

Prof. Dr. Jan Niehues

Adenauerring 2
76131 Karlsruhe, Germany

Phone: +49 721-608-46384
Fax: +49 721-608-44734
Email: jan.niehues@kit.edu
Web: ai4lt.anthropomatik.kit.edu

Official in charge:

Our reference:

Date: 29.05.2024

Re: Review of Doctoral thesis Mr. Dominik Macháček

This is my review of the doctoral thesis “Multi-Source Simultaneous Speech Translation” by Mr. Dominik Macháček.

Content

The thesis presented by Mr. Macháček “Multi-Source Simultaneous Speech Translation” is 205 pages long and well-structured in 9 chapters. In addition, the author provides an extensive bibliography. The structure of the thesis guides the reader from the specified task to a detailed analysis of the individual challenges in the task and finally to different proposed solutions.

The thesis starts with an abstract motivating the work as well as summarizing the main contributions of his work.

The first chapter gives an introduction to the topic of the thesis. The author provides a very clear motivation of his work and points out the main contributions of the work. The impact of the work is highlighted through his impressive number of scientific contributions. Finally, an outline of the thesis is provided.

In the second chapter (“Motivation”, p9.), he investigates the context in which the work happens. It should be positively noted that an extensive literature research has been conducted and especially also important work from related fields (e.g. Interpretation) has been included. Furthermore, a detailed description of possible impacts and risks is given.

In the next chapter (“Focus”), the author gives a detailed and extensive description of the task and all its subtasks. This chapter provides the reader with all necessary background knowledge to understand the research conducted within this thesis.

Starting from chapter three (“Direct ST Quality and Efficiency”, 53 pages), Mr. Macháček presents his actual contributions. The first important work of the thesis is the creation of a new data resource, the “ESIC – European Parliament Simultaneous Interpreting Corpus”. The author gives a very detailed description of the creation of the resource and all design decisions allowing a reproduction of his findings. The created corpus will have a long-standing impact in the research community as a valuable resource for many different research questions.

Based on the resource described in the last chapter, in chapter 5 (“Interpreting Analysis”), a detailed analysis of the challenges of interpretation is carried out. Again, the author starts with an extensive review of the background research and then performs a detailed analysis of how humans perform interpretation with respect to complexity, content preservation and latency. It is very nice to see that the author does not just try a new approach, but first analyzes how humans are performing interpretations and uses these findings to develop AI solutions.

In Chapter 6, a multi-source simultaneous speech translation system is developed and evaluated under restricted conditions. In addition to proposing several interesting solutions to perform multi-source speech translation, the author performs a detailed analysis of possible strengths and weaknesses and investigates many aspects that could influence the quality.

Since the evaluation of the proposed idea is challenging and no existing evaluation protocol is available, the author developed a dedicated solution to evaluate different translation and interpretation solutions for live interpretation (“Continuous rating”). The evaluation is analyzed in detail and the author presents interesting findings, especially by analyzing both human and AI generated interpretations.

In Chapter 8 (“Multi-sourcing in reality”) the author investigates the usage of multi-source live translation in realistic conditions and identifies several challenges. The author gives a very detailed description of a technical solution for low-latency, state-of-the-art automatic speech recognition. Furthermore, it performs an analysis under realistic conditions. However, in this chapter some additional techniques to address the identified solutions would have been nice.

Finally, the thesis is concluded by chapter 9 with a conclusion. In this the author highlights the number of contributions he has made to the field.

Contributions and assessment

During his doctoral studies, Mr. Macháček has contributed substantially to the research community in speech translation. This is documented by numerous publications at well-established conferences as well as the valuable creation of new resources, both data sets and software repositories.


Mr. Macháček handed in a well-structured thesis that highlights his contributions to the challenges. It is impressive that he thereby addresses the whole process of building dedicated AI solutions for a new task. He first gives a very broad, interdisciplinary overview on state-of-the-art and related work. Secondly, he created a new data set to evaluate a new solution. Thirdly, he performs a detailed analysis of the challenges and develops well-motivated approaches to the task. Finally, he performs an extensive evaluation of the approach and points out several future research questions.

It should be especially highlighted that Mr. Macháček did not perform this on a well-established task, but he addresses practical solutions for a new task. As pointed out in the thesis, this comes with many additional practical issues that arise. He analyzed these challenges and described well-motivated solutions. Therefore, the findings of this thesis are not only valuable to other researchers on speech translation, but are also very helpful when investigating new tasks in related fields.

Recommendation

To conclude, the thesis describes solid and significant work and establishes Mr. Macháček as a researcher in the field of speech translation. Based on his significant improvements and detailed analysis of specific challenges, Mr. Macháček has published a significant amount of work in top-tier conferences of the field. In summary, I therefore suggest that Mr. Macháček's thesis should be accepted.

Yours sincerely,



Prof. Jan Niehues