198:443 Digital Imaging and Multimedia
Spring 2008

Assignment 5
Due Date – in class: Thursday *May 1st.* There will be no further extensions.

The assignment builds over assignment 4. The goal is to enhance the hand written character recognition system you developed in assignment 4. You can use any approach you think is sensible to achieve good results. Here are some ideas for enhancing:

- Automate the threshold selection process: in assignment 4 the threshold was picked manually.
- Use morphology might be useful if you have fragmented characters
- Investigate different shape and region descriptors to extract invariant features to be used in the recognition. E.g., more invariant moments, shape profiles, contour descriptors, …
- Use a better classifier: in assignment 4, we used the nearest neighbor from the features' database for recognition. Instead you can find the k-nearest neighbors (k is small number 3, 5, 7) and do a majority vote.

Please note that 40% of the grade of this assignment will be on the efforts you do in experimenting with sensible ideas, as will be explained later.

Evaluation:
You can use the test.jpg image to evaluate your progress. You should NOT use this image for training your system. It is just for validating your work. Your final code will be evaluated with a similar image, which will not be given to you.

You can use ImageJ or Matlab for your implementation. Useful implementations of ImageJ plugins can be found at:
http://rsb.info.nih.gov/ij/developer/source/

Grading:

40% of the grade will be on submitting a working code that gives results better than 65% on test.jpg. Remember, you should NOT use this image for training your system.

40% of the grade will be on the ideas you have tried and experiments you carried out

20% of the grade will be competitive based on the results your code scores on test images.

Submission instruction:

I – Report: You need to prepare a report containing all things you have tried and what did you find out. Even if you tried some idea and it doesn't work, you need to report it and report it's result. For any experiment you try please provide

a. A description of what your idea
b. The code you wrote to implement it
c. The results on the training data, as was done in part III of assignment 4, in terms of a confusion matrix.
d. The results on "test.jpg" in terms of a confusion matrix.
e. Your analysis of the result: what did you learn from this experiment? If you get bad results, why is that? if you get better results, from where does the improvement come?

II – Final Code: You need to submit your final version of the code for so that your TA can run it on some other images to evaluate performance. Create an easy calling/driver function that takes an image the input, runs your recognition algorithm on it, and produces results in a particular format (described below). A complete specification of this calling/driver function follows.

Name of the function: RunRecog(filename)

Input parameter: 'filename' is the name of the image file to be used as an input.

Output: After applying your recognition algorithm on file indicated by 'filename', your program should produce two outputs:

a. A text file named "results.txt" that contains a comma-separated list of the corner coordinates connected components and their labels in the following format:

(min_row, min_col, height, width, the recognized character code).

Use separate lines for different connected components. Arrange this list such that it gives the corner coordinates of the bounding boxes in a row major order. That is, start with first row, write the coordinates of the boxes from left to right until all the boxes are recorded; then go to second row and so on.

b. An image named "results.jpg" containing the input image with bounding boxes overlaid on the characters. Also, write the label (that you used to encode the character) of each character beside its bounding box.

The images for this assignment are available at
ftp://ftp.cs.rutgers.edu/pub/elgammal/CS443/HW4