

Новая технология Target Speech Hearing от Университета Вашингтона выделяет важные голоса среди окружающего шума.

Университет Вашингтона представил новую систему искусственного интеллекта, которая позволяет пользователям наушников выделять речь определенного человека среди окружающего шума. Эта технология, получившая название «Target Speech Hearing» (TSH), была представлена 14 мая на конференции ACM CHI Conference on Human Factors in Computing Systems в Гонолулу. Код системы доступен для дальнейших разработок, хотя сама система пока не продается.

Система TSH позволяет пользователю наушников «записать» голос интересующего собеседника, просто посмотрев на него в течение нескольких секунд. После этого система отсеивает все другие звуки и передает только голос выбранного человека, даже если пользователь наушников перемещается или поворачивает голову.

Разработка направлена на улучшение слухового восприятия в шумных условиях, таких как многолюдные улицы или офисы. По словам ведущего автора исследования, профессора Университета Вашингтона Шьяма Голлакота, эта система отличается от типичных интернет-чатботов, применяя ИИ для изменения слухового восприятия в реальном времени. «С нашими устройствами теперь можно четко слышать одного человека даже в шумной среде с большим количеством разговоров», — отметил Голлакота.

Для использования системы достаточно надеть обычные наушники с микрофонами и нажать кнопку, направив голову на говорящего человека. Звуковые волны от его голоса достигнут микрофонов с обеих сторон гарнитуры одновременно (с допустимой погрешностью в 16 градусов). Наушники передают этот сигнал на встроенный компьютер, где программное обеспечение обучается распознавать голос выбранного собеседника и продолжает воспроизводить его в реальном времени.

В ходе испытаний системы 21 участник оценивал четкость воспроизведения голоса. Средняя оценка ясности была почти вдвое выше по сравнению с необработанным звуком.

Разработка TSH основывается на предыдущих исследованиях команды в области «семантического слуха», которые позволяли пользователям выбирать определенные звуки, такие как голоса или пение птиц, и отменять остальные звуки окружающей среды.

В текущей версии системы можно зарегистрировать только одного говорящего за раз, и это возможно только при отсутствии громких звуков из того же направления, откуда идет голос цели. Если пользователь недоволен качеством звука, он может повторно зарегистрировать голос, чтобы улучшить четкость.

Команда планирует расширить возможности системы, адаптировав её для использования с вкладышами и слуховыми аппаратами.