COL 780
Computer Vision

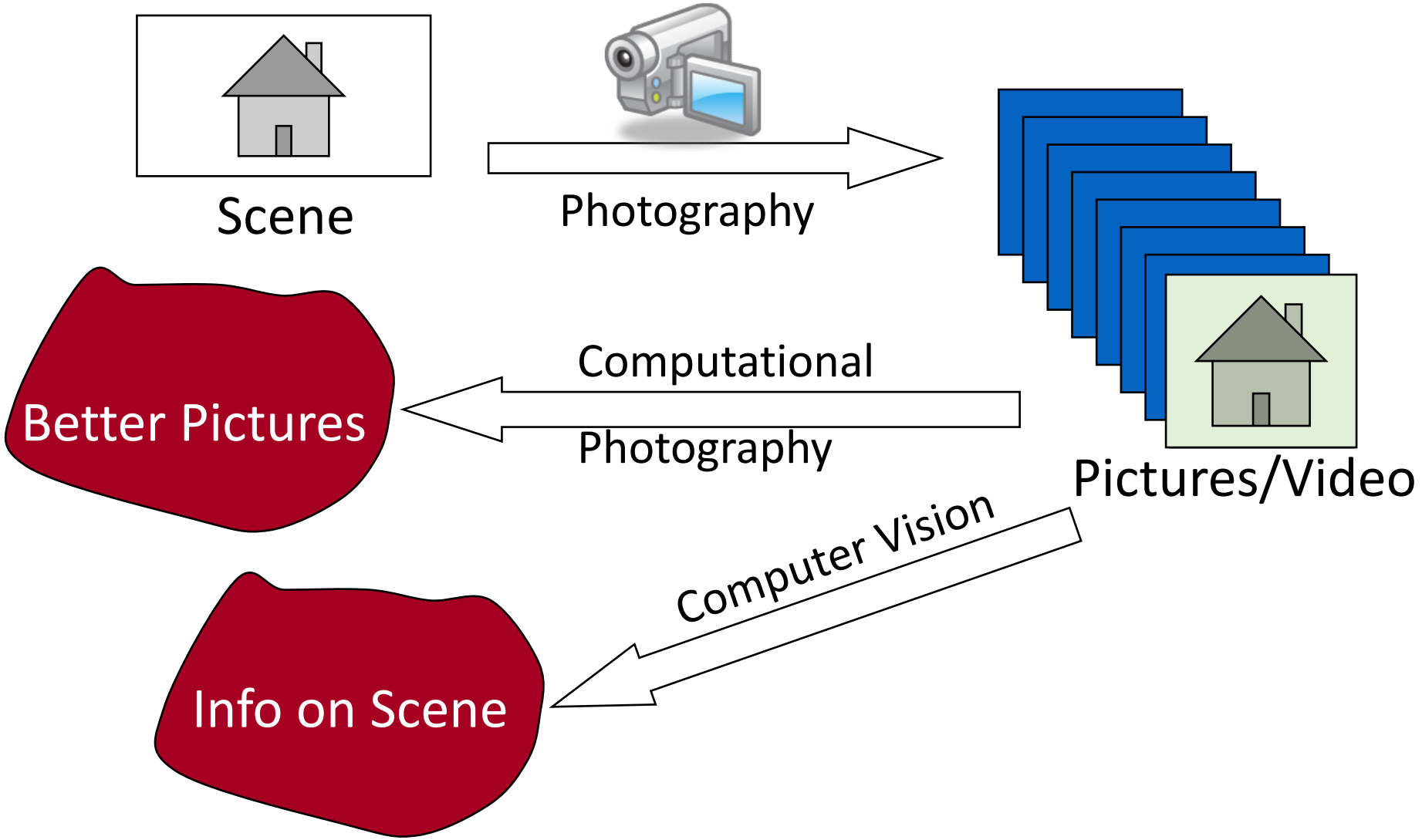
Chetan Arora

Chetan Arora
Department of Computer Science and Engineering, IIT Delhi

http://fortune.com/2015/10/16/how-tesla-autopilot-learns/

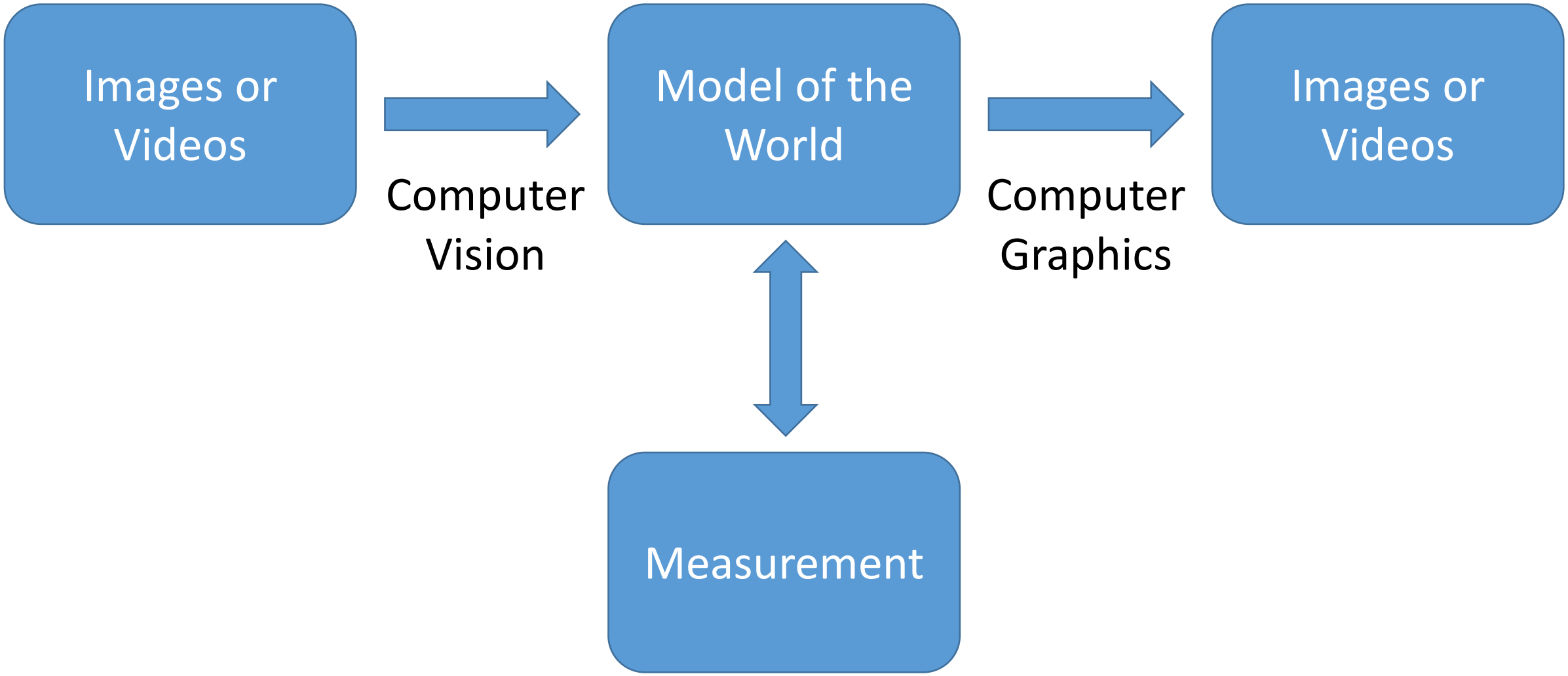
Chetan Arora
Department of Computer Science and Engineering, IIT Delhi

Computer Vision



Chetan Arora
Department of Computer Science and Engineering, IIT Delhi

Computer Vision Vs Graphics

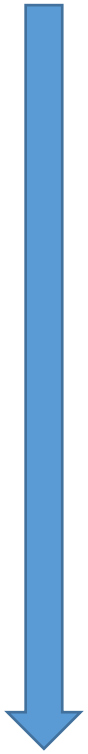


Chetan Arora
Department of Computer Science and Engineering, IIT Delhi

Stages in Computer Vision

- **Physics:** Image Formation (Light, Reflectance)
- **Physics:** Cameras: Optics (Lens), Sensors (CCD, CMOS)
- **Image Processing:** Coding (Transmission, Compression)
- **Comp. Photo.:** Enhancement (Noise, Resolution, Colors)
- **IP-CV:** Feature Detection (Objects, Actions, Motion)
- **Computer Vision:** Scene recovery (3D, Reflectance)
- **Computer Vision:** Object Classification / Recognition
- **Human and Machine Vision:** Visual Perception
- **Robotics:** Control Action (Autonomous Driving)

Low Level Vision



High Level Vision

Have You Used Computer Vision?

- **Laptop:** Biometrics auto-login (face recognition, 3D), OCR

- **Smartphones:** QR codes, computational photography (Android Lens Blur, iPhone Portrait Mode), panorama construction (Google Photo Spheres), face detection, expression detection (smile), Snapchat filters (face tracking), Google Tango (3D reconstruction), Night Sight (Pixel)

- **Web:** Image search, Google photos (face recognition, object recognition, scene recognition), Facebook (image captioning), Google maps aerial imaging (image stitching), YouTube (content categorization)

Have You Used Computer Vision?

- **VR/AR:** Outside-in tracking (HTC VIVE), inside out tracking (simultaneous localization and mapping, HoloLens), object occlusion (dense depth estimation)

- **Motion:** Kinect, full body tracking of skeleton, gesture recognition, virtual try-on

- **Medical Imaging:** CT / MRI reconstruction, assisted diagnosis, automatic pathology, endoscopic surgery

James Tompkin, Brown University

Have You Used Computer Vision?

- **Industry:** Vision-based robotics (marker-based), ANPR (number plates), surveillance, drones, shopping
- **Transportation:** Assisted driving (everything), face tracking/iris dilation for drunken-ness, drowsiness.
- **Media:** Visual effects for film, TV (reconstruction), virtual sports replay (reconstruction), semantics-based auto edits (reconstruction, recognition)


Chetan Arora
Department of Computer Science and Engineering, IIT Delhi

Nature: Vision ≡ Smart ≡ Moving

- Only smart and moving organisms can see!
  - Bacteria & Plants do not see

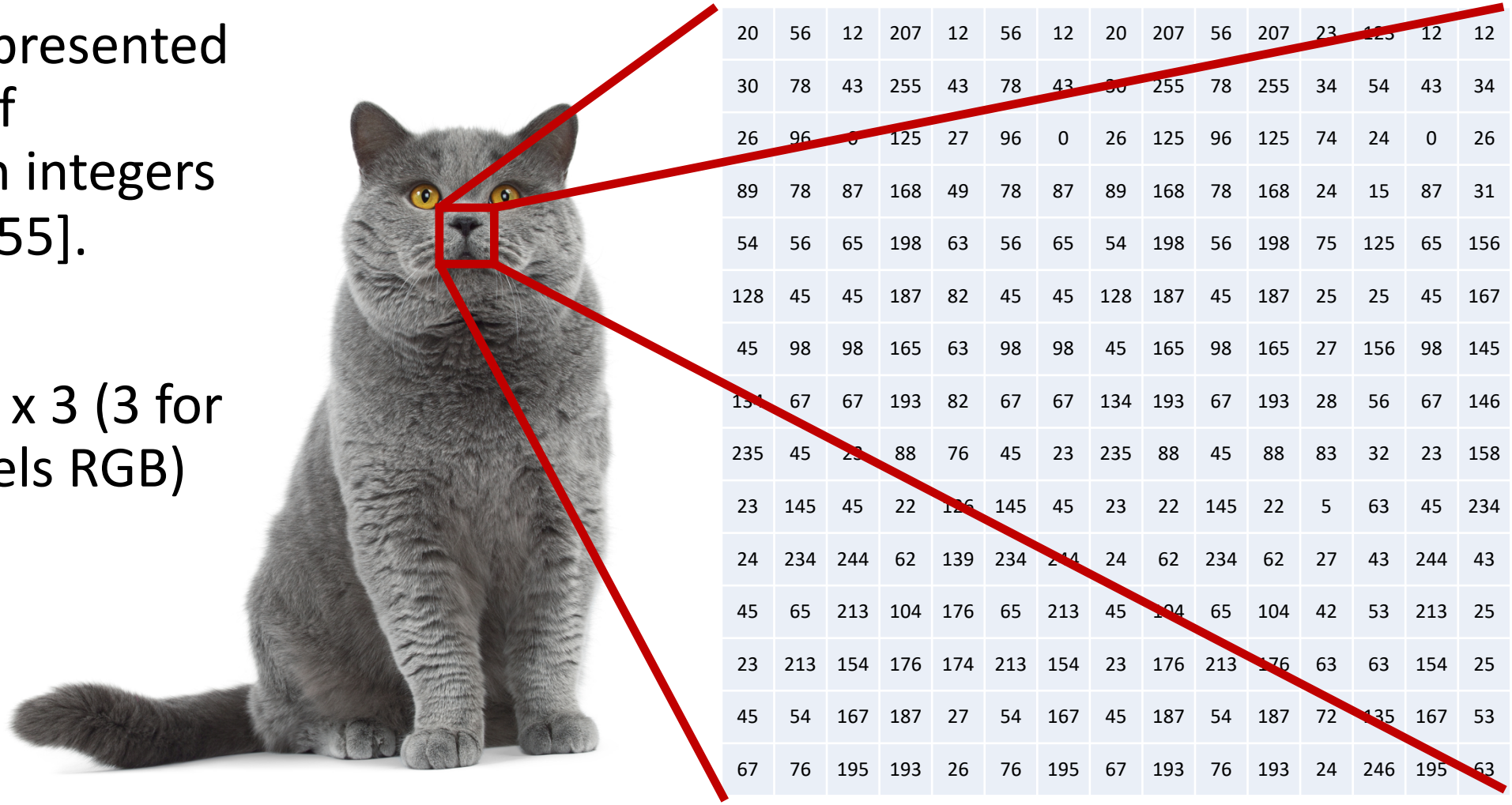
- Visual recognition at early development
  - Babies recognize and track the mother very early

- Most of the human brain is involved in vision processing.

Chetan Arora
Department of Computer Science and Engineering, IIT Delhi

Why is Computer Vision Hard: Semantic Gap

- Images are represented as 3D arrays of numbers, with integers between [0, 255].
- E.g. 300 x 100 x 3 (3 for 3 color channels RGB)

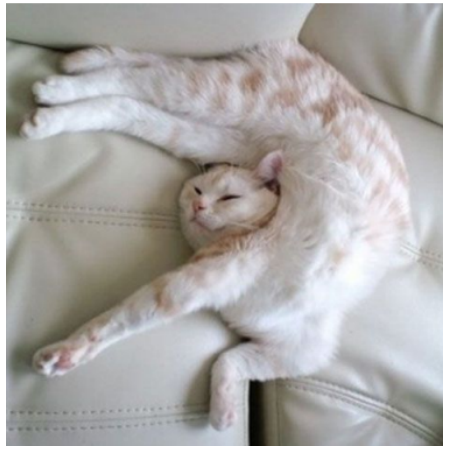


Chetan Arora
Department of Computer Science and Engineering, IIT Delhi

Visual Recognition Challenges

Intra class variation

Background Clutter



Deformation

Illumination

Occlusion

Size


Chetan Arora
Department of Computer Science and Engineering, IIT Delhi

Computer Vision Vs Deep Learning

- Just like for many other domains, deep learning is an enormous disruption to the computer vision.

- Since 2012, many of the state of the art computer vision techniques are based on deep learning models.

- For many of the problems like image classification or face recognition it has been shown to even surpass human performance.


Chetan Arora
Department of Computer Science and Engineering, IIT Delhi

Is there More to Computer Vision than DL?

- Not every problem is a learning problem. Geometry? Measurement?
- DL is data hungry. Not practical for many problems. Medical Imaging?
- Lots of data = lots of potential bias in the data.
  - Needs understanding of possible failures.
  - Responsible approach.
  - Techniques to overcome bias

Chetan Arora
Department of Computer Science and Engineering, IIT Delhi

Course Contents

Image Processing

- Camera model. Calibration, multi-views projective geometry and invariants.
- Feature detection, correspondence and tracking.

Applications

- 3D structure/motion estimation.
- Application of machine learning in object detection and recognition, category discovery, scene and activity interpretation.


Chetan Arora
Department of Computer Science and Engineering, IIT Delhi

Evaluation

- Mid-term: 10
- Major Exam: 20
- Assignments(3): 45
- Project(1): 25

Course Policies

- Scoring less than 40% marks in any of the exams/assignments will lead to failing the course.
- Any plagiarism detected in any of the assignments will lead to failing the course.
- No deadline extension in any submission.
- No auditing the course.

Chetan Arora
Department of Computer Science and Engineering, IIT Delhi

Course Feedback in Last Delivery

What did you dislike the most about this course?

- The assignment deadlines must be flexible instead of being hard.
- There was no late submission policy in the course.
- I think too many topics are covered to give a broader perspective of the subject; It was in one way necessary but that also means less time is allocated to some vital topics which increased the vagueness index
- Lectures were uploaded late. If you miss a class, you are badly screwed.
- Slightly weak presentation of the topics after introduction to ML. All the Deep architectures were taught vaguely
- The syllabus was huge and lectures very less comparatively.