

Segregation Stuttering Words and Make Continuous Sentence Using Segmentation Method Energy

Mrs.S. Poornima M.Sc(CS), M.Phil

Assistant Professor, Department of Computer Science, VLB Janakiammal College of Arts and Science,
Coimbatore, Tamil Nadu, India

ABSTRACT

Speech processing is an interesting area of signal processing where speaker identification and speaker recognition are widely used applications. Analyzing speech disorder is another significant application of speech processing. Stuttering is a kind of speech disorder start with child age to old age people. Identifying stuttering segment from normal speech signal will be a critical challenge for researchers since detection of this disorder depends on various factors. This research work segregate stuttering words using segmentation method energy.

Keywords : : Categories of Sound, Energy, Threshold

I. INTRODUCTION

Speech processing is a widely growing application of signal processing which is performed to identify the speech signals and learn about their processing method. Speech is a primary way of communication between human beings which is the process of transferring the information from speaker to listener. Human can express their feelings and thoughts through communication. Speech sounds have a various multi-layered temporal-spectral that express acoustic words, intention, expression, intonation, accent, speaker identity, gender, age, style of speaking and emotion [3].

II. SPEECH PROCESSING

Speech processing is used to learn about acoustic signals and their processing methods. It is an exciting application of signal processing which is used to recognize the words, syllables, voiced sounds, unvoiced sounds and silent from speech signals. An acoustic signal is a natural form of communication and also it is way to express human feelings and

thoughts. Speech signals are created by sequence of sounds. This speech sounds are produced from air vibration. This vibrating air is pushed from the lungs through the vocal cords along the vocal track. Acoustic will exist existed when vocal track is expanding to mouth by opening the vocal cord [9].

2.1 Production of speech

Speech is produced by breathing air is excluded from the lungs through the trachea and the oral and nasal cavities. It has mainly four processes between produced speech such as Initiation, phonation, oro-nasal process and articulation [11]. These speech production processes is shown in figure1.

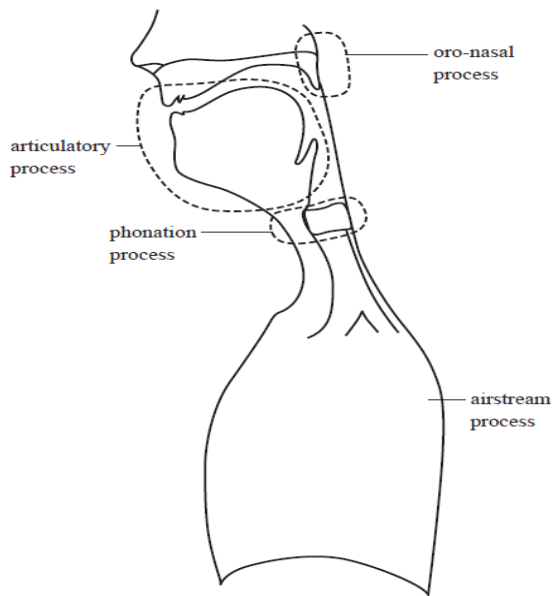


Figure 1. Processes of speech production

Initiation or air stream process is occurred when the breathing air is expelled from the lungs. The phonation process [10] is performing by the larynx. It is shown in figure 2.

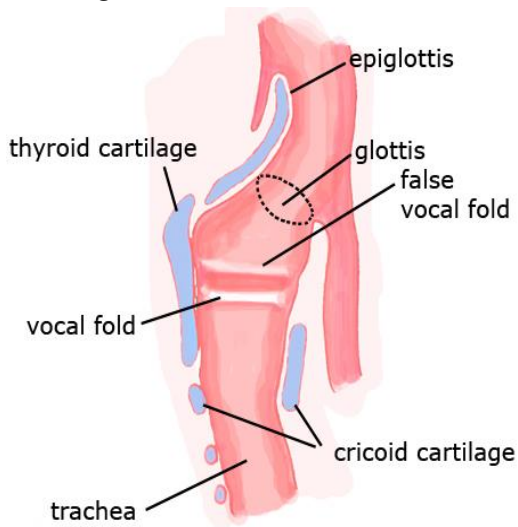


Figure 2. Larynx

The larynx has two horizontal folds of tissue in the way of air pressure called vocal folds. The gap between vocal folds called glottis. A thin opening of glottis can make vibrates on vocal folds which produce “voiced sounds”. The vocal folds vibration is reduced when the glottis have large opening and while producing “unvoiced sounds”. Air will not be passed when glottis is closed. A vibrating air is move into the nasal or oral cavity through larynx and

pharynx. This process is performed by velum. Oro-nasal process is shown in figure 4. The articulation process is final process of speech production which is processed through the mouth. The oral cavity acts as a resonator and the articulator in the mouth which can produce speech sounds from upper and lower lips, upper and lower teeth, tongue and roof of the mouth. The articulation is an important way to differentiate speech sounds [3].

2.2 Categorization of speech sounds

The production of speech sounds can be categorized into four types such as voiced sound, unvoiced sound, plosive sound and nasal sound.

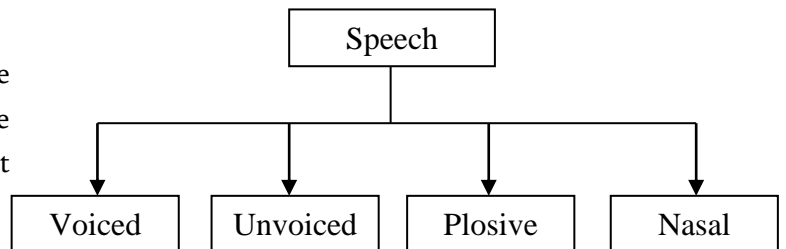


Figure 3. Types of Speech sounds

Voiced sounds are produced when vocal cord correctly vibrates. Example of voiced sounds is /a/, /e/, /i/. Unvoiced sounds are produced when vocal cords are not vibrated. Examples of unvoiced sounds are /s/, /f/. Plosive sounds are produced when the vocal cord completely closure near front of the vocal track. Examples for plosive sound are /p/, /t/. Another speech sound is nasal sound which produced when the sound gets radiation from nostril and lips. Examples of nasal sound are /m/, /n/, /ing/ etc. When air flow is coming to vocal cords from lungs without any block, the vocal cord vibration produces periodic sounds called vowels. But the airflow coming from block to vocal cords produces consonant sounds with or without vocal cord vibration [9].

III. STUTTERING AND SPEECH DISFLUENCY

There are reasons for stuttering speech. The first reason is genetic disorder or heredity. The next

reason is language problem due to stress and nervousness, it will make a delay in communication. It is the reason to get problem in producing acoustic signals. Brain plays a vital role in speech, and it is an important part to controls the complete function of human body. From the medical history, left side of the body is controlled by right side of the brain and right side of the body is controlled by left side of the brain. Speech also controlled and coordinated by human brain. Delivering speech signals having close relationship with cognitive behaviour of the brain. Abnormal or stuttering person brain will have higher impact on speech communication process. Brain disorders like epilepsy and brain disease like brain injury may be an another reason for stuttering issue [8]. Stuttering disfluencies is identified with some reaction of person such as (i) identifying repetition in communication, (ii) sound prolongation and (iii) block or hesitation. These speech disfluency problem includes interjection which is an unwanted extra sounds like “mmm.., ah..,uh, ah..”. Revision represent change in sentence or content for example instead of pronounce and mean a sentence as “It is bird and the name is rosy”, the revision problem may lead the person to change as “It is bird, not the cat name is rosy”. In incomplete phrases sentence or content will not be complete for example, “I want to go...”. Phrase repetition is pronouncing phrase repeatly like “I want..I want..I want.. to buy a pen” and word repetition such as “rose is... is... is...so beautiful” are the significant types of speech disfluency [5]. Prolongation sounds are part of natural talk mostly occurred in communication having consonant or vowels. Broken words are another issue related with speech disorder when person cannot complete some words. Hesitation is an another factor to be considered when speaker have long conversation. After lengthy communication hesitation is occurring with inserting unwanted words like “uh, mmm..” ect. [4].

IV. SEGREGATION WORDS AND MAKE CONTIINUEOUS SENTENCE

Word segmentation is an essential problem in many languages processing. Speech segmentation is used for removing unwanted silence and words from speech signal which is recognizable and meaningful. It includes different type of segmentation level such as language level segmentation, silence level segmentation, word level segmentation, sentence level segmentation. Language level segmentation is to segment a particular language speech from set of spoken languages. Silence level segmentation process which identifies the silence between speeches. Most important segmentation is word level segmentation that is used to segmenting word from speech signal. This segmentation is used to recognizing and removing unwanted words from the speech signal. Sentence level segmentation is similar to word level segmentation but it will segment a set of words from long speech signal [1].

These segmentation levels consists of different speech signal properties for segmenting speech signal such as zero crossing, power spectral density, energy, pitch, stress and intonation pattern. The noisy level Fuzzy logic systems are performed to detect the voiced segment and it is most suitable tool for finding good solution to voiced portion. It has ‘if-then’ systems and their membership function for classification [7]. Spectral Centroid is a frequency domain signal feature which specifies the centre of gravity of a spectrum. In digital signal processing it is used for measurement and it is characterized as a spectrum [41]. Power spectral density (PSD) shows the strength of energy frequency. It is also called Fast Fourier transform, and it is calculated by autocorrelation function. Stuttering speech has more silent and repeat words. It is removed by using speech signal properties and making continuous speech sentence. Autocorrelation function is represented in time domain speech signal future and it is used to identify the repetitions or periodicity in a

signals. Autocorrelation is also known as serial correlation in speech signal [8]. Zero crossing and energy is important properties for voiced and unvoiced segmentation. Zero crossing has high frequency and energy has low frequency when occurring unvoiced speech. Energy contains high frequency in voiced speech. These two properties are proved in segmenting voiced and unvoiced speech [2].

4.1 Energy

Energy is a one of the most important method for word segmentation and it is defined as the strength of the signal. It has higher influence for identifying noise, silent, voiced and unvoiced portion of speech signal. Generally speech signal amplitudes are vary with time. Energy reflect the amplitude variation using the equation

$$E_n = \sum_{m=-inf}^{inf} [x(m)W(n - m)]^2 \tag{1}$$

Where E_n is an energy of speech signal x and W is a window used for analysis. Energy based word segmentation for the speech sentence “Speech Processing is an interesting application” is given in below table 1 and table 2.

Threshold Value = 0.2

Serial No.	Word
1	Speech
2	processing
3	is
4	an
5	interesting
6	application

Table 1. Segmented words

Table 2 shows word segmentation for the speech sentence “I have AAlready completed this work”.

Threshold Value = 0.2

Serial No.	Word
1	I
2	have
3	Already
4	completed
5	this
6	work

Table 2. Segmented words

Prolongation word “processing” and “AAlready” is segmented based on energy and shown in table 3 and table 4.

Threshold value for prolongation word is 0.1 in second

Serial No.	Word
1	Proc
2	Ess
3	Ing

Table 3. Segmented Prolongation word “Processing”
Number of gaps=2

Serial No.	Word
1	A
2	already

Table 4. Segmented prolongation word “AAlready”

Number of gaps=1

Energy has high influence in voiced portion and low influence in voiceless portion. In silent portion energy is too low which is used to identify and remove the silent area [1].

V. CONCLUSION

Speech processing is used to learn about acoustic signals and their processing methods. Energy has high

influence in voiced portion and low influence in voiceless portion. In silent portion energy is too low which is used to identify and remove the silent area. This research successfully segregates stuttering word using energy method.

VI. CONCLUSION

Speech processing is used to learn about acoustic signals and their processing methods. Energy has high influence in voiced portion and low influence in voiceless portion. In silent portion energy is too low which is used to identify and remove the silent area. This research successfully segregates stuttering word using energy method.

VII. REFERENCES

1. AMilton, S.Ashitha Dayana, S.Tamil selvi, "Voiced and unvoiced classification of speech signal using Average Zero Crossing Index Difference Function", International Journal of Advanced Information Science and technology (IJAIST), ISSN: 2319:2682.
2. Bachu R.G., Kopparthi S., Adapa B., Barkana B.D. "Separation of Voiced and Unvoiced using Zero Crossing Rate and Energy of the Speech Signal", Advanced Techniques in Computer Science and Software Engineering 2010, pp:279-282.
3. Fernando Trujillo, "The Production of Speecg Sounds", English Phonetics and Phonology
4. Halina Kwasnicka, Blazei Zak, "Orator- An Intelligent System Supporting Stuttering Speech", ICASC, 2006.
5. Jeremy Kepner, James Lebak, "Software Technologies for High-Performance Parallel Signal Processing", Lincoln Laboratory Journal, Volume 14, Number 2, 2003.
6. Md.Mijanur Rahman, Md.A1-Amin Bhuiyan, "Continuous Bangla Speech Segmentation using Short-term Speech Features Extraction Approaches", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol.3, No. 11, 2012.
7. Paulraj M.P, Sazali Bin Yaacob, Ahamad Nazri Abdullah, Sathees Kumar Natraj, "Segmentation of voice portion for voice pathology classification using Fuzzy logic", Challenges and innovation in information technology, 2010.
8. S Poornima, J. Satheesh Kumar, "Feature Extraction and Signal Classification methods for Stuttering Speech Analysis", International Journal of Modern Computer Science and Applications (IJMCSA) ISSN: 2321-2632 (Online) Volume No.-1, Issue No.-5, November, 2013.
9. Shanthi Therese S., Chrlpa Lingam, "Review of Feature Extraction Technique in Automatic Speech Recognition", International Journal of Scientific Engineering and Technology, Volume no. 2, Issue no. 6, pp: 479-484, ISSN: 2277-1581, June 2013.
10. <http://www.ucumberlands.edu/academics/biology/faculty/kuss/courses/Respiratory%20system/LarynxTrachea.html>
11. <http://kodigz.blogspot.in/2013/11/lj-chapter-1-articulation-and-acoustics.html>