BVA2 – Advanced Image Processing OVERVIEW

GERALD ZWETTLER

Topics

Overview Lecture, Labs

Labs and Grading

Science Landscape Computer Vision

Outlook: Exercises

Labs

Exercises to all topics/chapters are handed out to be completed within 2 weeks (~3 weeks for larger exercises). From a total of 200 points, exercises for 100 points can be selected as 100%. Some exercises can be achieved in groups of 2 students.

During the labs, exercises are partially started and discussed together and can be worked on under guidance of the lecturer.

After grading of the exercises, one student/group will be selected to present and discuss the solution. This way one can gain insight into topics too, where the exercises were left out.

As toolset ImageJ/IntelliJ is used for Java solutions from scratch and OpenCV via Python wrappers and Spider IDE for comprehensive exercises where a given code basis is inevitable.

Labs cont’d

expected deliverables:

- **1 zip file per exercise**, naming: ex_(1|2|3|4|5|6|...|n)_name1.zip), e.g. *ex_47_zwettler.zip*
- all source files (no projects or binaries), well documented
- utilized testing images (if space permits, otherwise some relevant ones)
- documentation as PDF:
  - Theoretical background and literature if appropriate
  - Statement of the problem
  - Description of solution and implementation (if applicable)
  - Results, extensive test-runs (if applicable)
  - Discussion

Grading

both, labs and exam must be positive. Overall grade with 50% labs and 50% exam. Exam: 15min oral exam at the end of the term.

>87.5%: 1, >75%: 2, >62.5%: 3, >50%: 4, else 5.

Computer Vision

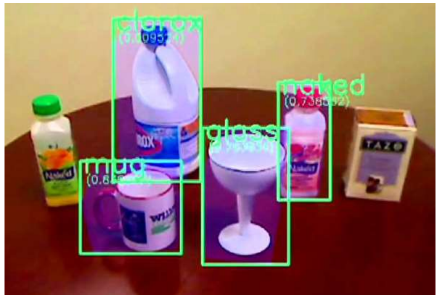
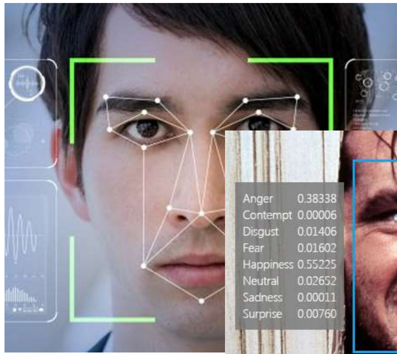
Definition

- *Computer Vision as extension of Image Processing refers to complex and advanced image analysis, detecting/segmenting desired structures as basis for full image and scene understanding.*
  - input data:
    - single monocular images
    - video streams
    - 3D images, additional sensor data, 3D Surface Models
    - more often huge amount of input data (tagged or untagged) in the sense of big data – images, i.e. *big visual data* as dominant information available from the internet
  - some of the involved disciplines are:
    - Image Processing (filtering, image reconstruction, color transformations, simple segmentation tasks,...)
    - Optics and Imaging Systems (camera calibration, distortion, 3D reconstruction,...)
    - Computer Graphics (geometric modelling, shape models, reflections, lightening models, test data generation, 3D reconstruction, result representation,...)
    - Pattern Recognition (Texture, SIFT, VSLAM,...)
    - Robotics (VSLAM, autonomous driving, production inspection)
    - Artificial Intelligence and Heuristics (feature based classification, registration, symbolic representation,...)
    - Mathematics, Numerics (projective geometry, probabilities, statistics, signal processing,...)

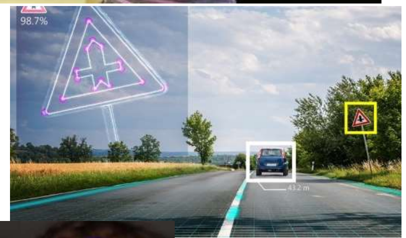
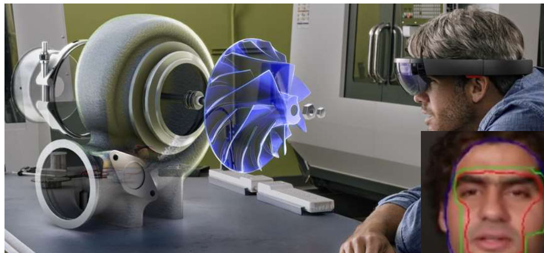
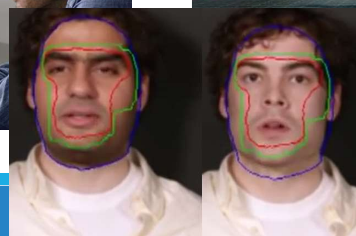
Computer Vision

Fields of Application

- 3D Reconstruction from planar images
- recognition
  - marker, objects
  - face detection / recognition, iris, fingerprint, gesture, handwriting, human pose and activity recognition
  - optical character recognition
- classification/segmentation
  - object contour segmentation – extract and characterize desired objects
  - semantic visual classification of entire images
  - complex reasoning from image data utilizing AI
- navigation
  - autonomous driving – land, underwater, space
- mixed reality
  - synthesizing real-world and computer-graphics objects
- synthetic images
  - morphing, computer generated imagery (movies)

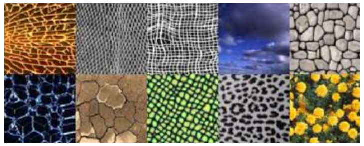


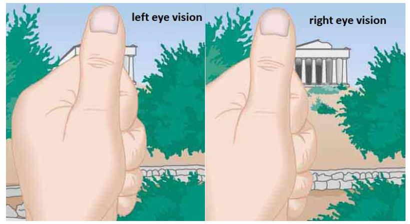





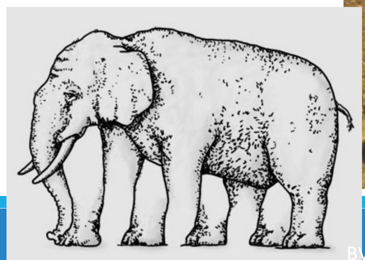
Computer Vision

How do Humans Recognize and Interpret Images?

- up-to-now human outperformed AI in image recognition tasks with complex semantic interpretation – meanwhile changing due to incorporation of Deep Learning cNN for Computer Vision Tasks. The bastions of human superiority are on the decline (*chess* ➔ *AlphaZero, GO* ➔ *AlphaGo,...*)
  - human visual recognition neural network input [Bramao et al. 2011; DiCarlo et al. 2012;Fleming 2014]:
    - (surface) color
    - 3D vision due to parallax of human eyes
    - lighting / shading for 3D shape and distance approximation
    - image texture
    - shape and surface properties
    - a priori knowledge of learned objects, size and their 3D structure
    - stereotype reasoning as short-circuit for human recognition system



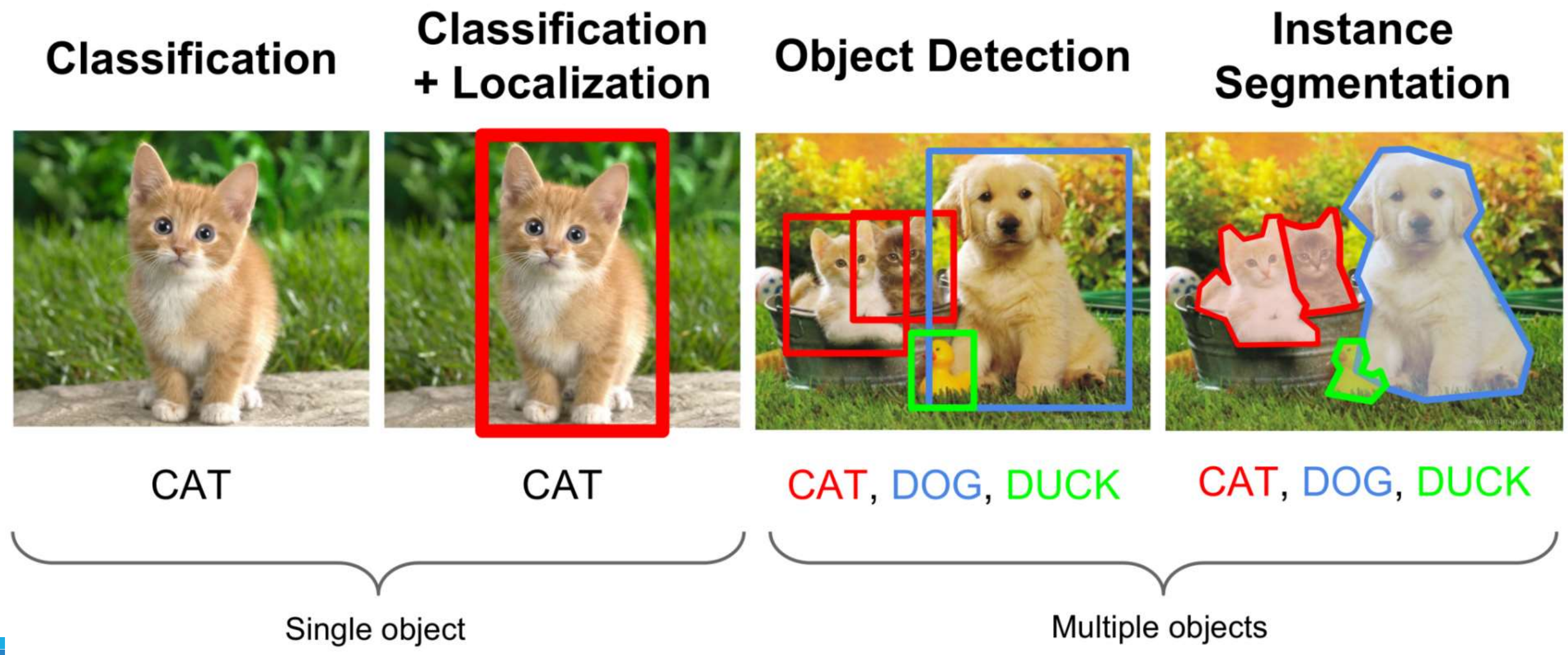




Computer Vision

Basic Terminology

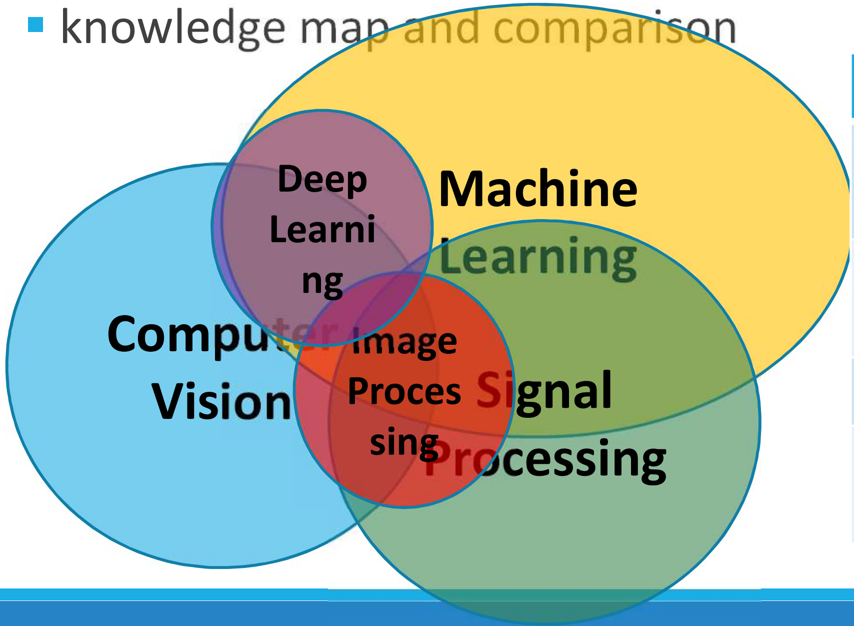
- semantic interpretation of images often a prerequisite
  - tagging of the entire image by one “word” ➔ *Classification*
  - additionally detecting the ROI around the potentially occluded key object ➔ *Object detection* with Classification and *Localization* for images with 1 or more objects involved
  - marking the precise visible or occluded pixel region referring to the classified object ➔ *Instance segmentation*



Computer Vison

Computer Vision and image processing algorithms knowledge map and comparison

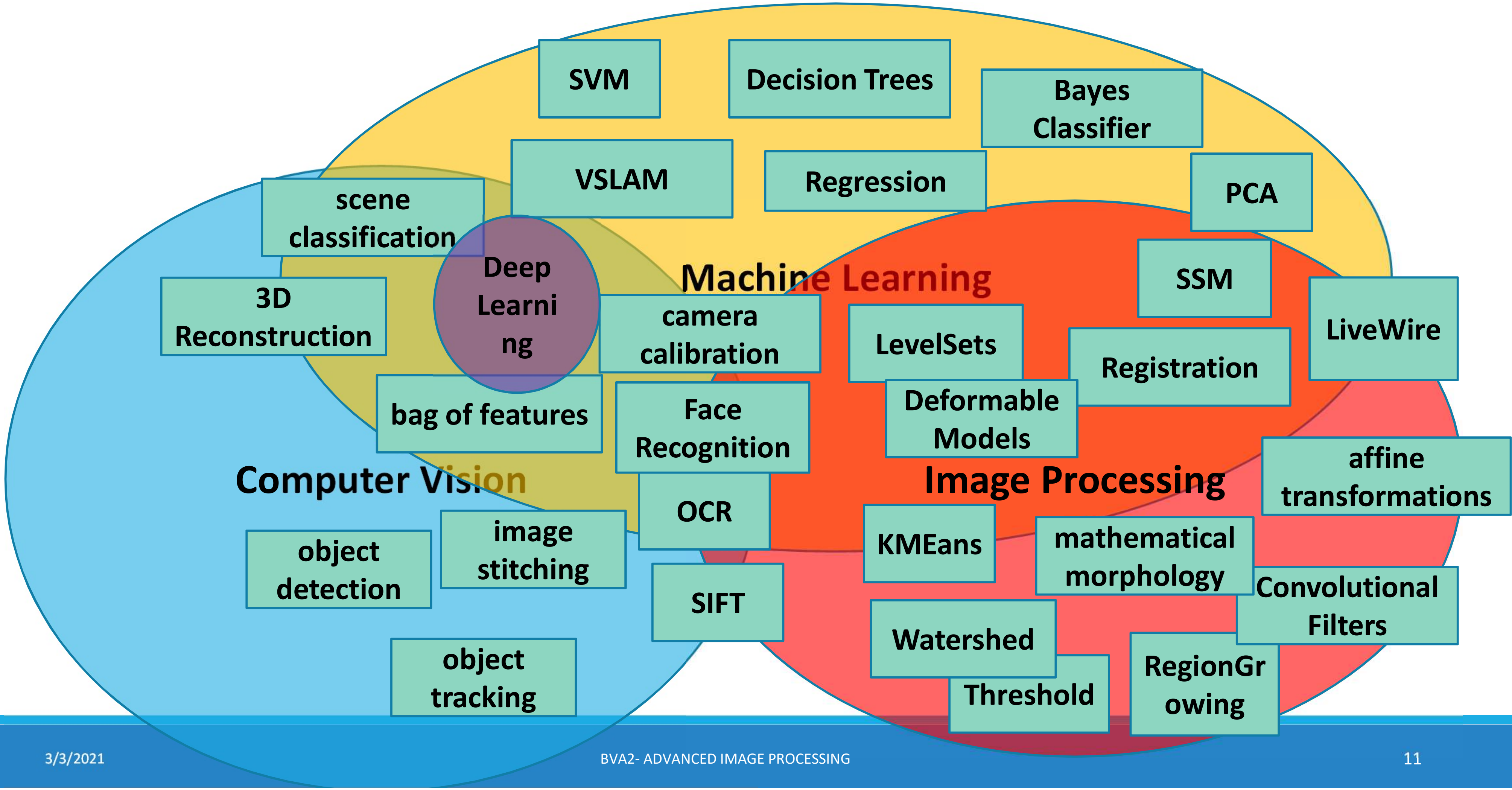
- *machine learning*, *deep learning*, *computer vision*, *signal processing* and *image processing* domains closely related
  - image processing thereby a specialization of common signal processing.
  - image processing and computer vision to be differentiated according to the level of semantic perception
  - computer vision and deep learning as specialization od CV necessitate image processing pre-processing steps
- knowledge map and comparison



domain	input data	output data
image processing	image	modified image
signal processing	signal (EKG, image, EEG,...)	derived information, signal
computer vision	image / images / video	semantic classification
machine learning	features of signal from any domain	classification, model parameters, numeric solution

Computer Vision

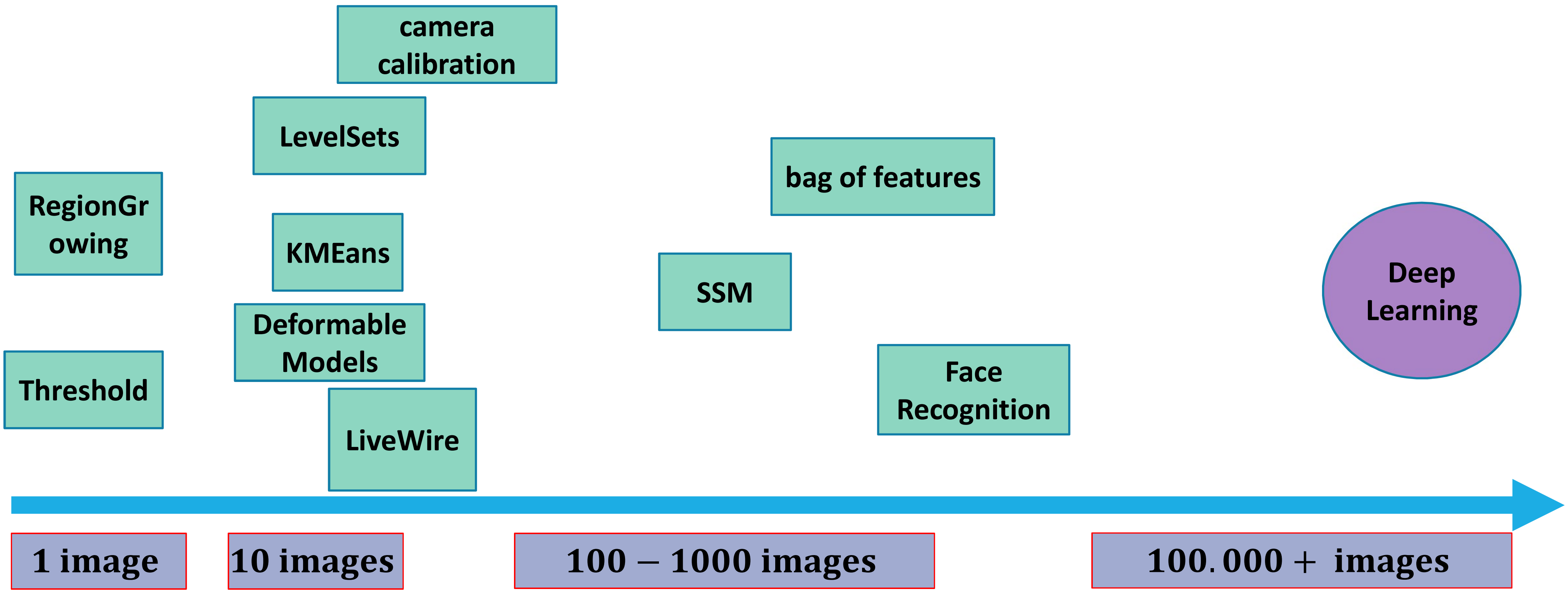
Classification of relevant algorithms and concepts



Computer Vision

Computer Vision and image processing algorithms – qualitative comparison

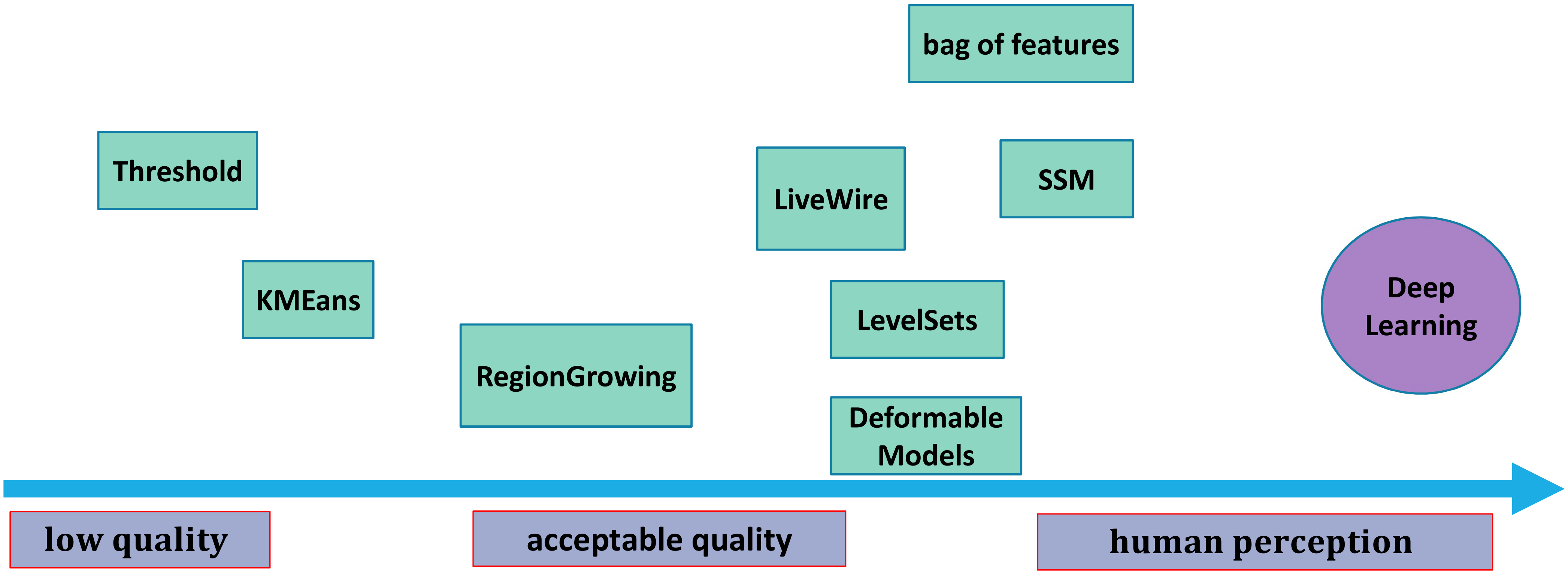
- required **training data** for **segmentation** approaches



Computer Vision

Computer Vision and image processing algorithms – qualitative comparison cont'd

- achievable **segmentation/classification accuracy**



References

- [DiCarlo et al. 2012] DiCarlo, J.J., Zoccolan, D., and Rust, N.C., 2012. “How does the brain solve visual object recognition?” . In: *NEURON 73(3)*.
- [Bramao et al. 2011] Bramao, I., Faisca, L., Petersson, K.M., and Reis, A. 2011. *The Contribution of Color to Object Recognition*. In: Tech, Rijeka, Croatia, available from: http://pubman.mpdl.mpg.de/pubman/item/escidoc:1468179/component/escidoc:1468180/Bramao-InTech-12.pdf
- [Fleming 2014] Fleming, R.W., 2014. “Visual perception of materials and their properties”. In *Vision Reserach 94*, pp. 62-75.

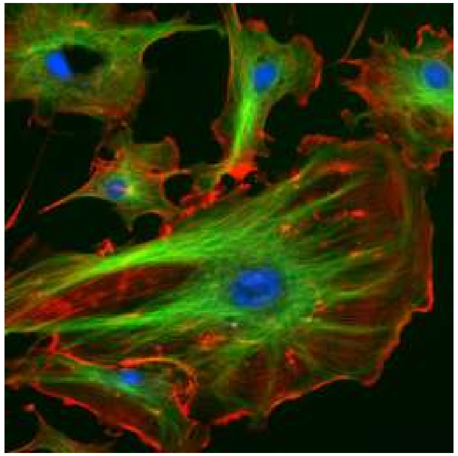
Outlook Exercises

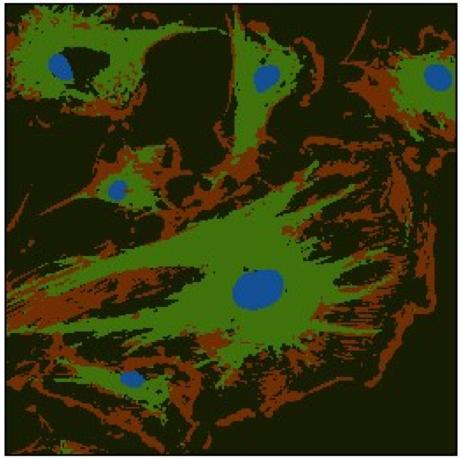
1. Linear Imaging Systems : camera calibration, video segmentation (heat map)

https://docs.opencv.org/master/dc/dbb/tutorial_py_calibration.html

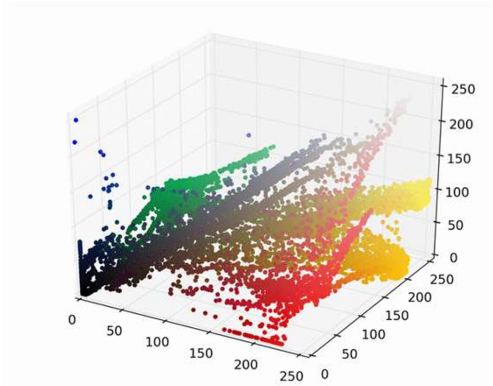
Outlook Exercises

2. Segmentation and Classification: KMeans Clustering, MeanShift





input image showing solid colors of similar chromaticity and luminosity



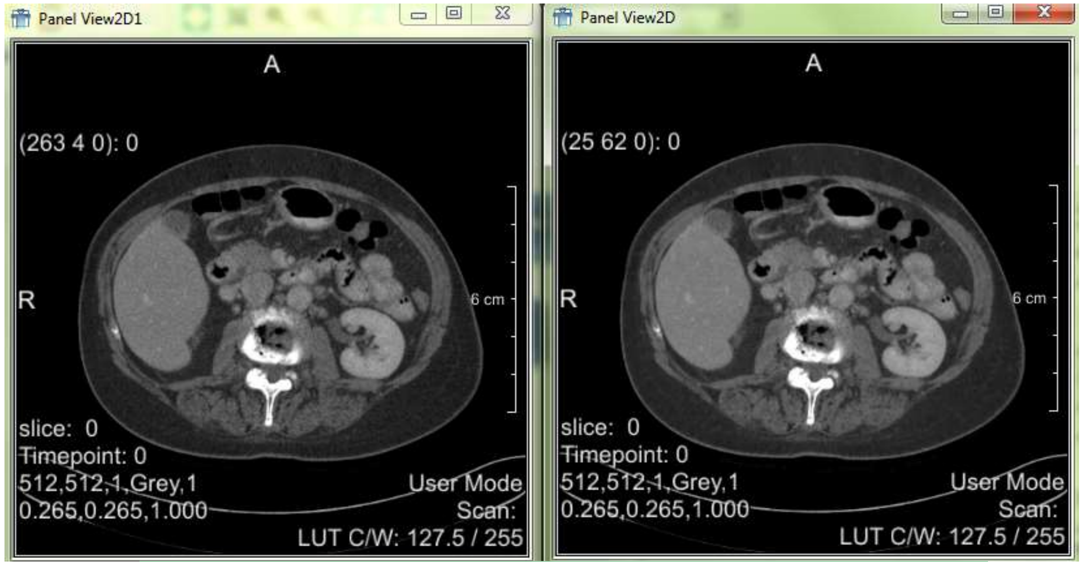
pixels as 3D particles in 3D color space

final classification with n=7 clusters with kernel radius k=25 leading to fragmentation into 7 color segments

Outlook Exercises

3. Image Restauration and Nonlinear Filtering: Anisotropic Diffusion, Richardson-Lucy Deconvolution

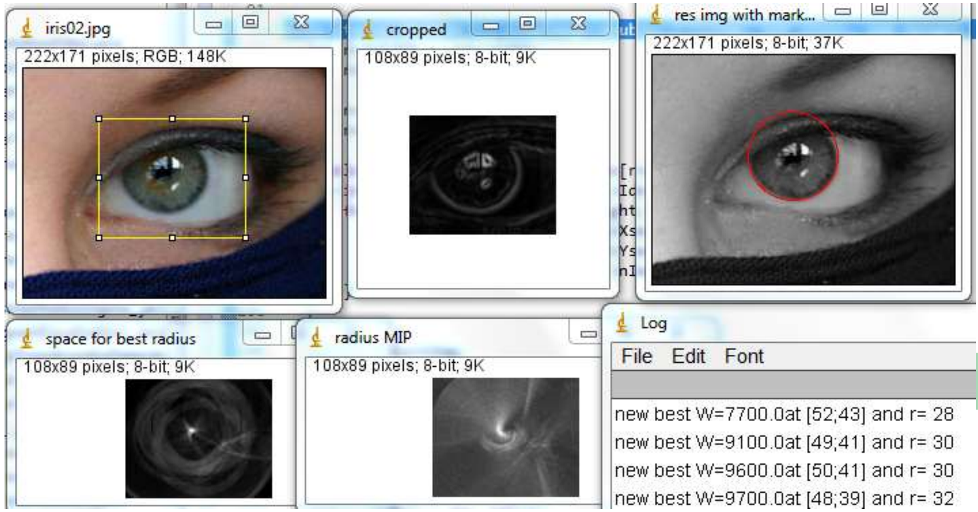
RLD: before and after



anisotropic diffusion – edge conserving low-pass filtering

Outlook Exercises

4. Localization: Hough Circles for Iris Detection, SIFT Feature Detection, Image Stitching



Hough Circles

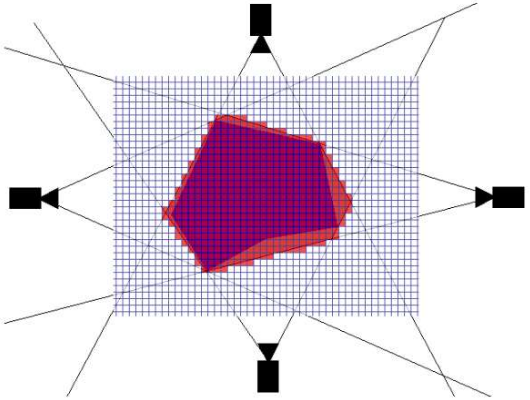
SIFT Feature Matching

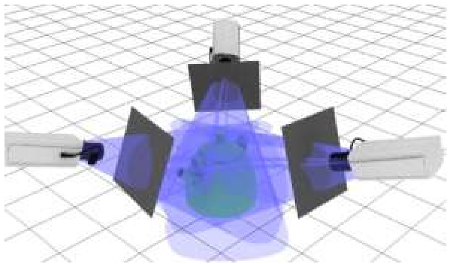
http://matthewalunbrown.com/autostitch/autostitch.html

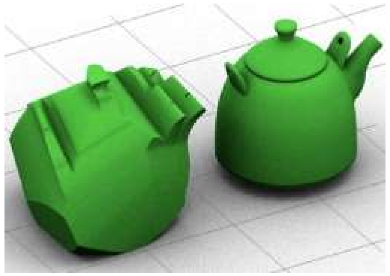
Image Stitching

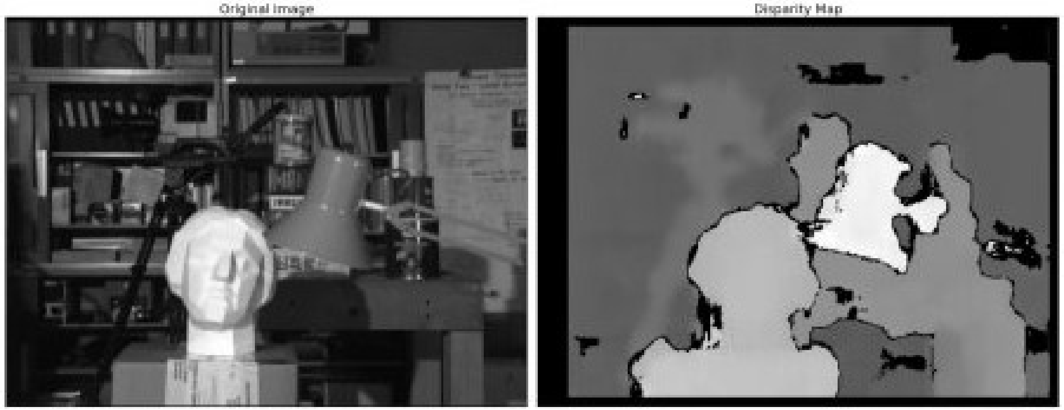
Outlook Exercises

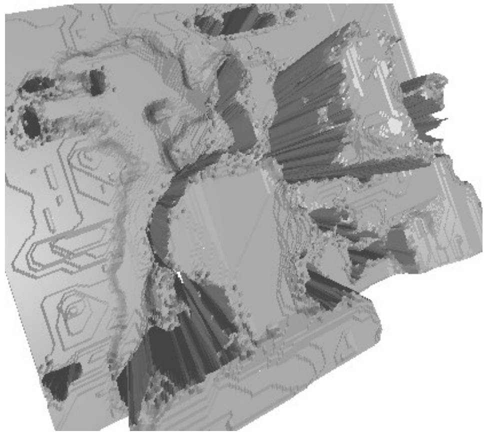
5. 3D Reconstruction: Stereo Matching and Visual Hull from Silhouette Reconstruction









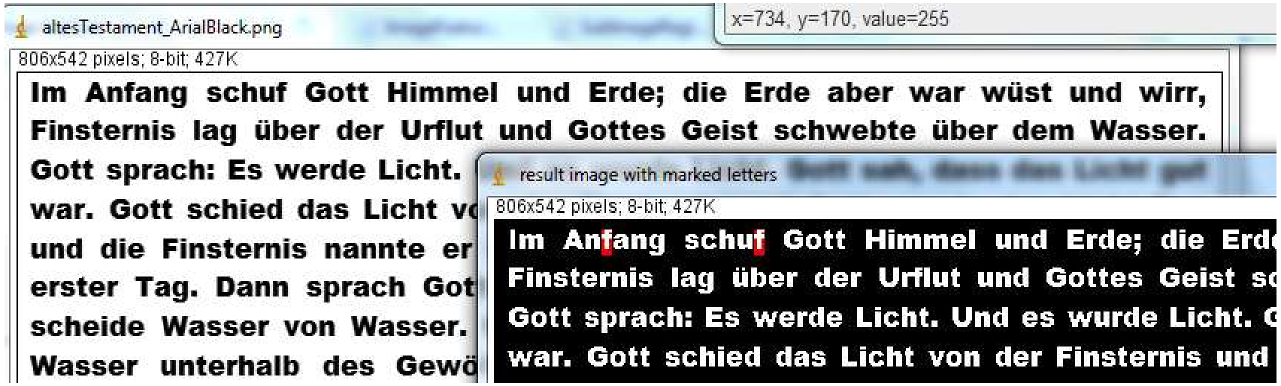


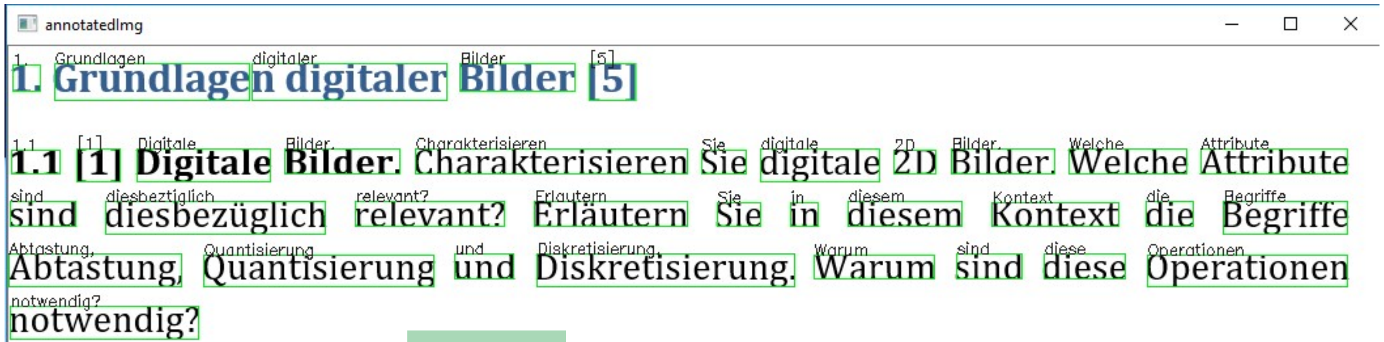
from https://www.cvg.ethz.ch/teaching/compvis/2011/lecture/vision08.pdf
https://ars.els-cdn.com/content/image/1-s2.0-S0141933112000750-gr1.jpg

https://docs.opencv.org/3.4/dd/d53/tutorial_py_depthmap.html,
https://imagetostl.com/

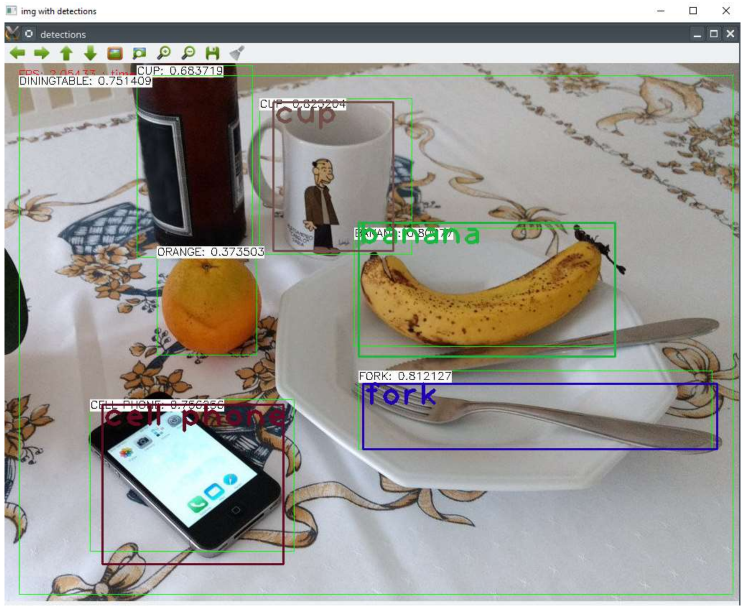
Outlook Exercises

6. Computer Vision and Machine Learning: OCR from scratch and using Tesseract, Object Detection & Classification





OCR



Yolo Object Detection