Image Understanding and Pattern Recognition
Prof. Dr. Thomas M. Breuel

08.07.2009
Dr. Faisal Shafait

Document and Content Analysis (SS 2009)

Exercise Sheet 5

to be submitted by E-Mail to faisal.shafait@dfki.de by: 20.07.2009

Exercise 5.1 (Text Line Segmentation)
A common method for text line segmentation is upper contour segmentation. The upper contour is the y-coordinates of the uppermost edge of a character. Segmentation is performed by looking for local minima in the upper contour and cutting apart characters at those locations. Implement upper contour segmentation. Generate both a color segmentation of the binary input image and a list of isolated character images. You can use the Text Line Segmentation worksheet as an example.

A collection of text line images is present in lines.tgz

Be sure to create non-trivial test cases and examples that demonstrate that your code actually works as expected.

Exercise 5.2 (Weighted Finite State Acceptors)
Please have a look at the worksheet on Weighted Finite State Acceptors. Based on the class defined in that worksheet, implement two functions:

- A $compose(u, v)$ that composes to WFSAs. That is, if (string,weight1)$\in u$ and (string,weight2) $\in v$, the result of compose should contain (string,weight1+weight2).
- A bestpath(u) function that, given a WFSA, computes a (string,weight) pair such that the weight is lower than that of any other string in u.

Be sure to create non-trivial test cases and examples that demonstrate that your code actually works as expected.
