CSE 291: Advances in 3D Vision
Manmohan Chandraker

Lecture 0: Introduction
Computer vision in living rooms

Microsoft Kinect Xbox

Sportvision first down line
Vision to explore the world
Vision to explore other worlds

- Panorama stitching
- Stereo imaging
- Navigation
- ....
Vision to explore all worlds

Including virtual ones!

*The Matrix* movies, ESC Entertainment, XYZRGB, NRC
Broad classes of vision applications

- Sense
- Understand
- Interface

- Reconstruct
- Recognize
- Reorganize
Broad classes of vision applications

- Sense
- Understand
- Interface

Scenes

People

Overcrowding
Mass stagnation

Percentage indicates the density of people in the area

Alerts indicate people flow

11%
Broad classes of vision applications

- Sense
- Understand
- Interface

- Human-Human
- Human-Machine
- Machine-Machine
Some applications span many classes

- Sense
- Understand
- Interface

Advanced Driver Assistance Systems
Deep learning is revolutionizing AI

- Tic-tac-toe (1952)
- Checkers (1994)
- Chess (1997)
- Atari (2015)
- Go (2016)
Computer vision is also riding the wave

- Autonomous driving (Google, Tesla, Mobileye, ....)
- Augmented reality (HoloLens, Oculus, MagicLeap)
- Social networks (Google, Facebook, ....)
- Mobile applications
- Surveillance
3D Computer Vision
Image: 2D projection of 3D world

What have we lost?

- Angles
- Distances (lengths)
3D Reconstruction

- Relate geometric information across multiple images
- Relate photometric information across multiple images
- Relate semantic information across a dataset of images

(figure courtesy of Michael Cohen)
Some 3D sensors do exist
Images are everywhere around us

Source: Domo
Autonomous navigation

Source: Wired
Photo-tourism

Source: Snavely et al.
Reconstructing building interiors

Source: Xiao and Furukawa
Games
Movies

_Avatar_ movie, Zoe Saldana emotes Neyitri (Fox Movie Channel)
A Few Challenges in 3D Vision
Challenge: High accuracy desired

Reconstruction on LFW dataset

Recognition on LFW dataset

State-of-art CNN: 95.9%
Humans: 97.5%
Challenge: Limited views

- Want to relate information across multiple images
- With limited views
  - Have to “guess” missing coverage
  - Correspondence across wide baseline
  - Deal with appearance changes
Challenge: Lack of texture
Challenge: Complex appearance
Challenge: Low resolution
Challenge: Non-rigidity
Challenge: Complex deformations
New Approaches to Tackle Challenges
Recognition has progressed rapidly

Error rates on ImageNet Visual Recognition Challenge, %

Sources: ImageNet; Stanford Vision Lab
Deep learning has led to large gains

- Hierarchical and expressive feature representations
- Trained end-to-end, rather than hand-crafted for each task
- Remarkable in transferring knowledge across tasks
Deep learning has opened new areas

- Availability of large-scale image and video data
- Availability of computational power
  - Better and cheaper GPUs
  - Cloud computing resources
- Better understanding of how to train deep neural networks
- These advantages are available for 3D reconstruction too!
  - Data-driven priors for 3D reconstruction
  - End-to-end training rather than expensive feature design.
3D vision also exploits deep learning

Zhu et al., CVPR 2016

Kanazawa et al., CVPR 2016

Su, Qi et al., ICCV 2015
New devices

- Time-of-flight sensors
- Structured light systems
- Light field cameras
- Coded apertures
Large-scale reconstructions

• Internet images pose challenges of scale and outliers
• Reconstructions with millions of images
• Choices to handle data
• Specific optimization approaches

Figure from Agarwal et al.
Real-time 3D vision

- Mobile platforms, embedded systems (IoT devices)
- Stringent demands on computational resources
- Low power platforms (wattage) for automobile ECUs
- Carefully designed and multithreaded architectures

Song and Chandraker, CVPR 2014

Newcombe et al., CVPR 2015
Course Details
Course details

• Each class will cover papers from a topic in 3D vision

• Examples of topics
  – Correspondence estimation
  – Structure from motion
  – Multiview stereo
  – Optical flow
  – Photometric reconstruction
  – RGBD reconstruction
  – Single-view reconstruction
  – Non-rigid and articulated shapes
  – Semantic reconstruction
  – Shape generation
Course details

• Presentation instructions
  – One student to present in each class
  – Discuss topic with instructor one week in advance
  – Send slides by 9pm the night before the class
  – Allow for speaking time of 50 minutes (about 40 slides)
  – Presentation should be well-organized and thoughtful
  – Ask questions and encourage discussions along the way

• Each student does 1-2 presentations
Course details

• Presentation contents
  – Summarize the topic and how the papers address it
  – Why the topic is interesting, or difficulty of the problem
  – Motivate with applications
  – Key technical ideas, why they are interesting
  – Strengths and weakness of proposed methods
  – Detailed analysis of experiments
  – If possible, include own analysis based on author code
  – Open problems, extensions, likely follow-up papers
Course details

• Class participation
  – Day before class, send a brief review of 1 paper
  – Summary and main contribution of the paper
  – Strengths and weaknesses
  – Critique of experiments
  – Suggestions for improvement
  – At least one possible extension or follow-up

• Presenter need not send in review for that class

• Ask questions, answer them, engage in discussions
Course details

• Final project
  – Pick any topic related to 3D reconstruction
  – Discuss with instructor to finalize topic and scope
  – Groups of two are allowed
  – Deadline: week 4

• Options for project
  – Pick a research topic and implement
  – Extend existing papers in interesting ways
  – Improve implementation of existing code from papers
  – Conduct critical survey, identifying possible extensions.
Course details

• Grading
  – 40% class presentations
  – 20% class participation
  – 40% final project

• Aim is to learn together, discuss and have fun!
Good papers to discuss

• Mostly recent works (but interesting older ones fine too)
• Not mere application of CNNs to Problem X
• Using deep learning to solve problems otherwise hard
• Exploit progress in deep learning to open up new problems

• Feel free to discuss papers not based on deep learning
  – By no means are advances limited to deep learning
  – But focus on hard problems or new problems is still required

• Not-so-good papers for discussion
  – Applied hammer X to nail Y
  – Improved x% with great parameter tuning
  – Both very important research skills, but limited classroom time