WHITE PAPER

Hands-Free Communications—Making the Right Design Choices

Dean Morgan
Applications Engineer, Voice Processing
Zarlink Semiconductor

As published in Network Systems Design Line,
December 2006

6ZS378

Hands-free performance has become a standard function on most desktop and residential phones as well as car kits. The quality of hands-free communication depends on the technology used for speakerphone function, as well as plastic enclosure design choices.

This article will look at some of the technology and design challenges of hands-free communication systems, paying particular attention to how plastic enclosure design can significantly affect speakerphone performance.

THE BASICS: HALF-DUPLEX AND FULL-DUPLEX OPERATION

The voice quality and performance of hands-free communication tools can vary widely, depending on the technology used for the speakerphone function. The simplest and still most common approach used is the half-duplex algorithm. This approach allows only one party to talk, cutting off the other party in the call. Only the party with the highest volume is heard.

This half-duplex approach causes the conversation to be broken as the algorithm determines the difference in volume and changes to the active party. Although the performance is poor, half-duplex algorithms tolerate low-quality plastic enclosure designs. Because speech is in only one direction at a given time the algorithm hides plastic enclosure deficiencies, such as distortion, excessive echo return and plastic vibration.

A key performance goal for speakerphone manufacturers is to make voice communication more natural. The only way to achieve this quality is with a full-duplex algorithm. A full-duplex algorithm allows both parties to talk at the same time and only cancels the returned echo. While the full-duplex algorithm provides a more natural conversation without cutting off any party, it is more susceptible to impairment in the echoed audio path. The performance of a full-duplex algorithm depends on the linearity of the echo signal. Non-linearities in the audio path degrade the voice performance.

The biggest challenge faced by most speakerphone manufactures is converting their current half-duplex design to superior full-duplex operation. Manufacturers frequently don't consider the effects of the plastic enclosure design in the overall performance of the speakerphone. Echo return loss and non-linearity (or distortion) are two areas where plastic enclosure design can significantly affect the acoustic echo cancellation performance.

ECHO RETURN

Echo return is the amount of reflected audio from the speaker that is picked up by the microphone. The reflected signal is composed of the direct acoustic coupling (the amount of audio the microphone picks up directly from the speaker) and the audio reflections from the surfaces in the room. The goal of the speakerphone designer is to minimize the echo return. This is done in part through the plastic enclosure design. Even with a well designed plastic enclosure it is impossible to entirely eliminate echo, hence the need for an acoustic echo canceller (AEC). To achieve the highest performance, an AEC tries to minimize this returned signal.

Gain in the Echo Path

Figure 1 shows the different gain locations in relation to the AEC. The speaker and microphone gain pads affect the echo path. These gains are necessary but should be set appropriately to achieve high performance echo cancellation. For optimal performance, an AEC should be selected based on its ability to cancel the highest level of acoustic echo return without freezing adaptation (should continuously track changes in the echo path) or using switched attenuation (should not fall back to half-duplex).

Many commercial devices available today can cancel up to 0dB or 6dB of echo return (from Rout to Sin as outlined in Figure 1) without impacting performance. Designers need to ensure that the echoed signal level (at Sin) does not exceed the speaker audio signal (at Rin) by the echo return of the device (0dB or 6dB). Speaker driver gain, microphone gain and acoustic coupling comprise echo return. To achieve optimal hands-free performance, a balance must be met between the loudness of the speaker and the sensitivity of the microphone while maintaining a maximum 0dB to 6dB acoustic echo path.

Both the speaker gain and microphone gain are needed to deliver the speaker volume and ensure the microphone signal is sufficient to perform acoustic echo cancellation. These gains should be set to a minimum for the system and all other gains should be done outside the echo path. The maximum volume should be set with the maximum signal from Rout (not average) and the microphone gain should be set using the maximum input signal (not average). If these settings do not provide sufficient speaker volume or microphone sensitivity then the gain in the echo path can be offset by the gain on the network side. For example, if the audio echo return needs to be reduced by 6dB to achieve the requirements of the AEC then the microphone gain can be reduced by 6dB and the network Tx gain can be increased by 6dB (giving an overall gain of 0dB).

It is important to note that speaker driver gain and microphone gain in the echo path should never be changed dynamically during a call. This will cause the AEC to see an abrupt echo path change and trigger a re-convergence, creating an echo for the far-end caller. Any dynamic speaker volume control must be done at the network Rx gain block.

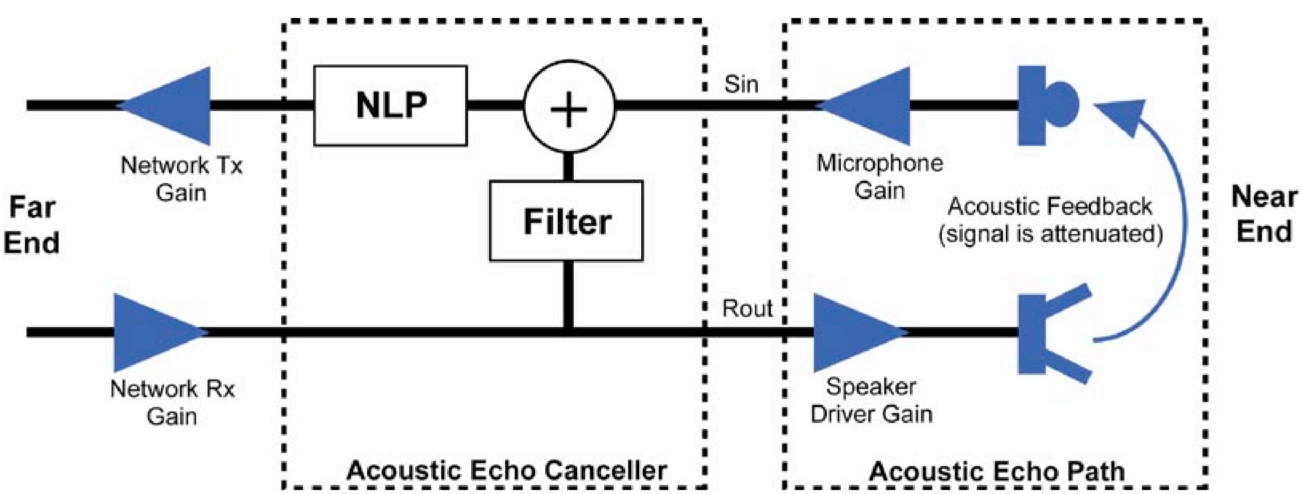

Figure 1: Components of ERL

Speaker Microphone Placement

Performance can be further improved by maximizing the acoustic separation between the speaker and microphone. Figure 2 illustrates the suggested speaker/microphone placement for a desktop phone. The speaker and microphone should be placed at opposite sides of the plastic enclosure. In addition, the microphone should be oriented to 90 degrees with respect to the speaker. This will minimize the amount of acoustic coupling external to the plastic enclosure. The speaker should be enclosed in a separate acoustic cavity to reduce the effect of acoustic coupling within the plastic enclosure. This is achieved either by using sound baffles or having a solid barrier between the speaker cavity and the rest of the plastics.

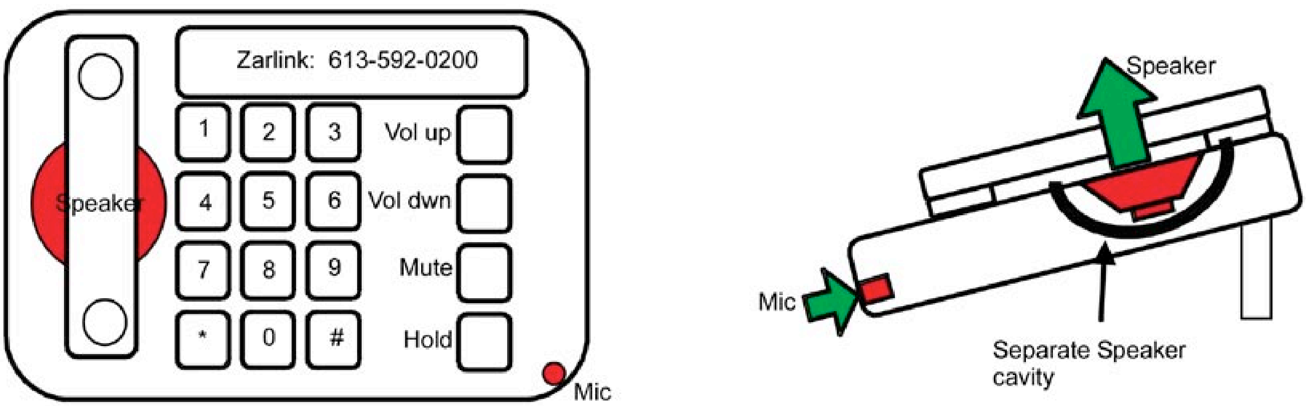

Figure 2: Microphone and speaker placement

DISTORTION

Distortion in the acoustic echo path is a major contributor to poor acoustic echo cancellation. Once the echoed audio signal is corrupted, the AEC has no way of modeling the echo path and canceling echo. There are many contributors to distortion, including microphone and speaker distortion, speaker and microphone amplification design, and plastic design.

Speaker and Microphone Selection

Speaker selection is critical to achieving high performance hands-free operation. The maximum speaker rating is usually given at a distortion rating of 10% total harmonic distortion (THD). For good quality AEC performance the THD should be less than 2%. Generally, 2% THD occurs at approximately half the maximum power rating. If the system design requires a 1W speaker driver then a 2W speaker or higher should be chosen.

This is not as significant an issue for electret microphones used in most applications as they are inherently linear across the operational range. Omnidirectional microphones, since they tend to be more linear, should be used over unidirectional microphones.

Speaker Driver and Microphone Amplifier Design

Similar to speaker and microphone selection, care must be taken when designing the speaker driver and microphone amplifier. Saturation occurs when either the microphone or the microphone amplifier circuit are saturated and is a significant concern in high-noise environments, such as hands-free car kits and industrial applications.

The most common cause of microphone saturation in a car kit is wind noise. Placing foam covering over the microphone to prevent direct airflow can reduce this significantly. If the sound pressure in the room is too loud for the microphone (which is unlikely), then select a microphone that has a lower sound pressure to volts rating.

The microphone amplifier circuit is easier to control. If the microphone is not saturated and the input to the analog-to-digital converter (ADC) is overloaded, then the amplifier gain must be reduced. It is important to note that by reducing the gain at the microphone in noisy environments the dynamic range will be reduced in quieter environments. A general rule is the greater the dynamic range the better the audio quality. Since audio quality becomes noticeable in quieter environments, the degradation of audio is more also noticeable.

A standard automatic gain control (AGC) on the microphone amplifier circuit can't be used to compensate for different signal levels. Since it is in the echo path, every time the AGC changes the gain there is an abrupt echo path change and a noticeable re-convergence time. Where possible, use a device that has an advanced AGC that changes the analog gain and updates the AEC algorithm to allow automatic maximum dynamic range over a wide range of noise environments.

Clipping in the speaker driver can also severely affect the AEC performance. It is very noticeable and can be adjusted by the user accordingly. What is less obvious is the non-linearity as the driver approaches clipping. The designer should take care to ensure the speaker driver can drive the required power to the speaker with less than 2% THD. This ensures that both the speaker and driver are operating in the linear range. If the speaker size and power is fixed and more volume is needed then an AEC with non-linear echo control can be used to compensate for the non-linearities in the speaker.

NON-LINEAR ECHO CONTROL

The plastics used in speaker design are often chosen based on esthetic and real estate considerations, not electrical characteristics. As a result it is sometimes difficult to achieve the desired volume and maintain appropriate level on THD (below 2%) for good AEC performance. In this situation an AEC with non-linear echo control will help. The non-linear echo control examines the residual echo from the linear echo canceller, compares it to the reference signal in the frequency domain, and then it subtracts the estimated frequency components from the echo signal. Although the performance is greatly improved during double-talk the overall performance is still dependant on the THD (the lower the THD the better the performance).

Plastic Design

Plastic design is important in heightening or minimizing distortion in any acoustic system. The goal is to minimize plastic vibration and the amount of vibration noise the microphone picks up from other sources, such as mechanical switches.

Microphone Placement

The microphone should be housed in a rubber grommet with no air gaps to minimize vibration coupling. The microphone should not be in contact with either the circuit board (connect to the circuit board using wires) or the plastic enclosure. It should not be placed at the seam of the plastic enclosure.

Care should also be taken to ensure that any moving parts (handset, keypad, switch hook) don't create extra distortion that could be picked up by the microphone. This can be achieved using soft rubber keypads or soft keys on a touch screen.

Speaker Placement

Not only do designers need to take care in speaker placement, they also need to ensure that the speaker is mounted appropriately to the plastic enclosure. It is important to minimize the amount of vibration in the plastic enclosure. The speaker should be mounted to the plastic enclosure using vibration-dampening materials, such as soft silicone or foam rubber. This should be placed between both the plastic enclosure and speaker as well as the speaker mounting bracket and the speaker (see figure 3).

For many phone plastics designed for half-duplex algorithms, placing a low-cost rubber gasket between the speaker and plastic enclosure and mounting bracket will greatly enhance the acoustic properties of the plastic.

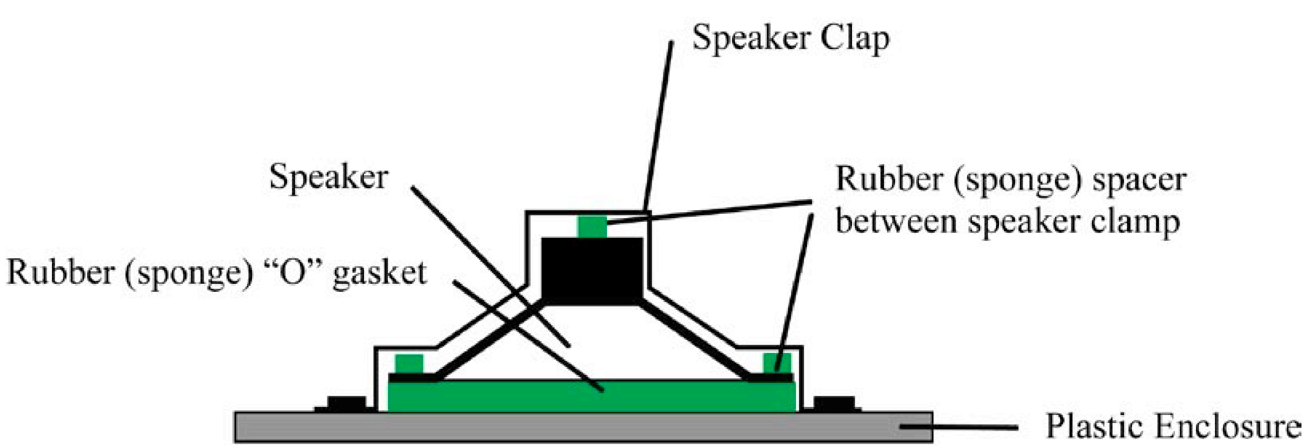

Figure 3: Suggested speaker mounting

EQUALIZATION

In cases where existing, poorly designed plastic enclosures must be used it is possible to compensate for errors by using an equalizer in the speaker path. For example, Figure 4 illustrates the total harmonic distortion (THD) verses frequency of a desktop phone. The results show that there is excessive distortion at lower frequencies, due to the poor frequency response of the speaker at lower frequencies. There is also a resonance frequency at approximately 420Hz, 820Hz and 1400Hz. These distortion points will degrade the overall performance of the echo canceller.

To counter this, an equalizer can be used to pre-condition the signal and minimize the effects of this resonance before it reaches the speaker. An equalizer with appropriately sized and located frequency bins can be used to attenuate those frequencies and reduce the resonance peaks. This may add a small amount of distortion to the speaker (which is modeled by the AEC). However, this distortion is unusually not noticeable and is a reasonable trade-off between the cost of re-spinning the plastic enclosure design and overall AEC performance.

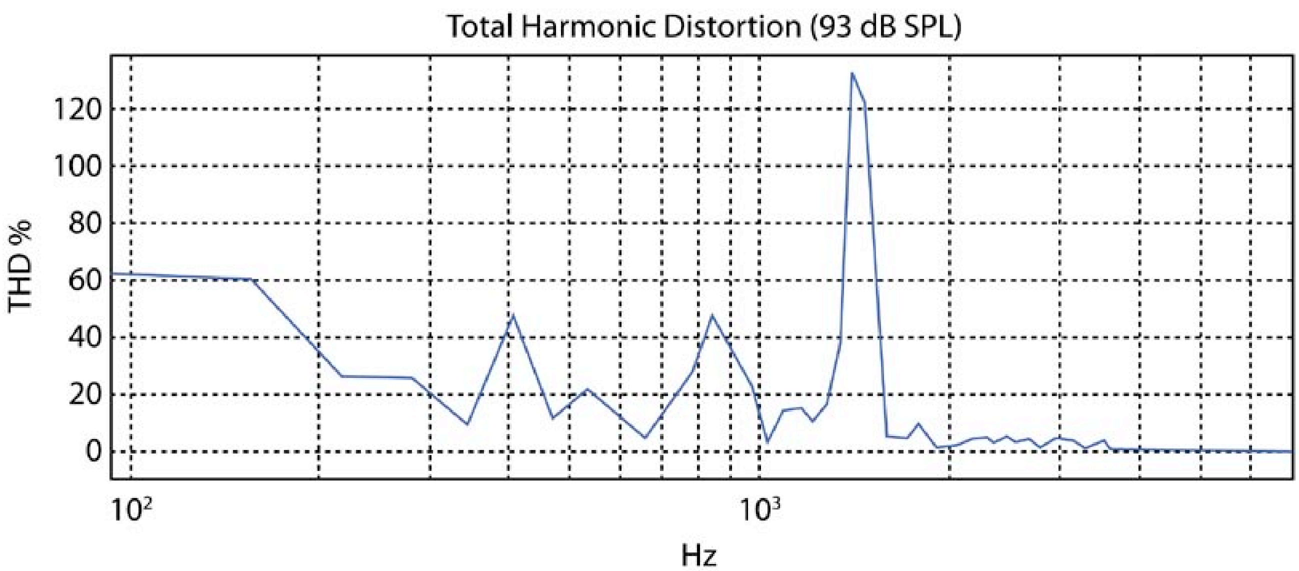

Figure 4: Total harmonic distortion versus frequency

AEC HINTS ON TUNING

When tuning an AEC there will always be some compromise between speaker volume and microphone sensitivity. The better the plastic enclosure and system design are, the more flexibility there will be in terms of trading-off speaker volume for microphone sensitivity (and vice-versa). To initially tune the AEC, the system designer must choose what is more important to the end user: speaker volume or microphone sensitivity. Once this is established the following steps can be used to tune the AEC.

1) Set the microphone sensitivity, ensuring the settings are adjusted in the intended environment. The microphone gain should be set to achieve the maximum dynamic range without clipping the incoming signal.

2) Set the user gain to 0dB and increase the analog speaker driver gain slowly until the AEC fails to cancel the echo.

3) If the speaker volume is too low, add 6dB of attenuation to the microphone amplifier and compensate for this attenuation with 6dB gain on the Sout path. Remember that the NLP threshold will have to be adjusted to compensate for these gain settings.

4) Increase the speaker gain by 6dB. If the speaker volume is still not loud enough then repeat step 3 with a 9dB attenuation/gain value. This method works for attenuation/gain settings of less than 9dB.

5) Once the desired speaker volume is reached then the microphone sensitivity can be tuned by adjusting the Sout gain.

6) If the AEC is working properly for most signals except the peaks, then either the speaker or speaker driver is causing some distortion. This can be addressed by limiting the Rout level. For a 16-bit codec the maximum PCM value is 7FFFh. By reducing this to 3FFFh the output to the speaker will not have peaks and the amount of distortion due to the speaker and speaker driver is reduced.

CONCLUSION

Although voice algorithms determine good acoustic echo cancellation, echo return and distortion are two significant factors that influence the final performance of the AEC. Both of these factors must be minimized before any AEC can work properly.

Plastic design plays a large role in the performance of the AEC. Plastics can contribute to echo return by producing high levels of acoustic coupling. Careful placement of the speaker and microphone and the addition of sound baffles or a separate speaker cavity can reduce the amount of acoustic coupling.

Vibrations through the plastics can also contribute to distortion. Soft mounting both the speaker and microphones can minimize vibrations. It is also possible to improve plastics that were designed for half-duplex algorithms. Designers can work with a full-duplex algorithm but should add vibration-absorbing material between the speaker/microphone and the plastic enclosure.

Another cause of reduced AEC performance is poor speaker and speaker driver selection. The speaker/speaker driver must have less than 2% THD to achieve good AEC performance.

Dean Morgan is a system architect with Zarlink Semiconductor's Voice Processing group. Dean holds a bachelor of engineering from Lakehead University. He can be reached at dean.morgan@zarlink.com.