Linear Prediction Algorithm for Voiced and Unvoiced Consonants of Human Speech for Speaker Recognition

Chetan Solanki[1], Chhaya Kinkar[2]
[1,2]*Sagar Institute of Research Technology and Science Bhopal (MP),India*

Abstract - Linear prediction analysis is very important part of any speaker recognition system. this paper describes analysis of speech signal over the various models of LP, these models are LP autocorrelation, LP coefficients and LP residue. For all of these models we are basically require to develop some algorithms, these algorithms are nothing but the set of various instructions and statements which we needed to provide interface between the user and used software i.e. SSCILAB. In this paper we presented the result an experiment which is perform on 2 different persons to recognize the individual one.

Keywords - Speaker recognition, Linear prediction (LP), autocorrelation, residual.

1. INTRODUCTION

With the advent of latest technologies, speech analysis has changed the potential in many security and confidential systems. Such systems are employed in automatic recognition system who is speaking on the basis of individual information included in speech waves. This technique enhances the chances of identity and control access to services such as voice dialing, banking by telephone, telephone shopping, database access services, information services, voice mail, security control for confidential information areas, and remote access to computers [1].

Speech signal can be classified into voiced, unvoiced and silence regions. The near periodic vibration of the vocal folds is excited for the production of voice speech. The random like excitation is present for unvoiced speech. There is no excitation during the silence reign. Majority of speech regions is voiced in nature that includes vowels, semivowels and other voiced components. In this paper we are considering both voiced and unvoiced portion of human speech for analysis and synthesis of speaker recognition [2].

2. PREVIOUS WORK

A considerable number of speaker-recognition activities are being carried out in industries, national laboratories and universities. Several enterprises and universities have carried out intense research activities in this domain and have come up with various generations of speaker-recognition systems. Those institutions include AT&T and its derivatives (Bolt, Beranek, and Newman)[3].The Dale Mole Institute for Perceptual Artificial Intelligence (Switzerland); MIT Lincoln Labs; National Tsing Hua University (Taiwan); Nippon Telegraph and Telephone (Japan); Rutgers University and Texas Instruments (TI) [4].Sandia National Laboratories, National Institute of Standards and Technology, the National Security Agency etc. have conducted evaluations of speaker-recognition systems. It is to be noted that it is difficult to make reasonable comparison between the text-dependent approaches and the usually more difficult text-independent approaches [5].

3. LINEAR PREDICTION

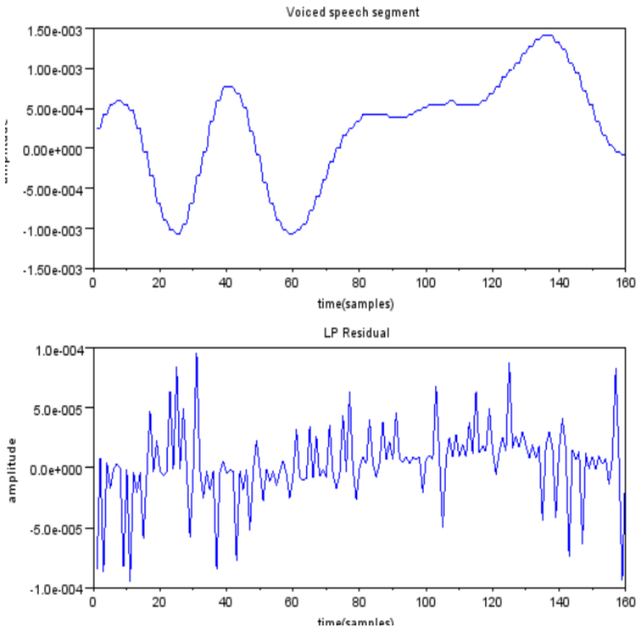
Linear prediction (LP) forms an integral part of almost all modern day speech coding algorithms. The fundamental idea is that a speech sample can be approximated as a linear combination of past samples. Within a signal frame, the weights used to compute the linear combination are found by minimizing the mean-squared prediction error; the resultant weights, or linear prediction coefficients (LPC), are used to represent the particular frame. [6]

Linear prediction algorithm is the most preferable model in modern day human voice analysis and synthesis. Here in Linear prediction the basic concept used is the speech sample can be approximated as a linear combination of past samples [7]. Thus, LP is an identification technique where parameters of a system are found from the observation. The basic assumption is that speech can be modeled as an AR signal, which in practice has been found to be appropriate. LP can also be viewed as a redundancy removal procedure where information repeated in an event is eliminated. After all, there is no need for transmission if certain data can be predicted. By displacing the redundancy in a signal, the amount of bits required to carry the information is lowered, therefore achieving the purpose of compression [8].

Estimating the parameters is the responsibility of the encoder. The decoder takes the estimated parameters and uses the speech production model to synthesize speech.

The approach of utilizing noise to generate the output signal is somehow mystifying. The approach throws away all phase information of the original waveform, preserving only the magnitude of the frequency spectrum.

Here we have a sample graph of Linear prediction of voice signal.



4. PROPOSED METHOD

Here we use some experimental setup and assumption for the keen observation of the human voice, we analyze the speech signals on various parameters like LP autocorrelation, LP coefficient, and LP residue analysis by using Linear prediction algorithm. For all such special and modern analysis we require to make a best suitable platform so we make some special arrangements for recording of voice signal till the final plotting of resultant graphs.

For the voice recording we just use the voice recording device. The noise avoidation is the very important aspect while recording voice signal, and for that we are make some special arrangements such as-

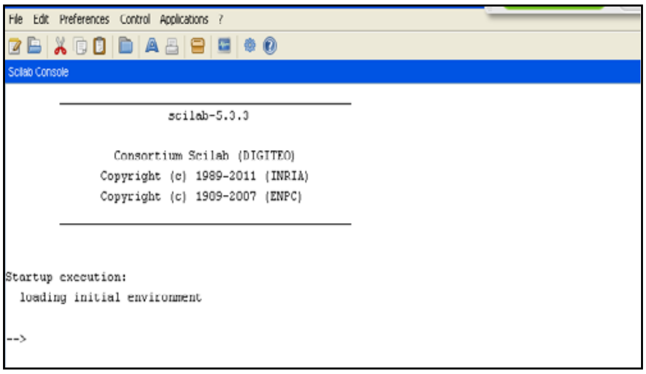
1) We record voice by using quality microphone.
2) We provide 2 second gap of empty slot in starting as well as at the end and we also tried to avoid human noise like coughing, sneezing, whispering etc.
3) We also put the microphone at an appropriate distance from mouth for avoiding unnecessary morganing.
4) For avoiding noise we provide adequate voice cutting also.

5. SOFTWARE PLATEFORM USED

There are few software's used in the analysis of different human voice such as SCILAB, audio converter and cutter.

SCILAB is stated as scientific laboratory which basically software package for numerical computation for engineers and scientists. SCILAB is endowed with powerful tools and easy syntax. Matrix being the basic fundamental object for calculation matrix manipulation can be easily handled [9].

It can be used for signal processing, statistical analysis, image enhancement, fluid dynamics simulations, numerical optimization, and modeling, simulation of explicit and implicit dynamical systems and (if the corresponding toolbox is installed) symbolic manipulations. SCILAB is one of several open source alternatives to MATLAB. SCILAB is released as open source under the CECILL license (GPL compatible), and is available for download free of charge. Scilab is available under GNU/Linux, Mac OS X and Windows XP/Vista/7/8 [10].



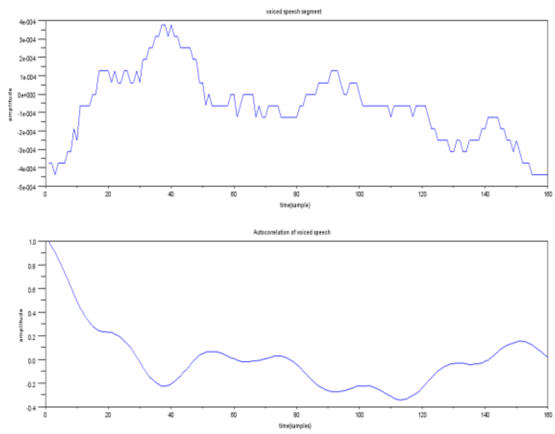
6. EXPERIMENTAL RESULT

The Linear prediction analysis is the feature extraction process. On the basis of the Linear prediction analysis we have the resultant graphs of two different persons.

The resultant graphs for LP autocorrelation, LP coefficient, and LP residual are plotted below.

Linear prediction estimation for person 1-

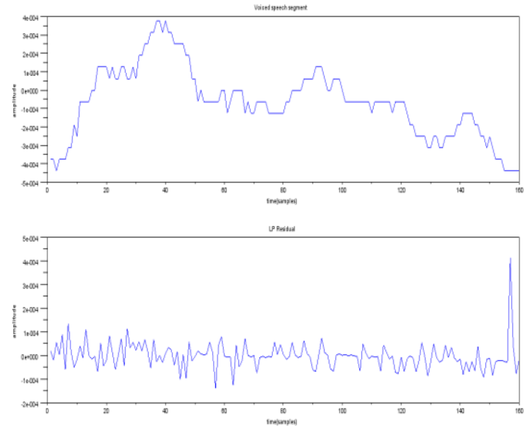
Person 1- Estimation from Code 1
LP Autocorrelation-



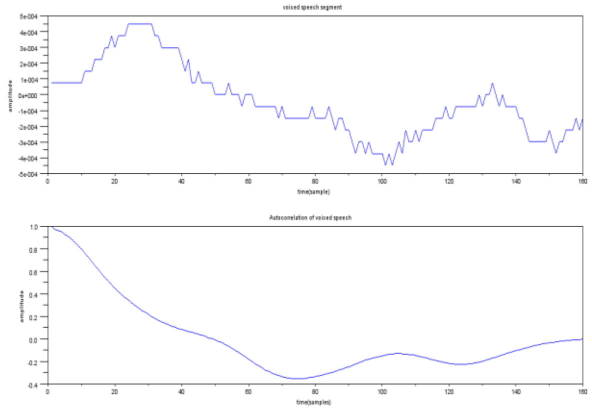
LP Coefficient-

- 0.8504332 - 0.2513562 0.1317704 - 0.0397921 0.0390939 - 0.0023926 - 0.0178927 0.0709137 - 0.0327843 0.0252964

LP Residual-



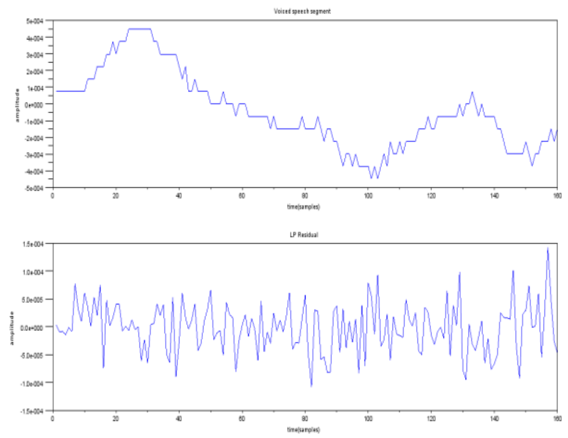
Linear prediction estimation for person 1-
Person 1- Estimation from Code 2
LP Autocorrelation-



LP Coefficient-

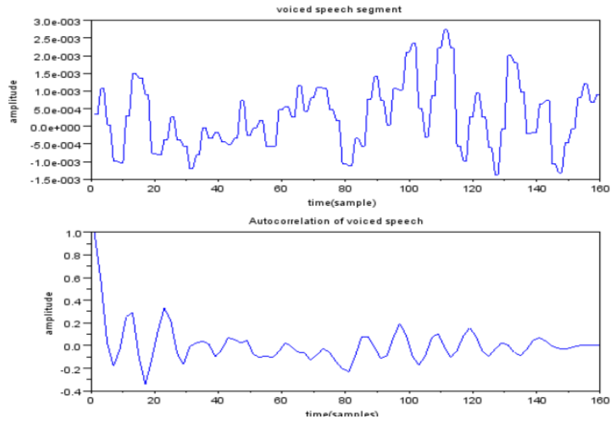
- 0.6129222 - 0.2905318 - 0.3094031 0.2629045 - 0.1642107 0.0102373 - 0.0870522 0.1799532 - 0.0977498 0.1467699

LP Residual-



Person 2- Estimation from Code 1

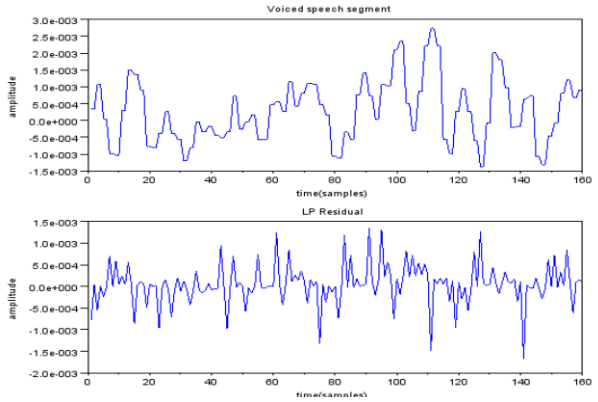
LP Autocorrelation-



LP Coefficient-

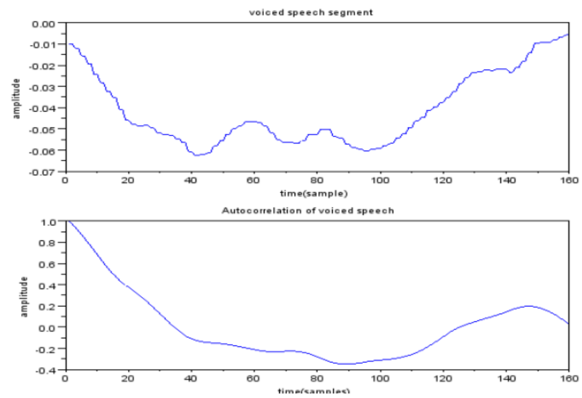
- 0.9565796 0.0330983 - 0.0181850 0.5098639 - 0.4751228 0.3249864 - 0.2902453 0.3252794 - 0.3103660 0.0434204

LP Residual-



Person 2- Estimation from Code 2

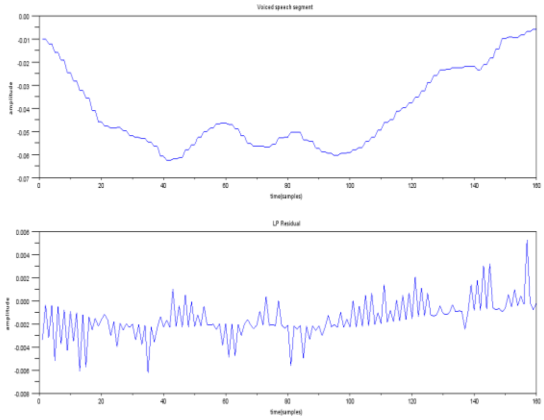
LP Autocorrelation-



LP Coefficient-

- 0.9772757 - 0.0904886 0.0887238 - 0.0369623 0.0356681 - 0.0212849 0.0199907 0.0110613 - 0.0128261 0.0227243

LP Residual-



7. CONCLUSION

This paper involves feature extraction process using freeware SCILAB for students so that they can achieve same efficient as that of MATLAB. The objective of this paper was to create a system which is efficient enough in comparison to different expensive software's, and making speech analysis easier for an unknown speaker. By investigating the extracted features of the unknown speech and then compare them to the stored extracted features for each different speaker in order to identify the unknown speaker. The feature extraction is done by using LPC method. We hope that our experience in analyze the human voice uniqueness and some special characteristics may be useful and inspiring for other people or groups who will face a similar task.

REFERENCES

[1] Sunil Sharma, Ravi Jain, Isha Suwalka topic on "Speech Analysis and Feature Extraction Using Scilab", in International Journal of Modern Communication Technologies & Research (IJMCTR) ISSN: 2321-0850, Volume-1, Issue-1, March 2013.

[2] Speech Coding Algorithms Foundation and Evolution of Standardized Coders. WAI C. CHU Mobile Media Laboratory DoCoMo USA Labs San Jose, California.

[3] Zoltan Tuseke, Pavel Golik, Ralf Schlliter, Fri R Drepper Non Stationary feature extraction for automatic speech recognition" Human language technology and recognition computer science dept, RWTH Aachen University.

[4] Roucos, S. Berouti, M. Bolt, Beranek and Newman, Inc.Cambridge "The application of probability density estimation to text independent speaker identification " IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '82. Volume : 7, On page(s): 1649-1652. Publication Date: May 1982.

[5] Yong Lu, Haining Huang "Research on a kind of Noisy Tibetan speech recognition algorithm based on WNN" 2011 IEEE.

[6] Resmi K, Satish Kumar, H.K. Radhika Chhabra "Graphical Speech Training System for Hearing Impaired"2011 IEEE.

[7] Thiang and Suryo Wijoyo topic on "Speech Recognition Using Linear Predictive Coding and Artificial Neural Network for Controlling Movement of Mobile Robot" in International conference on information and electronics engineering IPCSIT vol 6 2011 LACSIT press Singapore.

[8] Lawrence Rabiner, and Biing Hwang Juang, "Fundamentals of Speech Recognition", Prentice Hall, New Jersey,1993.

[9] SCILAB manual.

[10] SCILAB speech processing manual.

AUTHOR'S PROFILE

Chetan Solanki has received his Bachelor of Engineering degree in Electronics and communication Engineering from Sagar Institute of Research Technology and Science, Bhopal(MP) in the year 2011. At present he is pursuing M.Tech. with the specialization of Digital Communication in Sagar Institute of Research Technology and Science, Bhopal(MP). His area of interest is DSP, Signal and Systems, Digital communication.

Chhaya Kinkar has received her M.tec in Electronics & Communication Engineering from Govt. Engineering college Pune (MH) in the year 2005. At present she is working as an Associate Professor at Sagar Institute of Research Technology and Science, Bhopal (MP). Her areas of interests are Signal and Systems, Digital Signal Processing, Data communication.