Interactive tools for transcribing scanned handwritten documents

Duration: 1 year

Team: Loki (Inria Lille – Nord Europe & CRISTAL)

Advisors: Thomas Pietrzak (thomas.pietrzak@univ-lille.fr) & Stéphane Huot (stephane.huot@inria.fr)

Project: GeneaLire

Description

This postdoc is part of a larger project which aims at designing helping tools for transcribing ancient documents. Automatic methods based on statistical analyses require a knowledge database with ground truth information. In our context, such a database does not exist. Worse, a new database is required for every document corpus because the language and writing is different from one another. Therefore the knowledge database must be built by hand, which is a tedious task. We propose to design a transcription tool that helps manual transcription of document in order to build ground truth knowledge databases for automatic transcribing methods. This tool will also leverage automatic methods as the knowledge database is growing to facilitate manual work. With this method, not only the system will gain efficiency at transcribing documents, but users will also gain expertise in this.

The tool will consist in selecting words on scanned documents, and associate tags, such as family name, location, date or occupation for example. The user will also have the possibility to tag the textual transcription, corrected transcription, as well as translations in other languages. The work will consist in designing, developing and evaluating an interactive text selection tool, and an annotation tool for scanned handwritten documents.

Classical selection tools such as free forms and various magic wands are not adapted. Our approach is a selection brush with 4 degrees of freedom: x-y position, brightness threshold and selection size. Our first investigations are promising [1], but we still have to evaluate it with a controlled experiment. The mapping of 4 degrees of freedom is complicated with a keyboard + mouse/touchpad settings. We will investigate a pen + touch setting on a tablet [2].

Tags have several advantages: they are non-hierarchical, and multiple tags can be attached to a single item. The system will feature different visualization techniques to highlight the structure of the document, and the progress of transcription. Thanks to these techniques, the researcher will be able to identify words with glossaries, or with machine learning techniques. These techniques will most likely be more efficient with knowledge about the context.

This work will consist in:

- Studying pixel selection tools, and annotation tools in particular for text manuscripts.
- Defining a design space for pixel selection techniques and document annotation techniques.
- Evaluating interaction techniques.
- Adapt the technique for pen + touch interaction on a tablet.

Depending on the progress of the project, these results may be submitted as a research paper.

Candidate

The ideal candidate holds a PhD in Human-Computer Interaction. The candidate must have experience or a strong interest in software development. Knowledge on machine learning methods is a valuable asset. Creativity, independence, team work as well as great communication skills are valuable advantages. It is not required to speak French.

Working environment

The internship will take place in the Loki team in Lille, France, joint between Inria – Lille Nord Europe and the CRISTAL (UMR CNRS 9189) laboratory of the University of Lille. Supervisors: Thomas Pietrzak and Stéphane Huot.

Bibliography
