POSSIBILITIES OF AUTOMATED ASSESSMENT OF /S/

A. Stráník, R. Čmejla

Department of Circuit Theory, Faculty of Electrical Engineering, CTU in Prague

Abstract

In dysarthria testing, there are several phonorespiration tasks that evaluate the ability to control exhalation stream. Some of these tasks are related to the fricative consonant /s/. The aim of this paper is to introduce the possibilities of automated assessment of two of these tasks: *"Ability to sustain /s/ on exhalation"* and *"Ability to repeat series of /s/"*.

1 Introduction

Dysarthria is a motor speech disorder resulting from neurological injury. It is often characterized by poor articulation, because of dysfunction of a part of the vocal tract (most often tongue, lips or vocal folds) due to wrong innervation.

All over the world there are several dysarthria tests that evaluate the level of dysarthria disorder, for example [1], [3]. Every task in these tests is evaluated subjectively by therapists and thus cannot give comparable results. We introduce automatic evaluation of some of these tasks to objectify it. In every dysarthria test there are a few tasks related to the fricative consonant /s/. This paper focuses on two of them: *"Ability to sustain /s/ on exhalation"* and *"Ability to repeat series of /s/"*.

We use MATLAB to design algorithms which will be capable of making the decision whether the patient passes the task or not and in the future we want to implement these algorithms in C#.NET to evaluate the whole test in real-time.

2 Database

For the training of automatic algorithms, a speech database consisting of two sub databases was used. The first database was recorded in co-operation with the Department of Neurology 1st Faculty of Medicine and General Teaching Hospital. This database contains 13 recordings of patients with Parkinson's disease (**PD**) in the mean age of 64.9 years and a control group (**HC**) of 10 healthy people in the mean age of 56.2. The second database consists of 24 listeners of CTU FEE (**FEE**) in the mean age of 25.6 years.

Table 1: Structure of speech database used for training on automatic classifier.

database	**male**	**female**	∅ age
PD	10	3	64,9
HC	4	6	56,2
FEE	17	7	25,6
total	31	16	–

3 Evaluated tasks

Fricative consonants are produced by narrowing a supraglottal part of the vocal tract and such constriction generates a typical turbulence noise. In the case of the consonant /s/ the narrowed part occurs in mouth and it is created by an elevation of the tongue up while axial sides of the tongue are clamped to the upper teeth ridge. The slot remains between the dorsum tip of the tongue and the front part of the alveola. The lip slot is often quite narrow. The typical spectrogram of fricative consonants /s/ and /š/ is depicted in Fig. 1.

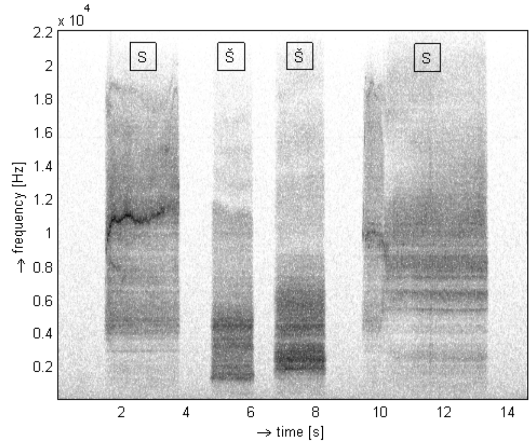
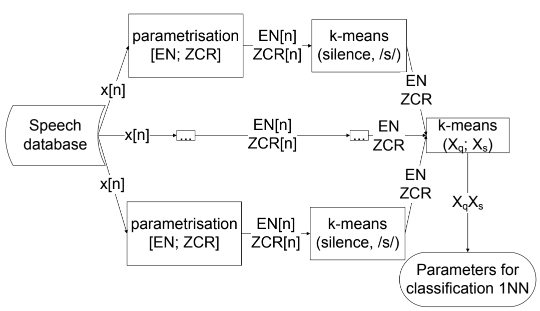


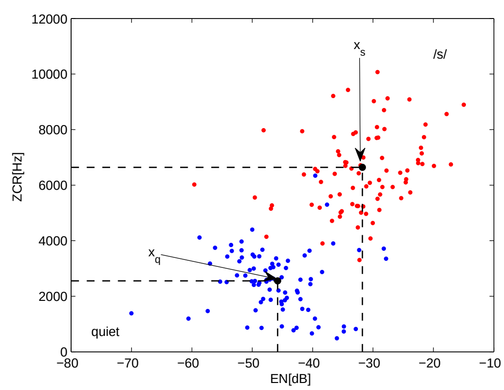
Figure 1: Typical spectrogram of fricative consonants /s/ and /š/.

3.1 Detection of /s/

Detection of /s/ is based on energy (EN) and zero crossing (ZCR). The main classifier uses 1 Nearest Neighbour algorithm (1NN) which is from the class of supervised classifiers. As samples which are used as a reference by the algorithm, "universal silence" (X_q) and "universal /s/" (X_s) were used. To find these values the training database consists only of records of task *"Ability to sustain /s/ on exhalation"* that is why there are only /s/ and quiet. The process of finding these values is shown in Fig. 2(a). In Fig. 2(b) universal values from every record[1] and both final "universal" values X_q and X_s are shown.



(a) estimation of X_q and X_s used for detection of /s/ by 1 Nearest Neighbour algorithm



(b) visualisation of results in <EN; ZCR> space

Figure 2: Process of determination of "universal silence" and "universal /s/" in each record and the final X_q and X_s.

3.2 Ability to sustain /s/ on exhalation

The description of this task is the following: *Take a deep breath and start to exhale with /s/ as long as you can. The intensity of /s/ should be uniform.* In this task the therapist evaluates the ability to manage the exhalation stream, the uniformity of intensity of /s/ during the whole exhalation and can also see the position of the patient's tongue and lips.

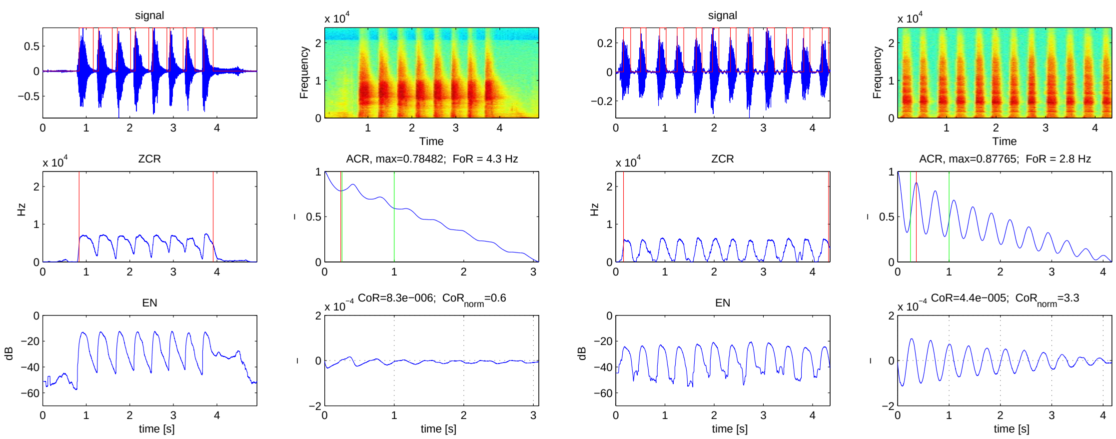
In this study the length of the exhalation is the only observed parameter.

[1] Blue dots are silence, red dots are /s/

3.3 Ability to repeat series of /s/

The description of this task is the following: *Take a deep breath and start to repeat series of prolonged /s/. The series should be repeated 2 or 3 times per second and the length of each serie should be the same.* In this task the therapist evaluates the patient's ability to start and stop the exhalation stream immediately as well as the length of every serie.

Patients with Parkinson's disease are often not able to start and to stop the exhalation stream immediately and that is the reason why there is s blur in the spectrum as shown in Fig. 3(a) in contrast to Fig. 3(b), in which the typical result of healthy control group is depicted.



(a) Parkinson disease – is not able to stop exhalation immediately; CoR = $8.3 \cdot 10^{-6}$; FoR = 4.3 Hz

(b) healthy – starts and stops exhalation stream immediately; CoR = $4.4 \cdot 10^{-5}$; FoR = 2.8 Hz

Figure 3: Differences between healthy control and Parkinson's disease record.

To evaluate this task automatically two parameters were defined:

- Frequency of Repetition (**FoR**) – the rate of series per second, a value between 2 and 3 is considered normal;
- Coefficient of Regularity (**CoR**) – describes the ability to start and to stop the exhalation stream immediately; the original task does not take this into account

The parameter FoR is simply determined by the autocorrelation function and the definition of CoR is the following:

$$\mathrm{CoR} = \sigma\left(\mathrm{diff}(\mathrm{ACR}(\mathrm{ZCR}(x)))\right) \begin{cases} > 1.365\ 10^{-5} & \text{healthy,} \\ \leq 1.365\ 10^{-5} & \text{disease,} \end{cases} \tag{1}$$

where σ is the standard deviation, 'ACR(ZCR(x))' is the autocorrelation function computed from zero crossing function between the first and the last serie of /sss/.

The greater the CoR, the more precise is the patient's ability to separate the series.

4 Results

The number of analysed records can differ from the number of patients in databases because one patient can have more records. The databases **HC** and **FEE** were merged together in order to classify only between the healthy and the Parkinsonians.

4.1 Ability to sustain /s/ on exhalation

As expected, the healthy control group is able to exhale longer than the Parkinsonians.

Table 2: Summary results of task *"Ability to sustain /s/ on exhalation"*.

	t [s]	
	PD	**HC+ FEE**
$\varnothing \pm \sigma$	12,86 ± 8,22	15,79 ± 9,02
median	9,58	16,10
min	0,00	0,01
max	29,81	48,60

4.2 Ability to repeat series of /s/

The parameter CoR described in Section 3.3 can be used as a threshold for linear discriminative analysis. For the threshold set to $1,356 \cdot 10^{-5}$, the classification results are shown in Tab. 3 and the distribution is shown in Fig. 4.

Table 3: Classification results of task *"Ability to repeat series of /s/"* according to CoR parameter.

database	**records total**	**pass**	**% pass**
PD	14	4	28%
HC	15	15	100%
FEE	24	22	92%

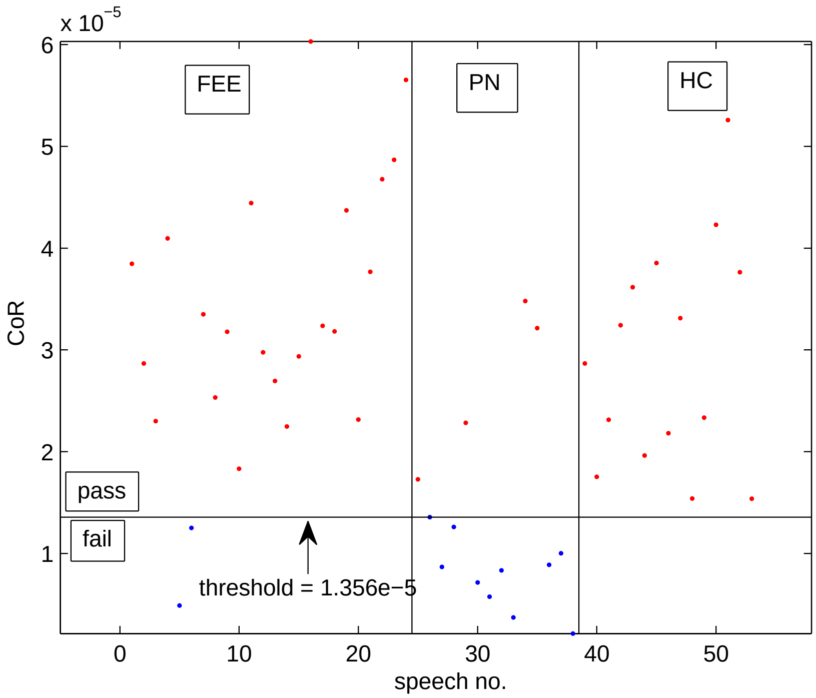


Figure 4: Visualisation of CoR based classification.

5 Conclusions

This paper deals with automatic assessment of two standard tasks from the dysarthria test. These tasks are *"Ability to sustain /s/ on exhalation"* and *"Ability to repeat series of /s/"*. An automatic detector of /s/ in speech based on energy and zero crossing function is introduced. This detector uses 1 Nearest Neighbour algorithm. The analysis reveals, that Parkinsonians are not able to exhale with /s/ as long as the healthy control group (see Tab. 2). A new parameter Coefficient of Regularity (CoR) useful for classifying task *"Ability to repeat series of /s/"* was introduced. Setting this parameter to $1,356 \cdot 10^{-5}$, a linear discriminative classifier classifies 88% of the records correctly.

Acknowledgements

This work has been supported by: **GACR102/08/H008** *Biological and Speech Signal Modelling*, **SGS10/180/OHK3/2T/13** *Assessment of voice and speech impairment*, **MSM6840770012** *Transdisciplinary Research in Biomedical Engineering.*

References

[1] HEDÁNEK, Jiří, ROUBÍČKOVÁ, Jaroslava. *Dysartrický profil - Test 3F*. Praha : Desk-Top Publishing Filozofické fakulty Univerzity Karlovy, 1997. 71 s.

[2] PUTNAM, Anne. H. B., HIXON, Thomas J. Respiratory Kinematics with Motor Neuron Disease. In *The Dysarthrias : Physiology, Acoustics, Perception, Management.* [s.l.] : [s.n.], 1984. s. 37-67.

[3] ROBERTSON, Sandra J., TANNER, Barbara, YOUNG, Fay. *Dysarthria Sourcebook.* [s.l.] : [s.n.], 1984. 140 s.

[4] ROBERTSON, Sandra J., TOMSON, Fay. *Working with Dysarthrics: A Practical Guide to Therapy for Dysarthria.* [s.l.] : [s.n.], 1986. 100 s.

[5] HAHN, Aleš, et al. *Otorinolaringologie a foniatrie v současné praxi.* Praha : Grada Publishing, a. s., 2007. 392 s.

[6] SMITH, Steven W. *Digital Signal Processing.* 2nd edition. San Diego, California : California technical Publishing, 1999. 848 s. Dostupný z WWW: <http://dspguide.com/>.

Adam Stráník
stranada@fel.cvut.cz

Roman Čmejla
cmejla@fel.cvut.cz