SEMARAK ILMU
PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage: https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Autoregressive Based Vocal Tract Shape Modelling of Vowels in Speech Processing

Balakrishnan Sivakumar[1], Praveen Kadakola Biligirirangaiah[2,*]

[1] Department of ETE, Dr. Ambedkar Institute of Technology., Bangalore, India
[2] Department of Electronics Engineering, PRIST University, Thanjavur., Bangalore, India

ARTICLE INFO

Article history:
Received 16 April 2023
Received in revised form 28 June 2023
Accepted 28 July 2023
Available online 21 August 2023

Keywords:
Speech Processing; Pitch; vocal; vowels; Acoustic

ABSTRACT

In the present research article, which concentrates on the estimation of the vocal tract shape for the five vowels of Indian English in the current growing interest in the field of speech processing, with practical limitation of the data that can be collected and analysed. Speakers without the degree of control on articulators cannot produce the desired data. Here, we aim to design an approximate phonatory model by improving the present approach for acoustic calculation with a fully aerodynamic simulation, explicitly accounting for the propagation of sound along the tract, generating patterns of movement of simulated vocal tract articulators, and specifying the temporal relations among dynamically defined gestures that lead to a time-varying vocal tract filter function, and an acoustic waveform. The principle of phonetic distinctiveness and formants frequency spread is one of the crucial parameters that are measured and compared on intra, inter-speaker in actual speaker identification & forensic applications.

1. Introduction

For the formal communication in daily life, speech is the most important aid for expressing the words and needs of all the human beings. Hence it is considered as the most information rich full content, these speech signals can be upgraded with multi-layered differentiation of the spectral features depending on the kind of speech we are expressing. In a normal speech which seem to be ordinary speech is duly embedded with the intonation, accent, age of the speaker, kind of speaking, different expression, health condition of the speaker and intention of communication including the gender of the speaker [1].

Speech signals are generated conventionally by means of the vibrations generated with the atmospheric pressure representation by means of movement of vocal cords there by vocal tract by pushing the encountered air through lung cavity via nostrils. The generated air modulated with the prominent vibrations by the combination of the closure and opening of the mouth [2].

* *Corresponding author.*
E-mail address: praveenkb.et@drait.edu.in

https://doi.org/10.37934/araset.32.1.3245

2. The Speech Chain

Figure 1 shows the entire method of manufacturing and perceiving speech from the formulation of a message within the brain of a speaker, to the creation of the speech signal, and eventually to the understanding of the message by attender. In their classic introduction to speech science, Dense and Pinson capably cited this method because of the "speech chain". The method starts within the higher left as a message drawn somehow within the brain of the speaker. The message data may be thought of as having variety of various representations throughout the method of speaking (the higher path in Figure 1) [4].

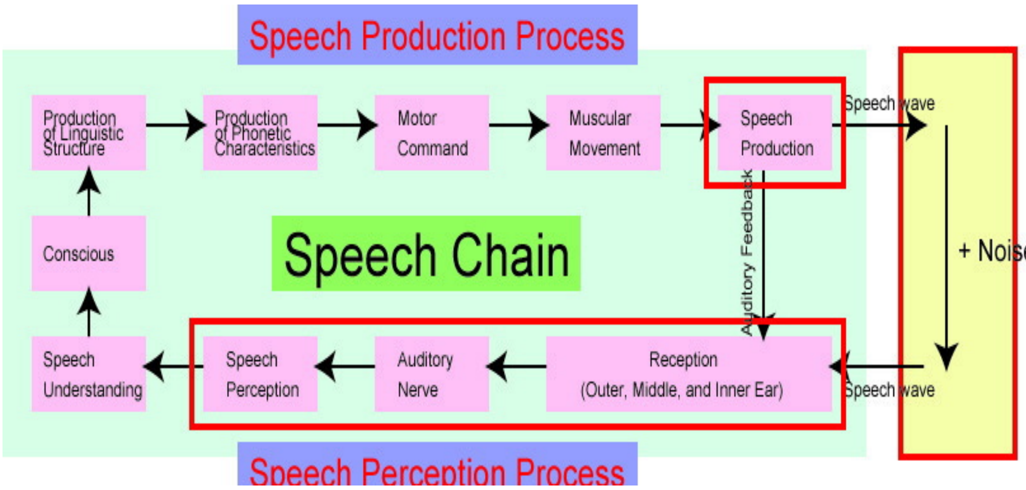

Fig. 1. The Speech Chain: from message to speech signal, to understanding

The complete speech chain comprises of a speech production/ generation model, of the sort mentioned higher than, further as a speech perception/recognition model, as shown planning to the left within the bottom half-part Figure 1. The sound perception model shows the series of steps from capturing speech at the ear to understanding the message encoded within the speech signal [5].

3. Problem Statement

The vocal tract estimation for any speech flag of any speaker is a testing issue in speech preparation. There is some proof that the phonetic peculiarity and speaker personality are profoundly instilled in the vocal tract module that is evaluated from the vowels utilizing the parameters of speech like formants, pitch, power, and zone work approximation.

Our proposed look into utilizes a vocal tract show and decides the arrangement of vocal tract module for same vowel expression, at various interim of times in various conditions. These diverse conditions bring about variety of acoustic highlights will show up in speech because of progress of status of the speaker, feeling, speaker wellbeing and fundamentally speech condition [6].

From the writing study it plainly demonstrates that speaker variety brings about the spread of ghostly amplitudes, pitch, formant recurrence and so forth they are normally described by wellspring of voice and channel, higher formants convey speaker particular data, Gross vocal tract module estimation from the lower formants causes the biggest spread among vowels and the individual speaker [7].

The consistent state vowels of grown-up male and female speaker are recorded at normal conditions. The changeability of the subsequent vocal tract module and formant spread as for formant recurrence proportion with the contiguous recurrence are measured on intra speaker premise. To explore the inconstancy of vocal tract module concerning length and tube number [8].

4. Objectives

The prime objective is to contemplate the variety in vocal tract shapes and formants spread for the male and female speakers, unwavering quality and set up the utilization of this variety in vocal tract shapes and formants. Additionally, it is utilized to show the performance vowels, non-logically, getting from the model spectrograms, formants, pitch, and vocal tract shape data. The examination is improved the situation a vowel in the above arrangement, to get the vocal tract shape for the vowel /a/of the male speaker by taking 30 tests of 30 subjects at various circumstances.

The primary destinations of our examination targets are:
(1) To assess the vocal tract shape for the different speakers for the vowels talked by Indian English speakers.
(2) To check the impact of formants and at last for vowel union process. These outcomes are extremely valuable for actualizing the speech acknowledgment frameworks.

5. Database

We have made a voice database with various south Indian male and female speakers. We have record by speakers mimicking all the five different Indian English vowels. Amid recording similar vowels are recorded under similar conditions it is expressed. The recorded examples are done under typical room condition, at the inspecting recurrence of 24 kHz, 16 bit utilizing mono chronicle gadget at the base separation of 10cm from mouth and spared as .wav arrange which will be useful for MATLAB condition [9].

5.1 The Mechanism of Speech Production

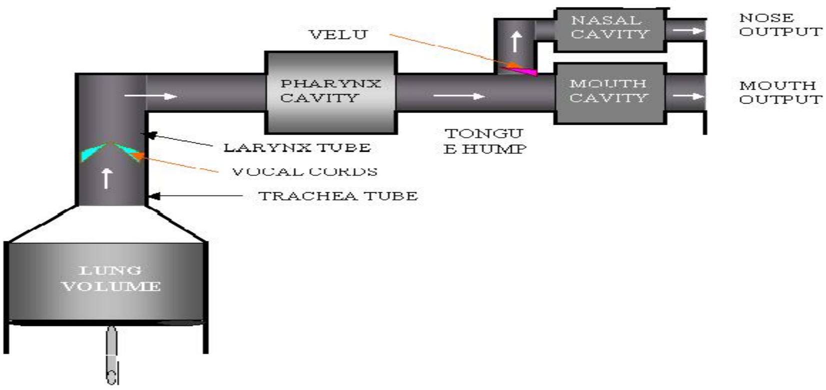

Fig. 2. The Mechanism of Speech Production

5.2 Lungs

The purpose of the lungs is the inhalation and exhalation of air. once we inhale, we tend to enlarge the cavum by increasing the skeletal structure encompassing the lungs and by lowering the diaphragm that sits at the lowest of the lungs and separates the lungs from the abdomen; this action lowers the atmospheric pressure within the lungs, so inflicting air to rush in through the vocal tract and down the trachea into the lungs. The trachea is typically spoken because the "windpipe," is a few one-2 cm long and 1.5–2 cm diameter pipe which matches from the lungs to the cartilaginous structure. The cartilaginous structure could be a tiny mass, or "switch," which, throughout swallowing and consumption, deflects food removed from coming into the trachea. Once we exhale, we tend to scale back the degree of the cavum by getting the muscles within the skeletal structure, so increasing the respiratory organ atmospheric pressure. This increase in pressure then causes air to flow through the trachea into the vocal organ [10].

5.3 Vocal Tract

The vocal tract is comprised of the mouth from the vocal organ to the lips and the nasal passage that's coupled to the oral tract by manner of the velum. The oral tract takes on many various lengths and cross sections by moving tongue, teeth, lips, and jaw and has a median length of seventeen cm during a typical man and shorter for females, and a spatially varying crosswise of up to 20cm2 [12].

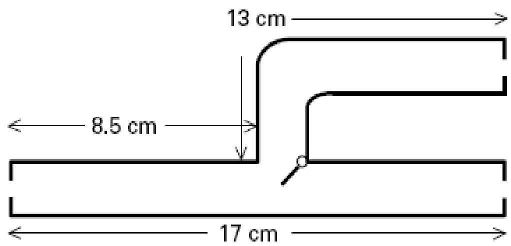

Fig. 3. Vocal Tract shape

A formant could be the concentration of acoustic energy around a specific frequency within the speech wave. There is area unit many formants, every at a distinct frequency, roughly one in every 1000Hz band. every formant corresponds to a resonance within the vocal tract. Formants modification with completely different vocal tract configurations [13].

Wakita has compared the LPC based assay adjustment with the lip actuation acknowledgment method. According to him, although different connected breadth action cannot be bent from a accent arresting due to its bandage limitation, a different detached breadth action can be acquired from band-limited speech. He has appropriate that articulate amplitude characteristics can be acquired from the accent articulation agnate to bankrupt glottis action for removing the aftereffect of alien antecedent characteristic [15].

He has approved that activity accident due to articulate amplitude bank accordance affect lower formant frequencies and appropriately by transforming the abstinent formant frequencies application a acceptable about-face chart, aftereffect of accident on appearance admiration could be taken affliction of. Even 0.5 cm change in the admiration of articulate amplitude breadth appreciably affects the estimated shape. Hence the breadth of the articulate amplitude has to be abstinent accurately either evidently or acoustically [3].

5.4 Linear Prediction and the Speech Model

Our appetite is to abstruse emphasis abuttals of the archetypal by appraisal of the emphasis signal. It is accustomed to acquire structures (or representations) for both the activity artist and the beeline system. One such archetypal uses an added abounding representation of the activity in acceding of absent anterior generators for authentic and blurred emphasis as credible in Figure 4.

The aloft archetypal is generally alleged the LPC Model. The archetypal says that the agenda accent arresting is the achievement of agenda filter. The aesthetics of beeline anticipation is carefully accompanying to the basal accent amalgam archetypal in which it was apparent that the sampled accent arresting was modelled as the achievement of a linear, boring time-varying arrangement aflame by either quasi-periodic impulses (during authentic speech), or accidental babble (during blurred speech) [4]. The beeline anticipation adjustment provides a able-bodied reliable and authentic adjustment for ciphering the ambit that characterize the linear, time-varying system. Over abbreviate time intervals, the beeline arrangement is declared by an all-pole arrangement function.

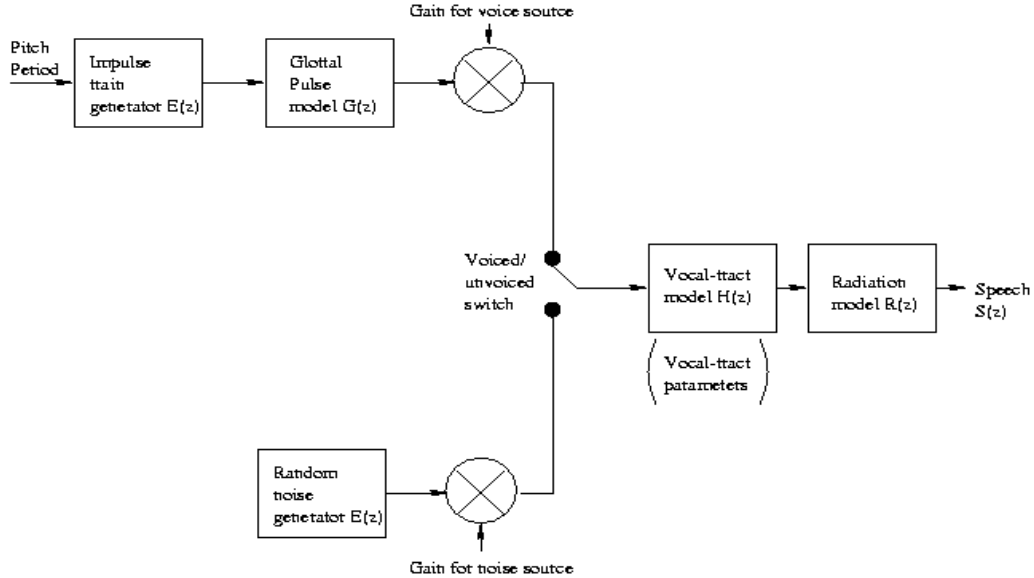

Fig. 4. Simplified model for speech production.

The blended spectrum furnishings of radiation, articulate tract, and glottal action are represented by a time-varying agenda clarify whose steady-state arrangement action is of the form.

$$H(Z) = \frac{S(Z)}{U(Z)} = \frac{G}{1 - \sum_{K=1}^{p} a_k z^{-k}} \quad (1)$$

The speech signal analysis for the vocal tract shape estimation can be broadly classified with the analysis and structured as shown in figure 5, which includes five different level of analysis such as speech signal analysis, pre-emphasis, hamming window technique, auto correlation method, parcor coefficients extraction and final estimation of the vocal tract shape [17].

5.5 Pre-emphasis

It should be acicular out that the breadth action acquired application absorption coefficients cannot be said to be the breadth action of the animal articulate tract. If pre-emphasis is acclimated above-mentioned to beeline predictive assay to abolish the furnishings due to the glottal beating and radiation, again the consistent breadth functions are generally actual agnate to articulate amplitude configurations that would be acclimated in animal speech. From the accent assembly archetypal it is accepted that the accent undergoes an ashen angle of -6dB/oct.

5.6 Window Technique

This technique of action is a selection of the speech segment which has to be selected from the running speech for the exact analysis with application of the w(k-n), where k defines the which particular speech for the selection. The ideal action of window is to access a large amount of information to simulate and analyze from the main lobe of information with increased resolution for the focus of the information. This action will differentiate the exact window in an fixed phenomena by adding particular weights and the process of adding the credits for the discrete speech access is termed as Gibbs phenomenon [18].

Various window operations have been considered for the speech segment analysis such as Hamming window, Bartlett, Blackmann, Hanning, Kaiser and etc, with the final processing of the data with the dependence from the outcomes with the dependable responses, in this present research work linear predictive method that processes the frames with the hamming window has been adopted. It is window size of 20 to 30ms which are spaced equally for the analysis. In the window system, it operates with the Discrete Fourier Transform (DFT) for multiple applications in the considered systems [19].

5.7Autocorrelation Method

Multiple parameters are normalized for the attainment of the linear predictive analysis using auto correlation method. It is one of the best methods to deal with the speech signal for the predictive analysis by addressing the issues related to the speech signal with the application of the Fourier transform and processing it with the Toeplitz matrix structure. In this method of analysis, it will be analysing the similarities with the existing speech frame and previous speech frame termed as inter and intra frame analysis.

5.8 PARCOR Coefficients

After the analysis from the autocorrelation method of speech analysis, it is required to normalize the reflection coefficients in the limited boundary ±1. in which the available speech signal spread over the large area will be limited to the minimal boundary by considering the reflection coefficients, acts as the quantization criteria in communication systems. These coefficients are modelled such a way that exactly the structure of the vocal tract and this vocal tract is the concatenated version of the irregular tubes depends on the amount of the air pressure and reflection coefficients with the usage of levinson Durbin algorithm for the estimation of the reflection coefficients.

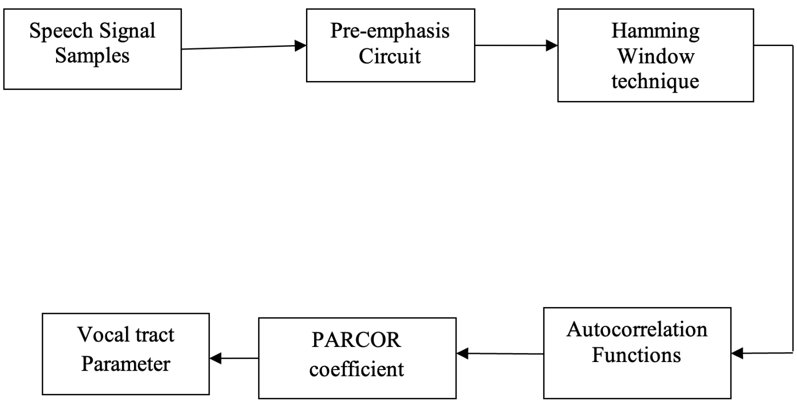

Fig. 5. Block diagram for vocal tract shape calculation

5.9 Concatenated tube model

The entire vocal tract module can be modelled as the concatenation of the irregular small tubes associated together as vocal tract as shown in figure 6. The cross-section view can be viewed as the 17cm for the male speakers and 15cm for the female speakers in a conventional method of measurement from the lips to glottis. Depending on the amount air pressure from the lungs, reflections will take place for the approximation of the vocal tract which includes different participants such as breath, jaws, tongue, lips, nostrils, and other parts of human body such as the articulators for the generation of the sound [20, 21].

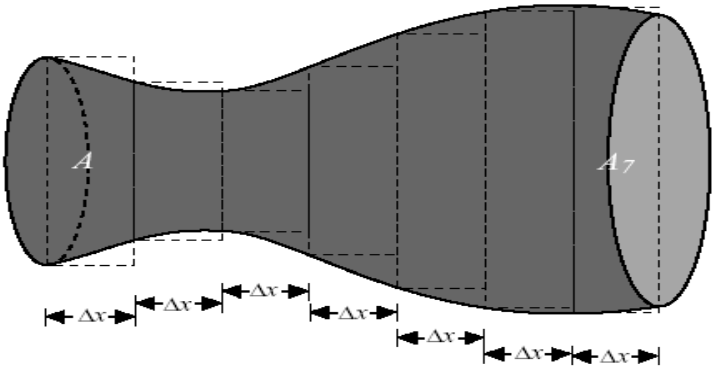

Fig. 6. Concatenated tube model

5.10 Pre-emphasis

The speech signal which is subjected to analysis by the area function with reflection coefficients. The area functions are calculated using the coefficients obtained with the reflection operations through the vocal tract. This estimation of the area function requires the preprocessing of the data in order to suppress the effects of the radiation at the lip end and glottal pulse which will result in the flattening of the response which results in the unusual response. It is highly essential to the deal with the pre-emphasis prior to the linear predictive analysis for the area function approximation in the subject under consideration, this pre-emphasis is set to process at 6 dB/octave, which results in the upscaling of the amplitude of 6 dB/octave. The resulting process of the pre-emphasized signal is represented in figure 7, shows that the speech signal is flat, and it is linear to estimation with reflection coefficients using LPC.

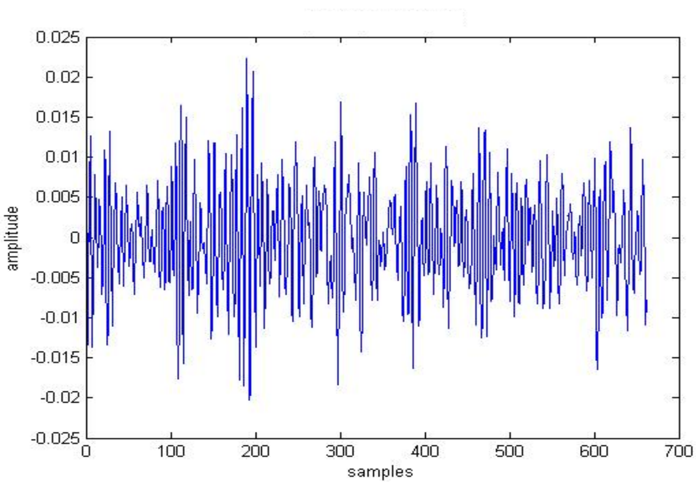

Fig. 7. Pre-emphasized Signal

5.11 Window Analysis

In the linear predictive analysis method, the linear prediction is carried out on the number of frames with the weighted sum using the hamming window, and the data frames are processed passed through the window w(n). this window w(n) is the good selection to deal with the attenuation at main lobe and side lobes and the windowing can be represented as follows:

$$w(n) = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{N}\right) & ;\ 0 \le n \le N-1 \\ 0 & ;\ otherwise \end{cases} \tag{2}$$

Figure 8 represents the resultant of the windowed signal speech waveform under consideration, it is observed clearly that when compared with figure 7 the major effects of the noise are minimized without disturbing the originality of the signal and its contents.

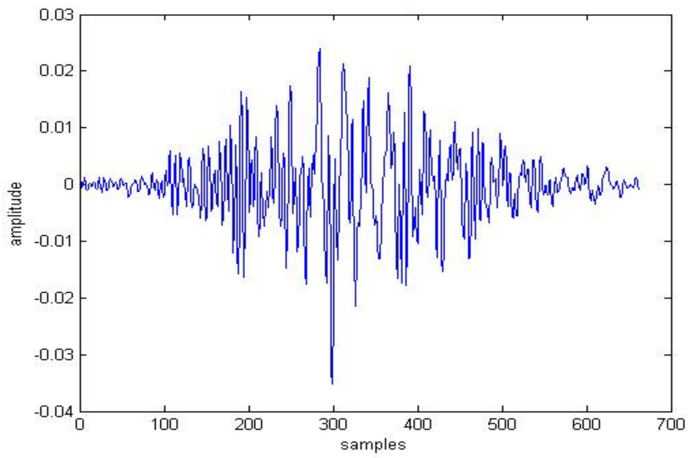

Fig. 8. Windowed signal

5.12 Voiced Signal

The voiced signal in the speech is considered as the signal embedded with the multiple information in single wave but requires the filtering operation to explore the features in it, during the exploring of the features filtering is one of the necessary operations. For the filtering operation the signal exited with sound source produced with the vocal chords. These generated signals are the combination of the voiced, unvoiced, and other possible noise. This filtering operation acts as the vocal tract filter for operation with the periodic signal for the selection of the voiced speech which are produced with the larynx, vocal folds, and glottis.

5.13 Noise Elimination of Voiced Signal

The deviation of the various amplitude levels depending on the various utterances for the same speech sample with the same speaker. To eliminate this variation in the utterance of speaker samples noise elimination is much needed for the analysis.

The sample output for the speech signal under consideration with the databases as mentioned in the initial para of the article, the estimation of the vocal tract module is done for all vowel samples under consideration with the undergoing of all different stages to carry out starting from the consideration of the speech signal to the vocal tract module estimation. To represent the various processes for operation of the speech signal, it is mentioned with figure 10 as the original speech signal, figure 11 represents the normalized speech of the waveform considered, figure 12 is used for the indication of the pre-emphasized speech signal.

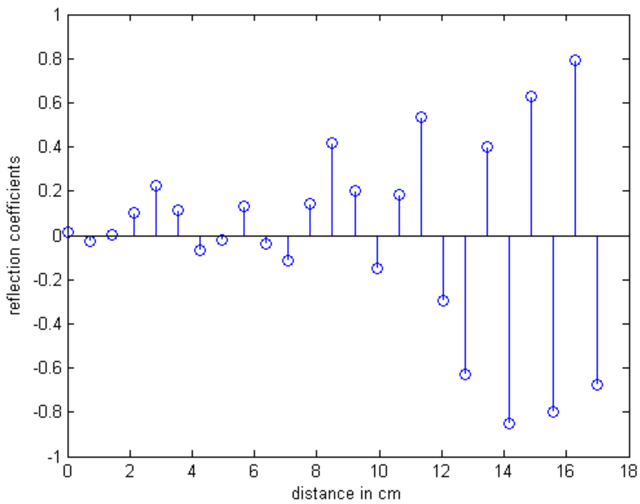

Fig. 9. Reflection (PARCOR) co-efficient for a frame

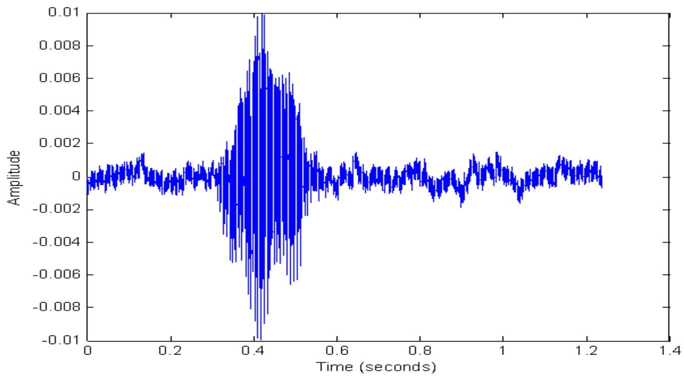

Fig. 10. Speech waveform for vowel /a/

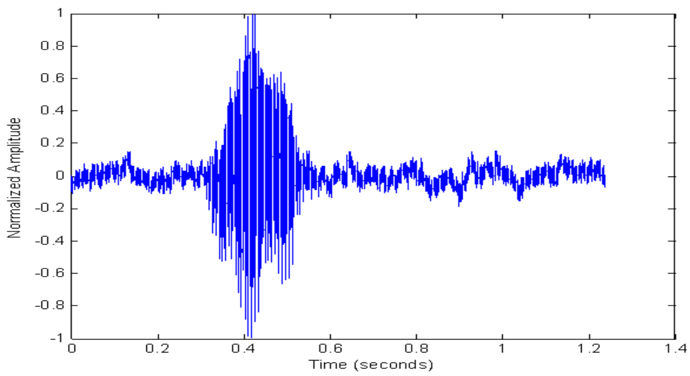

Fig. 11. Normalized speech waveform for vowel /a/

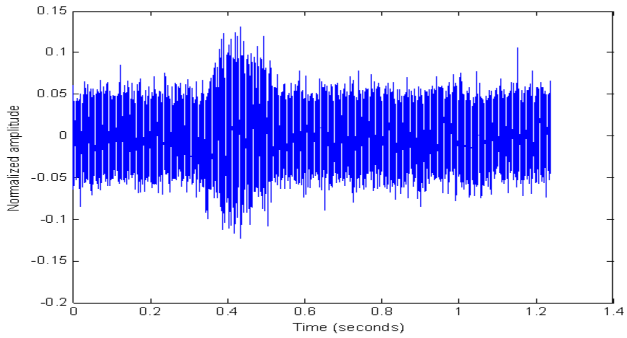

Fig. 12. The pre-emphasis signal for vowel /a/.

Figure 13 is dealing with the resultant filtered voiced signal, figure 14 is the continuation of the noise elimination voiced signal with suitable filtering technique, figure 15 is the representation of the reflection coefficients for the speech signal as par-cor coefficients. The final dynamic modelling of vocal tract module with the parcor coefficients is represented in figure 16 for the vowel /a/.

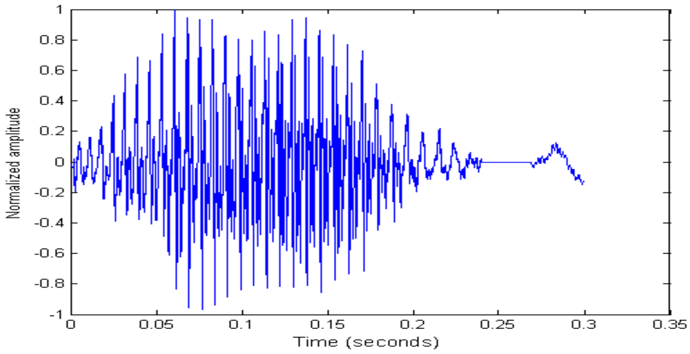

Fig. 13. The voiced signal for vowel /a/.

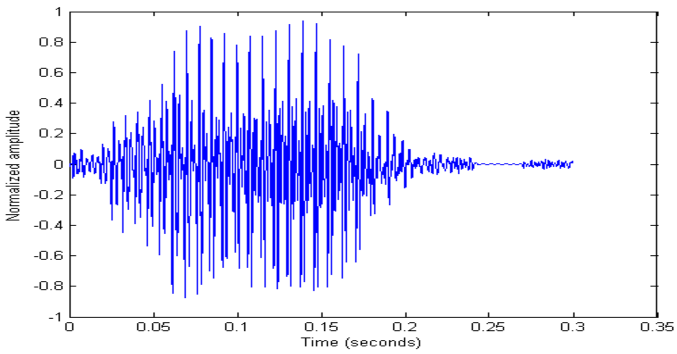

Fig. 14. The voiced signal with noise elimination for vowel /a/.

Similar process is carried out for all the vowels and vocal tract is modelled for the vowels and they are represented in figure 17, figure 18, figure 19 and figure 20 for the vowels such as /e/,/i/, /o/ and /u/ respectively.

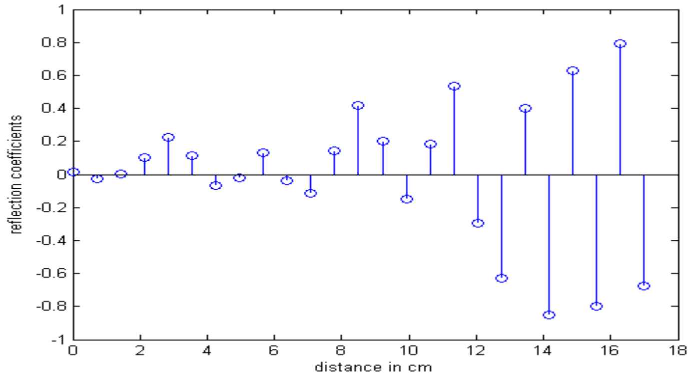

Fig. 15. Reflection co-efficients for vowel /a/

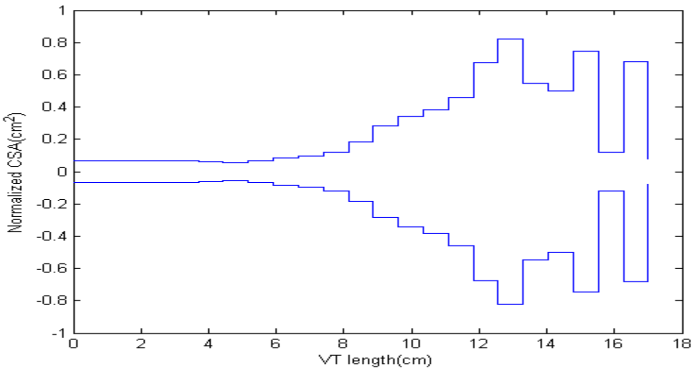

Fig. 16. Dynamic vocal tract model for vowel /a/

From Figure 16, it is observed that the male speaker sample under consideration hence the length of the vocal tract shape is 17cm and the representation normalized cross sectional area plot. As an observation it can be noted for the first vowel /a/ the vocal tract opening remain same for the lip end and glottis end, but the vocal tract module widens at the region measuring between the 12cm to 15cm and comes to same portion of opening at the lip closure.

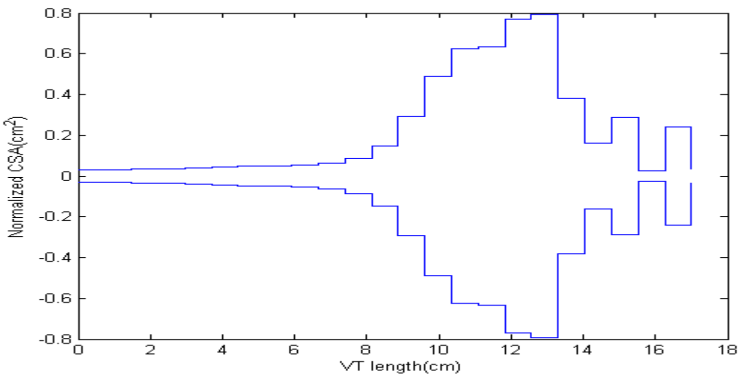

Fig. 17. Dynamic vocal tract model for vowel /e/

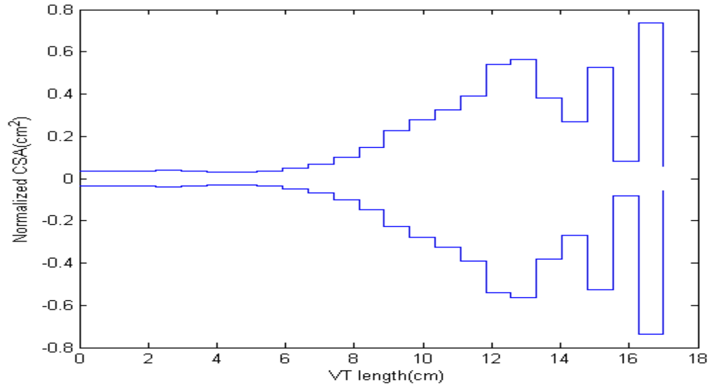

Fig. 18. Dynamic vocal tract model for vowel /i/

It can be observed with figure 17 and figure 18 that the vocal tract module for the vowel /e/ and vowel /i/ respectively. the analysis with figure 17 can be observed narrow opening at the glottis end and narrow closure at the lip end, maximum value of the cross-sectional area is achieved at the length 12 to 14cm thereby resulting in the sounding of the causal effect of the output sound. Whereas in figure 18 for the vowel /i/ the normalized cross-sectional area is minimum at the glottis end till the 10cm and gradually increased and comes to original position at 15cm region.

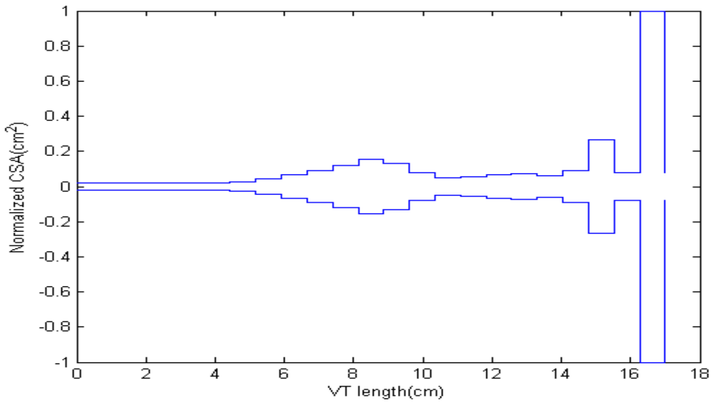

Fig. 19. Dynamic vocal tract model for vowel /o/

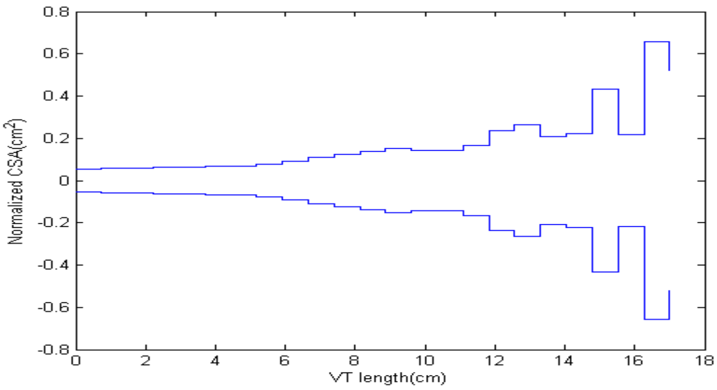

Fig. 20. Dynamic vocal tract model for vowel /u/

Figures 19 and 20 are dedicated to the vocal tract representation of the vowels /o/ and /u/ respectively. It is noted that the /o/ attains the maximum cross-sectional area at the lip end i.e.,17cm as it is a rounded vowel condition and reset all responses remain flat. The vowel /u/ represents the wide opening at the lip edn due to the unrounded vowel and narrow opening at the glottis.

Comparison of vowels normalized cross sectional area with distance of vocal track from the lip.

Table 1
Range in cm of vocal track widen distance measured from the lip

Sl no.	Vowels	Vocal tracks widen range
1	/a/	12-15cm
2	/e/	12-14cm
3	/i/	12-14cm
4	/o/	14-15cm
5	/u/	16-17cm

Table 1 shows the distance, in centimeters, at which the vocal tract becomes wider which highlights that the vowel /u/ has a wider range of widening compared to other vowels. This means that the distance at which the vocal tract becomes wider is greater for /u/ than for other vowels. This is because the tongue position for /u/ is further back in the mouth, which results in a larger opening at the back of the vocal tract.

The wider range of widening for /u/ is significant as it contributes to the distinctive sound of this vowel. The larger opening at the back of the vocal tract amplifies lower resonant frequencies, which gives /u/ its characteristic deep and rounded sound. In contrast, vowels with a narrower range of widening, such as /i/ or /e/, tend to have higher resonant frequencies and a more closed sound.

6. Conclusion and Discussion

Analysis of the vowels of the Indian English is done on the ambit of formant frequency, the articulate amplitude appearance is estimated from the absorption coefficients acquired application LPC based coefficient for every speech sample under consideration. The accent arresting is disconnected into frames of breadth 30msec with an overlapping of 10msec anatomy breadth with sampling amount of 22, 100 Hz with the LPC clarify adjustment as 14 in the assay anniversary anatomy consists of 'N' samples. The linear Predictive coding assay is done application AR adjustment to account the absorption coefficients which are as well accepted as PARCOR coefficients. The better absorption coefficients occur breadth the about changes in the formant frequency, the articulate amplitude appearance air headedness is greatest. The breadth action of a being taken altered times and altered ambience for the Indian English vowels (a, e, i, o, u) for both male and female speakers. Using Auto Regressive modelling of accent processing the formant frequencies are affected and by extracting PAR-COR coefficients from the accent database is considered. By autoregressive modelling of articulate amplitude appearance admiration from absorption co-efficient.

References

[1] Kirby, Simon. "The evolution of language." (2007).

[2] Shannon, Claude Elwood. "A mathematical theory of communication." *The Bell system technical journal* 27, no. 3 (1948): 379-423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

[3] Honda, Masaaki. "Human speech production mechanisms." *NTT Technical Review* 1, no. 2 (2003): 24-29.

[4] Osborne, Harold S. "Alexander Graham Bell." *Biographical Memoirs of the National Academy of Sciences* 23 (1943): 20-29.

[5] P. B. Denes; E. N. Pinson, “The Speech Chain: The Physics &Biology of Spoken Language”, Sigma Xi, The Scientific Research Society,p336A, Vol. 52, No. 3, SEPTEMBER 1964.

[6] Arai, Takayuki. "Education system in acoustics of speech production using physical models of the human vocal tract." *Acoustical science and technology* 28, no. 3 (2007): 190-201. https://doi.org/10.1250/ast.28.190

[7] Saravanan, D., and J. Surendiran. "A new framework for video data retrieval using hierarchical clustering technique." *International Journal of Engineering and Advanced Technology (IJEAT)* 8, no. 6S3 (2019). https://doi.org/10.35940/ijeat.F1389.0986S319

[8] Sarma, Mousmita, Kandarpa Kumar Sarma, Mousmita Sarma, and Kandarpa Kumar Sarma. "Speech processing technology: basic consideration." *Phoneme-Based Speech Segmentation using Hybrid Soft Computing Framework* (2014): 21-45. https://doi.org/10.1007/978-81-322-1862-3_2

[9] Surendiran, J., and M. Meena. "Analysis and Detection of Glaucoma from Fundus Eye Image by Cup to Disc Ratio by Unsupervised Machine Learning." In *2022 IEEE International Conference on Data Science and Information System (ICDSIS)*, pp. 1-6. IEEE, 2022. https://doi.org/10.1109/ICDSIS55133.2022.9915887

[10] Wolfe, Joe, Maëva Garnier, and John Smith. "Vocal tract resonances in speech, singing, and playing musical instruments." *HFSP journal* 3, no. 1 (2009): 6-23. https://doi.org/10.1109/ICDSIS55133.2022.9915887

[11] Surendiran, J., B. Mary Nathisiya, A. RadhaKrishnan, J. Mary Stella, M. Marimuthu, and N. Balaji. "Elimination of Data Modification in Sensor Nodes of WSN Using Deep Learning Model." In *2022 International Conference on Computer, Power and Communications (ICCPC)*, pp. 35-38. IEEE, 2022. https://doi.org/10.1109/ICCPC55978.2022.10072261

[12] Rao, Preeti, and A. Das Barman. "Speech formant frequency estimation: evaluating a nonstationary analysis method." *Signal Processing* 80, no. 8 (2000): 1655-1667. https://doi.org/10.1016/S0165-1684(00)00099-2

[13] Meenakshi, N., I. Juvanna, and J. Surendiran. "Diagnosis of Diabetes using Reverting Algorithm." In *2022 1st International Conference on Computational Science and Technology (ICCST)*, pp. 1-6. IEEE, 2022. https://doi.org/10.1109/ICCST55948.2022.10040290

[14] Story, Brad H. "Mechanisms of voice production." *The handbook of speech production* (2015): 34-58. https://doi.org/10.1002/9781118584156.ch3

[15] Vidhya, R. G., J. Surendiran, and G. Saritha. "Machine Learning Based Approach to Predict the Position of Robot and its Application." In *2022 International Conference on Computer, Power and Communications (ICCPC)*, pp. 506-511. IEEE, 2022. https://doi.org/10.1109/ICCPC55978.2022.10072031

[16] Turner, Richard, and Roy D. Patterson. "Inferring speaker size from speech formant data: A theoretical basis for vocal-tract length estimation."

[17] Obuli, Saravanan, T. P. Udhayashankar, R. Uma Mageswari, G. Saritha, J. Surendiran, and R. G. Vidhya. "An Analysis of Smart Clinical Trial Investigation Methods in Research Medicine Industries." In *2022 International Conference on Computer, Power and Communications (ICCPC)*, pp. 39-44. IEEE, 2022. https://doi.org/10.1109/ICCPC55978.2022.10072260

[18] Mahendru, Harish Chander. "Quick review of human speech production mechanism." *International Journal of Engineering Research and Development* 9, no. 10 (2014): 48-54.

[19] Sehgal, Prayank, and R. Kumar Jain. "Speech Processing." *International Journal of Engineering Sciences and Emerging Technologies* 2 (2013): 83-87.

[20] Sharon, M., D. Sharmila, Rachel Evangeline Christian, Kavya K. Shree, and J. Surendiran. "A Format Based Smart Object Tracking Model for Enhance the Quality of Digital Images." In *2022 International Conference on Computer, Power and Communications (ICCPC)*, pp. 29-34. IEEE, 2022. https://doi.org/10.1109/ICCPC55978.2022.10072053

[21] Reddy, Gopampallikar Vinoda, Kongara Deepika, Lakshmanan Malliga, Duraivelu Hemanand, Chinnadurai Senthilkumar, Subburayalu Gopalakrishnan, and Yousef Farhaoui. "Human Action Recognition Using Difference of Gaussian and Difference of Wavelet." *Big Data Mining and Analytics* 6, no. 3 (2023): 336-346. https://doi.org/10.26599/BDMA.2022.9020040