Technical Article

Advancements in audio technology help humanoids hear and speak in the real world

Jonathan Huang, Ramsey Foote, and Christian Greeff, Texas Instruments

Humanoid robots are moving from the lab into factories, hospitals, and homes, but those environments can be loud, cluttered, and unpredictable. Enabling a robot to reliably hear commands and respond naturally is one of the hardest engineering challenges in robotics design. Today, advances in microphone arrays, beamforming, and high-efficiency amplification are making it possible, opening the door to audio interaction that feels less mechanical and more human.

Microphone-based audio input lets humanoids receive and act on spoken commands, making interaction with humans more natural and intuitive without the need for touchscreens or pre-programmed routines. Onboard microphones also allow robots to monitor environmental audio cues such as a ringing alarm, a shout or a collision warning, enabling context-aware behavior in real time. Achieving this in noisy, real-world situations such as the situation in Figure 1 requires microphone arrays and beamforming – a signal-processing technique that focuses pickup in a specific direction – to isolate the source of sound and capture far-field audio clearly.

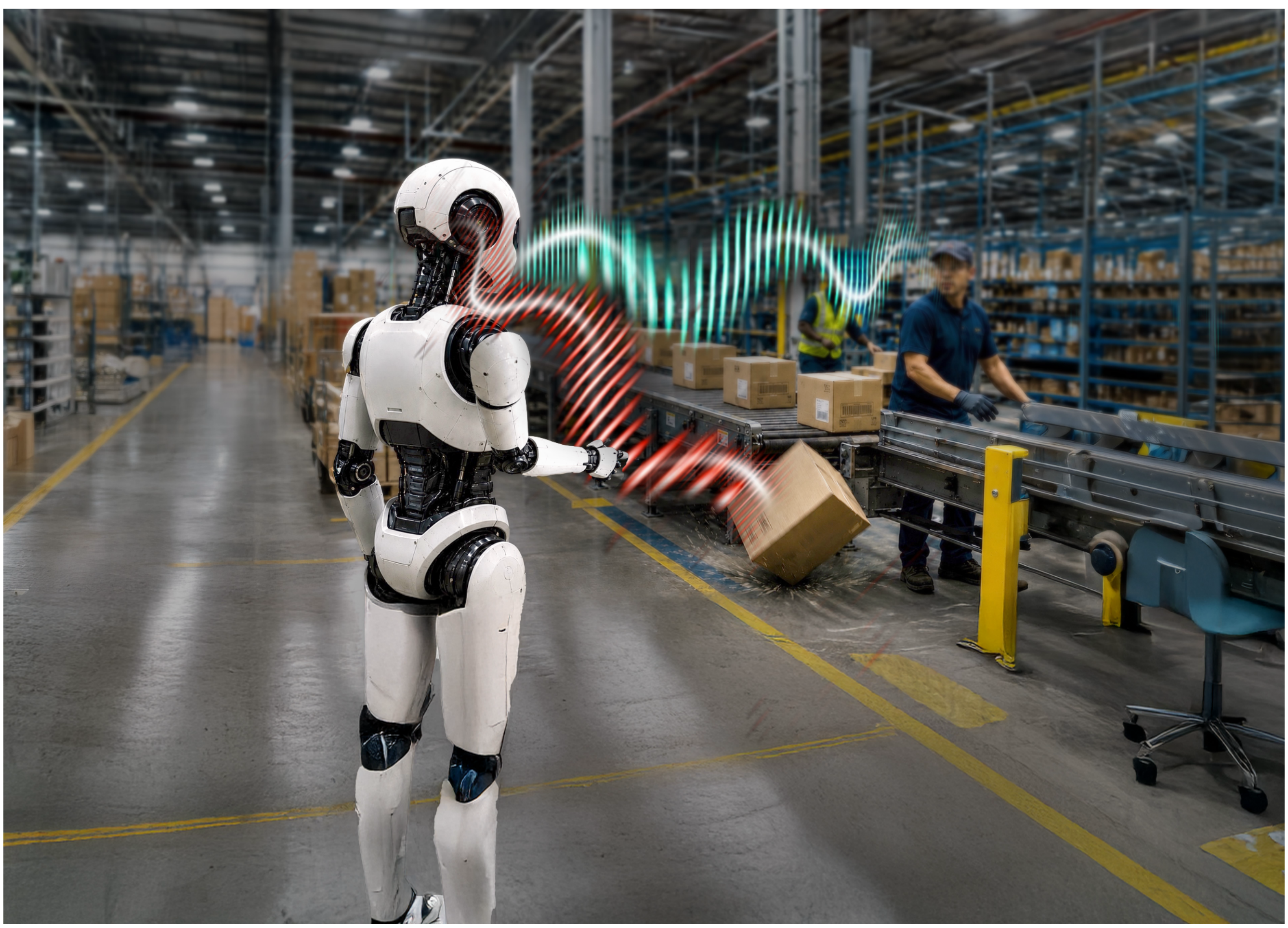

Figure 1. A humanoid robot audibly warns factory workers about an unsafe situation

Speaker output is equally critical for safety. Humanoids must produce clear, high-quality speech, whether responding to a user, providing a status update or broadcasting an emergency alert so that people nearby can stay informed and respond appropriately. Clear, intelligible output also makes two-way interaction feel more natural, building trust with humans in collaborative environments. For a robot to listen and speak at the same time without degrading either audio signal, the system must employ noise suppression and acoustic echo cancellation, techniques that keep incoming audio clean and outgoing audio intelligible even when both are active simultaneously.

Delivering all of these capabilities requires a complete audio signal chain, from microphone aggregation supporting analog and digital inputs to high-efficiency Class-D amplifiers that maximize output power without draining the battery.

How these features are enabled

Record path algorithm

Humanoid robots that communicate effectively in real-world environments typically rely on beamforming, which uses a microphone array of four or more microphones positioned around the robot's head. The beamforming algorithm processes the audio data from each microphone to identify the direction of the strongest audio signal. Once identified, the algorithm applies precisely calculated time delays to each microphone's input to electronically steer the pickup pattern toward the desired audio source. Designers can apply additional tuning to target specific profiles, such as human speech or other defined sounds.

Beamforming lets the microphone array focus on audio from a specific direction while rejecting unwanted sounds from other locations. The multi-microphone approach also improves far-field audio capture by summing the signals across channels, which reinforces audio that is common to all microphones while reducing uncorrelated noise. Noise-suppression algorithms applied to the beamformed audio signal further improve the overall audio quality before the signal reaches the host processor.

High power, high efficiency

Humanoid robots are often space- and thermally constrained. High-power motors and processors are packed into a compact space with limited airflow demand highly efficient amplifiers that are rated to operate in high temperatures. Integrated circuits (ICs) that eliminate the need for external heat sinks offer a clear advantage in these tight designs such as shown in Figure 2. TI's TAS6511-Q1 is a mono amplifier that delivers 50W of output power and supports operating temperatures up to 125°C, all in a pad-down 4mm-by-4mm QFN package. For designs using analog input amplifiers, the TPA3132D2 analog-input Class-D amplifier supports 50W into a mono speaker output and does not require any I2C configuration. Where less output is needed, such as in a smaller enclosure or lower ambient noise environment, the TAS6501-Q1 supports 15W of output power.

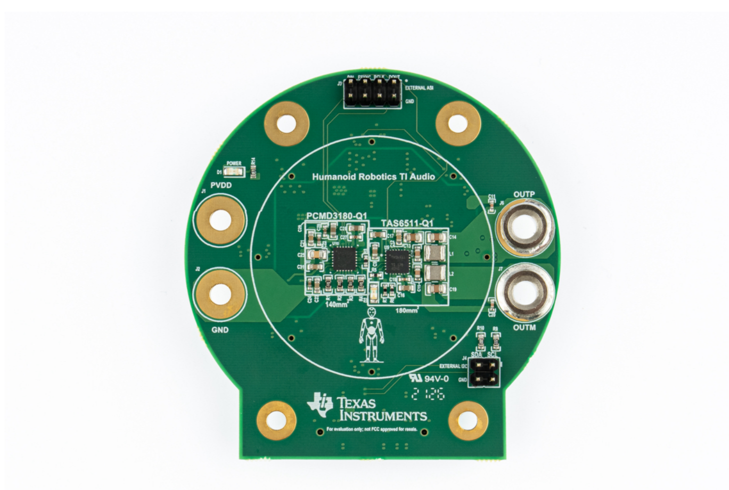

Figure 2. PDM-to-TDM audio converter and audio amplifier in TI's humanoid robotics audio subsystem reference design (TIDA-060054)

How are these features being implemented today?

Microphone array aggregators

To effectively process audio from its environment, a humanoid robot must aggregate all microphone data into a unified format that the onboard host processor can act on. A time division multiplexing (TDM) audio bus is the typical implementation, serving as a common language between the microphone array and processor.

Achieving this requires bridging the gap between how microphones capture sound and how the host processor expects to receive it. Analog microphones record analog data; digital microphones record pulse density modulation (PDM), neither of which is the TDM format required by the host processor. An audio converter between the microphones and the processor aggregates and translates these inputs into TDM, enabling seamless, real-time audio processing at the system level as shown in Figure 3.

TI offers solutions for both signal types. The TLV320ADC3140 is a high-performance audio ADC for analog microphones, while the PCMD3180 is a PDM-to-TDM audio converter for digital microphones. Both ensure that every microphone input arrives at the host processor in a TDM format.

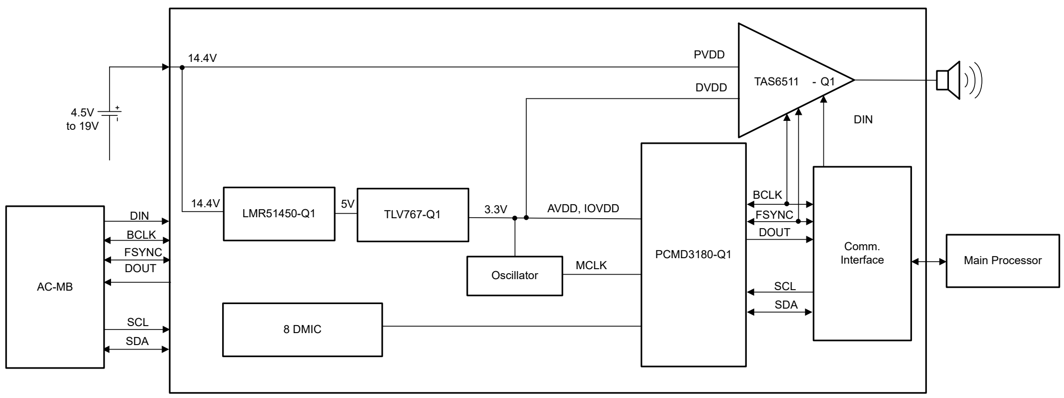

Figure 3. Typical audio block diagram of a humanoid robot

Small solution size and low BOM cost

In some humanoid robots, the audio system is housed in the head unit, where microphones and speakers must fit within a tightly constrained volume before routing data to the main processor near the chest. This space-constrained environment places significant demands on the size and integration of the chosen ICs.

Small, low profile IC packages such as the 4mm-by-4mm quad flat no-lead (QFN) chips used by the PCMD3180 audio converter, TAS6511-Q1 Class-D amplifier and TAC5212 low-noise audio codec can sit side by side without crowding the board. A slightly larger 5mm-by-5mm QFN device, the TPA3132D2 Class-D amplifier operates without an external I2C configuration.

Using pad-down (bottom-contact) packages eliminates the need for bulky heat-sinks, helping devices to manage thermal dissipation and fit within the limited space in the head unit. The TAS6511-Q1 operates at a 2.1MHz switching frequency, which significantly reduces the size of the output-filter inductors from 10µH to 3.3µH, allowing the entire audio signal chain to remain compact without sacrificing efficiency or output quality.

High efficiency and low power draw

Battery runtime is one of the most critical considerations in humanoid robot design. Motor drivers, the host processor, perception sensors, audio and a host of other systems all draw from the same power source simultaneously, so every component must minimize its energy consumption. High efficiency and low idle power consumption are foundational to making a humanoid viable for real-world operation.

TI addresses this directly across its audio portfolio. The PCMD3180 converter draws just 2.5mW at idle, while the TAS6511-Q1 amplifier consumes less than 0.5W in the same state, minimizing the audio system's drain on the battery when full output isn't needed. At full output, the TAS6511-Q1 delivers best-in-class 90% efficiency at 50W output power, minimizing draw on the battery.

Turnkey solutions

Many humanoid robotics companies have limited design engineering resources to directly dedicate to audio development. With resources spread across motion, perception, and AI, audio components must be quick to find and easy to implement.

TI designs its audio portfolio with this in mind. Each IC, including the PDM-to-TDM converter, Class-D amplifiers and audio codec, ships with application notes, evaluation modules and support through Pure Path Console, TI's audio configuration software. Designers can use these resources to configure, test, and debug their audio system quickly, without requiring deep audio expertise.

Conclusion

As humanoid robots move from controlled environments into factories, hospitals and homes, precision audio quality will define how well they function alongside people. The ability to hear clearly, respond intelligently and communicate safely in real-world noise is foundational to whether a robot can be useful and reliable. TI's compact, efficient and easy-to-integrate audio technology gives robotics design teams the tools to meet that challenge without sacrificing Sespace, power or development time, so designers can focus on what comes next: building robots that safely operate where people live and work.

Additional resources

- **Find your parts:** Browse product pages for the PCMD3180, TAS6511-Q1, TAC5212, TPA3132D2 and TI's robotics product overview.
- **Optimize your design:** Download Pure Path Console and review power consumption app notes for the TAC5212 and PCMD3180.
- **Ask an expert:** Post your audio design questions on the TI Audio E2E Forum.

About the authors

Jonathan Huang

Jonathan Huang is a product marketing engineer at TI with a focus on audio converters. He holds a bachelor's degree in electrical engineering from Georgia Tech. In his free time, he enjoys trying new recipes and traveling internationally.

Ramsey Foote

Ramsey Foote is a systems engineer at TI with a focus on audio converters. He holds a master's degree in electrical engineering from the University of Illinois Urbana Champaign, and spends his free time designing audio equipment and rock climbing.

Christian Greeff

Christian Greeff is a robotics systems engineer at TI, focusing on humanoid robot solutions. He earned a degree in electrical engineering from Florida Agricultural and Mechanical University. He designs personal projects including e-bikes, electric go karts, custom electric ski lift winches, and a 20kW electric personal aircraft.

Trademarks

All trademarks are the property of their respective owners.

IMPORTANT NOTICE AND DISCLAIMER