Document Analysis
Exercises

Univ.Prof. Dr. Robert Sablatnig

Schedule   The exercise must be solved in groups of 2 students. The exercise should be written in Matlab (a student license can be bought at the LMZ). Important dates are listed below:

- 07.05.2018 deadline of assignment 1
- 11.06.2018 deadline of assignment 2
- 25.06.2018 final presentation

Abgabe   The assignments must be sent via mail to: davu@cvl.tuwien.ac.at. Each assignment should be zipped (groupXX.zip) and contain the subsequently enumerated items:

- **main.m** the main routine
- **assignmentX.pdf** the report

All files must contain the name of all group members and the group number. The report should explain the assignments and answer all questions. It should be a comprehensive but short (e.g. 6 pages) explanation of the algorithms used, their advantages and drawbacks.

Assignment 1 – Page Detection

Page detection is the task of detecting and segmenting the page outlines of a document within an image. The *ICDAR2015Competition on Smartphone Document Capture and OCR (SmartDoc)* [1] is the first competition for (a) document page detection and (b) Smartphone OCR.

Your task is to implement either one of the participating methods (e.g. Carlinet and Geraud [2]) and compare your implementation results with the published one, or your own methodology which competes with methods presented at the competition. To assess the performance of your implementation or your own methodology the Jaccard Index should be used as measure [1].

Figure 1 shows an example image of the dataset and the detected page and the Ground Truth (GT).
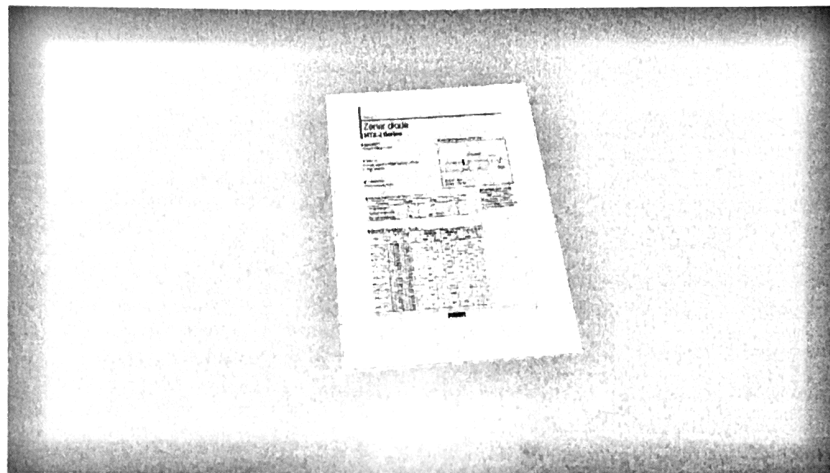

Fig. 1: Example image of dataset *Background 01* with a correct page detection ($JI > 0.99$).

The dataset of [1] is provided at `https://cloud.cvl.tuwien.ac.at/s/7kOg6MS4YH3MTRy` It is not allowed to publish the dataset without the consent of the authors of [1].

Assignment 2 - OCR

Optical Character Recognition (OCR) is needed for the accessibility of digitized documents. Subsequently, a typical OCR workflow is presented:

- Binarize image
- Deskew document (estimate orientation)
- Compute (binary) features for each blob (character)
- Classify features
- Assign labels/label probabilities
- Dictionary correction

Your task is to implement an OCR method. As training material you can create your own images. Evaluate your method using a proper evaluation metric (e.g. edit distance).

Literatur

[1] J. Burie, J. Chazalon, M. Coustaty, S. Eskenazi, M. M. Luqman, M. Mehri, N. Nayef, J. Ogier, S. Prum, and M. Rusinol. Icdar2015 competition on smartphone document capture and ocr (smartdoc). In *13th International Conference on Document Analysis and Recognition (ICDAR)*, pages 1161–1165, Aug 2015.

[2] Edwin Carlinet and Thierry Géraud. MToS: A tree of shapes for multivariate images. *IEEE Transactions on Image Processing*, 24(12):5330–5342, December 2015.