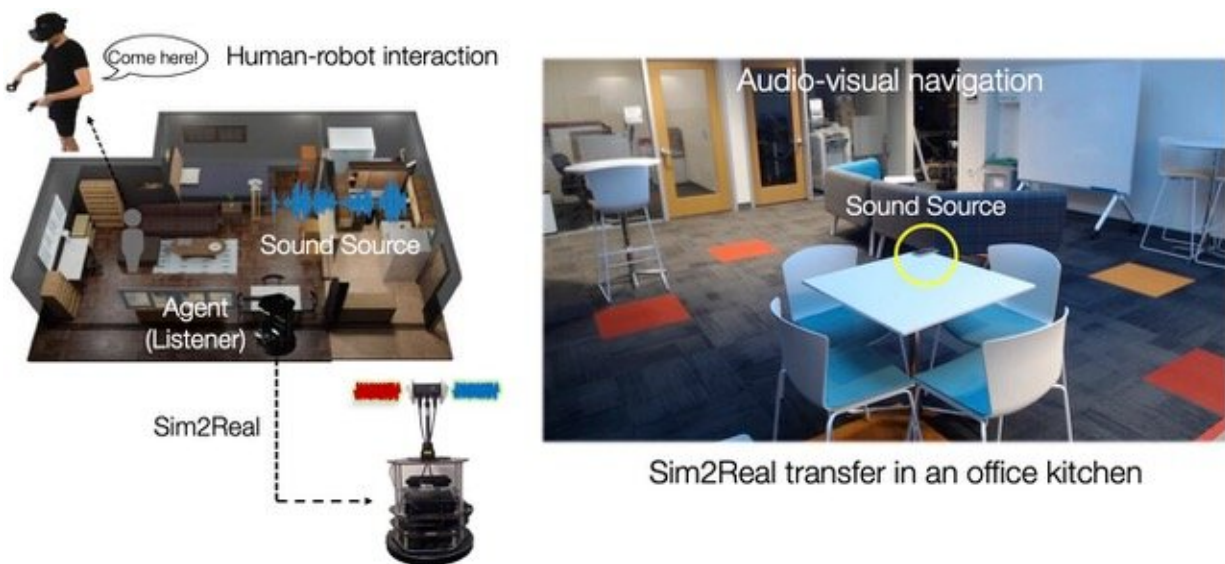


# A multisensory simulation platform to train and test home robots

June 16 2023, by Ingrid Fadelli



In the Sonicverse embodied AI simulator, the agent in the environment can act as a listener to receive directional information about the sound source and perform tasks that require audio-visual perception. The platform also features an audio-visual virtual reality interface that enables human-robot interaction through voice commands. The team trained audio-visual embodied navigation agents in their simulator and successfully performed Sim2Real transfer in real-world environments such as an office kitchen. Credit: Gao et al

AI-powered robots have become increasingly sophisticated and are gradually being introduced in a wide range of real-world settings, including malls, airports, hospitals and other public spaces. In the future,

these robots could also assist humans with house chores, office errands and other tedious or time-consuming tasks.

Before robots can be deployed in [real-world](#) environments, however, the AI algorithms controlling their movements and allowing them to tackle [specific tasks](#) need to be trained and tested in simulated environments. While there are now many platforms to train these algorithms, very few of them consider the sounds that robots might detect and interact with as they complete tasks.

A team of researchers at Stanford University recently created Sonicverse, a simulated environment to train embodied AI agents (i.e., robots) that includes both visual and auditory elements. This [platform](#), introduced in a paper presented at [ICRA 2023](#) (and currently available on the *arXiv* preprint server), could greatly simplify the training of algorithms that are meant to be implemented in robots that rely on both cameras and microphones to navigate their surroundings.

"While we humans perceive the world by both looking and listening, very few prior works tackled embodied learning with audio," Ruohan Gao, one of the researchers who carried out the study, told Tech Xplore. "Existing embodied AI simulators either assume that environments are silent and the agents unable to detect sound, or deploy audio-visual agents only in simulation. Our goal was to introduce a new multisensory simulation platform with realistic integrated audio-visual simulation for training household agents that can both see and hear."

Sonicverse, the simulation platform created by Gao and his colleagues, models both the visual elements of a given environment and the sounds that an agent would detect while exploring this environment. The researchers hoped that this would help to train robots more effectively and in more "realistic" virtual spaces, improving their subsequent performance in the real world.

"Unlike prior work, we hope to demonstrate that agents trained in simulation can successfully perform audio-visual navigation in challenging real-world environments," Gao explained. "Sonicverse is a new multisensory simulation platform that models continuous audio rendering in 3D environments in real time, It can serve as a testbed for many embodied AI and human-[robot](#) interaction tasks that need audio-visual perception, such as audio-visual navigation."

To evaluate their platform, the researchers used it to train a simulated version of TurtleBot, a robot created by Willow Garage, to effectively move around in an indoor environment and reach a target location without colliding with obstacles. They then applied the AI trained in their simulations to a real TurtleBot and tested its audio-visual navigation capabilities in an office environment.

"We demonstrated Sonicverse's realism via sim-to-real transfer, which has not been achieved by other audio-visual simulators," Gao said. "In other words, we showed that an agent trained in our simulator can successfully perform audio-visual navigation in real-world environments, such as in an office kitchen."

The results of the tests ran by the researchers are highly promising, suggesting that their [simulation](#) platform could train robots to tackle real-world tasks more effectively, using both visual and auditory stimuli. The [Sonicverse platform](#) is now available online and could soon be used by other robotics teams to train and test embodied AI agents.

"Embodied learning with multiple modalities has great potential to unlock many new applications for future household robots," Gao added. "In our next studies, we plan to integrate multisensory object assets, such as those in [our recent work ObjectFolder](#) into the simulator, so that we can model the multisensory signals at both the space level and the object level, and also incorporate other sensory modalities such as tactile

sensing."

**More information:** Ruohan Gao et al, Sonicverse: A Multisensory Simulation Platform for Embodied Household Agents that See and Hear, *arXiv* (2023). [DOI: 10.48550/arxiv.2306.00923](https://doi.org/10.48550/arxiv.2306.00923)

© 2023 Science X Network

Citation: A multisensory simulation platform to train and test home robots (2023, June 16) retrieved 8 September 2024 from <https://techxplore.com/news/2023-06-multisensory-simulation-platform-home-robots.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.