Lecture 3: CS677

September 1, 2020

Admin

- Office Hours
  - Professor Tu,Th 4:30-6:00 PM
  - Xuefeng Hu, M 2:00-3:30 PM
  - Jiaxing Cheng Tu,Th 10:30AM- 12:00pm
  - Rijul Dhir W, F 2:00-3:30pm
- Prefer to make 15 min appointments at this link
  - https://docs.google.com/spreadsheets/d/1vqgLOSXbrg8xstRKMA5HY3VgWh5zeaF15V6aei1Mz48/edit?usp=sharing
  - Appointments outside office hours also available, especially for students outside LA

Miscellaneous

- Slack Channel

  Fall20-csci-677-30228

  DEN students will be invited to join the same
- Exam Dates
  - Exam1: October 13
  - Exam2: November 24
  - Term paper: December 3 (waiver has been requested)
- Waiting list
  - 20 students were added last week
  - Few more may be accommodated
    - Need to email to me confirming interest
- Interaction
  - Visual feedback is weak, rely on speech and chat

Review

- Previous class
  - Some problems of vision
  - Model-based and Data-driven approaches
  - Syllabys
- Today's objective
  - Some example state-of-art apps
  - Human visual system (very briefly)
  - Image formation

Why is Vision Hard?

- Scene complexity and scene segmentation
- Issues of representation
  - What is a chair?
- Intra-class object variations
  - Inherent to the class
  - Due to imaging conditions such as viewpoint, illumination and occlusion
- Ambiguities due to projection from 3-D to 2-D
- Additional complexities in dynamic scene analysis and activity recognition
- …

Topics to be studied in this class (*updated*)

- **Introduction (1 week)**
  Background, requirements and issues, human vision.
- **Image formation: geometry and photometry (1.5 weeks)**
  Geometry, brightness, quantization, camera calibration, photometry
- **Image segmentation (1.5 weeks)**
  Region segmentation, Edge and line finding
- **Multi-view Geometry (1.5 weeks)**
  Shape from stereo and motion, feature matching, Active ranging
- **Object Recognition: Traditional Methods (1.5 weeks)**
  HoG/SIFT features, Bayes classifiers, Linear Regression, SVM classifiers
- **Neural Network Basics (1 week)**
  Neural nets, CNNs, Backprop, SGD, Batch Normalization
- **Object Recognition: (2.5 weeks)**
  Image classification, object detection, semantic segmentation, Human pose estimation
- **Adversarial Attacks and Defense (.5 week)**
- **Motion Analysis and activity Recognition (1 week)**
  Optical flow, motion features, classification network
- **Selected Topics (1 Week)**
  Face Identification, Vision and language …

Current state of the art

- The following slides show some examples of what current vision systems can do
  - Many taken from the class page of Prof. Seitz/Szeliski at U of Washington

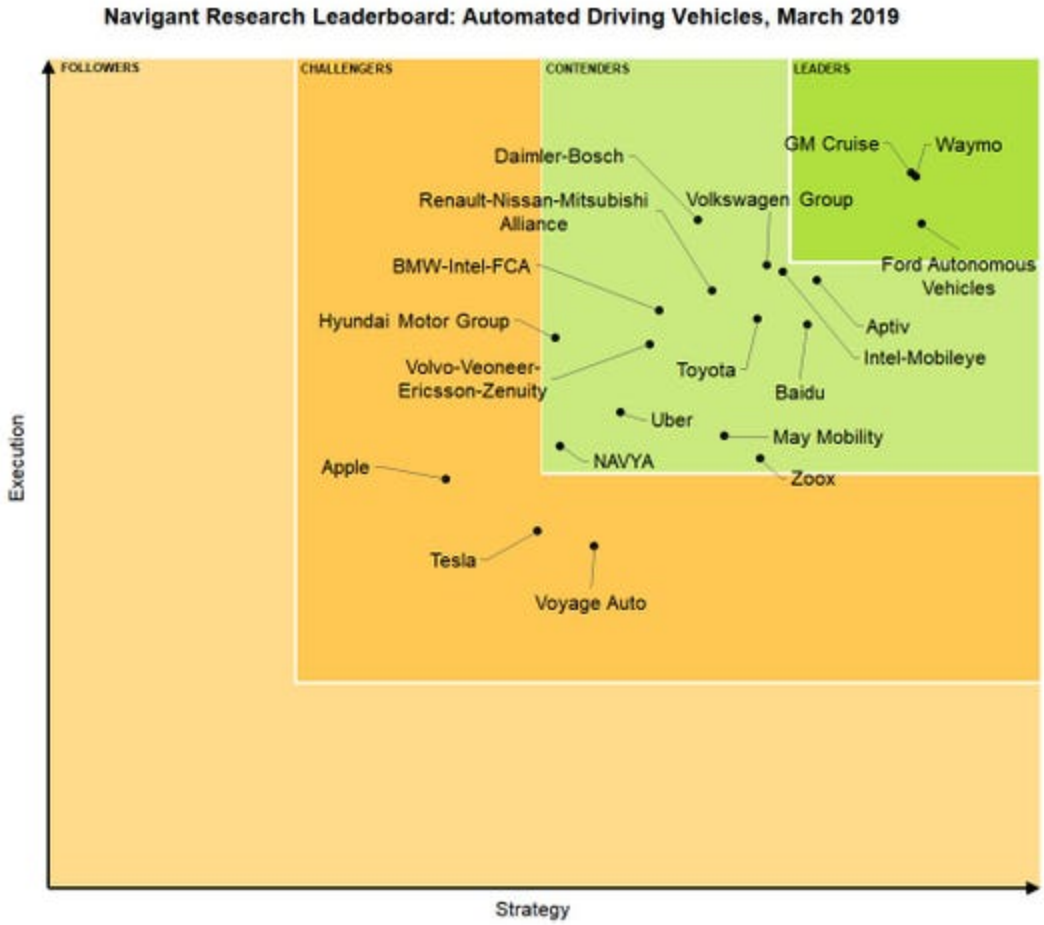
Driving Scene

From MobilEye

Self-Driving Cars

- A short video showing some visual needs and capabilities
  - https://www.youtube.com/watch?v=42rmGs0Rvtw
- A long talk on status of self-driving cars (watch on your own)
  - https://www.youtube.com/watch?v=GJ82mk99Agw
- A business analysis of participants in self-driving technology
  - http://www.businessinsider.com/the-companies-most-likely-to-get-driverless-cars-on-the-road-first-2017-4/#1-ford-18

Autonomous Driving Leaders



Top 10 Vendors

Waymo
Ford Autonomous Vehicles
Cruise
Baidu
Intel-Mobileye
Aptiv-Hyundai
Volkswagen Group
Yandex
Zoox
Daimler-Bosch

https://www.freep.com/story/money/cars/mark-phelan/2019/03/13/waymo-ford-chrysler-apple-tesla-autonomous-car/3142974002/
https://guidehouseinsights.com/reports/guidehouse-insights-leaderboard-automated-driving-vehicles?utm_content=buffereb595&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer

Manufacturing and Inspection

COGNEX

Products Industries Applications Support How to Buy Resources Company

> LEARN MORE

FEATURED PRODUCTS | FEATURED APPLICATIONS | EVENTS AND TRAINING | NEWS

Companies across the world use Cognex products to eliminate defects, verify assembly, automate production, track and identify parts.

KIA Motors

Achieves 99% read rates with Cognex barcode readers

Maclean Vehicle Systems

Eliminates Rejection of Good Parts in Inspecting Fasteners

Renault

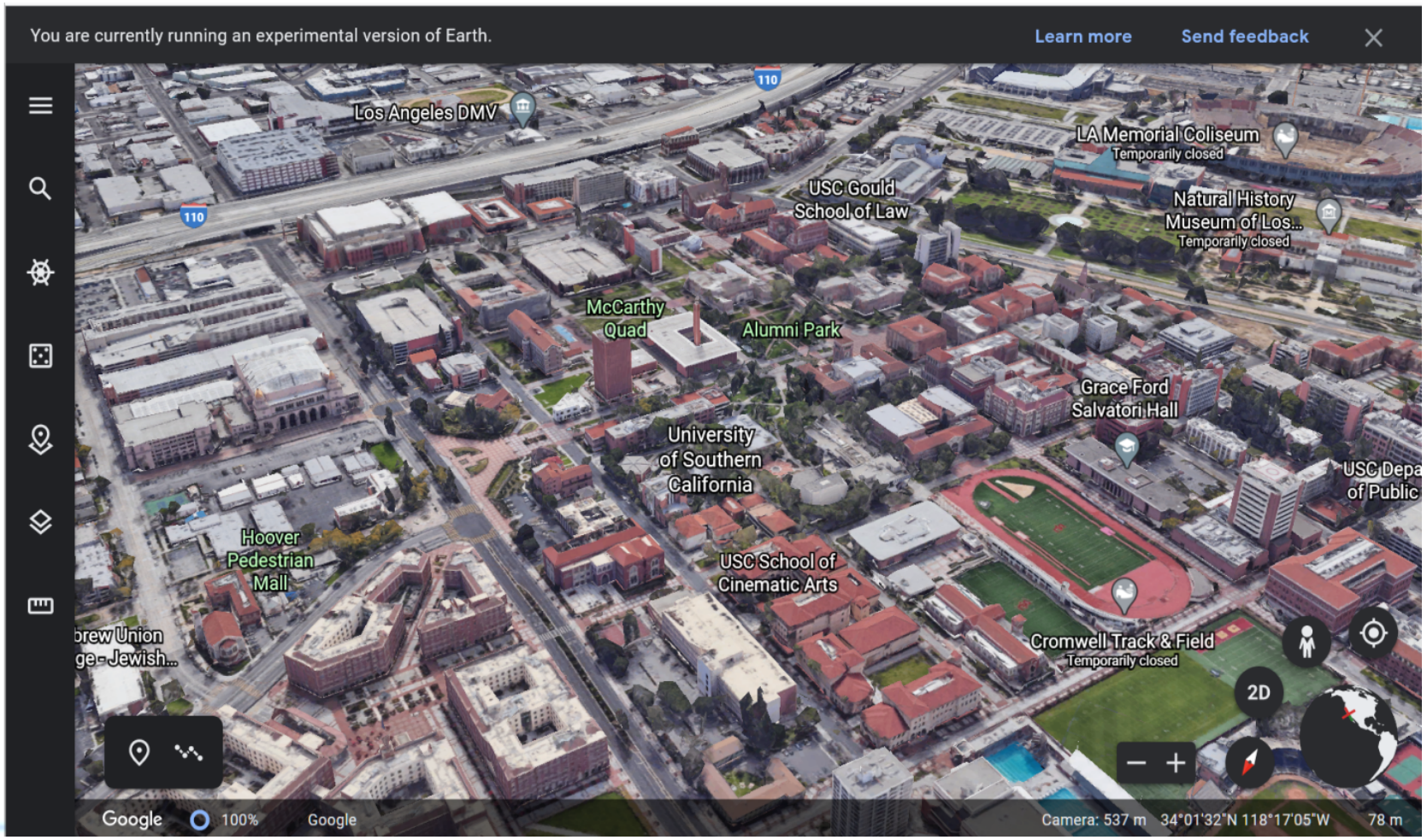
Assembly line upgrade boosts profitability

BMW

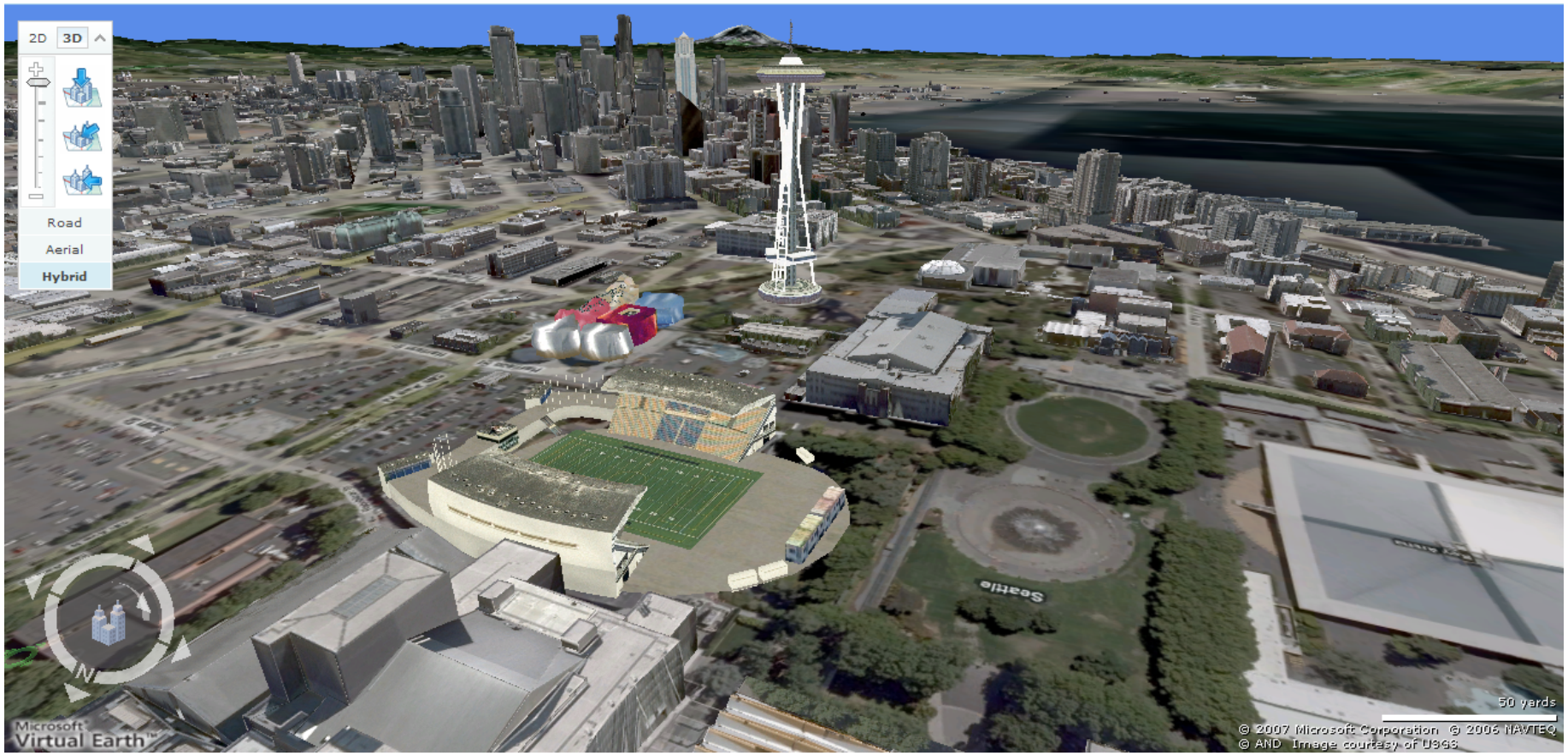
Quality control and inspection for head-up displays

iPhone PANO Images

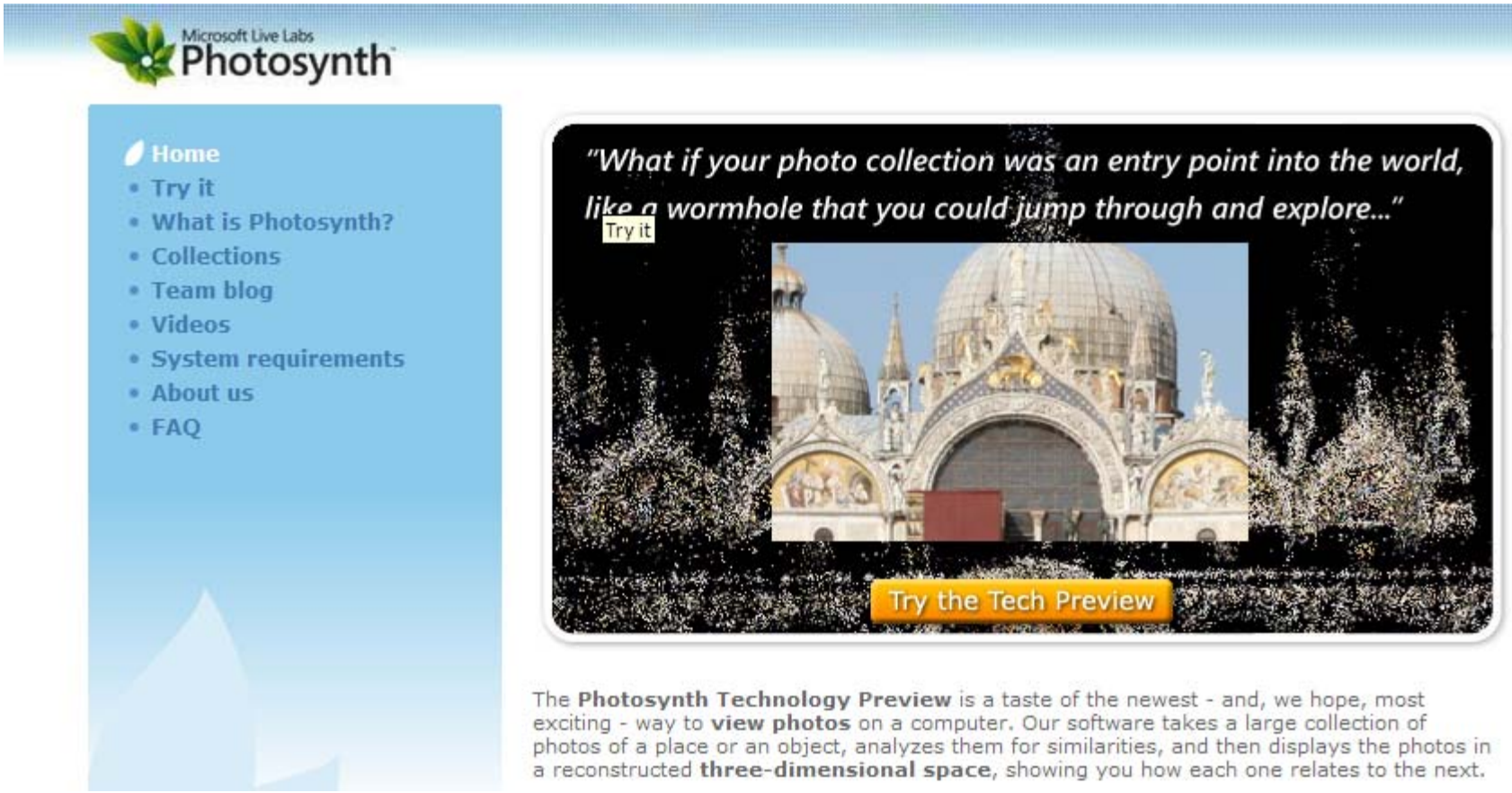
Google Earth



Virtual Earth (Microsoft)



Photosynth



http://labs.live.com/photosynth/

Based on Photo Tourism technology

Optical character recognition (OCR)

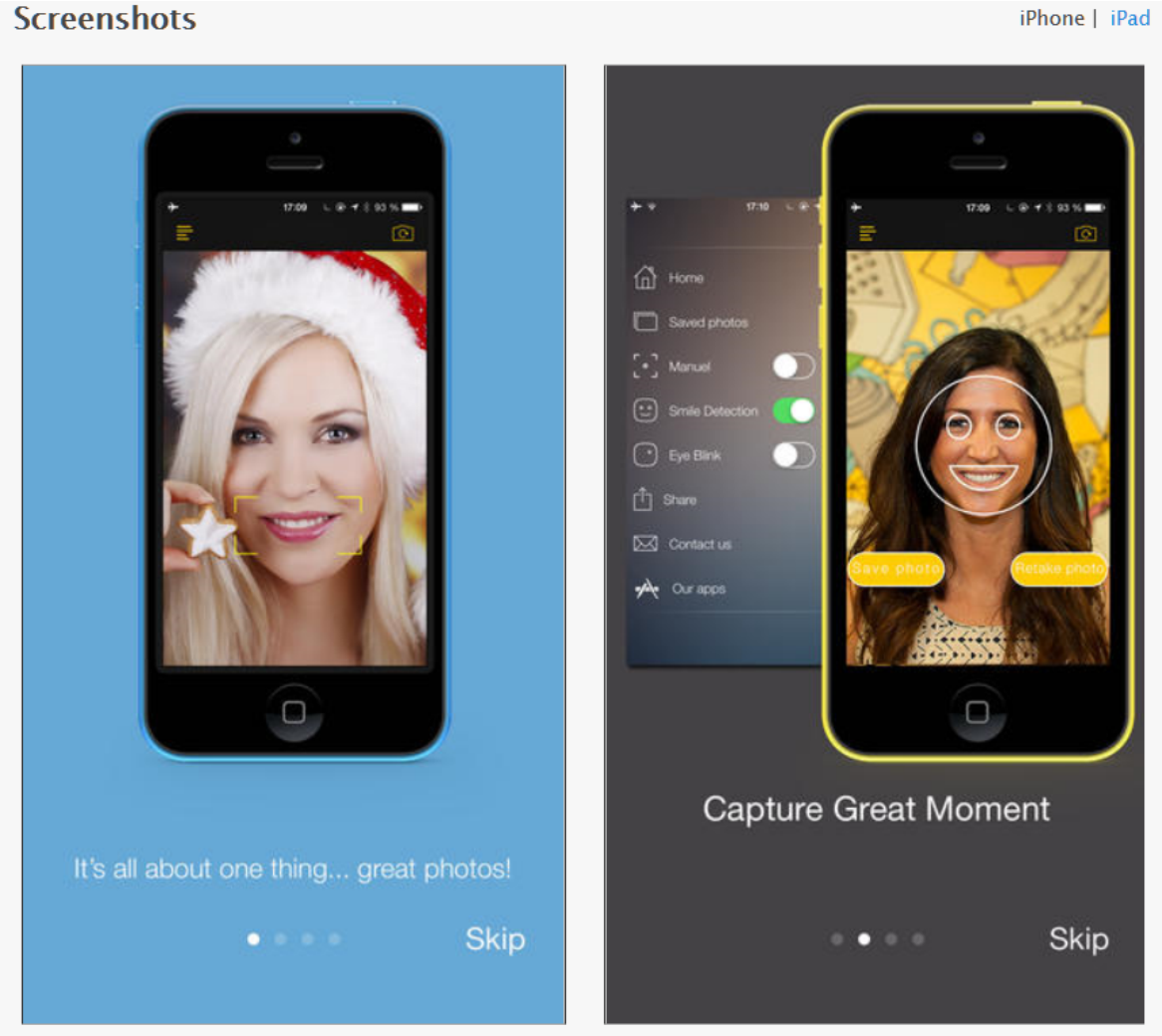
- Technology to convert scanned docs to text
  - If you have a scanner, it probably came with OCR software

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

Face detection

- Most new digital cameras detect faces

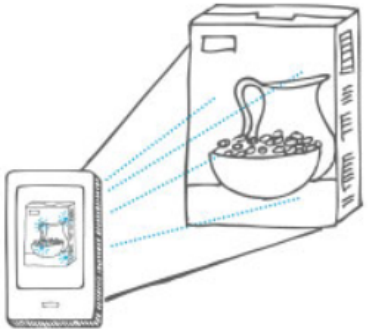
Smile Detector (from Quanticapps)

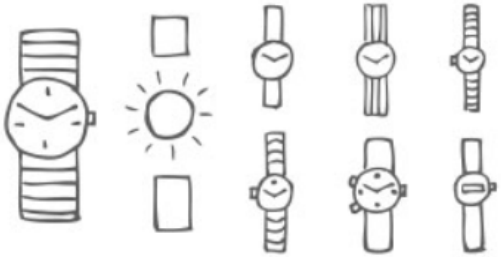


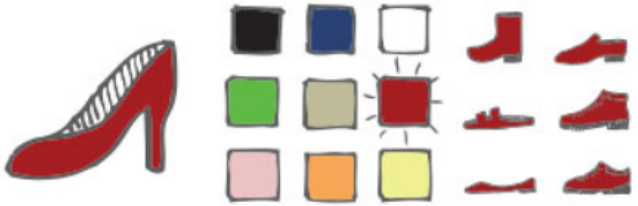
Cognitec Face Recognition

A9.com

Visual Search also develops computer vision solutions that support Amazon initiatives along the entire product delivery pipeline: from the time a new product is photographed and added to our catalog to the time an item is bought and shipped to the customer.







Amazon Go

- https://www.youtube.com/watch?v=NrmMk1Myrxc

Special Effects

- From movie “Avatar”; image from www.rockying.com

Sports

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

Vision-based interaction (and games)

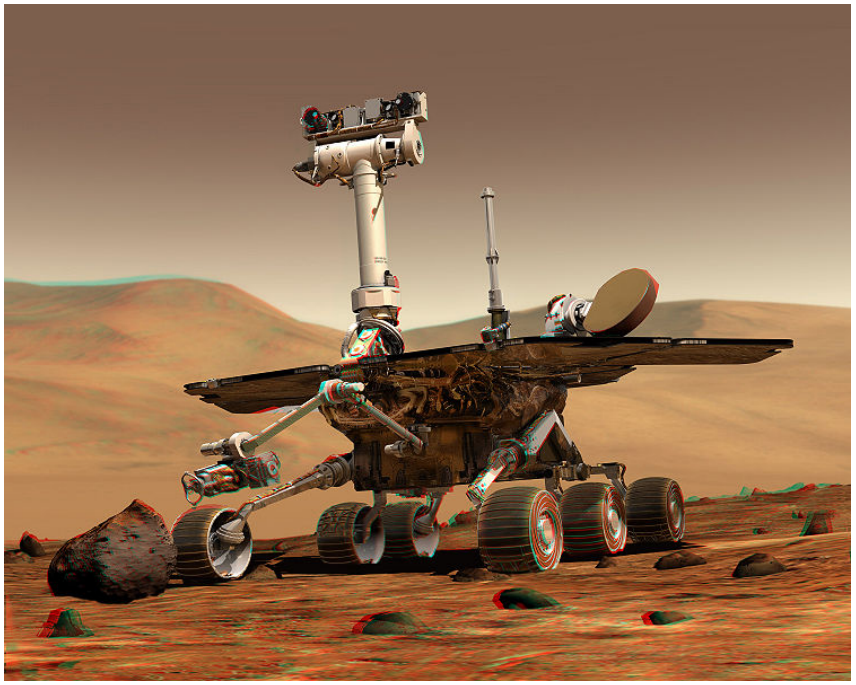
Nintendo Wii has camera-based IR tracking built in.

Microsoft Kinect
(no images here)

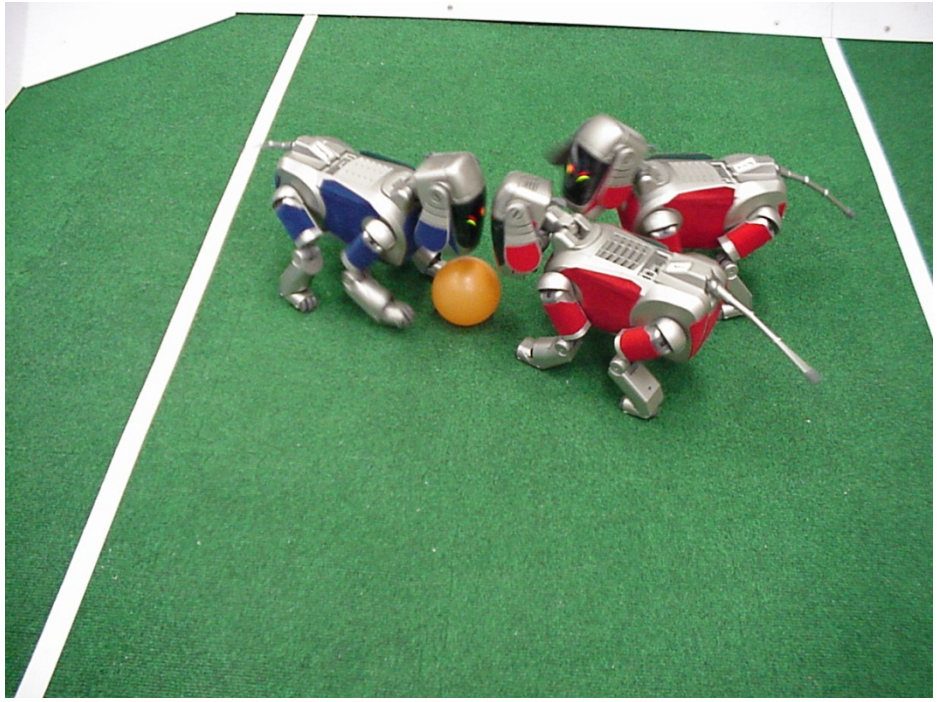
Digimask: put your face on a 3D avatar.

"Game turns moviegoers into Human Joysticks", CNET
Camera tracking a crowd, based on this work.

Robotics

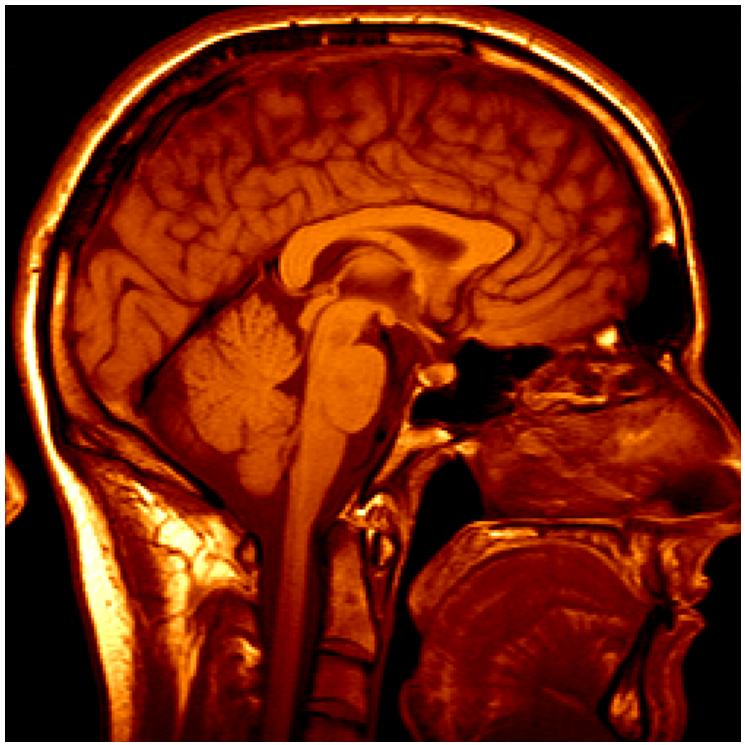


NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



http://www.robocup.org/

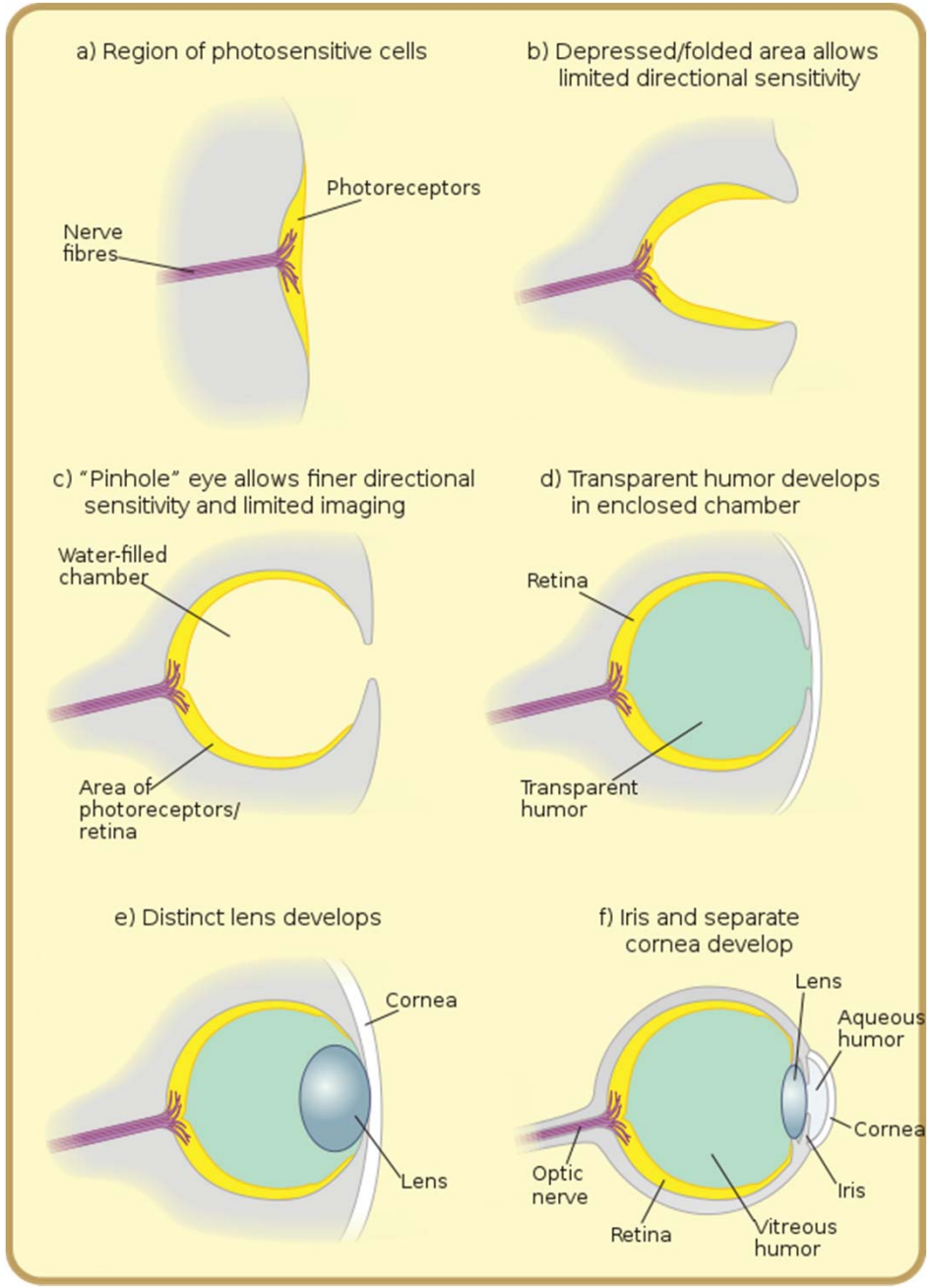
Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Evolution of Eyes

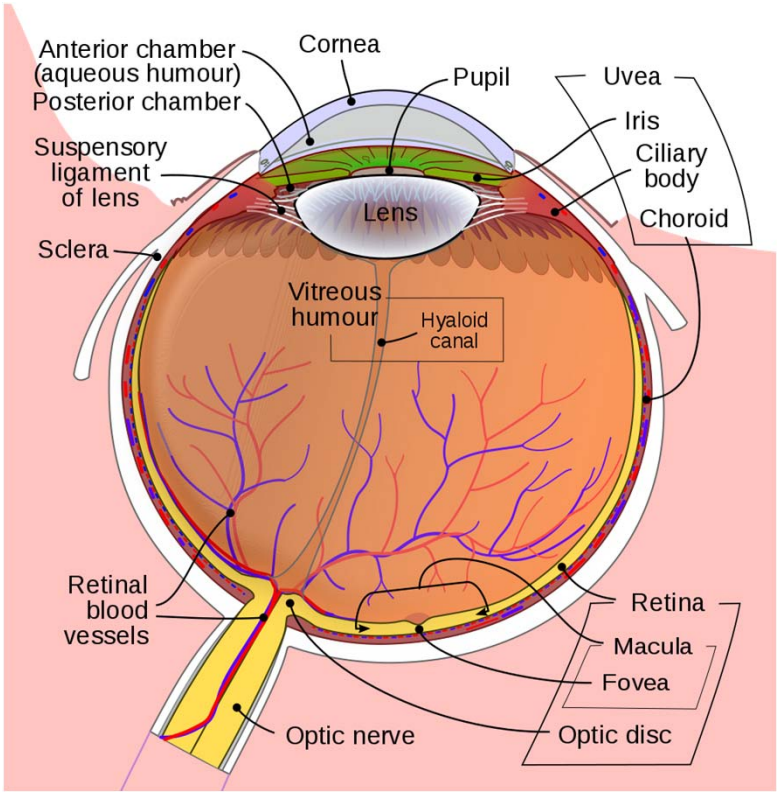


http://www.wikiwand.com/en/Evolution_of_the_eye

Human Eye

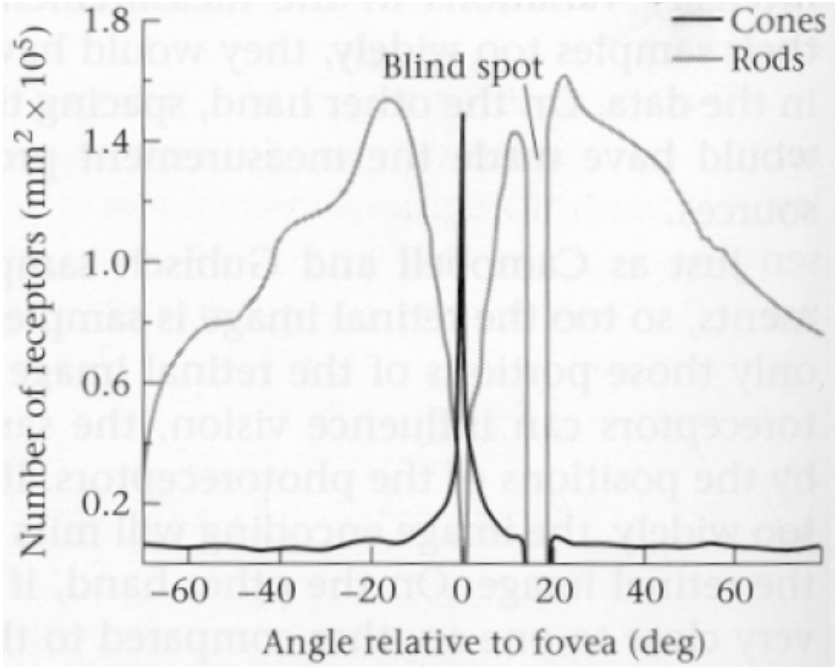
- Like a camera
  - Lens, pupil (iris), focus by *accommodation*
- Image formed on back of eye (retina)
- Optic nerve sends *data* to brain (cortex)
  - Blind spot (where optic nerve comes out)

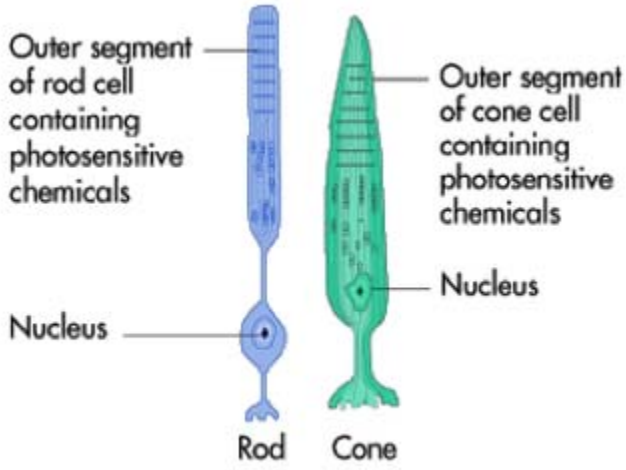
From Wikipedia



Retina

- Two types of photoreceptors
  - Rods: highly sensitive to light, not used for color vision, ~ 100M rods
  - Cones: 3 different types with different spectral sensitivities, less sensitive to light, ~ 5M cones
    - Explains why *color* is not seen at night
- Distribution is not uniform
  - High concentration of cones in fovea (0.5 minute visual angle)
  - Fixation (*foveation*) to get high resolution everywhere

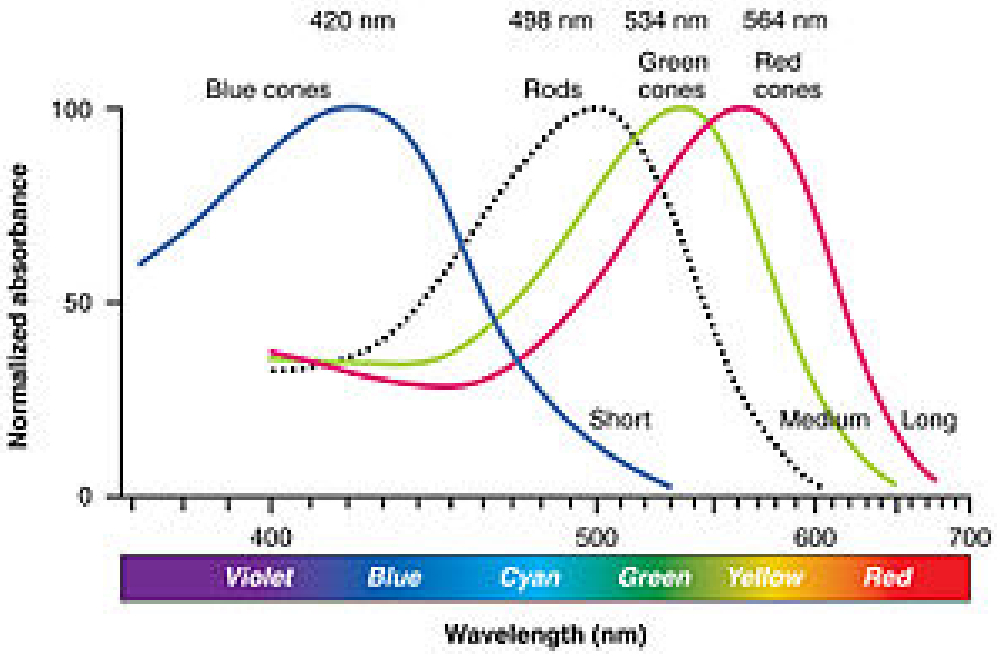




http://ionabio.weebly.com/

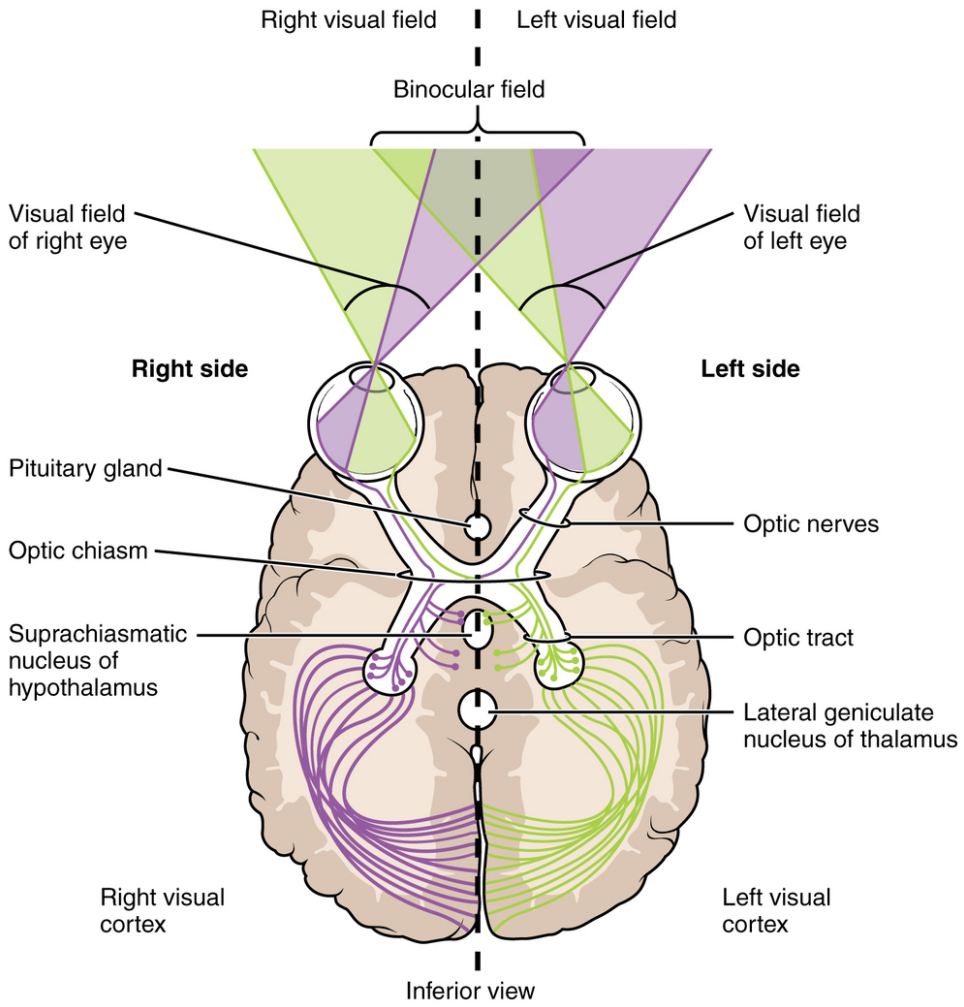
Color Sensor Response

- Eyes do not have built in color spectrometer
- Rather, we have 3 sensors with different responses to lights of different color
- Perceived color depends on relative responses of three sensors



https://en.wikipedia.org/wiki/Photoreceptor_cell

Cortex schematic



From lindsayoptometric.com

Optical nerve carries signals from retina to cortex

- ~100:1 ratio of nerve fibers to receptors: some processing performed at this level

Optical *chiasma*: optic nerve fibers split to two halves of the brain

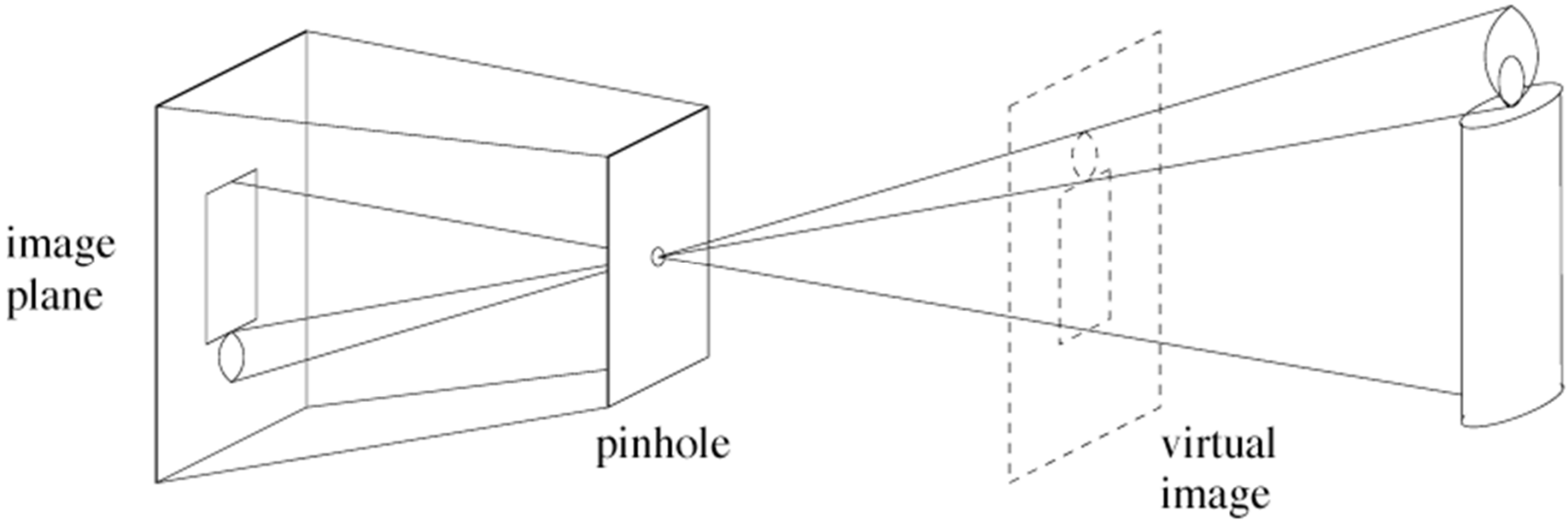
Many functional areas (V1, V2,….): knowledge about them is limited

Image Formation

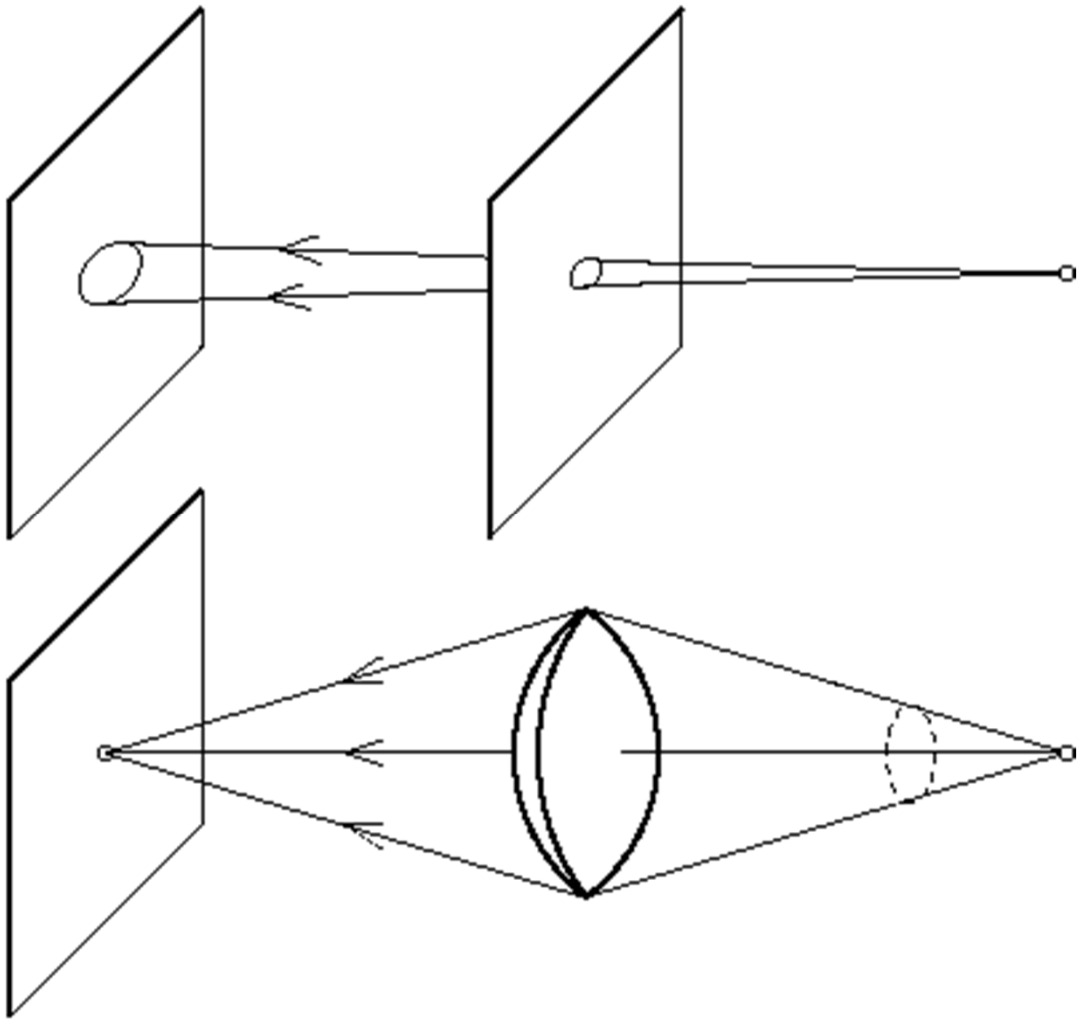
- Geometry
  - Where is the image of a point formed?
- Photometry/Colorimetry
  - How bright is the point?
  - What is its *color*?
- Ideal camera models
- Real lenses

Pinhole cameras

- Abstract camera model - box with a small hole in it
- Note inverted image
- Pinhole cameras work in practice, ignoring diffraction

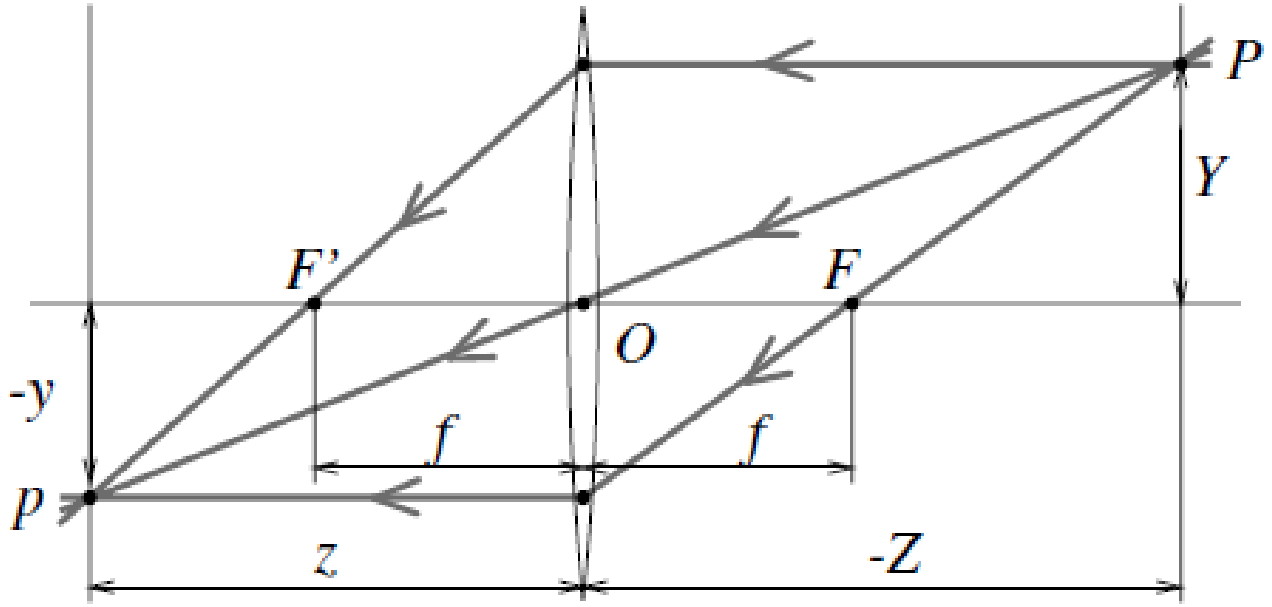


The reason for lenses



Sharper focus,
more light

The thin lens



$$\frac{1}{z} - \frac{1}{Z} = \frac{1}{f}$$

Note: Z is a negative number, z-axis points towards the camera

Thin Lens Properties

- Points at different depth focus at different positions of the image plane
  - With a fixed image plane, not all points will be in focus
  - "Depth of field (DOF)", *i.e.* distance over which focus is "acceptable" depends on the *aperture* size
  - Larger aperture captures more light but has lower DOF
  - Defocus property can be used to infer depth
    - Limited accuracy
- Field of view: depends on imaging surface size, not lens aperture size (see next slide)

Field of View (FoV)

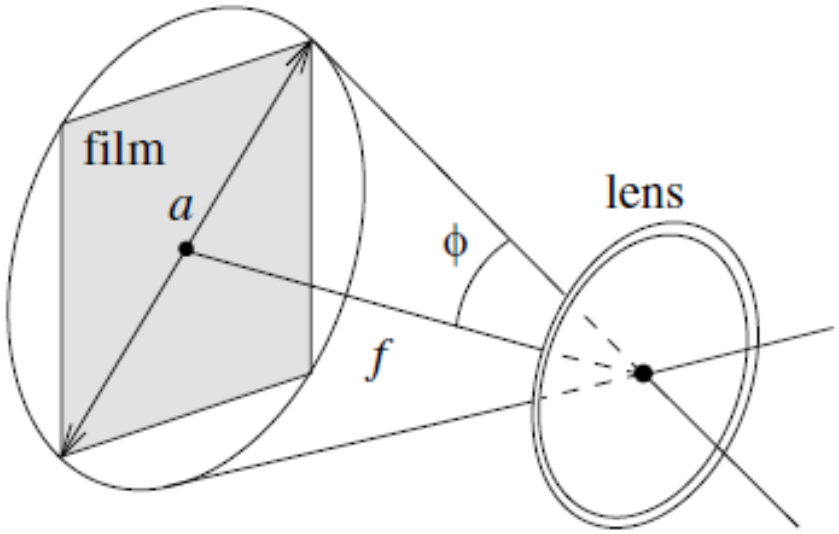


FIGURE 1.9: The field of view of a camera. It can be defined as 2ϕ, where $\phi \stackrel{\text{def}}{=} \arctan \frac{a}{2f}$, a is the diameter of the sensor (film, CCD, or CMOS chip), and f is the focal length of the camera.

Lens Distortions

- Real lenses suffer from various errors/distortions
- Chromatic aberration (not all wavelengths focus at the same point)
- Geometric distortions: complex lens systems used to reduce distortion
- Usually we will assume that complex lenses behave as ideal pinhole models but without the negative effects
  - No diffraction effects, sufficient light collection, all points in focus

Distortion Illustrations

(a)

(b)

(c)

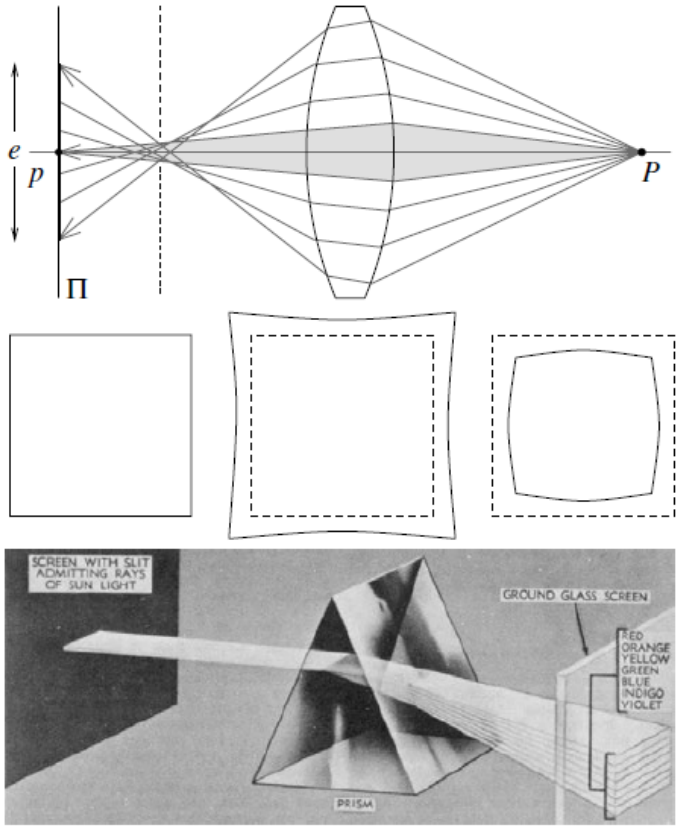
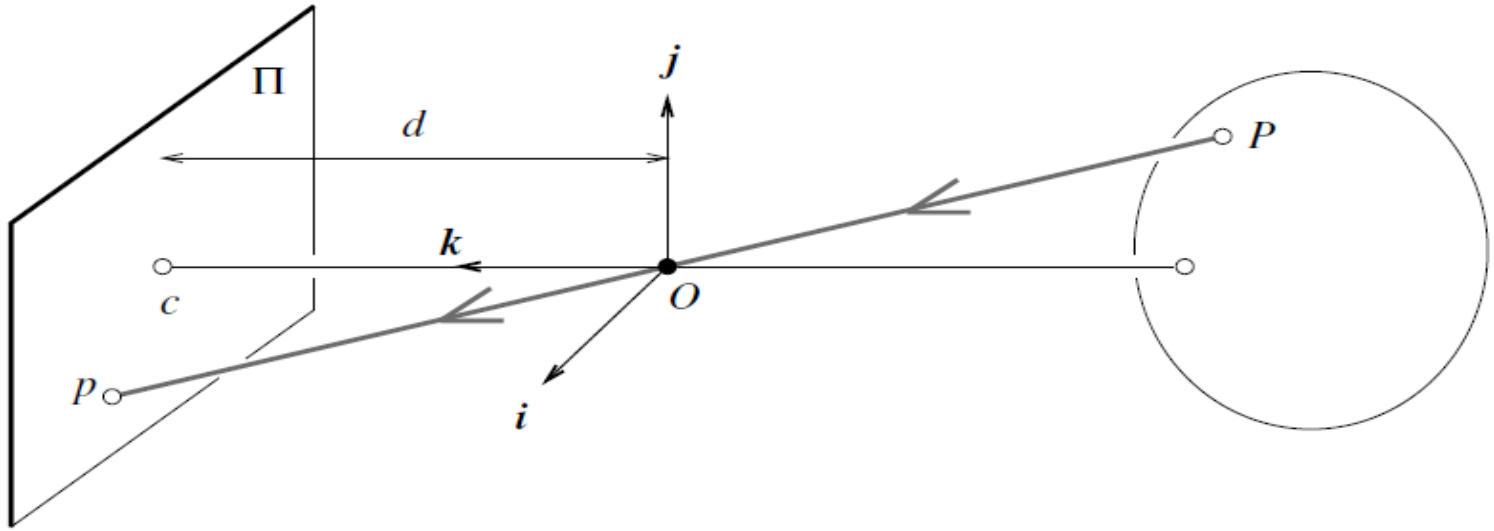


FIGURE 1.11: Aberrations. (a) Spherical aberration: The gray region is the paraxial zone where the rays issued from P intersect at its paraxial image p. If an image plane π were erected in p, the image of p in that plane would form a circle of confusion of diameter e. The focus plane yielding the circle of least confusion is indicated by a dashed line. (b) Distortion: From left to right, the nominal image of a fronto-parallel square, pincushion distortion, and barrel distortion. (c) Chromatic aberration: The index of refraction of a transparent medium depends on the wavelength (or color) of the incident light rays. Here, a prism decomposes white light into a palette of colors. *Figure from US NAVY MANUAL OF BASIC OPTICS AND OPTICAL INSTRUMENTS, prepared by the Bureau of Naval Personnel, reprinted by Dover Publications, Inc. (1969).*

The equation of projection

- Note: *k*-axis *towards* the camera (right handed coordinate system $\boldsymbol{k} = \boldsymbol{i} \times j$).



Let P = (X,Y,Z), p = (x, y, z)

- We know that z= d, find values of x and y
- Op = λ.OP for some λ, λ = d/Z

hence:

$$\begin{cases} x = d\dfrac{X}{Z}, \\ y = d\dfrac{Y}{Z}. \end{cases}$$

Next Class

- FP: Sections 1.1,1.2,1.3