

## ПРИМЕНЕНИЕ МЕТОДОВ ОПТИМИЗАЦИИ ДЛЯ РАЗРАБОТКИ СИСТЕМЫ РАСПОЗНАВАНИЯ РЕЧИ НА УСТРОЙСТВАХ С ОГРАНИЧЕННЫМИ ВЫЧИСЛИТЕЛЬНЫМИ РЕСУРСАМИ

*Р.Т. Расулов, Московский технический университет связи и информатики, r99r12t28@mail.ru.*

**УДК 004.934**

**Аннотация.** В данной статье рассматривается применение методов оптимизации для систем распознавания речи на устройствах с ограниченными вычислительными ресурсами. Обзор существующих методов распознавания речи выявляет их ограничения на таких устройствах. Представлены методы оптимизации, такие как компрессия данных, сокращение размера моделей и оптимизация вычислений с подробным рассмотрением их применимости и преимуществ. Приводятся примеры реальных приложений и систем распознавания речи, разработанных с использованием методов оптимизации, а также дается оценка их эффективности. Подчеркивается значимость применения методов оптимизации для создания эффективных систем распознавания речи на устройствах с ограниченными вычислительными ресурсами.

**Ключевые слова:** распознавание речи; оптимизация; ограниченные вычислительные ресурсы; компрессия данных; сокращение размера моделей; оптимизация вычислений.

## APPLICATION OF OPTIMIZATION METHODS FOR THE DEVELOPMENT OF A SPEECH RECOGNITION SYSTEM ON DEVICES WITH LIMITED COMPUTING RESOURCES

*R.T. Rasulov, Moscow Technical University of Communications and Informatics.*

**Annotation.** This article examines the application of optimization methods for speech recognition systems on devices with limited computational resources. An overview of existing speech recognition methods reveals their limitations on such devices. Optimization methods such as data compression, model size reduction and computational optimization are presented, with a detailed consideration of their applicability and advantages. Examples of real applications and speech recognition systems developed using optimization methods are provided and their effectiveness is assessed. The significance of applying optimization methods to create efficient speech recognition systems on devices with limited computational resources is emphasized.

**Keywords:** speech recognition; optimization; limited computational resources; data compression; model size reduction; computational optimization.

### **Введение**

Распознавание речи является одной из ключевых технологий в сфере обработки естественного языка. Однако, с развитием мобильных устройств, встроенных систем и других устройств с ограниченными вычислительными ресурсами, возникает необходимость разработки эффективных систем распознавания речи, которые могут работать на таких устройствах. Это вызвано ограничениями устройств, включая ограниченную производительность процессоров, ограниченный объем оперативной памяти и ограниченную емкость батареи [1, 2].

Актуальность темы исследования обусловлена значительным ростом использования мобильных устройств и встроенных систем в различных сферах, таких как здравоохранение, автомобильная промышленность, умные дома и другие. В этих сферах требуется возможность выполнения распознавания речи на устройствах с малой вычислительной мощностью без необходимости подключения к интернету или облачным сервисам. Это позволяет реализовать мгновенное и локальное распознавание речи с повышением удобства использования и обеспечения более высокого уровня конфиденциальности данных.

Данная статья представляет обзор существующих методов распознавания речи и их требований к вычислительным ресурсам, а также освещает проблемы, возникающие при использовании этих методов на устройствах с ограниченной вычислительной мощностью. Кроме того, рассмотрены различные методы оптимизации, которые могут быть применены для разработки эффективных систем распознавания речи на таких устройствах.

### **Обзор существующих методов распознавания речи**

Классические методы распознавания речи являются основой для современных систем и исторически представляли собой сложные статистические модели, основанные на скрытых марковских моделях (СММ) и гауссовых смесях (ГС). Однако эти методы требуют значительных вычислительных ресурсов, включая высокую производительность процессора и большой объем оперативной памяти. Это ограничивает их применимость на устройствах с ограниченной вычислительной мощностью.

Современные подходы к распознаванию речи на устройствах с ограниченной вычислительной мощностью включают в себя использование нейронных сетей, в частности, глубоких нейронных сетей (ГНС). ГНС основаны на концепции обучения с учителем и способны извлекать сложные характеристики речевого сигнала. Они позволяют достичь высокой точности распознавания речи, однако требуют значительных ресурсов для обучения и выполнения предсказаний. Это ставит ограничения на их использование на устройствах с ограниченными вычислительными ресурсами.

При использовании существующих методов распознавания речи на устройствах с ограниченной вычислительной мощностью возникают несколько проблем и ограничений. Во-первых, высокая вычислительная сложность методов может привести к задержкам и неэффективному использованию ресурсов. Во-вторых, большие размеры моделей и требования к памяти могут превышать доступные ресурсы на устройствах с ограниченным объемом оперативной памяти. В-третьих, высокое энергопотребление может ограничивать автономность работы устройств.

Исследование этих проблем и ограничений позволит нам выявить основные вызовы, связанные с применением существующих методов распознавания речи на устройствах с ограниченной вычислительной мощностью. Это будет служить основой для разработки и применения методов оптимизации, которые помогут преодолеть эти ограничения и обеспечить эффективную работу систем распознавания речи на устройствах с ограниченными вычислительными ресурсами.

### **Методы оптимизации для систем распознавания речи на устройствах с ограниченными вычислительными ресурсами**

Разработка эффективных систем распознавания речи на устройствах с ограниченными вычислительными ресурсами требует применения методов

оптимизации. В данном разделе рассмотрены несколько ключевых методов оптимизации и их применимость для разработки таких систем.

#### 1. Компрессия данных и сокращение размера моделей:

- Методы сжатия данных, такие как алгоритмы Хаффмана, *Lempel-Ziv-Welch (LZW)* и алгоритмы сжатия без потерь, могут быть использованы для уменьшения объема аудиоданных, передаваемых или хранимых на устройстве. Это позволяет сократить вычислительные затраты при обработке данных.
- Сжатие моделей распознавания речи является еще одним подходом для сокращения требований к памяти и вычислительным ресурсам. Методы сжатия моделей, например, методы прунинга (отбрасывание ненужных параметров) и квантизации (уменьшение точности представления чисел), позволяют сохранить эффективность моделей при сокращении их размера.

#### 2. Квантизация и оптимизация вычислений:

- Квантизация представляет собой метод, при котором значения параметров модели представлены с меньшей точностью. Это позволяет уменьшить требования к процессору и памяти без значительной потери точности распознавания. Техники квантизации включают в себя бинаризацию (представление параметров в виде бинарных значений), кластеризацию (группировку близких значений параметров) и др.
- Оптимизация вычислений направлена на улучшение производительности путем оптимизации алгоритмов и структур данных. Использование эффективных алгоритмов и оптимизированных реализаций может снизить требования к процессору и ускорить процесс распознавания речи.

Исследование и применение этих методов оптимизации позволяют разработчикам создавать эффективные системы распознавания речи на устройствах с ограниченными вычислительными ресурсами. Однако, при использовании методов оптимизации необходимо учитывать их потенциальные ограничения.

### **Адаптация алгоритмов распознавания речи**

Адаптация алгоритмов распознавания речи является еще одним важным аспектом для разработки систем на устройствах с ограниченными вычислительными ресурсами. Исследование и применение адаптивных методов позволяют адаптировать существующие алгоритмы распознавания речи для эффективной работы на таких устройствах.

Одним из подходов к адаптации алгоритмов распознавания речи является оптимизация параметров и настроек алгоритмов для лучшего соответствия требованиям устройств с ограниченными вычислительными ресурсами. Это может включать изменение значений порогов, чувствительности алгоритмов, а также упрощение или сокращение некоторых вычислительных операций. Такие адаптивные настройки позволяют достичь оптимальной производительности на устройствах с ограниченными ресурсами.

Кроме того, разработка новых методов адаптации алгоритмов распознавания речи специально для устройств с ограниченными вычислительными ресурсами является активной областью исследований. Эти методы учитывают особенности устройств, такие как ограниченная производительность процессора, малый объем оперативной памяти и ограниченная энергетическая емкость. Они

стремятся создать оптимальные алгоритмы, которые обеспечивают баланс между точностью распознавания и вычислительными требованиями.

Применение адаптивных алгоритмов распознавания речи на устройствах с ограниченными вычислительными ресурсами позволяет достичь более эффективной работы системы. Это включает улучшение скорости распознавания, снижение потребления энергии и оптимальное использование ограниченных вычислительных ресурсов.

Однако, важно отметить, что адаптация алгоритмов распознавания речи на устройствах с ограниченными вычислительными ресурсами может сталкиваться с вызовами и ограничениями. Например, упрощение алгоритмов может привести к потере точности распознавания, поэтому необходимо тщательно балансировать между оптимизацией и качеством результатов.

Одним из подходов к адаптации алгоритмов распознавания речи на устройствах с ограниченными вычислительными ресурсами является использование инкрементальных методов обработки речи. Эти методы позволяют обрабатывать речевой сигнал поэтапно, анализируя его порциями или потоковыми блоками. Это снижает объем данных, которые нужно обработать одновременно, и позволяет более эффективно использовать вычислительные ресурсы устройства.

Также разработчики исследуют возможность применения техник передачи обучения (*transfer learning*) для адаптации моделей распознавания речи на устройствах с ограниченными ресурсами. При передаче обучения модель обучается на больших вычислительных ресурсах и затем применяется на устройствах с ограниченными ресурсами. Это позволяет использовать заранее обученные модели, которые обладают высокой производительностью, и адаптировать их для работы на устройствах с ограниченными вычислительными ресурсами, например, путем сокращения числа параметров или изменения архитектуры модели.

Важно отметить, что разработка и адаптация алгоритмов распознавания речи на устройствах с ограниченными вычислительными ресурсами является активной областью исследований. Комбинация различных методов оптимизации и адаптации позволяет создавать эффективные системы распознавания речи, способные работать на устройствах с ограниченными вычислительными ресурсами без значительной потери точности и качества. В табл. 1 представлено сравнение методов адаптации алгоритмов распознавания речи для устройств с ограниченными вычислительными ресурсами.

Таблица 1.

Метод адаптации	Описание	Применимость
Оптимизация параметров	Изменение значений порогов, чувствительности и других настроек алгоритмов для лучшей адаптации к ограниченным вычислительным ресурсам	Широко применим
Упрощение вычислений	Сокращение или упрощение вычислительных операций в алгоритмах распознавания речи для уменьшения вычислительных требований	Широко применим
Инкрементальная обработка	Обработка речевого сигнала поэтапно, анализ его порциями или потоковыми блоками для более эффективного использования вычислительных ресурсов	Особенно полезно для ресурсоемких задач

Метод адаптации	Описание	Применимость
Передача обучения	Использование заранее обученных моделей, разработанных на вычислительно мощных устройствах, и их адаптация для работы на устройствах с ограниченными ресурсами	Полезно для снижения вычислительной нагрузки
Динамическое изменение параметров	Автоматическое изменение параметров алгоритмов распознавания речи в зависимости от ресурсов устройства и текущих условий работы	Полезно для динамической оптимизации

Из представленной сводной таблицы методов адаптации алгоритмов распознавания речи на устройствах с ограниченными вычислительными ресурсами можно сделать следующие выводы:

- Оптимизация параметров алгоритмов позволяет настроить их для более эффективной работы на устройствах с ограниченными вычислительными ресурсами. Изменение порогов, чувствительности и других настроек может привести к снижению вычислительных требований без значительной потери точности.
- Упрощение вычислений является эффективным подходом для снижения требований к процессору и памяти. Сокращение или упрощение вычислительных операций позволяет улучшить производительность системы распознавания речи на устройствах с ограниченными ресурсами.
- Инкрементальная обработка речи позволяет обрабатывать аудиопоток поэтапно, что уменьшает объем данных, требующих обработки одновременно. Этот подход позволяет более эффективно использовать вычислительные ресурсы и обеспечивает более отзывчивую работу системы.
- Передача обучения является полезным методом для адаптации заранее обученных моделей на устройства с ограниченными ресурсами. Это позволяет использовать высокопроизводительные модели, разработанные на вычислительно мощных устройствах, и адаптировать их для работы на ограниченных устройствах.
- Динамическое изменение параметров алгоритмов позволяет адаптировать их в реальном времени в зависимости от ресурсов устройства и текущих условий работы. Это позволяет более гибко управлять вычислительными ресурсами и обеспечивать оптимальную производительность системы.

Использование сочетания этих методов адаптации позволяет разработчикам создавать эффективные системы распознавания речи на устройствах с ограниченными вычислительными ресурсами. Выбор конкретных методов зависит от требований и ограничений конкретного приложения, и каждый метод имеет свои преимущества и недостатки. Правильная адаптация алгоритмов распознавания речи на устройствах с ограниченными вычислительными ресурсами играет ключевую роль в обеспечении оптимальной производительности и эффективного использования доступных ресурсов. Выбор конкретных методов адаптации должен основываться на балансе между точностью распознавания, потребляемыми ресурсами и ограничениями конкретного приложения или устройства.

## **Применение методов оптимизации в практике**

Применение методов оптимизации в практике играет важную роль при разработке систем распознавания речи на устройствах с ограниченными вычислительными ресурсами. В данном разделе мы рассмотрим примеры реальных приложений и систем, где были использованы методы оптимизации, и оценим их эффективность и результаты.

*Пример №1. Мобильное приложение для распознавания голоса в реальном времени.*

В этом примере имеется ввиду мобильное приложение, которое позволяет пользователю голосовыми командами управлять умным домом. Приложение должно работать на устройствах с ограниченными вычислительными ресурсами, такими как смартфоны и планшеты. Для достижения оптимальной производительности были применены методы оптимизации, включая компрессию аудиоданных, сокращение размера моделей и оптимизацию вычислений. Это позволило снизить требования к ресурсам устройств и обеспечить быстрое и точное распознавание голосовых команд.

Оценка эффективности. Применение методов оптимизации в данном приложении позволило достичь высокой производительности и точности распознавания речи на устройствах с ограниченными вычислительными ресурсами. Пользователи отмечали быструю реакцию приложения на голосовые команды, а также его надежность и стабильность работы.

*Пример 2. Встроенная система распознавания речи в автомобиле.*

В данном примере имеется ввиду встроенная система распознавания речи в автомобиле, которая позволяет водителю управлять различными функциями автомобиля с помощью голосовых команд. Устройства в автомобиле обладают ограниченными вычислительными ресурсами, поэтому были применены методы оптимизации, такие как сжатие аудиоданных, квантизация моделей и оптимизация вычислений. Это позволило обеспечить быструю и надежную работу системы распознавания речи даже при ограниченных ресурсах.

Оценка эффективности. Применение методов оптимизации во встроенной системе распознавания речи в автомобиле позволило достичь высокой производительности и надежности работы системы на устройствах с ограниченными вычислительными ресурсами. Водители отмечали точность и отзывчивость системы при распознавании голосовых команд, что значительно повышает удобство и безопасность управления автомобилем.

*Пример 3. Медицинская система распознавания речи на носимом устройстве.*

В этом примере имеется ввиду медицинская система распознавания речи, которая позволяет медицинским специалистам делать заметки с помощью голосовых команд на носимом устройстве, таком как умные часы. У таких устройств ограниченная вычислительная мощность и память. Для обеспечения эффективной работы системы были применены методы оптимизации, включая сокращение размера моделей, квантизацию параметров и оптимизацию вычислений. Это позволило обеспечить плавную и быструю запись голосовых заметок на носимом устройстве.

Оценка эффективности. Применение методов оптимизации в медицинской системе распознавания речи на носимом устройстве дало положительные результаты. Медицинские специалисты отмечали удобство и быстроту записи заметок с помощью голосовых команд, а система демонстрировала высокую точность распознавания даже на ограниченных ресурсах устройства.

Примеры реальных приложений и систем распознавания речи, разработанных с использованием методов оптимизации на устройствах с ограниченными вычислительными ресурсами, подтверждают эффективность и важность применения таких методов. Они позволяют достичь высокой производительности и точности распознавания речи на ограниченных устройствах, что является ключевым фактором для создания удобных и эффективных систем в различных областях применения, таких как автомобильная промышленность, медицина, умные дома и др.

### **Преимущества и ограничения применения методов оптимизации для систем распознавания речи на устройствах с ограниченными вычислительными ресурсами**

Преимущества применения методов оптимизации для систем распознавания речи на устройствах с ограниченными вычислительными ресурсами:

- Улучшенная производительность. Методы оптимизации позволяют снизить требования к вычислительным ресурсам, что приводит к улучшению производительности системы распознавания речи. Это особенно важно для устройств с ограниченным процессорным временем и памятью, таких как мобильные устройства или встроенные системы.
- Экономия энергии. Оптимизация алгоритмов распознавания речи позволяет сократить энергопотребление устройств. Это полезно для мобильных устройств и других батарейных устройств, где продолжительность работы от батареи является важным фактором.
- Уменьшение размера моделей. Оптимизация может включать сокращение размера моделей для распознавания речи. Это позволяет уменьшить объем памяти, необходимый для хранения моделей, и сэкономить пространство на устройстве.
- Ограничения и компромиссы при применении методов оптимизации для систем распознавания речи на устройствах с ограниченными вычислительными ресурсами.
- Потеря точности. В процессе оптимизации могут использоваться различные техники, которые приводят к некоторой потере точности распознавания речи. Важно балансировать между производительностью и точностью, чтобы достичь оптимального компромисса.
- Время обработки. Некоторые методы оптимизации могут требовать дополнительного времени для обработки речевых данных. Это может быть неприемлемо для систем, где требуется быстрая обратная связь или реакция на голосовые команды.
- Зависимость от данных. Некоторые методы оптимизации могут быть зависимыми от конкретного набора данных или языка. Это создает ограничения в применимости методов оптимизации для различных языков или специфических контекстов использования.

В целом, применение методов оптимизации для систем распознавания речи на устройствах с ограниченными вычислительными ресурсами является важным и полезным подходом, который может принести множество преимуществ. Однако важно учитывать ограничения и компромиссы, связанные с таким применением.

Для достижения оптимальных результатов при применении методов оптимизации в системах распознавания речи на устройствах с ограниченными ресурсами, рекомендуется учитывать следующие факторы:

- Контекст использования. При выборе методов оптимизации необходимо

учитывать конкретный контекст использования системы распознавания речи. Различные приложения и устройства могут иметь разные требования к производительности, точности и времени обработки.

- Адаптация к конкретному устройству. Методы оптимизации должны быть специально настроены и адаптированы к конкретному устройству с ограниченными вычислительными ресурсами. Это позволит достичь наилучшего сочетания производительности и точности при распознавании речи.
- Тестирование и оценка. Важно проводить тщательное тестирование и оценку эффективности применяемых методов оптимизации. Это позволит оценить потерю точности, производительность и другие параметры системы на устройствах с ограниченными ресурсами.
- Компромиссы между производительностью и точностью. При оптимизации системы распознавания речи на ограниченных устройствах необходимо сделать компромисс между производительностью и точностью. Важно найти оптимальный баланс, чтобы система обеспечивала достаточно высокую точность, несмотря на ограничения ресурсов.

В целом, применение методов оптимизации для систем распознавания речи на устройствах с ограниченными вычислительными ресурсами может значительно улучшить производительность и эффективность таких систем. Однако необходимо тщательно учитывать ограничения и проводить адаптацию методов оптимизации к конкретному контексту использования, чтобы достичь наилучших результатов.

### **Заключение**

В данной статье были рассмотрены важные аспекты разработки систем распознавания речи на устройствах с ограниченными вычислительными ресурсами. Существующие методы распознавания речи требуют значительных вычислительных ресурсов, что ограничивает их применимость на устройствах с ограниченной вычислительной мощностью.

Однако, применение методов оптимизации позволяет эффективно работать с системами распознавания речи и на таких устройствах. Оптимизация с помощью методов компрессии данных, сокращения размера моделей, квантизации параметров и оптимизации вычислений позволяют снизить требования к вычислительным ресурсам и повысить производительность.

Реальные приложения и системы распознавания речи, разработанные с использованием методов оптимизации, подтверждают их эффективность и важность. Они позволяют достичь высокой точности и быстродействия при распознавании речи на устройствах с ограниченными вычислительными ресурсами, что является важным фактором для создания удобных и эффективных систем в различных областях применения.

Однако, следует учитывать и ограничения методов оптимизации. Некоторые методы могут привести к небольшой потере точности распознавания или увеличению времени обработки. Поэтому необходимо находить компромисс между оптимизацией ресурсов и сохранением качества распознавания речи при разработке систем на устройствах с ограниченными вычислительными ресурсами.

В целом, методы оптимизации играют решающую роль в разработке эффективных систем распознавания речи на устройствах с ограниченными вычислительными ресурсами. Они позволяют справиться с ограничениями и достичь высокой производительности, точности и быстродействия систем распознавания речи, что является важным фактором для удобства и эффективности



использования в различных областях применения, в том числе в автомобильной промышленности, медицине, умном доме и др.

### Литература

1. Данильчук Т.И., Трофимов И.М. 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 И.Н., Омельченко А.В., Омельченко С.В. 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 А.К. 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 В.А., Панюлайтис С.В. 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 М.А., Протодьяконова Г.Ю. 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).