## VR TECHNOLOGIES AS A TOOL FOR DEVELOPING DIGITAL SKILLS

***Lazutin Artem Antonovich,***
*Master's Student, Department of Economics and Finance*
*Tyumen State University, Tyumen, Russia,*
***Supervisor: Yumanova Natalia Nikolaevna,***
*Candidate of Economic Sciences, Associate Professor,*
*Department of Economics and Finance, Faculty of Digital Economy,*
*Tyumen State University, Tyumen, Russia*

*Annotation.* The article proposes a solution to the problem of personnel shortage in industry through the implementation of a VR platform for career guidance. It describes the relevance of strategically demanded professions and examines the causes of staff shortages, such as the popularization of intellectual professions and stereotypes about factory work as physically demanding and low-paid. The perception of industrial occupations by young people as outdated and unpromising due to insufficient awareness is analyzed.

*Key words:* personnel shortage, virtual reality, VR technologies, education, industry.

## КАК ТЕХНОЛОГИИ МЕНЯЮТ ПОДХОДЫ К РАСПОЗНАВАНИЮ И СИНТЕЗУ ЖИВОЙ РЕЧИ

***Лашин Денис Андреевич,***
*аспирант кафедры теоретической и прикладной лингвистики*
*факультета лингвистики и межкультурной коммуникации*
*Юго-Западного государственного университета, г. Курск, Россия*

*Аннотация.* В статье рассматриваются этапы развития, методы и достижения технологий распознавания и синтеза живой речи. Выявляется ряд проблем и сложностей, с которыми сталкиваются исследователи в сфере технологий распознавания и синтеза речи. Намечаются перспективы и будущее развитие технологий распознавания и синтеза живой речи.

*Ключевые слова:* распознавание живой речи, синтез живой речи, нейронные сети, акустические модели, искусственный интеллект.

### Введение

Технологии играют ключевую роль в эволюции человечества, влияя на все аспекты нашей жизни, включая общение. С развитием цифровых технологий и искусственного интеллекта традиционные методы взаимодействия между людьми претерпевают значительные изменения. Одним из наиболее востребованных направлений является область распознавания и синтеза живой речи. В данной статье мы рассмотрим, каким образом современные технологии трансформируют процессы восприятия и генерации человеческой речи, а также обсудим перспективы дальнейшего развития этих технологий.

### История развития технологий распознавания и синтеза речи

Распознавание речи – это процесс преобразования акустических сигналов, представляющих человеческую речь, в цифровой формат для последующего анализа и обработки [Mohri, 1997, с. 560]. Эта технология имеет долгую историю, начиная от первых попыток машинного перевода текста в середине XX века до современных систем, основанных на нейронных сетях и глубоком обучении.

Первые попытки автоматического распознавания речи были предприняты еще в 1950-х годах. Первые системы распознавания речи могли воспринимать лишь цифры, что вполне оправдано, учитывая сложность языка и необходимость сосредоточенности инженеров именно на этой задаче. Лаборатория Bell, являющаяся одной из самых известных исследовательских организаций в мире, основанная в США в 1925 году как совместное предприятие AT&T и Western Electric, сыграла важную роль в развитии телекоммуникаций, электроники и вычислительной техники. За свою историю лаборатория внесла значительный вклад в различные области науки и технологий, включая разработку транзисторов, лазеров, операционных систем Unix и C++, а также разработала систему под названием «Audrey», способную различать цифры, произносимые одним человеком. Спустя десять лет, в 1962 году, компания IBM представила свою разработку – устройство «Shoebox», которое могло распознавать уже 16 английских слов.

Исследовательские центры в США, Японии, Великобритании и Советском Союзе создали несколько аналогичных устройств, которые могли различать отдельные звуки речи, включая четыре гласных и девять согласных звуков. Хотя качество звучания оставляло желать лучшего, эти ранние эксперименты заложили основу для дальнейшего развития технологии, особенно если учесть ограниченные возможности компьютеров тех времен [Тампель, 2015, с. 958].

Современное развитие технологий распознавания речи и синтеза находится на этапе стремительного прогресса, что оказывает значительное влияние на различные отрасли и повседневную жизнь. Эти технологии трансформируют способы взаимодействия между людьми и машинами, делая их более удобными и эффективными.

Технологии распознавания речи сегодня основаны на мощных алгоритмах машинного обучения. Среди них: глубокие нейронные сети, которые позволяют точно интерпретировать речь даже в сложных условиях, таких как шумная среда. В результате появились такие популярные решения, к примеру голосовые ассистенты и смарт-колонки, которые обеспечивают удобство и комфорт в управлении домашними устройствами и поиске информации. Синтез речи также сделал огромный скачок вперед благодаря использованию продвинутых моделей, таких как WaveNet, которые создают речь, близкую к натуральной человеческой. Это открывает новые возможности в образовании, медицине и развлекательной индустрии, позволяя создавать интерактивные и доступные материалы. Эти технологии продолжают развиваться, решая проблемы, связанные с конфиденциальностью, точностью и адаптацией к разным языкам и культурам. Будущее обещает ещё больше инноваций, которые сделают взаимодействие с технологиями ещё более естественным и продуктивным.

**Основные методы распознавания и синтеза речи**

Основные методы распознавания речи представляют собой совокупность подходов и алгоритмов, направленных на преобразование акустических сигналов, соответствующих устной речи, в цифровую форму для последующей интерпретации и обработки. Ниже приведены ключевые методы, используемые в современных системах распознавания речи:

**1. Акустические модели**

Акустическая модель играет центральную роль в процессе распознавания речи, поскольку именно она осуществляет первичное преобразование звуковых волн в цифровые параметры, которые затем подвергаются дальнейшей обработке. Акустические модели строятся на основе обширных корпусов речевых данных, что позволяет им осуществлять эффективное распознавание сложных фонетических и просодических структур [Гаврилович, 2016, с. 206-207].

**2. Нейронные сети**

Одним из ключевых прорывов в развитии технологий распознавания речи стало использование нейронных сетей, особенно глубоких архитектур (Deep Neural Networks, DNN). Глубокие нейронные сети обладают способностью к самообучению на больших массивах данных, что делает их идеальными для обработки сложных речевых паттернов. Современные системы часто применяют рекуррентные нейронные сети (RNN) и их разновидности, такие как LSTM (Long Short-Term Memory) и GRU (Gated Recurrent Unit), чтобы учитывать временные зависимости в последовательностях речевого сигнала.

**3. Методы глубокого обучения**

Глубокое обучение основывается на многослойных нейронных архитектурах, которые позволяют извлекать высокоуровневые признаки из входных данных. В контексте распознавания речи глубокое обучение обеспечивает возможность построения моделей, которые способны адаптироваться к различным языкам, диалектам и стилям речи, обеспечивая высокую точность даже в условиях зашумленной среды.

**4. Контекстная обработка**

Важным элементом современных систем распознавания речи является способность учитывать контекст, что значительно повышает точность интерпретации высказываний. Контекстная обработка включает использование лингвистических знаний, таких как грамматика, семантика и прагматика, для улучшения понимания смысла произнесенных фраз. Это достигается путем интеграции языковых моделей, которые оценивают вероятность последовательности слов в заданном контексте.

**5. Методы обработки временных рядов**

Распознавание речи требует учета временной структуры сигнала, что обеспечивается методами обработки временных рядов. Временные ряды анализируются с использованием окон Фурье (FFT) и других спектральных преобразований, что позволяет выделять важные частотные компоненты и оценивать динамику изменения амплитуды и частоты речевого сигнала.

**6. Методы уменьшения размерности**

Для эффективного представления речевой информации используются методы снижения размерности, такие как Principal Component Analysis (PCA) и Linear Discriminant Analysis (LDA). Эти методы позволяют уменьшить избыточность данных, сохранив при этом основную информативность, что существенно ускоряет обработку и улучшает производительность системы.

**7. Байесовский подход**

Байесовская теория вероятности лежит в основе многих современных систем распознавания речи. Байесовский подход предполагает вычисление условных вероятностей для определения наиболее вероятной

последовательности символов или слов на основании имеющихся данных. Этот метод применяется как для оценки акустических моделей, так и для выбора оптимальной гипотезы при распознавании речи.

**8. Гибридные модели**

Современные системы распознавания речи часто интегрируют различные подходы, создавая гибридные модели, сочетающие преимущества нескольких методов. Например, сочетание акустических моделей и нейронных сетей позволяет добиться высокой точности распознавания в широком диапазоне условий [Гаврилович, 2016, с. 206-207].

Также существует несколько основных методов синтеза речи.

**1. Формантный синтез (Formant synthesis)**

Этот метод основан на синтезировании отдельных формант – характерных резонансных частот, присущих гласным звукам. В рамках данного подхода моделируется форма вокального тракта (гортань, рот, нос) и производятся соответствующие гармонические колебания. Формантный синтез относительно прост в реализации и требует небольших вычислительных мощностей, однако получаемая речь звучит механически и неестественно [Меджидов, 2024, с. 161].

**2. Конкатенационный синтез (Concatenative synthesis)**

Конкатенационный синтез базируется на объединении заранее записанных сегментов речи (обычно фонем, дифонов или полных слов). Такой подход позволяет достичь высокого уровня естественности звучания, поскольку используются реальные записи человеческой речи. Основной проблемой является необходимость иметь большой банк данных с фрагментами речи, а также трудности в обеспечении плавного перехода между сегментами.

**3. Модель на основе скрытых марковских процессов (Hidden Markov Model, HMM)**

HMM-модели позволяют моделировать временные и спектральные свойства речи. Каждая фонема представляется как последовательность состояний, переходы между которыми управляются вероятностными распределениями. Это позволяет учесть динамические характеристики речи, такие как скорость произнесения и интонационные особенности. HMM-системы требуют значительного объема тренировочных данных и часто нуждаются в пост-обработке для повышения качества звучания.

**4. Синтез на основе глубоких нейронных сетей (DNN-based synthesis)**

Этот метод использует глубокие нейронные сети для прямого предсказания параметров синтеза речи на основе текстового ввода. Глубокие нейронные сети могут быть обучены на большом количестве данных и способны учитывать разнообразные аспекты речи, такие как интонация, темп и эмоциональная окраска. Данный подход демонстрирует высокий уровень реалистичности и гибкости, хотя и требует значительных вычислительных ресурсов.

**5. Генеративные состязательные сети (GANs)**

Генеративные состязательные сети (Generative Adversarial Networks) используются для создания синтетической речи, близкой к настоящей. Модель генератора создает образцы речи, а дискриминатор оценивает их реалистичность. Постоянное соревнование между этими двумя компонентами

приводит к улучшению качества генерируемой речи. GANs позволяют получать результаты, близкие к естественным, но требуют тщательной настройки и значительного объема обучающих данных [Меджидов, 2024].

**Современные технологии распознавания речи**

В настоящее время технологии распознавания речи находят всё большее число применений, начиная от программ, переводящих речь в текст, и заканчивая системами управления автомобилями. Далее приводится (почему именно эти объяснить).

1. Компания SPIRIT разрабатывает программное обеспечение для цифровой телефонии, компрессии речи, определения личности говорящего, а также для технологий VoIP и GPS. Движок распознавания речи (ASR engine) от SPIRIT предназначен для восприятия голосовых команд и используется в различных приложениях, таких как управление устройствами с помощью голоса, голосовой ввод на устройствах hands-free и ввод PIN-кодов в системах безопасности. Ядро легко интегрируется в любые платформы на базе DSP или RISC и предоставляется в форме объектного кода [Титов, 2006, с. 571-573].

2. Распознавание речи находит применение не только для набора текста или выполнения команд, но и для специализированных задач. Компания «Центр Речевых Технологий» создает и выпускает программные продукты, технологии и технические средства для нужд различных государственных структур, таких как МВД, ФСБ, Министерство юстиции, МЧС, Минобороны, службы экстренной помощи, контактные центры и другие организации, которым важна регистрация и обработка речевой информации [Титов, 2006, с. 571-573].

Среди разработок компании выделяются следующие приложения: «ИКАР Лаб» – инструментальный комплекс для криминалистического анализа звукозаписей речи, «Трал» – автоматизированная система для идентификации дикторов в записях телефонных разговоров, а также «Территория» – система для автоматической диагностики диалектов и акцентов в устной русской речи.

3. Для внедрения функций распознавания речи в разнообразные устройства, роботов и игрушки создаются специальные аппаратные решения. Американская компания Sensory Inc. разработала микросхему Voice Direct™ 364, которая способна распознавать ограниченное количество команд (до 60) после предварительной настройки. Перед использованием модуля его нужно обучить всем необходимым командам. Эти команды сохраняются во внешней памяти в виде образов объемом 128 байт. В процессе работы новый образ команды сравнивается с эталонами, хранящимися в памяти, и на основании этого делается вывод о совпадении [Мырадов, 2024, с. 301-305].

Это лишь несколько примеров использования технологий распознавания и синтеза речи, которые используются в современном мире для выполнения широкого спектра задач, однако этот ряд примеров уже привнес в современный мир множество инноваций.

**Проблемы и сложности в распознавании речи**

Вместе с тем существует внушительное количество сложностей, с которыми может столкнуться специалист при работе с системами распознавания и синтеза речи.

1. Проблема устойчивости системы распознавания речи к помехам требует решения по двум ключевым направлениям. Во-первых, необходимы меры по устранению шумов, помех и искажений, влияющих на речевой сигнал. Во-вторых, учитывая ограниченные возможности очистки речевых сигналов от шума на практике, возникает потребность в разработке более эффективных методов выделения полезной речевой информации из зашумленной акустической среды [Мин Зо, 2018, с. 28-32].

2. Современные системы распознавания речи часто заявляются как независимые от диктора, они способны распознавать отдельные слова из ограниченного словаря объемом до 500 слов с точностью около 95%. Тем не менее, обработка и распознавание непрерывной речи зачастую требуют длительной и сложной адаптации системы под конкретного говорящего.

3. Проблемы надежности распознавания отдельных речевых фрагментов. Несмотря на то, что системы распознавания речи в настоящее время все более совершенствуются, ни одну из них нельзя считать на данный момент совершенной, единственно точной и дающей наиболее оптимальный результат [Мин Зо, 2018, с. 28-32].

**Заключение**

В данной статье были рассмотрены ключевые аспекты влияния современных технологий на подходы к распознаванию и синтезу живой речи. Проанализированы эволюцию алгоритмов и методов, применяемых в этой области, начиная от ранних попыток машинного понимания речи и заканчивая новейшими достижениями в области нейронных сетей и глубокого обучения. Эти инновационные технологии позволяют значительно повысить точность распознавания речи, улучшить качество синтеза и адаптироваться к различным языкам и диалектам. Однако, несмотря на значительные успехи, перед исследователями всё ещё стоят важные вызовы. Среди них – необходимость повышения устойчивости систем к шумовым воздействиям, улучшение дикторонезависимого распознавания и повышение надёжности идентификации отдельных речевых фрагментов. Решение этих проблем открывает путь к созданию ещё более мощных и универсальных инструментов для работы с речью. Таким образом, развитие технологий не только меняет подходы к распознаванию и синтезу речи, но и способствует трансформации многих областей человеческой деятельности, включая образование, медицину, промышленность и повседневное взаимодействие людей с цифровыми устройствами.

## Список литературы

Гаврилович Н.В. Методы распознавания речи и их классификация / Н.В. Гаврилович, С.Н. Сейтвелиева // Таврический научный обозреватель. – 2016. – №6 (11). – С. 206-212.

Меджидов М.М. Современный подход к синтезу речи / М.М. Меджидов, Г.М. Ковалёв // Научный аспект. – 2024. – №6. – С. 161.

Мырадов М.Т. Современные технологии распознавания речи / М.Т. Мырадов, С.Н. Назарова // Наука и мировоззрение. – 2024. – №23. – С. 301-305.

Тампель И.Б. Автоматическое распознавание речи – основные этапы за 50 лет // Научно-технический вестник информационных технологий, механики и оптики. – 2015. – Т. 15. – № 6. – С. 957–968.

Титов Ю.Н. Современные технологии распознавания речи // Вестник российских университетов. Математика. – 2006. – №4. – С. 571-574.

Хеин Мин Зо Современное состояние проблемы обработки, анализа и синтеза речевых сигналов // Computational nanotechnology. – 2018. – №2. – С. 28-32.

Mohri M. Speech recognition with weighted finite-state transducers: In Springer Handbook of Speech Processing / M. Mohri, F. Pereira, M. Riley. – Pittsburg, AT$T Labs – Research, 1997. – 560 p.

## HOW TECHNOLOGIES ARE CHANGING APPROACHES TO RECOGNITION AND SYNTHESIS OF LIVE SPEECH

***Lashin Denis Andreevich,***
*Postgraduate student, the department of theoretical and applied linguistics*
*Southwest State University, Kursk, Russia*

*Annotation.* The article discusses the stages of development, approaches and achievements of technologies for recognition and synthesis of live speech. It identifies a number of problems and difficulties faced by researchers in the field of speech recognition and synthesis technologies. The prospects and future development of technologies for recognition and synthesis of live speech are outlined.

*Key words:* live speech recognition, live speech synthesis, neural networks, acoustic models, artificial intelligence.

## ПСИХОЛОГО-ПЕДАГОГИЧЕСКИЕ ОСНОВЫ ЯЗЫКОВОГО ОБРАЗОВАНИЯ

***Логвиненко Дарья Григорьевна,***
*студент кафедры дошкольного и специального (дефектологического) образования*
*факультета дошкольного начального и специального образования*
*педагогического института*
*НИУ «БелГУ», г. Белгород, Россия*
***научный руководитель: Смирнова Дарья Евгеньевна,***
*старший преподаватель кафедры иностранных языков*
*факультета иностранных языков*
*педагогического института НИУ «БелГУ», г. Белгород, Россия*

*Аннотация.* В статье рассматриваются психологические и педагогические основы обучения языкам. В ней анализируются разные подходы к обучению, включая бихевиористский, конструктивистский и коммуникативный. Также обсуждается, как на процесс изучения языка влияют мотивация, эмоциональный интеллект и возрастные особенности учащихся. Особое внимание уделяется современным технологиям и интерактивным методам, таким как игровые техники, проектная работа, а также значению обратной связи и самооценки в обучении.

*Ключевые слова:* психологические и педагогические основы, обучение языкам, бихевиоризм, конструктивизм, коммуникативный подход, методы обучения.