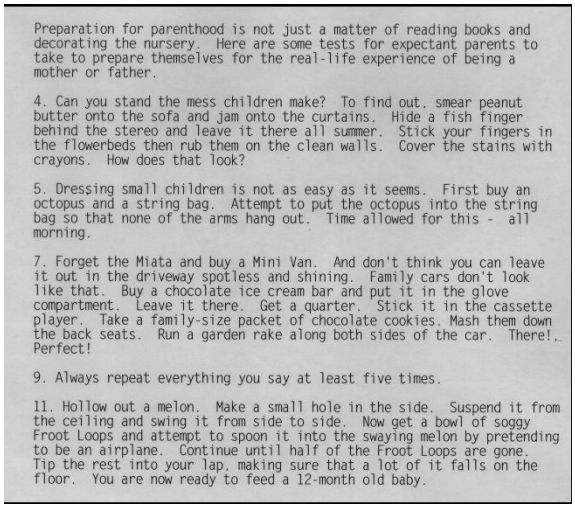
ECE 6310: Introduction to Computer Vision - Lab 2 Report
Optical Character Recognition

Objective: To implement a matched filter (normalized cross correlation) to recognize letters in an image of text, create an ROC curve from the program output.

Implementation and steps:

1. Original input image and template image were read and checked if read correctly. Template image contained a sample of letter ‘e’ (lowercase) from the original parenthood.ppm image. The template image size was 9*15.



Preparation for parenthood is not just a matter of reading books and decorating the nursery. Here are some tests for expectant parents to take to prepare themselves for the real-life experience of being a mother or father.

4. Can you stand the mess children make? To find out, smear peanut butter onto the sofa and jam onto the curtains. Hide a fish finger behind the stereo and leave it there all summer. Stick your fingers in the flowerbeds then rub them on the clean walls. Cover the stains with crayons. How does that look?

5. Dressing small children is not as easy as it seems. First buy an octopus and a string bag. Attempt to put the octopus into the string bag so that none of the arms hang out. Time allowed for this - all morning.

7. Forget the Miata and buy a Mini Van. And don't think you can leave it out in the driveway spotless and shining. Family cars don't look like that. Buy a chocolate ice cream bar and put it in the glove compartment. Leave it there. Get a quarter. Stick it in the cassette player. Take a family-size packet of chocolate cookies. Mash them down the back seats. Run a garden rake along both sides of the car. There!. Perfect!

9. Always repeat everything you say at least five times.

11. Hollow out a melon. Make a small hole in the side. Suspend it from the ceiling and swing it from side to side. Now get a bowl of soggy Froot Loops and attempt to spoon it into the swaying melon by pretending to be an airplane. Continue until half of the Froot Loops are gone. Tip the rest into your lap, making sure that a lot of it falls on the floor. You are now ready to feed a 12-month old baby.

*Fig. 1 parenthood.ppm (649*657)*

Fig. 2 parenthood_e_template.ppm

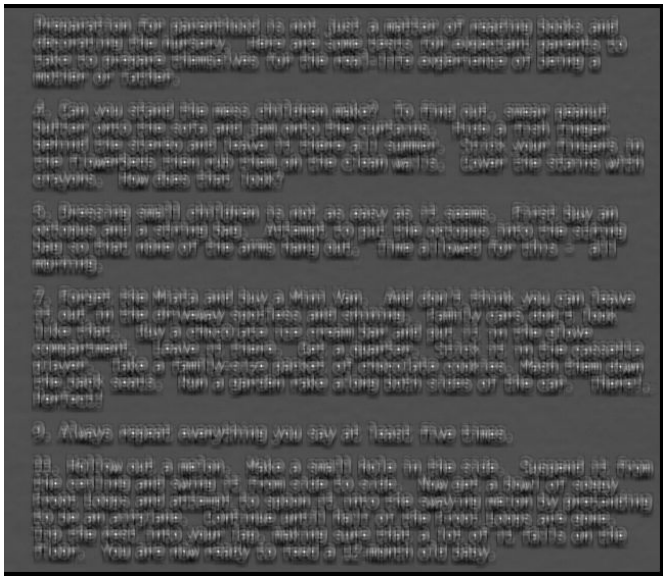
2. Template image was mean centered by subtracting mean value of 9*15 template from each of the pixels in template image. Mean was found to be 65.39.

3. MSF image was built using the cross correlation (multiplication of original image with mean centered image), skipping the border points. Code snippet is given below:

```
for (r2=-7;r2<=7;r2++)
        for(c2=-4;c2<=4;c2++)
                msf_sum+= (ori_image[(r+r2)*COLS+(c+c2)]) *
                        (mean_centered_template[(r2+7)*COLS1+(c2+4)]);
MSF[r*COLS+c]=msf_sum;
```

This was not a normalized MSF image, re-scaling to 0-255 range was required.

4. After normalizing, the final MSF image was obtained as follows:



In this image, the highest pixel value was assigned to those points where template matched more accurately. (Most of these were the center locations of letter ‘e’ with the brightest pixels)

Fig. 3 Normalized MSF image

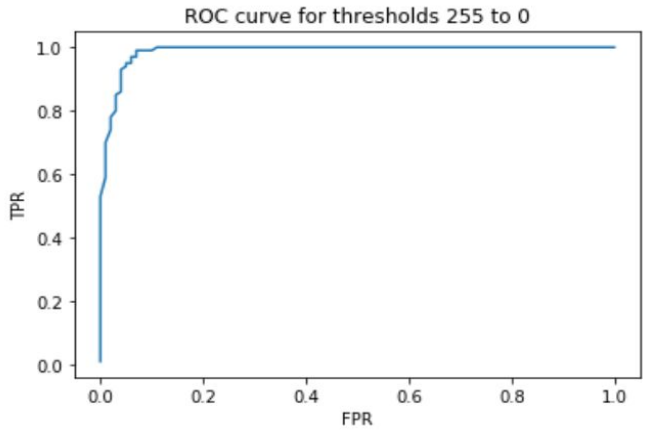
5. Ground truth file (parenthood_gt.txt) was read and row, column coordinates for centers of letters were stored in 2 arrays.

6. To check whether the letter is correctly detected or not, the following steps were performed, looping through 255 to 0 threshold values:

i) Read and checked MSF image
ii) Binary MSF image was created by thresholding the normalized MSF image (values > threshold were assigned a value: 255, and 0 was assigned to all other values)
iii) By looping through ground truth center locations, checked a 9 x 15 pixel area centered at each location and took a sum of pixels in this region. If any pixel in the binary msf image was found greater than the threshold (sum >=255), considered the letter “detected”. If none of the pixels in the 9 x 15 area were greater than the threshold, considered the letter “not detected”.
iv) To check if detections were actually true positive or false positive, the ground truth value was checked with each of the detected value, if the letter was ‘e’, it was a case of true positive, else, it was a false positive
v) Similarly, if the letter was not detected and it was ‘e’ from ground truth, it was a false negative, if program output and ground truth both had a value other than ‘e’, it was a true negative.
vi) The detections and values of TP(True Positive), TN(True Negative), FP(False Positive), FN(False Negative), TPR (True positive rate), FPR(false positive rate) were calculated for each threshold.

7. ROC Curve with TPR and FPR was plotted to get the optimal threshold value.

Results and evaluation:



Using the TP, FP, TN, FN values,
TPR and FPR are calculated as follows,

TPR = TP / (TP+FN)
FPR = FP / (FP+TN)

Sensitivity= TPR
Specificity= 1 - FPR

Fig 4. ROC Curve for threshold values from 255 to 0

Based on ROC curve above, optimal value of threshold would be somewhere at the top left, where TP is close to maximum (all possible detections) and FP is also a reasonable value, not the maximum one. Usually, optimal value of threshold is chosen at the knee of the curve, that establishes a balance between TP and FP values, and still closer to 1 compared to other points on the ROC curve. Ideal case is FP 0 and TP 1.

However, in some cases, where detecting all the positive values as positive correctly is important, even if there is a large number of false positives, we need to choose a value with Highest TP. In that case, the rightmost value on knee of the curve will be chosen. TPR and FPR will be high in this case.

In another case, if we want to detect the negative values correctly, to avoid large number of false positives, we need to choose a value with low FP, even if we compromise on the TP value. Both TPR and FPR will be lesser in this case compared to the first one.

For example, if we have 2 fish species and we want to detect which one is poisonous, detecting the poisonous one correctly as poisonous is more important than not detecting a non-poisonous as poisonous. In this case, we want the lowest False Negatives, high TP, irrespective of FP.

Overall, the optimal value selection depends on the end application and how critical it is to get a detection correct for positive case or negative case.

Optimal Threshold Selection:

Since, in this template matching example, it is not that critical to have either very high TPR or very low FPR, the optimal value for threshold I have chosen is **T= 208**.

At this threshold, **TP count: 144, FP count : 64, TN count: 1047, FN count: 7.**
TPR: 0.95, FPR: 0.05

Actual count of e in the image was 151, and count for letter not e was 1111.

```
modjoul1@LAPTOP-7FILMC9D:/mnt/c/Users/Netra Inamdar/Documents/fall_19_courses/Intro_to_cv/assignments$ ./lab2
Threshold:208   detected:208    not detected:1054       tp count:144    fp count:64     fn count:7      tn count:1047
TPR:0.9536424   FPR:0.0576058
True e count:151        True Not e count:1111
modjoul1@LAPTOP-7FILMC9D:/mnt/c/Users/Netra Inamdar/Documents/fall_19_courses/Intro_to_cv/assignments$
```

Fig 5. Results for T = 208

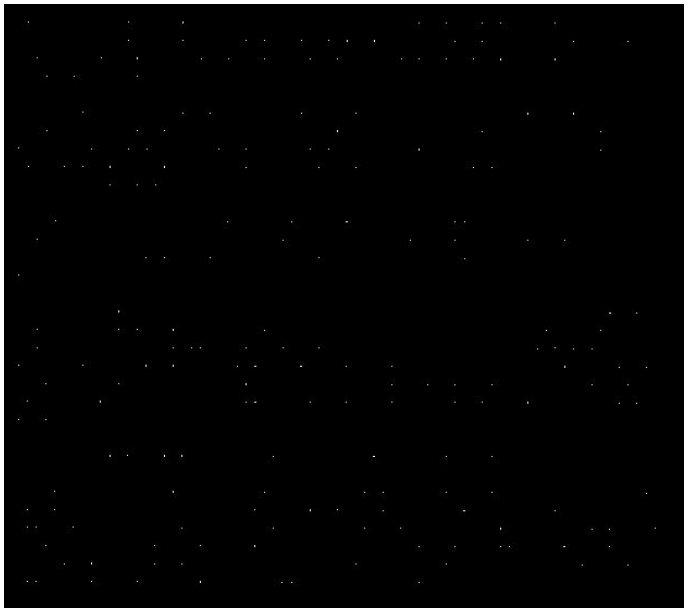


Fig 6. Binary image at Threshold = 208