

«УТВЕРЖДАЮ»  
Ректор Ярославского  
государственного университета  
им. П.Г. Демидова

А.И. Русаков

«24» 06 2015 г.

## ЗАКЛЮЧЕНИЕ

Федеральное государственное бюджетное образовательное учреждение  
высшего профессионального образования  
Ярославский государственный университет им. П.Г. Демидова

Диссертация «Разработка и исследование коллективных нейросетевых алгоритмов дикторонезависимого распознавания речевых сигналов» выполнена на кафедре динамики электронных систем ФГБОУ ВПО «Ярославского государственного университета им. П.Г. Демидова» (ЯрГУ им. П.Г. Демидова).

В период подготовки диссертации соискатель Сагациян Максим Владимирович работал в ООО ПСЦ «Электроника» в должности инженера-проектировщика и руководителя работ, в ЯрГУ им. П.Г. Демидова в должности преподавателя специальных дисциплин.

В 2011 г. окончил с отличием Государственное образовательное учреждение высшего профессионального образования «Ярославский государственный университет им. П.Г. Демидова» по специальности «Радиофизика и электроника» с присуждением квалификации радиофизик.

Удостоверение о сдаче кандидатских экзаменов выдано в 2015 г. ЯрГУ им. П.Г. Демидова.

Соискатель обучается в аспирантуре Ярославского государственного университета по очной форме по специальности 05.12.04 Радиотехника, в том числе системы и устройства телевидения (срок окончания обучения – 31 октября 2015 года).

Научный руководитель – доктор технических наук, профессор, заведующий кафедрой динамики электронных систем ЯрГУ им. П.Г. Демидова Брюханов Юрий Александрович.

По результатам рассмотрения диссертации «Разработка и исследование коллективных нейросетевых алгоритмов дикторонезависимого распознавания речевых сигналов» принято следующее заключение.



*В диссертационной работе изложена* методика разработки и результаты исследования коллективных нейросетевых алгоритмов дикторонезависимого распознавания речевых сигналов.

*Наиболее существенные научные результаты, полученные лично соискателем:*

– Разработан алгоритм bagging-коллектива на основе перцептронов Розенблатта с обучением SCG и его модификация для решения задачи дикторонезависимого распознавания русскоязычных речевых сигналов.

– Созданы речевые базы «КРИПТОН-01» с размерностью 10 сигналов и «КРИПТОН-02» с размерностью 102 сигнала для тестирования нейросетевых алгоритмов.

– Проведен анализ параметров bagging-коллектива многослойных перцептронов Розенблатта с обучением SCG, в результате чего определено, что рациональнее: выбирать размер bagging-коллектива 10; использовать 10 обучающих дикторов; устанавливать по 12 слоев в каждом нейросетевом распознавателе и использовать размер словаря не больше 10. При данных параметрах получена вероятность дикторонезависимого распознавания русскоязычных речевых сигналов 97,1 %, что на 4,1 процентных пункта выше существующих результатов. Учитывая доверительный интервал полученного значения  $\pm 2,8$  процентных пункта, следует, что с вероятностью 0,95 точность распознавания речевых сигналов также лучше существующих результатов.

– Разработана научно-исследовательская программа, с помощью которой можно проводить анализ алгоритмов дикторонезависимого распознавания русскоязычных речевых сигналов путем математического моделирования данных алгоритмов, обученных и протестированных на речевых базах «КРИПТОН-01» и «КРИПТОН-02». Авторские права защищены свидетельством о государственной регистрации программы для ЭВМ.

– Проведены исследования по определению параметров для нейросетевого bagging-коллектива дикторонезависимого распознавания русскоязычных речевых сигналов.

– Проведены исследования модифицированных алгоритмов на основе двух разновидностей нейронных сетей: 10 перцептронов Розенблатта с обучением SCG и 10 сетей Эльмана с обучением GDX. Модифицированный bagging-коллектив на основе 10 сетей Эльмана распознал 102 речевых сигнала с вероятностью распознавания 91,5 % при времени обучения данного алгоритма 3030 секунд и времени тестирования 380 секунд. Модифицированный bagging-коллектив на основе 10 перцептронов Розенблатта распознал 102 речевых сигнала с вероятностью распознавания 95,7 % при времени обучения данного алгоритма 2688 секунд и времени тестирования 381 секунд, что на 5,29 процентных пункта выше существующих результатов. Учитывая доверительный интервал  $\pm 3,2$  процентных пункта для полученного значения вероятности распознавания модифицированного bagging-коллектива на основе перцептронов Розенблатта с обучением SCG, следует, что с вероятностью 0,95 точность распознавания



речевых сигналов также лучше существующих результатов.

– Представлены коллективный и модифицированный коллективный нейросетевые алгоритмы с блоком шумоподавления для решения задачи дикторонезависимого распознавания русскоязычных речевых сигналов в условиях шумов.

– Проведены исследования работы коллективного и модифицированного коллективного нейросетевых алгоритмов с блоками шумоподавления дикторонезависимого распознавания речевых сигналов в условиях шумов. В результате этого показаны высокие результаты распознавания в условиях слабых шумов.

***Достоверность полученных результатов*** подтверждена согласованностью результатов математического моделирования разработанных алгоритмов и экспериментальной проверки в условиях полунатурного моделирования на реальных речевых сигналах, апробацией в печати и на научно-практических конференциях различного уровня.

#### ***Научная новизна полученных результатов***

Впервые получены следующие научные результаты:

– Разработан нейросетевой алгоритм bagging-коллектива на основе перцептронов Розенблатта с обучением масштабируемых сопряженных градиентов (Scaled Conjugate Gradient Backpropagation, SCG), позволяющий решать задачу дикторонезависимого распознавания русскоязычных речевых сигналов для малого словаря с вероятностью распознавания 97,1 %, что на 4,1 процентных пункта выше существующих результатов.

– Предложена модификация коллективного нейросетевого алгоритма, позволяющая результативно решать задачу дикторонезависимого распознавания русскоязычных речевых сигналов.

– Разработан модифицированный коллективный нейросетевой алгоритм на основе перцептронов Розенблатта с обучением SCG, позволяющий решать задачу дикторонезависимого распознавания русскоязычных речевых сигналов для большого словаря с вероятностью распознавания 95,7 %, что на 5,29 процентных пункта выше существующих результатов.

– Разработан коллективный и модифицированный коллективный нейросетевые алгоритмы с блоками шумоподавления дикторонезависимого распознавания русскоязычных речевых сигналов, работающие в условиях шумов.

#### ***Практическая значимость результатов исследования***

– Предложенная модификация коллективного нейросетевого алгоритма расширяет возможности нейросетевых алгоритмов дикторонезависимого распознавания русскоязычных речевых сигналов.

– Вероятность распознавания речевых сигналов для разработанного нейросетевого алгоритма bagging-коллектива на основе перцептронов Розенблатта с обучением SCG с блоком шумоподавления для малого словаря в



интервале от 5 до 20 дБ равняется 93,5 % при использовании алгоритма шумоподавления Скалара на основе винеровской фильтрации.

– Вероятность распознавания речевых сигналов для разработанного модифицированного нейросетевого алгоритма bagging-коллектива на основе перцептронов Розенблатта с обучением SCG с блоком шумоподавления для большого словаря в интервале от 15 до 20 дБ равняется 93,6 % при использовании алгоритма шумоподавления на основе бинарных масок, использующего критерий статистического детектирования на основе апостериорного отношения сигнал/шум.

– Разработана программа «NN-SCG speech recognition» (Свидетельство о регистрации программы для ЭВМ № 2015616920), с помощью которой можно проводить анализ различных алгоритмов нейросетевого дикторонезависимого распознавания русскоязычных речевых сигналов.

Результаты работы внедрены в соответствующие разработки ООО «ПАНТЕОН» (г. Ярославль) и ООО «А-Вижн» (г. Ярославль). Все результаты внедрения подтверждены соответствующими актами.

**Ценность научных работ соискателя** состоит в разработке и исследовании коллективных нейросетевых алгоритмов дикторонезависимого распознавания речевых сигналов.

#### ***Специальность, которой соответствует диссертация***

Представленная Сагацияном М.В. диссертационная работа соответствует специальности 05.12.04 Радиотехника, в том числе системы и устройства телевидения.

*Направление диссертационной работы соответствует области исследований:*

– Разработки методов приема, обработки, отображения и хранения информации. То есть, в диссертационной работе исследуется разработка методов приема, обработки, отображения и хранения информации дикторонезависимого распознавания русскоязычных речевых сигналов в радиотехнических устройствах.

– Разработки перспективных информационных технологий, в том числе цифровых, в радиотехнических устройствах. То есть, с помощью систем обработки, выполняющих дикторонезависимое распознавание русскоязычных речевых сигналов, возможно повысить эффективность радиотехнических устройств, таких как, например, радиоприемника, рации, мобильного телефона, телевизионного устройства, сканера магнитно-резонансной томографии, рентгеновского сканера и др.

**Полнота изложения материалов диссертации в опубликованных работах**

По теме диссертации соискателем опубликовано 16 научных работ, из них 2 статьи в журналах, рекомендованных ВАК, получено 1 свидетельство о регистрации программы для ЭВМ.

**Публикации в рецензируемых изданиях, рекомендованных ВАК РФ**

1. Сагациян, М.В. Разработка и исследование нейросетевого алгоритма дикторонезависимого распознавания речевых команд / М.В. Сагациян, А.В. Куликов, Г.С. Тупицин // Вестник Поволжского государственного технологического университета. Сер.: Радиотехнические и инфокоммуникационные системы. – Йошкар-Ола: 2014. – № 1(20). – С. 62–68.

2. Сагациян, М.В. Анализ эффективности нейросетевых алгоритмов в задаче дикторонезависимого распознавания речевых команд / М.В. Сагациян, Г.С. Тупицин // Информационные системы и технологии. – Орел: 2015. – № 3. – С. 16–26.

Диссертация «Разработка и исследование коллективных нейросетевых алгоритмов дикторонезависимого распознавания речевых сигналов» Сагацияна Максима Владимировича рекомендуется к защите на соискание ученой степени кандидата технических наук по специальности 05.12.04 Радиотехника, в том числе системы и устройства телевидения.

Заключение принято на заседании кафедры динамики электронных систем.

Присутствовало на заседании 10 чел.

Результаты голосования: «за» – 10 чел., «против» – 0 чел., «воздержалось» – 0 чел., протокол № 14 от «24» 06 2015 г.

Председательствующий на заседании  
кафедры динамики электронных  
систем ЯрГУ им. П.Г. Демидова



А.Л. Приоров

