Computer Vision

0. Overview

I-Chen Lin
College of Computer Science,
National Yang Ming Chiao Tung University

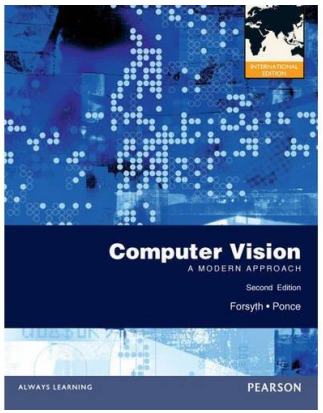
About the course

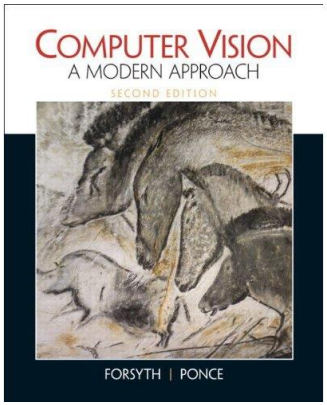
- Course title:
  - Computer Vision
- Lectures:
  - EC114, 10:10~12:00(Tues.) + asynchronous video
- Pre-requisites:
  - Computer programming skills in **Python or C/C++**.
  - Moderate levels in handing **data structures** and **algorithms**.
  - Essential Knowledge about **calculus** and **matrix computation**.
  - (optional) related courses: e.g. computer graphics, image processing, pattern recognition.

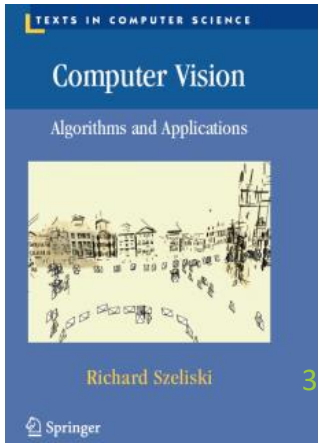
- Teacher:
  - I-Chen Lin (林奕成), Associate Professor
  - Email: ichenlin@cs.nctu.edu.tw
  - Office: EC 704 (工程三館)

About the course (cont.)

- TAs:
  - 傅信瑀 (karta2155802.cs05@nctu.edu.tw)
  - 涂仲謙 (a3102a123@gmail.com)
  - 林裕晟 (t871005y@gmail.com)
  - Office: EC229b
  - Phone ext: 56676
- Course web page:
  - Temp. Announce Page: https://reurl.cc/DdqNbE
  - E3 (Official): https://e3.nycu.edu.tw/course/view.php?id=35230
- Textbook
  - David A. Forsyth and Jean Ponce, Computer Vision: A Modern Approach, Prentice Hall, New Jersey. (1st or 2nd ed.)
- Reference book
  - Richard Szeliski, Computer Vision: Algorithms and Applications, Springer Verlag London, 2011.







About the course (cont.)

- References
  - IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI).
  - Intl. J. Computer Vision (IJCV).
  - Proc. Intl. Conf. Computer Vision (ICCV).
  - Proc. Intl. Conf. Computer Vision and Pattern Recognition (CVPR).
  - Proc. Euro. Conf. Computer Vision (ECCV).
  - IEEE Trans. Image Processing/Multimedia/CSVT/…

- Credits of several slides go to the reference lecture notes:
  - Prof. D.A. Forsyth, Computer Vision, UIUC.
  - Prof. T. Darrell, Computer Vision and Applications, MIT.
  - Prof. J. Rehg, Computer Vision, Georgia Inst. of Tech.
  - Prof. D. Lowe, Computer Vision, UBC, CA.
  - Prof. S. Seitz and P.Heckbert, Image-based modeling and rendering, CMU.

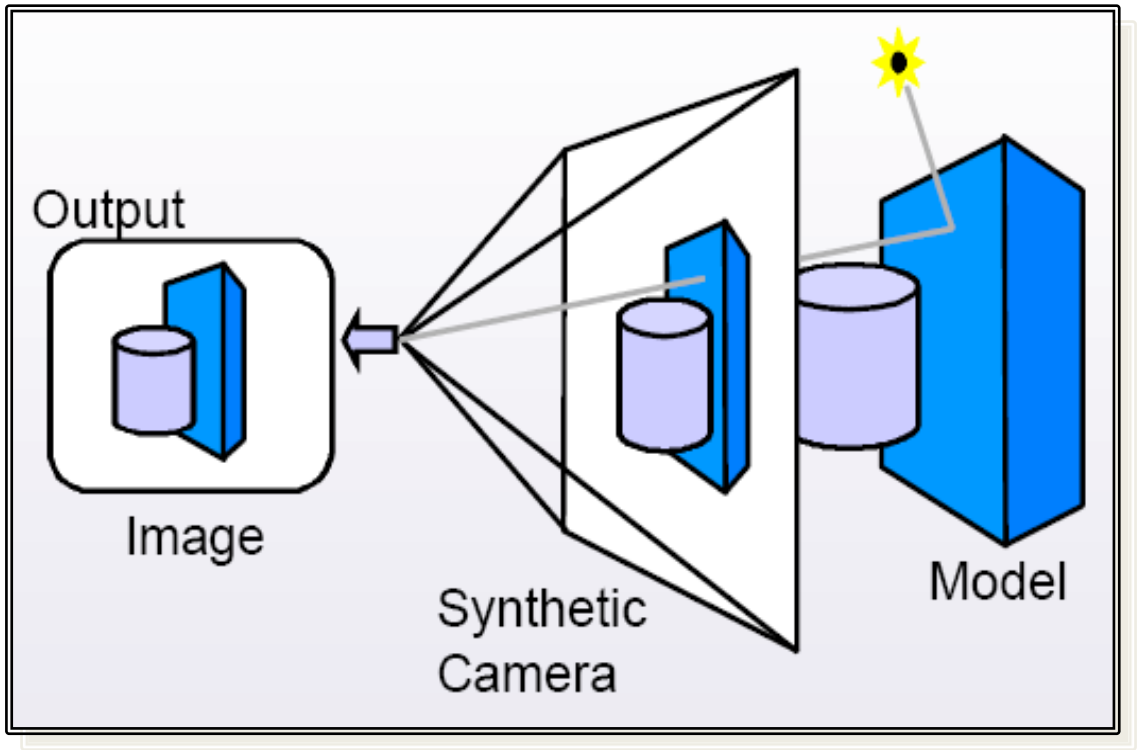
What's computer vision?

- The science of extracting information about the world from images.
  - "How to discover from images what is present in the world, where things are, what actions are taking place." (Marr 1982)
- One of the most challenging mysteries in Computer Science!
- Closely related fields:
  - Image processing
  - Artificial intelligence and machine Learning
  - Computer graphics
  - Human computer interaction
  - Augmented reality/virtual reality

Vision and related fields

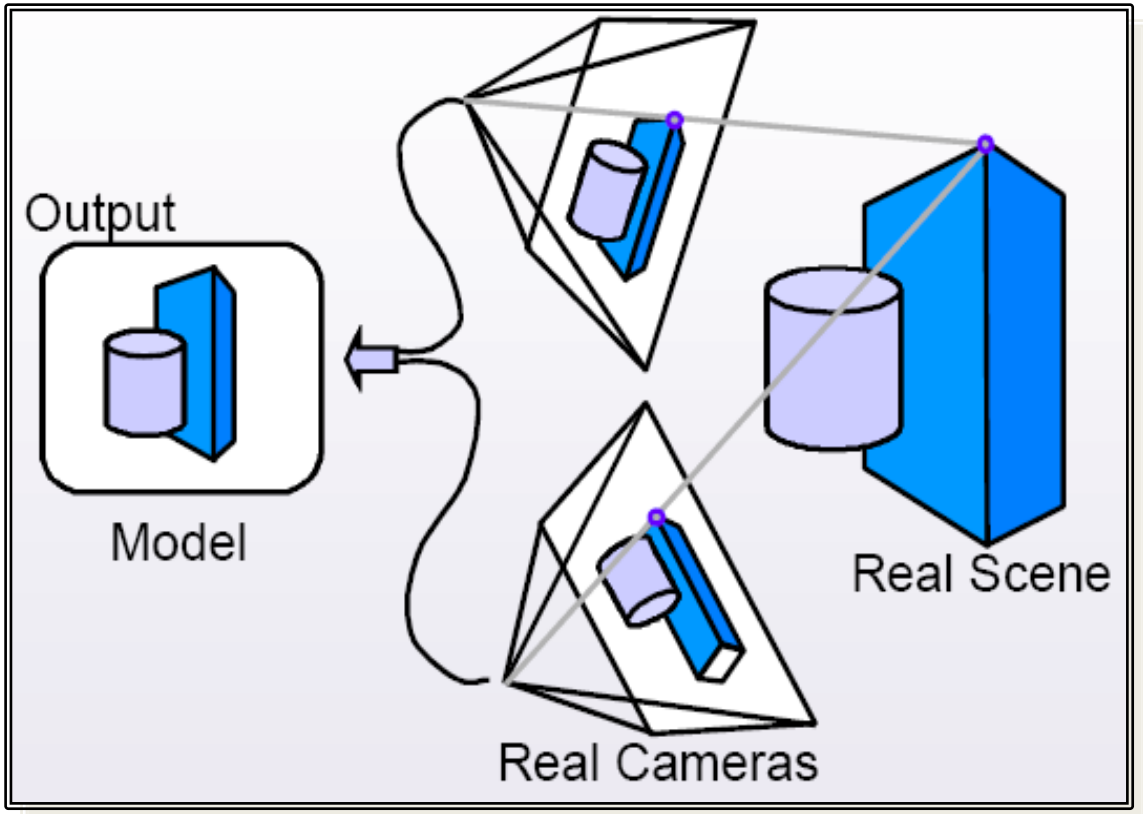
Input \ Outputs	descriptions	images
descriptions		Computer Graphics
images	Computer Vision & Pattern Recognition	Image Processing

Computer graphics



Figures from SIGGRAPH'99 Course Notes "IBMR"

Computer vision



Figures from SIGGRAPH'99 Course Notes "IBMR"

Knowing the scene

- Those common (or even trivial) abilities for humans can be quite difficult for computers.

Vision is fundamentally Ill-Posed

- There are an infinite number of possible scenes that could result in the pixels in a captured image.
  - Geometry
  - Photometry

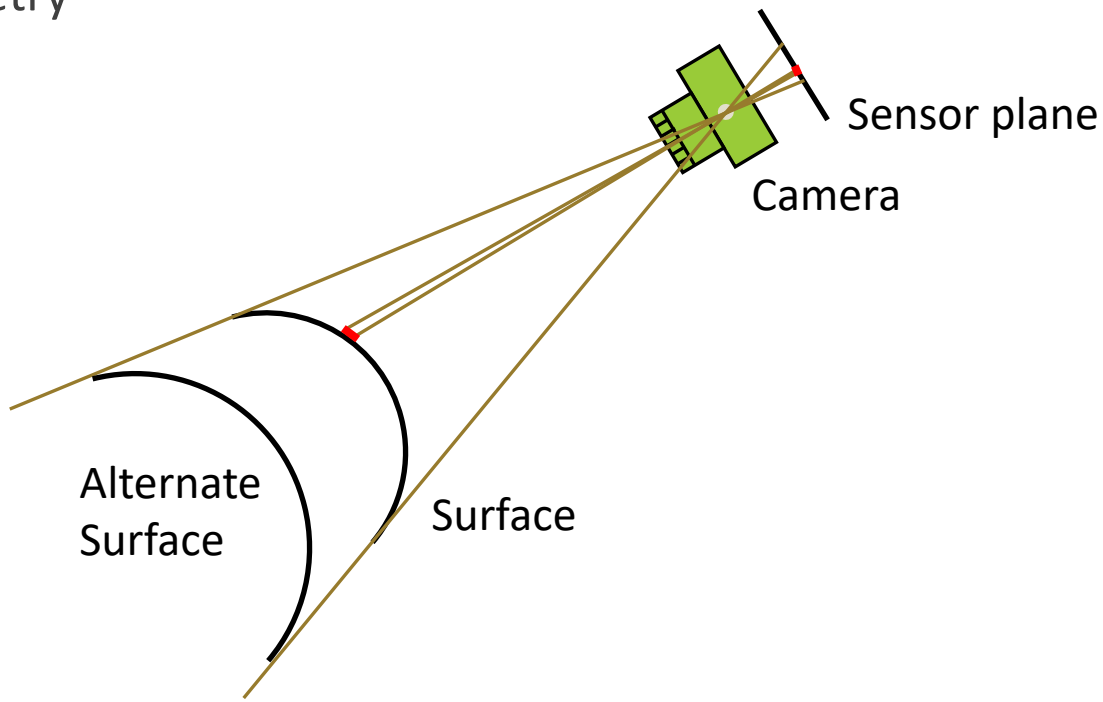
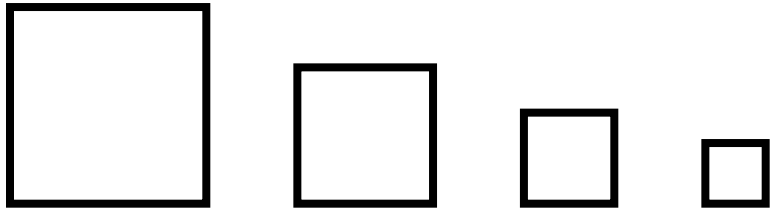


Figure from J. Rehg's lecture note: Computer Vision, Georgia Inst. Tech.

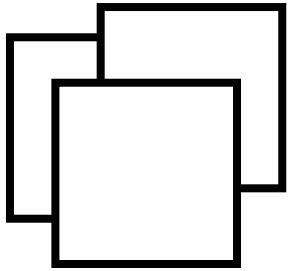
Monocular, static cues

Let's take the depth estimation as an example!

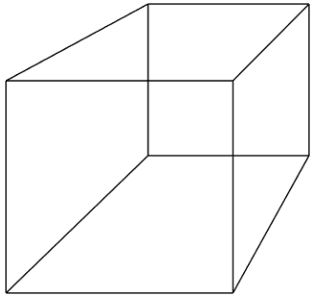
- Human perception makes use of **prior** knowledge about the shape and:
- Relative size



- Occlusion



- Perspective
  - Linear
  - Aerial

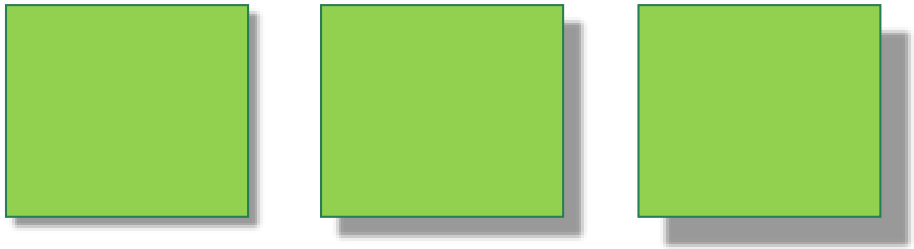


Monocular, static cues

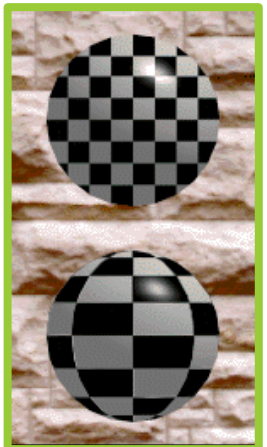
Taking depth estimation as an example!

- Lighting

- Shadow



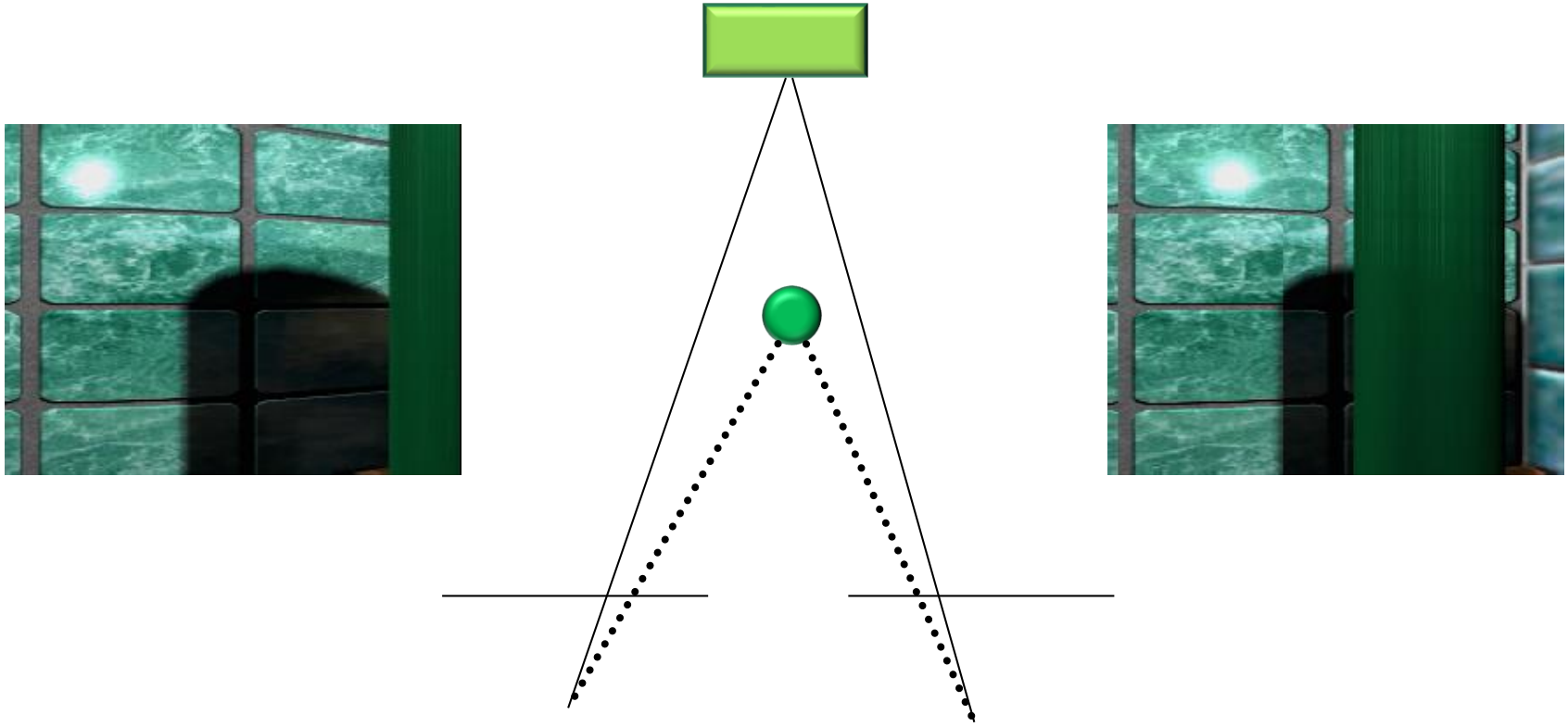
- Texture gradients



Monocular, dynamic cues

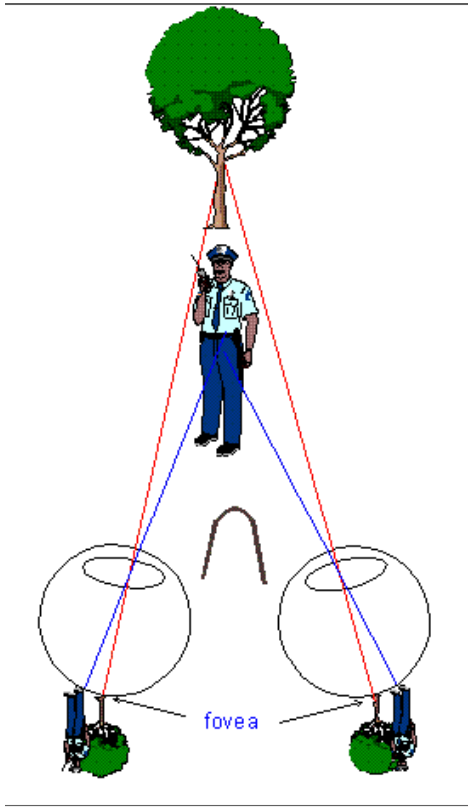
Taking depth estimation as an example!

- Motion parallax

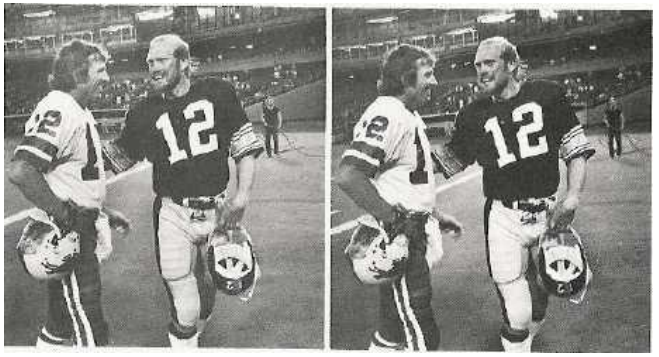


Binocular cues

Taking depth estimation as an example!



http://www.yorku.ca/eye/disparit.htm



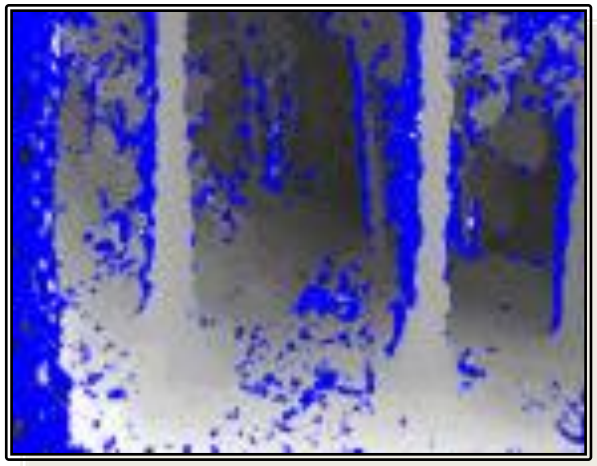
http://www.csus.edu/indiv/w/wickelgren/psyc110/Stereopsis.jpg

Stereo triangulation

Taking depth estimation as an example!

Left view

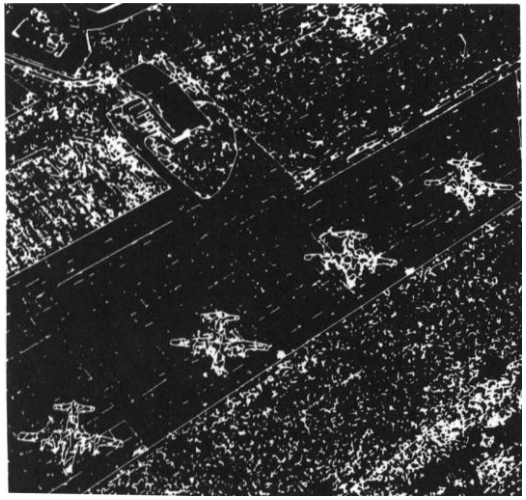
Right view

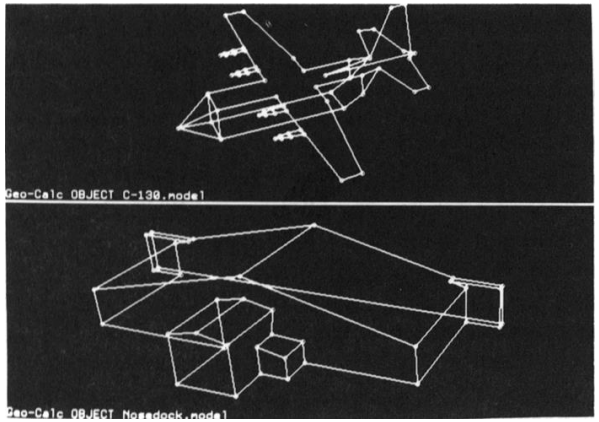


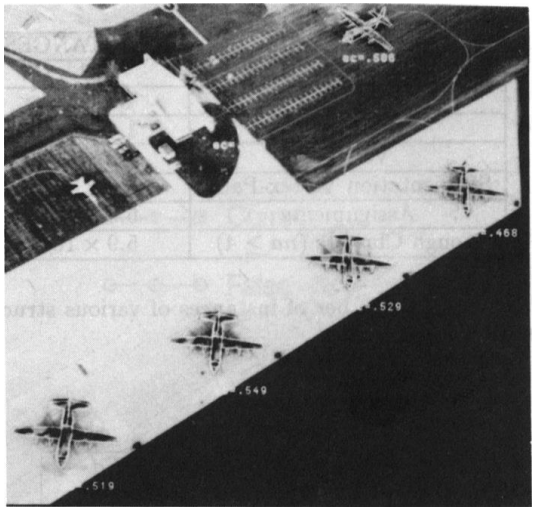
The estimated depth image (map)

Synthetic view with texture mapping

Evolution: early age







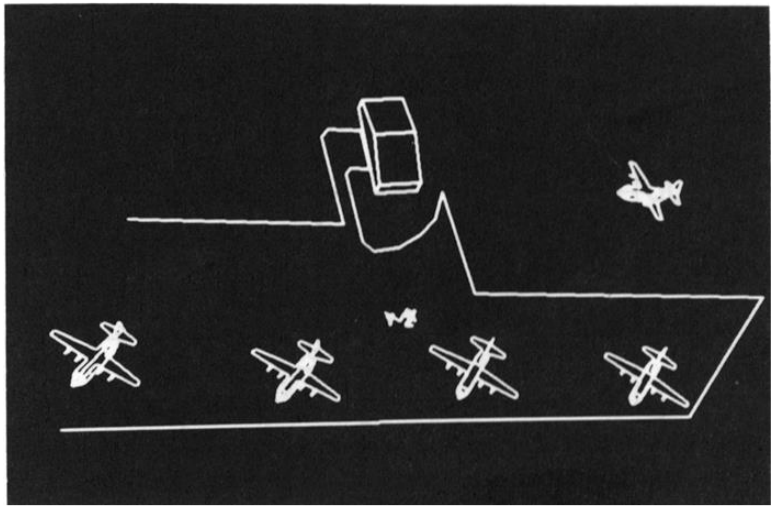
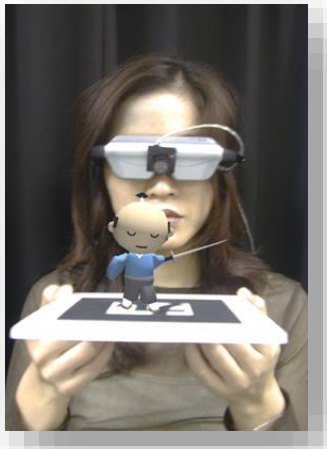
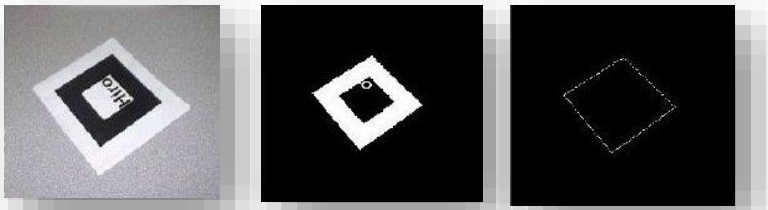


Fig. from D.A. Forsyth, Computer Vision course slides.

Evolution: specific applications with filters

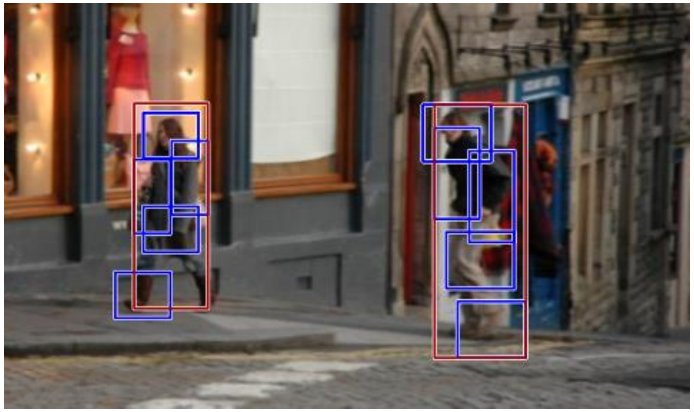




https://www.pyimagesearch.com/2020/09/21/opencv-automatic-license-number-plate-recognition-anpr-with-python/

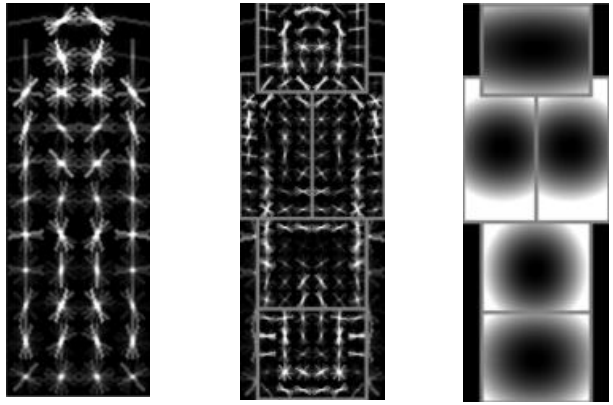
Kato and Billinghurst, ARToolKit

Evolution: robust features



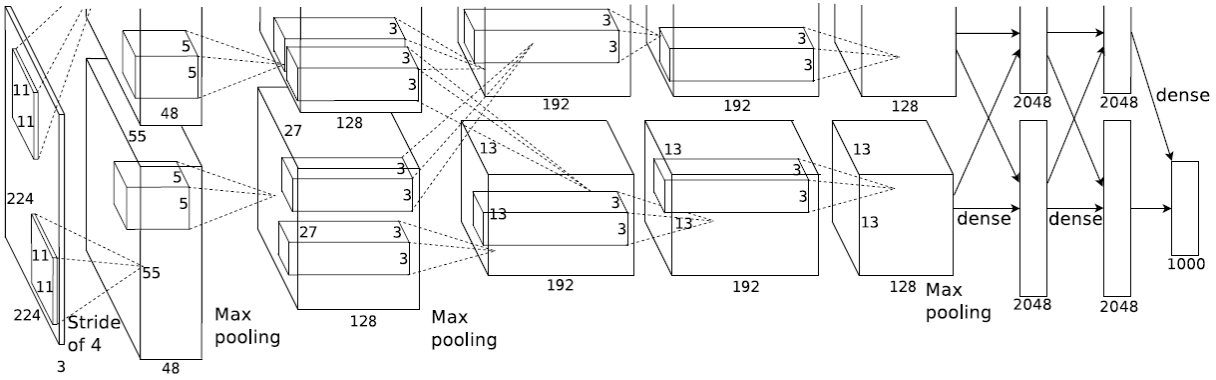
D. Lowe, Scale-invariant features (SIFT)

E.g. Large data applications, Snavely et al., “Photo Tourism”



Felzenszwalb et al., Deformable part model (DPM)

Evolution: deep learning



Krizhevsky et al., ImageNet Classification with Deep Convolutional Neural Networks, NeurIPS 2012.

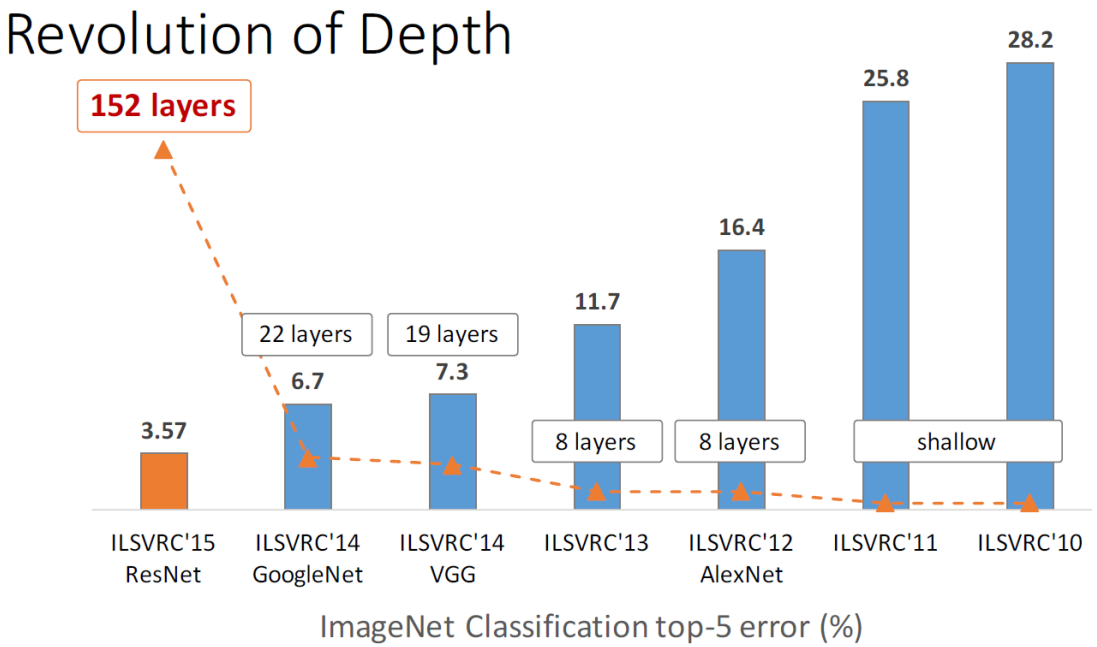
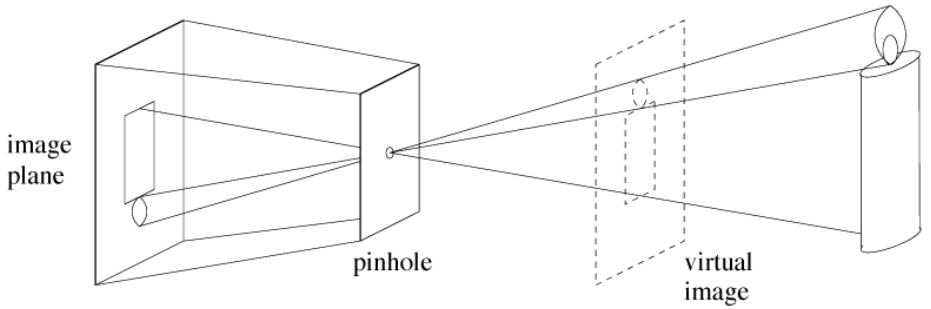


Fig. from Kaiming He, "Deep Residual Networks: Deep Learning Gets Way Deeper", ICML'16 tutorial

Syllabus

- Perspective, lens and camera



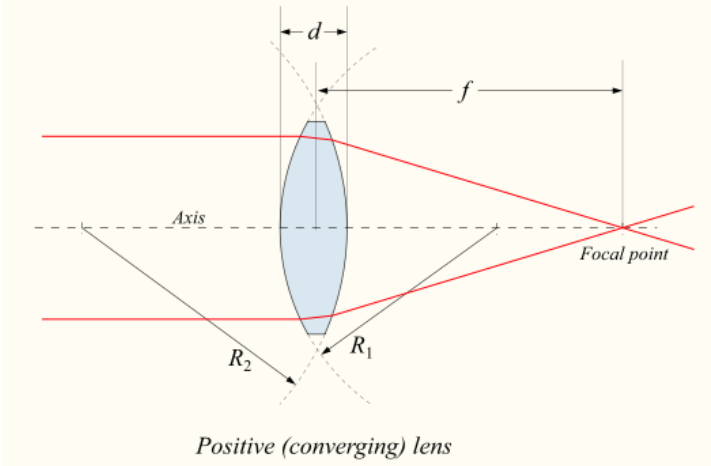
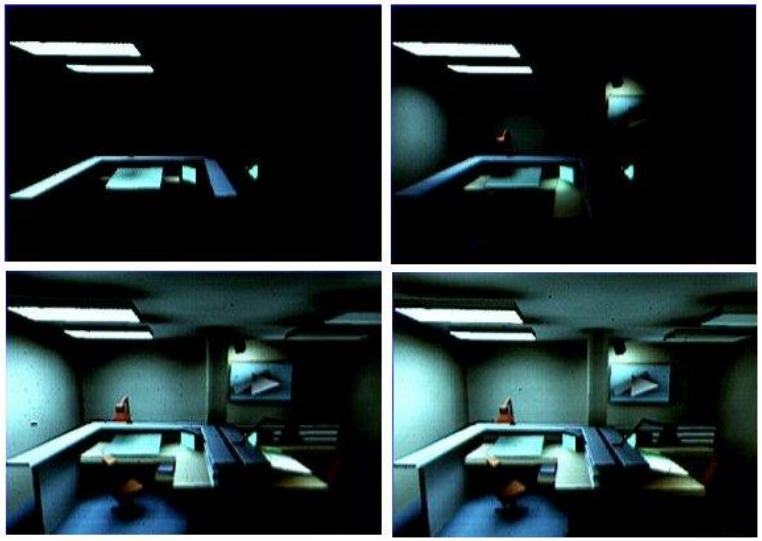


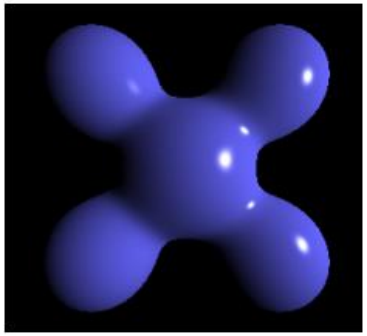
Figure from http://en.wikipedia.org/wiki/File:Lens1.svg

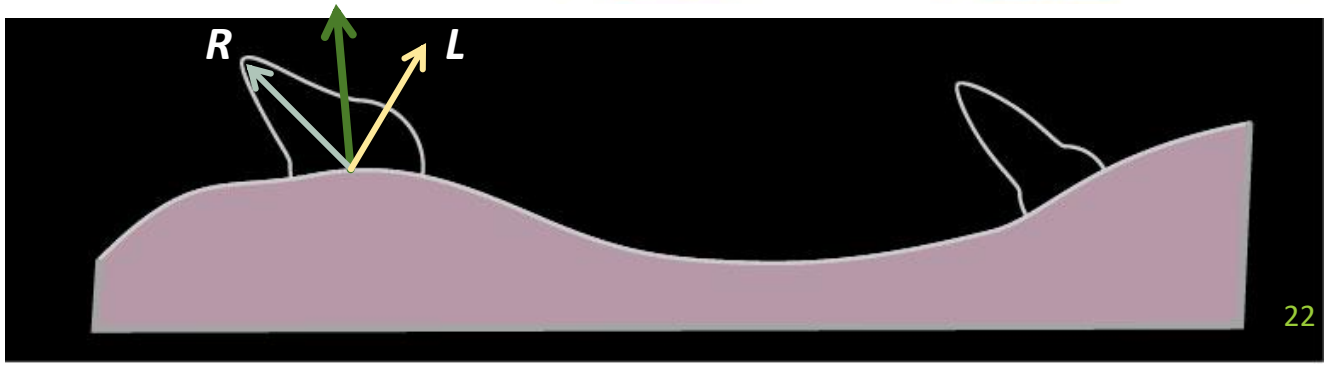
Syllabus (cont.)

- Radiometry
  - Illumination and reflectance models

Kettle, Mike Miller, POV-Ray



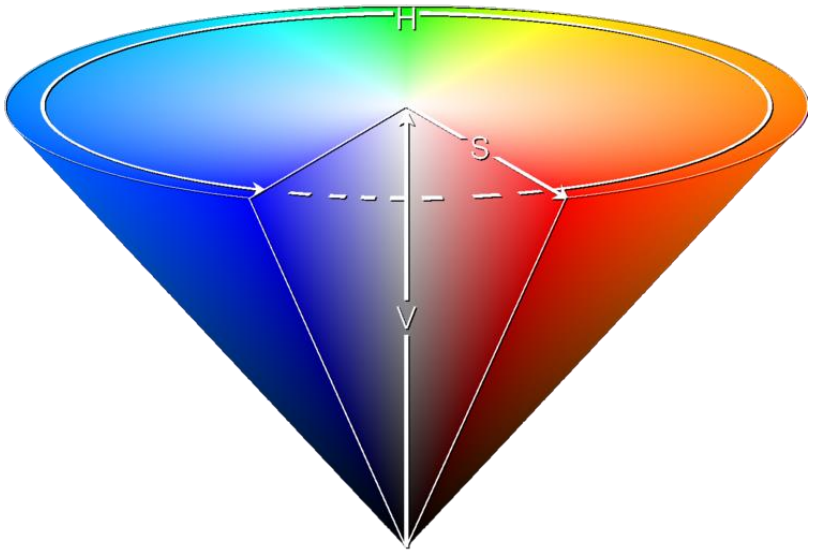




Syllabus (cont.)

- Color

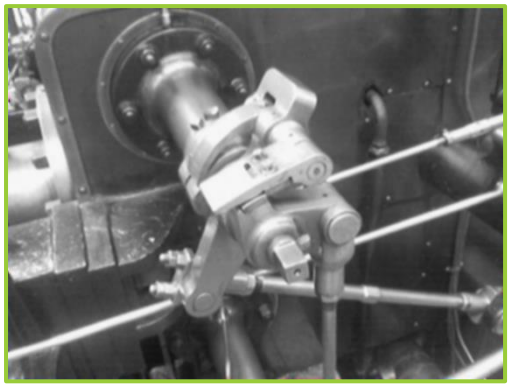
http://www.normankoren.com



http://en.wikipedia.org/wiki/HSV_color_space

Syllabus (cont.)

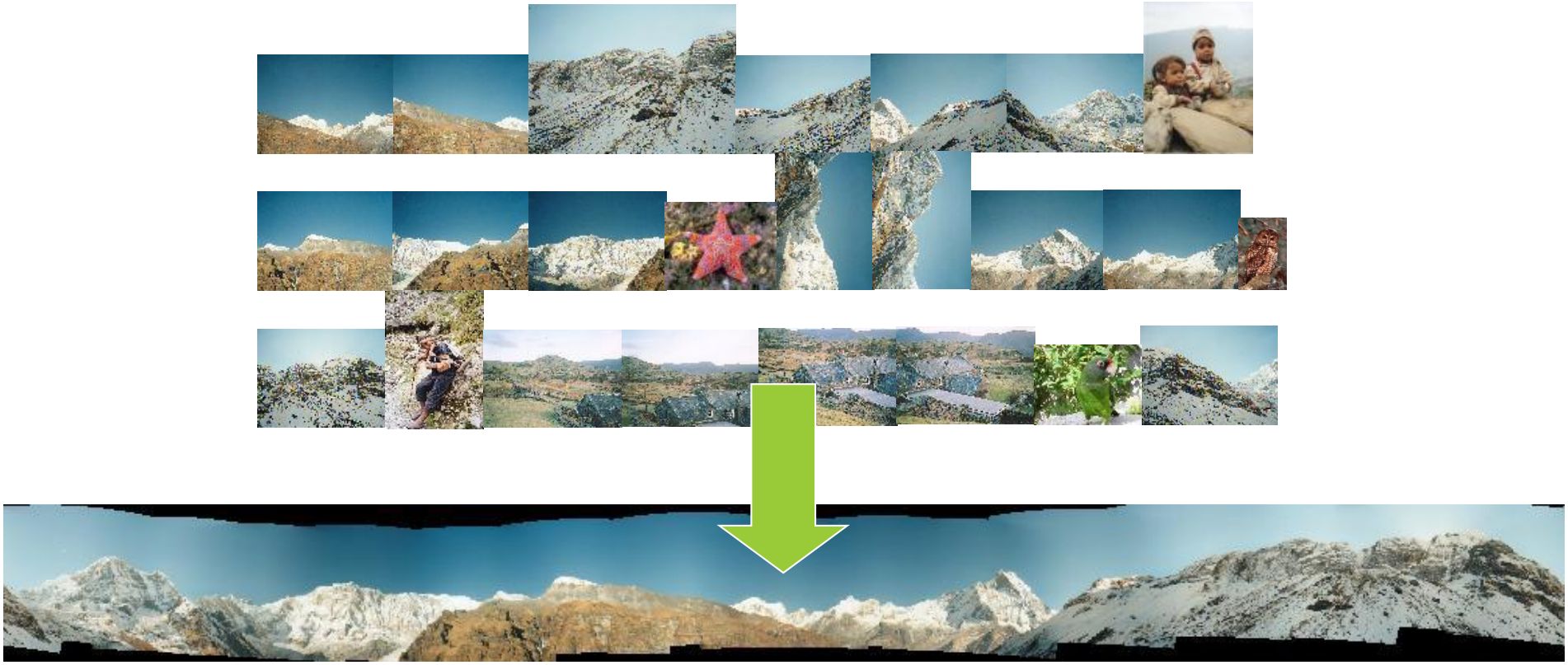
- Feature extraction: edge, corner, SIFT, etc.



•D. Frolova, D. Simakov, Slides of “Matching with Invariant Features”.

Syllabus (cont.)

- Image matching and panorama

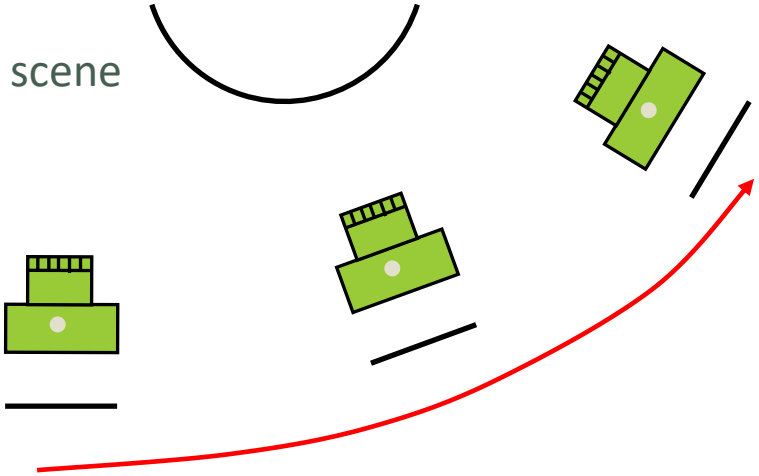


M. Brown and D.G. Lowe, Automatic Panoramic Image Stitching using Invariant Features, IJCV 2007

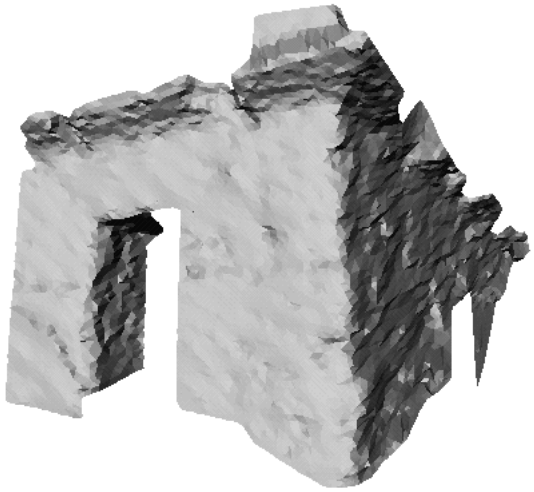
Syllabus (cont.)

- Structure from motion

Static scene



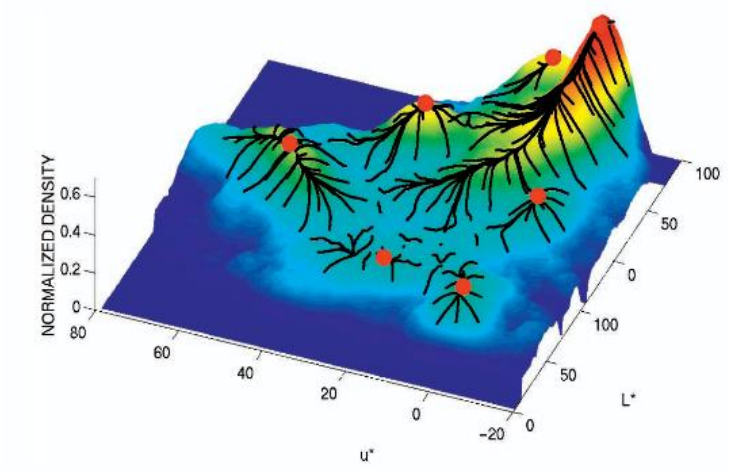
Shaded model



3D model reconstructed by Luc Van Gool et al.

Syllabus (cont.)

- Clustering and segmentation
  (dependent on the schedule)



Mean-shift clustering

(a) A woman from a village

(b) A church in Mozhaisk (near Moscow)

Graph-cut segmentation

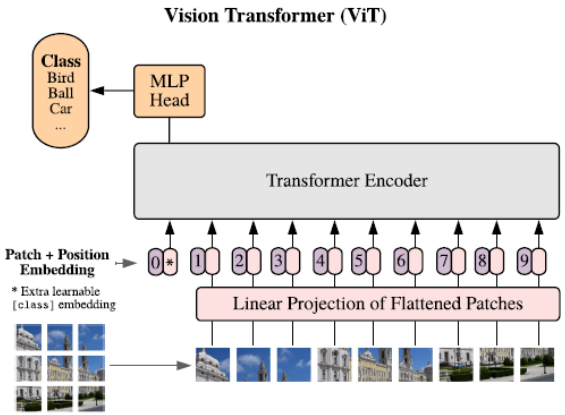
Syllabus (cont.)

- Advanced topics (dependent on the schedule)

E.g. learning-based methods

Shotton et al., Real-Time Human Pose Recognition in Parts from Single Depth Images, CVPR'11.

He et al., Mask R-CNN, ICCV'17.

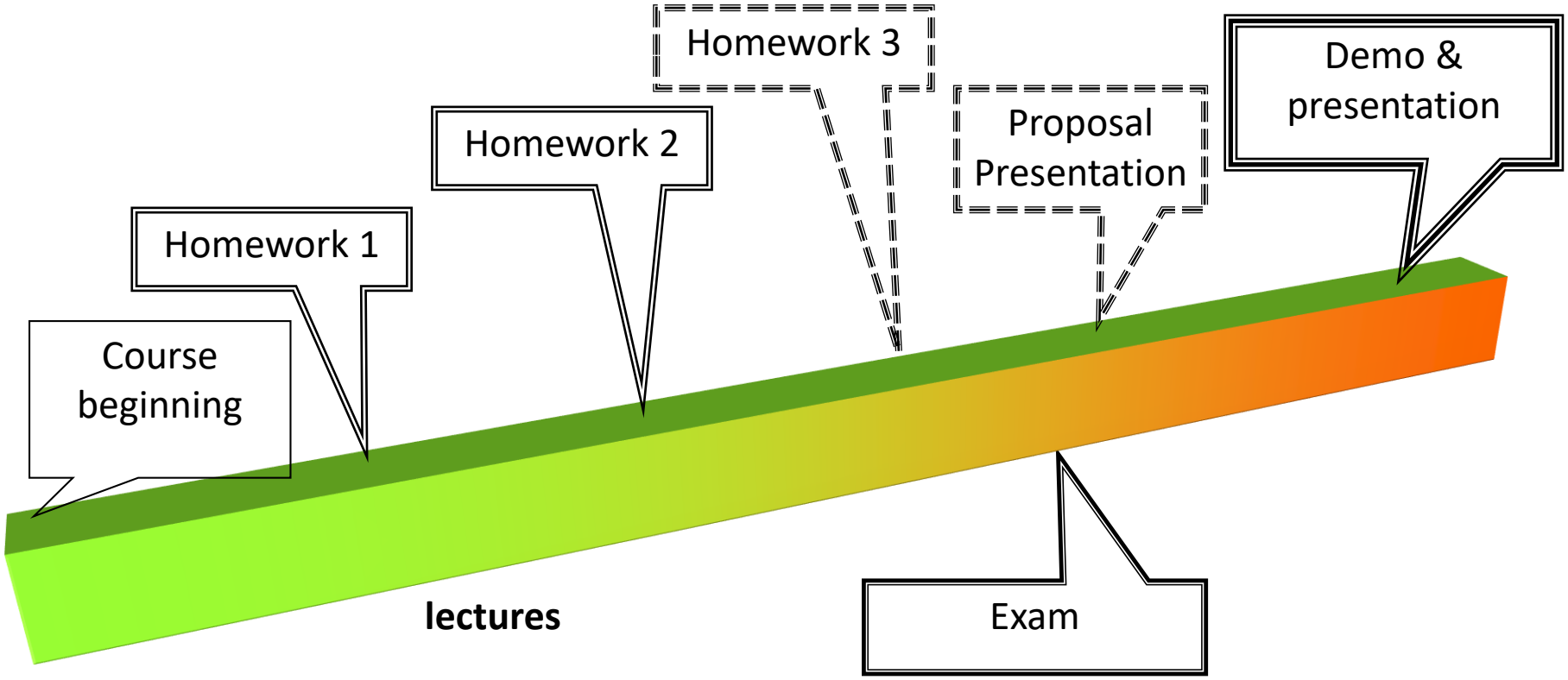


Dosovitskiy et al., Vision Transformer, ICLR'21

About the course (cont.)

- Grades: (temporary)
  - Exams (15~40%)
  - Homework (30~60%)
    - Photometric stereo
    - keypoint-related, e.g. photo alignment
    - …….
  - Term project (30~50%)
    - 1~3 members per group
    - Paper and proposal presentation (if possible)
    - Demo & final presentation.
  - Class participation (0~10% or bonus)

The schedule



Announcement

- Our classes for **at least the first two weeks** will be taught **online**.
- We only have limited quotas to add the course.
  - Application to add this course. https://reurl.cc/k78VpG
  - **CS Ph.D. students for qualifying**, **students of the CS college** and **enrolled students belonging to specific programs** will have higher priority to add the course.
- Slides and project announcements will be in **English**, but students are welcome to discuss or ask questions in Chinese.