CS6550 Computer Vision

• **Class Meeting:**
  M7M8 (3:30pm – 5:20pm), R6 (2:20pm – 3:10pm).
  Rm 106 Delta Bldg.,
  台達館 106室

• **Instructor:** Prof. Shang-Hong Lai, Rm. 636 Delta Bldg.,
  賴尚宏, 台達館636室,
  Tel: ext. 42958, Email: lai@cs.nthu.edu.tw,
  URL: http://www.cs.nthu.edu.tw/~lai

• **Office Hours:** R7R8 or by appointment

• **Teaching Assistant:**
  李東穎、蘇宏任, 台達館 720、721, CV lab.
  Email: d9562818@oz.nthu.edu.tw; suhongren@gmail.com
Prerequisite

- Linear Algebra
- Probability and Statistics
- Basic Programming
Course Description

• This course is to provide an introductory background in computer vision for graduate students to start research in this field.

• We will focus on teaching representative computer vision algorithms in class.

• You will need to implement some algorithms with computer programs for the homeworks and the final project.
Course Contents

1. Image Formation (1 week)
2. Image Features (2 weeks)
3. Image Segmentation (2 weeks)
4. Camera Calibration (1 week)
5. Two-View Geometry (1 week)
6. Stereo Reconstruction (1 week)
7. Image Matching (1 week)
8. Motion Analysis (1 week)
9. Object Recognition (1 week)
10. Augmented Reality (1 week)
Textbooks

• Primary:
  http://szeliski.org/Book/

• Secondary:

• Lecture slides distributed in class.
Sample Contents

2. Image Formation

3. Image Processing

4. Features

11. Stereo

5. Segmentation

6-7. Structure from Motion

8. Motion

14. Recognition
Grading

Midterm Exam. (11/26) 30%
Final Project 20%
Homeworks (4) 40%
Class Participation 5%
Quizzes 5%
Homework Policy

• Discussion of homework is encouraged, but you have to write your own. No copying is strictly enforced.

• Homework should be delivered before the announced due time, normally before the lecture. Late homework will be degraded by 25% per day.

• No make-up homeworks 4 days after the deadline.
Course Webpage

- It contains the course slides, basic course information, and class announcement.
- Important course announcement will also be posted on this webpage.
Class Participation

• Class attendance is required and treated as the basic requirement for class participation.
• Asking questions is strongly encouraged.
• Extra credit will be given for finding mistakes or asking questions.
CS 6550 Classroom Rule

• No eating is permitted in class.
• No sleeping is allowed in class.
• Disturbance to others in class should be minimized.
• Cell phone should be turned off during the class.
Computer Vision

Make computers understand images and video.

What kind of scene?
Where are the cars?
How far is the building?
...

[Image of a scene with a building, trees, and cars]
What is Computer Vision?

• To extract useful information about real physical objects and scenes from sensed images/video.
  – 3D reconstruction from images
  – Object detection/recognition
• Automatic understanding of images and video
  – Computing properties of the 3D world from visual data (measurement)
  – Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities. (perception and interpretation)
Vision for measurement

Real-time stereo

Structure from motion

Multi-view stereo for community photo collections

Pollefeys et al.

Goesele et al.

Slide credit: L. Lazebnik
Vision for perception, interpretation
Related Disciplines

- Artificial intelligence
- Machine learning
- Cognitive science
- Computer vision
- Algorithms
- Image processing
- Graphics
Inverse problems: analysis and synthesis.
Why computer vision?

• As image sources multiply, so do applications
  – Relieve humans of boring, easy tasks
  – Enhance human abilities: human-computer interaction, visualization
  – Perception for robotics / autonomous agents
  – Organize and give access to visual content
Why computer vision?

• Images and videos are everywhere!

- Personal photo albums
- Movies, news, sports
- Surveillance and security
- Medical and scientific images

Slide credit; L. Lazebnik
Why computer vision matters?

- Safety
- Health
- Security
- Comfort
- Fun
- Access
Again, what is computer vision?

- Mathematics of geometry of image formation?
- Statistics of the natural world?
- Models for neuroscience?
- Engineering methods for matching images?
- Science Fiction?
Very brief history of computer vision

• 1966: Minsky assigns computer vision as an undergrad summer project
• 1960’s: interpretation of synthetic worlds
• 1970’s: some progress on interpreting selected images
• 1980’s: ANNs come and go; shift toward geometry and increased mathematical rigor
• 1990’s: face recognition; statistical analysis in vogue
• 2000’s: broader recognition; large annotated datasets available; computational photography starts

Guzman ‘68
Ohta Kanade ‘78
Turk and Pentland ‘91
Applications of Computer Vision

• Robot Vision / Autonomous Vehicles
• Biometric Identification / Recognition
• Industrial Inspection
• Video Surveillance
• Digital Camera
• Medical Image Analysis/Processing
• Remote Sensing
• Multimedia Retrieval
• Augmented Reality
Consumer Applications

(a) image stitching: merging different views (Szeliski and Shum 1997) (b) exposure bracketing: merging different exposures.
Real-Time Stereo Camera

Point Grey Research makes video rate stereo camera (640 x 480 at 30 fps).

Bumblebee
3D Reconstruction from Images
Earth viewers (3D modeling)

Image from Microsoft's [Virtual Earth](https://www.bing.com/vodot) (see also: [Google Earth](https://www.google.com/earth))
“What if your photo collection was an entry point into the world, like a wormhole that you could jump through and explore...”

The Photosynth Technology Preview is a taste of the newest - and, we hope, most exciting - way to view photos on a computer. Our software takes a large collection of photos of a place or an object, analyzes them for similarities, and then displays the photos in a reconstructed three-dimensional space, showing you how each one relates to the next.

http://photosynth.net/
Object Detection
Optical Character Recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software

Digit recognition, AT&T labs
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition
Face Detection

- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, …
Smile detection?

The Smile Shutter flow
Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.

Sony Cyber-shot® T70 Digital Still Camera
Face Detection and Recognition

Face detection algorithms, coupled with color-based clothing and hair detection algorithms, can locate and recognize the individuals in this image (Sivic, Zitnick, and Szeliski 2006)
Biometric Recognition

Who is she?
Vision-based Biometrics

How the Afghan Girl was Identified by Her Iris Patterns?
http://www.cl.cam.ac.uk/~jgd1000/afghan.html
Login without a password...

Fingerprint scanners on many new laptops, other devices

Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/
Object Recognition (in mobile phones)

- This is becoming real:
  - Google goggles
  - Point & Find, Nokia
Special effects: shape capture

*The Matrix* movies, ESC Entertainment, XYZRGB, NRC
Sports – Augmented Reality

Sportvision first down line
Nice explanation on www.howstuffworks.com
Google street view
Google street view
Smart Cars

• **Mobileye**
  – Vision systems currently in high-end BMW, GM, Volvo models
Assisted Driving

Pedestrian and car detection

• Collision warning systems with adaptive cruise control,
• Lane departure warning systems,
• Rear object detection systems,

Lane detection
Google Autonomous Car

• The U.S. state of Nevada passed a law in June 2011 concerning the operation of driverless cars in Nevada.

• The Google Driverless Car combines information gathered from Google Street View, video cameras inside the car, a LIDAR sensor on top of the vehicle, radar sensors on the front of the vehicle and a position sensor attached to one of the rear wheels.
Vision-based Interaction

Nintendo Wii has camera-based IR tracking built in.

Control games with your own body motion/gesture and create immersive experiences by combining 3D personal image into the game scene.
Vision in Space

NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
Robotics

NASA’s Mars Spirit Rover

Darpa’s Robotics Challenge
Medical Imaging

3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT
Augmented Reality

• AR allows the user to see the real world, with virtual objects superimposed upon or composited with the real world. Therefore, AR supplements reality, rather than completely replacing it.

• Google Glasses is a research and development program to develop an augmented reality head-mounted display (HMD).
Virtual Dressing Room

TryLive™ Eyewear

TryLive™ Watches
Things to Do

• Read Chap. 1 (Szeliski)
• Next classes
  – Introduction to Matlab programming
  – Image formation (Chap. 2, Szeliski)