Computer Vision

CSE P576 Autumn 2021

Vitaly Ablavsky

These slides were developed by Dr. Matthew Brown for CSEP576 Spring 2020 and adapted (slightly) for Fall 2021

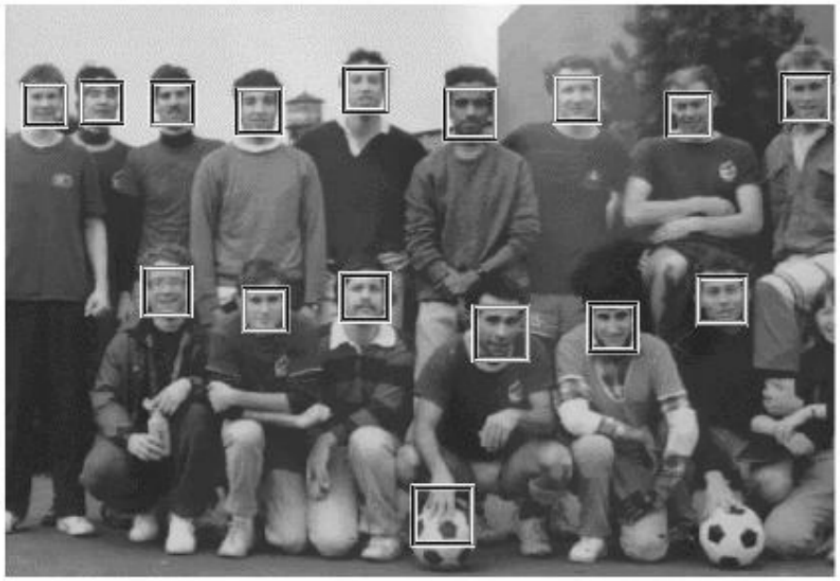
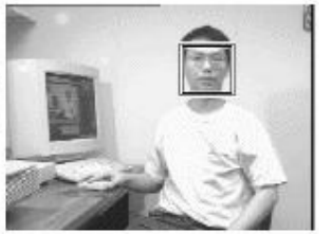
credit → Matt

blame → Vitaly

The Course

- People
  - Vitaly Ablavsky
  - TAs: Kalyani Marathe, Svetoslav Kolev
- Time and location
  - Lectures: Wednesdays 6:30-9:20pm
  - Office hours: Zoom, TBD
- Evaluation
  - 4 projects, equally weighted
- Resources
  - https://courses.cs.washington.edu/courses/csep576/21au/
  - Ed Discussion=Discussion board, Canvas=Assignments
  - Book 1: "Computer Vision", Szeliski,
  - Book 2: "Deep Learning", Goodfellow et al.
  - Stanford CS231N (CNNs for Vision)

Face Detection
(the Early Aughts Version)

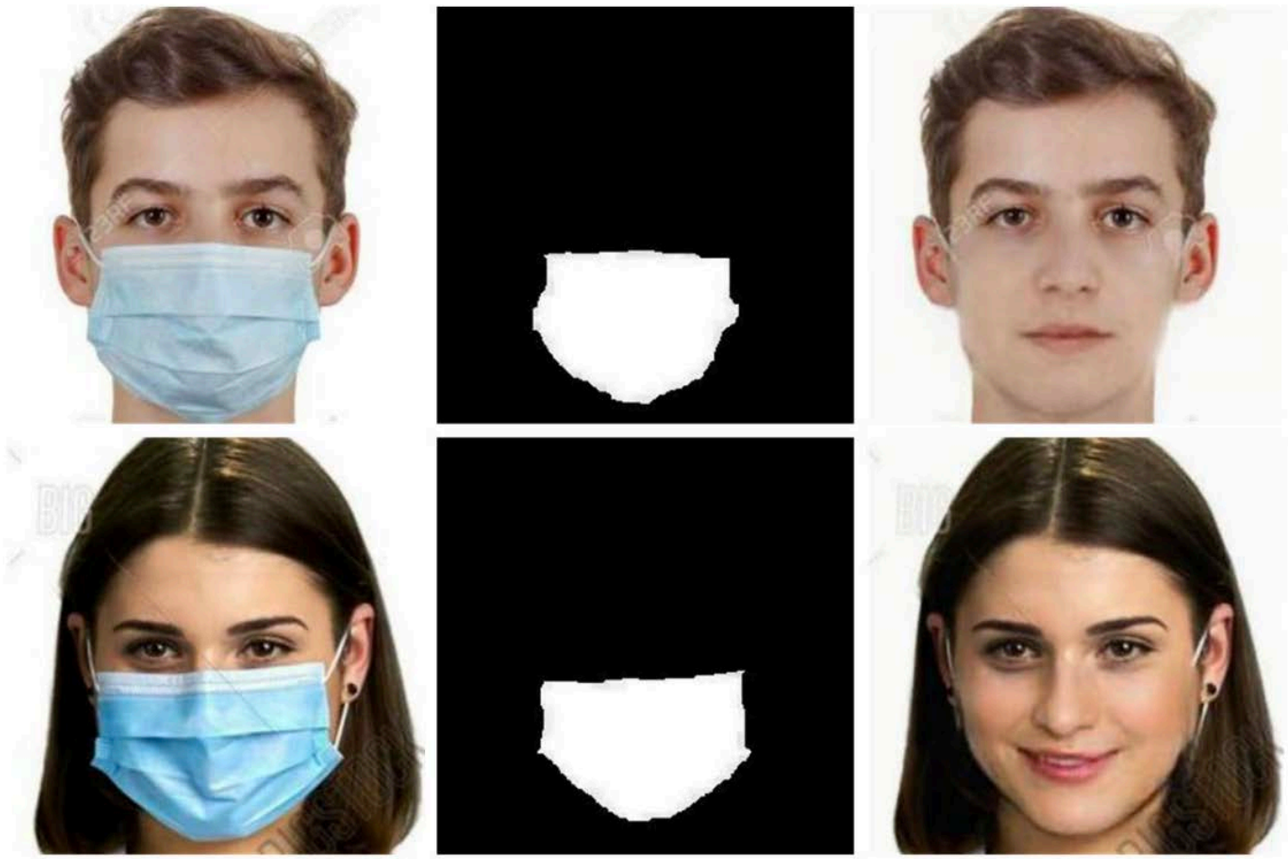




[Viola and Jones, CVPR 2001]

Tracking Animals

[Branson et al. CVPR 2005]

Virtual Un-masking (literally)



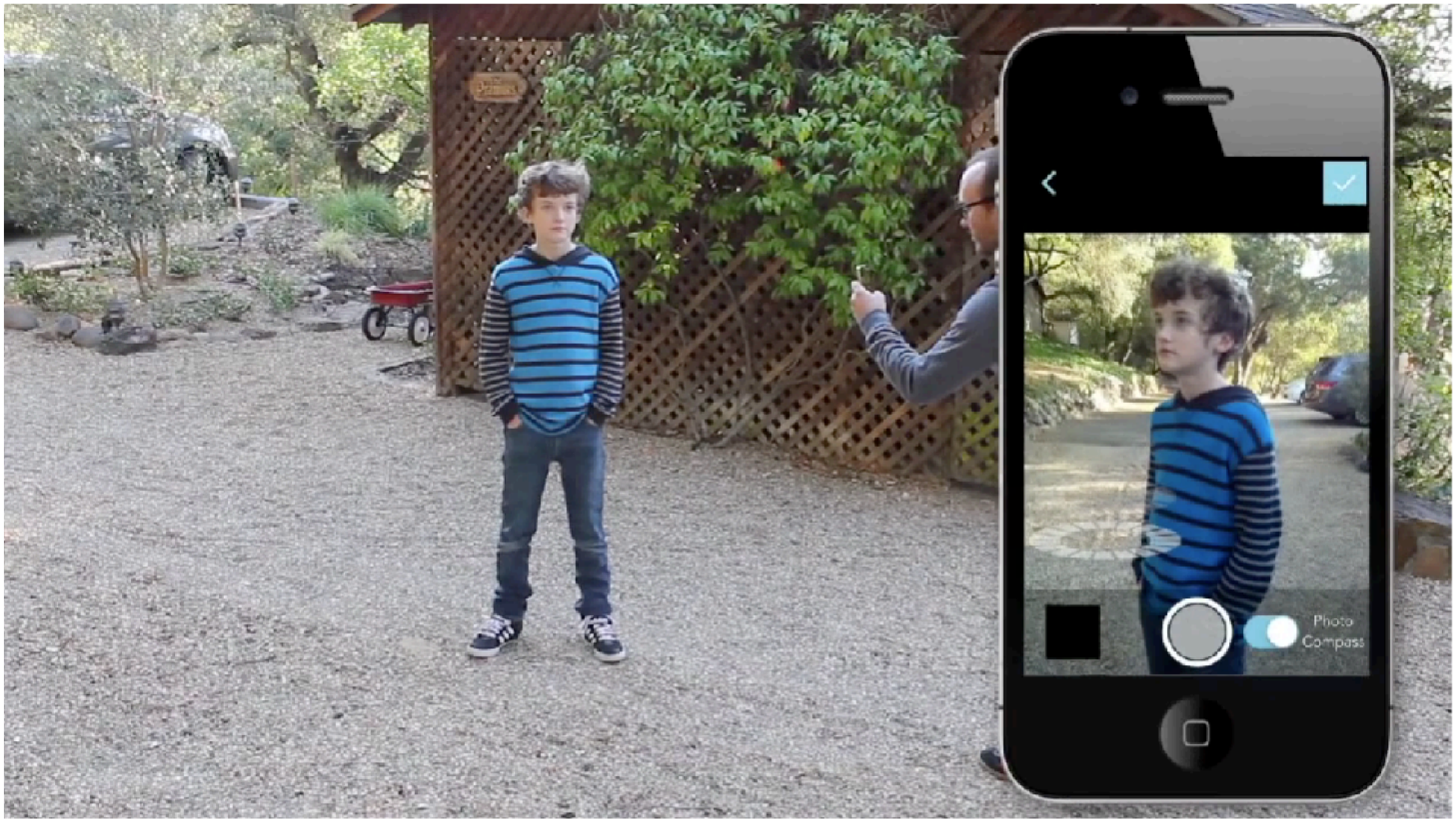
"Trend of wearing masks in public is growing in recent years all over the world. First, people wear masks to guard themselves from pollution. Second, some people are self-conscious about their look and they want to hide their face and emotions from the public. Removing the mask object that covers almost half of the face might be of help in guessing one's identity."

[Nizam Ud Din et al., A Novel GAN-Based Network for Unmasking of Masked Face, IEEE Access Magazine, March 2020]

Face Detection

[Motorola]

3D Reconstruction

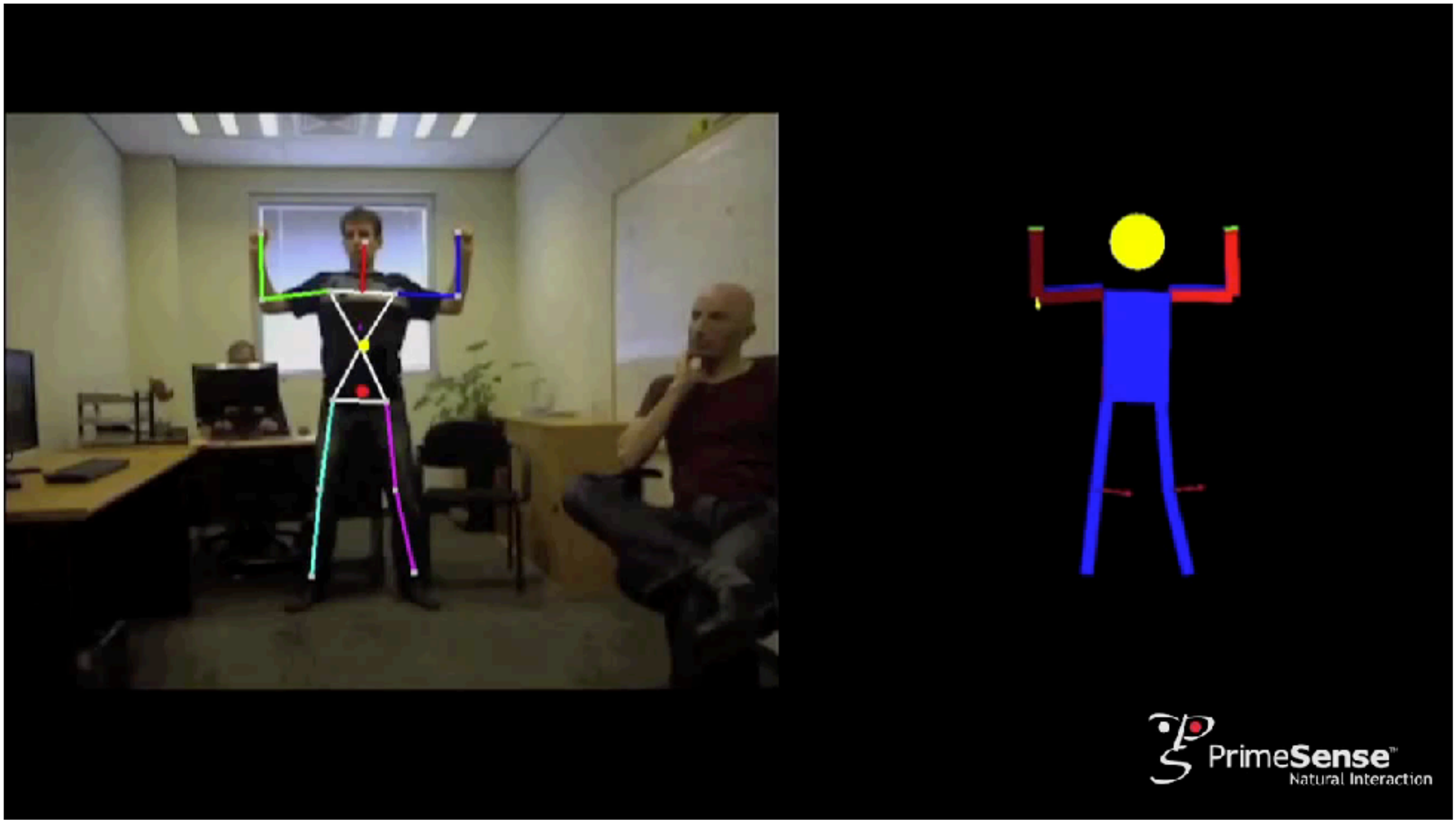


[Autodesk 123D Catch]

Body Pose Tracking

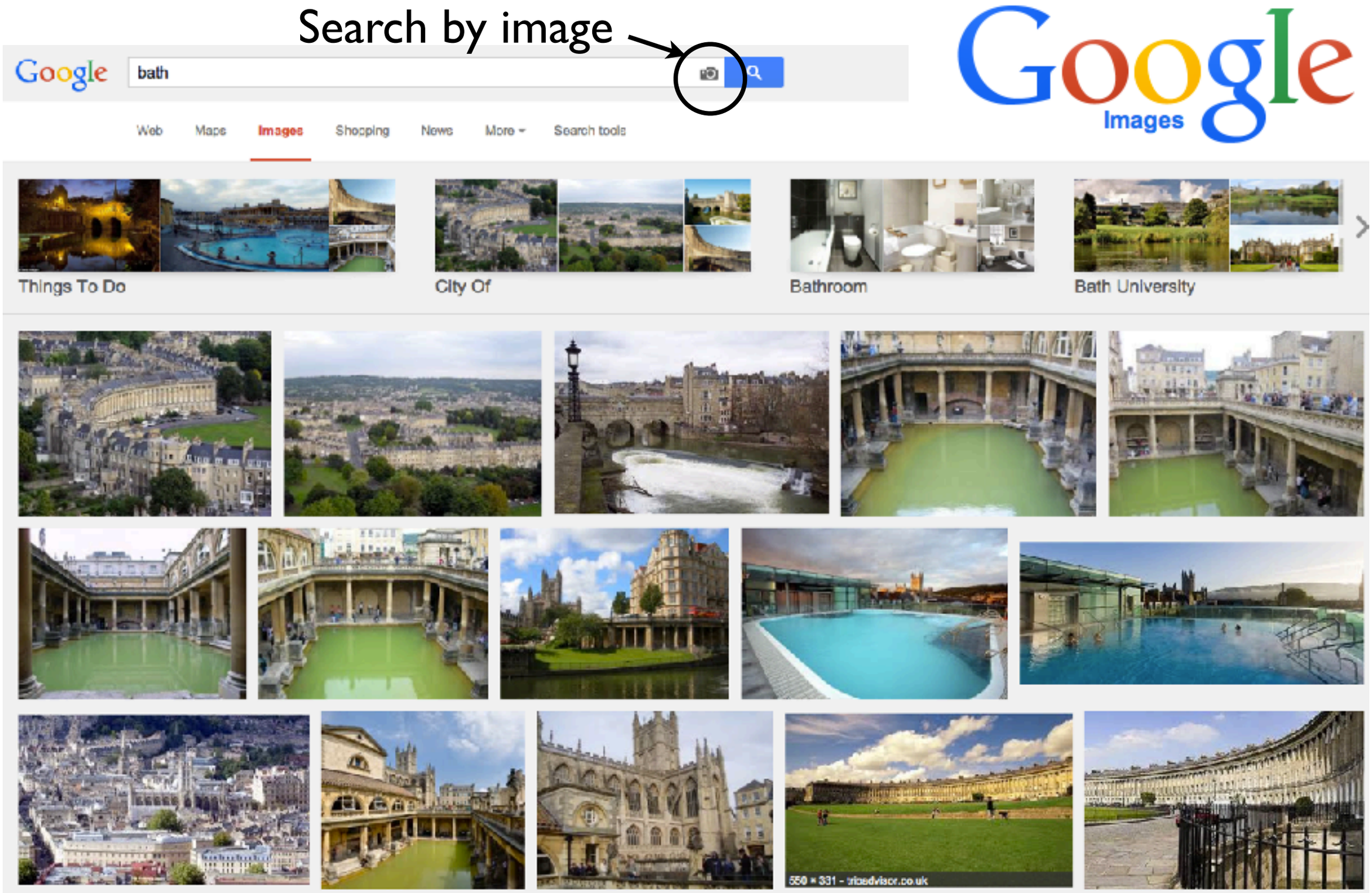
[Microsoft Xbox Kinect]

Body Pose Tracking



[PrimeSense]

Image Recognition and Search



Self Driving Cars

[Google]

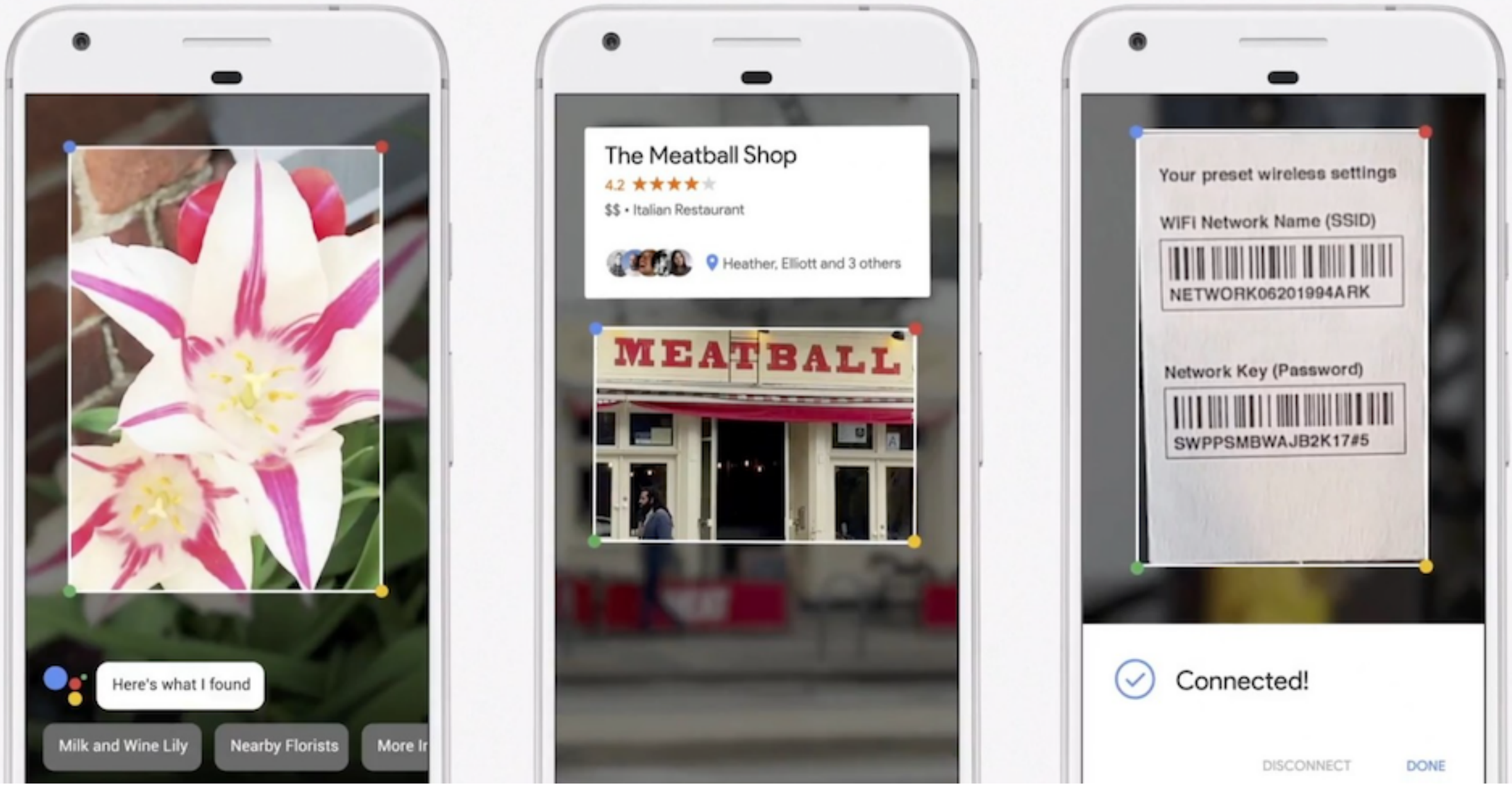
Flying Vehicles

www.skydio.com

AR/VR

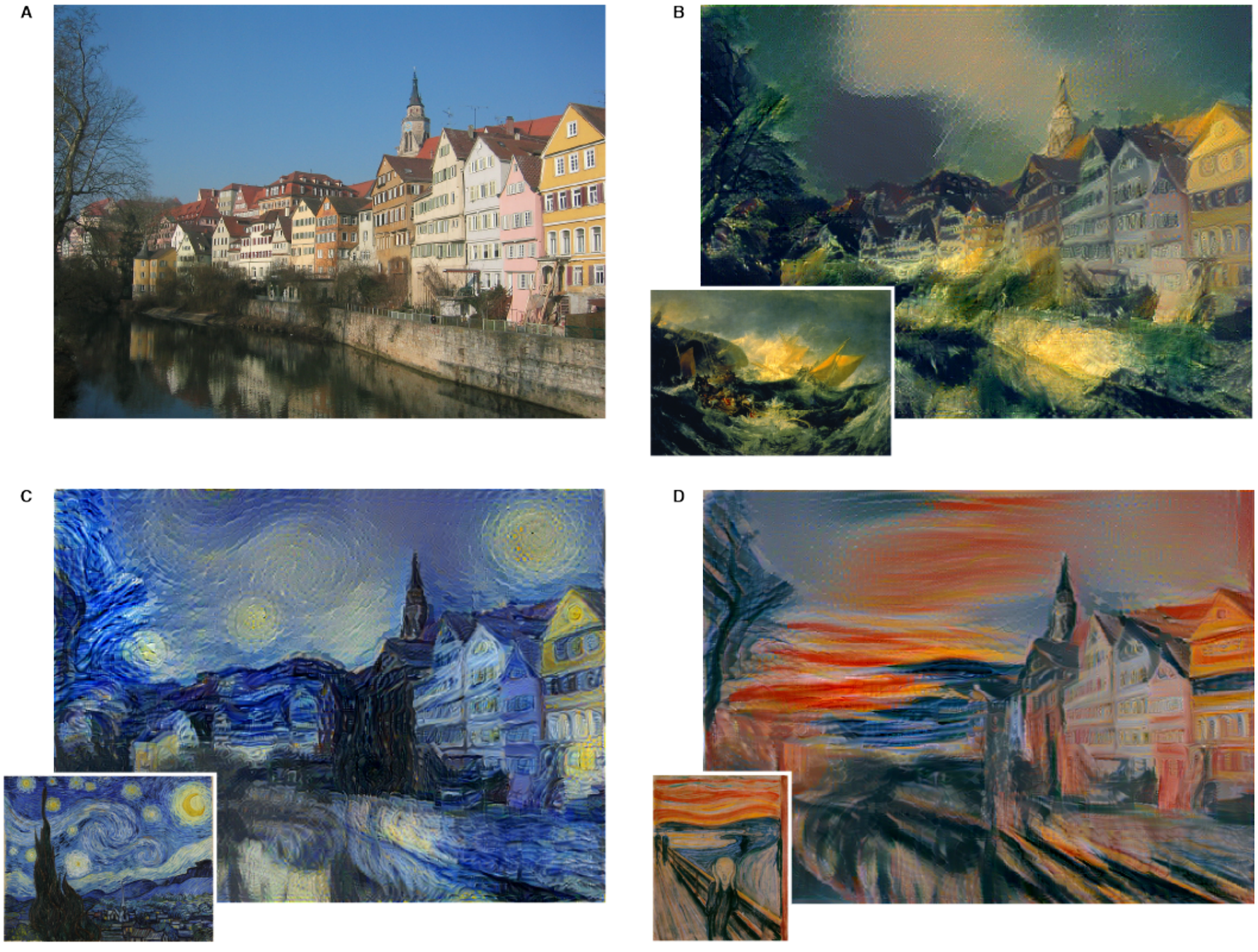
[Microsoft Hololens]

Mobile Apps



[Google Lens]

Art



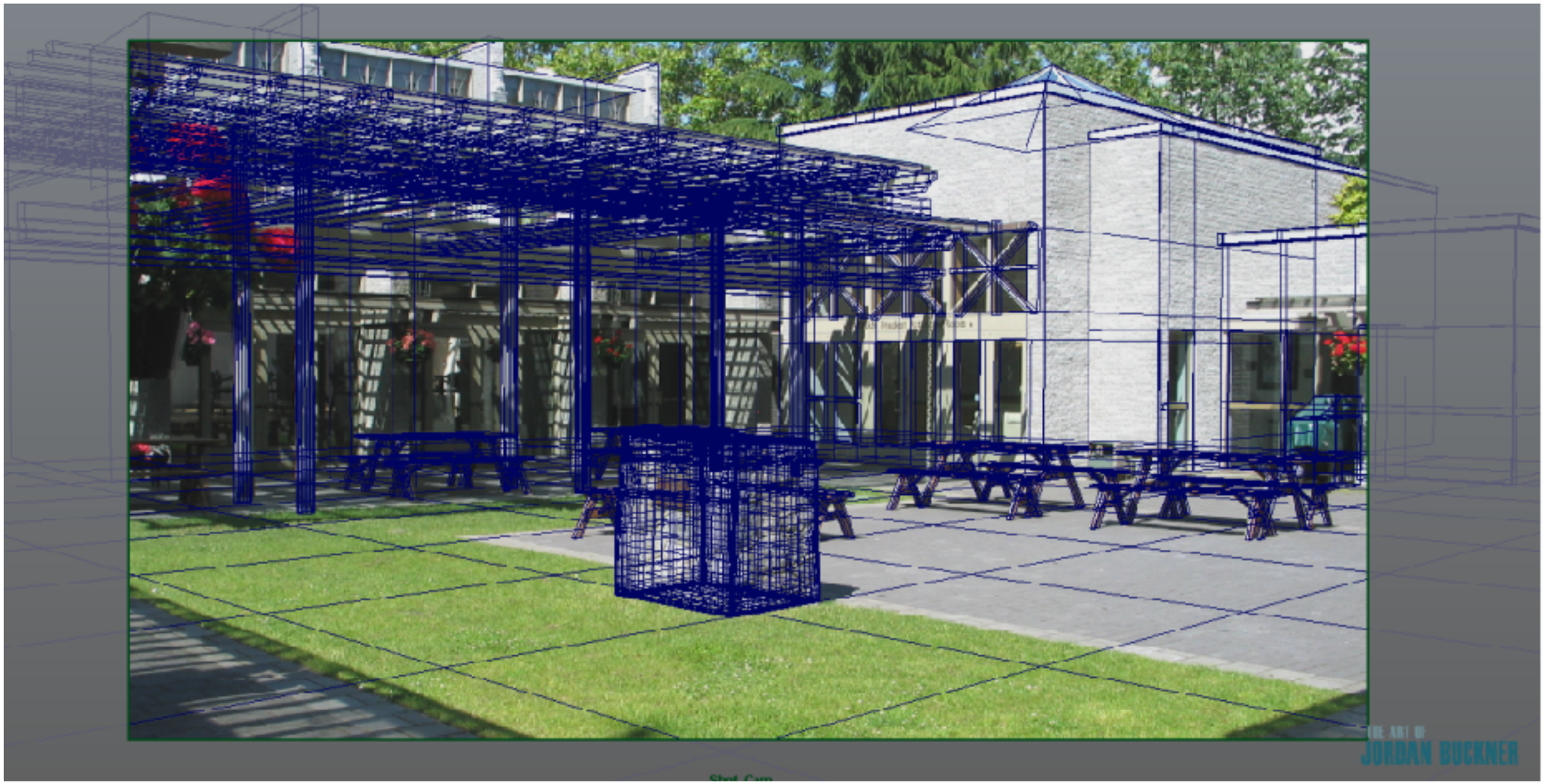
[Gatys, Ecker, Bethge 2015]

Applications of Computer Vision

- Digital Entertainment + Consumer
  - Camera tracking, 3D reconstruction, visual effects, virtual reality, augmented reality, product recognition
- Science and Medicine
  - Visual data analytics, anatomical measurement/analysis, tumour detection
- Engineering and Industry
  - Robotics, self driving cars, reverse engineering, visual servoing, industrial part inspection, OCR, precision agriculture
- Photography/Videography and Editing
  - Face detection, scene recognition, video stabilisation, drone camera, gap filling, image blending, panorama stitching, high dynamic range
- Mapping and Environmental
  - Image registration, 3D building modelling, streetview, numberplate recognition, landmark recognition, species identification

Definitions of Computer Vision #1

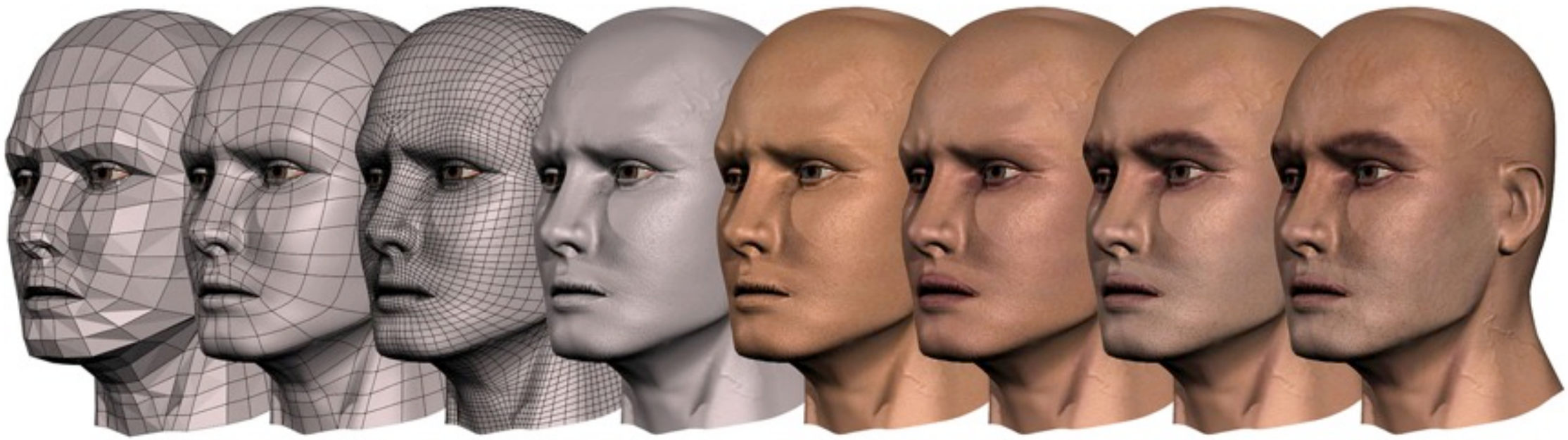
“Inverse Computer Graphics”



[J. Buckner]

Definitions of Computer Vision #1

"Inverse Computer Graphics"

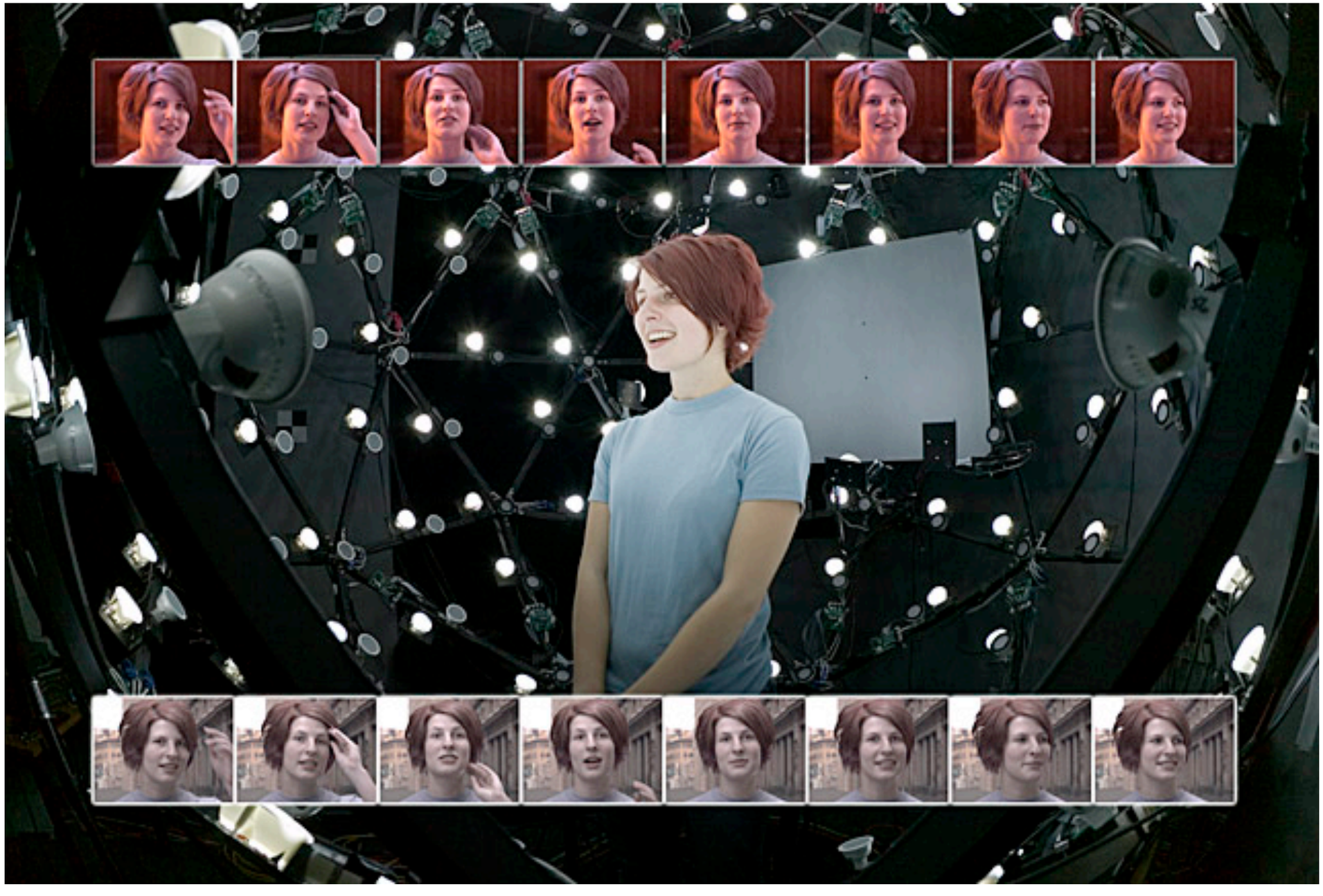


Graphics

Vision

Photometric Capture

- Capture reflectance as well as geometry ("Light Stage")

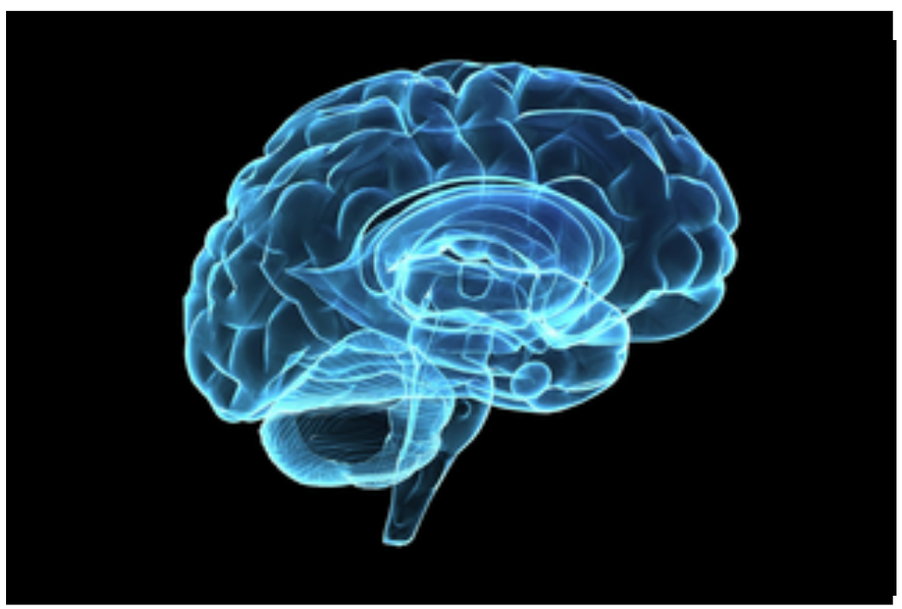


[Paul Debevec, USC]

Definitions of Computer Vision #2

“Replicate Human Vision”

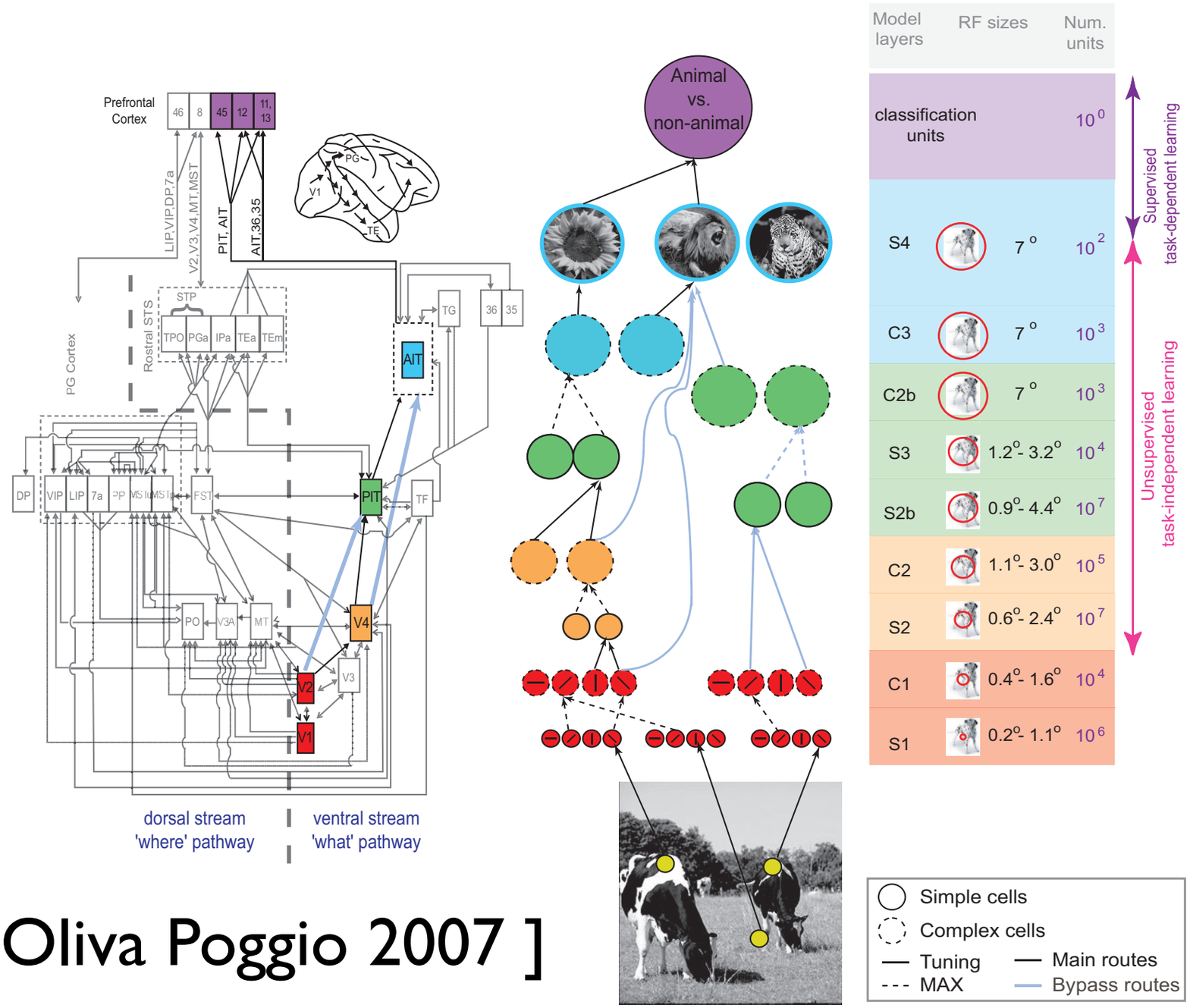
=



=

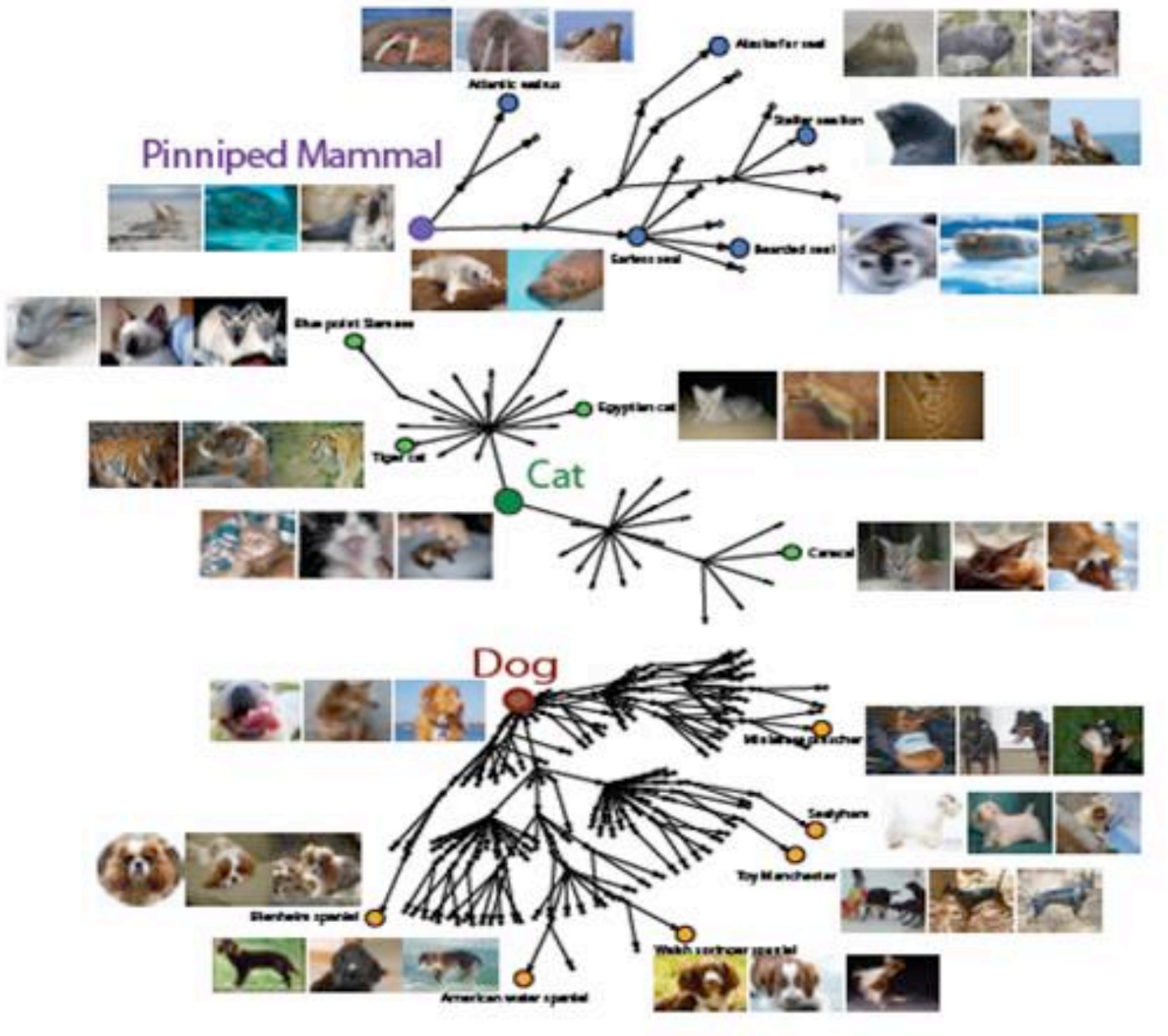
Definitions of Computer Vision #2

“Replicate Human Vision”



[Serre Oliva Poggio 2007]

ImageNet



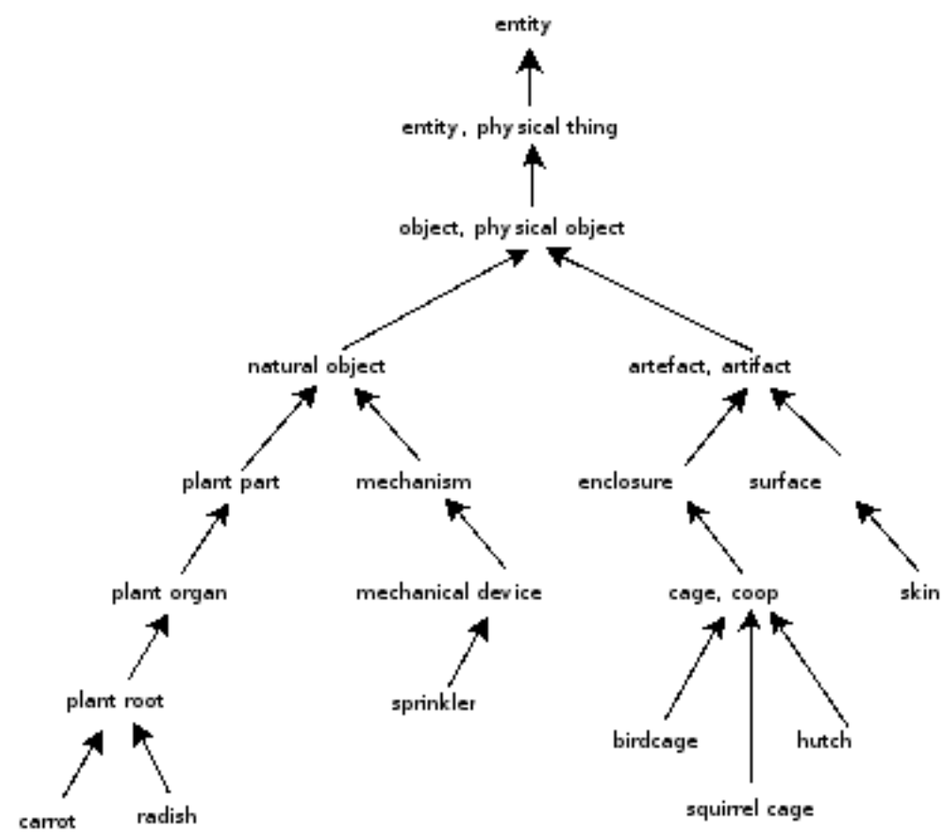


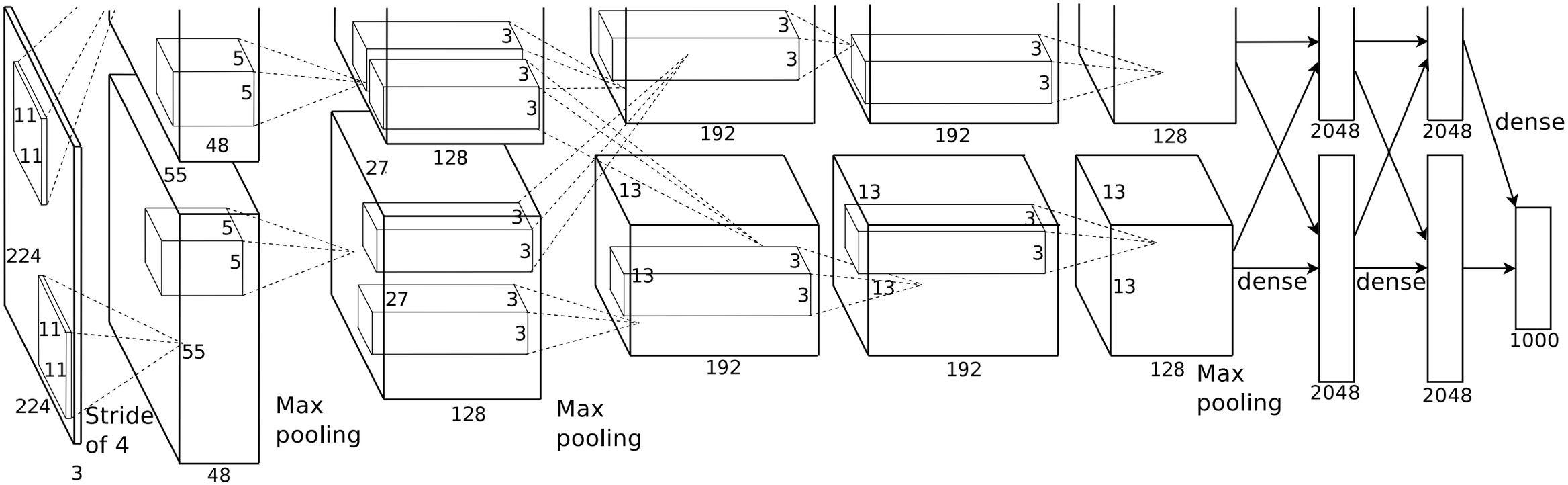
Figure 1. "is a" relation example

15 million images in 22,000 categories

[F. F. Li et al., 2009]

ImageNet Classification via CNN

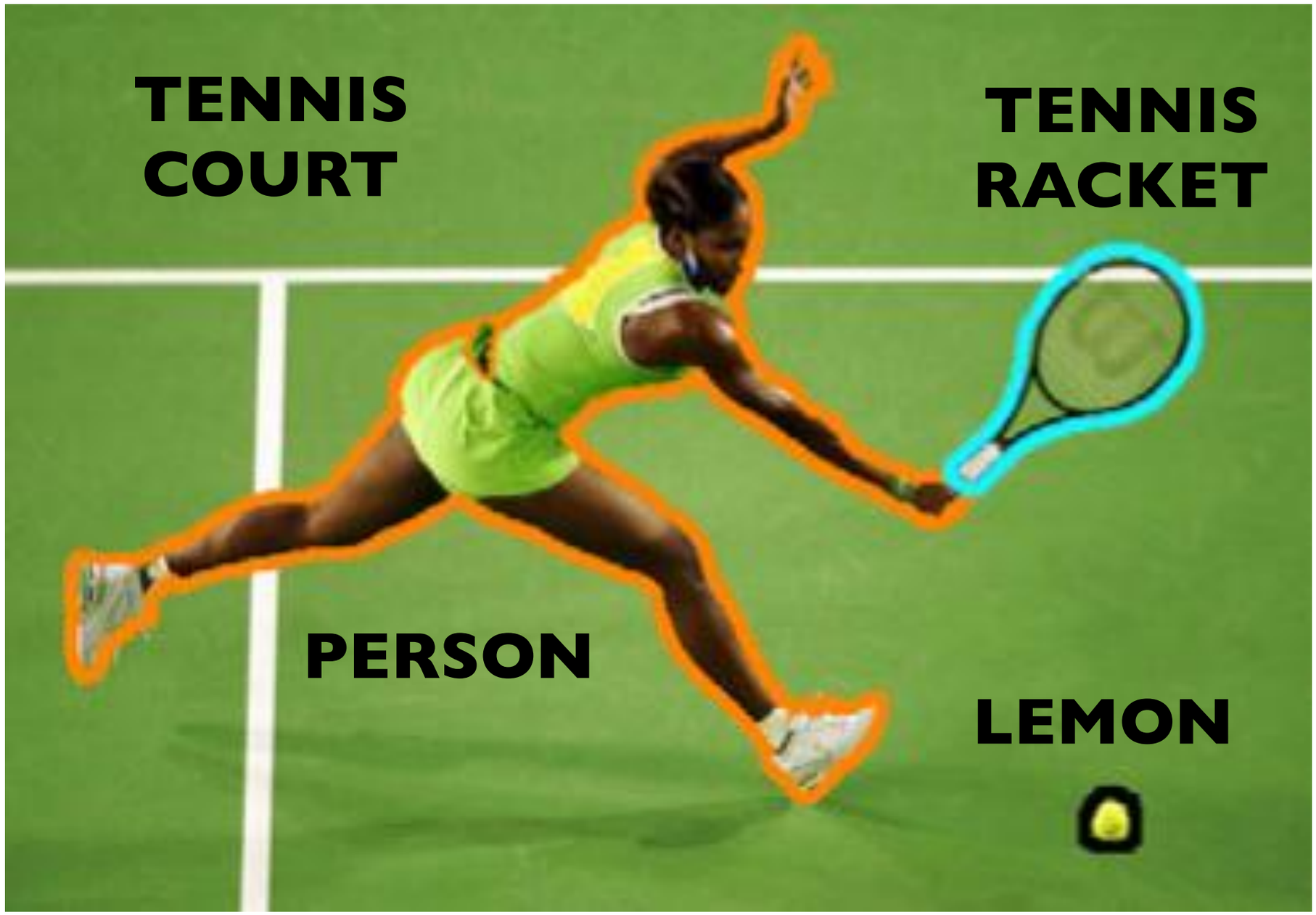
- "Alexnet" gave breakthrough results on the ImageNet 2012 Large Scale Visual Recognition Challenge (ILSVRC 2012)



[Krizhevsky, Sutskever, Hinton 2012]

Definitions of Computer Vision #3

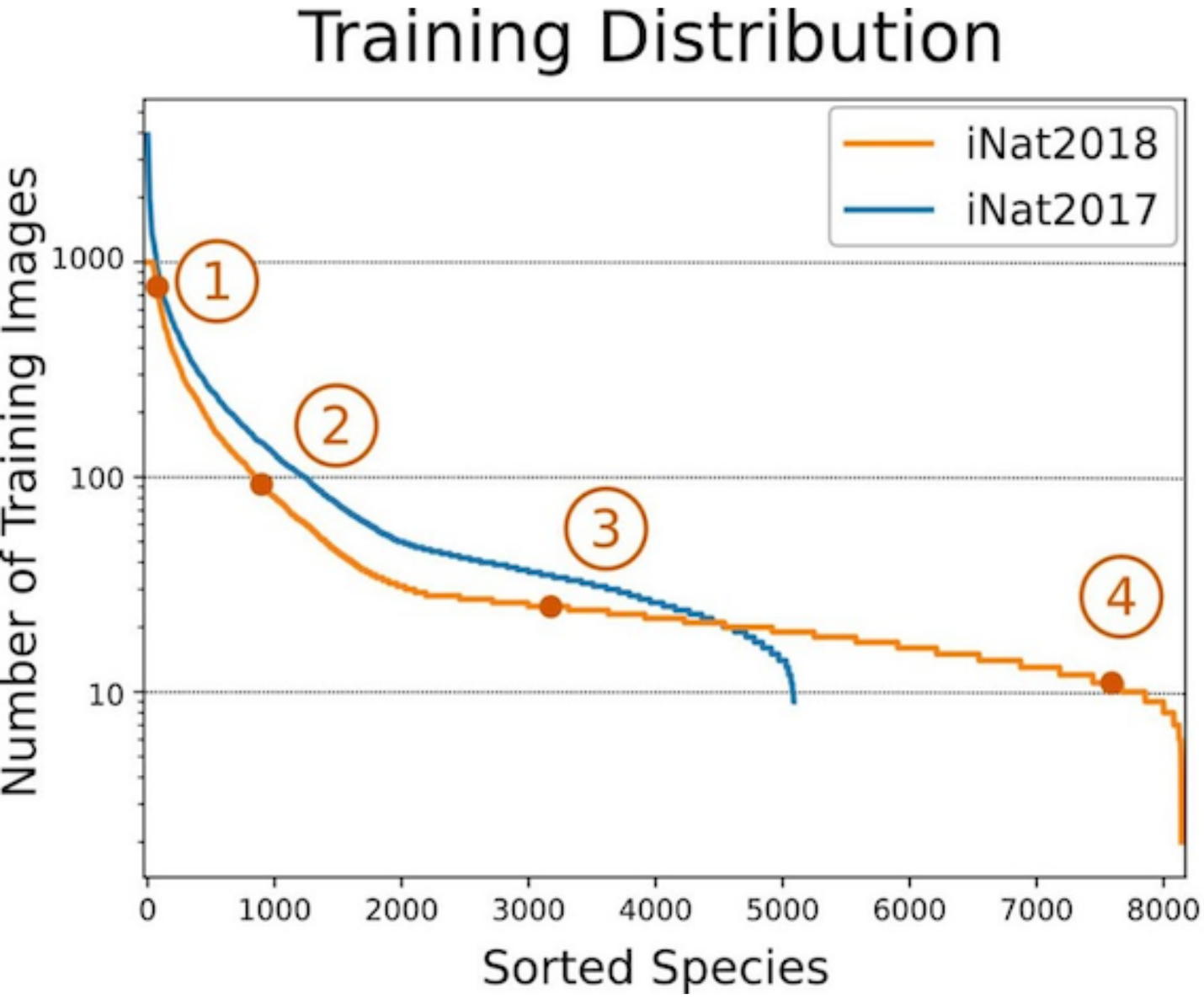
"Image/Video Understanding"



[Rabinovich, Galleguillos, Wiewiora, Belongie 2007]

Definitions of Computer Vision #3

“Image/Video Understanding”



iNaturalist Challenge 2018

1 Cooper's Hawk

2 American Bison

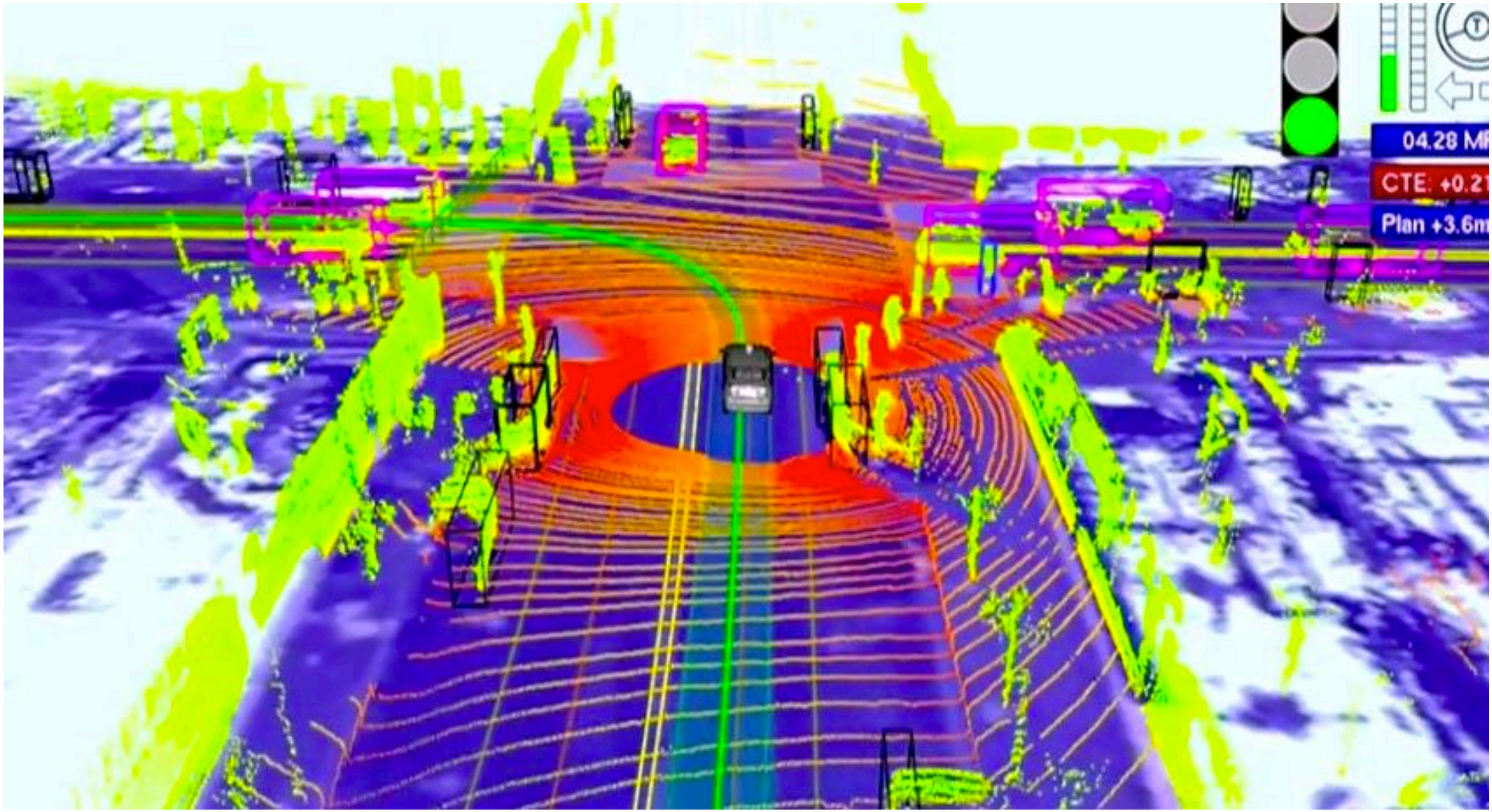
3 Mallow Bindweed

4 Island Fox

[Van Horn, Mac Aodha]

Definitions of Computer Vision #3

"Image/Video Understanding"



[Thrun Urmson Google]

This Course

- Computer Vision, with emphasis on **visual geometry + learning** (roughly 50-50 split between the two)
- 10 lectures, + office hours
- 4 projects, equally weighted
- **Project 1**: Feature Extraction and Matching
- **Project 2**: Panoramic Image Stitching
- **Project 3**: Image Classification using CIFAR10
- **Project 4**: Pose Estimation
- Projects will use iPython notebooks (e.g., Jupyter, Colab)
- Numpy for numerics
- PyTorch/Tensorflow for machine learning

Schedule

Date	Lecture	Description	Notes and Resources
9/29	Introduction		[CVA2] Ch.1
	Image Formation	Geometric and Photometric Image Formation, Pinhole Camera, Lenses, Sensors, Colour, Gamma, DCT, Image Coding	[CVA2] Ch.2
10/6	Filtering and Pyramids	Linear + Non-Linear Filtering, Correlation, Convolution, Gaussian + Laplacian Pyramids, Sampling and Aliasing	**Project 1 start**
	Features and Matching	Detection, Correspondence, Edges, Corners, Regions, Patch Matching, SIFT, Shape Context, Learning Features	
10/13	Planar Geometry	2D Transforms: Euclidean, Similarity, Affine, Projective, Camera Models: Perspective, Projective, Linear, Viewing planes, Lines and Camera Rotation	
	RANSAC	Least Squares 2-view Alignment, Outliers, Robust Line Fitting, RANSAC, Minimal Subsets	
10/20	Epipolar Geometry	Epipolar Lines, Plane Constraint, Fundamental/Essential Matrix, 8 point algorithm, Triangulation, 2-view SFM	**Project 2 start**
	Multiview Alignment and SFM	Multiview Alignment, Residuals, Error Function, Structure from Motion, Bundle Adjustment, Pose Estimation, Triangulation	
10/25	**Project 1 due**		

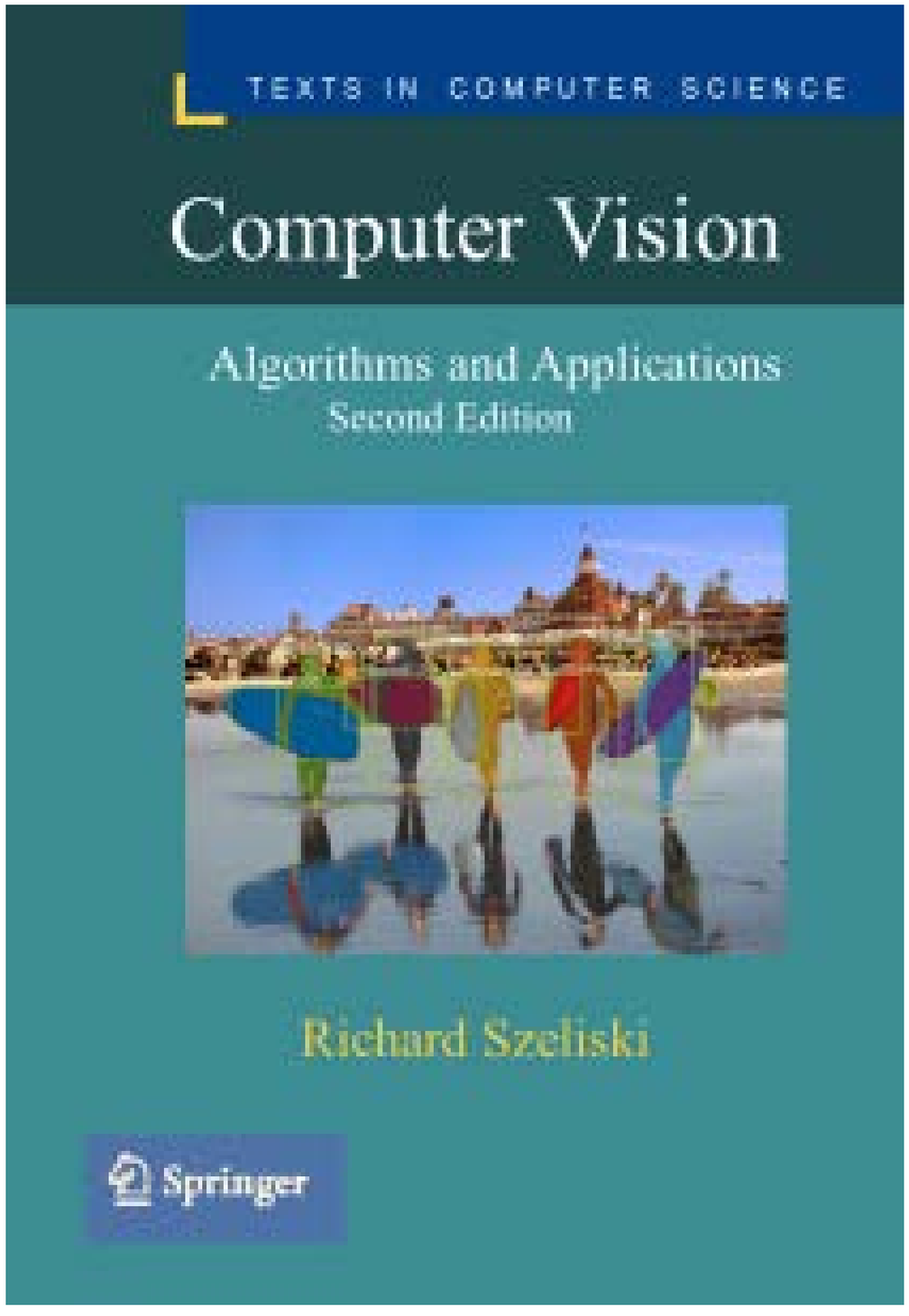
Schedule

10/27	Stereo	Stereo matching, local + global, multiview stereo, plane sweep, volumetric, depth map merging, photometric stereo	
	Depth + Flow	Depth imaging + fusion, signed distance functions, non-rigid matching, optical flow, Lucas Kanade algorithm	PlaneSweep ipynb, LucasKanade ipynb. Notebooks by Steven Lovegrove, Richard Newcombe
11/3	Linear Classification	Visual classification intro, object recognition, instance, category, classification vs detection, linear classification, 2-class, N-class, linear and softmax regression	**Project 3 start**
	Visual Classification 2	Fundamentals and Pre-Deep Learning Classification, Bayesian classifiers, Gaussian distributions, PCA, LDA, Decision Forests, Visual words, SVMs	
11/8	**Project 2 due**		
11/10	Neural Networks	Feature extraction, end to end learning, multiple linear layers, activation functions, biological neurons, space warping, universal approximation, convex optimization	
	Backpropagation	Chain rule, computational gradients, forward/reverse mode autodiff, upstream/local gradients, flat backprop, modular design, scalar/vector /tensor backprop, matrix multiplication example	
	Convolutional Networks	Convolutional layers, activation maps, dimension mappings, receptive fields, strides, pooling, LeNet5 example	

Schedule

11/17	Advanced CNNs	CNN building blocks, dropout, batch norm, factorized convolutions, residual connections, popular architectures: AlexNet, VGG, GoogLeNet, Resnet, MobileNet, SE-Net	**Project 4 start**
	Object Detection	Motivation + applications, sliding windows, anchor based detection, single-stage and two-stage architectures, evaluation metrics, IoU, precision-recall, mAP, practical tips	
11/22	**Project 3 due**		
11/24	Segmentation	Dense prediction, semantic, instance, panoptic segmentation, keypoint estimation, fully convolutional nets, atrous, transpose convolution	
	Single-View Depth, Superres, Colorization	Pixel labelling, single-view depth estimation, direct, self-supervision, super-resolution, colorization, image translation	
12/1	Deep Learning in 3D	Single-view, 2-view, multi-view depth, deep learning with points, meshes, voxels, SDFs, neural scene representation and rendering	
	Image Generation and GANs	Loss functions: L2, VGG, adversarial, texture synthesis, style transfer, generative adversarial nets, image generation, conditional GANs, image translation, pix2pix	
12/8 class???	**Project 4 due**		

Recommended Text 1



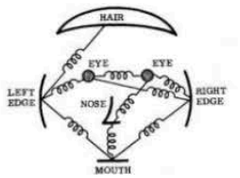
Computer Vision: Algorithms and Applications, 2nd edition

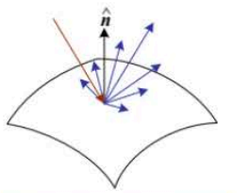
Richard Szeliski

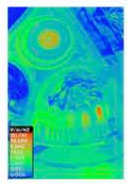
http://szeliski.org/Book

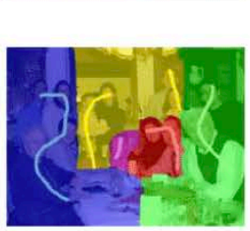
Core textbook for the course. Good coverage of most topics, oriented around practical applications

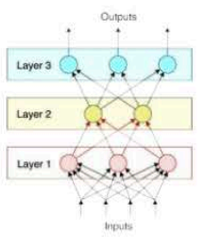
Computer Vision: Szeliski, 2nd edition

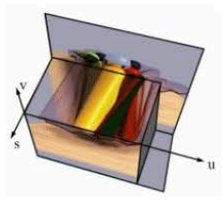




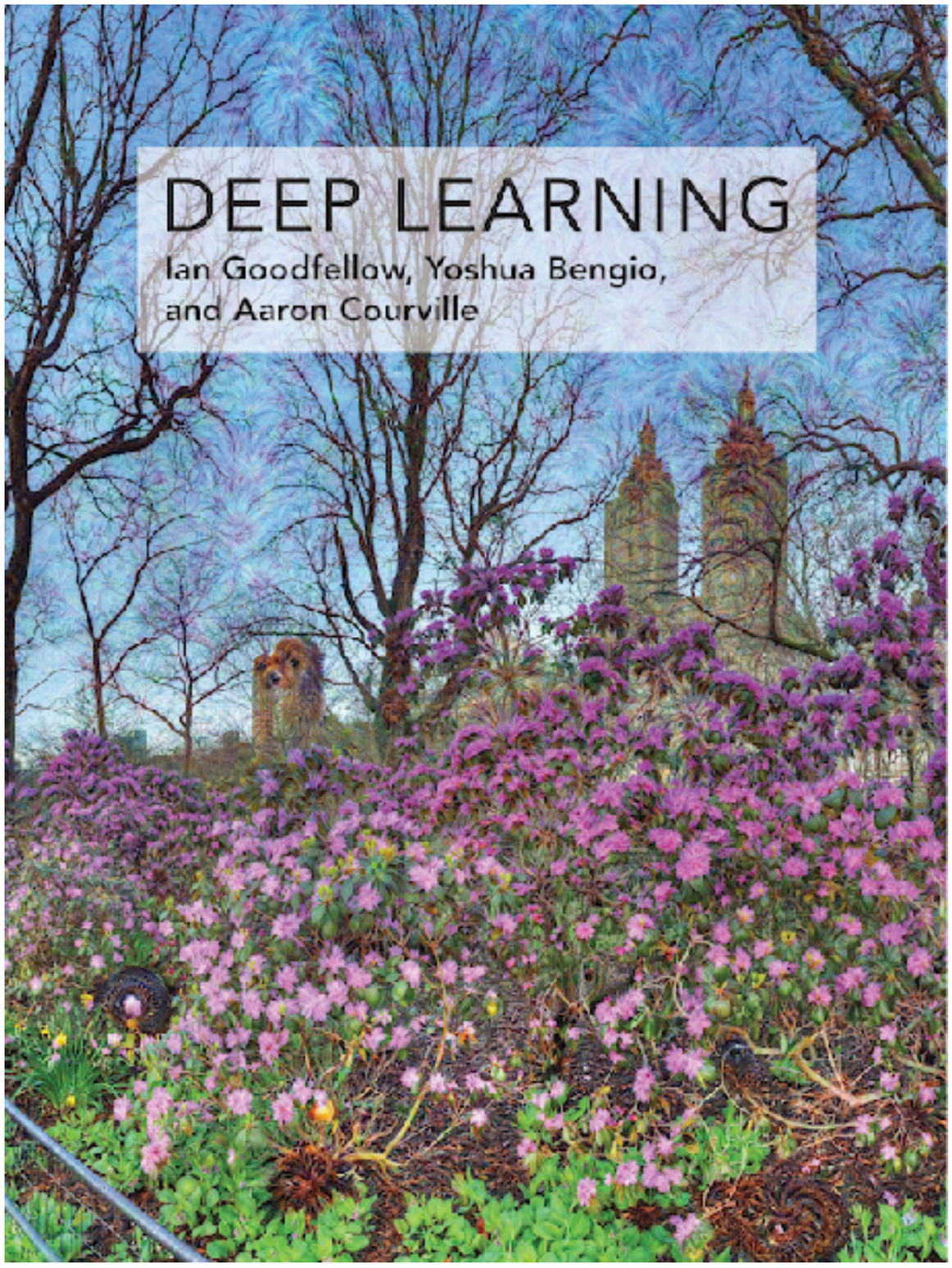








Recommended Text 2



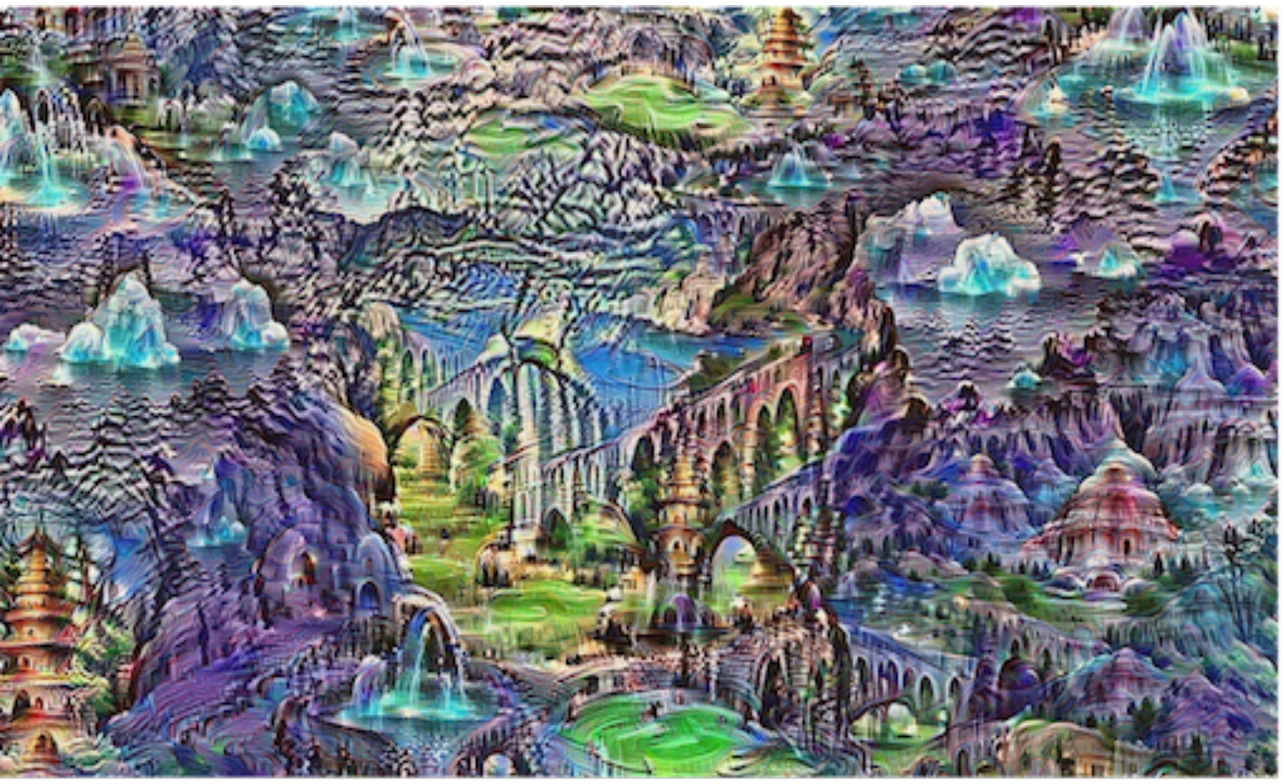
Deep Learning: Goodfellow, Bengio, Courville

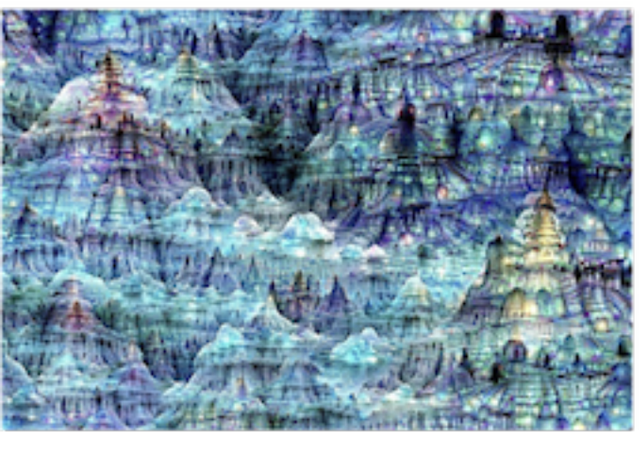
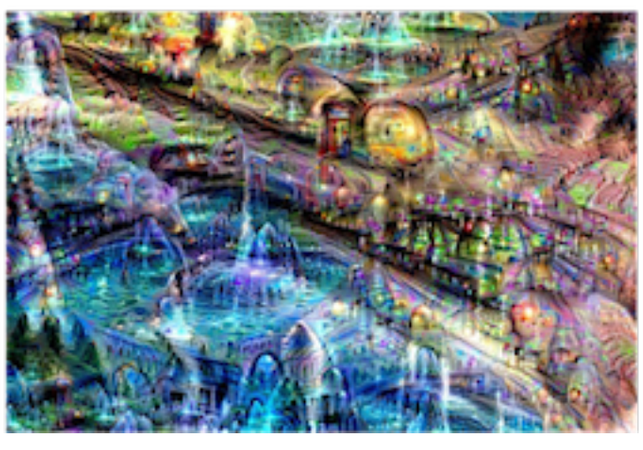
deeplearningbook.org

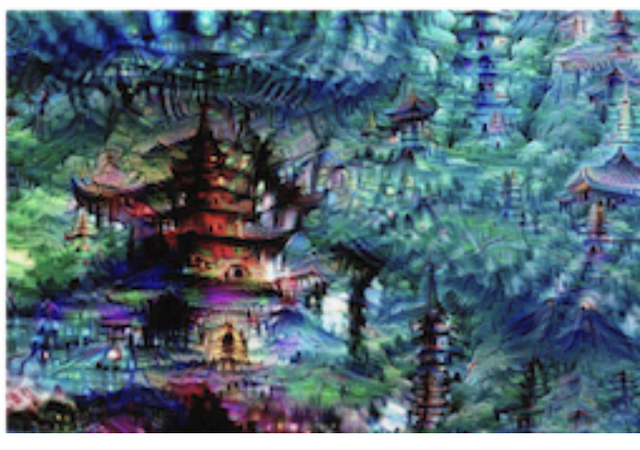
Background maths + probability, practical deep nets, deep learning research

Also cs231n.stanford.edu
— CNNs for Vision

Inceptionism





[Mordvintsev, Olah, Tyka 2015]

Next Lecture

- Cameras + Image Formation

Try getting Jupyter/Colab up and running,
and work through Justin Johnson's Python intro:
http://cs231n.github.io/python-numpy-tutorial