Know-How Transfer on Speech Technologies

In July-September 2015, several meetings with potential Users of the AComIn Speech Lab were held. The meetings included demonstrations of the speech processing prototypes, developed within AComIn, and discussions of potential applications in various scenarios of practical importance. The applications can be grouped in three main directions.

Topic 1

Large Vocabulary Continuous Speech Recognition of Bulgarian language for professional dictation transcription services

The task for generating automatic dictation transcriptions is of significant interest for professionals working in the law and medical areas, where large amount of documents have to be written every day. In the technology seminar sessions on this topic a prototype of a LVCSR system for Bulgarian for the law domain has been demonstrated. The prototype demonstrated accuracy over 93% over new law texts dictated. The demonstration included also adaptation to a new speaker (from the invited participants) and evaluation of the perceived accuracy for the new speaker. The discussed topics included the technological background, application areas, accuracy of transcription and implementation issues. It was concluded that the demonstrated technology is already mature for wider industrial usage.

Session: 17.06.2015
Invited parties:

- CIELA NORMA AD

Eleonora Saykova
Veselin Petrov

Session: 25.06.2015
Invited parties:

- Sofia Municipality

Ivaylo Ivanov

- Civic participation and direct democracy NGO

Svetlina Dragova

Topic 2

Large Vocabulary Continuous Speech Recognition for Bulgarian for automatic subtitle generation during real time TV Broadcasting

Currently there is a rising need for real-time subtitles generation during Television broadcasting. There exist new European-wide directives which require that the national televisions provide this

service. The challenge of this task is (i) the unrestricted language domain, (ii) lack of speaker adaptation and (iii) environment noise. In the Technology seminar sessions on this topic a prototype of a LVCSR system for Bulgarian has been demonstrated. The prototype was demonstrated on specialized domains and on unrestricted domains, with and without speaker adaptation and in the presence of environmental noise. The accuracy was judged for each of the setups. The discussed topics included unsupervised real-time speaker adaptation, adaptive language model and noise robustness. It was concluded that the demonstrated technology has to be further developed for wider industrial usage.

Session: 24.06.2015
Invited parties:
- Convergent Media EOOD

Hristo Tuchev

Session: 30.06.2015
Invited parties:
- Doli Media Studio EOOD

Dobromir Chochov

Topic 3

High-quality, natural Text to Speech synthesis for Bulgarian

The speech synthesis systems are a crucial technological tool for the visually impaired people. With the help of a personal computer or mobile device the technology is enabling the visually disabled people with the possibility to read books, to surf web pages, to use office software like text processing and table calculation, street navigation and many others. This session was devoted to the presentation of technologies for the synthesis of naturally sounding voice using advanced concatenative methods, statistical methods for automatic phonological and prosodical speech description and the integration of the technology into new devices.

Session: 30.09.2015
Invited parties:
- Union of the Blinds in Bulgaria

Vasil Dolapchiev
Stoyan Vassev