AMME4710: COMPUTER VISION AND IMAGE PROCESSING
WEEK 9

Dr. Mitch Bryson

School of Aerospace, Mechanical and Mechatronic Engineering, University of Sydney

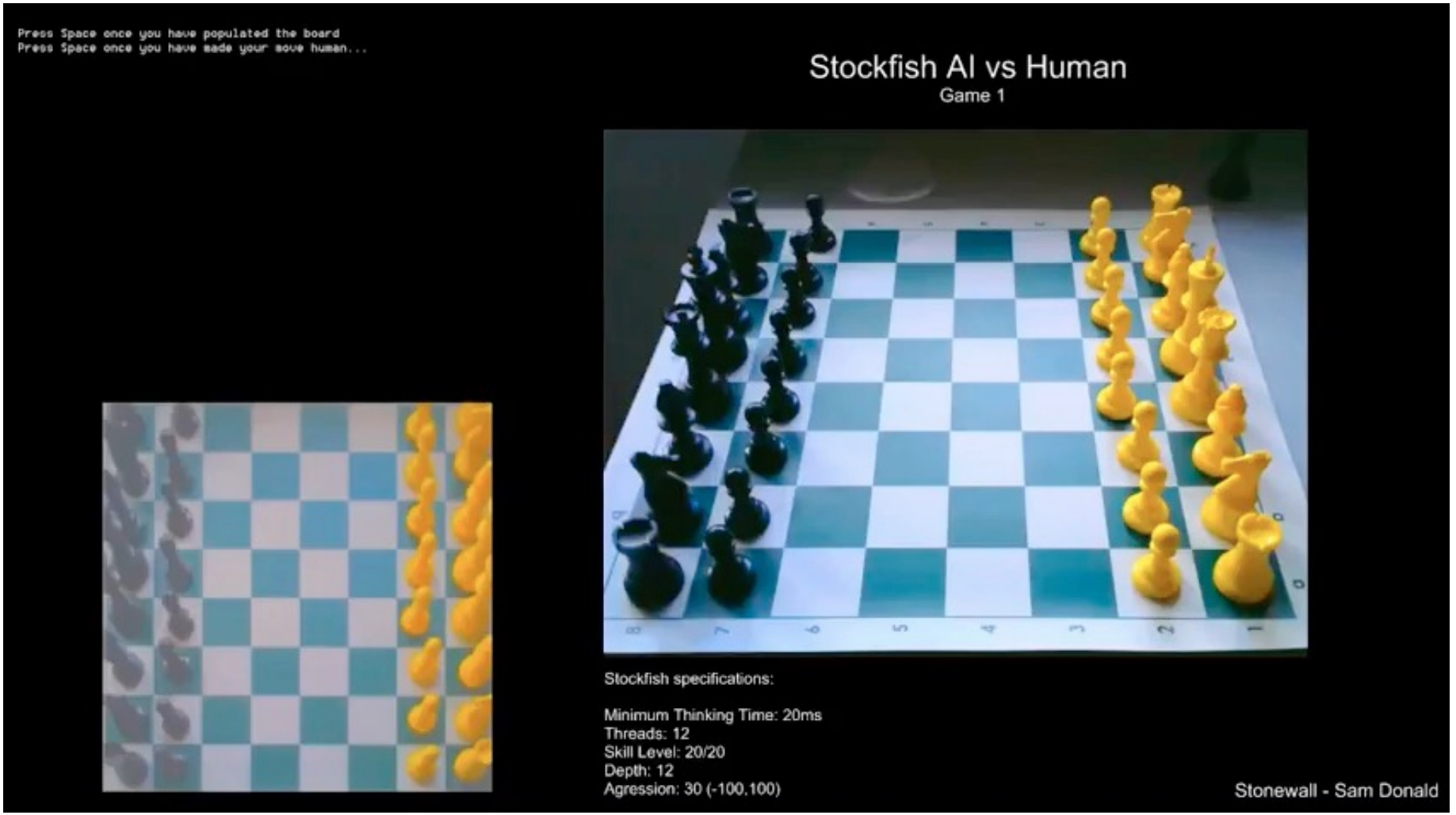
Major Design Project (due Week 13)

- Team project with 3 members, work on a computer vision/image processing problem of your choice
- Apply the theory and algorithms learnt in the course so far (plus other research) to solve a real-world problem
- Give a 10-minute team Zoom presentation on your results (Tuesday/Friday Week 13) (40%)
- Write an 8-page team report (in the format of a CVPR conference paper) (due Sunday Week 13) (60%)

Project A: Board Game Tracking

- Develop a computer vision system to track the exact moves made by players in a board game of your choice (i.e. Chess, Checkers, Go, Chinese Checkers, Monopoly)
- Use a video stream coming from a webcam watching the game
  - How robust is the system to occlusions?
  - Can the system identify/classify and track individual types of pieces? (i.e. a rook vs. a pawn)
  - How much cumbersome/clumsy calibration is needed before using the system?

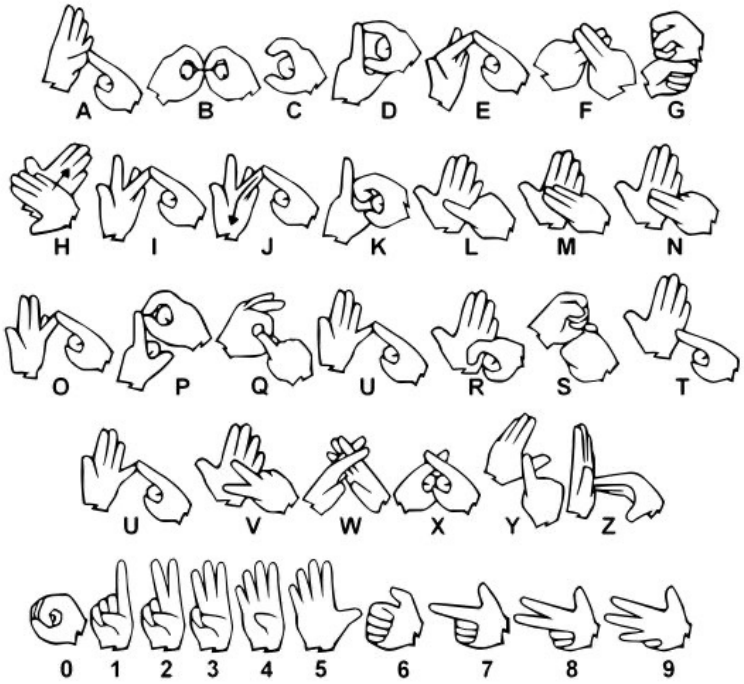
Project A: Board Game Tracking



https://www.youtube.com/watch?v=uDw_vCNbx-I

Project B: Human hand gesture recognition and sign-language translator

- More than 1% of the world's population are unable to hear:
  - Many of these people use sign language as their primary means of communication: in Australia, the Auslan system is used
- Develop a system for recognising Auslan hand gestures from a video stream that can translate or recognise basic sign language letters, numbers or words:
  - How does the person signing have to be oriented with respect to the camera position?
  - Is the system robust to different clothing that the person might be wearing or backgrounds (i.e. does the person need to be positioned in front of a specific background)?
  - How slowly must the gesture be performed?



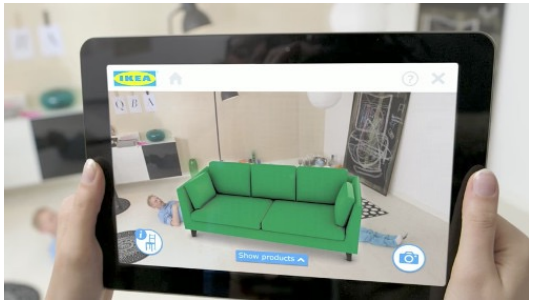
Project B: Human hand gesture recognition and sign-language translator

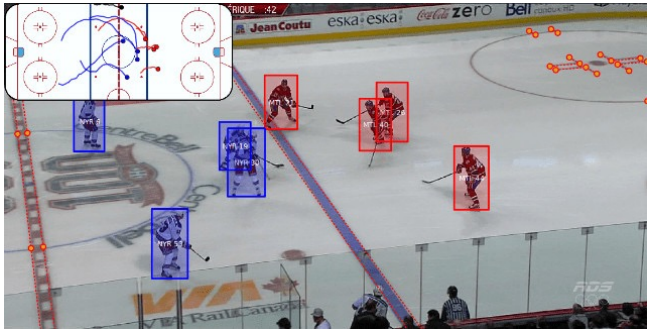
https://www.youtube.com/watch?v=sjf8f__UsdQ

Project B: Human hand gesture recognition and sign-language translator

https://www.youtube.com/watch?v=ITeSHLVNXN8

Project C: Augmented Reality Systems



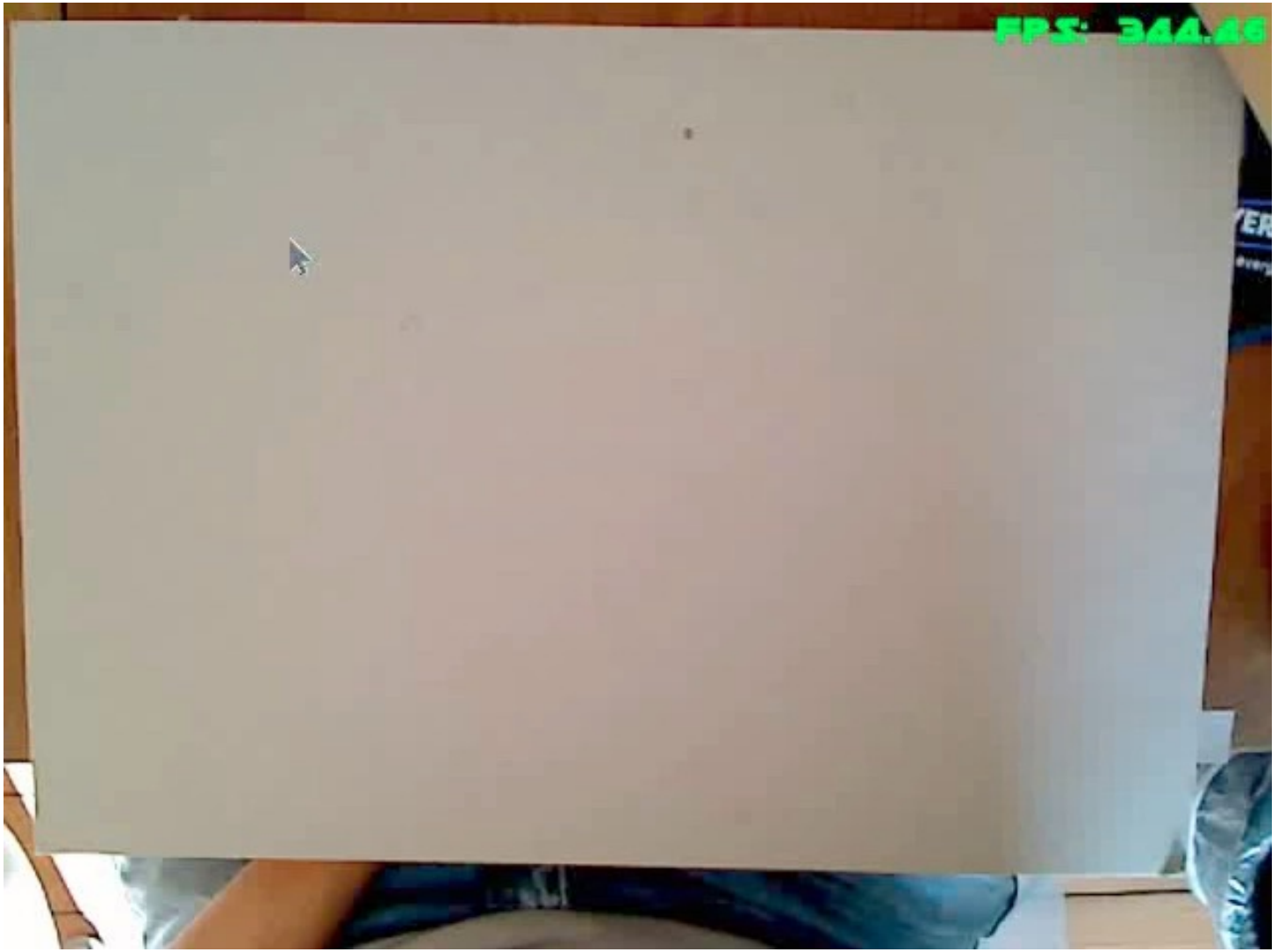


- Augmented reality: the superposition of virtual or computer-generated data onto real-time imagery to create a composite scene composed of real and virtual objects
- Develop an augmented reality system for something fun:
  - Ikea furniture visualiser
  - An augmented reality game
  - Enhancement of sport video by annotated game information
- How are 3D spatial relationships extracted from the video/images?
- Does the method rely on any targets or other markers to be present in the image?
- How do you deal with the interaction between real-world and virtual objects? Which object is occluding which?

Project C: Augmented Reality Systems

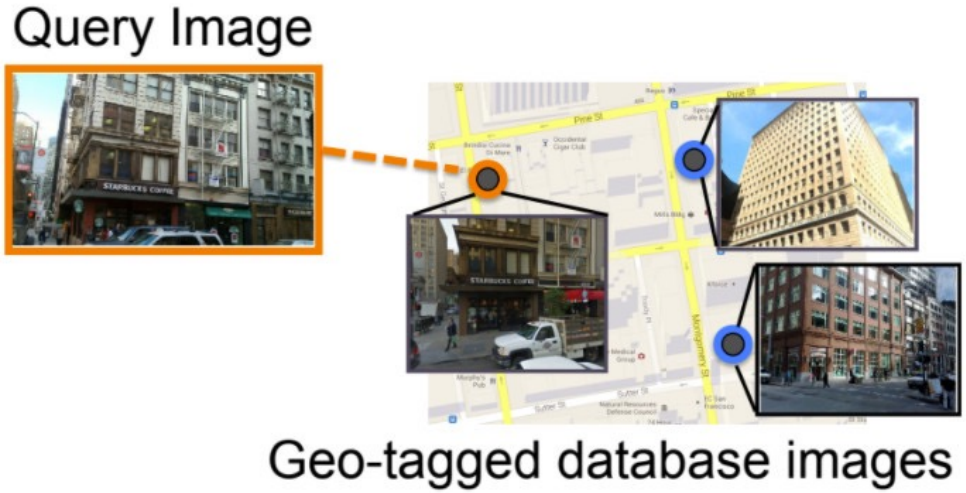
https://www.youtube.com/watch?v=vDNzTasuYEw

Project C: Augmented Reality Systems



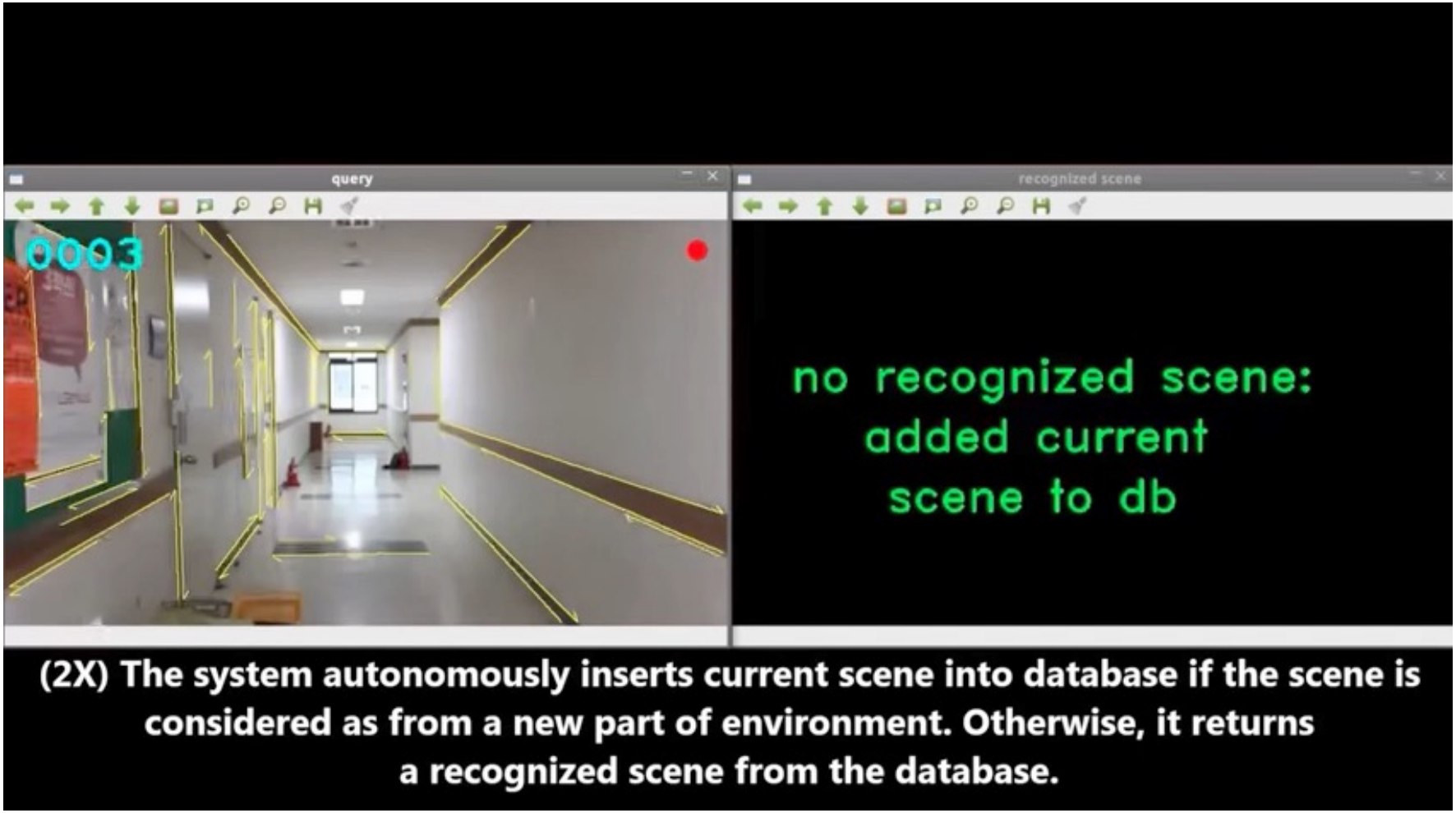
https://www.youtube.com/watch?v=OE9hPGuT8U4

Project D: Virtual Guide/Visual Place Recognition



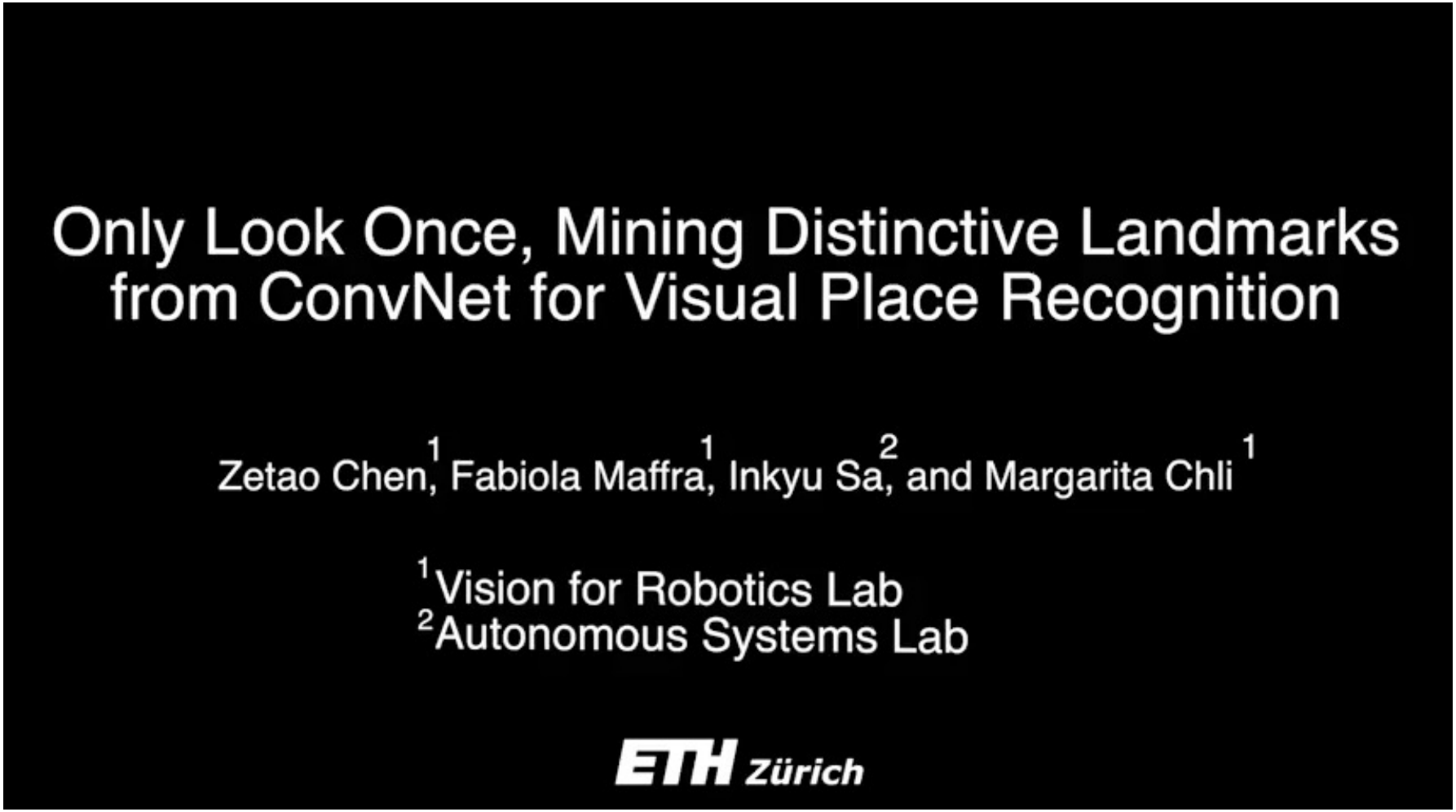
- Develop a system that can work out a user's location on a map using an image of the immediate surroundings
- Use a database that contains geotagged images of an environment, and to use techniques from feature matching and stereo vision to estimate the location from which the current photo was taken based on how it matches to the database
  - How do you account for changes in the perspective between query photos and the image database?
  - How do you account for differences in lighting and appearance that occur, for example due to the time of day (shadows etc.) or day vs. night-time navigation?

Project D: Virtual Guide/Visual Place Recognition



https://www.youtube.com/watch?v=dgbmTI9S-Ko

Project D: Virtual Guide/Visual Place Recognition



https://www.youtube.com/watch?v=s1z0oJz08Sk

Assessment

- Team Presentation (40%)
  - 10-minute presentation to be delivered Tuesday/Friday Week 13
  - Slides due 5pm Monday Week 13

- Team Report (60%)
  - In the format of an 8-page conference paper for CVPR (see template available on canvas site)
  - Due 11:59pm Sunday of Week 13

Journals and conferences

- Major international conferences on computer vision or image processing:
  - **International Conference on Computer Vision** (http://dblp.uni-trier.de/db/conf/iccv/index.html)
  - **European Conference on Computer Vision** (see 2016: http://www.eccv2016.org/proceedings/)
  - **Computer Vision and Pattern Recognition** (http://ieeexplore.ieee.org/xpl/conhome.jsp?punumber=1000147)
  - **International Conference of Computational Photography** (http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=7486620)
- Major computer vision/image processing journals:
  - **International Journal of Computer Vision** (https://link.springer.com/journal/11263)
  - **IEEE Trans. on Image Processing** (http://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=83)
  - **Computer Vision and Image Understanding** (https://www.journals.elsevier.com/computer-vision-and-image-understanding/)
  - **IEEE Trans. on Pattern Analysis and Machine Intelligence** (http://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=34)

Example Image Datasets and other Resources

- Caltech Computer Vision Archives:
  - Contains example image datasets of everyday objects, faces, landscapes, pedestrians etc.
  - http://www.vision.caltech.edu/archive.html
- Auslan Signbank: contains videos and images of signs
  - http://www.auslan.org.au/
- Google Streetview Database
  - https://andrewpwheeler.wordpress.com/2015/12/28/using-python-to-grab-google-street-view-imagery/ : example using python to automatically grab streetview images
  - Google API for accessing streetview data: https://developers.google.com/maps/documentation/javascript/streetview#StreetViewService

Alternative Software Packages

- MATLAB Image Processing Toolbox:
  - Used predominantly in this course and you can continue to use this
- OpenCV:
  - Probably the most-widely used computer vision software toolkit
  - Library interface for C/C++, binding to functions in python, MATLAB and others
- Fiji/ImageJ:
  - Used extensively in scientific and medical image processing, GUI interface to customisable plug-ins in JAVA
- Many others: e.g. Octave/Scilab (alternatives to MATLAB) each have extensive image processing toolboxes

5 minute break

Introduction to OpenCV/Keras

- Live online Notebook via Google Colab (*required Google account to run):
- https://colab.research.google.com/drive/1xvOfm8ZdMAPvb-Sj8euA3LGGR6XUNYNh?usp=sharing

- Download the Python Notebook and run on your own device ("AMME4710_CV_Week9.ipynb" under Modules/Week 9)
  - Requires you first install python (https://www.python.org/)
  - Then install modules for OpenCV, TensorFlow and Keras:
    - pip install opencv-python
    - pip install tensorflow
    - pip install keras

Next Week

- Next Week:
  - Structure-from-motion and 3D reconstruction using images