

Информационные технологии

УДК 004.4

Лю Чао

Амурский государственный университет
г. Благовещенск, Россия

E-mail: liuchao0805@yandex.com

Самохвалова Светлана Геннадьевна

Амурский государственный университет
г. Благовещенск, Россия

E-mail: sgs@amursu.ru

Liu Chao

Amur State University
Blagoveshensk, Russia

E-mail: liuchao0805@yandex.com

Samokhvalova Svetlana Gennadievna

Amur State University
Blagoveshensk, Russia

E-mail: sgs@amursu.ru

СИСТЕМЫ РАСПОЗНАВАНИЯ РЕЧИ РОБОТОВ: ЗАРУБЕЖНЫЕ ИССЛЕДОВАНИЯ**ROBOTS SPEECH RECOGNITION SYSTEMS: FOREIGN RESEARCH**

Аннотация. Статья представляет собой обзор исследований системы распознавания речи для интеллектуального робота. Проанализирован текущий статус исследований в области технологии распознавания речи в зарубежных изданиях, демонстрирующих перспективы применения технологии распознавания речи в сфере «интеллектуального производства».

Abstract. This article is an overview of the research into a speech recognition system for an intelligent robot. The article analyzes the current status of research in the field of speech recognition technology in foreign studies, demonstrating the prospects for the application of speech recognition technology in the field of smart manufacturing."

Ключевые слова: робот, распознавание речи, искусственный интеллект.

Key words: robot, speech recognition, artificial intelligence.

Введение

В наступившую эру искусственного интеллекта интеллектуальные роботы становятся все более популярными. Предварительно они загружаются программным обеспечением и могут автоматически выполнять определенную работу в соответствии с встроенной программной системой. Роботы способны заменить людей во многих сферах деятельности. Интеллектуальные робототехнические системы используются для выполнения опасных и трудоемких

задач, помогая снизить риск в ежедневной деятельности людей. Однако большинство таких роботов могут принимать только предварительно подготовленные встроенные программы, их движения фиксированные и механичные. Хотя в настоящее время уже существуют отдельные типы роботов, способных взаимодействовать с операторами. Операторы вводят информацию и команды в основном с клавиатуры, что ограничивает взаимодействие между ними и роботом.

Естественный язык является наиболее удобным средством взаимодействия между людьми, а поэтому есть основания полагать, что естественный язык в будущем станет одним из наиболее эффективных средств взаимодействия между роботами и людьми. За последние несколько лет технические специалисты приложили немало усилий для изучения и применения технологии распознавания речи. Использование технологии распознавания речи для управления механическими и электрическими устройствами стало одним из передовых исследовательских направлений, что позволило достичь значительного прогресса в управлении промышленными роботами.

Прежде чем выпустить на рынок роботов с низкой стоимостью и высокой производительностью, нужно, чтобы они действительно вошли в повседневную жизнь людей, а для этого необходимо исследовать и разработать высокопроизводительную систему распознавания речи.

Анализ текущего состояния исследований за рубежом

История распознавания электронными машинами человеческой речи началась в 1927 г., когда в США был создан робот *Televox*. Он представлял собой несложное устройство, в котором различные реле реагировали на звуки разной частоты. У робота было три камертона, каждый из которых отвечал за свою тональность. В зависимости от того, какой камертон срабатывал, включалось то или иное реле.

Следующим ключевым изобретением, открывшим путь к настоящему распознаванию человеческой речи, стала машина *Audrey*, разработанная в 1952 г. в кузнице инноваций *Bell Labs*. Огромная *Audrey* потребляла много электроэнергии и была размером с хороший шкаф, но вся ее функциональность сводилась к распознаванию произнесенных цифр от нуля до девяти. Несмотря на революционные для своего времени результаты, *Audrey* не нашла практического применения. Предполагалось, что систему можно приспособить вместо телефонных операторов, однако услуги человека оказались удобнее, быстрее и гораздо надежнее.

В начале 1960-х гг. работы по созданию машин для распознавания речи велись в Японии, Великобритании, США и СССР, где изобрели очень важный алгоритм динамической трансформации временной шкалы, с помощью которого удалось построить систему, знающую около 200 слов. Но все наработки были похожи друг на друга, а общим недостатком стал принцип распознавания: слова воспринимались как целостные звуковые отпечатки и затем их сверяли с базой образцов (словарем). Любые изменения скорости, тембра и четкости проговаривания слов значительно влияли на качество распознавания.

На данный момент в технологии распознавания речи уже имеются значительные успехи, что связано с различными теоретическими подходами – такими как теория линейного предсказания, теория динамического программирования, теория векторного квантования и др. Перечисленные подходы стали теоретической основой для прикладных исследований в области технологии распознавания речи. В практическом плане появились различные прило-

жения, основанные на скрытых марковских моделях (НММ) и искусственных нейронных сетях (ANN) для распознавания речи. Например, это приложения, направленные на распознавание непрерывной речи от неопределенных пользователей – такие как система распознавания речи SPHINX, а также система распознавания речи Vi-aVoice, ориентированная на распознавание речи в новостях, и др. Эти приложения активизировали исследования в области искусственного интеллекта.

В большинстве стран основное внимание уделяется разработке и исследованию робототехники. Многие используют государственные инвестиции для стимулирования развития индустрии интеллектуальных роботов. Например, в 2013 г. правительство США создало «Национальную программу по роботам», вложив в нее 22 млрд долларов, технологии интеллектуальных роботов были названы ключевыми для военных нужд. В Европе Германия и Великобритания руководят исследованиями в области робототехники, вкладывая средства прежде всего в разработку технологий обслуживания и медицинских операций. В Японии, лидере в области электроники, акцент делается на разработке технологий робототехники, имитирующих человека, а также технологий роботов для развлекательных целей [4]. Область исследований в робототехнике становится все разностороннее и разнообразнее.

С начала XXI в., по мере старения населения в мире, растет спрос на интеллектуальных роботов. В развитых странах (США и страны Европы) инвестиции в исследования и разработки в области интеллектуального производства составляют 10%. Это стимулирует взрывной рост интеллектуальных роботов с разнообразными формами и функционалом – роботы для эмоциональной поддержки, ухода за пожилыми людьми, образования и тренировок, домашние помощники, а также роботы для обеспечения общественной безопасности и тушения пожаров. Системы распознавания речи для роботов быстро развиваются, обладают всё более высокой способностью логического мышления, организовывается язык для ответа на вопросы людей с высокой точностью в различных сценариях. Верится, что в ближайшем будущем, с уменьшением стоимости производства таких роботов, они станут доступными для обычных семей.

На нынешнем этапе одним из ключевых направлений исследований в области робототехники является технология облачных вычислений для роботов, а вторым по значимости – технология биоинспирированного компьютеринга мозга. США, Япония, Бразилия и другие страны рассматривают облачных роботов как одно из будущих направлений исследований в области робототехники. Кроме того, зарубежные исследования проблем искусственного интеллекта включают разработку открытых архитектур для роботов, платформ для взаимосвязи их в сети, разработку алгоритмов и систем обработки изображений для роботов, а также исследования в области инфраструктуры сети для облачных роботов.

Исследования по технологии распознавания речи Китай начал еще в 50-х гг. XX в. С 1987 г. страна запустила программу «863», создала экспертную группу по искусственному интеллекту «863» и лабораторию распознавания образов, организованную совместными усилиями университета Цинхуа и Института автоматизации Китайской Академии наук. Эта лаборатория специализируется на исследованиях технологии распознавания речи, в частности на китайском языке. КНР достигла мирового уровня в технологии распознавания речи на китайском языке. Разработанные системы распознавания речи на китайском языке и системы

голосового взаимодействия на китайском языке имеют точность более 96%, а уровень отклика системы превышает 90%.

В июле 2000 г. в Зоологическом музее Пекина была представлена выставка «Панда». В ее рамках была использована разработанная в Китае система для неконтролируемого распознавания речи. Посетители могли общаться с большой пандой на месте и узнавать о жизненных привычках, структуре пищи, физиологической структуре больших панд. В условиях выставочного зала, с его высоким уровнем шума, распознаваемость речи все равно оставалась выше 98%, что соответствует требованиям для практического использования. Этот живой и увлекательный способ презентации привлек много посетителей, заставив их останавливаться, смотреть и общаться.

В Китае используется технология самостоятельной разработки встраиваемых чипов для распознавания речи, включая специализированную систему чипов с функциями распознавания, кодирования и синтеза речи. Система состоит из 8-битного микроконтроллера, низкочастотного фильтра, аналого-цифрового и цифро-аналогового преобразователей, усилителя мощности, ППЗУ, ОЗУ, модулей ШИМ и др. Она способна распознавать 30 специфических голосовых команд, скорость кодирования речи достигает 16 кб/с, а точность распознавания превышает 95%. Эта технология может использоваться в умных голосовых игрушках, голосовых телефонах и других устройствах. Она уже находится в фазе коммерциализации и достигла мирового уровня.

6 апреля 2016 г. под руководством Министерства промышленности и информационных технологий, Национальной комиссии по развитию и реформам, Министерства финансов и других ведомств во главе с Ассоциацией беспилотных летательных аппаратов города Шэньчжэнь был создан альянс по развитию китайской робототехнической индустрии искусственного интеллекта. На совещании был утвержден «План развития» робототехнической индустрии, разработанный совместно Минпромторгом, Национальной комиссией по развитию и реформам, а также другими ведомствами. Этот план – руководство здоровым, быстрым и устойчивым развитием робототехнической индустрии Китая.

В «Плане развития» указаны основные проблемы в области исследований по распознаванию речи и голосовому управлению в Китае.

1. Исследования в области речевых технологий требуют дополнительных прорывов. Количество данных по речевой информации велико, и она обладает временной изменчивостью. Речевые шаблоны различны для разных людей, и даже для одного и того же человека они могут быть разными.

2. Проблемы мониторинга и обработки точек разрыва. Когда человек говорит, паузы между словами могут быть как долгими, так и короткими. Эта нечеткость обработки и является одним из основных факторов, влияющих на точность сопоставления слов.

3. Проблемы распознавания речи для неопределенных лиц. Чтобы интеллектуальные роботы действительно стали практичными и коммерчески успешными, необходимо решить проблему создания модели распознавания речи для неопределенных лиц. Это также одна из ключевых проблем для преодоления в области технологии распознавания речи.

4. Реальное время распознавания речи. Интеллектуальные голосовые роботы требуют возможности работы в реальном времени, поэтому системы распознавания речи должны удо-

влетворять требованиям по реальному времени.

5. Проблема обработки нечеткости в речи. Люди из разных регионов произносят звуки по-разному, а кроме того, в китайском языке множество иероглифов многозначно. При произнесении слов существует неопределенность в звучании, поэтому при распознавании речи необходимо ориентироваться на контекст, лексические и синтаксические правила, чтобы обеспечить точность распознавания речи.

Заключение

Исследования по рассмотренной теме имеют следующие основные аспекты:

1) тема исследования касается обеих сторон – аппаратного и программного обеспечения и требует объединения технологий управления роботом и распознавания речи. Это одна из актуальных тем исследования, имеющая не только теоретическую, но и сильную практическую ценность;

2) распознавание речи – комплексное применение технологии, пересекающей несколько дисциплин (информатика, инженерия, лингвистика, психология и др.). Достижение беспрепятственного языкового взаимодействия между роботами и людьми всегда было мечтой человечества, а технология распознавания речи – ключ к этой мечте, вот почему данная тема является одной из важных областей научных исследований;

3) выбор этой темы – стимулирование интереса молодежи, позволяющее привлечь отдельных студентов к исследованиям в области интеллектуальных роботов, в том числе к изучению систем распознавания речи для роботов, что впоследствии может привести к формированию исследовательской группы по интеллектуальным роботам.

1. Rabiner, L. Переход к видению. Рассмотрение звука и аудиообработки[J] // Технический отдел AT&T, 2012. – №4.

2. Jelinek, F. Непрерывное распознавание речи методами статистики[J] // Труды IEEE. – 2013. – №11.

3. Ян Си. Состояние и тенденции развития искусственного интеллекта за рубежом[J] // Распознавание образов и искусственный интеллект. – 2015. – №2.

4. Чжу Шичжан. Текущее состояние и тенденции развития технологии тактильного восприятия роботов[J] // Журнал Ханчжоу электронной промышленности. – 2013. – №8.

5. Сунь Ие. Текущее состояние и тенденции развития искусственного интеллекта[J] // Инженерия стоимости. – 2015. – №6.

6. Сун Янань. Дизайн и реализация эксперимента по распознаванию речи для робота[J] // Компьютерная технология и развитие. – 2013. – №4.

7. Ян Сюэсонг. Исследование системы распознавания речи для робота с использованием микроконтроллера[D] – Пекин: Пекинский университет. – 2011.

8. Чен Сун. Проектирование робота с распознаванием речи на основе микроконтроллера[J] // Робототехника. – 2012. – №4.

9. Лю Гуйбин. Исследование системы распознавания речи для встроенного робота[D]. – Китайский нефтяной университет. – 2014.