CS4501: Introduction to Computer Vision

Introduction

Various slides from previous courses by:

D.A. Forsyth (Berkeley / UIUC), I. Kokkinos (Ecole Centrale / UCL). S. Lazebnik (UNC / UIUC), S. Seitz (MSR / Facebook), J. Hays (Brown / Georgia Tech), A. Berg (Stony Brook / UNC), D. Samaras (Stony Brook) . J. M. Frahm (UNC), V. Ordonez (UVA).

Today's Class

Who am I?

What is Computer Vision?

Why is Computer Vision Hard?

Cameras

Questions

About Me

Vicente

About Me

Assistant Professor
2016 - Now

Visiting Professor
2019

Adobe Research

Visiting Researcher
2015 - 2016

ALLEN INSTITUTE
for ARTIFICIAL INTELLIGENCE

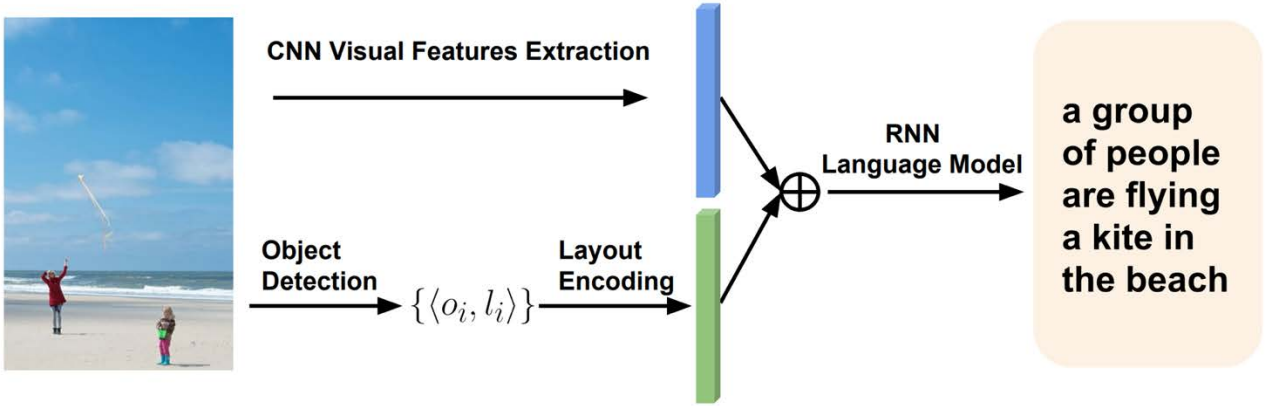
MS, PhD in CS,
2009-2015

THE UNIVERSITY
of NORTH CAROLINA
at CHAPEL HILL

Stony Brook University

… also spent time at:

Describing Images with Language



NEW! Obj2Text: Generating Visually Descriptive Language from Object Layouts
Xuwang Yin, Vicente Ordonez.
Empirical Methods in Natural Language Processing. **EMNLP 2017**. Copenhagen, Denmark. September 2017.

Large Scale Retrieval and Generation of Image Descriptions
V. Ordonez, X. Han, P. Kuznetsova, G. Kulkarni, M. Mitchell, K. Yamaguchi, K. Stratos, A. Goyal, J. Dodge, A. Mensch, H. Daume III, A.C. Berg, Y. Choi, T.L. Berg.
International Journal of Computer Vision. **IJCV 2015**. [August 2016 Issue]. [pdf] [link] [bibtex]

Im2Text: Describing Images Using 1 Million Captioned Photographs
Vicente Ordonez, Girish Kulkarni, Tamara L. Berg.
Advances in Neural Information Processing Systems. **NIPS 2011**. Granada, Spain. December 2011.

Naming Objects

Superordinates: animal, vertebrate
Basic Level: bird
Entry Level: bird
Subordinates: American robin

Superordinates: animal, vertebrate
Basic Level: bird
Entry Level: penguin
Subordinates: Chinstrap penguin

From Large Scale Image Categorization to Entry-Level Categories
Vicente Ordonez, Jia Deng, Yejin Choi, Alexander C. Berg, Tamara L. Berg.
IEEE International Conference on Computer Vision. **ICCV 2013**. Sydney, Australia. December 2013.

Learning to Name Objects
Vicente Ordonez, Wei Liu, Jia Deng, Yejin Choi, Alexander C. Berg, Tamara L. Berg.
Communications of the ACM. March 2016 (Vol. 59, No. 3). *(~Research Highlight)*

Predicting Entry-Level Categories
Vicente Ordonez, Wei Liu, Jia Deng, Yejin Choi, Alexander C. Berg, Tamara L. Berg.
International Journal of Computer Vision - Marr Prize Special Issue. **IJCV 2015**.

Recognizing Commonly Uncommon Situations

Query

Similar in imSitu train set






Predicted situations

falling				0.58372
agent	source	goal	place	
person	horse	land	outdoors	

Commonly Uncommon: Semantic Sparsity in Situation Recognition
Mark Yatskar, Vicente Ordonez, Luke Zettlemoyer, Ali Farhadi.
Intl. Conference on Computer Vision and Pattern Recognition. **CVPR 2017**. Honolulu, Hawaii. July 2017.

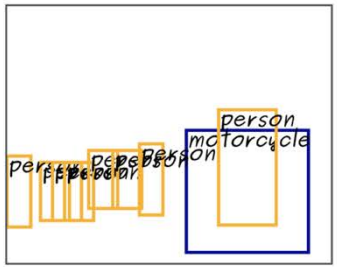
http://imsitu.org/demo/

Synthesizing Images from Text

Mike is surprised at the duck. The duck is standing on the grill. Jenny is running towards Mike and the duck.

Abstract Scene[1]

A guy on a motorcycle with some people watching.



Object Layout[2]

Several elephants walking together in a line near water.

Synthetic Image[2]

NEW! Text2Scene: Generating Compositional Scenes from Textual Descriptions
Fuwen Tan, Song Feng, Vicente Ordonez.
Intl. Conference on Computer Vision and Pattern Recognition. **CVPR 2019**.
Long Beach, California. June 2019. [arxiv] [bibtex] *(~Oral presentation + Best Paper Finalist -- top 1% of submissions)*

Vision, Language and Learning @ UVA

vislang | University of Virginia

home people demos publications

The **vis**ion, **lang**uage and learning lab, *vislang*, at the University of Virginia pursues fundamental research at the intersection of computer vision, natural language processing and machine learning. We aim to create intelligent systems that can learn from vast amounts of visual and textual information, that can integrate and enhance human experiences, and that can resolve complex tasks that typically require human intelligence.

Read about some of our work on bias in visual recognition in WIRED and Glamour. Some of our recent work on analyzing movies on TechXplore, and our work on generating images from text in the blogs of IBM and NVIDIA.

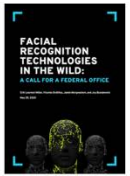
https://vislang.ai/

News and Announcements

- 12/2020. Paola and Fuwen had Curriculum Labeling accepted to AAAI.
- 09/2020. Tianlu, Ziyan and Leticia got papers accepted to EMNLP and Findings of EMNLP 2020 [publications].
- 06/2020. Andrew Ng's deeplearning.ai has a blog post highlighting our ACL 2020 paper with Salesforce Research on Double-Hard Debias [link].
- 06/2020. TechXplore features our work on regulation of Face Recognition in Researchers call for new federal authority to regulate facial recognition tech.
- 05/2020. With a group of colleagues and funding from the MacArthur Foundation, we released the whitepaper Face Recognition Technologies in the Wild: A Call for a Federal Office.
- 05/2020. Our group recently received a Facebook Research Award 2020 and gift funding from eBay Research and Adobe Research.

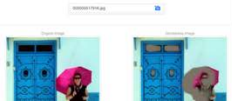
Facial Recognition Technologies in the Wild

With colleagues Erik-Learned Miller, Jamie Morgenstern, and Joy Buolamwini we make the case that benchmarks fall short in assessing Face Recognition Technologies (FRTs) for deployment and we propose a more holistic approach.



Genderless

This demo attemps to make it difficult for a model to predict gender from an image by modifying it so that this task



Welcome to Introduction to Computer Vision

UNIVERSITY of VIRGINIA

CS4501: Introduction to Computer Vision | Spring 2021

Instructor: Vicente Ordóñez-Román (vicente at virginia.edu), Office Hours: TBD
Class Time: Tuesdays & Thursdays between 12:30PM and 1:45PM (ET).
Discussion Forum: TBD

Course Description: Computer Vision is about empowering computers to visually explore and reason about the world. In this course we will study how images are represented in a computer, how to manipulate them, and how to extract information (features) from images for various applications including image matching, reconstruction, and recognition. We will particulary study basic techniques to recover geometry from images, camera calibration, automatic image alignment, boundary detection, object recognition, image classification, and image retrieval/search. We will develop the intuitions, study the foundations, techniques, methods, and underlying concepts, and learn to use these in practical computer vision systems.

Learning Objectives: (a) Develop intuitions between human vision and computer vision, (b) Understanding the basics of 2D and 3D techniques, (c) Become familiar with the techniques such as registration, matching, and recognition, and (d) Obtain practical experience in the implementation of computer vision applications.

Prerrequisites: It is recommended that students have a basic command of linear algebra, calculus, and statistics. Students are encouraged to complete this [Primer on Image Processing] and familiarize themselves with Jupyter notebooks or Google Colab. Python programming is also a requirement.

Textbook: No required textbook, but students are strongly encouraged to read chapters from: "Computer Vision: Algorithms and Applications" by Richard Szeliski. The book is available for free online or available for purchase. Also optional readings from David Forsyth & Jean Ponce's "Computer Vision: A Modern Approach".

Schedule

Date	Topic
Tue, Feb 2nd	Introduction to Computer Vision
Thu, Feb 4th	Cameras and Image Formation
	Assignment 1: Image Formation, Image Processing, and Image Filtering.
Tue, Feb 9th	Projective Geometry and Light
Thu, Feb 11th	Human Vision and Image Processing
Tue, Feb 16th	Image Filtering and Image Frequencies
Thu, Feb 18th	Frequencies and Edges
	Assignment 2: Detecting Corners and Lines, Hough Transform.
Tue, Feb 23th	Interest Points: Corners and Blobs

https://www.vicenteordonez.com/vision/

About the Course

CS4501-008: Introduction to Computer Vision

- Instructor: Vicente Ordóñez
- Email: vicente@virginia.edu
- Website: http://vicenteordonez.com/vision/
- Class Location: **Virtual (Zoom) – Links are on UVA Collab**
- Class Times: **Tuesday-Thursday 12:30pm – 1:45pm (ET)**
- Discussion Forum: Piazza? Campuswire?
- Office hours: TBD

Objectives

- Develop Intuitions Between Human Vision and Computer Vision
- Understanding the Basics of 2D and 3D Computer Vision
- Become familiar with the technical approaches in computer vision such as registration, matching, and recognition
- Obtain practical experience in the implementation of computer vision applications.

Pre-requisites

- Python programming skills
- Calculus / Linear Algebra / Probability

Grading

- Assignments: 100% (25%, 25%, 25%, 25%)
- Class Participation: +5% (extra points)
- Quiz: +5% (extra points)

Grade cutoffs: A+ (95%) AND complete all assignments on time and with good grades.

A (90%), A- (85%), B+ (80%), B (75%), B- (70%), C+ (65%), C (60%), C- (55%), D+ (50%), D (45%), D- (40%).

Textbook

- *Computer Vision: Algorithms and Applications*
  by Richard Szeliski. http://szeliski.org/Book/

Computer Vision: A Modern Approach, Forsyth and Ponce, Prentice Hall 2011.

Introductory Techniques for 3D Computer Vision, Trucco and Verri, Prentice Hall 1998.

Computer Vision: Models, Learning, and Inference by S. Prince, Cambridge University Press, 2012: draft at http://www.computervisionmodels.com/

What is Computer Vision?

Make computers understand images and video

What kind of scene?

Where are the cars?

How far is the building?

…

Slide by James Hays.

Why computer vision matters

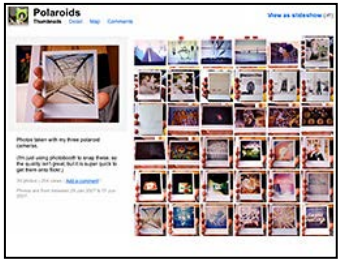
Safety

Health

Security

Comfort

Fun

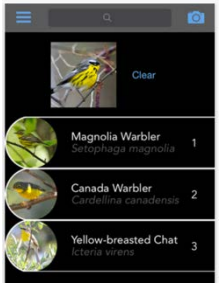


Access

kaggle

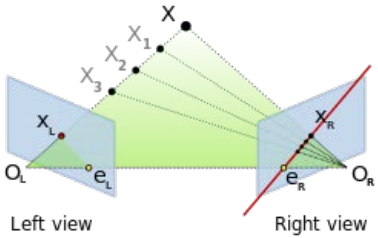
Create an algorithm to distinguish dogs from cats

Birdsnap



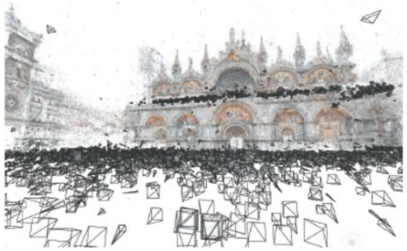
Face Detection in Cameras

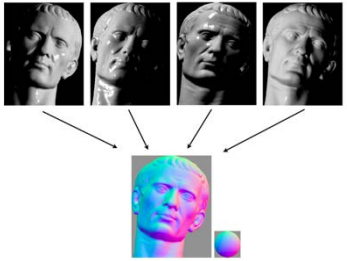
Computer Vision

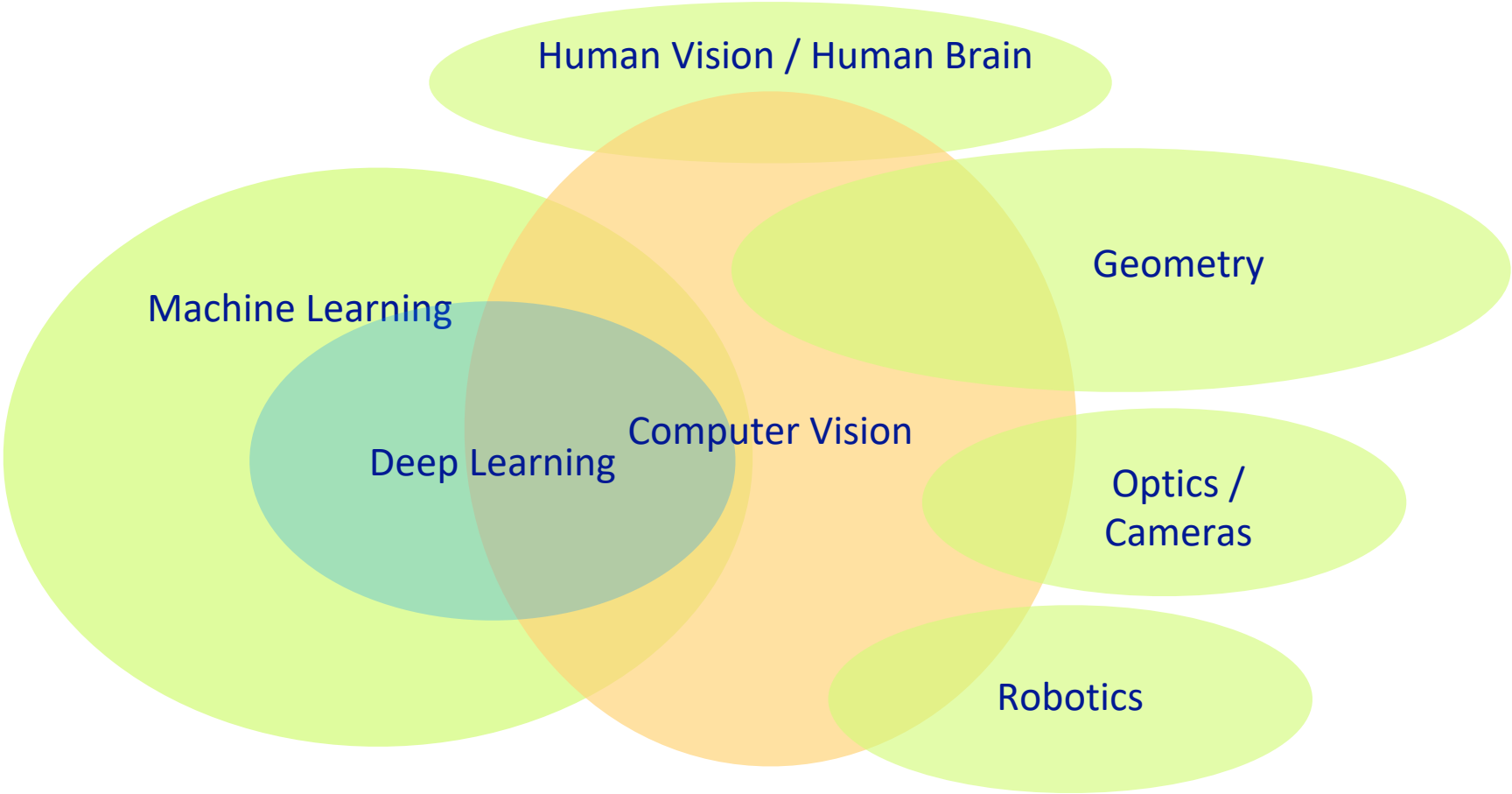


Left view

Right view

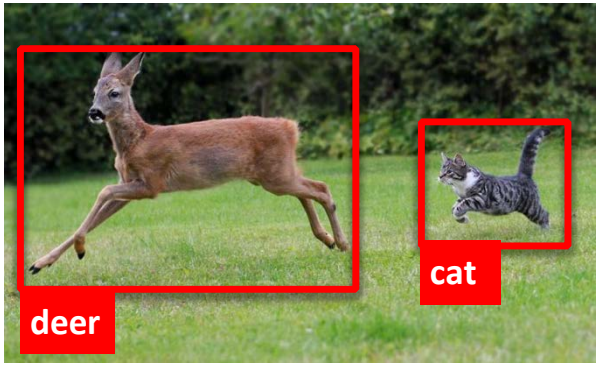





Human Vision / Human Brain
Machine Learning
Geometry
Deep Learning
Computer Vision
Optics /
Cameras
Robotics

Relationship with Other Fields

- Computer Vision: Image ⟶ Knowledge



Relationship with Other Fields

- Image Processing: Image → Image

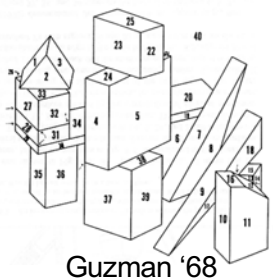
Relationship with Other Fields

- Computer Graphics: Knowledge ⟶ Image

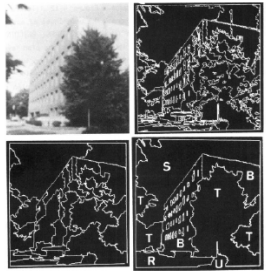
Vertices, Locations, Objects,
Shapes, Colors, Material properties,
Lighting settings, Camera settings, etc. ⟶

Ridiculously brief history of computer vision

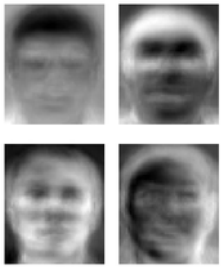
- 1960’s: interpretation of synthetic worlds
- 1970’s: some progress on interpreting selected images
- 1980’s: Neural Networks come and go; shift toward geometry
- 1990’s: face recognition; statistical analysis in vogue
- 2000’s: broader recognition; large annotated datasets available; video processing starts
- 2010’s: Deep learning with ConvNets
- 2030’s: ?



Guzman ‘68



Ohta Kanade ‘78



Turk and Pentland ‘91

How vision is used now

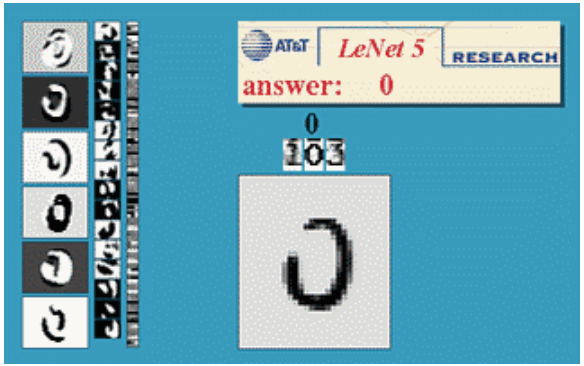
- Examples of real world applications

Some of the following slides by Steve Seitz

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



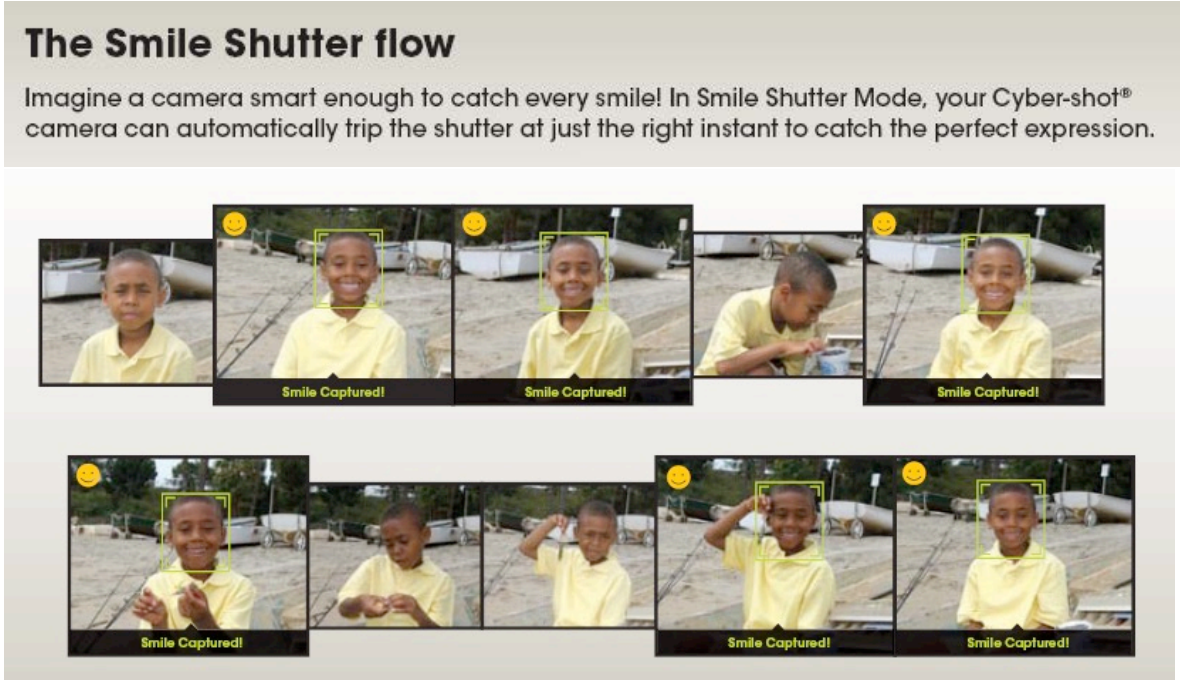
Digit recognition, AT&T labs
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Face detection

- Digital cameras detect faces

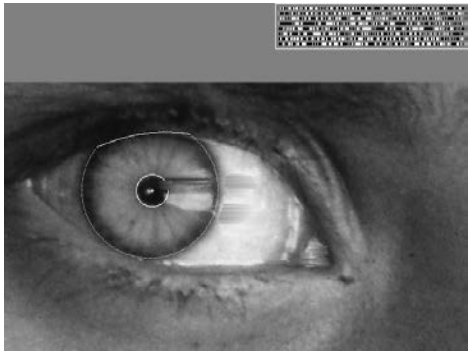
Smile detection

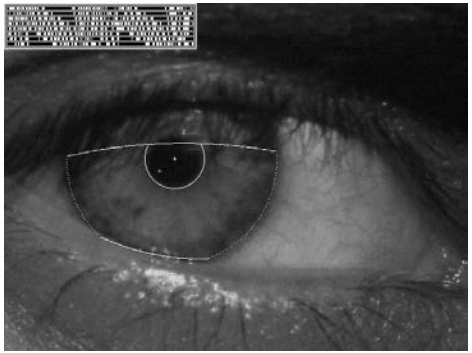


[Sony Cyber-shot® T70 Digital Still Camera]

Vision-based biometrics

"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the story wikipedia





Login without a password…

Fingerprint scanners on many new laptops, other devices

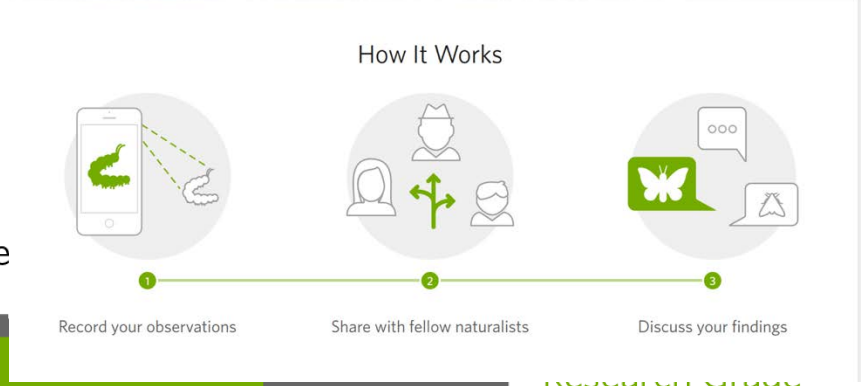
Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

Object recognition (in mobile phones)

[Point & Find](), [Nokia]()

[Google Goggles]()

iNaturalist



https://www.inaturalist.org/pages/computer_vi

Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Special effects: motion capture

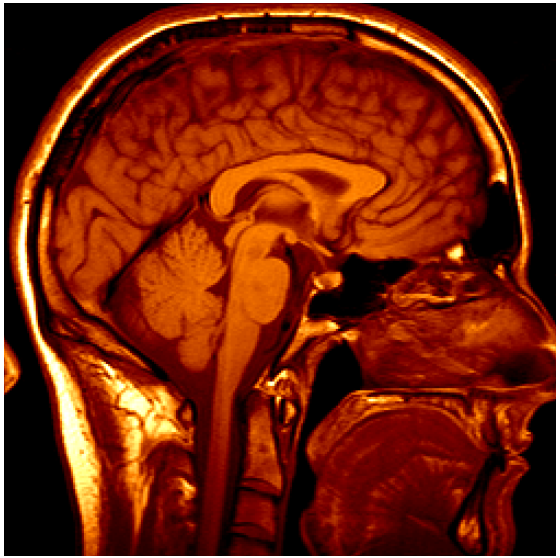
Pirates of the Carribean, Industrial Light and Magic

Sports

Sportvision first down line
Nice explanation on www.howstuffworks.com

http://www.sportvision.com/video.html

Medical Imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Slide content courtesy of Amnon Shashua

Smart cars

- Mobileye
  - Market Capitalization: 11 Billion dollars

Self-driving Cars e.g. Google's Waymo

Oct 9, 2010. "Google Cars Drive Themselves, in Traffic". *The New York Times*. John Markoff
June 24, 2011. "Nevada state law paves the way for driverless cars". *Financial Post*. Christine Dobby
Aug 9, 2011, "Human error blamed after Google's driverless car sparks five-vehicle crash". *The Star* (Toronto)

Ford acquires SAIPS for self-driving machine learning and computer vision tech

Posted Aug 16, 2016 by Darrell Etherington (@etherington)

Ford outlined a few of the ways it's aiming to ship driverless cars by 2021, and part of the plan involves acquisitions. CEO Mark Fields revealed at a press event in Palo Alto today that the automaker acquired SAIPS, an Israeli company focusing on machine learning and computer vision. It's also partnering exclusively with Nirenberg Neuroscience, to bring more "humanlike intelligence" to machine learning components of driverless car systems.

SAIPS' technology brings image and video processing algorithms, as well as deep learning tech focused on processing and classifying input signals, all key ingredients in the special sauce that makes up autonomous vehicle tech. This company's expertise should help with on-board interpretation of data captured by sensors on Ford's self-driving cars, and turning that data into usable info for the car's virtual driver system. SAIPS' offerings include detection of anomalies, persistent tracking of objects detected by sensors, and much more. The company's past clients include HP and Trax, but its partner group doesn't appear to have included much in the way of driving-specific applications.

CrunchBase

Ford Motor Company

FOUNDED
1903

OVERVIEW
Ford is an automotive company that develops, manufactures, distributes, and services vehicles, parts, and accessories worldwide. It operates through two sectors: automotive and financial services. The automotive sector offers vehicles primarily under the Ford and Lincoln brand names. This sector markets cars, trucks, parts, and accessories through retail dealers in North America and distributors ...

LOCATION
Dearborn, MI

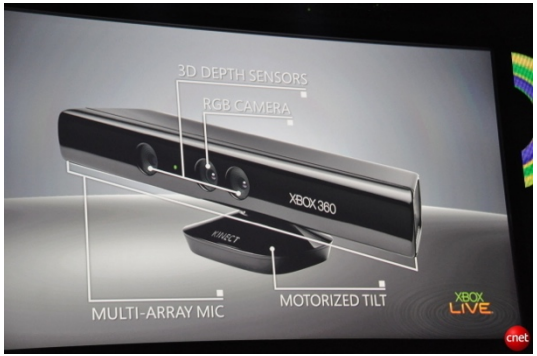
CATEGORIES
Automotive

WEBSITE
http://www.ford.com/

Full profile for Ford Motor Company

Interactive Games: Kinect – (Maybe)

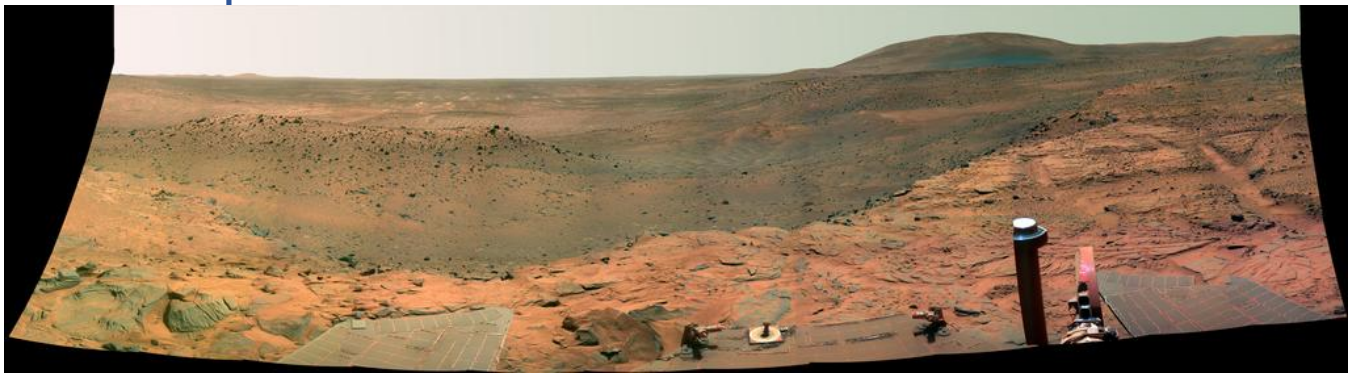
- Object Recognition: http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o
- Mario: http://www.youtube.com/watch?v=8CTJL5lUjHg
- 3D: http://www.youtube.com/watch?v=7QrnwoO1-8A
- Robot: http://www.youtube.com/watch?v=w8BmgtMKFbY



Industrial robots

Vision-guided robots position nut runners on wheels

Vision in space



NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

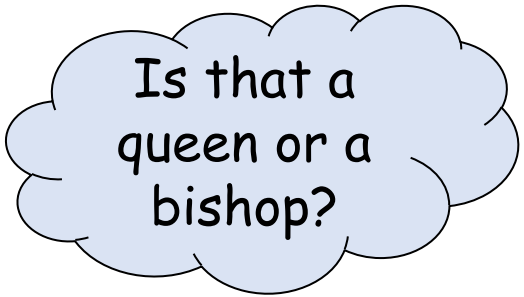
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

Augmented Reality and Virtual Reality

Magic Leap, Oculus, Hololens, etc.

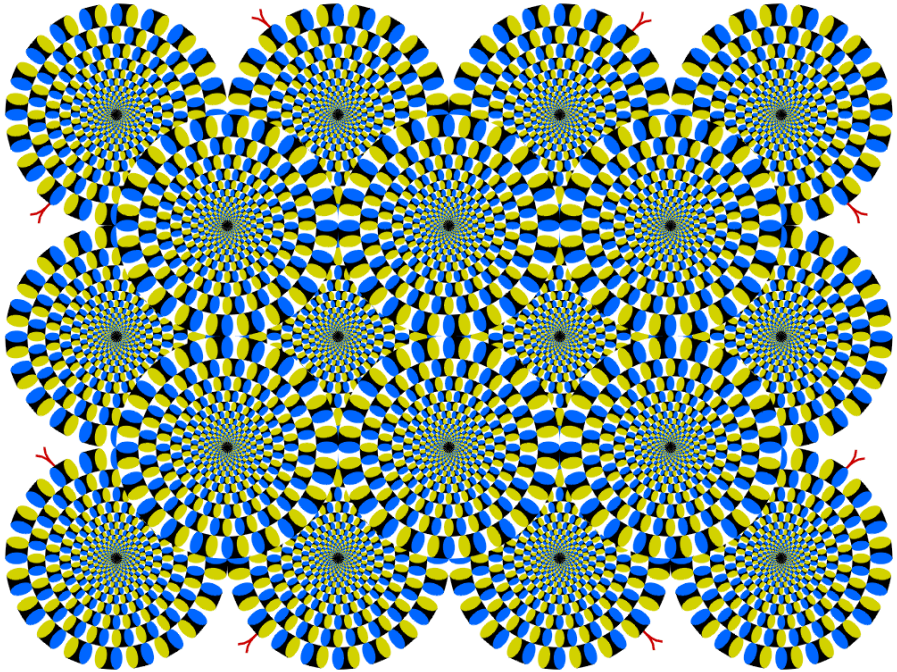
Vision is really hard

- Vision is an amazing feat of natural intelligence
- Visual cortex occupies about 50% of Macaque brain
- One third of human brain devoted to vision (more than anything else)

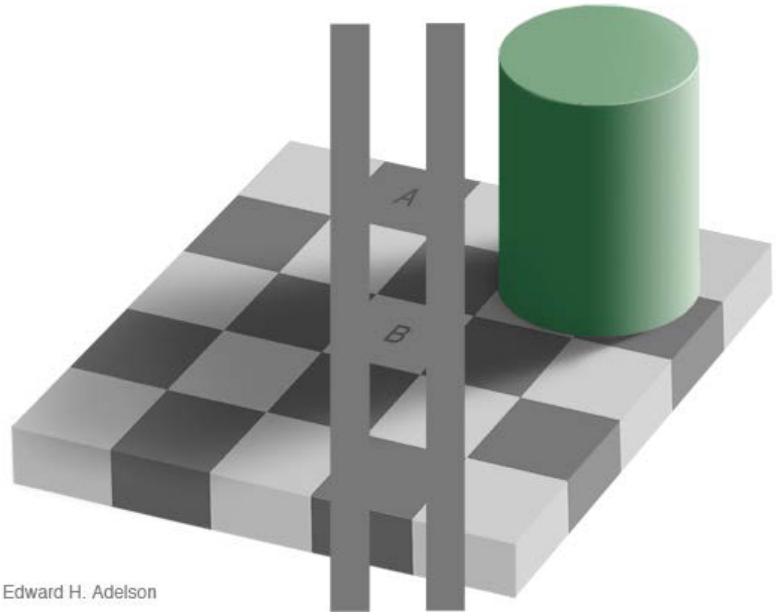


Slide by James Hays.

Is seeing trivial?

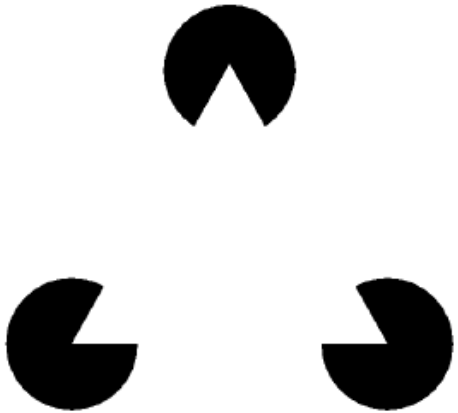


Is seeing trivial?



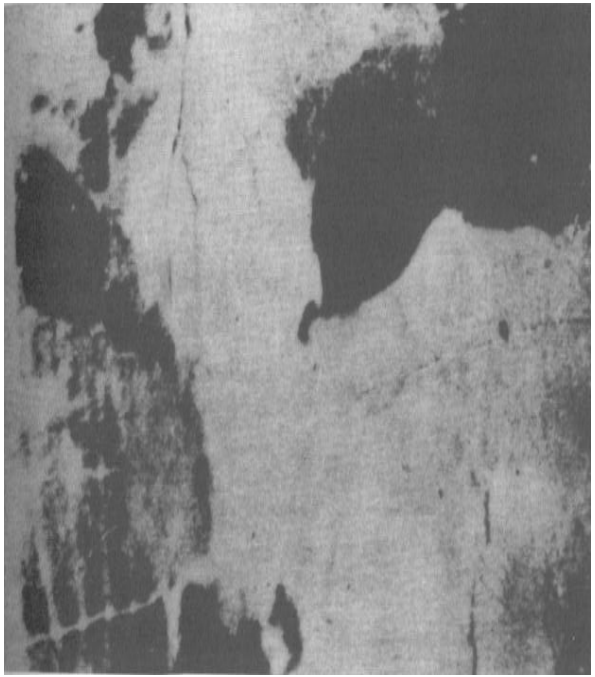
http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html

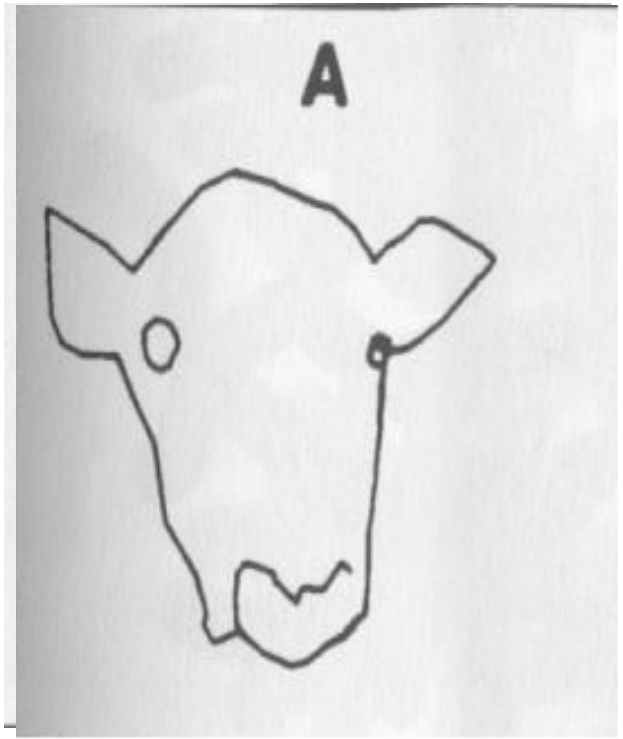
Is seeing trivial?



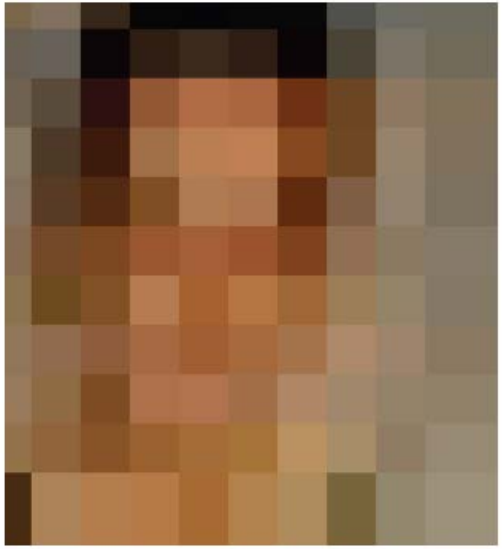
Is seeing trivial?

Is seeing trivial?

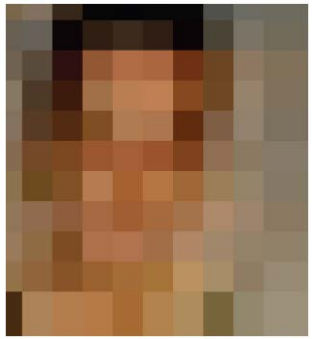




Face or non-face?

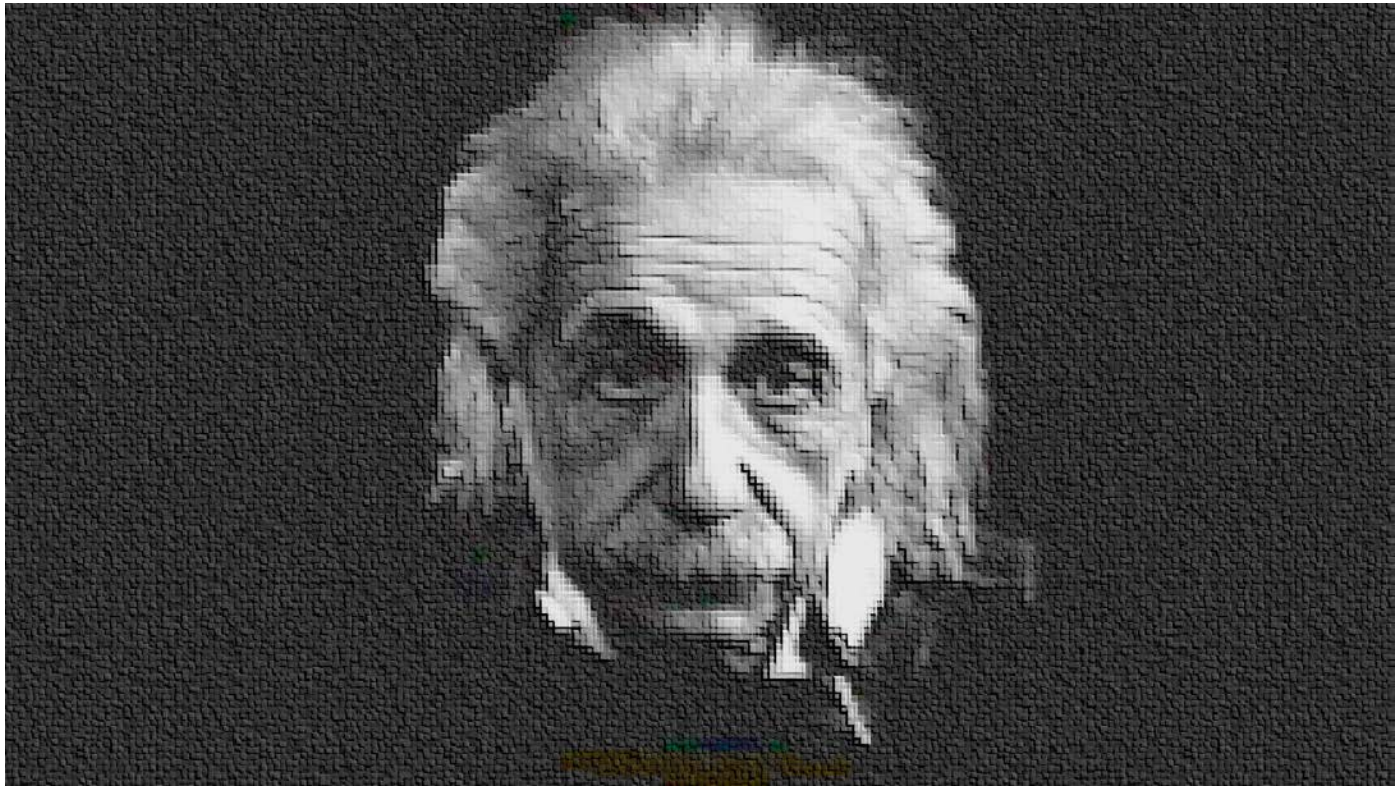


Face or non-face?

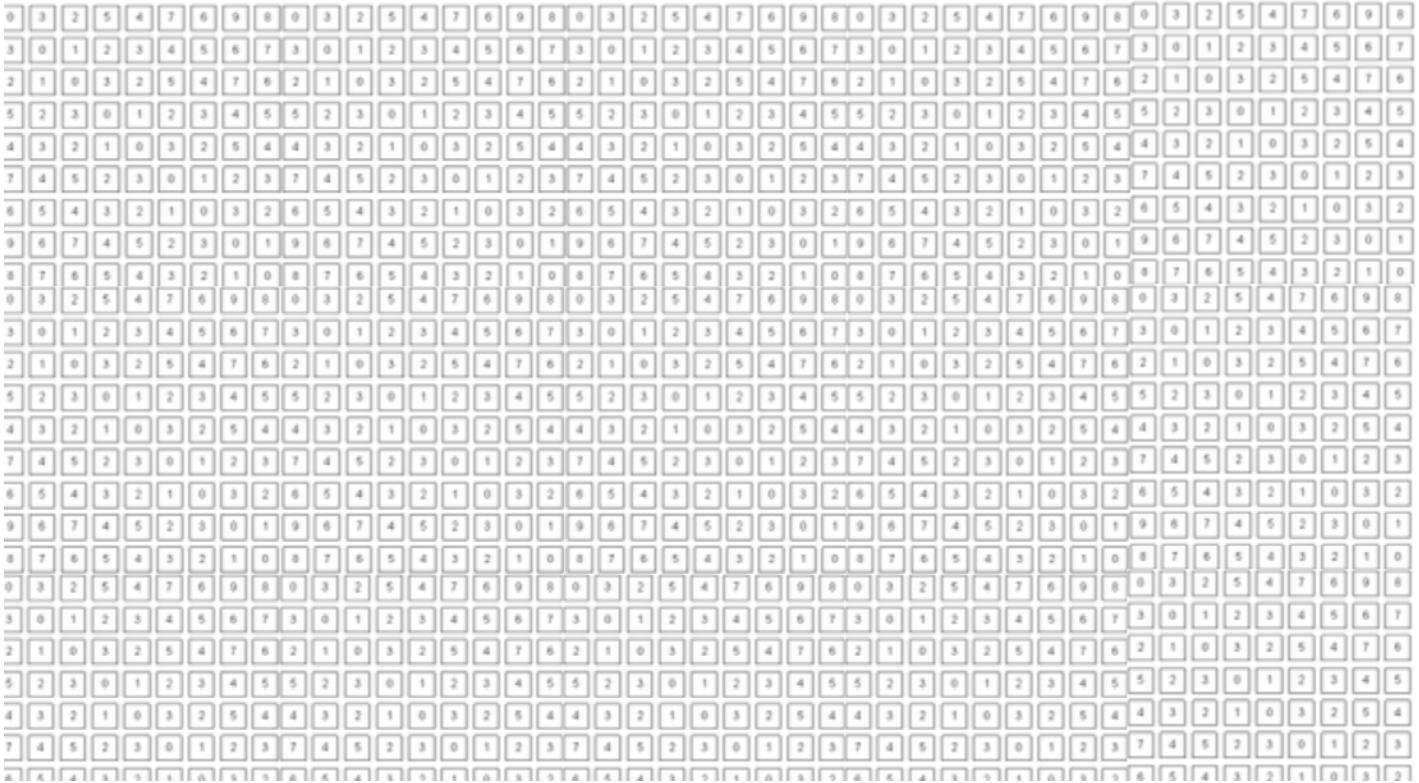


Why is vision so hard?

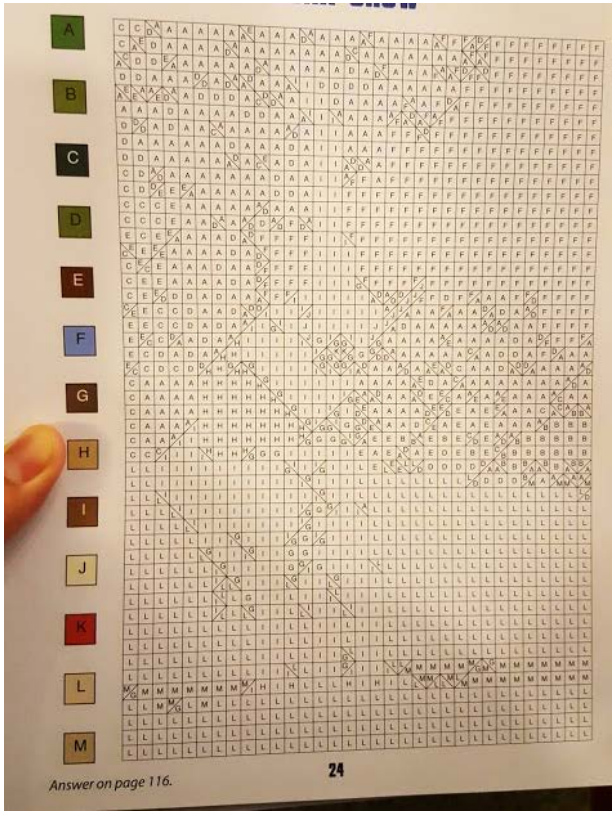
This is an image to us:

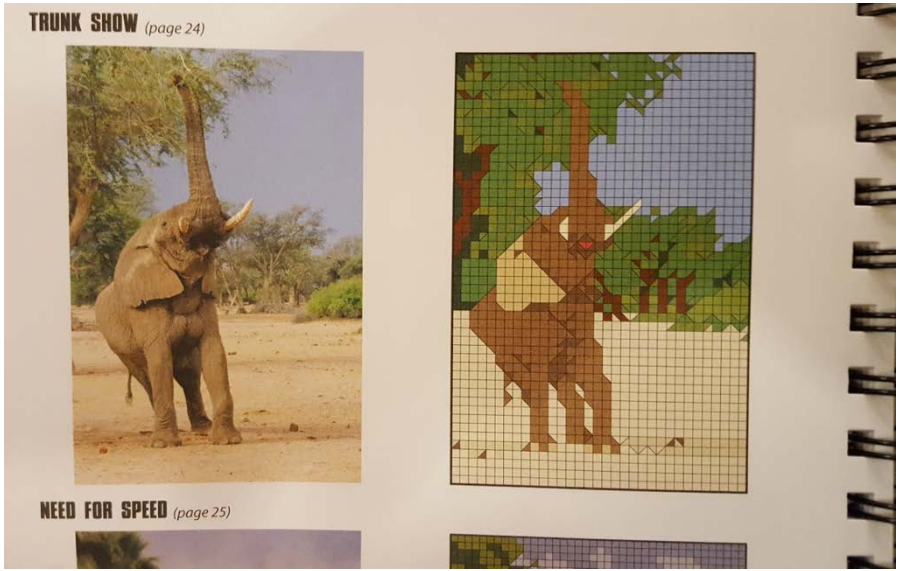


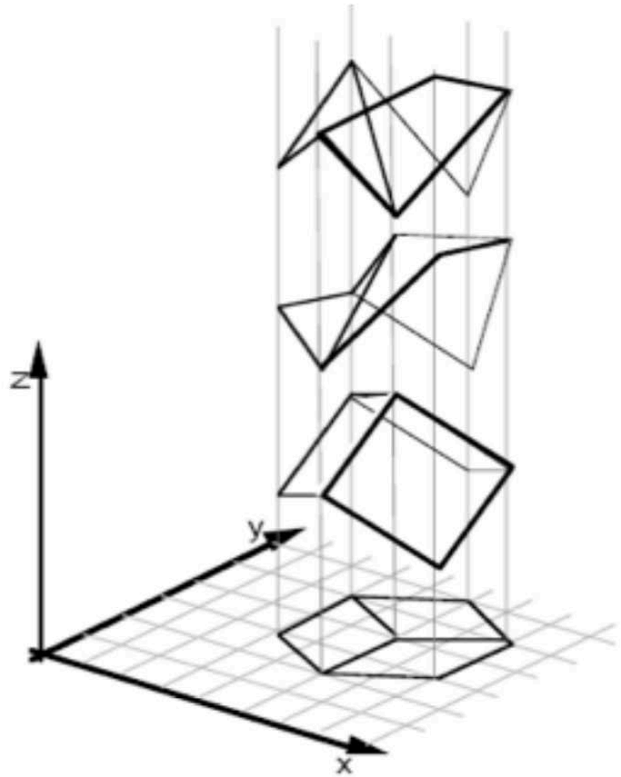
This is an image to a computer:



Vision is Hard







[Sinha and Adelson 1993]

View Points

Michelangelo 1475-1564

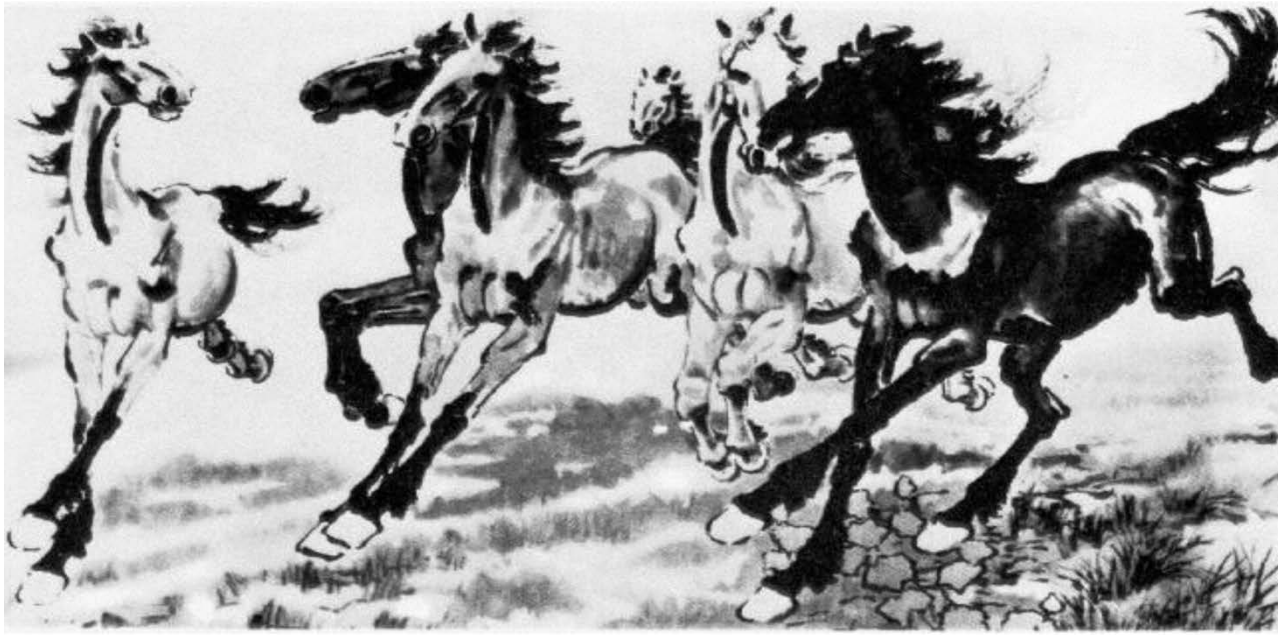
slide by Fei Fei, Fergus & Torralba

Illumination

slide credit: S. Ullman

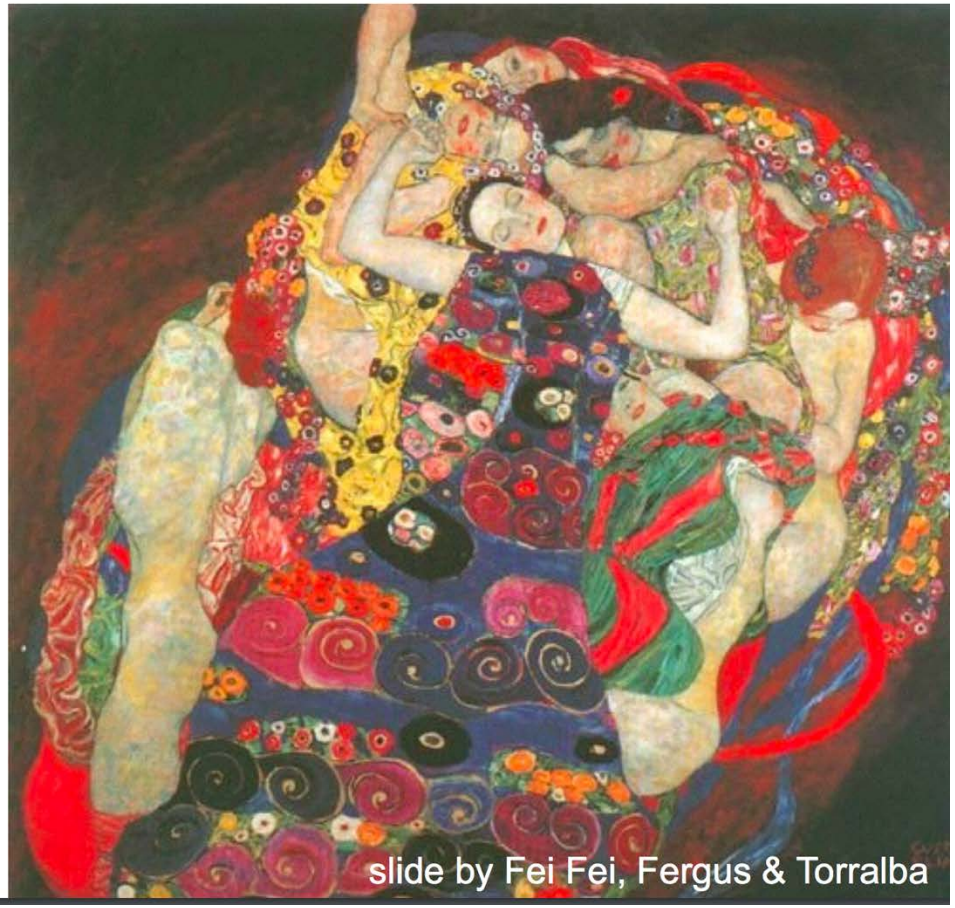
Occlusions

Deformation



Xu, Beihong 1943

Background clutter



Klimt, 1913

slide by Fei Fei, Fergus & Torralba

Intra-class variation

slide by Fei-Fei, Fergus & Torralba

What is the state of the art today?

- Given enough training data Computer Vision systems are surprisingly robust to the previously outlined challenges e.g. illumination changes, intra-class variation.

- Still not at the same level as humans, despite the hype.

- Still many open challenges, such as few-shot learning, transfer learning, and unsupervised learning.

Deep Learning and Vision

- Deep Learning has been a great disruption into the field of Computer Vision. Has made a lot of new things work!
- Many deep learning methods being applied to vision these days.
- This is not a deep learning course. We will study the important pre-deep learning methods, and then some deep learning.

Cameras

Cameras

Polaroid
$100

EOS Rebel T6i
$900

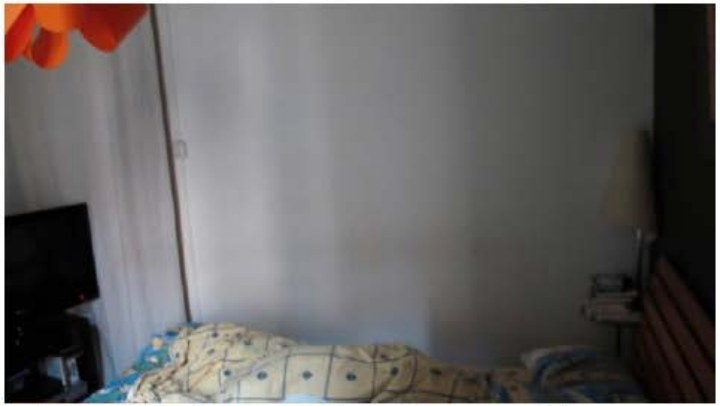
Canon EOS C300
$40,000

What do you need to make a camera from scratch?

Accidental Cameras

Accidental Pinhole and Pinspeck Cameras
Revealing the scene outside the picture.
Antonio Torralba, William T. Freeman

Accidental Cameras



a) Input (occluder present)

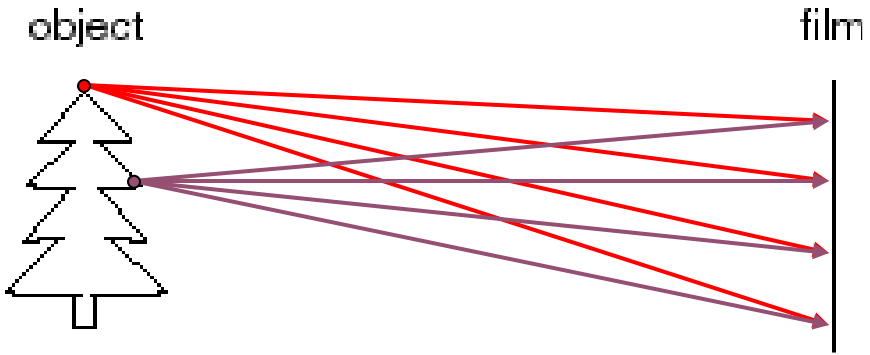
b) Reference (occluder absent)

c) Difference image (b-a)

d) Crop upside down

e) True view

Image formation

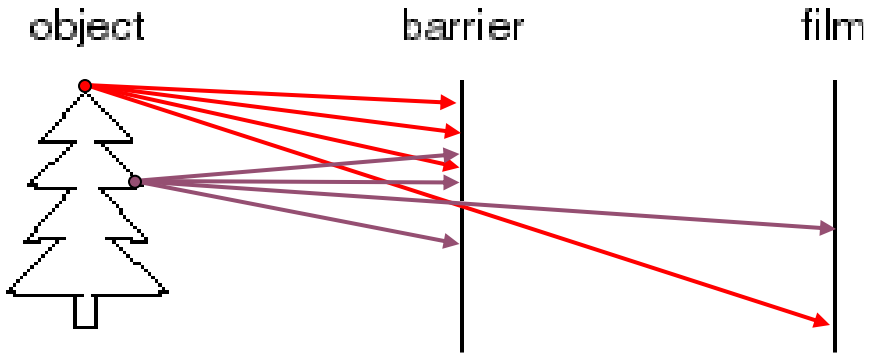


Let’s design a camera

- Idea 1: put a piece of film in front of an object
- Do we get a reasonable image?

Slide source: Seitz

Pinhole camera

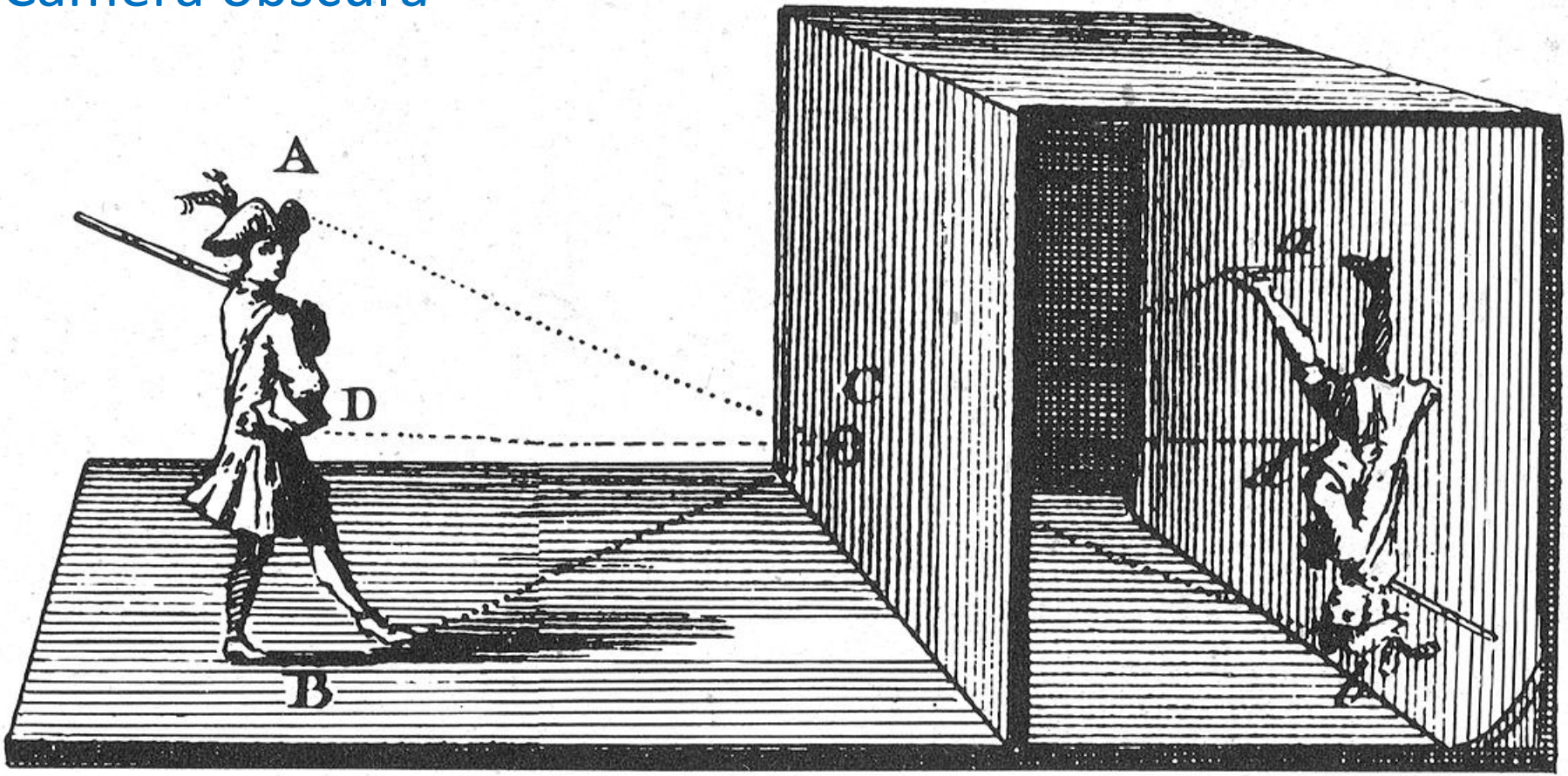


Idea 2: add a barrier to block off most of the rays

- This reduces blurring
- The opening known as the **aperture**

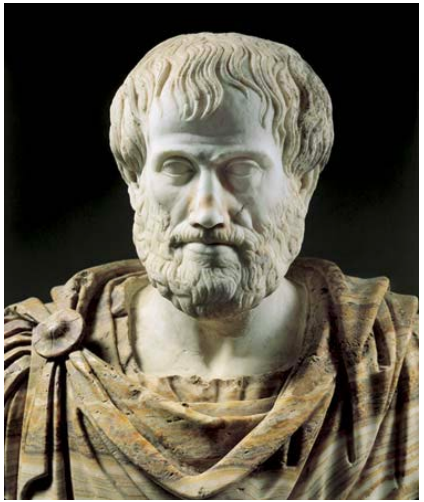
Slide source: Seitz

Camera obscura



Known by the Greeks and the Chinese 470BC-322BC

Recorded in writings by Chinese Philosopher Mozi (墨子)
470-390BCE



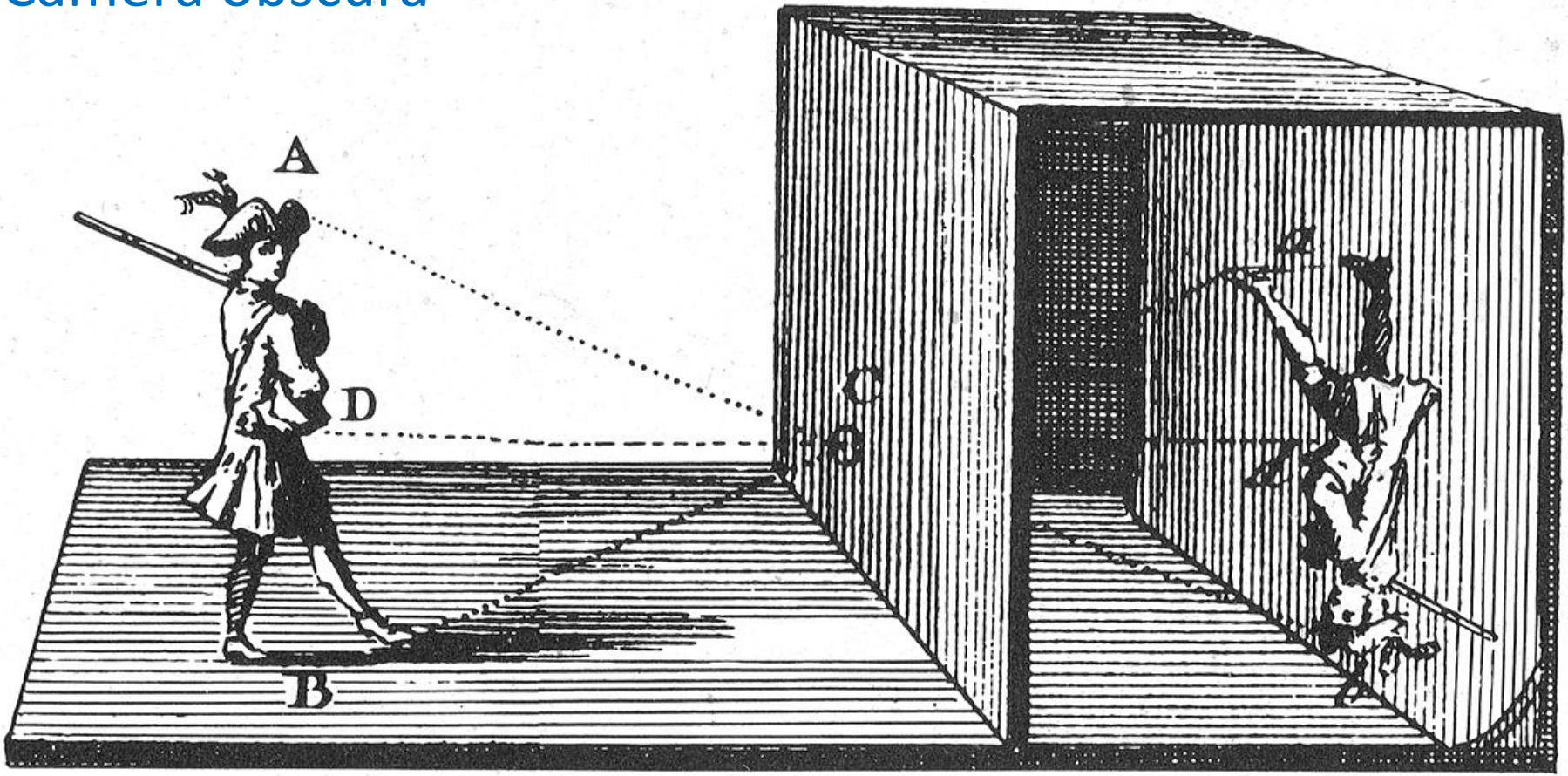
Recorded in writings by Aristotle or one of his disciples
384-322BCE

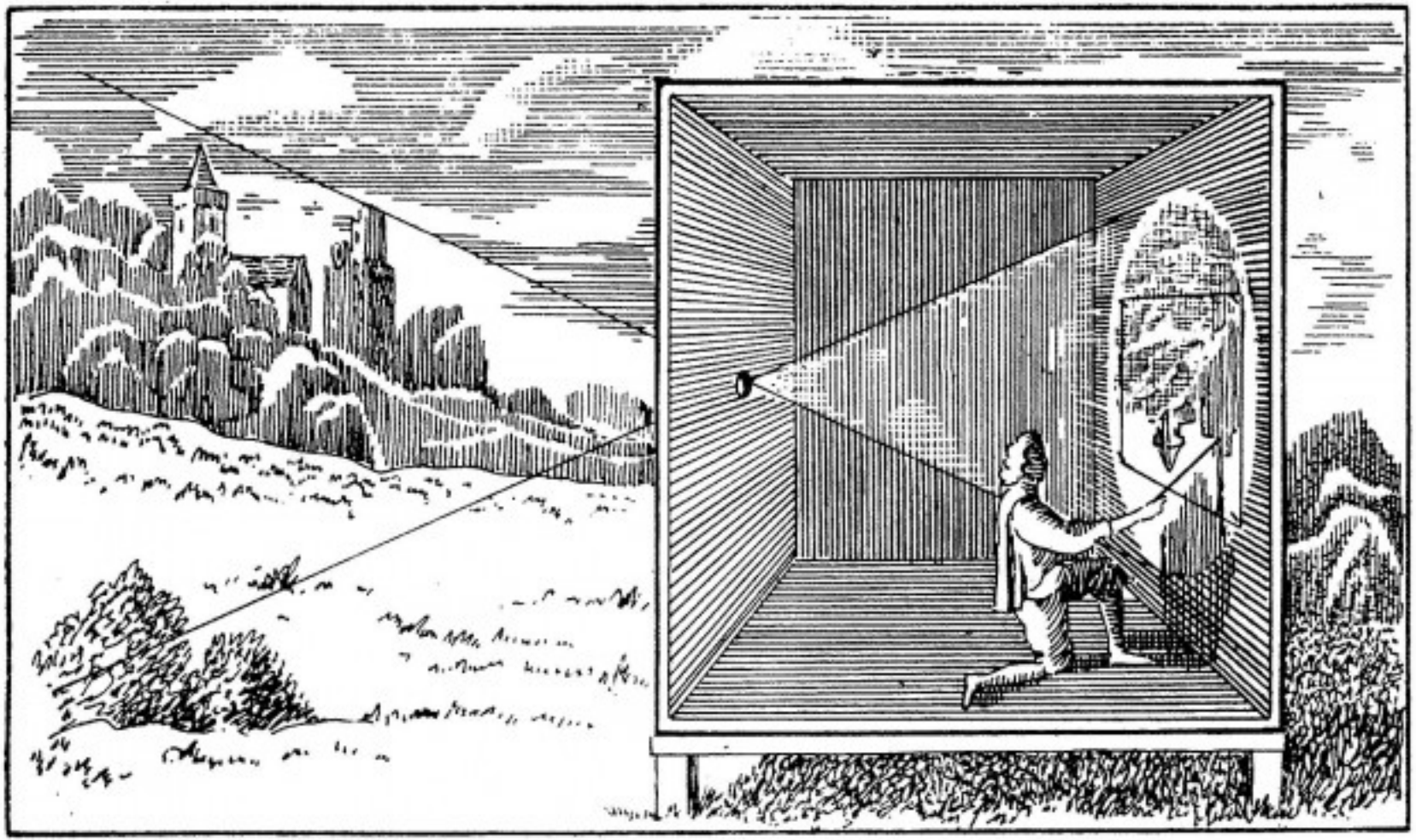
Camera obscura: the pre-camera

Freestanding camera obscura at UNC Chapel Hill

Photo by Seth Ilys

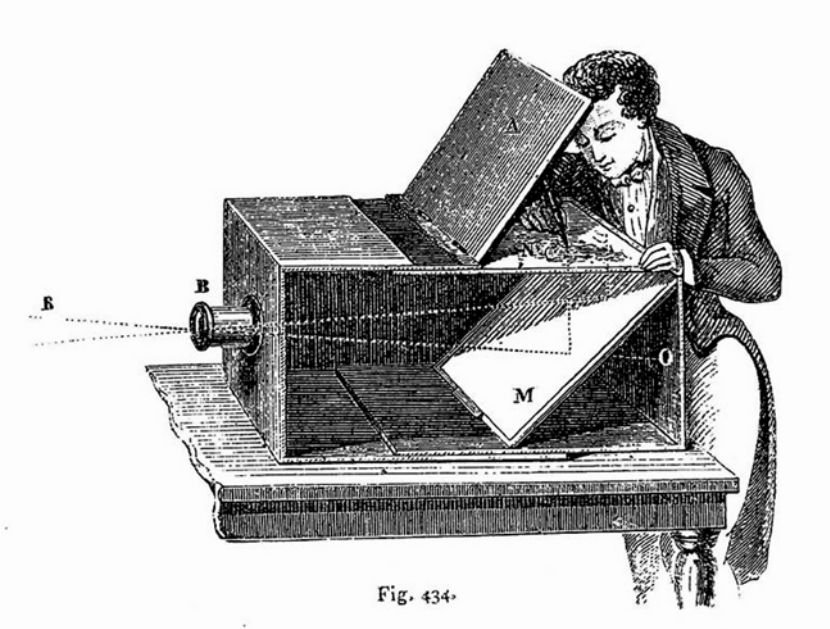
Camera obscura





https://pixsylated.com/blog/stepping-inside-havana-camera-obscura/

Camera Obscura used for Tracing



Lens Based Camera Obscura, 1568

First Photograph

Oldest surviving photograph

- Took 8 hours on pewter plate

Joseph Niepce, 1826

Photograph of the first photograph

Stored at UT Austin

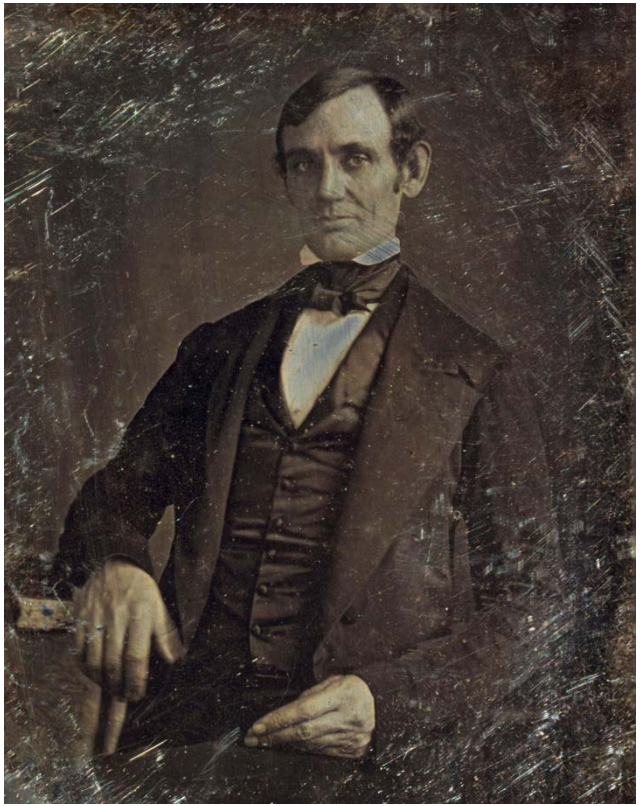
Bitumen of Judea: Naturally occurring asphalt that is photo-sensitive
https://en.wikipedia.org/wiki/Bitumen_of_Judea

From Joseph Niepce to Louis Daguerre

- Act 1: Joseph Niepce tells his nice idea to Louis Daguerre
- Act 2: Good friend Louis Daguerre improves idea and names it Daguerrotypes
- Act 3: Louis Daguerre makes history (and money).

1846 Daguerrotype of a young Abraham Lincoln

Daguerreotype camera built by La Maison Susse Frères in 1839, with a lens by Charles Chevalier



Hercules Florence's *Photographie*

- Brazilian painter and inventor
- Before Daguerre but after Niepce
- Included the idea of negatives

George Eastman 1885 (Rochester, NY)

- Founder of pioneering Eastman Kodak Company (Kodak)
- Popularization of film photography (nitrate film)

Should look familiar if you were born in the 80's or earlier or if you are a true modern-day hipster!

- Photo negatives

So, who invented cameras?

Maybe the wrong question to ask?

Questions?