

Optimal Method for Pitch Period Estimation in SCILAB

Chetan Solanki

*PG Scholar, Electronics and comm. Engineering,
Sagar institute of research technology and science Bhopal, Madhya pradesh, India*

Chhaya Kinkar,

*Asst. professor, Electronics and comm. Engineering,
Sagar institute of research technology and science Bhopal, Madhya pradesh, India*

Abstract— Speech is one of the natural forms of communication. Speech processing is emerged as one of the important application area of digital signal processing. 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 is one of the most useful and popular biometric recognition techniques in the world especially related to areas in which security is a major concern. It can be used for authentication, surveillance, speaker recognition and a number of related activities. Many systems have been made which costs a lot due to expensive software which checks the voice characteristics of the input utterance, using an automatic speaker recognition system. This paper depicts analysis side of the system implemented using freeware language SCILAB. . In this paper we presented the result an experiment which is perform on 5 different persons to recognize the individual one. Recognition platform is implemented on SCILAB for recognition purpose pitch period or individual person is calculated. the signal processing application of SCILAB has become boon in comparison to MATLAB for all programmers.

Keywords – Pitch Period, Pitch frequency, voiced and unvoiced sound.

I. INTRODUCTION

Speaker recognition is the process of recognizing automatically who is speaking on the basis of individual information included in speech waves. This technique uses the speaker's voice to verify their identity and provides control access to services such as voice dialing, database access services, information services, voice mail, security control for confidential information areas, remote access to computers and several other fields where security is the main area of concern [1].

Speech is a complicated signal produced as a result of several transformations occurring at several different levels: semantic, linguistic, articulatory, and acoustic. Differences in these transformations are reflected in the differences in the acoustic properties of the speech signal. Besides there are speaker related differences which are a result of a combination of anatomical differences inherent in the vocal tract and the learned speaking habits of different individuals. In speaker recognition, all these differences are taken into account and used to discriminate between speakers [2].

The pitch or frequency of voice signal is the most important parameter in speech analysis, synthesis and coding the application. Pitch frequency is basically depends on the voicing characteristics of the person, and internal structure of the throat[3].The range of pitch frequency is different for male and female voice, i.e. for male voice pitch frequency is of range in between 50 Hz to 250Hz, while for the female this range usually falls between the range 120 Hz to 500 Hz. And in the terms of the period, the range for a male is 4 ms to 20 ms, while for a female it is 2 ms to 8 ms . Pitch period of the voice must be estimated at every frame. By comparing a frame with previous frame, it is possible to identify the period in which the signal repeats itself, resulting in an estimate of the actual pitch period the voice[4].

In this paper we will going to study about pitch period estimation of the different person of the different age group.

II. LITERATURE SURVEY

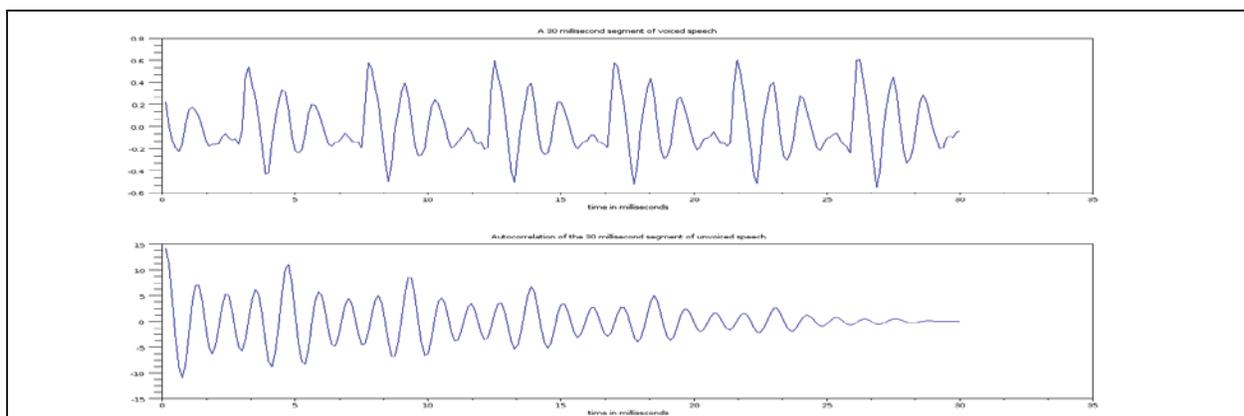
A considerable number of speaker-recognition activities are being carried out in industries, national laboratories and universities. Several enterprises and universities have carried out intense research activities in this domain and have come up with various generations of speaker-recognition systems. Those institutions include AT&T and its derivatives (Bolt, Beranek, and Newman)[5],[6].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)[7].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 [8].

III. PITCH PERIOD ESTIMATION

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 include vowels, semivowels and other voiced components. The voiced regions look like a near periodic signal in the time domain representation. In a short term, we may treat the voiced speech segments to be periodic for all practical analysis and processing. The periodicity associated with such segments defined is 'pitch period T_0 ' in the time domain and 'Pitch frequency or Fundamental Frequency ' F_0 ' in the frequency domain. Unless specified, the term 'pitch' refers to the fundamental frequency ' F_0 '. Pitch is an important attribute of voice speech. It contains speaker-specific information. It is also needed for speech coding task. Thus estimation of pitch is one of the important issues in speech processing.

Pitch period estimation of speech and music signals is an essential component in various speech processing applications such as speech coding, speaker identification and verification, and speech segregation. The existing pitch detectors are broadly classified into either event detection pitch detectors or nonevent detection pitch detectors. Event detection pitch detectors estimate the pitch period by locating the instant at which the glottis closes (called an event) and then measuring the time interval between two such events. The non-event based pitch detectors estimate pitch period by a direct approach like autocorrelation or cepstrum method [9].

For instance, if $T_0 = 10$ m sec, $F_s = 8000$ Hz , then T_0 in samples= $10 \times 8 = 80$. $F_0 = 8000/80 = 100$ Hz



IV. 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 pitch and frequency analysis by using pitch estimation 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 avoidance is the very important aspect while recording voice signal, and for that we are make some special arrangements such as-

- 1) We preferred midnight time because at that time the external noise was very low, as compared to day time.
- 2) We record voice by using quality microphone.
- 3) 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.
- 4) We also put the microphone at an appropriate distance from mouth for avoiding unnecessary morganing.
- 5) For avoiding noise we provide adequate voice cutting also.

V. SOFTWARE PLATFORM 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, [10].

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.[11]

```

File Edit Preferences Control Applications ?
Scilab Console
-----
scilab-5.3.3
Consortium Scilab (DIGITEO)
Copyright (c) 1989-2011 (INRIA)
Copyright (c) 1989-2007 (ENPC)
-----
Startup execution:
  loading initial environment
-->

```

VI. RESULT

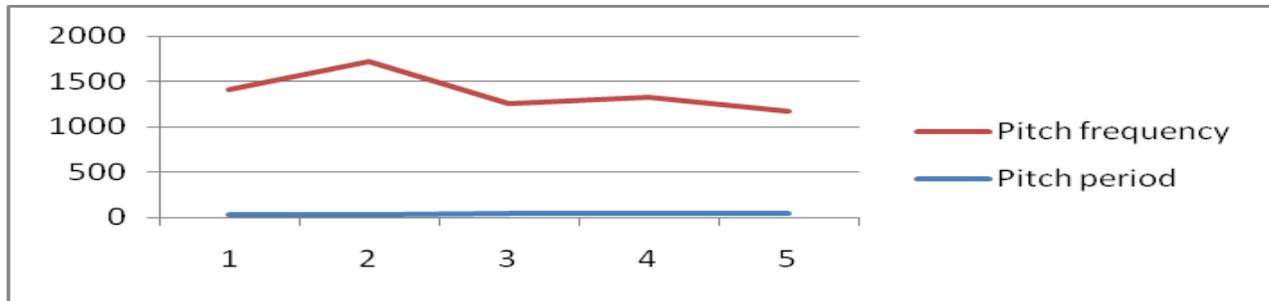
The pitch period and pitch frequency analysis is the main factor of the pitch estimation of the human voice. On the basis of the pitch period and pitch frequency we have the resultant graphs of each person.

The line graphs for pitch period and pitch frequency are plotted below.

For code 1 the pitch period and pitch frequency of different person:-

Different person	person1	person2	person3	person4	person5
Pitch period	32	26	36	34	39
Pitch frequency	1378.12	1696.15	1225	1297.05	1130.76

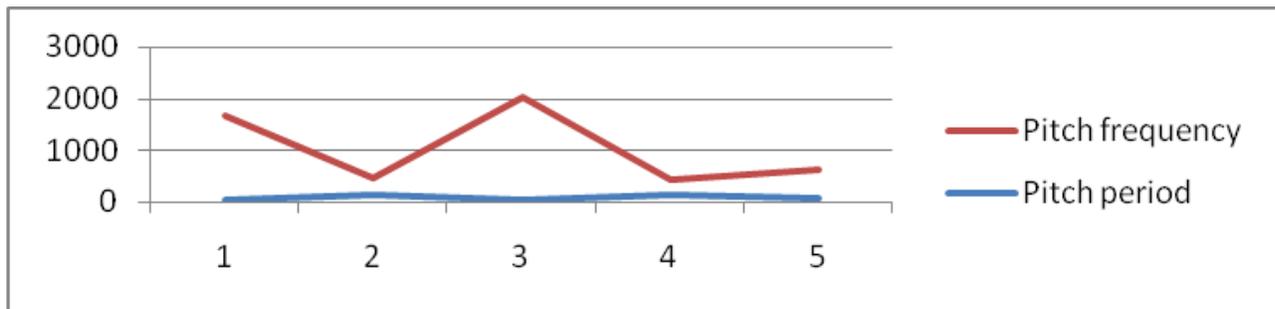
Line graph



For code 2 the pitch period and pitch frequency of different person:-

Different person	person1	person2	person3	person4	person5
Pitch period	27	134	22	151	80
Pitch frequency	1633.33	329.10	2004.54	292.05	551.25

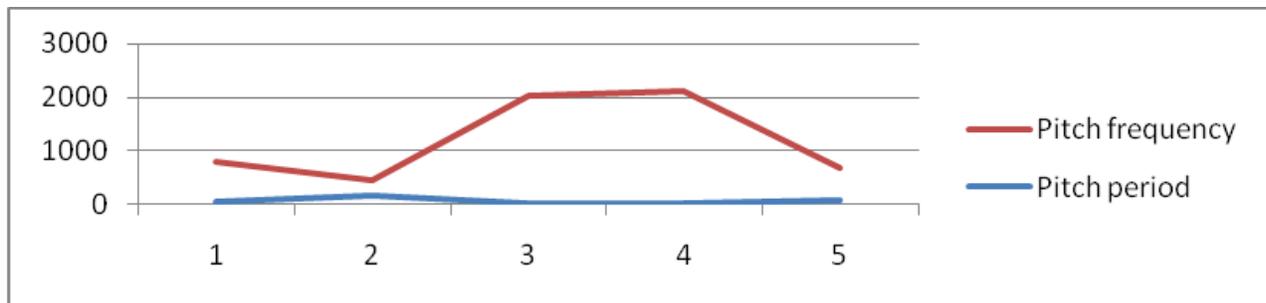
Line graph



For code 3 the pitch period and pitch frequency of different person:-

Different person	person1	person2	person3	person4	person5
Pitch period	62	158	22	21	74
Pitch frequency	711.29	279.11	2004.5	2100	595.94

Line graph



VII. CONCLUSION

In this paper we focused on the theoretical and practical characteristics of human voice. Here we offer several approaches for analysis of that concept, from a very simple one that can be used as a starting point in creating an annotated speech database to more complex methods that allow for a fully automated processing for a large speech corpus. For each method we have made a qualitative evaluation of its segmentation accuracy, which shows that even the simpler methods provide results that are acceptable in speech analysis/ recognition tasks.

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] Fu Zhonghua; Zhao ; “An overview of modeling technology of speaker recognition”, IEEE Proceedings of the International Conference on Neural Networks and Signal Processing Volume 2, Page(s):887 – 891, Dec. 2003.
- [2] .Mauge Sundermeyer, Ralf Schliter, Hermann Ney, “On the Estimation of discount parameters for language model smoothing” Human language technology . earn recognition cs dept, RWTH Aachen University
- [3] Roucos, S. Berouti, M. Bolt, Beranek and Newman, Inc., Cambridge, MA; “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.
- [4] Castellano, P.J., slomka, S. Sriram, S.; “Telephone based speaker recognition using multiple binary classifier and Gaussian mixture models” IEEE International Conference on Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 Volume 2, Page(s) 1075 – 1078 April 1997.
- [5] 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.
- [6] SPEECH CODING ALGORITHMS Foundation and Evolution of Standardized Coders. WAI C. CHU Mobile Media Laboratory docomo USA Labs San Jose, California.
- [7] Yong Lu, Haining Huang “Research on a kind of Noisy Tibetan speech recognition algorithm based on WNN” 2011 IEEE.
- [8] Resmi K, Satish Kumar, H.K. Sardana , Radhika Chhabra “Graphical Speech Training System for Hearing Impaired” 2011 IEEE.
- [9] Shing-Tai Pan, Ching-Fa Chen, Wei-Der Chang, Yi-Heng Tsai “Performances Comparison between Improved DHMM and Gaussian Mixture HMM for Speech Recognition” 2011 IEEE.
- [10] SCILAB manual.
- [11] SCILAB speech processing manual.