REPORT ON THE STATE DOCTORAL EXAMINATION THESIS

Title: Segmentation Strategies for Passage Retrieval in Audio-Visual Documents

Author: Petra Galuščáková

Opponent: Ing. Filip Jurcicek, Ph.D.
Charles University in Prague
UFAL, Faculty of Mathematics and Physics
Malostranske namesti 25
118 00 Prague 1
Czech Republic

The aim of segmentation is to split large documents, either textual or audio visual, into smaller parts so that the resulting segments can be later used by information retrieval techniques to find desired information. Different techniques of segmentation mainly differ in what the type of information (features) they use from the original documents, e.g. lexical, prosodic, or structural features, and the assumption on the size or type of the output segments, e.g. short overlapping window segments or paragraph like structures.

The topic of the presented work is segmentation for passage retrieval in audio-visual documents. The work starts with a brief introduction of the topic. Then, the next chapter explains what passage retrieval is and what it is used for. The third chapter describes numerous types and techniques of text and audio-visual segmentation. In addition, several evaluation metrics are discussed. The fourth chapter presents quite large number of experiments run by the PhD candidate using her baseline segmentation system. In these experiments many settings of standard techniques were carefully evaluated. Final two chapters briefly conclude the work and present a plan for future work.

The thesis is written in concise language and it is easy to read and understand. I would like say the main purpose of the work was accomplished since the PhD candidate written a comprehensive review of techniques used in the field. I do not have more comments on the content since the work does not include any "controversial" content where multiple explanations of either the results or the presented techniques could be possible.

However, there is something could have been improved. The purpose of the work was not only to prepare a review of the current stat-of-the-art techniques but also set the course for the rest of the doctoral study. The conclusion of this work could be improved by providing a proper discussion on the achieved results. In my opinion, what is really important is the researcher's insight into the problem and suggesting why some techniques are better than others. Also the future work would benefit from more detailed plan of what and/or how the presented techniques will/should be modified to achieve better results. In general, I can only agree with the future work plan. However, it is too general so that I am not able to provide any sensible comments.

Overall evaluation: **I recommend this work for defence at the state doctoral examination**. I expect that insight into why some methods work better and more detailed research plan with concrete steps will be presented during the state doctoral examination.

In Prague, 29. 1. 2013.

Filip Jurcicek