Computer Vision

Introduction

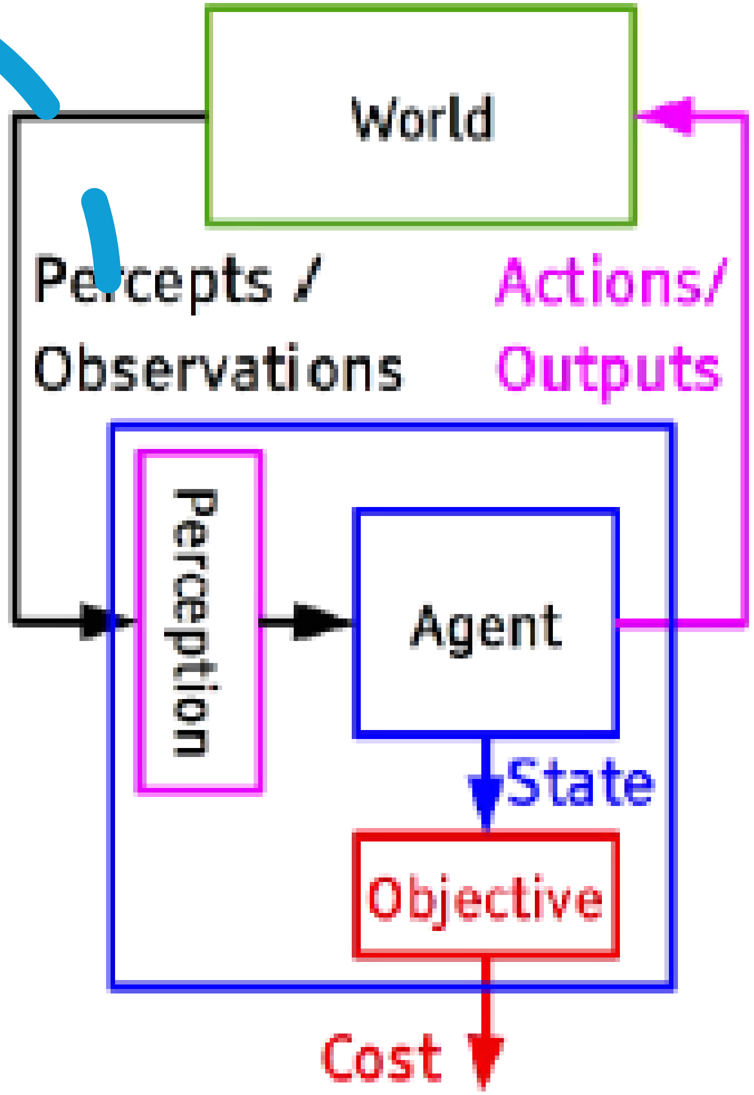
Slides Credit : Gege Gao
James Tompkin

Artificial Intelligence

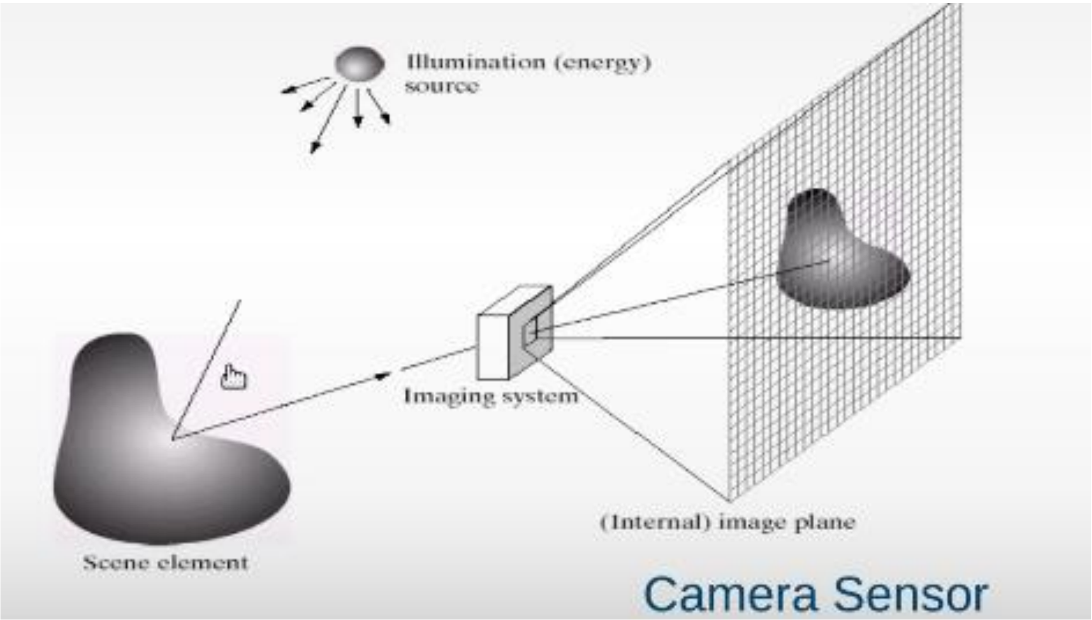
- "An attempt will be made to and how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves."

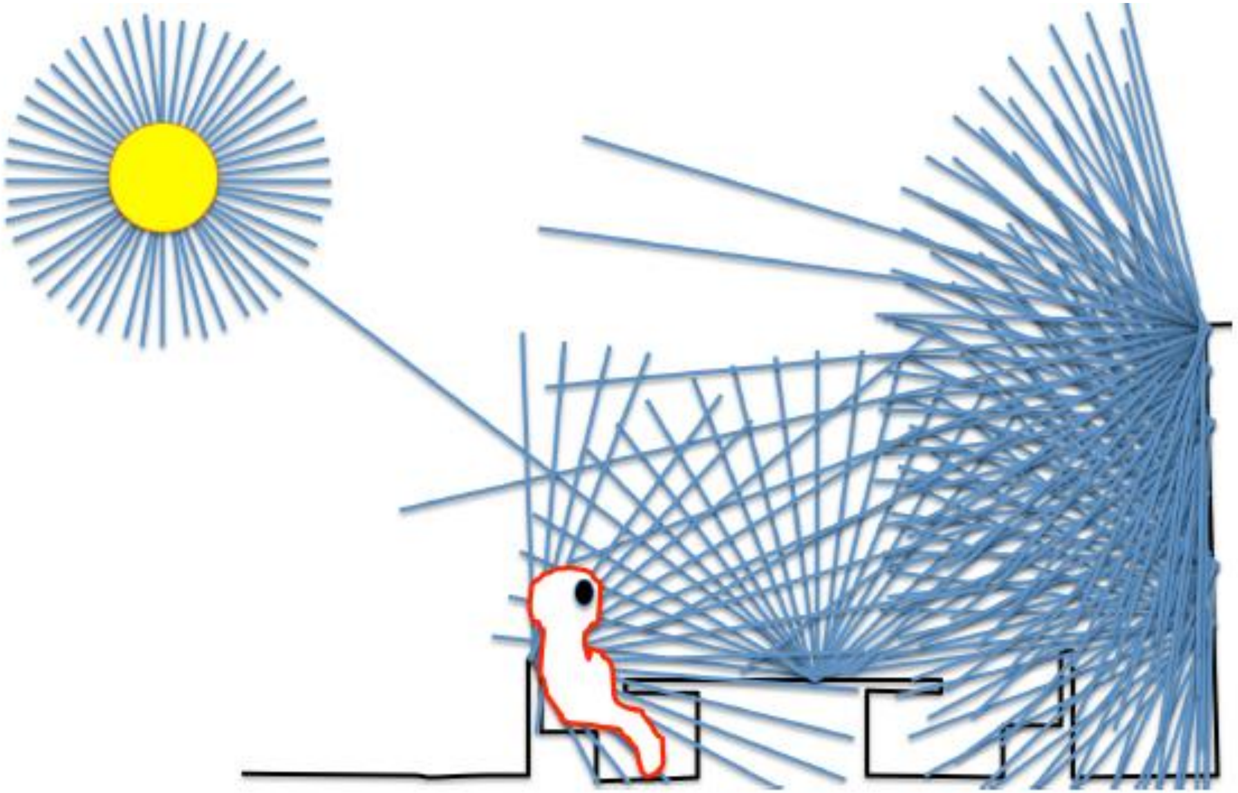
  [John McCarthy]

- Machine Learning
- Computer Vision
- Computer Graphics
- Natural Language Processing
- Robotics and Control
- Art, Industry 4.0, Education, ..



Computer Vision





- Goal of Computer Vision is to **convert light into meaning** (geometric, semantic, . . .)

Image Credits: Antonio Torralba

Computer Vision

- What does it mean, to see? The plain man's answer (and Aristotle's, too) would be, to know what is where by looking.

- To discover from images what is present in the world, where things are, what actions are taking place, to predict and anticipate events in the world.

Slide Credit: Torralba, Freeman, Isola

Applications

- Jitendra Malik, UC Berkeley
- Three ‘R’s of Computer Vision
- “The classic problems of computational vision:
  - reconstruction
  - recognition
  - (re)organization.”

Computer Vision

Slide Credit: Torralba, Freeman, Isola

Applications

- Laptops/Desktops: Biometrics auto-login (face recognition, 3D)

- Smartphones: QR codes, computational photography (Android Lens Blur, iPhone Portrait Mode), panorama construction (Google Photo Spheres), Night Sight (Pixel), iPhone Pro 3D scanning (LiDAR), body workout form detection, face filters, FaceID (iPhone)

- Web: Image search, Google photos (face recognition, object recognition, scene recognition, geolocalization from vision), Facebook (image captioning), Google maps aerial imaging (image stitching), YouTube (content categorization), Photoshop, PowerPoint (captioning, design suggestions)

Applications

- Virtual Worlds:VR/AR head tracking (Meta Quest, Apple Vision Pro), simultaneous localization and mapping, person tracking (Kinect), gesture recognition, virtual try-on, digital humans
- Medical imaging: CAT / MRI reconstruction, assisted diagnosis, surgery planning, automatic pathology, connectomics
- Industry: Vision-based robotics (human+robot spaces in Amazon warehouses), online shopping (Amazon, Walmart), machine-assisted tools (routers, jigs), OCR (USPS), ANPR (number plates for tolls), drones

Applications- Face Detection

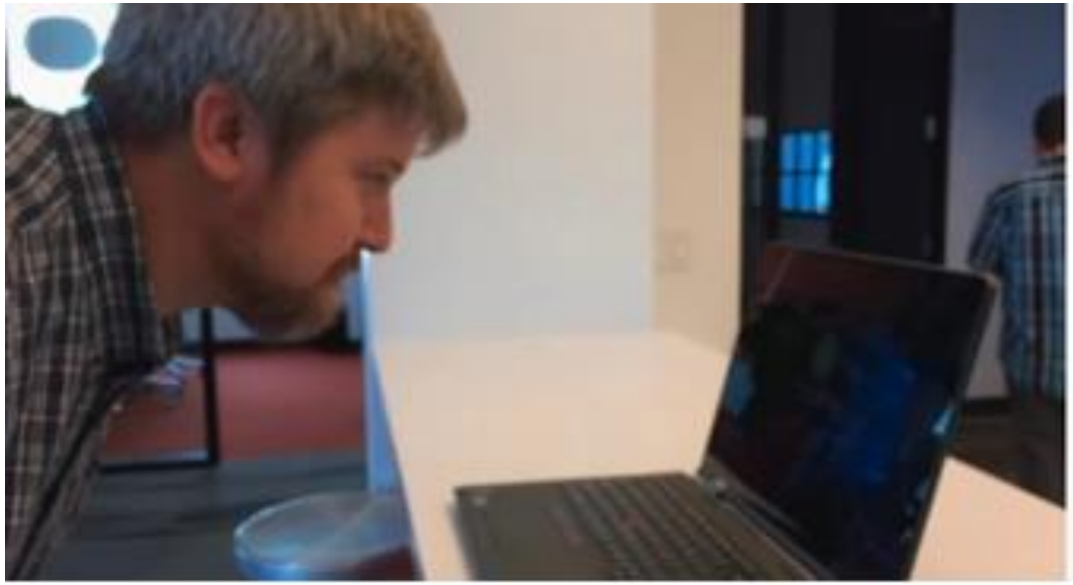

Applications- Face Swap

Applications- Smile Detection

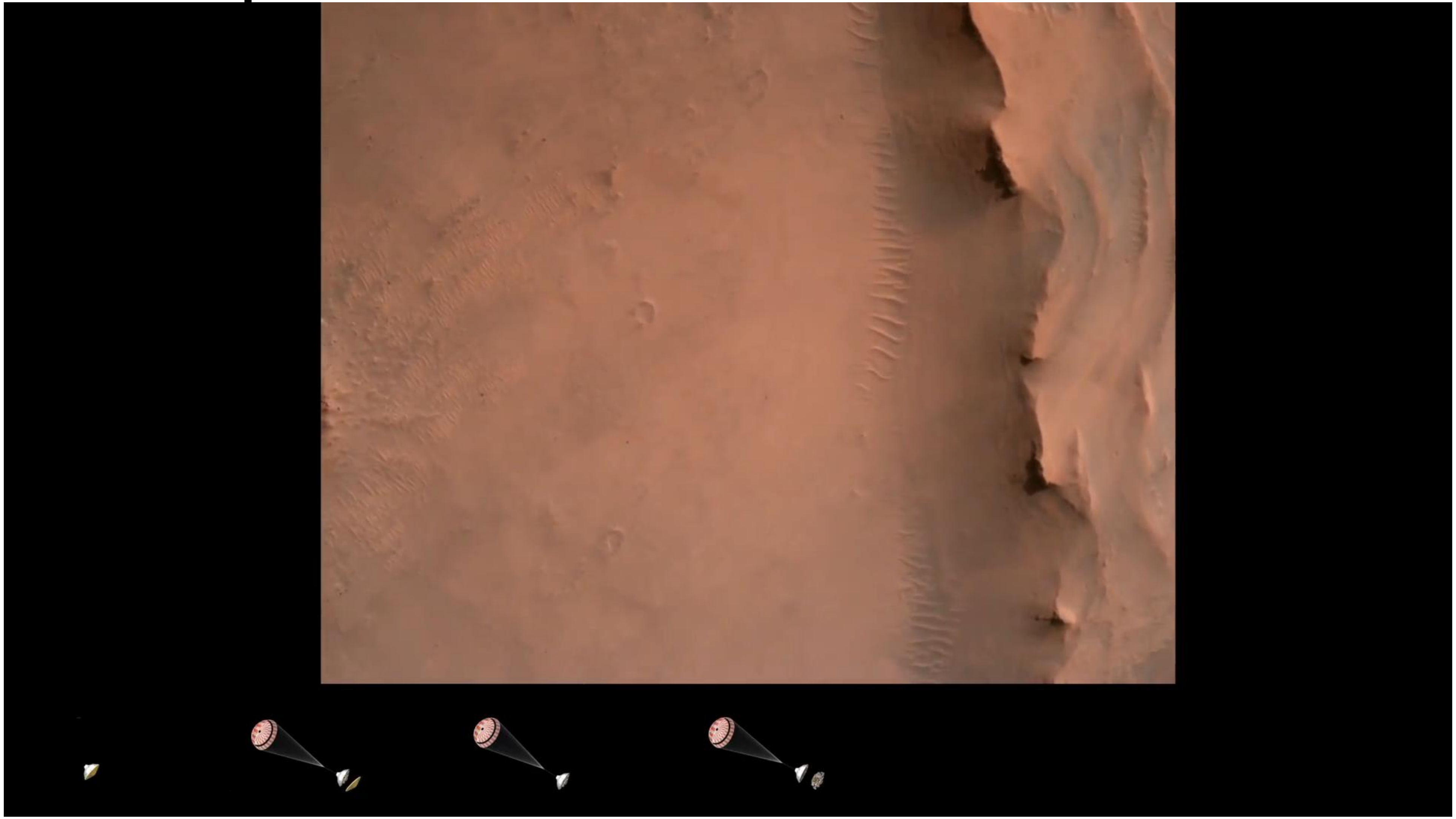
Applications- Object recognition in Supermarkets

Applications- Biometrics

Applications- Login Without Password



Space Exploration



NASA. https://www.youtube.com/watch?v=4czjS9h4Fpg

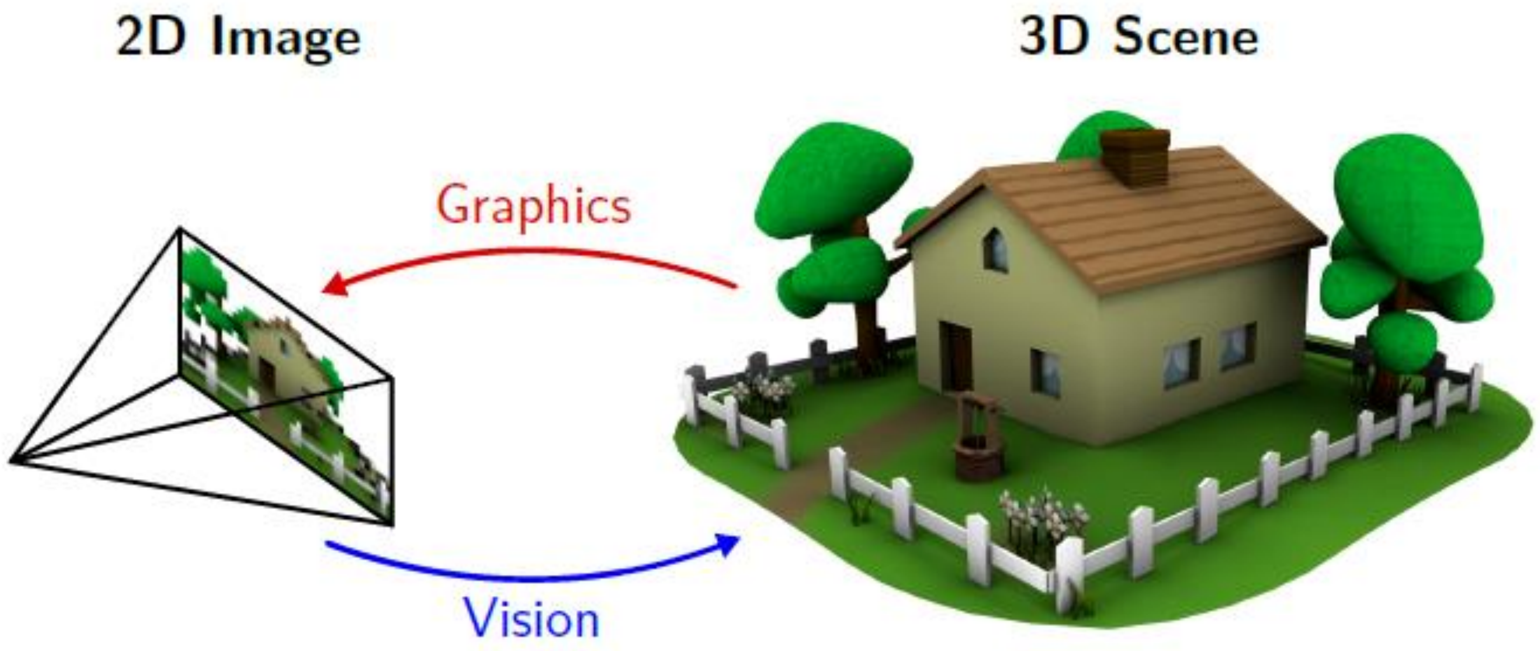
Vision as a poster child for success in AI

Insurance Institute for Highway Safety (US) – non-profit. 2024 review of autonomous emergency braking systems in Honda CR-V (350k sold in 2024)

Vision as a poster child for success in AI

Boston Dynamics (2017)

Computer Vision vs. Computer Graphics

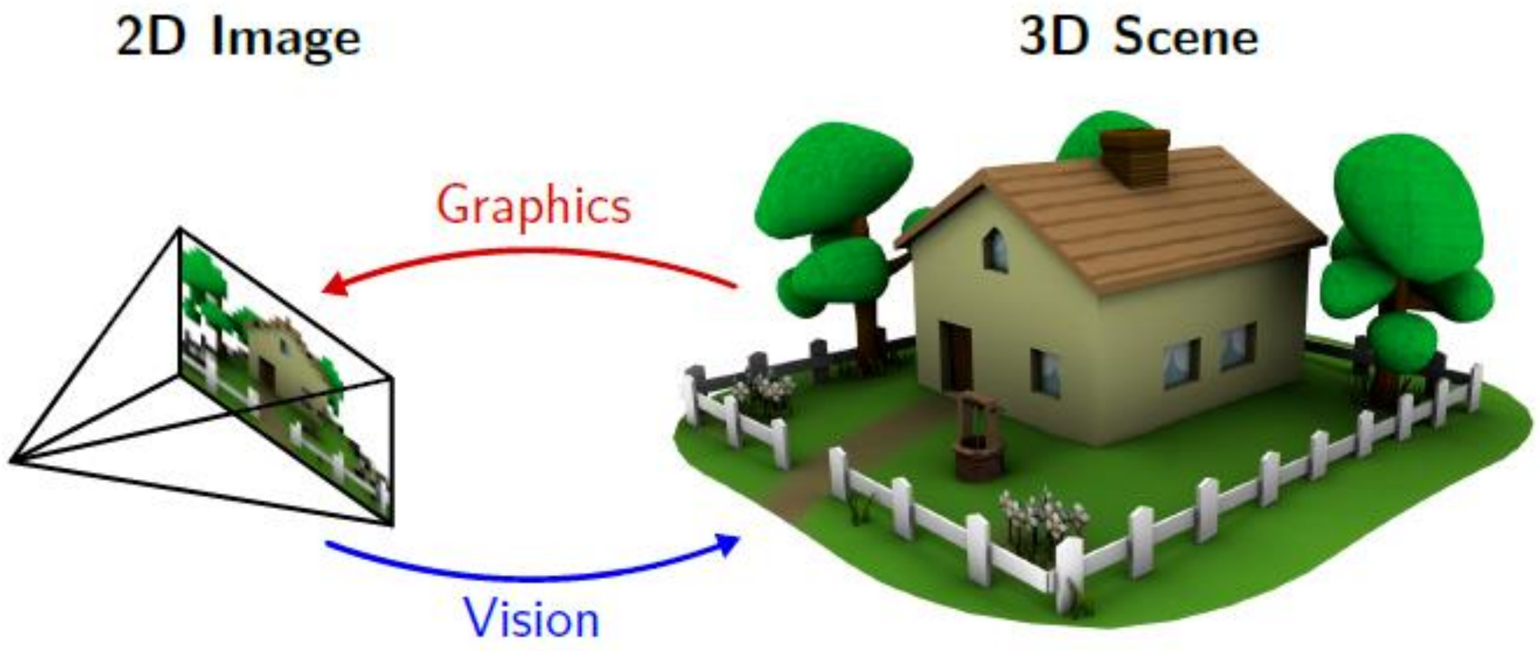


Pixel Matrix

217	191	252	255	239
102	80	200	146	138
159	94	91	121	138
179	106	136	85	41
115	129	83	112	67
94	114	105	111	89

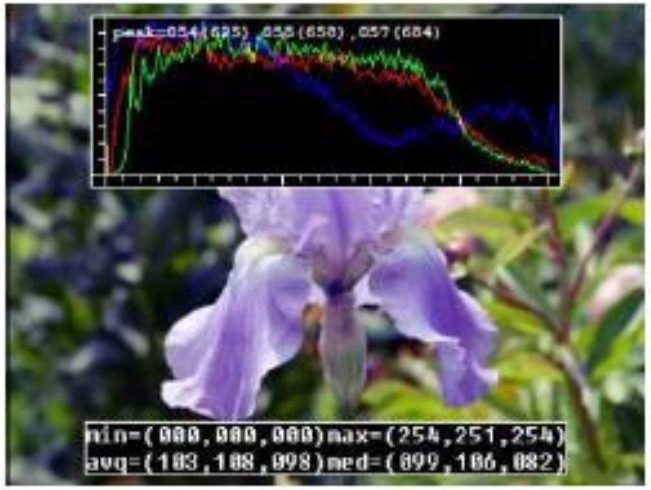
Objects Material
Shape/Geometry Motion
Semantics 3D Pose

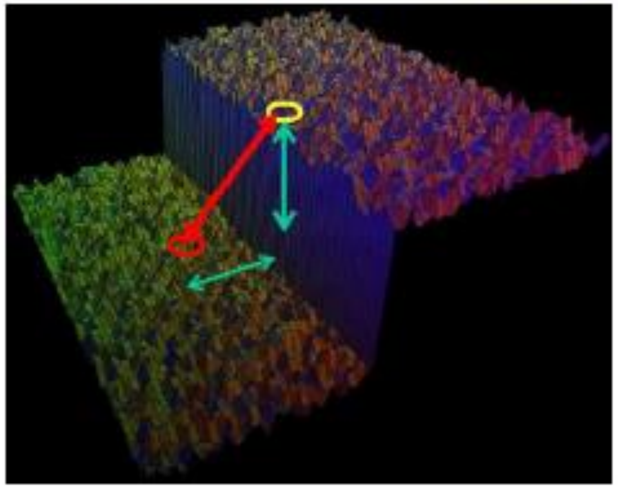
Computer Vision vs. Computer Graphics

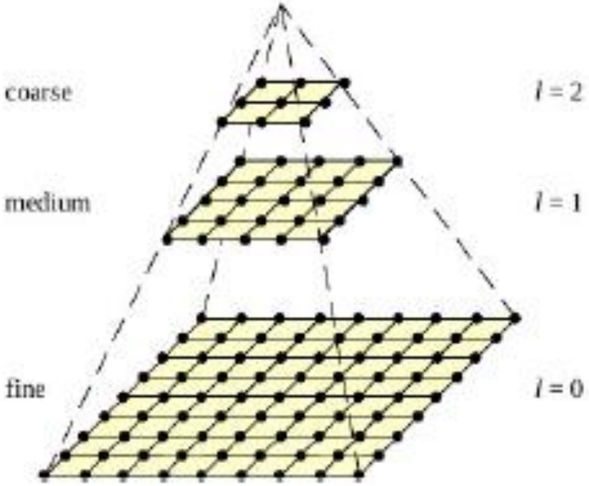


- **Computer Vision is an ill-posed inverse problem:**
  - Many 3D scenes yield the same 2D image
  - Additional constraints (knowledge about world) required

Computer Vision vs. Image Processing



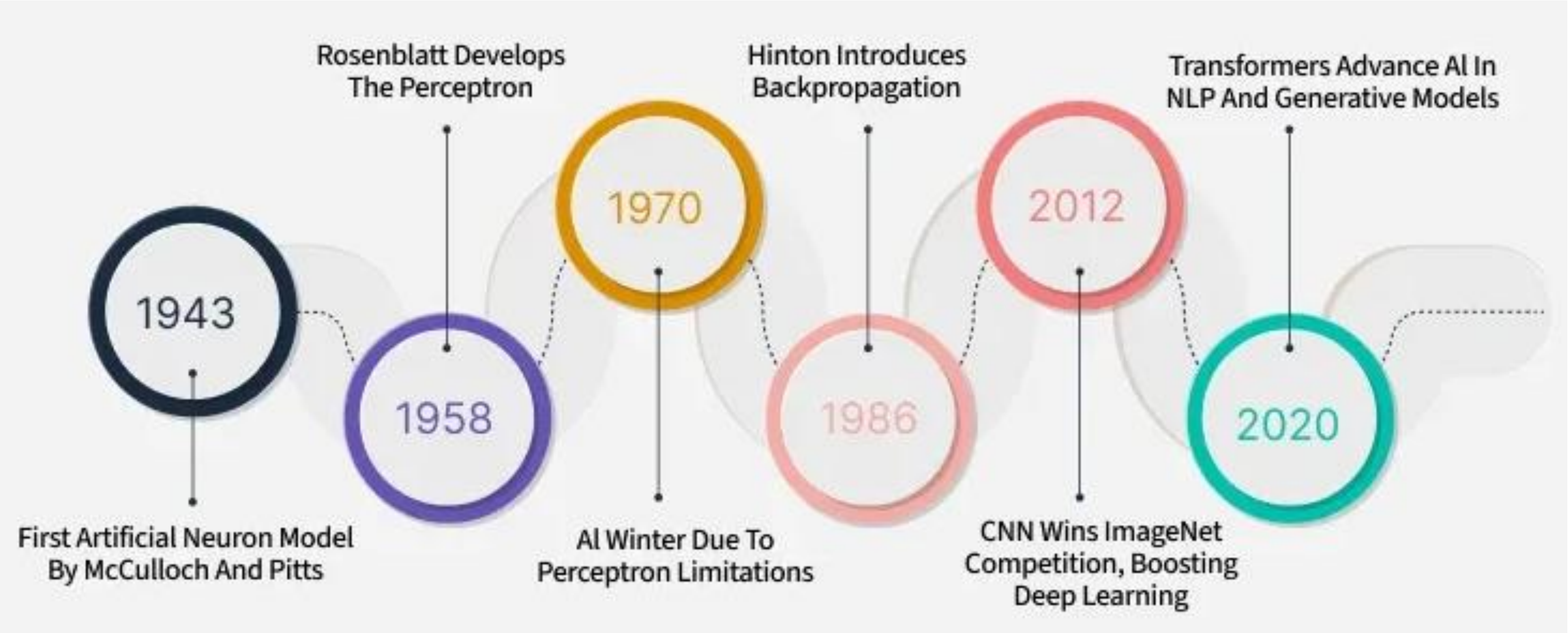




Slide Credits: Rick Szeliski

Computer Vision vs. Machine Learning

The Deep Learning Revolution



Why is Visual Perception hard?

- Paper: The Summer Vision Project. MIT AI Memos, 1966.

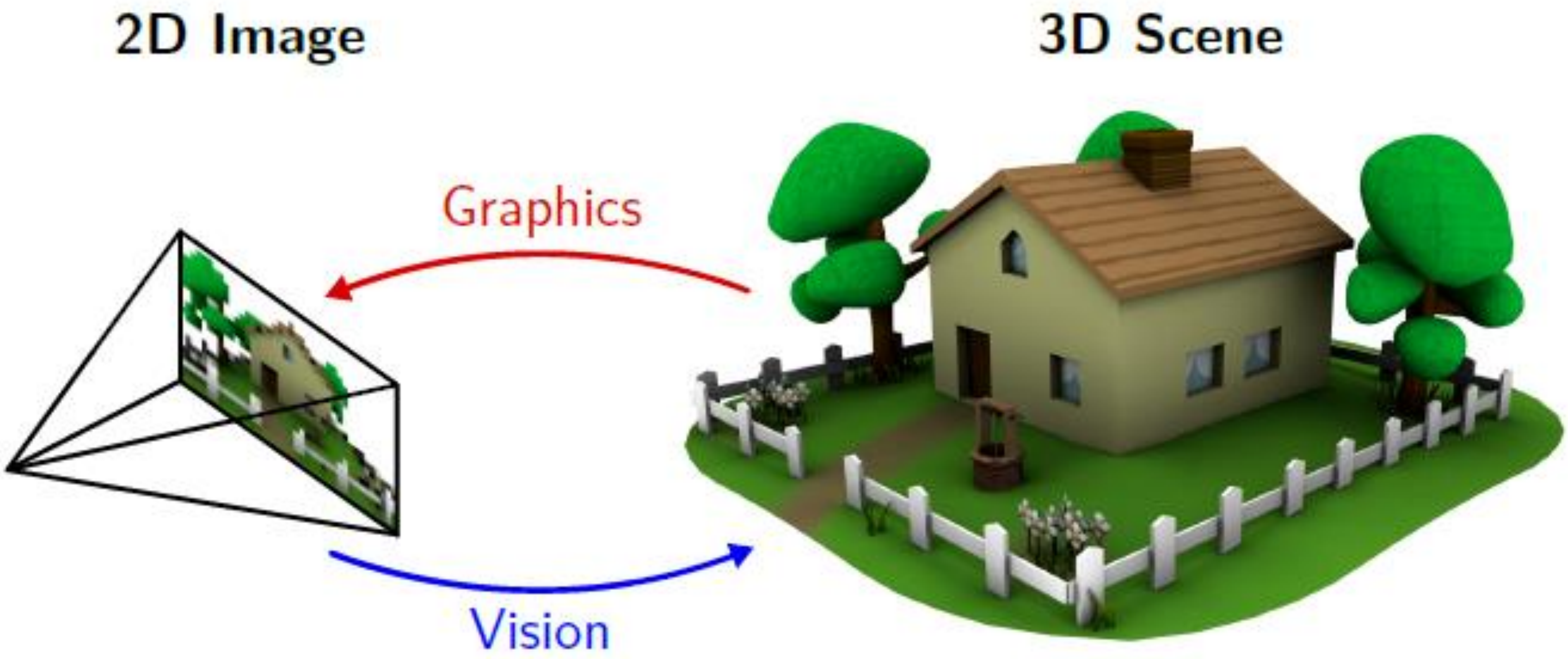
Why is Visual Perception hard?

What we see

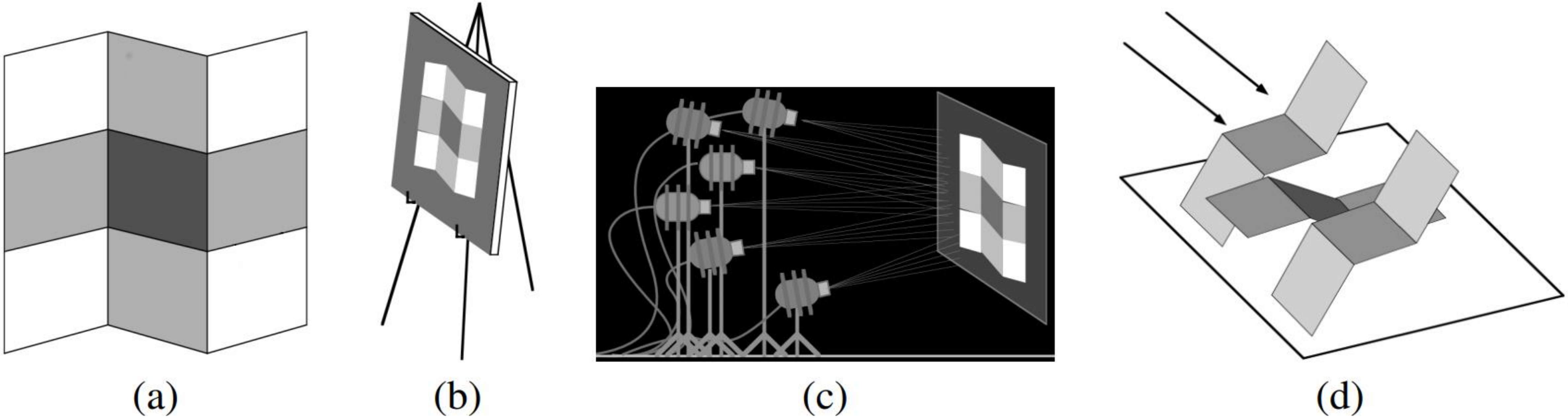
200	133	110	103	117	90	47	30	32	79	68	65
197	122	123	138	98	100	46	45	22	11	43	55
140	116	165	159	90	56	58	47	26	13	54	102
132	148	119	108	123	57	64	46	21	22	79	94
125	121	80	143	101	55	61	38	20	21	81	65
50	71	74	63	52	39	41	39	32	26	97	66
51	59	62	44	40	40	36	28	27	31	29	44
59	62	70	50	48	35	34	35	26	21	24	32
49	59	65	64	58	34	40	28	26	21	23	124
39	45	47	64	54	34	40	24	19	47	133	207
37	42	39	38	39	50	75	74	105	170	197	167
37	47	33	35	50	108	162	184	184	157	125	112
45	48	35	37	75	148	183	156	63	91	91	116
49	48	54	50	75	158	110	66	74	128	155	149
48	51	57	50	65	91	79	92	101	105	132	132
51	58	66	55	58	52	91	91	88	115	158	174
57	60	61	52	56	61	60	55	92	146	188	190
65	50	54	56	57	51	54	56	80	115	177	187
67	40	40	61	65	48	39	30	36	75	151	181
53	32	36	35	61	43	37	26	29	35	126	189
29	42	107	20	28	41	40	26	30	36	113	200
30	21	32	24	34	37	33	23	25	39	105	171
32	28	19	23	29	36	47	89	132	169	183	128
31	25	62	54	47	44	81	190	227	231	206	155
44	66	99	72	67	63	89	128	127	115	109	157
53	47	47	41	29	32	25	20	41	81	89	175
38	44	61	73	54	48	37	87	90	111	126	189
39	41	83	97	86	91	74	134	131	153	143	185
42	56	98	102	112	111	94	137	121	141	146	181
94	114	114	114	122	113	77	117	117	154	149	169
157	176	116	121	130	139	103	161	148	180	145	125
143	178	182	178	139	153	129	168	175	187	170	152
127	183	203	197	153	164	143	180	195	182	165	211
88	107	127	125	101	107	100	123	149	186	167	215

What the computer sees

Challenges: Images are 2D Projections of the 3D World



Challenges: Images are 2D Projections of the 3D World



- **Adelson and Pentland's workshop metaphor:**
  - To explain an image (a) in terms of reflectance, lighting and shape, (b) a painter, (c) a
  - light designer and (d) a sculptor will design three different, but plausible, solutions.

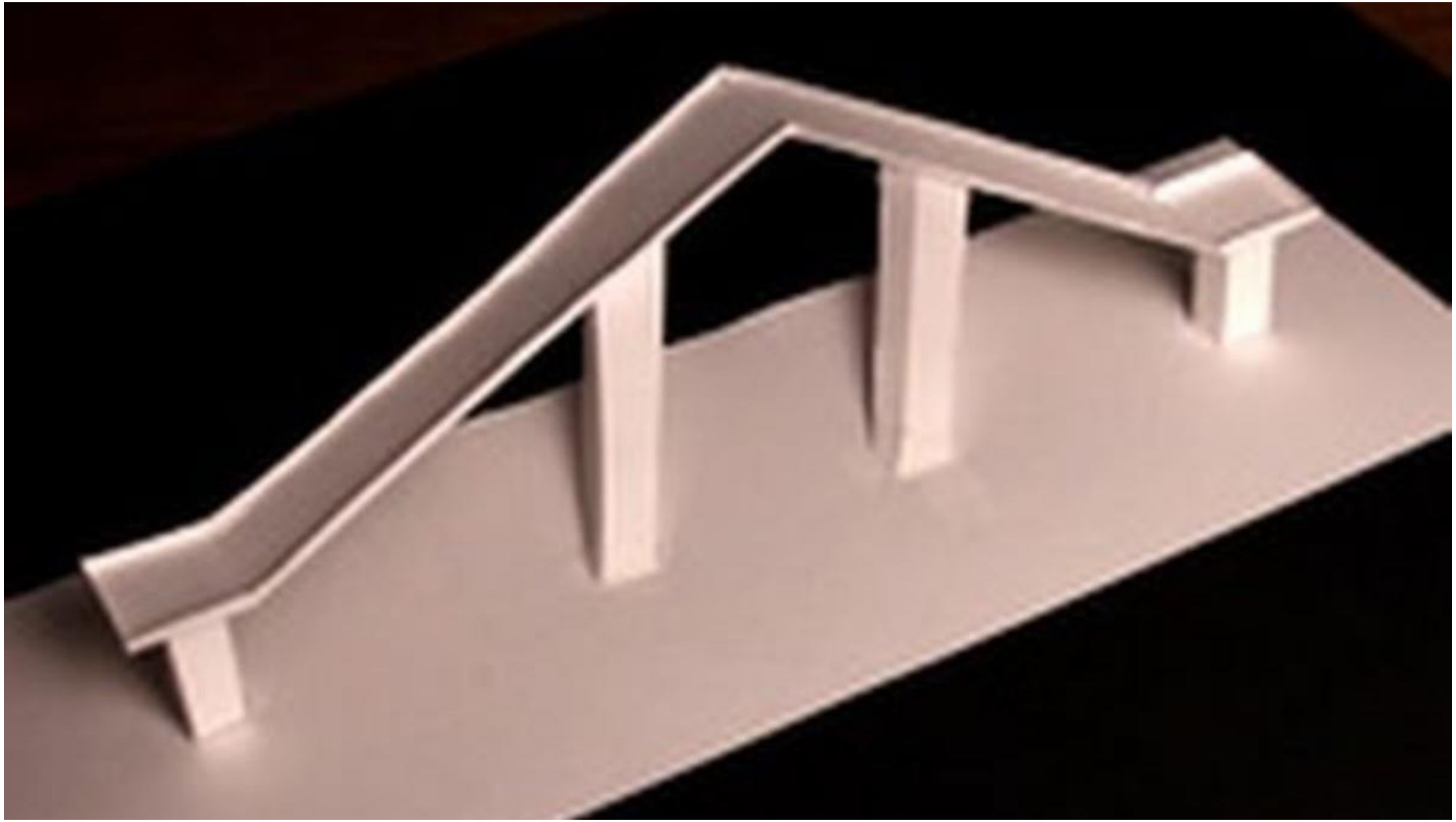
Adelson and Pentland: The perception of shading and reflectance. Perception as Bayesian inference, 1996.

Ames Room Illusion

Ames Room Illusion

https://www.youtube.com/watch?v=gJhyu6nlGt

Perspective Illusion

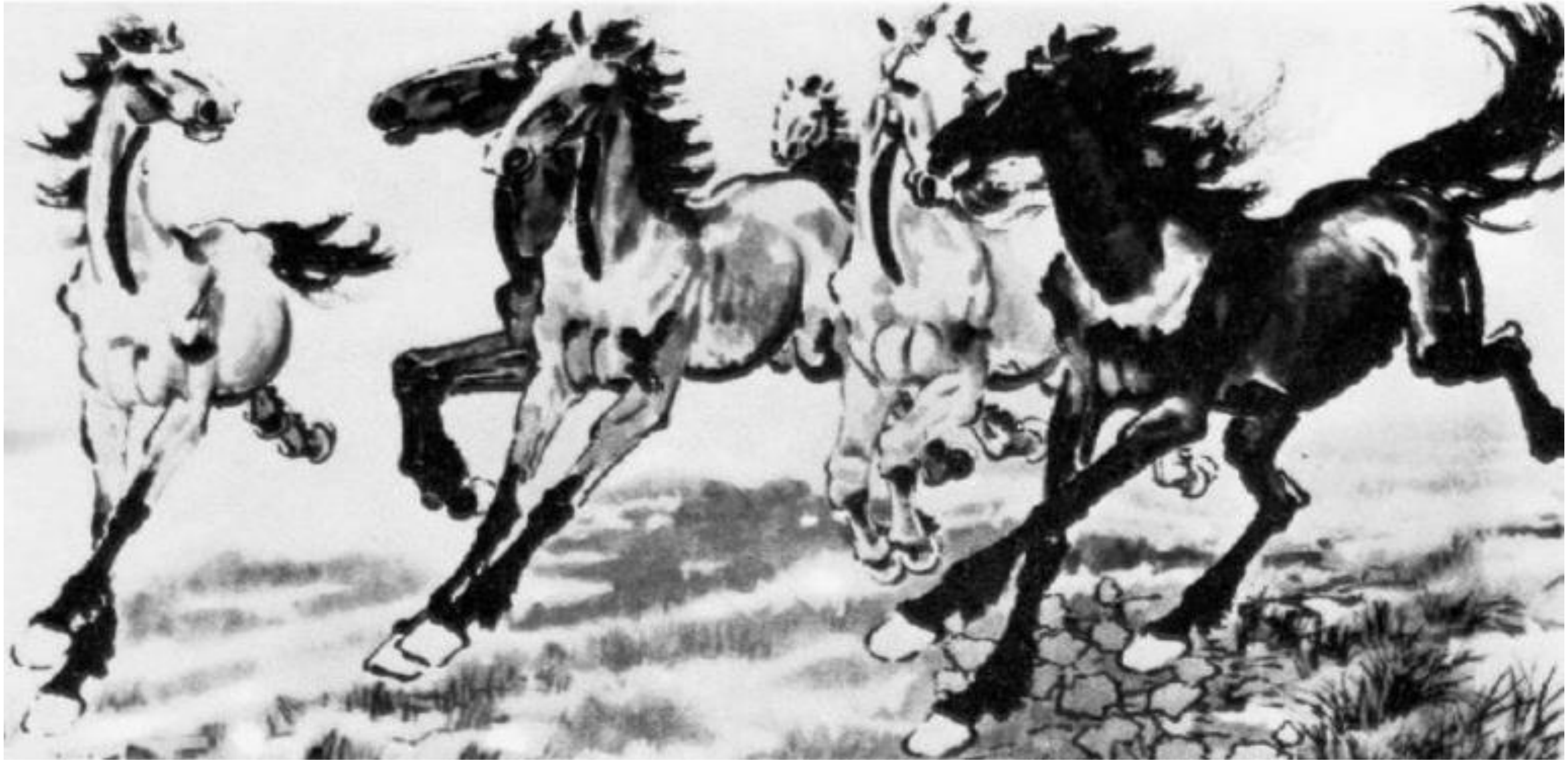


https://www.youtube.com/watch?v=vmkaVoLoFEU

Challenges: Viewpoint Variation

Michelangelo (1475-1564)

Challenges: Deformation



Michelangelo (1475-1564)

Challenges: Illumination

Slide Credits: Antonio Torralba

Challenges: Illumination

Slide Credits: Antonio Torralba

Challenges: Illumination

Slide Credits: Antonio Torralba

Challenges: Illumination

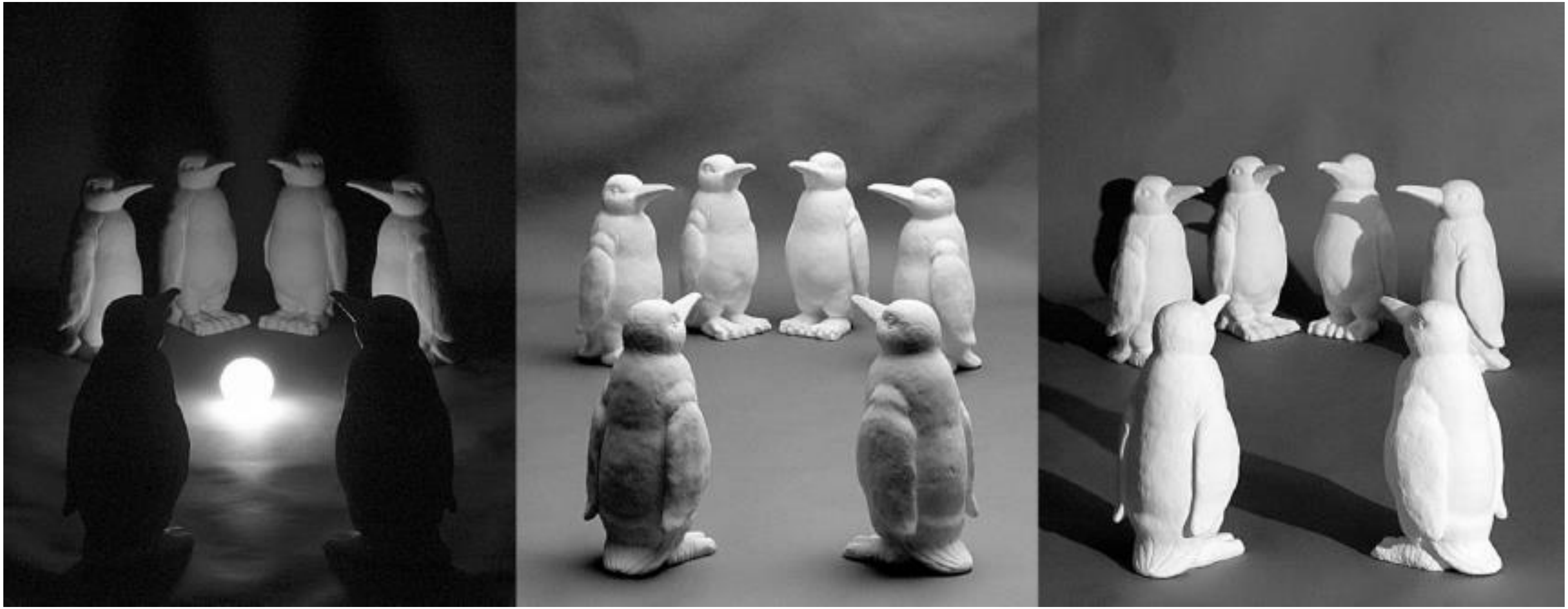
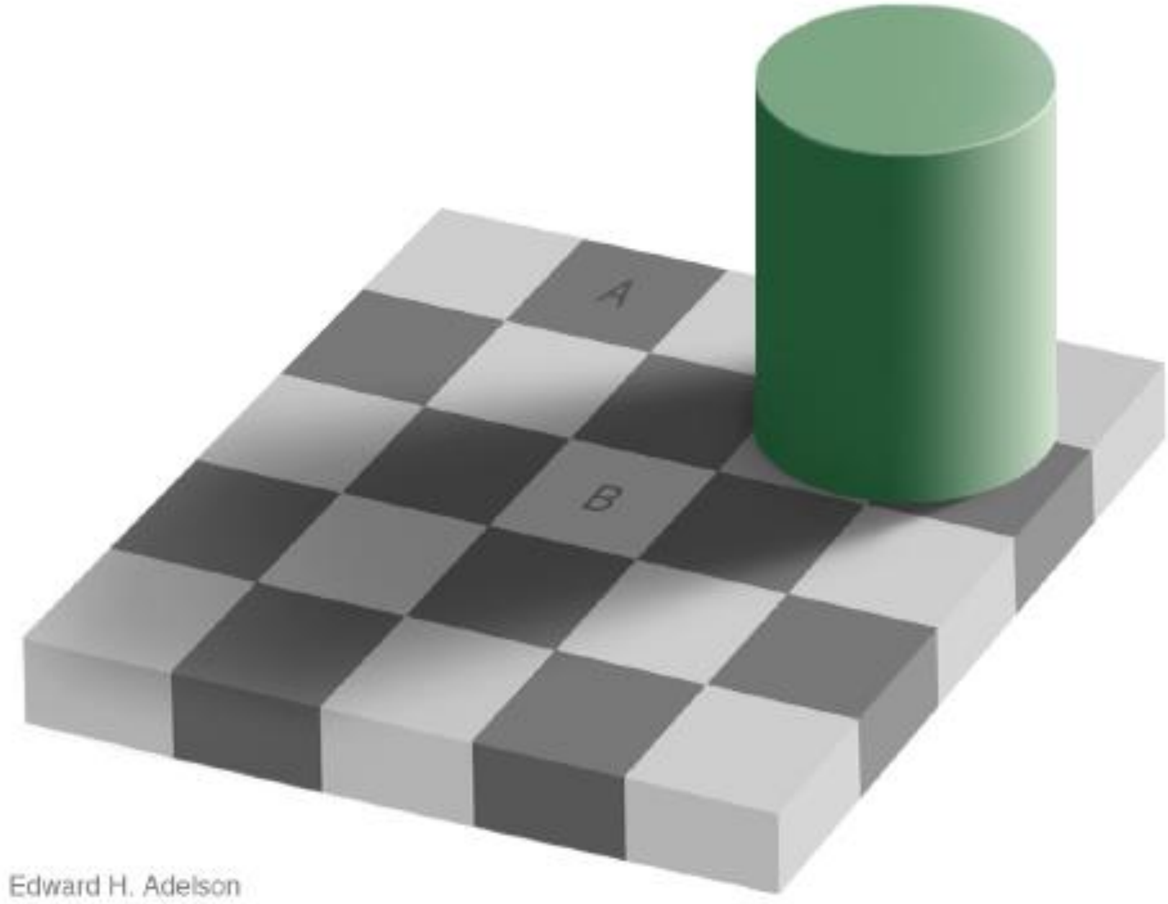


Image Credits: Jan Koenderink

Challenges: Motion

Challenges: Perception vs. Measurement



http://persci.mit.edu/gallery/checkershadow

Image Credits: Edward H. Adelson

Challenges: Perception vs. Measurement

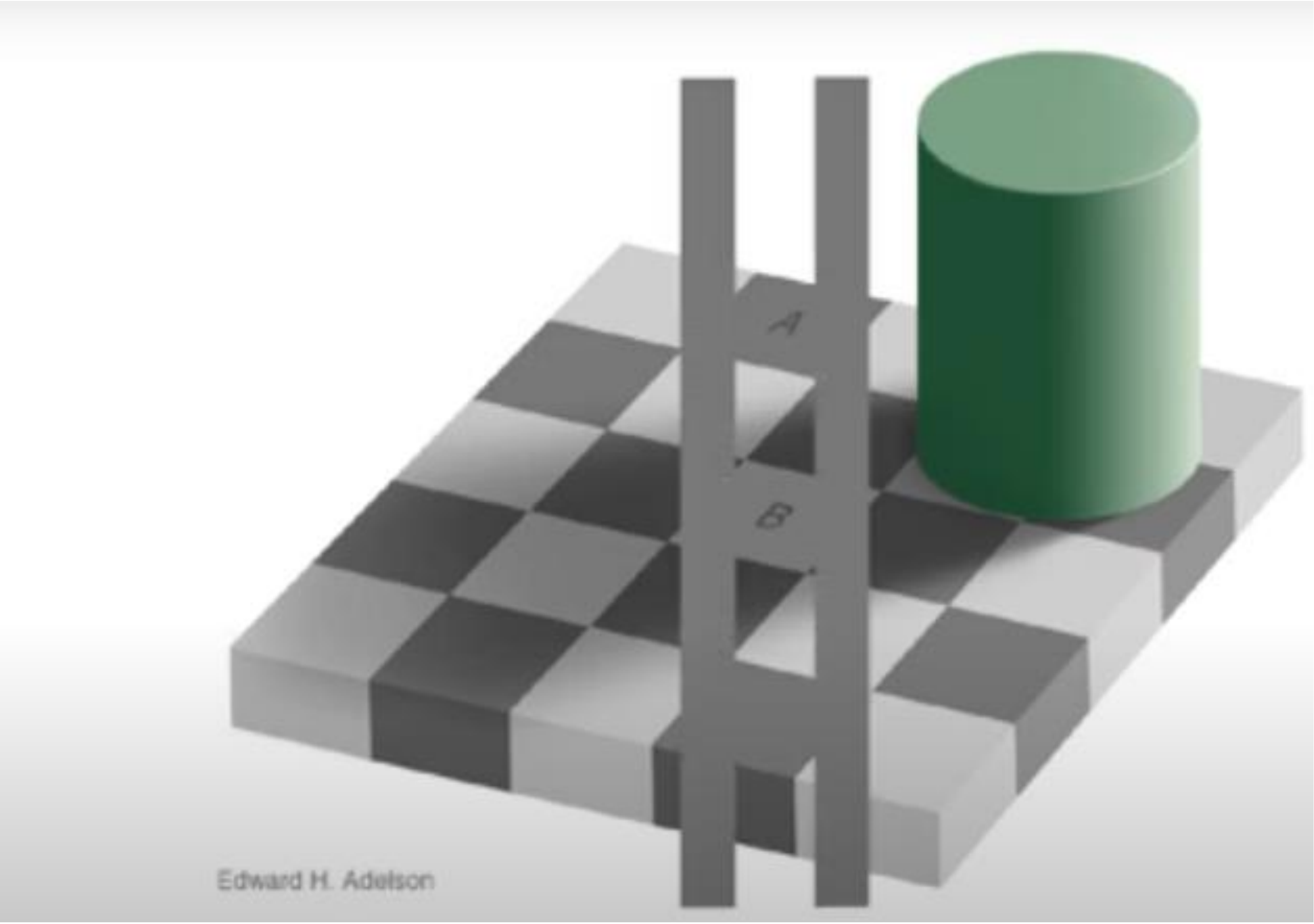
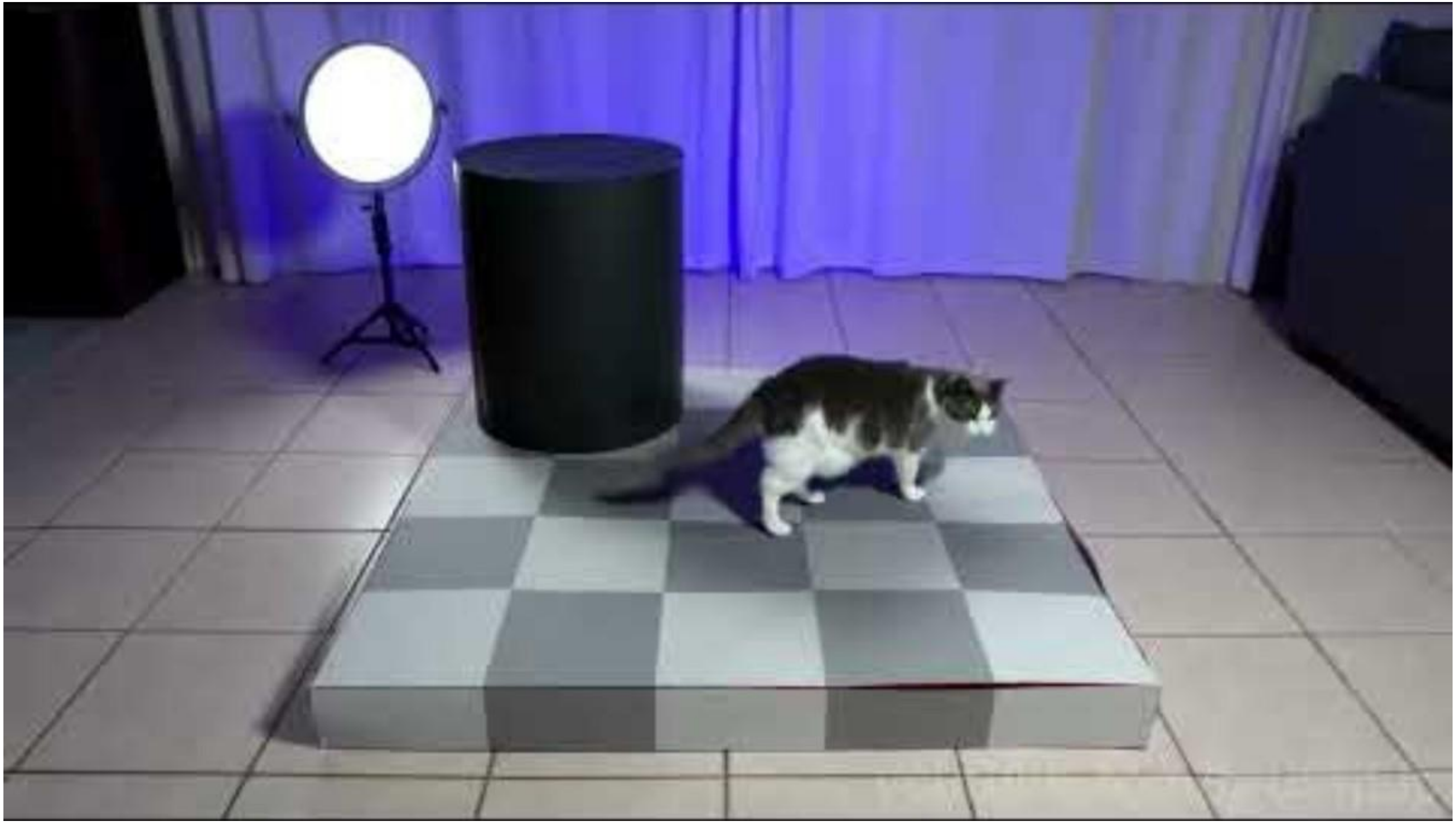
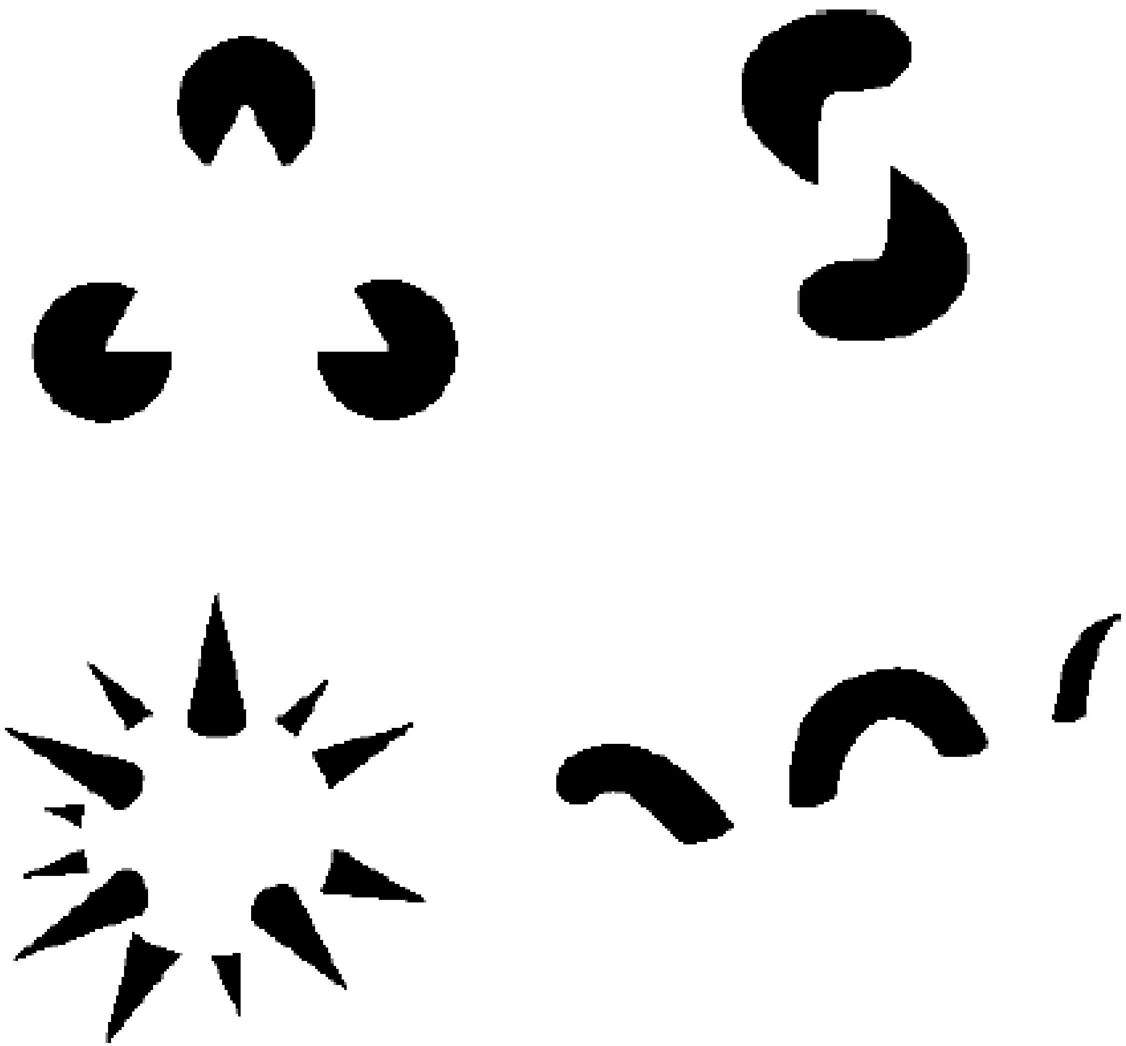


Image Credits: Edward H. Adelson

Challenges: Perception vs. Measurement



Challenges: Perception vs. Measurement



Gaetano Kanizsa (1976)

Challenges: Local Ambiguities

Image Credits: Antonio Torralba

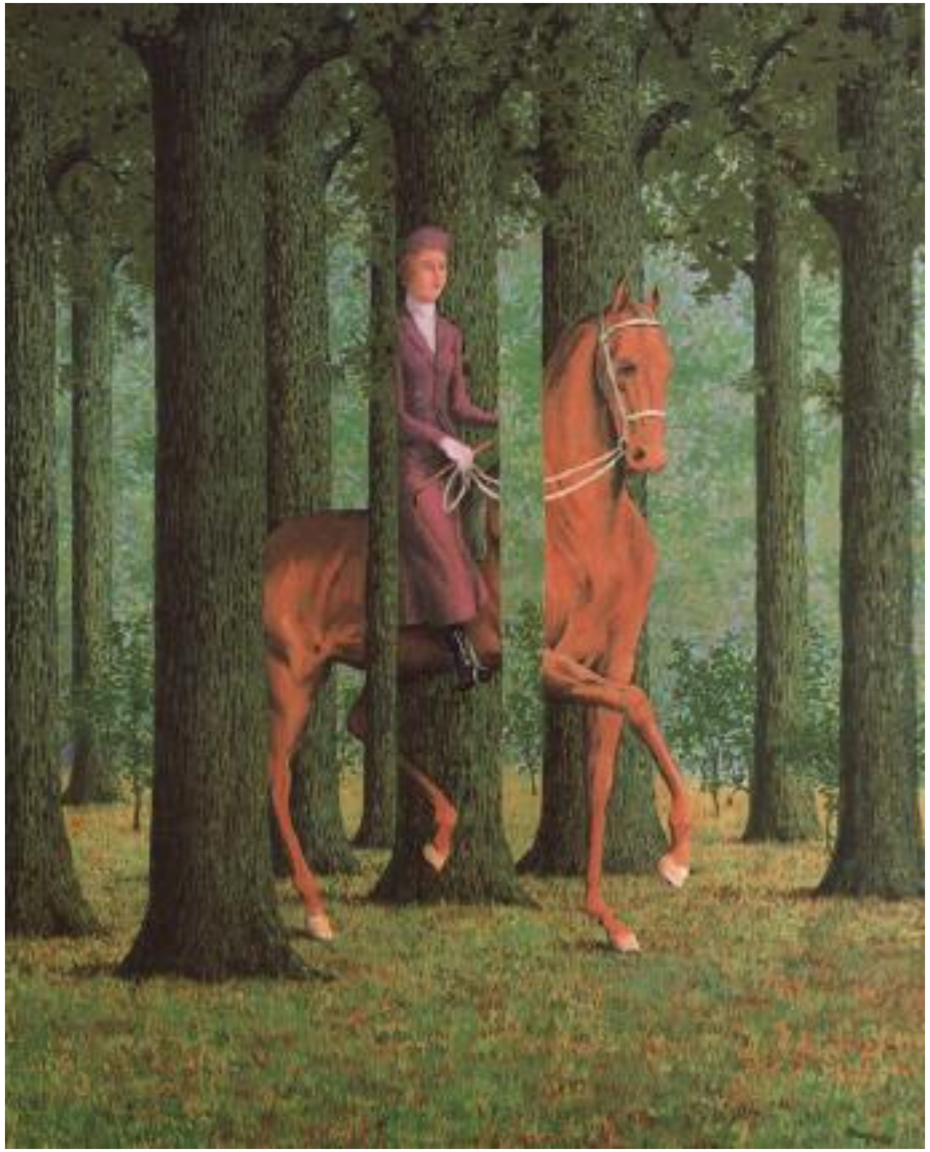
Challenges: Local Ambiguities

Image Credits: Antonio Torralba

Challenges: Local Ambiguities

Image Credits: Antonio Torralba

Challenges: Occlusion



Rene Magritte (1957)

Challenges: Intra Class Variation

Image Credits: Antonio Torralba

Challenges: Number of Object Categories

Image Credits: Antonio Torralba