#### VKGanalyzer: objective analysis of vocal fold vibrations

#### Medical Imaging & Biomedical Diagnostics

Author: Barbara Zitova

Coauthor(s): Adam Novozamsky1, Ales Zita 1, Michal Sorel 1, Barbara Zitova 1, Jan G. Svec 2, Jitka Vydrova 3; 1 The Czech Academy of Sciences, Institute of Information Theory and Automation, Prague, Czech Republic; 2 Voice Research Lab, Department of Biophysics, Faculty of Sciences, Palacky University, Olomouc, Czech Republic; 3 Voice Centre Prague, Medical Healthcom, Ltd., Czech Republic

Funding Acknowledgment: The work was supported by the Technology Agency of the Czech Republic under the project no TA04010877.

*Introduction*: The vocal folds and their vibrations are participating in the formation of human voice. Any disorder of this organ can potentially affect ability to communicate, important human skill, disruption of which can lead to economical losses in the order of hundreds of billions (USA, 2000) per year, not speaking about the decrease of quality of life. There are many existing methodologies for diagnostics and treatment of communication disorders, however many of them are still subjective in terms of an examiner. In our work we have focused on the objective evaluation of vocal fold vibrations. Their irregularities are often responsible for voice disorders. Videostroboscopy is the base approach for the vibrations examination, but it can fail analyzing non-periodical patterns, which is often the case in pathological voices. Videokymographic cameras, having the same time resolution as high speed cameras (HSC), taken as the first choice, scan just one line of the scene and thus capture much longer time interval than HSC and with better image quality. An objective evaluation of collected data can be improved by means of computer aided analysis, when characteristics of vibration patterns are estimated using digital image processing and feature detection. Results: We have proposed a novel system where videokymographic data are processed using software based on algorithms for scene analysis. Image processing helps to facilitate the examination and then to objectively estimate characteristics of the vibration patterns. During the data acquisition, uncomfortable process for the patient due to the inner position of a laryngoscope, the content of the video stream is automatically evaluated and information-rich parts are preselected, marked, and enhanced in order to suppress unwanted artifacts. The analysis starts with detection of vibration structures, which correspond to vocal chords borders. Then a set of parameters, capturing present phase, amplitude, left-right correspondence etc. is computed. These characteristics are classified into corresponding categories. Numerical representation of vocal fold behavior enables to objectively follow the therapy progress and to quantify the grades of disease, which can increase the insight into the dynamics of the regeneration. Individual steps of parameter computation were compared to the performance of experts (18 sets) to verify robustness of the proposed system.

*Future work*: We plan to include higher level features into the dataset such as mucosal waves to even better characterize vocal fold tissue elasticity. The complexity of new features enforces solution based on convolutional neural network, able to discern the complex disease specific patterns, difficult to be characterized by simpler traditional classifiers. The classification into categories will be enriched by inclusion of audio parameters.

*Conclusions*: We have developed a novel software tool based on videokymography which increases the objectivity of the vocal fold examination and enables to follow the therapy progress. The deterministic approach to parameter evaluation will increase the applicability of this type of examination because it lessens the uncertainty in the tissue categorization.