Automotive Sensor

Camera

Automotive Intelligence Lab.

Copyright © AI Lab. All rights reserved.

Contents

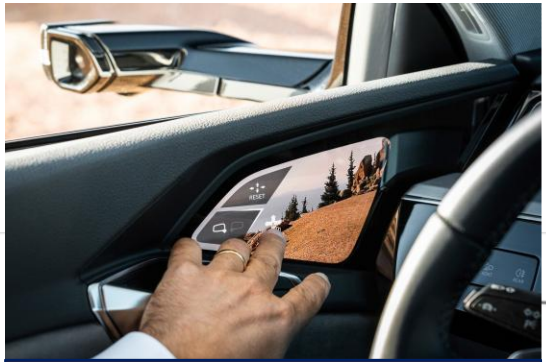
Copyright © AI Lab. All rights reserved.

Camera in autonomous vehicle

Copyright © AI Lab. All rights reserved.

Cameras in Autonomous Vehicle

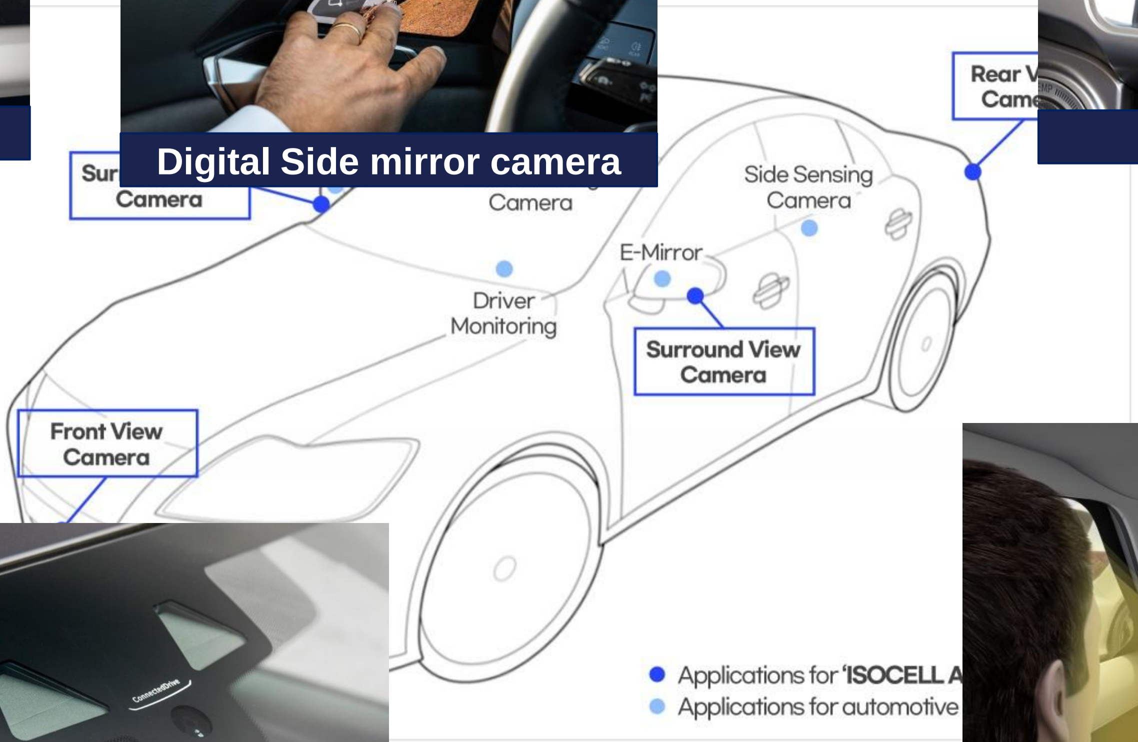
Surrounding view camera



Digital Side mirror camera

Rear view camera

Rear-side view camera



Front view camera

Driver monitoring camera

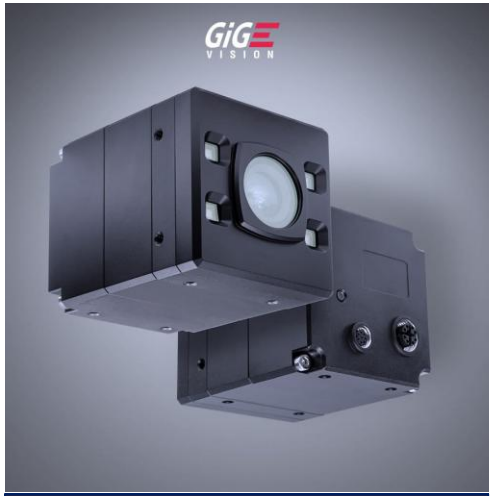
Copyright © AI Lab. All rights reserved.

Various Types of Camera Sensors

- Mono camera
- Stereo camera
- Time of Flight (ToF) camera
- Infrared (IR) camera


Monocular camera


Stereo camera



ToF camera

Infrared camera

Copyright © AI Lab. All rights reserved.

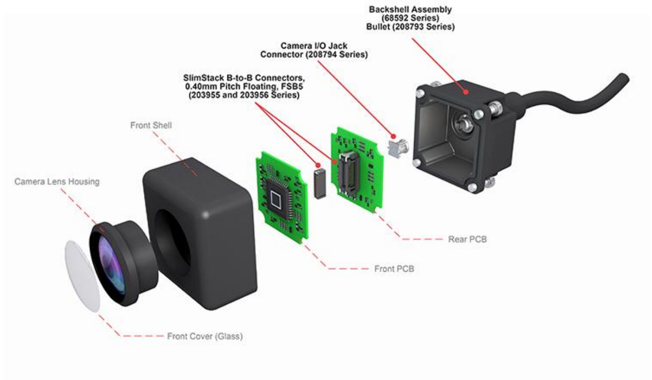
Mono Camera

- **An optical instrument that can capture an image**

- **Pros**
  - Passive
  - Low cost
  - Small and lightweight
  - High spatial resolution
  - Provides rich feature such as color and texture

- **Cons**
  - Limited range
  - Can only operate during day-time
  - Affected by dirt, poor illumination and bad weather

Mono camera



Structure of mono camera

Example of mono camera data

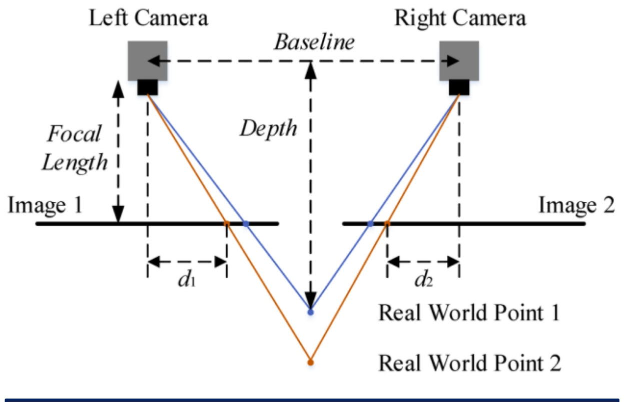
Copyright © AI Lab. All rights reserved.

Stereo Camera

- A type of camera with two or more lenses with a separate image sensor or film frame for each lens

- **Pros**
  - Passive
  - Provide depth information
  - High spatial resolution
  - Low cost compared to its active counterparts

Stereo camera



Principle of stereo camera

- **Cons**

피사체에 따라 영향 받음

  - Can only operate during day-time
  - Require extensive calibration and expertise
  - Inferring depth is computationally heavy
  - Depth inference with this sensor usually depends upon image containing shapes and texture

Stereo images and estimated depth map

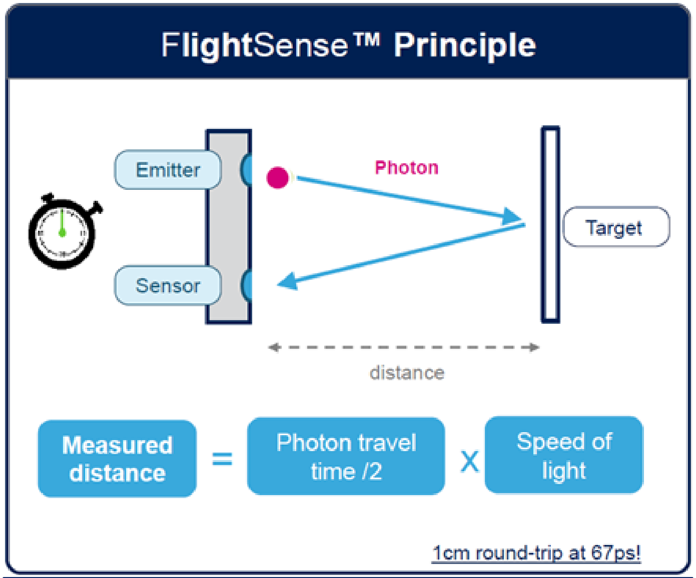
Copyright © AI Lab. All rights reserved.

Time of Flight (ToF) Camera

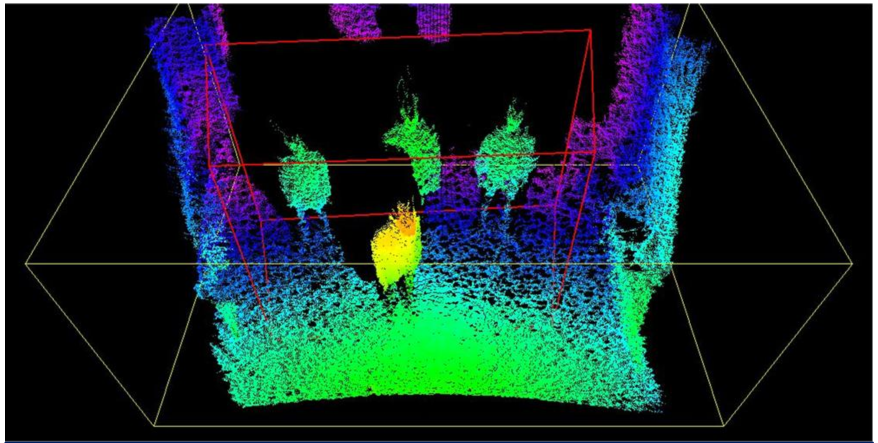
- **A sensor that uses Time-of-Flight to measure depth.**
  - Time-of-Flight technique is used to calculate the distance between

- **Pros**
  - High accuracy
  - Rapid data acquisition
  - Less computing power than stereo camera

- **Cons**
  - Lower spatial resolution than stereo camera
  - Cannot easily detect glass
  - Performance degrades due to strong ambient light (e.g. sunlight), scattered light scenarios.



Principle of ToF camera



ToF data example

Copyright © AI Lab. All rights reserved.

Infrared (IR) Camera

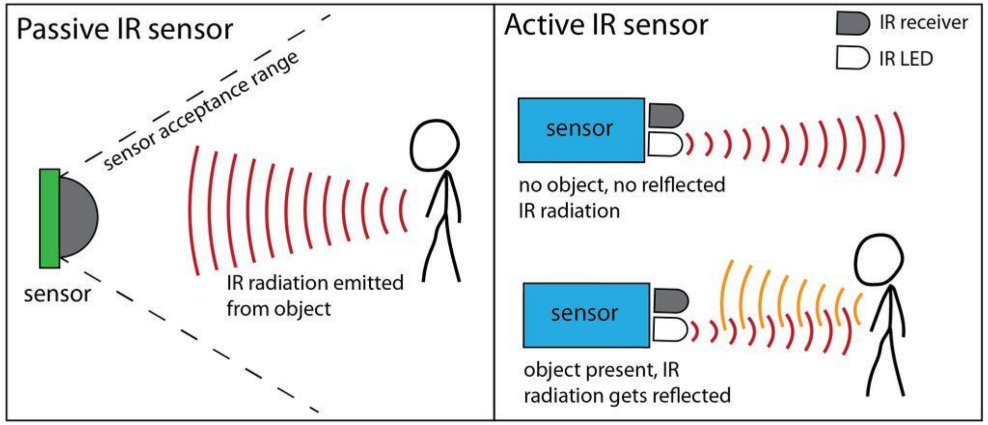
Passive IR camera

- Do not use an infrared light source, instead they capture thermal radiation already emitted by the objects, using a thermographic camera
- Pros
  - Higher contrast for living objects
  - Suitable for both day and night operation
  - Do not affected by oncoming headlights, smoke, haze, dust, etc.
- Cons
  - Works poorly in warmer weather conditions

Active IR camera

- Use an infrared light source to illuminate the road ahead with light that is invisible to humans
- Pros
  - Superior picture of inanimate objects
  - Better resolution than passive IR camera
  - Suitable for both day and night operation
- Cons
  - Does not work as well in fog or rain
  - Lower contrast for animals

IR Camera



Types of IR Cameras

BMW Night Vison

Copyright © AI Lab. All rights reserved.

Advantages and Disadvantages of Camera

Pros

- High-resolution in pixels and color across the full width of its field of view
- Constant 'frame-rate' across the field of view
- Capability to provide stereoscopic 3D vision with two camera
- 'Passive' system
  - There are no co-existence problems with other vehicles' transmissions

Cons

- Performance in poor lighting conditions (i.e. at night)
- Performance in bad weather
  - What happens when the lenses get coated with dirt or ice?

Copyright © AI Lab. All rights reserved.

Camera Model & Calibration

Copyright © AI Lab. All rights reserved.

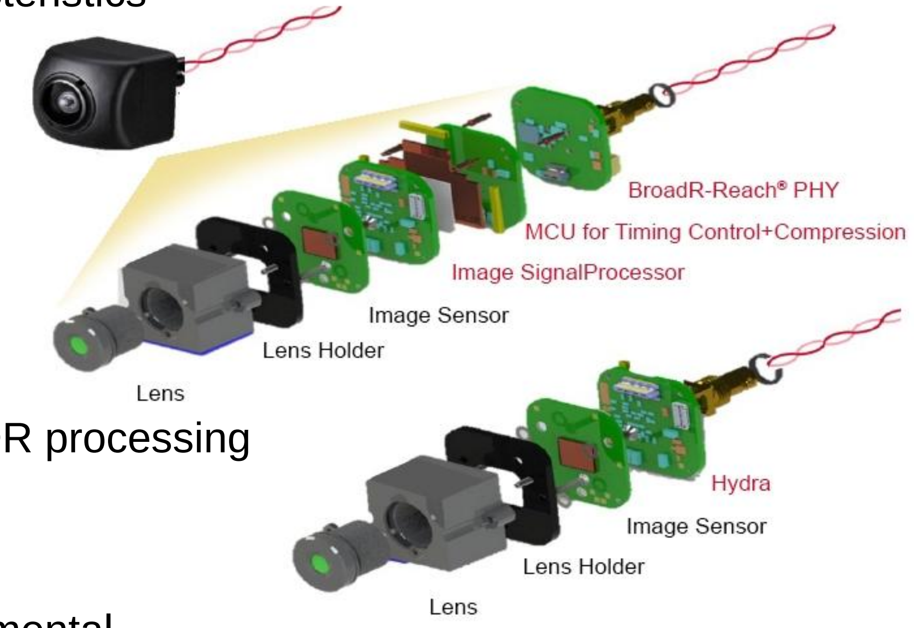
Component of Automotive Camera

Lens
- Collects and directs light onto the image sensor
- Determines focal length, field of view (FOV), and distortion characteristics

Image Sensor
- Converts optical signals into electrical signals
- Typically uses CMOS sensors (high resolution, rapid processing)
- Key parameters: sensitivity, noise characteristics, dynamic range

Image Signal Processor (ISP)
- Processes raw sensor data into clear images
- Performs noise reduction, white balance, color correction, and HDR processing
- Essential for image quality and reliability

Vision Processing Module
- Runs computer vision algorithms for object detection and environmental perception
- Enables autonomous driving features (lane detection, object tracking, distance estimation, etc.)
- Often utilizes high-performance GPUs or dedicated AI processors (e.g., NVIDIA, Mobileye)



Copyright © AI Lab. All rights reserved.

Image Representation (1/3)

Image Acquisition Process

색수차
(색마다 굴절률 다름)

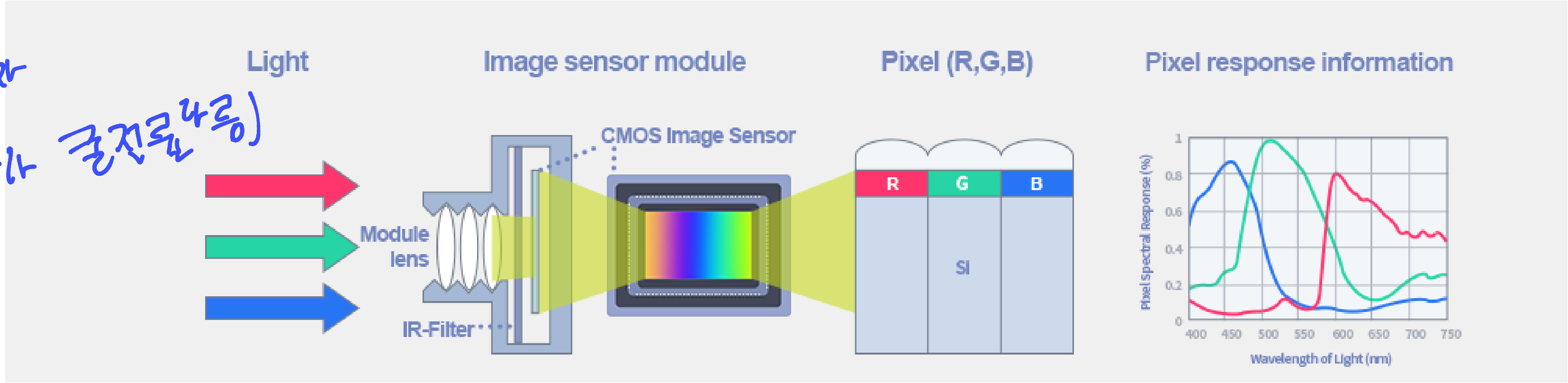
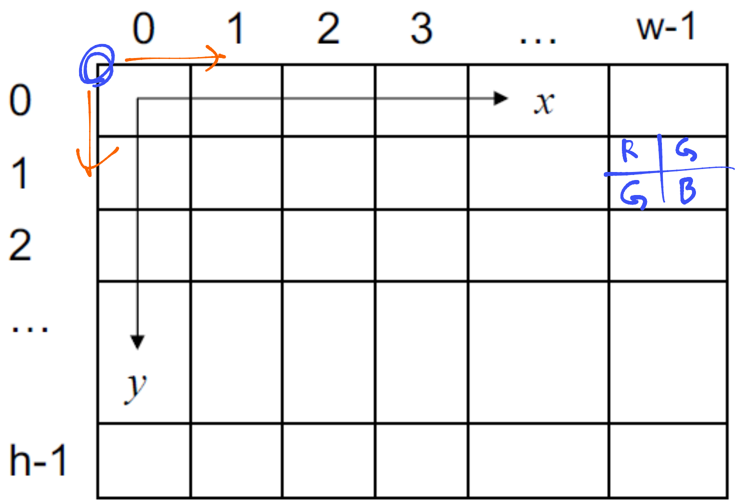


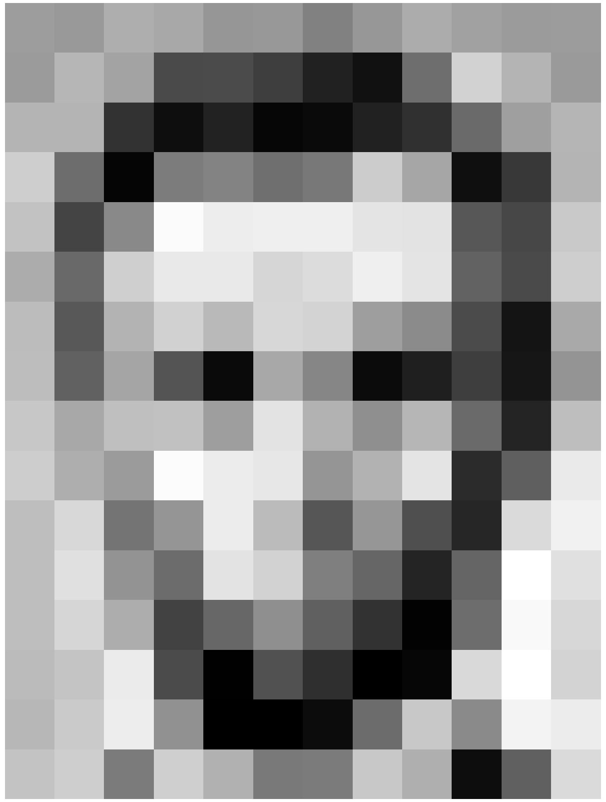
Image Representation

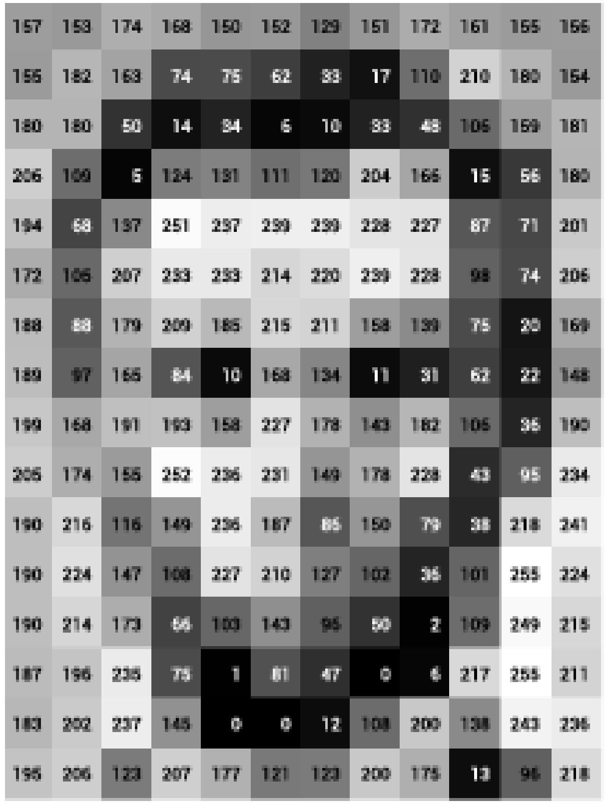
- Pixel
  - The smallest unit that composes an image
  - An image is an array of pixels
- Image coordinates
  - Normally, the top left is (0,0)
  - f(x,y): pixel values at coordinates (x,y)



Copyright © AI Lab. All rights reserved.

Image Representation (2/3)





157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	106	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
206	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	236	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

Copyright © AI Lab. All rights reserved.

Image Representation (3/3)

Grayscale image

- Pixel values are expressed in 8 bits (256 levels): 0~255
- 0: black, 255: white

Color image

- Expresses pixel values with a combination of red, green, and blue
- (R, G, B): (0, 0, 0) ~ (255, 255, 255)

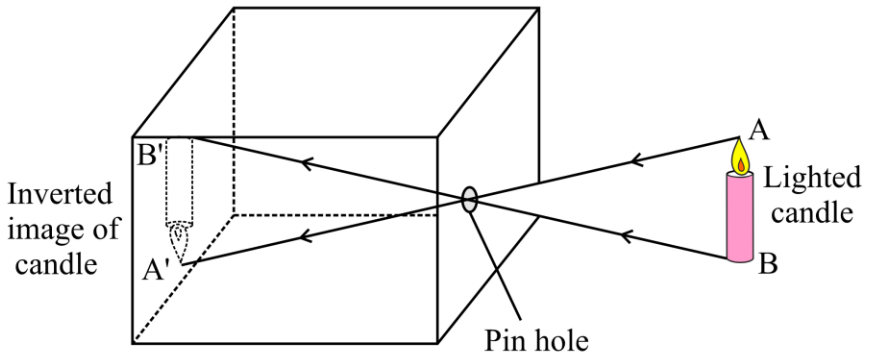
Grayscale image

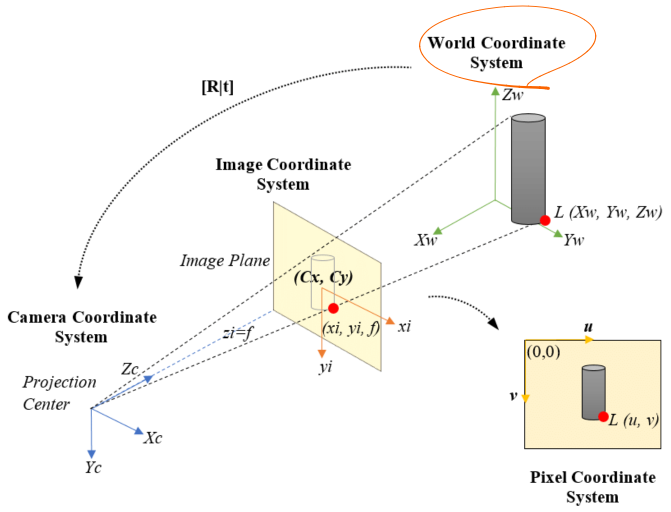
Color image

Copyright © AI Lab. All rights reserved.

Camera Model

- **The process of projecting a 3D point (in meters) to a 2D image plane (in pixels) can be described by a geometric model.**
- **The projection of a 3D point onto a 2D image plane can be modeled using the simple pinhole camera model.**
  - Due to the camera lens, distortion occurs during this projection, necessitating the pinhole model combined with a distortion model.
  - The pinhole model is similar to how a small hole projects an inverted image of a candle onto a surface inside a dark box.
  - This simplified analogy helps explain how a camera transforms a three-dimensional scene into a two-dimensional image.

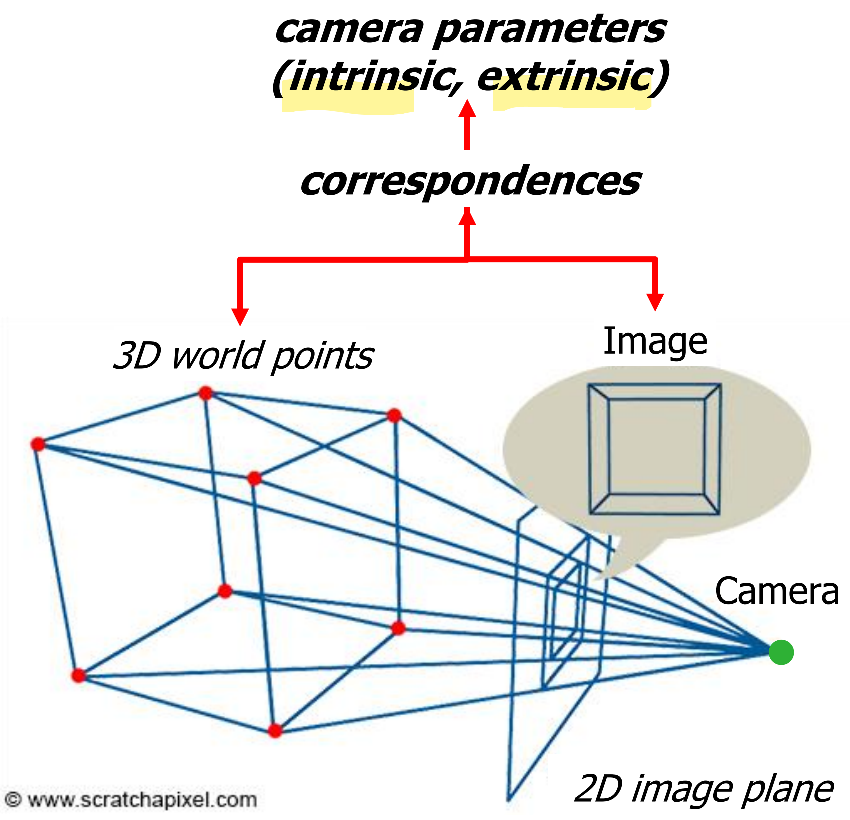




Copyright © AI Lab. All rights reserved.

What is Camera Calibration?

- ***Geometric camera calibration***
  - **Estimation of the parameters about an image sensor**
  - You can use these parameters to determine an accurate **relationship** between a **3D** point in the real world and its corresponding **2D** projection (pixel) in the image

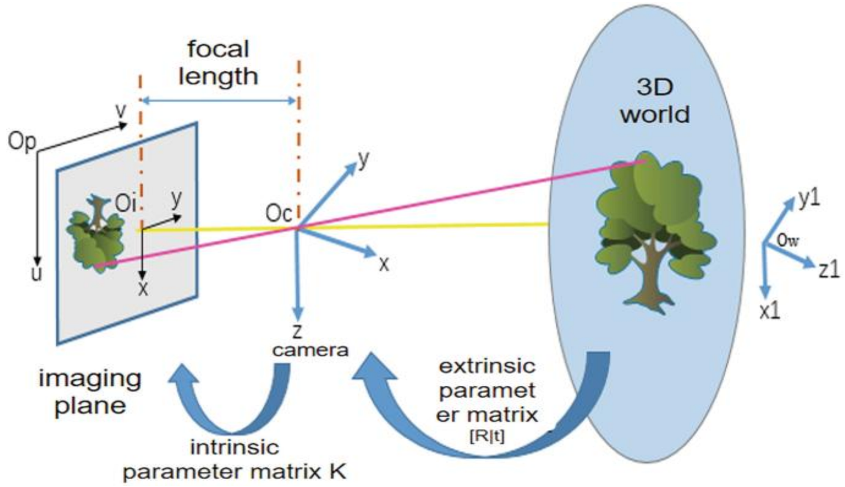


Copyright © AI Lab. All rights reserved.

Two Types of Camera Parameters

- **Extrinsic parameter** → visual odometry
  - The extrinsic parameters, also known as external parameters or camera pose, are the parameters used to describe the transformation between the camera and its external world.

- **Intrinsic parameter**
  - The intrinsic parameters, also known as internal parameters, are the parameters intrinsic to the camera itself, such as the focal length and lens distortion.



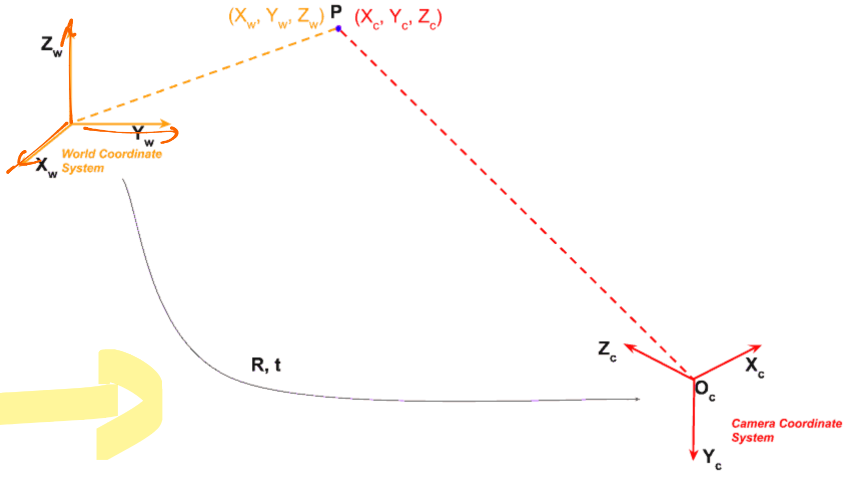
Intrinsic parameter **K**, Extrinsic parameter **[R | t]**

Copyright © AI Lab. All rights reserved.

Extrinsic Parameter (World-to-Camera)

- **To understand the process of calibration, we first need to understand the <u>geometry of image formation</u>.**
  - We will cover the math behind how a point in 3D gets projected on the image plane.
  - We want to find the pixel coordinates (u, v) of 3D point P in the image taken by the camera.

- **World coordinate system**
  - The world coordinate system and the camera coordinate system are related by a rotation and a translation.
  - These **six parameters (3 for rotation, and 3 for translation)** are called the **"extrinsic parameters"** of a camera.



- **Camera Coordinate System**
  - Camera is located at some **arbitrary location.**
    - Camera coordinate is **translated** with respect to the world coordinates.
  - Camera may be also looking in some **arbitrary direction.**
    - Camera is **rotated** with respect to the world coordinate system.

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \underset{\text{Rotation}}{\mathbf{R}} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \underset{\text{translation}}{\mathbf{t}}$$

Copyright © AI Lab. All rights reserved.

Intrinsic Parameter (Camera-to-Image) – (1/2)

Image Coordinate System

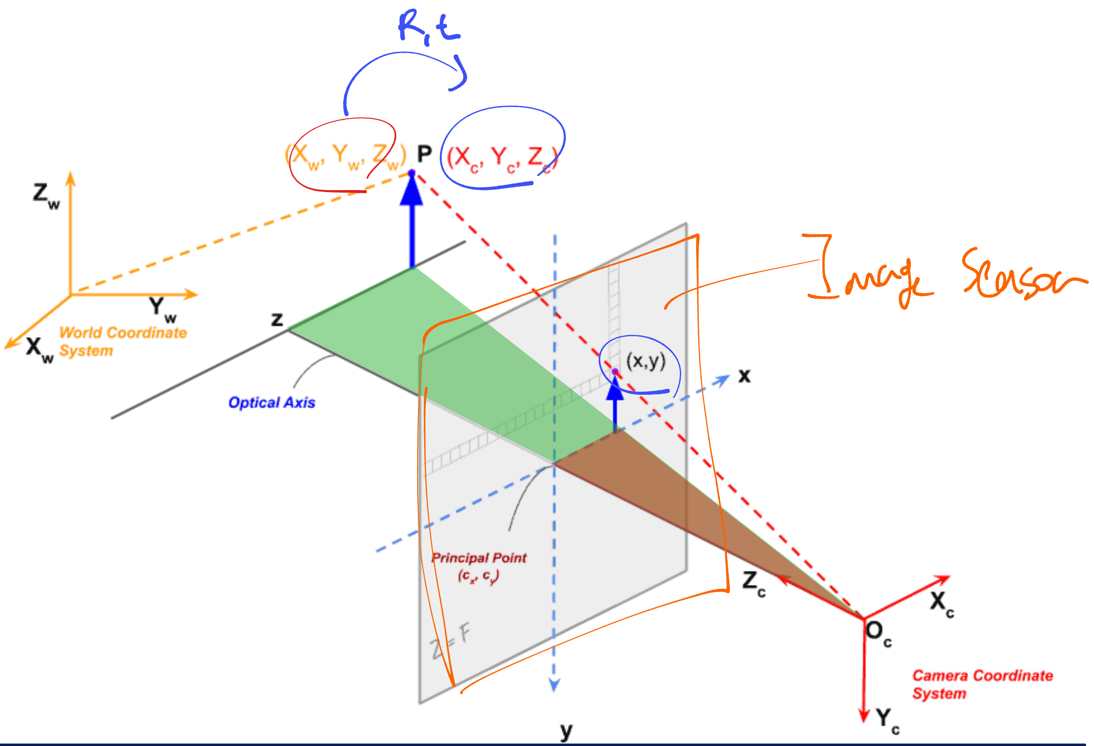
- The image plane is placed at a distance f (focal length) from the optical center.
- Using similar triangles geometry, we can show the project image (x, y) of the 3D point (X_c, Y_c, Z_c) is given by

$$x = f\frac{X_c}{Z_c}$$

$$y = \frac{Y_c}{Z_c}$$

- The above two equations can be rewritten in matrix form as follows

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & f \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$$



Camera projection in case of a simple pinhole camera

Copyright © AI Lab. All rights reserved.

Intrinsic Parameter (Camera-to-Image) – (2/2)

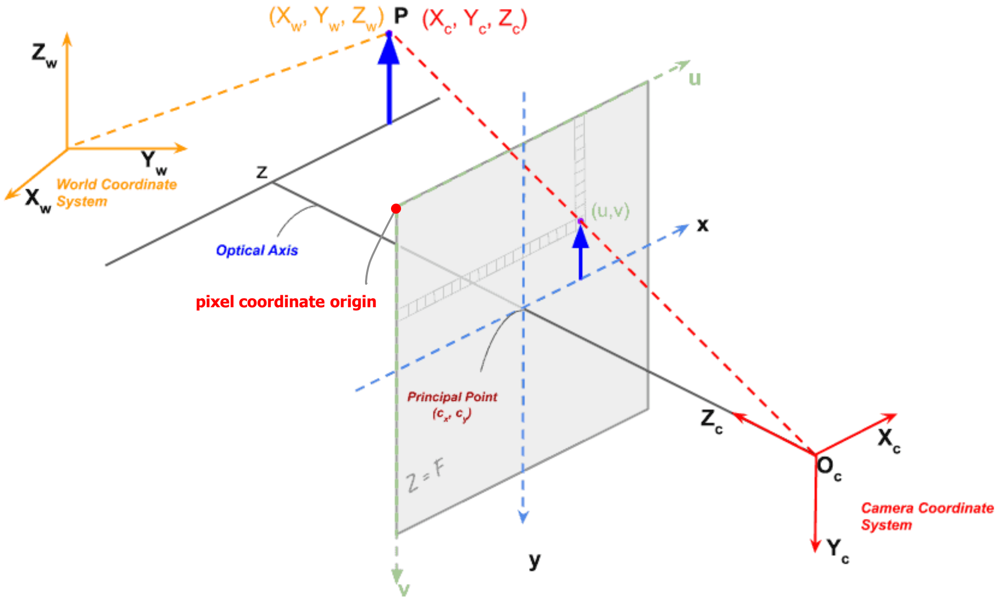
Image Coordinate System

- However, the pixels in the image sensor may not be square, and so we may have two different focal lengths f_x and f_y.
- The optical center (c_x, c_y) of the camera may not coincide with the center of the image coordinate system.
- In addition, there may be a small skew γ between the x and y axes of the camera sensor.
- Taking all the above into account, the camera **"intrinsic parameters"** can be re-written as:

$$\boldsymbol{K} = \begin{bmatrix} \boldsymbol{f_x} & \boldsymbol{\gamma} & \boldsymbol{c_x} \\ \boldsymbol{0} & \boldsymbol{f_y} & \boldsymbol{c_y} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{1} \end{bmatrix}$$

- Representing the image coordinates by (u, v):

$$\begin{bmatrix} \boldsymbol{u'} \\ \boldsymbol{v'} \\ \boldsymbol{w'} \end{bmatrix} = \begin{bmatrix} \boldsymbol{f_x} & \boldsymbol{\gamma} & \boldsymbol{c_x} \\ \boldsymbol{0} & \boldsymbol{f_y} & \boldsymbol{c_y} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{1} \end{bmatrix} \begin{bmatrix} \boldsymbol{X_c} \\ \boldsymbol{Y_c} \\ \boldsymbol{Z_c} \end{bmatrix} \qquad \boldsymbol{u} = \frac{\boldsymbol{u'}}{\boldsymbol{w'}} \quad \boldsymbol{v} = \frac{\boldsymbol{v'}}{\boldsymbol{w'}}$$



More realistic scenario

Copyright © AI Lab. All rights reserved.

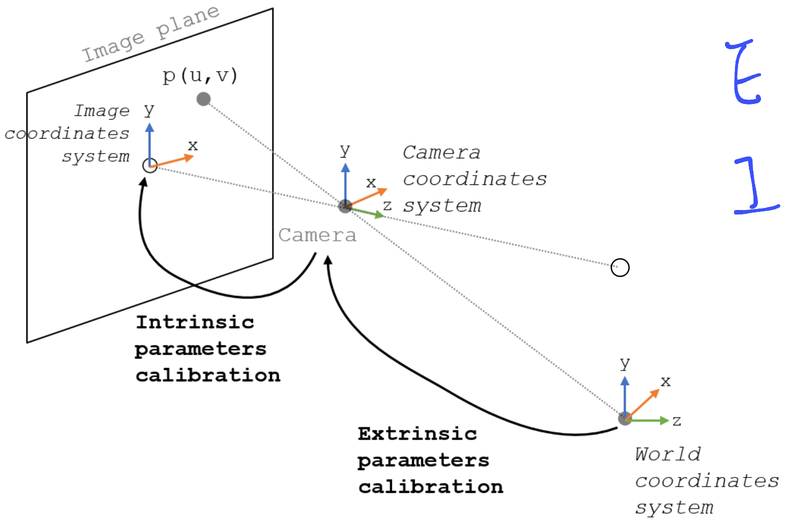
Camera Calibration

The goal of the camera calibration

- To find the 3×3 intrinsic matrix $\mathbf{K}$, the 3x3 rotation matrix $\mathbf{R}$, and the 3×1 translation vector $\mathbf{t}$ using a set of known 3D points (X_w, Y_w, Z_w) and their corresponding image coordinates $(\boldsymbol{u}, \boldsymbol{v})$.
- When we get the values of intrinsic and extrinsic parameters the camera is said to be calibrated.

In summary, camera calibration algorithm has the following inputs and outputs.

- **Inputs** : A collection of images with points whose 2D image coordinates and 3D world coordinates are known.
- **Outputs**: The 3×3 camera intrinsic matrix, the rotation and translation of each image.



Camera calibration

E : 3→ 3D

I : 3D→ (U,V)

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = R \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + t$$

$$\begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \begin{bmatrix} f_x & \gamma & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$$

Known (what we know)

Unknown (want to know)

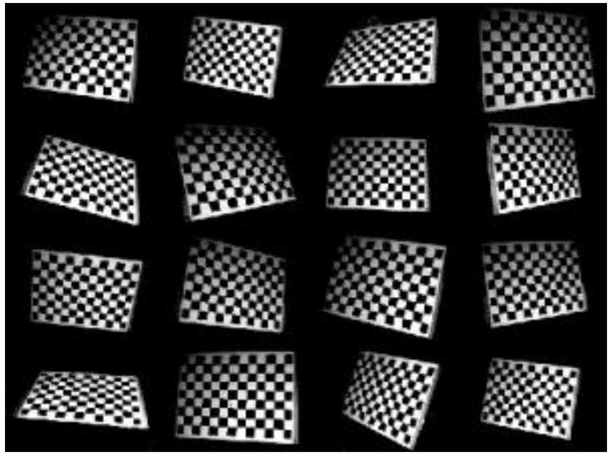
$$u = \frac{u'}{w'} \quad v = \frac{v'}{w'}$$

Equation of camera calibration (Up: extrinsic / Down: intrinsic)

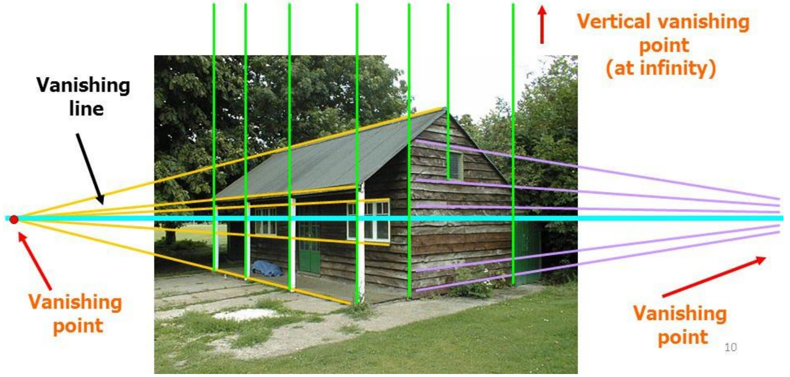
Copyright © AI Lab. All rights reserved.

Different Types of Camera Calibration Methods

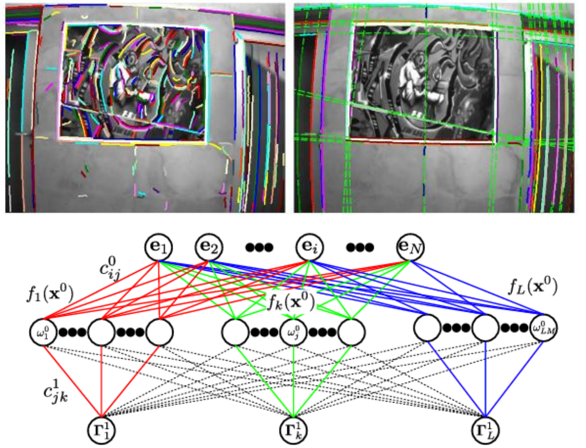
- **Major types of camera calibration methods**
  - **Calibration pattern**
    - When we have full control over the imaging process, calibration is performed by capturing images of a known pattern (e.g., checkerboard) from multiple viewpoints.
  - **Geometric clues**
    - When the scene contains geometric features like straight lines or vanishing points, these can be used to estimate camera parameters.
  - **Machine Learning based**
    - When there is very limited control (e.g., only a single image), camera calibration can still be estimated using machine learning methods.



Calibration pattern



Geometric clues



Machine learning based

Copyright © AI Lab. All rights reserved.

Basic of computer vision

1. Image processing
2. Machine learning

Copyright © AI Lab. All rights reserved.

Computer Vision

Definitions of Computer vision

- An interdisciplinary field that deals with how computers can be made to gain high-level understanding from digital images or videos
- Seeks to automate tasks that the human visual system can do
- Concerned with the automatic extraction, analysis and understanding of useful information from a single image or a sequence of images

Goal of computer vision

- To write computer programs that can interpret images

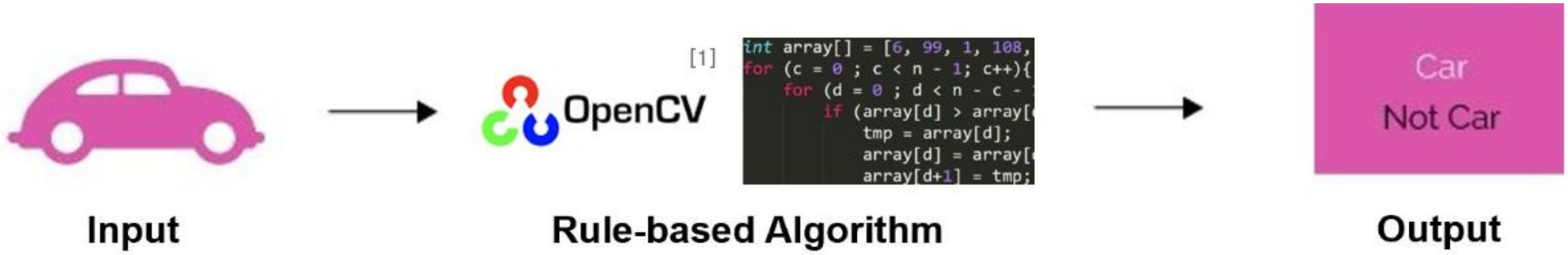
Computer graphics vs. computer vision

- Computer graphics (forward)
  - Physical model → image or video
- Computer vision (inverse problem)
  - Image or video → useful information

Copyright © AI Lab. All rights reserved.

Traditional Image Processing Methods

- **Utilize classical mathematics such as convolution, Fourier transforms, and morphological operations.**
- **Perform tasks including noise reduction, edge detection, and image enhancement without training data.**
- **Algorithms are deterministic, interpretable, and computationally efficient.**
- **Suitable for predictable environments requiring consistent performance.**



Copyright © AI Lab. All rights reserved.

Convert Color to Gray

- Grayscale conversion transforms a color image into a black and white image.
- This process reduces the amount of data while maintain important structural information.

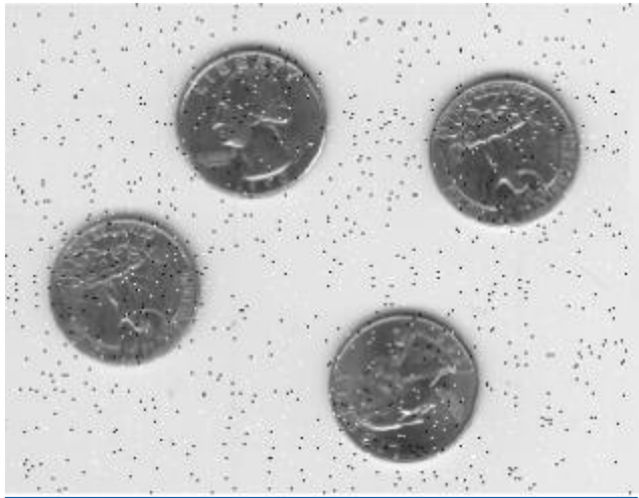
Rgb2Gray image

Img2gray image

Noise Removal

- Removing noise from images clarifies important information.
- Noise removal improves the accuracy of environment perception and supports safe driving

Original image



Imnoised image

Medfilt2 image

Copyright © AI Lab. All rights reserved.

Copyright © AI Lab. All rights reserved.

Invert Image

- Image inversion increases contrast by reversing the brightness levels of an image, typically used under specific conditions to enhance visibility.

- This can help identify objects more clearly in low-light conditions.

Original image

Imcomplemented image

Adjust Image Contrast

- Adjusting image contrast enhances feature recognition.
- This can help identify objects more clearly in low-light conditions.


Original image


Imadjusted image

Copyright © AI Lab. All rights reserved.

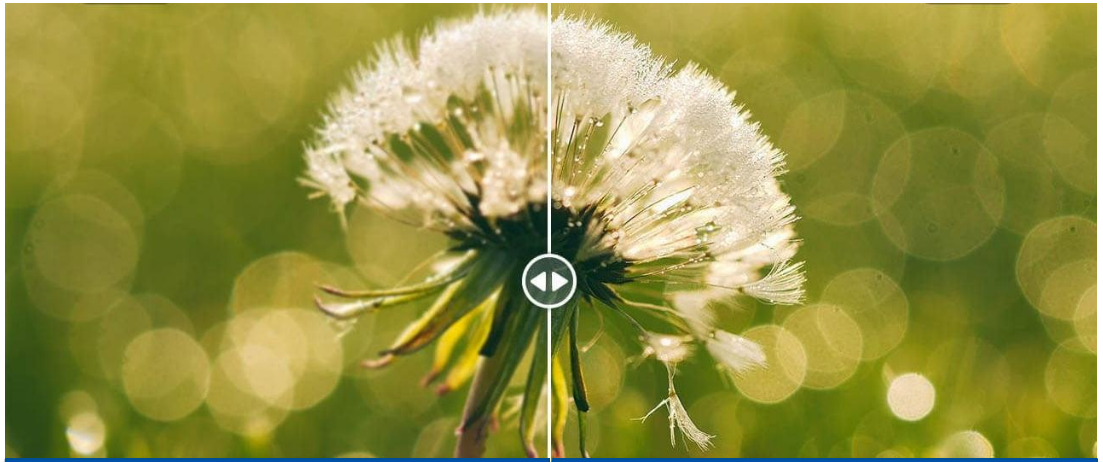
Copyright © AI Lab. All rights reserved.

Blurring & Sharpening

- Blurring reduces image details to minimize noise, while sharpening enhances the edges of images.

- These techniques optimize image analysis by emphasizing important parts or excluding less important information.

Original image | Imgaussfiltered image



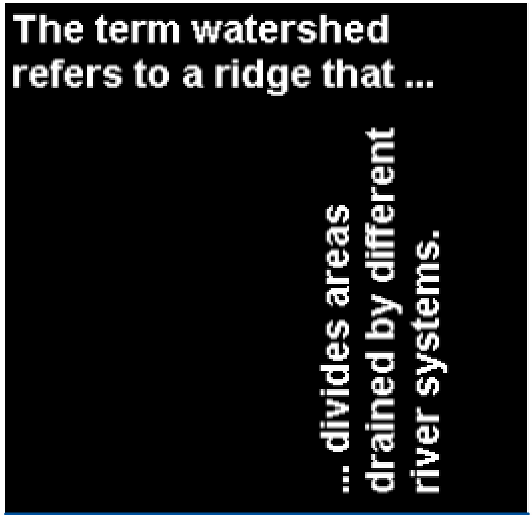
Original image | Sharpened image

Copyright © AI Lab. All rights reserved.

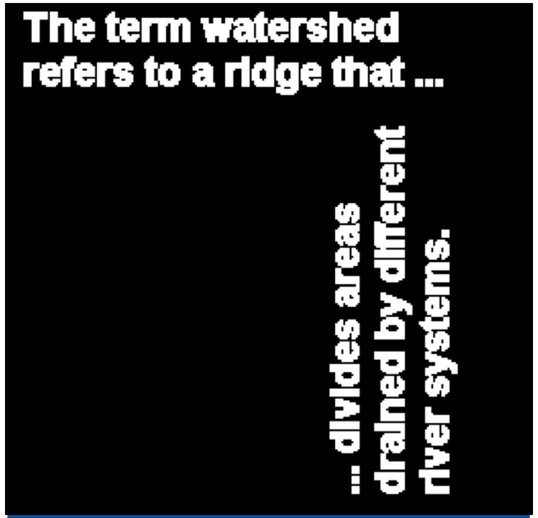
Morphology

- Morphological operations manipulate the structure of an image to emphasize morphological features.

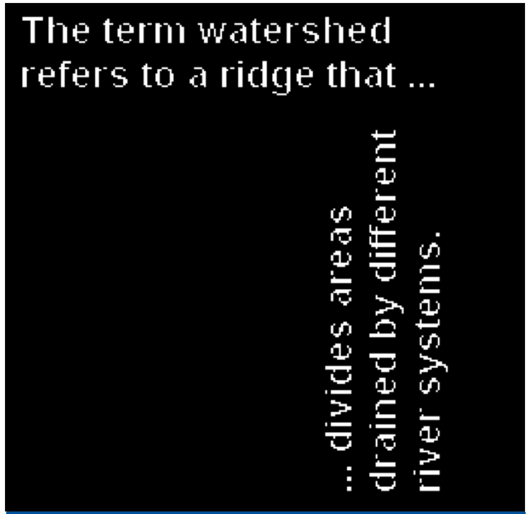
- Dilation expands the boundaries of objects, and erosion shrinks them.



Original image



Dilated image



Eroded image

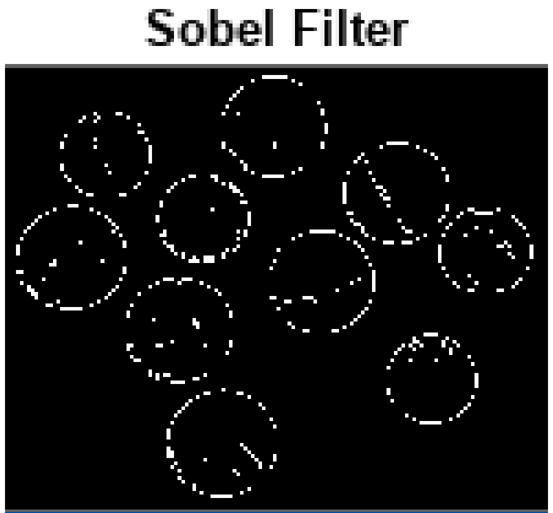
Copyright © AI Lab. All rights reserved.

Edge Detection

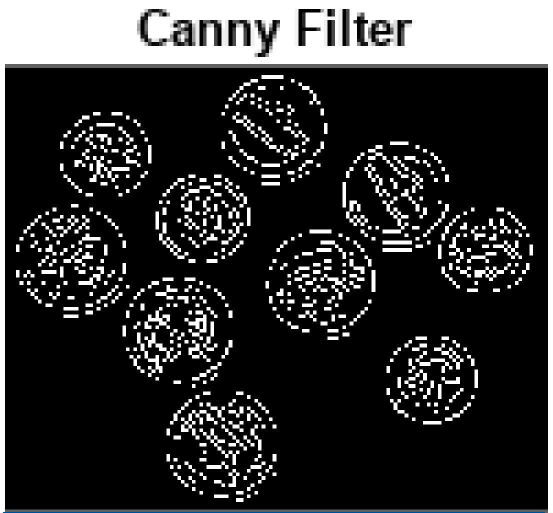
- **Edge detection identifies the boundaries of objects within an image.**

- **In autonomous vehicles, it's crucial for identifying lanes, road boundaries, and separations from other vehicles.**

Original image



Sobel filtered image

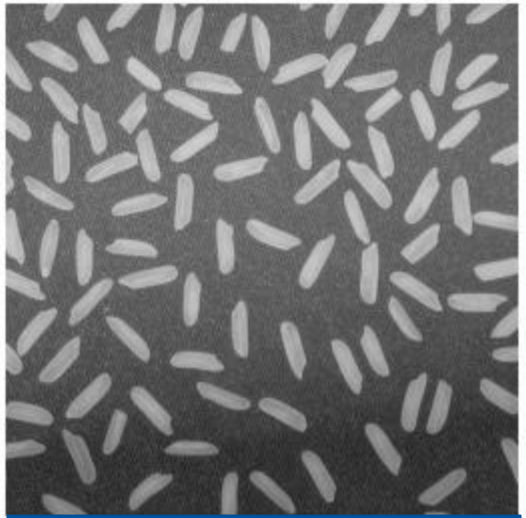


Canny filtered image

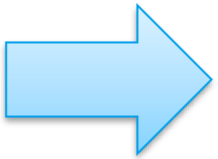
Contour Detection

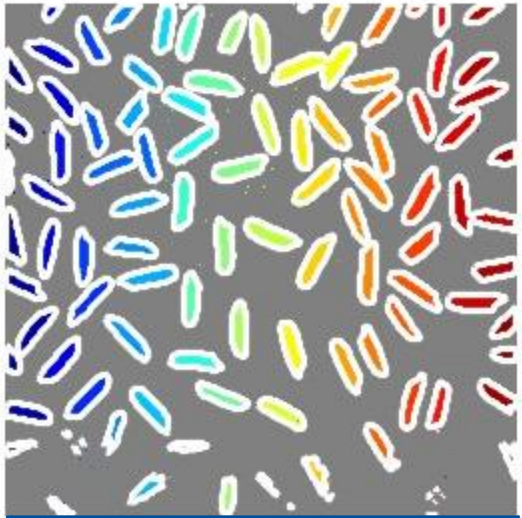
Copyright © AI Lab. All rights reserved.

- Contour detection finds continuous boundaries in an image.
- In autonomous vehicles, it's essential for distinguishing road structures, obstacles, and pedestrians.
- Analyzing the shape and size of objects contributes to more accurate environmental perception and decision-making.



Original image



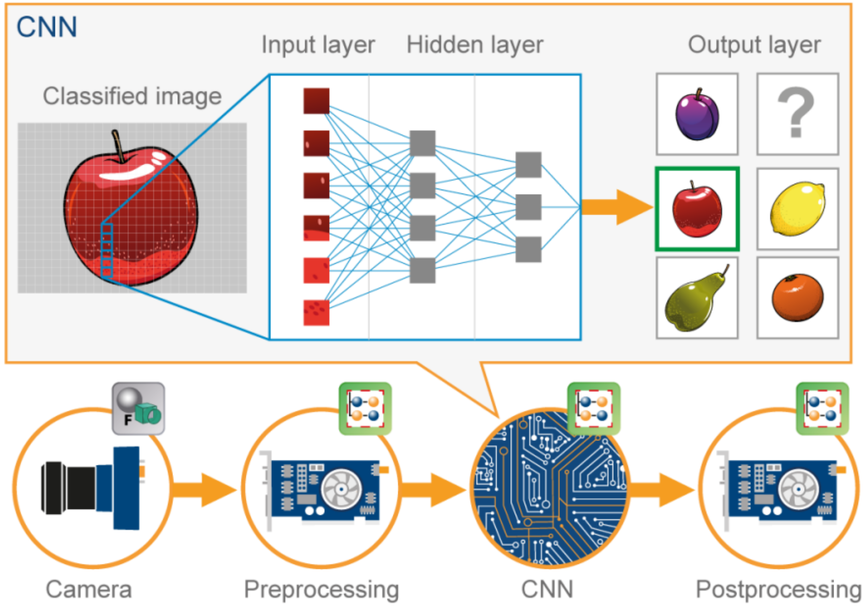


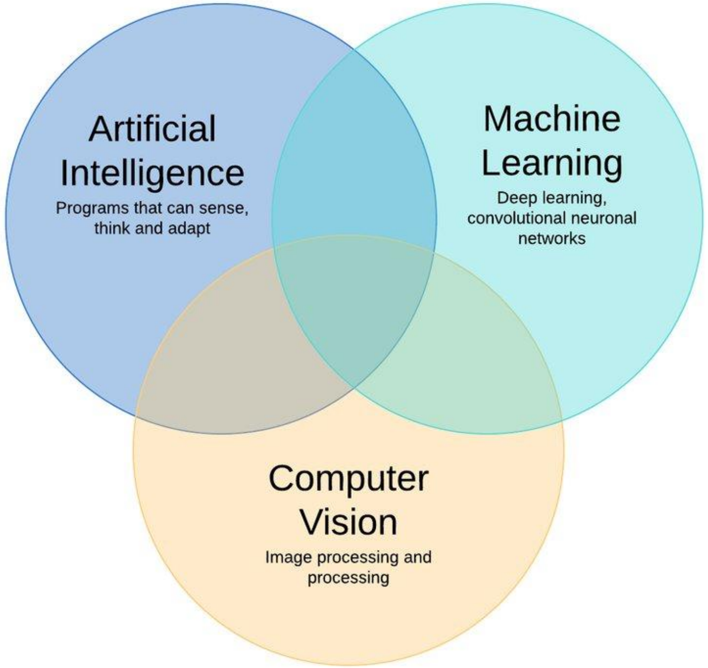
Contour detected image

Copyright © AI Lab. All rights reserved.

Machine Learning in Vision

- Utilizes algorithms like Convolutional Neural Networks (CNNs) to automatically extract features from images.
- Enables complex tasks such as object detection, image classification, and semantic segmentation.
- Learns from large datasets to improve accuracy and adaptability over time.
- Widely applied in areas like autonomous vehicles, medical imaging, and facial recognition systems.





Copyright © AI Lab. All rights reserved.

Automotive Vision Applications

Copyright © AI Lab. All rights reserved.

HDMap Construction

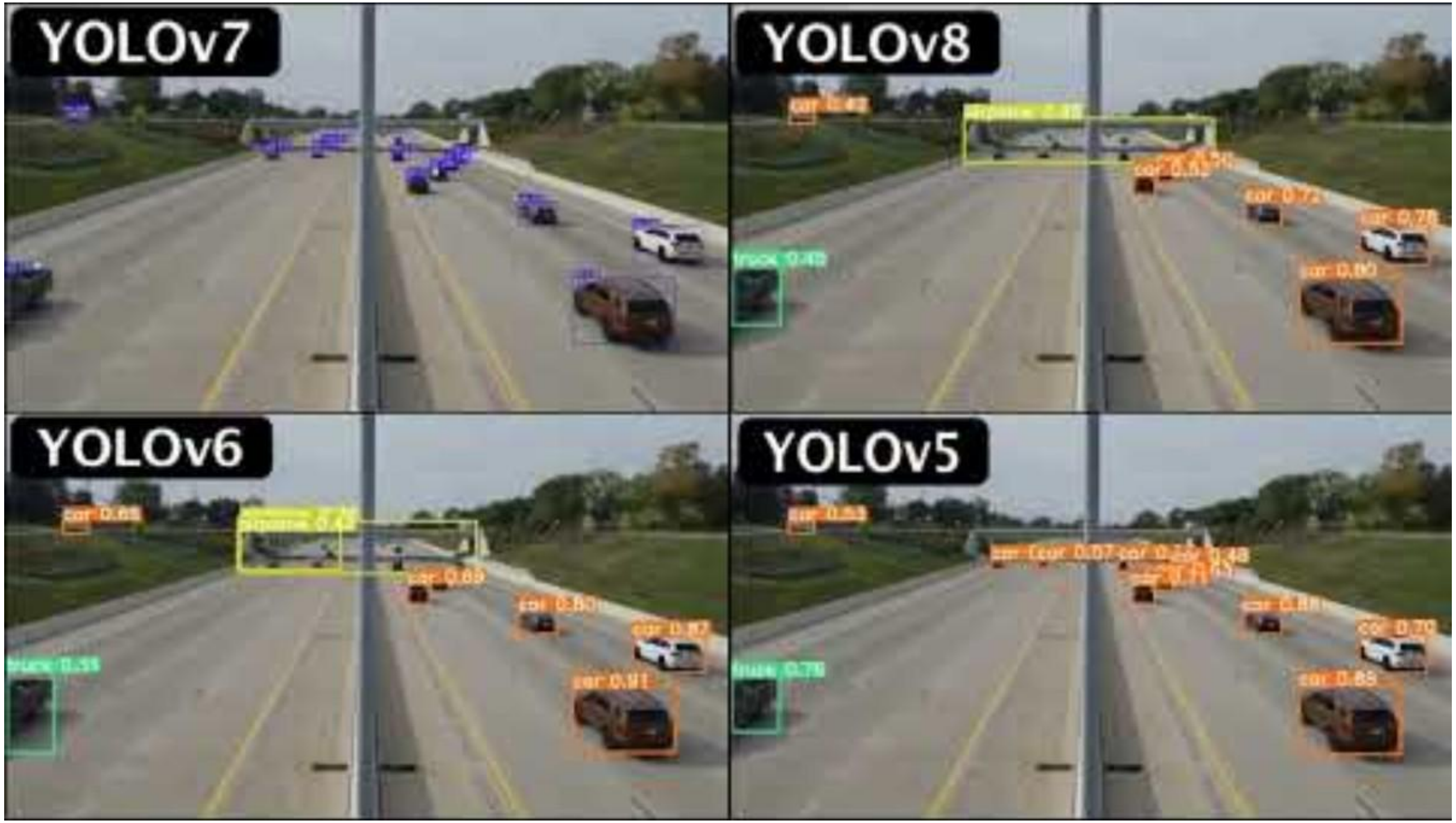
HDMap

- Highly detailed map used in autonomous driving, containing precise information such as lane divider, pedestrian crossing, road boundaries, etc. to support accurate localization and safe path planning.

Copyright © AI Lab. All rights reserved.

2D Object Detection

2D Object Detection using YOLO(You Only Look Once) model

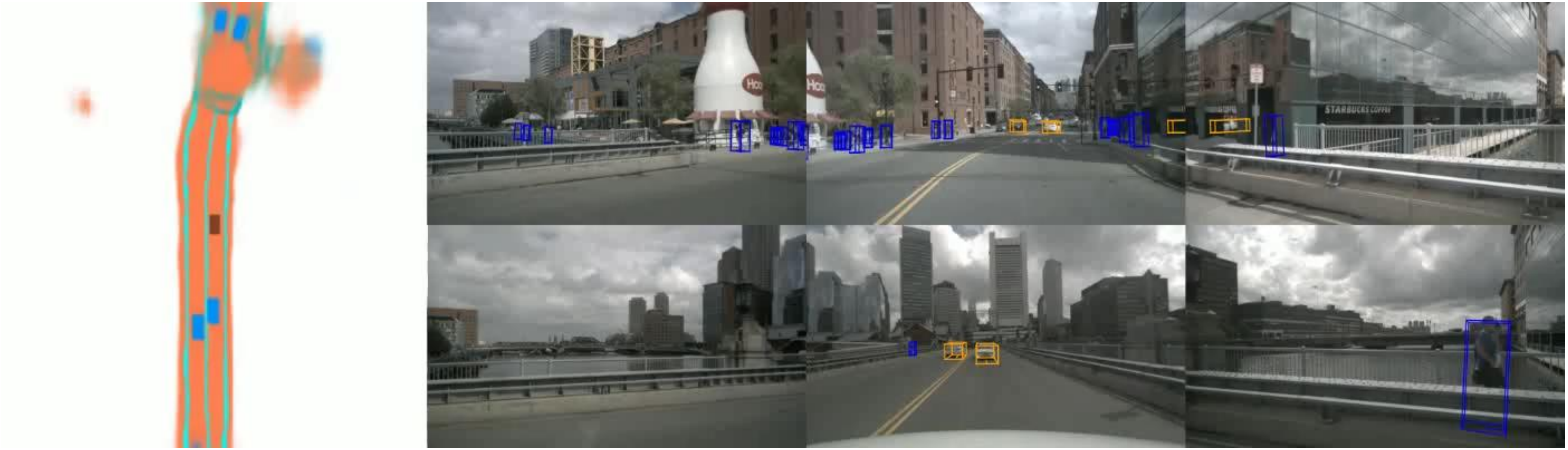


Copyright © AI Lab. All rights reserved.

3D Object Detection

Camera-based 3D Object Detection

- BEVFormer: Learning Bird's-Eye-View Representation from Multi-Camera Images via Spatiotemporal Transformers, ECCV 2022

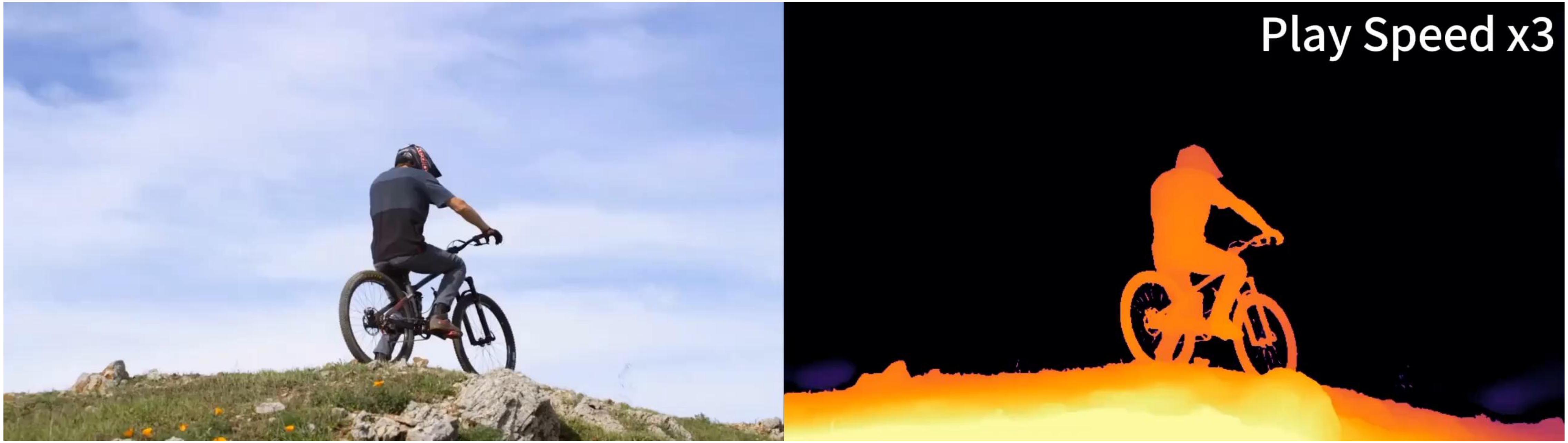


Copyright © AI Lab. All rights reserved.

Depth Estimation

Video Depth Estimation

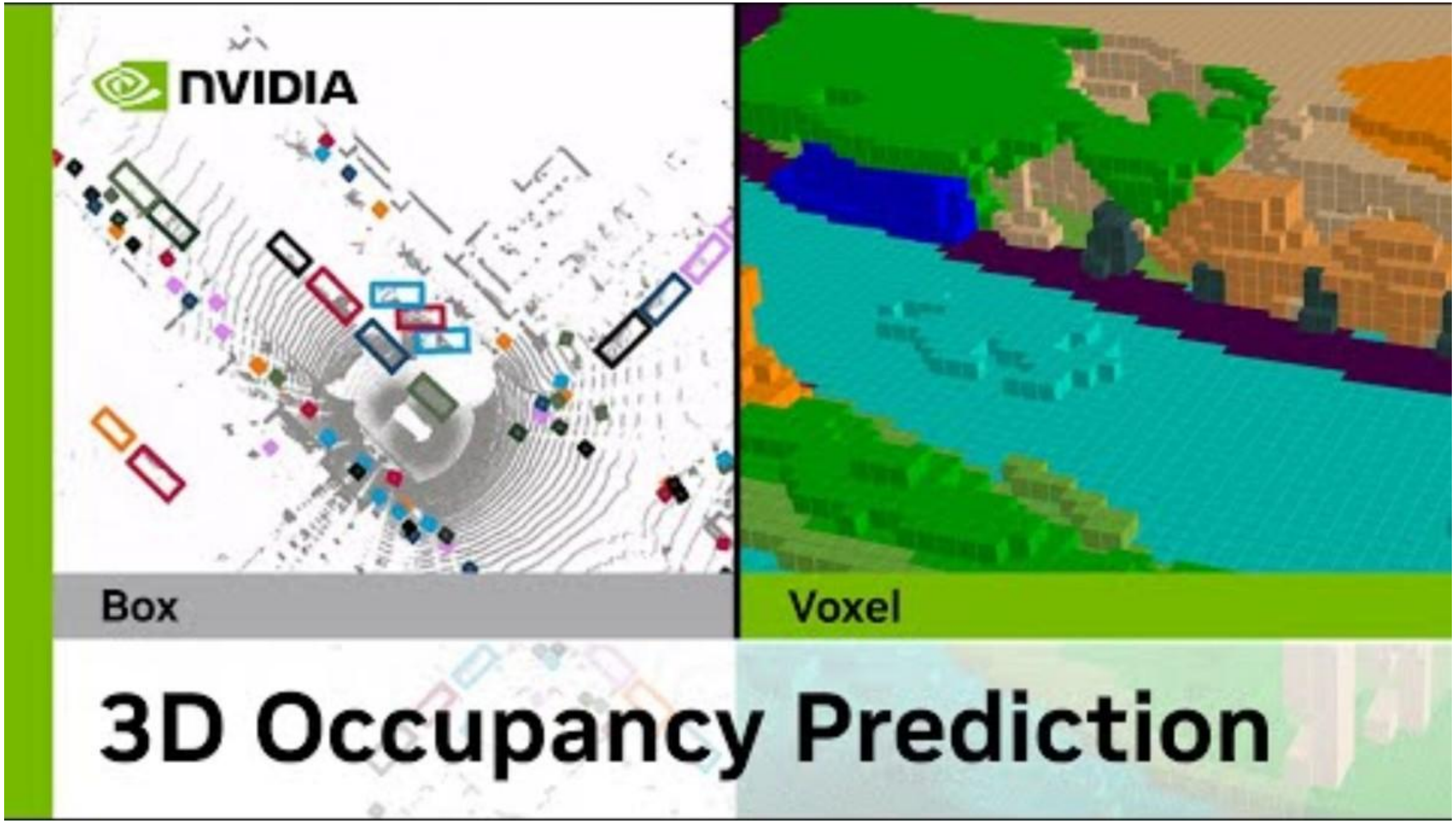
- Video Depth Anything: Consistent Depth Estimation for Super-Long Videos, CVPR 2025



Copyright © AI Lab. All rights reserved.

3D Occupancy Prediction

■ 3D Occupancy Prediction using 6 cameras



Semantic Segmentation

Copyright © AI Lab. All rights reserved.

- Use semantic segmentation to classify each pixel into one of 19 categories

Copyright © AI Lab. All rights reserved.

Panoptic Segmentation

- Panoptic Segmentation (Semantic Segmentation and Instance Segmentation)

Copyright © AI Lab. All rights reserved.

Visual SLAM

- **Visual simultaneous localization and mapping (SLAM)**
  - SLAM means estimating pose of the sensor while creating a map of the environment
  - Visual SLAM uses only vision(camera) sensor for SLAM

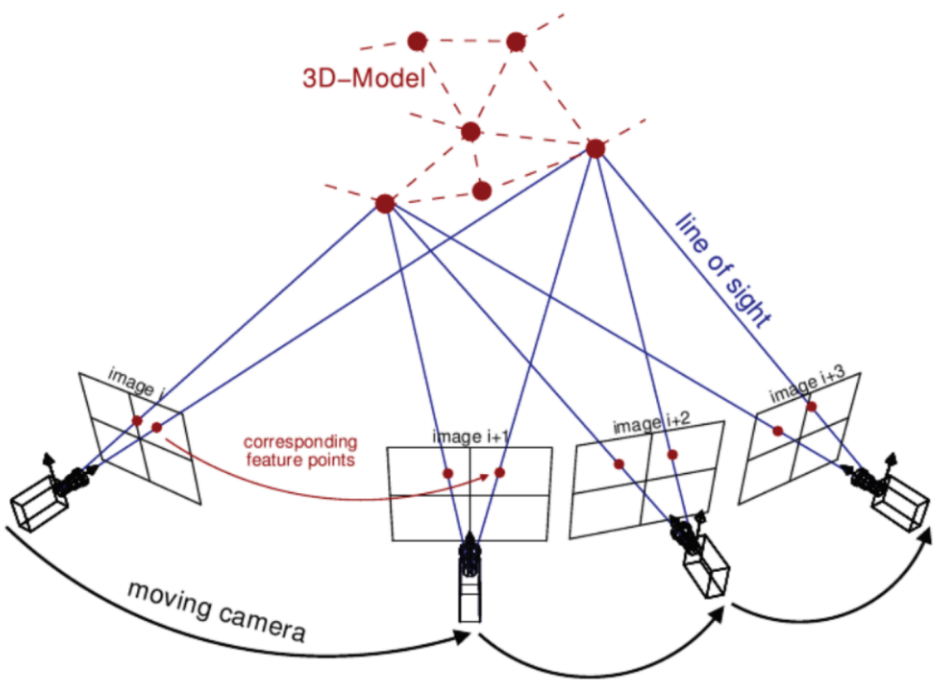
ORB SLAM Demo Video

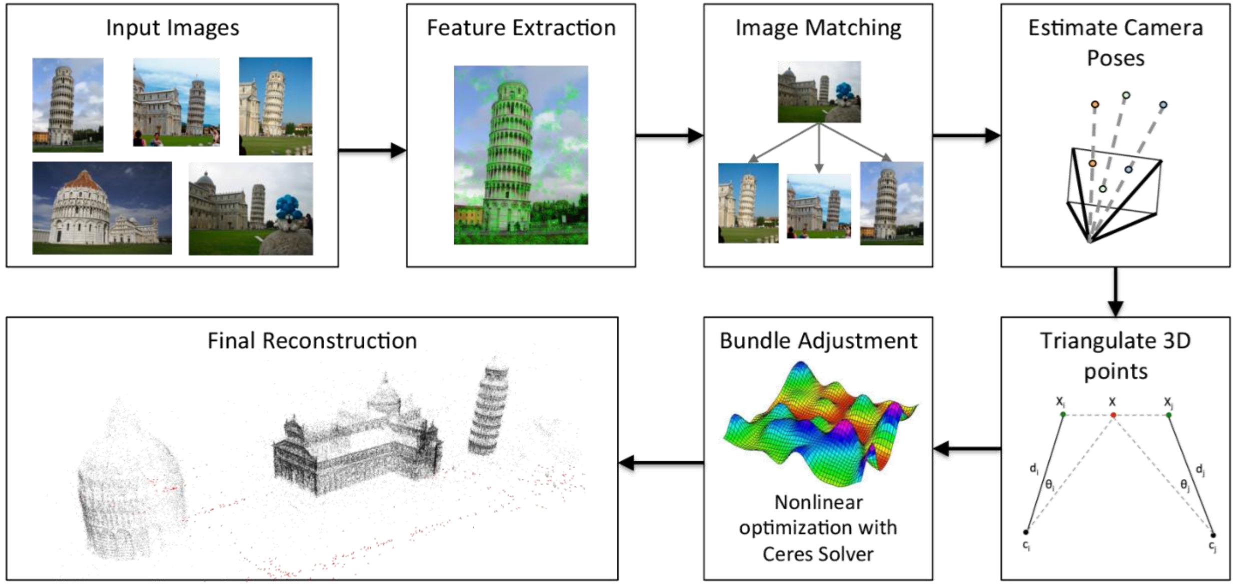
Copyright © AI Lab. All rights reserved.

Structure from Motion (SfM)

SfM with monocular camera

- From 2D-image to 3D-model with monocular camera
- A photogrammetric range imaging technique for estimating three-dimensional structures from two-dimensional image sequences that may be coupled with local motion signals.





Copyright © AI Lab. All rights reserved.

Structure from Motion (SfM) Demo

- SfM (Structure from Motion) demo video

Copyright © AI Lab. All rights reserved.

View Synthesis

■ Construct 3D model using single images

- ▶ **View synthesis** aims to create new views of a specific subject starting from a number of pictures taken from given point of views.

3D Gaussian Splatting

Copyright © AI Lab. All rights reserved.

Video Generation

- **Understanding and modeling real-world driving scenarios from images**
  - **Generate** realistic driving scenarios while offering fine-grained control over ego-vehicle behavior and scene features.

Gaia-1 Demo Video

Copyright © AI Lab. All rights reserved.

Mobileye

Copyright © AI Lab. All rights reserved.

Tesla Autopilot

Copyright © AI Lab. All rights reserved.

센서원리

센서 이론

실습

활용!!

연습!!

THANK YOU
FOR YOUR ATTENTION