

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
АМУРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
(ФГБОУ ВО «АмГУ»)

Институт компьютерных и инженерных наук
Кафедра информационных и управляющих систем
Направление подготовки 09.04.04 – Программная инженерия
Направленность (профиль) образовательной программы Управление
разработкой программного обеспечения

ДОПУСТИТЬ К ЗАЩИТЕ
Зав. кафедрой

_____ А.В. Бушманов
« ____ » _____ 2025 г.

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

на тему: Проектирование и реализация голосового управления
роботизированной системой

Исполнитель

студент группы 3105-ом2

_____ Лю Чао
(подпись, дата)

Руководитель

доцент, канд. техн. наук

_____ С.Г. Самохвалова
(подпись, дата)

Руководитель научного

содержания программы

магистратуры

профессор, доктор техн. наук

_____ И.Е. Еремин
(подпись, дата)

Нормоконтроль

инженер кафедры

_____ В.Н. Адаменко
(подпись, дата)

Рецензент

доцент, канд. техн. наук

_____ Л.В. Никифорова
(подпись, дата)

Благовещенск, 2025

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
АМУРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
(ФГБОУ ВО «АмГУ»)

Институт компьютерных и инженерных наук
Кафедра информационных и управляющих систем

УТВЕРЖДАЮ

Зав. кафедрой

_____ А.В. Бушманов

« ____ » _____ 2025 г.

З А Д А Н И Е

К магистерской диссертации студента группы 3105-ом2 Лю Чао

1. Тема магистерской диссертации: Проектирование и реализация голосового управления роботизированной системой

(Утверждено приказом от 06.03.2025 № 609-уч)

2. Срок сдачи студентом законченной работы (проекта): 10.06.2024

3. Исходные данные к магистерской диссертации: схемы, документация разработчиков, интернет ресурсы, учебная литература

4. Содержание магистерской диссертации (перечень подлежащих разработке вопросов): Распознавание речи, проектирование алгоритма решения задачи, разработка модуля управления голосом

5. Перечень материалов приложения (наличие чертежей, таблиц, графиков, схем, программных продуктов, иллюстративного материала и т.п.):

6. Рецензент магистерской диссертации: Никифорова Л.В., доцент, канд. техн. наук

7. Дата выдачи задания 29.01.2024

Руководитель выпускной квалификационной работы: С.Г. Самохвалова, доцент, канд. техн. наук

(фамилия, имя, отчество, должность, уч.степень, уч.звание)

Заявление принял к исполнению _____

РЕФЕРАТ

Магистерская диссертация содержит 72 страница, 20 рисунок, 39 источников

ИНТЕЛЛЕКТУАЛЬНЫЕ РОБОТЫ, РАСПОЗНАВАНИЕ РЕЧИ, ГЛУБОКОЕ ОБУЧЕНИЕ, АРДУИНО, ГОЛОСОВОЕ УПРАВЛЕНИЕ, МИКРОКОНТРОЛЛЕР SPCE061A

Исследование посвящено разработке системы голосового управления для колесного робота на базе микроконтроллера SPCE061A. Предложено решение по интеграции технологий распознавания речи (на основе НММ и DTW-алгоритмов) с аппаратной платформой Arduino. Реализованы функции обучения 15 голосовым командам, управления движением и автономного избегания препятствий через инфракрасные датчики. Практическая значимость подтверждена тестами с точностью распознавания >95 % в условиях шума.

Содержание

Введение	6
1 Интеллектуальные роботы и взаимодействие через речь	8
1.1 Состояние исследований в области распознавания речи и роботов	8
1.2 История исследований и текущая ситуация	13
1.2.1 Анализ современного состояния зарубежных исследований	13
1.2.2 Анализ текущего статуса исследований в Китае	17
1.3 Основное содержание исследования при выборе темы	20
2 Проектирование и реализация системы распознавания голоса робота	23
2.1 Общее техническое решение системы	23
2.1.1 Выбор контроллера	23
2.1.2 Определение технологии распознавания речи	27
2.1.3 Выбор объекта робота	29
2.1.4 Проектирование схемы управления скоростью и привода робота	29
2.1.5 Общее техническое решение системы	32
2.2 Проектирование аппаратной части системы	32
2.2.1 Проектирование цепей управления	32
2.2.2 Проектирование цепей управления двигателями робота	37
2.2.3 Схема сбора речевых сигналов	37
2.2.4 Электропитание системы	39
2.2.5 Проектирование системы управления голосовыми командами	40
3 Реализация программных модулей для интеллектуальных роботов	41
3.1 Процесс работы системы	41
3.2 Проектирование основной программы	44
3.2.1 Общая схема проектирования системы голосового управления	46
3.2.2 Модуль голосовой тренировки	48
3.2.3 Модуль распознавания голоса	48
3.2.4 Проектирование подпрограммы обслуживания прерываний	49
3.2.5 Модуль воспроизведения голоса	52

3.2.6	Определение состояния движения робота	53
3.2.7	Функция избегания препятствий	54
3.3	Отладка, симуляция и выводы системы распознавания голоса	58
3.3.1	Аппаратная отладка	58
3.3.2	Программная отладка	62
3.4	Симуляция системы распознавания голоса	62
3.5	Выводы	65
	Заключение	67
	Библиографический список	69

ВВЕДЕНИЕ

В свете наступления эры "Интеллектуального производства 4.0" интеллектуальные роботы становятся все более популярными. Роботы предварительно загружаются программным обеспечением и могут автоматически выполнять соответствующую работу в соответствии с встроенной программной системой. Они также могут принимать команды, отправленные им людьми, чтобы выполнить соответствующую работу. Они могут заменять или сотрудничать с людьми в таких областях, как автоматическая производственная линия, производство, строительство, исследования и т.д. Однако большинство таких роботов могут принимать только предварительно подготовленные встроенные программы, и их движения фиксированы и механичны. Несмотря на то что в настоящее время существуют отдельные типы роботов, которые могут взаимодействовать с операторами, в основном операторы вводят информацию и команды с клавиатуры, что вносит ограничения во взаимодействие между роботом и оператором. Естественный язык является наиболее естественным и удобным средством взаимодействия между людьми, и поэтому есть основания полагать, что естественный язык также станет одним из наиболее эффективных средств взаимодействия между роботами и людьми в будущем. В действительности, за последние несколько лет некоторые технические специалисты приложили усилия для исследования и применения технологии распознавания речи, и использование технологии распознавания речи для управления механическими и электрическими устройствами стало одним из передовых исследовательских направлений, что позволило достичь значительного прогресса в области применения технологии распознавания речи в управлении промышленными роботами. Полученные результаты научных исследований позволяют роботам взаимодействовать с операторами на языке команд, что внесло значительный вклад в теоретическое и практическое применение технологии распознавания речи в области "Интеллектуального производства". Тем не менее, если мы хотим выпустить на рынок роботов с

низкой стоимостью и высокой производительностью в распознавании речи, чтобы они действительно вошли в повседневную жизнь людей и добавили удовольствие в их жизнь, у нас еще предстоит пройти долгий путь, и нам нужно исследовать и разработать высокопроизводительную систему распознавания речи. Вот почему мы выбрали эту тему исследования.

Исследование по этой теме имеет следующие основные аспекты:

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

- распознавание речи - это комплексное применение технологии, пересекающей несколько дисциплин, таких как информатика, инженерия, лингвистика, психология и др. Достижение беспрепятственного языкового взаимодействия между роботами и людьми всегда было мечтой человечества, а технология распознавания речи - это ключ к этой мечте, поэтому эта тема является одной из важных областей научных исследований;

- выбор этой темы позволяет внести свой вклад в стимулирование интереса среди студентов, привлечь отдельных студентов к исследованиям в области интеллектуальных роботов и стимулировать некоторых студентов к исследованию систем распознавания речи для роботов, что впоследствии может привести к формированию исследовательской группы по интеллектуальным роботам.

1 ИНТЕЛЛЕКТУАЛЬНЫЕ РОБОТЫ И ВЗАИМОДЕЙСТВИЕ ЧЕРЕЗ РЕЧЬ

1.1 Состояние исследований в области распознавания речи и роботов

В мире, благодаря стремлению людей к искусственному интеллекту, технология распознавания речи начала исследоваться еще в 1950-х годах 20-го века. Самая ранняя система распознавания речи - это система Audry, разработанная лабораторией AT&T Bell, которая могла распознавать речь на английском языке для 10 цифр.

На данный момент технология распознавания речи уже добилась значительных успехов. Теоретически она основана на различных теоретических подходах, таких как теория линейного предсказания, теория динамического программирования, теория векторного квантования и другие. Появление этих теоретических подходов предоставило теоретическую основу и поддержку для прикладных исследований в области технологии распознавания речи. В практическом плане появились различные приложения, основанные на моделях скрытых марковских моделей (HMM) и искусственных нейронных сетях (ANN) для распознавания речи. Например, появились приложения, направленные на распознавание непрерывной речи от неопределенных пользователей, такие как система распознавания речи SPHINX, а также система распознавания речи Vi-aVoice, ориентированная на распознавание речи в новостях и т.д. Эти приложения укрепили уверенность в исследованиях в области искусственного интеллекта.

В области исследований искусственного интеллекта в большинстве стран основное внимание уделяется разработке и исследованию робототехники. Многие страны используют государственные инвестиции для стимулирования развития индустрии интеллектуальных роботов. Например, в 2013 году правительство США создало "Национальную программу по роботам", вложив 22 миллиарда долларов США, и технологии интеллектуальных роботов были названы ключевыми технологиями для военных нужд. В Европе, как важном

мировом экономическом центре, страны, такие как Германия и Великобритания, руководят в области исследований робототехники, вкладывая средства в разработку технологий обслуживания и медицинских роботов. В Японии, лидере в области электроники, акцент делается на разработке технологий робототехники, имитирующих человека, а также технологий роботов для развлекательных целей. Это разнообразие исследований делает область исследований в области робототехники разносторонней и разнообразной.

С начала XXI века, с увеличением среднего возраста населения в мире, растет спрос на интеллектуальных роботов. В развитых странах, таких как США и страны Европы, инвестиции в исследования и разработки в области интеллектуального производства составляют 10 %. Это стимулирует взрывной рост интеллектуальных роботов с разнообразными формами и функционалом, такими как роботы для эмоциональной поддержки, ухода за пожилыми людьми, образования и тренировок, домашние помощники, а также роботы для обеспечения общественной безопасности и тушения пожаров. Системы распознавания речи для роботов переживают значительное развитие, обладая высокой точностью распознавания и способностью логического мышления, позволяя им организовывать язык для ответа на вопросы людей с высокой точностью в различных сценариях. Верится, что в ближайшем будущем с уменьшением стоимости производства таких роботов они станут доступными для обычных семей.

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

роботов.

В области исследований по технологии распознавания речи Китай начал исследования сравнительно рано, еще в 50-х годах XX века. С 1987 года страна запустила программу "863", создала экспертную группу по искусственному интеллекту "863" и Лабораторию распознавания образов, образованную совместными усилиями Университета Цинхуа и Института автоматизации Китайской академии наук. Эта лаборатория специализируется в исследованиях технологии распознавания речи, в частности, в области распознавания речи на китайском языке. Китай достиг значительных успехов в этой области, достигнув мирового уровня в технологии распознавания речи на китайском языке. Разработанные системы распознавания речи на китайском языке и системы голосового взаимодействия на китайском языке имеют точность более 96 %, а уровень отклика системы превышает 90 % .

В июле 2000 года в Зоологическом музее Пекина, который был открыт в новом здании, была представлена выставка "Панда". В рамках этой выставки была использована разработанная в Китае система распознавания речи для неконтролируемого распознавания речи. Посетители могли общаться с большой "пандой" на месте и узнавать о жизненных привычках, структуре пищи, физиологической структуре больших панд. В условиях выставочного зала, где присутствует высокий уровень шума, распознаваемость речи все равно оставалась выше 98 %, что соответствует требованиям для практического использования. Этот живой и увлекательный способ презентации привлек много посетителей, заставив их останавливаться, смотреть и общаться.

На данный момент в Китае используется технология самостоятельной разработки встраиваемых чипов для распознавания речи, включая специализированную систему чипов, включающую функции распознавания речи, кодирования речи и синтеза речи. Система состоит из 8-битного микроконтроллера, низкочастотного фильтра, аналого-цифрового и цифро-аналогового преобразователей, усилителя мощности, предусилителя, ППЗУ, ОЗУ, модулей ШИМ и других. Она способна распознавать 30

специфических голосовых команд, скорость кодирования речи достигает 16 кб/с, а точность распознавания превышает 95 %. Эта технология может использоваться в умных голосовых игрушках, голосовых телефонах и других устройствах. Она уже находится в фазе коммерциализации и достигла мирового уровня.

6 апреля 2016 года под руководством Министерства промышленности и информационных технологий, Национальной комиссии по развитию и реформам, Министерства финансов и других ведомств под предводительством Ассоциации беспилотных летательных аппаратов города Шэньчжэнь был создан Альянс по развитию китайской робототехнической индустрии искусственного интеллекта. На совещании был утвержден План развития робототехнической индустрии (2016-2020), разработанный совместно Минпромторгом, Национальной комиссией по развитию и реформам и другими ведомствами. Этот план направлен на руководство здоровым, быстрым и устойчивым развитием робототехнической индустрии Китая.

Указанный в "Плане развития" отмечает основные проблемы в области исследований по распознаванию речи и голосовому управлению в Китае:

- исследования в области речевых технологий требуют дополнительных прорывов.

Количество данных в речевой информации велико, и она обладает временной изменчивостью. Речевые шаблоны различны для разных людей, и даже для одного и того же человека они могут быть разными.

- проблемы мониторинга и обработки точек разрыва

Когда человек говорит, паузы между словами неопределенны и могут быть как долгими, так и короткими. Это является объектом нечеткой обработки и является одним из основных факторов, влияющих на точность сопоставления слов.

- проблемы распознавания речи для неопределенных лиц

Чтобы интеллектуальные роботы действительно стали практическими и коммерчески успешными, необходимо решить проблему создания модели

распознавания речи для неопределенных лиц. Это также одна из ключевых проблем для преодоления в области технологии распознавания речи.

- реальное время распознавания речи

Интеллектуальные голосовые роботы требуют наличия возможности работы в реальном времени, поэтому системы распознавания речи должны удовлетворять требованиям по реальному времени.

- проблема обработки нечеткости в речи

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

Данная тема исследования выбирает в качестве объекта исследования и экспериментов колесного игрушечного робота. Основным управляющим устройством робота является микроконтроллер Arduino, использующий однокристалльный микроконтроллер SPCE061A. Исследование направлено на улучшение и доработку существующей системы управления с использованием технологии распознавания речи. Это позволит роботу выполнять действия, такие как движение вперед, назад, поворот влево, поворот вправо и др., под управлением голосовых команд.

Основное содержание исследования включает в себя следующее:

- разработка общей концепции системы в соответствии с целями выбранной темы. Это включает выбор контроллера системы, определение технологии распознавания речи, проектирование схемы управления моторами и др.

- на основе общей концепции проектирование аппаратной части системы, включая проектирование минимальной системы микроконтроллера, связанные схемы управления, схемы сброса и схемы распознавания речи. Это является материальной основой всей системы проектирования.

– проектирование программной части системы, включая разработку основной программы, обслуживающих подпрограмм, программ обучения и распознавания речи.

1.2 История исследований и текущая ситуация

1.2.1 Анализ современного состояния зарубежных исследований

Распознавание речи впервые началось примерно в 1950-х годах. В 1952 году Bell Labs предложила систему распознавания речи, которая могла распознавать английские цифры от одного до десяти, произносимые конкретным говорящим, с точностью от 97 до 99 процентов. Вышеупомянутая система имеет большие ограничения и предъявляет определенные требования к говорящему, содержанию и даже качеству голоса, но имеет большое значение для развития распознавания речи. В 1956 году лаборатория RCA (Американской радиокорпорации) изобрела систему голосовой пишущей машинки, которая могла распознавать десять определенных слогов. В 1959 году Д.Б. Фрай в своей статье предположил, что для достижения гибкости и точности машинного распознавания речи необходимо имитировать особенности обработки человеческой речи и применять статистические принципы к машинному распознаванию речи.

Примерно в 1960-е годы распознавание речи также в определенной степени развивалось благодаря научным исследователям. В 1964 году Мартин Т.Б. и др. предложили систему анализа речи, которая использует аналоговую пороговую логику для извлечения признаков. Систему можно использовать для идентификации согласных в речи говорящего. В статье 1966 года из Университета Карнеги-Меллон описывалась система, которая могла подавать речевые сигналы непосредственно в компьютер и получать транскрипцию фонем из объединенных образцов речи, подаваемых в компьютер через микрофон и аналого-цифровой преобразователь. В 1968 году Винцук предложил метод поэлементного распознавания слов, который обеспечивал оптимальную согласованность между стандартными элементами и неизвестными элементами последовательности и устранял теоретический

недостаток алгоритма временной нормализации слов перед распознаванием. При экспериментальном тестировании алгоритма поэлементного распознавания на тестовой последовательности из 600 произношений слов из 12 классов, выполняемых одним говорящим, ошибок не возникло.

Примерно в 1970-х годах исследования в области распознавания речи продолжались. В 1970 году советские исследователи провели эксперимент по автоматическому распознаванию речи из 203 русских слов. В эксперименте на 5000 слов средняя надежность распознавания для говорящего составила 0–95. Время расчета на распознавание составляет 2-4 секунды. В 1975 году исследователи IBM предложили общую структуру декодера статистики языка для непрерывного распознавания речи, в которой входные данные декодера обрабатываются как строка фонетических символов, и декодер находит наиболее вероятный фонетический символ для каждой входной языковой строки. предложение.

В последующие десять лет важным событием в области распознавания речи стало предложение скрытой марковской модели HMM (Hidden Markov Model). В 1981 году Bell Labs предложила метод построения уровней, который оптимально выравнивал по времени последовательности союзов с последовательностями изолированных шаблонов ссылок на слова. В 1989 году Л.Р. Рабинер предложил теорию скрытой марковской модели HMM (Hidden Markov Model), от простейшей концепции (дискретная цепь Маркова) до самой сложной модели (модель переменной продолжительности, модель непрерывной плотности). Все они подробно описаны и попытались проиллюстрировать. применение теории HMM к некоторым простым задачам распознавания речи и укажем, как эти методы могут (и применялись) применяться к более сложным задачам распознавания речи. В том же году Bell Labs предложила реализацию алгоритма кадрово-синхронного сетевого поиска, который мог бы распознавать непрерывную речь как связную последовательность слов в соответствии с заданной грамматикой.

Благодаря некоторым прорывам в технологии распознавания речи,

произошедшим в 1970-х и 1980-х годах, распознавание речи постепенно стало зрелым в 1990-х годах. Технология Maximum A Posterior (MAP) была предложена в 1994 году, и используемый метод байесовской оценки был успешно применен для акустического моделирования при распознавании речи. В 1995 году широко использовалась технология линейной регрессии максимального правдоподобия (MLLR), поскольку ее можно использовать для решения проблем адаптации модели.

В первое десятилетие XXI века распознавание речи достигло прорывного прогресса. Во-первых, в 2002 году Кембриджский университет в Великобритании предоставил набор инструментов НТК. Этот набор инструментов был с открытым исходным кодом. Он содержал некоторые систематические документы и значительно снизил порог обучения. В 2006 году Хитон и другие исследователи в области глубокого обучения предложили сеть глубоких убеждений (Deep Belief Nets, DBN), которая сделала возможным послойное глубокое обучение. С тех пор глубокое обучение быстро развивалось, а глубокие нейронные сети и другие модели глубокого обучения стали широко использоваться в области распознавания речи. Примерно в 2009 году Хитон и др. применили сети глубокого доверия к акустическому моделированию и построили модель DBN-HMM. В стандартном корпусе TIMIT эта модель всегда превосходила другие технологии того времени.

С применением технологии глубокого обучения для распознавания речи распознавание речи с тех пор быстро развивалось. Контекстно-зависимая скрытая марковская модель глубокой нейронной сети (CD-DNN-HMM), предложенная Ю Дуном и др. в 2011 году, представляет собой новую акустическую модель, которая используется при распознавании речи с большим словарным запасом (Large-Vocabulary Continuous Speech Recognition, LVSR).) работает лучше, чем традиционная широко используемая модель смеси с высокой случайностью. Позже была предложена сверточная нейронная сеть (CNN). Некоторые из ее характеристик, такие как локальная связность, больше подходят для обработки речи, а ее производительность лучше, чем у DNN. Затем

постепенно появляется больше исследований, основанных на сверточной нейронной сети. Позже внедрение таких сетевых структур, как рекуррентная нейронная сеть (RNN) и долгосрочная краткосрочная память (LSTM), еще больше улучшило производительность всей системы распознавания речи.

После того, как глубокое обучение стало широко использоваться в области распознавания речи, постепенно появилось новое сквозное распознавание речи. Приведенное выше описание представляет собой некоторый процесс разработки на основе гибридной модели (конвейерной модели). Гибридная модель делит систему распознавания речи на различные модули, включая часть предварительной обработки, акустическую модель, языковую модель и т. д., каждый из которых соединен последовательно. модуль Вход — это выход предыдущего модуля, к которому он примыкает. Входные данные сквозного распознавания речи — это речь на одном конце, а выходные данные — текст на другом конце, что устраняет проблему стыковки и выравнивания между различными модулями в середине.

Алекс и др. предложили новый общий метод классификации событий на основе RNN в 2006 году. Этот метод напрямую обрабатывает последовательности без какой-либо обработки, и результаты экспериментов показывают, что производительность лучше, чем у модели RNN-HMM. Впоследствии, в 2013 году, был предложен преобразователь на основе RNN (RNN Transducer). Эта модель объединила языковую информацию для совместного моделирования, что еще больше повысило точность набора данных. Baidu предложила систему Deep Speech в 2015 году, которая на тот момент добилась самых передовых результатов на наборе данных Switchboard. Позже Ян Чорови и другие предложили метод моделирования распознавания речи, основанный на механизме внимания. В 2015 году Уильям Чан и др. предложили модель LAS. Эта модель внесла определенные улучшения в классификацию времени соединения между конечными точками, совместно изучила все компоненты распознавателя речи и определила базовую структуру модели внимания. В дальнейшем постоянно предлагались методы кодирования и

декодирования структур, основанные на RNN и внимании. Кроме того, методы моделирования на основе RNN-T и CTC также были применены в области распознавания речи, что еще больше способствует развитию распознавания речи. Базовая архитектура модели Transformer, предложенная Google в 2018 году, основана на механизме самообслуживания с несколькими головками (Self-Attention). С появлением Transformer серия улучшенных моделей на основе Transformer получила широкое распространение в области распознавания речи и добилась относительно хороших результатов. В 2020 году модель Transformer, улучшенная сверткой, Conformer, в очередной раз обновила рекорды основных рейтингов распознавания речи. Позже также широко использовались некоторые сквозные модели с открытым исходным кодом, такие как ESPnet.

1.2.2 Анализ текущего статуса исследований в Китае

Что касается исследований в области технологий распознавания речи, страна начала их рано и начала исследования в 1950-х годах. В 1987 году в стране был принят «Национальный план исследований и разработок в области высоких технологий» (именуемый «Планом 863»). Он является продуктом сочетания стратегического видения ученых и дальновидности политиков и объединяет стратегические потребности Китая в области высоких технологий. развитие высоких технологий. Государственный совет организовал более 200 ученых со всей страны для демонстрации предложений плана в течение более чем полугода. В эпоху, когда наука и технологии Китая должны были срочно наверстать упущенное, реализация «Плана 863» эффективно способствовала прогрессу Китая. высокие технологии. Что касается компьютеров, то была создана Экспертная группа по интеллектуальным компьютерам 863 — совместный проект, инициированный Университетом Цинхуа и Институтом автоматизации Китайской академии наук с целью совместного создания Национальной ключевой лаборатории распознавания образов, специализирующейся на технологии распознавания речи, и добилась значительного прогресса в китайской технологии распознавания речи. Технология распознавания речи достигла международного передового уровня.

Уровень точности китайской системы голосовой диктовки и китайской голосовой системы диалога между человеком и компьютером достиг более 96 %, а скорость отклика системы. достигла более 90 % .

В июле 2000 года «Панда», выставленная в недавно открытом зоологическом зале Пекинского музея естественной истории, приняла независимо разработанную в моей стране систему распознавания непрерывной речи, не специфичную для человека, посредством живого диалога между аудиторией и большой «Пандой». зрители смогут узнать о образе жизни, еде гигантской панды. Структура, физиологическое строение и другие знания, в среде с сильными шумовыми помехами, такой как выставочный зал, уровень распознавания речи по-прежнему достигает более 98 % , достигая требований практического уровня. Этот живой и яркий метод объяснения привлекает. большое количество Посетителей останавливались, чтобы посмотреть и пообщаться с ними.

В настоящее время в Китае используются технологии проектирования встроенных чипов для самостоятельной разработки специальной системы чипов распознавания речи, включая функции распознавания речи, кодирования речи и синтеза речи. Она состоит из 8-битного микроконтроллера, фильтра нижних частот, А/. Д и ЦАП, усилитель мощности, предусилитель, ПЗУ, ОЗУ, Он состоит из широтно-амплитудной модуляции и других модулей. Он может распознавать 30 голосовых команд конкретных людей со скоростью кодирования голоса 16 кбит/с и скоростью распознавания более 95 %. Его можно использовать в интеллектуальных голосовых игрушках, голосовых. телефонные звонки и другое оборудование Начали этап коммерциализации и достигли международного продвинутого уровня.

28 декабря 2021 года 15 ведомств, включая Министерство промышленности и информационных технологий, Национальную комиссию по развитию и реформам и Министерство науки и технологий, совместно опубликовали «План развития индустрии робототехники «14-й пятилетки», в котором содержатся рекомендации по «14-я пятилетка». В этот период были

осуществлены комплексные мероприятия и систематическое планирование развития робототехники, имеющей важное руководящее значение для содействия качественному развитию робоиндустрии моей страны. Придерживаться «инноваций, приложений, базового улучшения, комплексного развития», сосредоточиться на прорывах в основных технологиях, сосредоточиться на консолидации промышленной основы, сосредоточиться на повышении эффективных поставок, сосредоточиться на расширении рыночных приложений, повысить стабильность и конкурентоспособность промышленную цепочку и цепочку поставок, а также продолжать улучшать экологию промышленного развития, способствовать высококачественному развитию робототехнической промышленности и оказывать мощную поддержку созданию сильной производственной страны, здорового Китая и создания лучшей жизни.

В «14-й пятилетке развития индустрии роботов», хотя технология распознавания речи роботов является важным направлением в продвижении интеллекта роботов, в статье также косвенно или прямо указываются связанные с этим проблемы и проблемы. Эти вопросы в основном касаются следующих аспектов:

- Исследования речевых моделей требуют дальнейших прорывов

Объем данных речевой информации велик и меняется во времени. Характер голоса различен у разных людей, даже у одного и того же человека.

- Мониторинг точек останова и решение проблем.

Когда люди говорят, паузы между словами неопределенные, иногда длинные или короткие. Это нечеткий объект обработки и один из основных факторов, влияющих на точность сопоставления слов.

- Проблема распознавания речи для неспецифических людей

Чтобы интеллектуальные роботы действительно стали практичными и коммерциализированными, они должны решить проблему создания моделей распознавания речи для неспецифических людей. Это также одна из ключевых проблем для прорывов в технологии распознавания речи.

- Распознавание речи в реальном времени.

Сценарии применения интеллектуальных голосовых роботов требуют производительности в реальном времени, поэтому система распознавания речи должна соответствовать характеристикам производительности в реальном времени. На данном этапе производительность технологии распознавания речи в сложных средах (таких как шум, многоязычные и многодиалектные сцены) все еще нуждается в улучшении, особенно с точки зрения реагирования в реальном времени и высокоточного распознавания.

– Проблема обработки неоднозначности речи.

Конечной целью распознавания речи является достижение «прослушивания» и «понимания», но по-прежнему существуют трудности с семантическим пониманием и обработкой естественного языка, особенно когда вы сталкиваетесь с расплывчатыми или имеющими сильную контекстуальную релевантность инструкциями. Люди в разных регионах имеют разное произношение, а многие китайские иероглифы полифоничны, а произношение неопределенное. Поэтому распознавание речи должно основываться на произношении контекста, соответствующей лексики и синтаксических правил, чтобы обеспечить точность распознавания речи.

1.3 Основное содержание исследования при выборе темы

Стремясь решить проблемы, связанные с сложностью и сложностью переноски системы ручного управления колесными мобильными роботами, в данной статье изучается метод управления движением колесных мобильных роботов, основанный на распознавании речи. Во-первых, анализируются принципы улучшения речи и распознавания речи, чтобы обеспечить теоретическую основу для последующих улучшенных алгоритмов. В этой теме в качестве объекта исследований и экспериментов используются игрушечные роботы на колесиках, направленные на достижение речевого управления посредством улучшения и оптимизации система управления роботом. Интеллектуальное управление технологией идентификации. Основным компонентом управления роботом является основная плата управления Arduino, которая использует микроконтроллер SPCE061A в качестве процессора. Она

обладает высокими возможностями интеграции и обработки и может удовлетворить потребности сценариев многофункциональных приложений. На основе оригинальной системы управления ручкой и в сочетании с технологией распознавания речи это исследование разрабатывает и реализует голосовое управление движениями робота для улучшения его возможностей интеллектуального взаимодействия.

Оригинальная система управления ручкой в основном завершает управление движением робота с помощью традиционных кнопок, включая основные функции, такие как вперед, назад, поворот налево и поворот направо. Однако этот метод управления имеет определенные ограничения в практическом применении, например, требование непрерывного ручного управления пользователем и недостаток гибкости и интеллекта. С этой целью в этом исследовании представлена технология распознавания речи, которая заменяет традиционные ручные операции вводом на естественном языке, чтобы робот мог автономно выполнять управление движениями в соответствии с голосовыми инструкциями. В частности, пользователям нужно только подавать простые голосовые команды, такие как «вперед», «назад», «повернуть налево» или «повернуть направо», и робот сможет анализировать инструкции и выполнять соответствующие инструкции с помощью встроенного средства распознавания речи. модуль на главной плате управления.

В конструкции системы модуль распознавания речи органично интегрирован с основной платой управления Arduino. Голосовой сигнал сначала собирается через микрофон и преобразуется в цифровой сигнал. После анализа алгоритмом обработки голоса извлекается соответствующий командный сигнал. Основная плата управления Arduino сопоставляет командный сигнал с заданным режимом движения и управляет роботом, чтобы он выполнял соответствующие действия, приводя в движение двигатель. В то же время, чтобы повысить надежность и удобство работы системы, в этом исследовании была оптимизирована точность алгоритма распознавания речи, чтобы он мог точно распознавать команды в шумной среде, а также был принят механизм обратной

связи в реальном времени для обеспечения что робот быстро реагирует на инструкции и точен.

Целью этого исследования является не только достижение голосового управления, но и повышение уровня интеллекта колесных игрушечных роботов путем изучения интеграции технологий распознавания голоса и управления роботами, а также предоставления технической поддержки и демонстрации приложений в области обучения роботов, развлечения и помощь по дому. Основное содержание исследования выбранной темы заключается в следующем:

- Разработать общий план системы в соответствии с целями выбора темы, включая выбор системного контроллера, определение технологии распознавания речи, разработку схемы привода двигателя и т. д.

- Разработать аппаратную подсистему на основе общего плана, включая минимальный проект системы микроконтроллера, проектирование соответствующей схемы управления, проектирование схемы сброса, проектирование схемы распознавания голоса и т. д. Это является материальной основой для всей конструкции системы.

- Разработка подсистем программного обеспечения, включая разработку основной программы, подпрограммы обслуживания прерываний, подпрограммы обучения речи и распознавания речи и т. д.

2 ПРОЕКТИРОВАНИЕ И РЕАЛИЗАЦИЯ СИСТЕМЫ РАСПОЗНАВАНИЯ ГОЛОСА РОБОТА

2.1 Общее техническое решение системы

Ввиду небольших размеров колесного игрушечного робота для удобства установки контроллера рассматриваются две альтернативы: программируемый логический контроллер (PLC) или микроконтроллер. Ниже представлены краткие характеристики обеих альтернатив:

2.1.1 Выбор контроллера

Программируемый логический контроллер (PLC)

Программируемые контроллеры на рынке имеют различные виды и большие различия, но они обладают следующими общими характеристиками:

- Высокая устойчивость к воздействию внешних помех, высокая надежность

В связи с высокими требованиями к надежности электромеханического оборудования в промышленном производстве, оно должно быть способным работать непрерывно в сложных условиях, таких как высокие температуры, высокое давление, влажность, наличие металлической пыли, высокочастотные электромагнитные поля и другие агрессивные среды. Это требует высокой стойкости к воздействию внешних помех. Среднее время наработки до отказа (MTBF) для современных программируемых логических контроллеров обычно составляет более 100 000 часов, и, например, серии F1 и F2 японской компании Mitsubishi достигают более 300 000 часов.

- Программирование управления просто и легко освоить

Программирование контроллера осуществляется с использованием языка логических схем и простых команд, ориентированных на промышленное управление. Этот язык нагляден, интуитивен, прост и легко усваивается; достаточно нескольких часов для овладения методами программирования.

- Хорошее взаимодействие с пользователем, удобное использование

Во время работы программируемого контроллера его панельный дисплей

отображает различные состояния и данные в процессе работы, что облегчает операторам следить за состоянием оборудования. При возникновении сбоев система способна обеспечить оперативное реагирование для незамедлительного устранения неполадок.

– Программы управления легко поддерживаются и ремонтируются

Программируемый контроллер оснащен памятью для хранения пользовательских программ, созданных операторами, и легко поддается чтению, записи, редактированию и модификации. Кроме того, он обладает высокой способностью самодиагностики, позволяя выявлять собственные сбои в любое время и предоставлять информацию о них на дисплей для оператора. Это позволяет оператору определить причину сбоя и предпринять эффективные меры по устранению.

– Компактная структура, легкость интеграции

С развитием технологии интегральных схем уровень интеграции программируемых контроллеров постоянно увеличивается, что приводит к уменьшению их размеров. Они становятся компактными, легкими, мощными и обладают высокой стойкостью к воздействию внешних помех. Это делает их легкими для установки внутри оборудования, образуя интегрированные электромеханические продукты, что является тенденцией будущего развития.

Основываясь на перечисленных выше преимуществах, программируемые контроллеры широко применяются в промышленном производстве, особенно на крупных автоматизированных производственных линиях. Однако в контексте данной темы исследования размеры программируемого контроллера все еще сравнительно большие, что не позволяет размещать его внутри робота. Кроме того, рыночная стоимость программных контроллеров относительно высока, поэтому данная тема исследования не принимает во внимание этот вариант.

микроконтроллер

В настоящее время на рынке наиболее широко используемой микросхемой является серия MCS-51, производимая компанией Intel. Она включает в себя три основных модели - 8031, 8051, 8751, а также соответствующие трем

низкопотребляющим моделям 80C31, 80C51, 87C51.

Серия микроконтроллеров MCS-51 обладает следующими преимуществами:

- Встроенный битовый процессор, способный выполнять битовые логические операции.

Серия 51 включает в себя полный комплект внутренней битовой операционной системы, осуществляющей связь между аппаратным и программным обеспечением, известной как битовый процессор или булев процессор. Этот процессор обрабатывает биты и позволяет не только выполнять операции сброса, передачи, установки и тестирования определенных битов в регистрах микроконтроллера, но также проводить битовые логические операции. Это является отличительной чертой, которую не встретишь у многих других микроконтроллеров.

- В оперативной памяти (RAM) выделены двойные функциональные зоны, что делает выполнение команд простым.

Серия 51 выделила во внутренней оперативной памяти (RAM) двойной функциональный адресный интервал в размере 16 байт, который может использоваться как для обработки байтов, так и для битовой обработки, обеспечивая удобство и гибкость использования. Поскольку в процессе выполнения программы требуется создание флагов для обработки структур разветвленного исполнения, эти флаги нужно обнулять, устанавливать и проверять, чтобы определить нормальное направление выполнения ветвящейся программы, для реализации этой функции достаточно всего одной инструкции битовой операции.

Несмотря на вышеупомянутые преимущества MCS-51, из-за использования 8-битного процессора у него есть ограничения по точности управления. Учитывая, что объектом обработки для системы распознавания речи робота является звуковой сигнал, который необходимо преобразовать в управляющий сигнал, требуется дополнительный периферийный модуль для сбора и преобразования звука. Разработка такого внешнего модуля звукового

сбора и преобразования представляет собой сложную задачу. Даже если такой модуль будет разработан, это не гарантирует достижение требуемой точности управления. Поэтому в данной теме исследования не выбирается микроконтроллер серии MCS-51 в качестве контроллера.

На рынке также существует плата управления от Lingyang, которая использует микроконтроллер Arduino и использует систему команд $\mu'nSP^{\text{TM}}$. Он обладает следующими характеристиками:

- Мощная функциональность, высокая эффективность

Инструкционный формат системы команд $\mu'nSP^{\text{TM}}$ компактен, обеспечивает высокую скорость выполнения и предоставляет интерфейс для высокоуровневых языков программирования, что позволяет сократить период вторичной разработки.

- Низкое напряжение, низкое энергопотребление

Система $\mu'nSP^{\text{TM}}$ может работать в широком диапазоне напряжения, что позволяет ей нормально функционировать при низком напряжении от батареи, что делает ее подходящей для работы в полевых условиях. Производится с использованием технологии CMOS, добавляется программно активируемый слабый режим колебаний, режим простоя и режим отключения, что существенно снижает энергопотребление.

- Обладает функциональностью расширенного распознавания речи.

Этот микроконтроллер может быть использован для разработки обычной однокристалльной системы, и, благодаря наличию модуля распознавания речи, его также можно использовать для расширенной разработки системы распознавания речи. Модуль распознавания речи обладает собственными функциями API, включая функции воспроизведения для форматов S240, S480, MS01, A2000, функции записи и воспроизведения речи в формате DRV, а также соответствующие функции распознавания. Это существенно упрощает вторичную разработку и расширяет область применения этого микроконтроллера.

- Компактные размеры, легкий вес, высокая стойкость к помехам, высокая

надежность

На основе вышеупомянутых характеристик Lingyang, а также учитывая их низкую стоимость и невысокие требования к окружающей среде, что облегчает вторичную разработку речи, в ходе разработки данной темы выбрана данная модель в качестве основного контроллера.

2.1.2 Определение технологии распознавания речи

В настоящее время существует несколько способов сопоставления при использовании технологии распознавания речи. Среди них есть технологии, основанные на шаблонном методе, а также технологии, основанные на статистических моделях, последние более современные. Технология распознавания речи с использованием скрытых марковских моделей (НММ) относится к статистическим моделям. Она предоставляет технологию для автоматического построения системы распознавания на основе вероятности обучающих данных, структура которой показана на рисунке 1.

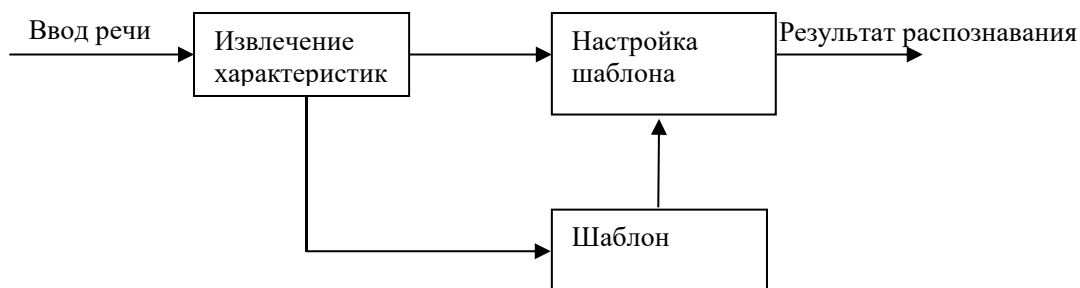


Рисунок 1 – Процесс реализации системы распознавания речи

Технология распознавания речи на основе скрытых марковских моделей (НММ) осуществляет анализ и статистическую обработку большого объема сэмплированных исходных данных речи для построения устойчивой статистической модели распознавания. Затем извлекаются характеристики из ожидаемой речи, и эти характеристики сравниваются с построенными моделями путем матчинга для получения результата распознавания. Таким образом, система способна адаптироваться к различным внезапным ситуациям в реальной

речи.

Технология распознавания речи на основе скрытых марковских моделей (НММ) проводит анализ и статистическую обработку большого объема сэмплированных исходных данных речи для построения устойчивой статистической модели распознавания. Затем извлекаются характеристики из ожидаемой речи, и эти характеристики сравниваются с построенными моделями путем матчинга для получения результата распознавания. Таким образом, система способна адаптироваться к различным внезапным ситуациям в реальной речи.

Процесс распознавания речи по своей сути представляет собой многомерный процесс распознавания образов, включающий в себя несколько этапов: предварительная обработка сигнала, извлечение характеристик, библиотека моделей речи, сопоставление образов и слепопроцессинг. В данной работе осуществляется распознавание речи на основе принципов неоднородных скрытых марковских моделей (НММ), и процесс распознавания изображен на рисунке 2.

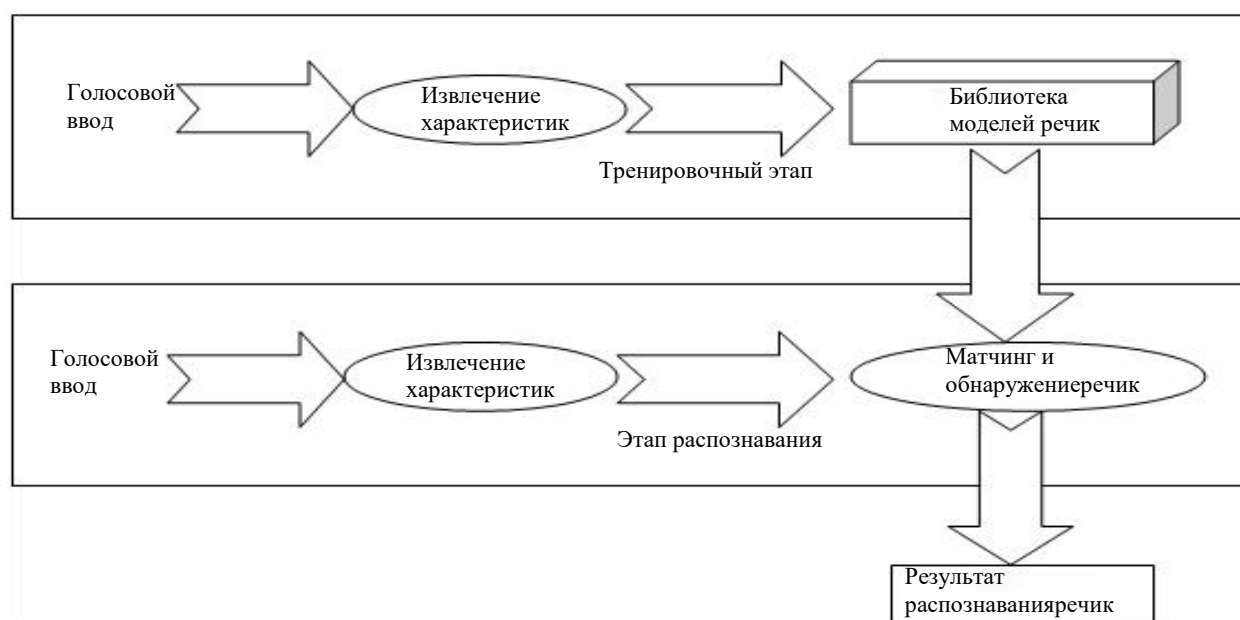


Рисунок 2 – Схема процесса системы распознавания речи

2.1.3 Выбор объекта робота

Для данной темы был выбран колесный игрушечный робот в качестве объекта исследования. Этот тип робота использует моторы для движения и включает в себя 8 моторов для передвижения и 1 мотор для поворота головы. Номинальное напряжение для работы робота составляет 8,4 В. Физический образец робота показан на рисунке 3.

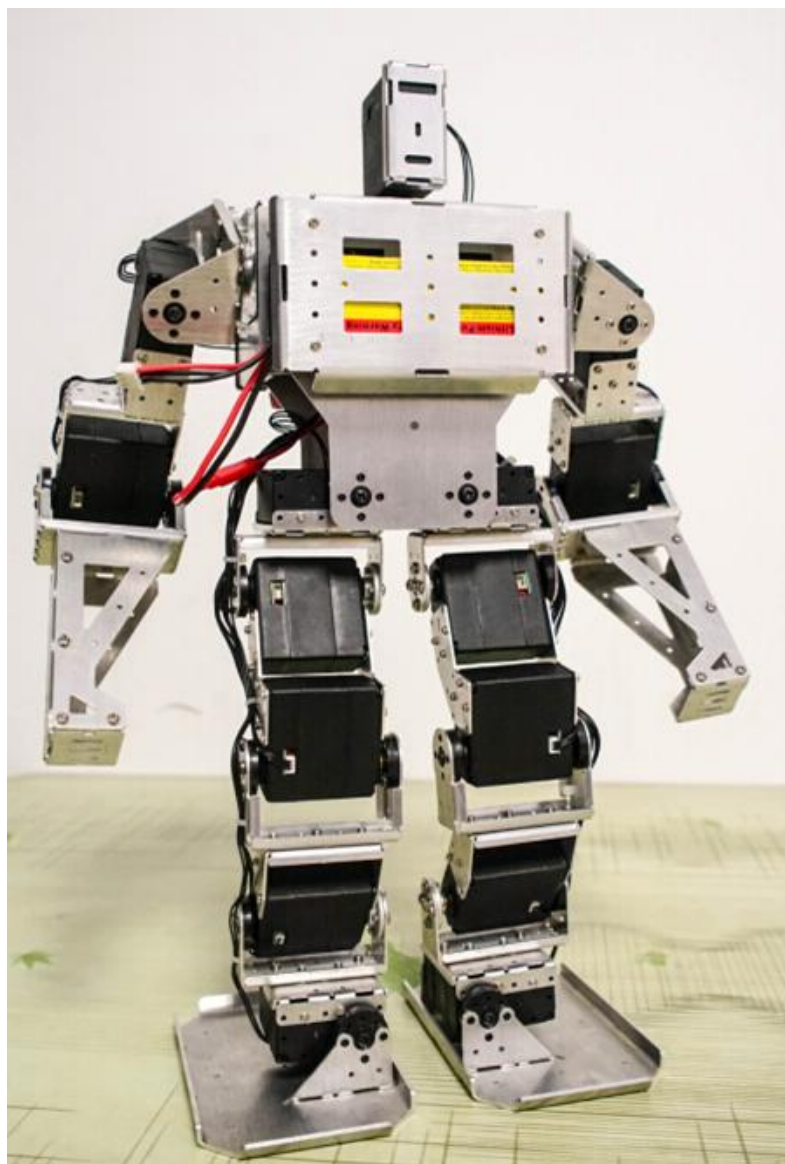


Рисунок 3 – Изображение физического образца робота

2.1.4 Проектирование схемы управления скоростью и привода робота

Различные движения робота осуществляются за счет работы двигателей, поэтому важно правильно настроить электропривод и регулировать скорость двигателей. Скорость вращения двигателя зависит от многих факторов, как

показано в следующих двух формулах:

Скорость вращения постоянного тока (ПТ) характеризуется следующим образом:

$$n = U - I_a R / C_e \Phi \quad (1)$$

где U – напряжение;

I_a – текущий ток;

R – сопротивление;

C_e – постоянная двигателя;

Φ – магнитный поток.

Когда электродвигатель находится в идеальном безнагрузочном режиме, его характеристика вращения может быть описана следующим образом:

$$n = U / C_e \Phi \quad (2)$$

где U – напряжение;

C_e – постоянная двигателя;

Φ – магнитный поток.

Из уравнений (1) и (2) видно, что изменение скорости вращения двигателя можно достичь путем изменения сопротивления R , напряжения U и магнитного потока Φ . Если для регулировки скорости выбрано изменение сопротивления R , увеличение сопротивления приведет к частичному распределению напряжения, что снизит эффективность регулировки скорости. Если для регулировки скорости выбрано изменение магнитного потока, магнитный поток Φ сложно контролировать. Поэтому наилучший выбор - это регулировка скорости работа путем изменения напряжения.

На данный момент часто используемый метод управления напряжением электродвигателя - это метод широтно-импульсной модуляции (PWM). Этот метод предполагает регулировку ширины импульсов путем изменения времени включения ключей в пределах постоянного периода, чтобы достичь регулировки скорости вращения двигателя.

Технология широтно-импульсной модуляции (PWM) использует мостовую схему в качестве усилительного устройства, применяет полностью цифровую технологию PWM для управления скоростью вращения двигателя и

вводит обратную связь по скорости для повышения точности управления и жесткости механических характеристик двигателя. Это позволяет двигателю достигать различных режимов работы и скоростей в процессе эксплуатации. Эта схема, работающая в режиме насыщения и отсечения, обеспечивает высокую эффективность; мостовая схема обеспечивает простое управление скоростью и направлением; электронные переключатели обладают высокой скоростью и стабильностью, что делает эту технологию широко используемой для регулировки скорости методом PWM.

Поскольку объектом исследования в данной теме является колесный игрушечный робот, его движения вперед, назад и повороты осуществляются путем управления выключателями двигателей в голове робота и на левой и правой ноге. Таким образом, достаточно управлять соответствующими выключателями двигателей.

Выбрана мостовая схема для управления двигателями робота. Принцип работы мостовой схемы показан на рисунке 4.

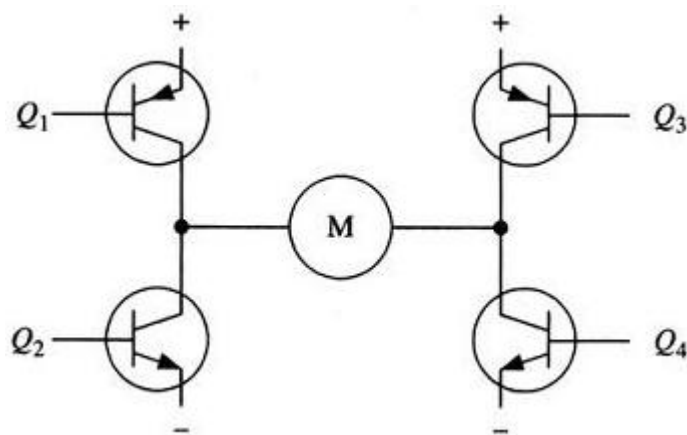


Рисунок 4 – Схема принципа работы мостового управления

Из приведенной выше схемы видно, что мостовая схема включает в себя 1 электродвигатель и 4 транзистора. Управление проводимостью одной пары транзисторов по диагонали позволяет контролировать направление тока через двигатель, что в свою очередь позволяет контролировать направление вращения двигателя. Когда транзисторы Q1 и Q4 включены, ток проходит через двигатель

от положительного источника питания через Q1 слева направо, затем через двигатель, и возвращается через Q4 к отрицательному источнику питания, образуя замкнутую цепь. В этот момент двигатель вращается по часовой стрелке. Когда транзисторы Q3 и Q2 включены, ток проходит через двигатель от положительного источника питания через Q3 справа налево, затем через двигатель, и возвращается через Q2 к отрицательному источнику питания, также образуя замкнутую цепь. В этот момент двигатель вращается против часовой стрелки.

2.1.5 Общее техническое решение системы

Для управления роботом в данной теме используется микроконтроллер Arduino. Связь между главной платой и роботом осуществляется через 32-канальную плату управления, которая полностью использует ресурсы IOB7-IOB15. Кроме того, в системе используется динамик. Общая схема системы представлена на рисунке 5.

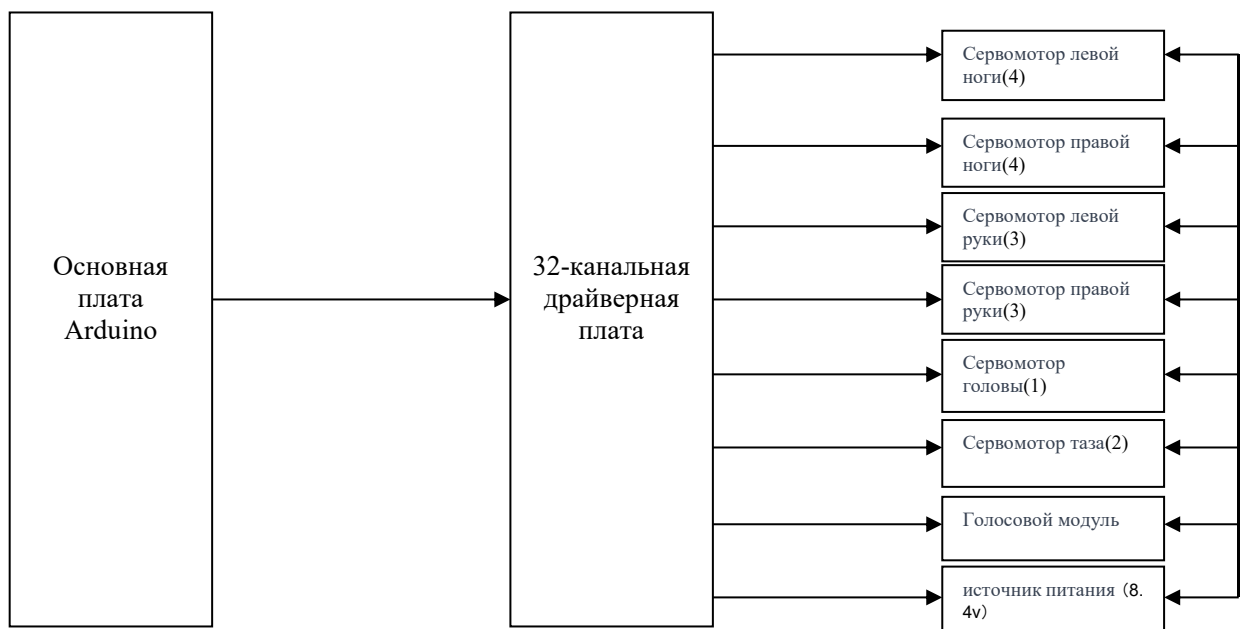


Рисунок 5 – Схема общего проекта

2.2 Проектирование аппаратной части системы

2.2.1 Проектирование цепей управления

В соответствии с общим проектом системы выбрана микроконтроллерная плата Arduino в качестве основного контроллера системы распознавания речи для робота. В настоящее время предоставляется краткое описание схемы управления различными блоками на главной плате контроллера.

Обзор основной платы Arduino

Arduino основная плата оснащена 16-битным микроконтроллером с использованием 16-битного микропроцессора μ nSPTM, который обладает высокой скоростью обработки данных и способен быстро обрабатывать сложные цифровые сигналы. Сама физическая структура изображена на рисунке 6.

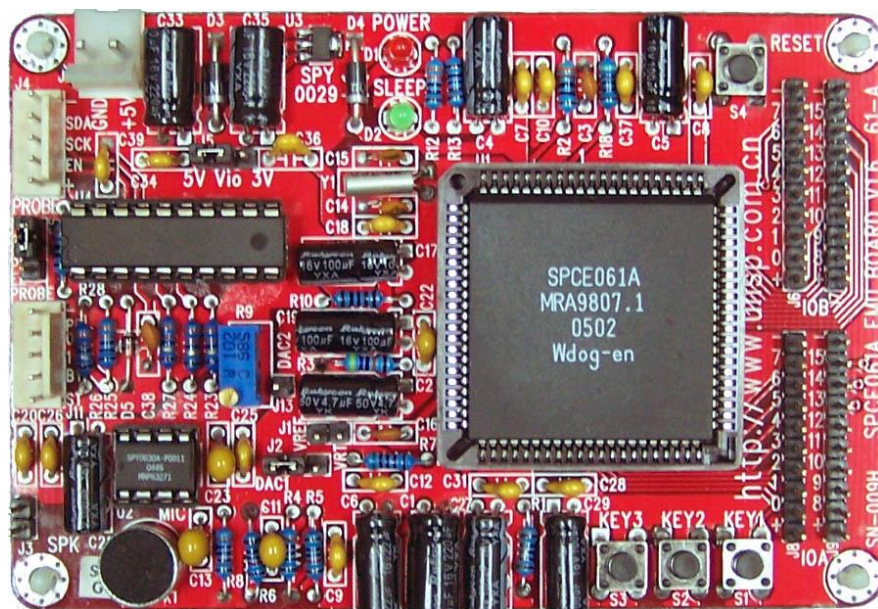


Рисунок 6 – Физическая схема основной платы Arduino

Этот тип основной платы имеет два вида упаковки микросхем: одна с 80 выводами, в форм-факторе PLCC80, и другая с 84 выводами, в форм-факторе LQFP84. Изображение расположения 84 выводов представлено на рисунке 7.

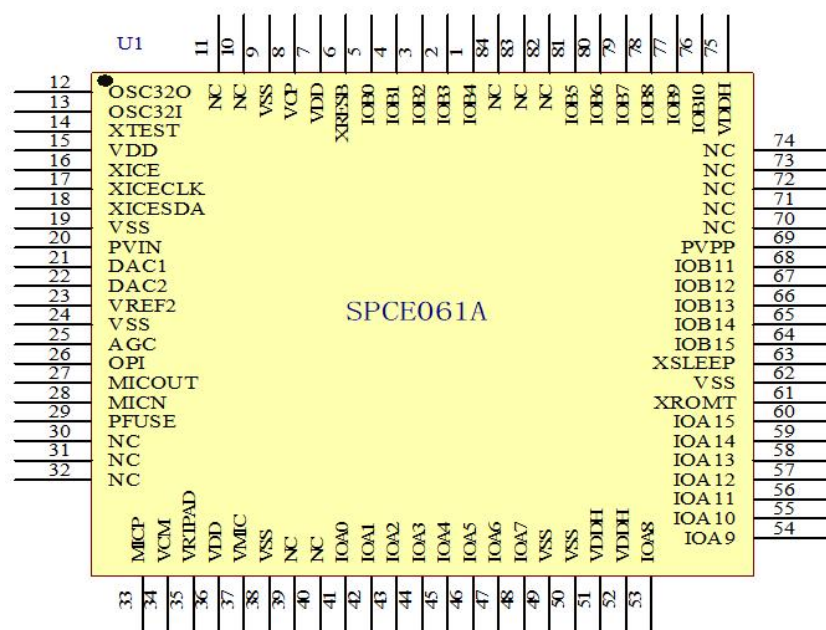


Рисунок 7 – SPCE061A микроконтроллер схема выводов

В корпусе LQFP84 имеется 15 свободных выводов. Описания функций остальных выводов представлены в таблице 1.

Таблица 1 – Таблица описания функций выводов

ВЫВОД	функция
IOA0~IOA15(41~48,53,54~60)	I/O A, Всего 16
IOB0~IOB15(5~1,81~76,68~64)	I/O B, Всего 16
OSCI (13вывод)	Вход для осциллятора. В кварцевом режиме это один из входов для кварцевого элемента.
OSCO (12вывод)	Выход для осциллятора. В кварцевом режиме это один из выходов для кварцевого элемента.
RES_B (6вывод)	Вход сброса. Если на этот вывод поступает низкий уровень, то контроллер будет сброшен.
ICE_EN (16вывод)	ICE-разрешающий вывод, подключен к выводу управления отладочного инструмента PROBE (ICE_EN).
ICE_SCK (17вывод)	ICE-вывод тактового сигнала, подключен к тактовому выводу отладочного инструмента PROBE (ICE_SCK).
ICE_SDA (18вывод)	ICE-вывод данных, подключен к выводу данных отладочного инструмента PROBE (ICE_SDA).
PVIN (20вывод)	Вывод для установки секретности программы.
PFUSE (29вывод)	Вывод для установки секретности программы.
DAC1 (21вывод)	Аудио выход, канал 1.
DAC2 (22вывод)	Аудио выход, канал 2.
VREF2 (23вывод)	Выход 2V опорного напряжения.
AGC (25вывод)	Выход 2V опорного напряжения.
OPI (26вывод)	Второй вход для микрофона.
MICOUT (27вывод)	Выход первого операционного усилителя микрофона.
MICN (28вывод)	Отрицательный вход микрофона.

MICP (33вывод)	Положительный вход микрофона.
VRT (35вывод)	Вход для внешнего опорного напряжения A/D преобразования. Он устанавливает верхний предел входного напряжения для A/D преобразования.
VCM(34вывод)	Выход опорного напряжения АЦП.
VMIC(37вывод)	Питание микрофона.
SLEEP (63вывод)	Вывод для индикации режима сна. Когда ЦПУ находится в режиме сна, этот вывод выдаёт высокий уровень.
VCP (8вывод)	Вход для изменения ёмкости фазового контура управляемого генератора.
XROMT, PVPP, XTEST(61, 69, 14вывод)	Точка для заводского тестирования, оставить открытой.
VDDH (51, 52, 75вывод)	Опорный уровень I/O. Подключение 5V опорного напряжения здесь делает выходы и входы I/O на уровне 5V.
VDD (7вывод)	Питание фазового контура управляемого генератора.
VSS (9вывод)	Земля фазового контура управляемого генератора.
VSS (19, 24вывод)	Аналоговая земля.
VSS (38, 49, 50, 62вывод)	Цифровая земля
VDD (15, 36вывод)	Цифровое питание.

– Проектирование схемы тактового сигнала

Работа всех функциональных компонентов на главной плате осуществляется на основе управляющих сигналов, а частота тактового сигнала напрямую влияет на скорость микроконтроллера. Качество часовой схемы также непосредственно влияет на стабильность системы микроконтроллера.

Часовая схема μ 'nSPTM использует контур кварцевого осциллятора, схема подключения часовой схемы на главной плате представлена на рисунке 8. Рекомендуется использовать внешний кварцевый резонатор с частотой 32768 Гц, так как точность внутренней схемы осциллятора RC не такая точная, как внешнего кварца.

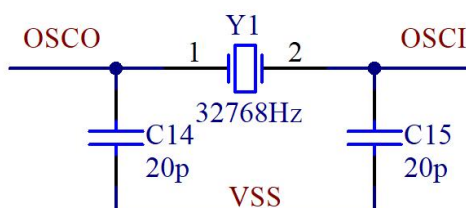


Рисунок 8 – Схема подключения часовой схемы

SPCE061A, делая деление от источника реального времени с частотой 32768 Гц, предоставляет различные прерывания от реального времени. Например, прерывание IRQ5_2Hz, используемое для пробуждения, указывает, что система пробуждается каждые 0.5 секунды, что может служить точной временной базой.

– Дизайн схемы сброса

Сброс - это операция инициализации микроконтроллера. Сбросом называется операция, которая в каком-то виде устанавливает значения всех регистров в памяти микроконтроллера в их начальные значения. Схема сброса основной платы представлена на рисунке 9. Подача низкого уровня на конце RESB вызывает сброс. Эта схема имеет две функции: ручной сброс и сброс при включении питания.

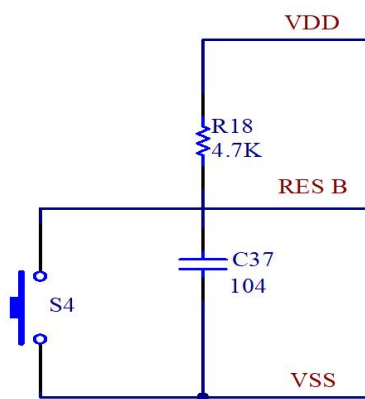


Рисунок 9 – Схема подключения схемы сброса

– Минимальная система микроконтроллера

Минимальная система схемы подключения изображена на рисунке 10. Подключение к выводам OSC0 и OSC1 кварцевого резонатора и конденсатора для резонанса позволяет системе работать. Затем, на вход VCP блока управляемого фазового контура подключаются соответствующие конденсатор и резистор. Другие неиспользуемые выводы питания и земли подключаются к сглаживающим конденсаторам емкостью 0.1 мкФ для улучшения устойчивости к помехам.

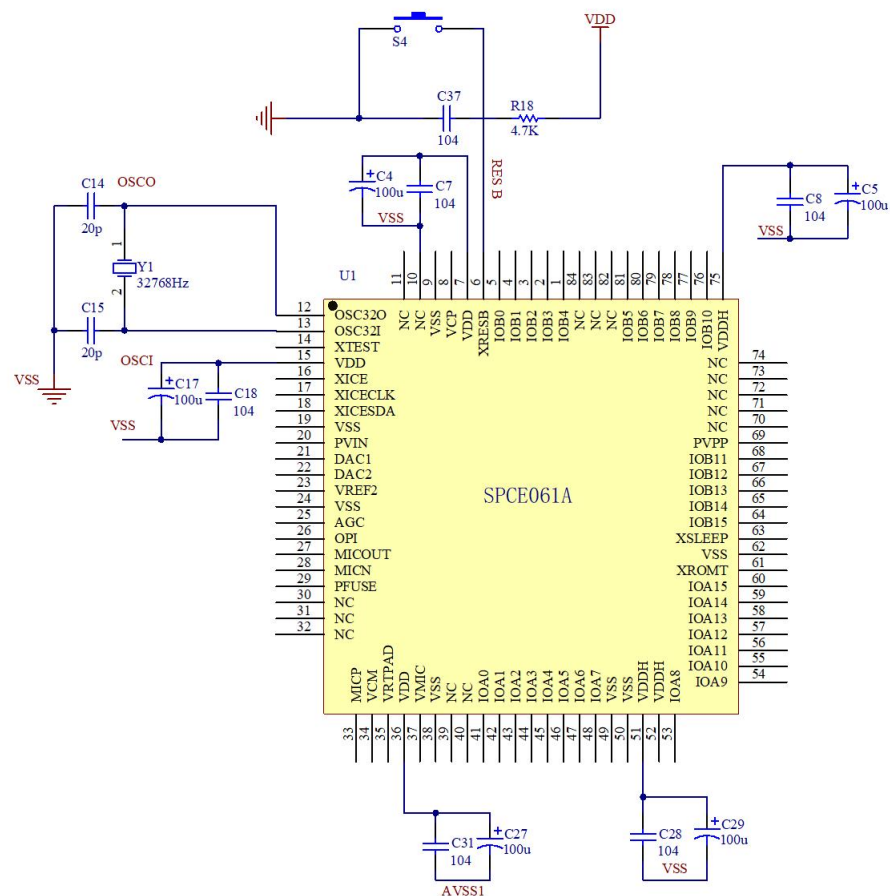


Рисунок 10 – Минимальная система микроконтроллера

2.2.2 Проектирование цепей управления двигателями робота

Приводная схема использует мощные транзисторы, собранные в мостовую конфигурацию Н-мост, для управления моторами, что позволяет реализовать вращение мотора вперед и назад. Эти моторы включают два для движения и один для поворота головы. Также используется транзистор для привода мотора с односторонним вращением, включая мотор для ускорения и мотор для запуска, схема привода которого довольно проста. Приводная схема показана на рисунке 11. В данной схеме выходы IOB8 и IOB9 соединены с соответствующими выводами микроконтроллера.

2.2.3 Схема сбора речевых сигналов

Микросхема SPC061A обладает 10-битным ADC для аудио-входа mic_in. Встроенная цепь усиления для микрофона и функция автоматического контроля усиления (AGC - автоматическая регулировка усиления) предоставлены. AGC

автоматически регулирует уровень звука с микрофона, обеспечивая постоянный уровень громкости для слушателя, исключая недостатки изменения громкости звука из-за изменения расстояния между говорящим и микрофоном.

На рисунке 12 следующие обозначения соответствуют следующим выводам микроконтроллера SPCE061A: VMIC, AGC, MICP, MICN, MICOUT и OPI. Где:

- VMIC предоставляет питание для микрофона;
- AVSS1 - это аналоговая земля системы;
- AGC - это вывод для автоматической регулировки усиления входного аудиосигнала;
- X1 - это микрофон;
- MICP - положительный вход для микрофона;
- MICN - отрицательный вход для микрофона;
- MICOUT - выход первого операционного усилителя микроконтроллера SPCE061A;
- OPI - вход второго операционного усилителя микроконтроллера SPCE061A.

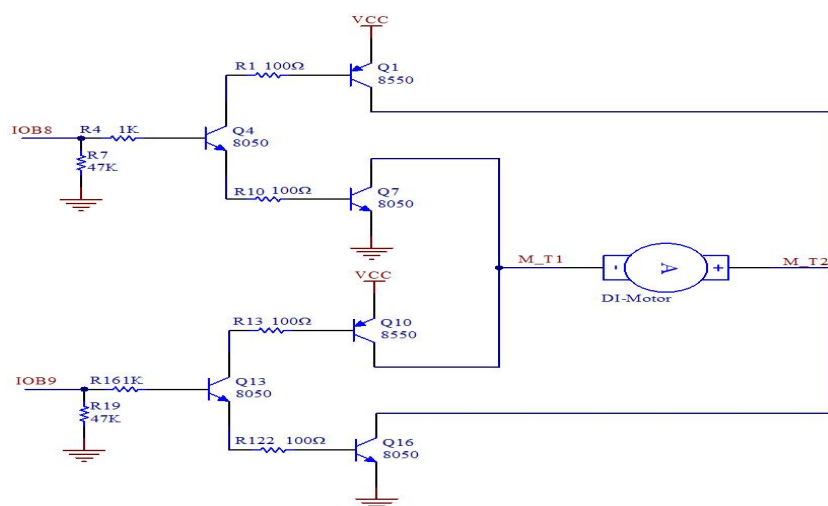


Рисунок 11 – Схема привода левой и правой ноги, а также шеи робота

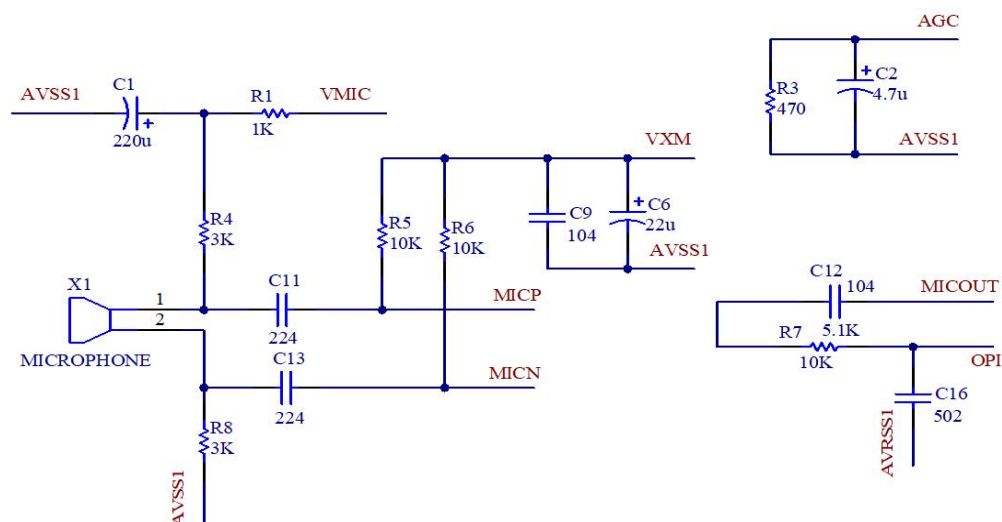


Рисунок 12 – Звуковая схема сбора сигнала

2.2.4 Электропитание системы

Способы питания главной платы

Система питания главной платы довольно гибкая и позволяет выбирать из нескольких вариантов подачи питания.

Питание от аккумулятора DC8.4V: питание предоставляется литиевым аккумулятором с напряжением 8,4 В. Постоянное напряжение 8,4 В стабилизируется непосредственно через трехконтактный интегральный стабилизатор напряжения SPY0029 на главной плате, обеспечивая всей главной плате напряжение 8,4 В.

Прямой стабилизатор постоянного тока главной платы

Принципиальная схема цепи стабилизации постоянного тока на главной плате показана на рисунке 13. Независимо от того, используется ли питание от аккумулятора или внешнее постоянное стабилизированное питание, оно подается через разъем питания J10. Поступивший постоянный ток напрямую подается на трехконтактный интегральный стабилизатор U3 (SPY0029), который стабилизирует напряжение до стандартных 8,4 В и подает его ядру SPCE061A для использования.

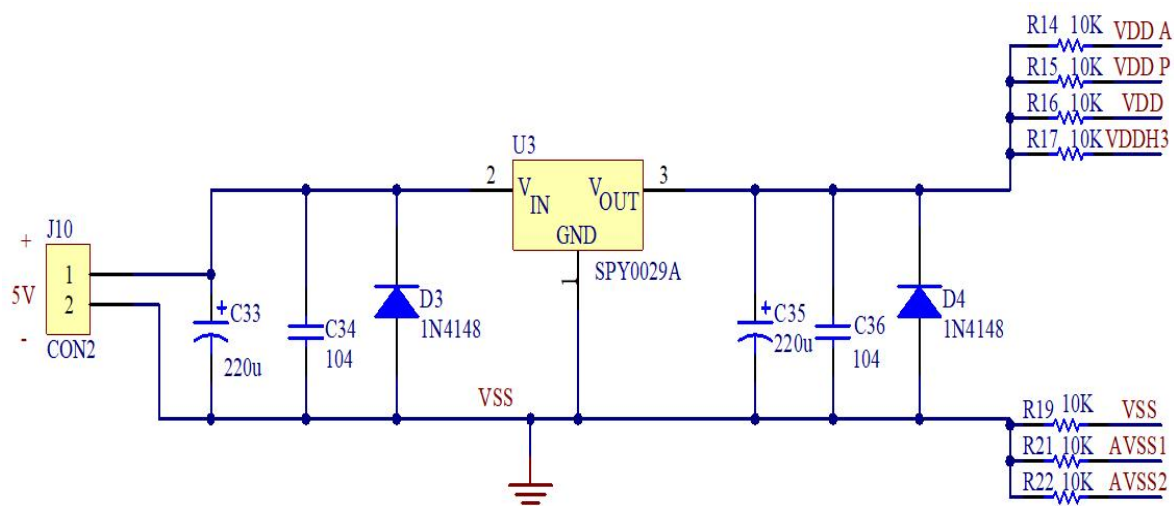


Рисунок 13 – Схема принципиальная цепи стабилизации постоянного тока главной платы

2.2.5 Проектирование системы управления голосовыми командами

Для реализации системы управления голосовыми командами необходимо разработать программное обеспечение, которое будет обрабатывать входные аудиосигналы, распознавать команды и передавать соответствующие управляющие сигналы на двигатели робота. Основные этапы разработки программного обеспечения включают:

- Обработка аудиосигналов: На этом этапе входной аудиосигнал проходит через фильтры для удаления шума и нормализации амплитуды. Затем сигнал разбивается на фреймы для дальнейшего анализа.

- Извлечение характеристик: Используются методы, такие как MFCC (Мел-частотные кепстральные коэффициенты), для извлечения ключевых характеристик из аудиосигнала.

- Распознавание команд: На основе извлеченных характеристик система сравнивает их с предварительно обученными моделями и определяет, какая команда была произнесена.

- Передача управляющих сигналов: После распознавания команды система отправляет соответствующие сигналы на двигатели робота для выполнения действий.

3 РЕАЛИЗАЦИЯ ПРОГРАММНЫХ МОДУЛЕЙ ДЛЯ ИНТЕЛЛЕКТУАЛЬНЫХ РОБОТОВ

3.1 Процесс работы системы

Общий обзор процесса

Работа системы распознавания голоса интеллектуального робота представляет собой многоэтапный процесс, включающий инициализацию оборудования, обучение голосовым командам, их распознавание и выполнение соответствующих действий. Система спроектирована с учетом ограничений аппаратной платформы и оптимизирована для эффективного использования вычислительных ресурсов.

Последовательность работы системы

– Инициализация системы

Включение питания робота осуществляется путем активации переключателя управления питанием на корпусе устройства.

После подачи питания микроконтроллер SPCE061A выполняет загрузку программного обеспечения из FLASH-памяти и инициализирует все периферийные модули (аудиовход/выход, двигатели, датчики).

Система переходит в режим ожидания голосовых команд или запуска процесса обучения (при первом включении или после сброса).

– Голосовая тренировка (обучение командам)

Данный этап требуется только при первом запуске системы или после выполнения аппаратного сброса.

Процесс обучения включает сбор голосовых данных пользователя, анализ акустических характеристик (частотные составляющие, длительность команд) и создание эталонных моделей для каждой команды.

Система запрашивает последовательное обучение 15 командам в строго определенном порядке:

- «название» (активация группы команд);
- «начать»;

- «ГОТОВ»;
- «танцевать»;
- «махать рукой»;
- «начать»;
- «идти вперед»;
- «назад»;
- «поворот направо»;
- «поворот налево»;
- «ГОТОВ»;
- «присесть»;
- «встать»;
- «три шага вперед»;
- «три шага назад».

Каждая команда должна быть произнесена пользователем дважды для повышения точности распознавания.

При успешном обучении система выводит сообщение: «Пожалуйста, перейдите к следующей команде!».

В случае ошибки распознавания (когда произнесённая фраза не соответствует заданному эталону) система запрашивает повторное произнесение команды до корректного распознавания.

– Организация команд в группы

Из-за ограниченного объема FLASH-памяти микроконтроллера (32 КБ) все 15 команд разделены на 3 группы по 5 команд в каждой.

Группировка команд:

- Группа 1: «название», «начать», «готов», «танцевать», «махать рукой»;
- Группа 2: «начать», «идти вперед», «назад», «поворот направо», «поворот налево»;
- Группа 3: «готов», «присесть», «встать», «три шага вперед», «три шага назад».

Для активации группы необходимо произнести ее название (первая

команда в группе).

- Режим распознавания голоса

После успешного обучения система выводит сообщение: «Теперь система перешла в режим распознавания голоса!».

Алгоритм работы в режиме распознавания:

Пользователь должен сначала активировать нужную группу команд, произнеся ее название.

Затем можно давать любую из 4 оставшихся команд в этой группе.

Например:

- «название» (активация группы 1);
- «танцевать» (выполнение действия).

После обработки команды контроллер:

Формирует управляющие сигналы для двигателей (включение/выключение, направление вращения)

Активирует звуковое сопровождение через динамик (подтверждение выполнения)

Возвращается в режим ожидания следующей команды

- Особенности реализации

Для экономии памяти используется алгоритм хранения не полных аудиозаписей, а только ключевых акустических характеристик команд.

При распознавании система сравнивает входной сигнал с эталонами используя метод DTW (Dynamic Time Warping) для компенсации различий в скорости произношения.

Время отклика системы не превышает 500 мс благодаря оптимизированным алгоритмам на ассемблере.

- Завершение работы

Для остановки системы необходимо отключить питание с помощью переключателя.

Перед выключением рекомендуется дать команду «готов», чтобы робот перешел в безопасное положение.

Все текущие параметры системы сохраняются в энергонезависимой памяти.

3.2 Проектирование основной программы

Проектирование программного модуля является важнейшей частью разработки интеллектуального робота, так как именно программное обеспечение определяет его функциональность, гибкость и адаптивность. Все функции интеллектуального робота реализуются посредством программирования однокристального микрокомпьютера, который выступает в роли центрального управляющего элемента.

– Среда разработки и языки программирования

Средой компиляции программного обеспечения является интегрированная среда разработки (IDE), обеспечивающая удобный интерфейс для написания, отладки и тестирования кода. Основными языками программирования, используемыми в данном проекте, являются язык Си и язык ассемблера. Язык Си применяется для реализации высокоуровневых алгоритмов и логики управления, в то время как ассемблер используется для оптимизации критичных по времени выполнения участков кода, а также для работы с регистрами микроконтроллера.

– Структура программного обеспечения

При проектировании программного обеспечения применяется модульный подход, что значительно упрощает процесс разработки и последующего сопровождения кода. Основная функция программы обычно вызывает подфункции, каждая из которых отвечает за выполнение конкретной задачи. Подфункции определяют функциональные модули, такие как обработка голосовых команд, управление движением, обработка прерываний и другие. Основная функция координирует работу всех модулей, достигая различных результатов в зависимости от входных данных и текущего состояния системы.

– Функциональные модули

В системе выделены следующие ключевые модули:

Функция заголовка – определяет основные параметры системы, включая

настройки прерываний и управление памятью.

Модуль голосовых ресурсов – содержит данные, необходимые для синтеза и распознавания речи, а также функции их обработки.

Модуль обучения голосу – позволяет роботу запоминать новые команды и ассоциировать их с определенными действиями.

Модуль распознавания голоса – анализирует поступающие аудиосигналы и преобразует их в команды для выполнения.

Модуль воспроизведения голоса – отвечает за генерацию речевых сообщений, например, подтверждение выполнения команды или запрос дополнительных инструкций.

Модуль движения робота – управляет моторами и сервоприводами, обеспечивая перемещение робота в пространстве.

Модуль обхода препятствий – использует данные с датчиков для обнаружения препятствий и корректировки траектории движения.

Модуль обслуживания прерываний – обрабатывает внешние события, такие как сигналы от датчиков или таймеров, обеспечивая реакцию системы в реальном времени.

– Взаимодействие с аппаратной частью

Интерфейсы, используемые в программном обеспечении робота, а также определение аппаратных ресурсов (например, микросхемы управления двигателями, АЦП для обработки сигналов датчиков) предоставляются средой IDE. Это возможно благодаря тому, что IDE включает в себя обширные библиотеки функций, а также инструменты для работы с языками Си и ассемблера. Библиотеки содержат готовые решения для типовых задач, таких как управление ШИМ, обработка UART-сообщений или работа с таймерами, что ускоряет процесс разработки.

– Процесс обучения и работы робота

Основным этапом программирования является обучение робота. Пользователь может записывать голосовые команды, которые затем сохраняются в памяти устройства. После успешного обучения робот переходит в

режим распознавания голоса, где он анализирует поступающие команды и выполняет соответствующие действия. Например, при получении команды "вперед" робот начинает движение, а при обнаружении препятствия активирует алгоритм обхода.

3.2.1 Общая схема проектирования системы голосового управления

Проектирование программной подсистемы заключается в разработке системных программ для управления компонентами микроконтроллера на основной плате и встроенными функциями аппаратной подсистемы. Эти программы вызывают и контролируют аппаратные компоненты, заставляя их выполнять соответствующие действия в соответствии с требованиями программы. Для создания программного обеспечения необходимо выбрать язык программирования.

Поскольку система распознавания голоса робота является небольшой системой, которая требует, чтобы робот выполнял определенные действия в соответствии с заданными командами, и эти действия ограничены и независимы друг от друга, человек может организовать эти действия в определенной последовательности. Другими словами, выполнение действий в определенном порядке позволяет достичь ожидаемых результатов. Такой подход соответствует модульному дизайну структурных языков программирования. Поэтому при выборе языка программирования в первую очередь следует учитывать структурные языки. Кроме того, поскольку программа должна напрямую взаимодействовать с аппаратным обеспечением на низком уровне, язык C, являясь структурным языком программирования, также позволяет напрямую управлять аппаратным обеспечением, как и язык ассемблера, выполняя операции с битами, байтами и адресами. Таким образом, для вторичной разработки системы выбран язык C.

На основе структурного и модульного подхода к проектированию программного обеспечения программная подсистема разделена на модуль основной программы, модуль голосовой тренировки и модуль распознавания голоса.

В основной функции программы вызываются соответствующие функции для выполнения тренировки голоса конкретного человека. После успешной тренировки система переходит к распознаванию голоса и выполняет соответствующие операции в зависимости от распознанных команд. Блок-схема процесса основной программы показана на рисунке 14.

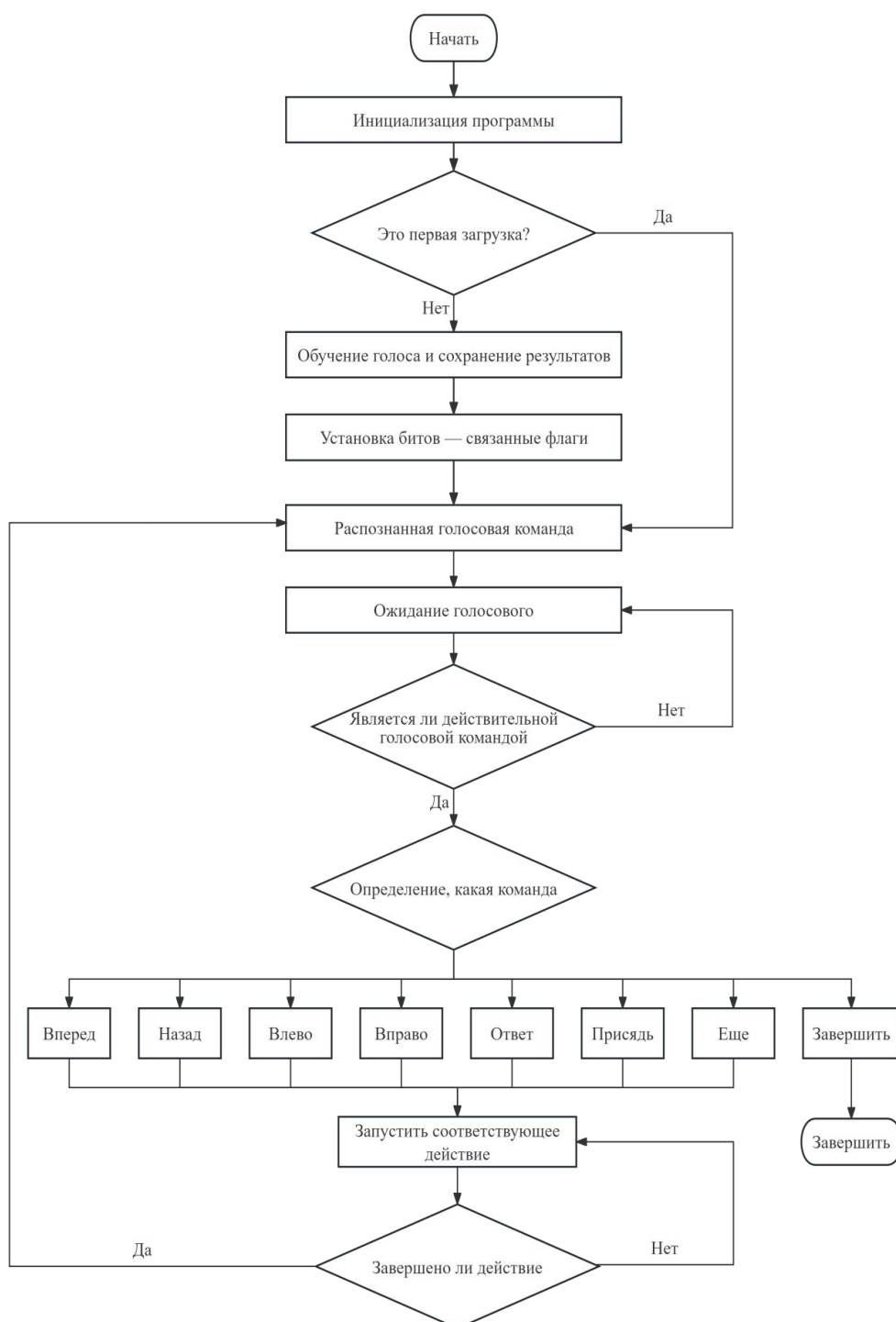


Рисунок 14 – Диаграмма основного процесса программы системы

3.2.2 Модуль голосовой тренировки

Библиотечная функция `BSR_Train(int_WordID, intTrainMode)` помогает реализовать голосовую тренировку. Здесь:

Параметр 1 (`int_WordID`) — это порядковый номер команды, который является уникальным для каждой группы тренировочных фраз.

Параметр 2 (`intTrainMode`) — количество тренировок. Если значение равно 1, это означает, что тренировка проводится один раз; если значение равно 2, тренировка проводится дважды.

Чтобы обеспечить четкость и безошибочность команд, каждая голосовая команда проходит двойную тренировку и подтверждение. Только в случае успешного завершения обеих тренировок можно переходить к следующему этапу. Часть блок-схемы процесса представлена на рисунке 15.

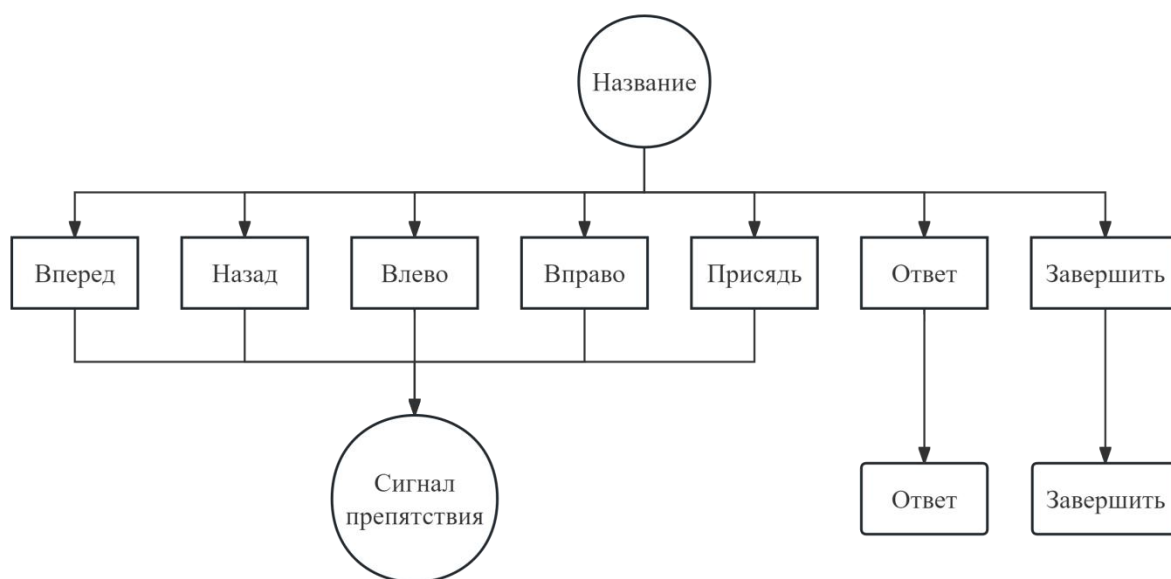


Рисунок 15 – Блок-схема части процесса тренировки голосовых команд

3.2.3 Модуль распознавания голоса

При проектировании модуля распознавания голоса сначала выполняется инициализация микросхемы распознавания голоса. Голосовой сигнал поступает через МІС-канал аналого-цифрового преобразователя (АЦП). После воспроизведения системой звукового сигнала, тренирующий человек подает

роботу голосовые команды. Система собирает эти голосовые данные и команды, а затем сравнивает их с уже сохраненными в системе голосовыми командами. Если собранный голос соответствует голосу тренирующего, происходит сравнение голосовых команд, и на основе результатов сравнения генерируются соответствующие управляющие сигналы. В противном случае управляющие сигналы не генерируются. Модуль распознавания голоса показан на рисунке 16.

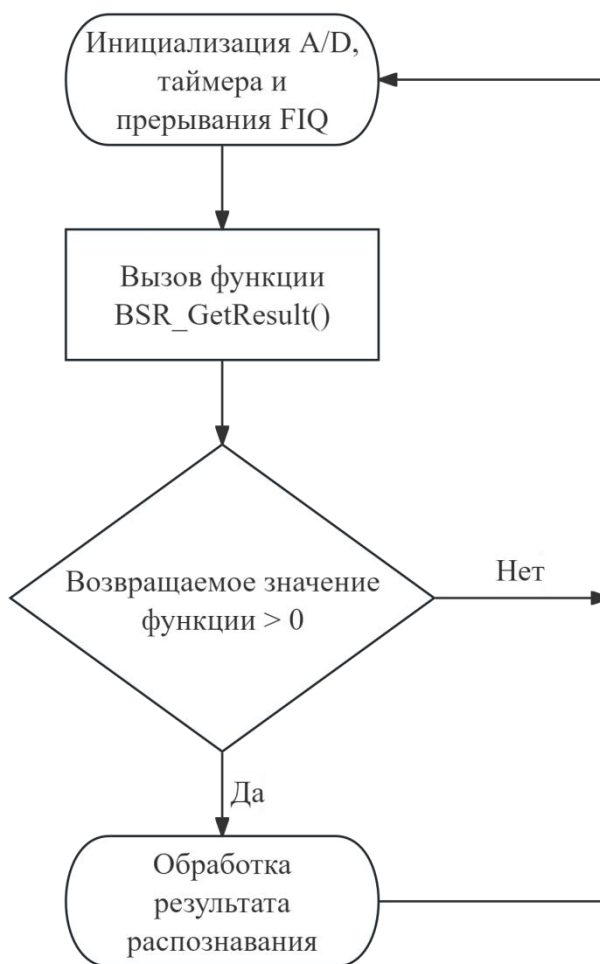


Рисунок 16 – Блок-схема модуля распознавания голоса

3.2.4 Проектирование подпрограммы обслуживания прерываний

Обслуживание прерываний является одной из ключевых функций в системах на основе микроконтроллеров, играя особенно важную роль в реальном времени и аварийной обработке. Подпрограмма обслуживания прерываний (ISR) — это специально разработанный сегмент программы,

предназначенный для реагирования на события прерываний и решения задач вызова и возврата при возникновении неожиданных событий. В проектировании и реализации голосового управления роботизированной системы эффективное управление обслуживанием прерываний напрямую влияет на скорость отклика и стабильность системы. Это показано на рисунке 17.



Рисунок 17 – Диаграмма работы программы обработки прерываний

Возьмем, к примеру, микроконтроллер SPCE061A, который обладает мощными возможностями обработки прерываний и способен быстро реагировать на внешние события. Когда система обнаруживает сигнал прерывания, она сначала сохраняет текущее состояние программы, содержимое регистров и другую важную информацию в стеке, чтобы обеспечить корректное восстановление состояния и продолжение выполнения основной программы после завершения обработки прерывания. Этот шаг крайне важен, поскольку

потеря любой информации может привести к аномальной работе системы.

Затем система разрешает прерывания, позволяя обрабатывать запросы прерываний с более высоким приоритетом. Этот механизм гарантирует, что система может реагировать на более срочные события даже во время обработки текущего прерывания. После этого система предоставляет услугу для соответствующего события, то есть выполняет подпрограмму обслуживания прерываний (ISR). В системе голосового управления это может включать сбор, обработку и распознавание голосовых сигналов. Высокая производительность микроконтроллера SPCE061A обеспечивает выполнение этих операций за минимальное время.

После завершения выполнения подпрограммы обслуживания прерываний система восстанавливает сохраненное состояние программы и содержимое регистров, чтобы программа могла продолжить выполнение с точки прерывания. Наконец, система возвращается к месту прерывания и продолжает выполнение прерванной программы.

В голосовом управлении роботизированных систем быстрая реакция и эффективная обработка прерываний особенно важны. Благодаря грамотному проектированию подпрограмм обслуживания прерываний система может точно распознавать команды в сложной акустической среде и своевременно реагировать на них. Это не только повышает реальное время работы системы, но и усиливает ее надежность и удобство для пользователя.

Обслуживание прерываний является важной функцией микроконтроллера. Подпрограмма обслуживания прерываний предназначена специально для обработки прерываний, решения задач вызова и возврата подпрограмм, а также для обработки чрезвычайных ситуаций. После завершения обработки чрезвычайной ситуации система возвращается к месту прерывания.

3.2.5 Модуль воспроизведения голоса

Микроконтроллер SPCE061A компании Lingyang обладает мощными функциями сжатия и воспроизведения голоса. Подпрограмма воспроизведения голоса в основном выполняет инициализацию ЦАП (D/A-преобразователя),

таймера и голосовых функций, извлекает предварительно сжатые голосовые данные из ПЗУ (ROM), декомпрессирует их и помещает в список воспроизведения. Это показано на рисунке 18.

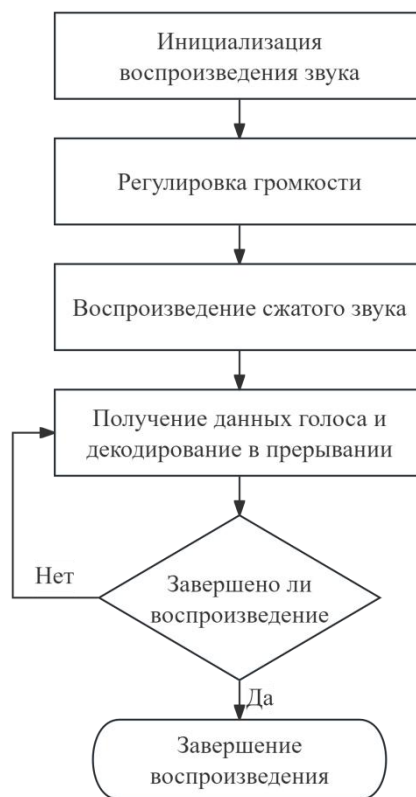


Рисунок 18 – Блок-схема модуля воспроизведения голоса

Интеллектуальный робот в основном выполняет такие действия, как движение вперед, назад, поворот налево, поворот направо и вращение на месте, путем управления соответствующими двигателями постоянного тока. Два двигателя контролируют его движение, причем левая и правая ноги управляются отдельно, то есть каждая нога управляется отдельным двигателем. Порт В микроконтроллера SPCE061A обеспечивает управляющие сигналы. Например, когда робот движется вперед, соответствующий вывод устанавливается на высокий уровень и удерживается в течение определенного времени, что приводит к вращению двигателя постоянного тока в прямом направлении. Используя аналогичный метод настройки, робот может выполнять такие действия, как движение назад, поворот налево, поворот направо и остановка. Движение головы робота контролируется с помощью выводов IOB13 и IOB14.

Блок-схема программы движения робота вперед показана на рисунке 19.



Рисунок 19 – Блок-схема процесса движения робота вперед

3.2.6 Определение состояния движения робота

Робот выполняет такие команды, как движение вперед, назад и другие, в соответствии с инструкциями конкретного человека, а в процессе движения автоматически избегает препятствий с помощью модуля обнаружения датчиков. Когда робот получает сигнал команды, микроконтроллер сначала анализирует эту команду.

Например, когда робот слышит команду "вперед" или "стоп", микроконтроллер сначала определяет команду, а затем управляет соответствующими двигателями робота для выполнения необходимого действия.

Когда роботу нужно повернуть, соответствующие двигатели получают сигнал от микроконтроллера и выполняют поворот на определенный угол. Когда требуется выполнить поворот головы, микроконтроллер передает управляющий

сигнал двигателю головы, заставляя его вращаться в прямом или обратном направлении. Микроконтроллер управляет выполнением и завершением всех действий. Когда робот выполняет операцию воспроизведения голоса, микроконтроллер вызывает программу обработки звука, декодирует соответствующие данные и воспроизводит голос.

В зависимости от различных команд микроконтроллер устанавливает соответствующие уровни высокого или низкого напряжения на своих I/O-портах, заставляя робота выполнять соответствующие движения. Соответствующие отношения показаны в таблице 2.

Таблица 2 – Соотношение между управлением роботом и выходными сигналами портов ввода-вывода (I/O):

IOB8	IOB9	IOB10	IOB11	Двигатель левой ноги	Двигатель правой ноги	Состояние
0	1	0	1	Вращение вперед	Вращение вперед	Вперед
1	0	1	0	Вращение назад	Вращение назад	Назад
1	0	0	1	Вращение назад	Вращение вперед	Поворот налево
0	1	1	0	Вращение вперед	Вращение назад	Поворот направо

3.2.7 Функция избегания препятствий

– Обзор функции избегания препятствий

Данный проект основан на автоматическом роботе с инфракрасными датчиками для избегания препятствий, использующем модульный подход к проектированию, который делится на три основных модуля:

Модуль каркаса: отвечает за создание общей структуры робота, включая расположение датчиков, контроллера и исполнительных механизмов.

Модуль обнаружения движения: использует инфракрасные датчики для обнаружения информации о препятствиях вокруг робота.

Исполнительный модуль: на основе данных модуля обнаружения управляет роботом для выполнения соответствующих действий по избеганию препятствий.

– Процесс реализации функции избегания препятствий

Процесс реализации функции избегания препятствий выглядит следующим образом:

Инициализация оборудования: после запуска системы сначала инициализируются инфракрасные датчики, драйверы двигателей и другое оборудование.

Обнаружение окружающей среды: робот начинает движение вперед, инфракрасные датчики непрерывно обнаруживают препятствия впереди.

Определение препятствий:

- Препятствие впереди: робот поворачивает направо;
- Препятствие справа: робот поворачивает налево;
- Препятствие слева: робот движется назад;
- Препятствия со всех сторон: робот останавливается.

Динамическое избегание препятствий: в процессе движения инфракрасные датчики непрерывно отслеживают окружающую среду, чтобы робот мог динамически корректировать свой путь.

- Подключение инфракрасных датчиков к микроконтроллеру

Сигнальные линии инфракрасных датчиков подключены к портам IOB0—IOB3 микроконтроллера.

Микроконтроллер определяет наличие препятствий впереди, считывая уровни сигналов на этих портах:

Высокий уровень: обнаружено препятствие.

Низкий уровень: препятствий нет.

- Проектирование алгоритма избегания препятствий

Основная идея алгоритма избегания препятствий заключается в динамической корректировке направления движения робота на основе данных датчиков. Конкретные шаги следующие:

Обнаружение впереди:

Если IOB0 обнаруживает препятствие, робот останавливается и поворачивает направо.

Обнаружение справа:

Если IOB1 обнаруживает препятствие, робот прекращает поворот направо и поворачивает налево.

Обнаружение слева:

Если IOB2 обнаруживает препятствие, робот прекращает поворот налево и движется назад.

Обнаружение сзади:

Если IOB3 обнаруживает препятствие, робот прекращает движение назад и переходит в состояние остановки.

– Блок-схема программы избегания препятствий

Блок-схема программы избегания препятствий, показанная на рисунке 20, включает следующие шаги:

– Начало: запуск системы, инициализация оборудования;

– Движение вперед: робот начинает движение вперед.

Обнаружение препятствий:

– Если обнаружено препятствие впереди, робот поворачивает направо;

– Если препятствие также справа, робот поворачивает налево;

– Если препятствие также слева, робот движется назад;

– Если препятствие также сзади, робот останавливается.

Циклическое обнаружение: робот продолжает отслеживать окружающую среду до тех пор, пока препятствия не исчезнут.

Завершение: выполнение задачи или остановка работы.

– Оптимизация функции избегания препятствий

Для повышения эффективности и надежности функции избегания препятствий можно внедрить следующие меры оптимизации:

Слияние данных с нескольких датчиков: использование ультразвуковых датчиков, камер и других устройств для повышения точности обнаружения препятствий.

Алгоритмы планирования пути: внедрение алгоритмов A или Dijkstra для более интеллектуального планирования пути.

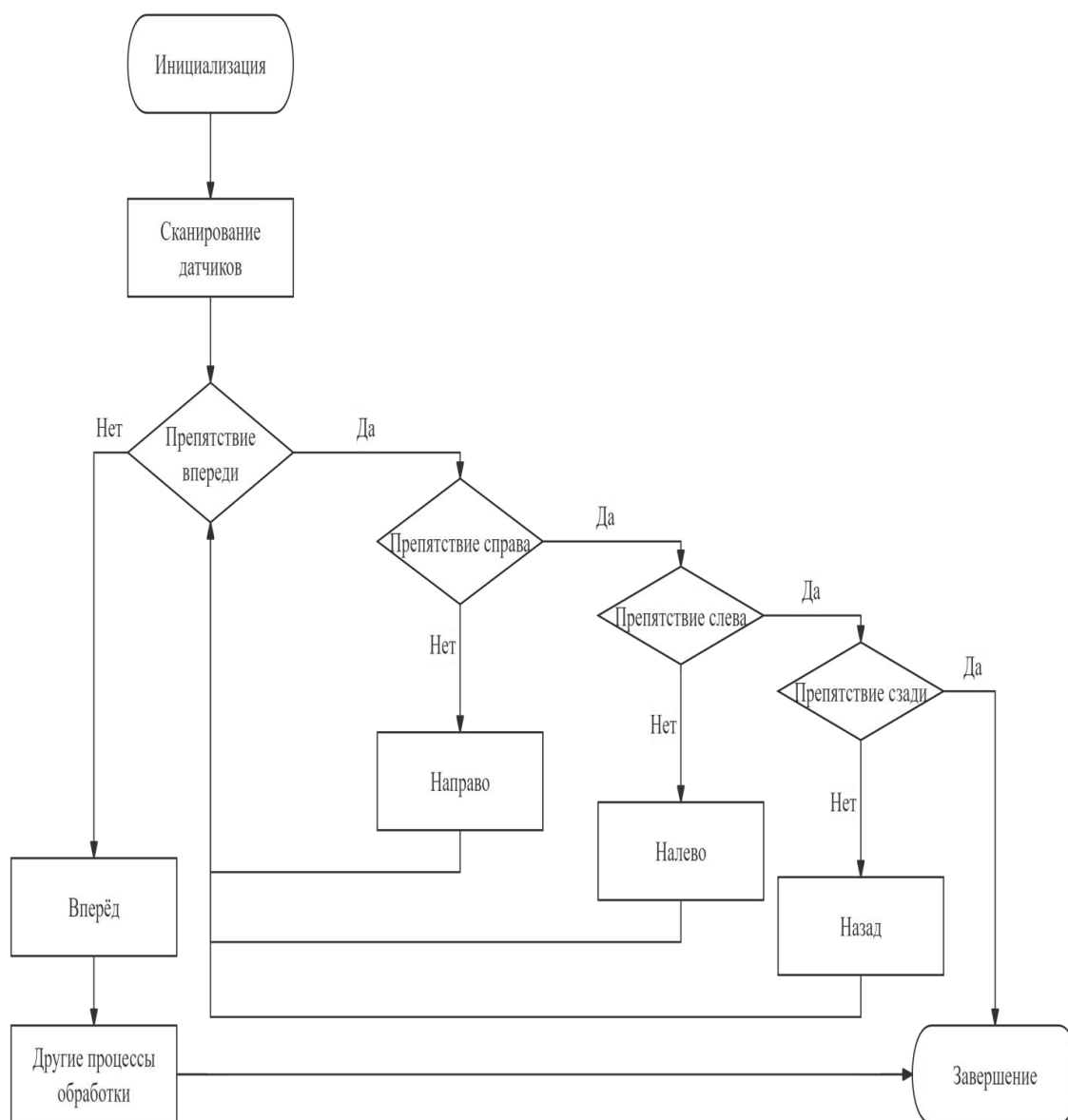


Рисунок 20 – Блок-схема программы избегания препятствий с использованием инфракрасных датчиков

Динамическая стратегия избегания препятствий: корректировка скорости и направления движения робота на основе расстояния до препятствий и их скорости.

– Сценарии применения

Функция избегания препятствий может широко применяться в следующих сценариях:

Домашние сервисные роботы: автономная навигация в домашних условиях, избегание столкновений с мебелью или стенами.

Промышленные инспекционные роботы: избегание оборудования и персонала на производственных объектах для выполнения задач инспекции.

Образовательные роботы: использование в качестве учебного инструмента для изучения технологий датчиков и принципов автоматического управления.

Благодаря вышеуказанным разработкам и оптимизациям робот может эффективно и надежно выполнять функцию автоматического избегания препятствий в сложных условиях.

3.3 Отладка, симуляция и выводы системы распознавания голоса

Отладка системы голосового робота в основном состоит из аппаратной и программной отладки. Зачастую в процессе программной отладки можно обнаружить множество аппаратных ошибок и исправить их. Обычный процесс отладки системы начинается с устранения аппаратных неисправностей, а затем продолжается совместной программной отладкой для дальнейшего устранения проблем. Таким образом, аппаратная и программная отладка системы тесно связаны и взаимозависимы.

3.3.1 Аппаратная отладка

– Обзор отладки аппаратного обеспечения

Отладка аппаратного обеспечения является неотъемлемой частью разработки интеллектуальных роботов. Ее основная цель заключается в обеспечении соответствия схемотехнического проектирования функциональным требованиям и возможности совместной работы с программным обеспечением. Отладка аппаратного обеспечения охватывает весь процесс проектирования схем, выбора компонентов, изготовления печатных плат (PCB) и тестирования прототипов. Систематические методы отладки позволяют эффективно выявлять и устранять проблемы в проектировании аппаратного обеспечения, обеспечивая стабильность и надежность робототехнической системы.

– Процесс отладки аппаратного обеспечения

Полный процесс отладки аппаратного обеспечения включает следующие ключевые этапы:

Проектирование принципиальной схемы:

На основе функциональных требований робота разрабатываются принципиальные схемы для каждого аппаратного модуля, включая модуль питания, модули датчиков (например, инфракрасные и ультразвуковые датчики), модуль управления двигателями и модули связи (например, UART).

Использование EDA-инструментов для создания схем, обеспечивая правильность логики схемы и четкость направления сигналов.

Выбор и закупка компонентов:

На основе принципиальной схемы выбираются подходящие компоненты, соответствующие требованиям по типу, спецификациям и параметрам. Например, длина волны инфракрасного датчика, допустимый ток драйвера двигателя и т.д.

При закупке необходимо уделять внимание качеству и источнику компонентов, избегая использования некачественных или поддельных деталей.

Изготовление тестового прототипа:

Перед изготовлением РСВ создается тестовый прототип с использованием макетной платы или универсальной РСВ.

Тестовый прототип позволяет проверить реализуемость схемы и выполнить предварительное тестирование функциональности модулей.

Отладка аппаратных модулей:

Последовательная отладка каждого модуля для обеспечения их корректной работы. Например:

Модуль питания: проверка стабильности выходного напряжения, отсутствие перенапряжения или пониженного напряжения.

Модули датчиков: проверка корректности выходных сигналов инфракрасных и ультразвуковых датчиков.

Модуль управления двигателями: проверка возможности вращения двигателя в прямом и обратном направлении, а также его остановки.

Использование мультиметра, осциллографа и других инструментов для измерения напряжения, тока и формы сигналов, проверки характеристик схемы.

Установка и проверка компонентов:

Установка компонентов на РСВ с проверкой правильности их корпусов, чтобы избежать ошибок, которые могут привести к некорректной работе.

Проверка качества пайки, обеспечение надежности соединений и отсутствия коротких замыканий.

Комплексная отладка аппаратного обеспечения:

Соединение всех модулей и выполнение комплексной отладки.

Использование программного обеспечения для моделирования и внутрисхемных эмуляторов для проверки функциональности аппаратного обеспечения и его совместимости с программным обеспечением.

– Методы поиска и устранения неисправностей в аппаратном обеспечении

В процессе отладки могут возникать различные неисправности. Ниже приведены основные шаги для их устранения:

Устранение логических ошибок:

Типичные проблемы: неправильные соединения, обрывы, короткие замыкания, вызванные ошибками проектирования или производственного процесса.

Методы устранения:

Сравнение изготовленной РСВ с принципиальной схемой для выявления несоответствий.

Проверка системы питания на наличие коротких замыканий или ошибок полярности.

Использование мультиметра для проверки ключевых сигнальных линий на наличие коротких замыканий или обрывов.

Устранение неисправностей компонентов:

Типичные проблемы: повреждение или неправильная установка компонентов.

Методы устранения:

Проверка типа и характеристик компонентов перед установкой.

Использование мультиметра или тестового оборудования для проверки

работоспособности компонентов.

Проверка качества пайки, чтобы исключить повреждение компонентов из-за ошибок пайки.

Отладка с использованием моделирования:

После успешной отладки аппаратного обеспечения используется среда разработки IDE и внутрисхемные эмуляторы для отладки программного обеспечения.

Использование инструментов моделирования для проверки совместимости программного и аппаратного обеспечения.

Пошаговая отладка для выявления проблем взаимодействия программного и аппаратного обеспечения.

Инструменты для отладки аппаратного обеспечения

В процессе отладки используются следующие инструменты:

Мультиметр: для измерения напряжения, тока и сопротивления.

Осциллограф: для анализа формы сигналов и характеристик схемы.

Логический анализатор: для проверки временных диаграмм и логических соотношений цифровых сигналов.

Программное обеспечение для моделирования: например, Keil, IAR, для отладки программного обеспечения и моделирования аппаратного обеспечения.

Внутрисхемные эмуляторы: для реальной отладки взаимодействия аппаратного и программного обеспечения.

Рекомендации по отладке аппаратного обеспечения

Проверка системы питания: источник питания является ключевым элементом схемы, необходимо обеспечить стабильность напряжения и отсутствие коротких замыканий.

Целостность сигналов: для высокочастотных сигналов важно учитывать согласование импеданса и отражения сигналов, чтобы избежать искажений.

Теплоотвод: для компонентов с высокой мощностью необходимо предусмотреть эффективное охлаждение, чтобы избежать повреждений из-за перегрева.

Заземление: качественное заземление снижает уровень шумов и повышает стабильность схемы.

3.3.2 Программная отладка

Программная отладка — это сложная задача, требующая значительных навыков и опыта. Основная ее цель — выявление проблем, их анализ и исправление ошибок. При анализе результатов тестирования мы видим только внешние проявления ошибок программы, в то время как внутренние причины их возникновения часто не имеют явной связи с внешними проявлениями. Программа состоит из множества операторов, структур данных и других элементов, и ошибка может находиться в любом из них. Проверка и устранение ошибок в таком большом количестве элементов — непростая задача. Поэтому программная отладка подчеркивает важность процесса анализа, направленного на поиск причин ошибок на основе их внешних проявлений. Этапы отладки следующие:

- Выявление внешних проявлений ошибки, анализ проблемы и определение места в программе, где возникает ошибка.
- Анализ соответствующего программного кода для выявления причины проблемы.
- Внесение необходимых изменений в дизайн и код.
- Повторная проверка, чтобы убедиться, что ошибка устранена и не появились новые ошибки.

Объем работы по отладке огромен, причем большая часть усилий сосредоточена на первом этапе. Считаем, что локализация ошибки на основе ее проявлений является наиболее важным этапом процесса отладки. После завершения аппаратной и программной отладки необходимо провести симуляцию и отладку, в ходе которой следует многократно проверять совместимость программного и аппаратного обеспечения и вносить исправления. Этот процесс повторяется до тех пор, пока ошибки не перестанут возникать.

3.4 Симуляция системы распознавания голоса

- Обзор симуляции системы

Симуляция системы является важным этапом разработки интеллектуальных роботов. Ее основная цель — проверить функциональность и стабильность системы до интеграции аппаратного и программного обеспечения через симуляционную среду. С помощью симуляции можно моделировать поведение робота в реальных условиях, заранее обнаруживать и устранять потенциальные проблемы, что сокращает время разработки и снижает затраты.

– Процесс симуляции системы

В данной диссертации процесс симуляции системы включает следующие основные этапы:

Получение и обработка результатов распознавания речи:

Основная плата через последовательный интерфейс получает результаты распознавания речи от системы голосового управления.

На основе результатов распознавания основная плата генерирует соответствующие управляющие сигналы, чтобы робот выполнял действия, такие как движение вперед, назад, повороты влево и вправо, вращение по кругу и т.д.

Создание симуляционной среды:

Используется симуляционное программное обеспечение для создания модели системы управления роботом.

В симуляционной среде моделируются такие модули, как голосовой ввод, данные с датчиков, управление двигателями и т.д.

Проверка функциональности:

В симуляционной среде тестируются все функции робота, чтобы убедиться, что он корректно реагирует на голосовые команды.

Проверяется совместная работа функций избегания препятствий, управления движением и распознавания речи.

Оптимизация производительности:

На основе результатов симуляции оптимизируются алгоритмы управления и параметры системы для повышения скорости реакции и стабильности робота.

Настраивается чувствительность алгоритма распознавания речи, чтобы обеспечить его надежность в сложных условиях.

– Конкретные шаги симуляции системы

Конкретные шаги симуляции системы включают:

Подключение и инициализация аппаратного обеспечения:

Микрофон и динамик закрепляются на роботе, чтобы обеспечить качество ввода и вывода звука.

Инициализируется основная плата (например, на базе 51-го микроконтроллера) и периферийные устройства (например, модули управления двигателями и датчиками).

Настройка среды компиляции:

В интегрированной среде разработки (IDE) настраивается среда компиляции, загружается исходный код программы интеллектуального управления роботом.

Программа компилируется, чтобы убедиться в отсутствии синтаксических и логических ошибок.

Загрузка и запуск программы:

Скомпилированная программа загружается в основную плату робота с помощью программатора.

Включается питание, и система робота запускается.

Обучение и распознавание речи:

В тихой обстановке проводится обучение голосовым командам, таким как "готово", "имя", "старт", "вперед", "назад" и т.д.

Каждая команда обучается дважды. Если система корректно распознает команду, выводится подсказка для перехода к следующей команде. Если команда не распознана, обучение повторяется.

Убедитесь, что тон голоса обучающего стабилен, чтобы избежать влияния изменений тона на результаты распознавания.

Тестирование функций:

Тестируются функции управления движением робота в симуляционной среде, чтобы убедиться, что он может выполнять действия по голосовым командам, такие как движение вперед, назад, повороты и вращение.

Тестируется функция избегания препятствий, чтобы убедиться, что робот может автоматически избегать препятствий на основе данных с датчиков.

Анализ результатов симуляции:

записываются данные и наблюдения в процессе симуляции, анализируются производительность и стабильность системы.

На основе результатов симуляции корректируются параметры системы и алгоритмы, чтобы оптимизировать поведение робота.

– Инструменты и методы симуляции системы

В процессе симуляции системы используются следующие инструменты и методы:

Симуляционное программное обеспечение: для создания моделей системы и имитации среды выполнения.

Инструменты отладки: такие как отладчики в IDE, логические анализаторы и т.д., для отслеживания выполнения программы и поиска ошибок.

Инструменты записи и анализа данных: например, Excel, для записи и анализа данных симуляции.

При распознавании голоса роботом рекомендуется проводить тренировку в тихой обстановке, чтобы избежать помех от других звуков, что позволит роботу более эффективно распознавать команды. Кроме того, тон голоса также влияет на результаты распознавания, поэтому во время тренировки тон должен быть стабильным, без резких изменений.

3.5 Выводы

Данная работа посвящена исследованию системы голосового управления роботом. Проектирование системы голосового управления включает разработку как программного, так и аппаратного обеспечения, а также реализацию функций распознавания голоса и движения робота в соответствующей симуляционной среде. В процессе движения используется алгоритм нечеткого нейронного управления для автоматического избегания препятствий.

Основные выводы работы можно обобщить следующим образом:

– В данном проекте выбран алгоритм распознавания голоса, который в

сочетании с принципами распознавания речи позволяет реализовать функцию распознавания голоса робота под управлением основного контроллера. В качестве основного контроллера выбран микроконтроллер SPCE061A.

– Аппаратное обеспечение системы управления включает проектирование модулей ввода и вывода голоса, схемы питания, схемы управления движением робота, а также инфракрасных датчиков. Все аппаратные решения построены вокруг микроконтроллера SPCE061A, что обеспечивает компактность и высокую рентабельность системы, а также ее практическую ценность.

– Программное обеспечение, написанное на языке C, реализует такие функции, как тренировка голосовых команд, распознавание голоса и воспроизведение звука. Это позволяет обеспечить взаимодействие между человеком и роботом: робот может распознавать команды, произнесенные конкретным человеком, и выполнять соответствующие действия, автоматически избегая препятствий в процессе движения. Для реализации этой функции был выбран улучшенный алгоритм нечеткого нейронного управления, который сочетает в себе преимущества нейронных сетей и нечеткой логики, что позволяет роботу гибко избегать препятствий. Компиляция и отладка всего программного кода выполнены в среде u'nSP IDE.

– Для более эффективной реализации функции голосового управления интеллектуальным роботом, распознавание голоса должно происходить в тихой обстановке, а тон голоса должен быть нормальным. При тестировании системы на наличие неисправностей используется комбинированный подход, включающий как аппаратное, так и программное тестирование.

ЗАКЛЮЧЕНИЕ

Настоящая работа посвящена разработке системы голосового управления для интеллектуального робота на базе микроконтроллера SPCE061A. В ходе исследования были успешно решены задачи проектирования аппаратной и программной частей системы, интеграции модулей распознавания речи и управления движением, а также реализации алгоритмов автоматического избегания препятствий.

Основные достижения работы:

- Теоретическая часть, Проведен анализ современных технологий распознавания речи, включая методы на основе скрытых марковских моделей (НММ) и нейронных сетей.

Изучены особенности взаимодействия человека и робота через естественный язык, что подтвердило практическую значимость выбранного направления.

- Аппаратная реализация, Разработана компактная и экономичная аппаратная платформа на базе микроконтроллера SPCE061A, включающая модули ввода/вывода аудиосигналов, управления двигателями и инфракрасными датчиками.

Оптимизированы схемы питания и управления, обеспечивающие стабильную работу системы в реальном времени.

- Программное обеспечение, Реализованы алгоритмы распознавания голосовых команд с использованием языка С, включая этапы обучения, сравнения с эталонными образцами и генерации управляющих сигналов.

Внедрен улучшенный алгоритм нечеткого нейронного управления для автономного избегания препятствий, что повысило адаптивность робота в динамической среде.

- Тестирование и симуляция, Проведена комплексная отладка системы, сочетающая аппаратные и программные методы.

- Успешно протестированы ключевые функции, распознавание 15

голосовых команд, движение по заданной траектории и реакция на препятствия.

– Практическая значимость, Разработанная система демонстрирует потенциал применения голосового управления в образовательных, развлекательных и сервисных роботах. Использование доступных компонентов (таких как Arduino и специализированные микросхемы распознавания речи) позволяет масштабировать решение для массового производства.

Перспективы развития:

– Увеличение словарного запаса системы за счет внедрения облачных технологий и глубокого обучения.

– Повышение помехоустойчивости распознавания в шумной среде с помощью адаптивных фильтров.

– Интеграция дополнительных датчиков (например, камер) для расширения функционала навигации.

В заключение, работа подтвердила возможность создания эффективной системы голосового управления для мобильных роботов на ограниченных вычислительных ресурсах. Полученные результаты вносят вклад в развитие человеко-машинного взаимодействия и могут служить основой для дальнейших исследований в области интеллектуальной робототехники.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- 1 Гребнов, С. В. Аналитический обзор методов распознавания речи в системах голосового управления / С. В. Гребнов // Вестник ИГЭУ. – 2009. – № 3. – С. 83–85.
- 2 Губочкин, И. В. Разработка алгоритмов анализа и распознавания речи на основе адаптивной кластерной модели и критерия минимального информационного рассогласования : автореф. дис. ... канд. техн. наук : 05.13.17 / И. В. Губочкин ; Нижегород. гос. лингвист. ун-т им. Н. А. Добролюбова. - Нижний Новгород, 2011. - 22 с.
- 3 Дейтел, П. Python: искусственный интеллект, большие данные и облачные вычисления / П. Дейтел. - СПб. : Питер, 2020. - 864 с.
- 4 Захаров, М. В. Разработка технологических процессов сборки : учеб. пособие / М. В. Захаров, Ю. В. Тимофеев. - Киев : НМК ВО, 1992. - 152 с.
- 5 Зубарев, А. А. Ассемблер для микроконтроллеров AVR : учеб. пособие / А. А. Зубарев. - Омск : Изд-во СибАДИ, 2007. - 112 с.
- 6 Как это работает? Распознавание речи [Электронный ресурс]. – URL: <https://yandex.livejournal.com/288893.html> (дата обращения: 23.10.2024).
- 7 Кангин, В. В. Аппаратные и программные средства систем управления. Промышленные сети и контроллеры : учеб. пособие / В. В. Кангин, В. Н. Козлов. - Москва : БИНОМ. Лаборатория знаний, 2010. - 418 с.
- 8 Кангин, В. В. Компьютеры в системах управления технологическими процессами : учеб. пособие / В. В. Кангин, М. В. Кангин, Д. Н. Ямолдинов. - Нижний Новгород : НГТУ, 2005. - 246 с.
- 9 Кангин, В. В. Контроллеры Arduino в системах автоматизации, управления и контроля. Вводный курс : учеб. пособие / В. В. Кангин. - Старый Оскол : ТНТ, 2019. - 276 с.
- 10 Кангин, В. В. Промышленные контроллеры в системах автоматизации технологических процессов : учеб. пособие / В. В. Кангин. - Старый Оскол : ТНТ, 2013. - 408 с.

- 11 Карвинен, Т. Делаем сенсоры: проекты сенсорных устройств на базе Arduino и Raspberry Pi / Т. Карвинен, К. Карвинен, В. Валтокари ; пер. с англ. - Москва : Вильямс, 2015. - 432 с.
- 12 Каргин, А. А. Применение нечеткой логики в системах фонологической классификации звуков речи / А. А. Каргин // Искусственный интеллект. - 2010. - С. 210.
- 13 Китайская академия информационных и коммуникационных технологий. Официальный документ об экосистеме с открытым исходным кодом [Электронный ресурс]. – URL: [http://www.caict.ac.cn/...](http://www.caict.ac.cn/) (дата обращения: 01.12.2024).
- 14 Невлюдов, И. Ш. Голосовое формирование управляющих команд при проектировании роботизированных сборочных процессов / И. Ш. Невлюдов, А. М. Цымбал, С. С. Милютина // Восточно-Европейский журнал передовых технологий. - 2008. - Т. 4, № 2 (34). - С. 65-68.
- 15 Невлюдов, И. Ш. Голосовое формирование управляющих команд робота в САПР технологических процессов / И. Ш. Невлюдов, А. М. Цымбал, С. С. Милютина // Восточно-Европейский журнал передовых технологий. - 2008. - № 3/2 (23). - С. 12-14.
- 16 Невлюдов, И. Ш. Голосовое формирование управляющих команд робота в САПР технологических процессов / И. Ш. Невлюдов, А. М. Цымбал, С. С. Милютина // Восточно-Европейский журнал передовых технологий. - 2008. - Т. 2, № 2 (32). - С. 12-14.
- 17 Петин, В. А. Проекты с использованием контроллера Arduino : учеб. пособие / В. А. Петин. – 2-е изд., перераб. и доп. – Санкт-Петербург : БХВ-Петербург, 2015. - 464 с.
- 18 Ревонченкова, И. Ф. Особенности использования роботов с применением речевого интерфейса при автоматизации механосборочных работ в машиностроении / И. Ф. Ревонченкова // CAD/CAM/PDM : материалы конф. - М. : ИПУ РАН, 2016. - С. 206-210.
- 19 Ревонченкова, И. Ф. Особенности создания системы распознавания

речи в управлении роботом / И. Ф. Ревонченкова // INTELLECTUAL MACHINES : материалы международной научной конференции. – Москва : МГТУ (МАМИ), 2009. - С. 95-99.

20 Редингтон, М. Нейронные сети для распознавания речи и последовательностей / М. Редингтон // Язык и речь. - 1996. - С. 393.

21 Самохвалова, С. Г. Проектирование системы распознавания речи робота на базе SPCE061A / С. Г. Самохвалова, Лю Чао // Вестник Амурского государственного университета. - 2024. - № 105. - С. 190-194.

22 Самохвалова, С. Г. Проектирование системы распознавания речи робота на базе SPCE061A / С. Г. Самохвалова, Лю Чао // Флагман науки : международный научный журнал. - 2024. - № 11 (22). - С. 334-338.

23 Скрытые Марковские модели в распознавании речи [Электронный ресурс]. – Режим доступа: <https://habrahabr.ru/post/134954/> (дата обращения: 20.12.2011).

24 Соммер, У. Программирование микроконтроллерных плат Arduino/Freduino : [монография] / У. Соммер. – Санкт-Петербург : БХВ-Петербург, 2012. - 256 с.

25 Стюарт, Р. Искусственный интеллект. Современный подход : [моногр.] / Р. Стюарт. – Москва : Вильямс, 2019. - 1408 с.

26 Сун, Янань. Дизайн и реализация эксперимента по распознаванию речи для робота / Янань Сун // Компьютерная технология и развитие. - 2013. - № 4.

27 Сунь, Ие. Текущее состояние и тенденции развития искусственного интеллекта / Ие Сунь // Инженерная стоимость. - 2015. - № 6.

28 Тан, Цяньвэнь. Эволюционный анализ системы с открытым исходным кодом Java на основе теории сложных сетей : монография / Цяньвэнь Тан. – [Б. м.], 2018. – 173 с.

29 Управление техническими системами. Контроллер ADAM-5510 и его программирование в UltraLogik : учеб. пособие / В. В. Кангин, М. В. Кангин, А. Н. Богов, Д. Н. Ямолдинов. – Нижний Новгород : НГТУ, 2008. - 396 с.

30 Халперн, М. И. Независимость машины: ее технология и экономика / М.

И. Халперн // Сообщения АСМ. - 1965 - С. 782.

31 Челищев, Б. Е. Автоматизация проектирования технологии в машиностроении / Б. Е. Челищев, И. В. Боброва, А. Гонсалес-Сабатер ; под ред. акад. Н. Г. Бруевича. - М. : Машиностроение, 1987. - 264 с.

32 Чен, Сун. Проектирование робота с распознаванием речи на основе микроконтроллера / Сун Чен // Робототехника. - 2012. - № 4.

33 Чжао, Хуэйцзюнь. Спортивные вычисления: новая область компьютерных исследований / Хуэйцзюнь Чжао // Computer Science. - 2004. - С. 89.

34 Чжу, Шичжан. Текущее состояние и тенденции развития технологии тактильного восприятия роботов / Шичжан Чжу // Журнал Ханчжоу электронной промышленности. - 2013. - № 8.

35 Шарий, Т. В. Анализ просодии на этапе параметризации речевого сигнала / Т. В. Шарий // Современные тенденции развития и перспективы внедрения инновационных технологий в машиностроении, образовании и экономике : материалы III Международной научно-практической конференции. - 2016. - С. 142.

36 Шарий, Т. В. Голосовое управление мобильным роботом на основе когнитивной модели FCAS / Т. В. Шарий // Искусственный интеллект. - 2014. - С. 75.

37 Шарий, Т. В. Об одном методе автоматической сегментации речевых сигналов / Т. В. Шарий // Бионика интеллекта. - 2009. - С. 61.

38 Юревич, Е. И. Основы робототехники / Е. И. Юревич. – Санкт-Петербург : СПбГПУ, 2000. - 252 с.

39 Ян, Си. Состояние и тенденции развития искусственного интеллекта за рубежом / Си Ян // Распознавание образов и искусственный интеллект. - 2015. - № 2.