Karaoke Machine Implementation and Validation using Out of Phase Stereo Method along with LMS

[1]Sneha Jadiye, [2] Aarti Bakshi, [3]Manjiri Gogate
[1] Student, [2]Assistance Professor, [3]Assistance Professor
[1]Electronics and Telecommunications,
[1],Shri L. R. Tiwari College of Engineering, Thane, India

_____

Abstract **- This paper presents a new approach for the implementation of Karaoke Machine using LMS filter by outcome the filter coefficients that narrate to creating the smallest mean squares of the fault signal. It is a stochastic slope descent method in that the filter is only modified based on the error at the current time. The basic idea behind LMS filter is to method the optimum filter weights by updating the filter weights in a manner to unite to the best filter weight. The process starts by presumptuous a minor weights (zero in most cases), and at each step, by finding the slope of the mean square error, the bulks are informed. The basic use of LMS filter is to remove undesirable noise from the Music signal by using OOPS model.**

Index Terms **- Karaoke Machine, MATLAB, Out of Phase Stereo (OOPS), Simulink, LMS**
_____

I. INTRODUCTION

Entertainment is a form of movement that clasps the attention and interest of an listeners, or gives delight and pleasure. It can be an impression or a duty, but is more likely to be one of the happenings or proceedings that have established over thousands of years specifically for the purpose of custody of an audience's consideration, most forms are recognizable and familiar are Music, Drama, Dance and different kinds of concert exist in all cultures. From Japanese kara, "vacant" or "cancelled"(and ōkesutora, "orchestra")[1] is a form of showbiz in which users sing along with recorded music using a microphone and a PA system. The music is typically a well-known song in which the voice of the unique singer is detached or condensed in volume. Karaoke machines find many applications in studios, theatres, in DVD and VCD players, music or voice recording as well as in games and reality shows and singing contests. In karaoke, the quality of the accompanying sound track directly affects the aesthetic experience of the "user singer". Reproduction of an accompaniment requires lots of time and labour, yet the music is probably not close enough to the original mix. Thus, removing the vocal or lead instrument track from the original recording remains an attractive topic.

So the document covers the methods like applying Equalizers to the original recording attenuating the vocal frequency bands of the spectrum for voice suppression. To support this theory, multiple filters of varying pass band frequencies are also present in the document.

II. CRITICALITY & COMPEXITY OF KARAOKE MACHINE

1. Implementing filtering method for karaoke machine was tried in the following manner:
   - Initial approach of suppressing vocal was Digital filter Digital filter [2] (FIR or IIR): Probably a band stop digital filter with stop band of approx.100Hz to around 3 KHz. Digital filter is shown in figure no 1.
   - A high pass digital filter with cut off frequency of about 3.3 KHz.
   - Initially Building different Simulink models in MATLAB using FDA tool of various speech frequencies was tried.

Fig. 1. Filter design using FDA tool

2. Following results were achieved through the filtering method:
   - Filtering changed/distorted the pitch of the original voice.
   - However, the voice was not removed or not even suppressed. On the other hand bass component was suppressed and distorted.

3. Problem with Karaoke was overcome by implementation of following.
   − Subtraction Method (OOPS method)[3].
   − To split a recorded stereo sound track into two channels.
   − Thus generating two mono sound tracks: left and right
   − Followed by that, one channel of original recording is subtracted from another.

III. KARAOKE WITH OOPS CONCEPT – DESIGN PROCESS

When a song is recorded in a music studio in stereo form, there are usually two channels: left and right. The artist who is standing at the center of the studio stage; his voice is simultaneously and equally fed into both these channels hence the vocal component is almost identical in both the left and right tracks of the audio song; whereas when the music of multiple instruments is recorded, the stereo sound output would have different musical components in left and right channels. Therefore the approach for reducing voice component involves splitting a recorded stereo sound track into two channels thus generating two mono sound tracks (left and right) and then by subtracting the left and right channel of the original recording.

Voice of the unique artist is efficiently reduced and if recorded music is honestly composed, voice is repressed totally. Karaoke Machine can be implemented using OOPS method. OOPS stands for Out Of Phase Stereo. This technique is used to process the two channels of current stereo recordings into a "new", third channel[3]. Block diagram of OOPS method is shown in below figure no 2.
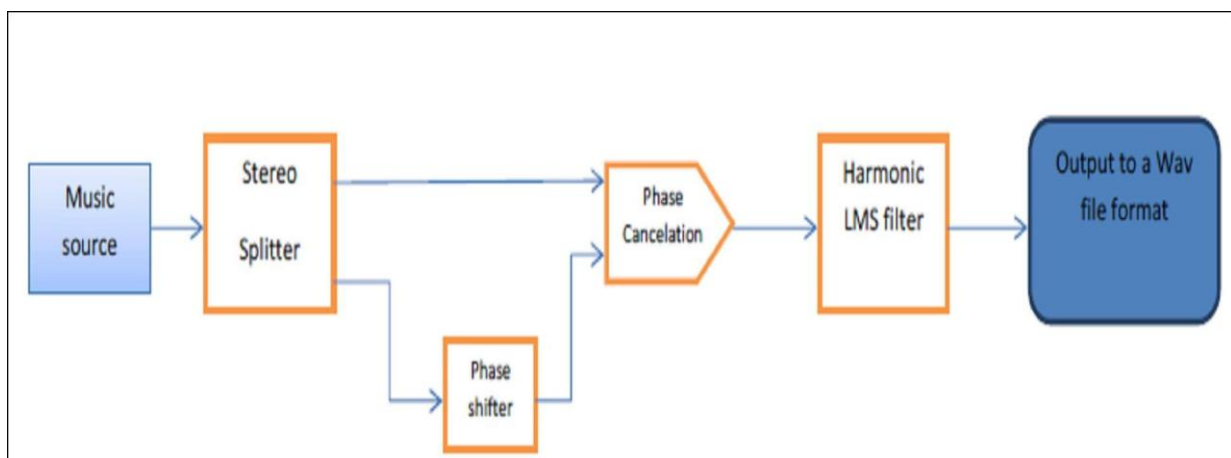

Fig 2. Model to analyze Karaoke Machine using OOPS method

The speakers (or headphones) of stereo system change an electrical signal into sound, by affecting the speaker cone in relative to the positive and negative waves in the signal. Following figure no 3 shows a single tone.
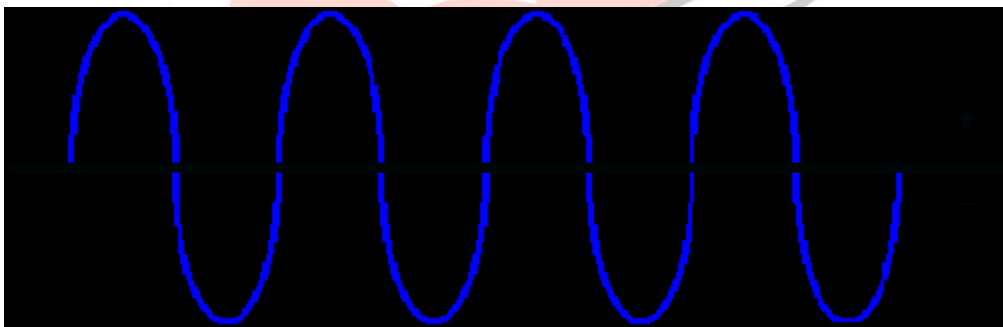

Fig 3.Diagram of Single tone

In the humble example of a clean tone, the electrical signal kinds the speaker "vacillate" (move back and forward speedily) by disturbing one way, outwards for example, on the positive shares of the electrical waves, and move back the other way on the negative shares. Recorded sound is through up of many of these electrical waves in various mixtures relating to the frequencies in the sounds presence recorded.

Imagine two signals shown in figure no 4 below, that are alike, except they are "out of phase" with each other, that is, that when one signal is having a positive wave peak, the other is having an identical negative peak.
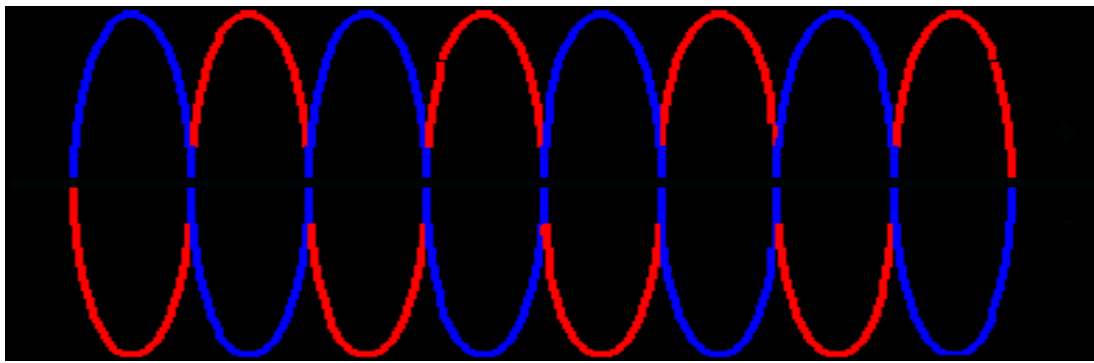
Fig 4. Diagram for Double tone opposite phase tone

If these signals are assorted together, they will cancel out each other. The first signal tries to make ready for speaker move of positive top, but as it does, it is countered by the equal but opposite negative top from the second signal, which efforts to pull the speaker cone back in, hence no any sound gives results.

Because the mixing done throughout the recording course, some of the sounds are on both the channels, which is very common in modern stereo recordings. Now, if one channel from above example is taken into the consideration, and then by combining it with OUT OF PHASE (plus to minus and minus to plus) with the another channel, whatever was same in both channels would be cancel out, and after that would hear only what was not same in both channels[4]. In this example, One would cancel the vocal, and hear only the instrumental complement.

OOPS is a technique used to process the two channels of recent stereo recordings into a "fresh", third channel. It gives us "unseen" sounds in stereo recordings. The speakers (or headphones) of your stereo system translate an electrical signal into sound, by moving the speaker cone in relation to the positive and negative waves in the signal.

In the simple example of a clean tone, the electrical signal makes the speaker "fluctuate" (move back and forth rapidly) by moving one way, outwards for example, on the positive (+) halves of the electrical waves, and move back the other way on the negative halves. The speaker cone movement styles the air in front of it move, and this normally oscillating air choices your ears as sound. This single wave is a "frequency" and you receive it as a tone. Verified sound is made up of many of these electrical waves in several combinations telling to the frequencies in the sounds actually recorded.

OOPS method is implemented using Simulink as shown in figure no 5. In simulink model audio file is read and then split into two parts. One part is left channel and another is right channel. Then we multiply left channel by 1 and Right channel by -1[5]. So left channel will remain as it and right channel will we inverted. After doing this we add both channels. So same component well be removed and different component will remain in the output. Now this output can be added with user's voice to create new song in user's voice.
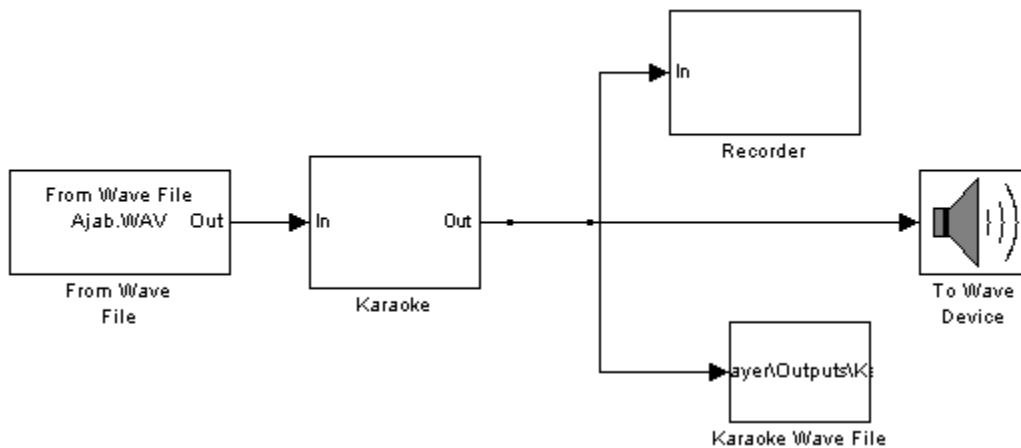

Fig 5. Simulink Model for Implementation of Karaoke Machine

IV. LEAST MEAN SQUARES FILTER

Least mean squares (LMS) algorithms are a type of adaptive filter used to impressionist a desired filter by finding the filter factors that narrate to constructing the least mean squares of the fault signal[6]. It is a stochastic gradient descent way in that the filter is only altered based on the error at the present time. The block diagram of LMS filter is as shown in figure no 6 below.
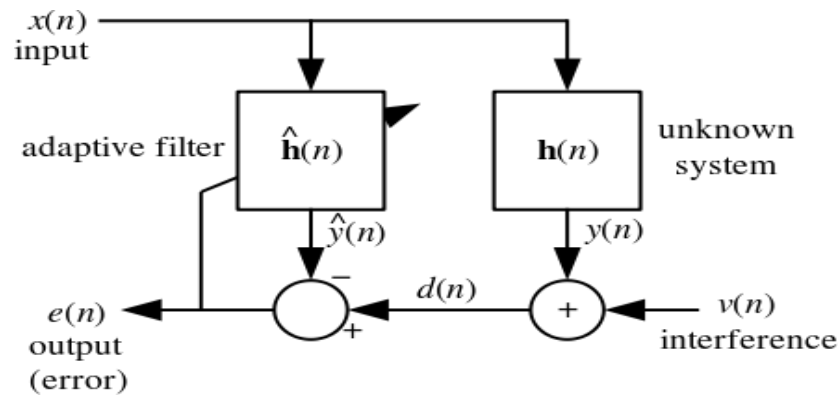
Fig 6. LMS Filter Block Diagram

The elementary idea in arrears of LMS filter is to line of attack the optimal filter weights $R^{-1}P$ [7], by informing the filter weights in a mode to unite to the optimum filter weight. The procedure starts by considering a slight weights (zero in most cases), and at every step, by governing the slope of the mean square error, the weights are updated. That is, if the gradient is positive, it suggests, the error would reserve increasing positively, if the same weight is used for extra repetitions, which means we want to decrease the weights. Similarly, if the gradient is negative, we have to increase the weights. So, the basic weight update equation is

$$W_{n+1} = W_n - \mu \nabla \varepsilon[n]$$

Where ε denotes the mean-square blunder. The negative sign shows that, we have to vary the weights in opposite direction to that of the gradient slope.

The mean-square error, as a role of filter weights is a quadratic role which means it has only one dangerous that minimizes the mean-square error, which is the ideal weight. The LMS thus, slants towards these ideal weights by arising/sinking down the mean-square-error vs. filter weight curve.

The LMS procedure for a pth order algorithm can be shortened as

Parameters: $p =$ filter order
$\mu =$ step size

Initialization: $\hat{\mathbf{h}}(0) = \text{zeros}(p)$

Computation: For $n = 0, 1, 2, \dots$

$$\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-p+1)]^T$$
$$e(n) = d(n) - \hat{\mathbf{h}}^H(n)\mathbf{x}(n)$$
$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \mu\, e^*(n)\mathbf{x}(n)$$

The key drawback of the "clean" LMS algorithm is that it is delicate to the scaling of its input $x(n)$. This marks it very hard (if not impossible) to pick a learning rate μ that assurances firmness of the algorithm[7]. The Normalised least mean squares filter (NLMS) is a different of the LMS algorithm that solve this problem by normalizing with the power of the input[8]. The NLMS algorithm can be précised as:

Parameters: $p =$ filter order
$\mu =$ step size

Initialization: $\hat{\mathbf{h}}(0) = \text{zeros}(p)$

Computation: For $n = 0, 1, 2, \dots$

$$\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-p+1)]^T$$
$$e(n) = d(n) - \hat{\mathbf{h}}^H(n)\mathbf{x}(n)$$
$$\hat{\mathbf{h}}(n+1) = \hat{\mathbf{h}}(n) + \frac{\mu\, e^*(n)\mathbf{x}(n)}{\mathbf{x}^H(n)\mathbf{x}(n)}$$

V. SOFTWARE CAN BE USED

1. **MATLAB:** MATLAB is a high-level language and collaborating location for numerical calculation, conception, and software design. Using MATLAB, we can examine data, improve algorithms, and create simulations and applications. The

language, tools, and built-in math jobs permit us to discover various approaches and spread a solution quicker than with spreadsheets or outdated software design languages, such as C/C++ or Java. We can practice MATLAB for a series of applications, counting signal processing and communications [8], image and video processing, control systems, check and dimension, computational finance, and computational biology. More than a billion engineers and experts in business and university use MATLAB, the language of practical computing.

2. **XILINX:** Xilinx is the world's foremost benefactor of all Programmable FPGAs, SoCs and 3D ICs. These industry-leading strategies are joined with a next-generation strategy environment and IP to serve a wide range of buyer needs[9], from programmable sense to programmable systems mixing. With over 3,500 rights and 60 industry firsts, Xilinx is known for its significant successes counting the overview of the first FPGA and the start of the fabless model[10][11].

VI. SIMULATION RESULTS

After simulation of an original song file in simulink, we get Following results:

The following figure no 6 represent the output taken from the left channel which is present into the Karaoke machine. In this the two graphs represents, one is the original wave in which noise as well as the music is present and in another graph only the balanced wave is presence in the left channel.

The figure no 7 represent the output taken from the right channel. In this the graphs represents the original wave as well as the balanced wave of the right channel.
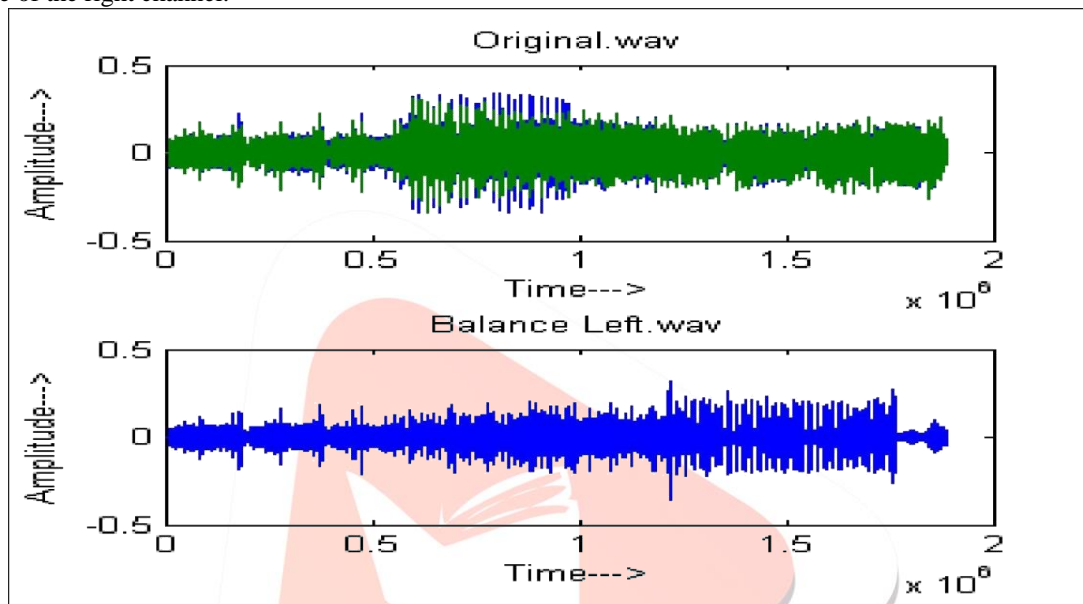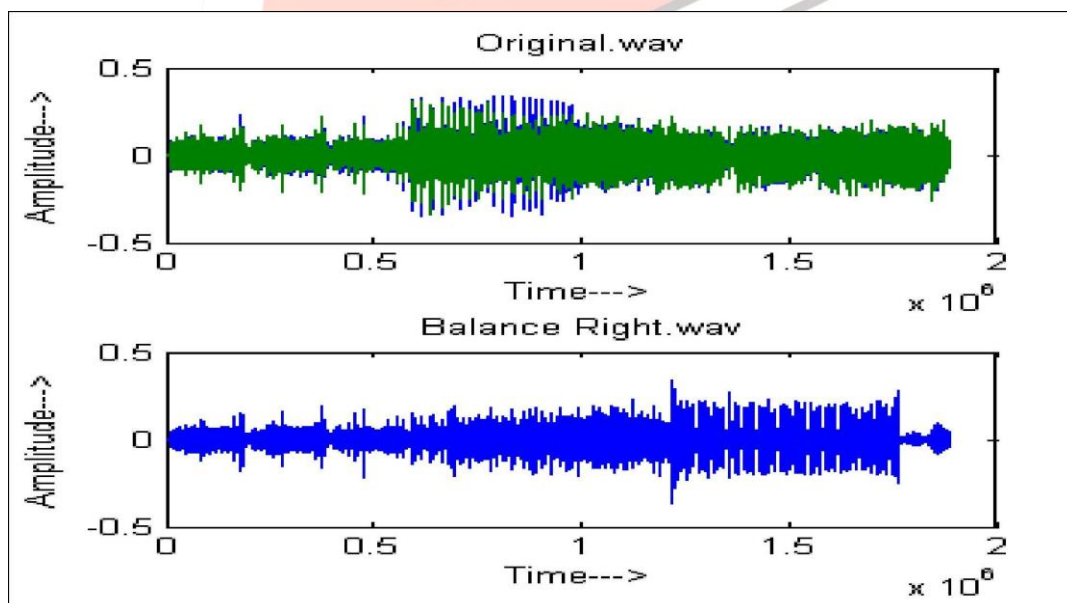

Fig 6. Original song and Balance left channel

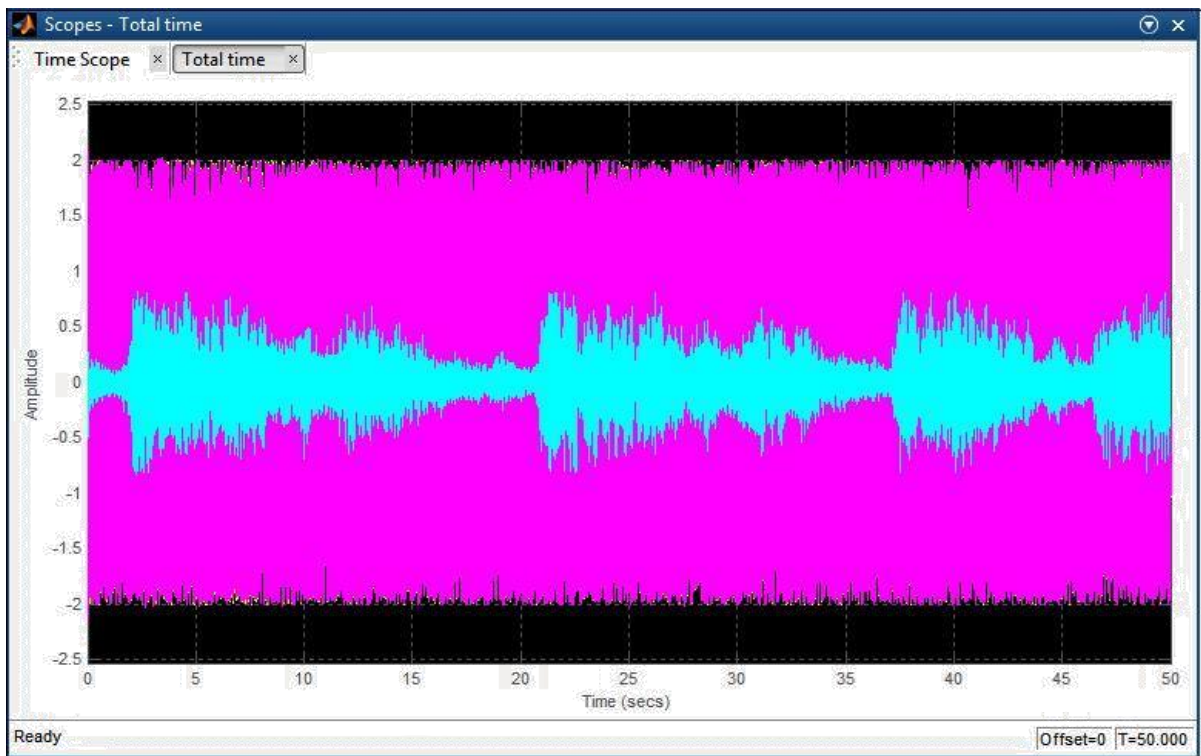Fig 7. Original song and Balance Right channel

Fig 8. LMS Filter Output OR Karaoke song

The main conclusion of this paper is by implementing a OOPS algorithm helps us to achieve voice cancellation form music signals and it can be applied in real time. The main purpose of karaoke machine is to obtain clear voice of user along with unsuppressed music signal, initially we used different filter but we obtained noisy and somewhat suppressed music components like bass because of that the quality of song is decreases[12]. So through this paper we used OOPS model with LMS filter to remove harmonic distortion and reduce original artist's voice from original song. For that we will use MATLAB language in order to calculate different values and indications of signal graph of voice signal. For future work we will implement Microprocessor chip with a fast special and containing special instruction set which include functions of LMS filter for calculation.

VII. ACKNOWLEDGMENT

REFERENCES

[1] Nirmal R Bhalani, Prof.Jaikaran Singh, Prof.Mukesh Tiwari,"*Karaoke Machine Implementation and validation Using Out of Phase Stereo Method" in* 2012 International Conference on Communication, Information & Computing Technology (ICCICT), Oct. 19-20, Mumbai, India.

[2] "*Recursive and Non-reeursive Realizations of Digital Filters Designed by Frequency SamplingTechniques"* in IEEE TRANSACTIONS ON AUDIO AND ELECTROACOUSTICS VOL. Au-19, NO. 3 SEPTEMBER 1971 LAWRENCE R. RABINER, Member, IEEE RONALD W. SCHAFER, Member, IEEE Bell Telephone Laboratories, Inc. Murray Hill, N. J. 07974

[3] Out of phase stereo (OOPS) available at http://en.wikipedia.org/wiki/Out_Of_Phase_Stereo_(OOPS)

[4] Udo Zoelzer (Ed). DAFX:Digital Audio Effects. J. Wiley & Sons, Chichester, 2002

[5] D. Jang, T. Lee, Y. Lee, and J. Yoo, "*A personalized preset-based audio system for interactive service,*" in *Proc. 121st AES Conv.*, San Francisco, CA, 2006, Preprint 6904.

[6] Consideration of Interactive Music Service Apr. 2008, ISO/IEC JTC1/ SC29/WG11 (MPEG), Archamps, Document M15390.

[7] J. Herre and S. Disch, "New concepts in parametric coding of spatial audio: From SAC to SAOC," in *Proc. Int. Conf. Multimedia Expo*, Jul. 2007, pp. 1894–1897.

[8] J. Engdegard, B. Resch, C. Falch, O. Hellmuth, J. Hilpert, A. Hoelzer, L. Terentiev, J. Breebaart, J. Koppens, E. Schuijers, and W. Oomen,, "*Spatial audio object coding (SAOC) The upcomingMPEG standard on parametric object based audio coding,*" in *Proc. 124th AES Conv.*, Amsterdam, The Netherlands, 2008, Preprint 7377.

[9] O. Hellmuth, H. Purnhagen, J. Koppens, J. Herre, J. Engdegard, J. Hilpert, L. Villemoes, L. Terentiv, C. Falch, A. Holzer,M. L. Valero, B. Resch,H. Mundt, andH.Oh, "*MPEG spatial audio object coding—The ISO/MPEG standard for efficient coding of interactive audio scenes*," in *Proc. 129th AES Conv.*, 2010, Preprint 8264.

[10] J. Park, K. Kim, J. Seo, and M. Hahn, "*Modified spatial audio object coding scheme with harmonic extraction elimination structure for interactive audio service*," in *Proc. Interspeech '10*, Sep. 26–30, 2010, pp. 2906–2909.

[11] Jan Vanus,Viteslav Styskala, "*Applications of optimum settings of the LMS adaptive filter for speech signal processing*" in Proceedings of the International Multiconference on Computer Science and Information Technology pp. 767–774.

[12] Wei-Ho Tsaiand Hsin-Chieh Lee, "*Automatic Evaluation of Karaoke Singing Based on Pitch, Volume, and Rhythm Features*" in IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 20, NO. 4, MAY 2012.