UNIVERSITÄT BONN

Juergen Gall

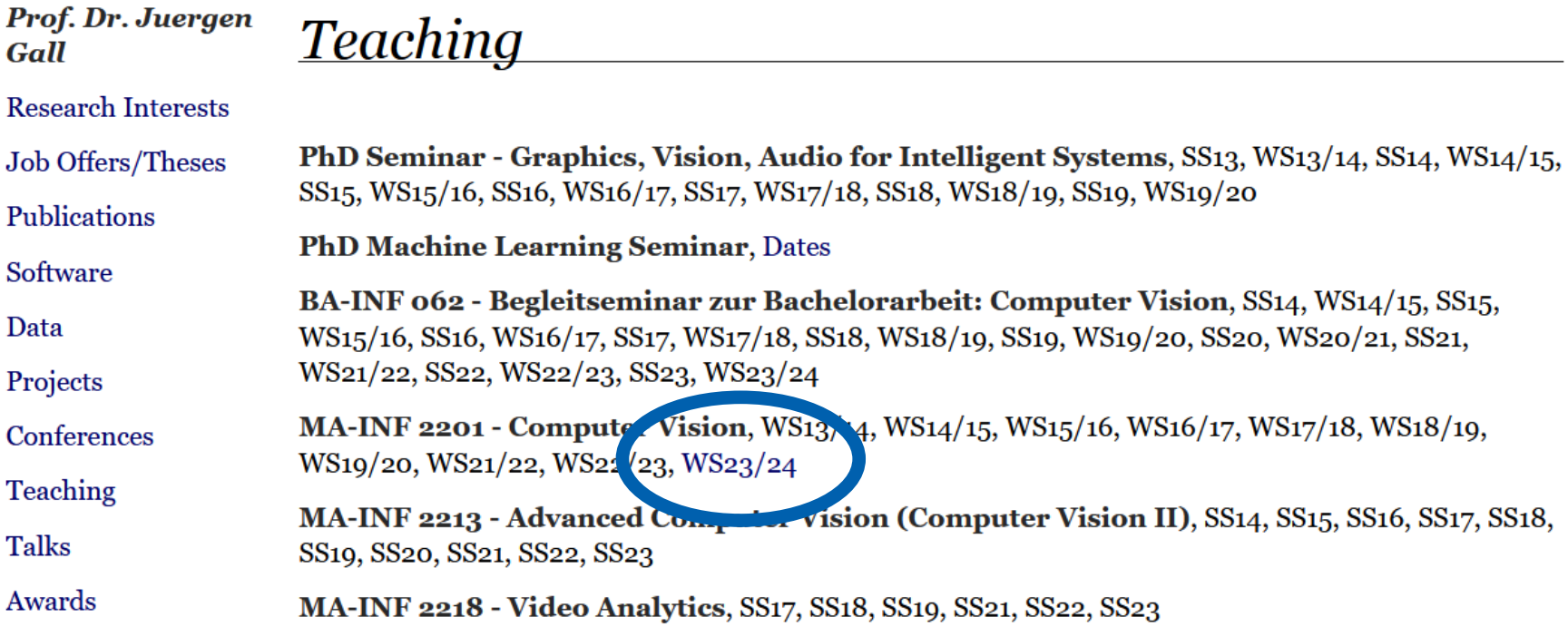
Introduction

MA-INF 2201 - Computer Vision

WS24/25

Organization

- Lecturer:
  - Prof. Dr. Juergen Gall
  - http://gall.cv-uni-bonn.de/
- Teaching:



Prof. Dr. Juergen Gall

Research Interests

Job Offers/Theses

Publications

Software

Data

Projects

Conferences

Teaching

Talks

Awards

Teaching

PhD Seminar - Graphics, Vision, Audio for Intelligent Systems, SS13, WS13/14, SS14, WS14/15, SS15, WS15/16, SS16, WS16/17, SS17, WS17/18, SS18, WS18/19, SS19, WS19/20

PhD Machine Learning Seminar, Dates

BA-INF 062 - Begleitseminar zur Bachelorarbeit: Computer Vision, SS14, WS14/15, SS15, WS15/16, SS16, WS16/17, SS17, WS17/18, SS18, WS18/19, SS19, WS19/20, SS20, WS20/21, SS21, WS21/22, SS22, WS22/23, SS23, WS23/24

MA-INF 2201 - Compute[r] Vision, WS13/[14], WS14/15, WS15/16, WS16/17, WS17/18, WS18/19, WS19/20, WS21/22, WS2[2]/23, WS23/24

MA-INF 2213 - Advanced Comp[uter] Vision (Computer Vision II), SS14, SS15, SS16, SS17, SS18, SS19, SS20, SS21, SS22, SS23

MA-INF 2218 - Video Analytics, SS17, SS18, SS19, SS21, SS22, SS23

Organization

UNIVERSITÄT BONN

- Slides: http://gall.cv-uni-bonn.de/teaching/Lectures/cv23.html

MA-INF 2201 - Computer Vision

4L + 2E, WS23/24

Lecturer

Juergen Gall

Content

Tentative: Linear filters, Edges, Derivatives, Hough Transform, Segmentation, Graph Cuts, Mean Shift

Background Subtraction, Temporal Filtering, Active Appearance Models, Shapes, Optical Flow, 2D Tra

Estimation, Articulated Pose Estimation, Deformable Meshes, RGBD Vision

Material

The slides and recordings are available at Slides/Recordings.

Prerequisites

Basic knowledge of linear algebra, analysis, probability theory, Python programming

Organization

- Slides via sciebo (https://hochschulcloud.nrw/en/):

- Password: MA-INF2201

Organization

UNIVERSITÄT BONN

- Structure: 4+2 SWS
- Lecture:
  - Tuesday, 10:15-11:45, HSZ / HS 3
  - Friday, 10:15-11:45, HSZ / HS 3
  - Lectures will be recorded
- Exercise:
  - Start: 16.10.
  - Wednesday, 12:15-13:45, 2.025, Informatikzentrum
  - A second slot for the exercise will be announced

Credit Points

- 9 CP for students from University of Bonn
- 8 CP for students from B-IT (Media Informatics), but
  - Contact: Exam Office (RWTH / H-BRS)
  - After exam you get certificate of achievement for the course from the lecturer stating the name, credits, area, and grade. That certificate of achievement then has to be presented to the exam office (RWTH / H-BRS) so that it gets registered into the system.

Organization

UNIVERSITÄT BONN

- Feedback welcome!
- There is time to repeat parts of previous lecture (only on request)
- Two recap lectures
- No explicit exam preparation (attend lecture and solve exercise sheets)
- Written exam
- Contact per email:
  - gall@iai.uni-bonn.de
  - Email with subject: [MA-INF 2201]

Exercise

UNIVERSITÄT BONN

- Small implementation exercises
- Exercises are not exam questions, but understanding → Able to implement it
- No programming skills → No job as computer scientist
- OpenCV and Python (http://opencv.org/)

Exercise

UNIVERSITÄT BONN

- You can use Github Copilot, Claude, …
  …but you need to understand your code.
- Remember: In a job interview you need to program without these tools
- If the tutors have some doubts they might ask you to explain your solution.
- Quality of code is relevant for grading

At least 50% of the points are required
to qualify for exam

Organization

UNIVERSITÄT BONN

- Registration for exercise today!

Exercises

Theory and programming. At least 50% of the exercise points are required to qualify for exam. The programming e
the exercises is on implementing algorithms presented in the lecture. If you are not comfortable with OpenCV, we r
exercise sheet, which is an introduction to OpenCV, will be released on the date of the first lecture.

Schedule

Wednesday, 12:15-13:45, 2.025, Informatikzentrum

Start: Wednesday, 18.10.

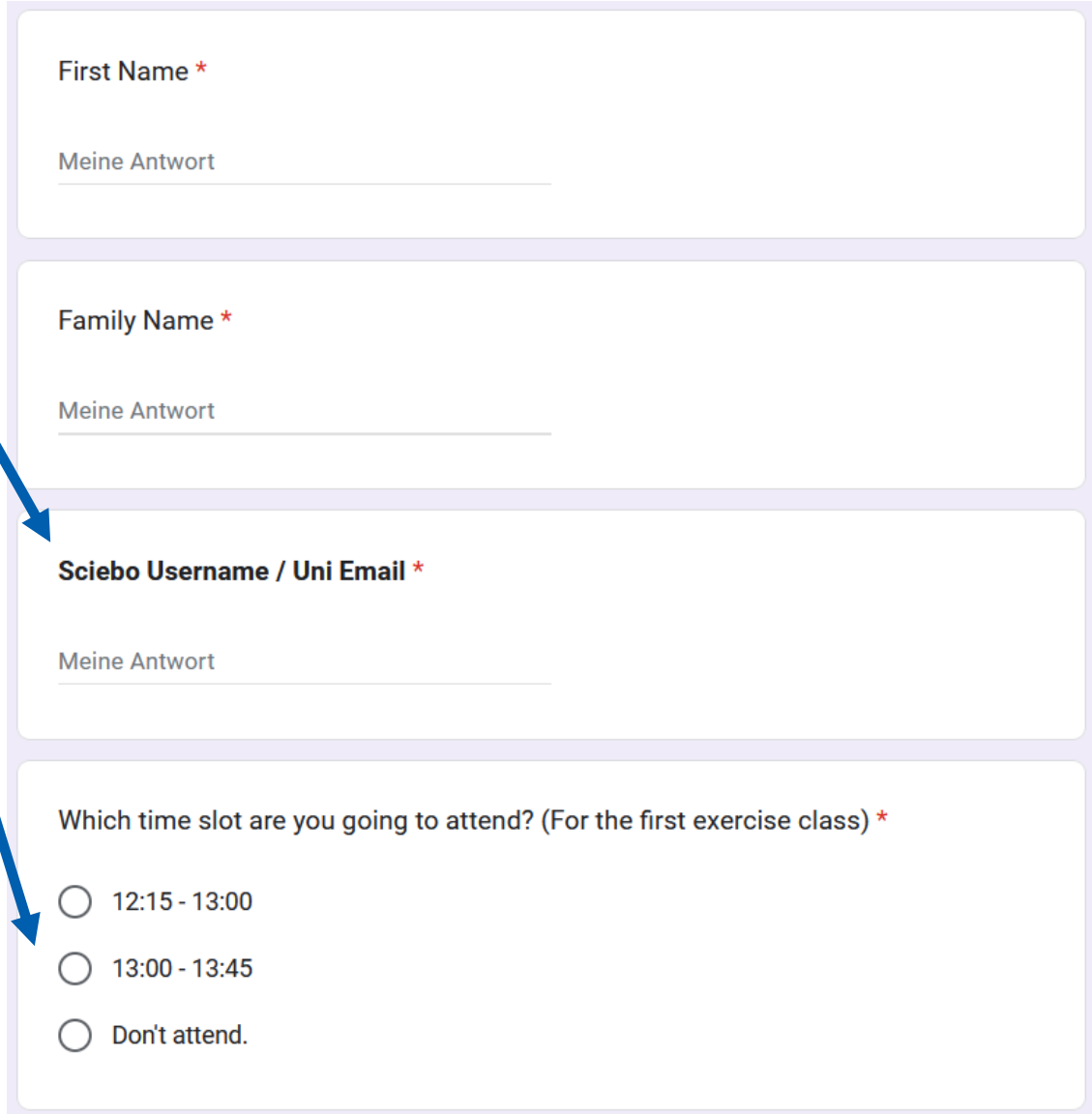
Register until 10.10. for the exercise groups: registration form.

If you did not manage to register until 10.10., please contact Shuai Li as soon as possible.

- Use email address of sciebo account (university email)

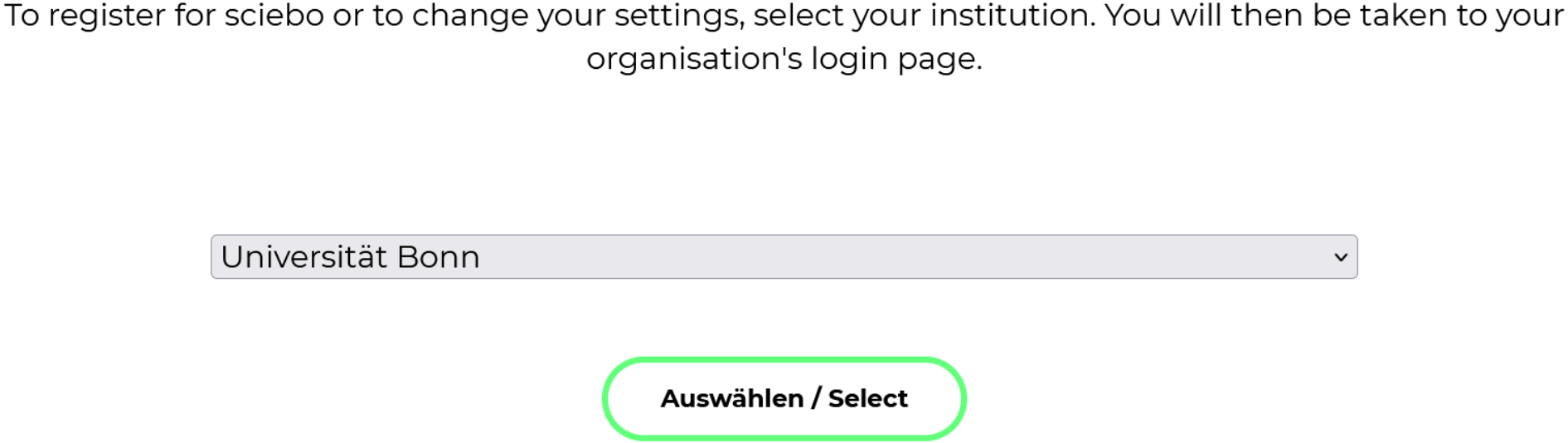
Organization

- Use email address of sciebo account (university email)
- There will be two time slots for the first exercise



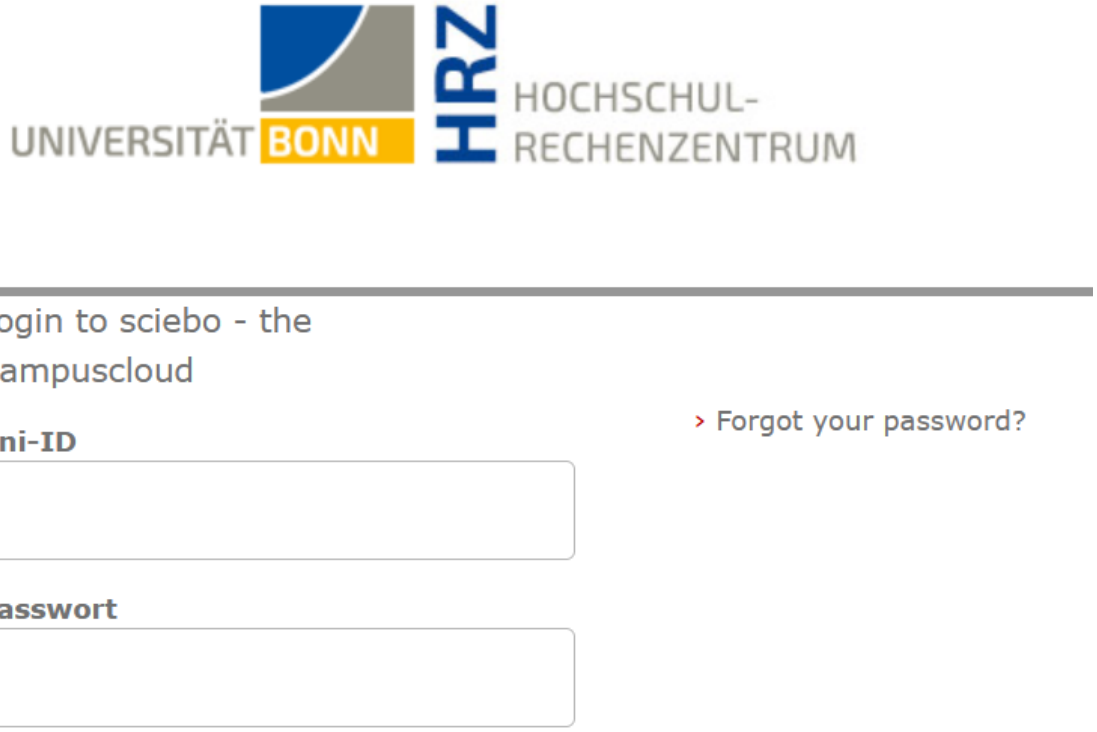
Sciebo

- https://hochschulcloud.nrw/en/
- Registration / Select University



Sciebo

- https://hochschulcloud.nrw/en/
- Registration / Select University
- Login with your university email



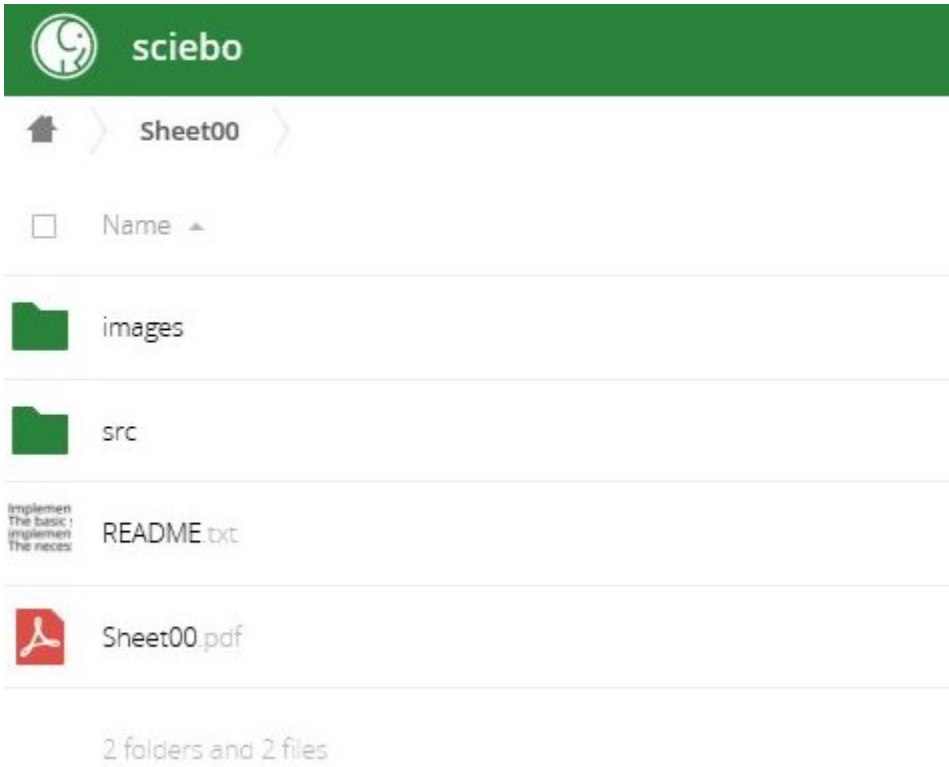
Organization

UNIVERSITÄT BONN

- Registration for exercises today
- Submission in groups of up to 2 persons allowed
- Submission deadline Sunday midnight
- First exercise sheet will be released tomorrow
- Submission for first sheet is Monday, 14.10., midnight
- How to submit: You will get details per email after registration

Organization

- Exercise sheets via sciebo (https://www.sciebo.de):



- Create sciebo account (university email)

Organization

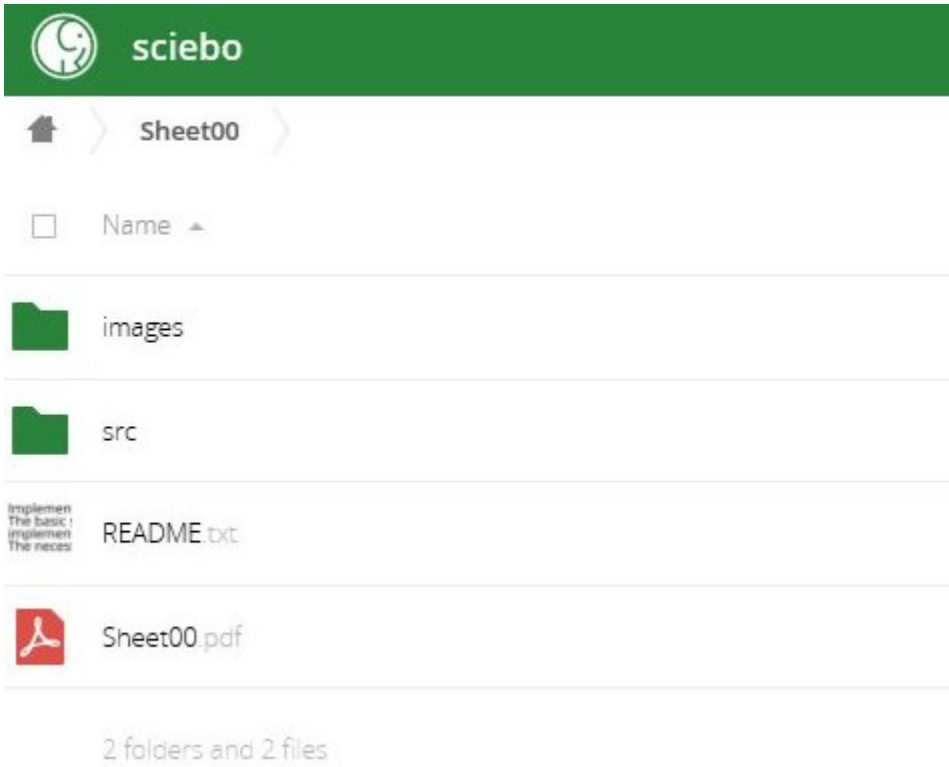
- Exercise sheets via sciebo (https://www.sciebo.de):

Data
(do not upload for submission)

Template
(modify and submit only template and readme; do not change paths)

Readme
(add group members)

Exercise Sheet



Exercises

UNIVERSITÄT BONN

- Use Python and OpenCV (Versions will be specified)
- Submit README.txt with your group members for all exercise sheet (name, student id, email)
- Only one submission per group
- Submit modified template

Exercises

UNIVERSITÄT BONN

- Never submit code that does not run (zero points)
- Add comments to the code
- Use readable variable and function names (e.g. `opticalflow_x` instead of `a`)
- Points can be removed per sheet for unreadable programming style

Exercises

UNIVERSITÄT BONN

Paper exercise:

- Submit as single PDF
- Hand writing → Photo → Image → Save as PDF (gimp)
- You can also use Latex or other tools to create PDF

Registration for Exercise groups

If you have troubles to register or to submit, please contact

Emad Bahrami

bahrami@iai.uni-bonn.de

How to attend a lecture

UNIVERSITÄT BONN

- Attend lecture
- After lecture go through the slides on the same or next day (2-3 hours)
- If something is unclear, check recording / additional reading material
- If something is still unclear write me an email and ask to repeat part of previous lecture
- Try to solve all exercise sheets

Mailinglist

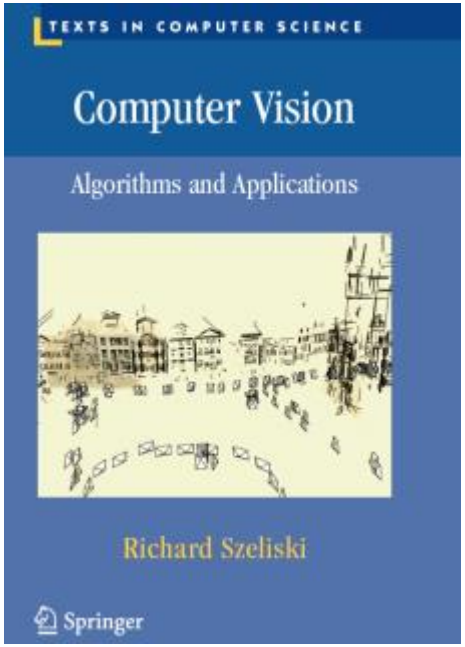
- After registration, you will be added to a mailinglist
- You can unsubscribe at any time

Literature

UNIVERSITÄT BONN

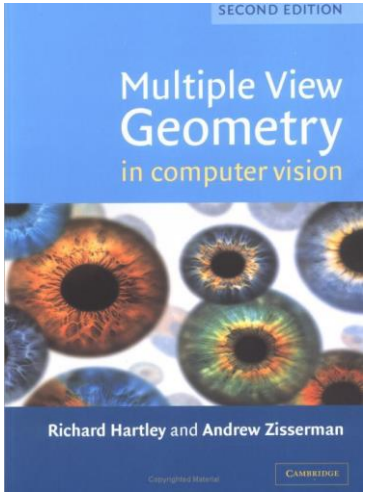
Basic material:

- http://szeliski.org/Book/
- http://www.robots.ox.ac.uk/~vgg/hzbook/



Computer Vision - Algorithms and Applications, Szeliski, Richard, Springer, 2011

Multiple View Geometry in Computer Vision, Second Edition, Richard Hartley and Andrew Zisserman, Cambridge University Press, 2004.



Literature

Basic material:

- http://web4.cs.ucl.ac.uk/staff/s.prince/book/book.pdf

S. Prince. Computer Vision: Models, Learning, and Inference. Cambridge University Press 2012

Literature

Additional material as references on slides.

Main conferences:

- IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
- IEEE International Conference on Computer Vision (ICCV)
- European Conference on Computer Vision (ECCV)

Main journals:

- IEEE Transactions on Pattern Analysis and Machine Intelligence
- International Journal of Computer Vision

Impact of Computer Vision

UNIVERSITÄT BONN

- Impact of CVPR is comparable to Nature or Science

	Publikation	h5-Index	h5-Medianwert
1.	Nature	488	745
2.	IEEE/CVF Conference on Computer Vision and Pattern Recognition	440	689
3.	The New England Journal of Medicine	434	897
4.	Science	409	633
5.	Nature Communications	375	492
6.	The Lancet	368	678
7.	Neural Information Processing Systems	337	614
8.	Advanced Materials	327	420
9.	Cell	320	482
10.	International Conference on Learning Representations	304	584

[https://scholar.google.com/citations?view_op=top_venues]

Impact of Computer Vision

UNIVERSITÄT BONN

- CVPR is in the top 5 of scientific journals or proceedings with highest impact
- Most important field in computer science: Computer Vision
- Second most important field in computer science: Machine Learning

[https://scholar.google.com/citations?view_op=top_venues]

What is Computer Vision?

UNIVERSITÄT BONN

Convert visual data into “meaningful” numbers

Images

Stereo

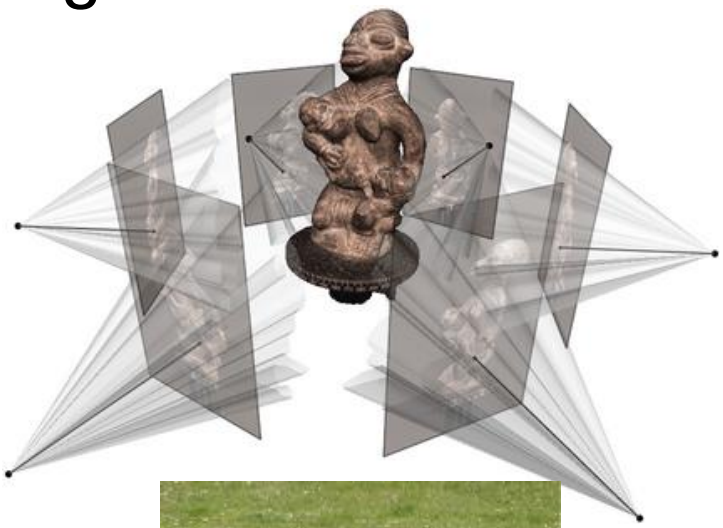
Videos

Time-of-Flight

What are “meaningful” numbers?

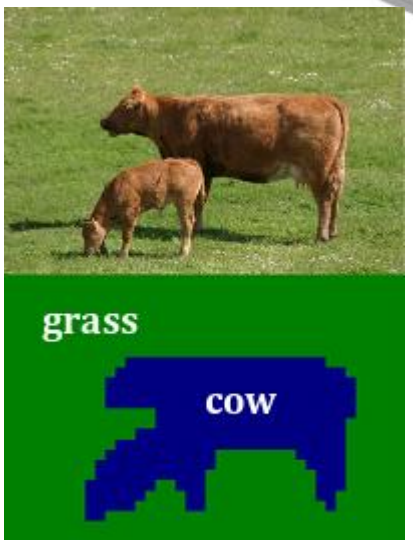
From low-level (pixel values) to a high-level representation:

– Metric measurements



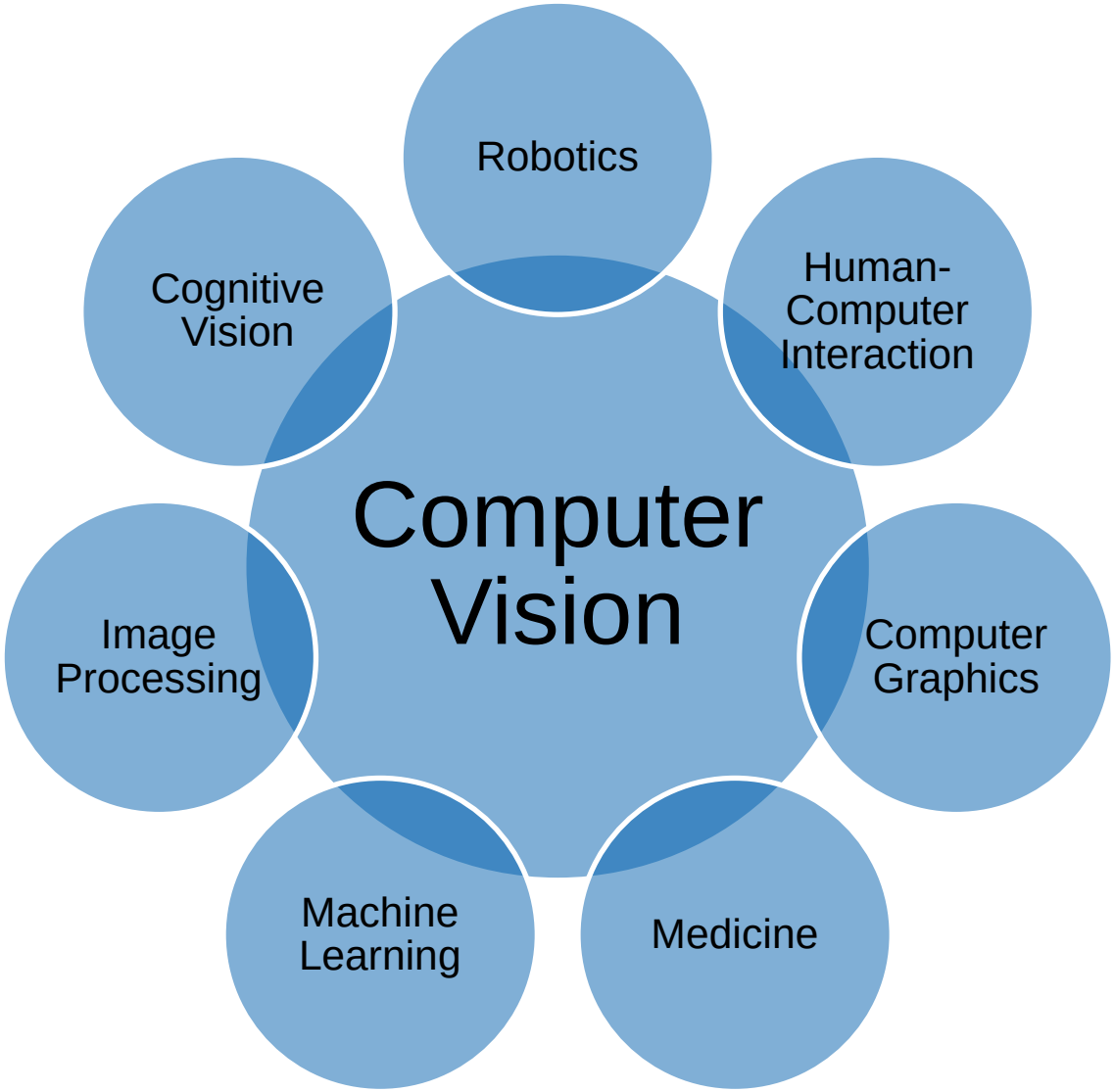
– Models

– Labels



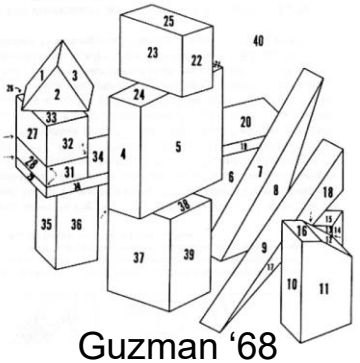
Connections to other disciplines

UNIVERSITÄT BONN

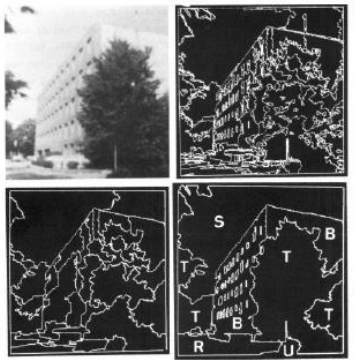


Brief history of computer vision

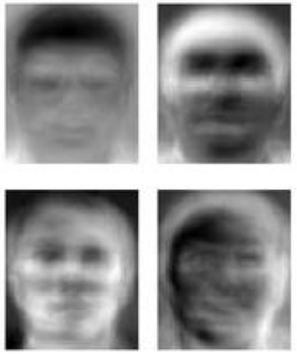
1960’s: interpretation of synthetic worlds
1970’s: some progress on interpreting selected images
1980’s: shift toward geometry and increased mathematical rigor
1990’s: face recognition; statistical analysis in vogue
2000’s: broader recognition; large annotated datasets available; video processing starts
2010’s: deep learning; many commercial applications and successful spin-offs



Guzman ‘68



Ohta Kanade ‘78



Turk and Pentland ‘91

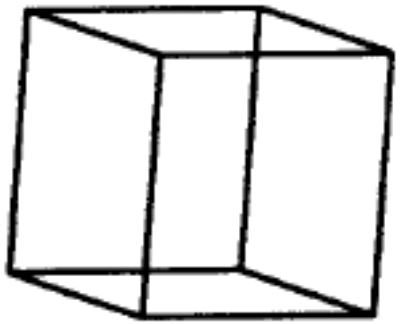
Source: J. Hays

Inherent ambiguity of the problem

Many different 3D scenes could have given rise to a particular 2D picture

Source: S. Lazebnik

A simple idea to recover 3D shapes from line drawings



Task: Given the set of image coordinates of the vertices (x_i, y_i), recover the world coordinates (X_i, Y_i, Z_i).

We will assume orthographic projection so that $X_i = x_i$ and $Y_i = y_i$

Then, the problem is: recover the missing depth Z_i

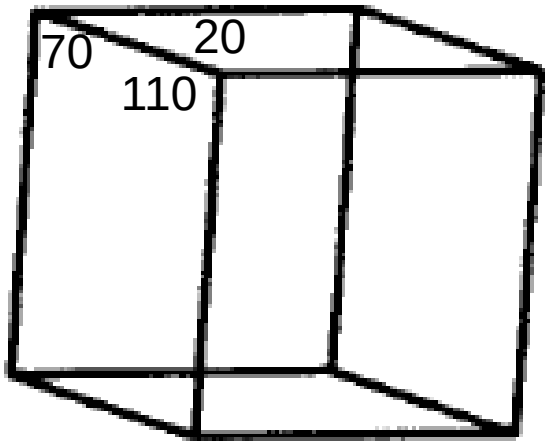
Marill, AI-Memo-1136, 1989

Source: A. Torralba

A simple idea to recover 3D shapes from line drawings

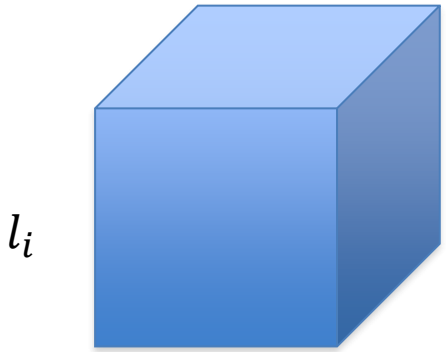
Heuristic: find Z_k that minimize the standard deviation of angles in the 3D object

Possible solution 1



$Z_k = 0$ for all k

Possible solution 2



All 90 degree angles

Minimize standard deviations of angles:

$$\underset{Z_k}{\operatorname{argmin}} \frac{1}{N-1} \sum \left(\alpha_{ij} - \frac{1}{N} \sum \alpha_{ij} \right)^2$$

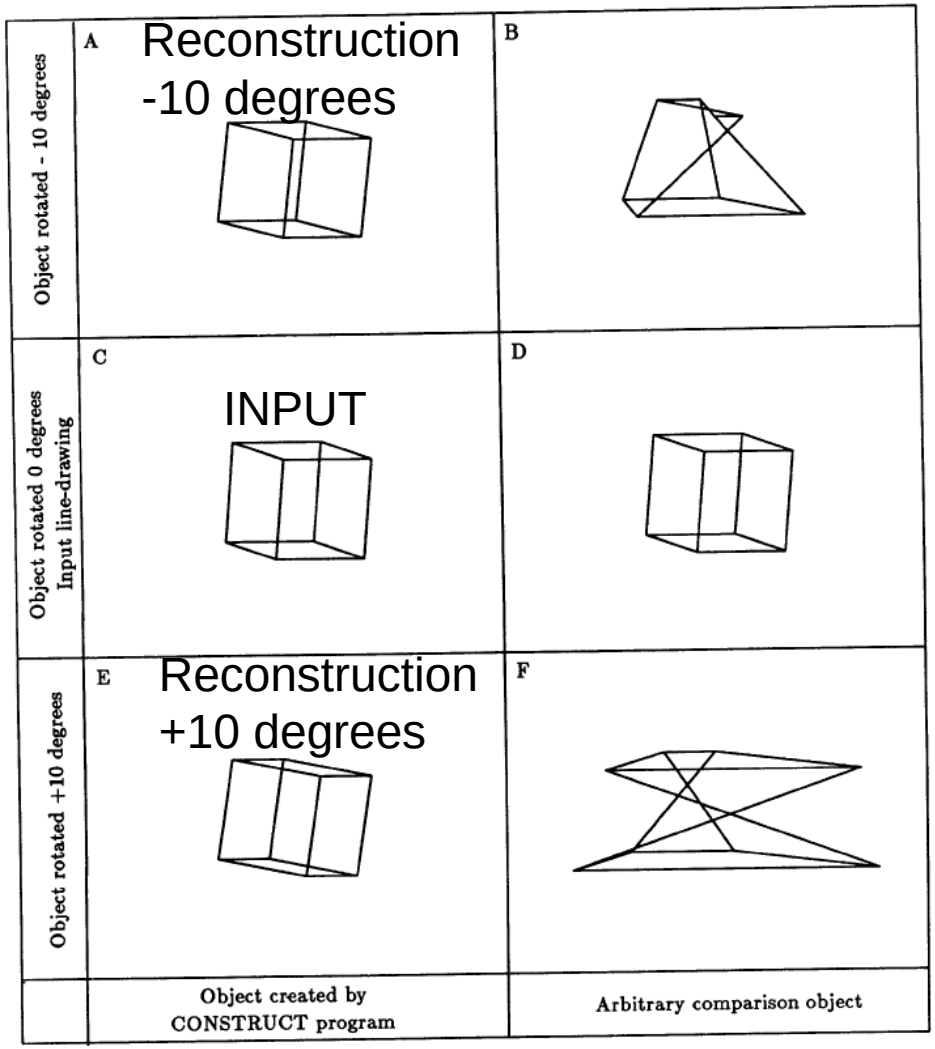
$$P_k = (X_k, Y_k, Z_k)$$

$$\alpha_{ij} = \cos^{-1}(\mathbf{u_i} \cdot \mathbf{u_j})$$

unit length vectors: $u_i = \frac{l_i}{\|l_i\|}$

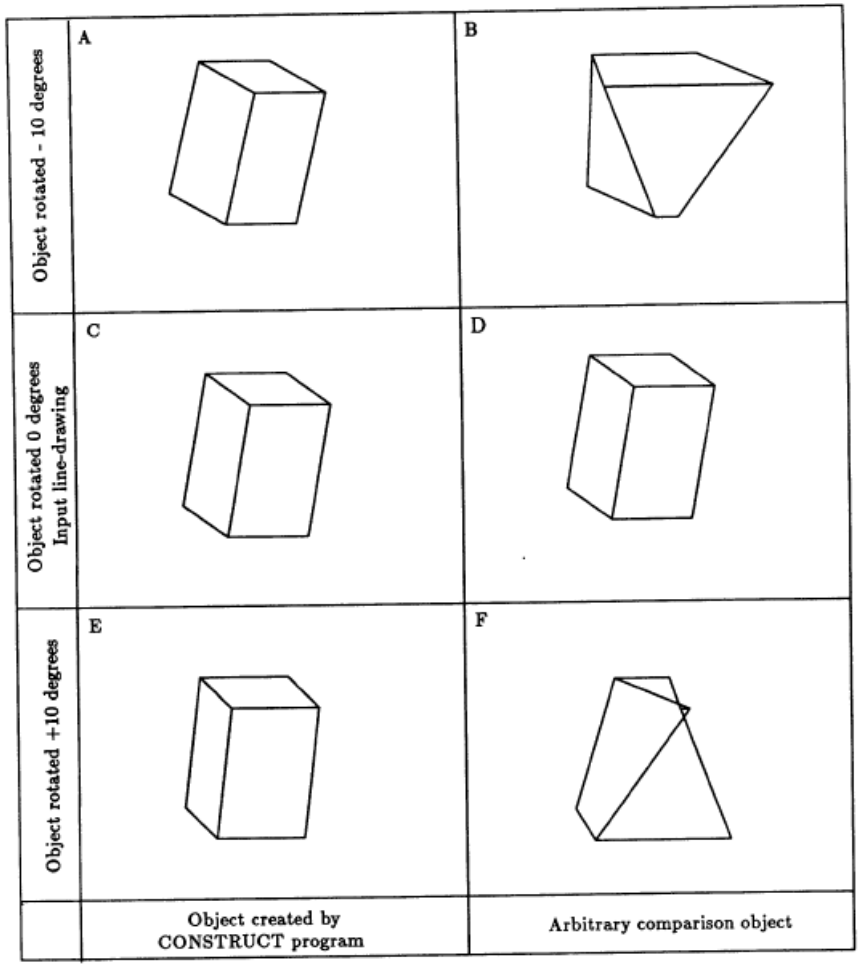
Source: A. Torralba

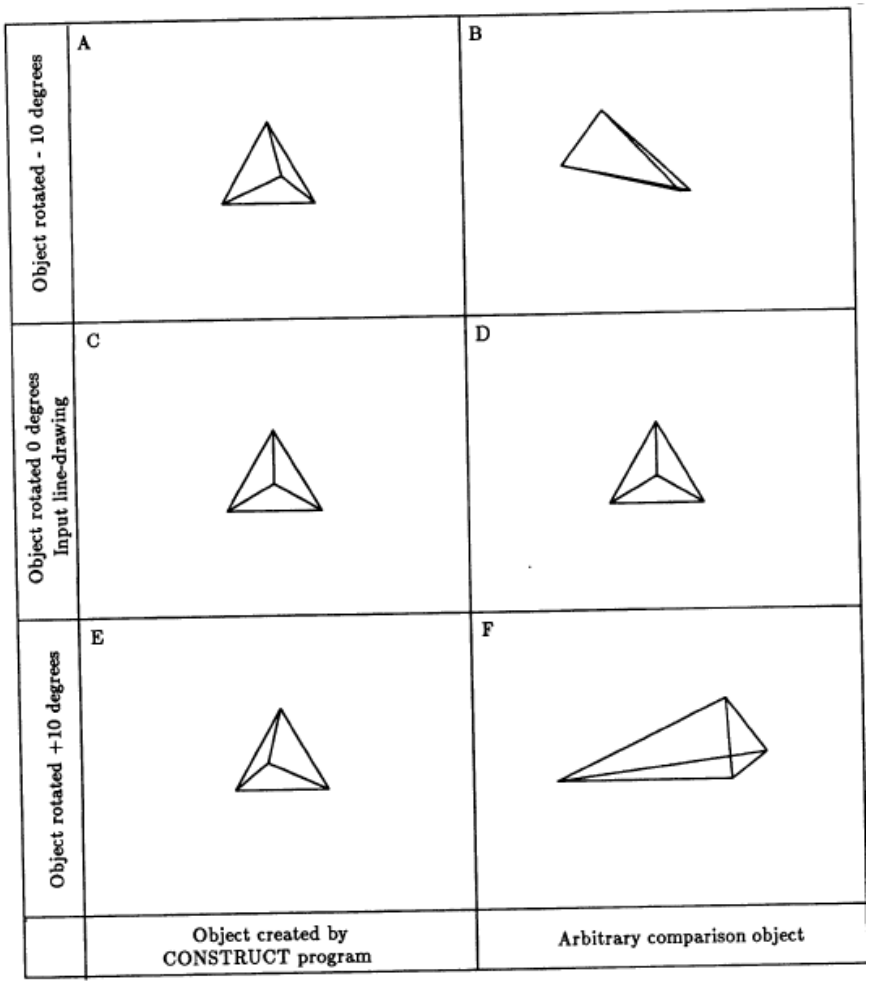
A simple idea to recover 3D shapes from line drawings



Source: A. Torralba

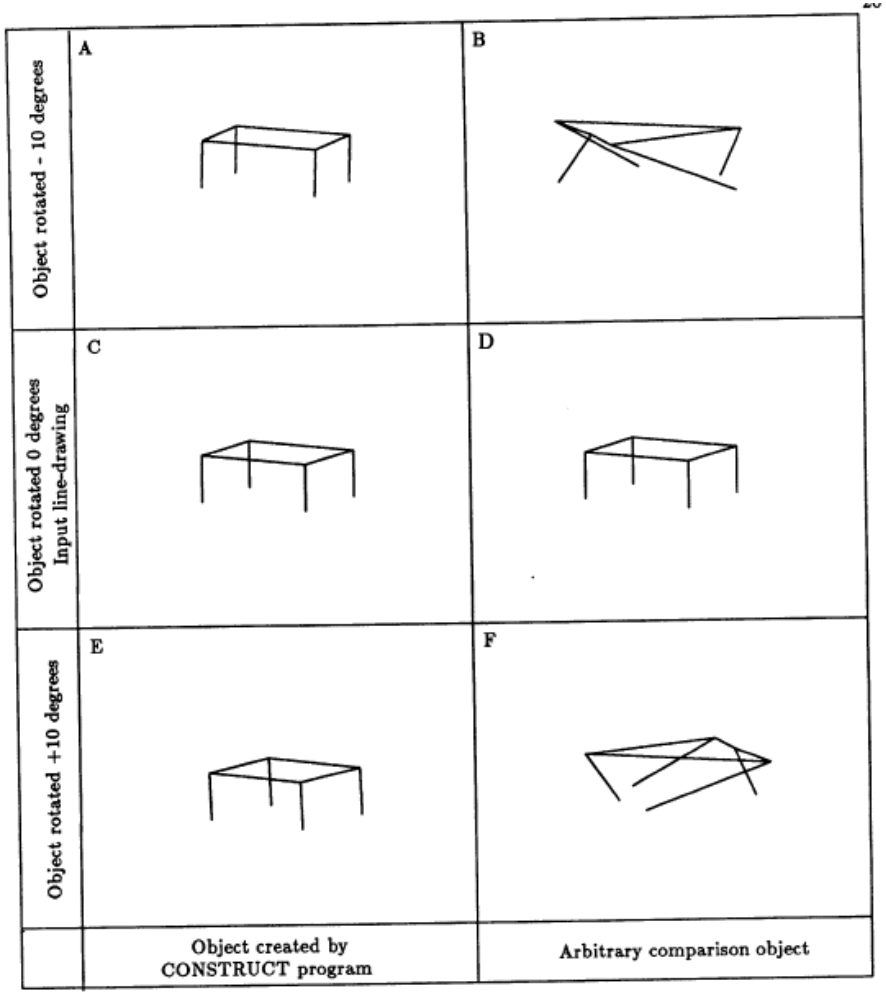
A simple idea to recover 3D shapes from line drawings

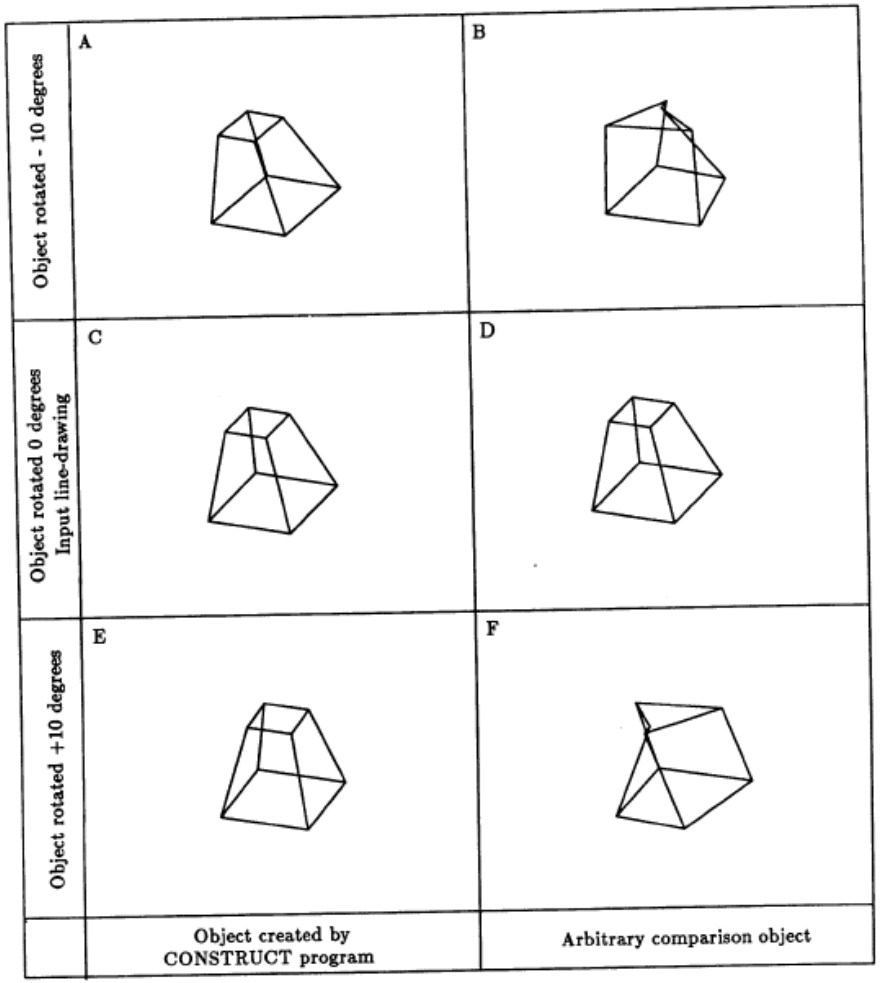




Source: A. Torralba

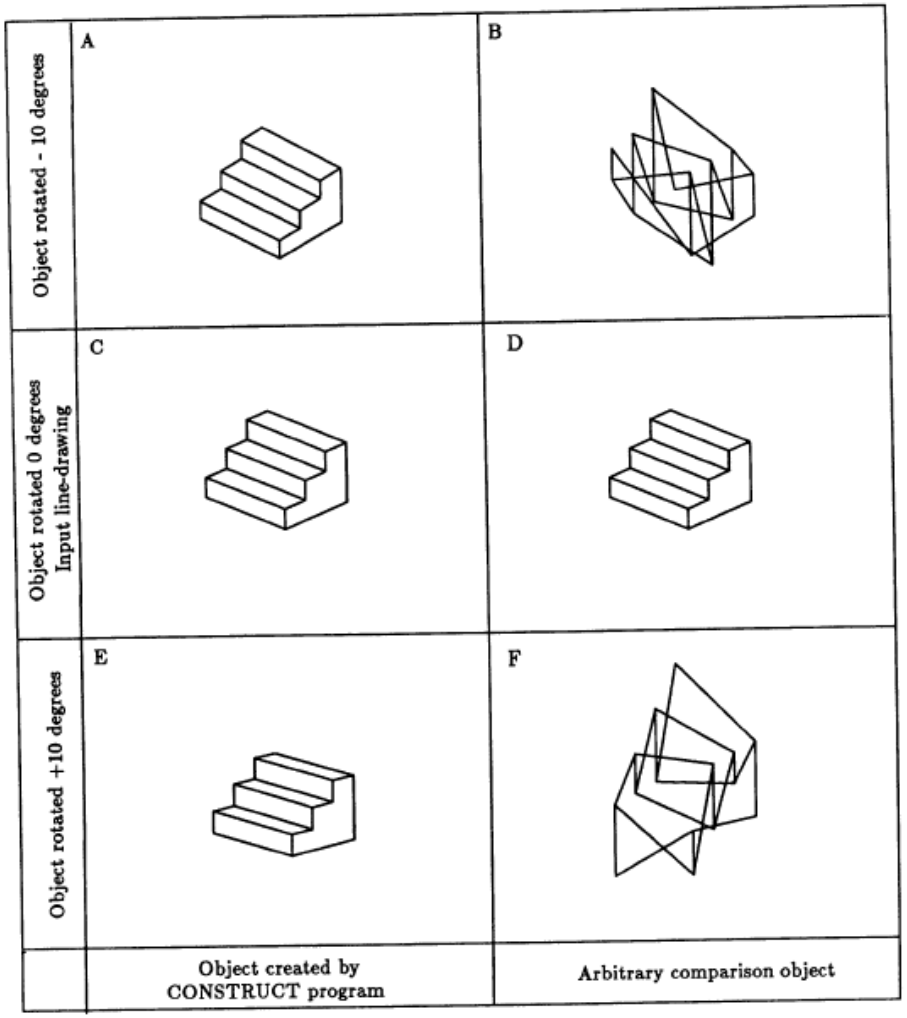
A simple idea to recover 3D shapes from line drawings

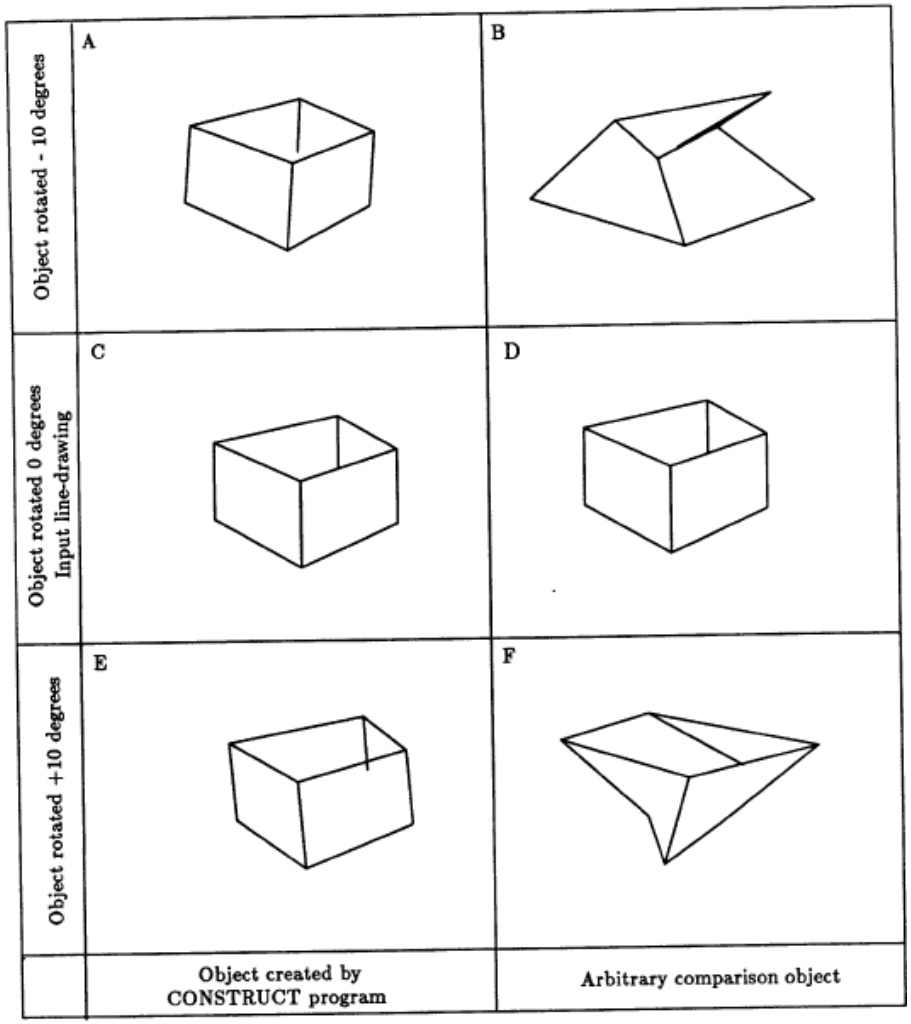




Source: A. Torralba

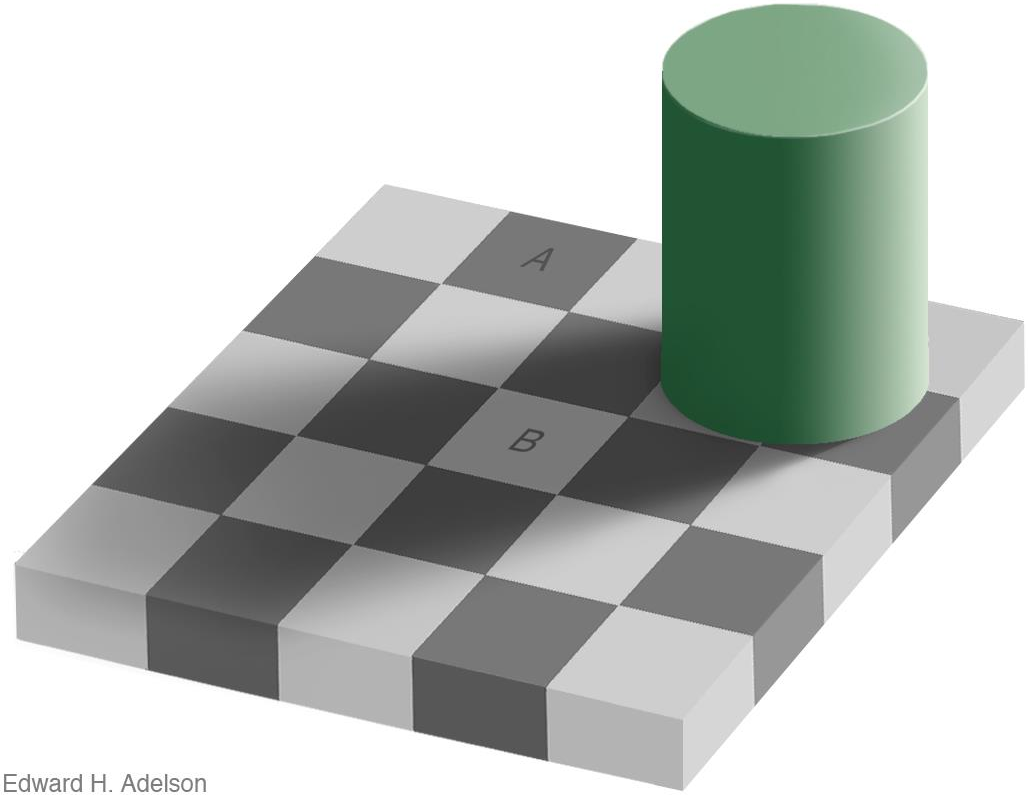
A simple idea to recover 3D shapes from line drawings





Source: A. Torralba

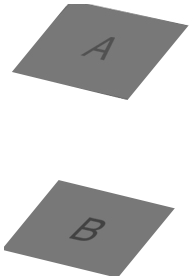
Measuring light vs. measuring scene properties



Edward H. Adelson

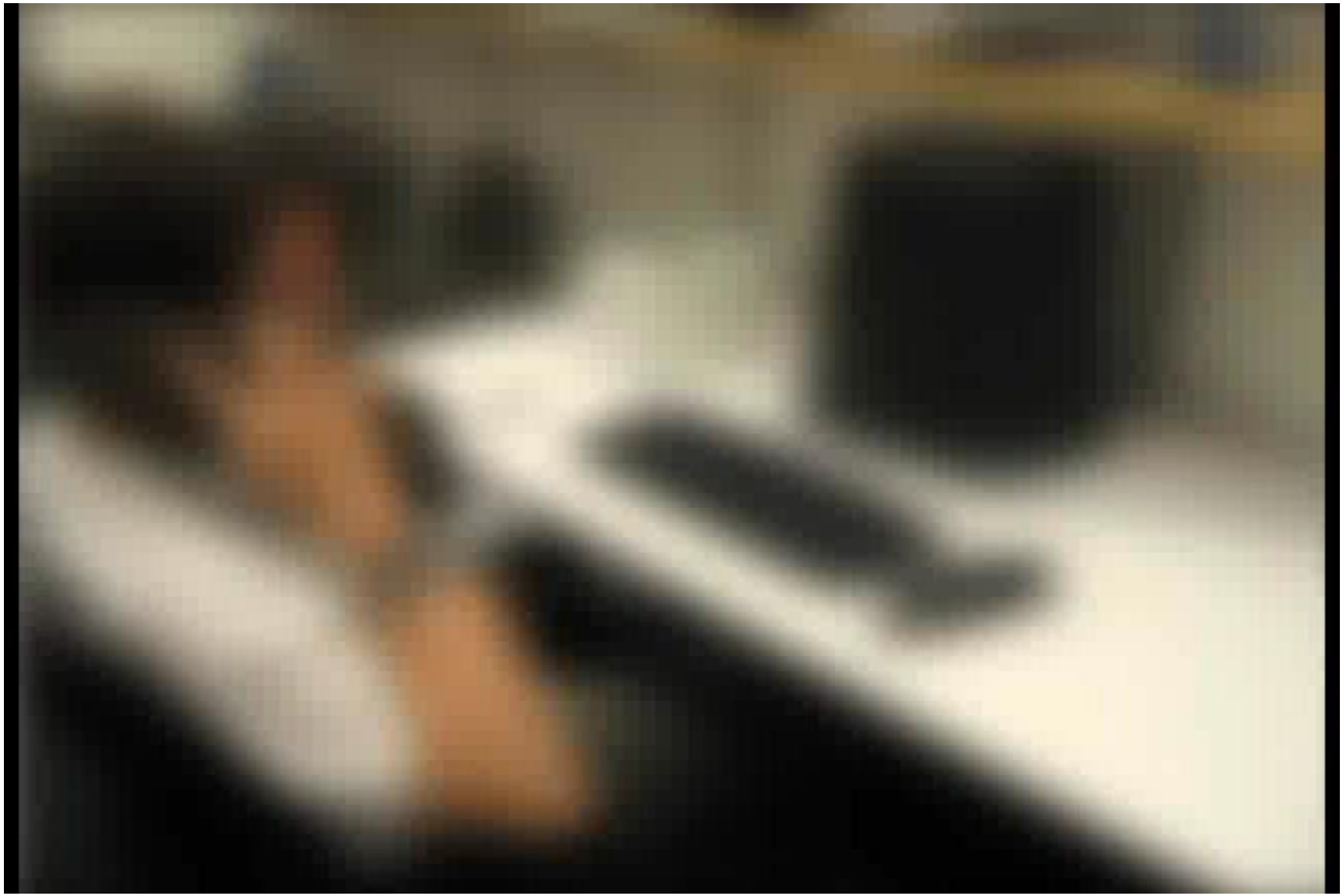
Source: A. Torralba

Measuring light vs. measuring scene properties



Source: A. Torralba

Context



Source: A. Torralba

Humans work with strong priors

Source: A. Torralba

Conclusion

- Human vision relies not **only** on data term (pixel information)
- Priors are important
- They explain the **most likely** scene if data is ambiguous:

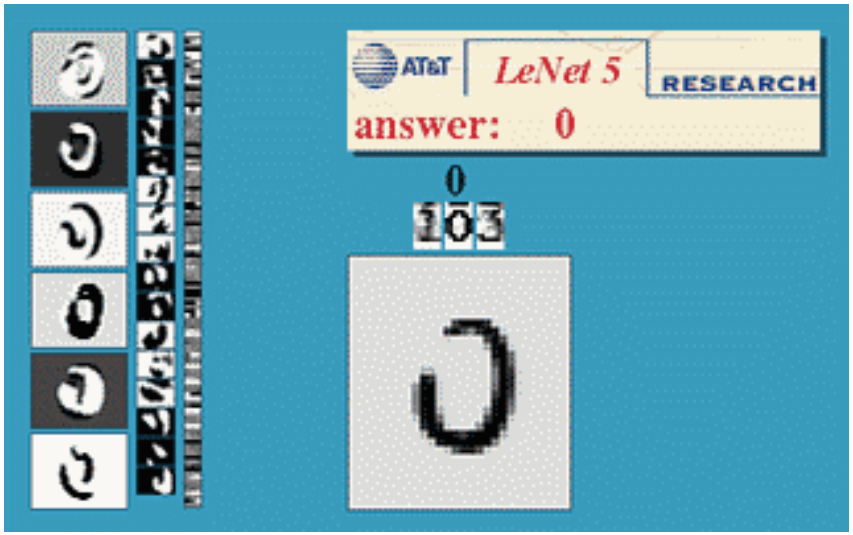
$$\operatorname*{argmax}_y P(y|I) = \operatorname*{argmax}_y P(I|y)P(y)$$

- Bayes formula:

$$P(y|I) = \frac{P(I|y)P(y)}{P(I)}$$

Posterior = Likelihood x Prior

Optical character recognition (OCR)



Digit recognition

License plate readers

Automatic check processing

Source: S. Seitz

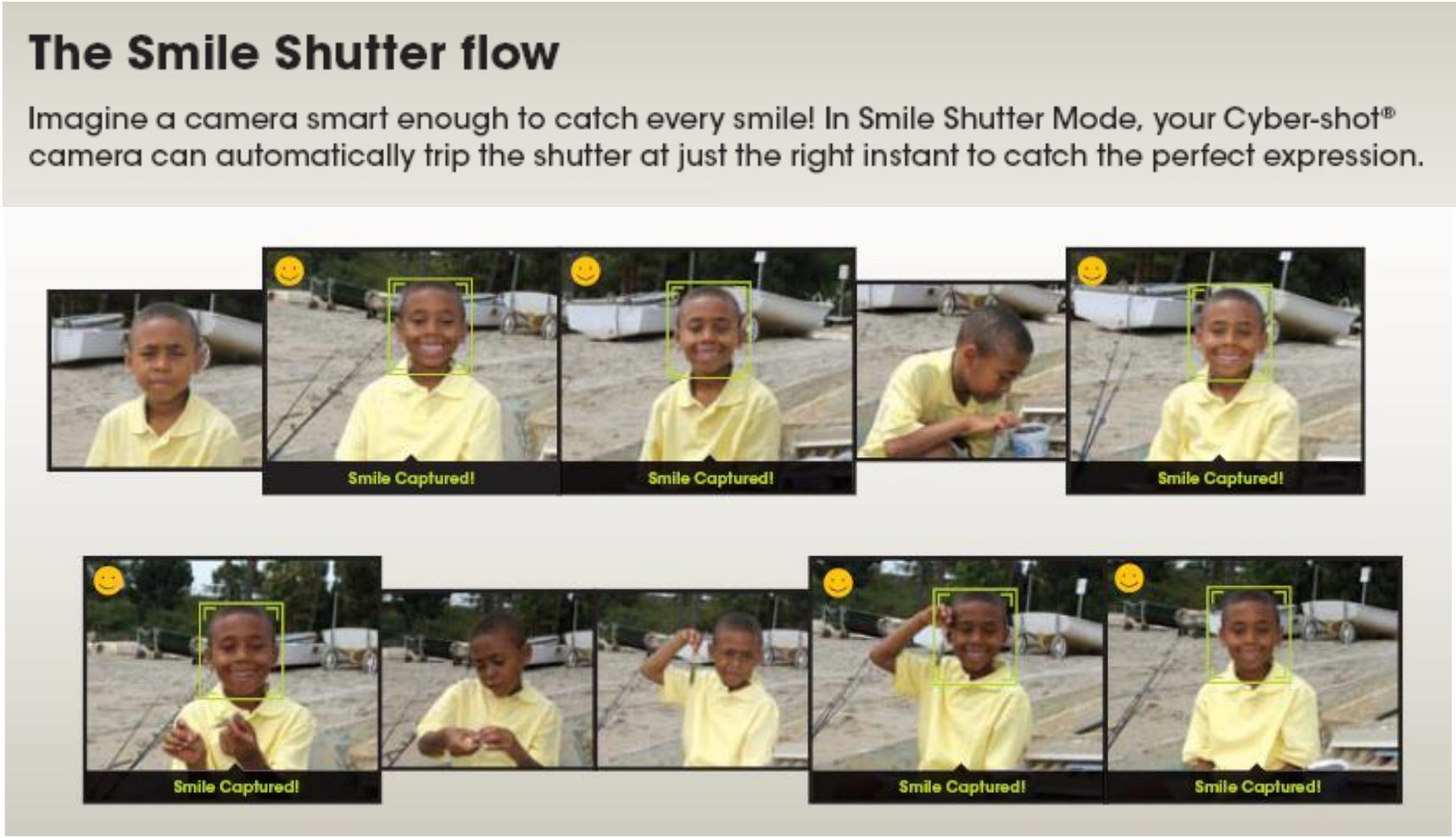
Face detection

UNIVERSITÄT BONN

Digital cameras detect faces

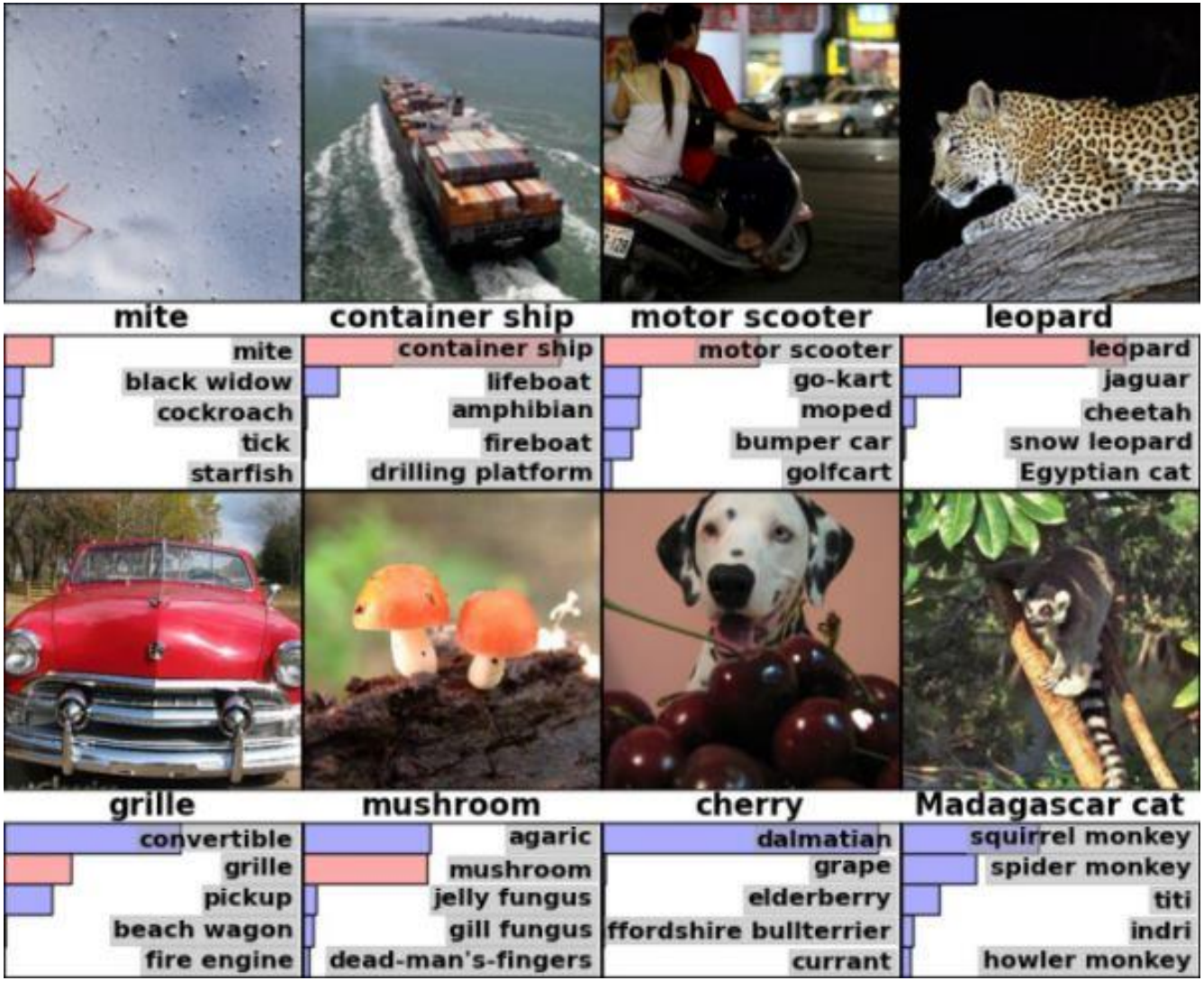
Source: S. Seitz

Smile detection



Source: S. Seitz

Image classification



http://www.clarifai.com/

Automotive safety

UNIVERSITÄT BONN

Pedestrian collision warning, forward collision warning, lane departure warning…

Vision-based interaction: Xbox Kinect

Movie production

Generative AI










Vision for robotics, space exploration

NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- Matthies et al. Computer Vision on Mars IJCV'07

Source: S. Seitz

Jobs, jobs, …


facebook
Adobe
Microsoft
amazon
intel
Google
Athena eyes
智慧眼
MITSUBISHI ELECTRIC
Changes for the Better
MITSUBISHI ELECTRIC
RESEARCH LABORATORIES, INC
NVIDIA
A9
图森
tu Simple
Prism
dji 大疆創新
TOYOTA
RESEARCH INSTITUTE
megvii
Face++
BOSCH
Invented for life
DiDi
More than a journey
iRobot
SKYDIO
UBER
ADVANCED TECHNOLOGIES CENTER
Baidu 百度
DAQRI
SenseTime
商湯集團
Sighthound
Linkface
COGNEX
IBM Research
ebay
here
spare 5
360
www.360.cn
HISCENE
亮风台信息科技
MathWorks
Disney Research
SIEMENS

Curriculum

UNIVERSITÄT BONN

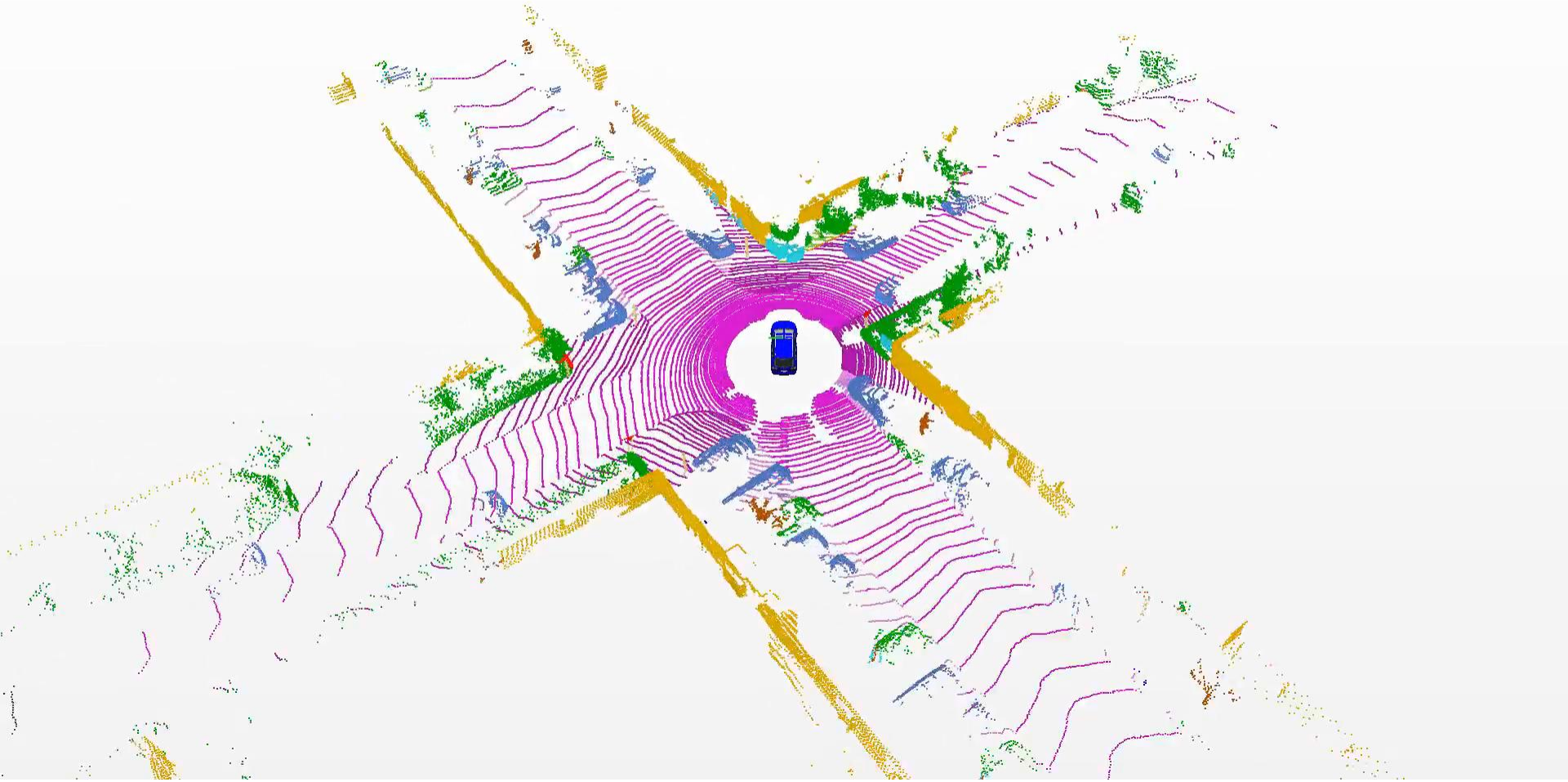
1. Semester
Computer Vision Lecture (WS)

2. and 3. Semester
Lab Vision (SS+WS)
Seminar (SS+WS)
Advanced Computer Vision Lecture (SS)

4. Semester
Master thesis

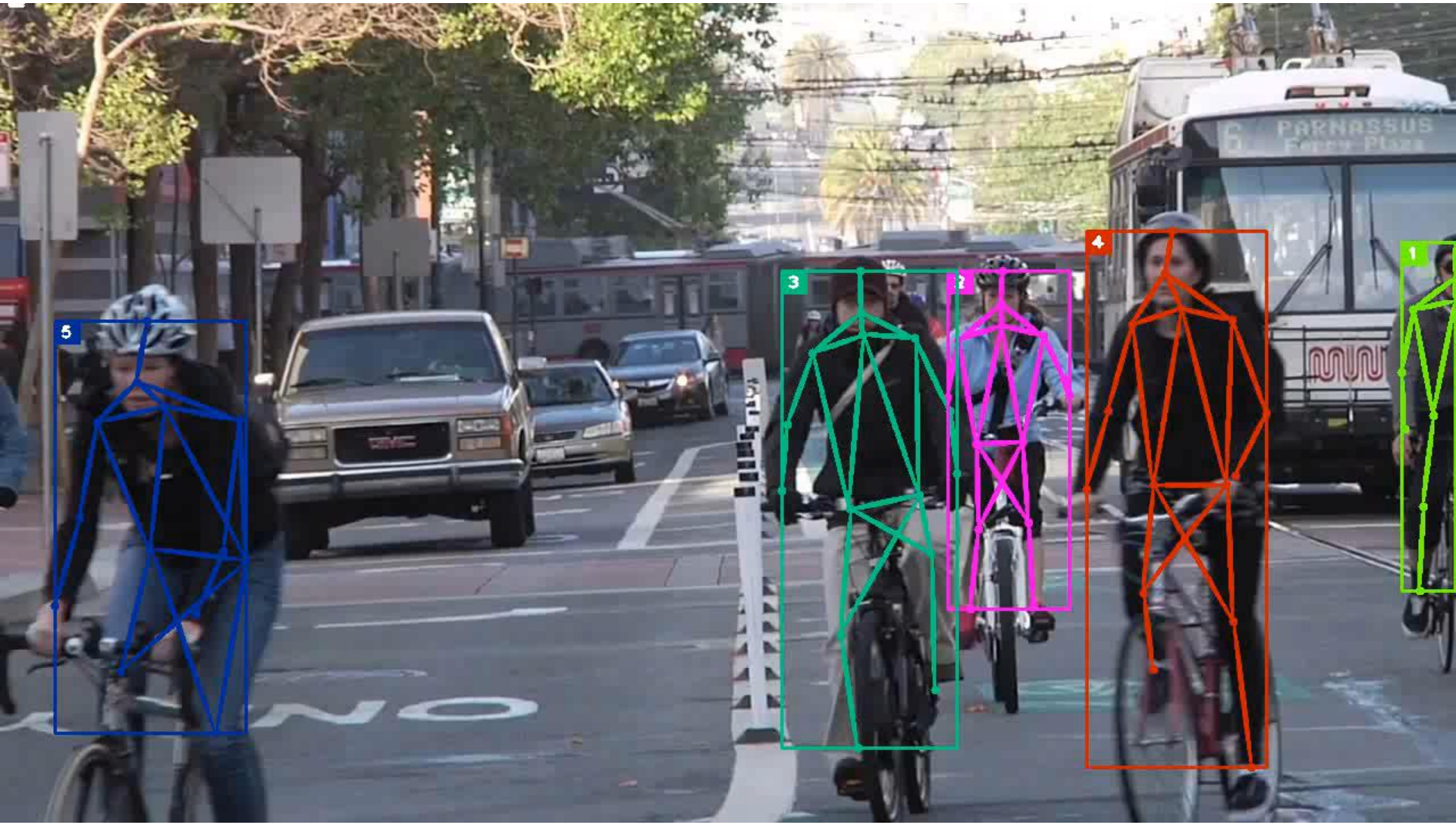
Semantic Segmentation

UNIVERSITÄT BONN



[J. Behley et al. **SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences.** ICCV 2019]

Human Pose Estimation



[A. Doering et al. **Joint Flow: Temporal Flow Fields for Multi Person Tracking.** BMVC 2018]

Video Understanding

UNIVERSITÄT BONN

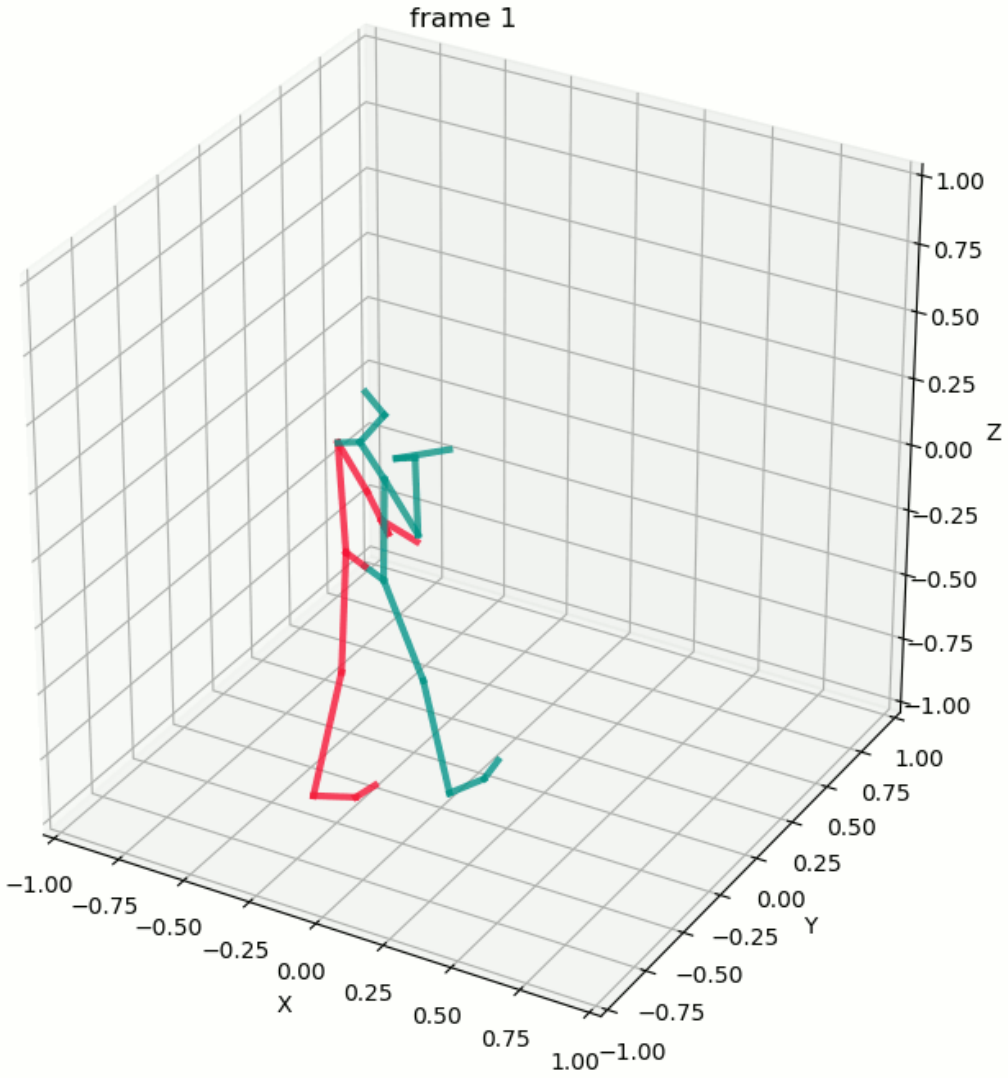
Video Captioning Task Results

Action Segmentation

Anticipating Actions

[Y. Abu Farha et al. **When will you do what? Anticipating Temporal Occurrences of Activities.** CVPR 2018]

Forecast Human Pose



[Ruiz et al. **Human Motion Prediction via Spatio-Temporal Inpainting.** ICCV 2019]

Research Unit - Anticipating Human Behavior

[https://for2535.cv-uni-bonn.de]

Cluster of Excellence - PhenoRob

UNIVERSITÄT BONN

[http://www.phenorob.de]

Organization

UNIVERSITÄT BONN

- Slides:

  Password: MA-INF2201

- Registration for exercises today
- Registration of sciebo account
- Submission for first sheet (5 bonus points) is Monday, 14.10.

Thank you for your attention.