

11 Myths About Noise Cancellation

[Electronic Design](#)

[John Yoon](#)

Fri, 2016-09-23 08:54



Voice-controlled technology isn't new, but as we take our devices more places and voice becomes prevalent in more devices, the need becomes greater for progress to be made in managing background noise and improving speech-recognition systems.

A new approach to old problems—with the help of deep neural networks—can make background noise a thing of the past. This approach is shattering long-standing myths.

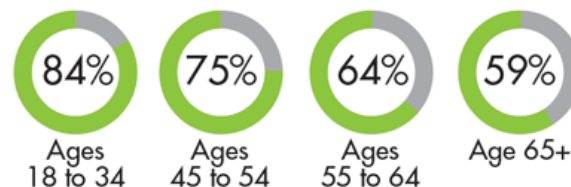
1. People don't really care as much about audio quality anymore.

A 2016 poll commissioned by [Cypher Corp.](#) (Fig. 1), and administered by Harris Poll, surveyed 1,875 U.S. adults and found that nearly three quarters (74%) of mobile-phone owners would be interested (very/somewhat) in a new offering that allows them to control background noise if their call recipient can hear such noise.



Nearly three quarters of mobile-phone owners would be interested in a mobile-phone function/feature that allows control of whether a call recipient can hear the background noise.

Interest in background noise control functionality by age group



2. Young people are moving away from voice.

Nearly half of all phone users today use their mobile phones as their primary voice connection, a number sure to grow. Mobile phones, by design, are used in many different environments: planes, trains, and automobiles; at sporting events, offices, factories, and shopping centers; on playgrounds; and (yeah, that guy) in public restrooms. Consider the noises encountered walking through an urban setting, near a construction site, or in an airport lounge.

Related

[Challenges Await In Noise-Cancelling Headset Design](#)

[Advancing the Audio Interface for Mobile and Mobile-Influenced Designs](#)

[A: Managing Background Noise for Speech-Recognition Systems](#)

Of the 74% of mobile-phone owners interested in an offering that allows them to control background noise, the 18-34 age group (84%) is more receptive than those age 45-54 (75%), 55-64 (64%), and 65+ (59%). Also, 90% of students would be interested in this technology.

3. I have a latest-generation smartphone, so I have the best noise reduction out there.

Current smartphones are a marvel. They enable you to do so many things: know the weather, play your music, snap the perfect photo, check on your finances, and book you a plane ticket. It's just that along the way, they stopped being especially good phones. It's as if in their haste to emphasize "smart," they forgot about being phones. In fact, half of the Harris Poll respondents agree that "modern mobile phones are trying to do so much they've forgotten about voice quality."

With new approaches to noise reduction using neural-network technology, we see improvements that offer over three times more voice isolation than the best available smartphone.

Remember, this used to be state of the art (*Fig. 2*). Until it wasn't.



4. Languages are different, so it's difficult to use voice isolation for noise control when you're developing products to be used worldwide.

We don't care about language—only the sounds of the human voice. We use technology to break down speech into its most basic elements. In effect, it's not constrained by the language spoken, but by the sounds that a human can make. That's a finite set. By comparing these basic phonetic elements to the background noise, we can make a very fast, very accurate decision as to what is speech and what is not.

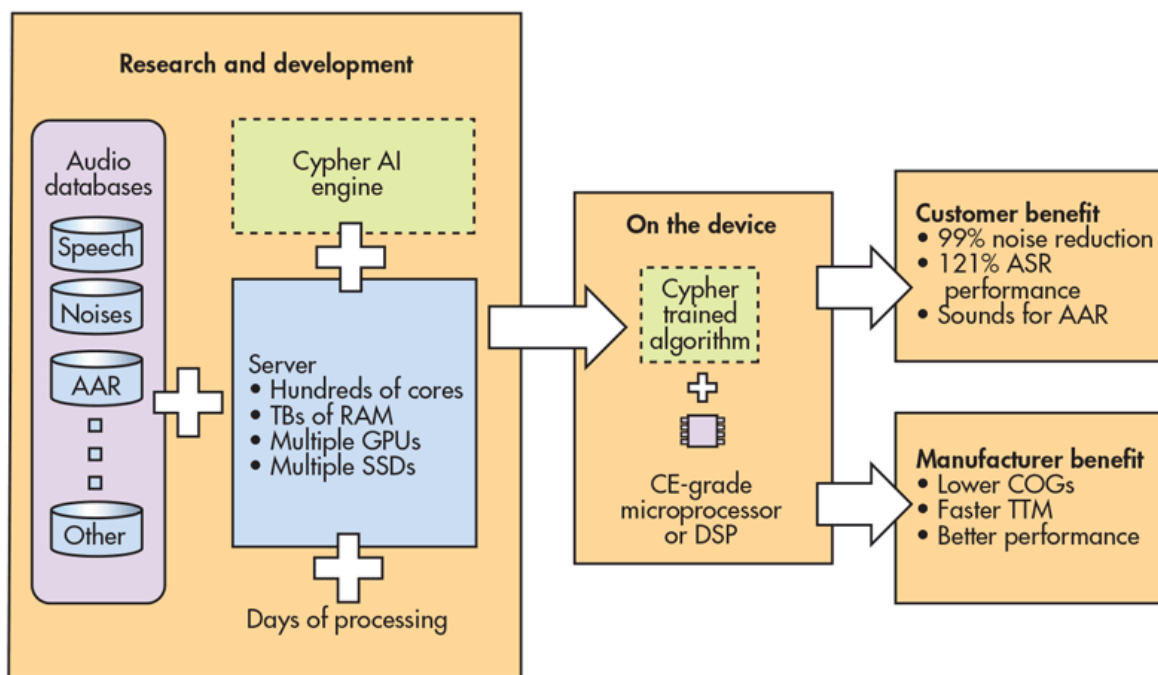
5. Dedicated hardware is always better than software solutions for noise cancellation.

This one is tough to argue in the abstract. If you could ignore cost, size, power consumption, and a lot of other factors at the core of today's consumer-electronics value proposition, then you might be able to make a better

tion with dedicated hardware. Of course, that project proposal would be stalled before it even got off the ground.

Or, you could take a pure software solution using new technology, run it on existing hardware, and get a 300% improvement over the dedicated hardware solutions of today.

Many attempts have been made to improve voice signal transmission, using everything from phase cancellation to beamforming microphones to HD voice technology to bandwidth upgrades. Unlike current technologies, Cypher's neural-net-based solution doesn't rely on the acoustical design of hardware or network bandwidth upgrades (*Fig. 3*). The approach harnesses the smartphone's processing power by seamlessly integrating the device. As a result, it's significantly less expensive and simpler to implement than hardware solutions while providing superior performance.



6. It's nearly impossible to deliver effective noise cancellation in a city environment, where noises come from all around and are unpredictable.

Cities are an interesting place for sounds. Many of us have had the experience of wearing top-of-the-line noise-canceling headphones, and we still hear all sorts of random, impulsive noises leaking through.

This is because traditional background-noise-cancellation methods like those used in noise-canceling headphones and today's smartphones are really best at eliminating lower frequency, steady noise sources. Think jet engine noise, not the passenger two rows back whom will not stop hitting on his seatmate.

With the availability of new software-based options to boost clarity and improve speech recognition, it's also now possible to effectively and affordably add this technology well into the planning process. In Cypher's case, we learned early on that we wouldn't be able to isolate and block every random sound. That led us to the discovery and development of a unique neural-network and machine-learning approach that enabled us to isolate all of the sounds of the human voice, and only allow those through. In other words, we block virtually all background noise by isolating and only letting through the human voice.

7. Neural networks are too big to run in consumer devices; thus, they're too cost- and space-

Inhibitive to use effectively for noise cancellation.

Training a neural network requires several things:

- A large set of target data (in this case lots of human voice recordings).
- A large set of things that aren't the target (lots of background noise recordings).
- A well-designed neural network.
- Lots of server time.

It's not practical to do this on your smartphone or your Internet appliance. Luckily, you don't have to—we did that already. The result is actually very small, very fast, and very accurate, and you can put THAT on the consumer device.

Printing presses are large. Newspapers are not.

8. Pre-training is required.

We have all had experience with systems that learn your voice. Typically, to get the best user experience, you have to speak to the device in order to train it to your voice. On systems where this isn't required, like Apple's Siri or Amazon's Alexa, a connection to a large computing platform in the cloud helps these systems deal with the issues of pre-training.

The neural-net approach is different in that it's not training to a particular voice, but rather the characteristics that comprise human speech. That's the training we do BEFORE we port the software to the device. Therefore, in effect, the algorithm is already trained—not on your voice, but on an amazing array of humans.

9. Wideband audio will make noise suppression unnecessary.

Current cellphones transmit only a fraction of the sound that you can hear. This is done for several reasons. It is similar to the way we set up wired phone technology. It's bandwidth-efficient, covering the frequency range that has the most voice energy.

All of that is changing as communication is now digital. With more sophisticated encoding, we can pass more data efficiently. As a result, new system upgrades will double or even quadruple the audio frequencies transmitted. This will make for a richer, more complete recreation of the speaker's voice. Unfortunately, it also allows for more noise to be picked up and transmitted.

While it is true that additional audio-processing steps are part of wideband audio, our ability to pre-filter for voice will be even more important as with the potential doubling of background noise.

10. I use Skype/Facetime/Google Talk, so noise cancellation isn't necessary.

That's great. These services typically have better audio due to the high bandwidth they can use in either a wired or wireless network. That doesn't change the fact that Skyping from a crowded coffee shop can still be misery for the person on the other end of the call.

Even with the high bandwidth of an Ethernet or Wi-Fi network and the directionality of a laptop or tablet camera, background noise is still going to be an unwanted part of the experience. Neural-net technology is designed to stop noise at this very edge of the call. In effect, the algorithm doesn't care if the filtered speech is then sent on its way via a cellular network or a data network.

Cloud-based service is required for effective noise-cancellation solution.

Wireless networking speed and coverage only recently reached a point where cloud services can accommodate an automated-speech-recognition (ASR) vocabulary library of 200 trillion+ possible word combinations, as well as allow for real-time learning with contextual clues. And while this is a nice to have option, it's not necessarily required. For instance, we can do all our training before the algorithm is ported to the device. No network? No problem.

Now, if we are filtering noise for another service like Siri or Alexa, those services typically need a connection back to a server. However, the cloud isn't required for noise-cancellation solutions to improve voice communication and ASR accuracy.

Neural-net-based software mutes the background sounds, letting the user's voice through without distraction. In recent standards-based testing conducted by [Spirent](#), Cypher's approach demonstrated up to 17 times more noise reduction than the most popular smartphones on the market.



Looking for parts? Go to sourceesb.com

Source URL: <http://electronicdesign.com/embedded/11-myths-about-noise-cancellation>