

УДК 004.934

ОПТИМИЗАЦИЯ РАСПОЗНАВАНИЯ РЕЧИ НА УСТРОЙСТВАХ С ОГРАНИЧЕННОЙ ВЫЧИСЛИТЕЛЬНОЙ МОЩНОСТЬЮ

Расулов Руслан Тахирович

Студент магистратуры

2 курс, факультет «Информационные технологии»

Кафедра «Математическая кибернетика и информационные технологии»

Московский технический университет связи и информатики

e-mail: r99r12t28@mail.ru

Басыров Ильдар Ильшатovich

Студент бакалавриата 4 курс, факультет «Сети и системы связи»

Кафедра «Сети связи и системы коммутации»

Московский технический университет связи и информатики

e-mail: basyrov.ildar@mail.ru

Богач Егор Владимирович

Студент бакалавриата 4 курс, факультет «Сети и системы связи»

Кафедра «Сети связи и системы коммутации»

Московский технический университет связи и информатики

e-mail: egorka.bogach.2001@mail.ru

Аннотация

Статья рассматривает проблему ограниченности вычислительной мощности устройств при распознавании речи и предлагает пути ее решения. В статье описываются методы и технологии, которые могут быть использованы для оптимизации процесса распознавания речи на устройствах с ограниченной вычислительной мощностью. Рассмотрены преимущества и недостатки каждого метода, а также эффективность на устройствах с ограниченной вычислительной мощностью. В результате проведенного анализа были предложены пути решения проблемы ограниченности вычислительной мощности, которые могут быть использованы для создания более эффективных и функциональных устройств распознавания речи.

Ключевые слова: распознавание речи, вычислительная мощность, устройства с ограниченной мощностью, оптимизация, методы и технологии, скрытые марковские модели, глубокое обучение, фильтрация и дискретизация сигнала, сжатие данных.

OPTIMIZING SPEECH RECOGNITION ON DEVICES WITH LIMITED COMPUTATIONAL POWER

Ruslan T. Rasulov

Master's degree student

2nd year, Faculty of "Networks and Communication Systems"
Department of "Communication Networks and Switching Systems"
Moscow Technical University of Communications and Informatics

Ildar I. Basyrov

Bachelor's student 4th year
Faculty of "Networks and Communication Systems"
Department of "Communication Networks and Switching Systems"
Moscow Technical University of Communications and Informatics

Egor V. Bogach

Bachelor's student 4th year
Faculty of "Networks and Communication Systems"
Department of "Communication Networks and Switching Systems"
Moscow Technical University of Communications and Informatics

ABSTRACT

The article addresses the problem of limited computational power of devices for speech recognition and proposes ways to solve it. The article describes methods and technologies that can be used to optimize the process of speech recognition on devices with limited computational power. The advantages and disadvantages of each method, as well as their efficiency on devices with limited computational power, are considered. As a result of the analysis, solutions to the problem of limited computational power have been proposed, which can be used to create more efficient and functional speech recognition devices.

Keywords: speech recognition, computational power, devices with limited computational power, optimization, methods and technologies, hidden Markov models, deep learning, signal filtering and discretization, data compression.

Введение

Распознавание речи – это процесс преобразования аудио-сигнала речи в текст. Современные системы распознавания речи используют различные алгоритмы и методы для достижения высокой точности и скорости распознавания. Однако, на устройствах с малой вычислительной мощностью, таких как смартфоны и микроконтроллеры, использование стандартных алгоритмов распознавания речи может быть невозможным из-за ограниченных вычислительных ресурсов [1].

Целью данной статьи является рассмотрение основных проблем, связанных с разработкой систем распознавания речи на устройствах с малой вычислительной мощностью, и предложение путей их решения.

Теоретическая часть

Основными методами распознавания речи в современных системах являются скрытые марковские модели, нейронные сети и глубокое обучение. Стандартные алгоритмы распознавания речи могут работать на вычислительно мощных серверах, но на

устройствах с малой вычислительной мощностью они могут работать слишком медленно или вообще невозможно [2, 3].

Одним из основных принципов работы систем распознавания речи на устройствах с ограниченными ресурсами является использование методов обработки сигналов. Такие методы позволяют уменьшить объем данных, обрабатываемых при распознавании речи. Например, можно использовать методы фильтрации и дискретизации сигнала, которые позволят уменьшить объем данных и повысить скорость обработки. Также можно использовать алгоритмы сжатия данных, которые помогут уменьшить объем информации, передаваемой для обработки [4].

Для успешного распознавания речи необходимо выполнение нескольких этапов. Один из ключевых этапов – это предобработка аудио-сигнала речи, включающая в себя фильтрацию шумов, нормализацию громкости, преобразование временного сигнала в частотный спектр и другие операции. После предобработки аудио-сигнала следует этап извлечения признаков, на котором основывается последующее распознавание. В качестве признаков, используемых в распознавании речи, часто выступают мел-частотные кепстральные коэффициенты (MFCC), которые представляют собой спектральные характеристики речи на определенных временных интервалах [5].

Важно отметить, что распознавание речи является задачей многоклассовой классификации, так как каждое слово может быть распознано как один из нескольких возможных вариантов. Для этого могут применяться различные алгоритмы классификации, такие как метод k -ближайших соседей (k -NN), метод опорных векторов (SVM) и другие.

Одним из вызовов, связанных с разработкой систем распознавания речи на устройствах с малой вычислительной мощностью, является необходимость баланса между точностью распознавания и скоростью работы системы. Для достижения этого баланса могут применяться различные методы оптимизации, такие как оптимизация гиперпараметров, уменьшение размерности данных и другие [6] (табл. 1).

Таблица 1. Сравнение методов и технологий для оптимизации распознавания речи на устройствах с ограниченной вычислительной мощностью

Метод / технология	Описание	Преимущества	Эффективность на устройствах с ограниченной вычислительной мощностью
Метод фильтрации и дискретизации сигнала	Методы обработки сигнала, которые позволяют уменьшить объем данных и повысить скорость обработки.	Уменьшение объема данных, увеличение скорости обработки.	Эффективен при работе на устройствах с ограниченной вычислительной мощностью.
Метод сжатия данных	Алгоритмы сжатия данных, которые позволяют уменьшить объем информации, передаваемой для обработки.	Уменьшение объема данных.	Эффективен при работе на устройствах с ограниченной вычислительной мощностью.
Глубокое обучение	Метод машинного обучения, который позволяет обучать нейронные сети на больших объемах данных.	Высокая точность распознавания.	Требует высокой вычислительной мощности и большого объема данных для обучения, но может быть эффективен при использовании предобученных моделей.
Скрытые марковские модели	Модели, которые используются для описания последовательности событий во времени.	Эффективны при работе с короткими аудиофрагментами.	Требуют высокой вычислительной мощности для обучения и высокой точности выравнивания последовательностей.
Kaldi	Комплекс программных средств для распознавания речи, включающий в себя различные методы обработки сигналов и алгоритмы машинного обучения.	Хорошо документирован, поддерживает множество языков и гибок в настройке.	Может быть сложным в использовании и требовать высокой вычислительной мощности для работы.

Постановка проблемы

Одной из основных проблем, связанных с использованием современных технологий распознавания речи на устройствах с ограниченными ресурсами, является ограниченность вычислительной мощности устройства. Распознавание речи требует больших вычислительных ресурсов, поэтому на устройствах с малой вычислительной мощностью процесс может быть слишком медленным или вообще невозможным. Это ограничение

может ограничить возможности устройств, таких как голосовые помощники и голосовые устройства управления [7].

Однако, на устройствах с ограниченными ресурсами, таких как смартфоны, голосовые помощники и другие устройства IoT (Internet of Things), требования к вычислительным ресурсам значительно ниже. В связи с этим, необходимо разработать специализированные алгоритмы и технологии, которые бы могли обеспечивать высокую точность распознавания речи, не увеличивая при этом нагрузку на вычислительные ресурсы устройств.

В частности, ограниченность вычислительной мощности может снизить качество распознавания речи, так как для повышения точности необходимо проводить более сложные вычисления. Это может привести к ошибкам распознавания и снижению качества работы системы [8].

Кроме того, наличие ограниченности вычислительной мощности ограничивает возможности устройств, так как процесс распознавания речи может быть слишком медленным или вообще невозможным на устройствах с ограниченными ресурсами. Это может быть особенно проблематично для голосовых помощников и других устройств управления, которые должны быстро обрабатывать и отвечать на запросы пользователя.

В целом, проблема ограниченности вычислительной мощности устройств является серьезной проблемой, которая затрудняет разработку эффективных и точных систем распознавания речи на устройствах с ограниченными ресурсами.

Пути решения проблемы

Одним из решений для проблемы ограниченности вычислительной мощности может быть использование специализированных алгоритмов распознавания речи, которые были разработаны специально для работы на устройствах с ограниченными ресурсами. Такие алгоритмы могут быть менее точными, чем стандартные алгоритмы распознавания речи, но они могут работать быстрее и требовать меньше вычислительной мощности.

Другим решением может быть использование методов машинного обучения для улучшения точности распознавания речи на устройствах с малой вычислительной мощностью. Методы машинного обучения могут помочь улучшить точность распознавания речи и оптимизировать процесс обработки данных.

Также можно использовать гибридные подходы, которые комбинируют различные методы и технологии. Например, можно использовать методы обработки сигналов в сочетании с алгоритмами машинного обучения для достижения высокой точности и скорости распознавания речи на устройствах с ограниченными ресурсами.

Заключение

Разработка систем распознавания речи на устройствах с малой вычислительной мощностью является сложной задачей, но с использованием специализированных методов обработки сигналов, алгоритмов машинного обучения и гибридных подходов можно достичь хороших результатов. Важно учитывать ограничения вычислительной мощности устройства и выбирать технологии и методы, которые оптимизируют процесс распознавания речи и уменьшают нагрузку на устройство.

Дальнейшие исследования в этой области могут быть направлены на разработку новых методов и технологий для распознавания речи на устройствах с ограниченными ресурсами, а также на оптимизацию и улучшение существующих методов и технологий.

В заключении стоит подчеркнуть, что разработка систем распознавания речи на устройствах с малой вычислительной мощностью имеет большое значение для развития голосовых технологий и улучшения пользовательского опыта. Такие системы могут помочь

увеличить эффективность работы и повысить комфортность использования устройств с ограниченными ресурсами.

Список литературы:

1. Данильчук Татьяна Ивановна, Трофимов Иван Михайлович Обработка и анализ речевых фраз с использованием математических алгоритмов уровней сигналов // Baikal Research Journal. 2013. №6. URL: <https://cyberleninka.ru/article/n/obrabotka-i-analiz-rechevyh-fraz-s-ispolzovaniem-matematicheskikh-algoritmov-urovney-signalov> (дата обращения: 23.04.2023).
2. Садыкова А. А., Амиргалиев Е. Н. Изучение применения автоматического распознавания речи // Colloquium-journal. 2020. №11 (63). URL: <https://cyberleninka.ru/article/n/izuchenie-primeneniya-avtomaticheskogo-raspoznavaniya-rechi> (дата обращения: 23.04.2023).
3. Титов Ю. Н. Современные технологии распознавания речи // Вестник российских университетов. Математика. 2006. №4. URL: <https://cyberleninka.ru/article/n/sovremennye-tehnologii-raspoznavaniya-rechi> (дата обращения: 23.04.2023).
4. Пресняков Игорь Николаевич, Омельченко Анатолий Васильевич, Омельченко Сергей Васильевич Автоматическое распознавание речи в каналах передачи // Радиоэлектроника и информатика. 2002. №1 (18). URL: <https://cyberleninka.ru/article/n/avtomaticheskoe-raspoznavanie-rechi-v-kanalah-peredachi> (дата обращения: 23.04.2023).
5. Алимуратов Алан Казанферович Параметры и классификация систем распознавания речи // Модели, системы, сети в экономике, технике, природе и обществе. 2014. №1 (9). URL: <https://cyberleninka.ru/article/n/parametry-i-klassifikatsiya-sistem-raspoznavaniya-rechi> (дата обращения: 23.04.2023).
6. Малиновкин В.А., Валуйских Н.В., Шведов Н.Н., Кенин С.Л., Гребенникова Н.И. СРАВНИТЕЛЬНЫЙ АНАЛИЗ СРЕДСТВ ГОЛОСОВОГО ИНТЕРФЕЙСА И ТЕХНОЛОГИЙ РАСПОЗНАВАНИЯ РЕЧИ // Вестник ВГТУ. 2022. №1. URL: <https://cyberleninka.ru/article/n/sravnitelnyy-analiz-sredstv-golosovogo-interfeysa-i-tehnologiy-raspoznavaniya-rechi> (дата обращения: 23.04.2023).
7. Егунов Виталий Алексеевич, Панюлайтис Станислав Владимирович РАСПОЗНАВАНИЕ РЕЧЕВЫХ КОМАНД С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ // Прикаспийский журнал: управление и высокие технологии. 2020. №3 (51). URL: <https://cyberleninka.ru/article/n/raspoznavanie-rechevyh-komand-s-ispolzovaniem-neyronnyh-setey> (дата обращения: 23.04.2023).
8. Дьячковская Мария Андреевна, Протодяконова Галина Юрьевна ИССЛЕДОВАНИЕ РАСПОЗНАВАНИЯ ЧЕЛОВЕЧЕСКОГО ГОЛОСА С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ // Автоматика и программная инженерия. 2016. №4 (18). URL: <https://cyberleninka.ru/article/n/issledovanie-raspoznavaniya-chelovecheskogo-golosa-s-pomoschyu-neyronnyh-setey> (дата обращения: 23.04.2023).

References:

1. Danilchuk Tatyana Ivanovna, Trofimov Ivan Mikhailovich Processing and analysis of speech phrases using mathematical algorithms of signal levels // Baikal Research Journal. 2013. No. 6. URL: <https://cyberleninka.ru/article/n/obrabotka-i-analiz-rechevyh-fraz-s-ispolzovaniem-matematicheskikh-algoritmov-urovney-signalov> (date of access: 04/23/2023).
2. Sadykova A. A., Amirgaliev E. N. Studying the application of automatic speech recognition // Colloquium-journal. 2020. No. 11 (63). URL: <https://cyberleninka.ru/article/n/izuchenie-primeneniya-avtomaticheskogo-raspoznavaniya-rechi> (date of access: 04/23/2023).
3. Titov Yu. N. Modern speech recognition technologies // Bulletin of Russian Universities. Mathematics. 2006. No. 4. URL: <https://cyberleninka.ru/article/n/sovremennye-tehnologii-raspoznavaniya-rechi> (date of access: 04/23/2023).
4. Presnyakov Igor Nikolaevich, Omelchenko Anatoly Vasilievich, Omelchenko Sergey Vasilievich Automatic speech recognition in transmission channels // Radioelectronics and informatics. 2002. No. 1 (18). URL: <https://cyberleninka.ru/article/n/avtomaticheskoe-raspoznavanie-rechi-v-kanalah-peredachi> (date of access: 04/23/2023).
5. Alimuradov Alan Kazanferovich Parameters and classification of speech recognition systems // Models, systems, networks in economics, technology, nature and society. 2014. No. 1 (9). URL: <https://cyberleninka.ru/article/n/parametry-i-klassifikatsiya-sistem-raspoznavaniya-rechi> (date of access: 04/23/2023).
6. Malinovkin V.A., Valuyskikh N.V., Shvedov N.N., Kenin S.L., Grebennikova N.I. COMPARATIVE ANALYSIS OF VOICE INTERFACE AND SPEECH RECOGNITION TECHNOLOGIES // Vestnik VGTU. 2022. №1. URL: <https://cyberleninka.ru/article/n/sravnitelnyy-analiz-sredstv-golosovogo-interfeysa-i-tehnologiy-raspoznavaniya-rechi> (date of access: 04/23/2023).
7. Yegunov Vitaly Alekseevich, Panyulaitis Stanislav Vladimirovich RECOGNITION OF SPEECH COMMAND USING NEURAL NETWORKS // Caspian journal: management and high technologies. 2020. No. 3 (51). URL: <https://cyberleninka.ru/article/n/raspoznavanie-rechevyh-komand-s-ispolzovaniem-neyronnyh-setey> (date of access: 04/23/2023).
8. Dyachkovskaya Maria Andreevna, Protodyakonova Galina Yuryevna RESEARCH OF HUMAN VOICE RECOGNITION WITH THE HELP OF NEURAL NETWORKS // Automation and software engineering. 2016. No. 4 (18). URL: <https://cyberleninka.ru/article/n/issledovanie-raspoznavaniya-chelovecheskogo-golosa-s-pomoschyu-neyronnyh-setey> (date of access: 04/23/2023).