

СЕГМЕНТАЦИЯ РЕЧЕВЫХ СИГНАЛОВ НА ОСНОВЕ СУБПОЛОСНОГО АНАЛИЗА¹

Жиляков Е.Г., Фирсова А.А.

*Национальный исследовательский университет
«Белгородский государственный университет»,
г. Белгород*

В настоящее время возрастает интерес к разработке разнообразных речевых технологий, в том числе к созданию методов и алгоритмов автоматического распознавания речи. Обработке при этом подвергаются речевые сигналы (РС), которые являются результатами регистрации значений электромагнитных колебаний на выходе микрофонов при воздействии акустических колебаний на их входах, возникающих в результате речевого обмена. Одной из важных задач является сегментация РС, то есть разбиение их на отрезки, которые порождаются разными звуками речи или при их отсутствии (паузы речи). В данной работе эта задача рассматривается без идентификации порождающих акустические колебания звуков речи.

Основные рассматриваемые аспекты: обнаружение переходов речь/пауза и пауза/речь; моменты смены одного звука речи другим; выделение отрезков РС, порождаемых звуками речи с почти периодичной структурой, прежде всего вокализованных.

Известно, что все звуки русской речи обладают свойствами концентрации энергии в пределах малой доли частоты дискретизации. Поэтому, адекватным подходом к сегментации РС является применение субполосного анализа, когда их характеристики соотносятся некоторым разбиением области нормированных частот.

В результате проделанной работы был предложен новый метод селекции отрезков РС, порождаемых квазипериодическими звуками русской речи, основанный на введенном в работе понятии нормированной субполосной корреляции. Данный метод позволяет осуществлять селекцию отрезков РС, порождаемых квазипериодическими звуками русской речи с вероятностью ошибочного принятия решения не более 0,12, а также определить значение периода основного тона с погрешностью не более 0,02. Разработана новая решающая функция сегментации РС по границам звуков речи, основанная на использовании введенного в работе понятия субполосного расстояния. Предложенная решающая функция позволяет определять границы звуков русской речи с вероятностью ошибочного принятия решения не более $0,15 \div 0,20$.

¹ Работа выполнена в рамках проекта № 8.2251.2011 Государственного задания Министерства образования и науки РФ подведомственным ВУЗам на выполнение НИР в 2013 году