# 1A-L1 Intro

- 2017/11/01 20:59
- Summary
  - CV is
    - to analyze images and to extract content from images/videos
      - So essentially the goal is image and video understanding which means labeling interesting things in an image and also tracking them as they move.
    - to introduce the mathematical and computational methods to provide you with core concepts of how can a computer be built to interpret images.
    - computational photography vs. Image Analysis vs. CV
    - interplay between the theory, the algorithm, and the images.
  - policy
    - Write you own code so as to undertand the algorithm.
  - why study CV
    - become ubiquitous in both production and consumption
    - It is a really deep and cool set of problems!
  - what are the state-of-art of CV now
    - OCR
    - Face detection
    - objection detection in mobile device
    - special effect
    - Smart cars
    - Sports
    - Vision based interaction and games
    - Security and Surveillance
    - Medical Imaging
  - SW
    - Matlab is a little bit easier to implement, while python/C++ is closer to the code you'll write in you future job. Both ok here

---

- Goal
  - covering the foundational aspects of how **to analyze images and to extract content from images.**
    - i.e. how can we build a computer or a machine that can see and interpret an image.
    - foundational ?

- **the mathematical and computational methods** to provide you with core concepts of how can a computer be built to interpret images.
    - interpret
        - **In Computer Vision we are interested in extracting information.** Many want to **go beyond processing an image** to really knowing what is inside the image, what's the content of the image.
    - So we will learn the math and the basic concepts how to compute with an image and extract information from it.

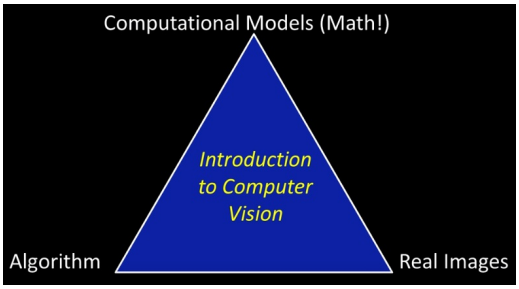
- what is CV
    - Computer Vision is really about **analyzing images and videos to extract information from them,**
        - **i.e. to know what is it representative of. input an image, output something that has some meaning to it or the understanding of it.**
    - So essentially the goal is image and video understanding which **means labeling interesting things in an image and also tracking them as they move.**
    - image type
        - images are of real scenes
        - X-ray
    - computational photography vs. Image Analysis vs. CV
        - Computational photography is really about **capturing a light from a scene to record a scene into a photograph** or such other related novel artifact that showcases the scene.
        - Image analysis is done **to support the capture and display of the scene** in novel ways.
        - Computer vision is really about **interpreting and analysis of the scene.** That is what is **the content of the image**

- why study CV
    - Images (and movies) have become ubiquitous in both production and consumption.
    - Therefore applications to manipulate images (movies) are becoming core.
    - As are systems that extract information from imagery
        - • Surveillance
        - • Building 3D representations
        - • Motion capture assisted
    - But most of all…It is a really deep and cool set of problems!

- what are the state-of-art of CV now
    - Sum:

        - these are all invented in less than 10 years, some even in less than 5 years. so the field is rapidly changing and there're tremendous amount of oppotunity right now where you can make lot of money
    - OCR: optical character recognition
        - ZIP code
        - check
    - face detection
        - face login
        - face location in the camera
        - blink, smile detection
    - objection detection in mobile device
        - #requrie a lot of compute power, and the computer power gets smaller
        - AR
        - Lane Hawk, supermaket
        - google glasses
    - special effect
        - shape capture
            - The Matrix
        - motion capture
            - Pirate of the Caribbean
            - #created by CGI(Computer-generated imagery)
        - earth viewers - 3D modeling
            - Microsoft Virtual Earth, Google Earth
    - Smart cars
        - autonomous driving
            - Stanley, Sebastian
        - detection
            - pedestrian
            - traffica light/sign
            - lane
            - car
    - Sports
        - football, lines
    - Vision based interaction and games
        - Nintendo Wii introduced CV into games
        - real game changer: Microsoft Kinect
            - produce
                - depth image, using gray scale to show the distance of objects to the camera

        - skeletal descriptions: recover the skeleton geometry of people from depth image
            - one frame one time, no track requred. robust!
        - a powerful interface for taking info into the system. open source
    - Security and Surveillance
        - people or port surveillance
    - Medical Imaging
        - 3D imaging MRI, CT
        - image guided surgery
- Why CV is hard
    - $ searching
        - A computer has to learn how to do it, or has to be programmed. But programmation implies that the system will be able to cope with already known patterns and context ; this is currently done in industry and works very well. Without context, even AI systems are not able to cope with an image they do not know already, as a human can.
    - #A photometer, generally, is an instrument that measures light intensity or the optical properties of solutions or surfaces
    - #A photon is a type of elementary particle. The photon has zero rest mass and always moves at the speed of light within a vacuum.
- Vision is NOT Image Processing
    - • In the previous example, the two squares have exactly the same measurement of intensity.
    - • So, seeing is not the same as measuring properties in the image.
    - • Rather, **“seeing” is building a percept of what is in the world based upon the measurements made by an imaging sensor.**
    - But actually, what happened there is that, there’s ambiguity as to what the ball is doing, and your brain is creating the story, it is making the description, and **the difference between straight image processing and computer vision is building that description.**
- Course Overview




    - You can think of computer vision as being a relationship between sort of three ways of thinking about what goes on: computational

models, algorithms, and real images
        - computational models, often the math models, are the theory, expressing how to address the problem
        - algorithms is the solution based on the math models to solve these problems
        - real images with gound trurh help to measure the performance of the models and algorithms
    - The idea of the problem sets is to get you to understand sort of the interplay between the theory, the algorithm, and the images.
- Topic outline
    - 1. INTRODUCTION
    - 2. IMAGE PROCESSING FOR COMPUTER VISION
    - 3. CAMERA MODELS AND VIEWS
    - 4. FEATURES AND MATCHING
    - 5. LIGHTNESS AND BRIGHTNESS
    - 6. IMAGE MOTION
    - 7. MOTION AND TRACKING
    - 8. CLASSIFICATION AND RECOGNITION
    - 9. MISCELLANEOUS OPERATIONS
    - 10. HUMAN VISION
    - 8 problem sets
- Plicies
    - When you have a problem/question, go to the Piazza to ask or search for it.
    - Write you own code
        - it's the only way to let you get a good understanding about the algorithm i.e. that is why it is that certain code make things happen. Even if there's libraries, it may not work or work well on your specific problems. You'll know why when they don't work, if you've understood them before.
- Software
    - Matlab + image processing tool box / Octave + image processing tool box
    - or python / C++ with OpenCV
    - Matlab is a little bit easier to implement, while python/C++ is closer to the code you'll write in you future job. Both ok here.
    - Matlab
        -