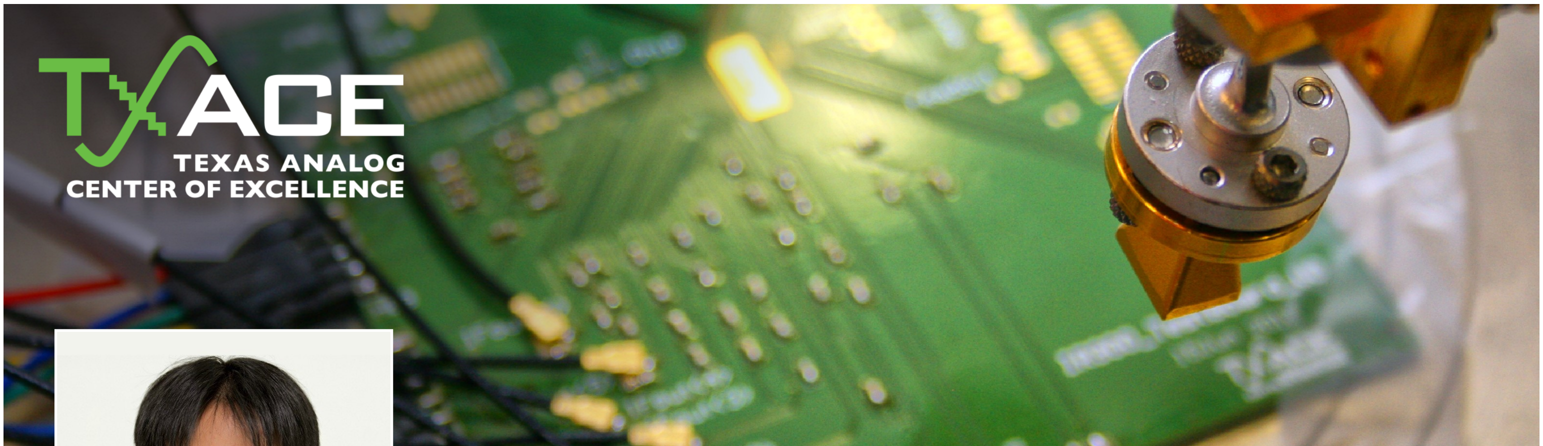

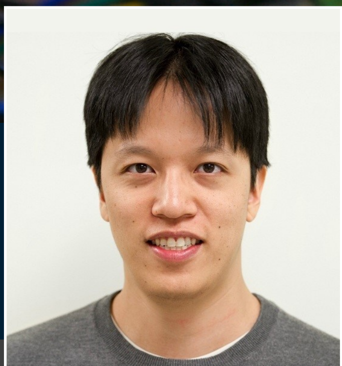

Sub-Microwatt End-to-End Keyword Spotting Chip that is Robust to Background Noise

Mingoo Seok
Associate Professor
Columbia University

BIO

Mingoo Seok is an associate professor of Electrical Engineering at Columbia University. He received the B.S. from Seoul National University, South Korea, in 2005, and the M.S. and Ph.D. degree from the University of Michigan in 2007 and 2011, respectively, all in electrical engineering. His research interests are various aspects of VLSI circuits and architecture, including ultra-low-power integrated systems, cognitive and machine-learning computing, an adaptive technique for the process, voltage, temperature variations, transistor wear-out, integrated power management circuits, event-driven controls, and hybrid continuous and discrete computing. He won the 2015 NSF CAREER award and the 2019 Qualcomm Faculty Award. He is the technical program committee member for multiple conferences, including IEEE International Solid-State Circuits Conference (ISSCC). In addition, he has been as an associate editor for IEEE Transactions on Circuits and Systems Part I (TCAS-I) (2014-2016), IEEE Transactions on VLSI Systems (TVLSI) (2015-present), IEEE Solid-State Circuits Letter (SSCL) (2017-present), and as a guest associate editor for IEEE Journal of Solid-State Circuits (JSSC) (2019).

ABSTRACT

Keyword spotting is one of the most familiar features of mobile devices. A mobile phone spots a keyword such as “Hi Siri,” “Hello Google,” and “Hi Bixby,” and once it spots it, it starts to pay attention to the following speech command and perform more complex speech recognition task. Today, a mobile phone consumes milliwatts (10^{-3}) for this task, which is acceptable for its given battery size. However, smaller devices, such as wearables and Internet-of-Thing (IoT) devices, often find such power consumption unacceptable because it has an orders-of-magnitude smaller battery size. For the last few years, therefore, several Ph.D. students and postdocs in my research group at Columbia University have been working on reducing the power consumption of a keyword spotting IC chip down to 100’s of nanowatts (10^{-7}). A significant reduction was realized with multiple novel architecture and circuits design in all three stages of acoustic signal processing, namely analog front end, data conversion, and neural-network-based inference. Power optimization, however, usually comes with reduced robustness. For example, we employ a simpler neural network inference model to minimize power consumption, but such a small model tends to degrade keyword-spotting accuracy under background noise. Not to lose the robustness while reducing power, we have employed a noise-tolerant mechanism called divisive energy normalization (DEN), commonly found in many biological acoustic sensory systems. Our chips can detect a keyword under severe background noise by adopting DEN and even though the background noise is unknown during the neural network training time.

Friday, January 21, 2022 at 1:00 – 2:00 p.m.
Virtual

UT DALLAS | Erik Jonsson School of Engineering and Computer Science