A RULE-BASED CLASSIFICATION OF LARYNGOPATHIES BASED ON SPECTRUM DISTURBANCE ANALYSIS

An Exemplary Study

Krzysztof Pancerz[1], Wiesław Paja[1], Jarosław Szkoła[1], Jan Warchoł[2] and Grażyna Olchowik[2]
[1]*Institute of Biomedical Informatics, University of Information Technology and Management in Rzeszów, Rzeszów, Poland*
[2]*Department of Biophysics, Medical University of Lublin, Lublin, Poland*

Keywords: Laryngopathy, Classification, Spectrum analysis.

Abstract: Our research concerns data derived from the examined patient's speech signals for a non-invasive diagnosis of selected larynx diseases. The paper is devoted to the rule-based classification of patients on the basis of a family of coefficients reflecting spectrum disturbances around basic tones and their multiples. The paper presents a proposed procedure for feature selection and classification as well as the experiments carried out on real-life data.

1 INTRODUCTION

Our research concerns designing methods for classification of patients with selected larynx diseases using a non-invasive diagnosis. Two diseases are taken into consideration: Reinke's edema (RE) and laryngeal polyp (LP). In general, the classification is based on selected parameters of a patient's speech signal (phonation). Gathering parameters in this way is convenient for the patient because a measurement instrument (microphone) is located outside the voice organ. This enables free articulation. Moreover, different physiological and psychological patient factors impede making a diagnosis using direct methods.

There exist various approaches to analysis of biomedical signals (cf. (Semmlow, 2009)). In general, we can distinguish three groups of methods according to a domain of the signal analysis: analysis in a time domain, analysis in a frequency domain (spectrum analysis), analysis in a time-frequency domain (e.g., wavelet analysis). Therefore, in our research, we are going to build a specialized computer tool for supporting diagnosis of laryngopathies based on a hybrid approach. A decision support system will have a hierarchical structure based on multiple classifiers working on signals in time and frequency domains.

A series of papers published earlier has presented approaches designed for the computer tool being developed. Approaches presented in (Warchoł et al., 2010) and (Pancerz et al., 2011) have been based on the speech spectrum analysis. However, in the approach presented in (Szkoła et al., 2010b), (Szkoła et al., 2010a), (Szkoła et al., 2011a), and (Szkoła et al., 2011b), we have detected all non-natural disturbances in articulation of selected phonemes using the modified Elman-Jordan networks that are recurrent neural networks. They can be used for pattern recognition in time series data due to their ability to memorize some information from the past. The approach initialized in (Pancerz et al., 2011) is extended in this paper.

2 CLASSIFICATION PROCESS

In the proposed approach, selected features (parameters) reflecting patient's speech spectrum disturbances around a basic tone and its multiples (harmonics) are calculated. Clinical experience shows that harmonics in the speech spectrum of a healthy patient are distributed approximately steadily. However, larynx diseases may disturb this distribution (cf. (Warchoł, 2006)). Therefore, the analysis of a degree of disturbances can support diagnosis of larynx diseases.

Disturbances are expressed by a family of coefficients computed for neighborhoods of a basic tone f_0 and its four multiples (f_1, f_2, f_3, f_4). In a real situation frequencies f_1, f_2, f_3, etc. are not distributed steadily (cf. (Warchoł, 2006)). It means, that we need to find a real distribution of harmonics. In the presented approach, it is done on the basis of the resul-

Pancerz K., Paja W., Szkoła J., Warchoł J. and Olchowik G..
A RULE-BASED CLASSIFICATION OF LARYNGOPATHIES BASED ON SPECTRUM DISTURBANCE ANALYSIS - An Exemplary Study.
DOI: 10.5220/0003874304580461
In *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing* (BIOSIGNALS-2012), pages 458-461
ISBN: 978-989-8425-89-8

tant spectrum (a sum of spectrums calculated for selected N-point time windows into which an original speech signal is divided). For each original frequency f, a maximum magnitude is searched in the interval $[f-d_1, f+d_1]$. This maximum value is assumed as a real harmonic. The original basic tone f_0 has been obtained for each patient from histogram created in the Multi-Dimensional Voice Program (MDVP). It is a software tool for quantitative acoustic assessment of voice quality, calculating various parameters on a single vocalization (see (MDVP, 2011)). On the basis of f_0, its harmonics (for ideal case) have been calculated.

Each coefficient expresses the distribution of a spectrum around a given frequency f. We can distinguish two types of coefficients: the regularity coefficient R determining a degree of slenderness of this distribution and the deviation coefficient D determining a relative difference between a real multiple derived from the spectrum and a multiple calculated on the basis of the basic tone f_0.

Let a speech signal of a given patient (a vector of samples $S=[s_1,s_2,\ldots,s_n]$) be given. Our approach can be expressed as a sequence of the following steps:

1. Normalizing signal sample values to the interval $[-1.0,1.0]$, i.e., for each sample $s \in S$, a normalized sample s_{norm} is equal to $s_{norm} = \frac{s}{\max\limits_{i=1}^{n}|s_i|}$.
2. Dividing the speech signal S into a family $\mathbf{W}$ of N-point disjoint time windows, i.e., $\mathbf{W} = [W_1,W_2,...,W_m]$, where $W_1 = S[1,\ldots,N]$, $W_2 = S[N+1,\ldots,2N]$, etc.
3. Calculating harmonics (first f_1, second f_2, third f_3, and fourth f_4) of a patient's basic tone f_0, i.e., $f_1 = 2f_0$, $f_2 = 3f_0$, $f_3 = 4f_0$, and $f_4 = 5f_0$.
4. Calculating a resultant discrete spectrum $\mathbf{Sp}$ for the family $\mathbf{W}$ of time windows excluding some initial and final ones, based on the Discrete-Time Fourier Transform (DTFT), see e.g. (Semmlow, 2009), i.e., $X[k] = \sum\limits_{n=1}^{N} W[n]e^{\frac{-2\pi jkn}{N}}$, where $k = 0,1,\ldots,N-1$. The spectrum $Sp(W)$ is a vector of magnitudes $|X[k]|$, i.e., $Sp(W) = \left[|X[0]|,|X[1]|,\ldots,|X[\frac{N}{2}]|\right]$. The resultant discrete spectrum $\mathbf{Sp}$ is calculated as $\mathbf{Sp} = \sum\limits_{W\in\mathbf{W}} Sp(W)$.
5. Finding real multiples (f_1^r, f_2^r, f_3^r, and f_4^r) of the basic tone f_0. A real multiple f_k^r is assumed to correspond to the maximum value of $\mathbf{Sp}$ in a neighborhood of f_k, i.e., in the interval $[f_k-d_m, f_k+d_m]$, where d_m is the input parameter and $k=1,\ldots,4$.
6. Calculating the regularity coefficient R_k, which can also be treated as some kind of shape parameter, for each frequency f_k (the basic tone f_0 or its multiple f_1^R, f_2^r, f_3^r, or f_4^r). Two discrete integrals I_1 and I_2 of the spectrum are calculated, the first one for the frequency interval $[f-d_2, f+d_2]$, the second one for the frequency interval $[f-d_3, f+d_3]$, where d_2 and d_3 are input parameters and $d_2 < d_3$. The discrete integral I of a fragment (between points k_1 and k_2) of the spectrum $\mathbf{Sp}$ is equal to $I = \sum\limits_{j=k_1}^{k_2} |\mathbf{X}[j]|$. The ratio of I_1 to I_2 constitutes the regularity coefficient R_k, i.e., $R_k = \frac{I_1}{I_2}$. If both integrals are equal (ideal case), then the regularity coefficient $R_k = 1$. In a real situation, some fuzziness of a spectrum around a given harmonic frequency can be observed. It causes that the regularity coefficient $R_k < 1$, because $I_1 < I_2$. It is easy to see, that the greater the fuzziness of the spectrum, the smaller the regularity coefficient (the smaller the slenderness of the spectrum distribution).
7. Calculating the deviation coefficient D_k for each multiple f_k of the basic tone f_0. The deviation coefficient D_k is the ratio $D_k = \frac{|f_k^r - f_k|}{f_k}$, where $|x|$ denotes the absolute value of x.

After execution of the procedure consisting of steps presented above, we obtain a decision table which is an input for classification algorithms. A sample of the decision table is presented in Table 1.

In the decision table being a training set of cases for classifiers, we assign to each patient one of the two classes: *norm* - a norm - for the patient from the control group, i.e., without disturbances of phonation confirmed by a phoniatrist opinion, *path* - pathology - for the patient either with laryngeal polyp, or with Reinke's edema (both clinically confirmed).

The approach presented in this paper has been tested using classification algorithms available in the popular data mining and machine learning software tools: WEKA (Witten and Frank, 2005), Rough Set Exploration System (RSES) (Bazan and Szczuka, 2005), NGTS (Hippe, 1997). Four of them are rule-based algorithms: exhaustive (RSES) (Bazan et al., 2000), LEM2 (RSES) (Grzymala-Busse, 1997), Genetic (RSES) (Wróblewski, 1998), and NGTS (Hippe, 1997). Two of them are decision-tree based algorithms: J48 (WEKA) - an implementation of C4.5 (Quinlan, 1993), CART (WEKA) (Breiman et al., 1993).

Values of features can be treated as continuous quantitative data. Building classification rules for such data can be difficult and/or highly inefficient. Therefore, for RSES generation algorithms, the so-called discretization was a necessary preprocessing

Table 1: A sample of the decision table.

Patient ID	R_0	R_1	R_2	R_3	R_4	D_1	D_2	D_3	D_4	*CLASS*
#1	0.90	0.89	0.89	0.82	0.75	0.01	0.00	0.00	0.00	*norm*
#2	0.85	0.82	0.75	0.59	0.58	0.04	0.03	0.08	0.08	*path*
...	...	...	...	...	...	...	...	...	...	...

step (Cios et al., 2007). Its overall goal was to reduce the number of values by grouping them into a number of intervals. Some discretization techniques based on rough sets and Boolean reasoning have been presented in (Bazan et al., 2000).

The sets of rules obtained using NGTS were improved using the RuleSEEKER system (Błajdo et al., 2004). The main optimizing process was based on an exhaustive application of a collection of generic operations (Paja and Hippe, 2005): finding and removing redundancy, finding and removing incorporative rules, merging rules, finding and removing unnecessary rules, finding and removing unnecessary conditions, creating missing rules, discovering hidden rules, rule specification, selecting final set of rules.

3 EXPERIMENTS

In the experiments, sound samples were analyzed. The experiments were carried out on two groups (Warchoł, 2006). The first group included persons without disturbances of phonation - the control group (CG). They were confirmed by a phoniatrist opinion. All persons were non-smoking, so they did not have contact with toxic substances which can have an influence on the physiological state of vocal folds. The second group included patients of Otolaryngology Clinic of the Medical University of Lublin in Poland. They had clinically confirmed dysphonia as a result of Reinke's edema (RE) or laryngeal polyp (LP). Experiments were carried out by a course of breathing exercises with instruction about the way of articulation. The task of all examined patients was to utter separately different Polish vowels with extended articulation as long as possible, without intonation, and each on separate expiration.Each sound sample was recorded on MiniDisc MZ-R55 (Sony). Effectiveness of such an analysis was confirmed by Winholtz and Titze in 1998 (Winholtz and Titze, 1998).

For experiments, we have tested various combinations of input parameters. The best classification results have been obtained for the following parameters: $N = 8192$ - the number of points (samples) taken for DTFT, $d_m = 12$ - deviation for searching maximum, $d_2 = 4$, $d_3 = 8$ - deviations for calculating spectrum regularity coefficients.

To determine accuracy of generated rules by classification algorithms a cross-validation method was used. Cross-validation is frequently used as a method for evaluating classification models. It comprises of several training and testing runs. First, the data set is split into several, possibly equal in size, disjoint parts. Then, one of the parts is taken as a training set for rule generation and the remainder (sum of all other parts) becomes the test set for rule validation. In our experiments, a standard 10 cross-validation test was used (CV-10). Each classification algorithm was evaluated via testing the classification accuracy of unseen cases. Results of evaluation are collected in Table 2.

Table 2: Results of experiments: classification accuracy.

Algorithm	Classification accurracy
Exhaustive (RSES)	0.8430
LEM2 (RSES)	0.8700
Genetic (RSES)	0.8430
J48 (WEKA)	0.8441
CART (WEKA)	0.7922
NGTS	0.6786
NGTS (partial matching)	0.7786
NGTS (after optimization)	0.7214
NGTS (after optimization, partial matching)	0.8071

Almost all popular rule-based algorithms of machine learning and data mining manifest very similar classification accuracy. The obtained result can be treated as promising. In general, the considered problem is not simple. An important role is played by the quality of speech recording. The quality of results is also dependent on chosen preprocessing methods (e.g. filtration) and signal processing methods, e.g. Discrete-Time Fourier Transform (DTFT). An especially difficult task is the extraction of the correct signal.

4 CONCLUSIONS

On the basis of experiments described in the paper, we can notice that a family of coefficients, calculated on the basis of the analysis of spectrum shapes and deviations around a basic tone and its multiples for the examined patient's speech signal, may be helpful for classification of a patient. In general, the speech analysis of patients with larynx diseases is a difficult task. Therefore, the obtained results seem to be promising.

In the future, we plan to tune parameters reflecting regularity and deviations in the speech spectrum, especially by selecting proper regions of interest of the spectrum and we are going to add some new coefficients characterizing spectrum disturbances as well. The proposed approach based on analysis of speech signals in a frequency domain can be a part of a computer tool based on multicriteria decision making process. It should be treated as a supplementary element for other techniques. An important challenge is to design methods enabling distinction between different larynx diseases (for example, laryngeal polyp and Reinke's edema). The approach presented in this paper does not enable us to make this distinction.

ACKNOWLEDGEMENTS

This research has been supported by the grant No. N N516 423938 from the Polish Ministry of Science and Higher Education.

REFERENCES

Bazan, J. G., Nguyen, H. S., Nguyen, S. H., Synak, P., and Wroblewski, J. (2000). Rough set algorithms in classification problem. In Polkowski, L., Tsumoto, S., and Lin, T. Y., editors, *Rough Set Methods and Applications*, pages 49–88. Physica-Verlag, Heidelberg, Germany.

Bazan, J. G. and Szczuka, M. S. (2005). The Rough Set Exploration System. In Peters, J. and Skowron, A., editors, *Transactions on Rough Sets III*, pages 37–56. Springer-Verlag, Berlin Heidelberg.

Błajdo, P., Grzymała-Busse, J., Hippe, Z., Knap, M., Marek, T., Mroczek, T., and Wrzesień, M. (2004). A suite of machine learning programs for data mining: chemical applications. In Debska, B. and Fic, G., editors, *Information Systems In Chemistry 2*, pages 7–14. University of Technology Editorial Office, Rzeszow.

Breiman, L., Friedman, J., Olshen, R., and Stone, C. (1993). *Classification and Regression Trees*. Chapman & Hall, Boca Raton.

Cios, K., Pedrycz, W., Swiniarski, R., and Kurgan, L. (2007). *Data mining. A knowledge dis-covery approach*. Springer, New York.

Grzymala-Busse, J. (1997). A new version of the rule induction system LERS. *Fundamenta Informaticae*, 31:27–39.

Hippe, Z. (1997). Machine learning a promising strategy for business information processing? In Abramowicz, W., editor, *Business Information Systems*, pages 603–622. Academy of Economics Editorial Office, Poznan.

MDVP (2011). Multi-dimensional voice program (MDVP). http://www.kayelemetrics.com.

Paja, W. and Hippe, Z. (2005). Feasibility studies of quality of knowledge mined from multiple secondary sources. I. Implementation of generic operations. In Klopotek, M., Wierzchon, S., and Trojanowski, K., editors, *Intelligent Information Processing and Web Mining*, pages 461–465. Springer-Verlag, Berlin Heidelberg.

Pancerz, K., Szkoła, J., Warchoł, J., and Olchowik, G. (2011). Spectrum disturbance analysis for computer-aided diagnosis of laryngopathies: An exemplary study. In *Proc. of the International Workshop on Biomedical Informatics and Biometric Technologies (BT'2011)*, Zilina, Slovak Republic.

Quinlan, J. (1993). *C4.5. Programs for machine learning*. Morgan Kaufmann Publishers.

Semmlow, J. (2009). *Biosignal and Medical Image Processing*. CRC Press.

Szkoła, J., Pancerz, K., and Warchoł, J. (2010a). Computer-based clinical decision support for laryngopathies using recurrent neural networks. In Hassanien, A. et al., editors, *Proc. of the ISDA'2010*, pages 627–632, Cairo, Egypt.

Szkoła, J., Pancerz, K., and Warchoł, J. (2010b). Computer diagnosis of laryngopathies based on temporal pattern recognition in speech signal. *Bio-Algorithms and Med-Systems*, 6(12):75–80.

Szkoła, J., Pancerz, K., and Warchoł, J. (2011a). Improving learning ability of recurrent neural networks: Experiments on speech signals of patients with laryngopathies. In Babiloni, F. et al., editors, *Proc. of the BIOSIGNALS'2011*, pages 360–364, Rome, Italy.

Szkoła, J., Pancerz, K., and Warchoł, J. (2011b). Recurrent neural networks in computer-based clinical decision support for laryngopathies: An experimental study. *Computational Intelligence and Neuroscience*, 2011. Article ID 289398.

Warchoł, J. (2006). *Speech Examination with Correct and Pathological Phonation Using the SVAN 912AE Analyser (in Polish)*. PhD thesis, Medical University of Lublin.

Warchoł, J., Szkoła, J., and Pancerz, K. (2010). Towards computer diagnosis of laryngopathies based on speech spectrum analysis: A preliminary approach. In Fred, A., Filipe, J., and Gamboa, H., editors, *Proc. of the BIOSIGNALS'2010*, pages 464–467, Valencia, Spain.

Winholtz, W. and Titze, I. (1998). Suitability of minidisc (MD) recordings for voice perturbation analysis. *Journal of Voice*, 12(2):138–142.

Witten, I. H. and Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann.

Wróblewski, J. (1998). Genetic algorithms in decomposition and classification problem. In Polkowski, L. and Skowron, A., editors, *Rough Sets in Knowledge Discovery 2*, volume 2, pages 471–487. Physica-Verlag, Heidelberg, Germany.