Computer Vision Tutorial

basic introduction

Liyuan Cao Haidong Gu

Industrial and Systems Engineering
Lehigh University

OptML, Sep 4, 2019

Table of Contents

Table of Contents

What is Computer Vision?

Objective of Computer Vision

The objective of computer vision is to make computers see and interpret the world like humans and possibly even better than us.

What is Computer Vision?

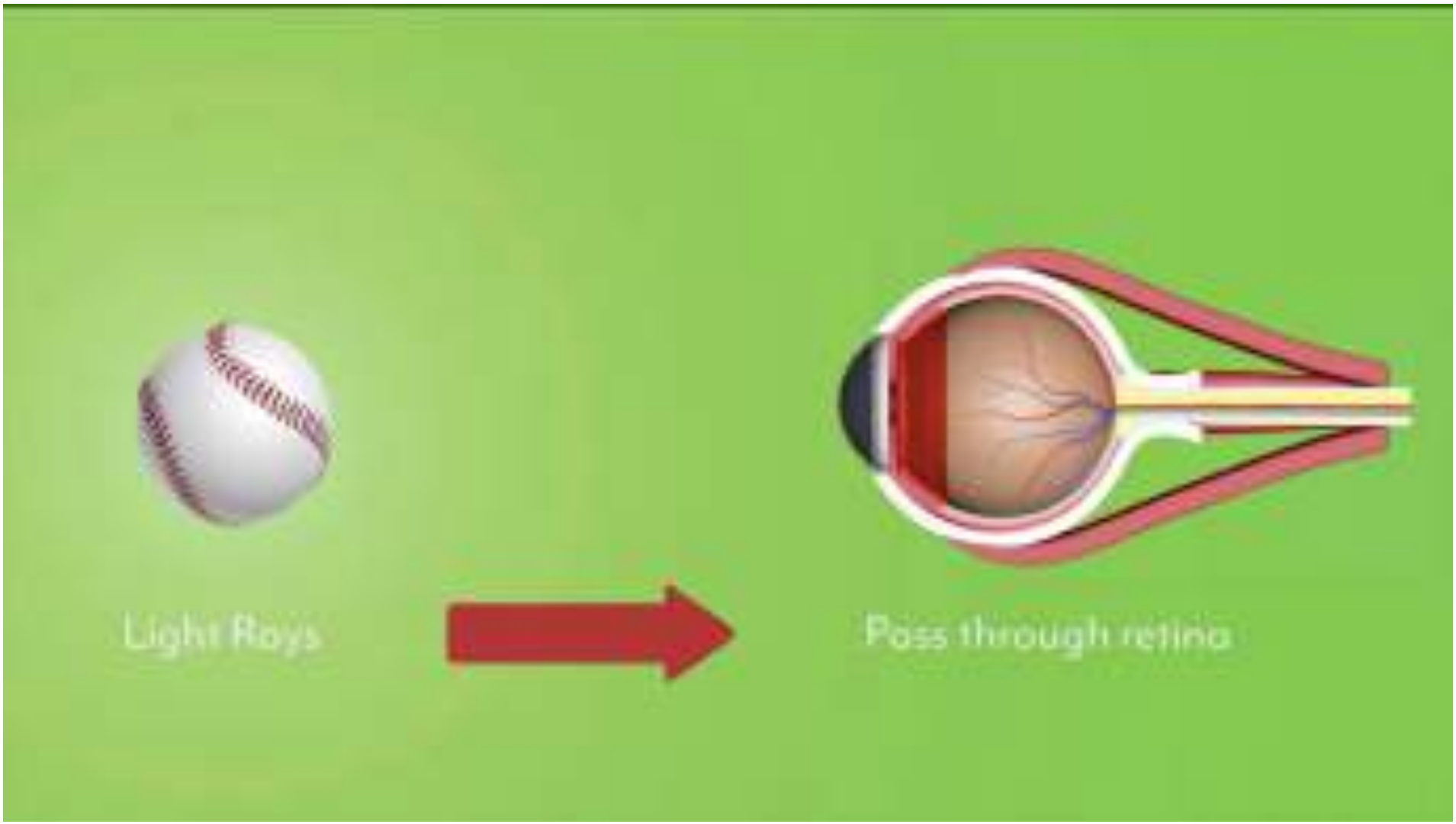
Human vision performs multiple visual tasks quite effortlessly and effectively.

- How is visual information processed and understood in biological systems?
- What is the nature of computation involved in visual tasks?
- How might we build machines that can see?

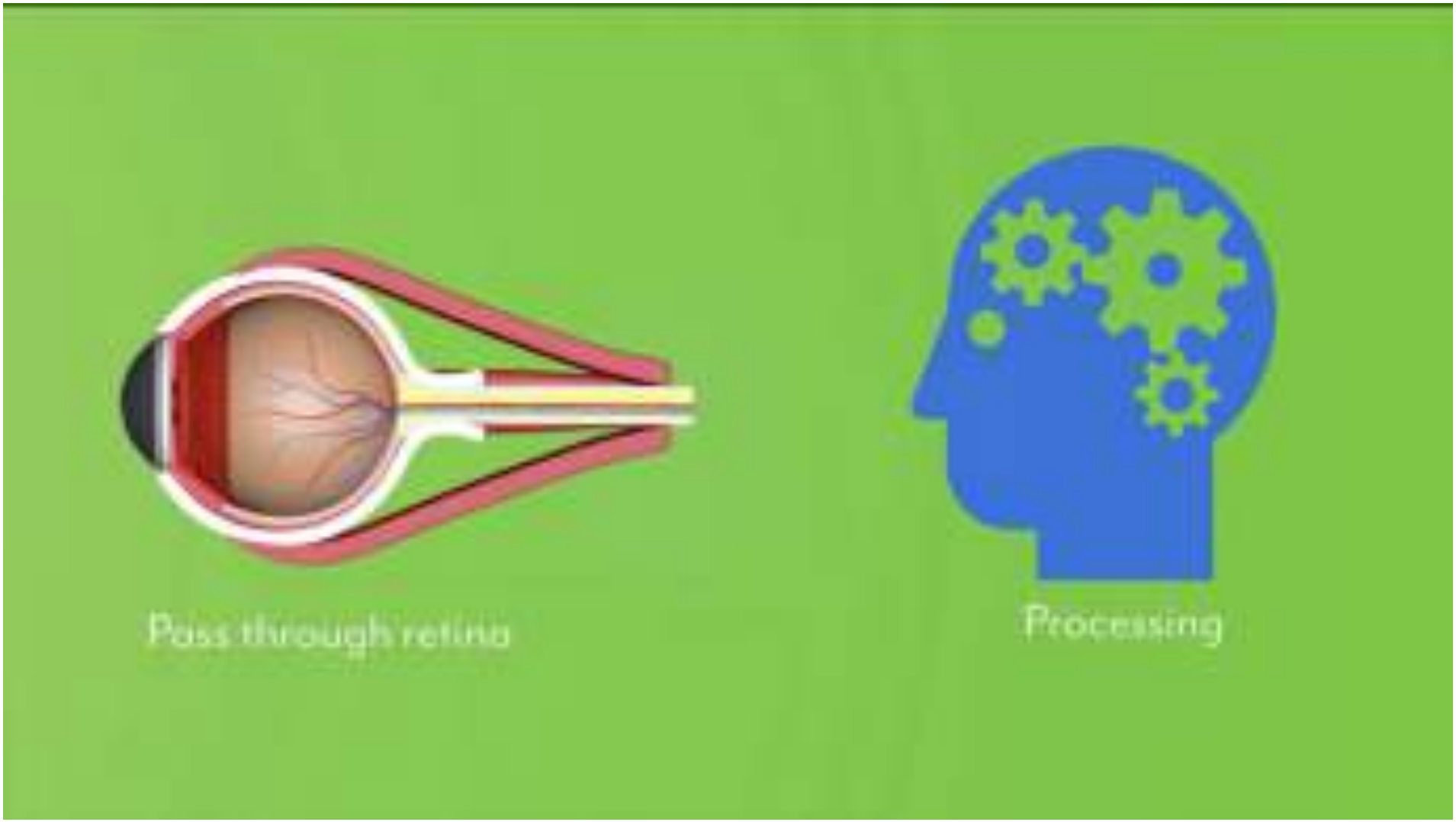
What is Computer Vision?

Figure: Ball Catch

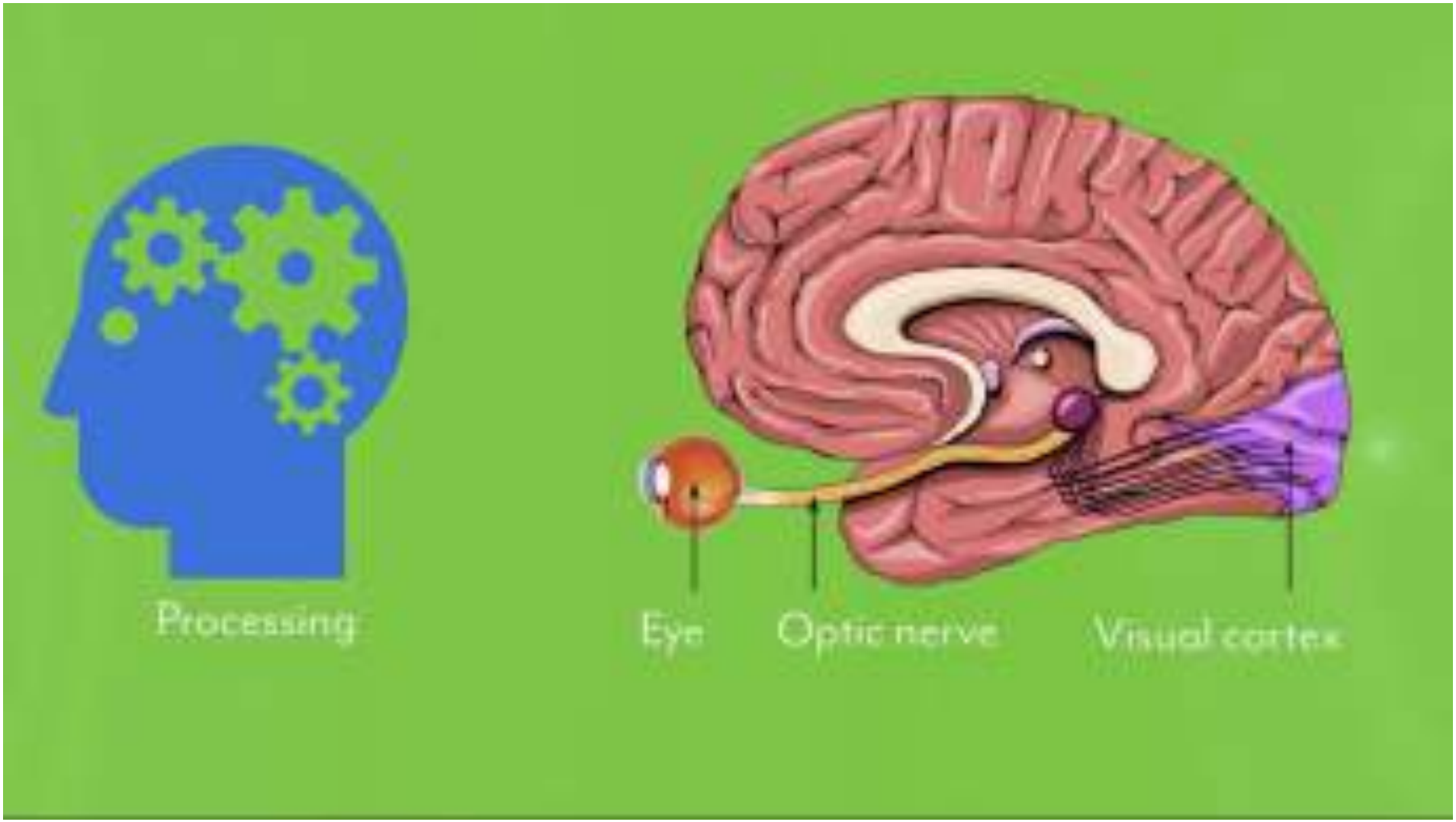
What is Computer Vision?



What is Computer Vision?



What is Computer Vision?



What is Computer Vision?

What is Computer Vision?

What is Computer Vision?

Recreating human vision isn't just a hard problem, it's a set of them, each of which relies on the other.

Definition of Computer Vision

Computer Vision is concerned with the automatic extraction, analysis and understanding of useful information from a single image or a sequence of images.

Computer Vision

theoretical and algorithmic basis to
achieve automatic visual understanding

Computer Vision

high-level understanding from digital images or videos

From the biological science point of view,

Computer Vision

computer vision aims to come up with computational models for human visual system

From the engineering point of view,

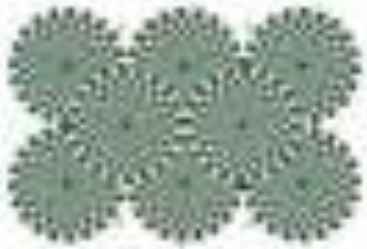
Computer Vision

computer vision aims to build autonomous systems to perform some of the tasks which the human visual system can perform and even surpass it in many cases.

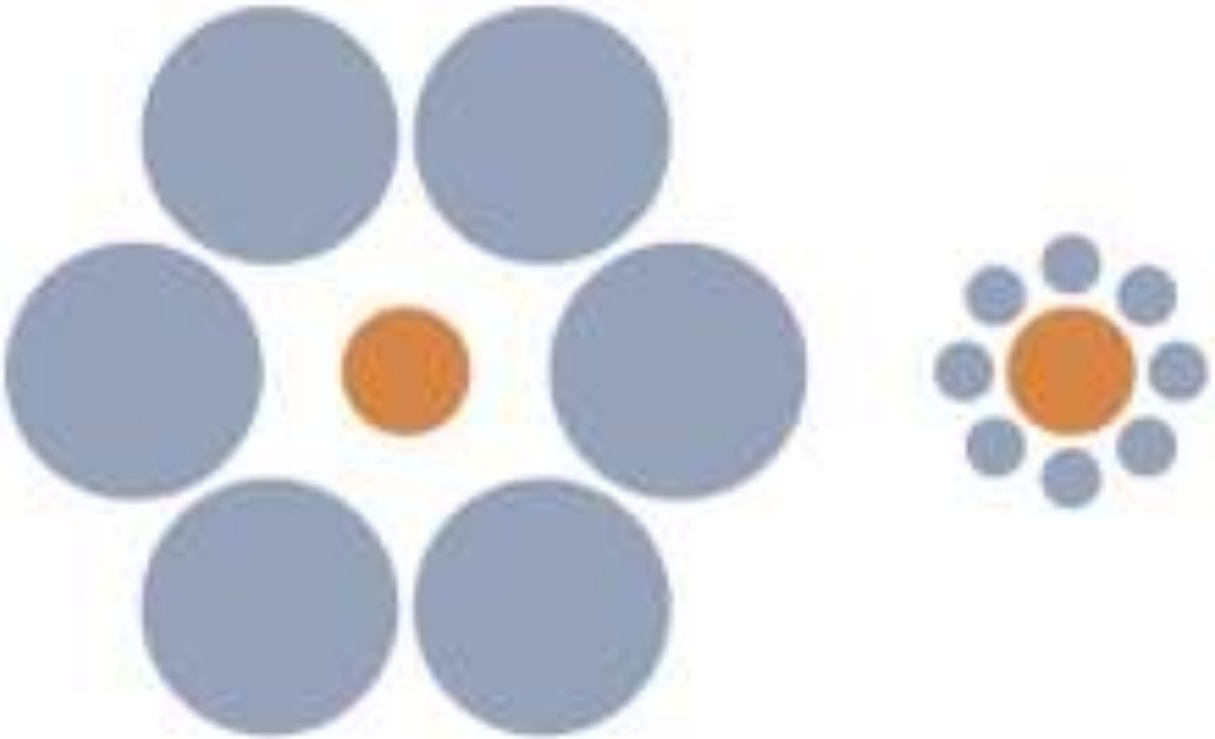
What is Computer Vision?

What is Computer Vision?

Optical Illusions



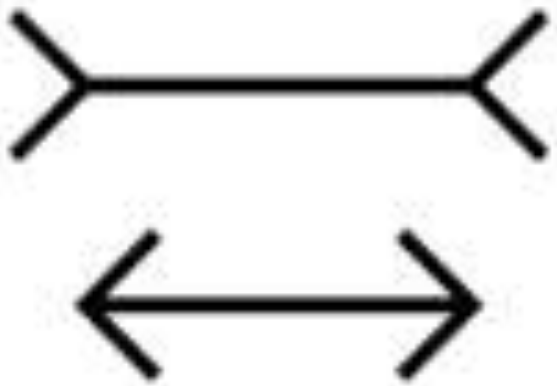
What is Computer Vision?



The Ebbinghaus illusion

What is Computer Vision?

What is Computer Vision?



The Muller-Lyer illusion

What is Computer Vision?

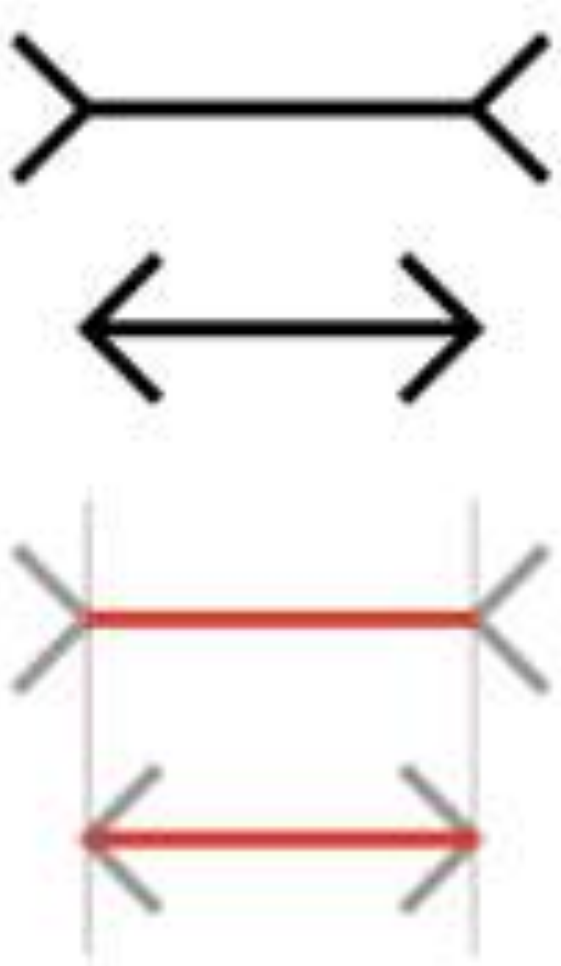
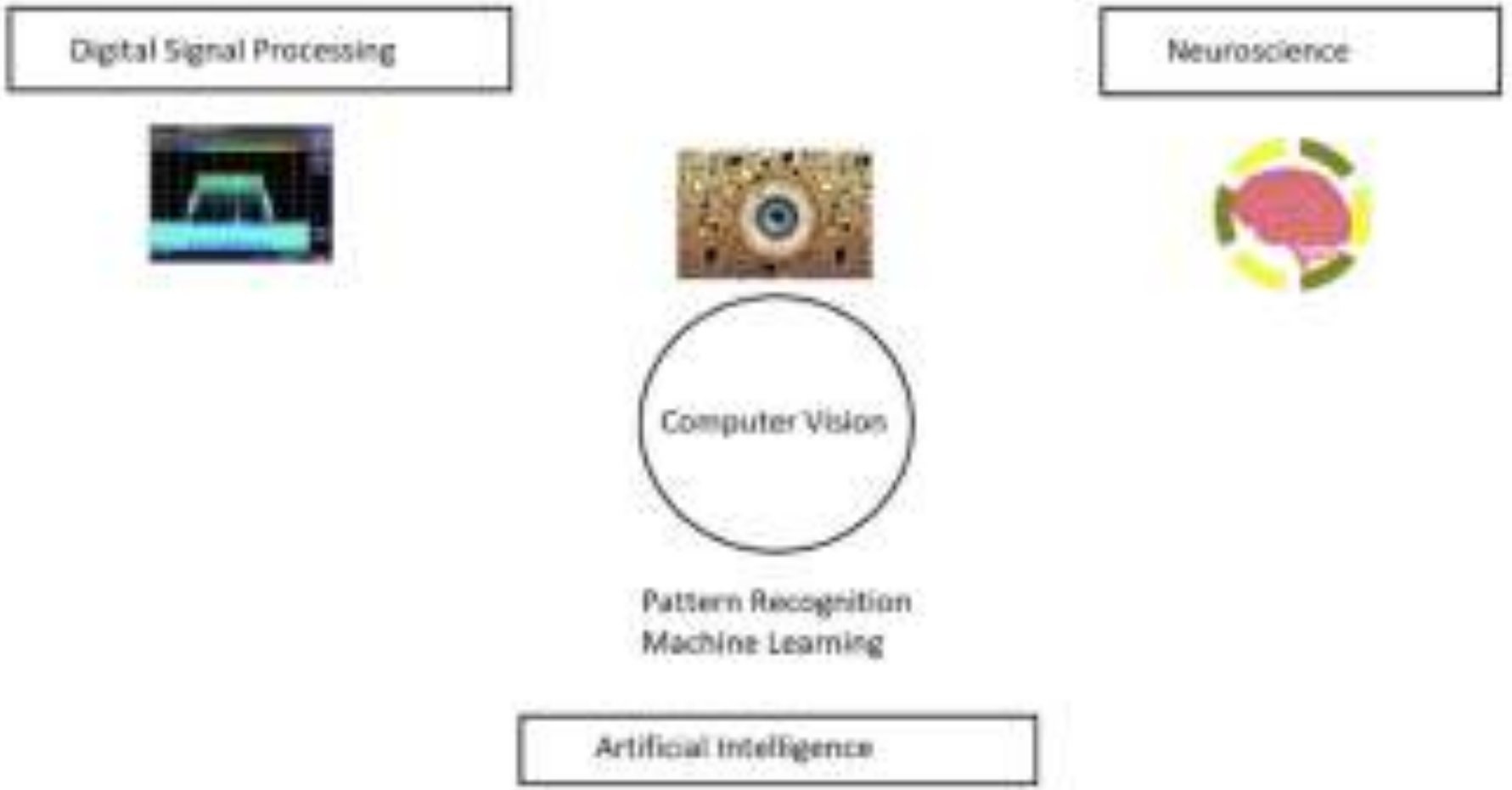


Table of Contents

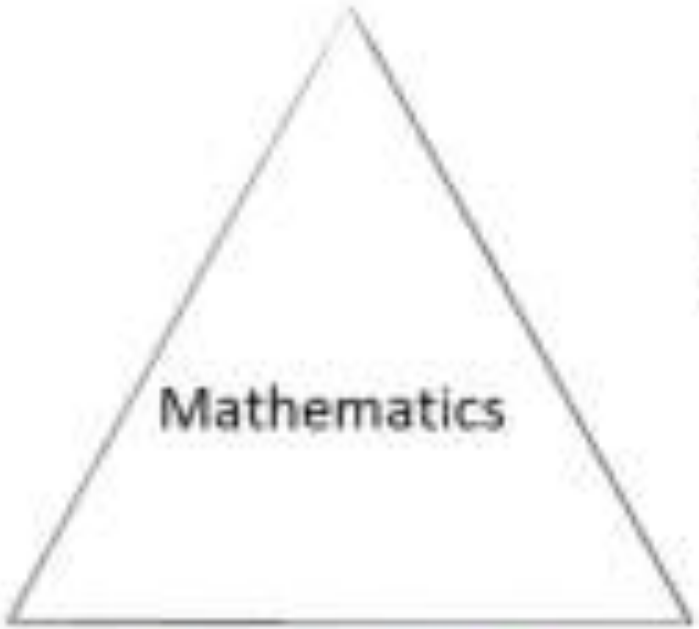
Related Fields of Computer Vision



Computer Vision

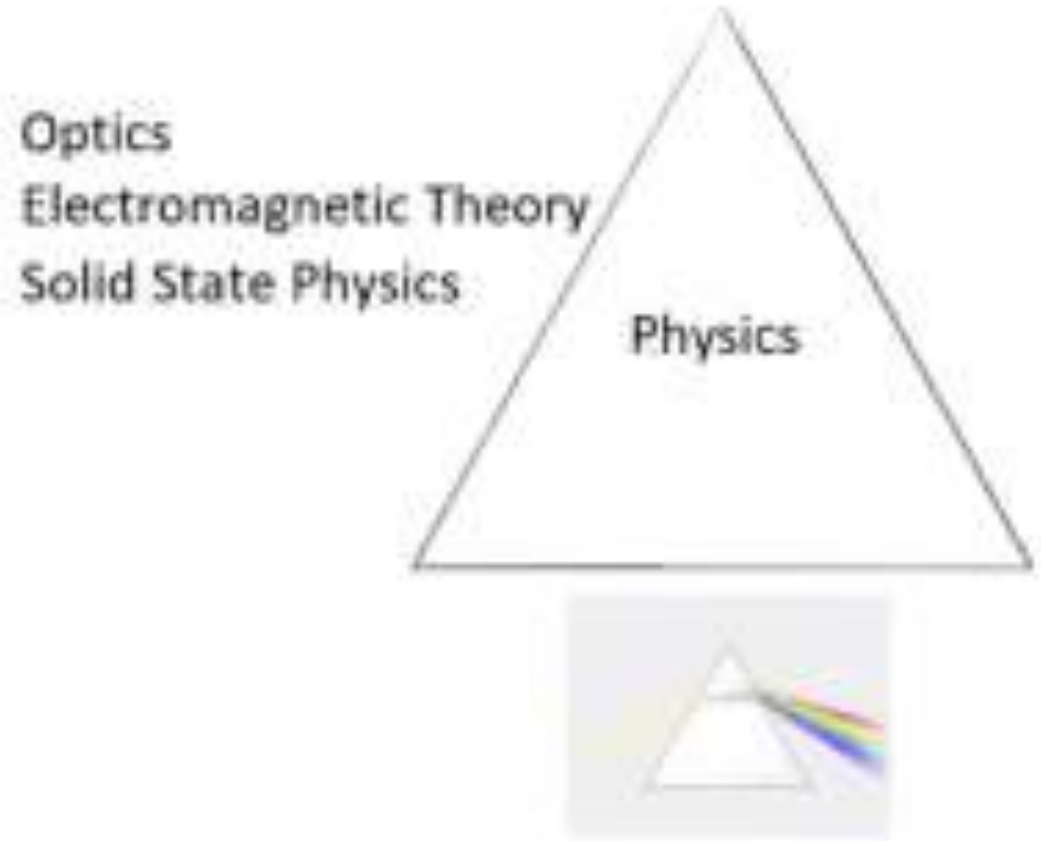
can be studied from a purely mathematical point of view

Related Fields of Computer Vision



Statistics
Optimization
Geometry

Related Fields of Computer Vision



Computer vision systems rely on image processing techniques to pre-process the image data for robust high-level analysis.

Digital Image Processing

image compression

image restoration

image enhancement

The robust high-level analysis is the next major task in computer vision pipeline. It is the area where neuroscience plays an important role. Specifically, the study of the biological vision system.

Neuroscience

(a) eye

(b) neuron

(c) brain structure

Related Fields of Computer Vision

Table of Contents

Computer vision is transitioning from a nascent stage and is proving to be incredibly useful in several application areas.

Computer Vision for Security

Computer Vision Applications

Figure: visual surveillance

Computer Vision Applications

Figure: fingerprint-based identification and authentication

Computer Vision Applications

Figure: iris identification and authentication

Computer Vision Applications

Figure: face recognition

Computer Vision in Multimedia and Entertainment

Computer Vision Applications

Figure: face recognition

Computer Vision Applications

Figure: augmented reality game

Computer Vision for Navigation

Computer Vision Applications

Figure: robot navigation

Computer Vision Applications

Figure: autonomous driving

Computer Vision in Retail

Computer Vision Applications

Figure: automated supermarket

Computer Vision Applications

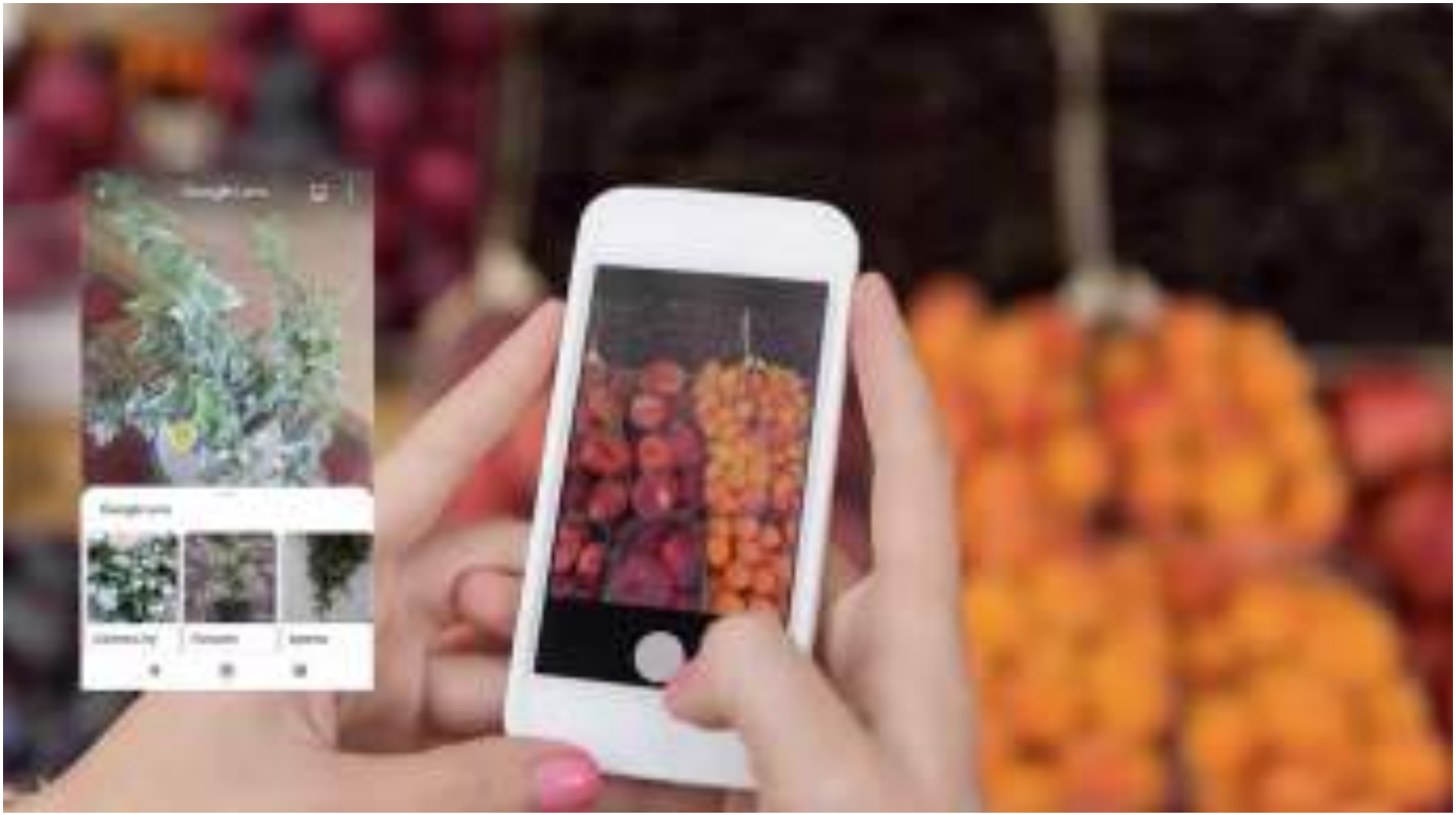


Figure: visual search on smartphone cameras

Computer Vision in Industry

Computer Vision Applications

Figure: real-time box analytics

Computer Vision Applications

Figure: auto cycle counting

Computer Vision Applications

Figure: inspect products

Computer Vision for Social Causes

Computer Vision Applications

Table of Contents

Computer vision concepts
can be broadly categorized as
low, mid and high level vision techniques

Low Level Vision

image processing techniques, feature detection and matching and early segmentation

Three-Level Paradigm

Image features

Three-Level Paradigm

Mid Level Vision

image segmentation, things start to come together

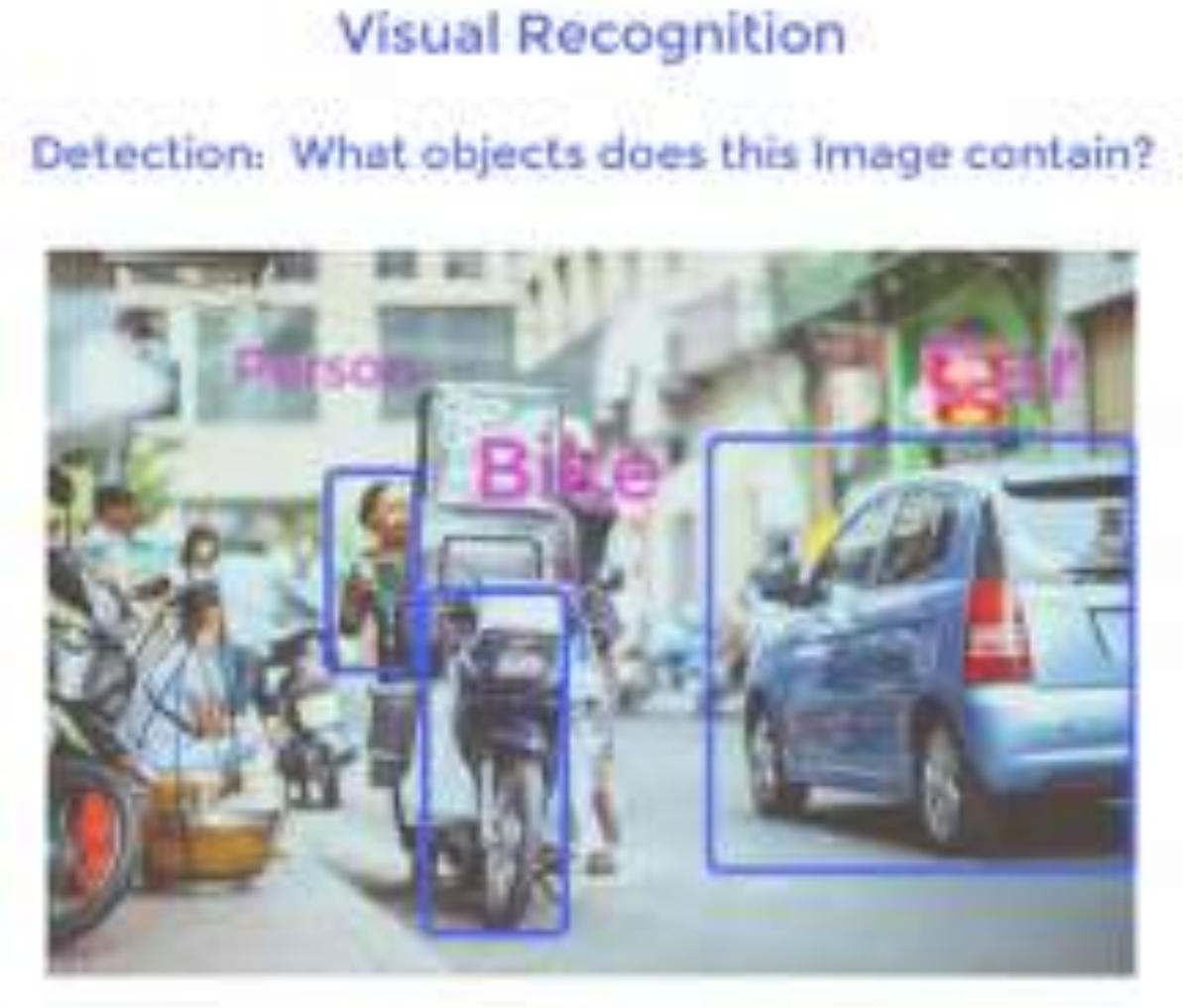
attributing meaning

Three-Level Paradigm

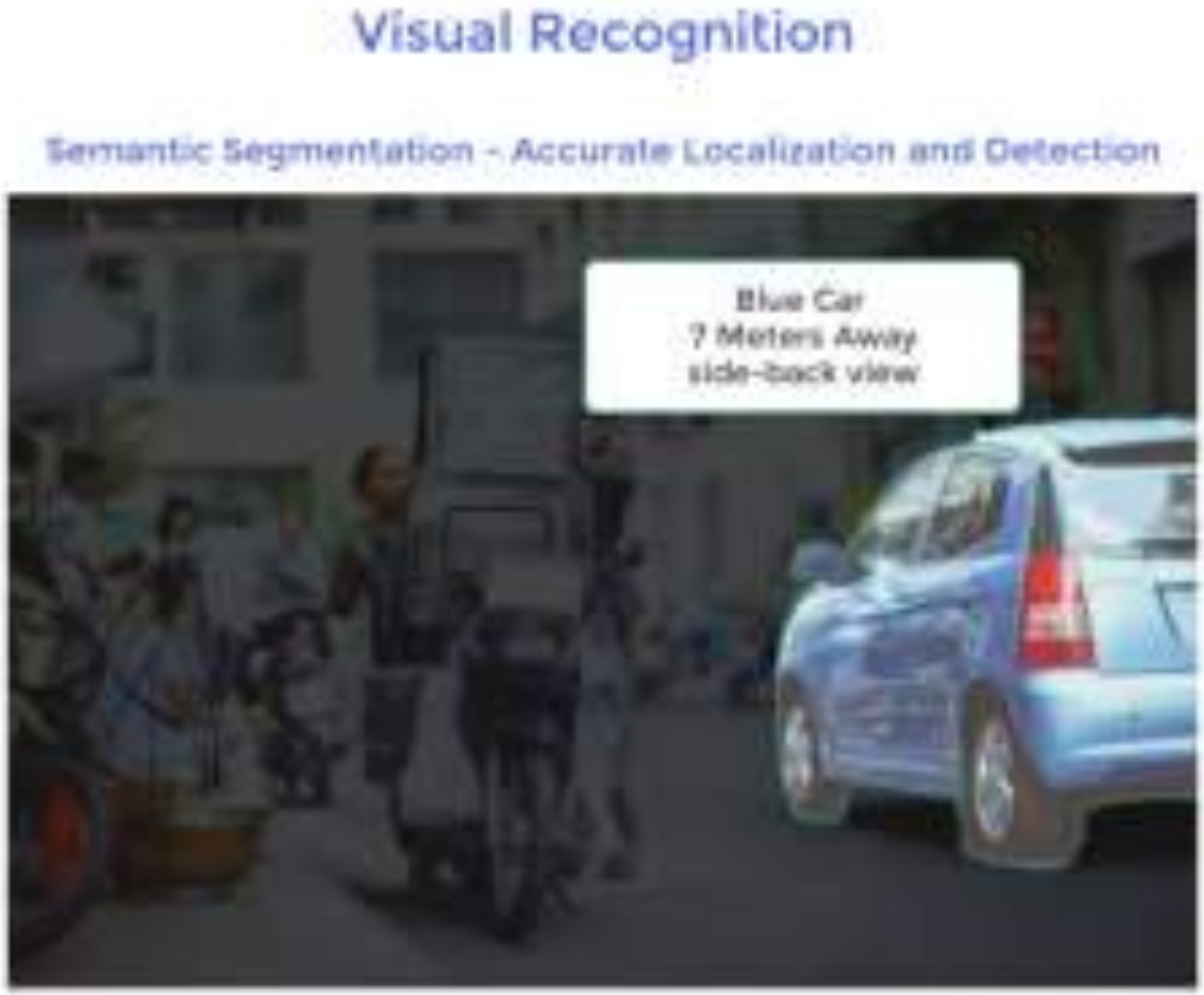
High Level Vision

Visual recognition and scene understanding, algorithms which makes sense of the visual content and make computer vision live up to the capabilities of human vision

Three-Level Paradigm



Three-Level Paradigm



Three-Level Paradigm

Visual Recognition

Event Recognition

Three-Level Paradigm

Low-Level or Early Vision

Considers local properties of an image

"There are several edges in this image!"

Three-Level Paradigm

Mid-Level Vision

Grouping and Segmentation

"There is an object and a background. Let us try to identify which pixels belong to which one."

Three-Level Paradigm

High-Level Vision

"It's an Aeroplane!"

Table of Contents

Types of CV Problems

Recognition determining whether or not the image data contains some specific object, feature, or activity

Motion Analysis an image sequence is processed to produce an estimate of the velocity either at each points in the image or in the 3D scene, or even of the camera that produces the images

Scene Reconstruction computing a 3D model of the scene

Image Restoration The aim of image restoration is the removal of noise (sensor noise, motion blur, etc.) from images.

Recognition: Object Recognition, Identification, Detection

Figure: **Object Recognition**: one or several pre-specified or learned objects or object classes can be recognized, usually together with their 2D positions in the image or 3D poses in the scene

Figure: **Identification**: an individual instance of an object is recognized

Detection: the image data are scanned for a specific condition.

Motion Analysis: Egomotion, Tracking, Optical flow

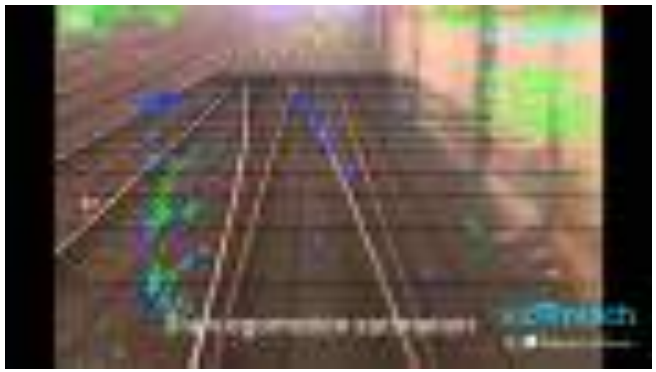



Figure: **Egomotion**: determining the 3D rigid motion (rotation and translation) of the camera from an image sequence produced by the camera. (visual odometry)


Figure: **Tracking**: following the movements of a (usually) smaller set of interest points or objects (e.g., vehicles, humans or other organisms) in the image sequence.

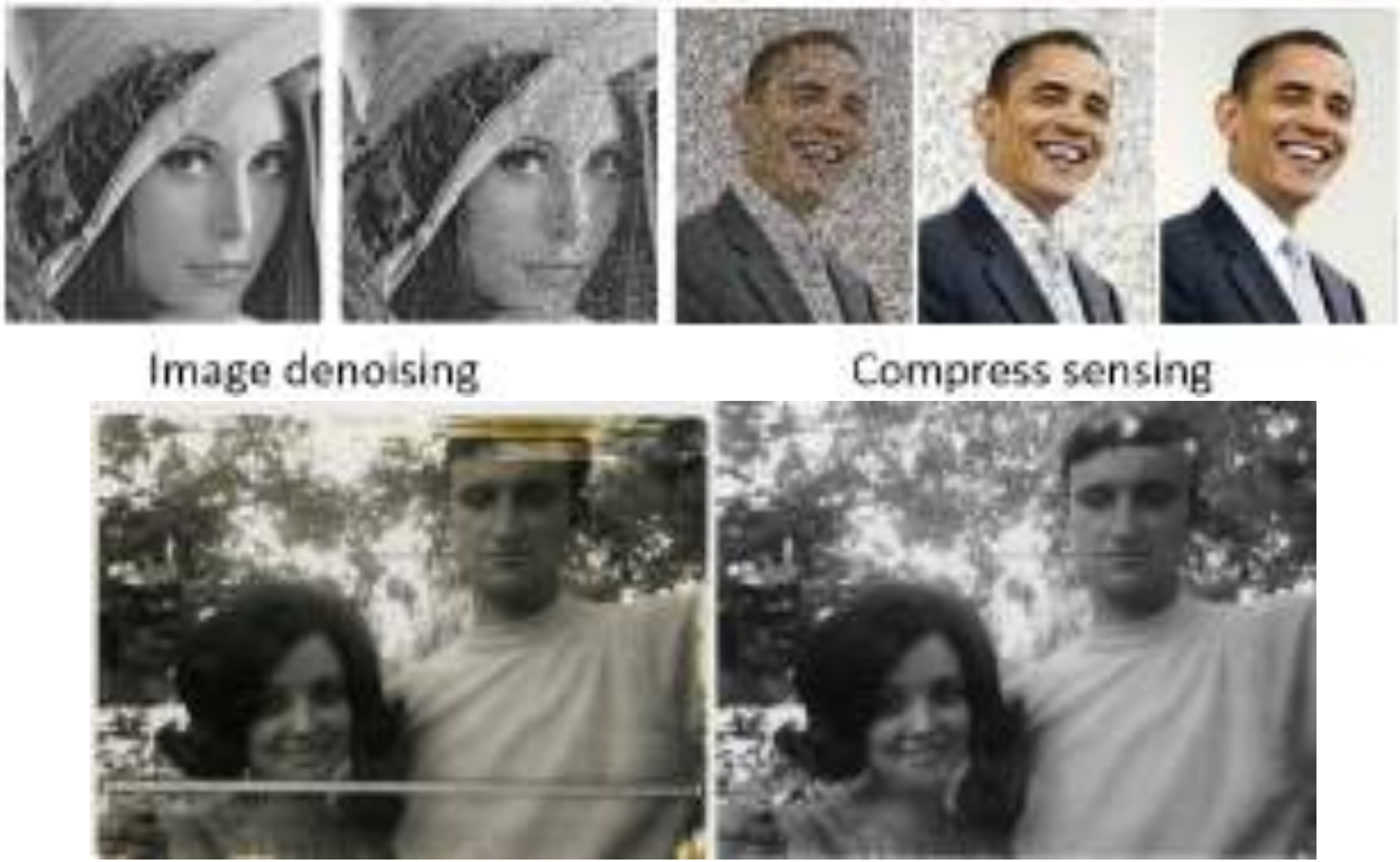
Motion Analysis: Egomotion, Tracking, Optical flow

Figure: **Optical Flow**: to determine, for each point in the image, how that point is moving relative to the image plane, i.e., its apparent motion. This motion is a result both of how the corresponding 3D point is moving in the scene and how the camera is moving relative to the scene.

Scene Reconstruction

Figure: **Scene Reconstruction**: Given one or (typically) more images of a scene, or a video, scene reconstruction aims at computing a 3D model of the scene.

Image Restoration



(a) inpainting

Figure: **Image Restoration**