Software Implementation of Automatic Gain Controller for Speech Signal

Fitzgerald J. Archibald

An Automatic Gain Controller (AGC) for speech signals embedded in additive noise requires Voice Activity Detection (VAD) to avoid noise amplification, a peak level detector for computing gain, and a gain controller for adjusting gain. This paper describes a low computational-intensive software AGC for use in handheld devices. The AGC provides options for static and dynamic noise floor estimation in a VAD module. Further, this paper describes analog and digital gain adjustment with gain curve selection to allow for distance perception during the AGC operation.

Index Terms— AGC, ALC, Automatic Volume Controller, Digital Gain Controller, Digital amplification, Gain curve, VAD.

Contents

1 Introduction 1
2 Algorithm Design 2
3 Results 13
4 Summary 14
5 References 15

List of Figures

1 AGC Block Diagram 2
2 AGC Flowchart 3
3 Dynamic VAD Flowchart 5
4 Binary Search for Gain Selection 6
5 Input Versus Output (Non-Linear) 7
6 Input Versus Output (Pseudo Linear) 8
7 Input Versus Output (Linear) 9
8 Input Versus Gain (Pseudo Linear) 10
9 Input Versus Gain (Constant) 11
10 Input Versus Gain (Linear) 12
11 AGC Internal and External Signal Plot 14

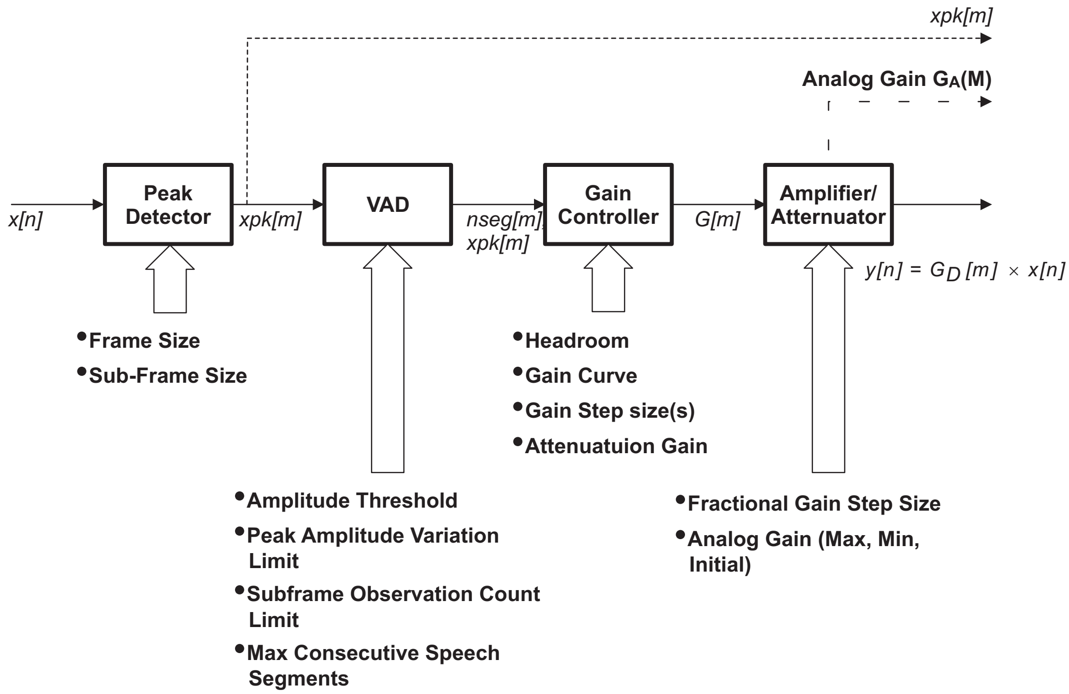
1 Introduction

In Digital Still Camera (DSC), sound is recorded along with captured video frames for movie capture application. The sound signal is converted to an electrical signal by a microphone and converted to a digital signal by an Analog to Digital Converter (ADC). The need for an Automatic Gain Controller (AGC) is to amplify speech segments to an intelligible sound level, while not amplifying noise only segments [4], [5], [6]. This paper describes an AGC using a low computation intensive Voice Activity Detector (VAD) for detecting speech embedded in stationary noise. In addition, the paper explains the use of gain curve to preserve relative sound level of different speech segments to retain distance perception contained in the audio signal. The paper provides three different means of gain correction; namely, analog, digital and a combination of analog and digital gain.

2 Algorithm Design

The AGC for the speech signal consists of a peak detector, VAD, gain controller, and amplifier/attenuator. Figure 1 illustrates the block diagram of the AGC. The peak detector detects the peak signal envelope [6]. The VAD detects if a sub-frame of the input signal is speech or noise. the gain controller computes the gain required to enhance the speech signal amplitude. The Amplifier/Attenuator changes the speech signal amplitude by changing the analog gain of the ADC or applying the digital gain to the Pulse Code Modulation (PCM) samples.

Figure 1. AGC Block Diagram



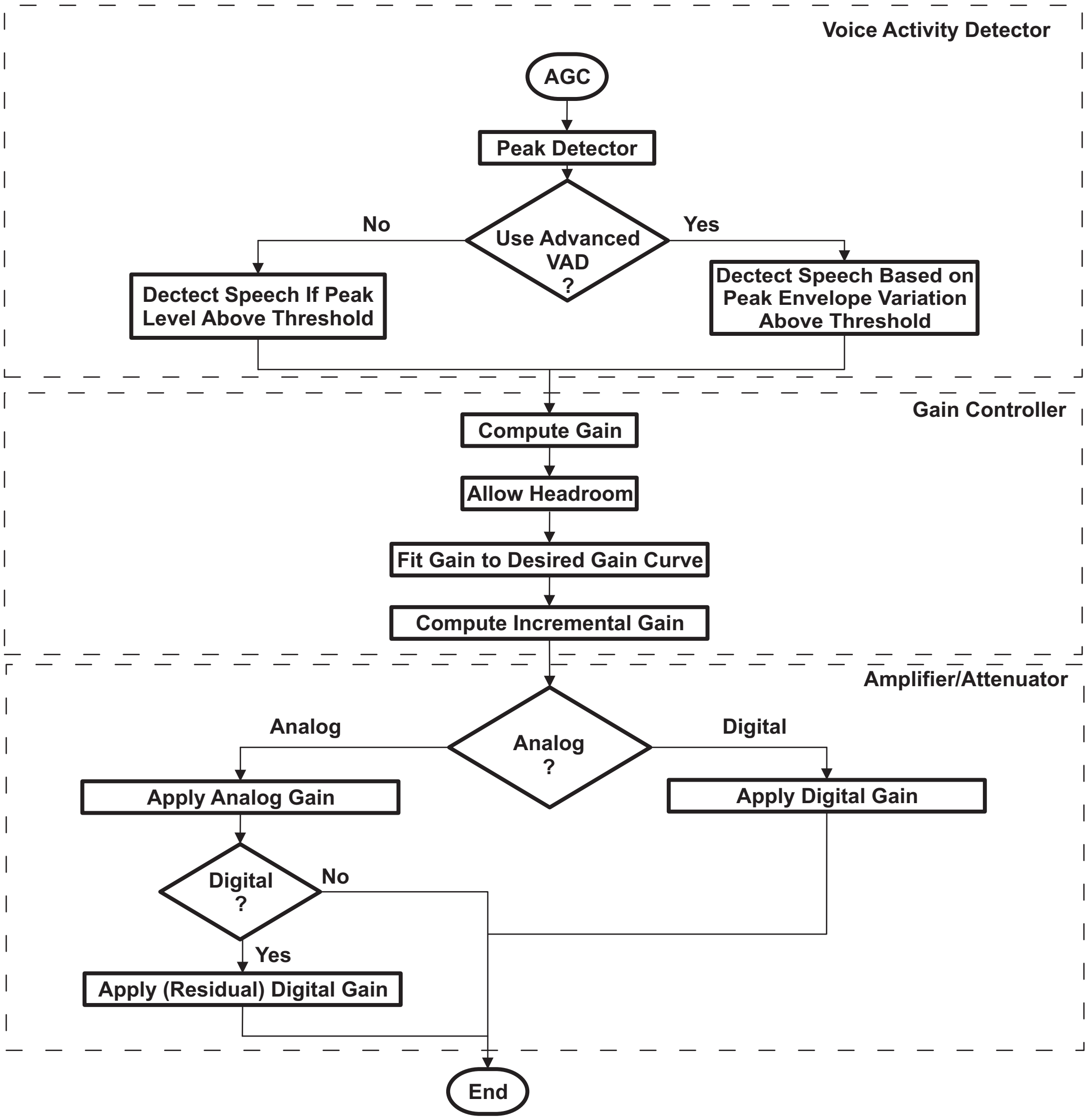
The inputs for AGC are:

- Input data stream
- Algorithm selection parameters such as advanced VAD, gain curve enable, and analog, digital or analog/digital volume control
- Algorithm tuning parameters such as amplitude threshold, peak amplitude variation limit, sub-frame observation count limit, maximum allowed consecutive speech segments, headroom level, gain curve, gain step size, attenuation limit for noise segment, analog gain limits and initial gain of ADC

The input data stream is divided into sub-frames and frames. Sub-frames consist of multiple speech samples. The frame consists of multiple sub-frames [6]. The outputs from the AGC are the digital gain applied linear PCM samples, and/or analog gain and peak signal in current sub-frame.

Figure 2 shows the top level flowchart of AGC. The VAD based on peak signal variation is used if advanced VAD is selected. The speech segment has a larger variation compared tothe stationary noise segment. Otherwise, the signal segment whose peak falls below the threshold is considered as noise. If a specific segment/sub-frame is speech, the gain needed for normalizing the signal just below the headroom level is computed. The computed gain is mapped to the target gain based on a predefined gain curve. Depending on the previous gain value that was applied, the reason for the gain change and the amount of the gain change, the incremental gain change in integral numbers of decibel is computed. The reason for gain change can be attributed to speech to noise segment transition, noise to speech segment transition, headroom attenuation, clip removal, noise attenuation, and signal amplification. If the analog gain is allowed, the analog gain is computed and returned to the system. The ADC driver can be programmed to adjust the ADC analog gain using the gain returned by the AGC. Any residual gain, gain that cannot be applied in the analog stage, is applied digitally if needed. Alternately, digital gain can be selected instead of analog, or analog and digital combination. The integral gain steps are divided into fractional gain steps in digital domain to minimize zipper noise.

Figure 2. AGC Flowchart



2.1 Peak Detector

The peak sample xpk[k] in the current and previous frame is the input for the gain computation and VAD [6].

2.2 VAD

The VAD takes the output of the peak detector as input. The simplest VAD consists of a comparison of the peak signal amplitude and threshold. This is referred to as static VAD since the threshold or noise floor is fixed. The advanced VAD, though of low computational load, uses the amplitude variation of the signal and threshold to determine whether a sub-frame is speech or noise. This is referred to as dynamic VAD, since the noise floor is estimated dynamically.

2.2.1 Static VAD

If the absolute magnitude of the peak input signal in the current sub-frame is above the threshold XPKTH, the signal segment is detected as speech (va[k] =1) as given by (1). If the peak signal is below the threshold, the segment is detected as noise (va[k] =0). When analog gain control is used with static VAD, the threshold needs to be raised or lowered depending on the current and initial analog gain. The gain difference between current and initial is applied to the threshold. This is because the noise floor gets amplified or attenuated depending on the analog gain of the ADC.

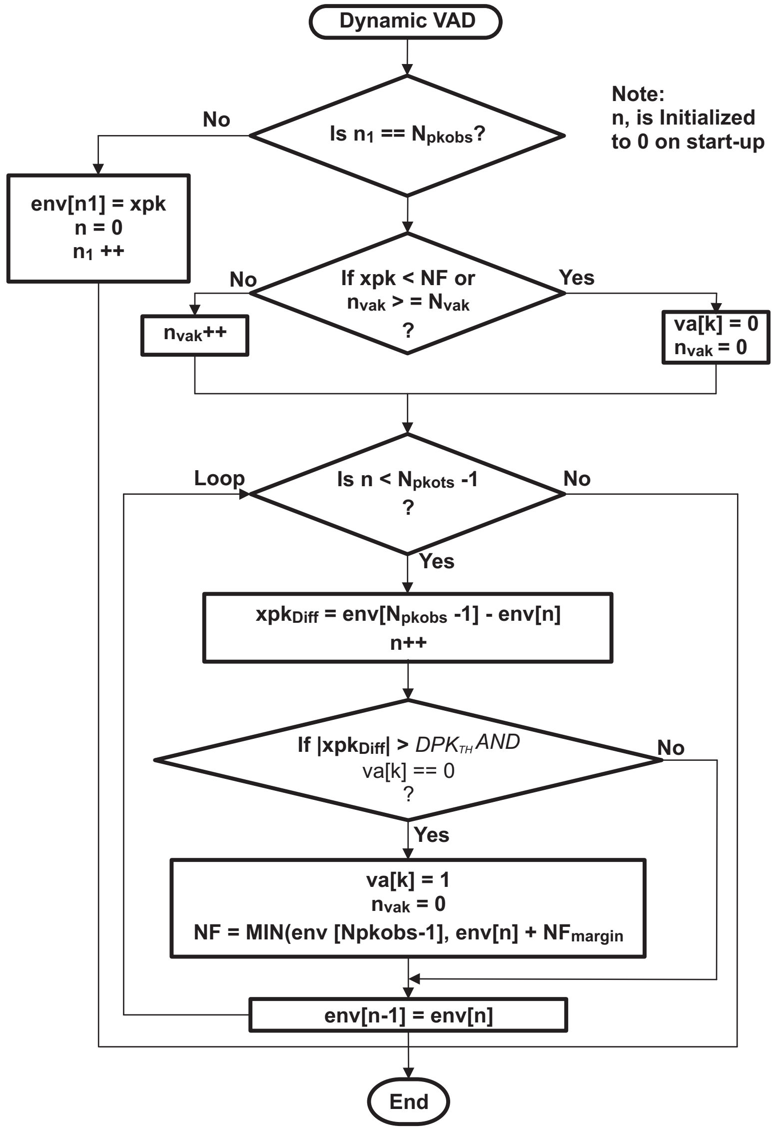
$$\begin{aligned} va[k] &= 1, \text{ if } xpk[k] > XPK_{TH} \\ va[k] &= 0, \text{ if } xpk[k] < XPK_{TH} \end{aligned} \tag{1}$$

The disadvantage of the static VAD is that any change in noise floor level would cause the VAD to output erroneous result. Thereby, AGC can cause amplification of noise.

2.2.2 Dynamic VAD

The stationary noise results in a flat envelope in the absence of a speech signal, regardless of the noise floor level. The speech signal envelope with or without the presence of stationary noise will show variation. This property can be applied to xpk[k] to detect voice activity accurately even when the noise floor level changes. The flowchart for dynamic VAD is shown in Figure 3.

Figure 3. Dynamic VAD Flowchart



Initially, the input signal is assumed to be noise. The VAD algorithm stores past Npkobs number of peak samples xpk[k]. Once Npkobs peak samples are available for the VAD algorithm, the variation of the peak signal across the sub-frames is calculated. If any of the absolute differences between the stored peak samples are more than threshold DPKTH, speech segment is detected (va[k] =1). The lowest xpk[k] of the Npkobs samples is used as the noise floor, NF. The noise floor can be raised by small amount to allow for an additional safety margin. If the signal falls below the noise floor, the signal is detected as noise (va[k] =0). Also, VAD forces the current sub-frame to be noise if speech is continuously detected for more than Nvak sub-frames. This is done to force peak amplitude variation to correct decision errors, if any.

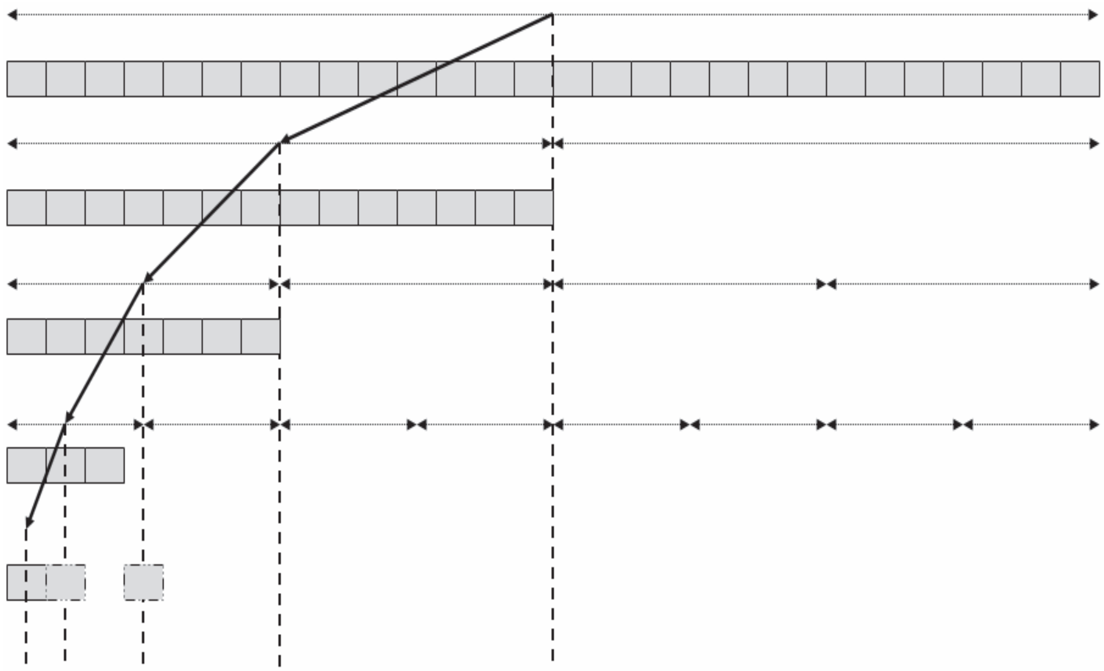
2.2.3 Gain Controller

The gain controller detects the gain change that is to be applied to attain the desired signal level [6]. The gain controller consists of raw gain computation, headroom creation, gain shaping, and incremental gain computation based on the cause for gain change. The incremental gain change is in integral numbers at this stage.

2.2.3.1 Raw Gain Computation

The gain required to level the input signal to the maximum amplitude, 0 decibels Full Scale (dBFS), is stored in lookup tables with the memory address index mapping to gain and the memory content holding the peak level. The table is small in size since the dynamic range of 16-bit numbers is limited to 96 dB. By using xpk[k], the raw gain Graw is fetched from the table. Binary search is used for speeding up the gain selection. The table is divided into two equal halves. The half which contains the gain value is further divided into equal halves. The division is continued till the gain Graw is selected.

Figure 4. Binary Search for Gain Selection



2.2.3.2 Headroom Creation

The signal level should be less than the headroom level, which is below the maximum amplitude (0 dBFS). The headroom level Ghr is subtracted from the raw gain Graw. This can result in attenuation of the signal reaching above headroom. The attenuation can be disabled if the signal segment reaching the headroom region is to be unaltered in magnitude. The resultant Grawh is clipped so that the gain never exceeds the maximum allowable gain. The clipping can be performed in the gain shaping (curve) processing also.

2.2.3.3 Gain Shaping

The gain curve is stored in a lookup table. The gain Grawh is mapped to the target gain Gs using the lookup table. This step aids in shaping the target gain for different needs like preserving perception of distance, leveling of speech segments, and clipping of gain. If gain curve is not specified, gain mapping step is skipped.

In the case of ALC, Figure 5 will remain flat to level the speech signals. The flatness depicts the output signal level being the same for varying input level signals. The amplifier gain will decrease linearly with increasing input signal level as shown in Figure 8 to achieve constant output level. The maximum allowed gain of 40 dB is used to illustrate unintentional gain saturation resulting in linearly changing output signal level (instead of constant output level). The maximum amplification gain value of 20 dB is nominal for the AGC intended for leveling speech signals.

To maintain distance perception, the input and output signal levels have to map linearly as shown in Figure 7. The corresponding gain curve is shown in Figure 10. By selecting gain in such a way that the output signal level is linearly changing with respect to the input signal level, the relative amplitude of the different speech levels is still retained and thus perception of distance is achieved. The disadvantage of this approach is the low energy signals are not amplified adequately to produce intelligible speech.

Figure 6 and Figure 9 illustrate the trade off of distance perception and constant energy level of the output signal. Low energy signals below a threshold are amplified uniformly to give a feel of distance perception. Thus, the low energy signals are amplified for intelligibility and at the same time distance perception is retained. Amplification is reduced with increasing energy level for high energy signals to avoid clipping. This results in high energy signals and/or signals closer to microphone losing the property of distance perception.

Figure 5. Input Versus Output (Non-Linear)

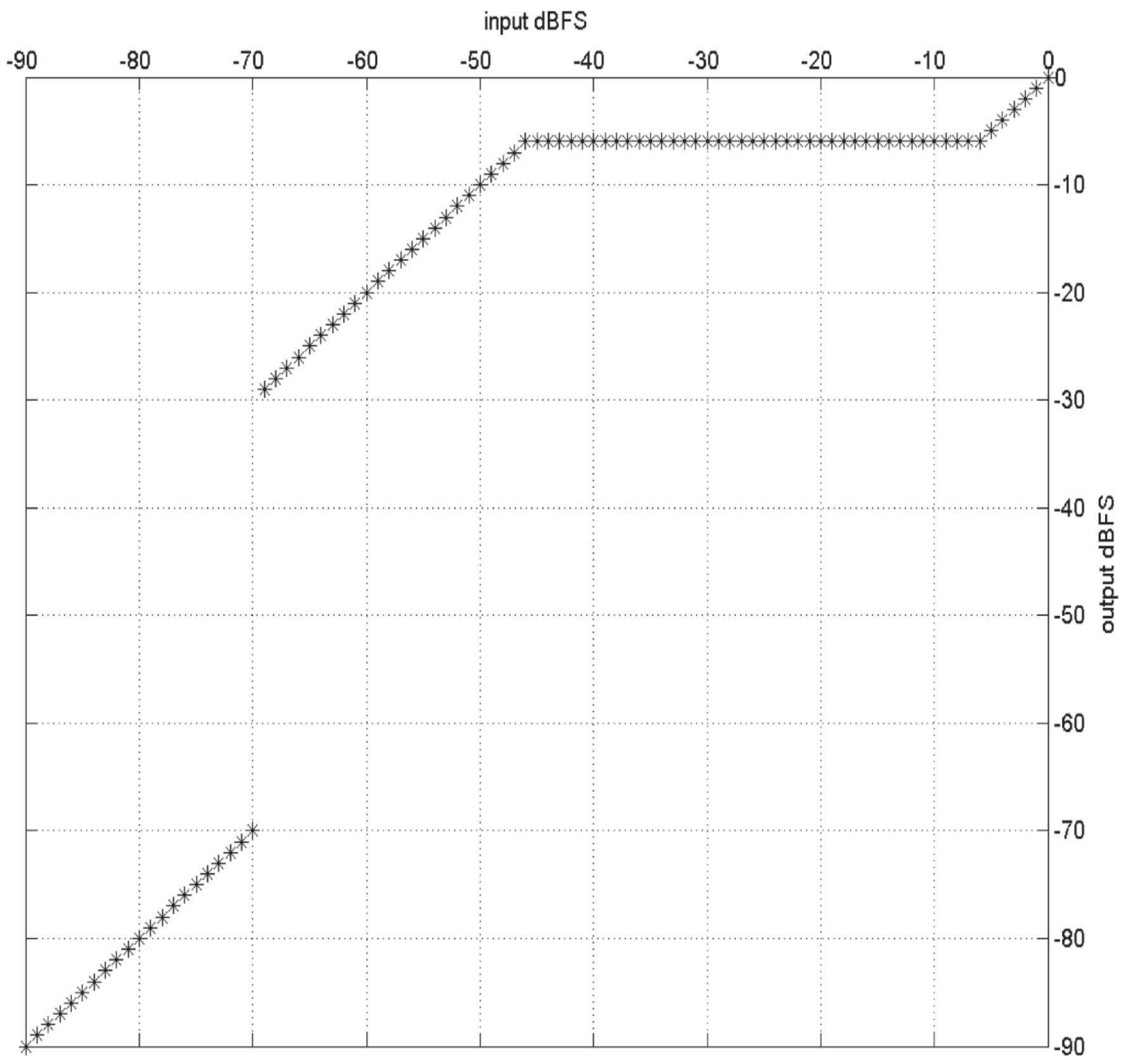


Figure 6. Input Versus Output (Pseudo Linear)

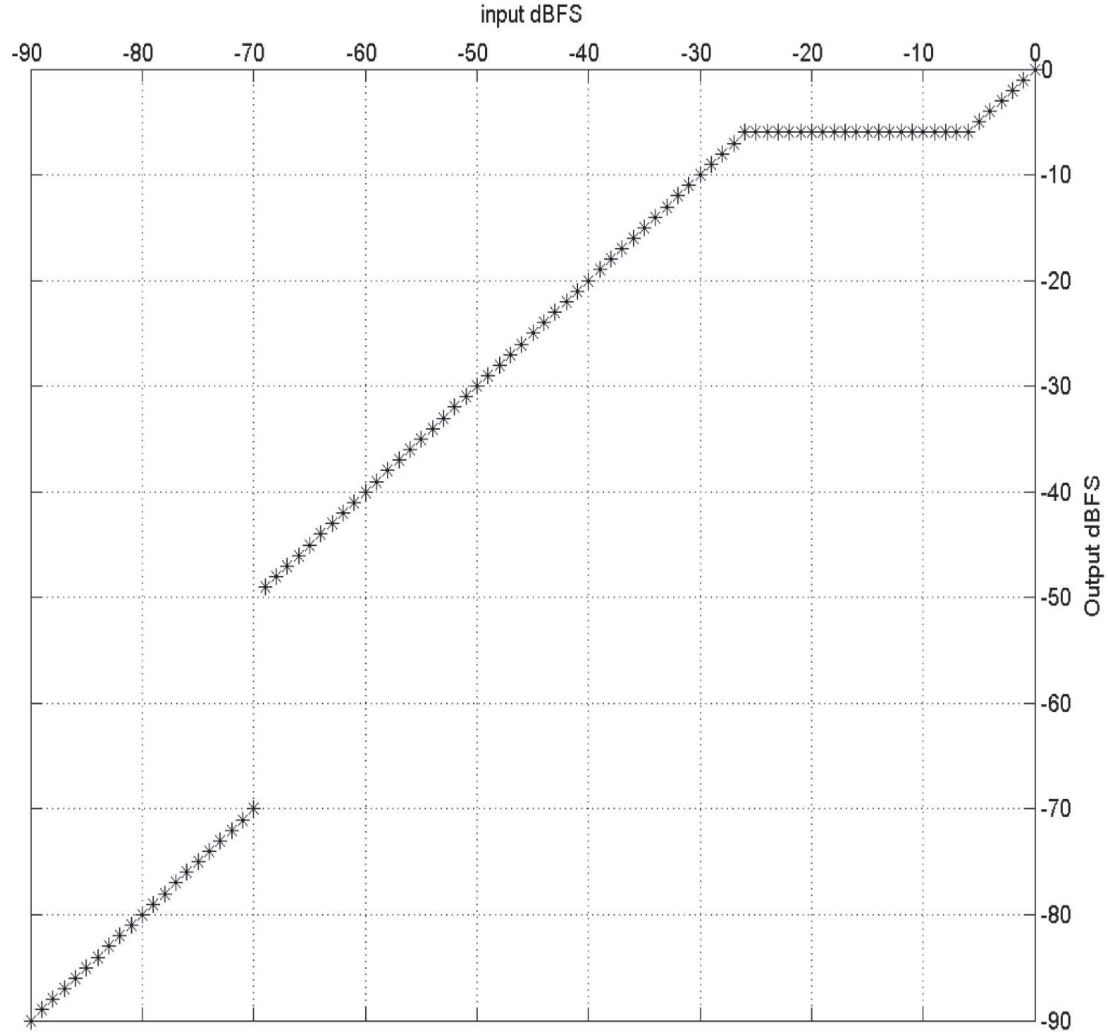


Figure 7. Input Versus Output (Linear)

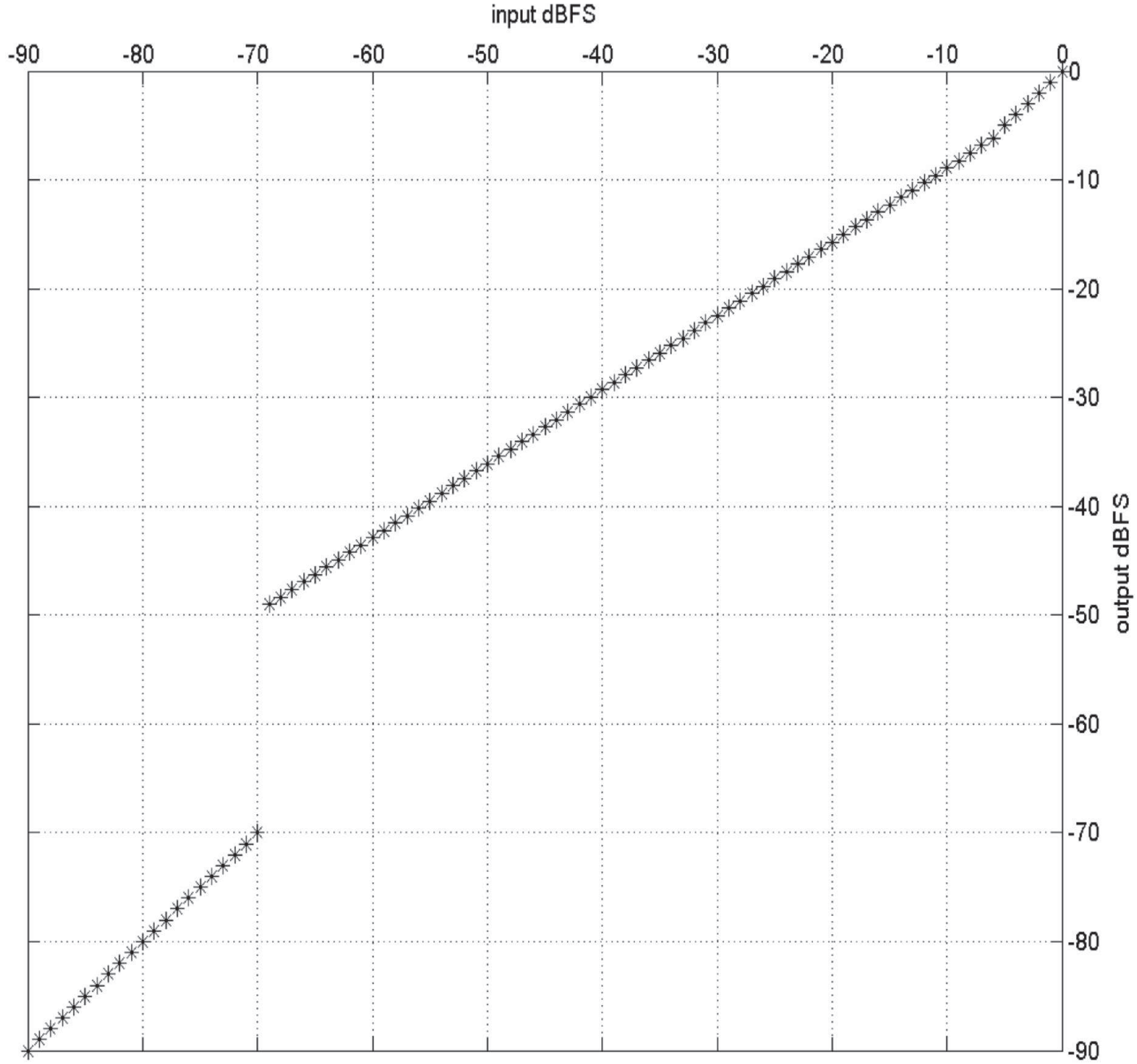


Figure 8. Input Versus Gain (Pseudo Linear)

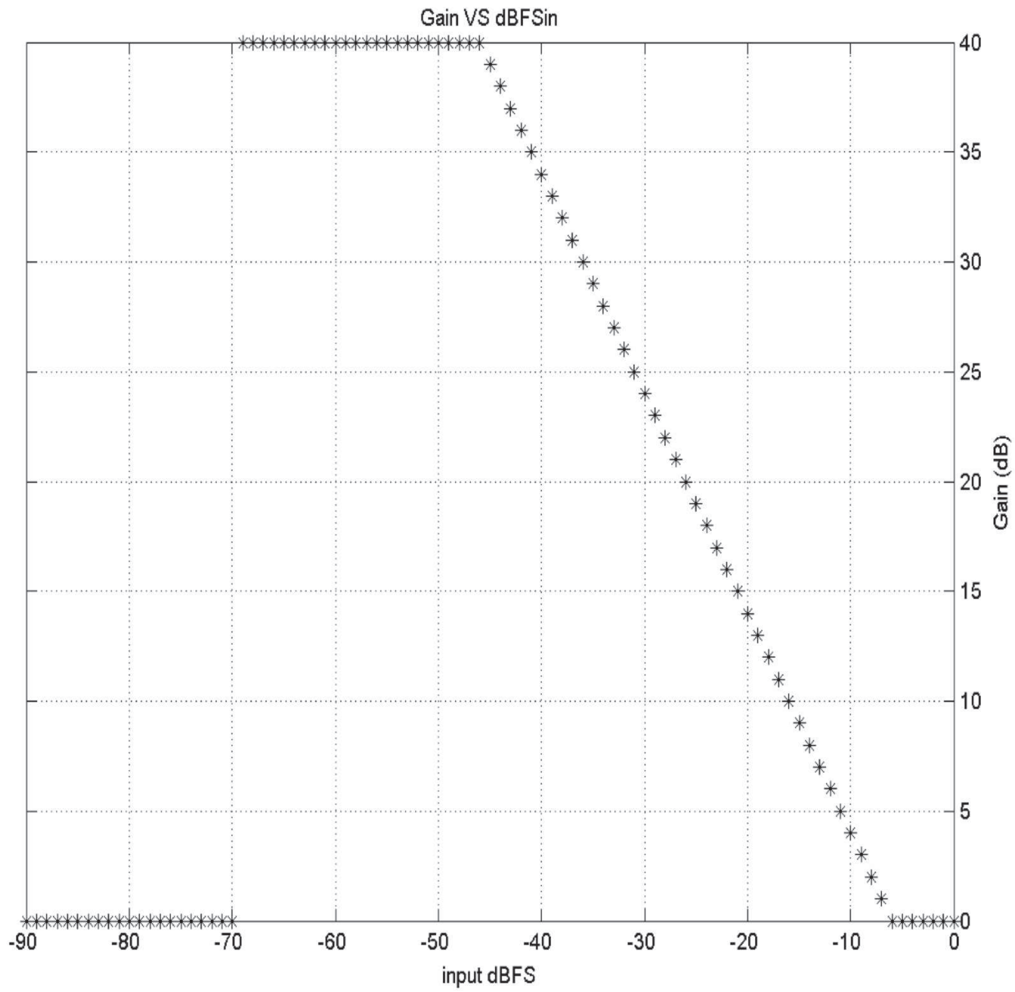


Figure 9. Input Versus Gain (Constant)

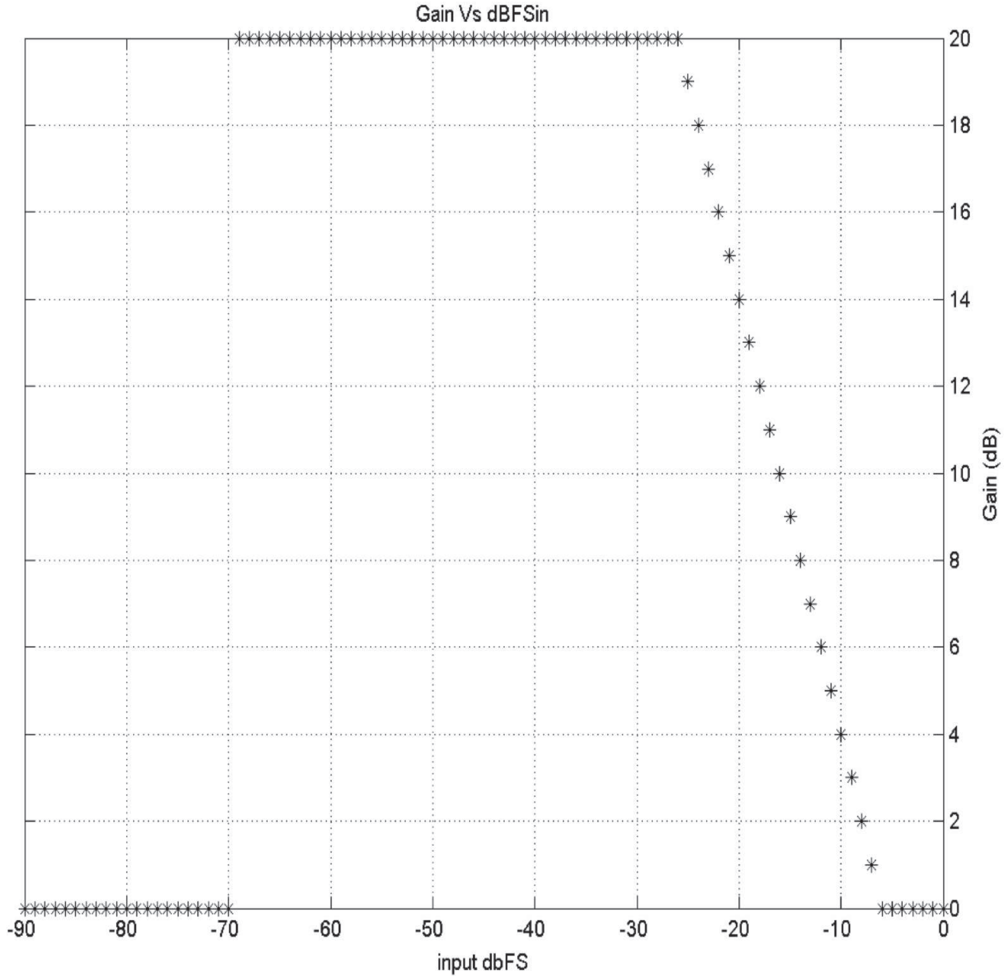
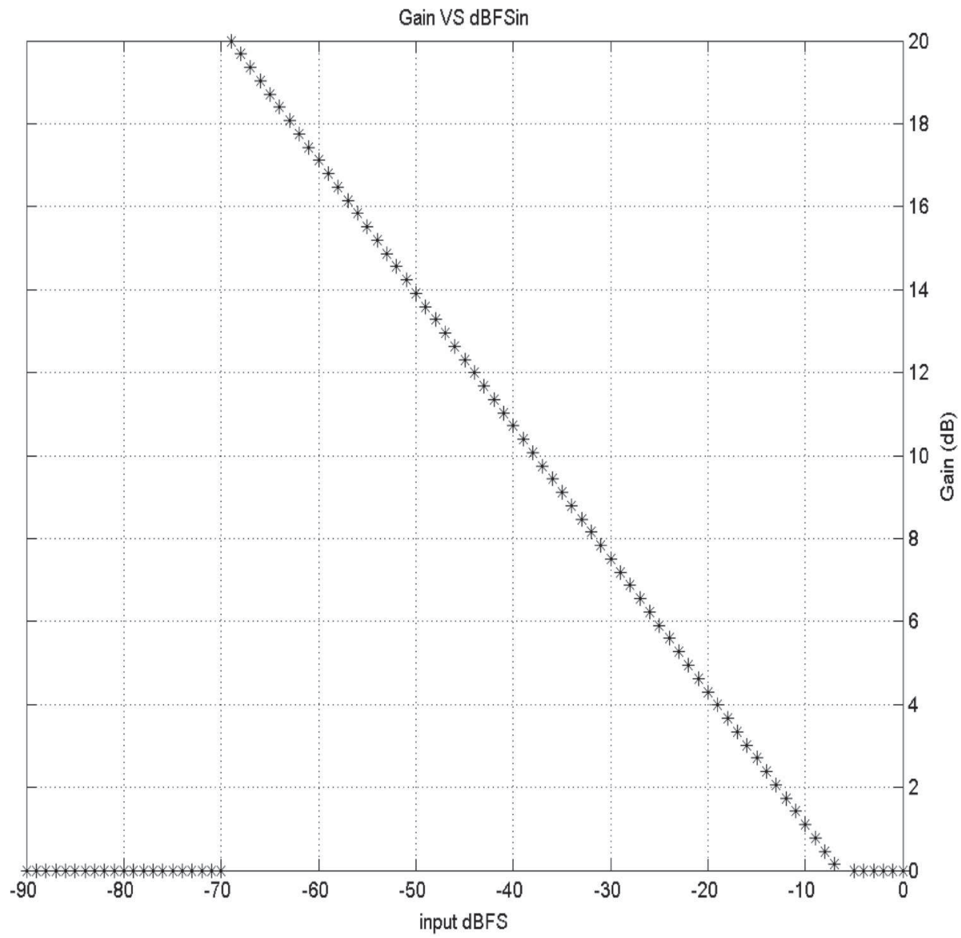


Figure 10. Input Versus Gain (Linear)



Further, no gain is applied for signals below the noise floor and signals above the headroom to prevent noise amplification and possible clipping, respectively. Figure 5, Figure 6, and Figure 7 show this region as linear to indicate no output level change with respect to input level. The headroom region can be attenuated to bring the signals in the headroom region below the headroom level. This feature is especially useful in the case of analog amplification or attenuation in the ADC. Figure 8, Figure 9, Figure 10 and show flat gain of 0 dB to illustrate no amplification or attenuation in headroom region.

2.2.3.4 Clip Detection

In the case where the difference between the previous gain being applied and the current gain being computed (Gact[k-1] – Gs[k]) is greater than the positive threshold GDTH, there will be possible clipping of the signal if the gain change is not carried out rapidly. In that case, a flag (flagClip) is activated to inform the amplifier/attenuator to perform a fast gain change.

2.2.3.5 Noise Attenuation

Since the signal is assumed to be noise at the start, the attenuation feature is disabled till the detection of speech segment for the first time. The gain Gs for noise segment is reduced to the specified attenuation level Gatt.

2.2.3.6 Incremental Gain Computation

If the current computed gain Gs[k] is greater than the previously applied gain Gact[k-1], the incremental change in gain is computed. If the current computed gain Gs[k] is lesser than previously applied gain Gact[k-1], gain reduction is desired. If it is possible the clipping flag flagClip is active, a arger decremental gain is computed to avoid clipping. Otherwise, the normal decremental gain is computed to reduce amplification.

2.2.4 Amplifier and Attenuator

The gain can be applied either on the digital data or analog signal.

2.2.4.1 Analog Gain of ADC

Analog Programmable Gain Array (PGA) on ADC can be used for gain control. The buffering of PCM samples between the ADC hardware and the AGC software has to be bare minimal (for example, two buffers of size 20 ms each). If the delay is not small, the gain is determined from the past samples by the AGC and gain is applied on future samples (with respect to AGC) at ADC. The delay in gain control, if not minimized, will cause undesired saturation and attenuation of the input signal.

If analog-only amplification or attenuation is chosen, the incremental gain is applied by the ADC. The integral gain is programmed into the ADC by a firmware driver. Often ADCs handle zipper noise problem by gradually reaching the target gain or by applying the gain change on zero crossings. The current gain Gact[k-1] is updated with the desired gain Gs[k].

2.2.4.2 Digital Gain

Gain can be applied on digitized PCM samples with a low computational load by fixed point multiplication. The scale factors for applying amplification or attenuation is stored in lookup tables. The fixed point scale factor corresponding to the target gain Gs[k] is fetched from lookup table. The fractional part of the difference between the previous and current scale factor is applied on the input PCM samples to produce output. The fractional part is selected based on reason for gain change, similar to integral gain change in the gain controller. During the sign change of gain being applied, amplification to attenuation or attenuation to amplification transition, the current amplification or attenuation is gradually removed before proceeding with desired attenuation or amplification respectively.

Digital gain eliminates the need for a low delay between the ADC and AGC, as the gain can be applied on the sub-frame for which the gain is computed. Digital gain has the disadvantage that precision does not increase with applied gain as is the case with analog gain adjustment.

2.2.4.3 Analog and Digital Gain

A combination of analog and digital gains can be used for:

- Applying residual gain digitally in case the amount of analog gain change is limited
- Minimizing effect of control delay and/or improving precision of output PCM samples
- Both the above

Analog amplification is applied as first preference. If there is any residual gain to be applied due the maximum allowed analog gain, digital amplification or attenuation of input samples is used. Any reduction in amplification or increase in attenuation is applied first on analog to avoid clipping. Once the signal is clipped at ADC, the digital gain reduction can not reverse the clipping since the clipping operation is non-linear.

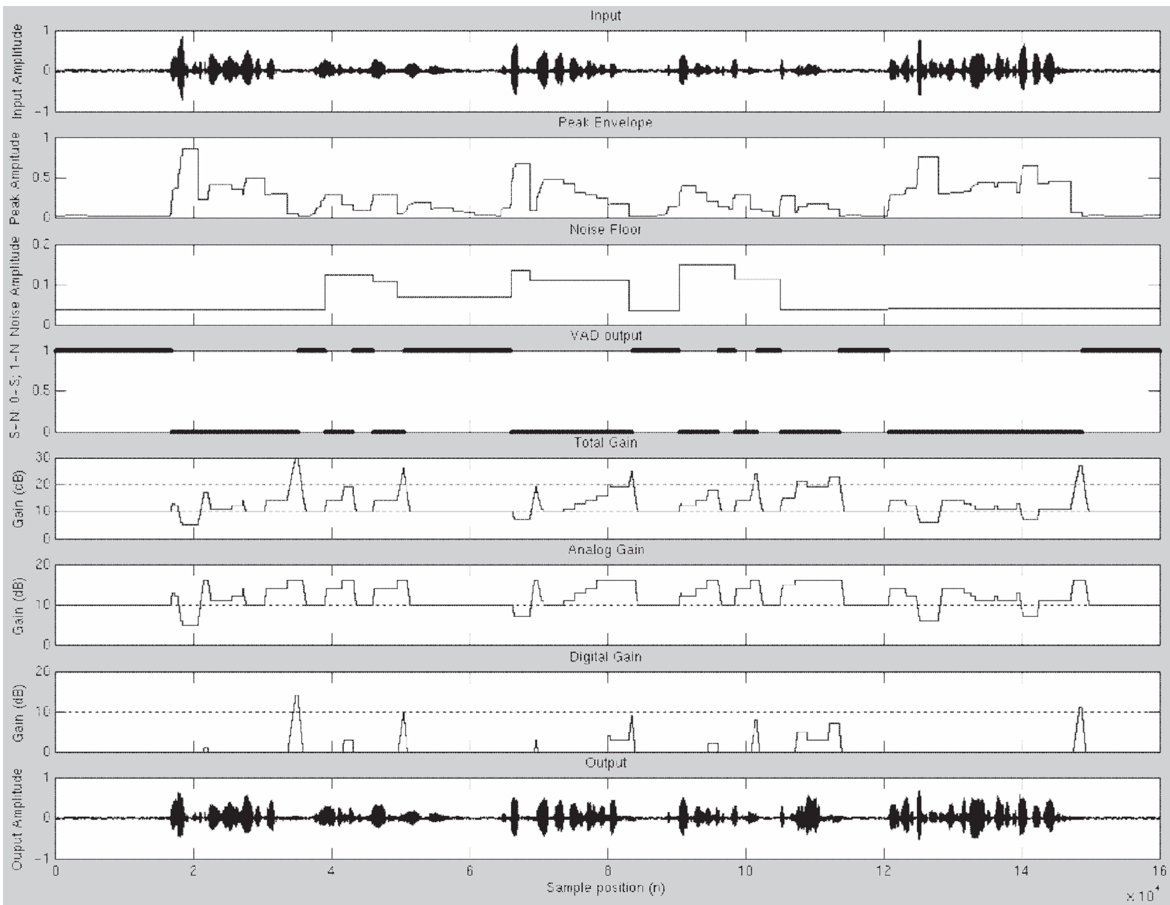
3 Results

The CPU cycle intensive part of the AGC algorithm is peak detection since computation is done using all the samples in the current sub-frame. The rest of the computation in AGC is performed on the sub-frame and frame level. The AGC on the ARM9EJ processor consumes less than 1 MHz.

Figure 11 shows internal waveforms of the AGC algorithm for a speech signal embedded in colored noise, with varying SNR. The peak signal envelope along with the S-N signal when provided to the gain controller results in the gain curve shown in the graphs. The gain curve consists of analog, residual digital, and sum of analog and digital gains. The variation of peak signal envelope is used to estimate noise floor. The noise floor and peak signal level is used for estimating S-N signal. The S-N signal shows some of the speech segments as noise. This error in detection can be eliminated by adjusting the threshold for noise detection. However, it is better to classify speech as noise than vice versa. In the case of a noise segment being detected as speech, noise amplification would result.

The Band Pass Filter (BPF) is used to band-limit the input signal to the speech spectrum [3]. The BPF filter takes about 1 MHz on ARM9EJ. BPF is not required if the input signal is already band-limited.

Figure 11. AGC Internal and External Signal Plot



4 Summary

The peak amplitude variation used for dynamic noise floor estimation is computationally efficient. The noise floor adjustment based on the applied analog gain is effective for the static noise floor threshold. The detection of a voice activity or speech segment in the input signal is less error prone with the use of dynamic noise floor. The dynamic noise floor is effective in eliminating noise amplification. The use of analog and residual digital gain minimizes control delay problem in analog and/or lack of precision change in digital gain. The fractional digital gain application based on the reason for the gain change eliminates or minimizes zipper noise and the possibility of clipping. The integral gain change in analog based on signal characteristic minimizes clipping and zipper noise. The gain curve can be selected to retain distance perception while amplifying weak signals.

5 References

1. *PID control system analysis, design, and technology* by K. H. Ang, G. Chong and Y. Li, Control Systems Technology, IEEE Transactions, Volume: 13, Issue: 4, pp. 559- 576, July 2005.
2. *PID control* by K. J. Astrom and T. Hagglund, In Levine, Ed., The Control Handbook, CRC Press and IEEE Press, pp. 198-209, 1996.
3. *An Efficient Audio Noise Filter* by F. J. Archibald, submitted for publication.
4. *Automatic gain control* by G. Kang and M. Lidd, Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP. Volume 9, pp. 120 – 123, Mar 1984.
5. *Digital Signal Processing In Automatic Gain Control Systems* by G. R. Steber, Industrial Electronics Society, 14 Annual Conference, Volume 2, pp. 381 – 384, 24-28 Oct 1988.
6. *Automatic Level Controller for Speech Signal* by F.J Archibald, submitted for publication

Fitzgerald Archibald earned a B.E (Electronics and Communication Engineering) from PSG College of Technology, Coimbatore, Tamil Nadu, India in 1996. He worked on control systems software development for geo-synchronous satellites from 1996 to 1999 in ISRO Satellite Centre, Bangalore, India. In 2001-2002, he worked on speech decoder, real-time kernel, and audio algorithms for DVD audio team in Sony Electronics, San Jose, USA. While in Philips Semiconductors (Bangalore, India, and Sunnyvale, USA) in 1999-2001 and 2002-2004, he worked on audio algorithms and systems for STB, DTV, and internet audio. He is part of the Personal Audio Video and Digital Imaging groups in Texas Instruments Inc, Bangalore, India from 2004-till date working on audio, video, and imaging systems and algorithm development. Interests include multimedia and control algorithms and systems.

Mr. Archibald is member of AES.

IMPORTANT NOTICE

Following are URLs where you can obtain information on other Texas Instruments products and application solutions:

Products		**Applications**	
Amplifiers	amplifier.ti.com	Audio	www.ti.com/audio
Data Converters	dataconverter.ti.com	Automotive	www.ti.com/automotive
DSP	dsp.ti.com	Broadband	www.ti.com/broadband
Clocks and Timers	www.ti.com/clocks	Digital Control	www.ti.com/digitalcontrol
Interface	interface.ti.com	Medical	www.ti.com/medical
Logic	logic.ti.com	Military	www.ti.com/military
Power Mgmt	power.ti.com	Optical Networking	www.ti.com/opticalnetwork
Microcontrollers	microcontroller.ti.com	Security	www.ti.com/security
RFID	www.ti-rfid.com	Telephony	www.ti.com/telephony
RF/IF and ZigBee® Solutions	www.ti.com/lprf	Video & Imaging	www.ti.com/video
		Wireless	www.ti.com/wireless

Mailing Address: Texas Instruments, Post Office Box 655303, Dallas, Texas 75265