

Санкт-Петербургский государственный университет

Программная инженерия

Группа 24.М71-мм

# Разработка и оптимизация системы голосового управления для автоматизированного заполнения медицинской документации

*Альшаеб Басель*

Отчёт по производственной практике  
в форме «Решение»

Научный руководитель:  
ст. преподаватель, к.ф.-м.н. С. С. Сысоев.

Санкт-Петербург  
2026

# Оглавление

<b>1. Введение</b>	<b>3</b>
1.1. Цель и задачи исследования . . . . .	5
1.2. Практическая значимость работы . . . . .	6
<b>2. Постановка задачи</b>	<b>7</b>
<b>3. Анализ существующих решений</b>	<b>8</b>
<b>4. Архитектура разрабатываемой системы</b>	<b>9</b>
4.1. Клиентская часть системы . . . . .	9
4.2. Механизм активации системы по голосовой команде . . .	10
4.3. Серверная часть системы . . . . .	11
4.4. Семантическая интерпретация команд . . . . .	11
4.5. Отличие предлагаемого подхода от традиционных решений	12
<b>5. Алгоритмы обработки речевого сигнала</b>	<b>13</b>
5.1. Захват и предварительная обработка аудиосигнала . . .	13
5.2. Детекция речевой активности . . . . .	13
5.3. Выделение участков тишины . . . . .	14
5.4. Активация системы по ключевому слову . . . . .	14
5.5. Оптимизация вычислительных ресурсов . . . . .	15
5.6. Выводы по главе . . . . .	15
<b>6. Алгоритмы обработки команд и автоматического заполнения медицинской карты</b>	<b>16</b>
6.1. Разделение речевых фрагментов и управляющих команд	16
6.2. Формирование структурного представления команды . .	17
6.3. Использование языковой модели для интерпретации команд	17
6.4. Ограничение пространства решений и фильтрация команд	17
6.5. Особенности заполнения стоматологической карты . . .	18
6.6. Интеграция с медицинской информационной системой .	18
6.7. Сравнение с традиционным подходом . . . . .	19
6.8. Выводы по главе . . . . .	19

<b>7. Экспериментальная оценка и анализ эффективности системы</b>	<b>20</b>
7.1. Условия проведения экспериментов . . . . .	20
7.2. Временные характеристики обработки команд . . . . .	20
7.3. Анализ вычислительной нагрузки . . . . .	21
7.4. Сравнение с традиционным подходом . . . . .	21
7.5. Качественная оценка удобства использования . . . . .	22
7.6. Выводы по главе . . . . .	22
<b>8. Заключение</b>	<b>23</b>

# 1. Введение

В процессе эксплуатации медицинской информационной системы *Stommis*, предназначенной для сопровождения работы стоматологической клиники, была выявлена практическая проблема, связанная с взаимодействием врача с системой во время приёма пациента. В существующем варианте системы основная часть действий врача (заполнение полей медицинской карты, выбор диагнозов и процедур из справочников, навигация по разделам) выполняется с использованием традиционных средств ввода: клавиатуры и мыши. Данный подход является стандартным для многих медицинских информационных систем, однако в условиях стоматологического приёма он приводит к существенным ограничениям.

Специфика стоматологической практики заключается в том, что врач одновременно выполняет осмотр, использует инструменты, поддерживает стерильность и взаимодействует с пациентом. При этом необходимость переключаться на управление интерфейсом медицинской системы создаёт дополнительные паузы, повышает когнитивную нагрузку и может снижать общую эффективность приёма. Кроме того, работа в перчатках и требования к санитарной безопасности затрудняют использование клавиатуры и мыши, а также делают нежелательным частое касание внешних устройств ввода.

В системе *Stommis* ряд сценариев заполнения карты пациента требует выбора значений из структурированных иерархических списков (например, диагнозов, симптомов, процедур), а также точного указания параметров, привязанных к зубной формуле. Ручная навигация по таким структурам занимает заметное время и отвлекает врача от лечебного процесса. Таким образом, возникает необходимость внедрения альтернативного интерфейса управления системой, позволяющего сократить число действий, выполняемых вручную, и обеспечить более естественное взаимодействие в процессе работы.

Одним из перспективных направлений решения данной задачи является использование голосового управления. Однако прямое приме-

нение систем распознавания речи, ориентированных на непрерывную транскрипцию речевого потока, не является оптимальным. Во-первых, постоянная передача и обработка аудиопотока приводит к увеличению вычислительных затрат и эксплуатационных расходов (в случае использования облачных сервисов распознавания). Во-вторых, речь врача может включать нерелевантные комментарии, паузы, фоновые шумы и фразы, не предназначенные для внесения в медицинскую карту. В-третьих, даже корректно распознанный текст не обеспечивает автоматизации заполнения карты, если результатом является лишь сохранение неструктурированного текстового фрагмента.

Анализ существующих решений, доступных на рынке, показывает, что многие продукты ориентированы преимущественно на задачу преобразования речи в текст, после чего ответственность за структурирование информации и выбор корректных элементов медицинской карты остаётся на пользователе. Для стоматологической практики этого недостаточно, поскольку врачу важно не только получить текстовую расшифровку, но и быстро выполнить конкретное действие: выбрать диагноз из справочника, установить состояние конкретного зуба, добавить процедуру и заполнить соответствующие поля карты пациента.

В настоящей работе предлагается подход, отличающийся многоэтапной обработкой голосового ввода. Речь рассматривается не как конечный результат, а как источник управляющих команд. На ранних этапах выполняется выделение участков речи и тишины непосредственно на стороне клиента (в браузере), что позволяет снизить объём передаваемых данных и уменьшить нагрузку на сервер. Далее применяется механизм активации системы по ключевому слову, обеспечивающий запуск полноценной обработки только в те моменты, когда врач действительно произносит команду. На серверной стороне выполняется распознавание речи с использованием сервиса автоматического распознавания, после чего применяются методы семантической интерпретации команд на основе языковых моделей для преобразования текста в структурированное действие в рамках заданного множества допустимых операций. Итогом работы системы является не сохранение транскрипции речи,

а автоматизированное внесение изменений в структурированные поля медицинской карты.

## 1.1. Цель и задачи исследования

Целью данной магистерской диссертации является разработка и исследование программной системы голосового управления, обеспечивающей распознавание команд врача и автоматизированное заполнение структурированных данных медицинской карты пациента в информационной системе *Stommis*.

Для достижения поставленной цели необходимо решить следующие задачи:

- разработать механизм захвата звукового потока в браузере и предварительной обработки аудиосигнала с минимальными вычислительными затратами;
- реализовать алгоритм детекции речевой активности для выделения участков речи и условной тишины;
- разработать механизм активации по ключевому слову, снижающий количество ложных срабатываний и уменьшающий число обращений к модулю распознавания речи;
- реализовать распознавание речи с использованием внешнего сервиса и обеспечить совместимость аудиоформатов;
- разработать алгоритмы интерпретации распознанных команд и преобразования их в структурированные действия в системе;
- обеспечить интеграцию результатов обработки команд с медицинской информационной системой и справочниками предметной области.

## 1.2. Практическая значимость работы

Практическая значимость работы определяется возможностью внедрения разработанного подхода в систему *Stommis* и повышения эффективности работы врачей за счёт уменьшения числа ручных действий, сокращения времени заполнения документации и снижения отвлечения внимания врача от лечебного процесса.

## 2. Постановка задачи

Целью данной магистерской диссертации является разработка и исследование программной системы голосового управления, предназначенной для автоматизированного распознавания команд врача и структурированного заполнения электронной медицинской карты в процессе приёма пациента.

Для достижения поставленной цели в работе необходимо решить следующие задачи:

- разработать механизм захвата и предварительной обработки звукового сигнала в браузере врача с минимальными вычислительными затратами;
- реализовать алгоритм детекции речевой активности, позволяющий выделять участки речи и условной тишины;
- разработать метод активации системы распознавания по голосовой команде без использования физических элементов управления;
- реализовать модуль распознавания речи с использованием облачных сервисов;
- разработать алгоритмы интерпретации распознанных команд и сопоставления их с элементами структуры медицинской карты;
- обеспечить интеграцию результатов обработки речи с существующей информационной системой медицинского учреждения.



### 3. Анализ существующих решений

Существующие системы голосового ввода в медицинских информационных системах в большинстве случаев ограничиваются задачей преобразования речи в текст. При этом распознанная речь сохраняется в неизменном виде без выполнения семантического анализа и без привязки к конкретным элементам медицинской карты пациента.

В отличие от данных решений, предлагаемый в работе подход предполагает многоэтапную обработку голосового ввода. На первом этапе выполняется фильтрация звукового потока и выделение речевых команд. На втором этапе осуществляется распознавание речи, после чего применяется семантический анализ с использованием языковых моделей для определения типа команды и соответствующего действия в системе.

Таким образом, система ориентирована не на сохранение речи как текстового комментария, а на автоматизированное заполнение структурированных данных медицинской карты, включая выбор диагнозов, симптомов и процедур из заранее определённых медицинских справочников.

## 4. Архитектура разрабатываемой системы

Разрабатываемая система голосового управления предназначена для использования в условиях медицинского приёма и ориентирована на минимизацию вмешательства в рабочий процесс врача. В связи с этим архитектура системы проектировалась с учётом следующих требований:

- минимальное использование вычислительных ресурсов на стороне клиента;
- возможность работы в браузере без установки дополнительного программного обеспечения;
- масштабируемость и независимость компонентов системы;
- возможность интеграции с существующими медицинскими информационными системами.

В соответствии с указанными требованиями система реализована в виде распределённого клиент-серверного решения и включает следующие основные компоненты:

- клиентское веб-приложение врача;
- сервер обработки аудиоданных и команд;
- модуль распознавания речи;
- модуль семантической интерпретации команд;
- модуль интеграции с медицинской информационной системой.

### 4.1. Клиентская часть системы

Клиентская часть системы представляет собой веб-приложение, запускаемое в стандартном браузере врача. Основной задачей клиентского компонента является захват аудиопотока, предварительная фильтра-

ция звуковых данных и передача релевантных фрагментов речи на сервер для дальнейшей обработки.

Для захвата звука используется стандартный Web API браузера, обеспечивающий доступ к микрофону пользователя. Обработка аудиопотока осуществляется в режиме реального времени с применением алгоритма детекции речевой активности (Voice Activity Detection, VAD), который позволяет разделять звуковой поток на участки речи и условной тишины.

Использование VAD на стороне клиента позволяет существенно снизить объём передаваемых данных и уменьшить нагрузку на сервер, поскольку на обработку отправляются только сегменты, содержащие речь врача.

## **4.2. Механизм активации системы по голосовой команде**

Одной из ключевых особенностей разрабатываемой системы является отказ от физического управления процессом записи речи. В условиях медицинского приёма врач часто работает в перчатках и не имеет возможности взаимодействовать с интерфейсом посредством кнопок или сенсорных элементов.

Для решения данной проблемы в системе реализован механизм активации по голосовой команде. Клиентская часть непрерывно анализирует аудиопоток с использованием VAD, однако полноценная запись и обработка речи запускается только после обнаружения короткой голосовой команды активации.

Активация системы осуществляется по принципу “ключевого слова”. При обнаружении короткого речевого сегмента он передаётся на сервер для проверки на соответствие ключевой фразе. В случае успешного распознавания команды активации система переходит в режим обработки основной команды врача.

### **4.3. Серверная часть системы**

Серверная часть системы реализована в виде веб-сервиса и выполняет функции централизованной обработки аудиоданных и интерпретации команд. Сервер принимает аудиофрагменты, переданные с клиентской части, и выполняет их дальнейшую обработку в несколько этапов.

На первом этапе выполняется нормализация и, при необходимости, перекодирование аудиосигнала в формат, совместимый с используемым сервисом распознавания речи. Для обеспечения совместимости применяется конвертация входных данных в формат Ogg Opus.

На втором этапе выполняется распознавание речи с использованием внешнего сервиса автоматического распознавания речи. Результатом данного этапа является текстовая транскрипция голосовой команды врача.

### **4.4. Семантическая интерпретация команд**

Полученная текстовая транскрипция не используется напрямую для заполнения медицинской карты. Вместо этого в системе реализован дополнительный этап семантического анализа команд, целью которого является определение типа действия и соответствующих параметров.

Для решения данной задачи используется языковая модель, способная интерпретировать естественный язык и сопоставлять распознанные фразы с заранее определёнными действиями системы. В результате семантического анализа команда врача преобразуется в структурированное представление, пригодное для автоматического заполнения полей медицинской карты.

Такой подход позволяет фильтровать нерелевантные фрагменты речи и снижает вероятность некорректного заполнения данных, что особенно важно в медицинских информационных системах.

## **4.5. Отличие предлагаемого подхода от традиционных решений**

В отличие от традиционных систем голосового ввода, ориентированных на сохранение полной текстовой расшифровки речи врача, разрабатываемая система использует многоуровневую фильтрацию и интерпретацию команд.

Речь врача рассматривается не как конечный результат, а как источник управляющих сигналов, на основании которых выполняются структурированные изменения в медицинской карте пациента. Такой подход особенно эффективен в стоматологической практике, где врачу необходимо быстро фиксировать состояние пациента без отвлечения от лечебного процесса.

## **5. Алгоритмы обработки речевого сигнала**

Обработка речевого сигнала в разрабатываемой системе осуществляется в несколько этапов, каждый из которых направлен на снижение вычислительной нагрузки, уменьшение количества ошибок распознавания и повышение удобства использования системы в реальных условиях медицинского приёма.

Основная идея алгоритмического подхода заключается в том, что непрерывный звуковой поток не передаётся целиком на сервер и не подвергается постоянному распознаванию. Вместо этого используется последовательная фильтрация аудиоданных, позволяющая выделять только информативные участки речи и игнорировать фоновые шумы и нерелевантные звуки.

### **5.1. Захват и предварительная обработка аудиосигнала**

Захват аудиосигнала осуществляется на стороне клиента с использованием стандартных средств Web Audio API. Входной сигнал представляет собой непрерывный поток аудиоданных, поступающих с микрофона пользователя.

На этапе предварительной обработки к аудиосигналу применяются базовые методы улучшения качества записи, включая подавление шума, автоматическую регулировку усиления и компенсацию эха. Данные методы позволяют стабилизировать амплитуду сигнала и повысить устойчивость последующих этапов обработки к изменениям громкости речи и внешним акустическим условиям.

### **5.2. Детекция речевой активности**

Для разделения непрерывного аудиопотока на участки речи и условной тишины в системе используется алгоритм детекции речевой активности (Voice Activity Detection, VAD). Задачей данного алгоритма является определение моментов начала и окончания речевых сегментов в

звуковом сигнале.

Использование VAD позволяет отказаться от постоянной записи и передачи аудиоданных на сервер, что существенно снижает объём обрабатываемой информации и нагрузку на вычислительные ресурсы. Алгоритм VAD работает в режиме реального времени и принимает решение о наличии речи на основе анализа кратковременных характеристик сигнала.

### **5.3. Выделение участков тишины**

Завершение речевой команды определяется на основе анализа продолжительности интервалов тишины между фрагментами речи. После обнаружения речевого сегмента система продолжает наблюдение за входным сигналом и фиксирует момент завершения команды при достижении заданной длительности условной тишины.

Использование интервалов тишины в качестве критерия завершения команды позволяет отказаться от явных маркеров окончания речи и делает взаимодействие с системой более естественным для пользователя. Такой подход особенно удобен в медицинской практике, где речь врача может сопровождаться паузами, не являющимися признаком завершения команды.

### **5.4. Активация системы по ключевому слову**

Для предотвращения случайной активации системы и снижения количества ложных срабатываний используется механизм активации по ключевому слову. В неактивном состоянии система выполняет только детекцию речевой активности и анализ коротких речевых фрагментов.

При обнаружении короткого речевого сегмента он передаётся на сервер для проверки на соответствие ключевой фразе активации. Только после подтверждения ключевого слова система переходит в активный режим и начинает обработку основной команды врача.

Данный подход позволяет существенно сократить количество обращений к сервису распознавания речи и, как следствие, снизить эксплу-

атационные затраты и задержки обработки.

## **5.5. Оптимизация вычислительных ресурсов**

Одним из ключевых требований к разрабатываемой системе является минимизация вычислительных затрат при сохранении приемлемой точности распознавания. Для достижения данной цели применяется многоуровневая фильтрация аудиосигнала.

На клиентской стороне выполняется первичная фильтрация и сегментация речи с использованием VAD. На сервер передаются только короткие аудиофрагменты, потенциально содержащие команды или ключевые слова. Полноценное распознавание речи и семантическая интерпретация выполняются исключительно для подтверждённых сегментов.

Таким образом, большая часть нерелевантных данных отбрасывается на ранних этапах обработки, что обеспечивает эффективное использование вычислительных ресурсов и повышает масштабируемость системы.

## **5.6. Выводы по главе**

В данной главе был рассмотрен алгоритмический подход к обработке речевого сигнала, основанный на последовательной фильтрации аудиоданных и активации системы по ключевому слову. Предложенные алгоритмы позволяют обеспечить устойчивую работу системы в реальных условиях медицинского приёма при ограниченных вычислительных ресурсах.

Использование детекции речевой активности и анализа интервалов тишины делает взаимодействие с системой естественным и не требует от врача дополнительного обучения или изменения привычного рабочего процесса.



## **6. Алгоритмы обработки команд и автоматического заполнения медицинской карты**

После завершения этапа распознавания речи система получает текстовую транскрипцию голосовой команды врача. Однако прямое использование данного текста для заполнения медицинской документации является неэффективным и может приводить к ошибкам, поскольку речь врача часто содержит уточнения, комментарии и фразы, не предназначенные для прямого сохранения в медицинской карте.

В связи с этим в разрабатываемой системе реализован отдельный этап интерпретации команд, целью которого является преобразование неструктурированной текстовой информации в формализованное представление, пригодное для автоматизированного внесения данных в электронную медицинскую карту пациента.

### **6.1. Разделение речевых фрагментов и управляющих команд**

Речь врача в процессе приёма пациента может содержать как управляющие команды, адресованные системе, так и пояснительные комментарии, не имеющие прямого отношения к заполнению медицинской карты. Для корректной обработки голосового ввода необходимо различать данные типы речевых фрагментов.

В предлагаемом подходе речевая команда рассматривается как управляющее высказывание, содержащее указание на конкретное действие системы, например выбор диагноза, добавление симптома или заполнение определённого поля медицинской карты. Все остальные речевые фрагменты рассматриваются как нерелевантные для автоматической обработки и отбрасываются на данном этапе.

## 6.2. Формирование структурного представления команды

Для обеспечения однозначной интерпретации команд врача используется структурное представление команды, включающее тип действия и набор параметров. Каждая команда преобразуется в формализованную структуру следующего вида:

$$(тип\_команды, объект, параметры)$$

Тип команды определяет общее направление действия, например добавление записи, выбор значения или подтверждение состояния пациента. Объект команды указывает на элемент медицинской карты, к которому относится действие, а параметры содержат дополнительные данные, необходимые для выполнения команды.

## 6.3. Использование языковой модели для интерпретации команд

Для преобразования текстовой транскрипции в структурированное представление используется языковая модель, способная выполнять семантический анализ естественного языка. Языковая модель получает на вход текст команды и контекст текущего состояния медицинской карты пациента.

На основе анализа входных данных модель определяет наиболее вероятный тип команды и соответствующие параметры. Использование языковой модели позволяет учитывать вариативность формулировок команд, характерную для естественной речи, и снижает зависимость системы от строго фиксированного набора шаблонов.

## 6.4. Ограничение пространства решений и фильтрация команд

Для повышения надёжности работы системы и предотвращения некорректных действий языковая модель не имеет доступа к произ-

вольному пространству решений. Вместо этого интерпретация команд выполняется в рамках заранее определённого множества допустимых действий и объектов.

В частности, выбор диагнозов, симптомов и процедур осуществляется исключительно из предопределённых медицинских справочников и классификаторов. Такой подход позволяет снизить вероятность логических ошибок и обеспечивает соответствие результатов обработки требованиям медицинских информационных систем.

## **6.5. Особенности заполнения стоматологической карты**

Стоматологическая медицинская карта обладает высокой степенью структурированности и включает множество взаимосвязанных параметров, таких как состояние отдельных зубов, наличие патологий и выполненные процедуры. Ручной выбор соответствующих значений из иерархических списков является трудоёмкой задачей и отвлекает врача от лечебного процесса.

В разрабатываемой системе голосовые команды используются для навигации по структуре стоматологической карты и выбора необходимых элементов без прямого взаимодействия с пользовательским интерфейсом. Это позволяет врачу сосредоточиться на пациенте и существенно ускоряет процесс заполнения документации.

## **6.6. Интеграция с медицинской информационной системой**

После формирования структурированного представления команды выполняется её интеграция с медицинской информационной системой. На данном этапе проверяется корректность параметров команды и соответствие текущему состоянию медицинской карты пациента.

При успешной проверке система автоматически вносит изменения в соответствующие поля медицинской карты. В случае возникновения

неоднозначностей или ошибок команда отклоняется, а врач получает уведомление о необходимости уточнения запроса.

## **6.7. Сравнение с традиционным подходом**

В традиционных системах голосового ввода распознанная речь сохраняется в медицинской карте в виде текстового комментария без дальнейшей обработки. Такой подход не позволяет автоматизировать работу с данными и требует последующего ручного анализа.

В отличие от этого, предлагаемый в работе метод ориентирован на интерпретацию речи как источника управляющих команд. Результатом обработки является не текст, а структурированные изменения в медицинской карте, что обеспечивает более высокий уровень автоматизации и снижает нагрузку на медицинский персонал.

## **6.8. Выводы по главе**

В данной главе был рассмотрен алгоритмический подход к интерпретации голосовых команд врача и автоматизированному заполнению медицинской карты пациента. Использование языковой модели в сочетании с ограничением пространства допустимых решений позволяет обеспечить надёжную и контролируемую обработку команд в медицинской информационной системе.

## **7. Экспериментальная оценка и анализ эффективности системы**

Целью экспериментальной части работы является оценка эффективности разработанной системы голосового управления в условиях, приближенных к реальной медицинской практике. В рамках экспериментов анализируются временные характеристики обработки команд, вычислительная нагрузка на клиентскую и серверную части, а также сравнительная эффективность предложенного подхода по отношению к традиционным решениям голосового ввода.

### **7.1. Условия проведения экспериментов**

Экспериментальные исследования проводились в условиях, моделирующих рабочее место врача-стоматолога. Клиентская часть системы запускалась в современном веб-браузере на персональном компьютере, оснащённом стандартным микрофоном. Серверная часть системы функционировала в локальной вычислительной сети медицинского учреждения.

Для оценки устойчивости системы использовались различные сценарии речевого взаимодействия, включающие как корректные голосовые команды, так и нерелевантные речевые фрагменты, фоновые шумы и паузы в речи. Это позволило оценить работу системы в условиях, приближенных к реальной клинической среде.

### **7.2. Временные характеристики обработки команд**

Одним из ключевых параметров системы голосового управления является задержка между произнесением команды врачом и её фактическим выполнением в медицинской информационной системе. Общая задержка обработки команды складывается из следующих компонентов:

- время детекции речевой активности и завершения команды;

- время передачи аудиофрагмента на сервер;
- время распознавания речи;
- время семантической интерпретации команды;
- время внесения изменений в медицинскую карту.

Результаты измерений показали, что использование клиентской детекции речевой активности и активации по ключевому слову позволяет существенно сократить среднее время обработки команд за счёт уменьшения объёма передаваемых и обрабатываемых аудиоданных.

### **7.3. Анализ вычислительной нагрузки**

Для оценки эффективности использования вычислительных ресурсов была проанализирована нагрузка на клиентскую и серверную части системы. На стороне клиента основная вычислительная нагрузка связана с выполнением алгоритма детекции речевой активности и предварительной обработки аудиосигнала.

Экспериментальные результаты показали, что использование VAD в браузере не оказывает заметного влияния на отзывчивость пользовательского интерфейса и не требует значительных вычислительных ресурсов. На серверной стороне нагрузка существенно снижается за счёт того, что распознавание речи и семантический анализ выполняются только для подтверждённых речевых команд.

### **7.4. Сравнение с традиционным подходом**

Для оценки преимуществ предложенного решения было выполнено сравнение с традиционным подходом, при котором вся речь врача непрерывно передаётся на сервер и подвергается распознаванию.

Результаты сравнения показали, что использование многоэтапной фильтрации аудиосигнала позволяет существенно сократить количество обращений к сервису распознавания речи. В среднем количество

распознаваемых аудиофрагментов уменьшается за счёт отбрасывания нерелевантной речи и фоновых звуков.

Кроме того, предложенный подход обеспечивает более высокую степень автоматизации, поскольку результатом обработки является не текстовая расшифровка речи, а структурированное изменение медицинской карты пациента.

## **7.5. Качественная оценка удобства использования**

Качественная оценка системы проводилась на основе сценариев использования, характерных для стоматологической практики. Особое внимание уделялось возможности использования системы без физического взаимодействия с пользовательским интерфейсом.

Результаты оценки показали, что голосовое управление позволяет врачу сосредоточиться на лечебном процессе и снижает необходимость переключения внимания на экран компьютера. Использование голосовых команд особенно эффективно при заполнении структурированных разделов медицинской карты, требующих выбора значений из иерархических списков.

## **7.6. Выводы по главе**

В результате экспериментальных исследований было показано, что предложенная система голосового управления обеспечивает эффективную обработку команд врача при умеренных вычислительных затратах. Использование детекции речевой активности и активации по ключевому слову позволяет снизить нагрузку на сервер и уменьшить задержки обработки.

Сравнение с традиционными решениями голосового ввода подтвердило целесообразность использования многоэтапной обработки речи и семантической интерпретации команд для автоматизированного заполнения медицинской документации.

## 8. Заключение

В данной магистерской диссертации была рассмотрена задача разработки и оптимизации системы голосового управления, предназначенной для автоматизированного заполнения медицинской документации в процессе приёма пациента. Актуальность исследования обусловлена высокой нагрузкой на медицинский персонал и ограничениями традиционных интерфейсов взаимодействия в клинических условиях.

В ходе работы был предложен многоэтапный алгоритмический подход к обработке речевого сигнала, включающий детекцию речевой активности на стороне клиента, активацию системы по ключевому слову, распознавание речи и семантическую интерпретацию голосовых команд. Данный подход позволил существенно сократить объём обрабатываемых данных и снизить вычислительные затраты по сравнению с традиционными системами непрерывного распознавания речи.

Особое внимание в работе было уделено интерпретации команд врача и автоматическому заполнению структурированных элементов медицинской карты. Использование языковой модели в сочетании с ограничением пространства допустимых действий позволило обеспечить надёжную и контролируемую обработку команд без риска некорректного внесения данных.

Экспериментальные исследования подтвердили эффективность предложенного решения в условиях, приближенных к реальной стоматологической практике. Результаты показали снижение задержек обработки, уменьшение нагрузки на серверную часть системы и повышение удобства использования по сравнению с традиционными решениями голосового ввода.

Полученные результаты могут быть использованы при разработке медицинских информационных систем, ориентированных на повышение эффективности работы врачей и снижение когнитивной нагрузки при заполнении медицинской документации. В дальнейшем возможным направлением развития работы является расширение набора поддерживаемых команд, адаптация системы к другим медицинским специаль-



ностям и исследование методов локального распознавания речи.