Computer Vision – Lecture 1

Introduction

09.04.2019

Bastian Leibe
Visual Computing Institute
RWTH Aachen University
http://www.vision.rwth-aachen.de/

leibe@vision.rwth-aachen.de

Organization

- Lecturer
  - Prof. Bastian Leibe (leibe@vision.rwth-aachen.de)
- Teaching Assistants
  - Istvan Sarandi (sarandi@vision.rwth-aachen.de)
  - Dan Jia (jia@vision.rwth-Aachen.de)
- Course webpage
  - http://www.vision.rwth-aachen.de/courses/
    → Computer Vision
  - Slides will be made available on the webpage
  - There is also a moodle electronic repository
- Please subscribe to the lecture on RWTH Online!
  - Important to get email announcements and moodle access!

Language

- Official course language will be English
  - If at least one English-speaking student is present.
  - If not… you can choose.
- However…
  - Please tell me when I'm talking too fast or when I should repeat something in German for better understanding!
  - You may at any time ask questions in German!
  - You may turn in your exercises in German.
  - You may answer exam questions in German.

Organization

- Structure: 3V (lecture) + 1Ü (exercises)
  - 6 EECS credits
  - Part of the area "Applied Computer Science"
- Place & Time
  - Lecture: Mon 10:30 – 12:00 TEMP2
  - Lecture/Exercises: Tue 14:30 – 16:00 H03
- Exam
  - Written exam
  - Dates will be communicated soon

Exercises and Demos

- Exercises
  - Typically 1 exercise sheet every 2 weeks (numpy/TensorFlow)
  - Hands-on experience with the algorithms from the lecture.
  - Send in your solutions the night before the exercise class.
  - No admission requirement to qualify for the exam this year!
- Teams are encouraged!
  - You can form teams of up to 3 people for the exercises.
  - Each team should only turn in one solution.
  - But list the names of all team members in the submission.

Course Webpage

Course Schedule

Date	Title	Content	Material
Mon, 2019-04-01	--	no class (RWTH DIES)	
Tue, 2019-04-02	--	no class (RWTH DIES)	
Mon, 2019-04-08	--	no class	
Tue, 2019-04-09	Introduction	Why vision? Applications, Challenges, Image Formation	
Mon, 2019-04-15	Image Processing I	Linear Filters, Gaussian Smoothing, Multi-scale Representations	
Tue, 2019-04-16	TBD	TBD	
Mon, 2019-04-22	--	no class (Easter Monday)	
Tue, 2019-04-23	Image Processing II	Image Derivatives, Edge detection, Canny	
Mon, 2019-04-29	Structure Extraction	Line Fitting, Hough Transform, Gen. Hough Transform	
Tue, 2019-04-30	Segmentation I	Segmentation as Clustering, k-means, EM, Mean-Shift Segmentation	
Mon, 2019-05-06	Exercise 1	Derivatives, Edges, Hough Transform	
Tue, 2019-05-07	Segmentation II	Segmentation as Energy Minimization, Markov Random Fields, Graph Cuts	
Mon, 2019-05-13	Categorization I	Sliding Window-based Object Detection, HOG, SVMs, Viola-Jones detector, AdaBoost	

http://www.vision.rwth-aachen.de/courses/

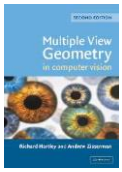
Textbooks

- No single textbook for the class.
- Basic material is covered in the following two books.

R. Szeliski
Computer Vision – Algorithms and Applications
Springer, 2010

(available in the library's "Handapparat")

R. Hartley, A. Zisserman
Multiple View Geometry in Computer Vision
2nd Ed., Cambridge Univ. Press, 2004



- Additional material will be given out for some topics.
  - Tutorials and deeper introductions.
  - Application papers

How to Find Us

- Office:
  - UMIC Research Centre
  - Mies-van-der-Rohe-Strasse 15, room 124

- Office hours
  - If you have questions to the lecture, come to us.
  - Send us an email before to confirm a time slot.

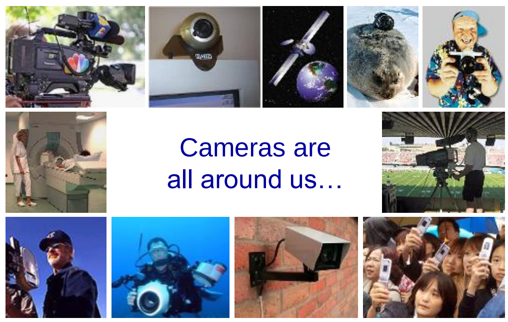
Questions are welcome!

Topics of Today's Lecture

- What is computer vision?
- What does it mean to see and how do we do it?
- How can we make this computational?

- First Topic: Image Formation
  - Details in Forsyth & Ponce, chapter 1.

Why Computer Vision?



Cameras are
all around us…

Images and video are everywhere…

Personal photo albums

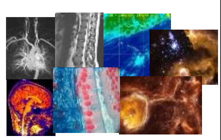
Movies, news, sports

Google Picasa flickr webshots picsearch YouTube kooaba

Internet services

Surveillance and security

Mobile and consumer applications



Medical and scientific images

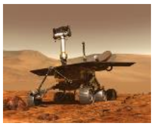
What is Computer Vision?

- Goal of Computer Vision
  - *Enable a machine to "understand" images and videos*

- Automatic understanding
  - Computing properties of the 3D world from visual data *(measurement)*
  - Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities. *(perception and interpretation)*

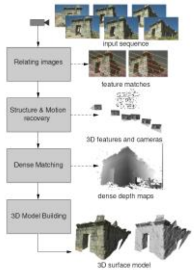
Vision for Measurement

Real-time stereo



Pollefeys et al.

Structure from motion



Multi-view stereo for community photo collections

Goesele et al.

Slide credit: Svetlana Lazebnik

B. Leibe

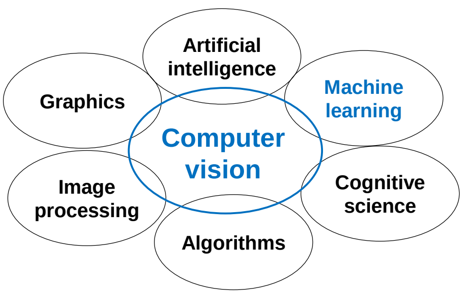
13

Vision for Perception, Interpretation

Slide credit: Kristen Grauman

14

Related Disciplines



B. Leibe

15

Directions to Computer Vision

- Science
  - Foundations of perception. How do WE see?
- Engineering
  - How do we build systems that perceive the world?
- Many applications
  - Medical imaging, surveillance, entertainment, graphics, …

B. Leibe

16

Applications: Faces and Digital Cameras

Setting camera focus via face detection

Camera waits for everyone to smile to take a photo [Canon]

Automatic lighting correction based on face detection

Slide credit: Kristen Grauman, Rob Fergus

B. Leibe

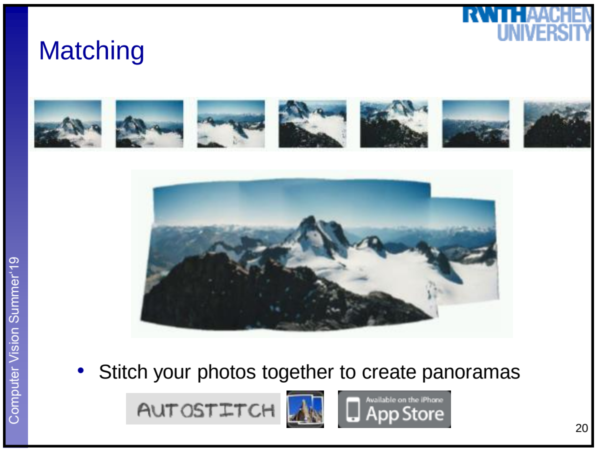
17

Segmentation

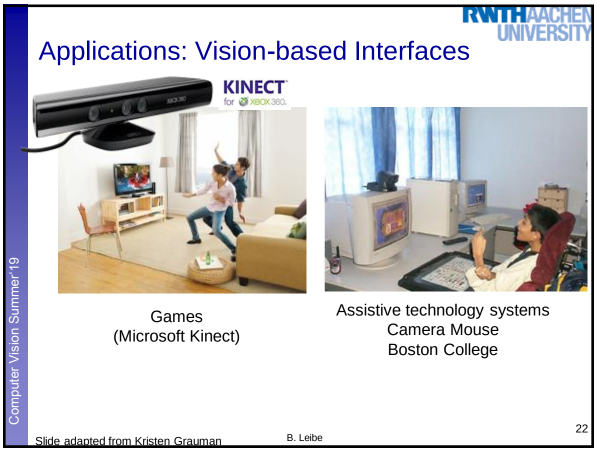
- Automatic background removal from images
  - Functionality is included in Microsoft Office 2010…

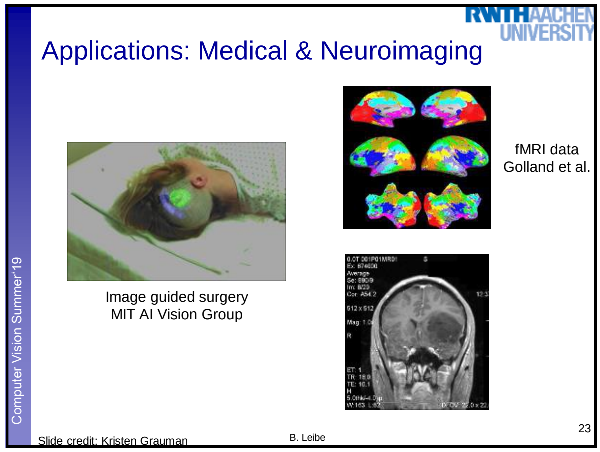
Microsoft Office 2010

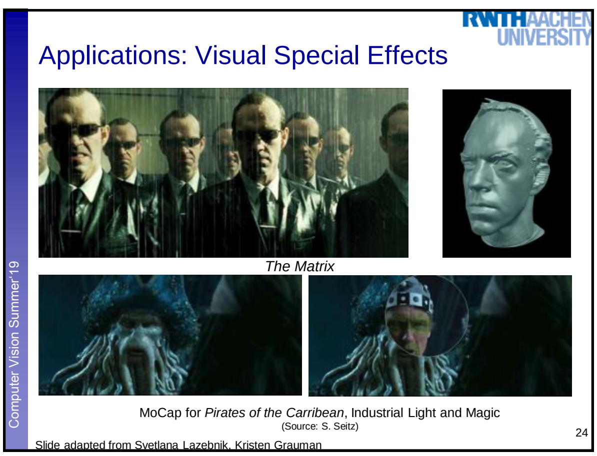
19

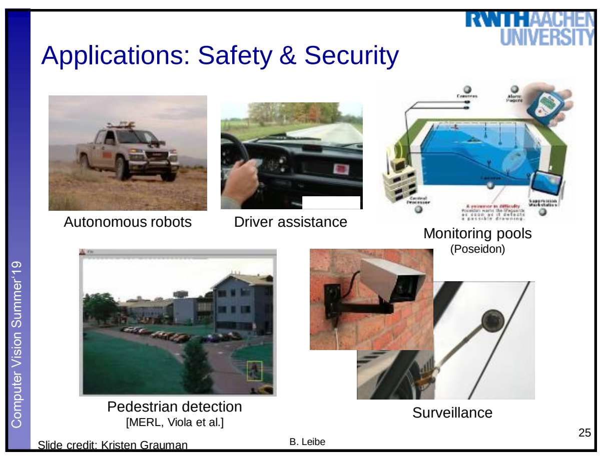

RWTH AACHEN UNIVERSITY
Computer Vision Summer'19
Matching
• Stitch your photos together to create panoramas
AUTOSTITCH
Available on the iPhone App Store
20


RWTH AACHEN UNIVERSITY
Computer Vision Summer'19
Applications: Vision for Mobile Phones
Google Goggles in Action
Click the icons below to see the different ways Google Goggles can be used.
Google goggles labs
Landmark
Book
Contact Info.
Artwork
Places
Wine
Logo
• Take photos of objects as queries for visual search
Slide credit: Svetlana Lazebnik


RWTH AACHEN UNIVERSITY
Computer Vision Summer'19
Applications: Vision-based Interfaces
KINECT
for XBOX 360.
Games
(Microsoft Kinect)
Assistive technology systems
Camera Mouse
Boston College
Slide adapted from Kristen Grauman
B. Leibe
22


RWTH AACHEN UNIVERSITY
Computer Vision Summer'19
Applications: Medical & Neuroimaging
fMRI data
Golland et al.
Image guided surgery
MIT AI Vision Group
Slide credit: Kristen Grauman
B. Leibe
23


RWTH AACHEN UNIVERSITY
Computer Vision Summer'19
Applications: Visual Special Effects
The Matrix
MoCap for Pirates of the Carribean, Industrial Light and Magic
(Source: S. Seitz)
Slide adapted from Svetlana Lazebnik, Kristen Grauman
24


RWTH AACHEN UNIVERSITY
Computer Vision Summer'19
Applications: Safety & Security
Autonomous robots
Driver assistance
Monitoring pools
(Poseidon)
Pedestrian detection
[MERL, Viola et al.]
Surveillance
Slide credit: Kristen Grauman
B. Leibe
25

Ok, Let's Do It – Any Obstacles?

- 1966: Seymour Papert directs an undergraduate student to solve "the problem of computer vision" as a summer project.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

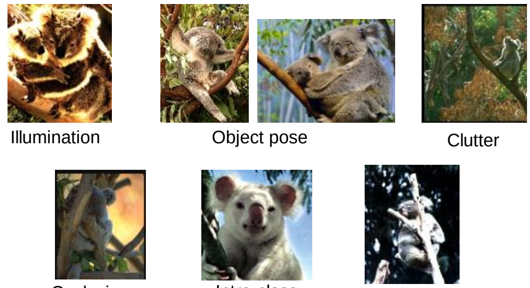
Artificial Intelligence Group July 7, 1966
Vision Memo. No. 100.

THE SUMMER VISION PROJECT
Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

- Obviously, computer vision was too difficult for that…

Challenges: Many Nuisance Parameters



Illumination | Object pose | Clutter

Occlusions | Intra-class appearance | Viewpoint

Slide credit: Kristen Grauman

Challenges: Intra-Category Variation

Slide credit: Fergus, FeiFei, Torralba

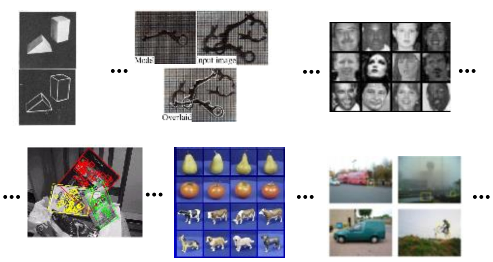
Challenges: Complexity

- Thousands to millions of pixels in an image
- 3,000-30,000 human recognizable object categories
- 30+ degrees of freedom in the pose of articulated objects (humans)
- Billions of images indexed by Google Image Search
- 18 billion+ prints produced from digital camera images in 2004
- 295.5 million camera phones sold in 2005
- About half of the cerebral cortex in primates is devoted to processing visual information [Felleman and van Essen 1991].

Slide credit: Kristen Grauman

So, Should We Give Up?

- NO! Very active research area with exciting progress!



Slide credit: Kristen Grauman

Things Are Starting to Work…

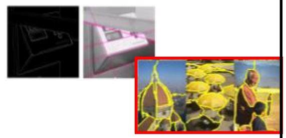
Computer Vision in realistic scenarios is becoming feasible!

Course Outline

- Image Processing Basics
- Segmentation
- Local Features & Matching
- Object Recognition and Categorization
- 3D Reconstruction

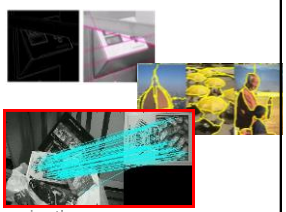
Course Outline

- Image Processing Basics
- Segmentation
- Local Features & Matching
- Object Recognition and Categorization
- 3D Reconstruction



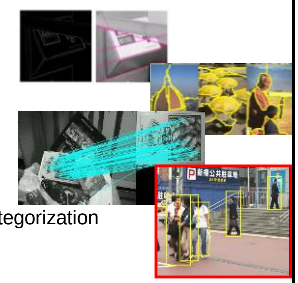
Course Outline

- Image Processing Basics
- Segmentation
- Local Features & Matching
- Object Recognition and Categorization
- 3D Reconstruction



Course Outline

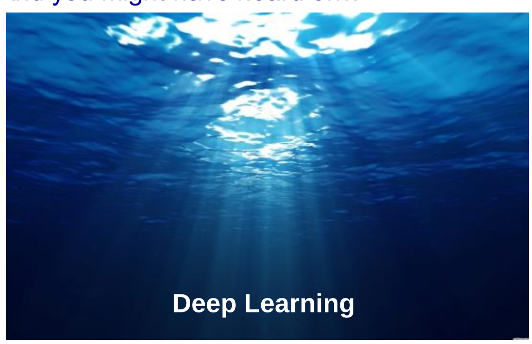
- Image Processing Basics
- Segmentation
- Local Features & Matching
- Object Recognition and Categorization
- 3D Reconstruction



Course Outline

- Image Processing Basics
- Segmentation
- Local Features & Matching
- Object Recognition and Categorization
- 3D Reconstruction

And you might have heard of…



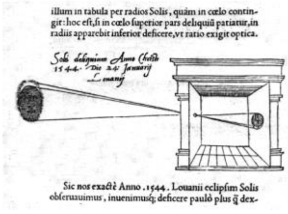
Topics of Today's Lecture

- What is computer vision?
- What does it mean to see and how do we do it?
- How can we make this computational?

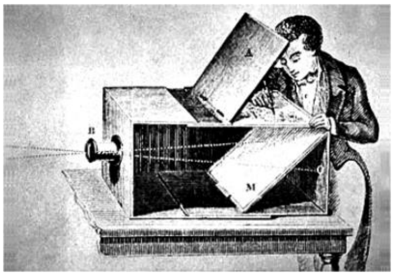
- First Topic: Image Formation
  - Details in Forsyth & Ponce, chapter 1.

Camera Obscura

- Around 1519, Leonardo da Vinci (1452 – 1519)
  - "When images of illuminated objects … penetrate through a small hole into a very dark room … you will see [on the opposite wall] these objects in their proper form and color, reduced in size … in a reversed position owing to the intersection of the rays"



Camera Obscura



- Used by artists (e.g. Vermeer 17th century) and scientists

Camera Obscura

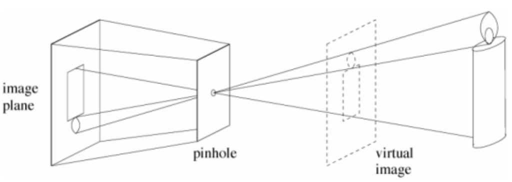
Jetty at Margate England, 1898.

An attraction in the late 19th century

http://brightbytes.com/cosite/collection2.html

Pinhole Camera

- (Simple) standard and abstract model today
  - Box with a small hole in it
  - Works in practice

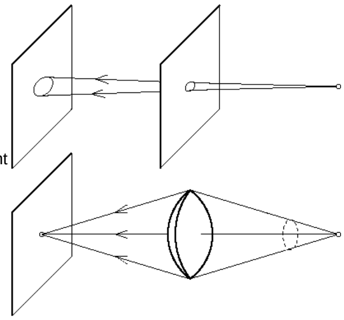


Pinhole Size / Aperture

- Pinhole too big – many directions are averaged, blurring the image
- Pinhole too small – diffraction effects blur the image
- Generally, pinhole cameras are *dark*, because a very small set of rays from a particular point hits the screen.

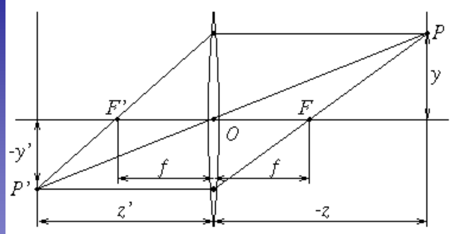
The Reason for Lenses

- Keep the image in sharp focus while gathering light from a large area



Source: Forsyth & Ponce — B. Leibe — 45

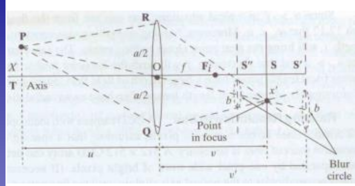
The Thin Lens



$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}$$

Source: Forsyth & Ponce — B. Leibe — 46

Focus and Depth of Field



Thin lens: scene points at distinct depths come in focus at different image planes.

(Real camera lens systems have greater depth of field.)

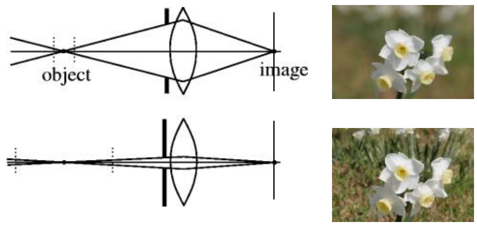
"circles of confusion"

- Depth of field: distance between image planes where blur is tolerable

Source: Shapiro & Stockman — B. Leibe — 47

Focus and Depth of Field

- How does the aperture affect the depth of field?

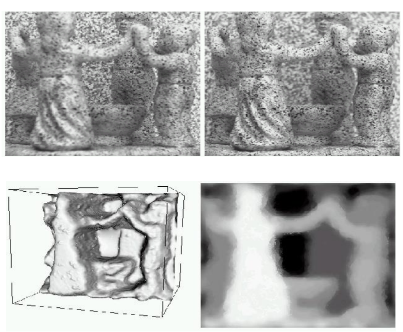


- A smaller aperture increases the range in which the object is approximately in focus

Flower images from Wikipedia http://en.wikipedia.org/wiki/Depth_of_field

Slide from S. Seitz — 48

Application: Depth from (De-)Focus



Images from same point of view, different camera parameters

3D Shape / depth estimates

Slide credit: Kristen Grauman — B. Leibe — [figs from H. Jin and P. Favaro, 2002] — 49

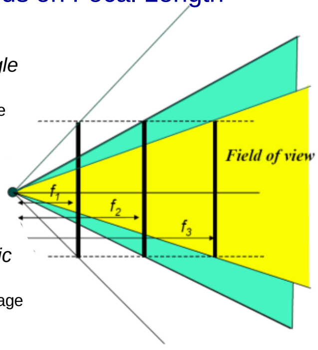
Field of View

- Angular measure of the portion of 3D space seen by the camera

Slide credit: Kristen Grauman — B. Leibe — Images from http://en.wikipedia.org/wiki/Angle_of_view — 50

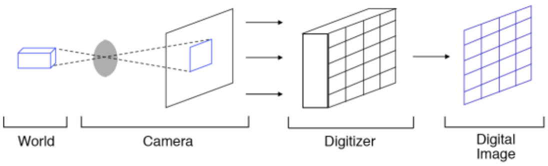
Field of View Depends on Focal Length

- As *f* gets smaller, image becomes more *wide angle*
  - More world points project onto the finite image plane
- As *f* gets larger, image becomes more *telescopic*
  - Smaller part of the world projects onto the finite image plane



from R. Duraiswami

Digital Images



- Film is replaced by a sensor array
- Current technology: arrays of *charge coupled devices* (CCD)
- *Discretize* the image into pixels
- *Quantize* light intensities into pixel values.

Image source: Michael Black

Resolution

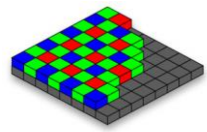
- Sensor: size of real world scene element that images to a single pixel
- Image: number of pixels
- Influences what analysis is feasible, affects best representation choice

Slide credit: Kristen Grauman

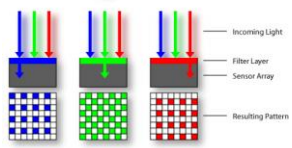
[figs from Efros et al., Mori et al.]

Color Sensing in Digital Cameras

Bayer grid



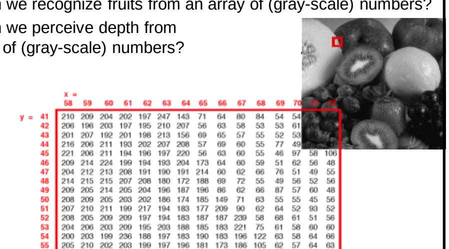
Estimate missing compo -nents from neighboring values (demosaicing)



Source: Steve Seitz

Grayscale Image

- Problem of Computer Vision
  - How can we recognize fruits from an array of (gray-scale) numbers?
  - How can we perceive depth from an array of (gray-scale) numbers?
  - …



  - How do we humans do it? How can we make a computer do it?

Slide credit: Michael Black

Next Lectures

- First few lectures: low-level vision
  - Filtering operations
  - Edge and structure extraction
  - Segmentation and grouping
- Next week: Linear Filters

Computer Vision Summer'19

Questions ?