Computer Vision

Term Project

I-Chen Lin

College of Computer Science

National Yang Ming Chiao Tung University

Expected schedules

- W12: **5/7 lecture + video** (Ch8 camera calib.; Ch9 epipolar geom.)
  - TAs will announce how to **find teammates** and **form groups** on E3.
- W13: **5/14 lecture + video** (Ch9; Ch10 Rect. Depth)
  - Finalize the **term-project groups** (TBA).
- W14: **5/21 lecture (or + video)**: (Ch12) Selected topics about learning-based methods
  - **Submit** your **proposal video** (TBA)
- W15: no class
  - **Browse** a limited number of **proposal videos** and **QA**
- W16: **6/4 Early-bird final demo**
- W17: **6/11 Regular final demo**

Term project schedule

- **Before May 14 (2pm)**: Submit your team member list , the (temporary) project topic. (website will be announced later)
  - 1-person teams: OK, and your are encouraged to find teammates on the forum on E3. (proposal: 10~20min)
  - 2- or 3-person teams: recommended. (proposal: 15~25min)
  - 4-person teams: OK, and you are encouraged to do something more extensive or deeper. (proposal: 20~30min)
- **Before May 25:**
  - 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.
- **Before June 2:**
  - Students have to browse a number of videos and provide commends or ask questions.

Term project schedule (cont.)

- **June 4:** online early final demo (with a bonus)
- **June 15** : regular final demo
  - Presentation time per group: (temporarily, less than 10 minutes)
  - Briefly introduce your *goal/problem*, *related work*, *system/method*
  - Emphasize on *uniqueness or solutions to difficulties*, *results/demo*, and/or *comparison*.

- **Before June 18**:
  - 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?
    - *If there is a confidential problem, please discuss with Teacher in advance*.

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).

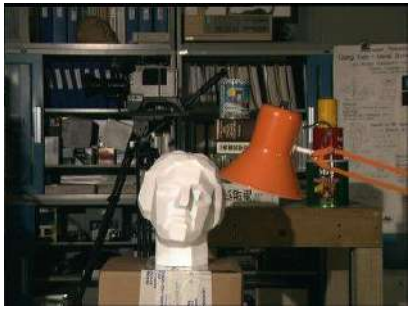
Topics (references)

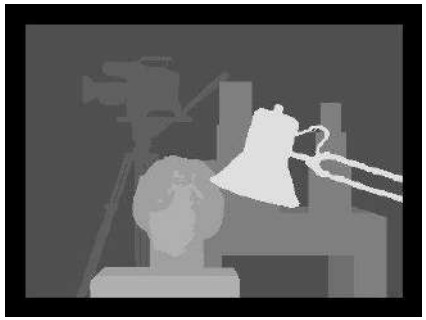
- Full papers of related renowned conferences:
  (Usually more than 7 pages)
  - Proc. Neural Information Processing Systems (NeurIPS)
  - Proc. ACM Multimedia
  - Proc. ACM SIGGRAPH, SIGGRAPH Asia.
  - ……

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

Topics (Extension of course slides)

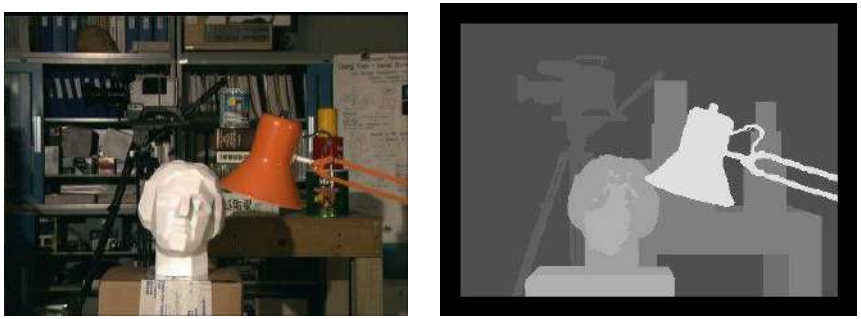
- *Depth map from stereo images*.
- Data
  - Acquired by yourselves (calibration and rectification)
  - Middlebury http://vision.middlebury.edu/stereo/data/
  - Other indoor scene datasets
  - Prof. Chuang's calibration methods: https://github.com/jessie0915/Geometry-based-Camera-Calibration-Using-Closed-form-Solution-of-Principal-Line
- *Classic 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….



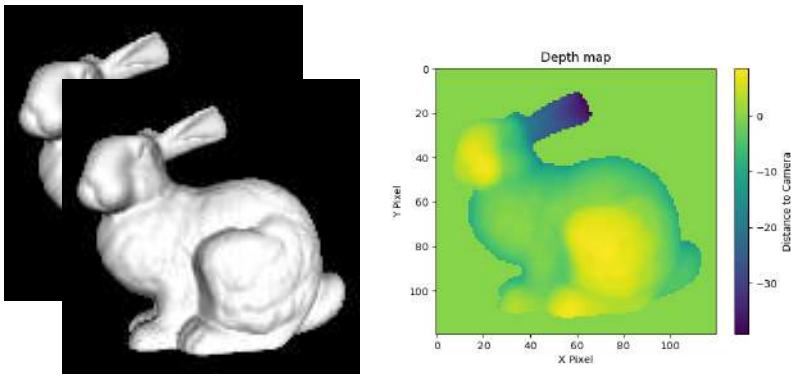


Topics (Extension of course slides)

- *Depth map from stereo images + refinement with photometric stereo*



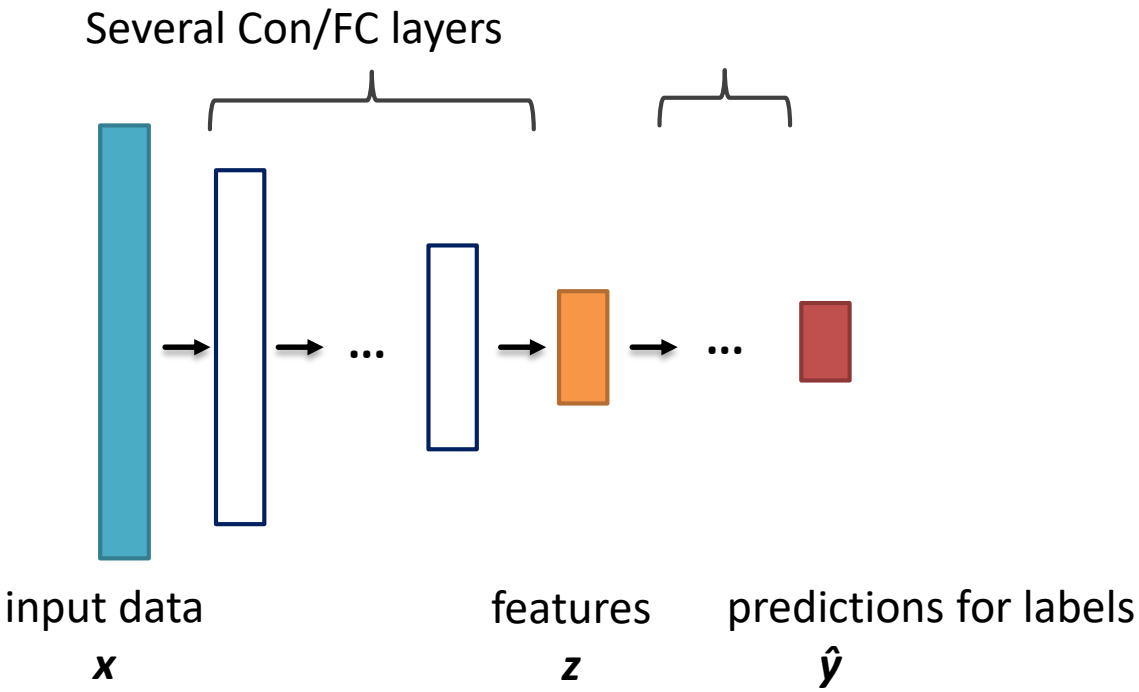
Structure from motion



Photometric stereo

- How to take their advantages and refine the surface details ?
  - Blending the depth?
  - Conditional Filtering?
  - Optimization with certain constraints?

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.))

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

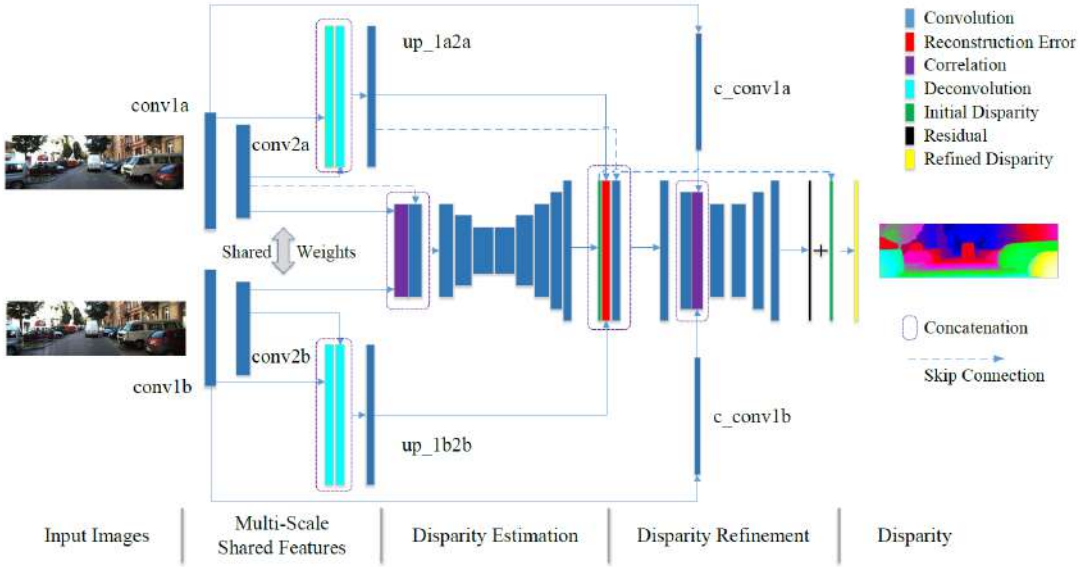
Classic examples:

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

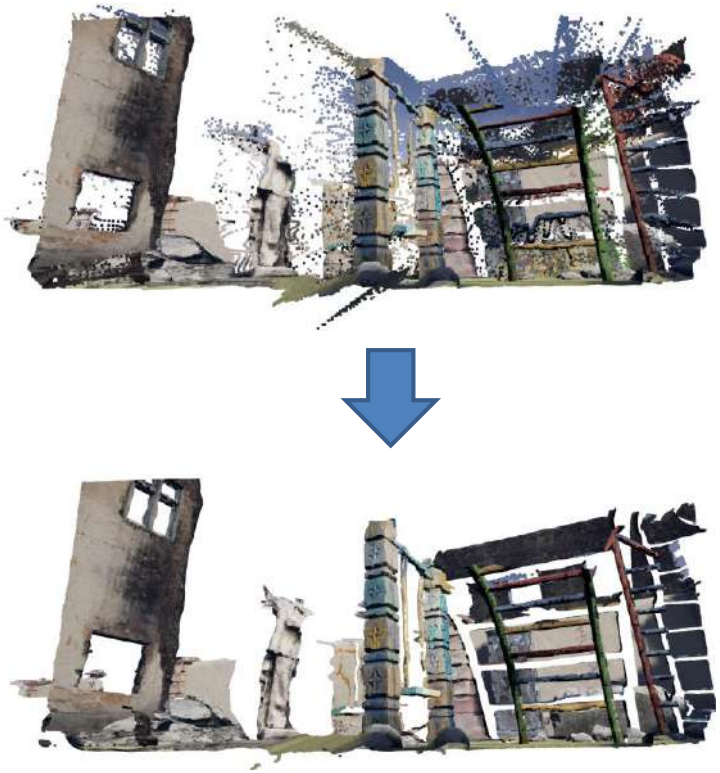
- Huiyu Zhou, et al., “Object tracking using SIFT features and mean shift”, Computer Vision and Image Understanding (CVIU) 2009.

Topics (Modern approaches for classic problems)

- E.g. Disparity or depth estimation



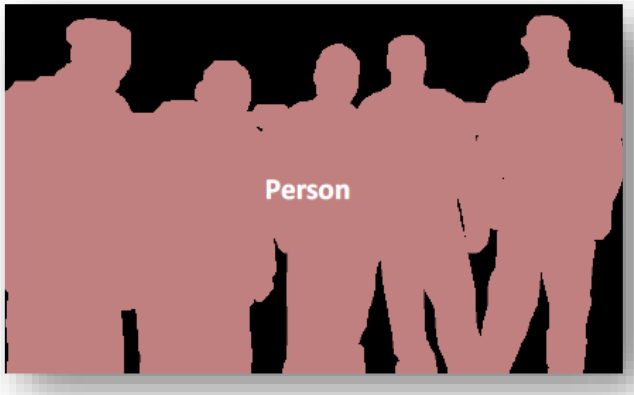
Liang et al, Learning for Disparity Estimation through Feature Constancy, CVPR'18.



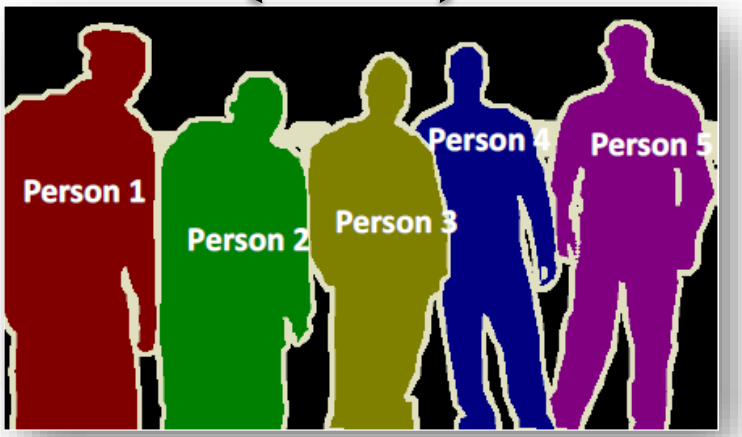
Tosi et al., SMD-Nets: Stereo Mixture Density Networks, CVPR'21

Topics (Modern approaches)

Object detection



Semantic segmentation



Instance segmentation

E.g. He et al., Mask RCNN, ICCV 2017.

Wang et al., SOLO and v2, ECCV2020, NeurIPS 2020.

Liu et al., Swin Transformer and v2, ICCV21, CVPR22.

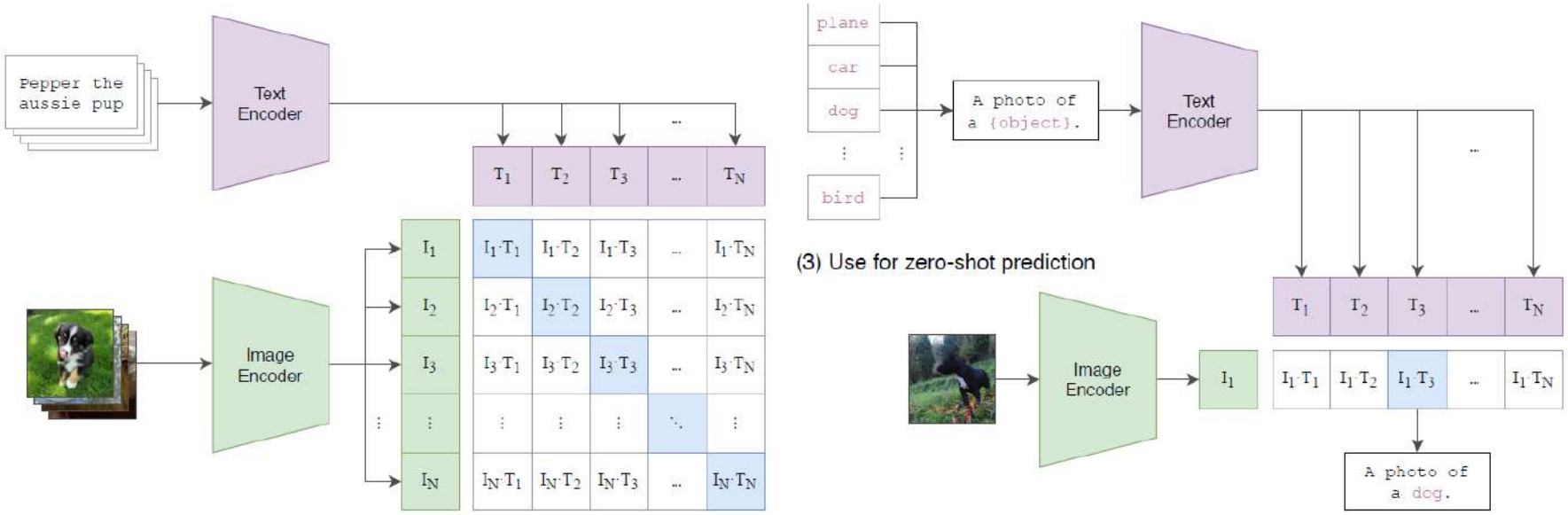
Topics (Applications with extracted maps)

- Special effects with specific maps extracted from images or videos.
  - E.g. depth maps, semantic maps, saliency maps, and so forth.

Examples of Ufocus, HTC application

Topics (Work with vision-language pretraining)

CLIP



E.g.

Radford et al., Learning Transferable Visual Models From Natural Language Supervision, ICML'21

Li et al., BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation, ICML'22

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