Proposal for a Competition in ICFHR'2010

On-Line Arabic Handwriting Recognition Competition

ADAB: Arabic DAta Base, for on-line recognition of the cursive Arabic handwritten word

Our proposal: On line recognition of the cursive Arabic handwritten words, aims to contribute in the evolution of on line Arabic handwriting recognition research. Since 2009 the freely available (ADAB data base) is used by some groups all over the world to develop on line Arabic handwriting recognition systems. This database was the basis for the last year competition of ICDAR'2009 for systems that are specialized in on line recognition of the cursive Arabic handwritten words. This ICFHR competition uses as a next step the same background of the ADAB database but now with and extended collected data of freely written words. A comparison and discussion of different algorithms and recognition methods should give a push in the field of on line Arabic handwritten word recognition.

Evaluation Process

The object is to run each Arabic handwritten word recognizer (trained on a part of version 2.0 of the ADAB-database) on an already published part of the ADAB-database and on a test set not included in the published part. The recognition results on word level of each system are compared on the basis of correct recognized words, i.e. there correspondent consecutive Numeric Character References (NCR). A dictionary can be used in the recognition process.
A recognizer may return up to 10 candidates for each classification that not only the first ranked result can be used for comparison but also the correct result between the 5 or 10 candidates will be used for comparison.

Running a Recognizer

Note that the evaluation process of all systems will be released in our laboratory REGIM: Group of Research on Intelligent Machines. We run your recognizer (called myrec) by invoking it from the command line as follows:
myrec input.txt output.txt

input.txt

The input file is just a list of relative paths to each *.inkml online trace to be recognized. For example:

```
word/1.inkml
word/2.inkml
...
```

output.txt

The output file should have one line as result for each input file. Each line should show the name of the online trace file that was recognized, followed by the responses (sequence of NCRs code) for that file.
Each response is given as a pair of values: the text, followed by the confidence. In the following example the first line shows that for the file word/1.inkml the recognizer has produced two word hypotheses (بوذر and نبّر) with confidences of 1.0 and 0.3 respectively.

```
word/1.inkml UFE91 UFEEE U0630 U0631 1.0 UFEE7 UFE92 U0651 UFEAE 0.3
word/2.inkml U062A U0627 U0644 U0629 0.7 U0645 U0627 U0631 U062B 0.4 U0634 U064A U0627 U0628 0.2
```

Important dates

Deadline for submission of systems: May 1, 2010

Organizers

Monji Kherallah and **Adel M. Alimi**
Emails: {monji.kherallah, adel.alimi }@ieee.org
REGIM: Research Group on Intelligent Machines.
University of Sfax, National School of Engineers (ENIS) BP "W", Sfax, 3038, Tunisia
Haikal El Abed and **Volker Märgner**
Emails: { el-abed , maergner }@ifn.ing.tu-bs.de
Institute for CommunicationsTechnology (IfN), Braunschweig Technical University, Germany