International Journal of Speech and Language Processing
Volume 1, Issue 1, 2011, pp-07-09
Available online at: http://www.bioinfo.in/contents.php?id=112

Analyzing Children's Speech for Biometric Identification

Rohini Bhujangrao Shinde
Department of Computer Science, ollege of Computer Science & Information Technology, Latur
e-mail: rvmali007@gmail.com

Abstract—This paper reports on the recording and planned research activities on recognition of children's speech in the Marathwada region. The task is quite more difficult than recognition of adult speech for several reasons. High fundamental frequency and formant frequencies change the spectral shape of the speech signal. Also the pronunciation and the use of language differ from adult speech. One more objective is there it is very difficult to identify the voice of male child & female child we dealing these points.

I. INTRODUCTION

In daily life, the new speech processing applications are used in many purposes especially in Telecom-munications, Security and Entertainment businesses. Speech recognition for children is a key to create automatic speech training, language earning, reading tutors, information retrieval, and entertainment applications for children. Children are also a vital resource in the research of adaptation and normalization techniques since their speech differs significantly from that of adults not only in the acoustic domain, but also on the lexical and grammatical levels. The aim of this project is the development of a program which discriminates gender of the recorded voice. Which is a one element can be used for speech recognition system.

II. DIFFICULTIES

A few main reasons can be attributed for the lower recognition performance of children's speech. The first one is physical size. Children have shorter vocal tract and vocal folds compared with adults. This results in higher position of formants and fundamental frequency. The high fundamental frequency is reflected as a large distance between the harmonics, resulting in poor spectral resolution of voiced sounds. The difference in vocal tract size results in a non-linear increase of the formant frequencies. The higher formants of children may also fall outside the bandwidth of the transmission channel. This is particularly a problem with telephone speech, since a childs third formant may be higher than 4 kHz and hence lost.

A second problem is that younger children may not have a correct pronunciation. Sometimes they have not yet learnt how to articulate certain phonemes. A special case is around the age of six when they lose their front milk teeth, which will cause difficulties in producing certain dental consonants, such as.

III. RECORDINGS SPECIFICATIONS

A. Technical Specifications

A headset and an directional microphone are used for the recordings. The directional unit is placed on a table at a distance of around 50 – 100 cm from the subject. The recordings are stored directly onto computer files via an external A/D converter connected to the USB port. By using the external module instead of the computer's sound card, the quality of the recordings is always the same, regardless of which computer is used.

The low frequency limit of the table microphone is around 10 Hz. Several environmental sounds, such as door closures, foot steps, ventilation, etc. can have high intensity in this frequency region, much higher than the speech signal. It is necessary to set the gain low enough to avoid clipping of these signals, which would otherwise severely distort the speech signal as well.

This is unfortunate since there will be less resolution for the speech signal in this way. we have been able to use a more simple solution, by using the higher A/D resolution of the USB converter (24 bits), The hardware filter can thus be replaced by a software algorithm or just by discarding the low frequency range in the FFT spectrum during the analysis.

The sampling frequency is set to 32 kHz, which is considerably higher than what is required for high performance adult speech recognition. However, as mentioned previously, the smaller vocal tract size of children expands the spectrum. The first three formants are shifted upwards by around 65% for 5-7 year old children compared to adults.

B. Recording Conditions and Texts

The children are recorded at after-school. A separate room is used with one child present at a time. The child repeats utterances that are printed on a paper, they reads texts presented on a paper. The text material for each child consists of alphabets, digits & some words. The younger children of 4 and 5 years age,

C. Definition of Biometrics

Any Automatically measurable, robust and distinctive physical characteristics or personal trait that can be used to identify an individual or verify the claimed identity of an individual. Biometrics is the automatic recognition of a person using distinguishing traits.

Biometrics refers to the automatic identification of a person based on his/her physiological or behavioral characteristics. This method of identification gives several advantages over traditional methods involving ID cards (tokens) or PIN numbers (passwords) for various reasons:

- The person to be identified is required to be physically present at the point-of-identification.
- Identification based on biometric techniques remove the need to remember a password or carry a token. With the increased integration of computers and Internet into our everyday lives, it is necessary to protect sensitive and personal data. By replacing Pins, biometric techniques can potentially prevent unauthorized access to ATMs, cellular phones, laptops, and computer networks

A biometric system is essentially a pattern recognition system which recognizes a user by determining the authenticity of a specific physiological or behavioral characteristic possessed by the user. Several important issues must be considered in designing a practical biometric system. First, a user must be enrolled in the system so that his biometric template can be captured. This template is securely stored in a central database or a smart card issued to the user. The template is retrieved when an individual needs to be identified. Depending on the context, a biometric system can operate either in verification (authentication) or an identification mode.

IV. EXPERIMENT RESULTS

To have an idea about the performance improvement due to the gender inclusion we have done the followings:

- A speech of alphabets like “A”, “B” up to “Z” was constructed using pooled examples of both male and female speakers.
- Another two similarly constructed models of the alphabet “A”, “B”, up to “Z” were also constructed; one based on male speakers’ examples, and the other based on female speakers’ As a comparison of the state clustering of the three prepared models the Gaussian ellipsoids

The following table illustrates that speech of male child & female child may be similar to heard but when they are recorded and digitized in that case they obtains different values for same alphabet pronunciations.

TABLE 1

Sample speech	Pronounced word	Mean value	Std
Male child	A	4.2137e-006	0.0421
Female child	A	5.3068e-006	0.0562

The following figure illustrates that speech of male child & female child may be similar to heard but when they are recorded and digitized in that case they obtains different graphs for same alphabet.

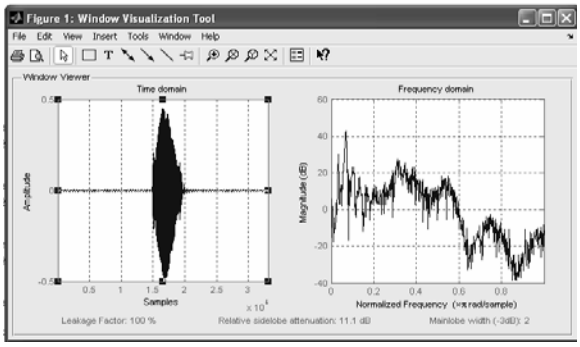

Fig.1: Graphical Representation for Alphabet “A” Pronounced by Female Child

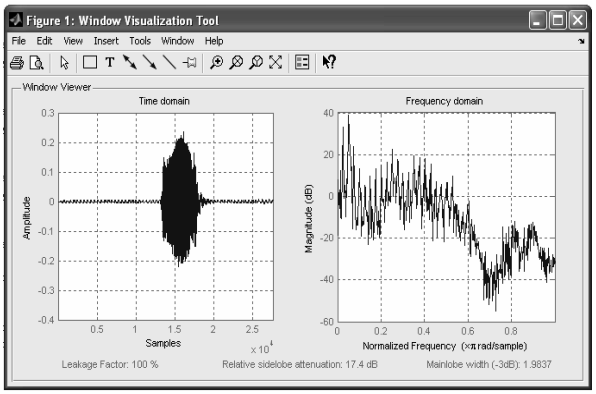

Fig.2: Graphical Representation for Alphabet “A” Pronounced by Male Child

V. CONCLUSION

The aim of this paper is to develop a element for speech recognition system which is useful for children’s identification it is very difficult to differentiate the speech of male & female child of age 4-6. This paper has described Marathwada region children’s speech. The first comprises over 14 hours of read data from 50 children, while the second includes 1 hour and 23 minutes of isolated speech from 30 children. A partition of the data into training, evaluation and test sets has been proposed, and the results of ‘baseline’ speech recognition experiments are under development for future work.

REFERENCES

[1] Collection and recognition of children’s speech in the PF-Star project Mats Blomberg and Daniel Elenius

International Journal of Speech and Language Processing
Volume 1, Issue 1, 2011, pp-07-09
Available online at: http://www.bioinfo.in/contents.php?id=112

[2] Lee, S., Potamianos, A., and Narayanan, S. "Acoustic of children's speech: Developmental changes of temporal and spectral parameters", Journal of Acoust. Soc. Amer., 105(3):1455– 1468, 1999.

[3] Daniel Elenius and Mats Blomberg "Comparing speech recognition for adults and children" Department of Speech, Music and Hearing, KTH, Stockholm Proceedings, FONETIK 2004, Dept. of Linguistics, Stockholm University

[4] Mohamad Adnan Al-Alaoui, Lina Al-kanj, jimmy Azar , & Elias Yaacoub "Speech Recognition Using Artificial Neural Networks & Hidden Markov Models". IEEE Multidisciplinary engineering education Magazine, VOL.,3, No. 3, September 2008.

[5] Avram lev Robinson- Mosher and brain Scassellati Department of Computer Science Yale University "Prosody Recognition in Male Infant-directed speech" Proceedings of 2004 IEEE/ rsj international conference on intelligent Robots & system Sep-28/ Oct-2,2004 Sedni, Japan.

[6] Ingegerd Eklund & Hartmut Traunmuller –Institution for lingvistik, Stockholms university-"A comparative study of male and female whispered and phonated version of the long vowels of Swedish"proceeding THM-QPSR- 2/1996