Revista Facultad de Ingeniería Universidad de Antioquia
ISSN: 0120-6230
revista.ingenieria@udea.edu.co
Universidad de Antioquia
Colombia

Vásquez Correa, Juan Camilo; Orozco Arroyave, Juan Rafael; Arias-Londoño, Julián David; Vargas Bonilla, Jesús Francisco; Nöth, Elmar

New computer aided device for real time analysis of speech of people with Parkinson's disease

Revista Facultad de Ingeniería Universidad de Antioquia, núm. 72, septiembre-, 2014, pp. 87-103
Universidad de Antioquia
Medellín, Colombia

New computer aided device for real time analysis of speech of people with Parkinson's disease

Nuevo dispositivo para análisis de voz de pacientes con enfermedad de Parkinson en tiempo real

Juan Camilo Vásquez Correa[1], Juan Rafael Orozco Arroyave[1,2], Julián David Arias-Londoño[1], Jesús Francisco Vargas Bonilla[1], Elmar Nöth[2]*

[1]Facultad de Ingeniería, Universidad de Antioquia. Calle 70 N.° 52-21. AA. 1226. Medellín, Colombia

[2]Pattern Recognition Lab., Friedrich-Alexander-Universität Erlangen-Nürnberg. Martensstraße 3. D-91058. Erlangen, Alemania.

(Received September 27, 2013; accepted May 15, 2014)

Abstract

Parkinson's disease (PD) is a neurodegenerative disorder that affects the coordination of muscles and limbs, including those responsible of the speech production. The lack of control of the limbs and muscles involved in the speech production process can generate intelligibility problems and this situation has a negative impact in the social interaction of the patients. It is already demonstrated that constant speech therapy can improve the communication abilities of the patients; however, the measurement of the recovery progress is done subjectively by speech therapists and neurologists. Due to this, it is required the development of flexible tools able to asses and guide the speech therapy of the patients. In this paper the design and deployment of a new device for the real time assessment of speech signals of people with PD is presented. The processes of design and deployment include the development on three platforms: first, a graphic user interface is developed on Matlab, second the first prototype is implemented on a digital signal processor (DSP) and third, the final device is developed on a mini-computer. The device is equipped with an audio codec, storage capacity and the processing unit. Besides, the system is complemented with a monitor to display the processed information on real time and with a keyboard enabling the interaction of the end-user with the device.

Different acoustics and nonlinear dynamics measures which have been used in the state of the art for the assessment of speech of people with PD are

* Corresponding author: Juan Rafael Orozco Arroyave, e-mail: rafael.orozco@udea.edu.co; phone number: +57 4 2198627 (J. Orozco)

implemented on the three mentioned platforms. In accordance with the state of the art, the designed platforms show an increment in the variation of the fundamental period of speech (commonly called pitch) of people with PD. Additionally, the decrease of the vocal space area is validated for the case of patients with PD. These results indicate that the designed device is useful to perform the assessment and monitoring of the speech therapy of people with PD.

----------*Keywords:* Parkinson's disease, portable device, real time speech assessment, vocal space area, jitter, shimmer, pitch, nonlinear dynamics

Resumen

La enfermedad de Parkinson (EP) es un desorden neurodegenerativo que afecta la coordinación de músculos y extremidades, incluyendo aquellos responsables de la producción del habla, generando alteraciones en la inteligibilidad de la señal de voz. Está demostrado que el ejercicio terapéutico constante puede mejorar las habilidades de comunicación de los pacientes; sin embargo, el diagnóstico acerca del avance en el proceso de recuperación es realizado de forma subjetiva por los fonoaudiólogos o neurólogos. Debido a esto se requiere el desarrollo de herramientas flexibles que valoren y guíen la terapia fonoaudiológica de los pacientes. En este artículo se presenta el diseño e implementación de un sistema embebido para el análisis en tiempo real de la voz de pacientes con EP. Para esto se desarrollan tres plataformas; primero, se construye una interfaz gráfica en Matlab; luego, se crea un primer prototipo basado en un DSP TMS320C6713 de Texas Instruments. La aplicación final es desarrollada sobre un mini-ordenador que cuenta con un códec de audio, capacidad de almacenamiento, y una unidad de procesamiento. El sistema además se complementa con un monitor LCD para desplegar información en tiempo real, y un teclado para la interacción con el usuario.

En todas las plataformas se evalúan diferentes medidas usadas comúnmente en la valoración de la voz de pacientes con EP, incluyendo características acústicas y de dinámica no lineal. En concordancia con otros trabajos del estado del arte donde se analiza la voz de personas con EP, la plataforma diseñada muestra un incremento en la variación del pitch en la voz de los pacientes, además de un decremento en el valor del área del espacio vocálico. Este resultado indica que la herramienta diseñada puede ser útil para hacer la evaluación y seguimiento de la terapia fonoaudiológica de pacientes con EP.

----------*Palabras clave*: enfermedad de Parkinson, dispositivo portátil, evaluación de voz en tiempo real, área del espacio vocal, jitter, shimmer, pitch, dinámica no lineal

Introduction

Parkinson's disease (PD) is the second more prevalent neurodegenerative disorder after Alzheimer's [1]. Patients with PD exhibit a neurologic disorder that is produced by the progressive loss of the dopaminergic cells in the substantia nigra. Such disorder inhibit the right control of the muscles and limbs, including those involved in the speech production process and generating intelligibility problems for the patients. Speech impairments appear in about 90% of the patients, and the symptoms experienced by them are commonly called hypokinetic dysarthria, which is perceived as monotonic speech, low tone, low intensity, with inappropriate stops, imprecise consonants production and prosodic problems [2].

With the aim of finding solutions to the speech impairments or to reach an earlier diagnose, there exist different tools and methodologies focused on the characterization and classification of speech of people with PD. Different laboratories and research centers have addressed the problem of measuring the affection of the speech of people with PD. In [3] the articulation capability of the PD patients is evaluated by means of the vocal space area (VSA) and vocal articulation index (VAI). The authors report that for people with PD the values of the VAI are lower than those obtained with speech of healthy controls. The study considered a total of 68 patients and 32 healthy controls and both groups were balanced by age and gender. In [4], different measures that are used to classify speech signals from people with PD in different states of the disease are presented. The authors highlight the increment on the jitter of speech signals of PD patients; additionally, they describe the ability of VSA and formant centralization ratio (FCR) to measure the articulation capability of the patients. On the other hand, in [5] a multimodal analysis that considers the tremor, gait and speech of people with PD is presented in order to perform an early diagnosis of the disease. The authors indicate that nonlinear dynamics (NLD) features are important for such diagnosis, and highlight the capability of correlation dimension (CD) and largest Lyapunov exponent (LLE) to classify between PD patients and healthy controls. In the same way, considering not only different NLD features, but also acoustics and noise measures, the authors in [6] perform the evaluation of speech of people with PD and the monitoring of their state during 6 months. Considering a set with 42 patients, the authors report that it is possible to assess the neurological state of the patients considering only their speech analysis. On the other hand, different works have been focused on the analysis of speech of people with PD considering spectral measures. In [7] the authors use linear prediction coefficients (LPC), cepstral LPC (LPCC), mel-frequency cepstral coefficients (MFCC), perceptual LPC (PLP) and relative spectra coefficients (RASTA) to perform the automatic classification of speech signals of people with PD and healthy controls. The database used includes 20 patients with PD and their respective controls (balanced in age and gender) who uttered the five Spanish vowels in a sustained manner. According to the reported results, considering different sets of spectral features, it is possible to reach accuracy rates of about 76%. In [8] the low frequency zone of the speech spectrum is also analyzed by means of the Teager energy operator (TEO) and the modified group delay functions (MGDF). The authors consider the same database of [7] and use these low frequency features to classify between speech signals uttered by people with PD and by HC with accuracies of about 92.5%. Respect to the NLD analysis, in [9] different measures are used recently, including CD, LLE, Lempel-Ziv complexity and Hurst exponent. The authors report accuracies of up to 77% considering only NLD features calculated on the same database of [7].

The research community has also shown interest in the development of devices and prototypes to perform real time analysis of speech signals. In [10] the authors present a portable instrument developed to control the volume of the speech of people with PD through auditory feedback. Few years later, in [11] the authors develop a portable device to control the intensity of the speech of patients with PD including not only the auditory feedback that is presented in [10] but also other visual mechanisms to give audio-visual feedback to the patient.

There exist also different patents related to the bio-feedback process. In [12] the authors present the details of a bio feedback method to train people in their self-regulation of physiological functions. The method is based on displaying the results of the analysis, giving useful information to correct the parts of the process that the patient is performing badly, or in order to display information that explains to the patient what is actually happening on his/her organism. In [13] a system designed for monitoring the speech therapy of people with stuttering problems is presented. The device is able to display information useful to both the patient and the therapist in order to know how much the therapy has advanced. Recently, in [14] a system to monitor the intensity of the speech of the patients while they are out of the hospital is presented. The device captures the signal by a contact microphone that is sticked to the neck of the patient while the results are displayed on a LCD monitor. Also in [15] is presented a device equipped with an accelerometer that is sticked to the neck of the patient to capture the movements of the skin while speaking. In the same device, the sound pressure level is also measured along with the fundamental frequency and the duration of the phonation. Additionally, in [16] the authors present a portable device for the assessment of speech signals on real time. The aim of the device is to identify different speech disorders by means of the analysis of the fundamental frequency and other features that are measured using one accelerometer sticked to the neck of the patient.

It is important to note that the methods presented in [14-16] are all invasive and based on the use of external instruments that are added to the body, which can be uncomfortable for the patients. Although there are different works that are focused on the analysis of speech of people with PD, there is a lack of portable devices that allow the speech analysis by means of a common microphone, not adhering external things to the neck of the patients, and providing the bio-feedback to the user on real time.

This paper presents a portable device that is able to assess the speech of people with Parkinson's disease. The device evaluates the speech signal and provides the bio feedback to the patients on real time, helping the patients to perform their speech therapy correctly, and thereby improving the outcome of the exercises. With the aim of performing a complete assessment of the speech, different measurements that have been already tested in the state of the art are included in the device. The set of implemented measures includes the fundamental frequency [17], its variations on time and amplitude (jitter and shimmer, respectively) [18], vocal formants, VSA and FCR [19]. Two different NLD features are also included in the device: CD and LLE.

All of the mentioned features are programmed on three different platforms. First, one graphic user interface (GUI) is built in Matlab. This GUI receives and process the speech signal to display the obtained values of the measures mentioned above. Once the algorithms are designed and tested on the GUI, they are implemented on the DSP TMS320C6713 from Texas Instruments. Finally, after the tests performed on the GUI and the DSP, the final prototype is deployed on a minicomputer called Odroid u2 [19]. This minicomputer has one processing unit, different peripheral ports, one audio codec and one storage unit. The system is complemented with an LCD monitor for the real time displaying of the processed information and with a keyboard that enables the interaction of the user with the device. Note that the Odroid minicomputer is chosen for the final device because it provides portability, and has all of the required peripherals to interact with the display, the computer, and the user. These peripheral characteristics are ready to use in the Odroid chipset, while for the DSP this is not the case, and the installation of such peripherals requires the development of additional electronic adjustments.

The obtained values from acoustic measures that are implemented on the portable device are also compared to those obtained with the software Praat [20], which is a software that is widely used by the community that peform acoustic analysis of speech signals.

The rest of the paper is organized as follows: section 2 describes the methodology that we have follow to design the system and also presents the experiments that were made to validate the proper operation of the system. Section 3 includes the results of the comparison of the measurements when are computed on the three different platforms, and finally in section 4 the conclusions and the future work derived from this study are presented.

Experiments

The measures deployed on the three platforms are presented in this section. First, the methods and algorithms developed to calculate features are introduced, second, the different platforms developed to compute the measures are described and finally, the database used to test the operability of the different platforms is presented.

Characterization

Due to its importance in the evaluation of dysarthric speech signals, the fundamental frequency is estimated on the three platforms along with other important acoustics measures such as jitter, shimmer, vocal formants, VSA and FCR. Pitch, jitter and shimmer are included to perform the evaluation of the phonation capability of the patients, while the coefficients VSA and FCR allow the evaluation of the articulation capability. Along with the acoustic features, two NLD measures are also introduced: CD and LLE.

Fundamental frequency: pitch

Pitch is defined as the fundamental frequency of the speech signal and it is associated to the vibration period of the vocal folds [17]; in people with speech disorders developed due to the presence of PD, this feature shows an unstable behavior and a decrease of its range, producing a monotonic speech [5]. The process to estimate the pitch is as follows: the energy of the normalized speech signal is calculated, only the frames of the speech that exceed a threshold are considered to calculate the autocorrelation function. Negative part of the autocorrelation function is eliminated (clipping) and finally, the pitch is estimated as the distance between the first two peaks in the clipped autocorrelation function [21]. The procedure is summarized in figure 1.

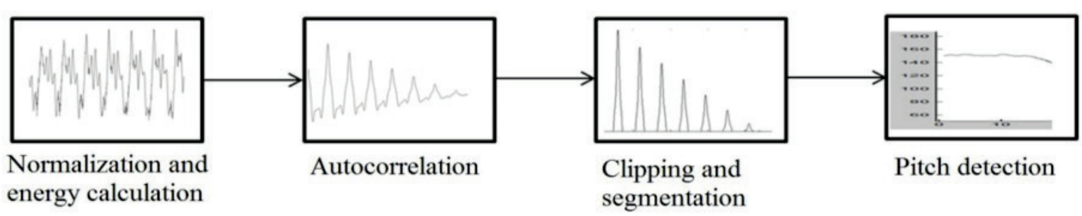

Figure 1 Estimation of the pitch value based on the autocorrelation function

Jitter

This feature gives information about temporal variation of the pitch during the phonation [18]. To calculate the jitter, equation (1) is applied, where *Mp* is the maximum value of the pitch, and it is updated while the phonation is running, acting as the reference to the estimation of the jitter.

$$Jitter(\%) = \frac{100}{N * Mp} \sum_{k=1}^{N} |pitch(k) - Mp| \tag{1}$$

Shimmer

It is used to measure the amplitude variation of the pitch during the phonation [18], to calculate the shimmer it is necessary to apply the equation (2), where *Ma* corresponds to the maximum amplitude of the signal in every pitch period.

$$Shimmer(\%) = \frac{100}{N*Ma} \sum_{k=1}^{N} |A(k) - Ma| \tag{2}$$

Vocal formants

In order to analyze the articulatory capability of the patients, the vocal formants are also calculated. As it is well known, vocal formants are able to model the resonances in the vocal cavity and these features can be found through linear predictive filtering (LP analysis). Such filter is represented by an all-pole transfer function where each of the conjugate pole pairs corresponds to a formant frequency (resonance). The estimation of the LP coefficients can be addressed by means of the autocorrelation method, which takes advantage of the Toeplitz symmetry that can be found in the autocorrelation function. The autocorrelation method is based on the Levinson-Durbin algorithm [22] and allows the efficient estimation of the vocal formants as the peaks of the espectral envelope that is formed by the estimated LP filter. Once the filter is built, one can find different vocal formants as the peaks on such envelope. In this work we are only focused on the first two formants, because these are used for the estimation of VSA and FCR.

Vocal space area: VSA

Once the first two formants, $F1$ and $F2$, are calculated for the vowels /a/, /i/ and /u/, the VSA can be calculated. The process begins with the construction of the vocal space, which is formed by $F1$ in the horizontal axis and $F2$ in the vertical axis of the Cartesian plane. With the values of the two formants for the vowels /a/, /i/ and /u/ a triangle is formed on such plane and the area of such triangle is called VSA [23]. According to [4], VSA is lower in people with PD than in healthy controls; this reduction indicates the loss of the articulation capability by the PD patients.

Formant centralization ratio: FCR

This is another feature that is measured to characterize the articulation capability of the PD patients. It was introduced in [19] and it is calculated according to the expression (3) after being estimated the first two formants for the vowels /a/, /i/ and /u/.

$$FCR = \frac{F_{2u} + F_{2a} + F_{1i} + F_{1u}}{F_{2i} + F_{1a}} \tag{3}$$

Correlation dimension: CD

Although pathological speech signals have been classically analyzed by acoustic and spectral features, there are cases where the level of the pathology is very high and the classical analysis is not 100% reliable [24]. CD is one of the NLD features and it is conceived to assess the intrinsic dimensionality of a time series i.e. speech signal. The estimation of CD begins with the reconstruction of the phase space of the signal; such space is also called attractor. This space, also known as embedding space, is built by a set of vectors that are formed with different delayed versions of the speech signal, as it is indicated in (4). Where $X[k]$ is the speech signal, τ is known as the embedding delay, and m is the embedding dimension [25].

$$X[k] = \{x[k], x[k+\tau], x[k+2\tau], \dots \\ x[k+(m-1)\tau]\} \tag{4}$$

The feature can be estimated according to the process that was proposed by Grassberger and Procaccia in [26]. First the correlation sum is defined as in (5), where Θ is the Heaviside function and N is the number of points in the speech frame. By means of this expression the distances between X_i and X_j points is evaluated. Only when such distance is less than a threshold given by a hyper-sphere of radius ε, the sum is increased.

$$C(\varepsilon) = \lim_{N\to\infty} \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \Theta(\varepsilon - |X_i - X_j|) \quad (5)$$

According to [26], for small values of ε it is possible to approximate $C(\varepsilon)$ to the expression presented in (6), which allows the estimation of CD as is expressed in (7).

$$C(\varepsilon) = \lim_{\varepsilon\to 0} \varepsilon^{CD} \quad (6)$$

$$CD = \lim_{\varepsilon\to 0} \frac{\ln(C(\varepsilon))}{\ln(\varepsilon)} \quad (7)$$

Largest Lyapunov exponent: LLE

This NLD feature quantifies the exponential divergence between neighbor trajectories in the embedded space (the phase space or the attractor). Through this feature it is possible to measure the "aperiodicity" in sustained phonations, which is considered as one level indicator of pathological speech signals [27].

To estimate LLE the first step is to find the reconstructed attractor, after that, the nearest neighbors to each point on every trajectory are found. One nearest neighbor X_j is one that minimizes the Euclidean distance to a reference point X_j, as is expressed in (8), where d_j (0) indicates the initial divergence between the point X_j and its nearest neighbor X_j. ‖·‖ denotes the Euclidean distance between points.

$$d_j(0) = \|X_j - X_j\| \quad (8)$$

With the aim of ensure that the neighbor points belong to different trajectories in the embedded space, one additional condition must be imposed. The distance between X_j and X_j should be greater than the average period of the speech signal. After estimating the divergence, the LLE is calculated as the average separation rate follows the Oseledec's theorem [28] which is expressed in (9). Where λI corresponds to LLE, $d(t)$ is the average divergence when it is measured at the instant t, and C is a normalization constant.

$$d(t) = Ce^{\lambda It} \quad (9)$$

Designed platforms

As it is stated above, three different platforms are designed in order to perform the real time analysis of speech signals of people with PD. First the GUI that is designed in Matlab is presented, second the algorithms that are adapted to operate on the DSP TMS320C6713 are shown and finally, the portable device that is deployed on the Odroid U2 is presented.

Matlab graphic user interface

The main window of the designed GUI is depicted in figure 2. This interface allows the estimation and the bio feedback on real time of the features described in section 2.1. The only exception for the real time visualization is for VSA and FCR which need the analysis of the three phonations /a/, /i/ and /u/.

Besides the estimation and displaying of the features, through this interface it is possible to make the registration of the patients allowing the comparison of their results through the time. The platform generates a report on a portable document format (pdf) with the therapy results, so both the therapist and the patient can control their progress.

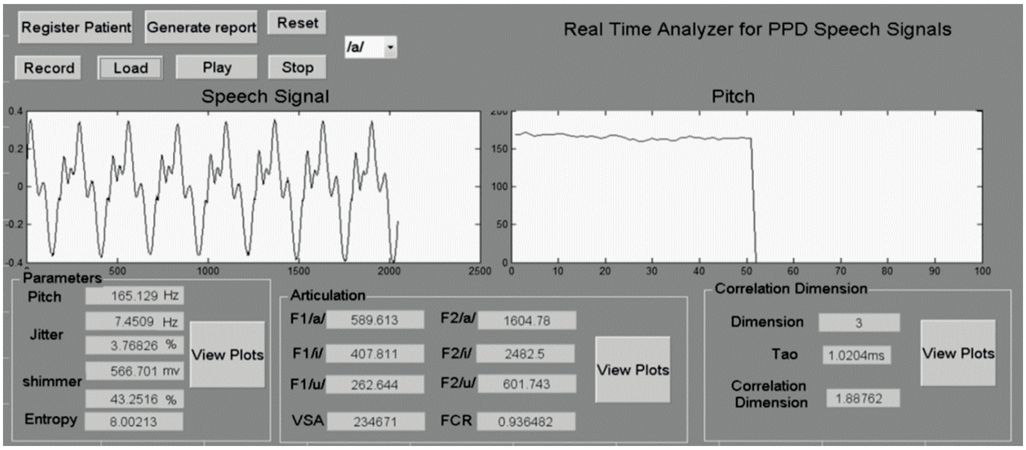

Figure 2 Matlab graphic user interface

DSP TMS320C6713

This platform is equipped with a 225MHz processor, a float point processing unit, 16MB of RAM memory, audio input/output ports and a 16bits audio codec.

On this device only the acoustic features are programmed and the process is made according to the following strategy [29]: three processing windows are formed; the first is used to store the input data, the second is used to process the data that is in the first window and the third is used to display the results. The details of the methodology are given in figure 3.

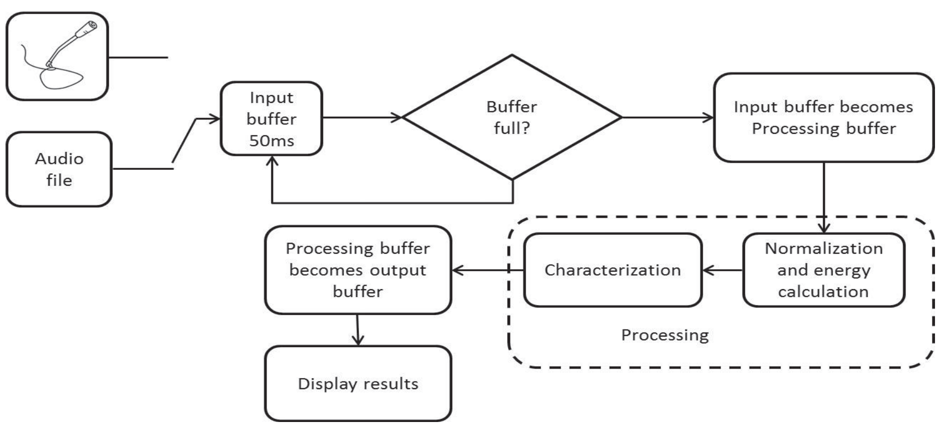

Figure 3 Blocks diagram of the followed process for the implementation on the DSP

Final prototype: the minicomputer Odroid U2

The designed device is based on the minicomputer Odroid U2 [30], which is a small and powerful processing unit. The board is equipped with an ARM Cortex-A9 quad core processor with a clock of 1.7GHz, one 3D graphic accelerator, 2GB of RAM memory and a micro SD port, which it is possible to store both operating system and data.

The designed device is not only composed by the minicomputer, there is also one 7" LCD monitor to display the results and one keyboard for the communication with the end-user and a storage unit. Figure 4 depicts a blocks diagram with the functional parts of the device. The software that is running in the device is developed on Python because it is an efficient scientific programming language than can be complemented with a varied set of libraries for acquisition, processing and displaying [31]. The platform is operating on Ubuntu linaro 12.10, which is a Linux distribution. With the aim of optimize the performance of the platform and the velocity in the characterization of the speech signals, some of the routines are developed on C++ using the scipy.weave library [32], additionally, some parts of the code are developed according to multiprogramming and multithreading techniques [33]. Figure 5 illustrates the followed process for the real time estimation of the acoustic and NLD features. After capturing one voice frame, the estimation of the features is executed, while the results of the estimated features from the last voice frame are displayed.

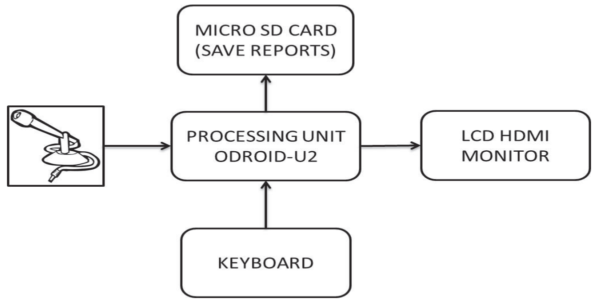

Figure 4 Blocks diagram of the designed device

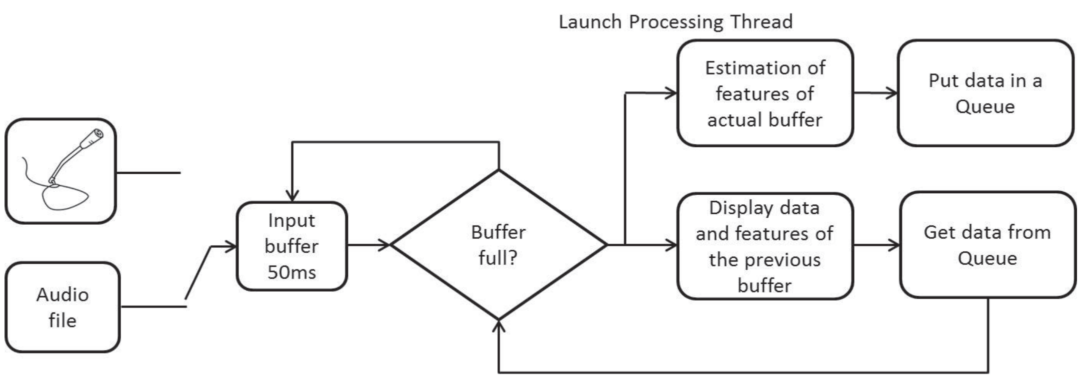

Figure 5 Process for the estimation of features on real time using the embedded system

As in the case of the Matlab GUI presented above, the device deployed in the embedded system has a graphic interface for the interaction with the user. The main window of such interface is showed in figure 6. This interface has buttons to enable different functions such as the registration of patients, generation of reports with the results of the therapy, recording of speech signals, playing of a previously stored recording, displaying of previously analyzed results, etc. Figure 7 illustrates the designed platform with its peripheral components.

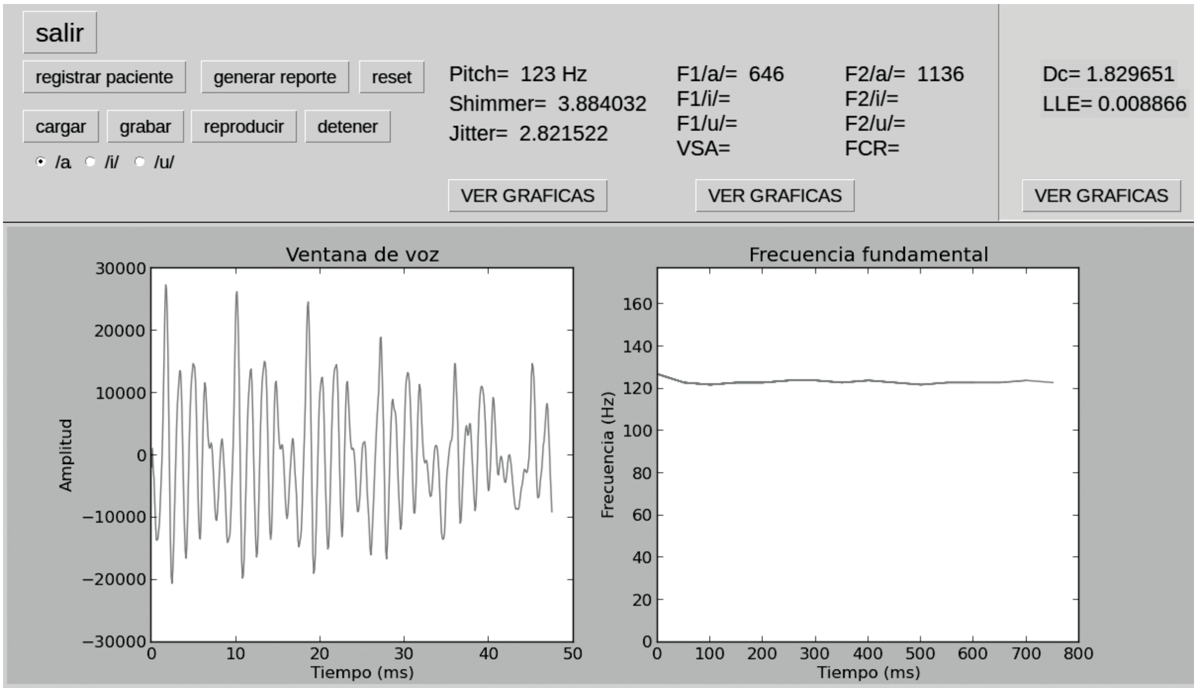

Figure 6 Graphic user interface deployed on the embedded system

Figure 7 Portable device for the analysis of speech of people with PD

Database and performed tests

The database that is used to test the reliability of the designed platforms is an extended version of the one that is used in [7-9]. This version includes speech recordings of 50 people with PD and 50 healthy controls, 25 men and 25 women on each group and all of the recordings are balanced by age and gender. The recordings are sampled at 44100Hz with 16 resolution bits. The age of the men with PD range from 33 to 77 years old (mean 62.2 ± 11.2), the age of the women with PD range from 44 to 75 years old (mean 60.1 ± 7.8). For the case of the healthy controls, the age of the men range from 31 to 86 (mean 61.2 ± 11.3) and the age of the women range from 43 to 76 years old (mean 60.73 ± 7.7).

All of the patients were diagnosed by neurology experts and none of the healthy control had symptoms associated to PD or any other movement disorder. All of the developed algorithms are tested on this database with the aim of assure the good performance of the platforms and to validate the obtained results. The acoustic measures are also compared to the results obtained in Praat, which is widely used for the acoustic analysis of speech signals [20].

Results and discussion

As it is well documented in the state of the art, people with dysarthric speech show instability in their pitch envelope, this phenomenon is validated on the prototype (Odroid U2). Left side of figure 8 illustrates the pitch envelope obtained from the recordings of five different PD patients. The right side of figure 8 shows the same feature measured on recordings of five healthy controls. Note that the pitch values are more stable in the case of healthy people, while in the case of the patient, the pitch contour presents jumps that reflect the stability problem mentioned above.

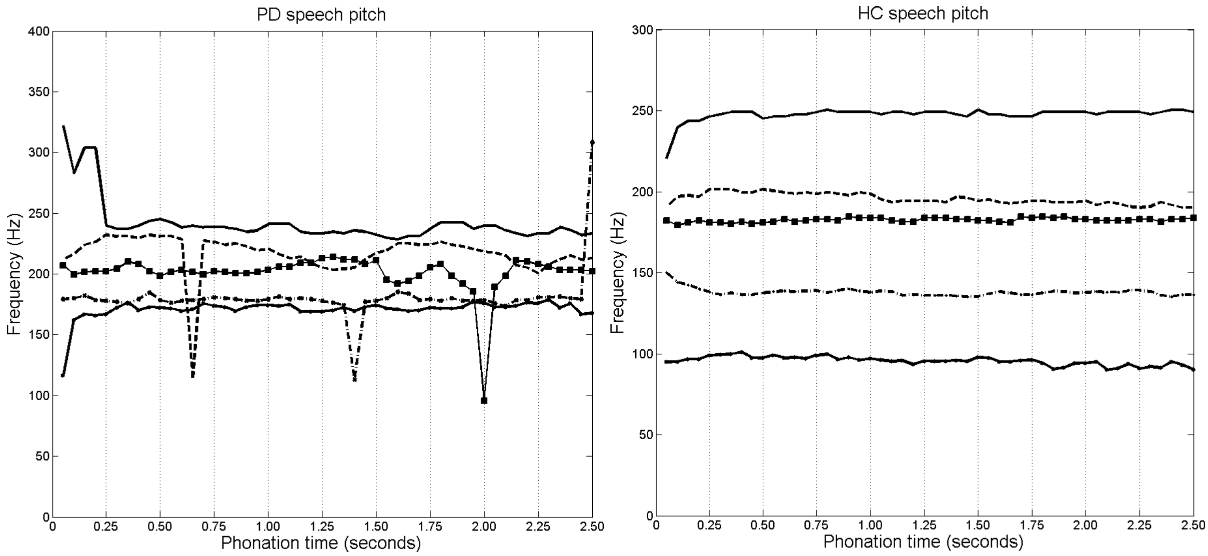

Figure 8 Left: pitch calculated for a PD patient. Right: Pitch calculated for a HC person

Another important result that is validated on the database is the behavior of the vocal triangle of PD patients and healthy controls. Figure 9 shows the difference between the obtained vocal triangles for three PD patients, respect to the triangle obtained from the mean values of all healthy controls (triangle with thicker lines). The compression of the VSA in the case of the PD patient is due to the loss of their articulatory capability.

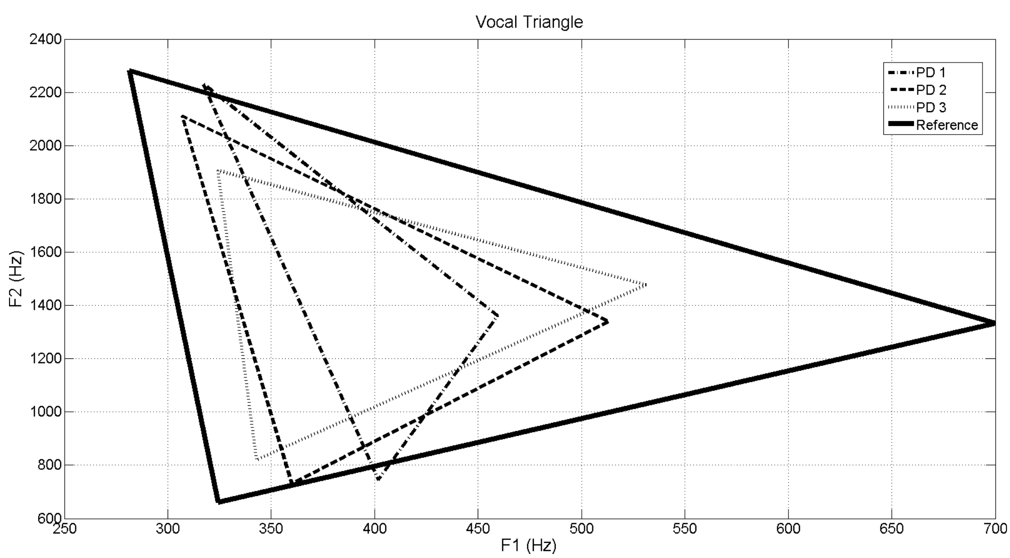

Figure 9 Vocal space area calculated on the Odroid U2

Respect to the NLD analysis, it is already demonstrated that the more pathologic is the speech signal, the more complex is the associated attractor [27]. It means that one can expect more chaotic attractors in the case of speech signals of people with PD than of healthy ones. Left side of figure 10 depicts the obtained attractor for one PD patient and right side of the same figure shows the result for one healthy person, both are 63 years old.

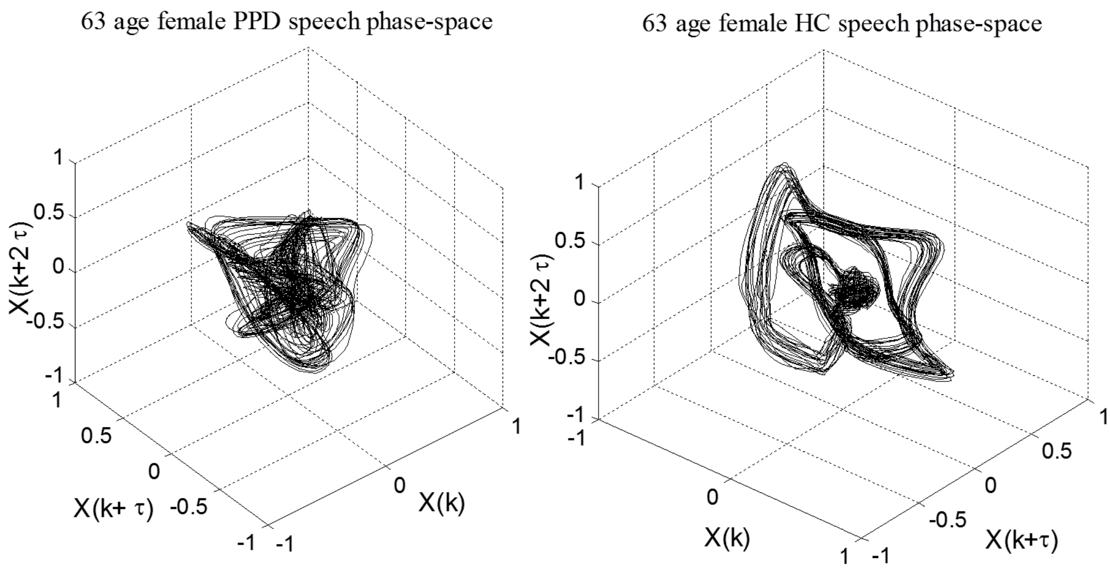

Figure 10 Phase space obtained from voice recordings (PD: left, HC: right)

Besides the graphical validation of speech disorders suffered by people with PD, in this paper the numerical differences of the features calculated for PD patients and healthy controls are also analyzed. figure 11 depicts box-plot diagrams where it is possible to appreciate the differences between the results obtained from the recorded PD patients and HC evaluated on the Odroid platform. For this analysis the participants of each class (PD and HC) are divided into two groups, one containing people younger than 61 and other with people older than 61. This division allows the analysis of possible age dependency in features. The threshold for the age division is set in 61 which is the median of the age distribution in the data.

Note that FCR and CD show the differences clearer than other features; however, as the systems presented here are conceived for the monitoring of patients with PD, the reference to compare each measure on one patient will be taken respect to his/her own speech recorded previously. Additionally, the inclusion of other measures is motivated by their proven utility in the analysis of disordered speech signals. Other important point is related to the medical interpretation of each measure; jitter, shimmer, VSA and FCR can be associated to different part of the speech production process. CD is included because of its discrimination capability, as can be observed in figure 11, and also in other works [6]; however, more research is required to find clinical interpretation of such kind of measures.

figure 11 also confirms that people with PD have higher jitter and shimmer values, showing the existence of instability in the control of the vocal folds vibration process. Note also that VSA values in PD patients are slightly lower than those exhibited in healthy people, this behavior also confirms the loss of articulatory capability of the people with PD.

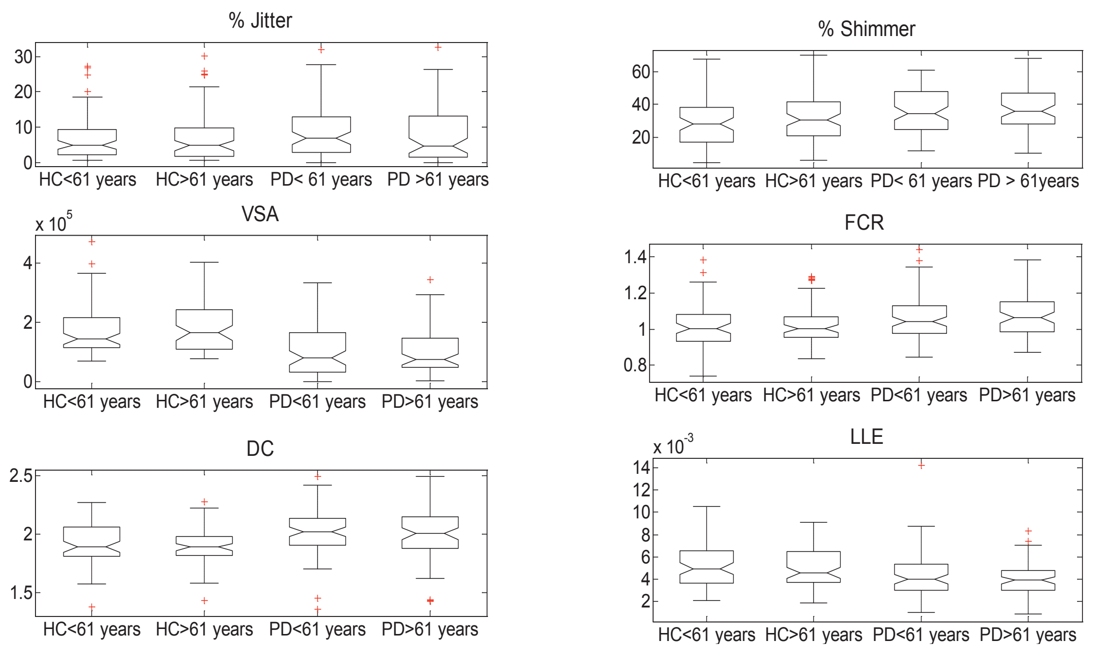

Figure 11 Differences of the features measured on PD and HC, calculated with the Odroid U2

The result of the evaluation of jitter and shimmer on each platform is presented in table 1, besides the results obtained with Praat are also included. Note that higher values of jitter and shimmer are observed for PD patients. The behavior of jitter values is also associated to the inability of the patients to hold the larynx muscles in a stable position for long periods of time [4] while the values of shimmer are high due to the breathing problems also present in people with PD [4].

Table 1 Mean and standard deviation of jitter and shimmer measured on each platform

FEATURE	***PLATFORM***							
	HC Matlab	***HC DSP***	***HC Odroid***	***HC Praat***	***PD Matlab***	***PD DSP***	***PD Odroid***	***PD Praat***
Jitter mean	8.38	7.65	6.98	5.31	12.92	12.16	9.82	8.94
Jitter Std. Dev	12.24	6.49	7.19	3.04	13.95	12.05	9.90	7.27
Shimmer mean	30.04	28.63	30.77	40.95	35.47	30.11	36.58	59.34
Shimmer Std. Dev	15.03	13.88	14.81	20.52	13.52	11.15	13.19	33.46

Table 2 shows the values of the articulation features, confirming that VSA value is lower in PD patients than in healthy people. This behavior can be observed on the four tested platforms: Matlab, DSP, Odroid and Praat. Respect to the FCR, as it is also observed in figure 11, people with PD exhibit values slightly higher than those obtained for HC.

Table 2 Mean and standard deviation of VSA and FCR measured on each platform

FEATURE	PLATFORM							
	HC Matlab	HC DSP	HC Odroid	HC Praat	PD Matlab	PD DSP	PD Odroid	PD Praat
VSA mean	158,530	176,960	178,920	151,780	123,210	91,502	102,880	100,040
VSA Std. Dev	101,600	94,660	84,197	85,106	107,630	80,596	80,516	96,802
FCR mean	0.9841	0.9913	1.0149	1.0146	1.0744	1.0539	1.0671	1.0612
FCR Std. Dev	0.1023	0.1176	0.1133	0.0823	0.1478	0.1138	0.1176	0.1137

The values obtained for the NLD features are presented in table 3. The discriminative capability of CD can be easily observed, while LLE is not able to discriminate between PD and HC recordings.

Table 3 Mean and standard deviation of CD and LLE measured on each platform

FEATURE	PLATFORM			
	HC Matlab	HC Odroid	PD Matlab	PD Odroid
CD mean	1.7951	1.9107	2.1474	2.0100
CD Std. Dev	0.3619	0.1731	0.2869	0.2083
LLE mean	0.0016	0.0021	0.0019	0.0021
LLE Std. Dev	0.0011	0.0019	0.0013	0.0018

In addition to the discriminative capability of each feature, the differences of the values obtained when features are calculated on different platforms are also analyzed. Table 4 shows the mean squared error obtained when the values of each feature are compared respect to the values obtained on the Odroid platform; comparison is made considering each class of recordings (HC and PD) separately. Note that in general, there are not important differences between the results on each platform. Comparisons among the results obtained on Odroid and Praat or DSP are not included for CD and LLE because these measures were not calculated on these platforms.

Table 4 Mean squared error between the values obtained with each platform and Odroid

Measure	% MSE between platforms		
	Odrdoid-Matlab	Odroid-DSP	Odroid-Praat
pitch HC	0.3556	0.3678	0.6983
pitch PD	0.7068	2.5774	1.8376
jitter HC	0.8620	0.5846	4.0236
jitter PD	0.8734	0.2435	0.3430
shimmer HC	0.013	1.2171	6.9484
shimmer PD	0.0205	0.4576	7.4644
VSA HC	1.6185	4.7003	5.4545
VSA PD	0.5671	2.1809	4.4031
FCR HC	0.7617	4.9223	1.3933
FCR PD	1.6362	4.8168	1.9757
CD HC	0.7907	-	-
CD PD	0.3095	-	-

It is important to note that the differences in the values of features when are measured using different platforms are mainly related to the precision that is used to represent such values on each system. In Matlab the values are computed in float format, while the system developed on the Odroid platform uses integer with 32 bits of precision [34].

Table 5 shows the p-value of difference mean tests (t-test) between HC and PD on each platform. These tests evaluate the discriminative capability of each feature in all of the considered platforms.

Table 5 P-value of t-test between HC and PD evaluated on each platform

Measure	*Odroid*	*Matlab*	*DSP*	*Praat*
Jitter	0.0498614	0.00600349	1.73E-06	4.03E-08
Shimmer	0.00039345	0.00113738	6.35E-07	2.34E-08
VSA	2.90E-14	0.00372965	6.29E-11	0.0193866
FCR	0.00011208	5.18E-05	3.66E-05	6.14E-05
DC	1.11E-05	0	-	-
LLE	1.76E-05	0.03341846	-	-

Conclusion

A new system for the automatic evaluation of speech signals of people with Parkinson's disease is presented. The process for the design and testing of such system is completed in three steps: first, a Matlab graphic user interface is designed; second, the digital signal processor TMS320C6713 is used; and third, the prototype which is based on the Odroid-U2 is presented.

Different acoustic and perturbation features that are well known for the analysis of speech of people with PD are implemented on the designed system. Differences between the values obtained from speech recordings of PD patients and HC are observed in several features. The problems on articulation, phonation and breathing of people with PD are validated in the results.

Considering that the portable device is running on real time and that it is also equipped with a LCD monitor, the patient will receive a bio-feedback about his/her speech therapy. This functionality will improve the effectiveness of the therapy and will motivate the patient to perform the exercises in the best way every day.

From the point of view of the speech therapists, this device will provide them with a new tool for the continuous monitoring of their patients. Since the device allows recording the speech signals and the obtained results after doing the exercises, the therapist will be able to monitor the evolution of the speech therapy in their patients objectively.

The future work will be focused on the robustness of the system, in order to improve its accuracy and usability. Likewise, the inclusion of other information sources such as gait and writing analysis will be considered in the future in order to provide a complete multimodal platform for the analysis of patients with movement disorders.

Acknowledgments

Juan Rafael Orozco Arroyave is under grants of "Convocatoria 528 para estudios de doctorado en Colombia 2011" financed by COLCIENCIAS. This work is partially funded also by COLCIENCIAS through the project # 111556933858. Juan Camilo Vásquez Correa is granted by the program of young researchers and innovators 2014, financed by COLCIENCIAS.

References

1. M. de Rijk, L. Launer, K. Berger, M. Breteler, J. Dartigues, M. Baldereschi, L. Fratiglioni, A. Lobo, J. Martinez, C. Trenkwalder, A. Hofman. "Prevalence of parkinson's disease in europe: a collaborative study of population-based cohorts". *Journal of Speech Language and Hearing Research*. Vol. 54. 2007. pp. 21-23.

2. F. Martínez. "Trastornos del habla y la voz en la enfermedad de Parkinson". *Neurología*. Vol. 51. 2010. pp. 542-550.

3. S. Skodda, W. Visser, U. Schlegel. "Vowel articulation in Parkinson's diease". *Journal of Voice*. Vol. 25. 2011. pp. 467-472.

4. J. Mekyska, I. Rektorova, Z. Smekal. *Selection of Optimal Parameters for Automatic Analysis of Speech Disorders in Parkinson's Disease*. Proceedings of the 34th International Conference on Telecommunications and Signal Processing (TSP). Budapest, Hungría. 2011. pp. 408-412.

5. O. Geman. *Data Processing for Parkinson's Disease: Tremor, Speech and Gait Signal Analysis*. Proceedings of the 3rd International Conference on E-Health and Bioengineering (EHB). Iaşi, Romania. 2011. pp. 1-4.

6. A. Tsanas, M. Little, P. McSharry, L. Ramig. "Accurate Telemonitoring of Parkinson's Disease Progression by Noninvasive Speech Tests". *IEEE transactions on Biomedical Engineering.* Vol. 57. 2010. pp. 884-893.

7. J. Orozco, J. Arias, J. Vargas, E. Nöth. "Perceptual Analysis of Speech Signals from People with Parkinson's Disease". J. Ferrández, J. Álvarez, F. de la Paz, F. Toledo (editors). *Natural and Artificial Models in Computation and Biology*. *Lecture Notes in Computer Science.* Vol. 7930. Ed. Springer Verlag. Mallorca, Spain. 2013. pp. 201-211.

8. E. Belalcazar, J. Orozco, J. Vargas, J. Arias, E. Noth. "New Cues in Low-Frequency of Speech for Automatic Detection of Parkinson's Disease". J. Ferrández, J. Álvarez, F. de la Paz, F. Toledo (editors). *Natural and Artificial Models in Computation and Biology*. *Lecture Notes in Computer Science.* Vol. 7930 . Ed. Springer Verlag. Mallorca, Spain. 2013. pp. 283-292.

9. J. Orozco, J. Arias, J. Vargas, E. Noth, "Analysis of Speech from People with Parkinson's Disease through Nonlinear Dynamics". T. Drugman, T. Dutoit (editors). *Advances in Nonlinear Speech Processing*. *Lecture Notes in Computer Science.* Vol. 7911 . Ed. Springer Verlag. Mons, Belgium. 2013. pp. 112-119.

10. J. Zicker, W. Tompkins, R. Rubow, J. Abbs. "A Portable Microprocessor-Based Biofeedback Training Device". *IEEE transactions on Biomedical Engineering.* Vol. 27. 1980. pp. 509-515.

11. R. Rubow, E. Swift. "A microcomputer-based wearable biofeedback device to improve transfer of treatment in parkinsonian dysarthria". *Journal of Speech and Hearing Disorders.* Vol. 50. 1985. pp. 178-185.

12. L. Ochs. *Self adjusting bio-feedback method and apparatus*. US Patent 4.461.301, 24 Jul. 1984 Washington, USA. 1984. pp.1-10

13. J. Dembowski, B. Watson. "An instrumented method for assessment and remediation of stuttering: A single-subject case study". *Journal in fluency disorder.* Vol. 16. 1991. pp. 241-273.

14. A. S. AlMejrad. "Design of an intelligent system for speech monitoring and treatment of low and excessive vocal intensity". *Artif Life Robotics.* Vol. 15. 2010. pp. 320-324.

15. M. Wirebrand. *Real-time monitoring of voice characteristics using accelerometer and microphone measurements.* Master Thesis, Linköping University. Linköping, Sweden. 2011. pp.1-46.

16. A. Carullo, A. Vallan. "Design Issues for a Portable Vocal Analyzer". *IEEE Transactions on instrumentation and measurement.* Vol. 62. 2013. pp. 1084-1093.

17. B. Resch, M. Nilsson, A. Ekman, W. Kleijn. "Estimation of the Instantaneous Pitch of Speech". *IEEE Transactions on Audio, Speech, and Language Processing.* Vol. 15. 2007. pp. 813-822.

18. C. Shahnaz, W. Zhu. *A new Technique for the Estimation of Jitter and Shimmer of Voiced Speech Signal*. Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering (CCECE). Ottawa, Canada. 2006. pp. 2112-2115.

19. S. Sapir, L. Ramig, J. Spielman, C. Fox. "Formant Centralization Ratio (FCR): A proposal for a new acoustic measure of dysarthric speech". *Journal of Speech, Language and Hearing Research.* Vol. 53. 2010. pp. 114-133.

20. P. Boersma, D. Weenink. *Praat: doing phonetics by computer*. Available on: http://www.fon.hum.uva.nl. Accessed: Sept. 18, 2013.

21. P. Boersma. "Accurate Short-Term Analysis of the Fundamental Frequency and the Harmonics-to-Noise Ratio of a Sampled Sound". *Institute of Phonetic Sciences, U. of Amsterdam.* Vol. 17. 1993. pp. 97-110.

22. M. Kammoun, D. Gargouri, M. Frikha, A. Hamida. *Linear Prediction Method Evaluation in Speech Formant Frequencies Estimation*. Proceedings of the Third International Conference on Systems, Signals & Devices. Sousse, Tunisia. 2005. pp. 1-5.

23. R. Kent, Y. Kim. "Toward an acoustic typology of motor speech disorders". *Clinical Linguistics & Phonetics.* Vol. 17. 2003. pp. 427-445.

24. H. Herzel, D. Berry, I. Titze. "Analysis of Vocal Disorders With Methods From Nonlinear Dynamics". *Journal of Speech and Hearing Research.* Vol. 37. 1994. pp. 1008-1019.

25. F. Takens. "On the numerical determination of the dimension of an attractor". *Lecture Notes in Mathematics.* Vol. 1125. 1985. pp. 99-106.

26. P. Grassberger and I. Procaccia. "Characterization of Strange Attractors". *Physical review.* Vol. 50. 1983. pp. 346–349.

27. J. Jiang, Z. Yu, M. Clancy. "Chaos in voice, from modeling to measurement". *Journal of Voice.* Vol. 20. 2006. pp. 2-17.

28. V. Oseledec. “A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems”. *Transactional. Moscow Mathematical. Society.* Vol. 19. 1968. pp. 197-231.

29. R. Chassaing, D. Reay. *Digital Signal Processing and Applications with TMS320C6713 and TMS320C6416 DSK*. 2nd ed. Ed. Wiley Interscience. New Jersey, USA. 2008. pp. 283-290.

30. Hardkernel, *Hardkernel ODROID*. Available on: www.hardkernel.com. Accessed: Sept. 15, 2013.

31. H. Langtangen. *Python Scripting for Computational Science.* 3rd ed. Ed. Springer Publishing Company. Berlin, Germany. 2009. pp. 1-70.

32. R. Johansson. *Lectures on scientific computing with Python*. Available on: https://github.com/jrjohansson/scientific-python-lectures. Accessed: Sept. 15, 2013.

33. docs, Python. *Multiprocessing Process-based threading interface.* 2013 Available on: https://docs.python.org/2/library/multiprocessing.html. Accessed: Jul. 21, 2013.

34. J. Glover, V. Lazzarini, J. Timoney. *Python for audio signal processing*. Proceedings of the Linux Audio Conference. Maynooth, Ireland. 2011. pp 1-8.