

5. Trukhin, A., et al. "Virtual Reality in Laboratory Learning." Educational Research Review, 2018, vol. 10, pp. 100–115

6. Зұлпыхар А. VR технологияларын Қазақстан мектептерінде қолдану тәжірибесі // Білім және технологиялар журналы. – 2025. – № 2. – Б. 12–25.

**Цих А.А., Крецу Д.В.** Карагандинский национальный исследовательский университет имени академика Е.А.Букетова, Физико-технический факультет, гр. РТ-22-1р, студент, Факультет математики и информационных технологий, гр.РиМ-22-2р

*(Научный руководитель — профессор кафедры радиофизики и электроники Афанасьев Д.А.)*

## АВТОНОМНАЯ ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ГОЛОСОВОГО УПРАВЛЕНИЯ И МОНИТОРИНГА СРЕДЫ

### Введение и актуальность

В последние годы наблюдается стремительный рост распространения голосовых интерфейсов. Однако подавляющее большинство коммерческих решений построено на облачной архитектуре: аудиоданные передаются на удалённые серверы, где осуществляется распознавание и обработка. Такой подход формирует зависимость от сети и создаёт потенциальные риски утечки информации.

Согласно концепции периферийных вычислений (edge computing), перенос обработки данных непосредственно на устройство позволяет снизить задержку отклика, повысить автономность и обеспечить независимость от внешней инфраструктуры [3, с. 18–21]. Кроме того, локальная обработка исключает передачу конфиденциальной информации третьим лицам.

Целью данной научно-исследовательской работы стала разработка полностью автономной интеллектуальной системы голосового управления и мониторинга среды, способной выполнять распознавание речи, семантический анализ и синтез ответа без подключения к облачным сервисам.

### Аппаратная платформа и возможности одноплатных компьютеров

Ключевым фактором, сделавшим возможной реализацию подобной системы, стало развитие современных одноплатных микрокомпьютеров.

Ещё несколько лет назад выполнение нейросетевых алгоритмов требовало либо персонального компьютера, либо специализированных графических ускорителей. Однако современные одноплатные решения обладают многоядерными ARM-процессорами, достаточным объёмом оперативной памяти и поддержкой высокоскоростных интерфейсов хранения данных.

Как отмечает М. Гук, современные одноплатные компьютеры представляют собой полноценные встраиваемые вычислительные системы, способные работать под управлением Linux и выполнять ресурсоёмкие задачи обработки данных [1, с. 12–15]. Их производительность в сочетании с низким энергопотреблением позволяет использовать их в составе интеллектуальных устройств реального времени. [Рисунок 1]

Дополнительно следует отметить экономический аспект. Стоимость одноплатного компьютера с 8 ГБ оперативной памяти сопоставима со стоимостью коммерческих голосовых колонок, которые фактически представляют собой микрофон, динамик и сетевой модуль. Однако в разработанной системе вычисления выполняются локально, без привлечения удалённых дата-центров.

Таким образом, благодаря прогрессу микроэлектроники и оптимизации операционных систем Linux для ARM-архитектуры стало возможным создание автономных интеллектуальных устройств сравнительно небольшой стоимости [1, с. 27–30; 13, с. 45–48].



Рисунок 1. Одноплатный микрокомпьютер OrangePi 4 Pro

### Акустическая система и конструктив корпуса

Корпус устройства спроектирован с учётом требований электроакустики. Геометрия влияет на формирование диаграммы направленности, распределение стоячих волн и характер внутренних резонансов.

Как указывает Ф. Тул, конструкция корпуса акустической системы оказывает прямое влияние на линейность частотной характеристики и уровень паразитных вибраций [4, с. 145–152]. В связи с этим в работе реализована комбинированная конструкция: пластиковый корпус с внутренней гироидной матрицей, заполненной демпфирующим композитом.

С точки зрения механики ячеистых структур гироид обеспечивает близкую к изотропной жёсткость и равномерное распределение нагрузок [6, с. 73–78]. Заливка демпфирующего состава увеличивает внутренние потери и снижает амплитуду структурных резонансов. [Рисунок 2]

В результате удалось уменьшить паразитные призвуки и повысить акустическую стабильность корпуса без значительного увеличения массы.

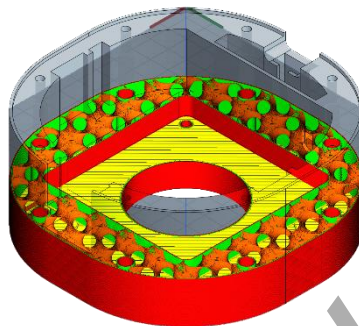


Рисунок 2. Внутренняя структура заполнения гироидной матрицы для заливки демпфирующего состава

#### Усилительный тракт и энергоэффективность

В системе используется усилитель класса D. Принцип его работы основан на широтно-импульсной модуляции сигнала, что обеспечивает высокий коэффициент полезного действия.

Согласно Р. Эриксону, импульсные преобразователи и усилители класса D достигают КПД свыше 85–90 %, существенно снижая тепловые потери по сравнению с линейными схемами [8, с. 412–418]. Это особенно важно для компактных устройств с ограниченным теплоотводом.

Применение импульсного источника питания дополнительно повышает энергоэффективность системы и уменьшает габариты блока питания [8, с. 95–101].

Таким образом, аудиотракт спроектирован с учётом требований динамического диапазона, тепловой стабильности и компактности.

#### Архитектура распознавания речи

Система автоматического распознавания речи реализована по классической схеме: предварительная обработка сигнала, выделение признаков и декодирование на основе вероятностной модели.

Как отмечают Л. Рабинер и Б. Джуанг, ключевым этапом является преобразование временного сигнала в спектральное представление и вычисление параметров MFCC, отражающих особенности человеческого слухового восприятия [17, с. 82–89].

Дальнейшая обработка осуществляется акустической и языковой моделями. В современных системах всё чаще используются нейронные сети, способные аппроксимировать сложные зависимости в речевом сигнале [17, с. 312–318].

В рамках работы реализована локальная модель распознавания, работающая непосредственно на одноплатном компьютере без обращения к облачным сервисам.

#### Семантический поиск и обработка ошибок распознавания

Особенностью голосового интерфейса является наличие ошибок распознавания. Для их компенсации реализована двухэтапная система поиска.

Первый этап основан на полнотекстовой индексации, что соответствует классическим принципам информационного поиска, описанным Мэннингом и соавторами [14, с. 35–41].

Второй этап использует нечеткое сравнение строк на основе расстояния Левенштейна. Данный алгоритм позволяет учитывать вставки, удаления и замены символов, что особенно актуально при фонетических искажениях [16, с. 3–7].

Комбинирование этих методов обеспечивает устойчивую работу системы даже в условиях зашумленной среды.

#### Нейросетевой синтез речи

Синтез речи основан на современных нейросетевых архитектурах, включающих акустическую модель и вокодер.

Как отмечается в работе В. Сзе и соавторов, оптимизация нейронных сетей позволяет выполнять сложные вычисления на встраиваемых платформах при разумном энергопотреблении [3, с. 112–118].

Предварительная загрузка модели в оперативную память обеспечивает низкую задержку генерации ответа, что делает взаимодействие с системой естественным.

Заключение. В ходе выполнения работы разработана и реализована автономная интеллектуальная система голосового управления и мониторинга среды.

Продемонстрировано, что благодаря развитию одноплатных микрокомпьютеров и оптимизированных нейросетевых моделей стало возможным выполнение полного цикла обработки речи — от захвата сигнала до генерации ответа — непосредственно на локальном устройстве.

Система сочетает:

- акустически оптимизированный корпус,
- энергоэффективный усилительный тракт,
- локальное распознавание речи,
- двухэтапный семантический поиск,
- нейросетевой синтез речи.

Полученные результаты подтверждают перспективность автономных edge-решений в области интеллектуальных голосовых интерфейсов.

#### Литература:

1. Гук М. Одноплатные компьютеры и встраиваемые Linux-системы. — СПб.: БХВ-Петербург, 2020.
2. Сзе В., Чен Ю., Янг Т. Эффективные вычисления глубоких нейронных сетей. — М.: ДМК Пресс, 2022.
3. Лав Р. Программирование в Linux. Системный подход. — СПб.: Питер, 2014.
4. Тул Ф. Звук и акустика громкоговорителей и помещений. — М.: ДМК Пресс, 2019.
5. Гибсон Л., Эшби М. Ячеистые материалы: структура и механические свойства. — М.: Техносфера, 2016.
6. Эрикссон Р., Максимович Д. Основы силовой электроники. — М.: ДМК Пресс, 2021.
7. Мэннинг К., Рагхаван П., Шютце Х. Введение в информационный поиск. — М.: Вильямс, 2011.
8. Левенштейн В. И. Двоичные коды с исправлением выпадений, вставок и замещений символов // Доклады АН СССР. — 1965.
9. Рабинер Л., Джуанг Б. Основы распознавания речи. — М.: Техносфера, 2012.