# Usage of Phase Space Diagram to Finding Significant Features of Rhinomanometric Signals

AndriyYerokhin, Alina Nechyporenko, Andrii Babii, Oleksii Turuta,

Ihor Mahdalina

*Abstract*– **Active anterior rhinomanometry is important method for diagnosis of rhinological disorders. This paper presents the new approach for feature extraction based on chaos theory for tasks of rhinology. It has been demonstrated that rhinomanometric signals have a fractal properties. The usage of phase space diagram for feature extraction for rhinomanometric data was proposed.**

*Keywords*—**Biomedical signals, Rhinomanometry, Hurst parameter, Feature extraction, Classification.**

## I. INTRODUCTION

Patients with diseases of the nose and paranasal sinuses in most cases have nasal obstruction. A lot of pathologies relate to nasal obstruction. There are sinusitis, rhinitis, rhinosinusitis and others. For example, sinusitis can be divided into many different categories [1, 2]. Such Rhinology disorders have a wide range of symptoms. Medical data are complex, with low accuracy. Internal dependences are hidden, their distributions are often not known.

Medical diagnosis in rhinology is very complex process. There is a wide variety of methods for evaluating the parameters of nasal breathing. There are acoustic rhinometry, rhinomanometry, peak nasal flow, nasal spirometry and other. An overview of objective measures for functional diagnostics of nasal breathing is given in [3]. Rhinomanometry is one of the most widely used methods of objective evaluating of nasal breathing function [4]. In work [5] the spectral parameters derived from sound signal of nasal breathing. They were used for obstruction identification. Logistic regression, linear and piecewise linear regression were used to identify relationship between clinical symptoms, nasal resistance coefficients and anthropometric data of patients [6]. All of these methods have common disadvantage: lack of objective criteria for differential diagnosis of rhinological pathologies.

In recent years, there has been a dramatic increase in the use of computation-intensive methods to analyze biomedical signals. The general approach falls under the methods of artificial intelligence or machine learning for decision-making in medicine. Such methods require a dataset of significant features that will be fully representative of underlying biological processes.

Rhinomanometric signals are time series data. In this case the data scientists usually use two approaches: analysis of global integral statistical properties of signals or analysis of significant parts of signal [7]. The probabilistic-statistical methods for diagnosis of nasal breathing are not suitable because the law of the distribution of the test data is not determined [6].

In paper [8] the study was conducted on features which could be derived from airflow waveform. This study analyzed flow patterns for identifying of patients with abnormal spirometry. The method of fuzzy approximation based on F-transform for preprocessing of rhinomanometric signals was used for feature extraction [9].

However, physiological processes are all nonstationary and highly nonlinear. The study of biomedical processes, which are heavily depended on observations, is crucially important for analysis. The usage of highly sensitive sensors is always associated with the registration of noises. Preprocessing stage is necessary to informative features of signals extraction. Also the ‘’hidden information” in time series can be generated by complex biological systems. Loss of complexity may be a generic, defining feature of pathologic dynamics and the basis of new diagnostic, prognostic, and therapeutic approaches [10]. It’s very important to find the criteria derived from the data itself.

Applications of nonlinear signal processing methods allow analyzing test results in decision support for variety of illness. The list of applications includes automated electrocardiogram (ECG) or electroencephalogram (EEG) analysis for cardiovascular or neurological disorder diagnosis [11, 12]. Several techniques using non-linear chaos features of the signal have been proposed for classification [13, 14].

.

## II. MATERIALS AND METHODS

Rhinomanometric data are the result of the measurement according to active anterior rhinomanometry (AAR) method. This method is based on simultaneous registration of two parameters: differential pressure $\Delta p$ and an airflow rate $Q$ through a nasal cavity. Rhinomanometric data were recorded by a system for rhinomanometric measurements [15]. The measurements result are the values of the differential pressure $\Delta p$ [Pa] and the airflow rate $Q$ [cm$^3$/s] on a time t [s], presented in Fig.1.

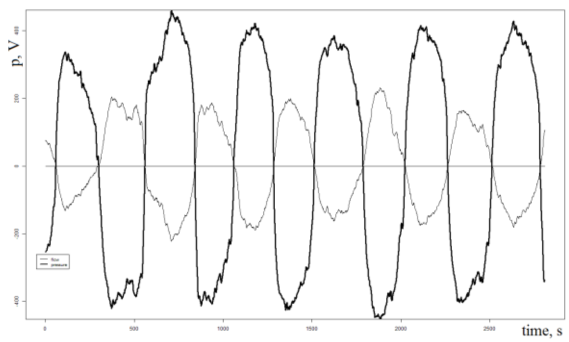


Fig.1.The dependence of pressure and airflow rate on a time

*A. Fractal properties of rhinomanometric signals*

Rhinomanometric signals are quasi periodic, nonstationary and nonlinear. In work [16] several chaotic properties of rhinomanometric signals were investigated. The main properties of a chaotic system are aperiodicity, determinism, confinement and sensitive dependence on initial conditions. For additional validation of existence of fractal properties of signals, we propose a calculation of Hurst parameter and fractal dimension. Hurst parameter is suggested as a measure of the degree of self-similarity of a data series because it allows the evaluation of the bursty nature of a data [17]. For calculation we use (1).

$$H = \frac{\log\left(\frac{R}{S}\right)}{\log\left(\frac{\tau}{2}\right)}, \tag{1}$$

where R – delta between maximal and minimal values of deviation $X(t,\tau)$, $\tau$ - is amount of elements of time series. Deviation $X(t,\tau)$ is calculated using $X(t,\tau) = \sum_{u=1}^{t} \{y(u) - \langle y \rangle_\tau\}$. The mean value $\langle y \rangle_\tau$ is defined by:

$$\langle y \rangle_\tau = \frac{1}{\tau}\sum_{t=1}^{\tau} y(t),$$

where $S = \sqrt{\frac{1}{\tau}\sum_{t=1}^{\tau}\{y(t) - \langle y \rangle_\tau\}^2}$, where parameter t has discrete integer values. Fractal dimension is calculated using (2):

$$D = 2 - H \tag{2}$$

Parameter D depends on the number of elements of the series, so we have used an amount of elements equal to 10000 and above [18]. Processed dataset contains 1076 measurements of rhinomanometric signals. Result of calculation is within the range of [0.16,0.20] and the range of [1.80,1.84] for H and D respectively. If $P = (H, D)$ belongs to this range, the rhinomanometric signals will have a fractal properties and will be anti-persistent.

*B. Phase Space Reconstruction method*

Phase space reconstruction is a standard procedure when analyzing chaotic systems. It shows the trajectory of the system in time. The phase space diagram of the rhinomanometric signal is constructed as follows [19]: the differential function $\dot{x} = \partial x / \partial t$ is plotted on the Y-axis, and the original function $x$ - on the X-axis of the phase plane. Phase diagrams for norm and septal deviation are shown in Fig. 2, 3.

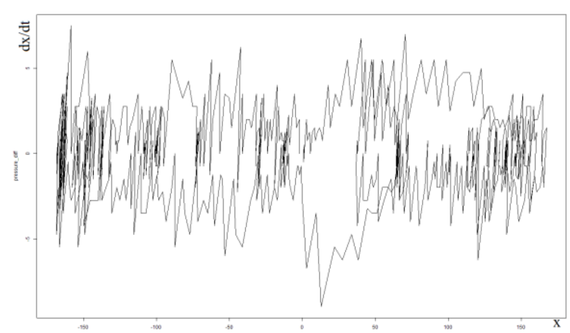


Fig.2.The phase diagram of pressure for 'norm'

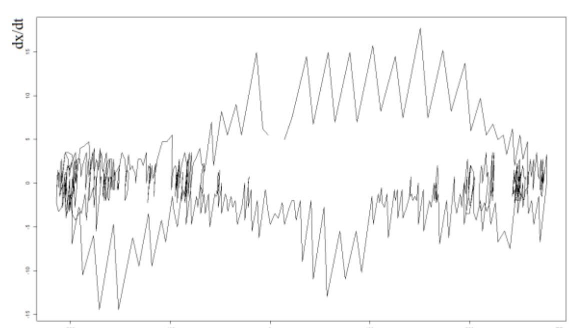


Fig.3.The phase diagram of pressure for 'septal deviation'

Processing and analysis of this data have one problem related with nature and technology of data receiving – amount of points in the source data differs for each measurement. We propose solution of this problem.

Let $i = 1..N$, where $N$ - amount of calculated values of phase diagram.

Represent each point with coordinates $(x_i, y_i)$ of the shape as a complex number

$$z_i = x_i + y_i j,$$

where $x_i$ as real part and $y_i$ as imaginary part.

We will apply Discrete Fourier transform (DFT) to the vector $Z = z_1 ... z_N$ using method from [20]

As result we will receive components

$$F = [F_0, F_1, ..., F_{N-1}]$$

Let the $K$ - count of pairs of Fourie components $F$, which will be used in reduced Fourie component list $F_r$

$$F_r = [F_0, F_1, F_2 ... F_K, F_{N-K}, ..., F_{N-2}, F_{N-1}]$$

If we perform Inverse Discrete Fourie Trasform (IDFT) to this reduced components list, we will receive approximated representation of the phase space diagram. It allows making visualization which could be useful in exploratory analysis of the data.

Also, we apply scale normalization, using approach represented in [21]

$$F_{sc} = \frac{F_r}{C_s}, \text{ where } C_s = \sqrt{|F_1|^2 + |F_{N-1}|^2}$$

Sample of visualization for the 40 Fourie components using IDFT is shown in Fig. 4

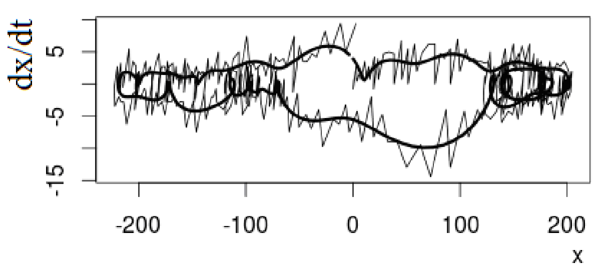


Fig. 4. Approximated phase diagram (bold line) and source phase diagram (thin line).

Using this approach we can receive fixed count of features ($F_r$) invariant to scale, which can be passed to supervised learning algorithms which have restriction for fixed count of input values.

Also, usage of restricted amount of Fourie components allows doing decrease level of hi-frequency noise in phase diagram.

In this preprocessing method $K$ shows how many Fourie components will be selected. Approximation result will be smooth or more close to initial data depends on value of the $K$ variable. This value is discreet. Selection of this value should be preformed to receive best learning values using learn/test/validation data set. For current research best $K$ value is 43.

This processing method should be applied to the both differential pressure and airflow rate.

## III. EXPEREMENTS AND RESULTS

We have used the data set with rhinomanometric signals. The signals have been collected from database of complex for objective evaluating of nasal breathing “Optimus”. Each measurement stores information about differential pressure and airflow rate. Amount of elements in data set is 1076 measurements, which were classified by otolaryngologists to the ‘norm’ and ‘deviation’ classes.

The classifiers implemented in this research were Support Vector Machine (SVM) [22] and Random Forest Approach (RF) [23]. Set of features consists of approximated phase diagrams. For each classification method the set of features has been performed such that the optimal classification results are achieved.

Learning set takes 85% from all measurement numbers, test set takes 15% from all measurement numbers. Best K value was selected equal to 43. Error rates for different learning methods are shown in Tabl.1.

TABLE I. ERROR RATES FOR DIFFERENT LEARNING METHODS

| K | Learning | | Test | |
|---|---|---|---|---|
| | RF | SVM | RF | SVM |
| 43 | 95,6 | 96,4 | 11,8 | 11,7 |

## IV. CONCLUSION

The paper demonstrates the potential of using the methods of nonlinear dynamics for processing of time series in biomedicine. The phase space diagram to finding significant features of rhinomanometric signals approach was proposed. Approach is based on DFT for generation of fixed amount of initial features of rhinomanometric signals. Supervised learning algorithms SVM and RF were used for classification on two classes ‘norm’ and ‘pathology’: The best value of K is equal to 43. Future investigation can be related with usage of methods of signal approximation.

## REFERENCES


[1] D. Wallace, M. Dykewicz, D. Berstein at al., “The diagnosis and management of rhinitis: an updated practice parameter”, *J Allergy Clin Immunol*, no 122, pp. 81–84, 2008.

[2] D. Demirbas, C. Cingi, H. Cakli, E. Kaya, “Use of rhinomanometry in common rhinologic disorders”, *Expert Rev. Med. Devices*, no. 8(6), pp. 769–777, 2011.

[3] C. Chaves, C. Ribeiro de Andrade, C. Ibiapina, “Objective measures for functional diagnostic of the upper airways: practical aspects”, *Rhinology, Vol. 52,* no 2, pp. 99-103, 2014.

[4] K. Vogt, A. A. Jalowayski, W. Althaus, C. Cao, D. Han, W. Hasse, H. Hoffrichter, R. Mosges, J. Pallanch, K. Shah-Hosseini, K. Peksis, K. D. Wernecke, L. Zhang and P. Zaporoshenko, “4-Phase- Rhinomanometry (4PR) – basics and practice 2010”, *Rhinology* Suppl. 21, pp. 1-50, 2010.

[5] R. Tahamiler, Y. Alimoglu, S. Canakcioglu, “Comparision of Odiosoft-Rhino and rhinomanometry in evaluation of nasal patency”, *Rhinoplogy*, Vol. 49, pp. 41-45.

[6] H.L. Thulesius, “Rhinomanometry in clinical use. A tool in the septoplasty decision making process”, Ph.D. dissertation, Department of Otorhinolaryngology, Head and Neck Surgery, Clinical Sciences, Lund University Sweden, 2012.

[7] N. Bogunovic, A. Jovic, “iomedical nonlinear signals by data mining methods”, *Proceedings of 17th Conference on Systems, Signals and Image Processing IWSSIP*, pp. 276-280, 2010.

[8] J.S. Reynolds, W.T. Goldsmith, J.B. Day, A.A. Abaza, A.M. Mahmoud, A.A. Afshari, J.B. Barkley, E.L.Petsonk, M.L. Kashon, D.G. Frazer “Classification of voluntary cough airflow patterns for prediction of abnormal spirometry”, *Journal of biomedical and health informatics*, pp. 2-8, 2015.

[9] A. Yerokhin, A. Nechyporenko, A. Babii, A. Turuta, “Usage of F-transform to finding informative parameters of rhinomanometric signals”, *Proc. of the International Conference on Computer Sciences and Information Technologies*, Lviv, Ukraine, 14-17 September, pp. 129-132, 2015.

[10] C.K. Peng, M. Costa, A.L. Goldberg, “Adaptive data analysis of complex fluctuations in physiologic time series”, Adv Adapt Data Anal. 2009 Jan 1; no 1(1), pp. 61–70.

[11] L.S. Fainzilberg, T.P. Potapova, “Computer amalysis and recognition of cognitive space electro-cardio graphic image”, Proc. of the 6th Intern. Conf. on Computer analysis of Images and Patterns (CAIP’95) *Prague, pp.* 668 -673, 1995.

[12] G. Saprykin, J. Kobytska, L. Kirichenko, “Using the characteristics of chaotic dynamics in the knowledge base of decision support system”, *Proc. of the International Conference on Computer Sciences and Information Technologies*, Lviv, Ukraine, 14-17 September, pp. 162-165, 2015.

[13] N. Wessel at al., “Short-term forecasting of life threatening cardiac arhytmias based on symbolic dynamics and finite-time growth rules”, *Phys. Rev. no*. 61(1), pp. 733-739, 2000.

[14] M. Costa, A.L. Goldberg, C.K. Peng, “Multiscale entropy analysis of complex physiologic time series”, *Phys., Rev., Lett, no*., 2, p. 89, 2002.

[15] A. L. Yerokhin, A. S. Nechyporenko, O.G. Garyuk, V.V. Chmovzh,“Software and hardware systems for rhinomanometric research «Optimus»”, *Proceedings of the International Ukrainian-Japanese Conference on Scientific and Industrial Cooperation*, Odessa. Oct. 24 – 25 pp. 74-76, 2013

[16] A. Nechyporenko, “Rhinomanometric signal processing for selection of formalized diagnostic criterion in rhinology”, Telecommunications and Radio Engineering, Vol 74, no. 14, P. 1285-1294, 2015.

[17] A.L. Goldberger, L. Amaral, J.M. Hausdorff, P. Ivanov, C.K. Peng, H.E. Stanley, “Fractal dynamics in physiology: Alterations with disease and aging”, *Proc. Natl. Acad. Sci.* (USA), vol. 99 (Suppl 1), pp. 2466–2472, 2002.

[18] Р.Э. Пащенко, “Фрактальный анализ процессов, структур и сигналов. Коллективная монография”, *Харьков: ХООО НЭО «Экоперспектива»*, 348 с., 2006.

[19] Л.С. Файнзильберг, “Оценка функционального состояния оператора на рабочем месте по фазовому портрету электрокардиограмм”, *Кибернетика и вычислительная техника*, Вып. 155. С. 3-17, 2007.

[20] E. Kamen, B. Heck, “Fundamentals of signals and systems using Mathlab*”, N.Y. Prentice Hall*, 1997.

[21] Dinh Nghia Do, S. Osowski, “Shape recognition using FFT processing and neural network”, *Compel*, Vol.17, No 5/6. pp. 658-666, 1998.

[22] C. Cortes and V. Vapnik, ”Support-vector networks” , *Machine Learning*, Vol. 20, no. 3, pp. 273-297, 1995.

[23] L. Breiman, “Random forests. Machine Learning”, no. 45(1), pp. 5–32, 2001.