

**Д. Г. Шумилов**  
(ГГТУ имени П. О. Сухого, Гомель)  
Науч. рук. **Л. К. Титова**, ст. преподаватель

## **РАСПОЗНАВАНИЕ ГОЛОСОВЫХ КОМАНД МЕТОДОМ MEL-КОЭФФИЦИЕНТОВ**

Распознавание голосовых команд – это процесс автоматического определения и интерпретации произносимых слов или фраз в звуковом сигнале. Распознавание широко используется в различных областях, включая телефонию, контроль голосовых устройств, медицину и т.д.

В общем случае, распознавание речи включает в себя следующие шаги:

- захват звука: звуковой сигнал, содержащий речь, записывается и декретирован в цифровую форму;
- предварительная обработка: сигнал может быть фильтрован, шумоподавлен и нормализован для улучшения качества распознавания;
- извлечение признаков: из сигнала извлекаются характеристики, которые описывают звучание речи. Эти характеристики могут включать MFCC (Mel-frequency cepstral coefficients) и другие;
- моделирование: полученные признаки подаются на вход модели машинного обучения, которая обучена распознавать речь на основе предварительно известных данных;
- распознавание: модель выполняет классификацию и определяет, какое слово или фраза была произнесена в звуковом сигнале.

MEL-коэффициенты (Mel-frequency cepstral coefficients, MFCC) – это особый вид коэффициентов, которые используются для извлечения характеристик из аудиосигнала речи. Они являются представлением спектра сигнала, преобразованного в мел-шкалу, которая лучше соответствует человеческому восприятию частотных характеристик. MFCC можно использовать для извлечения признаков из звукового сигнала, которые затем используются в алгоритмах машинного обучения для распознавания речи.

Допустим имеется фрейм размером 256 элементов. Из данных об аудио формате известно, что частота звука в данной фрейме 16000 hz. Предположим, что человеческая речь лежит в диапазоне от [300; 8000] hz. Количество искомым мел-коэффициентов положим  $M = 12$ .

Для того, чтобы разложить такой спектр по mel-шкале, потребуется создать фильтры. По сути, каждый mel-фильтр – это треугольная оконная функция, которая позволяет просуммировать количество энергии на определённом диапазоне частот и тем самым получить

mel-коэффициент. Зная количество mel-коэффициентов и анализируемый диапазон частот можно построить набор фильтров.

Число фильтров должно быть больше, чем число коэффициентов, которые рассчитываются. Предположим, идёт расчёт 10 коэффициентов, в это случае необходимо указать 12 фильтров. И важно учитывать, что фильтры должны постоянно возрастать. Для первого коэффициента фильтр будет минимальным, для последнего максимальным.

Сам алгоритм расчёта выглядит следующим образом:

- необходимо произвести шумоподавление в входном файле. Это делается для снижения уровня фоновых и широкополосных шумов, с минимальным снижением качества сигнала. На этом этапе может использоваться любой алгоритм шумоподавления, например среднеквадратичное шумоподавление;

- нормализация сигнала. Этот процесс ответственен за выравнивание частотных характеристик. Существует два типа нормализации звука: пиковый и средний;

- исходную запись необходимо разбить на части, называемые фреймами. Стандартный размер фрейма 128 миллисекунд. Разбиение может иметь как стандартный вид, так и может происходить внахлёт, то есть каждый фрейм имеет в начале себя конец предыдущего;

- применение оконной функции. В стандартном методе расчёта применяется треугольная оконная функция;

- преобразование Фурье, которое применяется для каждого фрейма, расчёт происходит по формуле 1;

- произвести оставшиеся расчёты по формулам 2, 3 и 4 для каждого фрейма.

$$X_a[k] = \sum_{n=0}^{N-1} x[n] e^{\frac{-2\pi i}{N} kn}, \quad (1)$$

где  $x[n]$  – исходный речевой сигнал.

$$H(m) = \begin{cases} 0, & k < f[m-1] \\ \frac{(k-f[m-1])}{(f[m]-f[m-1])}, & f[m-1] \leq k < f[m] \\ \frac{(f[m+1]-k)}{(f[m+1]-f[m])}, & f[m] \leq k \leq f[m+1] \\ 0, & k > f[m+1] \end{cases}, \quad (2)$$

где  $f$  – набор опорных точек (фильтров),

$k$  – фрейм, для которого производится вычисление.

$$S[m] = \ln(\sum_{k=0}^{N-1} |X_a[k]|^2 H_m[k]) ; \quad (3)$$

$$c[n] = \sum_{m=0}^{M-1} S[m] \cos(\pi n(m + 0.5)/M) . \quad (4)$$

Исходя из данных записей можно сделать вывод, что расчёт коэффициентов является максимально гибким и зависит от методов, что были выбраны для расчёта.

К данному моменту задача с распознаванием голосовых команд остаётся не решённой, для её завершения необходимо обучить классификатор на тестовых данных. В данном случае тестовые данные представляют из себя голосовую команду и коэффициенты к данной команде.

Для классификации можно использовать любую обученную модель, которая может быть создана на выбранном языке программирования.

**N. A. Aksionova, V. A. Lahvinets**

(F. Skorina GSU, Gomel)

Scientific adviser **A. V. Varuyeu**, Ph.D. in technics, associate professor

## **DEPLOYING AUTOMATED KNOWLEDGE CONTROL SYSTEM PROJECT TO VERCEL**

Deploying a Next.js project like the Automated Knowledge Control System (AKCS) to the Vercel platform is a straightforward process that requires a few key steps. AKCS is an educational content management system designed for schools to conduct online tests and is built using Next.js.

The first step to deploying a Next.js project like AKCS to Vercel is to create an account on the Vercel platform. To create an account, simply visit the Vercel website and sign up using your GitHub or GitLab account. Once the account is created, the next step is to connect the project's repository to Vercel. This can be accomplished using either the Vercel CLI or the Vercel web interface. To connect the repository using the web interface, simply select the «Import Project» option and choose the repository you wish to deploy. After the repository is connected, it is necessary to configure the project's settings. This includes setting the project's name, build command, and any necessary environment variables. Additionally, deployment settings such as the deployment branch and build command need to be configured as well. It's important to note that Next.js projects require a build step in order to generate the static assets necessary for deployment. The build command for a Next.js project is typically "npm run build". This command tells Vercel