How Digital Signal Processing Can Enhance the Utility and Performance of Microphones

Smartphones and the Internet of Things have made microphones much more vital in today's technology. In this paper, the author describes a technology for digital processing of microphone signals that requires no coding skills or DSP expertise.

By Dr. Paul Beckmann, DSP Concepts, Inc.

Use of microphones has exploded with the popularity of smartphones, and is increasing even more with the advent of the Internet of Things (IoT). Sales of component microphones (the microphone capsules that go into products, not complete microphones of the type found in recording studios) are now growing at a rate of 17% per year.

The applications in which microphones are finding new utility go far beyond traditional uses such as voice reception. Mobile applications demand that microphones be able to isolate desired sounds in noisy environments. Security and IoT applications use microphones as sensors, where the signals they produce are processed to detect abnormal conditions such as glass breaking or a patient's breath stopping. Most microphones must be extremely tiny to fit the form factors of mobile products, yet are still expected to deliver high performance.

This paper explains how engineers can use digital signal processing (DSP) to solve all of these problems and many more. It will also profile a DSP software package that allows engineers without coding skills and DSP expertise to optimize microphone performance in their applications.

Modern applications for microphones

- *Phones/Tablets:* Half of U.S. customers have no landline, making sound quality of smartphones increasingly important. Many phones and tablets have multiple mics to accompany their built-in cameras.
- *Cameras:* Almost all digital cameras now shoot video and thus need a microphone, preferably with resistance to wind noise. Action cams (i.e., GoPro) and webcams are also increasing in use.
- *Wearables:* Health sensors often use microphones to detect body functions, such as breath and sleep. Almost all smartwatches incorporate a microphone.
- *Headphones:* Most mass-market headphone models include a built-in microphone for taking phone calls. Wind and noise rejection is useful in this application. Many headphones also use microphones for noise cancelling.
- *Bluetooth speakers:* A common feature in Bluetooth speakers is a speakerphone function, which requires one or more microphones.
- *Automotive:* All but the least expensive cars now include voice command and Bluetooth for phone calls, and thus need microphones. Road noise rejection is critical here. Some luxury cars use microphones for active road noise cancelling through the audio system.

- *Security:* IP-based security cameras have become inexpensive and ubiquitous, and most include microphones. In addition, microphones are increasingly used to detect dangers such as gunshots and glass breaking.

Many of these applications require multiple microphones, which can enhance sound reception, help to cancel noise and wind, and can be used to create directional arrays with steerable pickup patterns.

General challenges in incorporating microphones

Engineering microphones into products is typically complicated by problems inherent to microphones and to many of their applications.

One problem inherent to most microphones is poor signal-to-noise ratios (SNRs). We usually expect SNRs of 90 dB or better with audio electronics, but a typical microphone might have an SNR in the range of 60 dB.

Compounding this problem is that microphones often must cover a large dynamic range. A mic might be expected to pick up sounds as quiet as 30 dB SPL, then a fraction of a second later handle a human voice yelling at 110 dB SPL at a 2cm distance, or a drum playing as loud as 130 dB at close distance. This high dynamic range can make it difficult to achieve a gain level that accommodates the loudest peaks but does not produce excessive noise. The high-gain preamp circuitry required for most microphones can also create excessive noise.

Placement of the microphone relative to the sound source is unpredictable in applications such as speakerphones or digital cameras, where the source might be as close as 1cm or as distant as 100 meters. The environment is sometimes unpredictable; it may be a quiet room at 40 dB SPL or a crowded club at 100 dB SPL. Often, the same brand and model of mic capsule may be specified across numerous applications and thus must accommodate all of these conditions and more.

Most of the applications for microphones are in portable devices, which present additional challenges. Portable products tend to be compact, and their form factor often limits designers' options in mic placement, thus compromising performance. Most of these products are battery-powered, with only single-digit voltage supplies on hand and often a run time of just a few hours.

Specific challenges in incorporating microphones

In addition to the general challenges noted above, many specific microphone designs and applications also present challenges to audio engineers. These include:

- *DC offset:* Many microphones, especially MEMS and condenser types, produce a steady-state DC voltage at their output, which must be removed because it can damage downstream components.
- *Non-flat frequency response:* Many microphones show considerable roll-off or boost at high and/or low frequencies, and mounting microphones in enclosures tends to

produce large non-linearities, many of which are in the midrange and thus readily audible.

- *Output mismatch with multiple mics:* Because of manufacturing variances, microphones can exhibit significant differences in output voltage from sample to sample. With single microphones, this can cause performance variances in the finished product. Output mismatch also reduces the efficacy of technologies such as noise/wind cancelling and directional sound reception that rely on multiple microphones.
- *Environmental noise:* Microphones often must operate in noisy environments, and in many cases are required to reject environmental noise with minimal effect on the desired sounds.
- *Essing and popping:* Nearfield voice applications require the ability to deal with the high air velocity of sibilant and plosive sounds (the letters B, D, G, K, P and T), which can produce distortion in downstream circuitry.
- *Wind noise:* In most mobile applications, air often rushes across the microphone, due either to wind or motion of the microphone. This produces mid- and high-frequency noise as well as low-frequency rumble that can obscure desired sounds.
- *Interference from nearby speakers:* Microphones often have to operate in proximity to or in conjunction with speaker systems, whose sound can leak into a microphone and create interference or feedback.

Solutions to problems encountered with microphones

Each of the problems cited above can be solved through signal processing. In most cases, each requires its own specific solution.

DC offsets

DC offsets can be blocked using a high-pass filter set to a frequency low enough not to affect the audio range. Typically the filter will be 2nd order (12 dB/octave) or steeper, usually with a -3 dB point around 10 Hz. The filter should be tuned to suit the application. Higher frequencies and gentler slopes can be used if deep bass response is not a concern. To reject microphone handling noise and low-frequency rumble at the expense of some bass response, the filter can also be set to a frequency in the audio band, typically 80 Hz.
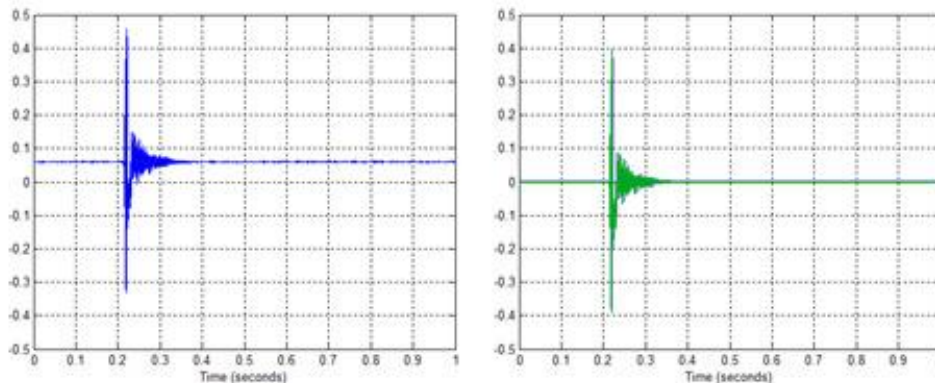
Typical DC offset of a MEMS microphone before correction (left) and after high-pass filtering (right)

Non-flat frequency response

Frequency response nonlinearities inherent to a microphone or its enclosure can be removed through the use of filters. Low-pass and/or high-pass filters can correct for rising or falling bass or treble response. Bandpass and band-reject filters can correct response peaks and dips, which are often caused by sound reflections and/or resonances in the enclosure.
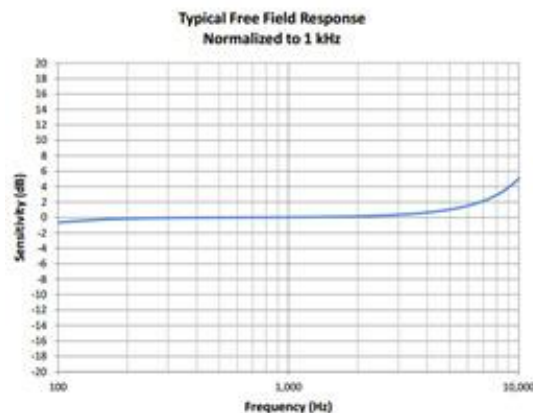
Frequency response of Knowles SPM0437HD4H-B MEMS microphone

Filters also serve many other purposes. For example, to enhance voice reception and intelligibility, a midrange boost filter (typically a bandpass filter centered at about 2 kHz) is often used. Filters can also be tuned to optimize the subjective sound quality of a device, and can also create selectable sound modes, such as a reduction in bass or treble response to suit the characteristics of specific singers. If the reproduction chain is known, filters can be tuned to optimize microphone response for the capabilities of the playback device, such as reducing the bass response when a small speaker will be used.

Output level mismatch

Mismatch of microphone output levels due to manufacturing inconsistency can be solved in two ways using signal processing: through calibration or using automatic gain control.

Factory calibration is best suited to finished products that incorporate microphones. It involves playing a test tone, measuring each microphone's output, calculating the proper gain for that microphone, then loading that gain setting into the firmware of the product in which that microphone is installed. Field calibration is appropriate for sound reinforcement and home theater applications where several microphones, often purchased separately, are used. The microphones are placed together, a test tone is played, and the output of each microphone is processed so they all match.

Automatic gain control monitors the signal level from each microphone and adjusts the gain so all of the microphones produce the desired average or peak level. It's especially appropriate for applications where the environment and/or the character of the sounds to be received are unknown, such as in multichannel sound recording.

Dynamic range management

As stated previously, there are many situations in which it is difficult for microphones and their accompanying electronics to capture the full dynamic range of sounds. With gain set too high, an unexpectedly loud sound can cause distortion. With gain set too low, an analog-to-digital converter (ADC) connected to the microphone cannot use its maximum number of bits and will thus produce more noise than is necessary.
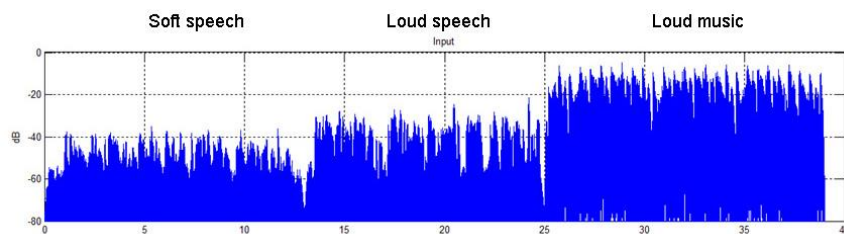

Typical dynamic range encountered by microphones in voice and music applications

One solution for dynamic range management is the automatic gain control discussed in the previous section. The level of the microphone(s) is monitored and automatically adjusted so that levels are high enough to minimize noise, but never exceed 0 dBFS and thus do not cause clipping.

A related solution is a compressor/limiter, which uses fixed signal processing to optimize sound and recording quality. In a limiter, signals exceeding a predetermined threshold are either reduced in level so they do not exceed the threshold, or the rate of their rise once past the threshold is reduced by a certain ratio, which is usually adjustable. A compressor combines a limiter with a function that raises the level of weak

signals; this combination results in a signal with reduced dynamic range that runs much less risk of overloading an ADC.
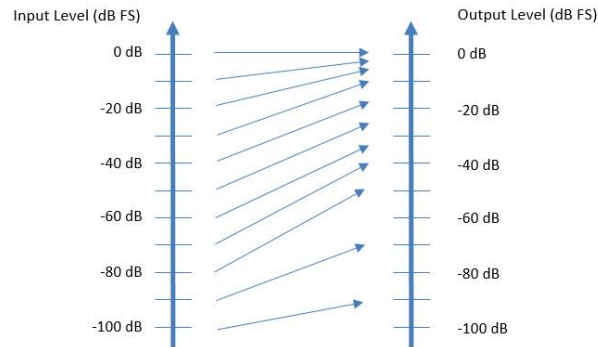

Operating principle of dynamic range compressor

Compressor/limiters can also be used to optimize quality of sounds picked up by a microphone, producing a higher average signal level that, for example, helps voices be heard and understood in noisy environments. Compressing a microphone signal can also help reduce stress on and maximize output of downstream components, such as a small speaker that does not have its own compressor/limiter processing.

Any compressor/limiter can be made frequency-dependent by adding a filter. For example, if loud bass noises need to be controlled but midrange and treble sounds do not, a low-pass filter can be used before the compressor/limiter, and the output of the compressor/limiter recombined with a high-pass filtered, uncompressed signal.

Noise/hiss

Noise and hiss are common problems with microphones because of their low SNRs and because they tend to pick up environmental noise. This noise can be distracting or annoying to a listener. To lessen or eliminate this problem, a noise gate can be used. A noise gate mutes an audio signal when it falls below a certain level, typically the level at which no important or audible sounds are present in the signal. The threshold of the noise gate—the signal level below which the signal is muted—must be adjusted so it doesn't cut out unwanted sounds or allow intrusion of hiss or excessive noise.
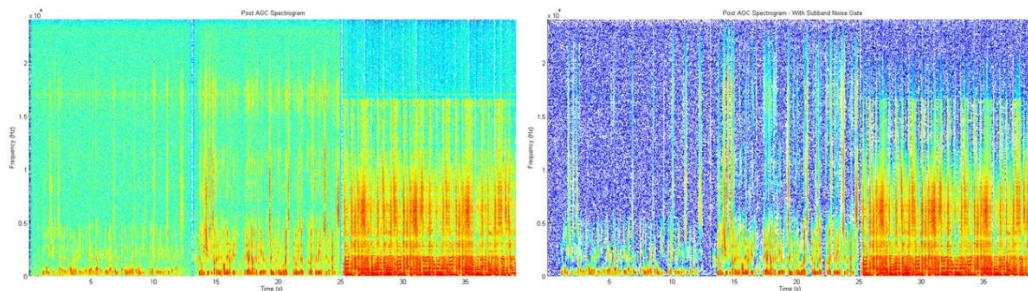
Microphone signal before (left) and after (right) noise gating

As with compressor/limiters, a noise gate can be made frequency-dependent by preceding it in the signal chain with a filter. If, for example, low-frequency rumble from machinery would be sufficient to keep the noise gate open at inappropriate times, the low-frequency component can be filtered out of the noise gate's level-detection input, so the gate's operation is affected only by higher-frequency sounds.

Environmental and wind noise
Many different microphone signal processing techniques allow desired sounds to be heard easily over loud environmental or wind noise. These techniques include everything from simple filtering to noise cancelling to advanced multi-microphone arrays.

Filtering can be tuned to reject the type of noise that is expected. For example, a microphone intended for outdoor applications (i.e., action cams) might be high-pass filtered to reduce the low-frequency components of wind and road noise. If the range of desired sound is fairly narrow—such as the human voice, which ranges mostly from 100 Hz to 3 kHz, or a gunshot with most of its sonic energy centered at 1 kHz—a bandpass filter can be used to emphasize that range.

Using multiple microphones can cancel noise more effectively. When two microphones—a primary mic positioned near the subject with a second mic further away—are used on a speakerphone, any sound that appears in equal volume at both microphones can be attenuated, thus increasing the relative level of the desired sound.

If the position of the subject relative to the microphone is known, an array of multiple microphones can produce a directional pickup pattern that helps isolate the subject from the environment. Sounds arrive at the various microphones in the array at slightly different times. By using all-pass filters to shift the phase on certain microphones in the array, sounds arriving from a specific direction are amplified while sounds arriving from other directions are attenuated.

Essing and popping
Essing and popping in vocal applications are typically minimized through the use of two methods: a compressor/limiter that reduces the loud peaks produced by plosive and sibilant sounds, and a filter or filters that reduce output at troublesome frequencies. A high-pass filter at roughly 100 Hz can reduce the bass component of the "pops" from plosive sounds without significantly affecting voice quality. A band-reject filter in the 2 to 3 kHz range can be used to reduce sibilance. Both of these filters must be adjusted to maximize their effect without substantially reducing sound quality.

Interference from speakers
If microphones must be used in the presence of loudspeakers, the pickup of sound coming from the speakers can be substantially reduced if the signal feeding the loudspeakers is known (such as in a P.A. system or a Bluetooth speaker). The signal feeding the speakers is phase-reversed, then combined at an appropriate level with the

signal from the microphone to create an acoustic echo cancelling function. Sounds from the speakers are then attenuated or canceled, so they do not interfere with the desired sounds being picked up by the microphone.
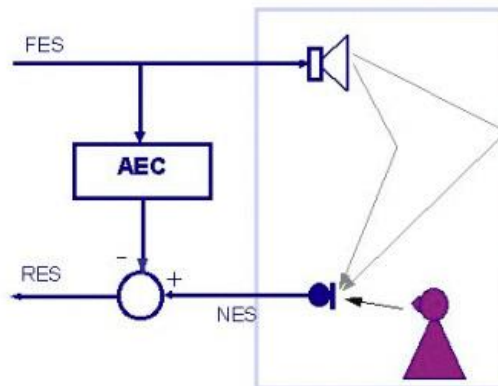
Block diagram of an acoustic echo cancelling circuit

Difficulties in implementing microphone processing

For optimum performance, most applications using microphones require some combination of the above techniques. For audio engineers, such a combination can be challenging to implement.

Creating multiple mic signal processing stages using analog circuitry is usually impractical, because multiple stages of analog processing tend to create excessive noise; development and tuning of the circuits is too time-consuming; and the resulting product might be too expensive.

Digital signal processing (DSP) can solve these problems, but many of the functions described above require sophisticated algorithms not available in the stock programming interfaces for DSP chips, and require considerable DSP coding skill to implement. Furthermore, the sophisticated algorithms required to execute these processing techniques demand powerful DSP chips, preferably with 32-bit processing capability.

The DSP chips running these algorithms must in most cases be able to run on low voltage from batteries or an energy-efficient power supply. Some applications require the DSP to run at all times—for example, security and health monitoring equipment, and voice-command systems used in cars and consumer electronics such as TVs and the Amazon Echo smart Bluetooth speaker. In these applications, the core audio processing components of a product must always be on so they can detect trigger words from the user that will bring the product to full power and allow it execute commands. Because of energy efficiency concerns, it is desirable to keep the power consumption at idle to very low levels, often just a few milliwatts.

Designing and tuning products that use microphones requires quite a different set of skills than those used for DSP programming. Only a few DSP engineers possess the engineering knowledge and experience to design audio products and optimize their performance. Yet few audio engineers possess DSP coding skills.

A solution to the challenges of DSP for microphones

Audio Weaver modular DSP software was born out of years of product development work for many customers developing and tuning audio products. Audio Weaver allows complex, sophisticated audio processing algorithms to be created and optimized in a small fraction of the time it would take to write them into DSP code, saving companies hundreds of thousands in engineering cost and getting products to market faster while reducing development risk. The modular nature of the software allows audio engineers to build signal processing chains through block diagrams shown on a graphical interface. Thus, an audio engineer can assemble multiple DSP functions by arranging and connecting processing components onscreen, much as they might assemble a P.A. or home theater system.
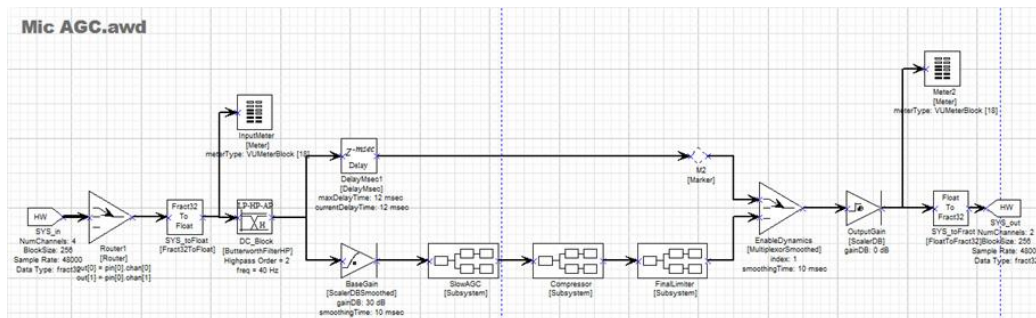
Automatic gain control implemented in Audio Weaver's graphical interface

The software's agile development environment allows practically any combination or configuration of the more than 400 signal processing modules now available. The modules use familiar, analog-style graphical controls, so an audio engineer can adjust, say, a compressor/limiter module the same way she or he would adjust a traditional analog compressor/limiter in a recording studio. No coding skills or knowledge are required.

Audio Weaver is now available for most major DSP families and many embedded cores, including:

- Tensilica HiFi
- ARM Cortex M4 and M7
- Texas Instruments C67x and C66x
- Analog Devices SHARC
- ARM Cortex-A

- Intel x86/x64

For audio companies without the manufacturing capability to incorporate DSP, Audio Weaver capability is also available in finished DSP processing boards from ST Microelectronics and Danville Signal. These boards incorporate digital and analog input/output plus a USB programming interface.

Conclusion

The rapid growth in the use of microphones in consumer and industrial products presents great opportunities, but also difficult challenges. Audio Weaver can meet any of these challenges through its flexible, agile development environment. The software can be operated by any audio engineer, using a graphical interface for rapid configuration and adjustment. Audio Weaver is easy to incorporate into new or existing audio product designs, regardless of a company's manufacturing resources and capabilities.