Microcontroller Based System for Voice Disorders Detection

I Nyoman Kusuma Wardana
Sekolah Tinggi Manajemen Informatika dan Teknik Komputer (STMIK) STIKOM Bali
Jalan Raya Puputan No.86 Renon, Telp: (0361)244445, Denpasar-Bali
Email: kusumawardana@stikom-bali.ac.id

Abstrak

Penelitian ini terkait pada usaha untuk menciptakan alat monitoring aktivitas vokal berbasis mikrokontroler. Selanjutnya, alat ini dapat digunakan sebagai pendeteksi dini kekacauan suara (voice disorders). Penggunaan alat ini dirasa penting untuk pengguna yang berkepentingan. Aktivitas utama penelitian adalah mendesain sebuah sistem data logger, baik yang mencakup perangkat keras dan perangkat lunak. Perangkat keras mencakup sistem logger, untai pengkondisian sinyal, dan untai catu daya.Data yang diperoleh oleh mikrokontroler selanjutnya akan dianalisis menggunakan software MATLAB. Alat dirancang untuk bekerja pada mode normal dan kalibrasi. Analisis mencakup kalibrasi terhadap Sound Pressure Level (SPL) dan estimasi frekuensi fundamental (f0) yang merupakan parameter dasar untuk memonitor aktivitas vokal. Berdasarkan penelitian yang telah dilakukan, diperoleh konstanta kalibrasi K sebesar 1.76 Pa/V (dengan standar deviasi 0.12%) dan rata-rata error untuk ECM sebesar -1.8dB (dengan standar deviasi 4.6 dB) terhadap mikrofon standar. Selanjutnya, setelah diperoleh data secara secara teknis, diperlukan kerjasama berbagai pihak yang terkait untuk merepresentasikan data tersebut kedalam tingkat kelelahan suara.

Kata kunci:*voice disorder, mikrokontroler, frekuensi fundamental, sound pressure level*

Abstract

The research was concerned with creating a device for vocal signal monitoring based on microcontroller application. The device was implemented to monitor the vocal signal in order to prevent a negative impact so called voice disorder. The availability of a device for vocal signal monitoring can be a benefit to the concerned users. The principal work was creating a suitable logger system, including software and hardware parts. The hardware part consists of logger system, conditioning circuitry, and power circuitry. The data obtained from the microcontroller was analyzed off-line using MATLAB. The device was organized to work in both normal and calibration modes. The analysis consisted of Sound Pressure Level (SPL) calibration and fundamental frequency (f0) estimation which are the most important parameters used in voice monitoring systems. According to the experiment, the calibration constant K was 1.76 Pa/V (0.12% standard deviation) and error average of the ECM was -1.8 dB (4.6 dB standard deviation) compared to the reference microphone.Collaboration works between engineers and physicians are advisable for a correct use of the estimated parameters.

Keywords*: voice disorder, microcontroller, fundamental frequency, sound pressure level*

1. Introduction

1.1 Voice Disorders

A voice disorder can be defined as a problem involving abnormal pitch, loudness or quality of the sound produced by the larynx [1] and it is closely related to disturbances of the dynamic patterns of vocal folds [2]. The cause of a voice disorder may be attributable to a wide range of structural, medical, neurological, or behavioral conditions, but it can simply be categorized as follow:

a. vocal abuse or misuse, including excessive talking, throat clearing, coughing, smoking, etc.
b. organic causes, including vocal cord paralysis, structural anomalies, etc.
c. functional causes including allergies, respiratory diseases, etc.

Almost every disorder may present in more than one symptom and one single symptom cannot be associated with one specific voice disorder. However, the common symptoms of the voice disorders are hoarseness or breathiness, decrease in pitch range and loudness, loss of voice, increased strain to speak, and tension in neck muscle [1].

Approximately 3% to 10% of the general population will experience a voice disorder in their lifetime, but certain individuals are at greater risk, including professional voice users (i.e. teachers, singers, clergy, coaches, attorneys), smokers, individuals who have had surgery to the head or neck, or individuals who have neurological disorders (i.e. stroke, spasmodic dysphonia, Parkinson's Disease) [3],[4],[5].

Many voice disorders are underestimated or even ignored by most people, many of whom are usually unaware of the risks or possible illnesses, e.g. the presence of nodules on the vocal folds. The impact of a disturbance in voice production can affect normal speech communication and result in consequences at a social, professional and personal level.

1.2 Vocal Signal Monitoring

The voice disorder can be identified early to avoid the further risks. By conducting this research, the issues related to the application of microcontrollers for vocal signal monitoring can be discovered. It is considerably useful to give suggestions for the following research. Based on the facts explained above, a prototype of the vocal signal monitoring that allows voice disorder to be identified is advisable to be studied. Through this research, the capabilities of the microcontroller are explored. The main objectives of the research are obtaining a prototype device for vocal signal monitoring based on microcontroller, obtaining the appropriate system for logging the vocal signal, and obtaining the calibration functions for the logger system.

The issues about creating devices for vocal signal monitoring can bring to varies works. The final product is restricted to build a prototype for vocal signal monitoring based on microcontroller.Therefore, the principal work is to create a suitable logger system, including software and hardware parts.The created software is dedicated to control the microcontroller. Once the data is obtained and saved in a storing device, it can be off-lined processed with another software to estimate the useful parameters.

1.3 Hardware Requirements

Apart from the software sections, the hardware parts strongly determine the overall systems. Both software and hardware sections should be carefully organized. Since the main purpose of the experiment is to create a portable device for vocal analyzer, the final product should be user friendly (wearable, comfort, small, light, etc.). The sensor bandwidth has to be large enough for the voice spectrum and has a capability to sense the input voice and converts it to acceptable voltages for the conditioning circuit. The conditioning circuit supports a suitable system for input voice from the sensor and it properly provides signal outputs for the logger system (microcontroller and storage device).The logger system has a capability to sample the input signals with a proper sampling rate and saves the results on the storage device.

The logger system consists of a microcontroller and a storing device. Due to its functional capabilities and physical properties, microcontroller can be a solution for any portable device that requires programmable tasks. To support its tasks, the microcontroller is equipped with several peripheral devices such as communication ports, power jack, storing device, etc. However, specific for this research, the required capabilities of the microcontroller and its supporting devices can be summarized as follows:

1. The microcontroller has to be equipped with a built-in Analog-to-Digital Converter (ADC). Moreover, ADC has a capability to sample the input data up to 40 kSPS. This value can be reach when only single channel is performed. Due to the calibration process, the dual-channel sampling is required, and this sampling rate value will reduce about a half of the single-channel sampling rate. The value of 40kSPS is considered as an acceptable value.
2. The microcontroller board has to be equipped with an external storing device. The external storing device is used to store the large amount of possible data of the vocal signals so they can be processed later on (off-line).

1.4 Software Requirements

The software provides instructions for the microcontroller to acquire the voice signals from the sensor(s), andprocesses the information obtained from the microcontroller. This work deals with the task on how to provide appropriate instructions for the microcontroller, how to sample the raw data and how to collect and manage them to the storing device. In the other hand, the second functionality of the software is to process and to interpret the raw data obtained from the microcontroller.

One of the the unique requirement for the software algorithm is the ADC programming. The voice signals have to be sampled fast enough by the ADC. The sampling theorem states that for a limited bandwidth (band-limited) signal with maximum frequency f_{max}, the equally spaced sampling frequency f_s must be greater than twice of the maximum frequency f*max*, i.e.,

$$f_s > 2\cdot f_{max} \tag{1.1}$$

in order to have the signal be uniquely reconstructed without aliasing. The frequency $2\cdot f_{max}$ is called the *Nyquistsampling rate*. Half of this value, f_{max}, is sometimes called the *Nyquist frequency*[6],[7].

2. Research Method

2.1 Electronics Prototyping Platform

According to the hardware requirements, the selected device for the logger system was Arduino board. Arduino is an open-source electronics prototyping platform [8], [9]. It is designed as a physical or embedded computing platform that has ability to interact with its environment through the use of hardware and software. Among many variants of the Arduino boards, the Arduino Ethernet is selected. It is based on the Atmel’s ATMEGA328 and equipped with onboard microSD reader. It supports SPI communications to connect the microcontroller and the SD card.

2.2 Conditioning Circuitry

The conditioning circuit provides suitable outputs for the inputs of the microcontroller (0-5 Volts range). It consists of a 12 Volts phantom power for electret condenser microphone (ECM) *laryngophone*, amplifier with proper gain, and a low-pass filter and a high-pass filter. The overall conditioning circuit for the ECM and schematic of calibration circuit are shown in Figure 1 and Figure 2, respectively.

For the *Sound Pressure Level* (SPL) calibration purpose, a condenser microphone Behringer ECM8000 is used as a reference. Slightly modification for conditioning circuit is done. Instead of a 12 Volts voltage source shown in Figure 1, the external MicroPOWER PS400 is used as a phantom power for the microphone. An AD622 from Analog Devices, as seen in Figure 1, is chosen as an amplifier circuit. The AD622 is a low cost, moderately accurate instrumentation amplifier that requires only one external resistor to set any gain between 2 and 1000.

The output signals produced by the conditioning circuit in Figure 1 will be centered on 0 Volts. Since the output voltages of the microphone are bipolar signals (contains positive and negative signals), the amplification will also be bipolar signals. These values are not suitable as inputs for the microcontroller. To fix this problem, the amplifier outputs should be shifted around the proper reference. The value of 2.5 Volts is chosen as a reference because the acceptable input range for the microcontroller is 0 to 5 Volts.

REF03 from Analog Devices is used as a reference voltage IC. It is mentioned on the datasheet that the REF03 voltage reference provides a stable 2.5 Volts output with minimal change in response to variations in supply voltage, ambient temperature or load conditions [10]. The REF03 is designed to operate with input voltage range 4.5 to 36 Volts.

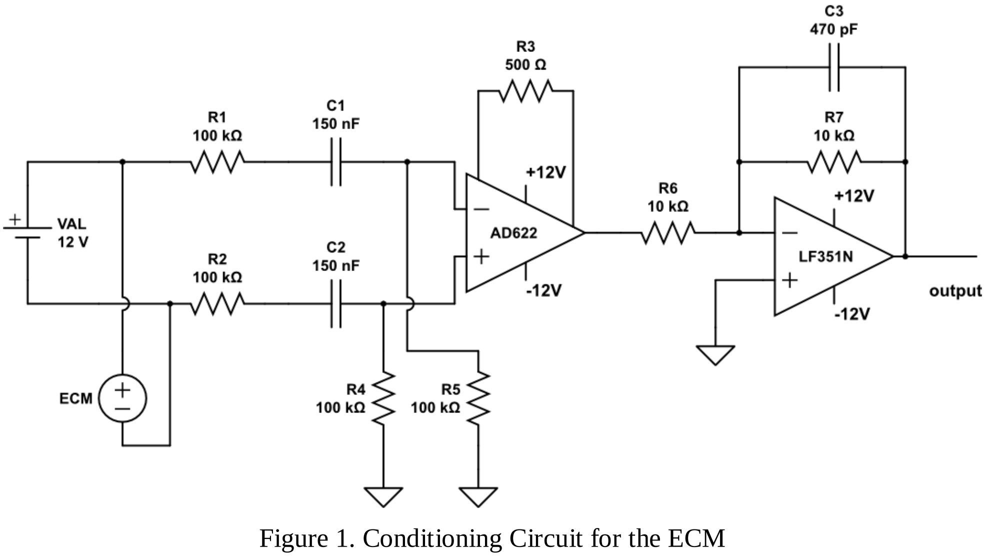

Figure 1. Conditioning Circuit for the ECM

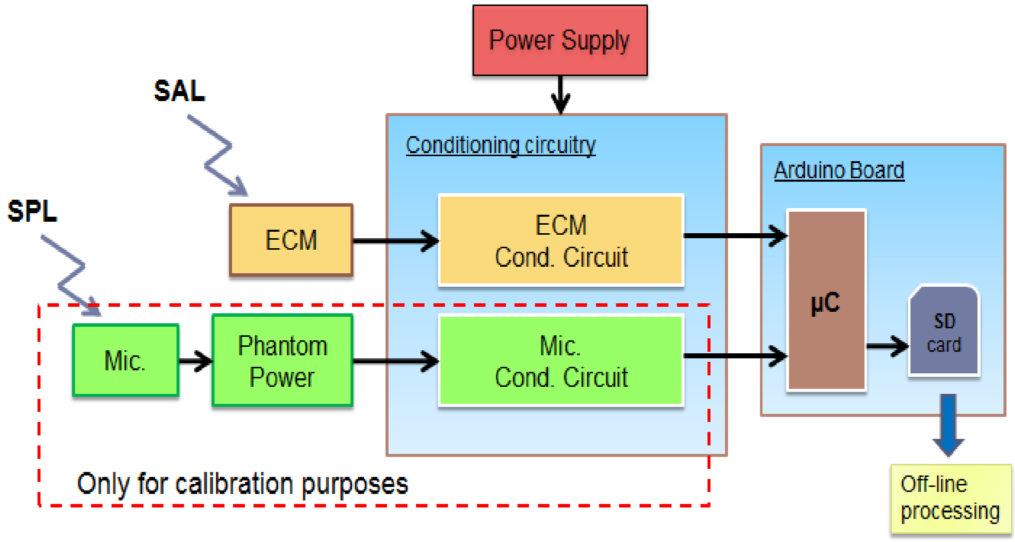

Figure 2. Schematic for Calibration Circuit

2.3 Sensors and Transducers

Microphone is an acoustic-to-electric transducer or sensor that converts sound into an electrical signal. Most microphones today use electromagnetic induction (dynamic microphone), capacitance change (condenser microphone), piezoelectric generation, or light modulation to produce an electrical voltage signal from mechanical vibration. Microphones are referred to by their transducer principle, such as condenser, dynamic, etc., and by their directional characteristics [11]. In this research, the electret microphone (ECM) is used. An electret microphone is a type of capacitor microphone that the externally applied charge under condenser microphones is replaced by a permanent charge in an electret material. An electret is a ferroelectric material that has been permanently electrically charged or polarized [11].

In this experiment, a commercial lightweight laryngophone ALBRECTH AE 38 S2 is used. The modification is done for the sake of experimental purposes, especially for the microphone connector and rubber neck. Laryngophone lets users communicate by exploiting the vocal vibration technique. The microphone is placed on the side of *jugular notch* to exploit vocal cord vibrations. In order to conduct the SPL calibration, there is another microphone used as reference. A condenser microphone Behringer ECM8000 is used for this purpose. Another possibility device to sense the vocal vibration is accelerometers. One of the main reasons the electret microphone is investigated as an alternative to the accelerometer is its expected larger bandwidth [12].

2.4 Power Supply Management

The power source for all components is designed to be supplied by a single battery. There are two specifications of the supplied voltage, ±12 and 5 Volts. The value of ±12 Volts is intended powering the parts of the acquisition board (ECM phantom power and Op-Amp ICs), whereas the 5 Volts is used to supply the Arduino board. To achieve these purposes, a single battery is used with nominal voltage of 15 Volts. The battery is chosen by a consideration that the power source has to be higher than the consuming device and it is purchasable from the supplier. Apart from the voltage specification, the available current provided by the power source has to be taking into account. Arduino board and conditioning circuit consume a certain number of current. In this research, the rechargeable Li-ion MGL2809 battery pack from Enix Energies is selected. The battery has the nominal capacity of 2.2 Ah, 15 V output voltage and the maximum charge current of 3 A.Two DC/DC converters (±15/±12V and ±15/±5V) are applied to convert the voltage down to ±12 and ±5 Volts. The MEA1D1512DC and the MER1S1505SC from Murata are selected for 15/12 Volts and 15/5 Volts DC/DC converter, respectively.

2.5 Software Algorithm for Normal and Calibration Modes

The device is organized to work in both **normal** and **calibration** modes. Calibration mode is used only at the first time to calibrate the device for different users. Once the calibration procedure is completed, the normal mode is used henceforth. Normal mode means that the device samples the input signals and stores them in a device memory (SD card) by occupying only one analog input of the microcontroller. In this research, there are two assigned buffers with array size of 400 each. These buffers are used to support the **dual-buffer technique** which is applied in storage management system. In calibration modes, the device will occupy two analog inputs of the. A calibration process requires sampling two devices simultaneously. One device is a standard device and another device is a non-calibrated device.

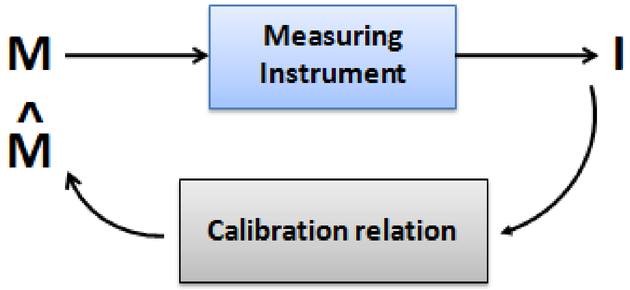

Figure 3. Simple Model of Measuring Instrument

A measuring device has to be calibrated against reference standards in order to provide traceable measurements. Figure 3 shows a simple model of measuring instrument with single input and single output. Input M is unknown measurand whereas output I is instrument indication. A condition for the employment of an instrument is therefore the knowledge of the output/input relation, the so called **calibration relation**, i.e. a relationship that allows an estimation $\widehat{M}$ of the measurand to be obtained from the indication I [13].

The results stored in SD card by the microcontroller will be analyzed. The analyzing process begins by finding the value of root mean square (rms) from the obtained signal. The rms is the square root of the mean of the squares of the data. Mathematically, the rms value can be formulated as

$$v_{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} {v_i}^2} \tag{2.4}$$

where v_i the voltage value for the *i*-th frame and N is the number of frames. The frame length is set to 30ms.

2.6 Data Storing Management

The data obtained from the ADC will be stored in a temporary buffer. Once the buffer is full, the data will be written to the SD card. However, the buffer used in this system cannot be a single buffer. It is because the ADC collects the data continuously and runs in *free-running mode*, as shown in Figure 4.

A free-running mode means that the ADC runs continuously without stopping. Once a single conversion has been finished, the next conversion will be immediately started. It this case, to maintain the correct data, it is not possible to directly write the data to the SD card when the buffer is full. A set of data will lose because writing process needs a certain time. Therefore, the *dual-buffer technique* is applied.The dual-buffer technique means that there are two buffers involved in this system as illustrated in Figure 4. In the first step, the ADC is managed to load buffer 1. There is no writing process to SD card during this period. Once the buffer 1 is full, the ADC is switched to load the buffer 2. During loading the buffer 1, the writing process to SD card is started. Generally, when the ADC loads one buffer, there will be a writing process for another buffer. The important requirement needed for the dual-buffer method is the timing constraint. The time to load the buffer must be less than the time to write another buffer to SD card [14].

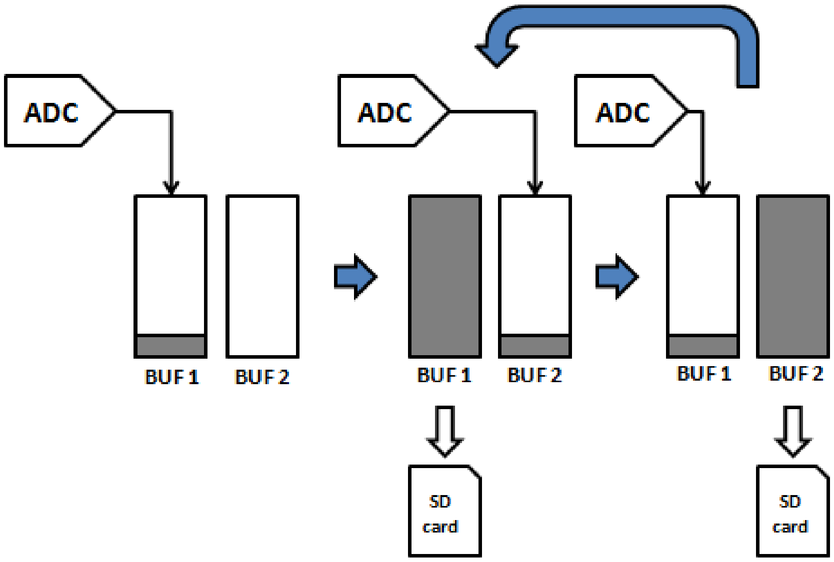

Figure 4. Illustration of the Dual-technique Process

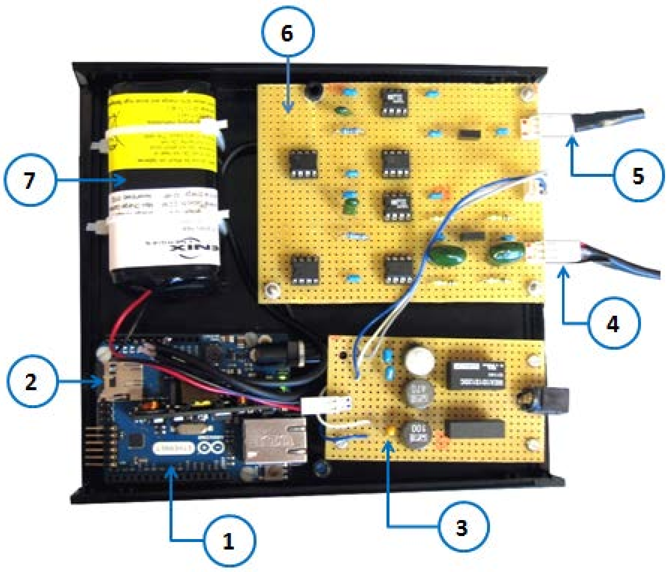

Figure 5. Final hardware consists of : 1) Microcontroller board, 2) SD card, 3) Power board, 4) ECM connection, 5) Mic connection, 6) Signal conditioning board and 7) Battery [11]

3. Results and Analysis

3.1 The Final Product

The final product of the research is a portable device for vocal signal monitoring. As shown in Figure 5, the hardware consists of microcontroller board, power supply board, signal conditioning board, and battery. All boards are arranged and fixed together such that fit in a plastic case.

3.2 Calibration and Verification Processes

There are four steps to conduct the calibration and verification, as follow [11]:

1. Microphone calibration against pressure calibrator
2. ECM calibration against microphone
3. ECM verification against microphone
4. Normal use of the portable vocal logger

3.2.1Microphone Calibration Against Pressure Calibrator

The rms values for with and without signal conditions is shown in Figure 6. From these values, the average value for no-signal condition is 12.1 mV and the condition when the signal exists is 576.2 mV. The SPL can be calculated as follow:

$$SPL = 20 \cdot \log_{10}\left(\frac{P}{P_{ref}}\right).$$

The value of the SPL is known from the calibrator specification (94.1 dB), and the value of P_{ref} is 2.10^{-5} Pa. Therefore, the value of P is 1.014 Pa.

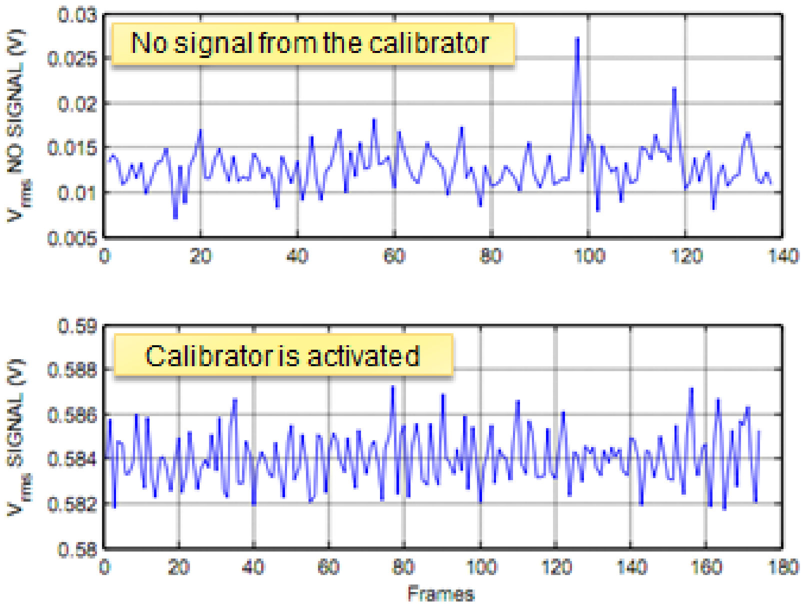

Figure 6. The rms values of the microphone

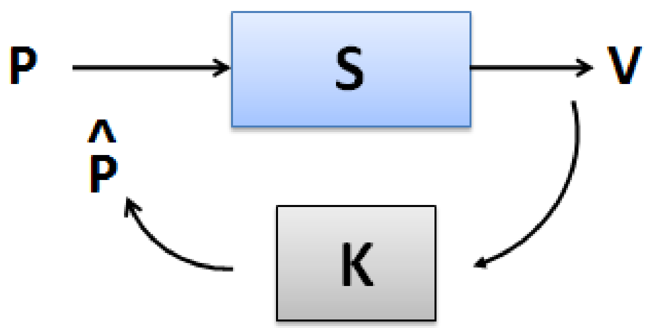

Figure 7 Model for SPL Calibration

The constants S and K are introduce in Figure 7 to simplify Figure 3. The measuring instrument S constant can be obtained by

$$S = \frac{V_{rms}}{P}$$

and the result is 0.57 V/Pa. Therefore, the calibration constant K can be calculated (K = 1/S) and the result is 1.76 Pa/V (0.12% standard deviation).

3.2.2 ECM Calibration Against Microphone

Once obtaining the calibration constant, the next step can be done. In this section, however, the signal source is produced from the talker instead of the sound level calibrator. It is important to note that due to the different physical and biological characteristics of the vocal generator and the vocal track, different talker will produce a unique behavior in terms of fundamental frequency, rms value and other voice parameters. Both ECM and microphone will be used and microcontroller is set to acquire data in dual-channel mode. The talker is assigned to wear the ECM and matched it on the jugular notch, as shown in Figure 8.

.

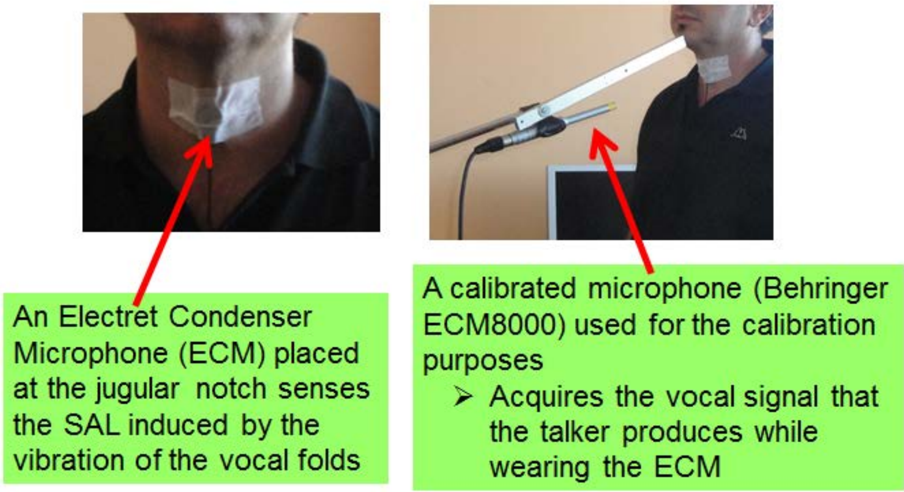

Figure 8. ECM and Microphone Calibration Process

During this step, the talker produces the vowel /a/ at different intensity while the device is logging the data at its two inputs (ECM and Microphone). Figure 9 shows the results. The upper part of the figure shows the rms value produced by the ECM whereas the lower part by the reference microphone. Zero rms means unvoiced condition. In general, the microphone produces the higher rms value than the ECM.

The average value of each peak is calculated, resulting only a single value of rms voltage. It is shown in Figure 10. By applying the linear fitting for $V_{rms,ECM}$, the relation between P and $V_{rms,ECM}$ is obtained as P = 1.735 x $V_{rms,ECM}$ - 0.15.

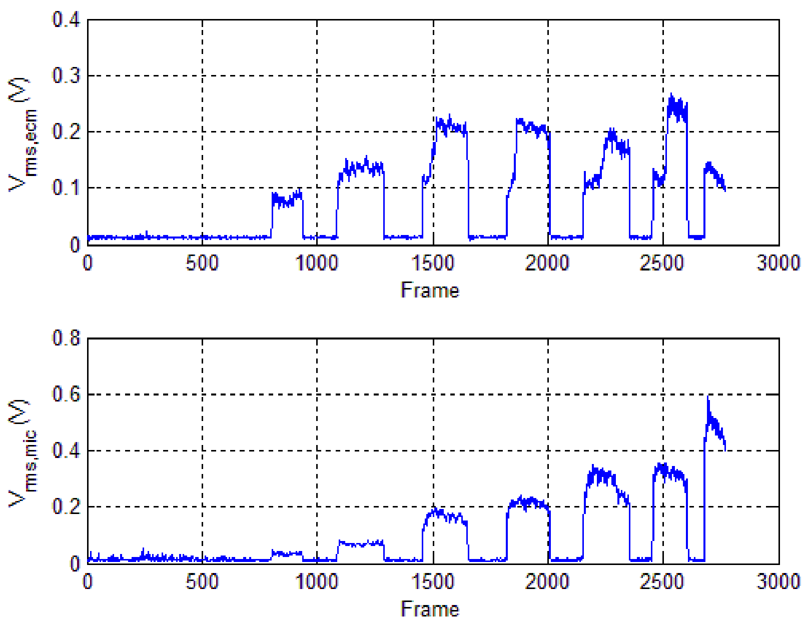

Figure 9 The rms values of the microphone and ECM

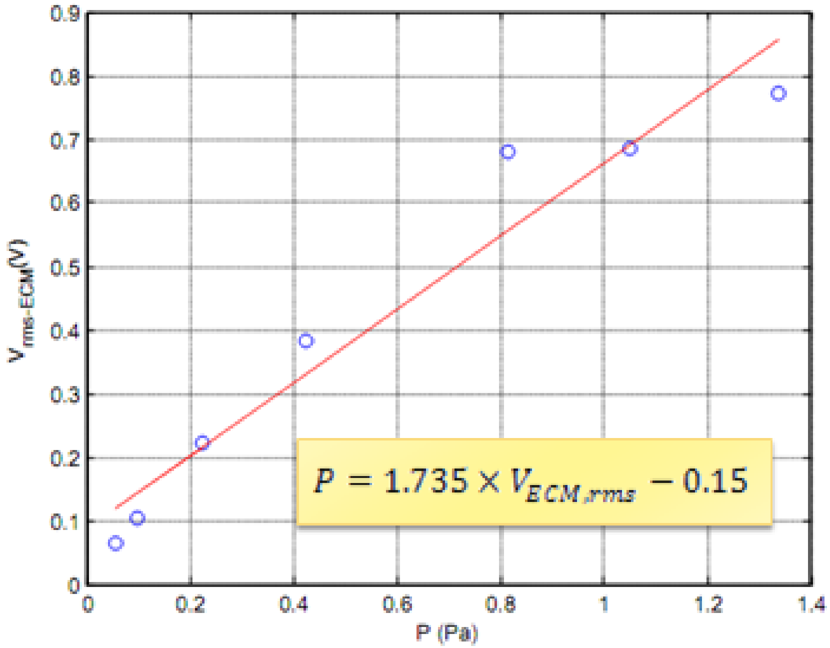

Figure 10. Relation between Pressure of the Microphone and ECM Voltage.

3.2.3ECM Verification Against Microphone

At this step, the verification between ECM and Microphone is done. This step is based on the relation function of the ECM. The verification is done by reading a passage for about 80 seconds and comparing SPL_{mic} and SPL_{ecm}. The result obtained from microphone and ECM is shown in Figure11a . Average error is obtained about -1.8 dB with standard deviation is about 4.6 dB. The error figure is shown in Figure 11b.

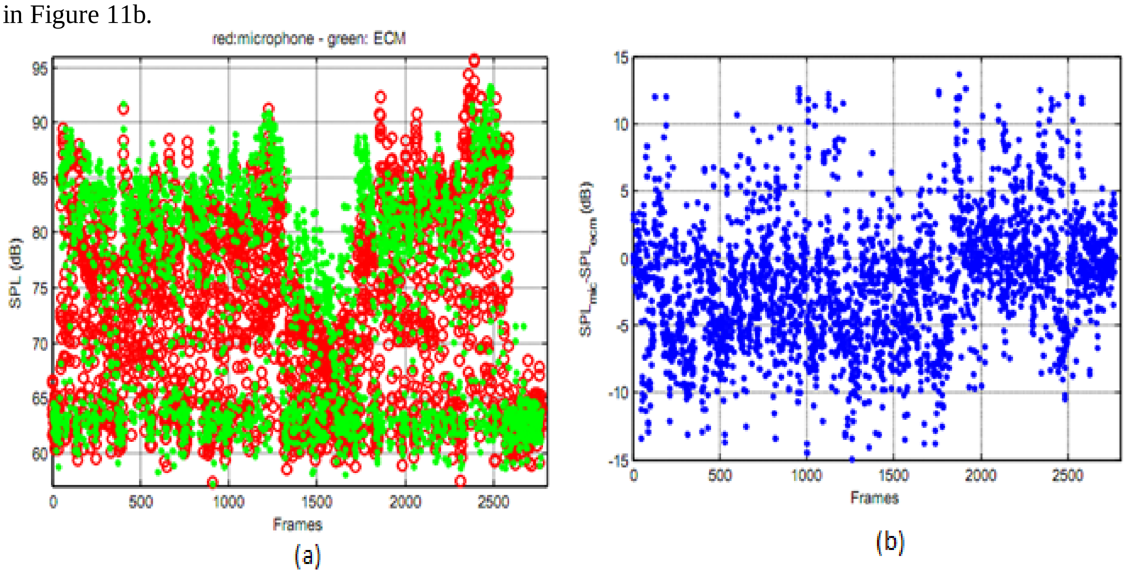

Figure 11.(a) Verification Result of SPL ECM against SPL Microphone, (b) SPL ECM Error

3.2.3 Normal Use as a Portable Device

The last step is done as a real application.The device is tested and the ECM is worn by the male talker. The talker does their usual activities about 45 minutes. The occurance of SPLs in dB and bar chart of Fundamental Frequencies (F0) and SPL is shown in Figure 11 (a) and (b), respectively.

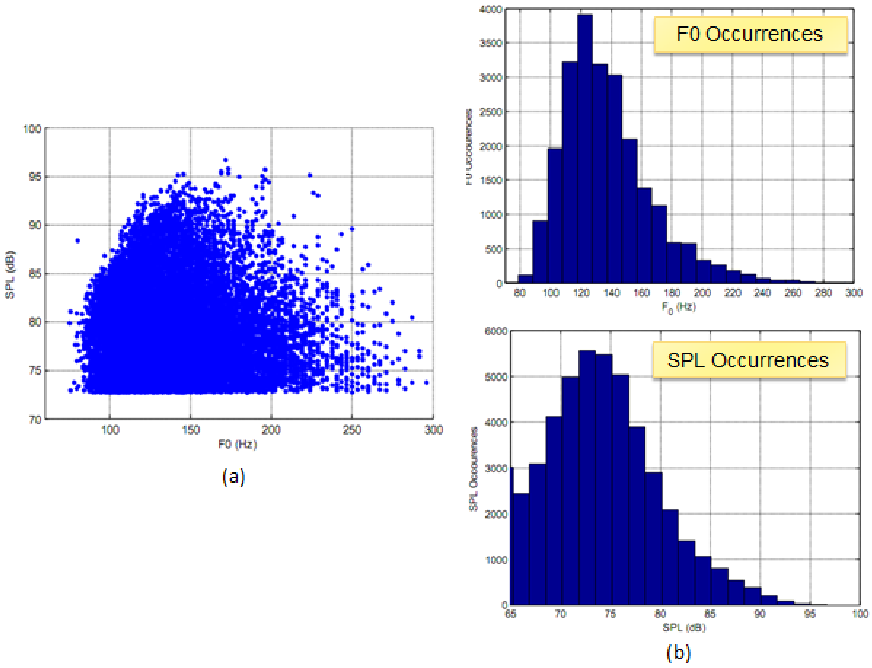

Figure 12.(a) SPL Obtained as Normal use and (b) Bar chart of F0 and SPL

4. Conclusion

As a preliminary research, a portable microcontroller-based device for vocal signal monitoring has been developed. The calibration and verification of the device have been done. Error average of the ECM is -1.8 dB (std. dev. 4.6 dB) compared to the reference microphone. A frequency characterization of the ECM will be useful to correct its effect and then improve the device performance. For the future work, the possibility to improve the functionality, not only as a logger but also as a processor of the vocal signals. Different type of microcontrollers can be a solution. Futhermore, the user-friendly aspects can be considered for the future work. Collaboration works between engineers and physicians are advisable for a correct use of the estimated parameters.

References

[1] Camilleri, Norma.*Voice Disorder.* Association of Speech and Language Pathologists.2012

[2] Lohscheller, J. *Towards Evidence Based Diagnosis of Voice Disorders Using Phonovibrograms*, 2nd International Symposium on Applied Sciences in Biomedical and Communication ISABEL 2009. ISBN: 978-1-4244-4641-4

[3] Anonym. *Dance Voice Disorder Brochure*.The Voice Center Greater Baltimore Medical Center

[4] Asgari M. and Shafran I.*Predicting Severity of Parkinson's Disease from Speech, 32nd* . Annual International Conference of the IEEE EMBS pp 5201. 2010

[5] Anonym.*Voice Disorder*. American Speech-Language-Hearing Association (ASHA). 2012

[6] Anonym. *Analog to Digital Conversion, Data Acquisition PDF Handbooks of Measurement Computing™*.Measurement Computing corp. 2013

[7] Anonym.*AVR120: Characterization and Calibration of the ADC on an AVR*.Atmel. 2013

[8] McRobert M.*Beginning Arduino*, United States of America. Apress Publication. 2010

[9] Arduino official website: http://www.arduino.cc/. Arduino corp. 2013

[10] Anonym. *REF03 Datasheet*. Analog Device. 2013

[11] Kusuma W.*Microcontroller Based System for Vocal Signal Monitoring*. Master Thesis. Turin: Politecnico di Torino; 2012

[12] A. Carullo et.al. *A Portable Analyzer for Vocal Signal Monitoring*.IEEE Instrumentation and Measurement Technology Conference (I2MTC). 2012.

[13] Carullo A.*Calibration and Traceability*, Electronic Engineering Course material, Politecnico di Torino 2010.

[14] Kusuma W.Studi Penerapan Mikrokontroler sebagai Data Logger Sinyal Suara. Proceeding of Applied Science for Technology Innovation(Astechnova). Yogyakarta. 2012