Proceedings of the World Congress on New Technologies (NewTech 2015)
Barcelona, Spain – July 15 - 17, 2015
Paper No. 117

Discriminating Between Normal and Pathological Voices of Professors by Using Acoustic Analysis

C. J. Pérez, Y. Campos-Roca, F. Calle, J. Martín, A. Moreno
University of Extremadura, Avda. de la Universidad s/n, 10003, Cáceres, Spain
carper@unex.es; ycampos@unex.es; fcalonso@unex.es; jrmartin@unex.es; morenog@unex.es

Extended Abstract

Vocal fold nodules are recognized as an occupational disease in many countries for people who use voice as a working tool (teachers, singers, telephone operators, broadcasters…). Teachers and professors constitute a vulnerable group (Roy et al., 2004).

This work analyzes the feasibility of using acoustic features extracted from voice recordings as a tool to assess the risk of vocal fold nodules in university professors. There is a current research line on the analysis of acoustic signals for clinical diagnosis and evaluation of voice disorders (Baghai-Ravary and Beet, 2013).

In this work, a database composed of 90 professors (59 men and 31 women) has been collected. The 90 subjects underwent a periodic medical examination within the Program for Health Prevention of the University of Extremadura between March and July 2014. A protocol (approved by the Bioethical Committee from the University of Extremadura) was applied to all the individuals. This protocol consisted of a physical examination, a survey and some voice recordings.

Each subject was asked to perform three sustained phonations of /a/ vowel at comfortable pitch and loudness during 5 seconds. Information about each subject's vocal health was obtained from the physical examination and survey. In addition, data from 19 patients suffering nodules and 53 healthy controls were available from the commercial database MEEI (Massachusetts Eye and Ear Infirmary) from KayPentax.

After preprocessing, features were extracted from the speech recordings by using linear and nonlinear signal processing algorithms (Jitters, Shimmers, Harmonics-to-Noise Ratio, Detrended Fluctuation Analysis, Recurrence Period Density Entropy, Pitch Period Entropy…). A feature vector was assigned to each subject. Each vector was composed of average feature values obtained from the three recordings.

By using the features extracted from MEEI as the learning database, different classification algorithms were applied to assign the professors to one of two groups (healthy and suffering nodules). From the 75 different classifiers applied, 34 of them provided accuracies higher than 85% and 5 of them higher than 90%. The best result was obtained by using Support Vector Machine with Minimal Sequential Optimization, providing an accuracy of 92.22%, sensitivity of 92.20% and specificity of 78.70%.

The developed techniques are non-invasive and inexpensive, providing a certain degree of objectivity. In addition, their effectiveness has been demonstrated in the sample under study. Nevertheless, in future works MEEI database will be replaced by an own database, collected under the same conditions than the professors' database. This will enable the use of these techniques in a voice prevention program avoiding negative consequences on professors' health as well as economic costs.

This work has been supported by MAPFRE Foundation through "Ignacio H. de Larramendi" funding programme, Ministerio de Economía y Competitividad (Projects MTM2011-28983-C03-02 and MTM2014-56949-C3-3-R), Gobierno de Extremadura (Projects GR15052 and GR15106), and European Union (European Regional Development Funds).

Baghai-Ravary, L., Beet, S. W. (2013). *Automatic Speech Signal Analysis for Clinical Diagnosis and Assessment of Speech Disorders.* Springer.

Roy, N., Merrill, R. M., Thibeault, S., Parsa, R. A., Gray, S. D., Smith, E. M. (2004). Prevalence of Voice Disorders in Teachers and the General Population. *Journal of Speech, Language, and Hearing Research,* 47, 281-293.