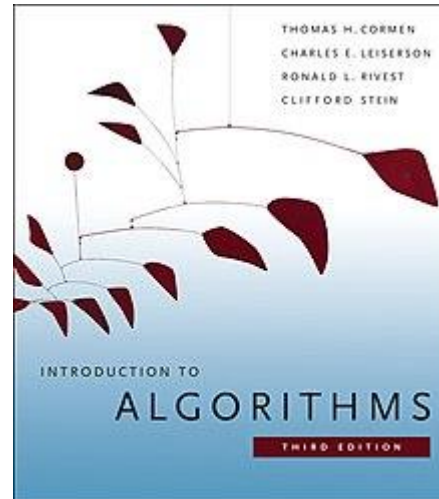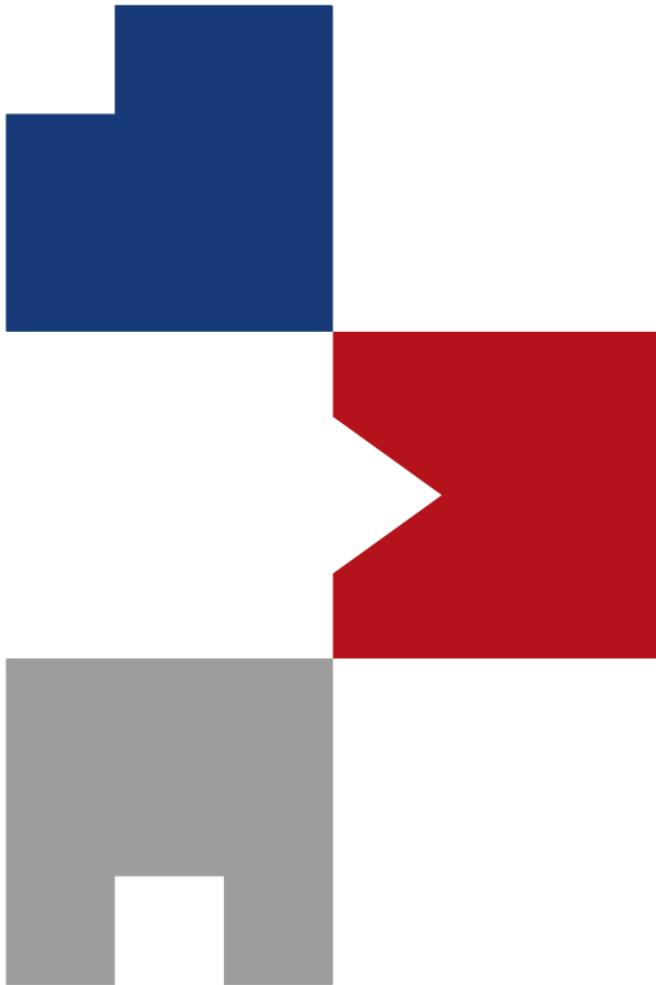# Introduction to 3D Vision

1312 pages

Sunglok Choi, Assistant Professor, Ph.D.

Computer Science and Engineering Department, SEOULTECH

sunglok@seoultech.ac.kr | https://mint-lab.github.io/

An Invitation
~~Introduction~~ to 3D Vision
: A Tutorial for Everyone

Sunglok Choi, Assistant Professor, Ph.D.

Computer Science and Engineering Department, SEOULTECH

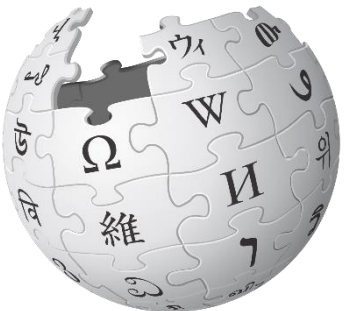**sunglok@seoultech.ac.kr | https://mint-lab.github.io/**

What is Computer Vision?

- **Computer vision is an interdisciplinary field that deals with how computers can be made to gain high-level understanding from digital images or videos.**
- From the perspective of engineering, it seeks to automate tasks that the human visual system can do.[1][2][3]
- "Computer vision is concerned with the automatic extraction, analysis and understanding of useful information from a single image or a sequence of images.
- It involves the development of a theoretical and algorithmic basis to achieve automatic visual understanding."[9]

Reference: Wikipedia (Computer Vision)

What is Computer Vision?

- **Computer vision is an interdisciplinary field that deals with how computers can be made to gain high-level understanding from digital images or videos.**

- From the perspective of engineering, it seeks to automate tasks that the human visual system can do.[1][2][3]

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

- It involves the development of a theoretical and algorithmic basis to achieve automatic visual understanding."[9]
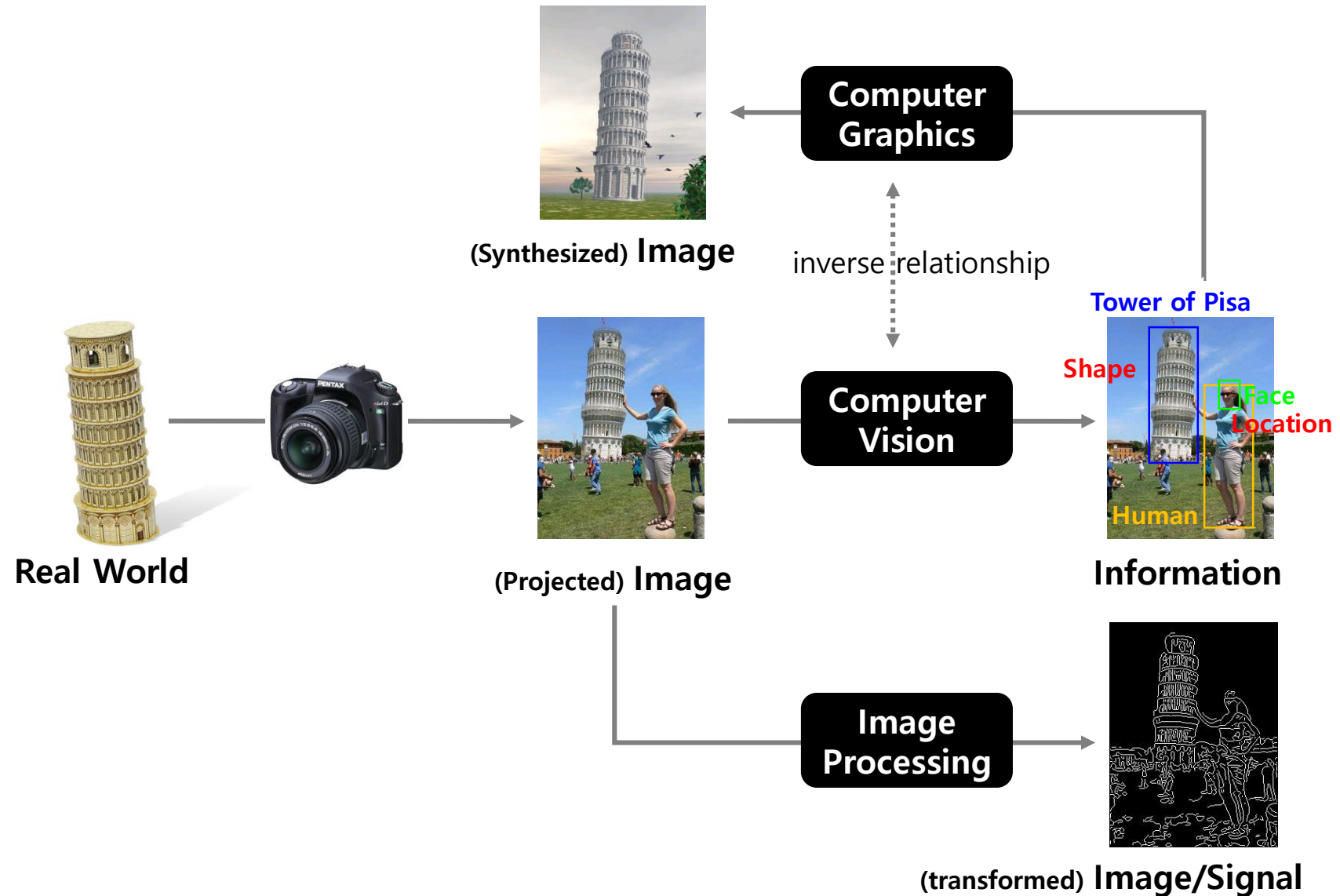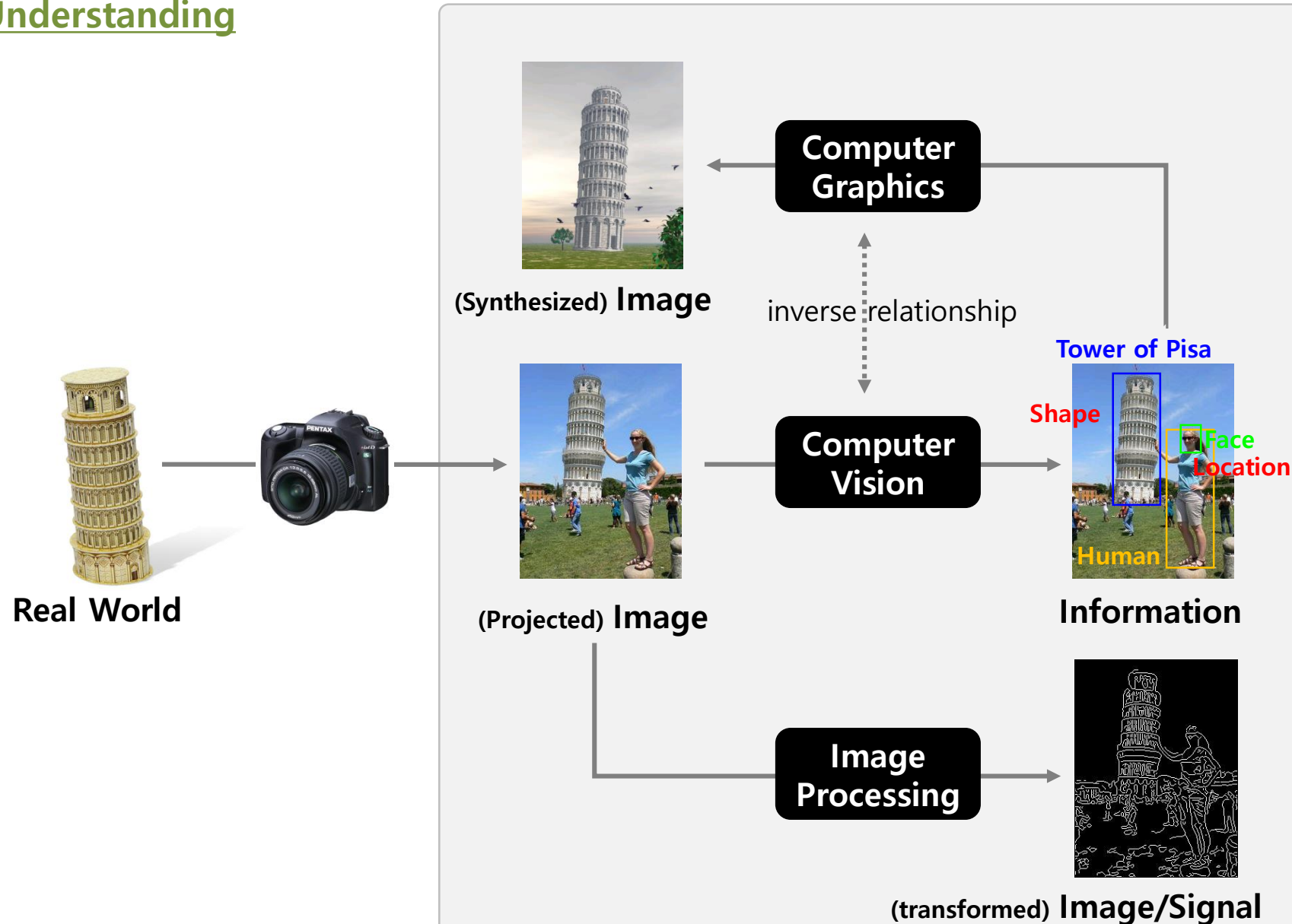
What is Computer Vision?

Image Understanding

(Synthesized) **Image**

Computer Graphics

inverse relationship

Computer Vision

Image Processing

Real World

(Projected) **Image**

Information

Tower of Pisa

Shape

Face

Location

Human

(transformed) **Image/Signal**

What is **Computer** **Vision**?

Image **Understanding**

Computer Vision

(Synthesized) **Image**

inverse relationship

Computer Graphics

Real World

(Projected) **Image**

Computer Vision

Tower of Pisa

Shape

Face

Location

Human

Information

Image Processing
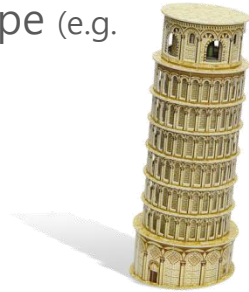
(transformed) **Image/Signal**

What is Computer Vision?

Computer Vision

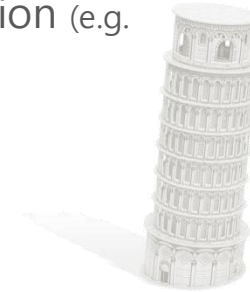**What** is it?

- Label (e.g. Tower of Pisa)
- Shape (e.g.   )

Where am I?

- Place (e.g. Piazza del Duomo, Pisa, Italy)
- Location (e.g.   )

(84, 10, 18) [m]
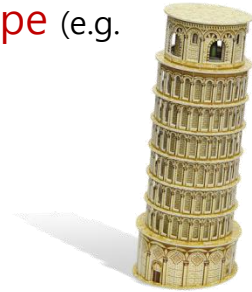
What is 3D Vision?

Visual Geometry

Multiple View Geometry

Geometric Vision

Computer Vision

What is it?

- Label (e.g. Tower of Pisa)

- Shape (e.g.   )

Where am I?

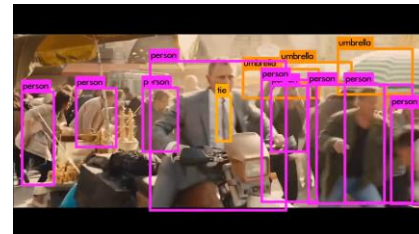- Place (e.g. Piazza del Duomo, Pisa, Italy)

- Location (e.g.   )
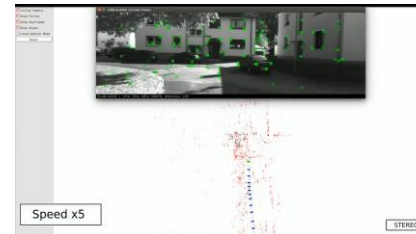
(84, 10, 18) [m]

Recognition Problems v.s. Reconstruction Problems

Stanford CS231n:

CNN for Visual Recognition

Stanford CS231A:

**Computer Vision,
From 3D Reconstruction to Recognition**

YOLO v2 (2016)   ORB-SLAM2 (2016)

What is 3D Vision?

image

depth image, range data, point cloud, polygon mesh, ...

3D Vision v.s. **3D Data Processing**

Perspective Camera

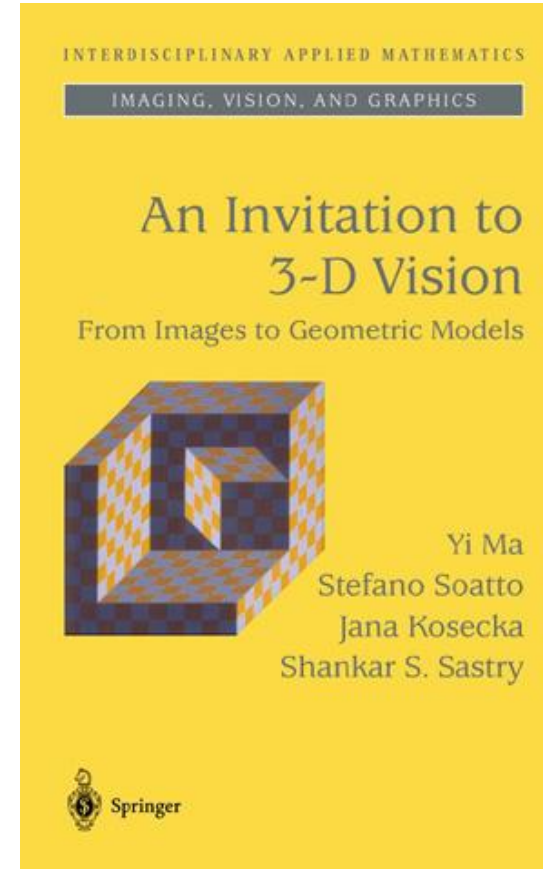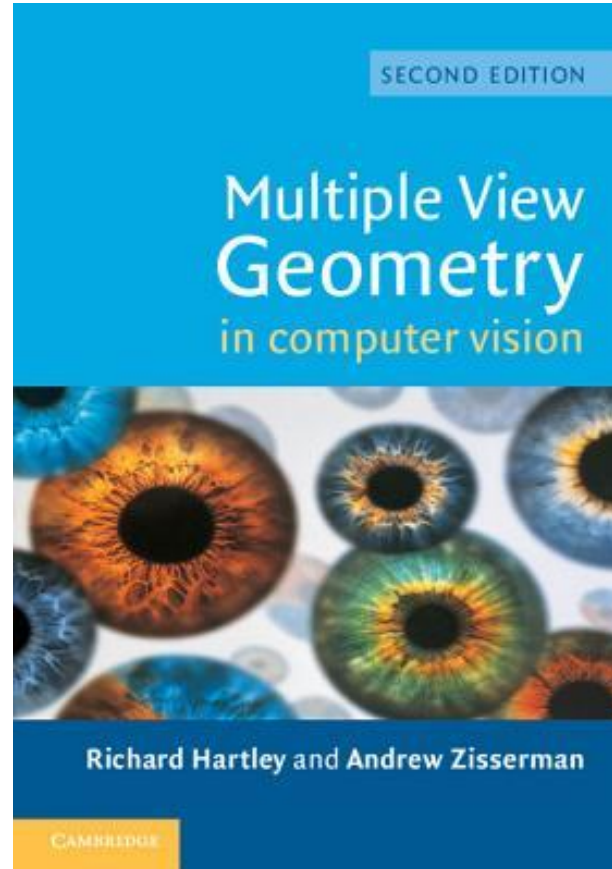**Omni-directional Camera**

RGB-D Camera
(Stereo, Structured Light, ToF, Light Field)

Range Sensor
(LiDAR, RADAR)

What is 3D Vision?

- Reference books

What is **3D Vision?**

- All example codes are available at https://github.com/mint-lab/3dv_tutorial.

  - All example codes are mostly **less than 100 lines** and based on **recent OpenCV (> 3.0.0)**.

  - Note) OpenCV (Open Source Computer Vision)
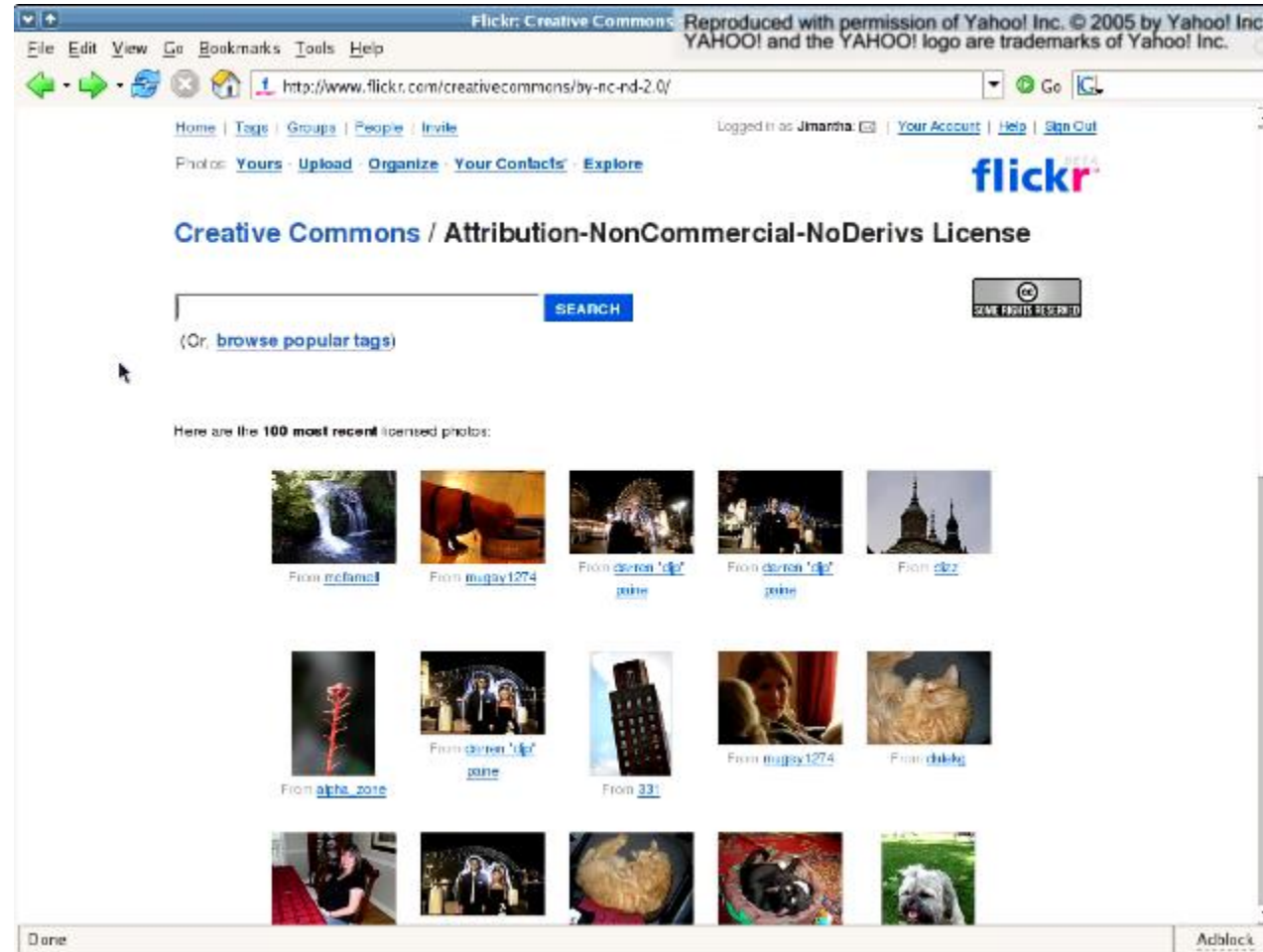
OpenCV v4.8.0 main modules:
- core. Core functionality
- imgproc. Image Processing
- imgcodecs. Image file reading and writing
- videoio. Video I/O
- highgui. High-level GUI
- video. Video Analysis
- calib3d. Camera Calibration and 3D Reconstruction
- features2d. 2D Features Framework
- objdetect. Object Detection
- dnn. Deep Neural Network module
- ml. Machine Learning
- flann. Clustering and Search in Multi-Dimensional Spaces
- photo. Computational Photography
- stitching. Images stitching
- gapi. Graph API

OpenCV v5.0.0-pre main modules:
- core. Core functionality
- imgproc. Image Processing
- imgcodecs. Image file reading and writing
- videoio. Video I/O
- highgui. High-level GUI
- video. Video Analysis
- 3d. 3d
- stereo. Stereo Correspondence
- features2d. 2D Features Framework
- calib. Camera Calibration
- objdetect. Object Detection
- dnn. Deep Neural Network module
- ml. Machine Learning
- flann. Clustering and Search in Multi-Dimensional Spaces
- photo. Computational Photography
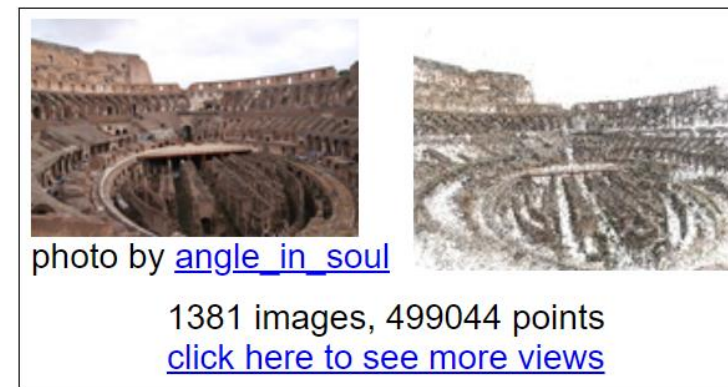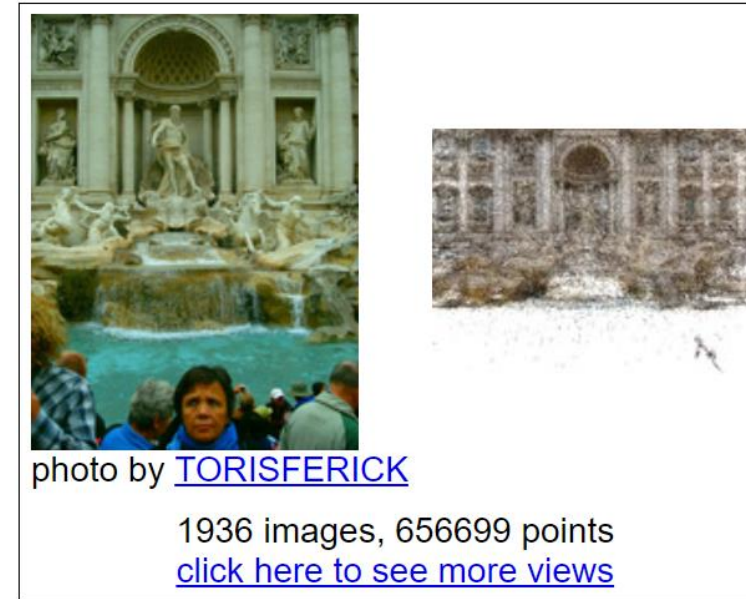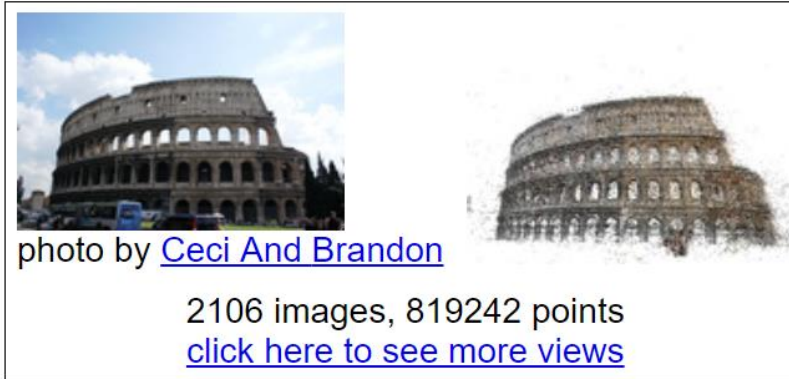- stitching. Images stitching
- gapi. Graph API

Applications) Photo Browsing

- [Photo Tourism](#) (2006)

Reference: Snavely et al., **Photo Tourism: Exploring Photo Collections in 3D**, SIGGRAPH, 2006

12

Applications) 3D Reconstruction

- [Building Rome in a Day](#) (2009)

photo by Ceci And Brandon
2106 images, 819242 points
click here to see more views

photo by TORISFERICK
1936 images, 656699 points
click here to see more views

photo by act2win
1815 images, 422593 points
click here to see more views

photo by angle_in_soul
1381 images, 499044 points
click here to see more views

Reference: Agarwal et al., **Building Rome in a Day**, ICCV, 2009

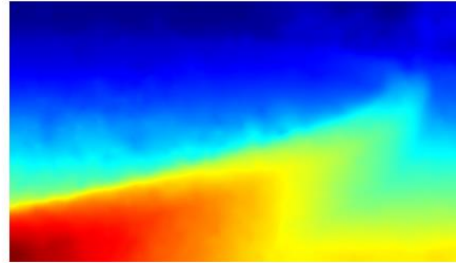# Applications) Depth Estimation from Cellular Phones

- [Structure from Small Motion](#) (SfSM; 2015)

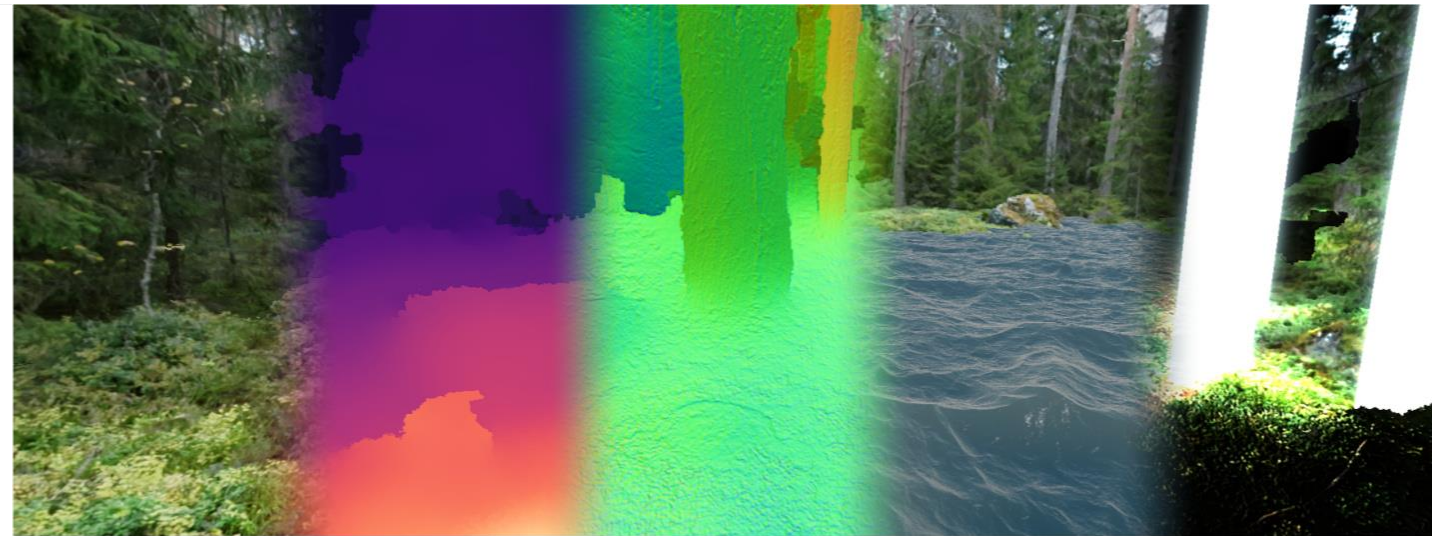(a) Reference images   (b) SfSM results   (c) Depth maps   (d) Our 3D meshes

- [Casual 3D Photography](#) (2017)

Casual 3D photo capture   Color — Depth — Normal map   Geometry-aware — Lighting

Reconstruction   Example Effects

Reference: Im et al., **High Quality Structure from Small Motion for Rolling Shutter Cameras**, ICCV, 2015
Reference: Hedman et al., **Casual 3D Photography**, SIGGRAPH Asia, 2017

14

Applications) Real-time Visual SLAM

- **ORB-SLAM** (2014)

Reference: Mur-Artal et al., **ORB-SLAM: A Versatile and Accurate Monocular SLAM System**, T-RO, 2015
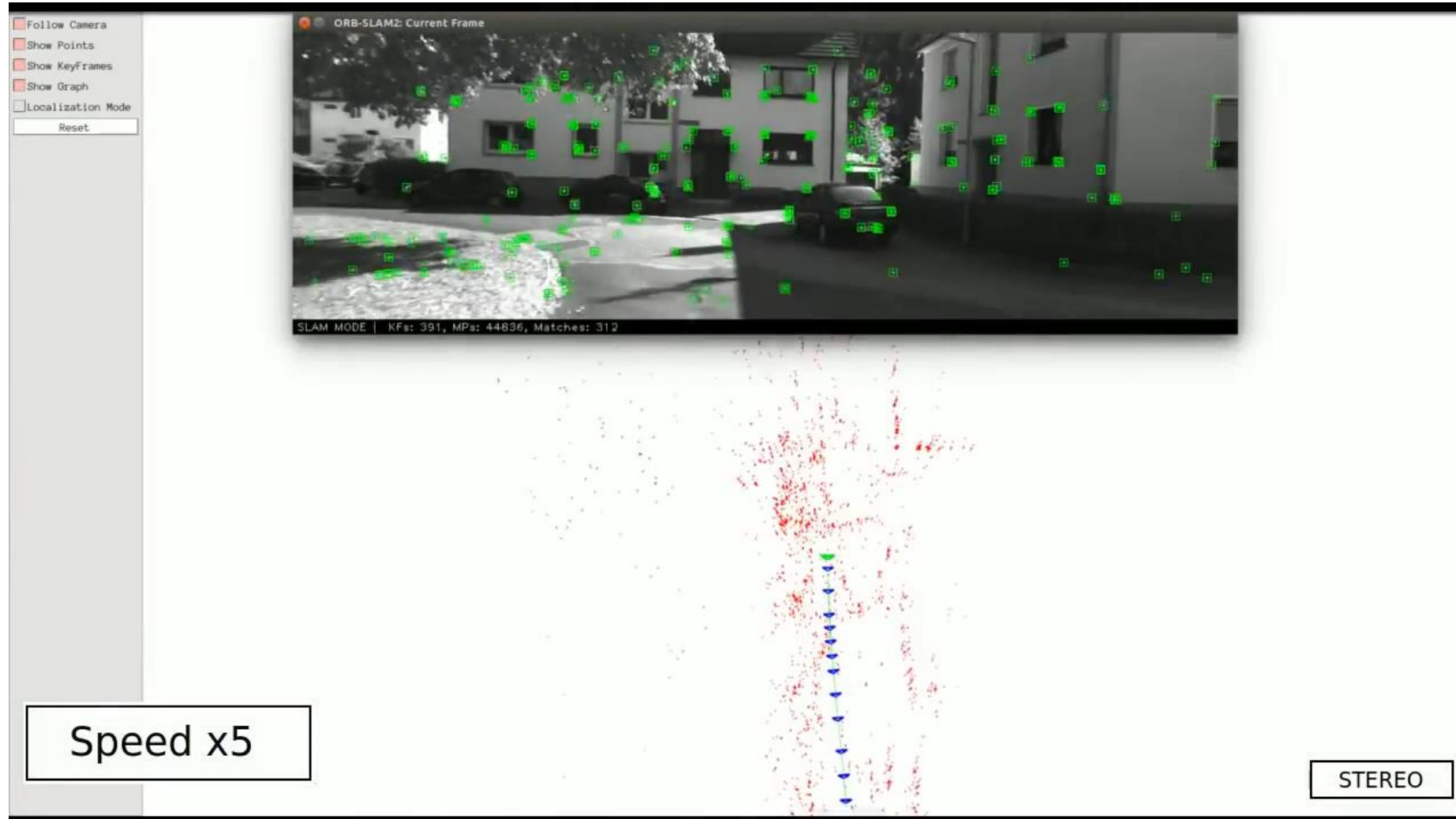
Applications) Augmented Reality

- [PTAM: Parallel Tracking and Mapping](#) (2007)

Reference: Klein and Murray, **Parallel Tracking and Mapping for Small AR Workspaces**, ISMAR, 2007
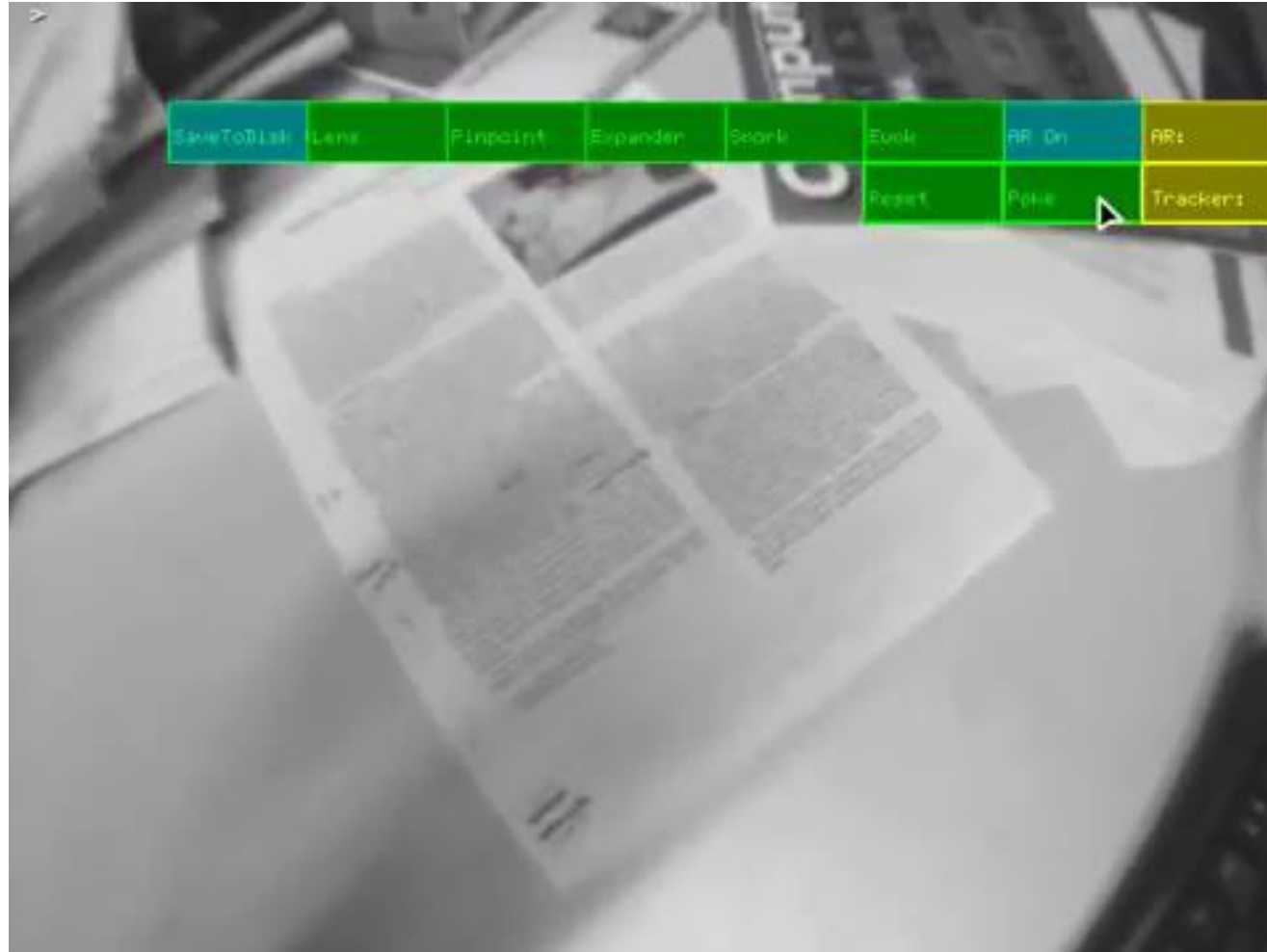
Applications) Virtual Reality

- [Oculus Quest](#) (2019)

Image: [TechSpot](#)

Applications) Mixed Reality

- [Microsoft Hololens 2](#) (2019)
  - Head tracking: 4 x visible light cameras

HoloLens image sensors

depth camera

short and long-throw IR illuminators

4 gray-scale cameras

Color video camera
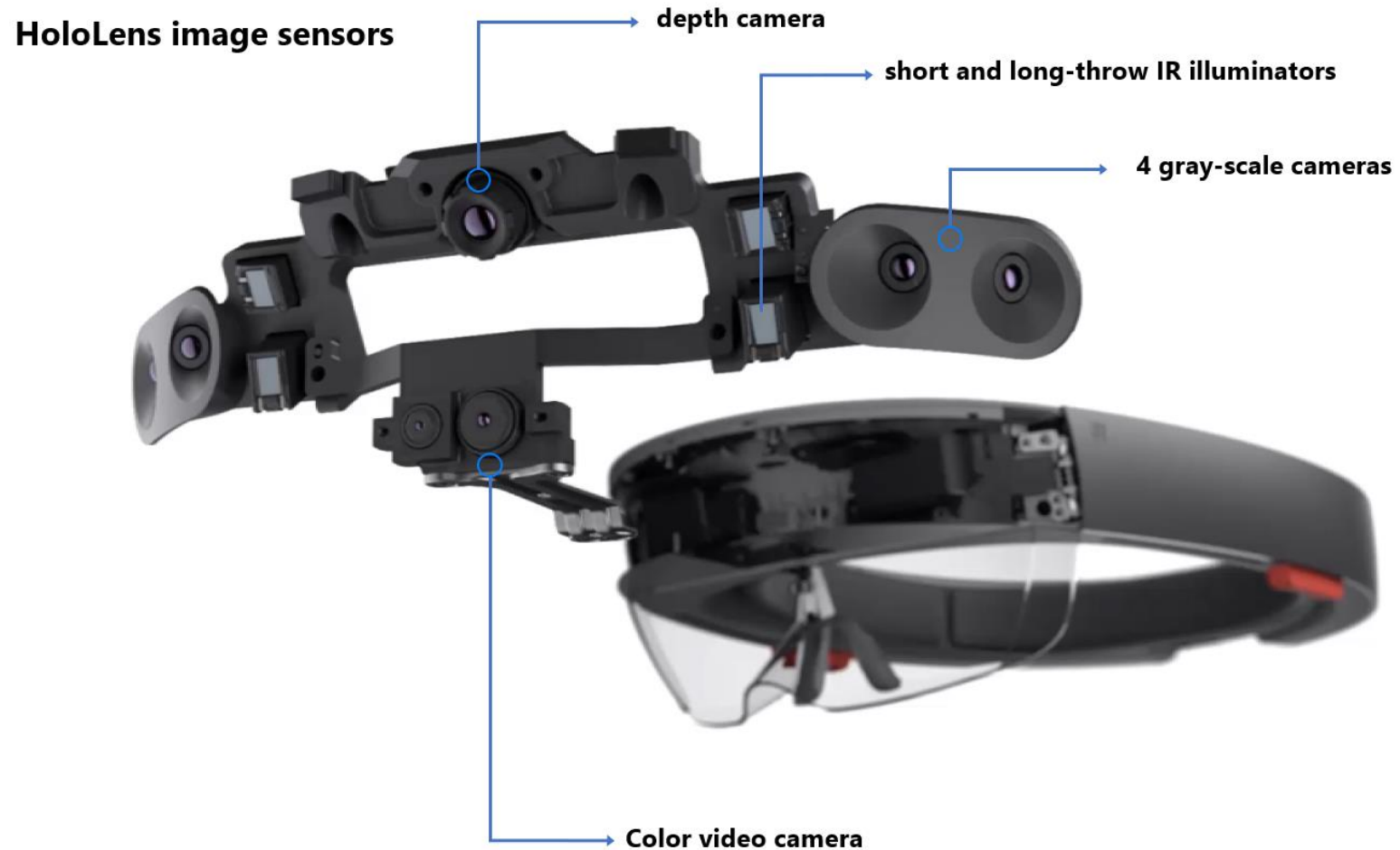
Summary

- **What is Computer Vision?**

- **What is 3D Vision?**

  – What? Recognition problem vs. Reconstruction problem

    • Note) Generation problem vs. Reconstruction problem

  – Why? Applications

Next Topics

- **Single-view Geometry**

- **Two-view Geometry**

- **Solving Equations**

- **Finding Correspondence**

- **Multiple-view Geometry**

- **Bayesian Filtering**

- **Visual SLAM and Odometry**