Voice Analysis for Detecting Parkinson's Disease

Conference Paper · December 2011

1 author:

Reza Arefi
Islamic Azad University
4 PUBLICATIONS   **55** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Voice Analysis for Detecting Parkinson's Disease View project

functional near infrared spectroscopy View project

Voice Analysis for Detecting Parkinson's Disease Using Genetic Algorithm and KNN Classification Method

R. Arefi Shirvan, E. Tahami

Abstract—**Parkinson's disease is a disorder of central nervous system. It is estimated that 90 percent of people with parkinson's disease suffer from speech and voice disorders. Vocal folds are usually undermined by this disease which would lead to creation of an improper voice in the patient's speech. In this paper, various features have been extracted from the voice signals of healthy people and people suffering from parkinson's disease. Afterwards, optimized features that influenced the process of data classification were detected using genetic algorithm and ultimately, based on various numbers of optimized features, the data classification was done using KNN classification method. It was shown that a classification accuracy percent of 93.7 per 4 optimized features, an accuracy percent of 94.8 per 7 optimized features and an accuracy percent of 98.2 per 9 optimized features could be achieved which is a notable result compared to other studies.**

I. INTRODUCTION

PARKINISON's disease was first described by the British scientist James Parkinson in 1817 so the disease was named after him. After the Alzheimer, parkinson's disease is said to be the most common neurodegenerative disorder [1]. It results from the death of dopamine-containing neurons in the substantia nigra. These neurons produce dopamine which carries messages from midbrain to another part of the brain called corpus striatum. These messages balance the body movements. While the dopaminergic neurons in substantia nigra die, other controlling units of body movements work irregularly. The disturbances in movement controlling units of brain cause the symptoms of parkinson's disease [2]. The disease is a brain disorder which is more common in people over 60 years of age. The prevalence of the disease is about 1% of the people over 60 years old. Although 5-10% of cases are younger people. Some of the disease's symptoms are: tremor(specifically when the limbs are stable), rigidity and slowness of movement in the whole body, dropping of the fore foot leading to an unnatural space between the two legs and improper walking, facial changes, changes in voice(the voice weakens), depression,... [1]. Patients suffering from the disease have a shaky rough breathy voice when speaking and they usually speak slowly and loosely. Their voice tone is decreased and the words can't be pronounced well. All of

these affect the voice clarity. These disturbances occur when the speech controlling muscles do not work properly. The vocal folds are weakened by this disease and their improper vibrating leads to creation of an undesired voice [3]. Speech disorders gradually isolate the patient from the society and affect his/her personal and social life. It is so hard for many of patients with parkinson's disease to visit clinical centers for treatment purposes. On the other hand, major developments and advances in telemedicine systems have made it possible for some of the patients to be treated without going to a doctor and via telemedicine systems [4]. Although applying this technology needs highly reliable and applicable tools. It is estimated that about 90 % of people suffering from parkinson's disease are prone to vocal disorders. In other words we could say that speech disorders may be among the early stage symptoms of the disease [5], [6]. Voice signal recording is easy and non-invasive so it is a notable parameter for detecting and tracking the symptoms progression.

Voice analysis has been used in many researches to detect various illnesses specifically the parkinson's disease. In these researches, the subject's voice is recorded in an standard experiment using a microphone and then specific features will be extracted from the recorded signal and will be analyzed by different methods [7]-[11].

M. Ene [10], 2008, extracted usual nonlinear features out of the same data used in this paper and the maximum classification accuracy of 81.28% using 22 features is achieved. M. A. Little and his colleagues [8] ,2008, extracted features similar to those used in [10] and achieved a classification accuracy of 91.4 after selecting four optimized features based on the correlation equation and using the SVM method. M. F. caglar and his colleagues [9], 2010, using MLP, RBF and ANFC classification methods and extracting features similar to those extracted in [10] and succeeded to achieve classification accuracy of 93.1% with four features.

In this paper the data is described first and after that various standard characteristics will be extracted from it. In order to decrease the feature vector dimension, optimized features are chosen using genetic algorithm and finally the data from healthy people and people with parkinson's disease is classified using the KNN method.

R. Arefi Shirvan, Department of Electrical Engineering, Bardaskan Branch, Islamic Azad University, Bardaskan, Iran. TeleFax: +98-5327226204; e-mail: Arefi.reza@iaubc.ac.ir.

E. Tahami, Department of Biomedical Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran. TeleFax: +98-5116629467; e-mail: tahami@mshdiau.ac.ir.

II. MATERIALS AND METHOD

In this research data from [8] have been used. The data include 192 voice signal recordings from 32 male and female subjects. Each subject has had 6 voice signal recordings. 23 subjects suffer from parkinson's disease and the rest are healthy people. Subjects were about 46-85 years old (with a standard deviation of 9.8 and mean age of 65.8). The duration of each recording is about 3 seconds. There are many vocal tests in order to achieve proper information for detection of speech disorders. Two types of them are commonly used: 1-The person is trained to make a vowel and try to keep the bass and tremble as steady as possible. 2- The person is trained to express an standard sentence and then needed information will be extracted. In this test every person is asked to smoothly create the sound of the word "A" for 3 seconds. The voice signal recording is done in an special voice recording room using an AKGC420 microphone, which is 8 centimeters away from lips. The voice signal has been recorded with a sampling frequency of 44.1 Hz and a resolution of 16 bits. Recorded signals have been normalized digitally before feature extraction. Fig. 1 shows an example of recorded signals from a healthy subject and a subject with parkinson's disease.
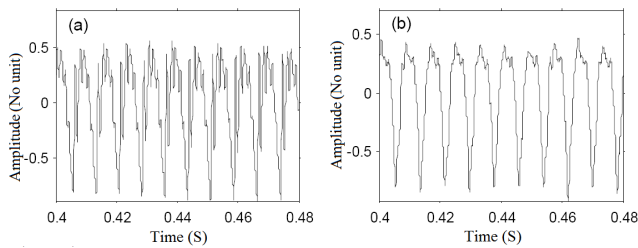

Fig. 1. Two selected examples of speech signals: (a) healthy, (b) subject with Parkinson's disease. The horizontal axis is time in seconds, the vertical axis is signal amplitude (no units) [8].

III. FEATURE EXTRACTION

In this research, the most important and most common factors of voice signal have been used to extract the features. The factors include: F0 (fundamental frequency or pitch), jitter, shimmer and noise to harmonics ratio. Studies show that the change in these factors is notable in people with parkinson's disease compared to healthy people. 13 different features have been extracted from each recorded voice signal. The praat software [12] has been used to calculate the features. Each feature is explained below [13]:

--Fo (Hz): Mean fundamental frequency.
--Fhi (Hz): maximum fundamental frequency.
--Flo (HZ): minimum fundamental frequency.
--Jitter (%):This is the average absolute difference between consecutive periods of fundamental frequency, divided by the average period (expressed as a percentage)

$$Jitter(\%) = \frac{\frac{1}{N}\sum_{i=1}^{N-1}|T_i - T_{i+1}|}{\frac{1}{N}\sum_{i=1}^{N}T_i} \tag{1}$$

Where Ti are the period of fundamental frequencies of window number "i" and N is the total number of windows.

--Jitter (ABS): This is the average absolute difference between consecutive periods of fundamental frequency, in microseconds (μs)

$$Jitter(ABS) = \frac{1}{N}\sum_{i=1}^{N-1}|T_i - T_{i+1}| \tag{2}$$

--Jitter (RAP): This is the Relative Average Perturbation, the average absolute difference between a period of fundamental frequency and the average of it and its two neighbors, divided by the average period.

--Jitter (DDP): This is the average absolute difference between consecutive differences between consecutive periods, divided by the average period.

--Shimmer: This is the average absolute difference between the amplitudes of consecutive periods, divided by the average amplitude

$$Shimmer = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|A_i - A_{i+1}|}{\frac{1}{N}\sum_{i=1}^{N}A_i} \tag{3}$$

Where Ai are the peak to peak amplitude of window number "i" and N is the total number of windows.

--Shimmer (dB): This is the average absolute base-10 logarithm of the difference between the amplitudes of consecutive periods, multiplied by 20

$$Shimmer(dB) = \frac{1}{N-1}\sum_{i=1}^{N-1}|20\log(A_i - A_{i+1})| \tag{4}$$

--Shimmer (APQ3): This is the three-point Amplitude Perturbation Quotient, the average absolute difference between the amplitude of a period and the average of the amplitudes of its neighbors, divided by the average amplitude.

--Shimmer (APQ5): This is the five-point Amplitude Perturbation Quotient, the average absolute difference between the amplitude of a period and the average of the amplitudes of it and its four closest neighbors, divided by the average amplitude.

--Shimmer (DDA): This is the average absolute difference between consecutive differences between the amplitudes of consecutive periods.

--HNR: harmonics to noise ratio.

IV. GENETIC ALGORITHM FOR SELECTING OPTIMIZED FEATURES

Genetic algorithm is one of the most influential methods in the process of data classification which is effectively used to select optimized features. This algorithm is inspired by the natural evolution. The algorithm suggests various optimized solutions for the optimization problems [14]-[16]. The process of running this algorithm in order to select the optimized feature (pattern) is explained below [14] (Fig. 2):

--Step 1: First, each pattern's entropy is calculated by using equation (5) and output (target) vector's entropy is measured using equation (6)

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i) \tag{5}$$

$$H(Y) = -\sum_{i=1}^{n} p(y_i) \log p(y_i) \tag{6}$$

Where x is the vector of features and y is the vector of targets, p(x) and p(y) are respectively density probability
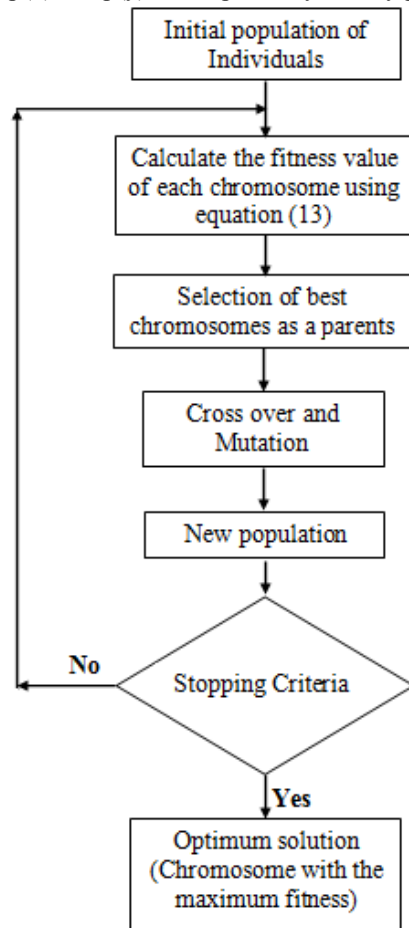

Fig. 2. Block diagram of genetic algorithm to select the optimal features.
function of features and targets.

Mutual information between each pattern and every single output (target) is measured via equation (7)

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$$
$$= H(X) + H(Y) - H(X,Y) = H(X,Y) - H(X|Y) - H(Y|X) \tag{7}$$

In this equation, the patterns' entropy (H(X)), the target vector's entropy (H(Y)) and H(X,Y) are measured by using equation (8)

$$H(X,Y) = \sum_{i=1}^{N}\sum_{k=1}^{N} -\log(P(x_i, y_k)).P(x_i, y_k) \tag{8}$$

And

$$H(Y|X) = H(X,Y) - H(X) \tag{9}$$

And ultimately H(Y|X) is measured via equation (10)

$$H(Y|X) = \sum_{x \in X} p(x) H(Y|X=x) = \sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \log \frac{1}{p(y|x)} \tag{10}$$
$$= -\sum_{x \in X}\sum_{y \in Y} p(x,y) \log(y|x) = -\sum_{x \in X, y \in Y} p(x,y) \log p(y|x)$$

--Step 2: Initial population of genetic algorithm is produced randomly using $200 \times n$ chromosomes, n is the number of features that need to be selected. So each chromosome consists of n genes where the feature's number is placed randomly and it is possible for the feature number to be repeated randomly in a chromosome.

--Step 3: The amount of relevance among patterns and targets is measured for each chromosome using equation (11)

$$V = \frac{1}{n}\sum_{i=1}^{n} I(X_i;Y) \tag{11}$$

Where I is the mutual information between features and targets.

The amount of redundancy among patterns and targets is measured for each chromosome using equation (12)

$$P = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} I(X_i;X_j) \tag{12}$$

--Step 4: The fitness value is assigned to each chromosome using equation (13)

$$\varnothing = V - P \tag{13}$$

Purpose of the suggested genetic algorithm is to maximize the fitness function of equation (13).

--Step5: The chromosomes are rearranged according to the given fitness function.

--Step 6: Elite chromosomes are selected as a parents.

--Step 7: In this Step, Crossover and mutation are done and a new population is produced. The chromosomes which can maximize the fitness function will remain and the rest will be removed and then Steps 1-5 are repeated and this process continues as long as the changes in chromosomes' fitness is less than 0.02 or the algorithm reaches the predetermined number of iterations which is supposed to be 80 in this paper. Finally the chromosome with the maximum fitness is chosen and the number of features in that chromosome is considered as selected features.

V. CLASSIFICATION BASED ON THE K-NEAREST NEIGHBOR (KNN) METHOD

In pattern recognition, the K-nearest neighbor (KNN) is a method for classifying objects based on closest training examples in the feature space. The KNN is a basic simple method in pattern recognition and when the information about the data distribution is not enough the method is amongst the first classification methods chosen [16]. This method includes two parts: a) determining K close neighbors, b) determining class type using these close neighbors. Suppose the training data space D described as equation (14)

$$D = \{X_1, X_2, ..., X_n\} \qquad (14)$$

Which includes n samples and each sample Xi is defined by f features as

$$X_i = (x_{i1}, x_{i2}, ..., x_{if}) \qquad (15)$$

And the whole data include C different classes. For determining the X' data class, first of all, its distance from all the data in space D is measured and then K data are determined from space D which are in the nearest neighborhood of X'. Class type of all K data is determined and so the X' belongs to the class which has had the most iteration between all K data. There are various criteria to evaluate the distance between X' and Xi but the most common one is the Euclidean distance criteria. Euclidean distance between X' and Xi along with the feature dimension, f, is defined as equation (16)

$$d = \sqrt{(x_{i1} - x'_1)^2 + (x_{i2} - x'_2)^2 + ... + (x_{if} - x'_f)^2} \qquad (16)$$

In order to use this method, 70 percent of the data was considered for training and the remaining 30 percent was considered as test data.

VI. RESULTS

In order to run optimized feature selection algorithms and also KNN classification method, MATLAB software has been used. After selection of optimized features by genetic algorithm, for various numbers of optimized features, data classification was done between the data of healthy people and those of patients with parkinson's disease, using the KNN classification method. Table I shows data classification accuracy per different numbers of optimized features.

TABLE I
DATA CLASSIFICATION ACCURACY PER DIFFERENT NUMBERS OF OPTIMIZED FEATURES

Number of Features (N)	Classification Accuracy (%)
9<N<13	98/2
7<N<8	94/8
4<N<6	93/7
N<4	75/8

The results from table I show that the maximum amount of classification accuracy which is approximately 98.2 % has been achieved for having at least 9 optimized features. The results also show that a classification accuracy of 94.8

% for at least 7 optimized features, a classification accuracy of 93.1 % for at least 4 optimized features and a classification accuracy of 75.8 % for less than 4 optimized features have been achieved. Table II has outlined the selected optimized features. Fig. 3 shows the diagram of chromosome changes versus the maximum fitness in each iteration. As it is shown in the figure, when the number of desired features is 10, the best fitness amount is -0.28373 which is achieved in 13th iteration and after that the algorithm has stopped.

TABLE II
NAME OF SELECTED OPTIMIZED FEATURES

Number of Features (N)	Feature	Classification Accuracy (%)
N=9	F0(Hz), Fhi(Hz), Flo(Hz), Jitter(%), Jitter(ABS), Jitter(RAP), Shimmer, Shimmer(APQ5), HNR	98/2
N=7	F0(Hz), Fhi(Hz), Flo(Hz), Jitter(ABS), Jitter(RAP), Shimmer, Shimmer(APQ5)	94/8
N=4	Fhi(Hz), Flo(Hz), Jitter(RAP), Shimmer(APQ5)	93/7

Some of the features discussed by genetic algorithm as the optimized features, have been drawn versus each other in Fig. 4 and Fig. 5.
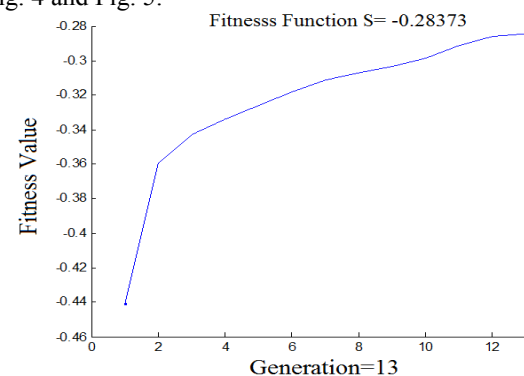

Fig. 3. Changes in fitness values of chromosomes per generations

Fig. 4 shows Jitter(RAP) versus Shimmer(APQ5) and Fhi(Hz) versus Shimmer and Fig. 5 shows Fo(Hz) versus HNR and Jitter(Abs) versus Shimmer. As it is shown in the figures, there is little correlation between data of healthy people and those ones with parkinson's disease which means that the genetic algorithm has performed well in detecting optimized patterns.

Considering the research results, it is seen that a classification accuracy of 93.7% has been achieved using 4 optimized features which is obviously a prominent result compared to other researches. Other results from this research include: 94.8% of classification accuracy for 7 optimized features and 98.2% of classification accuracy for 9 optimized features. No similar researches could have achieved such results so it shows that the genetic algorithm has been successful in selecting the best features that have had the most influence on data classification process between healthy people data and the data from patients suffering from the disease.
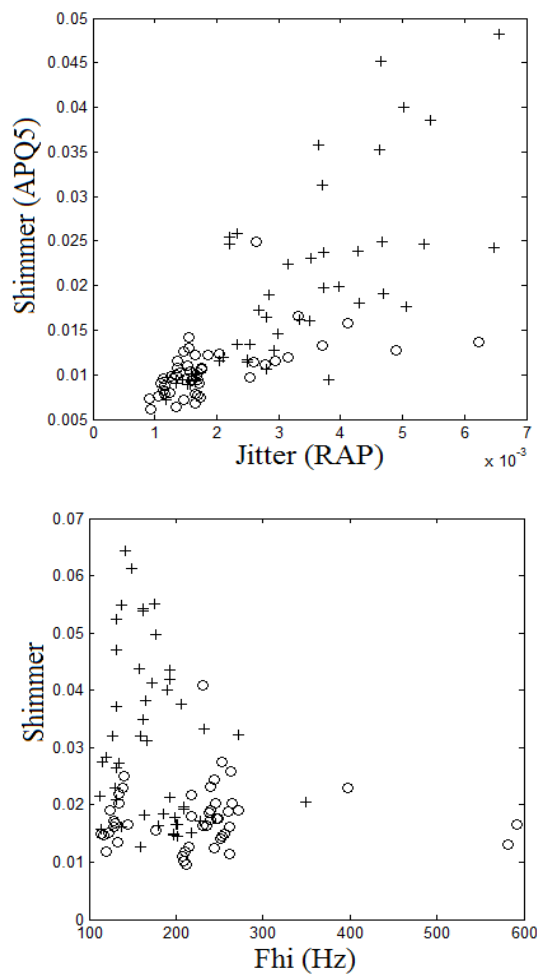
Fig. 4. Plots of pairs of features. A) Demonstration of Shimmer(APQ5)-Jitter(RAP) dataset, B) Demonstration of Shimmer-fhi(Hz) dataset. The "O" marks are for healthy subjects, the "+" marks for Parkinson's subjects.
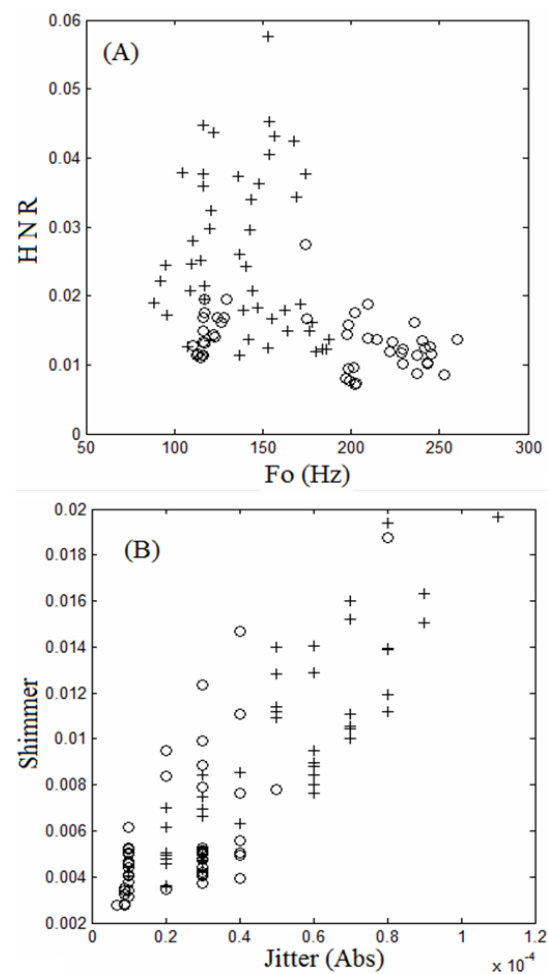

Fig. 5. Plots of pairs of features. A) Demonstration of HNR-Fo(Hz) dataset, B) Demonstration of Shimmer-Jitter(Abs) dataset. The "O" marks are for healthy subjects, the "+" marks for Parkinson's subjects.

VII. CONCLUSION

In this research, it was tried to extract useful information out of recorded voice signals using different feature extraction methods. In order to reduce the complexity of measurements, optimized and influential features were detected using genetic algorithm and finally, data classification was performed using KNN classification method. Recently, many studies have been done to detect illnesses such as parkinson's disease using voice signal analysis. In the research done, only usual voice signal features have been extracted and nonlinear features that include more complex calculations have been ignored. One noticeable advantage of this research could be using the genetic algorithm to find optimized features. Another advantage of this research is the usage of a simple classification method with non-complex measurements (KNN) and achieving noticeable results. Results from this research show that parkinson's disease notably affects the human voice. And also that the disease could be detected using voice analysis which could be used for various applications specifically for telemedicine purposes.

REFERENCES

[1] National Institute of Neurological Disorders and Stroke, http://www.ninds.nih.gov/disorders/parkinsons_disease/detail_parkinsons_disease.htm, (2009.06.21).

[2] WebMD,http://www.webmd.com/parkinsonsdisease/guide/parkinsons-disease-cause , (2009.06.21).

[3] K. M. Rosen, R. D. Kent, A. L. Delaney and J. R. Duffy, "Parametric quantitative acoustic analysis of conversation produced by speakers with dysarthria and healthy speakers," J Speech Lang Hear Res, vol. 49, pp. 395-411, 2006.

[4] C. Ruggiero, R. Sacile, and M. Giacomini, "Home telecare," J Telemed Telecare, vol. 5, pp. 11-7, 1999.

[5] A. K. Ho, R. Iansek, C. Marigliani, J. L. Bradshaw and S. Gates, "Speech impairment in a large sample of patients with Parkinson's disease," Behav Neurol, vol. 11, pp. 131-137, 1998.

[6] J. A. Logemann, H. B. Fisher, B. Boshes and E. R. Blonsky, "Frequency and Co-Occurrence of Vocal-Tract Dysfunctions in Speech of a Large Sample of Parkinson Patients," J Speech Hear Disord, vol. 43, pp. 47-57, 1978.

[7] A. Izworski, P. Augustyniak, T. Orzechowski, "PROCESSING AND ANALYSIS OF VOICE ANOMALIES IN COURSE OF PARKINSON'S DISEASES," Proceeding of the Eighth IASTED International Conference SIGNAL AND IMAGE PROCESSING, vol. 10, pp. 354-357, 2006.

[8] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, L. O. Ramig, "Suitability of dysphonia measurements for telemonitoring of

Parkinson's disease," IEEE Transactions on Biomedical Engineering, vol. 56, pp. 1015-1022, 2006.

[9]   M. F. Caglar, B. Cetisli, I. B. Toprak, "Automatic Recognition of Parkinson's Disease from Sustained Phonation Tests Using ANN and Adaptive Neuro-Fuzzy Classifier, " Journal of Engineering Science and Design, Vol. 1, No. 2, pp. 59-64, 2010.

[10] M. Ene, "Neural network-based approach to discriminate healthy people from those with Parkinson″ s disease, " Annals of University of Craiova, *Math. Comp. Sci. Ser.,* Vol. 35, pp. 112-116, 2008.

[11] G. Lee and S. Lin, "Changes of Rhythm of Vocal Fundamental Frequency in Sensorineural Hearing Loss and in Parkinson's Disease", Chinese Journal of Physiology, Vol. 1, pp. 446-450, 2009.

[12] P. Boersma and D. Weenink, "Praat, a system for doing phonetics by computer, " *Glot Int*, vol. 5, pp. 341-345, 2001.

[13] M. S. Ardakani, Z. Soleymani, F. Torabinejad, S. M. Khoddami, M. A. Heydari, "Fundamental frequency, jitter, and shimmer of adult stutters` and nonstutters` voice, " Audiology, Vol.16, No.2, 2008.

[14] O. Ludwig, U. Nunes, R. Araujo, L. Schnitman, H. A. Lepikson, "Applications of information theory, genetic algorithms, and neural models to predict oil flow," Communications in Nonlinear Science and Numerical Simulation,Vol. 14, Issue 7, pp. 2870-2885, 2009.

[15] H. Peng, F. Long, C. Ding, "Feature Selection Based on Mutual Information:Criteria of Max-Dependency, Max-Relevance and Min-Redundancy," IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 27, NO. 8, pp. 1226-1238, 2005.

[16] P. Devijver, andJ. Kittler, "PATTERN RECOGNITION: ASTATISTICAL APPROACH," Prentice Hall, 1982.