

Bildverarbeitung

Vorlesung (2+2)

Vorlesung: 2 SWS
Praktikum: 2 SWS

Dozent: Thomas Haenselmann

1. Kickoffveranstaltung vor Ort

Danach online hier:

<https://b-lb.hs-mittweida.de/b/rooms/65y-0eo-wim-sak/join>

It. Stundenplan. Dieses Semester neuer Modus Operandi:

Vorlesung und Praktikum in beiden Terminen gemischt.

Idee: Wissen dann verfestigen, wenn es noch frisch im Kopf ist.



Klaus D. Tönnies
**Grundlagen der
Bildverarbeitung**

2005, Pearson Studium
ISBN 3-8273-7155-4

Foley, vanDam, Feiner, Hughes
Computer Graphics:
Principles and Practice
2nd Edition oder später
Addison-Wesley
(nur f. Farbräume)

Bernd Jähne
Digitale Bildverarbeitung
6. Auflage, Springer Verlag
ISBN 3-540-24999-0

(vertiefend, insbes. zu digit. Filtern, über die
Hochschulbib. online als Volltext verfügbar!)

Nicht alles durch Lehrbücher abgedeckt.
Nutzen Sie

- die Videos
- verlinkte Originalliteratur
- Bibliothek, Google Scholar, KI, bel.
Eigeninitiative erlaubt!

Prokrastination

Viele Menschen schieben Aufgaben gern auf.

Bis zur Klausur ist noch so viel Zeit, keine Panik.

- Nächste Woche hat sich der Stoff schon verdoppelt, übernächste verdreifacht.
- Lassen Sie dem Rückstand keine Chance. Lernen Sie den Stoff dieser Woche in dieser Woche.

Heute lernen fühlt sich komisch an. Morgen bin ich aber hochmotiviert.

- Die Erfahrung zeigt: Morgen ist die Motivation so hoch wie heute, die Arbeit hat sich lediglich vermehrt.
- Bleiben Sie realistisch: Jetzt sofort 20 Minuten oder 1/2 Stunde lernen ist vielleicht nicht genug aber viel besser als aufschieben.

Themen im Überblick

- Bildgebende Geräte, Kamera-Rohdaten
- Farbräume
- digitale Kurven
- Neuronale Netze
- Die Hough-Transformation
- Kurvenrepräsentationen in Skalenräumen
- Objektsegmentierung
- Fouriertransformation, Basistransformationen, DCT, Idee der Wavelets
- digitale Filter, Gradienten
- Histogramme und Entropie von Grauwertverteilungen
- Textur (falls Zeit)
- (Kalmanfilter falls Zeit)

Digitale Bildverarbeitung

- 2D/3D Bild wird als Eingangssignal aufgefasst
- Dann folgt eine Form der Signalverarbeitung
- Ausgabe ist abgeleitetes Bild/ Merkmale des Bildes

Bildbearbeitung

- Ziel ist eher die (evtl. künstlerische) Manipulation des Bildes
- Meist mit Hilfe eines Grafikprogramms (z.B. GIMP*), das sich oft Verfahren der Bildverarbeitung bedient

(*) www.gimp.org

Maschinelles Sehen (Computer Vision)

- Eingabe ebenfalls Bild, Ziel jed. Bedeutung von Bildinhalten zu finden (was ist zu sehen)
- Als Eingabe dienen häufig vorverarbeitete Bilder aus der Bildverarbeitung

Computergrafik

- Ziel ist Erzeugung von Grafik
- Nicht nur Pixel- sondern auch 2D Vektor- und 3D Grafik
- Computational Photography
- Bilder aus Bildern berechnen (z.B. Panoramabilder, Image Morphing)

Überschneidungen zw. den vier Gebieten sind möglich.

Bezug zur Forensik

Bezug zur Forensik

Reversible Transformationen

Beispiel: Deformationen invertieren

Das durch einen Bildstrudel verfremdete Gesicht des Verdächtigen konnte durch die inverse Funktion weitgehend rekonstruiert werden.

Problem der hochspezifischen Anwendungen: Sie treten kaum wiederholt auf.



Strudel als Beispiel einer Bildmanipulation, die (ausreichend) reversibel ist..

Bezug zur Forensik

Copy-Move Attacken

Ein-/Ersetzen von Bildausschnitten

Analoges historisches Beispiel (links), digitales (rechts)



Die Light Sphere (Idee)

Nachträgliche Anpassung bereits gedrehten Materials:

Hohe Framerate mit unterschiedlichen Beleuchtungsrichtungen/Farben, die später interpoliert werden können.

Zweite Kamera im gleichen optischen Brennpunkt berechnet infrarot-basierte Alphamap.



P. Debevec, A. Wenger, C. Tchou et al.
A Lighting Reproduction Approach to Live-Action Compositing
SIGGRAPH 2002, San Antonio, July 21-26, 2002

Bezug zur Forensik

Image compositing

Die Light Sphere (Idee)

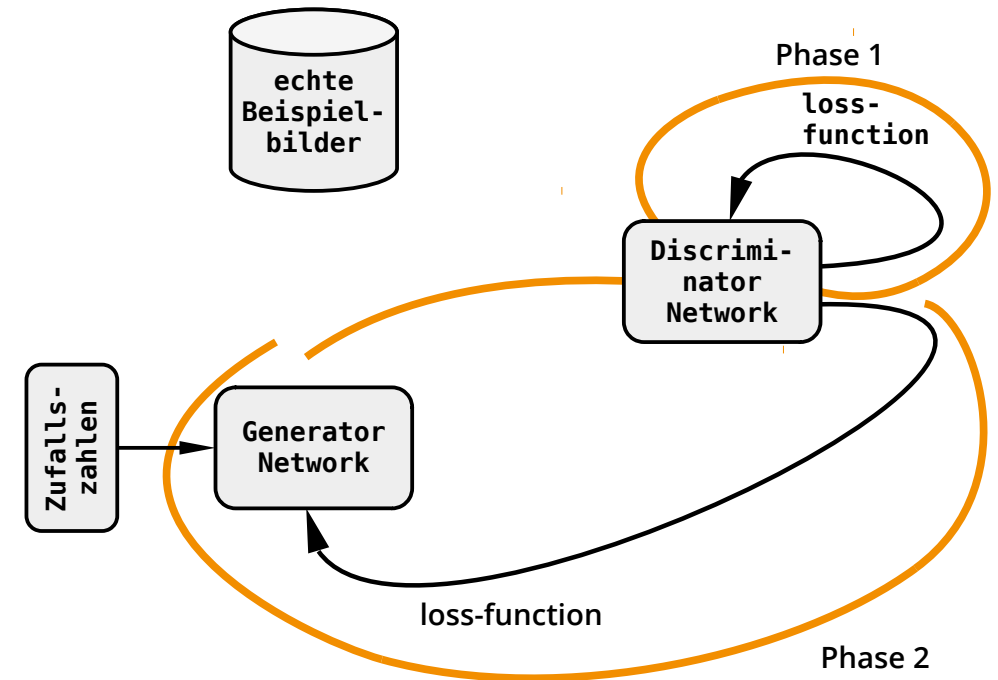


Bezug zur Forensik

Generative Adversarial Networks

GANs

(Lern-)Phase 1 verbessert die Fähigkeit des Discriminators zur Unterscheidung echter Bilder von erzeugten, Phase 2 die des Generators zur Erzeugung von guten synthetischen Bildern. Beide Phasen wechseln sich ab.



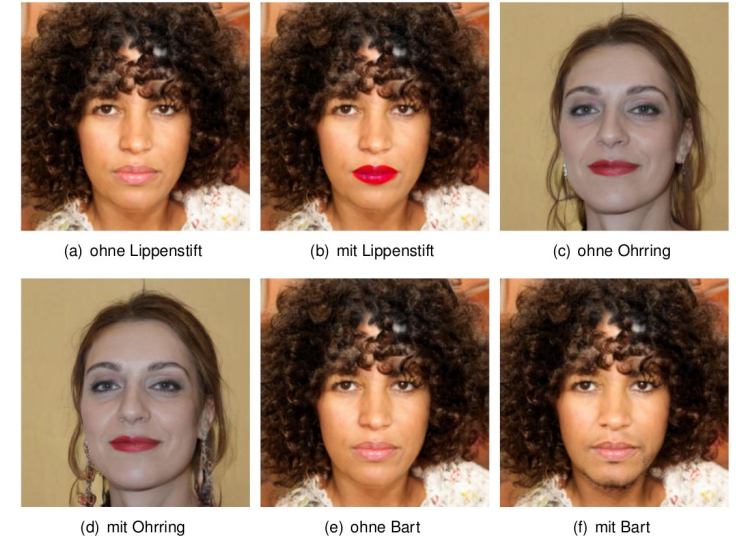
Aus Wikipedia (Artikel: GAN) unter Creative Commons Lizenz (User: Bod Inga klang)

Bezug zur Forensik

Generative Adversarial Networks

StyleGAN

Ein Open Source Projekt von Nvidia zur Generierung von Gesichtsbildern mittels eines vortrainierten GANs, gesteuert durch einen Stylevektor.



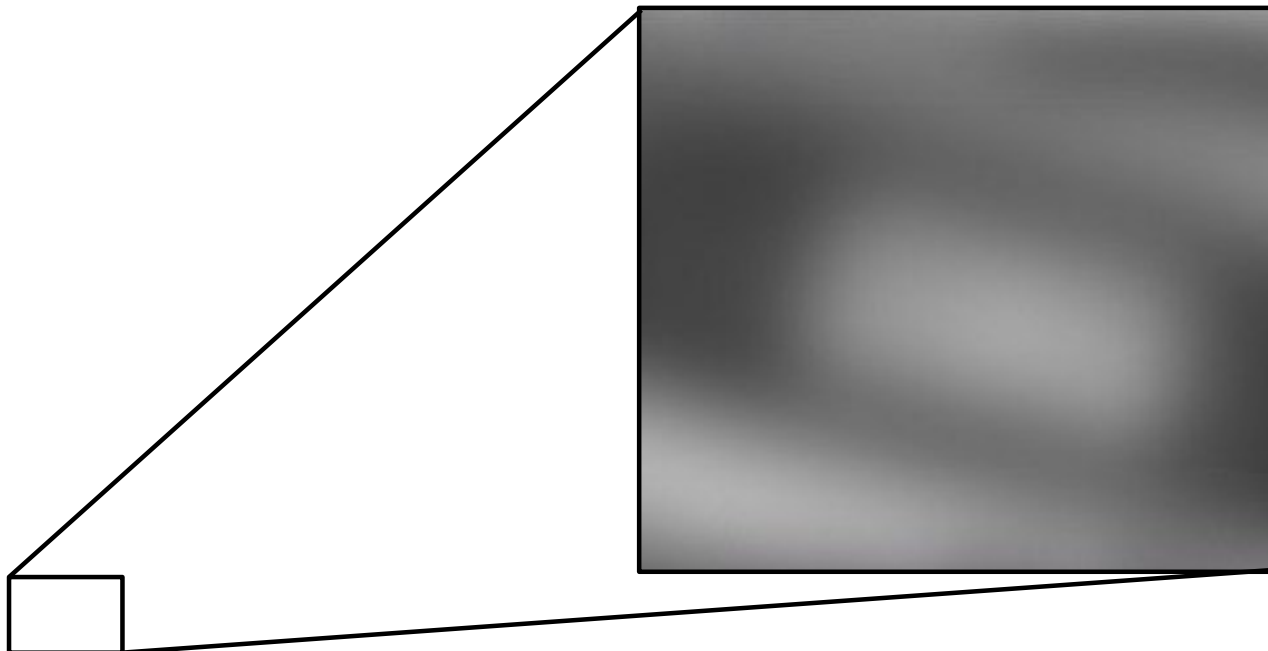
Bezug zur Forensik

Video Superresolution



Bezug zur Forensik

Video Superresolution



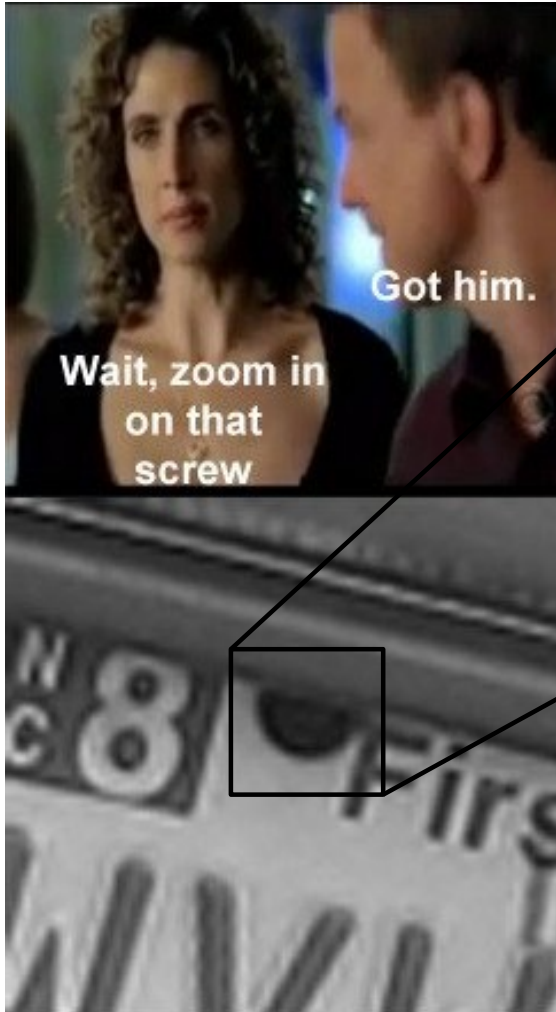
Bezug zur Forensik

Video Superresolution



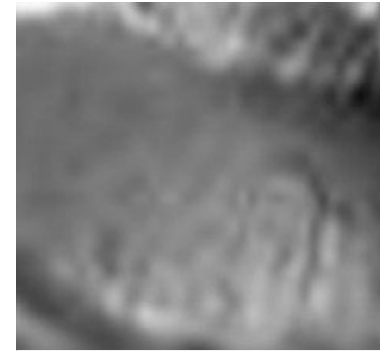
Bezug zur Forensik

Video Superresolution



Bezug zur Forensik

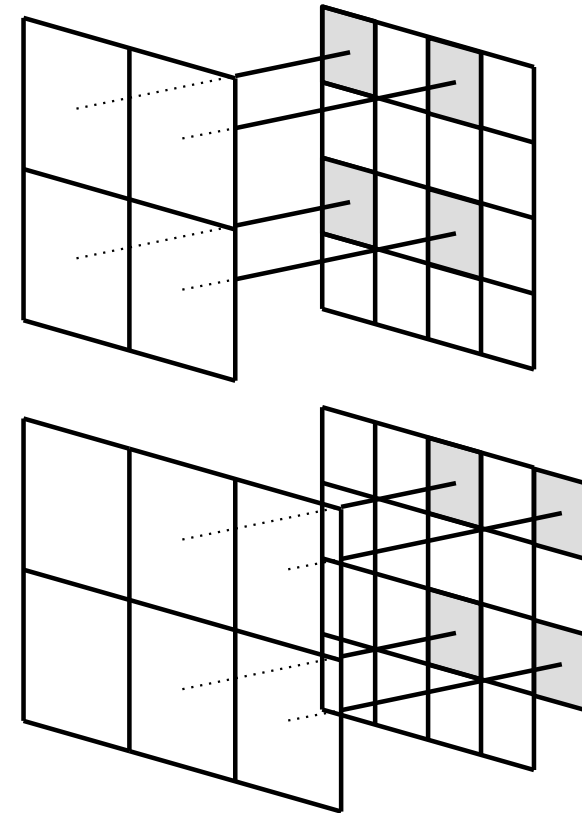
Video Superresolution



Bezug zur Forensik

Video Superresolution

Projektion eines realen Bildes auf ein 4-fach besser aufgelöstes. Die Verschiebung der Kamera um einen ganzen Pixel erzeugt eine weitere Spalte. Sie erweitert das hochaufgelöste Bild, bringt aber keine höhere Detailtiefe.

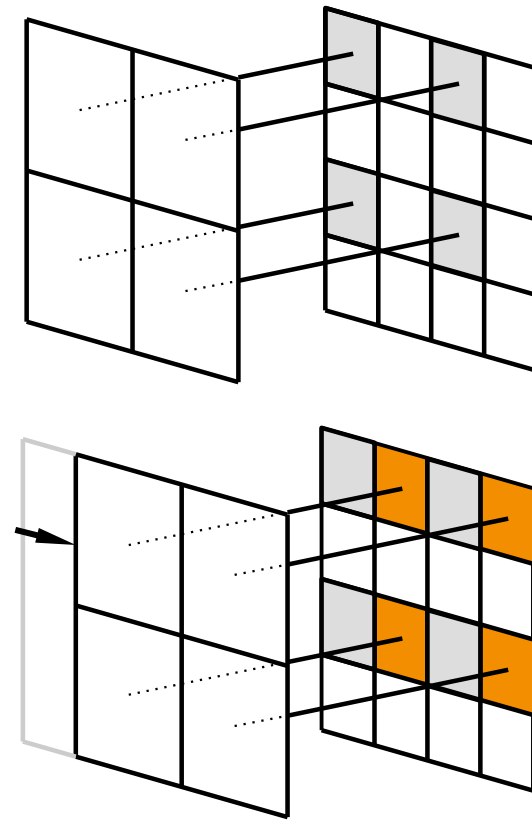


Bezug zur Forensik

Video Superresolution

Die Verschiebung der Kamera um einen Sub-Pixel (hier $1/2$ Pixel in X-Richtung) fügt dem hoch aufgelösten Bild **neue** Information hinzu.

Um die Verschiebung der Kamera zu berechnen bedient man sich unter anderem dem dem sog. *Optischen Fluss*.



Typische Anwendungen

Anwendungen

Med. Anwendungen

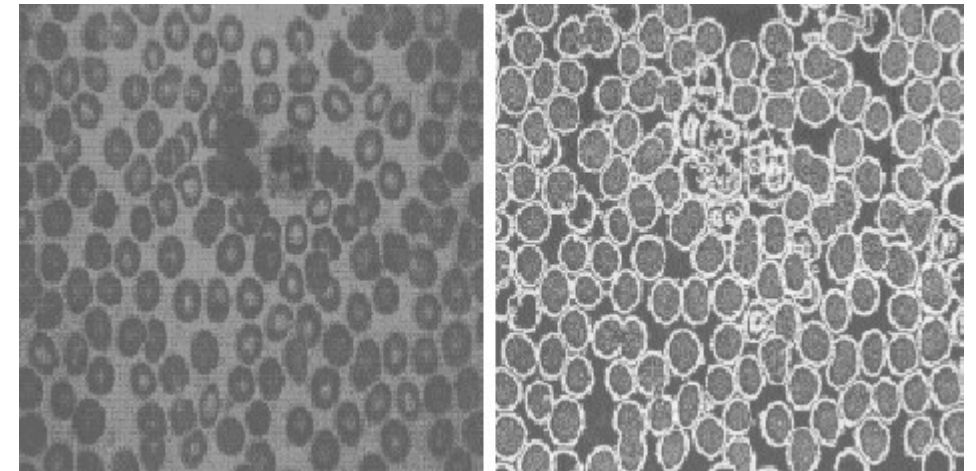
Bildvorbereitung zum Zählen von roten Blutkörperchen. Meist nutzt man bekannte Voraussetzungen aus:

Aufgabenteil Bildverarbeitung:

- Verbessere Bild für die Zählung der Zellen z.B. durch *Berechnung des Gradientenbildes*.

Aufgabenteil Computer Vision:

- Identifiziere und zähle Zellen bei bekannter Varianz in der Größenstreuung (z.B. +/- 20%)



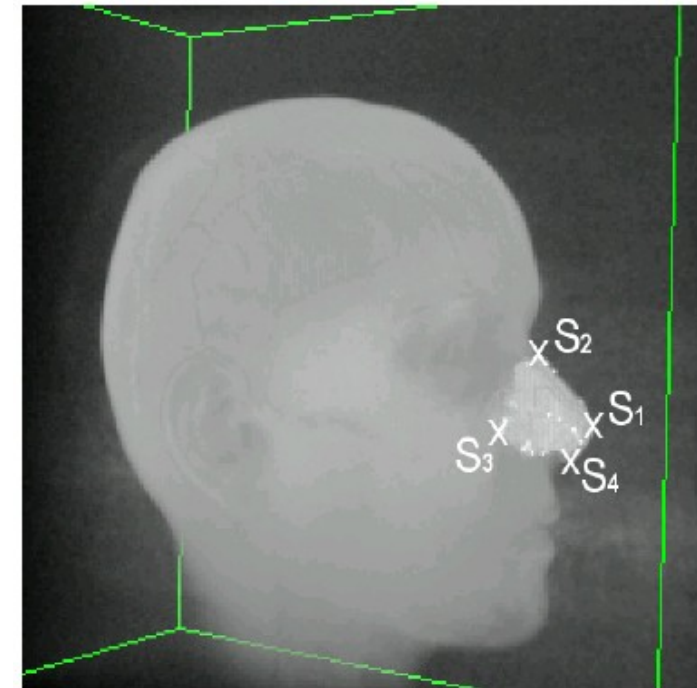
Anwendungen

Med. Anwendungen

Medizinische Volumendaten

Viele Algorithmen lassen sich auch auf mehrdimensionale Signale (hier 3D Volumendaten) anwenden. In 3D werden dabei nicht Pixel, sondern Voxel (Volume Elements) verarbeitet.

Einige Algorithmen können auf einer „Scheibe“ im Volumen arbeiten => klassische 2D Bildverarbeitung. Einige Operatoren müssen auf die dritte Dimension angepasst werden und einige Fragestellungen in 3D unterscheiden sich von 2D Bildverarbeitung.



Segmentierung in 3D Volumen

Anwendungen

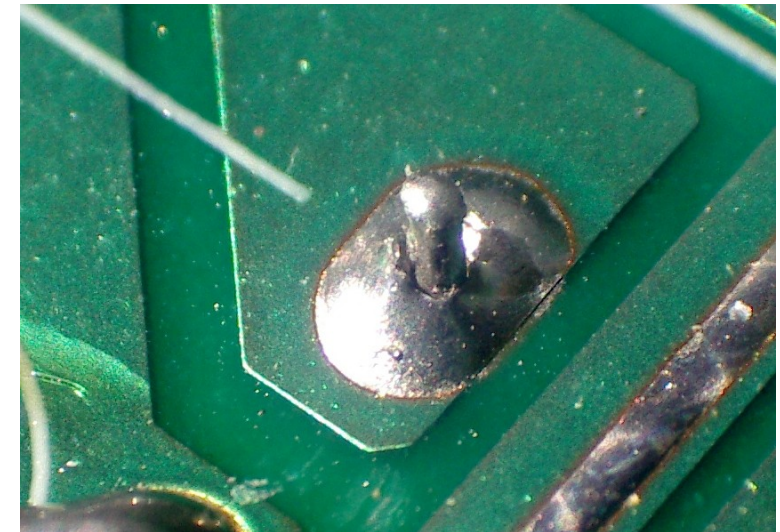
Industrielle Anwendungen

Industrielle Bildverarbeitung

Elektrische Prüfungen erkennen nicht alle Fehler (z.B. kalte Lötstellen)

Zusätzliche visuelle Prüfung kann helfen die Qualität zu verbessern.

Charakteristisch für industrielle Anwendung: Beleuchtung, Kamerawinkel etc. können so optimiert werden, dass die Aufgabe für die Bildverarbeitung so einfach wie möglich wird (z.B. optimale Perspektive, Spiegelung von Licht nutzen/vermeiden). Erst dann müssen die Algorithmen optimiert werden.



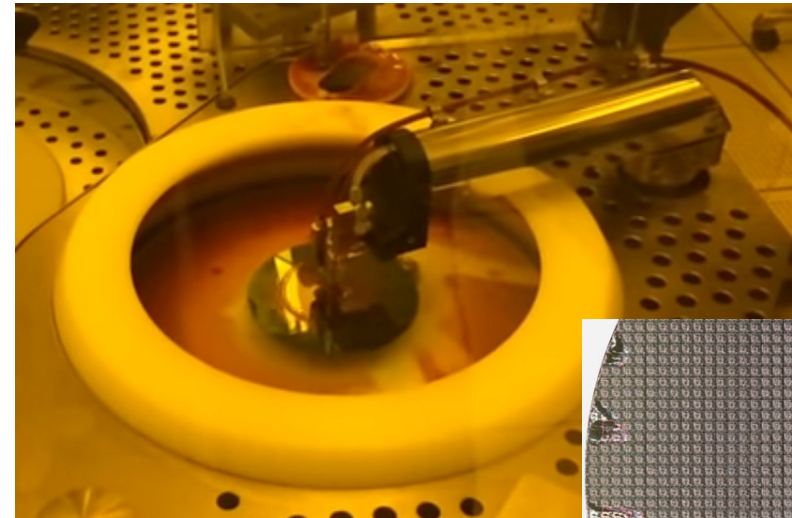
Kalte Lötstelle: Elektrisch manchmal nicht feststellbar, optisch eindeutiger zu erkennen.

Anwendungen

Industrielle Anwendungen

Waferfertigung

Visuelle Prüfung des Belackungsprozesses in der Waferfertigung. Im ungünstigsten Fall hinterlässt der Lack am Rand trockene Stellen, die viele Dies auf den äußeren Radien wertlos machen würden. Erkannte fehlerhafte Wafer können abgewaschen und neu belackt werden.



Auftrag von Lack in der Waferfertigung.



Ungleichmäßig verlaufener Lack.

Bildgebende Verfahren

Zeilenkamera/Zeilenscanner

Findet Anwendung in

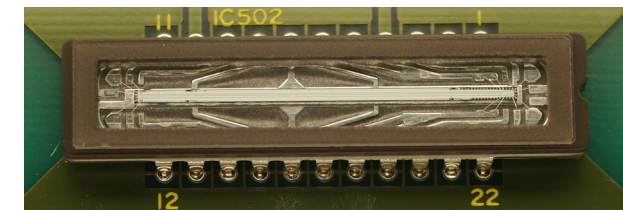
- Scannern, Kopierern, Barcode Scannern (EAN 13) im Einzelhandel, früher Handscanner
- Erzeugung von Luft- und Satellitenbildern
- Industrielle Bildverarbeitung an Produktionsbändern

Spaltenauflösung > 10k Pixel möglich, Zeilenauflösung abhängig von Auslesegeschwindigkeit und Vertikalbewegung.

Ungleichmäßige Vertikalbewegung (Flugzeug wackelt, Handscanner läuft ungleichmäßig) => Bild in eine Richtung verzerrt. Synchronisation von Abtastung mit Bewegung nötig.



Handscanner mit Zeilensensor (*1)



Einzlener Zeilensensor (*2)

(*1) Aus Flickr.com, User BJTsunami

(*2) Aus Wikipedia (Artikel: Zeilensensor) unter Creative Commons Lizenz

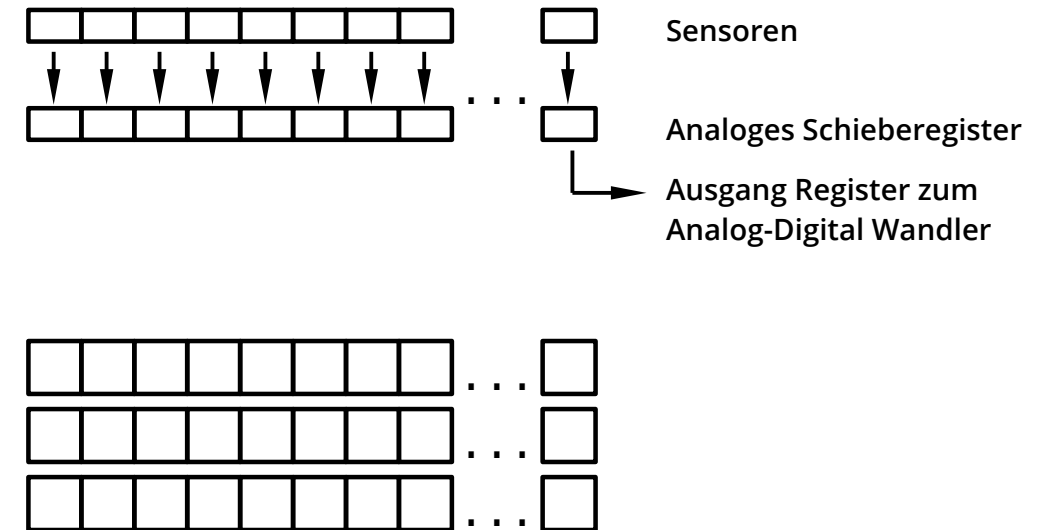
Bildgebung

Zeilensensoren

Die Zeile besteht aus einem Array lichtempfindlicher Sensoren und einem Analogen Schieberegister.

Lichtempfindlichkeit Pixel auf CCD Sensor abhängig von Größe.
Kritisch bei Zeilensensoren mit 10k Pixeln und mehr. Lösung:

Hintereinander schalten von mehreren CCD Sensoren und Mitteln der ausgelesenen Werte:
Bei n Zeilen reduziert sich dabei der Fehler um den Faktor $\sqrt{n}$.



(*1) Aus Flickr.com, User BJTsunami

(*2) Aus Wikipedia (Artikel: Zeilensensor) unter Creative Commons Lizenz

Bildgebung

1-Pixel Kamera

Die 1-Pixel Camera (Rice-University)

Ganze Bilder werden nur mittels einer einzelnen Photodiode aufgenommen.

Vorteil: Photodiode für unterschiedliche Spektren leicht herzustellen

(Infrarot, Röntgen, Funkwellen) – als 2D Bildsensor nur schwer realisierbar

Nachteil: 1-Pixel Bilder langweilig

Lösung: Brauche Spiegel-Array

Besteht aus einer Matrix kleiner Spiegel (Digital Micromirror Device DMD). Jeder Spiegel kann in eine Richtung ausgelenkt werden.



Micro-Spiegelarray wie z.B. in Beamern verbaut

Aufbau der 1-Pixel Kamera

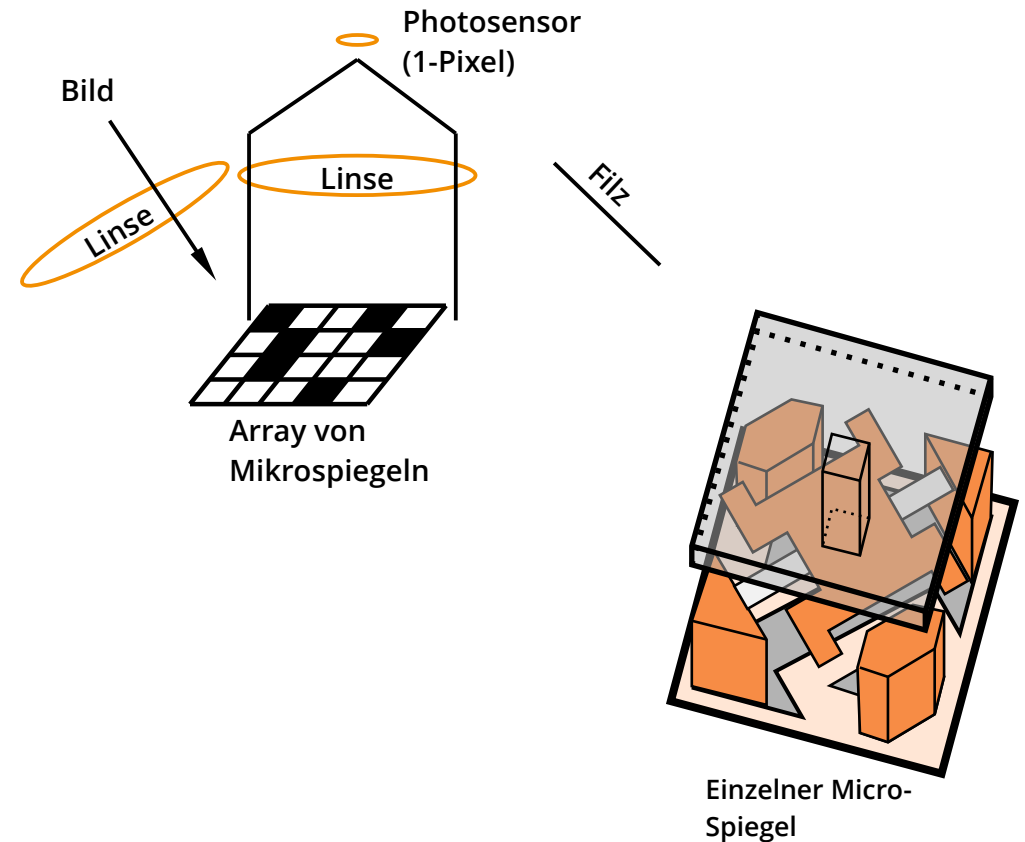
Bild wird durch Linsen(-system) auf den Spiegel-Array fokussiert. Eine zweite Linse bündelt das ausfallende Licht auf den Photosensor.

Ein Spiegel kann seinen Strahl auf den Photosensor lenken oder auf ein schwarzes Filz.

Spiegel-Array

Einzelne Mikrospiegel auf einer drehbaren Achse aufgehängt. Drehung geht nur ganz oder gar nicht, jedoch sehr schnell.

(Halbspiegel durch schnelles Wechseln der Spiegelstellung möglich.)



Bildgebung

1-Pixel Kamera

Wie entsteht das Bild? (vereinfachte Idee)

Fast alle Spiegel lenken ihre einfallenden Lichtstrahlen auf das Filz, außer dem ersten. Die Photodiode „sieht“ nun das Licht des ersten Spiegels bzw. Pixels. In der nächsten Messung wird nur der zweite Pixel korrekt ausgelenkt usw.

Folge: Wir brauchen so viele Messungen wie Pixel.

Mathematische Notation: Die Einheitsmatrix.

Das Bild lässt sich unmittelbar aus den Messungen ablesen. D.h, die erste Messung entspricht auch dem ersten Pixel etc.

$$\begin{array}{c} \text{Spiegelstellung} \\ \left(\begin{array}{cccccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \dots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right) \times \begin{array}{c} \text{Bild} \\ \left(\begin{array}{c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \right) \end{array} = \begin{array}{c} \text{Messung} \\ \left(\begin{array}{c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \right) \end{array} \\ \text{Photosensor} \end{array}$$

Bildgebung

1-Pixel Kamera

Erweiterte Variante: Erzeuge eine **zufällige** Pixelstellung und speichere diese als Zeile einer Matrix. Nimm einen Messwert. Erzeuge weitere zufällige Pixelstellungen.

Löse das Gleichungssystem nach den einzelnen Pixeln auf – wenn möglich.

Problem: Brauche immer noch mindestens so viele Gleichungen (Bedingungen) wie Pixel – eher etwas mehr.

$$\begin{array}{c} \text{Spiegelstellung} \\ \left(\begin{array}{cccccccccccc} 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ \dots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \times \begin{array}{c} \text{Bild} \\ \left(\begin{array}{c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \right) = \begin{array}{c} \text{Messung} \\ \left(\begin{array}{c} \cdot \\ \cdot \\ \cdot \\ \cdot \end{array} \right) \end{array} \end{array} \quad \begin{array}{c} \text{Photosensor} \end{array}$$

Bildgebung

1-Pixel Kamera

Eigentliches Verfahren

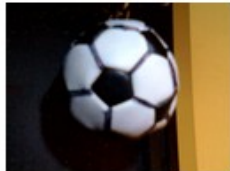
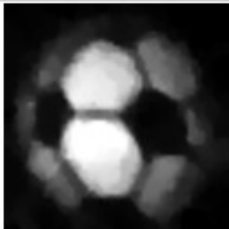
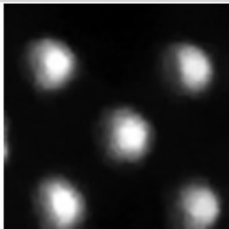
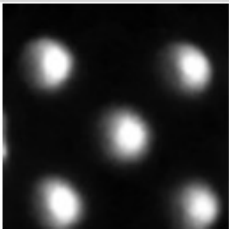
(nach der Idee des *Compressive Sensing*):

Verfahren macht erst Sinn, wenn ich mit wenigen Messungen auskomme.

Lösung: Löse das Gleichungssystem als unterdefiniertes GS. Dazu braucht man ein Optimierungskriterium.

Kriterium: Ein Pixel soll von seinem Nachbarn so wenig wie möglich abweichen. Dies gilt auch für natürliche Bilder.

Dadurch entsteht unter den gegebenen Bedingungen (wenigen Messungen) die wahrscheinlichste Lösung.

		
Original Object	4096 Pixels 800 Measurements (20%)	4096 Pixels 1600 Measurements (40%)
		
Original Object	4096 Pixels 800 Measurements (20%)	4096 Pixels 1600 Measurements (40%)

Rice Single-Pixel Camera Project (compressive sensing)
<https://archive.org/details/cscameradata>

Bildgebung

1-Pixel Kamera

Der Spiegelarray sei im folgenden Beispiel winzig. Er enthalte insgesamt nur $2 \times 2 = 4$ Pixel.

- a) Wie viele Bilder muss die 1-Pixel Kamera machen, damit ein Bild (vollständig) ermittelt werden kann?
- b) Wie sieht die Matrix aus, wenn die Kamera pro Bild nur genau 1 Spiegel durchschaltet (d.h. dessen Licht auf die Fotozelle ablenkt)?

- c) Bisher wurden drei 1-Pixel Bilder mit diesen drei Pixelstellungen gemacht:

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ ? & ? & ? & ? \end{pmatrix}$$

Finden Sie eine vierte Spiegelstellung, die nicht hilfreich wäre das Bild (perfekt) zu rekonstruieren.

- d) Erläutern Sie, welche Eigenschaft die letzte Spiegelstellung haben sollte, damit ein Bild rekonstruiert werden kann.

Bildgebung

1-Pixel Kamera

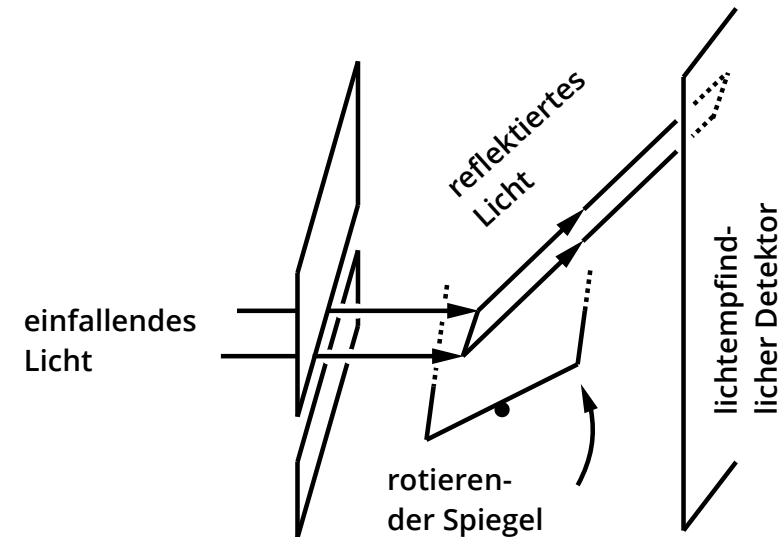
- e) Finden Sie nun eine Spiegelstellung die dazu geeignet ist das Bild zu rekonstruieren. Wie sind Sie darauf gekommen? Sie können ein Online Tool für eine Rechnung verwenden.
- f) Ermitteln Sie die Transformation die vier Messungen wieder in Pixelwerte zurück transformiert. Sie können ein Online Tool für die Rechnung verwenden.
- g) Könnte es noch eine weitere Spiegelstellungen geben?

Femto-Fotographie

Aufnahme von Bildern mit sehr kurzen Belichtungszeiten (1 Femtosekunde = 1×10^{-15} = eine Billionstel-Sekunde)

Anwendung: Visualisieren

- der Verbreitung von Licht oder
- von schnellen Schmelzprozessen beim Lasersintern (3D Druck).

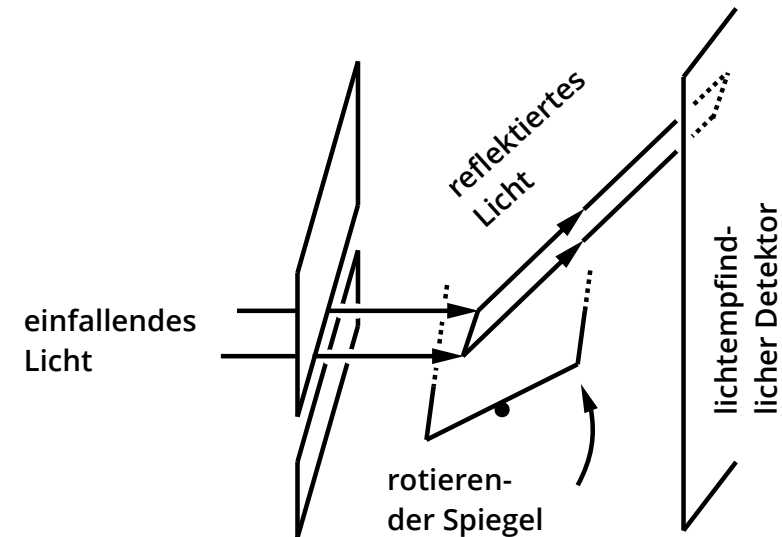


Idee einer superschnellen Kamera

Einfallendes Licht wird durch einen schnell rotierenden Spiegel abgelenkt und trifft auf eine Detektor-Oberfläche.

Früh eintreffendes Licht trifft auf einen flachen Spiegel und wird weiter unten auf den Detektor reflektiert. Später ankommendes Licht trifft durch einen schon weiter rotierten Spiegel höher auf den Detektor.

Unterschiedliche Zeitpunkte auf dem Lichtstrahl werden somit in unterschiedliche Höhen übersetzt.



Idee einer superschnellen Kamera

Somit „rast“ der Strahl aber nur einmal über den Detektor auf. Der Beobachter würde das nicht wahrnehmen. Daher wiederholt man das Experiment jedes mal, wenn der Spiegel wieder in der Anfangsstellung angekommen ist.

Interpretation des Detektorbildes: Wir sehen eine Zeile eines Kamerabildes (x-Richtung) zu unterschiedlichen Zeiten (y-Richtung). So kann jedoch nur eine Bildzeile aufgenommen werden.

Wie erstellt man Videos als Vollbild?

Prinzip: Das Experiment wird wiederholt. Dabei wird in jedem Experiment die Kamera weiter nach oben versetzt (technisch gelöst durch einen weiteren Spiegel). Die einzelnen Zeilenvideos werden dann zusammengesetzt.

Weiteres Problem: Man benötigt einen extrem kurzen Lichtimpuls. Dies leistet ein sog. Femto-Laser.

Weitere Anwendung: Sehen um die Ecke

Bildgebung

Femtofotographie

Femto-Fotographie

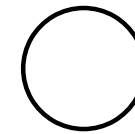
Fotographie um die Ecke

Gesendet wird ein kurzer Laserimpuls der z.B. auf eine diffus reflektierende Wand auftrifft.

Von dort aus breitet sich eine Kugelwelle in alle Richtungen aus. Diese trifft irgendwann auf ein Objekt, welches das Licht zurückwirft.

Die reflektierte Welle kann irgendwann wieder auf der Wand beobachtet werden. Die Reflexion erlaubt eine 3D-Rekonstruktion der verdeckten Objekte.

Objekt



Hindernis

