EE290T : 3D Reconstruction and Recognition

Acknowledgement

Courtesy of Prof. Silvio Savarese.

Introduction

“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 …”

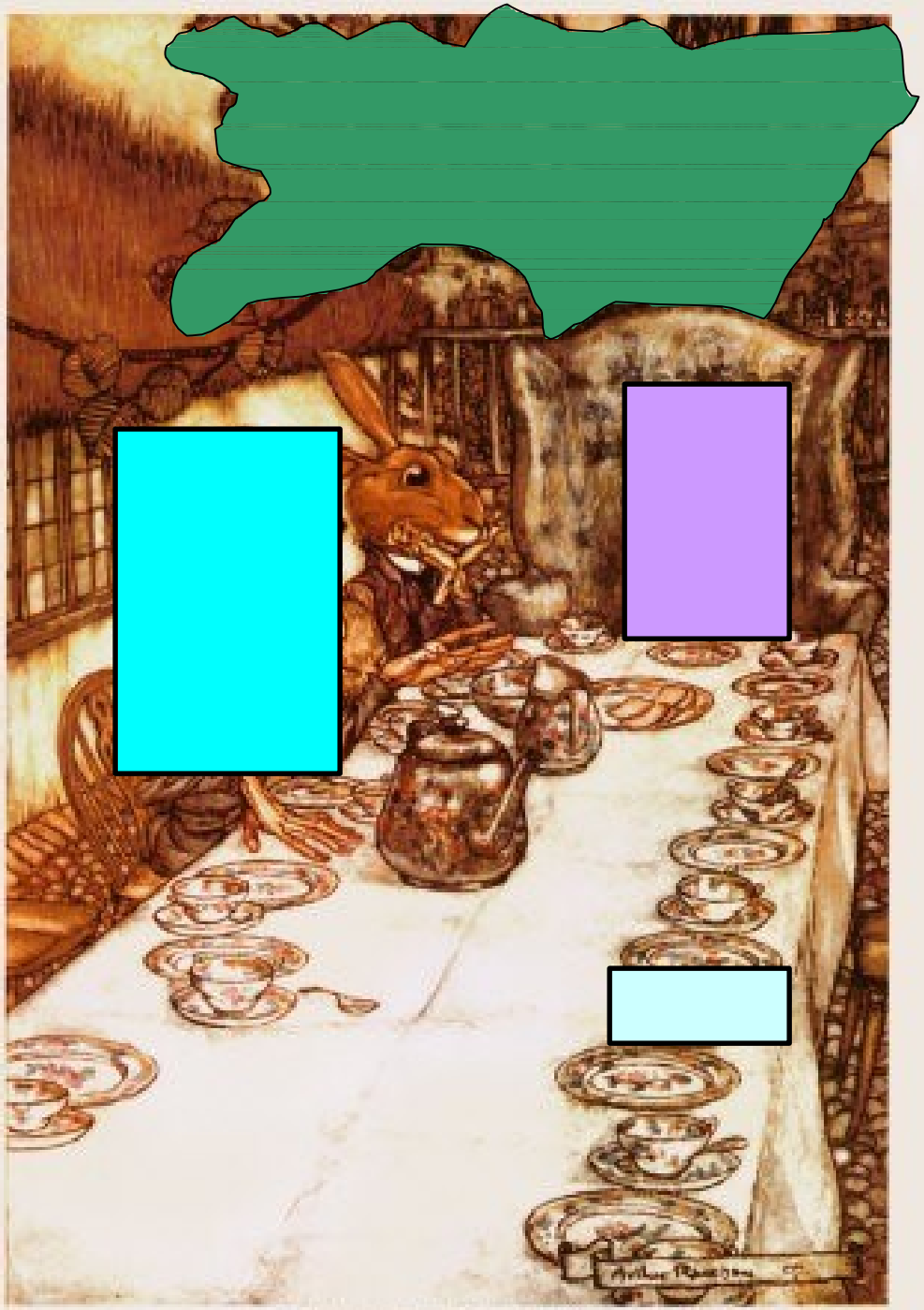
From “A Mad Tea-Party”
Alice's Adventures in Wonderlan
by
Lewis Carroll

Illustration by Arthur Rackham

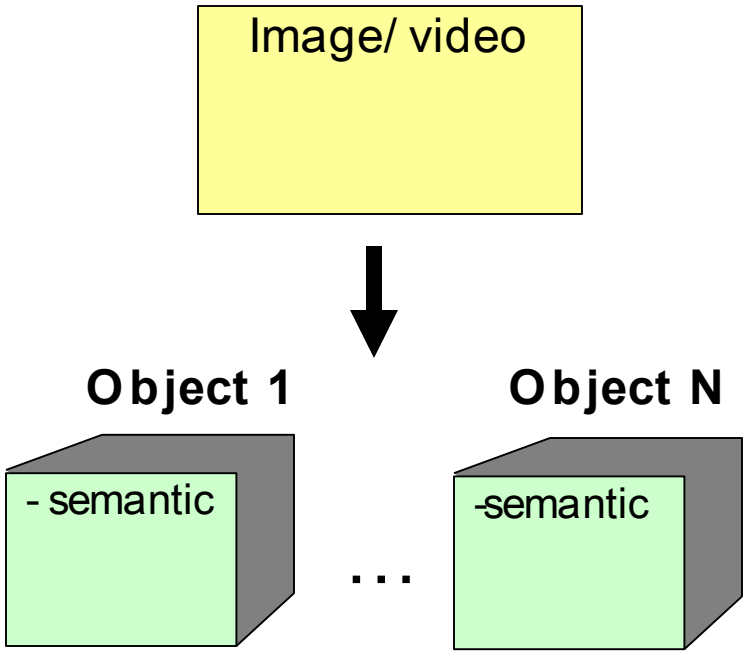
"There was a table set out under a tree in front of the house, and the March Hare and the Hatter were having tea at it."

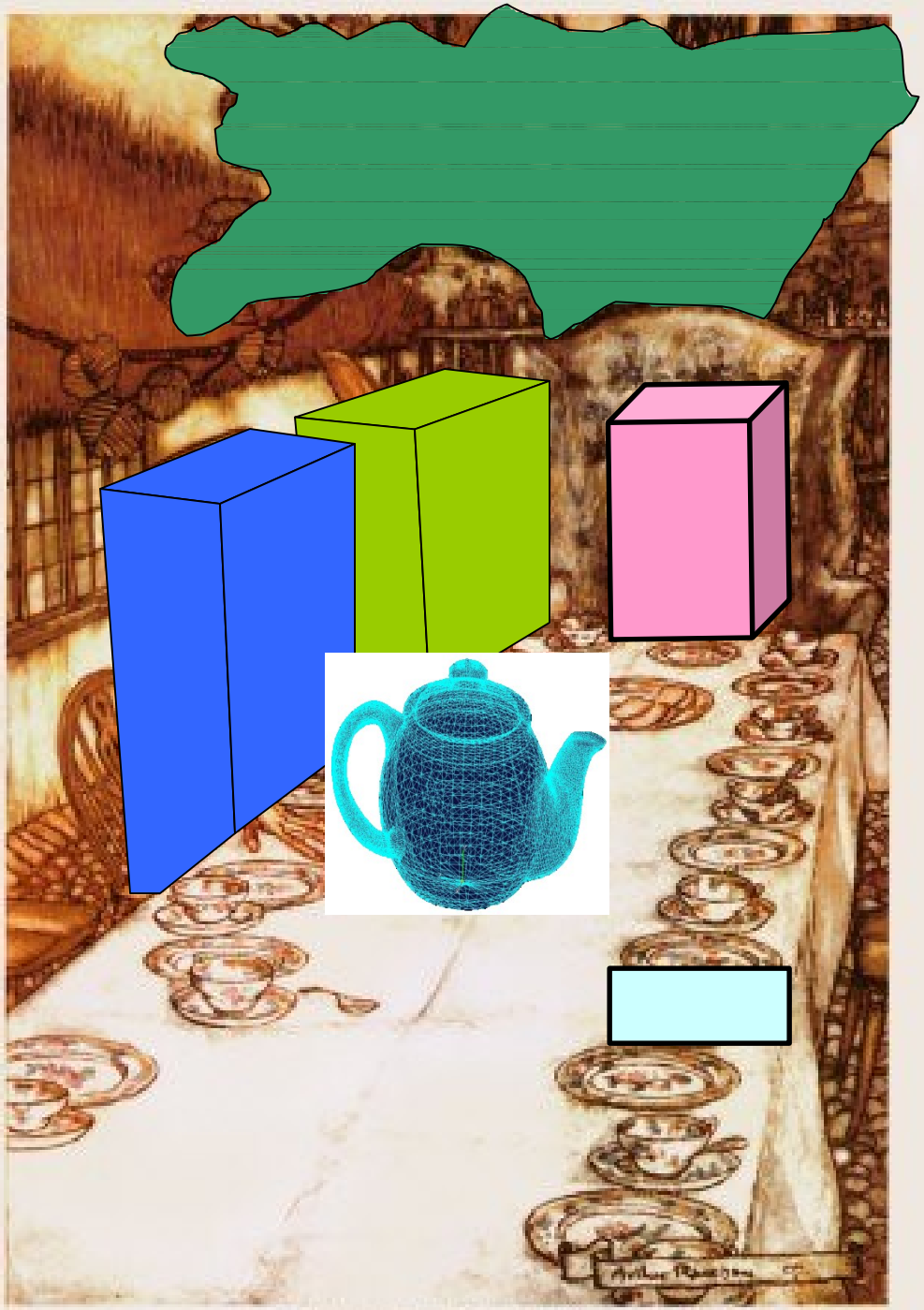
"The table was a large one, but the three were all crowded together at one corner of it …"

From "A Mad Tea-Party"
Alice's Adventures in Wonderlan
by
Lewis Carroll



Computer vision





Computer vision

Image/ video

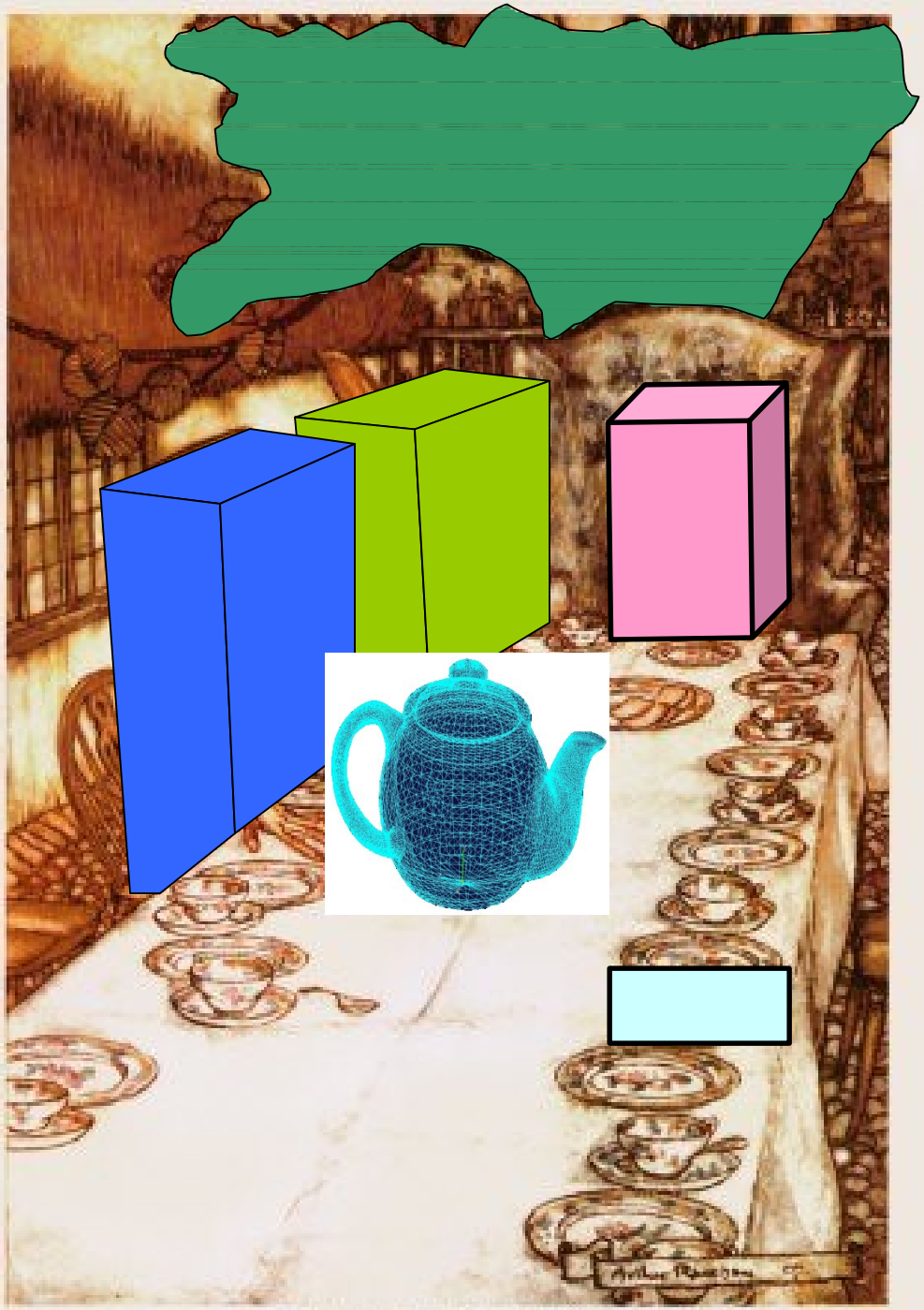
Object 1

- semantic
-geometry

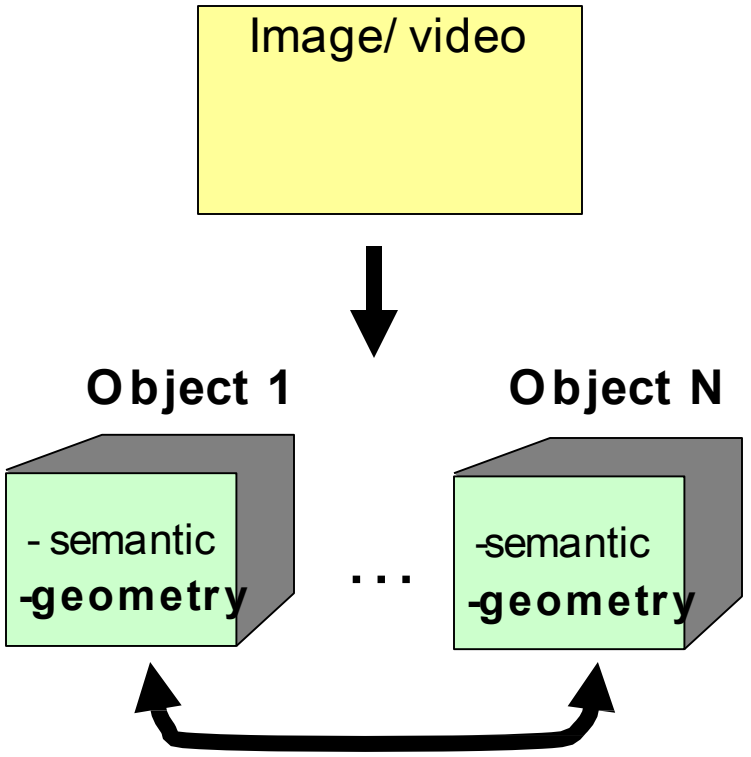
. . .

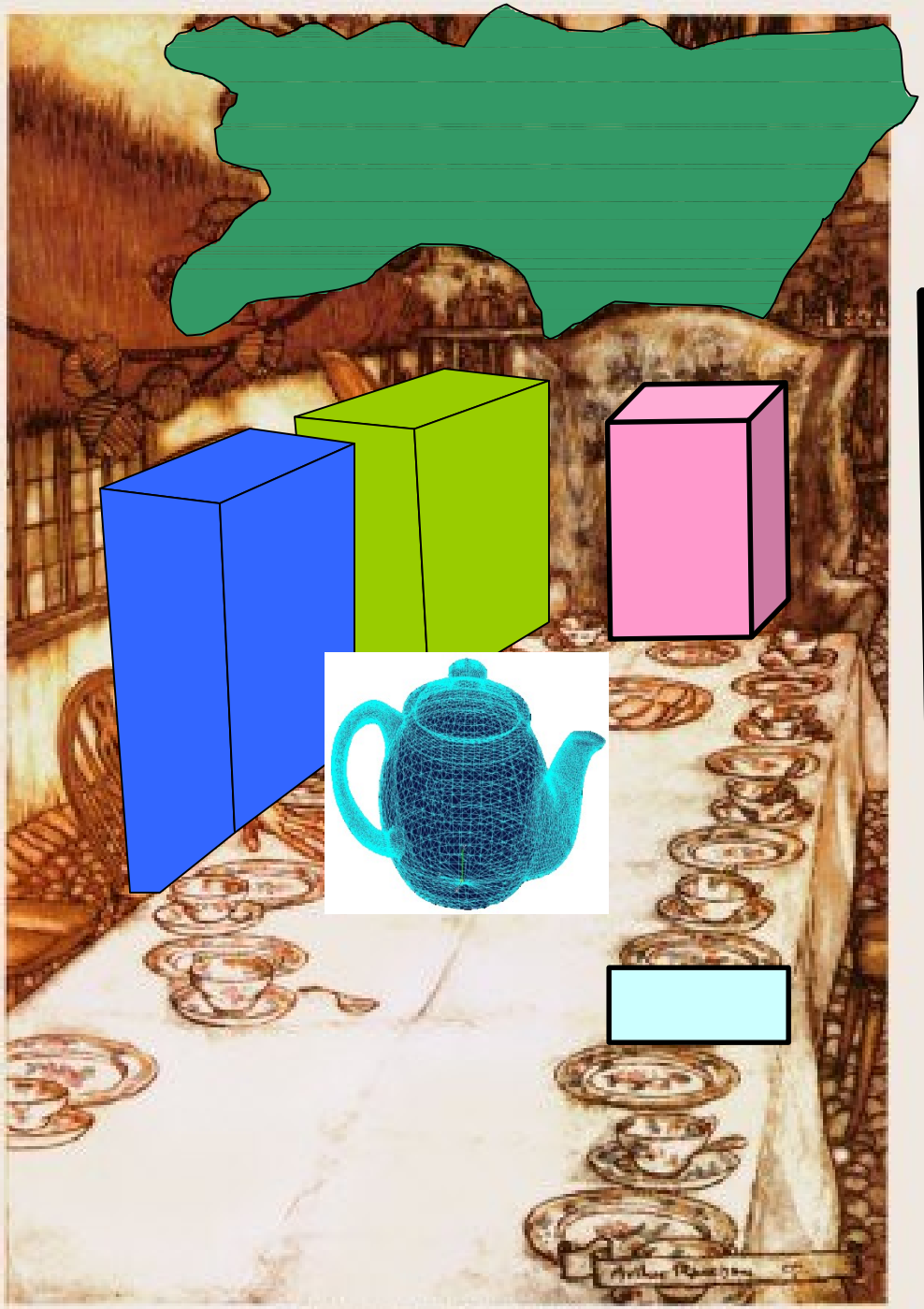
Object N

-semantic
-geometry



Computer vision





Computer vision

Image/ video

Object 1

- semantic
-geometry

...

Object N

-semantic
-geometry

spatial & temporal relations

Scene

-Semantic
- geometry

Computer vision

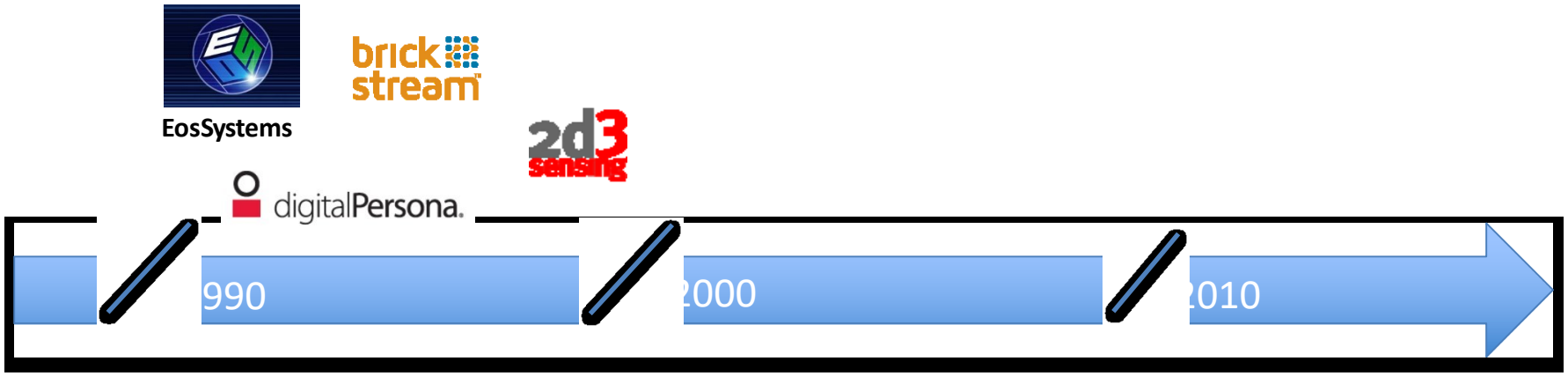
Sensing device

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

digitalPersona.

Augmentation with 3D computer graphics

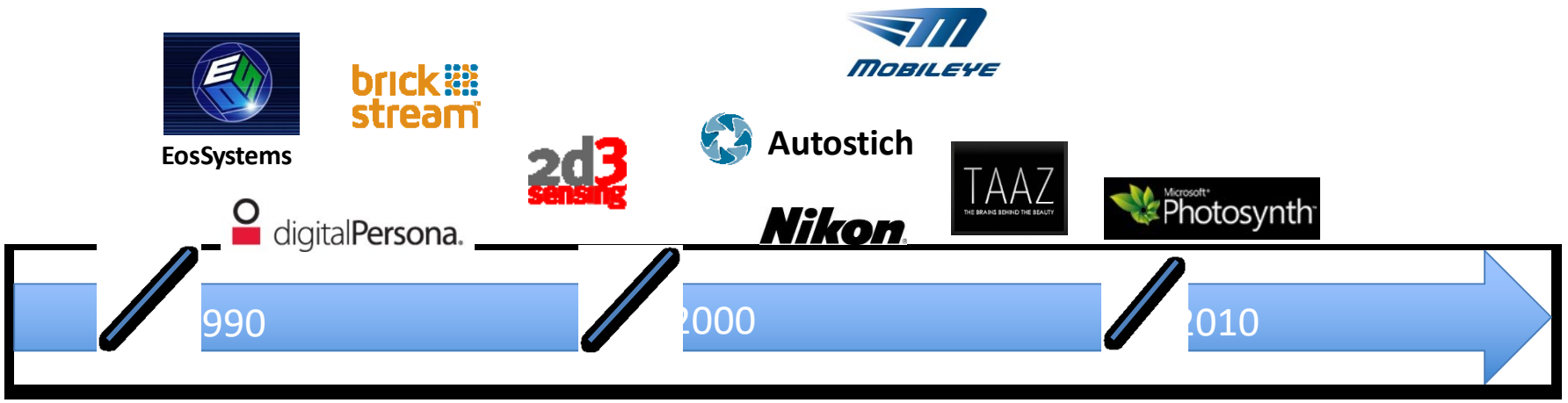
2d3 sensing

3D object prototyping

EosSystems

Photomodeler

Computer vision and Applications



Face detection

BBC NEWS

UK version | International version | About the versions | L

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

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

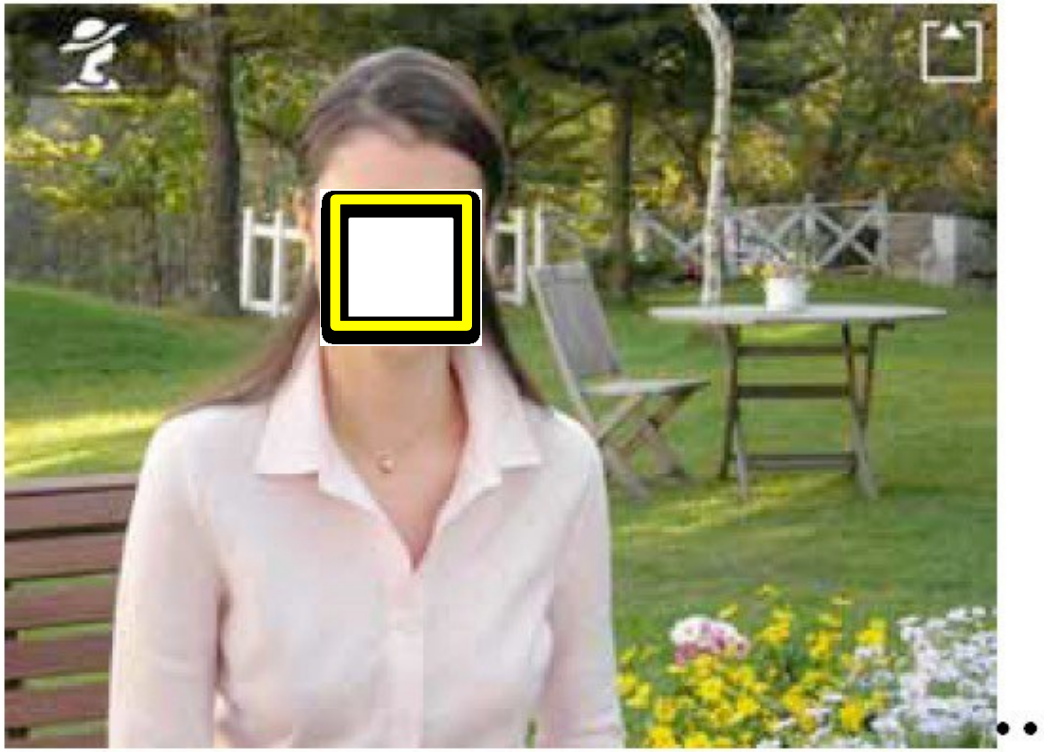
E-mail this to a friend Printable version

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

Face detection



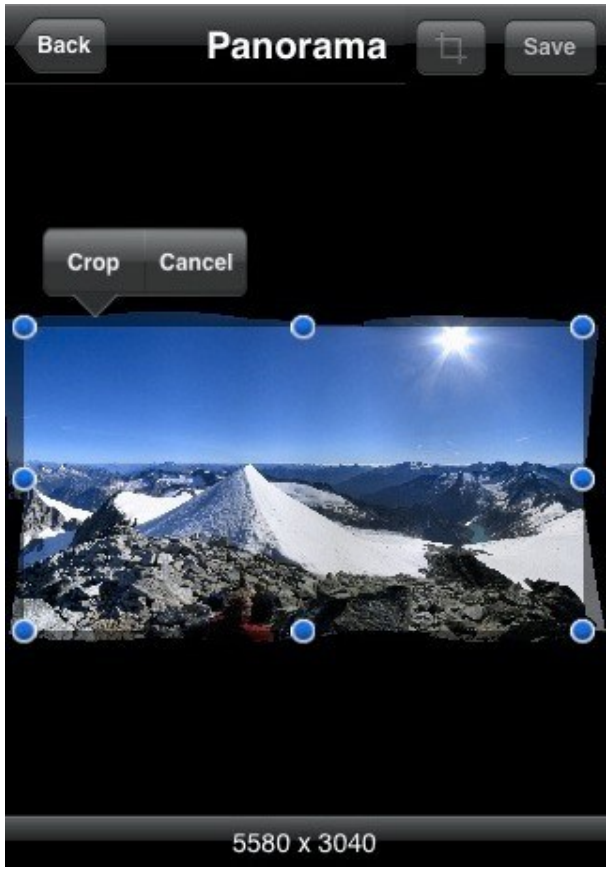
Sample image: Subject as seen on the COOLPIX 5900 camera's color LCD and when using Nikon's Face-priority AF function.

Web applications

TAAZ
THE BRAINS BEHIND THE BEAUTY

Photometria

Panoramic Photography

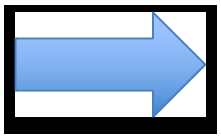


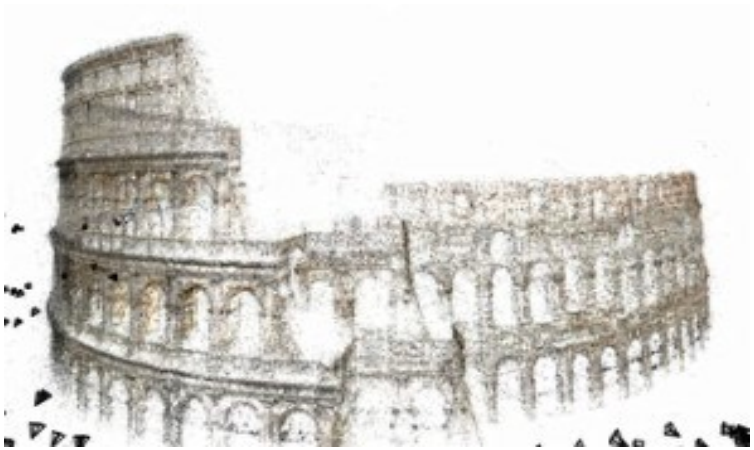
Cloudburst RESEARCH

HumanEyes Print Out of the Box

kolor

3D modeling of landmarks

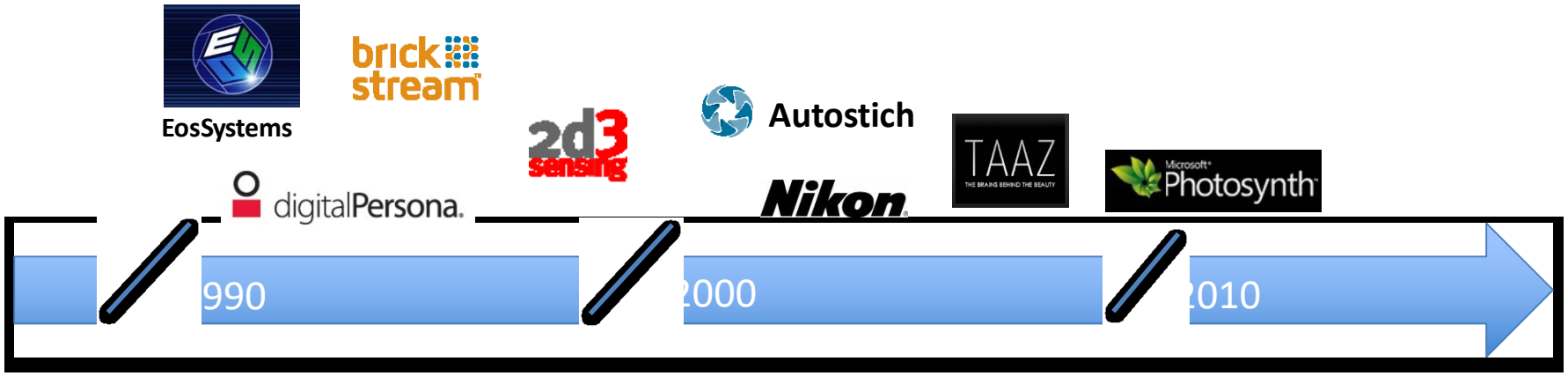




Microsoft Photosynth

Computer vision and Applications

- Efficient SLAM/SFM
- Large scale image repositories
- Deep learning (e.g. ImageNet)



Computer vision and Applications

Project Tango

Mirriad

KINECT

AUGMENT

ebay research labs

A9

MOBILEYE

Google Goggles

kooaba IMAGE RECOGNITION

EosSystems

brick stream

2d3 sensing

Autostich

TAAZ

Microsoft Photosynth

digitalPersona.

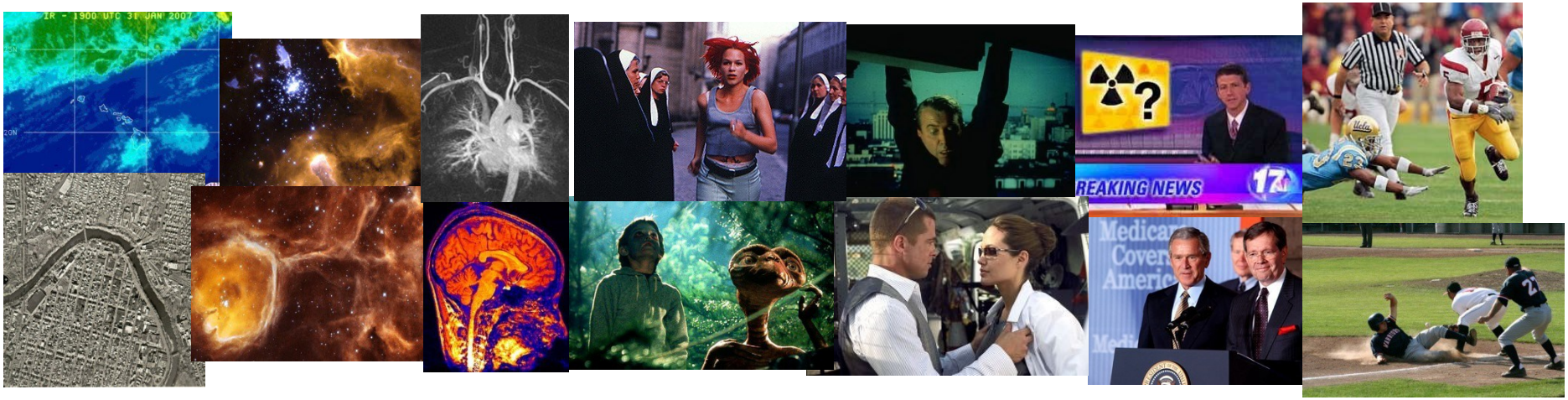
Nikon

990

2000

2010

Image search engines



Google Image Search™

Picasa™

flickr™

webshots beta

bing

YouTube Broadcast Yourself™

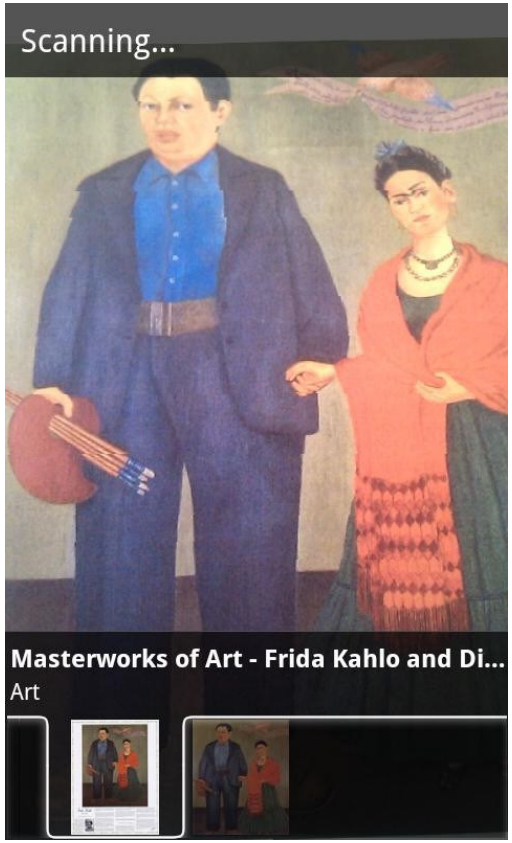
Incogna

LTU technologies LTU

picsearch™

YAHOO!®

Visual search and landmarks recognition



Google Goggles

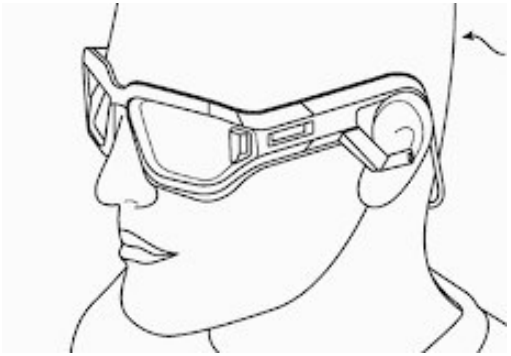
Visual search and landmarks recognition

kooaba
IMAGE RECOGNITION

Augmented reality

Mirriad
Advertising for the Skip Generation

AUGMENT



- Magic leap
- Daqri
- Meta
- Etc…

Motion sensing and gesture recognition

Autonomous navigation and safety

Mobileye: Vision systems in high-end BMW, GM, Volvo models

But also, Toyota, Google, Apple, Tesla, Nissan, Ford, etc….

Source: A. Shashua, S. Seitz

Personal robotics

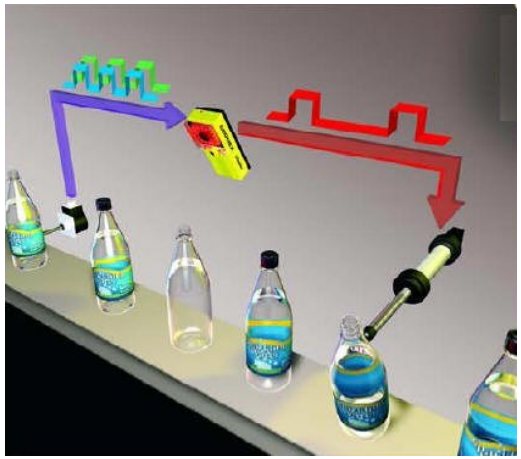
jibo

FIVE ELEMENTS ROBOTICS

BLUE FROG robotics

ECOVACS

Computer vision and Applications



Factory inspection

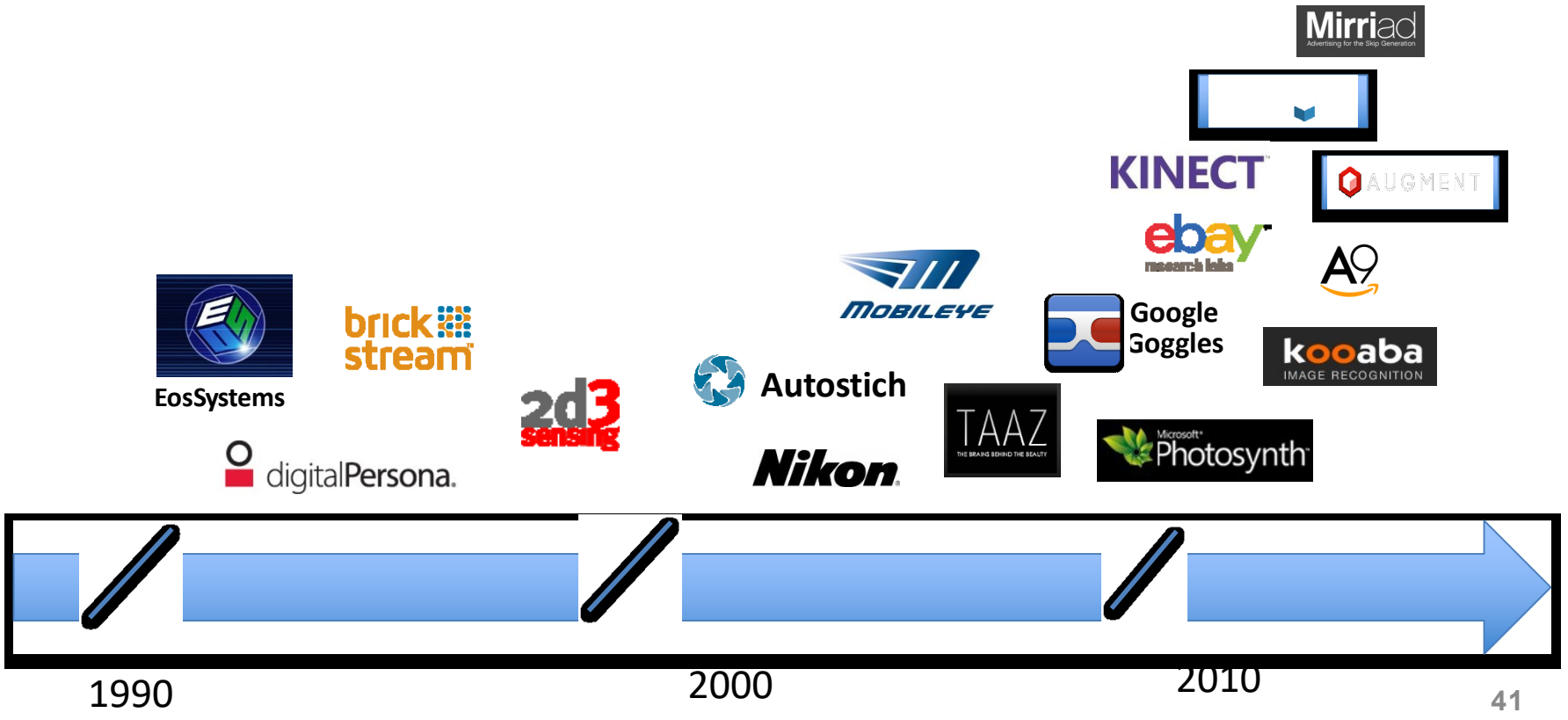
Assistive technologies

Surveillance

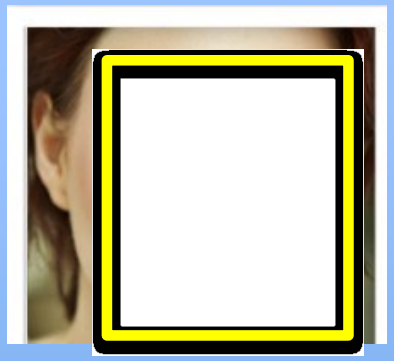
Vision for robotics, space exploration

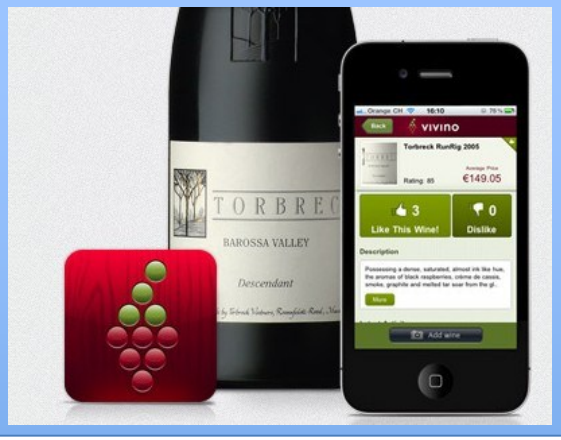
Security

Computer vision and Applications



Computer vision and Applications





3D

Microsoft Photosynth

AUGMENT

2D

digitalPersona.

brickstream

Nikon

TAAZ
THE BRAINS BEHIND THE BEAUTY

Google Goggles

A9

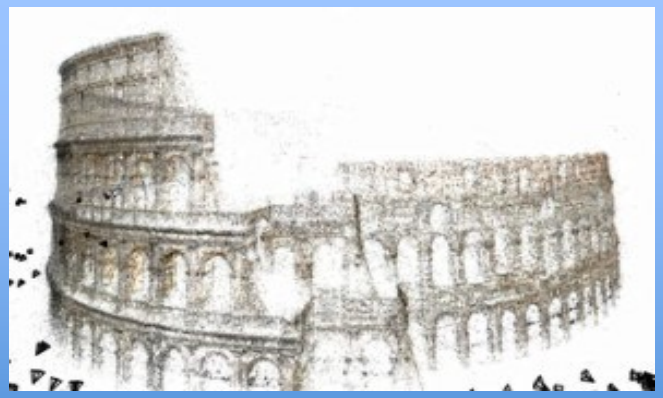
kooaba
IMAGE RECOGNITION

1990

2000

2010

Computer vision and Applications



3D

2D

1990

2000

2010

Current state of computer vision

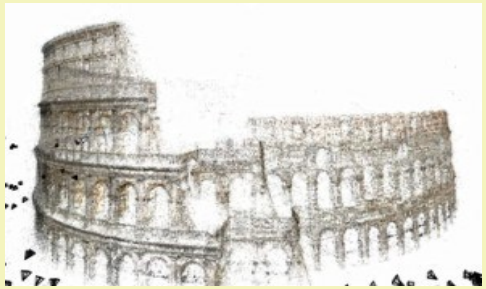
3D Reconstruction

- 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

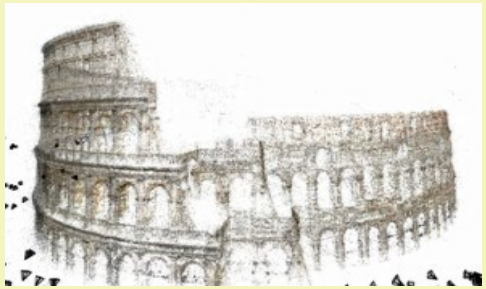
Snavely et al., 06-08

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
Rusinkiewic et al., 02
Nistér, 04
Brown & Lowe, 04
Schindler et al, 04
Lourakis & Argyros, 04
Colombo et al. 05

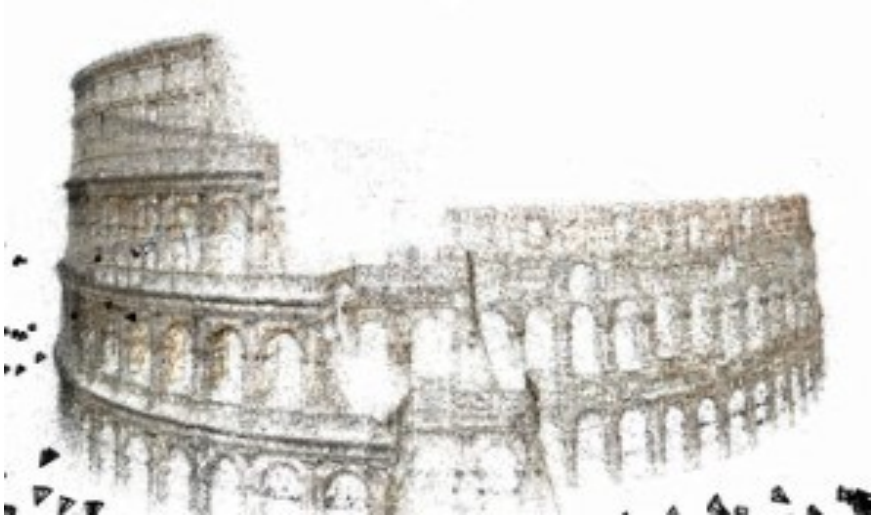
Golparvar-Fard, et al. JAEI 10
Pandey et al. IFAC , 2010
Pandey et al. ICRA 2011
Savarese et al. IJCV 05
Savarese et al. IJCV 06
Microsoft's PhotoSynth
Snavely et al., 06-08
Schindler et al., 08
Agarwal et al., 09
Frahm et al., 10

Current state of computer vision



3D Reconstruction

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



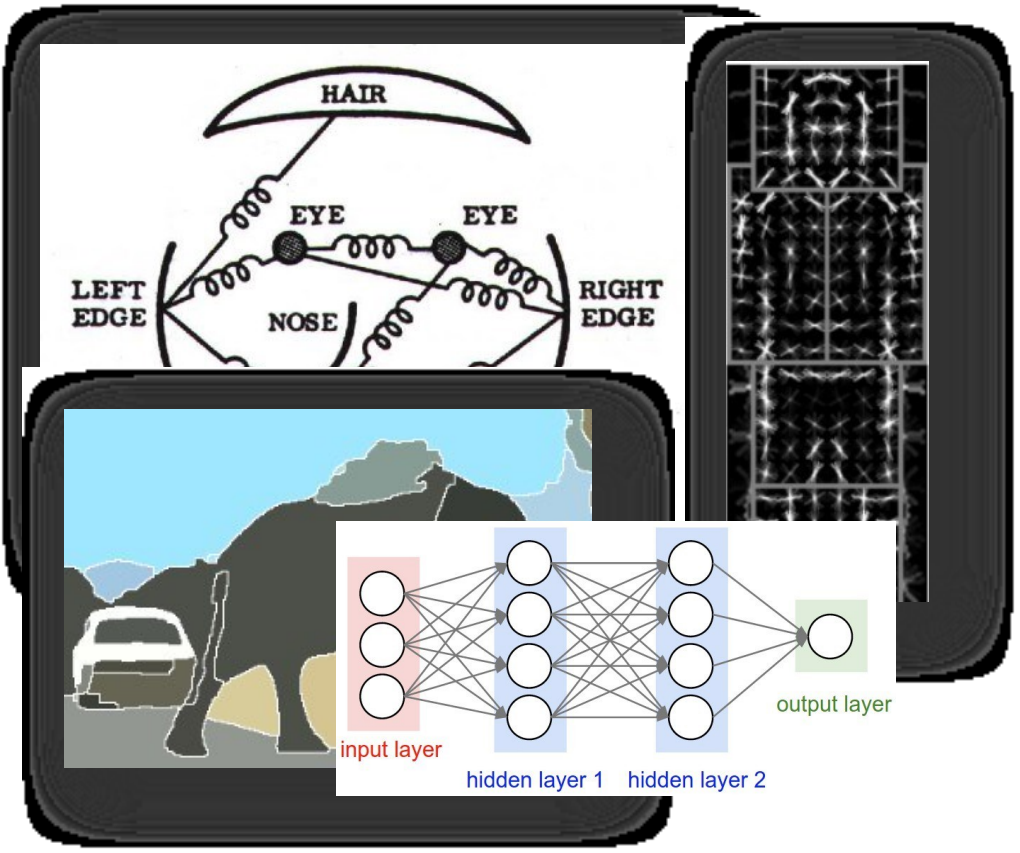
Snavely et al., 06-08

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
Rusinkiewic 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

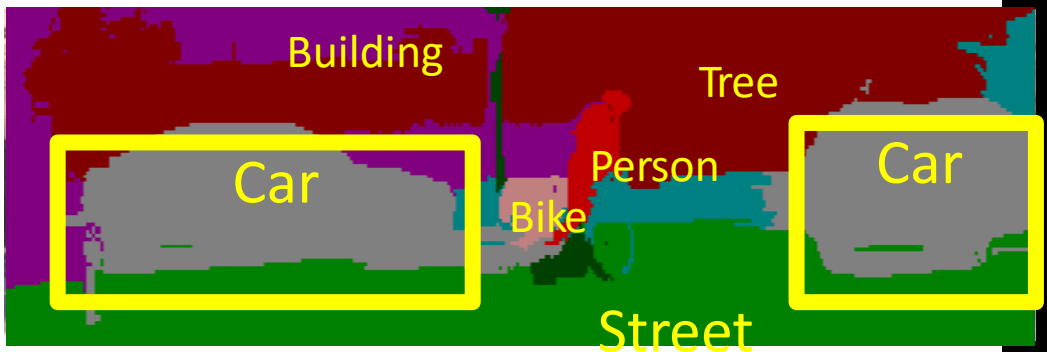
- Object detection
- Texture classification
- Target tracking
- Activity recognition

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
Gonfaus 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!

Course overview

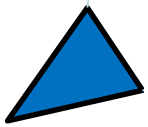
1. Geometry
2. Semantics

Geometry:

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

Camera systems

Establish a mapping from 3D to 2D



How to calibrate a camera

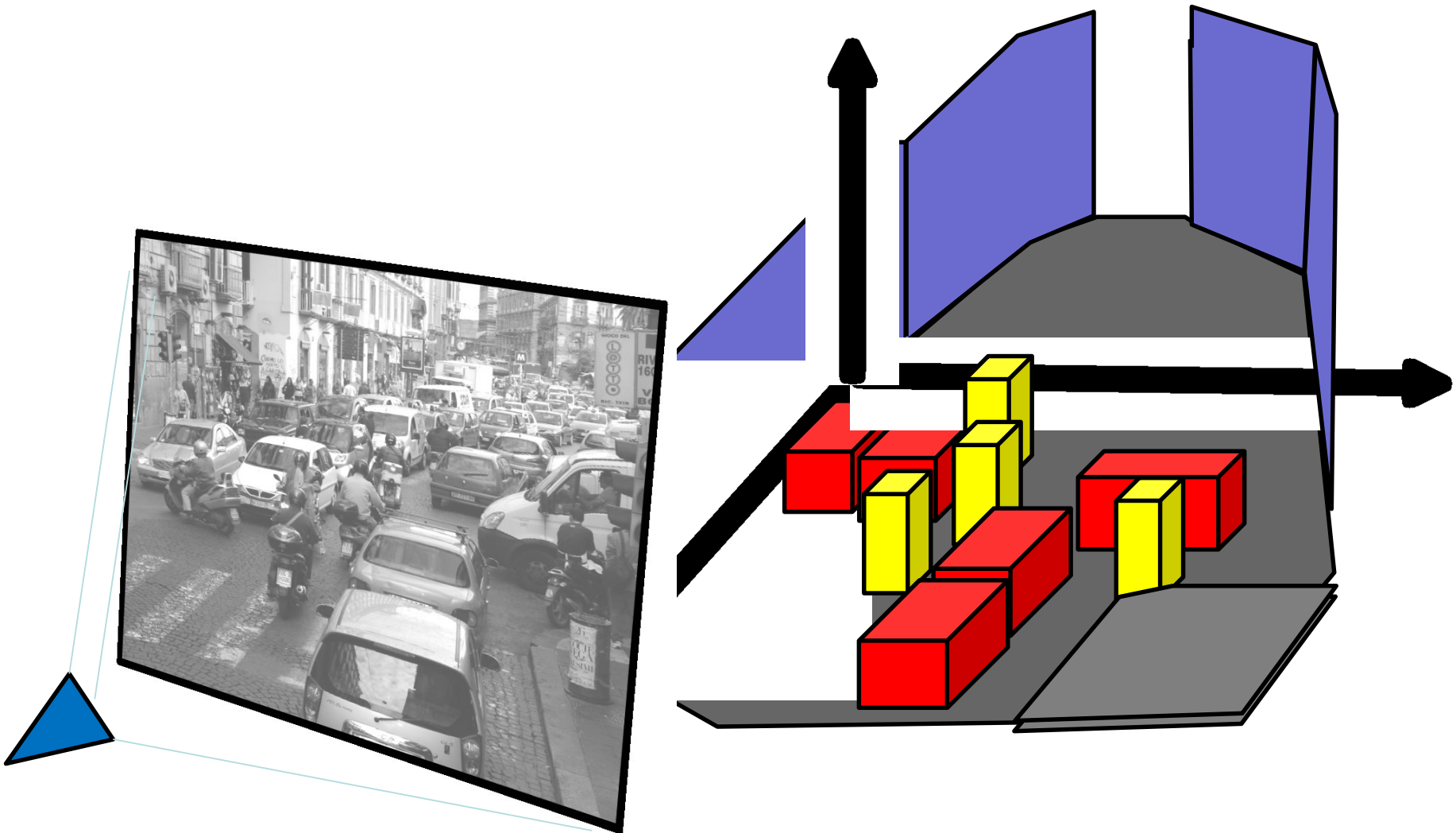
Estimate camera parameters such as pose or focal length

Single view metrology

Estimate 3D properties of the world from a single image

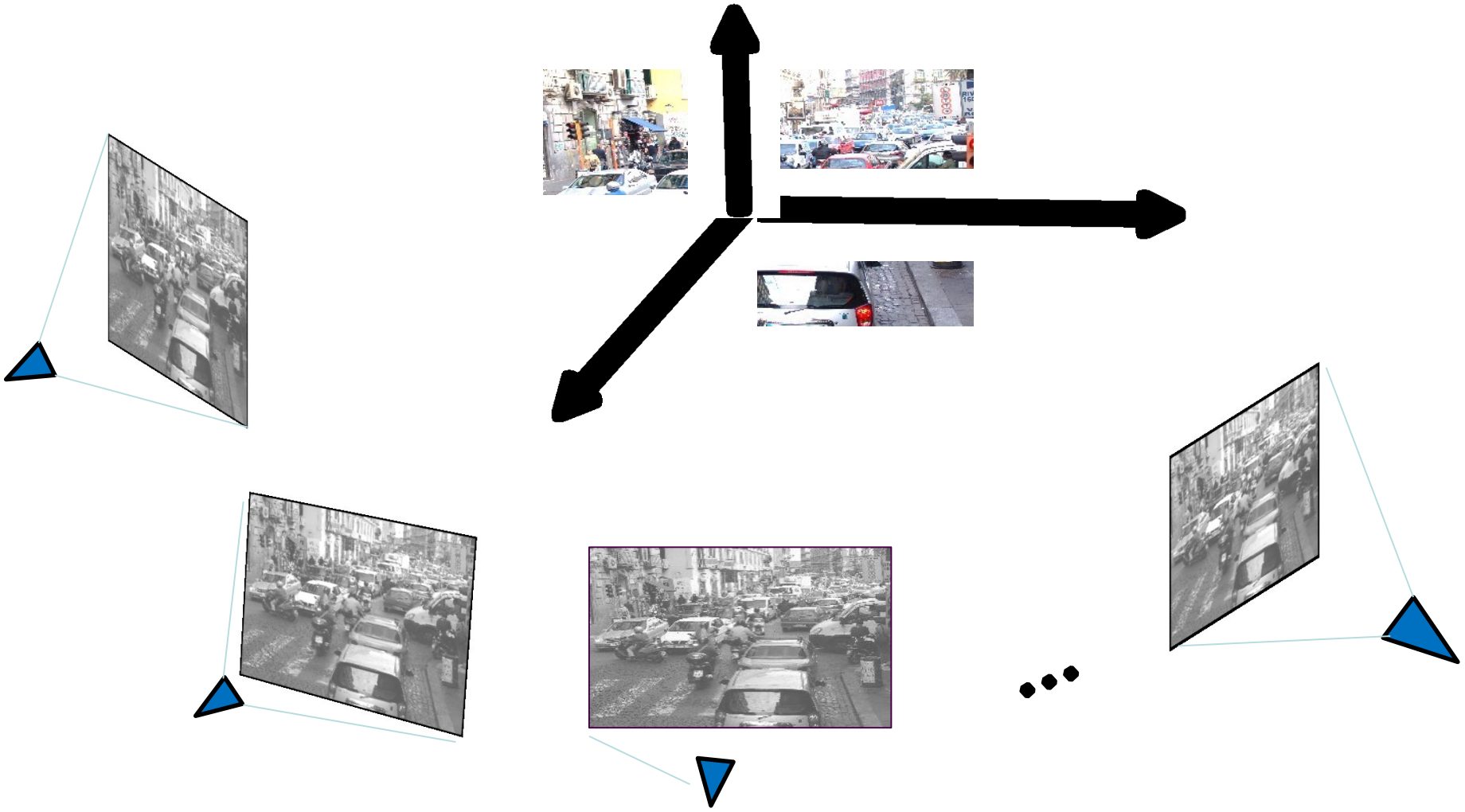
Single view metrology

Estimate 3D properties of the world from a single image

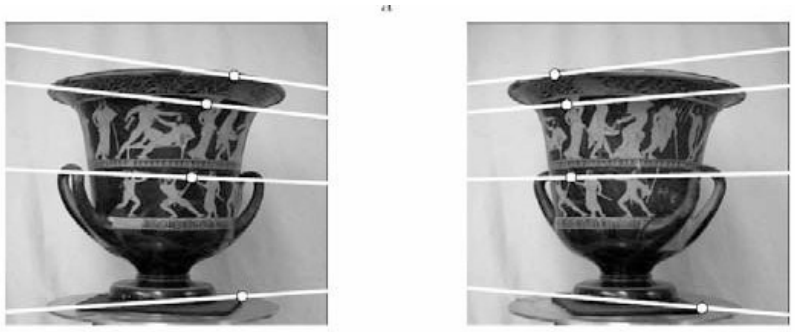


Multiple view geometry

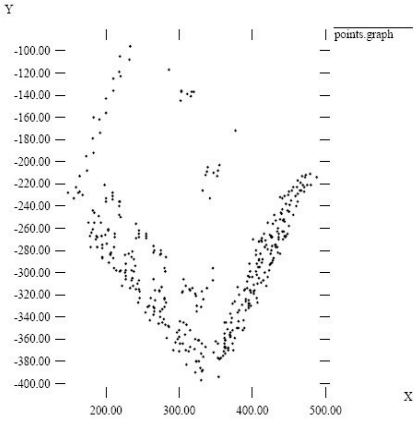
Estimate 3D properties of the world from multiple views

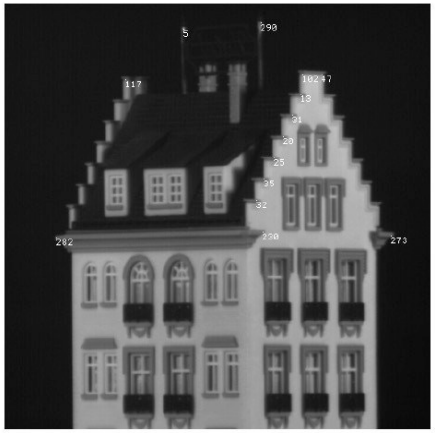


Mathematical tools



Epipolar geometry



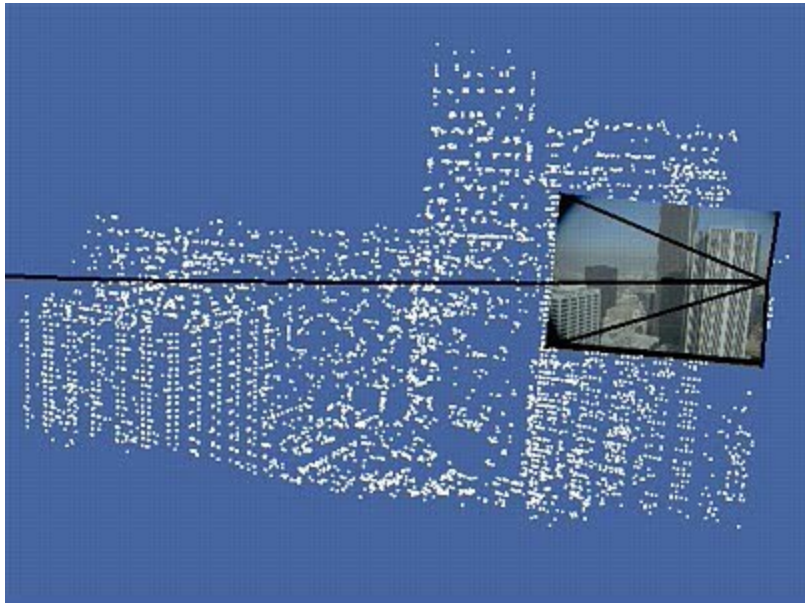


Tomasi & Kanade (1993)

Драконъ, видимый подъ различными углами зрѣнія
По гравюрѣ на мѣди изъ „Oculus artificialis teledioptricus“ Цана. 1702 года.

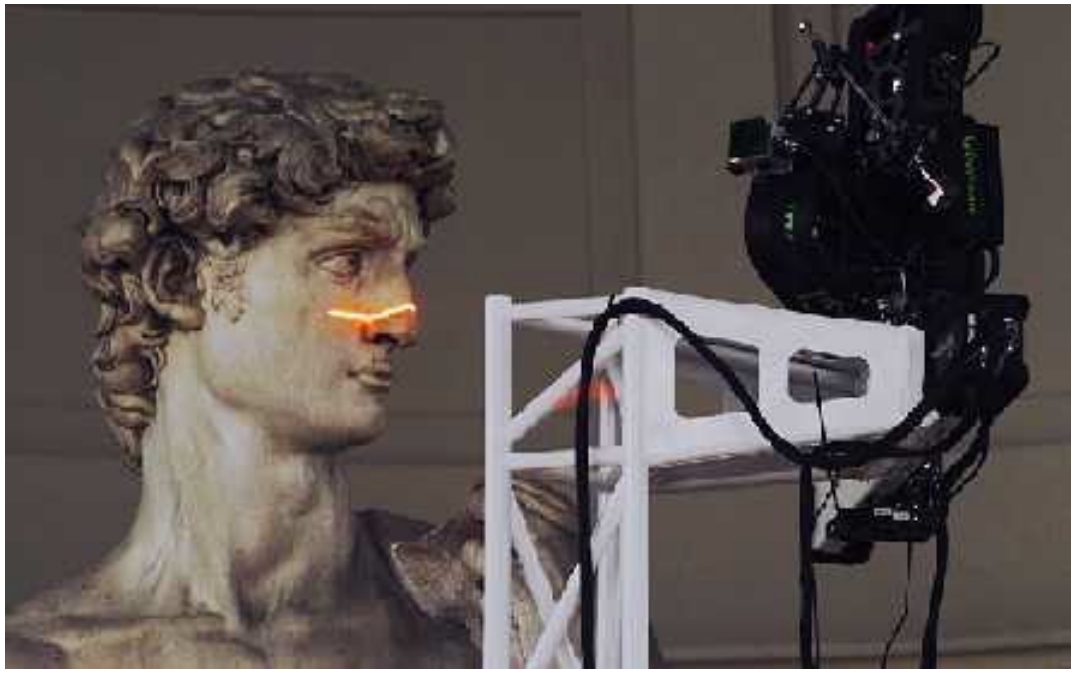
Photoconsistency

Structure from motion



Courtesy of Oxford **Visual Geometry Group**

Structure lighting and volumetric stereo



Scanning Michelangelo's *"The David"*

- The Digital Michelangelo Project
  - http://graphics.stanford.edu/projects/mich/
- 2 BILLION polygons, accuracy to .29mm

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 (pose, size, shape…)

Object recognition and categorization

Classification:

Is this an forest?

Classification:

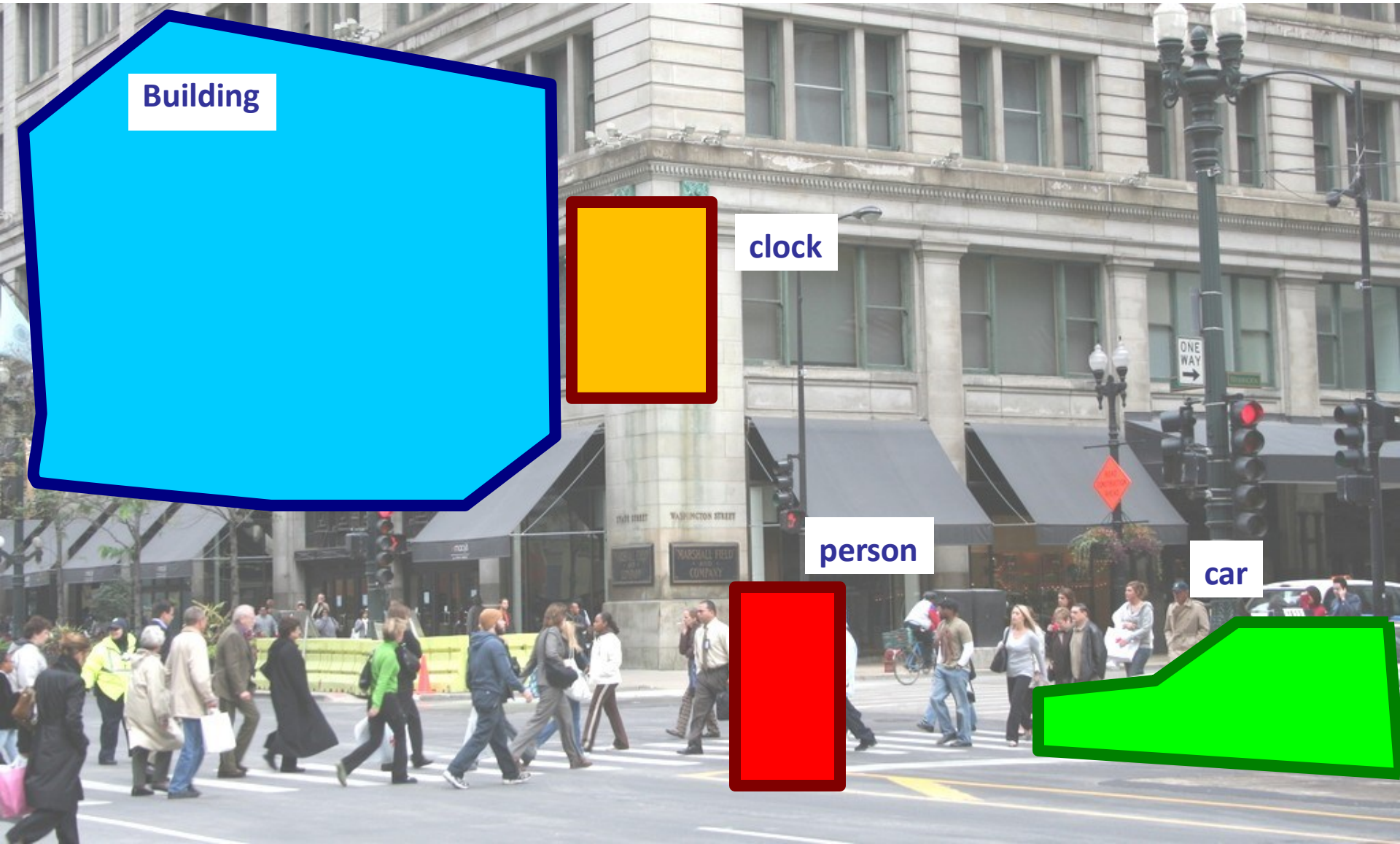
Does this image contain a building? [yes/no]

Detection:

Does this image contain a car? [where?]

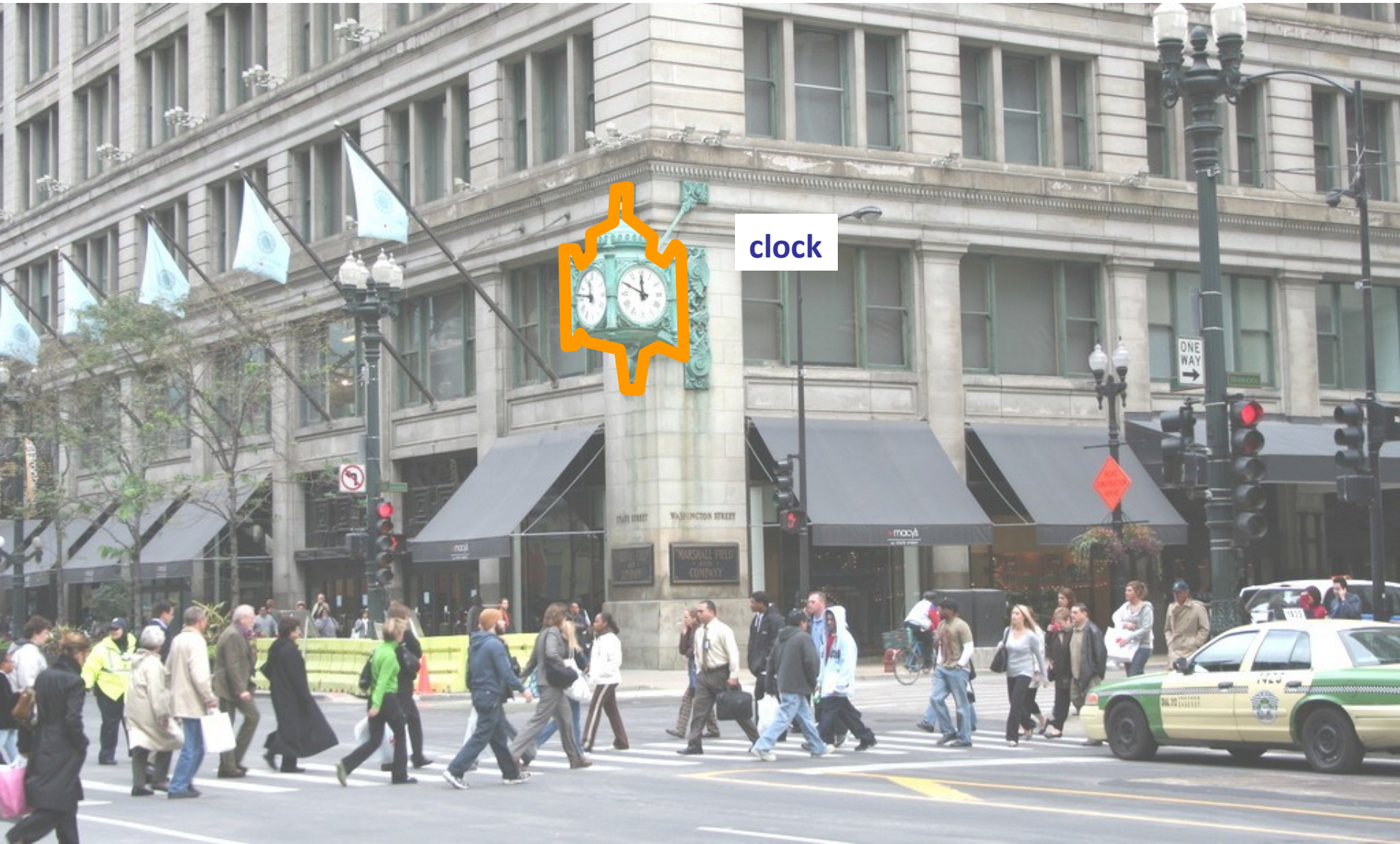
Detection:

Which objects do this image contain? [where?]



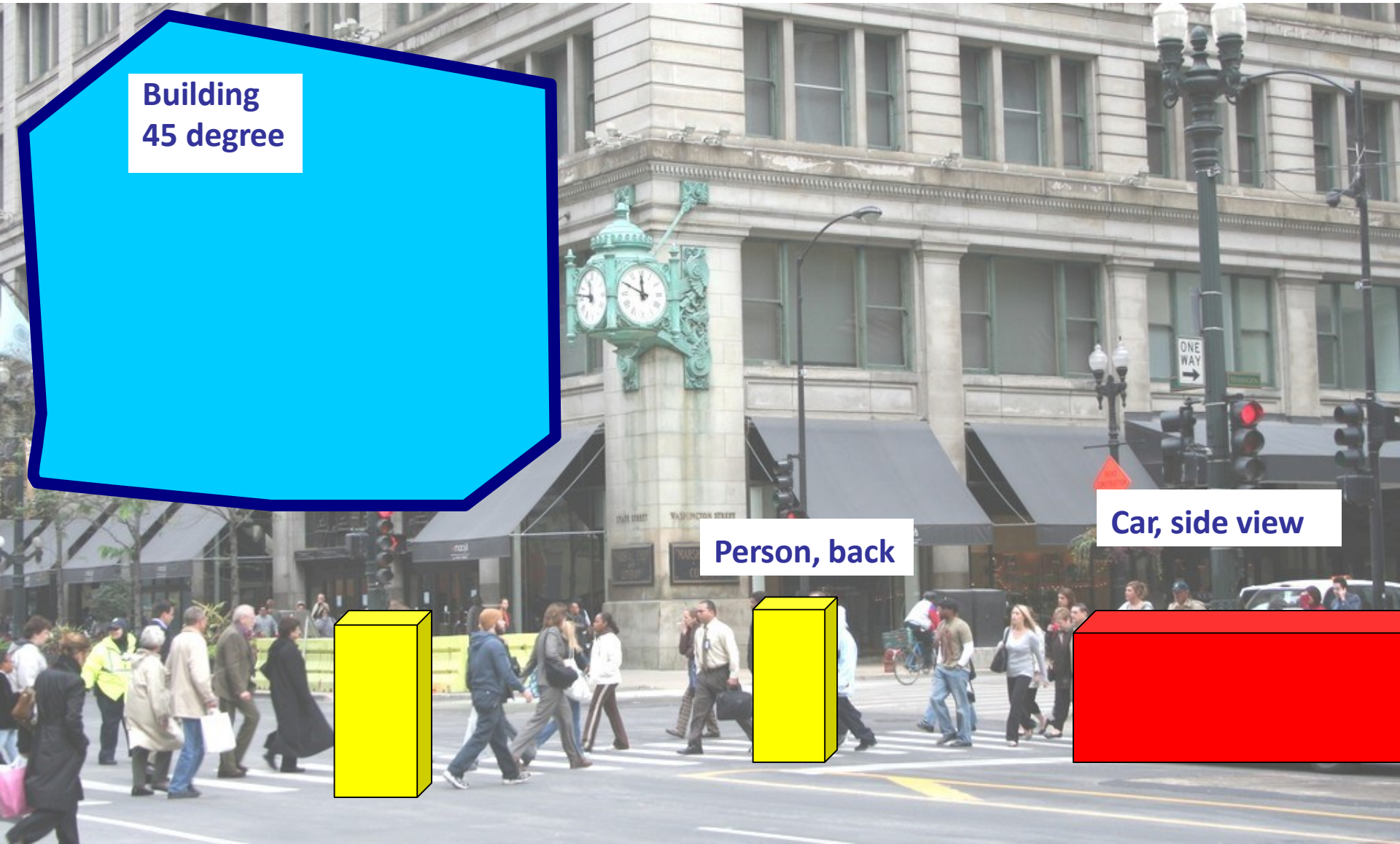
Detection:

Accurate localization (segmentation)



Detection:

Estimating 3D geometrical properties



Challenges: viewpoint variation

slide credit: Fei-Fei, Fergus & Torralba

Challenges: illumination

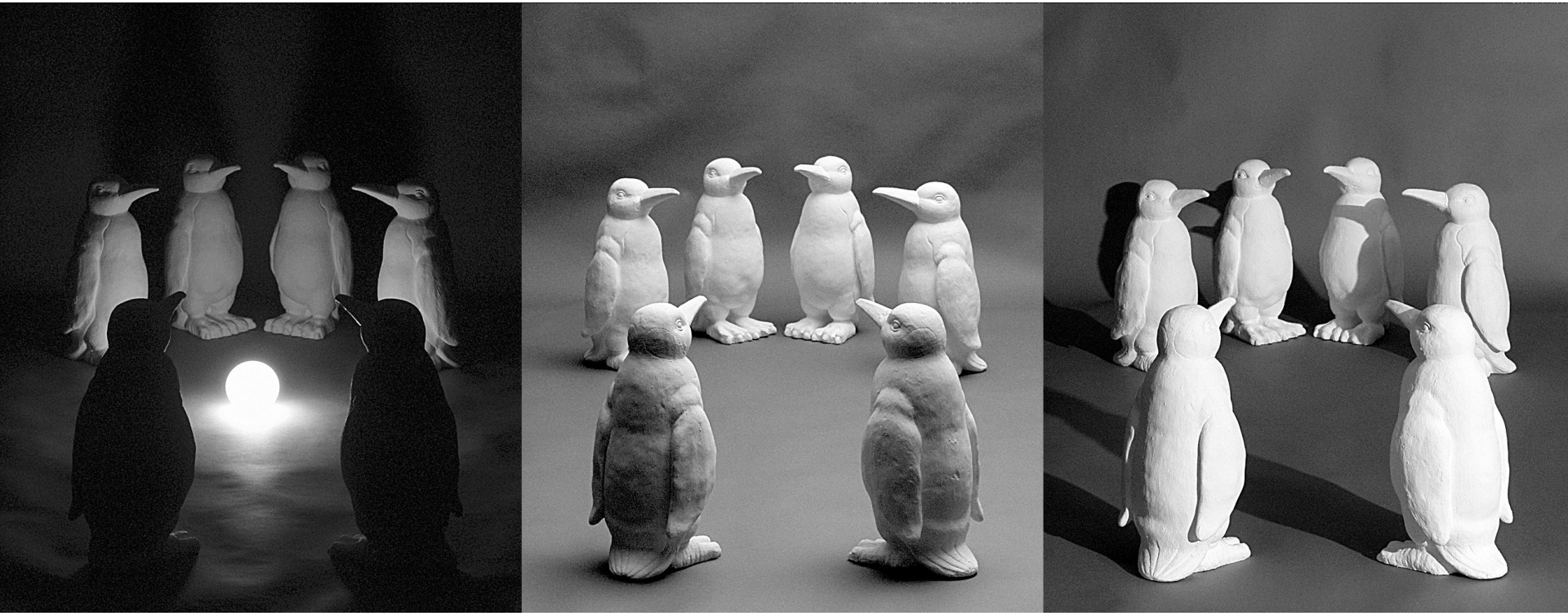


image credit: J. Koenderink

Challenges: scale

slide credit: Fei-Fei, Fergus & Torralba

Challenges: deformation

Challenges: occlusion

Magritte, 1957

slide credit: Fei-Fei, Fergus & Torralba

Challenges: background clutter

Kilmeny Niland. 1995

Challenges: object intra-class variation

slide credit: Fei-Fei, Fergus & Torralba

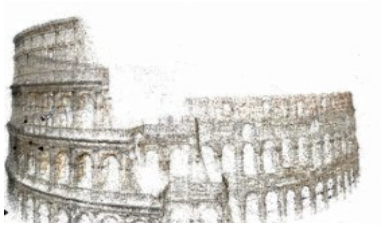

~10,000 to 30,000

Course overview

1. Geometry
2. Semantics

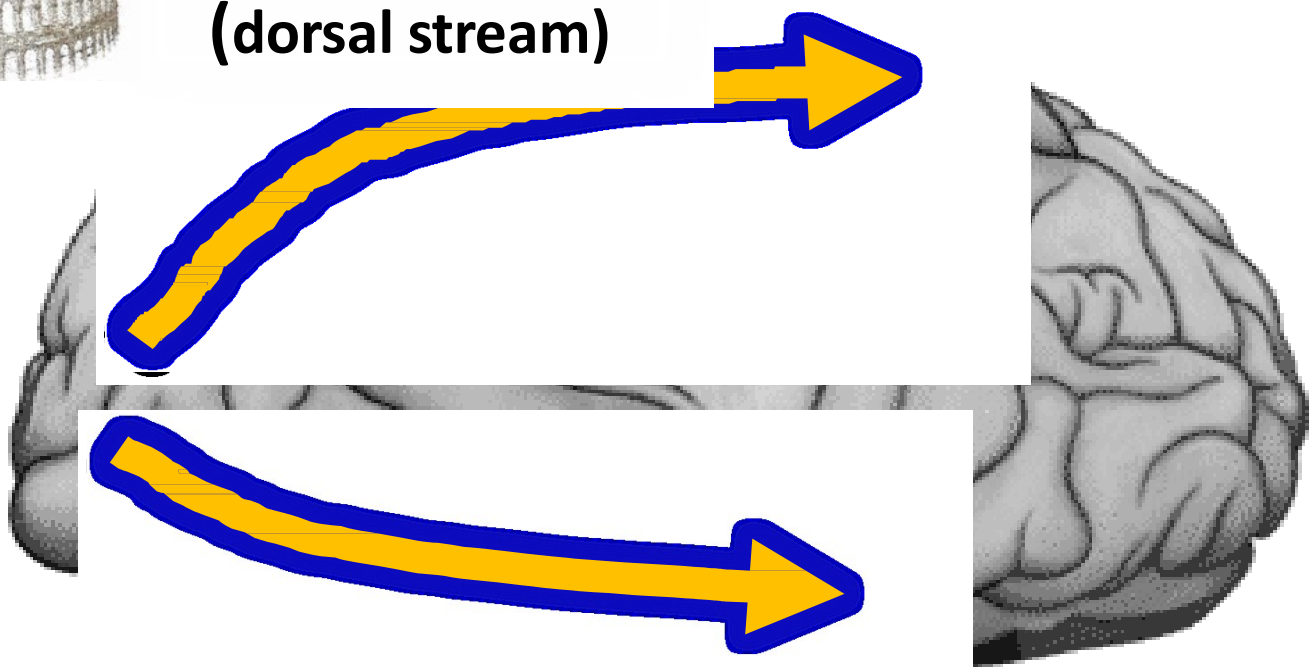
Joint recovery of geometry and semantics!

Visual processing in the brain



where pathway
(dorsal stream)

The ventral stream (also known as the "what pathway") is involved with object and visual identification and recognition.

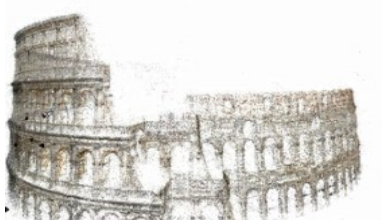


V1

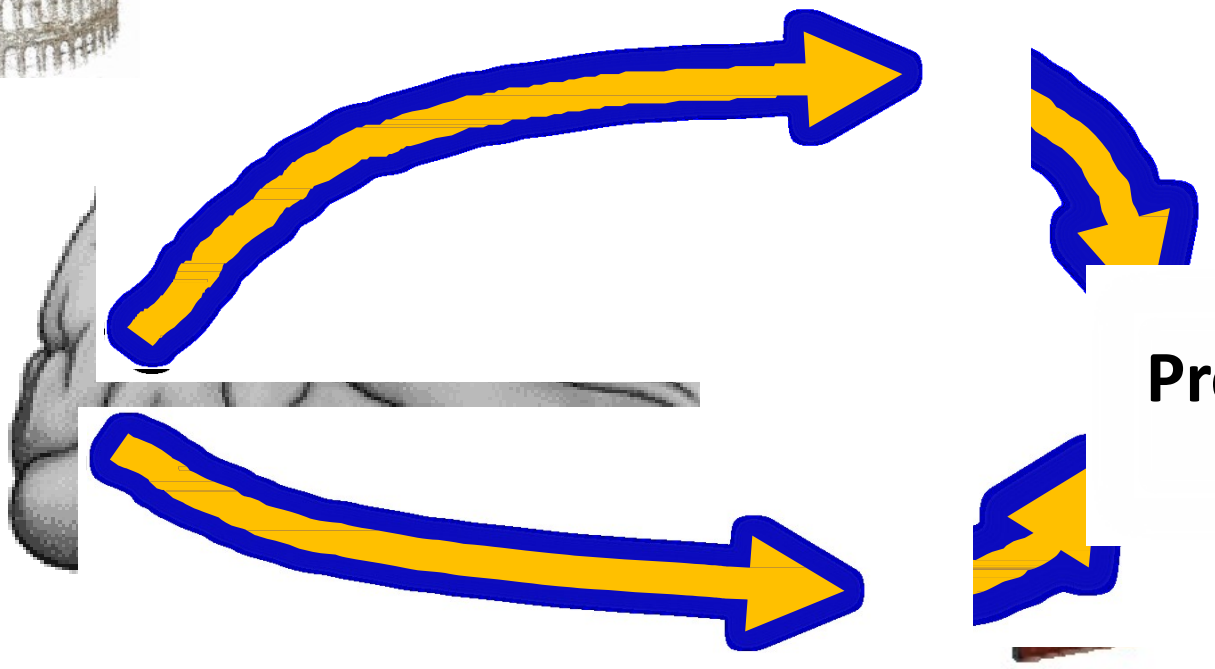
what pathway
(ventral stream)

The dorsal stream (or, "where pathway") is involved with processing the object's spatial location relative to the viewer and with speech repetition.

Visual processing in the brain



where pathway



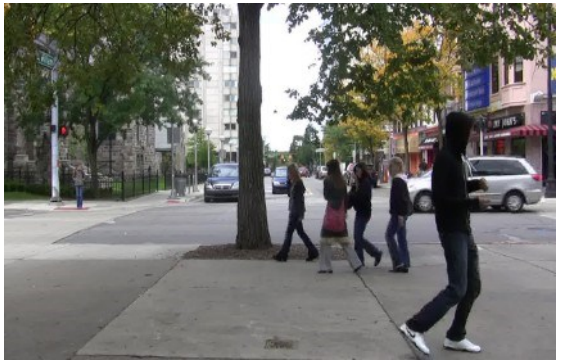
V1

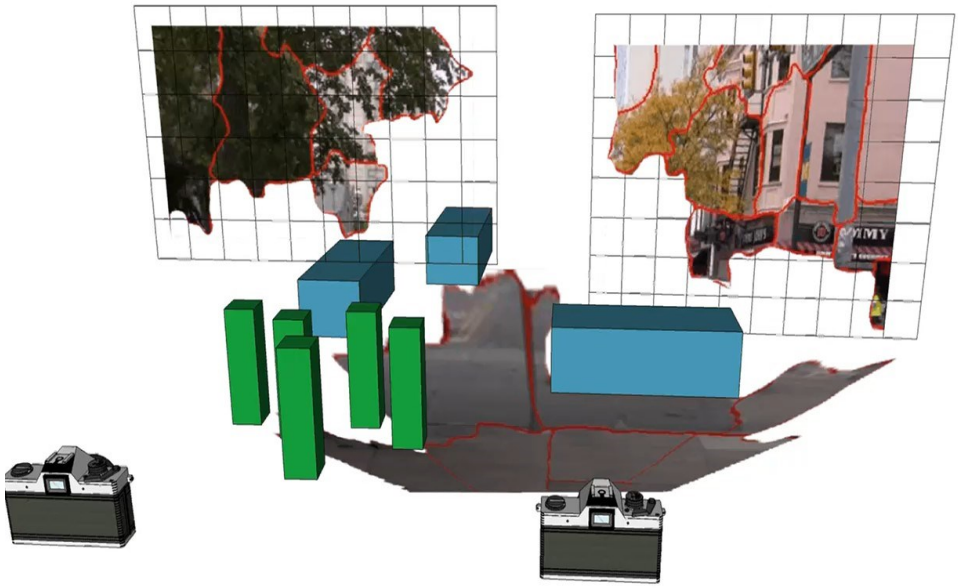
Pre-frontal cortex

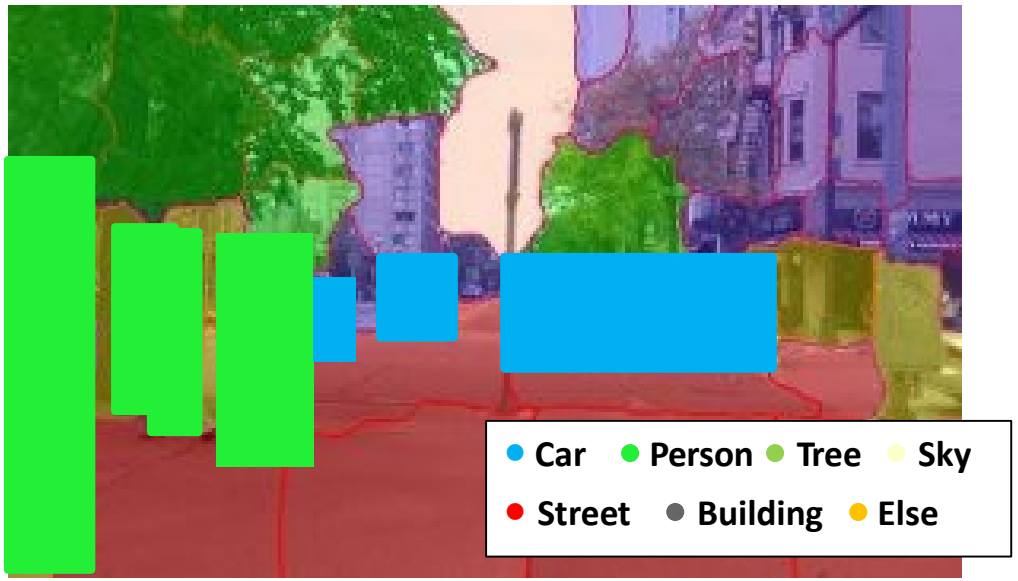
what pathway
(ventral stream)

Joint reconstruction and recognition

Input images

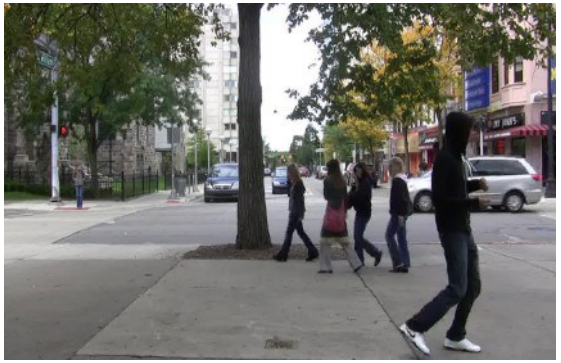


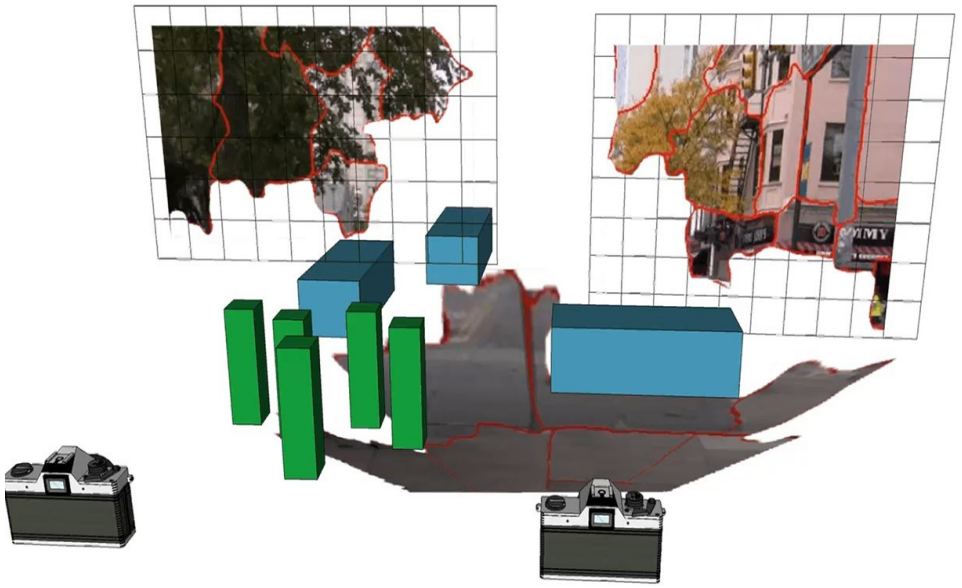


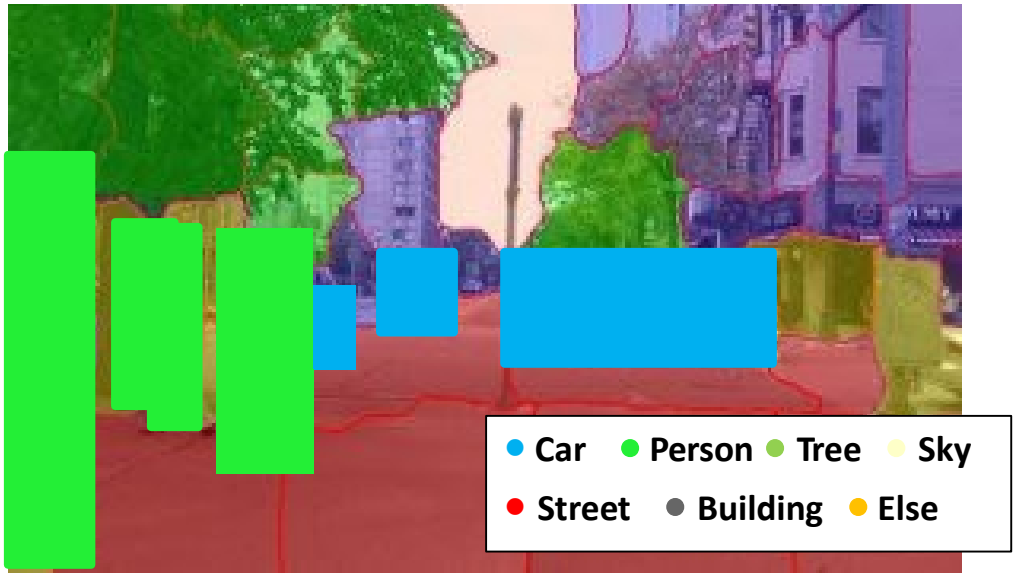


Joint reconstruction and recognition

Input images







“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

Syllabus

Lecture	Topic	
1	Introduction	3D geometry
2	Camera models	
3	Camera calibration	
4	Single view metrology	
5	Epipolar geometry	
6	Multi-view geometry	
7	Structure from motion/ SLAM	
8	Volumetric stereo	
9	Fitting and Matching	Recognition
10	Detector and Descriptors	
11	Intro to Recognition; Object classification I	
12	Object classification II	
13	Scene understanding & segmentation	
14	Visual Representation Learning by NNs	
15	3D Object recognition	
16	3D Scene understanding	

Project presentations