CPSC 425: Computer Vision  
(Jan-April 2007)  

David Lowe

Prerequisites: 4th year ability in CPSC  
Math 200 (Calculus III)  
Math 221 (Matrix Algebra: linear systems)  
Useful: Numerical analysis
Why study Computer Vision?

• Images and video are everywhere
• Fast-growing collection of useful applications
  – matching and modifying images from digital cameras
  – film special effects and post-processing
  – building representations of the 3D world from pictures
  – medical imaging, household robots, security, traffic control, cell
    phone location, face finding, video game interfaces, ...
• Various deep and attractive scientific mysteries
  – what can we know from an image?
  – how does object recognition work?
• Greater understanding of human vision and the brain
  – about 25% of the human brain is devoted to vision
Vision is inferential: Illumination

http://web.mit.edu/persci/people/adelson/checkershadow_illusion.html
• **Course requirements**
  – 4 to 5 homework assignments (Matlab and written exercises)  [25% of final mark]
  – Midterm exam (75 minutes, during class)  [25%]
  – Final exam (2.5 hours, scheduled by the registrar)  [50%]

• **My expectations**
  – Read assigned textbook sections and readings in advance
  – Ask questions
  – Complete all assignments on time
  – Never claim credit for work done by others
Textbook

*Computer Vision* by Forsyth and Ponce

- Available in the bookstore now
- Readings will be assigned with each class
- Only one edition is available, so second-hand copies are fine

**Reading for next class:**

Chapter 1
Applications of Computer Vision: Texture generation

Input image

Simple repetition

New texture generated from input

We will do this for a homework assignment
Application: Football first-down line

Requires (1) accurate camera registration; (2) a model for distinguishing foreground from background

www.sportvision.com
Application: Augmented Reality

Application areas:
– Film production (the “match move” problem)
– Heads-up display for cars
– Tourism
– Architecture
– Training

Technical challenges:
– Recognition of scene
– Accurate sub-pixel 3-D pose
– Real-time, low latency
Application: Medical augmented Reality

Visually guided surgery: recognition and registration
Application: Automobile navigation

Lane departure warning

Pedestrian detection

Mobileye (see mobileye.com)

• Other applications: intelligent cruise control, lane change assist, collision mitigation

• Systems already used in trucks and high-end cars
Course Overview

Part I: The Physics of Imaging

• How images are formed
  – Cameras
    • What a camera does
    • How to tell where the camera was (pose)
  – Light
    • How to measure light
    • What light does at surfaces
    • How the brightness values we see in cameras are determined
Course Overview
Part II: Early Vision in One Image

• Representing local properties of the image
  – For three reasons
    • Sharp changes are important in practice -- find “edges”
    • We wish to establish correspondence between points in different images, so we need to describe the neighborhood of the points
    • Representing texture by giving some statistics of the different kinds of small patch present in the texture.
      – Tigers have lots of bars, few spots
      – Leopards are the other way
Course Overview
Part III: Vision in Multiple Images

• The geometry of multiple views
  – Where could it appear in camera 2 (3, etc.) given it was here in 1?
  – Stereopsis
  – What we know about the world from having 2 eyes

• Structure from motion
  – What we know about the world from having many eyes
    • or, more commonly, our eyes moving.

• Correspondence
  – Which points in the images are projections of the same 3D point?
  – Solve for positions of all cameras and points.
Course Overview
Part IV: High Level Vision

• Model based vision
  • find the position and orientation of known objects

• Using classifiers and probability to recognize objects
  – Templates and classifiers
    • how to find objects that look the same from view to view with a classifier
  – Relations
    • break up objects into big, simple parts, find the parts with a classifier, and then reason about the relationships between the parts to find the object
Course Overview
Object and Scene Recognition (my research)

• **Definition:** Identify objects or scenes and determine their pose and model parameters

• **Applications**
  – Industrial automation and inspection
  – Mobile robots, toys, user interfaces
  – Location recognition
  – Digital camera panoramas
  – 3D scene modeling
Invariant Local Features

- Image content is transformed into local feature coordinates that are invariant to translation, rotation, scale, and other imaging parameters.
Examples of view interpolation
Recognition using View Interpolation