ECE 493/693 Introduction to Computer Vision

Lab #2 – optical character recognition

In this project each student must implement a matched filter (normalized cross-correlation) to recognize letters in an image of text.

The following materials are supplied at the web site and should be used:

<table>
<thead>
<tr>
<th>Image Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>parenthood.ppm</td>
<td>input image</td>
</tr>
<tr>
<td>parenthood_e_template.ppm</td>
<td>template image</td>
</tr>
<tr>
<td>parenthood_gt.txt</td>
<td>ground truth</td>
</tr>
</tbody>
</table>

The ground truth file lists all the letters and image pixel coordinates of text in the image. The pixel coordinates are for the center point of each letter.

The program should perform the following steps:
1. Read the input image, template image, and ground truth files.
2. Calculate the matched-spatial filter (MSF) image (remember, this is not 8-bits).
3. Normalize the MSF image to 8-bits.
4. Loop through the following steps for a range of T:
   a. Threshold at T the normalized MSF image to create a binary image.
   b. Categorize and count the detected letters as TP and FP.
   c. Output the total TP and FP for each T.

Using any desired program, you must create an ROC curve from the program output.

For purposes of evaluating true positives (TP) and false positives (FP), use the following definitions. For each letter, consider a 9x15 window centered at its ground truth location. If any pixel in the MSF image is above the given threshold within this window, then the letter is considered “detected”. It does not matter how many pixels in the MSF image are above the threshold, the letter can only be detected once. If the detected letter is an “e” then it is a TP, otherwise it is a FP.

You must write a brief report that includes the code and the ROC curve. Show your MSF image. Identify the optimal threshold and TP and FP values at that threshold.

The report is due date is given at the web site. Reports will be collected in class.