ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUL1771

CNC Mechanical Machine and Musical Sound Analysis of Zero Crossing Rates (ZCR) by Artificial Intelligence Based Tools.

Islam Md Shafikul
Schwartzburg Leonid Efraimovich Kamal Joshi

Abstract:- In our regular lives, sound plays an important role on various sides. There is a valuable effect on communications, emotions, and affections. Humans and animals are not the only sources of sounds. Machines and engines also generate a wide range of sounds. Every sound has different characteristics according to its internal format. Sound source and production method are the key factors in these differences. In our article, we showed the differences in zero crossing rates between mechanical machines (CNC milling) and music sounds using the artificial intelligence-based tool LibROSA. At the end of the results, we estimate that the human or musical voice has a lower zero crossing rate than mechanical machine sounds.

Keywords:- *Sound Analysis, CNC Milling Machine, Artificial Intelligence, Sound Zero Crossing Rate (ZCR).*

I. INTRODUCTION

In recent years, it has been demonstrated that zero-crossing rates (ZCR) of sound waves are useful for speech sound segmentation, analysis, and recognition. ZCR measures appear to be less speaker dependent than spectrum data, are more suitable for digital processing, and are essentially independent of talker volume [1]. By employing the ZCR, Peterson [2] calculated the resonance frequency of a periodic signal that an RLC circuit was running on. ZCR and amplitude of sound signals were used by Reddy [3]– [5] for segmentation and for the first and last categorization of the segmented sounds.

Reddy and Vincens [6] used the ZCR and amplitude of low-pass and high-pass filtered audio to segment speech. Bezdel and Chandler [7] used the ZCR of speech to classify five different vowels. Subsequently, Bezdel and Bridleused the ZCR of low-pass and high-pass filtered speech to recognize the numbers 1 through 9 with 90% accuracy. With some degree of success, Scan [SI] was able to connect the ZCR of voiced to unvoiced formant frequencies by approximating the number of times a deterministic periodic signal crosses zero in a given time interval.

The research described in this work aimed to further explore the usage of ZCR of unfiltered, filtered, and differentiated speech waveforms for analysis and recognition, as well as to express ZCR in terms of the more often used spectral characterizations. In terms of different sound waveforms, ZCR is expressed as a sample function from an ergodic random process. When some contextual environment information is available, the ZCR has been shown to be useful for classifying unvoiced and noise. The density of amplitude frequency, figure out the sound's representations of mel spectrogram. That are related to zero crossing rate of sounds over time.

II. RESEARCH METHODS

In the research, there are combined zero crossings rate and a mel spectrogram. Zero-crossing rate is an important parameter for used as a part of the front-end processing in automatic speech recognition system. The zero-crossing count is an indicator of the frequency at which the energy is concentrated in the signal spectrum. Voiced speech is produced low zero-crossing count [9], whereas the unvoiced speech is produced high zero-crossing count. Mel spectrogram is a variation of the spectrogram that is commonly used in speech processing and machine learning tasks. It is similar to a spectrogram in that it shows the frequency content of an audio signal over time, but on a different frequency axis.

- *Implementation of a Block Diagram:*

In our research, there are input, analysis and output steps. For input we use microphone for sounds recording around one minute as wav format. Then we use that sounds for analysis on the LibROSA software tools. LibROSA is a widely used better tools for sounds analysis based on Artificial intelligence. At the end of the analysis figure out the output results.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUL1771

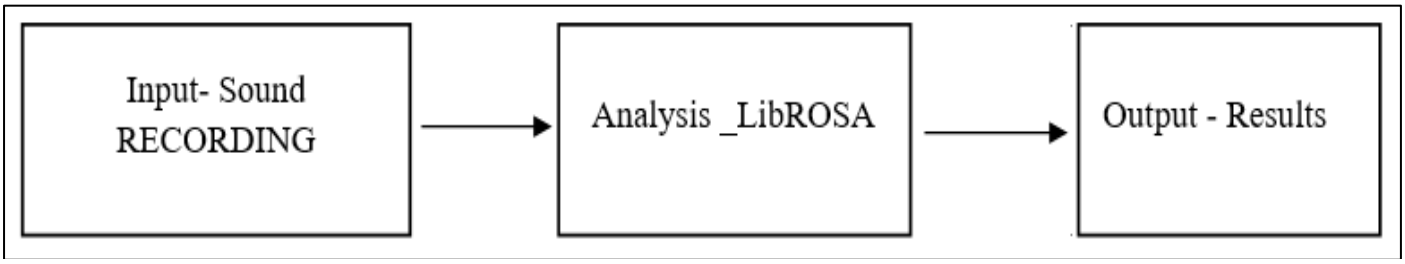

Fig 1: Block Diagram of the Process

- **Mechanical machine (CNC milling):** During the cutting of Aluminium, CNC milling machine produces sound. For collecting data for aluminium cutting, We recorded aluminium cutting sound from CNC milling machine by the systems sound recorder as wav format. Besides, the sound implements in artificial intelligence-based machine learning algorithms for beat retrievals by the LibROSA tools. Then find out frequency of time frame (amplitude), Zero Crossing rate of sounds per milliseconds, sample rate of sounds length and Time duration of sounds.

Fig 2: Aluminium Cutting Image of CNC Milling Machine

The information is displayed in the following tables. There are 6 different frequency data collecting from 1 – 4000, 4000 – 8000, 8000 – 12000, 12000 – 16000, 16000 – 20000 and 1 – 1300000 Hz. Subsequently, find out different zero crossing rate according to 771, 801, 645, 613, 622, 380869. Where the Sample rate of sounds is same 22050 and time duration 61 seconds. According to total sound length our frequency of time frame amplitude is 1345050 and total zero crossing rate 427071.

Table 1: CNC Milling Machine Sound Information Data

Frequency of time frame (amplitude) Hz	Zero Crossing rate (ZCR)	Sample Rate	Time Duration(seconds)
1 - 4000	771	22050	62 seconds
4000 - 8000	801		
8000 - 12000	645		
12000 - 16000	613		
16000 - 20000	622		
1 - 1200000	380869		
Total = 1345050 Hz	Total = 427071		

In this article, standard frequency of time frame amplitude 1 to 1200000 Hz is accepted from the whole frequency length. The figure for CNC milling aluminium cutting sound below according to amplitude (Hz) and time.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUL1771

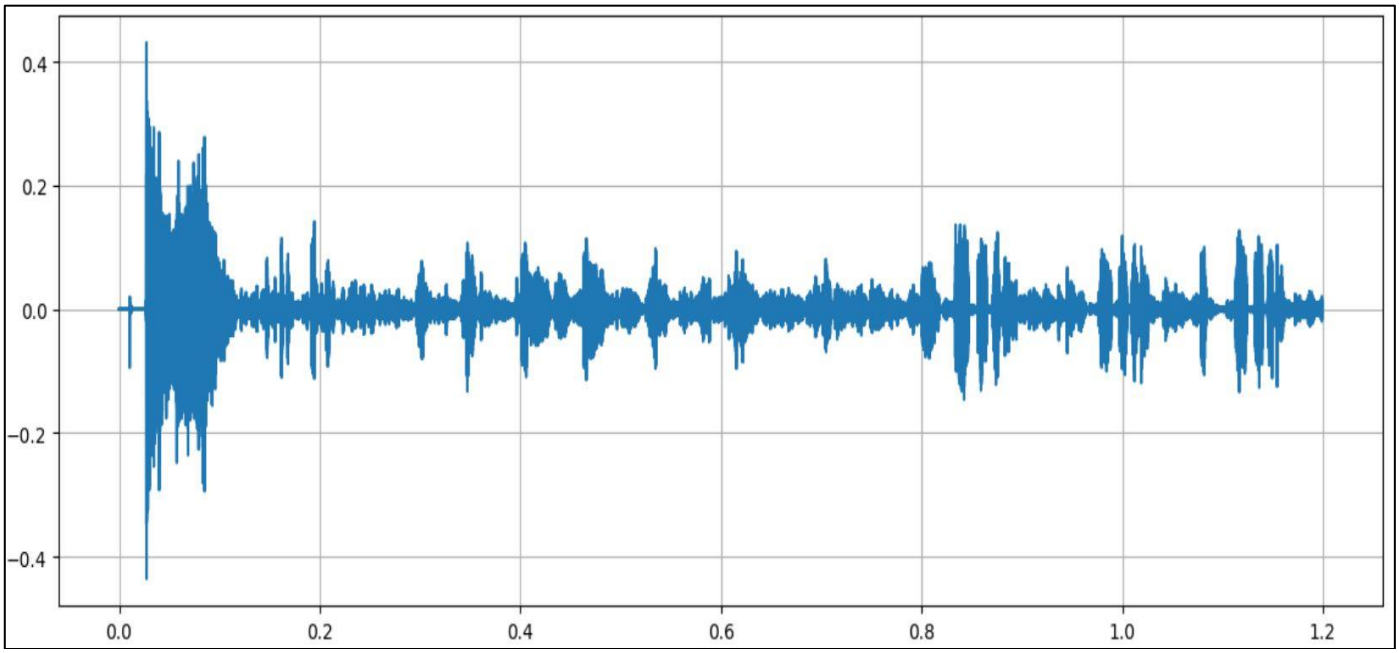

Fig 3: Frequency of Time Frame Amplitude of Aluminium Cutting Sound Vertical (Hz) vs Horizontal (Time)

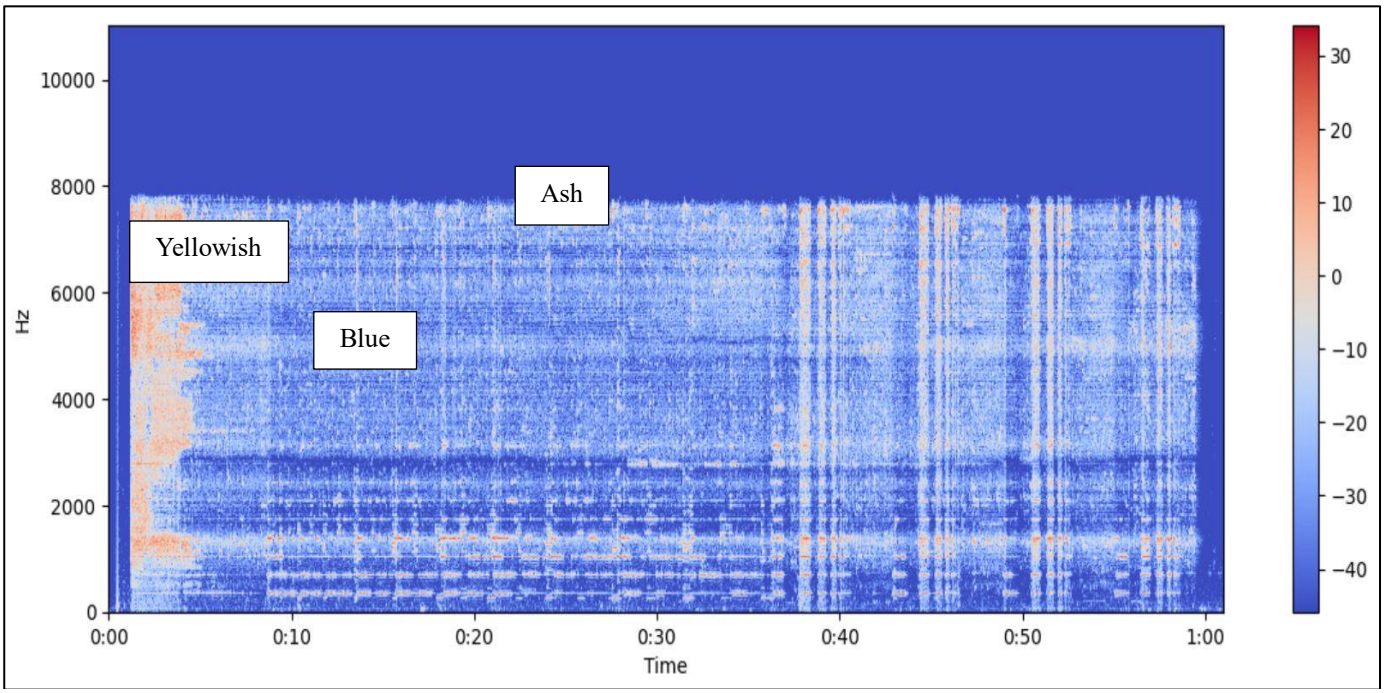

Fig 4: Mel Spectrogram of Aluminium Cutting Sound.

In the mel spectrogram figure of aluminium cutting sound showed, there are different colour shadow such as yellowish, blue and ash. The yellowish shadow indicates the high sounds density. The ash colour indicates that there are sounds available there but les density. The blue colour indicates that there is very low sounds density and some time there is no sounds. In this figure we noticed the difference among high, medium and low sound frequencies of density. The figure is generated sounds amplitude verses over time duration.

- **Music Sound:** For collecting data for music sound, we recorded musical and vocal sounds from royalty-free music on our sound recorder in wav format.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUL1771

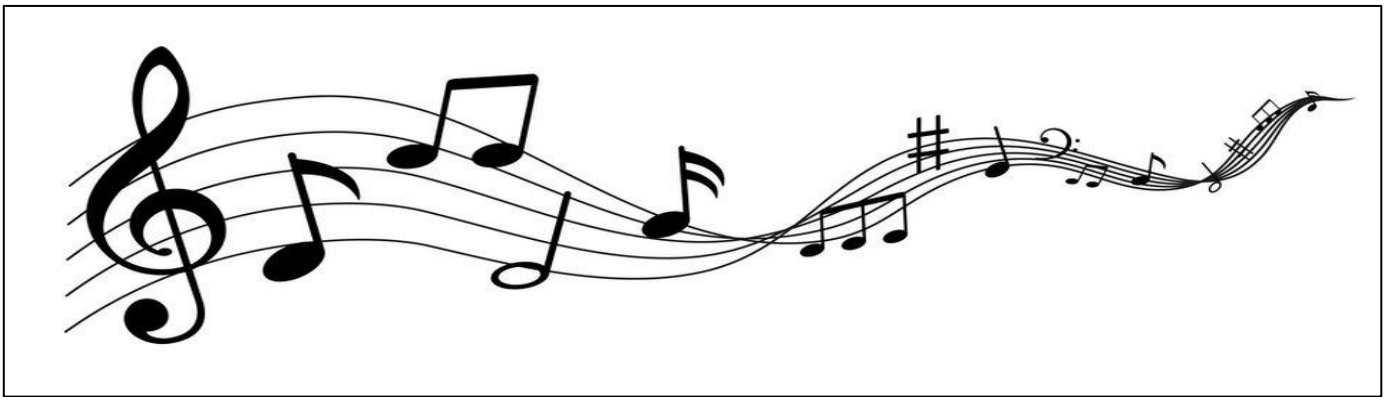

Fig 5: Music Sounds

Then implement artificial intelligence-based machine learning algorithms for beat retrievals using the LibROSA tools. Hereafter, find out the frequency of the time frame (amplitude), the zero-crossing rate of sounds per millisecond, the sample rate of sound length, and the time duration of sounds. The information is bellowing tables are shown.

In the same way, there are 6 different frequency data points collected from 1–4000, 4–8000, 8000–12000, 12–16000, 16–20000, and 1–1300000 Hz. Subsequently, find out the different zero crossing rates according to 1306, 673, 683, 1043, 1030, and 294541. Where the sample rate of sounds is the same 22050 and the time duration is 65 seconds. According to total sound length, our frequency of time frame amplitude is 1345050 and our total zero crossing rate is 427071.

Table 2: Music Sound Information Data

Frequency of Time Frame (Amplitude) Hz	Zero Crossing Rate (ZCR)	Sample Rate	Time Duration (Seconds)
1 - 4000	1306	22050	65 seconds
4000 - 8000	673		
8000 - 12000	683		
12000 - 16000	1043		
16000 - 20000	1030		
1 - 1200000	294541		
Total = 1653750 Hz	Total = 396785		

From the total length of frequency, in this article, a standard frequency of time frame amplitude 1 to 1200000 Hz is accepted. Afterwards, get the figure for the music sound below according to amplitude (Hz) and time.

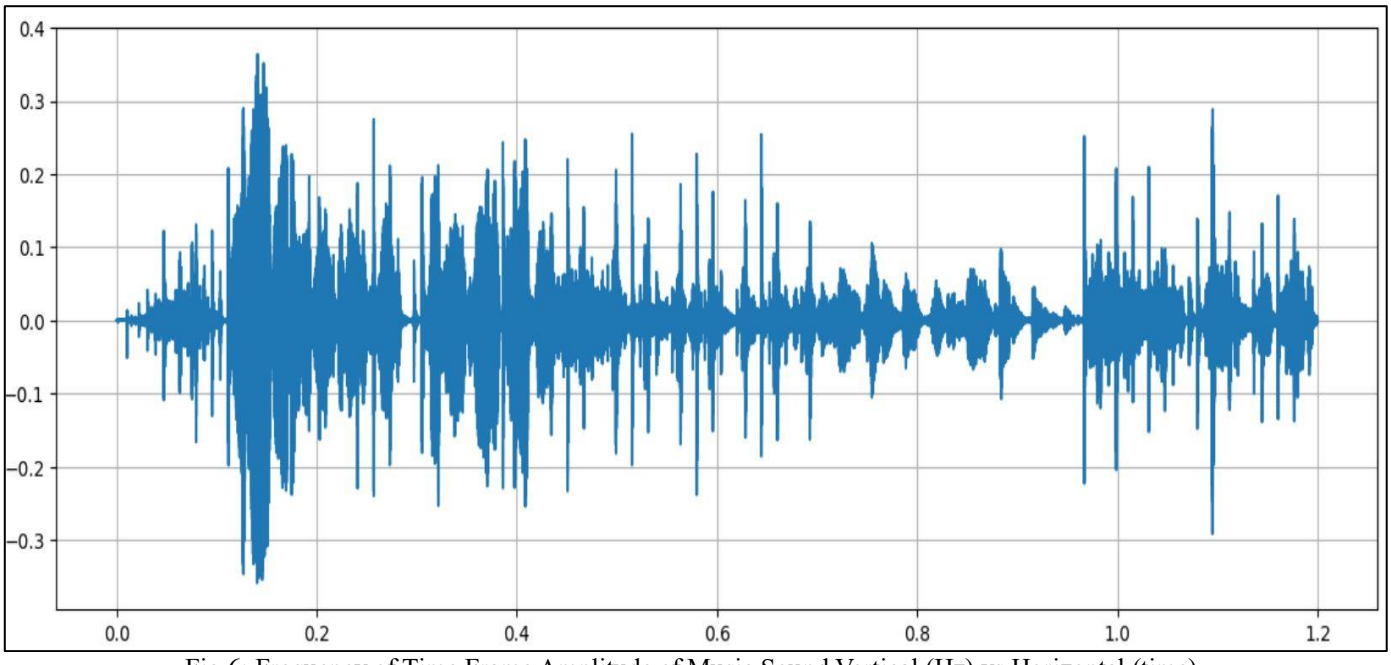

Fig 6: Frequency of Time Frame Amplitude of Music Sound Vertical (Hz) vs Horizontal (time).

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUL1771

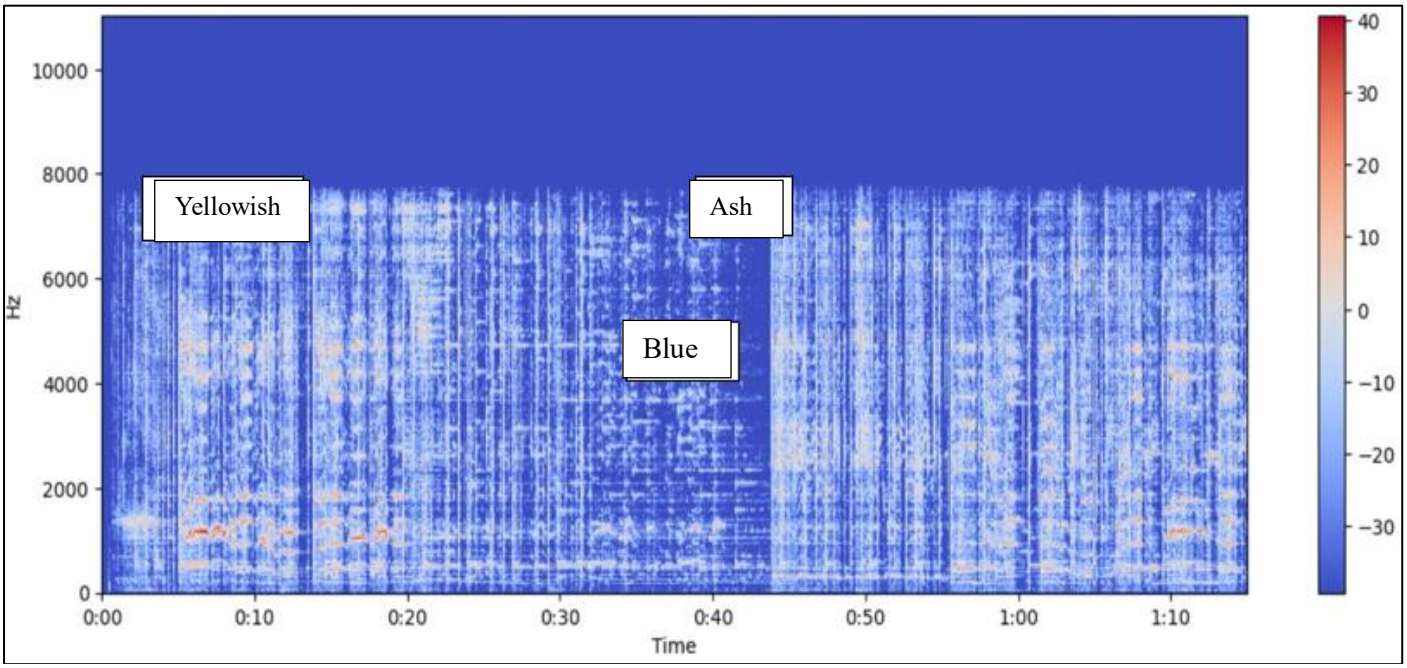

Fig 7: Mel Spectrogram of Music Sound

In the mel spectrogram figure of music sound, there are different colour shadows, such as yellowish, blue, and ash. The yellowish shadow indicates the high sound density. The ash colour indicates that there are sounds available there, but with less density. The blue colour indicates that there is very low sound density, and sometimes there are no sounds. If we notice carefully, we can see the balanced colour representations of the music sounds. The density of sounds is not very high or low during this time. This figure is generated sound amplitude versus time duration.

A very simple way for measuring smoothness of a signal is to calculate the number of zero-crossing within a segment of that signal. A voice signal oscillates slowly - for example, a 100 Hz signal will cross zero 100 per second - whereas an unvoiced fricative can have 3000 zero crossing per second. An implementation of the zero-crossing for a signal x_h at window k is

$$ZCR_{\bar{k}} \sum_{h=kM}^{kM} |sign(x_h) - sign(x_h - 1)|$$

where M is the step between analysis windows and N the analysis window length.

To calculate of the zero-crossing rate of a signal you need to compare the sign of each pair of consecutive samples. In other words, for a length N signal you need O(N) operations. Such calculations are also extremely simple to implement, which makes the zero-crossing rate an attractive measure for low-complexity applications.

III. CONCLUSION

According to our analysis of two different sounds, we can see and measure the results depending on the data number of the zero-crossing rate. In the data table, there is a difference of 4000 Hz in frame amplitude. And there are noticeable differences. We can see the zero-crossing rate is different at different frequency levels. But when you measure the 1 to 1200000 frame amplitude, we get the zero-crossing rate 380869 for CNC aluminium cutting machine sounds and 294541 zero-crossing rates for musical sounds. According to my article, it's very clear that the human or musical voice has a lower zero crossing rate than mechanical machine sounds. Simultaneously, the mel spectrogram represents the balanced sound amplitude over time instead of a more fluctuating density.

REFERENCES

[1]. J. K. Lee, C. D. Yoo, "Wavelet speech enhancementbased on voiced/unvoiced decision", Korea AdvancedInstitute of Science and Technology The 32nd International Congress and Exposition on NoiseControl Engineering, Jeju International ConventionCenter, Seogwipo, Korea, August 25-28, 2003.

[2]. B. Atal, and L. Rabiner, "A Pattern RecognitionApproach to Voiced-Unvoiced-Silence Classificationwith Applications to Speech Recognition," IEEETrans. On ASSP, vol. ASSP-24, pp. 201-212, 1976.

[3]. S. Ahmadi, and A.S. Spanias, "Cepstrum-Based PitchDetection using a New Statistical V/UV ClassificationAlgorithm," IEEE Trans. Speech Audio Processing,vol. 7 No. 3, pp. 333-338, 1999.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24JUL1771

[4]. Y. Qi, and B.R. Hunt, "Voiced-Unvoiced-SilenceClassifications of Speech using Hybrid Features and aNetwork Classifier," IEEE Trans. Speech AudioProcessing, vol. 1 No. 2, pp. 250-255, 1993.
[5]. L. Siegel, "A Procedure for using Pattern ClassificationTechniques to obtain a Voiced/Unvoiced Classifier",IEEE Trans. on ASSP, vol. ASSP-27, pp. 83- 88, 1979.
[6]. T.L. Burrows, "Speech Processing with Linear andNeural Network Models", Ph.D. thesis, CambridgeUniversity Engineering Department, U.K., 1996.
[7]. D.G. Childers, M. Hahn, and J.N. Larar, "Silent andVoiced/Unvoiced/Mixed Excitation (Four-Way)Classification of Speech," IEEE Trans. on ASSP, vol. 37 No. 11, pp. 1771-1774, 1989.
[8]. J. K. Shah, A. N. Iyer, B. Y. Smolenski, and R. E.Yantorno "Robust voiced/unvoiced classification usingnovel features and Gaussian Mixture model", SpeechProcessing Lab., ECE Dept., Temple University, 1947N 12th St., Philadelphia, PA 19122-6077, USA.
[9]. J. Marvan, "Voice Activity detection Method andApparatus for voiced/unvoiced decision and PitchEstimation in a Noisy speech feature extraction",08/23/2007, United States Patent 20070198251.
[10]. T. F. Quatieri, Discrete-Time Speech SignalProcessing: Principles and Practice, MIT LincolnLaboratory, Lexington, Massachusetts, Prentice Hall,2001, ISBN-13:9780132429429.
[11]. F.J. Owens, Signal Processing of Speech, McGrawHill, Inc., 1993, ISBN-0-07-04795550.
[12]. L. R. Rabiner, and R. W. Schafer, Digital Processing ofSpeech Signals, Englewood Cliffs, New Jersey,Prentice Hall, 512-ISBN-13:9780132136037, 1978.