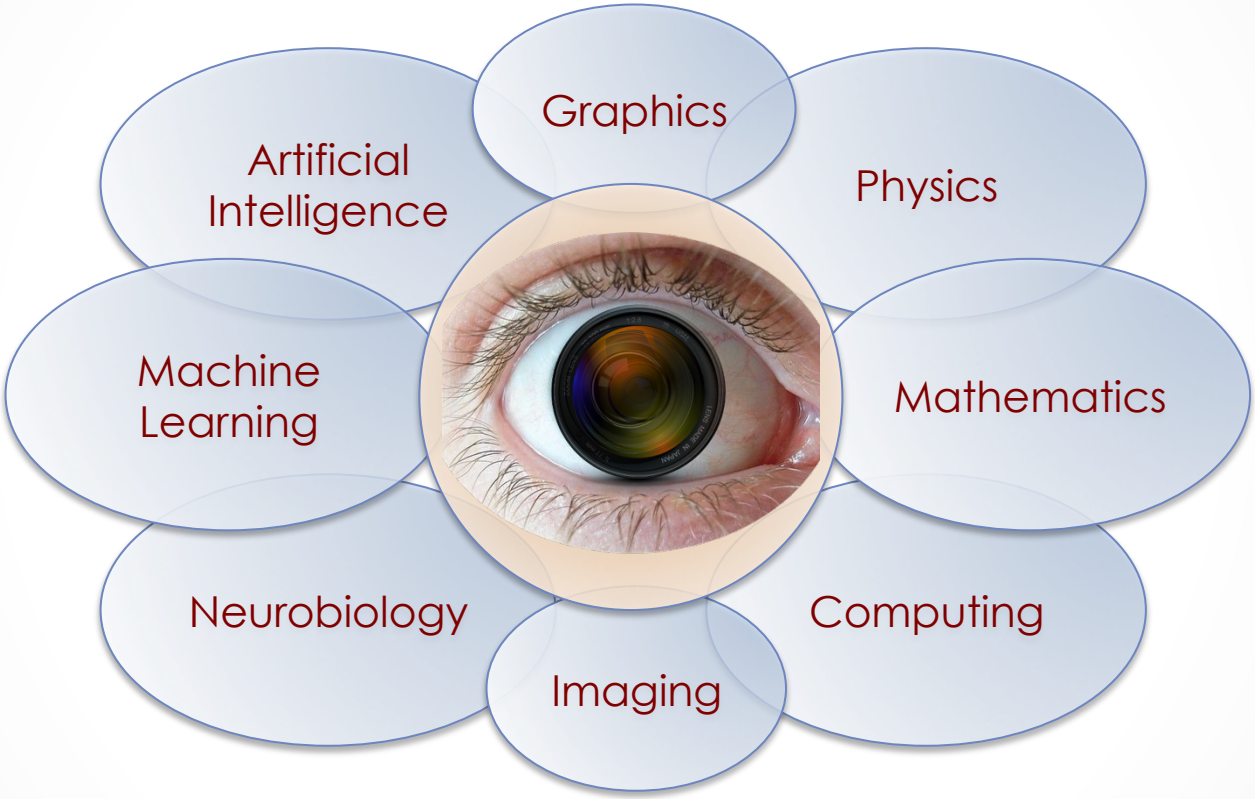
CSE578: Computer Vision

Spring 2017: Introduction



Anoop M. Namboodiri

Center for Visual Information Technology

IIIT Hyderabad, INDIA

Topics Covered

Computer vision has grown really fast recently!

We will sample some aspects in this course.

- Geometry of imaging:
  - Camera, projection, and estimation
  - Geometry of one, two, and three views
  - Camera and structure recovery from images
- Segmentation and grouping
  - Optimization methods in vision
- Visual recognition
  - Object detection
  - Object recognition & retrieval
  - Scene classification

Pre-Requisites for the Course

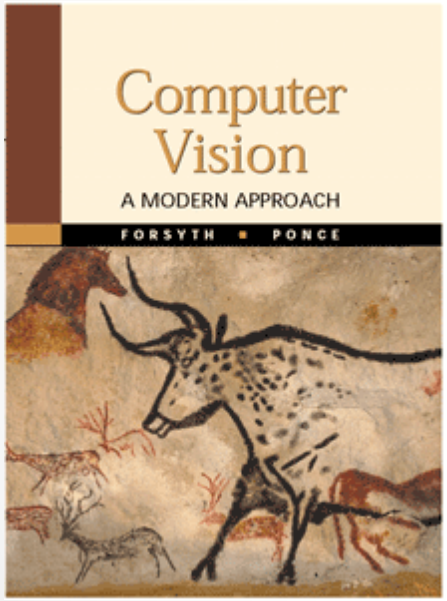
- Linear algebra and good mathematical outlook
  - Vectors, matrices, vector spaces, eigenvalues, singular values
  - 2D/3D geometry: Lines and planes in 2D and 3D.
- Image/Signal processing
  - Filtering, edge detection, segmentation
  - Transforms, analysis
- Pattern Analysis, Algorithms, Programming
  - Features, classifiers
  - Training, testing, validation
  - Matlab, OpenCV, C++

Brush up these topics if you aren't certain. A reading list of online material will be prepared for the preliminaries

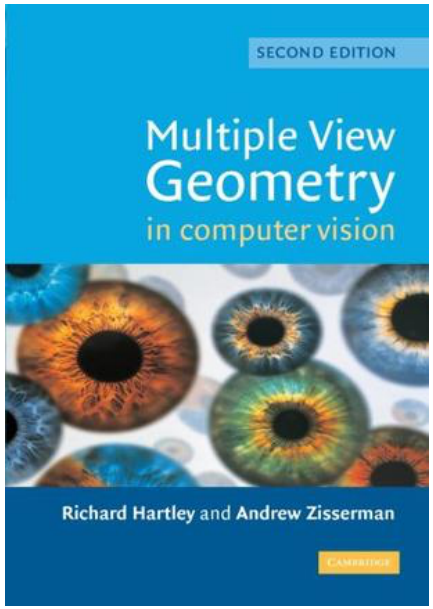
Reference Books

- No single textbook

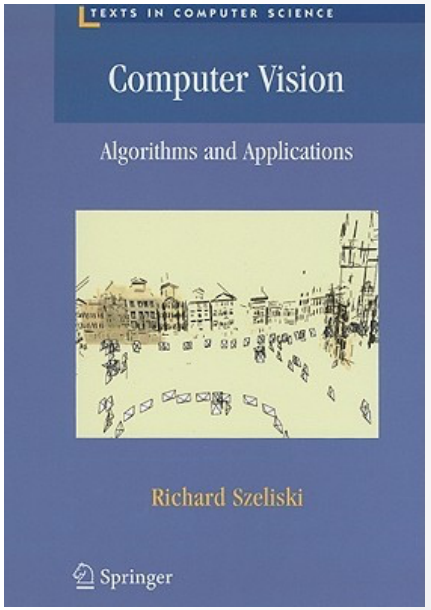
Forsyth & Ponce
Indian Edition



Hartley & Zisserman
Indian Edition



Rick Szeliski
Free online



And read many recent papers

Administrivia

- Grade Distribution (Tentative)
  - Mid-Sem Exam: (20%)
  - End-Sem Exam: (30%)
  - Assignments: Programming Assignments (30%)
  - Project: 1 project in groups of 2 or 3 (20%)
  - Bonus Project: For Fun and Extra credit (tentative)
- This is an advanced elective that you opted for
  - We expect you to work hard to learn it well.
  - Class participation lifts the level of the class
  - We don't want credit-seekers or resume-padders here
  - The more your effort, the more everyone benefits!

Class Etiquettes

- Be in the class before 2:00pm
  - No attendance for late comers
- Switch off your cell phones
  - Use of cell phone is strictly prohibited in the class.
  - If you are caught checking your phone/tablet, it will be confiscated.
- If you have a doubt, ask. Others are likely to have the same doubt

What is Computer Vision?

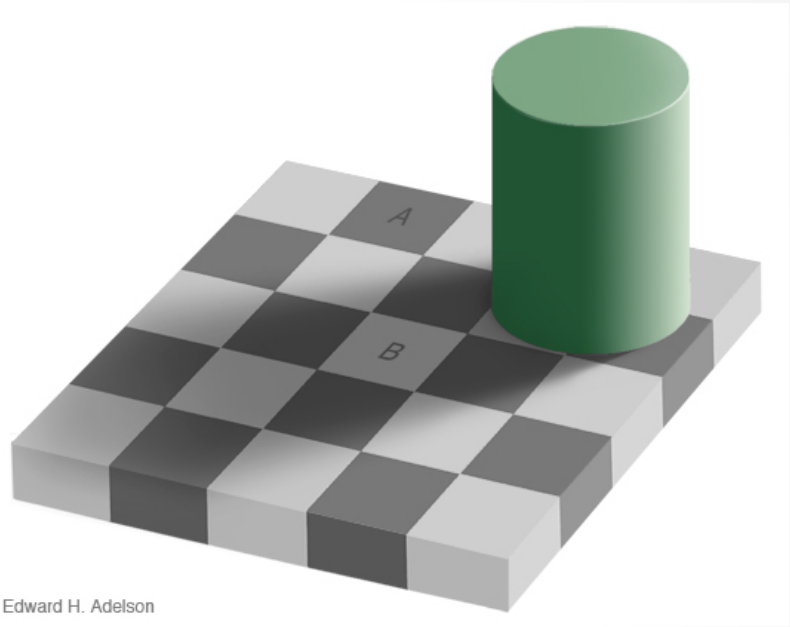
- Computer processing of visual inputs: images and videos.
- Making sense out of them. Describing them.
- Does computer vision mimic the human vision?
  - Certainly in many of its goals
  - Why? Human vision is among the best!
  - Sophisticated and efficient but not understood well
- Should computers process visual inputs like humans?

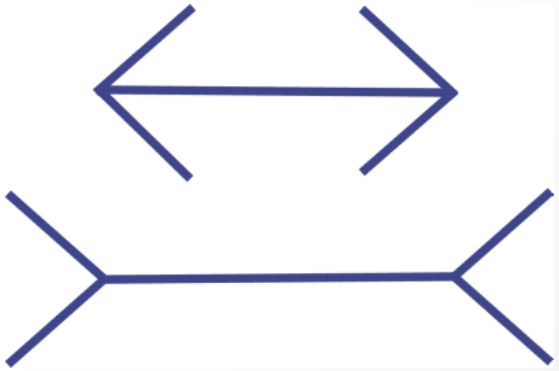
Not necessarily!

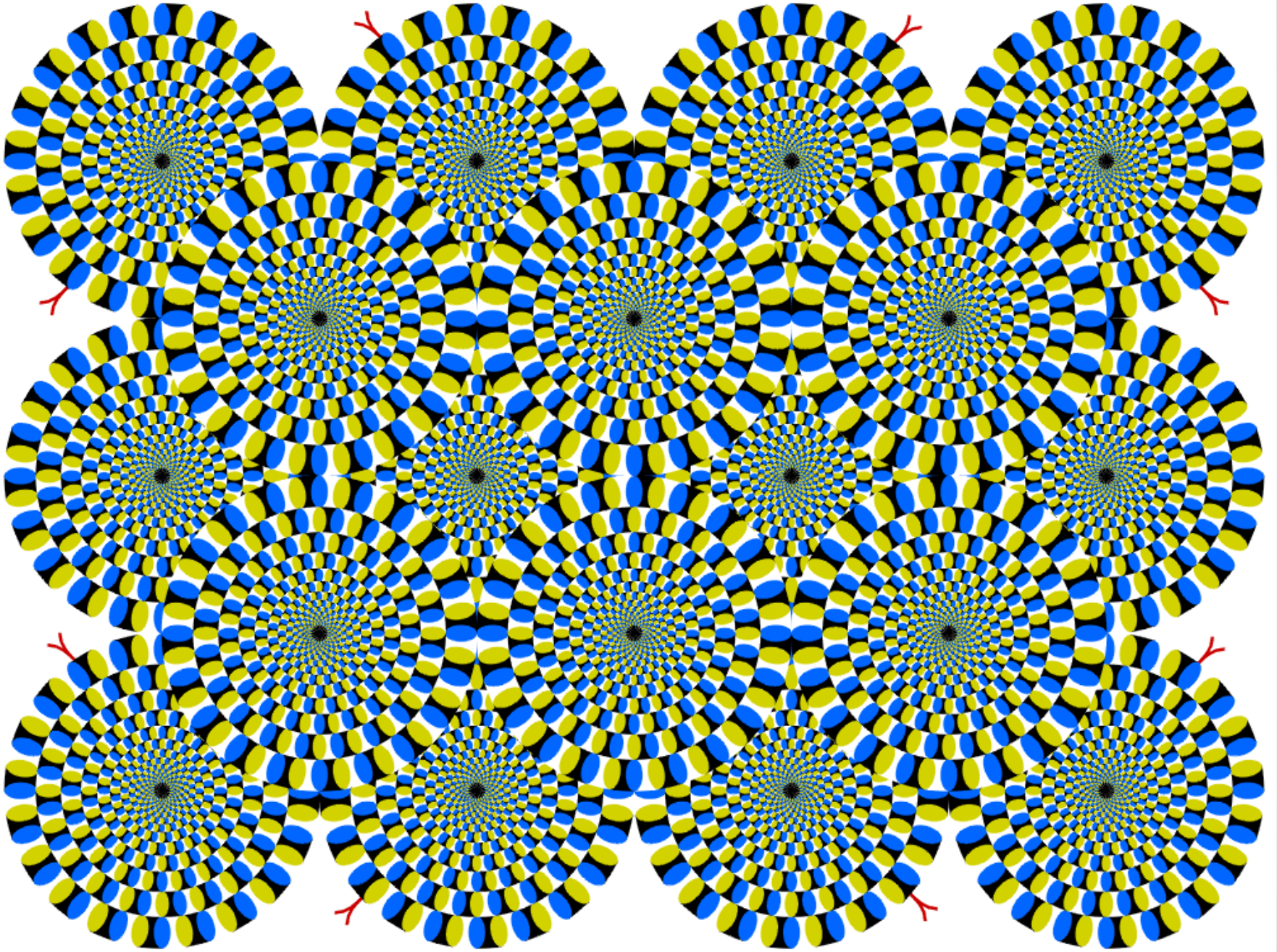
  - Human visual system need not limit computer vision
  - We draw inspiration from it as often as is convenient

Human perception has its shortcomings…

Sinha and Poggio (Image: Ron Rensick)







Three "Urges" on seeing a Picture*

1. **To group** proximate and similar parts of the image into meaningful "regions".

   Called segmentation in computer vision.

2. **To connect to memory** to recollect previously seen "objects".

   Called recognition in computer vision.

3. **To measure** quantitative aspects such as number and sizes of objects, distances to/between them, etc.

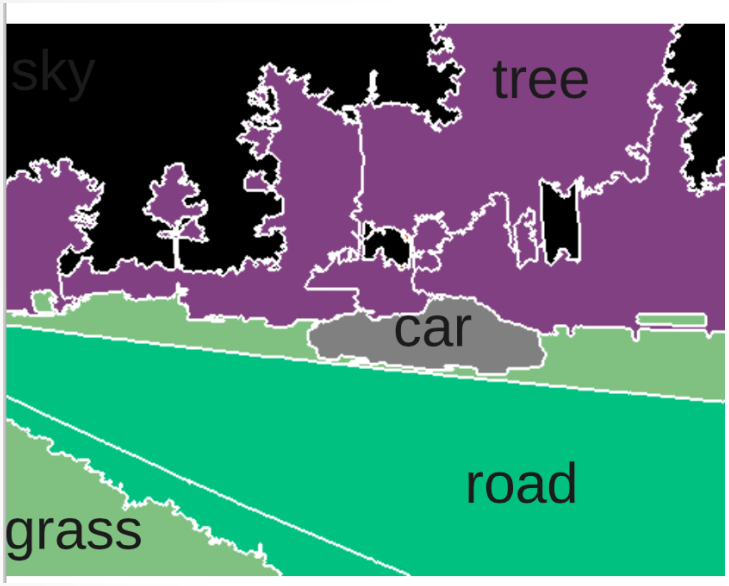
   Called reconstruction in computer vision.

**Jitendra Malik; Mysore Park, Dec. 2011*

Urge to Group

- We don't see individual pixels (like the computer does!).
- We see groups of pixels together.
- What is the basis for "correct" grouping?

Urge to Group



- Group similar pixels together as objects.
- Group semantically meaningful pixels together as objects.
- Is appearance similarity the same as semantic similarity?

Urge to Touch Memory







Recognizing objects from (visual) input is fundamental to human cognition of the world.

Urge to Measure

We measure number, distance, colour, etc, from images. Sometimes correctly too!

Scene Interpretation

Segmentation and Labeling

1.Hand-carved Shesham wooden screen

2.Wooden flowers

3.Wicker basket

4.Pair of hand-carved Thai candlesticks

5.Indonesian rattan screen

6.Dhurry covered armchair

7.Hand-painted chest

8.Striped wooden Indian candlestick

9.Stone terracotta Thai

10.Moroccan ceramic candlestick

11.Blue Egyptian glass decanter

12.Bronze goblet-shaped candlesticks

13.Painted wooden Indian elephant

14.Blue Egyptian glass goblets

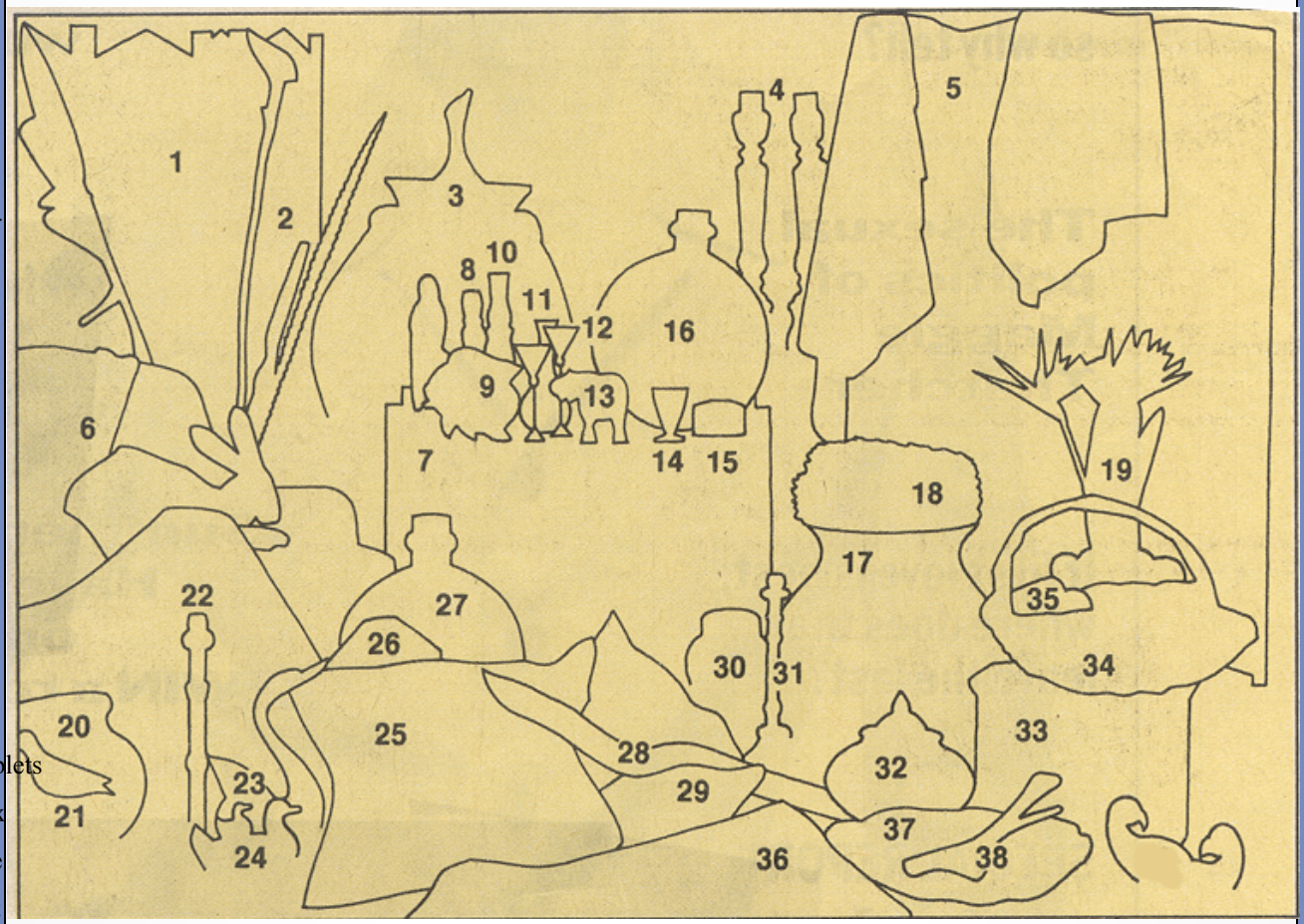
15.Indian brass filigree box

16.Painted Indian oil bottle

17.Large African water pot

18.Philippino twig basket

19.Philippino bamboo covered urn



20. African cooking pot

21. Decoy bird

22. Painted candlestick

23. Thai wooden swan

24.Carved wooden duck

25. Embroidered mirror cushion covers

26. Green hexagonal Indian box

27. Painted Indian oil bottle

28. Joint wooden snake

29. Black embroidered cushion

30.Moroccan ceramic jar

31.Painted wooden candlestick

32.Thai pot with lid

33.Octagonal Indian box

34.Shallow twig baskets

35.Mexican paper mache fake fruit and vegs

36.Nakshe Kantha Bengali wall-hanging

37.Wooden shell bowl

38. Wooden servers

Computer Vision

- Goal: Extract all possible information about a visual scene by computer processing

  What? When? Where? Who? How? How many?

- Over 50% of the brain is devoted to vision for humans. – Must be important to us!
- Why is it difficult?

Chairs and Chairs

- Which are chairs?
- Large intra-class variations
- How do we describe a chair?
- Basic property: Sittability!
- We infer a lot from pictures. Can we instruct a computer to do the same?
- Do we understand how we infer?



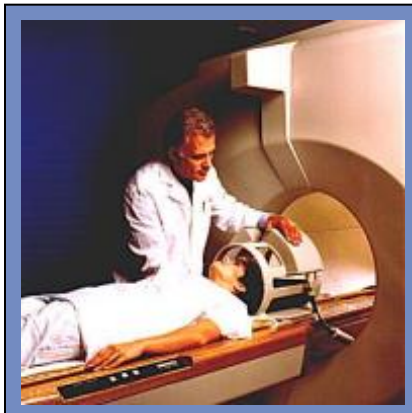


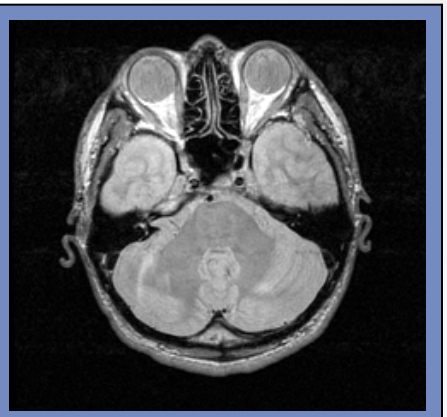




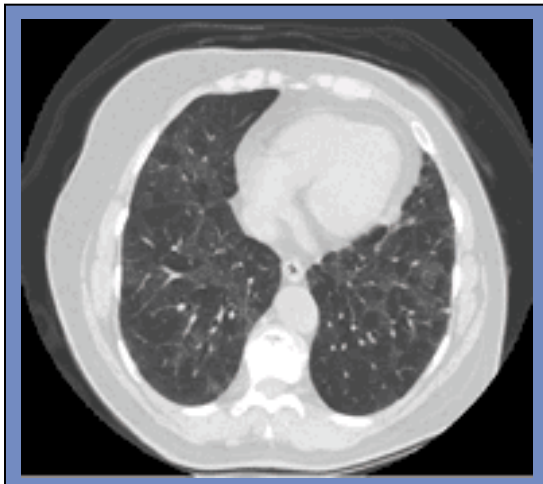
Applications: Medical

Computer Assisted Surgery



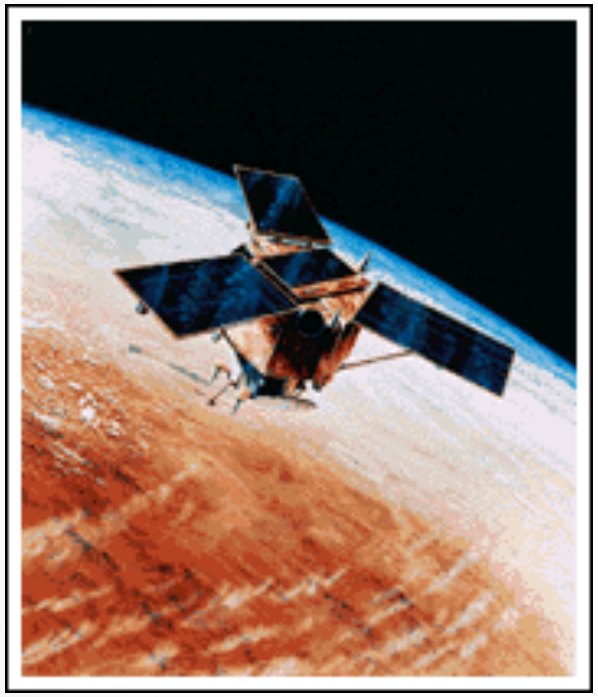


CT Scan

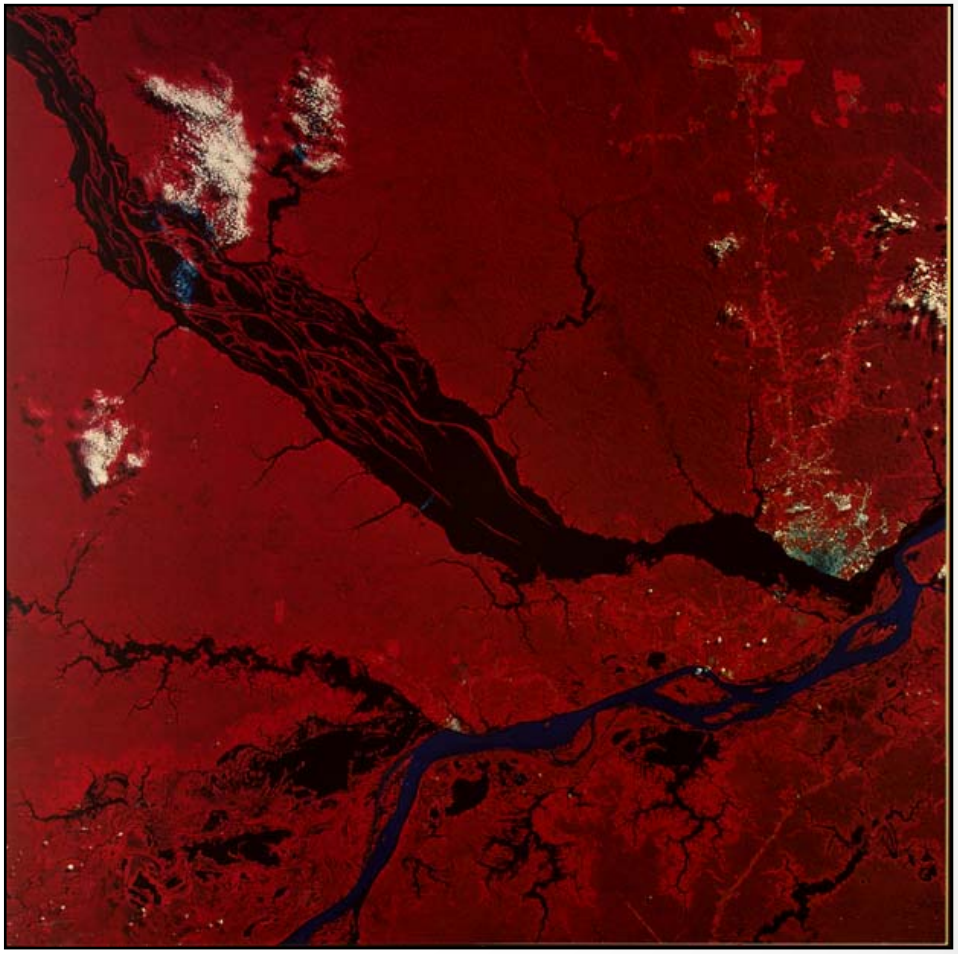


Segmentation

Applications: Space Imaging



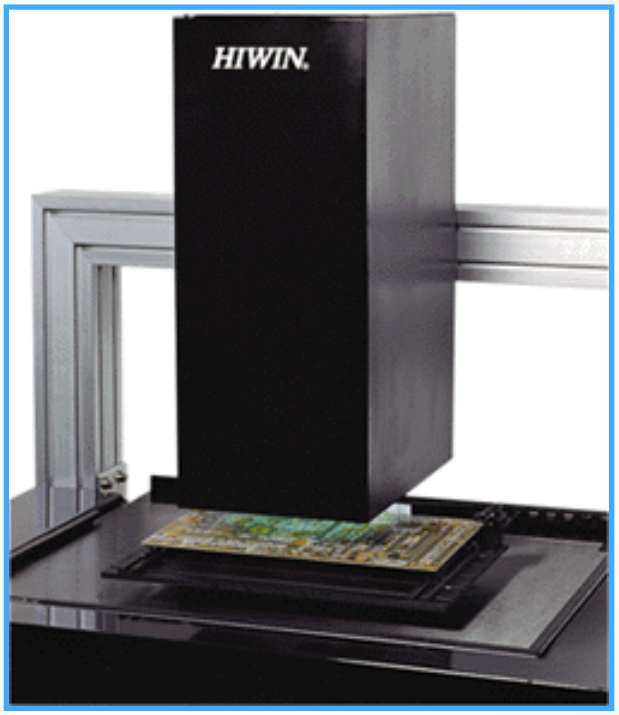
Ikonos



Rio Negro (black) meets Amazon (blue)

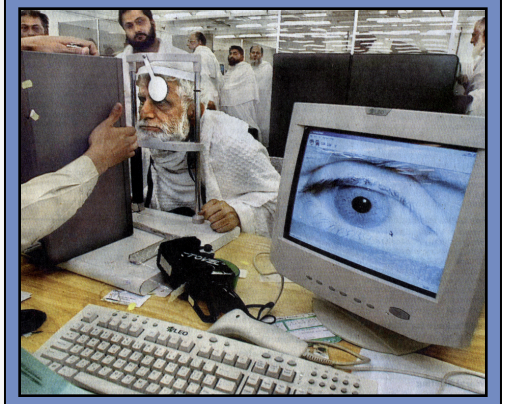
Applications: Automated Inspection

Manual PCB Inspection

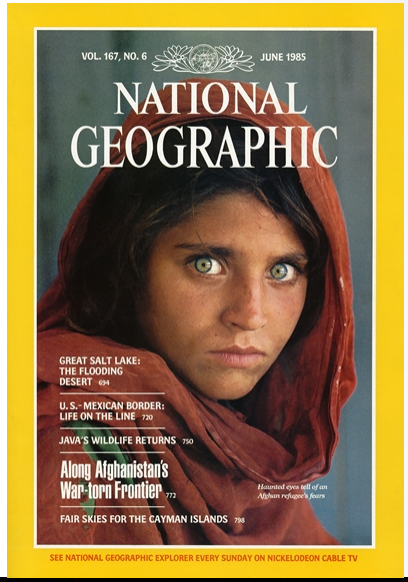


Automated PCB Inspection

Applications: Biometrics

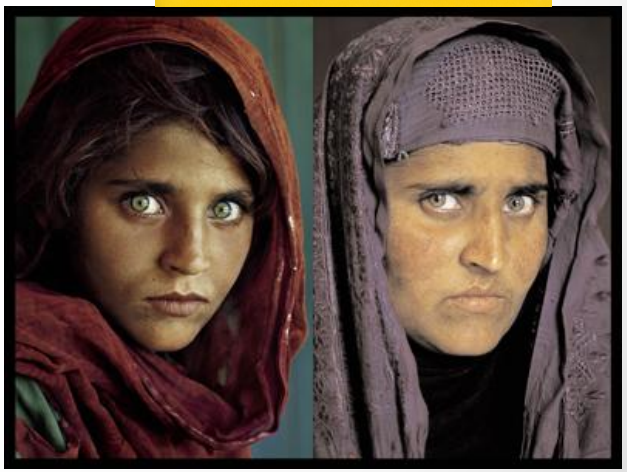


Travel



Computer Access

Disney Land



Applications: Broadcasting

Virtual Replay

Chroma Keying: Replacing Backgrounds

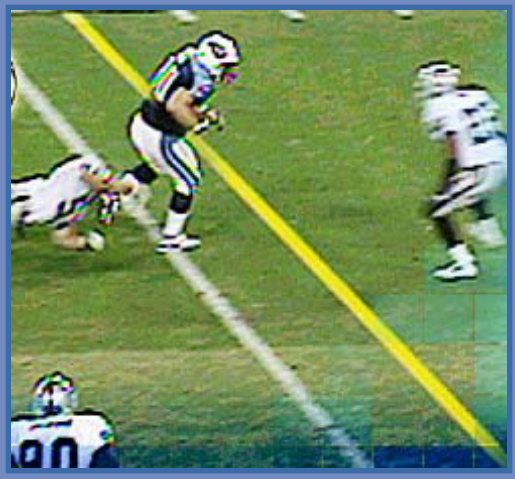
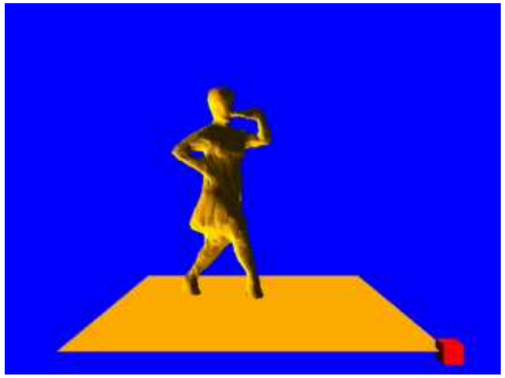


Image Overlay

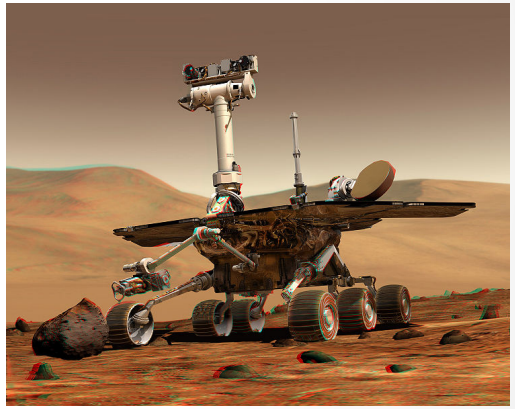
3D Shape and Motion Recovery

- Structure light scanner, laser range finder
- Multicamera stereo, structure recovery
- Reverse Engineering
- Virtualized/Augmented reality



Applications: Others

- Surveillance
- Automated Assembly
- Mail Sorting
- Face detection (photography)
- Robot Navigation
- Content-Based Image Retrieval
- Entertainment
- And many more… with your help…



Why Automated Vision?

1. High reliability
2. High repeatability
3. More objective evaluation
4. Lower cost
5. Higher speed
6. Ability to operate in hazardous environments

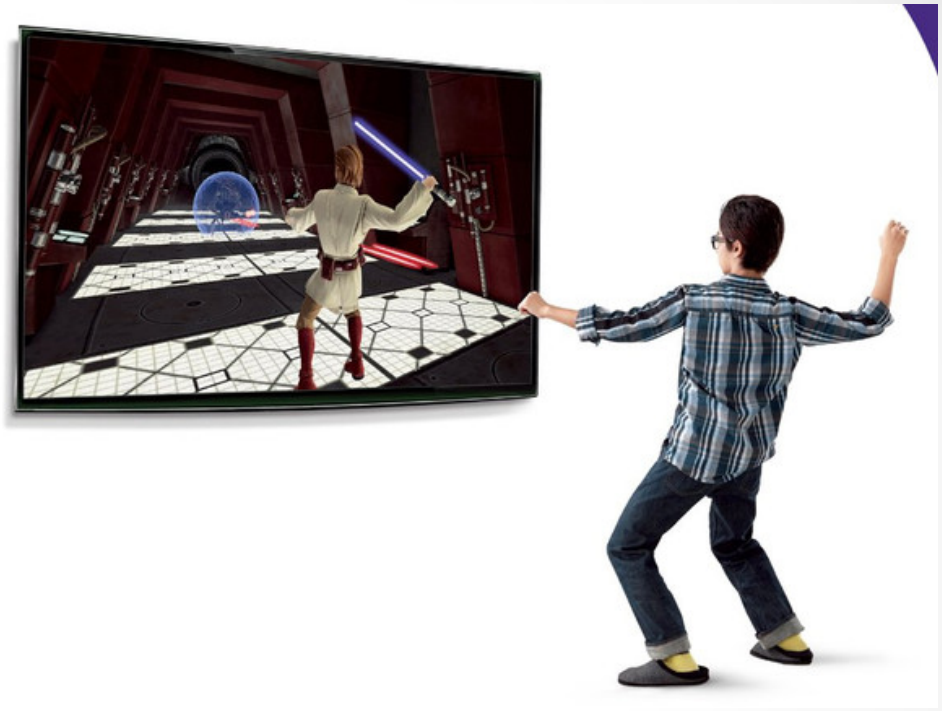
No general purpose machine vision system exists.

Recent: Structure from Motion

- Approximate 3D structure from an unstructured collection of images! [PhotoTourism, SIGGRAPH2006]
- PhotoSynth
- Autodesk 123D: Your pictures to model
- And many more to follow soon

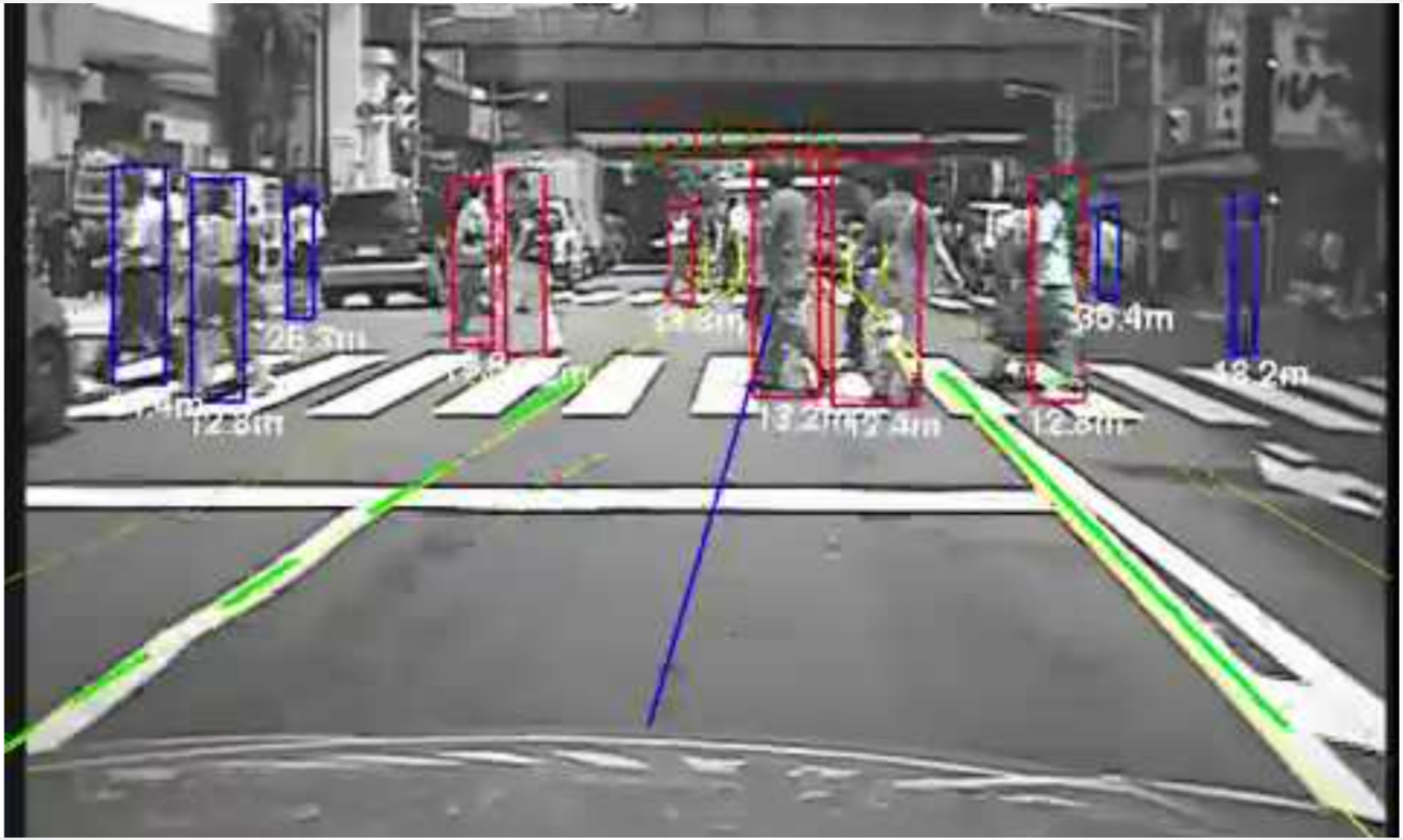
Recent: Natural Gaming

Microsoft Kinect



- You are the controller. Interact naturally with the game.
  - Fastest Selling Electronic Device Ever: 80 lakh units in 60 days!!
- Finding great use in Computer Vision, Robotics, etc.

Recent: Automotive Safety



Can help avoid accidents greatly!

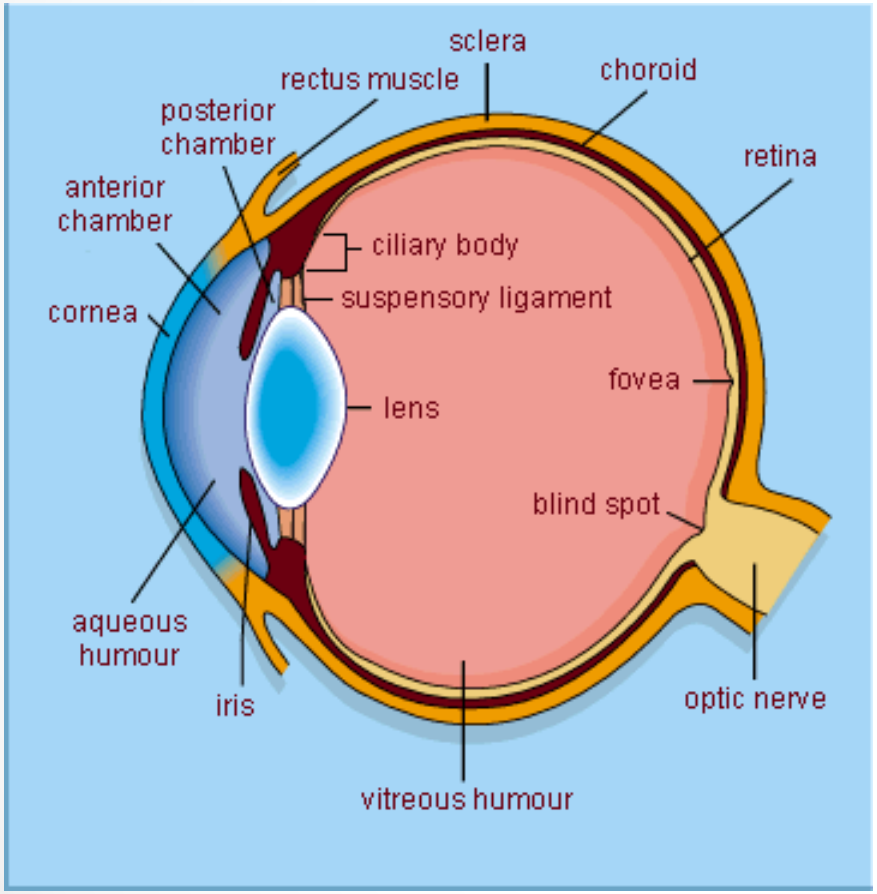
The Real Problem

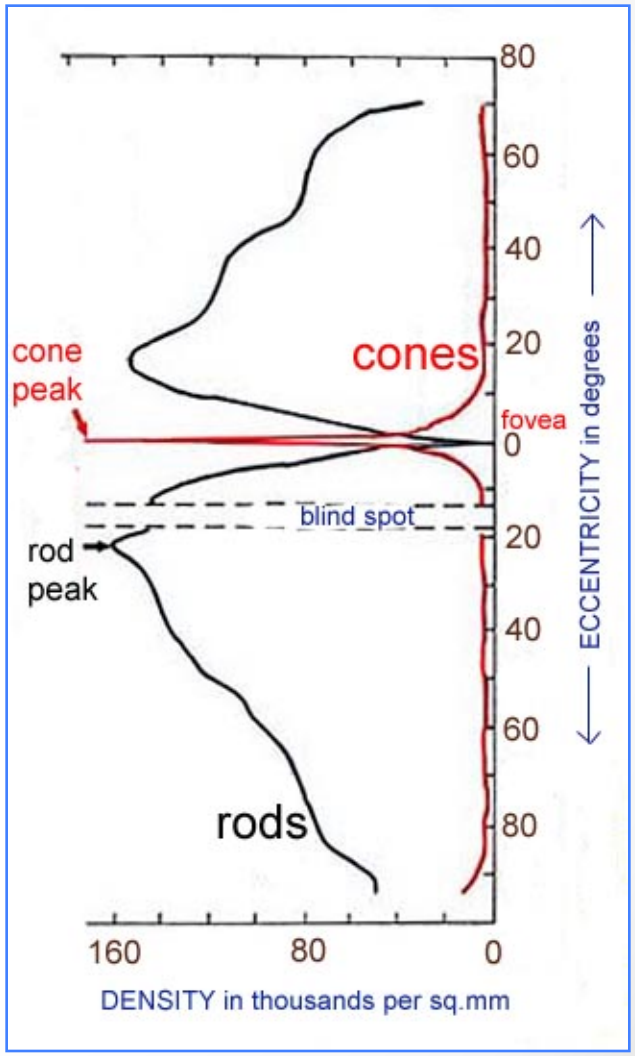
Develop something similar for Indian roads!

What More is Possible?

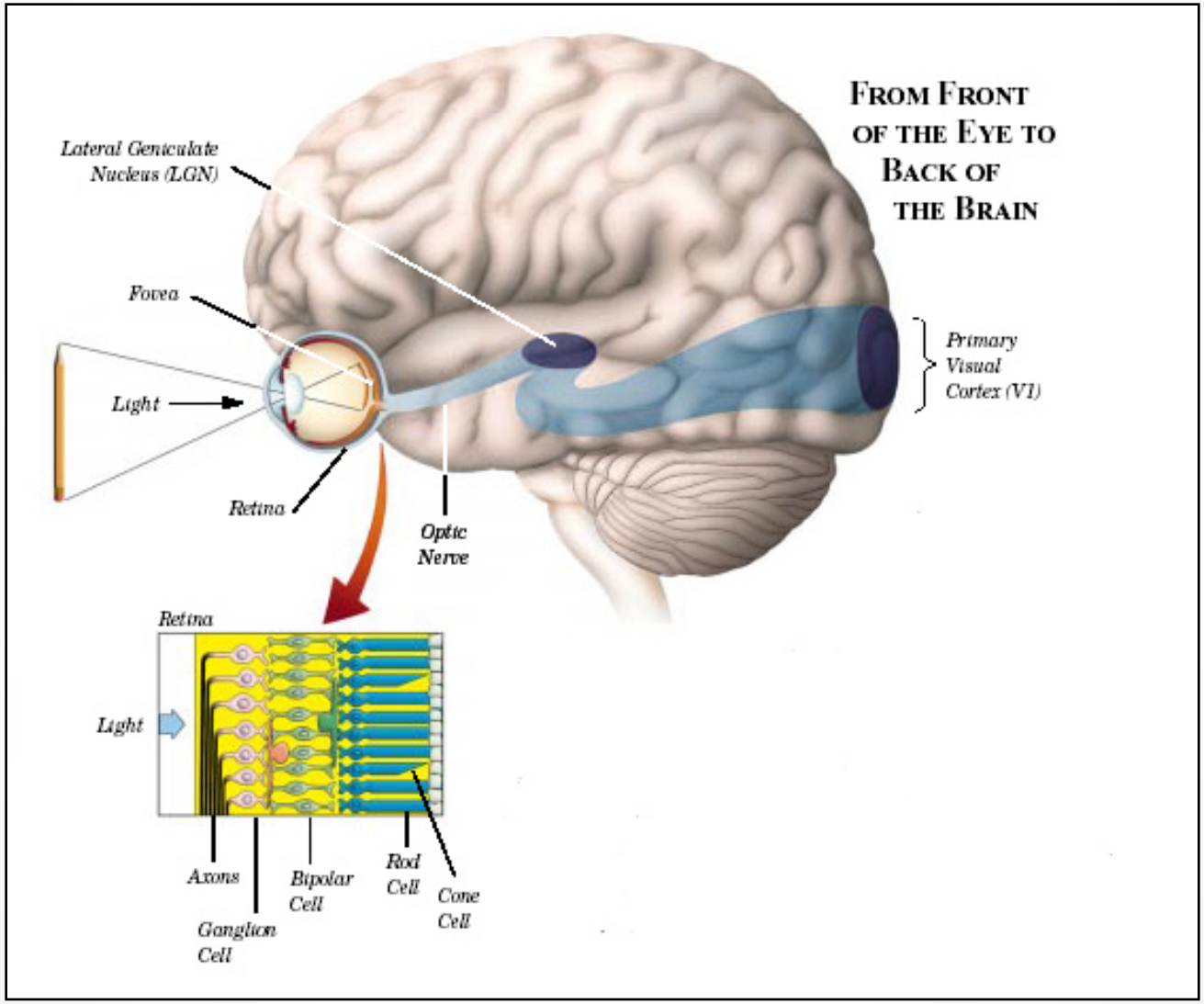
- Much much much more .....
- The journey has just begun for computer vision.
- Large amount of data, high computing power, machine learning algorithms continue to transform computer vision.
- Big things are yet to come, however.

Human Eye


sclera
choroid
rectus muscle
posterior chamber
retina
anterior chamber
ciliary body
suspensory ligament
cornea
fovea
lens
blind spot
aqueous humour
iris
optic nerve
vitreous humour

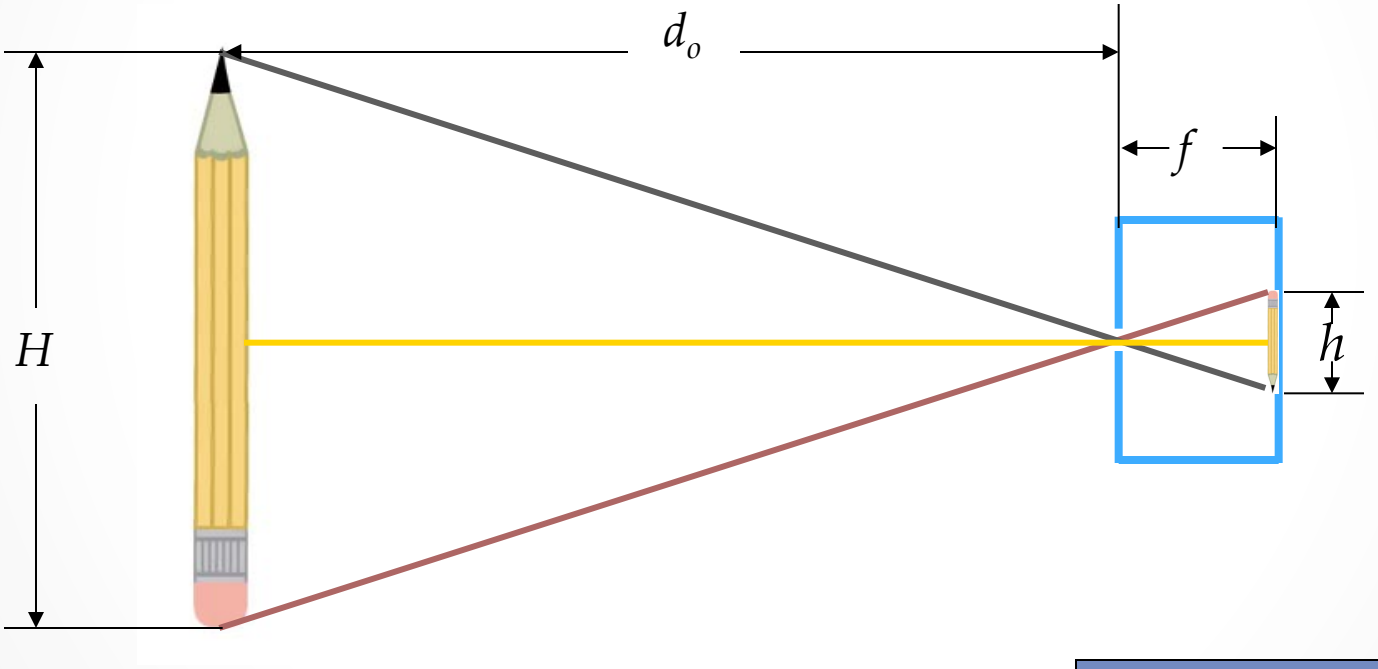

80
60
40
20
0
20
40
60
80
cone peak
cones
fovea
blind spot
rod peak
rods
ECCENTRICITY in degrees
160
80
0
DENSITY in thousands per sq.mm

Human Visual System



http://www.hhmi.org/senses

The Pinhole Camera



$$h = f \,.\, H / d_o$$

Pinhole Camera in practice