EE150
Homework 2
Due: Thurs, April 14 (Before Class)

1 Find a Project

Several project descriptions have been posted. Please look at these by Tues, April 12. Remember, the main portion of this class will be spent implementing and working on your project. Please spend several hours looking over the papers and the project descriptions to identify projects that interest you.

2 Eigenfaces vs. Fisherfaces

Read the Paper Belhumeur et. al, Eigenfacs vs Fisherfaces: 'Recognition Using Class Specific Linear Projection' located on the website. It presents a different view on both PCA and Fisher from that which was presented in class. Both of these techniques occur frequently within vision and pattern classification applications.

1. General Fisher vs. PCA questions: Describe, in your own words, the subspaces which Fisher and PCA project to. Fisher is a 'supervised' technique. It requires an extra piece of information associated with each data-point which PCA does not necessarily require. What is this extra information?

2. Why does the image for a face lie in a 3D linear subspace if we consider it a Lambertian surface?

3. Why do the authors achieve better performance when they ignore the first 3 principle components? What do these components represent? Would you expect the same improvement if you removed the first 3 Fisher components?

3 Male and Female Faces

In class we suggested the challenge of achieving the highest performance on Male vs Female faces (the data-set is online via the class website). You have been given code which replicates the experiments which were presented in class. You have also been given the data-sets which were used. Try to obtain the best possible classification results!

1. What is over-fitting? When do we observe over-fitting? Assume we are fitting a high dimensional parametric model to our data. Draw a plot of how the test and train performance would vary as a function of the number of training examples. Explain your plot. Describe (or show) how these curves would change if the amount of over-fitting were increased or decreased.

2. Try at least 3 different techniques which you feel are reasonable to obtain the best possible classification performance on Males vs Females. Use portions of the code from class if you wish as a framework for your own code.

3. These approaches (Fisher and PCA) are often termed 'global' approaches to image recognition and classification. They are not well suited for handling occlusion and/or background clutter. Test how well your system performs if you occlude part of the image. Specifically, create a plot indicating performance as a function of the amount of image occluded. Only occlude images in the *test* set. It is up to you how to generate the occlusion, but it should vary from image to image within the test set. Show us some of the occlusions you generate.