

«УТВЕРЖДАЮ»

Первый заместитель генерального директора-
технический директор

ОАО «Ярославский радиозавод»

В.Л. Филимонов

«09» сентября 2015 г.



ОТЗЫВ

ведущей организации ОАО «Ярославский радиозавод»

на диссертацию Сагацияна Максима Владимировича

«Разработка и исследование коллективных нейросетевых алгоритмов дикторонезависимого распознавания речевых сигналов», представленную на соискание учёной степени кандидата технических наук по специальности 05.12.04 Радиотехника, в том числе системы и устройства телевидения

Актуальность проведённых исследований

Диссертационная работа Сагацияна М.В. посвящена одному из актуальных направлений исследований в области проектирования речевых интерфейсов для различных устройств, в том числе радиотехнических. Исследуемая задача направлена на повышение качества дикторонезависимого распознавания русскоязычных речевых сигналов. Она возникает при построении более удобного интерфейса для различных технических устройств.

В настоящее время распознавание речевых сигналов находит все новые области применения, начиная от приложений, осуществляющих преобразование речевой информации в текст и заканчивая бортовыми устройствами управления транспортными средствами.

С помощью вспомогательных технологий распознавания речевых сигналов специалисты, с ограниченными возможностями по зрению, могут выполнять многие виды работ наравне со своими коллегами.

Развитие данной тематики в диссертационной работе имеет важное теоретическое и практическое значение для радиотехники и ряда смежных областей.

Выполненный автором анализ существующих алгоритмов и предложенные методы распознавания речевых сигналов позволили создать программно-аппаратный комплекс, направленный на решение задач дикторонезависимого распознавания русскоязычных речевых сигналов.

Научная новизна проведенных исследований и полученных результатов

Наиболее значимые научные результаты диссертационной работы состоят в следующем:

1. Разработан нейросетевой алгоритм bagging-коллектива на основе персептронов 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.

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

3. Разработан модифицированный коллективный нейросетевой алгоритм на основе персептронов 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.

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

Новизна и достоверность научных положений и выводов, сформулированных в диссертации, подтверждены апробацией на международных и всероссийских научных семинарах и конференциях.

По теме диссертационной работы опубликовано 17 научных работ, из них 3 статьи в журналах, рекомендованных ВАК. Получено свидетельство о государственной регистрации программы для ЭВМ.

Практическая значимость полученных в диссертации результатов

Практическая значимость диссертационной работы подтверждается следующими положениями:

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

2. Вероятность распознавания речевых сигналов для разработанного нейросетевого алгоритма bagging-коллектива на основе персептронов Розенблатта с обучением SCG с блоком шумоподавления для малого словаря в интервале от 5 до 20 дБ достигает 93,5 % при использовании алгоритма шумоподавления Скалара на основе винеровской фильтрации.

3. Вероятность распознавания речевых сигналов для разработанного модифицированного нейросетевого алгоритма bagging-коллектива на основе персептронов Розенблатта с обучением SCG с блоком шумоподавления для большого словаря в интервале от 15 до 20 дБ достигает 93,6 % при использовании алгоритма шумоподавления на основе бинарных

масок, использующего критерий статистического детектирования на основе апостериорного отношения сигнал/шум.

4. Разработана программа «NN-SCG speech recognition» (свидетельство о государственной регистрации программы для ЭВМ № 2015616920), с помощью которой проведен 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.

Перечисленные результаты являются хорошей основой для разработки перспективных систем распознавания речевых сигналов для различных технических устройств, в том числе радиотехнических.

Замечания по диссертационной работе

1. Для исследования нейросетевых алгоритмов применялись речевые базы с ограниченным количеством обучающих и тестирующих дикторов. Не произведено исследование разработанных нейросетевых алгоритмов при количестве обучающих дикторов более 12 и тестирующих дикторов более 12.

2. Не исследуется работа коллективных и модифицированных коллективных нейросетевых алгоритмов при различных видах шумов, кроме белого гауссовского шума.

3. Не понятно, почему выбрано три определенных алгоритма шумоподавления.

4. Из приведенного в работе описания не ясно, насколько разные дикторы включались в эксперимент при записи собственных речевых баз.

5. Результаты диссертационной работы не реализованы в программно-аппаратной части для указанных в работе радиотехнических устройств, таких как, например, радиостанция или блок управления

магниторезонансным томографом. Реализация произведена только на персональном компьютере.

Выводы

Диссертация «Разработка и исследование коллективных нейросетевых алгоритмов дикторонезависимого распознавания речевых сигналов» является законченной научно-квалификационной работой, выполненной на высоком научном уровне. Автореферат отражает основное содержание диссертационной работы. Она соответствует требованиям ВАК к кандидатским диссертациям, а ее автор Сагацян Максим Владимирович заслуживает присуждения ученой степени кандидата технических наук по специальности 05.12.04 Радиотехника, в том числе системы и устройства телевидения.

Заместитель технического директора по
НИОКР – начальник ИНТЦ
ОАО «Ярославский радиозавод»



С.С. Поелуев