An Approach to Reduce Noise in Speech Signals Using an Intelligent System: BELBIC

Edet Bijoy K edetbijoyk@ieee.org
Department of ECE
MES College of Engineering
Kuttippuram, 679573, Kerala, India

Musfir Mohammed mohammed.musfir@ieee.org
Department of ECE
MES College of Engineering
Kuttippuram, 679573, Kerala, India

Abstract

The two widespread concepts of noise reduction algorithms could be observed are spectral noise subtraction and adaptive filtering. They have the disadvantage that there is no parameter to distinguish between the speech and the noise components of same frequency. In this paper, an intelligent controller, BELBIC, based on mammalian limbic Emotional Learning algorithms is used for increasing the speech quality from a noisy environment. Here the learning ability to train the system to recognize and the output thus obtained would be the fundamental frequency of the speech spectrum thus reducing the noise level to minimum. The parameters on which the reduction of noise from the input speech spectrum depends have also been studied. The real time implementations have been done using Simulink and the results of the analysis thus obtained are included in the end.

Keywords: BELBIC, Spectral Noise, Adaptive Filtering, Fundamental Fequency, Simulink

1. INTRODUCTION

In recent years, two widespread concepts of noise reduction algorithms could be observed: spectral noise subtraction and adaptive filtering. The former has the drawback of generating residual noise with musical character, the so-called musical noise, while the latter distorts the frequency and phase response of speech signals [1,2]. In addition, both methods fail to enhance speech recordings disturbed by loud hum (50 Hz or 60 Hz) because their analysis windows enlarge the line spectrum by significant artificial side lobes. Here an attempt is made to cancel the noise from the surrounding environment to improve the quality of the speech and the avail of speech analysis in its fundamental frequency using a self learning Brain Emotion Learning system which works by mimicking the action of the mammalian limbic system.

Emotional Learning is a psychologically motivated algorithm which is a family of intelligent algorithms. Recently, biologically motivated intelligent computing has been successfully employed for solving different types of problems. The greatest different of an intelligent system from a traditional one is the capability of learning. A common attribute of the learning process is the adaptation of the system parameters to better tackle the changing environment. An evaluation mechanism is necessary that any learning algorithm assesses the operating condition of the system. One type of evaluation is based on emotional cues, which evaluate the impact of the external stimuli on the ability of the system both to function effectively in the short term and to maintain its long term prospects for survival [3]. Emotional learning is one of the learning strategies based on emotional evaluations. In mammalian brains, this learning process occurs in the brain Limbic System. The paper includes a short overview of the biological concepts followed by a detailed study of the architecture including the algorithm of the model. The results obtained from the real time implementation in Simulink are discussed towards the end.

2. BELBIC

Moren and Balkenius [4,5] presented a neurologically inspired computational model of the Amygdala and the Orbitofrontal Cortex in the Limbic System. Based on this model, a control algorithm called Brain Emotional Learning Based Intelligent Controller (BELBIC) has been suggested [6]. The two approaches of applying the brain emotional learning model into control systems are direct approach and indirect approach. The former uses BEL model as the controller block, while the latter utilizes BEL model to tune the controller parameters. In this work the BEL model itself is used as the controller. In real time control and decision systems, Emotional Learning is a powerful methodology due to its simplicity, low computational complexity and fast training where the gradient based methods and evolutionary algorithms are hard to be applied because of their high computational complexity [7].The BELBIC has been designed for SISO applications and for MIMO systems, one must employ each controller for generating one control output. The relevant researches have indicated that BELBIC has a good robustness and performance.

2.1 Architecture of the Limbic System

The Limbic System, as part of the mammalian creatures' brain, is mainly in charge of the emotional processes. The Limbic System located in the cerebral cortex consists mainly of following components: Amygdala, Orbitofrontal Cortex, Thalamus, Sensory Cortex, Hypothalamus, Hippocampus and others of which the Amygdala and Orbitofrontal Cortex are of importance in our model. Fig. 1 illustrates the anatomy of the main components of Limbic System [8].

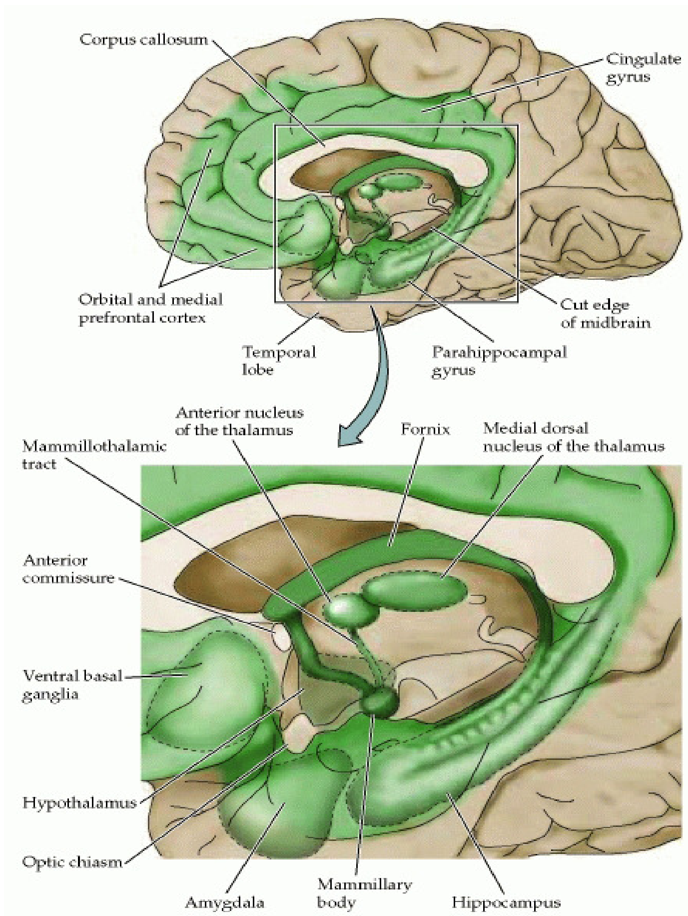

FIGURE 1: The major brain structures associated with the Limbic System

In this section these main components are discussed in detail. The first sign of affective conditioning of the system appears in Amygdala which is a small almond-shaped in sub-cortical area. This component is placed in a way to communicate with all other Sensory Cortices and areas within the Limbic System. The Amygdala connections to/from other components are illustrated in Fig. 2 [5]. The studies show that a stimulus and its emotional consequences are associated in the Amygdala area [9]. In this region, highly analyzed stimuli in the Sensory Cortices, as well as coarsely categorized stimuli in the Thalamus are associated with an emotional value.

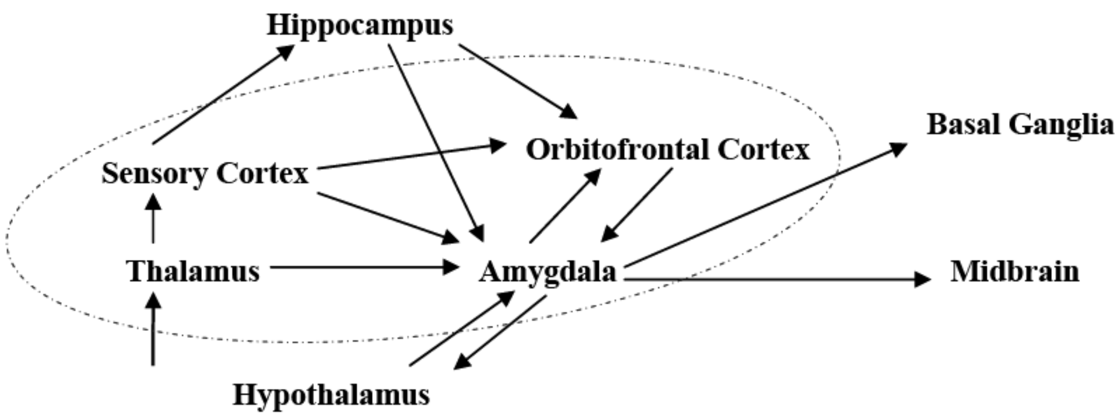

FIGURE 2: Connections of the Amygdala with other components of the Limbic System.

The Orbitofrontal Cortex, as of another component of the brain system, interacts with the Amygdala. The main interrelated function of this component is: Working Memory, Preparatory Set and Inhibitory Control [9]. The current and recent past events are represented in the Working Memory. The Preparatory Set is the priming of other structures in anticipation of impending action. Inhibitory Control is the selective suppression of areas that may be inappropriate in the current situation. More specifically, the Orbitofrontal Cortex takes action in omission of the expected reward or punishment and control the extinction of the learning in the Amygdala [9].

Another component in this area is Thalamus which lies next to the basal ganglia. It is a non-homogeneous sub-cortical structure and a way-station between cortical structures and sub-cortical. Moreover, various parts of the Thalamus also relay the majority of sensory information from the peripheral sensory systems to the Sensory Cortices. The Thalamus signal going to the Amygdala evades the processes involved in the Sensory Cortex and other components of the system. Therefore, Amygdala receives a non-optimal but fast stimulus from the Thalamus which among the input stimuli is often known as a characteristic signal.

The next component is the Sensory Cortex close to the Thalamus which receives its input from the latter one. In fact, Sensory Cortex processes the information from the sensory areas. The Sensory Cortex sends highly analyzed input to the Amygdala and Orbitofrontal [9]. Generally, the mammalians use these areas of their Limbic System for higher perceptual processing.

2.2 Computational Model of the BEL

Moren and Balkenius [4,5] developed a computational model that mimics Amygdala, Orbitofrontal Cortex, Thalamus, Sensory Input Cortex and generally those parts of the brain thought responsible for processing emotions. Fig. 3 shows the computational model of emotional learning [4]. The model is divided into two parts: the Amygdala and the Orbitofrontal cortex. The Amygdala part receives inputs from the Thalamus and from cortical areas, while the Orbitofrontal obtains inputs from the cortical areas and the Amygdala. The system also receives a reinforcing signal (Primary Reward).

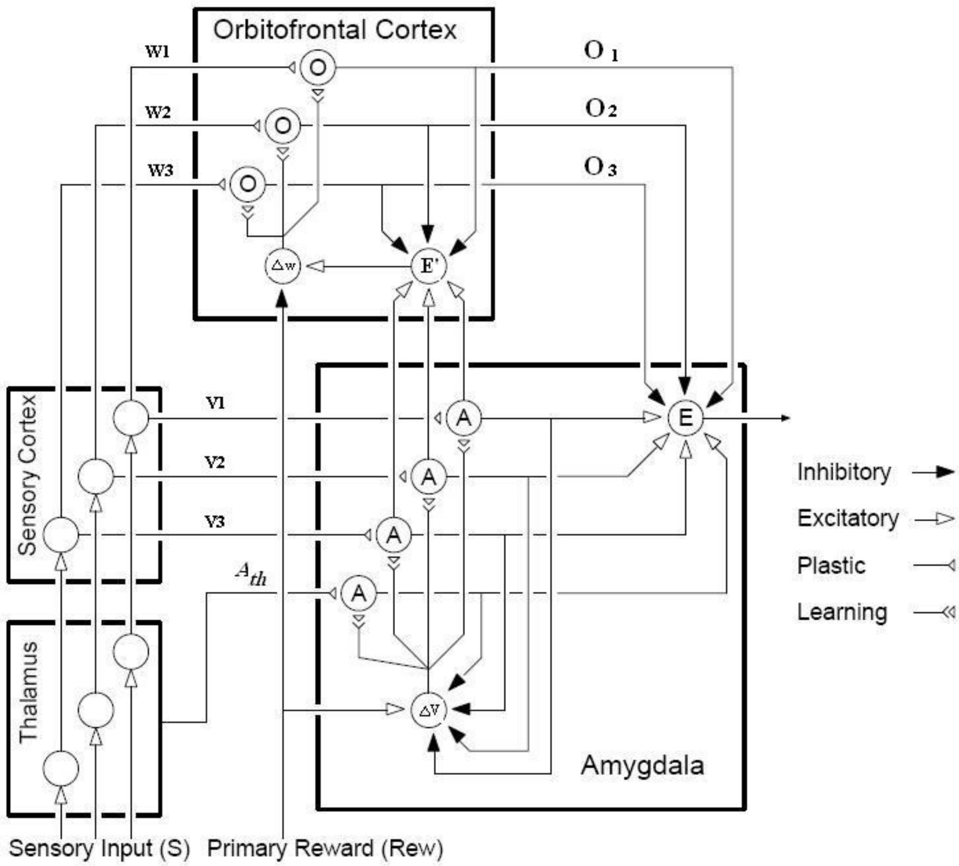

FIGURE 3: Graphical depiction of the Brain Emotional Learning (BEL) process

The vector *S* shows stimuli inputs to the system. There is one *A* node for each stimulus *S*. A_{th} is another input to the Amygdala part which is the maximum of stimuli inputs(S):

$$A_{th} = \max\ (S_i) \tag{1}$$

There is a plastic connection weight *V* for each *A* node. The output of each node obtains by multiplying any input with the weight *V*.

$$A_i = S_i V_i \tag{2}$$

The V_i is adjusted proportionally to the difference between the activation of the *A* nodes and the reinforcement signal *Rew*. The α term is a constant used to adjust the learning speed:

$$\Delta V_i = \alpha(S_i \max\left(0, Rew - \textstyle\sum_j A_j\right)) \tag{3}$$

The weights V cannot decrease. It is good reasons for this design choice because once an emotional reaction is learned, this should be permanent and cannot be unlearned. It is the task of the Orbitofrontal part to inhibit this reaction when it is inappropriate. The Orbitofrontal learning rule is very similar to the Amygdala rule but the Orbitofrontal connection weight can both increase and decrease. The O nodes behave analogously, with a connection weight W employed to the input signal to create an output.

$$O_i = S_i W_i \qquad (4)$$

β is another learning rate constant. ΔW_i is calculated as:

$$\Delta W_i = \beta(S_i(E' - Rew)) \qquad (5)$$

The E node sums the outputs from the A nodes and then subtracts the inhibitory outputs from the O nodes. The result is the output from the model. The A nodes give outputs proportionally to their contribution in predicting the reward Rew, while the O nodes inhibit the output of E as necessary. The E' node is sums of the outputs from A except A_{th} and then subtract from inhibitory outputs from the O nodes.

$$E = \textstyle\sum_i A_i - \sum_i O_i \ (including\ A_{th}) \qquad (6)$$

$$E' = \textstyle\sum_i A_i - \sum_i O_i \ (not\ including\ A_{th}) \qquad (7)$$

Based on the above equations the BELBIC model was designed as a cognitively open loop model by C. Lucas et al [6]. Fig 4 is the implementation of the foresaid model [6]. The BELBIC is essentially an action generation mechanism based on sensory inputs and emotional cues (Reward signals). The BELBIC equations are the mentioned formulas of (1) - (7). The main issue in using the model for different applications is defining the sensory and emotional signals in such a way that properly represent the state and objectives of the system.

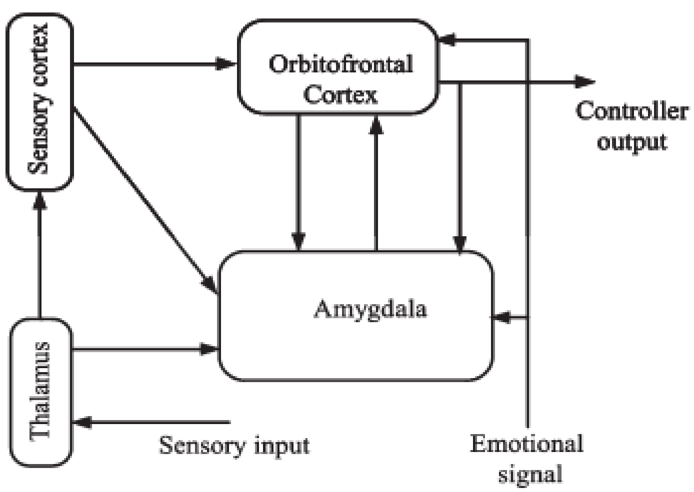

FIGURE 4: Basic block structure of emotional controller

Fig. 5 demonstrates a reasonable candidate for embedding the BELBIC model within a typical feedback control block diagram [6]. The implemented functions in emotional cue and sensory input blocks should be defined for each application.

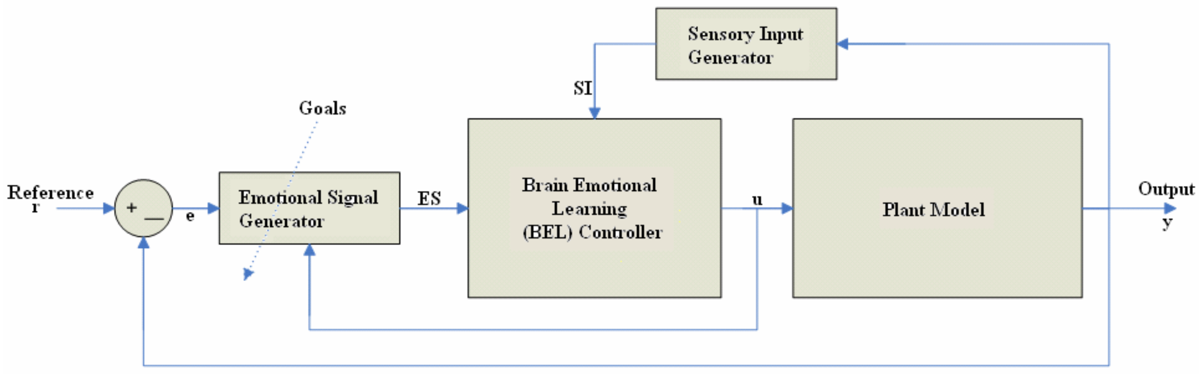

FIGURE 5: Control system configuration using BELBIC

3. APPLICATION OF BELBIC FOR NOISE REDUCTION

The above model is implemented in the obtaining of the speech signals from a noisy environment. The real time implementation is done using Matlab2006b. Fig 6 shows the block structure implemented in Simulink.

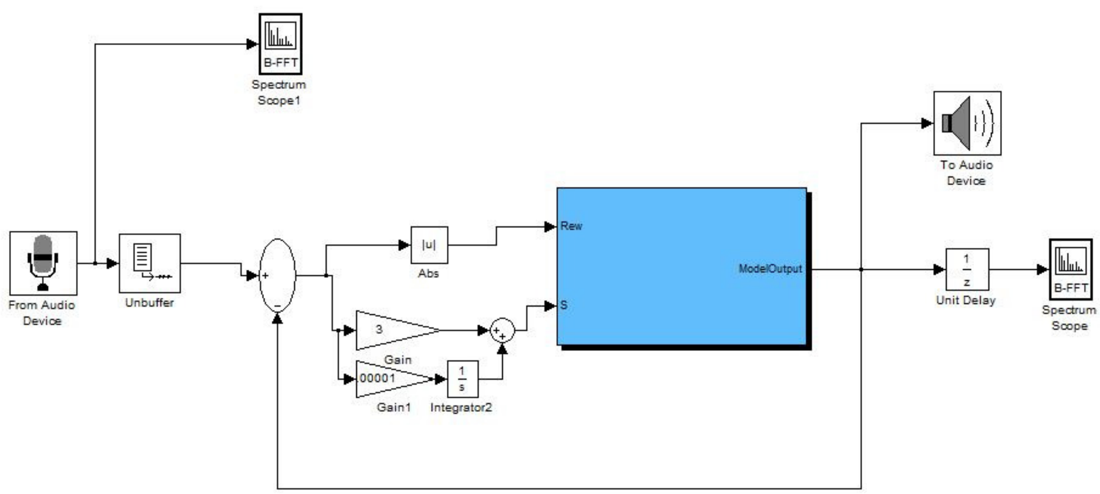

FIGURE 6: Implementation of noise elimination system in Simulink

Male voice sampled at a frequency of 1000Hz is given as the speech input through the microphone from a room environment. Noisy environment is ensured with other sound sources at varying frequencies. The Reward signal consisted of the absolute value of the speech input. The sensory input is taken as the sum of the speech input and it's integral. The negative feedback is provided for the stability of the system. The system tries to attain more of the reward from the input by giving the output of the most prominent fundamental frequencies of the speech signal and suppressing the other. Thus the hum produced due to the harmonics is very negligible. The gain control is achieved due to the negative feedback present in the system and the integral component to the sensory input.

3.1 Results

The simulation results are as shown for varied gain parameters of the speech and the integral input respectively.

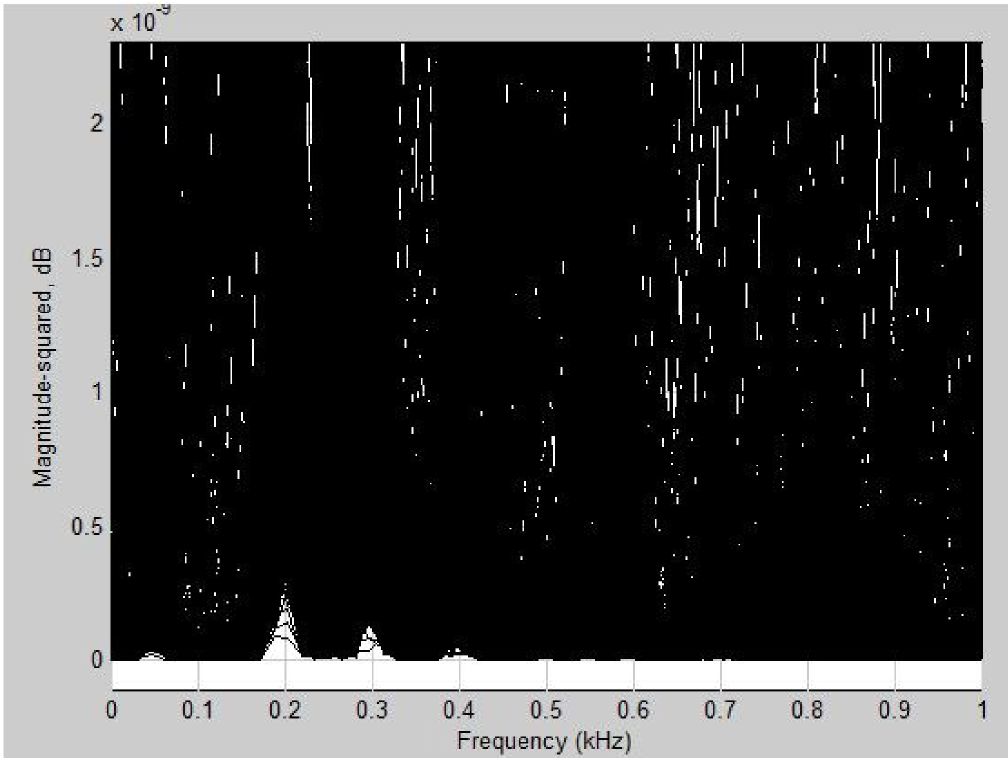

FIGURE 7: Spectrum of the speech input to the system

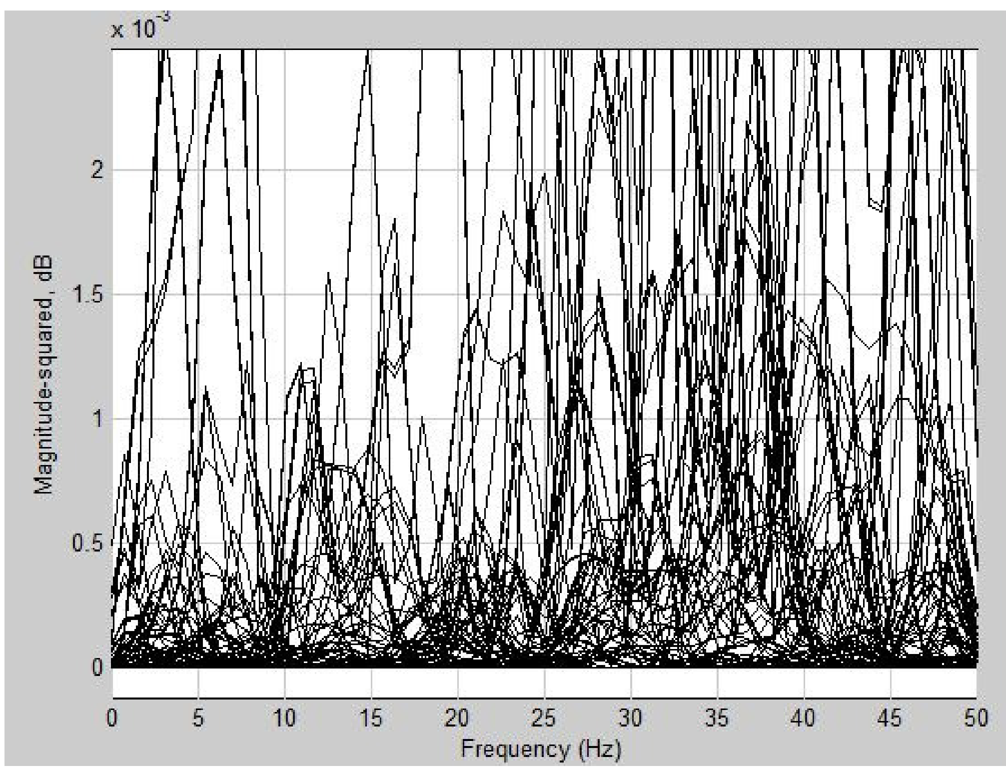

FIGURE 8: Output for gains 1 and 0

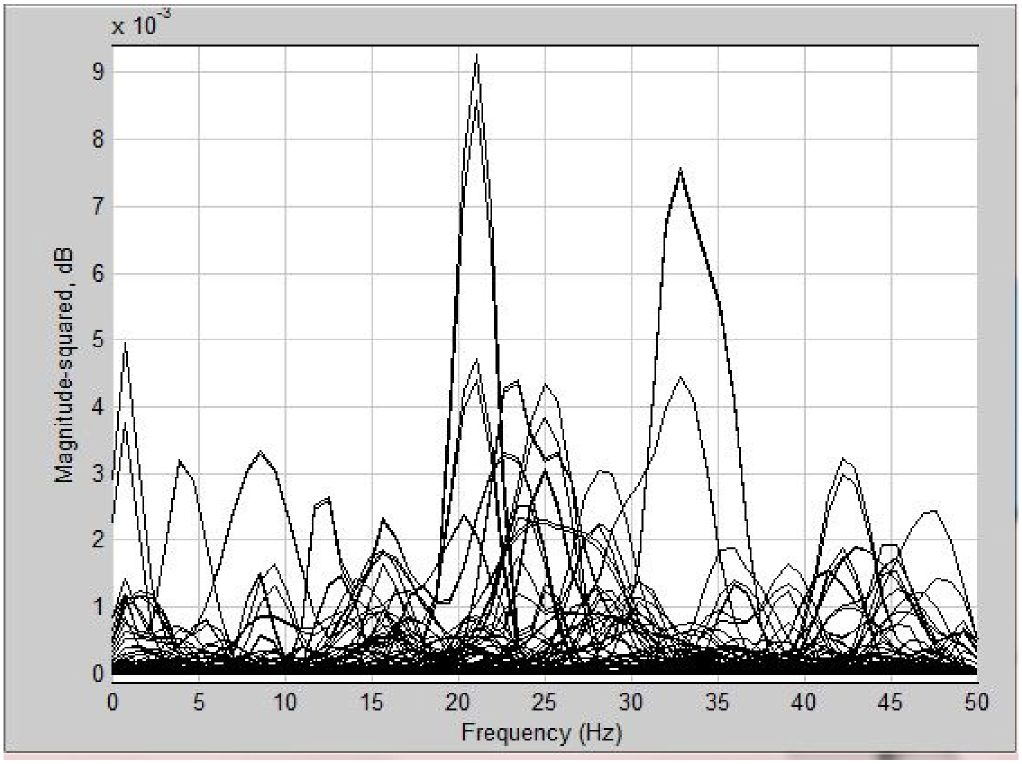

FIGURE 9: Output for gains 1 and 0.00001

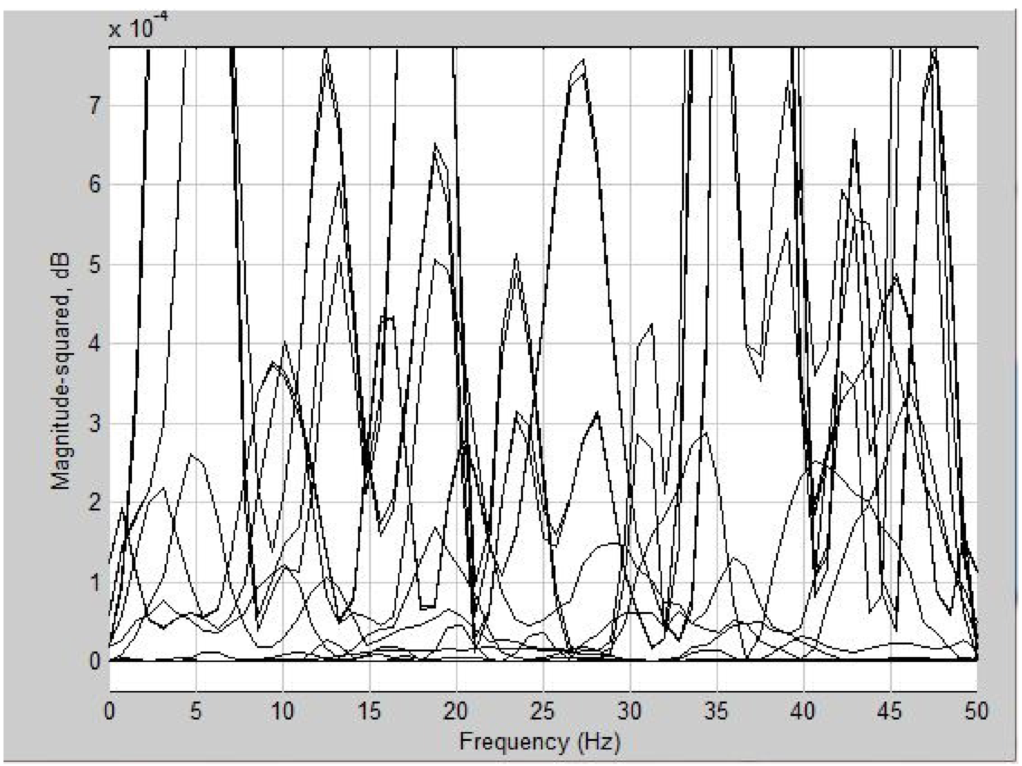

FIGURE 10: Output for gains 1 and 10

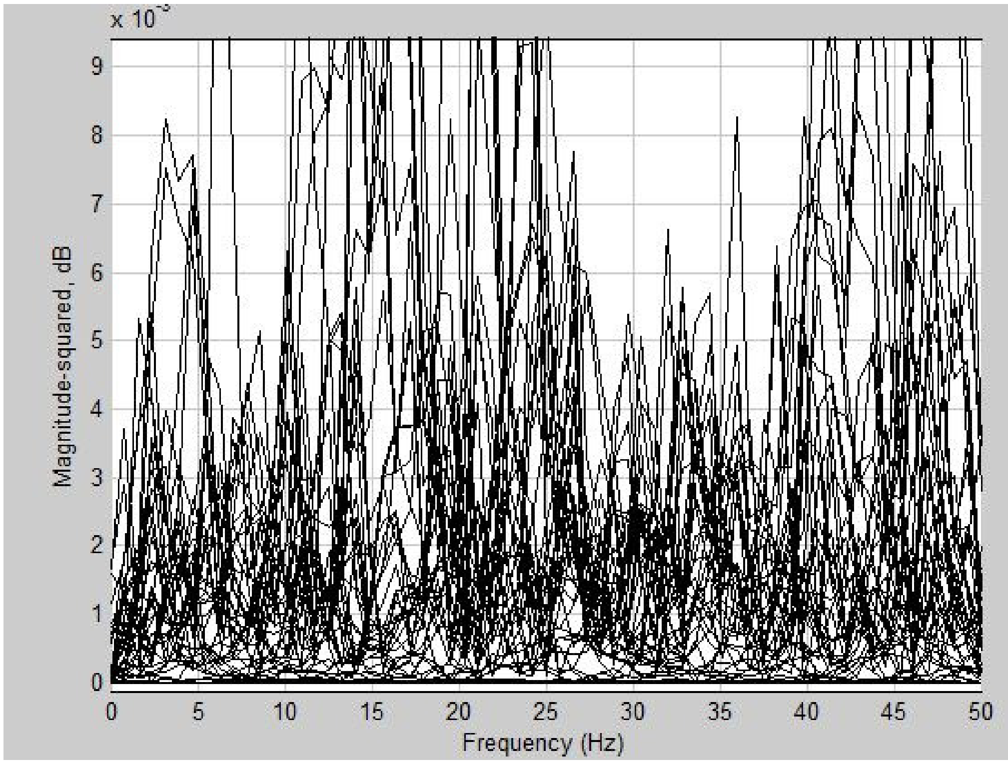

FIGURE 11: Output for gains 100 and 2

Fig. 7 shows the spectrum of the input speech in which the noise and the harmonics are very prominent. Fig. 8 is the output without the integral component to the sensory input. The output quality is further increased when the integral component gain is .00001 as in fig 9. Fig. 10 shows the output for a gain of 10 for the integral component. Fig 11 is when the speech component has a gain of 100 and the integral component 2. From the above results it can be seen that the quality of noise elimination increases with the increase in integral gain. This cannot be increased above a particular limit as the learning ability of the system reduces i.e. the learning period increases thus causing a delay in the transmission of the output. The easy learning property of the system is to be exploited in this system. The increase in the gain of the speech component increases the learning capability over an undesirable rate and thus causes the repetition of the speech thus affecting the quality.

The inputs given have a noisy component which is the original speech signal and the integrated signal which contains the lower frequency component. The gain of this lower frequency signal is the determining parameter as it is the signal with lesser noise and more of speech component. Further the integrated signal also functions in tuning the learning ability of this system at a desirable speed. A higher gain for the low frequency signal ensures a higher quality in the speech output but is limited by the learning capability of the system.

Noise reduction using this system requires only preliminary knowledge of the elements of speech spectrum. There have been approaches made using analogous controllers where a Voice Activity Detection (VAD) is necessary to prevent noise amplification. VAD being computationally intensive requires a very complex algorithm [10]. Using a system of high learning ability such as BELBIC saves this complexity. Furthermore as the learning method is emotional learning the algorithm is very simple. The efficiency of the BELBIC controller over the analogous counterparts such as the PID controller is very high due to the high intelligence. Thus the responses of this controller are faster and with less overshoots compared with those of the PID counterparts [3].

4. CONCLUSION

BELBIC is a highly versatile intelligent system. It has a very high degree of accuracy and disturbance handling ability due to appropriate learning ability compared to analogous classical controllers. In this work as the fundamental frequency of speech is prominent in the output this can be implemented in the voice recognition systems and other high end speech applications. The algorithm being based on emotion learning is very versatile and easy to implement. The work is to be extended for implementation in DSK TMS320C6713. Applications of this type of systems can be in hearing aids, speech synthesis systems communication systems in noisy environment etc.

5. REFERENCES

[1]. Yoma, N. B.; McInnes, F. R.; Jack, M. A. (1997). “Spectral subtraction and mean normalization in the context of weighted matching algorithms.” *In Proc. of EUROSPEECH ’97, vol. 3, pp. 1411–1414, Rhodes; Greece.*

[2]. Pfitzinger, H. R. (1998). “The collection of spoken language resources in car environments.” *In ICLRE ’98, vol. 2, pp.1097–1100, Granada; Spain.*

[3]. D. Shahmirzadi, “Computational Modeling Of The Brain Limbic System And Its Application In Control Engineering”, *Master dissertation, Texas A&M University, U.S.A. , (2005).*

[4]. J. Moren, C. Balkenius, "A Computational Model of Emotional Learning in the Amygdala", *Cybernetics and Systems, Vol. 32, No. 6, (2000), pp. 611- 636.*

[5]. J. Moren, “Emotion and Learning – A Computational Model of the Amygdala”, *PhD dissertation, Lund University, Lund, Sweden, (2002).*

[6]. C. Lucas, D. Shahmirzadi, N. Sheikholeslami, "Introducing BELBIC: Brain Emotional Learning Based Intelligent Controller", *International Journal of Intelligent Automation and Soft Computing, Vol. 10, NO.1, (2004), pp. 11- 22.*

[7]. R. Ventura, and C. Pinto Ferreira, "Emotion based control systems", *Proc. Of IEEE Int. symp. On Intelligent control/Intelligent systems and semiotics, Cambridge, MA, (1999), pp. 64-66.*

[8]. D. Purves, G. J. Augustine, D. Fitzpatrick, L. C. Katz, A. LaMantia, J. O. McNamara, S. M. Williams, “Neuroscience” , *Sinauer Associates, (2001).*

[9]. E. T. Rolls, “A Theory of Emotion and Consciousness, and Its Application to Understanding the Neural Basis of Emotion.” *Cambridge, MA: MIT Press. (1995).*

[10]. G.S. and Lidd, M.L., “Automatic gain control. Acoustics, Speech, and Signal Processing Kang”, IEEE International Conference on ICASSP. Volume 9,1984, pp. 120 – 123.