What is computer vision?

“To know what is where by looking.”

— Aristotle (300 BC)

To know what is where
by looking.

by looking.

The study of recovering useful properties of the world by looking.

The study of recovering useful properties of the world

The study of recovering useful properties of the world from one or more images

The study of recovering useful properties of the world from one or more images with an algorithmic level of specification.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group — July 7, 1966
Vision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

What makes computer vision HARD?

210	210	210	210	210	210	210	210	210	210	210	210	210	210	210	210	210	2
210	210	210	210	210	210	210	210	210	210	210	210	210	210	210	210	210	2
211	211	211	211	211	211	211	211	211	211	211	211	211	211	211	211	211	2
211	211	211	211	211	211	211	211	211	211	211	211	211	211	211	211	211	2
212	212	212	212	212	212	212	212	212	212	212	212	212	212	212	212	212	2
212	212	212	212	212	212	212	212	212	212	212	212	212	212	212	212	212	2

211	212	212	213	214	214	211	211	211	212	212	211	212	212	213	214	214	2
211	212	212	213	214	214	213	213	212	211	211	211	212	212	213	214	214	2
211	212	212	213	213	213	215	214	212	210	209	211	212	212	213	213	213	2
212	212	212	212	212	212	214	212	210	208	208	212	212	212	212	212	212	2
212	212	212	212	212	212	210	209	209	208	209	212	212	212	212	212	212	2
213	212	212	211	211	211	209	209	209	210	210	213	212	212	211	211	211	2
213	212	212	211	210	210	210	211	212	213	211	213	212	212	211	210	210	2

Images are a complex function of several variables

Images are a complex function of several variables

scene geometry

Images are a complex function of several variables

scene geometry

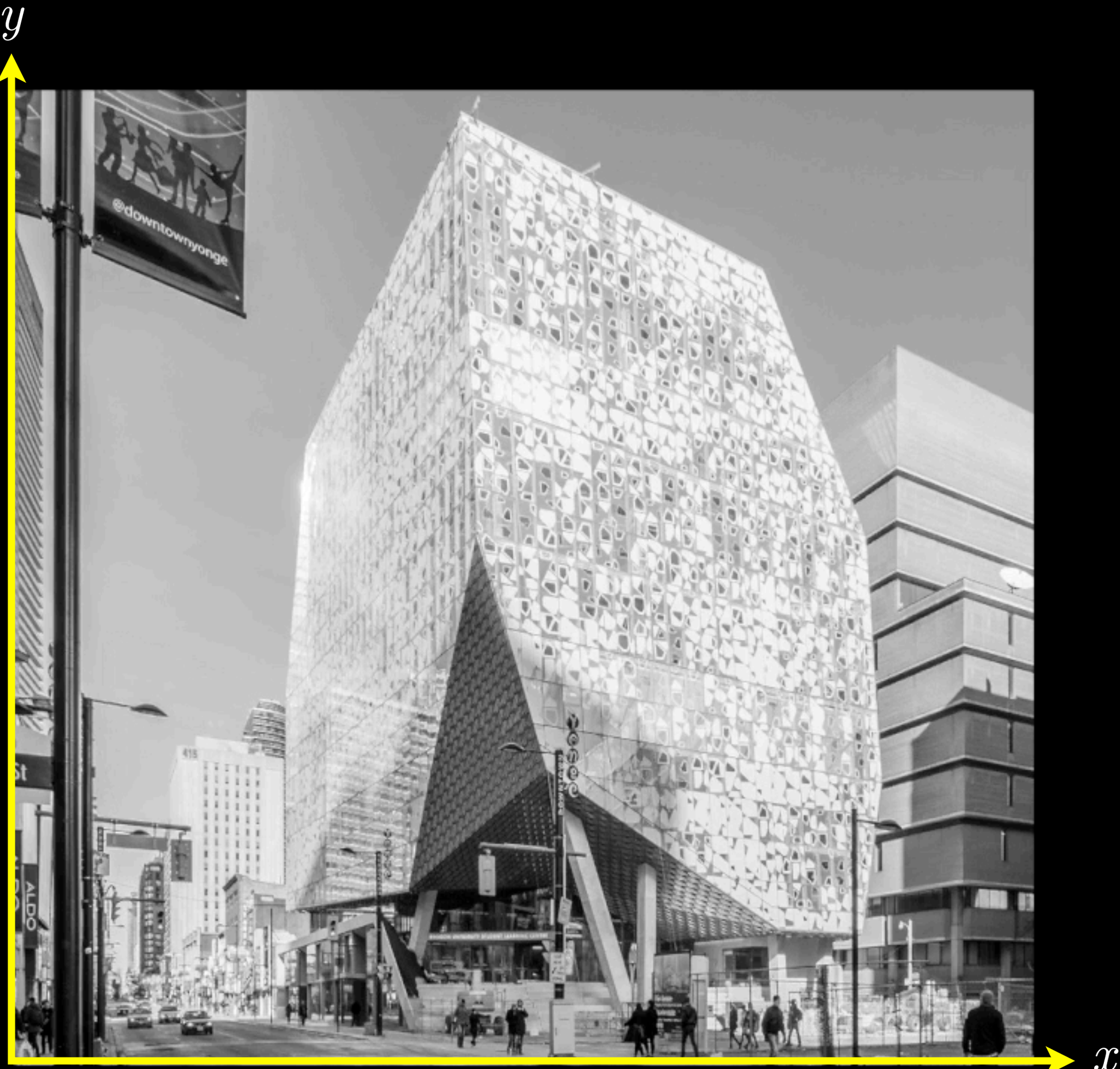
scene surface microstructure

Images are a complex function of several variables

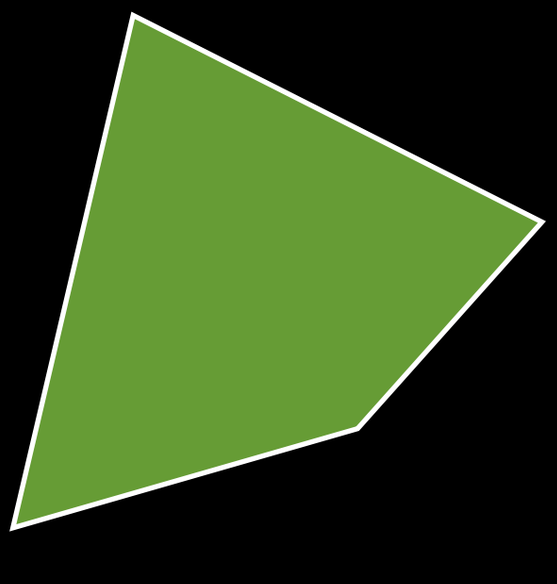
scene geometry

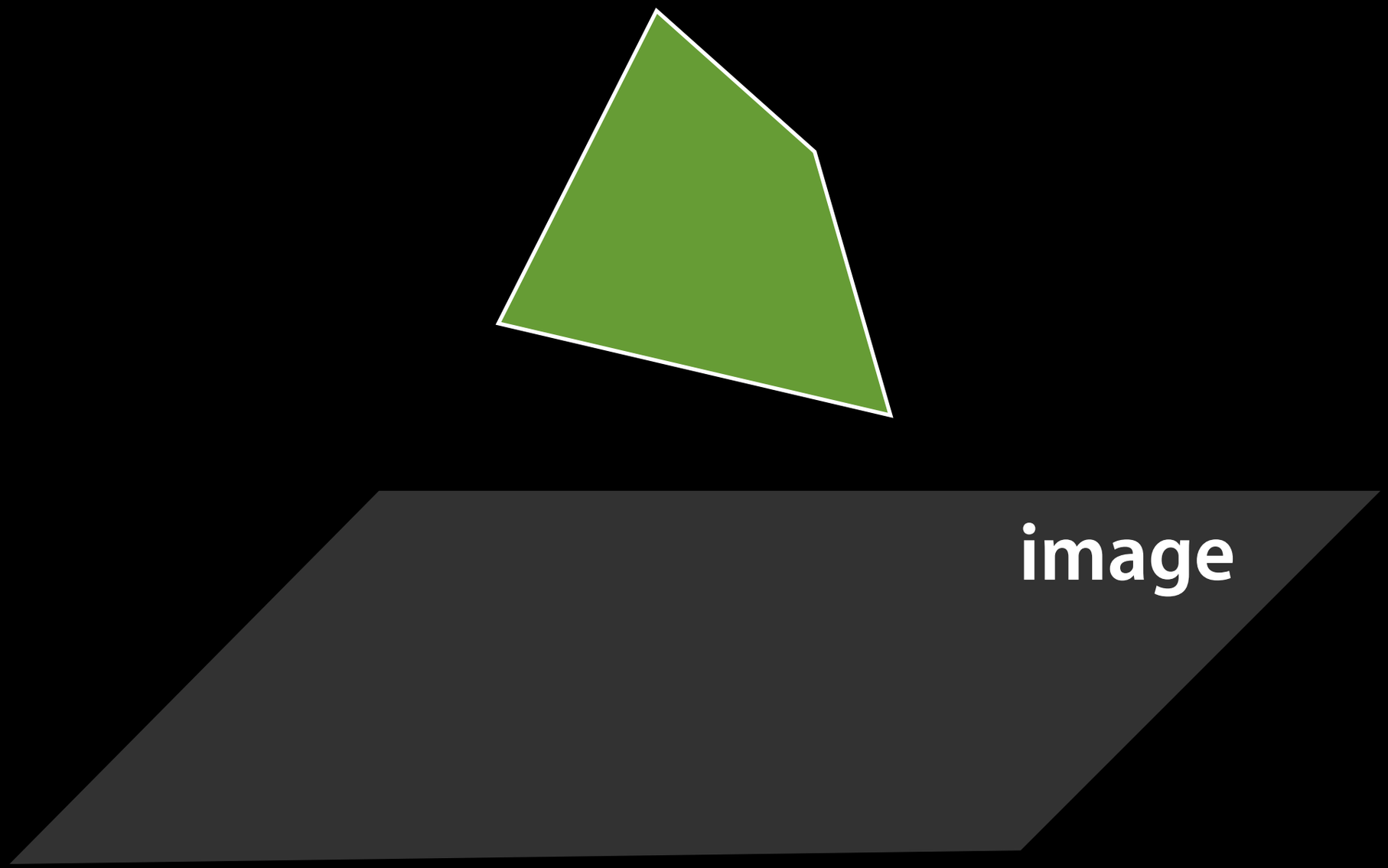
scene surface microstructure

scene illumination

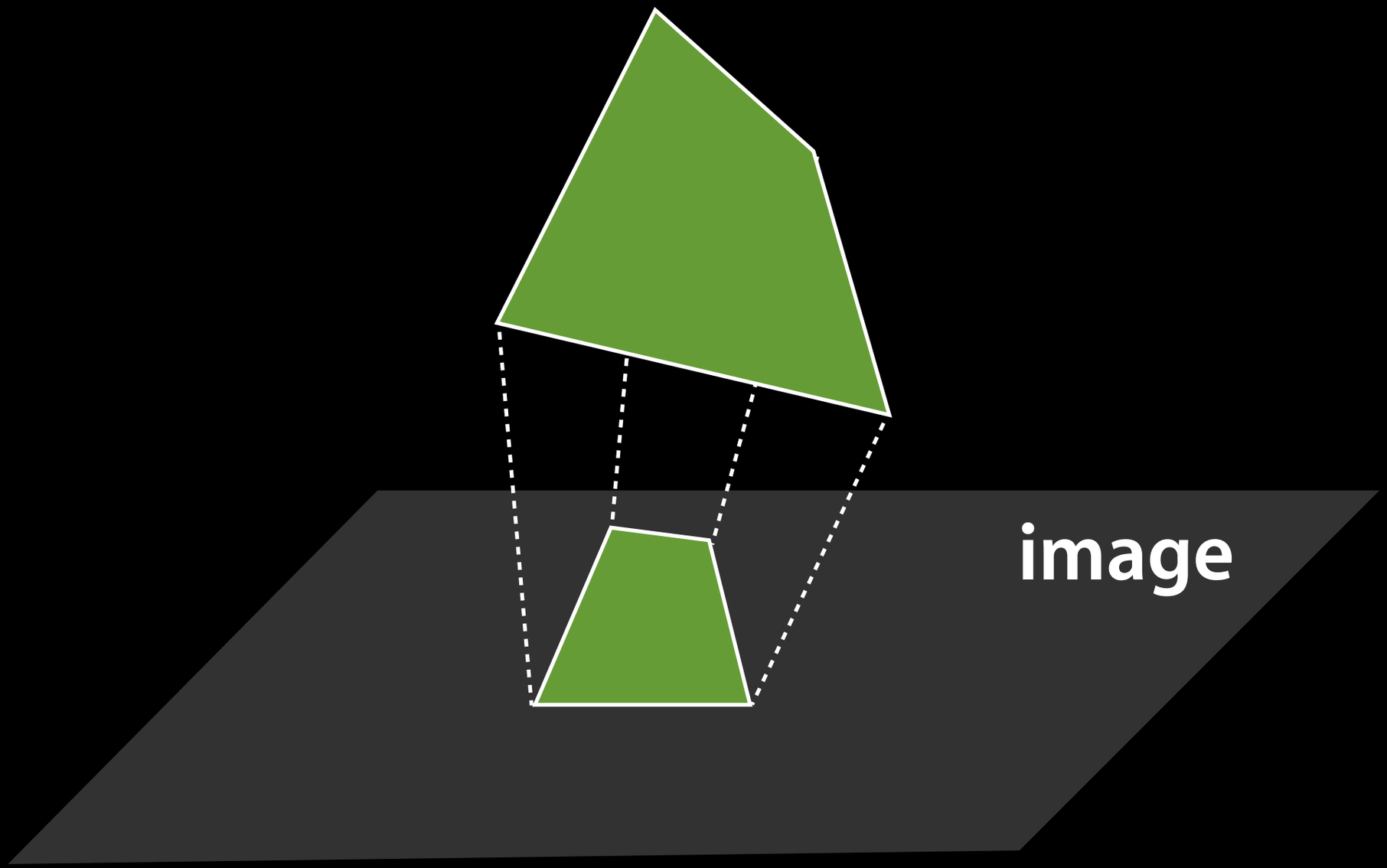

y
@downtownyonge
ALDO
x


@downtownyonge
ALDO

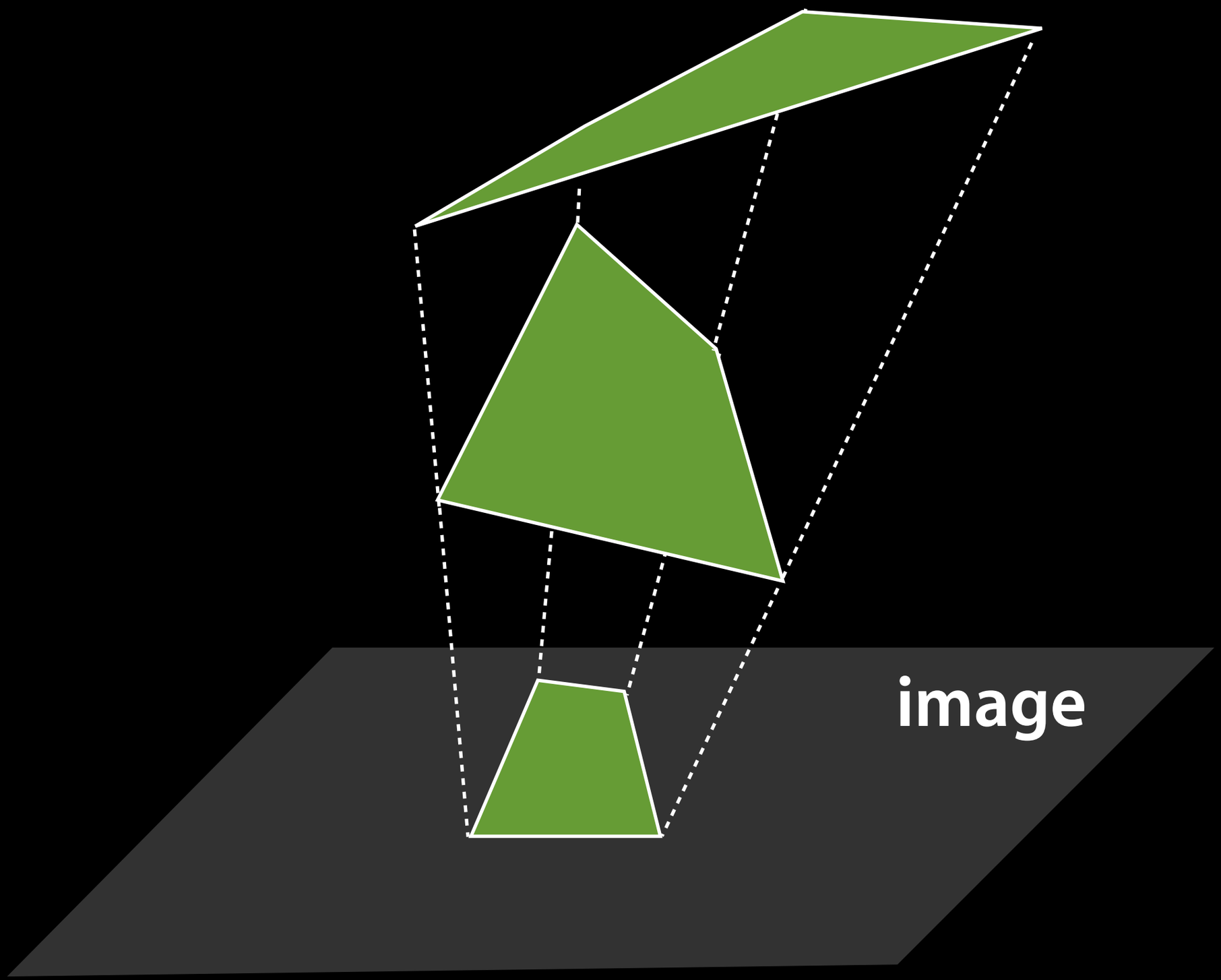




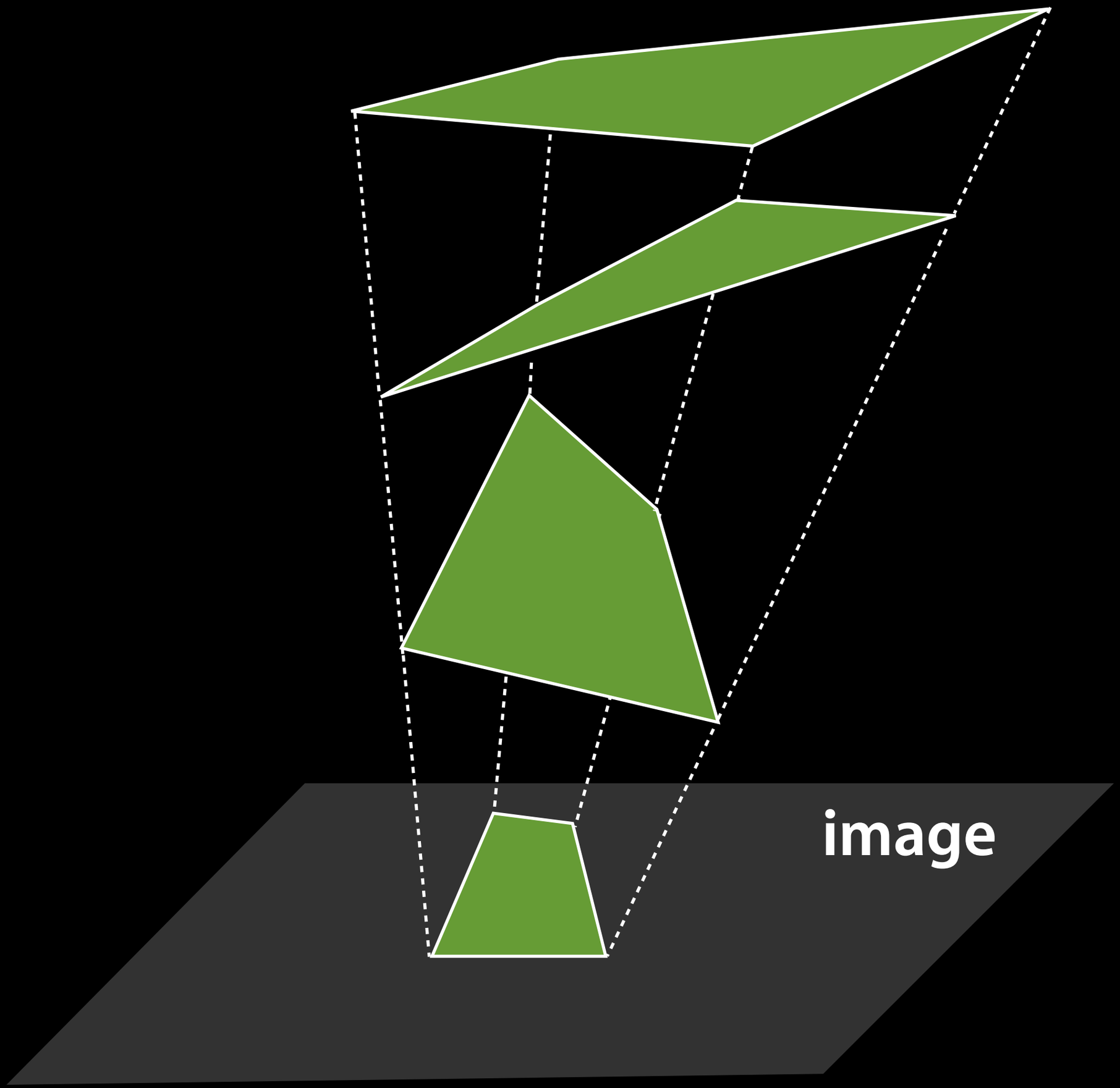
image



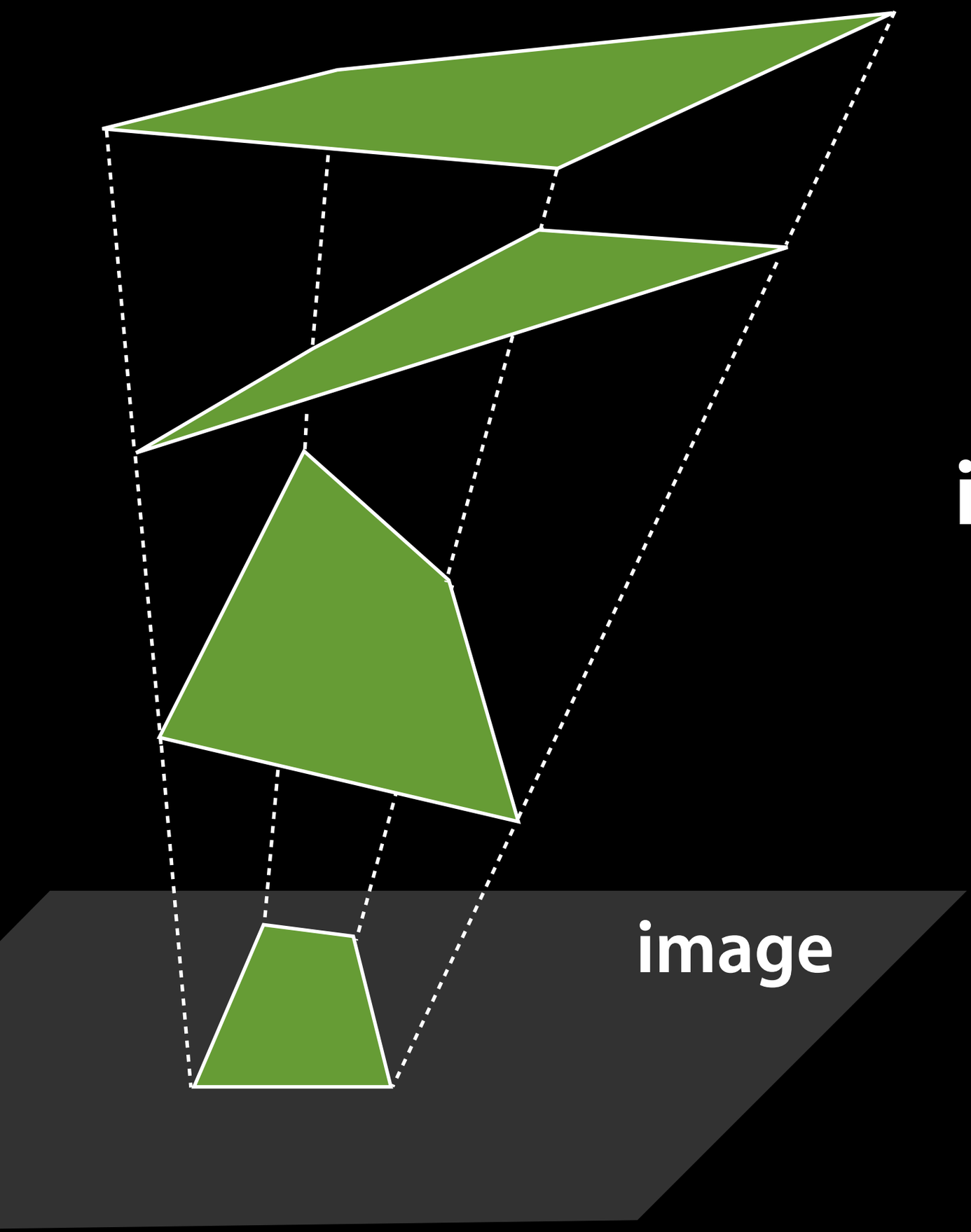
image



image



image



infinite number of possible shapes

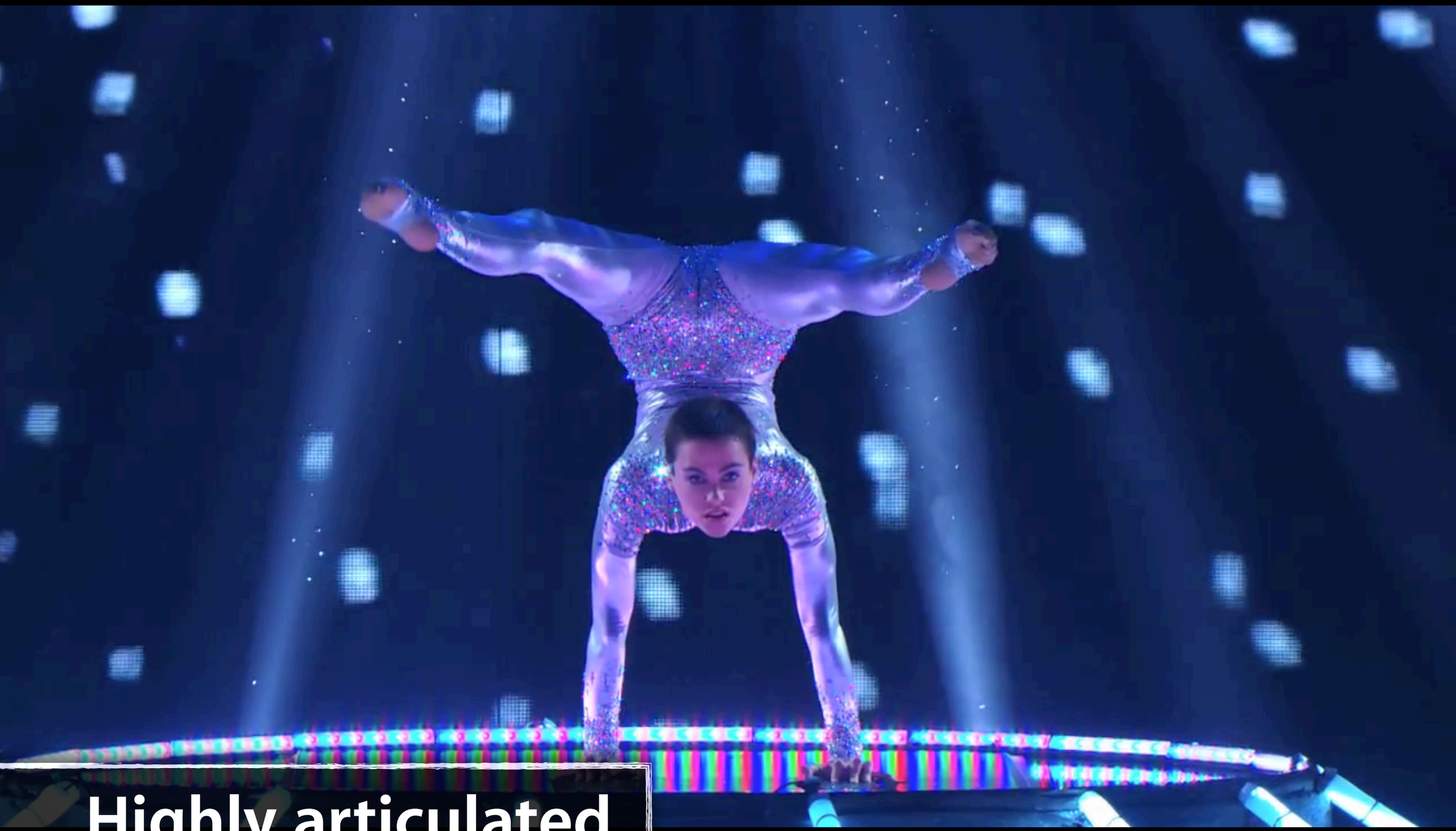
Objects have highly variable appearance

Viewpoint differences

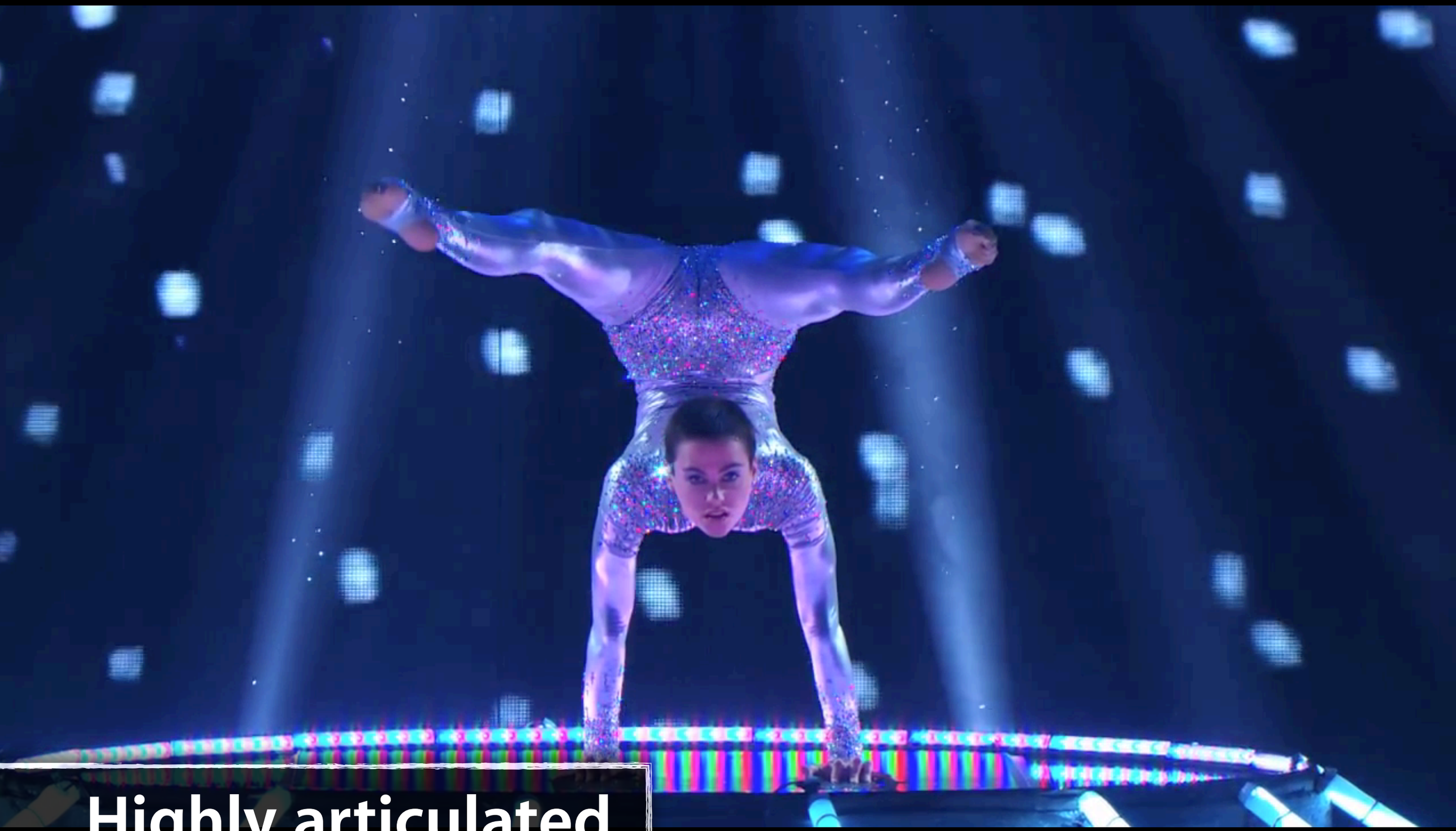
Varying SCALES

Varying instances

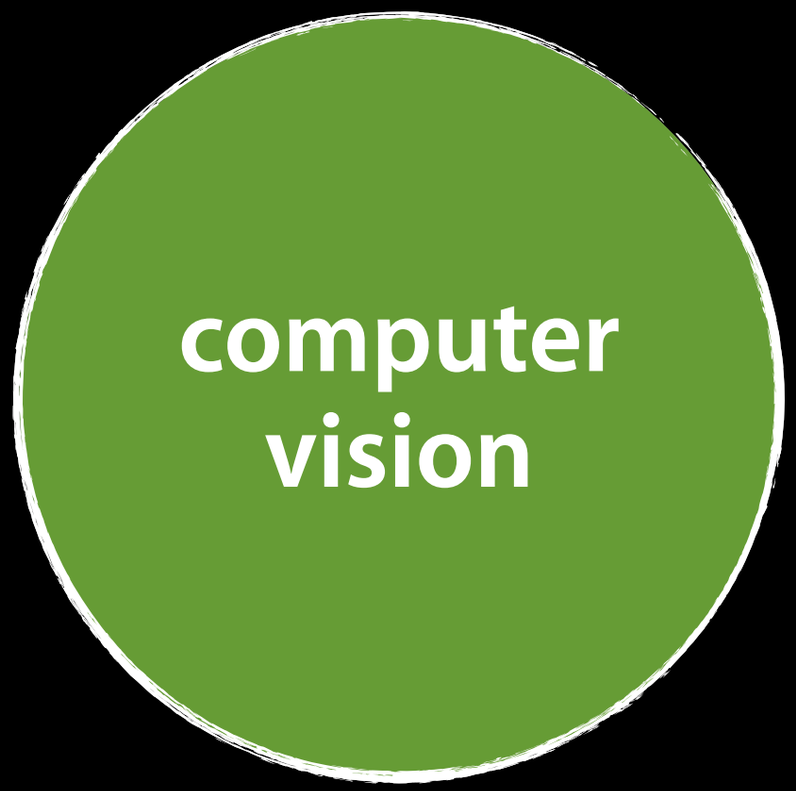
Occlusion



Highly articulated



Highly articulated


computer
vision

computer graphics

computer vision

computer graphics

computer vision

image processing

computer graphics

computer vision

image processing

computational photography

HDR

Exposure value (EV): -7

EV: 0

EV: +5

Final Result

Research areas

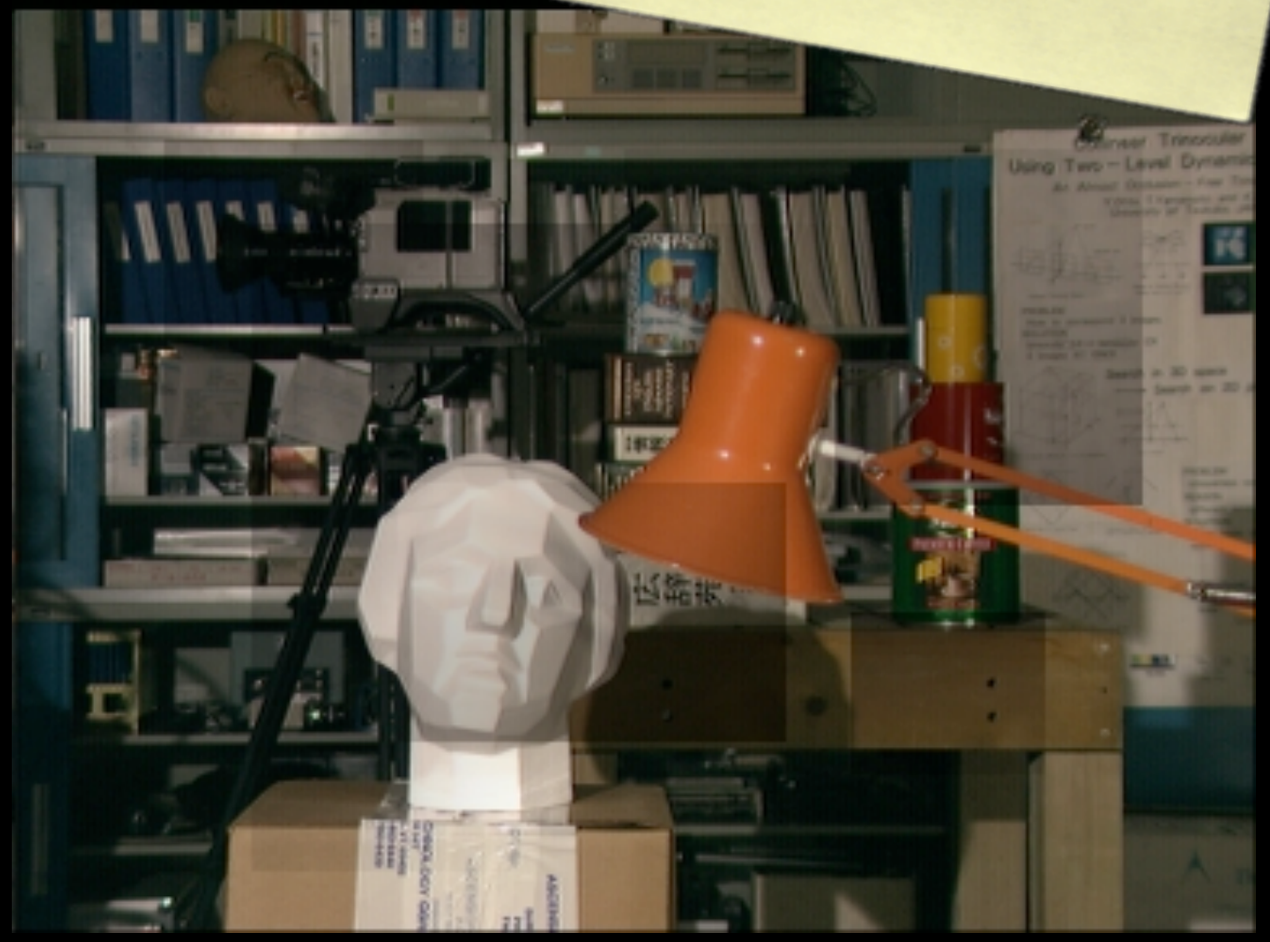
Feature
Detection


Feature
Detection


Shape-from-X

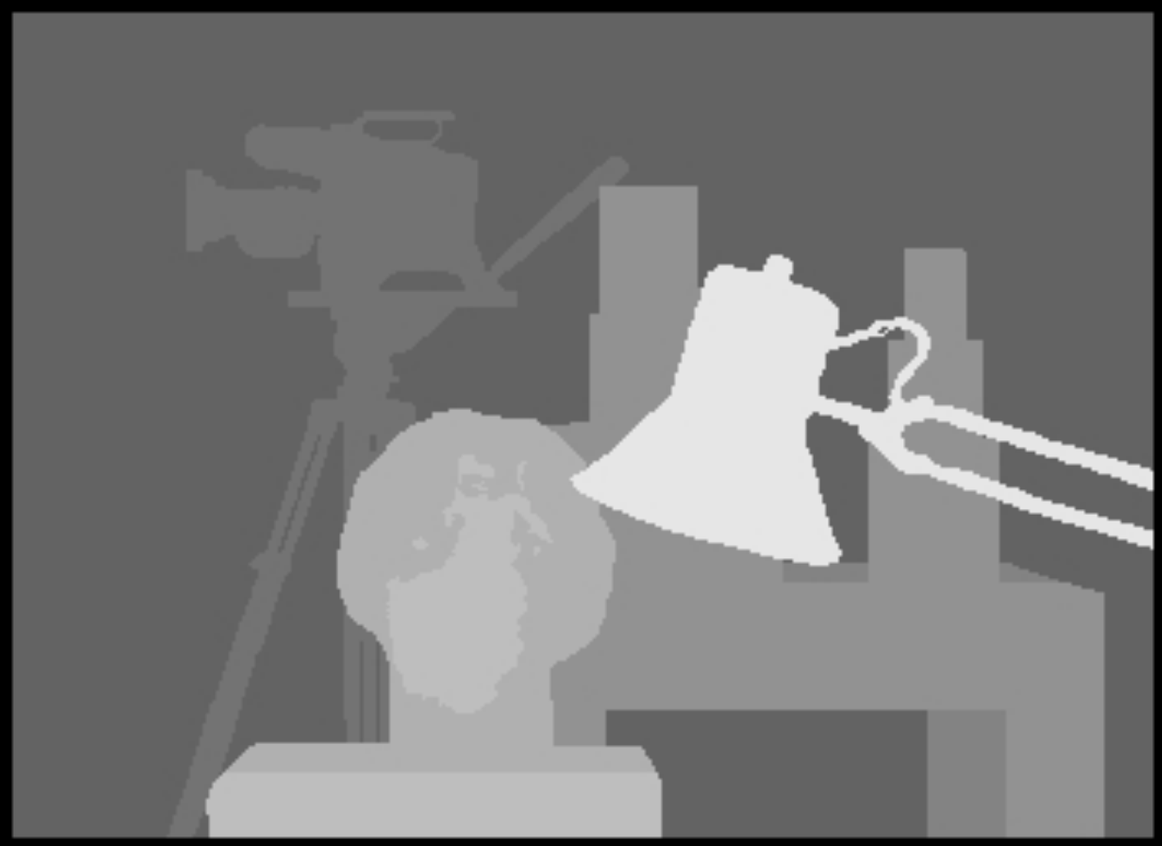
Stereopsis

left view



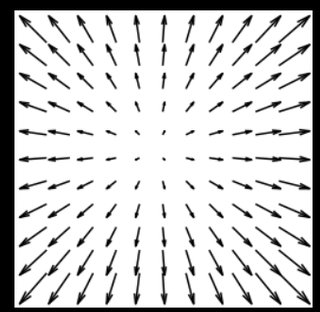
right view

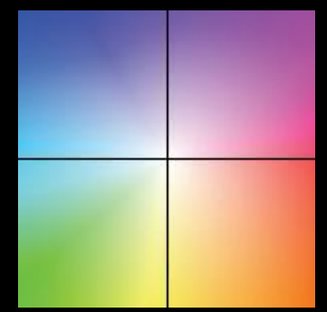
Stereopsis



depth

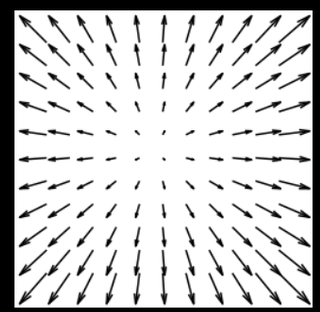
Motion Analysis

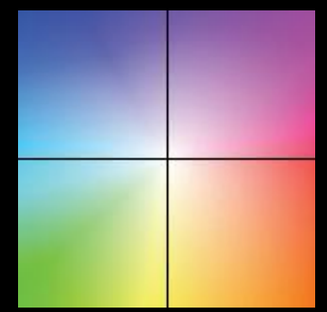




www.gpu4vision.org

Motion Analysis





www.gpu4vision.org

Tracking

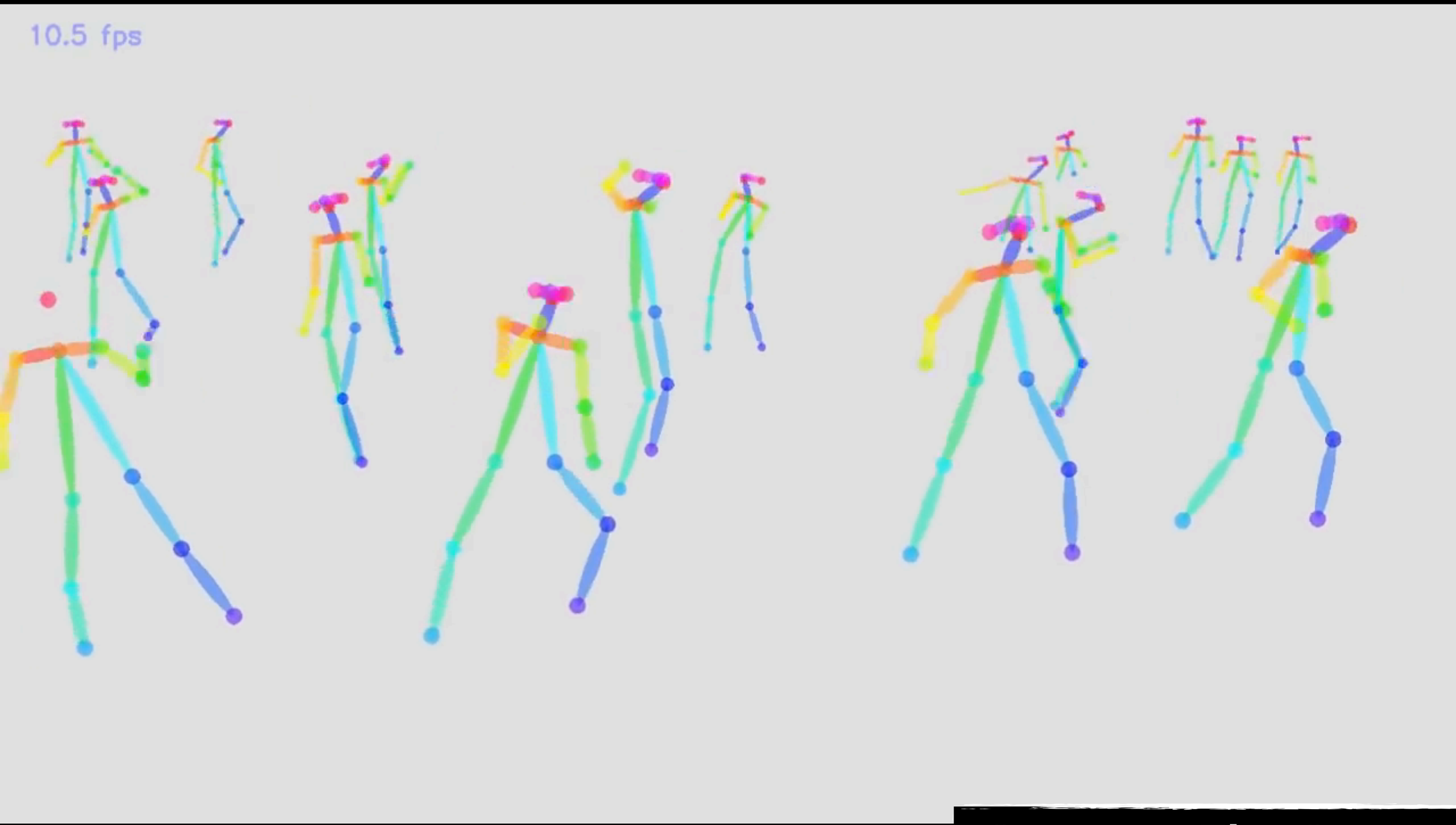
Courtesy: Anton Andriyenko

Tracking

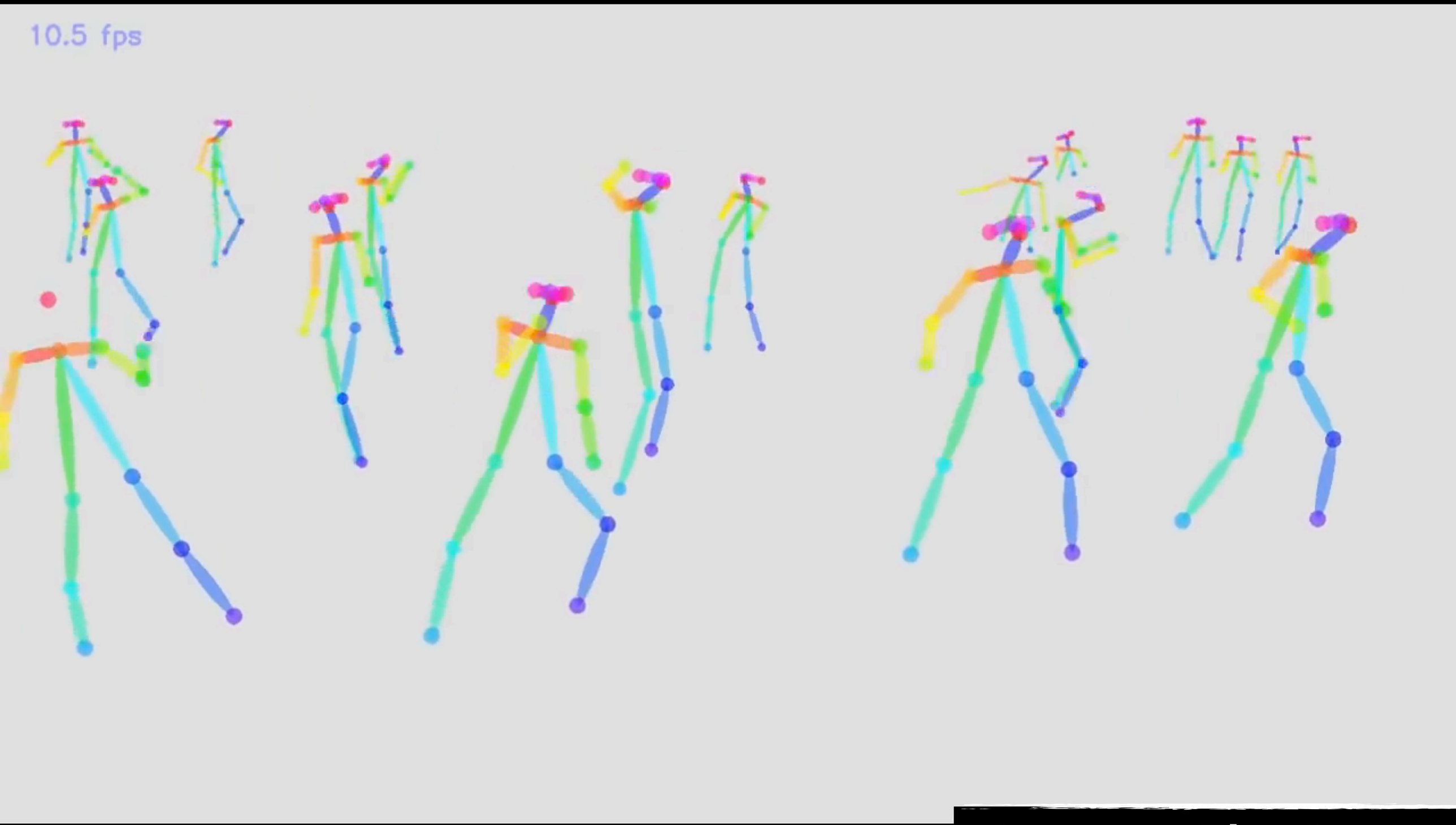
Courtesy: Anton Andriyenko


15 msec
Pitch : 2.9
Yaw : 10.6
Roll : 4.5
Tracking
Courtesy: Akshay Asthana


15 msec
Pitch : 2.9
Yaw : 10.6
Roll : 4.5
Tracking
Courtesy: Akshay Asthana

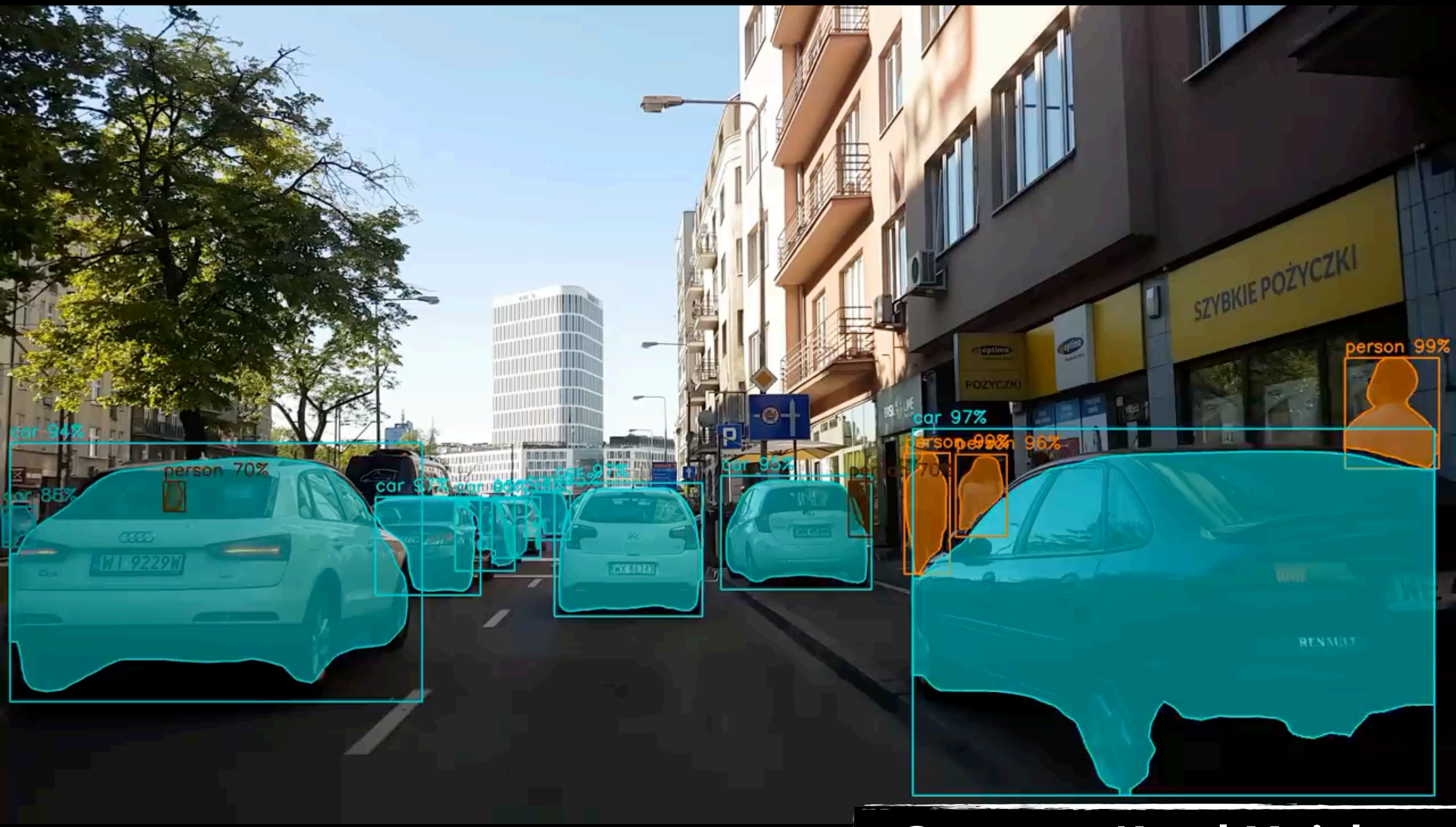


Courtesy: Zhe Cao

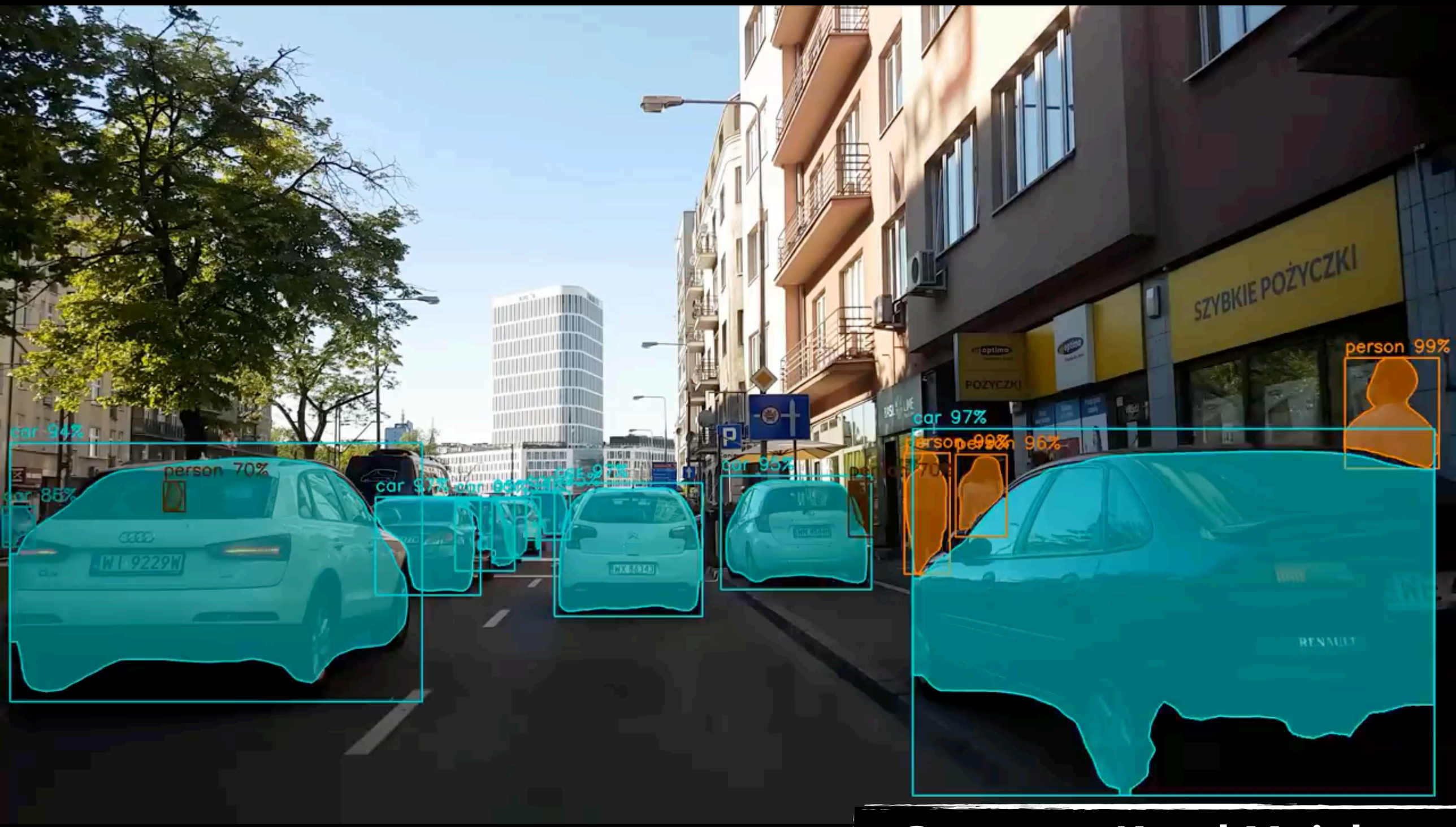


Courtesy: Zhe Cao

IN CS, IT CAN BE HARD TO EXPLAIN THE DIFFERENCE BETWEEN THE EASY AND THE VIRTUALLY IMPOSSIBLE.

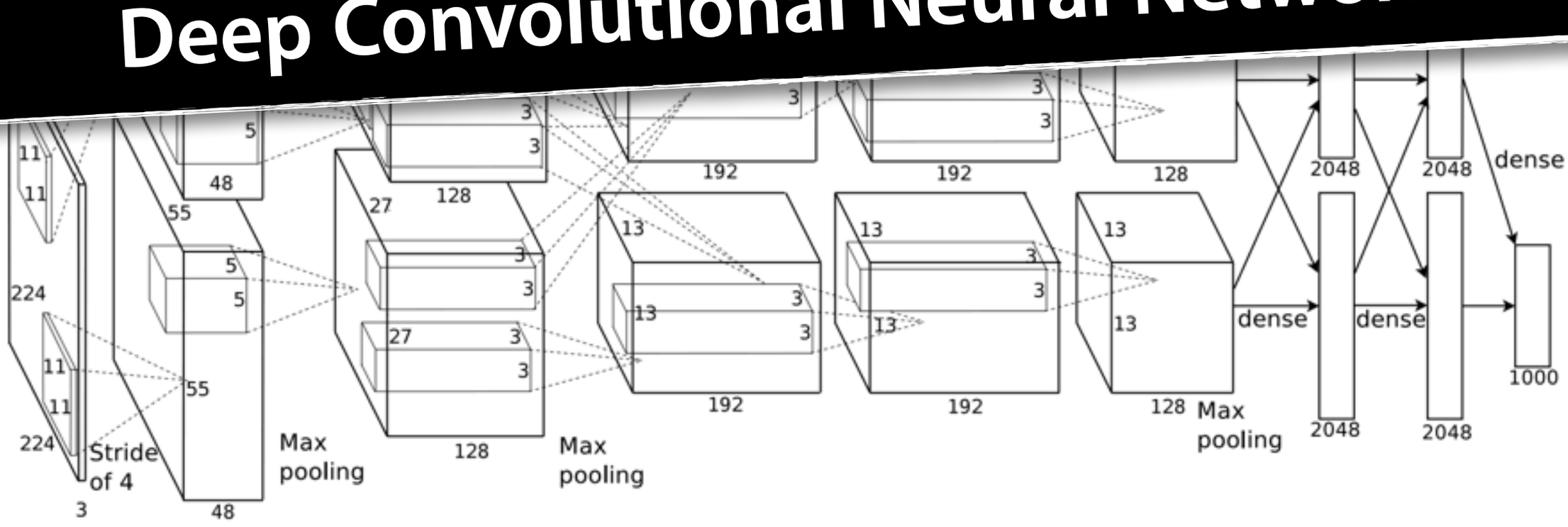


Courtesy: Karol Majek



Courtesy: Karol Majek

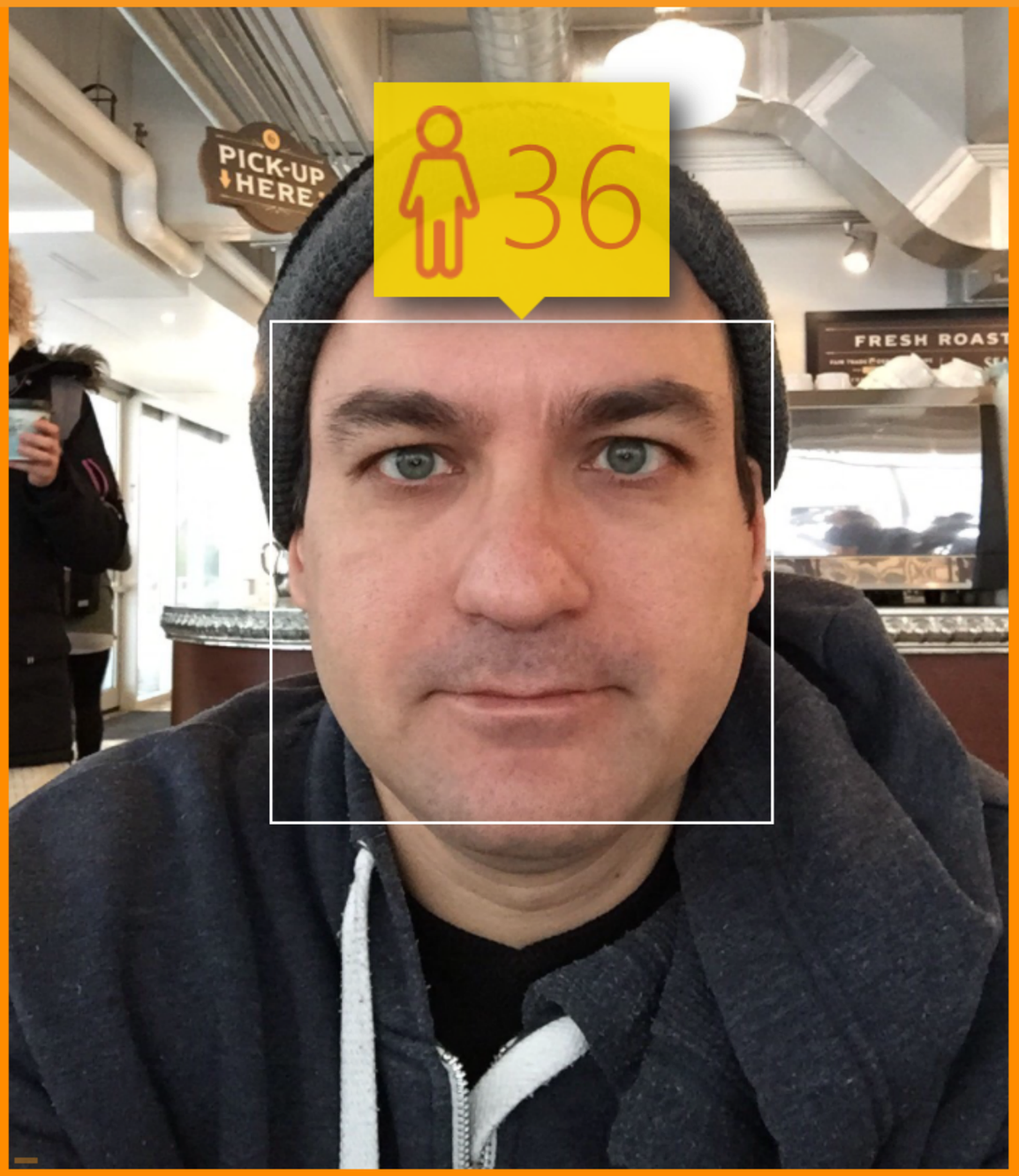
Deep Convolutional Neural Network



Applications

Face Detection

How Old Do I Look?



Sorry if we didn't quite get it right - we are still improving this feature..

Get a human opinion


9:41 AM
100%
IntraFace
Sad
Disgust
Neutral
Surprise
Happy


9:41 AM
100%
IntraFace
Sad
Disgust
Neutral
Surprise
Happy

Face Swap Live

Face Swap Live

Image Style Transfer Using Convolutional Neural Networks

Leon A. Gatys
Centre for Integrative Neuroscience, University of Tübingen, Germany
Bernstein Center for Computational Neuroscience, Tübingen, Germany
Graduate School of Neural Information Processing, University of Tübingen, Germany
leon.gatys@bethgelab.org

Alexander S. Ecker
Centre for Integrative Neuroscience, University of Tübingen, Germany
Bernstein Center for Computational Neuroscience, Tübingen, Germany
Max Planck Institute for Biological Cybernetics, Tübingen, Germany
Baylor College of Medicine, Houston, TX, USA

Matthias Bethge
Centre for Integrative Neuroscience, University of Tübingen, Germany
Bernstein Center for Computational Neuroscience, Tübingen, Germany
Max Planck Institute for Biological Cybernetics, Tübingen, Germany

Abstract

Rendering the semantic content of an image in different styles is a difficult image processing task. Arguably, a major limiting factor for previous approaches has been the lack of image representations that explicitly represent semantic information and, thus, allow to separate image content from style. Here we use image representations derived from

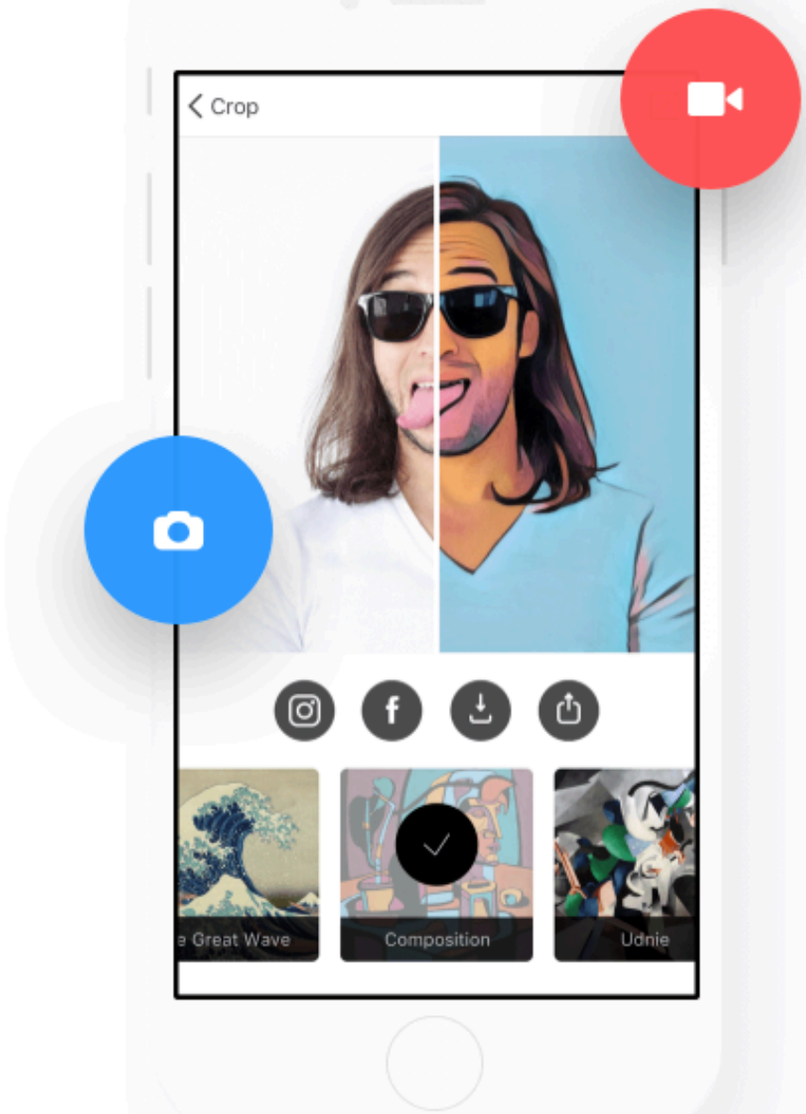
there exist a large range of powerful non-parametric algorithms that can synthesise photorealistic natural textures by resampling the pixels of a given source texture [7, 30, 8, 20]. Most previous texture transfer algorithms rely on these non-parametric methods for texture synthesis while using different ways to preserve the structure of the target image. For instance, Efros and Freeman introduce a correspondence map that includes features of the target image such as im-

CVPR 2016

prisma-ai.com

PRISMA

Follow on Instagram



Turn Memories into Art

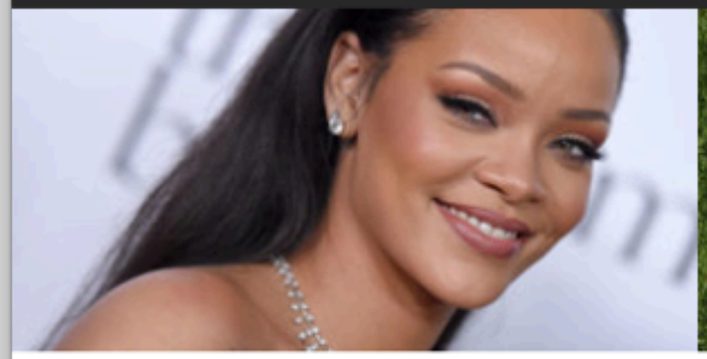
Using Artificial Intelligence

Prisma transforms your photos and videos into works of art using the styles of famous artists: Van Gogh, Picasso, Levitan, as well as world famous ornaments and patterns. A unique combination of neural networks and artificial intelligence helps you turn memorable moments into timeless art pieces.

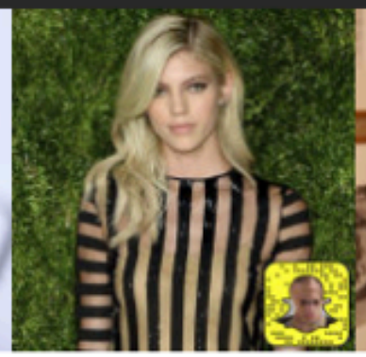
Tweet ··· Like 142K

Available on the App Store

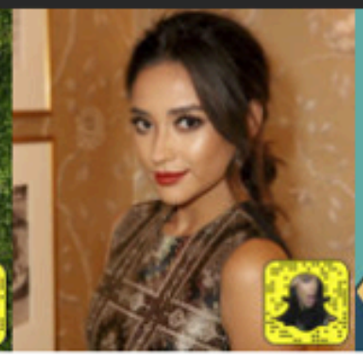
GET IT ON Google Play



Rihanna Is Still in the Studio Working on Her New...



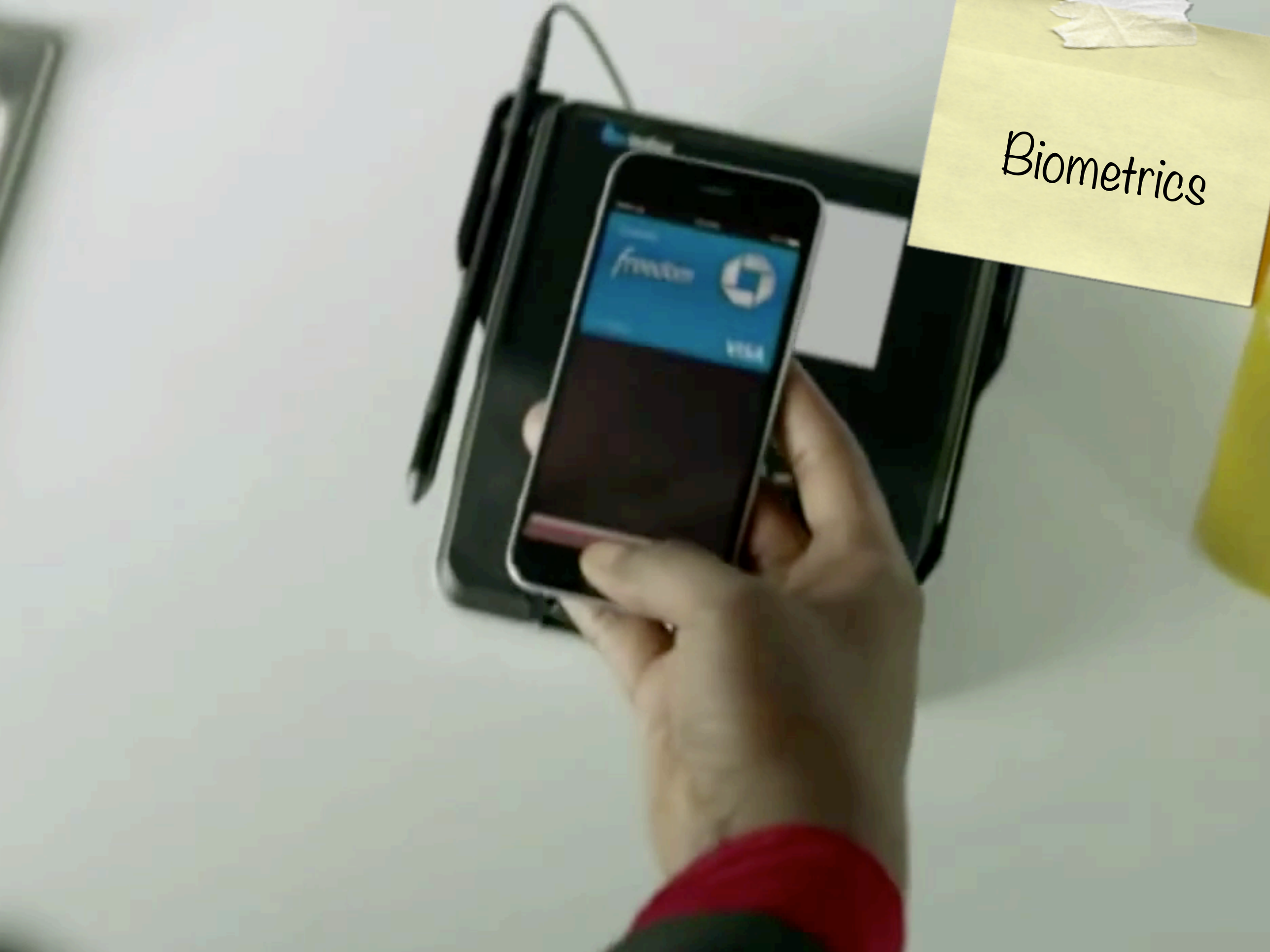
How to Kill It On Snapchat

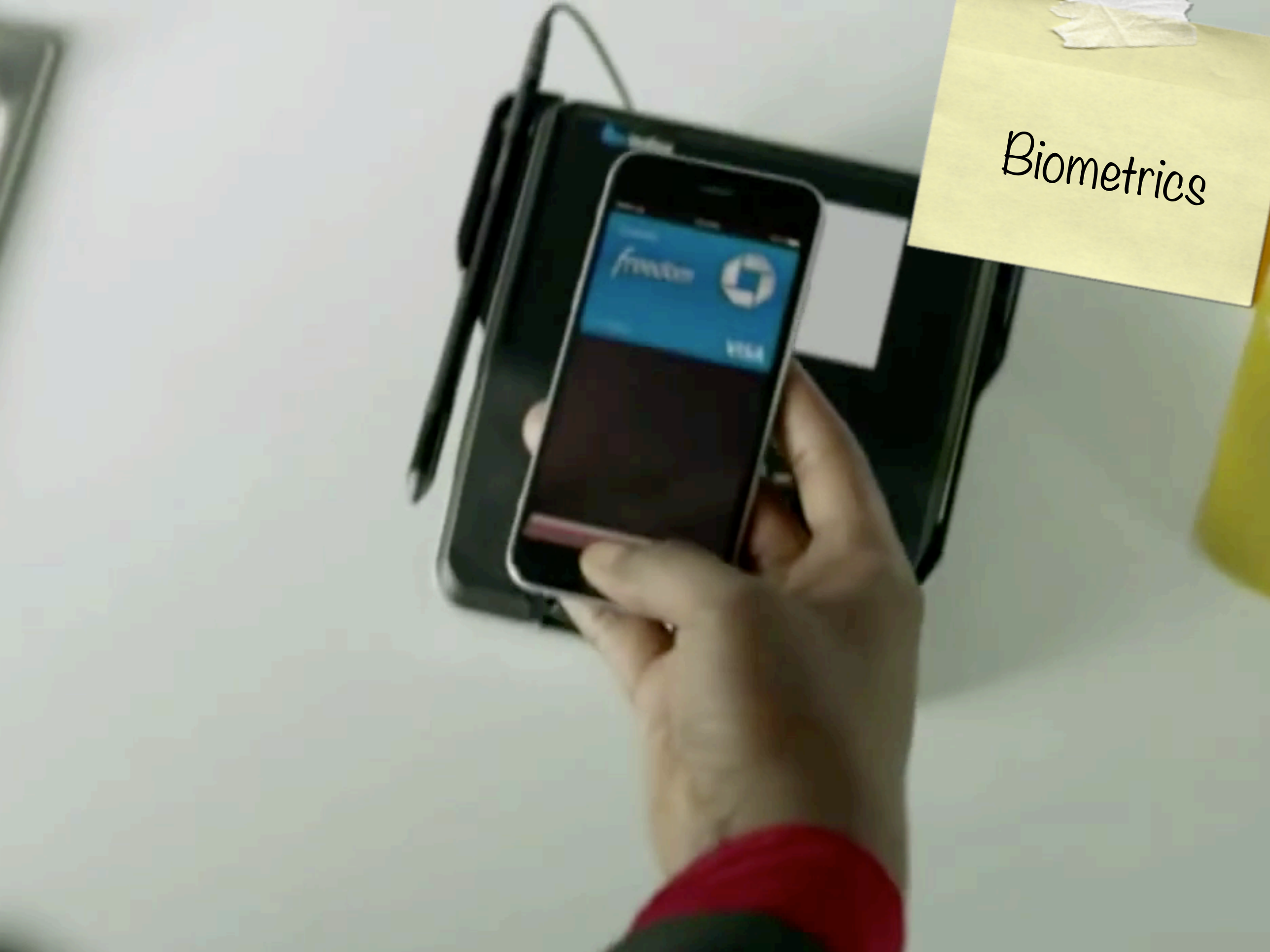


Aquarius: Your 2016 Year Ahead

NEW ALGORITHM FOR INSTAGRAM WILL TELL YOU HOW TO DRESS BETTER

For maximum likes.

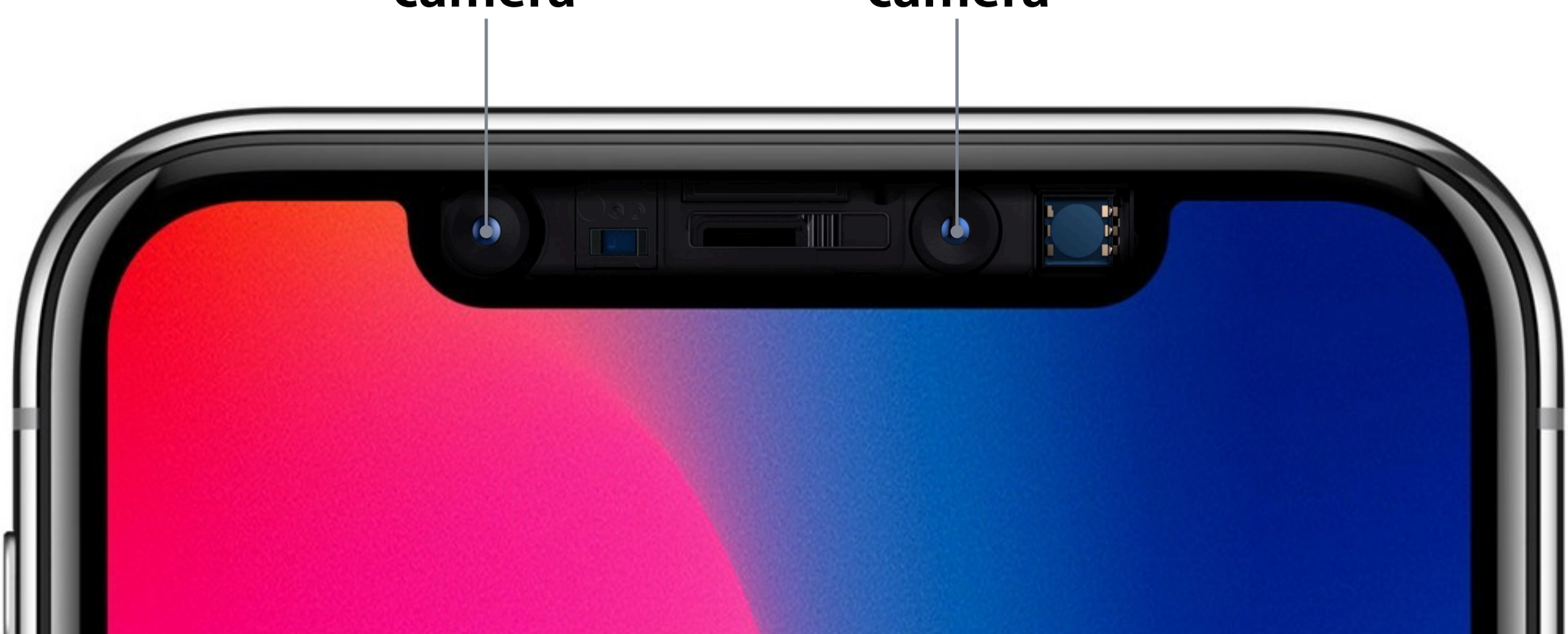

Biometrics

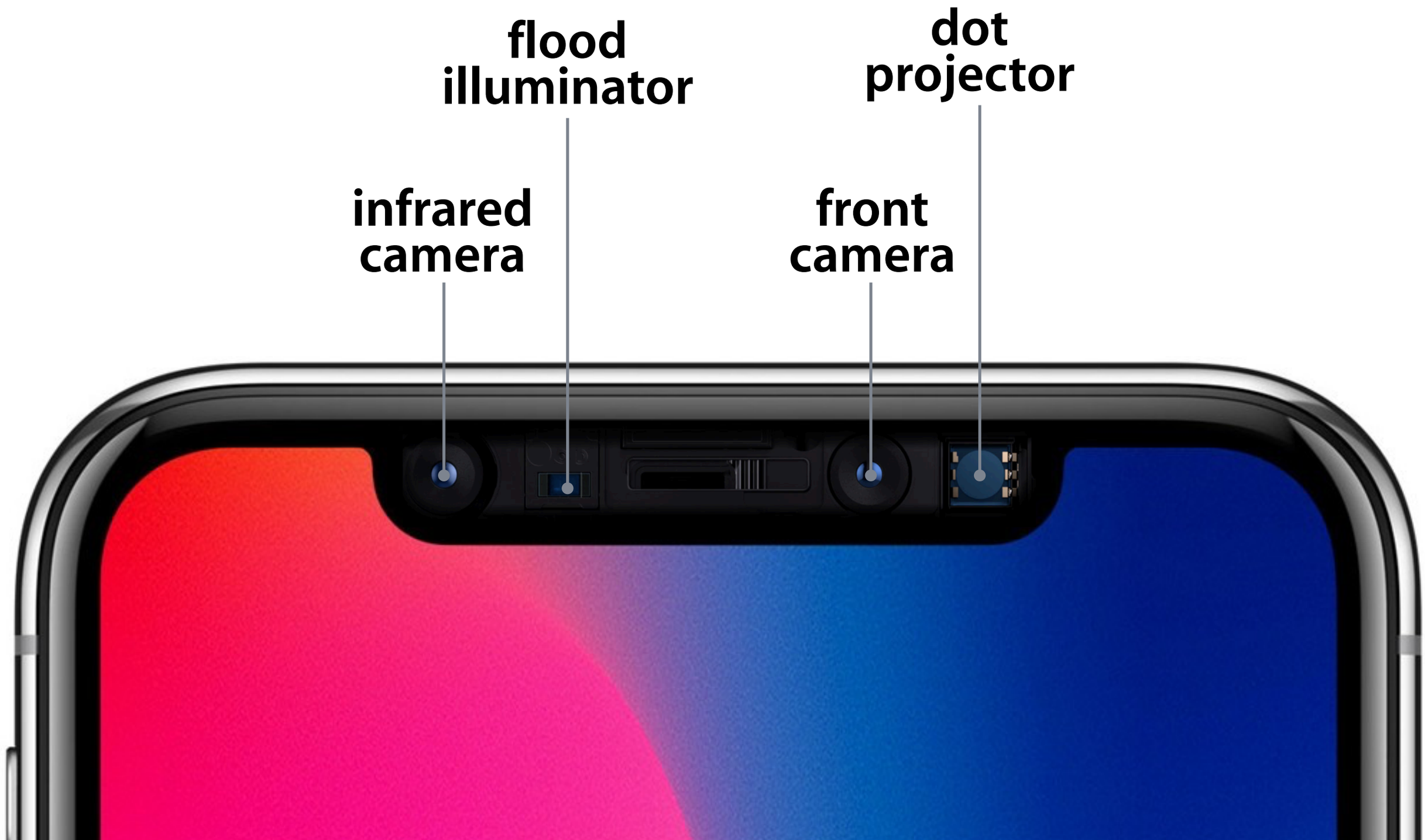

Biometrics

iPhone X

infrared camera

front camera




flood
illuminator
dot
projector
infrared
camera
front
camera

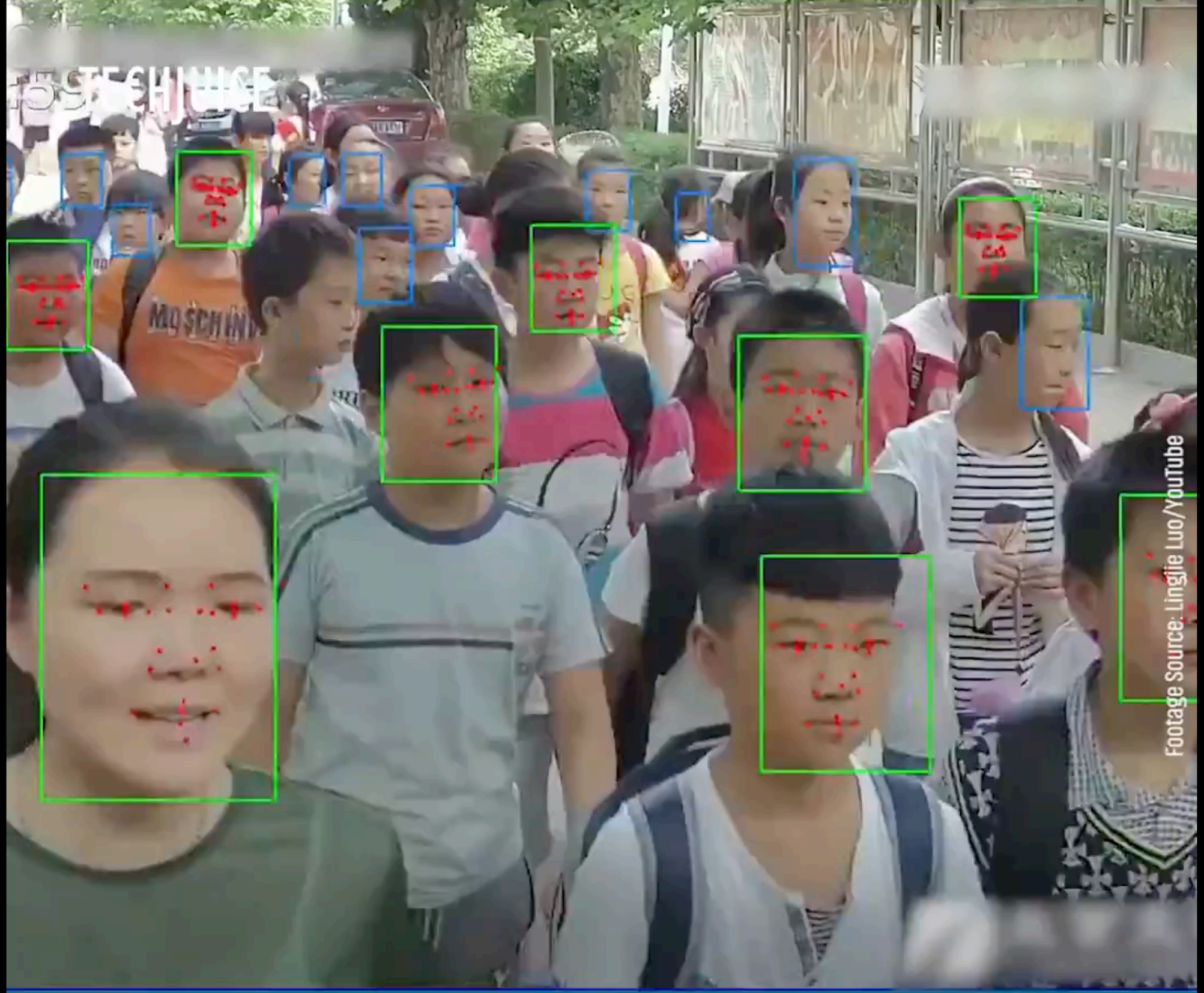
Source: Minority Report (2002)

BUSINESS INSIDER

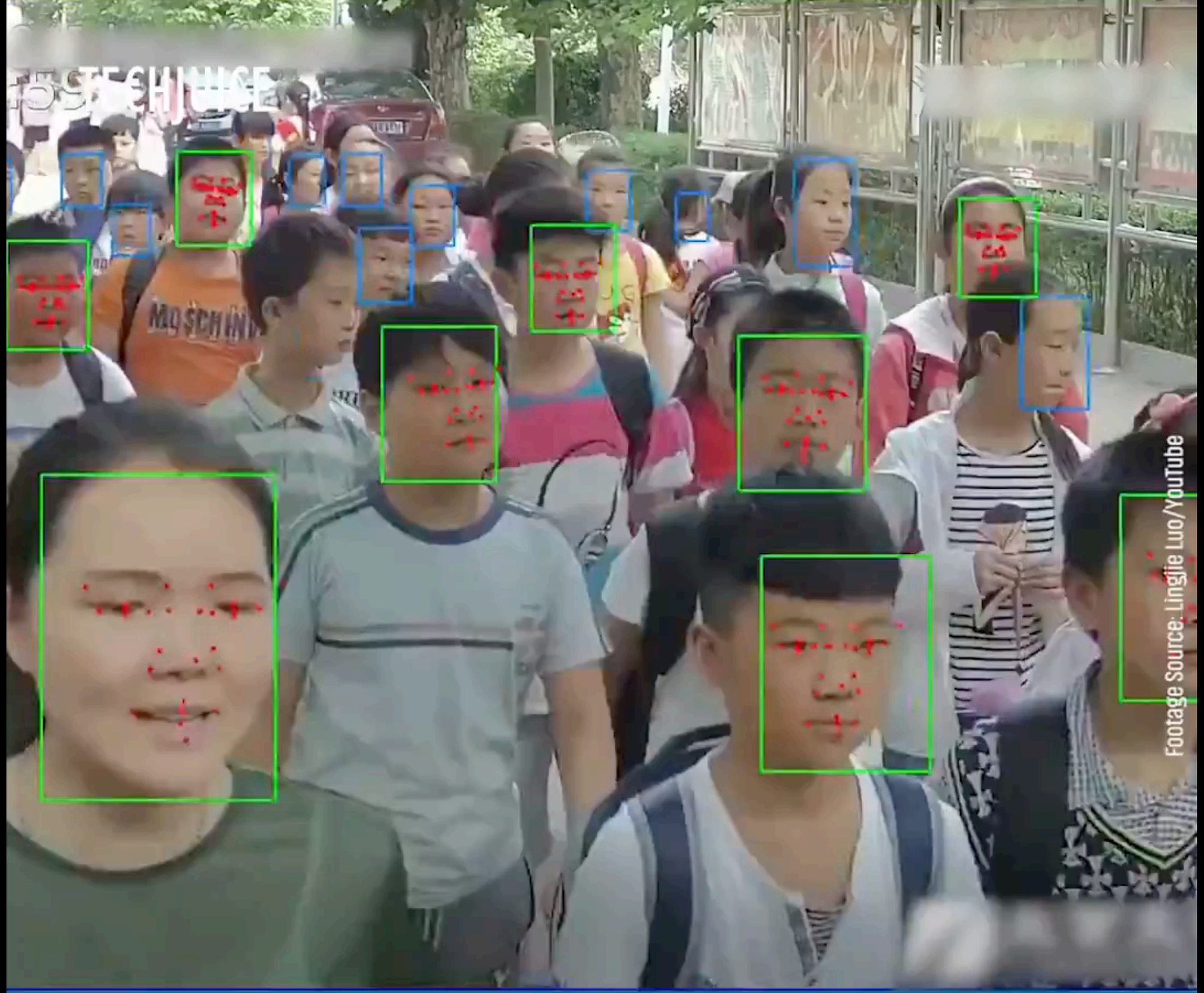
Beijing police are using facial-recognition glasses to identify car passengers and number plates

Tara Francis Chan

Mar. 12, 2018, 2:29 AM 3,899


Footage Source: Lingjie Luo/YouTube


ID60 Girl
17:06:58
ID07 Boy
17:06:58
ID55 Boy
17:06:58
ID58 Gi
17:06:5

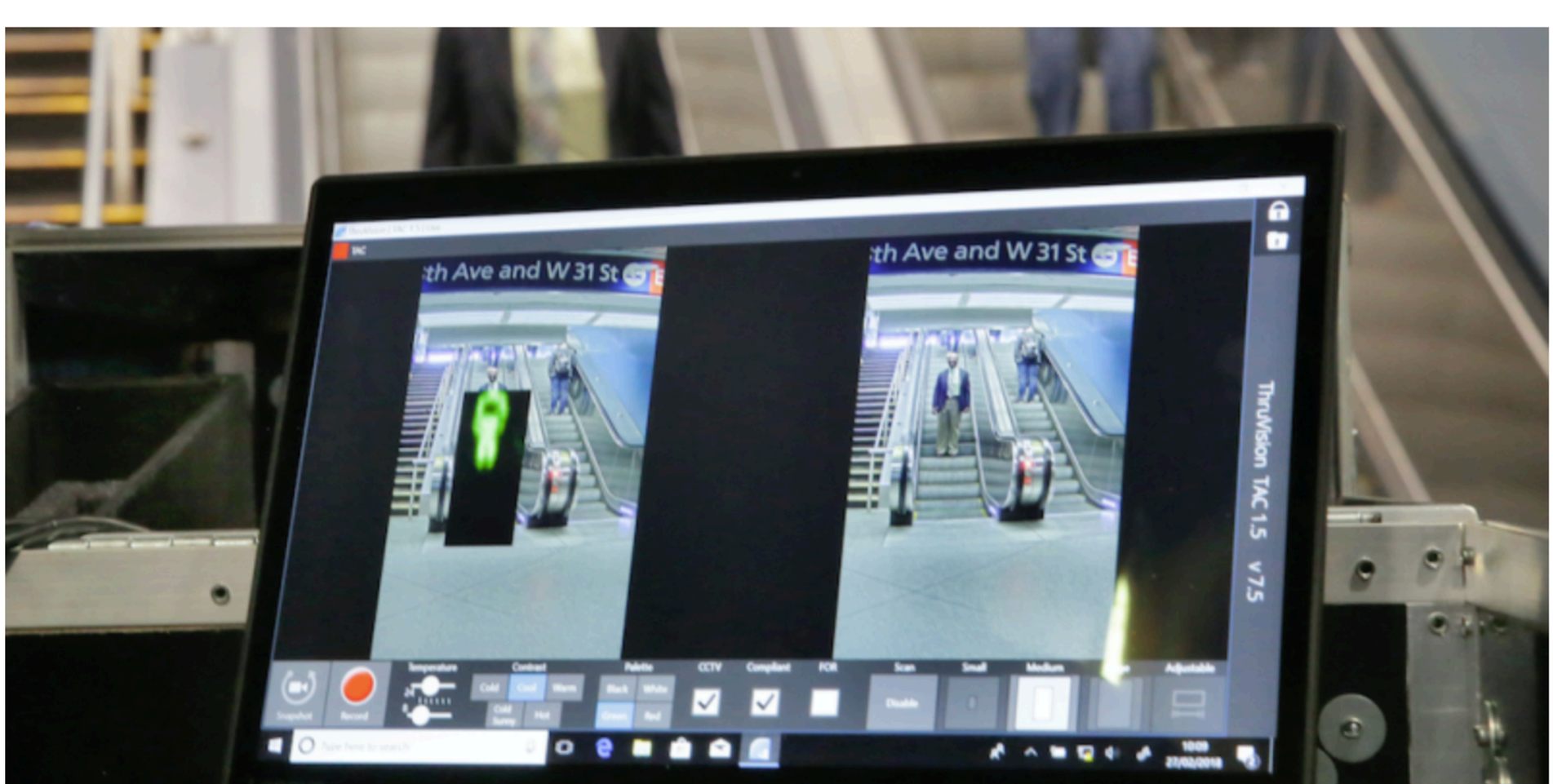

Footage Source: Lingjie Luo/YouTube


ID60 Girl
17:06:58
ID07 Boy
17:06:58
ID55 Boy
17:06:58
ID58 Gi
17:06:5

PRIVACY AND SECURITY

Madison Square Garden Has Been Secretly Using Face Recognition Tech: Report

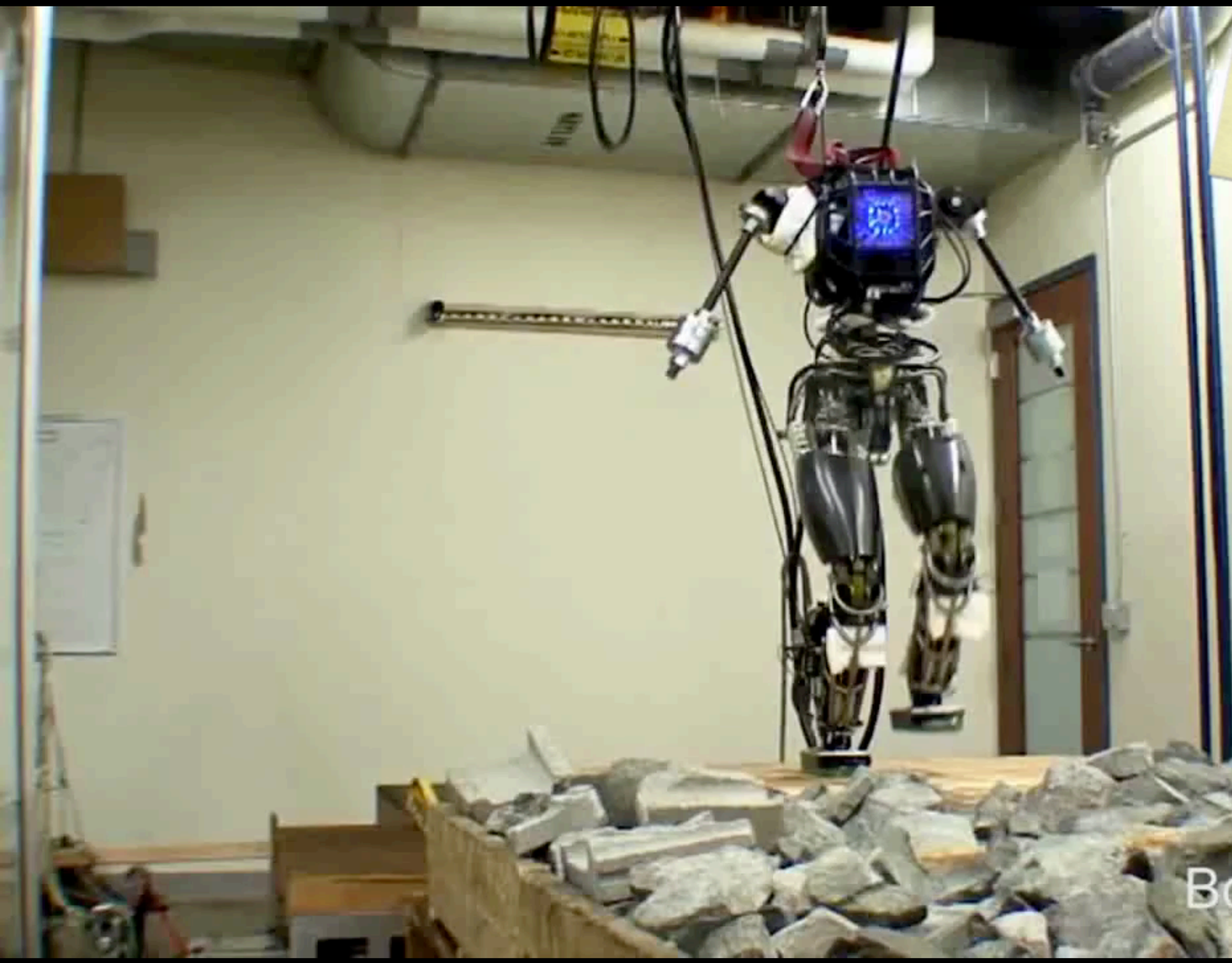
Sidney Fussell
3/13/18 6:50pm • Filed to: SURVEILLANCE



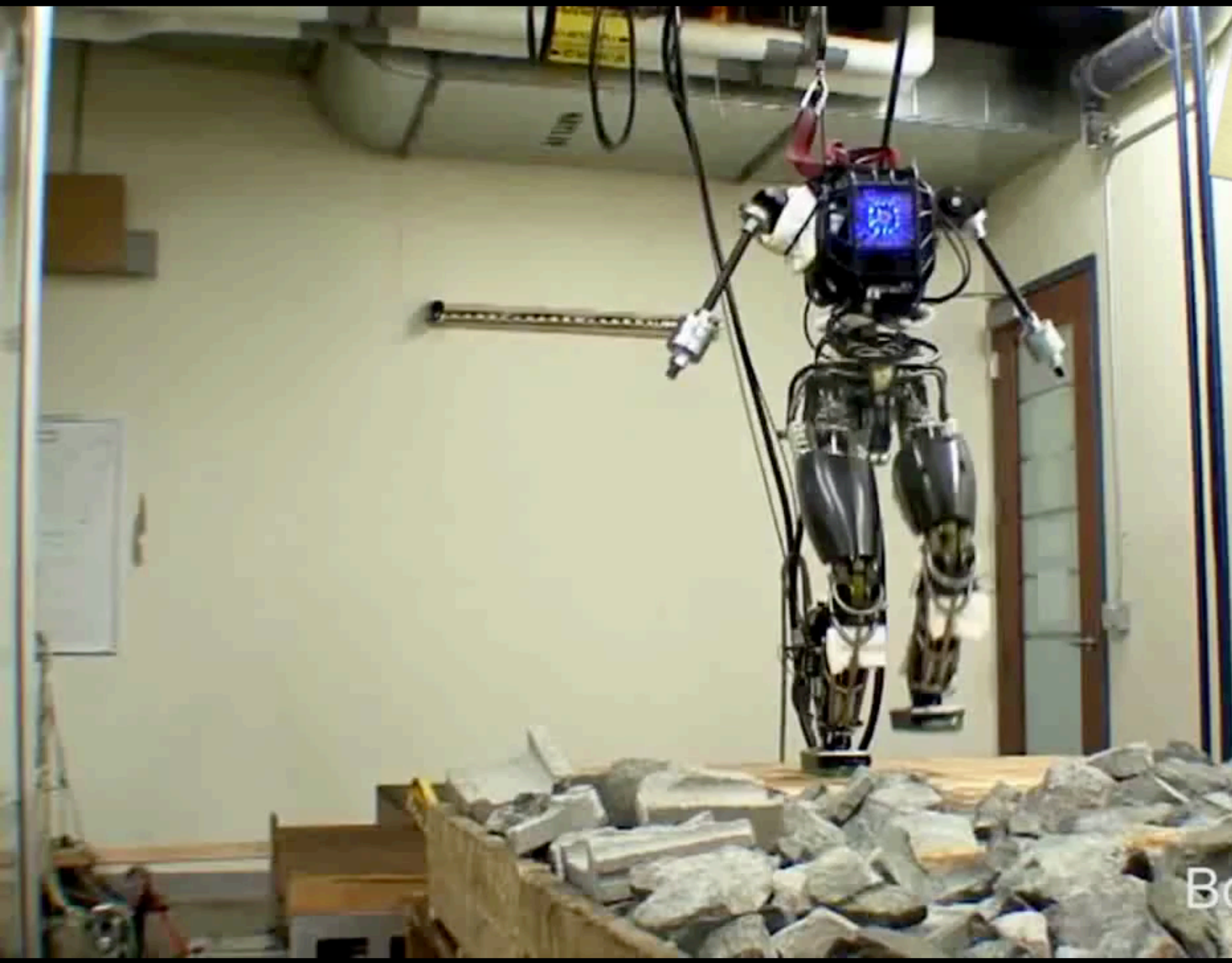
Product
Search

Product
Search

Robot Perception



Robot Perception




Boston Dynamics

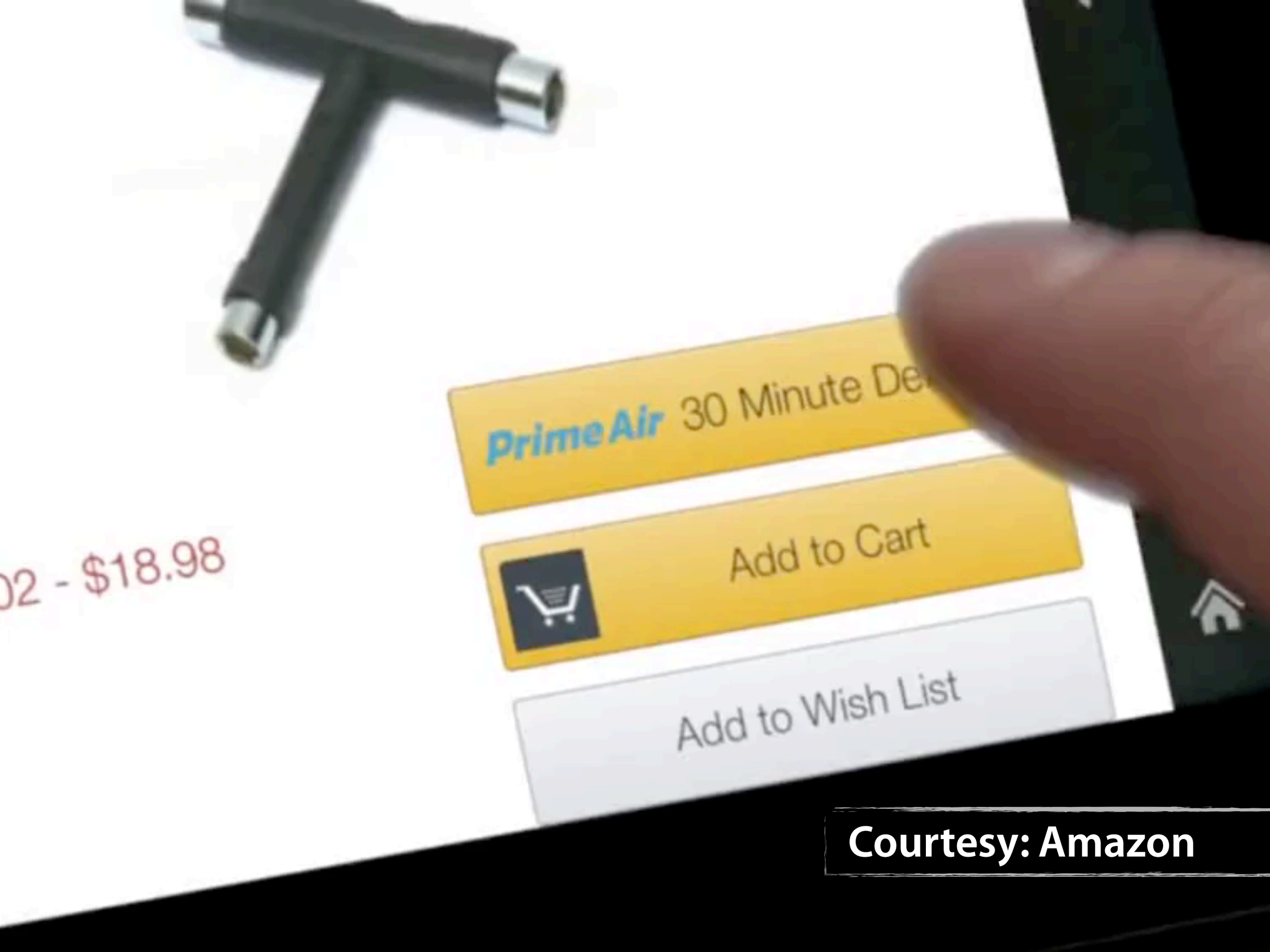

Boston Dynamics

dyson 360 eye

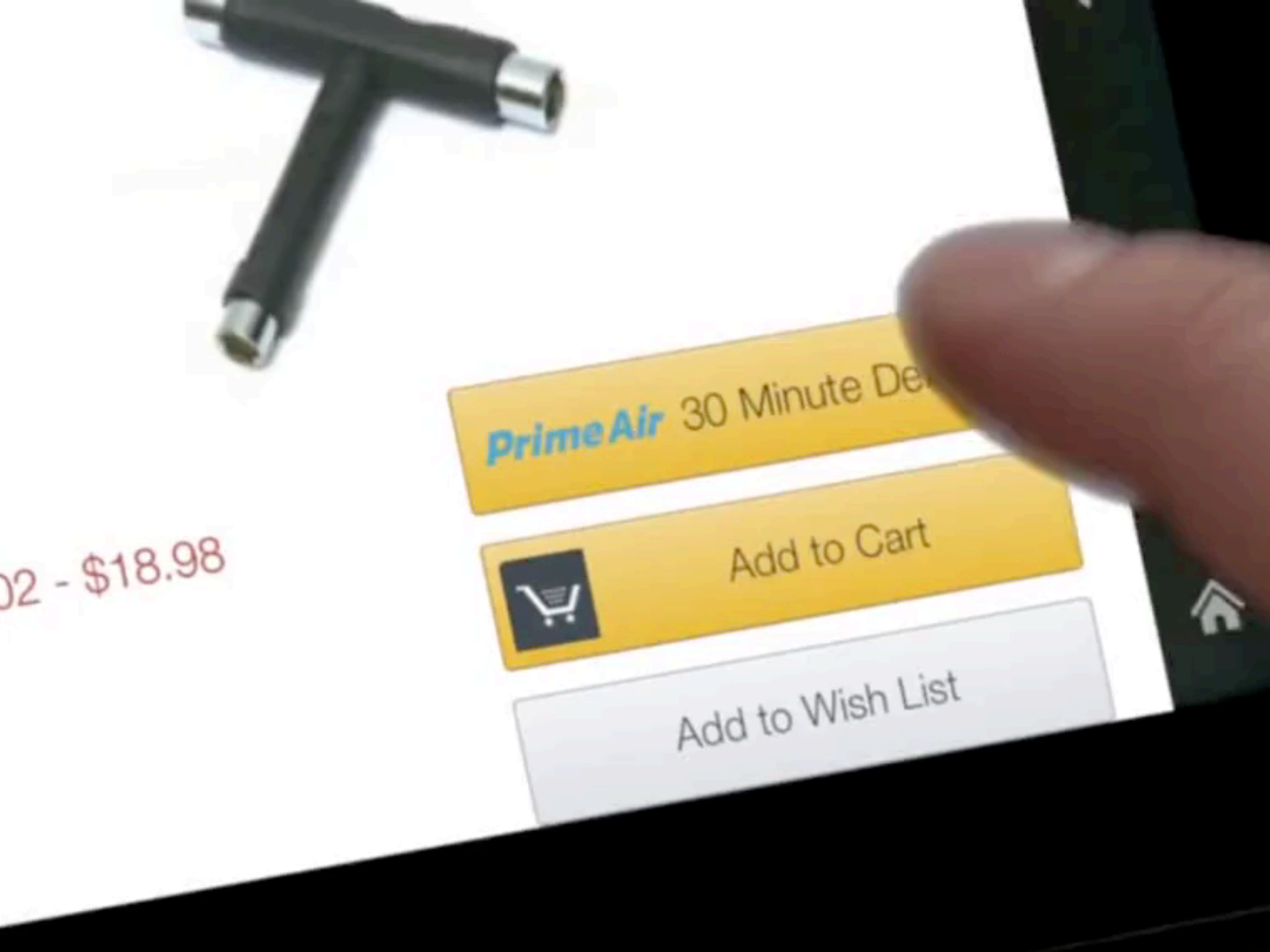
360 degree camera


360 Vision System
Infrared sensors


360 Vision System
Infrared sensors



Courtesy: Amazon


Prime Air 30 Minute De
Add to Cart
Add to Wish List
02 - $18.98

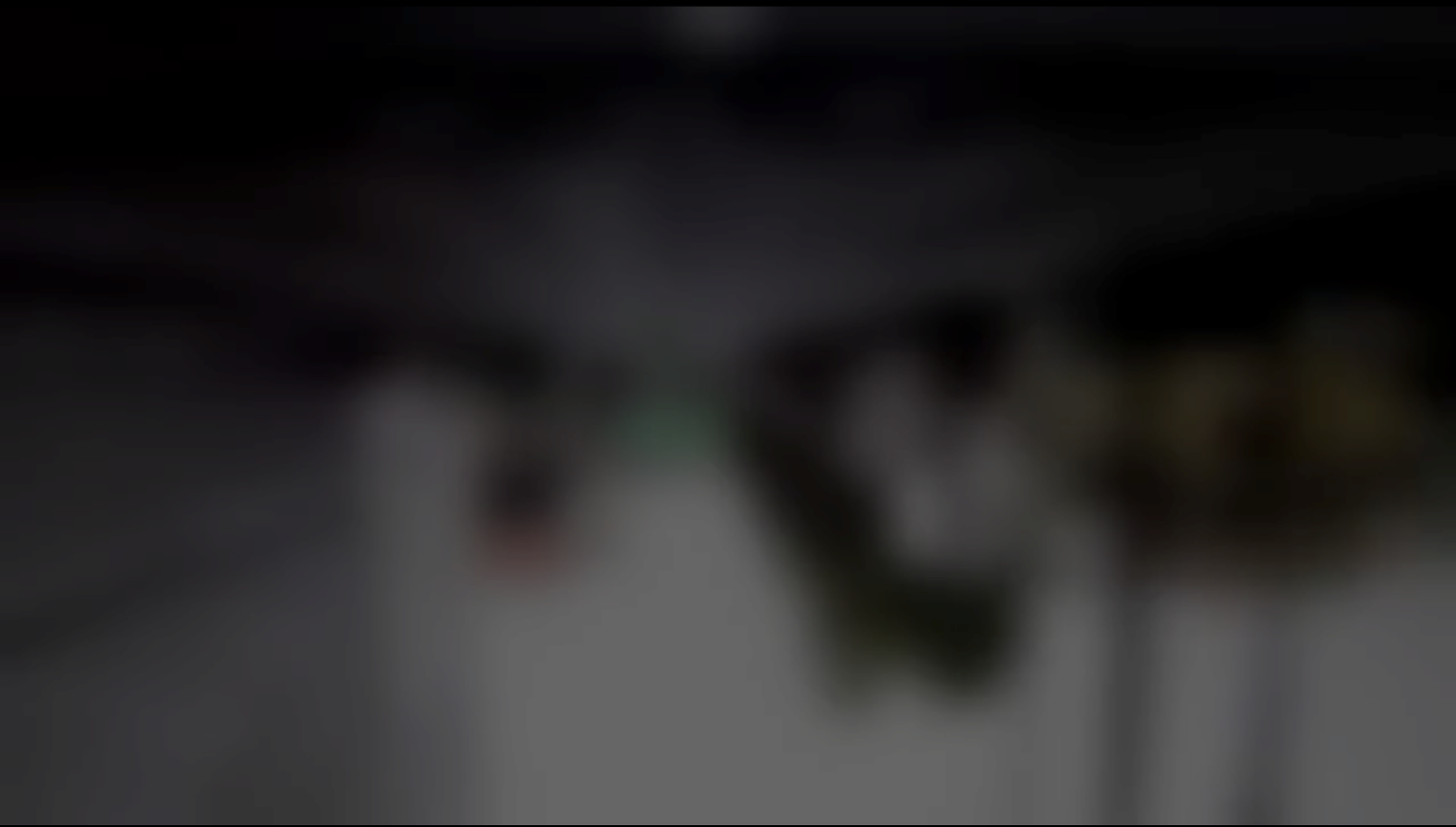

amazon go


amazon go

Self-Driving Cars

Courtesy: Nvidia


PASSENGER CAR



1.24 million road traffic deaths

Source: World Health Organization (WHO) 2010

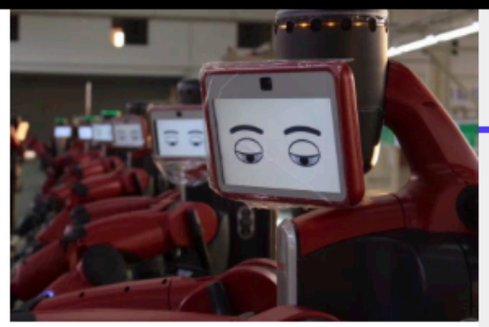
65.6 minutes commute time in Toronto

Source: Canadian Index of Wellbeing (CIW) 2014


國1 北 268K+410 水上路段
爆料公社
edge case (mis-)handling

國1 北 268K+410 水上路段

爆料公社



Search ...

RECENT POSTS

Peer Review

What Networks Will Co-Evolve With AI and Robotics?

Predictions Scorecard, 2020 January 01

Guest Post by Phillip Alvelda: Pondering the Empathy Gap

AGI Has Been Delayed

CATEGORIES

Dated Predictions

Essays

Quick Takes

Reviews

Whimsy

ARCHIVES

May 2020

April 2020

January 2020

May 2019

March 2019

January 2019

POST: EDGE CASES FOR SELF DRIVING CARS

JUNE 17, 2017 — ESSAYS

Edge Cases For Self Driving Cars

rodneybrooks.com/edge-cases-for-self-driving-cars/

Perhaps through this essay I will get the bee out of my bonnet[1] that fully driverless cars are a lot further off than many techies, much of the press, and even many auto executives seem to think. They will get here and human driving will probably disappear in the lifetimes of many people reading this, but it is not going to all happen in the blink of an eye as many expect. There are lots of details to be worked out.

In my very first post on this blog I talked about the unexpected consequences of having self driving cars. In this post I want to talk about about a number of edge cases, which I think will cause it to be a very long time before we have level 4 or level 5 self driving cars wandering our streets, especially without a human in them, and even then there are going to be lots of problems.

First though, we need to re-familiarize ourselves with the generally accepted levels of autonomy that every one is excited about for our cars.

Here are the levels from the autonomous car entry in Wikipedia which attributes this particular set to the SAE (Society of Automotive Engineers):

- *Level 0: Automated system has no vehicle control, but may issue warnings.*
- *Level 1: Driver must be ready to take control at any time. Automated system may include features such as Adaptive Cruise Control (ACC), Parking Assistance with automated steering, and Lane Keeping Assistance (LKA) Type II in any combination.*
- *Level 2: The driver is obliged to detect objects and events and respond if the automated system fails to respond properly. The automated system executes accelerating, braking, and steering. The automated system can deactivate immediately upon takeover by the driver.*
- *Level 3: Within known, limited environments (such as freeways), the driver can safely turn their attention away from driving tasks, but must still be prepared to take control when needed.*
- *Level 4: The automated system can control the vehicle in all but a few environments such as severe weather. The driver must enable the automated system only when it is safe to do so. When enabled, driver attention is not required.*

rodneybrooks.com

Edge Cases For Self Driving Cars – Rodney Brooks

Search ...

RECENT POSTS

Peer Review

What Networks Will Co-Evolve

Overcoming the Empathy Gap

AGI Has Been Delayed

CATEGORIES

Dated Predictions

Essays

Quick Takes

Reviews

Whimsy

ARCHIVES

May 2020

April 2020

January 2020

May 2019

March 2019

January 2019

POST: EDGE CASES FOR SELF DRIVING CARS

JUNE 17, 2017 — ESSAYS

Edge Cases For Self Driving Cars

rodneybrooks.com/edge-cases-for-self-driving-cars/

Perhaps through this essay I will get the bee out of my bonnet[1] that fully driverless cars are a lot further off than many techies, much of the press, and even many auto executives seem to think. They will get here and human driving will probably disappear in the lifetimes of many people reading this, but it is not going to all happen in the blink of an eye as many expect. There are lots of details to be worked out.

tinyurl.com/edge-cases

[illegible] without a human in them, and even then there are going to be lots of problems.

First though, we need to re-familiarize ourselves with the generally accepted levels of autonomy that every one is excited about for our cars.

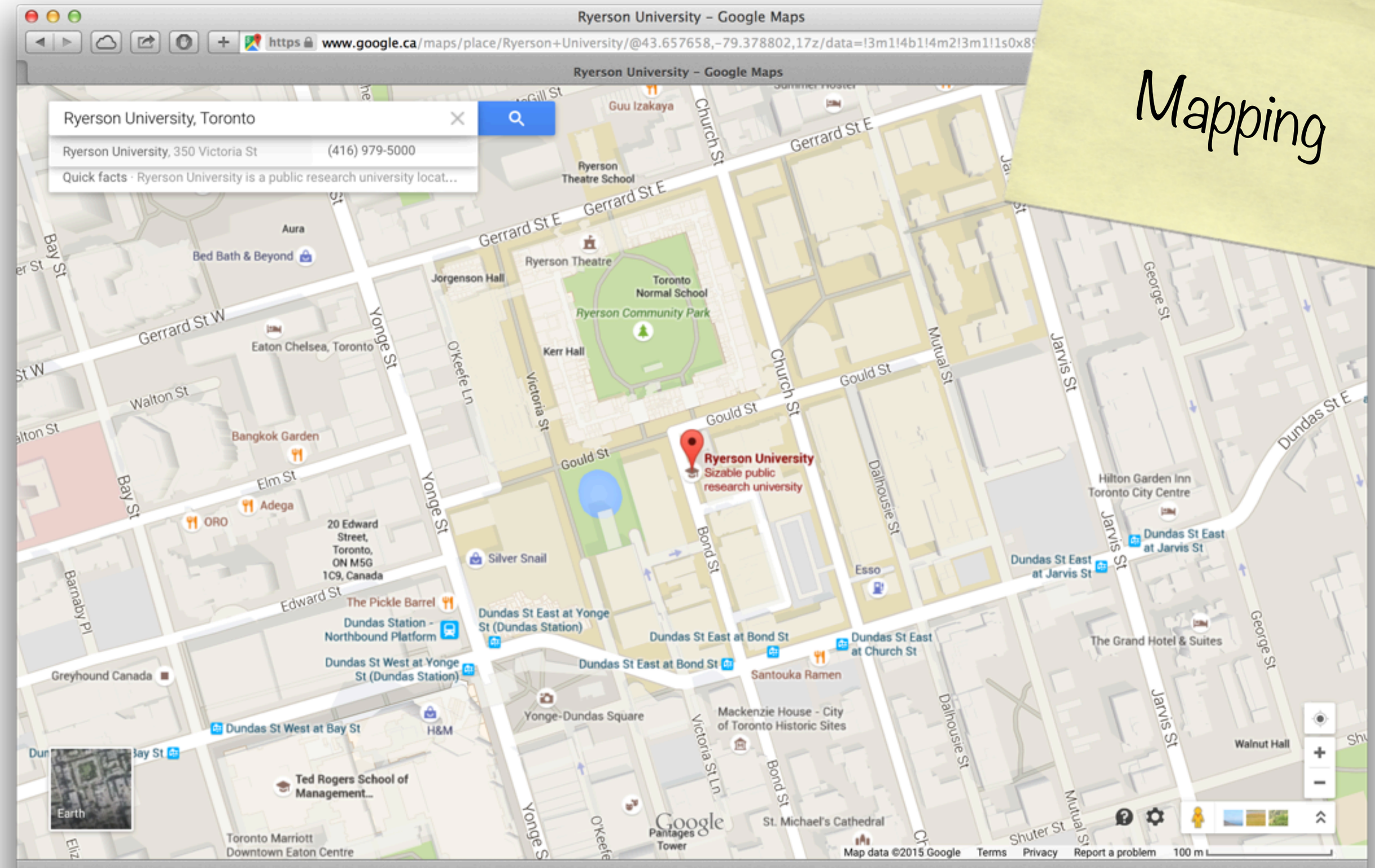
Here are the levels from the autonomous car entry in Wikipedia which attributes this particular set to the SAE (Society of Automotive Engineers):

- *Level 0: Automated system has no vehicle control, but may issue warnings.*
- *Level 1: Driver must be ready to take control at any time. Automated system may include features such as Adaptive Cruise Control (ACC), Parking Assistance with automated steering, and Lane Keeping Assistance (LKA) Type II in any combination.*
- *Level 2: The driver is obliged to detect objects and events and respond if the automated system fails to respond properly. The automated system executes accelerating, braking, and steering. The automated system can deactivate immediately upon takeover by the driver.*
- *Level 3: Within known, limited environments (such as freeways), the driver can safely turn their attention away from driving tasks, but must still be prepared to take control when needed.*
- *Level 4: The automated system can control the vehicle in all but a few environments such as severe weather. The driver must enable the automated system only when it is safe to do so. When enabled, driver attention is not required.*

Produce
Sorting

Produce
Sorting

Mapping

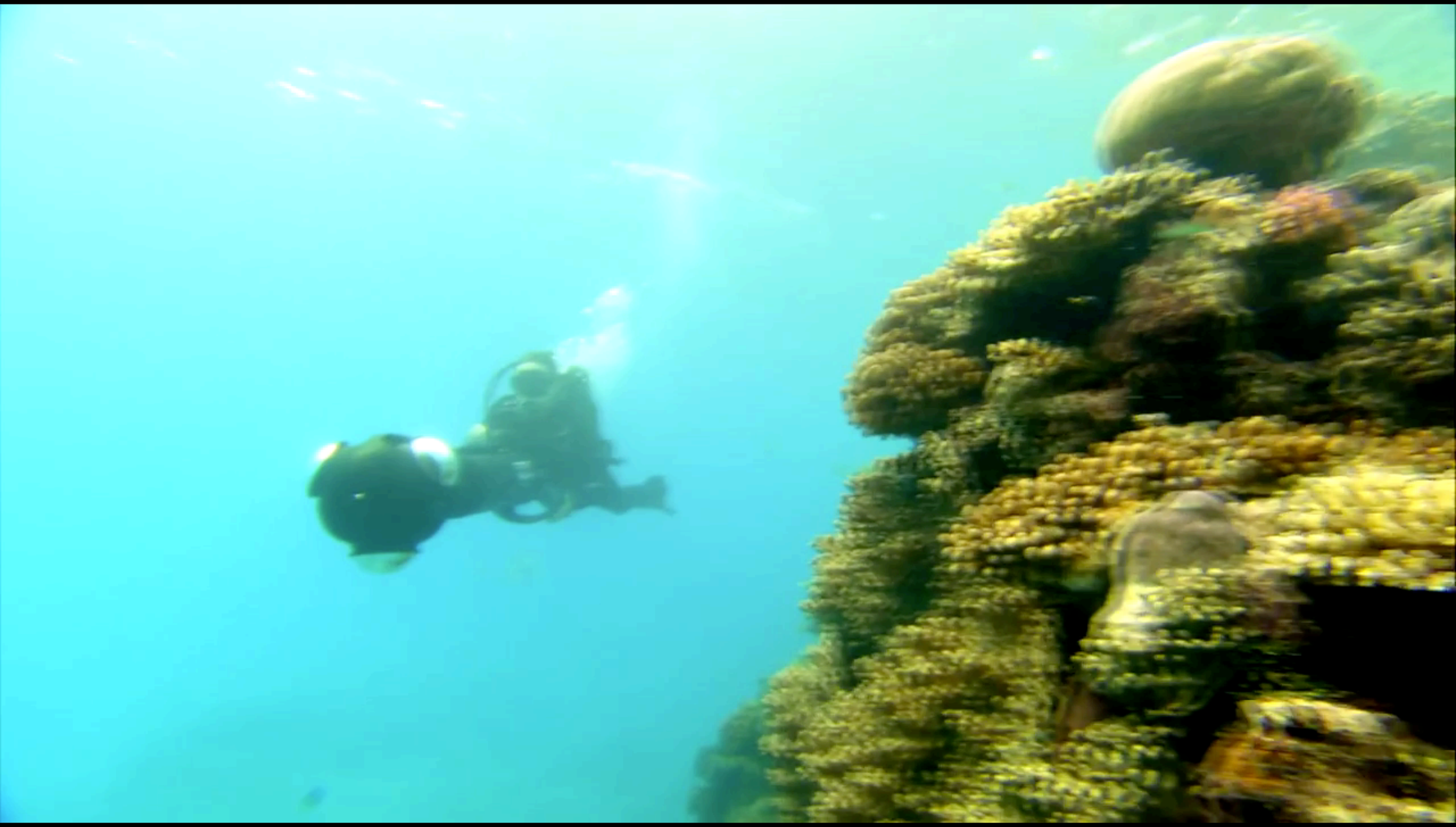

Ryerson University - Google Maps
https www.google.ca/maps/place/Ryerson+University/@43.657658,-79.378802,17z/data=!3m1!4b1!4m2!3m1!1s0x8
Ryerson University - Google Maps
Ryerson University, Toronto
Ryerson University, 350 Victoria St (416) 979-5000
Quick facts · Ryerson University is a public research university locat...
Guu Izakaya
Church St
Gerrard St E
Ryerson Theatre School
Aura
Bed Bath & Beyond
Bay St
Jorgenson Hall
Ryerson Theatre
Toronto Normal School
Ryerson Community Park
Gerrard St W
Eaton Chelsea, Toronto
Yonge St
O'Keefe Ln
Victoria St
Kerr Hall
Walton St
Bangkok Garden
Elm St
Adega
ORO
20 Edward Street, Toronto, ON M5G 1C9, Canada
Barnaby Pl
Edward St
Silver Snail
The Pickle Barrel
Dundas Station - Northbound Platform
Dundas St East at Yonge St (Dundas Station)
Dundas St West at Yonge St (Dundas Station)
Greyhound Canada
Dundas St West at Bay St
H&M
Ted Rogers School of Management...
Toronto Marriott Downtown Eaton Centre
Earth
Gould St
Ryerson University
Sizable public research university
Bond St
Dundas St East at Bond St
Yonge-Dundas Square
Mackenzie House - City of Toronto Historic Sites
St. Michael's Cathedral
Pantages Tower
Google
Esso
Santouka Ramen
Dundas St East at Church St
Mutual St
Dalhousie St
Jarvis St
George St
Dundas St East at Jarvis St
Hilton Garden Inn Toronto City Centre
The Grand Hotel & Suites
Walnut Hall
Dundas St E
Shuter St
Map data ©2015 Google
Terms
Privacy
Report a problem
100 m

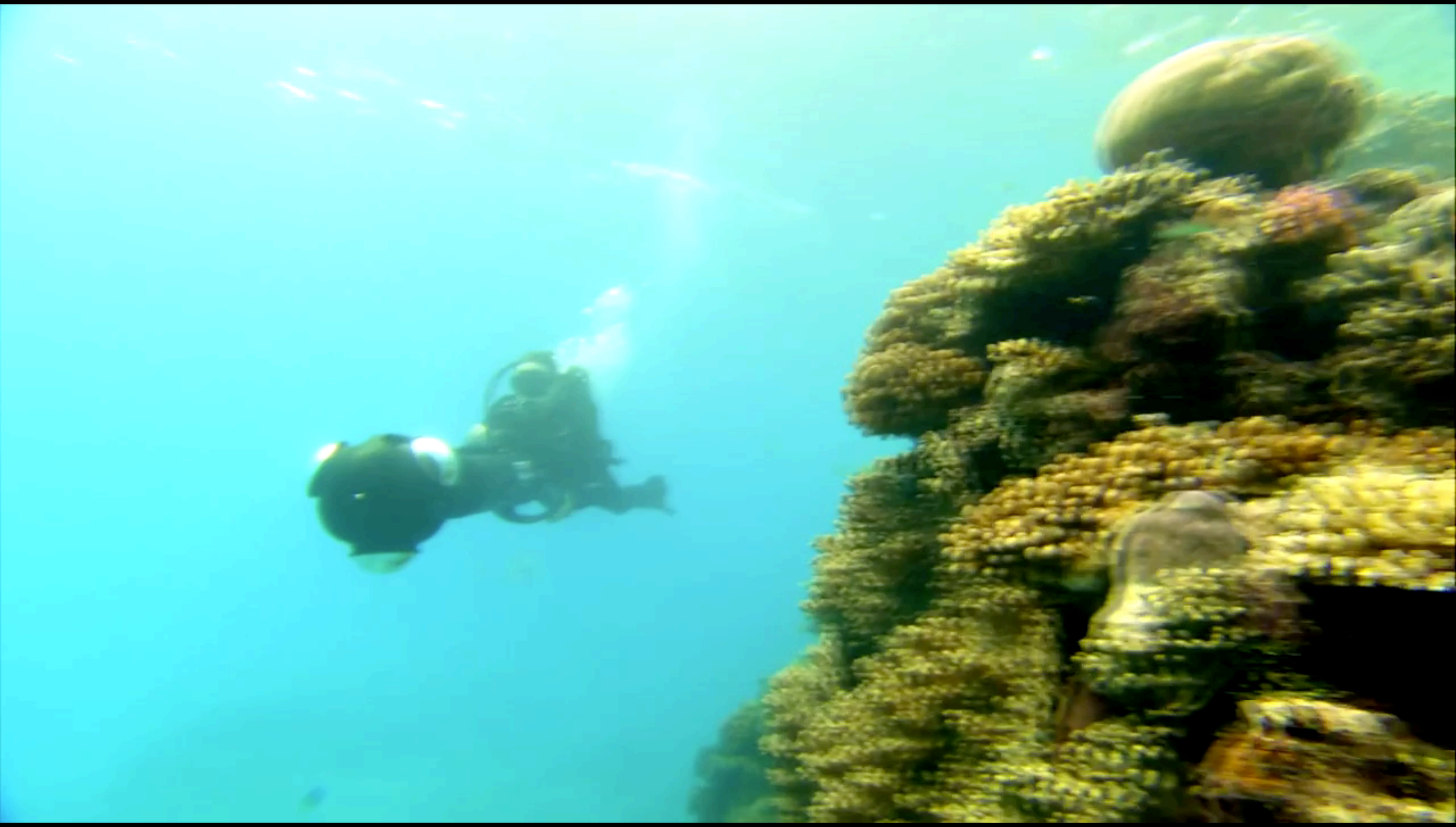
Google Street View


JESSE JAMES IS A DEAD MAN
Foot Locker
TOSHIBA
THE LION KING
ENTRANCE ON 45TH STREET
Budweiser
Justin Drew Pictures
Broadway
7th Ave
Blecher Mendel & Fedele
Bella Vita Pizzeria
New Jersey
New York
Central Par
Midtown East
Flatiron District
25
495


Google maps
Street View
google.com/streetview


W 42 ST
ONE WAY
citibank



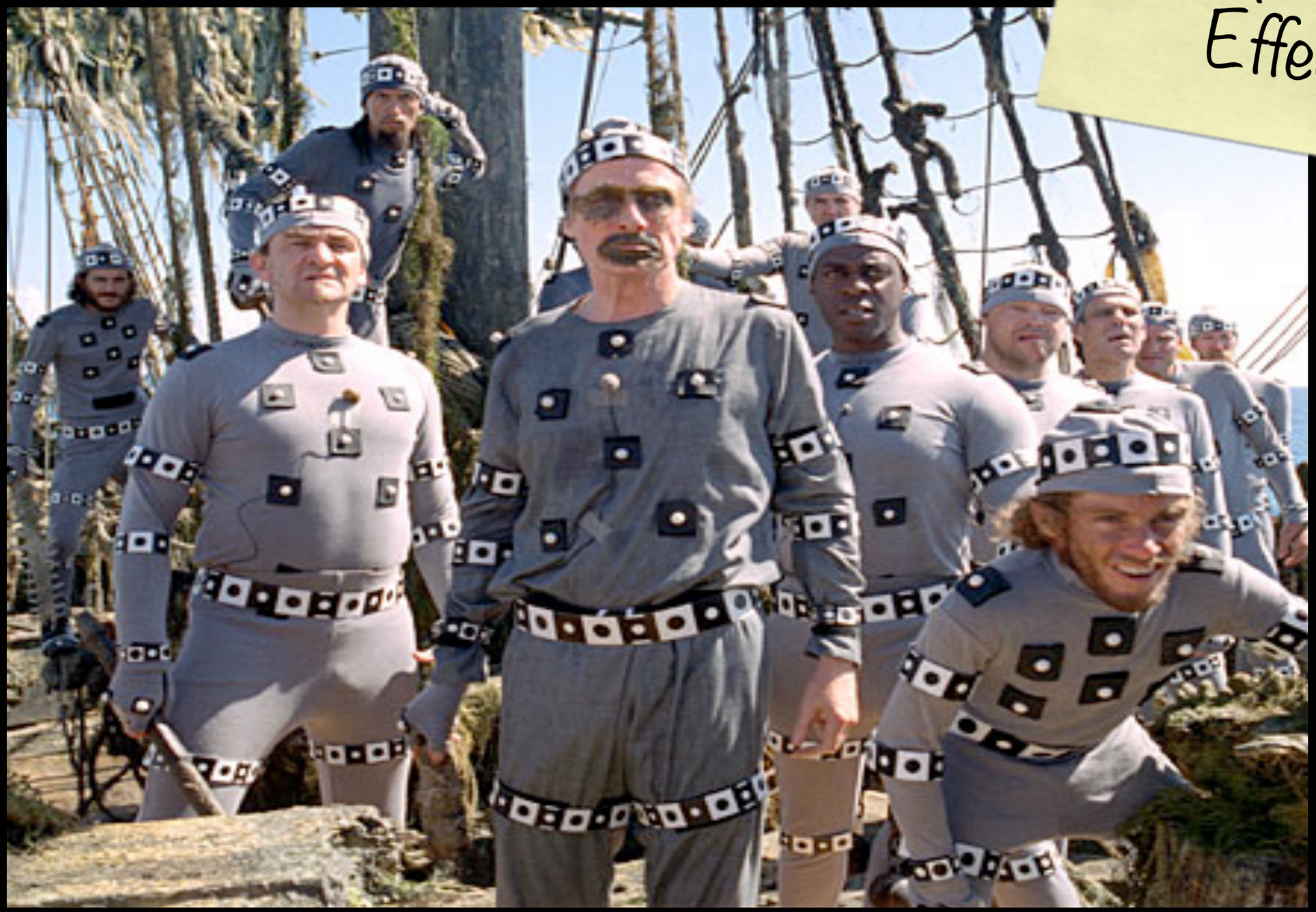



Marriott


Marriott

Special Effects

Special Effects



Source: Avengers (2010)

Source: Avengers (2010)

CAESAR

Source: Dawn of the Planet of the Apes (2014)

CAESAR

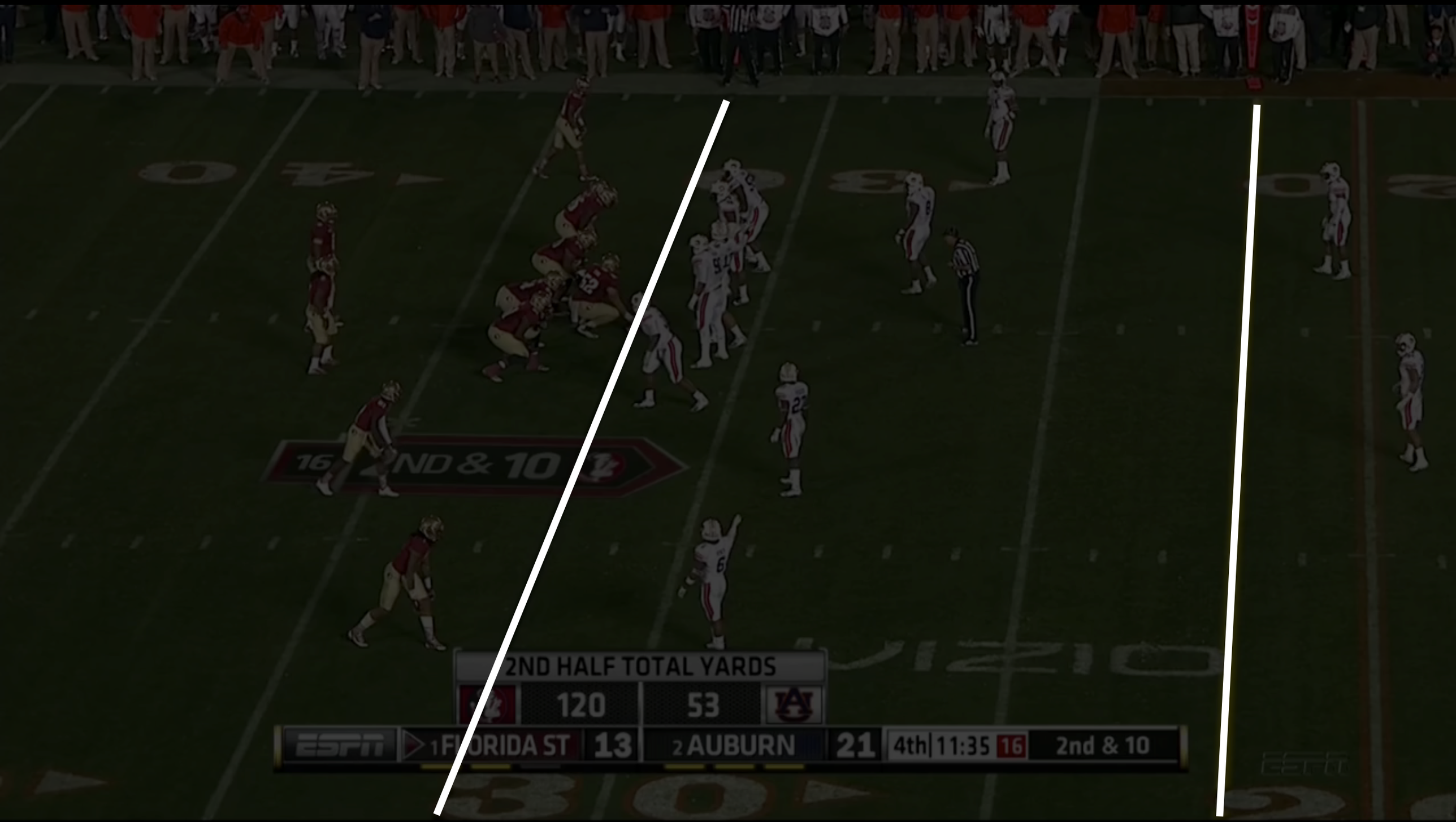
Sports

16 2ND & 10

2ND HALF TOTAL YARDS

120		53

ESPN | 1 FLORIDA ST 13 | 2 AUBURN 21 | 4th 11:35 16 | 2nd & 10


16
2ND & 10
2ND HALF TOTAL YARDS
120
53
ESPN
1 FLORIDA ST
13
2 AUBURN
21
4th
11:35
16
2nd & 10

16 2ND & 10

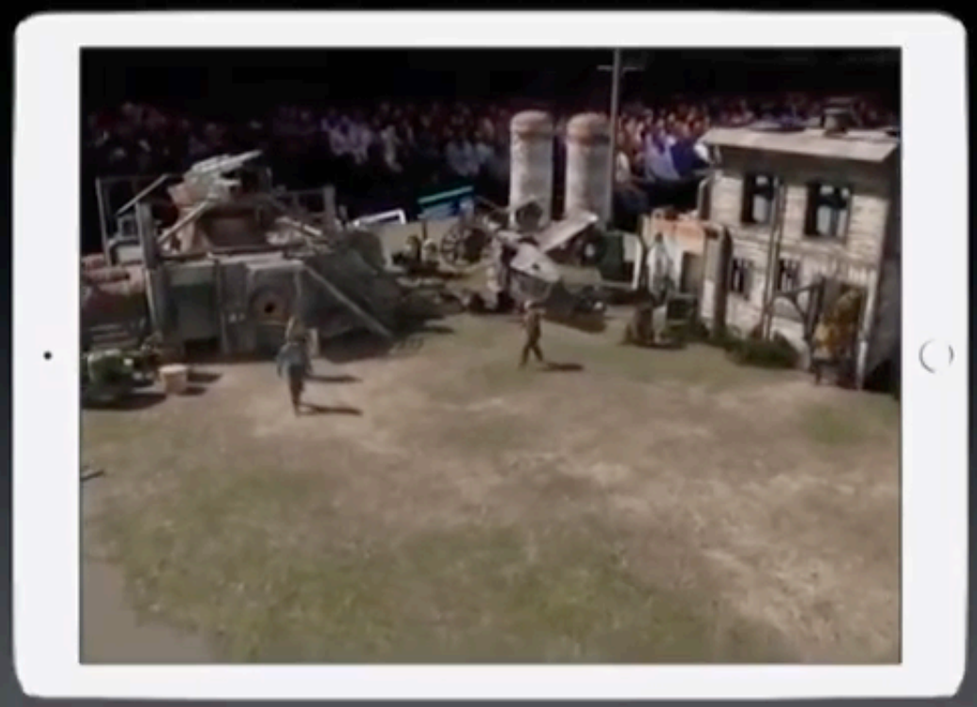
2ND HALF TOTAL YARDS

120		53

ESPN | 1 FLORIDA ST 13 | 2 AUBURN 21 | 4th 11:35 16 2nd & 10

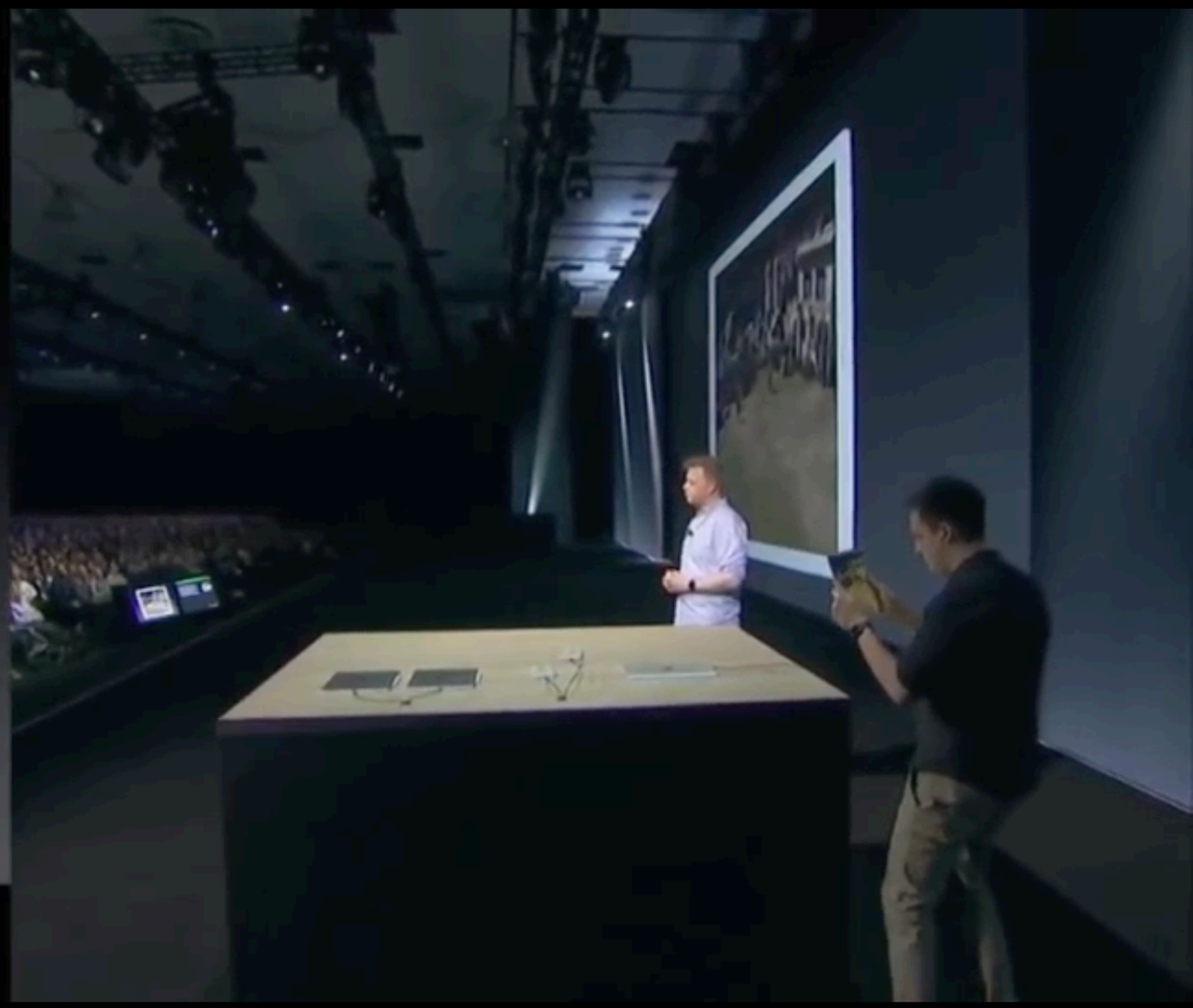
60 MINUTES
sports

SHOWTIME



Gaming

Courtesy: Apple and Wingnut AR



Google AR Stickers


EXIT
911

KINECT

Microsoft
HoloLens


cameras

Vision-based tracking

Facebook Reality Labs and Oculus


oculus

Courtesy: M. Rubinstein et al.

Courtesy: M. Rubinstein et al.

Courtesy: M. Rubinstein et al.


00;00;00;11


00;00;00;11

Human pulse

Respiratory motion


156 bpm


156 bpm

The New York Times

ROBOTICS

Computer Eyesight Gets a Lot More Accurate

By John Markoff

August 18, 2014 8:01 pm

Just as the Big Bad Wolf promised Little Red Riding Hood that his bigger eyes were “the better to see you with,” a machine’s ability to see the world around it is benefiting from bigger computers and more accurate mathematical calculations.

The improvement was visible in contest results released Monday evening by computer scientists and companies that sponsor an annual challenge to measure improvements in the state of machine vision technology.

Started in 2010 by Stanford, Princeton and Columbia University scientists, the Large Scale Visual Recognition Challenge this year drew 38 entrants from 13 countries. The groups use advanced software, in most cases modeled loosely on the biological vision systems, to detect, locate and classify a huge set of images taken from Internet sources like Twitter. The contest was sponsored this year by Google, Stanford, Facebook and the University of North Carolina.

Contestants run their recognition programs on high-performance computers based in many cases on specialized processors called G.P.U.s, for graphic processing units.

This year there were six categories based on object detection, locating objects and classifying them. Winners included the National University of Singapore, the Oxford University, Adobe Systems, the Center for Intelligent Perception and Computing at the Chinese Academy of Sciences, as well as Google in two separate categories.

Accuracy almost doubled in the 2014 competition and error rates were cut in

MIT Technology Review

Facebook Creates Software That Matches Faces Almost as Well as You Do

Facebook's new AI research group reports a major improvement in face-processing software.

By Tom Simonite on March 17, 2014

Asked whether two unfamiliar photos of faces show the same person, a human being will get it right 97.53 percent of the time. New software developed by researchers at Facebook can score 97.25 percent on the same challenge, regardless of variations in lighting or whether the person in the picture is directly facing the camera.

That's a significant advance over previous face-matching software, and it demonstrates the power of a new approach to artificial intelligence known as deep learning, which Facebook and its competitors have bet heavily on in the past year (see "Deep Learning"). This area of AI involves software that uses networks of simulated neurons to learn to recognize patterns in large amounts of data.

"You normally don't see that sort of improvement," says Yaniv Taigman, a member of Facebook's AI team, a research group created last year to explore how deep learning might help the company (see "Facebook Launches Advanced AI Effort"). "We closely approach human performance," says Taigman of the new software. He notes that the error rate has been reduced by more than a quarter relative to earlier software that can take on the same task.


IBM Revenue Slides, but Cloud Business Grows


Review: Public Wi-Fi's Fast Future


Local M&A Gains Traction in Israeli High-Tech

TECH

Driverless Cars Steal the Show

Makers and suppliers highlight commitment to autonomous vehicles at Consumer Electronics Show

Gill Pratt, head of Toyota's artificial intelligence company, spoke at the Consumer Electronics Show in Las Vegas on Tuesday. *PHOTO: REUTERS*

By MIKE RAMSEY

9 COMMENTS

Jan. 5, 2016 5:12 p.m. ET

LAS VEGAS—Autonomous cars are taking over the Consumer Electronics Show.

In the early days of the tech event in Las Vegas, car makers and nascent auto suppliers, such as chip maker Nvidia Corp., have made announcements about deepening commitments to driverless vehicles.

Toyota Motor Corp. has committed $1 billion to an artificial

Most Popular Videos

1. Familiar Auto Features Go the Way of the Dodo

2. The Future of Public WiFi: What You Need to Know

3. Why Global Demand for Oil Is Slowing Down

4. Microsoft Builds 'Minecraft' Classroom Edition

5. Creating a Blue Rose: The Holy Grail of Horticulture


Most Popular Articles

1. Thousands Apply to U.S. to Forgive Their Student Loans

2. Opinion: Hillary Is in Big Trouble

3. Opinion: Taking Sanders Seriously

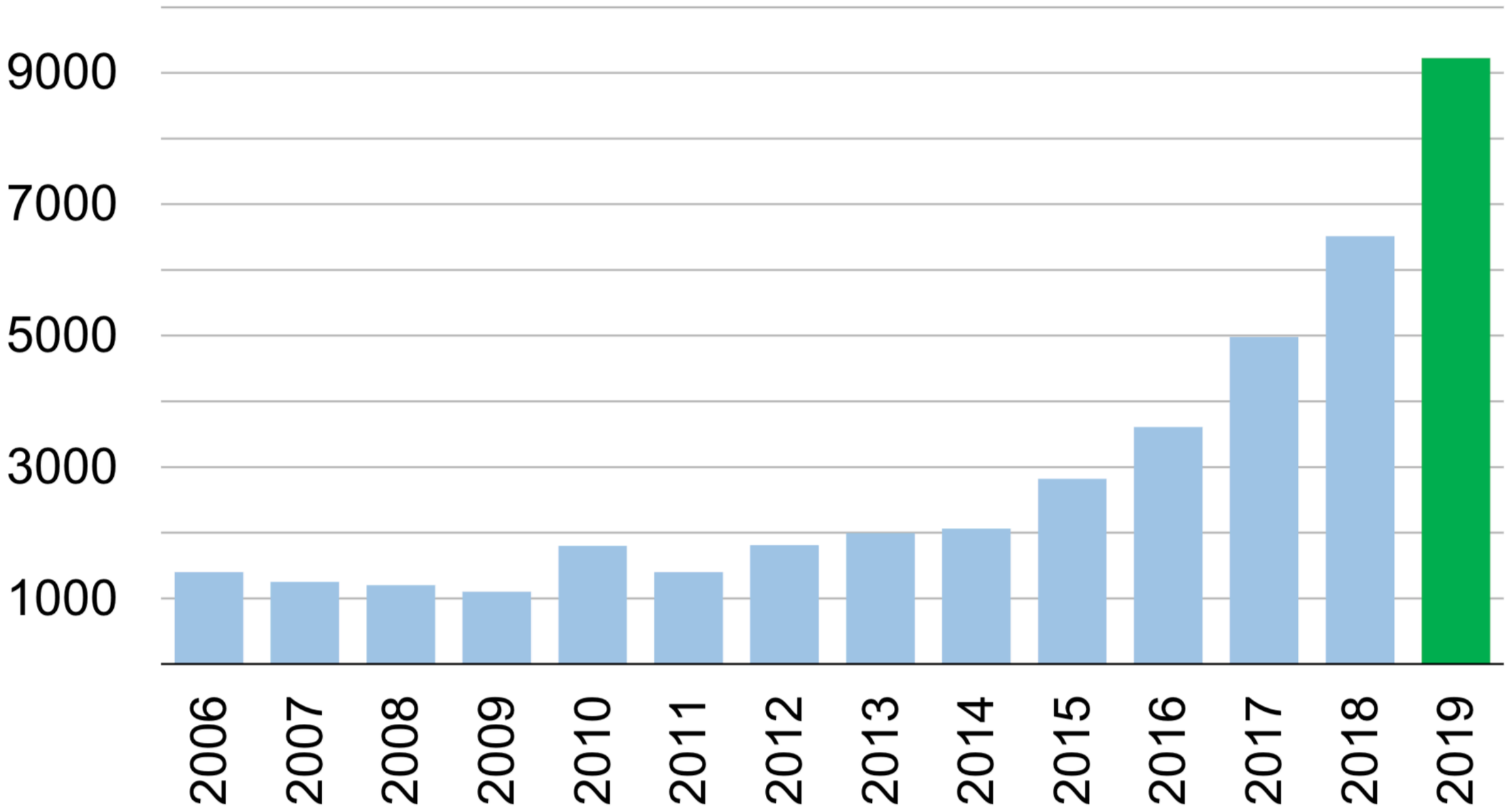

CVPR 2019, Long Beach

CVPR 2019, Long Beach

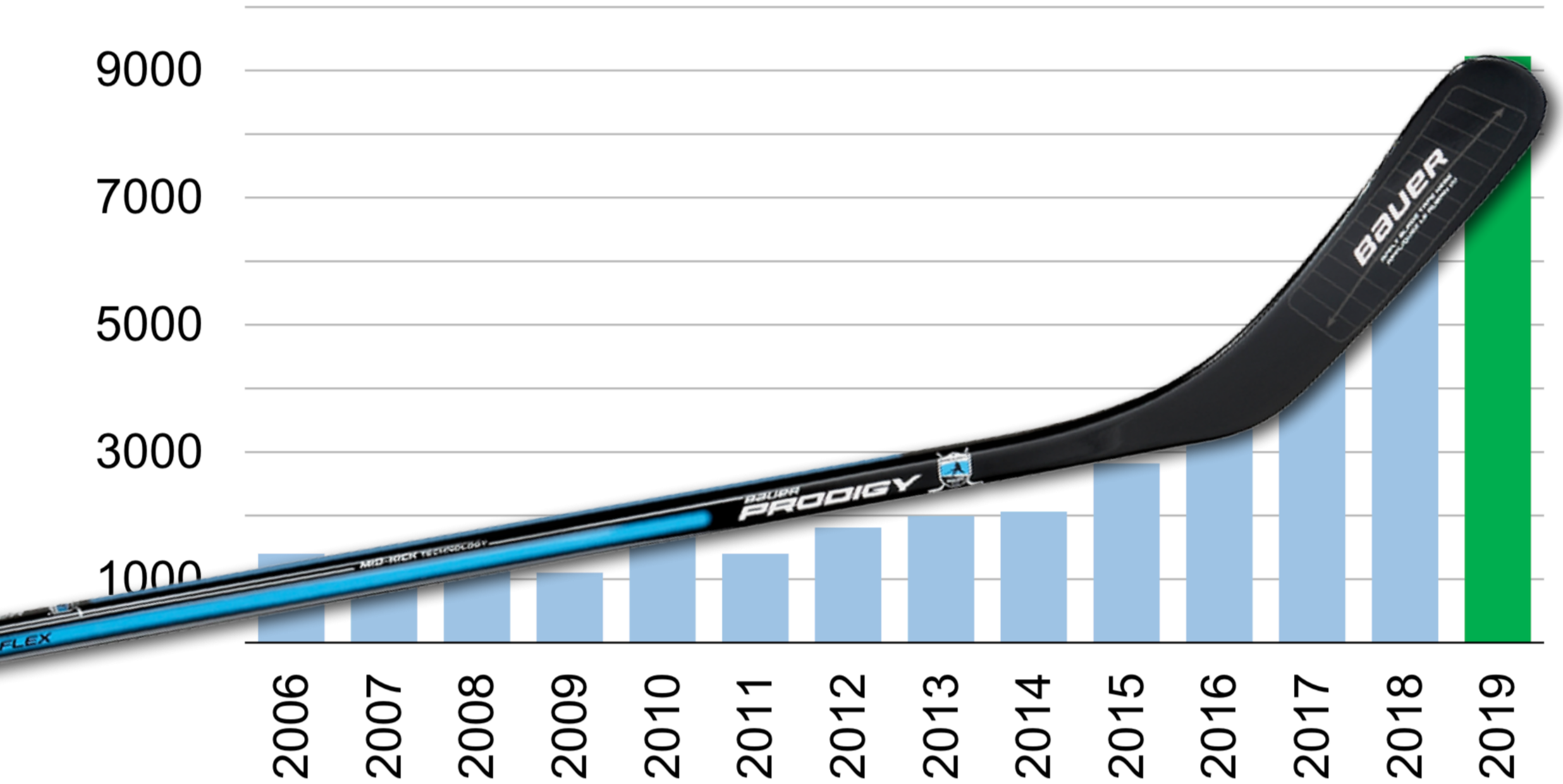
9200

attendees

CVPR attendance



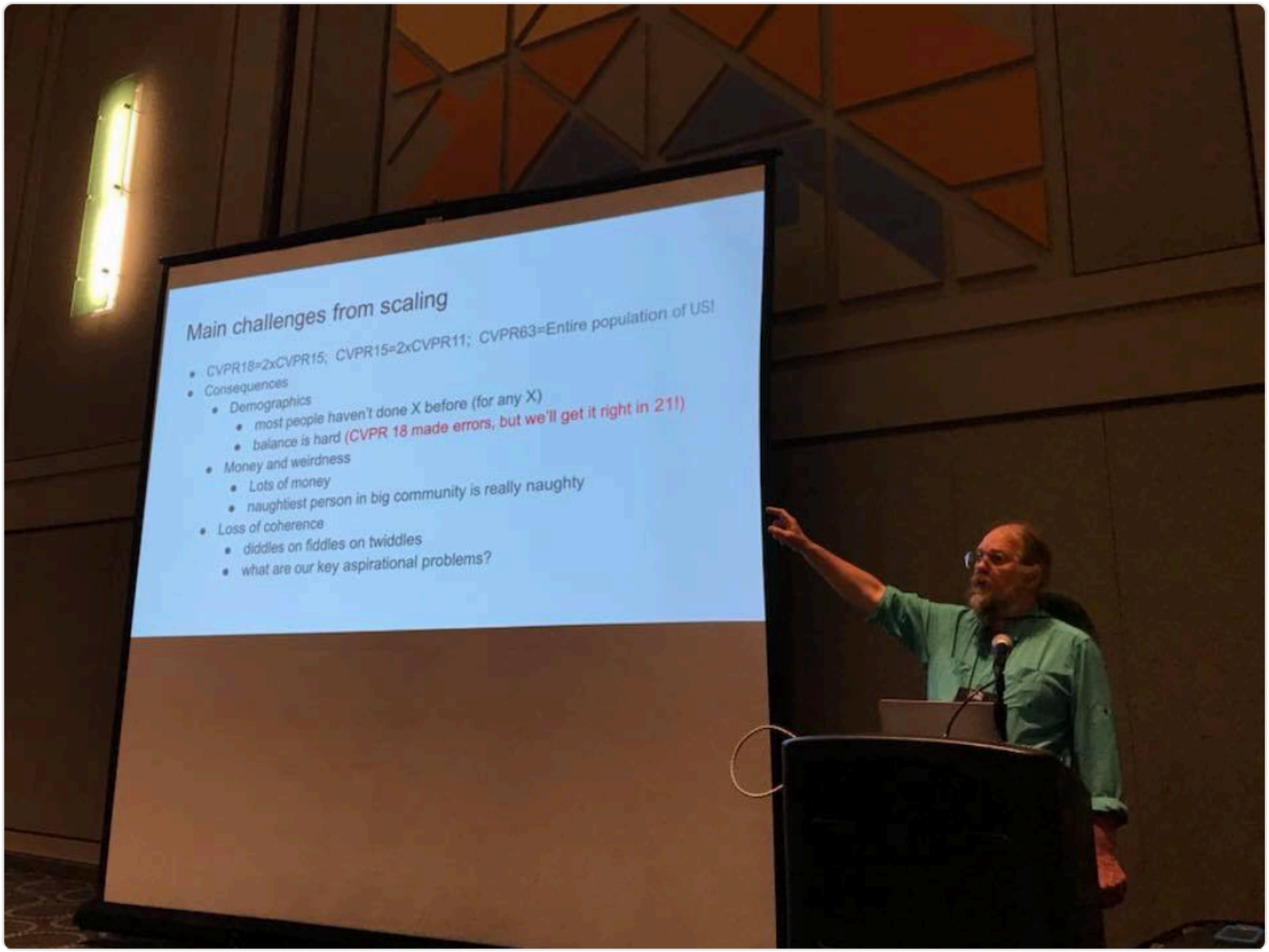
CVPR attendance

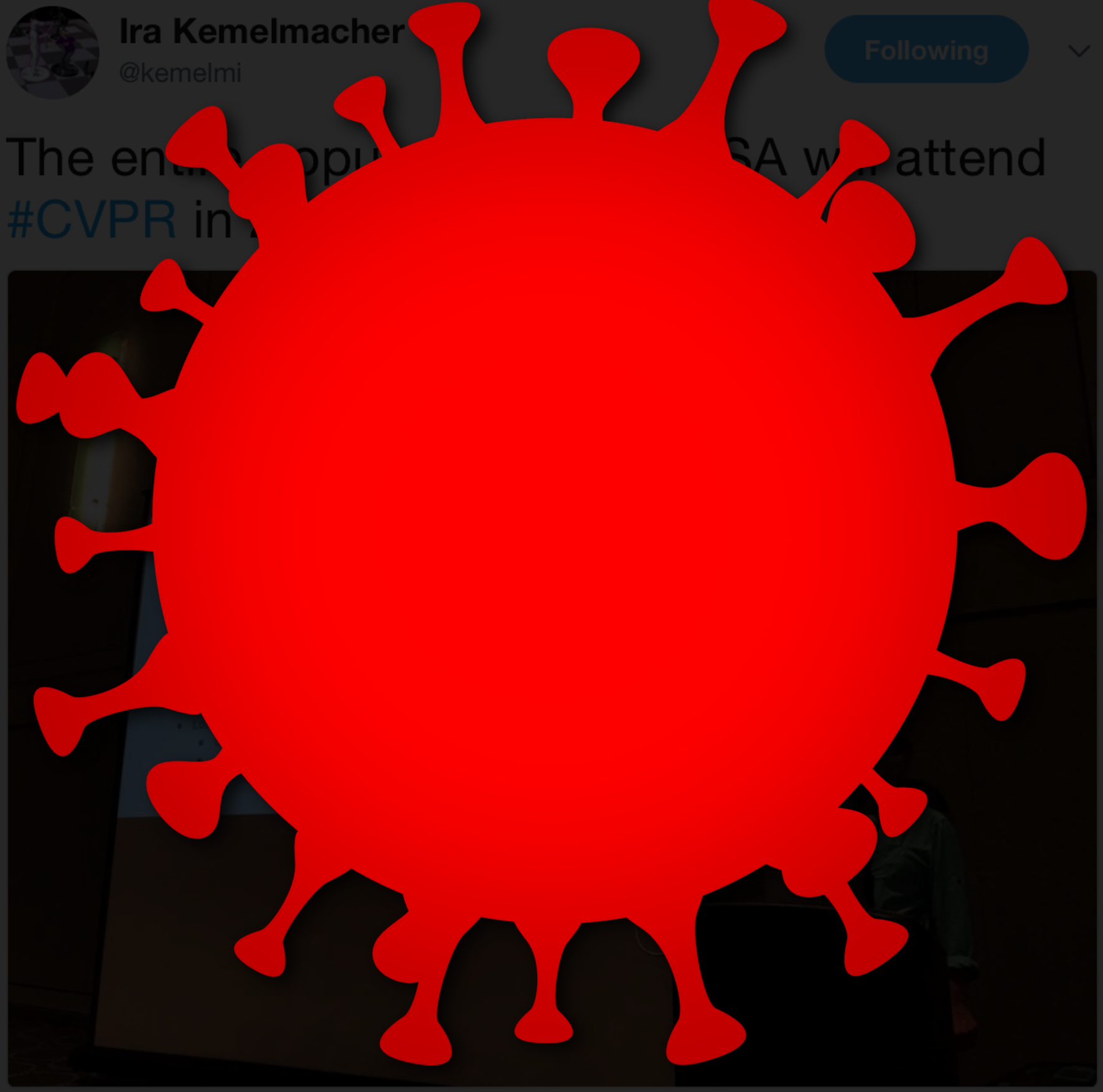


Ira Kemelmacher
@kemelmi

Following

The entire population of the USA will attend #CVPR in 2063 :-)




Ira Kemelmacher
@kemelmi
Following
#CVPR

Welcome
to
CVPR 2020

Deep Learning for Geometric Computing

CVPR 2020 Workshop and Challenge

Novel View Synthesis: From Depth-Based Warping to Multi-Plane Images and Beyond

How to write a good review?

ACTIVITYNET

Large Scale Activity Recognition Challenge 2020

Introduction

Google

Microsoft

facebook

SIEMENS

Baidu 百度

JPL

Jet Propulsion Laboratory

amazon.com

iRobot

Adobe

SRI International

IBM

xerox

MITSUBISHI ELECTRIC

intel

SAMSUNG

GE

Computer vision industry will grow from $1.1 billion in 2016 to $26.2 billion by 2025

Source: Tractica (2020)

156 bpm

Road ahead

Foundational analyses of visual information processing

Foundational analyses of visual information processing

Algorithms exploiting these analyses

Topics

Image formation

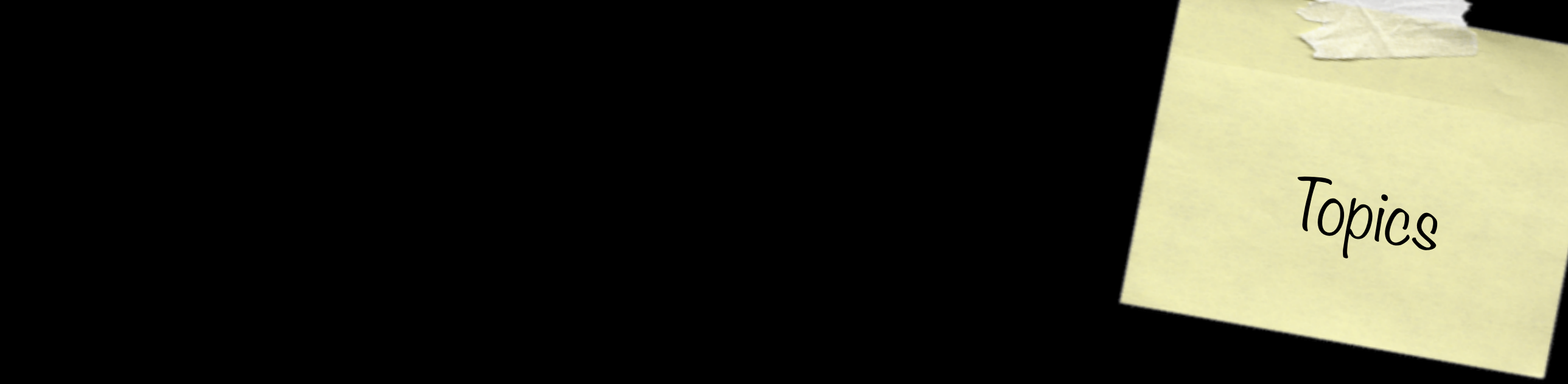


Image formation

Image representation

Image formation

Image representation

Feature detection

Image formation

Image representation

Feature detection

Stereopsis

Image formation

Image representation

Feature detection

Stereopsis

Motion analysis

textbook

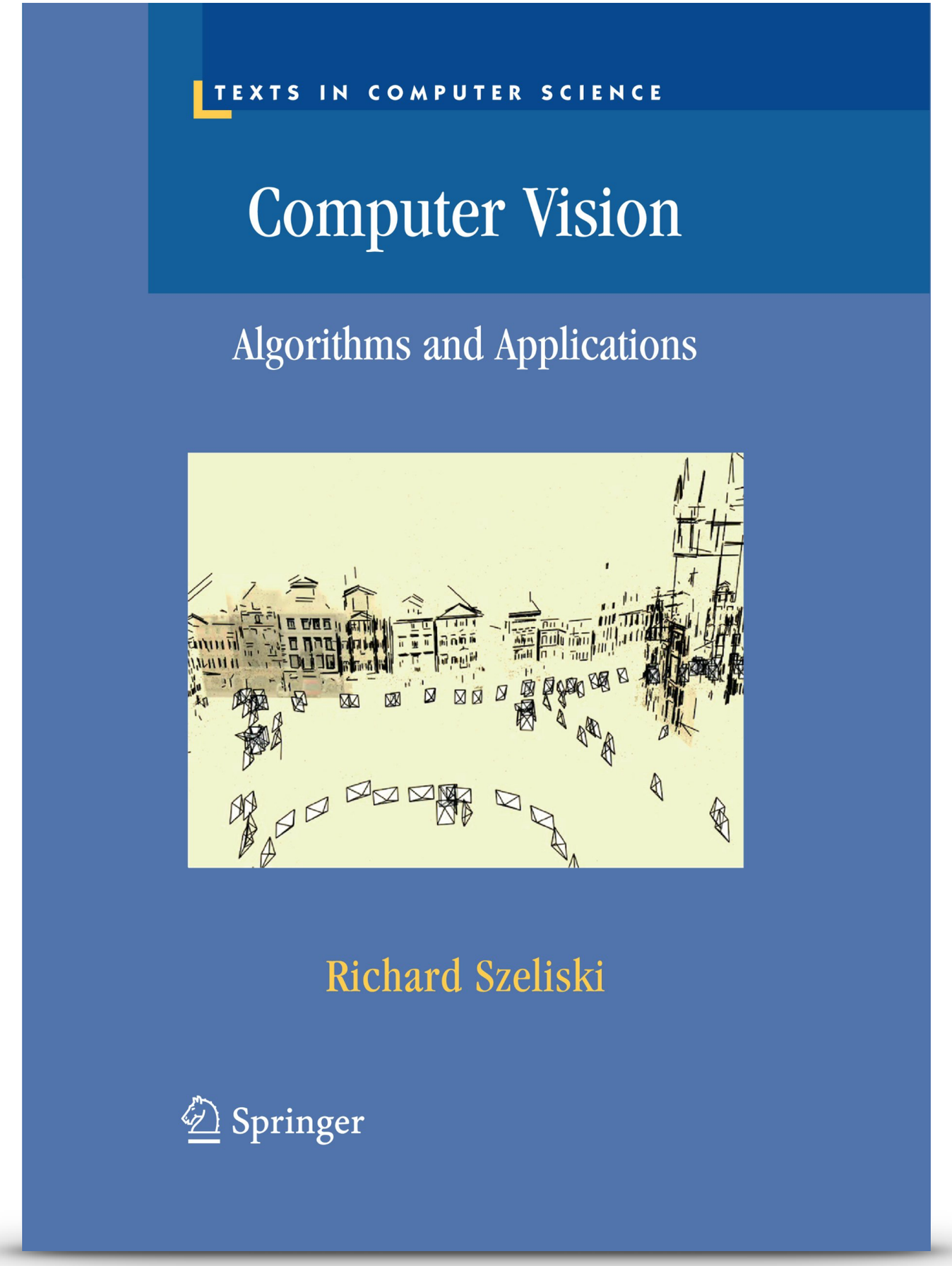
TEXTS IN COMPUTER SCIENCE

Computer Vision

Algorithms and Applications

Richard Szeliski

Springer



available online for free

szeliski.org/Book