

**ПРИЛОЖЕНИЕ ДЛЯ АНАЛИЗА ЗВУКОВОЙ ИНФОРМАЦИИ,
СОДЕРЖАЩЕЙ ДЕФЕКТЫ РЕЧИ ЛЮДЕЙ С ОГРАНИЧЕННЫМИ
ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ**

Столер В.А.

*Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь
stoler@bsuir.by*

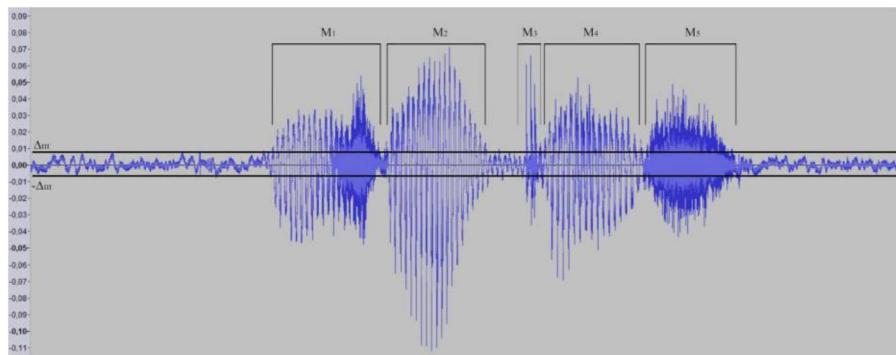
В статье рассмотрено компьютерное приложение, для распознавания дефектов слов при их произношении людей с ограниченными возможностями здоровья, используемое для анализа и исправления этих дефектов. Рассмотрены способы оптимизации программы и пути повышения точности исправления дефектов речи.

Ключевые слова: дефекты произношения слов, звуковая модель, обработка звука, алгоритмы программы, компьютерное приложение.

В настоящее время все чаще обращается внимание на людей с ограниченными возможностями здоровья и помощи им. В повседневной жизни, люди с такими заболеваниями испытывают значительные трудности при общении. В данной работе рассматривается приложение, позволяющее людям имеющие дефекты речи, возникшие от рождения, в результате травмы или болезни, при общении с окружающим миром передавать и получать качественную информацию. Приложение нацелено на преобразование информации в формате аудио, когда звук в реальном времени обрабатывается приложением, и человек получает исправленную звуковую информацию в онлайн режиме.

Предлагается алгоритм построения программы для распознавания ключевых слов, содержащие дефекты произносимой речи. В терминах цифровой обработки сигналов процесс шумоочистки представляет собой преобразование входного сигнала, содержащего как полезный сигнал – речь, так и аддитивный сигнал-помеха – шум, в выходной сигнал, содержащий только речь. Поскольку создание систем, в частности удовлетворяющих данному условию, невозможно, задачу очистки сигнала от шума упрощают [1].

На первом этапе зададим погрешность для шума, оставшегося после обработки. Определим локальные максимумы графика. Интервал M_i , на котором график выходит за пределы заданного значения $\Delta_{ш}$, имеет локальный максимум и ограничен локальными минимумами, является буквой в слове (рисунок 1).

Рисунок 1– Выбор интервалов M_i на графике записанной речи

На втором этапе сопоставляем эталонный график и график записанной речи. Выберем количество точек N для сравнения на заданном интервале M_i и зададим погрешность для точек, в пределах сравниваемых интервалов графиков, Δ_t . При этом эталонными считаются графики слов, записанные на профессиональном оборудовании, и содержащие минимальное количество шумов. Данные условия необходимы для минимизации числа ошибок при сравнении слов.

При равенстве точек $[x_j; y_j]$ графика записанной речи точкам $[x_j; y_k]$ эталонного графика с учетом выбранной погрешности Δ_t можно говорить о равенстве произнесенных букв. Из равенства всех интервалов M_i сравниваемых звуковых сигналов следует равенство произнесенных слов.

В результате был разработан обобщенный алгоритм программы для распознавания ключевых слов в записанной речи. Для оптимизации алгоритма необходимо было нормализовать входной сигнал, т. е. задать верхнюю границу значения y . Таким образом, если разница между минимальным (максимальным) значением y первого графика и минимальным (максимальным) значением y второго графика больше заданной погрешности точки Δ_t , можно говорить о неравенстве символов, т.е. о неравенстве слов [2].

Для корректного определения M_i необходимо учесть, что значение y локального максимума символа много больше заданного значения погрешности шума Δ_{sh} . Изменение значений погрешности шума Δ_{sh} и погрешности точки Δ_t позволяет задать необходимую точность сравнения. При разработке алгоритма были учтены следующие возможные ситуации: количество символов M в записанном слове больше, чем у эталонного; интервал M в записанном слове вмещает в себя большее количество точек N , чем эталон. Большее число символов M в записанном слове может означать не только более длинное слово, но и дефект произношения (например, «Привет»). Для обработки данной ситуации необходимо предусмотреть проверку, в которой у записанного слова будут попарно сравниваться соседние интервалы M_i и M_{i+1} . Если соседние интервалы равны между собой, значит M_i необходимо исключить из сравнения.

В заключении необходимо обратить внимание на такую ситуацию, при которой длина интервала M_i в записанном слове больше длины соответствующего интервала M_i в эталонном слове, т.е. количество вмещаемых интервалом точек N различно. Такое различие может возникнуть не только при неравенстве символов, но и при более длинном произношении буквы (напр., «Ммама»). При систематическом повторении значений y в пределах рассматриваемого интервала, можно говорить о более длинном произношении символа и исключить часть интервала из сравнения для уравнивания значений N .

В результате исследований был выполнен анализ основных способов взаимодействия пользователя с интерфейсами, используемых в программах распознавания речи. Предложен алгоритм обработки звука, учитывающий дефекты произношения слов, для построения компьютерной программы распознавания речи, используемой при звуковом онлайн общении. Рассмотрены способы оптимизации программы и повышения точности распознавания речи.

Литература

- Vishnyakov I.E., Masyagin M.M., Odintsov O.A., Sliusar V.V. Methods and algorithms for real

time voice noise cleaning. Proc. Univ. Electronics, 2021, vol. 26, no. 2, pp. 184–196.

2. Jayashree Padmanabhan, Melvin Jose, Johnson Premkumar. Machine Learning in Automatic Speech Recognition: A Survey. Proc. IETE Technical Review, 2015, vol. 32, no. 5, pp. 240–251.

APPLICATION FOR ANALYZING SOUND INFORMATION CONTAINING SPEECH DEFECTS OF PEOPLE WITH DISABILITIES

Stoler V.A.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

This article examines a computer application for recognizing speech imperfections in the speech of people with disabilities, using it to analyze and correct these imperfections. Methods for program optimization and ways to improve the accuracy of speech imperfection correction are discussed.

Keywords: pronunciation defects, sound model, sound processing, program algorithms, computer application.