Speech segmentation
and quasi-linguistic information in speech processing

Bartosz Ziółko[1], [2] and Magdalena Igras[1]
[1]*AGH University of Science and Technology,*
Department of Computer Science, Electronics and Telecommunications,
al. Mickiewicza 30, Kraków, Poland,
www.dsp.agh.edu.pl, bziolko@agh.edu.pl
[2]*Techmo,*
Ul. Pychowicka 18G/20, Kraków, 30-364, Poland
Techmo.pl

Fuzzy set theory application to the evaluation of audio signals segmentation with high resolution and accuracy [1] and processing of sentence boundaries [2], pauses [3], accents [4] emotions [5] and irony [6] in speech will be presented. The segmentation evaluation is based on comparing automatically found boundaries with ground truth. Hence, the method is more accurate and able to grasp the evaluation problem in a way more similar to the evaluation conducted by a human being. Traditional methods often fail on grading segmentation algorithms, particularly those of relatively similar qualities. We define a fuzzy membership function that measures the degree to which the segments obtained by an automatic procedure are similar to the results of a correct segmentation. To identify a pair of equivalent segments, we set a fuzzy alignment function that points the pairs of segments obtained by an automatic segmentation with the corresponding segments from a correct segmentation.

Beside the mainstreams of speech technology - automatic speech and speaker recognition - a study of paralinguistics has gained an increasing attention recently. It covers all the nonlinguistic or quasi-linguistic information conveyed in speech signal - the form and the content of speech that lies beyond linguistic message. This presentation summarizes the investigation of Polish speech prosody in terms of different functions - starting from paralinguistic aspects like accents or sentence boundaries, ending with non-linguistic information like speaker emotions and attitudes (on the example of irony expression).

References

[1] B. Ziółko, "Fuzzy precision and recall measures for audio signals segmentation", *Fuzzy Sets and Systems,* 2015 vol. 279, pp. 101–111.

[2] M. Igras, B. Ziółko, "Detection of sentence boundaries in Polish based on acoustic cues", *Archives of Acoustics*, in print.

[3] M. Igras, B. Ziółko, "Different types of pauses as a source of information for biometry", In: *Models and analysis of vocal emissions for biomedical applications*: 8th international workshop : Firenze, Italy, 2013, pp. 197–200.

[4] M. Igras, B. Ziółko, 'The role of acoustic features in marking accent and delimiting sentence boundaries in spoken Polish', *Acta Physica Polonica A*, 126(6), 2014, pp. 1246–1257.

[5] M. Ziółko, P. Jaciów, M. Igras, "Combination of Fourier and wavelet transformations for detection of speech emotions", In: *7th International Conference on Human System Interactions (HSI)*, 2014, Costa da Caparica, USA, pp. 49–54.

[6] M. Igras, B. Ziółko, "Akustyczne korelaty intonacji ironicznej" (Eng. Acoustic correlates of ironic intonation), In: *Sens i brzmienie (Eng. Meaning and sound)*, Warszawa: Wydawnictwo UKSW, 2015. pp. 46–48.