

СИСТЕМА АВТОМАТИЧЕСКОГО ФОРМИРОВАНИЯ РЕЧЕВОГО КОРПУСА ДЛЯ СИСТЕМЫ СИНТЕЗА РЕЧИ ПО ТЕКСТУ

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

Ахраменко И.О

Азаров И.С. – д.т.н., доцент

В настоящее время системы синтеза речи имеют широкое распространение. Синтез речи находит своё применение в таких отраслях как:

- 1) информационно-справочные системы
- 2) авиация
- 3) диалог человека с компьютером
- 4) робототехника

Корпусный метод синтеза речи использует базу данных естественной речи, которая представляет собой речевые сегменты различной длины. Основой базы данных является предложения в текстовой форме и их звуковая запись.

Для автоматической сегментации речи можно использовать скрытые марковские модели. Так как размер базы данных при корпусном методе синтеза речи достаточно велик, для поиска необходимых сегментов целесообразно использовать хеш-таблицы, что поможет значительно сократить время поиска необходимых сегментов. В лучшем случае алгоритм поиска в хеш-таблице имеет сложность $O(1)$, в худшем случае (когда все ключи имеют одинаковое значение хеш-функции) – $O(n)$.

Структура алгоритма формирования речевого корпуса представлена на рисунке 1.

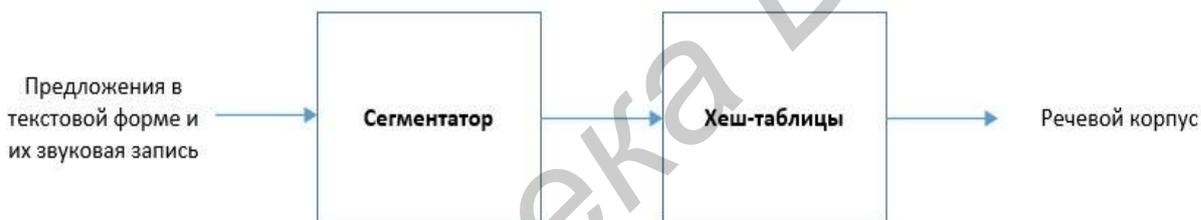


Рисунок 1 – Схема формирования речевого корпуса

Сегментатор преобразует исходные данные в набор речевых сегментов различной длины, которые затем заносятся в хеш-таблицы. В дальнейшем, при синтезе речи, из этих хеш-таблиц извлекаются необходимые элементы, которые преобразуются в целостную речь.

Использование скрытых марковских моделей для сегментации речи позволит разграничить необходимые сегменты с достаточно высокой точностью, а использование хеш-таблиц обеспечит быстрый доступ к элементам корпуса несмотря на его большой размер.

Список использованных источников:

1. Томас Х. Кормен. Алгоритмы: вводный курс.
2. Е. В. Пышкин. Структуры данных и алгоритмы: реализация на C/C++

МОДУЛЬ ОБРАБОТКИ СИГНАЛА ЦИФРОВОГО СЛУХОВОГО АППАРАТА

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

Бабок Е.И.

Вашкевич М.И. – к.т.н., доцент

Приблизительно 5% населения земли страдают от потери слуха, однако лишь небольшой процент от них фактически использует слуховой аппарат. Есть несколько факторов, влияющих на это. Во-первых, это неудовлетворенность качеством работы устройств, не оправдывающими ожиданий. Во-вторых, это стоимость,

связанная с производством новых слуховых аппаратов. Также одной из главных жалоб от владельцев слуховых аппаратов является отсутствие универсальности устройств - они усиливают все звуки (частоты), а не только те, которые хочет слышать пользователь.

Разрабатываемый модуль обработки сигнала для цифрового слухового аппарата состоит из 3 блоков:

- бинауральный эквалайзер;
- адаптивное шумоподавление;
- амплитудное сжатие;

Входной сигнал последовательно проходит через каждый из этих блоков.

Бинауральный эквалайзер реализуется с помощью двух банков фильтра, по одному для каждого уха и коэффициентов усиления для каждой полосы пропускания в банках фильтра. В банк фильтра входит от 1 до 50 фильтров с конечной импульсной характеристикой (FIR). Каждый из этих КИХ-фильтров имеет порядок от 50 до 200.

Блок адаптивного шумоподавления реализуется с помощью фильтра предсказаний и алгоритма RACE (Real-Time Adaptive Correlation Enhancer). RACE оценивает значения функции автокорреляции входных данных с использованием уравнения, заданного автокорреляционными оценками входного $x(n)$ при значениях k и временами n и $n - 1$ соответственно, а β – константа, значение которой находится между 0 и 1. Коэффициенты автокорреляции оцениваются для значений запаздывания k , варьирующихся от $-L$ до $+L$, где L называется запаздыванием. Это приводит к уникальному набору из $(2L + 1)$ коэффициентов автокорреляции. Обычно L выбирают между 5 и 7.

Скорость конвергенции или постоянную времени RACE можно определить, анализируя уравнение представленное ниже:

$$R(n, k) = \beta \cdot R(n - 1, k) + (1 - \beta) \cdot x(n) \cdot x(n + k)$$

Результат работы алгоритма, на примере зашумленного сигнала, показан на рисунке 1.

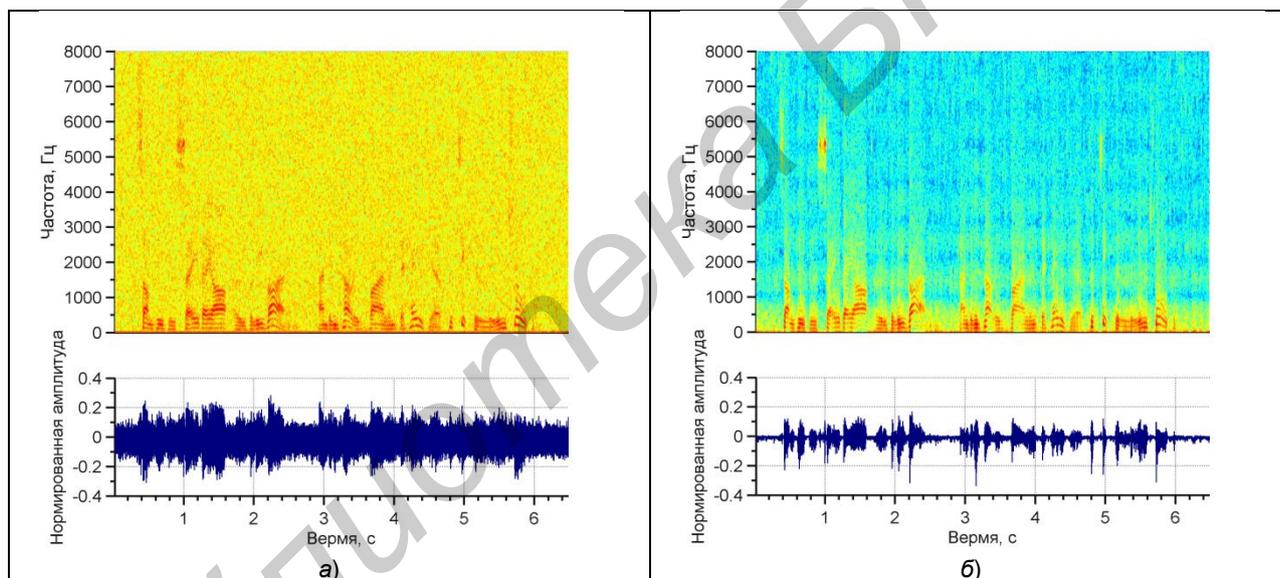


Рисунок 1 – Пример работы блока адаптивного шумоподавления: а) зашумленный сигнал до обработки; б) результат обработки сигнала

Помимо бинаурального эквалайзера и адаптивного шумоподавления, модуль также позволяет осуществлять амплитудное сжатие в режиме реального времени. Амплитудное сжатие достигается применением коэффициента усиления менее одного к сигналу, когда входная мощность превышает заданный порог.

Главной особенностью и плюсом данного модуля является то, что он может усиливать какой-то диапазон частот который хочет слышать пользователь или назначил врач, а не все звуки.

Список использованных источников:

1. T. Stetzler, N. Magotra, P. Gelabert, P. Kasthuri, S. Bangalore, Low-Power Real-Time Programmable DSP Development Platform for Digital Hearing Aids // Texas Instruments Application report (SPRA657 – April 2000).
2. А. Оппенгейм, Р. Шафер, Цифровая обработка сигналов // Москва: Техносфера, 2006 г. – 856 с.
3. Р. Лайонс Цифровая обработка сигналов: Второе издание // Москва: ООО «Бином-Пресс», 2006 г. – 656 с.