Introduction
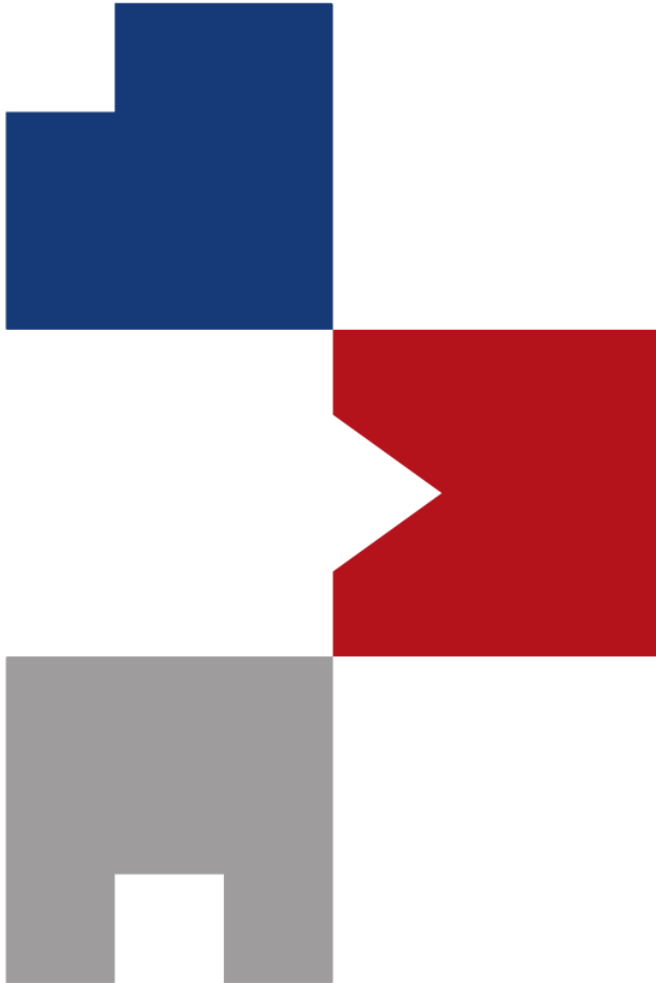
Sunglok Choi, Assistant Professor, Ph.D.

Computer Science and Engineering Department, SEOULTECH

sunglok@seoultech.ac.kr | https://mint-lab.github.io/
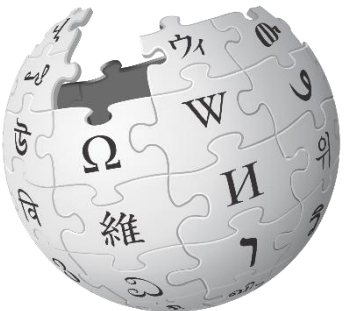
What is Computer Vision?

- **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]
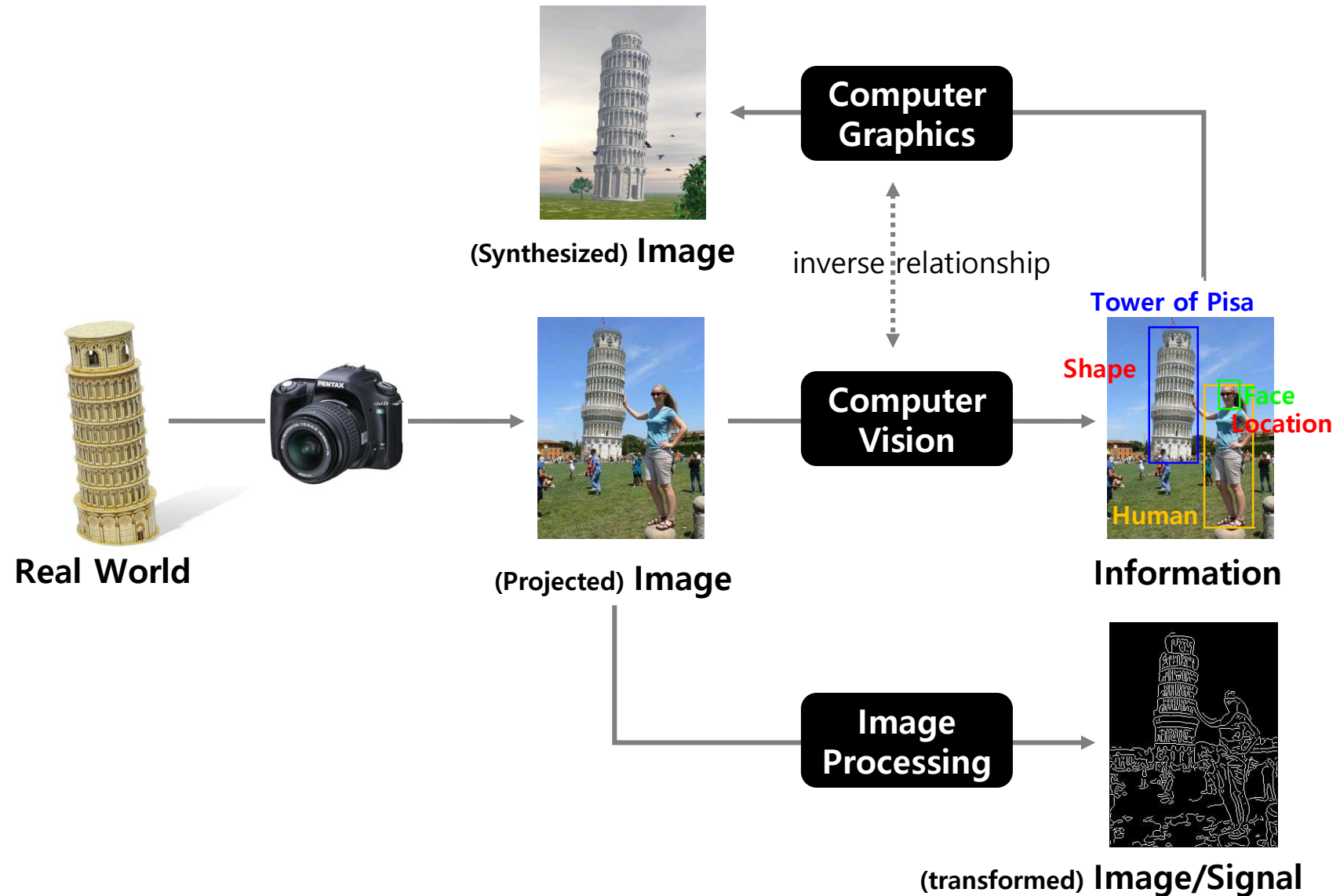

Reference: Wikipedia

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]
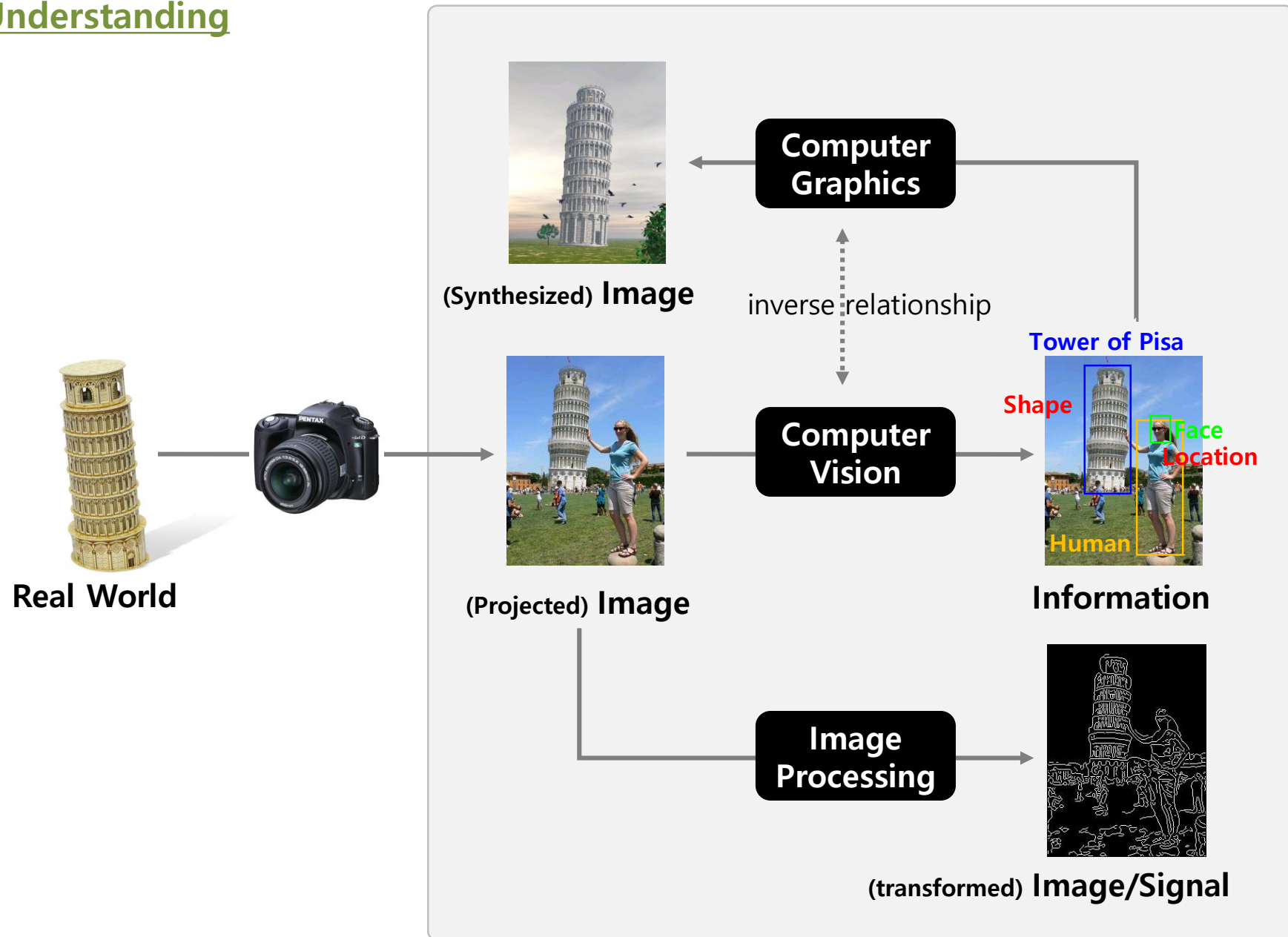
Reference: Wikipedia

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

Image **Understanding**

(Synthesized) Image

Computer Graphics

inverse relationship

Tower of Pisa

Shape

Face

Location

Human

Real World

(Projected) Image

Computer Vision

Information

Image Processing

(transformed) Image/Signal

4

What is Computer Vision?

Image Understanding

Computer Vision

Real World

(Synthesized) **Image**

Computer Graphics

inverse relationship

(Projected) **Image**

Computer Vision

Tower of Pisa

Shape

Face

Location

Human

Information

Image Processing

(transformed) **Image/Signal**
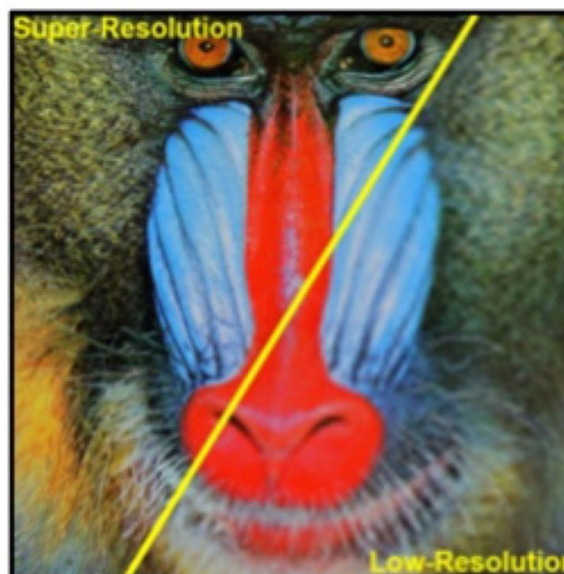
Examples: Image Editing and Generation

- **MMEditing** (2022)
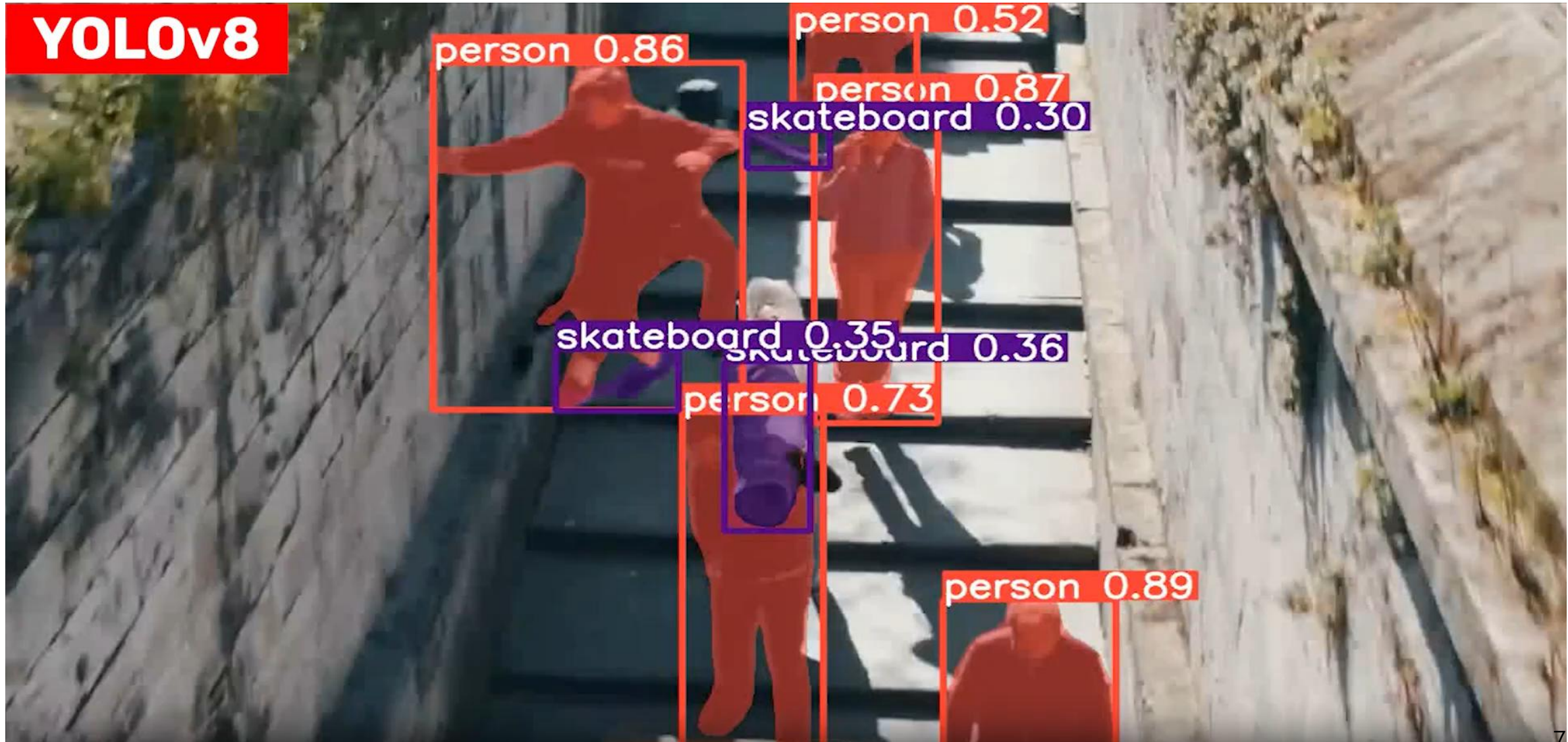

Inpainting
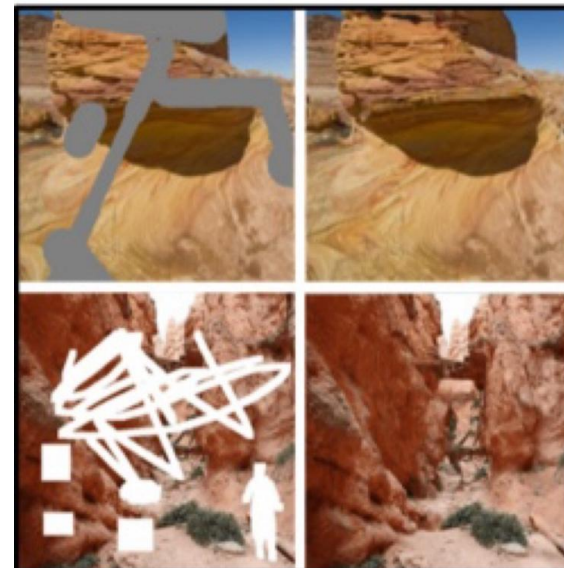
Matting

Super Resolution

Frame Interpolation

$t = 0s$

$t = 0.5s$ New Frame

$t = 1s$

Generation

zebra → horse

Generated Animal Faces
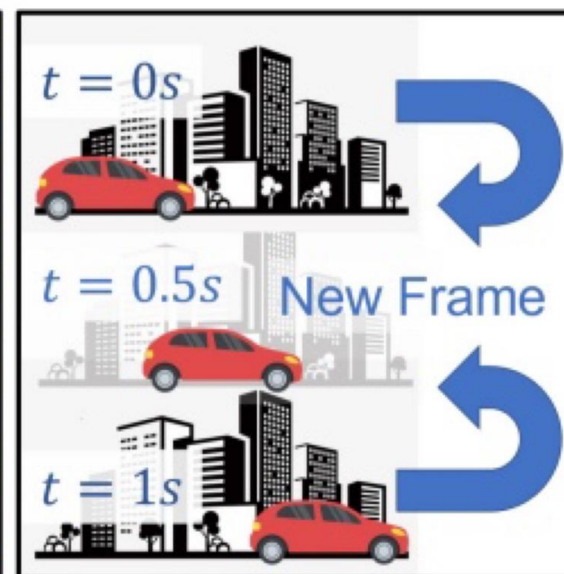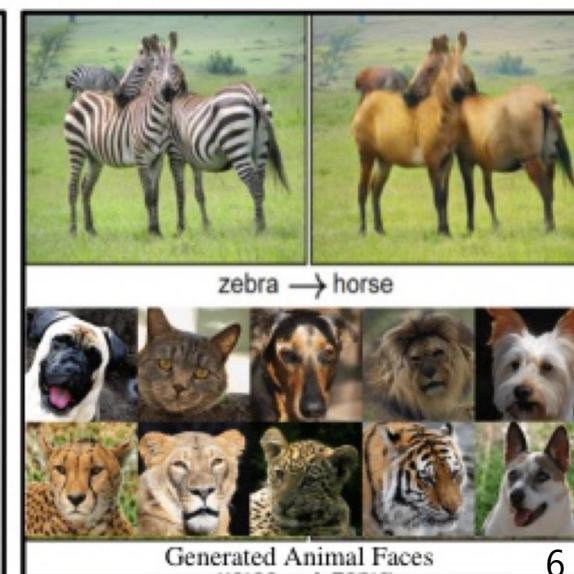
Examples: Object Detection (+ Instance Segmentation)

- **[YOLOv8](#)** (2023)

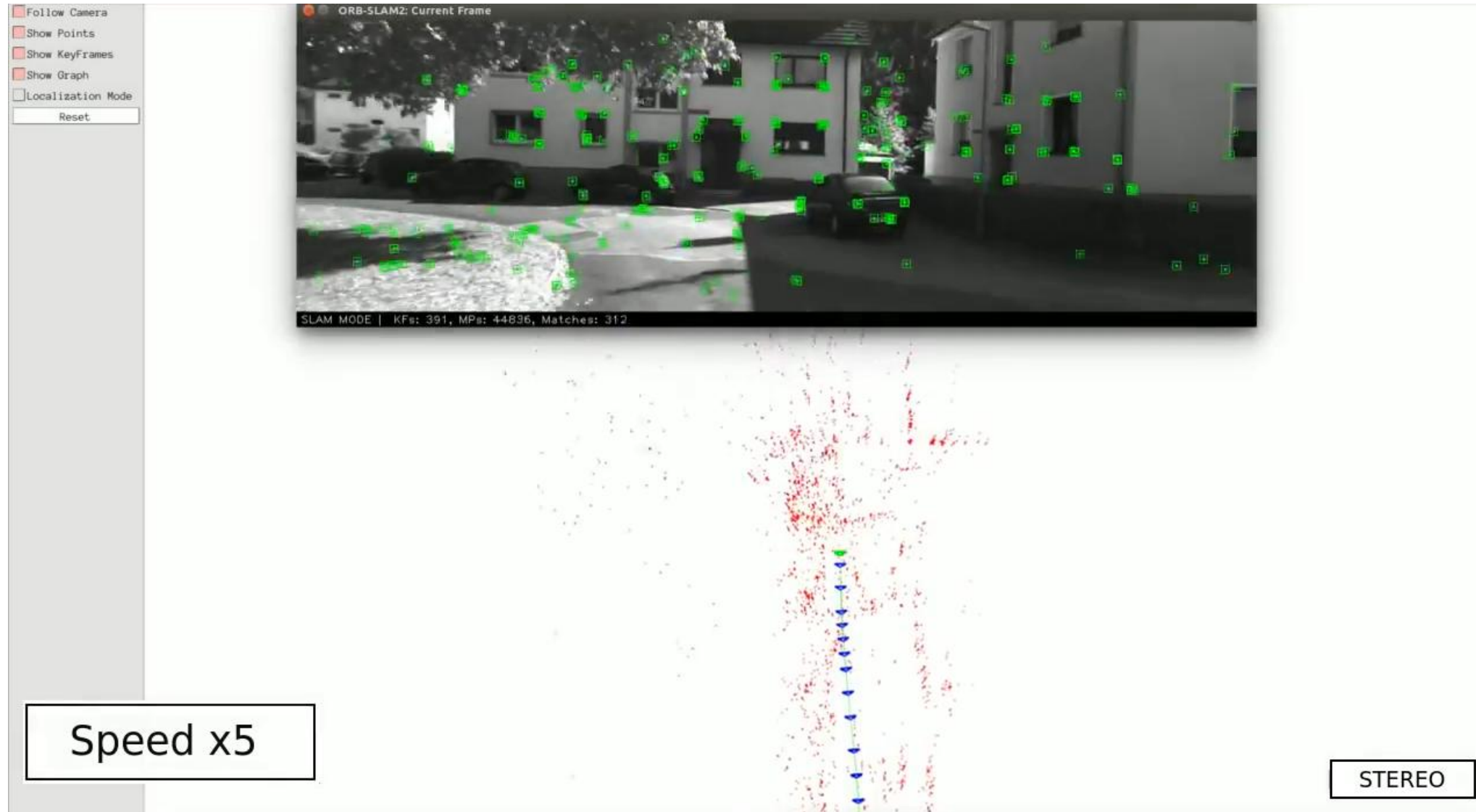# Examples: Multi-Object Tracking (MOT)

- **ByteTrack** (2022)

Examples: Human Pose Estimation

- **[MediaPipe](link) Holostic** (2019)

Examples: Localization and Mapping

- **ORB-SLAM3** (2021)

Examples: 3D Reconstruction

- **COLMAP** (2016)

A sparse model of *central Rome* using 21K photos produced by COLMAP's SfM pipeline

Dense models of several landmarks produced by COLMAP's MVS pipeline
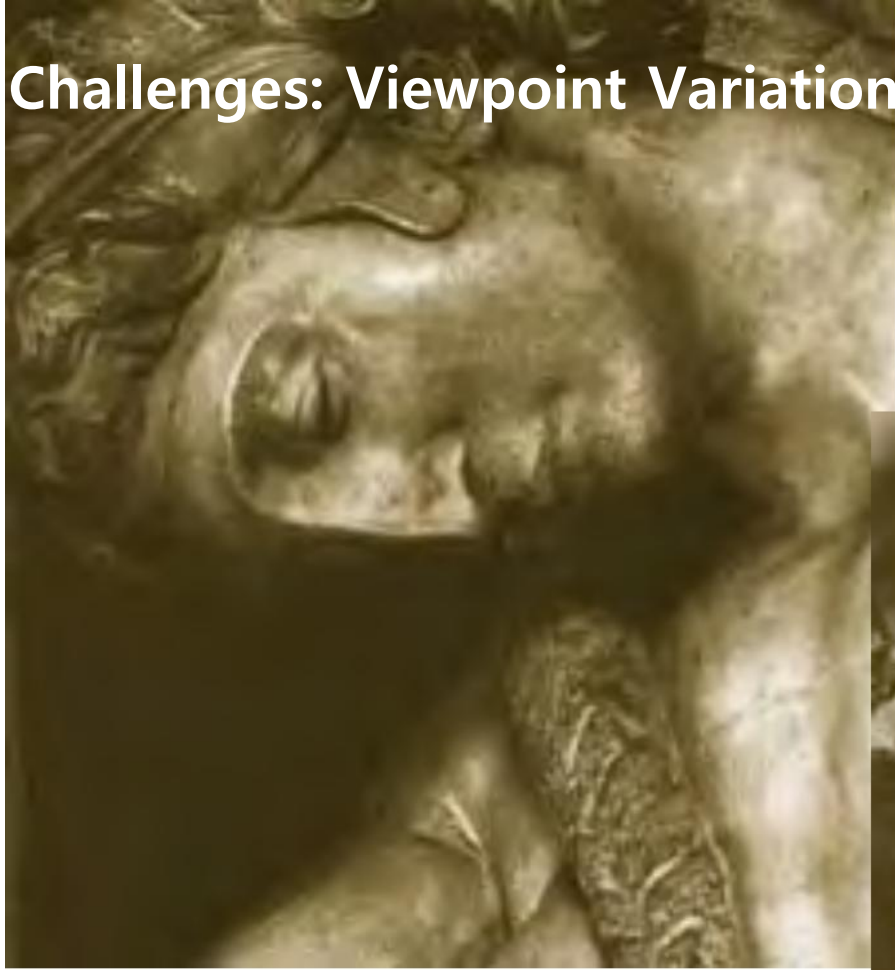
Examples: New View Synthesis

- **NeRF** (2020)

Why Do We Study Computer Vision?

- Vision is *useful*.

- Vision is *interesting*.

- Vision is *difficult*.

    – The half of primate cerebral cortex is devoted to visual processing.

    – Achieving human-level visual perception is probably "AI-complete".
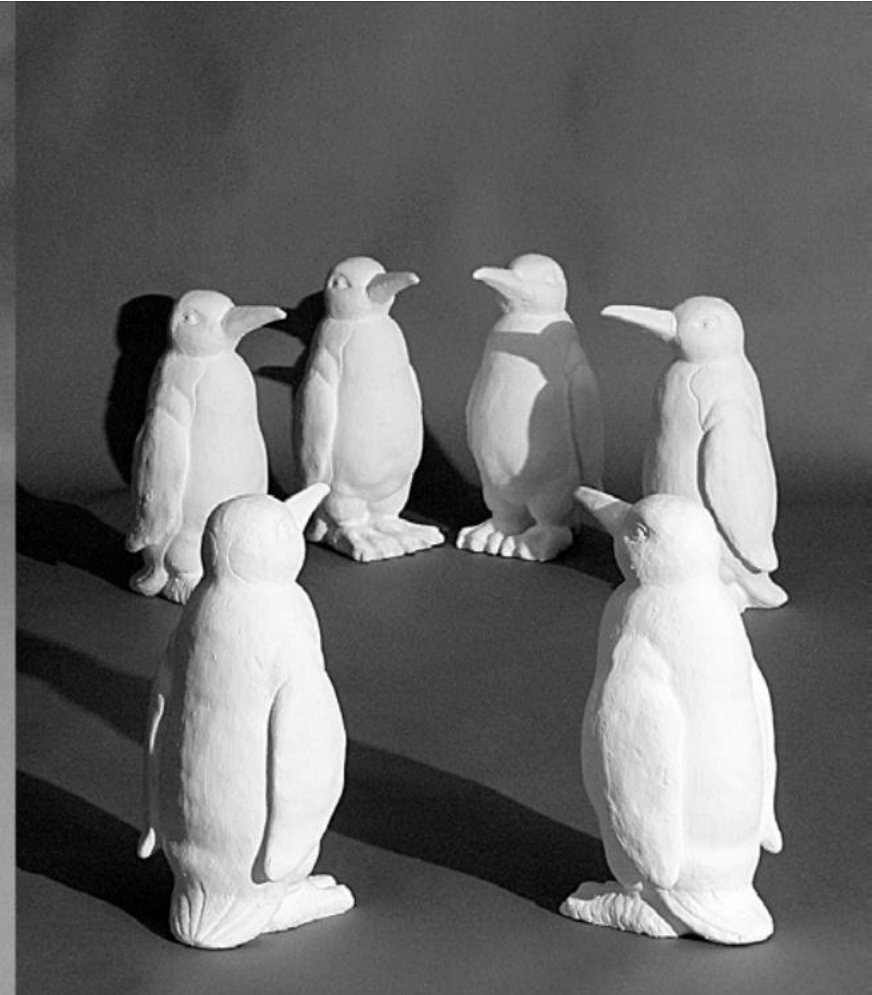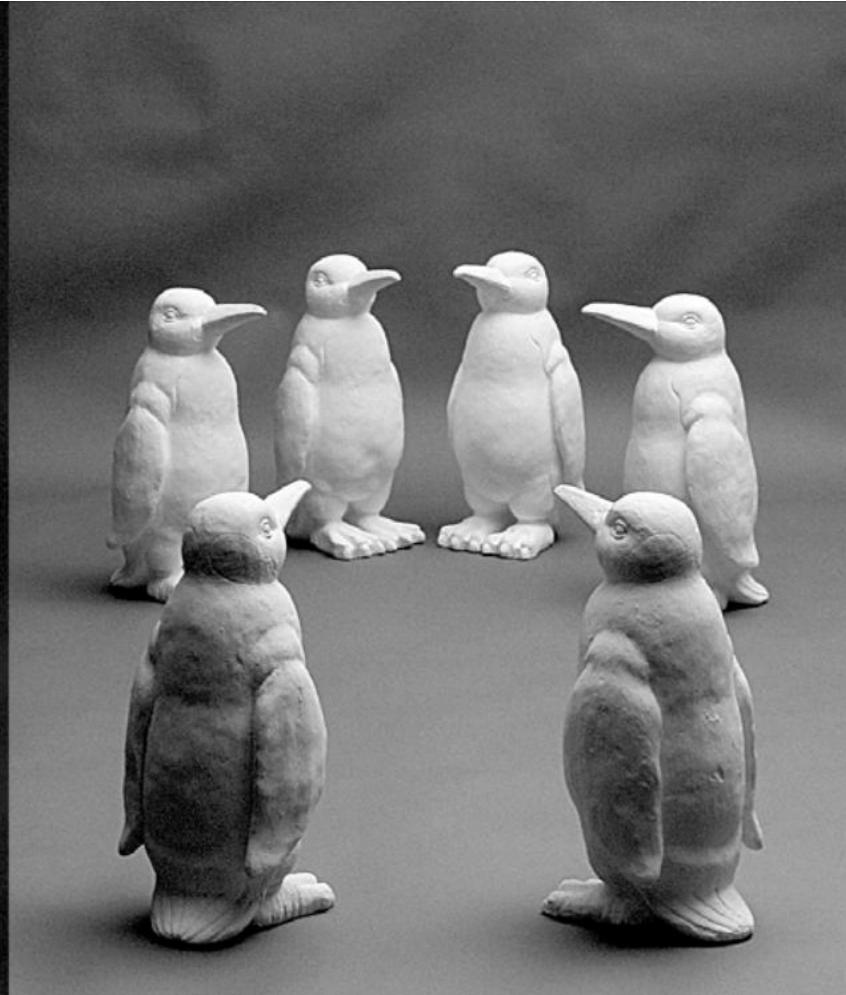
Challenges: 2D Projection

Image: Floor trick arts

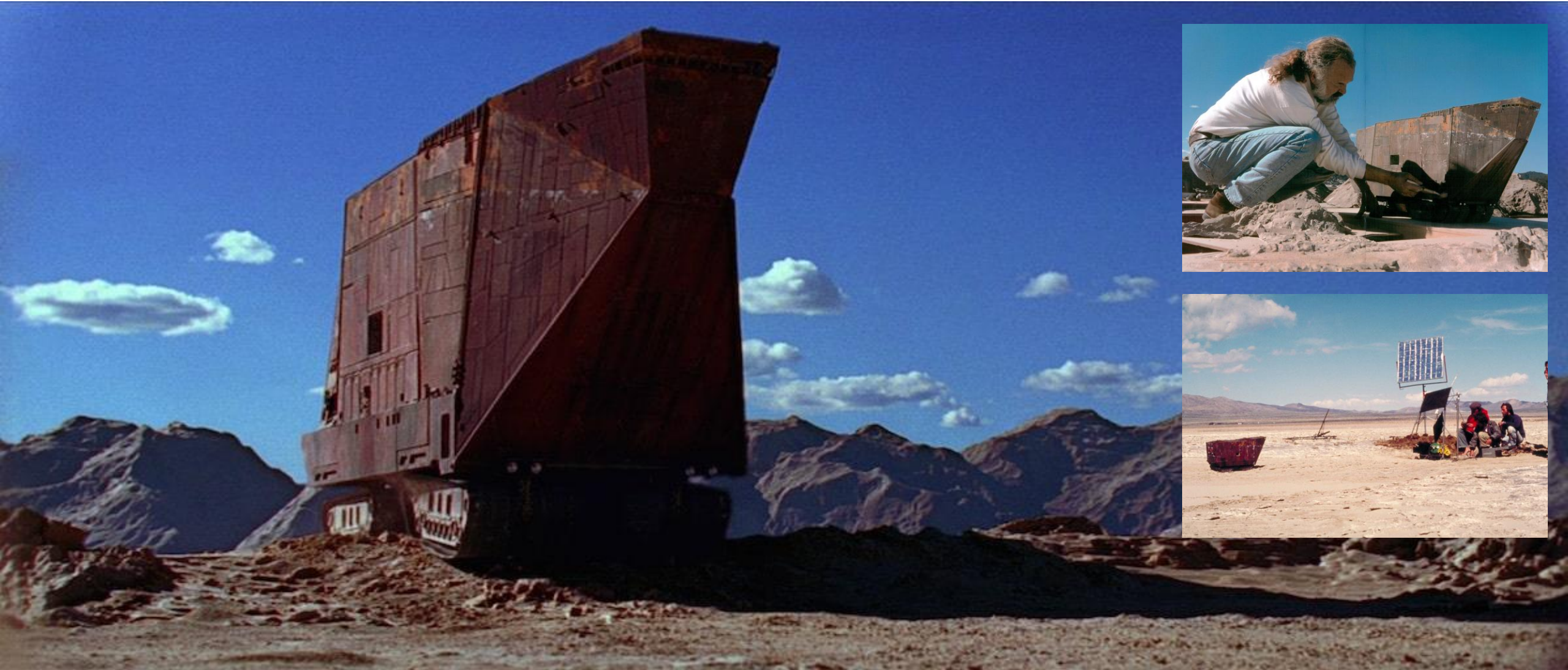# Challenges: Viewpoint Variation

Michelangelo (1475-1564)

Challenges: Illumination Change

J. Koenderink

Challenges: Scale Ambiguity

The Sandcrawler @ Star Wars IV: A New Hope (1977)

Challenges: Deformation

Xu, Beihong (1943)

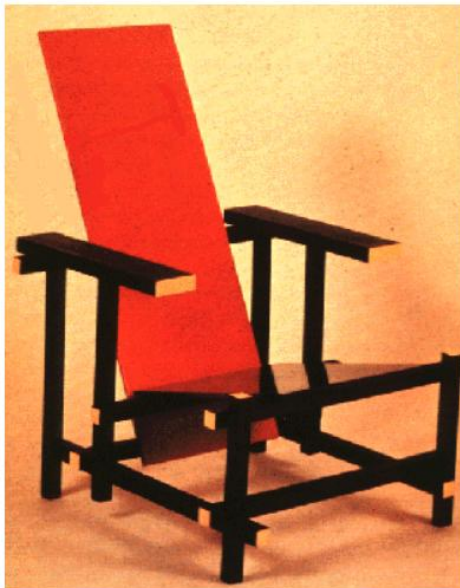# Challenges: Occlusion

Magritte (1957)

19

Challenges: Background Clutter
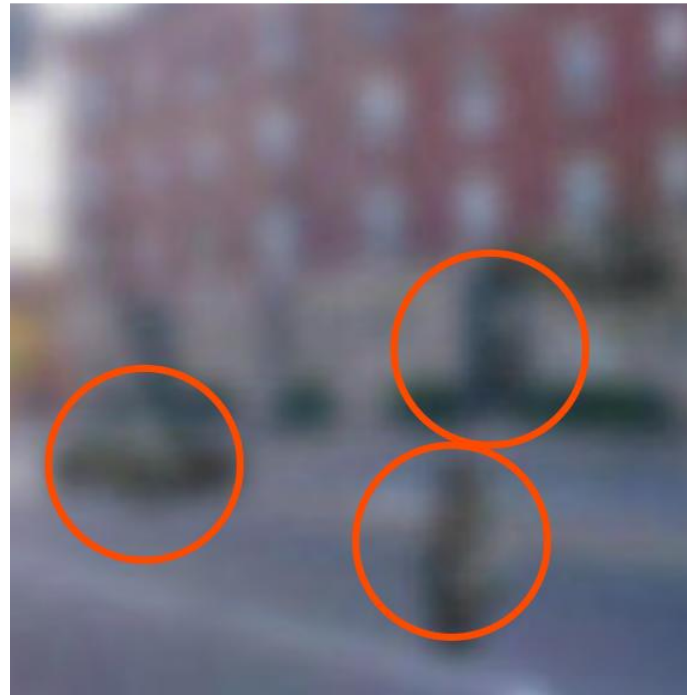
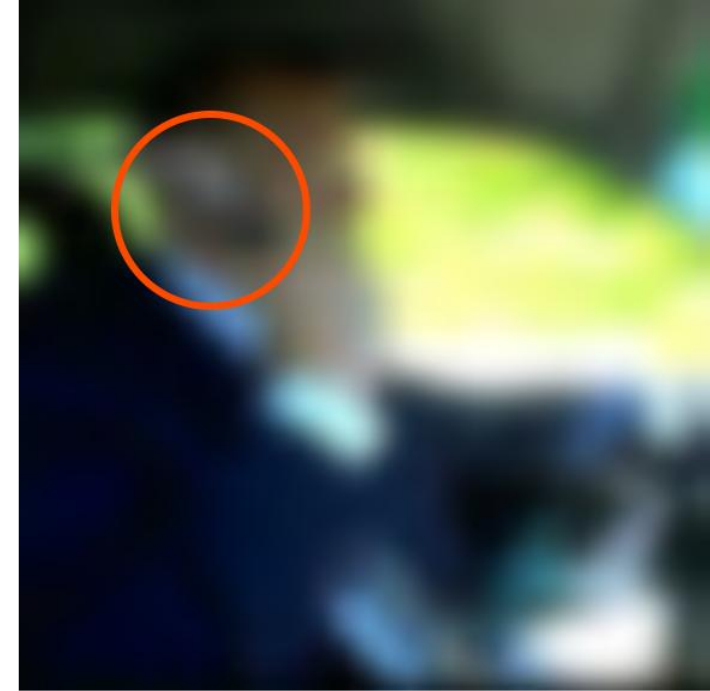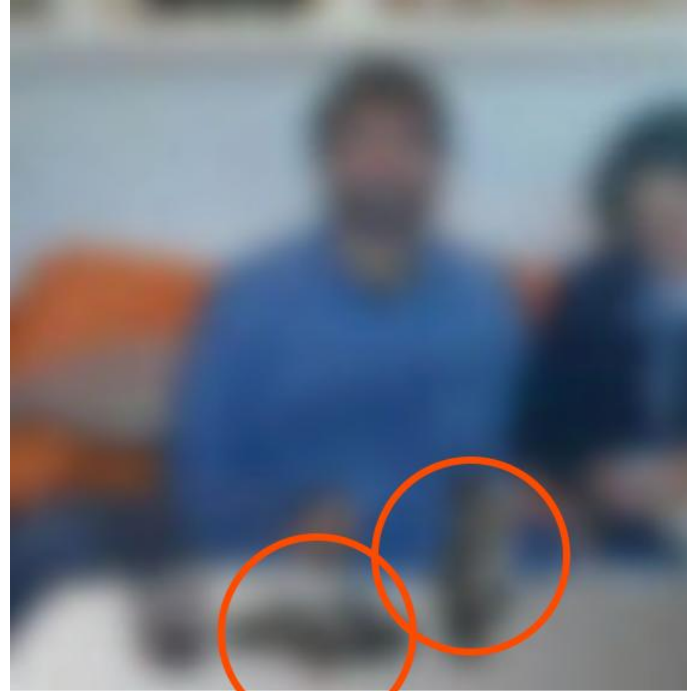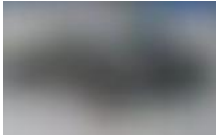Tim Laman (National Geographic)

Challenges: Motion

Challenges: Object Intra-class Variation

22

Challenges: Local Ambiguity

23

Summary

- **What is computer vision?**
  - Automatic visual understanding

- **Examples**
  - Object detection, …, 3D reconstruction, …

- **Why do we study computer vision?**
  - Vision is *useful*.
  - Vision is *interesting*.
  - Vision is *difficult*.

- **Challenges**
  - 2D projection, …, object intra-class variation, …