EECS 442 – Computer Vision – fall 2012

• Instructor
  – Silvio Savarese
  – silvio@eecs.umich.edu
  – Office: ECE Building, room: 4435
  – Office hour: Tues 4:30-5:30pm or under appoint.

• GSI:
  – Johnny Chao (ywchao125@gmail.com)

• Class Time & Location
  – Tu Th 3:00PM - 4:30PM -- G906 COOL

• Conversation hour [it’s part of the course!]
  – Wed 3:30PM - 4:30PM -- 1013 DOW

http://www.eecs.umich.edu/vision/teaching/EECS442_2011/eecs442.html
If you plan to audit this class, please signup your name on the mailing list
Agenda

• Administrative
  – Grading policy
  – Project

• What is computer vision?

• Syllabus
Grading policy

• Homeworks: 40%
  – 5 homeworks

• Mid term exam: 10% [end of October]

• Course project: 45%
  – progress report 5%
  – final report 30%
  – presentation 10%

• Attendance and class participation: 5%
  – Questions, answers, remarks…

• Late policy home works:
  – If 1 day late, 50% off the grade for that homework
  – Zero credits if more than one day.
  – A “48-hours one-time late submission bonus” is available; that is, you can use this bonus to submit your HW late (i.e. after at most 48 hours). This is one time bonus: After you use your bonus, you must adhere to the standard late submission policy. No exceptions will be made.

• Late policy project:
  – If 1 day late, 25% off the grade for the project
  – If 2 days late, 50% off the grade for the project
  – Zero credits if more than 2 days
  – No "late submission bonus" is allowed when submitting your exam or project.

• Collaboration policy
  – Read the student code book, understand what is ‘collaboration’ and what is ‘academic infraction’.
  – Discussing project assignment with each other is allowed, but coding must be done individually
  – Home works or class project coding policy: using on line code or other students/researchers’ code is not allowed in general. Exceptions can be made and individual cases will be discussed with the instructor.
Course Project

• Replicate an interesting paper
• Comparing different methods to a test bed
• A new approach to an existing problem
• Original research

• Write a 8-page paper summarizing your results
• Release the final code
• Give a presentation

• We will introduce projects in two weeks

• Important dates: look up class schedule
Course Project

• Form your team:
  – 1-3 people
  – the quality is judged regardless of the number of people on the team
  – be nice to your partner: do you plan to drop the course?

• Evaluation
  – Quality of the project (including writing)
  – Final ~15 minutes project presentation in class – students will vote your presentation!
  – For final code and paper due dates please consult webpage
Agenda

• Administrative
  – Grading policy
  – Project

• What is computer vision?

• Syllabus
“There was a table set out under a tree in front of the house, and the March Hare and the Hatter were having tea at it.”

“The table was a large one, but the three were all crowded together at one corner of it …”

From “A Mad Tea-Party” Alice's Adventures in Wonderland by Lewis Carroll
Computer vision

Image/video

Object 1
- semantic
- geometry

... Object N
- semantic
- geometry
Computer vision

Image/video

Object 1

- semantic
- geometry

... Object N

- semantic
- geometry

Spatial & temporal relations
Computer vision

Image/video

Object 1
- semantic
- geometry

Object N
- semantic
- geometry

spatial & temporal relations

Scene
- Semantic
- geometry
Computer vision

Computer vision studies the **tools and theories** that enable the design of machines that can **extract useful information from imagery data** (images and videos) toward the goal of **interpreting the world**.

- **Scene**
- **Objects**
- **People**
- **Actions**

**Information**: visual cues, 3D structure, motion flows, etc…

**Interpretation**: recognize objects, scenes, actions, events
Have we reached humans?

… not yet

- computer vision is still no match for human perception
- but catching up, particularly in certain areas
Optical character recognition (OCR)

Technology to convert scanned docs to text

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

Digit recognition, AT&T labs

Source: S. Seitz
Biometrics

Fingerprint recognition
Biometrics

Iris Patterns recognition

How the Afghan Girl was Identified by Her Iris Patterns

Source: S. Seitz
Biometrics

Face recognition systems
http://www.sensiblevision.com/

Source: S. Seitz
Face-hunting cameras boost Nikon

Japanese camera maker Nikon has tripled its profits on the back of strong sales of digital cameras that automatically focus on human faces.

Face recognition cameras like the Coolpix L1 are popular.
Consumer applications

Face detection

• Many new digital cameras now detect faces
  – Canon, Sony, Fuji, …

http://www.apple.com/ilife/iphoto/

Source: S. Seitz
Consumer applications

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

Source: S. Seitz
Special effects movies - videogames

Source: S. Seitz
3D urban modeling

Source: S. Seitz

Bing maps, Google Streetview
3D urban modeling: Microsoft Photosynth

http://labs.live.com/photosynth/

Source: S. Seitz
Toy & Robots

Credit slide: Fei-Fei Li
Mobile visual search:

Google Goggles in Action

Click the icons below to see the different ways Google Goggles can be used.
Vision in supermarkets

LaneHawk by EvolutionRobotics

Source: S. Seitz
Vision-based interaction (and games)

Microsoft’s Kinect
Sony EyeToy
Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read “Computer Vision on Mars” by Matthies et al.
Applications of computer vision

Medical Imaging

Assistive technologies

Driver assistance
(collision warning, lane departure warning, rear object detection)

Surveillance

Autonomous driving,
robot navigation

Security

Sources: K. Grauman, L. Fei-Fei, S. Laznebick
Computer vision

- Automatic control
- Robotics
- Signal processing
- Compression
- Non linear SP
- Multi-variate SP
- Robot vision
- Neurobiology
- Physics
- Imaging
- Data mining
- Image retrieval
- Visual pattern recognition
- Statistics
- Geometry
- Optimization
- Machine learning
- Artificial intelligence
- Applied math
- Computer graphics
- Machine learning
- Applied math
- Artificial intelligence
EECS 442 course overview

1. Geometry
2. Low & Mid-level vision
3. High level vision
EECS 442 course overview

1. Geometry
2. Low & Mid-level vision
3. High level vision

Geometry:
- How to extract 3d information?
- Which cues are useful?
- What are the mathematical tools?
Visual cues: texture • shading • contours • shadows • reflections
Visual cues:  texture  • shading  • contours  • shadows  • reflections
Visual cues: texture • shading • contours • shadows • reflections
Visual cues: texture • shading • contours • shadows • reflections
Visual cues: texture • shading • contours • shadows • reflections
Visual cues: texture • shading • contours • shadows • reflections

Number of observers: monocular • multiple views
Vision techniques

Number of observers: monocular • multiple views

Visual cues: texture • shading • contours • shadows • reflections
Stereo

Epipolar geometry

Tomasi & Kanade (1993)

Structure from motion

Image sources: S. Laznebick
Structure from motion

Courtesy of Exford Visual Geometry Group
Visual cues:  texture • shading • contours • shadows • reflections

Number of observers:  monocular • multiple views

Active lighting:  laser stripes • structured lighting patterns
3D Scanning

Scanning Michelangelo’s “The David”

- The Digital Michelangelo Project
- 2 BILLION polygons, accuracy to .29mm

Courtesy of Stanford computer graphics lab
3D Digital models

Architecture

Entertainment

Medicine
The ‘yellow line’ in Superbowls

THE VIRTUAL FIELD
The computer-generated map of the field appears as a blue grid on the computers used. It is manipulated to fit the cameras’ views.

THE LOOK OF THE LINE
The size and appearance of the line can be changed. It can look like paint on artificial surfaces or like chalk on grass fields.

CAMERA POSITIONS
Three cameras are used in the process. One is at the 50-yard line, and the others are at about the 25-yard lines.
The ‘yellow line’ in Super Bowls
Course overview

1. Geometry

2. Low & Mid-level vision

3. High level vision

Mid-level vision:
- Extract useful building blocks
- Region segmentation
- Motion flows
Extract useful building blocks
Extract useful building blocks

Alignment

Extract planar regions

Object segmentation
Image enhancement

Image Inpainting, M. Bertalmío et al.
http://www.iua.upf.es/~mbertalmio/restoration.html
Automatic Panorama Stitching

Sources: M. Brown
Automatic Panorama Stitching
Image Modeling
Tracking and 3D modeling
Camera tracking and V.R. insertions

Courtesy of Exford Visual Geometry Group
Course overview

1. Geometry
2. Low & Mid-level vision
3. High level vision

High level operations
- Recognition of objects and people
- Places
- Actions & events
Object recognition and categorization

- Building
- Downtown chicago
- Pedestrians crossing street
- clock
- person
- car
Challenges: viewpoint variation

Michelangelo 1475-1564

slide credit: Fei-Fei, Fergus & Torralba
Challenges: illumination

image credit: J. Koenderink
Challenges: scale
Challenges: deformation
Challenges: occlusion
Challenges: background clutter

Kilmeny Niland. 1995
Challenges: object intra-class variation

slide credit: Fei-Fei, Fergus & Torralba
Face recognition
Recognizing scenes
Sorting out millions of images/videos
Detecting and tracking people
Recognizing human activities

- Biking
- Walking
- Walking

Credit slide: Song & Perona
“There was a table set out under a tree in front of the house, and the March Hare and the Hatter were having tea at it.”

“The table was a large one, but the three were all crowded together at one corner of it …”

From “A Mad Tea-Party” Alice's Adventures in Wonderland by Lewis Carroll
Syllabus

See webpage
Next lecture

- Review of linear algebra for multi-view geometry
- Basic image transformations