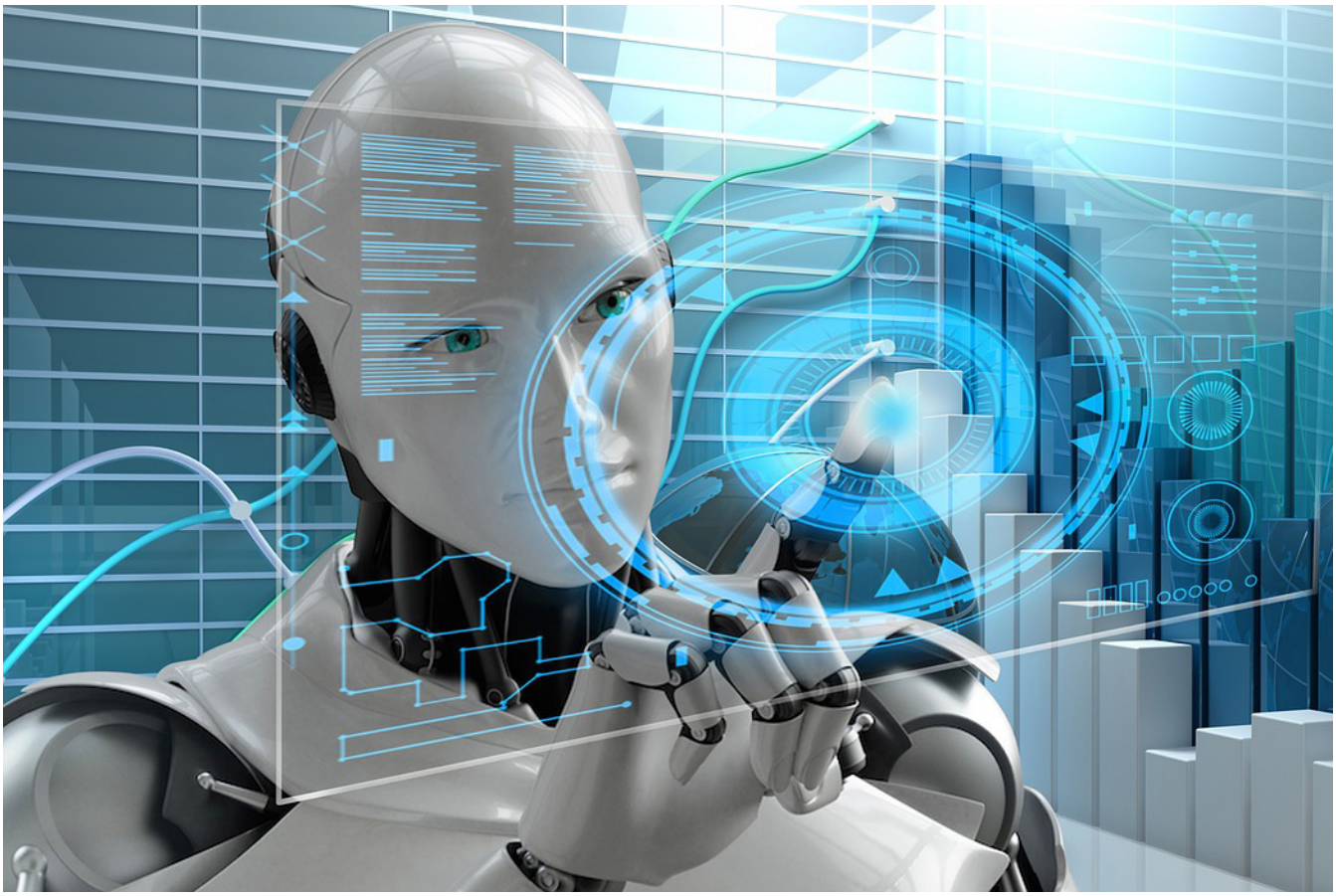
Ученые Политеха наделят искусственный интеллект человеческим слухом

Цифровая обработка и анализ акустических сигналов – одна из актуальных на сегодняшний день задач для ученых. Специалисты неплохо продвинулись в этом направлении, однако, чем больше проблем решается, тем более сложные остаются. Глобальная задача – цифровая обработка речи в реальных звуковых условиях. Это значит, что звуковой сигнал подвержен шумовому воздействию разных источников звука, например, при одновременном разговоре нескольких людей. И пока еще этот вопрос остается открытым.



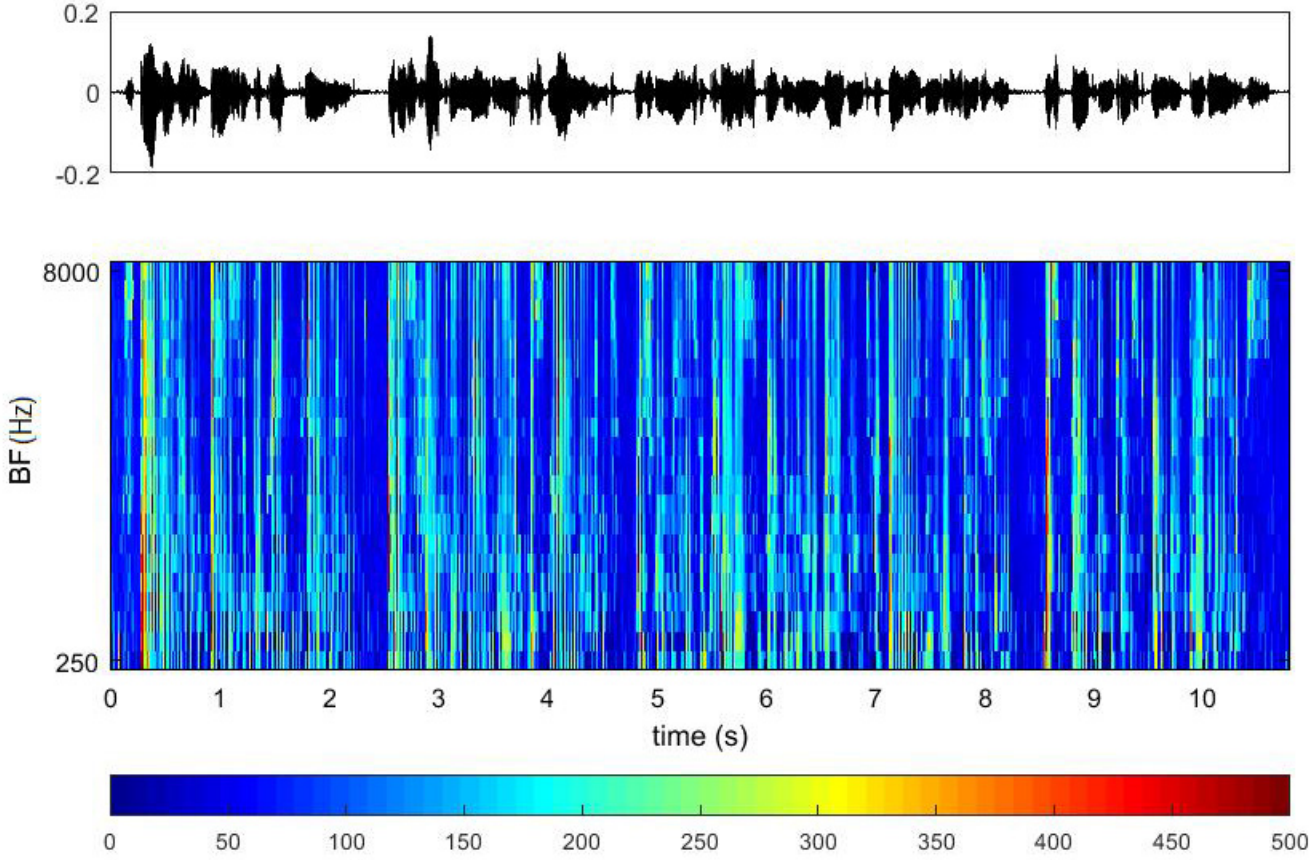
Специалисты кафедры «Измерительные информационные технологии» ИКНТ СПбПУ предложили свой подход к решению проблемы, в основе которого лежит моделирование процесса рецепторного кодирования звуков путем имитации работы периферической части нервной системы человека, отвечающей за слух.

Нервная система обрабатывает информацию в форме электрических сигналов. Восприятие внешней среды обеспечивается комплексными

подсистемами – анализаторами, рецепторы и периферические нервы которых обеспечивают преобразование энергии стимула в поток электрических импульсов и их проведение в центральные отделы для дальнейшей обработки. В результате таких базовых принципов работы слуховой системы, человек может без каких-либо значительных усилий распознать голос собеседника в сильно зашумлённых условиях. При этом, автоматизированные системы, созданные на данный момент, не представляют столь эффективного решения и требуют мощных вычислительных ресурсов.

Проект, направленный на решение данной задачи получил поддержку РФФИ, в рамках фундаментальных исследований, выполняемых молодыми учеными (№ 18-31-00304). В процессе его реализации ученые разрабатывают методы обработки и анализа акустических сигналов на базе периферического кодирования. Говоря проще, специалисты частично воспроизведут процессы, выполняемые нервной системой при обработке информации и синтезируют их с модулем принятия решений, который определяет, что за сигнал был получен.

По словам руководителя проекта, Антона Александровича ЯКОВЕНКО, глобальная цель – приблизить возможности машины к способностям человека, то есть обеспечить машинное восприятие акустических сигналов в реальных условиях.



На текущем этапе работы специалистами смоделирована значительная выборка образцов нервной деятельности для гласных фонем и предложен подход, сочетающий самоорганизующиеся нейронные сети и теорию графов. Обработка данных производится специальным алгоритмом, который осуществляет их структурный анализ для выявления «паттернов», описывающих каждую фонему. Анализ реакции модели слухового нерва позволил идентифицировать гласные фонемы при значительном шумовом воздействии и превзошел по качеству наиболее распространенные на сегодняшний день методы параметризации акустических сигналов. Соответствующие результаты данного исследования приведены в научной статье.

Главной особенностью подхода является разрозненность входных данных, которые никак не структурированы. Это отличает предлагаемый политехниками метод от ныне существующих, где структура данных известна изначально, а алгоритму необходимо ее «выучить» и впоследствии применить. В итоге, машина «самообучается» так же, как человек, который впервые воспринимает те или иные звуковые сигналы.

Получаемые в ходе исследования результаты носят в значительной степени фундаментальный характер и прежде всего имеют ценность для междисциплинарных областей науки. Разрабатываемые методы могут лечь в основу нового поколения нейрокомпьютерных интерфейсов, а также обеспечить более качественное человеко-машинное взаимодействие. В связи с этим потенциал практического применения представляется значимым в разнообразных сферах деятельности: кохлеарная имплантация, разделение звуковых источников, создание новых бионических методов распознавания голоса, речи и акустических событий.

«Реализуемые в рамках исследования алгоритмы обработки и анализа больших данных сами по себе являются универсальными и могут применяться для решения прикладных задач, не связанных с акустическими сигналами, – подытожил Антон Александрович. – *Например, один из методов был успешно применен в задаче обнаружения аномалий сетевого трафика. А сенсорное кодирование – универсальный язык нервной системы, поэтому разработку гипотетически можно адаптировать под другие модальности, такие как зрение или осязание, и более широко использовать в нейропротезировании»*.

Материал подготовлен Сектором научных коммуникаций СПбПУ. Текст: Мария ГАЙВОРОНСКАЯ

Дата публикации: 2019.01.28

>>Перейти к новости

>>Перейти ко всем новостям