**Syllabus for 4476 (Introduction to Computer Vision)**

**Course Summary**

This course provides an introduction to computer vision including fundamentals of image formation, camera imaging geometry, feature detection and matching, stereo, motion estimation and tracking, image classification and scene understanding. We’ll develop basic methods for applications that include finding known models in images, depth recovery from stereo, camera calibration, image stabilization, automated alignment, tracking, and recognition.

The focus of the course is to develop the intuitions and mathematics of the methods in lecture, and then to learn about the difference between theory and practice in the problem sets.

The difference between the undergraduate version of the class (CS4476) and the graduate version (CS6476) will be the requirements on the problem sets. In particular, more challenging variations of the problems will be extra credit for CS4476 but required for CS6476. Graduate students will also be studying a few recent techniques from papers. The Advanced Computer Vision course (CS7476) will build of CS6476 and deal with advanced and research related topics in Computer Vision, including Machine Learning and Robotics topics that impact Computer Vision.

**Why Take This Course?**

Images have become ubiquitous in computing. Sometimes we forget that images often capture the light reflected from a physical scene. This course gives you both insight into the fundamentals of image formation and analysis, as well as the ability to extract information much above the pixel level. These skills are useful for anyone interested in operating on images in a context-aware manner or where images from multiple scenarios need to be combined or organized in an appropriate way.

**Prerequisites and Requirements**

* Data structures: You'll be writing code that builds representations of images, features, and geometric constructions.
* Programming: A good working knowledge of programming environments that support image and video analysis. Recently, this includes MATLAB and/or Python with NumPy. The lectures typically use MATLAB for discussing algorithms and the occasional demonstration. Problem sets can be done in Matlab or Python.
* Math: Linear algebra, vector calculus, linear algebra, probability and linear algebra (that is not a typo).
* No prior knowledge of vision is assumed though any experience with Signal Processing is helpful.

**Learning Objectives**

Upon completion of this course, students will be able to:

1. Become familiar with both the theoretical and practical aspects of computing with images building on the image processing approaches
2. Describe the foundation of image formation and image analysis.
3. Become familiar with theoretical foundations of the major technical approaches involved in computer vision based image analysis.
4. Understand basics of measurements and robust detection of features in images.
5. Describe various methods used for registration, alignment, and matching in images.
6. Understand the basics of 2D and 3D Computer Vision.
7. Get an exposure to advanced concepts leading to object and scene categorization from images.
8. Be able to connect issues from Computer Vision to Human Vision
9. Develop practical skills that are necessary for building computer vision applications.

**Assignments and Exams**

The primary assessment is done through problem sets that require implementing algorithms and applying them to provided images. Problem sets account for 85% of the grade. There iwill be an EXAM that comprises the remaining 15%.

**Attendance**

Attendance is required and periodical attendance polls will be taken without prior notice.

**Academic Integrity**

Academic dishonesty will not be tolerated. This includes cheating, lying about course matters, plagiarism, or helping others commit a violation of the Honor Code. Some exams (when specifically announced in class) allow the use of self-prepared supporting information (one sheet of paper, either typed or handwritten, could be double-sided); no other support materials are allowed at tests. Plagiarism includes reproducing the words of others without both the use of quotation marks and citation. Students are reminded of the obligations and expectations associated with the Georgia Tech Academic Honor Code and Student Code of Conduct, available online at www.honor.gatech.edu.

**Learning Accommodations**

If needed, we will make classroom accommodations for students with documented disabilities. These accommodations must be arranged in advance and in accordance with the ADAPTS office (<http://www.adapts.gatech.edu>).

**Course materials**

The principal resource for the course are the slides used in lecture. These are provided as PDFs at the time of lecture. In addition the following text is considered required as a resource:

* Forsyth & Ponce, *Computer Vision: A Modern Approach (2nd Edition)*, Prentice Hall, 2011, ISBN-10: 013608592X, ISBN-13: 978-0136085928 (on [Amazon](http://www.amazon.com/Computer-Vision-Modern-Approach-Edition/dp/013608592X/ref=sr_1_1?ie=UTF8&qid=1345401093&sr=8-1&keywords=forsyth+ponce" \t "_blank))

For graduate students there is another book of interest by Rick Szeliski. That book is a great reference but really more of a modern review of the state of the art methods – more appropriate for the graduate section:

* Richard Szeliski, *Computer Vision: Algorithms and Applications* (book [Web site](http://szeliski.org/Book/" \t "_blank) )

#### **Required Software:**

Problem sets will be done in Matlab or Python with OpenCV.

**Syllabus**

A brief outline of units of the most recently offered section of the class is given below, grouped into 10 parts. Some topics may be added or removed as connects to ongoing developments in the field and in industry.

**1 Introduction**

* 1A Introduction

**2 Image Processing for Computer Vision**

* 2A Linear image processing
* 2B Model fitting
* 2C Frequency domain analysis

**3 Camera Models and Views**

* 3A Camera models
* 3B Stereo geometry
* 3C Camera calibration
* 3D Multiple views

**4 Image Features**

* 4A Feature detection
* 4B Feature descriptors
* 4C Model fitting

**5 Lighting**

* 5A Photometry
* 5B Lightness
* 5C Shape from shading

**6 Image Motion**

* 6A Overview
* 6B Optical flow

**7 Tracking**

* 7A Introduction to tracking
* 7B Parametric models
* 7C Non-parametric models
* 7D Tracking considerations

**8 Classification and Recognition**

* 8A Introduction to recognition
* 8B Classification: Generative models
* 8C Classification: Discriminative models
* 8D Action recognition

**9 Useful Methods**

* 9A Color spaces and segmentation
* 9B Binary morphology
* 9C 3D perception

**10 Human Visual System**

* 10A The retina
* 10B Vision in the brain