Design of intelligent system for speech monitoring and treatment of low and excessive vocal intensity

Ali S. AlMejrad

Department of Biomedical Technology, King Saud University
P.O. Box 10219, Riyadh 11433, Kingdom of Saudi Arabia
(amejrad@ksu.edu.sa)

Abstract: Number of people and especially children suffer from speech problems such as too low or often excessive vocal intensity which is one of the behavioral problems that may be encountered in people with intellectual disability. Also it may happen due to the aggressive behaviors which produce disturbance which is rejected by others. The persons with this behavior may often be unaware of their problems. To limit these negative outcomes an intelligent system has been designed for monitoring, to provide a patient with information about speech intensity outside the clinic and for controlling of the treatment of excessive and low vocal intensity without the need to the continuous therapeutic follow-up in speech pathology clinic. The use of the intelligent system proves to be an effective due to the many features of this system. The technical design considerations, the system features and system evaluation will be discussed.

Keywords: - intelligent system, speech monitoring, vocal intensity

I. INTRODUCTION

There are common vocal cords disorders include vocal cord nodules, vocal cord polyps, contact ulcers, laryngitis, vocal cord tumors, vocal cord paresis in addition to vocal loading [1]. Speech impairments such as low or reduced vocal intensity which present in patient with Parkinson's disease is generally been classified as a dysarthria which refers to motor defects of speech resulting from lesions of the nervous system [2, 3]. Others have described this symptom as inadequate volume, soft voice, and weak voice. Many parkinsonian patients thus have problems normally regulating the volume of their speech. However, if instructed to speak louder, the patient can regulate the speaking volume but unfortunately such improvement generally is not long-term one. On the other side, excessive vocal intensity due to speaking loudly in addition to the vocal abuse and misuse may result in vocal cord pathologies such as vocal nodules. This pathology occurs in more than 50% of children with voice disorders [4, 5]. This pathology is one of the behavioral problems that may be encountered in people with intellectual disability [6]. Also it may happen due to the aggressive behaviors which produce disturbance which is rejected by others.

Usually the persons with the above behaviors may often be unaware of their problems, so intervention program is required to limit these negative outcomes. Speech monitoring is necessory in the first instance in addition to the use of biofeedback training. Biofeedback is that technology which can learn to exercise some control over a physiological process if information related to that process immediately available. Compared to the simple verbal prompting by therapist such as (" you are speaking too low or too loudly") which failed to regulate their speech the use of biofeedback instead is more effective. Zicker et el. used feedback instrument for speech therapy with Parkinson's disease patients [7]. This system is a general-purpose feedback training device which is based on the microcomputer and some additional hardware interfaces of 1 kg weight. Prosek et al. and Schliesser use the biofeedback training employing electromyography (EMG) to voice disorders [8, 9]. McGillivray et al. developed a biofeedback analog device to reduce excessive vocal intensity and treat a case of a child with soft nodules on her vocal cords by interrupt her conversation with a loud tone when her speech exceeded an intensity level determined by the therapist [10]. This device is simple and has limited functions of only tone. The use of a portable auditory-feedback and vibratory-feedback devices was used by Lancioni [11, 12] for reducing excessive vocal loudness in persons with mental retardation and a deaf woman [13].

To limit these negative outcomes an intelligent system has been proposed for speech monitoring, to provide a patient with information about speech intensity outside the clinic and for controlling of the treatment of both excessive and low vocal intensity without the need to the continuous therapeutic follow-up in speech pathology clinic. The intelligent system which will be based on microcontroller technology will be an effective due to the many features of this system such as portability and wearily due to its small size, ease of programmability, patient alert, automatic tracking of up normal occasions during therapeutic sessions. The technical design considerations, the system features and system evaluation will be discussed.

II. MATERIALS AND METHODS

Hardware: The proposed intelligent speech monitoring system in this work consists of many components including: transducer, amplifiers, filters, rectifier, unity gain buffer, comparators, and integrator in addition to a low-cost microcontroller; auditory and vibratory feedback. Capacitors and headers were added to the circuit design and operational amplifiers were selected to make the system battery operated as shown in Figure 1.

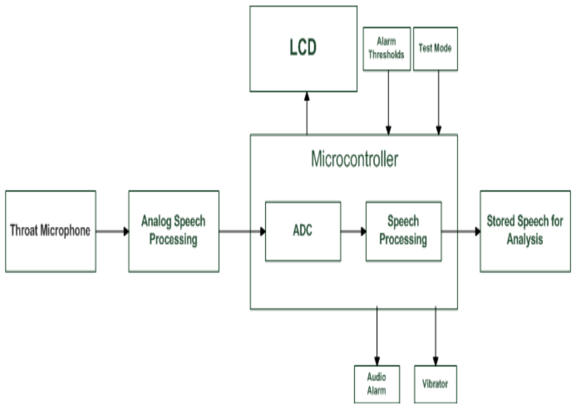

Figure 1. Block diagram of the intelligent speech monitoring system

The objective of the analog circuitry is to take the speech signal coming from the microphone, band pass filter it, amplify it, and make every output value positive [14]. The analog signal conditioning consists of filter to eliminate unwanted noise, main Amplifier with an adjustable gain to make the signal bigger, rectifier to turn negative signal values to positive (absolute value) and integrator to evaluate voice intensity over a set time period and determines whether or not alarm triggers. As a test only two comparators are used to compare input (speech) with specific threshold(s). The most important issue was determining the values for the corner frequency of the filter. A low pass filter with a 500 Hz corner frequency was considered because this frequency was low enough to include the volume or intensity aspects of the voice. It would cut off the high frequencies that are associated with pitch, quality, or uniqueness of a particular speech. Since the typical values expected were on the order of 10 mV and the circuitry components needed to work with volts, the gain of the amplifier was designed to be 100. The operational amplifier is selected based on its minimal operating voltage requirement due to

Speech signals from a throat microphone is sampled and passed through analog circuitry to filter, amplify and rectify the speech signal to convert it into appropriate and useable input for the analog-to-digital converter (A/D) as shown in Figure 2.

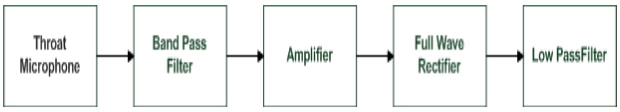

Figure 2. Block diagram of the Analog speech signal processing

Limited space permitted for batteries. The LM358 was selected because it required only 3V operating voltage. It was necessary that the signal prior to the integrator be rectified in order to prevent the negative and positive values from equaling zero. The integrator is designed using a setup resembling a low pass filter to allow the signal logarithmically to serve the intended purpose. For test only two separate analog comparators are used to properly compare the input values to the set thresholds.

After conditioning the speech signal using analog circuitary it passes through the microcontroller 8-bit A/D converter to digitize the speech signal for processing and performing speech intensity level calculations on the sampled speech signal using microcontroller based system (PIC16F877A from Microchip Technology Inc.). The feedback consisting of an audio alarm (a beeb sound) and a vibrator (a vibrating motor of weight of 2.5 g from Korea Partner IND. CO.. LTD.) are used and controlled by the microcontroller. The purpose of the feedback device is to warn the patient when his or her speech intensity had gone below or above the specified threshold specified by therapist. With a minimal amount of feedback, the patient would be notified that vocal changes and regulation of speech were needed to be made. It is believed that tactile feedback devices are more appropriate to use in this case than non-tactile devices. Therefore, a vibrating motor, similar to those used in commercial pagers, is used. The pager is selected because of its smallest size, most rigid, mounting capability and cost. The pager could be worn on the waist that gives off enough vibration to alert the user that changes in the speech volume need to be made.

Finally, the microcontroller is programmed to process the speech data gathered and to determine when feedback to the patient is necessary.

Software: a usable algorithm, a step-by-step problem-solving procedure, has been written with C and encoded into assembly program that the microcontroller can read and run as shown in the flow chart below (Figure 3).

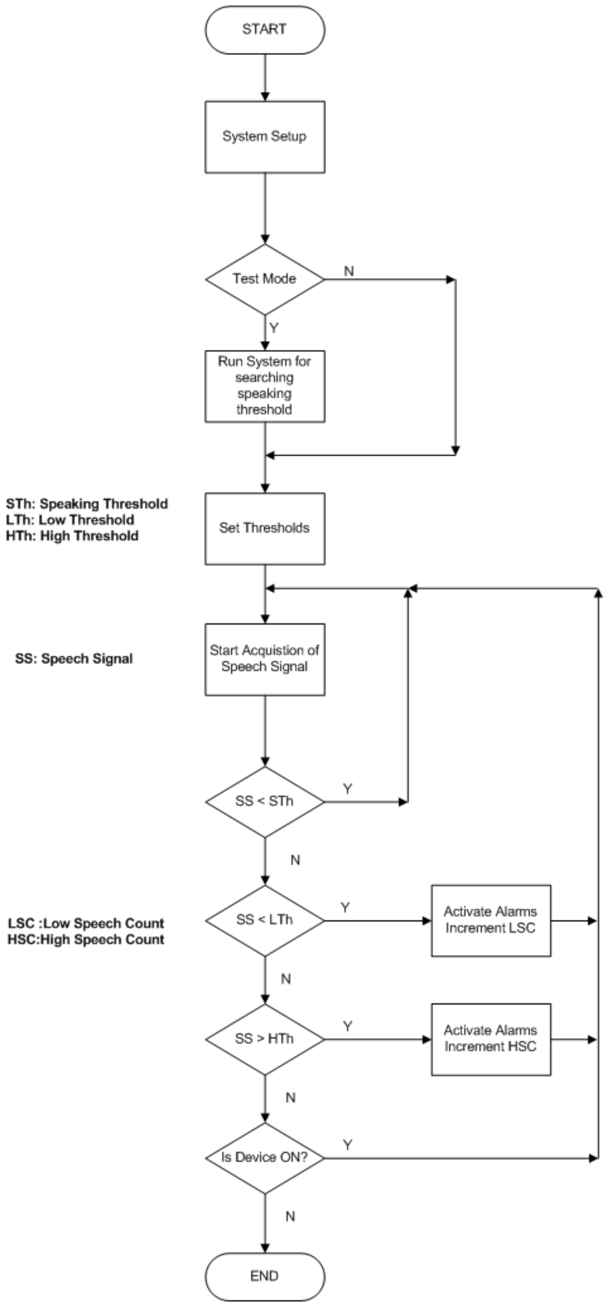

Figure 3. Speech processing of vocal intensity and biofeedback activation flowchart

The microcontroller receives one input signal and can send one output signal. This input is a steady varying voltage signal, which is representative of the patient's speech intensity. This speech signal is then converted from an analog signal (measured in volts) to a digital signal (in binary) using A/D. The lower or louder the patient speaks, the lower or greater the intensity measured by the microcontroller. The microcontroller then utilizes the algorithm to process the inputted data derived by the voltage signal and decides whether or not feedback should be sent to the patient. Feedback of an audio and vibration alarms is the result of the patient speaking too softly or loudly for an extended period of time, which at this time is assumed to be one-half second.

The speech processing algorithm required by the microcontroller of the speech feedback device needed to be flexible and accurate [15]. This means the device could be configured to each patient's needs in order to minimize the amount of false feedback given to the patient. The algorithm implemented in this design aims to meet these criteria. In order to avoid giving false responses, the user must, before using the device, sample the background noise level by pushing a button on the device that activates the test mode. While in the test mode, the intensity of the current ambient noise is determined, which determines a minimum speaking threshold-equivalent to the intensity measurement at which the user is not speaking. In other words, when the noise intensity measured by the device is less than this minimum threshold, the user is not speaking and all readings below this intensity are disregarded. Only those measurements in which the speech intensity is above the speaking threshold are used for further processing.

An alarm threshold, preset by a speech therapist, based on the patient's requirements is determined and programmed into the device prior to use. This alarm threshold is the intensity level the patient needs to speak above or below. An integrator and counter have been used in order to determine when the device must send feedback to the user. During times in which the user is not speaking, the counter is held constant, while speaking below the threshold, the counter is decreased by one, and while speaking above the threshold; the counter is increased by one. When the user's speech intensity falls below the low alarm threshold for duration of time that allows the counter to reach the preset alarm threshold, a signal is sent to the feedback device that alerts the patient that they must speak with greater intensity. The same process is repeated when the user's speech intensity exceeds the high alarm threshold for duration of time that allows the counter to reach the preset alarm threshold, a signal is sent to the feedback device that alerts the patient that they must speak with lower intensity The counter is reset after every two seconds to be reevaluated during testing time frame in which the microcontroller is inputting the patient's speech intensity.

At first, the code will determine a speaking threshold when the button is pushed. It will try to get the maximum intensity of the background noise over a range of time. The speaking threshold is then set around 80 mV above this maximum value as the speaking threshold. In addition to that, the speaking threshold is made to be at max of a certain value. This is for anticipating if there is a sudden and very high intensity noise, which is unusual.

After that, it will start processing the signal received. If the signal received is below the speaking threshold, it will repeat taking another value of signal before it try to compare it with the appropriate speech alarm thresholds. Each signal value obtained will be stored in the additional RAM. Further more, if the speech signal value is below or above the appropriate speech alarm threshold, it will increase a counter. On the other hand, when it is normal, it will decrease the counter. If the counter is beyond a certain limit, it will set the alarm to be high. The method that we are using to fill the addresses in the memory is called first in first out (FIFO). In this way the memory is filled and refilled circularly thus reusing the allocated memory addresses. After processing and storing the first set of samples, it will loop back to process and store another set. The data replaced by a new value will be read and compared to the alarm appropriate speech threshold to update the counter of alarm into a value which the counter should be at, if the data is not there. Then, it will take another sample to define the new set and then determine whether to set the alarm or not. This loop repeats until the person turns off the device.

III. TESTING AND RESULTS

In this study a complete simulation for evaluation was performed. To test the system that should work in theory, an initial prototype was created using a breadboard to evaluate the connections to each unit of the system. Each stage of the intelligent system was implemented, tested and calibrated separately. The processing stage using the programmable microcontroller was simulated firstly using different levels of speech intensity of speech-like signals. Secondly it was tested by different normal subjects who were asked to speak by reading many statements for few minutes using different levels of speech intensity and test to speak softly or loudly a few seconds and check if the biofeedback activates the audio and the vibrator alarms attached to the patient. To make accurate tests a sound level meter above the speech monitoring system was used to monitor the volume of sounds and speech intensity to adjust the biofeedback to be activated when low or excessive speech intensity. In addition to that two comparators are used to compare input (speech) with specific threshold(s) for test only. After calibration, the system was tested for various levels of speech intensity and in each test the system was successful.

IV. CONCLUSION AND FURTHER WORK

The proposed intelligent speech monitoring system developed as a prototype achieved the requirement of this work such as the portability, programmability and low power consumption. Our intelligent speech monitoring system performed satisfactorily as a prototype system for monitoring, to provide a patient with information about speech intensity outside the clinic and for controlling of the treatment of excessive and low vocal intensity without the need to the continuous therapeutic follow-up in speech pathology clinic. The system is able to collect speech intensity signals through a throat microphone and able to discriminate between speech and silence and is only active when the subject is speaking. The device activities an alarm and vibration as a biofeedback to the subject whenever the speech intensity falls below or exceeds above an adjustable thresholds for a preset time interval. The system is able to record and count up normal occasions and display them in real-time.

The system can be enhanced by additional memory to store information concerning time dependent data of the user's speech intensity when alarms occurred. Following use, this data could then be downloaded to a personal computer (PC) where it can be analyzed by a speech therapist. This additional information would allow not only real time feedback, but also the speech therapist could use it to monitor the patient's progress and make changes to the therapy program if it is necessary.

Clinical evaluations of the developed system based on specific procedures using different patients who have speech disorders such as Parkinson's disease and patients suffering from the problem of excessive speech are required to test its feasibility to be used as monitoring and therapeutic device.

ACKNOWLEDGEMENT

This work is supported by the grant from Applied Medical Research Center, the Deanship of Scientific Research, King Saud University. Grant Code: PS/39/26

REFRENCES

[1] Greene, M., Lesley M., (2001), *the* Voice and its Disorders, John Wiley & Sons; 6th Edition.

[2] de Swart BJ, Willemse SC, Maassen BA, Horstink MW, (2003), Improvement of voicing in patients with Parkinson's disease by speech therapy, Neurology. 60(3):498-500.

[3] Spielman J, Ramig LO, Mahler L, Halpern A,

Gavin WJ., (2007), Effects of an extended version of the lee silverman voice treatment on voice and speech in Parkinson's disease, Am J Speech Lang Pathol.;16(2):95-107.

[4] Hooper CR., (2004), Treatment of voice disorders in children, Lang Speech Hear Serv Sch.; 35(4):320-6.

[5] Wilson, D., 1987, Voice problems of children, Williams and Wilkins Co., pp. 33-35.

[6] Lancioni GE, van Houten K, ten Hoopen G., (1997), Reducing excessive vocal loudness in persons with mental retardation through the use of a portable auditory-feedback device, J Behav Ther Exp Psychiatry.;28(2):123-8.

[7] Zicker JE, Tompkins WJ, Rubow RT, Abbs JH, (1980), A portable microprocessor-based biofeedback training device, IEEE Trans Biomed Eng.;27(9):509-15.

[8] Prosek, R, Montgomery, A. Walden, B., and Schwartz, D., (1978), EMG biofeedback in the treatment of hyperfunctional voice disorders, J. Speech Hear. Disorders, XLIII, pp. 282-294.

[9] Schliesser, H., (1987), EMG biofeedback as a function of voice-quality change, J. Percept. Motor Skills, 64, pp. 719-724.

[10] McGillivray R, Proctor-Williams K, McLister B., (1994), Simple biofeedback device to reduce excessive vocal intensity, Med Biol Eng Comput.; 32(3):348-50.

[11] Lancioni GE, Brouwer JA, Markus S., (1995), A portable visual-feedback device for reducing excessive vocal loudness in persons with mental retardation, Percept Mot Skills;81(3 Pt 1):851-7.

[12] Lancioni GE, Markus S., and Behrendt, M. (1998), A portable vibratory-feedback device for reducing excessive vocal loudness: A case study, Behavioural and Cognitive Psychotherapy, 26, 371-376.

[13] Lancioni GE, Markus S, (1999), a deaf woman learning to control her excessive vocal loudness through a portable feedback system, Percept Mot Skills.; 88(3 Pt 2):1347-9.

[14] Webster, J., (1997), Medical Instrumentation Application and Design, 3rd Edition, Wiley.

[15] Rainer, L. and Schafer, R., (1978), Digital processing of speech signals, Englewood Cliffs, NJ, PH.