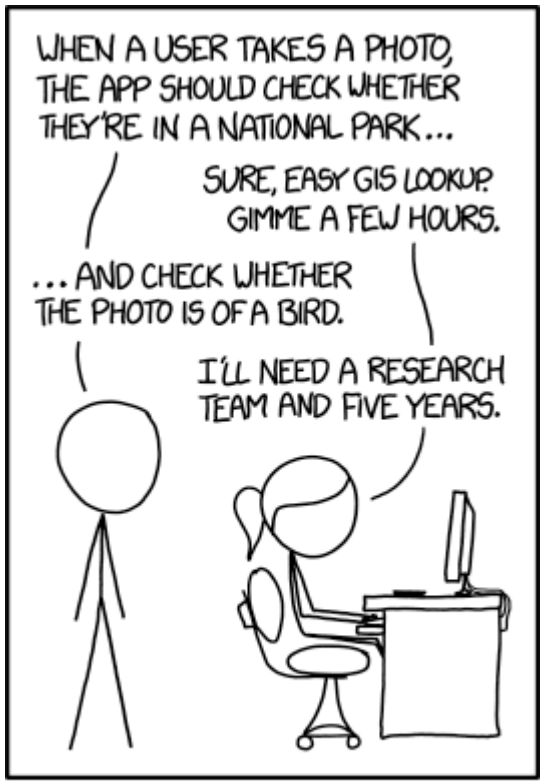
Computer Vision

Introduction

Technische Hochschule Rosenheim
Winter 2024/25
Prof. Dr. Jochen Schmidt

Object Recognition ...

... used to be hard:



... and sometimes still is:

Chihuahua or Muffin?

Object Recognition ...

Cat or ice cream?

© Twitter / Karen Zack / @teenybiscuit,
https://twitter.com/teenybiscuit/status/709538333219618816

Sloth or chocolate roll?

© Twitter / Minh Le / @GoosemanCS,
https://twitter.com/GoosemanCS/status/709589513832570880

Contents

Technische Hochschule Rosenheim

Main Course

1. Introduction
2. Image acquisition (camera sensors)
3. Camera calibration
4. Color
5. Image pre-processing: Threshold and filter operations (histograms and thresholds, linear and non-linear filters)
6. Object classification: Convolutional Neural Networks (CNN)
7. Object localization and detection (R-CNN, SSD, YOLO)
8. Image segmentation - classic methods (contour detection)
9. Image segmentation - with deep learning (Mask R-CNN, U-Net/Autoencoder)
10. Stereo

AAI Add-On

- Stereo: epipolar & fundamental matrix, 8-point algorithm
- Outlier Elimination: RANSAC, LMedS
- Structure-from-Motion
- SLAM
- Feature tracking

Literature

Technische Hochschule Rosenheim

Especially recommended

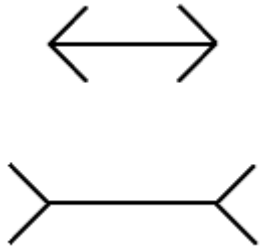
- Beyerer, J., Puente Leon, F., Frese, Ch.: ***Machine Vision***, Springer, 2016.
- Beyerer, J., Puente Leon, F., Frese, Ch.: ***Automatische Sichtprüfung***, Springer Vieweg, 2. Auflage, 2016.
- Szeliski, R.: ***Computer Vision: Algorithms and Applications***, Springer, 2nd edition, 2022.
- Goodfellow, I., Bengio, Y., Courville, A.: ***Deep Learning***, MIT Press, 2017. http://www.deeplearningbook.org/

Additional literature

- A. Nischwitz, M.W. Fischer, G. Socher: ***Computer Graphics and Image Processing***, Volume 2 Image Processing, Springer Vieweg, 4th edition, 2020.
- Gonzalez, R.C., Woods, R.E.: ***Digital Image Processing***, Pearson, 4th edition, 2017.
- Frochte, J.: ***Maschinelles Lernen - Grundlagen und Algorithmen in Python***, Hanser, 3rd edition, 2020.
- Aggarwal, Ch.C.: ***Neural Networks and Deep Learning: A Textbook***, Springer, 2018.
- Hartley, R., Zisserman, A.: ***Multiple View Geometry in Computer Vision***, Cambridge University Press, 2nd edition, 2004.

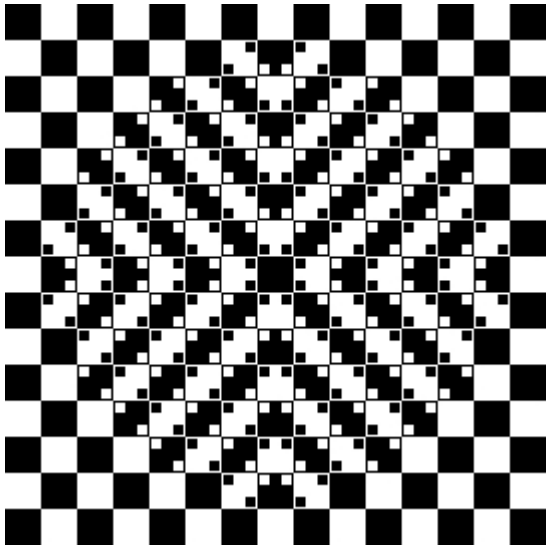
Human Vision: Optical Illusions

Lines



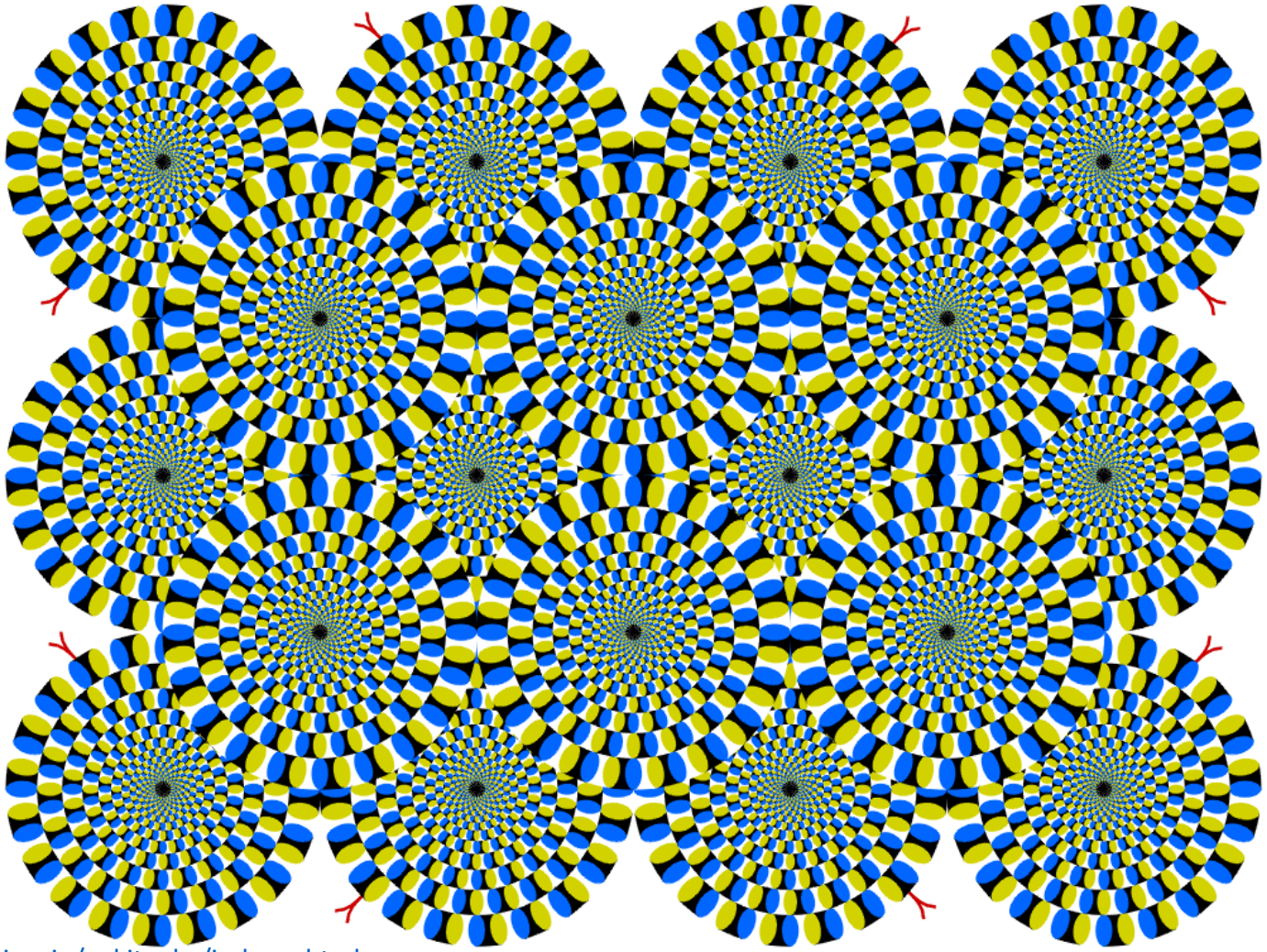
Are the two lines the same length?

http://www.ritsumei.ac.jp/~akitaoka/index-e.html



All fields are exact squares
All lines are exactly horizontal or vertical

Illusion of Movement

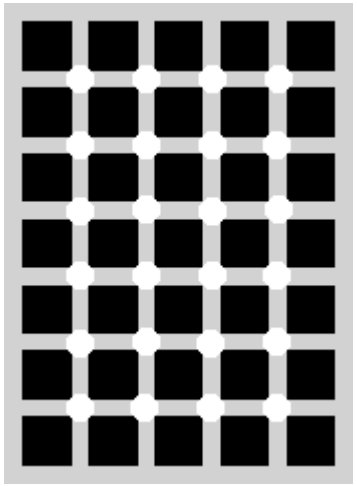


http://www.ritsumei.ac.jp/~akitaoka/index-e.html

Gray Values

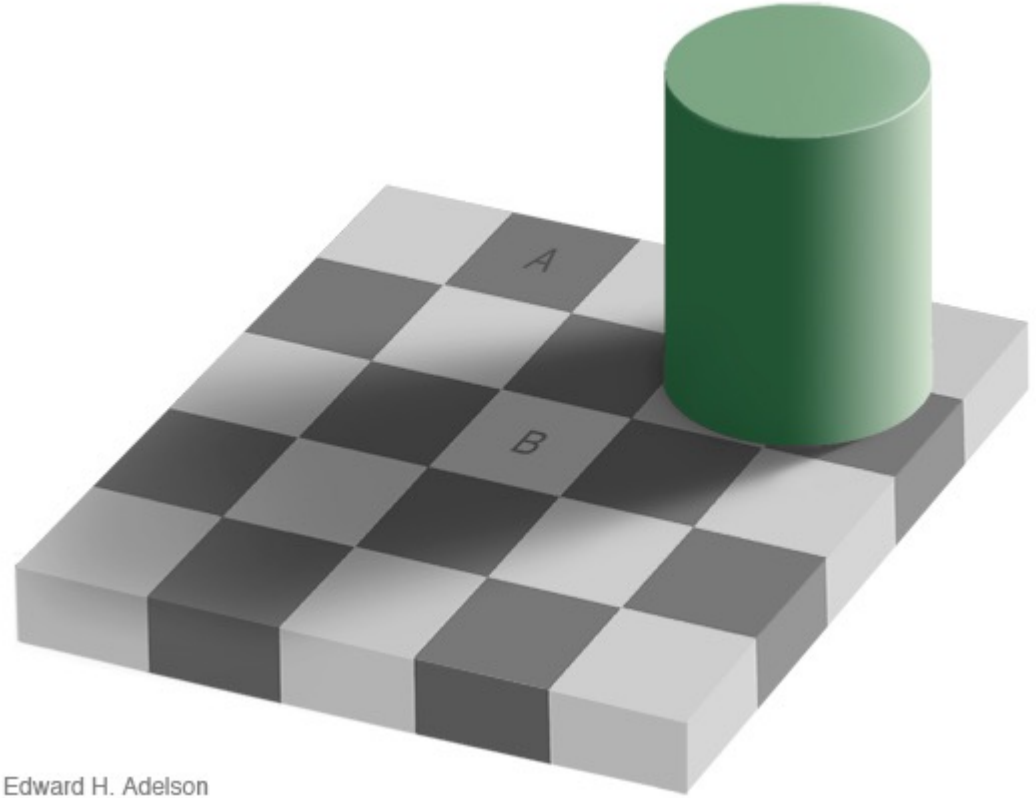
Technische Hochschule Rosenheim

http://www.cs.dartmouth.edu/~farid/Hany_Farid/Illusions/Entries/2011/6/6_Hermann_Grid.html



dark circles with
eye movement

http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html



Edward H. Adelson

Field A and B have **exactly** the same brightness
(the same gray value: 120)

Terminology

Computer Vision / Image Analysis

Technische Hochschule Rosenheim

- Image Processing
  - Images are transformed to images, often as a pre-processing step
- Image Analysis
  - Analytical processes that typically lead to a considerable reduction in data and a more compact description of the image
- Scene analysis and image understanding
  - Connection between recognized objects and the meaning for the user. Example: automatic navigation of autonomous vehicles. Artificial intelligence (AI) methods are often used here.
- Computer Vision
  - only vaguely defined
  - includes image and scene analysis as well as methods for 3D reconstruction and depth calculation (which can of course also play a role in scene analysis)

Pattern recognition

Technische Hochschule Rosenheim

- Pattern Recognition
  - Includes more than just image processing, e.g. also speech processing and other signals. In general: the processing of *patterns*.
  - Classification of simple patterns
  - Analysis of complex patterns

- Pattern
  - an element of the problem area under consideration (e.g.: an image of a figure or an object)
  - A pattern is a function

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{pmatrix} f_1(x_1, x_2, \ldots, x_n) \\ f_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ f_m(x_1, x_2, \ldots, x_n) \end{pmatrix}$$

Examples: Patterns

- Speech signal: $f(t)$
- Gray image: $f(x, y)$
- RGB color image:

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{pmatrix} f_r(x, y) \\ f_g(x, y) \\ f_b(x, y) \end{pmatrix}$$

- Image sequence (video):

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{pmatrix} f_r(x, y, t) \\ f_g(x, y, t) \\ f_b(x, y, t) \end{pmatrix}$$

Patterns

Technische Hochschule Rosenheim

- Simple pattern
  - Pattern can be classified as a whole
  - Class name is sufficient as a description
  - Examples:
    Character recognition,
    single spoken word

- Complex pattern
  - Classification as a whole pattern not possible
  - Class name not sufficient
  - Example:
    Description of a street scene,
    Interpretation of a CT scan

Classification System