

Отзыв официального оппонента

ведущего специалиста АО «КОБРА», кандидата технических наук,
на диссертационную работу Аль-Дайбани Абдулгани Мохаммеда Салеха
«Исследование методов и разработка алгоритмов обработки сигналов для
систем автоматического распознавания телефонной речи в республике
Йемен», представленной на соискание ученой степени кандидата
технических наук по специальности
05.12.13 – «Системы, сети и устройства телекоммуникаций»

На данный момент времени телефонная сеть в Йемене достаточно
развита, в связи с чем возникает достаточно большой спрос на
автоматизированные справочные и регистрационные системы.
Необходимость в них, а также высокий уровень вычислительной техники
создают предпосылки для организации данных систем, что обеспечит
простой доступ широких слоев населения к справочной информации и
позволит повысить уровень жизни.

Особенностью разговорной арабской речи является наличие множества
диалектов даже в пределах одной страны, что в значительной мере приводит
к снижению качества распознавания. Указанная особенность обуславливает
сложность создания систем автоматического распознавания речи, ввиду чего
количество разработок в этой области для арабских стран крайне мало. В
Йемене подобные разработки вообще отсутствуют.

Отдельно следует отметить проблемы, связанные с тем, что реальных
условиях речь сопровождается большим количеством помех, а при попытке
подавления помех искажаются полезные сигналы и снижается достоверность
распознавания. Также сильное влияние на качество распознавания оказывает
частотная характеристика канала связи.

Таким образом, в настоящее время существует актуальная техническая и
научная задача исследования существующих методов обработки речевых
сигналов, а также разработки эффективных алгоритмов обработки данных

сигналов, предназначенные для арабской республики Йемен. Поэтому решаемая в диссертационной работе задача повышения устойчивости систем автоматического распознавания речи к воздействию внешних факторов на сигнал является **актуальной**.

Новизна полученных результатов заключается в следующем:

1. В ходе работы были получены выражения, позволяющие оценить вероятность ошибки идентификации диалекта при произнесении контрольных слов. При идентификации используются акустические модели произнесений контрольных слов, созданные для каждого диалекта.

2. В результате эксперимента получены значения достоверности распознавания при идентификации Йеменских диалектов. Показано, что при использовании акустических моделей произнесений контрольных слов, в значительной мере повышается достоверность распознавания.

3. Получены выражения, определяющие зависимость значений мел-частотных кепстральных коэффициентов, нормализованных по среднему значению, от вида оконной функции, используемой при дискретном преобразовании Фурье, и неравномерности АЧХ канала связи.

4. В результате экспериментального исследования влияния различных оконных функции на значения нормализованных мел-частотных кепстральных коэффициентов, были выделены оконные функции, позволяющие уменьшить влияние частотной характеристики канала связи на нормализованные параметры речевого сигнала.

5. В результате экспериментального исследования возможностей спектрального вычитания и фильтра Винера по подавлению помех при автоматическом распознавании речи, были оценены их возможности снижения влияния аддитивных помех на достоверность распознавания речи в Йемене.

Практическими результатами работы являются:

1. Для повышения помехоустойчивости систем распознавания речи разработаны методика и соответствующий ей алгоритм оценки возможно-

стей спектрального вычитания и фильтра Винера по снижению влияния помех на значения параметров речевых сигналов и точность распознавания речи.

2. Разработаны методика и соответствующий алгоритм оценки эффективности нормализации мел-частотных кепстральных коэффициентов для снижения влияния частотной характеристики канала связи на достоверность распознавания речевых сообщений Йеменского диалекта.

3. Разработано и прошло государственную регистрацию программное обеспечение, реализующее разработанные алгоритмы, которое позволяет обеспечить оптимальную настройку средств подавления влияния помех и частотной характеристики канала связи на работу систем распознавания речи.

4. Предложен алгоритм идентификации диалектов, который обеспечивает относительную ошибку идентификации равную 0,24%. что позволяет повысить достоверность распознавания арабских названий цифр Йеменского диалекта, как минимум, на 7%.

5. Составлены и обработаны выборки аудиозаписей для обучения и тестирования систем автоматического распознавания арабской речи Йеменского диалекта.

Достоверность полученных автором результатов не вызывает сомнений, т.к. определяется корректным использованием математического аппарата и имитационным моделированием, а также практическим использованием, о чем свидетельствуют акты внедрения полученных диссертантом результатов в ООО «Центр речевых технологий» (г. Санкт-Петербург), а также в учебный процесс на кафедре радиотехники и радиосистем Владимирского государственного университета имени Александра Григорьевича и Николая Григорьевича Столетовых (ВлГУ).

Материалы диссертации прошли **апробацию** в виде выступлений автора на четырех научно-технических конференциях, и в виде публикаций в сборниках трудов и журналах, в т.ч. 3 статьи в журналах, рекомендованных

ВАК, 7 - на международных конференциях (одна работа - в издании IEEE, индексируемом SCOPUS). Автором получено 4 свидетельства о государственной регистрации программ для ЭВМ.

Диссертационная работа состоит из введения, четырёх глав, заключения, библиографического списка, включающего 100 наименований, списка сокращений и 7 приложений. Общий объём диссертации 130 страниц, в том числе 111 страниц основного текста.

Во введении содержится постановка проблемы, сформулированы цель исследования и решаемые задачи, показаны научная новизна и практическая значимость полученных результатов, определены положения, выносимые на защиту.

В первом разделе диссертантом рассмотрены особенности и проблемы построения систем автоматического распознавания арабской речи, поскольку арабский язык характеризуется большим разнообразием диалектов. Выявлена необходимость включения идентификатора диалекта в состав системы распознавания речи. Сделаны выводы, что достаточно высокая точность идентификации достигается при использовании акустических моделей, на основе скрытых Марковских моделей и при использовании мел-частотных кепстральных коэффициентов в качестве параметров речевых сигналов. Также в данной главе показано, что для повышения устойчивости мел-частотных кепстральных коэффициентов к изменениям частотной характеристики канала связи используют нормализацию коэффициентов по их среднему значению. Большое количество ошибок, возникающих при автоматическом распознавании речи, обусловлено влиянием акустических помех, которые сопровождают речевой сигнал.

Для очистки зашумленной речи применяют спектральное вычитание и фильтр Винера. Исследованы возможности подавления помех с помощью спектрального вычитания и фильтра Винера на примере автоматического распознавания названий цифр на диалектном языке Йемена

Во второй главе рассмотрено использование идентификации диалекта для повышения достоверности распознавания. Рассмотрена возможность использования акустических моделей произнесений тестовых слов для оперативной идентификации одного из трех йеменских диалектов в процессе обращения пользователя к автоматической телефонной справочной системе на примере арабских названий отдельных цифр. Автор работы делает вывод о целесообразности идентификации диалектов, поскольку вероятность правильного распознавания увеличивается с 91,6 % до 98,1 % - 98,3 % (в зависимости от диалекта).

В третьей главе исследована возможность применения нормализации параметров сигнала для подавления влияния частотной характеристики канала связи на работу системы автоматического распознавания речи. Рассмотрено влияние вида оконной функции и неравномерности амплитудно-частотной характеристики канала связи на степень стабилизации нормализованных параметров сигнала. Получено выражение для разности нормированных логарифмов спектров сигнала для случаев обучения и тестирования системы автоматизированного распознавания речи. С помощью имитационного моделирования доказано, что точность распознавания увеличивается с использованием нормализации параметров речевого сообщения. Использование окон Чебышева и Ханна, которые имеют малый уровень боковых лепестков, удаленных от основного лепестка, обеспечивают большую достоверность распознавания по сравнению с окном Хэмминга

В четвертой главе представлено описание программного комплекса, предназначенного для оценки помехоустойчивости систем распознавания речи в телефонии, а именно:

- возможность комплекса оценить эффективность подавления аддитивных помех с помощью спектрального вычитания, а также путем использования фильтра Винера;

- возможность комплекса оценить нормализацию параметров речевого сигнала по их среднему значению во времени, что позволяет уменьшить

влияние изменчивости частотной характеристики канала связи на точность распознавания.

В заключении сформулированы основные результаты работы, сделаны выводы о ее практической значимости.

В приложениях представлены результаты экспериментов классификации диалектов на три группы и на две группы, характеристики исследуемых шумов и результаты их влияния на достоверность системы распознавания, Результат тестирования систем автоматического распознавания речи, акты внедрения, свидетельства о регистрации программ для ЭВМ..

Основные положения и выводы, сформулированные в работе, обоснованы, проверены и отражены в научных публикациях. Автореферат в достаточной степени отражает содержание диссертации.

В качестве замечаний по диссертационной работе следует отметить следующее:

1. В диссертационной работе не проведена оценка влияния внешних шумов на распознавание диалекта Йеменской группы диалектов.

2. В работе не приведена оценка производительности требуемых вычислительных систем для реализации методов и соответствующих алгоритмов, разработанных автором.

3. При исследовании влияния шума на достоверность распознавания не рассмотрен шум в виде посторонней фоновой речи, как один из наиболее распространенных в повседневной жизни.

Несмотря на указанные замечания, диссертация Аль-Дайбани Абдулгани Мохаммеда Салеха является завершенной научно-исследовательской работой, содержащей исследование методов, разработку алгоритмов и программных средств для обработки сигналов с целью повышения помехоустойчивости телефонных систем автоматического распознавания арабской речи в республике Йемен.

На основании изложенного считаю, что диссертационная работа удовлетворяет требованиям п.п. 9, 10, 11, 13, 14 Положения о присуждении ученых степеней, а ее автор **Аль-Дайбани Абдулгани Мохаммед Салех** заслуживает присуждения ученой степени кандидата технических наук по специальности 05.12.13 «Системы, сети и устройства телекоммуникаций».

Официальный оппонент

Ведущий специалист

АО «Конструкторское Опытное Бюро

РадиоАппаратуры» (АО «КОБРА»),

г. Владимир,

кандидат технических наук



02.12.2019

С.И. Кучин

600026, г Владимир, ул. Лакина, д. 1-Г,

телефон: (4922) 43-14-14,

Email: Kuchin@kobra.su

Подпись Кучина С.И. заверяю,

менеджер по персоналу



О.Ю. Лоева

