Computer Vision Project

Summer 2021

Vincent Christlein, Martin Mayr, Thomas Köhler,
Mathias Seuret, Prathmesh Madhu
November, 2020
Pattern Recognition Lab (CS 5)

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Today‘s Outline

- Organization of this project
- Motivation
- What is computer vision?
- Today‘s topic: **Image Processing for Grayscale and Distance Images**

Organization

Organization – Waiting List

- Has everyone joined the StudOn-course?
  - Course member
  - Waiting list

Organization – Grades and Certificates

- Project for Computer Science and Medical Engineering (https://lme.tf.fau.de/teaching/curriculum-courses/lv_id/21623404/)

Hochschulpraktikum	Hochschul- & Forschungspraktikum
Lectures Exercises (5 ECTS)	Lectures Exercises Research project (10 ECTS)

- Grade for Hochschulpraktikum, if neccessary, based on short talk (about a state-of-the-art research paper)
- The research project will be related to one of the discussed topics
- **Forschungspraktikum only:** contact us or any other PhD student at the lab for a research project (Overview: https://lme.tf.fau.de/teaching/thesis/)

Organization

Who are you?

Organization – Exercises

- Exercises are coded in Python
- You can work in groups of two
- Tasks are complementary to the lecture of the previous week. It is mandatory to **complete them all**
  - **At home**
  - It is strongly recommended to begin with the programming tasks before the exercise!
- Results have to be presented to one of the supervisors
- All tasks have to be finished by their individual deadlines (more later)

Organization – Exercises

- Prepare questions or problems **before** the exercise
- Zoom-Meeting:
  https://fau.zoom.us/j/94348907556?pwd=NmtDMDNjci8xRjlROVJHVDFWa0lQdz09
  Meeting-ID: 943 4890 7556
  Password: 987411
- Zoom Breakout Rooms
- Lecture and exercises on face recognition will be held via Skype.
- **Own virtual hardware**
  Virtual CIP-Pools: https://remote.cip.cs.fau.de/ (Xpra)
- You can ask additional questions via:
  Studon Forum

Organization - Topics

- Basic image processing of grayscale and distance images
- Image Processing using Seam Carving
- Writer Identification/Retrieval
- Face Recognition
- Image Recognition for Humanities

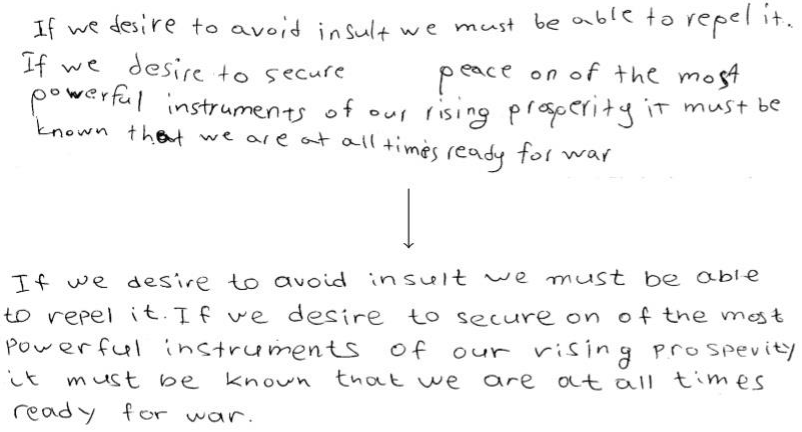
Distance image

Organization - Topics

- Basic image processing of grayscale and distance images
- Image Processing using Seam Carving
- Writer Identification/Retrieval
- Face Recognition
- Image Recognition for Humanities

Organization - Topics

- Basic image processing of grayscale and distance images
- Image Processing using Seam Carving
- Writer Identification/Retrieval
- Face Recognition
- Image Recognition for Humanities



Organization - Topics

- Basic image processing of grayscale and distance images
- Image Processing using Seam Carving
- Writer Identification/Retrieval
- Face Recognition
- Image Recognition for Humanities

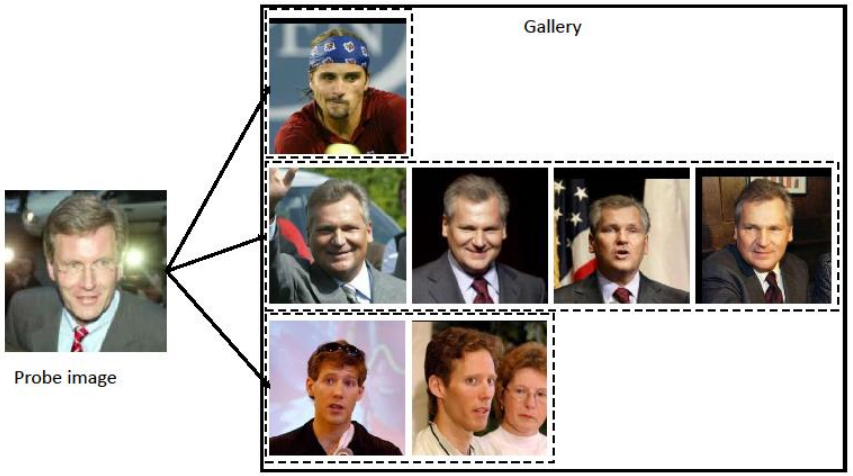


Figure: Supervised learning (face identification)

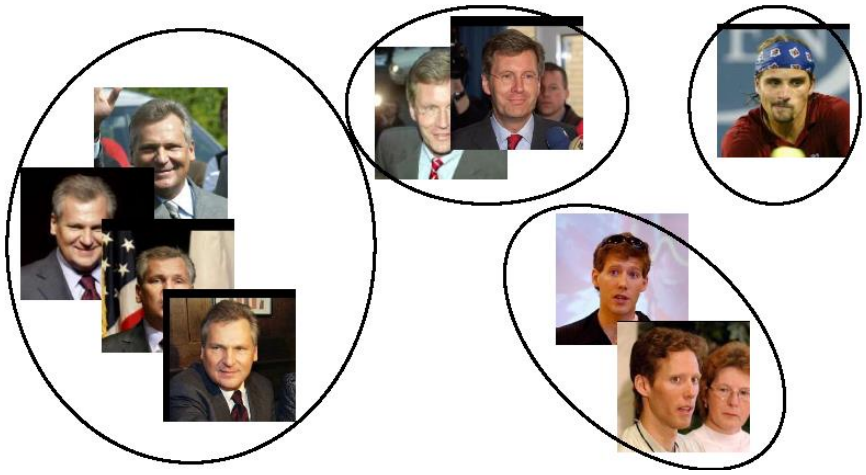


Figure: Unsupervised learning (face clustering)

Organization - Topics

- Basic image processing of grayscale and distance images
- Image Processing using Seam Carving
- Writer Identification/Retrieval
- Face Recognition
- Image Recognition for Humanities

Organization – Lectures and Exercises

- Lecture and exercise will alternate from week to week
- After each Lecture there is an optional **question section** for the current exercises

Preliminary schedule	
12.04.	TODAY – Lecture 1
19.04.	Exercise 1
26.04.	Lecture 2
03.05.	Exercise 2 **(Deadline ex 1)**
10.05.	Lecture 3
17.05.	Exercise 3 **(Deadline ex 2)**
31.05.	Lecture 4
07.06.	Exercise 4 **(Deadline ex 3)**
14.06.	Exercise 4
21.06.	Lecture 5
28.06.	Exercise 5 **(Deadline ex 4)**
05.07.	Exercise 5
12.07.	**(Deadline ex 5)**

Organization

- Are there any questions?

Motivation: Why are you here?

Motivation: Collision Avoidance

- Use cameras to observe a human/machine environment
- Avoid collisions by fusing camera and robot position information

Motivation: Autonomous Driving

Motivation: Hand Tracking

Computer Vision Primer

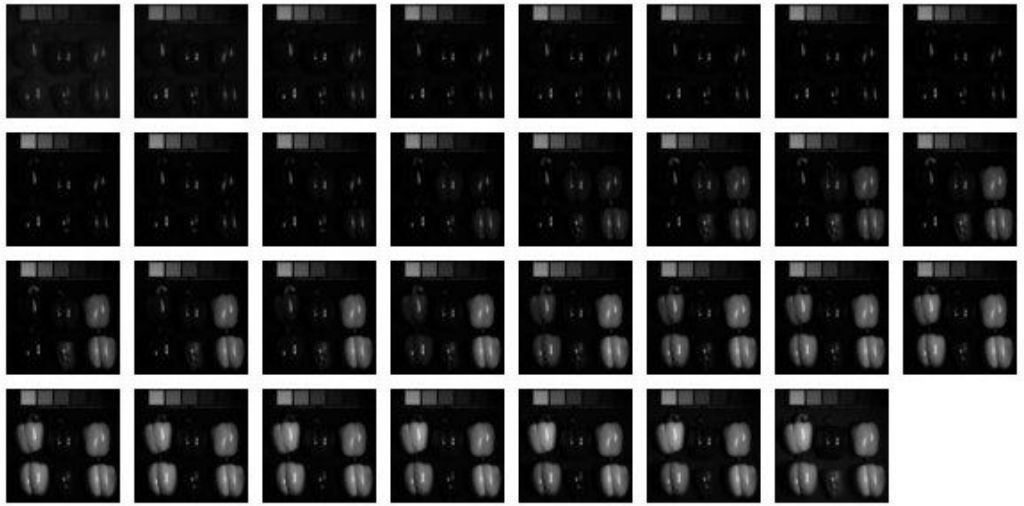
- Computer Vision is about extracting information from images
  - **Grayscale**
  - RGB
  - Multispectral
  - 2.5D and 3D

Computer Vision Primer

- Computer Vision is about extracting information from images
  - Grayscale
  - **RGB**
  - Multispectral
  - 2.5D and 3D

Computer Vision Primer

- Computer Vision is about extracting information from images
  - Grayscale
  - RGB
  - **Multispectral**
  - 2.5D and 3D



Computer Vision Primer

- Computer Vision is about extracting information from images
  - Grayscale
  - RGB
  - Multispectral
  - **2.5D and 3D**

Dataset by Johannes Schauer, Andreas Nüchter from the University of Würzburg, Germany

Computer Vision Primer

- Computer Vision is about extracting information from images
- What information?
  - Content: object detection, face identification
  - Text: sign detection, text extraction
  - Movement: tracking, motion capturing
  - Geometry: multi-camera setups, shape extraction, SLAM
  - Medical imaging: segmentation, tumor detection
  - Quality control: find production errors
  - ...

Image Processing on Grayscale and Distance Images

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

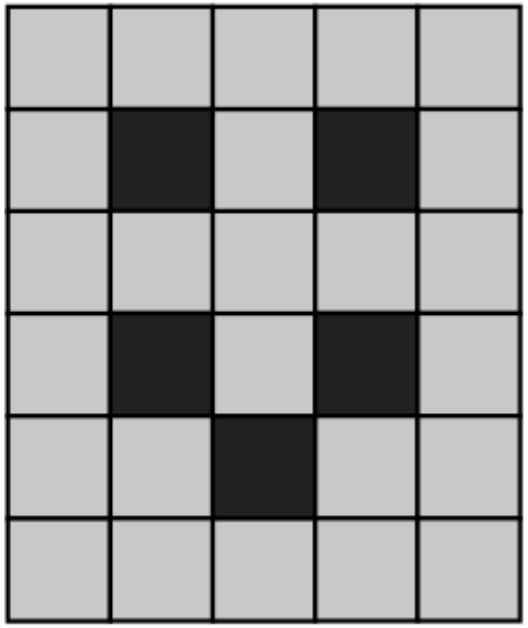
TECHNISCHE FAKULTÄT

Image Formation

Images

- Images can be represented as 2D-arrays of a certain type, e.g. byte for grayscale images.

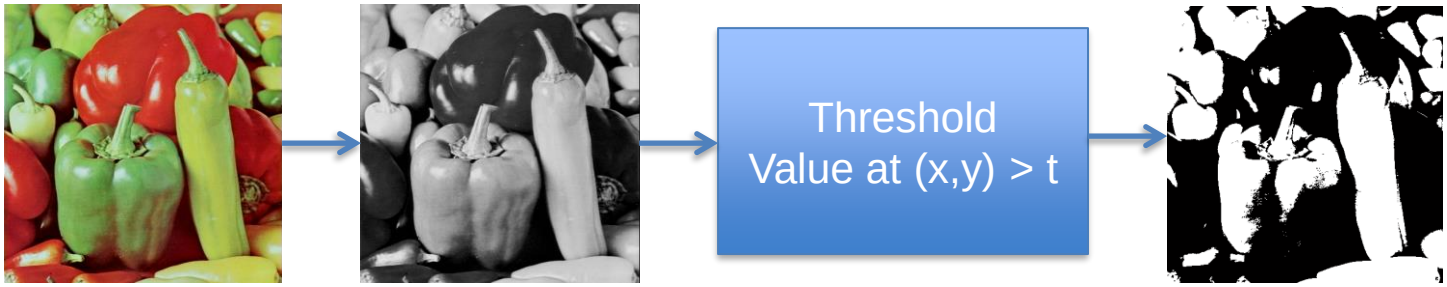
200	200	200	200	200
200	35	200	35	200
200	200	200	200	200
200	35	200	35	200
200	200	35	200	200
200	200	200	200	200



Gray scale image

Thresholding

- Simple operation that is used to separate regions of different „colors“
- Returns a mask of the same size as the image

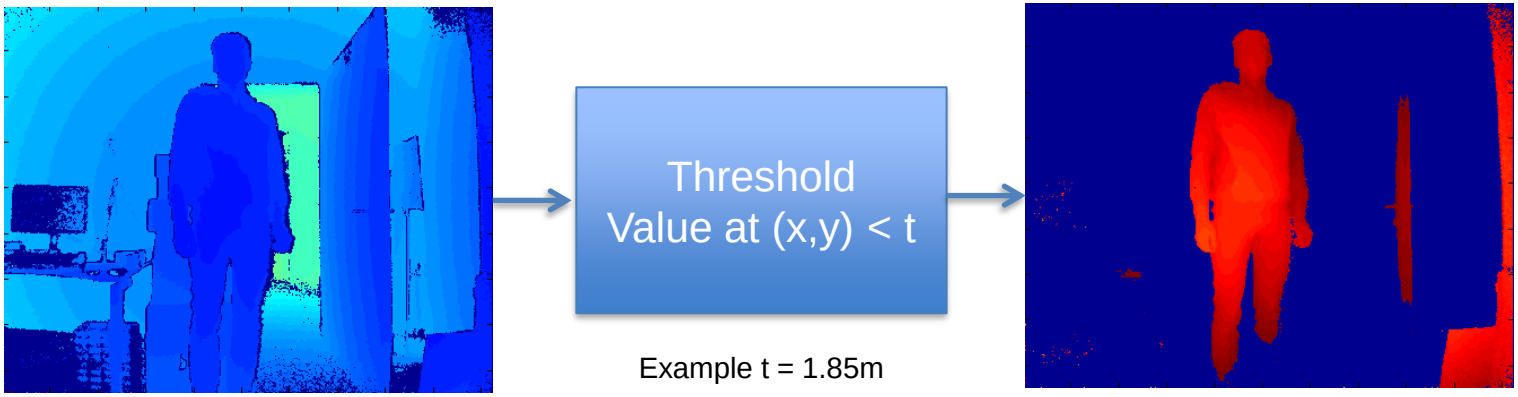


Input | Green channel | Example t = 150 | Resulting mask

Green peppers

Thresholding on Distance Images

- Pixels contain distance information. „Color“ encodes the distance.
- Example: On distance images thresholding can be used to separate fore- from background

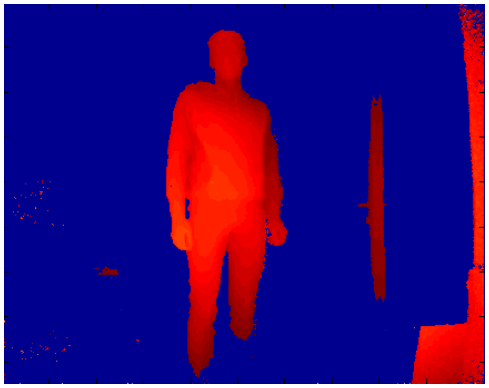


Distance image with person

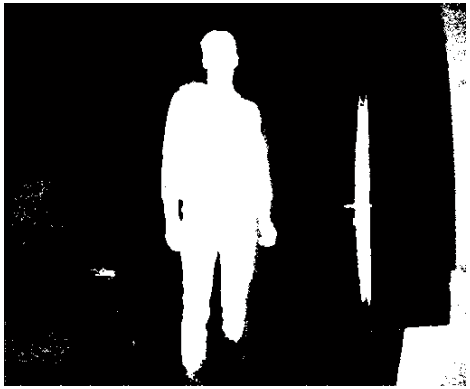
Scene foreground

Connected Components

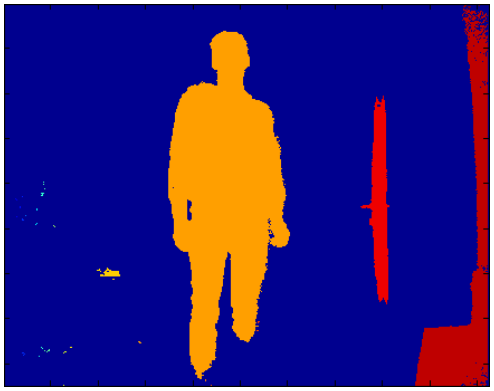
- Connected components are regions in an image that are connected
- A pixel is connected to another pixel if they are in the same neighborhood
- The neighborhood can be defined arbitrarily. Common neighborhoods are:
  - 4-connected neighborhood
  - 8-connected neighborhood



Input image



Mask image

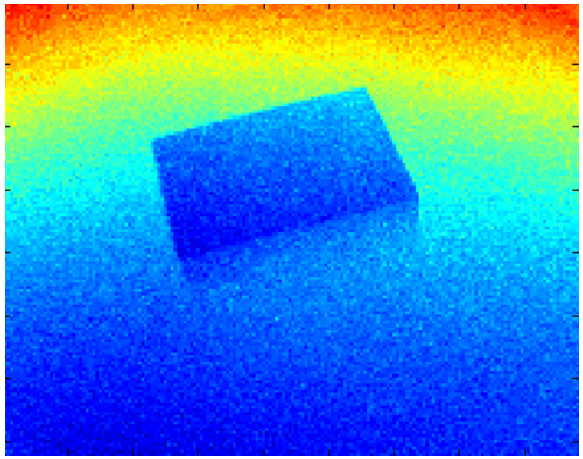


Connected Components

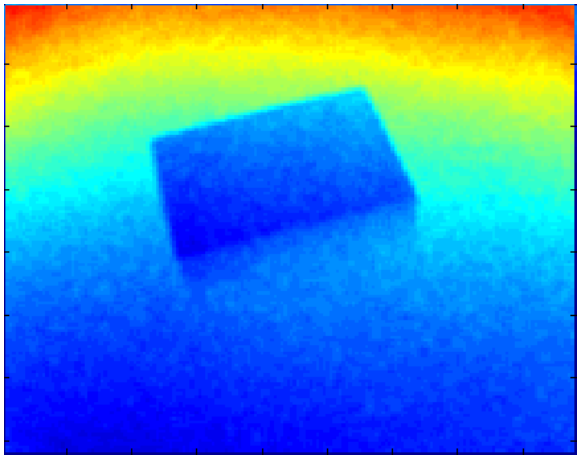
Smoothing: Mean Filter

- Smoothes the image by averaging pixel values within a certain window
- Commonly represented as a convolution
- A simple example is a 3x3 mean filter

$$I' = I * H = I * \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$



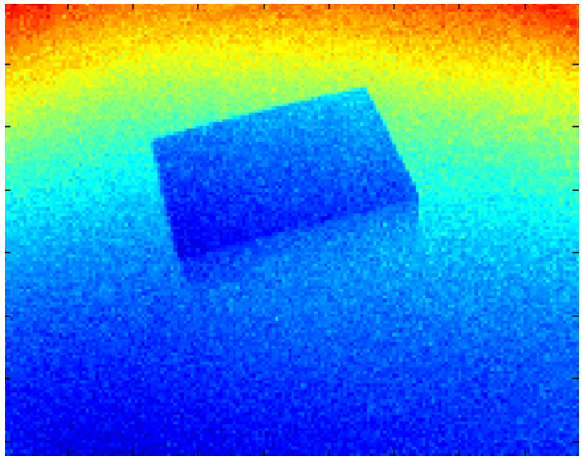
Distance image with a box



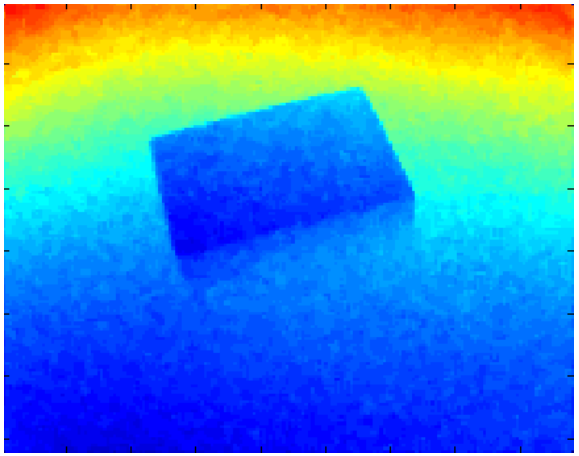
Distance image after mean filtering

Smoothing: Median Filter

- Sorts all pixel values within a certain window and uses the median of the sorted list as new pixel value
  - No convolution – nonlinear filter
  - Popular to remove salt-and-pepper noise



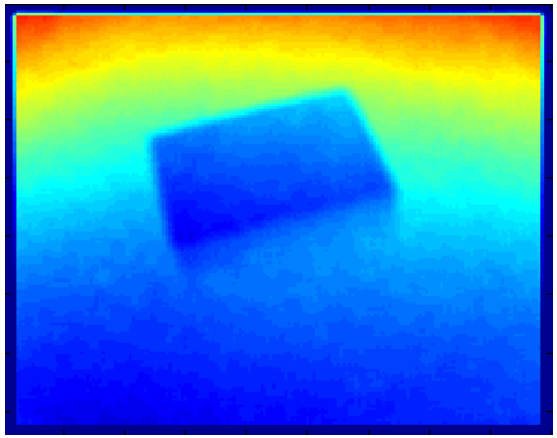
Distance image with a box

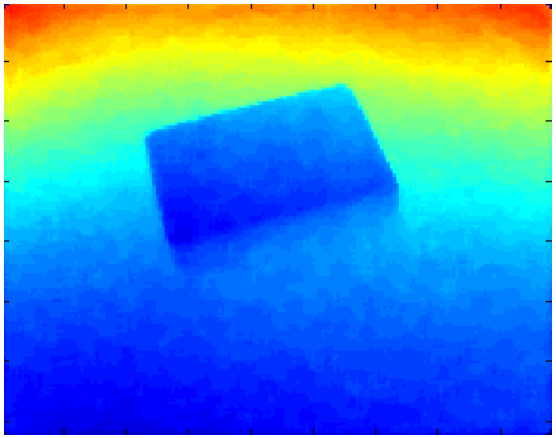


Distance image after median filtering

Smoothing: Mean vs. Median Filtering

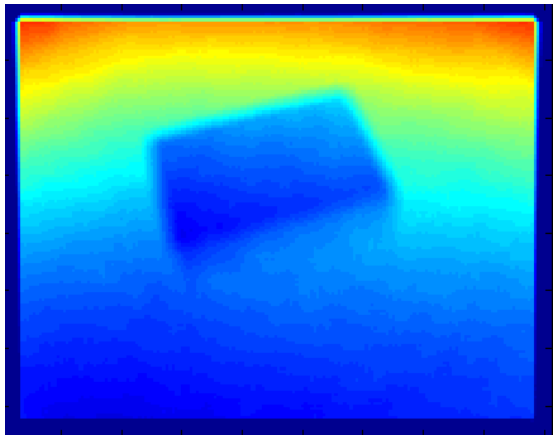
Mean 5x5

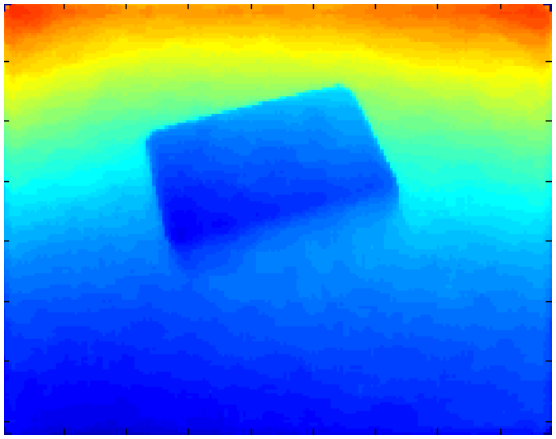




Median 5x5

Mean 7x7





Median 7x7

Smoothing

- There exist more smoothing filters and algorithms out of which some are based on convolutions while others are not.
  - Gaussian smoothing
  - Bilateral filtering
  - Guided filtering
  - ...

Morphological Operators

- Morphological operators typically work on binary images (e.g. masks)
- Analyzes pixels within a certain window/structuring element B
- Basic operators are: erosion, dilation, opening and closing

- Erosion $\ominus$: a pixel is 1 if all pixels in B are 1 in the input image
- Dilation $\oplus$: a pixel is 1 if at least one of the pixels in B is 1.

Green color mask | Input | Erosion (3x3) | Dilation (3x3)

Morphological Operators

- Opening is obtained by eroding the input first and applying dilation on the result:

$$A' = (\mathrm{A} \ominus B) \oplus B$$

- Closing is obtained by applying dilation first and erosion on the result:

$$\mathrm{A}' = (\mathrm{A} \oplus B) \ominus B$$

Green color mask | Input | Opening (3x3) | Closing (3x3)

Morphological Operators

(a)

(b)

(c)

(d)

(e)

http://szeliski.org/Book/

- (a) Input image
- (b) Dilation (5x5)
- (c) Erosion (5x5)
- (d) Opening (5x5)
- (e) Closing (5x5)

Further Reading

There exist different filters for a multitude of tasks

- Edge detection (Sobel, Scharr, Canny)
- Second order derivatives (Laplacian of Gaussian)
- Corner detectors [Harris88]
- Circle detectors [Loy03]
- ...

Edges: Sobel

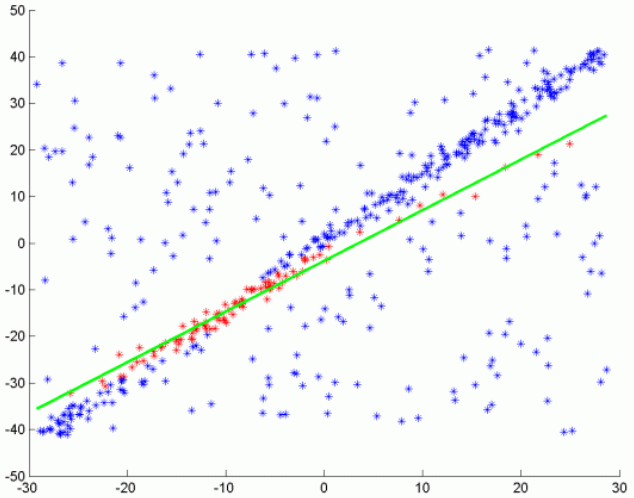
Edges: Canny

Corners: Harris

Optimization with RANSAC

- RANdom Sampling And Consensus is a simple optimization strategy for noisy data
- Assumes that a data set contains inliers and outliers (noisy samples)
- Can be used for all kinds of models: lines, planes, rotations, projection matrices, ...

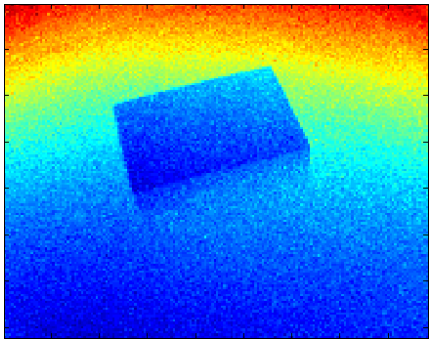
```
max_inliers = 0
best_model = null
while (i < N) {
        samples = minimal_sample(k, all_points)
        model = estimate_model(samples)
        n_inliers = computeInliers(all_points, model)
        if (n_inliers > max_inliers)  {
                best_model = model
                max_inliers = n_inliers
        }
}
```

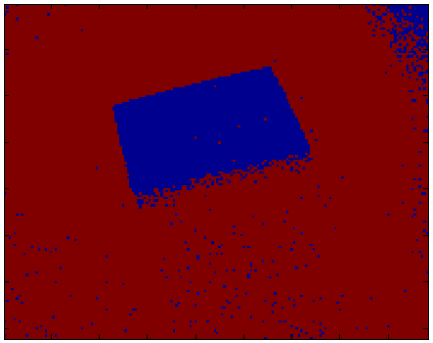


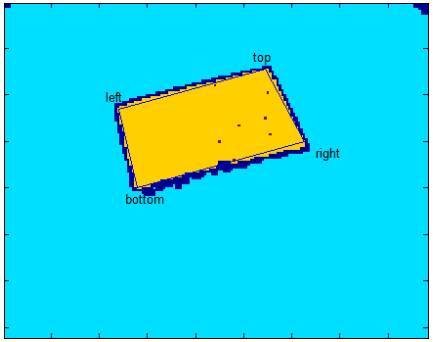
https://de.wikipedia.org/wiki/RANSAC-Algorithmus

Exercise 1: Box Detection

- What you get
  - A distance image of a simple box on the floor
  - And its corresponding point cloud
- Your goal is to
  - Find the floor plane
  - Find the box
  - Estimate the size of the box
- ...with the tools you have learned today

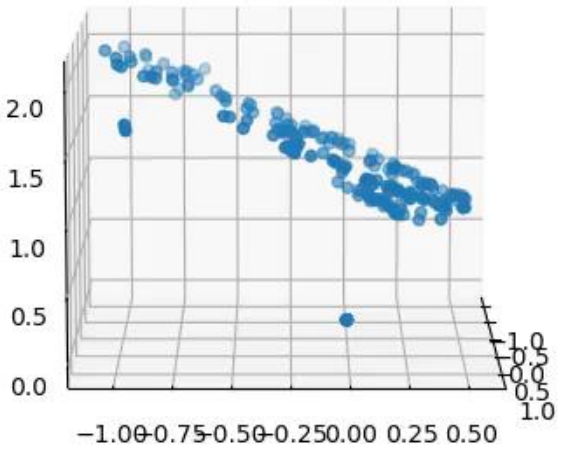




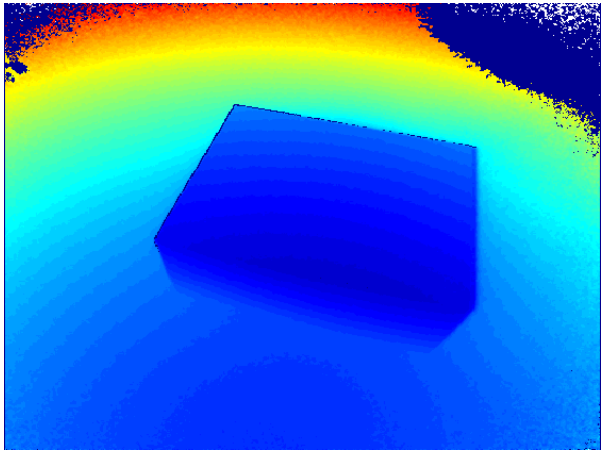


Exercise 1: Box Detection

Point cloud



Distance image



Questions?

References

[Harris88] C. Harris and M. Stephens (1988). "*A combined corner and edge detector*“, 1988

[Loy03] G. Loy, „*Fast radial symmetry for detecting points of interest*,” 2003