DWANI-AN ARTIFICIAL VOICE BOX

[1]Mr.Ravi M V, [2]Ramkumar K H, [3]Aayush R Nayak, [4]Venu V, [5]Shubhanjar Mahapathra
[1]Asst.prof, Department of electronics and communication Engineering,SJCIT
email:ravimvreddy86@gmail.com
[2,3,4,5]Students, Department of Telecommunication Engineering,SJCIT
[2]ramkumarkh5@gmail.com, [3]aayushnayak1999@gmail.com
[4]venugoogly415@gmail.com, [5]shubhankar25.m@gmail.com,

Abstract: In a word, there are about 600,000 people whose larynxes have been removed. In India also every year more than 17000 peoples suffered from Laryngeal cancer so these people lose their larynx, these people are not able to make voice, so they have problems in communication. For example, they cannot communicate their will enough with communication by writing or gestures alone, thereby avoiding communication. However, they can move the mouth and tongue normally. Therefore, they speak using alternative vocalizations. Electrical artificial larynx (EL) is one of the alternative vocalizations. When the vibrator machine is pressed against the lower part of the jaw, vibration sound is transmitted in the sound path. This speech method is easy to learn, but this voice sound is of bad quality because it has no intonation and it is noisy.To Improvements of EL speech quality have been divided into two parts: improving the sound quality of the EL device by applying different noise reduction algorithms to reduce the radiated and the additive noise, and implementing a pitch-control function to the EL with advanced technology.

Keywords— Larynxes, EL, artificial Larynx, alternative vocalization

1. INTRODUCTION

Speech is for conveying some kinds of information or meaning, it is essential for our life. However, those who have lost their vocal folds due to diseases such as laryngeal cancer and injuries due to traffic accidents cannot communicate using voice. Such people are called the laryngectomee or a laryngeal patient. The larynx must be removed. In laryngectomy surgery, the larynx, where the respiratory and digestive systems intersect, is removed. Humans produce sound by utilizing the voice box which is called larynx. The voice confinement is arranged in throat at the highest point of the windpipe. The human voice box contains two tendons known as vocal cords. Vocal cords present over the larynx stretch so that it leaves a limited space between them for the entry of air. At the point when the human talks, muscles present in the larynx get extended and the opening becomes smaller. When air is made to go through the cut, the vocal lines vibrate. As vibrations increases, higher intensity of sound is delivered. Total laryngectomy is the elimination of larynx by surgery. People who have lost their larynx lack the ability to produce sound through the normal sound production mechanism. [1]

The electrolarynx has some benefits compared with other approaches. Electrolarynx is a small unit, powered by battery and hand held. The gadget's vibrating coupler plate is held opposite the neck. The coupler circle vibration or signal is transmitted into the vocal tract which routes the sign along those lines to generate sound. The vibrations are produced by a vibrator powered by an external battery (known as an electrolarynx or an artificial larynx) that is usually placed on the cheek or under the jaw. It creates a humming vibration that reaches the user's throat and mouth. The person then uses his / her mouth to modify the sound to articulate the sounds of speech, since the frequency of the EL's vibration is constant; the EL voice is flat and noisy like a buzzer sound. [2]

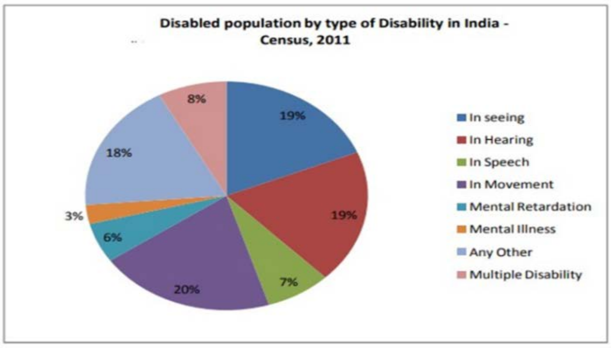

Fig1.diasbled population by type of disability in india census,2011

In a word, there are about 600,000 people whose larynxes have been removed. In India also every year more than 17000 peoples suffered from Laryngeal cancer so these people lose their larynx.

.

Fig2: Larengeal cancer patients India compared to word.

These people are not able to make voice, so they have problems in communication. However, they can move the mouth and tongue normally. Therefore, they speak using alternative vocalizations. Electrical artificial larynx(EL) is one of the alternative vocalizations.

EL speech identification and following conversion to normal speech makes the EL speech more intelligible, some of the noise filtering algorithm and wave path identification algorithm has been used to convert EL speech to normal speech.

2. PROBLEM STATEMENT

Some of the laryngectomies mentioned the fallowing problems they are facing: -They cannot communicate and they will communicate by writing or gestures alone. If EL communication is not good, it will lead to the reluctance and the avoidance of contact with friends and acquaintances. Some of the laryngectomy people told they spend their daily life alone; they are not able to participate in any activity effectively. The speech from the existing EL machine is not pleasant to hear, which is very noisy and not clearly audible.

3. METHODOLOGY

This device consists of a DSP processor, analog to digital converter, amplifier circuit, DAC circuit, and speaker. The speech from the EL machine is sampled by using the A-D converter and it passed to the DSP processor. The output from the EL machine consists of a lot of noise which can be removed by using the noise reduction algorithms in Matlab software.

Then it passes through the speech enhancement algorithm to control the pitch in the Matlab software. DAC is used to convert the digital data from the DSP processor to the Analog form. Then this data can be amplified and passes to the speaker complete action occurs in real- time.

4. BLOCK DIAGRAM

The below fig shows the Block diagram of the system

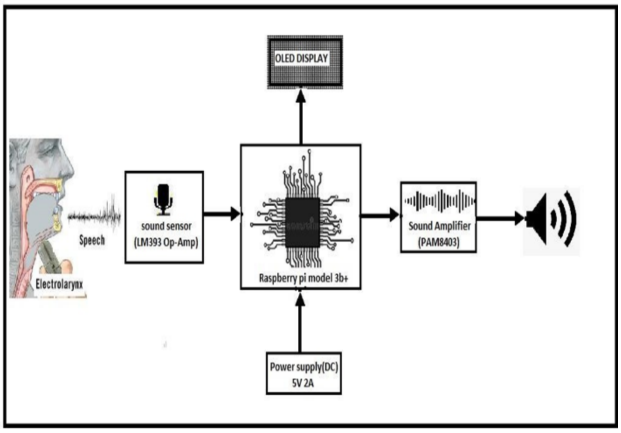

Fig3:Block diagram of proposed system.

Analog to digital conversion can be achieved with the help of sampling techniques. The sampled data send over to the DSP processor for processing. it has some noise reduction and pitch control tools to improve the EL speech. The processing data will be converted to analog data when it passes through the DAC circuit, then it is amplified and passed to the speaker. Electrolarynx also called as “Throat Back” is a medical device about the size of a small electric razor used to produce clearer speech by those people who have lost their voicebox, usually due to cancer of the larynx. The electrolarynx works by inducing vibrations of oral or pharyngeal mucosa by an external device, generally, at a constant fundamental frequency. The choice of device is dependent on anatomical factors and patient preference.

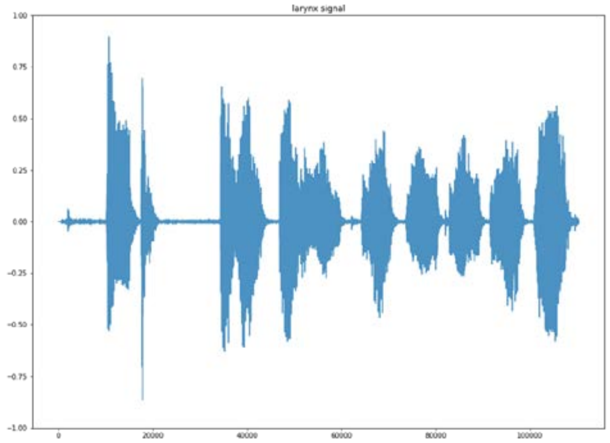

Fig4:Sampled EL speech.

The speech produced by the voice box in the form of wave is sensed by a sound sensor[LM393 Op-Amp]. Then the LM393 Op-Amp senses the sound and provides an easy way to detect sound and is generally used for detecting sound intensity. This module can be used for security, switch, and monitoring applications. Its accuracy can be easily adjusted for the convenience of usage. The above Waveform shows the electrolarynx speech waves which then gets detected by a sound sensor LM393. Sound detection sensor module detects the intensity of sound where sound is detected via a microphone and fed into an LM393 op-amp. It comprises an onboard potentiometer to adjust the setpoint for sound level.

5. HARDWARE REQUIREMENTS.

The fallowing Hardware components are used to build a prototype of the proposed system .

SL.NO	COMPONENT	SPECIFICATION	QUANTITY
01	Raspberry pi 3	B+	01
02	Microphone	LM383	01
03	Amplifier	PAM8403	01
04	Speaker	8 Ohm	01
05	Li-Ion battery	2500mah	01

Fig5.List of Hardware devices.

5.1. Raspberry pi 3b+

Rasbery Pi is a series of small single-board computers developed in the United Kingdom by the Raspberry Pi Foundation. Several generations of Raspberry Pis have been released. All models feature a Broadcom system on a chip (SoC) with an integrated ARM- compatible central processing unit (CPU) and on-chip graphics processing unit GPU).

Processor: Broadcom BCM2837B0, Cortex-A53 64-bit SoC @ 1.4GHz, Memory: 1GB LPDDR2 SDRAM, Connectivity: 2.4GHz and 5GHz IEEE 802.11.b/g/n/ac wireless LAN, Bluetooth 4.2, BLE, Gigabit Ethernet over USB 2.0 (maximum throughput 300Mbps)4 × USB 2.0 ports. Access: Extended 40-pin GPIO header.Multiedia: H.264, MPEG-4 decode (1080p30); H.264 encode (1080p30); OpenGL ES 1.1, 2.0 graphics.

S D Card Support: Micro SD format for loading operating system and data storage. Input Power:5V/2.5A DC via micro USB connector ,5V DC via GPIO header ,Power over Ethernet (PoE)–enabled (requires separate PoE HAT). Environment: Operating temperature, 0–50°C.

5.2. Audio Amplifier(PAM8403)

An amplifier, electronic amplifier or (informally) amp is an electronic device that can increase the power of a signal (a time-varying voltage or current). It is a two- port electronic circuit that uses electric power from a power supply to increase the amplitude of a signal applied to its input terminals, producing a proportionally greater amplitude signal at its output.

- Dual channel stereo output 3 w + 3 w power Class D
- Works with 2.5V-5v power supply
- High amplification efficiency 85%
- can directly drive 4 Ω/8 Ω small speakers
- Good sound quality & noise suppression
- Unique without LC filter class D digital power board
- Can use computer USB power supply directly
- Small Size, 1.85 x 2.11 cm can easily fit in a variety of products.

Then the converted and amplified speech is pronounced out with the help of speakers. Without the disturbing and unwanted voices.

5.3.Mic Transducer(LM393)

A microphone, colloquially named mic or mike is a device a transducer – that converts sound into an electrical signal. Several types of microphone are used today, which employ different methods to convert the air pressure variations of a sound wave to an electrical signal.

- Operating voltage 3.3V-5V
- Operating current : 4-8mA
- sensitivity (1Khz): 52-48dB
- Impedance: 2.2KΩ
- Frequency: 16-20Khz

5.4.Speaker(1W)

Wireless speakers are loudspeakers that receive audio signals using radio frequency (RF) waves rather than over audio cables. Wireless speakers are composed of two units: a main speaker unit combining the loudspeaker itself with an RF receiver, and an RF transmitter unit. The transmitter connects to the audio output of any audio devices such as hi- fi equipment, televisions, computers, MP3 players, etc. An RCA plug is normally used to achieve this. The receiver is positioned where the listener wants the sound to be, providing the freedom to move the wireless speakers around without the need of using cables.

- The speaker with resistance of 8 ohm.
- power rating equals to 1W.

5.5.Li Ion Battery

A lithium-ion battery or Li-ion battery is a type of rechargeable battery. Lithium-ion batteries are commonly used for portable electronics and electric vehicles and are growing in popularity for military and aerospace applications. In the batteries, lithium ions move from the negative electrode through an electrolyte to the positive electrode during discharge, and back when charging.

- Capacity :2600mAh
- Voltage:3.7v
- Max charge voltage:4.7v
- Standard charging current:0.5 Amps
- Max charging current : 1 Amps

5.5.OLED Module

An organic light-emitting diode (OLED or organic LED), also known as organic electroluminescent (organic EL) diode, is a light-emitting diode (LED) in which the emissive electroluminescent layer is a film of organic compound that emits light in response to an electric current. This organic layer is situated between two electrodes typically, at least one of these electrodes is transparent. OLEDs are used to create digital displays in devices such as television screens, computer monitors, portable systems such as smartphones, handheld game consoles and PDAs.

- 128*32 high resolution
- OLED Driver IC: SSD1306
- Working temperature: -30°C ~ 70°C
- Interface: I2C
- The display is self-illuminating

6. SOFTWARE REQUIREMENTS

The fallowing software full fill our software requirements to build proposed system:

6.1. Matlab Software

MATLAB (an abbreviation of "matrix laboratory") is a proprietary multi- paradigm programming language and numerical computing environment developed by MathWorks. MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages. Although MATLAB is intended primarily for numeric computing, an optional toolbox uses the MUPAD symbolic engine allowing access to symbolic computing abilities. An additional package, Simulink, adds graphical multi-domain simulation and model-based design for dynamic and embedded systems. As of 2020, MATLAB has more than 4 million users worldwide. MATLAB users come from various backgrounds of engineering, science, and economics.

MATLAB can call functions and subroutines written in the programming languages C or Fortran.[39] A wrapper function is created allowing MATLAB data types to be passed and returned. MEX files (MATLAB executables) are the dynamically loadable object files created by compiling such functions. Since 2014 increasing two-way interfacing with Python was being added.

6.2.Raspbian OS (2020)

Raspberry Pi OS (formerly Raspbian) is a Debian-based operating system for Raspberry Pi. Since 2015 it has been officially provided by the Raspberry Pi Foundation as the primary operating system for the Raspberry Pi family of compact single- board computers. The original Raspbian OS was created by Mike Thompson and Peter Green as an independent project. The initial build was completed in June 2012.

Previous Pi OS versions have been 32bit and based on Raspbian core, taking the name Raspbian. Since recent 64bit versions no longer use the Raspbian core, the name has been changed to Raspberry Pi OS for both 64bit and 32bit versions. As of 1 August 2020, the 64-bit version is a beta and is not suitable for general use

.

6.3.Python:

Python is an interpreted, high-level and general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant whitespace. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects.

Python is dynamically typed and garbage-collected. It supports multiple programming paradigms, including structured (particularly, procedural), object-oriented and functional programming. Python is often described as a "batteries included" language due to its comprehensive standard library.

Python 3.0, released in 2008, was a major revision of the language that is not completely backward-compatible and much Python 2 code does not run unmodified on Python 3. With Python 2's end-of-life, only Python 3.6.x and later are supported, with older versions still supporting e.g. Windows 7 (and old installers not restricted to 64-bit Windows).

7. ALGORITHM

The electrolarynx artificial speech is extracting from the condenser microphone. Condenser (or capacitor) microphones are commonly used in studios to pick up sounds with great detail and accuracy. This is accomplished with a lightweight membrane (referred to as the diaphragm) suspended by a fixed plate. Sound pressure against the diaphragm causes it to move, which in turn creates electrical output.

continuous time signal can be represented in its samples and can be recovered back when sampling frequency fs is greater than or equal to the twice the highest frequency component of message signal. i. e. fs≥2fm.where fs is sampling frequency and fm is audio signal frequency. We sample the signal g(t) instantaneously at a uniform rate of fs once every Ts sec. Thus, we can write:

$$g_\delta(t) = \sum_{n=-\infty}^{\infty} g(nT_s)\delta(t - nT_s),$$

where gδ(t) is the ideal sampled signal and where δ(t-nTs) is the delta function positioned at tim

t = nTs.

Digital audio is the system in which we store, recreate, and manipulate audio information in a computer system. Certain characteristics of an analog sound wave, like the frequency and amplitude, are converted to data computer software can read. This allows us to manage, edit, and arrange audio in a software-based context. The system takes these measurements at a speed called the audio sample rate, measured in kilohertz. The audio sample rate determines the range of frequencies captured in digital audio. In most DAWs, you'll find an adjustable sample rate in your audio preferences. This controls the sample rate for audio in our project. sampling period is given by T≥1/2W . The most common audio sample rate you'll see is 44.1 kHz, or 44,100 samples per second. This is the standard for most consumer audio, used for formats like CDs.

Fig4:Algorithm Steps 1 To 9

We are converting the audio segment data into numpy array because Arrays are a collection of elements/values, that can have one or more dimensions. ... NumPy arrays are called ndarray or N-dimensional arrays and they store elements of the same type and size. It is known for its high-performance and provides efficient storage and data operations as arrays grow in size. NumPy is very useful for performing mathematical and logical operations on Arrays. It provides an abundance of useful features for operations on n-arrays and matrices in Python. Fast Fourier Transformation(FFT) is a mathematical algorithm that calculates Discrete Fourier Transform(DFT) of a given sequence. The only difference between FT(Fourier Transform) and FFT is that FT considers a continuous signal while FFT takes a discrete signal as input. DFT converts a sequence (discrete signal) into its frequency constituents just like FT does for a continuous signal. In our case, we have a sequence of amplitudes that were sampled from a continuous audio signal. DFT or FFT algorithm can convert this time-domain discrete signal into a frequency-domain.

Fig5:Alogorithm steps 10 to 18.

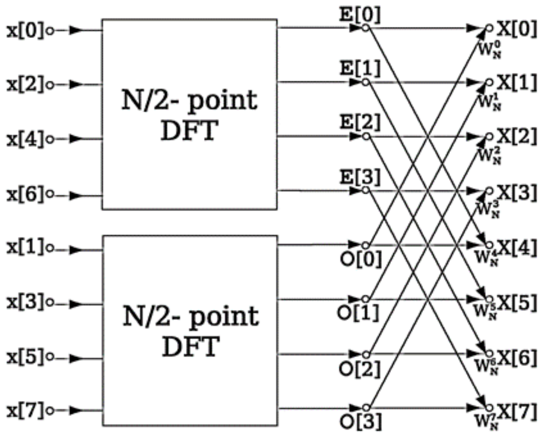

Fig6: FFT algorithm overview.

The FFT is a fast algorithm for computing the DFT. If we take the 2-point DFT and 4-point DFT and generalize them to 8-point, 16-point, ..., 2r -point, we get the FFT algorithm. To compute the DFT of an N-point sequence using equation (1) would take O(N2) multiplies and adds. The FFT algorithm computes the DFT using O(N log N) multiplies and adds. There are many variants of the FFT algorithm. We'll discuss one of them, the "decimationin-time" FFT algorithm for sequences whose length is a power of two (N = 2r for some integer r). Above is a diagram of an 8-point FFT, where W = W8 = e−iπ/4 = (1 − i)/√2.

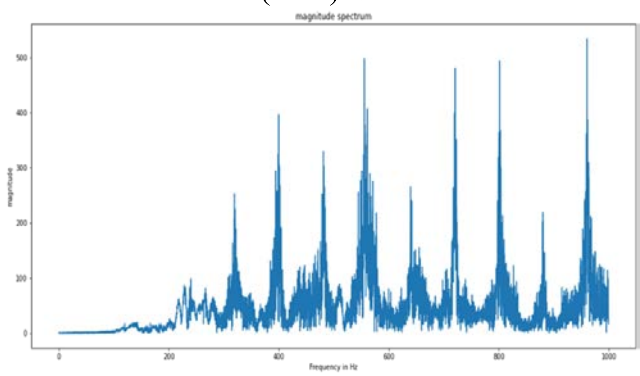

Fig6.Magnitude spectrum of the Speech signal.

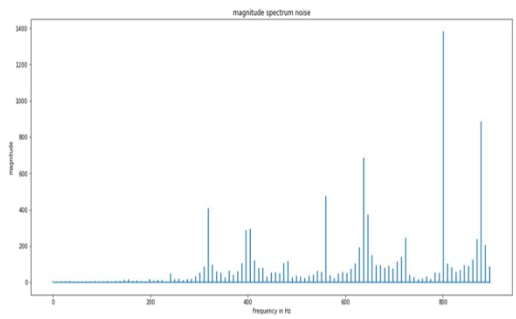

Fig7.Magnitude spectrum of the noise signal.

Removing the unwanted noise component on the using the bandpass filter and capturing the amplitude envelope of the audio signal.

Inverse Fast Fourier transform (IDFT) is an algorithm to undoes the process of DFT. It is also known as backward Fourier transform. It converts a space or time signal to signal of the frequency domain. The DFT signal is generated by the distribution of value sequences to different frequency component. Working directly to convert on Fourier transform is computationally too expensive. So, Fast Fourier transform is used as it rapidly computes by factorizing the DFT matrix as the product of sparse factors. As a result, it reduces the DFT computation complexity from O(N2) to O(N log N). And this is a huge difference when working on a large dataset. Also, FFT algorithms are very accurate as compared to the DFT definition directly, in the presence of round-off error.

This transformation is a translation from the configuration space to frequency space and this is very important in terms of exploring both transformations of certain problems for more efficient computation and in exploring the power spectrum of a signal. This translation can be from xn to Xk. It is converting spatial or temporal data into the frequency domain data.

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(x)e^{-ixt}\,dx$$

8. RESULT AND CONCLUSION

Theoverall performance of the device reaches our expectations.we successfully extract 86% noise-free signal from the EL speech signal. The quality of the connection between the Bluetooth audio receiver and our device is also good.

The below shown spectrogram shows the audio signal coming out of the electro larynx device which contains both the original message signal and the noise. This audio signal is then given to the analog to digital converter to convert it to the samples. This digital audio is stored in the system which is recreated and manipulated in the computer system. Certain characteristics of the analog sound wave like the frequency and amplitude are converted to the data that computer software can read. This audio segment data is converted to the numpy array collection of element less values. This audio signals noise is separated through computer languages such as Mat lab and Python.

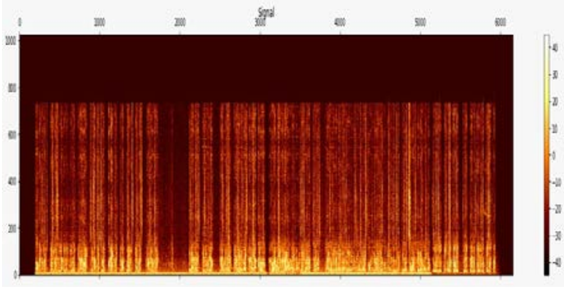

Fig8.Output from amplifier

Here the output of each stage is represented in terms of a spectrogram A spectrogram is a visual representation of the spectrum of frequencies of a signal as it varies with time.When applied to an audio signal, spectrograms are sometimes called sonographs, voiceprints, or voicegrams.

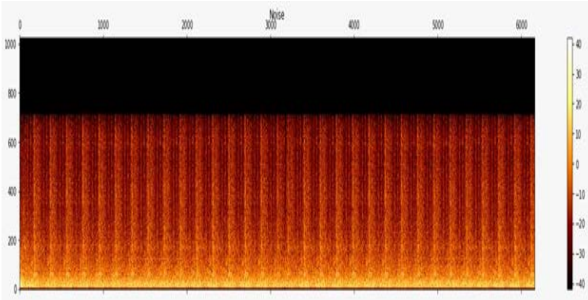

Fig9.Noise signal

This above shown spectrogram represents the range of frequencies of the noise signal that is present in the audio signal that is been captured from the electro larynx. In this above shown spectrogram the vertical and the horizontal axis represent the frequency in herds and time respectively.

The collection of signals that are currently blocked is called the signal mask. Each process has its own signal mask. When you create a new process (see Creating a Process), it inherits its parent's mask. You can block or unblock signals with total flexibility by modifying the signal mask. The unmasked threshold is the quietest level of the signal which can be perceived without a masking signal present. The masked threshold is the quietest level of the signal perceived when combined with a specific masking noise. The amount of masking is the difference between the masked and unmasked thresholds.

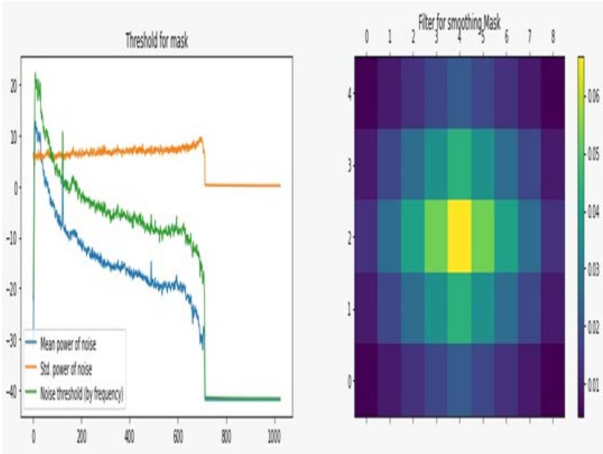

Fig9.The Mask applied Data

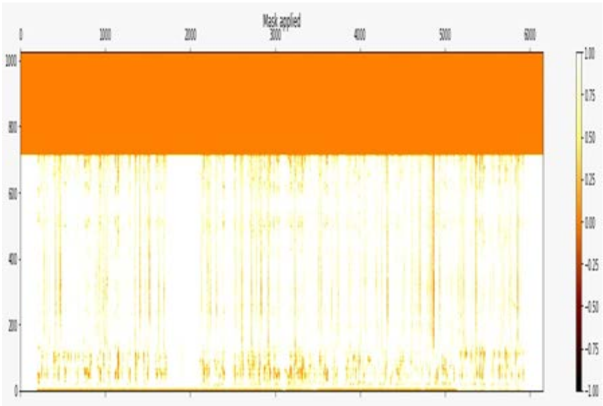

Fig10.spectrogram of mask

If two sounds of two different frequencies are played at the same time, two separate sounds can often be heard rather than a combination tone. The ability to hear frequencies separately is known as frequency resolution or frequency selectivity.

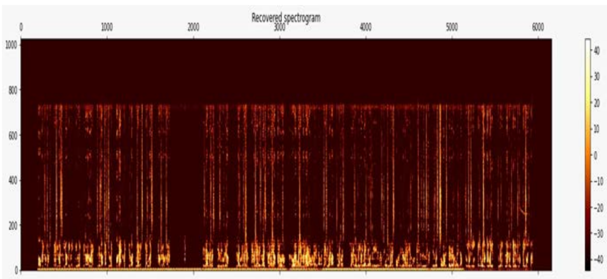

Fig11.Spectrogram of extracted signal.

The above is the output signal that is extracted by removing the noise signal from the signal captured from the electro larynx. The noise is filtered out and the envelope of the audio signal is extracted. FFT algorithms are used in order to convert the signal into frequency domains and extract its amplitude so that the noise from the captured amplitude of the signal is removed.

The following signal is the output that is coming out of the created device.

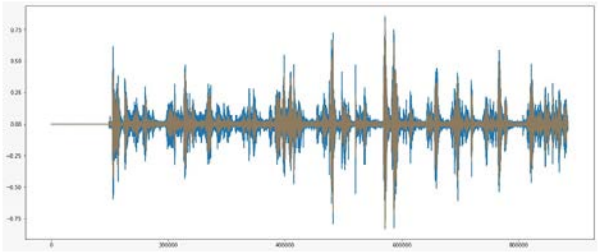

Fig.12.Final output waveform

9. FUTURE SCOPE

- The overall performance of the device reaches our expectations.
- The overall performance of the device reaches our expectations.

10. REFERENCES.

[1] Etsuro Urabe, Rin Hirakawa, Hideaki Kawano and Kenichi NakashioshihisaNakatoh, Electrolarynx System Using Voice Conversion Based on WaveRNN, IEEE International Conference on Consumer Electronics , ICCE , 2020.

[2] Urabe, Rin Hirakawa, Hideki Kawano, Kenichi Nakashi and Yoshihisa Nakatoh, Enhancement of Electrolarynx speech based on WaveRNN. ACIT, May, 2019.

[3] R. Nandana, Nissie Mary Johnson, V. Jubily. B.S. Reji, Shreelakshmi, Ancy Anselam, Lani Rachel Mathew International Journal of Engineering Research & Technology (IJERT) Electrolaryngeal SpeechIdentification using GMM , Vol. 9 Issue 06, June 2020.

[4] Hanjun Liu, M. L. NgAuris Nasus Larynx 34 Electrolarynx in voice rehabilitation 2007.

[5] Ramya Devi Ra , D Pugazhenthib Quaid-E-Millath Ideal Sampling Rate to reduce distortion in Audio Steganography Government College for Women, Chennai , Tamil Nadu, India.

[6] Ali Al-Haj A dual transform audio watermarking algorithm Springer Science+Business Media New York 2013.

[7] Ajeet Thakur , Rajesh Mehra and Rajesh Kumar , Computationally Efficient Sampling Rate Converter for Audio Signal Application Department of Electronics and Communication Engineering, NITTR , Nov 2011.

[8] Jean Jiang Audio Processing with Channel Filtering using DSP Techniques Department of Engineering
Technology Purdue University Northwest Westville, IN, USA ,2013 .

[9] Maximo Cobos Sandra Roger SART3D: A MATLAB toolbox for spatial audio and signal processing , sept 2011 .

[10] Nivea Sharma, Meghna Prabhu,Simran Kaur and Dr. Reena Kumbhar Implementation of Sound Effects Audio Signals. SARDAR PATEL institute of technology,June 2012.

[11] L. David William Raj1 , K. Santhosh, S. Subash, C. Sujin and P. Tharun Voice Controlled Door Lock System Using Matlab and Arduino , Feb 2001.

[12] Pitch Shifter Jake Garrison and Jisoo Jung 443 DSP Capstone 6 June 2016.

[13] https://cutt.ly/mjIE0zX accessed on 07-11-2020.

[14] https://cutt.ly/7jIE991 accessed on 07-11-2020.

[15] https://cutt.ly/7jIE991 accessed on 07-11-2020.

[16] https://cutt.ly/hjITWn7 accessed on 07-12-2020.

[17] https://cutt.ly/YjIUY8X accessed on 09-12-2020.

[18] https://ph.element14.com/buy-raspberry-pi accessed on 09-12-2021.

[19]https://www.dnatechindia.com/image/cache/catalog/pam8403-500x500.jpg accessed on 09-01-2021.

[20] https://bit.ly/2LI2zbV accessed on 09-01-2021.

[21] https://bit.ly/3o0ZsZW accessed on 09-01-2021.

[22] https://bit.ly/2LGYwww accessed on 09-01-2021.

[23] https://bit.ly/3iksmms accessed on 09-01-2021Several