

ПРОГРАММНЫЙ МОДУЛЬ КЛАССИФИКАЦИИ РЕЧИ

А.С. Райкевич

В современных условиях информационного общества с каждым днем все более актуальным становится использование речевых технологий, таких как распознавание, анализ речи, голосовое управление сложными техническими системами, а также автоматизированная постановка произношения. Автоматическое распознавание речи является динамично развивающимся направлением в области искусственного интеллекта. Распознавание речи является задачей классификации образов акустических характеристик речевых сигналов.

Архитектура системы включает два основных модуля – модуль предобработки сигнала, который предназначен для выделения информативных акустических характеристик речевого сигнала и формирования акустического сигнала как набора характеристик и модуль постобработки сигнала.

На этапе предварительной обработки исходный сигнал преобразуется в векторы признаков, на основе которых затем будет произведена классификация. Этот этап включает в себя следующие шаги:

- процесс ввода речевого сигнала;
- применение фильтров для подавления шумов;
- выделение границ речи;
- нарезка речевого сигнала перекрывающимися кадрами;
- выделение признаков сигнала.

Выделение признаков сигнала происходит по методу мел-частотных кепстральных коэффициентов. Вычисление мел-частотных кепстральных коэффициентов включает в себя следующие шаги:

- разделение исходного сигнала на кадры, кадры накладываются друг на друга;
- к каждому кадру применяется преобразование Фурье;
- к каждому кадру применяется блок мел-фильтров – треугольных пересекающихся фильтров, расположенных наиболее плотно в области нижних частот;
- полученные энергии логарифмируются;
- применяется дискретное косинусное преобразование;

На вход программной оболочки попадает звуковой файл форматом WAV. На экране отображаются сигналы, соответствующие этапам обработки, а также параметры преобразования. В результате обработки на выходе образуется массив фонетических единиц.

В результате предложена программная оболочка для обработки речевых сигналов для системы распознавания речи.