CS 479/679 Pattern Recognition
Spring 2014 – Prof. Bebis
Programming Assignment 4 - Due: 5/14/2014 – Extended!

In this assignment, you will experiment with two different classifiers for gender classification: SVMs and the Bayesian classifier.

Data Set and experiments: The dataset to be used in your experiments contains 400 frontal images from 400 distinct people, representing different races, with different facial expressions, and under different lighting conditions. The 400 images have been equally divided between males and females. Histogram equalization has been applied to each normalized image to account for different lighting conditions. The data, which is available from the course's webpage, comes into two sizes: 16x20 and48x60; you should only use the largest size (i.e., 48x60) in your experiments. For each classifier, you need to report the **average** error rate using a three-fold cross-validation procedure. To do this, we have randomly divided the database three times as follows:

Fold 1: Training (69M, 65F), Validation (73M, 60F), Test (58M, 75F)
Fold 2: Training (62M, 72F), Validation (58M, 75F), Test (80M, 53F)
Fold 3: Training (71M, 63F), Validation (67M, 66F), Test (62M, 71F)

The validation set is typically use for parameter optimization. You will not need to optimize any parameters here so combine the validation with the test sets for testing. For each image, we have pre-computed its eigen-face representation (see **README** file for more information). You should be training/testing your classifiers using the first 30 eigen-features (i.e., corresponding to the top 30 eignevectors).

Experiment 1: Apply Support Vector Machines (SVMs) for gender classification. You will be using the LibSVM implementation. Experiment with different kernels (polynomial and RBF) and C values. Report your best results.

Experiment 2: For comparison purposes, apply the Bayes classifier for the same problem. Model the male and female classes using a Gaussian distribution and use ML estimation to estimate the parameters for each class. Use equal prior probabilities (e.g., $P(\omega_1)= P(\omega_2)$). Compare your results with those obtained using SVMs.

FORMAT FOR SUBMITTING PROJECT REPORTS

Page 1: Cover Page. The cover page should contain Project title, Project number, Course number, Student's name, Date due, and Date handed in.

Page 2: Technical discussion. This section should include the techniques used and the principal equations (if any) implemented.

Page 3 (or 4): Discussion of results. A discussion of results should include major findings in terms of the project objectives, and make clear reference to any figures generated.

Appendix: Program listings. Includes listings of all programs written by the student. Standard routines and other material obtained from other sources should be acknowledged by name, but their listings should not be included