

Computer Vision

Rao Muhammad Umer
Lecturer,

CS and IT Department,
The University of Lahore.

Web: raoumer.com



My Background

- **Name: Rao Muhammad Umer**
- **Teaching Experience**
 - Oct. 2016 up-till now
 - DSA, PF-I, DLD, ITC,...
- **MS (Computer Science)**
 - PIEAS (2014-2016)
 - PIEAS Fellowship
 - Area of Research: Data Science and Machine Learning
- **BS (Computer System Engineering)**
 - UCET, IUB (2010-2014)
 - National ICT Scholarship
 - Area of Research: AI, Parallel & Distributed Computing, and Digital Image Processing
- Visit my personal website for more information about me on following link:

raoumer.com



Prerequisites

- This course requires knowledge of **Data Structures & Algorithms**, **Calculus**, **Linear Algebra**, **Analytical Geometry**, **Probability & Statistics** as well as decent **programming skills**.
- We will leverage concepts from **low-level image processing** (e.g., linear filters, edge detectors, corner detectors, etc...) and **machine learning** (e.g., SVM, clustering, neural networks, etc...).
- I will provide tutorials to background material related to image processing and machine learning, so students can refresh or study those topics if needed.



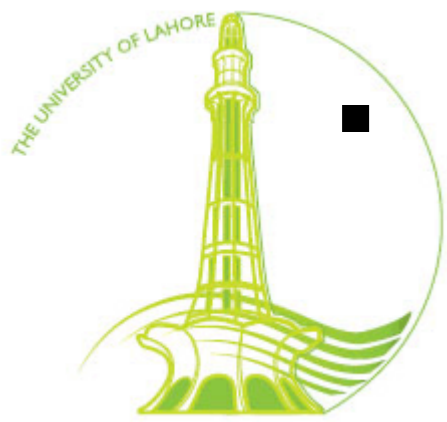
Text Books

- **Required:**

- Szeliski R., Computer Vision: Algorithms and Applications, Springer, 2010.
- Forsyth D. A. and Ponce J., Computer Vision: A Modern Approach, Prentice Hall, 2011.

- **Recommended:**

- Goodfellow, Bengio, and Courville , Deep Learning.
- Fisher et al , Dictionary of Computer Vision and Image Processing.
- Hartley R, and Zisserman A., Multiple View Geometry in Computer Vision, 2nd Ed., Cambridge University Press, 2004.
- Trucco and Verri A., Introductory Techniques for 3-D Computer Vision, Prentice Hall, 1998.
- R. C. Gonzalez & R. E. Woods, Digital Image Processing.



Course Assignments

- **4 assignments** (programming based) + **1 assignment** (theoretical)
- **1 project**



Course Projects

- **Course Projects**
 - Replicate an interesting paper or real time project
 - Comparing different methods to a test bed
 - A new approach to an existing problem or comparison between existing approaches
 - Write a atleast 5-page technical report summarizing your results
 - Release the final code (via Github repository)
 - Give a final in-class presentation
 - We will introduce projects in 2-3 weeks



Course Projects

- **Form your team:**
 - 1-3 people
 - The larger is the team, the more work we expect from the team
 - Be nice to your partner: do you plan to drop the course? Please, take this decision early.
- **Evaluation**
 - Quality of the project (including writing plus code)
 - Final project in-class presentation



Grading Policy

No.	Assessment	Percentage
1.	Final Exam	50%
2.	Mid Exam	30%
3.	Assignments+ Quiz	10%
5.	Project	10%
Total		100%





Illustration by **Arthur Rackham**

“There was a table set out under a tree in front of the house, and the March Hare and the Hatter were having tea at it.”

“The table was a large one, but the three were all crowded together at one corner of it ...”

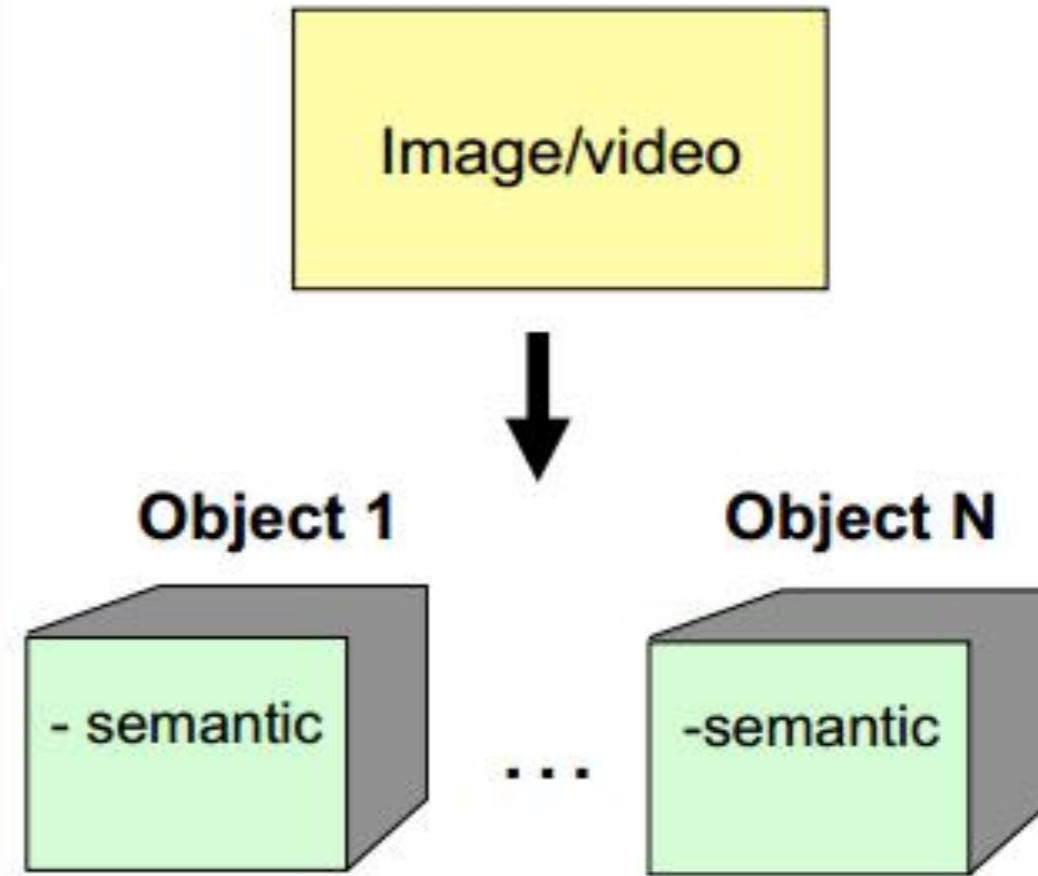
From “**A Mad Tea-Party**”
Alice's Adventures in Wonderland
by
Lewis Carroll

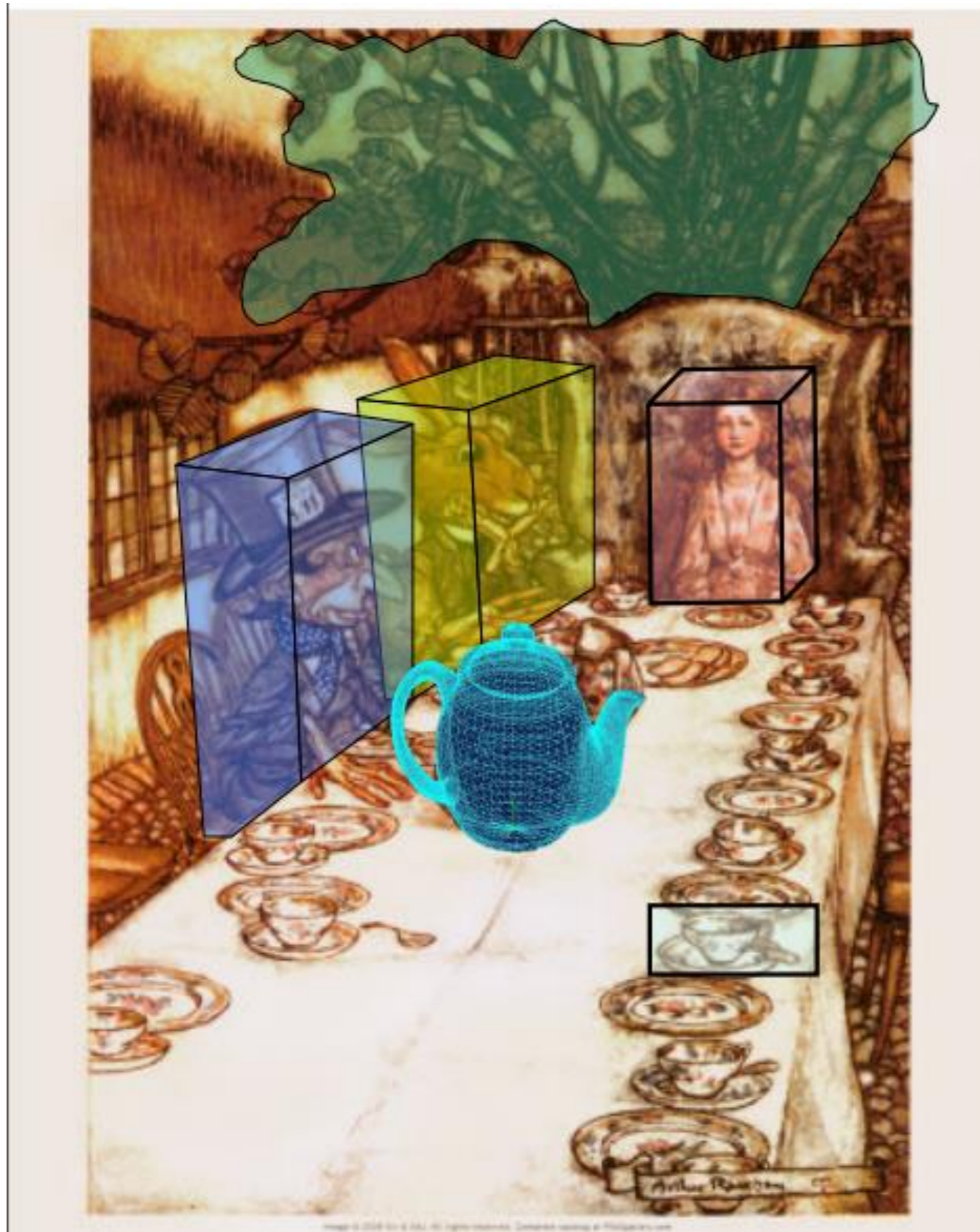




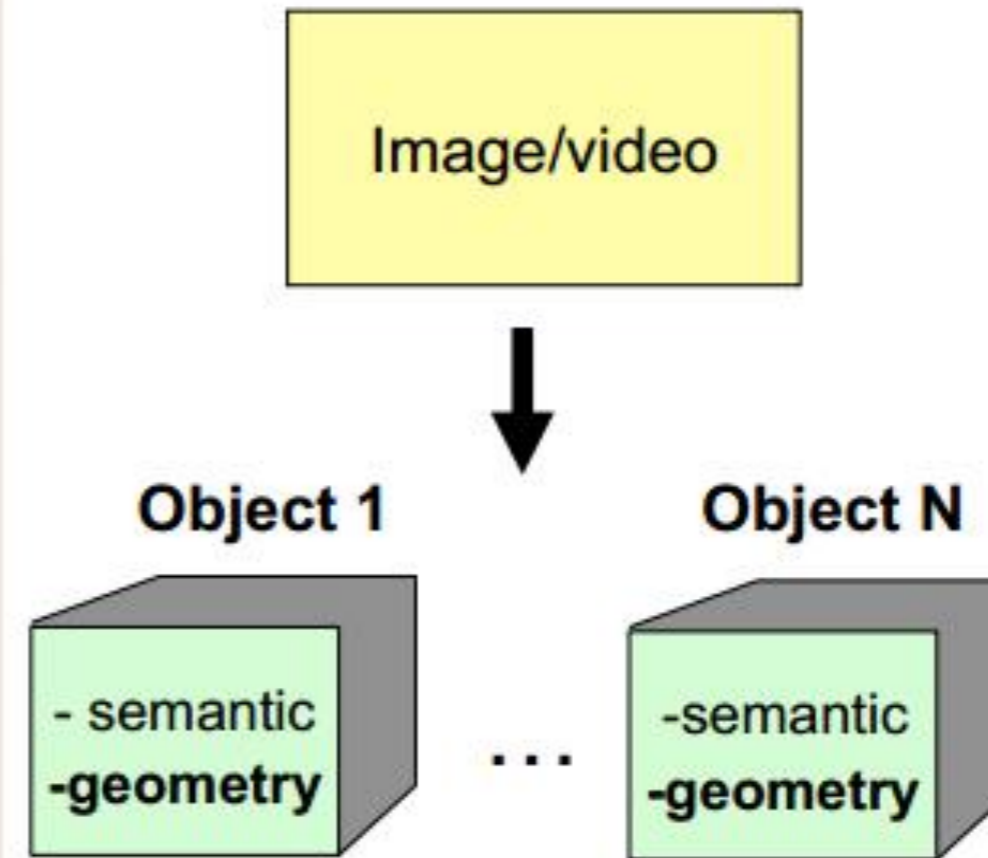
Computer Vision

Computer vision



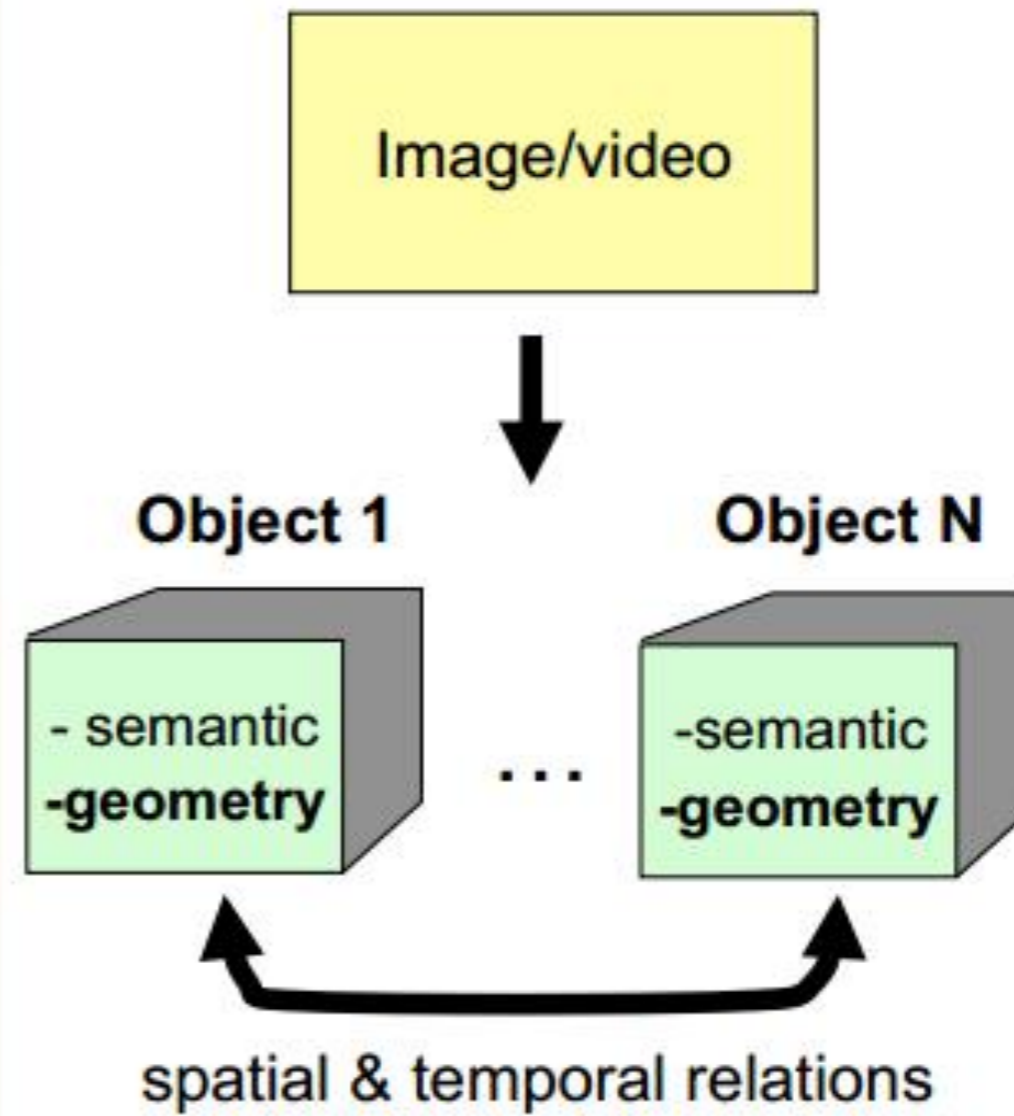


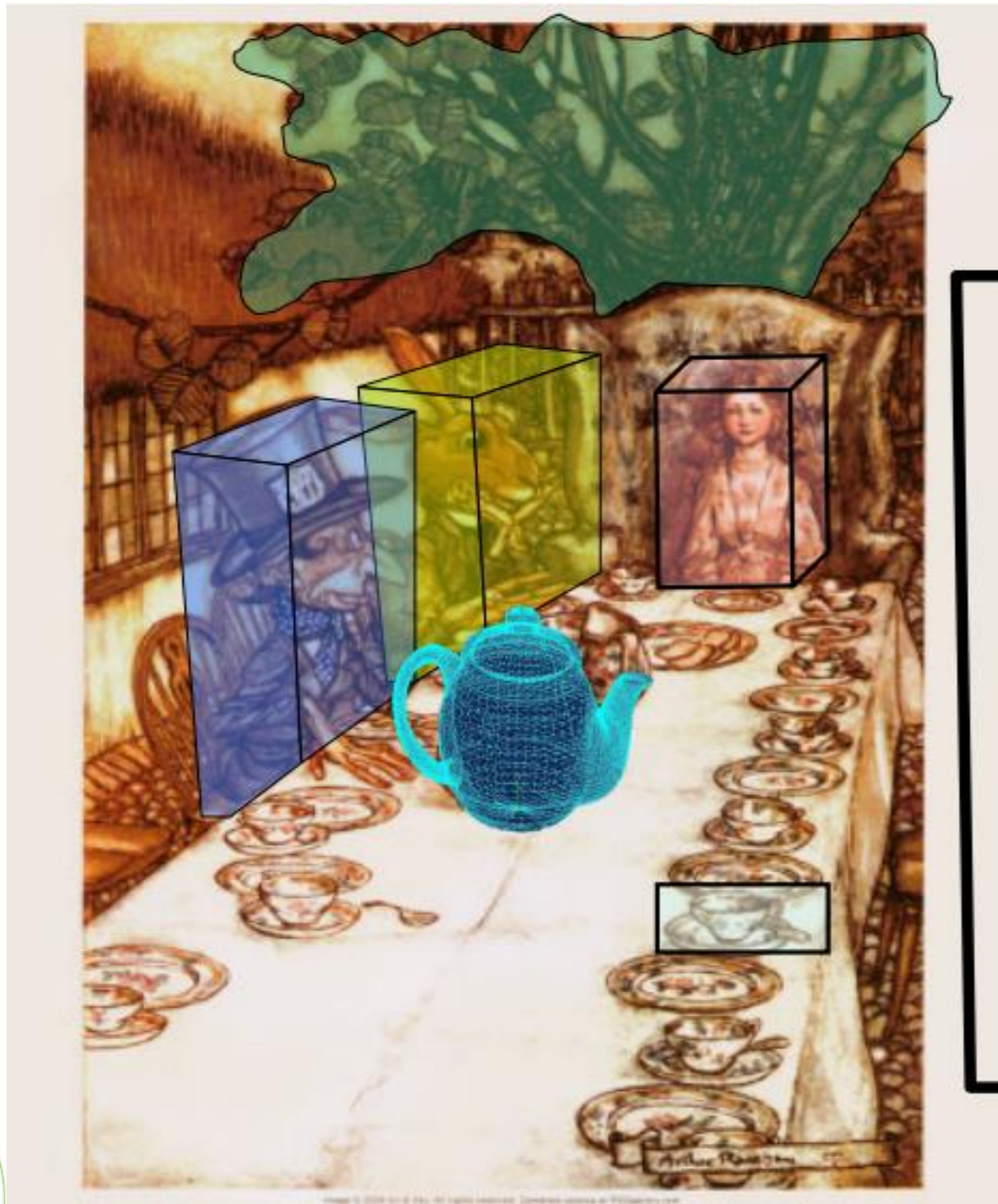
Computer vision



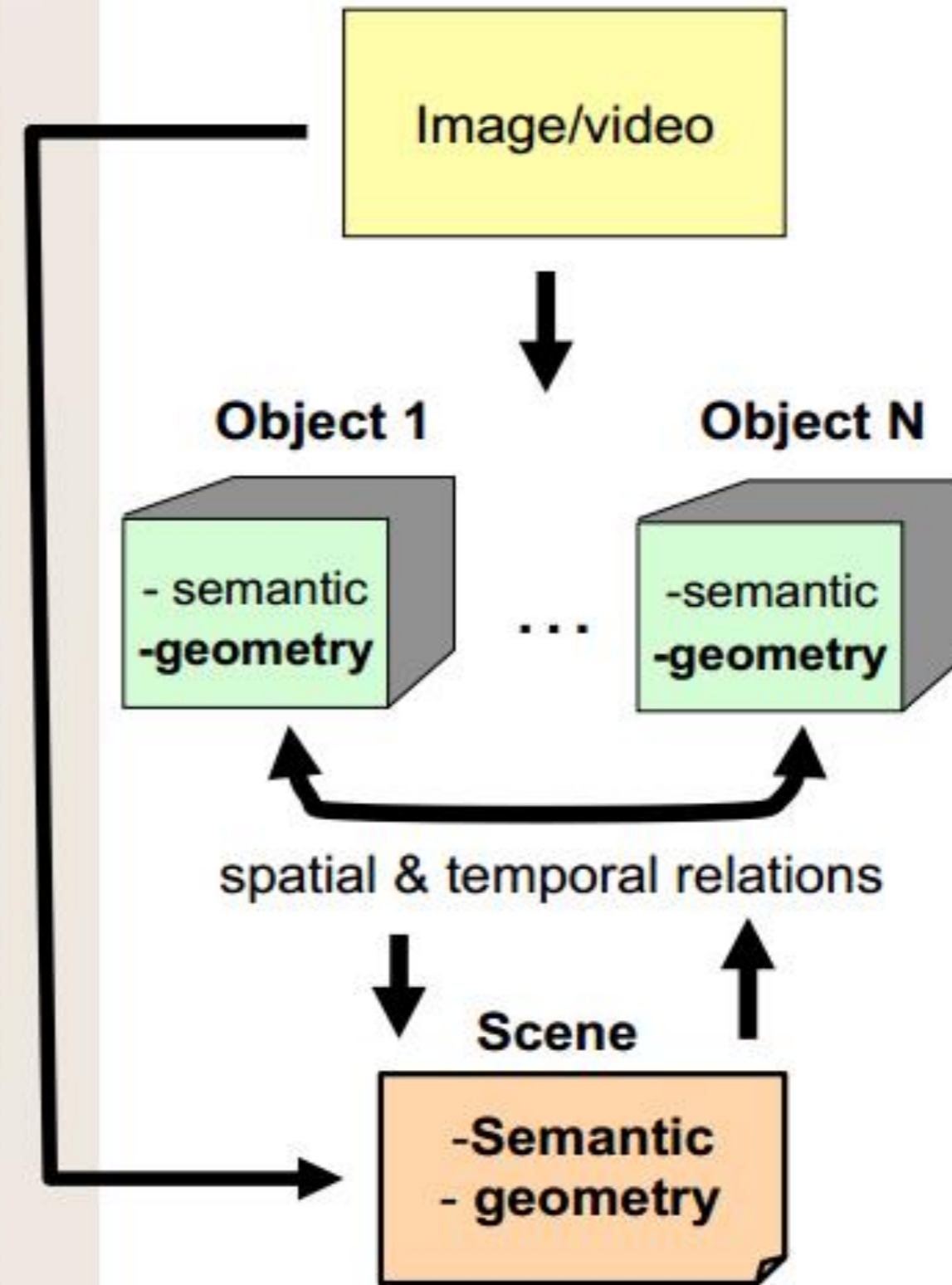


Computer vision



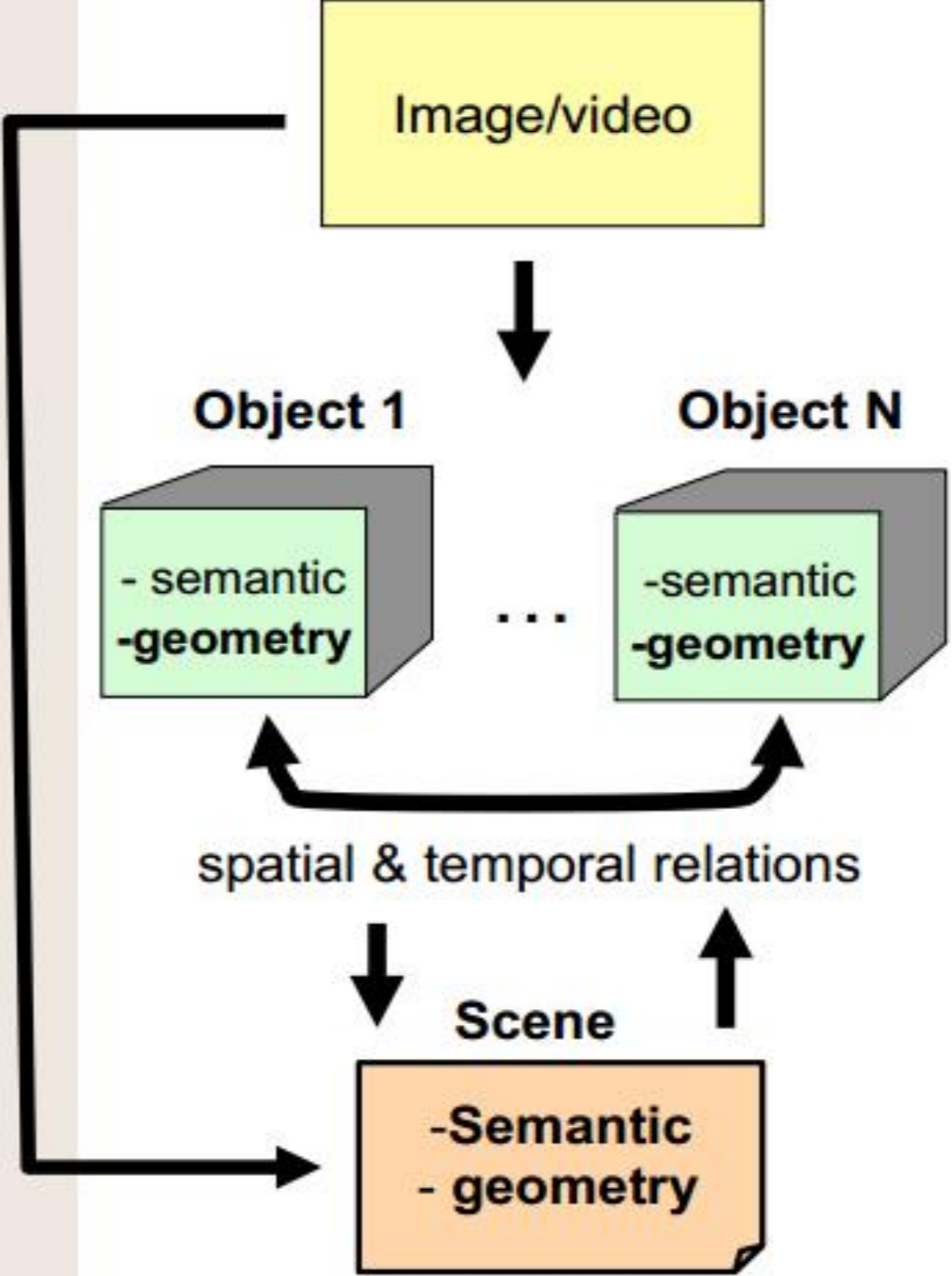


Computer vision





Computer vision



Computer Vision



Sensing device



Computational device



- Information extraction
- Interpretation

1. Information extraction: features, 3D structure, motion flows, etc...

2. Interpretation: recognize objects, scenes, actions, events



Computer Vision and Applications



Fingerprint biometrics



Augmentation with 3D Computer Graphics

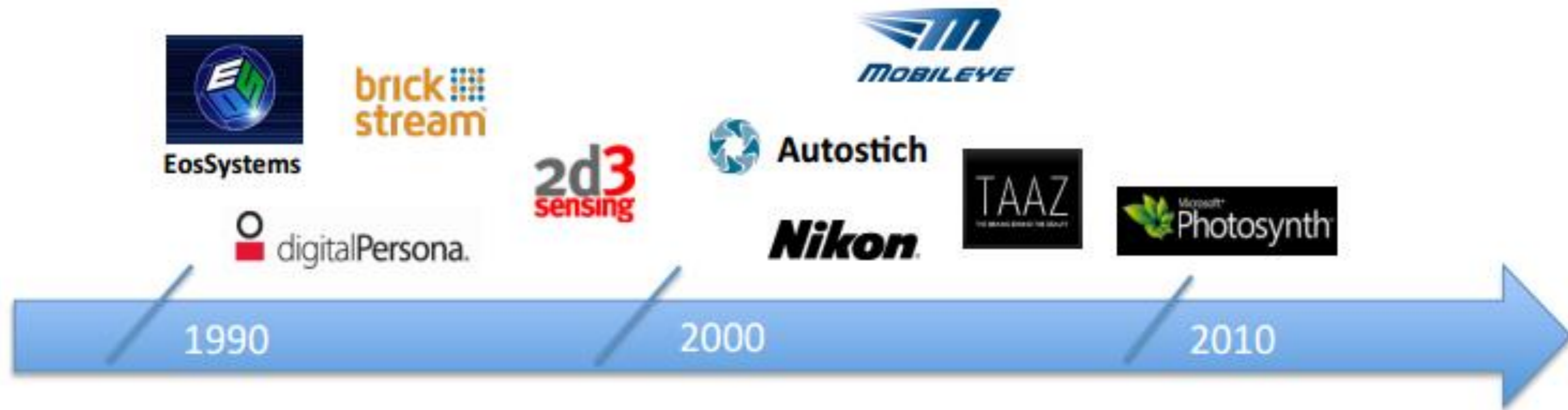


2d3
sensing



Computer Vision and Applications

- New features detector/descriptors
- CV leverages machine learning



Face Detection



BBC NEWS

UK version International version [About the versions](#) | [Log out](#)

Last Updated: Monday, 6 February 2006, 14:29 GMT

[E-mail this to a friend](#) [Printable version](#)

Face-hunting cameras boost Nikon

Japanese camera maker Nikon has tripled its profits on the back of strong sales of digital cameras that automatically focus on human faces.

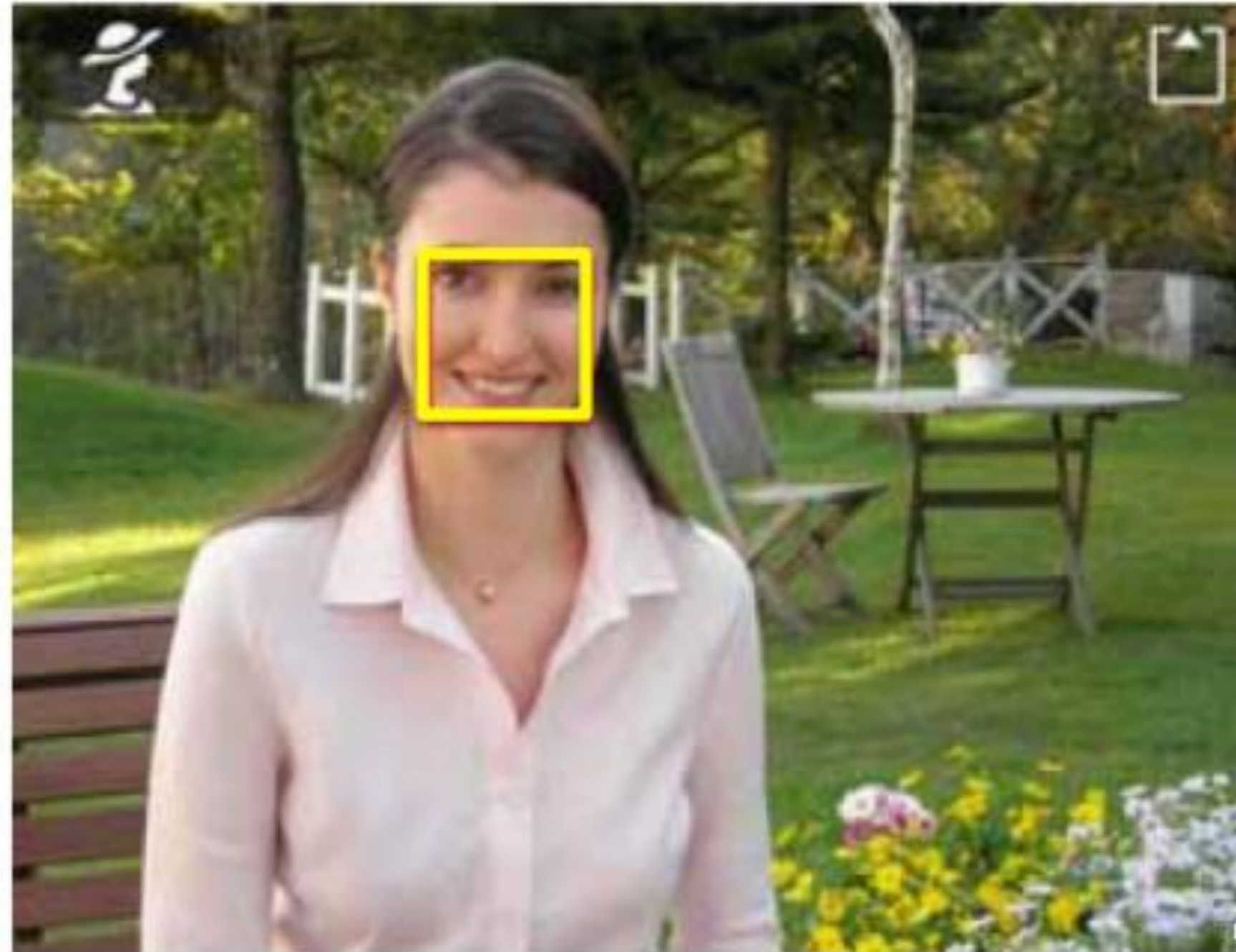


Face recognition cameras like the Coolpix L1 are popular.

News Front Page
World
UK
England
Northern Ireland
Scotland
Wales
Business
Market Data
Your Money
E-Commerce
Economy
Companies
Politics
Health
Education



Face Detection



Web applications

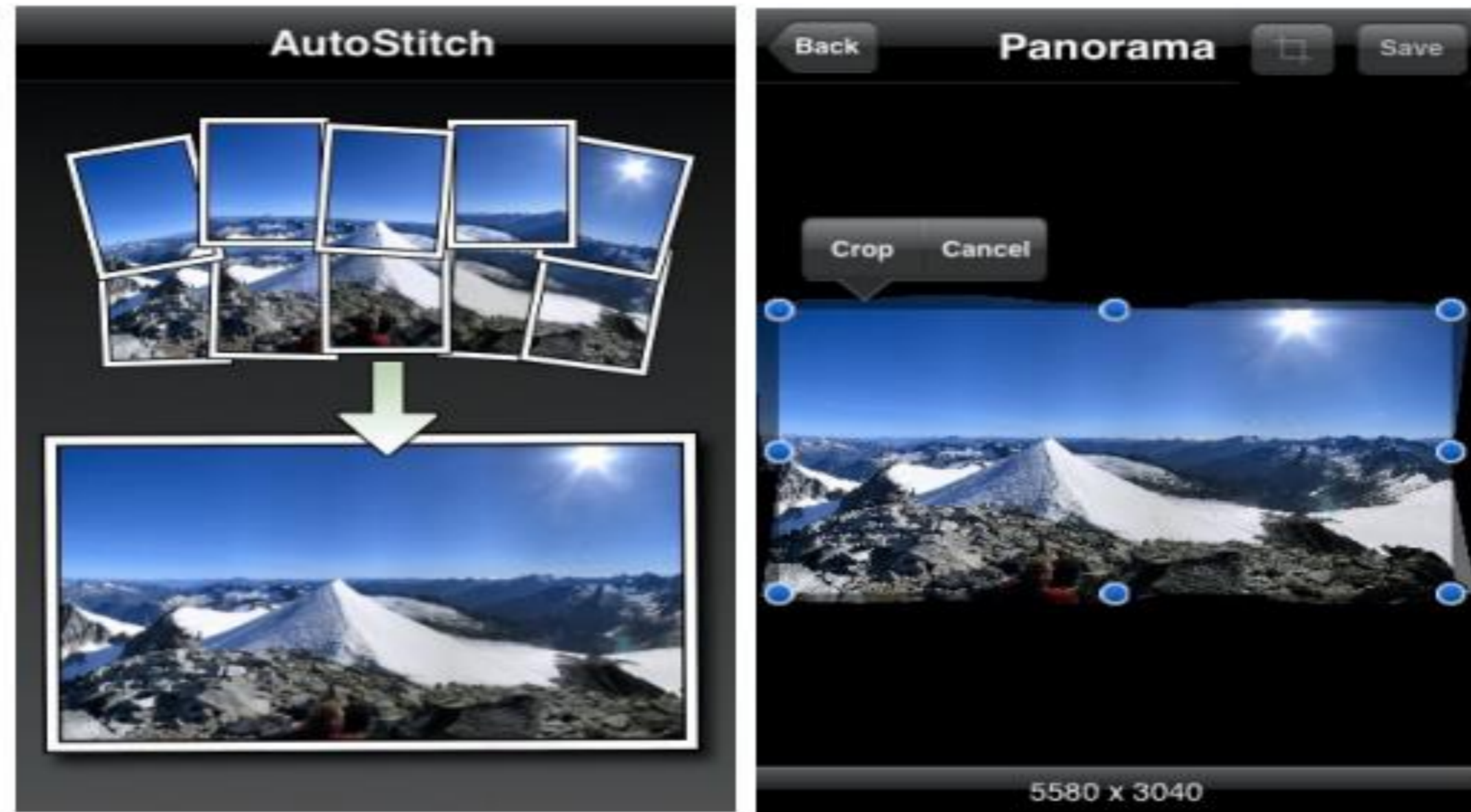


Photometria

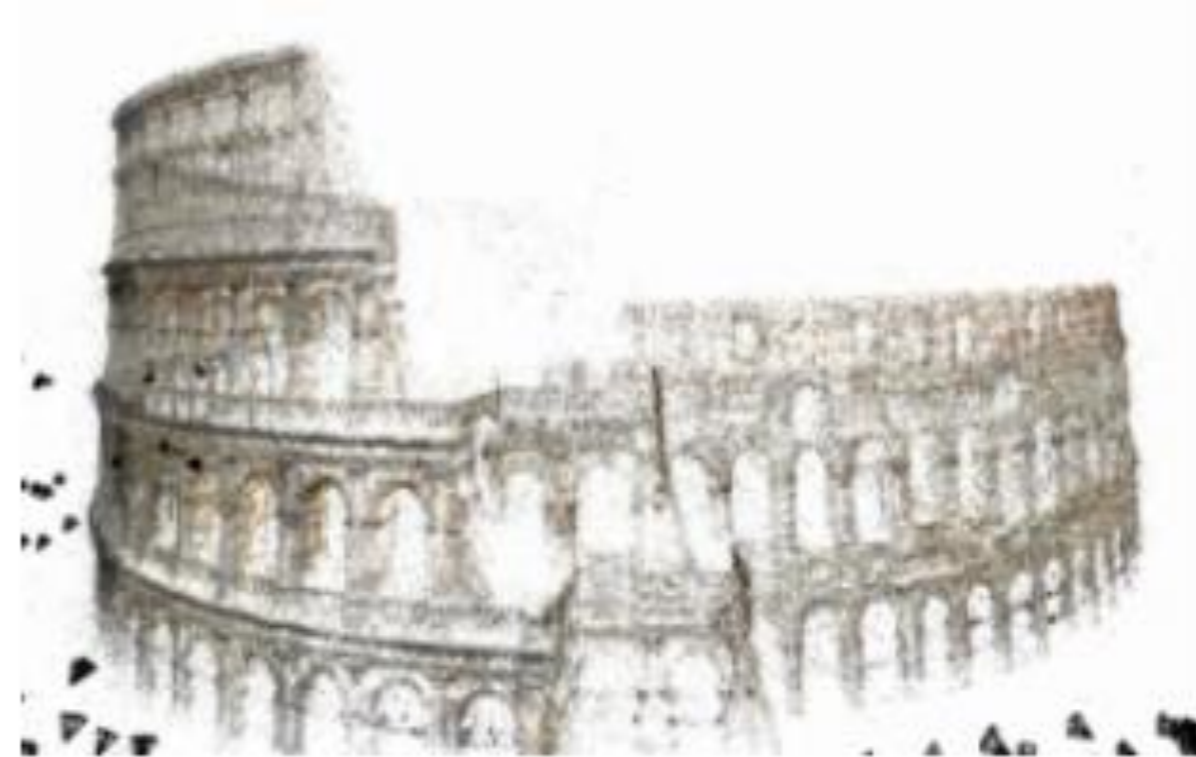


Computer Vision

Panoramic Photography



3D modeling of landmarks



Computer Vision and Applications

- Efficient SLAM/SFM
- Large scale image repositories
- Deep learning (e.g. ImageNet)
- Better clouds ☺
- More bandwidth
- Increase computational power

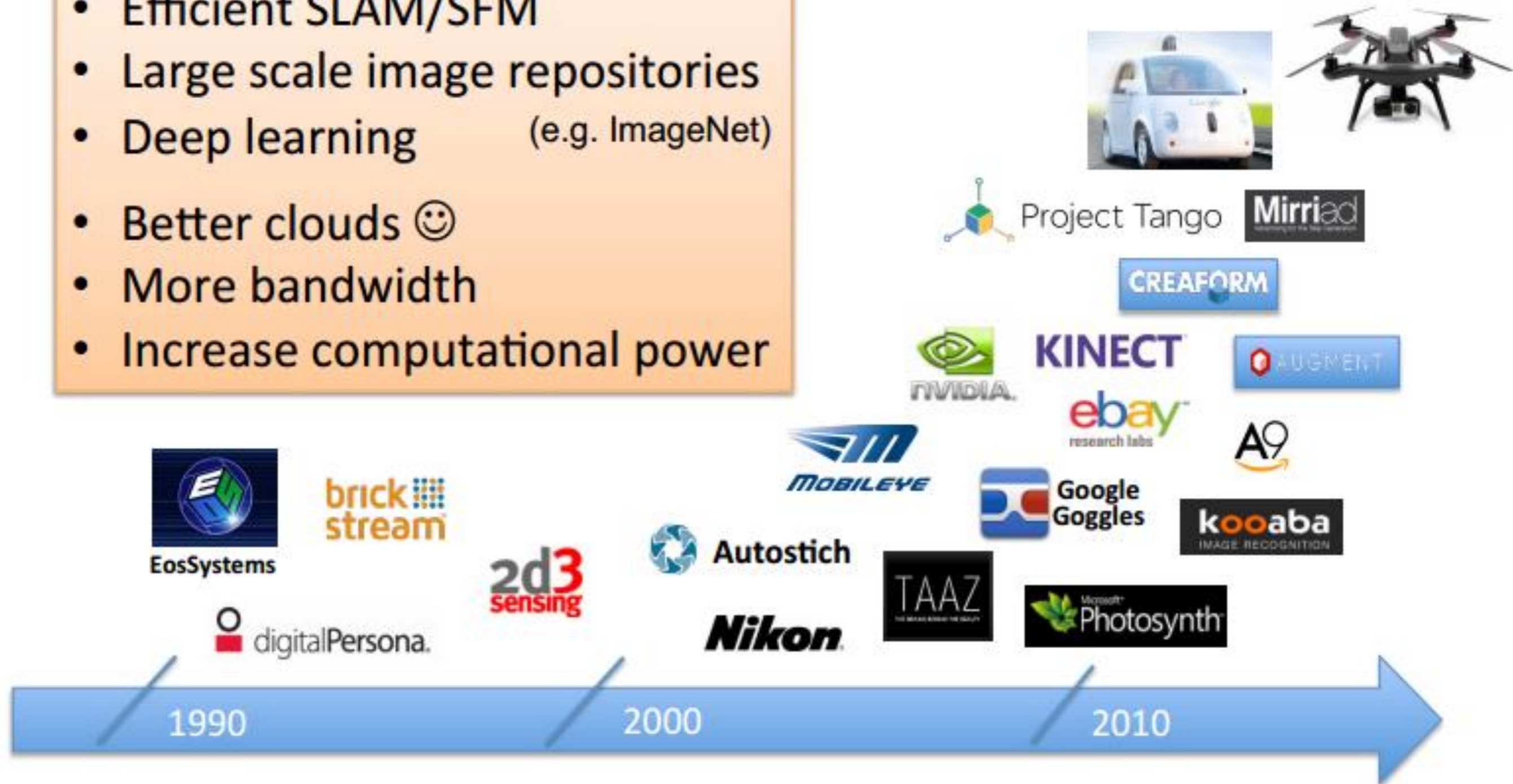
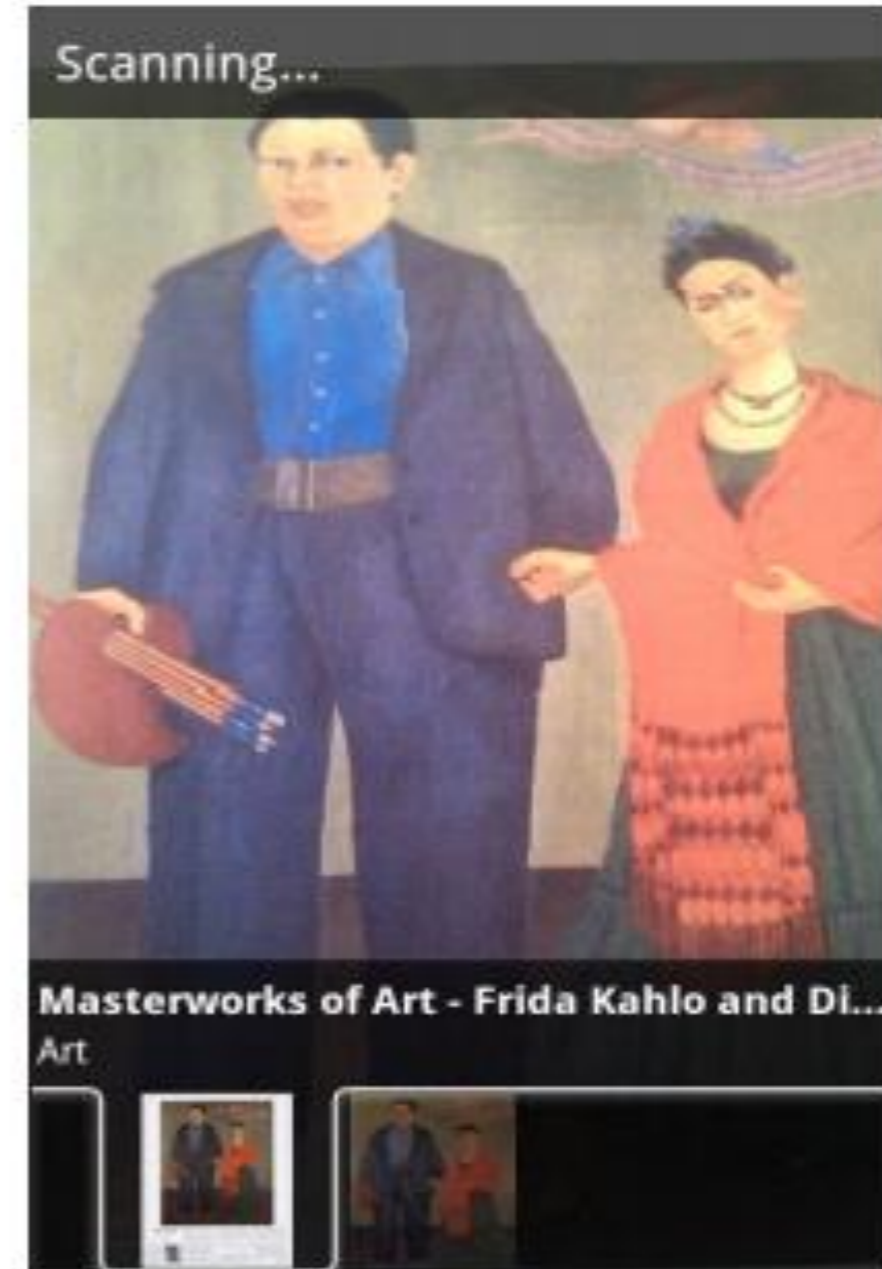


Image search engines



Visual search and landmarks recognition



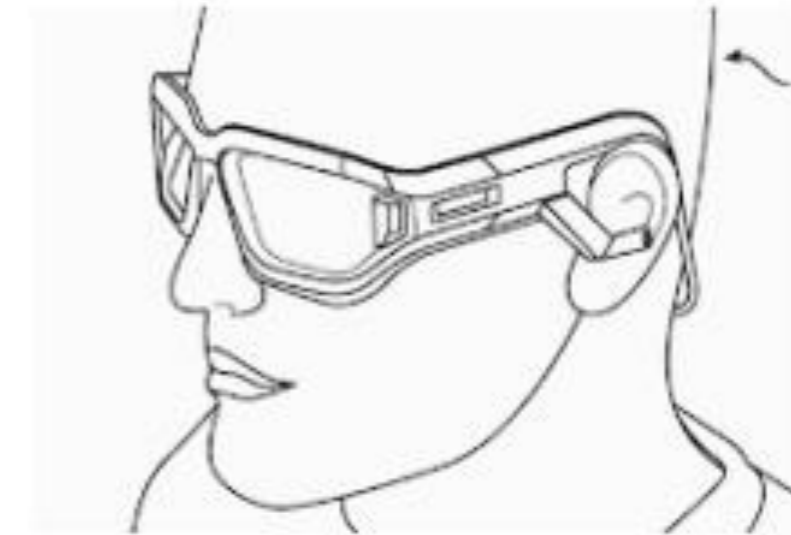
Google Goggles



Visual search and landmarks recognition



Augmented reality



- Magic leap
- Daqri
- Meta
- Etc...



Motion sensing and gesture recognition



Computer Vision



Autonomous navigation and safety

The screenshot displays the Mobileye website interface. At the top, there are navigation tabs for "manufacturer products" and "consumer products". The main header reads "Our Vision. Your Safety." Below this, a central image shows a car with three camera fields of view: "rear looking camera", "side looking camera", and "forward looking camera". To the right, a "News" section features two headlines: "Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System" and "Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end". Below the news is an "Events" section with two items: "Mobileye at Equip Auto, Paris, France" and "Mobileye at SEMA, Las Vegas, NV". At the bottom, three product highlights are shown: "EyeQ Vision on a Chip" with an image of a chip, "Vision Applications" with an image of a pedestrian and text "Road, Vehicle, Pedestrian Protection and more", and "AWS Advance Warning System" with an image of a car on a radar screen.

Mobileye: Vision systems in high-end BMW, GM, Volvo models But also, Toyota, Google, Apple, Tesla, Nissan, Ford, etc....



Personal Robotics



jibo

a

ROBOTICS FIVE ELEMENTS

ECOVACS

BLUE FROG robotics




















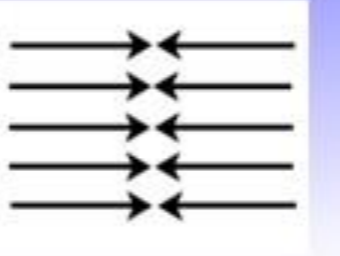

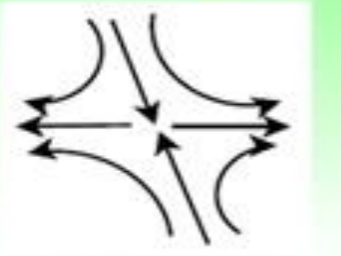
Crowd Management: Counting



Computer Vision



Crowd Management: Behavior

				
				
				
Bottlenecks	Fountainheads	Lanes	Arches	Blockings
				



Crowd Management: Anomaly



Tracking: Camouflage Objects



Computer Vision



Minimal Invasive Surgery



Computer Vision



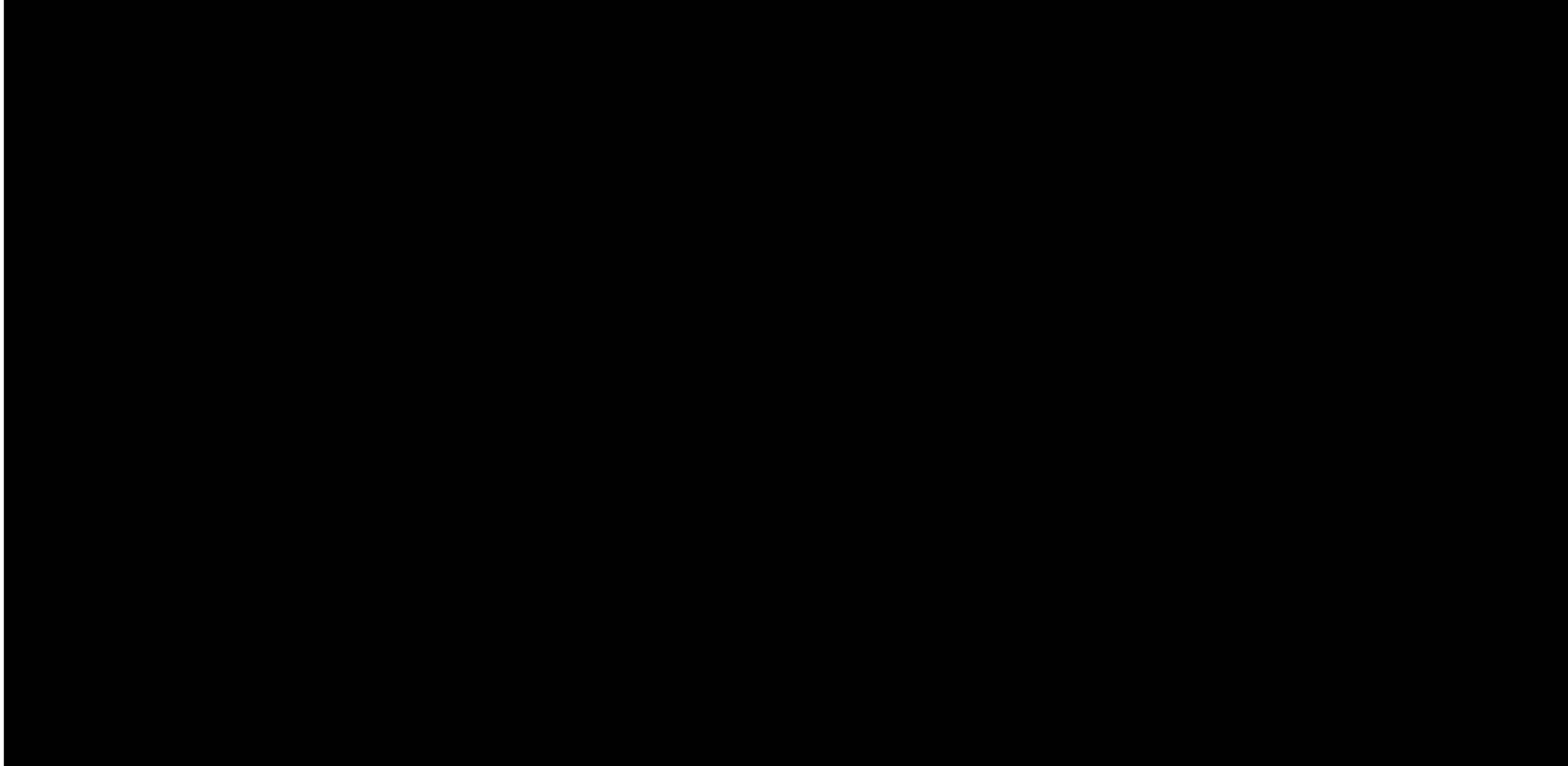
Future Operating Room



Computer Vision



Digital Cosmetics



Virtual Dressing Room

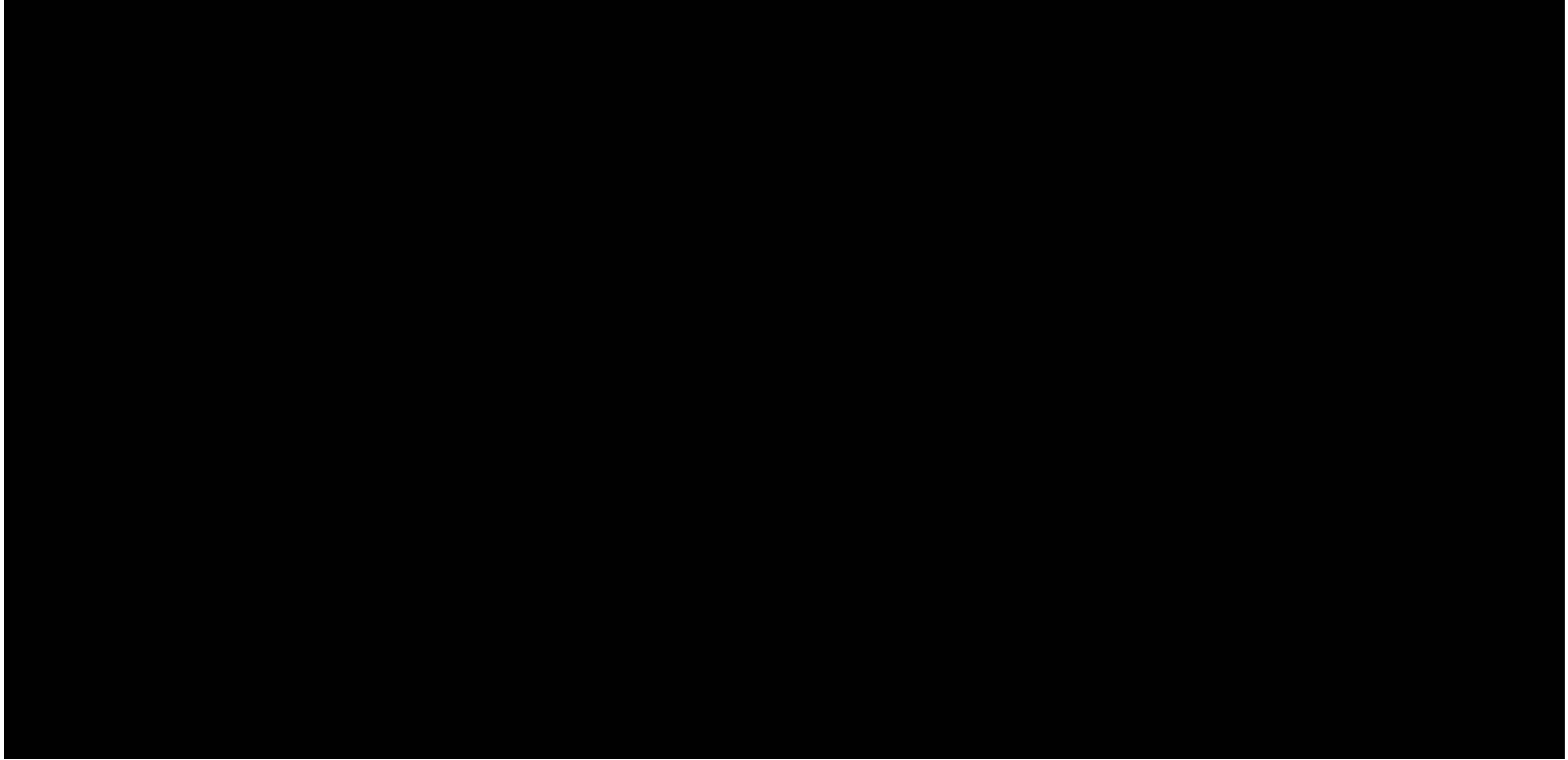
NEWTEMPO
新节奏智能科技



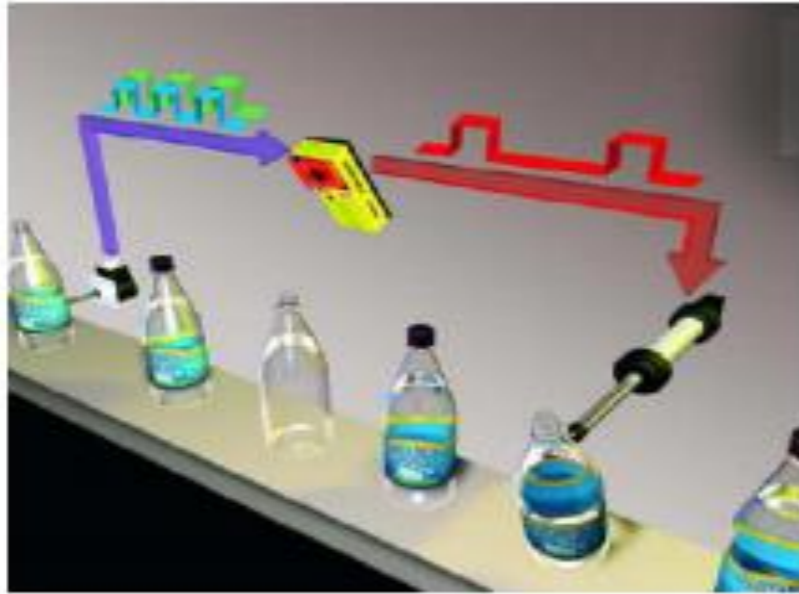
Virtual Furniture



Sports



Computer Vision and Applications



Factory inspection



Assistive technologies



Surveillance



Vision for robotics, space exploration



Security

Computer Vision



Why Computer Vision?

- Images and video are everywhere!



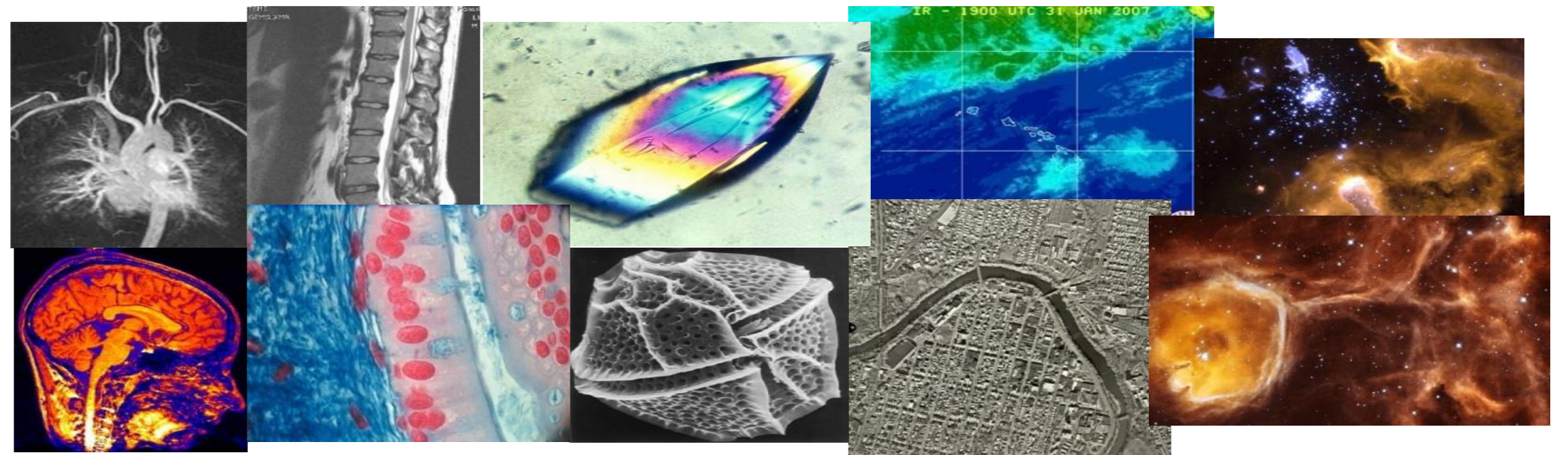
Personal photo albums



Movies, news, sports



Surveillance and security



Computer Vision

Medical and scientific images

Current state of computer vision



3D Reconstruction

- 3D shape recovery
- 3D scene reconstruction
- Camera localization
- Pose estimation



2D Recognition

- Object detection
- Texture classification
- Target tracking
- Activity recognition



Current state of computer vision



3D Reconstruction

- 3D shape recovery
- 3D scene reconstruction
- Camera localization
- Pose estimation



Snavely et al., 06-08

Lucas & Kanade, 81
Chen & Medioni, 92
Debevec et al., 96
Levoy & Hanrahan, 96
Fitzgibbon & Zisserman, 98
Triggs et al., 99
Pollefeys et al., 99
Kutulakos & Seitz, 99

Levoy et al., 00
Hartley & Zisserman, 00
Dellaert et al., 00
Rusinkiewicz et al., 02
Nistér, 04
Brown & Lowe, 04
Schindler et al., 04
Lourakis & Argyros, 04
Colombo et al., 05

Golparvar-Fard, et al. JAEI 10
Pandey et al. IFAC, 2010
Pandey et al. ICRA 2011
Savarese et al. IJCV 05
Savarese et al. IJCV 06
Microsoft's PhotoSynth
Snavely et al., 06-08
Schindler et al., 08
Agarwal et al., 09 45
Frahm et al., 10



Current state of computer vision



3D Reconstruction

- 3D shape recovery
- 3D scene reconstruction
- Camera localization
- Pose estimation



Snavely et al., 06-08

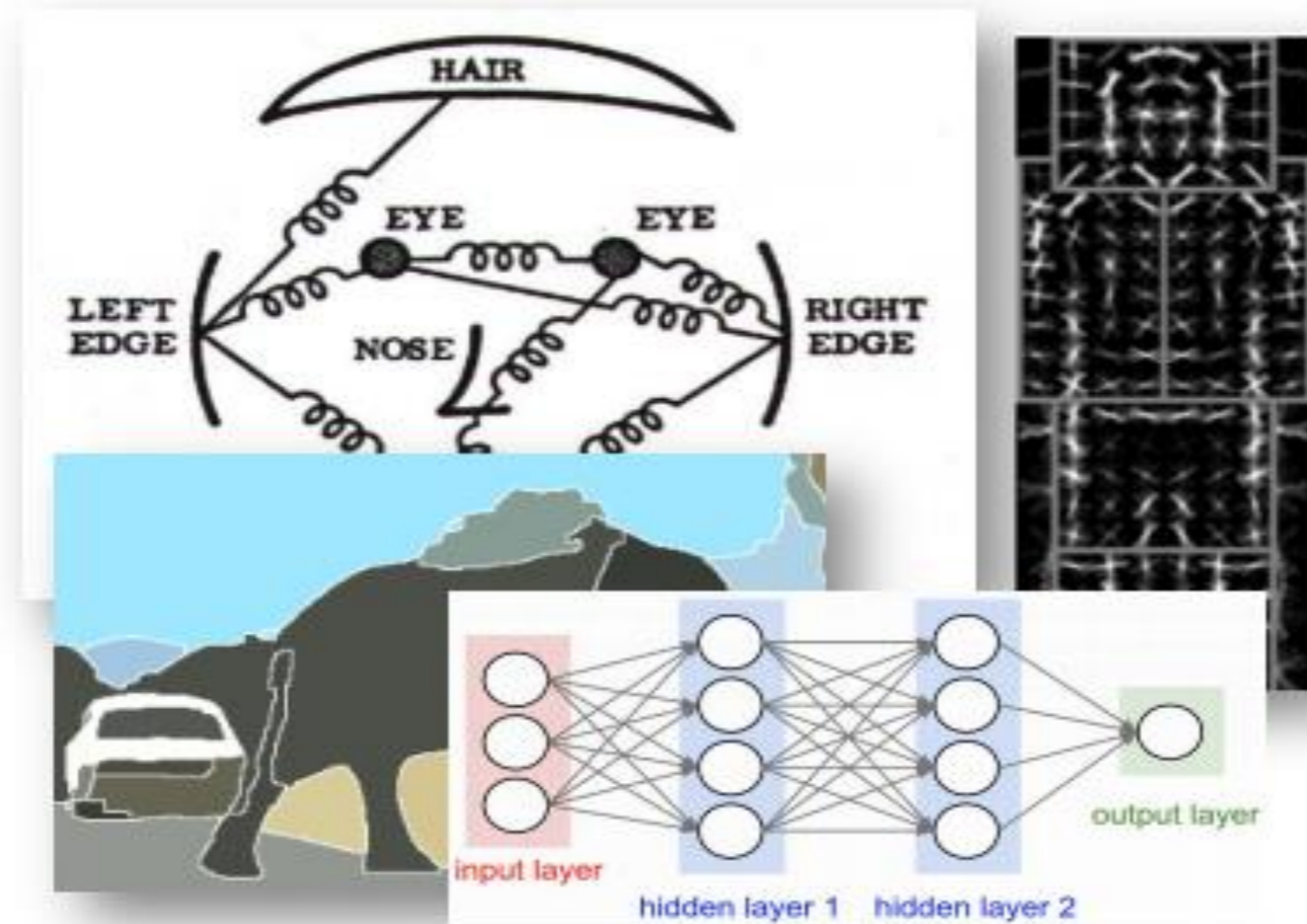
Lucas & Kanade, 81
Chen & Medioni, 92
Debevec et al., 96
Levoy & Hanrahan, 96
Fitzgibbon & Zisserman, 98
Triggs et al., 99
Pollefeys et al., 99
Kutulakos & Seitz, 99

Levoy et al., 00
Hartley & Zisserman, 00
Dellaert et al., 00
Rusinkiewicz et al., 02
Nistér, 04
Brown & Lowe, 04
Schindler et al., 04
Lourakis & Argyros, 04
Colombo et al., 05

Golparvar-Fard, et al. JAEI 10
Pandey et al. IFAC, 2010
Pandey et al. ICRA 2011
Savarese et al. IJCV 05
Savarese et al. IJCV 06
Microsoft's PhotoSynth
Snavely et al., 06-08
Schindler et al., 08
Agarwal et al., 09
Frahm et al., 10



Current state of computer vision



2D Recognition

- Object detection
- Texture classification
- Target tracking
- Activity recognition

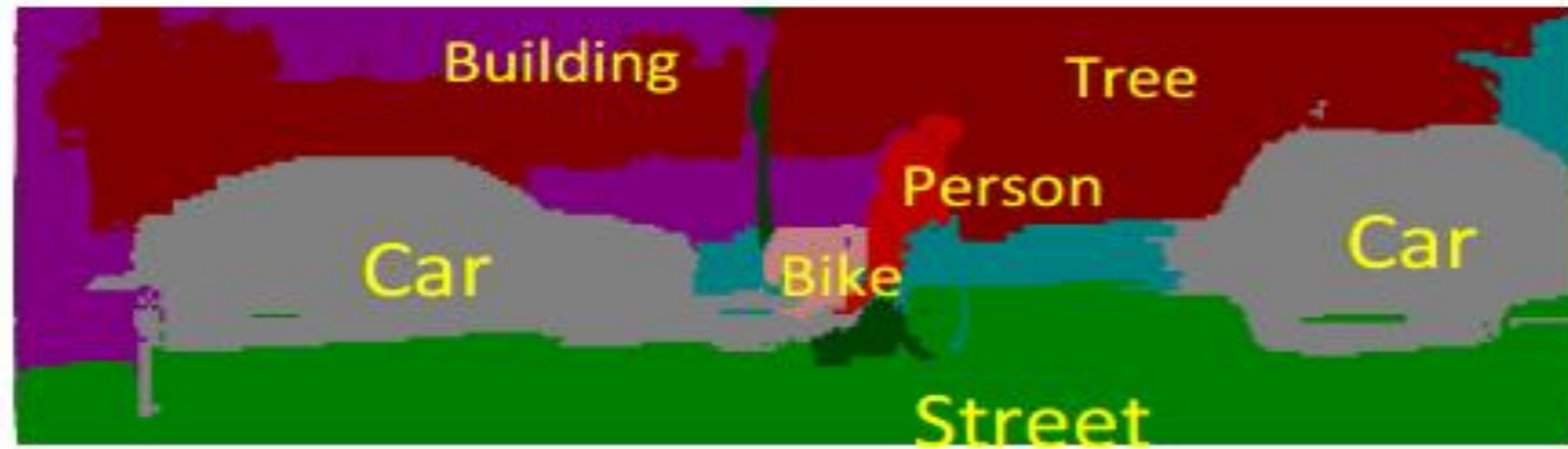
Turk & Pentland, 91
Poggio et al., 93
Belhumeur et al., 97
LeCun et al. 98
Amit and Geman, 99
Shi & Malik, 00
Viola & Jones, 00
Felzenszwalb & Huttenlocher 00
Belongie & Malik, 02
Ullman et al. 02

Argawal & Roth, 02
Ramanan & Forsyth, 03
Weber et al., 00
Vidal-Naquet & Ullman 02
Fergus et al., 03
Torralba et al., 03
Vogel & Schiele, 03
Barnard et al., 03
Fei-Fei et al., 04
Kumar & Hebert '04

He et al. 06
Gould et al. 08
Maire et al. 08
Felzenszwalb et al., 08
Kohli et al. 09
L.-J. Li et al. 09
Ladicky et al. 10,11
Gonfaus et al. 10
Farhadi et al., 09
Lampert et al., 09



Current state of computer vision



2D Recognition

- Object detection
- Texture classification
- Target tracking
- Activity recognition

Turk & Pentland, 91
Poggio et al., 93
Belhumeur et al., 97
LeCun et al. 98
Amit and Geman, 99
Shi & Malik, 00
Viola & Jones, 00
Felzenszwalb & Huttenlocher 00
Belongie & Malik, 02
Ullman et al. 02

Argawal & Roth, 02
Ramanan & Forsyth, 03
Weber et al., 00
Vidal-Naquet & Ullman 02
Fergus et al., 03
Torralba et al., 03
Vogel & Schiele, 03
Barnard et al., 03
Fei-Fei et al., 04
Kumar & Hebert '04

He et al. 06
Gould et al. 08
Maire et al. 08
Felzenszwalb et al., 08
Kohli et al. 09
L.-J. Li et al. 09
Ladicky et al. 10,11
Gonfous et al. 10
Farhadi et al., 09
Lampert et al., 09



Current state of computer vision



3D Reconstruction

- 3D shape recovery
- 3D scene reconstruction
- Camera localization
- Pose estimation



2D Recognition

- Object detection
- Texture classification
- Target tracking
- Activity recognition

Perceiving the World in 3D!

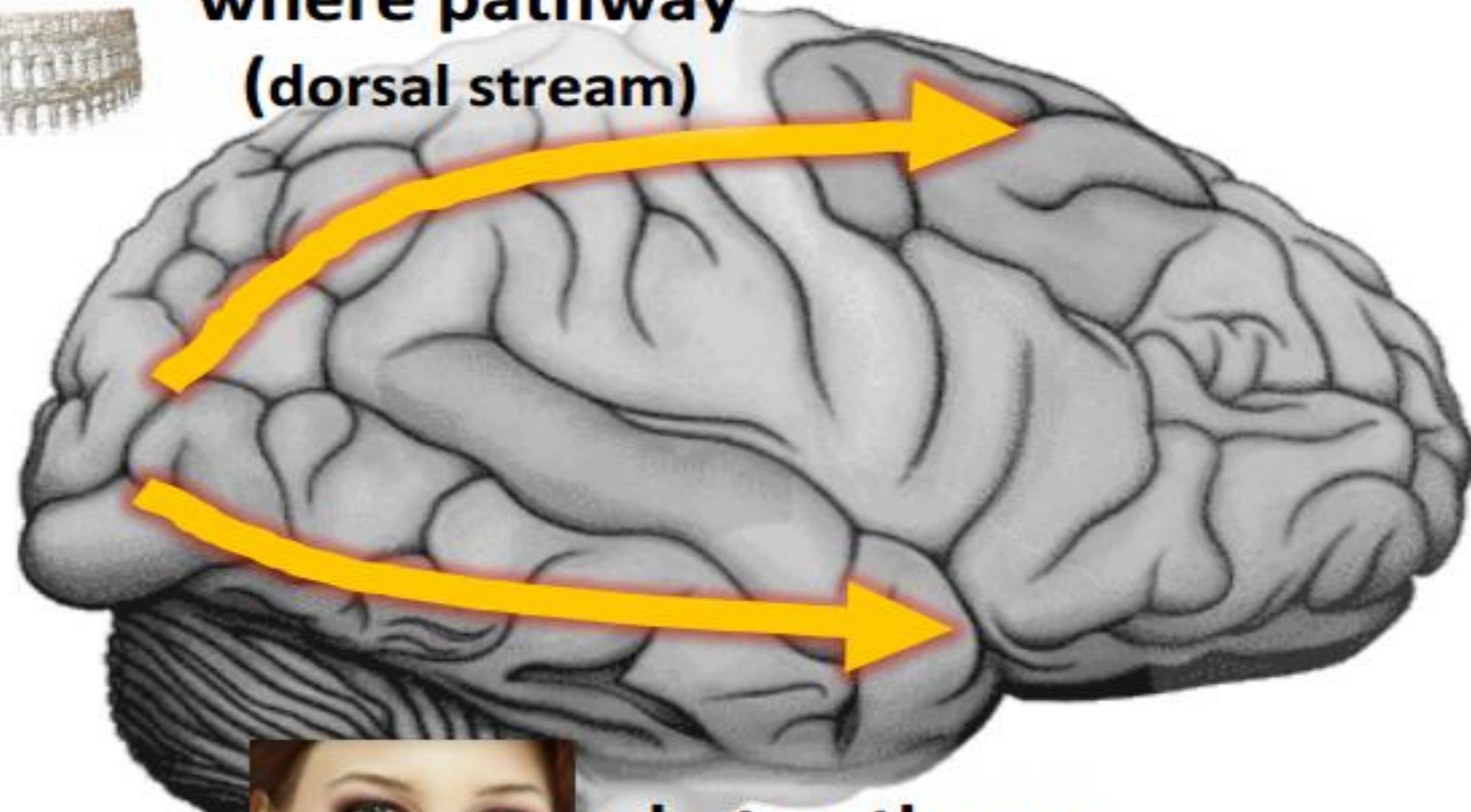


Visual Processing in the brain



where pathway
(dorsal stream)

V1



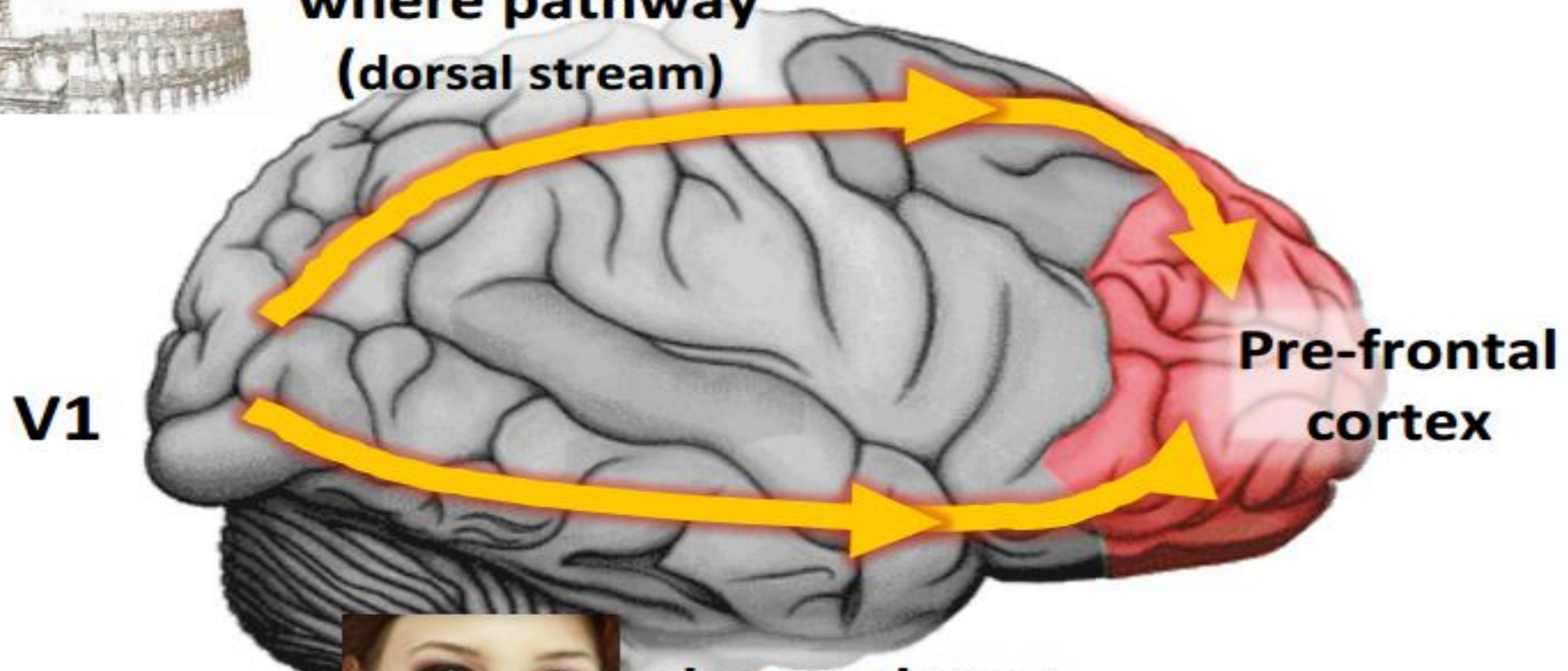
what pathway
(ventral stream)



Visual Processing in the brain



where pathway
(dorsal stream)



what pathway
(ventral stream)



Course Overview

1. Geometry

2. Semantics

Geometry:

- How to extract 3d information?
- Which cues are useful?
- What are the mathematical tools?



Camera Systems

Establish a mapping from 3D to 2D



Computer Vision



How to calibrate a camera

Estimate camera parameters such pose or focal length



Single view metrology

Estimate 3D properties of the world from a single image

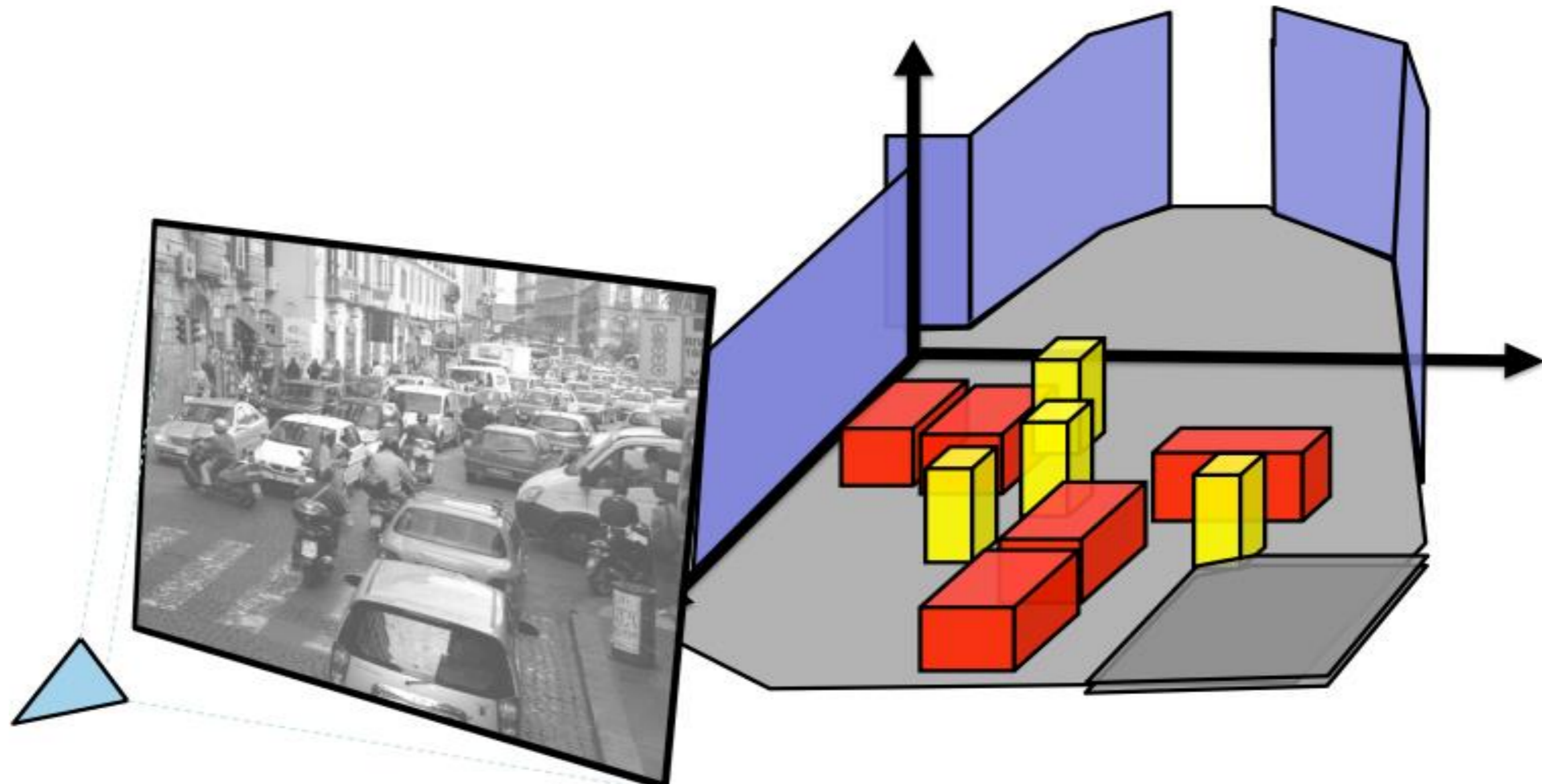


Computer Vision



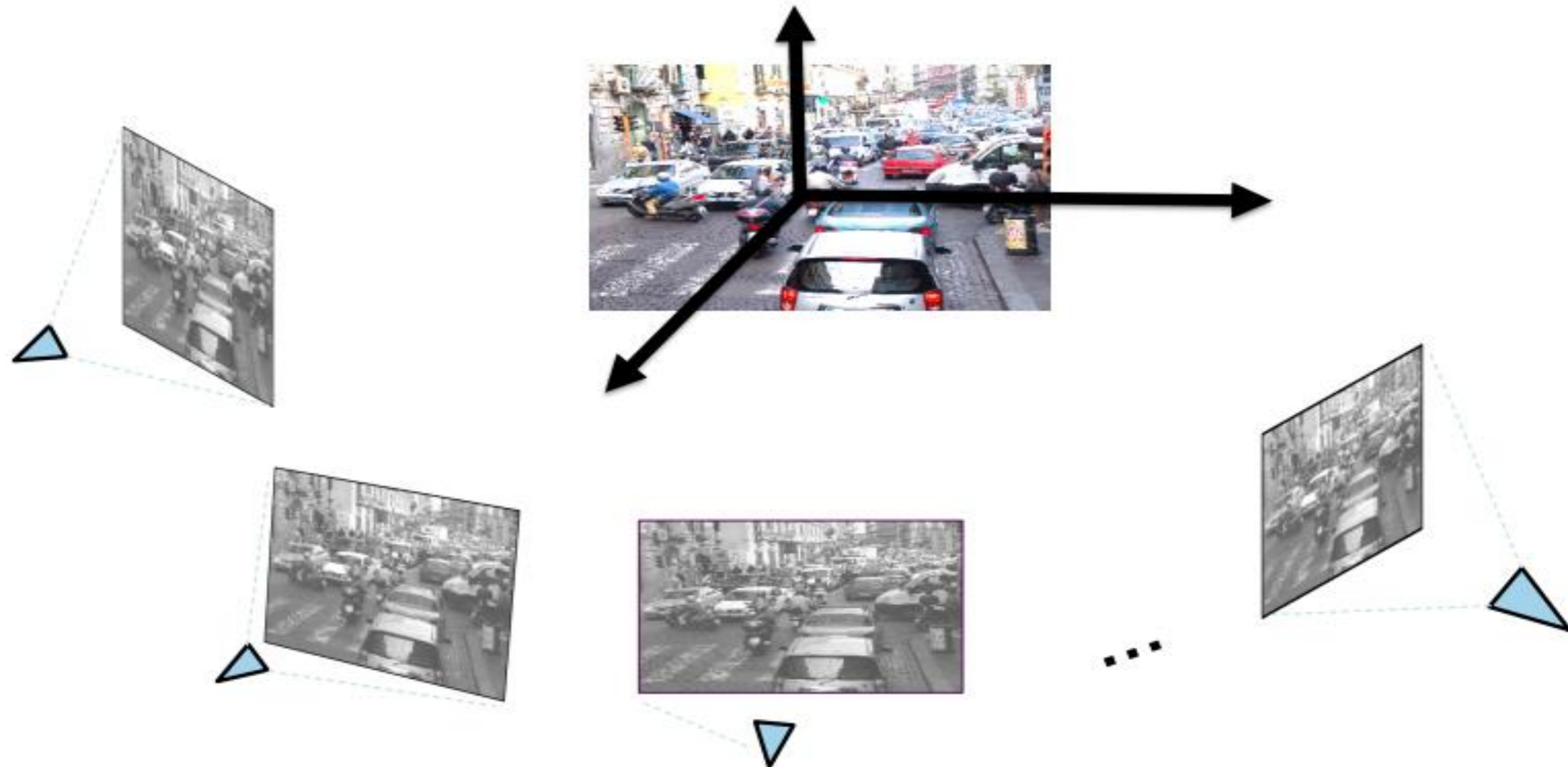
Single view metrology

Estimate 3D properties of the world from a single image



Multiple view geometry

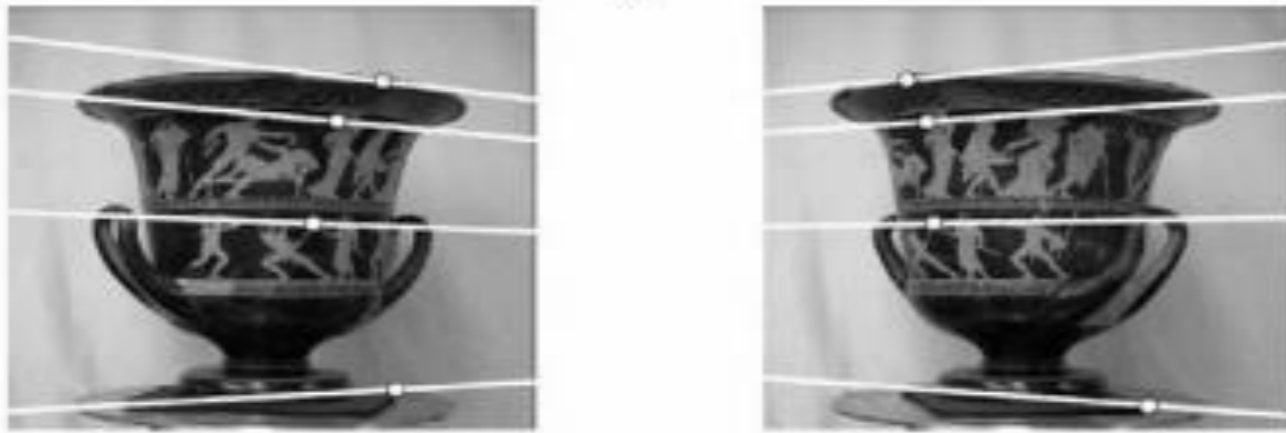
Estimate 3D properties of the world from multiple views



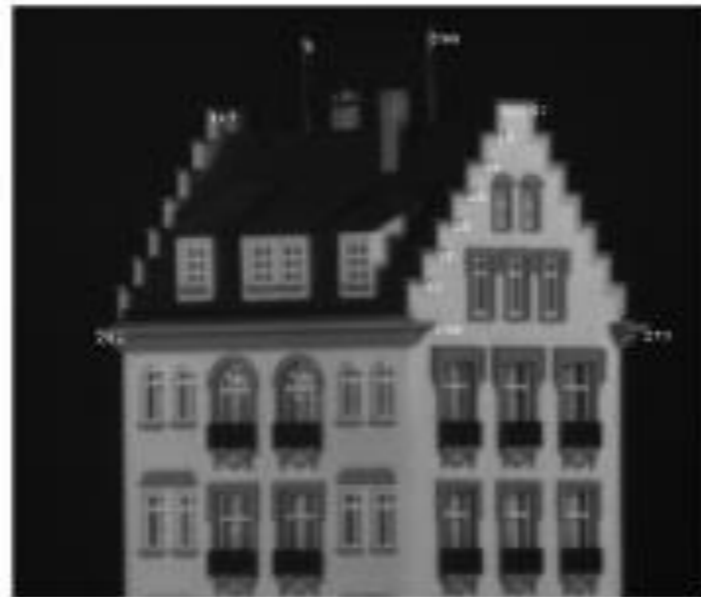
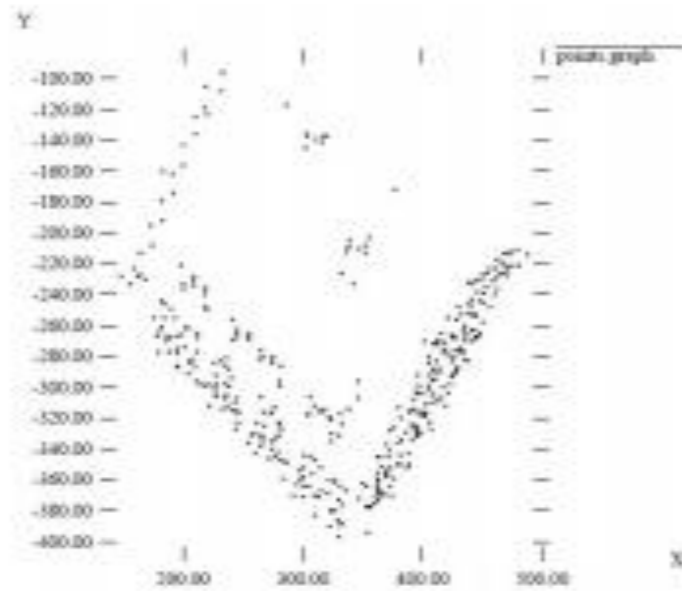
Computer Vision



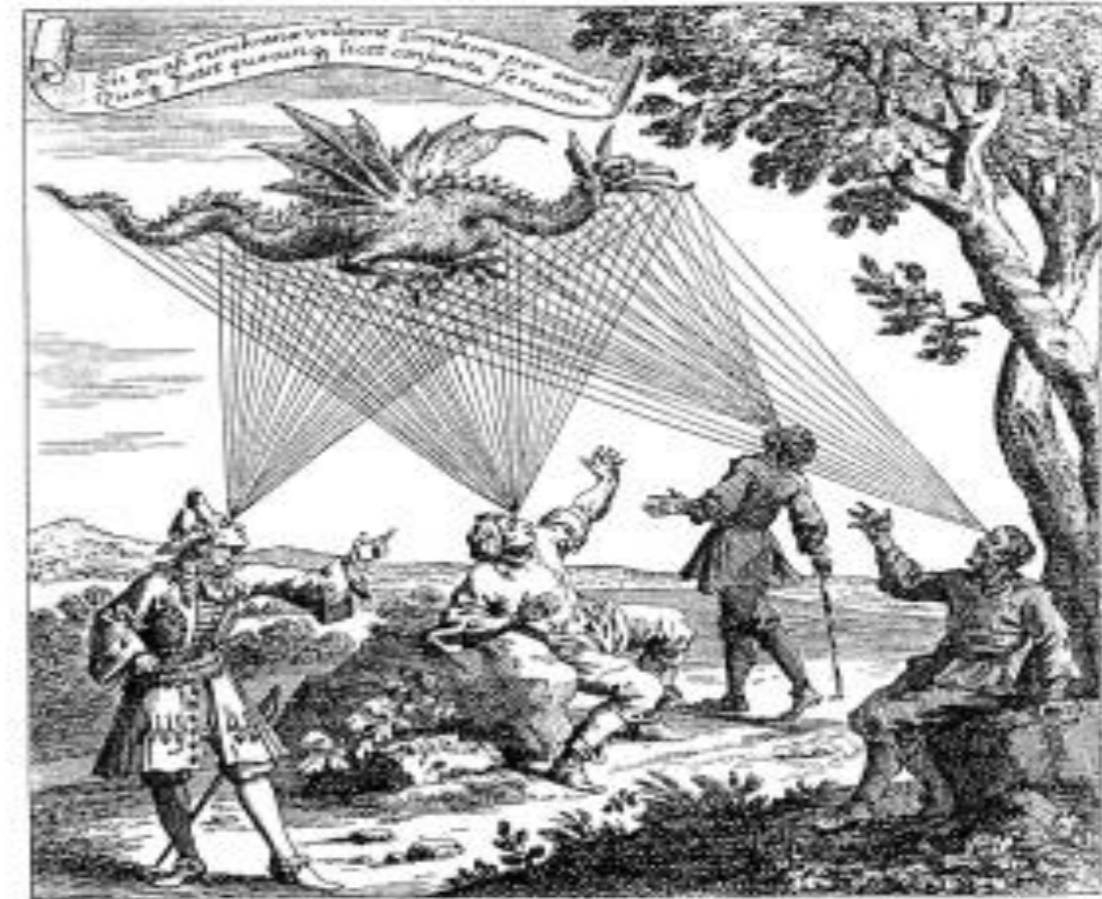
Mathematical tools



Epipolar geometry



Tomasi & Kanade (1993)

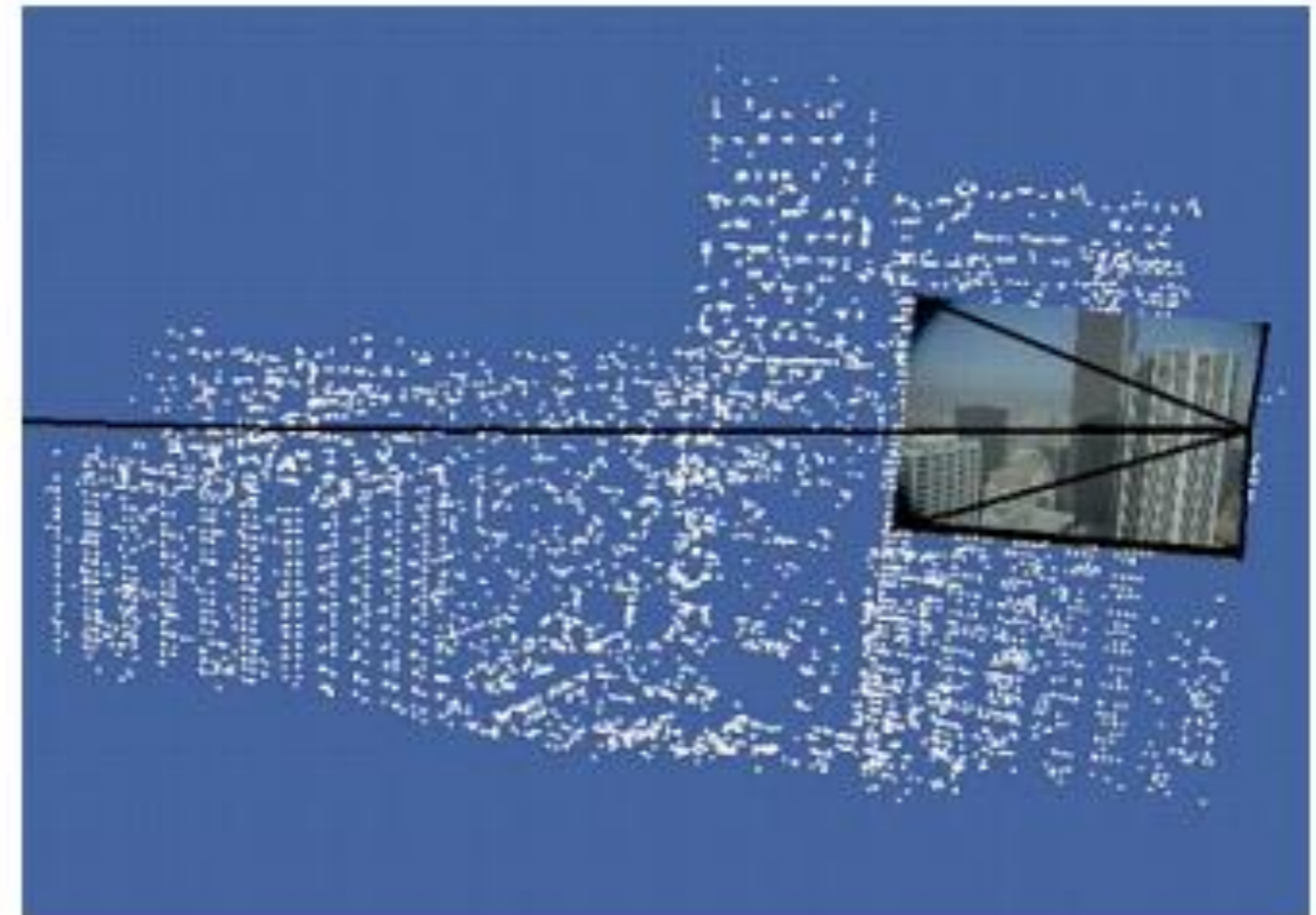


Драконъ, виденный подъ различными видами зрителя
По способу не иначе как „общимъ артифактомъ телескопическимъ“ Нави, 1790 года

Photoconsistency



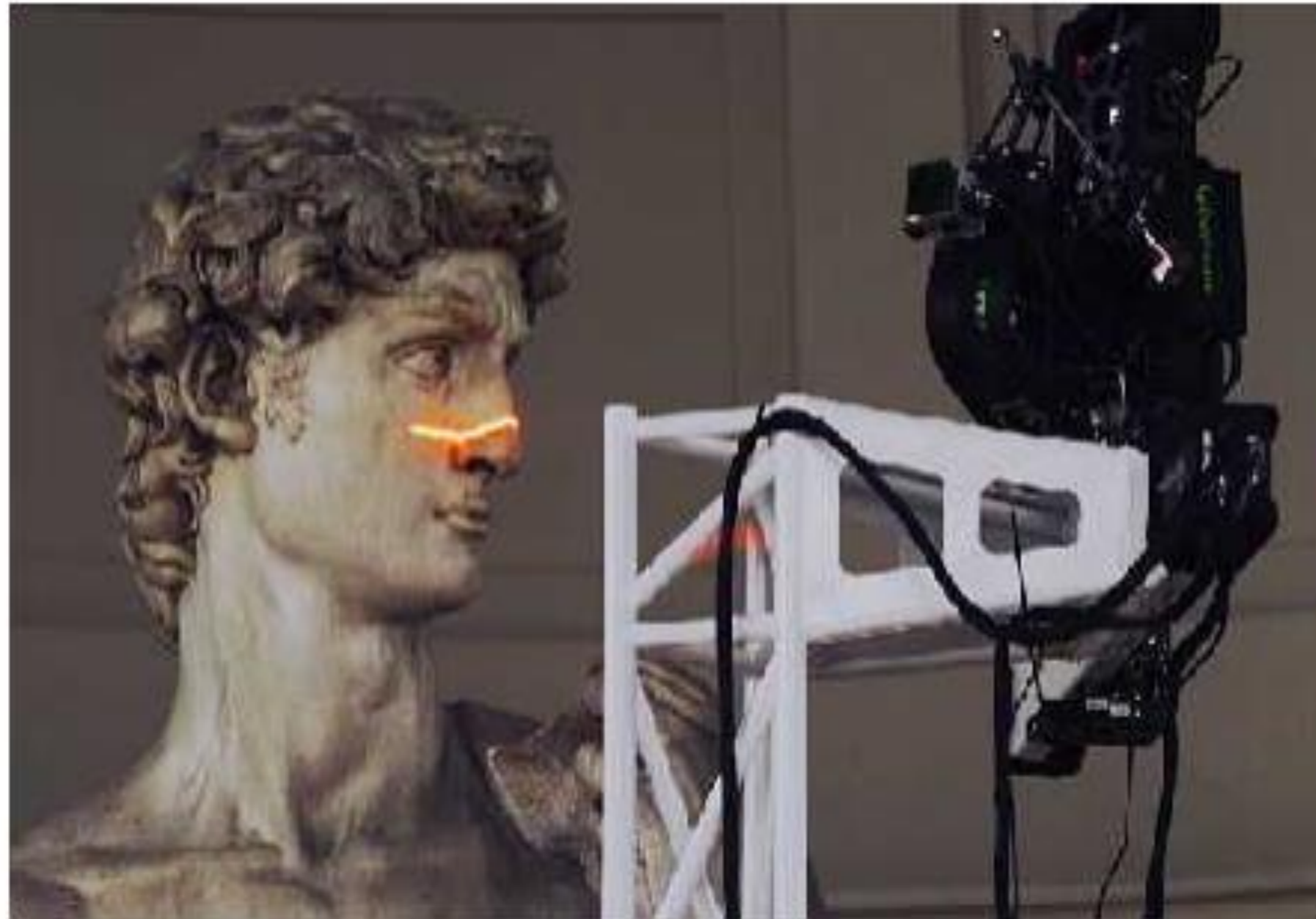
Structure from motion



Courtesy of Oxford **Visual Geometry Group**



Structure lighting and volumetric stereo



Course Overview

1. Geometry

2. **Semantics**

Semantics:

- How to recognize objects?
- How to classify images or understand a scene?
- How to segment out critical semantics?
- How to estimate 3D properties (i.e. pose, size, shape...)?



Object Recognition and Categorization



Classification: Is this an forest?



Classification:

Does this image contain a building? [yes/no]



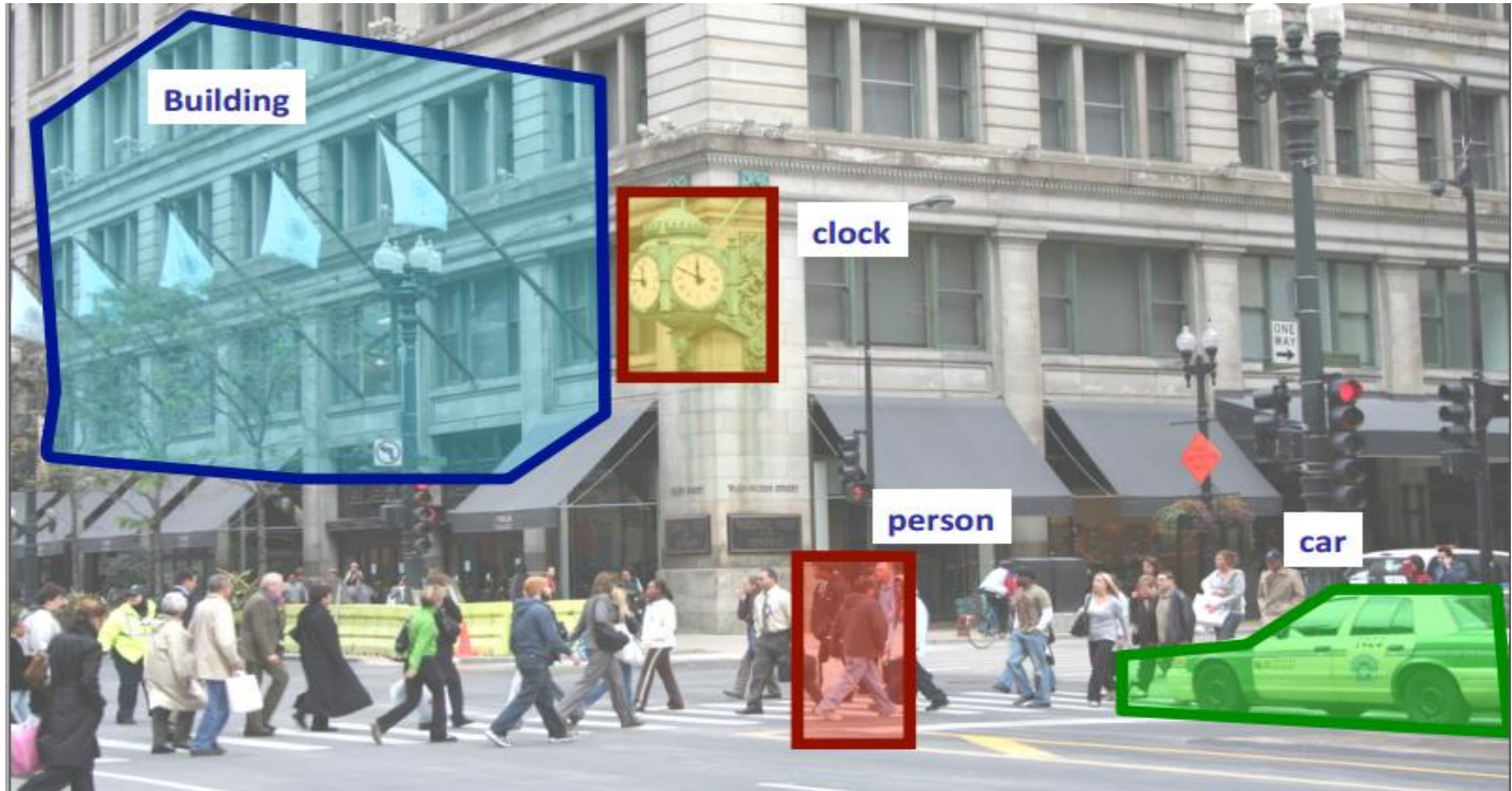
Detection:

Does this image contain a car? [where?]



Detection:

Which objects do this image contain? [where?]



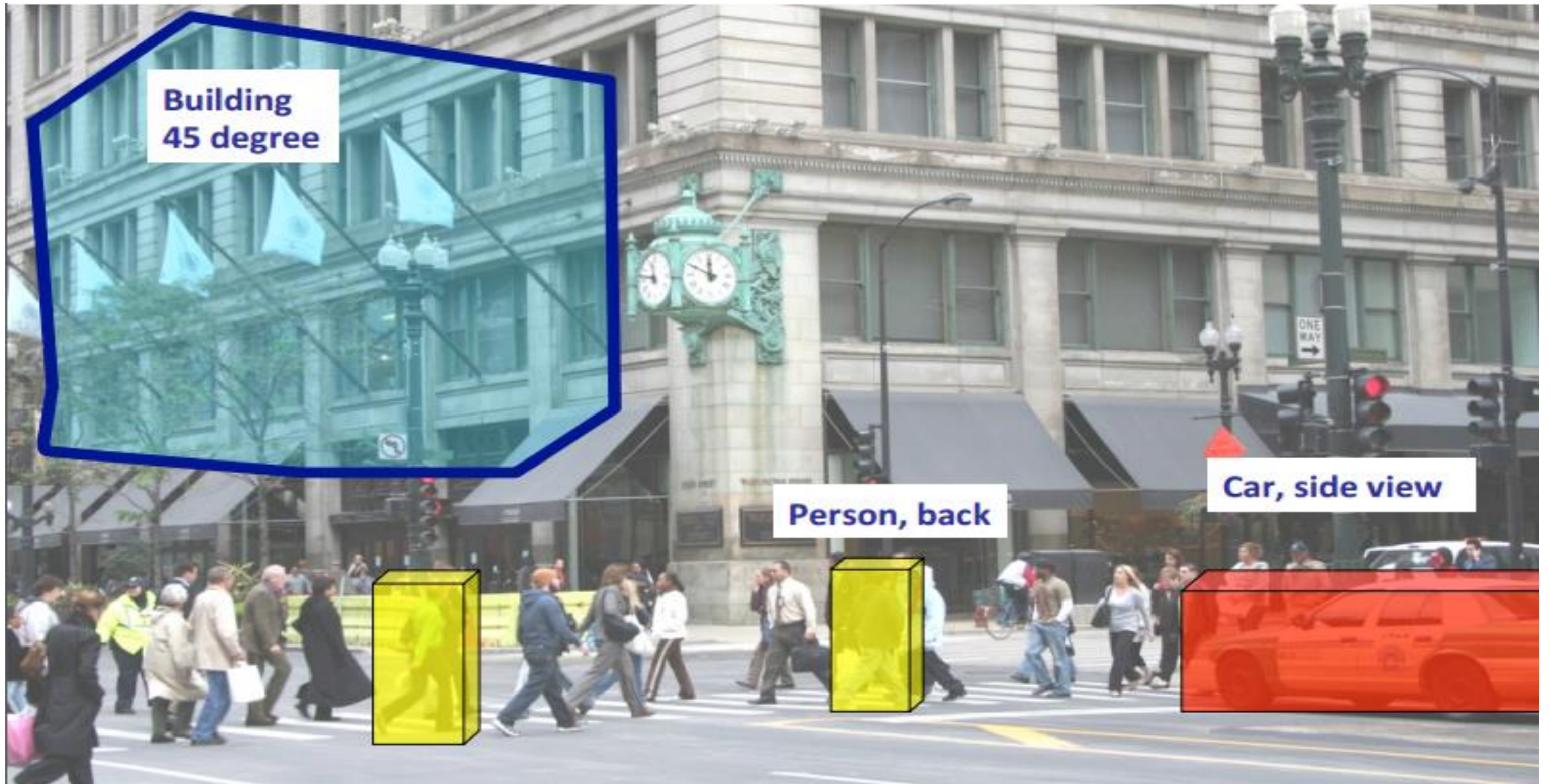
Detection:

Accurate localization (segmentation)



Detection:

Estimating 3D geometrical properties



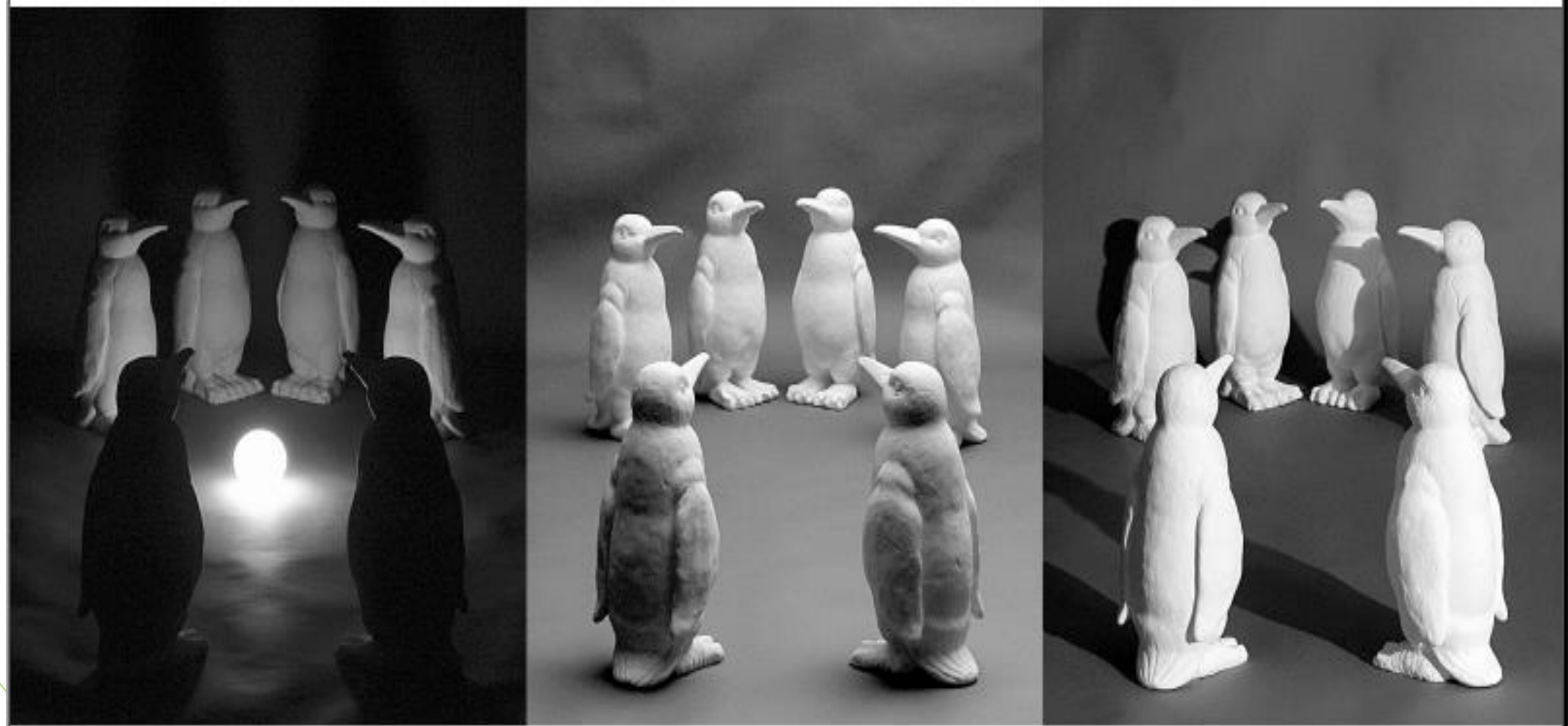
Challenges: viewpoint variation



slide credit: Fei-Fei, Fergus & Torralba



Challenges: illumination



Challenges: scale



slide credit: Fei-Fei, Fergus & Torralba

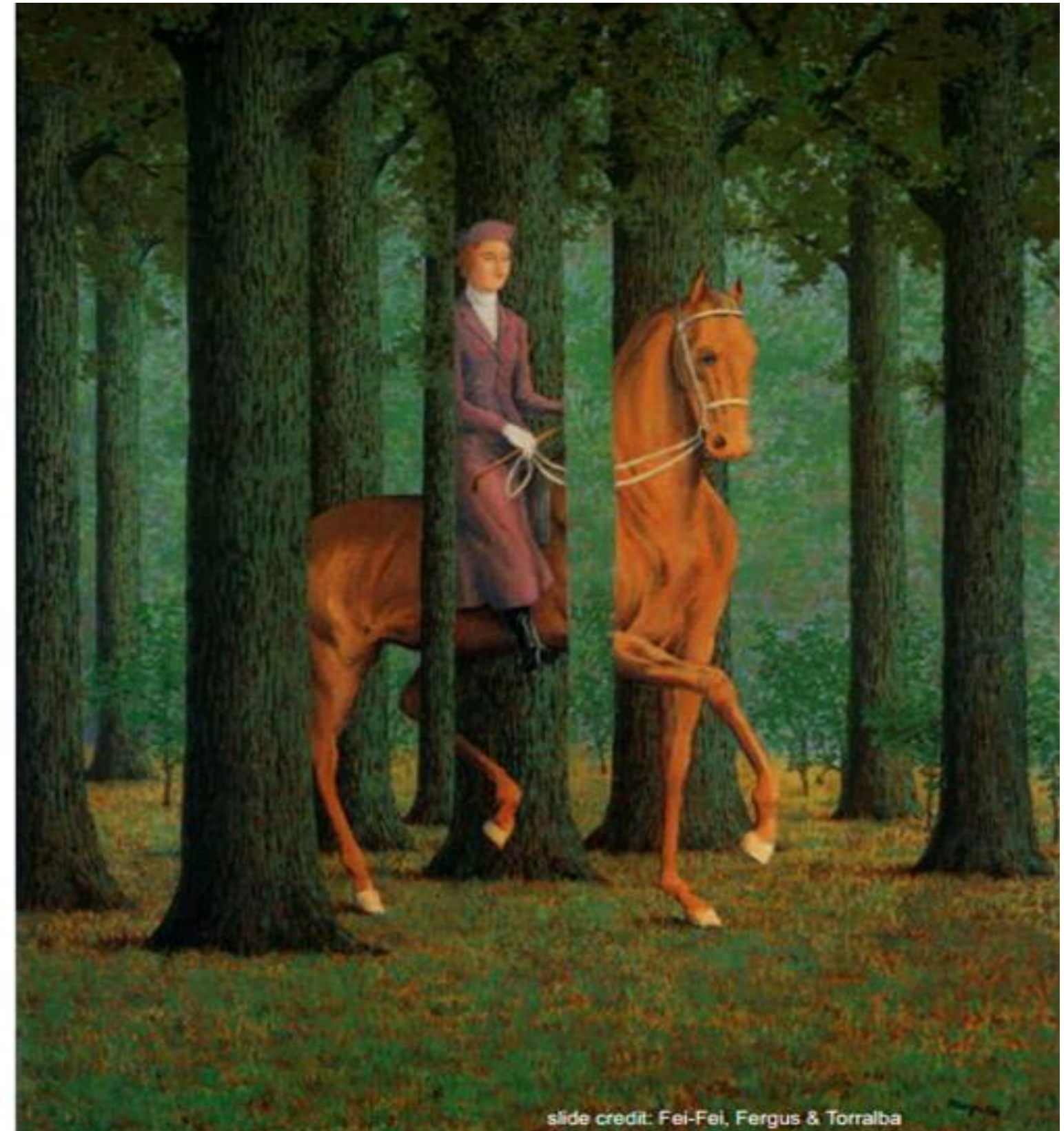
Computer Vision



Challenges: deformation



Challenges: occlusion



Magritte, 1957

slide credit: Fei-Fei, Fergus & Torralba

Computer Vision



Challenges: background clutter



Kilmeny Niland. 1995

Computer Vision



Challenges: object intra-class variation



slide credit: Fei-Fei, Fergus & Torralba

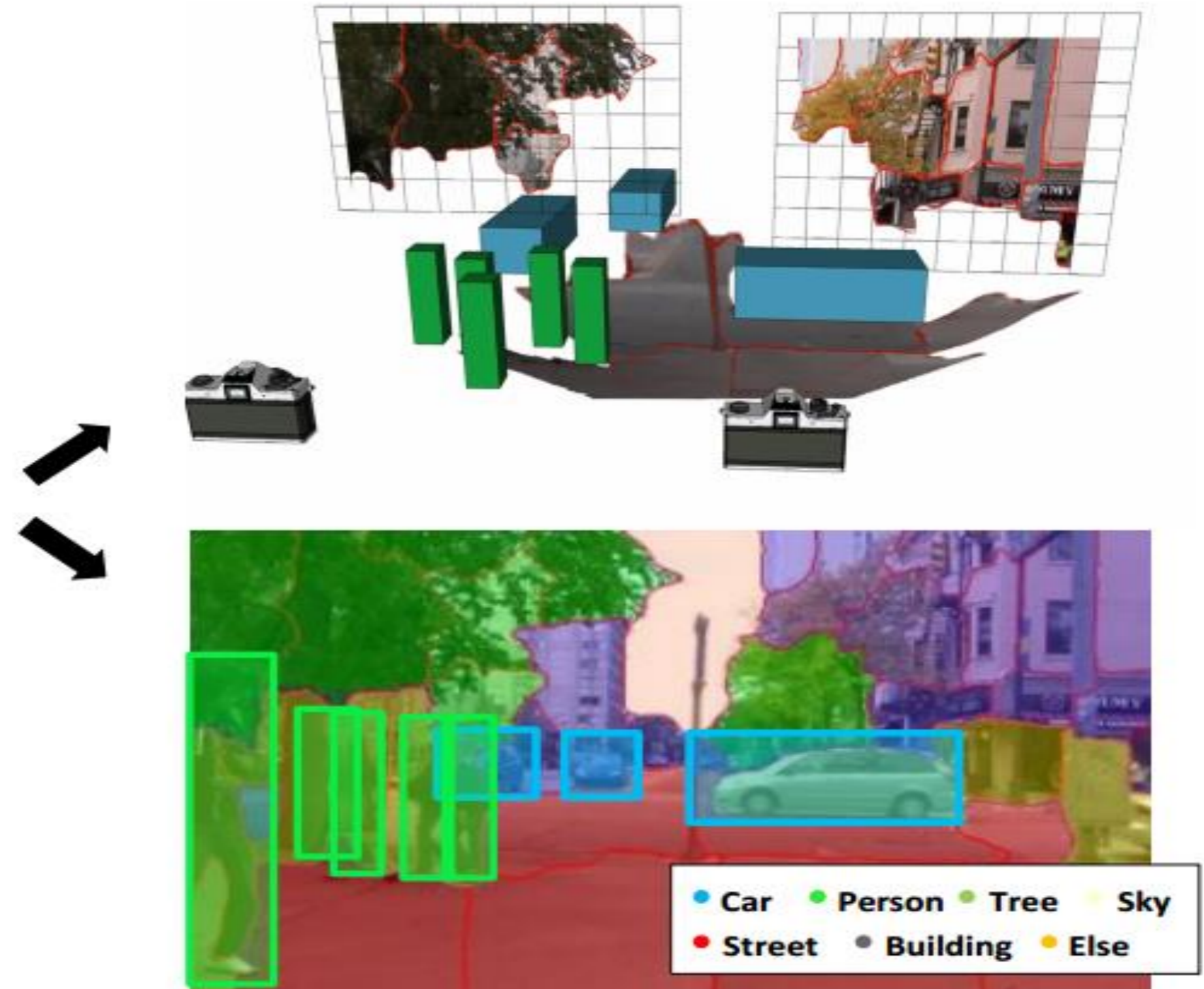




~10,000 to 30,000

Joint Reconstruction and Recognition

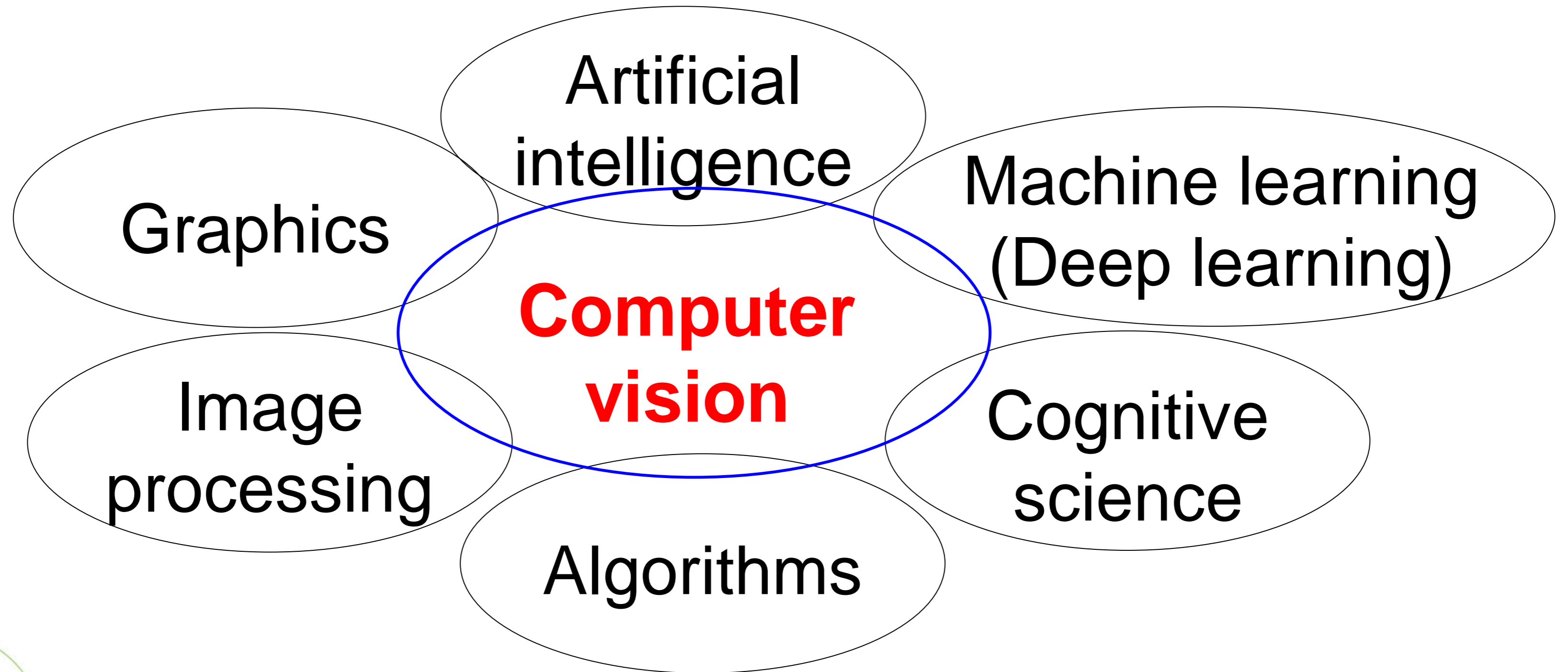
Input images



Computer Vision



Related disciplines



Resources (Web)

Course Website.

- Course Info. / Content .
- Lecture slides.
- Programming Assignments.
- Home Exercises.
- Exams.

- Please register, signup or send me an email on
 - muhammad.umer@cs.uol.edu.pk
 - Please use a single email address for all interaction and be sure to check it daily for updates.



<https://piazza.com/uol.edu.pk/fall2017/cs4455/home>

▪ **Attendance: CS & IT, UOL Policy (75%)**



To Do

- **Python**

- Install **Anaconda Python** Distribution in Windows or Spyder in Ubuntu Linux on following link: <https://www.anaconda.com/download/>

- **Python Tutorial:**

- Available at: <https://piazza.com/uol.edu.pk/fall2017/cs4455/resources>

- **Reading**

- Required

- https://en.wikipedia.org/wiki/Computer_vision



Acknowledgement

- Mostly Slides taken from “**CS231A: Computer Vision, From 3D Reconstruction to Recognition**” course by Silvio Savarese, Stanford University.

