6.801/6.866 Machine Vision
Fall 2002
Tuesday and Thursday from 2:30-4:00 in 2-105

Announcements
• The first class will be held on Thursday, September 5th.

Course Information
• Syllabus
• Grading and Requirements
• Internet Resources

Contacts
All offices are on the sixth floor of 400 Technology Square (Forrester building).

Instructors
Professor Bill Freeman
wtf@ai.mit.edu
NE43-V601
(617) 253-8828
Professor Trevor Darrell
trevor@ai.mit.edu
NE43-V605
(617) 253-8966

Office Hours
TBD
TBD

Teaching Assistants
Erik Sudderth
esudder@mit.edu
NE43-V609
(617) 253-3497
Louis-Philippe Morency
lmorency@ai.mit.edu
NE43-V603
(617) 253-4278

Office Hours
TBD
TBD

http://www.ai.mit.edu/courses/6.801/
Map showing 400 Technology Square

The building says "Forrester" on the side. Only the parking garage side building entrance is unlocked. (After normal business hours, the elevator to our floor and the building itself are both locked.) Exiting the elevator on the 6th floor, you'll see a pair of glass doors on one side. Enter the left glass door, then turn right at every opportunity to find my office, room 601.

back to my home page, Sept., 2002.
6.801/6.866 Machine Vision

Course Requirements

Prerequisites

- Signals and Systems (6.003)
- Familiarity with linear algebra (18.06) and probability (6.041)
- Some Matlab experience

Textbooks

Both texts are available at Quantum Books:

Grading

- Two take-home exams
- Four problem sets with lab exercises in Matlab
- No final exam
- Final project for graduate students (6.866)

Problem sets may be discussed, but all written work and coding must be done individually. Take-home exams may not be discussed. Individuals found submitting duplicate or substantially similar materials due to inappropriate collaboration may get an F in this class and other sanctions.

<table>
<thead>
<tr>
<th></th>
<th>Take-Home Exams (2)</th>
<th>Problem Sets (4)</th>
<th>Final Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.801</td>
<td>60%</td>
<td>40%</td>
<td>N/A</td>
</tr>
<tr>
<td>6.866</td>
<td>40%</td>
<td>30%</td>
<td>30%</td>
</tr>
</tbody>
</table>

The final project may be

- An original implementation of a new or published idea
- A detailed empirical evaluation of an existing implementation of one or more methods
- A paper comparing three or more papers not covered in class, or surveying recent literature in a particular area

A project proposal not longer than two pages must be submitted and approved before the end of October.
6.801/6.866 Machine Vision

Internet Resources

Matlab Links

Mathworks Online Documentation
Online versions of the official Matlab documentation, including a general command reference and information about the image processing toolbox.

Matlab on Athena
Guide to using Matlab at MIT. Athena consulting's stock answers may also be useful.

University of Colorado Matlab Tutorials
A decent collection of Matlab tutorials, including one focusing on image processing.

Matlab Image Processing Tutorial
A short introduction to the manipulation of images in Matlab, including an introduction to principal components analysis via eigenfaces.

Matlab Educational Sites
An index of various Matlab tutorial websites.

Computer Vision Homepage
# 6.801/6.866 Machine Vision

## Syllabus

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>9/5</td>
<td>Course Introduction</td>
<td>Pset #0 (not collected)</td>
</tr>
<tr>
<td>9/10</td>
<td>Cameras, Lenses, and Sensors</td>
<td></td>
</tr>
<tr>
<td>9/12</td>
<td>Radiometry and Shading Models</td>
<td>Pset #1 Assigned</td>
</tr>
<tr>
<td>9/17</td>
<td>Color</td>
<td></td>
</tr>
<tr>
<td>9/19</td>
<td>Epipolar Geometry</td>
<td></td>
</tr>
<tr>
<td>9/24</td>
<td>Stereo</td>
<td>Pset #1 Due</td>
</tr>
<tr>
<td>9/26</td>
<td>Shape from Shading</td>
<td>Pset #2 Assigned</td>
</tr>
<tr>
<td>10/1</td>
<td>Filtering I</td>
<td></td>
</tr>
<tr>
<td>10/3</td>
<td>Filtering II</td>
<td></td>
</tr>
<tr>
<td>10/8</td>
<td>Intro to Bayesian Vision</td>
<td>Pset #2 Due</td>
</tr>
<tr>
<td>10/10</td>
<td>Edge Detection and Optic Flow</td>
<td>Exam #1 Assigned</td>
</tr>
<tr>
<td>10/15</td>
<td>Columbus Vacation (NO LECTURE)</td>
<td></td>
</tr>
<tr>
<td>10/17</td>
<td>Direct SFM</td>
<td></td>
</tr>
<tr>
<td>10/22</td>
<td>Models of Texture</td>
<td>Exam #1 Due</td>
</tr>
<tr>
<td>10/24</td>
<td>Statistical Classifiers I</td>
<td>Pset #3 Assigned</td>
</tr>
<tr>
<td>10/29</td>
<td>Statistical Classifiers II</td>
<td></td>
</tr>
<tr>
<td>10/31</td>
<td>Multi-view Geometry</td>
<td></td>
</tr>
<tr>
<td>11/5</td>
<td>Affine SFM</td>
<td>Pset #3 Due</td>
</tr>
<tr>
<td>11/7</td>
<td>Projective Geometry vs. Perspective Projection</td>
<td>PSet #4 Assigned</td>
</tr>
<tr>
<td>11/12</td>
<td>Clustering &amp; Segmentation</td>
<td></td>
</tr>
</tbody>
</table>
### 6.801/6.866 Machine Vision

**Syllabus**

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Due Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/12</td>
<td>Clustering &amp; Segmentation</td>
<td></td>
</tr>
<tr>
<td>11/14</td>
<td>EM</td>
<td></td>
</tr>
<tr>
<td>11/19</td>
<td>Hough Transforms, RANSAC, &amp; Voting Methods</td>
<td>Pset #4 Due</td>
</tr>
<tr>
<td>11/21</td>
<td>Tracking &amp; Density Propagation</td>
<td>Exam #2 Assigned</td>
</tr>
<tr>
<td>11/26</td>
<td>Model-Based Vision</td>
<td></td>
</tr>
<tr>
<td>11/28</td>
<td>Thanksgiving (NO LECTURE)</td>
<td></td>
</tr>
<tr>
<td>12/3</td>
<td>Visual Surveillance / Activity Monitoring</td>
<td></td>
</tr>
<tr>
<td>12/5</td>
<td>Image Databases</td>
<td>Exam #2 Due</td>
</tr>
<tr>
<td>12/10</td>
<td>Image-Based Rendering</td>
<td>Projects Due</td>
</tr>
</tbody>
</table>
Vision

• What does it mean, to see? “to know what is where by looking”.

• How to discover from images what is present in the world, where things are, what actions are taking place.

from Marr, 1982
Computer vision class, fast-forward
Images and image formation
Cameras, lenses, and sensors

- Pinhole cameras
- Lenses
- Projection models
- Geometric camera parameters

Figure 1.16 The first photograph on record, *la table servie*, obtained by Nicéphore Niepce in 1822. *Collection Harlinge–Viollet.*

Radiometry

Bidirectional Reflectance
Distribution Functions: Causes

Wolfgang Lucht, 1997

Mirror BRDF: specular reflectance
Rough water surface BRDF: sunglint reflectance

Volume scattering BRDF: leaf/vegetation reflectance
Gap-driven BRDF (Forest): shadow-driven reflectance

http://geography.bu.edu/brdf/brdfexpl.html
4.1 **NEWTON'S SUMMARY DRAWING** of his experiments with light. Using a point source of light and a prism, Newton separated sunlight into its fundamental components. By reconverging the rays, he also showed that the decomposition is reversible.

From Foundations of Vision, by Brian Wandell, Sinauer Assoc., 1995
Low-level vision
Image filtering

- Review of linear systems, convolution
- Bandpass filter-based image representations
- Probabilistic models for images

Oriented, multi-scale representation
Non-linear filtering, and applications

Sample session of television viewing. (a) Television is off, but searching for the trigger gesture. (b) Viewer shows trigger gesture (open hand). Television set turns on and hand icon and graphics overlays appear. (c) The hand icon tracks the user’s hand movement. User changes controls as with a mouse. (d) User has moved hand icon to change channel. (e) User closes hand to leave control mode. After one second, the hand icon and controls then disappear.

Models of texture

A parametric model

A Non-parametric model

A Parametric Texture Model based on Joint Statistics of Complex Wavelet Coefficients


Shape from shading

Shape from shading, Horn and Brooks, MIT Press, 1986
Bayesian framework for vision

“Good lord, Holmes! How did you come to know I'd seafood for lunch?”

Gahan Wilson’s Still Weird, Forge, 1994
Bayesian framework for vision

Coincidental appearance of face profile in rock?

http://www.cs.dartmouth.edu/whites/old_man.html
Bayesian framework for vision

Coincidental appearance of faces in rock?

http://bensguide.gpo.gov/3-5/symbols/print/mountrushmore.html
Edge detection and optical flow

Edge detection and optical flow

Images

True optical flow

Estimated optical flow

Learning and vision
Statistical classifiers

- MIT Media Lab face localization results.
- Applications: database search, human machine interaction, video conferencing.
Support vector machines and boosting

Large-margin classifier

www.support-vector.net/nello.html
Support vector machines and boosting

“The kernel trick”

www.support-vector.net/nello.html
Applications

• Computer vision for computer games
Computer vision applications as ocean-going vessels
Game: Decathlete
Optical-flow-based Decathlete figure motion analysis
Decathlete 100m hurdles
Decathlete javelin throw
Decathlete javelin throw
video
Nintendo Game Boy

- Several million sold (most of any digital camera). Imaging chip is Mitsubishi Electric’s “Artificial Retina” CMOS detector.
video
The course, in broad categories

• Images and image formation
• Geometry
• Low-level vision
• High-level vision
• Implementations and applications