Quality of voice signal after ENT surgery

Wieslaw Wszolek 1,a, Maciej Modrzejewski 2, Ryszard Tadeusiewicz1,b, Tadeusz Wszolek1,a

1 University of Mining and Metallurgy, Kraków, Al. Mickiewicza 30, Poland
   a Department of Mechanics and Vibroacoustic
   b Department of Automatic
2 Collegium Medicum Jagiellonian University, Kraków, ul.sw.Anny 12

Abstract

In the presented article, there are fragments of research done by its authors concerning the evaluation of the quality of a patient’s voice following a surgical procedure in the area surrounding the voice channel.

Keywords: speech analysis, pathological speech, surgical treatment

1. Introduction

When encountering several problems of a medical diagnosis, as well as medical therapy and rehabilitation, the evaluation of pathological speech quality is necessary. In the present work a selection has been shown from studies concerning the application of modified acoustic signal processing methods to the task of evaluation and classification of larynx surgery effects, by using analysis of the speech signal. The standard methods of speech signal processing and classification, used in the speech or voice recognition, totally fail in this problem. Also standard techniques of speech signal parameterisation, like the linear prediction coefficients or cepstral coefficients, cannot satisfactorily describe the pathological speech, because of its specific phonetic and acoustic structure, different from the correct speech, and also because of totally different aim of the recognition study. While in the standard speech recognition studies the goal is to reveal the semantic aspects of the pronounced text, in the tasks of medical diagnosis employing the speech signal analysis the semantic aspects are insignificant. The required signal characteristics should be as sensitive as possible to small deformations of the layers directly related to the voice functioning and the structure of vocal tract. Too often the above mentioned fact is disregarded or is not accounted for properly. In connection with the above functional studies have been carried out for patients after ENT surgery, with special attention being paid to the study of voice quality after the operation. The goal of the work is presentation of voice quality after various surgical treatments, performed in the ENT area. The evaluation has been carried out both for voice quality after larynx surgery (because of tumour and non-tumour recommendations) as well as voice quality after surgical treatment of resonance cavities (nose, paranasal sinussis) and also after standard intubations (under general anesthesia)

2. The material and method

In the present work an attempt has been made to apply the contemporary (state-of-the-art) techniques of acoustic signal processing, making use of the artificial intelligence methods, to the evaluation of speech deformation degree after a surgical treatment. The studies of speech articulation have been carried out for persons treated by ENT surgery. The final acoustic material has been collected from 175 persons divided into two groups:
The reference group (standard group), 25 persons with correct pronunciation

The group of patients (150 persons) treated by the following surgery methods:

- intubation (including prolong intubation) (20 persons)
- resection of the septum of the nose and paranasal sinus surgery (30 persons)
- removing of inflammation polyps of the larynx (25 persons)
- partial surgery of the larynx (75 persons)

The registration of acoustic signal has been carried out in anechoic chamber, where the time dependence of the acoustic pressure signal during the pronunciation has been registered on a magnetic digital tape recorder.

Both the patients and the persons from the reference group pronounced the same text (three times), which consisted of vowels, words containing vowels, and a test sentence. The set of words has been chosen on the acoustic basis (the words contain all the phonemes expected to exhibit the speech deformation, being the result of the operation). Examined probes are the samples of speech signals recorded during reading.

On the basis of multiple previous work, which were aimed at construction of specific diagnostic systems for specific applications (both technological and biological, particularly medical), the authors in the present work have made an attempt of creating a general methodology for diagnostic system construction, which would be capable (according to the needs) of automated evaluation of selected functions of the voice tract. The study was oriented towards the construction of systems based on the analysis of objectively registered acoustic signals of pathologically deformed speech. In construction of methodology of building systems for evaluation of speech quality it was assumed that the required evaluation accuracy will be comparable to the accuracy achieved by a human by means of his/her senses. This is in fact a limiting and simplifying assumption. Objective methods of measurement of acoustic parameters allow the extraction from the signal its more subtle diagnostic features, undetectable for human senses. However at the present stage of development of signal classification methods it proves to be impossible to achieve a fully automated recognition better than the one which is offered by a well-trained specialist (but as a rule the automated system classifies signals better than a person without much practical experience in that field).

During improvement of the described methodology of construction of medical diagnostic systems special attention has been paid to the following problem scopes: preliminary signal processing, extraction of its features and finding optimal methods of automated classification. The topic of preliminary processing will not be considered in the present work, as it was subject of many other works of the same group of authors.

3. Results

Selected results are presented below. The results of partial surgery of the larynx are presented on Figures 1 to 3.
Fig. 1. Degree of dispersion and distribution along the axis marking the distance from the standard in the case of generalized Hamming formula.

Fig. 2. Surgery of glottis, mean distance from the standard, (distance from the standard, type of mathematical formula).
Fig. 3. Surgery of the glottis, distribution on the Jitter-Shimmer scale. The results of removing of inflammation polyps of the larynx are presented on Figures 4 to 6.

Fig. 4. Degree of dispersion and distribution along the axis marking the distance from the standard in the case of generalized Euclides formula – examination before operation.

Fig. 5. Degree of dispersion and distribution along the axis marking the distance from the standard in the case of generalized Euclides formula – examination after operation.
3. Discussion

In the case of an operation on a section of the larynx, based on the results of tests discussed previously, we can decide on the type of operation (hemilaryngectomy - enlarged cordecomy). Based on the automated evaluation of a patient’s voice prior to performing a surgical treatment of resonance cavities (nose, paranasal sinusses), the advisable procedure can also be determined. When removing a non-cancerous tumour in the vocal cords, patients gain (aside from the oncological aspect) a quick recovery of voice quality, depending on the size and area of the region surgically treated and the operation technique. We practically cannot find any available current research to confirm our results and findings. We believe that acoustic analysis is very advantageous in determining the method of the operation on patients who undergo surgical treatment due to throat diseases. The information gained from the acoustic analysis helps both the patient and the surgeon to obtain optimal quality of life following surgical treatment.

4. Conclusion

The presented methodology can be directly applied in the monitoring examinations for patients after larynx surgery as well as the operations of the nose and paranasal sinusses and intubation in general anesthesia. The results of the present study seem to be helpful in proper qualification of patients for the specific types of operations, in order to provide at the same time the maximal therapeutic effect and minimal speech deformation degree after the operation.

References

3. Reroń E., Tadeusiewicz R., Modrzejewski M., Wszołek W.; Application of Neural Networks and Pattern Recognition Methods to the Evaluation of Speech Deformation Degree for Patients Surgically Treated
