International Conference on Biomedical Engineering (Biomed) 2006
Kuala Lumpur, Malaysia, December 11-14,2006.

DETECTING THE VOCAL DISORDER BY EXTRACTING THE PITCH IN THE PHONETICS OF INDIAN REGIONAL MARATHI LANGUAGE NUMERICAL

Pramod B. Patil Dr. V. T. Ingole
Research Scholar Director
PRM Institute of Technology and Research, Badnera, Amravati, India

Abstract:

The paper presents detection of the vocal disorder suffered due to the reaction of antibiotics during the course of medical treatment by extracting pitch information of the speech. Extraction of pitch of the speech is an important task due to the presence of background noise. Primarily start and end points of speech is detected and thereafter pitch boundaries are recognized using autocorrelation technique. This paper emphasized on accurate end point analysis for detection of the vocal disorder suffered due to the reaction of antibiotics during the course of treatment by extracting pitch information in the phonetics of Indian regional Marathi language numerical.

Keywords: *Disorder, End Point, Pitch Information, and Autocorrelation*

I. Introduction:

In the production of speech, there are a number of sources that generate acoustic energy in the vocal track, periodic sources include aspiration at the glottis; frication, generated further forward in the vocal track; and transient bursts produced by the rapid release of complete constrictions. The periodic source in the speech is created by vibration of the vocal folds creating periodic energy in the glottis. These sources are filtered by the vocal track to generate an output signal. [1]-[4]. Due to the reaction of the antibiotics during the course of medical treatment, patient suffers vocal disorder affecting the speech. The vocal disorder is detected by extracting pitch of the speech. The samples of the phonetics of Indian regional Marathi language numerical are used for extracting the pitch of the speech.

II. Detecting the speech

An important problem in speech processing is to detect the presence of speech in a background noise. [5][6] The accurate detection of words start & end points allows the subsequent processing of the desired data, which reduces the memory requirement & reduces the processing time. The start & end points detection of speech improve the detection of vocal disorder.[7][8]

The method uses two measures of the signal – the zero crossing rate and the energy.

Three thresholds are computed

1. ITU – Upper energy threshold.
2. ITL – Lower Energy Threshold.
3. IZCT- Zero crossings rate threshold.

Algorithm is as follows:

1. Search from the beginning until the energy crosses ITU.
2. Back off towards the signal beginning until the first point at which the energy falls below ITL is reached results into provisional beginning N1.
3. N2 (the end point) is selected in similar way.
4. For beginning point, examine the previous 250ms of the signal's zero crossing rates.
5. If the measure exceeds the IZCT threshold, N1 is moved to the first point at which IZCT threshold is exceeded.
6. Perform similar method for the end point N2.

Fig 1 shows the flowchart for detecting start & end points of speech.

III. Method

The phonetics of Indian regional Marathi language numerical samples for patients before and after the course of medical treatment is recorded. At every frame, the method for detecting the pitch boundaries is important, since detecting start and end point is very critical. The pitch period is estimated by autocorrelation technique by detecting the peak amplitude in the function. The second largest peak relative to the true delay of the correlation is used to determine the pitch period. The peak in the autocorrelation function tends to fall off linearly starting from the first peak [9]-[12]

III. Methodology:

The phonetics of Indian regional Marathi language numerical samples of patients before and after the course of medical treatment is recorded. The speech data is recorded at 8KHz sampling frequency with 8 bits representing each sample. Figure 2 shows block-processing model for detecting the pitch boundaries.

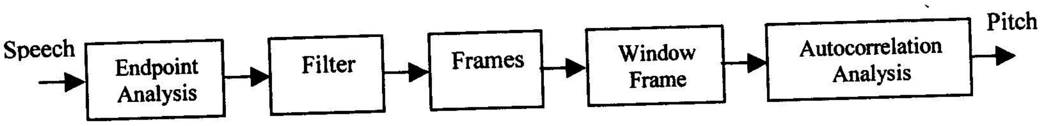

Fig 2. Block Processing Model for Detecting Pitch Boundaries to control the Robot Activities

The data is processed in frame of 200 samples and a vector of pitch is measured. The vector is obtained by pre-emphasizing speech data to flatten it using first order digital filter with coefficient a= 0.97.The signal is processed in the steps of 100 samples to provide smoothing between vectors of pitch. For pitch extraction, the hamming window and autocorrelation technique with order of 10 is used. In the start & End point analysis of speech, each window samples are classified as voiced speech, unvoiced speech or silence, based entirely on measurements made on the signal during the prescribed interval [7]-[11].

IV. Result:

The results with the samples of the phonetics of Indian regional Marathi language numerical //ek// for patients are obtained. The autocorrelation system order selected is equal to 10 for analyzing the better performance. Fig 2,3 shows the start & end point analysis for detection of the vocal disorder by extracting pitch information using autocorrelation technique. Table 1 shows the start and end frames of spoken word "ek".

Spoken Word= //ek//	Starting Frame	Ending Frame
Normal	26	50
Disorder	32	80

Table 1. Showing the Start & End frames of spoken word //ek//

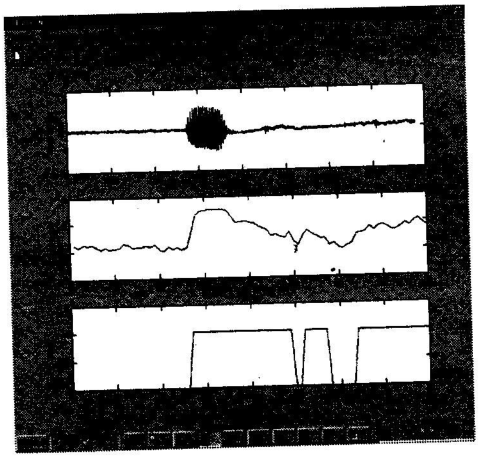

Fig 2 Pitch for the Normal with//ek//

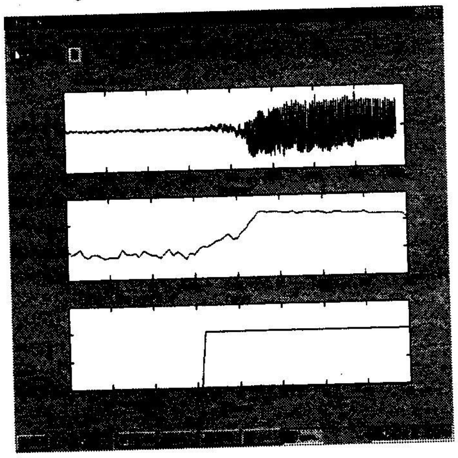

Fig 3 Pitch for disorder detected in//ek//

International Conference on Biomedical Engineering (Biomed) 2006
Kuala Lumpur, Malaysia, December 11-14,2006.

VI. Flowchart for Start & End points Detection

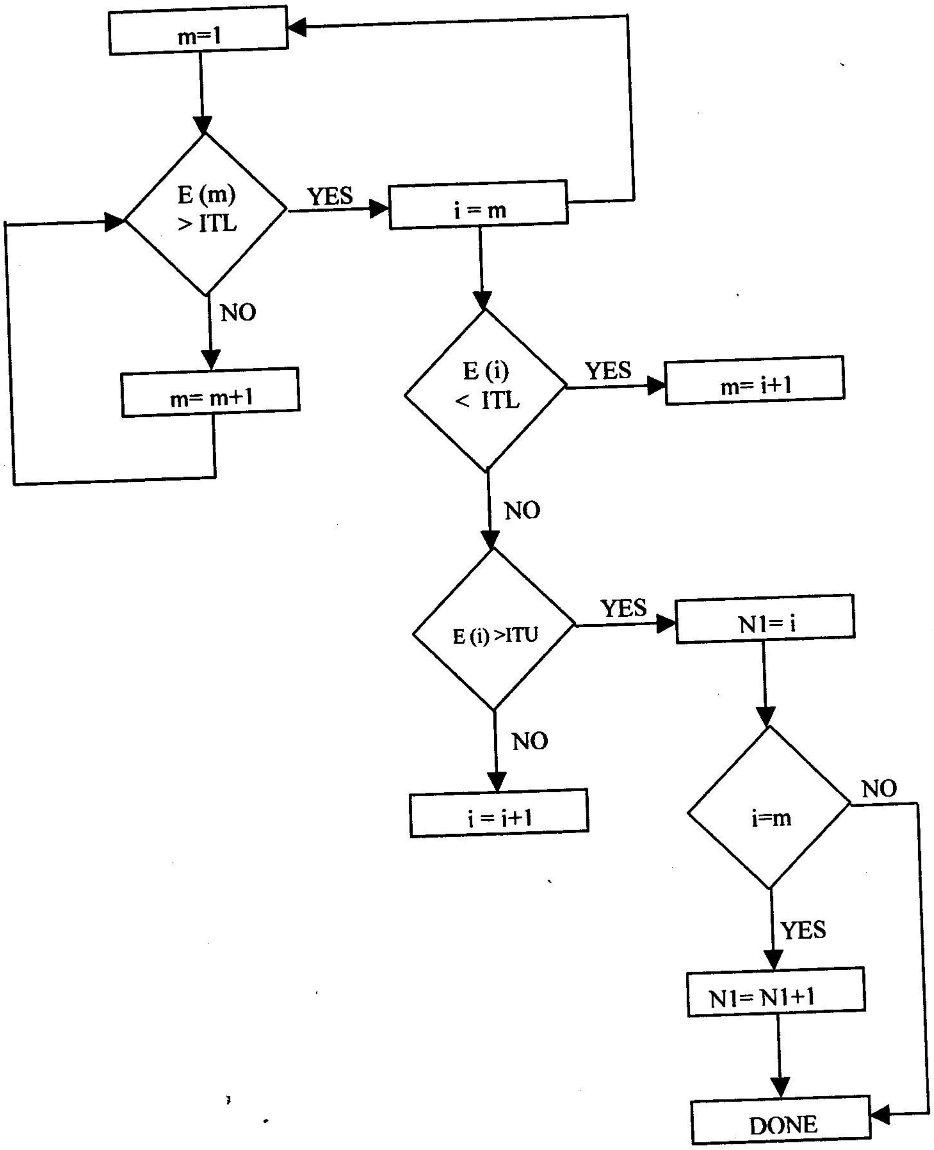

Fig 1. Flowchart for Start & End points detection

VII. References:

1. Samad S, Hussain A, Fah L K, " Pitch Detection of Speech Signals using the Cross Correlation Technique", Proceedings of IEEE on Speech, Audio and Signal Processing, pp 283 - 286,2000.
2. M.J.Ross, H.L.Shaffer, A. Cohen, R. Freudberg, and H.J. Manley, "Average Magnitude Difference Function Pitch Extractor", IEEE Trans. Acoustic, Speech, and Signal Processing, pp. 353-362 Oct. 1974.
3. D. Takin, "A Robust Algorithm for Pitch Tracking (RAPT)", Speech Coding and Synthesis, Netherlands: Elsevier Science. 1995.
4. LA. Atkinson, M. Kondoz and B.G. Evans, "Time Envelop Vocoder, A New LP Based Coding Strategy for Use of Bit-Rate 2.4kb/s and Below", IEEE Journal on Selected Areas on Communications, Vol. 13, No. 2, Feb. 1995.
5. Pramod B. Patil, Dr. V.T. Ingole, "Cross-Correlation Technique for the Detection and Reduction of Environmental Noise in Speech Signal", Journal of Pradushan Nirmulan, Vol 1, No2 pp 29 – 31, April 2004. (ISSN -0972-8902)
6. A. M. Kondoz, Digital Speech: Coding for Low Bit Rate Communications Systems, Wiley, England. 1995.
7. B.Gold and L. Rabiner, "Parallel Processing Techniques for Estimating Pitch Periods of Speech in the Time Domain", Journal of Acoustics. Society, America, Vol. 46, pp. 442-448, Aug. 1969.
8. Rabiner L R, Levinson S E., "Isolated and Connected Word Recognition – Theory and Selected Applications ", IEEE Transaction on Communications, Vol 29, No 5, 1981
9. Rabiner L R, Sambur M R, " Speaker Independent Recognition Of Connected Digits", Bell System Technical Journal Vol 54, pp 202 –205, 1972.
10. Rabiner L R, Schafer R W, Digital Processing of Speech Signals, Englewood Cliffs, NJ Prentice Hall, 1978.
11. Rabiner L R, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", Proceedings of IEEE, Vol 77, No2, pp 257-285, 1989.
12. Rabiner L R, Sambur M R, "An Algorithm for Determining the Endpoints of Isolated Utterances" Bell System Technical Journal, Vol 54, pp 297 –315, 1975.