

## О Т З Ы В

на автореферат диссертации Томчука Кирилла Константиновича на тему:  
«Сегментация речевых сигналов для задач автоматической обработки речи»,  
представленной на соискание ученой степени кандидата технических наук  
по специальности 05.12.13 – Системы, сети и устройства телекоммуникаций

Речевые интерфейсы обладают неоспоримыми преимуществами при реализации диалога человека с ЭВМ, не требуя со стороны человека больших трудозатрат по вводу информации, а также повышая скорость такой коммуникации. В то же время, реализация любой системы, работающей с естественной речью, относится к сложным техническим задачам. Одним из важных этапов решения таких задач является автоматическая сегментация речевого сигнала. В свою очередь, понятие сегментации подразумевает обширную группу алгоритмов, в которую входят алгоритм разделения речевого сигнала на интервалы активной речи и свободных от речи пауз, алгоритмы выделения вокализованных фрагментов, шумных, взрывных звуков и т.д. Однако критическая оценка существующих алгоритмов данной группы позволяет выявить в них ряд проблем, проявляющихся в недостаточной эффективности, дикторнезависимости, шумоустойчивости.

В работе Томчука К.К. предложен новый комплекс алгоритмов сегментации, направленных на применение в задачах автоматической обработки речи: от передачи речевых сообщений до автоматического распознавания и синтеза. Разработанный соискателем алгоритм сегментации на периоды колебаний голосовых связок отличается от аналогов высоким быстродействием. В предложенном в работе алгоритме сегментация речь/пауза учитывается ряд известных закономерностей звуковосприятия человека (компенсация отсутствия основной гармоники гармониками кратных частот, психофизиологический закон интенсивности ощущения Вебера-Фехнера). Также предложен метод увеличения шумоустойчивости оценки мел-частотных кепстральных признаков, основанный на применении модели одновременной слуховой маскировки из стандарта MPEG-1 Layer 1 и модификации спектральной плотности мощности сигнала на частотах кратных гармоник основного тона.

Перечисленный комплекс алгоритмов был впервые применен к задаче модификации темпа произнесения речи с сохранением голоса диктора и естественного соотношения длительностей разных типов сегментов.

ГУАП

№ 74-1187/17-0-0  
от 01.06.2017



Автореферат диссертационного исследования отражает ключевые результаты, полученные автором, имеет целостную структуру и оформлен в соответствии с требованиями. Пункты новизны и основные защищаемые положения сформулированы грамотно и четко. Выводы к работе свидетельствуют о решении поставленных задач и достижении заявленной цели. Дополнительно можно отметить, что данная работа крайне полезна для использования в преобразовании сигналов речевой информации в сигналы управления воздушным движением летательных аппаратов.

Из автореферата можно отметить как недостаток. Недостаточно полно описаны детали работы предлагаемого алгоритма модификации темпа речи: не приведены конкретные методы обработки, применяемые к сегментам разных типов, и не обозначены соответствующие парциальные коэффициенты;

Указанный недостаток не снижает общего положительного результата, достигнутого в работе.

Результаты, полученные диссертантом, имеют теоретическое и практическое значение, сформулированные выводы и рекомендации обоснованы в полной мере. Работа отвечает требованиям ВАК, предъявляемым к кандидатским диссертациям, а ее автор Томчук Кирилл Константинович заслуживает присуждения ученой степени кандидата технических наук по специальности 05.12.13 – Системы, сети и устройства телекоммуникаций.

Доктор технических наук,  
профессор,  
Почетный работник высшего  
профессионального образования РФ

И.А. Вельмисов

Подпись И.А. Вельмисова удостоверяю.

Ученый секретарь АО «ВНИИРА»,  
д.т.н., профессор,  
Заслуженный деятель науки РФ,  
Лауреат Государственных,  
Правительства СПб и Международной (ИКАО)  
премии



Ю.Г. Шатраков