Laryngeal pathology detection by means of class-specific neural maps

Article *in* IEEE Transactions on Information Technology in Biomedicine · April 2000

DOI: 10.1109/4233.826861 · Source: PubMed

3 authors, including:

Stefan Todorov Hadjitodorov
Bulgarian Academy of Sciences
58 PUBLICATIONS **913** CITATIONS

SEE PROFILE

LARYNGEAL PATHOLOGY DETECTION BY MEANS OF CLASS-SPECIFIC NEURAL MAPS[1]

Stefan Hadjitodorov[2], Boyan Boyanov, Bernard Teston*

Center on Biomedical Engineering,
Bulgarian Academy of Sciences,
Acad. G. Bonchev Street, Block 105, 1113 Sofia, BULGARIA
tel: +359 2 979 3653 fax: +359 2 723 787
E-mail:STHADJ@ARGO.BAS.BG

*Laboratoire Parole et Langage, E.S.A. 6057 - CNRS, Universite de Provence, 29,
Av.Robert Schuman, 13621 Aix-en-Provence Cedex, FRANCE

[1] Work supported by National Fund "Scientific Research" - contract I-804/98

[2] Corresponding author

ABSTRACT

Most of the existing systems and methods for laryngeal pathology detection are characterized by a classification error. One of the basic problems is the approximation and estimation of the probability density functions of the given classes. In order to increase the accuracy of laryngeal pathology detection and to eliminate the most dangerous error - classification of a patient with laryngeal disease as a normal speaker, here an approach based on modeling of the probability density functions (pdf) of the input vectors of the normal and pathological speakers by means of two prototype distribution maps (PDM) respectively is proposed. The pdf of the input vectors of an unknown normal or pathological speaker is also modeled by such a prototype distribution neural map - PDM(X) and the pathology detection is done by means of a ratio of specific similarities rather than by a direct comparison of some type of distance/similarity with a threshold. The experiments show an increased classification accuracy and that the proposed method can be used for screening the laryngeal diseases. The method is applied in a consulting system for clinical practice.

INDEX TERMS: Pathology detection, Neural networks, Self-organizing map, Prototype distribution map.

I. INTRODUCTION

Most of the laryngeal pathologies produce a change in the voice of the patient. In clinical practice, the otolaryngologist listens to the patient's voice as a first test of whether or not a normal voice is present, followed by direct or indirect laryngoscopy. The initial step in the diagnostics is listening to the patient's voice and it depends on many subjective factors such as: experience, training, emotional and physiological state, etc. The most dangerous disease of the larynx, the cancer, has to be diagnosed and treated during its early stages. However the patient's opinion of his/her laryngeal state is highly subjective and a physician may only be consulted when the pathology is already well advanced. The initial symptoms of the laryngeal disorders may include difficulties in breathing, pain and discomfort in speaking, or the feeling that the voice sounds different. An effective and non invasive method for early diagnostics of the diseases of the voice-producing system is the acoustical analysis of the voice. That is why several systems and approaches [1,3,6,7,10,11,15] for laryngeal pathology detection have been developed in the last few years. The researchers have used various parameters describing the pathological voices - pitch period (To), various statistics of To, noise in the residual signal from the LPC analysis, ratio harmonics-to-noise, etc., and different recognition methods. In [11] the pattern recognition is realized by means of the self-organizing map of Kohonen (SOM) and in [6] a fuzzy logic method is used. In the screening system [10] the pathology detection is realized by means of linear discriminant analysis. The method described in [1] classifies by calculating the Mahalanobis distance between the patient/normal speaker and the classes for normal and for laryngeal pathology voices. The approach [15] is based on recognition by means of multiple regression. In the two-level system described in [7] the linear discriminant analysis and K-NN rule are applied.

One of the main drawbacks of these systems and approaches is the presence of a classification error and in particular the most dangerous error - classification of a patient with laryngeal disease as a normal speaker, the so called “false negative”. One of the basic problems is the approximation and estimation of the probability density functions of the given classes. In order to increase the accuracy of laryngeal pathology detection, an approach based on modeling of the probability density functions (pdf) of the input vectors of the normal and pathological speakers by means of two prototype distribution maps (PDM) respectively is proposed here. The pdf of the input vectors of an unknown speaker/patient is also modeled by such a prototype distribution neural map - PDM(X) and the pathology detection is done by means of a ratio of specific similarities rather than by a direct comparison of some type of distance/similarity with a threshold.

This paper is organized as follows. In Section II the theoretical backgrounds of the proposed neural nets based method are described. In Section III the simulated annealing (SA) procedure for determination and adjustment of the necessary coefficients is developed. In Section IV the experimental research and the results are given and discussed. In Section V the conclusions are made.

II. THEORETICAL BACKGROUND

The prototype distribution map (PDM) is introduced and described in details in [8]. These maps are based on the SOM and as result have the following SOM's advantages:

1. In speech and image recognition SOMs are used to compress the amount of data that must be processed without loosing significant information [12,13].

2. By means of SOM the probability density function (pdf) of the input vectors could be estimated no matter how complex the form of the pdf is [9]. Features do not need to be assumed independent. More generally, there is no need for strong assumptions about the statistical distributions of the input features.

3. The algorithm for SOM formation is characterized by an asymptotic convergence (detailed proof in [14]).

4. The SOM algorithm can perform relatively well in the presence of a noise because [13]:

a) the number of neurons (weights) is fixed;

b) the weights adapt slowly;

c) the adaptation stops after training.

That means that the noise will be taken into consideration during the training of the map. This property is very useful for pathological voice analyses where the input signals are usually noise corrupted.

The second property is of great importance. On its base and on the base of consequent PDM formation we can achieve better approximation of the pdf for a given class (in our case two classes - NORMA and PATHOLOGY).

II.1. Speech analysis

The digitized speech signals are analyzed as follows: first, the stable zones are determined. In these zones the generation of the pitch period (To) is stable, i.e. the pitch period is nearly the same. Second, each stable zone is divided into segments (each segment consists of nearly 2048 samples from the digitized speech signal). Every segment

is represented by a feature vector. That means that each speaker's sustained vowel phonation (described in Section IV) is usually represented by 30 to 100 feature vectors which are the speaker's input vectors.

For the pathology detection, all of the below described voice parameters are used as features in the input vectors, with the exception of To. The following voice parameters are evaluated:

a) pitch period (To), by means of the method [4]. Using the To the stable zones are determined;

b) deviations of the pitch periods (PPQ) and of the amplitudes of the pitch pulses (APQ) [5]. PPQ and APQ are determined for the entire phonation and they retain the same values for all the input vectors corresponding to the different segments analyzed in the speaker's phonation;

c) stability of To generation (STAB) during vowel phonation [5]. STAB is determined for the entire phonation and it retains the same value for all the input vectors of the speaker's phonation;

d) the degree of dissimilarity of the shape(DISS) of the pitch pulses [5]. DISS is determined for each stable zone and it retains the same value for all the segments in the stable zone;

e) low-to-high energy ratio(LHER) [3]. LHER is determined for every segment analyzed in the given stable zone;

f) noise - to- harmonics ratio(NHR) in spectral domain by means of the method [17].NHR is determined for each segment analyzed in the given stable zone;

g) harmonics-to-noise ratio(HNR) in time domain by means of a modification (described in [5]) of the method of Yumoto[16]. HNR is determined for each stable zone and it retains the same value for all the input vectors corresponding to the different segments analyzed in the given stable zone;

h) ratio (energy concentrated in the pitch impulse - in cepstra)-to-total cepstral energy [2]. This parameter is determined for each segment analyzed in the given stable zone.

II.2. SOM formation

Let $y(t), t = 0,1,..., N$ be a sequence of n-dimensional input vectors representing the speech of M different classes, with $N_S, s = 1,..., M$ vectors available for each class, $\sum_s N_s = N$. These vectors are projected as neurons on a two-dimensional ($q \times q$) square map. Each neuron is defined by a n-dimensional model (weight) vector $\{m_{i,j}\}$, where $m_{i,j}$ corresponds to the (i,j)-th neuron. At the beginning of the self-organization process the

algorithm is initialized with random values for the $m_{i,j}(0)$. The placements of the input vectors on the map are optimized by iterative corrective steps. At each step the model vector $m_{a,b}(t)$ that is closest to $y(t)$ is determined according to the Euclidean distance [9,12]:

$$\left\|y(t)-m_{a,b}(t)\right\| = \min_{i,j}\left\|y(t)-m_{i,j}(t)\right\| \quad (1)$$

The best matching model vector $m_{a,b}(t)$ and the model vectors of the surrounding locations (neurons) are corrected by the rule:

if (i, j) is one of the neurons of neighborhood of (a,b), i.e.

$$\begin{aligned} i &\in [a-r(t), a+r(t)], \\ j &\in [b-r(t), b+r(t)], \end{aligned} \quad (2)$$

then

$$m_{i,j}(t+1) = m_{i,j}(t) + \alpha(t)\left[y(t)-m_{i,j}(t)\right], \quad (3)$$

otherwise

$$m_{i,j}(t+1) = m_{i,j}(t), \quad (4)$$

where $\alpha(t)$ is a monotonically decreasing, scalar - valued gain coefficient and $0<\alpha(t)<1$; $r(t)$ is a monotonically decreasing radius.

II.3. Formation of the prototype distribution map (PDM)

The feature vectors of speakers belonging to a given class $s, s=1,...,M$ are passed again through the already joint (commonly) trained SOM. As a result for each class at each neuron in SOM the frequency of activation ($f_{i,j}; i, j=1,...,q$), i.e. the number of input vectors which activate that neuron (according to the minimum of Euclidean distance), is obtained. Thus for each class a new map, containing the frequencies of activation, is formed. The values of $f_{i,j}$ are normalized by the number of vectors of all the speakers in the respective class. This map is named prototype distribution map -PDM [8]. The PDMs are used because they have the following useful properties[8]:

1. The PDM's neurons try to imitate and to approximate the pdf of the input signals, however complex the form of the pdf is. The PDM's estimation and approximation of the pdf gives better correspondence between the network and the pdf. This property is due to the facts that the PDM is formed on the basis of the already trained SOM and that the elements

of the PDM contain the frequencies of activation, which exactly corresponds to the usual presentation of the pdf.

2. The PDM allows dimensionality reduction - a two-dimensional SOM with n-dimensional weight vectors is transformed into a two-dimensional map with one dimensional weight vectors (frequencies of activation).

3. Less significant neurons in the PDM can be eliminated by filtering of the map according to the following rule:

if $0 \le f_{i,j} < k.f_{\max}$ then $f_{i,j} = 0$;

if $k.f_{\max} \le f_{i,j} \le f_{\max}$ then $f_{i,j} = f_{i,j}$, (5)

where $0 < k < 1$ is a filter coefficient adjusted by the "simulated annealing" (SA) procedure described in Section III ; $f_{\max}$ is the maximal value of $f_{i,j}$.

II.4. Training stage

The two classes **Norma** (Normal speakers) and **Pathology** (Patients with laryngeal diseases) are represented by means of two sample PDMs (PDM_{norma} and PDM_{pat}).

First step - Using the n-dimensional (here n=8) training input vectors of the speakers in class NORMA (y_{norma} (t)) and these of the patients (y_{pathol}(t)) a joint (common) two-dimensional SOM is trained.

Second step - Tuning of the SOM. In order to increase the accuracy of modeling of the input vectors by means of SOM one long term fine tuning of the SOM by means of the learning vector quantisation algorithm LVQ3 (detailed description in [9]) is done.

Third step - Formation of two sample PDMs (PDM_{norma} and PDM_{pat}) for the two classes (NORMA and PATHOLOGY) at the training stage. The following procedure is proposed and applied:

1. The input vectors y_{norma}(t) and y_{pat}(t) are passed separately again through the already commonly trained SOM. As a result PDM_{norma} and PDM_{pat} are obtained.

2. The PDMs are filtered in order to eliminate sporadically activated locations (noncharacteristic neurons) by means of (5).

The components of the speaker's input vectors are normalized (by dividing by the norm of the vector) before being subject to the training and decision making stages.

II.5. Decision making stage (pathology detection)

To increase the accuracy of laryngeal pathology detection the following procedure is proposed and used:

1. The data vectors $y_X(t)$ of the person (X) with unknown health state are passed through the already commonly formed (joint) sample SOM. As a result the PDM(X) for the normal or pathological speaker is obtained.

2. Filtering of the PDM(X) (the value of the filter coefficient ***k*** is the same as on step 3 of the training procedure).

3. Calculation of the similarities between PDM(X) and the two sample PDMs by means of the following cross-correlation type measure[8]:

$$D_{nor} = \sum_{i=1}^{q}\sum_{j=1}^{q}(f_{i,j}^{Nor} \cdot f_{i,j}^{X})(d + \left| f_{i,j}^{Nor} - f_{i,j}^{X} \right|)^{-1} \qquad (6)$$

$$D_{pat} = \sum_{i=1}^{q}\sum_{j=1}^{q}(f_{i,j}^{pat} \cdot f_{i,j}^{X})(d + \left| f_{i,j}^{pat} - f_{i,j}^{X} \right|)^{-1} \qquad (7)$$

where: $\{f_{i,j}^{Nor}\}$ - frequency of activation (value of a neuron) in PDM_{norma};

$\{f_{i,j}^{Pat}\}$ - frequency of activation (value of a neuron) in PDM_{pat};

$\{f_{i,j}^{X}\}$ - frequency of activation (value of a neuron) in PDM(X);

d - $0 < d < 1$ is a constant tuned by the SA procedure described in Section III.

4. Pathology detection and elimination of the most dangerous error (false negative -a person with laryngeal disease to be classified as a normal speaker) is carried out by the algorithm:

Step 1. Calculation of the ratio D_{nor} - to - D_{pat}:

RNP = D_{nor} / D_{pat} (8)

Step 2. Robust classification, done by the rules:

1. Classification of the unknown speaker in class NORMA if:

RNP > 1 + k_{nor}, (9)

where: $k_{nor} < 1$ is a coefficient determined by the SA procedure described in Section III.

2. Classification of the unknown speaker in class PATHOLOGY if:

RNP < 1 - k_{pat}, (10)

where: k_{pat} is a coefficient adjusted by the SA procedure described in Section III.

To avoid the above mentioned most dangerous error the following inequality is introduced and kept:

$k_{pat} < k_{nor} < 1$. (11)

3. Refusal of classification (no classification is done) if:

$1 - k_{pat} \leq RNP \leq 1 + k_{nor}$ (12)

III. SIMULATED ANNEALING PROCEDURE

Simulated annealing (SA) [18, 20] is implemented as the adjustment and tuning algorithm for the coefficients *k, d,* k_{nor} and k_{pat}. We have chosen this algorithm because it can optimize the parameters with respect to the classification accuracy directly, i.e. with classification accuracy as criterion function.

Let *W* be the vector $W = (w_1, \ldots, w_4) = (k, d, k_{nor}, k_{pat})$ of all parameters tuned at a time and *J(W)* be the criterion function value, given *W*. We form *J(W)* as the counting estimate of the classification accuracy, i.e. the proportion of correctly classified cases from the training set. Briefly, the algorithm operates as follows:

1. Generate randomly a parameter vector *W* (each parameter takes a value in the respective range). Calculate *J(W)*. Fix the initial value of the temperature $T = T_{ini} > 0$. Set the iteration counter *t* to 0.
2. Conduct l_T trials at the current *T*
   a) generate a neighbor state *W'* of *W* according to the current temperature spanning limitation
   b) calculate *J(W')*
   c) calculate $\delta = J(W) - J(W')$
   d) if $\delta<0$ then *W*:= *W'* , else if *random*(0,1)< $\exp(-\delta/T)$ then *W*:= *W'* otherwise *W*:= *W*
3. $t:=t+1; T:=\eta T, (0<\eta<1)$

The procedure stops when a predefined number of iterations t_{max} is reached. During the procedure in step 2.a) the inequality (11) is kept.

IV. EXPERIMENTAL RESEARCH AND RESULTS

IV.1. Speech material

The voices of 400 persons, which were recorded in the Phoniatric department of the University Hospital in Sofia, Bulgaria have been analyzed. The database consisted of:

-normal voices - 100 persons ;

-pathological voices (laryngeal pathology) - 300 patients.

The diagnosis was made by a council of three physicians using the routine subjective and objective examination methods – anamnesis, indirect laryngoscopy, audiometry, perceptual acoustical analysis, videolaryngostroboscopy, electromyography, phonetography, sonagraphy, biopsy and computer acoustical analysis. These 400 persons were cases for whom the diagnosis was set by consensus between the three physicians. We have not analyzed persons for whose diagnosis no agreement has been reached among the doctors. The distribution of the pathological patients according to their diagnosis is shown in Table 1. The voice signals were quantized directly into the computer's memory in order to avoid the distortions caused by the tape recorders. The station DSP Sonagraph Model 5500 of Kay Elemetrics was used at sampling rate 20480 Hz with 12 bits/sample.

Only sustained vowels were analyzed because they allow separation of the normal from the pathologic voices and are used in almost all systems for pathological voice analysis. One speaker's pronunciation of the sustained vowel "a" was analyzed.

IV.2. Training of SOM and PDM and tuning of parameters

During the training stage the voices of 150 patients and 50 normal speakers were used (200 phonations). The test of the proposed approach was realized over the remaining 150 patients and 50 normal speakers (200 phonations). Ten independent training and testing sessions were carried out, i.e. the entire data set was divided 10 times into various training and test sets of the mentioned size. Using the training sets the joint SOM, PDM_{norma} and PDM_{pat} are obtained and the parameter vector W was tuned by means of the SA procedure. For better approximation of the pdf the SOM (respectively PDMs) size was q=15. The experiments with other sizes of SOM have shown slightly worse results.

The means and the variances for each of the optimized parameters over the ten training sessions are given in Table 2.

IV.3. Classifier performance

In order to make comparison between the proposed approach and some classical recognition methods, the laryngeal pathology detection was realized by means of: the proposed approach, the K- nearest neighbors (K-NN) method, the linear discriminant analysis (LDA) and the classical SOM method. Common and well known classifier performance metrics have been used [19] for comparing the results obtained by the various classification methods. For a given decision of a classification method, four possible alternatives exist:

1. True positive (TP) - the method has classified a pathological patient (according to the verified physicians diagnosis) as belonging to class PATHOLOGY;
2. True negative (TN) – the method has classified a normal case as belonging to class NORMA;
3. False positive (FP) - the method has classified a normal case as belonging to class PATHOLOGY;
4. False negative (FN) - the method has classified a pathological case as belonging to class NORMA;

The sensitivity *SE* (likelihood that an event will be detected given that it is present) and the specificity *SP* (likelihood that the absence of an event will be detected given that it is absent) are determined as follows:

$$SE = 100xTP/(TP+FN) \quad (13)$$

$$SP = 100xTN/(TN+FP) \quad (14)$$

IV.4. Results and discussion

The averaged (over the ten runs) classification accuracy, accuracy variance, sensitivity and specificity in [%] for the laryngeal pathology detection by the K-NN, LDA, SOM and the proposed methods are shown on Table 3. Using the proposed method at an average of 3.2 patients (out of 150) were classified as a normal speaker. However for at an average of 6.6 persons (2.8 patients and 3.8 normal speakers) the classification was not done. The overall classification accuracy of the proposed method was 95.1%. The incorrectness of the method was 4.9% - 1.6% pure errors (in our case FNs) and 3.3% refusals (1.4% FNs and 1.9% FPs). Using that method the confidence of the classification decision made was increased – the pure accuracy (without refusals) was 98.4% and for the 3.3% the inequality (12) was fulfilled. In this case more parameters of the speech signals should be analyzed in order to make correct classification. Probably these errors were due to the fact that the classes were overlapping and there were no functions allowing complete separation of such classes. The sensitivity and the specificity of the method proposed were higher than those of the other methods, i.e. the correct detection of the pathology and normal state was higher. It should be noted also that there were no FNs among the patients with laryngeal precancerouses and cancer - for the other methods there was at an average of one patient erroneously classified as normal. Testing with a paired t -test showed the significance of the averaged accuracy differences between the proposed and other experimented methods.

V. CONCLUSIONS

The proposed classifier may be used for laryngeal pathology detection. Concerning the pure classification errors - at an average of 1.6% of the patients were classified as normal speakers. For other 3.3% of the tested persons the classification was not done, i.e. the method did not classify some “conflict” cases. Consequently, the confidence of the classification decision made, is increased. This fact is very important, because the most dangerous error is significantly minimized, leading to the supposition that the method may be used for screening the laryngeal diseases and that is why the method is applied in a consulting system for laryngeal pathology detection in the clinical practice. The increased accuracy of the proposed approach is due to:

1. Modeling the pdf of the input vectors of the normal and pathological speakers by means of two prototype distribution maps respectively.

2. Modeling the pdf of the input vectors of the unknown normal or pathological speaker also by a prototype distribution map.

3. Laryngeal pathology detection by means of ratios rather than by direct comparison of some type of distance/similarity with a threshold.

Our future investigations will be devoted to the differential diagnosis of the laryngeal pathologies and to the analysis of the voices of patients for whose diagnosis no agreement has been reached among the physicians.

REFERENCES

[1] G. Banci, S. Monini, A. Falaschi and N.Sario, “Vocal fold disorder evaluation by digital speech analysis”, *J. Phonetics,* vol. 14, pp. 495-499, 1986.

[2] B. Boyanov and G.Chollet, “Pathological voice analysis using cepstra, bispectra and group delay functions”, In *Proc. ICSLP'92*, Banff, Alberta, Canada, Oct.1992, pp. 1039-1042.

[3] B. Boyanov,D. Doskov,T. Ivanov and S.Hadjitodorov, “PC based system for analysis the voice of patients with laryngeal diseases”, *ENT Journal* , vol.70, No 11 , pp. 767-772, 1991.

[4] B. Boyanov, S. Hadjitodorov, T. Ivanov and G. Chollet, “Robust Hybrid Pitch Detector”, *Electronic Letters,* vol.29, No 22, pp. 1924-1926, 1993.

[5] B. Boyanov, “Analysis of pathological voice”, research report under contract ERB-CIPA-CT-92-0170 with the European Community, 1992.

[6] R. Gubrynowicz, “A fuzzy approach of pitch period analysis for evaluation of functional status of the larynx source”, *J. Phonetics,* vol.14, pp. 525-530, 1986.

[7] S. Hadjitodorov, T. Ivanov and B. Boyanov, "Analysis of Dysphony Using Objective Voice Parameter", In *Proc.of 2-nd Balkan Conf. on Operational Research,* Thessaloniki, Greece, pp. 911-917, Oct. 1993.

[8] S. Hadjitodorov, B. Boyanov, T. Ivanov and N.Dalakchieva, "Text-independent speaker identification using neural nets and AR-models", *Electronics Letters,* vol.30 No 11,pp. 838-840, 1994.

[9] T. Kohonen, "The self-organizing map", *Proc. IEEE,* vol.78,No 9, pp. 1464-1480, 1990.

[10] J. Laver, S. Hiller, J. Mackenzie and E. Rooney, "An acoustic screening system for the detection of laryngeal pathology", *J. Phonetics* vol. 14, pp. 517-524, 1986.

[11] L. Leinonen, J.Kangas, K. Torkkola and A. Juvas , "Dysphonia Detected by Pattern Recognition of the Spectral Composition", *J. Speech & Hear. Res.* , vol.35, pp. 287-295, 1992.

[12] R. Lippmann , "Pattern classification using neural networks", *IEEE Communications Magazine,* No 11,pp. 47-64, 1989.

[13] R. Lippmann, "An Introduction to Computing with Neural Nets", *IEEE ASSP Magazine* No 4, pp. 4-22, 1987.

[14] Z-P Lo, Y. Yu and B. Bavarian, "Analysis of the Convergence Properties of Topology Preserving Neural Networks", *IEEE Trans. Neural Networks* ,vol.4 ,pp. 207-220, 1993.

[15] C. Ludlow, C. Bassich, N. Connor, D. Coulter and Y. Lee, "The validity of using phonatory jitter and shimmer to detect laryngeal pathology"*, In: Laryngeal function in phonation and respiration*, Boston, Brown & co, 1987, pp. 492-508.

[16] E. Yumoto, " The quantitative evaluation of hoarseness. A new harmonics to noise ratio method", *Archives of Otolaryngology,* vol.109, pp. 48-52, 1983.

[17] M. Yunik and B. Boyanov, " Method for evaluation the noise-to-harmonics ratios in pathologic and normal voices", *Acustica* vol.70,pp. 89-91, 1990.

[18] R. Eglese, "Simulated annealing: A tool for operational research", *European Journal of Operational Research*, vol. 46, pp. 271-281, 1990.

[19] R.Eberchart and R.Dobbins; *Neural Network PC Tools. A practical Guide*. Academic Press, New York, 1990.

[20] Kuncheva, L. and S. Hadjitodorov. An RBF Network with tunable function shape. *Proc. of 13-th Int. Conf. on Pattern Recognition*, August 1996, Vienna, Austria, pp. 645-649.

Table 1. Distribution of the patients into laryngeal pathology classes

Pathological class (disease)	**Number of patients**
Hyperkinetic dysphonia	39
Hypokinetic dysphonia	24
Spastic dysphonia	10
Inflammatory reactive tumors	35
Chronical laryngitis	92
Laryngeal paralyses	50
Laryngeal precancerouses	32
Laryngeal cancer	18
Total	300

Table 2. Means and variances of the tuned parameters

Parameter	Mean	Variance
k	0.115	0.00225
d	0.0175	0.000013
k_{nor}	0.253	0.0017
k_{pat}	0.167	0.0006

Table 3. Averaged classification accuracy, accuracy variance, sensitivity and specificity in [%]

Method	Accuracy	Variance	Sensitivity	Specificity
K-NN	90.05	3.136	90.16	89.6
LDA	89.95	3.581	90.04	89.6
SOM	90.45	1.636	90.23	91
PROPOSED	95.1 (pure errors - 1.6%) (refusals – 3.3%)	0.433 (0.211) (0.122)	96.0	92.4