Computer Vision

Term Project

I-Chen Lin

College of Computer Science

National Yang Ming Chiao Tung University

Expected new schedules

- W12: **5/03** lecture: (Ch9) Two views(II) + (Ch10) Rect. and depth
- W13: **5/10** lecture: (Ch11) Clustering and compact representation
  - No midterm exam
  - Register your team for the term project
- W14: **5/17** lecture: (Ch12) Selected topics about learning-based methods
- W15~W17: no class
  - Before 5/28 upload video presentation
- W18: **6/14** morning and afternoon, online final demo
  - 6/13 afternoon, online early final demo (with 1.5pt bonus)

Term project schedule

- **Before May 10**: Submit your team member list , the (temporary) project topic. (website will be announced later)
  - (Recommended) 2~3 members per team.
  - 1-person teams: we will ask whether you are willing to team up with others (e.g. with other 1-person teams of similar topics)
  - 4-person teams: you have to address the scope of your project.
- **Before May 28:**
  - Upload video presentation about your proposal and research survey
    - List your *project title*, *member info.*
    - Talk about your ***goal/problem***, ***survey of related work***, and your ***expected method/framework***.
    - Allocate more time to the survey, and focus on one or two important related methods.
  - Students will have to browse a number of videos and provide commends or ask questions.

Term project schedule (cont.)

- **June 13:** afternoon, online early demo (with 1.5pt bonus)
- **June 14** : morning and afternoon, online final demo
  - Presentation time per group: (temporarily, 10 minutes per team)
  - Briefly introduce your *goal*/*problem*, *method*, and *difficulty/uniqueness*
  - Emphasize on *results*, *demo* and/or *comparison*.

- Upload Report (no more than 6 pages) and codes (with comments)
  - Describe the goal/problem, method, difficulty/uniqueness, results and/or comparison.
  - List libraries and open sources used in your project.
  - List the contribution (e.g. work items) of each member?

Topics (references)

- Renowned computer vision journal and conferences:
  - Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR).
  - Proc. Intl. Conf. Computer Vision (ICCV).
  - European Conf. Computer Vision (ECCV).
  - IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI).
  - Intl. J. Computer Vision (IJCV).

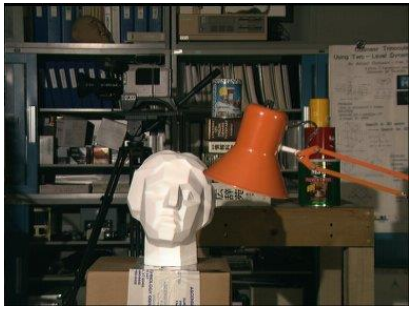
Topics (references)

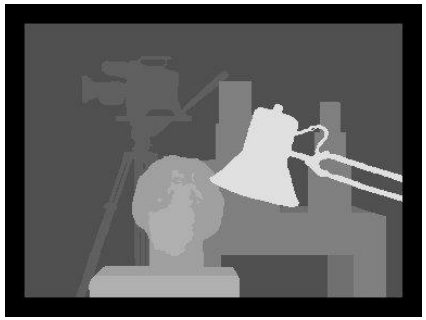
- Full papers of related renowned conferences:
  (Usually more than 7 pages)
  - Proc. Neural Information Processing Systems (NeurIPS)
  - Proc. Intl. Conf. Machine Learning (ICML)
  - Proc. ACM SIGGRAPH, SIGGRAPH Asia.
  - Proc. ACM Multimedia
  - ……

- Related renowned ACM or IEEE journals,
  e.g. TIP, TMM, TOG, TVCG, TCSVT, etc.

Topics (Extension of course slides)

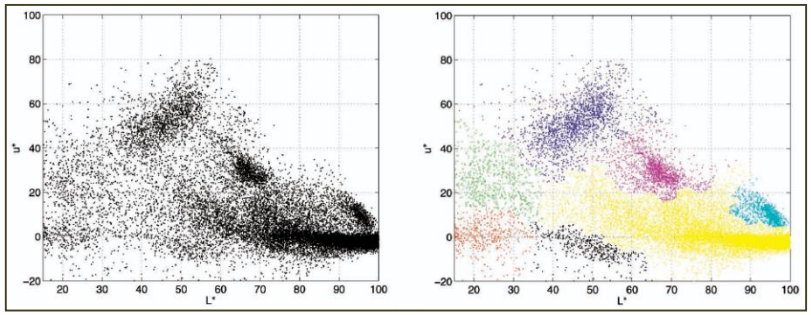
- *Depth map from stereo images*.
- Data
  - Acquired by yourselves (calibration and rectification)
  - Middlebury http://vision.middlebury.edu/stereo/data/
  - Other indoor scene datasets
- *Methods: (you have to implement parts of the method by yourselves)*
  - Block matching
  - Dynamic programming (DP: class slides)
  - Optimization methods
    - V. Kolmogorov et al., Computing Visual Correspondence with Occlusions via Graph Cuts, Proc. ICCV'01.
    - J. Sun et al, Stereo Matching Using Belief Propagation, IEEE T. PAMI, 2003.
  - Incorporating SIFT, DNN features.

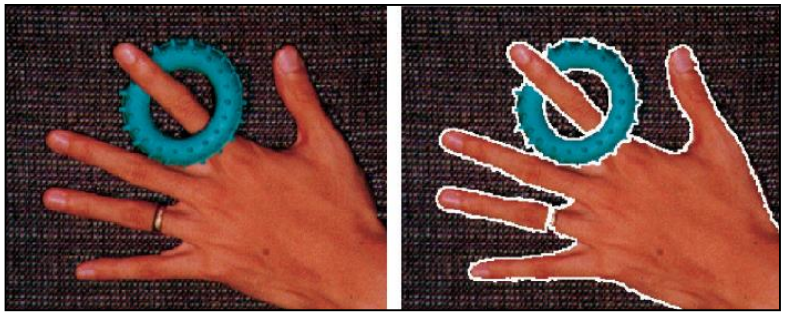




Topics (Extension of course slides)

- *Application with clustering methods.*
  - D. Comaniciu and P. Meer, “Mean Shift: A Robust Approach toward Feature Space Analysis”, IEEE T. PAMI, 2002.





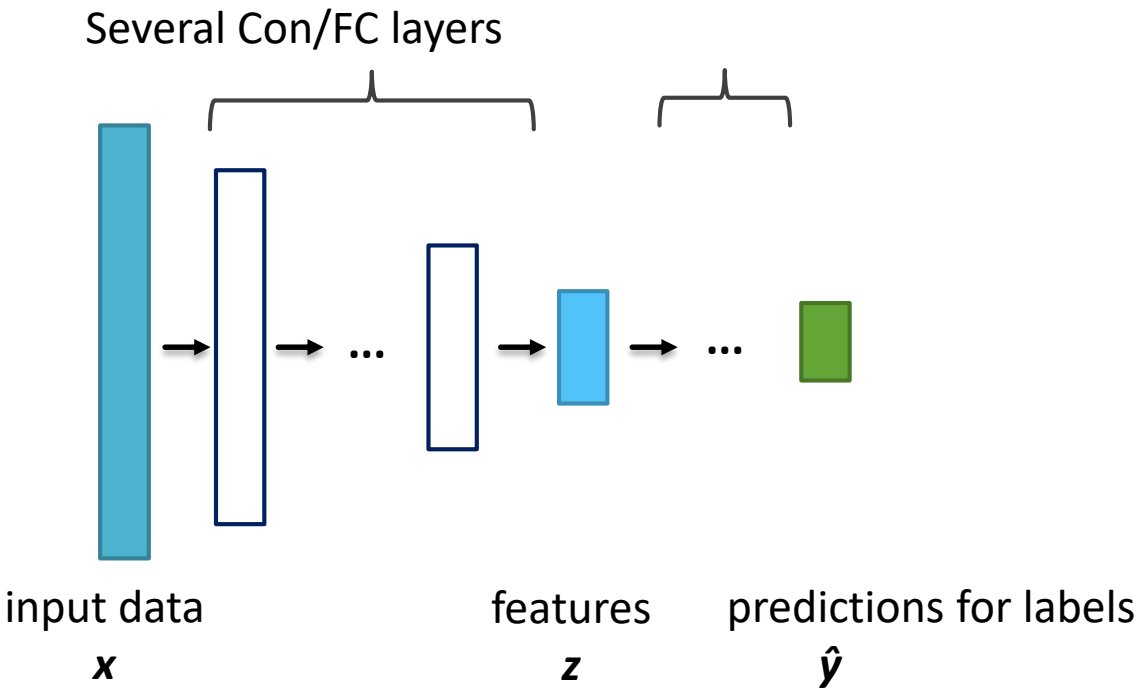
  - D. Comaniciu, V. Ramesh, P. Meer, “Kernel-Based Object Tracking,” IEEE Trans. Pattern Analysis Machine 2003.

Topics (Extension of course slides)

- *Special effects with depth and clusters.*

  (e.g. UFocus of HTC, Google camera)

Topics (Incorporating modern features)



E.g. Simonyan and Zisserman, Very Deep Convolutional Networks for Large-scale Image Recognition, Proc. ICLR'15.

VGG19
input
conv3-64
conv3-64
maxpool
conv3-128
conv3-128
maxpool
conv3-256
conv3-256
conv3-256
conv3-256
maxpool
conv3-512
conv3-512
conv3-512
conv3-512
maxpool
conv3-512
conv3-512
conv3-512
conv3-512
maxpool
FC-4096
FC-4096
FC-1000
soft-max

Topics (Incorporating modern features (cont.))

- DNN features at higher layers contain some sematic information.

reference

source 1

source 2

Nearest neighbor field (NNF)

*ReluL*_1, *L* = 5, . . . , 1 in VGG19

Fig. from He et al., “Progressive Color Transfer With Dense Semantic Correspondences”, ACM Trans. Graph. 2019.

Topics (Incorporating modern features (cont.))

- Incorporating DNN features for object detection or tracking.
  - Comparison with SIFT or other classic features.

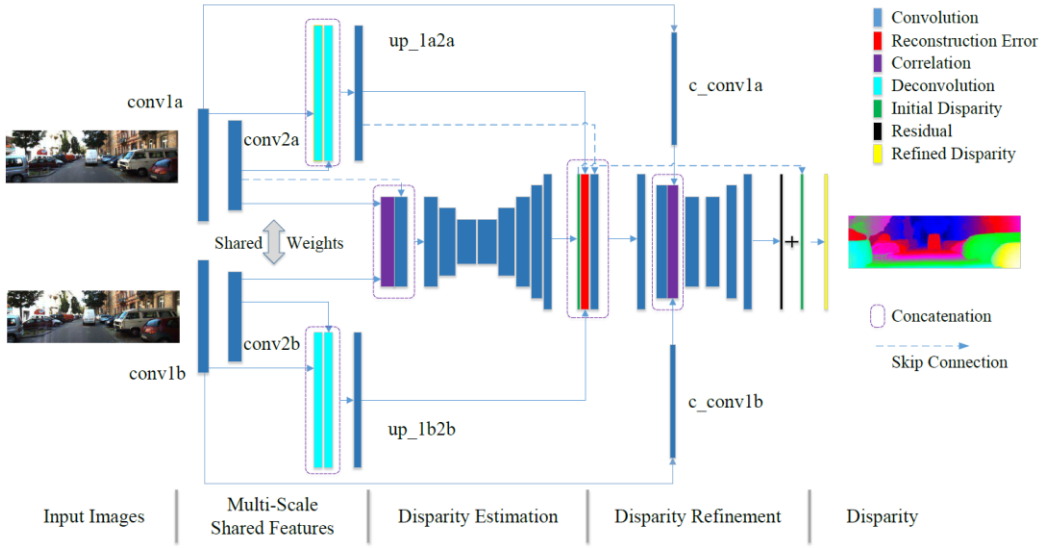
E.g.

- D. Lowe, “Distinctive Image Features from Scale-invariant Keypoints,” Intl. J. Computer Vision (IJCV), 60(2):91-110, 2004.

- Huiyu Zhou, et al., “Object tracking using SIFT features and mean shift”, Computer Vision and Image Understanding (CVIU) 2009.

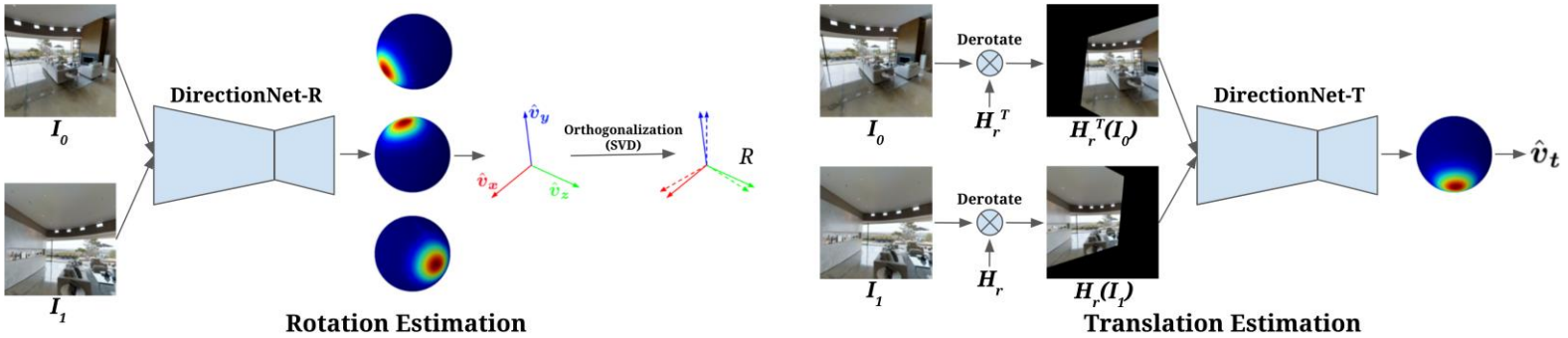
Topics (Modern approaches for classic problems)

- E.g. Disparity estimation



Liang et al, Learning for Disparity Estimation through Feature Constancy, CVPR'18.

- Camera pose estimation



Chen et al., Wide-Baseline Relative Camera Pose Estimation with Directional Learning, CVPR'21.

Topics (Your expertise)

- You are encouraged to work on CV-related topics that you or your lab specializes in.
- Your project can also be developed based on some work from your lab.
- However, please explicitly describe your contribution in this CV term project.

About usage of public codes/libraries

- You can use public codes or libraries, but you have to clearly point out your sources during your presentation and in your report.

- It is now allowed to hand over a project in which only one or two pretrained models are directly applied.

- You have to do at least one of the followings:
  - Enhance/modify certain parts of the codes
  - Integrate one or multiple public codes into your system for a new goal.
  - Compare and discuss the pros and cons of multiple methods