

The logo for DİRAC features the letters 'D', 'R', 'A', and 'C' in a bold, blue, sans-serif font. The letter 'I' is replaced by a stylized purple line graph with three peaks of varying heights. Above the first peak of the graph is a small circle containing the Greek letter sigma ( $\sigma$ ), with an arrow pointing to the right towards a second circle containing the letter 'O'.

# DİRAC

**Detection and Identification of Rare Audiovisual Cues**

## **Visualization for a demonstration of OOV detection**

Confidence measures and classifying techniques are widely used for the recognition error detection task in LVCSR (Large Vocabulary Continuous Speech Recognition). But in many recognition scenarios the amount of words not included in the dictionary (e.g. real names, neologisms) lead to so-called OOV (Out Of Vocabulary) errors which increase the WER (Word Error Rate) even more.

We investigate further improvements of an OOV detection task performed by combining strong and weak phone posterior features using neural networks based on [ICASSP08] and the use of phone context.

The reason why OOV words usually decrease overall performance so badly is that LVCSR systems commonly operate on N-Gram based Language Models. Those spread the error due to their contextual nature to previous and following words. Therefore a well-performing OOV detection seems promising for helping to improve existing LVCSR systems.

One approach to detect OOV words can be done by watching the distribution of frame-by-frame phone posteriors [ICASSP08].

## **REFERENCES**

[ICASSP08] Burget, L. et al: Brno University of Technology, accepted to ICASSP 2008: “Combination of Strongly and Weakly Constrained Recognizers for reliable detection of OOVs”, Las Vegas, NV, USA

[Wessel01] F. Wessel, R. Schlüter, K. Macherey and H. Ney: “Confidence measures for large vocabulary continuous speech recognition”, IEEE Trans. Speech and Audio Processing, vol. 9, no. 3, pp. 288–298, Aachen, DE, 2001.

Location: Brno University of Technology, Faculty of Information Technology, Brno, Czech Republic