RESEARCHERID
THOMSON REUTERS

ISSN: 2277-9655
[Narayanam * *et al.*, 6(9): September, 2017] Impact Factor: 4.116
IC™ Value: 3.00 CODEN: IJESS7

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

VOICED AND UNVOICED SEPARATION IN SPEECH AUDITORY BRAINSTEM RESPONSES OF HUMAN SUBJECTS USING ZERO CROSSING RATE (ZCR) AND ENERGY OF THE SPEECH SIGNAL

Ranganadh Narayanam*
Assistant Professor, Eece, Its, Ifhe - India

DOI: 10.5281/zenodo.892088

ABSTRACT

In speech signals two activities - voiced & unvoiced - are prominently observed, in both the cases of "production speech" from mouth, and "hearing speech" passed through ears-brainstem-and brain. These speech signals are broadly categorized into these two regions: Voiced- nearly periodic in nature; Unvoiced – random noise like in nature. For many speech applications it is most important to distinguish between Voiced and Unvoiced speech. We have collected Speech Auditory Brainstem Responses (ABR), from healthy human subjects, for "consonant-vowel" stimulus, using single electrode EEG – having Brainstem Speech Evoked Potentials data of voiced and unvoiced combined regions. For this Speech ABR we have proposed two simple & best approaches to detect the Voiced & Unvoiced regions in the EEG Data - first approach is Zero Crossing Rate (ZCR) & Second approach is Short Time Energy (STE). We have collected real time data from 20 different healthy human subjects in an audiology lab of University of Ottawa. We have collected the data at the sampling frequency of 3202 Hz (3.202 KHz). For this research article, we did the Voiced/Unvoiced separation experiment on the Auditory Evoked Potentials (AEP) data for 2 different human subjects. We observed that even for Speech Auditory Brainstem Responses the combined algorithm of ZCR & STE is very good solution for the separation of voiced/unvoiced parts

KEYWORDS: ABR, ZCR, EEG, Auditory Evoked Potentials, Voiced, Unvoiced, Audiology.

I. INTRODUCTION

Speech processing is one of the largest growing research areas in signal processing. Each year billions of pounds are being spent on supporting research in speech processing. The ultimate aim of this research is to provide an interactive man-machine communication [1, 2, 3]. The human apparatus concerned with speech production and perception is complex and uses many important organs - The lungs, mouth, nose, ears controlling muscles and the brain. It is remarkable that this apparatus has developed to enable not only the speech production but also serves other purposes such as breathing or eating. It was discovered that various specific areas in the brain are regarded to be of prime importance for speech and language [2]. These are called the speech centers - damages to any of these areas causes disruption to speech. The vocal tract and vocal cord play a major role in speech production. The vocal tract consists of several organs and muscles which are regularly monitored and carefully controlled by the speech centers. The precise controlling is achieved by internal feedback in the brain [2]. As an example auditory feedback helps us to ensure that we are producing the correct speech sounds and that they are of the correct intensity for the environment. During the past few years, the vast number of research and development in speech processing brought up changes in our everyday life. There are commercially available products which are based on Automatic Speech Recognition, Speaker Verification, Speaker Identification and Speech Synthesizer [5, 7].

A. Motivation

The Scientists and Engineers have understood the basic concepts behind the anatomy and physiology of speech production and perception. **But the lack of understanding of the interaction of the brain with vocal tract and auditory apparatus prevents Engineers from designing machines, which will be able to understand and speak like ordinary human beings [2, 4, 7].** Due to this we found it to be an excellent area of research to find the techniques, algorithms & applications of mouth production speech, whether they are applicable to

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

auditory speech processing or not, and if not, then to find the required modifications to make them applicable to auditory speech processing.

B. Production Speech

Speech is an acoustic signal produced from a speech production system. In many speech analysis/synthesis systems it is critical to make the decision of voiced and unvoiced. Speech can be divided into numerous voiced and unvoiced regions. For many speech processing applications such as speech enhancement, speech synthesis the voiced/unvoiced classification provides a preliminary acoustic segmentation [5, 6].

For the speech production system if the input excitation is nearly periodic impulse sequence, then the corresponding speech looks nearly periodic and is termed as **Voiced speech**. Voiced speech consists of constant frequency tones of some duration, made when vowels are spoken. During the production of voiced speech, the air exhaling out of lungs through the trachea is interrupted periodically by the vibrating vocal folds. Due to this, the glottal wave is generated that excites the speech production system in the voiced speech [6, 8, 9]. About two-thirds of speech is voiced. For intelligibility of speech, voiced speech is most important. Because of its periodic nature, Voiced speech can be identified and extracted.

If the excitation for speech production is random noise like, then the resulting speech will also be random noise-like without any periodic nature and is termed as **Unvoiced Speech**. During the production of unvoiced speech, the air exhaling out of lungs through the trachea is not interrupted by the vibrating vocal folds. However, starting from glottis, somewhere along the length of vocal tract, total or partial closure occurs which results in obstructing air flow completely or narrowly. This modification of airflow results in stop or frication excitation and excites the vocal tract system to produce unvoiced speech [6, 8, 9]. Hence, unvoiced speech is non-periodic, random like sounds, caused by air passing through a narrow constriction of the vocal tract as when consonants are spoken.

C. Hearing speech: Speech Auditory Brainstem Responses Data from Single Electrode EEG.

When speech is heard, three tiny bones called the hammer, anvil, and stapes (or stirrup) in the skull detect eardrum vibrations and pass them on to a snail-shaped organ called the cochlea, which is filled with fluid and tiny hairs called cilia. The sound vibrations make the fluid in the cochlea wash back and forth, agitating the cilia. The cilia detect those vibrations and send electrical signals and excites auditory nerves and it carries the signals to brainstem and to the brain, where it is processed which we hear as sounds of different frequencies [7, 10].

In recent years, the interest of neuroscientists and neural engineers for recording Auditory Brainstem Responses for speech stimuli (Speech ABR) is proliferating manifolds, due to the reason that there is evidence that they are useful in the diagnosis of central auditory processing disorders [11]. The frequency content of natural speech is neither concentrated in time nor in frequency, it takes tens of minutes for recording sufficient quality speech ABR [12]. In our experiment [11, 12, 13], it was required several minutes for recording even with a synthetic consonant-vowel stimulus. Speech ABR is believed to originate in neural activity that is phase locked to the envelope or harmonics of the stimulus. As a result, the recorded responses are remarkably speech-like. In fact, speech ABR is quite intelligible if played back as a sound [14]. As a result, all the methods such as VAD techniques, speech enhancement etc can be applicable to Speech ABR also. In this research we have collected synthetic speech signals of "consonant-vowel sounds" for an experiment in an audiology lab, of university of ottawa. This has been heard by various human subjects under test through earphones (experimenting on one human subject at a time), in a completely quiet sound proof auditory room setup of the audiology lab. During this time the experimental setup of single electrode EEG has been setup properly and Speech Auditory Brainstem Responses data collections have been done successfully from 20 different healthy human subjects [11, 12, 13, 14, 15].

Our Previous Research work on this Single Electrode EEG collected Speech ABR Data:

On these Brainstem Speech Auditory Evoked Potentials data we have done two experiments:

1) **Voice Activity Detection (VAD) in Speech ABR [16, 17, 26]:** In this research we have put into application 3 different VAD algorithms (a) Linear-Interpolation zero-crossing rate algorithm [18] (b) Two statistical algorithms [19] based on a statistical approach that has become the standard for detecting harmonic components in a related evoked response, the auditory steady-state response (ASSR) (c) our own newly proposed Novel VAD algorithm (based on the property that vowels have

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

distinctive spectral peaks) which depends on a binary weighting of the spectral components of the signal under test – Signal to Noise Ratio Peak Valley Difference Detection VAD [16,17,20, 26]. In this research we could be able to detect the Voice in this Speech ABR successfully and SNRPVD VAD is working much better than the standard Statistical algorithms, and could be able to detect even when statistical algorithm could not detect [16, 17, 26], which is a highly novel result in the world of Speech ABR.

2) **Signal to Noise Ratio (SNR) improvement in Speech ABR [21, 22, 26]:** In this research we put into application various filtering techniques such as Yule-Walker multiband filter, Comb filter & cascaded filter of both, Wavelet de-noising filters, FASTICA filter. Then, for further improvement in SNR, we proposed 2 innovative techniques of special interest which were never put into application for Speech ABR signals: a) Translation Invariant Wavelets (TIW) Filtering Technique b) Combined TI Wavelets filter & FASTICA filter – "Cycle Spin TI-ICA – CSTIICA filter". We found TI and CSTIICA methods are working far better than conventional approaches [21, 22, 26], and in two of them CSTIICA is better than TIW [22].

In the research of this paper, the third experimentation of interest is to do the separation of Voiced & Unvoiced parts in the single electrode EEG collected Speech ABR data. In speech processing there has been considerable research done in the recent years in solving the problem of classifying speech into voiced/unvoiced parts. A pattern recognition approach and statistical and non statistical techniques, non parametric methods based on multi-layer feed forward network [23] has been applied for deciding whether the given speech signal segment of a speech signal should be classified as voiced speech or unvoiced speech.

In solving the problem of Voiced/Unvoiced separation in our current research of interest of Neural speech signals of Speech ABR data collected from single electrode EEG, we have chosen a simple and fast approach – Using Zero-Crossing-Rate (ZCR) & Short Time Energy (STE) for voiced and unvoiced separation in the Speech ABR signals. In this we are providing the results for two different subjects to evaluate the performance of the combined Zero-Crossing-Rate (ZCR) & Short Time Energy (STE) based Voiced/Unvoiced separation algorithm [6].

In section 2 ZCR is briefly introduced, section 3 introduces STE, section 4 gives the model system & methodology of the voiced/unvoiced separation algorithm using combined ZCR&STE, section 5 result analysis, with section 6 we conclude the research, and in section 7 future scope of the research is given.

II. ZERO CROSSING RATE (ZCR)

Zero crossing rate is a measure of number of times that the amplitude of the speech signals passes through a value of zero in a given time interval, as in Figure 1. The meaning of "The signal x(n) has a zero-crossing at n_0" is that

$$x(n_0)\, x(n_0+1) < 0.$$

Unvoiced signals oscillate much faster, so they will have a much higher rate of zero-crossings. Zero-Crossing rate is an indicator of the frequency at which the energy is concentrated in the signal spectrum [24].

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

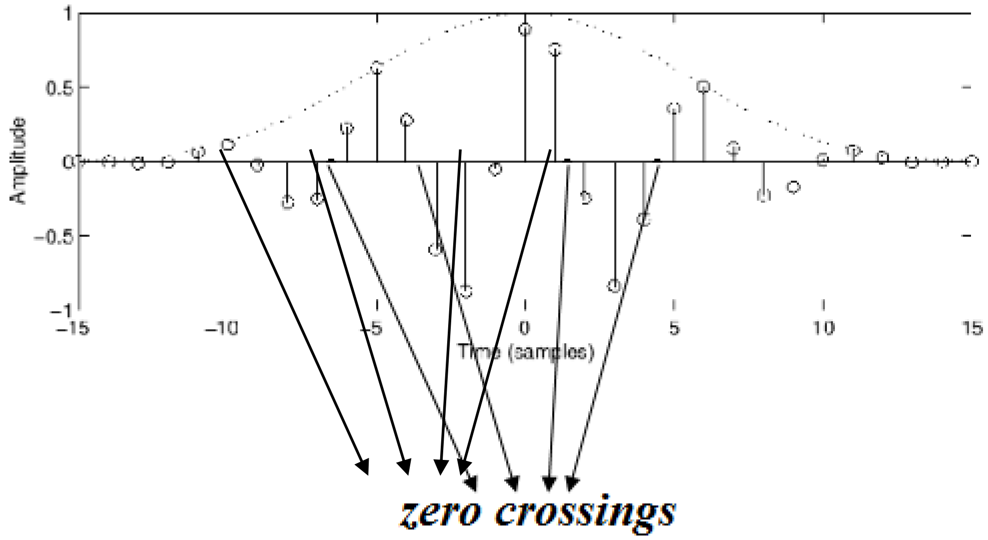

Figure 1 Zero Crossings Definition

A definition for zero-crossings rate is

$$Zn = \sum_{m=-\infty}^{\infty} \left| sgn[x(m)] - sgn[x(m-1)] \right| w(n-m)$$

Where,

$$sgn[x(n)] \begin{cases} = 1 & x(n) \geq 0 \\ = -1 & x(n) < 0 \end{cases}$$

W(n) is the windowing function with a window size of N samples

$$W \begin{cases} = 1/2N & 0 \leq n \leq N-1 \\ = 0 & \text{otherwise} \end{cases}$$

In general, a speech production model spectrum suggests that voiced speech energy is concentrated at lower frequencies, and for unvoiced speech most of the energy found at higher frequencies [6]. As high frequency means high zero crossings rates, and low frequencies means low zero crossing rates. So from this we can correlate strongly between zero-crossing rate and energy distribution with frequency. So, if Zero-crossing rate is high the speech signal is unvoiced, while if the zero-crossing rate is low, the speech signal is voiced [24].

III. SHORT TIME ENERGY (STE)

The amplitude of unvoiced segments is noticeable lower than that of the voiced segments. The short time energy of speech signals reflects the amplitude variation [6]. The speech signal energy represents these amplitude variations. Most of the short time processing techniques that give time domain features (Qn), can be mathematically represented as

$$Qn = \sum_{m=-\infty}^{\infty} T[x(m)]w(n-m)$$

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

Where T[] is the transformation matrix which may be either liner or nonlinear, X(m) represents the data sequence and W(n-m) represents a limited time window sequence (frame duration). The energy of the discrete time signal is defined as

$$E = \sum_{m=-\infty}^{\infty} X^2(m)$$

The short time energy of speech signal provides a convenient representation that reflects the amplitude variation and can be defined as

$$En = \sum_{m=-\infty}^{\infty} [x(m)W(n-m)]^2$$

From this it can be observable that voiced speech part contains much higher energy than unvoiced speech content. This Energy function (En) is highly useful for speech analysis purposes.

The choice of window determines the nature of the short-time energy representation. As hamming window gives much greater attenuation outside the band-pass than the comparable rectangular window, we have chosen Hamming Window in our research. An example framing of the given samples of speech data can be seen in the following Figure 2.

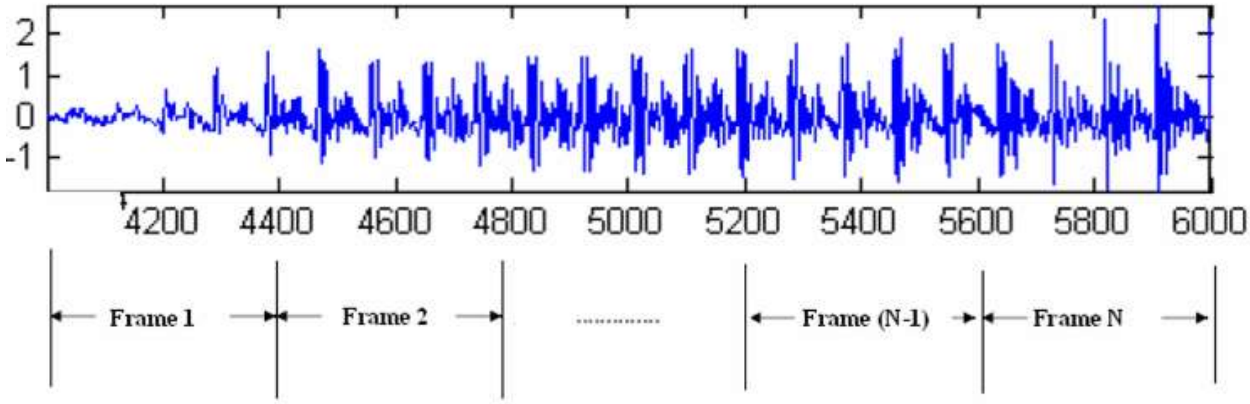

Figure 2 Frame by frame processing of speech signal

Then if we combine both ZCR and STE features and if we make decision, we can make sure of the separation of Voiced and unvoiced parts of the given speech signal.

IV. SYSTEM FOR VOICED & UNVOICED PARTS SEPARATION OF SPEECH AUDITORY BRAINSTEM RESPONSES

The following system is the system which uses the features of both ZCR & STE for solving the problem of Voiced & Unvoiced separation in single electrode EEG collected Auditory Evoked Potentials (AEP) for consonant-vowel stimulus. The data is collected from 20 healthy human subjects. Here we are providing the data and results of 2 different human subjects. One of the important parameters for voiced/unvoiced classification is Zero Crossing Rate, used as front end processing in automatic speech recognition system [6]. As explained in section I, Voiced speech is periodic in nature and hence it has low Zero Crossing Rate (at low frequencies in spectrum), and Unvoiced speech is of noise-like & random in nature with high rate of Zero Crossings and hence it has high Zero Crossing Rate. The Short Time Energy is another parameter for classifying Voiced/Unvoiced which is a parameter in terms of amplitude of the signal. So Energy is high for periodic, high amplitude Voiced part of the speech signal; and Energy is very low for non periodic random noise like Un-voiced part of the speech signal.

In our experiment for EEG collected Speech Auditory Brainstem Responses we have evaluated this combined ZCR & STE algorithm for Voiced & Un-Voiced parts separation. The algorithm is given in following Figure 3.

With consonant-vowel speech stimulus, EEG collected Speech ABR signal is divided into frame by frame. For Exact frame selection of set of samples hamming window has put into application. For each individual frame we found STE and ZCR and if STE is high and ZCR is low then Voiced part, if STE low and ZCR high then Unvoiced part. If in any frame if results are uncertain then we will subdivide the frame and apply the algorithm for sub-frames and make the voiced/unvoiced decision in the main frame.

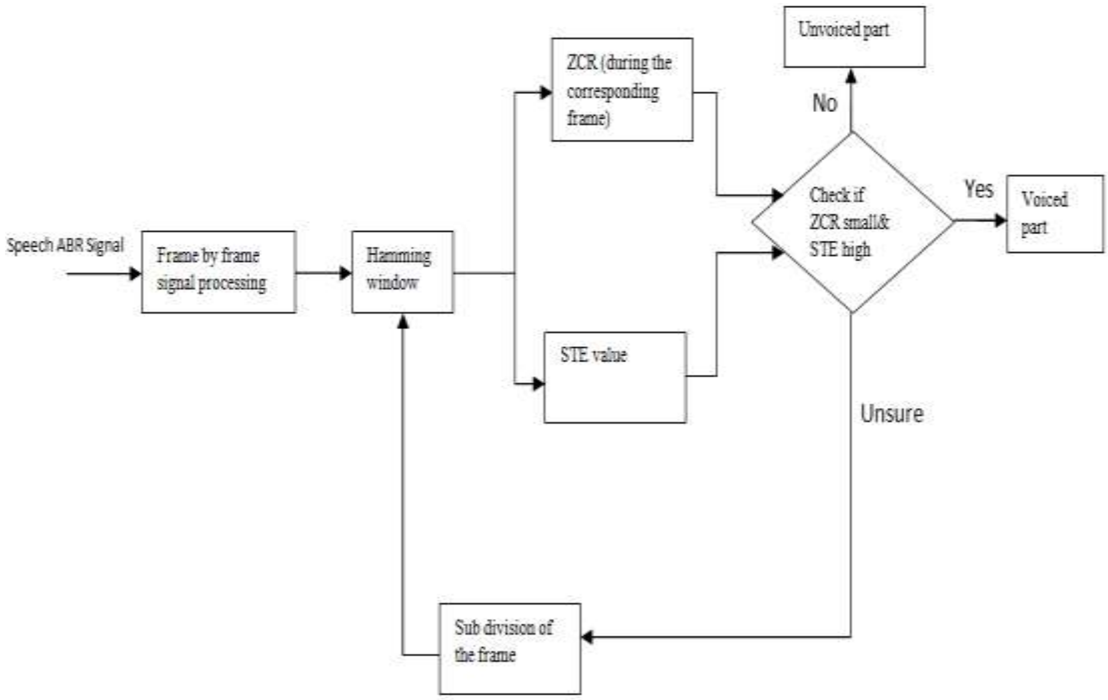

Figure 3 voiced/Unvoiced classification algorithm using ZCR & STE for EEG collected Speech ABR, for "Consonant-Vowel stimulus".

V. RESULT ANALYSIS

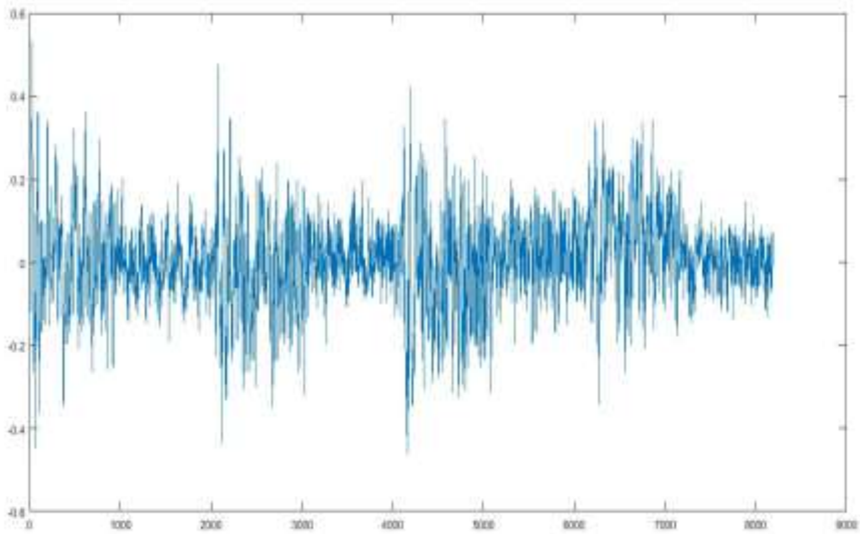

Figure 4 For consonant-vowel stimulus, Single electrode EEG collected Speech ABR data of human subject 1.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

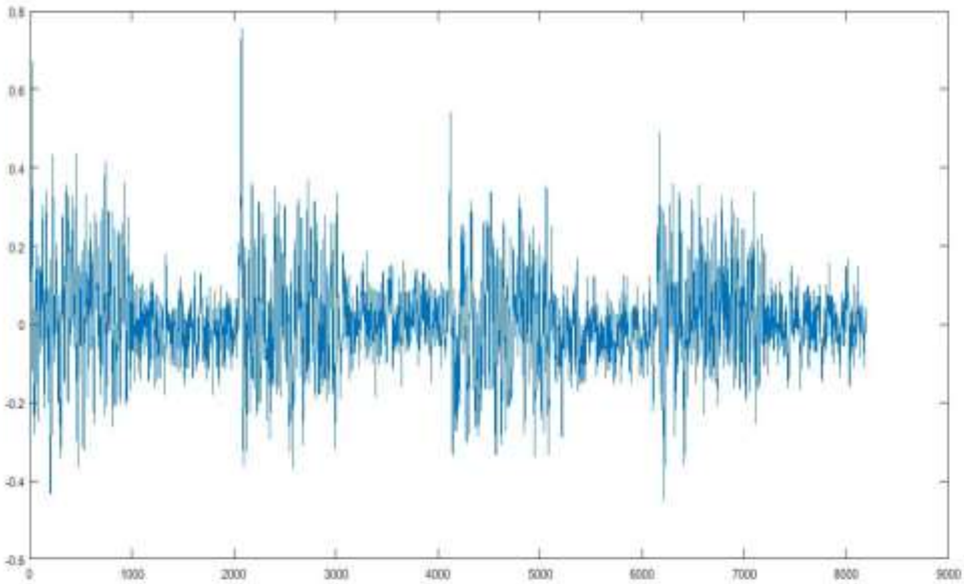

Figure 5 For consonant-vowel stimulus, Single electrode EEG collected Speech ABR of human subject 2.

The data is collected at the sampling rate of 3202 hz (3.202 kHz) for both subject 1 and subject 2. Total of 8192 samples are collected for each subject. But for framing purposes we are using up-to 8000 samples. As per the algorithm, the speech ABR data for each subject is being divided into 400 frames. Total of 20 frames are evaluated for each subject. We have verified in the research of this paper, with ZCR&STE combined algorithm for voiced part & unvoiced part identification in time domain.

The tables Table 1 & Table 2 containing the experimental results of Voiced/Unvoiced separation. As per the combined algorithm of ZCR&STE, the Tables show exactly the locations of Voiced and Unvoiced parts of the Speech ABR data of the 2 human subjects. So this algorithm also very useful in voiced part identification, in addition to Un-Voiced part identification of the Speech ABR signals. So, this combined STE & ZCR is found to be a good tool for Voiced/Unvoiced parts separation in single electrode EEG collected Speech Auditory Brainstem Responses, for consonant-Vowel stimulus. The obtained results are proving that our data is containing Voice Activity, confirming to our experiment 1 of Section 1 [16, 17, 26].

Table 1 Voiced/Unvoiced Decisions using ZCR-STE algorithm for human subject 1. For Consonant-Vowel stimulus, single electrode EEG collected speech Auditory Brainstem Responses for Human subject 1.

Frame number	Samples/frame	ZCR	Energy (J)	Decision
1	400			
	200	21	289.3015	Voiced
	200	58	0.3456	Unvoiced
2	400			
	200	59	0.1871	unvoiced
	200			
	100	25	123.4098	voiced
	100	21	130.4568	voiced
3	400	68	0.01789	unvoiced
4	400	51	0.7896	unvoiced
5	400	49	0.0098	unvoiced
6	400			

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

	200	22	220.4567	voiced
	200	48	0.01235	unvoiced
7	400	79	0.45678	unvoiced
8	400	88	0.20100	unvoiced
9	400	98	0.3468	unvoiced
10	400	109	0.1209	unvoiced
11	400			
	200	19	27.1098	voiced
	200	87	0.00189	unvoiced
12	400			
	200	21	56.9087	voiced
	200	18	49.0965	voiced
13	400	67	0.1789	unvoiced
14	400	70	0.6543	unvoiced
15	400	85	0.0876	unvoiced
16	400			
	200	18	95.1098	voiced
	200	27	189.0987	voiced
17	400	76	0.00987	unvoiced
18	400	78	0.1290	unvoiced
19	400	98	0.0987	unvoiced
20	400	120	0.6789	unvoiced

Table 2 Voiced/Unvoiced Decisions using ZCR-STE algorithm for human subject 2. For Consonant-Vowel stimulus, single electrode EEG collected speech Auditory Brainstem Responses for Human subject 2.

Frame number	Samples/frame	ZCR	Energy (J)	Decision
1	400			
	200	23	210.0981	Voiced
	200	25	100.8709	Voiced
2	400			
	200	29	98.0112	Voiced
	200	19	29.0987	Voiced
3	400	154	0.1789	Unvoiced
4	400	118	0.0987	Unvoiced
5	400	107	0.0154	Unvoiced

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

6	400			
	200	21	12.0123	Voiced
	200	129	0.21345	Unvoiced
7	400	189	0.19871	Unvoiced
8	400	201	0.00987	Unvoiced
9	400	154	0.6543	Unvoiced
10	400	128	0.4567	Unvoiced
11	400			
	200	17	13.0987	Voiced
	200	84	0.09876	Unvoiced
12	400	54	0.17658	Unvoiced
13	400	59	0.23147	Unvoiced
14	400	55	0.34560	Unvoiced
15	400	69	0.21090	Unvoiced
16	400			
	200	23	15.1234	Voiced
	200	101	0.11191	Unvoiced
17	400	75	0.27899	Unvoiced
18	400	60	0.21033	Unvoiced
19	400	71	0.10012	Unvoiced
20	400	99	0.45690	Unvoiced

VI. CONCLUSION

In this research we proposed a novel algorithm of a combination of Zero Crossing Rate (ZCR) & Short Time Energy (STE), for Voiced/Unvoiced separation in Single Electrode EEG collected data of Speech Auditory Brainstem Responses, for Consonant-Vowel stimulus. We could be able to find the Voiced and Unvoiced parts of the data, also confirming to our experiment of VAD in speech ABR in section 1 [16, 17, 26]. We found this approach is highly useful in making the decision even in the case of neurological speech signals also, which is relatively new approach of application for Auditory Evoked Potential data.

VII. FUTURE RESEARCH

As a future research, we would like to add some very high noise levels to the data and observe the applicability of this algorithm. In such case of very high levels of noise, we would like to apply various de-noising filters for these Brainstem Speech Evoked Potentials, and then we would like to verify how this ZCR&STE combined algorithm will work. Then we would like to compare the locations of Voiced/Unvoiced decisions before de-noising & after de-noising, with those of the locations of the Voiced/Unvoiced decisions of the results of the current research paper. Then we can check the efficacy of this algorithm.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

VIII. REFERENCES

[1] Thmoas W.Parsons, “Voice and Speech Processing”, McGraw-Hill, Inc. 1986.
[2] W.J.Hardcastle, “Physiology of Speech production”, Academic press Inc (London).1976.
[3] Sadaoki Furui, “Digital Speech Processing, Synthesis and Recognition”, Marcel Dekker, Inc.1989.
[4] Mc Caffrey, Patrick, CMSD 620 Neuroanatomy of speech, Swallowing and language (Neuroscience on the web, California State university, Chico, Feb 2009).
[5] Jaber Marvan, “Voice Activity detection Method and Apparatus for voiced/unvoiced decision and Pitch Estimation in a Noisy speech feature extraction”, 08/23/2007, United States Patent 20070198251.
[6] D.S.Shete, Prof. S.B. Patil,Prof. S.B. Patil, “Zero crossing rate and Energy of the Speech Signal of Devanagari Script”, IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 4, Issue 1, Ver. I (Jan. 2014)
[7] Harish Chander Mahendru, “Quick Review of Human Speech Production Mechanism”, International Journal of Engineering Research and Development, Volume 9, Issue 10 (January 2014), PP. 48-54.
[8] Inge Timmers, Bernadette M. Jansma, M. Estela Rubio-Gozalbo, “From Mind to Mouth: Event Related Potentials of Sentence Production in Classic Galactosemia”, Journal of PLOS neglected tropical diseases, December, 2016.
[9] Rupali V. Pawar, Dr. R.M. Jalnekar “Review on Speech Production Model”, International Journal of Engineering and Innovative Technology (IJEIT) Volume 3, Issue 9, March 2014
[10] JERKER RÖNNBERG, “Hearing With Your Ears, Listening With Your Brain”, APS association for psychological science, 2015
[11] – 7 : Johnson, K.L., Nicol, G.T., Kraus, N. 2005. Brain Stem Response to Speech: A Biological Marker of Auditory Processing. Ear & Hearing 26, 424-434.
[12] – 6 Dajani, R.H., Purcell, D., Wong, W., Kunov, H., Picton, T.W. 2005. Recording Human Evoked Potentials That Follow the Pitch Contour of a Natural Vowel. IEEE Transactions on Biomedical Engineering 52, 1614-1618.
[13] – 8 Russo, N., Nicol, T., Musacchia, G., Kraus, N. 2004. Brainstem responses to speech syllables. Clinical Neurophysiology 115, 2021-2030.
[14] – 9 Galbraith GC, Arbagey PW, Branski R. Intelligible speech encoded in the human brain stem frequency-following response. NeuroReport 1995; 6: 2363-2367.
[15] – 10 M.S. John, T.W. Picton, MASTER: a Windows program for recording multipleauditory steady-state response. Computer Methods and Programs in Biomedicine 61 (2000) 125–150, Elsevier.
[16] Ranganadh Narayanam, “Robust detection of speech auditory brainstem responses using Voice Activity Detection (VAD) algorithms”, IEEE CAMAN international conference, 2012
[17] Ranganadh Narayanam, “AN EFFICIENT PEAK VALLEY DETECTION BASED VAD ALGORITHM FOR ROBUST DETECTION OF SPEECH AUDITORY BRAINSTEM RESPONSES”, CS & IT-CSCP 2013.
[18] – 5 GBrson Eduardo Mog, Eduardo Parente kbeiro, “Zero Crossing determination by linear interpolation of sampled sinusoidal signal.
[19] – 10 R. Nicole, J. Sohn, N.S. Kim, and W. Sung, “A statistical Model Based Voice Activity Detection,” IEEE Signal Process. Lett., vol. 6, 1999, pp. 1-3.
[20] In-Chul Yoo and Dongsuk Yook, “Robust voice activity detection using the spectral spectral peaks of vowel sounds”. ETRI Journal,Volume 31,Number 4,August 2009.
[21] ranganadh narayanam, “Brain-Activity-Filters: Efficient performance of Translation-Invariant (TI) Wavelets approach for Speech-Auditory Brainstem Responses of human subjects”, International Journal of Scientific and Research Publications, Volume 4, Issue 9, September 2014.
[22] ranganadh narayanam, “Efficient De-noising Performance of a Combined Algorithm of Translation Invariant (TI) Wavelets and Independent Component Analysis over TI Wavelets for Speech-Auditory Brainstem Responses”, Elsevier international journal Procedia Computer Science Volume 54, 2015, Pages 829-837, Sciencedirect.
[23] Y. Qi, and B.R. Hunt, “Voiced-Unvoiced-Silence Classifications of Speech using Hybrid Features and a Network Classifier,” IEEE Trans. Speech Audio Processing, vol. 1 No. 2, pp. 250-255, 1993.
[24] Bachu R.G., Kopparthi S., Adapa B., Barkana B.D., "Voiced/Unvoiced Decision for Speech Signals Based on Zero-Crossing Rate and Energy", IEEE International Joint Conferences on Computer, Information, and Systems Sciences, and Engineering (CISSE'2008).
[25] Rabiner, L. R., and Schafer, R. W., Digital Processing of Speech Signals, Englewood Cliffs, New Jersey, Prentice Hall, 512-ISBN-13:9780132136037, 1978.

ISSN: 2277-9655
Impact Factor: 4.116
CODEN: IJESS7

[26] Ranganadh Narayanam, “Developing ‘standard novel ‘VAD’ technique’ and ‘noise free signals’ for speech auditory brainstem responses for human subjects”, Thomson Reuters ENDNOTE – ResearcherID, IJESRT international journal, 5(6), June 2016..

CITE AN ARTICLE

Narayanam, R. (2017). VOICED AND UNVOICED SEPARATION IN SPEECH AUDITORY BRAINSTEM RESPONSES OF HUMAN SUBJECTS USING ZERO CROSSING RATE (ZCR) AND ENERGY OF THE SPEECH SIGNAL. *INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY*, *6*(9), 370-380. Retrieved September 15, 2017.