

## О Т З Ы В

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

Развитие, реализация, внедрение и эксплуатация современных речевых систем выдвигают на передний план решение вопросов эффективной предобработки речевых сигналов, важную роль среди которых играет временная сегментация.

Однако существующие подходы к сегментации не достаточно систематизированы, а некоторые и вовсе имеют неравномерный характер решающих функций. Кроме того, речевой сигнал человека отличается от искусственных сигналов высокой вариабельностью, что отражается в недостаточной эффективности существующих частных алгоритмов сегментации.

Таким образом, актуальность диссертации Томчука К. К., в которой разрабатываются методы и алгоритмы сегментации речевых сигналов для последующего применения результатов сегментации в речевых задачах, не вызывает сомнения.

В ходе диссертационного исследования автором была определена и решена совокупность частных научных задач исследования, при этом получен ряд новых научных результатов:

- база данных по реализациям звуков речи, которая позволяет хранить и извлекать по многокритериальным условиям информацию о значениях сигнальных параметров звуков, о дикторе, относительному положению звука в слове, и может быть использована для настройки и обучения алгоритмов сегментации;

- алгоритм сегментации на периоды основного тона, который за счёт работы во временной области только с отсчетами локальных экстремумов сигнала обладает высоким быстродействием и может быть успешно применен вместо менее производительных аналогов при условии воздействия внешних акустических помех с ОСШ более 5 дБ;

- модифицированный алгоритм параметризации речевого сигнала коэффициентами MFCC, который, позволяет получить среднее улучшение пословной точности распознавания на уровне 12% при воздействии внешних акустических помех;

ГУАП

№ 74-1099/17-0-0

от 23.05.2017



- метод модификации темпа произнесения речи, основанный на разработанных в диссертации алгоритмах многоуровневой временной сегментации речевых сигналов, позволяет получать ускоренную или замедленную в несколько раз речевую фонограмму с сохранением голоса диктора и низким количеством артефактов звучания.

Обоснованность и достоверность научных положений, основных выводов и результатов диссертационной работы обеспечивается всесторонним анализом современного состояния исследований в проблемной области, подтверждается корректностью предложенных результатов, достаточной апробацией основных теоретических положений диссертации в печатных трудах и докладах на российских и международных научных и научно-практических конференциях и семинарах.

Автореферат написан корректно и чётко, оформлен в соответствии с требованиями ГОСТ и соответствует специальности 05.12.13 – «Системы, сети и устройства телекоммуникаций».

В качестве недостатков следует отметить следующее:

1. В автореферате недостаточно внимания уделено вопросам применения результатов сегментации в конкретных речевых задачах.
2. Можно было провести исследование не только на базе фонограмм TIDIGITS.

Данные замечания не снижают общей оценки диссертации. Диссертационная работа удовлетворяет требованиям, предъявляемым к кандидатским диссертациям, а её автор, Томчук К.К. заслуживает присуждения учёной степени кандидата технических наук.

Диссертация Томчука К.К. является законченной научно-квалификационной работой и соответствует требованиям п. 9 «Положения о присуждении учёных степеней».

« 19 » Мая 2017 года

Ведущий инженер по внедрению  
программного обеспечения  
кандидат технических наук  
197101, г. Санкт-Петербург,  
ул. Большая Монетная 16-1, лит. В  
тел.: 8 (812) 336-55-33  
e-mail: [pavel.balakshin@firstlinesoftware.com](mailto:pavel.balakshin@firstlinesoftware.com)

Балакшин Павел Валерьевич

Подпись Балакшина П.В. подтверждаю:  
Менеджер по персоналу



Лысак Алина Анатольевна