

ПРОГРАММНО-АППАРАТНЫЙ КОМПЛЕКС АНАЛИЗА ГОЛОСОВЫХ И РЕЧЕВЫХ СИГНАЛОВ ДЛЯ ЗАДАЧ МЕДИЦИНСКОЙ ДИАГНОСТИКИ

Субботенко О.Р., студент

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Порхун М.И. – маг. техн. наук, старший преподаватель

В работе рассматривается реализация программно-аппаратного комплекса анализа голосовых и речевых сигналов для задач медицинской диагностики. Функционал устройства включает в себя запись звукового сигнала и его последующую обработку. В процессе обработки рассчитываются следующие параметры: частота основного тона (ЧОТ), джиттер и шиммер. Функционал устройства и результаты его работы отображаются в графическом интерфейсе.

Программно-аппаратный комплекс анализа голосовых и речевых сигналов для задач медицинской диагностики реализован на базе одноплатного компьютера Raspberry Pi. На рисунке 1 изображена структура устройства.

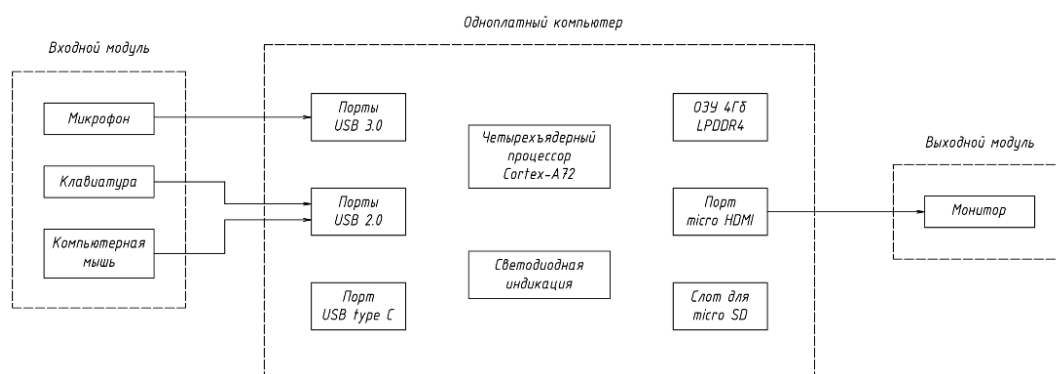


Рисунок 1 – Структурная схема программно-аппаратного комплекса анализа голосовых и речевых сигналов для задач медицинской диагностики

Комплекс реализован на одноплатном компьютере Raspberry Pi. К нему подключены следующие периферийные устройства: микрофон, клавиатура, компьютерная мышь (по интерфейсу USB) и монитор (по интерфейсу HDMI).

Алгоритм работы программной части комплекса приведен ниже:

5. Подключение необходимых библиотек.
6. Создание экземпляра класса QApplication (создание приложения).
7. Создание главного окна.
8. Описание интерфейса главного окна.
9. Запись нового аудиофайла или загрузка существующего.
10. Анализ аудиофайла (расчет ЧОТ, джиттера, шиммера, представление сигнала во временной области и расчет мел-спектрограммы).
11. Обновление интерфейса в соответствии с рассчитанными значениями и построенными графиками.

Для расчета частоты основного тона необходимо произвести грубую разметку сигнала на периоды основного тона. Сначала из основного сигнала выделяется фрагмент длительностью 50 мс и вычисляется его автокорреляционная функция (АКФ). Затем производится поиск первого пика АКФ, индекс которого применяется при вычислении частоты грубой разметки. Для обнаружения индексов грубой разметки сигнал пропускается через КИХ-фильтр нижних частот, после чего индексы корректируются с учетом групповой задержки фильтра. Далее производится вычисление ЧОТ согласно методу локальных максимумов. Суть данного метода заключается в нахождении отрицательных пиков в фрагментах между точками грубой разметки. На следующем шаге к точкам найденных отрицательных пиков применяется интерполяция второго порядка [1]. После выполненных вычислений определяется ЧОТ для каждого фрагмента:

$$F_0 = \frac{Fs}{P(i) - P(i - 1)}, \quad (1)$$

где Fs – частота дискретизации, P – массив индексов отрицательных пиков на каждом фрагменте.

После того, как получено значение ЧОТ, можно приступить к расчету джиттера и шиммера. Джиттер – это мера вариативности периода основного тона, иными словами, он отражает стабильность голоса [2,3]. Измеряется в процентах и вычисляется по следующей формуле:

$$J = \frac{\frac{1}{N-1} \sum_{i=2}^N |T_i - T_{i-1}|}{\frac{1}{N} \sum_{i=1}^N |T_i|}, \quad (2)$$

где N – количество периодов основного тона в анализируемом сигнале, T_i – продолжительность i -го периода основного тона.

Шиммер – это мера вариативности амплитуды акустических колебаний в процессе фонации, то есть характеризует изменение амплитуды голоса во времени. Измеряется также в процентах и вычисляется по следующей формуле:

$$S = \frac{\frac{1}{N-L+1} \sum_{i=1+(L-1)/2}^{N-(L-1)/2} |A_i - \frac{1}{L} \sum_{n=i-(L-1)/2}^{i+(L-1)/2} A_n|}{\frac{1}{N} \sum_{i=1}^N |A_i|}, \quad (3)$$

где N – количество периодов основного тона в анализируемом сигнале, A_i – амплитуда i -го периода основного тона, L – количество периодов, используемых для вычисления «скользящего» среднего.

Функционал устройства позволяет записать новый аудиофайл, предварительно указав его имя, или же выбрать уже существующий. После успешной записи (загрузки) файла становится возможным его анализ, по окончании которого выводятся рассчитанные значения ЧОТ, джиттера, шиммера, а также представление сигнала во временной и частотной областях.

Рассчитанные параметры, а также представление сигнала во временной и частотной областях выводятся в графическом интерфейсе. На рисунке 2 показано окно интерфейса приложения и результаты анализа звука \a\.

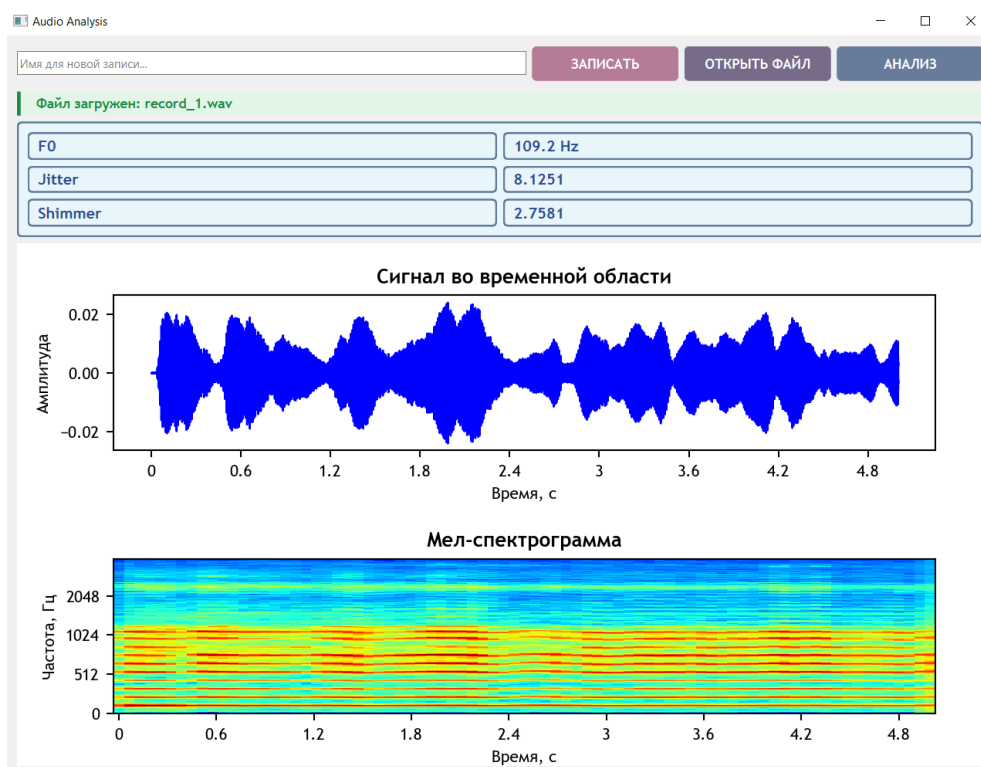


Рисунок 2 – Интерфейс программно-аппаратного комплекса

Представленные в интерфейсе результаты подтверждают корректность работы алгоритмов обработки звука и комплекса в целом, что предоставляет возможность его применения для предварительного скрининга голосовых нарушений.

В процессе разработки программно-аппаратного комплекса были пройдены такие этапы, как создание структуры устройства, разработка его функционала, написание алгоритма программной части, создание графического интерфейса. Результатами работы являются макет устройства, а также рассчитанные параметры и построенные графики, которые отображаются в графическом интерфейсе.

Список использованных источников:

5. Пастернак В. В., Стецко В. Ю. Исследование методов сегментации голосовых сигналов. – 2022.
6. Вашкевич М. И., Рушкевич Ю. Н. Детектирование речевых нарушений у пациентов с БАС на основе анализа голосового сигнала // Цифровая обработка сигналов. – 2019. – №. 4. – С. 33-41.
7. Саламатов К. А., Смирнов А. А. Извлечение формант и других характеристик и классификация состояния здоровья голоса методами машинного обучения // Международный научно-исследовательский журнал. – 2025. – №. 10 (160) S. – С. 1-6.