WHITE PAPER

# Moving Noise – Voice Processing for Hands-Free Car Kits

**Andre Coucopoulos and Dean Morgan**

May 2006

ZARLINK SEMICONDUCTOR

Demand for higher quality in-vehicle hands-free voice systems is growing. Many governments have already passed or are considering legislation that prohibits the use of handheld mobile phones while driving. In addition, car manufacturers are responding to consumer demand for higher performance in-car communication and audio systems.

However, moving vehicles pose a hostile environment for mobile voice communications technology. In particular, acoustic echo coupled with high levels of road and car noise decreases the sound quality of a conversation. To boost the performance of hands-free car kits, designers must find a solution that cancels echo, reduces background noise and enhances double-talk through full-duplex operation.

This article provides a brief overview of the hands-free communication market, introduces the challenges faced by car kit designers, and looks at voice processing solutions available today.

## THE NEED FOR HANDS-FREE COMMUNICATION

Previous hands-free car kit solutions have met limited market success, primarily because of their inability to support high-quality voice communication. A poorly designed car kit will deliver voice service far worse than any cellular voice call. In general, lower cost has won out over performance in car kit design. For example, a low-quality car kit will support only half-duplex operation, meaning only one call participant can speak at a time.

This is now changing. As government legislation is implemented around the world, drivers will need to install a hands-free car kit if they want to communicate inside a moving vehicle. Subsequently, consumers will demand better performing hands-free communication technology.

While Bluetooth headsets have made some strides in the hands-free communication market, hands-free car kits offer some significant performance and safety advantages. A high-performance car kit will deliver better voice quality over a Bluetooth headset solution. In terms of safety and ease-of-use, little is required for the hands-free car kit user to receive a call. In comparison, using a Bluetooth headset may require searching for the headset and placing it in the right position to accept the call.

## CHALLENGES FOR DESIGNERS

Moving vehicles create complex noise conditions that must be overcome in the design of high-performance car kits. Many of the hands-free car kits available today are based on technology originally developed for speakerphones. However, noise conditions inside a moving vehicle are far more obtrusive and complex than anything encountered in an office meeting room.

### Acoustic Echo Sources

There are two sources of acoustic echo – direct acoustic coupling between the speaker and the microphone, and voice signal reverberation inside the passenger cabin.

The majority of acoustic echo results from direct acoustic coupling, which occurs when the microphone picks up the voice signal directly from the speakers. This is often exaggerated in a car environment, as the volume of the hands-free car kit is set loud to overcome engine and road noise.

The second source of acoustic echo is indirect coupling or cabin reverberation. The audio signal reflects off different surfaces inside the car cabin. The intensity of the echo depends on the type of materials used in the interior of the car. Harder surfaces reflect more audio while softer surfaces absorb more echo. For example, car seats will absorb more echo while windows will reflect more echo. As a further complication, the echo created by reverberation inside the vehicle is also delayed. The amount of echo delay will vary depending on the size of the car cabin and materials used. Figure 1 illustrates direct acoustic coupling and cabin reverberation echo.

An echo tail of 64 ms to 128 ms is normally sufficient to cancel echo inside the car cabin.

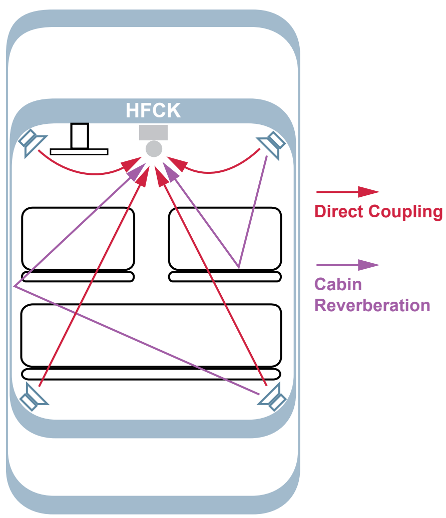


*Figure 1: Direct coupling and cabin reverberation are two sources of acoustic echo inside a vehicle*

## Noise Sources

Noise, whether created by the vehicle or the operating environment, is the main challenge that hands-free car kits have to overcome. Of particular concern are variable levels of noise. The most significant sources of noise inside a car are:

**Engine noise:** Noise generated when the car is operating can degrade the voice quality performance of a communication system. Engine noise reaches its peak when the car is accelerating, and is more noticeable in sports cars during fast acceleration.

**Road noise:** Friction between tires and the road surface generates noise. The intensity of the noise depends on the tire tread and the road material. For example, summer tires are quieter than winter tires and asphalt roads are quieter than concrete roads.

**Wind noise:** Air flowing over the car and hitting protrusions, for example rearview mirrors or roof racks, or driving with open windows creates wind noise. Wind noise varies from vehicle to vehicle. For example, a sleek sports car will have less wind noise than a SUV. Wind noise is also controlled by the amount of sound insulation installed by car manufacturers. High-end luxury cars have better sound insulation than lower cost compact cars. Driving with open windows also creates large pressure differences between the inside and outside of the car body.

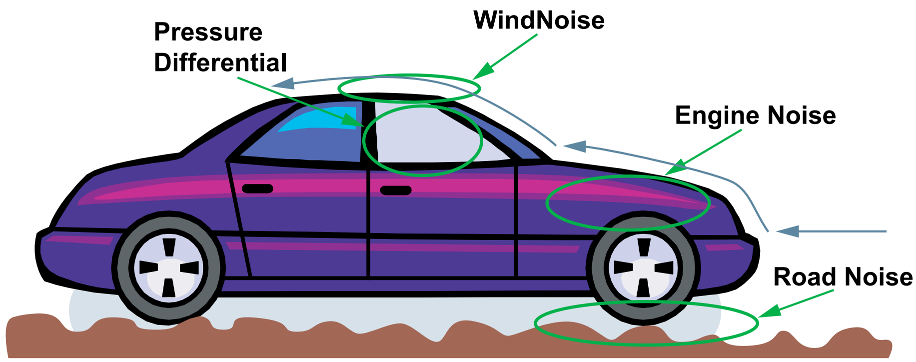


*Figure 2: Noise sources that challenge the design of in-car hands-free communication systems*

Noise reduction algorithms can be used to alleviate the effects of wind noise, road noise and engine noise on the sound quality of the communication system. For wind noise with high sound pressure, for example when car speed is over 100 km/h and the windows are open, the analog front-end circuitry will get saturated and clip the sound signals. Signal clipping degrades the overall performance of both the acoustic echo cancellation (AEC) and total noise reduction.

There are two options to solve this problem. The first option is to design the input circuit for worst case conditions and set the microphone gains to a level where the input never clips. The dynamic range of the analog-to-digital converter (ADC) is reduced, the signal-to-noise ratio (SNR) deteriorates and so does the overall performance of the AEC.

The degradation of AEC performance is acceptable in noisy environments because residual echo is masked by the background noise. In a quiet environment, residual acoustic echo is more noticeable and the degraded performance would not be acceptable. For example, if noise inside a car adds 30 dB of signal strength to the audio path, the designer could attenuate the input gain by 30 dB to compensate for noise. However, when the car stops and the noise source disappears the SNR will be 30 dB higher than the required level. The increase of SNR will reduce the level of acoustic echo cancellation.

Another issue with setting the gain level high is variation of speech levels during a conversation. When speaking, we adjust our voice level according to the noise conditions. Designing for worst-case noise conditions will result in good performance when the noise level is high inside a car but performance will degrade when the noise level is low.

The ideal solution is to dynamically change the microphone gain depending on the environment. If the system is based on a simple analog gain controller, detecting a change in the gain level of the microphone will be considered as a change in the echo path and will force the AEC to reconverge.

Car noise is generally low frequency. Many hands-free car kits use a high-pass filter to remove car noise. A high-pass filter removes the majority of noise but also impacts voice quality. To achieve high performance, the algorithm must be designed to distinguish between the properties of voice and noise signals and remove the noise from the total signal. A good noise reduction algorithm achieves a balance between removing noise and preserving the integrity of the speech signal.

To achieve the required performance, the system designer has to add an automatic gain control (AGC) on the microphone input (Sin). This AGC eliminates signal clipping while maintaining the overall gain setting determined by the user. Using the analog gain on the microphone input (controlled by the digital signal processor) and implementing algorithms that compensate for the abrupt echo path change (which would normally cause a reconvergence) will eliminate signal clipping.

This method allows the ADC to have maximum dynamic range over a broad range of environmental conditions. It also allows the call participants to speak at normal sound levels regardless of the noise conditions inside the vehicle.

## ECHO CANCELLATION AND FULL-DUPLEX OPERATION

Figure 3 illustrates a basic block diagram for a hands-free car kit. The line echo canceller eliminates echo generated by the network. If a cellular phone with an analog port is used to connect to the car kit, a line echo canceller is required to eliminate the echo caused by the side tone on the analog port. If a cellular phone with a Bluetooth connection is used to connect to the car kit, the side tone does not exist, therefore the line echo canceller is not required.

The acoustic echo canceller block reduces the echo and the non-linear processor eliminates residual echo. The noise reduction circuit block reduces background noise.

Other important features of a high-performance car kit are full-duplex operation and good double-talk (two people talking at the same time) performance. Full-duplex car kits allow both voice signals to pass simultaneously, supporting a natural conversation. Most commercial car kits have a half-duplex fallback where voice signals pass in only one direction. In the case of double-talk only the signal with the highest power level is passed and the conversation becomes stilted.

Most full-duplex algorithms today perform poorly during double-talk conditions. These duplex algorithms freeze adaptation and the algorithm then has to re-adapt to the echo environment when double-talk ceases. This can result in an audible burst of echo at the end of single-talk and during double-talk situations. To maintain high performance during double-talk, it is essential for the algorithm to continue to converge and track changes in the echo path.

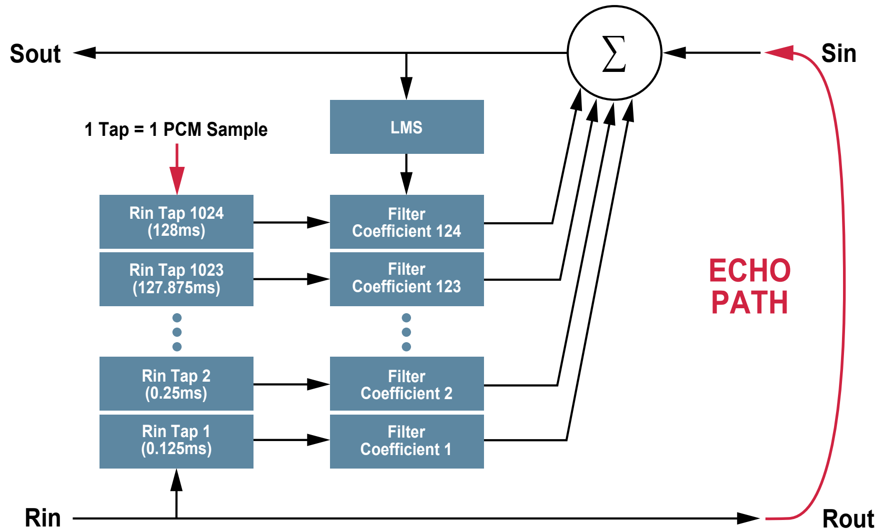


Figure 3: Hands-free car kit block diagram

Many commercial algorithms available today offer full-duplex operation but deliver variable performance. The key components of a full-duplex algorithm are the acoustic echo canceller and the non-linear processor. The interaction between these two blocks determines the quality of the full-duplex algorithm.

## Acoustic Echo Canceller

Figure 4 shows an acoustic echo canceller block diagram. The acoustic echo canceller uses a least mean square (LMS) filter, or similar type, to generate a model of the echo profile. The echo profile signal is used to estimate the echo signal that is subtracted from the input signal in the echo path. The algorithm continuously tracks changes in the echo profile and updates the estimated echo.

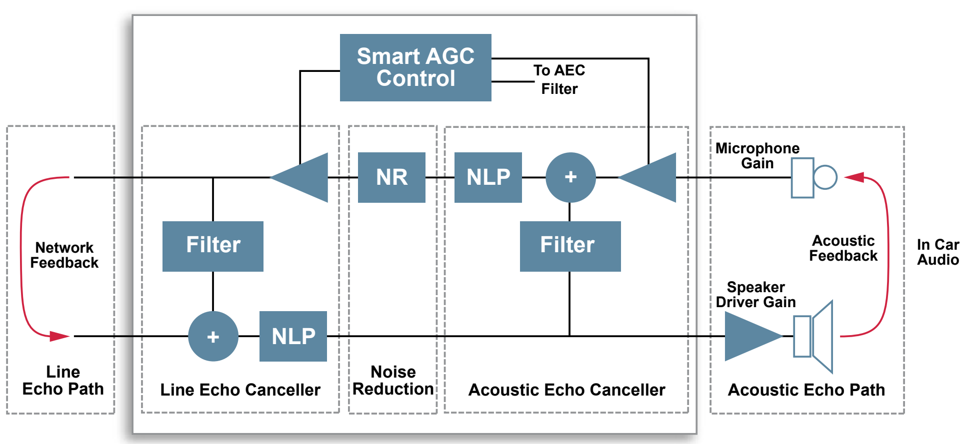


Figure 4: Acoustic echo canceller block diagram

Echo cancellation works well when the echo path is linear, but tends to diverge under non-ideal conditions (e.g. double-talk). To avoid divergence during double-talk, many commercially available full-duplex algorithms freeze adaptation and stop tracking the source but continue to cancel echo based on the profile prior to freezing adaptation. When the profile of the source signal changes while in double-talk, the algorithm can't track the source and the end-user hears residual echo or a burst of echo.

This approach leads to poor performance because any change in the echo path generates residual echo that will not be canceled by the echo canceller until the double-talk condition ceases. This results in higher residual echo during double-talk conditions, or bursts of echo at the beginning of single talk as the algorithm re-convergences.

For best-in-class echo cancellation performance, the algorithm must converge continuously during double-talk conditions.

## Non-Linear Processor (NLP)

Any acoustic echo canceller is not perfect. It is impossible to accurately estimate non-linearity in the echo path. The sources of non-linearity are quantization in the codec and clipping and distortion in the echo path. Typically an acoustic echo canceller can cancel up to 30 dB of echo. To achieve better performance an NLP is required to reduce echo and handle non-linearity.

A simple NLP is an active switch. When activated during single-talk periods the NLP opens the path and injects comfort noise, which allows very good voice quality performance. During double-talk conditions the system relies solely on the echo canceller.

Many acoustic echo cancellers depend heavily on the NLP to achieve fast convergence and mask poor performance of the echo canceller. If the NLP does not turn off fast enough it causes speech to break up during double-talk. If the NLP is used to mask a poorly performing echo canceller there will be large residual echo during double-talk conditions. Normally, commercial echo cancellers specify the echo return loss (ERL) parameter for echo cancellation and NLP performance. This parameter does not describe the performance of the echo cancellation during double-talk.

## CLIPPING COMPENSATION

Car kit performance is judged on the quality of the voice signal for the end user. If a car kit is designed to meet worst-case conditions in a car travelling with the windows closed, the design will fail when the car is moving with the windows open. When the windows are open, wind noise increases the signal level significantly, the signal from the microphone gets clipped and the analog-to-digital converter is overloaded.

Figure 5 shows a signal that is severely clipped. The voice and background noise signals appear to be identical and the circuitry can barely distinguish the voice from the background noise. Under these conditions, the AEC would not be able to converge to the echo and the noise reduction algorithm would not be able to easily distinguish background noise and reduce it.

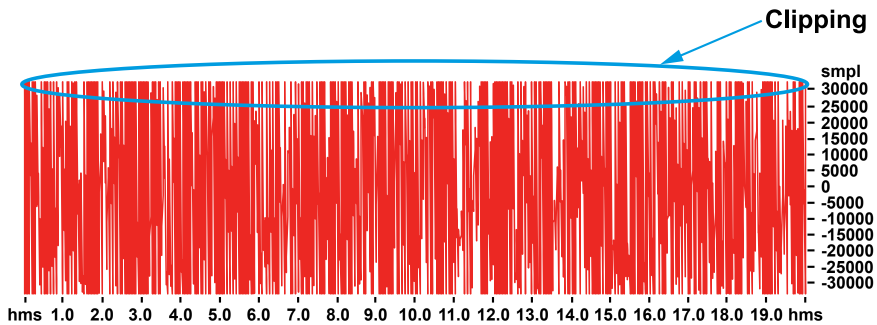


*Figure 5: Severely clipped input signal*

Figure 6 shows the same signal that has been processed with a noise reduction algorithm but without clipping compensation. The signal is still clipped and the noise level is still fairly high.

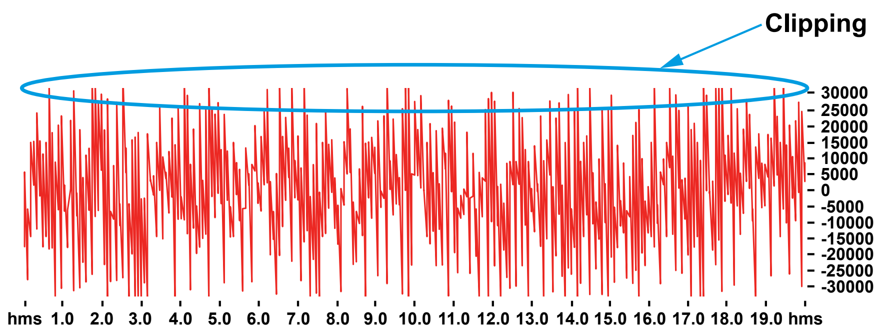


*Figure 6 – Severely clipped passed through a noise reduction algorithm*

Figure 7 shows same signal that has been processed with a noise reduction algorithm with clipping compensation. Under these conditions the AEC performs very well because signal clipping is eliminated.

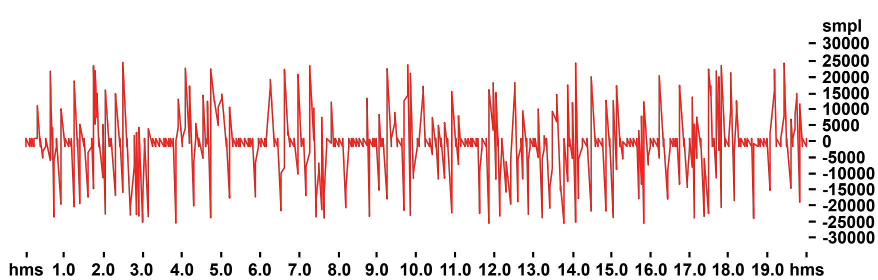


*Figure 7 – Noise reduction with clipping compensation*

## CONCLUSION

Governments around the world are enforcing the use of hands-free car kits as a means of communication inside moving vehicles for safety reasons. A high-performance car kit is judged on the sound quality that it delivers. With growing demand for hands-free communication systems supporting high-performance, vendors are concentrating on developing solutions that deliver superior voice quality.

The key elements of designing a high performance car kit are noise reduction, acoustic echo cancellation, double-talk performance and clipping compensation.

To achieve good noise reduction, the car kit must be capable of dynamically modifying the microphone gain using a smart automatic gain control circuit on the input of the system to eliminate signal clipping and preserve the settings of the end user.

Good double-talk performance relies on the AEC and the NLP designed to interact in such a way that the AEC tracks changes in the echo path and continuously converges during double-talk and the NLP conceals the imperfections of the AEC by reducing echo further and handling non-linear effects.

To compensate for signal clipping under severe noise conditions, for example a car driving at over 100 km/h with the windows open, a good clipping compensation circuit is required. Under these conditions, the microphone signal is clipped and the ADC is overloaded. The clipping compensation circuit prevents the signal from being clipped and allows the AEC to converge to echo and the noise reduction circuit to identify and reduce background noise.

Zarlink Semiconductor has developed a family of single-chip devices targeting high-performance hands-free communications. The company's newest hands-free solution, the ZL38004, is a dedicated voice processor that combines integrated dual-channel codecs with multiple interfaces and supports echo cancellation, noise reduction and clipping compensation. The chip's patented software algorithms continuously track changes in the echo path, even during double-talk conditions, and reduce background noise while preserving high voice quality. The ZL38004 platform's full-duplex operation, even when voice signals are low, enables natural two-way voice conversation.

This new platform builds on the company's successful hands-free voice processing chips, including the ZL38002 and ZL38003 products.  For more on Zarlink's hands-free communication technology, visit http://voiceprocessing.zarlink.com.

*Andre Coucopoulos is product marketing manager and Dean Morgan is system architect with Zarlink Semiconductor's Voice Processing product line. .*