VOCAL

Dual Microphone Adaptive Noise Reduction Software

White Paper

12/15/2015

Dual Microphone Adaptive Noise Reduction (ANR) Software

Introduction

In any telecommunication system, the microphone will pick up not only the signal of interest, speech, but also any other sound source located in the same acoustic environment.   For example, when using a bluetooth car phone, the car engine, tire noise and wind result in noisy speech signal, that lacks intelligibility and quality. In our continuously connected world, we have gained accessibility, mobility and convenience, but lost the overall quality in our communications system. However with the aid of speech enhancement and noise reduction techniques we can improve the quality of speech. This paper discusses the fundamentals of noise reduction and the benefits of dual microphone adaptive noise reduction (ANR) software.

The Evolution of Noise Reduction Algorithms

The one of the earliest algorithms for the noise reduction for speech applications is spectral subtraction. This was logical first step because the noise source is considered uncorrelated and independent of the speech signal. Once the noise spectral is estimated it can be subtracted from the captured signal, improving the SNR of the signal.

Over the years, this simple technique has been researched and improved extensively, however there are some limitations to this approach. Single channel suppression based solutions rely on being able to estimate noise spectrum directly in the presence of speech. Obtaining estimates of the noise spectrum requires long-term averaging and minimum statistics. This is problematic when the noise source quickly changes over time and is non-stationary. Ultimately resulting in lower SNR improvement and speech distortions.

However, the problem can be solved using a different approach with the addition of a second microphone. If the second microphone is able to sample the noise of the acoustic environment without the presence of speech, then that signal can be subtracted from the original microphone. This changes the problem from trying to estimate noise spectrum to estimating the transfer function between the two microphones.   This is a better solution because the noise spectrum can change quickly over time, but the transfer function typically changes much more slowly. Let's go back to the bluetooth car phone example. In this scenario the location of the engine and the tires, and microphones are all stationary. The RPMs of engine and the surface s the tires travel over will change, resulting in varying noise spectrums, while the transfer functions between the components will remain relatively stationary.

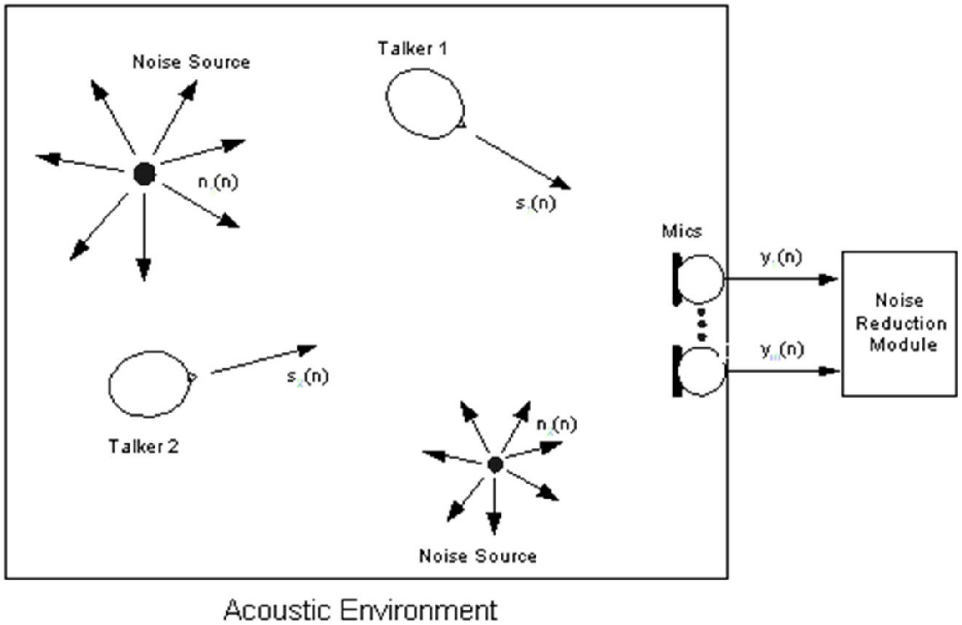

Figure 1. Noisy Acoustic Scene

The Fundamentals of Dual Channel Adaptive Noise Filtering

Dual channel ANR is an system identification adaptive filtering problem, very similar to line and acoustic echo cancellation. As shown in Figure below , one input source, y1(n), contains a desired signal and a linear transformed copy of an the noise source. The second input source, y2(n), contains a reference of the noise source signal. An adaptive filter can be applied to the y2(n) and subtracted from y1(n) to remove noise source from y1(n) to generate e(n). The noise source in signals y2(n) is correlated and linear transformed version of the noise source in y1(n), so optimal filter can be found which will minimize the output power of e(n).

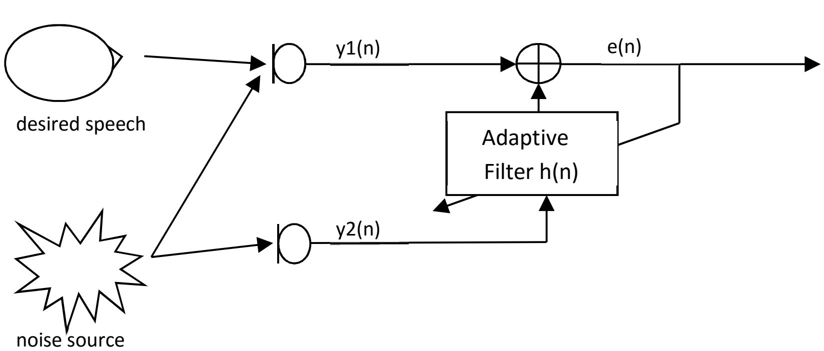

Figure 2. Adaptive Noise Filter

Any adaptive filtering algorithm can be used to determine the optimum filter, but the normalized least mean squares(NLMS) is most commonly used because it provides the best tradeoff between convergence speed and computational complexity. The equation for the filter update is as follows:

$$h(n+1) = h(n) + \frac{\mu \cdot \overleftarrow{y2}(n) \cdot e(n)}{\left\| \overleftarrow{y2}(n) \right\|^2}$$

The optimal filter is reached when the e(n) becomes zero.

The goal of any adaptive filtering algorithm is to converge fast and have a minimum misadjustment. The step-size parameter, μ, controls the adaptation rate in conjunction with the error. When error signal is large, a large step-size should be used, and as the error gets smaller the size-size should get progressively smaller too. To use a golf analogy, on a particular hole you might start 400 meters away from the hole, to quickly get to the hole, you first use your driver, then as you get closer you use a lesser club and don't swing as hard. If you swing hard and only use your driver you will never get close to the hole. Therefore, a variable step-size is used to achieve quick convergence and good steady-state cancellation.

Beside the adaptive algorithm itself, control logic is need around the adaptive filter. Training of the adaptive filter needs to be stopped when speech is present because the speech is uncorrelated with the noise signal and will cause the filter to diverge. The same methods use in echo cancellation for doubletalk detection (DTD) can be used in this adaptive noise reduction algorithm to prevent divergence. Voice activity detectors (VAD) are also necessary because the speech signal can potentially leak into the noise reference signal. Adaptation needs to be stopped during voice active periods to prevent self-cancellation of the speech.

In theory, an adaptive filter can perfectly cancel the noise, but in practice there are several limitations. Each microphone has their own inherit, uncorrelated noise source. In the spec sheets for microphones one of the characteristics of the microphone is the A-weighted SNR. Typical electret condenser microphones has an SNR of 60dB. This is the new bound, but when you also consider quantization noise and error, and the fact that the transfer function between the noise reference microphone and the speech microphone is not completely linear and stationary, real-world realizations of ANR achieve about 10 to 15 dB of cancellation.

The Real World Implementation of ANR

VOCAL's Dual Microphone Adaptive Noise Cancellation Software can remove additive noise even when the signal to noise ratio (SNR) is very low, about 0 dB. In the first example, a clean speech segment was mixed with noise at a SNR of 15dB. The algorithm accurately removed the noise, while leaving the speech signal largely intact. This is shown in Figure 3.

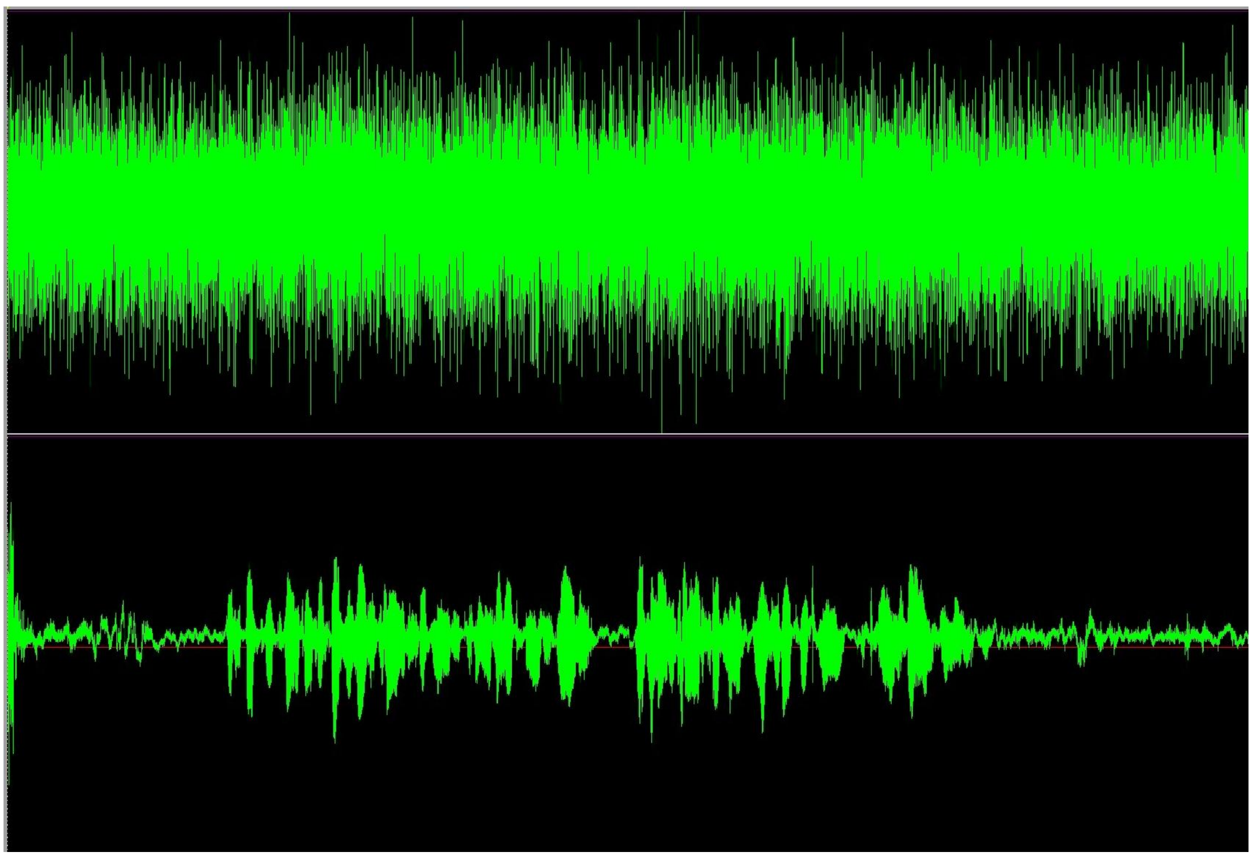

Figure 3. Noisy Input Signal (Top) and ANR Output Signal(Bottom)

In examining Figure 3, we see that the noise begins very high. This is illustrated by the large signal power to the left of the bottom trace. Quickly however, the signal power decreases, showing the software beginning to remove the noise. As the previous section noted, this is due to the adaptive filter converging to the true transfer function. An illustration of this performance is shown in Figure 4.

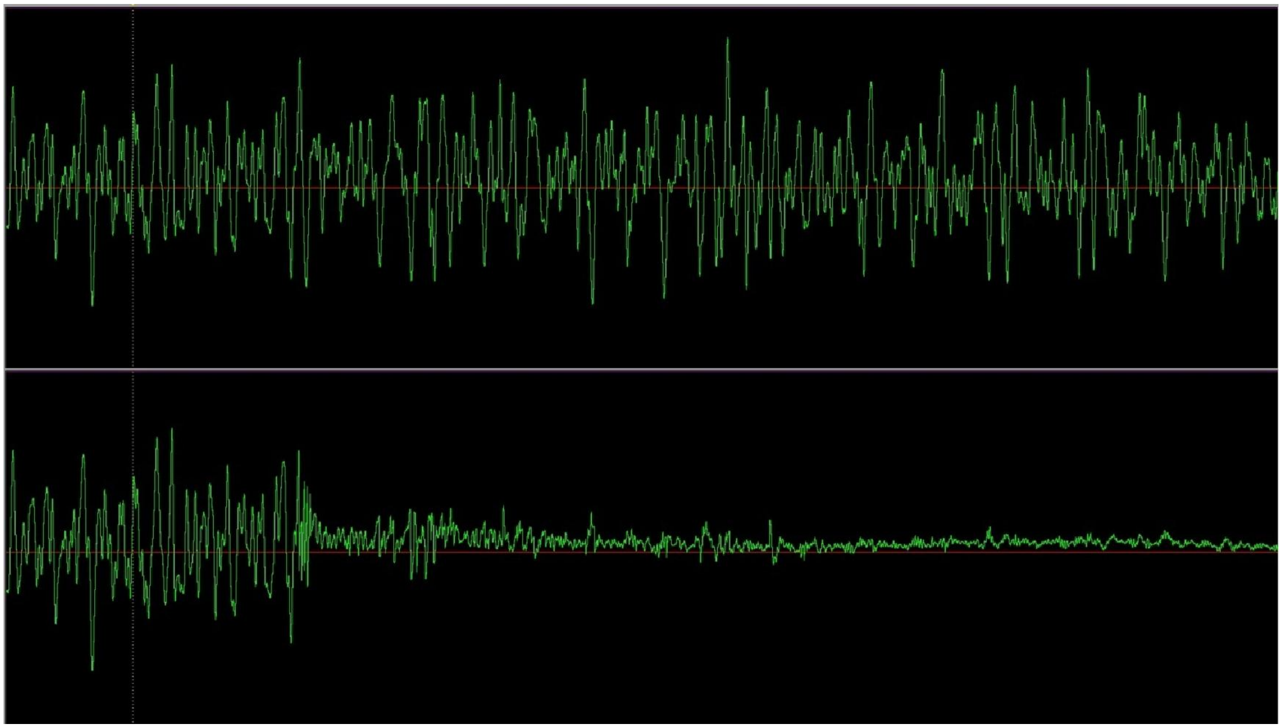

Figure 4. Initial Convergence -Noisy Speech Signal (Top) and ANR Output Signal (Bottom)

The top part of Figure 4 shows the reference noisy speech signal, while the bottom shows the error signal. As we can see, the filter adapts quickly to the environment, in less than 250 ms at 16kHz sampling the noise can be accurately cancelled.

The chart shows the amount of noise cancellation achieve over time. In 1000 samples (62.5ms at 16kHz) more than half the of noise energy has been reduced. Steady-state convergence of over 15dB is achieved in 8000 samples (500ms at 16kHz).

Time (ms)	Noise Reduction (dB)
62.5	0
125	4.71
250	13.63
500	15.44

This robust cancellation produces almost no reduction is speech quality or intelligibility and introduces minimal distortions. The before and after of the speech spectrum is shown in Figure 5.

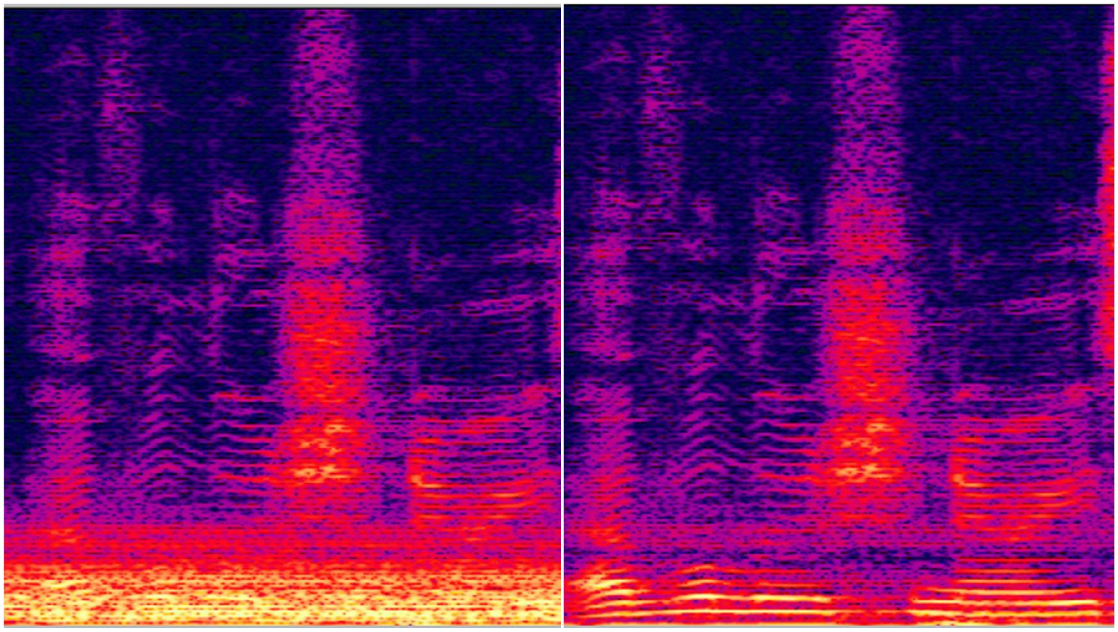

Figure 5. Speech Spectrum Before And After ANR

Before the noise is cancelled, the speech harmonics and lower formants are completely masked by noise. Formants are important for determining the type of vowel being spoken, and hence important for intelligibility, while the harmonics are important for speaker identification and perceptual richness. After the filter cancels the noise, both the harmonics, the formants, and their transitions within connected speech are preserved. Clearly, the algorithm has simultaneously eliminated the noise while offering no visible distortion in the speech spectrum.

The ANR Software Package

VOCAL's ANR is available as a standalone module or as part of a comprehensive VoIP System customizable to most platforms and environments. The ANR is available for several different platforms, including true DSPs such as Texas Instrument's 64x, 62x, 55x, and 54x families as well as ADI's Blackfin, SHARC and 218x families. General-purpose CISC/RISC architectures such as PowerPC, PowerQUICC, x86, ARM, and MIPS are also supported. MIPS and memory numbers of the ANR software on the TI C5515 DSP and the ARM Cortex M-4 are presented below, benchmarks for other platforms can be generated upon request.

Processor	ROM (bytes)	RAM (bytes)	MIPS
TI C5515 DSP	16,320	3,582	13.6
ARM Cortex M-4	13,991	3,592	14.4

Please note that the ANR is also available as platform-independent C Reference Code, using macros to implement MIPS and/or Memory sensitive operations. This greatly simplifies the porting and optimization of the ANR to new platforms. It uses a highly optimized Block NLMS algorithm, providing a short convergence time, and low processor requirements. Please contact us for more information or to arrange a demonstration.

Conclusion

The addition of a noise referencing microphone provides communication systems the opportunity to apply adaptive noise filtering to remove non stationary noise sources from noisy speech signals. Improving the speech quality and intelligibility while introducing minimal distortions. The principals of adaptive filtering allow the solution track changing noise environments to provide continuous noise reduction. The solution can be applied in almost any acoustic environment seamlessly. VOCAL's software implementation is optimized and portability to many platforms making it a great solution for many speech communication applications.