

Debbie Greenstreet



In a twist of irony, the massive technological expansion of the telephone system – which led to the creation of the internet and all of its features and benefits – enables millions of simultaneous transactions without requiring a single human conversation. Yet in this efficiency-focused 21st century world, the plethora of voice-activated smart home devices confirms that live speech has become the de facto medium for smart home virtual assistants. Consumers can still pick up a remote and press the volume button; they can use their phones to place an online order and their hands to flip a light switch. But it seems clear that voice as a user interface is here to stay, and real-time signal processing is the key enabler of this modality.

Real-time signal processing consists of converting and processing analog signals by performing complex math operations. Digital signal processors (DSPs) are the most efficient way to process math in real time. While all processors can perform real-time signal processing math calculations, the DSP architecture by design ensures that these processes will happen faster, with less energy and less generated heat than more generic processor architectures.

Voice as the user interface – a new era in speech processing

 [Read the white paper here](#)

Voice capture and speech recognition devices and applications are not new. However, properly recognizing and capturing speech amid television noise or conversations requires far more processing than simply capturing speech from a single microphone in a device that's close by and in a relatively quiet environment.

Key design factors in a smart speaker-like solution include accurate voice discernment (given the user's distance from the speaker), the amount of ambient noise in the area and the need for two-way speaker conversation. For near-field processing, a relatively simple system of three or fewer microphones, wake word detection, fixed beamforming and signal-noise reduction may be all that's required. It is possible to execute such a configuration easily on a microcontroller (MCU) within a low-power, memory and cost footprint.

However, depending on the application, the consumer may expect to command the smart speaker from both the near-field and far-field, and to have the speaker accurately discern their speech over noises from sources such as TVs, smartphones, background conversations, wind and other ambient sounds. To be effective, this complex environment typically requires between four and eight microphones that in turn require adaptive beamforming

and adaptive spectral noise reduction (ASNR) algorithms, along with multisourcing selection functionality. This significantly increases the real-time signal processing complexity.

Applications such as video doorbells expand the processing complexity one step further, requiring process-intensive acoustic echo cancellation (AEC) to improve the user experience. AEC, along with beamforming and ASNR, tips the scales in terms of exceeding MCU efficiency, but DSPs can effectively process voice as a user interface engine.

DSPs continue to be the most efficient means of processing real-time audio commands, especially amid the ambient sounds and noises commonly present in our environment. Just as in a smart home interface where consumers prefer using their voice because it's the most efficient, DSPs are preferable for far-field, four-or-more microphone, or two-way-speaker smart home solutions where size, power, performance and cost are key metrics.

Additional resources:

- Check out this [Audio Preprocessing System Reference Design for Voice-Based Applications](#)
- Learn more about all of [TI's audio and media processors](#).

IMPORTANT NOTICE AND DISCLAIMER

TI PROVIDES TECHNICAL AND RELIABILITY DATA (INCLUDING DATA SHEETS), DESIGN RESOURCES (INCLUDING REFERENCE DESIGNS), APPLICATION OR OTHER DESIGN ADVICE, WEB TOOLS, SAFETY INFORMATION, AND OTHER RESOURCES "AS IS" AND WITH ALL FAULTS, AND DISCLAIMS ALL WARRANTIES, EXPRESS AND IMPLIED, INCLUDING WITHOUT LIMITATION ANY IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT OF THIRD PARTY INTELLECTUAL PROPERTY RIGHTS.

These resources are intended for skilled developers designing with TI products. You are solely responsible for (1) selecting the appropriate TI products for your application, (2) designing, validating and testing your application, and (3) ensuring your application meets applicable standards, and any other safety, security, regulatory or other requirements.

These resources are subject to change without notice. TI grants you permission to use these resources only for development of an application that uses the TI products described in the resource. Other reproduction and display of these resources is prohibited. No license is granted to any other TI intellectual property right or to any third party intellectual property right. TI disclaims responsibility for, and you will fully indemnify TI and its representatives against, any claims, damages, costs, losses, and liabilities arising out of your use of these resources.

TI's products are provided subject to [TI's Terms of Sale](#) or other applicable terms available either on [ti.com](https://www.ti.com) or provided in conjunction with such TI products. TI's provision of these resources does not expand or otherwise alter TI's applicable warranties or warranty disclaimers for TI products.

TI objects to and rejects any additional or different terms you may have proposed.

Mailing Address: Texas Instruments, Post Office Box 655303, Dallas, Texas 75265
Copyright © 2023, Texas Instruments Incorporated