

Межрегиональное общественное учреждение
"Институт инженерной физики"
(Научное, образовательное и производственное учреждение)
(МОУ "ИИФ")

Большой Ударный пер., д. 1а, г. Серпухов, Московская обл., 142210
тел. 8(4967)353193; 8(4967)351371; 8-499-400-05-75; факс: 354420; e-mail: info@iifmail.ru; www.iifrf.ru
ОКПО 42232569, ОГРН 1035000009417, ИНН/КПП 5043014134/504301001

№ _____
на № _____ от _____

УТВЕРЖДАЮ

Первый Вице-резидент Института –
Главный конструктор
доктор технических наук, профессор



С.В. Смуров
2019 г.

ОТЗЫВ

на автореферат диссертации Аль-Дайбани Абдулгани Мохаммеда Салеха на тему: «ИССЛЕДОВАНИЕ МЕТОДОВ И РАЗРАБОТКА АЛГОРИТМОВ ОБРАБОТКИ СИГНАЛОВ ДЛЯ СИСТЕМ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ ТЕЛЕФОННОЙ РЕЧИ В РЕСПУБЛИКЕ ЙЕМЕН» по специальности 05.12.13 – Системы, сети и устройства телекоммуникаций

Развитая телефонная сеть в республике Йемен и современный уровень вычислительной техники создают предпосылки использования систем автоматического распознавания речи (САРР) в телефонии. Использование САРР обеспечивает простой доступ широких слоев населения к автоматическим справочным и регистрационным системам. Однако неспокойная политическая обстановка в стране и отсутствие инженерных и научных кадров соответствующей квалификации не позволяют до сих пор построить соответствующие системы.

Следует учесть, что особенности арабского языка повышают сложность создания САРР по сравнению с аналогичными системами,

используемыми, например, в Европе и США. В частности, разговорный арабский язык характеризуется множеством диалектов.

Снижение достоверности распознавания во многом обусловлено отличием частотной характеристики (ЧХ) канала связи, который использовался при создании звукозаписей, предназначенных для обучения САРР, от ЧХ каналов связи, которыми пользуются абоненты телефонных систем в процессе эксплуатации САРР. Указанное снижение достоверности обусловлено зависимостью параметров речевого сигнала (РС), которые используются при распознавании, от ЧХ канала связи. Для снижения зависимости используется нормализация параметров сигнала по их среднему значению. Однако отсутствуют исследования, связанные с оценкой влияния различных факторов на степень стабилизации значений нормализованных параметров РС при изменении ЧХ канала связи. Именно поэтому, диссертационная работа Аль-Дайбани А.М., направленная на разработку алгоритмов обработки речевых сигналов, является актуальной и имеющей практическое значение для систем САРР.

В ходе выполнения диссертационной работы автором получены результаты, обладающие научной новизной и практической значимостью:

1. Методика и соответствующий алгоритм оценки эффективности применения спектрального вычитания и фильтра Винера для повышения помехоустойчивости САРР.

2. Методика и соответствующий алгоритм оценки эффективности нормализации МЧКК для снижения влияния ЧХ канала связи на достоверность распознавания голосовых команд.

3. Алгоритм идентификации диалектов арабского языка республики Йемен, обеспечивающий относительную ошибку идентификации равную 0,24% и позволяющий повысить достоверность распознавания арабских названий цифр, как минимум, на 7%.

Практическая значимость результатов диссертационных исследований заключается в том, что полученные в ней научно-технические предложения доведены до уровня математического обеспечения соответствующих алгоритмов.

Судя по автореферату, результаты, выдвигаемые на защиту, нашли практическую реализацию в ООО «Центр речевых технологий» (г. Санкт-Петербург), а также в учебном процессе ВлГУ.

Содержание диссертационных исследований и результаты, выдвигаемые на защиту, докладывались на научно-технических конференциях различного уровня, достаточно полно опубликованы как в рецензируемых научных изданиях, так и в других изданиях научного плана по профилю заявленной научной специальности.

К недостаткам диссертации, вытекающим из анализа автореферата, следует отнести:

1) при использовании системы распознавания речи совместно с идентификацией диалекта возникает противоречие: для того чтобы идентифицировать диалект, необходимо знать речевую команду (РК), на основе которой производится идентификация, а для того чтобы с повышенной точностью распознать речь, необходимо знать диалект. В работе предварительная идентификация диалекта производится на основе известной РК, однако данный подход на практике не применим и не понятно, как на практике описанное выше противоречие будет разрешаться;

2) влияние вида оконной функции при формировании мел-частотных кепстральных коэффициентов (МЧКК) и неравномерности АЧХ проведено на малой группе слов, относящейся к определенному типу (наборы цифр). Вызывает определенное сомнение, что данные зависимости будут иметь место при анализе другого словарного запаса;

3) в работе упор сделан на акустические модели на основе скрытых Марковских моделей и МЧКК (в качестве параметров речевых сигналов), однако в текущий момент широко используются нейросетевые алгоритмы и другие варианты признакового пространства.

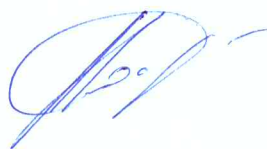
Однако указанные замечания не снижают положительной оценки по работе в целом.

Исходя из автореферата, можно сделать вывод о том, что диссертационная работа Аль-Дайбани А.М. является законченной научно-квалификационной работой, содержащей решение актуальной научной задачи, имеющей важное значение для дальнейшего развития теории и

практики цифровой обработки сигналов в республике Йемен. Диссертация отвечает требованиям п.п. 9, 10, 11 и 14 «Положения о присуждении ученых степеней», поэтому ее автор заслуживает присуждения ученой степени кандидата технических наук по специальности 05.12.13 – Системы, сети и устройства телекоммуникаций.

Отзыв составили:

Старший научный сотрудник управления АСУ и связи
кандидат технических наук



А.И. Косяк

Старший специалист управления АСУ и связи
кандидат технических наук



К.В. Карпочкин

Отзыв рассмотрен на заседании НТС МОУ «ИИФ» протокол №18/11/01
от «18» ноября 2019 года.