The George Washington University
The School of Engineering and Applied Science
Department of Electrical and Computer Engineering

Final Product Review

Functioning Model of the Larynx & Voice Analysis System

Onyekachi "Kachi" Odoemene
ECE 158 - 10
April 09 2009
Professor T.J. Manuccia
Professor A.B. Hancock, PhD

Abstract

The goal of the proposed project is to design and build a functioning biomechanical model of the larynx. This model will be used to simulate the natural vocalization process and characterize common vocal disorders such as vocal nodules, cysts, polyps and Reinke's edema. The model will include vocal fold structures that vibrate and produce sound in response to airflow. Having a functioning model of the larynx will help in understanding the mechanism of voice production and also the different vocal disorders. It can also be used as a teaching and/or voice training tool for students, and patients diagnosed with certain vocal disorders.

Accompanying the model is a voice analysis system that will be used to analyze voice signals to aid professionals in determining whether or not a voice is normal or disordered.

Table of Contents

List of Figures

List of Tables

List of Equations

Introduction

Background Information

The larynx, commonly known as the voice box, is responsible for protecting the airway from foreign objects and the production of human sound. The larynx houses the vocal folds which are responsible for production of human sound. Human voice is produced by vibration of the vocal folds when air from the lungs blows on them, this process is known as phonation. Vocal folds are a pair of muscle surrounded by pliable mucosal membranes that are attached anteriorly to the thyroid cartilage and posteriorly to the two arytenoid cartilages. The arytenoid cartilages adduct and abduct to move the vocal folds in and out of the airway.

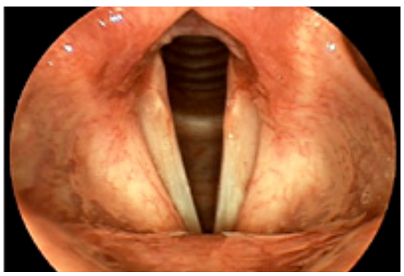


Figure 1: Adducted Vocal Folds

Figure 2: Abducted Vocal Fold

To begin voicing, the vocal folds approximate (or come close together), causing a constriction in the airway. The area between the two folds is the glottis. Subglottal pressure, pressure from the lungs, opens the approximated vocal folds from the bottom up. Because of the elastic nature of the vocal folds, they will want to return back to the approximated state. In returning, the vocal folds begin to close from the bottom, causing a decrease in pressure perpendicular to the airflow which in turn will suck the vocal folds back together (Bernoulli Effect). This cycle is repeated several times during the voicing and it is known as the mucosal wave. Below is a figure that illustrates the vibration of the vocal folds.

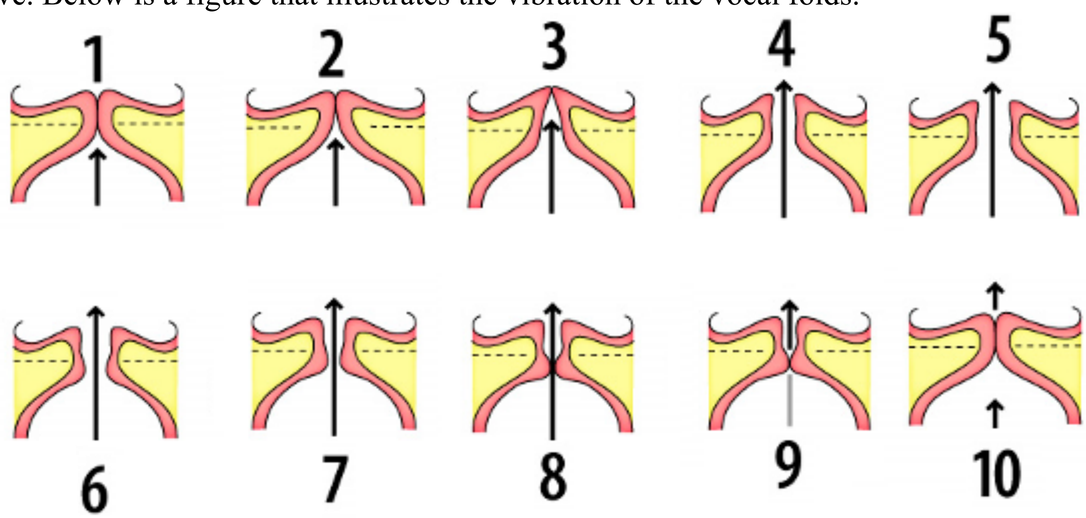


Figure 3: One Cycle of Vocal Fold Vibration

The vocal folds vibrate between 80-100Hz in males, 180-220 Hz in females and about 300 Hz in children. The pitch of the voice can be adjusted by increasing the tension on the vocal folds during vibration. The action of the thyroid cartilage moving anteriorly stretches the vocals to produce desired tension to adjust pitch. The volume of the sound produced is manipulated by increasing the airflow from the lungs which causes the vocal folds to vibrate at higher amplitudes.

The voice produced by the vibration of the vocal folds is often described as a "buzzy" sound which differs significantly from the voice one hears when they speak. This is because this buzzy sound is amplified by resonators in the vocal tract and modified by articulators which include the tongue, soft palate and lips[a] and this is then sound we hear when we speak. This model will not attempt to reproduce the phenomena that occur beyond production of sound by vocal fold vibration.

Any interruption in the normal pattern of the mucosal wave results in a vocal disorder. For the purpose of this project, the vocal disorders of interest are vocal nodules, polyps, cysts and Reinke's edema. These disorders are typically a result of vocal misuse or abuse which includes yelling, shouting, chronic throat clearing and smoking. Each of these disorders is a benign vocal lesion.

Vocal nodules are small collection of cells that develop on both sides of the folds. The vocal characteristics of a patient with vocal nodules are low pitch, hoarseness and breathiness. The nodules do not allow the vocal folds to close properly and the additional mass causes irregularities in vibration. Treatment for nodules includes voice therapy and surgical removal of nodules. Vocal fold polyps are fluid filled lesions that usually occur on one of the folds usually caused by extreme vocal abuse. Like its counterparts, nodules and cysts, they also affect the vibration of the folds and also cause breathiness.

Reinke's edema is a disorder characterized by the swelling of the Reinke's layer, one of the 4 layers of fiber covering the vocal fold muscles. This disorder is caused by smoking. Its primary vocal characteristic is hoarseness of the voice.

There are many more disorders not mentioned in this introduction but these are the disorders of interest for this project because the overall design is limited. .

Motivation and Applications

People diagnosed with mentioned voice disorders have difficulty understanding the mechanism of voice and voice disorders. This model will be used to provide a visual aid of how human vocalization occurs. It will also be used to simulate common voice disorders (specifically those mentioned in the previous section). Also there are hardly any functioning models of the larynx that accurately depict the vocalization process. Accompanying the model will be a system for recording and analyzing the acoustic signal produced by vibration of vocal folds. This will assist students in learning to classify different speech pathology.

[a] http://www.voiceproblem.org/anatomy/understanding.asp

Market Analysis

Although this model is not entirely novel, such biomechanical models have been built before for educational and experimental purposes but not for commercial purposes. Thus it's safe to assume that there is no competing model similar to the proposed design.

However, there is a lot of ongoing research in developing computational models of the vocal fold vibration mechanism and the disorders associated with them. Another competing model is stationary models of the larynx used in clinic settings.

A recent competitor has emerged since the PDR. It is the WT-7 (Waseda Talker No.7). It is an anthropomorphic robot built by engineers in Takinishi Laboratory in Waseda University in Tokyo, Japan. This is a highly sophisticated robot that has modeled all organs used in speech production including mechanical lungs, vocal cords based on human biomechanical structure, articulators, tongue, lips, teeth, soft palate among others. To the best of our knowledge, this device is not yet commercially available. The market value of this product is discussed in the Economic Analysis section of the report.

Many types of commercial voice analysis software exist. The leading voice analysis program is the Multi-Dimensional Voice Program produced by KayPentax. It calculates more than 22 parameters on a single vocalization. The software excels in extraction of parameters from disordered voices. This software is much more sophisticated than the software discussed in the project. MDVP will serve as the standard model for the software used in this project.

Overall Product Requirements

To reiterate, the main goal of this project is to simulate the human vocalization process through a device that will produce sound through vibration of its components. The following are the requirements for the mechanical model larynx.

1. The sound produced by larynx model must be of audible frequency range.

2. The vocal fold model will have similar elastic properties as the human vocal folds to allow it vibrate and produce sound.

3. The vocal folds must easily move in out the airway.
    a. The vocal folds will be stretched or relaxed to change the tension which will increase or lower the pitch.
    b. The volume of the sound produced must be regulated using air supply.

4. The overall system must have an input for an external microphone for capturing the audio signals for analysis.

5. The voice signal for analysis must be filtered and amplified

    a. Low pass filter to remove high frequency noise and high formant frequencies
    b. High pass filter to remove low frequency noise due to motion artifacts, wind pop noise, and other noise will distort signal being recorded.

6. User will be allowed to upload prerecorded signals for analysis

7. User will be allowed to select section of voice signal to analyze

8. The Analysis system give generate three waveforms: the sound produced by the model, the spectrogram and the frequency spectrum (FFT).

9. The model will simulate different vocal disorders by modification of the vocal fold structure.

10. The software will show graphic display of voice signal to allow user compare results of analysis to norms

11. Software for analysis of signals produced from the model must be easy to use for person knowledgeable of vocal anatomy and voice production.

12. The model must be portable including airflow system.

Overall Product Specifications

The following are specifications that support the project requirements.

1. Vocal fold vibration frequency range 80-120Hz.

2. Model vocal folds should have shore hardness less than 40 Shore A.

3. User will be allowed to control gain of amplifier.
   a. Varying it from 50V/V to 200 V/V to increase voice signal strength for analysis

4. Vocal folds will be manipulated manually to change tension and move laterally.

5. Analog low pass filter will cut off frequencies above 5+/- 1 kHz

6. Analog high pass filter will cut off frequencies below 30 Hz +/-5hz

7. Analysis software will produce provide report including parameters such as: Jitter%, Shimmer%, NHR and Average fundamental frequency.

8. Analysis system will analyze only 1 second sample which is determined by the user.

   a. User can select which one second segment to analyze

9. Analysis system will include user friendly Graphic User Interface.

10. Software for analysis of voice signals must be compatible with Windows PC.

11. Model vocal folds will take air as input. Airflow must be regulated. Air pressure must be variable with the range of 0-2kPa.

Design Approach

Current Design

The end goal of the proposed product is to produce a functioning model of the larynx. A functioning model implies that the model will contain an active component, in this case the vocal folds. The vocal folds will vibrate in response to airflow and as a result of vibration produce a sound. This is much similar to the function of the vocal folds found in nature. This model will be used to simulate natural vocalization and common voice disorders that occur in many people. To simulate natural voicing and common voice disorders, the proposed larynx model will have vocal folds that can be manipulated by moving them in and out of the airway, adjusting the tension to change pitch and respond to changes in mass.

The basic concept behind the design is illustrated in the figure below. Airflow causes a pair of pliable membranes in the larynx to vibrate and produce sound. The sound is captured and further analyzed.

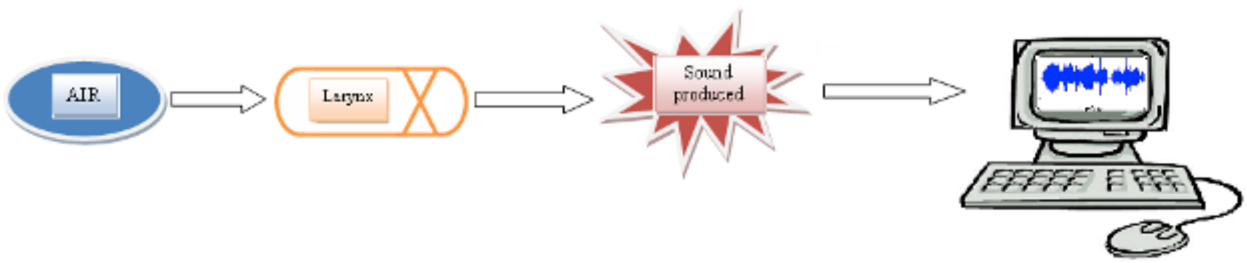


Figure 4: Context Level Diagram

This project can be easily broken into two major modules: the hardware module and the software module. The hardware module consists of both a mechanical and electronic component. The mechanical component is the functioning model of the larynx with active vocal folds, while the electrical component is a circuit that filters and amplifies vocal signals for analysis by the software module.

The functioning model of the larynx will function on the theory described in the introduction section. The electrical circuit will be used to process the voice signal before analysis by first filtering out high frequency noise in the signal and also low frequency noise due to motion artifacts such as tapping or dropping of the microphone. Because the acoustic signal of the human voice is a relatively weak electrical signal and the microphone used may vary, a variable gain amplifier is needed to amplify the signal to increase the voice signal's strength.

The active vocal fold model will be made using thermoplastic rubber molded in the shape of the human vocal folds. The rubber material shares similar elastic properties as that of the human vocal folds, thus it will vibrate in response to airflow. Using a mechanical apparatus the vocal folds will be manipulated to stretch to increase tension and ultimately change the pitch of the sound produced and moved laterally and medially from the where each fold meet. In the human larynx, the thyroid cartilage moves anteriorly and the arytenoid cartilages move posteriorly to stretch the vocal folds and change pitch. The model vocal folds are also designed to move medially and laterally to simulate voiced and voiceless sound. The designed mechanical apparatus for vocal fold manipulation attempts to achieve the same physiological actions performed on the vocal folds.

The software will acquire the processed voice signal and perform several functions to extract information necessary for computation of vocal parameters. The calculated vocal parameters will be used by the clinician user to determine if the voice is normal or disordered. The vocal parameters extracted are: average fundamental frequency of voice signal; jitter % which is a measure of cycle to cycle pitch variability; shimmer percent, cycle to cycle peak to peak amplitude variability; and noise to harmonic ratio (NHR), a general evaluation of noise present in the analyzed signal is defined as the average ratio of noise spectral energy to the harmonic spectral energy in the frequency range of 70-4200 Hz.

Each module and its respective sub-modules will be discussed in great detail in the following section.

Evolution to Current Design

The vocal fold model was the most challenging aspect of design, as it being a very complex system. The first prototype is below Figure 5.

Figure 5: First Larynx Model Prototype

The prototype was built last spring. Although a very crude model it was able to introduce the concept of vocal vibration and provide more insight into the mechanism. The model vocal folds used was a thin rubber latex strip that was sliced in the middle. The figure to the right demonstrates how the tension of the latex strips was adjusted. The bottom disk is fixed onto the tube and has a spring above it; the middle disk is free to move. The user then presses down on the middle disk which lowers the top disk causing the latex strip to stretch, thus changing the tension and ultimately the pitch of the sound.

The latex strip vibrates and produces sound when air is released into the model from the nozzle below. However, it required a great deal of pressure get the strips to vibrate and produce sound because there was plenty of leakage at the sides of the latex strip. The vibration of the latex strips produced a sound which was audibly similar to a buzzy sound, slightly comparable to that produced from a human vocal fold. The sound was masked by the high pressure source which added high frequency noise to the recorded signal. A slight increase in pitch was heard when the strip was stretched, and when a tube was placed above the vibrating strip the sound is amplified (resonated). This gave the idea to incorporate a resonator module in the new model larynx.

The prototype model demonstrated that the vocal folds needed to be made of thicker material to simulate the mucosal wave. It was also pointed out that for vibration of vocal folds low pressure is required; this would eliminate the high frequency noise that interferes with the sound from the vibration of the vocal folds. In the new model, we attempt to seal all possible sources for leakage. The prototype was also tested using the MDVP software system, and its measured parameters were well outside the range of normal vocal parameters. The fundamental frequency was about 400Hz. This was probably a result of the high pressure source and thickness of the vocal fold material used. The new model will have vocal folds that are thick and based on the human biological structure which will allow it to perform a mucosal wave like motion upon airflow. Both models manipulate the tension manually; however, the new model will have a more controlled way of adjusting the tension. The new model will also move the vocal folds in lateral position.

To obtain an accurate geometry of the human vocal folds, it was originally intended that the vocal fold shape be designed with the help of a digitized image of the larynx. The digitized image would be generated with the help of a laser scanning arm and an anatomy model of the larynx, like the one in the figure below.

Figure 6: Anatomy Larynx Model[b,c]

The larynx model can be opened from the middle and the different parts can be scanned. The model is fairly detailed and highlights the major component needed, especially the vocal folds. The anatomical model has the vocal folds in the open position. This will not be very helpful in mold making so once the anatomical model has been digitized into an image, the portion with vocal folds will need to be extracted to create a mold that will give us a similar geometry to the actual vocal folds. Below is a digitized image of the anatomical model.

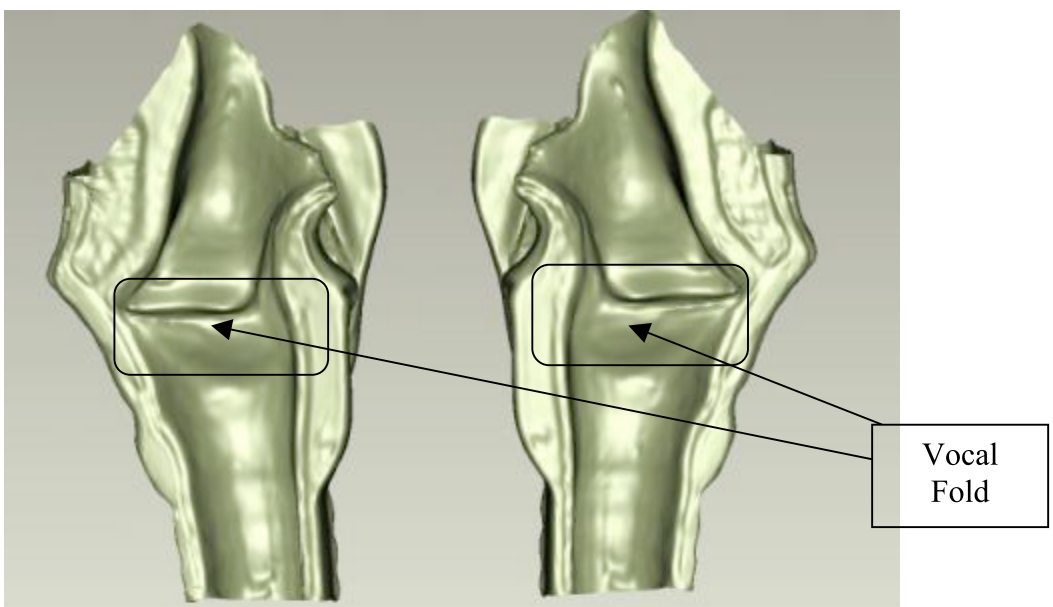


Figure 7: Digitized Image of Larynx: Vocal Fold Region

[b] http://www.montgomerycollege.edu/~wolexik/Thorax%20Model-Larynx-anterior-151.JPG
[c] http://www.montgomerycollege.edu/~wolexik/Thorax%20Model-Larynx-open-154.JPG

The regions that are boxed (vocal fold region) in the figure above will be extracted from the digitized image to generate a separate mold for the larynx. The software that will perform this extraction is the same one used to generate the scan data. The software is called Geomagic Studio. This software allows connection to a digital laser scanner, allowing the user to generate a 3D numerical model of the object which they are scanning, in this case the anatomical model of the larynx. Geomagic also allows the user to modify the scan data to generate a different model, which can then be used an input for a rapid prototyping machine. Using these features of Geomagic, the extracted region will be modified to make a mold for the model vocal fold. The actual mold will be made using the rapid prototyper. Once the mold is generated it will be used to make a negative mold out of porcelain clay, which would then be fired to dry. Thermoplastic rubber Septon® will be used to cast the clay mold to form a positive which would be the model vocal folds.

This design was abandoned because as one can notice it involved many complex steps, thus making the implementation of the final product very expensive and time intensive. It would have been difficult to shape the clay mold to resemble the vocal folds and one would also have to account for shrinkage of the clay when it is fired to dry.

An alternative to the vocal fold design was inspired by the Takanishi Laboratory in Japan. It is discussed further in the report.

Overall System Design

The project is centered on the sound generated by the functioning model of the larynx, which will be produced by the vibration of the model vocal folds in response to an external air supply. The sound produced will then be analyzed by the PC software.

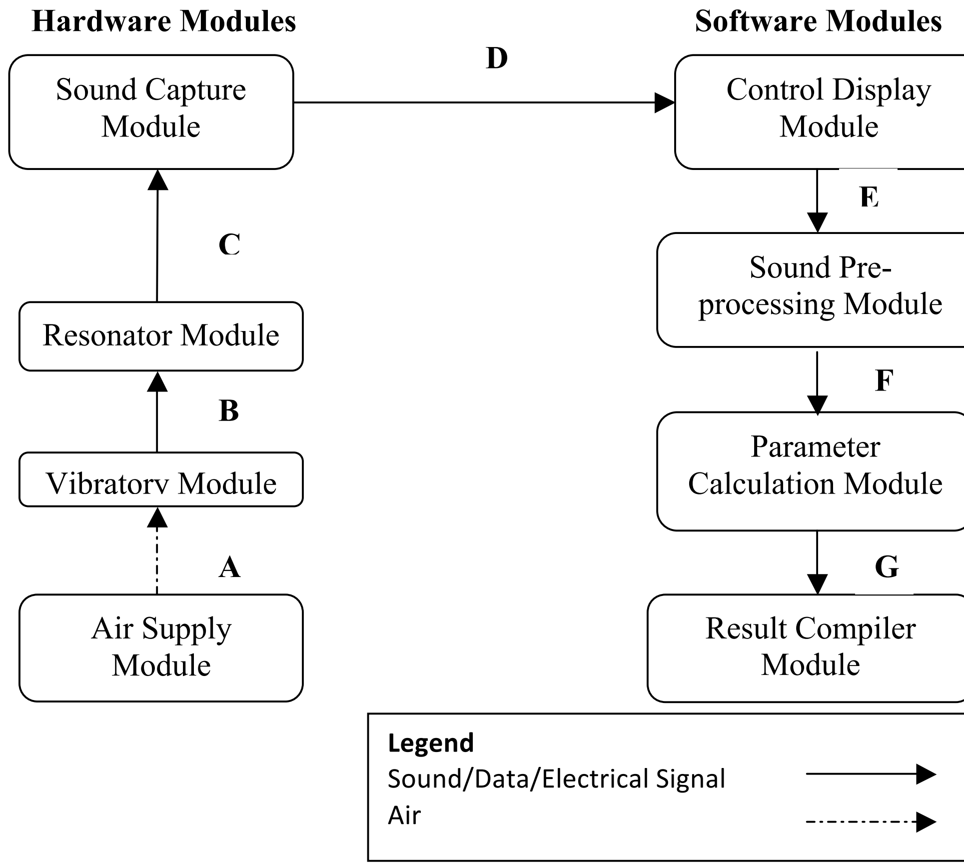


Figure 8: Logical Architecture Diagram

In this design, the hardware module is primarily responsible for generating and modifying the sound produced from the model larynx, whereas the software modules analyze the sound produced. In the figure above, D represents the interaction between these two major modules. The Sound Capture module, which consists of a microphone and an electronic filter and amplifier, will acquire the sound and send it for analysis by the software module.

Hardware Module

As a potential educational tool, it is important that the model closely mimic the physiology of human voice production. The mechanical module divides into the resonator module, vocal fold module and trachea modules. The resonator and airway module will essentially be separate tubes placed above and below the vocal fold module, respectively. The resonator module amplifies and slightly filters the sound produced from the vocal fold module through resonance. The vocal fold module receives its input, B, allowing it to generate sound through vibration of two pliable rubber structures.

The vocal fold module design is one of the most crucial aspects of the hardware module. It will be designed by creating a mold that gives an approximate shape and geometry of the vocal folds. The mold will be made using a CAD tool, which can be fed into a rapid prototyping machine. The rapid prototyper will then generate the 3-D shape of the mold and the mold can be casted with material that share similar elastic properties as the human vocal folds. The material of choice is thermoplastic rubber from Septon®.

An alternate approach to the larynx structure design would be to translating CT or MRI images of the larynx into a file format that can be fed into the rapid prototyper. This approach would be much more difficult to implement because the images would have to be converted into another file format. This will require a great deal of computational expertise to translate the file. Also, translation to another file format might lead to significant loss in information of the original larynx image data. The presented design is a much easier to implement and cuts out very many steps. However, one should anticipate many trial-and-errors to achieve the correct geometry.

The VF module also functions to manipulate the sounds produced from the vibration. Manipulations include tensing the VF structures to increase the pitch and movement of the vocal folds in and out the airway. This will be done with a mechanical device which is outlined in the detailed design further in the report.

The last major hardware module is the Sound Capture module. This module functions to record and amplify the sound produced from the larynx structure and inputs the sound to the software module. These five major modules encompass the hardware module of the design.

Software Module

The software module is used to analyze the sound produced from the hardware module, and compare the signal to known voice parameters. In implementing the software module, the software MATLAB will be used. A graphical user interface (GUI) will also be used to facilitate communication between user and analysis system. The software module is broken down into the 4 major submodules, as depicted in the figure 1.

The software module receives its signal D, from the Sound Capture module. The control display module begins the sound. It also will serve as the display module for all waveforms. The Sound Pre-processing module further filters the sound coming in via a digital low pass filter which will cut off at a frequency of 1.8 kHz. This is because higher frequencies are considered noise, since we expect a low frequency sound output from the model. In addition to effects of noise, the formant frequencies of the voice signal interfere with pitch detection and extraction. Output F is fed into the Parameter Calculation Module. The Calculation module computes the values of common vocal parameters: shimmer (%), jitter (%), noise to harmonic ratio (NHR), and the fundamental frequency of the sound signal. These parameters will be explained in further detail in the Parameter Calculation Module design. The next module, Result Complier, receives all the calculated parameters and provides a graphical display and text file containing the calculated parameters in relation to normative parameter values. From this, the user can determine if the voice signal is within normal limits. These are the four major modules that encompass the software module.

Modular Decomposition

The modular decomposition shows the hierarchical arrangement of all the modules involved in this project. In the next section, the design of the sub modules will be described in detail, including the specifications and requirements of each module.

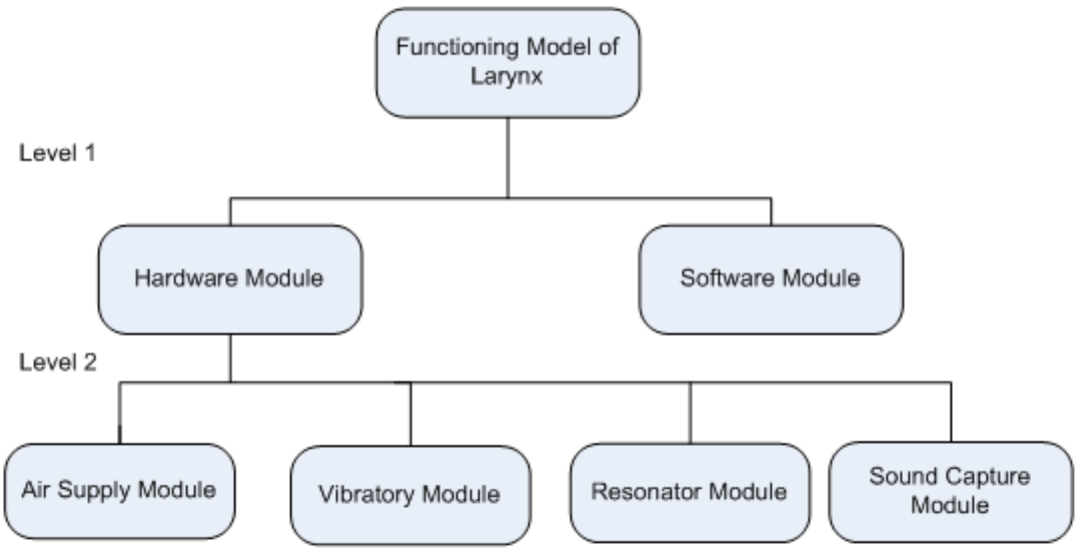


Figure 9: Hardware Modular Decomposition

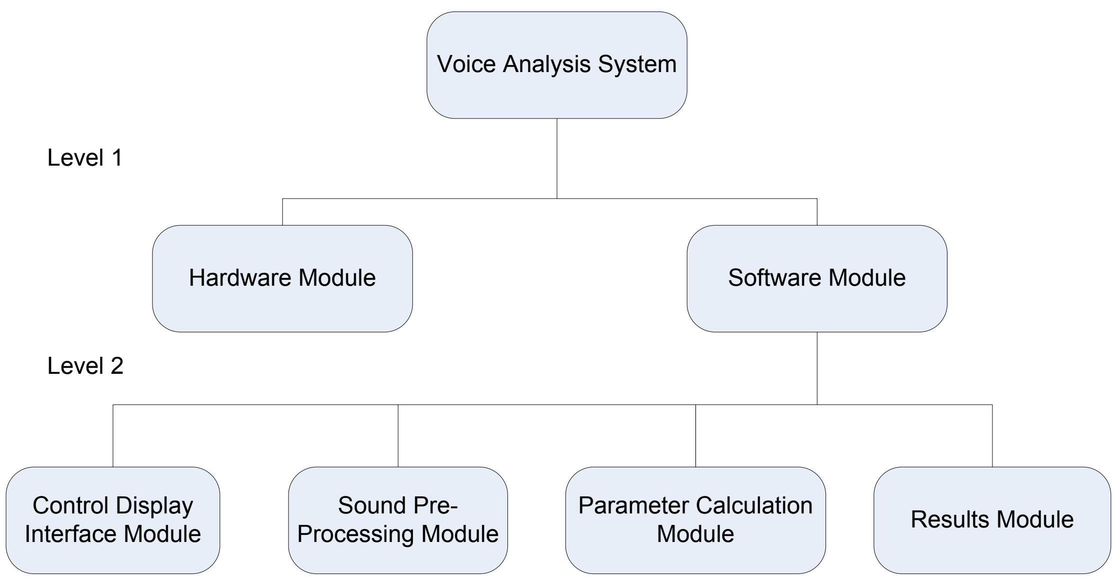


Figure 10: Software Module Decomposition

Module Level Design

Beginning with the Level 2 modules, this section will describe the design of the modules and their individual specifications and requirements. Also included in this section are submodules of respective level two modules.

Air Supply Module

Specifications

- Low pressure source with minimum pressure of 300 Pa
- Size
    - Outer diameter : 1.25 inches
    - Minimum length: 6 inches

Requirements:

- The air supply must be mobile
- User must be able to regulate the amount of pressure
- Will have connectors to attach to the model

The air supply module will function as the lungs of larynx model, providing regulated air to the vocal folds in order to vibrate them to produce sound. The human lung pressure required for phonation (production of sound via vibration of vocal folds) ranges from 0.3kPa-1.2kPa[1]; the lower limit being for very low conversations and the upper limit for very loud conversations. A variable air pressure supply will allow this model to change its volume from low to high and vice versa. To ensure air pressure specification and requirements, the vocal fold model must have no leaks and the connection between the vocal fold module and air supply module must be sealed.

For the potential application of this project, it would be essential for the air supply to be mobile. This module will be connected the vibratory module through a PVC tube with an inner diameter if 1.25".

Vibratory Module

The vibratory module will include the larynx structure and the model vocal folds. This module the center of the hardware design because it is responsible for production of sound.

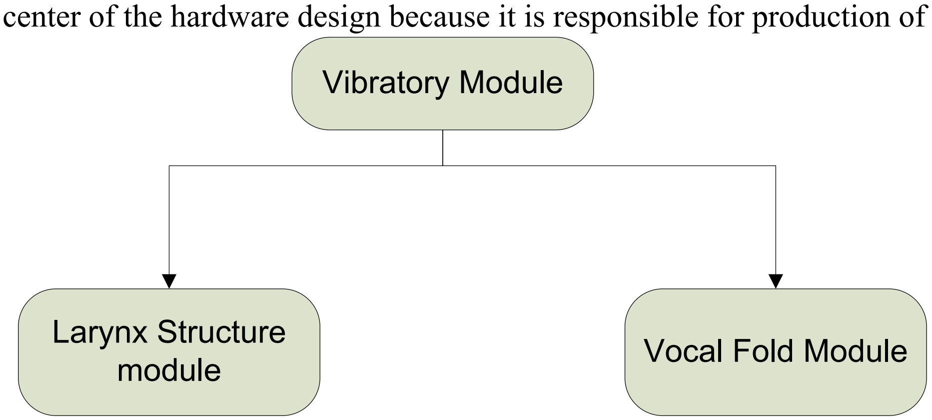


Figure 11: Vibratory Module

Requirements:

- Vocal folds will be active component, must respond to airflow from air supply system by vibrating.
- Vocal fold vibration will imitate similar vibration pattern of human vocal folds.
- Vibration of vocal folds will produce buzzy sound.
- Vocal folds will be manipulated to move laterally to simulate voiced and voice less sound.
- Vocal folds will be stretched to change the pitch of the sound.
- Vocal fold must be used to simulate different vocal disorders.
- Vocal folds must be incorporated into larynx structure.
- Larynx structure will include structures that manipulate vocal folds

Specifications:

- Vocal fold will vibrate at frequency above 60 Hz
    - Closely model adult male voice
- Vocal fold must be casted with thermoplastic rubber with shore A hardness range of 25 to 40.
- Different vocal fold molds should be generated to simulate different disorders.
- The larynx structure must include provision for insertion of model vocal folds.
- Vocal fold tension and opening and closing movement must be manipulated mechanically

This module is the most complex and will require a good deal of experimentation. As stated earlier, the state of the art of this larynx design used in the Waseda Talker no. 7 (WT-7). This is an anthropomorphic talking robot that is equipped with similar vocal mechanisms as human. It has vocal folds with human like biological structures, mechanical lungs that supply air, a resonator and articulators which modify the sound, including a tongue and a set of lips. Although their project surpasses the scope of this present project, the methods of developing the model vocals used in the WT-7 will be implored in this design. The model vocal folds were made from thermoplastic rubber Septon® by Kurary Co. Ltd. They created molds and used the thermoplastic rubber to cast the mold. This methodology of vocal fold formation will be adapted in this project. Below is an image of what the WT-7 vocal folds look like. The vocal folds used in this project will resemble a slight variation of WT-7 because of different molding pattern.

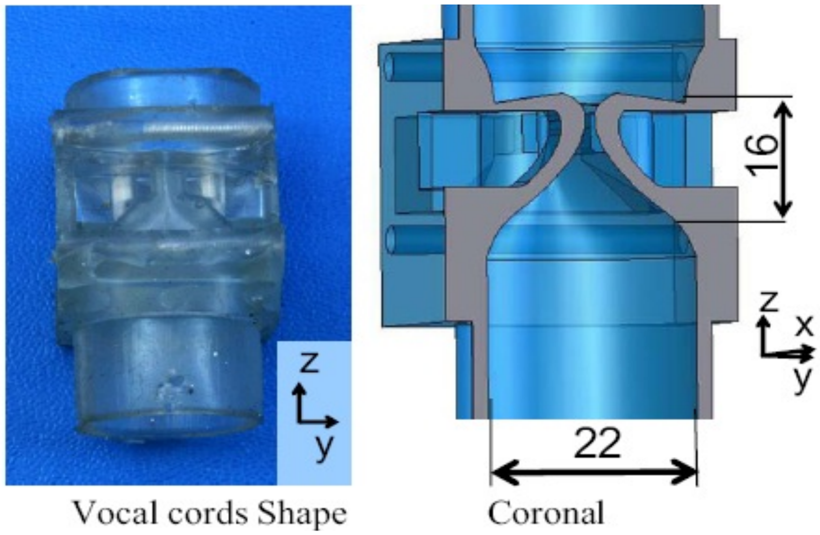


Figure 12: WT-7 Vocal Fold Shape[3]

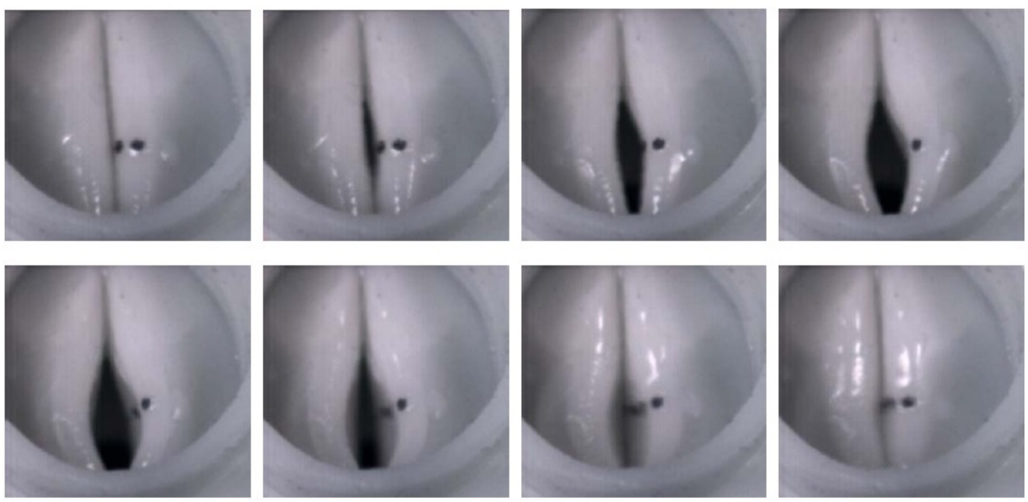


Figure 13: WT-7 Vocal Fold Vibration[3]

In Figure 12, you will notice that the vocal fold model is a hollow structure. This does not emulate the human focal folds because the human vocal folds are made of 5 layers that are generally lumped into 2 major layers. This is called the cover-body theory. The first major layer is made a very thin layer of epithelium tissue that encapsulates softer non-muscular tissue (fluid) like a balloon filled with water. The second major layer, known as the body, is made of collagen fibers and muscle[2].

The graduate student working on the WT-7 robot said that the vocal fold structure is hollow because of difficulty in modeling with cover-body theory approach. However, the vocal folds were able to vibrate almost like the human by having different phases in the upper and lower sections of the folds[3].

WT-7 vocal folds were manipulated with using a 4 degrees-of-freedom (DOF) linkage mechanism: 2-DOF to open and close the vocal folds and 2-DOF to control the pitch. They found that the fundamental frequency or pitch range varied with the mechanism with which the vocal folds were manipulated. The WT-5 model has 3 DOF (1-DOF to change tension and 2-DOF to open the glottis) and the pitch range was 100 to 110 Hz. The WT-6 model had a 5-DOF mechanism with a pitch range of close to 35 (117-152 Hz)[i]. In discussion with the graduate student working on these talking robots, the WT-7 robot was reported to have a pitch range of 120-200 Hz. The same vocal fold models were used in all three robots but the tension mechanism varied for each one. DC servo motors were used to implement the DOF linkage mechanisms. This will not be adopted in this project; model vocal folds presented in this report will be manipulated manually by the user.

The author of this report was able to replicate the design of the vocal folds used by the Waseda Talker-7 (Talking robot). Figure 14: Prototype Vocal Fold Mold DesignFigure 14 shows the CAD's made using PRO Engineering. These drawings fed as input to the rapid prototyping machine.

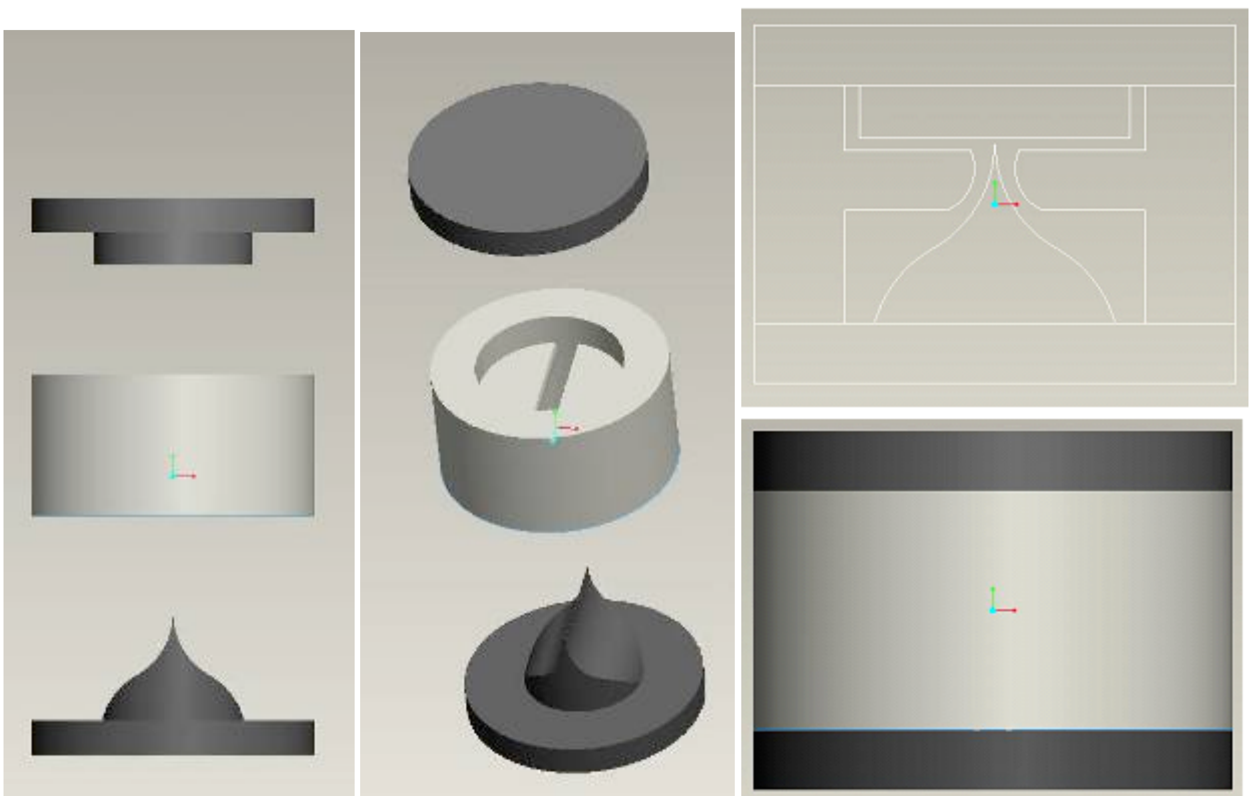


Figure 14: Prototype Vocal Fold Mold Design

The mold contains three major parts. The base of the mold gives the shape of the lower half of the vocal folds. It also separates the folds with a thin sheet at the apex. The middle piece gives the cylindrical shape to the vocal folds. Most importantly it provides the major shape of the vocal folds. As mentioned earlier the vocal folds are designed to be hollow. Once the mold has been generated via the rapid prototyper, it would be ideal to immediately cast the mold with thermoplastic rubber. However, the thermoplastic rubber must be heated and melted at 400°F in order for it to be useful as casting material and the material produced via rapid prototyping is plastic material called acrylonitrile butadiene styrene (or ABS). This will not work. Thus, the mold in Figure 14 will be machined using an NC milling machine out of Polytetrafluoroethylene (PTFE) commonly known as Teflon.

Vocal Fold Manipulation Design

Manipulating the model vocal folds will be done manually with the help of attachments onto the vocal folds. The human vocal folds are two bands joined together at one end, with the other end free to move laterally. From the top view it resembles a lambda Λ, with the point where the two lines meet being the anterior position of the vocal folds where they are fixed to the thyroid cartilage and the two separate ends being the posterior position of the vocal folds where they are attached to the arytenoids cartilage. The combination of the thyroid cartilage pivoting forward and the arytenoids rocking backwards stretches the vocal folds to increase tension, which causes a change in pitch. The arytenoids are sitting on a ring like cartilage called the cricoid where they glide from side to side to move the vocal folds laterally in the x-direction. This motion moves the vocal folds in and out the glottis to produce voiced and voiceless sound in response to airflow. These modes of movement will be adapted in manipulating the model vocal folds.

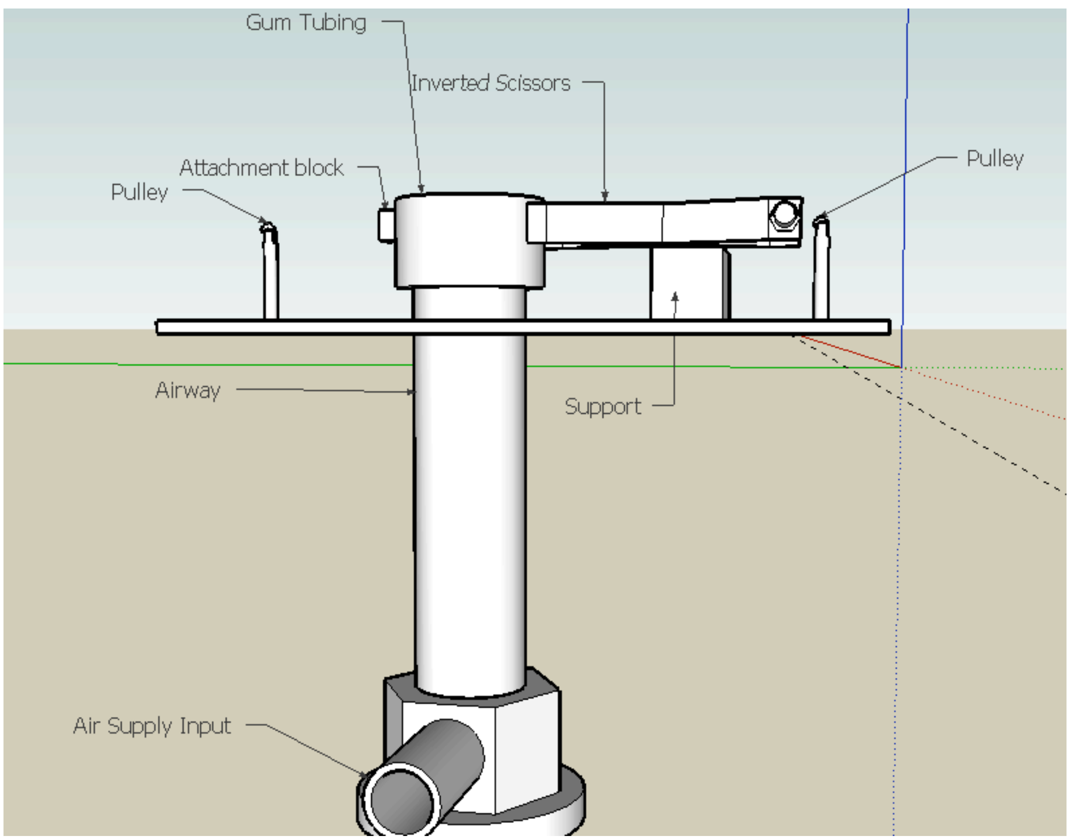


Figure 15: Vocal Fold Manipulation Mechanical Drawing

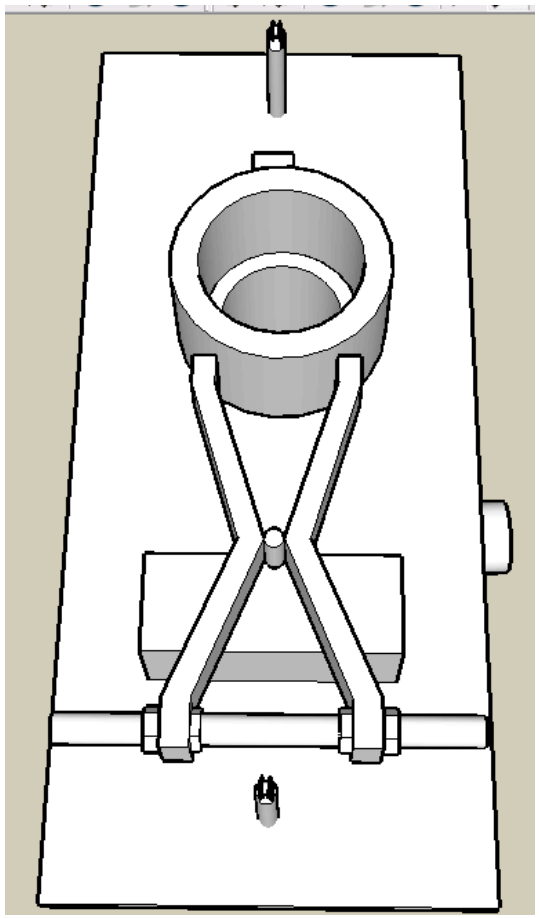


Figure 16: Vocal Fold Manipulation

The above figures illustrate the design for manipulating the model vocal folds. Detailed mechanical drawings with dimension can be found in the Appendix. The idea behind this design is to manipulate the vocal folds by deforming the material surrounding it. It is almost analogous to the mechanism of vocal fold manipulation discussed in earlier. The model vocal folds will be glued onto elastic rubber tubing. The scissors are attached to the scissors by an easily detachable, yet strong material such as Velcro. The inverted scissor like structure will function as the arytenoids that abduct and adduct to move the vocal folds in and out the air way. In the figure above the scissors are considered to be in the abducted position. So when the posterior ends of the scissors are brought closer together, the anterior end of the scissors adducts. This will cause the tubing to deform outward and effectively open the air way.

To tense the vocal folds to adjust pitch, notice in the anterior of the rubber tubing there is an attachment block (for material such Velcro). It is at this point that a string will be connected to a pulley. Likewise at the posterior end, a string connected to a pulley will attach to the inverted scissors. The user will pull on the strings which are guided by a rod not shown in figures above. Underneath the inverted scissors is a block of plastic material used to support the scissors from bending downwards.

As mentioned, the model vocal folds will be glued into the rubber tubing. This means that multiple vocal fold models and glued onto the rubber tubing. Provisions for attaching the manipulators must be made to allow easy transport and set up. Attachment through Velcro seems most suitable.

Resonator Module

Requirements:

- Will receive sound produced from vibration of vocal fold and amplify the sound.
- Will vary in size and length depending on the user preference
- Will be easy attached to larynx structure

Specifications:

- Length and size must be integer multiples of tube used to supply air to vocal fold model.
- Attachment with model must not disturb vocal fold model or prevent it from producing sound.
- Must be equipped with an adapter that fits into larynx structure.

The resonator module connects above the vibratory module and amplifies the sound produced from vibration of the model vocal folds. In the human voice process, the resonators are the throat, nasal passages and mouth cavity. The help give personal quality to a voice. In this design, tubes of vary length and size will be used simulate resonance of sound produced. Using tubes of vary length will demonstrate how the resonator adds a “personal quality” to the sound that is produced from the vibration of the model vocal folds. This module will be equipped with an adapter to fit into larynx structure.

Pre-Amplifier Module

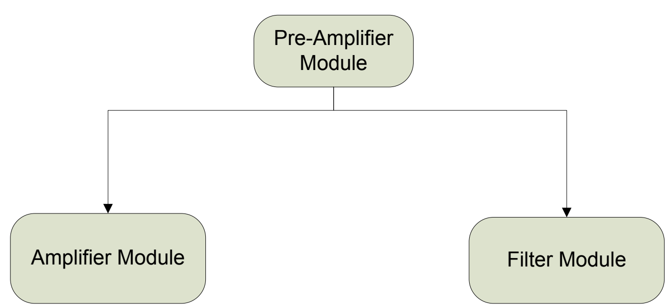


Figure 17: Pre-Amplifier Module

Requirements:

- Will receive an audio signal as an input from the microphone
- Will filter the sound signal to remove high and low frequency noise
- Will amplify the sound signal.
- Will output modified sound signal to the PC.

Specifications:

- Amplifier will produce a variable gain from 10V/V to 200V/V or enough to bring audio signal to or above voltage line level.

- Low pass filter will cut off frequencies above 5+/- 1 kHz
- High Pass filter cut off frequency 30 Hz +/-5 Hz
- Minimum Low pass filter roll off : 20dB+/- 10 per decade
- Minimum High pass filter roll off: 60dB
- Minimum power supply: 6V +/- 0.3V DC battery to allow easy transport.
- Circuit should include microphone jack for connection to microphone

The pre-amplifier or preamp module will receive the voice signal from the microphone and first amplify the signal then filter out frequencies out 30-5000Hz. The signal will then sent into the PC for analysis. Audio signals are generally low voltage in the range of a few millivolts. Microphones convert an audio signal into an electronic signal, thus depending on the type of microphone used the audio signal will need amplification. Thus the amplifier design requires an amplifier circuit with a variable gain from 10 to 200 V/V. This will ensure that the signal is amplified to a level sufficient for analysis.

The filters will prevent frequencies above 5 kHz from being analyzed. And it will also ensure that lower frequencies, below 30 Hz, usually due to moving the microphone or blowing into it do not interfere with the data analysis. The filter design selected to meet the requirements is the Sallen –Key topology Butterworth filter. This filter design was chosen because the Butterworth filter does not introduce a ripple in the passband, which is desirable for this project. The order of the low pass filter can be as low as a first order filter because in the software module there is a sharp filter that will perform low pass filtering also at 5 KHz. The high pass filter however needs to be very sharp because low frequency noise introduces very large spikes that will interfere will signal in the software module. Small capacitors should be included at inputs and outputs of each part of the circuit to eliminate low frequency noise.

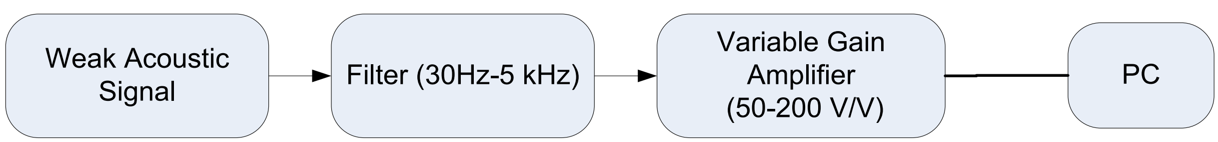


Figure 18: Signal Flow of Preamplifier Module

The figure above shows the signal flow of the preamplifier module. The weak sound signal into the microphone goes into a filter, reducing the frequency range of the signal. Then the signal is amplified and sent to the PC where analysis is performed. The figure below is what the schematic and simulation of the preamplifier module should look like.

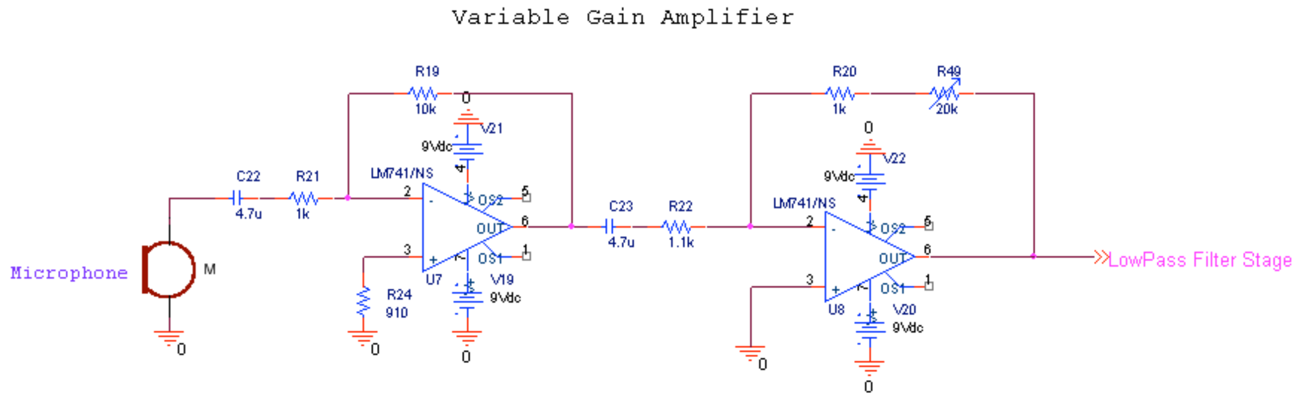


Figure 19: Amplifier Circuit Schematic

The amplifier circuit is a cascade of two inverting amplifiers that will produce a gain above 10V/V. The amplifier stage receives its input directly from the microphone and the output goes into the low pass filter stage. Blocking capacitors, C22, and C23 were used to cutoff low frequency noise before entering the filter stage. This reduces the effect of amplifying noise.

The filter module has two filters. A high pass filter and a low pass filter. The combined configuration results in a bandpass filter.

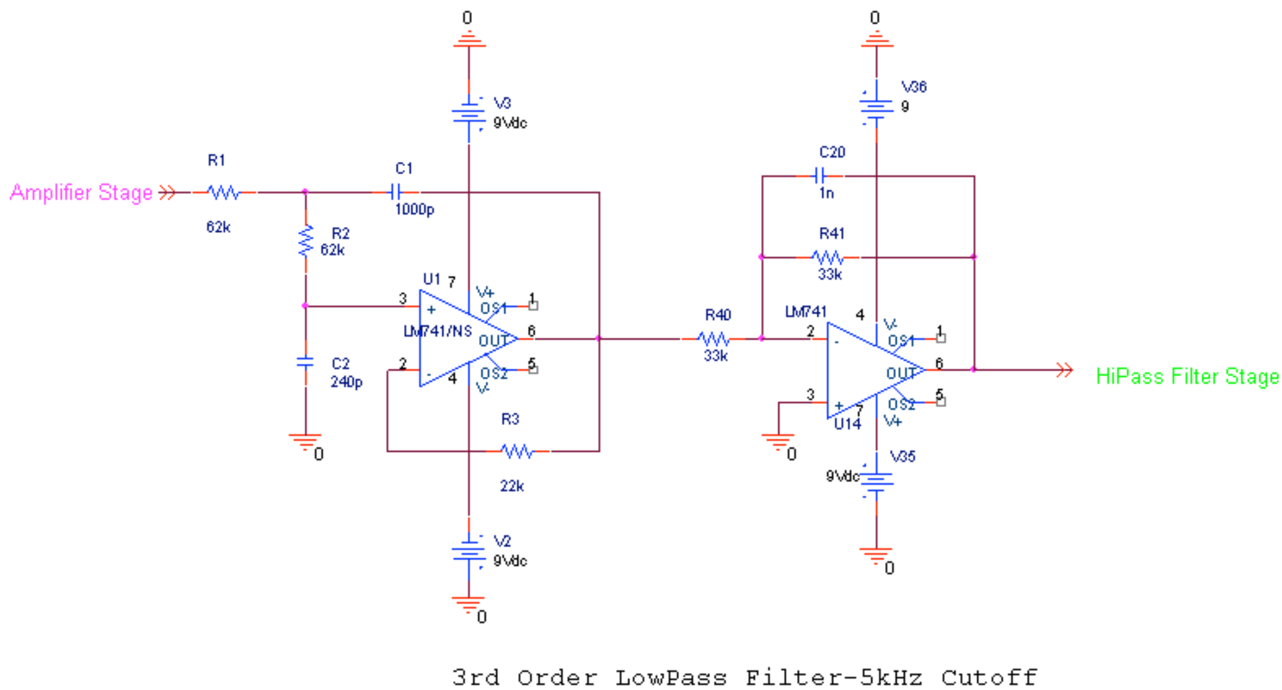


Figure 20: Low Pass Filter

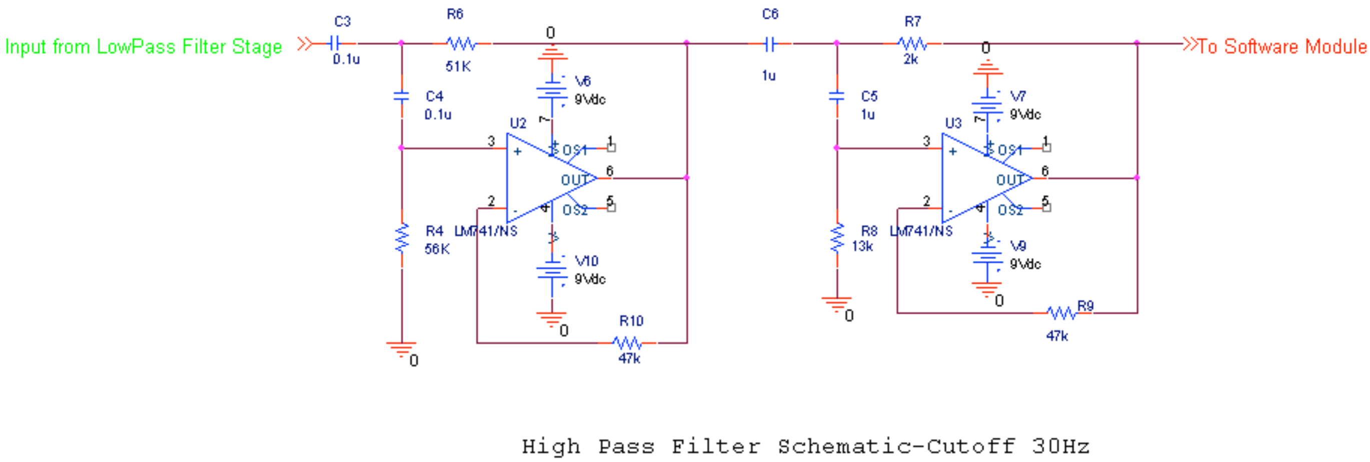


Figure 21: Hi pass Filter Schematic

This filter does not necessarily need to cut off precisely at 5 kHz because the next module will introduce a digital filter that will cut off also cut of 5 kHz. 5 kHz was randomly chosen to because higher frequencies are generally associated with noise. The human vocal folds vibrate within a frequency range of 80 to 300 Hz, sometimes higher during singing. Since this project is focused on human vocal fold vibration, cutting off around 5 kHz seems very reasonable. This filter does not necessarily need to cut off precisely at 5 kHz because the next module will introduce a digital filter that will cut off all frequencies above 5 kHz. For more convenience and to ensure less noise, this module will be powered with a 9V DC battery. An alternate design would to power the module with a power supply; however, building a power supply and joining it with the filter and amplifier stage may introduce some additional interference (noise) with the signal which may not get filtered out.

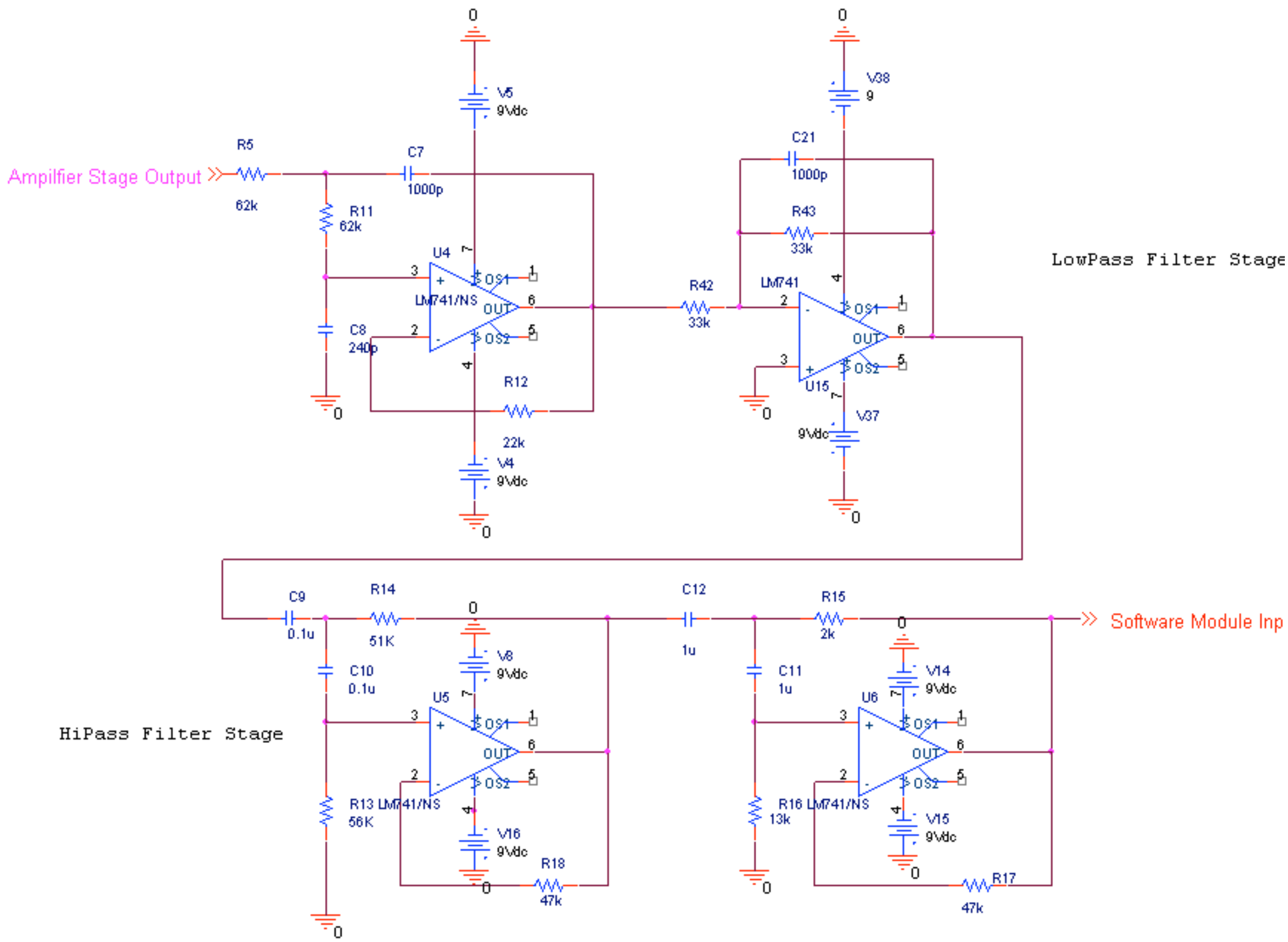


Figure 22: Preamp Module-Combined Filters and Amplifier

The above figure illustrates the cascade of both the filter and amplifier stages to form the preamp module.

Voice Analysis System Design

A top level execution flow diagram for the sequence of processes in the Voice Analysis system (Software Module) is shown below. Upon starting the program, the user will be prompted to provide a voice signal either through a WAV file of a previously recorded signal or a microphone. If no data is provided an error message will appear and it will prompt the user to start again. If sound data exists it will move through the sequence to perform the next functions as shown in the figure. At the final module it will compile and generate a report including the results of the calculated parameters along with normative values of measured parameters. This allows the clinician determine if the voice is normal or disordered. The software will not attempt to classify the nature of the voice. After displaying the results, the user will have the option to repeat the program or exit.

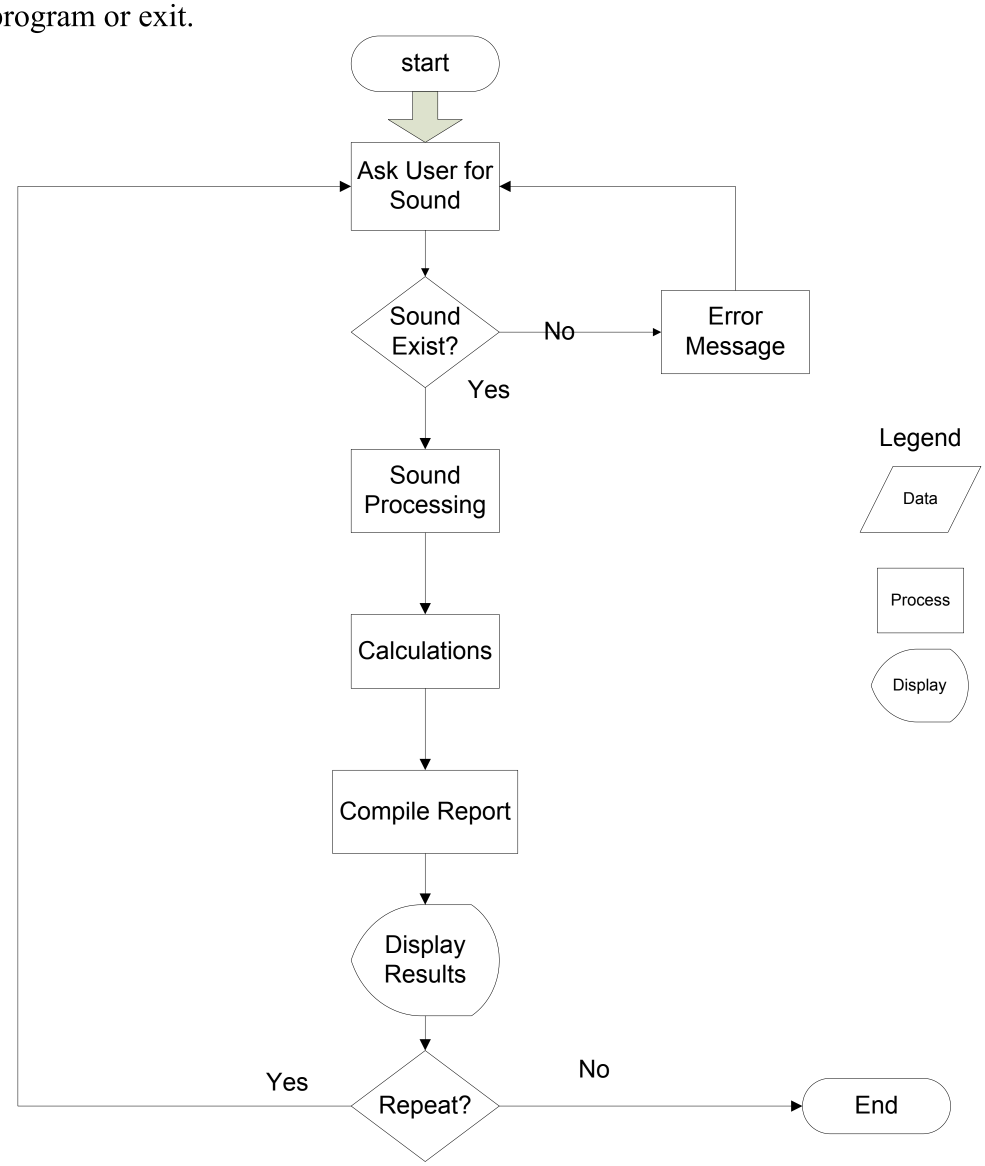


Figure 23: Top Level Execution Flow Diagram of Software Module

Control Display Interface Module

Requirements:

- Provide user friendly interface for viewing plots of the acquired audio signal, spectrogram, and frequency spectrum
- Allow user view results of analysis
- Allow user select and enter input data for report compiler

Specifications:

- Audio signal waveform will be a plot of amplitude vs. time
- Spectrogram waveform will be a plot of frequency vs. time
- Frequency spectrum should be decibels vs. frequency
- Export results in TXT file format
- User Input
  - Via microphone: 4 seconds voice data. 50-5000Hz bandwidth
  - Via sound file: WAV file type. 44.1KHz sampling frequency
  - Age: Integer value from 0-150

This module is responsible for communication between the user and the software. It is where the user will provide input necessary for analysis and view results of analysis. This module will plot all required signals: the audio signal, spectrogram and frequency spectrum. The start of this module is when the program starts and the sound has been acquired. It runs throughout the program updating the plots with the respective waveforms. This is also where the user enters data regarding the subject such as age and sex, which will be used in compiling the results.

Graphical User Interface (GUI) Architecture Design

The display module is effectively the graphical user interface or GUI of the software. The GUI will have options for the user to record and upload audio signals. It will display the signal that will be analyzed. The user will be able to highlight a region he/she is interested in analyzing and press Analyze to perform the signal analysis. The GUI will then display the frequency spectrum and spectrogram of the signal that has been selected for analysis.

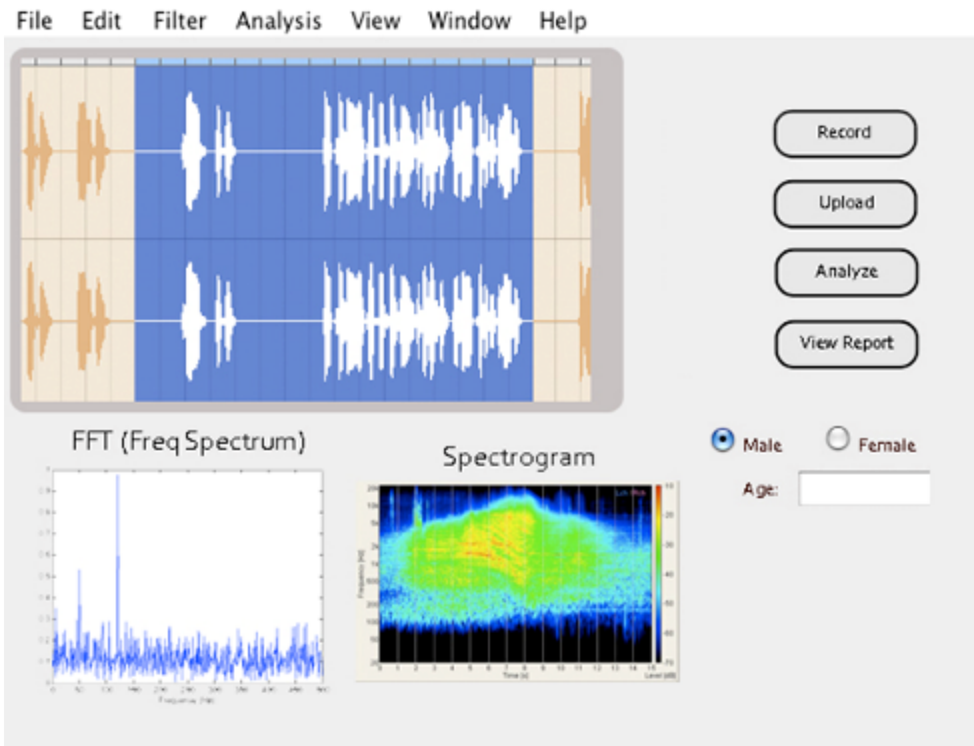


Figure 24: Mock Voice Analysis GUI Architecture[de]

Sound Pre-processing Module

Requirements:

- Receive voice signal from microphone that has been filtered and amplified
- Sample voice signal through PC Soundcard
- User can decide to analyze signal directly from microphone or previously recorded signal
- Filter the audio signal received.
- Allow user to select data range for analysis

Specifications:

- Sampling frequency of input signal must be 44.1 kHz
- Maximum signal acquisition length of 4 seconds
- Store acquired signal as WAV file format
- Low pass filter voice signal at a cut off frequency of 5 kHz.
- Maximum user data selection range of 1 second for analysis

---

[d] FFT waveform obtained from : http://www.mathworks.com/access/helpdesk/help/techdoc/ref/new_fft2.gif
[e] Spectrogram waveform from: http://www.ymec.com/hp/signal/images/sp002.JPG

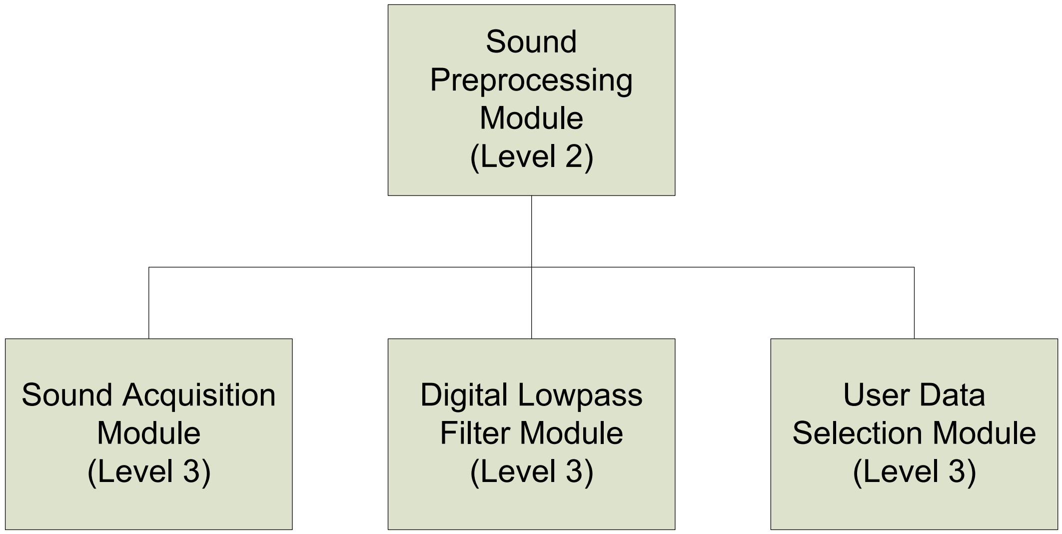


Figure 25: Sound Preprocessing Decomposition

The Sound Pre-Processing Module is primarily responsible for acquiring the voice data for analysis. It acquires the voice signal to be analyzed either through a microphone or from a WAV file of a prerecorded voice signal through the Sound Acquisition sub-module. The Digital Low pass Filter sub-module will filter the audio signal to ensure that no frequencies above 5 kHz are analyzed. This will serve to smooth the acquired audio signal, making it easier for the user to select one second region for analysis. This cutoff frequency will preserve a sufficient number of pitch harmonics for accurate pitch detection and other analyses. The User Data Selection Module allows the user select the range of data to be analyzed. The user can select up to 1 second of data.

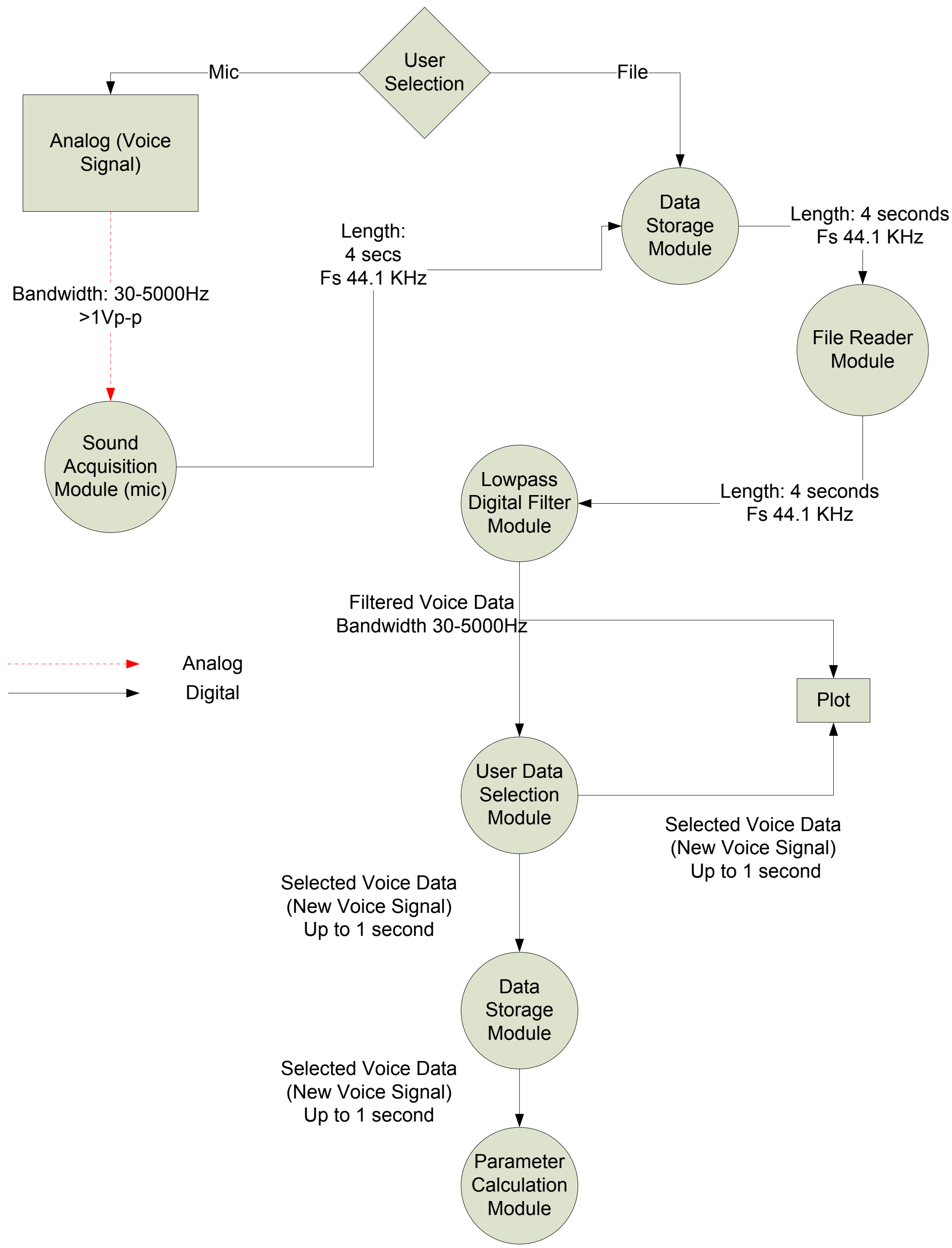


Figure 26: Sound Preprocessing Data Flow Diagram

Figure 25 shows the modular decomposition of the Sound Preprocessing Module. In the figure above, Figure 26, the interactions between these modules are outlined. Based on the user selection from the GUI, the sound preprocessing module will decide to take its input either from a prerecorded signal or directly from the microphone. The Sound Acquisition module receives voice input from both input types. However, the voice signal from microphone must be sampled through the PC soundcard at 44.1 KHz to convert the analog signal into digital. The signal from the microphone is amplified and filtered by analog circuits discussed earlier before reaching the Sound Acquisition module. After the voice input signal is converted into a digital signal it is stored. The user can access this stored data later to perform analysis. Prerecorded signals can be obtained from signals recorded via the microphone at an earlier time or signals from another recording device. However, the sampling frequency of the prerecorded signal must be 44.1 KHz; otherwise it will not undergo analysis.

Once the signal has been acquired, the Digital Low pass filter module acts to smooth the signal and remove high frequencies above 5 KHz. For signals from the microphone-amplifier-filter stage, this further complements the work of the analog filters which may not have been able to cut off frequencies above 5 KHz. It is easier to implement close to ideal filters digitally than in analog. Also a low order analog filter is designed to filter out high frequencies, which does not produce sharp cutoff frequency.

The Data Range Selection module will allow the user to select a one second range from the voice signal waveform for analysis. This will be implemented by two cursors placed on the signal waveform to select one second range. Once this module receives the filtered signal, it will call a function in Display module to plot the filtered data. The user will then be prompted to select a one second sample using the cursors and the selected range becomes the new voice data which will undergo analysis in the parameter calculation module.

Each sub module was reduced to its function prototype which is an outline of how each function will be implemented in software using MATLAB programming language.

Sound Acquisition Module

Requirements

- Allow user to chose whether to acquire signal from microphone input or from previously recorded file
- Convert analog signal from microphone input to digital signal
- Limited acquisition time

Specifications

- Based on user selection acquire signal from microphone or upload prerecorded signal
- Store acquired signal from microphone signal.
- Sample analog voice signal from microphone at sampling frequency of 44.1 KHz
- Acquisition time of analog signal from microphone: 4 seconds
- Input from microphone
    - Bandwidth: 30-5000Hz
    - Greater than 1 Vp-p
- Input from file
    - Sampling frequency 44.1 KHz
    - WAV file format
    - Length: 4 seconds

Function

- This module performs the following functions
    - Acquire analog voice signal from microphone
    - Read signal from WAV file
    - Convert acquired analog voice signal to digital signal

Digital Low-pass Filter Module

Requirements

- Filter audio signal to remove unwanted frequencies
- Smooth audio signal for display

Specifications

- Input
    - Sampled voice signal 44.1 KHz
    - Maximum signal length: 4 seconds
- Output
    - Filtered signal with bandwidth less than 5000Hz.

Data Range Selection Module

Requirements

- Will prompt user to select data range for analysis
- Update display with selected data range

Specifications

- Data range -1 second
- Allow selection of data range from waveform display of audio signal

This module will reduce the length of the input voice signal to less than or equal to 1 second for analysis. It will prompt the user to select from the audio signal waveform a one second sample that will be used as the new voice data input for analysis.

Parameter Calculation Module

Requirement:

- Will perform all necessary calculations: Shimmer %, Jitter%, Average Fundamental Frequency, and Noise to Harmonic Ration (NHR).
- Output calculated parameters to report compiler

Specifications:

- Input signal length up to 1 second or 44100 samples depending on sampling frequency
- Frequency range: 30-5000 Hz
- Analysis will be performed using both time and frequency domain processing
- Calculated results will be displayed to user

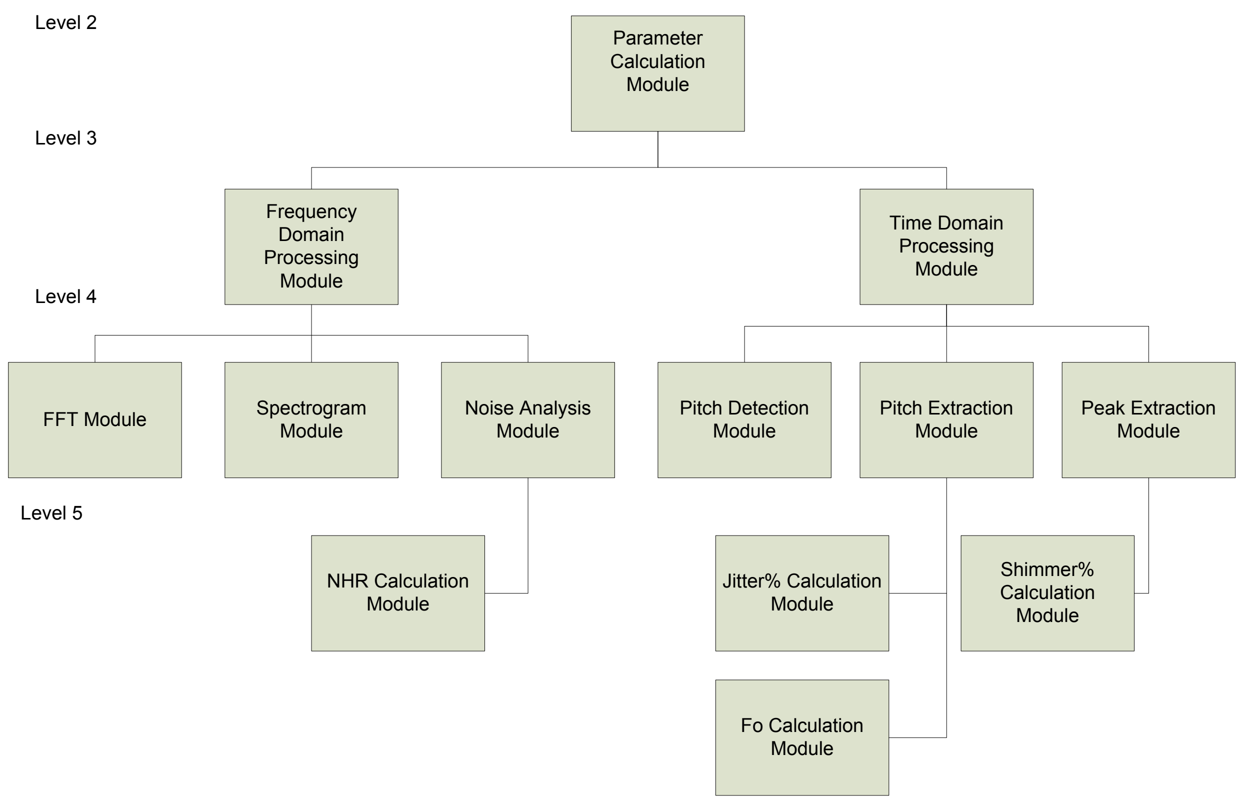


Figure 27: Parameter Calculation Modular Decomposition

The parameter calculation module is the bulk of the software module. It performs all major computations of parameters in the voice signal need by clinicians to make decisions about the quality of the voice. This module is split into two main parts: time and frequency domain module. The time domain analysis is computation performed on the preprocessed voice signal. The time domain processing will calculate the following: Shimmer %, Jitter%, and Average Fundamental Frequency. The formulas for these parameters as MDVP defines are discussed below[5]: The frequency domain processing performs analysis on the frequency content of the voice signal. It will compute the Fast Fourier Transform, the spectrogram, and the, and Noise to Harmonic Ration (NHR).

Time Domain Processing Module

Specifications:

- Input-voice data
    - Sampled at 44.1 KHz
    - Bandwidth : 30- 5000 Hz +/- 5%
    - Length of up to 1 second or 44100 points
- Output
    - Average fundamental frequency
    - Shimmer %
    - Jitter %

- Function-determine time domain based parameters
- Window length greater than 20ms

The execution flow diagram for this module is illustrated below. This is the sequence of how the functions in the module will be carried out. Once this module is called in the program (start) it receives voice data meeting specifications stated above. The input voice signal is partitioned into many 30ms-50ms windows (or segments). This allows for better analysis of the input signal. Think of it as looking at the data in 30-50ms chunks. Within each window a series of computation is then performed. First, a digital low pass filter that cuts off at 1800Hz is applied. The signal within the window is then coded into 1s, 0s and -1s through a technique known as non-linear sgn coding to simplify the voice data. This suppresses the effects of harmonic and sub-harmonic frequency components of the voice sound data. Following the non-linear sgn coding, autocorrelation is performed on the simplified signal. Autocorrelation will find patterns within the given signal and thus give a rough estimate of the fundamental frequency of the signal which will be verified. Pitch verification also involves autocorrelation but this time on the original signal. The verified pitch will then be used to extract all the periods from cycle to cycle in the original signal. Jitter %, Shimmer %, and the fundamental frequency F_0 are then calculated from the extracted period data.

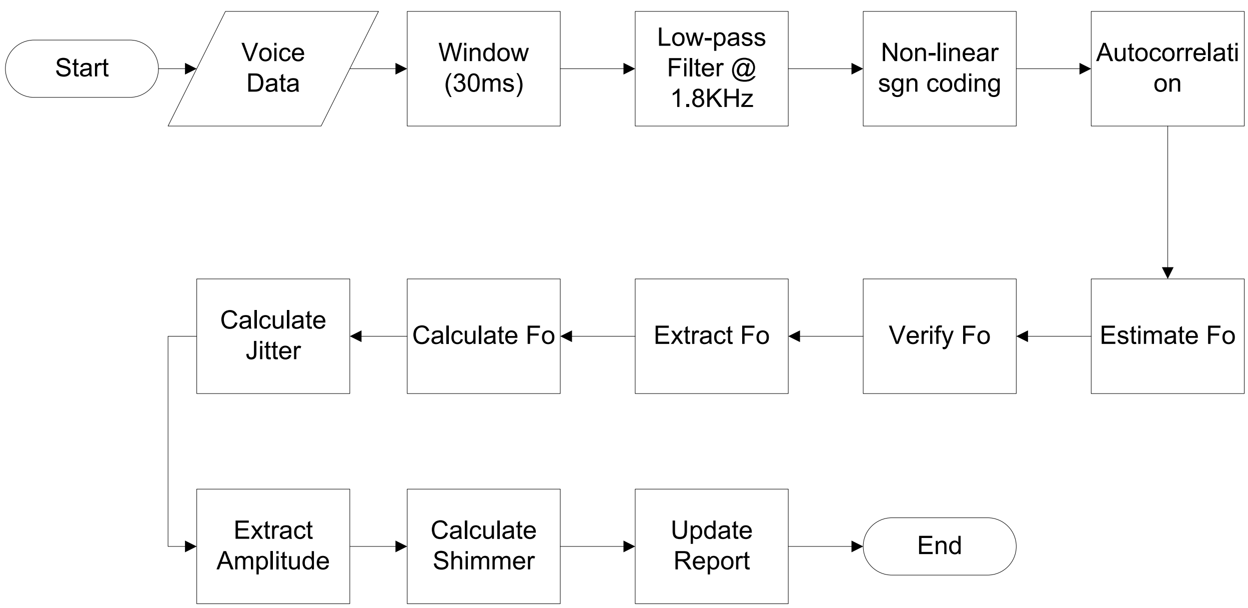


Figure 28: Time Domain Processing Execution Flow

The algorithm for pitch extraction was derived from the MDVP manual which can be found in Appendix A.1. In the figure below, the flow of data is shown. The Time domain analysis module receives input data from the Data Range Selection module. This one second long signal with a bandwidth of less than 5000Hz first undergoes a 30ms window. In each window, it is low pass filtered at 1.8 KHz. This is done to reduce effects of strong formant structure on the autocorrelation function[6]. In doing this, a sufficient number of pitch harmonics are still preserved. The filtered data serves as input for the Pitch determination module which simplifies the data to 1s, 0s and -1s then computes the autocorrelation function to determine the pitch. Using the calculated pitch result, the pitch and peak extraction modules can extract the cycle to cycle pitch and peak data and store the values in their respective arrays. The extractions are made from the voice signal from the Data Selection Range Module. The Jitter and Fundamental Frequency (F0) calculation modules use the pitch data array to compute the jitter percent value and average F0 respectively. The shimmer % calculation module uses the cycle to cycle peak array to calculate the shimmer percent value. The Results compiler module takes the calculated data from the different modules and arranges them into a report that will be displayed when user requests it via the Display module.

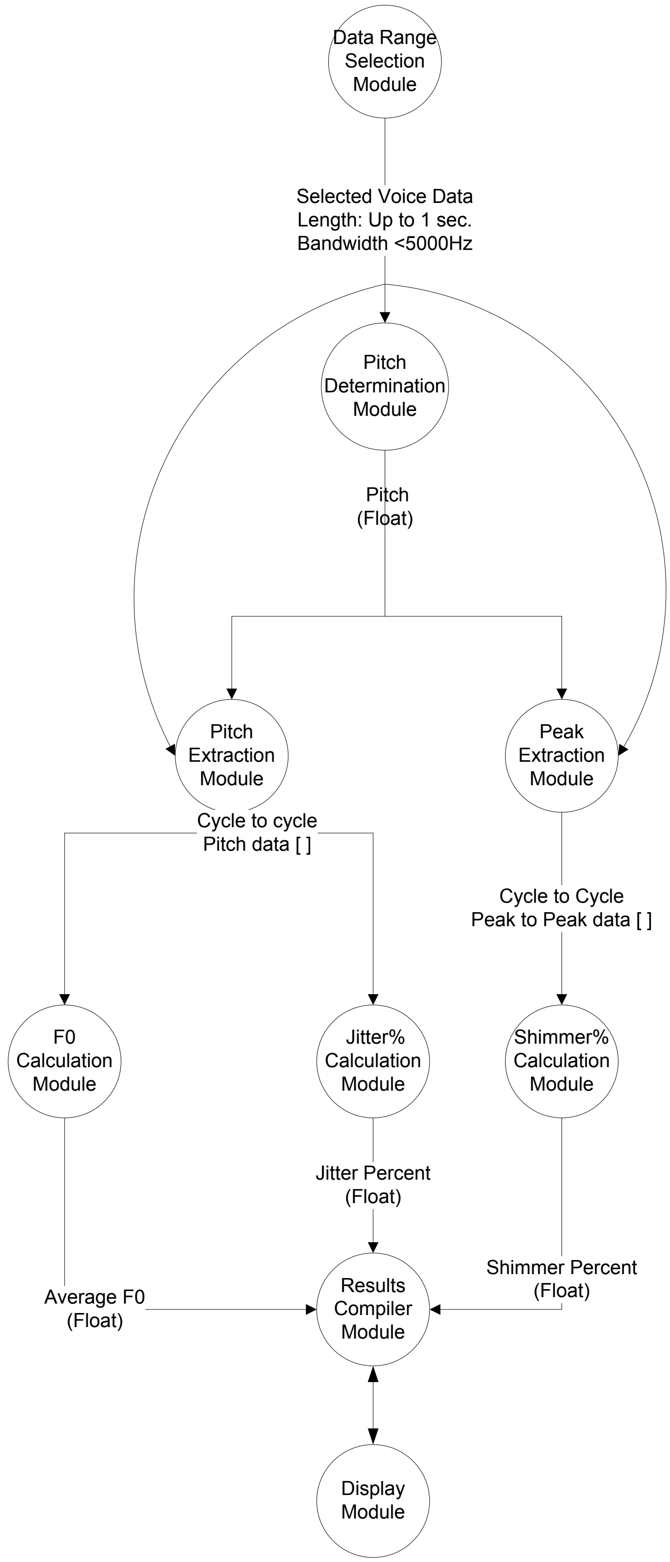


Figure 29: Time Domain Processing Module Data Flow

Pitch Determination Module

Specifications:

- Input-voice data
  - Sampled at 44.1 KHz
  - Frequency bandwidth 30-1800 Hz +/- 5%
  - Length of up to 1 second or 44100 points
- Output
  - Estimated fundamental pitch of voice data
    - Single float value
- Functions
  - Estimate pitch of voice signal
  - Verify estimate of pitch
  - Compute autocorrelation function of voice data
  - Provide basis for cycle to cycle pitch and peak extraction

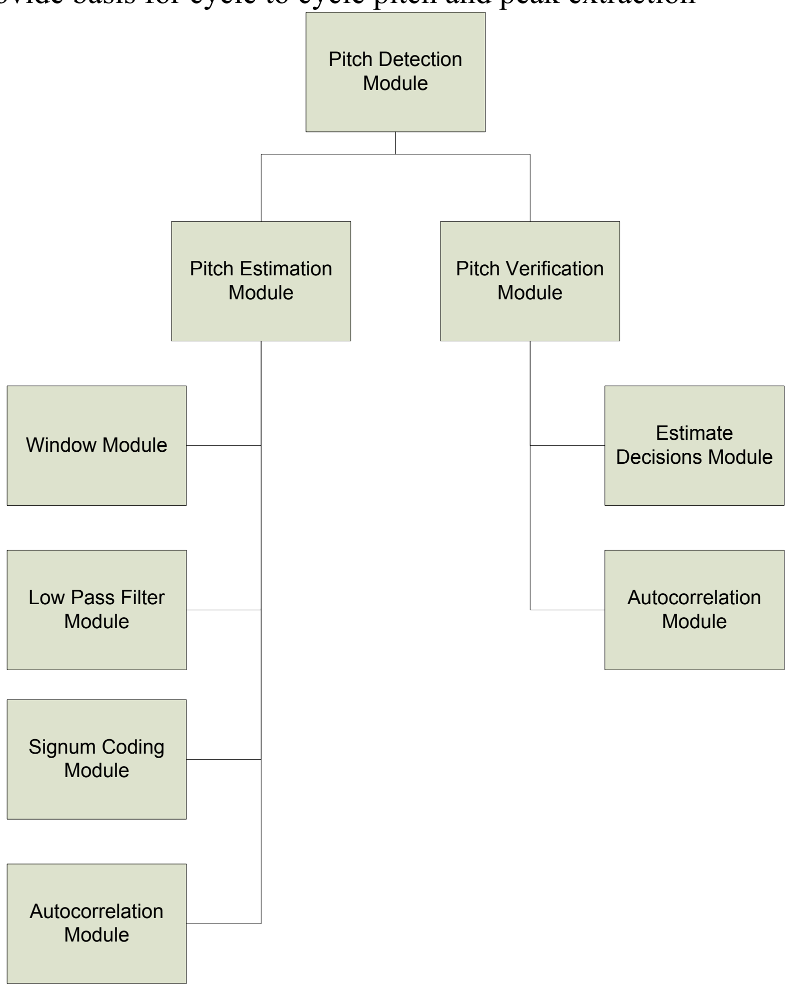


Figure 30: Pitch Determination Module Decomposition

From receiving the signal for analysis to calculation of fundamental frequency and perturbation (jitter and shimmer percent) involves four major steps: fundamental frequency estimation, verification, determination of cycle to cycle period data and extracted period data resolution enhancement. This module covers the first two major stages. In the pitch estimation sub-module, the input signal is subjected to a window of fixed length, usually above 30ms. The window size should be large enough to fit three or more periods of the signal. For example a voice signal of 100 Hz has a period of 10ms, thus a window size of 30ms or more would be ideal. The windowed signal is then filtered with a digital low pass filter to cut off frequencies above 1800 Hz. As mentioned earlier, to suppress the effects of strong harmonic and sub-harmonic components of the voice signal, a technique called non-linear sgn coding is performed on the windowed signal. This sgn coding is based on the Signum function. Basically what is does is apply a positive and negative threshold to the voice data and based on where points in the data fall a 1, 0, or -1 is assigned. For example, if the signal is above the positive threshold it is assigned a value of 1; if below the positive threshold but above the negative threshold a 0 is assigned and for values below the negative threshold, a -1 is assigned.

The autocorrelation function of the Signum coded signal within each window is computed. Looking at the positive half of the autocorrelation signal, the time period at which the maximum peak occurs is recorded as the fundamental period of the windowed signal. MATLAB has an inbuilt function that will perform the autocorrelation. A peak finding algorithm will need to be implemented to find the appropriate peak. This is done for each window of the signal. Once the pitch has been determined, a second round of signum coding is performed. However, the threshold value for the second round of signum coding is much lower than the first. This second signum coding suppresses the sub-harmonic components of the signal whereas the first suppresses the harmonic. Window based autocorrelation is performed once again on the newly signum coded signal and the pitch within the window is measured.

Once the two pitch estimates have been calculated, one from high threshold Signum coding and the other from low threshold Signum coding, the next step is to make some decisions as to what is the correct pitch. This is the start of the verification stage or sub-module. The Estimate Decisions module will compare the results of the two pitch estimate data by finding the median frequency between two estimates of the same window. This module will also correct incorrect pitch estimates based on some logical decision. Take for example the analysis of a signal that is one second long with a fundamental frequency of 120 Hz. And in window number 5, the first pitch estimate based on a high threshold Signum coding is 122Hz and the second pitch estimate based on the low threshold estimates a fundamental frequency of 119 Hz. Thus, median fundamental frequency of the two is 121.5 Hz. But in window number 10, the first pitch estimate maybe 365Hz and the second pitch estimate maybe 122 Hz. What this module will then do in this case would be to first check to see if pitch estimate from the first estimate is greater than twice that of the second estimate and vice versa. It will then replace the incorrect estimate with the other estimate. That is 365Hz will be replaced with 122Hz as the new pitch estimate. This correction is mainly in place for the first pitch estimate as it will not provide a very accurate pitch estimate in comparison to the second pitch estimate. Future enhancement of the present software can involve replacing the values erroneous estimates with the mean or median of the pitch estimate data. A correction should also be applied if either value is significant less than its respective counterpart and if either value is zero.

A third pitch estimate is then performed directly on the original voice signal and not the signum coded versions of the signal. This pitch estimate also uses a window based autocorrelation. However this window size is determined by the corrected pitch data obtained from the Estimate decisions module. The window size to be used should be ten times the inverse of the median pitch estimated from the previous sub module. The resulting pitch estimate from this sub module is coined the verified pitch, which is much more accurate than the first two estimates. Similar to the first two pitch estimators, the pitch in each window is also measured. The mean of the third pitch estimate data is used to find the cycle to cycle peaks and pitch. This is the premise of the Extraction Module.

Extraction Module

Specifications:

- Input
  - Estimated Pitch Data
    - Data type: double
    - Size: single value
  - Voice signal
    - Sampled at 44.1 KHz
    - Bandwidth 50- 1800 Hz +/- 5%
    - Length of up to 1 second or 44100 points
- Output
  - Cycle to Cycle peak amplitude
    - Size: Number equal to approximately the fundamental frequency
    - Data type: Doubles
  - Cycle to cycle pitch periods
    - Size: Original equal to approximately the fundamental frequency
      - Because of interpolation size increases by interpolation factor
- Function
  - Extraction of cycle to cycle pitch periods
  - Extraction of cycle to cycle peak to peak amplitude

The peak extraction and pitch extraction modules are lumped into the category of extraction module because they are performed in the same manner. The mean verified pitch (or third pitch estimate) is used to extract the cycle to cycle pitch data. The inverse of the estimated fundamental frequency is the approximate fundamental period of the signal under analysis. Therefore, an algorithm can be implemented to move in time blocks the size of the estimated period and find the maximum value or peak and then also find the time where the peak occurs. This is then repeated for the entire signal and the peak value and time of occurrence recorded and stored in an array. The collected cycle to cycle pitch data is interpolated to increase the resolution of the extracted pitch data. This data can then be used to perform the necessary parameter calculations.

Shimmer Percent Calculation

Shimmer is the relative evaluation of the period-to-period (very short term) variability of the peak to peak amplitude within the analyzed voice sample. Voice breaks are excluded. It is reported as a percent. Shimmer is computed from the extracted peak to peak amplitude data.

$$Shimmer = 100\% * \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|A^{(i)} + A^{(i+1)}|}{\frac{1}{N}\sum_{i=1}^{N} A^{(i)}}$$

Equation 1: Shimmer Equation

Where: $A^{(i)}$, i=1,2…N-extracted peak to peak amplitude data,
N=number of extracted impulses

Jitter Calculation

Jitter is the relative evaluation of the period-to-period (very short term) variability of the pitch within the analyzed voice sample. Voice breaks are excluded. It is reported as a percent. Jitter is computed from the extracted peak to peak pitch data.

$$Jitter = 100\% * \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}|T^{(i)} + T^{(i+1)}|}{\frac{1}{N}\sum_{i=1}^{N} T^{(i)}}$$

Equation 2: Jitter Equation

Where: $T^{(i)}$, i=1,2…N-extracted pitch period data,
N=number of extracted pitch periods

Fundamental Frequency Calculation

Average fundamental frequency is the average value of all extracted period to period fundamental frequency values. Voice breaks excluded. It is also computed from the extracted period to period pitch data.

$$F_0 = \frac{1}{N}\sum_{i=1}^{N} F_0{}^{(i)}$$

Equation 3: Average Fundamental Frequency Equation

$$Where: F_0 = \frac{1}{T^{(i)}} - period\ to\ period\ fundamental\ frequency$$

$T^{(i)}$, i=1,2…N-extracted pitch period data,
N=number of extracted pitch periods

Relative Amplitude Perturbation (RAP)[f]

Relative evaluation of the period to period variability of the pitch within the analyzed voice sample with smoothing factor of 3 periods. This is similar to the measuring Jitter percent.

[f] Newly introduced parameter calculation—FPR

$$RAP = \frac{\frac{1}{N-2}\sum_{i=2}^{N-1}\left|\frac{To^{(i-1)} + To^{(i)} + To^{(i+1)}}{3} - To^{(i)}\right|}{\frac{1}{N}\sum_{i=1}^{N} To^{(i)}}$$

where: $To^{(i)}$, $i=1,2...N$ - extracted pitch period data,

$N = PER$ - number of extracted pitch periods.

Frequency Domain Processing Module

Specifications:

- Input-voice data
  - o Sampled at 44.1 KHz
  - o Bandwidth :50- 1800 Hz +/- 5%
  - o Length of up to 1 second or 44100 points
- Output
  - o NHR
  - o Power Spectrum
  - o Spectrogram
- Function-convert voice data to frequency counterpart and compute parameters associated with frequency domain

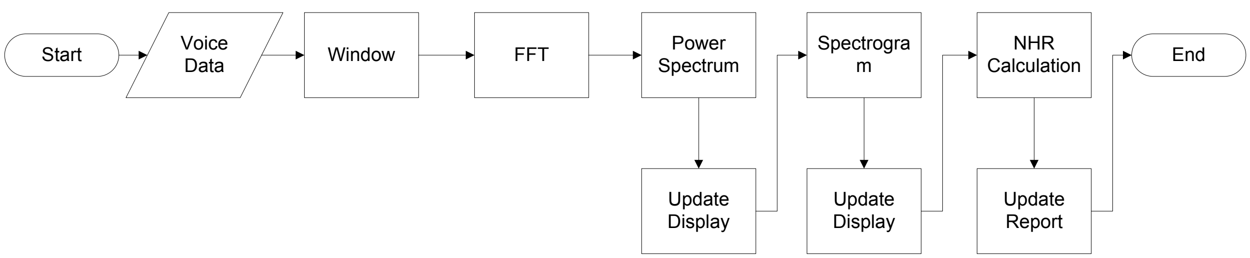


Figure 31: Frequency Domain Processing Execution flow

The frequency domain processing module is needed to calculate the noise to harmonic ratio, power spectrum and spectrogram. A window is first applied to the preprocessed voice data and the FFT is computed. From the FFT the power spectrum is computed and updated on the GUI display. The spectrogram is plotted next and calculations for the noise harmonic ratio are performed finally and the results are displayed in the report.

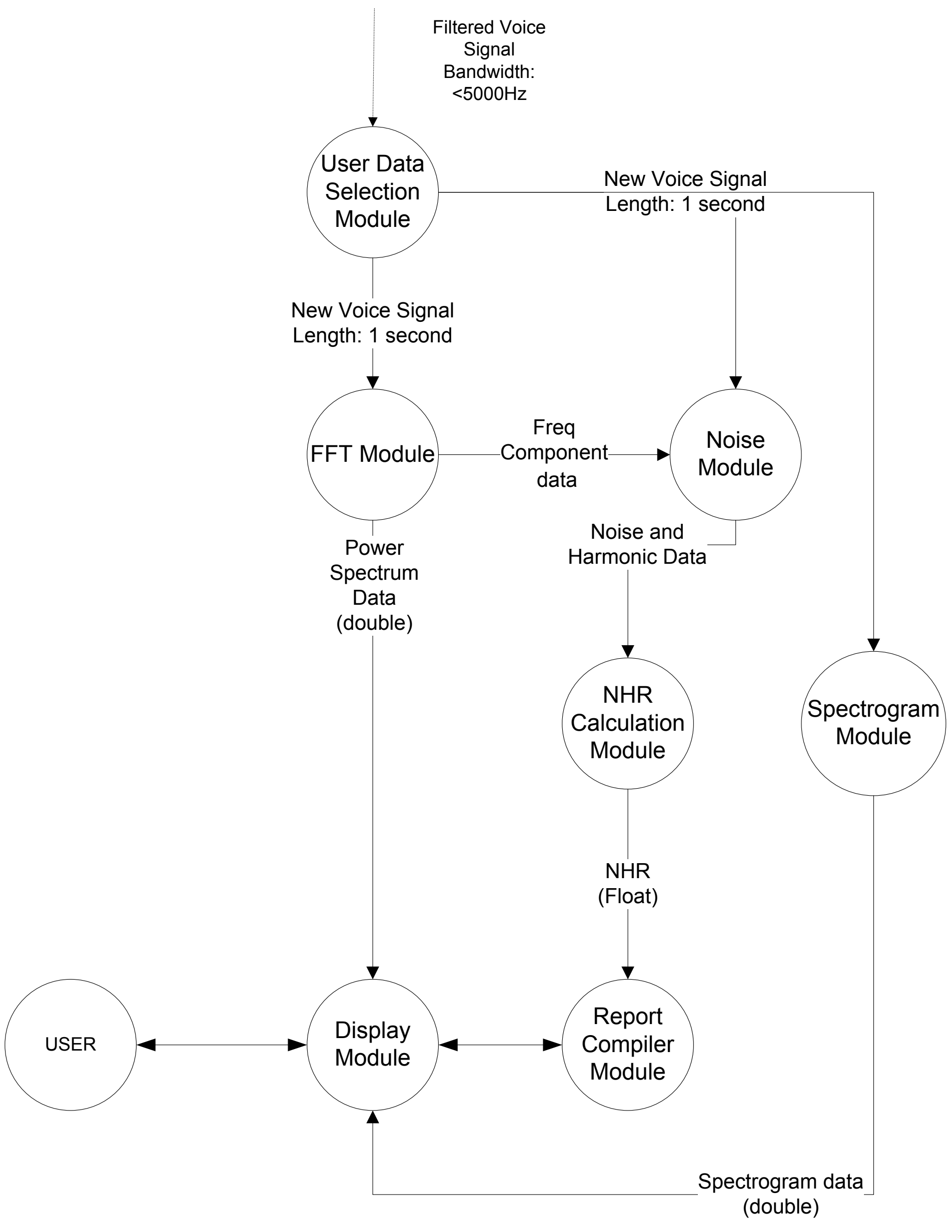


Figure 32: Frequency Domain Processing Data Flow

The data reduction module provides the input for the frequency domain processing module. The FFT module produces two outputs. It feeds one output to the Noise module and the second to the results complier which contains information for plotting the power spectrum of the voice signal. The result compiler compiles and displays the results to the user.

Noise to Harmonic Ratio Calculation

Noise to Harmonic Ration is the average ratio of inharmonic spectral energy to the harmonic spectral energy in the frequency range 70-4200Hz. This is a general evaluation of noise present in the analyzed signal. NHR is computed using a pitch-synchronous frequency domain method. The algorithm can be found in the appendix A2 of this document.

FFT Module

To produce the frequency spectrum the Fast Fourier Transform (FFT) technique will be used. This is the most common technique used in speech analysis. It is also very simple to implement. The FFT converts the signal from a time domain to a frequency domain. The frequency spectrum as specified earlier is a function of power in decibels vs. frequency. The number of points for the FFT should be 1024. This is the same number of points that the Multi-Dimensional Voice Program (MDVP) uses to perform its frequency spectrum. MDVP is a gold standard voice analysis program[g].

Spectrogram Module

The spectrogram is another common waveform used in voice analysis. It is a three-dimensional plot of the energy of the frequency content of the audio signal as it changes over time. In the most usual format, the horizontal axis represents time, the vertical axis is frequency and the intensity of each point in the image represents amplitude of a particular frequency at a particular time. Often the diagram is reduced to two dimensions by indicating the intensity with thicker lines, more intense colors or grey values[h]. This can be easily implemented in MATLAB.

Results Module

Requirement:

- Compile results from modules that return parameter values in tabular form along with normative parameters
- User should be able to export results

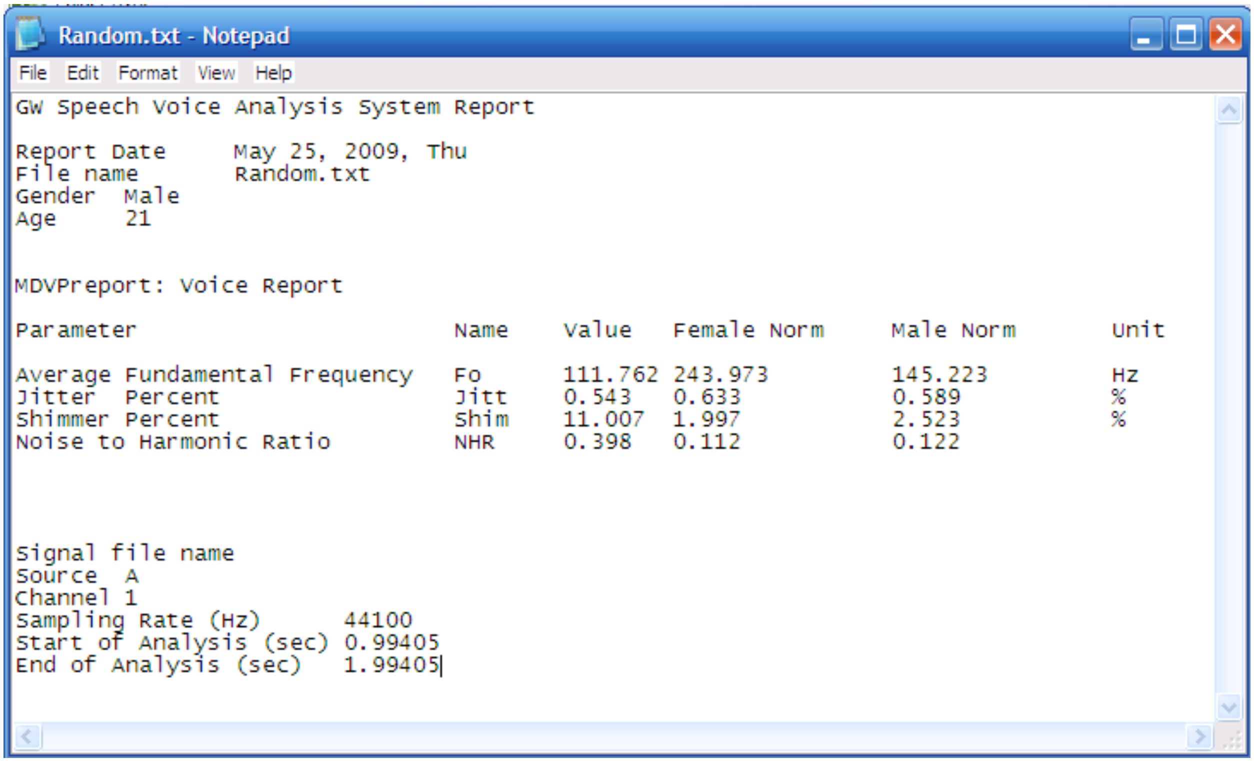
Specifications:

- Input
  - Calculated parameters from different modules
  - Normative parameters
- Output
  - Updated GUI
  - TXT file containing all results
  - Graphical representation of calculated results

This module compiles all the results from the modules performing calculations such as NHR, Jitter, Shimmer, and average fundamental frequency.

---

[g] http://www.kayelemetrics.com/Product%20Info/CSL%20Options/5105/5105.pdf
[h] http://en.wikipedia.org/wiki/Spectrogram



Random.txt - Notepad

File Edit Format View Help

```
GW Speech Voice Analysis System Report

Report Date     May 25, 2009, Thu
File name       Random.txt
Gender   Male
Age      21


MDVPreport: Voice Report

Parameter                       Name    Value   Female Norm     Male Norm       Unit

Average Fundamental Frequency   Fo      111.762 243.973         145.223         Hz
Jitter  Percent                 Jitt    0.543   0.633           0.589           %
Shimmer Percent                 Shim    11.007  1.997           2.523           %
Noise to Harmonic Ratio         NHR     0.398   0.112           0.122



Signal file name
Source   A
Channel 1
Sampling Rate (Hz)      44100
Start of Analysis (sec) 0.99405
End of Analysis (sec)   1.99405
```

Figure 33: Voice Analysis Mock Report[i]

The figure above is mock report generated by results module. The report includes all calculated parameters and normative values. The results module will also plot a bar graph representing the calculated parameters. Also included in the report is information about the signal such as sampling frequency, length of signal analyzed and microphone channel.

Key Component Selection

The heart of this design is the model vocal folds. The material used to develop this model will be crucial in that the specifications for the model states that vibration of the model vocal folds should be similar to that of a human. The graduate student who developed model vocal folds that share similar vibration pattern as the human vocal folds was reduced to experimenting because there was no reliable data giving specific elastic properties of the vocal folds. Many people have tried to measure the elastic property, specifically the elastic modulus, of the vocal folds and there data vary significantly from one experiment to the other because of mode of measurement, type of vocal folds (canine or human), state of vocal folds (phonation, dead, alive), and environment (*in vitro* or *in vivo*). There are too many parameters involved and the data vary by more than a factor of 2 which is not very reliable for the purpose of this project. In the table below, the different papers are summarized briefly with their values.

Paper	Authors	Youngs Modulus(kPa)	Shear Modulus (Pa)	Nature of VF	Mode of Measurement

[i] Modified report from MDVP Software

Elasticity of Canine VF Tissue	Perlman et al	13.5-94.6		Insitu, excised	Excised VF from canine (both dead and alive), many test conditions. Not conclusive results
In vivo measurements of the shear modulus of human vocal folds	Goodyer		701-2225	Invivo	Used device called laryngeal tensiometer-took measurements while patient under anesthesia
Elasticity of Human VF Measured In vivo using Color Doppler Imaging	Hsiao	M:30 to 120 & F: 120 to 300		Invivo	Used CDI to measure vibrating length then used fundamental frequency of vibrating string formula to calculate the tension which was used to find stress and subsequently Youngs Modulus
Measurement of Youngs Modulus in the Invivio Human	Tran		12.6-21.6	Invivo	Electrically stimulated laryngeal nerve, which caused VF to exert on force gauge

Table 1: Summary of Data Elastic Moduli Data

The state of the art will be the vocal folds developed by Mr. Fukui, the graduate student who is working on the talking robot. The material used to model the vocal folds was a compound called Septon which is a styrenic block copolymer. This compound can be purchased in form of pellets. It is heated to about 400°F into liquid form. The liquid can then be poured into a mold for casting. This is the technique used to make the model vocal folds. All other components can be substituted for something else, but as for the vocal fold material Septon will be the best way to go.

Testing Strategy

The major modules of this product will be tested to see if they meet the specifications and requirements set in the earlier section. Tests will be performed at each phase of the construction and integration processes to ensure that each module meets its requirement and specifications. More detailed testing strategies are discussed in the Product Tests Review section in page 61.

Hardware Modules

Sound Capture Module (including submodules-filter and amplifier)

i. Theory and Description: This module will be tested to ensure that the sound produced from the vibratory module is captured via the microphone, amplified and filtered.
ii. Protocol: Test each component of the preamplifier circuit individually, then as a unit. First to test the filter signal, connect the input to a function generator and apply a small signal of $20mV_{P-P}$ at 100 Hz. Use an oscilloscope to measure the output of the filter and make sure that it is equivalent to the input. Increase the frequency (in 100 Hz increments) on the function generator and observe the oscilloscope for change attenuation of the signal at or around 5 kHz. The same should be repeated for the high pass filter. Apply same signal from function generator to high pass filter and vary the frequency. Test to see if signal attenuates at frequencies below 30Hz. Power the filter with 9V DC from a power supply. After which use a DC battery.
To test the amplifier, use the function generator, applying the same signal settings as before for the input observe the oscilloscope (which is connected to the output of the amplifier stage) to see an increase amplitude of the signal. The ratio of the output to the input (gain) should be above 180V/V.
To test both stages as a unit, connect the output of filter stage to the input of the amplifier stage and connect the oscilloscope to the output of the amplifier. Use the same function generator settings and increase the frequency in increments of 100 Hz and observe attenuation of the signal outside the range of 30-5 kHz. Also observe that the voltage is amplified in the output signal.
iii. Equipment: Function generator, oscilloscope and Tektronix TDS460A Four Channel Digital Oscilloscope, and Agilent Triple Output Power Supply, or equivalent.
iv. Success Criteria: Cut off frequency of about 5 kHz +/- .5 kHz and at 30Hz. Variable gain between 50-200V/V for all frequencies in filter pass band.

Vibratory Module

i. Theory and Description: This module will be tested primarily by perception of sound and vibration. The model vocal folds should also vibrate in a frequency range of 80 to 200 Hz in response to air flow. The larynx structure should resemble some variation of the human larynx.
ii. Protocol: The fundamental frequency of the vocal fold vibration will be measured using known software for speech analysis MDVP. Air from the supply tank is released into

the model (and the vocal folds should be in closed position). The vocal folds will vibrate and produce an audible sound. Stretch the vocal folds to change the pitch.

iii. Equipment: Air supply and MDVP software

iv. Success Criteria: As regulated air is introduced into the larynx structure the vocal folds will begin to vibrate. As it vibrates it should produce an audible buzzy sound (like when one says “ahhh”). If the vocal folds are stretched this should also change the pitch of the sound. The vocal folds can be moved in and out the air way.

Software Modules

i. Theory and description: All software modules will be tested individually using the same protocol.

ii. Protocol: Generate a sinusoidal waveform that has similar properties as the human voice. This can be done by adding a sine wave to a random function (white noise). Then use the signal as an input to each module. After successfully doing so use a real audio signal as an input.

iii. Equipment: MATLAB software, MDVP software or equivalent

iv. Success Criteria: a) all necessary waveforms plotted b) calculation of vocal parameters and c) comparison of sound signal

Labor Cost

The labor cost graph displays the cumulative actual, estimated and projected project expenditures. As the project proceeds and actual costs are updated, the labor cost graph will change to express the projected project costs. This graph can show at a glance the progress of the project in terms of cumulative cost versus time. Any over-budgeting or under-budgeting issues will be clear from this graph by the projected costs curve. The estimated rise at week 8 is due to the high concentration of design and hardware engineers and purchasing of components. The rise at the final weeks is due to final testing and documentation costs.

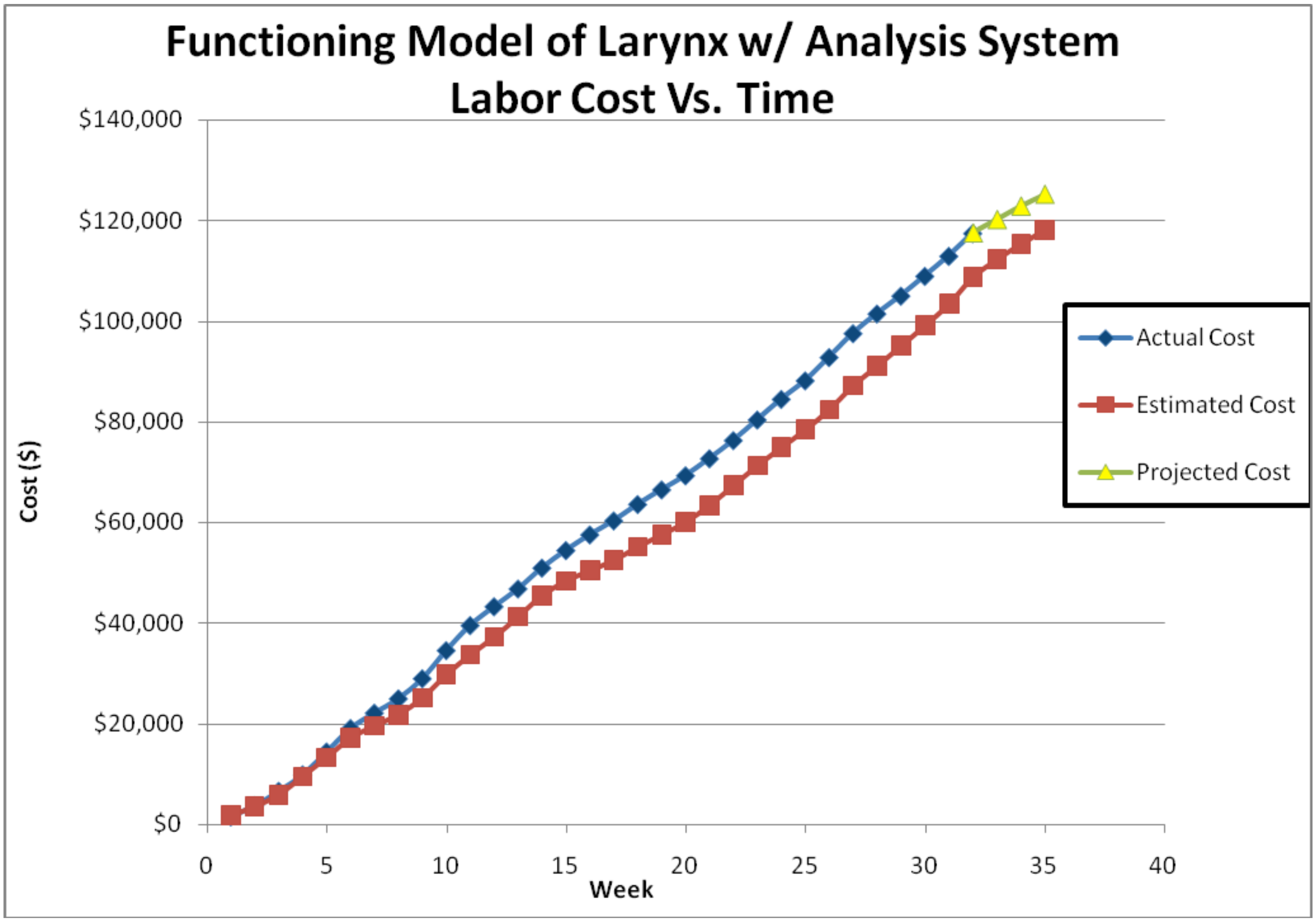


Figure 34: Labor Cost Graph

Economic Analysis

The functioning model of the larynx and software analysis system will be used as an education tool for students in Speech and Hearing Science and for patients suffering from common vocal disorders. It could also be used commercially. To the knowledge of the manufacturers, this product is unlike any product of its kind in the market. Other models of the larynx exist but they are stationary, without any functional parts such as the vocal folds which vibrate in response to air flow.

In the table below, is an outline of how much it will cost to produce and manufacture this model, including the cost of labor and parts.

PROTOTYPE COSTS:			
	Cost ($/hr)	Hours	Total Cost ($)
Project Manager	$66	350	$23,100
Design Engineer	$57	470	$26,790
Hardware Engineer	$48	519	$24,912
Software Engineer	$40	435	$17,400
Test Engineer	$36	411	$14,796
Technical Writer	$30	375	$11,250
Speech Pathologist (Consultant)	$100	25	$2,500
Overall Hours:	2585		
Overall Cost:	$120,748		
Multiplier:		2.8	
Labor Charged to Contract:		$338,094	
Parts Total:		$100.00	
Machine Shop, PCB fab. and pop.:		$50	
Sub-Total:		$150	
Pass-Through Fee:		1.05	
Total:		$158	
Cost of Labor:		$338,094	
Cost of parts & external services:		$158	
Total prototype cost:		$338,252	

Table 2: Prototype Cost

PRODUCTION COSTS:		
Manufacturing Process:		
Salary ($/hr)	20	
Hours	200	
Total:	$4,000	
Software Testing:		
Salary ($/hr)	15	
Hours	200	
Total:	$3,000	
Total Engineering	$7,000	
Multipler:		2
Production Labor Costs:		$14,000
Number of Units:		1,000

Parts Cost:	$120,000
Printing and Packaging Cost:	$120,000
Production Non-labor Costs:	$240,000
Overhead Cost Multiplier:	1.4
Profit Multiplier:	1.2
Total Production Cost:	$426,720
Production Cost per Unit:	$426.72

Table 3: Production Cost for 1000 units

PROJECT COST:	
Prototype Cost:	$338,252
Production Cost:	$393,120
Total Project Cost:	$731,372
Total Project Cost per Unit:	$731.37

Table 4: Project Cost for 1000 units

COST OF DISTRIBUTION:		
Cost per Unit:	$762.51	
Wholesaler Price:	$915.01	1.2 Multiplier
Retailer Price:	$1,372.51	1.5 Multiplier

Table 5: Cost of Distribution of 1000 Units

The economic analysis shows that a retailer price of $1,372.51 corresponds to 1,000 units produced. If 10,000 units are produced (which is estimated at $1,000,000 for parts and $120,000 for packaging), this price will be greatly reduced to $464.84.

Implementation Plan

The circuit for filtering and amplifying the voice signal will be built on a perfboard. This makes it easier to build, test, and redesign in comparison to using a printed circuit board (PCB). The circuit will be enclosed in a steel box to provide electrical shielding and prevent other external involves. The choice of material for the box is aluminum. The enclosure will have a power switch and LED to indicate that it is functioning.

The user will select the microphone of their choice for acquiring the voice signal. The circuit will use a standard XLR3 microphone connector to connect to the microphone. The XLR 3 connector is useful in sound recording because they reduce effect of external noise.

The airflow and resonator modules will be implemented with PVC pipes. These are cheap to use and can be easily cut to different sizes. The pipes will be transparent to increase visibility of vocal folds during operation. The supports holding the stand will be made using plastic from the rapid prototyping machine. This will make the model more attractive than using scrap metal from the machine shop which may also be heavy.

Implementation Progress Report

Work Anticipated for this Period

In the period between FDR until the submission of the IPR, I anticipated working on tying up loose ends in all modules of the design. In addition this, I wanted to familiarize myself with the CAD tool, Pro Engineering for creating drawings for the mold. Some minor issues arouse with arrangement of the circuit and schematic symbols used. I also anticipated work on the software design to make the design more fluid for the lay person by changing name conventions and specifying correct inputs and outputs including the function of the modules involved in software.

Work accomplished

Below is a block diagram illustrating the work accomplished. Green means complete, light green means over 65% complete, orange means in progress and the light blue or uncolored blocks mean not done.

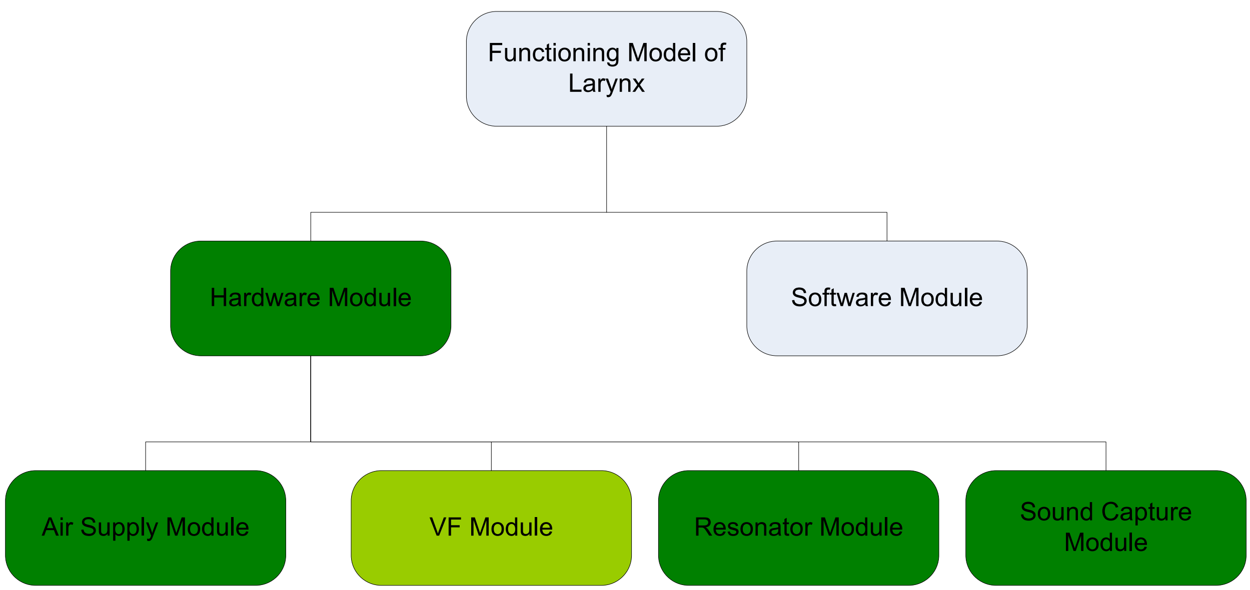
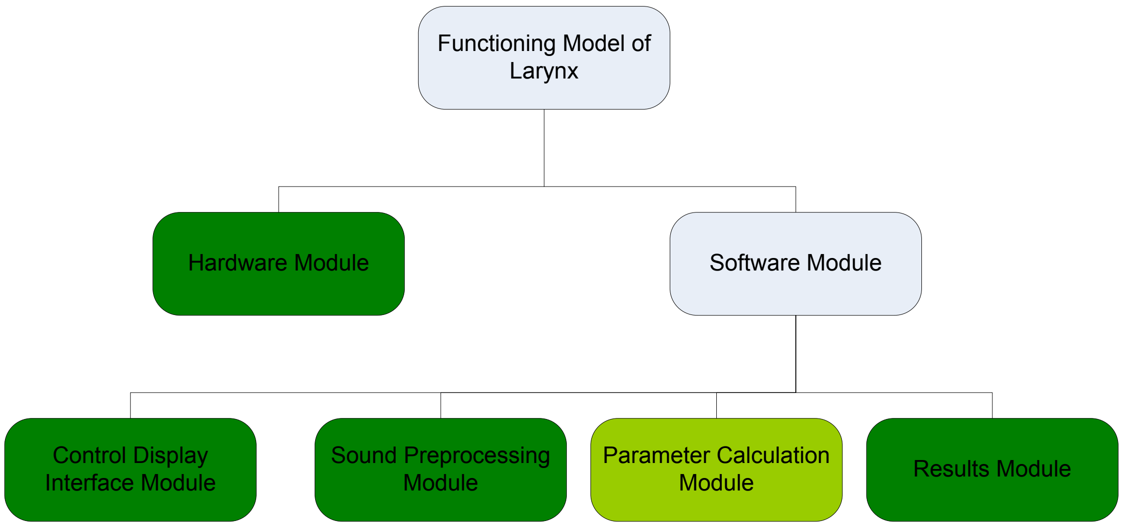


Figure 35: Work Completed in Hardware


Functioning Model of Larynx
Hardware Module
Software Module
Control Display Interface Module
Sound Preprocessing Module
Parameter Calculation Module
Results Module

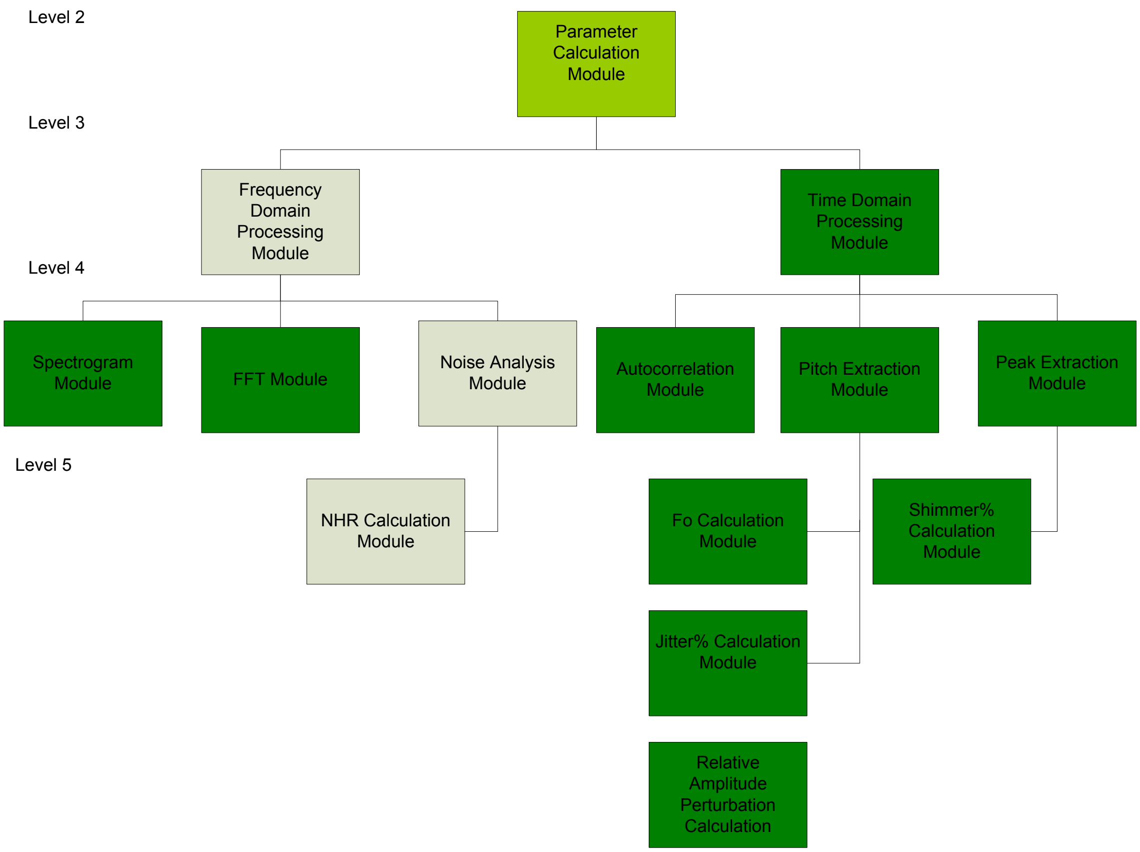


Figure 36: Work Completed in Software

Parts List

Part Description	Quantity	Manufacturer's name	Manufacturer Part No.	Suppliers Name	Supplier's Catalog No.	Unit Cost ($)
Thermoplastic Rubber-Cl 30	3lbs	Septon		GLS Corporation	CLS-30	$ 3.50
1.25" PVC pipe	1 ft	McMaster Carr	8749K44	McMaster Carr	8749K44	$ 9.83
1.25" (ID)Gum Rubber Tubing	1 ft	McMaster Carr	5546K53	McMaster Carr	5546K53	$ 5.54
					Total	$ 18.87

Table 6: Mechanical Parts List

Product Tests Review

Scopes of Tests

Modules which have been implemented and subjected to testing include:

- Vocal Fold Module
    - F0 calculation
    - Jitter % calculation
    - Shimmer % calculation and
    - Relative Amplitude Perturbation % calculation
- Time Domain Analysis Module
    - Pitch Detection Module
        - Windowing
        - Signum Coding
        - Peak Detection
        - Pitch Verification Module
    - Extraction Module
        - Cycle to cycle pitch extraction
        - Cycle to cycle peak extraction
        - F_0 calculation
        - Jitter % calculation
        - Shimmer % calculation
        - Relative Amplitude Perturbation calculation
    - Spectrogram Module
    - FFT/Power Frequency Spectrum Module

Testing Strategy

Vocal Fold Module

For testing of the design of the VF, the first vocal fold model was produced by molding of silicone rubber. The mold was developed using ABS material via rapid prototyping. This was a much simpler approach for testing in comparison to having machined the mold out of Teflon and heating the thermoplastic rubber to 400°F to create the vocal model. The silicon rubber model served as a very good testing tool.

Connected to the apex of the model vocal fold is a resonator (Plexiglass tube). The tube is connected to a four-way PVC connector/fitting which allows for simultaneous video and sound recording of sound produced by the vocal fold oscillation. The fourth opening serves as an extension to the resonator will also be attached to the connector.

To test the silicone rubber vocal fold model, regulated compressed air was supplied to the inferior end of the model and vibration of the vocal folds was observed visually. Three techniques were used to measure the fundamental frequency of the model VF vibration. The first method involved the use of a stroboscope. With the stroboscope, light of different frequencies is emitted onto the vibrating vocal fold model causing the vibration of the model to appear to be going very fast or very slow or motionless. When the model appears to slow down or vibrate in slow motion, this means that the frequency of the light being emitted from the stroboscope is synchronized with the oscillation of the model vocal folds. In this slowed down vibration one can actually see the opening and closing of the vocal fold vibration. When this occurs the frequency of the light is some multiple of the vocal fold vibration. So for example, let's say the stroboscope flashes light at 0°, 90°, 180° and 270° and the vocal fold is open at 90° and 270° and close at 0° and 180°. If the frequency of the light is equal to the fundamental, one should expect to see the opening and closing of the vocal fold oscillation because the light is on/off at the correct points. If it were twice the frequency, expect to see more of the vocal fold oscillation as it progresses from open to close and vice versa. If the frequency were half the fundamental, the observer will only see the closing of vocal folds. Thus a frequency sweep was performed using the stroboscope to find all frequencies at which the model vocal folds appeared to be slowing down. The data from this experiment was then tabulated in EXCEL to find the fundamental frequency which was determined to be the most lowest common multiple of all collected frequencies. Videos of the vibration were taking with a camera phone.

While the stroboscopic data was being collected, a dynamic microphone was used to record the acoustic signal of the vocal fold model. The recorded signal used as an input to an oscilloscope, which produced the waveform of the produced sound. The periods of the signal were then measured to give an estimate of the fundamental frequency. The third method used to measure the fundamental frequency of the vocal fold model was through the software algorithm discussed in this review. A one second segment of the produced sound was selected for analysis. The signal was used as an input to the Time Domain Analysis module (which is part of the Parameter Calculation Module). In addition to calculation of the fundamental frequency, jitter %, shimmer % and relative amplitude perturbation % were also measured.

Time Domain Processing Module

The software modules were implemented and tested in MATLAB. Many of the functions used in these modules are in-built functions in MATLAB. Normally distributed (or Gaussian) random noise was generated in EXCEL to serve as input data for testing the modules. This was achieve through a moving average of 10 randomly generated numbers which were then normalized by dividing the product of the standard deviation and square root of the number of variables. Equation 1Equation 4 is the central limit theorem, which describes how the simulated noise data was generated. Plotting the histogram or the probability density function of the random noise signal should produce a bell-shaped curve. The data for this can be found in the Appendix A.5 page **Error! Bookmark not defined.**. The reason why the random data is normalized is because EXCEL's inbuilt random function generates uniformly distributed random numbers and this is unrealistic since nothing in nature is completely uniform, especially the human voice. Therefore, normally distributed random data was generated to give a more realistic touch to the simulated data, allowing for more variations in the signal.

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}},$$

$Where\ n = number\ of\ variables, S_n = sum\ of\ n\ random\ variables, \mu = mean,$
$\sigma = standard\ devition, and\ Z_n = random\ variable = n_1 = n_2$

Equation 4: Central Limit Theorem

$$y = \left(1 + \alpha \cdot n_1(t)\right) \cdot \left(\sin\left(2 \cdot \pi\left(F_0 \cdot t + \beta \cdot n_2(t)\right)\right)\right)$$

Equation 5: Simulated Signal (Amplitude & Phase Noise)

$Where\ F_0 = fundamental\ frequency, t = time, \alpha = amplitude\ noise\ coefficient,$
$\beta = phase\ noise\ coefficient, y = input\ random\ signal$

By changing the values of α and β, the random noise signal allows for addition of amplitude and/or phase noise. For example if α and β are both zero, the result of the equation is a sine wave.

Pitch Detection Module

Pitch Estimation Sub module

According to the algorithm provided by MDVP, pitch detection is achieved by autocorrelation of the input signal (Appendix A.1 Pitch Extraction Algorithm). First the input signal is multiplied by a window function. Within the window, the signal is coded to ones, minus ones, and zeros by a Signum-like function based on a positive and negative threshold value. Note: The algorithm in the appendix says to also apply a low pass filter at 1800 Hz; this has not been incorporated in the current algorithm.

The window function was tested by multiplying the function by a vector of ones. It is expected that the output signal should be the window function. The window moves in increments of the window size throughout the entire signal. Testing results can be found in the Appendix A.5 on page **Error! Bookmark not defined.**. The signum-like function was developed by the author of this report and tested with a sine wave of randomly generated data from the equations above. The testing results can also be found in the data appendix page **Error! Bookmark not defined.**.

Once the signal has been coded autocorrelation is performed. The software proceeds to estimate the pitch. The pitch or fundamental frequency of the signal is considered to be the inverse of the time where the first peak in the autocorrelation function occurs. An algorithm for finding the peaks was obtained[10] and used to find all the peaks and determine where the first peak occurred. It was subjected to testing as well using the simulated noise data. The values of α and β were changed and subsequently the input variable for determining the difference in value between adjacent peaks was changed as well. Data for this testing can be found in the data appendix page **Error! Bookmark not defined.**.

Since the pitch estimation is based on autocorrelation of the input signal, and because the autocorrelation function can be computed using MATLAB's inbuilt function *xcorr* no focused testing of the autocorrelation function was performed. The results of the pitch estimates speak to the efficacy of the implemented autocorrelation function. It is important to note that in this sub module, the pitch was estimated within each window. In the next sub module, Pitch Verification, the results from the pitch estimate in this sub module is used to yield a better overall pitch estimate.

---

[10]MATLAB Peak finding algorithm http://www.billauer.co.il/peakdet.html

In this sub module, the pitch was estimated twice in each window for a given signal. The first time it is performed, the signal has been signum coded at a high threshold value (78% of the maximum peak in a particular window). The second pitch estimate is done on a signal that has been signum coded with a lower hold (45% of the maximum peak in a particular window). The purpose of the high and low threshold is to suppress the effects of the harmonic and sub harmonic components of the signal under analysis.

The combined modules that make up the Pitch Estimation Sub module were tested for accuracy and efficiency. To test the entire sub-module, the random signal data was used. The values of α and β were changed to simulate amplitude and phase variations respectively. Each variation was repeated 100 times and each time the pitch was estimated. A histogram of each was then generated using MATLAB. This was repeated for windows of different sizes (25ms, 50ms and 100ms) to see which window size would result to a better pitch estimator.

Pitch Verification Sub module

This entire sub module was tested using simulated noise data and real human and model vocal fold data. Similar to the previous sub module, histograms should be generated showing the verified pitch.

Extraction Module

The extraction module was first tested independently of previously discussed software modules. During testing the calculation of fundamental frequency and percents jitter, shimmer and RAP were also conducted simultaneously. Simulated noise data and real data were used in testing. In order to test this module independently the fundamental frequency of the input signal has to be known because the extraction of cycle to cycle pitch and peak data depends on first knowing an approximate fundamental frequency.

The real human data used to test this module was previously analyzed using MDVP thus the parameters of interest had been calculated. Thus the testing results from this module were compared to the results of MDVP. The errors between the results were tabulated in each case. Values for α and β were changed in the simulated data, and the parameters were calculated. Graphs were plotted to show correlation of the calculated parameters with values of α and β. In addition, histograms of the calculated parameters from simulated data were produced for different values of α and β and different frequencies. Histograms for six different frequencies were also produced to show the accuracy of the fundamental frequency calculation.

Another test performed to determine the appropriate number of points for interpolation of the extracted cycle to cycle pitch data. The number of points was increased from 3 points to 6 points and the point that yielded the best results (or least error) in comparison to results from MDVP was chosen.

Frequency Domain Processing Module

The testing of this module is not really necessary because the spectrogram plot is an inbuilt MATLAB function including the FFT. However, both modules were implemented in testing of the vocal fold model.

Testing Results and Assessments

Vocal Fold module

a. ***Progress/success for the test***

As discussed, three modes of fundamental frequency calculations were performed on the acoustic signal of the vocal fold model. The results from all three methods are conclusive and amenable. The fundamental frequency of the silicone rubber vocal folds was measured to be around 152-158 Hz.

The table below shows how the fundamental frequency was calculated from the data obtained from the stroboscopic experiment. The left most column represents the raw data which was read from the stroboscope. The raw data was converted to Hertz by dividing by 60. The yellow highlights the fundamental frequency, which appears the most common in the first five multipliers. The average fundamental frequency is 152 Hz. The cells in the green represent the second harmonic (average 304 Hz) and the blue represents the third harmonic (average 456 Hz).

strobe freq											
Revolutions per min	**per sec**	**ratio to fundamental**	**multiplier**								
(raw data)	**1.0**		**2.0**	**3.0**	**4.0**	**5.0**	**6.0**	**7.0**	**8.0**	**9.0**	**10.0**
1820	30.3	0.198	60.7	91.0	121.3	151.7	182.0	212.3	242.7	273.0	303.3
2280	38.0	0.248	76.0	114.0	152.0	190.0	228.0	266.0	304.0	342.0	380.0
3120	52.0	0.339	104.0	156.0	208.0	260.0	312.0	364.0	416.0	468.0	520.0
4400	73.3	0.478	146.7	220.0	293.3	366.7	440.0	513.3	586.7	660.0	733.3
9200	153.3	1.000	306.7	460.0	613.3	766.7	920.0	1073.3	1226.7	1380.0	1533.3
18300	305.0	1.989	610.0	915.0	1220.0	1525.0	1830.0	2135.0	2440.0	2745.0	3050.0

Table 7: Model Vocal Fold Fundamental Frequency Calculation from Stroboscopic Data

The second measurement was performed on an analog oscilloscope. Sound from the vocal fold model was recorded directly to an IPOD recorder through a dynamic microphone. By counting the peaks of the signal and time divisions, the period of the signal was determined. The measured frequency was 769Hz, which happens to very close to the 5th harmonic as seen in the table above. Because this value is much higher than the data measured from the stroboscope data, the likely possible explanation is the effect of resonance in a tube. The recording of the signal had taken place inside a tube structure (4-way connector), thus allowing sound amplification/resonance to occur. To confirm this conclusion, the equation for resonance in an open tube was used.

$$f_r = n * \frac{v}{2 * L}$$

Equation 6: Resonance in an Open Tube

Where n=is the integer multiple representing the resonant mode, v= speed of sound=343 m/s, L=length of tube=17(+/-0.5) inches, and f_r= resonant frequency=769 Hz. The values of v and L are known and f_r is a multiple (n) of the fundamental frequency which has been estimated to be 152 Hz. Plugging the numbers in and solving for n results in an n of about 5.

Using the software developed in MATLAB as part of this project, a third analysis was performed on the sound produced from the model vocal folds. Ten recordings were made and analyzed in three ways. The major difference between the recordings is the volume (or loudness) at which the sound was recorded. Only data for four of the ten recordings will be presented in this section. The data remaining can be found in the appendix of this report (**Error! Reference source not found.**).

The tables in the four sections below are measurements taken from the developed algorithm to detect pitch and calculate jitter %, shimmer %, and RAP %. For seven of the ten recordings, a fundamental frequency of approximately 158 Hz was measured via the algorithm discussed in this report. There were two seemingly anomalous results one of which is mentioned below (High Volume). The algorithm measured the highest fundamental frequency to be 44100Hz which is the sampling rate of the recorded signal. A correction was implemented in the algorithm to correct the error. The harmonics of the vocal fold vibration, especially the third harmonic which is approx. 467 Hz, are very strong. In addition to the analysis of the acoustic signal with the algorithm of the Time Domain Analysis Module, some Frequency Domain Analysis was performed including the power spectrum and the spectrogram. It can be seen from these analyses that the third harmonic is the most dominant. The fundamental frequency measured from the power spectrum ranges from 154-157 Hz.

The figure below is a color bar representing the intensities of a spectrogram. In the spectrogram, the intensity of the frequencies present in the sound is plotted over time. The louder the sound the greater the intensity of a particular frequency, this is why in the spectrogram of the very high volume, there is more red and orange at higher frequencies. In all the spectrograms, the most intense region (red) is located around 467 Hz, which is the dominant frequency. This effect is also seen in the power spectrum.

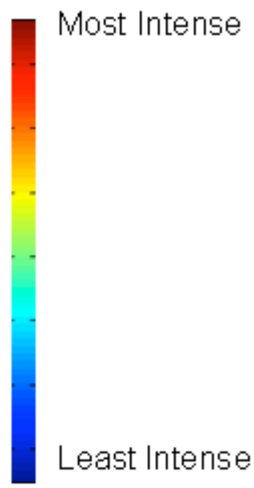


Figure 37: Spectrogram Color Scale

In summary, all the methods used to test the model vocal folds proved to have very similar results. It remains now to test the sound produced from the model with the MDVP software, which is the gold standard software for voice analysis. Although it is still preliminary to judge whether the measured parameters are too low or too high as the software's parameter calculation module has not been calibrated to fit the standard normative data, the developed software was able to measure percentages of RAP, Jitter and Shimmer. Nonetheless, the pitch estimates are accurate and show that the model has very strong harmonic components.

<u>Very High Volume</u>

Loudness	Max Amplitude	F0 (Hz)	Highest F0 (Hz)	Lowest F0 (Hz)	Jitter %	Shimmer %	RAP %		File Name
Very Hi	0.96	169.04	428.16	94.03	11.57	1.77	2.35	No Correction	20090318 220355 Very High.wav
		154.35	156.38	152.60	0.18	1.77	0.03	correction	

Table 8: Very High Volume VF Model —Calculation Results from Software Module

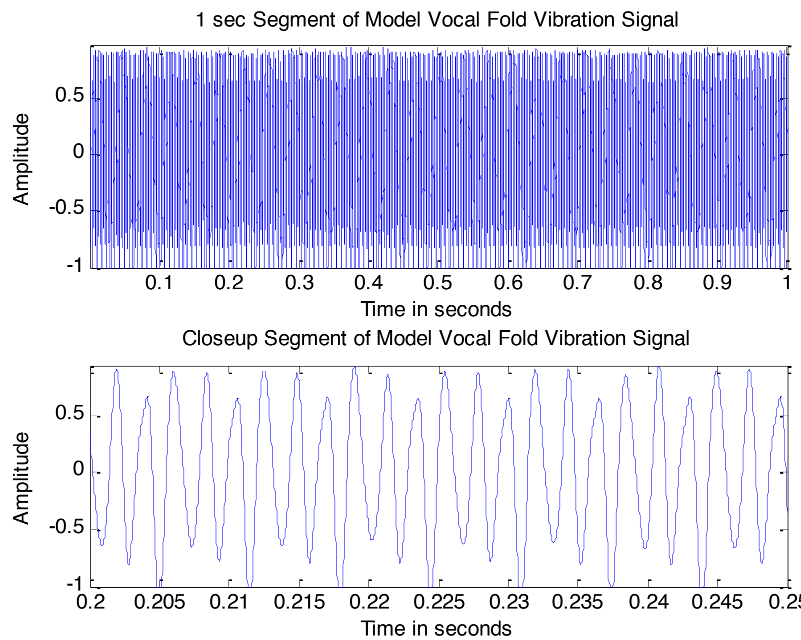


Figure 38: Very High Volume Model VF Plot & Close-up

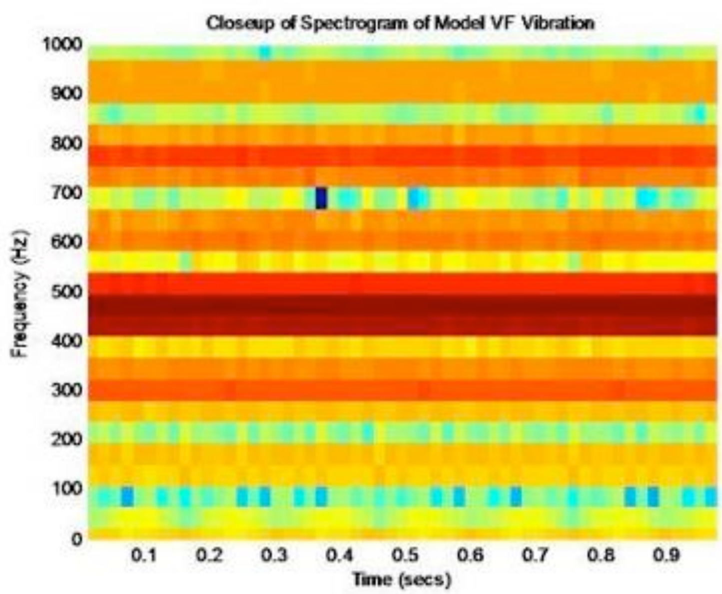


Figure 39: Spectrogram of Very High Volume Model VF

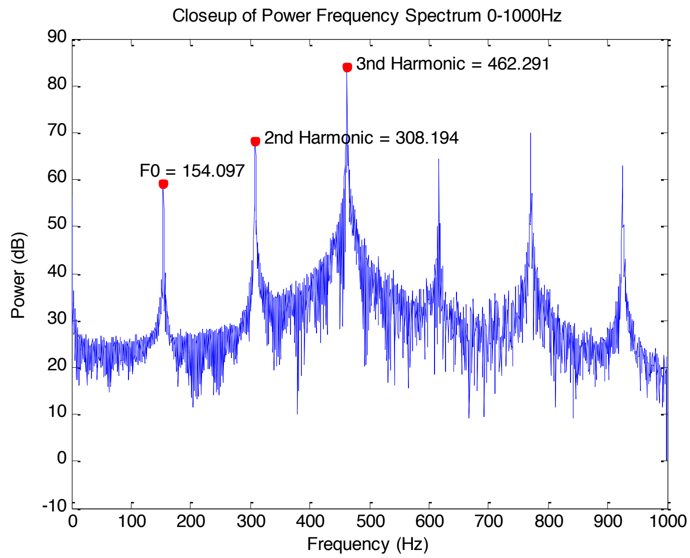


Figure 40: Frequency Power Spectrum of Very High Volume Model VF

High Volume

Loudness	Max Amplitude	F0 (Hz)	Highest F0 (Hz)	Lowest F0 (Hz)	Jitter %	Shimmer %	RAP %		File Name
High	0.60	818.33	44100.00	235.83	10.09	12.44	1.77	No Correction	20090318 221925 High.wav
		467.04	501.14	424.04	2.27	12.44	0.42	correction	

Table 9: High Volume Model VF—Calculation Results from Software Module

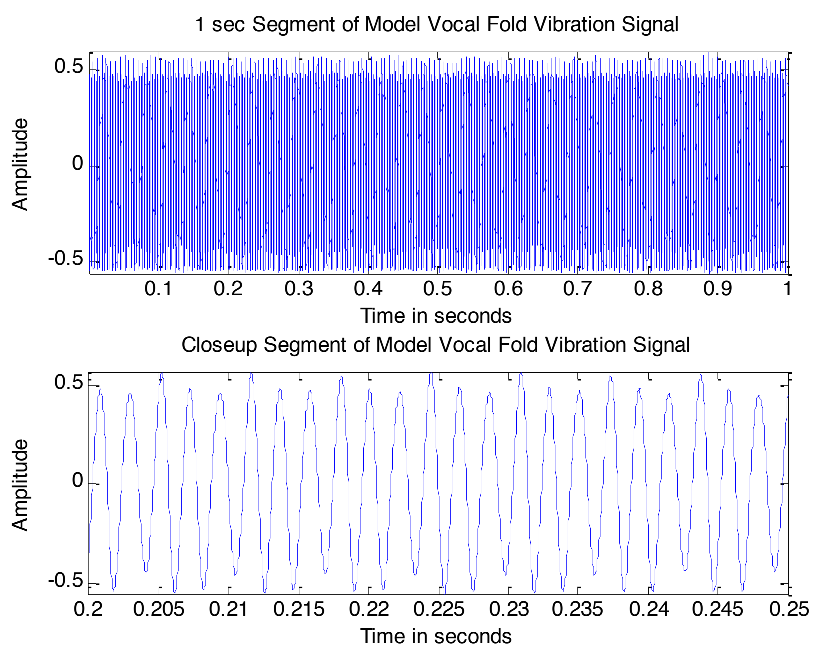


Figure 41: High Volume Model VF Plots

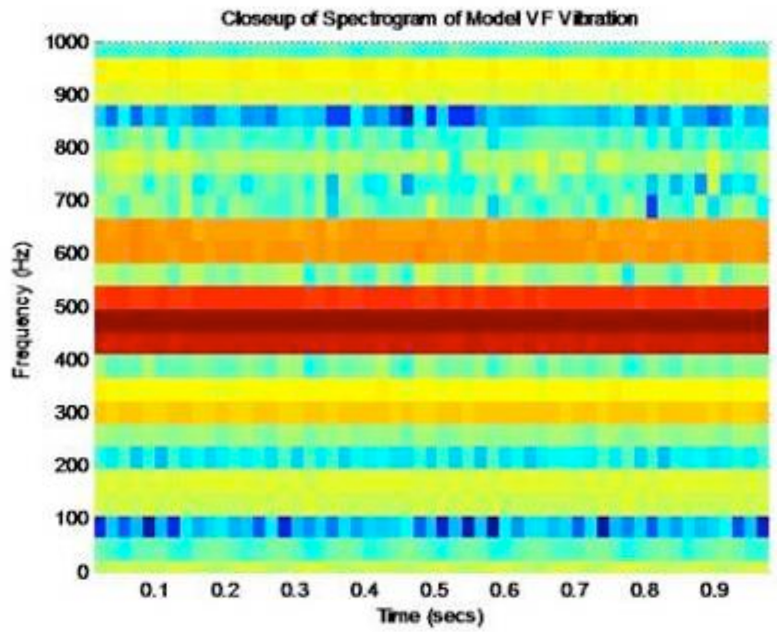


Figure 42: Spectrogram of High Volume VF Model

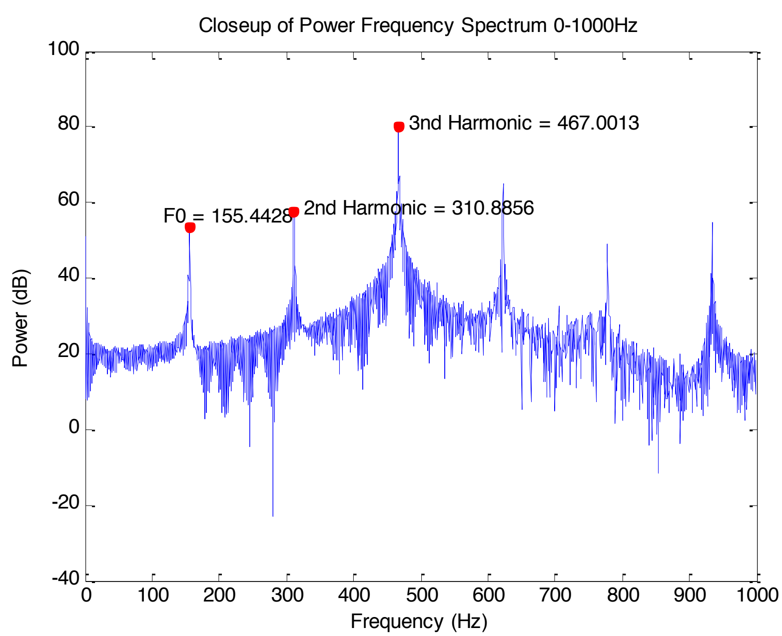


Figure 43: Frequency Power Spectrum of High Volume VF Model

Medium Volume

Loudness	Max Amplitude	F0 (Hz)	Highest F0 (Hz)	Lowest F0 (Hz)	Jitter %	Shimmer %	RAP %	File Name

Medium	0.36	157.50	158.06	156.94	0.09	1.44	0.02	No Correction	20090318 220439 Medium.wav
		157.50	158.06	156.94	0.09	1.44	0.02	correction	

Table 10: Medium Volume—Calculation Results from Software

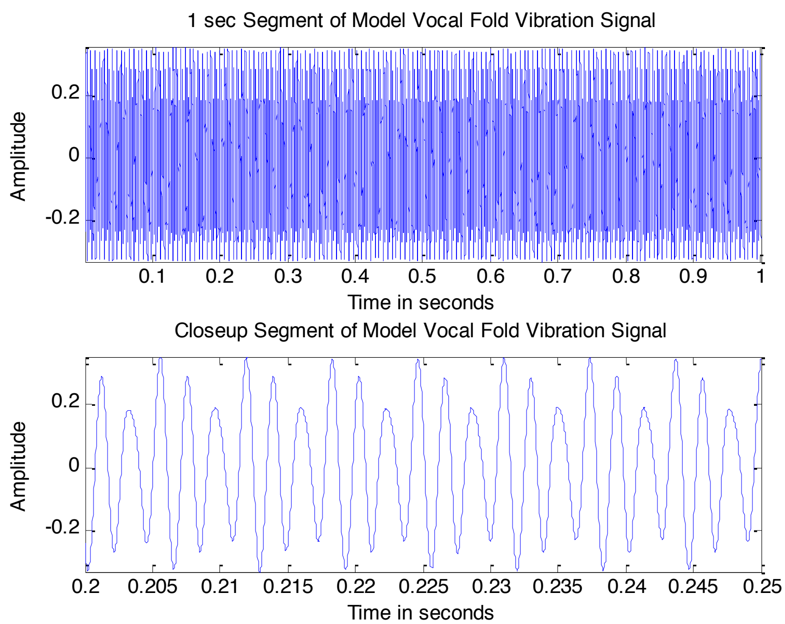


Figure 44: Low Volume Model VF Plots

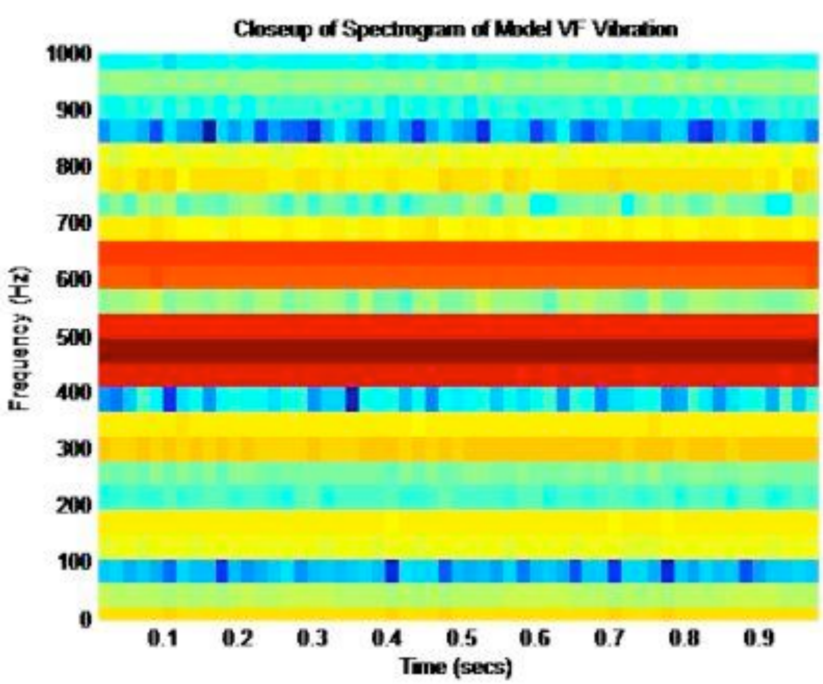


Figure 45: Spectrogram of Medium Volume Model VF

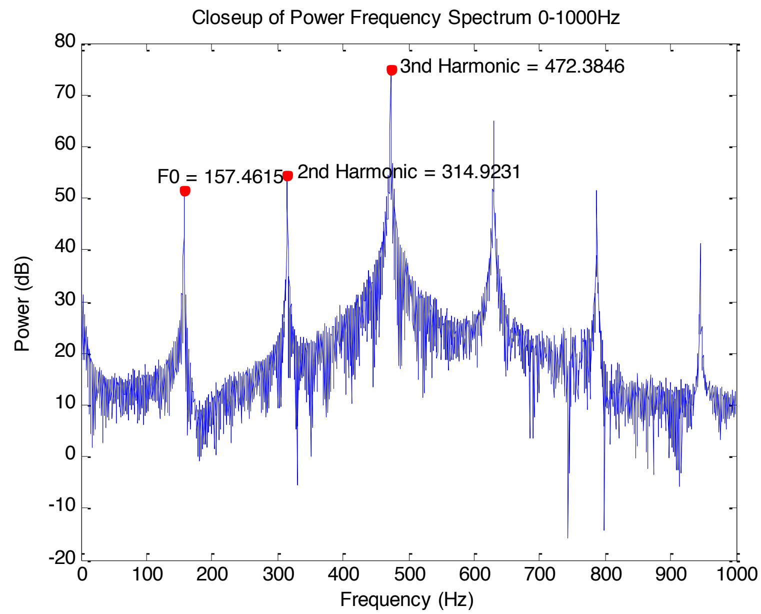


Figure 46: Frequency Power Spectrum of Medium Volume VF Model

Low Volume

Low Volume/20090318 222327 Low.wav

	Max Amplitude	F0 (Hz)	Highest F0 (Hz)	Lowest F0 (Hz)	Jitter %	Shimmer %	RAP %		File Name
Low	0.11	157.33	159.78	155.28	0.21	1.57	0.04	No Correction	20090318 222327 Low.wav
		157.33	159.78	155.28	0.21	1.57	0.04	correction	

Table 11:Low Volume—Calculation Results from Software module

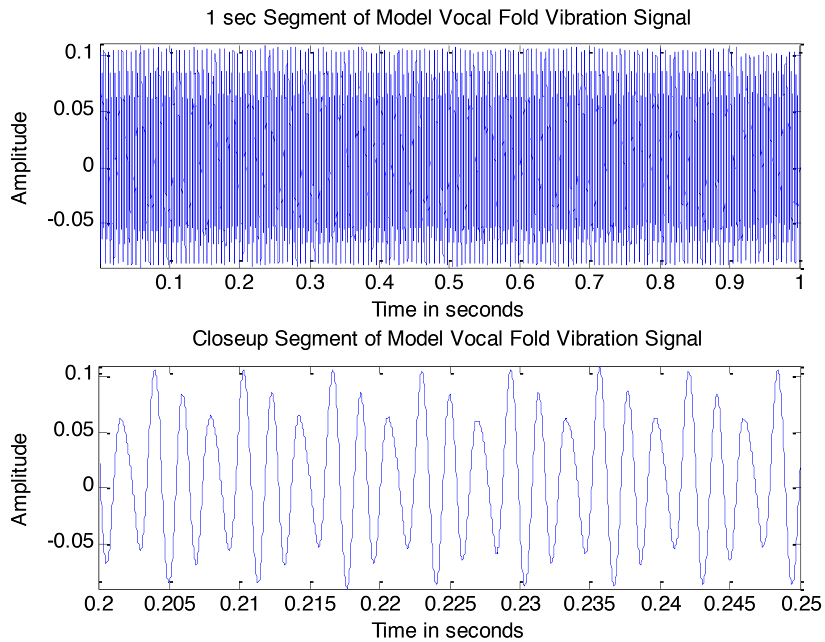


Figure 47: Low Volume VF Module Plots

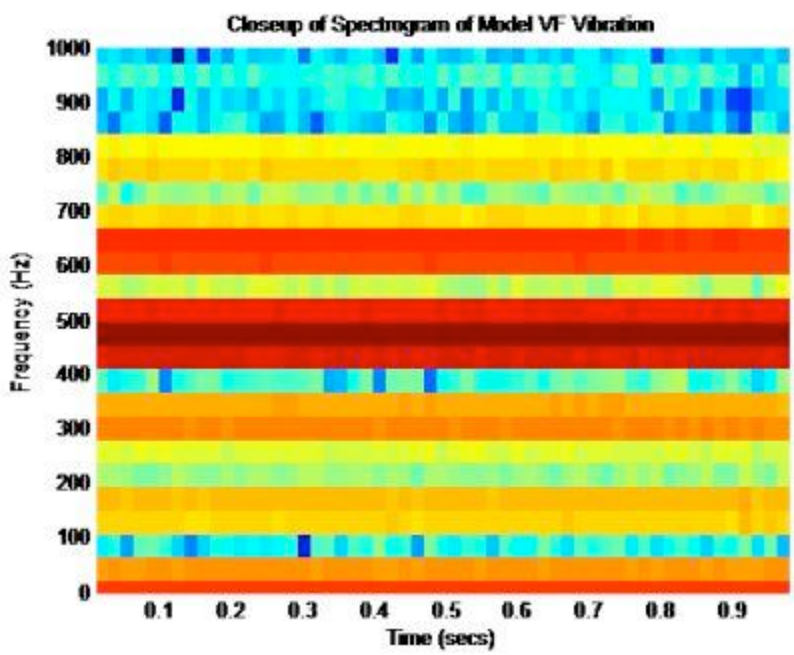


Figure 48: Spectrogram VF Model Low Volume

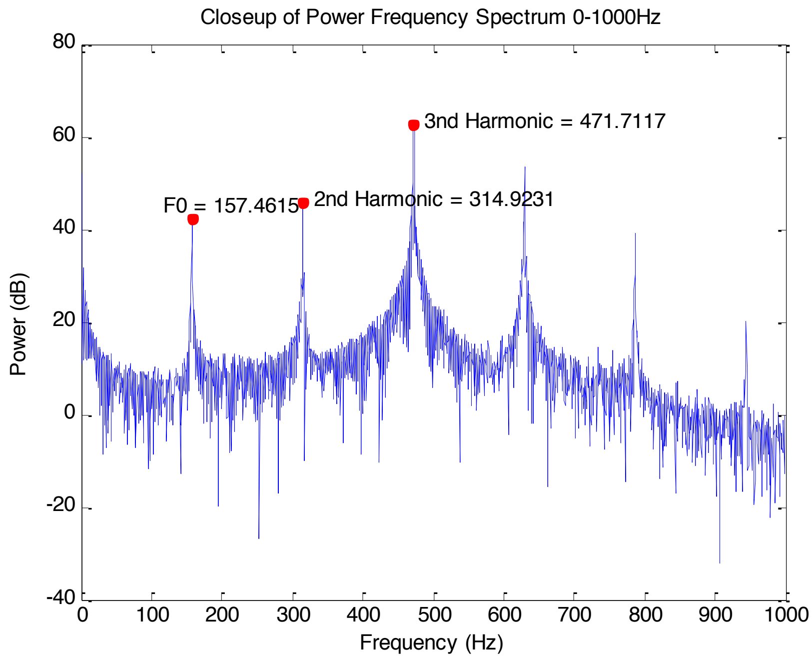


Figure 49: Frequency Power Spectrum Low Volume VF Model

Very Low Volume

This particular recording was taken simultaneously with during the stroboscopic measurement.

Loudness	Max Amplitude	F0 (Hz)	Highest F0 (Hz)	Lowest F0 (Hz)	Jitter %	Shimmer %	RAP %		File Name
Very Low**	0.03	157.82	162.13	153.13	0.53	2.38	0.11	No Correction	20090318 223304 Low.wav
		157.82	162.13	153.13	0.53	2.38	0.11	correction	

Table 12: Very Low Volume VF Model-Calculation Results from Software Module

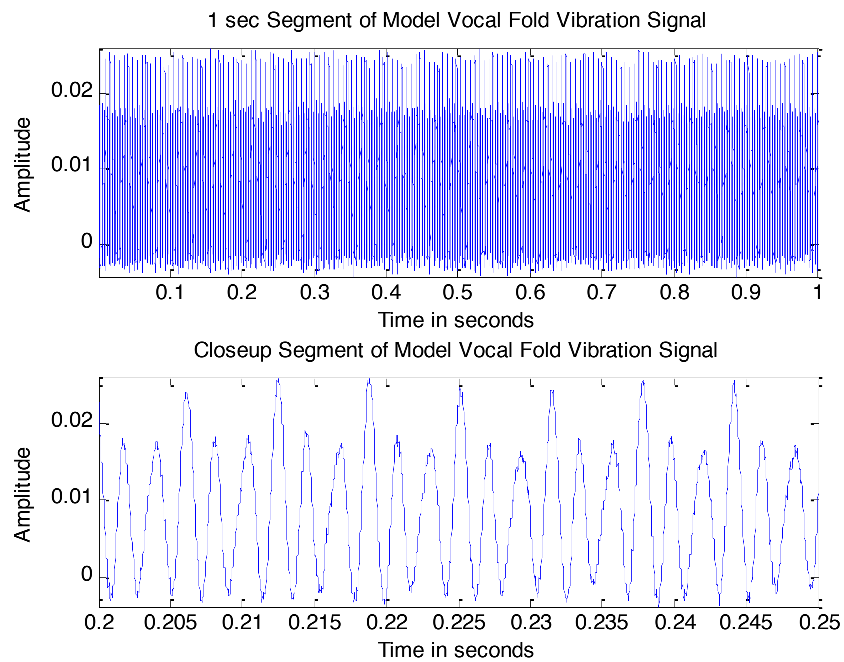


Figure 50: Very Low Volume Model VF Plots

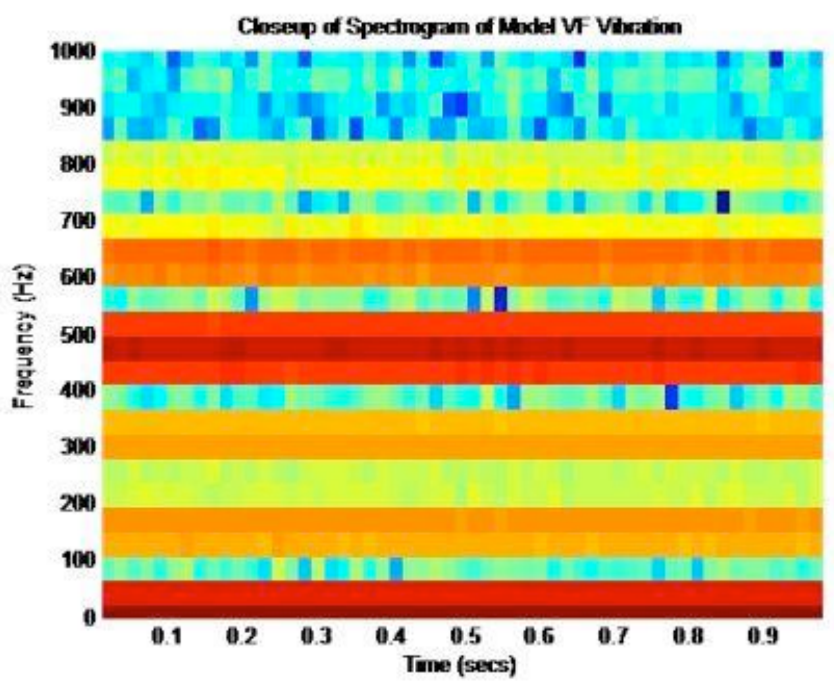


Figure 51: Spectrogram of Very Low Volume Model VF

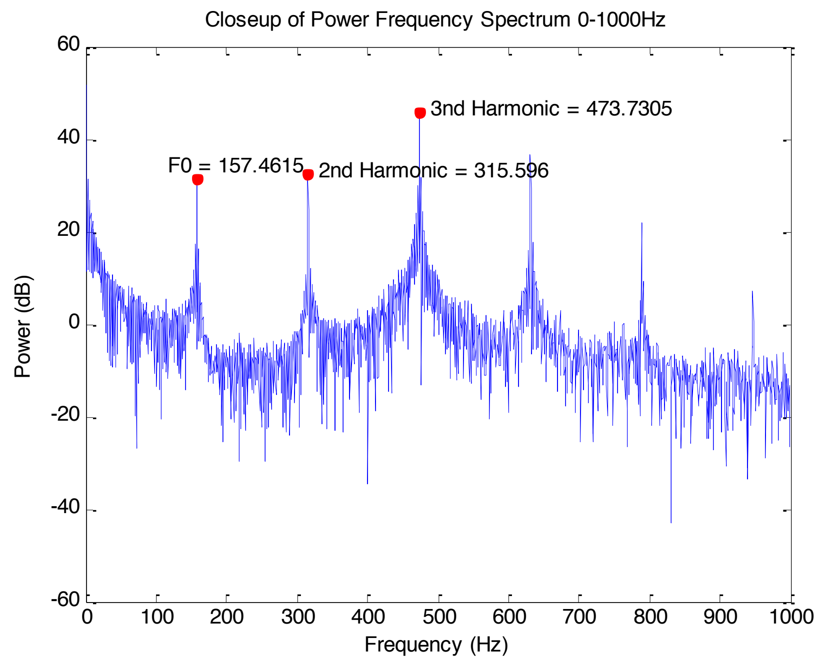


Figure 52; Frequency Power Spectrum--Very Low Volume VF Model

b. ***Problems encountered***
- Poor exposure of video image
- Slightly difficult to see vibration of vocal fold from the top

c. ***Problems solved/Solutions yet to be attempted***
- Using video camera with manual exposure setting
- Redesign of vocal mold to add curvature on the top surface of the vocal folds.

d. ***Testing work for next period***
- Use MDVP to test acoustic signal of vocal fold model and compare
- Design and implementation of different model vocal folds for simulation of vocal disorders and observe the associated effects.
- Implementation of vocal fold manipulation stage

Summary & Conclusions

PTR to FPR

Although the entire project is not complete all the essential modules have been implemented and tested and all are functioning as expected. Therefore it is safe to conclude that a solid foundation is in place for future enhancement of this project. It is as complete as it can be for the allocated time frame and size of the project. The testing results confirm that the Time Domain Processing Module, which is the core of the software (amounting to about 65%), is fully functional. The fundamental frequency calculation is very accurate and comparable to the gold standard software, MDVP. As for the calculated parameters: jitter, shimmer and RAP, it is uncertain at the moment as to the accuracy in measurements. They each respond to additive phase and amplitude noise thus validating their function as a voice perturbation measuring tool. A next step towards enhancement of the calculations would be calibration to fit the standard norms. Partial success was also achieved in the implementation of the model vocal folds. The vocal fold produced oscillations similar to that of human vocal folds. The fundamental frequency was calculated to be about 156 Hz, which is a high male, low female voice quality. It was originally anticipated that the vocal fold vibration frequency be much lower. More testing of can be done to confirm if the vibration produced is androgynous. Future work on the vocal fold model include: simulation of different voice disorders-polyps, nodules and cysts-and manipulation of model vocal folds to open and close and change the pitch of the produced sound. With a model in place this would be much easier to implement.

IPR to PTR

Much progress has been made in implementation of major components of the project. The goals for this review were to have one or two VF models produced, implementation of the design of the two major software modules (autocorrelation/pitch detection and the noise analysis module), and testing of major software modules. Thus far, I have been able to accomplish my goals. The noise analysis module is currently under implementation, the algorithm is a little tricky. The pitch estimation algorithm has been implemented and tested as seen in the earlier section. No work was done this time for the electrical module because it is ranked low priority. However, work will start on the electrical module once the noise analysis module has been completed. Aside from the major software modules, the vocal fold module of highest priority, I plan to carry out the experiment outlined above and hope to eventually simulate vocal disorders using the VF model.

FDR to IPR

Much stride has been made in all areas of this project. The software design can use more improvement to make it more specific for laymen (or non mathematical programmer). Mechanical drawings have been made for the molds and vocal fold manipulators. Mold making should begin shortly since the material for casting has been received. Over the next few week, software coding and mold making will begin in parallel for these are the major modules of this project and will require most time.

PDR to FDR

Since the Preliminary Design Review, there has not been much change in terms of the goals of this project. The ultimate goal of this product is to produce a functioning model of the larynx that will have active vocal folds that respond to air flow. The vocal folds will be used to simulate common vocal disorders that are characterized by addition of mass to the vocal folds. Recent progress has been made in determining the correct material to use to make the model vocal folds. This is a major accomplishment because the elastic properties of the vocal folds are very difficult to emulate. However there is much needed work to be done in terms of integrating the vocal fold and larynx components together and manipulating the vocal folds. Most of the software component of this project was slightly overlooked because most efforts went into finding the right material for the vocal folds. Now one of the major components of the hardware module has been resolved and some functions of the software module are more understood, the software component will be fully ironed out in time for the Final Design Review.

Qualifications of Key Personnel

The project involves several disciplines, including but not limited to materials engineering, digital signal processing, circuitry, speech pathology and software programming. As a biomedical engineering major, I have taken Introduction to Biomaterials, C and C++ programming, 2 courses in Digital Signal Processing, University Physics, and Circuit theory. I have 2 years of research experience where I am trying to develop a resistive strain gauge with gold nanoparticles that could be used for many sensor-type applications. I have been allowed to work autonomously on this project, under the mentorship of Dr. Mark Reeves and Dr. Jason Zara. I believe I am qualified to take on the task of this project because of my academic background and my research experience. My resume and transcript are included below.

Resume

Profile

- Seeking entry into PhD program to pursue career as a research scientist
- Interested in integration of biology with electronics
- Looking for hands on technical experience in engineering and related fields

Education

B.S, Biomedical Engineering Expected May 2009
The George Washington University (GW)

Research Accomplishments

Research Assistant, Research Interest Group (RIG), Emergency Medicine Division**,** University of Medicine & Dentistry of NJ Newark NJ **Summer 08**

- Worked on three individual projects under the tutelage of Dr. Sandra Scott
- Reviewed charts completed by attending physicians and residents as part of a quality improvement project
- Performed data analysis on a study to investigate the source of the Emergency Department's high patient walkout rate
- Collected data for ongoing palliative care study
- Shadowed Dr. Scott and residents in the Emergency Department

Undergraduate Research Assistant, GW Electrical and Computer Engineering (ECE) & Physics Departments ***Summer 07-Present***

- Continuing independent research project, under the guidance of Drs. Jason Zara and Mark Reeves. The goal of this project is to develop a strain gauge using gold nanoparticle wires and polyimide Kapton® films.

Walter Coulter Summer Fellow, GW ECE Departments ***Summer 07***

- Participated in ongoing research under the mentorship of Prof. Zara to develop imaging probes for endoscopic optical coherence tomography as part of a grant received from the Walter Coulter Foundation
- Tested and characterized different mirror devices for probes

Gamow Fellow, GW ECE & Physics Departments ***Sept06-May07***

- Conducted independent research, under the mentorship of Profs Zara and Reeves, to develop a potential chemical sensing device via integration of self-assembled gold nanoparticles wires onto Polyimide (PI or Kapton®) films with a bimorph actuator.
- Used techniques such as argon plasma etching to treat the surface of the PI films; and Vertical Colloidal Deposition to make the gold nanoparticle wires
- Presented research findings to University professors, faculty and students at the 5th Annual George Gamow Research Fellowship Symposium

Howard Hughes Scholar, GW Physics Department ***Summer06***

- Worked on an independent project, under the mentorship of Professor Reeves, to develop a technique for electrophoresis using a gold nanoparticle wires in place of gels.
- Used fluorescent microscopy to observe the diffusion of fluorescently tagged proteins across the nanoparticle wire in the absence of an electric field.
- Presented findings to University professors and students participating in the program

Relevant Skills

- Relevant courses include: Biology (I&II); General Chemistry (I&II); University Physics (I&II), Circuit Theory (with Lab); Introduction to C Programming; C++; Biophysics (Macro and Micro); Engineering Electronics; Biomedical Properties Lab; Digital Signal Processing and Three Seminars in Biomedical Engineering.
- Relevant Projects include: Successfully designed and built working music audio Amplifier using operational amplifiers and transistors, and a DC Power Supply; Many uses of C & C++ programs; 3 undergraduate research projects

Activities

- President, National Society of Black Engineers, GW Chapter May07-April08
- Departmental Tutor, GW Electrical and Computer Eng. Dept Feb08-Present
- Tutor, Catholic Charities June06-Present
- Tutor, GW Society of Physics Students Sept07-Present
- Mentor, GW School of Engineering and Applied Sciences Aug06-Present

Transcript

Intellectual Contribution

I will be building most of the components for this project. I will use a rapid prototyper to make the model of the larynx. The software for analyzing the signal produced by the model will be custom written. The vocal fold material will also be custom made and different parts may be purchased to get the right material. The electric motors for operating the movement of vocal fold will be purchased commercially or found in the Tompkins Machine Shop.

References

1. Siekel. Anatomy & Physiology for Speech, Language and Hearing Science. NY: Thomson Delmar Learning, 2005.

2. R.Titze, Ingo. Principles of Voice Production. Englewood Cliffs: Prentice-Hall Inc, 1994.

3. Figure 2c. Fukui, Kotaro. "New Anthropomorphic Talking Robot having Sensory Feedback Mechanism and Vocal Cords based on Human Biomechanical Structure." IEEE International Conference on Robotics and Automation 2006 13th ser. 8 (2006): 1095-1100.

4. Fukui, Kotaro. "New Anthropomorphic Talking Robot having a Three-dimensional Articulation Mechanism and Improved Pitch Range." Proceedings of the 2007 IEEE International Conference on Robots and Automations (2007): 2922-927.

5. Kay Elemetrics Corporation. *Multi-Dimensional Voice Program (MDVP) Model 5105: Software Instruction Manual*. Lincoln, NJ: Kay Elemetrics; 2005.

6. Rabiner, Lawrence. "On the User of Autocorrelation Analysis for Pitch Detection." IEEE Transactions on Acoustics, Speech, and Signal Processing Assp-25 (1977): 24-33.

Appendix

A.1 Pitch Extraction Algorithm

Obtained from MDVP Manual

Pitch Extraction

The amplitude and frequency demodulation curves of the voice signal contain information about the time-domain behavior of $a(n)$ and $\varphi(n)$. The period-to-period pitch extraction [10] is the classic type of demodulation used for evaluation of voice pathology [7, 8]. However the irregularity of the disordered voice makes the pitch extraction inaccurate, often impossible.

In order to provide reliable data an adaptive time-domain pitch-synchronous method for pitch extraction was developed. It consists of the following main steps: fundamental frequency (*Fo*) estimation, *Fo* verification, period-to-period *Fo*-extraction and computation of time-domain voice parameters.

The *Fo-estimation* provides preliminary information about the pitch. It is based on short-term autocorrelation analysis with non-linear *sgn*-coding [11] of the voice signal x(n)

$$R(\tau) = \sum_{n=0}^{N-\tau-1} x'(n)x'(n+\tau), \;\; 0 \leq \tau \leq N/2,$$

where: $x'(i)=0$ *if* $Pmin<x(i)<Pmax;$

$x'(i)=1$ if $x(i) \geq Pmax;$

$x'(i)=-1$ if $x(i) \leq Pmin$

and $Pmax=KpAmax;$

$Pmin=KpAmin;$

Amax and *Amin* - global extremes of the current window in the voice signal *x(n)*. The length of the autocorrelation window is 30ms or 10ms depending on the *Fo*-extraction range (67-625Hz or 200-1000Hz). The sampling rate is 50kHz and every window is low-pass filtered at 1800Hz before coding. The value of the coding threshold at this stage of the analysis is $Kp=0.78$ in order to eliminate the incorrect classification of *Fo*-harmonic components as *Fo* [12]. The current window is considered to be voiced with period $To=\tau max$ if the global maximum is $Rmax(\tau max)>KdR(\tau=0)$, where the voiced/ unvoiced threshold value id $Kd=0.27$ [12].

The *Fo-verification* procedure is similar to the *Fo*-estimation. The autocorrelation function is computed again for the same windows at $Kp=0.45$ in order to suppress the influence of components subharmonic to *Fo*. The results are compared to the previous step and the decision about the correct *To* is made for all windows where difference is discovered.

A *period-to-period Fo-extraction* is made on the original signal x(n) using a peak-to-peak extraction measurement. It is synchronous with the verified pitch and voiced/unvoiced results computed in the previous steps. A linear 5-point interpolation is applied on the final period-to-period *Fo*-data in order to increase the resolution. This increased resolution is necessary for meaningful frequency perturbation measurements. The peak-to-peak amplitude is also extracted for every period.

A.2 NHR Algorithm

Obtained from MDVP Manual

NHR

Definition: Noise-to-Harmonic Ratio - Average ratio of the inharmonic spectral energy to the harmonic spectral energy in the frequency range 70-4200 Hz. This is a general evaluation of noise present in the analyzed signal.

Method: NHR is computed using a pitch-synchronous frequency-domain method. In general terms, the algorithm functions as follows:

A. The signal to be processed must be captured using one of the sampling rates in the left column of the table below, and processing must include the **MDVPvoice** command, which identifies individual pitch periods. As a part of the processing, the signal is decimated to a lower sampling rate, as shown in the center column of the table.

Sampling Rate:	Decimated To:	1024-Pt. Block Length
25, 50, 75, 100, 125, 150 175 and 200 KHz	12.5 KHz	81.92 ms
32 KHz	16 KHz	64.00 ms
44.1 KHz	14.7 KHz	69.66 ms
48 KHz	12 KHz	85.30 ms

The decimated signal is divided into a sequence of 1024-point blocks, the duration of which is related to the sampling rate, as shown in the right column of the table. For every data block, the following steps apply:

1. Compute an unwindowed 1024-point Fast Fourier Transform (FFT) for the data and convert to a power spectrum.

2. Calculate the average fundamental frequency within the window synchronously using the pitch extraction results from the **MDVPvoice** command.

3. Separate the spectrum into the harmonic and inharmonic components synchronously with the average fundamental frequency of the current block.

4. Compute the Noise-to-Harmonic Ratio (NHR) of the current data block as the ratio of the inharmonic to the harmonic spectral energy in the frequency range 70-4200 Hz.

B. When all blocks have been processed, combine the NHR ratios for all blocks into the average value, which is reported as the Noise-to-Harmonic Ratio for the data.

A.3-Normative Data

Obtained from MDVP Manual

Female Norms

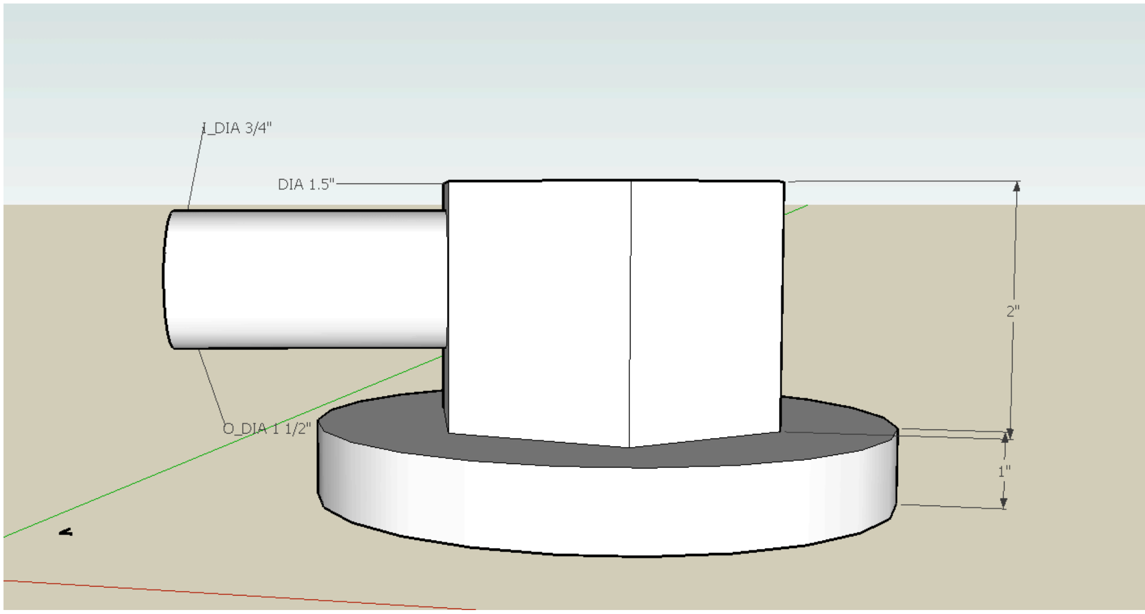
APQ=1.397	FO=243.973	NVB=0.200	SHDB=0.176	SPR2=0.000
ATRI=2.658	FTRI=0.304	PATR=0.000	SHIM=1.997	SPR3=0.000
DSH=0.200	JITA=26.927	PER=713.188	SPI=7.534	STD=2.722
DUV=0.200	JITT=0.633	PFR=2.250	SPLA=0.000	TO=4.148
DVB=0.200	MFO=241.080	PFTR=0.000	SPLH=0.000	TSAM=3.000
FATR=2.375	NHR=0.112	PPQ=0.366	SPLL=0.000	VAM=10.743
FFTR=3.078	NNE=0.000	RAP=0.378	SPLR=0.000	VFO=1.149
FHI=252.724	NSH=0.200	SAPQ=2.371	SPPQ=0.532	VTI=0.046
FLO=234.861	NUV=0.200	SEG=92.594	SPR1=0.000	

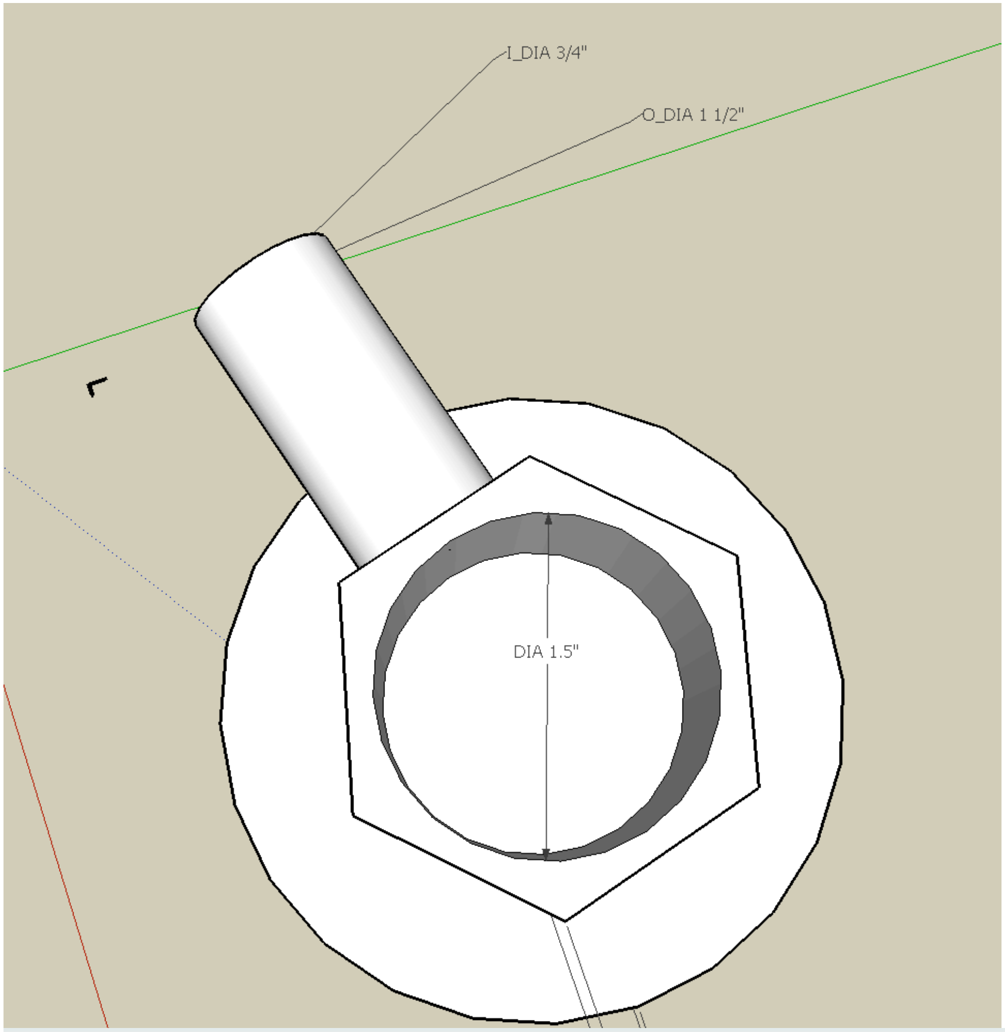
Male Norms

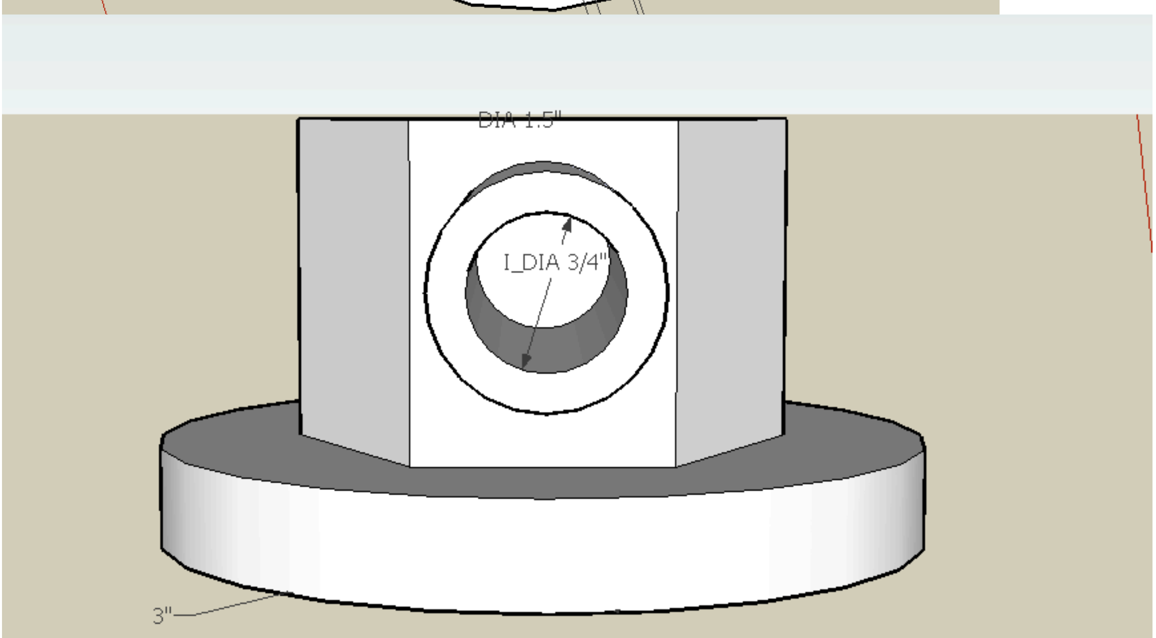
APQ=1.986	FO=145.223	NVB=0.200	SHDB=0.219	SPR2=0.000
ATRI=2.133	FTRI=0.311	PATR=0.000	SHIM=2.523	SPR3=0.000
DSH=0.200	JITA=41.663	PER=433.143	SPI=6.770	STD=1.349
DUV=0.200	JITT=0.589	PFR=2.095	SPLA=0.000	TO=7.055
DVB=0.200	MFO=141.743	PFTR=0.000	SPLH=0.000	TSAM=3.000
FATR=2.728	NHR=0.122	PPQ=0.338	SPLL=0.000	VAM=7.712
FFTR=3.655	NNE=0.000	RAP=0.345	SPLR=0.000	VFO=0.939
FHI=150.080	NSH=0.200	SAPQ=3.055	SPPQ=0.561	VTI=0.052
FLO=140.418	NUV=0.200	SEG=95.000	SPR1=0.000	

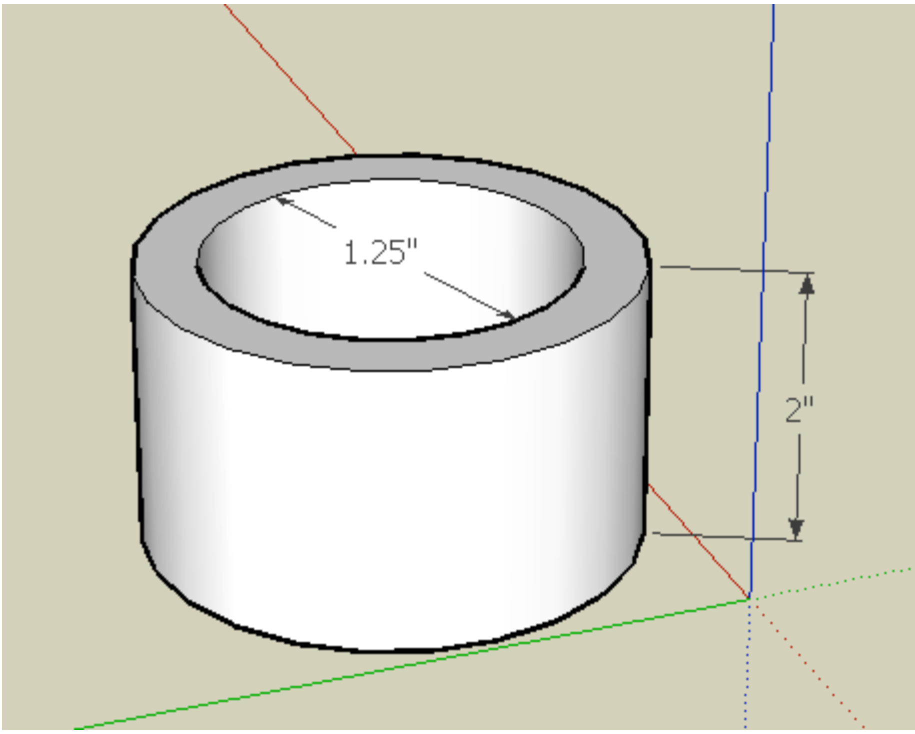
A.4 Mechanical Drawings

Manipulation Module

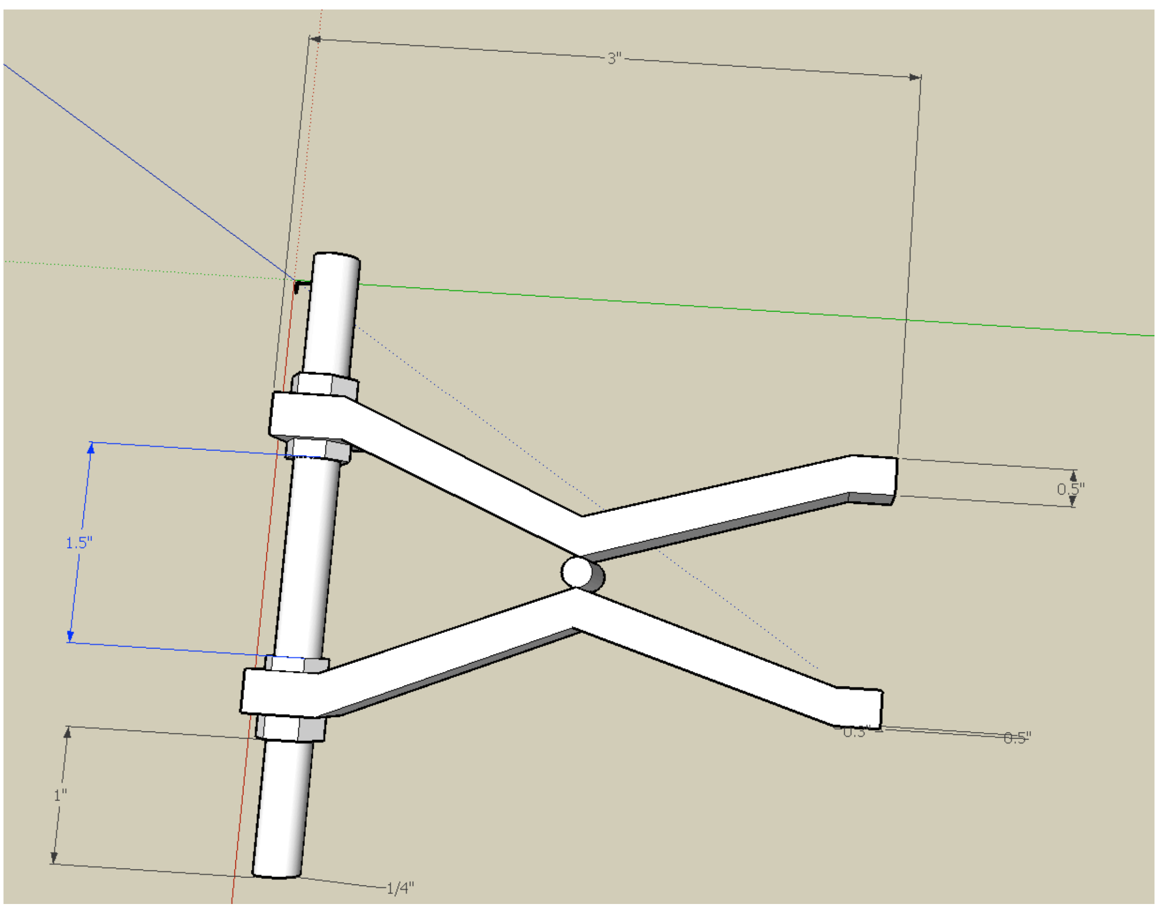



I_DIA 3/4"
O_DIA 1 1/2"
DIA 1.5"

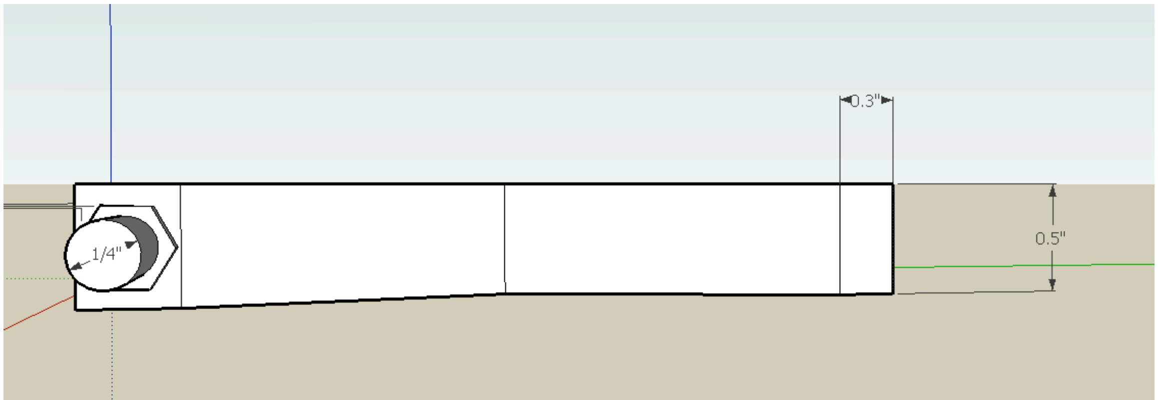

DIA 1.5"
I_DIA 3/4"
3"

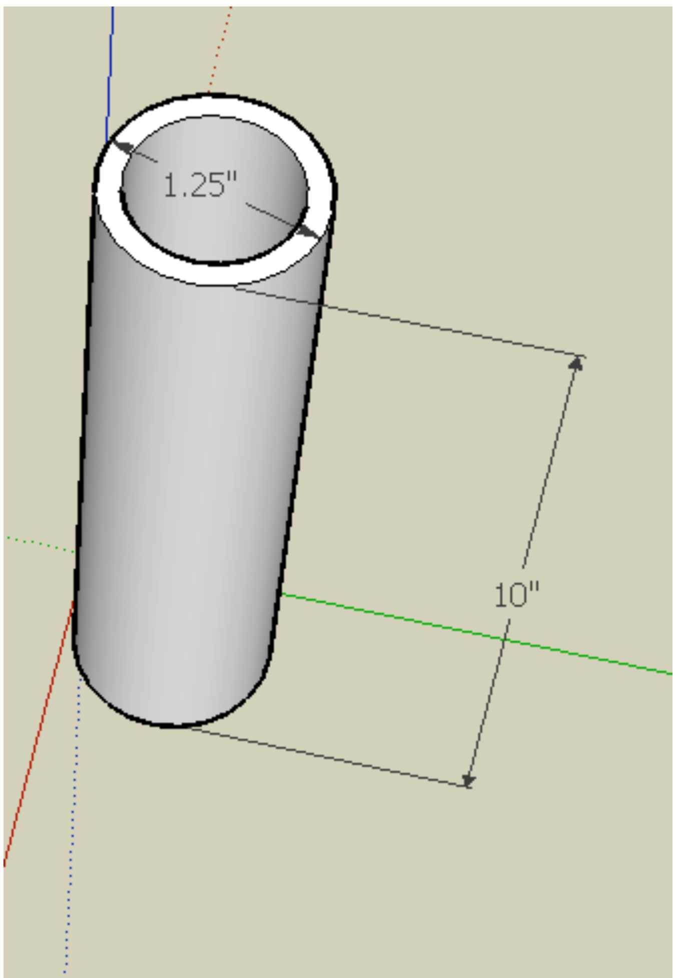


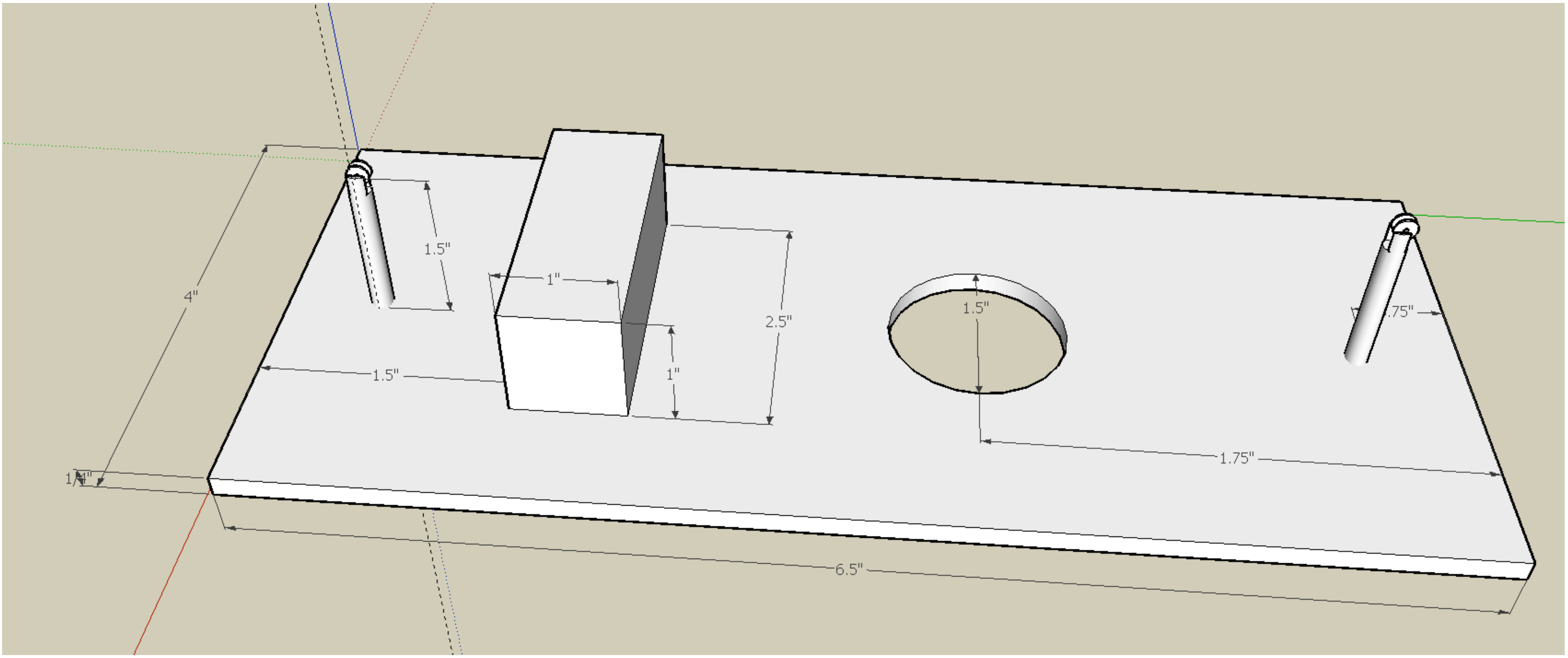
1.25"
2"

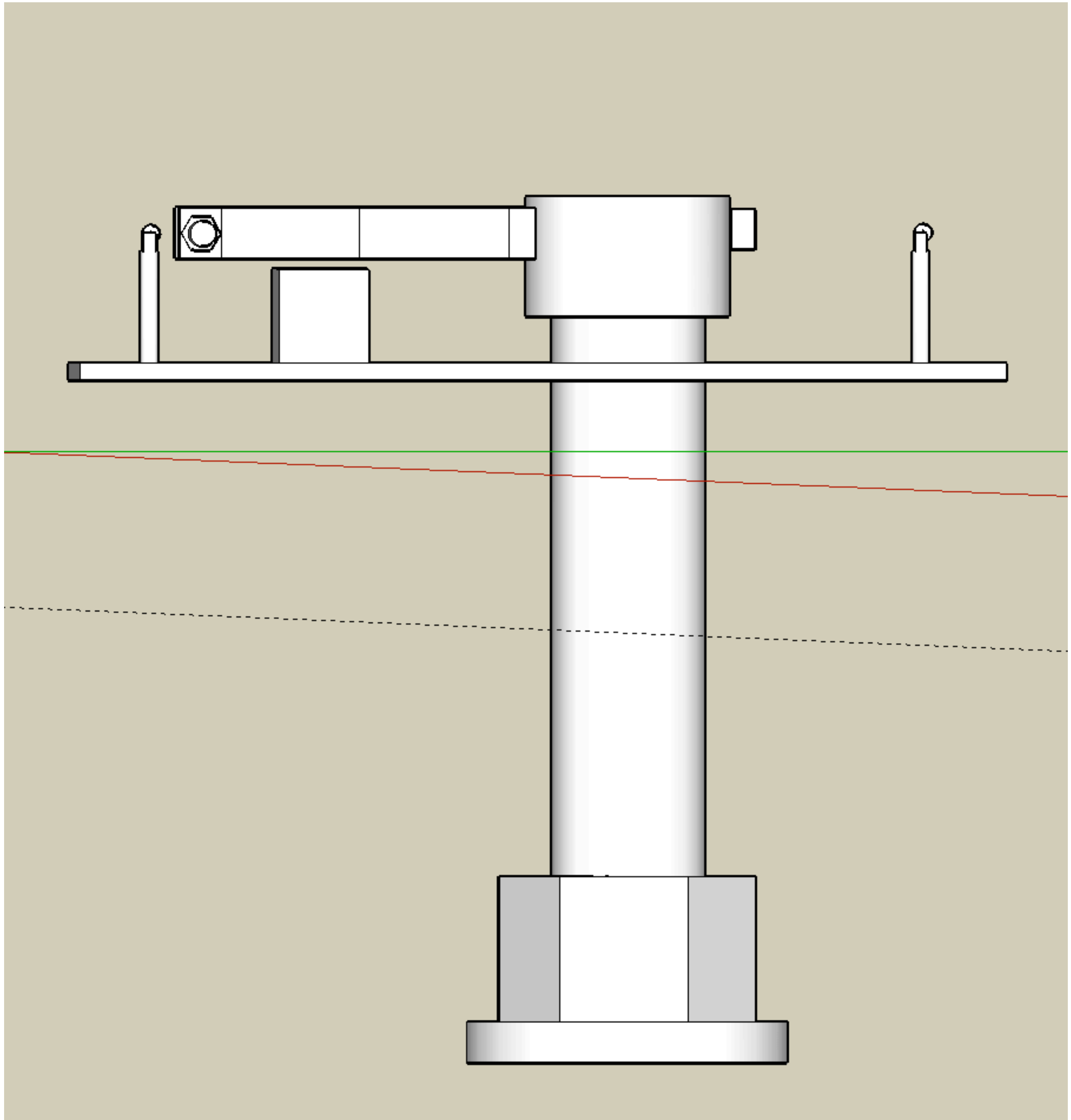


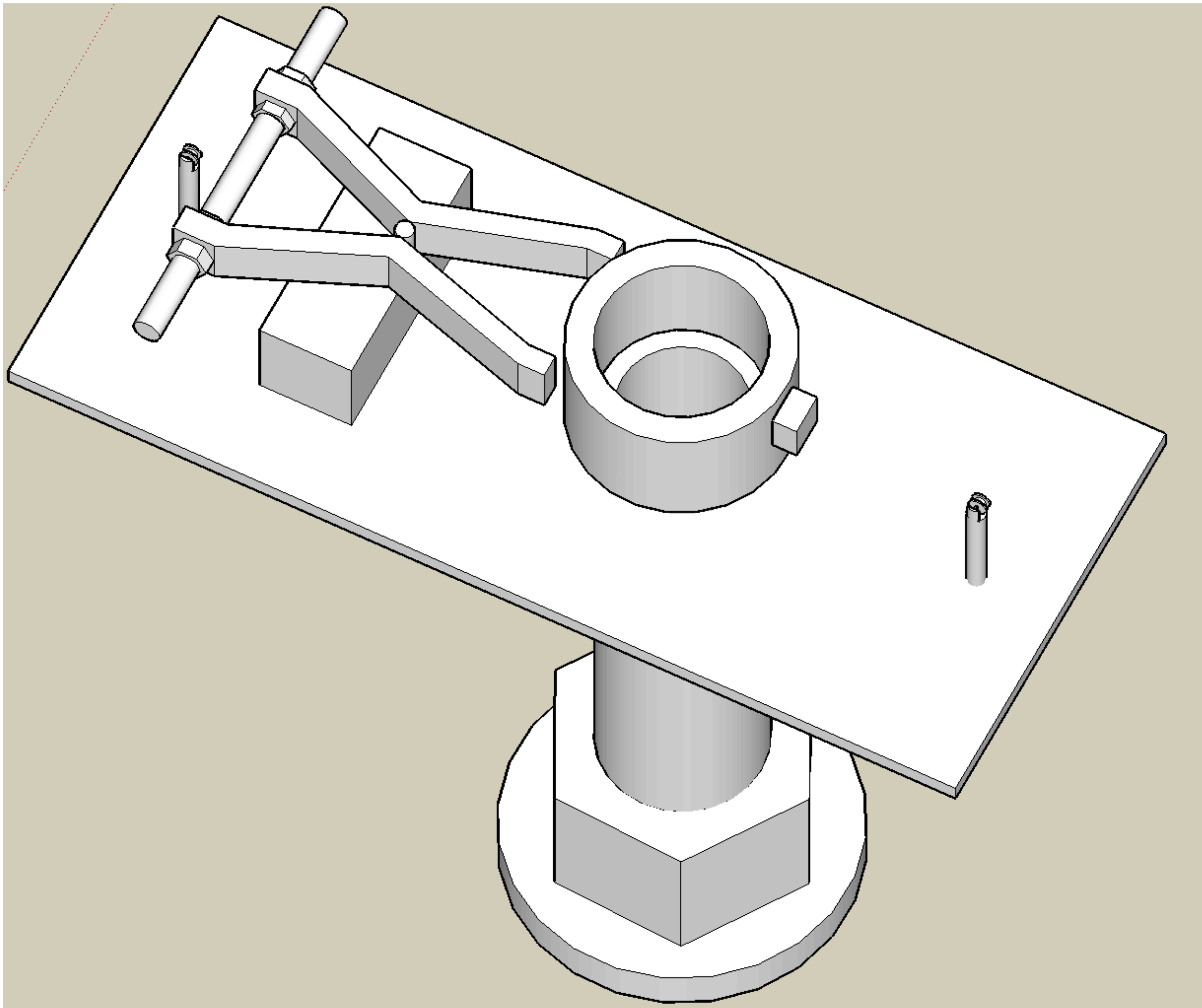
3"
0.5"
1.5"
0.5"
1"
1/4"

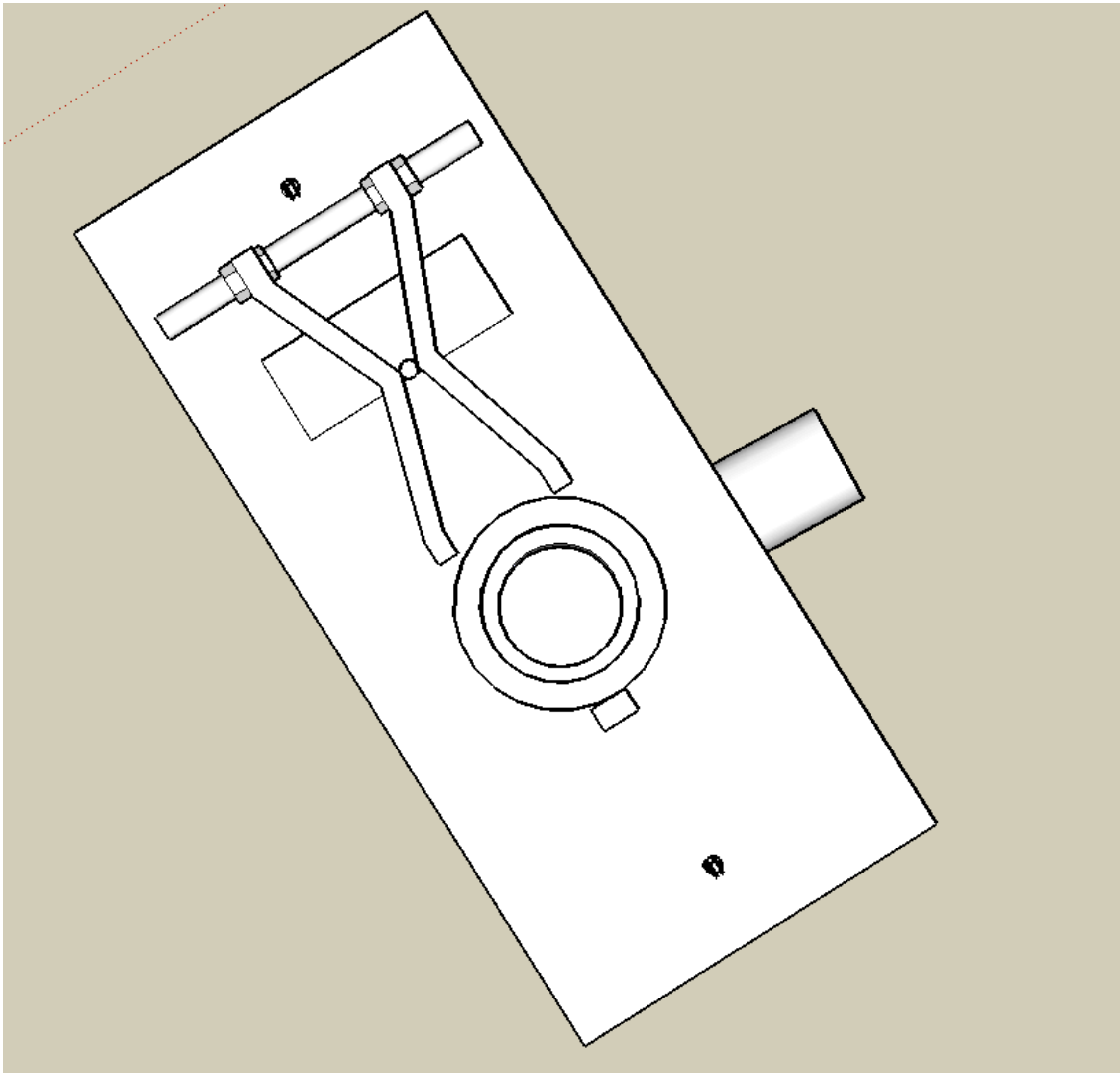

0.3"
0.5"
1/4"


1.25"
10"

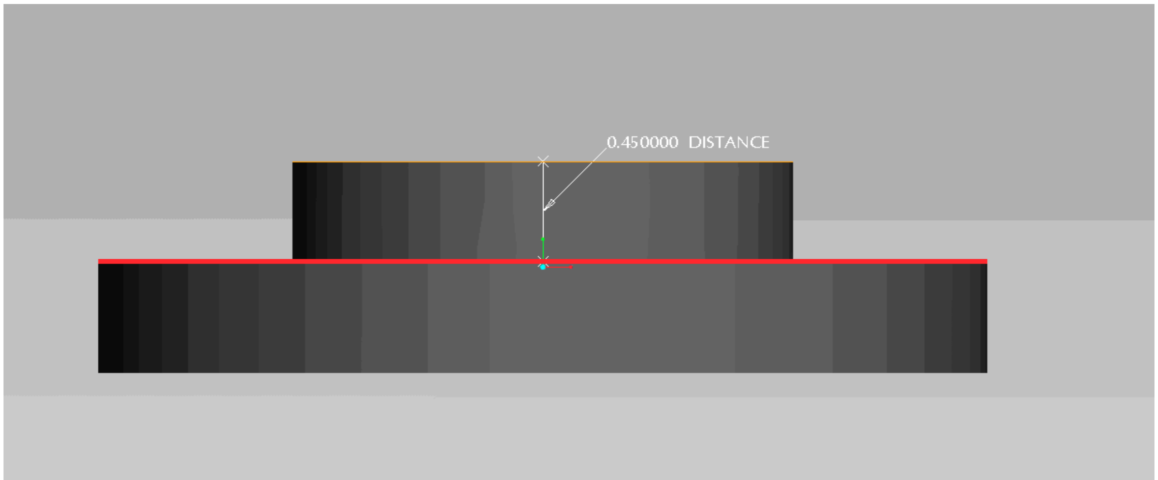

1.5"
1"
4"
2.5"
1.5"
.75"
1.5"
1"
1.75"
1/4"
6.5"



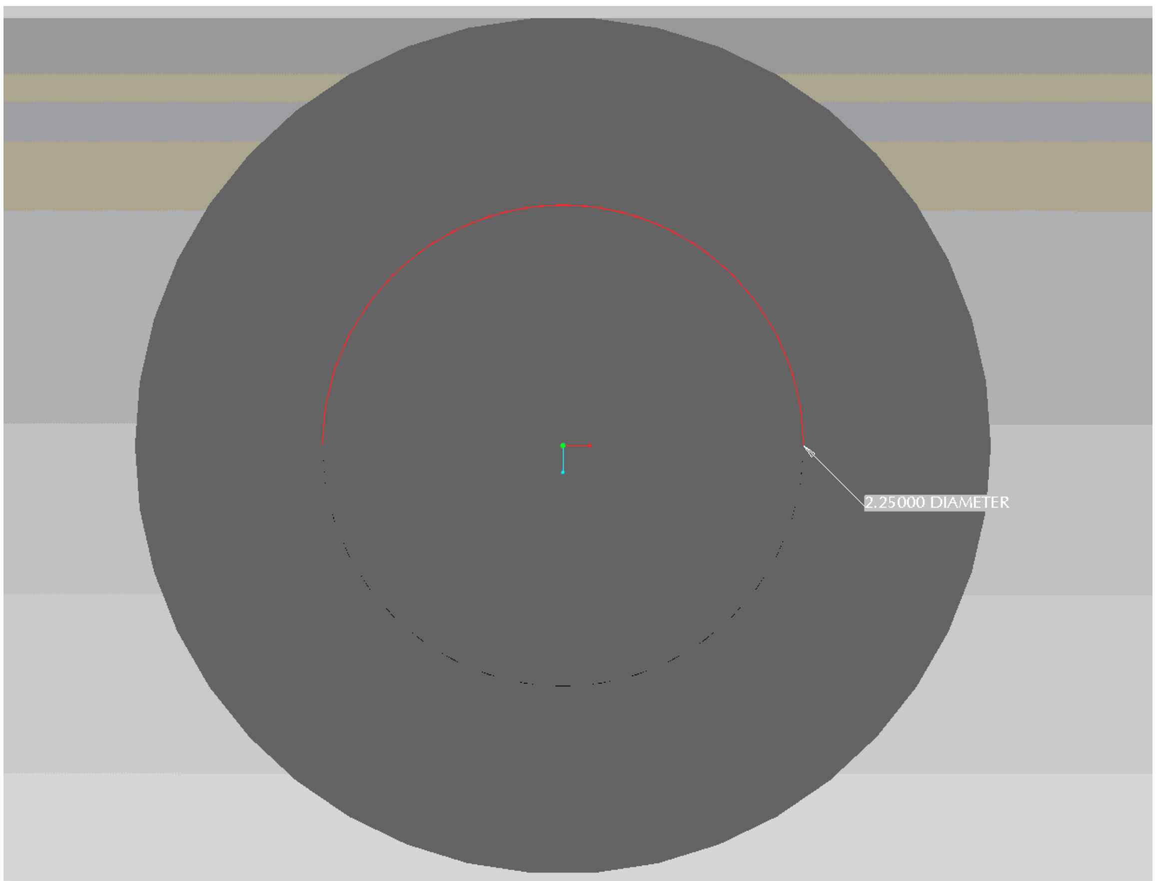




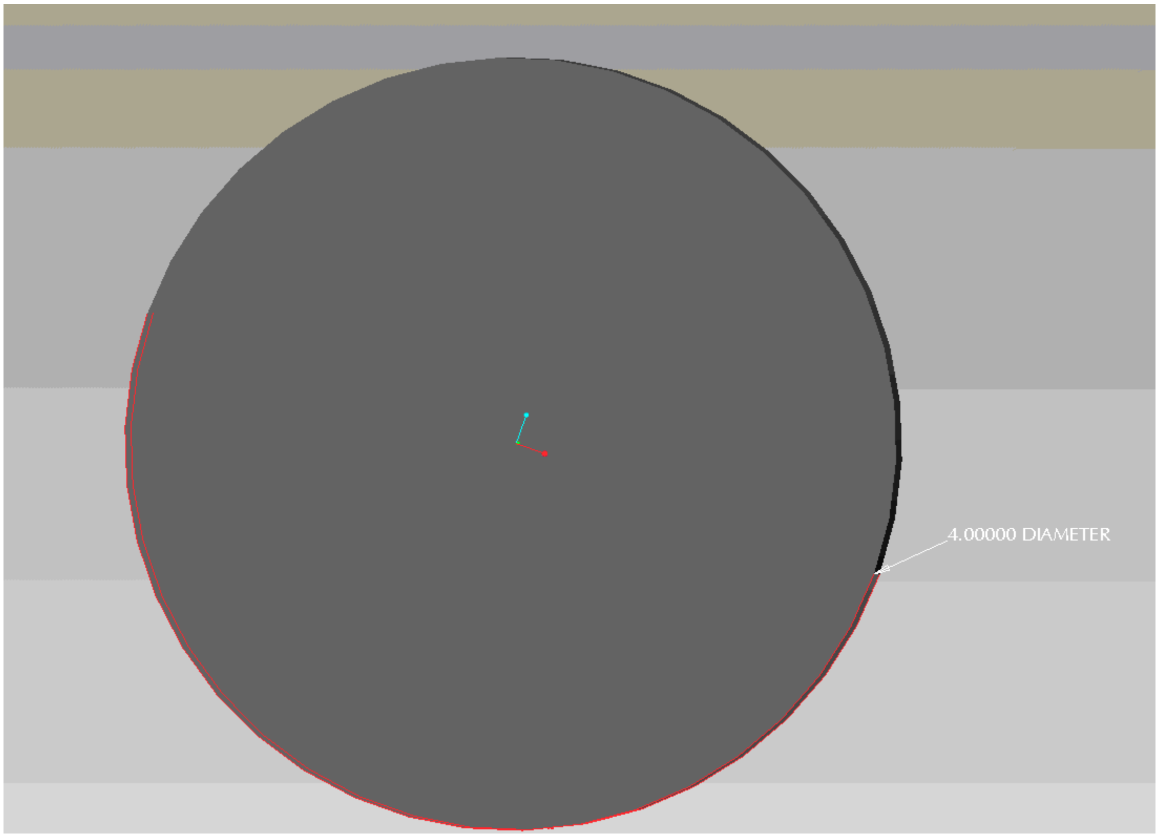
Vocal Fold Mold Mechanical Drawings



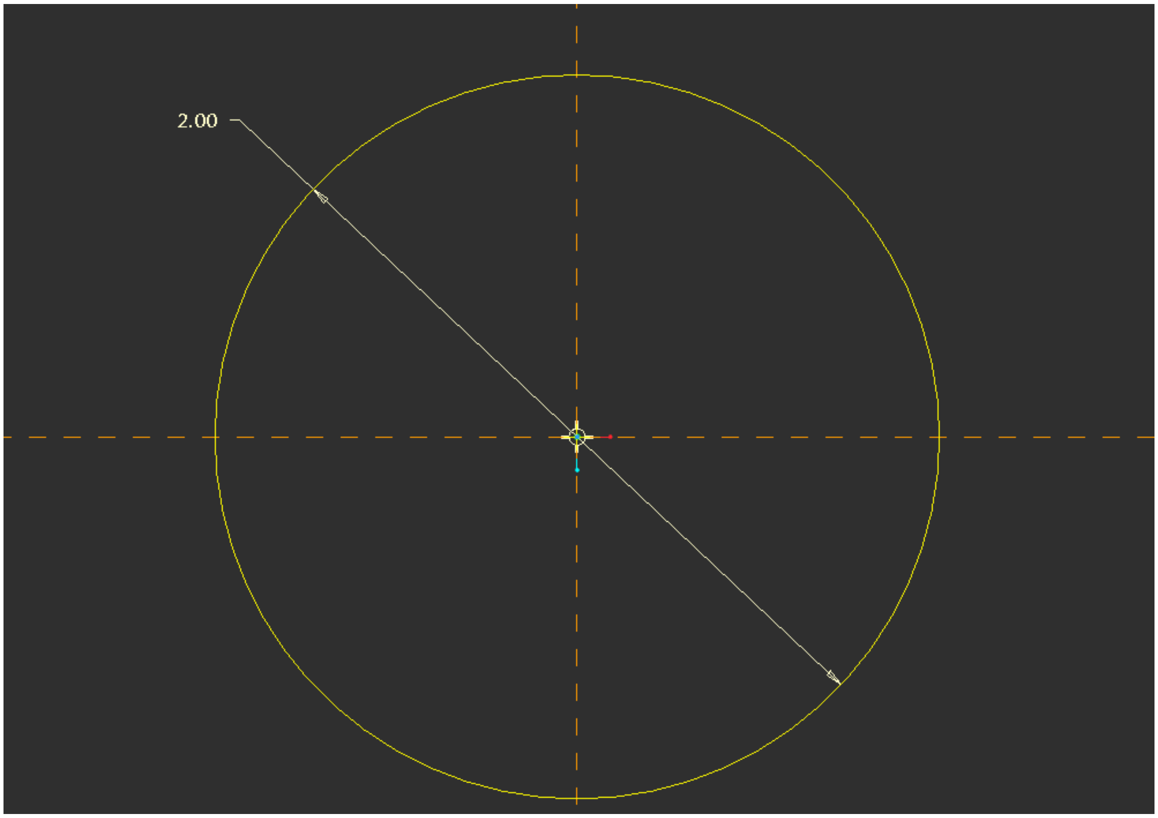
Cover Side

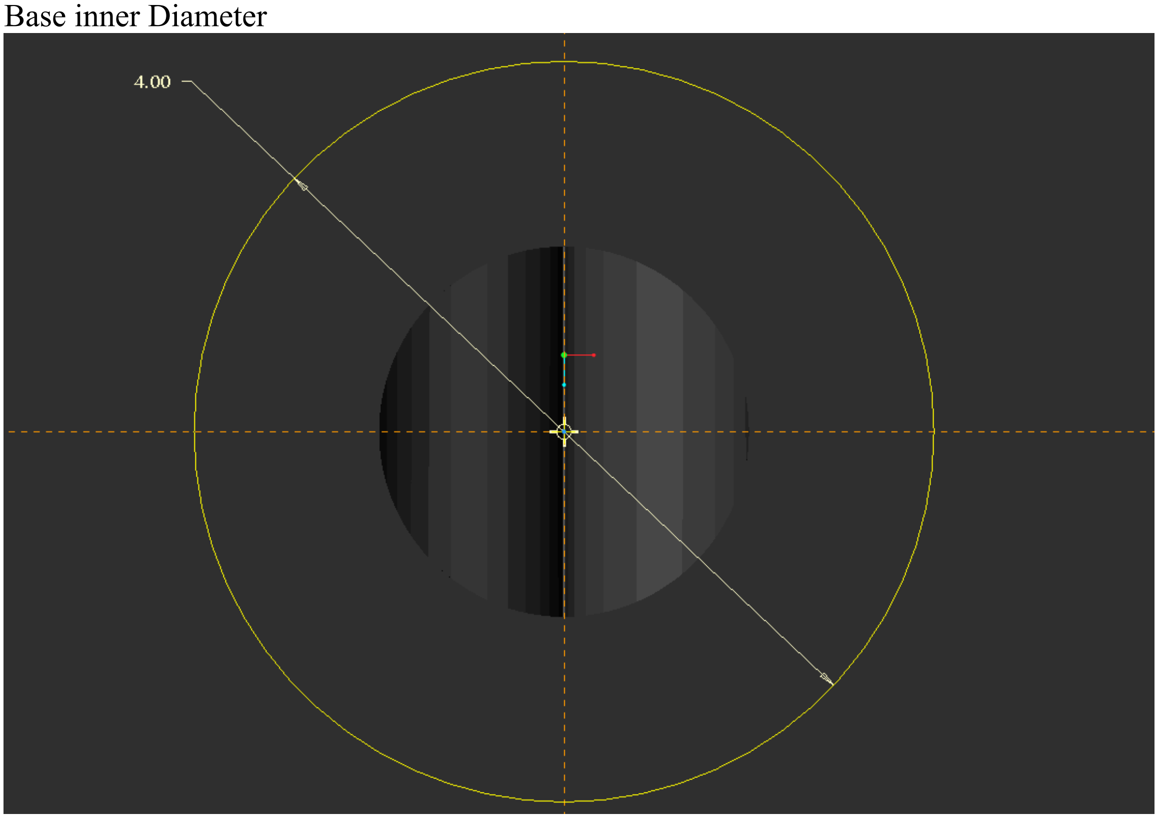


Cover Inner Diameter

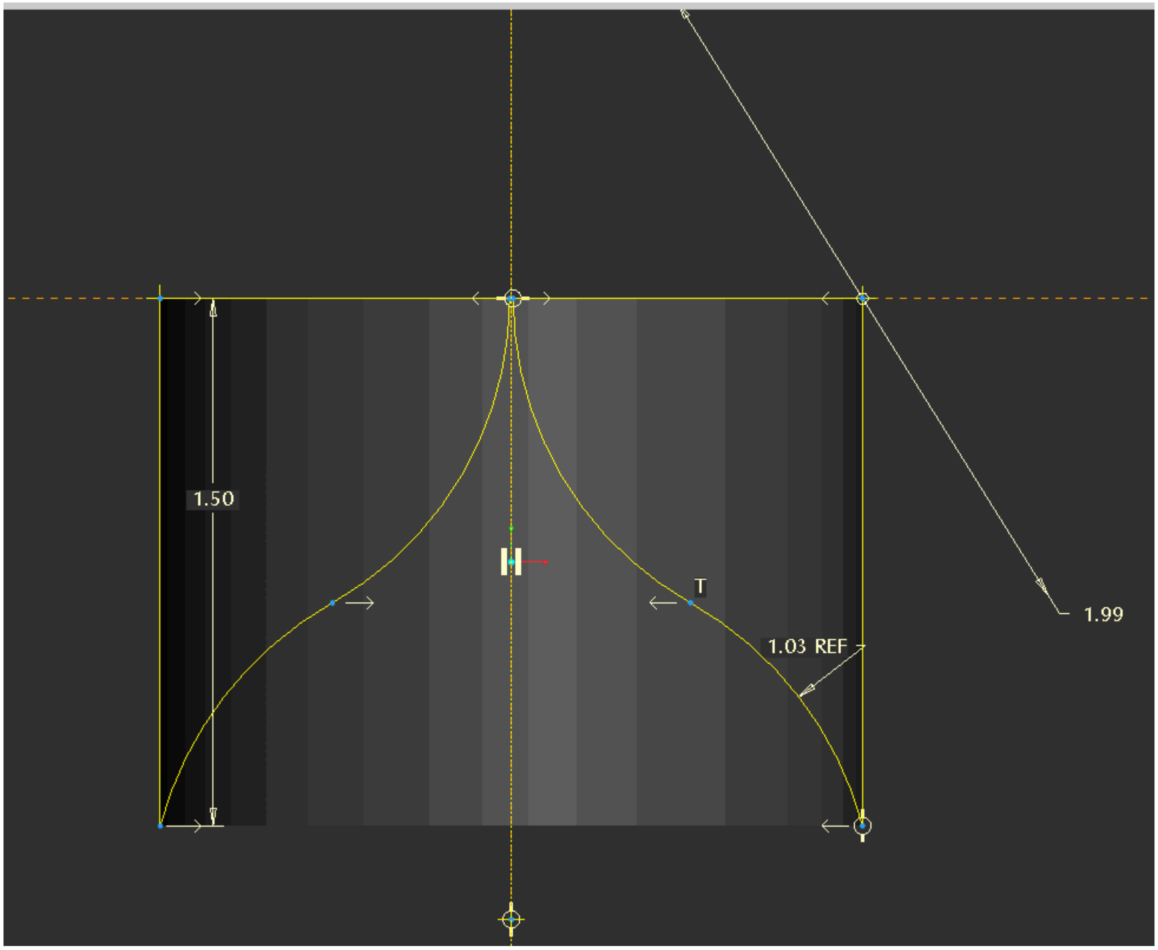


Cover OD

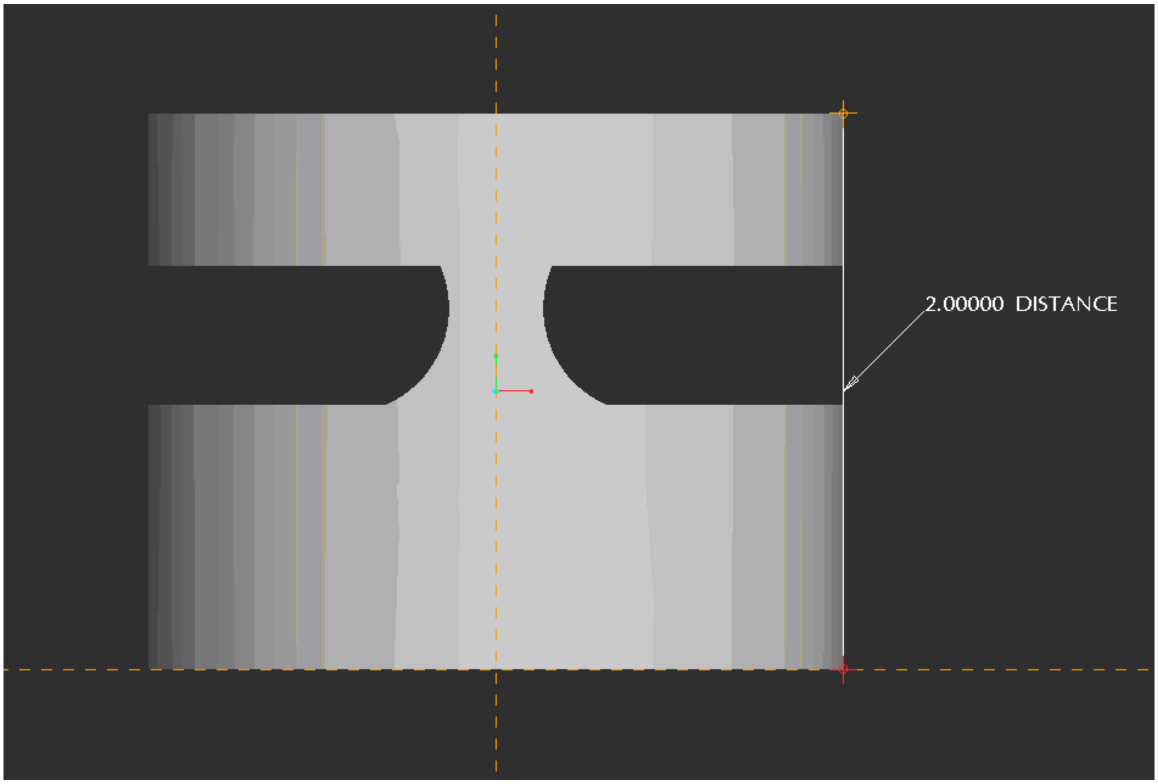


Base inner Diameter


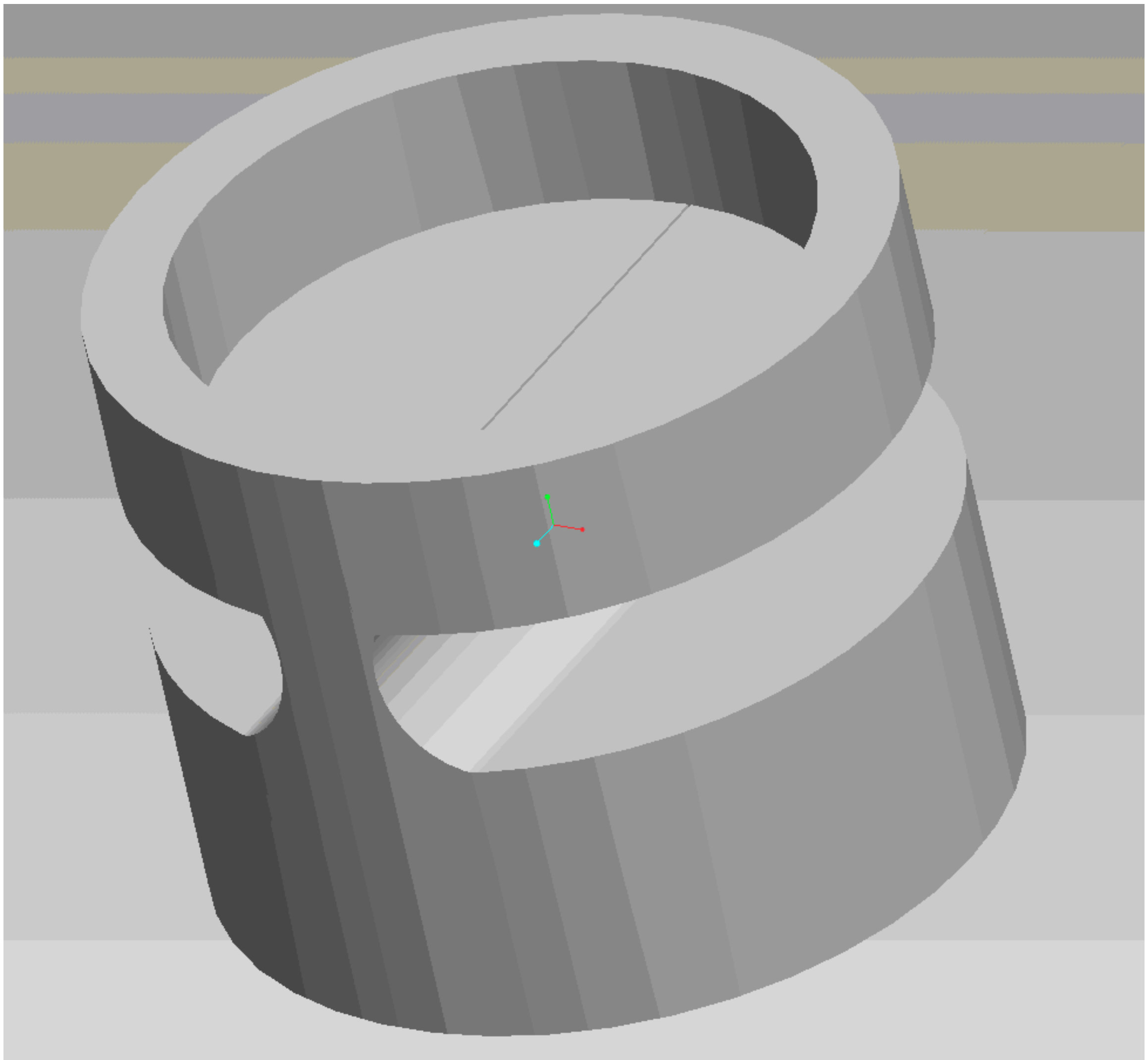
Base Outer Diameter



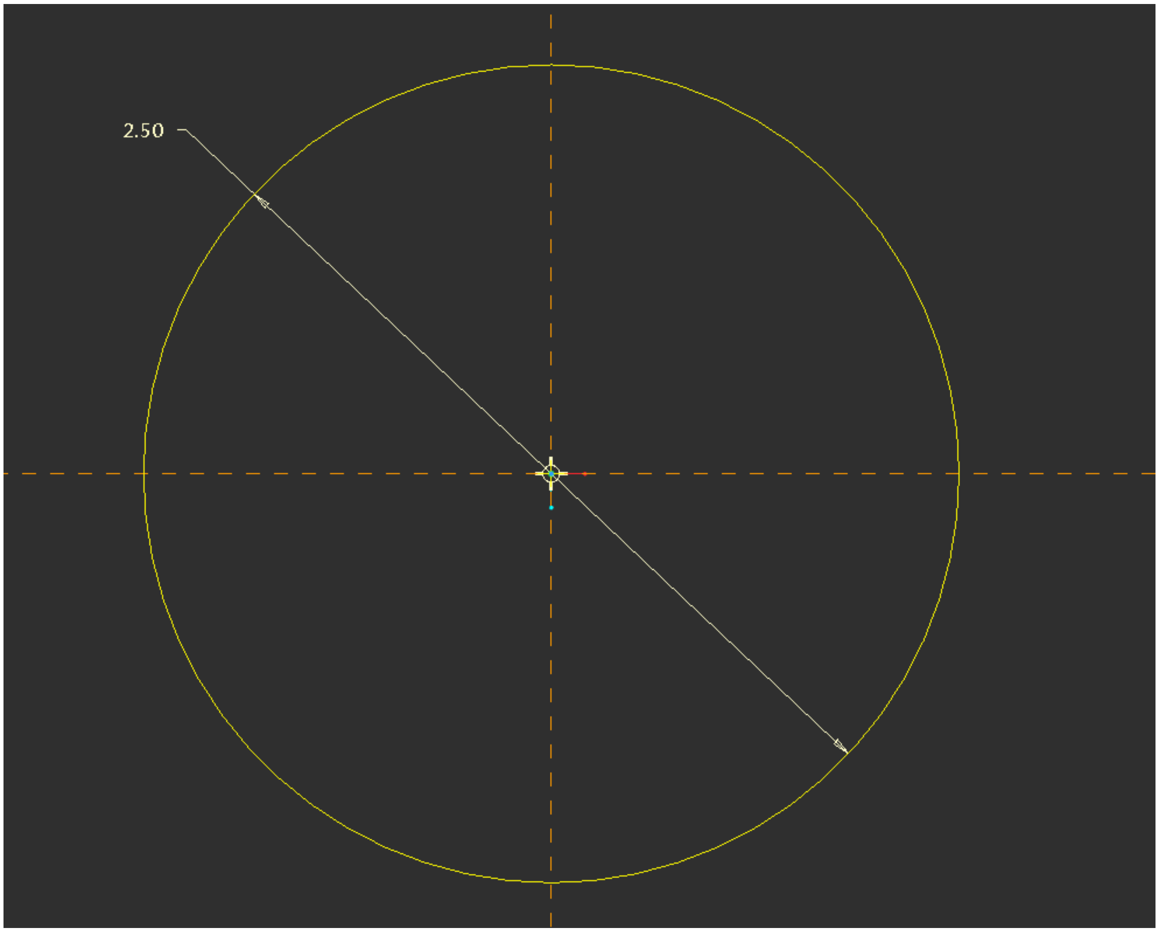
Middle Insert



Vocal fold Mold w/ Height

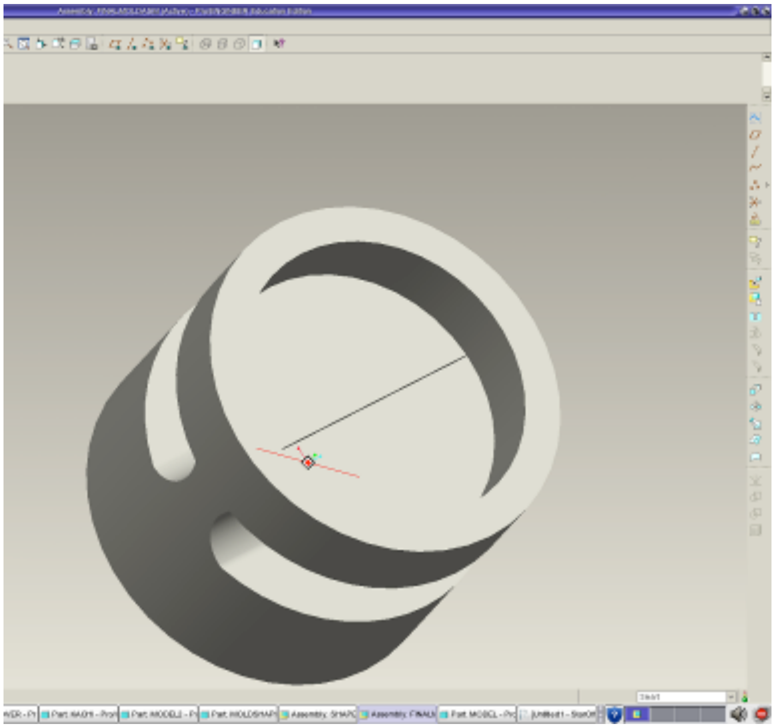


Final vocal fold model from mold

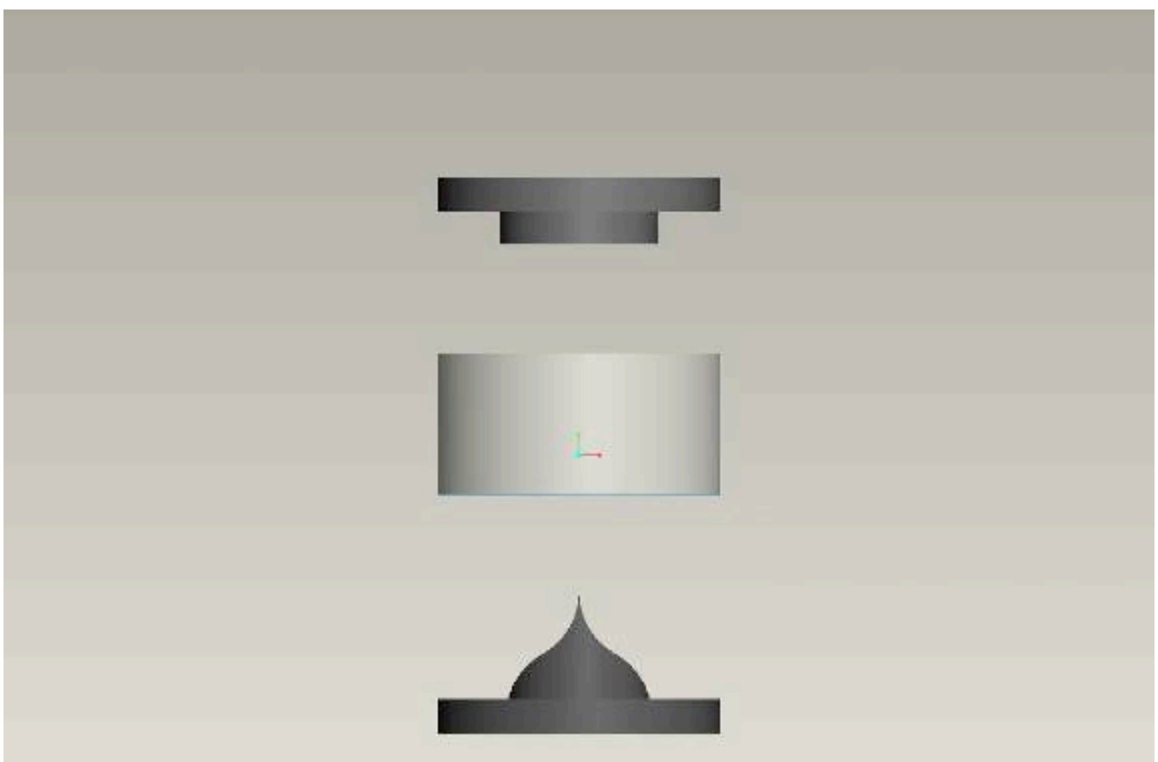


Vocal Fold model Outer Diameter

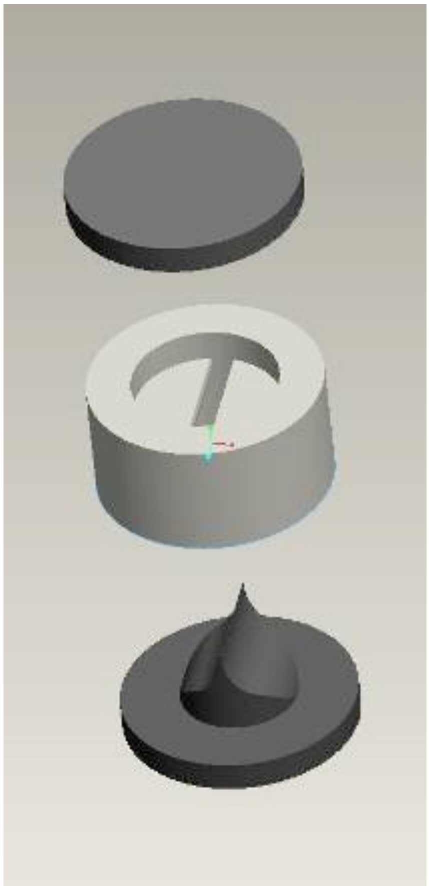
Vocal Fold model from mold

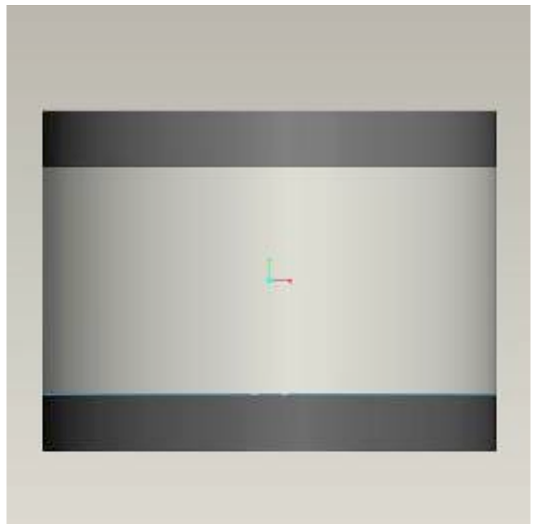


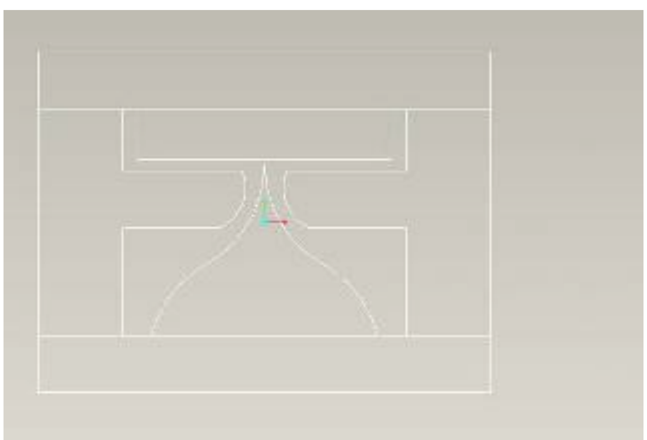
Vocal Fold Model from clay

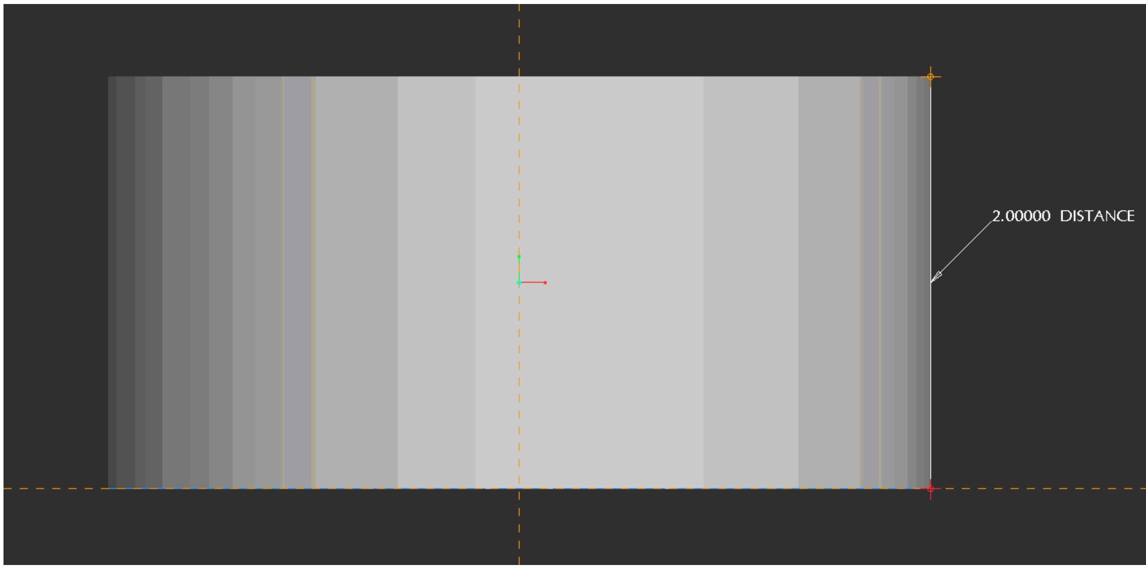


Mold Assembly

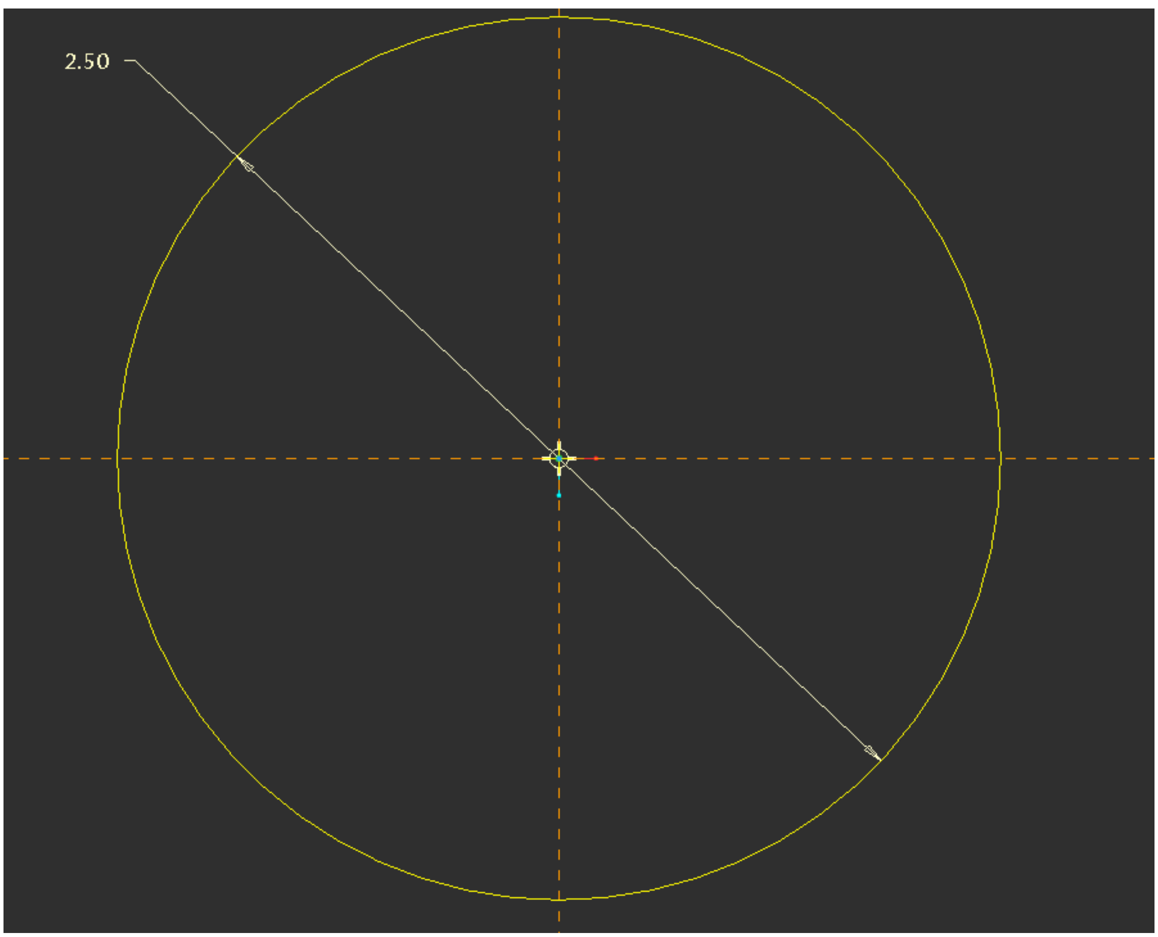




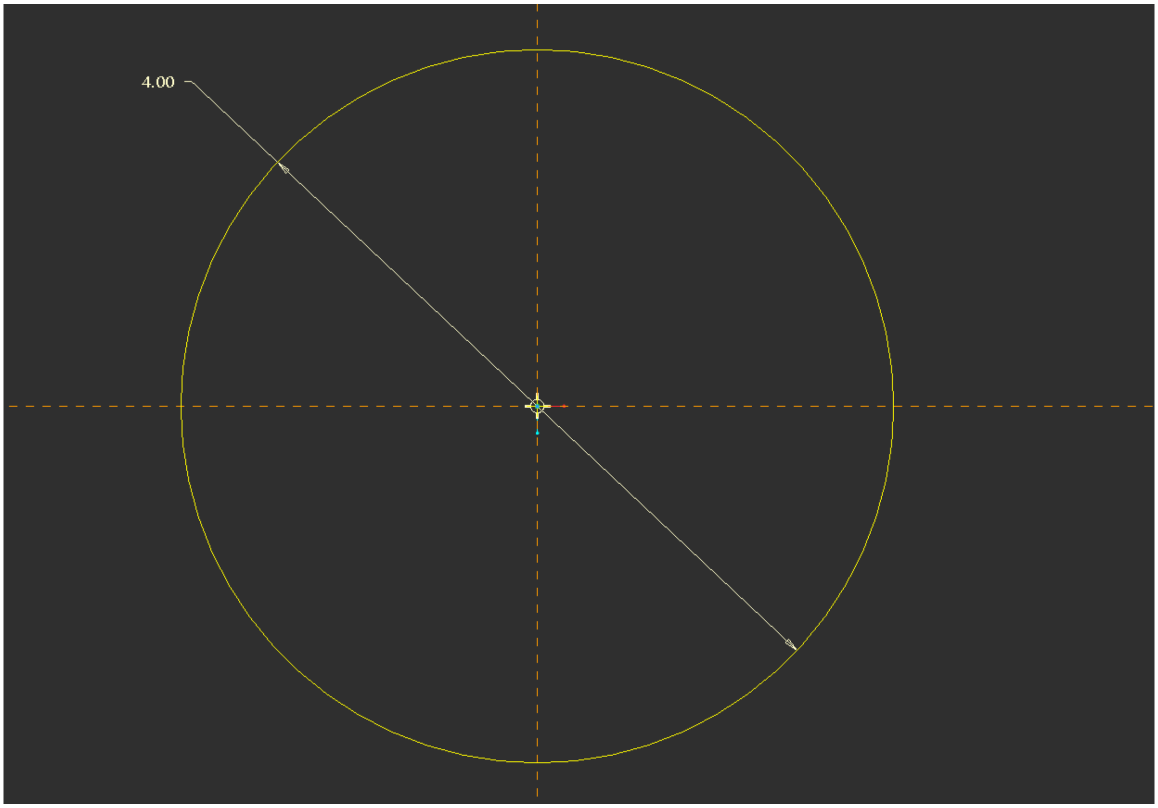




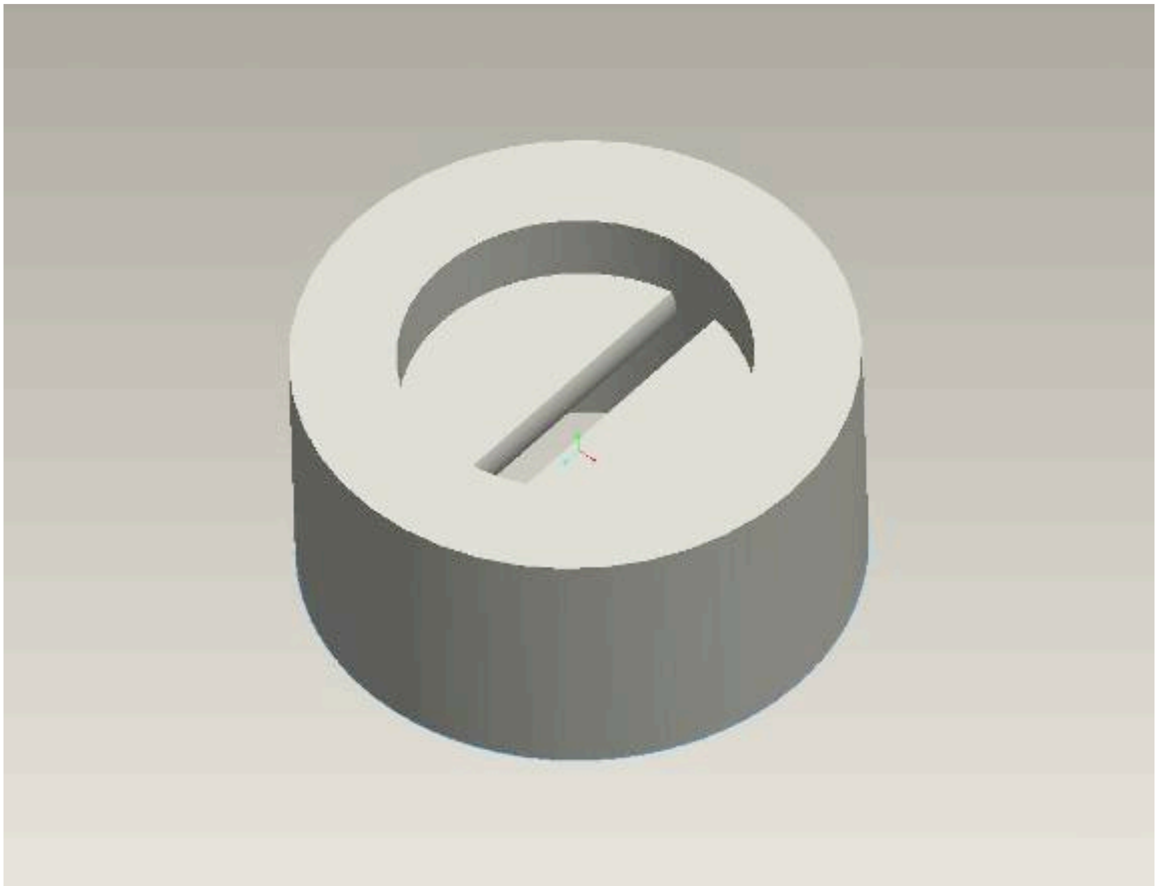
Mold Height



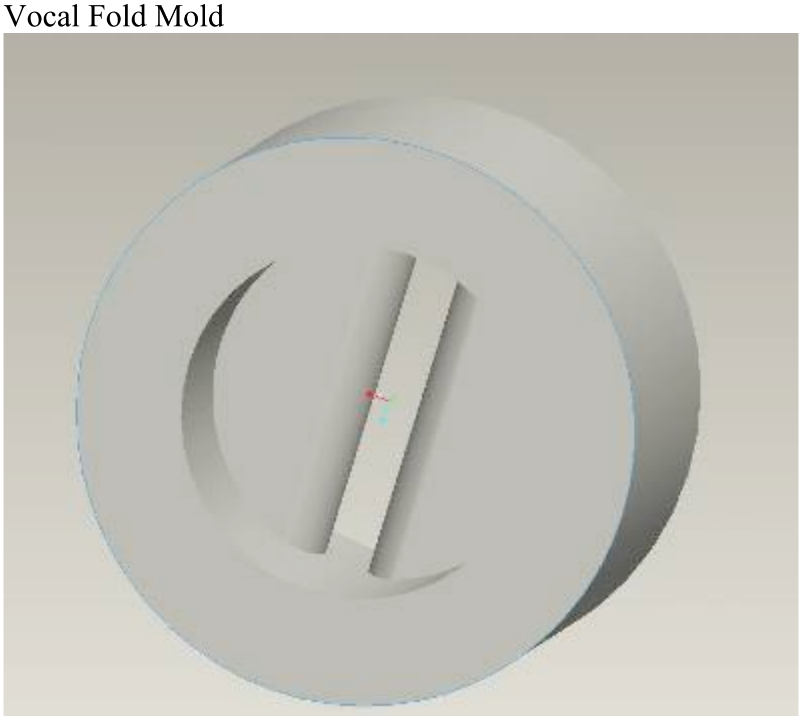
Mold Inner Diameter



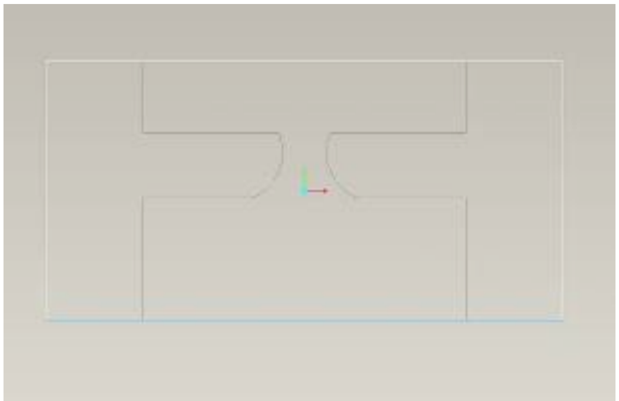
Vocal Fold Mold Outer Diameter



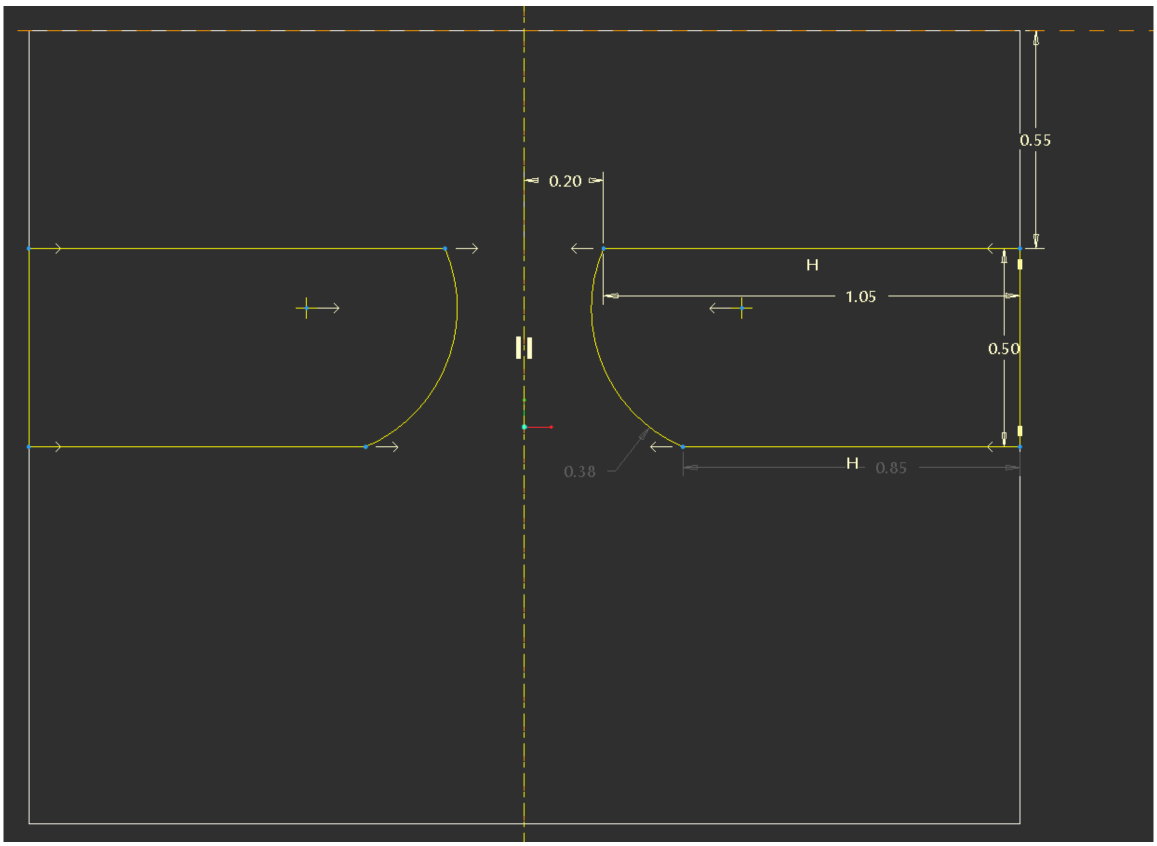
Vocal Fold Mold



Vocal Fold Mold



Vocal Fold Mold Wire View



Vocal Fold Mold w/ Dimensions

Additional Vocal Fold Module Testing Results

For more explanations of the data please see page 65.

High Volume/20090318 220608 High.wav

	Max Amplitude	F0 (Hz)	Highest F0 (Hz)	Lowest F0 (Hz)	Jitter %	Shimmer %	RAP %		File Name
Hi	0.72	485.90	44100.00	155.83	7.54	4.79	1.50	No Correction	20090318 220608 High.wav
		162.72	232.11	155.83	1.75	4.79	0.34	correction	

Table 13: High Volume VF Model —Calculation Results from Software Module

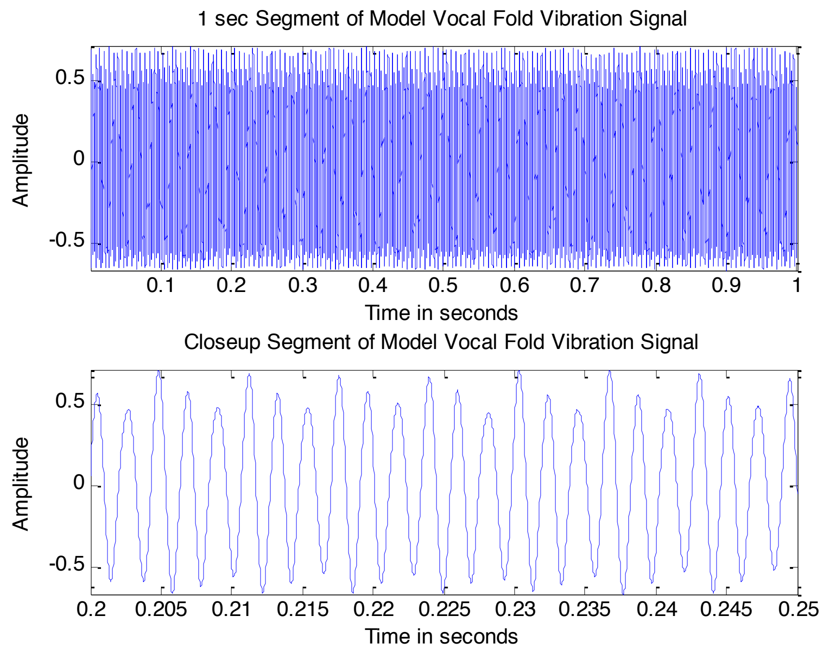


Figure 53: High Volume VF Model-Signal Plot

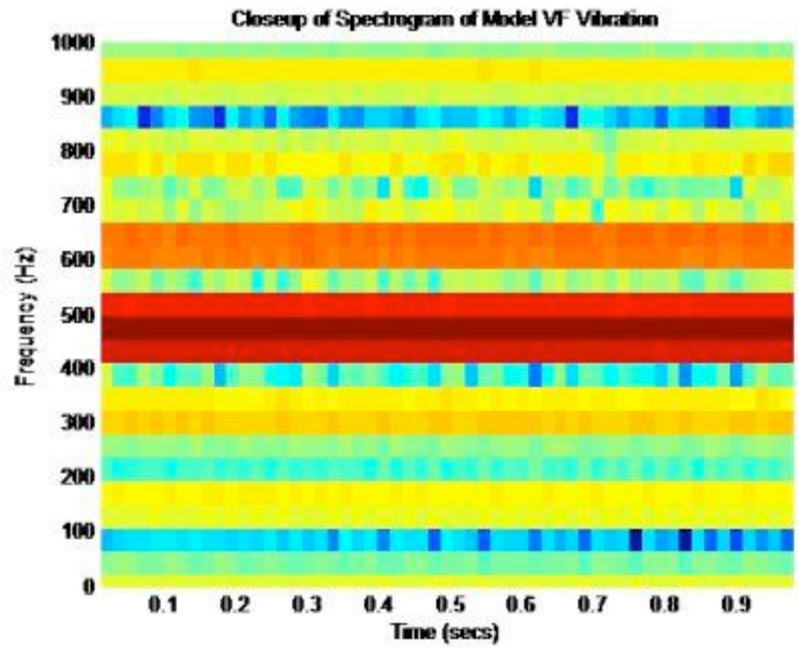


Figure 54: Spectrogram of High Volume VF Model

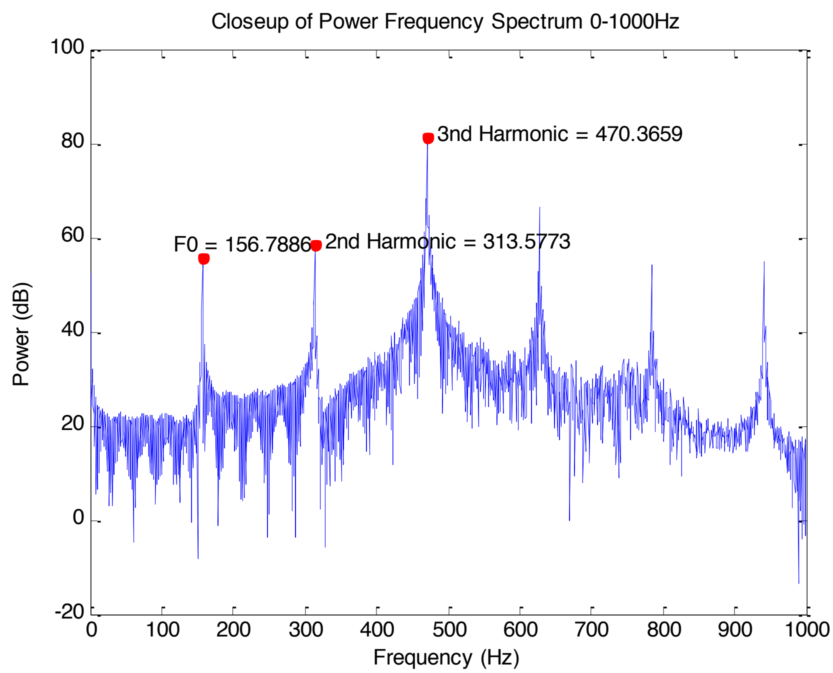


Figure 55: Power Frequency Spectrum of High Volume VF Model

Low Volume/20090318 221340 Medium.wav

	Max Amplitude	F0 (Hz)	Highest F0 (Hz)	Lowest F0 (Hz)	Jitter %	Shimme r %	RAP %		File Name
Lo w	0.38	157. 43	158.63	156.38	0.12	1.62	0.03	No Correction	20090318 221340 Medium.wav
		157. 43	158.63	156.38	0.12	1.62	0.03	correction	

Table 14: Low Volume VF Model —Calculation Results from Software Module

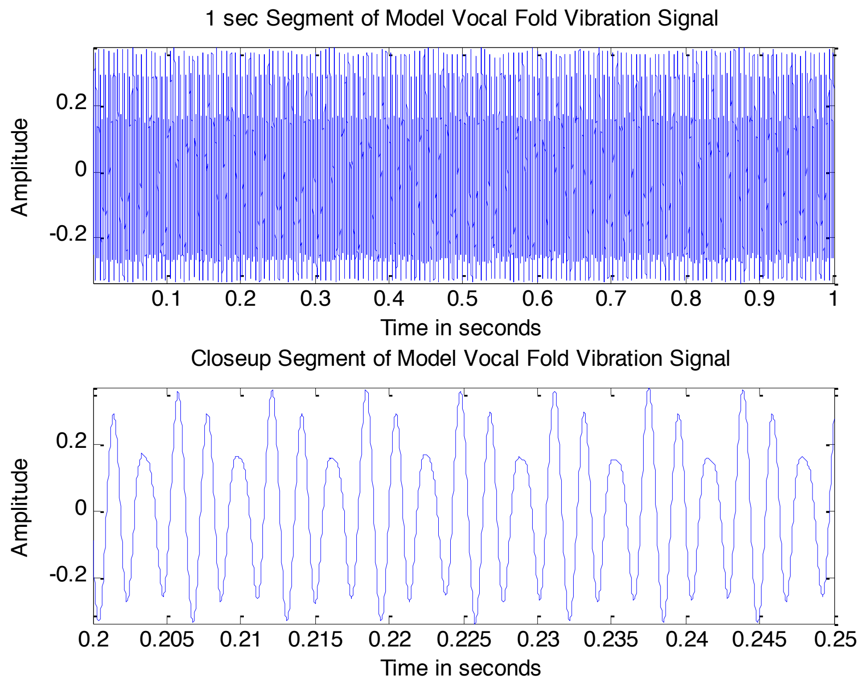


Figure 56: Low Volume VF Model Signal Plot

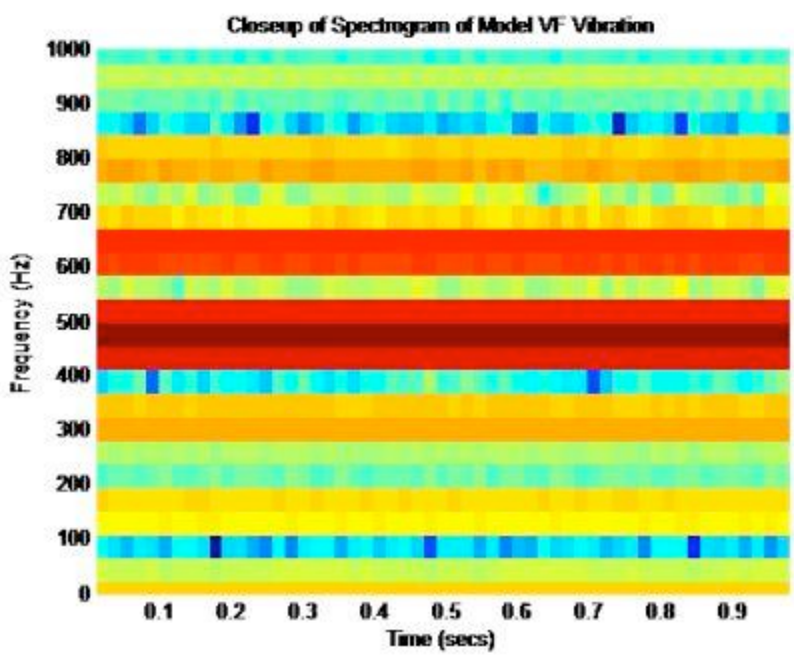


Figure 57: Spectrogram of Low Volume-VF Model

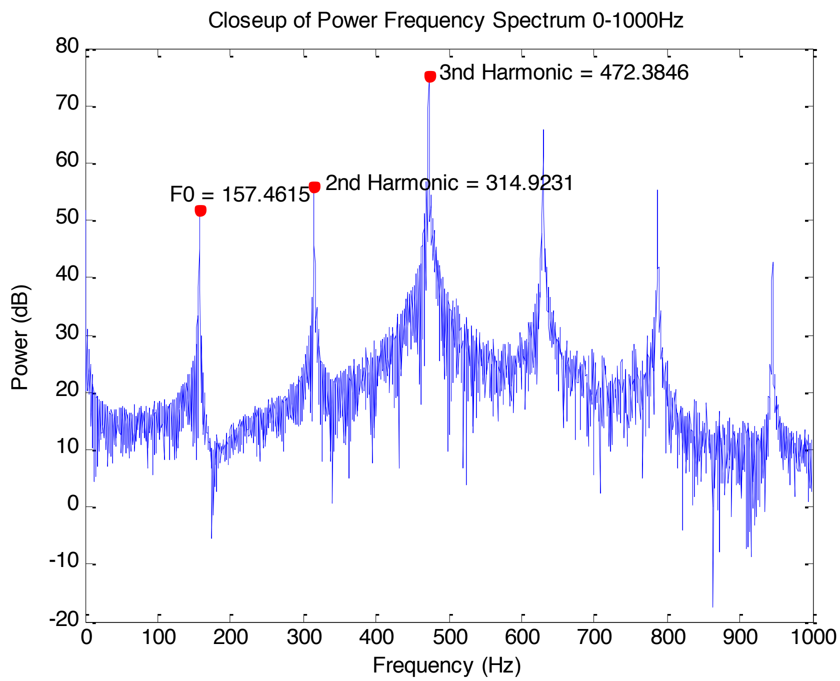


Figure 58: Power Frequency Spectrum of Low Volume-VF Model

Low Volume/20090318 221605 Low.wav

	Max Amplitude	F0 (Hz)	Highest F0 (Hz)	Lowest F0 (Hz)	Jitter %	Shimmer %	RAP %		File Name
Low	0.25	157.89	159.78	156.38	0.18	3.41	0.04	No Correction	20090318 221605 Low.wav
		157.89	159.78	156.38	0.18	3.41	0.04	correction	

Table 15: LowVolume VF Model —Calculation Results from Software Module

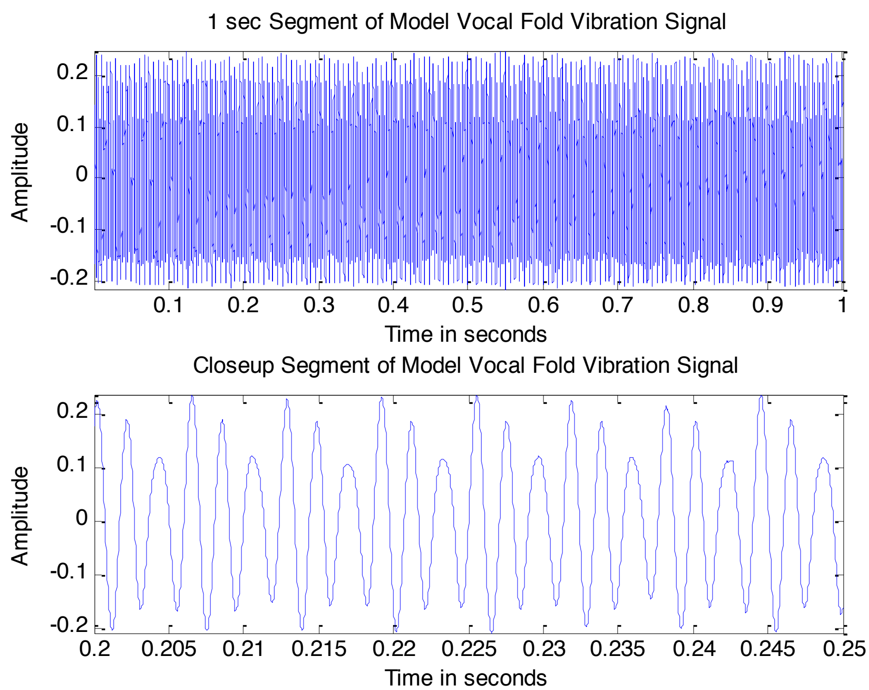


Figure 59: Low Volume VF Model Signal Plots

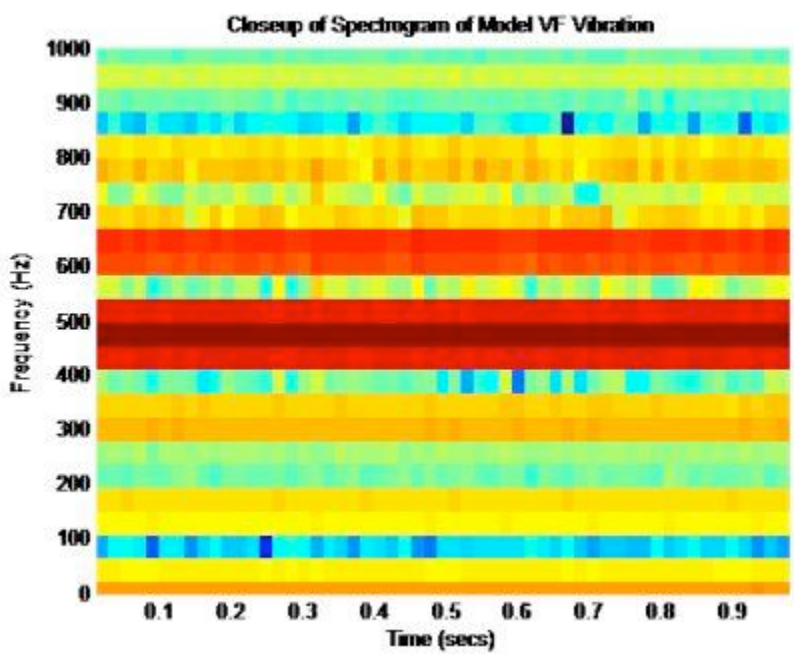


Figure 60: Spectrogram of Low Volume-VF Model

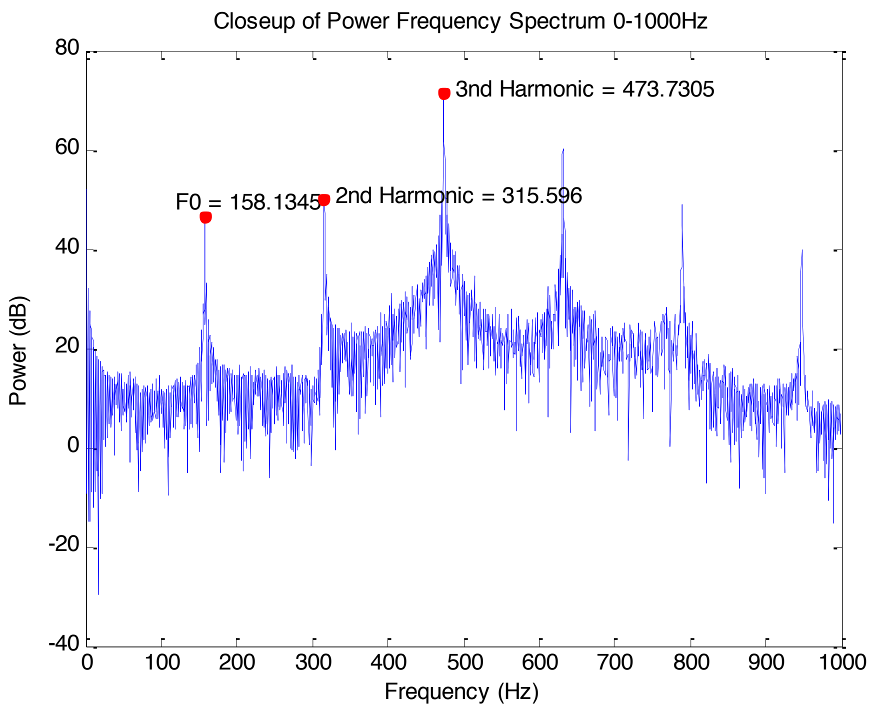


Figure 61: Power Spectrum of Low Volume-VF Model

Medium Volume/20090318 221957 Medium.wav

	Max Amplitude	F0 (Hz)	Highest F0 (Hz)	Lowest F0 (Hz)	Jitter %	Shimmer %	RAP %		File Name
Medium	0.51	798.39	44100.00	238.38	9.00	19.99	1.57	No Correction	20090318 221957 Medium.wav
		469.97	506.90	432.35	1.86	19.99	0.35	correction	

Table 16: MediumVolume VF Model —Calculation Results from Software Module

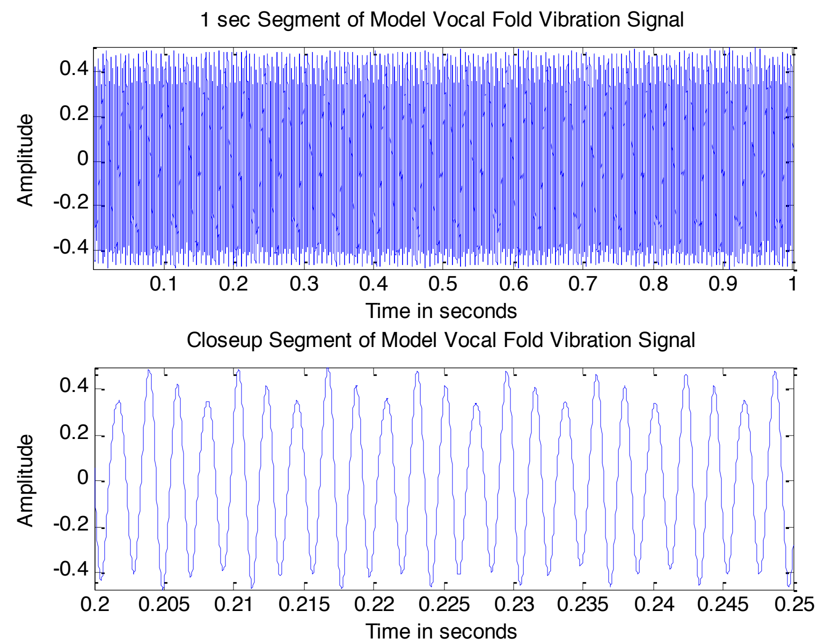


Figure 62: Signal Plots of Medium Volume VF Model

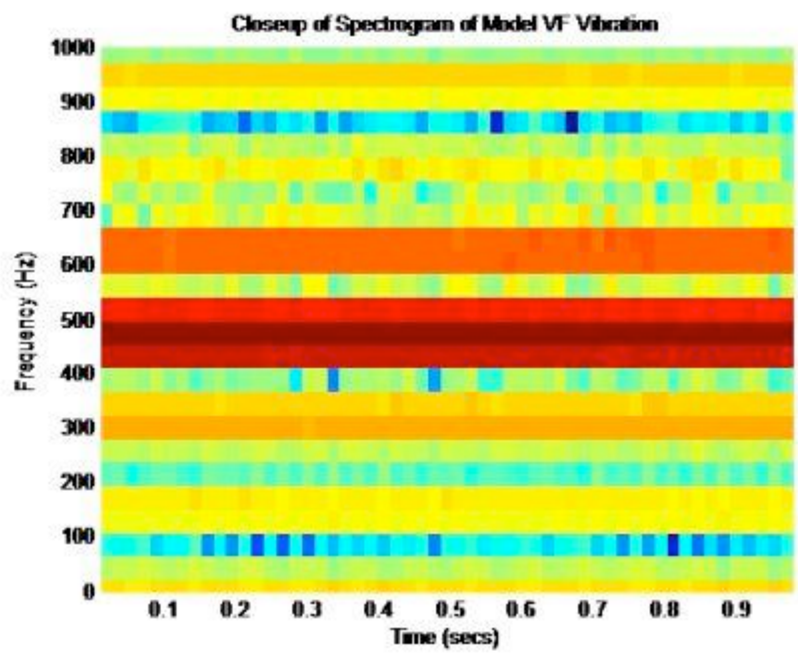


Figure 63: Spectrogram of Medium Volume-VF Model

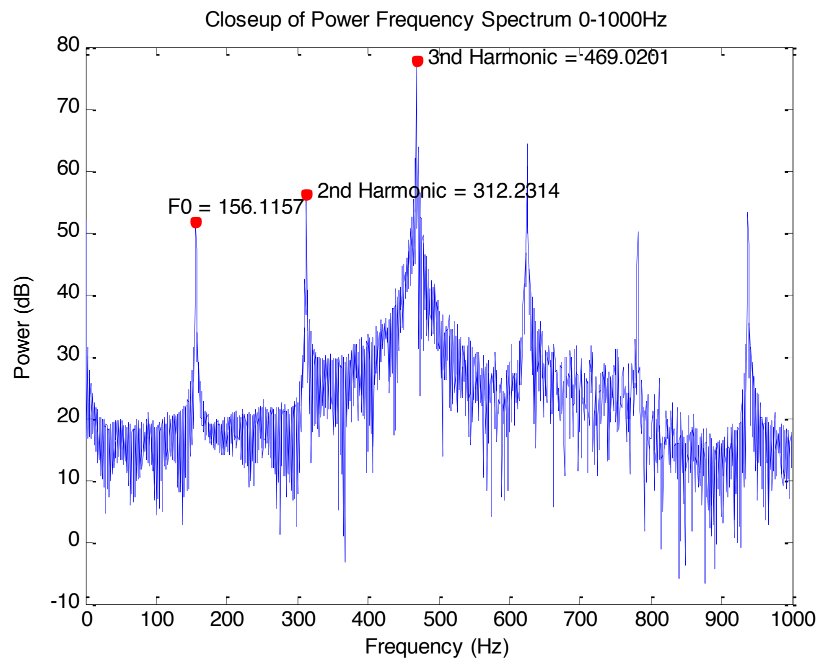


Figure 64: Power Frequency Spectrum Medium Volume-VF Model

Low Volume/20090318 222956 Low.wav

	Max Amplitude	F0 (Hz)	Highest F0 (Hz)	Lowest F0 (Hz)	Jitter %	Shimmer %	RAP %		File Name
Low	0.14	157.56	159.78	155.83	0.22	2.90	0.04	No Correction	20090318 222956 Low.wav
		157.56	159.78	155.83	0.22	2.90	0.04	correction	

Table 17: Low Volume VF Model —Calculation Results from Software Module

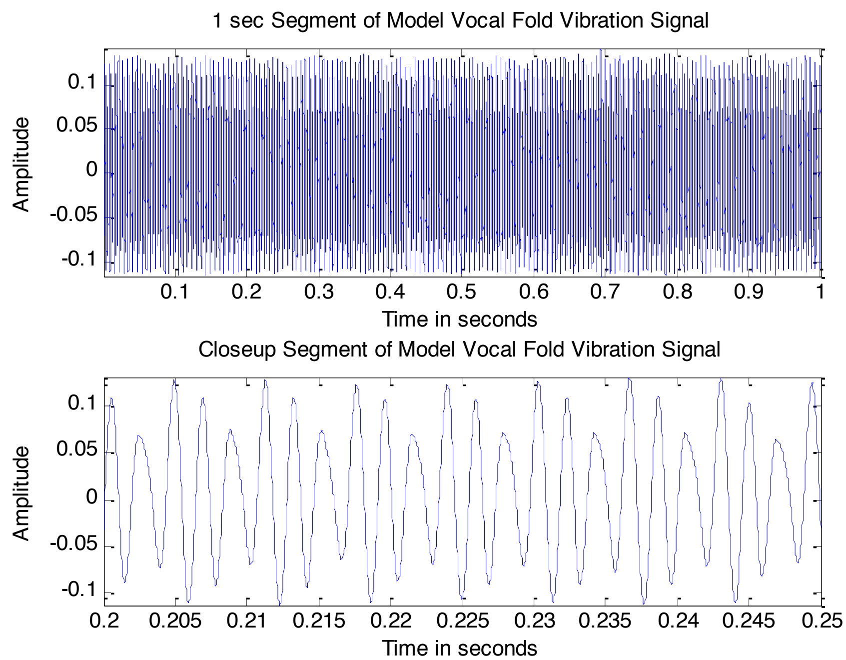


Figure 65: Low Volume VF Model Signal Plots

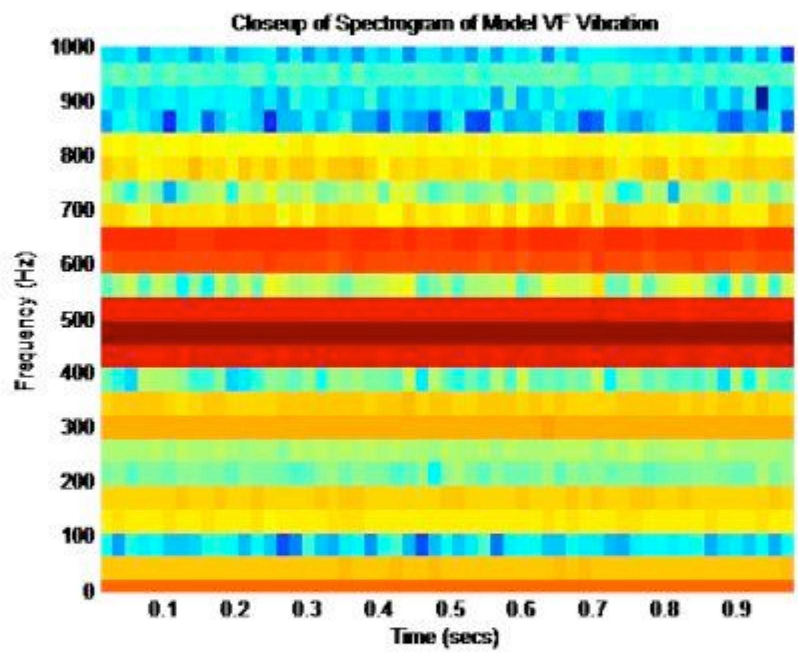


Figure 66: Spectrogram of Low Volume -VF Model

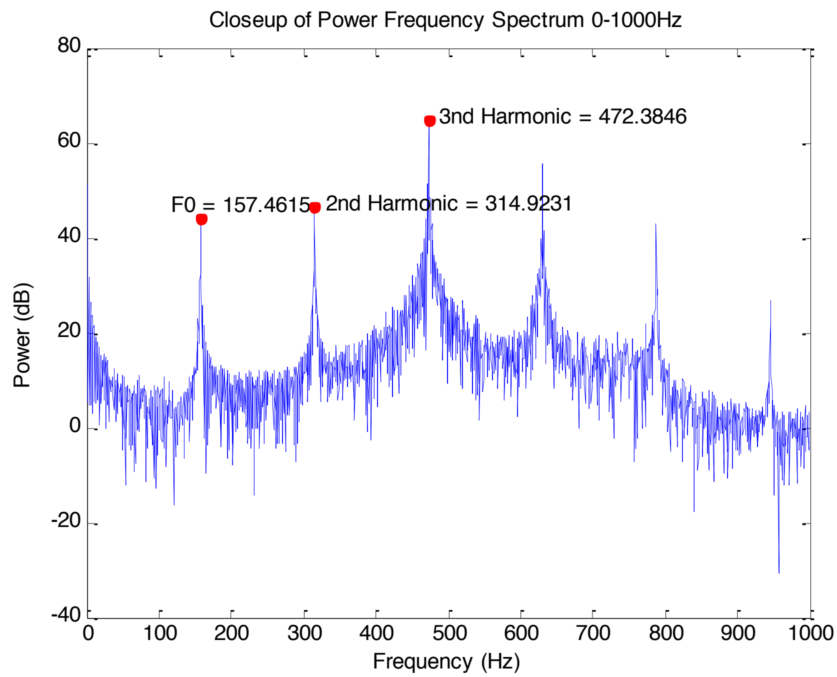


Figure 67: Power Frequency Spectrum of Low Volume VF Model