ECE-161C

Nuno Vasconcelos

ECE Department, UCSD
The course

- The course will cover the most important aspects of image processing and computer vision.
- We will cover a lot of ground, at the end of the quarter you will know how to implement a lot of things that may seem very complicated today.
- **Homework:** one problem set every week. Whenever possible it will include a small computational problem. It will count 20%, but really no chance without it.
- **Exams:** 1 mid-term, date TBA- 35%
  1 final – 45% (covers everything)
Homework policies

- Homework is individual
- OK to work on problems with someone else but you have to:
  - write your own solution
  - write down the names of who you collaborated with
- Homework is due one week after it is issued. No exceptions.
- All grading will be with respect to the class average
  - do not get discouraged if a problem is hard, just put some extra effort into solving it. It is your opportunity to shine…
  - if someone else is cheating, your grades will suffer. It is going to hurt you, not me or the TA
Resources

Course web page:
http://www.svcl.ucsd.edu/~courses/ece161c/
- all handouts, problem sets, code will be available there

TA:
- Sunhyoung Han, s1han@ucsd.edu,

Me:
- Nuno Vasconcelos, nuno@ece.ucsd.edu, EBU 1-5602

My assistant:
- Priscilla Haase (haase@ece.ucsd.edu), EBU1 - 3401, may sometimes be involved in administrative issues
Resources

► office hours:
  • for problem set questions, see the TA first
    • Wednesday, 3:30p-5:00PM, EBU1 5706
  • other questions, my office hours:
    • Tuesday, 6:30-8:00 PM, EBU1 5602

► this is a relatively large class, try to bring your questions ready

► discussion session:
  • Mon., 5:00p-5:50p, PETER102
Texts

We will follow:

• Very up-to-date, will follow closely

Various other good, but optional, texts:

• “Two Dimensional Signal and Image Processing”, J. Lim, P. Hall, 1990
• some lectures will be from this book

• “Pattern Classification”, Duda, Hart, Stork, Wiley, 2001
• “Vision science”, S. Palmer, MIT Press, 1999
Texts

stuff you must know inside out:

• “Discrete-time Signal Processing”, Oppenheim, Schaffer, Prentice, 1999
  • this is an absolute pre-requisite!!!
  • I will not have time to go over this stuff again
  • a lot of what we will do will build on it
  • you are assumed to know it and know it well!

also important

• “Linear Algebra”, Gilbert Strang, 1988
• or any other equivalent books
The course

Why computer vision?

Images are important
  • web, networks, images available everywhere
  • video driving force for many innovations (e.g. networking, cell-phones). You need to know about it no matter what you do.

Vision is a generalization of image processing
  • IP aims to make images look better (compressed, restored, etc.), vision aims to understand them.
  • if you can understand you can make them look better, not clear how to otherwise.

Vision gives you a broader exposure:
  • lies at the intersection of deterministic (filtering, convolutions, Fourier, etc) and statistical signal processing (expectations, covariances, classification error):
The course (cont’d)

- Vision is fundamentally **harder**,
  - in the absence of biological vision systems we would have given up a long time ago
  - more science to be done

- **Connections to cognitive science**
  - interest/familiarity with literature from other areas
  - broader exposure

- **Vision has a broader range of applications**
  - you will be more likely to get a job
  - e.g. vision is hot for homeland defense
  - you are more likely to shine after graduation
Applications

- there are many for vision
  - understanding actions
Applications

- object recognition
Applications

- navigation aids
Applications

- surveillance

and many other...
Topics

► We will cover five main areas
  • inspiration comes from biological vision

► Image formation:
  • how do images get captured?

► Low-level vision:
  • basically filtering
  • you will be shocked by how much information we can extract by filtering images in different ways.

► Mid-level vision:
  • what are the components of an image? (segmentation)
  • how can we fit models to our images/video?
  • what is the scene motion and how do we measure it?
Topics (cont’d)

High-level vision:
- a lot of statistics
- how do we find people, skin, various objects
- how do we classify scenes?

Compression (time permitted)
- the JPEG and MPEG standards
- it is one of the most important applications
Image formation

- What equations govern the projection from the 3D world to the 2D image plane?

- What types of light sources are there? Does that matter?
Image formation

- How do we measure the amount of light emitted by each object?

- How do we measure color? What is a color space?

- Pre-requisites:
  - basic calculus, integration, derivatives, etc.
Low level vision:

We will start with **2D signal processing**:

- 2D convolution
- 2D Fourier transforms
- 2D sampling
- 2D etc

**Useful thing to know**: one of two possibilities

- straightforward extension of 1D (e.g. all of the above)
- or intractable (e.g. factoring polynomials)

**For this reason we will spend little time on this**

**I assume you know the 1D case in all possible ways, shapes, and forms!!!**
Low-level vision (cont’d)

- The importance of scale, pyramids and multi-resolution

- Filtering to get derivatives, edges, corners, etc.
Low level vision (cont’d)

What is texture and how do we characterize/classify/segment it?

Pre-requisites: know your 1D signal processing well!
Mid-level vision

- From localized to global processing
- Tasks that are essential to achieve vision goals, but not goals by themselves
- We will study problems like grouping and segmentation
- Most are incredibly hard
Mid-level vision (cont’d)

- One common solution is to fit a model to the images
- How does one minimize the error of the fit?
Mid-level vision (cont’d)

Motion is an important cue for many tasks:
- surveillance
- compression
- action recognition
- registration
- segmentation

Various types:
- global (camera)
- local (clouds, leaves, etc)
- object (people, cars)
- rigid/non-rigid (cars vs people)

Many algorithms
Mid-level vision (conclusion)

▶ Pre-requisites:

▶ Basic probability and statistics:
  • do you know what a random variable is?
  • what about its expected value?
  • a Gaussian?
  • independence?

▶ Algebra, Algebra, Algebra, Algebra:
  • matrix/vector notation
  • four fundamental spaces of a matrix
  • eigenvectors/eigenvalues, diagonalization

▶ This week’s homework tests some of these
High-level vision

- Solves problems that are end-goals of vision.
- Involves lots of classification/categorization problems
- Mostly based on statistical modeling
- E.g. how do we find if there is human skin in the image?
High-level vision (cont’d)

What about peoples faces? Or cars?
High-level vision (cont’d)

To do this, we will learn about statistical classification methods

• linear regression
• principal component analysis
• Bayes decision rule
• support vector machines
• boosting

This will be a mini-course in pattern recognition and statistical learning

Pre-requisites:

• basic statistics,
• algebra, algebra, algebra
Compression

- We will discuss the fundamental ideas in compression
- JPEG, MPEG, and how all the rest applies to them
Things you will have learned

► 2D signal processing fundamentals
  • there is really no mystery
  • spectral/statistical properties of images
  • a little bit on brains and why they are like they are

► Statistics/stochastic processes
  • introduction/applied maximum likelihood, least squares, Bayesian methods
  • covariance, Gaussian models, principal components, the DCT
  • classification and optimal classifiers
Things you will have learned

- How do I design an optimal classifier?
- What is a support vector machine? Boosting?

This can be very useful even if you don’t care at all about images

- data mining
- all sorts of signal classification
Things you will have learned

- Image databases, classification, and retrieval
- How to design the new Google for pictures?

- How do we say that two images are similar?
Things you will have learned

- Object recognition
- Face recognition
- etc