

СЕГМЕНТАЦИЯ РЕЧЕВЫХ СИГНАЛОВ НА ОСНОВЕ АНАЛИЗА ОСОБЕННОСТЕЙ РАСПРЕДЕЛЕНИЯ ДОЛЕЙ ЭНЕРГИИ ПО ЧАСТОТНЫМ ИНТЕРВАЛАМ¹

Жиляков Е.Г., Прохоренко Е.И., Болдышев А.В., Фирсова А.А.,
Фатова М.В.

*Национальный исследовательский университет
"Белгородский государственный университет", г. Белгород*

Одной из проблем обработки речевых сигналов является сегментация сигналов на звуки. Точность алгоритмов сегментации определяет надежность и эффективность использования в дальнейшем таких алгоритмов, как распознавание речи, синтез, сжатие.

Сегментация может осуществляться в ручную или автоматически. Сегментация в ручную является надежным, но трудоемким способом. Также сегментация в ручную невозможна при реализации обработки сигнала в режиме реального времени.

Наиболее интересной представляется сегментация речевых сигналов автоматически. Эффективность алгоритма сегментации определяется точностью определения границы между различными звуками.

Каждый звук имеет свое особенное распределение долей энергии по частотному диапазону. Эта особенность может быть использована для сегментации речи.

Анализ распределения энергии отрезков сигналов по частотным интервалам предлагается проводить на основе точного метода.

В результате проведенных экспериментов было выявлено, что длина фонем изменяется в пределах 1000-4000 отсчетов и зависит от типа звуко сочетания: открытый слог, закрытый слог, ударный слог, безударный слог и т.д. Сравнение ударных и безударных слогов показало, что если гласный стоит под ударением, то длительность слога возрастает примерно в $\frac{1}{4}$ раза. Спектры соответствующих звуков в ударном и безударном слогах отличаются незначительно. Использование представленного метода позволяет выявить место перестройки речевого аппарата с согласной на гласную и с гласной на согласную. Таким образом, данный метод может быть использован как один из элементов алгоритма сегментации речевого сигнала на отдельные звуки.

¹ Исследования выполнены при поддержке гранта РФФИ № 10-07-00326-а.