ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

TIK
Technische Informatik und Kommunikationsnetze
Computer Engineering and Networks

Prof. L. Thiele

Semester or master thesis

Detecting Strong Prosodic Events

Generally, an automatic speech-to-speech translation system (in German: elektronischer Dolmetscher) consists of three subsystem: (1) a speech recognizer that transforms the input speech into a sequence of word, (2) a machine translation that translates the words from language L1 into language L2, and (3) a text-to-speech synthesizer that transforms the L2 word sequence into audible speech.

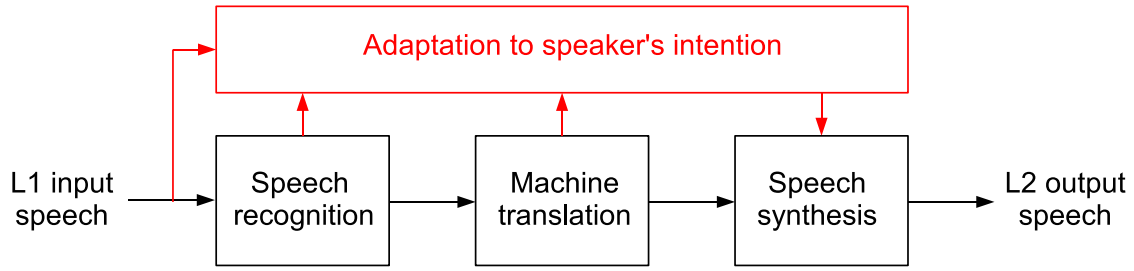

Figure 1: Augmented speech-to-speech translation system

In the framework of a new research project, such an S2ST system is to be improved as shown by the red extension in Figure 1. In order to better represent the speaker's intention in the translated output speech, not only a sequence of L2 words will be provided to the speech synthesis, but additional information such as if the speaker has put special emphasis on some words or has pronounced a part of the utterance in a particular manner.

The underlying intention of the speaker can hardly be detected from the input speech. But the speaker's intention is reflected in the prosody of the speech (melody, rhythm and loudness of the speech). Therefore, particularly strong intentions manifest in salient prosodic events, i.e. the prosody deviates considerably from the normal case.

The aim of this work is to develop and investigate methods, mainly based on statistical models, that allow to detect such salient prosodic events in speech signals. Using information from the speech recognition and the machine translation, the detected prosodic events can be assigned to the corresponding L2 words and the output speech is synthesized accordingly.

The work can be done in Matlab.

The work is suited as a semester or a master thesis for one or two students.

If you are interested in this topic or you want to know more about the work, please contact:

Beat Pfister pfister@tik.ee.ethz.ch, ETZ D97.6

Hui Liang liangh@tik.ee.ethz.ch, ETZ D97.4