



Computer Vision

Lecturer: Zhaopeng Cui
State Key Lab of CAD&CG



Outline Today

- Introduction of Computer Vision
- About the Course
- Camera



0. Introduction

What is Computer Vision

For 5 years old...



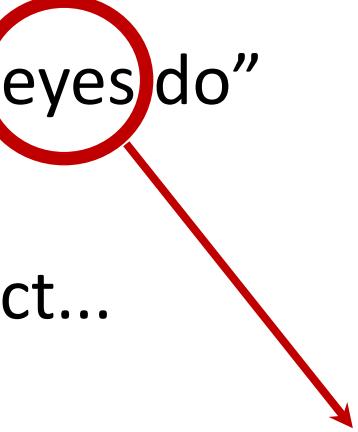


What is Computer Vision

- In layman's language:

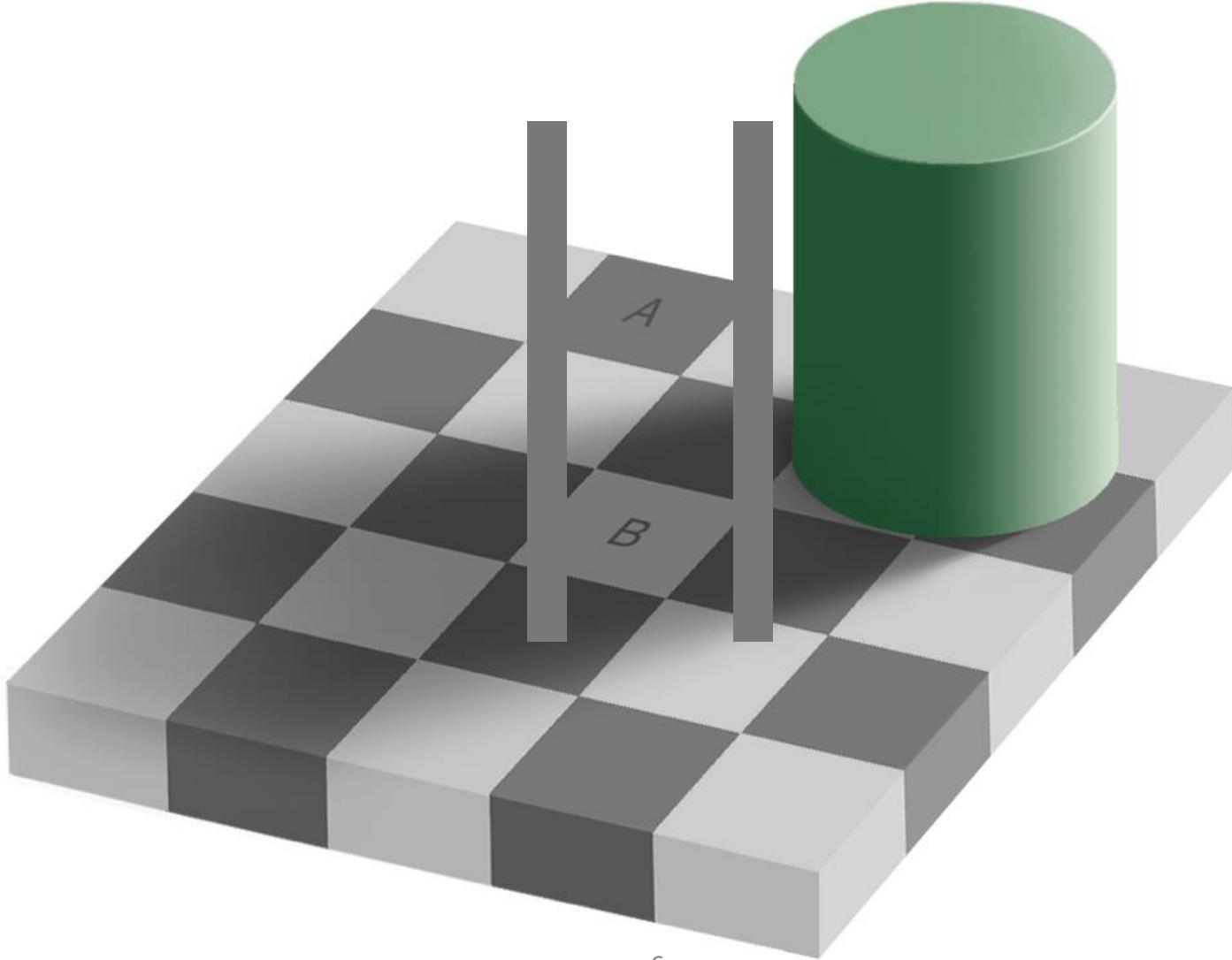
“Make computers do what our eyes do”

but this is not really correct...



visual system (eyes and brain)

Illusion 1



Illusion 2



[Camera Obscura Museum]

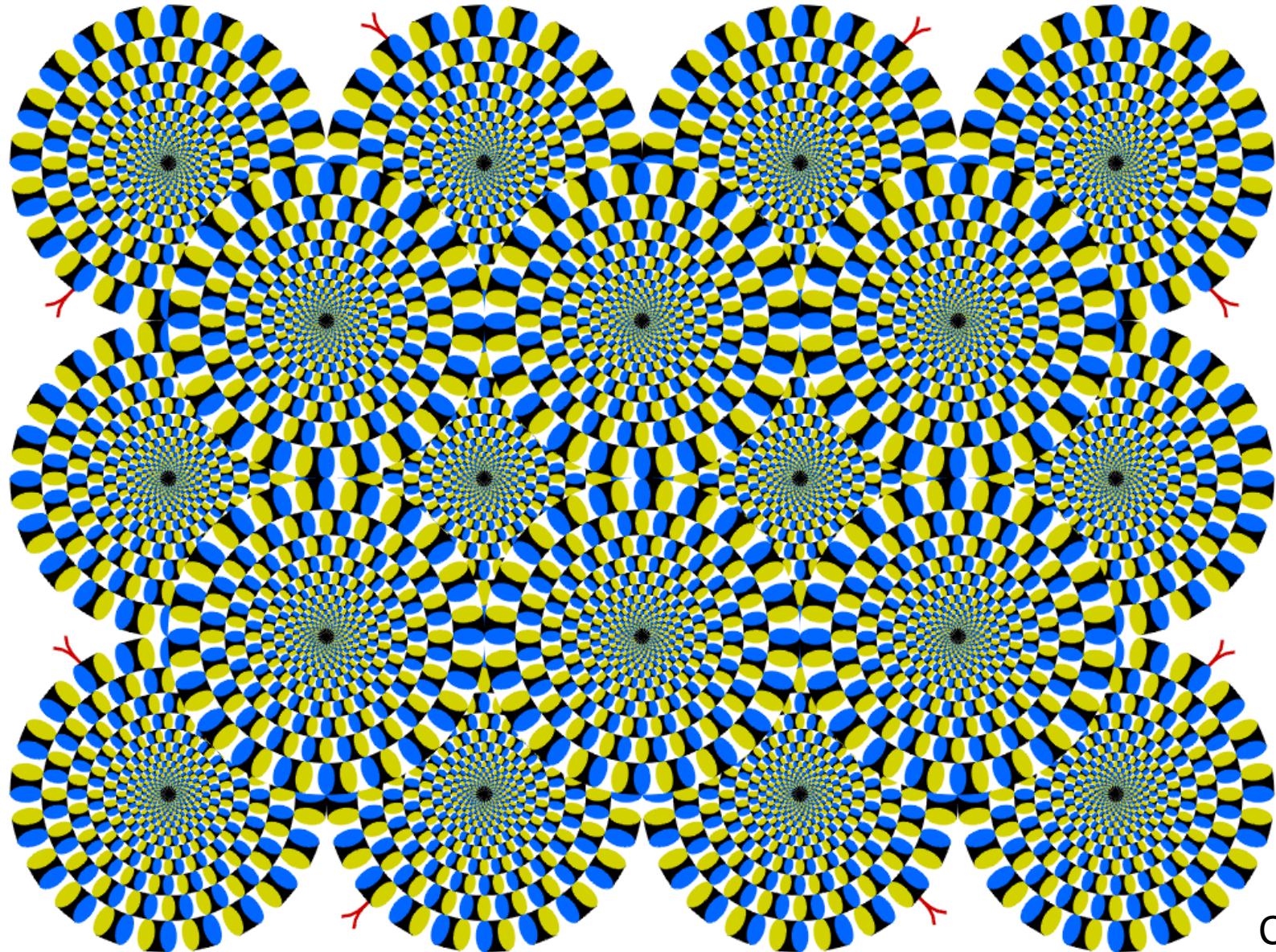
Illusion 3



Illusion 3



Illusion 4



Copyright [A.Kitaoka](#) 2003



What is Computer Vision

- A professional's comment:

In general, computer vision does not work
(except in certain situations/conditions)

It was true a few years ago...
Now,

Computer vision works in many important tasks



Face Capture and Retargeting

Face2Face: Real-time Face Capture
and Reenactment of RGB Videos

*Justus Thies¹, Michael Zollhöfer²,
Marc Stamminger¹, Christian Theobalt²,
Matthias Nießner³*

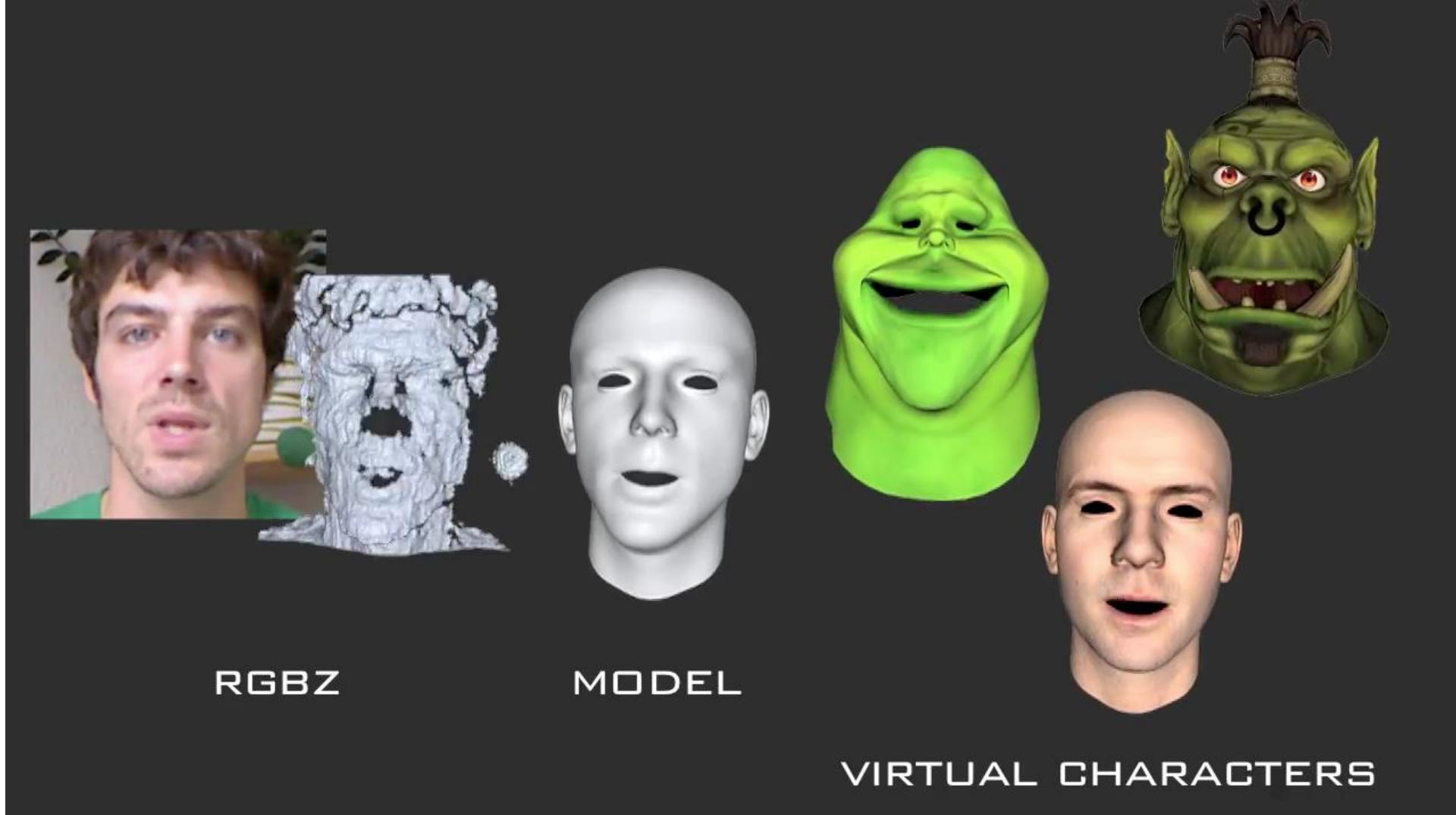
¹University of Erlangen-Nuremberg

²Max-Planck-Institute for Informatics

³Stanford University

CVPR 2016 (Oral)

Face Capture and Retargeting



Real-time Facial Motion Capture [FaceShift]
13

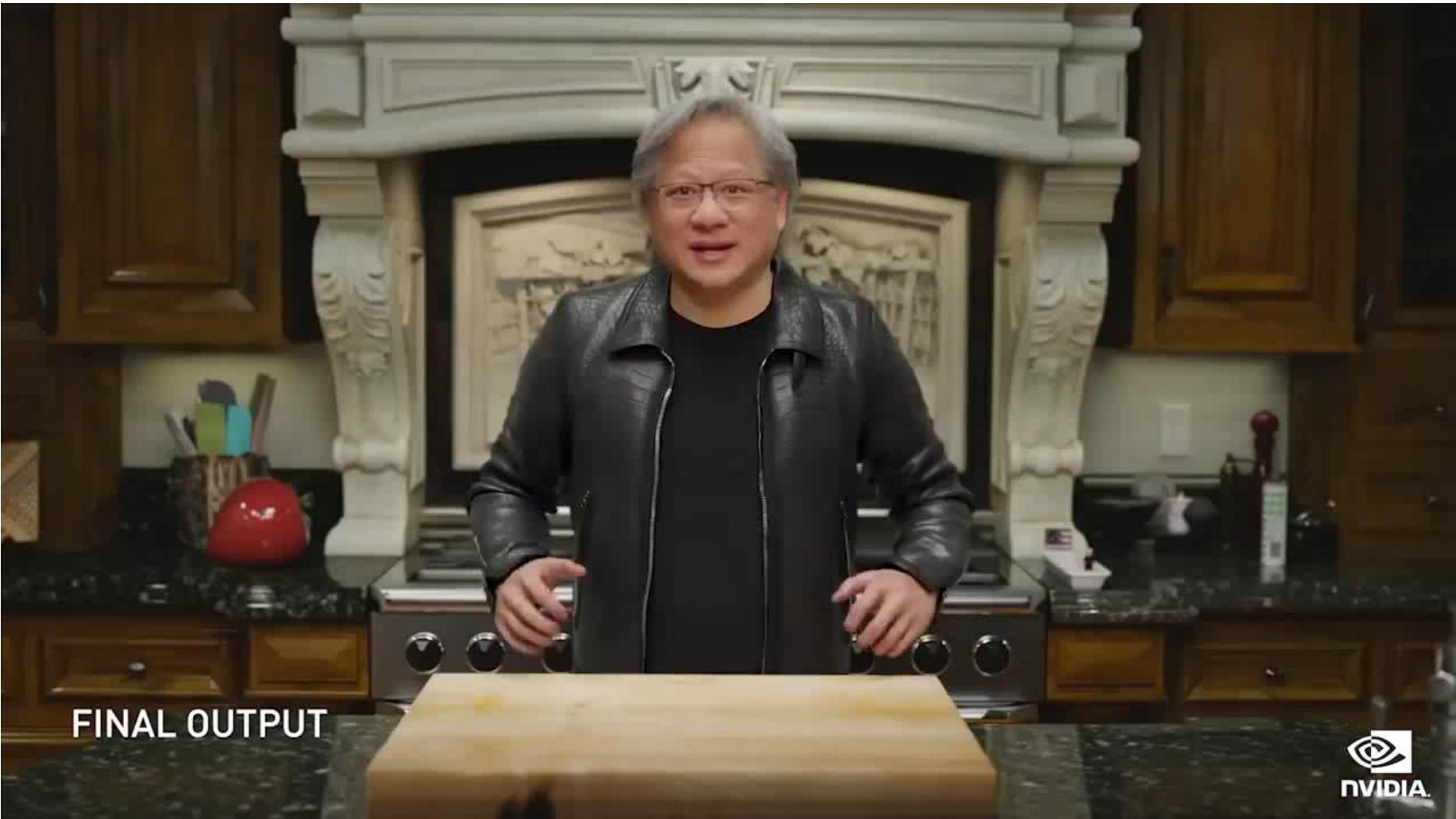
Human Pose Estimation



Human Pose Estimation



3D Avatar



3D Urban Modeling



Altizure

全自动大规模三维重建

广东某城市 A

52920 张 5000 万像素相片 = 2663 GP

无 GPS 信息

12 台服务器 10 天完成

Automatic Large-scale 3D Reconstruction

City A, Guangdong, China

52920 images @ 50MegaPixels = 2663 GigaPixels

Without GPS information

12 servers in 10 days

联系 support@altizure.com 获取大规模数据处理服务的相关信息

Contact support@altizure.com for a quotation

Autonomous Driving

Google driving to be driverless

Google's modified Toyota Prius uses an array of sensors to navigate public roads without a human driver. Other components, not shown, include a GPS receiver and an inertial motion sensor.

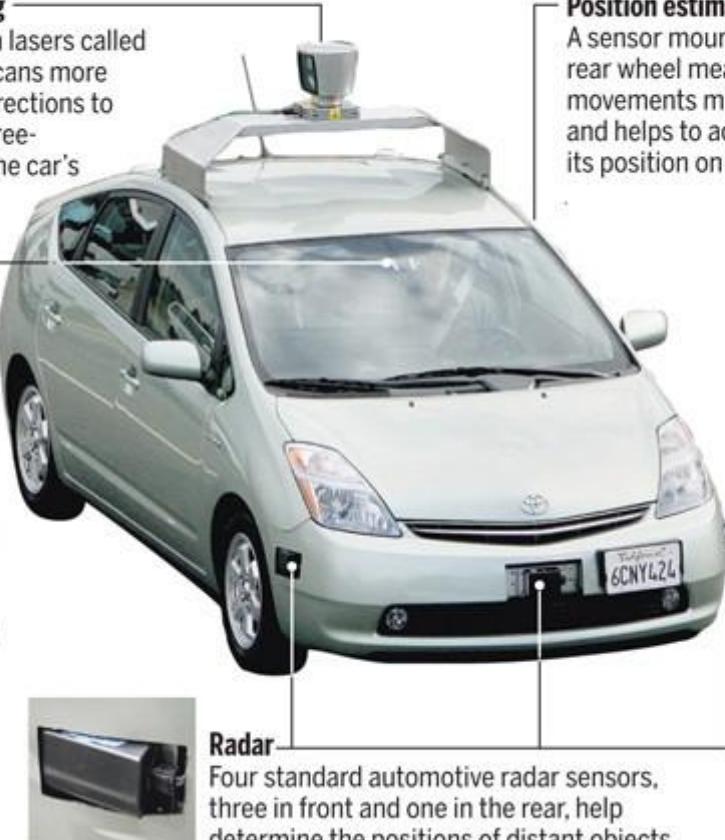
Laser-guided mapping

A rotating sensor with lasers called a LIDAR on the roof scans more than 200 feet in all directions to generate a precise three-dimensional map of the car's surroundings.

Video camera



A camera mounted near the rear-view mirror detects traffic lights and helps the car's onboard computers recognize moving obstacles—such as pedestrians and bicyclists.



Position estimator

A sensor mounted on the left rear wheel measures small movements made by the car and helps to accurately locate its position on the map.

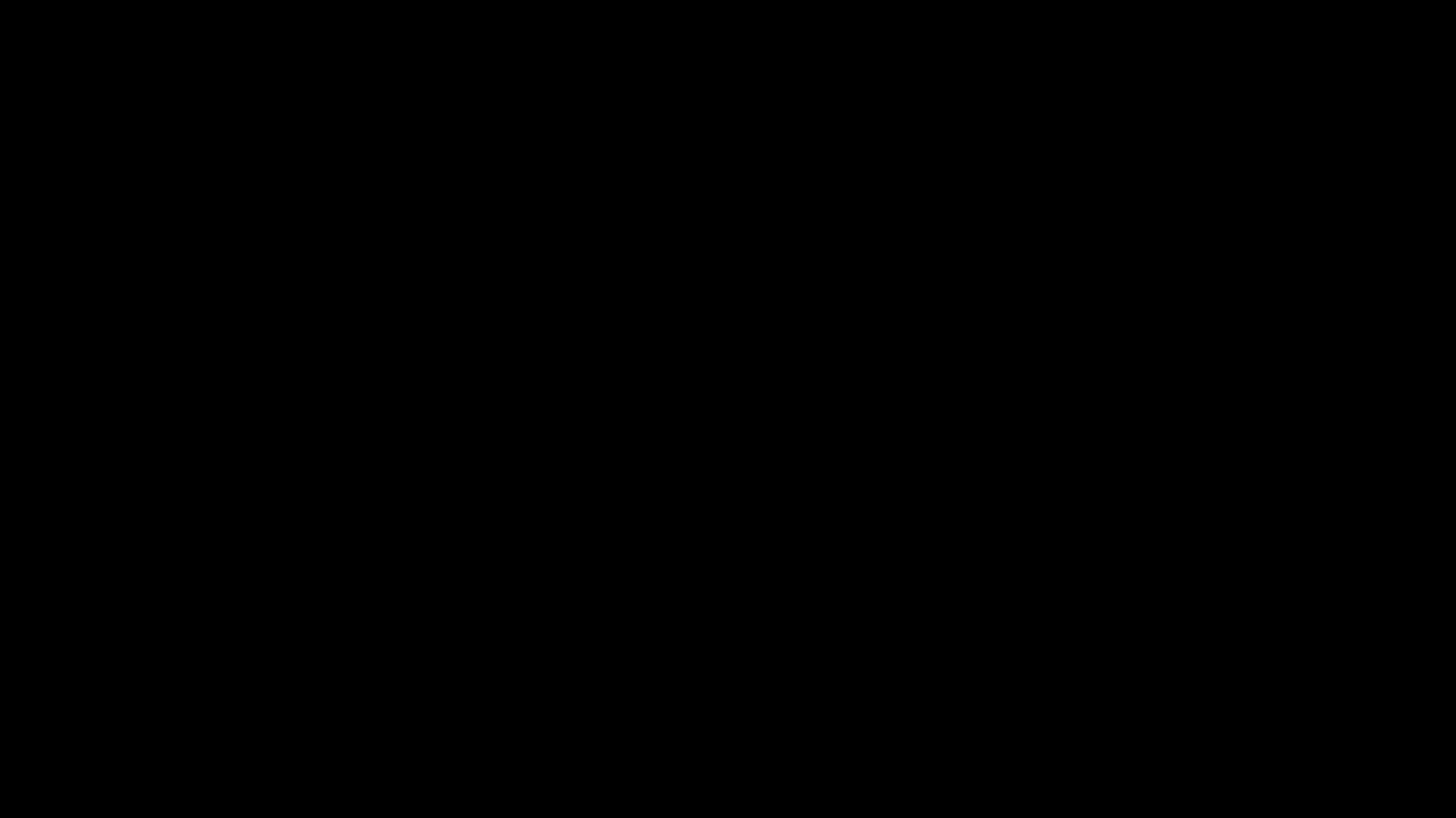


Source: Google

NEW YORK TIMES: PHOTOGRAPHS BY RAMIN RAHIMIAN FOR THE NEW YORK TIMES

Google Self-Driving Cars

Autonomous Driving



Autopilot Full Self-Driving

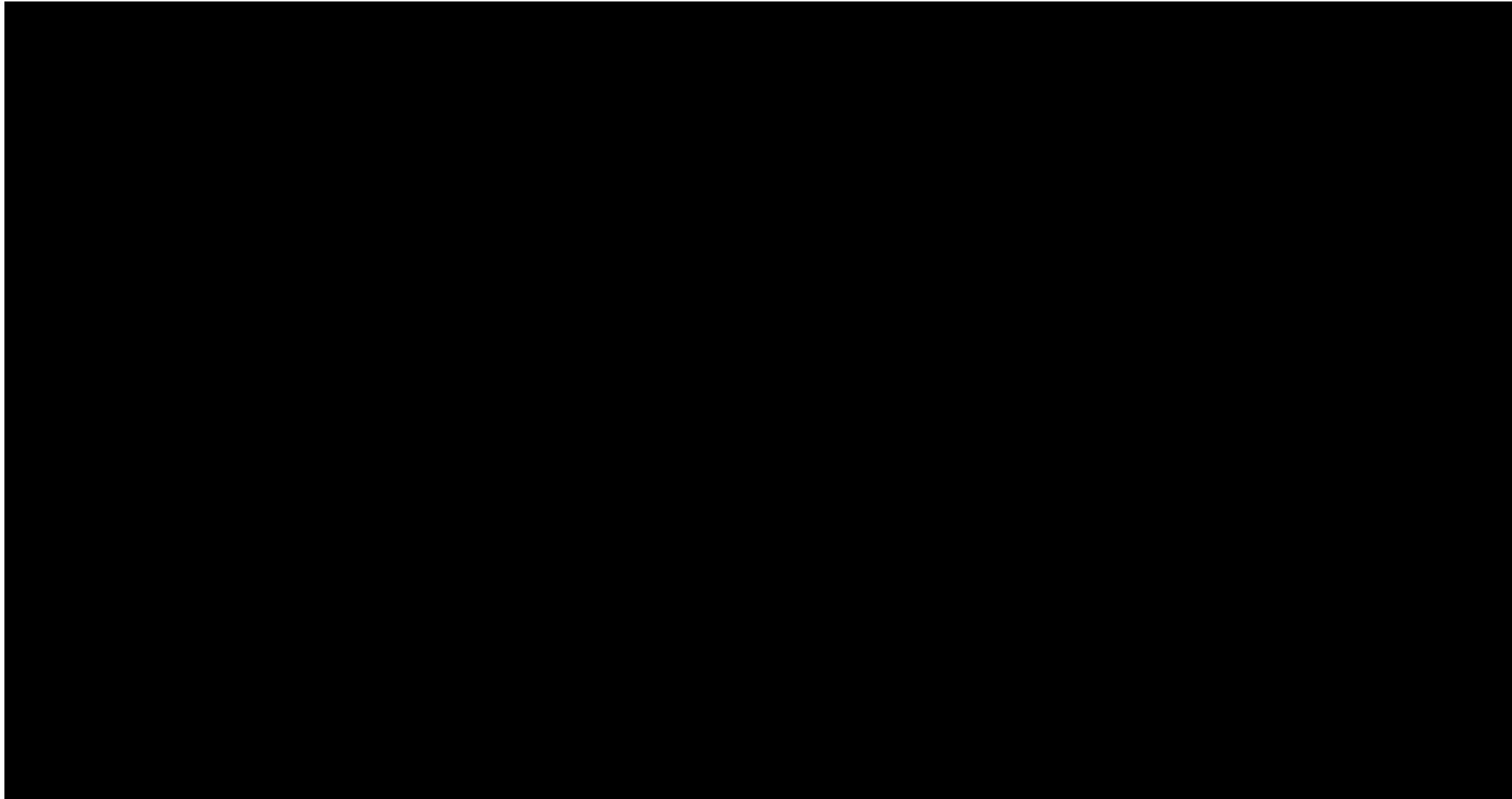
Robotics



3D Telepresence



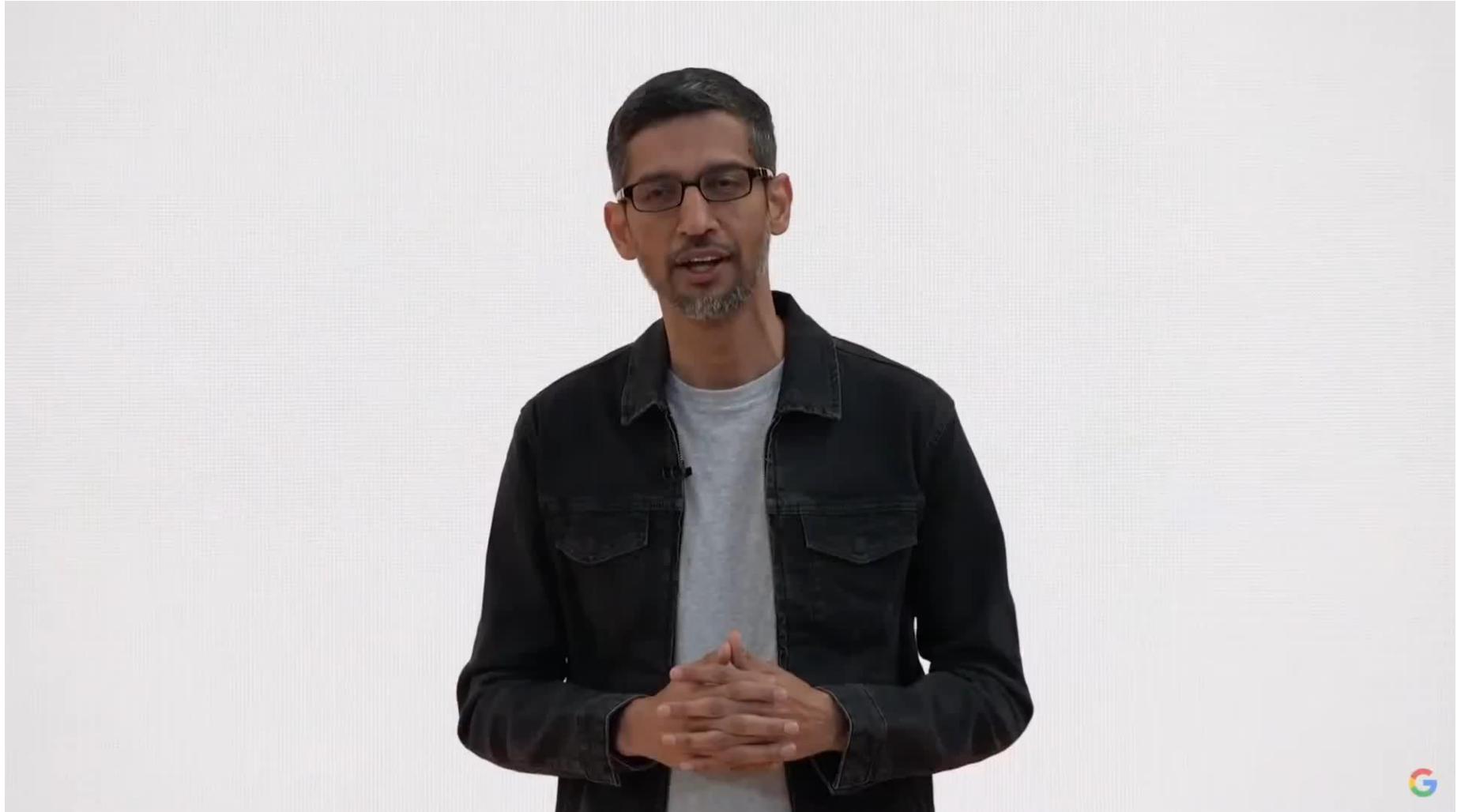
Mixed Reality (AR&VR)



Metaverse



Immersive Map



Google

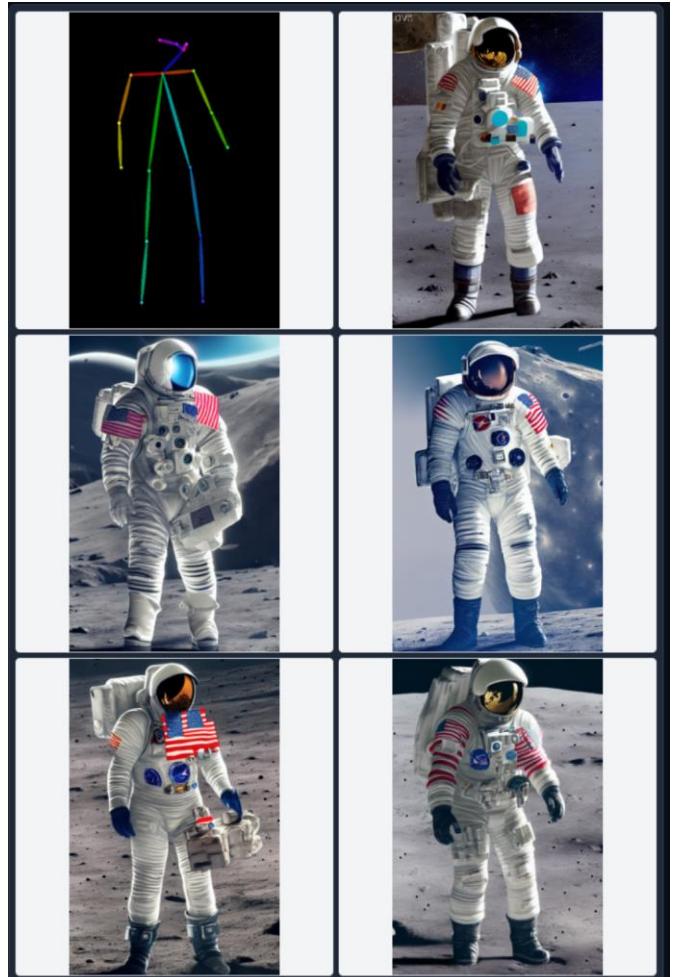
Text-to-Image



Midjourney
《太空歌剧院》



网易丹青约
《飞流直下三千尺》



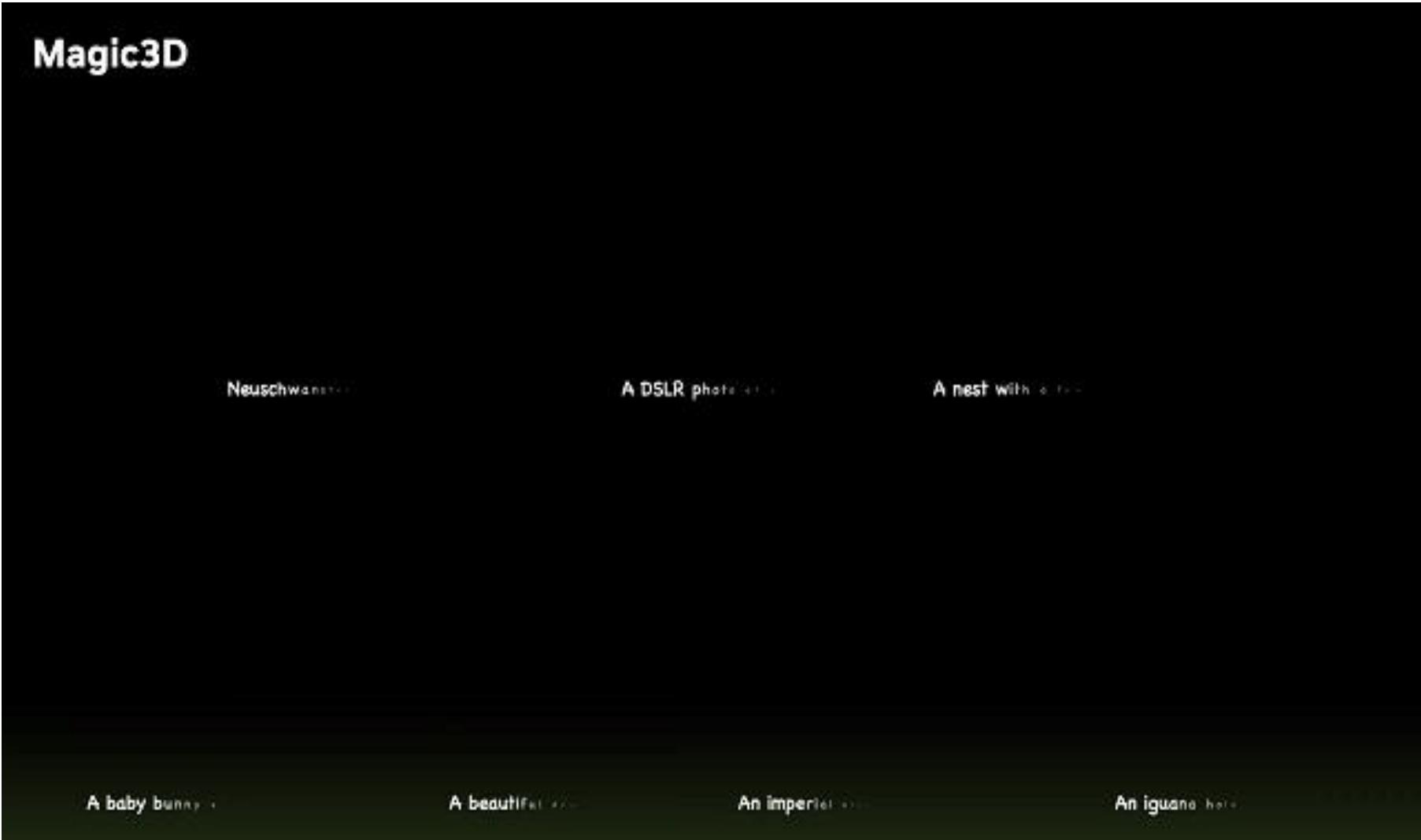
ControlNet
《月球上的宇航员》

Text-to-Video



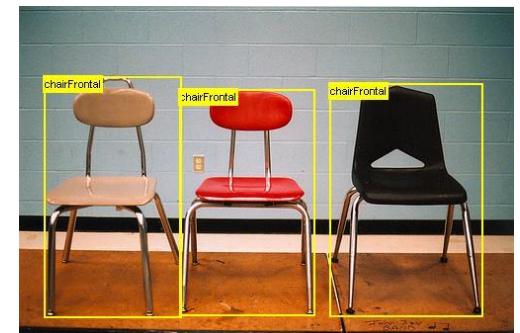
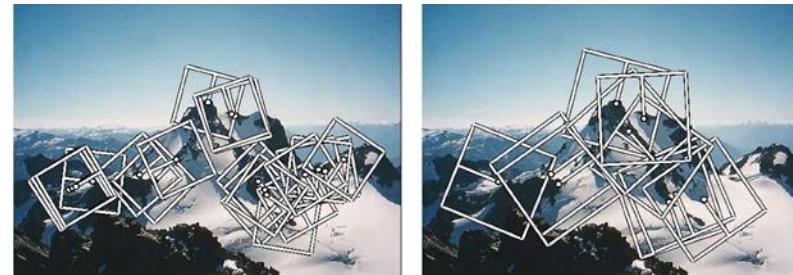
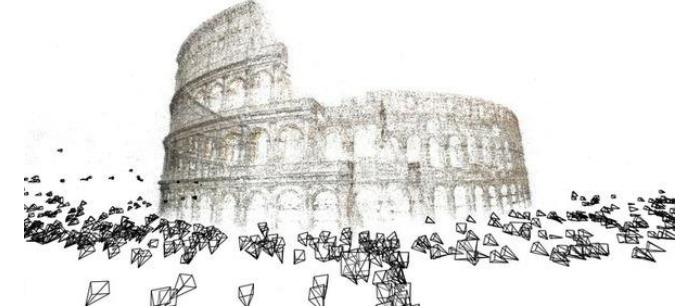
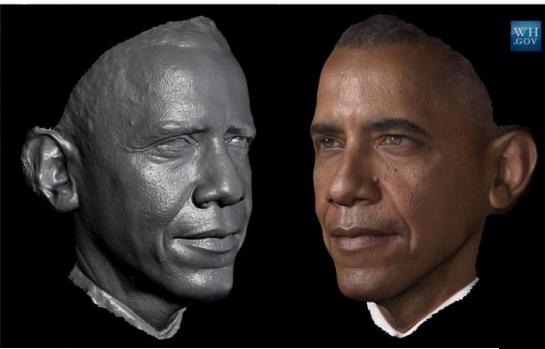
GEN-2

Text-to-3D

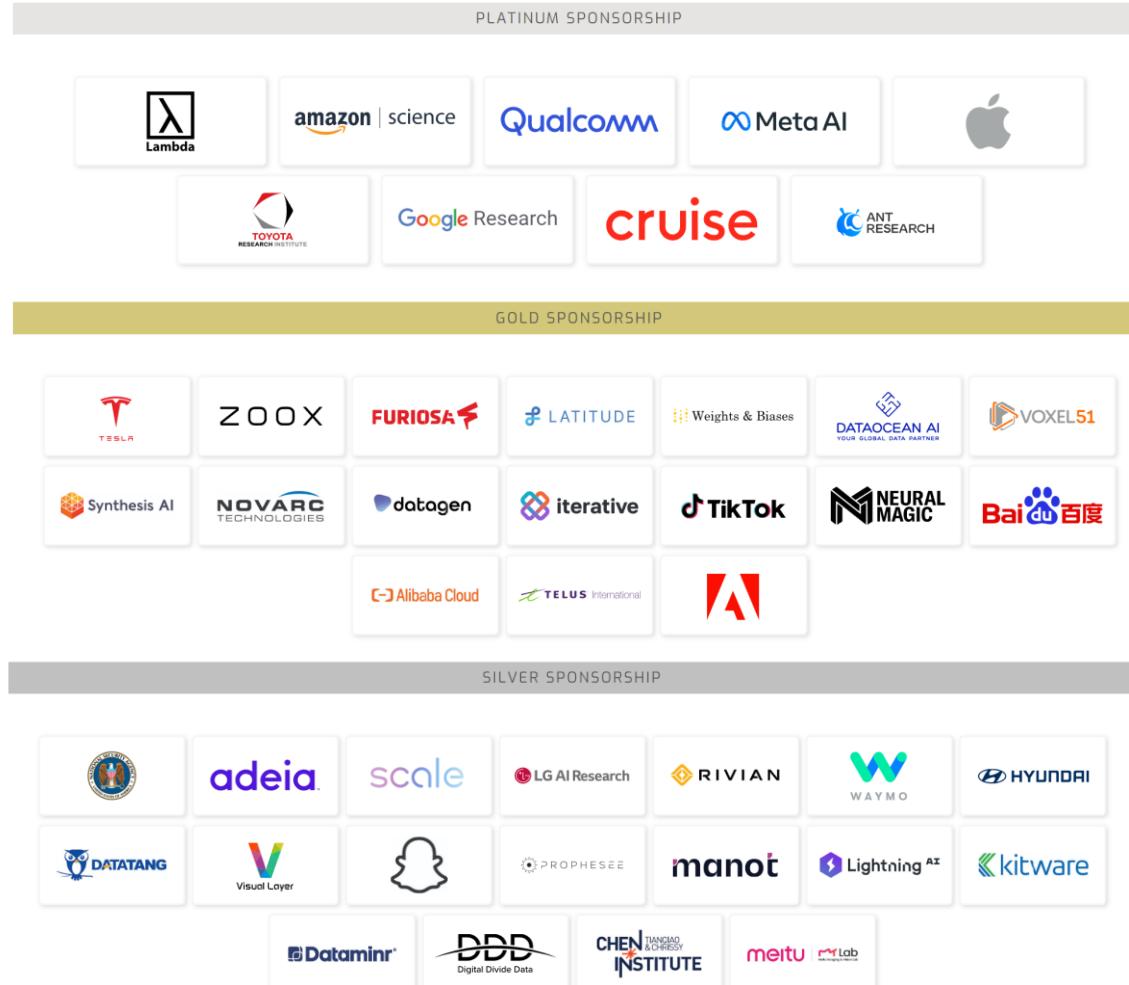


So, what is computer vision really?

- A discipline highly relevant to AI, robotics, computer graphics
- Contains many different sub-fields / problems
 - Physically-based computer vision: color, reflectance, photometric stereo, shape-from-shading, etc.
 - Geometric vision: camera models, epipolar geometry, structure-from-motion, stereo, multi-view stereo, etc.
 - Image/video processing: filtering, corner & edge detection, optical flow, tracking, segmentation, etc.
 - Image understanding: classification, detection, semantic segmentation, etc.
 - Statistics and psychophysics: robust estimation, saliency, etc.
 - DEEP LEARNING (started from classification, now applied everywhere)



It's a good time to do computer vision!



Microsoft at CVPR 2023: Pushing the boundaries of computer vision

Published June 20, 2023

By [Baining Guo](#), Distinguished Scientist; [Steve Lin](#), Senior Principal Research Manager

Share this page



Google at CVPR 2023

SUNDAY, JUNE 18, 2023

Posted by [Shaina Mehta](#), Program Manager, Google

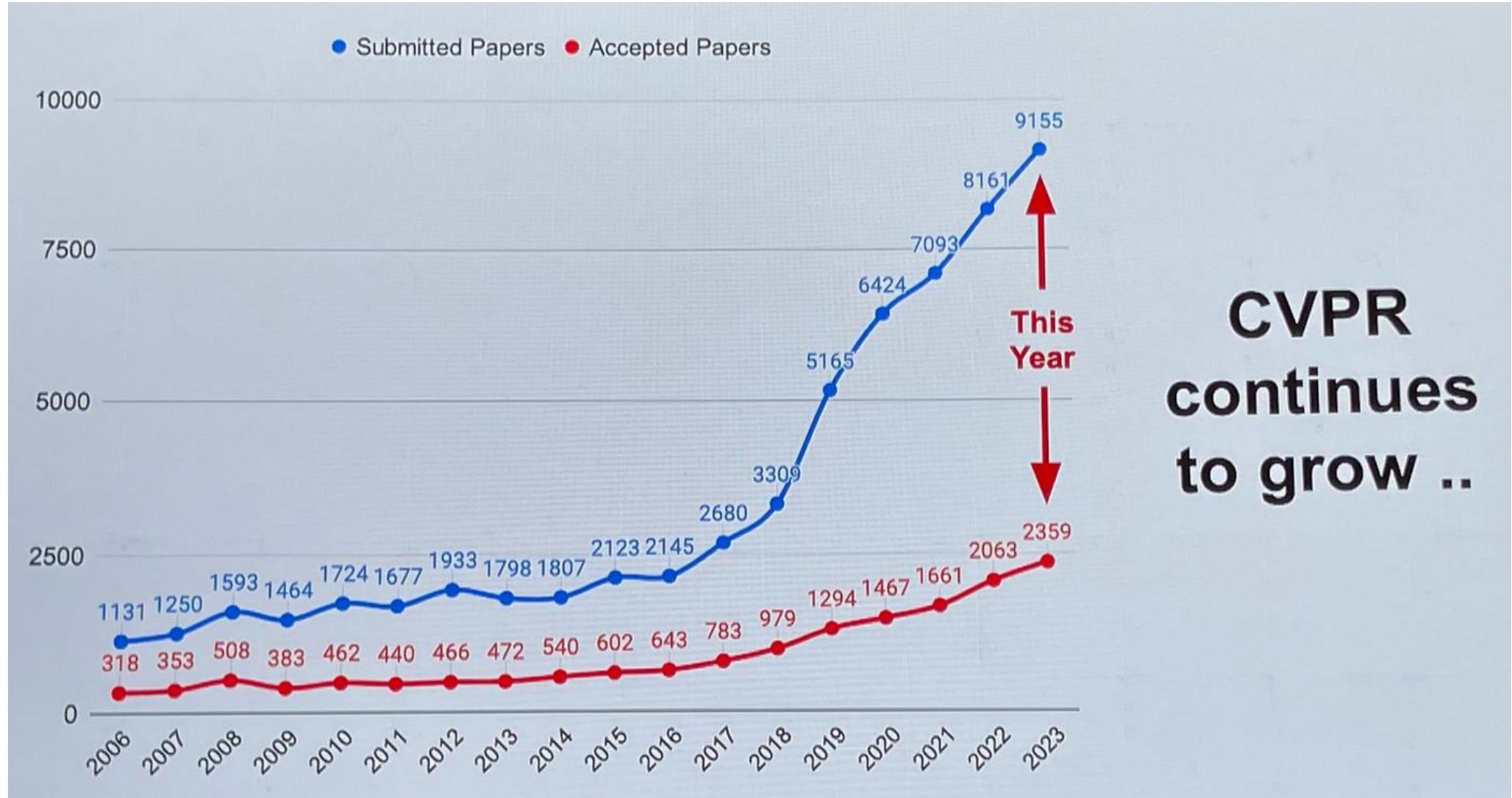
This week marks the beginning of the premier annual [Computer Vision and Pattern Recognition](#) conference (CVPR 2023), held in-person in Vancouver, BC (with additional virtual content). As a leader in computer vision research and a [Platinum Sponsor](#), [Google Research](#) will have a strong presence across CVPR 2023 with ~90 papers being presented at the [main conference](#) and active involvement in over 40 conference [workshops](#) and [tutorials](#).

If you are attending CVPR this year, please stop by our booth to chat with our researchers who are actively exploring the latest techniques for application to various areas of [machine perception](#). Our researchers will also be available to talk about and demo several recent efforts, including on-device ML applications with [MediaPipe](#), strategies for differential privacy, neural radiance field technologies and much more.

You can also learn more about our research being presented at CVPR 2023 in the list below (Google affiliations in bold).



It's a good time to do computer vision!





It's a good time to do computer vision!

≡ Top publications

Categories ▾ English ▾

	Publication	<u>h5-index</u>	<u>h5-median</u>
1.	Nature	<u>467</u>	707
2.	The New England Journal of Medicine	<u>439</u>	876
3.	Science	<u>424</u>	665
4.	IEEE/CVF Conference on Computer Vision and Pattern Recognition	<u>422</u>	681
5.	The Lancet	<u>368</u>	688
6.	Nature Communications	<u>349</u>	456
7.	Advanced Materials	<u>326</u>	415
8.	Cell	<u>316</u>	503
9.	Neural Information Processing Systems	<u>309</u>	503
10.	International Conference on Learning Representations	<u>303</u>	563

Questions?





Course Logistics

- Goal: Lead students into research in the field of computer vision
 - Provide students with a good understanding of computer vision and its applications.
 - The basic concepts and techniques will be studied in-depth.
 - Some practical algorithms will be explored through the exercises.
 - May not introduce the most advanced/recent algorithms
- Students are expected to be able to:
 - Understand most of the research papers at computer vision conferences
 - Apply computer vision techniques for various applications



Course Prerequisites

- Familiarity with linear algebra
 - Eigen-value, eigen-vectors, linear least square
- Familiarity with probability
 - Probability, conditional probability, Gaussian distribution
- Familiarity with Matlab or C++ (this is very important for you to finish course projects)



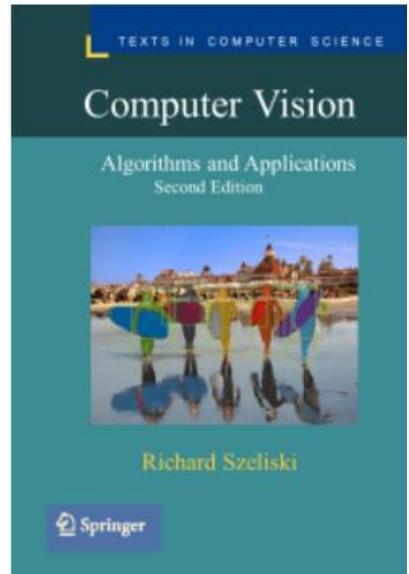
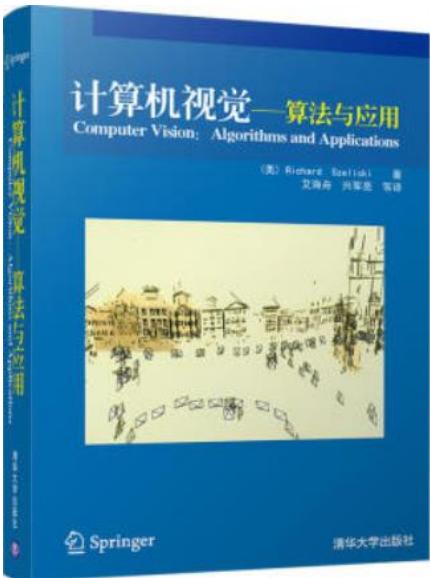
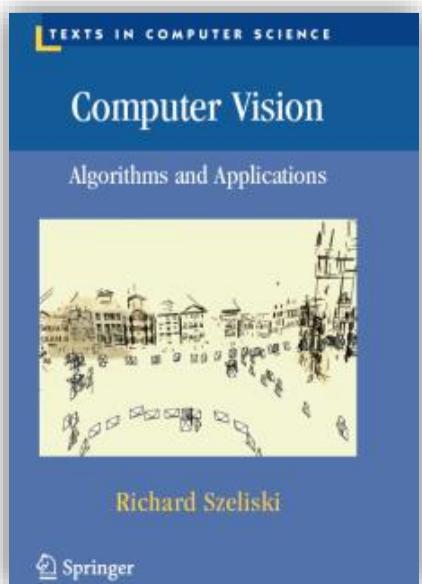
Complementary Courses

- Computer Graphics (计算机图形学)
 - Image Processing and Modeling(图像处理与建模)
 - Pattern Recognition and Machine Learning (模式识别与机器学习)
 - Virtual Reality (虚拟现实)
- ...



Recommended Textbook

- 《Computer Vision: Algorithms and Applications》 Springer
© 2010 Richard Szeliski, Microsoft Research (Now in Facebook)
<http://szeliski.org/Book/>



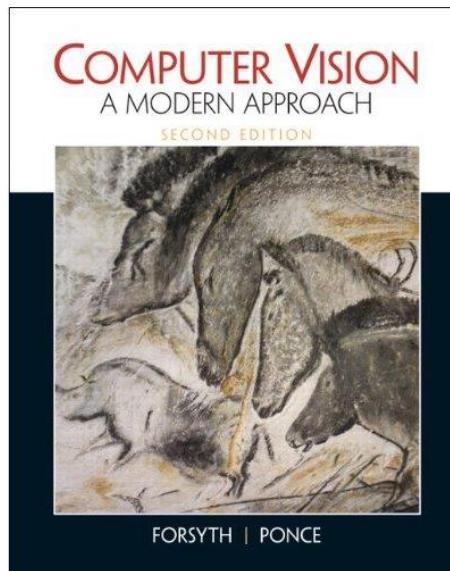
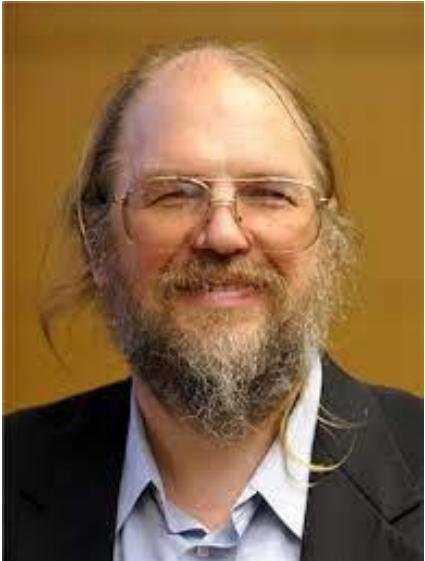
2nd edition

PDFs can be Downloaded from the Author's Website



Other Recommended Textbooks

- 《Computer Vision: A Modern Approach》 Prentice Hall
© 2010 David A. Forsyth (UIUC) and Jean Ponce (Inria)
<http://luthuli.cs.uiuc.edu/~daf/CV2E-site/cv2eindex.html>



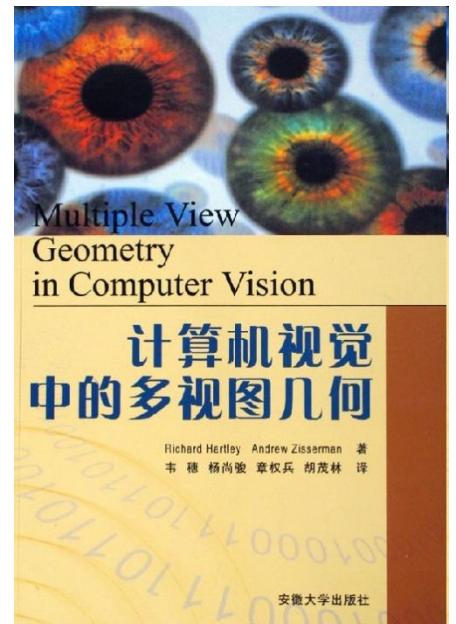
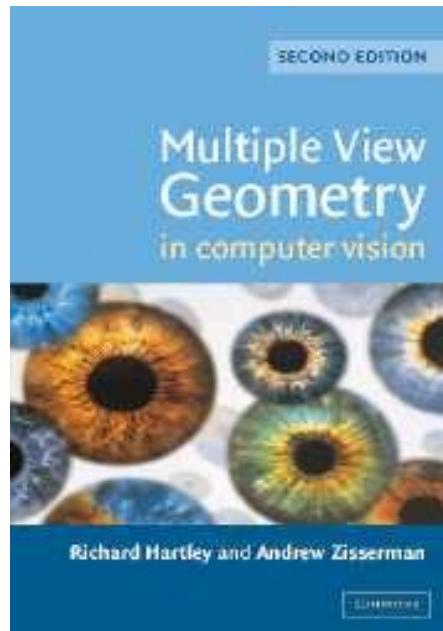
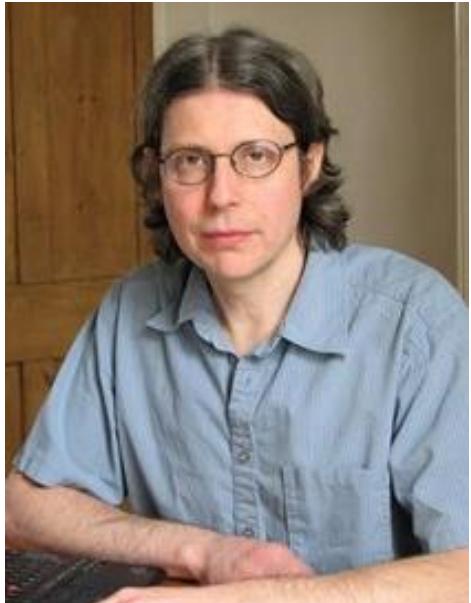
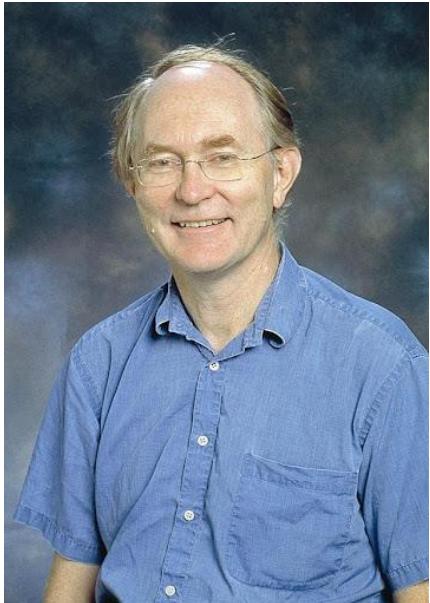


Other Recommended Textbooks

- 《Multiple View Geometry in Computer Vision》 Cambridge University Press

© 2004 Richard Hartley (ANU) and Andrew Zisserman (Oxford)

<http://users.rsise.anu.edu.au/~hartley/>





Top Journals and Conferences

- International Journal of Computer Vision (IJCV)
- IEEE Transaction on Pattern Analysis and Machine Intelligence (TPAMI)
- IEEE Computer Society International Conference on Computer Vision and Pattern Recognition (CVPR)
- IEEE International Conference on Computer Vision (ICCV)
- European Conference on Computer Vision (ECCV)
- ACM SIGGRAPH
- ACM SIGGRAPH Asia



Grading

- Three mini programming projects:
 - Homework1 (10%)
 - Homework2 (15%)
 - Homework3 (15%)
- Final: Team work, ≤ 3 , **Weight**: 1.0、0.9、0.8
 - Final Project (50%)
 - Deadline: At the very beginning of Spring Semester 2024 [Tentative: **March 8, 2024**]
 - Final Project Presentation (10%)
 - TBD



Submission & Late Policy

- Assignments are submitted via Learning in ZJU (学在浙大) .



- For a late submission, penalty of 20% score reduction per 12 hours is applied. If one submits an assignment within 12 hours after the deadline, the score becomes 80%. Within 24 hours, 60%, and so on.



Course Webpage

- <http://www.cad.zju.edu.cn/home/zhpcui/courses/compvis/>

Computer Vision Schedule Description

Computer Vision

Winter 2023, Friday

Description

Computer vision aims at modeling and understanding the real world from digital images, videos, or other visual inputs. Recently the computer vision techniques are applied to many areas including intelligent robotics, mixed reality, autonomous driving, and so on. The goal of this course is to provide students with a good understanding of these techniques. The basic concepts and techniques will be studied in-depth, and practical algorithms will be explored through the exercises.

Prerequisites:

- Linear Algebra
- Probability
- Programming with Matlab or C++

Schedule

Date	Lecture	Instructor	Notes
2023/11/17	Introduction & Camera	Zhaopeng Cui	

Homework

Topic	Content	Submission Deadline
-------	---------	---------------------

Slides Credit



- The slides are mainly from Ping Tan and Gang Pan, who further adopted many slides from Yasu Frukawa (SFU), Ioannia Gkioulekas (CMU), Kris Kitani (CMU), Fredo Durand (MIT), Kayvon Fatahalian (Stanford), Michael Brown (York) and Jiaya Jia (CUHK).

Who is Your Lecturer?



- Zhaopeng Cui (崔兆鹏)
- B.S. and MSc. from Xidian University, 2009 and 2012
- Ph.D. from Simon Fraser University, 2017
- Senior Researcher at ETH Zurich, 2017 – 2020
- Research Professor (百人计划研究员) at Zhejiang University, 2020.11 – present
- Research: computer vision, graphics, and robotics

Zhaopeng Cui Home Publications Contact



Zhaopeng Cui

College of Computer Science and Technology
Zhejiang University
Email: zhpcui AT zju.edu.cn
zhpcui AT gmail.com

[Google Scholar](#)

- Research Interest
- Structure-from-Motion
 - Multi-View Stereo
 - SLAM
 - 3D Deep Learning
 - Video and Image Processing

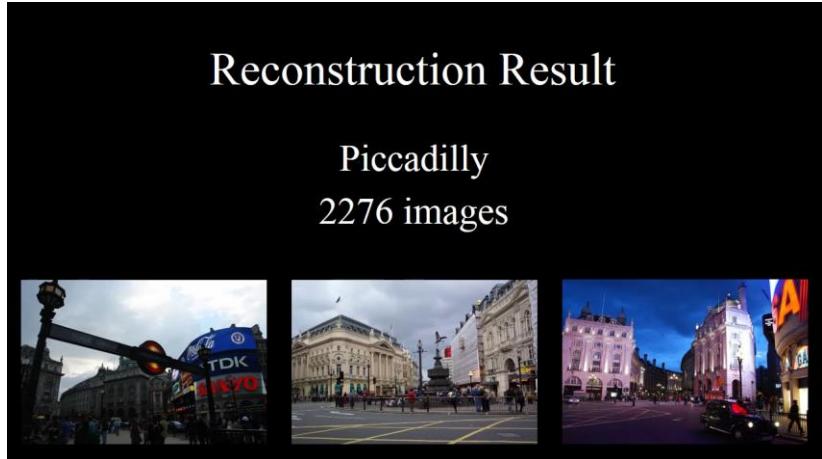
I am a Research Professor in the [College of Computer Science](#) at [Zhejiang University](#), affiliated with the [State Key Laboratory of CAD&CG](#). My research interests include computer vision, computer graphics, robotics, and machine learning.

Before joining Zhejiang University, I was a senior researcher working with Prof. Marc Pollefeys in the [Computer Vision and Geometry Group](#) at [ETH Zurich](#). I received my Ph.D. degree in computer science under the supervision of Prof. Ping Tan at [Simon Fraser University](#). Before that, I was a Ph.D. candidate supervised by Prof. Ping Tan at [National University of Singapore](#). I obtained my Bachelor degree and Master degree at [Xidian University](#) in 2009 and 2012 respectively.

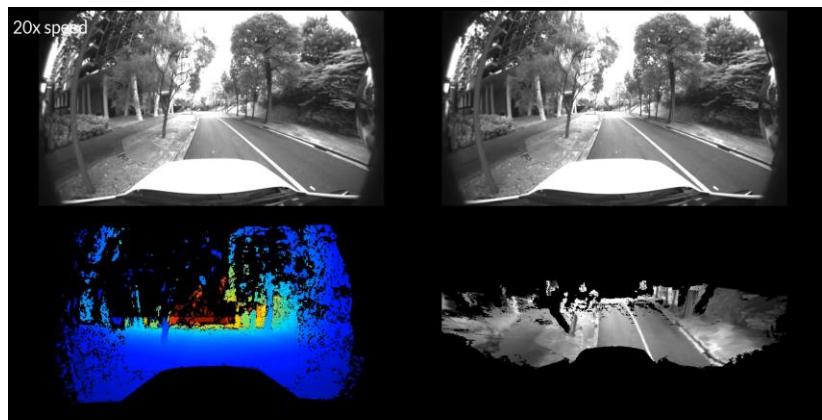
Who is Your Lecturer?



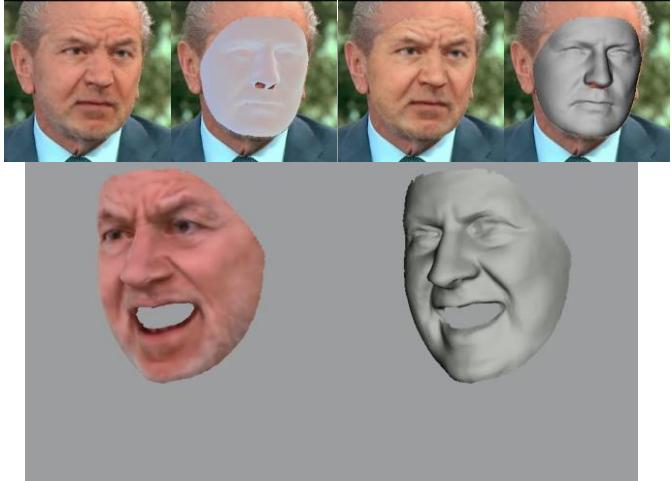
- Some of my past research projects:



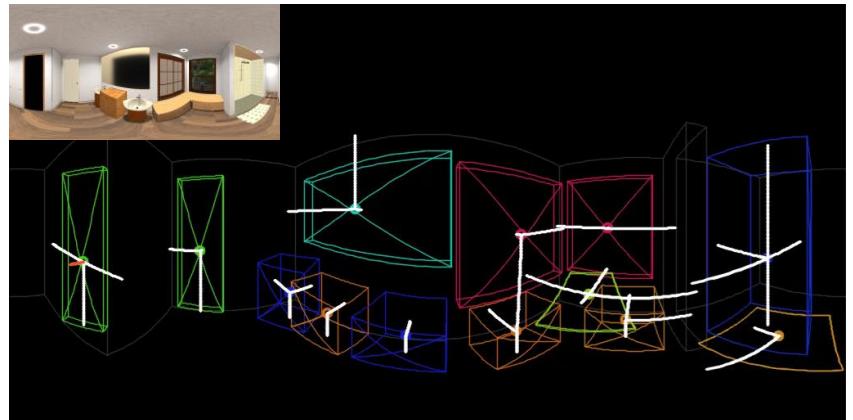
3D Reconstruction



3D Perception for Autonomous Driving



Human Modeling



3D Scene Understanding



Who is Your Lecturer?

- Some of my past research projects:

More examples of
Editable Scene Rendering



We can insert **virtual objects** into the **real-world** and tour in
the edited scene.

3D Perception with Neural Radiance Fields



Who are Your TAs?

- Xiyue Guo
- PhD student at ZJU
- 12121044@zju.edu.cn
- 紫金港校区东1A509
- Jiarui Hu
- PhD student at ZJU
- 12221071@zju.edu.cn
- 紫金港校区东1A509
- Mao Mao
- PhD student at ZJU
- 12221070@zju.edu.cn
- 蒙民伟楼415



Questions?

