

Signal Processing and Analysis of a Person on the Basis of Time, Age and Gender



Taiba Khalid, Bhavana Narain

Abstract: Nowadays, signal processing is very common in machine learning. Tools like MATLAB and PRAAT made this really easy. Best example is google search using voice, this can be mostly beneficial for illiterate people and children. I have collected the sample of speech of 160 people of different age group, different genders, and at different and constant time as databases and then these databases are analysed. Each database consists of samples of human speech uttering the same first 10 alphabets of hindi varnamala. I have used two parameters i.e. formants and intensity for analysing the results of database. I have also characterized it into the form of constant time and variant time. Our aim is to identify the voices of a person on the basis of age and gender through different parameters like frequency and formants

Keywords: varnamala, MATLAB, PRAAT.

I. INTRODUCTION

Signal processing is a widely used technology that prevails every aspect of our life. It is defined as functions which is used to give information about the characteristics or properties of some phenomenon such as sound, images and biological measurements etc. It includes arrangement of the signal to change its properties or take out the information to get the desired result. Signal processing is an empowering technology that surrounds the basic theory, applications, algorithms and implementations of processing or transforming information consisted in various physical, symbolic or abstract formats mainly named as signals. It implies mathematical, statistical, computational, heuristics and linguistic representations, formalism and techniques or Representational modelling, analysis, synthesis, discovery, recovery, sensing, acquisition, extraction, learning, security or forensics.

a. What is Signal

A signal describes how some physical quantity varies over time or space. It is a function that is used to transfer information about the characteristics or properties of some phenomenon. It is commonly used in the field of electronics and telecommunication, it is used as time varying with the respect of voltage, current or electromagnetic wave which is used to carry information.

Revised Manuscript Received on June 15, 2020.

* Correspondence Author

Miss Taiba Khalid*, Research Scholar, Department of MSIT, MATS University, Raipur (C.G.), India.

Dr. Bhavana Narain, Associate Professor, Department of MSIT, MATS University, Raipur (C.G.), India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

b. TYPES OF SIGNALS

ANALOG SIGNALS: Analog signal is a signal which is continuous in amplitude, of a continuous independent variable. For example: Time.

DISCRETE SIGNAL: Discrete signal is a signal which is continuous in amplitude, but it is defined only for the set of discrete values for the independent variable.

DIGITAL SIGNAL: Digital signal is a signal which is similar as discrete signal but it is quantized finite amplitude values.

II. CONCEPTS OF MACHINE LEARNING

Machine learning is one of the fastest growing application of artificial intelligence that focuses on learning from the past experience. The main aim of machine learning is the lowering the human involvement and making the data work automatically without any involvement of human being. So that we can say that computers are working similar as human. The elementary process of machine learning is the observation of previous data and makes better decision on the basis of this previously observed data.

It includes the scientific study of algorithms and statistical models. Computers depend upon patterns and inferences to perform specific task. Machine learning is also considered as the subsets of artificial intelligence. Algorithms of machine learning uses sample data known as training data to build a mathematical model. Algorithms of machine learning uses many applications like e-mail filtering and computer vision. Machine learning mainly focuses on prediction like computational statistics, mathematical optimisation adds methods, theory and application to machine learning. Data mining is also the field of machine learning which focuses on exploratory data analysis through supervised learning.

a. Types of machine learning

- **Supervised learning:** It is the task of inferring a function from the training data. The training data consist of a set of observations together with its outcomes. It can be further subdivided in regression and classification algorithms. Regression is used to predict numerical values and classification techniques are used to predict categorical values.
- **Unsupervised learning:** Set of algorithm used to draw inferences from data sets consisting of input data without using outcome. The most common unsupervised learning method is cluster analysis, which is used for exploratory data analysis to find hidden patterns or grouping in data.

- **Reinforcement learning:** It is the most ambitious type of learning. It is rewarded from the sequence of action.

III. ALGORITHMS OF MACHINE LEARNING

- a. Linear regression:** It is the supervised learning algorithm.
- It models relationship between observation and outcomes using a straight line.
 - Root mean square error and gradient descent is used to fit the best possible line.
 - Provides insights into the factors that have greater influence on the outcome.
- b. Polynomial regression:** It is also based on supervised learning algorithm.
- It is a form of regression analysis in which the relationship between the observation and the outcome is modelled as an nth degree polynomial.
 - This method is more reliable when the curve is build on the large number of observations.
- c. Logistic regression:** It is also based on supervised learning algorithm.
- It is a misleading name even though the name suggests regression but in reality it is a classification technique.
 - It is used to estimate the probability of a binary response
 - It can be generalised to predict more than to categorical values.
- d. K nearest neighbour:** It is also based on supervised learning algorithm([1]Anuja Bombatkar,Gaytri Bhojar et al.,2014).
- It is a classification technique where an object is classified by a majority vote of its neighbour.
 - The observation is assigned the class which is most common among its k nearest neighbours.
 - The best choice of k depends upon the data, generally larger values of k reduce the effect of noise in the classification.
- e. Decision tree:** It is based on supervised learning algorithm.
- It is a decision support tool that uses a tree like model of decisions and their possible consequences.
 - Decision trees aim to create a model that predicts by learning simple decision rules from the training data.
- f. K means clustering:** It is based on unsupervised learning algorithm.
- K-means clustering aims to partition observations into k clusters.
 - For instance, the items in a super market are clustered in categories(butter, cheese and milk are grouped in dairy products)
 - K-means algorithm does not necessarily find the most optimal configuration. The k-means algorithm is usually run multiple times to reduce this effect.
- g. Principal component analysis:** It is based on unsupervised learning algorithm.
- It is a technique used to emphasize variation and strong patterns in a dataset.

- First principal component has the largest variances (that is accounts for as much of the variability in the data as possible)

Each succeeding component in turn has the highest variance possible under the constraint that it orthogonal to the preceding components.

IV. PRAAT AND MACHINE LEARNING

The word PRAAT is a dutch word which means talk or speak. It was developed by Paul Boersma and David Weenik of the university of Amsterdam. PRAAT is a tool used for doing phonetics .PRAAT is an open software tool which is designed for various needs ,with an easy interface ,searchable manual and various possibilities of analysis, manipulation and labelling . It is a software tool used for the analysis and reconstruction of acoustic speech signals .We can manipulate, synthesize and analyze speech and we can also create high quality pictures for our articles and thesis. For most of the platforms this software tool is free and is easily available.

With the help of PRAAT we can generate waveforms, spectrogram of narrow and wide band ,contours of intensity and tracks of pitch ([2]P. Gangamohan, Sudarsana Reddy Kadiri and B. Yegnanarayana).We can make recordings ,edit a recorded sound and extract individual sounds for further analysis .We can also get the detailed information about pitch, intensity, formants ,pulses etc. We can also enhance the certain frequency regions ,segments and label words, syllabus or individual phonemes .And lastly ,it puts the work or graphical form for printing.

Following are the four acoustic properties of speech sounds ,they are:

a. FREQUENCY: Frequency is defined as the individual pulsations produced by vocal chord vibration for a unit of time ([3]V.K.Mittal,B.Yegnarayana..2014). There is a different frequency for every child ,adult male and female speech because it rate of vibrations depends on length ,thickness and tension of the vocal chords. There are two types of frequencies in speech sounds They are fundamental frequency and formant frequency .Fundamental frequency is a frequency(F_0) which reflects the rate of vocal chord vibration during phonation(pitch) and it also relates to function of vocal chords. And formant frequency is a frequency which gives the vocal tract configuration.

b. TIME: Time is that property of speech sound which gives the duration of a given sound.

c. Amplitude: The darkness of the band are marked by the amplitude.The darkness marked at the corresponding point on the screen depends on the greatness of the intensity of the sound energy present in a given time and frequency.

d. FORMANT: In a speech wave, a formant is a concentration of acoustic energy around a particular frequency .In every different frequency there are several formants roughly one in each 1000Hz band.Formants occur at roughly 1000Hz intervals to put it differently.

In the vocal tract each formant is corresponded by the resonance.

V. DATASETS AND METHODOLOGY

In this paper, we have presented the different datasets of voices of 100 people of different age group and gender at constant and different time which are collected during the research process([4]F.Burkardatt, A. Paeschke, M. Rolfes, W. Sendlmeier, B. Weiss4) .([5]Dimitrios vervidis,Constantine Kotropplous.....2004)

.In this chapter methodology of the research is also discussed

DATASETS

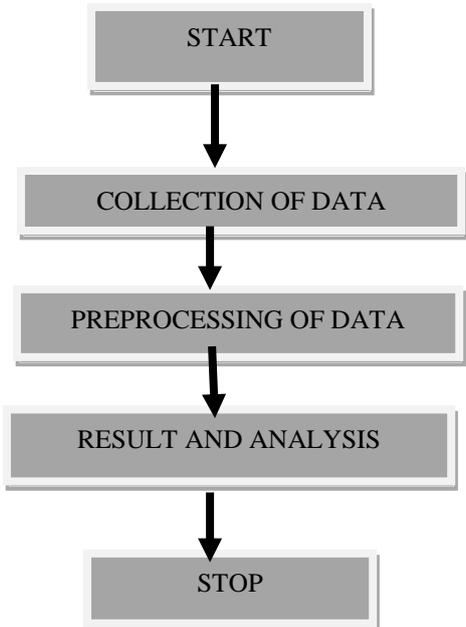
Data is been collected from both male and female of various age groups and at different and constant time.(11-20,21-30,31-40,41-50,51-60). 10 male and 10 female voices are collected from all the age groups and at different and constant time.In this way,100 voice samples are collected.Everyone uttered the same first 10 alphabets of hindi varnmala.

Data is collected in two ways:

- By taking constant time for various persons of same age group and same gender.
- By taking different time for different person of same age group and same gender.

Data is been recorded with the help of tool known as PRAAT.

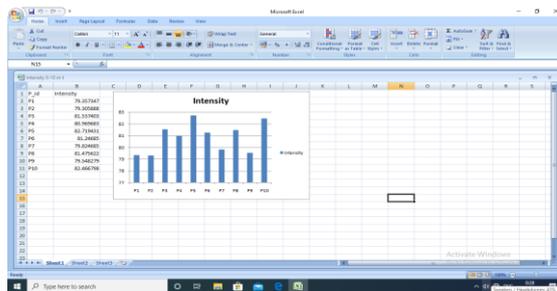
Flow chart of the methodology is stated below:



VI. RESULT

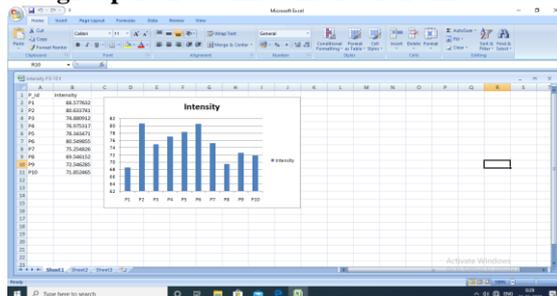
We have analyzed the data using different parameters like intensity and formants with the help of bar graphs. .

1. Graph of intensity of 10 different males of age group 10-20 at fixed interval of time



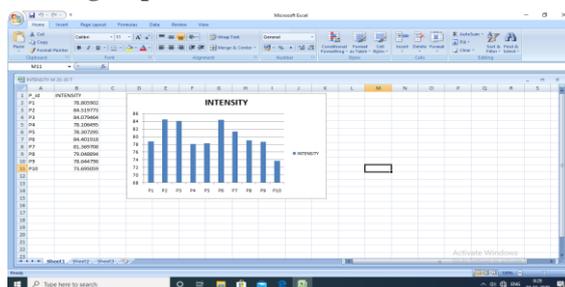
This graph shows that P5 has the maximum intensity.

2. Graph of intensity of 10 different females of age group 10-20 at fixed interval of time.



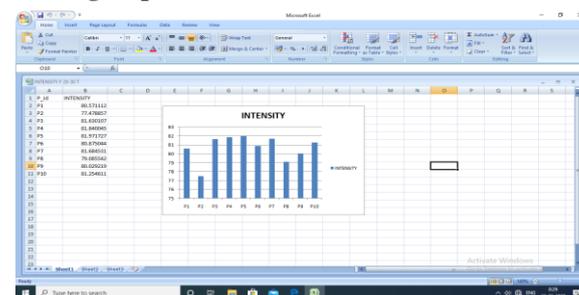
This graph shows that P2 has the maximum intensity.

3.Graph of intensity of 10 different males of age group 20-30 at fixed interval of time.



The graph shows that P2 has the highest intensity.

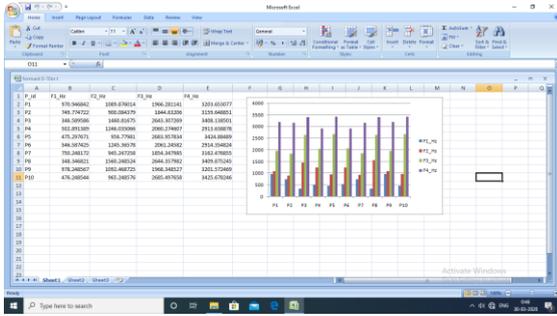
4.Graph of intensity of 10 different females of age group 20-30 at fixed interval of time.



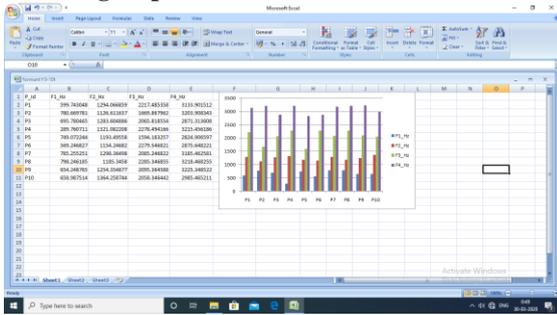
This graph shows that P5 has the highest intensity.

5..Graph of formant of 10 different males of age group 10-20 at fixed interval of time.

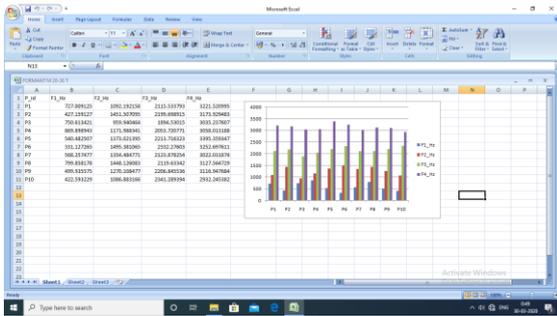
Signal Processing and Analysis of a Person on the Basis of Time, Age and Gender



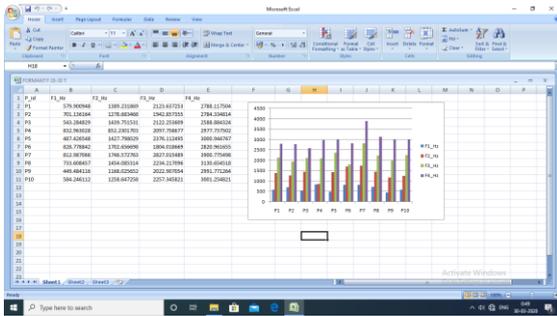
6. Graph of formant of 10 different females of age group 10-20 at fixed interval of time.



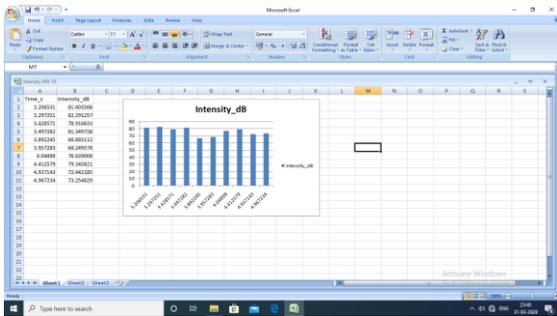
7. Graph of formant of 10 different males of age group 20-30 at fixed interval of time.



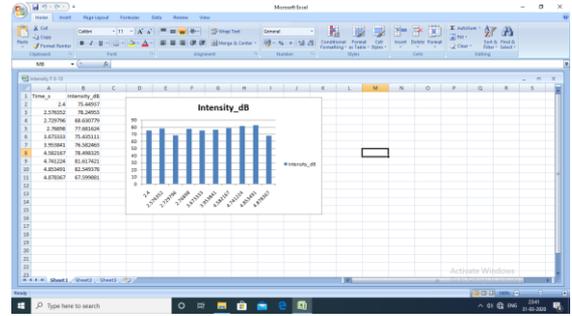
8. Graph of formant of 10 different females of age group 20-30 at fixed interval of time.



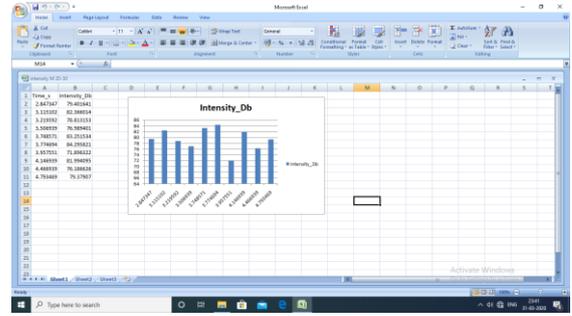
9. Graph of intensity of 10 different males of age group 10-20 at different interval of time.



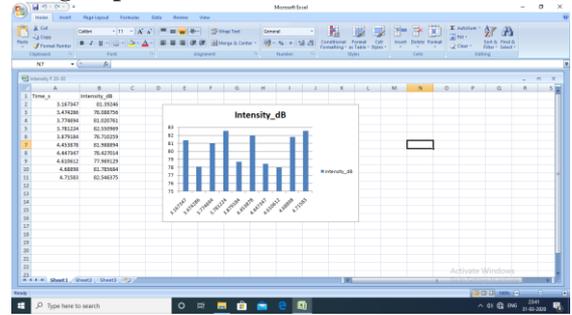
10. Graph of intensity of 10 different females of age group 10-20 at different interval of time.



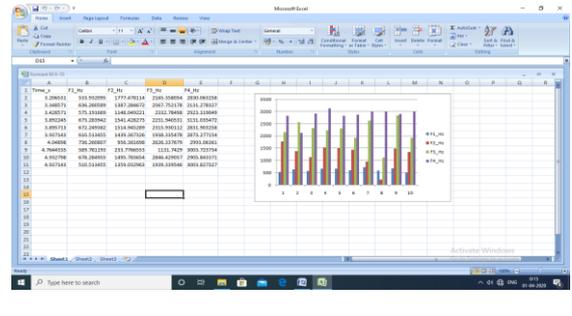
11. Graph of intensity of 10 different males of age group 20-30 at different interval of time.



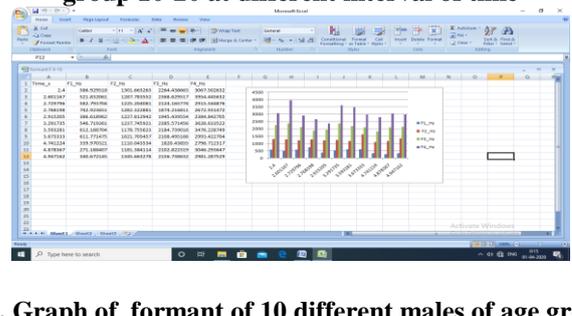
12. Graph of intensity of 10 different females of age group 20-30 at different interval of time.



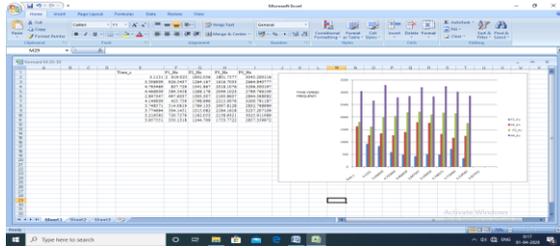
13. Graph of formant of 10 different males of age group 10-20 at different interval of time.



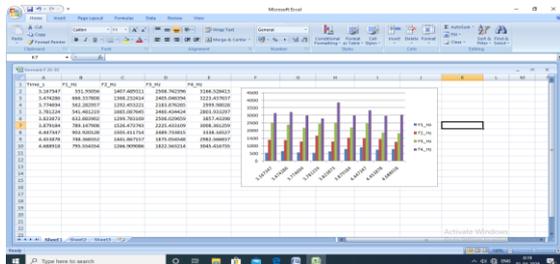
14. Graph of formant of 10 different females of age group 10-20 at different interval of time



15. Graph of formant of 10 different males of age group 20-30 at different interval of time



16. Graph of formant of 10 different females of age group 20-30 at different interval of time



VII. CONCLUSION

- I have concluded that:
 - Natural speech cannot be easily classified.
 - Intensity of female speech is lesser than the male speech.
 - There is no fixed or particular change in intensity when we change the timings
 - F3 has always highest frequency in both genders and in all age groups.
 - Frequency of female voice is higher than the male voice.
 - It cannot achieve a correct classification.
- So, there is a need for establishing a protocol that will resolve all these issues.

FUTURE WORK

- For future work I can do the following:
- Increase the datasets.
 - Increase the parameters.
 - Can develop the app from which we can search through voice in which speech is in hindi language and can work faster than google chrome.

REFERENCES

1. Anuja Bombatkar, Gaytri Bhojar, Khushbu Morjani, Shalaka Gautam, Vikas Gupta a review on “Emotion recognition using Speech Processing Using k-nearest neighbor algorithm”. 2014.
2. P. Gangamohan, Sudarsana Reddy Kadiri and B. Yegnanarayana a review on “Analysis of Emotional Speech”, 2014
3. V.K.Mittal, B. Yegnanarayana a review on “Analysis of production of laughter”, 2014
4. F. Burkardatt, A. Paeschke, M. Rolfes, W. Sendlmeier, B. Weiss4, a review on “A database of german emotional speech”.
5. Dimitriosvervidis, Constantine Kotroplous, a review on “emotion speech recognition: resources, features and methods”, 2004.

AUTHORS PROFILE



Miss Taiba Khalid is a research scholar of MPhil (computer science), MSIT Department of MATS University, Raipur (C.G.) She has completed her M.Sc (computer science) from department of CSIT from guru ghasidas central university. This is her first research paper on the basis of her MPhil dissertation which is on signal process and analysis.



Dr. Bhawna Narain is Associate Professor in MSIT Department of MATS University, Raipur (C.G.) Ph.D (Computer Application), MPhil (Computer Science). Her work experience is 15 years. Her area of interest is computer networks (ad hoc, mesh) and digital image processing. She has published 41 papers in national and international journals and conferences. Two books are published, she has worked as co-pi in two minor projects. She has been working as state student coordinator of computer society of INDIA and awarded as best teacher and researcher by national organization.