



Robot Vision

Introduction

Dr. Chen Feng

cfeng@nyu.edu

ROB-UY 3203, Spring 2024

Chen Feng

- *Appointments*

2018–now, Assistant Professor, **NYU**

Dept. Civil & Urban Engineering

Dept. Mechanical & Aerospace Engineering

Dept. Computer Science & Engineering

Center for Urban Science & Progress

2015–2018, Research Scientist, **Mitsubishi Electric Research Labs (MERL)**

Computer Vision Group

- *Education*

University of Michigan, Ann Arbor

2015 Ph.D. in Civil Engineering

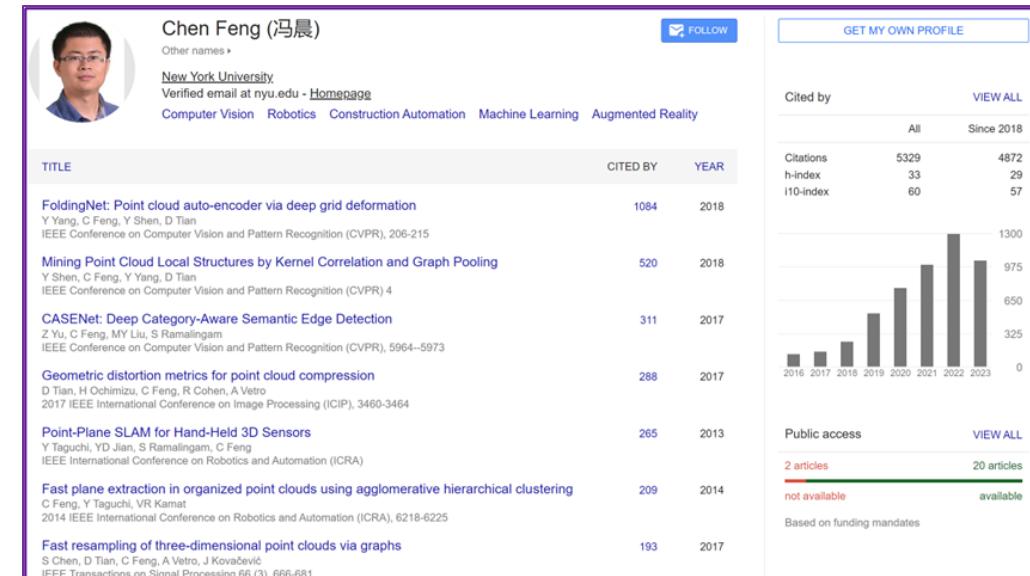
2013 M.S.E. in Electrical Engineering

Wuhan University

2010 B.E. in Photogrammetry

- *Publications*

- CVPR/ICCV/ICRA/IROS/NeurIPS/RA-L
 - #citation 5300+, h-index=33; 10+ patents



- Awards

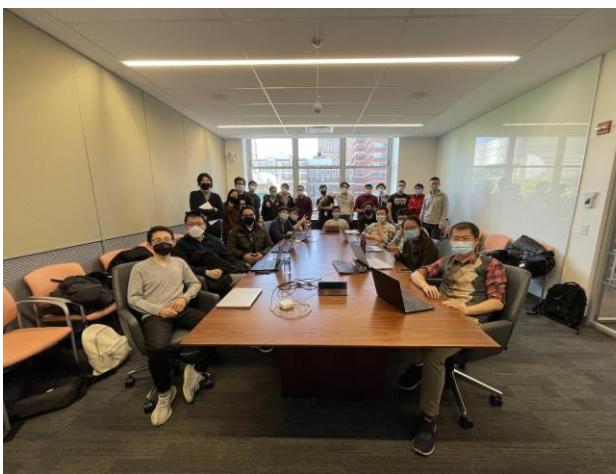
2023 NSF CAREER Award

2022 NYU Junior Faculty Research Award

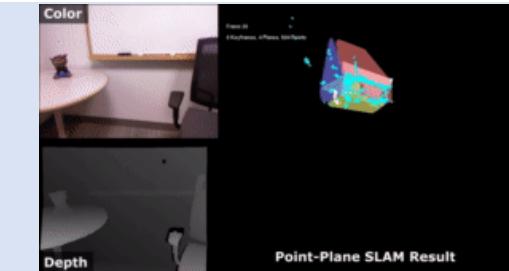
2019 NSF CPS, NRI, Future of Work

AI4CE Lab

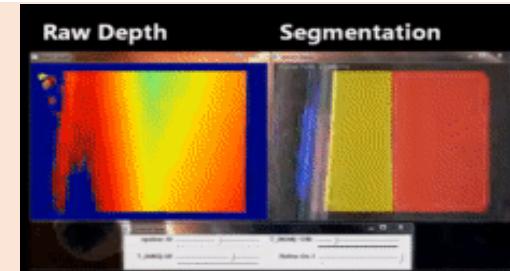
- Challenge:
 - To develop novel algorithms for **intelligent agents** to accurately and efficiently **interact** with **materials** and **humans** in **dynamic** and **unstructured** environments.
- Methodology: **multidisciplinary use-inspired research**
- Expertise
 - In Robotics & AI: **computer vision, robot perception, and machine learning**
 - In Engineering: **construction robotics, smart cities, intelligent transportation**



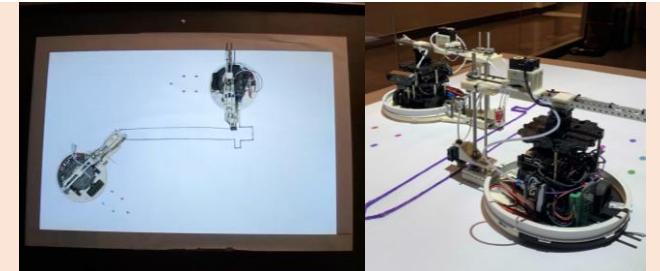
AI4CE Lab's Research in Spatial AI (Vision, Learning, & Robotics)



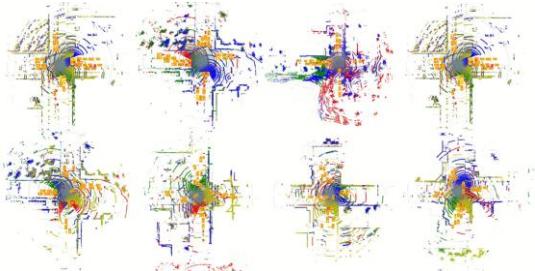
PP-SLAM, ICRA'13



PEAC, ICRA'14



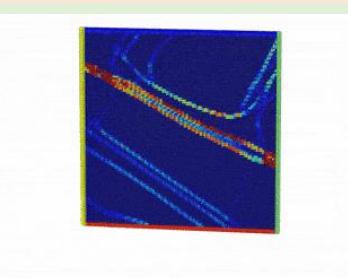
M3DP, ICRA'21



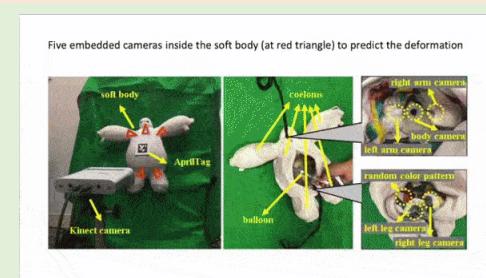
CoPerception, NeurIPS'21



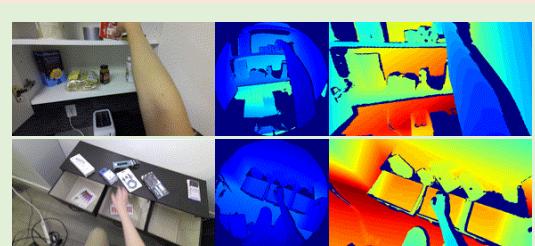
DeepExplorer, RSS'23



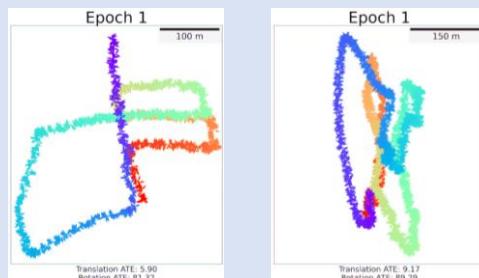
FoldingNet, CVPR'18



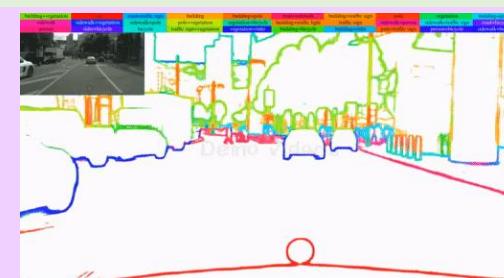
DeepSoRo, RAL'20



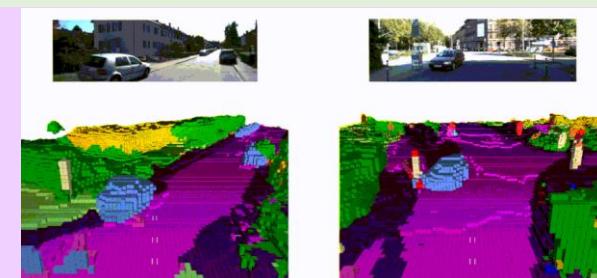
EgoPAT3D, CVPR'22



DeepMapping2, CVPR'23



CASENet, CVPR'17



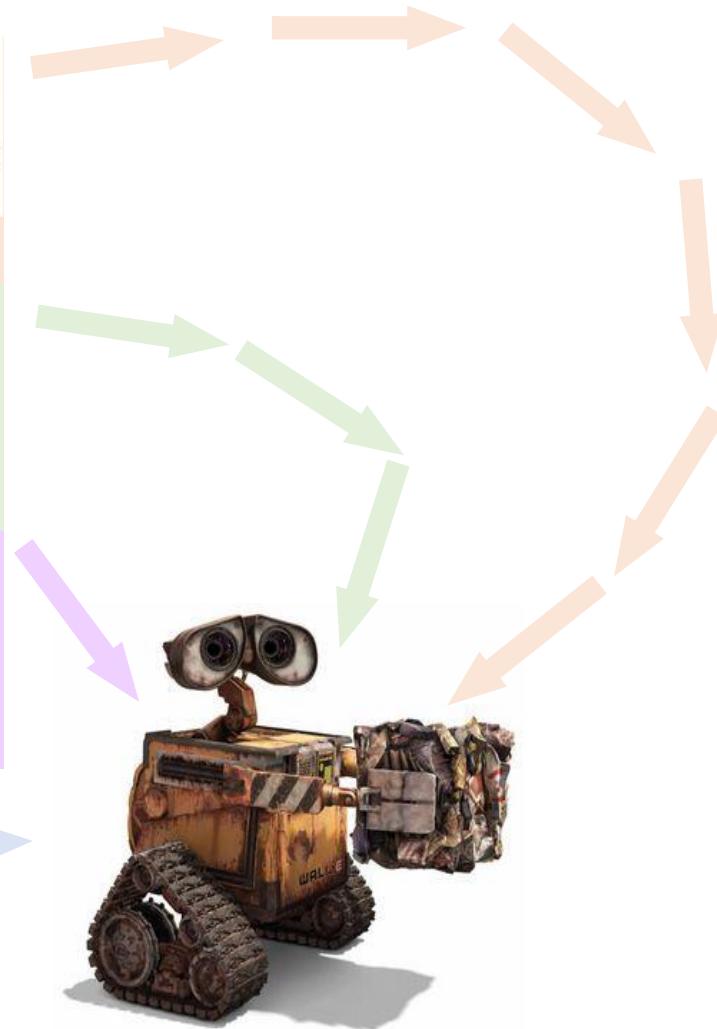
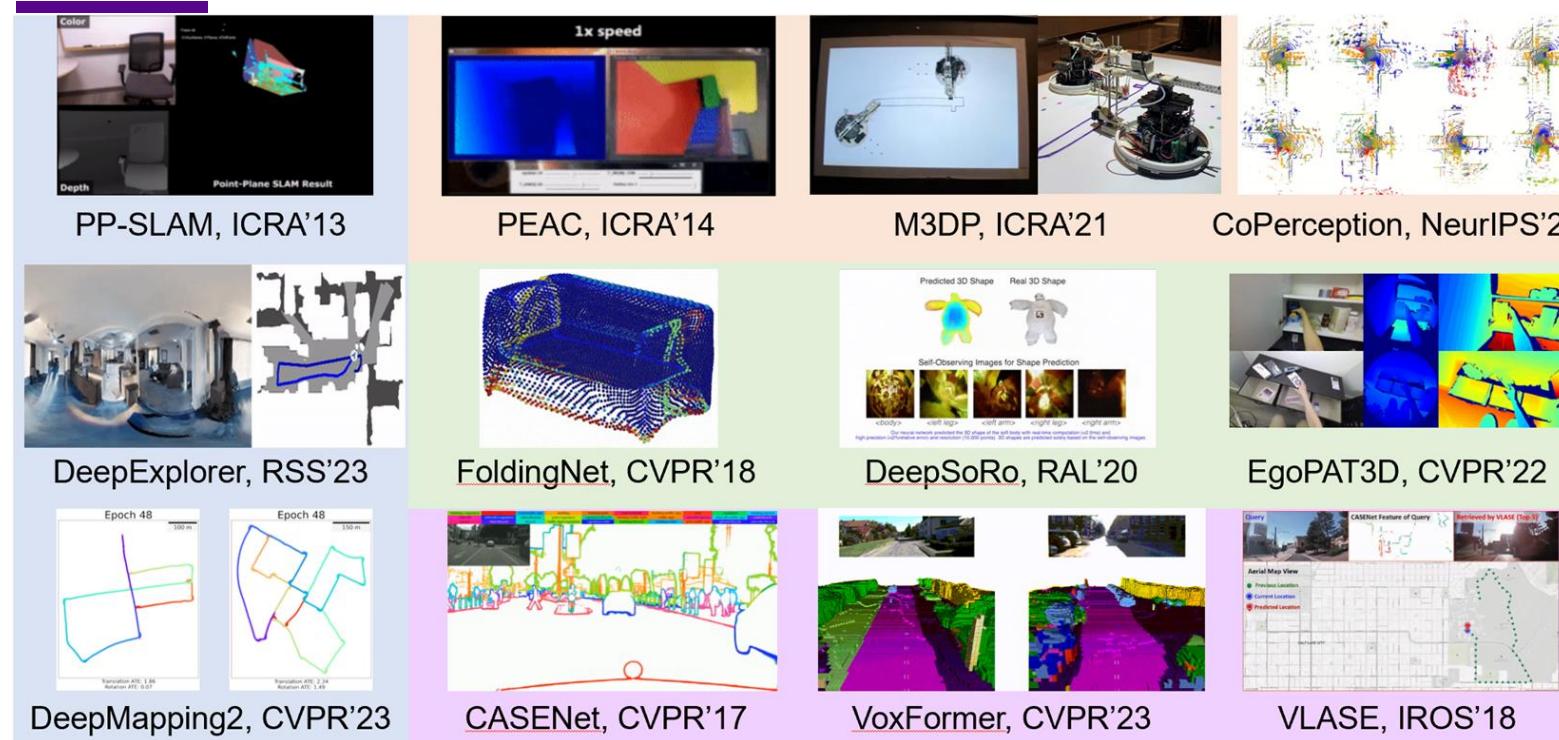
VoxFormer, CVPR'23



VLASE, IROS'18



Next 20–30 Years?



Autonomous Robots with Spatial AI



Course Information

- Class
 - Lecture (with [Zoom](#)) + [Slack](#) (Q&A)
 - Monday & Wednesday, 15:30 – 16:40
 - Regular Lectures + Hands-on/HW Lectures
 - Office Hours:
 - Slack channel #qna-spring24 (or direct message) is the best place for asking questions
 - 1on1 meetings could be arranged with the TA/me after Slack chat.
- Assistant Instructors
 - [Keifer Lee \(MS in ECE\)](#)
 - [Ashay Ajbani \(MS in Robotics\)](#)
- Current Demographics:
 - 2 sophomore, 26 junior, 22 senior students
 - 6 CE, 14 CS, 5 ECE, 12 EE, 8 ME, 5 Math
 - 3 AD, 47 TANDON



Prerequisites

- Must-have
 - Have interest in new technologies and how it works
 - Enjoy both learning the theories and acquiring hands-on experiences
- Should-have
 - Linear Algebra/Probability/Vector Calculus
 - Be comfortable to learn new software (GUI-based or command-line-based)
 - Know the basics of Python, and is comfortable coding with it
- Nice-to-have (**but not required, depends on your background**)
 - Image processing
 - Machine learning
 - Project: Python/MATLAB/C/C++/Java/...



Course Objectives

- To introduce the basic knowledge in robot vision
 - Main problems of interest
 - Fundamental theories
 - Classical solutions
- To enable the analysis and design of engineering projects by selecting and using robot vision properly
- To provide some hands-on robot vision experiences
 - Minimum: use existing software tools
 - Advanced: write your own robot vision codes
- To prepare the next generation of engineers/managers with the mindset, knowledge and skills of learning/using robot vision technologies



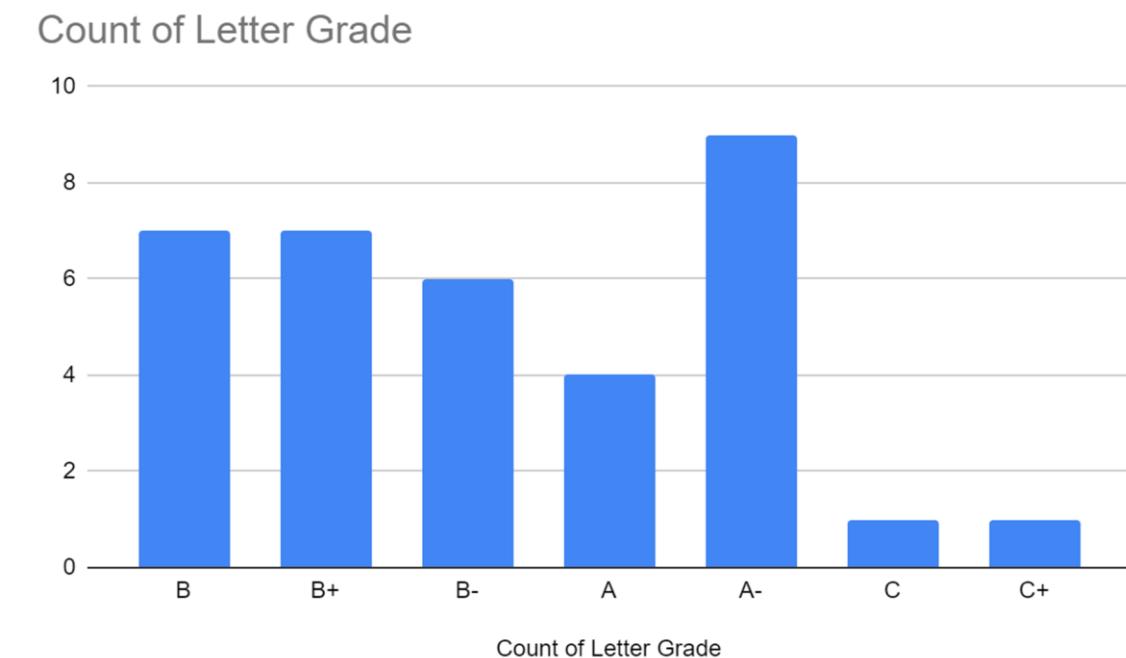
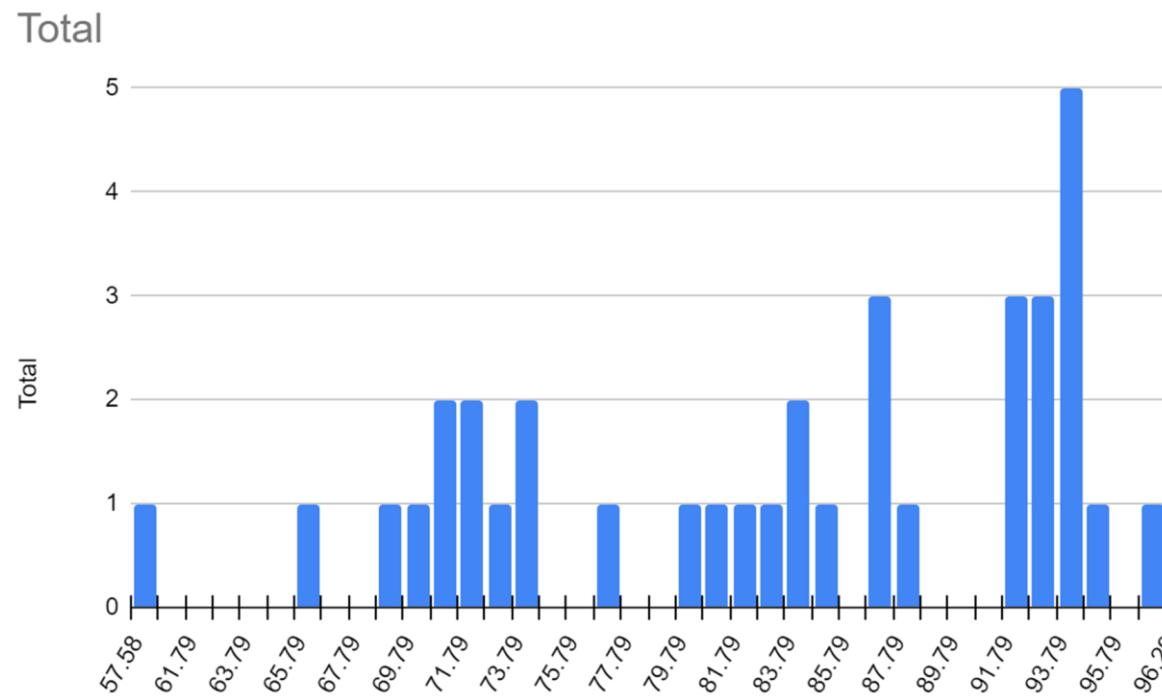
Assessment – Homework, Quiz, Project

- 40%: 10 in-class hands-on homeworks (4% each)
 - Due in 48 hours after the hands-on lecture (Either Pass or Fail)
 - Usually, these are simple variations of hands-on lecture tasks (theory + hands-on)
 - NYU Honor Code: discussion is allowed, but code/report must be written by yourself!
 - Potential “Copy & Paste” cases will be reported to the corresponding departments
 - All students involved in such cases will lose all of their points for that week
- 30%: 2 in-class online exams (based on Google Form)
 - Standardized tests, no coding needed, close-book, 1-page cheat sheet allowed
 - Any student in different time zones than NYC?
 - NYU Honor Code: **Discussion is NOT allowed!** Everything must be done by yourself ONLY!
- 30%: course challenge (two competitions)
 - Midterm competition (10%)
 - Final competition (20%)
 - video presentation on a virtual symposium via slack



Grading

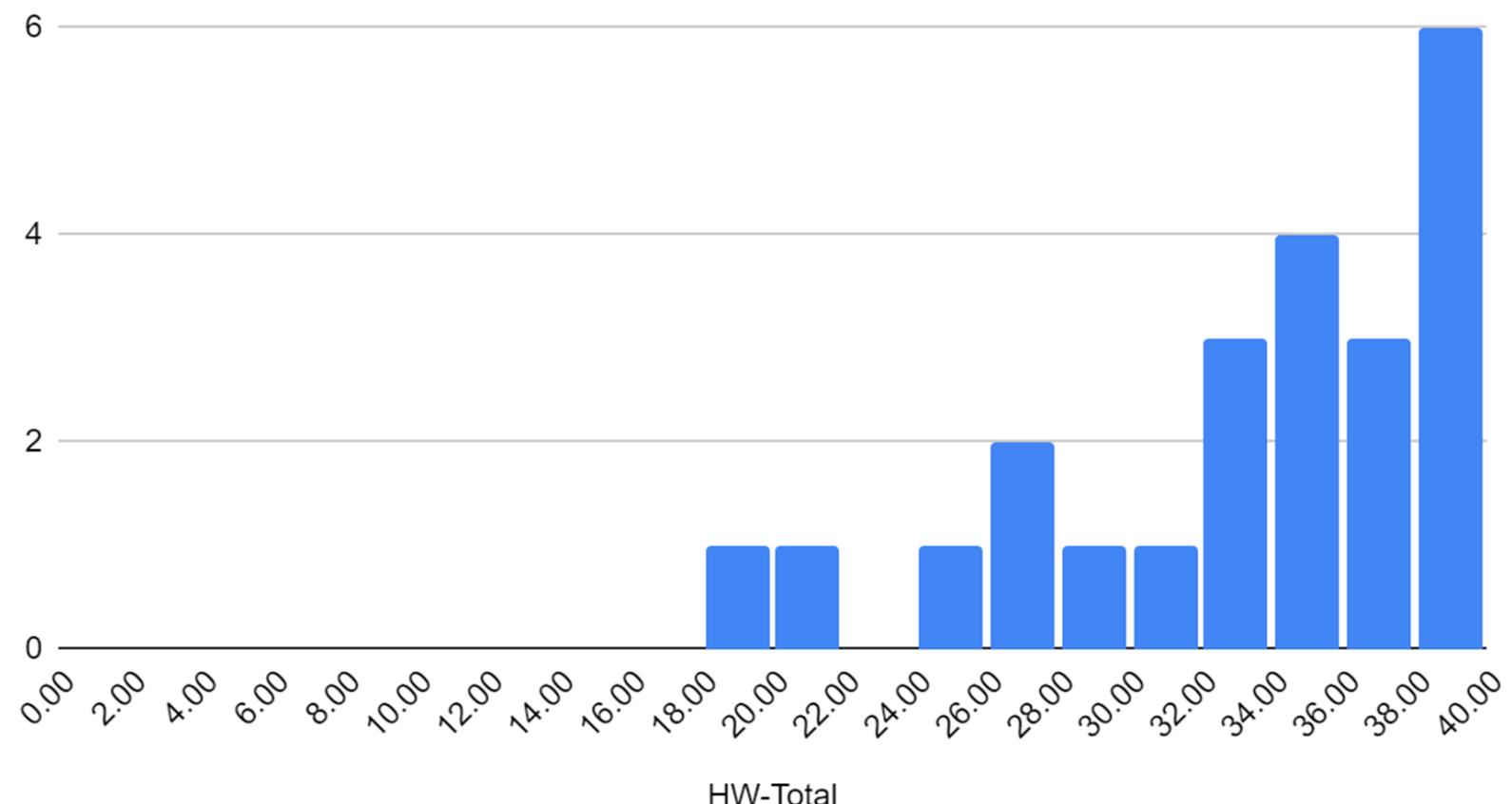
- Overall Course Score (HW*40%+Quiz*30%+Proj*30%) Distribution from Spring'22



Grading

- HW Grade (max=40) Distribution from Spring'22

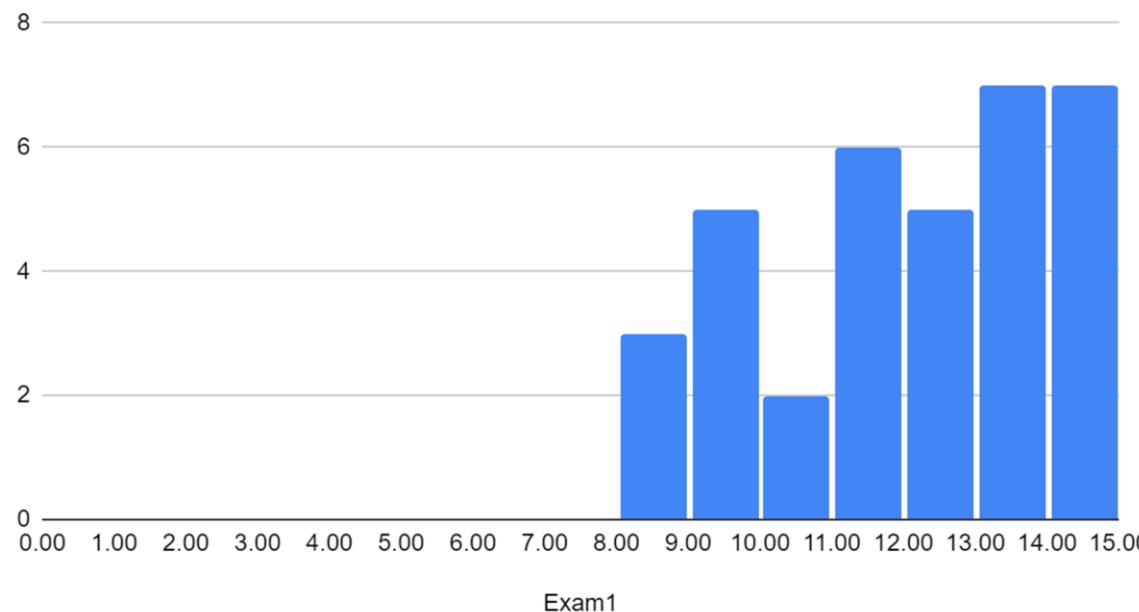
Histogram of HW-Total



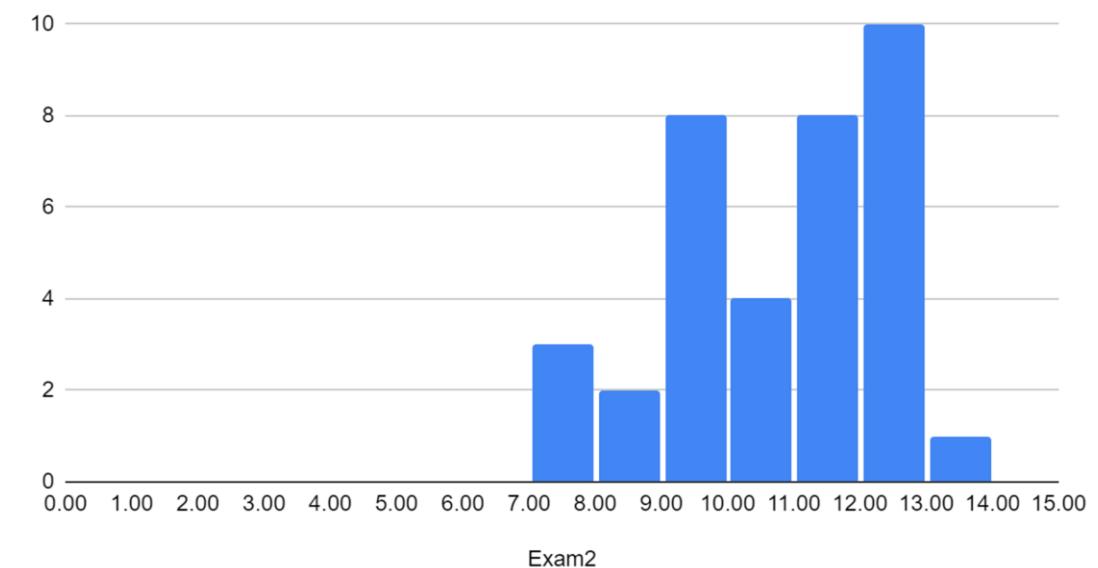
Grading

- Exams Grade Distribution (15% each) from Spring'22

Histogram of Exam1



Histogram of Exam2

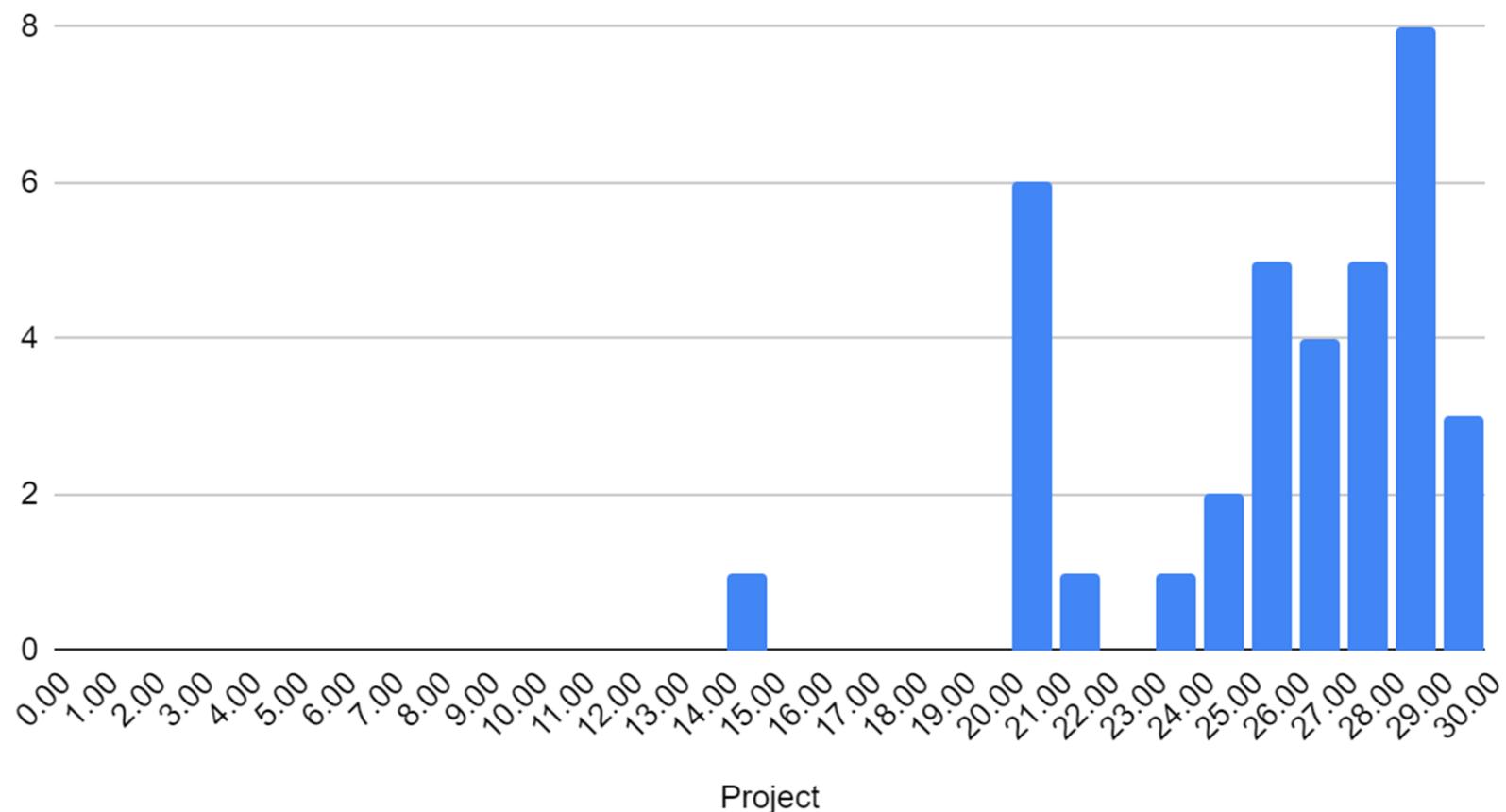




Grading

- Course Project Grade (max=30) Distribution from Spring'22

Histogram of Project





Text Books

- Main:

- [Ha2002] Hartley, R. and Zisserman, A. Multiple View Geometry in Computer Vision, Academic Press, 2002. [available online from NYU Library]
- [Fo2011] Forsyth, David A., and Jean Ponce. Computer Vision: A Modern Approach (2nd Edition). Pearson. 2011. [available online]
- [Co2017] Corke, P. Robotics, Vision and Control: Fundamental Algorithms in MATLAB. Springer, 2017. [available online from NYU Library]

- Supporting:

- [Sz2022] Szeliski, Richard, Computer Vision: Algorithms and Applications Springer, 2nd Edition, 2022. [\[available online\]](#)
- Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow, Second Edition, Sebastian Raschka, and Vahid Mirjalili, Packt Publishing, 2017. [\[available online from NYU Library\]](#)
- [Go2016] Goodfellow, Ian, et al. Deep learning. Cambridge: MIT Press, 2016. [\[available online\]](#)

Other Resources

- Wikipedia
 - Good way to learn new concepts
 - But do not trust it wholeheartedly
- CVonline: <https://sites.google.com/site/cvonlinewiki>
 - A dedicated list of wiki links to computer vision
 - And a list of vision-related online books
 - Some maybe outdated
- Google/Bing/Google-Scholar should be your best friend!
 - You need to learn how to search effectively and efficiently

Related Conferences

- CVPR: Computer Vision and Pattern Recognition
- ICCV/ECCV
- ICRA: IEEE International Conference on Robotics and Automation
- IROS
- 3DV
- ISARC: International Symposium on Automation and Robotics in Construction



Related Conferences

- Computer Vision
 - CVPR: Computer Vision and Pattern Recognition
 - ICCV: International Conference on Computer Vision
 - ECCV: European Conference on Computer Vision
- Robotics
 - ICRA: IEEE International Conference on Robotics and Automation
 - IROS: International Conference on Intelligent Robots and Systems
 - 3DV: International Conference on 3D Vision
 - ISARC: International Symposium on Automation and Robotics in Construction
- Machine Learning
 - NeurIPS: The Conference and Workshop on Neural Information Processing Systems
 - ICML: International Conference on Machine Learning
 - ICLR: International Conference on Learning Representations

Open-Source Libraries & Tools

- Ready-to-use software:
 - [MeshLab/CloudCompare](#)
 - [COLMAP](#)
 - [OpenVSLAM](#)
- Software development libraries:
 - OpenCV
 - PCL: Point Cloud Library
 - Open3D
 - NumPy/Eigen
 - Google Ceres, G2o
 - Scikit-learn
 - PyTorch, TensorFlow/Keras



Definition of Robot Vision in This Course

- Robot
 - A machine capable of carrying out a complex series of actions automatically
 - Regardless of how it looks
 - With a bit of a stretch in scale, an automated system also counts
- Vision: Visual Perception
 - The *awareness* of the elements of environment through visual sensing
 - The process of becoming *aware* of something using visual sensors

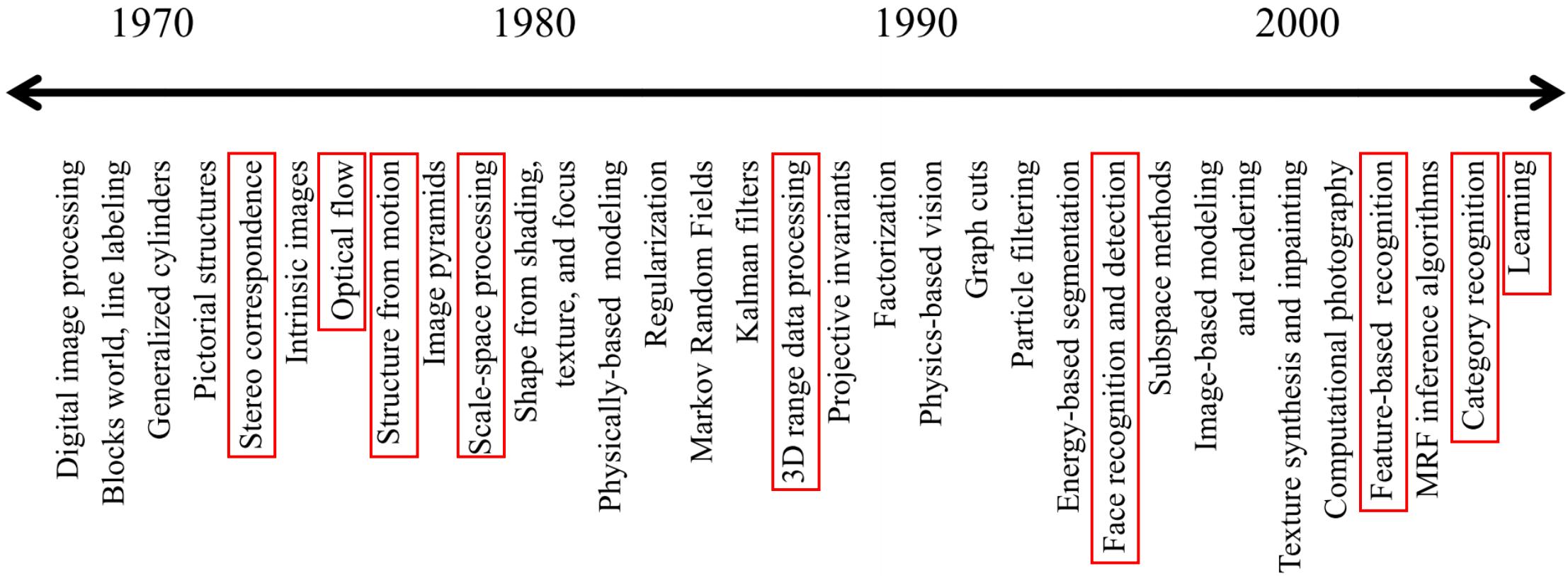


Definition of Robot Vision in This Course

- Robot Vision
 - The sensors and algorithms
 - to sense and process raw visual signals
 - to become aware of the states of a system itself and/or its surroundings
 - which enables the system to take proper actions to fulfil its design goals
 - Visual
 - Line-of-sight Sensors: cameras, LIDAR
 - Algorithms: **computer vision, machine learning**
 - The focus of this course
 - Tactile
-
- Shortened version
 - **Observing visual data to understand situations** for achieving system objectives

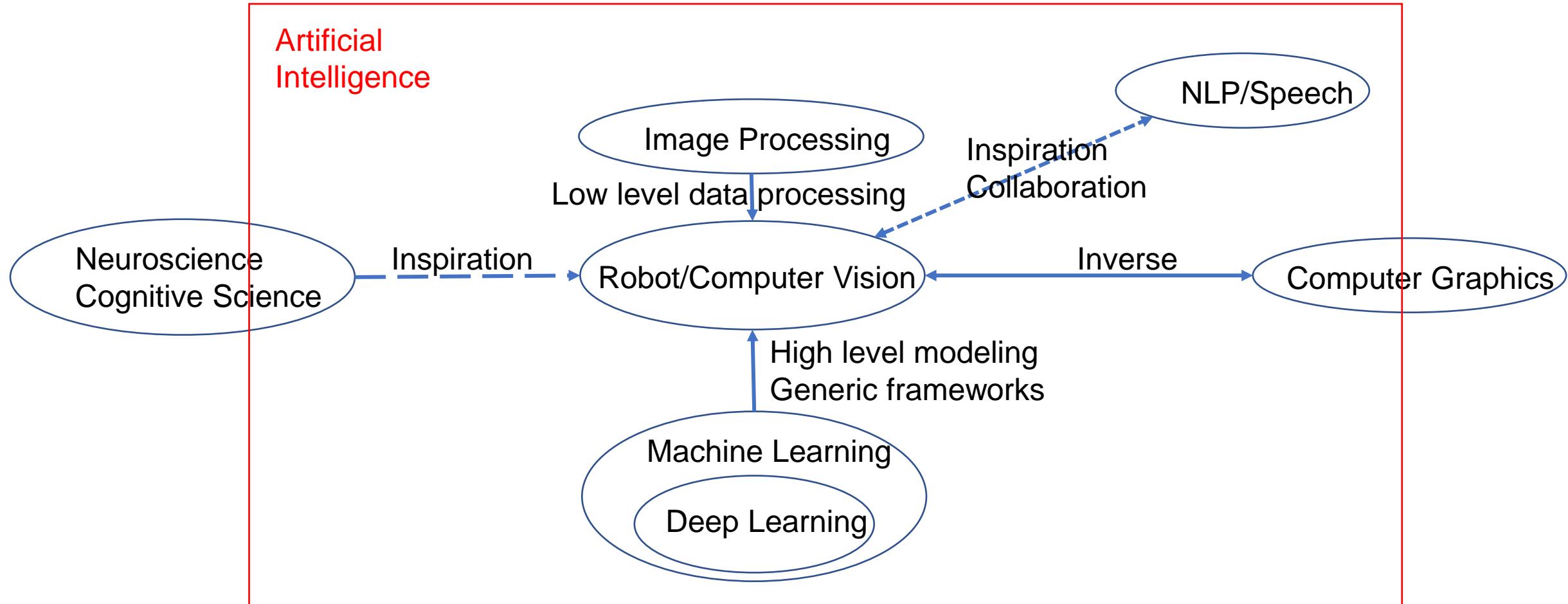


A Brief History of Computer Vision





Related Disciplines



Main Topics in Robot Vision

- Geometric:
 - Shape/Structure
 - Pose (position + orientation; posture for articulated objects)
- Semantics:
 - Segmentation
 - Detection
 - Tracking
 - Recognition

LinkedIn job posts



Robotics Engineer - New Grad

NASA Jet Propulsion Laboratory · Pasadena, CA, US

Posted 1 day ago · 82 views



2 company alumni work here

Robotics engineers use multi-disciplinary skills to develop robotic systems for space or terrestrial applications. These systems are remotely-controlled or autonomous machines that sense, understand, and take useful actions in unknown and unstructured environments. Typical sub-disciplines include computer vision and perception, multi-sensor estimation and control,

FANUC

Senior Engineer Machine Vision Products

FANUC America Corporation · 3900 W Hamlin Rd, Rochester Hills, MI 48309, US

Posted 1 week ago · 37 views



Robotics Perception Validation Engineer

Volt Workforce Solutions · Pittsburgh, Pennsylvania

Posted 2 weeks ago · 68 views



Be an early applicant



Robotics Research Hardware Engineer (NCG)

NVIDIA · Redmond, WA, US

Posted 5 days ago · 64 views

Save

Apply

experiments. Experience with core robotics areas such as manipulation, vision, control, etc. is a significant plus. Candidates should have deep and gauges, tactile sensing, cameras (depth, binocular, high-speed, high-resolution, calibration, communication, etc.), real-time



Systems QA Robotics Automation Engineer

Apple · Cupertino, CA, US

Posted 2 days ago · 30 views

- Developing and executing testing using computer vision.



Advanced Robotics R&D Engineer Intern Summer 2019

Amazon Robotics · North Reading, MA, US

Posted 1 week ago · 578 views

The Intern Will Be Responsible For

- Identifying creative solutions for challenging research problems in robotics and computer vision



Robot Vision: Introduction (cfeng@nyu.edu)

BIG CLOUD

Robotics Engineer - SLAM
Big Cloud · Sunnyvale, California
Posted 6 days ago · 28 views

[Save](#) [in Easy Apply](#)

Chris Bradbury can refer you
Get referred to [increase your chances](#) of landing an interview.

[Ask for a referral](#)

Job description
An exciting autonomous driving start-up is expanding its team. They need world-class machine learning researchers to help them solve complex problems in machine learning, computer vision, data engineering and robotics. They just closed an impressive series A and have an incredible technical leadership team.

They're looking for candidates with real-world experience of robot

How you match
Criteria provided by job poster

Skills

- Computer Vision
- Machine Learning
- Robotics
- SLAM

Senior Robotics Engineer
Quest Groups LLC · San Francisco Bay Area
New · Posted 20 hours ago · 75 views

[Save](#) [in Easy Apply](#)

Job description
We are building the future of construction.
Our mission is to take the latest sensors from self-driving cars, retrofit them into proven equipment from the job site, and develop a suite of autonomous software designed specifically for the requirements of construction and earthmoving. And over the last two years, with a team of a dozen dedicated engineers, roboticists, and construction experts, that's what we've done. It hasn't been easy—in fact, no one has ever done what we're doing—but with

Job Description

- As a robotics engineer on our team you'll be designing, developing, testing and improving our autonomous track loader (ATL). You'll be a generalist, doing everything from optimizing controllers to designing high-level planners for new to construction tasks to improving the accuracy of our object recognition systems. You'll have a significant impact on the product road map, and you'll deploy code every day.

Qualifications

- BS, MS or PhD in computer science or related field
- 2+ years of software development experience
- Proficiency in Python
- Experience with distributed systems
- Familiarity with ROS is a plus
- Knowledge of computer **vision** is a plus
- Mechanical engineering experience is a plus, but not required
- Controls experience is a plus, but not required

Compensation & Benefits

- Competitive equity and salary (we believe great engineers are worth it)
- Free catered lunch and snacks in the office
- In-office gym

Robotics Engineer
The Mackenzie Group, LLC · Detroit, Michigan, United States
Posted 1 week ago · 135 views

[Save](#) [in Easy Apply](#)

Job description
The Mackenzie Group has been retained by a Detroit area automation company to fill a ROBOTICS ENGINEERING position. If you are interested in this exciting opportunity, please respond.

Skills Required/Details:

- Strong and recent background in two of the following robots: Fanuc, ABB, Kuka, or Comau.

How you match
Criteria provided by job poster

Skills

- Computer Vision
- Programming
- Robotics
- Electrical Troubleshooting

Robotics Software Engineer
Fox Robotics · Austin, Texas
Posted 2 days ago · 1,368 views

[Save](#) [in Easy Apply](#)

Job description
Develop navigation and perception capabilities on a mobile robot. You will need to rapidly integrate and evaluate third party libraries, keep abreast of current research, and develop core autonomy capabilities.

Requirements

- strong C++ programming skills
- experience with localization and mapping algorithms
- experience with 3D obstacle avoidance and navigation algorithms

Bonus points for

- proven ability to ship working robotic applications

How you match
Criteria provided by job poster

Skills

- Computer Vision
- C++
- Mobile Robotics
- Machine Learning
- Robotics
- Algorithms
- Localization
- Computer Hardware

Computer Vision Engineer - Robotics
Harnham · Greater Pittsburgh Area
Posted 2 weeks ago · 271 views

[Save](#) [in Easy Apply](#)

Sam Brown can refer you
Get referred to [increase your chances](#) of landing an interview.

Job description
Computer Vision Engineer - Robotics
Pittsburgh, PA
\$150,000 - 180,000 + Benefits + Equity

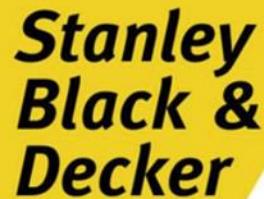
Harnham has been retained by a rapidly growing and well funded start up with over 40 employees globally and 10 million in initial funding. Focusing on localization and mapping as well as image processing, their mission is to develop robotic solutions, allowing robots to better perceive, understand and navigate in both indoor and outdoor environments.

Very much focused on C++ programming, you'll be an integral part of their

How you match
Criteria provided by job poster

Skills

- Research
- Computer Vision
- C++
- Robotics
- Algorithms
- SLAM
- Computer Hardware



Robotics Engineer Intern

Stanley Black & Decker, Inc. · Atlanta, GA, US

Posted 5 days ago · 165 views

Reporting to the Technical Manager of Robotic Systems, the Robotics Engineer engages in design and development of one or more areas related to advanced sensors and sensing technologies, algorithms development for perception, AI, robotic **vision**, control and data analysis, and integration of modern methods of navigation and path planning (specially with application to obstacle avoidance and autonomous operations.)

In this role, the ideal candidate is expected to have the following technical qualifications:

- BSc, MSc or PhD in Mechanical/Electrical/Robotic Engineering or Computer Science preferred, Bachelor's degree with appropriate additional years of experience also acceptable.
- Interest and/or experience developing robots and/or autonomous mobile systems. Hands on design/build experience of robot/UAV systems highly desired.
- Ability to do research in Robotics **vision** and or computational



Amazon Robotics Software Engineer - Autonomous Mobility

Amazon Robotics · North Reading, MA, US

Posted 1 week ago · 136 views

Job description

As an Amazon Robotics software engineer on the Autonomous Mobility team, you will own key components of navigation, **vision**, control, and coordination for single and multi-agent systems. You will collaborate cross-functionally to define clear requirements, deliverables, and test cases in an exciting, highly innovative environment.



Robotics Engineer (LiDAR / Computer Vision / C++)

MoTek Technologies · San Francisco Bay Area

Posted 3 weeks ago · 99 views

[Save](#)

[in Easy Apply](#)



Robotics Computer Vision Engineer

Bastian Solutions · Boise, ID, US

Posted 2 weeks ago · 193 views



Job description

Robotics Engineer

Location: Austin, TX (relocation assistance)

Salary: \$80,000 - \$120,000 + equity

Top Reasons to Work with Us

- Be part of the founding team.
- Influence the direction of the company and product.
- Experience with robotics
- Chance to live in amazing Austin, Texas!

What You Will Be Doing

- Writing our core software stack
- Have an interest in:
- Web Based Applications
- Navigation for dynamic door environments
- Machine Learning and data visualization
- Manipulation in semi-structured environments
- Computer vision on real systems

What You Need for this Position

- Experience with deploying algorithms on real robots
- ROS

Seniority Level
Entry level

Industry
Hospital & Health Care

Employment Type
Full-time

Job Functions
Information Technology



Software Engineer, Researcher / Robotics
MoTek Technologies · Palo Alto, California
Posted 3 weeks ago · 160 views

Save Easy Apply

Jerald Baker can refer you
Get referred to **increase your chances** of landing an interview.

Ask for a

Job description

Software Engineer, Researcher / Robotics - Level 5 Autonomous Driving

Our software allows our vehicles to perceive the world around them, make the right decision for every situation, and deliver people safely to their destinations. We think deeply and solve complex technical challenges in areas like robotics, perception, decision-making and deep learning, while collaborating with hardware and systems engineers. If you're a software engineer or researcher who's curious and passionate about Level 5 autonomous driving, we'd like to meet you.

How you match
Criteria provided by job poster

Skills

- Research
- Computer Vision
- C++
- Machine Learning
- Robotics
- Algorithms

Sensor Fusion
 Object Tracking
 Computer Hardware
 Automotive

Level of education

- Doctor of Philosophy

In this researcher/robotics role you'll:

- Work with world-class experts in the field of autonomous vehicles and advance the state of the art in areas such as computer vision, sensor fusion, machine learning, object tracking, and motion planning.

We'd like you to have:

- MS/PHD in Robotics, Computer Science, or equivalent practical experience.
- Experience in hands-on robotics research and expertise in one or more of the following: computer vision, LIDAR, object tracking, sensor fusion, perception, machine learning, motion planning, and control.
- Experience in data structures and advanced algorithms.
- Experience programming in C++ and with robust, safety-critical, efficient code.
- Experience with field robotics and systems design.

Jerald Baker 1st
Technical Recruiter

Message

Job Details

Software Engineer, Researcher/Robotics
Waymo · Ann Arbor, MI, US
Posted 1 week ago · 1,965 views

In This Researcher/Robotics Role You'll

- Work with world-class experts in the field of autonomous vehicles and advance the state of the art in areas such as computer vision, sensor fusion, machine learning, object tracking, and motion planning

Job description
Waymo is the self-driving company that makes it safe and easy for people to get where they need to go. Our mission is to make self-driving vehicles available to everyone. Waymo was the first fully self-driving vehicle to take passengers from A to B at scale. Our software allows us to do things that no one else can.

We'd Like You To Have

- MS/PHD in Robotics, Computer Science, or equivalent practical experience
- Experience in hands-on robotics research and expertise in one or more of the following: computer vision, LIDAR, object tracking, sensor fusion, perception, machine learning, motion planning, and control
- Experience in data structures and advanced algorithms
- Experience programming in C++ and with robust, safety-critical, efficient code
- Experience with field robotics and systems design

Robotics Research Engineer
CyberCoders · Pittsburgh, PA, US
Posted 1 day ago · 8 views

Be an early applicant

NICE

- Experience with distributed computing, inter-processor communication, or low-level networking
- Experience with voting architectures, safety-critical systems, reliable computing, or redundant processing
- Experience in developing spacecraft electronics systems or flight software
- Experience with applications such as GPS-denied navigation, localization, SLAM, computer vision, and deep learning
- Experience with basic systems engineering principles, risks, and requirements development

Motivating Examples

- Humanoid/Legged Robots
- Manufacturing
- Transportation
- Civil/Construction Engineering
- Space Exploration
- Architecture
- Agriculture
- Medical/Health
- Entertainment

DARPA Robotics Challenge



Video from: <https://youtu.be/wX0KagJ1du8>

Boston Dynamics



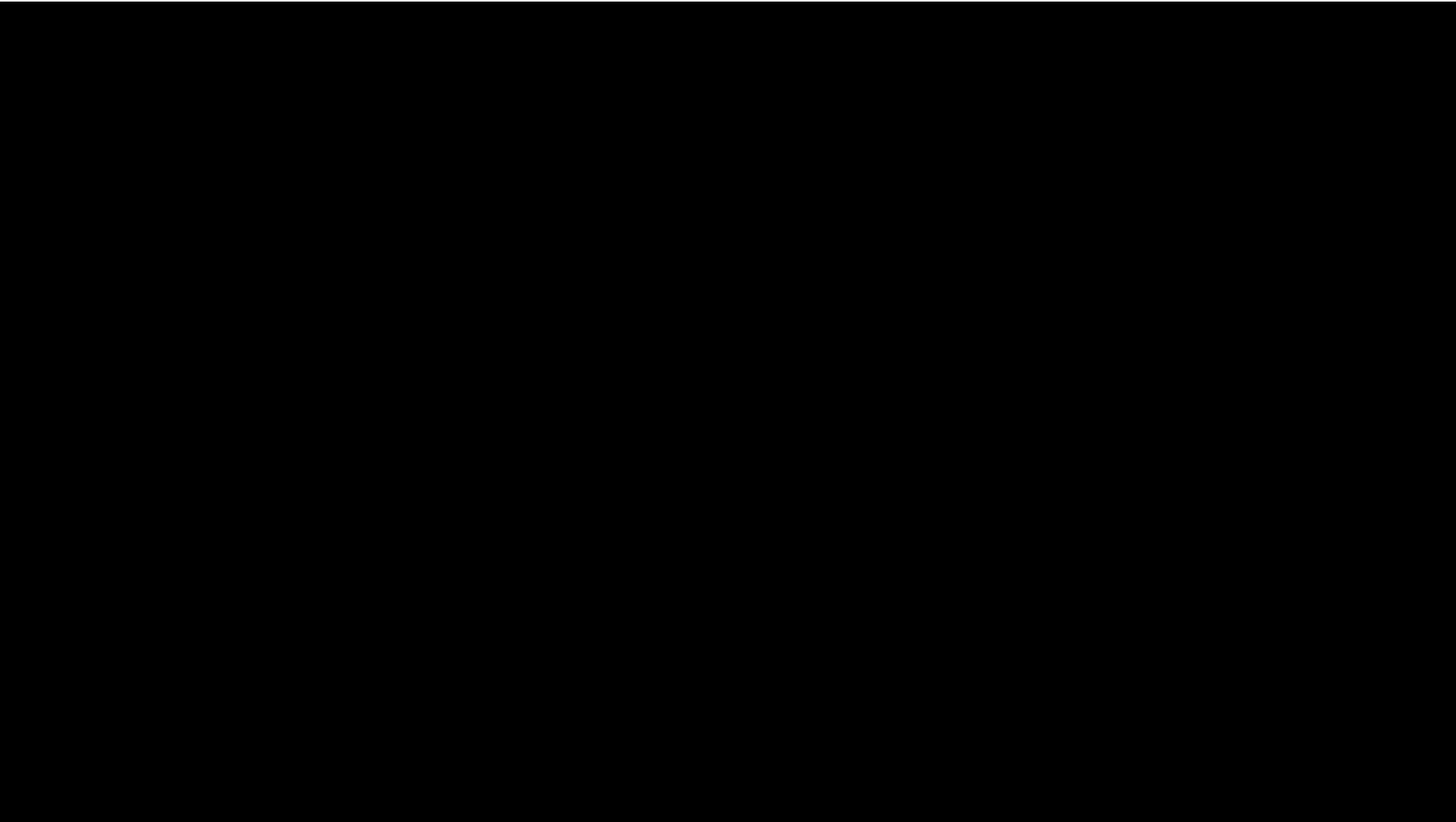
Video from: https://youtu.be/Ve9kWX_KXus

Drones in the lab



Video from: <https://youtu.be/w2itwFJCgFQ>

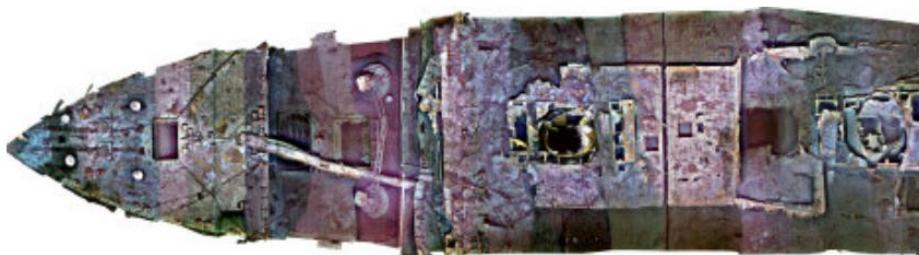
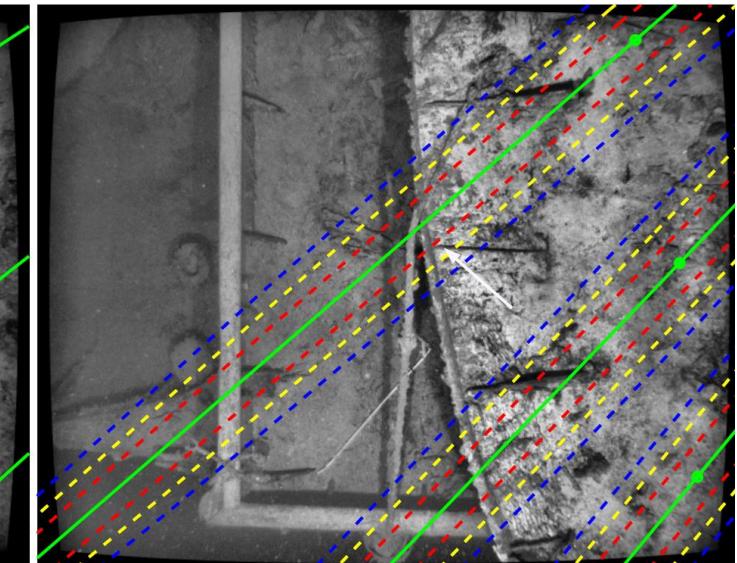
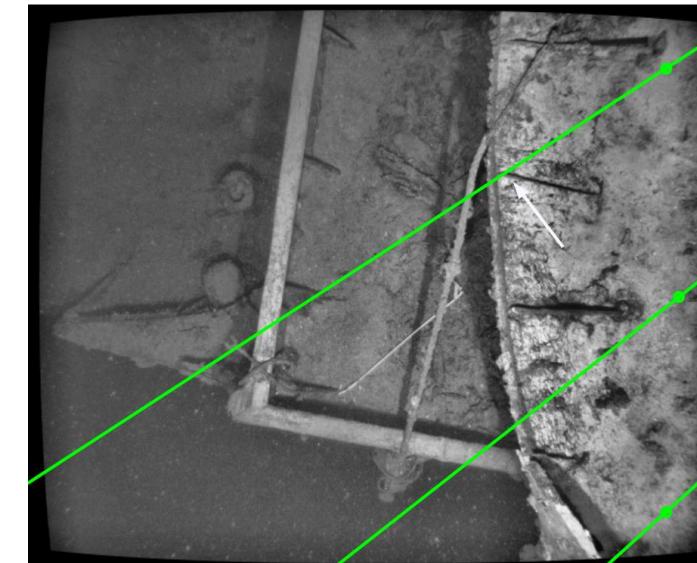
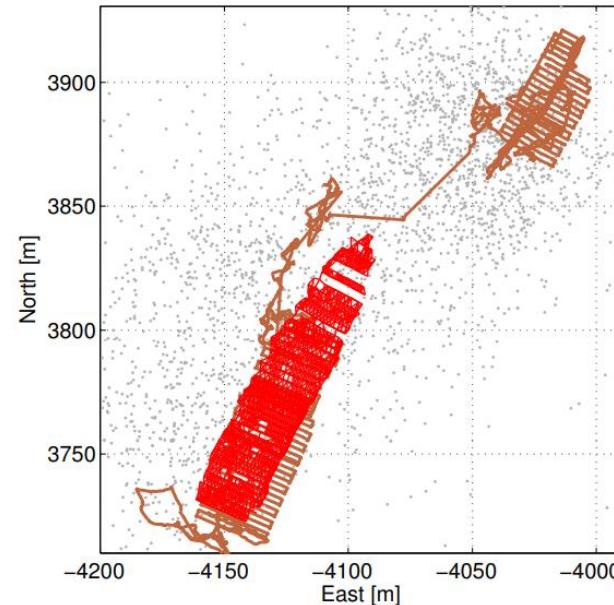
Drones in the Wild



Video from: https://youtu.be/p1d_ptE6yrc



Underwater Robots



Eustice, R., Singh, H., Leonard, J. J., Walter, M. R., & Ballard, R. (2005). *Visually Navigating the RMS Titanic with SLAM Information Filters*. In *Robotics: Science and Systems* (Vol. 2005, pp. 57-64).

Roomba



Video from: <https://youtu.be/tZ0bq-jlg-o>



Video from: <https://youtu.be/oj3Vawn-kRE>

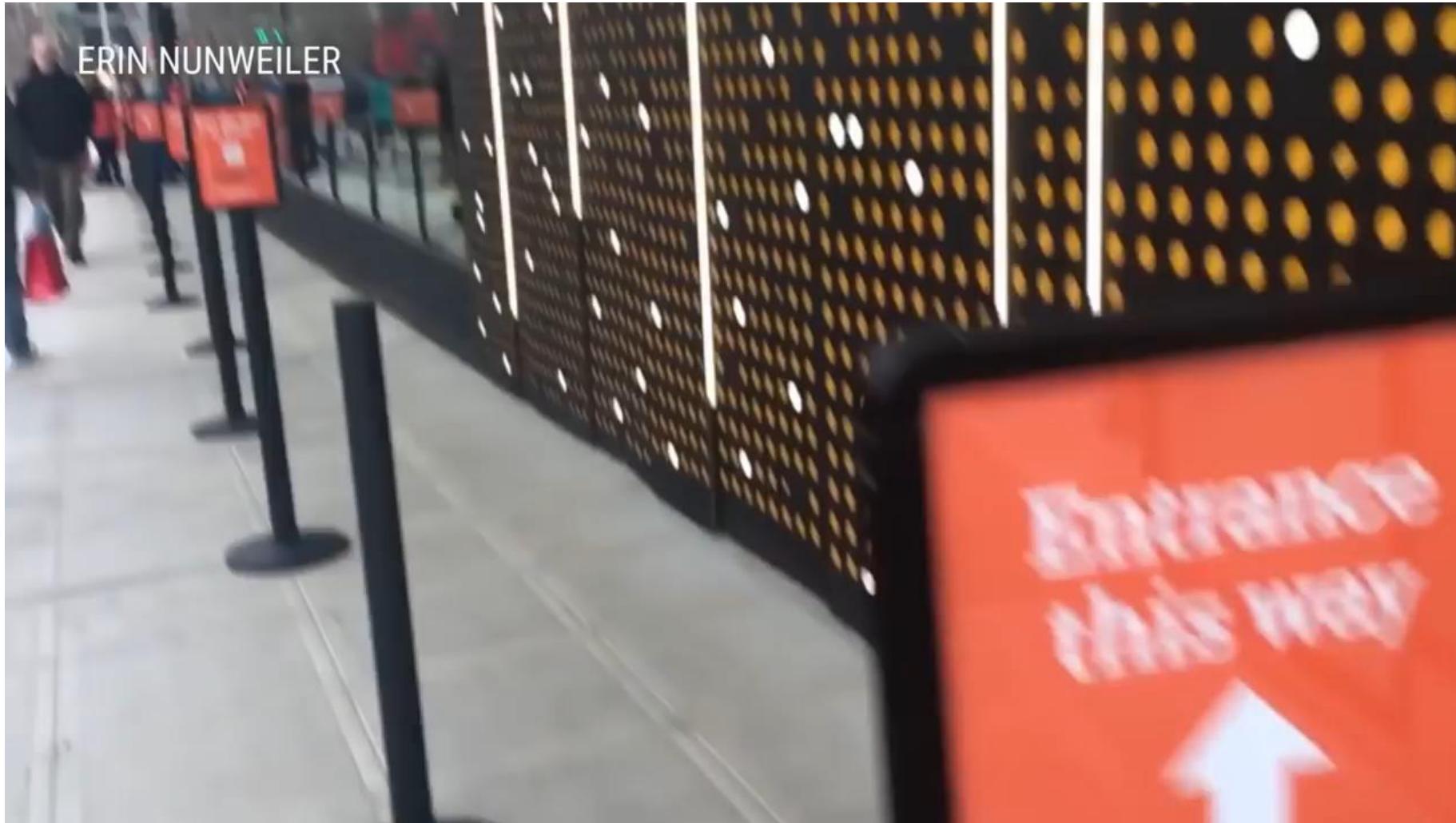
Amazon Robotics Challenge

Cartman: The low-cost Cartesian Manipulator that won the Amazon Robotics Challenge

D. Morrison, A.W. Tow, M. McTaggart, R. Smith, N. Kelly-Boxall, S. Wade-McCue, J. Erskine, R. Grinover, A. Gurman, T. Hunn, D. Lee, A. Milan, T. Pham, G. Rallos, A. Razjigaev, T. Rowntree, K. Vijay, Z. Zhuang, C. Lehnert, I. Reid, P. Corke, and J. Leitner

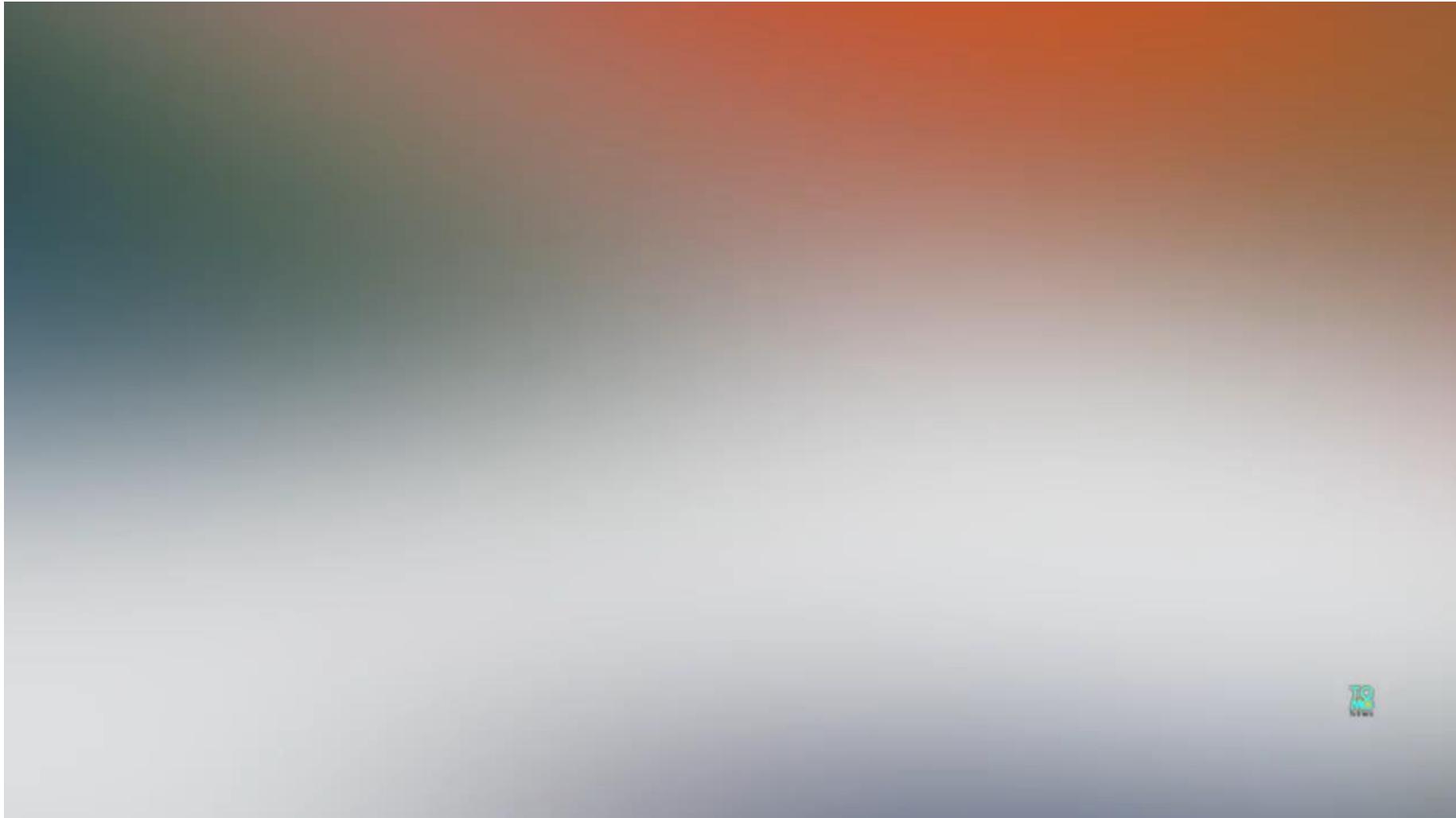


Amazon Go



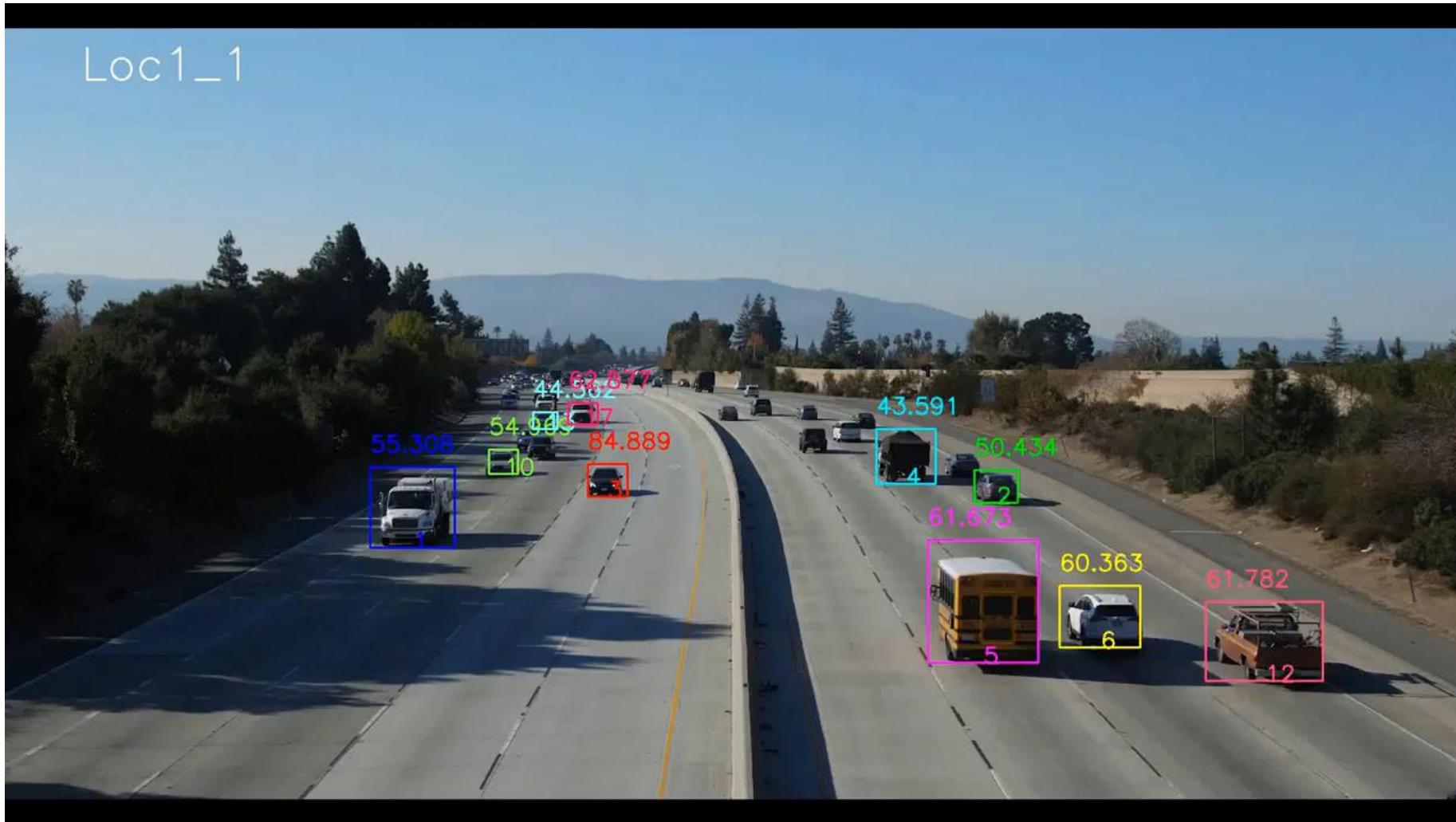
Video from: <https://youtu.be/zdbumR6Bhd8>

Autonomous Driving



Video from: <https://youtu.be/7oCe0aLye-U>

Transportation Surveillance



Video from: https://youtu.be/_i4numqiv7Y

Space Exploration



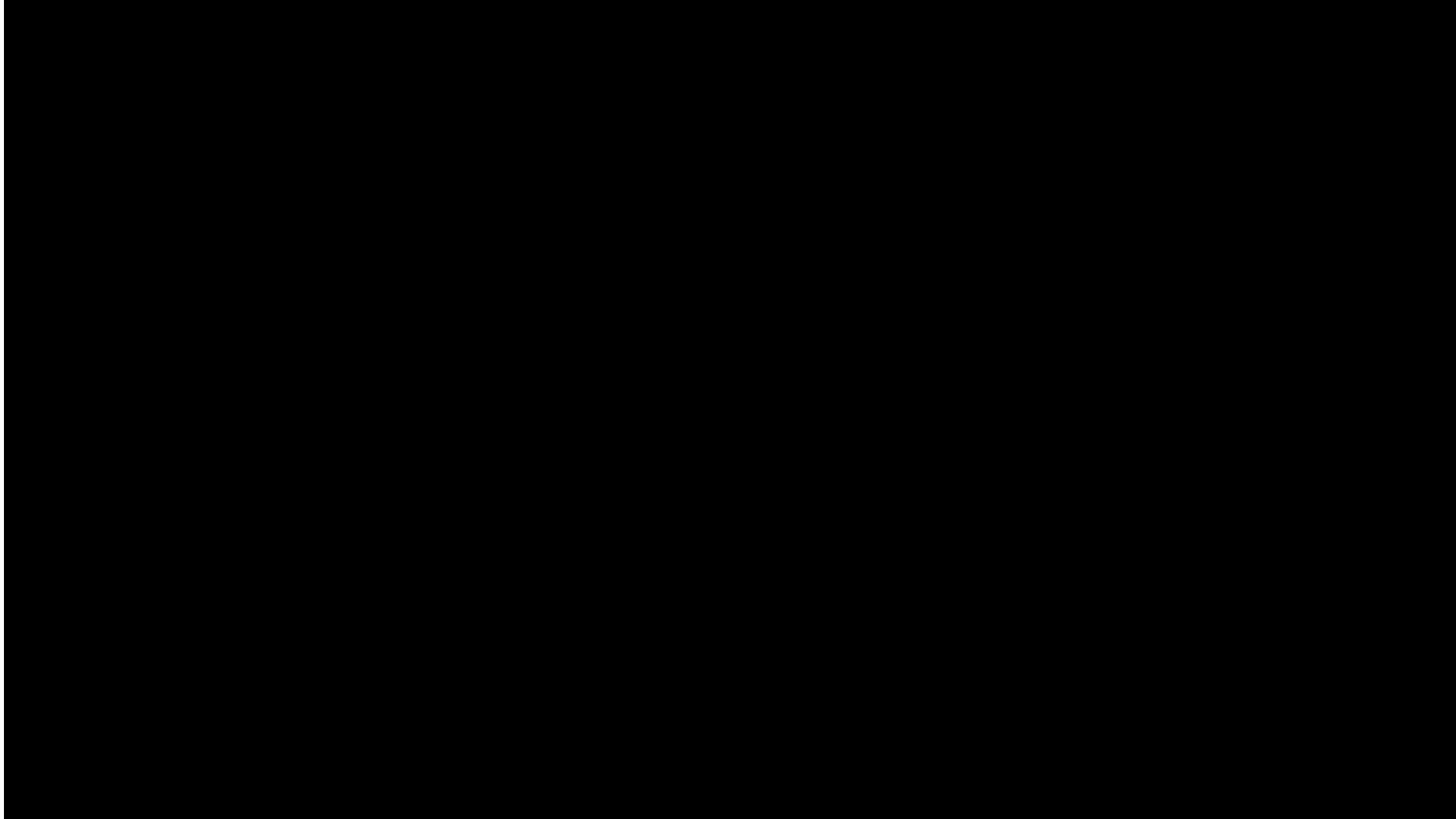
<https://mars.jpl.nasa.gov/mer/gallery/artwork/rover1browse.html>



Video from: <https://youtu.be/kr58r0b5LKM>



Smart Jobsite Surveillance



Video from: <https://youtu.be/pL-c00M2CnI>

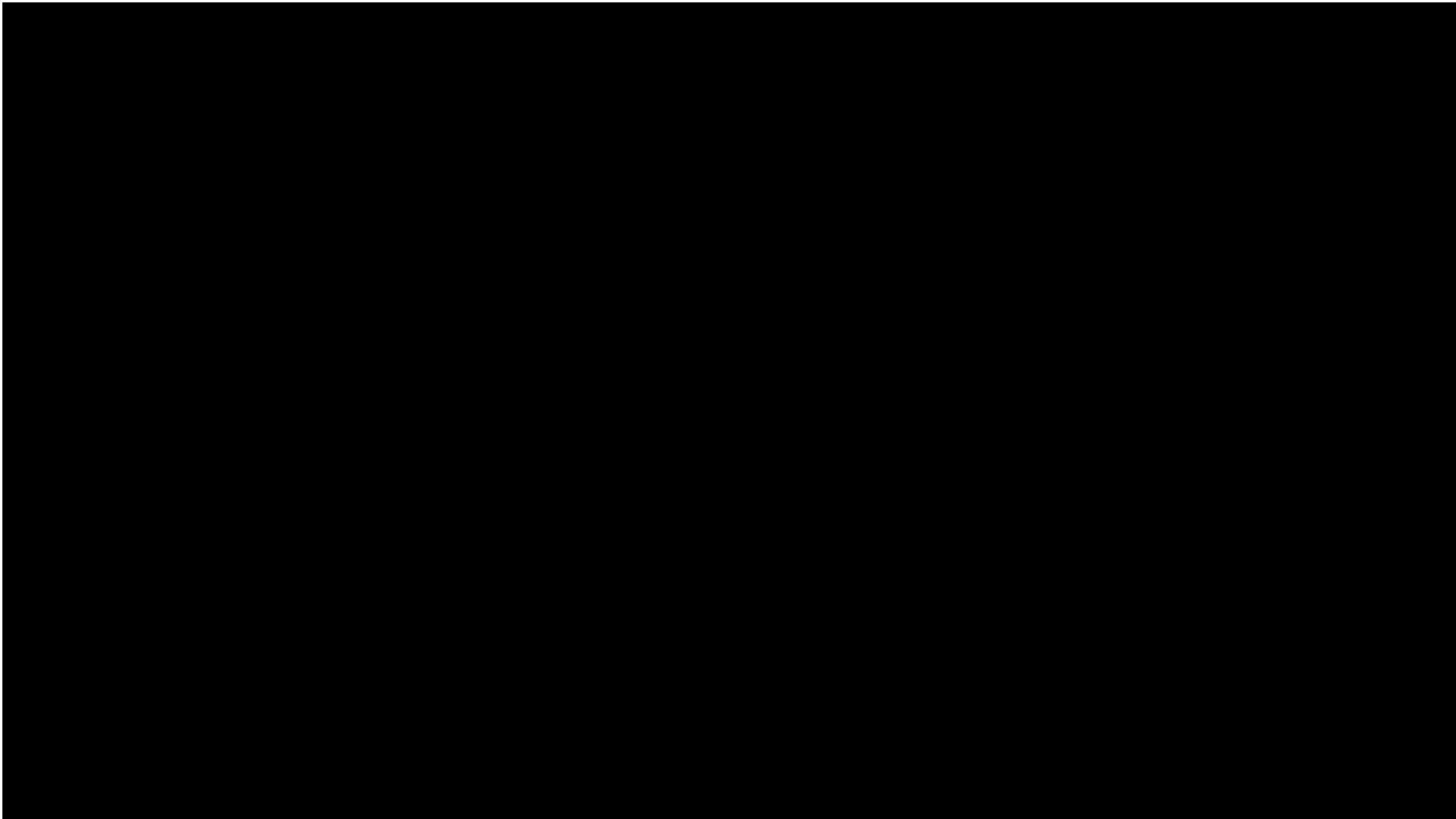
Robotic Jobsite Scanner



Video from: <https://www.doxel.ai>



Grit Blasting Robot



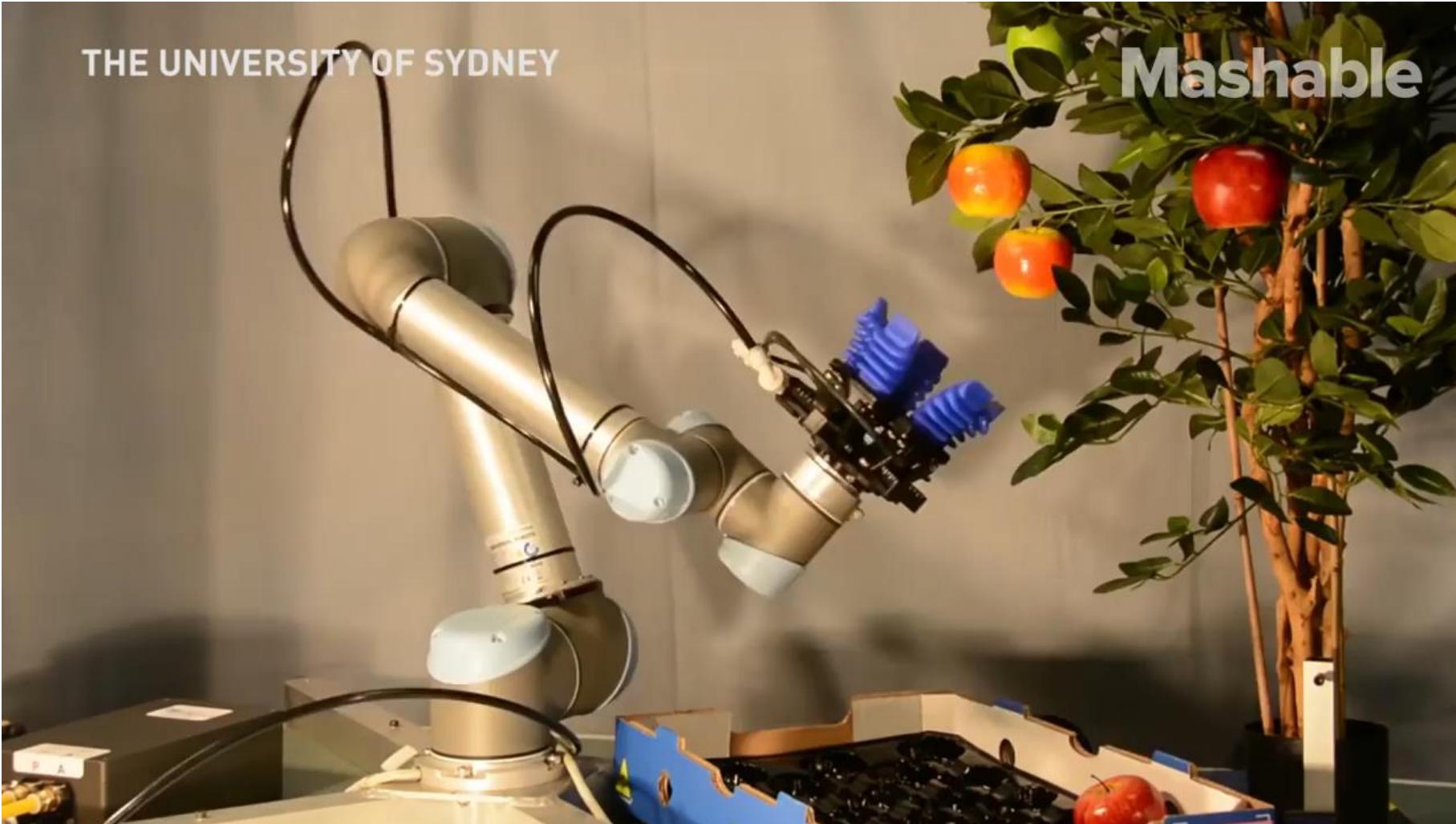
Video from: https://youtu.be/zREgx4NsA_Q

Architecture



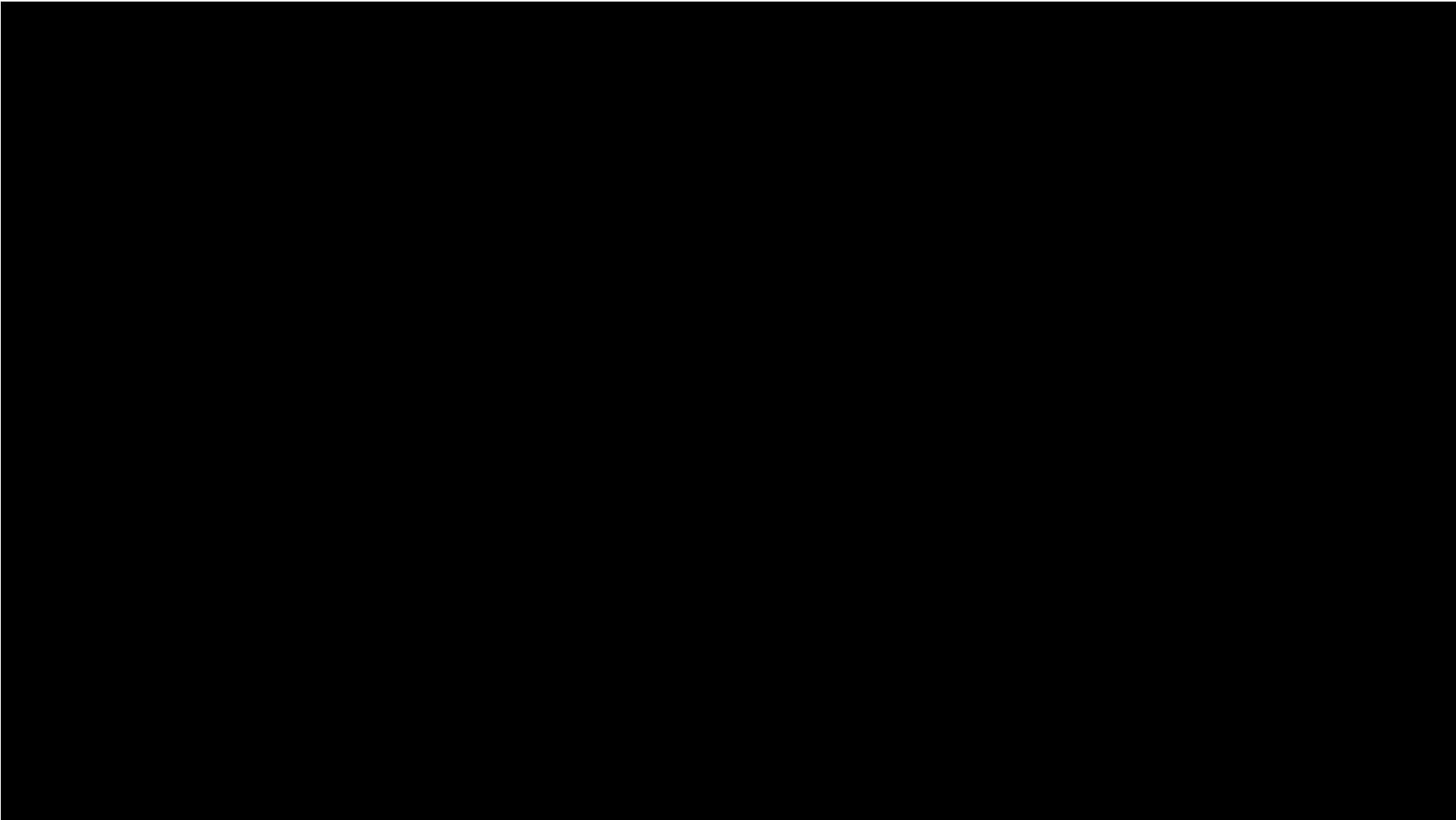
Video from: <https://youtu.be/TCJOQkOE69s>

Agriculture



Video from: <https://youtu.be/NO8PmqEI0cc>

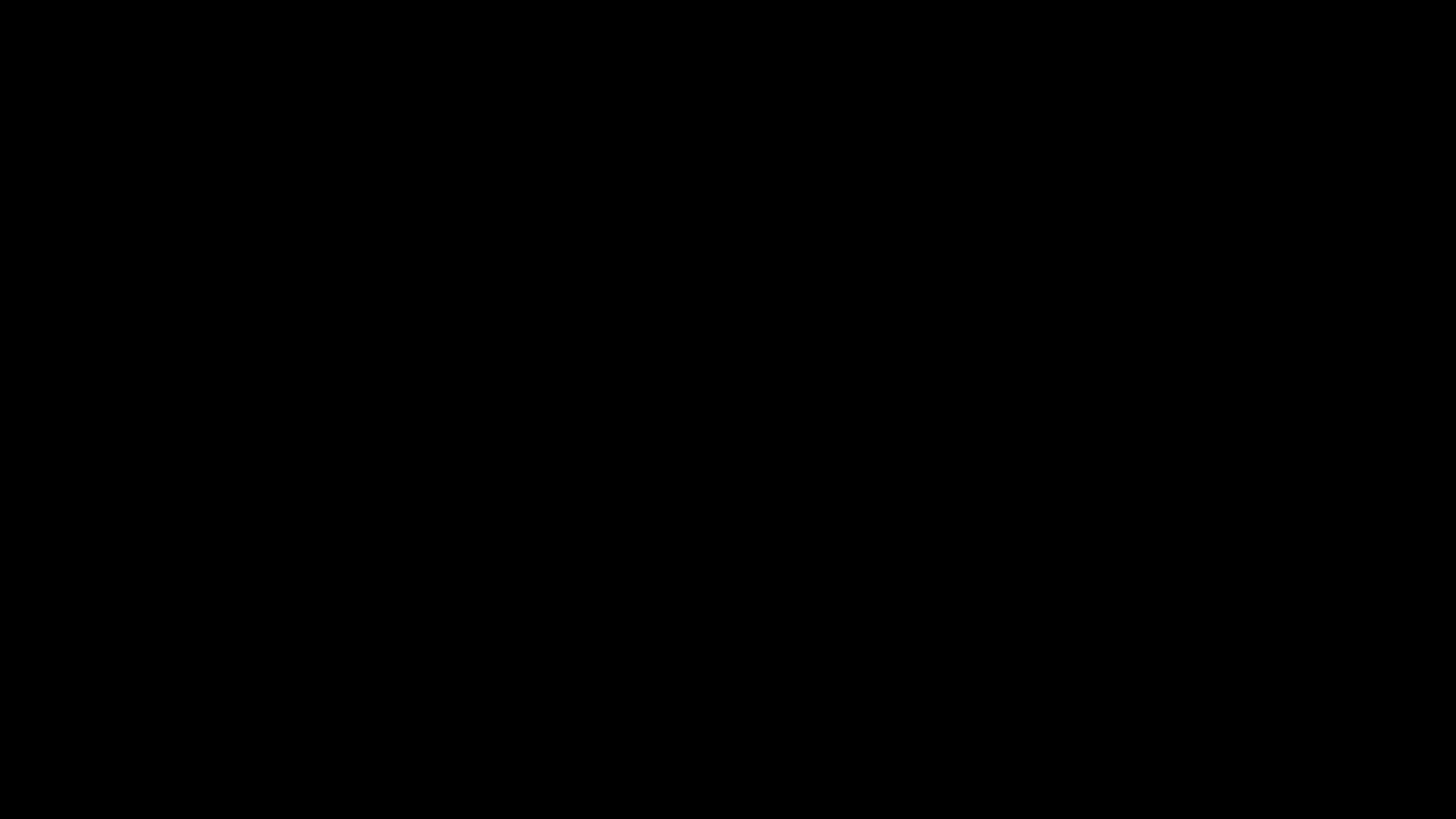
Surgical Robot



Video from: <https://youtu.be/R17lsilRjbM>



Hollywood AR Car



Video from: <https://youtu.be/OnBC5bwV5y0>

Vision for Tactile Sensing

GelSight

Retrographic sensing for touch, texture and shape

Micah K. Johnson, Edward H. Adelson and Alvin Raj

MIT Department of Brain and Cognitive Sciences
MIT Computer Science and Artificial Intelligence Lab



Massachusetts
Institute of
Technology

Presented at SIGGRAPH Emerging Technologies, 2009



Maybe one day we can see this...



References for Next Week

- Go2016
 - Section 6-6.1, 6.5, 9-9.4, 10-10.2, 14-14.2
- Sz2022
 - Section 5.3.1-5.3.2, 5.3.4-5.3.5, 5.4.1-5.4.4, 5.5.2, 5.5.4
- Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. “Deep residual learning for image recognition.” In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.
- <https://github.com/ajaymache/machine-learning-yearning>