ECE 532
Crystallography Project — Final Goals and Objectives

Due Sunday, May 1, 2005

The final project consists of two *challenge* problems.

Challenge 1. Two-Class Problem

$$\begin{aligned} \text{Class } 0 &= \text{crystal data types } 1,2,3 \\ \text{Class } 1 &= \text{crystal data types } 7,8,9 \end{aligned}$$

Design a classifier to distinguish between Class 0 and 1 as defined above. That is, the input to the classifier will be an image from the datasets 1, 2, 3, 7, 8 or 9, and you should output a label indicating a prediction of whether it is a type 1, 2, or 3 image or a type 7, 8, or 9 image. In your design, keep in mind that misclassifying Class 1 data is more undesirable than misclassifying Class 0 data.

Challenge 2. Three-Class Problem

$$\begin{aligned} \text{Class } 0 &= \text{crystal data types } 1,2,3 \\ \text{Class } 1 &= \text{crystal data types } 7,8,9 \\ \text{Class } 2 &= \text{crystal data type } 5 \end{aligned}$$

Design a classifier to distinguish between Class 0, 1, and 2 as defined above. Again, keep in mind that misclassifying Class 1 data is more undesirable than misclassifying Class 0 or Class 2 data.

The deliverables for the final project are

1. A detailed on-line report documenting theory, methods and results.

2. User-friendly Matlab functions for the two classification challenges above. The functions should be called `classifier2class.m` and `classifier3class.m`. The I/0 specifications for the functions are as follows:

```
label = classifierXclass(image)
        Input:  crystal image in standard JPEG format
        Output: label assigned to image
```

In addition to the two Matlab functions, each team should develop a script that automatically applies the classifiers to all images in the corresponding datasets and outputs the total number of errors (mistakes) made for each class.

The final report and Matlab code must be completed and posted on-line at your team website by 11:59pm, Sunday, May 1.

(continued on next page)

Evaluation and Best Classifier Awards: To evaluate the performance of each team's classifiers, I have a separate set of "test" images, which have been labeled by George Phillip's lab. I will make these images available after the final classifiers are submitted on-line. Then each team will be responsible for evaluating the performance of the classifiers from two other teams, according to the following assignments:

Teams 1 & 2 will evaluate Teams 3 & 4

Teams 3 & 4 will evaluate Teams 5 & 6

Teams 5 & 6 will evaluate Teams 1 & 2

Since each team's classifiers will be independently tested by two other teams, this will ensure a correct and unbiased evaluation. The evaluations are due by noon on Wednesday, May 4, and the results (error totals for each class) should be emailed to me.

The team or teams whose classifiers perform best will receive an award for their designs. The criterion to be employed is a weighted error measure. Misclassifications of Class 1 images will be weighted by a factor of 10, relative to other misclassifications (i.e., one Class 1 error is as costly as 10 errors on either of the other classes). The resulting error criteria for determination of the best classifier are:

Challenge 1: error $= 10 \times$ Class 1 errors + Class 0 errors

Challenge 2: error $= 10 \times$ Class 1 errors + Class 0 errors + Class 2 errors

In each case, the classifier that produces the minimum error will be declared the best.

The team whose classifier performs best on Challenge 1 will receive a cash prize of $120. The prize is to be split evenly among the winning team members. The prize will be split accordingly in the case of a tie.

Each member of the team whose classifier performs best on Challenge 2 will receive an extra 25 bonus points on their final grade. These bonus points are equivalent to an additional 25 points on their midterm exam. In the case of a tie, members from each winning team will receive the 25 bonus points.

The winners will be announced at our final class meeting on Thursday, May 5.