andere Weise bekannt wird, können aus den Pixelkoordinaten des korrespondierenden Bildpunktes die beiden übrigen Weltkoordinaten ebenfalls berechnet werden. Beispielsweise könnte bekannt sein, dass ein bestimmter Punkt, der im Bild gut detektierbar ist, 15 mm oberhalb der Arbeitsebene liegt. Dann ist seine Z-Koordinate ebenfalls bekannt: sie ist nicht null, sondern +15 mm, und die beiden übrigen Koordinaten X und Y können berechnet werden. Das ist übrigens regelmäßig bei der Szene aus Abbildung 1 der Fall, denn die Oberflächen der Schraubverschlüsse liegen nicht in einer Ebene mit der Montagefläche des Roboters, auf der üblicherweise die Kalibriervorlage platziert wird.

Das hier vorgestellte Kalibrierverfahren geht auf die Arbeiten von Lenz [1] und Tsai [2] zurück, die bereits 1987 publiziert wurden. Es ist bemerkenswert, dass diese Methoden schon vor 20 Jahren bekannt waren, aber erst vor wenigen Jahren vereinzelt kommerziell verfügbar wurden. Mittlerweile wird die Kamerakalibrierung von allen wesentlichen Bildverarbeitungsbibliotheken unterstützt und gehört zu den Grundlagen der Bildverarbeitung.

### Literatur

- [1] Lenz, R., Informatik-Fachberichte 149: Mustererkennung 1987, S. 212-216, Springer 1987
- [2] Tsai, R. Y., IEEE Journal of Robotics and Automation, Vol. RA-3 (4) 1987, 323-344

# Triangulation

Nicht für die Schule, sondern für das Leben ....

Triangulation ist ein trigonometrisches Verfahren zur Vermessung des Abstands eines Punktes von einem Referenzpunkt durch Anvisieren unter zwei verschiedenen Blickwinkeln. In der optischen Messtechnik versteht man unter einem Triangulationssensor einen Abstandstaster, der einen Laserstrahl unter einem definierten Einfallswinkel auf die Oberfläche eines Objekts projiziert und den Spot unter einem anderen Winkel auf eine Detektorzeile abbildet [1]. In Kombination mit einem Scanner oder einer Verschiebeeinheit kann man die gesamte Oberfläche eines Objekts abrastern und als Punktwolke im dreidimensionalen Raum darstellen. Der Artikel erläutert die Grundzüge des Verfahrens und einige grundlegende Konstruktionsprinzipien.

Jeder von uns hat das Grundprinzip der Triangulation schon in jungen Jahren in der Schule gelernt, verpackt in mehr oder weniger interessanten Fragestellungen, und ist später gelegentlich wieder darüber gestolpert. Wenn man wie in Abbildung 1 die Entfernung zwischen einem Punkt B und einem Objekt C nicht direkt messen kann, visiert man es von zwei verschiedenen Standpunkten A und B aus an, misst die Winkel, unter denen es in Bezug auf eine Referenzrichtung erscheint, und bestimmt den Abstand  $b$  zwischen den beiden Standpunkten, die Basisbreite. Mit ein wenig Trigonometrie kann man nun alle Stücke des Dreiecks ausrechnen, also auch den Abstand zwischen B und C. Wer sich schon einmal mit Stereo-Vision befasst hat, erkennt schnell, dass diese Konfiguration der Standard-Stereo-Geometrie mit zwei Kameras entspricht, deren optische Achsen parallel zueinander ausgerichtet sind [2]. Abbildung 1b zeigt einen Spezialfall, bei dem eine der beiden Blickrichtungen senkrecht auf der Basis steht. Im Triangulationssensor wird diese Blickrichtung durch einen Laserstrahl ersetzt, der unter einem definierten Beleuchtungswinkel einen Lichtfleck auf einer Oberfläche erzeugt. Vom Standpunkt A aus wird der Winkel bestimmt, unter dem der Laserspot auf der Oberfläche erscheint. Geometrisch ist diese Konfiguration vollkommen gleichwertig zu der Situation in Abbildung 1b. Gedreht um  $180^\circ$  entsteht die Geometrie von Abbildung 1d, die meist bei der Beschreibung eines Triangulationssensors verwendet wird: Ein Laserstrahl wird senkrecht auf eine Oberfläche gerichtet, und der Laserspot wird unter dem Triangulationswinkel  $\alpha$  von einem Linsensystem auf einen Detektor abgebildet.

## Die optische Abbildung

In Abbildung 2 ist die optische Abbildung des Laserspots vom Objekt in die Bildebene genauer dargestellt. Der Leuchtfleck wird mit einer Linse erfasst, deren optische Achse auf den Arbeitsabstand  $z$  des Sensors ausgerichtet ist. Wenn sich die Oberfläche längs der Ausbreitungsrichtung des Laserstrahls ver-

schiebt, der Abstand also größer oder kleiner wird, verschiebt sich auch die Position des Bildes in der Bildebene des Sensors. Die Verschiebung  $\Delta x$  in der Detektorebene ist in erster Näherung proportional zur Verschiebung  $\Delta z$  längs der Richtung des Laserstrahls:

$$\Delta x = \beta' \sin \alpha \Delta z$$

Dabei ist  $\beta'$  der Abbildungsmaßstab, der von der Brennweite  $f'$  der Optik und der Gegenstandsweite  $a$  bzw. der Bildweite  $a'$  abhängt:

$$\beta' = (a'/a) = f'/(a+f')$$

In der Bildebene eines Triangulationssensors muss folglich die Verschiebung  $\Delta x$  des Bildes des Laserspots in Bezug auf die optische Achse ermittelt werden. Dazu kann eine Detektorzeile mit diskreten Pixeln oder ein positionsempfindlicher Detektor, der ein Analogsignal liefert, verwendet werden. Der Proportionalitätsfaktor  $\beta' \sin \alpha$  ist in erster Näherung konstant. Abweichungen von der linearen Beziehung zwischen  $\Delta x$  und  $\Delta z$  können durch

Kalibrierung erfasst und in der Auswertung kompensiert werden. Der Arbeitsbereich, also die maximale Verschiebung  $\Delta z$  um den Arbeitsabstand herum, wird durch die maximale Ablage  $\Delta x$  des Bildpunktes auf dem Detektor begrenzt, also letztlich durch die Länge der Detektorzeile. Die Tiefenauflösung ergibt sich aus der Genauigkeit, mit der die Position des Bildflecks erfasst werden kann. Bei einer Detektorzeile ist das auf den ersten Blick der Pixelabstand. Grundsätzlich kann der Schwerpunkt eines Spots, der sich über mehrere Pixel erstreckt, jedoch mit Subpixelgenauigkeit bestimmt werden. Für einen Sensor, in dem eine Detektorzeile mit 1024 Pixeln verbaut ist, wird man folglich eine Tiefenauflösung in der Grö-

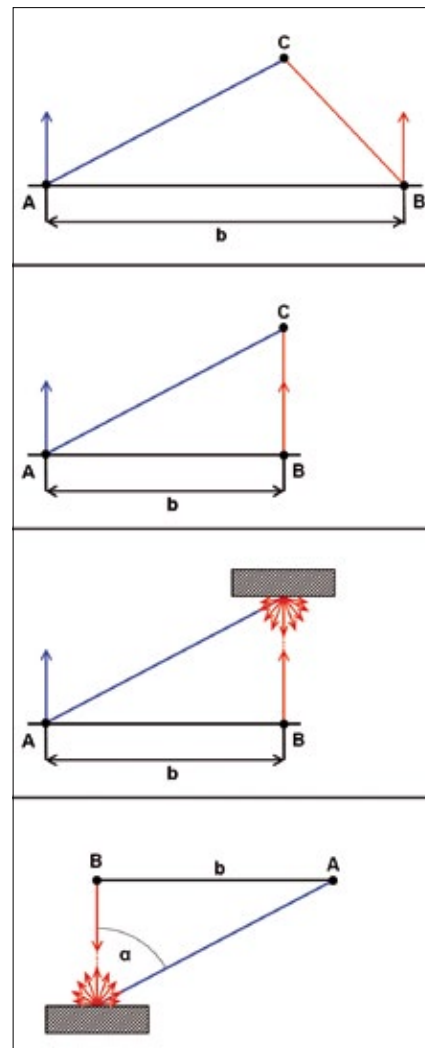


Abb. 1: Das Triangulationsprinzip

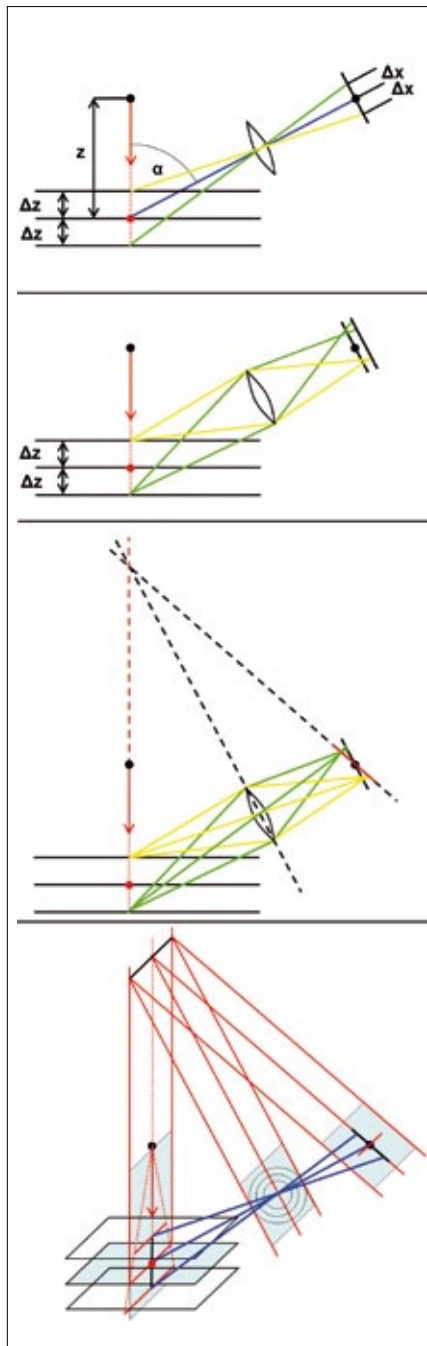
Benennung von einem Promille des Arbeitsbereichs erwarten.

## Die Scheimpflug-Bedingung

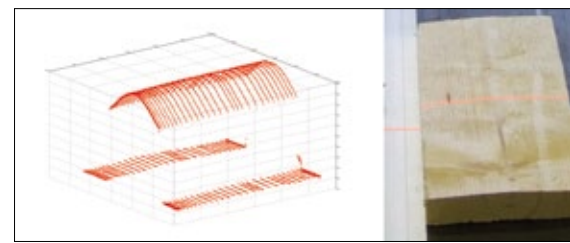
Aus Abbildung 2 geht auch hervor, dass die optische Abbildung unscharf wird, wenn die Objektebene sich längs der z-Richtung vom Arbeitspunkt aus nach oben oder unten verschiebt. Die Gegenstandsebene, die von der Linse scharf abgebildet wird, steht senkrecht auf der optischen Achse. Die Tiefenvariation  $\Delta z$  verläuft jedoch längs des Laserstrahls und ist damit gegenüber der Gegenstandsebene verkippt. Wenn der Detektor senkrecht auf der optischen Achse steht, ist das Bild über- oder unterfokussiert, sobald es aus der zentralen Position herauswandert. Die Unschärfe in der Bildebene kann man kompensieren, wenn man den Detektor um die zentrale Position herum verkippt. Für den Kippwinkel muss die sog. Scheimpflug-Bedingung [1] erfüllt sein: Bildebene, Linsenebene und Gegenstandsebene müssen sich in einer Linie bzw. in einem Punkt schneiden, siehe Abbildung 2c. Die Gegenstandsebene wird beim Triangulationssensor durch den einfallenden Laserstrahl definiert, denn man möchte, dass alle Punkte, die sich im Arbeitsbereich auf dem Laserstrahl befinden, scharf abgebildet werden. Damit ist auch der nächste Schritt zum sog. Lichtschnitt-Verfahren nahe liegend: Anstelle eines dünnen Strahlenbündels kann man auch eine Laserlinie auf das Objekt projizieren und mit der Empfängeroptik auf einen Array-Detektor abbilden. Das Bild der Linie enthält Punkt für Punkt die Abstandsinformation für den korrespondierenden Punkt der Laserlinie auf dem Objekt. Die Gegenstandsebene, die der Scheimpflug-Bedingung genügen muss, wird dann durch die Laserlinie und die Ausbreitungsrichtung des Mittenstrahls definiert, also durch den entstehenden Laserfächer, siehe Abbildung 2d. Solche Anordnungen werden als Lichtschnitt-Sensoren bezeichnet. Abbildung 3 zeigt ein Beispiel für eine Messung mit einem kommerziellen Lichtschnitt-Sensor. Die Laserlinie verläuft senkrecht zu einem Vor-schub, mit dem das Messobjekt unter dem Sensor hinweg bewegt werden kann. Für jeden Lichtschnitt ergibt sich ein Oberflächenprofil in der Ebene des Laserfächers. Die gesamte abgetastete Kontur des Objekts entsteht als Punktwolke aus der Aneinanderreihung der einzelnen Profile.

## Randbedingungen

Die Triangulation entsprechend Abbildung 2 funktioniert nur für diffus reflektierende Objekte. Der einfallende Laserstrahl muss zum Teil in Richtung der Nachweicoptik remittiert werden, damit ein Signal entsteht. Glänzende, spiegelnde Oberflächen sind daher für das Verfahren nicht geeignet. Dunkle Flächen mit geringem Remissionsgrad erzeugen möglicherweise ein zu geringes Signal für eine verlässliche Bestimmung der Bildposi-



**Abb. 2: Die optische Abbildung beim Triangulationssensor. Damit in der Detektorebene ein scharfes Bild entsteht, muss der Sensor so verkippt werden, dass sich Gegenstandsebene, Linsenebene und Bildebene in einer Linie schneiden (Scheimpflug-Bedingung)**



**Abb. 3: Eine Messung mit einem Lichtschnittsensor. Das Messobjekt wird senkrecht zur Laserlinie transportiert. Die Gesamtheit der Höhenprofile ergibt eine 3D-Darstellung der Oberfläche als Punktwolke.**

tion. Helle Objekte sind folglich besser für die Triangulation geeignet als dunkle Flächen. Für eine verlässliche Abtastung der Oberfläche muss die Remission tatsächlich an der Oberfläche erfolgen. Manche Materialien, z.B. einige Kunststoffe, sind jedoch Volumenstreuer, d.h. das Streulicht kommt aus einer ausgedehnten Zone im Innern des Materials. Das Streuverhalten kann außerdem stark von der Wellenlänge abhängen. Materialien, die im Sichtbaren stark streuen, können im NIR nahezu transparent sein [3]. Die Wellenlänge der Beleuchtung muss folglich bei der Anwendung des Sensors berücksichtigt werden. Bei der optischen Auslegung ist zu beachten, dass die Scheimpflug-Bedingung lediglich eine Korrektur erster Ordnung ist. Ein Objektiv mit starker Bildfeldwölbung und optischer Verzeichnung wird zusätzliche Unschärfe und weitere Änderungen des Abbildungsmaßstabs längs der Detektorzeile hervorrufen. Hinzu kommen perspektivische Verzerrungen durch die Zentralprojektion bei Standard-Objektiven. Wenn Störreflexe in die Optik gelangen können, finden sich womöglich plötzlich zwei Spots in der Bildebene, und die Unterscheidung zwischen Nutz- und Störsignal kann schwierig werden. Generell muss sich ein Laserfleck oder eine Laserlinie im Bild hinreichend vom Hintergrund abheben. Hier kommt die Verknüpfung mit der Bildverarbeitung ins Spiel. Sowohl die Strahlformung durch die Beleuchtungsoptik als auch die Algorithmen können speziell so ausgelegt werden, dass die sichere Erkennung und Auswertung der Laserspots im Bild unterstützt

wird. Auch die Projektionsoptik für die Beleuchtung muss optimiert werden, wenn die physikalisch mögliche Ortsauflösung quer zur Tiefenauflösung erreicht werden soll. Weiterbildungen des einfachen Lichtschnittprinzips sind Sensoren, die Gruppen von parallelen Linien oder Streifenmuster für die Beleuchtung verwenden.

## Literatur

- [1] A. Donges, R. Noll, Lasermesstechnik, Hüthig, Heidelberg 1993
- [2] s. INSPECT 1/2006, Stereo-Vision, S. 18
- [3] s. INSPECT 2/2006, NIR-Imaging, S. 14



# Radiometrie

## Strahlende Präsenz

Radiometrie ist die objektive Bewertung der Intensität von Strahlung im Hinblick auf Anwendungen in der optischen Messtechnik. Sobald das Signal in einem optischen Sensor in Kombination mit einer Beleuchtungseinrichtung quantifiziert oder abgeschätzt werden soll, kommen die Messgrößen und Begriffe der Radiometrie ins Spiel. Wer z. B. vorab wissen will, ob 10 LEDs zur Ausleuchtung eines Prüflings ausreichen, damit das Kamerasignal voll ausgesteuert wird, kommt um die Beschäftigung mit dem Themengebiet Radiometrie nicht herum. Dasselbe gilt für den Vergleich von Lichtquellen oder Beleuchtungseinrichtungen verschiedener Hersteller.

### Radiometrische Größen

Radiometrische Messgrößen bewerten die Energie von Strahlung. In diesem Artikel steht die Anwendung in der optischen Messtechnik im Vordergrund. Wir beschränken uns daher auf den sichtbaren Spektralbereich, das UV und den nahen Infrarotbereich, soweit diese mit Photodioden, CCD- und CMOS-Detektoren erfasst werden können. Die Strahlung soll unpolarisiert und inkohärent sein, d.h. die Welleneigenschaften der Strahlung sollen nicht in Erscheinung treten. Zur Vereinfachung gehen wir außerdem zunächst davon aus, dass wir es mit einer spektral schmalbandigen Lichtquelle zu tun haben. Eine typische Lichtquelle, die diese Bedingungen erfüllt, ist eine z. B.

rot leuchtende LED mit einer Zentralwellenlänge von z. B. 620 nm und einer spektralen Halbwertsbreite von vielleicht 20 nm. Wenn der Hersteller den LED-Chip in eine Kunststofflinse mit z. B. 5 mm Durchmesser eingebettet hat, wird die Lichtemission in einen Strahlkegel mit vielleicht 20° Öffnungswinkel gebündelt. Betrachtet man diese LED aus genügend großer Entfernung, z. B. aus 50 cm, sieht es so aus, als gehe das Strahlenbündel von einer winzigen emittierenden Fläche aus, und wir können die LED als Punktlichtquelle betrachten. Wenn in diesem Strahlenbündel nun eine Photodiode oder eine Kamera montiert wird, kann man mit Hilfe weniger Kenngrößen abschätzen, welches Signal im Detektor entstehen wird.

Eine elementare radiometrische Größe ist die gesamte Strahlungsenergie  $Q$  einer Lichtquelle. Damit ist die Energie gemeint, die über alle Emissionsrichtungen integriert von der Quelle innerhalb eines definierten Zeitraums abgestrahlt wird, z. B. die Strahlungsenergie einer Blitzlampe in einem Blitz.  $Q$  ist eine Energie und trägt die Einheit Joule. Meist wird nicht  $Q$ , sondern die Strahlungsleistung  $\Phi$  spezifiziert.  $\Phi$  ist die zeitliche Änderung von  $Q$ , also die über alle Emissionsrichtungen der Quelle integrierte abgestrahlte Energie pro Zeiteinheit, angegeben in der Einheit Watt. Generell kann man die von einem Strahlenbündel transportierte Strahlungsleistung quantifizieren, indem man die Strahlungsenergie bestimmt, die pro Zeiteinheit durch den gesamten Querschnitt des Bündels hindurchtritt.  $\Phi$  wird daher auch als Strahlungsfluss bezeichnet. Strahlungsenergie  $Q$  und Strahlungsfluss  $\Phi$  sind Erhaltungsgrößen, d.h. sie ändern sich entlang der Ausbreitungsrichtung des Strahlenbündels nicht, solange keine Verluste durch Absorption, Streuung oder Reflexion an Grenzflächen auftreten. Der Hersteller der LED könnte für definierte Betriebsbedingungen z. B. eine Strahlungsleistung von 10 mW spezifizieren, die in einem Kegel mit 20° Öffnungswinkel emittiert wird. Unabhängig vom Abstand von der Quelle treten immer 10 mW Strahlungsleistung durch den Querschnitt des Strahlenbündels. Mit zunehmendem Abstand von der Quelle nimmt jedoch der Querschnitt zu, so dass die Strahlungsleistung auf eine immer größer werdende Fläche verteilt wird. Diese Überlegung führt zur Definition der Strahlungsflussdichte, auch als Bestrahlungsstärke  $E$  bezeichnet.  $E$  ist der Strahlungsfluss  $\Phi$  bezogen auf die Fläche, durch die die Strahlung hindurch tritt:

$$E = d\Phi/dA$$

$E$  wird in der Einheit  $W/m^2$  angegeben. Beispielsweise erzeugt die Sonne außerhalb der Erdatmosphäre in der Flughöhe der internationalen Raumstation ISS eine Bestrahlungsstärke von  $1350 W/m^2$ . Unsere Beispiel-LED hingegen, senkrecht ausgerichtet auf ein weißes Blatt Papier in 50 cm Entfernung, würde eine kreisförmige Fläche mit 8,8 cm Radius ausleuchten und daher auf dieser Fläche von  $244 cm^2$  eine mittlere Bestrahlungsstärke von  $10 mW/244 cm^2$ , also etwa  $0,4 W/m^2$  liefern. Die Bestrahlungsstärke ist in Bezug auf die Fläche eine differentielle Größe. Sie kann daher für jeden Punkt eines Strahlungsfeldes angegeben werden. Durch Multiplikation mit der Abtastfläche  $dA$  eines Detektors oder auch eines Pixels ergibt sich dann der Strahlungsfluss  $d\Phi$ , der diese Abtastfläche durch-



setzt. Im Allgemeinen wird die Bestrahlungsstärke vom Ort abhängen, z.B. auf einer beleuchteten Fläche.

### Photonenfluss

Die Bestrahlungsstärke ist gut zur radiometrischen Charakterisierung eines Strahlenbündels an einer definierten Stelle im Strahlengang geeignet. Da der Strahlungsfluss im Strahlengang eine Erhaltungsgröße ist und eine Abbildungsoptik in erster Näherung lediglich die Geometrie des Bündels, also dessen Öffnungswinkel und Querschnitt verändert, kann man den Fluss durch die Optik rechnerisch von der Lichtquelle bis zur Detektorebene verfolgen und die Bestrahlungsstärke berechnen. Aus der Bestrahlungsstärke auf dem Detektor eines optischen Sensors oder auf dem Detektorarray einer Kamera kann man dann das Signal abschätzen. Bei Photodioden, CCD- und CMOS-Sensoren ist der Elementarprozess der innere Photoeffekt: Photonen werden im Detektormaterial absorbiert und erzeugen Elektronen, die entweder als Strom ausgekoppelt oder in Form von Ladungspaketen als Signal ausgelesen werden. Bei dieser Umwandlung von Licht in Ladung erzeugt jedes Photon höchstens ein Elektron. Die Wahrscheinlichkeit, mit der Photonen in Elektronen umgesetzt werden, ist die Quantenausbeute. Sie ist von der Wellenlänge der einfallenden Strahlung abhängig und wird bei guten Detektoren im Datenblatt spezifiziert. Man kann die Quantenausbeute daher bei Bedarf in die Abschätzung einbeziehen. Wesentlich bei dieser Überlegung ist aber ein anderer Aspekt. Die aufgeführten Detektoren erzeugen ihr Signal nicht dadurch, dass sie Strahlungsenergie in elektrische Ladung umsetzen, sondern sie setzen einzelne Photonen in einzelne Elektronen um. In erster Linie ist also wichtig, wie viele Photonen auf den Detektor fallen, und nicht, welche Lichtenergie dort deponiert wird. Selbstverständlich hängen diese beiden Größen miteinander zusammen. Ein Photon der Wellenlänge  $\lambda$  trägt die Energie:

$$E_{\text{photon}} = hc/\lambda$$

Dabei ist  $h$  das Planck'sche Wirkungsquantum und  $c$  die Lichtgeschwindigkeit. In der optischen Messtechnik verwendet man als Einheit für die Photonenenergie nur ungern das Joule, weil die Zahlenwerte für Photonenenergien damit unhandlich klein werden. Stattdessen ist die Einheit eV („Elektronenvolt“) üblich, das Produkt aus der Elementarladung und der Einheit Volt der elektrischen Spannung. 1 eV sind  $1,6 \times 10^{-19}$  J. Wer häufiger Photonenenergien ausrechnet, merkt sich irgendwann den Wert des Produktes  $hc$  und weiß, dass die Energie eines Photons der Wellenlänge  $\lambda$  sich aus der einfachen Beziehung:

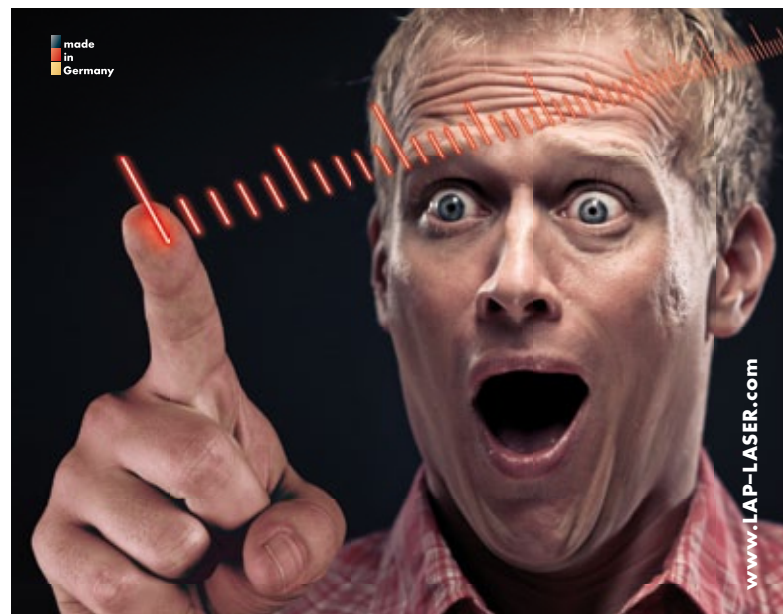
$$E_{\text{photon}} = (1240/\lambda) \text{ eV} \times \text{nm}$$

ergibt. 1240 dividiert durch die Wellenlänge in nm ergibt also die Photonenenergie in eV. Beispielsweise hat ein IR-Photon mit 1240 nm Wellenlänge gerade die Photonenenergie 1 eV. Ein rotes Photon mit 620 nm Wellenlänge wie bei unserer Beispiel-LED hat eine Energie von 2 eV, ein blaues Photon mit ca. 413 nm Wellenlänge hat eine Energie von 3 eV. Langwellige Strahlung besteht folglich bei gleicher Strahlungsenergie aus einer größeren Zahl von Photonen als kurzwellige Strahlung. Das Strahlenbündel unserer Beispiel-LED mit der Wellenlänge 620 nm und einem Strahlungsfluss von 10 mW transportiert pro Sekunde die ungeheure Zahl von etwa  $3 \times 10^{16}$  Photonen durch den Strahlquerschnitt. Bei einer LED mit 10 mW Strahlungsleistung, jedoch bei der Emissionswellenlänge 413 nm, wären es nur  $2 \times 10^{16}$  Photonen pro Sekunde, und das Detektorsignal

würde nur zwei Drittel des Signals bei 620 nm betragen, dieselbe Quantenausbeute bei beiden Wellenlängen vorausgesetzt. Für die Abschätzung des Detektorsignals sind daher die Photonenzahl  $N$  und der Photonenfluss  $dN/dt$  von Bedeutung. Der Photonenfluss ergibt sich aus der Strahlungsleistung:

$$dN/dt = \Phi/E_{\text{photon}}$$

Es ist offenkundig, dass diese Beziehung nur für annähernd monochromatische Strahlung verwendet werden darf. Beispielsweise kann aus der Angabe für die Bestrahlungsstärke von  $1350 \text{ W/m}^2$  aufgrund der Sonneneinstrahlung auf der Flugbahn der ISS nicht ohne weiteres der Photonenfluss pro Flächeneinheit berechnet werden, denn das Spektrum der Sonneneinstrahlung ist kontinuierlich vom UV bis ins IR. Erst die genaue Kenntnis der Spektralverteilung ermöglicht eine Umrechnung in spektrale Photonenflüsse und letztlich die Berechnung des gesamten Photonenflusses durch Integration über die spektralen Anteile. Wenn die Spektralverteilung auch noch variiert, wie das z.B. bei Anwendungen im Außenraum beim Sonnenlicht in Abhängigkeit von der Jahres- und Tageszeit der Fall ist, wird die Abschätzung des Signals zusätzlich erschwert. Schon bei Beleuchtungen mit modernen weißen LEDs unter kontrollierten Bedingungen werden radiometrische Berechnungen aufgrund des kontinuierlichen Spektrums erheblich komplizierter als bei quasi-monochromatischen LED-Beleuchtungen.



## UNHEIMLICH GENAU!



ATLAS



POLARIS



CALIX



## Strahlstärke

Zum Abschluss soll noch eine weitere radiometrische Größe erwähnt werden, die beim Vergleich von Lichtquellen eine große Rolle spielt. Unsere Beispiel-LED emittiert 10 mW in einen Kegel mit 20° Öffnungswinkel. Ohne Fokussierlinse würde die Strahlungsleistung in den gesamten Halbraum oberhalb des Chips abgestrahlt, so dass in einem bestimmten Abstand von z. B. 50 cm die Bestrahlungsstärke  $E$  nicht mehr etwa  $0,4 \text{ W/m}^2$  betragen würde, sondern erheblich weniger, weil sich die Strahlungsleistung nun auf eine Halbkugel mit 50 cm Radius verteilt und nicht mehr nur auf die Grundfläche des Strahlungskegels mit ca. 8,8 cm Radius im Abstand von 50 cm. Eine Größe, die die Divergenz der Abstrahlung und ihre Auswirkung auf die energetischen Verhältnisse des Strahlenbündels unabhängig vom Abstand zur Quelle beschreibt, wäre daher praktisch. Diese radiometrische Größe ist die Strahlstärke  $I$ , definiert als der Fluss  $d\Phi$  pro Raumwinkel  $d\Omega$ . Der Raumwinkel beschreibt den kegelförmigen Ausschnitt aus einer Kugel, dessen Spitze im Zentrum der Kugel liegt. Zur Quantifizierung betrachtet man eine Kugel mit Radius  $r$  und berechnet die Fläche  $A$  der Kugelkalotte, die vom Raumwinkelkegel aus der Oberfläche dieser Kugel ausgeschnitten wird. Der zugehörige Raumwinkel ist dann  $\Omega = A/r^2$  mit der Einheit Steradian (sr). Diese Begriffsbildung ist analog zum Bogenmaß für den ebenen Winkel als Verhältnis von Kreisbogen zu Radius in der Einheit Radian (rad). Der Raumwinkel für die Vollkugel ist gleich der Fläche einer Kugel mit Radius  $r$  dividiert durch  $r^2$ , also  $4\pi r^2/r^2 = 4\pi \text{ sr}$ , für die Halbkugel  $2\pi \text{ sr}$ , und für unsere Beispiel-LED etwa  $244 \text{ cm}^2/50^2 \text{ cm}^2$ , also etwa  $0,1 \text{ sr}$ . Für kleine Öffnungswinkel approximiert man meist  $d\Omega = dA/r^2$  und nimmt als Fläche  $dA$  die auf dem Radius senkrecht stehende ebene Kegelgrundfläche anstelle der Kugelkalotte. Als Strahlstärke ergibt sich damit im Mittel  $I = 10 \text{ mW}/0,1 \text{ sr} = 0,1 \text{ W/sr}$ . Da die Strahlstärke in Bezug auf den Raumwinkel eine differentielle Größe ist, kann man sie für jeden Punkt des Strahlungsfeldes angeben. Multipliziert mit dem jeweils erfassten Raumwinkelement ergibt sich daraus die zugehörige Strahlungsleistung. Im Unterschied zur Bestrahlungsstärke bleibt die Strahlstärke jedoch bei freier Strahlausbreitung für Punktquellen unabhängig vom Abstand zur Quelle konstant, während die Bestrahlungsstärke mit  $1/r^2$  abnimmt.

## Photometrische Größen

Hersteller von LEDs geben oft die Strahlstärke auf der Symmetrieachse als Gütemaß in der Einheit W/sr an. Beim Vergleich zweier LEDs sollte man beachten, dass dabei die Winkeldivergenz der Emission eingeht. Wenn es dem Hersteller gelingt, bei gleicher Strahlungsleistung die Winkeldivergenz von 20° Öffnungswinkel auf 10° zu halbieren, steigt die Strahlstärke um den Faktor 4, obwohl die LED nicht mehr Lichtleistung auf eine Probe einstrahlen kann. Bei gleichem Abstand zur Probe wird diese Lichtleistung lediglich auf einer kleineren Fläche deponiert, was für manche Anwendungen wesentlich sein kann, für andere hingegen vielleicht nicht. Hinzu kommt, dass LED-Hersteller leider nach wie vor gern „Intensitäten“ als Lichtstärke in der photometrischen Einheit Candela (cd) angeben. Die Lichtstärke ist begrifflich identisch mit der Strahlstärke, jedoch bewertet mit der Empfindlichkeitskurve des menschlichen Auges. Für messtechnische Zwecke ist die Sinnesempfindung des Menschen meist irrelevant. Das photometrische Analogon der Bestrahlungsstärke, also bewertet mit der Empfindlichkeitskurve des Auges, ist die Beleuchtungsstärke mit der Einheit Lux (lx). So findet man leider immer noch Empfindlichkeitsangaben für Kameras in der Form „0 lx“, was lediglich bedeutet, dass die Kamera im Infraroten empfindlich ist, denn Strahlung im IR ist für das menschliche Auge unsichtbar und führt daher photometrisch zur Beleuchtungsstärke 0 lx, unabhängig von der radiometrischen Bestrahlungsstärke in  $\text{W/m}^2$  dieses Strahlungsfeldes, selbst wenn es vielleicht eine Intensität von  $10 \text{ W/m}^2$  haben sollte. Für monochromatische Lichtquellen kann man im sichtbaren Spektralbereich die photometrischen Angaben anhand der genormten Empfindlichkeitskurve in radiometrische umrechnen. Bei spektral kontinuierlichen Lichtquellen ist eine allgemeine Umrechnung nicht möglich, sondern erfordert genaue Kenntnis der Spektralverteilung.

### Literatur

- [1] Snell, J.F., Radiometry and Photometry, in Driscoll, W.G. (ed.), Handbook of Optics, McGraw-Hill, 1978, p. 1–1 ff.
- [2] Pedrotti, Pedrotti, Bausch, Schmidt, Optik, Prentice Hall 1996, p. 29 ff.



# Kameraempfindlichkeit

## Dynamisches Rauschen

Welche Beleuchtungsstärke ist erforderlich, damit ein auswertbares Bildsignal entsteht? Wie hell dürfen die hellsten Stellen und wie dunkel die dunkelsten Zonen im Bild sein, ohne dass die Bildauswertung zusammenbricht? Wer sich mit diesen Fragen befasst, stellt schnell fest, dass die Angelegenheit nicht mit einem kurzen Blick in das Datenblatt einer Kamera erledigt ist. In diesem Artikel erläutern wir einige grundlegende Zusammenhänge, die für die weitergehende Beschäftigung mit diesem komplexen Thema [1] wichtig sind.

### Empfindlichkeit

Das Signal in CCD- und CMOS-Sensoren entsteht durch den inneren Photoeffekt. Photonen aus dem einfallenden Lichtstrom lösen im Innern eines Detektorpixels Elektronen aus. Die Wahrscheinlichkeit, mit der dies geschieht, wird als Quantenausbeute oder auch als Quantenwirkungsgrad bezeichnet. Sie hängt von der Wellenlänge der einfallenden Strahlung ab. Die Kamerachips sprechen also nicht auf die einfallende Lichtleistung oder die Lichtenergie an, sondern auf die Zahl der einzelnen Photonen  $N_{ph}$ , die während der Integrationszeit auf den Detektor fallen. Das Quellsignal ist folglich die Zahl  $N_e$  der Photoelektronen, die innerhalb der Belichtungszeit in einem Pixel erzeugt und dort gespeichert werden, also eine Ladung. Diese Ladung wird schließlich ausgelesen und in eine Spannung umgesetzt, die wiederum mit einem AD-Wandler in einen digitalen Grauwert  $g$  umgewandelt wird. Der Proportionalitätsfaktor  $K = g/N_e$  wird als Ausleseverstärkung (conversion gain) bezeichnet und ist eine wichtige Kenngröße einer Kamera. Zur besseren Verständlichkeit fügt man die Pseudoeinheit „DN/e“ an, „digital numbers per electron“, also Graustufen pro Elektron. Beispielsweise bedeutet die Angabe  $K = 0,25 \text{ DN/e}$ , dass vier Signalelektronen im Pixel eine Graustufe erzeugen. Wenn  $K$  über die gesamte Aussteuerung des Bildsignals konstant bleibt, ist das eine Aussage über die Güte des elektronischen Signalpfades in der Kamera. Es gibt Kameras mit erheblichen Abweichungen von diesem Ideal. Wenn zusätzlich die Quantenausbeute bekannt ist, kann man ausrechnen, wie viele Photonen erforderlich sind, um einen bestimmten Grauwert im digitalen Signal zu erzeugen. Umgekehrt kann man aus dem Strahlungsfluss, der auf ein Pixel fällt, und der Belichtungszeit die Zahl der einfallenden Photonen berechnen [2] und damit das Graustufensignal abschätzen. Als Empfindlichkeit kann man dann das Verhältnis  $s$  zwischen dem Grauwert und der Zahl der einfallenden Photonen in der Einheit DN/photon angeben. Alternativ kann man den Grauwert auf die Lichtenergie beziehen, die während der Integrationszeit auf ein Pixel fällt und eine Empfindlichkeit in der Einheit DN/J angeben, wahlweise auch bezogen auf die Pixelfläche, DN/(J/m<sup>2</sup>). Ein Anwender, der ein Grauwertsignal abschätzen möchte, ist folglich dankbar, wenn der Kamerahersteller im Datenblatt die Pixelfläche und die absolute spektrale

Empfindlichkeit in der Einheit DN/photon, DN/J oder DN/(J/m<sup>2</sup>) angibt und die Wellenlänge nennt, bei der dieser Wert gemessen ist. Außerdem muss die Empfindlichkeit oder die Quantenausbeute als Funktion der Wellenlänge angegeben sein, notfalls relativ, z. B. normiert auf das Maximum. Alternativ hilft auch die Angabe der Quantenausbeute und der Ausleseverstärkung  $K$  in der Einheit DN/e<sup>-</sup> weiter.

### Dunkelsignal und Full Well Capacity

Auch wenn kein Licht in die Kamera fällt, werden während der Integrationszeit in jedem Pixel Elektronen erzeugt und gespeichert. Dieses Dunkelsignal entsteht durch thermische Effekte im Innern des Detektorchips und kann lediglich durch Kühlung reduziert werden. Die Dunkelektronen sind nicht von den Photoelektronen unterscheidbar, so dass sich das Nutzsignal und das Dunkelsignal zum Gesamtsignal addieren,  $N_{ges} = N_e + N_{dark}$ . Ein Detektorpixel kann nur bis zu einer Obergrenze Signalelektronen speichern, der „full well capacity“ FW. Standardkameras sind so ausgelegt, dass sich die maximale Belichtungszeit aus der Framerate ergibt und das Dunkelsignal dann weit unter der Sättigungsgrenze bleibt. In jedem Fall reduziert das Dunkelsignal jedoch den verfügbaren Signalumfang für das Nutzsignal. Wenn die FW beispielsweise 16.000 Elektronen beträgt, das Dunkelsignal aber bereits 4.000 Elektronen beiträgt, bleibt für das Nutzsignal nur noch ein Bereich von 12.000 Elektronen übrig. Mit  $K = 0,25 \text{ DN/e}$  bleiben damit von insgesamt 4.000 Graustufen, also einem 12bit-Signal, nur noch 3.000 für das Nutzsignal. Bei zu geringer Helligkeit in der Szene kann man sich also nicht einfach damit helfen, dass man die Belichtungszeit so lange erhöht, bis das Signal hinreichend groß ist, denn das Dunkelsignal wächst mit und reduziert den effektiv nutzbaren Signalumfang. Hinzu kommt, dass die meisten Detektoren in der Nähe der FW nichtlinear werden, d. h. das Signal hängt nicht mehr linear mit der Bestrahlungsstärke zusammen. Viele Anwender meiden daher diesen Bereich.

### Rauschen

Das Signal in einem Pixel ist über die Zeit nicht konstant, sondern es rauscht. Auch unter sonst völlig gleichen Aufnahmebedingungen wird man in einer Serie von Bildaufnahmen sowohl für das Dunkelsignal als auch für das Nutzsignal eine zufällige Streuung um den jeweiligen Mittelwert feststellen. Die Wurzel aus der mittleren quadratischen Abweichung vom Mittelwert, das rms-Rauschen, ist ein Maß für diese Streuung. Für einen Teilchendetektor ist das Rauschen gleich der Wurzel aus der Zahl der Teilchen. Beträgt der Mittelwert des Signals also 16.000 Elektronen, ist diesem Mittelwert ein Fluktuationsband mit einem rms-Wert von etwa 125 Elektronen überlagert. Wenn man dieses Signal auf 12bit abbildet, also 16.000 Elektronen etwa auf den Grauwert 4.000, entspricht das rms-Rauschen von 125 Elektronen etwa 32 Graustufen. Auch die Subtraktion des Dunkelsignals hilft nicht weiter, denn im Ergebnis wird das Rauschen höher als bei jedem der beiden Einzelsignale, weil die rms-Rauschwerte sich quadratisch addieren. Je höher das Dunkelsignal wird, desto größer wird also

das Rauschen im Gesamtsignal und ggf. auch im Differenzsignal, bis es völlig in der Fluktuation untergeht. Die Graustufenauflösung, die der ADC anbietet, ist folglich unter Umständen gar nicht voll nutzbar. Wie viele Graustufen man tatsächlich sehen kann, hängt vom Rauschen ab und nicht von der Auflösung des AD-Wandlers.

### Dynamikbereich

Die Verhältnisse in der Nähe der Sättigung haben wir bereits betrachtet. Bei geringen Signalen muss außerdem das Ausleserauschen berücksichtigt werden. Auch bei tatsächlich identischer Zahl von Signalelektronen fluktuiert der Grauwert bei mehreren aufeinander folgenden Auslesevorgängen um einen Mittelwert. Der rms-Wert dieser Fluktuation ist das Ausleserauschen, und man muss damit rechnen, dass das Ausleserauschen größer ist als eine Graustufe, die der AD-Wandler erzeugt, z. B. fünf Elektronen. Als Kenngröße gibt man oft das Verhältnis zwischen dem Sättigungssignal und dem Ausleserauschen an und bezeichnet diesen Quotienten als den Dynamikbereich DR (dynamic range). In unserem Beispiel wäre der Dynamikbereich also  $16.000/5 = 3.200$  oder 70 dB. Für eine Industriekamera wäre das schon ein beeindruckender Wert. Das bedeutet aber nicht, dass man tatsächlich 3.200 Graustufen sehen kann, denn das Signalrauschen kommt hinzu. In der Nähe der Sättigung beträgt das Signalrauschen etwa 125 Elektronen, also das 25-fache des Ausleserauschs. Deshalb ist es auch irreführend, den Dynamikbereich als Signal-Rausch-Verhältnis (SNR) zu bezeichnen, wie dies in manchen Datenblättern geschieht. In der Nähe der Sättigung ist in unserem Beispiel das Signal-Rausch-Verhältnis lediglich etwa 125, und besser kann es über den ganzen Signalbereich auch nicht werden. Auch dieser Wert ist bereits ein Optimalwert, denn die Folgeelektronik kann (und wird meist auch) dem Signal weitere Rauschteile hinzufügen. Damit ein Signal sich verlässlich vom Rauschen unterscheidet, fordert man meist ein SNR von mindestens 10. In unserem Beispiel wäre bei 120 Elektronen das Gesamtrauschen bei Vernachlässigung des Dunkelsignals etwa 12 Elektronen, d. h. die Graustufe 30 wäre bei einem Rauschen von drei Graustufen sicher erkennbar. Ab Graustufe 400 hätte man mit  $SNR = 40$  bei einem Rauschen von etwa 10 Graustufen eine sehr gute Signalqualität. An diesem Beispiel wird bereits deutlich, dass die Überlegungen in Bezug auf die Auslegung von Algorithmen an dieser Stelle noch nicht beendet sind. Wenn man Graustufendifferenzen sicher erkennen

will, kommt es auf das absolute Rauschen an, so dass in dunkleren Bildbereichen andere Kriterien gelten als in sehr hellen Zonen.

### Fazit

Wer die Graustufenwerte im Kamerabild abschätzen will, sollte für die hellsten und die dunkelsten Stellen der Szene messen oder berechnen, welche Lichtenergie oder welche Zahl von Photonen während der Integrationszeit auf ein Pixel fallen wird. Dabei muss die Spektralverteilung beachtet werden. Aus der Quantenausbeute und der Ausleseverstärkung oder einer alternativen Empfindlichkeitsangabe ergeben sich dann die Grauwerte. Wenn das Signal in den dunklen Bereichen nicht ausreicht, kann man die Blende weiter öffnen oder mehr Licht zur Beleuchtung anbieten. Dabei muss man beachten, dass die hellen Bereiche nicht in Sättigung geraten dürfen. Auch längere Belichtungszeiten führen zu höheren Signalen, das Dunkelsignal steigt aber ebenfalls an. Man kann auch die Verstärkung erhöhen, das Rauschen an der Signalquelle wird jedoch ebenfalls verstärkt. Wenn die hellen Stellen in Sättigung sind, kann man mit denselben Mitteln das Signal reduzieren. Anschließend rechnet man aus, ob das Signal-Rausch-Verhältnis sowohl in den Lichtern als auch in den Schatten für die geplanten Algorithmen ausreicht. Falls nicht, sollte man nach einer Kamera mit höherer Quantenausbeute suchen, die Beleuchtung spektral so abstimmen, dass sie im Maximum der Quantenausbeute liegt, den Dunkelstrom reduzieren (z. B. durch kürzere Integrationszeit und abgestimmte Blitzbeleuchtung), oder ein Modell mit höherer FW und geringerem Ausleserauschen, also einem besseren Wert für den Dynamikbereich wählen. Abschließend sei ein Blick in die Norm EMVA 1288 [3] empfohlen. Dort werden weitere Effekte behandelt, die bei manchen Kameras wichtig sind, z. B. das fixed-pattern-Rauschen und in der Version 3.0 auch die Charakterisierung von Kameras mit nichtlinearer Kennlinie.

### Literatur

- [1] G. C. Holst, T. S. Lomheim, CMOS/CCD-Sensors and Camera Systems, SPIE-Press, 2nd ed. 2011
- [2] INSPECT 2/2010, S. 18, Strahlende Präsenz, Grundlagen der optischen Messtechnik: Radiometrie
- [3] [www.emva.org/standard1288](http://www.emva.org/standard1288)

## PRODUKTE

### Defekte auf matten Oberflächen finden

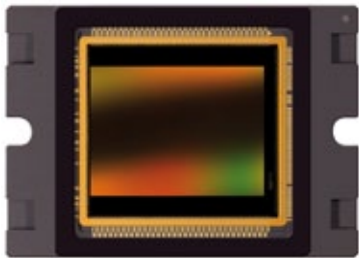
Für die Defekterkennung auf diffus reflektierenden Oberflächen bietet die INB Vision AG, ein Unternehmen der Micro-Epsilon Gruppe, das Messsystem Surface Control an. Dieses System arbeitet nach dem Verfahren der Streifenlichtprojektion. Damit können lokale Formfehler, wie Beulen und Dellen in Oberflächen erkannt und analysiert werden, die nur wenige Mikrometer vom Sollmaß abweichen. Auch genarbte Oberflächen, wie bei Interieurteilen von Automobilen üblich werden damit zuverlässig bewertet. Surface Control bietet verschiedene Messflächen in der Größe von ca. 150x100 mm<sup>2</sup> bis ca. 600x400 mm<sup>2</sup> und erfasst innerhalb weniger Sekunden die 3D-Daten der Oberfläche. Für die Auswertung stehen je nach Ausprägung der gesuchten Formabweichungen verschiedene Verfahren zur Verfügung. So kann für einen Vergleich aus den 3D-Daten eine feh-



lerfreie virtuelle Hülle berechnet und oder in Analogie zum Abziehstein in Presswerken ein digitaler Abziehstein eingesetzt werden. Diese Verfahren bieten die Möglichkeit einer reproduzierbaren, objektiven Bewertung von Abweichungen. Das Streifenlichtprojektionsverfahren eignet sich für alle Flächen, die mindestens einen Teil des Lichtes diffus reflektieren. Das sind z. B. Stahl, Aluminium, Kunststoffe oder Keramik. Bei spiegelnden

Objekten wird mit dem System Reflect Control von Micro-Epsilon auf das Messprinzip der Deflektometrie zurückgegriffen. [www.micro-epsilon.de](http://www.micro-epsilon.de)

### Massenproduktion des CMV12000 startet

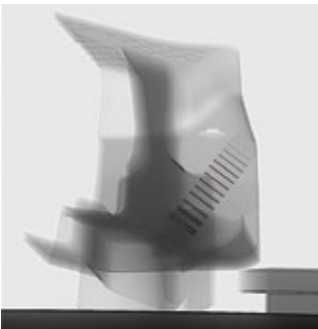


Cmosis beginnt mit der Volumenfertigung der monochromen wie der Color-Version seines 12-Megapixel Bildsensors CMV12000. Beide Ausführungen bieten Frame-Raten von 150 fps. Er bietet eine Auflösung von 12 Megapixeln (4.096 x 3.072) bei quadratischem Pixelformat mit 5,5 µm Kantenlänge. Seine Pixelgröße,

und damit die optische Performance, entsprechen den Bildsensoren CMV2000 und CMV4000. Der Spitzenwert der Quanteneffizienz erreicht mehr als 60 Prozent. Dies resultiert, zusammen mit dem Dynamikbereich von 60 dB, in einer Lichtempfindlichkeit von 4,64 V/Lux-s. Mittels Correlated Double Sampling (CDS) bei Betrieb mit Zentralverschluss reduziert die patentierte 8-Transistor Pixelzellenarchitektur das Dunkelrauschen und die FPN-Uniformitätsabweichungen der Sensormatrix. Bei der vollen Auflösung von 12 Megapixeln und mit 10-bit Auflösung des A/D-Wandlers liefert der CMV12000 eine Bildrate von 150 Full Frames pro Sekunde. Dies wird erreicht durch Verwendung von 64 LVDS-Ausgängen mit je 300 Mbit/s. Beide Varianten des Bildsensors CMV12000, als monochrome und RGB-Color-Version, sind in voller Produktion. Der Baustein kommt in einem keramischen 143-Pin PGA Package. Dies ermöglicht Betrieb im Temperaturbereich -30 bis +70°C. Der erweiterte Temperaturbereich macht den CMV12000 hervorragend geeignet für äußerst anspruchsvolle und raue industrielle Applikationen.

[www.cmosis.com](http://www.cmosis.com)

### Erhöhte Prüfqualität von Gussteilen durch Radioskopie



Martinrea Honsel Germany setzt seit Jahren Maßstäbe in der Qualitätsprüfung von Gusserzeugnissen. Dabei nimmt die Röntgenprüfung einen sehr hohen Stellenwert ein. Im Rahmen der erfolgreichen Zusammenarbeit, nutzt das Unternehmen eine Röntgenkabine des Typs Y.MU2000 von Yxlon. Diese ist derzeit mit analoger Bildanalysetechnik ausgerüstet. Bei einigen speziellen Prüfungen wurde schnell deutlich, dass die physikalischen Grenzen der Bildverstärkertechnik erreicht waren. Aus

diesem Grund entschied man sich, die Röntgenkabine mit der digitalen HDR-Bildanalyse und -auswertetechnik von Yxlon auszurüsten und das System somit auf den aktuellsten Stand der Technik zu bringen. HDR (Hoch dynamische Radioskopie) bezeichnet ein Verfahren, bei dem mittels digitalem Flachbilddetektor und einem speziellen Bildverarbeitungs-Filter ein durchleuchtetes Prüfteil in einem rauscharmen Livebild „wie aus Glas“ wirkt. Aufgrund dieser glasartigen Darstellung ist das Prüfteil mit einem Blick prüfbar, ohne die Röntgenparameter für dicke und dünne Bereiche anpassen zu müssen. Fehler werden auf Anhieb sichtbar und somit sicher detektiert. Das Unternehmen erschließt sich damit die Vorteile einer hohen Detailerkennbarkeit, Prüfsicherheit und Prüfgeschwindigkeit sowie der Erkennbarkeit aller Fehler auf einen Blick im Livebild und der Definition der räumlichen Lage und Form von Fehlern.

[www.yxlon.com](http://www.yxlon.com)

### Mobiles Fahrradmesssystem vorgestellt



In Zusammenarbeit mit dem Labor für Polymer- und Komposit-Technologie (LTC) der TH Lausanne (EPFL) hat Hexagon Metrology das Romer Bike Measurement System entwickelt, eine mobile 3D-Messlösung, die von der International Cycling Union (UCI) zur Messung von Fahrradrahmen bei Radrennen zugelassen wurde. Es basiert auf dem mobilen Romer Absolute Arm mit integriertem Laserscanner von Hexagon Metrology. Zum System gehört 3D-Ins-

pektionsoftware für benutzerunabhängige berührungslose Messungen der Geometrie des Fahrradrahmens, die für Transparenz und Einheitlichkeit im professionellen Radrennsport sorgen soll. Das System ist UCI zertifiziert. Dank der Mobilität des Systems kann die Prüfung an jedem Ort stattfinden. Während des Messvorgangs erfasst der integrierte Laserscanner pro Sekunde 30.000 Messpunkte des Rahmens. Damit wird ein hochpräzises 3D-Modell des Rahmens erstellt. Weil der Romer Absolute Arm keine Aufwärmzeit benötigt und die Software nur acht einfache Schritte für die Messung vorgibt, dauert es weniger als zehn Minuten, einen kompletten Fahrradrahmen zu messen, ihn mit dem entsprechenden CAD-Modell abzugleichen und einen abschließenden Prüfbericht zu generieren.

[www.hexagonmetrology.com](http://www.hexagonmetrology.com)

### Lasermesssysteme für die Stahlindustrie



LAP Lasermesssysteme eignen sich, um Abmessungen von Lang- und Flachprodukten, von Brammen, Knüppeln und Rohren in der Stahlproduktion zu messen. Für die gesamte Prozesskette von der Stranggießanlage bis zum fertigen Produkt stellt LAP lasergestützte Systeme für die berührungsfreie Messung von Abstand, Breite, Dicke, Länge, Durchmesser, Kontur und Ebenheit her. Mit dem Einsatz von LAP Messsystemen können darüber hinaus unnötige Zugaben, etwa zum Brammengewicht, vermieden werden. Diese helfen das Metergewicht zu garantieren, beispielsweise in der Blechproduktion.

[www.lap-laser.com](http://www.lap-laser.com)

### 3D-Technologie für die Fabrikautomation

LMI erweitert die Gocator Produktlinie durch zwei neue Sensor-Serien. Die 1000er-Serie, bestehend aus Hochgeschwindigkeitswegsensoren und die 3000er-Serie, bestehend aus 3D-„Snapshot“-Scannern. Alle Gocator Sensoren sind eigenständige, intelligente Sensoren mit Web-Schnittstelle, integrierten Messinstrumenten und verschiedenen I/O-Schnittstellen. Sie sind einfach über gängige Webbrowser konfigurierbar. Die Sensoren der 1000 Serie stellen die nächste Generation von Single-Point Messsensoren dar. Sie verfügen über eine hohe Messgeschwindigkeit und sind geeignet für die berührungslose Messung von Parametern wie Dicke, Höhe, Breite, Länge und Position. Die Sensoren der 3000 Serie verwenden Stereoskopie in Verbindung mit kodiertem, strukturiertem Licht, um komplette 3D-Flächen auf einmal zu erfassen. Sie eignen sich für Anwendungen, die eine Auswertung von 3D-Punktwolken stationärer Objekte erfordern. „Power User“ steht eine quelloffene Programmierschnittstelle (SDK) zur Verfügung.

[www.lmi3d.com](http://www.lmi3d.com)

**LED-Beleuchtungen  
für die Bildverarbeitung**

[www.falcon-illumination.de](http://www.falcon-illumination.de)

**FALCON**

FALCON ILLUMINATION MV GMBH & CO.

**LED-Beleuchtungen made in Germany**

●●IMAGING●LIGHT●TECHNOLOGY

**BÜCHNER**

[www.buechner-lichtsysteme.de/inspect](http://www.buechner-lichtsysteme.de/inspect)





## INDEX

<b>Allied Vision Technologies</b>	19, 57	<b>IDS Imaging Development Systems</b>	9, 38	<b>Schäfer + Kirchhoff</b>	31, 61
<b>Basler</b>	5, 37	<b>Isra Vision</b>	38	<b>Sony Europe</b>	39
<b>Baumer</b>	42, 57	<b>JAI</b>	25, 37	<b>Teledyne Dalsa</b>	38, 4, US
<b>Büchner Lichtsysteme</b>	77	<b>Kappa optronics</b>	35	<b>The Imaging Source Europe</b>	39
<b>Carl Zeiss</b>	31	<b>Keyence Deutschland</b>	BL	<b>Vision &amp; Control</b>	39, 57
<b>Chromasens</b>	56	<b>Kowa Optimed Deutschland</b>	29, 37	<b>Vision Components</b>	11, 38
<b>Cmosis NV</b>	56, 77	<b>LAP Laser Applikation</b>	73, 77	<b>Vision Ventures</b>	57
<b>EVT Eye Vision Technology</b>	39	<b>LMI Technologies</b>	59, 77	<b>Visuelle Technik</b>	55
<b>Falcon Illumination mv</b>	77	<b>Micro-Epsilon Messtechnik</b>	3, 76	<b>VRmagic</b>	37, 38
<b>Faro Europe</b>	39	<b>Mikrotron</b>	39	<b>Ximea</b>	37, 2, US
<b>Framos</b>	57	<b>Point Grey Research</b>	15, 38	<b>Xlcon International</b>	45, 77
<b>Hexagon Metrology</b>	67, 77	<b>Polytec</b>	23, 31		
<b>Hochschule Darmstadt</b>	6	<b>Qioptiq Photonics</b>	57		

## IMPRINT

### Herausgeber

GIT VERLAG  
Wiley-VCH Verlag GmbH & Co. KGaA  
Boschstraße 12, 69469 Weinheim, Germany  
Tel.: +49/6201/606-0

### Redaktion

Bernhard Schroth (Chefredakteur)  
Tel.: +49/6201/606-753  
bernhard.schroth@wiley.com

Anke Grytzka  
Tel.: +49/6201/606-771  
anke.grytzka@wiley.com

Stephanie Nickl  
Tel.: +49/6201/606-738  
stephanie.nickl@wiley.com

Andreas Grösslein  
Tel.: +49/6201/606-718  
andreas.groesslein@wiley.com

### Redaktionsassistentz

Bettina Schmidt  
Tel.: +49/6201/606-750  
bettina.schmidt@wiley.com

### Beirat

Roland Beyer, Daimler AG

Prof. Dr. Christoph Heckenkamp;  
Hochschule Darmstadt  
Gabriele Jansen, Vision Ventures  
Dipl.-Ing. Gerhard Kleinpeter, BMW Group  
Dr. rer. nat. Abdelmalek Nasraoui,  
Gerhard Schubert GmbH  
Dr. Dipl.-Ing. phys. Ralph Neubecker, Schott AG

### Anzeigenleitung

Oliver Scheel  
Tel.: +49/6201/606-748  
oliver.scheel@wiley.com

### Anzeigenvertretungen

Claudia Brandstetter  
Tel.: +49/89/43749678  
claudia.brandst@t-online.de

Manfred Höring  
Tel.: +49/6159/5055  
media-kontakt@t-online.de

Dr. Michael Leising  
Tel.: +49/3603/893112  
leising@leising-marketing.de

### Herstellung

Christiane Potthast  
Claudia Vogel (Sales Administrator)  
Ruth Herrmann (Layout)  
Elke Palzer, Ramona Kreimes (Litho)

### Leserservice / Adressverwaltung

Marlene Eitner  
Tel.: 06021/606-711  
marlene.eitner@wiley.com

### Sonderdrucke

Oliver Scheel  
Tel.: +49/6201/606-748  
oliver.scheel@wiley.com

### Bankkonto

Commerzbank AG, Darmstadt  
Konto-Nr. 01.715.50100,  
BLZ 50880050

Druckauflage: 20.000

Zurzeit gilt die Anzeigenpreisliste vom 1. Oktober 2011  
2012 erscheinen 7 Ausgaben  
„INSPECT“  
Druckauflage: 20.000  
(1. Quartal 2012)

### Abonnement 2012

7 Ausgaben EUR 45,00 zzgl. 7 % MWSt  
Einzelheft EUR 14,50 zzgl. MWSt+Porto

Schüler und Studenten erhalten unter Vorlage einer gültigen Bescheinigung 50 % Rabatt.

Abonnement-Bestellungen gelten bis auf Widerruf; Kündigungen 6 Wochen vor Jahresende. Abonnement-Bestellungen können innerhalb einer Woche schriftlich widerrufen werden, Versandreklamationen sind nur innerhalb von 4 Wochen nach Erscheinen möglich.

### Originalarbeiten

Die namentlich gekennzeichneten Beiträge stehen in der Verantwortung des Autors. Nachdruck, auch auszugsweise, nur mit Genehmigung der Redaktion und mit Quellenangabe gestattet. Für unaufgefordert eingesandte Manuskripte und Abbildungen übernimmt der Verlag keine Haftung.

Dem Verlag ist das ausschließliche, räumlich, zeitlich und inhaltlich eingeschränkte Recht eingeräumt, das Werk/den redaktionellen Beitrag in unveränderter Form oder bearbeiteter Form für alle Zwecke beliebig oft selbst zu nutzen oder Unternehmen, zu denen gesellschaftsrechtliche Beteiligungen bestehen, so wie Dritten zur Nutzung zu übertragen. Dieses Nutzungsrecht bezieht sich sowohl auf Print- wie elektronische Medien unter Einschluss des Internets wie auch auf Datenbanken/Datenträgern aller Art.

Alle etwaig in dieser Ausgabe genannten und/ oder gezeigten Namen, Bezeichnungen oder Zeichen können Marken oder eingetragene Marken ihrer jeweiligen Eigentümer sein.

### Druck

Frotscher Druck  
Riedstr. 8, 64295 Darmstadt

Printed in Germany  
ISSN 1616-5284



**www.inspect-online.com**

Die

**INSPECT**

ist

**online**

Das

**INTERNET-  
PORTAL**

für

- INDUSTRIELLE BILDVERARBEITUNG
- OPTISCHE MESSTECHNIK
- CONTROL

**online**



**INSPECT**

# DIE VIELSEITIGSTE GIGE VISION-KAMERA DER WELT IST DA

- 16 Modelle (VGA-bis-12MP)
- Verschiedene Belichtungszeiten für jedes einzelne Bild
- Auto-Iris (Video, DC oder motorisiert)
- Auto-Helligkeitsfunktion
- Erweiterter Dynamikbereich
- Integrierte Bildkompression
- Synchronisierung mehrerer Kameras über IEEE 1588
- Bildfilterkernel



**GiGE**  
VISION

## Genie TS VGA > 1.2M > 1.4M > 2M > 4M > 5M > 8M > 12M

Die **Genie TS** nimmt Bilder mit bis zu 12 Megapixel mit extrem hoher Qualität und Auflösung auf und erreicht Bildfrequenzen von bis zu 300 Bilder/s. Integrierte Funktionen wie Mehrfachbelichtung, Farbraumkonvertierung und Metadaten mit Bildtransfer "on-demand" sind in ein kompaktes Gehäuse eingebaut. Die Schnittstelle zum PC ist per GigE Vision-Standard implementiert.

## Mehr über Vision.

**Laden Sie unser technisches Dokument herunter:**

4 entscheidende Faktoren:

Verwendung von Gig E Vision in der industriellen Realtime-Bildverarbeitung (Englisch)

[www.teledynedalsa.com/a/i2012](http://www.teledynedalsa.com/a/i2012)

