

Detection of Phoneme [S] Mispronunciation for Stigmatism Diagnosis in Adults

Authors : Michal Krecichwost, Zauzanna Miodonska, Pawel Badura

Abstract : The diagnosis of stigmatism is mostly based on the observation of articulatory organs. It is, however, not always possible to precisely observe the vocal apparatus, in particular in the oral cavity of the patient. Speech processing can allow to objectify the therapy and simplify the verification of its progress. In the described study the methodology for classification of incorrectly pronounced phoneme [s] is proposed. The recordings come from adults. They were registered with the speech recorder at the sampling rate of 44.1 kHz and the resolution of 16 bit. The database of pathological and normative speech has been collected for the study including reference assessments provided by the speech therapy experts. Ten adult subjects were asked to simulate a certain type of stigmatism under the speech therapy expert supervision. In the recordings, the analyzed phone [s] was surrounded by vowels, viz: ASA, ESE, ISI, SPA, USU, YSY. Thirteen MFCC (mel-frequency cepstral coefficients) and RMS (root mean square) values are calculated within each frame being a part of the analyzed phoneme. Additionally, 3 fricative formants along with corresponding amplitudes are determined for the entire segment. In order to aggregate the information within the segment, the average value of each MFCC coefficient is calculated. All features of other types are aggregated by means of their 75th percentile. The proposed method of features aggregation reduces the size of the feature vector used in the classification. Binary SVM (support vector machine) classifier is employed at the phoneme recognition stage. The first group consists of pathological phones, while the other of the normative ones. The proposed feature vector yields classification sensitivity and specificity measures above 90% level in case of individual logo phones. The employment of a fricative formants-based information improves the sole-MFCC classification results average of 5 percentage points. The study shows that the employment of specific parameters for the selected phones improves the efficiency of pathology detection referred to the traditional methods of speech signal parameterization.

Keywords : computer-aided pronunciation evaluation, sibilants, stigmatism diagnosis, speech processing

Conference Title : ICSST 2017 : International Conference on Speech Science and Technology

Conference Location : Barcelona, Spain

Conference Dates : February 26-27, 2017