



# Tensilica HiFi 5 DSP

Voice control everywhere

We've come a long way since the only way to interact with a device was to touch a button or type something into a keyboard. It started in the mid-2000s with always-on applications that required specialized digital signal processors (DSPs), like the Cadence® Tensilica® HiFi Mini, which offloads the host processor, enabling low-power battery-operated devices that always listen for keywords to wake them up.

Early voice control products suffered from major problems recognizing words spoken in foreign accents and by young children (because their speech processing is still developing and they tend to shout instructions). Complicating this is the challenge of background noise. How does the device isolate the voice, and which voice should be isolated? Companies are working on algorithms that improve these challenges.

These algorithms need to run on sophisticated, yet very low-power processors. They require excellent AI processing capabilities, as now more AI capabilities are being used to give the best results. Thankfully, even more AI capabilities are being designed into processors, such as the Tensilica HiFi 5 DSP. With better, more efficient hardware capabilities to execute the most advanced AI algorithms, these processors will open up a much more enhanced user experience.

## Portable Devices

There are two major challenges that will be overcome by the next generation of voice-enabled systems. First, they need to be trained to understand a wider range of voices. This training requires a lot more processing power for the AI required to interpret all the possible voice variations than is in current systems. This training is often done with the computer power available in the cloud and then transferred to portable devices that you use at home.



Second, they need to process the voice on the device—not across the cloud—for immediate response. The processing of the AI algorithms must be very power efficient, so it can be employed by battery-operated devices. Currently, even the most advanced algorithms only work on products that are plugged into the wall to get the processing power they need.

From mobile phones to appliances, our devices will be listening for voice commands. Forbes projects that, by 2020, nearly half of all internet searches will be voice driven.<sup>1</sup> Speech recognition now works three times faster than humans can type.<sup>2</sup> And it's only going to get better, while our devices learn to process much more complex requests, like for shopping and controlling entire households.

## Noisy Environments

The challenge of voice recognition is compounded where there is a lot of background noise or overlapping speech at various distances from the microphones in the devices. Voice-recognition systems have trouble recognizing human voices over blenders, road noise, dogs barking, and TVs playing. Shouting louder at the device is not the solution!



The software algorithms for noise cancellation and far-field voice recognition capabilities are rapidly improving, which is another way of saying they are incorporating more machine learning and artificial intelligence. To do this recognition at speed, this means they need many times more hardware processing power than is available in today's devices. The primary voice needs to be isolated via a deep neural network that can determine which sound is the most important one. This will require a lot of neural network training, and that training will have to be downloaded to each device, so it can be independent of a Wi-Fi interface to the cloud to get immediate responses.

## Home Voice Control

Just about anything you manually interact with in your home today will be voice controlled in the future. This includes your thermostat, security cameras, TV, appliances in the kitchen and laundry, door locks, doorbells, and lighting. Devices will be connected using wireless connections and will require low-power DSPs to perform this voice recognition and control because many of them will be battery driven. You'll need a home automation hub to provide a common interface to these devices from different manufacturers.



As they get smarter, home automation hubs will be able to think for you. You will no longer need to communicate with each device individually. Imagine that you are going on vacation—you tell your hub that you'll be out of town, and it programs every enabled device in your home for your absence. You don't have to worry that you forgot to turn your water sprinklers off while you were gone, or that air conditioning is cooling an empty house.

## Next-Generation Voice Processors

Typically, general-purpose CPUs are not ideal for special-function processing, like voice or audio, because they consume too much power. That's why specialized processors have been in development for years. Specialized DSPs can process audio, voice, pixels, communications, and much more with their customized datapaths for the exact data widths and processing requirements. This specialization is essential to keep the power requirements as low as possible, as many applications are battery operated.



Some of the things you should look for when evaluating a new processor—particularly DSPs that you can add to your system-on-chip (SoC) design—include:

- Efficient 16-bit instructions optimized for voice and audio codecs
- An extensive library of software for audio and voice processing
- Proven support for neural network-assisted speech recognition (ASR) processing
- Fully programmable in C so you can add special software
- Performance to spare, as new applications are rapidly evolving and often need additional processing power

Even in the world of audio and voice processors, there's no one-size-fits-all solution. Look for a family of compatible processors so you have the flexibility to pick just the features you need for your application. Cadence has already introduced several processors—and you can expect that we will continue enhancing them to meet the computational requirements in the years ahead.

1. "Optimizing for Voice Search Is More Important than Ever," by Amine Bentahar, Forbes, <https://www.forbes.com/sites/forbesagencycouncil/2017/11/27/optimizing-for-voice-search-is-more-important-than-ever/>
2. "Voice Recognition Software Finally Beats Humans at Typing, Study Finds," by Aarti Shahani, NPR, <https://www.npr.org/sections/alltechconsidered/2016/08/24/491156218/voice-recognition-software-finally-beats-humans-at-typing-study-finds>

**cadence**<sup>®</sup>

Cadence software, hardware and semiconductor IP enable electronic systems and semiconductor companies to create the innovative end products that are transforming the way people live, work, and play. The company's System Design Enablement strategy helps customers develop differentiated products—from chips to boards to systems. [www.cadence.com](http://www.cadence.com)

© 2018 Cadence Design Systems, Inc. All rights reserved worldwide. Cadence, the Cadence logo, and the other Cadence marks found at [www.cadence.com/go/trademarks](http://www.cadence.com/go/trademarks) are trademarks or registered trademarks of Cadence Design Systems, Inc. SystemC is a trademark of Accellera Systems Initiative Inc. All other trademarks are the property of their respective owners. 11471 10/18 SA/LL/PDF