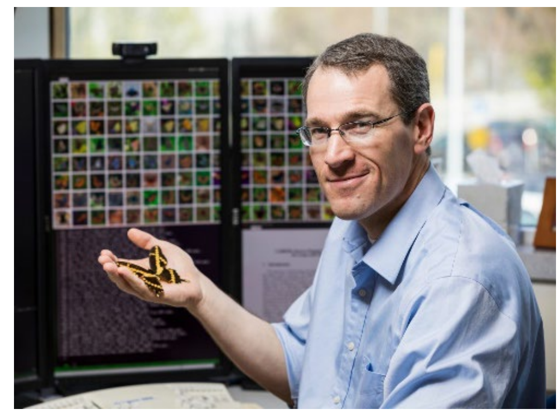
Computer Vision:
Understanding, Interpreting and Learning from Visual Data

DSSI
DATA SCIENCE
SUMMER INSTITUTE

Lawrence Livermore
National Laboratory

DAY ONE

July 25, 2019
July 30, 2019
July 31, 2019
August 1, 2019

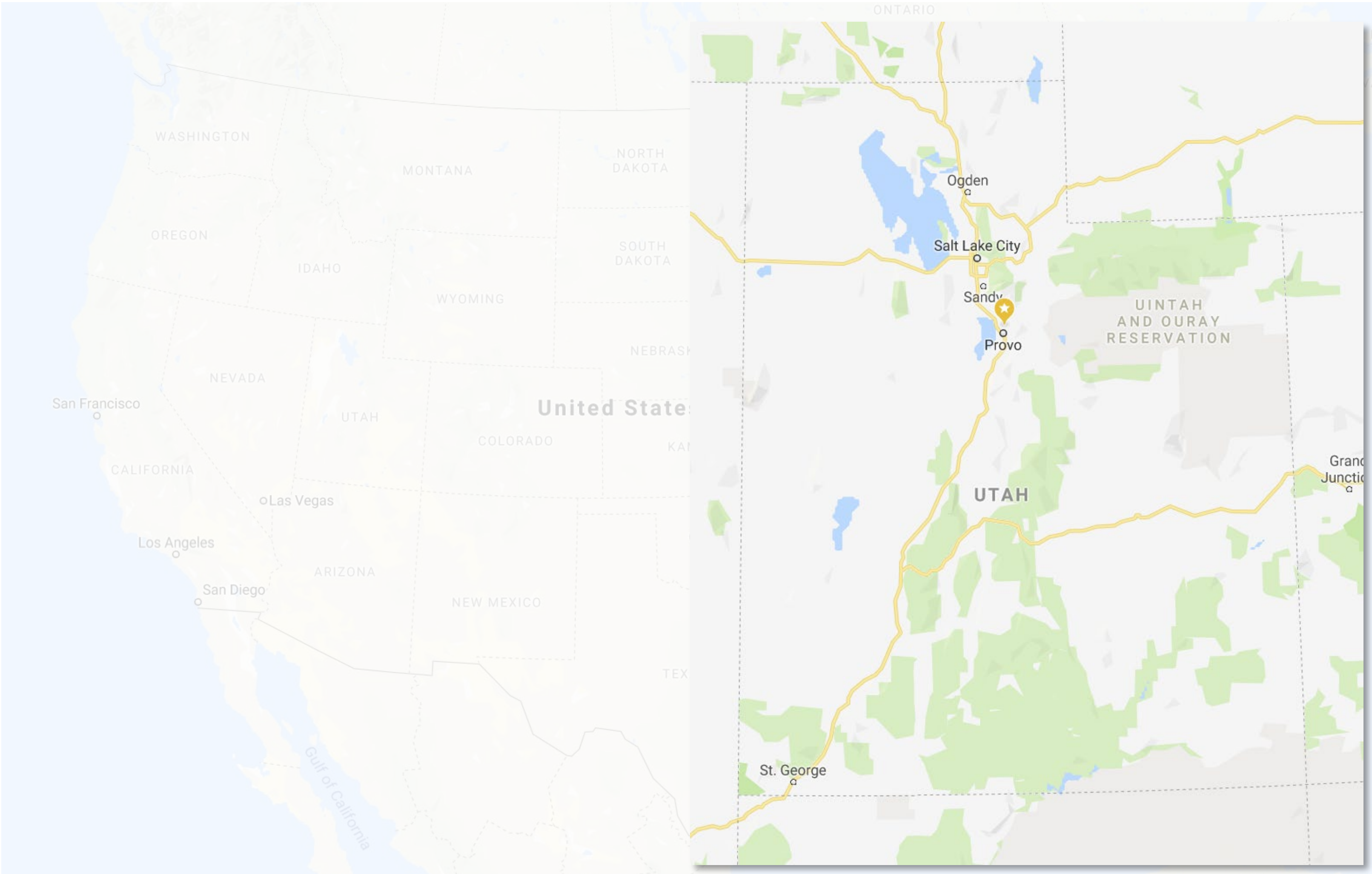


Ryan Farrell

BYU
BRIGHAM YOUNG
UNIVERSITY

Introductions

About Me


ONTARIO
WASHINGTON
MONTANA
NORTH DAKOTA
OREGON
IDAHO
SOUTH DAKOTA
WYOMING
NEBRASK
NEVADA
UTAH
United State
San Francisco
COLORADO
CALIFORNIA
Las Vegas
Los Angeles
ARIZONA
San Diego
NEW MEXICO
Gulf of California
Ogden
Salt Lake City
Sandy
Provo
UINTAH AND OURAY RESERVATION
UTAH
Grand Junctio
St. George
Cuba

UTAH

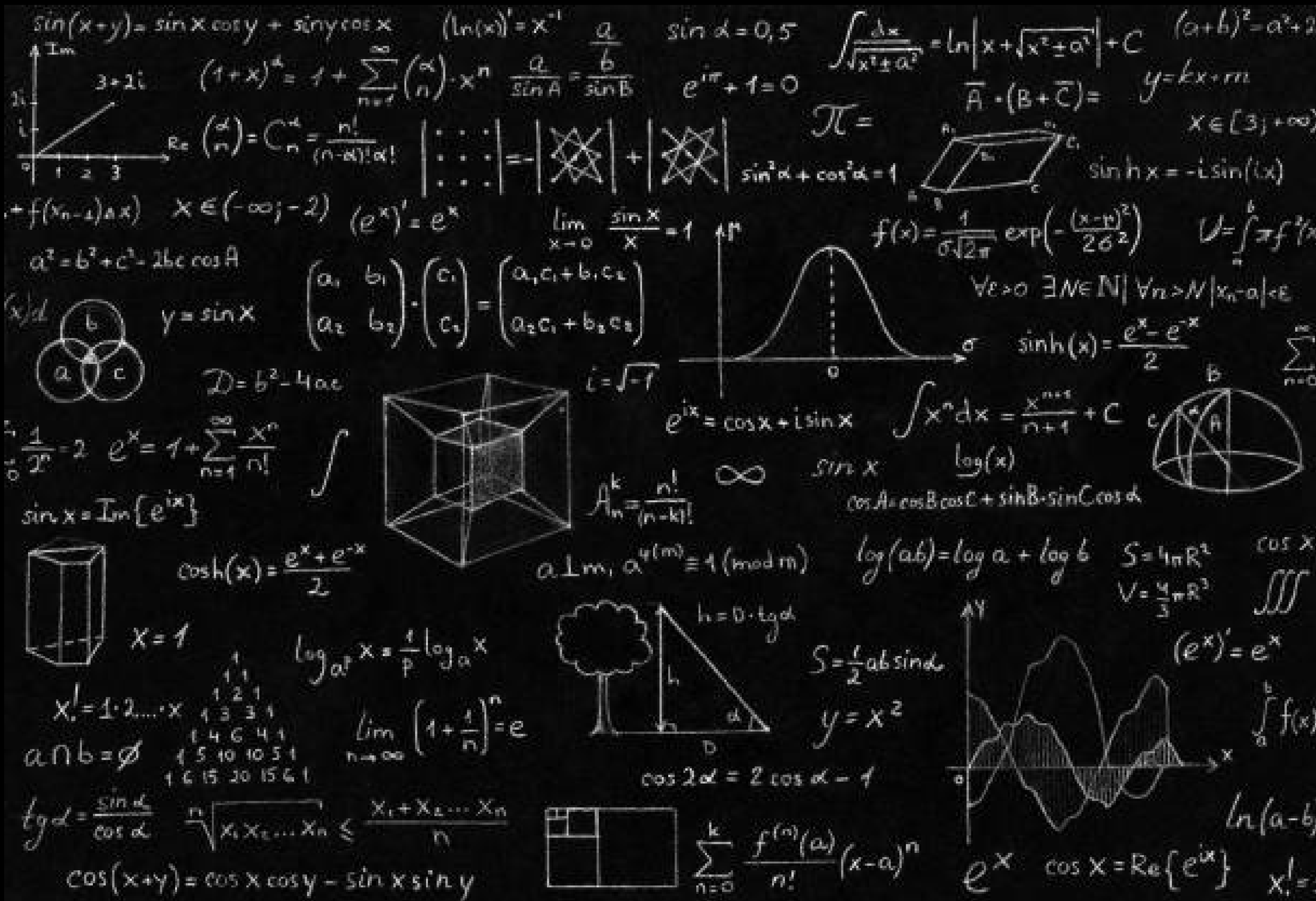
Brigham Young University (BYU)

By the Numbers

- 30,843 Undergraduates
- 2,790 Graduate Students
- 1,537 International Students (105 Countries)
- 65% of students speak a second language

Introductions

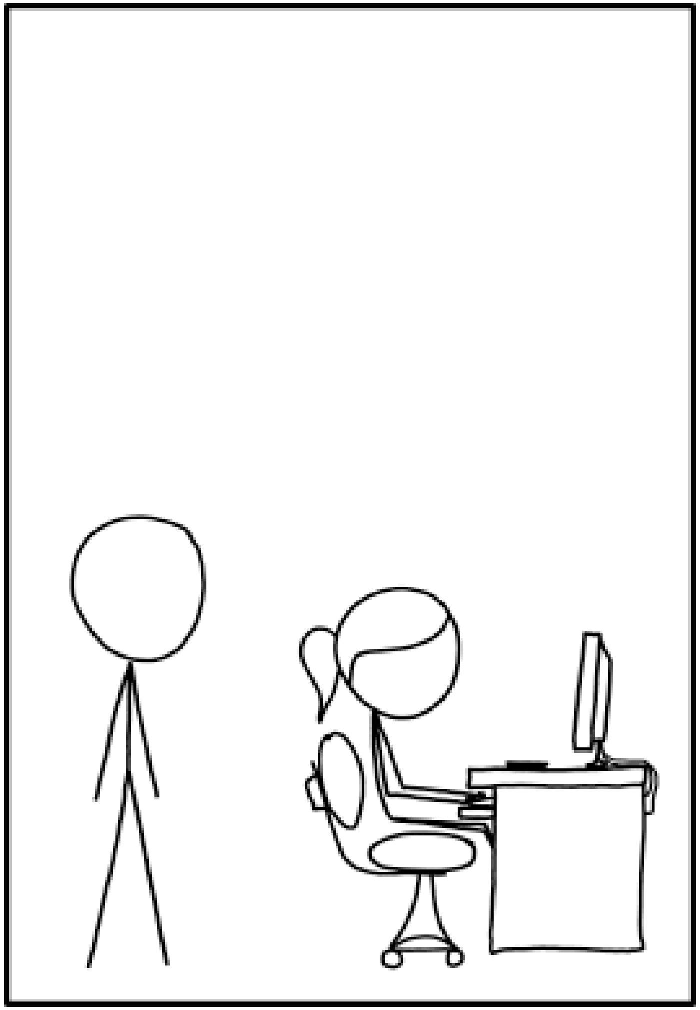
DAY ONE


sin(x+y) = sin x cos y + sin y cos x
sin α = 0,5
e^{iπ} + 1 = 0
y = kx + m
sin²α + cos²α = 1
(e^x)' = e^x
a² = b² + c² − 2bc cos A
y = sin x
D = b² − 4ac
i = √−1
e^{ix} = cos x + i sin x
log(x)
cos A = cos B cos C + sin B sin C cos α
cosh(x) = (e^x + e^{−x})/2
log(ab) = log a + log b
S = 4πR²
V = 4/3 πR³
X = 1
h = D·tg α
S = 1/2 ab sin α
y = x²
a ∩ b = ∅
cos 2α = 2 cos α − 1
ln(a−b)
cos x = Re{e^{ix}}
cos(x+y) = cos x cos y − sin x sin y

What is Computer Vision?

https://xkcd.com/1425/

In the 60s, Marvin Minsky assigned a couple of undergrads to spend the summer programming a computer to use a camera to identify objects in a scene. He figured they'd have the problem solved by the end of the summer. Half a century later, we're still working on it.



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

DSSI Computer Vision Course

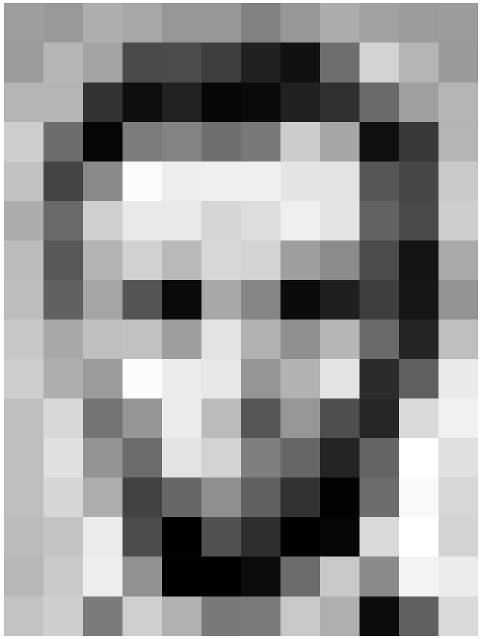
What do
you already
know?

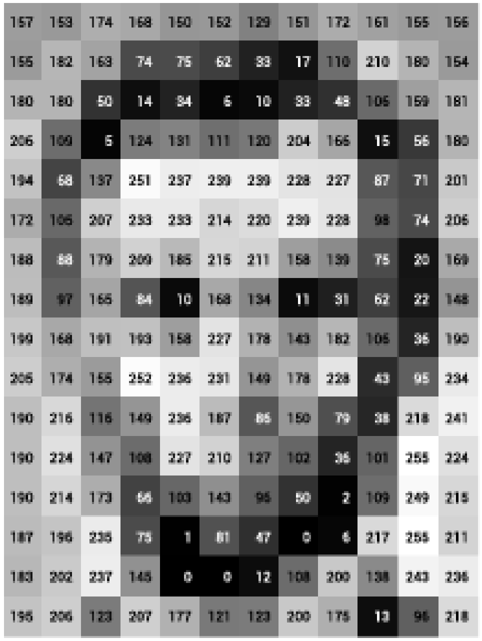
What do
you want to
know?

DSSI Computer Vision Course

- Day 1 (Today) – Intro to Computer Vision
- Day 2 (Tuesday) – Clustering / Metric Learning
- Day 3 (Wednesday) – Data Visualization / Deep Learning
- Day 4 (Thursday) – Fine-grained Recognition

Why is Vision Hard?





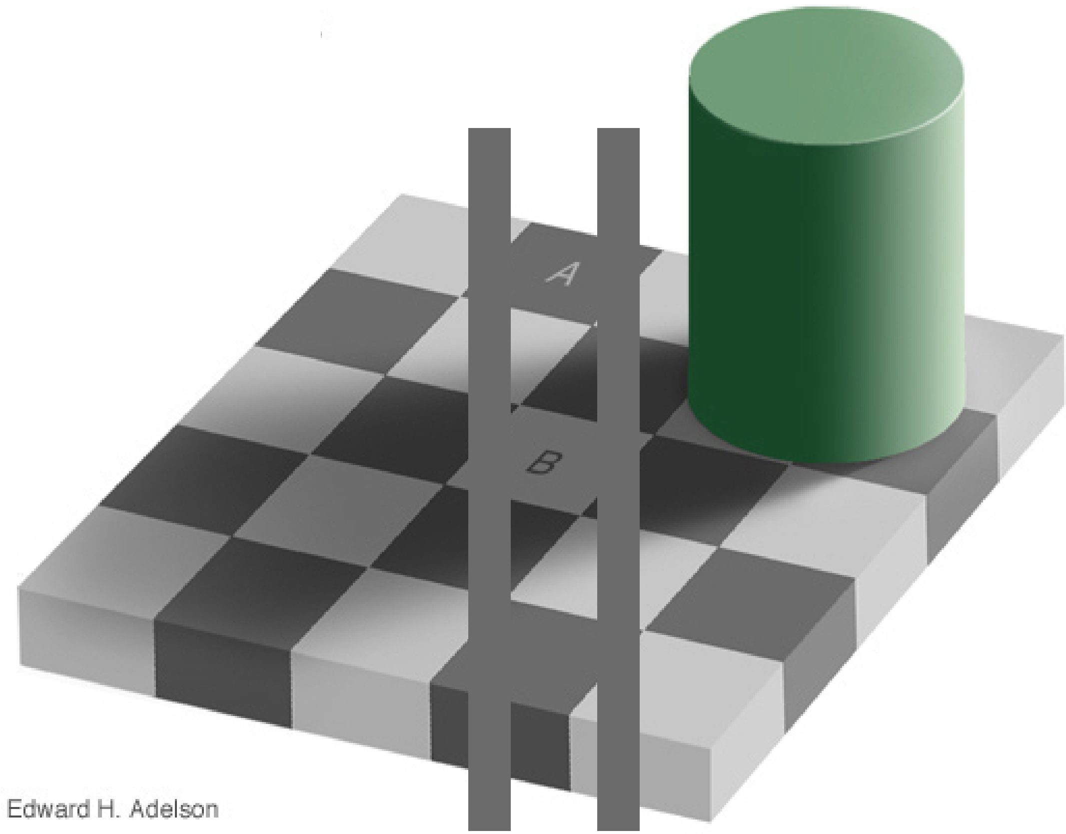
157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	95	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

Vision is Hard, Even for Humans

Vision is Hard, Even for Humans

Vision is Hard, Even for Humans

Vision is Hard, Even for Humans



Vision is Hard, Even for Humans

[Video Explaining]

Problems in Computer Vision

Fine-Grained Recognition

Photographs by Frode Jacobsen (CCUB NABirds)

Fine-Grained Recognition

Photographs by Frode Jacobsen (CCUB NABirds)

What is Recognition??

DETECTION (WHERE)

What is Recognition??

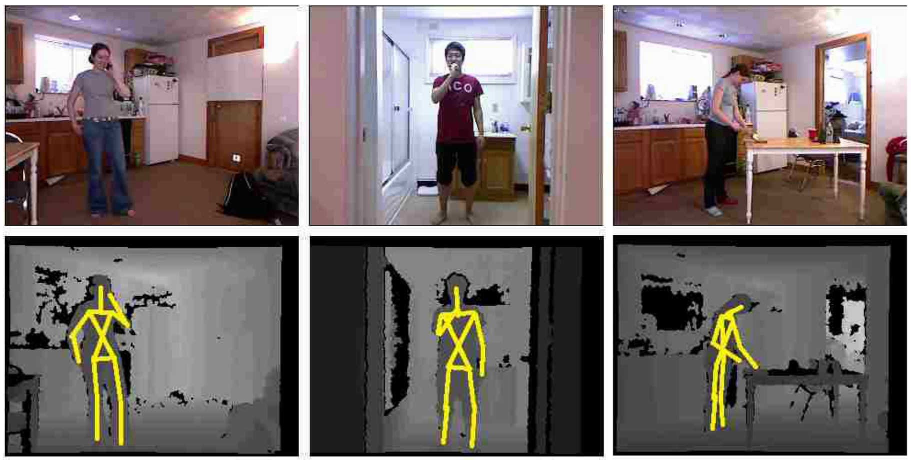
POSE ESTIMATION

What is Recognition??

CLASSIFICATION (WHO/WHAT?)

Action/Activity Recognition

Action/Activity Recognition



Segmentation

Biometrics

Faces

INVITED PAPER

Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About

Increased knowledge about the ways people recognize each other may help to guide efforts to develop practical automatic face-recognition systems.

By Pawan Sinha, Benjamin Balas, Yuri Ostrovsky, and Richard Russell

ABSTRACT | A key goal of computer vision researchers is to create automated face recognition systems that can equal, and eventually surpass, human performance. To this end, it is imperative that computational researchers know of the key findings from experimental studies of face recognition by humans. These findings provide insights into the nature of cues that the human visual system relies upon for achieving its impressive performance and serve as the building blocks for efforts to artificially emulate these abilities. In this paper, we present what we believe are 19 basic results, with implications for the design of computational systems. Each result is described briefly and appropriate pointers are provided to permit an in-depth study of any particular result.

KEYWORDS | Benchmarks; configuration; face pigmentation; face recognition; human vision; neural correlates; resolution; visual development

I. INTRODUCTION

Notwithstanding the extensive research effort that has gone into computational face recognition algorithms, we have yet to see a system that can be deployed effectively in an unconstrained setting, with all of the attendant variability in imaging parameters such as sensor noise, viewing distance, and illumination. The only system that does seem to work well in the face of these challenges is the human visual system. It makes eminent sense, therefore, to attempt to understand the strategies this biological system employs, as a first step towards eventually translating them into machine-based algorithms. With this objective in mind, we review here 19 important results regarding face recognition by humans. While these observations do not constitute a coherent theory of face recognition in human vision (we simply do not have all the pieces yet to construct such a theory), they do provide useful hints and constraints for one. We believe that for this reason, they are likely to be useful to computer vision researchers in guiding their ongoing efforts. Of course, the success of machine vision systems is not dependent on a slavish imitation of their biological counterparts. Insights into the functioning of the latter serve primarily as potentially fruitful starting points for computational investigations.

We have endeavored to bring together in one place several diverse results to be able to provide the reader a fairly comprehensive picture of our current understanding regarding how humans recognize faces. Each of the results is briefly described and, whenever possible, accompanied by its implications for computer vision. While the descriptions here are not extensive for reasons of space, we have provided relevant pointers to the literature for a more in-depth study. The results are organized along the following broad themes.

Recognition as a function of available spatial resolution

Result 1: Humans can recognize familiar faces in very low-resolution images.

Result 2: The ability to tolerate degradations increases with familiarity.

Manuscript received July 12, 2005; revised March 15, 2006.
P. Sinha, **B. Balas**, and **Y. Ostrovsky** are with the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: psinha@mit.edu; bjbalas@mit.edu; yostr@mit.edu).
R. Russell is with the Department of Psychology, Harvard University, Cambridge, MA 02138 USA (e-mail: rrussell@fas.harvard.edu).
Digital Object Identifier: 10.1109/JPROC.2006.884093

Sinha, et al. 2006

Faces

Equivalent to 7 x 10 pixels. Performance Ceiling at 19 x 27

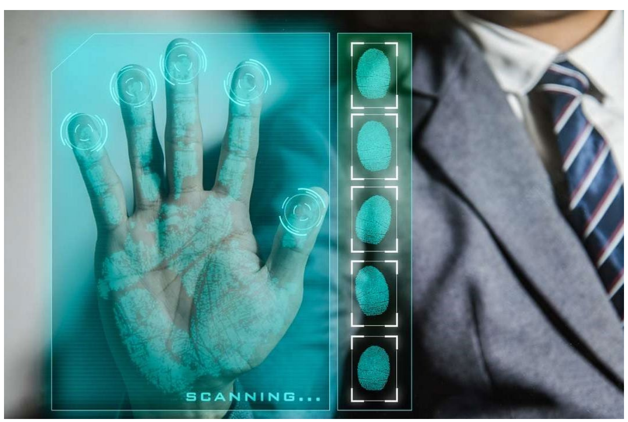
Fig. 1. ***Unlike current machine-based systems, human observers are able to handle significant degradations in face images. For instance, subjects are able to recognize more than half of all familiar faces shown to them at the resolution depicted here. Individuals shown in***

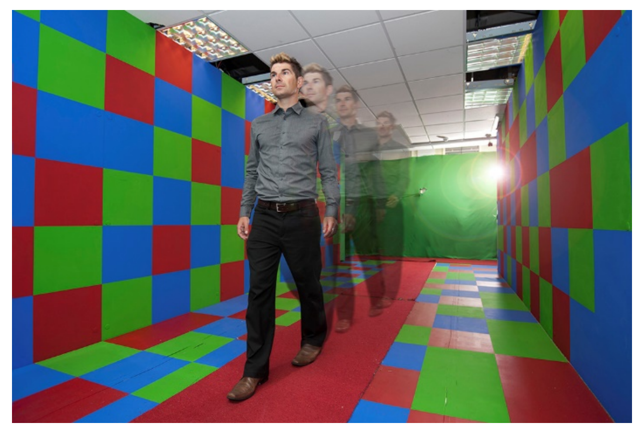
Faces

Faces

President Richard M. Nixon and actor Winona Ryder

Biometrics





Animal Biometrics

Animal Biometrics

Animal Biometrics

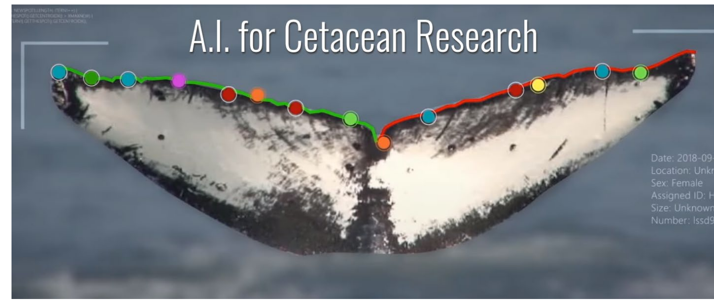
WILDBOOK

Animal Biometrics

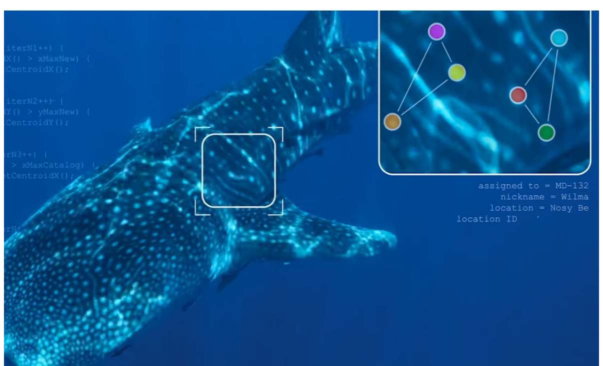
Giraffe SPOTTER
Wildbook for Giraffe

Animal Biometrics

INTERNET OF TURTLES
A WILDBOOK FOR SEA TURTLES



flukebook

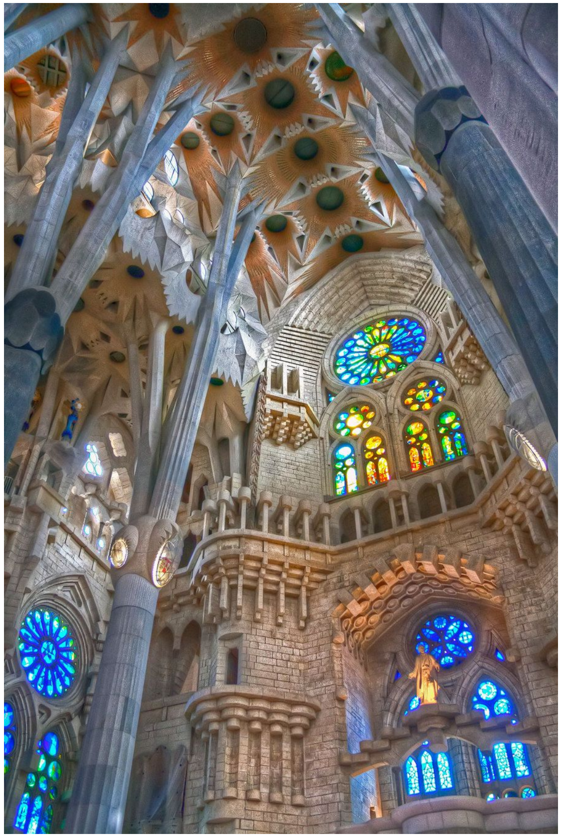


WILDBOOK
for Whale Sharks

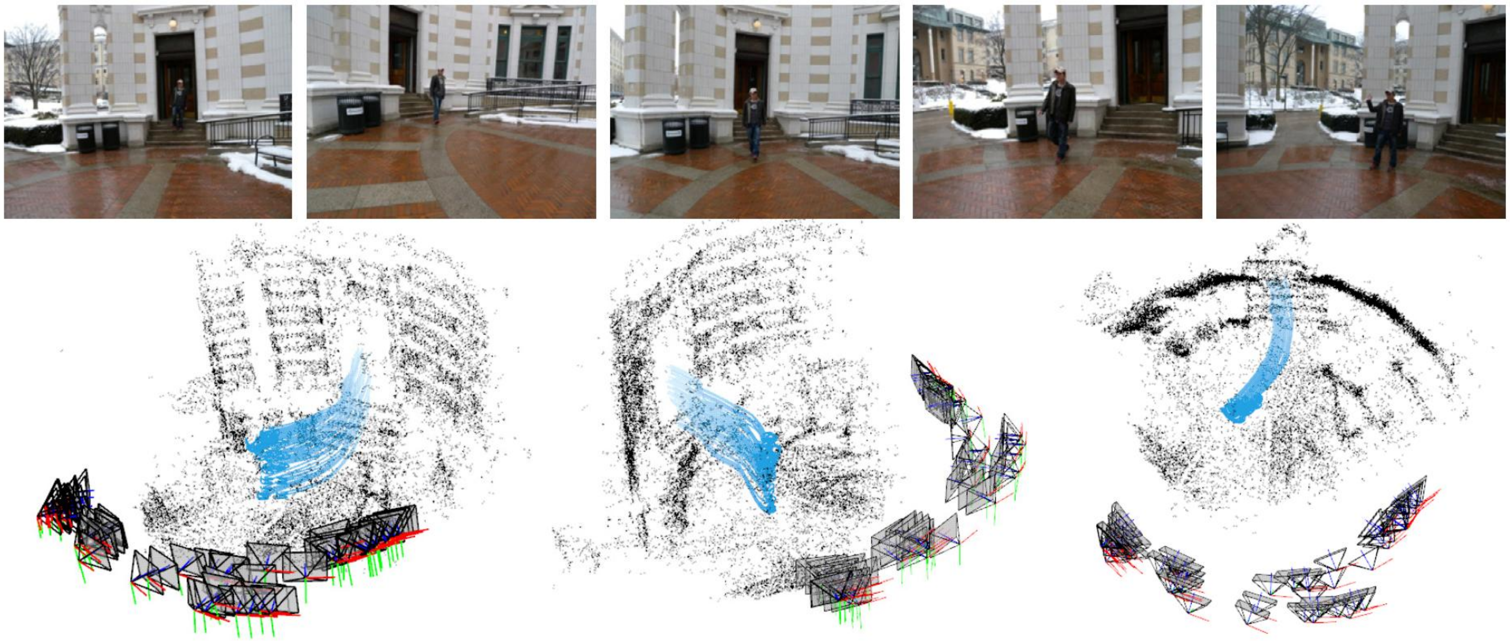
MantaMatcher
THE WILDBOOK FOR MANTA RAYS

Applications

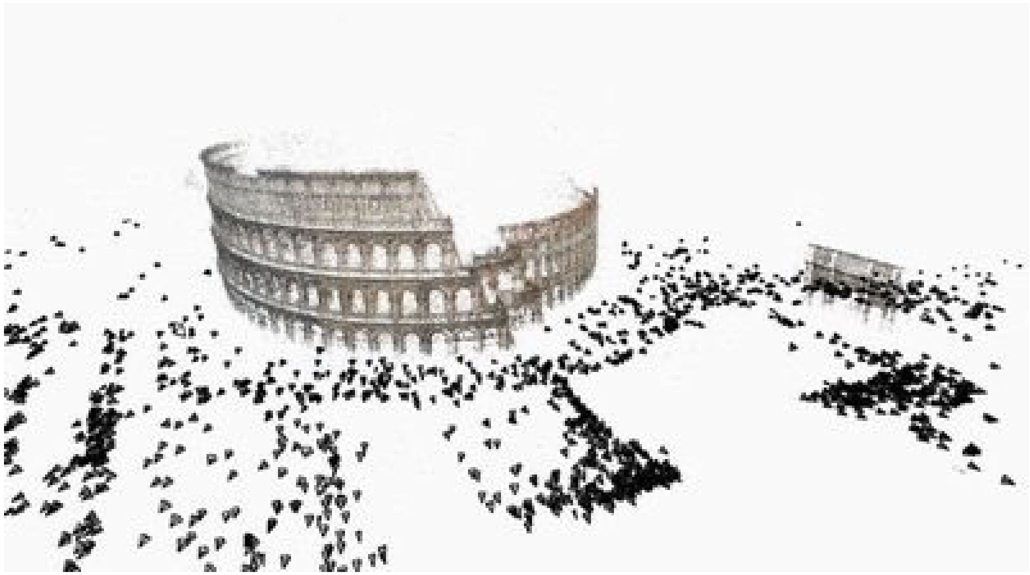
Augmented Reality



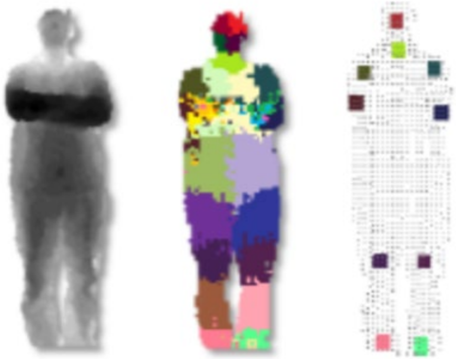
3D Reconstruction



3D Reconstruction

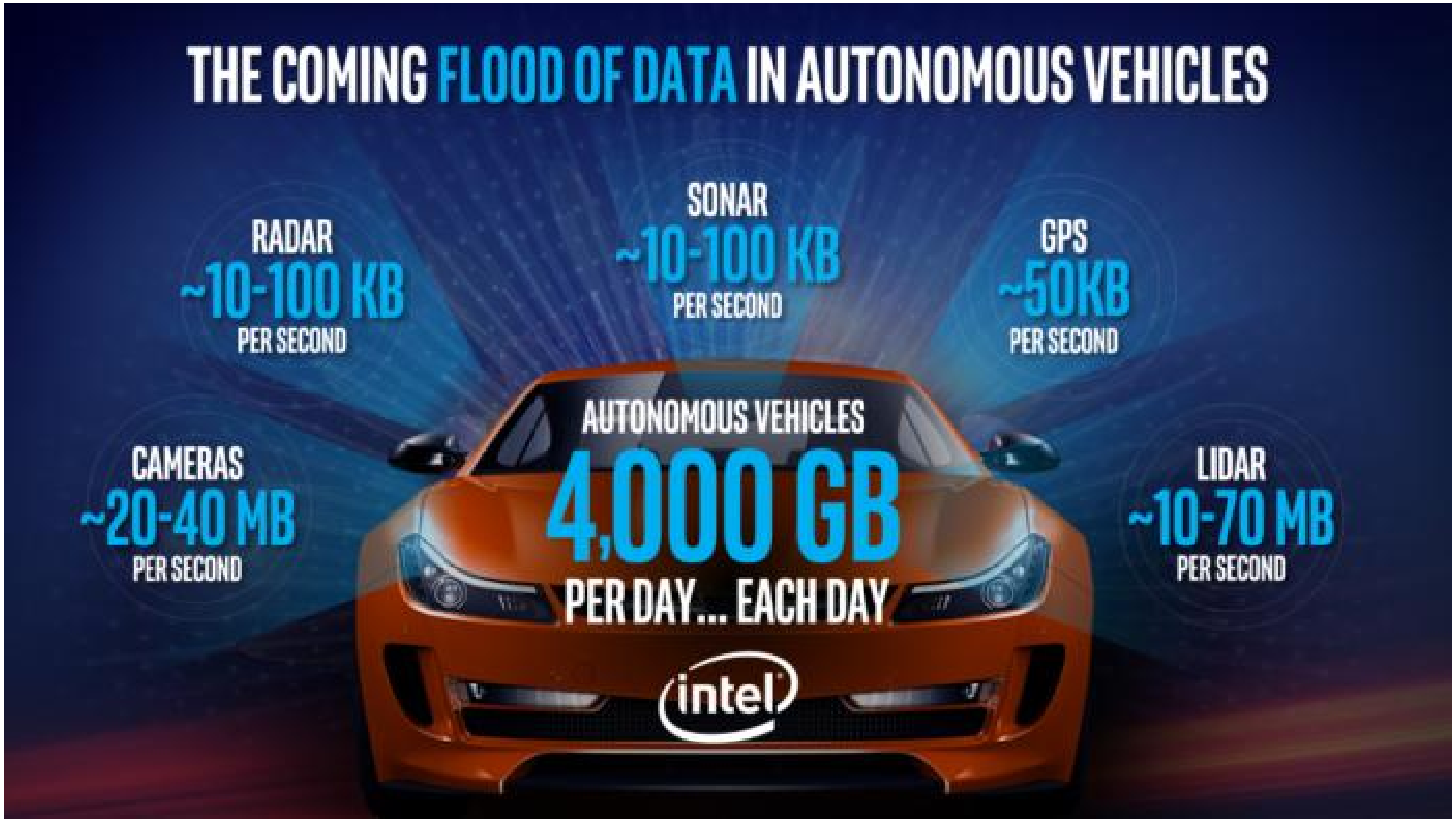


Kinect / Depth Cameras

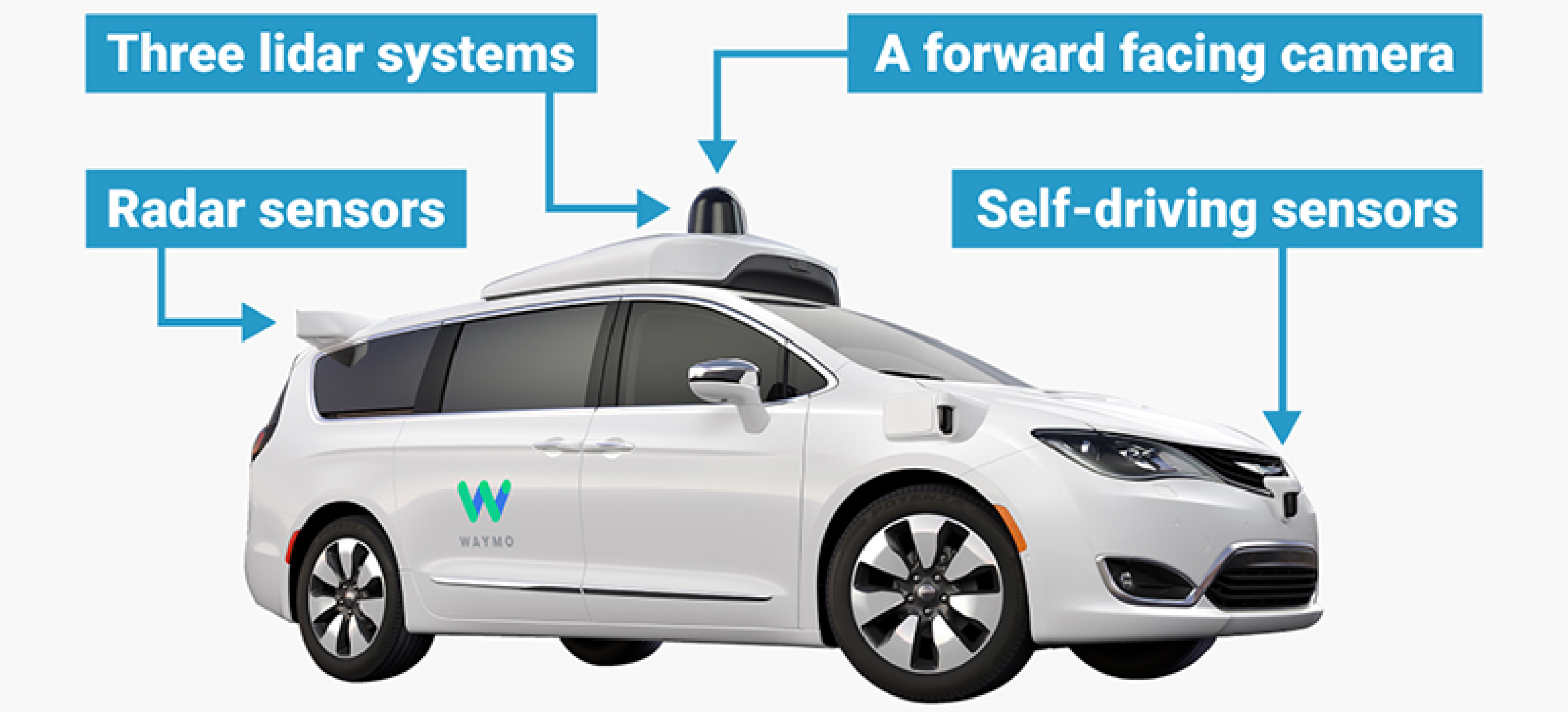


[Dot Pattern Video]

Self-driving / Autonomous Vehicles

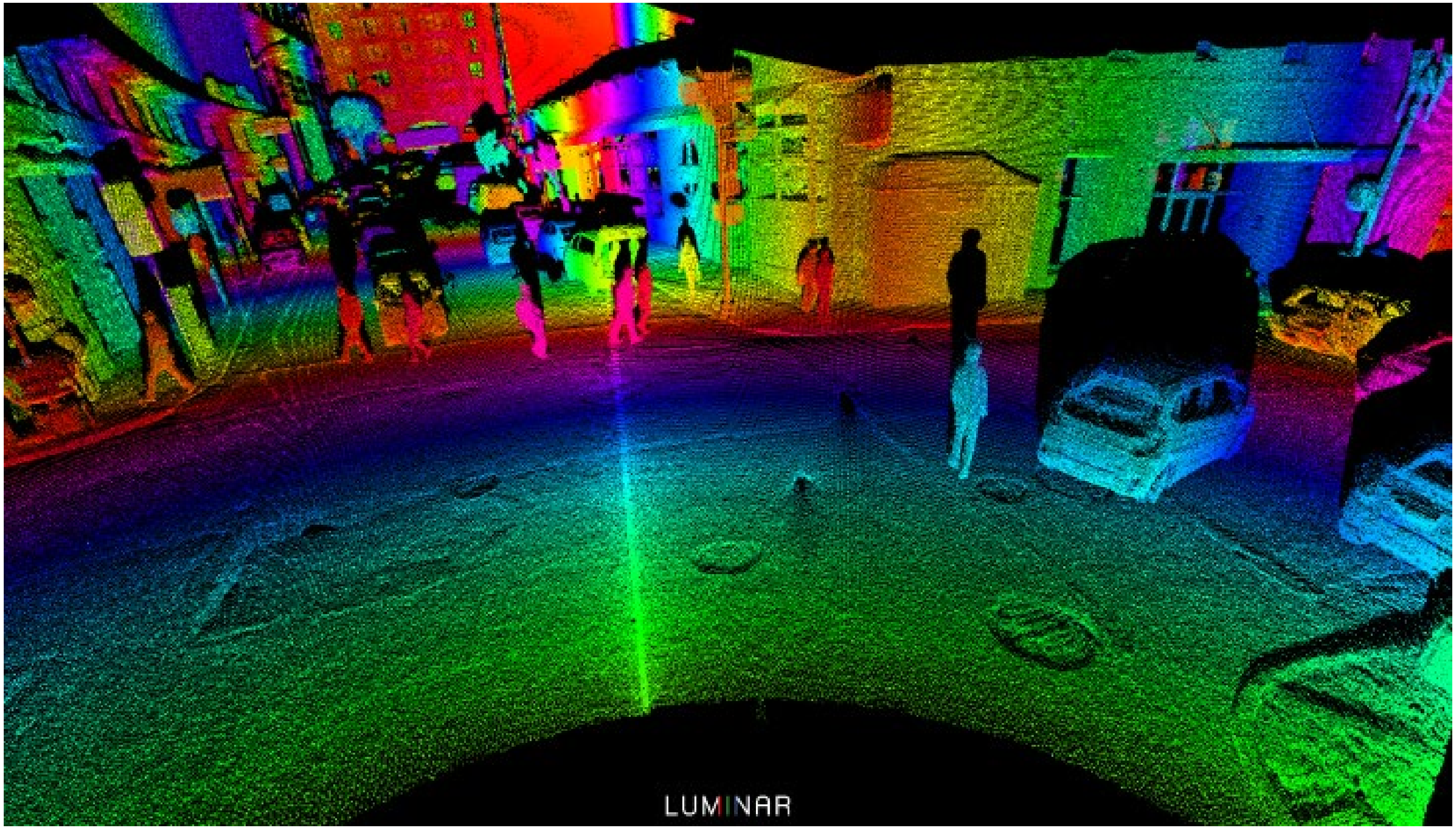


Self-driving / Autonomous Vehicles

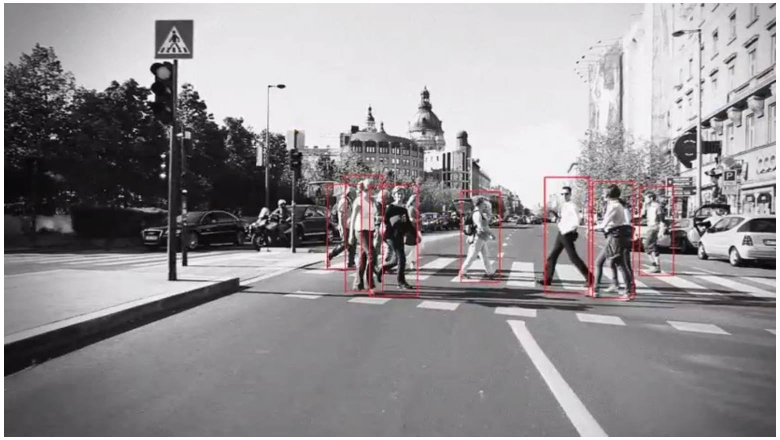


Self-driving / Autonomous Vehicles

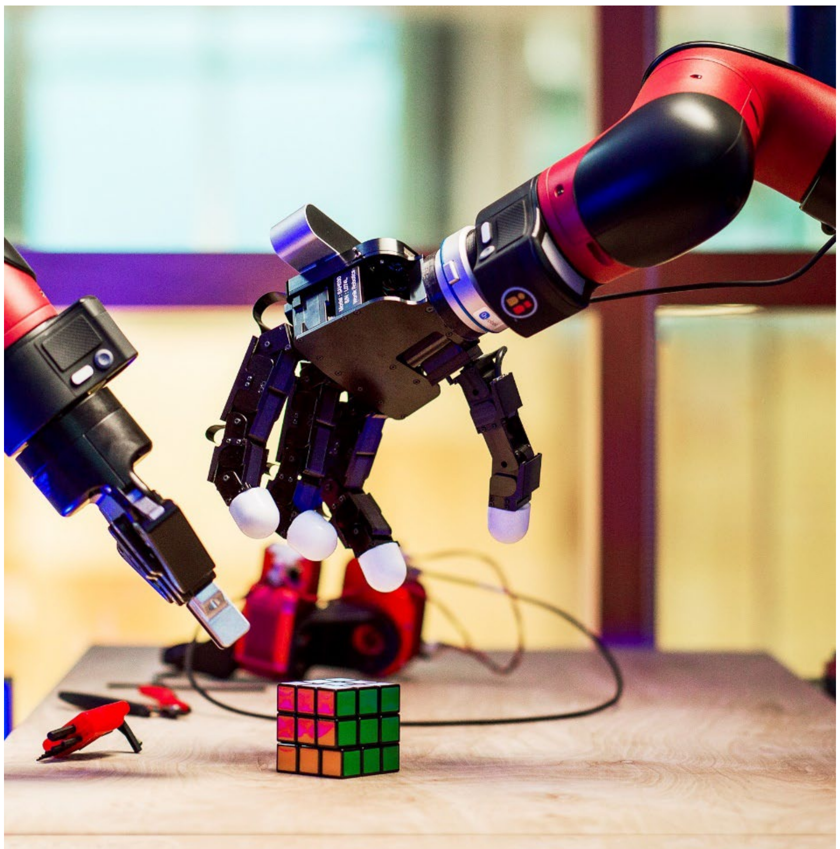
Self-driving / Autonomous Vehicles

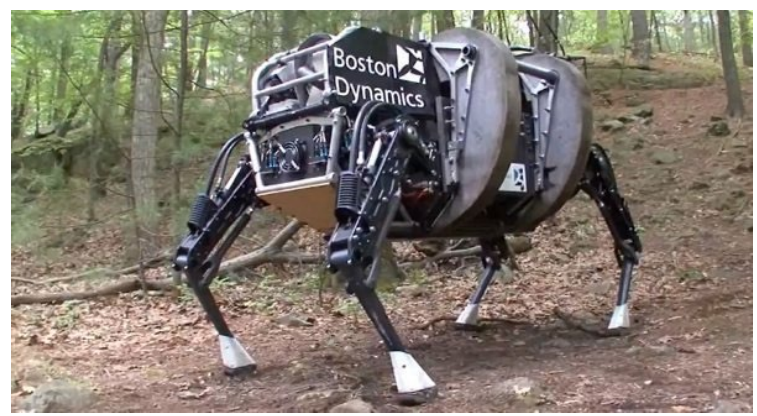


Self-driving / Autonomous Vehicles



Robotics





Robotics

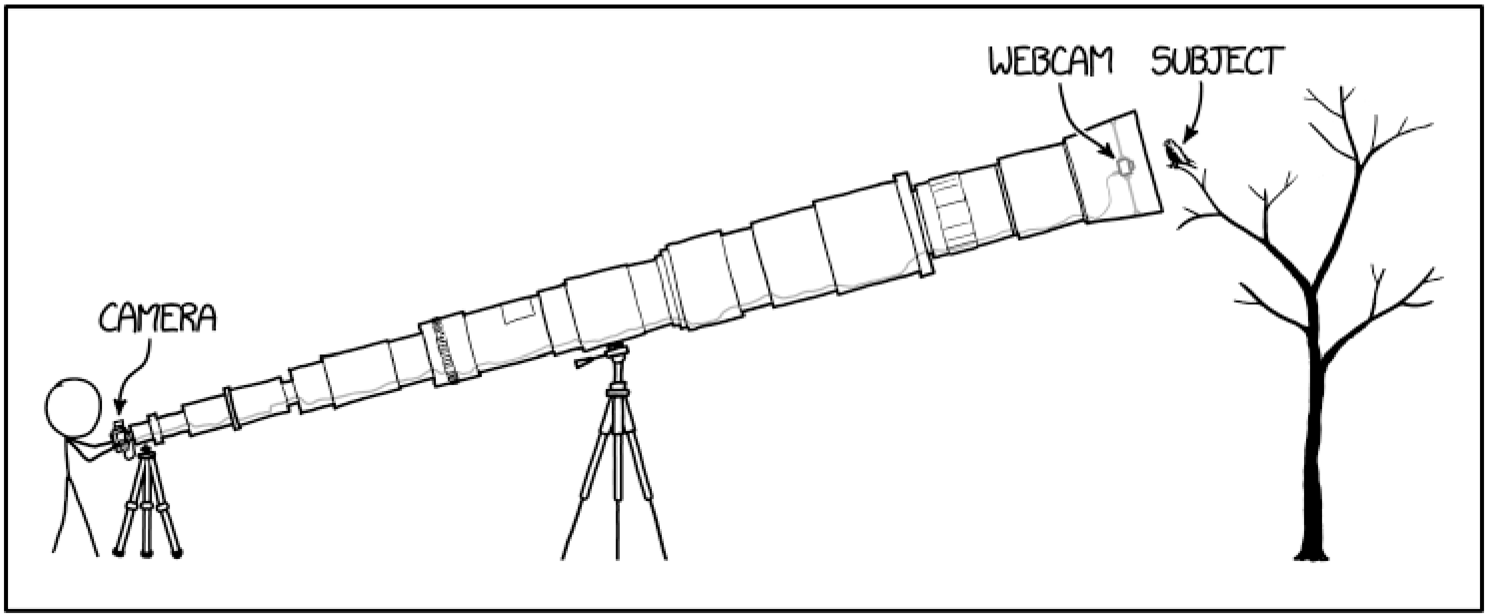
Our Everyday Lives

Our Everyday Lives

Our Everyday Lives

Our Everyday Lives

https://xkcd.com/1855/

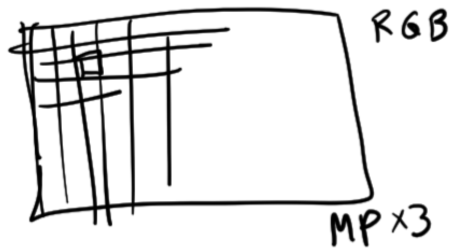


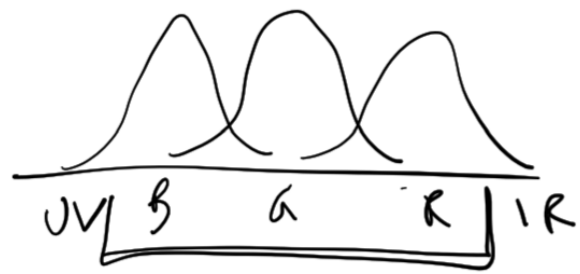
TELEPHOTO TIP: IF YOU ADD ENOUGH CONVERTERS AND EXTENDERS, YOU DON'T ACTUALLY NEED A FANCY LENS.

Representations of Visual Data

“WHITEBOARD”

PIXEL = PICTURE ELEMENT





SHAPE / CONTOURS

GROUPS OF PIXELS — SUPERPIXELS

CNNs — FILTERS (GABOR)

EDGES — DERIVATES [-1 0 1]

SIFT / MSER / LBP — HAND-CRAFTED FEATURES

128D

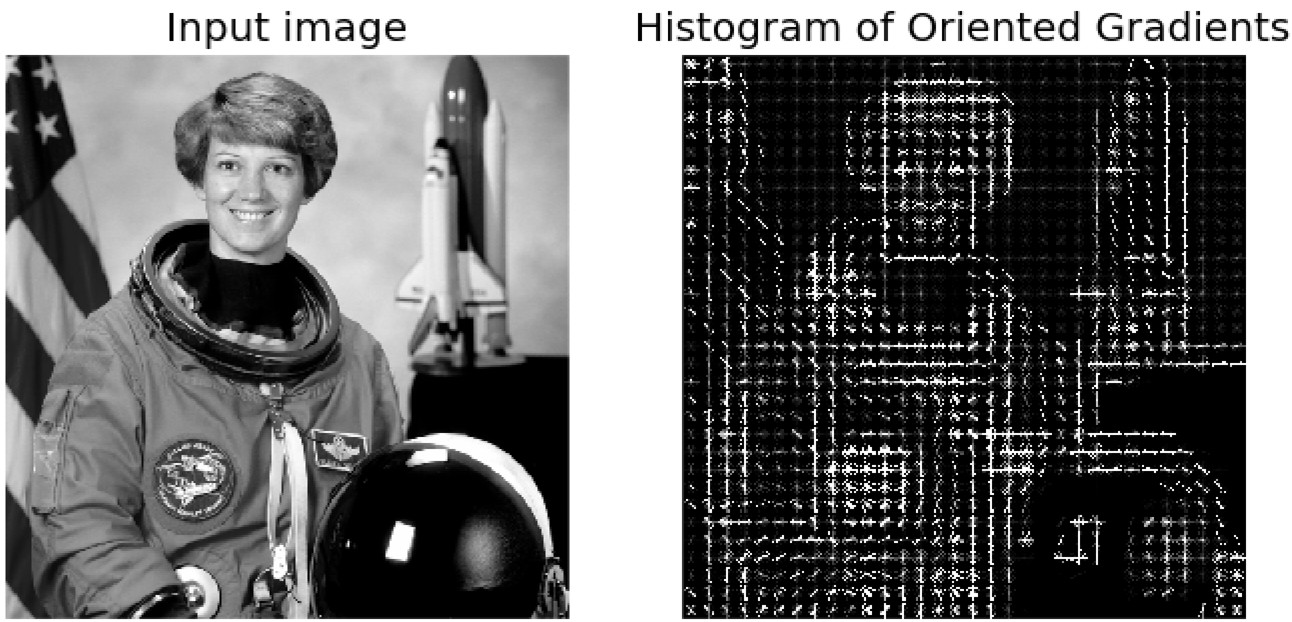
INTEREST POINTS — SIFT

CORNERS

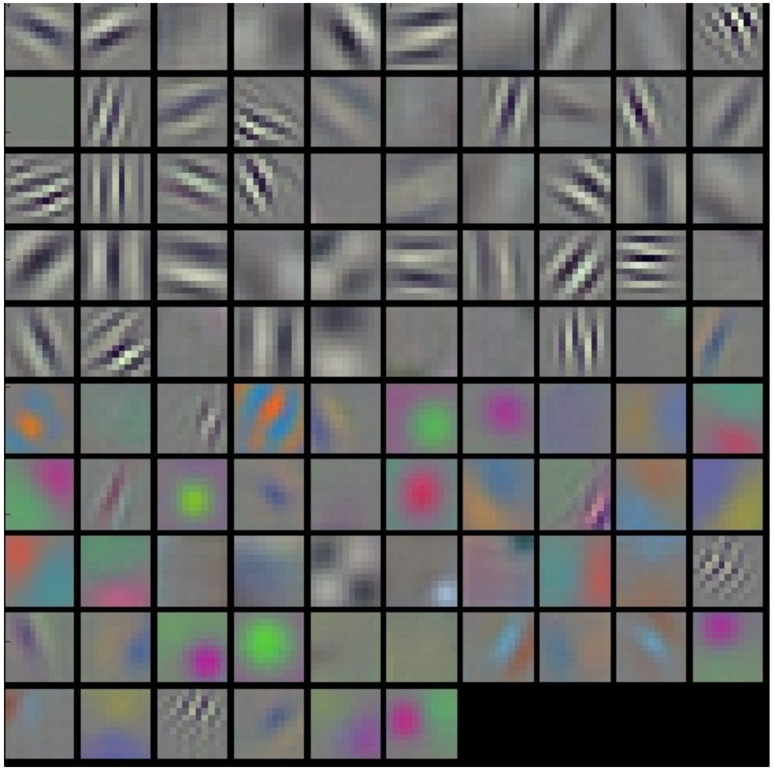
Data Representation

Superpixels

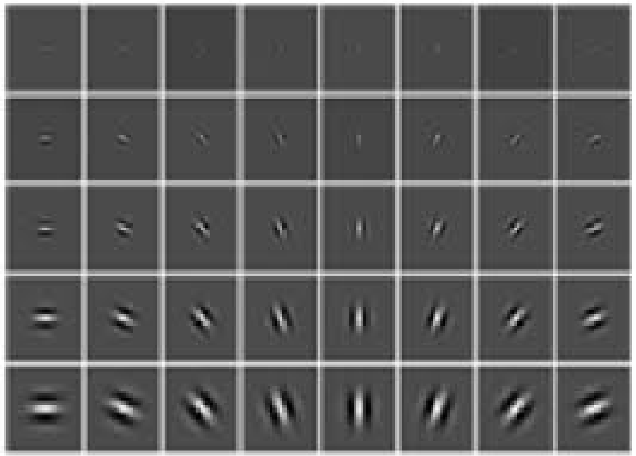
Data Representation



Data Representation

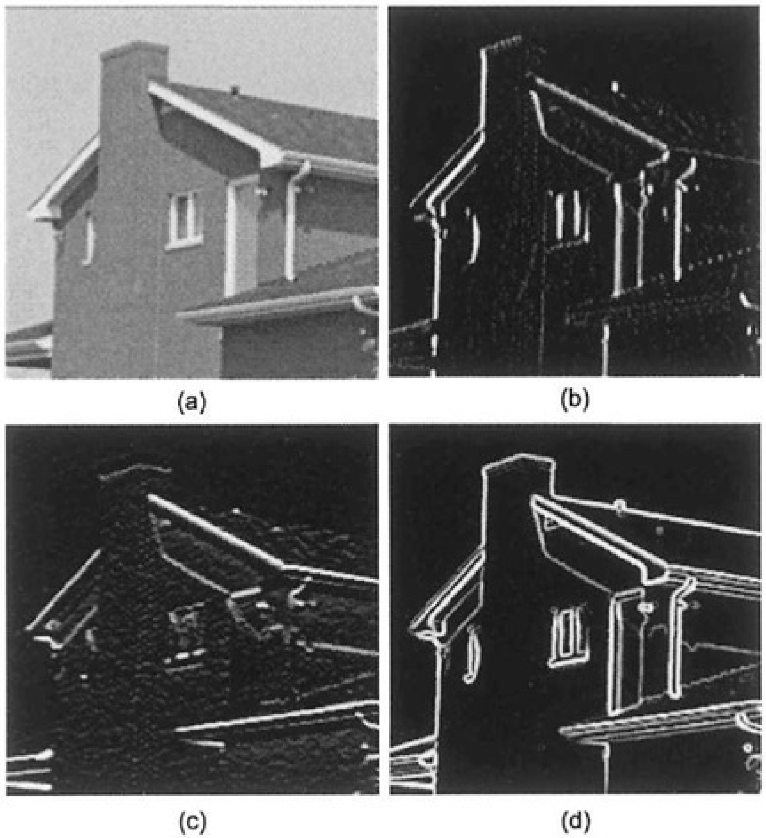


CNN Filters



Gabor Filters

Data Representation

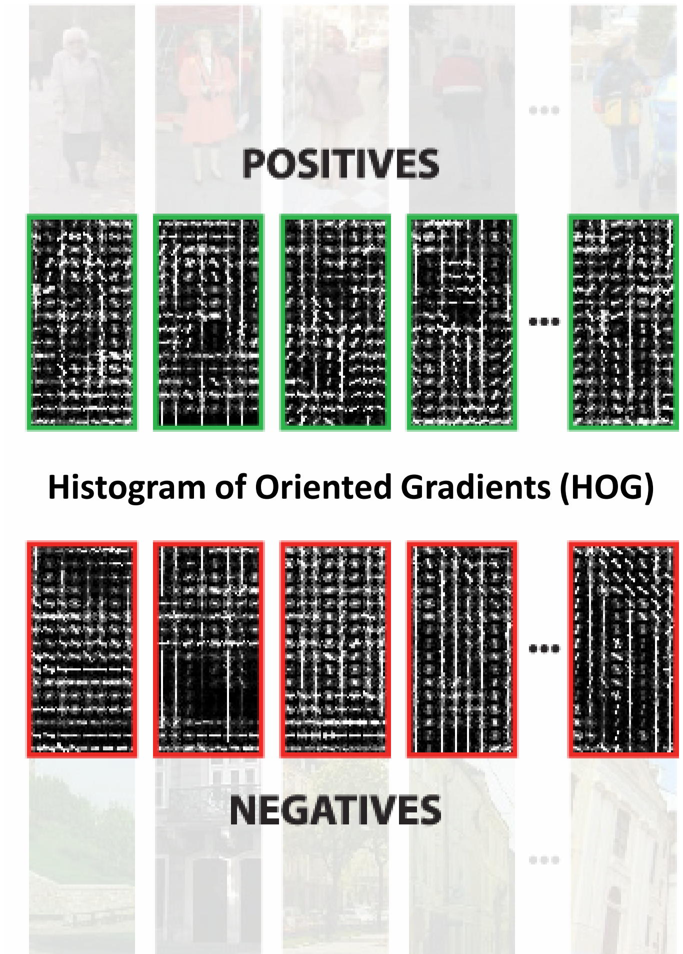


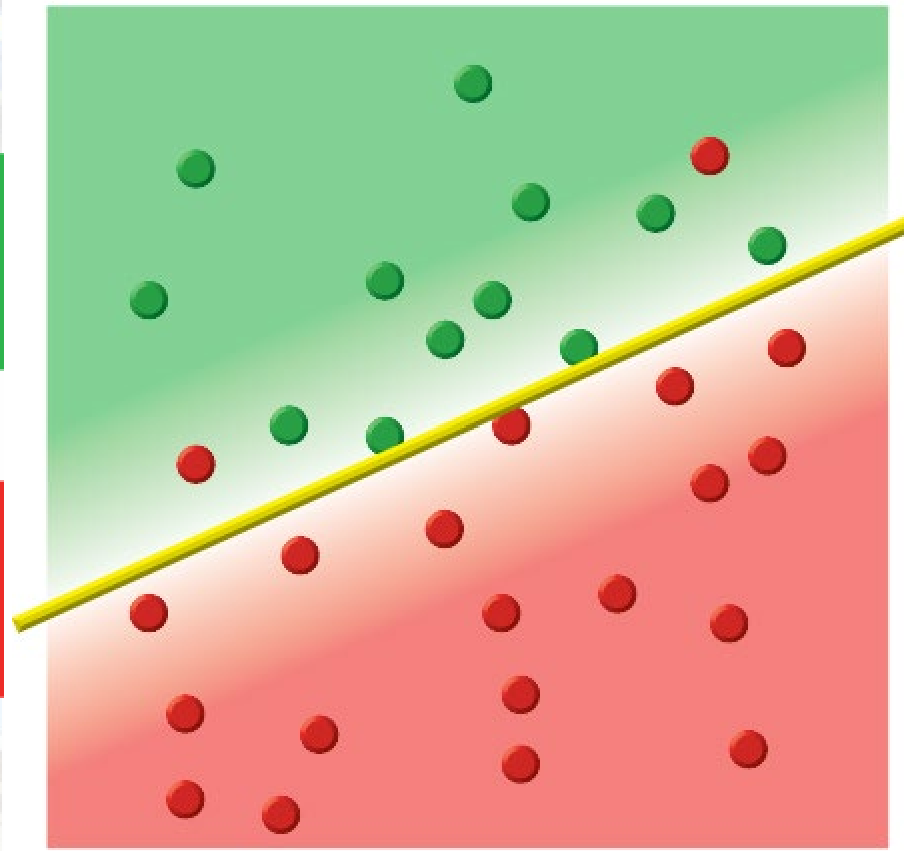
Derivative Responses

SIFT features/interest points

Object Recognition

Detection/Categorization as a Classification Problem

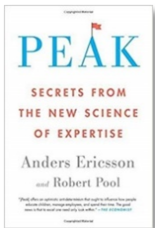




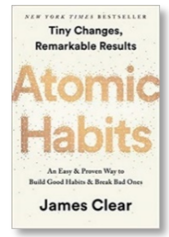
SVM (Support Vector Machine), Decision Tree / Random Forest, etc.

Advice for Students

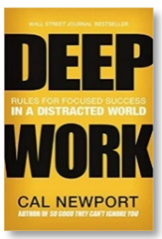
Recommended Books



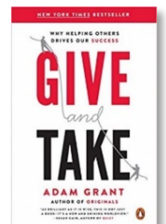
Peak
by *Anders Ericcson*



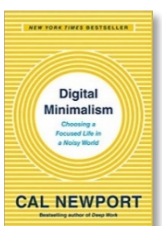
Atomic Habits
by *James Clear*



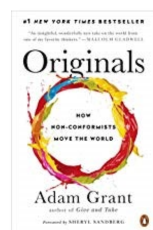
Deep Work
by *Cal Newport*



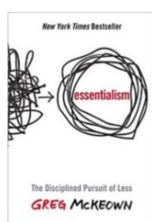
Give and Take
by *Adam Grant*



Digital Minimalism
by *Cal Newport*

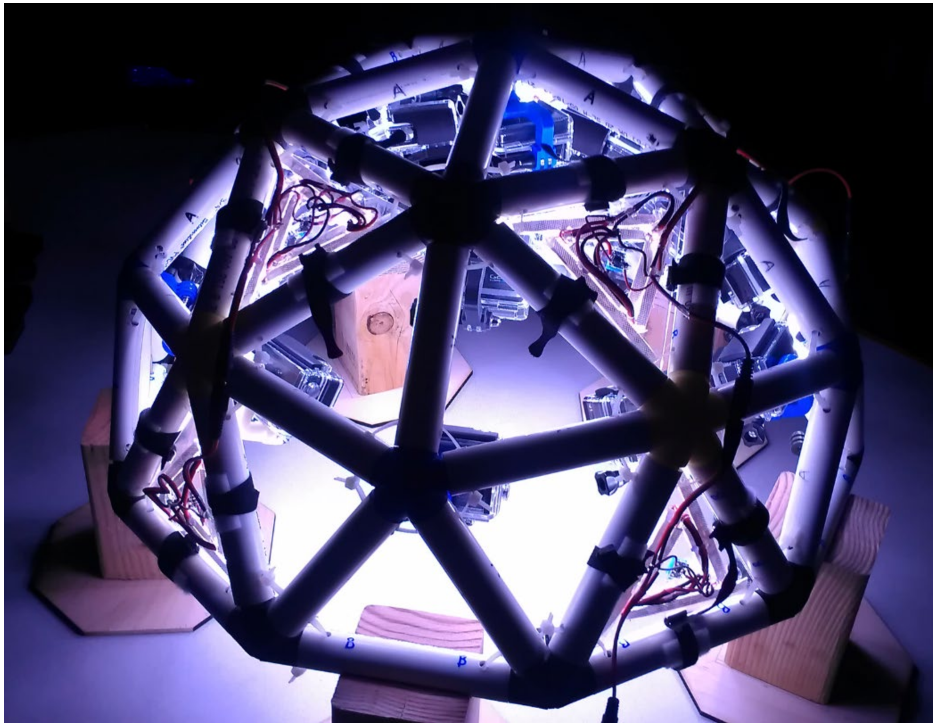


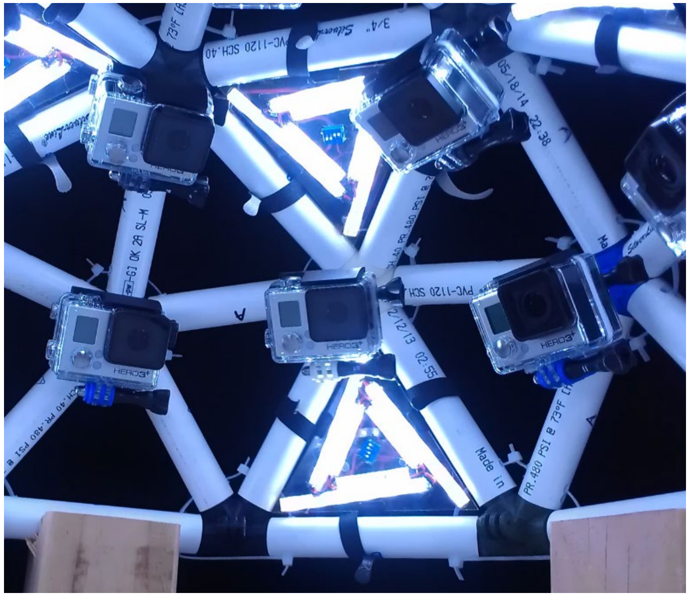
Originals
by *Adam Grant*

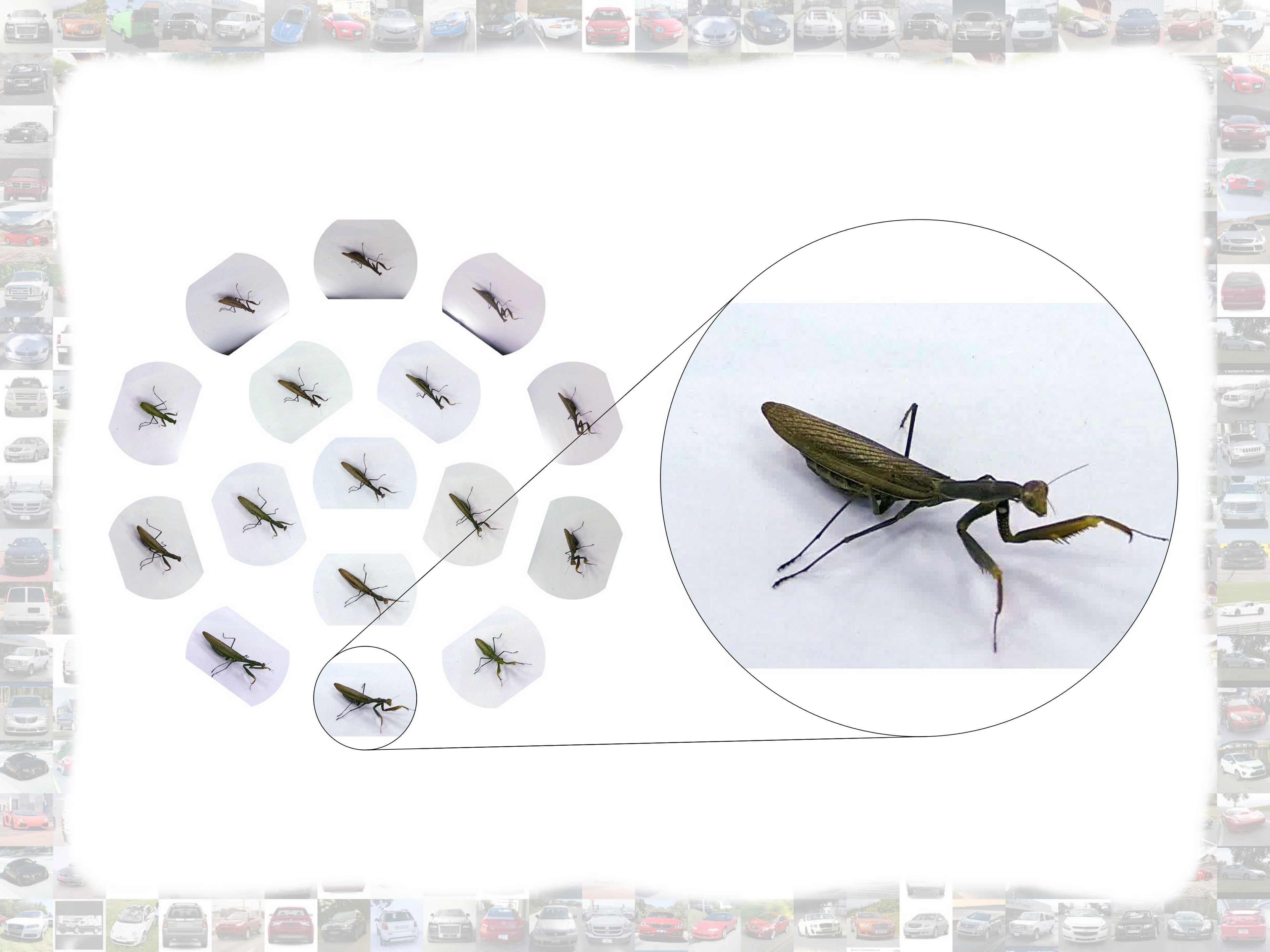


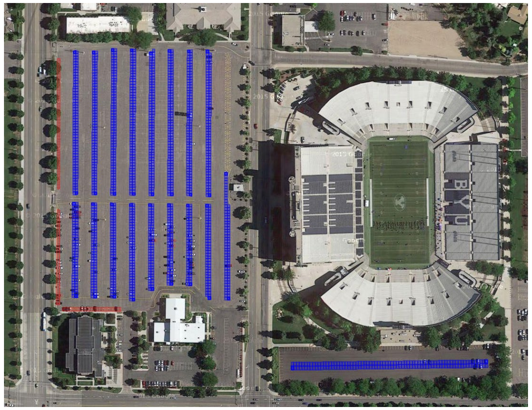
Essentialism
by *Greg McKeown*

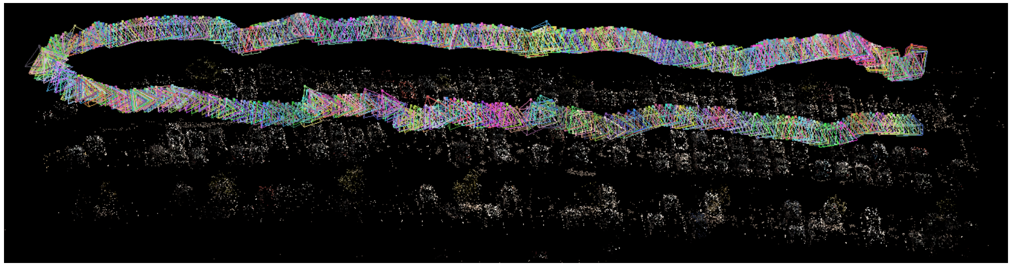
An Important Lesson about Success

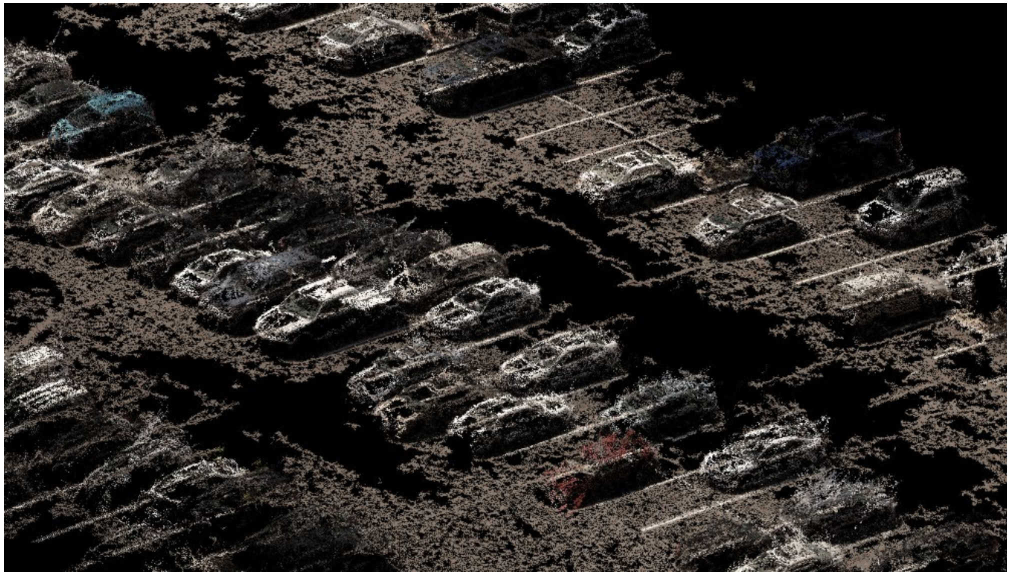


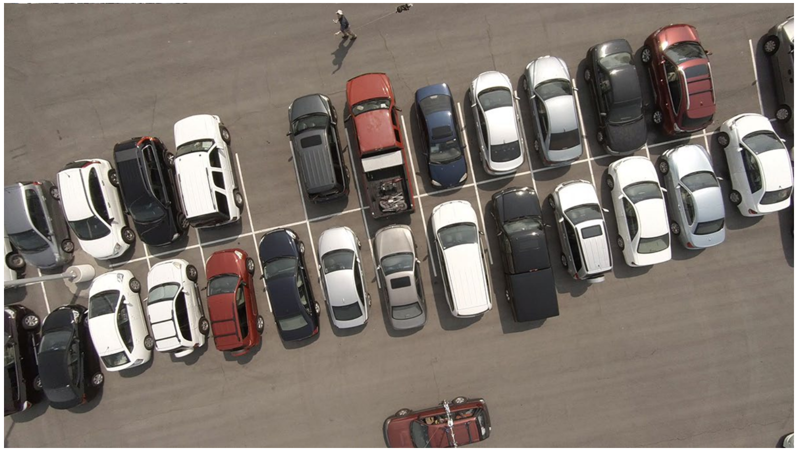

3/4"
PVC-1120 SCH.40
HERO3+
Made in
PR.480 PSI @ 73°F



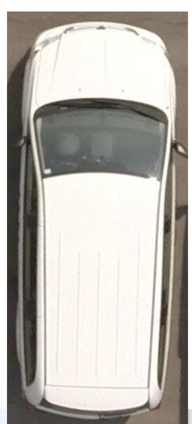








{
  "make": "**Nissan**",
  "style": "Sedan 4D",
  "vin": "XXXXXXXXXXXXXXXXXX",
  "model": "**Altima GLE / GXE / SE / SE-L / X**",
  "year": **1999**
}



{
  "make": "**Dodge**",
  "style": "Wagon 4D",
  "vin": "XXXXXXXXXXXXXXXXXX",
  "model": "**Grand Caravan S**",
  "year": **2013**
}

- Approximately 4000 vehicles imaged
- 20-25 Views per vehicle
- Approximately 100,000 total images

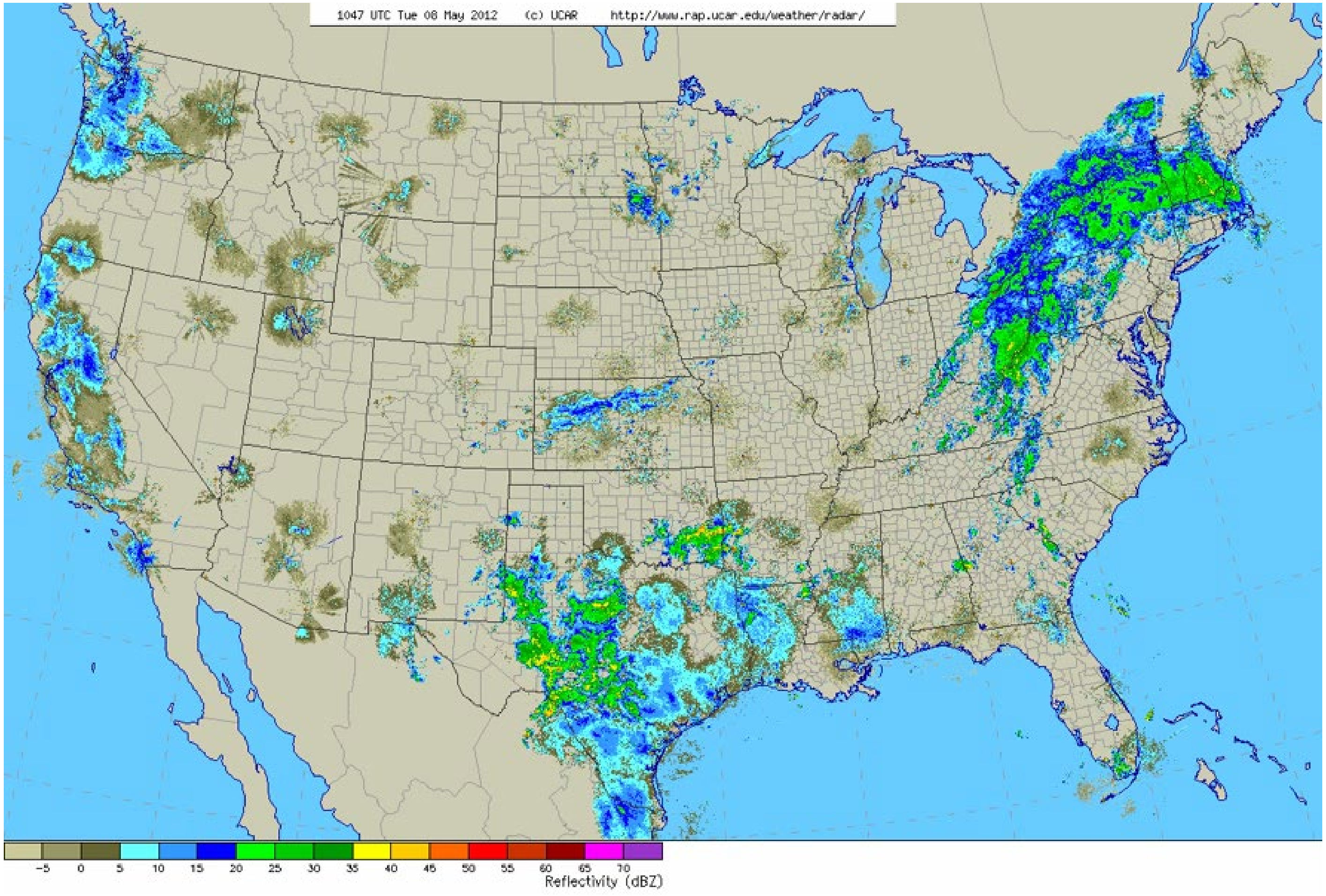
1997 Nissan Sentra

2011 Toyota 4Runner

"In the long run, men hit only what they aim at." – Henry Thoreau

"**Not failure, but low aim** would be the most severe indictment"
– Jeffrey R. Holland

"Other peoples see things and say, 'Why?' But I dream things that never were – and I say, **'Why not?'**"
– George Bernard Shaw


1047 UTC Tue 08 May 2012
(c) UCAR
http://www.rap.ucar.edu/weather/radar/
-5 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70
Reflectivity (dBZ)