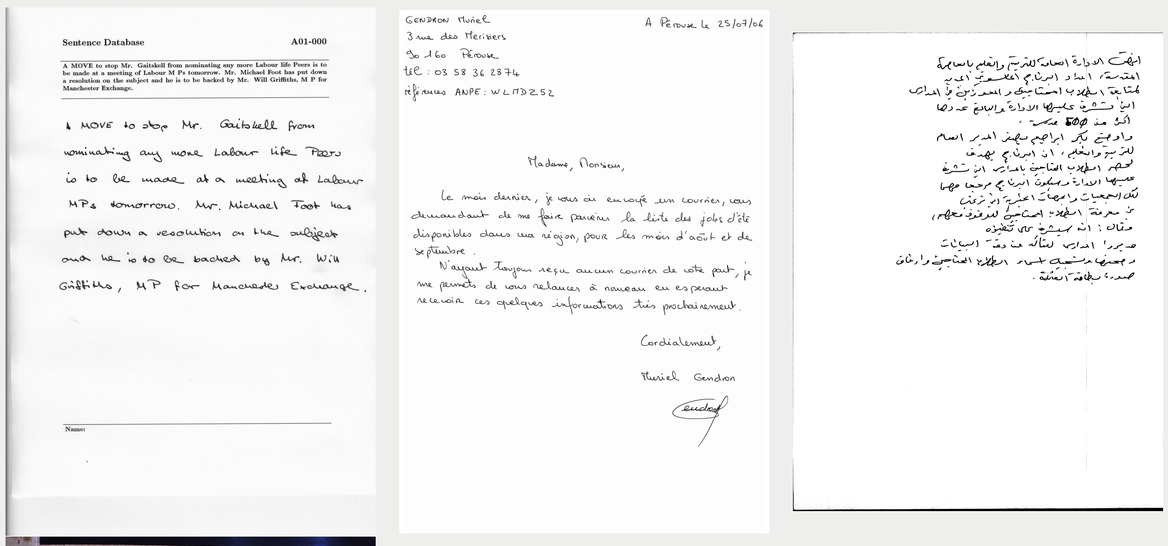
ON THE EVALUATION OF HANDWRITTEN TEXT LINE DETECTION ALGORITHMS

Bastien Moysset and Christopher Kermorvant - A2iA, Paris, France

Motivation

Revisit the evaluation protocol for text line segmentation algorithms : test 4 algorithms, on 4 databases, with 3 metrics.

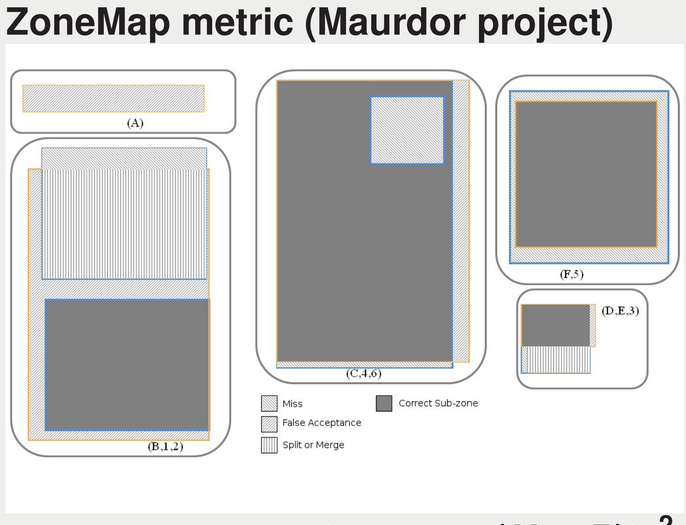
Handwritten Document Databases



IAM : English, 115 pages
RIMES : French, 200 pages
OpenHaRT : Arabic, 845 pages
Numen-RA : Historical Creole French, 665 pages

Evaluation Methods

ZoneMap metric (Maurdor project)



- Measures the error due to merges, splits, false acceptances, misses and unperfect matches separately.
- Hypothesis (H) and reference (R) boxes are grouped according to their strength f :

$$f(R,H) = \left(\frac{\texttt{Surface}(H \cap R)}{\texttt{Surface}(H)}\right)^2 + \left(\frac{\texttt{Surface}(H \cap R)}{\texttt{Surface}(R)}\right)^2$$

Metric of the ICDAR 2009 text line detection competition

- A matching score is computed between each reference box and each hypothesis box.

$$\texttt{matching\_score} = \frac{\texttt{Surface}(H \cap R)}{\texttt{Surface}(H \cup R)}$$

- The number of one to one matches is computed using a threshold on these matching scores.
- Global score is computed according to recognition accuracy (RA) and detection rate (DR).

$$\texttt{error} = 1 - \frac{2 * DR * RA}{DR + RA}$$

Recognition metric : closest to the final application

- Use of a Handwritten text line recognizer : MDLSTM Recurrent Neural Network.
- Word recognition error rate computed using sclite.

Perspectives

- Improve the geometric metrics to better reflect the impact of the different types of errors on the recognition
- Study the correlation of the ZoneMap metric and the recognition error rate on a larger number of databases

Line Segmentation Algorithms

Projection algorithm

- Simple baseline: horizontal projection of pixel intensity.

Rectangle-based filtering inspired from Shi(2009)

- The image is blurred using median filtering with a rectangular mask.
- Blobs corresponding to lines are extracted by binarisation.

Shredding method inspired from Nicolaou(2009)

- Lines are separated by following ”valleys” (in term of pixel gradient) in both directions.

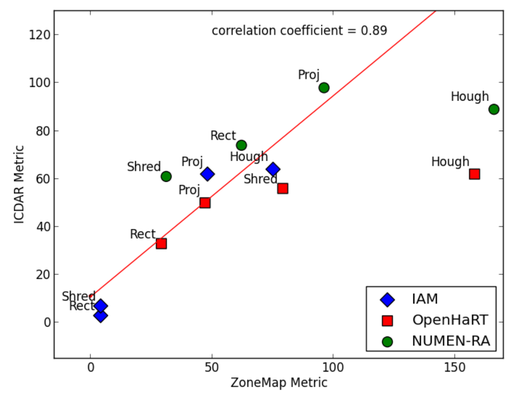
Hough-based method inspired from Louloudis(2009)

- Hough transform is used to find lines using connected components gravity centers as voting points.

- G. Louloudis, B. Gatos, I. Pratikakis, and C. Halatsis, “Text line and word segmentation of handwritten documents,” *Pattern Recognition*, 2009.
- A. Nicolaou and B. Gatos, “Handwritten Text Line Segmentation by Shredding Text into its Lines,” in *ICDAR*, 2009.
- Z. Shi, S. Setlur, and V. Govindaraju, “A Steerable Directional Local Profile Technique for Extraction of Handwritten Arabic Text Lines,” in *ICDAR*, 2009.

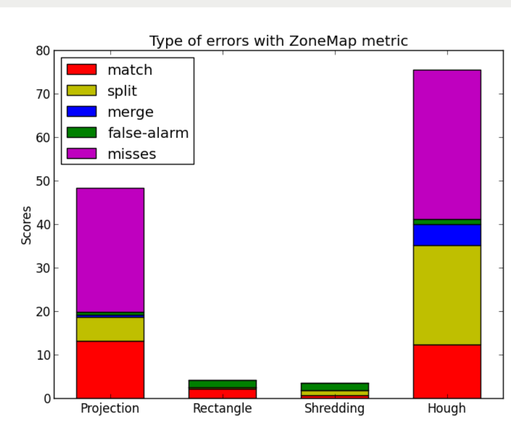
Experimental Results

Geometric metric comparison : ICDAR vs ZoneMap

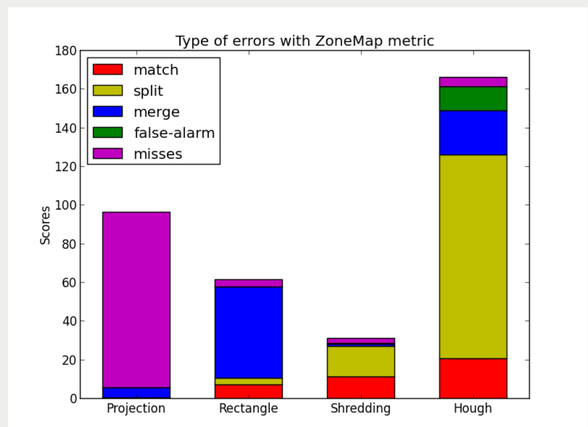


- The ZoneMap metric and the ICDAR metric are highly correlated
- but ZoneMap provides a more detailed error analysis

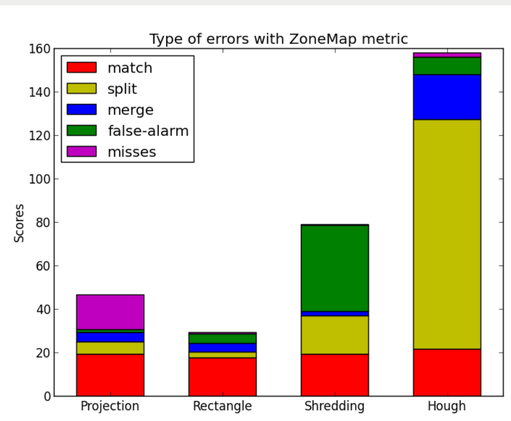
Line segmentation algorithms comparison on different databases



IAM



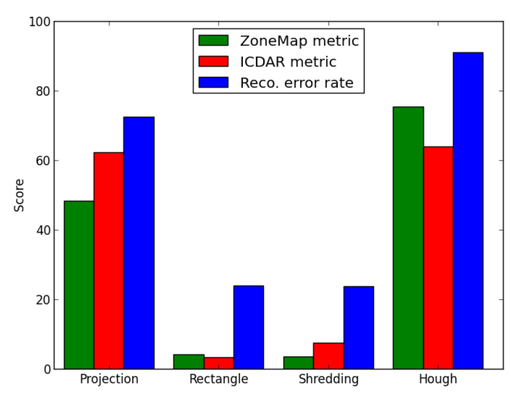
Numen-RA



OpenHaRT2010

- The rectangle based filtering and the shredding methods are generally better than the projection and hough-based methods.
- For each algorithm, the repartition of the different types of error depends on the database

Geometric versus Recognition metric



- The geometric metrics are correlated to the recognition error rate
- but ZoneMap seems to be more correlated

{bm,ck}@a2ia.com