

Introduction



16-705 Fundamentals of Computer Vision
Spring 2026, Lecture 1

Instructor: Ioannis (Yannis) Gkioulekas

You can call me Yannis.



Originally from Greece



National Technical University of Athens (2004-09)



Harvard University (2009-17)



Carnegie Mellon University (2017-now)

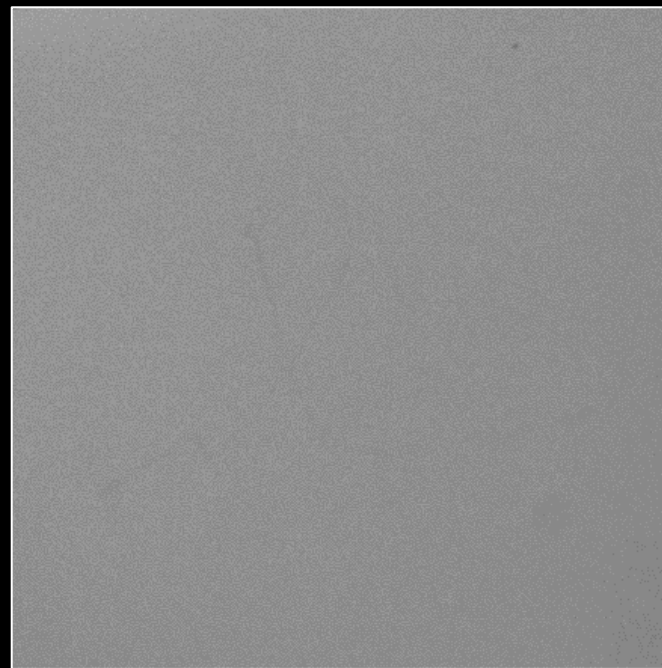
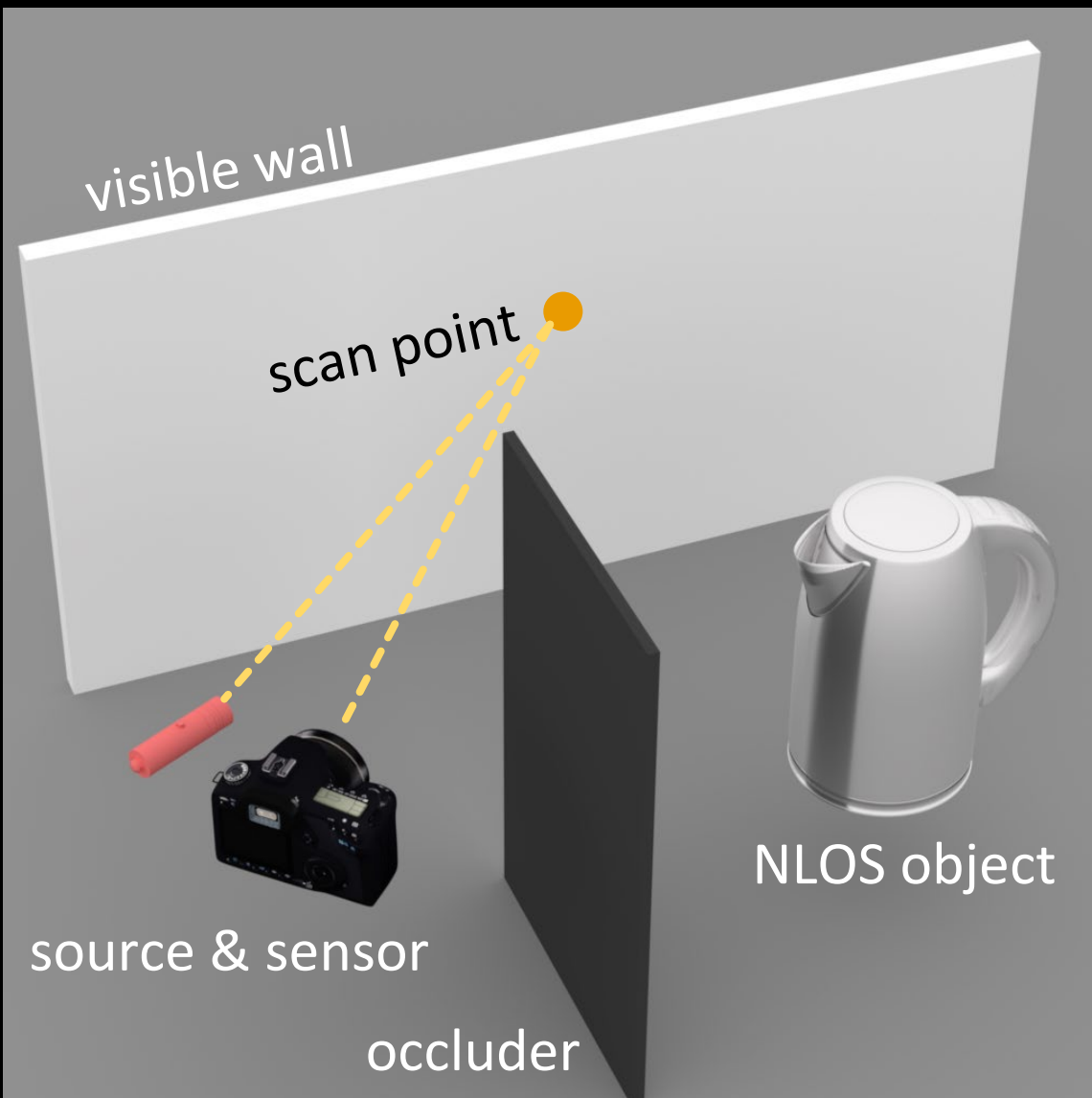


Yannis at Harvard in 2011

My website: <http://www.cs.cmu.edu/~igkioule>

See also: <http://imaging.cs.cmu.edu/>

Looking around corners

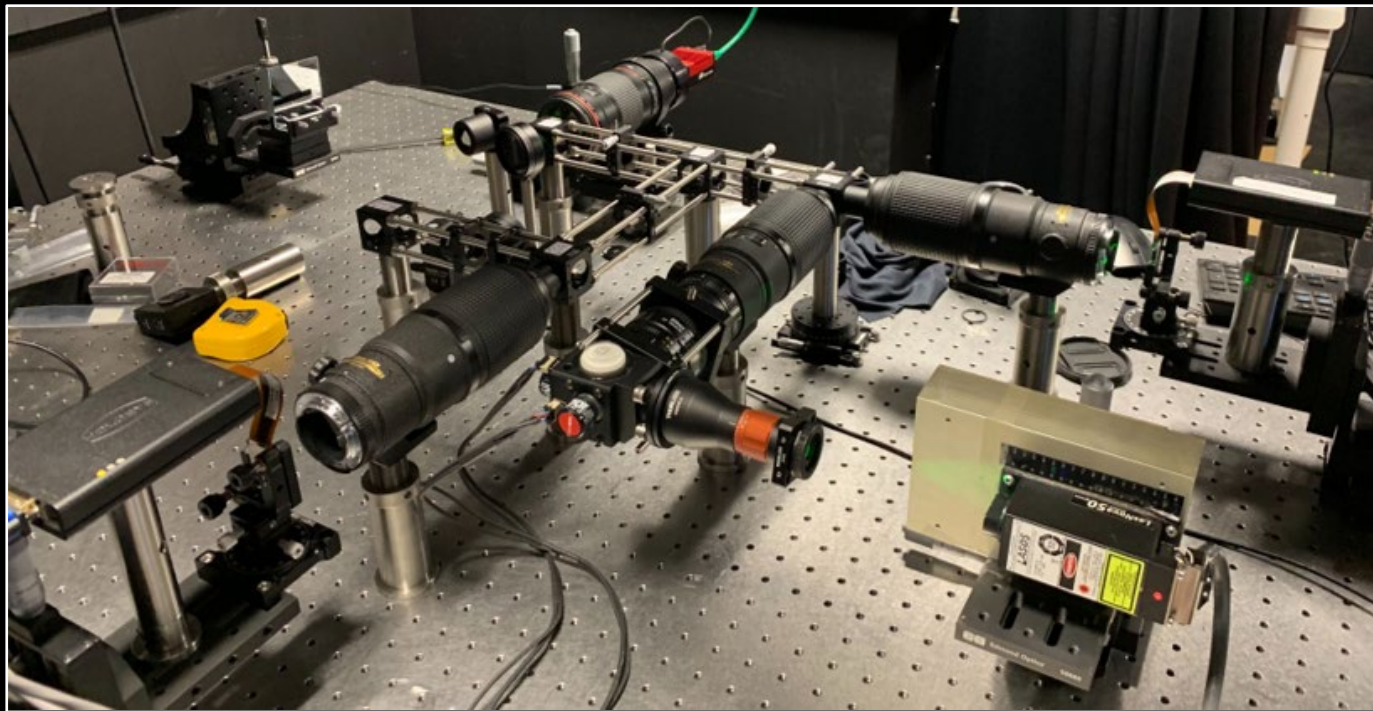


what a regular
camera sees

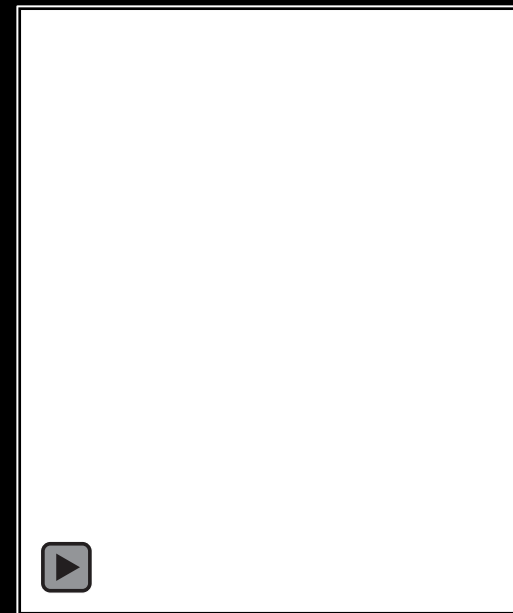


what we can
reconstruct

Seeing light in flight

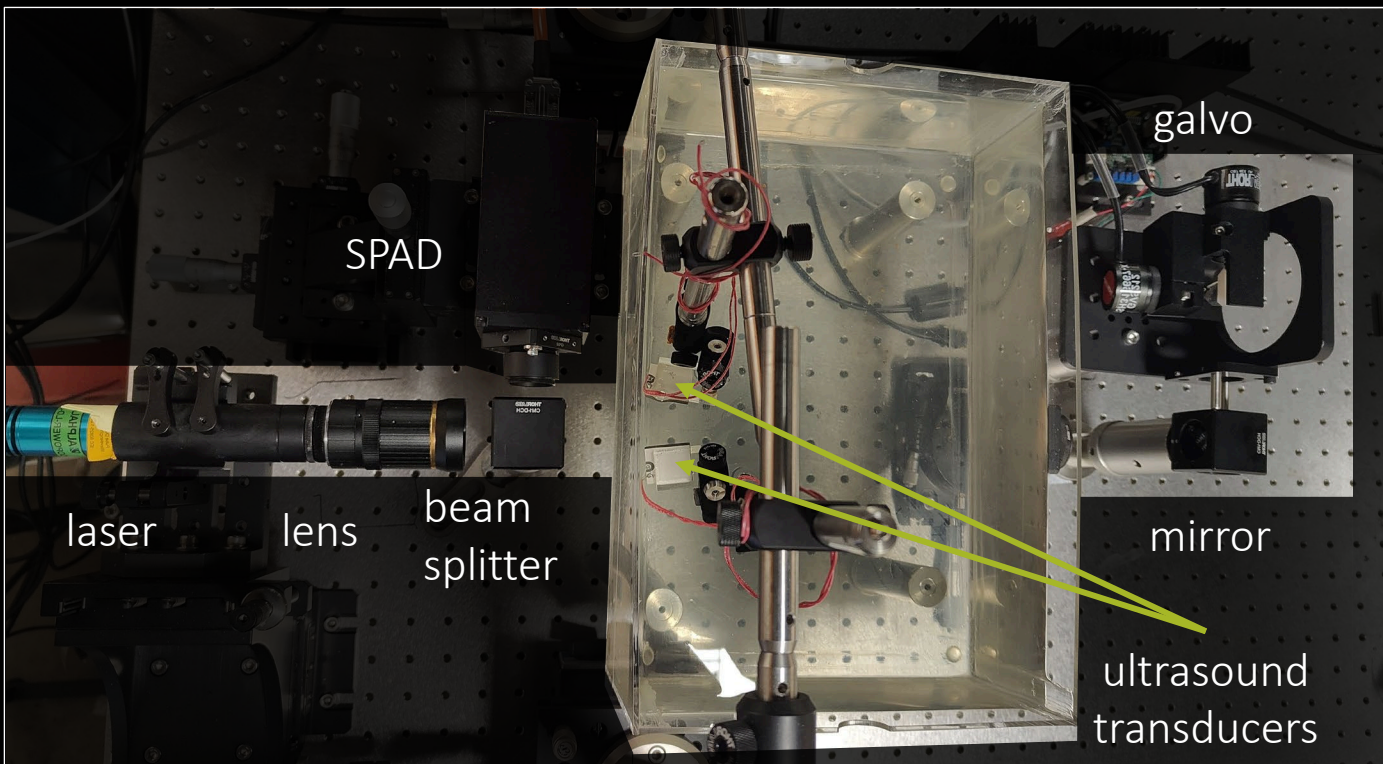


camera for capturing video at 10^{15} frames per second

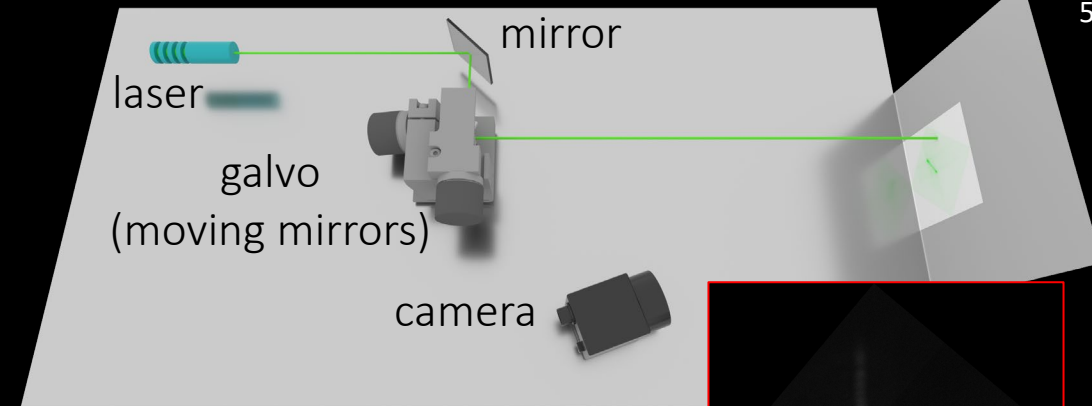


Ultrafast lidar

5

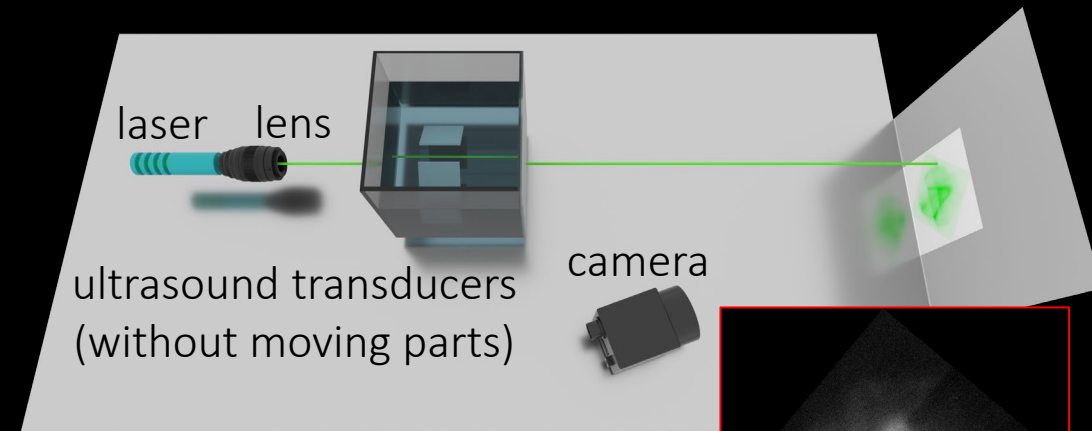
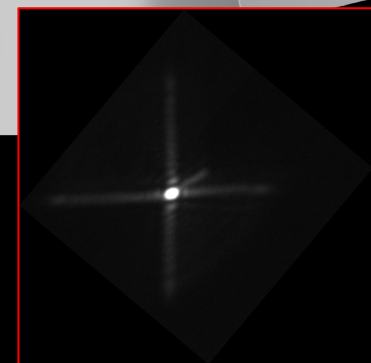


1000 × faster than a
commercial Velodyne lidar



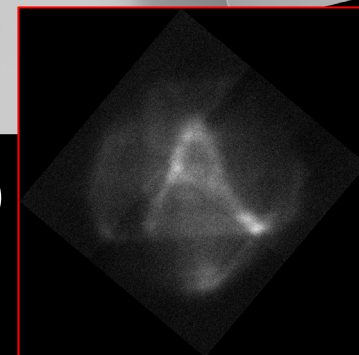
galvo steering (1 kHz)

1 ms exposure



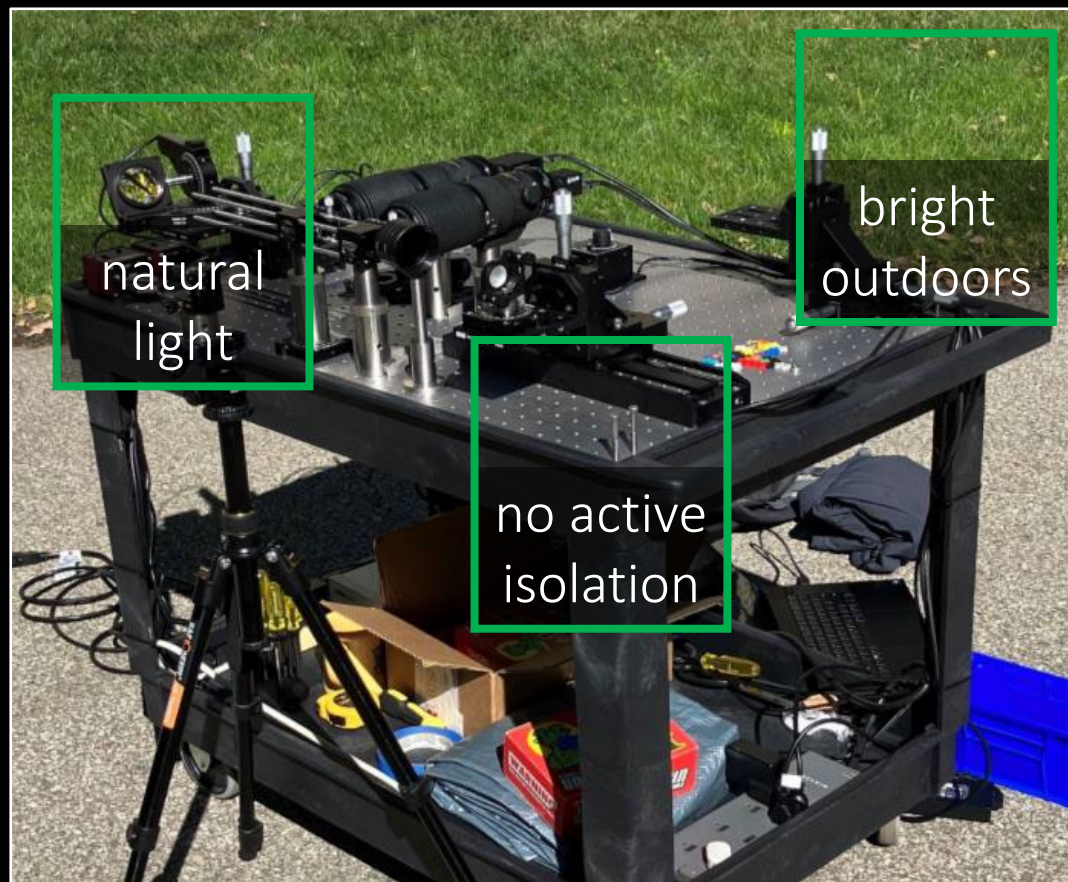
our acousto-optic steering (1 MHz)

1 ms exposure



<http://imaging.cs.cmu.edu/>

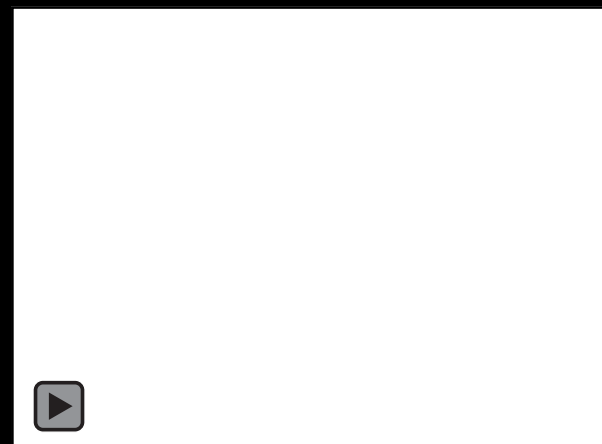
Sunlight lidar



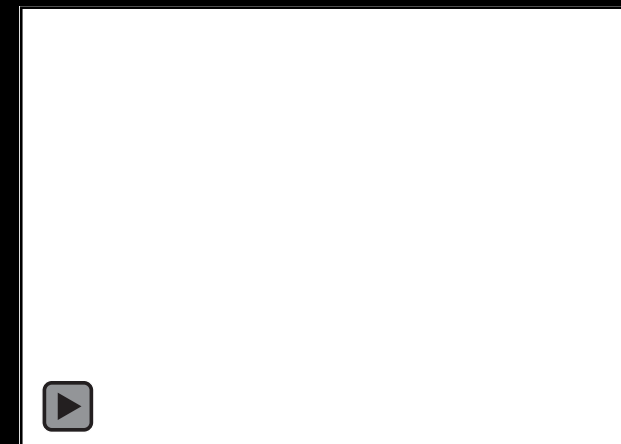
scene



depth map



input images



scene transient

What is
computer vision?



What can you tell me about this image?

www.shutterstock.com · 340880915

<https://www.shutterstock.com/image-photo/berlin-germany-aug-30-2015-flea-340880915>



What objects do you see in the scene?



What type of event is this?



Do you know which people are friends?



Where is the person taking the photo standing?



What does a computer see?



So how are we able to draw out so much information from just one image?

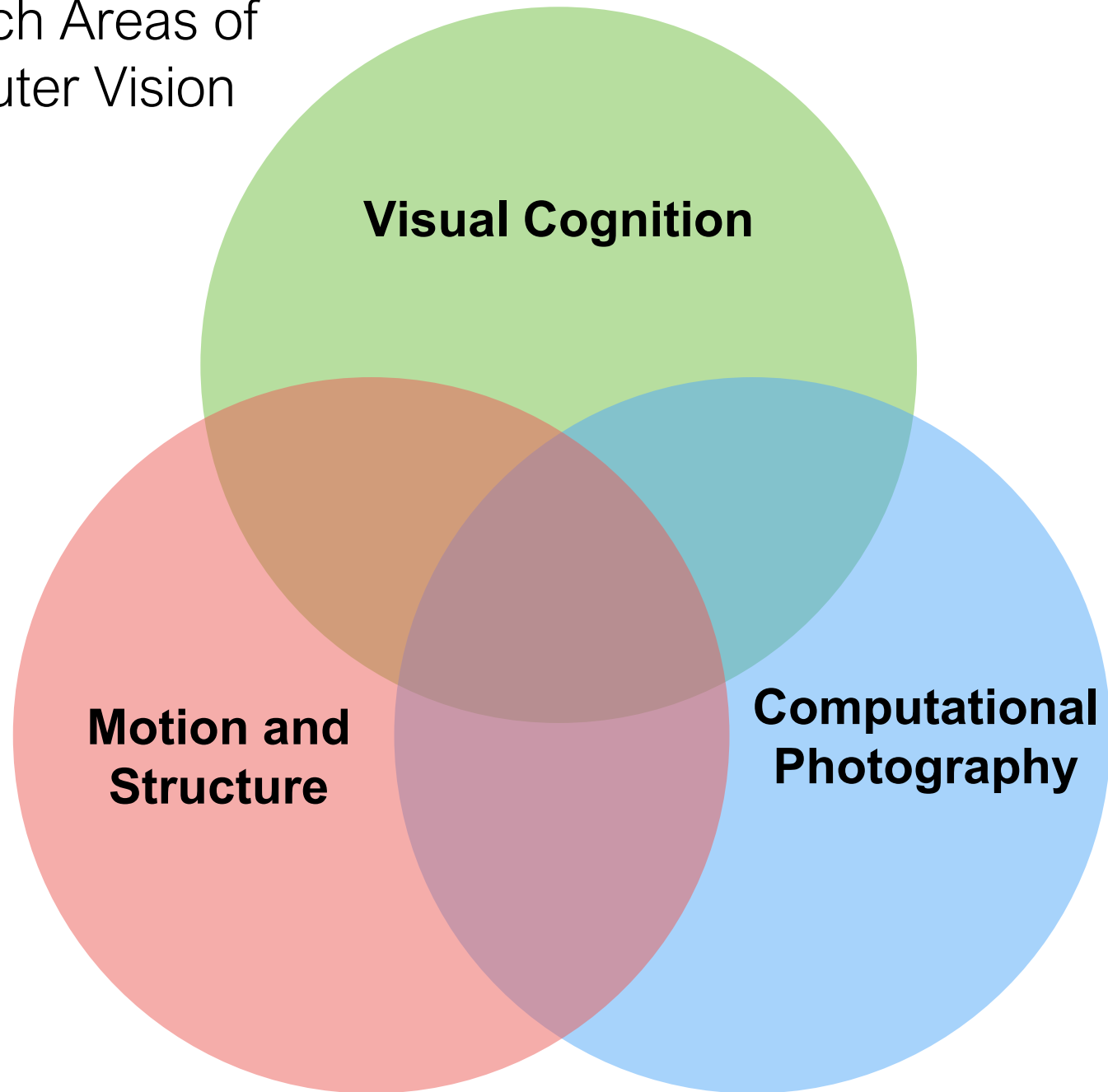


Can we get a computer to understand this image in the same way that we do?

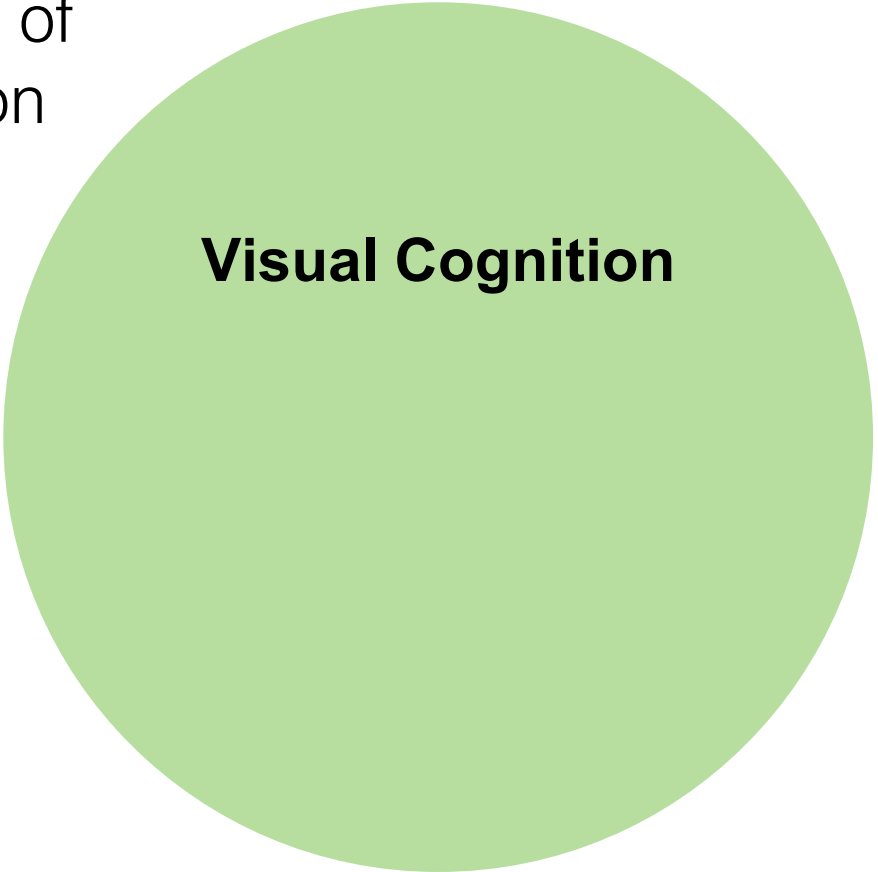
The goal of computer vision is
to give computers
(super) human-level visual perception

Overview of Research Areas in Computer Vision

Research Areas of Computer Vision



Research Areas of Computer Vision

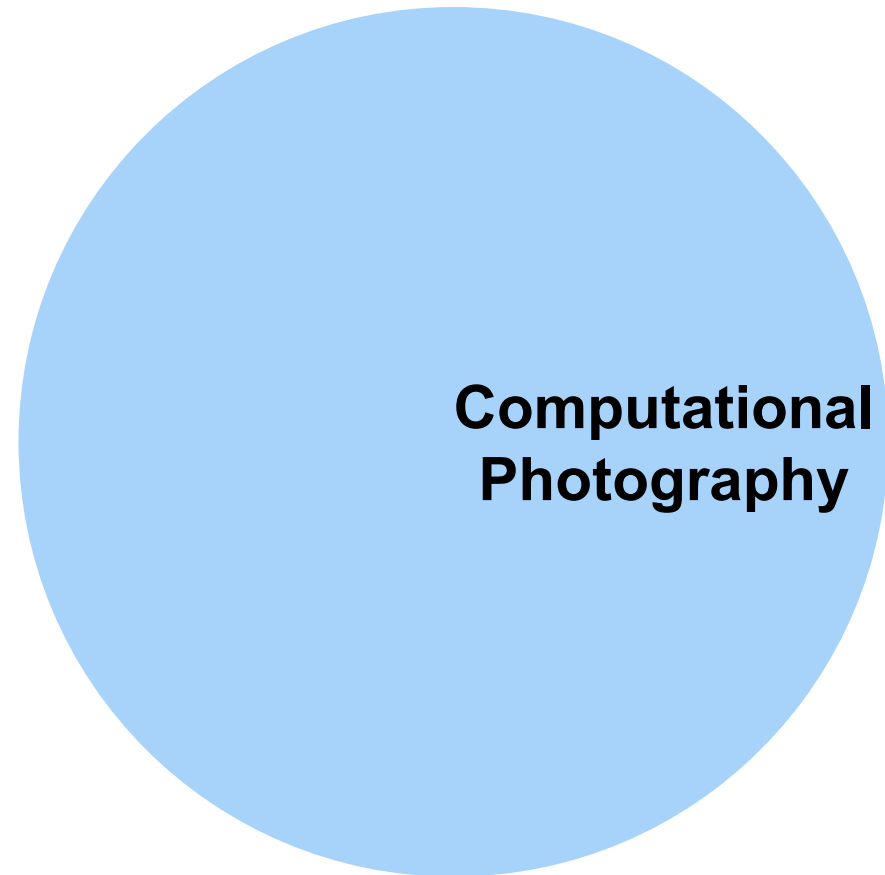


Visual Cognition

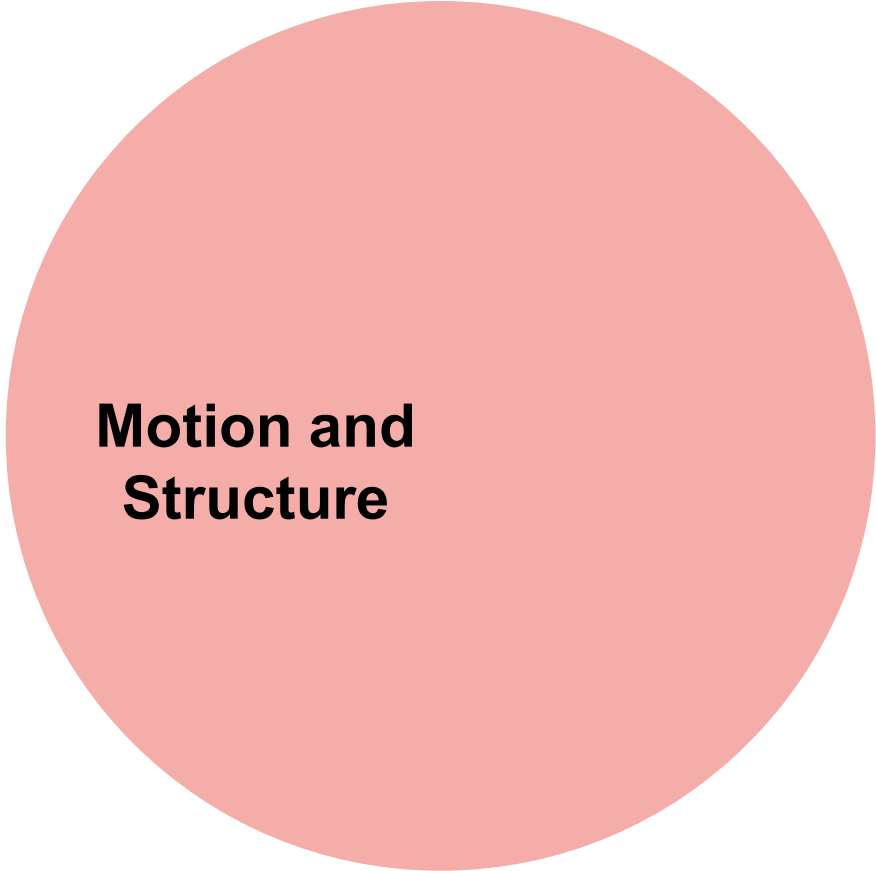
What is in an image?

Research Areas of Computer Vision

*How is an image
created?*

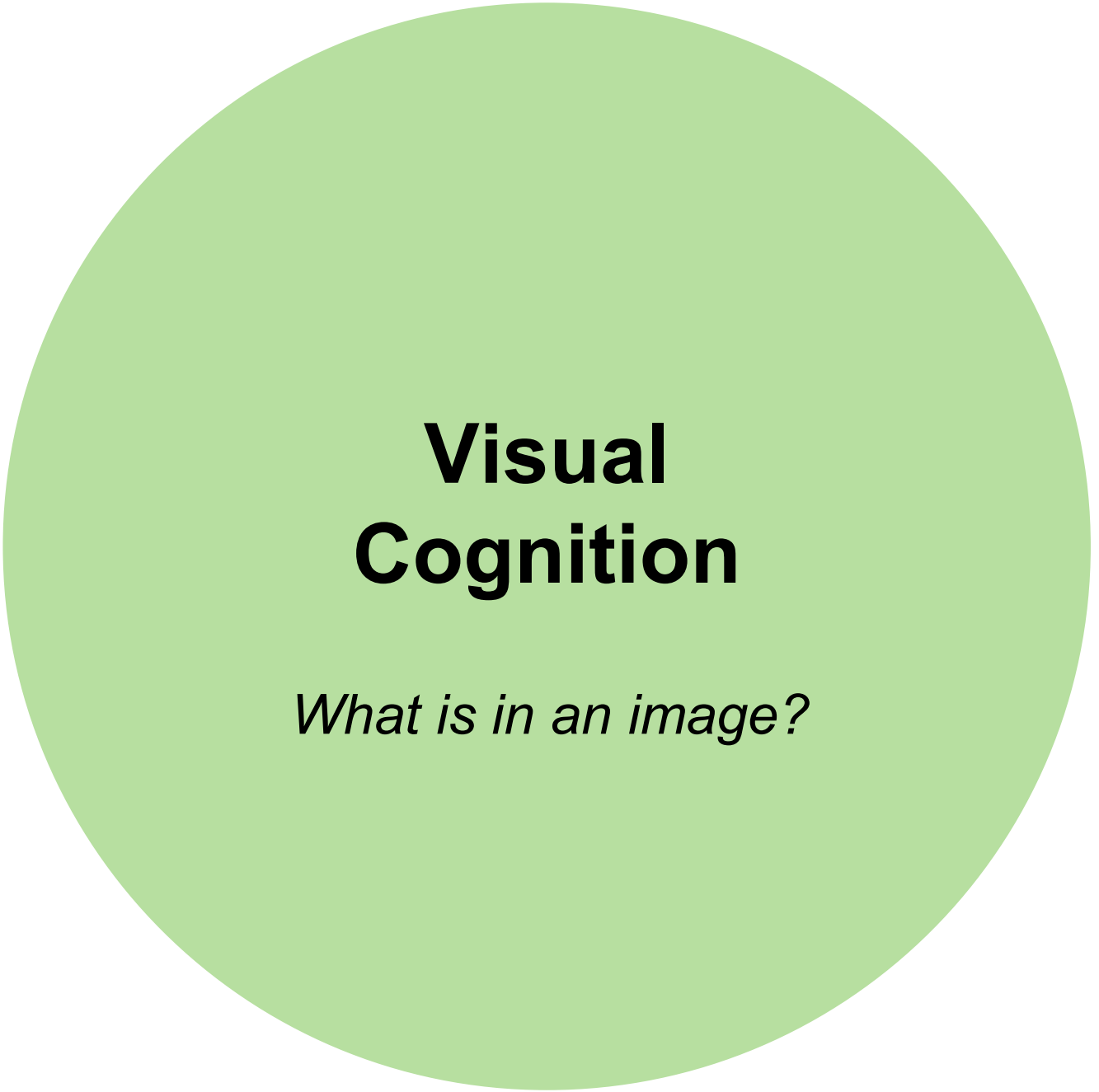


Research Areas of Computer Vision



**Motion and
Structure**

*How does an image
change as a result of
motion?*



Visual Cognition

What is in an image?

Segmentation and detection

Human detection

Object detection

Semantic segmentation

Super-pixels

Activity detection

Recognition

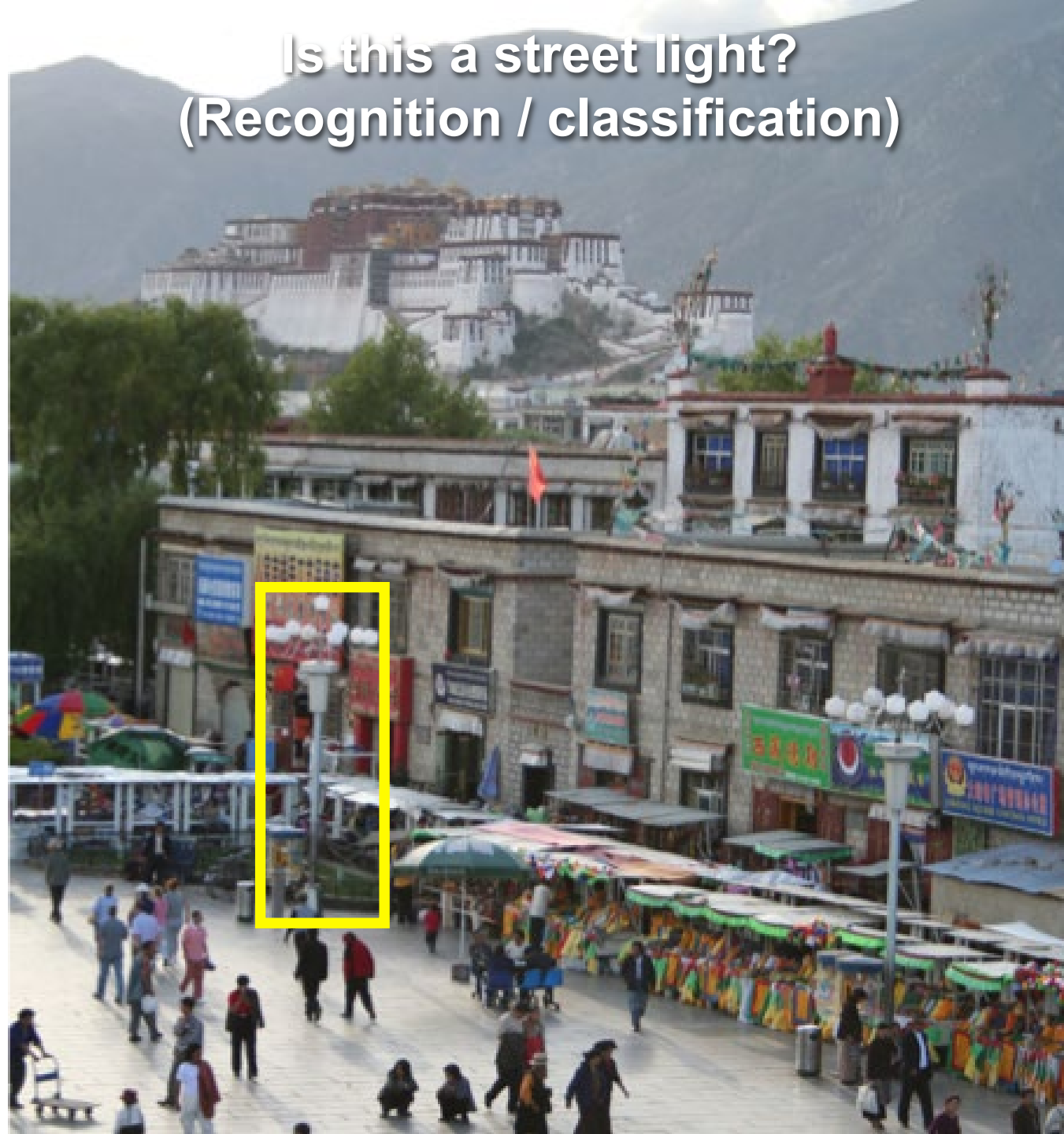
Human identification

Face recognition

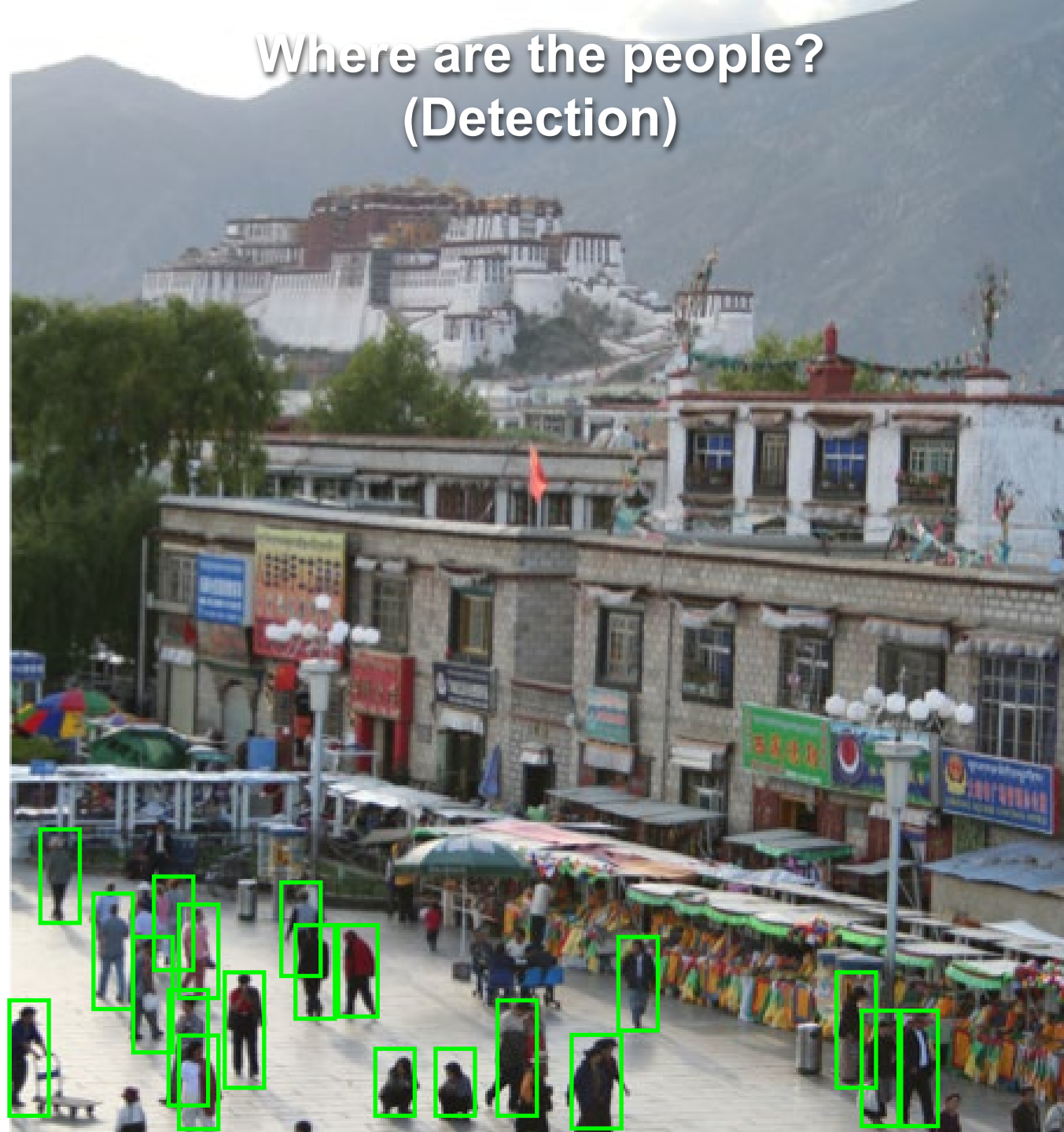
Object recognition

Activity recognition

Is this a street light?
(Recognition / classification)



Where are the people?
(Detection)



Is that Potala palace?
(Identification)



Sky

**What's in the scene?
(semantic segmentation)**

Mountain

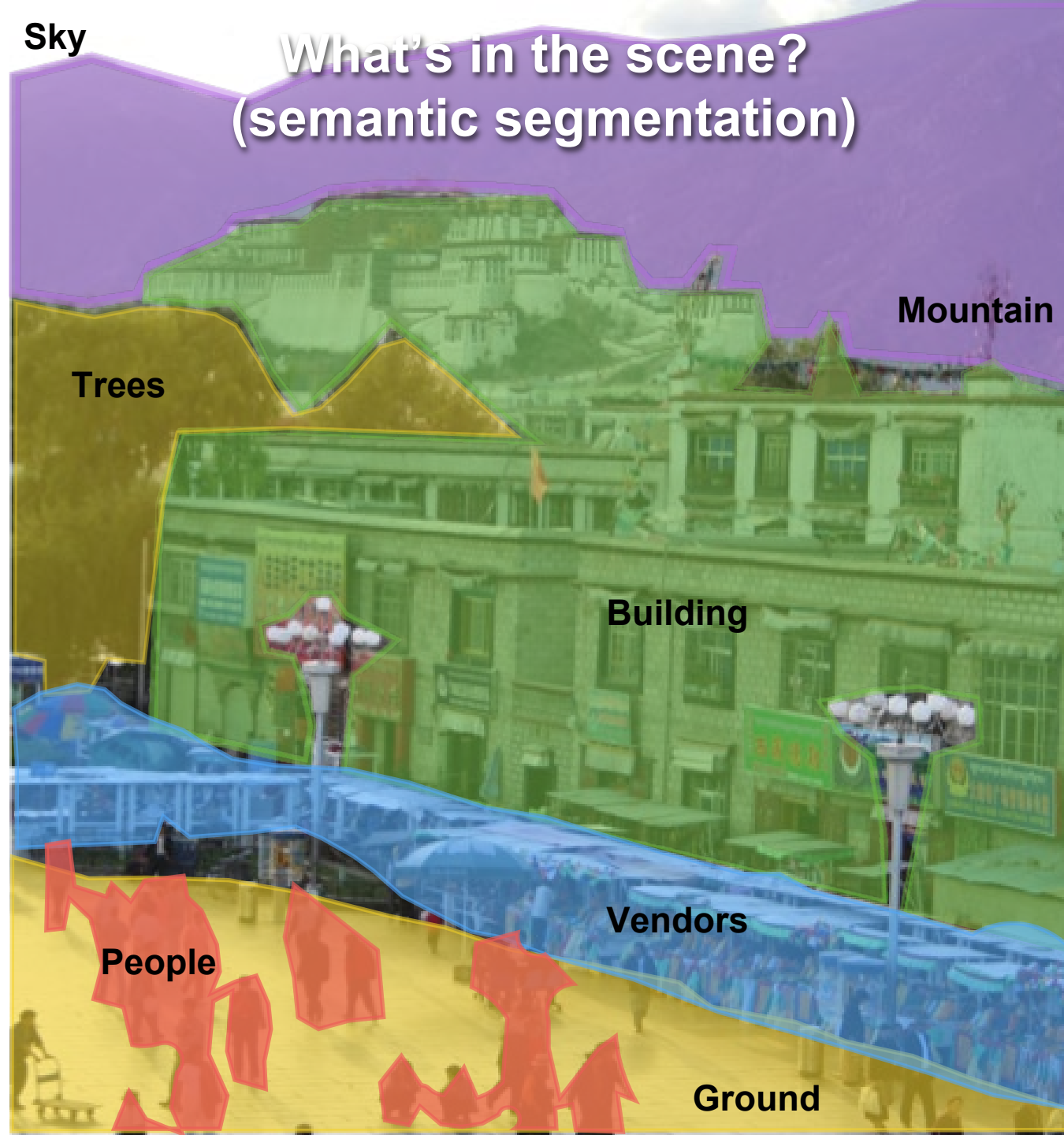
Trees

Building

Vendors

People

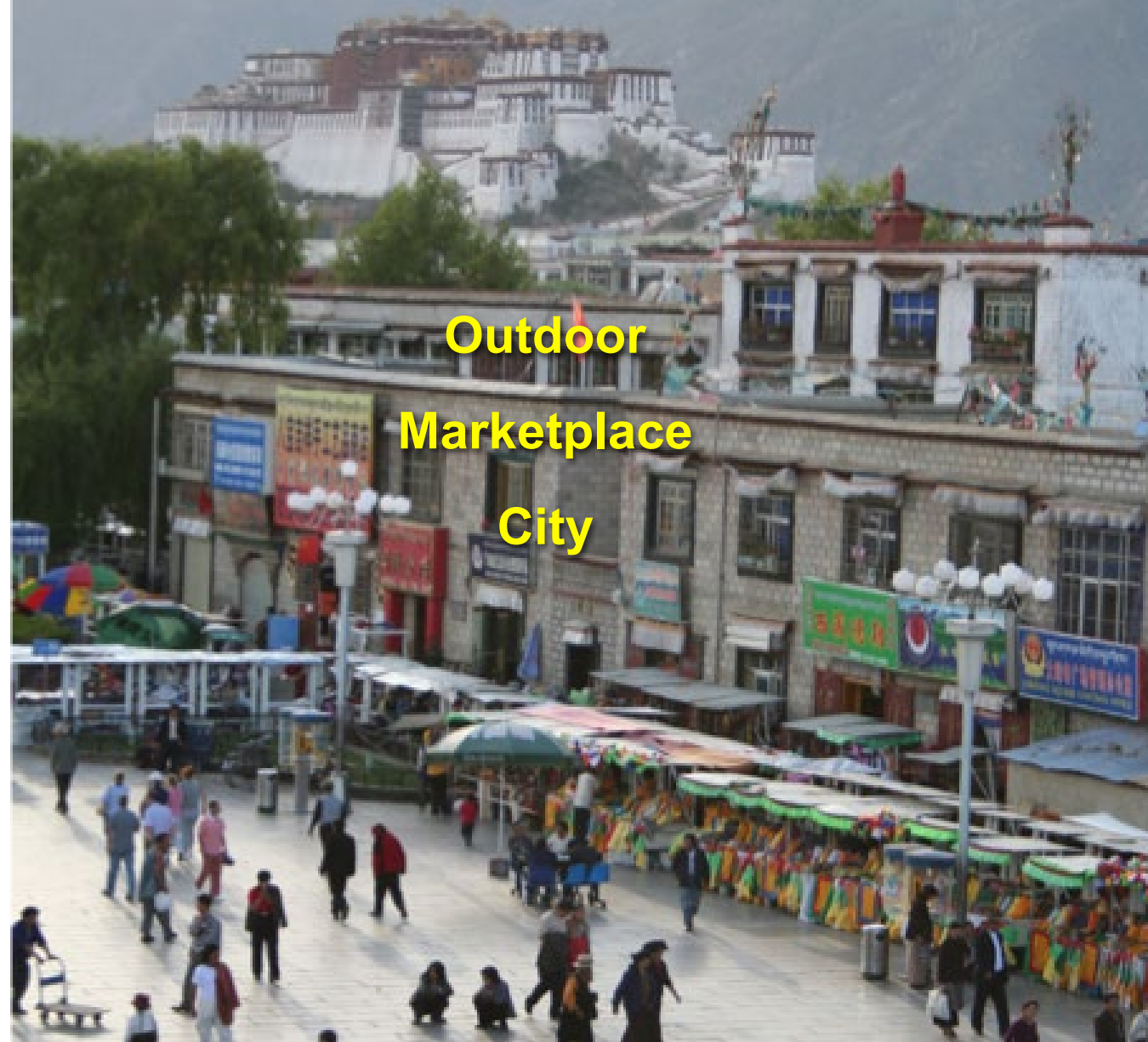
Ground



Object categorization



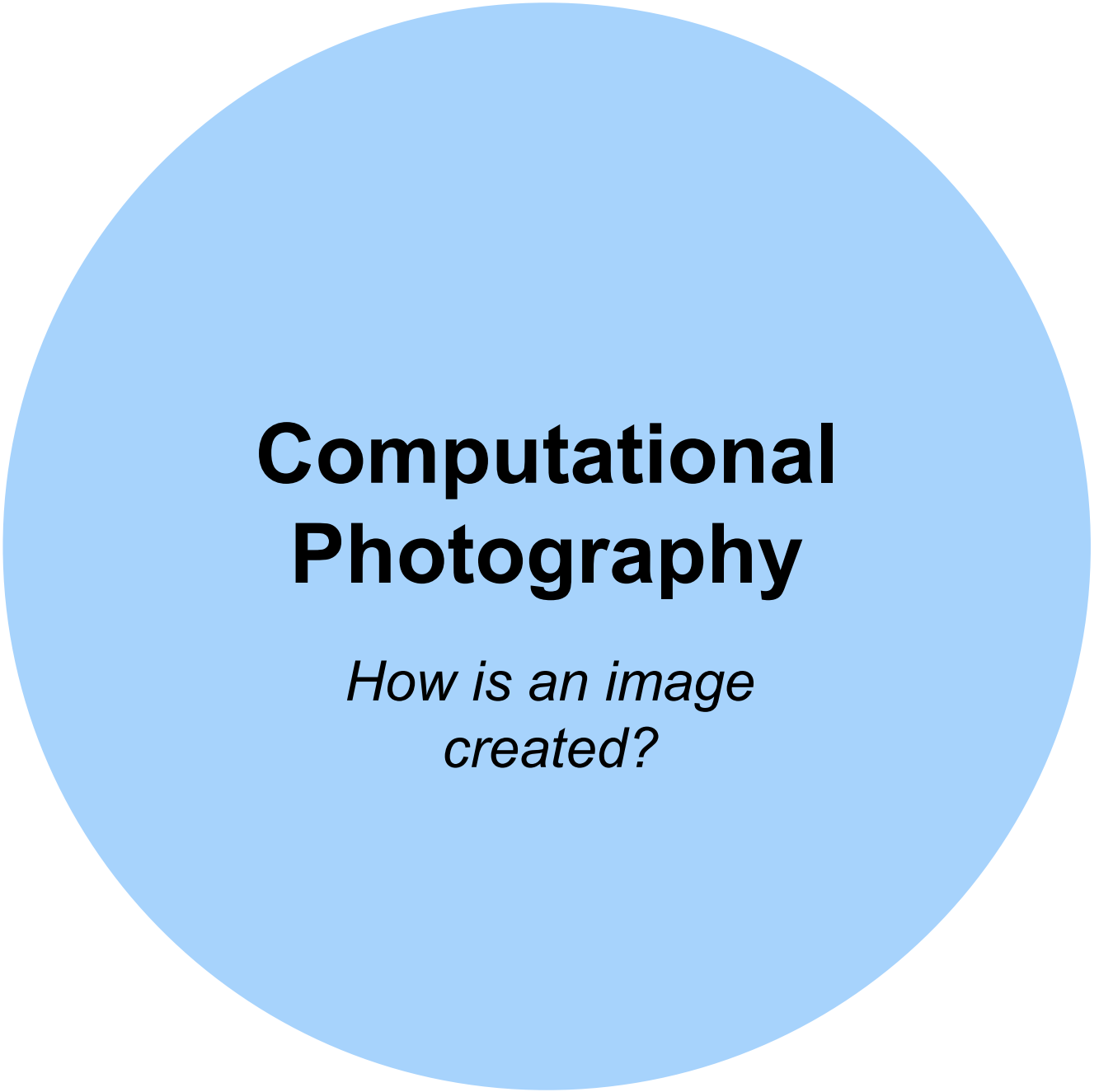
What type of scene is it?
(Scene categorization)



Outdoor
Marketplace
City

Activity / Event Recognition





Computational Photography

*How is an image
created?*

Computational Imaging

Light Field Cameras

Non-Line of Sight Imaging

Structured Light Cameras

Time of Flight Imaging

Computational Illumination and Display

Image and Video Processing

Computational Optics

Imaging Models

Event-based Cameras

High Dynamic Range Imaging

Synthesizing Image

In-painting

Denoising

Image Morphing

Image Restoration

Image Matting and Segmentation

Super-resolution and digital zoom

Colorization

Style Transfer

Image Generation

Image-based Rendering

Physics-based Rendering

Neural Rendering

Synthetic Defocus

Deblurring

Image Fusion

Synthesizing Images

In-painting

Denoising

Image Morphing

Image Restoration

Image Matting

Deblurring

Super-resolution

Colorization

Style Transfer

Image Generation

Image-based Rendering

Physics-based Rendering

Neural Rendering



www.shutterstock.com · 1059553952



www.shutterstock.com · 470889989



www.shutterstock.com · 1060356047



www.shutterstock.com • 1543718633



Computational Imaging

Light Field Cameras

Non-Line of Sight Imaging

Structured Light Cameras

Time of Flight Imaging

Computational Illumination and Display

Image and Video Processing

Computational Optics

Imaging Models

Event-based Cameras

High Dynamic Range Imaging





www.shutterstock.com · 1314016142

Motion and Structure

*How does an image
change when there is
motion?*

Multiple-view Geometry

Pose Estimation

Triangulation

Depth Estimation

Stereo correspondences

Camera calibration

3D reconstruction

Simultaneous Localization
and Mapping (SLAM)

Visual Odometry

Monocular Depth Estimation

Surface normal estimation

State Estimation

Pedestrian Tracking

Face Tracking

Hand Tracking

Multi-object Tracking

Registration

Optical Flow

Image Alignment

Active Appearance Models

Image Registration

Non-rigid Motion Estimation

Motion Modeling

Human Pose and Motion Estimation

3D hand Pose Estimation

Trajectory analysis

Multi-agent trajectory forecasting



www.shutterstock.com · 1793697937



www.shutterstock.com • 613624070



Multiple-view Geometry

Pose Estimation

Triangulation

Depth Estimation

Stereo correspondences

Camera calibration

3D reconstruction

Simultaneous Localization
and Mapping (SLAM)

Visual Odometry

Monocular Depth Estimation

State Estimation

Pedestrian Tracking

Face Tracking

Hand Tracking

Multi-object Tracking

Registration

Optical Flow

Image Alignment

Image Registration

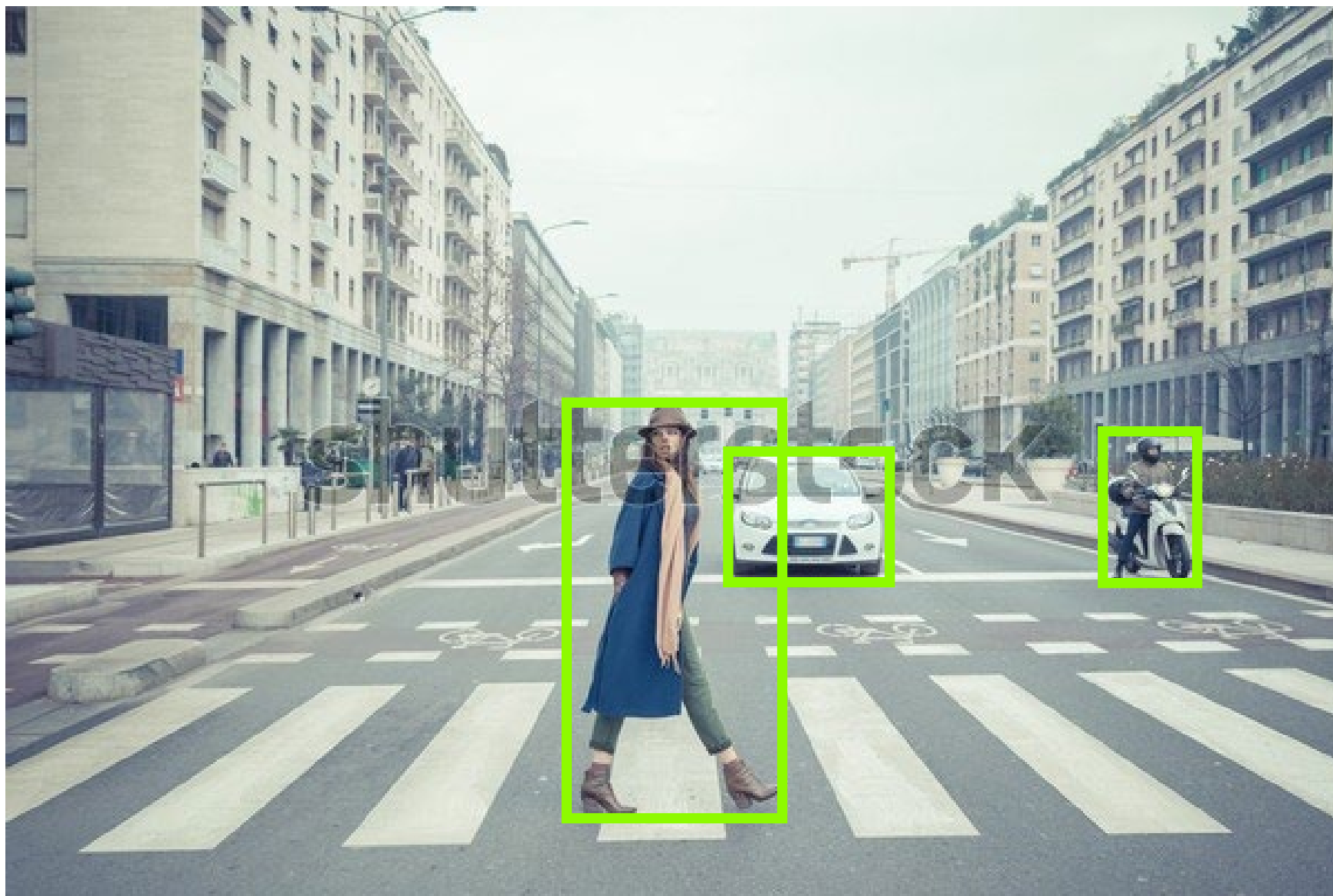
Non-rigid Motion Estimation

Motion Modeling

Human Pose and Motion Estimation

Trajectory analysis

Multi-agent trajectory forecasting

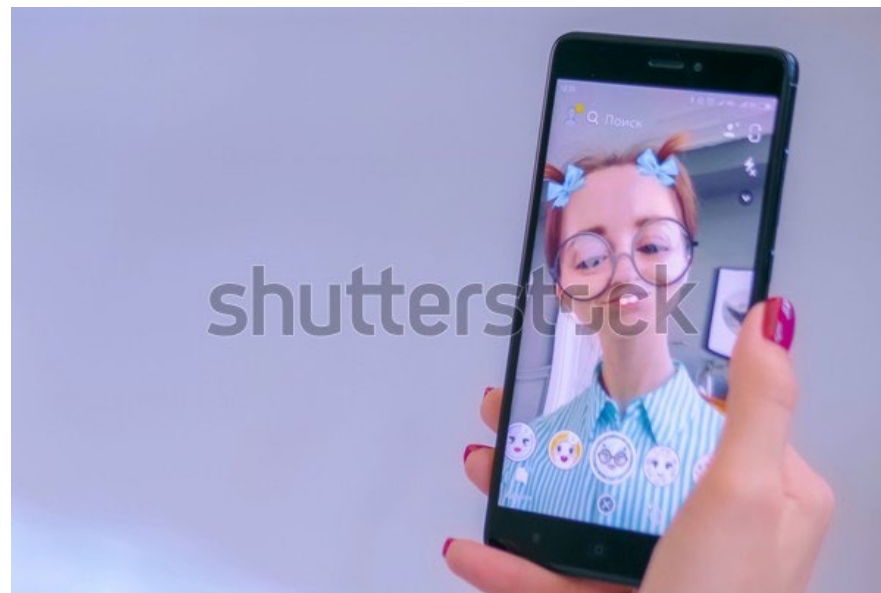




www.shutterstock.com · 445337020

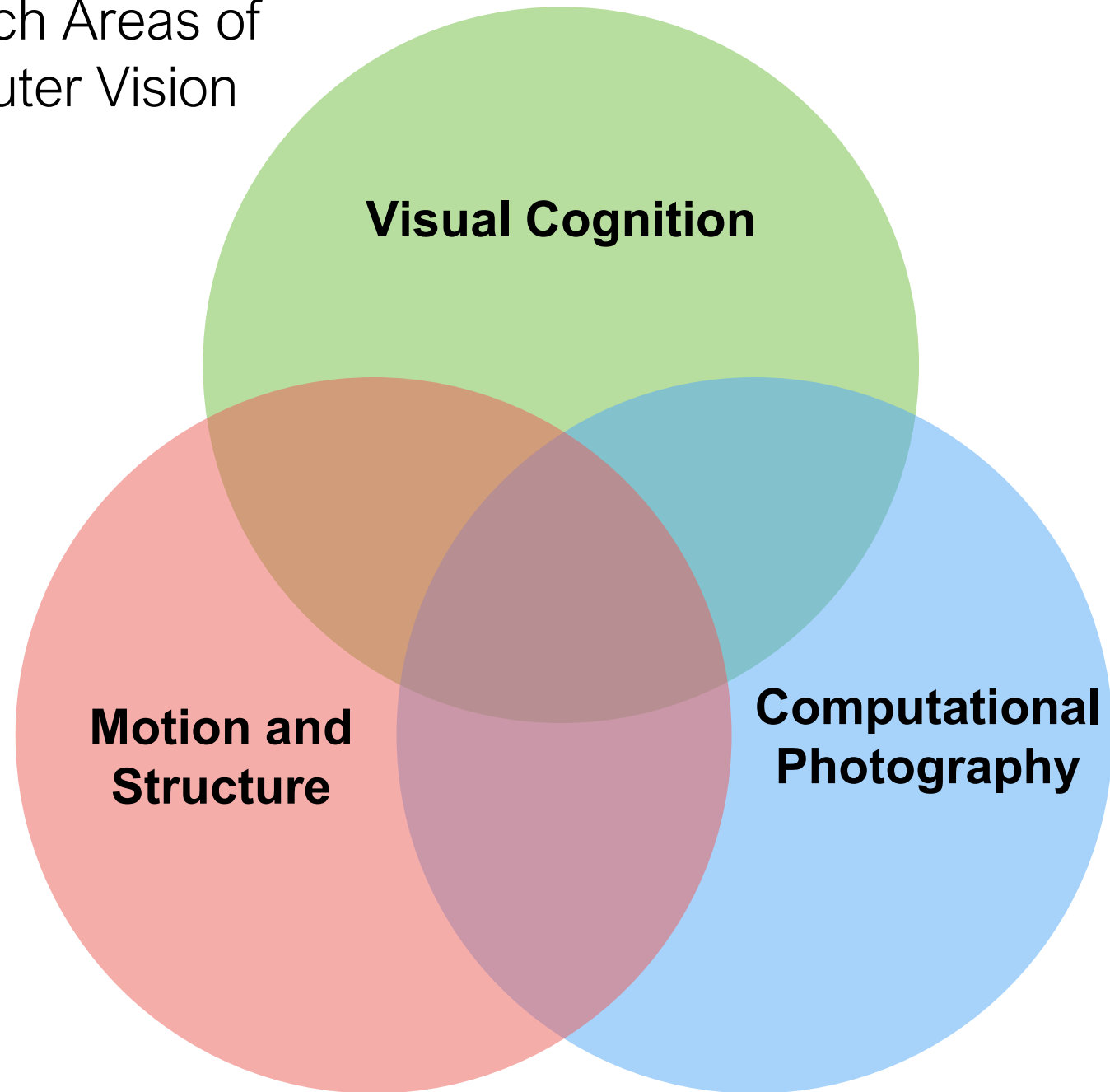


www.shutterstock.com · 1620865849



www.shutterstock.com · 1414050719

Research Areas of Computer Vision



Course logistics

Website



<https://cmu-ci-lab.github.io/16705/>

(includes links to Canvas and Piazza)

Assignments

Canvas

<https://canvas.cmu.edu/courses/52669/>

Discussion & notes



(you will receive an invite from me by tonight)

Topics to be covered

Image processing pipeline

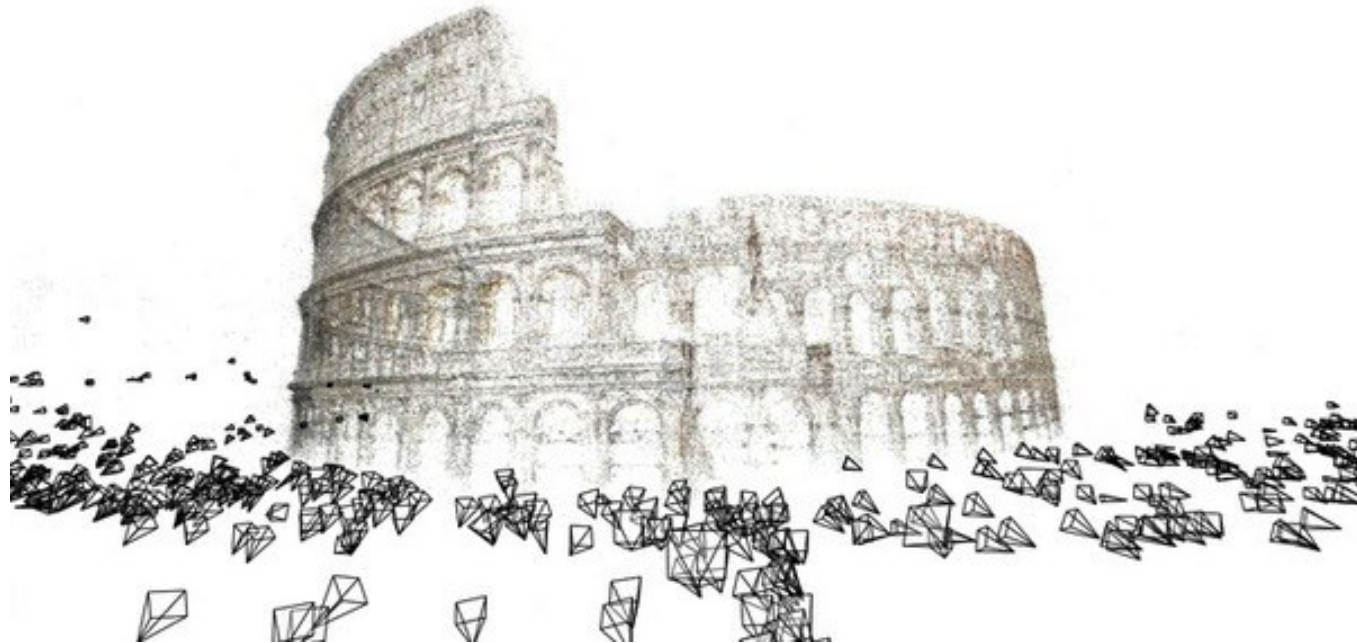
- How images are formed.
- Basics of imaging sensors.
- Standard post-capture processing.
- Color and noise.



Topics to be covered

Geometry-based vision:

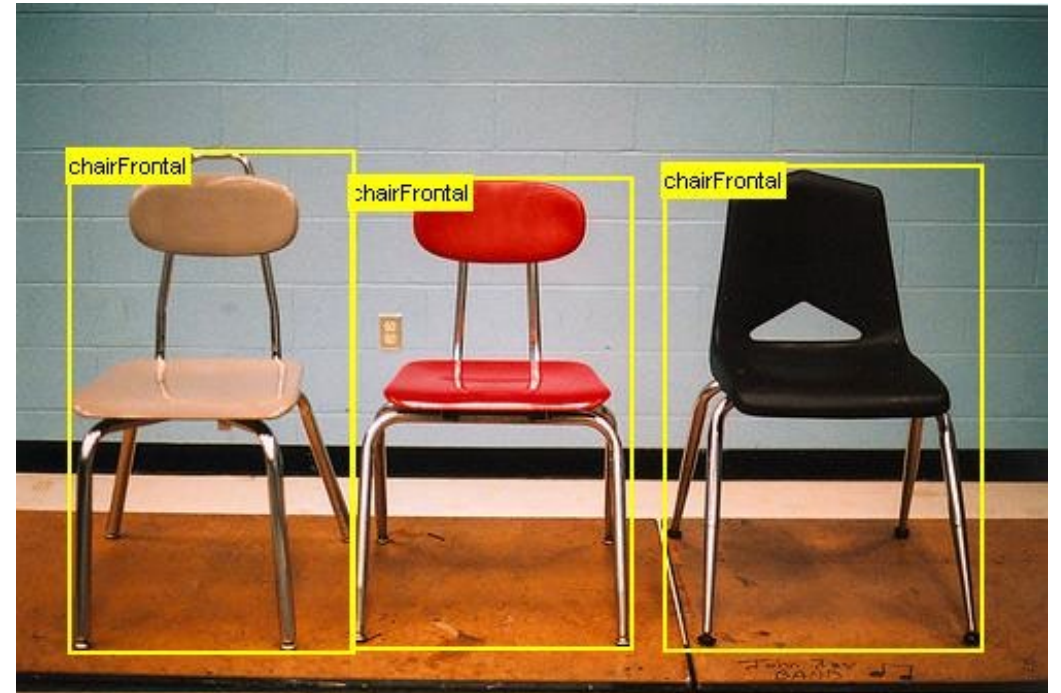
- Camera models.
- Fundamental matrix.
- Epipolar geometry and stereo.
- Structure from motion.



Topics to be covered

Learning-based vision:

- Convolutional neural networks.
- Training, backpropagation.
- Classification, segmentation.
- Modern architectures.



Topics to be covered

Dealing with motion:

- Optical flow.
- Image registration.
- Tracking.
- Applications in robotics.



Lecture etiquette

- Lecture slides (PDF) are posted on the course website before each lecture.
- Lectures are **not** recorded.
- You are expected to attend lectures in person.

Grading

- Two two-week programming assignments: 100%

Assignment 1 Structure from motion
Assignment 2 Detection and tracking

- A lot of programming in Python.
- Managed via Gradescope (will be available through Canvas).

Leniency

Deadlines are enforced strictly to the minute.

- Exceptions only for medical or serious family reasons.

Late days for programming assignments:

- 10% reduction of points per late day.
- No more than 3 late days per assignment.

Schedule

(Tentative) syllabus

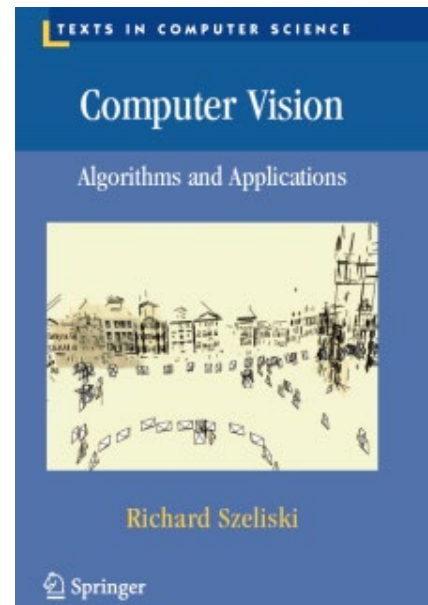
Dates and topics are likely to change during the semester. Slides will be uploaded on this website before each lecture.

- Tentative schedule on course website.
- Likely to change.
- Always check course website and Slack for updates!

Date	Topics	Homework
Mon, Mar 9	Introduction	
Wed, Mar 11	Image processing pipeline	
Mon, Mar 16	Camera models	
Wed, Mar 18	Two-view Stereo	
Mon, Mar 23	Stereo	
Wed, Mar 25	Structure from motion	
Fri, Mar 27		PA1 out
Mon, Mar 30	Classification	
Wed, Apr 1	Segmentation	
Mon, Apr 6	Neural networks	
Wed, Apr 8	Modern architectures	
Fri, Apr 10		PA1 due, PA2 out
Mon, Apr 13	Detection	
Wed, Apr 15	Tracking	
Mon, Apr 20	Generative models	
Mon, Apr 22	Wrap up	
Fri, Apr 24		PA2 due

Book

Mostly readings from this book.



PDF online

<http://szeliski.org/Book/>

Prerequisites

We assume familiarity with calculus, linear algebra, basic probability, and programming.

Contact and office hours

- Feel free to email us about *administrative questions*.
 - please use [16705] in email title!
- Technical questions should be asked on Slack.
 - we won't answer technical questions through email.
 - you can post anonymously if you prefer.
- Office hours will be determined next week.